GraphiCon 2018

28-я Международная конференция
по компьютерной графике
и машинному зрению

Труды конференции

GraphiCon 2018

The 28th International Conference
on Computer Graphics and Vision

Conference Proceedings

Национальный исследовательский Томский политехнический университет,
Томский государственный университет систем управления и радиоэлектроники
Томск, 24—27 сентября 2018 года
http://2018.graphicon.ru/

УДК 004.92
ББК 32.973.26-018.2
G73

G73 **GraphiCon 2018**: труды 28-й Междунар. конф. по компьютерной графике и машинному зрению Томск, 24–27 сент., 2018 г.) / Нац. исслед. Том. политех. ун-т. – Томск, 2018. – 512 с.

ISSN 2618-8317

Настоящий сборник содержит материалы 28-й Международной конференции по компьютерной графике и машинному зрению. В её подготовке и проведении участвовали специалисты ведущих организаций, отраслевых и академических научно-исследовательских институтов, а также высшей школы.

Цикл «GraphiCon» представляет собой регулярные ежегодные конференции. Он служит международным форумом для исследователей и разработчиков со всего мира, целью которого является обмен новыми идеями, знаниями и опытом в области компьютерной графики, машинного зрения, обработки изображений, систем визуализации и виртуального окружения, систем поддержки принятия решений, облачных вычислений и обработки больших массивов данных.

В сборник включены приглашенные, пленарные и секционные доклады конференции.

Адресуется сотрудникам научно-исследовательских и образовательных организаций, предприятиям ИТ-индустрии, аспирантам, студентам, а также широкому кругу читателей.

УДК 004.92
ББК 32.973.26-018.2

ISSN 2618-8317

Организационный комитет

Захарова Алёна Александровна
ТПУ

Галактионов Владимир Александрович
ИПМ им. М.В. Келдыша РАН

Берберова Мария Александровна
МФТИ

Бондарев Александр Евгеньевич
ИПМ им. М.В. Келдыша РАН

Волобой Алексей Геннадьевич
ИПМ им. М.В. Келдыша РАН

Евсютин Олег Олегович
ТУСУР

Ерёмченко Евгений Николаевич
ООО «Неогеография»

Завьялов Дмитрий Алексеевич
ТПУ

Конушин Антон Сергеевич
МГУ

Крылов Андрей Серджевич
МГУ

Ротков Сергей Игоревич
ННГАСУ

Турлапов Вадим Евгеньевич
ННГУ

Берберова Мария Александровна
МФТИ

Будак Владимир Павлович
МЭИ

Вехтер Евгения Викторовна
ТПУ

Гергет Ольга Михайловна
ТПУ

Губин Владимир Евгеньевич
ТПУ

Карпов Леонид Евгеньевич
ИСП РАН

Катаев Михаил Юрьевич
ТУСУР

Кочегуров Александр Иванович
ТПУ

Мамонтов Геннадий Яковлевич
ТПУ

Масленникова Алла Александровна
МГУ

Медовник Александр Владимирович
ТУСУР

Мещеряков Роман Валерьевич
ТУСУР

Небаба Степан Геннадьевич
ТПУ

Ризен Юлия Сергеевна
ТПУ

Серяков Вадим Александрович
ТПУ

Шкляр Алексей Викторович
ТПУ

Юрченкова Елена Анатольевна
ТУСУР

Яковлев Алексей Николаевич
ТПУ

Программный комитет

Авербух Владимир Лазаревич
ИММ УрО РАН

Арлазаров Владимир Викторович
Smart Engines

Баринова Ольга Вячеславовна
МГУ

Барладян Борис Хаимович
ИПМ им. М.В. Келдыша РАН

Бобков Валерий Александрович
ИАПУ ДВО РАН

Бондарев Александр Евгеньевич
ИПМ им. М.В. Келдыша РАН

Будак Владимир Павлович
МЭИ

Ватолин Дмитрий Сергеевич
МГУ

Волобой Алексей Геннадьевич
ИПМ им. М.В. Келдыша РАН

Велижев Александр
МГУ

Вехтер Евгения Викторовна
ТПУ

Визильтер Юрий Валентинович
ГосНИИАС

Выголов Олег Вячеславович
ГосНИИАС

Галактионов Владимир Александрович
ИПМ им. М.В. Келдыша РАН

Гергет Ольга Михайловна
ТПУ

Горбачев Вадим
ГосНИИАС

Давыдова Евгения Михайловна
ТПУ

Дебелов Виктор Алексеевич
ИВМиМГ СО РАН

Дышкант Наталья Фёдоровна
МГУ

Евсютин Олег Олегович
ТУСУР

Ерёмченко Евгений Николаевич
ООО «Неогеография»

Ерофеев Михаил Викторович
МГУ

Ерухимов Виктор
Itseez

Жданов Дмитрий Дмитриевич
ГОИ

Захарова Алёна Александровна
ТПУ

Знаменская Ирина Александровна
МГУ

Игнатенко Алексей Викторович
МГУ

Катаев Михаил Юрьевич
ТУСУР

Конушин Антон Сергеевич
МГУ

Копылов Андрей Валерьевич
ТулГУ

Крылов Андрей Серджевич
МГУ

Кустикова Валентина Дмитриевна
ННГУ

Местецкий Леонид Моисеевич
МГУ

Мещеряков Роман Валерьевич
ТУСУР

Михайлюк Михаил Васильевич
НИИСИ РАН

Насонов Андрей Владимирович
МГУ

Небаба Степан Геннадьевич
ТПУ

Николаев Дмитрий Петрович
ИППИ РАН

Подвесовский Александр Георгиевич
БГТУ

Потемин Игорь Станиславович
НИУ ИТМО

Ризен Юлия Сергеевна
ТПУ

Ринкевичюс Бронюс Симович
МЭИ

Ротков Сергей Игоревич
ННГАСУ

Рябинин Константин Валентинович
ПГНИУ

Середин Олег Сергеевич
ТулГУ

Серяков Вадим Александрович
ТПУ

Сорокин Дмитрий Васильевич
МГУ

Спицын Владимир Григорьевич
ТПУ

Столбова Ирина Дмитриевна
ПНИПУ

Тикунов Владимир Сергеевич
МГУ

Турлапов Вадим Евгеньевич
ННГУ

Усилин Сергей Александрович
ИСА РАН

Фаворская Маргарита Николаевна
СибГУ им. М.Ф. Решетнева

Фролов Владимир Александрович
МГУ

Чернявский Алексей
Samsung R&D Institute Rus

Чуприна Светлана Игоревна
ПГНИУ

Шапиро Лев Залманович
ИПМ им. М.В. Келдыша РАН

Шкляр Алексей Викторович
ТПУ

Юрин Дмитрий Владимирович
МГУ

Международный программный комитет

Timur Paltashev
AMD, USA

Marina Gavrilova
University of Calgary, Canada

Karol Myszkowski
MPI, Germany

Vladimir Savchenko
Hosei University, Japan

Alexei Sourin
NTU, Singapore

Olga Sourina
FhG IGD, Germany

Tom Sederberg
Brigham Young University, USA

Andre Gagalowicz
INRIA, France

Sabine Coquillart
INRIA, France

Andres Iglesias
University of Cantabria, Spain

Wolfgang Heiden
BRS, Germany

Andre Hinkenjann
IVC BRS, Gemany

Gianluca Mura
Politecnico Di Milano, Italy

Ekaterina Prasolova-Forland
NTNU, Norway

Samet Refik
Ankara Uni, Turkey

Bill Seaman
Duke Uni, USA

Skala Vaclav
University of West Bohemia, Czech Republic

Адрес: ГрафиКон-2018, Национальный исследовательский
Томский политехнический университет (НИ ТПУ),
Инженерная школа информационных технологий и робототехники (ИШИТР),
634034, Россия, г. Томск, ул. Советская, д. 84/3, оф. 216.
Тел.: +7 (382) 260-61-29 E-mail: 2018@graphicon.ru Website: http://2018.graphicon.ru/

Организаторы и спонсоры

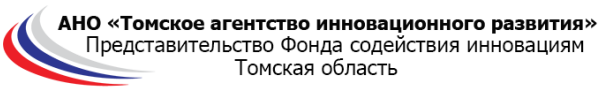
АНО Научное Общество «ГрафиКон»
Организатор

Национальный исследовательский Томский политехнический университет
Локальный организатор

Томский государственный университет систем управления и радиоэлектроники
Соорганизатор

ООО «С3Д Лабс»
Спонсор

Институт прикладной математики им. М.В. Келдыша РАН
Соорганизатор



АНО «Томское агентство инновационного развития»
Соорганизатор

Ассоциация некоммерческих организаций «Томский консорциум научно-образовательных и научных организаций»
Соорганизатор

Информация о спонсорах

C3D Labs

ООО «С3Д Лабс»

В 1995 году компания АСКОН начала разрабатывать собственное геометрическое ядро, заложив основу для системы трёхмерного моделирования КОМПАС-3D. На протяжении 20 последующих лет ядро развивалось в соответствии с запросами архитектурной и машиностроительной отраслей России и стран ближнего зарубежья.

В ноябре 2012 года математическое подразделение АСКОН было выделено в отдельную компанию C3D Labs, цель которой — развитие комплекта инструментальных средств C3D Toolkit для разработки инженерного программного обеспечения и выход с ним как с самостоятельным продуктом на мировой рынок PLM-компонентов.

Под руководством кандидата технических наук, Николая Николаевича Голованова, над математической и программной частью компонентов C3D Toolkit сегодня трудятся высоко квалифицированные специалисты — выпускники ведущих российских вузов, прошедшие строгий конкурсный отбор на знание профильных дисциплин и владение практическими навыками программирования в среде Visual Studio. Силами команды C3D Labs постоянно совершенствуются методы построения и обработки численных моделей геометрии, улучшается их программная реализация, пропагандируется клиентоориентированный подход к работе над проектами, а также предлагается гибкая система лицензирования C3D Toolkit, учитывающая собственную бизнес-модель каждого заказчика.

C3D Labs — резидент инновационного центра Сколково.

Дорогие участники Графикон-2018!

Организационный комитет 28-ой Международной конференции по компьютерной графике и машинному зрению ГрафиКон-2018 приветствует вас в Томске!

Графикон впервые ступил на сибирскую землю, расширив тем самым не только свою географию, но и границы замечательного профессионального сообщества. И, как Россия приросла Сибирью, так и Графикон прирастает. Особенно приятно отметить возросшее число молодых участников, что особенно важно для появления новых научных школ и развития уже заявивших о себе и добившихся результатов научных групп.

Сегодня Графикон стал еще интеллектуальнее и эстетичнее! Впервые появились две новые секции: «Автоматизация проектирования, тренажеры и симуляторы» и «Компьютерная графика, светотехника и дизайн».

Томский политехнический университет (ТПУ) и Томский государственный университет систем управления и радиоэлектроники (ТУСУР), как организаторы, рады всем участникам Графикона-2018 и желают продуктивно и с интересом провести «графиконовские дни». Пусть они станут для вас, молодые ученые, временем интересных встреч и знакомств, приглашаем вас стать частью нашего замечательного сообщества. А всем тем, кто уже графиконовец, желаем вдохновения для поиска новых идей, приятных встреч со своими партнерами и оппонентами.

О Сибири и Томске можно много писать… но передать колорит этих мест сложно. На страницах http://2018.graphicon.ru можно узнать о традициях нашего края и истории университетов-организаторов.

Поэтому приглашаю всех на страницы нашего сайта и заседания ГрафиКона-2018. Всем участникам и гостям желаю творческих и научных успехов, здоровья и благополучия.

От имени организационного и программного комитетов ГрафиКон-2018,

председатель оргкомитета ГрафиКон-2018, д.т.н.
Захарова Алёна Александровна

Содержание

Пленарные доклады

Е.Н. Ерёмченко, А.С. Крылов, Д.Ю. Васин, Т.Т. Палташев, М.Ю. Катаев, С.Е. Сляднев
Список пленарных докладов 15

Реалистичная графика

S.V. Ershov, D.D. Zhdanov, A.G. Voloboy
Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing 20

V.A. Frolov, V.A. Galaktionov, V.V. Sanzharov
Mixing ordinary and Markov chain Monte Carlo rendering techniques 26

V.V. Sanzharov, V.A. Frolov
Precomputed procedural content level of detail 31

Б.Х. Барладян, Е.Д. Бирюков, А.Г. Волобой, Л.З. Шапиро
Эффективный алгоритм визуализации предварительно рассчитанных лучей 36

А.М. Вин
Визуализация объёмных объектов с помощью структурированного оптического излучения 40

А.Д. Жданов, Д.Д. Жданов, Н.Н. Богданов, И.С. Потемин, В.А. Галактионов, И.Л. Лившиц
Использование методов виртуального прототипирования при разработке устройств виртуальной реальности, не подверженных проблеме конфликта вергенции и аккомодации зрения 43

А.С. Рудникович, К.В. Шипунова
ISO 12233 и программные средства измерения частотно-контрастных характеристик изображений 48

М.И. Свириденко
Съемка сферических панорам с каналом освещенности 52

В.Г. Соколов, Д.Д. Жданов, И.С. Потемин, Б.Х. Барладян, Н.Н. Богданов, А.Д. Жданов
Восстановление двунаправленной функции рассеяния света для среды с объемным рассеянием 56

М.И. Сорокин, Д.Д. Жданов
Использование свёрточных нейронных сетей для классификации видов искусственного освещения 61

Научная визуализация и визуальная аналитика

M. Demeshko, A. Demin, D. Oliva
Edge detection using neural network committee 66

K. Ryabinin, M. Kolesnik
Using IoT Devices Powered by Scientific Visualization Tools to Create Interactive Paleontological Museum Exhibitions 70

Н.В. Авербух
К вопросу о феномене присутствия в системах компьютерной визуализации на базе виртуальной реальности . . 74

В.Л. Авербух, И.Л. Гвоздарев, Г.И. Левчук
Визуализация программного обеспечения с использованием виртуальной реальности 77

С.В. Андреев, А.Е. Бондарев
Построение стереоанимаций на современных стереоустановках 82

С.В. Андреев, Н.А. Бондарева
Построение отображения текстовой информации в стереопрезентациях 86

К.Ю. Андросов, А.Н. Горбунов, С.В. Давыдов
Формирование пространства признаков для определения границ зерен структурных компонентов сплава по полихроматическому растровому изображению его микроструктуры 90

Р.А. Багутдинов, С.Г. Небаба, А.А. Захарова
Редуктивно-аккомодативный принцип обработки гетерогенных данных мультисенсорной системы 94

А.Е. Бондарев, В.А. Галактионов, Э.С. Клышинский, Л.З. Шапиро
Система процедур визуального анализа многомерных данных 98

П.А. Васёв, В.Л. Авербух, М.О. Бахтерев, Д.В. Манаков, Д.Ю. Филоненко, М.А. Форгани
Сцены визуализации в вычислительных блокнотах 103

А.М. Вин, Б.С. Ринкевичюс, А.Ю. Поройков
Количественная визуализация формы физических тел с помощью точечно структурированного оптического излучения 107

Т.П. Галкин, М.А. Григорьева, А.А. Климентов, Т.А. Корчуганова, И.Е. Мильман, В.В. Пилюгин, М.А. Титов
Кластерный анализ вычислительных задач системы управления загрузкой эксперимента ATLAS на LHC с использованием визуальной аналитики 111

Д.С. Гордеев
Модель визуализации изменений в графе внутреннего представления программ языка Cloud Sisal 115

Ю.Н. Дубнищев, В.А. Арбузов, Э.В. Арбузов, В.С. Бердников, С.А. Кислицын, О.С. Мелёхина
Оптическая диагностика конвективных структур в вертикальном слое воды, ограниченном теплообменными поверхностями при нестационарных граничных условиях 119

Ю.Н. Дубнищев, В.А. Арбузов, В.В. Лукашов, В.В. Леманов
Гильберт-диагностика пламён 123

А.А. Захарова, Е.В. Вехтер, А.В. Шкляр
Когнитивная интерпретация визуализации 127

И.А. Знаменская, Е.Ю. Коротеева, Ф.Н. Глазырин
Кросс-корреляционные методы анализа изображений в экспериментальной визуализации потоков 132

И. Знаменская, Д. Наумов, И. Дорощенко, А. Кузнецов, А. Луцкий
Цифровые технологии в анализе теневых изображений газоплазменных потоков 135

В.Н. Касьянов, Т.А. Золотухин
Визуализация графовых представлений потоковых программ 138

В.В. Князь, С.Ю. Данилов, О.В. Выголов, А.Н. Бордодымов
Визуализация препятствий на индикаторе на лобовом стекле с помощью оценки оптического потока 142

Д.В. Манаков
Визуальная аналитика и модели абстракции данных 146

С.В. Мещеряков, Д.А. Щемелинин
Методы визуализации и анализа больших данных в облачных информационных системах 151

С.В. Муравьев, М.А. Борисова
Анализ и визуализация данных энергоаудита методом агрегирования предпочтений 155

А.Г. Подвесовский, Р.А. Исаев
Критерии когнитивной ясности как основа для построения метафоры визуализации нечетких когнитивных карт 158

И.Л. Расковская, Б.С. Ринкевичюс, А.В. Толкачев, А.В. Тихомирова
Визуализация траекторий пузырьков газа в акустическом поле 163

К.В. Рябинин, С.И. Чуприна, К.И. Белоусов, С.С. Пермяков
Методы визуальной аналитики вариативности речевого поведения пользователей социальных сетей в зависимости от психологических черт личности 167

О.А. Салтыкова, И.В. Папкова, А.В. Крысько
Визуализация режимов колебаний балок с учетом их контактного взаимодействия 172

А.О. Трубаков, Т.Д. Праздникова
Методы реставрации потерянных и искаженных участков изображений 176

М.В. Харинов
Об СКО, методе K-средних и деревьях в приложении к обработке изображений 180

О.О. Шумская, О.О. Евсютин
Встраивание информации в сжатые JPEG-изображения с минимизацией проявления демаскирующих признаков 185

Компьютерное зрение

Д.С. Белобородов
Контурная глубина плоской фигуры: применение для определения симметрии и классификации форм 191

А.П. Береснев, И.В. Зоев, Н.Г. Марков
Исследование свёрточных нейронных сетей класса YOLO для мобильных систем детектирования объектов на изображениях 196

И.В. Зоев, Н.Г. Марков, А.П. Береснев, Т.А. Ягунов
Аппаратная реализация на ПЛИС свёрточных нейронных сетей для распознавания объектов на изображениях . . 200

И.А. Калиновский, Г.М. Лаврентьева
Обнаружение спуфинг-атак на систему лицевой биометрии 204

Д.А. Купляков, Е.В. Шальнов, В.С. Конушин, А.С. Конушин
Распределённый алгоритм сопровождения людей в видео 208

Д.Н. Лепихова, В.Ю. Гудков
Метод быстрого отбора идентичных контрольных точек на изображениях отпечатка пальца 214

Н.А. Ломов, С.П. Арсеев
Нейронные сети для распознавания формы по медиальному представлению 218

А.К. Мовчан, В.В. Капустин, М.И. Курячий
Методы и средства томографического видения пространства активно-импульсными телевизионными измерительными системами 222

Л.В. Новоторцев, А.Г. Волобой
Ускорение метода перебора плоскостей при трехмерной реконструкции строений по аэроснимкам 226

Е.С. Попова, В.Г. Спицын, Ю.А. Болотова
Программа и алгоритм сегментации и распознавания рукопечатных символов с помощью сверточных нейронных сетей 230

А.И. Соколова, А.С. Конушин
Обзор методов распознавания человека по походке в видео 234

В.Г. Спицын, Ю.В. Савицкий, А.Б. Казиев, Ю.А. Болотова
Нейросетевой метод обнаружения размытых изображений лиц 238

П.В. Хрушков, В.И. Шахуро, А.С. Конушин
Синтез изображений дорожных знаков с помощью условных порождающих противоборствующих нейросетей . . 242

Т.Н. Шолохова, Л.М. Местецкий
Классификация изображений таблиц на основе структурного подхода 247

Обработка и анализ биомедицинских изображений

V.V. Danilov, I.P. Skirnevskiy, R.A. Manakov, O.M. Gerget, F. Melgani
Image-based algorithm for the catheter recognition based on GLCM 252

A. Khvostikov, A. Krylov, O. Kharlova, N. Oleynikova, I. Mikhailov, P. Malkov
CNN-based histological images segmentation of mucous glands 258

A.A. Melnikova, A.S. Krylov
Computer analysis and restoration of motion blurred images 262

Д.А. Гаврилов, А.В. Мелерзанов, Н.Н. Щелкунов, Э.И. Закиров
Диагностика меланомы с использованием нейронной сети Inception v3 266

А.А. Довганич, А.С. Крылов, Д.В. Юрин
Алгоритм нелокального среднего основанный на модифицированном индексе структурного сходства 270

М.В. Дубровская, М.М. Новожилов
Лучевой метод реконструкции поверхности сердца человека по данным компьютерной томографии 275

Н.В. Мамаев, Д.В. Юрин, А.С. Крылов
Нахождение параметров метода шумоподавления с использованием вейвлет-преобразования на медицинских изображениях на основе анализа хребтовых структур 280

А.В. Насонов, И.С. Лузин, А.А. Насонова, А.С. Крылов, Н.С. Семенова, В.С. Акопян
Повышение качества изображений оптической когерентной томографии сетчатки глаза с использованием метода полной обобщённой вариации 285

М.М. Новожилов, М.В. Дубровская
Реконструкция поверхности по точкам контуров в ортогональных сечениях томограммы 289

С.А. Носова, В.Е. Турлапов
Машинное обучение по текстурным признакам в задаче детектирования нейронов на срезах мозга, окрашенных по Нисслю 292

Цифровая Земля и Большие Данные

О.О. Евсютин, А.С. Кокурина, Р.В. Мещеряков
Улучшенный алгоритм встраивания информации в область ДКП цифровых изображений на основе метода QIM с переменным шагом квантования 297

Е.Н. Ерёмченко
Визуализация и новое определение знака 301

Е.Н. Ерёмченко, В.С. Тикунов, А.А. Захарова
Цифровая Земля и «цифровые» инициативы современности 304

Д.А. Завьялов, А.А. Захарова
Анализ многомерных данных на основе комплексной модели нефтегазового месторождения 307

С. Исхаков, А. Исхаков, Р. Мещеряков
Визуализация больших данных с применением методов корреляции событий информационной безопасности в инфраструктуре интернета вещей 311

Л.И. Лебедев, А.О. Шахлан
Пространственный анализ гиперспектральных изображений 315

И.П. Урманов, А.Н. Камаев, А.А. Сорокин
Обнаружение термических аномалий на инфракрасных изображениях вулканов 319

Геометрическое моделирование. Компьютерная графика в образовании

R. Ziatdinov, V. Muftejev, R. Nabiyev, A. Mardanov, R. Akhmetshin
Techniques for modeling a high-quality B-spline curves by S-polygons in a float format 324

А.А. Бойков, А.М. Федотов
Использование чертежей с разметкой для проверки электронных чертежей и трехмерных моделей 328

Д.Ю. Васин, В.П. Громов, Е.А. Кустов, С.И. Ротков, Ю.В. Ясаков
Автоматизированная технология ввода графических документов со слабо формализованным описанием объектов 332

Д.Ю. Васин, В.П. Громов, С.И. Ротков
Модели представления растровых графических документов со слабо формализованным описанием объектов . . 337

Е.В. Вехтер, В.Ю. Радченко, Т.Д. Казакова
Концепция мультимедийного дидактического комплекта для морфологического анализа в обучении студентов-дизайнеров 348

Д.В. Волошинов
Об уточнении некоторых понятий конструктивной геометрии 350

О.С. Воронова
Геометрическое моделирование и оптимизация конструктивных характеристик котлоагрегата 354

Е.В. Конопацкий
Решение дифференциальных уравнений методами геометрического моделирования 358

А.Л. Липкина, Л.М. Местецкий
Классификация букв в изображениях на основе медиального представления 362

Т.В. Мошкова, С.А. Роменский, С.И. Ротков, В.А. Тюрина
Исследование временных характеристик алгоритма восстановления каркасной модели по проекционным изображениям . . . 369

К.Л. Панчук, Е.В. Любчинов, Т.М. Мясоедова
Циклография. Аспекты теории и практических применений . . . 372

Ф.Н. Притыкин, Д.И. Нефедов
Аналитическое задание четырехмерной области разрешенных конфигураций руки андроидного робота при наличии запретной зоны в рабочем пространстве . . . 376

И.Д. Столбова, М.А. Ширинкина
«Доступная» графика в довузовском образовании . . . 380

М.С. Тарских, А.В. Шкляр
Инструмент автоматизации создания и визуализации геометрии корпуса аудиосистемы рупорного типа . . . 385

Д.Е. Тихонов–Бугров, С.Н. Абросимов, В.А. Дюмин
Базовые принципы преподавания инженерной и компьютерной графики будущим специалистам . . . 389

В.А. Тюрина, М.М. Смычек, С.И. Ротков, Т.В. Мошкова
Синтез предварительной каркасной модели 3D-объекта по техническому чертежу, содержащему разрезы и сечения . . . 393

Компьютерная графика, светотехника и дизайн

А.Ю. Басов, В.П. Будак
Модель рассеивающего слоя с диффузной подложкой и френелевской границей . . . 399

А.А. Бородина, Т.С. Хруль
Освещение лесопарковой зоны. Чувственные аспекты темы . . . 402

И.А. Бринюк, В.Ю. Радченко
Особенности проектирования развивающей модульной мебели для детей . . . 406

В.П. Будак, А.В. Гримайло
Влияние учета поляризации света на результат при расчете многократных отражений . . . 409

В. Будак, В. Желтов, В. Чембаев, Т. Мешкова
Оценка качества внутреннего освещения в сценах с неравномерными блескими источниками . . . 411

И.В. Бычкова, В. Колина
Разработка концепции источника освещения на базе Arduino Uno, автоматически поддерживающего заданный уровень освещенности . . . 415

Д.Т. Валиев, Н.Е. Россомахина, Е.В. Тюлькин, Н.А. Агапов
Разработка математического обеспечения по расчету и оценке качества изображения оптической системы . . . 418

Д.Г. Грачёв, В.П. Будак
Моделирование световых приборов автомобилей с использованием трассировки лучей . . . 422

Ю.С. Дё, А.В. Шкляр
Снижение ресурсоемкости дизайн-проектирования . . . 424

Р.Г. Калинин, Д.Т. Валиев, Н.А. Агапов
Расчет и конструирование оптико-механической системы зрительной трубы Ньютона с ортоскопическим окуляром . . . 428

И.А. Канаева, Ю.А. Болотова
Создание панорамного видео . . . 432

Н.С. Канкеева, Е.М. Давыдова
Применение метода художественного формообразования в проектировании интерьерных часов «К вершине» . . . 436

С. Касаткин, В. Калайда
Формализация описания класса изображений, полученных в различных условиях освещения . . . 439

А.Ю. Ларионова, Ю.С. Ризен
Основы использования цвета в рамках геймдизайна . . . 442

А.М. Минчук, Е.М. Давыдова
Методы дизайн-проектирования портативного медицинского устройства 446

С.Г. Небаба
Методы оценки и подготовки изображений в видеопотоке к распознаванию объектов 450

М.М. Новиков, В.А. Князь, Р.М. Галеев
Реверсивный инжиниринг палеоантропологических объектов для визуализации и исследований 454

С.А. Романенко, Е.М. Шакиров, С.Б. Туранов
Подходы к созданию культиватора микроводорослей для домашнего использования 458

В.И. Сафьянникова, Е.В. Вехтер
Особенности дизайн-проектирования установки для переработки полимеров 460

М.Е. Сотников, Ю.С. Ризен
Создание развивающего инструмента для обеспечения безбарьерного инклюзивного обучения 463

Б.Б. Тарбаев, В.Т. Калайда
Выделение объекта в видеопотоке в различных условиях освещения 467

Ю.П. Хмелевский, В.А. Серяков, Г.Я. Мамонтов, Т.С. Хруль
Применение световых панно для преодоления депрессии 471

Л.Н. Черепова, А.А. Черепов, А.В. Шкляр
Решение задач эргономики с помощью трехмерной модели рабочей зоны 475

Автоматизация проектирования, тренажеры и симуляторы

A. Malyshev, A. Zhidkov, V. Turlapov
Adaptive mesh generation using shrink wrapping approach 479

М.С. Копылов, Е.Д. Бирюков, Б.Х. Барладян
Особенности использования сценариев в автоматическом тестировании систем оптического моделирования . . . 484

С.Е. Сляднев, А.С. Малышев, В.Е. Турлапов
Автоматизированное упрощение машиностроительных CAD-моделей и сборок без использования истории построения 488

С.Е. Сляднев, В.Е. Турлапов
Реконструкция параметрической модели лопатки турбины из набора аэродинамических сечений с использованием техники сглаживания 495

М.В. Титова, Т.Н. Томчинская
Разработка имитационной модели сократительной функции сердца в среде Autodesk Maya 500

Пленарные доклады

Plenary Talks

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Визуализация и новое определение знака

Е.Н. Ерёмченко
eugene.eremchenko@gmail.com
МГУ имени М.В. Ломоносова, Москва, Россия

Аннотация

В докладе предлагается к обсуждению новое определение знака, открывающее новые возможности для исследования семиотики визуализации.

Евгений Николаевич Ерёмченко - Руководитель группы «Неогеография», член совета ISDE, соучредитель DSRA, научный сотрудник Московского государственного университета им. Ломоносова.

Алгоритм нелокального среднего основанный на модифицированном индексе структурного сходства

А.А. Довганич, А.С. Крылов, Д.В. Юрин
a.dovganich@yandex.ru|kryl@cs.msu.ru|yurin@cs.msu.ru
Факультет Вычислительной математики и кибернетики
МГУ имени М.В. Ломоносова, Москва, Россия

Аннотация

Предлагается новый алгоритм шумоподавления на изображениях. Он является модификацией алгоритма нелокального среднего с использованием метрики основанной на индексе структурного сходства(SSIM). Производится анализ этой весо-вой функции с точки зрения статистических характеристик изображений, физиологии восприятия изображений и струк-туры наблюдаемой сцены. Индекс SSIM разбирается на отдельные составляющие и для каждой компоненты задается фи-зически обоснованная весовая функцию. Результаты модифицированного метода сравниваются с алгоритмом нелокального среднего, где веса вычисляются по метрике L2 и SSIM.

Крылов Андрей Серджевич – д.ф.-м.н., профессор кафедры математической физики, заведующий лабораторией математических методов обработки изображений, помощник декана факультета Вычислительной математики и кибернетики Московского государственного университета им. Ломоносова.

Модели представления растровых графических документов со слабо формализованным описанием объектов

Д.Ю. Васин[1], В.П.Громов[1], С.И. Ротков[2]
dm04@list.ru|gvp1941@bk.ru|rotkov@nngasu.ru
Нижний Новгород, Россия,
[1]ИТММ, Нижегородский Государственный Университет им. Н.И. Лобачевского
[2]Нижегородский Государственный Архитектурно-Строительный Университет

Аннотация

В докладе в рамках развития комбинаторно–геометрического подхода к обработке пространственно–распределенных данных изложены оригинальные модели представления растровых и векторных графических документов со слабо формализованным описанием объектов, предложены алгоритмы их формирования и поддержки, а так же разработанные на основе указанных моделей алгоритмы фильтрации растровых данных и распознавания объектов по растровым данным планов поэтажной застройки, символов на растровых изображениях текстовых документов.

Васин Дмитрий Юрьевич – к.т.н., ведущий научный сотрудник Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Аппаратные и программные платформы Advanced Micro Devices для высокопроизводительных вычислений, компьютерной графики и машинного интеллекта

AMD Hardware and Software Platforms for HPC, Computer Graphics and Machine Intelligence

Т.Т. Палташев
Timour.Paltashev@amd.com
AMD Radeon Technology Group

Abstract

New server processor EPYC for enterprises and data centers was introduced in 2017 and new version EPYC+ in summer 2018. It has several new architecture features and built on Zen CPU core which is connected to redesigned memory hierarchy with data encryption and virtualization protection. EPYC+ family has very capable I/O and power management subsystems making them a best price-performance server option for data centers and cloud computing. There is comprehensive IHV and ISV platforms ecosystem for EPYC+ and upcoming EPYC ROME with 64 Zen2 cores helping to build various configurations for applications in HPC, cloud gaming/computing and industrial supercomputing. EPYC+ server family is combined with Radeon Instinct GPU accelerators to deliver petaflop range performance in clusters like P47. Such powerful hardware platform is supported by Radeon Open Compute (ROCm) open source software stack for HPC and Machine Intelligence with various optimized libraries and profiling tools. Deep learning applications can be mapped via several frameworks supported by ROCm and MIOpen software stack.

Семейство графических процессоров VEGA: Решения для компьютерной графики, систем машинного интеллекта и высокопроизводительных вычислений

Vega GPU Family: A Solution for Graphics, Scalable Machine Intelligence Systems and HPC

Т.Т. Палташев
Timour.Paltashev@amd.com
AMD Radeon Technology Group

Abstract

New GPU Vega become a heart of several product lines as Radeon RX64 in gaming, Radeon PRO WX8200 in workstations and Radeon Instinct MI25 accelerators for deep learning applications. 64 nCU compute units to accelerate demanding workloads with up to 13.7 TFLOPS of peak SPFP to speed up professional graphics applications and premium games. 16GB of high bandwidth cache (HBC) memory supports massive textures and data sets. Removes the need for complicated memory management and large resources are not required to remain complete in local memory with active pages have prioritized residency in HBC, inactive pages are marked for migration to slower memory and automatically moved. Vega architecture has high performance FP16 and FP32 compute units supported by open software ROCm platform for HPC-class rack scale. Radeon Instinct™ MI25 accelerators have large BAR support for mGPU peer to peer communication for superior compute density and performance per node in combination with AMD EPYC™ server processors.

Палташев Тимур Турсунович - д.т.н., Senior Manager in Radeon Technology Group, Advanced Micro Devices.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Обработка и анализ изображений в задачах альянса «Техническое зрение»

Image processing and analysis in the tasks of the "Technical Vision" Alliance

М.Ю. Катаев[1], В.А. Ширшин[2]
kmy@asu.tusur.ru
[1]Томский университет систем управления и радиоэлектроники, Томск, Россия;
[2]ЗАО «Элекард Девайсез», Томск, Россия

Аннотация

Прогресс в области вычислительной и измерительной оптической техники, ее удешевление, привел к повсеместному распространению программно-аппаратных решений, связанных с использованием цифровых изображений. Можно с большой долей уверенности сказать, что численная обработка изображений стала не только предметом научных исследований, но и многих практических приложений, элементом технологий. Это позволило использовать результаты применения элементов технического зрения в искусстве, промышленности, медицине, спорте, образовании, охране и множества других приложениях. Техническое зрение можно рассматривать как комплексное программно-аппаратное направление научных и практических знаний, позволяющее решать разнообразные задачи от разработки оборудования получения изображений, оцифровки полученных изображений, процессорного ускорения алгоритмов обработки и анализа. В настоящее время многие задачи технического зрения решены на высоком научно-практическом уровне, однако, условия получения измерений, технические возможности аппаратуры, существенная сложность анализируемой сцены оставляет возможность для развития уже известным направлениям, и образоваться новым. Большинство задач технического зрения сейчас решается академической и вузовской наукой, ИТ компаниями изолированно друг от друга и объединения возникают лишь только на уровне воплощения конкретных задач. В данной работе рассматривается описание структуры компаний (альянс «Техническое зрение» [https://svtsk.ru/]), в той или иной мере решающих или использующих задачи технического зрения. Кратко рассказывается об основных направлениях деятельности компаний и возможные направления объединения усилий в рамках альянса. Деятельность альянса рассчитана на то, чтобы не изменяя внутренней деятельности ИТ компаний, быть полезными для решения постоянно расширяющегося круга задач, требующих привлечения разнообразных специалистов. Именно такая структура взаимоотношения внутри альянса, по мнению участников, позволит решать более крупные задачи, чем по силам какой-то отдельно взятой ИТ компании. Помимо практической деятельности, альянс планирует заниматься и образовательной деятельностью, расширяя круг специалистов, владеющих основами подходов технического зрения.

Abstract

Progress in the field of computational and measuring optical technology, its reduction in price, led to widespread distribution of software and hardware solutions related to the use of digital images. It can be said with great certainty that the numerical processing of images has become not only the subject of scientific research, but also many practical applications, an element of technology. This allowed us to use the results of applying elements of technical vision in art, industry, medicine, sports, education, protection and many other applications. Technical vision can be considered as a complex software and hardware direction of scientific and practical knowledge, which allows to solve various tasks from the development of the imaging equipment, the digitization of the images obtained, and the processor acceleration of the processing and analysis algorithms. At present, many problems of technical vision are solved at a high scientific and practical level, however, the conditions for obtaining measurements, the technical capabilities of the equipment, the essential complexity of the analyzed scene leaves an opportunity for the development of already known areas, and form a new one. Most of the tasks of technical vision are now being solved by academic and high school science, IT companies are isolated from each other and associations arise only at the level of the implementation of specific tasks. In this paper, we describe the structure of companies (the "Technical Vision" alliance [https://svtsk.ru/]), to some extent deciding or using the tasks of technical vision. Briefly describes the main areas of activity of companies and possible areas of uniting efforts within the alliance. The alliance's activities are designed to be useful in order not to change the internal activities of IT companies, to solve an ever-expanding range of tasks requiring the involvement of a variety of specialists. It is this structure of the relationship within the alliance that, in the opinion of the participants, will allow solving larger tasks than the forces of a single IT company. In addition to practical activities, the alliance plans to engage in educational activities, expanding the range of specialists who know the basics of approaches to technical vision.

Катаев Михаил Юрьевич – д.т.н., профессор Томского университета систем управления и радиоэлектроники, член-корреспондент САН ВШ.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Проблемы развития геометрического моделирования вне коммерческих САПР

С.Е. Сляднев, В.Е. Турлапов
sergey.slyadnev@gmail.com | vadim.turlapov@itmm.unn.ru
Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия

Аннотация

Рассматриваются современные вызовы геометрического моделирования с точки зрения независимого разработчика инженерного ПО. Исходя из опыта авторов, сформулированы и обсуждаются пять групп взаимосвязанных проблем:

1) Открытое геометрическое ядро нового поколения. Проблема создания открытого геометрического ядра, обладающего новыми качествами полноты не только по отношению к созданию геометрических форм, но и по отношению ко всему кругу задач реинжиниринга, инженерного анализа, оптимизации по геометрическим, функциональным и технологическим критериям.

2) Обмен инженерными данными. Проблема разработки открытого формата(ов), поддерживающего новое ядро и создания конвертеров, обеспечивающих надежное чтение и запись, усилиями всех сторон (академическим сообществом, производителями и пользователями ПО).

3) Прямое редактирование. Проблема распознавания конструктивных элементов модели и их редактирования без потери замысла проектировщика. Операторы прямого редактирования должны быть доступны в геометрическом ядре.

4) Параметрическая оптимизация. Проблема обеспечения использования в геометрическом ядре параметров и критериев (ограничений) адекватных функциональному назначению объекта проектирования, создающих возможность решения задач параметрической оптимизации.

5) Обратный инжиниринг. Проблема включения в состав ядра нового поколения операций для реинжиниринга из точечно-сеточного представления, возникающих в процессе проектирования по различным причинам.

Решение рассмотренных проблем видится в создании коллективными усилиями, академического сообщества и индустрии, открытого геометрического ядра нового поколения для нужд САПР. Демонстрируются возможности прототипа ядра, снабженного аппаратом визуальной аналитики и прямого (основанного на принципе feature recognition) редактирования геометрических данных.

Challenges in geometric modeling field out of commercial CAD-systems

S.E. Slyadnev, V.E. Turlapov
sergey.slyadnev@gmail.com | vadim.turlapov@itmm.unn.ru
Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation

Abstract

This review presents modern challenges of geometric modeling from the standpoint of an independent engineering software developer. Based on the applied experience of the authors, five groups of interrelated problems are formulated and discussed:

1) A new-generation open geometric kernel. The problem of creating an open geometric kernel that possesses new qualities of completeness, not only in relation to the creation of a shape, but also in relation to the whole range of tasks of reengineering, engineering analysis, and optimization according to geometric, functional and technological criteria.

2) Exchange of engineering data. The problem is to develop an open format(s) that support a new kernel and create converters that ensure reliable reading and writing by the efforts of all parties (the academic community, manufacturers, and users of the software).

3) Direct editing. The problem of recognizing the design features of the model and their editing without losing the design intent. The corresponding direct editing operators must be available in the geometric kernel.

4) Parametric optimization. The problem of ensuring the utilization of parameters and criteria (constraints) adequate to the functional designation of the design object, enabling the possibility of parametric optimization in the geometric kernel.

5) Reverse engineering. The problem of including in the kernel a set of operations for reengineering from a point-cloud/mesh representation that arises during the design process for various reasons.

The solution of the considered problems is seen in the creation of a new generation of an open geometric kernel for the needs of CAD by collective efforts of the academic community, and industry. The review demonstrates the capabilities of the prototype core, equipped with facilities for visual analytics and direct editing (based on feature recognition) of geometric data.

Сергей Евгеньевич Сляднев – OPEN CASCADE SAS, руководитель проектной группы, аспирант Нижегородского Государственного Университета им. Н.И. Лобачевского (Россия).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Реалистичная графика

Realistic Graphics

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing

S.V. Ershov[1], D.D. Zhdanov[2], A.G. Voloboy[1]
measure@spp.keldysh.ru|ddzhdanov@mail.ru|voloboy@gin.keldysh.ru
[1]Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, 125047 Miusskaya sq. 4, Moscow, Russia;
[2]St. Petersburg National Research University of Information Technologies, Mechanics, and Optics, St. Petersburg, Russia

In bi-directional Monte-Carlo ray tracing using photon maps the backward ray collects luminance when hitting a diffuse surface, and the number of diffuse events is limited by the so-called "backward diffuse depth", BDD. Its value strongly affects the resulting noise. The paper describes a method to choose BDD to reduce noise. The idea is that some surfaces while diffuse are processed as specular, not incrementing the diffuse event counter. Also not all illumination photons are collected at the hit points, and this is also affected by the proposed method. We show how this treatment allows to reduce noise both for surface and volumetric scattering.

Keywords: *bi-directional ray tracing, noise reduction, BDF.*

Abbreviations

MCRT = Monte-Carlo ray tracing

FMCRT = forward Monte-Carlo ray tracing.

BMCRT = backward Monte-Carlo ray tracing.

BDF = bi-directional scattering function. It describes surface luminance as a function of the illumination and observation direction

BDD = backward diffuse depth. It is a specific parameter of a hybrid ray tracing, when FMCRT calculates illumination and BMCRT is used to convert it to the observed luminance. In this method the backward ray usually has a limited "length" and terminates after BDD diffuse events.

1. Introduction

A powerful method of calculation of a virtual camera image is a bi-directional Monte-Carlo ray tracing using photon maps [4]. FMCRT creates photon maps that allow to estimate illumination of scene surfaces and then BMCRT converts them to the camera image. It traces rays from camera and wherever they hit a diffuse surface, they take the local illumination from the photon map, "convolve" it with the local BDF and add the resulting luminance to the image pixel. There are many modifications of this method [3], [8], [6], [1] which all follow the above general idea.

An important parameter of such an approach is the number of operators ("ray events") in that product series that are handled with BMCRT. Usually only diffuse events are counted, and their number is named "backward diffuse depth" (BDD). Efficiency of the approach, i.e. the rate of convergence (or noise level, which is more or less the same) strongly depends on that BDD, and its optimal value is specific for each scene.

A better approach is to let it be different across the scene [1] and even mix calculations with different BDD [1], and frequently it is possible to find it, automatically or manually, so that the calculations are quite efficient. But this is not always and in some cases changing BDD does not help and whichever value we choose the image is highly noisy. Here we also prove why an "adaptive choice" of BDD individually for each ray may result in a biased estimates (i.e. the calculated luminance converges to a wrong value).

We propose a simple method which helps in such cases. The idea of the method is that the scene objects (surface or volumetric) which have some diffuse scattering are separated into two sets: "genuine diffuse" and "quasi-specular". The latter are *usually* those with a narrow, nearly-specular scattering function, but formally the criterion can be *arbitrary*, e.g. one can treat a Lambert surface as quasi-specular. Usually the separation treats each scattering function (BDF for a surface or phase function for a medium) as a *sum* of quasi-specular and genuine diffuse components so that both are not zero.

A quasi-specular scattering does not increase the "event counter" (which terminates the ray when it exceeds the BDD). Also, it is necessary that quasi-specular component of BDF does not convolve with the *indirect* illumination (i.e. FMCRT rays that underwent a genuine diffuse scattering). This change to the BMCRT can be applied to both surface and volumetric scattering to reduce image noise and the memory used.

2. Global illumination equation

Light distribution in a scene is described by a "self-consistent field" equation which can be written in different forms that are named "global illumination equation", "rendering equation" [5] etc.: there is some light field in the scene; it illuminates its surfaces which scatter it. This transformation of the illumination I (light that falls onto a surface) into luminance L (light that goes away from a surface point $\boldsymbol{x}$ in direction $\boldsymbol{v}$) is described by the surface BDF f:

$$L(\boldsymbol{v};\boldsymbol{x}) = \int f(\boldsymbol{x};\boldsymbol{v},\boldsymbol{v}')I(\boldsymbol{v}',\boldsymbol{x})d^2\boldsymbol{v}'$$

Then this light, emitted *from* the surface, propagates over the scene and eventually illuminates its surfaces.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

This transformation of the luminance (light going away off a surface) into illuminance (light incident onto another surface point) is described by the transport operator. The explicit form of the scattering and transfer operators is inessential here so we use a compact notations:

$$L = \hat{F}I \quad (1)$$

$$I^{(i)} = \hat{T}L \quad (2)$$

Notice that here $\hat{F}$ includes *only diffuse* scattering and thus the "direct" component also contains light that underwent pure specular transformation.

Notice L is **the luminance of the scene surfaces**; the luminance of the **camera image** would be $\mathcal{L} = \hat{\mathcal{S}}L$ where $\hat{\mathcal{S}}$ describes the *pure* specular transformation between the camera and a scene surface (frequently $\hat{\mathcal{S}} = 1$). Here and below we shall only calculate the L. Its transformation into the camera image can be applied after that if needed.

The full illumination consists of indirect and direct components:

$$I = I^{(i)} + I^{(0)} \quad (3)$$

Combining the three above equations we arrive at the form of "rendering equation" by Kajiya [5] we shall use:

$$I = \hat{T}\hat{F}I + I^{(0)} \quad (4)$$

3. Iterative solution (Neumann series) of (4) and BMCRT

A widely used approach in computational optics is a combination of forward and backward ray tracing, when the forward part calculates illumination of the diffuse surfaces I and stores it as, say, photon maps [4], and then the backward part "converts" it in the camera image. Notice that the actual illumination we operate is noisy and this noise is then "transferred" to the image, and its final amplitude strongly depends on how the BMCRT part works.

The simplest way is that we trace rays from camera, terminating then at the first diffuse surface where the surface luminance under the full illumination ($= \hat{F}I$) is calculated and *added* to the pixel luminance. This provides estimate of the surface luminance (1), though the result is not perfect because the estimation of I from FMCRT is usually subjected to (spatial) filtering to reduce noise. Thus all fine illumination details such as highlights are usually lost.

Instead one can first apply the N-th iteration of (4) to the illumination, which leads to

$$L = \left(\hat{F}\hat{T}\right)^N \hat{F}I + \sum_{k=0}^{N-1} \left(\hat{F}\hat{T}\right)^k \hat{F}I^{(0)} \quad (5)$$

For the exact illumination field I this gives exactly the same result as the above simple $\hat{F}I$, but for the actual, noisy I the second form is frequently *better* because result in the lower level of the *image noise* thanks to the convolution with a power of $(\hat{F}\hat{T})$. The value of N is termed "**backward diffuse depth**", or **BDD** for short.

The term $\hat{F}I^{(0)}$ is the surface luminance under only the *direct* (including caustic) illumination. The indirect (diffusively scattered) illumination I is accounted only in the last hit point as $(\hat{F}\hat{T})^N\hat{F}I$.

The integral operators can be calculated using Monte-Carlo: we fire rays from camera, they hit a surface, are scattered (left operator $\hat{F}$), propagate the scene (operator $\hat{T}$; notice in the *backward* ray tracing the events counted from camera correspond to the operators *left to right*) and so on, until reach the N-th diffuse surface and terminate. In the k- th diffuse hit point (i.e. just *before* the k-th diffuse scattering) we calculate this surface luminance under the direct (and caustic) illumination ($\hat{F}I^{(0)}$) if $k < N$ and full illumination ($\hat{F}I$) if $k = N$. Then scale the result by the luminance transmission factor due to the specular transformation in $\hat{T}$, and *add* to the pixel luminance. The average over ensemble of camera ray converges to L.

4. How the BDD affects the noise level

Let us first consider a simple scene consisting of a diffuse plate illuminated by a cone light (which emits downwards) enclosed in a matte (diffuse) spherical shade. The camera sees only the plate while the lamp shade is out of the view area. If the lamp shade is small enough we can neglect interreflections.

Illumination of the plate is then entirely *indirect* (light from the diffuse sphere). In case of BDD=0 camera ray collects that illumination in the hit point. There is some moderate noise from the FMCRT (= from photon map), but no more, while for BDD=1 the situation becomes much worse.

Indeed, now in the first hit point (i.e. the plate) we take only the direct (and caustic) illumination which in this model is exactly 0. So *all* contribution to pixel luminance comes only from the *second* diffuse hit, which is only possible when the scattered camera ray hits the lamp shade. But this probability is very low, because the BDF (and thus the scattered light cone) is *wide* while the lamp shade is *small*. Therefore, it is very rare that camera ray brings any luminance to the image pixel, and this means high noise, see [2].

The case of BDD>1 is not better because it also requires camera ray hitting the lamp shade.

Therefore, in our model scene the best is BDD=0 while BDD>0 results in a *very* high noise.

Now let us replace the small lamp shade with something very large. The scene becomes the diffuse plate in the centre of the floor of a large room box with diffuse walls and ceiling illuminated by a grid of cone lights (emitting towards the nearest surface and not towards the room centre), so their light can not reach the plate directly. The floor outside the plate is black. The camera is inside the room and sees only the plate. The whole scene is as the one from Figure 1 *only with-*

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

out the light in lampshade. At last, unlike the previous model scene, the bottom plate has now a *sharp* (nearly specular BDF).

Like in the previous example, the plate's illumination is purely *indirect*. The light from the grid of cone emitters undergoes many inter-reflections and so creates a rather uniform illumination of the walls and ceiling, though it is not exactly constant and the areas near the light emitters are brighter.

In case BDD=0 we collect the indirect illumination (from FMCRT!) at the first hit point. Illumination comes from the large box which is illuminated rather uniformly, and thus makes a wide cone. Meanwhile, the BDFs is sharp, so it effectively senses only illumination in direction *close to the mirror reflection of the view ray*. Most of FMCRT rays are be effectively excluded and only few of them will contribute to the pixel luminance which leads to a strong noise.

In case BDD=1 the situation improves. At the first camera hit we collect only the direct (and caustic) illumination, but it is 0 there. After diffuse reflection by the plate the camera ray hits the box and there it takes the full (in fact, again indirect) illumination. Most of camera rays and most of FMCRT rays contribute here, so the noise is low, see the images in [2].

The larger values of BDD work *less good*. Since illumination is *indirect*, the contribution to pixel luminance comes only from the *last* hit point. It can be in the box or in the bottom plate with more or less close probabilities. When it is in the box, it is like for BDD=1 i.e. all works good. But when it hits a bottom plate, all is like for BDD=0 i.e. we take a wide cone illumination of a narrow BDF which results in strong noise.

5. When all ``is worse''. Introducing the quasi-specular method

The model scenes from Section 4 admit an optimal BDD for which the noise is low. But there are scenes in which whatever BDD we choose the image is highly noisy.

An example is a *combination* of the two model scenes from Section 4: the bottom plate now has BDF which is a sum of a wide (nearly Lambert) and a sharp (nearly specular component). To see which BDF does which contribution to the image, the Lambert component is pure red and the sharp component is nearly pure green;[1] their integral reflectance is 50% and 40%, respectively. Indirect illumination comes both the room walls and ceiling (illuminated by an array of cone lights), and also by a small "central" lamp shade with a cone light inside, see Figure 1.

Unlike the first scene from 4 the light from the lamp shade also results in interreflections, but when the bottom plate is small (recall the light which hits the room floor outside it is absorbed), it is inessential.

[1] Its normalized color is $(0.1, 1, 0.1)$

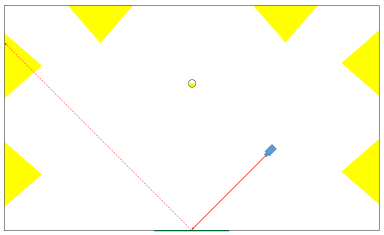


Figure 1. Combined model scene: indirect illumination both from the lamp shade and the room walls and ceiling. This is the side view in correct scale. The yellow cones show emission of lights; the camera view ray and its specular reflection are drawn with red arrows. The circle around the central light is the diffuse lamp shade; the yellow cone inside it mark its light emission (downwards). The thick green strip is the bottom plate.

Repeating the analysis from Section 4 one finds that BDD=0 is bad because of high noise from illumination of a narrow BDF part from the box and BDD=1 is bad because of high noise from illumination of a Lambert BDF from a small lamp shade, see Figure 2. Larger BDDs are not better.

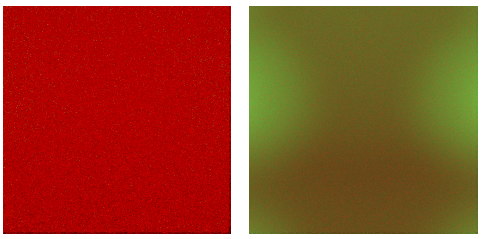


Figure 2. Camera images for scene from Figure 1, calculated during the same time (3000 sec) for BDD=0 (left) and BDD=1 (right). For the area of 100x100 pixels around the centre, the average RGB color is: $(150, 27, 2.71)$ and $(150, 31.6, 3)$. The noise level (relative to the photometric luminance) is 1000%, and 200% respectively.

This, however, does not work [1] which has been proved in [2]. That is, just making the near-specular and off-specular rays to collect illumination according to the ray's BDD produces wrong luminance, and one must apply a more sophisticated selection of the illumination components collected in each hit point.

The key idea of the approach though remains the same: the nearly specular narrow BDF is treated specially, not as a "genuine diffuse"; but in a sense closer to specular ones. Such narrow BDFs (and the whole method of their treatment in BMCRT) are thus named "quasi-specular". The result of calculation of the same model scene with this method is shown in the rightmost panel of Figure 2, see Section 9.1 for more explanations.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

6. Operator series in presence of quasi-specular BDFs

Now let us come to the formal derivation of what to do with a backward ray when the diffuse BDF is subdivided into the "genuine diffuse" and "quasi-specular" components:

$$\hat{F} = \hat{F}_d + \hat{F}_{qs}$$

The separation can be *arbitrary* (although some separations are advantageous as concerning the noise level and some are not), i.e. the calculated luminance all the same converges to the exact value.

Illumination is also subdivided into three components: **direct** (that was not scattered at all or scattered by pure specular surfaces), **quasi-caustic** (scattered at least once by a specular BDF and never by a diffuse BDF) and **indirect** (scattered at least once by a diffuse BDF and any times by a specular BDF or quasi-specular), i.e.

$$I = I^{(0)} + I^{(qc)} + I^{(i)}$$

Luminance of a surface point $\boldsymbol{x}$ is then:

$$L(\boldsymbol{v}, \boldsymbol{x}) = (\hat{F}_d I)(\boldsymbol{v}, \boldsymbol{x}) + (\hat{F}_{qs} I)(\boldsymbol{v}, \boldsymbol{x})$$

The above luminance expression can be written as

$$L = \hat{F} I = \hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right) + \hat{F}_{qs} I^{(i)} + \hat{F}_d I \qquad (6)$$

Substituting our decomposition of I into the global illumination equation (4) one arrives at

$$I^{(qc)} + I^{(i)} = \hat{T}\hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right) + \hat{T}\hat{F}_{qs} I^{(i)} + \hat{T}\hat{F}_d I$$

In the r.h.s., the term $\hat{T}\hat{F}_{qs} I^{(i)} + \hat{T}\hat{F}_d I$ comprises light that underwent at least one *diffuse* scattering while the term $\hat{T}\hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right)$ comprises light that underwent only specular scattering. In view of the decomposition into diffuse and quasi-caustic illumination, this means

$$I^{(i)} = \hat{T}\left(\hat{F}_d I + \hat{F}_{qs} I^{(i)}\right) \qquad (7)$$

$$I^{(qc)} = \hat{T}\left(\hat{F}_{qs} I^{(qc)} + \hat{F}_{qs} I^{(0)}\right) \qquad (8)$$

from what it follows that

$$I^{(i)} = \left(1 - \hat{T}\hat{F}_{qs}\right)^{-1} \hat{T}\hat{F}_d I \qquad (9)$$

$$I^{(qc)} = \left(1 - \hat{T}\hat{F}_{qs}\right)^{-1} \hat{T}\hat{F}_{qs} I^{(0)} \qquad (10)$$

We assume that if the backward ray underwent scattering by the quasi-specular component, this does *not* increment the diffuse event counter so the ray does not terminate and shall derive which illumination components must taken at which hit points so as "to be compatible" with the above behavior, i.e. so that the mathematical expectation of the BMCRT luminance to match the exact value.

Combining Eq. (6) with Eq. (9) and Eq. (10) we after some tedious yet trivial transformations see that the surface luminance from BMCRT with BDD=N is

$$\begin{aligned} L = {} & \hat{F}_{qs}\left(I^{(0)} + I^{(qc)}\right) \\ & + \sum_{k=0}^{N-1}\left(\left(1-\hat{Q}\right)^{-1}\hat{F}_d\hat{T}\right)^k \left(1-\hat{Q}\right)^{-1} \\ & \times \hat{F}_d(I^{(0)} + I^{(qc)}) \\ & + \left(\left(1-\hat{Q}\right)^{-1}\hat{F}_d\hat{T}\right)^N \left(1-\hat{Q}\right)^{-1}\hat{F}_d I \end{aligned} \qquad (11)$$

where

$$\hat{Q} \equiv \hat{F}_{qs}\hat{T} \qquad (12)$$

There is also an **alternative** form, which gives exactly the same result for the *exact* illumination while for a real noisy illumination may give different (in noise level) result:

$$\begin{aligned} L = {} & \left(1-\hat{Q}\right)^{-1}\hat{F}_{qs} I^{(0)} \\ & + \sum_{k=0}^{N-1}\left(\left(1-\hat{Q}\right)^{-1}\hat{F}_d\hat{T}\right)^k \left(1-\hat{Q}\right)^{-1} \\ & \times \hat{F}_d(I^{(0)} + I^{(qc)}) \\ & + \left(\left(1-\hat{Q}\right)^{-1}\hat{F}_d\hat{T}\right)^N \left(1-\hat{Q}\right)^{-1}\hat{F}_d I \end{aligned} \qquad (13)$$

Detailed derivation can be found in [2].

7. Integration by paths for BMCRT

Expanding $(1-\hat{F}_{qs}\hat{T})^{-1}$ in Eqs. (11) via Neumann series we see that e.g. for BDD=2

$$\begin{aligned} L = {} & \hat{F}(I^{(0)} + I^{(qc)}) \\ & + \sum_{k=1}^{\infty} \hat{Q}^k \hat{F}_d (I^{(0)} + I^{(qc)}) \\ & + \sum_{k,m=0}^{\infty} \hat{Q}^k \hat{F}_d \hat{T} \hat{Q}^m \hat{F}_d (I^{(0)} + I^{(qc)}) \\ & + \sum_{k,m,n=0}^{\infty} \hat{Q}^k \hat{F}_d \hat{T} \hat{Q}^m \hat{F}_d \hat{T} \hat{Q}^n \hat{F}_d I \end{aligned} \qquad (14)$$

In this expression a term like $\hat{Q}^k \hat{F}_d \hat{T} \hat{Q} \hat{F}_d (I^{(0)} + I^{(qc)})$ means that:

1. The ***final*** transformation ("before camera") light is subjected to $k \geq 0$ quasi-specular transformations with corresponding transport term, i.e. $\hat{Q}^k = (\hat{F}_{qs}\hat{T})^k$
2. ***Before*** that (i.e. "further from camera") light is subjected to the *diffuse* transformation with corresponding transport term, i.e. $\hat{F}_d\hat{T}$
3. ***Before*** that (i.e. more "further from camera") are light is subjected to $m \geq 0$ quasi-specular transformations with corresponding transport term, i.e. $\hat{Q}^m = (\hat{F}_{qs}\hat{T})^m$
4. And all the above is applied to the convolution of the direct and quasi-caustic illumination with the *diffuse* BDF i.e. to $\hat{F}_d\left(I^{(0)} + I^{(qc)}\right)$.

The action of the integral operators $\hat{F}_d$ and $\hat{F}_{qs}$ can be estimated with the BMCRT, tracing ray from camera. Then the first ray transformation events correspond

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

to the *leftmost* operators in a product and the last ray events — to the *rightmost* operators. That is, our term $\hat{Q}^k\hat{F}_d\hat{T}\hat{Q}\hat{F}_d(I^{(0)} + I^{(qc)})$ is estimated from the rays that **first** underwent $k \geq 0$ quasi-specular events, **then** one diffuse event, then $m \geq 0$ quasi-specular events and **after** that take the luminance of the *diffuse* BDF under direct and quasi-caustic illumination (i.e. the diffuse illumination part is ignored).

Detailed derivation can be found in [2].

The other terms in Eq. (14) can be interpreted and then estimated with BMCRT similarly. This leads to the following algorithm of processing BMCRT rays:

- before (and including!) the first not pure specular event we take $\hat{F}\left(I^{(0)} + I^{(qc)}\right)$;
- after the first quasi-specular event and up to the 2nd diffuse event we take $\hat{F}_d\left(I^{(0)} + I^{(qc)}\right)$;
- *after* the 2nd diffuse event we take $\hat{F}_d I$
- at the 3rd diffuse event we stop

The case of another BDD, as well as the alternative form i.e. Eq. (11) are considered similarly. When applying BMCRT, we trace rays from camera so that they undergo BDD+1 diffuse events (unless are absorbed prematurely) and then stop. Whenever a ray hits a surface which has some diffuse or quasi-specular BDF, the ray takes the convolution of *some part* of the BDF with *some part* of illumination, as summarized in the Table 1:

Table 1. In which hit points and for which illumination component the "genuine diffuse" BDF part $\hat{F}_d$ is used instead of the full $\hat{F}$. "QS" mean "quasi-specular".

	Main variant	**Alternative variant**
Direct, caustic	After the 1st QS event	After the 1st diffuse event
Quasi-caustic		Always
Indirect	Always	

8. Volumetric scattering

8.1 Standard method

Here all works like for surface scattering. Suppose for example that the camera is inside a turbid medium and BDD=1. Then camera ray propagates in the medium and when it undergoes the first volumetric scattering, it takes *direct* and *caustic* illumination and "convolves" it with the phase function. In the second volumetric scattering it takes *full* illumination and "convolves" it with the phase function.

Let us consider a model scene in which camera looks through a layer of turbid medium onto some illuminated object. Let also the turbid medium is absorption-free and has a large scattering coefficient, so that camera ray undergoes many (volumetric) scattering events before it penetrates the medium and reaches the object. Let, at last, phase function be sharp enough so that each scattering changes the ray direction only slightly.

For the "standard" method with a small BDD, the camera ray terminates close to the ray origin (i.e. camera), and the ray does not reach the object. Also, since the phase function is narrow, its convolution with that illumination which came from the scene surfaces results in strong noise, cf. Section 4.

We must set a *large* BDD so that camera ray to *leave* the medium, reach a scene surface behind and also not to collect indirect illumination in the volumetric scene points (which due to sharp phase function creates strong noise). But then we must keep in memory all the volumetric scattering points (because they still collect direct and caustic illumination), and this is usually too expensive.

8.2 "Quasi-specular medium"

Applying the quasi-specular approach to the volumetric scattering greatly improves the situation with the above scene.

Suppose that the scene *surfaces* are not quasi-specular. Let also the *whole* phase function be treated as quasi-specular, i.e. its "genuine diffuse" part $\hat{F}_d = 0$. Then, until the camera ray travels inside the medium, it undergoes only quasi-specular events and thus does not increment the "diffuse counter". As a result it penetrates the medium layer and reaches for a scene surface.

Indirect illumination now is the light after diffuse scattering by scene surfaces; *quasi-caustic* is the light which underwent at least one volumetric scattering (and any specular events, but no diffuse surface scattering).

Since $\hat{F}_d = 0$, the "main variant" from Table 1 implies that:

- the *indirect* illumination (=from scene surfaces) is effectively *ignored* inside the medium
- the *direct, caustic* and *quasi-caustic* illumination is taken only up to the *first* volumetric scattering, i.e. in fact for the *first* volumetric event only.

As a result, for *any* BDD the camera ray reaches the scene surfaces behind the medium layer; only *one* volumetric ray event (the first one) is remembered; there is no noisy convolution of the sharp phase function with the indirect illumination from scene surfaces.

There is, though, the convolution of phase function with the direct, caustic and scattered by the medium light. But they all do not have a *wide* angular distribution that would result in strong noise.

The "alternative" variant from Table 1 is worse because the direct and caustic illumination *will be* taken "up to the first *diffuse* event" and since the volumetric scattering is quasi-specular, that happens only *after* the camera rays leaves the medium. We must therefore

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

store *all* the ray scattering events inside the medium to take direct and caustic illumination in that points which is expensive in memory.

9. Results

9.1 Surface case

The model scene is from Figure 1. Calculation for the quasi-specular mode were performed for BDD=0 and the same other conditions as those for the "standard mode". The results for the standard mode are in Figure 2 and for the quasi-specular mode in Figure 3.

One can see the noise level is about threefold lower than for the best case without the quasi-specular method. And, which is more essential, the calculated image has a better estimation of color: for BDD=0 the image looks mainly red with rare bright green dots; for BDD=1 it looks mainly green with rare bright red dots, and it is only with quasi-specular method that we see a "mixture" of red and green. In the meanwhile the *averages* obtained in the three simulations are about the same, which means that the result e.g. for the "standard BDD=1" just has very bright red dots which "on average" would give the correct value. That is, in the standard method achieves the correct average by very rare very bright peaks, i.e. the worst sort of noise.

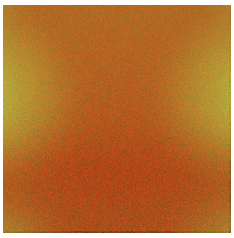


Figure 3. Camera image for scene from Figure 1, calculated for quasi-specular method with BDD=0. For the area of 100x100 pixels around the centre, the average RGB color is $(150, 31.4, 3)$. The noise level (relative to the photometric luminance) is 76%.

9.2 Turbid medium

The model scene consisted of a plane parallel plate of thickness 3 mm laid upon a paper sheet with a chessboard-like texture, illuminated by a self-emitting sphere above it. The plate's medium has refraction index 1.5 and scattering coefficient 7.5 mm^{-1}. The phase function is the Henyey-Greenstein one [7] with $g = 0.9$. The images calculated during the same time (and other settings) are presented in Figure 4; one can see that in the "standard mode" the boundaries of the texture squares seen through the plate are *sharp* while in reality they must blur. In the image calculated with the quasi-specular representation of the phase function these boundaries are smoothed (as they should be).

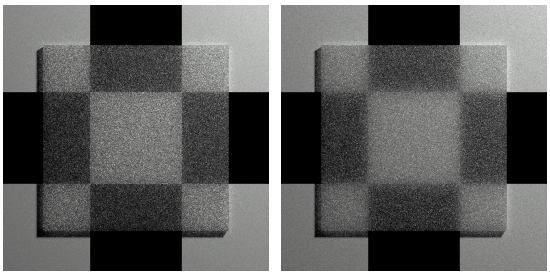


Figure 4. Camera images for the scene with a plate of turbid medium laid upon a chessboard-like texture. The left image is the "standard mode" of volumetric scattering, when it is treated as diffuse. The right panel is when the volumetric scattering is purely quasi-specular. In both cases BDD=1.

Acknowledgments

The work was partially supported by RFBR, Grant 18-01-00569.

10. References

[1] Д.Д. Жданов, С.В. Ершов, А.Г. Волобой. Адаптивный выбор глубины трассировки обратного луча в методе двунаправленной стохастической трассировки лучей. In *Proceedings of 25-th International Conference on Computer Graphics and Vision, September 22-25, 2015.*, pages 44–49, Protvino, 2015.

[2] S.V. Ershov, D.D. Zhdanov, A.G. Voloboy, M.I. Sorokin. Treating diffuse elements as quasi-specular to reduce noise in bi-directional ray tracing. *Preprint of the Keldysh Institute of Applied Mathematics*, No. 122, doi:10.20948/prepr-2018-122-e, 2018.

[3] P. Christensen, H.W. Jensen. High quality rendering using ray tracing and photon mapping. In *ACM SIGGRAPH 2007 Courses*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.

[4] H.W. Jensen. Global illumination using photon maps. In *Proceedings of the Eurographics Workshop on Rendering Techniques '96*, pages 21–30, London, UK, UK, 1996. Springer-Verlag.

[5] J.T. Kajiya. The rendering equation. In *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '86, pages 143–150, New York, NY, USA, 1986. ACM.

[6] G. Humphreys, M. Pharr. *Physically Based Rendering, Second Edition: From Theory To Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2010.

[7] G.C. Pomraning. On the Henyey-Greenstein approximation to scattering phase functions. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 39(2):109 – 113, 1988.

[8] H.W. Jensen T. Hachisuka. Stochastic progressive photon mapping. In *ACM SIGGRAPH Asia 2009 Papers*, SIGGRAPH Asia '09, pages 141:1–141:8, New York, NY, USA, 2009. ACM.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Mixing ordinary and Markov chain Monte Carlo rendering techniques

V.A. Frolov[1],[2]; V.A. Galaktionov[1]; V.V. Sanzharov[3]
[1]Keldysh Institute of Applied Mathematics RAS, Moscow, Russia;
[2]Moscow State University, Moscow, Russia;
[3]Gubkin Russian State University of Oil and Gas, Moscow, Russia;

In this work we propose way of combining ordinary and Markov chain Monte Carlo rendering techniques in image space. We used per-pixel mask to separate pixels (which we want to run Markov chain on) from the rest of the image. The mask was obtained from the ordinary Monte Carlo noise analysis. The proposed method was tested with combination of Multiplexed Metropolis Light Transport as Markov chain technique and two ordinary Monte Carlo rendering techniques - Instant Bidirectional Path Tracing and Light Tracing. As a result, our method allows us to get better accuracy in comparison to ordinary Monte Carlo and a better visual perception in comparison to Markov chain Monte Carlo techniques with the same rendering time.

Keywords: *Multiplexed Metropolis Light Transport, Markov chain Monte Carlo, bidirectional path tracing*

1. Introduction

Light transport algorithms have progressed significantly last 10 years. The evolution of light transport algorithms proceeded in two parallel ways: Ordinary Monte Carlo (OMC) and Markov Chain Monte Carlo (MCMC). Both ways have their own advantages and disadvantages.

2. Previous work

2.1 Ordinary Monte Carlo and MIS

OMC way includes such methods as Path Tracing (PT), Light Tracing (LT) and Bidirectional Path Tracing (BPT) [1, 15, 18], Photon Mapping (PM) [6, 9], Bidirectional Photon Mapping (BDPM) [20, 22], Vertex Connection and Merging (VCM) [4, 7] and other. These methods are often used in industry but they are not efficient for complex lighting phenomena in general (such efficiency is called «robustness» [18]).

Most robust OMC methods are based on Multiple Importance Sampling (MIS) technique which has strong mathematical foundation for variance reduction. The main idea of MIS is to use many different ways of sampling lighting integral with a posteriori weighing samples according to their probability density.

Unfortunately, Multiple Importance Sampling has strong disadvantage: the more robust algorithm becomes, the less average rendering speed it gains. This is the result of weighting samples since usually among many weights only a few are significantly different from zero [15]. At the same time sampling with many different strategies is expensive: N^2 shadow samples in BPT or density estimation on each bounce in BDPM are such examples.

2.2 Markov Chain Monte Carlo

In contrast to MIS (which samples proportionally to several parts of integrand by each strategy), Markov chain based methods [10, 11, 19, 21] create samples proportional to the final answer — lighting integral itself. This is achieved by representing lighting integral as multidimentional function projected to the image plane: $F(x, y, r_0, r_1, ...r_n) \xrightarrow{project} F(x, y)$. With such representation Markov chain places more samples (via Metropolis algorithm) in more complex regions of multidimentional space automatically and thus greatly reduces variance.

It should be noted that Markov chains and MIS nevertheless should be used together. The key to success here is to construct good integration space in which Metropolis algorithm and MIS will strengthen but not compete with each other. This is the main idea of Multiplexed Metropolis Light Transport (MMLT) algorithm [5]. MMLT uses Markov chain including for sampling strategy selection which completely eliminates discussed MIS disadvantage because low contribution strategies are rarely selected. More advanced Markov chain algorithms are constructed on top of MMLT framework or use similar principles [12, 14, 17].

Unfortunately, Markov chain methods have their own disadvantages. First, they have distinctive, visual unpleasent artifacts, especially in the begining of rendering. These artifacts are caused by high correlations of samples (excluding [17]) so that image shows the trajectories of Markov chains movements in the image space. Second, precision for short rendering time may be even less than for OMC methods due to the startup bias. Finally, Markov chain methods are inefficient for rendering direct light due to almost all samples would be placed at bright locations (for example even not noisy light surface!). MMLT amortizes last disadvantage in some degree (due to possibility of manual control chains number per each reflection depth) but does not solve same problem for caustics: large and bright areas always take most of computational resources even if these areas do not have true hard-sampling light transport phenomena.

2.3 Combining OMC and MCMC

At first, any practical implementation of MCMC rendering technique will separate direct and indirect

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

light calculations due to the reason discussed above. Second, normalization constant should be estimated and good starting points for Markov chain have to be selected to reduce startup bias. Thus, in practice, OMC and MCMC always work together.

Next, known ways of combining OMC and MCMC for the same lighting phenomena are implemented though Multiple Importance Sampling and have certain restrictions. Kelemen et al. in [11] has shown that large steps for Primary Sample Space Metropolis light transport (PSSMLT) can be used as a separate integration strategy and thus, OMC and MCMC can be combined together via MIS. Moreover, this method can be considered as a «shareware» due to the fact that large steps in PSSMLT will be computed anyway like other proposals. However, here are the limitations of Kelemen et al. approach (which we call «Kelemen MIS» further):

(1) Shareware Kelemen MIS, obviously, can not be used to combine rendering techniques with different basic algorithms. For example, if Metropolis algorithm is implemented on top of BPT framework, OMC sampling analogue will be exactly the same BPT. It does not sound as a serious disadvantage if both OMC and MCMC versions of the same rendering technique are good enough. However, if their efficiency differs significantly, Kelemen MIS becomes useless. This is exactly the case of Multiplexed Metropolis light transport where its OMC analogue is extremely inefficient due to end-points-only connections leaving few chances for random sample to hit high contribution region of multi-dimensional space.

(2) Bright areas (large caustics for example) are still oversampled by MCMC technique even if OMC can calculate them well enough. Moreover, OMC contribution in such areas will be almost discarded due to Kelemen MIS weights being evaluated from probability densities as shown further: $p_{omc} = 1, p_{mcmc} = I/b$, where I — MCMC image contribution function, and b — average image brightness. In this way if I is large enough, MIS weight for OMC contribution will tend to zero:

$$w_{omc} = \frac{p_{omc}^2}{p_{mcmc}^2 + p_{omc}^2} = \frac{1}{1 + (I/b)^2} \qquad (1)$$

Hachisuka et al. in [13] suggested combining SPPM and MLT as two completely different algorithms via machine learning. Training sets were obtained by classifying paths on their «path grammar». Resulting improvements in precision and visual perception turned out small. This is quite expected because such blending didn't go far from MIS ideas — use 2 strategies/algorithms (which simply doubles computational resources!) and then select best of them by weight. Moreover, both MLT and SPPM in general are good for the same lighting phenomena, so, their blending has limited application.

In this way, more efficient and general method of combining OMC and MCMC for rendering became our research area.

3. Suggested approach

Our key idea is similar to MMLT approach — let Markov chain to select best algorithm statistically. To combine OMC and MCMC we propose to use «a priori blending function» $\alpha(x, y)$ which marks image areas that can not be evaluated well with OMC and uses it as expression 2 shows:

$$Color_{res} = D_{OMC}(x, y) + I_{OMC}(x, y) * (1 - \alpha(x, y)) + I_{MCMC\alpha}(x, y). \qquad (2)$$

where (x, y) are screen space coordinates, $D_{OMC}(x, y)$ - direct light computed by OMC, $I_{OMC}(x, y)$ - indirect light computed by OMC and $I_{MCMC\alpha}(x, y)$ - priory weighted indirect light computed by MCMC. Note that $I_{MCMC\alpha}(x, y)$ is not multiplied with $\alpha(x, y)$ because it's already constructed proportional to $\alpha(x, y) * I(x, y)$ by Markov chain.

3.1 Weight functions details

We build $\alpha(x, y)$ by OMC noise analysis:

1. run OMC for short time (64-128 samples per pixel in average) to compute noisy image approximation;
2. compute for each pixel its difference with median in 5x5 or 7x7 window;
3. sort all pixels by this difference, take 10% most noisy to assign them values in range [0,1]. This assignment is done with sigmoid function relative to median value of these 10% pixels so that noisy pixels tend to one and not noisy tend to zero.
4. split the image into segments relied on the information from G-buffer (per pixel depth, normals, material id);
5. fill each segment with the value of it's most noisy pixel (fig. 1);
6. clamp values to epsilon, i.e. $\alpha(x, y) := max(\alpha(x, y), \epsilon)$;
7. assign everything greater than 0.5 to 1 and less than 0.5 to ϵ (so that our function became binary).

Let us discuss some of these items in details. Although many advanced methods of noise detection was developed recently [2, 3, 16, 23], median difference was enough for our experiments due to the observation that if Ordinary Monte Carlo fails for some phenomena, it gives significant ejection/spike. Segment filling was used based on similar considerations: if some region (for example, glossy or mirrored object) gives at least one strong spike, this region have complex light transport phenomena and it has to be computed using MCMC technique.

Next, it should be noted that weigh function $\alpha(x, y)$ must not be zero in any region where indirect light $I(x, y)$ is not zero. This is essential for correct evaluation of premultiplied $I(x, y) * \alpha(x, y)$ by

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

MCMC due to normalisation constant is estimated over $I(x,y) * \alpha(x,y)$ for the whole image. Otherwise MCMC result will have inconsistent average brightness and color with the rest of image as shown at fig 2.

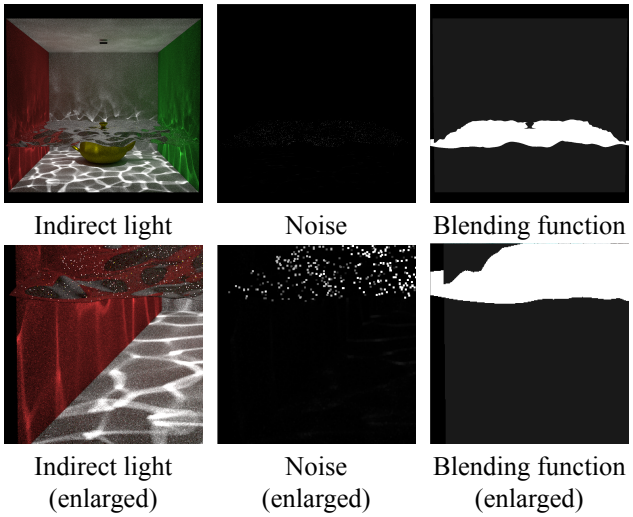


Figure 1. Constructing blending function from noise and G-buffer analysys.

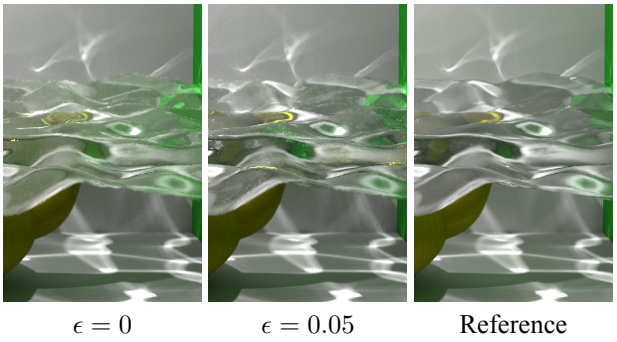


Figure 2. Inconsistent green color of water (computed by MCMC) when $min(\alpha(x,y),\epsilon)$ tends to zero in other (non water) regions on the image.

In this way we apply $\alpha(x,y) := max(\alpha(x,y),\epsilon)$ where ϵ is a parameter equal to 0.1 (so $\alpha(x,y)$ fits in range [0.1,1]). At last, we made our function binary — clamp everything greater than 0.5 to 1 and less equal 0.5 to ϵ. This shown visually more pleasant result in practice then blend of intermediate values for mirror and glossy objects.

Because our blending function became binary, it is neccesary to smooth border pixels to prevent aliasing. Fortunately, we have segments and their exact borders, so we apply FXAA algorithm for border pixels only to smooth them.

4. Implementation details and results

We tested our approach on 4 scenes (fig. 3) with MMLT as a Markov chain technique and two Ordinary Monte Carlo rendering methods - IBPT [1] and Light Tracing. All algorithms were implemented with our own C++ framework. As a result our method gives less error then Ordinary Monte Carlo (table 1) and has better visual perception in comparison to Markov chain Monte Carlo techiques with the same rendering time (fig. 4, 5, 6, 7). Our approach also outperforms Kelemen MIS combining (fig 8).

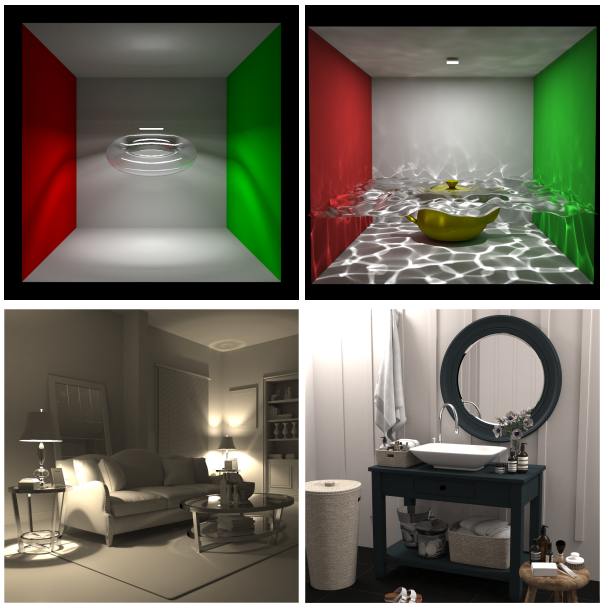


Figure 3. Tests scenes.

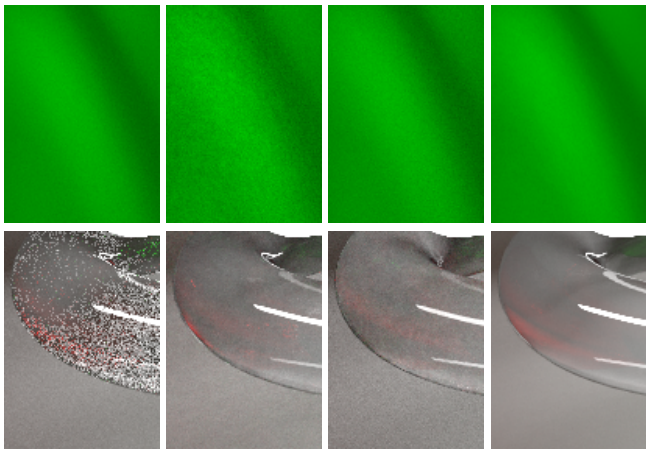


Figure 4. Comparison 1. Enlarged fragments of Cornell box with glass torus. 10 minutes.

5. Limitations

In practice we found that our approach is useful for scenes where true hard sampling lighting phenomena takes less then 30-40% of image pixels. Otherwise we recommend to render image completely with MCMC technique. This is quite expected because our method halves computational resources for both OMC and

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

MCMC methods for the fixed rendering time. Thus if OMC technique fails for most of image pixels it becomes useless. Fortunately, because we have $\alpha(x, y)$ in image space we can easily count per cent of complex pixels and switch to «MCMC only» rendering automatically.

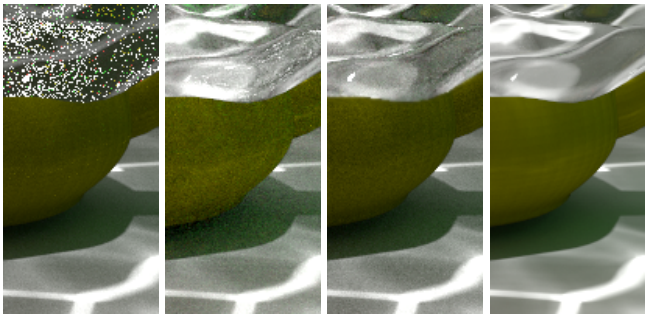


Figure 5. Comparison 2. Enlarged fragments of Cornell box with water. 10 minutes.

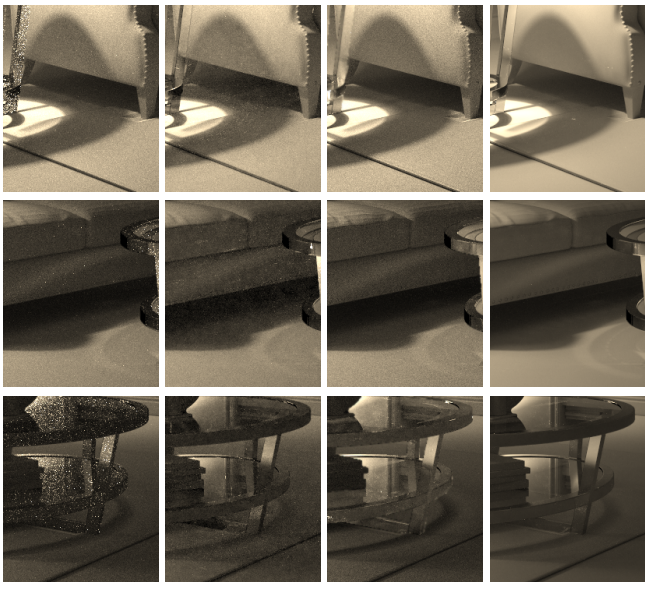


Figure 6. Comparison 3. Enlarged fragments of the room scene, 30 minutes.

Scene/Method	IBPT	MMLT	Ours	Ours combination
Thorus	15.4	7.28	12.5	IBPT + MMLT
Water	40.1	17.6	20.0	IBPT + MMLT
Room	33.5	15.6	28.1	LT + MMLT
Bathroom	58.8	32.7	45.4	LT + MMLT

Table 1. Mean Square Error (MSE); LDR space (0-255).

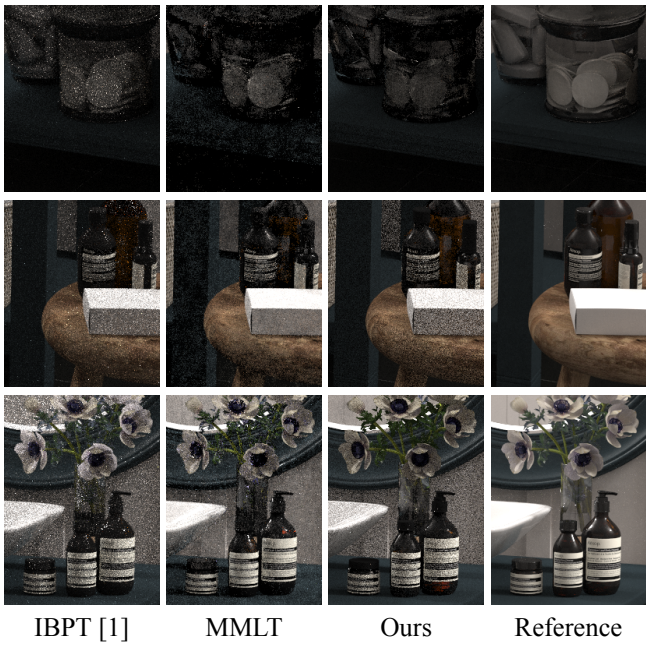


Figure 7. Comparison 4. Enlarged fragments of bathroom scene, 10 minutes.

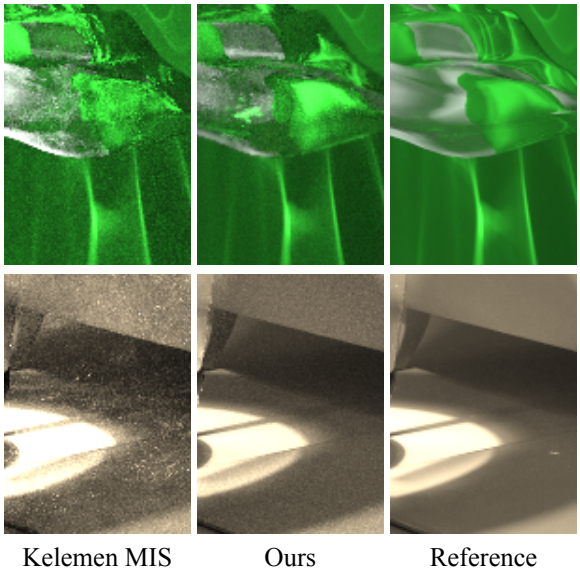


Figure 8. Comparison of our method to blending by Kelemen MIS weight [11].

6. Conclusions

Similar to previous approaches our algorithm halves computational resources for both blended methods (because we have to run 2 different algorithms with the same time). In contrast to previous work, we used a priory blending weight that allows us to evaluate complex image regions with MCMC more efficiently in comparison to rendering via MCMC without blending it with OMC. The indirect result of such approach is more pleasant view

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

(without MCMC artifacts) in other regions of image where we took result from OMC.

7. Acknowlegments

This work was sponsored by RFBR 16-31-60048 and 16-01-00552.

8. References

[1] Bogolepov D., Ulyanov D. *GPU-Optimized Bi-Directional Path Tracing.* // WSCG'2013. Conference on Computer Graphics, Visualization and Computer Vision.

[2] Constantin J., Bigand A., Constantin I., Hamad D. *Image noise detection in global illumination methods based on fast relevance vector machine.* // In Proceedings of the 12th international conference on Artificial Neural Networks: advences in computational intelligence. Springer-Verlag, Berlin, Heidelberg, 467-479.

[3] Granados M., Aydın T.O., Tena J.R., Lalonde J.-F., Theobalt C. *HDR image noise estimation for denoising tone mapped images.* // In Proceedings of the 12th European Conference on Visual Media Production (CVMP '15). ACM, New York, NY, USA, Article 7, 8 pages.

[4] Georgiev I., Křivánek J., Davidovič T., Slusallek P. *Light Transport Simulation with Vertex Connection and Merging.* // ACM Trans. Graph. (SIGGRAPH Asia 2012) 31, 6 (2012)

[5] Hachisuka T., Kaplanyan A.S., Dachsbache C. *Multiplexed Metropolis Light Transport* // ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2014

[6] Hachisuka T., Wann Jensen H. *Stochastic progressive photon mapping* // Proceeding SIGGRAPH Asia '09 ACM SIGGRAPH Asia 2009 papers Article No. 141

[7] Hachisuka T., Pantaleoni J., Wann Jensen H. *A path space extension for robust light transport simulation.* // Proceedings of ACM SIGGRAPH Asia 2012, Volume 31 Issue 6, November 2012 Article No. 191

[8] Jendersie K. Rohmer, F. Brüll, T. Grosch *Pixel Cache Light Tracing.* // Vision, Modeling, and Visualization (2017).

[9] Jensen H.W. *Global Illumination using Photon Maps.* // In Proc. of Eurographics Workshop on Rendering (EGWR) (1996), EGWR, Springer, pp. 21–30.

[10] Kaplanyan A. S., Hanika J., Dachsbacher C. *The Natural-constraint Representation of the Path Space for Efficient Light Transport Simulation.* // ACM Transactions on Graphics (Proceedings of SIGGRAPH) 33, 4, Article 102 (July 2014), 13 pages.

[11] Kelemen C, Szirmay-Kalos L., Antal G., Csonka F. *A Simple and Robust Mutation Strategy for the Metropolis Light Transport Algorithm* // EUROGRAPHICS 2002 / G. Drettakis and H.-P. Seidel. Volume 21, 2002, Number 3.

[12] Otsu H., Kaplanyan A. S., Hanika J., Dachsbacher C., Hachisuka T. *Fusing state spaces for markov chain Monte Carlo rendering* // ACM Trans. Graph. 36, 4, Article 74 (July 2017), 10 pages.

[13] Otsu H., Kinuwaki S. Hachisuka T. *Supervised Learning of How to Blend Light Transport Simulations.* // Preprint, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo (2017).

[14] Pantaleoni J.*Charted metropolis light transport.* // ACM Trans. Graph. 36, 4, Article 75 (July 2017), 14 pages.

[15] Popov S., Ramamoorthi R., Durand F., Drettakis G. *Probabilistic Connections for Bidirectional Path Tracing* // Eurographics Symposium on Rendering 2015, Volume 34 (2015), Number 4.

[16] Takouachet N., Delepoulle S., Renaud C., Zoghlamid N., Tavarese J.-M. *Perception of noise and Global Illumination: towards an automatic stopping criterion based on SVM.* // Computers and Graphics (2017).

[17] Tzu-Mao Li, Jaakko Lehtinen, Ravi Ramamoorthi, Wenzel Jakob, and Frédo Durand. 2015. Anisotropic Gaussian mutations for metropolis light transport through Hessian-Hamiltonian dynamics. ACM Trans. Graph. 34, 6, Article 209 (October 2015), 13 pages.

[18] Veach E. *Robust monte carlo methods for light transport simulation* // (1998) Ph.D. Dissertation. Stanford University, Stanford, CA, USA.

[19] Veach E., Guibas L. J. *Metropolis Light Transport.* // SIGGRAPH '97 Proceedings of the 24th annual conference on Computer graphics and interactive techniques. Pages 65-76.

[20] Vorba J. *Bidirectional Photon Mapping.* // Proceedings of CESCG 2011: The 15th Central European Seminar on Computer Graphics. Charles University, Prague, 2011.

[21] Wenzel J. *Light Transport on Path-Space Manifolds.* // Ph.D. Dissertation. Cornell University (2013).

[22] *Zhdanov D.D., Garbul A.A., Potemin I.S., Voloboy A.G., Galaktionov V.A., Ershov S.V., Sokolov V.G.* Photorealistic volume scattering model in the bidirectional stochastic ray tracing problem. Programming and Computer Software. 2015. Vol. 41. No. 5. pp. 295-301.

[23] Wang S.-S., Wu C.-H. *A new impulse detection and filtering method for removal of wide range impulse noises* // In Pattern Recognition, Volume 42, Issue 9, 2009, Pages 2194-2202, ISSN 0031-3203

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Precomputed procedural content level of detail

V.V. Sanzharov[1], V.A. Frolov[2,3]
vsan@protonmail.com|vova@frolov.pp.ru
[1]Gubkin Russian State University of Oil and Gas, Moscow, Russia;
[2]Keldysh Institute of Applied Mathematics RAS, Moscow, Russia;
[3]Moscow State University, Moscow, Russia

In this paper we address the problem of determining the detail level of precomputed procedural content for photo-realistic rendering. We propose a rasterization prepass with an algorithm based on mipmapping to calculate the detail levels. Suggested solution is applied to computing resolutions of procedural and image textures and mesh subdivision levels.

Keywords: *procedural textures, photo-realistic rendering, level of detail.*

Introduction

Procedural content is a basis of todays cinema and VFX (visual effects) industrial pipeline because it allows to reduce the cost of manual content creation. This is also important in recent applications of photo-realistic rendering to the synthesis of images and video sequences for the purpose of using them as training data for artificial intelligence algorithms in computer vision [25] where manual content creation is infeasible because of the dataset sizes.

The procedural approach in material modeling is the synthesis of texture maps, which can be used to specify the values of any material model parameters, geometry (for example, displacement maps), and light sources. Methods of procedural generation are often classified as implicit or explicit [5]. Explicit methods calculate the whole texture in advance. These textures can be used further by the lighting modeling system as conventional images. Implicit methods define the texture as a function of some arguments, for example, from the coordinate in the texture or world space and calculate it "on the fly". Basically, the algorithm responds with a value to a query about particular sample. This approach implies the creation of some algorithmic description of the desired visual result.

Procedural texture synthesis overview

Implicit texture synthesis techniques have been used for quite a long time, one of the first major breakthroughs being Perlin noise algorithm [23] back in 1985, and are continually being developed (for example [7, 14, 20]). These methods have several advantages - texture maps are not stored in memory and don't have fixed resolution (the texture is infinite and has no seams). Also, textures can be parametrized with arbitrary values (for example, world space position) and therefore applied to geometry without uv-unwrapping. By changing the input parameters it's possible to get a large number of different texture variants. Disadvantages include: the complexity of the process of algorithmic formalization, the need for careful manual adjustment of parameters to achieve a desired result and independent texture samples evaluation. This restricts the generalization possibilities and therefore limits the variability in resulting textures. Still, implicit methods are widely used in computer graphics and are implemented in many software products (Houdini, Blender, Maya, 3ds Max, etc.).

In explicit texture synthesis texel values are interdependent and therefore can't be evaluated independently. A detailed review of early explicit methods is presented in [26]. Example based and most statistical texture synthesis algorithms are classified as explicit. Example based methods are being actively developed by improving existing algorithms [10, 11], applying neural networks [6, 18] and combining existing algorithms with neural networks [4, 17]. One of the main advantages of these methods is that they don't suffer from the problem of formalizing the desired result which is represented by an example image passed as an input to the algorithm. Among the shortcomings is a considerable amount of computations usually required for these methods, which directly depends on the output texture resolution. As reported in [13] running time for texture synthesis algorithms ranges from seconds to minutes and may even reach hours for large resolutions.

Existing problems

Resolution influences the size of the texture in memory, time needed to execute texture synthesis algorithm and image quality in rendering. Thus, for photo realistic rendering it is crucial to determine the texture resolution which will provide the best perceived quality with the smallest possible size in memory and computation time. Another problem arises when rendering system needs to support existing implementations of implicit texture synthesis methods in closed-source software like Maya and 3ds Max. This usually means that the renderer should call a virtual function provided by the software API to get the texture sample value. So, if the address space where the renderer is running differs from the address space of the procedural texture implementation (for example, the renderer runs on the GPU or on a different computer), it's simply not possible to query these proce-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

dural textures during rendering like it should be done with implicit texture synthesis. Therefore, to support them the renderer developers need to take on a different approach. The first option is to implement analogous procedural textures in the rendering system and match their input parameters with existing software. This has to be done for every product the rendering system wishes to integrate into. And the second (and less expensive) option is to treat these procedural textures as explicit - that is precompute them before rendering. And again, this leads to the same problem - the need to determine the resolution.

In this paper we suggest an approach to estimate resolution for procedural textures precomputation for photo-realistic rendering systems. Proposed solution can also be used to resize bitmap textures and with some modifications applied to other preliminary computations like estimation of subdivision level for meshes subject to displacement.

Related work

Existing approaches for the most part are concerned with level of detail (and therefore resolution) estimation for image textures in real-time rendering. Conventional solutions to the problem of texture resolution selection are based around mipmapping [28] and involve storing several copies of an image with different resolutions - mipmap pyramid, out of which appropriate mip level is selected at rendering time usually based on the projected size of a textured object. There are methods enabling parallel synthesis of different miplevels for some of the explicit procedural textures [8, 16, 27]. Some methods focus on performing relatively small amount of preliminary computations which are then used to speed up subsequent synthesis of the whole image pyramid [3]. An optimization to mipmapping called clipmapping [24] is based on the idea that for a large texture not all data of a mipmap is used in rendering of a particular frame so it is sufficient to only store parts of a mipmap which are potentially visible - clipped mipmap. More complex approach - virtual texturing [1, 15, 22] is based around splitting texture into uniform tiles and loading into video memory only the tiles needed to render a particular frame using visibility information. Virtual texturing also received hardware support [2] and is available in Vulkan API and as OpenGL ARB_SPARSE_TEXTURE extension [30].

Limitations of existing solutions

Existing solutions to determine the texture resolution focus on application to real-time rendering and are designed to be implemented as a part of a rendering pipeline. Since we need to estimate resolution for precomputed procedural textures before rendering, these approaches don't solve the problem, at least by themselves. The other challenge lies with the rendering pipeline integration - to implement any of these methods in an existing rendering system, the rendering system itself would obviously require changes which can be substantial. For example, when the mesh has more than one texture applied to it, using virtual texturing requires using texture atlases [21]. Since rewriting already established parts of a complex system is often undesirable, we consider this as a disadvantage of direct implementation of these methods in the renderer. There are also other possible problems with the approaches mentioned earlier like visual artifacts caused mainly by filtering [1, 22]. For some texture types, for example glossiness textures and normal maps, mipmap filtering can lead to significant losses in detail [29]. This is not acceptable in photo-realistic rendering since details are one of the foundations of a realistic image.

Suggested approach

The essence of the suggested solution is to separate all the necessary precomputations from the rendering system into a simple scene rasterization prepass followed by texture synthesis. This prepass would determine the resolution of the each texture in a scene from mip levels. And mip levels are computed using a slightly modified implementation of the approaches suggested in [1] and provided in OpenGL API specification [30]. Now let's consider the whole proposed process.

Implementation details

First of all, we will need to store function pointers and input parameters for all procedural textures used in the scene we are going to render. This way we'll be able actually synthesize them after we determine the needed resolution.

The next step is to rasterize the scene geometry with the information on the assigned materials. Rasterized scene allows to determine texture coordinates gradient for all textured objects and thus the mip level. The equations for gradient (1, 2) are similar to those for mip level calculation in [1] and OpenGL API specification [30].

$$G_x(uv) = \frac{R_rast_x}{R_rend_x} * tex_res * \frac{\partial u}{\partial x} \quad (1)$$

$$G_y(uv) = \frac{R_rast_y}{R_rend_y} * tex_res * \frac{\partial v}{\partial y} \quad (2)$$

R_rast_x, R_rast_y – rasterization resolution, R_rend_x, R_rend_y – photo-realistic rendering resolution, tex_res – texture resolution, u, v – texture coordinates, x, y – rasterized image coordinates.

However, for procedural textures tex_res is unknown and it is actually the value we want to determine. So, let's assume for the moment that it is an arbitrarily high resolution which we want to reduce. Also we need to take into account ratio be-

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

tween rasterization resolution (R_rast_x, R_rast_y) and actual photo-realistic rendering resolution (R_rend_x, R_rend_y) since these can be different - rasterized image only serves a utility function and there is no need for it to be high resolution while photo-realistic rendering obviously can have arbitrary resolution.

The gradients are then used to compute the mip level mip_lvl (3) analogous to OpenGL specification [30] and [1]:

$$mip_lvl = log_2[max(G_x, G_y)] \quad (3)$$

After mip levels are calculated we find the maximum mip level for each texture in the scene - the same textures can have different materials mip levels in different materials and on different geometry. So, we basically retain only the maximum level from the mip pyramid for each texture. Required resolution R is then calculated with (4).

$$R = \frac{tex_res}{2^{mip_lvl}} \quad (4)$$

After substituting the mip_lvl in (4) with (3) and (1, 2) we get the following (5):

$$R = \frac{tex_res}{max(\frac{R_rast_x}{R_rend_x} * \frac{tex_res}{\frac{\partial u}{\partial x}}, \frac{R_rast_y}{R_rend_y} * \frac{tex_res}{\frac{\partial v}{\partial y}})} \quad (5)$$

It's obvious, that tex_res can be eliminated, yielding the final equation for the resolution (6):

$$R = max(\frac{R_rend_x}{R_rast_x} * \frac{1}{\frac{\partial u}{\partial x}}, \frac{R_rend_y}{R_rast_y} * \frac{1}{\frac{\partial v}{\partial y}}) \quad (6)$$

All that is left is to pass the obtained resolution to the stored texture synthesis functions.

The same procedure can be applied to the «ordinary» image textures to resize them to the 1-to-1 pixel-to-texel ratio for the sake of possible memory saves. It's straightforward to incorporate different maximum allowed mip levels for different texture types, for example, to force the reflection or bump textures to be generated in higher resolution than diffuse textures.

Limitations and special cases

There are several special cases which cannot be tackled by suggested approach directly. The first such case is a scene with transparent objects - with straightforward rasterization they can potentially occlude other objects and interfere with detail level calculation. This situation can be handled by performing two rasterization passes for detail level estimation - one for all non-transparent objects and the other for transparent objects.

The second more complex case concerns textured objects visible as reflections (possibly after multiple reflections/refractions). This is a very difficult situation for rasterization and multiple bounces are nearly impossible to handle. Within the framework of our approach it can be addressed heuristically - for a reflected object we can safely assume that it won't be larger than the screen resolution and resize the textures accordingly. Obviously, it's not very efficient but still allows for possible memory savings. Another heuristic would be to assign to not directly visible objects the level of detail computed for the largest visible reflective object (i.e. mirror) in the scene.

A promising way for solving the challenge of reflections more rigorously is using techniques like ray differentials [9] with an upcoming real-time ray tracing technology [31].

Application to other precomputed effects

Proposed solution can also be applied to other precomputed effects, in particular to compute subdivision level for meshes. Since in the essence, texture mip level and geometry subdivision level both represent the same notion of detail level. For example, in [19] clipmapping ideas were applied to mesh subdivision in terrain generation from a height map. The goal of precomputing mesh subdivision instead of evaluating it at render time is 1) to do the computation only once in the case of rendering on multiple machines and 2) to keep the rendering system implementation simpler. To utilize the rasterization prepass described earlier to calculate the per-mesh subdivision factor some modifications are required. First of all, the gradient should be calculated for the positions, not the texture coordinates. This way we can obtain the detail level per mesh. To get different subdivision levels on the same big mesh (like terrain), it needs to be split into regions, for example as proposed in [19].

Results and discussion

We tested our approach on 3 scenes (fig. 1) which use procedural and image textures. First we rendered scenes with image textures in their respective native resolution and all procedural textures with manually picked reasonable resolutions. Rendered images obtained this way served as references. Then, we rendered the same scenes with the same image and procedural textures but the rasterization prepass was executed to determine the resolution for textures before actual rendering. Scenes were rendered with our own C++/OpenCL path tracing implementation on the GPU. The comparison of the results with the references in terms of overall memory needed for textures (table 1) showed several times lower memory size with rasterization prepass. Visual perception comparison (fig. 2, fig. 3, fig. 4) shows no perceptible differences as confirmed by mean square error (MSE) values (table 2). We also tested our approach in application to computation of the subdivision level (fig. 5, fig. 6) using non-adaptive sqrt 3 [12] subdivision algorithm. First, terrain generated from height map was subdivided before displacement with the same level across all three parts (separate meshes). Next, modified rasterization prepass was used to determine the subdi-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

vision levels per mesh. Applying suggested prepass allowed to reduce the amount of memory needed for geometry from 185 Mb to 89 Mb while maintaining image quality.

Conclusion

Proposed approach allows for tangible memory savings compared to naive approach of simply synthesizing procedural textures with a same fixed resolution and passing images into rendering system as they are.

Scene, rendering resolution	Ref., Mb	Ours, Mb	Ratio
Arch, 1920 x 1080	276	156	1.77
Arch, 3840 x 2160	276	173	1.60
Bathroom, 1024 x 1024	521	142	3.67
Bathroom, 2048 x 2048	521	178	2.93
Alley, 1024 x 1024	527	129	4.09
Alley, 2048 x 2048	527	146	3.61

Table 1. Memory used by textures

Scene	MSE
Alley	1.67
Bathroom	3.59
Arch	3.02

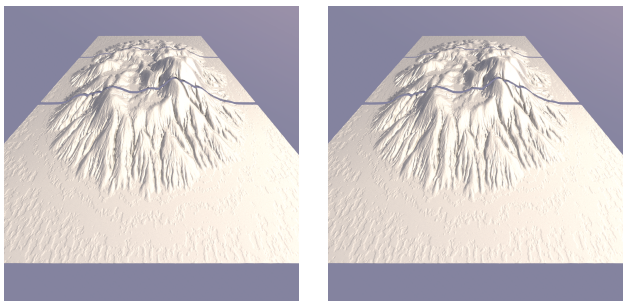
Table 2. Mean Square Error (MSE) for rendered images (computed after tone mapping, i.e. in low dynamic range)

Figure 1. Test scenes - Arch, Bathroom, Alley. Arch and Alley - primarily procedural textures, Bathroom - high resolution bitmaps

Figure 2. Reference vs With resized textures, Bathroom

Figure 3. Reference vs Difference vs With resized textures. Bathroom scene, in top row - bitmap diffuse texture, bottom row - bump map computed from bitmap

Figure 4. Reference vs Difference vs With resized textures. Alley scene, procedural noise-based texture on an object with one of the texture coordinates stretched out



geom. size in memory: 185 Mb 89 Mb

Figure 5. Terrain subdivision with uniform subdivision levels vs subdivision levels computed by prepass, terrain mesh generated from height map

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

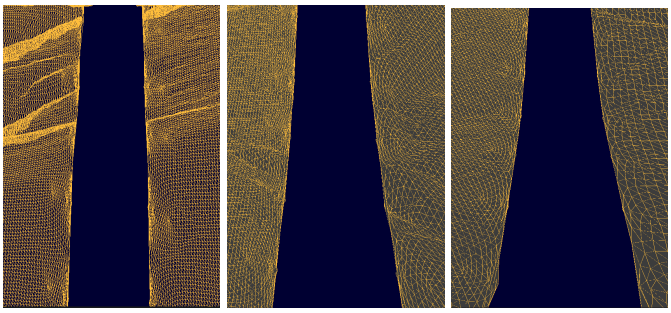


Figure 6. Neighboring mesh grids fragment. Left to right: no prepass applied - uniform subdivision levels; reduced by prepass - the closest and the middle meshes; reduced by prepass - the middle and the farthest meshes. The farther the mesh is from the camera - the smaller its subdivision levels and therefore more sparse mesh grid

Acknowledgments

This work was sponsored by RFBR 16-31-60048 «mol_a_dk».

References

[1] Barrett S. *Sparse virtual textures* http://silverspaceship.com/src/svt/

[2] Bilodeau B., Sellers G. and Illesland K. *Partially Resident Textures on Next-Generation GPUs* // Game Developers Conference, 2012.

[3] Dong Y., Lefebvre S., Tong X. and Drettakis, G. *Lazy Solid Texture Synthesis* // Computer Graphics Forum, 2008, 27: 1165-1174

[4] Dong J., Wang L., Liu J., Sun X. *A Procedural Texture Generation Framework Based on Semantic Descriptions* // arXiv:1704.04141, 2017

[5] Ebert D.S., Musgrave F.K., Peachey D., Perlin K., and Worley S. *Texturing and Modeling: A Procedural Approach.* // Morgan Kaufmann, 1998.

[6] Gatys L.A., Ecker A.S., Bethge M. *Texture Synthesis Using Convolutional Neural Networks* // arXiv:1505.07376, 2015

[7] Gilet G., Sauvage B., Vanhoey K., Dischler J.-M., and Ghazanfarpour D. *Local random-phase noise for procedural texturing* // ACM Trans. Graph., 2014, Vol. 33(6), Article 195

[8] Han C. et al. *Multiscale texture synthesis* // ACM Trans. Graph., 2008, Vol. 27(3), p. 51.

[9] Igehy H. *Tracing Ray Differentials* // Proceedings of ACM SIGGRAPH, 1999, pp. 179-186.

[10] Jamriška O., Fišer J., Asente P., Lu J., Shechtman E., and Sýkora D. *LazyFluids: appearance transfer for fluid animations* // ACM Trans. Graph., 2015, Vol. 34(4), Article 92

[11] Kaspar A., Neubert B., Lischinski D., Kopf J. *Self Tuning Texture Optimization* // Computer Graphics Forum 34(2), 2015

[12] Kobbelt L. *√ 3-subdivision* // Proceedings of ACM SIGGRAPH, 2000

[13] Kolář M. , Debattista K. and Chalmers A. *A Subjective Evaluation of Texture Synthesis Methods* // Computer Graphics Forum 36: 189-198, 2017

[14] Lefebvre L., Poulin P. *Analysis and synthesis of Structural Textures* // Proceedings of Graphics Interface 2000, pp. 77-86

[15] Lefebvre S., Darbon J., Neyret F. *Unified texture management for arbitrary meshes* // Technical Report RR5210-, INRIA, may 2004

[16] Lefebvre S., Hoppe H. *Parallel controllable texture synthesis* // ACM Trans. Graph., 2005, Vol. 24(3), pp. 777-786.

[17] Li C., Wand M. *Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis* // arXiv:1601.04589, 2016

[18] Li Y., Fang C., Yang J., Wang Z., Lu X., Yang M.-H. *Diversified Texture Synthesis with Feed-Forward Networks* // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, p. 266-274

[19] Losasso F., Hoppe H. *Geometry clipmaps: terrain rendering using nested regular grids.* // ACM Trans. Graph., 2014, Vol. 23(3), pp. 769-776.

[20] Maung D., Shi Y., Crawfish R. *Procedural textures using tilings with Perlin Noise.* // 60-65. 10.1109/CGames.2012.6314553

[21] Mayer A.J. *Virtual texturing* // Master's thesis, Institute of Computer Graphics and Algorithms, Vienna University of Technology, Oct. 2010.

[22] Mittring M. et al. *Advanced virtual texture topics* // ACM SIGGRAPH 2008, 2008, p. 23-51.

[23] Perlin K. *An Image Synthesizer* // Proceedings of ACM SIGGRAPH, pp. 287–296, 1985

[24] Tanner C.C., Migdal C.J., Jones M.T. *The clipmap: a virtual mipmap* // Proceedings of ACM SIGGRAPH, 1998, p. 151-158.

[25] Tsirikoglou A., Kronander J., Wrenninge M., Unger J. *Procedural Modeling and Physically Based Rendering for Synthetic Data Generation in Automotive Applications* // arxiv.org:1710.06270v2, 2017

[26] Wei L.-Y., Lefebvre S., V. Kwatra, G. Turk *State of the Art in Example-based Texture Synthesis* // Eurographics 2009, EG-STAR 2009

[27] Wei L. and Levoy M. *Order-independent texture synthesis* // Tech. Rep. TR-2002-01, Stanford University CS Dept.

[28] Williams L. *Pyramidal parametrics* // ACM SIGGRAPH Computer Graphics, vol. 17, no. 3, pp. 1–11, Jul. 1983

[29] Yan L.Q., Hašan M., Jakob W., Lawrence J., Marschner S., Ramamoorthi R.*Rendering glints on high-resolution normal-mapped specular surfaces* // ACM Trans. Graph., 2014, Vol. 33(4), p. 116.

[30] *Khronos registry* https://www.khronos.org/registry

[31] *Nvidia RTX* https://developer.nvidia.com/rtx

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Эффективный алгоритм визуализации предварительно рассчитанных лучей

Б.Х. Барладян, Е.Д. Бирюков, А.Г. Волобой, Л.З. Шапиро
Институт прикладной математики им. М.В.Келдыша РАН, Москва

Визуализация траекторий лучей света при моделировании его распространения в исследуемой сцене является эффективным средством анализа оптических систем, в частности, в задачах обнаружения паразитного света. В данной статье предложен алгоритм распараллеливания обработки карты лучей, предварительно рассчитанных методом стохастической трассировки, для многоядерных компьютеров. В большинстве практических случаев использование предложенного алгоритма позволяет обеспечить визуализацию лучей, выбранных согласно заданному критерию, в реальном времени.

Ключевые слова: *реалистичные изображения, глобальное освещение, стохастическая трассировка лучей, интерактивный анализ сцен, карты освещенности, трехмерные карты лучей.*

Effective algorithm for visualization of the pre-calculated rays

B.Kh. Barladian, E.D. Birukov, L.Z. Shapiro, A.G. Voloboy
Keldysh Institute of Applied Mathematics RAS, Moscow

Visualizing the trajectories of light rays in the lighting simulation in the investigated scene is an effective tool of analyzing optical systems, in particular, in problems of parasitic light detection. In this paper we propose an algorithm for parallelizing the processing of a ray map, previously calculated by the forward Monte Karlo ray tracing method, for multi-core computers. In most practical cases the use of the proposed algorithm makes it possible to visualize the rays selected according to a specified criterion in real time.

Keywords: *photorealistic images, global illumination, Monte Carlo ray tracing, interactive scene analysis, illumination maps, 3D ray maps.*

1. Введение

Использование физически корректных методов расчёта освещения, моделирующих распространение света и его взаимодействие с объектами сцены, находит все большее применение в различных областях науки и индустрии. Кроме классического использования этих методов для генерации реалистичных изображений [1] или для проектирования различных оптических приборов и освещающих устройств [2] они находят применение в новых областях. Например, перспективным является применение физически корректных синтезированных изображений и видео последовательностей в системах обучения искусственного интеллекта [3-5]. Корректность получаемых изображений является принципиально важной во всех этих применениях. В противном случае цель не будет достигнута, например, интеллектуальная система будет обучена неправильно. Однако эта физическая корректность может быть нарушена из-за ошибок в программе моделирования или ошибок пользователя при описании сцен, геометрии объектов, их оптических свойств. Одним из наиболее эффективных средств для обнаружения и анализа такого рода ошибок является визуализации распространения световых лучей [6-7].

2. Визуализация распространения света с использованием трехмерных карт лучей

В работе [8] было предложено сохранять в файл всю информацию, необходимую для дальнейшего анализа, при прямой стохастической трассировке лучей в файл. Эта информация содержит не только координаты начала и конца каждого сегмента, но и информацию об источниках света, оптических свойствах поверхностей и сред распространения света и объектах, с которыми происходило взаимодействие. Было предложено называть совокупность этих данных трехмерной картой лучей. В работе [8] были также предложены эффективные алгоритмы построения трехмерных карт лучей на многоядерных компьютерах. Эти трехмерные карты лучей используются для быстрого анализа световых характеристик на приемнике излучения с варьируемыми параметрами, для изучения особенностей распространения света в сцене и т.д. Одним из примеров использования этих карт является визуализация распространения световых лучей в задачах проектирования сложных оптических систем [6]. Для изучения деталей распространения света и получения различных статистических характеристик приходится использовать трехмерные карты очень большого размера, состоящие из десятков миллионов лучей и сотен миллионов сегментов вместе с большим объемом информации об оптических событиях, имевших место вдоль траектории луча. Размеры файлов для хранения этих карт могут достигать нескольких гигабайт.

При визуализации распространения световых лучей типичной задачей, критически важной с точки зрения времени обработки, является отбор и визуализация траекторий лучей, удовлетворяющих заданному критерию. Критерии могут быть самые разнообразные и достаточно сложные. Можно отметить следующие типичные события, происходящие вдоль траектории луча, которые могут представлять интерес при исследовании оптической системы [7], [9]:

- луч был выпущен заданным источником света;
- луч пересекся или не пересекся с заданной гранью объекта;
- луч пересекся с заданной частью геометрического объекта оптической системы и претерпел заданное оптическое преобразование;
- было зарегистрировано взаимодействие луча с заданным виртуальным измерительным прибором;
- луч пересекся с поверхностью, обладающей заданными оптическими свойствами, и претерпел заданное оптическое преобразование;

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

- луч претерпел заданное оптическое преобразование (зеркальное или диффузное отражение, поглощение и т.д.) на одном из объектов оптической системы;
- луч был зарегистрирован в заданной области виртуального измерительного прибора.

В общем случае из этих элементарных событий строится логическое выражение, представленное в виде дерева событий, изображенного на рис. 1, которое и является критерием отбора лучей для визуализации.

Выражение строится с помощью операций логического пересечения ($\cap$), объединения ($\cup$) и логического отрицания ($\neg$). Для каждого заданного оптического преобразования можно выбрать, какие именно преобразования имели место и сколько раз.

Для комфортной работы пользователя необходимо минимизировать время реакции системы при любых изменениях критерия отбора лучей для визуализации. Желательно обеспечить реакцию системы в реальном времени. В случае трехмерных карт лучей большого размера, полученных стохастической трассировкой, в случае спектральных вычислений и сложных критериев визуализации необходимо, чтобы время реакции системы не превышало нескольких минут. Поскольку практически все используемые в настоящее время компьютеры являются многоядерными, то представляется целесообразным распараллелить процесс анализа трехмерных карт лучей.

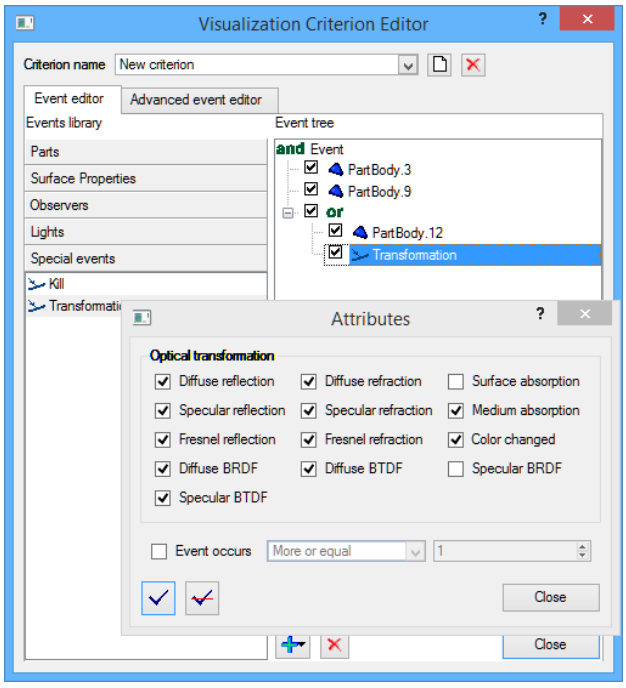


Рис. 1. Представление сложного критерия визуализации в виде дерева событий.

3. Многопотоковый алгоритм анализа трехмерных карт лучей.

Наш базовый алгоритм для моделирования распространения света в виртуальной сцене – это прямая трассировка лучей методом Монте-Карло. Как и во всех статистических методах, точность достигнутого решения зависит от количества экспериментов или, в нашем случае, от количества протрассированных лучей. Так как практические сцены, имеющие интерес для пользователя, обычно сложны, то для достижения высокой точности часто необходимо протрассировать десятки, а иногда и сотни, миллионов лучей. Размер файла, в котором хранится такая карта, может превышать несколько гигабайт. Как уже было показано в предыдущем разделе, пользователь может задавать достаточно сложные критерии отбора лучей для визуализации. Таким образом, получается, что чтение файла с сохраненными данными и выбор из них лучей, удовлетворяющих заданному критерию, является наиболее критичной процедурой для скорости визуализации распространения света. Поэтому именно алгоритм выбора лучей, удовлетворяющих критерию, должен быть реализован в многопотоковом варианте для обеспечения приемлемого времени реакции системы на действия пользователя.

В общем виде алгоритм выбора лучей, удовлетворяющих заданному критерию, представлен на рис. 2.

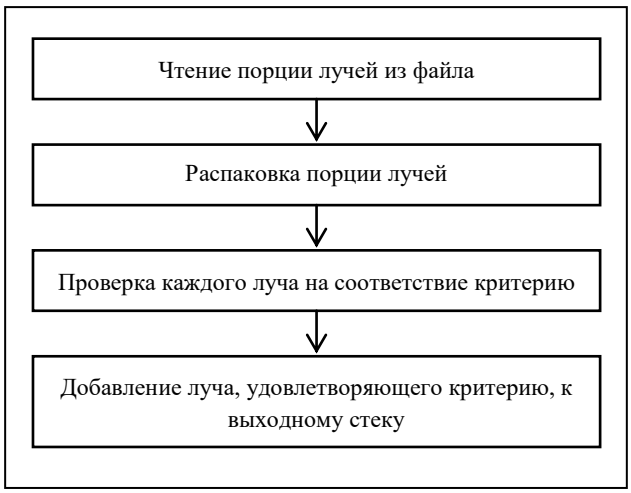


Рис. 2. Общий алгоритм выбора лучей по критерию.

На многоядерных компьютерах естественным подходом для ускорения обработки является параллельное использование нескольких потоков. Этот подход эффективно нами используется для прямой стохастической трассировки лучей [10]. Количество потоков выбирается, как правило, равным количеству виртуальных ядер компьютера. Ясно, что в процедуре, приведенной на рис. 2, только распаковка порции и проверка луча на соответствие критерию могут выполняться параллельно для разных порций лучей. Чтение порций лучей из файла должно выполняться последовательно, поскольку все потоки работают с одним файлом. Добавление найденного луча, удовлетворяющего критерию, в выходной стек также не может выполняться параллельно, поскольку стек является общим для всех потоков. Таким образом, мы приходим к многопотоковой схеме выбора лучей, удовлетворяющих заданному критерию, приведенной на рис. 3.

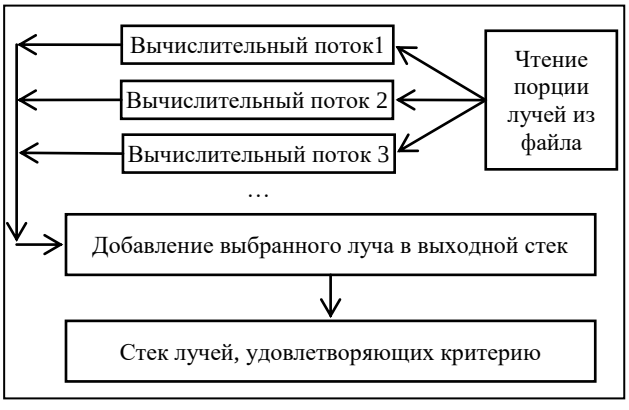


Рис. 3. Многопотоковой алгоритм выбора лучей по критерию

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Процедура анализа трехмерной карты лучей стартует после того, как пользователь указал необходимые параметры критерия выбора лучей посредством графического диалога, приведенного на рис. 4.

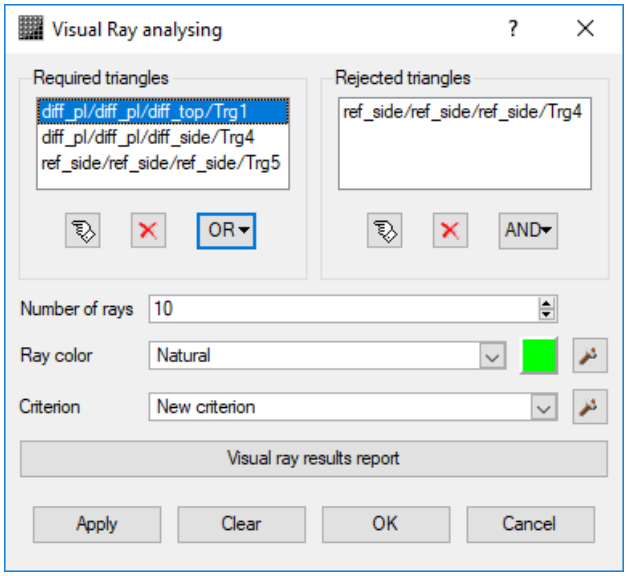


Рис. 4. Параметры процедуры анализа лучей.

В этом диалоге пользователь может установить треугольники, которые исследуемые лучи должны пересекать или не пересекать. Для выбранного списка треугольников можно установить условие **OR** или **AND,** то есть луч должен пересечь (или не пересечь) все треугольники из списка или хотя бы один. Дополнительное, более сложное условие критерия, может быть отредактировано в диалоге Visualization Criterion Editor (рис. 1). Пользователь может выбрать как естественный цвет отрезков лучей для визуализации так и искусственный, когда это более удобно с точки зрения восприятия информации.

Непосредственно процесс анализа и визуализации лучей начинается после нажатия кнопки **Apply**. Создаются все необходимые объекты для хранения результатов обработки данных, в частности, выходной стек лучей, открывается файл, в котором хранится карта лучей. А также создаются и запускаются вычислительные потоки для выбора лучей. Далее вся обработка данных осуществляется в вычислительных потоках, а главный поток только отражает прогресс этой обработки, указывая процент обработанных лучей.

Обработка лучей в каждом потоке осуществляется в соответствии со схемами алгоритма, приведенными на рисунках 2 и 3:

1. Считывается порция лучей из файла. Поскольку все потоки работают с одним файлом, то в этой функции для синхронизации потоков используется критическая секция. Поскольку чтение файла в современных компьютерах использует специальный сопроцессор ввода / вывода и кэш, то эта процедура практически не ограничивает работу потоков, которые распаковывают порции и проверяют критерии.
2. Полученная порция лучей распаковывается и осуществляется поиск лучей, удовлетворяющих заданному критерию.
3. Найденный луч, удовлетворяющий заданному критерию, добавляется в выходной стек в формате, пригодном для немедленной визуализации. Т.к. выходной стек является общим для всех вычислительных потоков, то это добавление также осуществляется с использованием критической секции. Поскольку эта процедура сводится, в основном, к добавлению в стек указателя на объект, содержащий массив сегментов, и достаточно редко требуется больше нескольких сотен таких лучей, то время ее выполнения мало и не задерживает работу основных вычислительных потоков.

4. Результаты

Эффективность построенной процедуры можно охарактеризовать следующими цифрами. Для файла, содержащего ~50 миллионов лучей (~450 миллионов сегментов), время полной обработки трехмерной карты лучей на компьютере Intel Core(TM) i7–4770 с четырьмя физическими ядрами и восемью потоками (с использованием Hyper Threading) составляет ~40 секунд. Файл в данном случае обрабатывался полностью, поскольку заказанное число лучей, удовлетворяющих заданному критерию, так и не было достигнуто. В файле содержались спектральные данные, т.к. само моделирование освещенности использовало 33 длины волны видимого спектра света, поэтому размер файла был ~6 Gb. В большинстве практических случаев достаточно обработать несколько миллионов лучей, тогда обработка и визуализация выбранных лучей осуществляется в реальном времени.

На рис. 5 приведена зависимость скорости обработки трехмерной карты лучей от количества используемых потоков для указанного выше компьютера. Видно, что предложенный алгоритм хорошо масштабируется при увеличении числа используемых процессоров. Даже использование виртуальных процессоров (добавление 5, 6, 7 и 8 потоков) дает заметное увеличение скорости обработки.

Рис. 5. Зависимость скорости обработки трехмерной карты лучей от количества используемых потоков.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

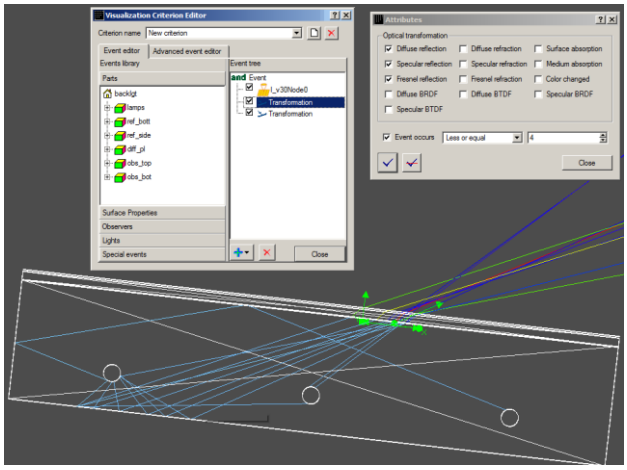


Рис. 6. Пример визуализации лучей.

На рис. 6 приведен пример визуализации выбранных лучей в соответствии с заданным критерием. Лучи меняют цвет, поскольку материалы сцены обладают дисперсией оптических характеристик, а стохастическая трассировка использует «русскую рулетку» для выбора длины волны после соответствующего события.

5. Заключение

Визуализация процесса моделирования освещенности посредством рисования траекторий лучей света в исследуемой сцене является востребованным средством анализа оптических систем, в частности, в задачах обнаружения паразитного света. Само моделирование требует больших вычислительных ресурсов, поэтому анализ в процессе моделирования не может выполняться в интерактивном режиме. Однако механизм предварительно рассчитанных трехмерных карт лучей позволяет проводить анализ сцены на этапе постобработки. И здесь конструкторы оптических систем хотели бы получать максимально быструю реакцию программы на заданные условия.

Предложенный алгоритм распараллеливания процесса постобработки лучей позволил обеспечить в большинстве практических случаев выборку лучей из трехмерных карт и их визуализацию в реальном времени. Только для очень больших карт размером в несколько гигабайт, как правило, спектральных, и при использовании сложных критериев, требующих обработку всего файла, время визуализации может достигать нескольких минут.

6. Литература

[1] Pharr M., Humphreys G.. Physically Based Rendering. Second Edition: From Theory to Implementation. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 2010.

[2] Bogdanov, N., Zhdanov, D., Potemin, I., & Zhdanov, A. (2017). Design of Ergonomic Illumination Systems for Cultural, Medical, Educational Facilities. The Educational Review, USA, 1(4), 85-90. http://dx.doi.org/10.26855/er.2017.04.001

[3] Tsirikoglolu A., Kronander J., Wrenninge M., Unger J. “Procedural modeling and physically based rendering for synthetic data generation in automotive applications,” arXiv preprint arXiv:1710.06270, 2017

[4] McCormac J., Handa A., Leutenegger S., Davison A. SceneNet RGB-D: 5M Photorealistic Images of Synthetic Indoor Trajectories with Ground Truth // The IEEE International Conference on Computer Vision (ICCV 2017).

[5] Rozantsev A, Lepetit V., Fua P. On rendering synthetic images for training an object detector // Comput. Vis. Image Underst. 137, C (August 2015), 24-37. DOI: http://dx.doi.org/10.1016/j.cviu.2014.12.006.

[6] Wernert E.. A unified environment for presenting, developing and analyzing graphics algorithms // Computer Graphics, vol. 31, № 3, 1997, pp. 26-28.

[7] Волобой А.Г., Галактионов В.А., Жданов А.Д., Жданов Д.Д.. Средства визуализации распространения световых лучей в задачах проектирования оптических систем // "Информационные технологии и вычислительные системы, № 4, 2009, с. 28-39.

[8] Barladyan B., Shapiro L., Voloboy A. Ray maps technique for effective interrogation of results of MCRT simulation // *Conference proceedings of 21-th International Conference on Computer Graphics and Vision GraphiCon-2011, Moscow State University, September 26-30, 2011, Moscow, Russia, p.46-49.*

[9] Kopylov E., Dmitriev K.. Light propagation visualization as a tool for 3D scene analysis in lighting design // Computers & Graphics, vol. 24, № 1, 2000, pp. 31-39.

[10] Барладян Б.Х., Шапиро Л.З., Денисов Е.Ю.,. Волобой. А.Г Эффективный многопотоковый алгоритм расчета глобальной освещенности // Программирование, № 4, 2017, с. 3-12.

Об авторах

Борис Хаимович Барладян, к.т.н., с.н.с., ИПМ им. М.В. Келдыша РАН.
E-mail: obb@gin.keldysh.ru

Елисей Дмитриевич Бирюков, м.н.с., ИПМ им. М.В. Келдыша РАН.
E-mail: peb@gin.keldysh.ru

Алексей Геннадьевич Волобой, д.ф.-м.н., в.н.с., ИПМ им. М.В. Келдыша РАН.
E-mail: voloboy@gin.keldysh.ru

Лев Залманович Шапиро, к.т.н., с.н.с., ИПМ им. М.В. Келдыша РАН.
E-mail: pls@gin.keldysh.ru

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Визуализация объёмных объектов с помощью структурированного оптического излучения

А.М. Вин
humburger49@gmail.com
Национальный исследовательский университет «МЭИ», Москва, Россия

В статье рассматриваются оптические методы визуализации объёмных объектов с помощью структурированного оптического излучения (СОИ), получаемого от источников когерентного излучения с помощью различных дифракционных оптических элементов (ДОЭ). Изображение фотографируется цифровым фотоаппаратом. Рассмотрен оптический метод визуализации плоских и объёмных объектов, описана экспериментальная установка и проведение измерений. Получены экспериментальные результаты, показывающие возможность определения параметров различных объемных объектов.

Ключевые слова: *структурированное оптическое излучение, дифракционный оптический элемент, исследуемые объемных объекты, фотографирование.*

Visualization of volumetric objects using structured optical radiation

A.M. Win
humburger49@gmail.com
National Research University “MPEI”, Moscow, Russia

The article deals with optical methods of visualization of three-dimensional objects using structured optical radiation (SOR) obtained from sources of coherent radiation by means of various diffraction optical elements (DOE). The image is photographed by a digital camera. The optical method of visualization of plane and volume objects is considered, the experimental setup and measurements are described. The experimental results show the possibility of determining the parameters of various volume objects.

Keywords: *structured optical radiation, diffraction optical element, investigated three-dimensional objects, photography.*

1. Введение

Особый интерес на практике, в промышленности, в научных исследованиях представляет определение формы поверхности, а также перемещения объектов бесконтактным способом. В данном методе решается обратная задача: по полученному изображению восстанавливается форма или перемещение поверхности.

На качественно новом научно–техническом уровне разработаны оптические методы диагностики неоднородных сред, которые используются для получения информации об оптической среде или потоке жидкости в выбранном сечении (2D-диагностика) и в объеме (3D-диагностика).

Принцип измерения основан на использовании структурированного оптического излучения (СОИ), получаемого с помощью оптических дифракционных элементов, зондирующего исследуемый экран и регистрации части пучка, прошедшего мимо экрана [1].

Оптическое излучение, пройдя через дифракционный оптический элемент (ДОЭ), даёт совокупность элементарных микропучков, которая может состоять из набора вертикальных и горизонтальных прямых. Назовём плоскость, перпендикулярную выходящей из ДОЭ совокупности элементарных пучков, исходной плоскостью.

Если поместить цифровую камеру под некоторым углом к исходной плоскости, то на изображении, получаемым фотокамерой, расстояние между соседними элементами дифракционной картины (например, вертикальными или горизонтальными прямыми) будет различным в зависимости от формы поверхности исследуемого объекта и угла между прямой, перпендикулярной плоскости камеры, и исходной плоскостью. Зная этот угол, и проанализировав полученные изображения, можно восстановить форму поверхности исследуемого объекта. По фотографии вычисляются координаты точек пересечения вертикальных и горизонтальных линий. Эти данные в дальнейшем используются для визуализации поверхности изучаемого объекта. При этом для съёмки совокупности элементарных микропучков, расположенных на поверхности предмета, требуются две камеры, расположенные под некоторым углом друг относительно друга (желательно, чтобы этот угол был равен 90˚).

В данном методе применяется СОИ, образующееся с помощью специальных оптических элементов непосредственно на выходе источника. В таком способе сохраняется высокая когерентность и малая расходимость пучков, поэтому для описания модели структурированного оптического излучения можно использовать представления геометрической оптики. Эта модель может быть представлена группой пучков, образующих поверхности в виде дискретного набора плоскостей [2].

Классические оптические элементы позволяют получать СОИ ограниченного вида: узкий пучок, рассходящийся пучок, плоскость и крест. Ситуация улучшилась с появлением компактных ДОЭ, изготавливаемых методами дифракционной оптики.

2. Оптическая визуализация трёхмерных объектов

При проведении съёмки с помощью простого фотоаппарата получается плоское изображение, при этом теряется объёмность. Для фотографирования объёмных предметов используется стереосъёмка. Стереосъёмка — это получение стереопары изображений фотографическим способом. Стереопара представляет собой сочетание двух изображений одного и того же объекта с различных точек зрения. Одно изображение предназначается для левого, а другое — для правого глаза. При рассматривании стереопары, например, в стереоскоп изображение объекта воспринимается с такой же локальной глубиной, что и сам объект.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Основой стереофотографии является изготовление двух подобных, но не идентичных фотоизображений, полученных с двух точек съёмки относительно объекта. Действительное расстояние между точками съёмки называется базовым расстоянием или базисом фотосъёмки; величина его определяется тем, как должен выглядеть объект на стереоизображении. В идеальном случае оси фотокамеры и объективов в этих двух положениях (точках) должны быть параллельны. Избыточная сходимость осей объективов приводит к заметному изменению масштаба вдоль горизонтали изображения; на парных снимках это изменение происходит в противоположных направлениях.

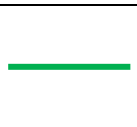
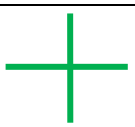
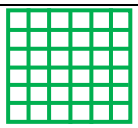
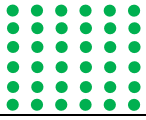
Одним из существующих методов визуализации трёхмерных объектов является измерение формы с использованием проецирования интерференционных полос [3]. Нелокализованная интерференционная картина, образующаяся при детектировании двух пучков света (например, картина, получаемая в интерферометре Майкельсона), может быть использована для формирования интерференционных полос на поверхности предмета. Полоса появляется каждый раз, когда поверхность предмета пересекает интерференционную полосу в световом пучке. Форма этих интерференционных полос зависит от формы поверхности и от направления наблюдения.

Для визуализации геометрических параметров диффузно рассеивающих плоских объектов в работах [4, 5] было предложено использовать СОИ, получаемое с помощью ДОЭ. Для данной задачи СОИ существенно проще, чем использование двухлучевого интерферометра, потому что ДОЭ позволяет получать СОИ с более узкими полосами, чем двухлучевые интерферометры, что повышает точность измерений.

3. Основные типы СОИ

СОИ представляет собой пространственно модулированное по амплитуде излучение, получаемое с помощью классических оптических элементов, ДОЭ или структурированных экранов.

Таблица 1. Основные типы СОИ, получаемые с помощью ДОЭ

Линия	Крест	Горизонтальные линии
		
Вертикальные линии	Сетка	Круговая линия
		
Точка	Точечная линия	6 на 6 матрица
		

Дифракционные оптические элементы (ДОЭ), имеют форму тонкой фазовой пластины со специальным фазовым рельефом, выгравированным в ней. Дифракция оптического излучения таким оптическим элементом производит различные виды пространственно модулированного излучения, известного как структурированное оптическое излучение. используется как с газовыми, так и полупроводниковыми лазерами, генерирующими высоко астигматические лучи. Основные элементы СОИ приведены в таблице 1.

4. Схема экспериментальной установки и результаты измерений

Схема созданной установки для визуализации исследуемых объёмных объектов показана на рис. 1. Излучение лазерного модуля 1 с длиной волны излучения 0,552 мкм, мощностью 20 мВт распространяется вдоль оси оптической системы, проходит через ДОЭ 2, которые формируют различные дифракционные картины, фокусируется объективом 3 на исследуемую поверхность 4. Рассеянное излучение регистрируются с помощью цифрового фотоаппарата 5.

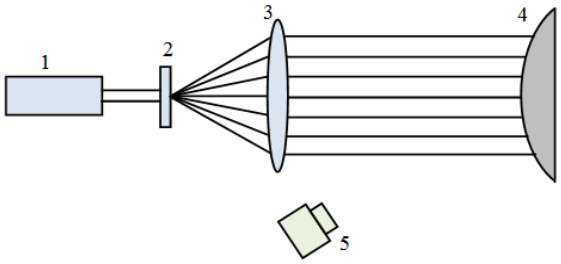
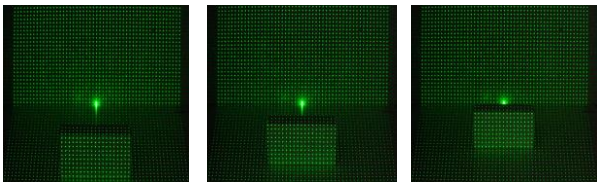


Рис. 1. Принципиальная схема оптического комплекса: 1– лазерный модуль, 2 – дифракционный оптический элемент (ДОЭ), 3 – объектив, 4 – исследуемая поверхность, 5 – цифровой фотоаппарат.

В начале были получены цифровые фотографии дифракционных картин, формирующихся после прохождения оптического излучения через ДОЭ, при их наблюдении перпендикулярно поверхности исследуемого объекта (рис. 1). Использовались 9 различных ДОЭ.

На рис. 2 приведены результаты визуализации прямоугольника с помощью измерительного комплекса при различных расстояниях от объектива до измеряемого объекта. Результаты измерений исследуемого прямоугольника с помощью штангенциркуля: высота 34 мм, длина 84 мм и площадь основания прямоугольника 2856 мм^2. Здесь l – расстояние от края СОИ до передней поверхности исследуемого объекта.



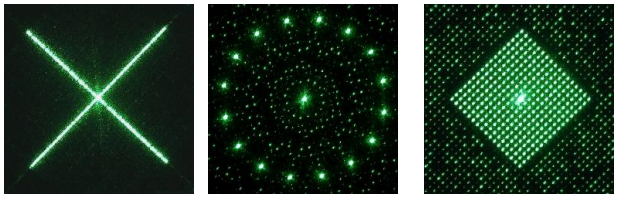
l_1= 1 мм l_2 = 12 мм l_3 = 27 мм

Рис. 2. Визуализация прямоугольника при различных расстояниях от края СОИ до его передней поверхности.

5. Виды СОИ

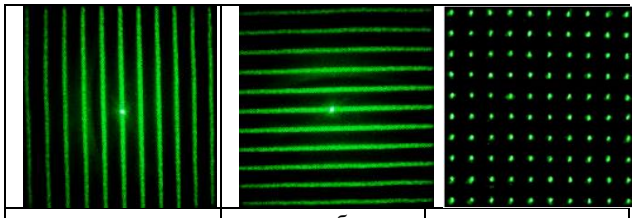
Представлены изображения СОИ на полупрозрачном плоском экране, установленном перпендикулярно направлению распространения рис. 3 и рис. 4. По полученным фотографиям можно найти угол поворота экрана. Расстояние (в пикселях) между крайними горизонтальными точками на картине СОИ. Фотографии были получены при одинаковых условиях эксперимента, а именно – одинаковое расстояние до плоскости падения СОИ, и при одних и тех же параметрах фотокамеры.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18



а б в

Рис. 3. Изображения СОИ на экране: а – крестовидный пучок, б – круговая точечная сетка, в – прямоугольная точечная сетка.



а б в

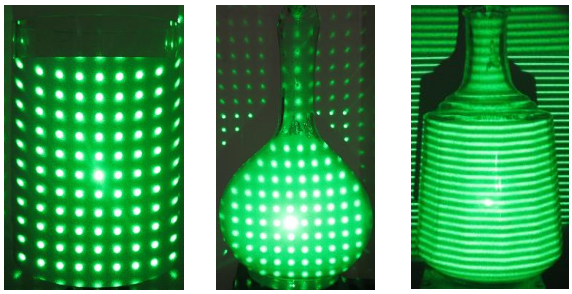
Рис. 4. Изображения СОИ на экране: а – вертикальные линии, б – горизонтальные линии, в – точечная сетка.

6. Визуализация объемных объектов

На рис.5 приведены изображения трёхмерных объектов при естественном освещении. На рис. 6 приведены изображения трёхмерных объектов с помощью СОИ.

а б в

Рис. 5. Изображение цилиндрических и сферических поверхностей при естественном освещении: а – цилиндр, б – сферический кувшин, в – цилиндрический кувшин.



а б в

Рис. 6. Визуализация цилиндрических и сферических поверхностей с помощью СОИ: а – цилиндр, б – сферический кувшин, в – цилиндрический кувшин.

На рис.7 приведены изображения трёхмерных объектов при естественном освещении и с помощью СОИ.

а б в

Рис. 7. Визуализация сферических поверхностей: а – освещение некогерентным излучением, б – освещение линейным оптическим излучением, в – освещение точечным оптическим излучением.

7. Заключение

Разработана экспериментальная установка, содержащая полупроводниковый лазер с мощностью излучения 20 мВт, 6 дифракционных элементов для получения пучков различной формы, цифровую фотокамеру. Освоена методика проведения экспериментальных исследований. Проведены эксперименты по измерению прямоугольной поверхности как при помощи интерферирующих пучков, так и при помощи СОИ, получаемого с помощью ДОЭ. Получено удовлетворительное совпадение расчетных значений угла поворота плоской поверхности с параметрами, задаваемыми при помощи поворотного столика.

Сделан краткий обзор методов визуализации трёхмерных объектов: стереофотография, интерференционный метод. Проанализирован способ визуализации трёхмерных объектов с помощью СОИ. В результате эксперимента получены изображения дифракционных картин на поверхности различных объектов (плоскости, при различных углах наблюдения, и поверхности сложного объекта).

8. Литература

[1] Rinkevichyus B.S., Evtikhieva O.A., Raskovskaya I.L. Laser Refractography. New York, Springer, 2011. – 189 p.

[2] Сойфер В.А., Безус Е.А., Быков Д.А., Досколович Л.Л., Ковалев А.А. и др. Дифракционная оптика и нанофотоника. М.: Физматлит, 2014. – 608 с.

[3] Джоунс Р., Уайкс К. Голографическая и спекл-интерферометрия. Пер. с англ. – М.: Мир, 1986. – 328 с.

[4] Вин А.М., Павлов И.Н., Ринкевичюс Б.С. Визуализация геометрических параметров диффузно рассеивающих плоских тел с помощью структурированного оптического излучения: Научная визуализация, 2017, Том 9, № 2. С. 43 – 48.

[5] Вин А.М., Павлов И.Н., Ринкевичюс Б.С. Оптико-электронный комплекс для визуализации формы шероховатой поверхности с помощью структурированного оптического излучения: Труды XIV Международной научно-технической конференции. [Электронный ресурс] – М. Издательство Перо, 2017. С. 389 – 393.

Об авторе

Аунг Мьо Вин, аспирант кафедры физики им. В.А. Фабриканта Национального исследовательского университета «МЭИ». E-mail humburger49@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Использование методов виртуального прототипирования при разработке устройств виртуальной реальности, не подверженных проблеме конфликта вергенции и аккомодации зрения

А.Д. Жданов[1], Д.Д. Жданов[1], Н.Н. Богданов[1], И.С. Потемин[1], В.А. Галактионов[2], И.Л. Лившиц[1]
andrew.gtx@gmail.com|ddzhdanov@mail.ru|nnbogdanov@corp.ifmo.ru|ipotemin@yandex.ru|vlgal@ginkeldysh.ru|ecenter-optica10@yandex.ru
[1]Университет ИТМО, Санкт-Петербург, Россия
[2] ИПМ им. М.В.Келдыша РАН, Москва, Россия.

Одной из основных проблем при проектировании систем виртуальной (VR), дополненной (AR) и смешанной (MR) реальностей является оценка зрительного восприятия результата, в частности дискомфорт, вызванный несоответствием вергенции и аккомодации зрения пользователя данного устройства. Для оценки дискомфорта зрительного восприятия стандартные оптические характеристики, такие как аберрации, функция контраста и другие, не применимы. С другой стороны, дискомфорт может быть оценен на реальном прототипе, однако его изготовление может оказаться трудоемким. Авторы данной работы разработали методы виртуального прототипирования визуальных оптических систем, которые могут быть использованы при проектировании AR, VR и MR устройств для анализа дискомфорта, в том числе вызванного рассогласованием вергенции и аккомодации зрения.

Ключевые слова: *виртуальное прототипирование, конфликт вергенции и аккомодации, виртуальная реальность, дополненная реальность, смешанная реальность, отслеживание зрачка, оптическое проектирование, обработка изображений.*

A virtual prototyping approach in designing the vergence-accommodation conflict-free VR devices

A.D. Zhdanov[1], D.D. Zhdanov[1], N.N. Bogdanov[1], I.S. Potemin[1], V.A. Galaktionov[2], I.L. Livshic[1]
andrew.gtx@gmail.com|ddzhdanov@mail.ru|nnbogdanov@corp.ifmo.ru|ipotemin@yandex.ru|vlgal@ginkeldysh.ru|ecenter-optica10@yandex.ru
[1]ITMO University, St. Petersburg, Russia
[2] Keldysh Institute of Applied Mathematics RAS, Moscow, Russia.

One of the main problems in the design of AR, VR and MR devices is an estimation of the visual perception of the design result by its end user, and in particular the discomfort caused by disagreement between the vergence and accommodation of the user eyes. Base optical characteristics like aberrations, contrast functions, etc. are not applicable for estimation of this discomfort. On the other hand, the discomfort can be estimated with the real prototype of such device, however creating the real prototype might be costly. Authors propose the virtual prototyping method for visual optical systems, that can be used as an element of design of AR, VR and MR devices to estimate the visual perception discomfort including the vergence-accommodation conflict at each design step.

Keywords: *virtual prototyping, vergence-accommodation conflict, virtual reality, augmented reality, mixed reality, eye tracking, optical design, image processing.*

1. Введение

Устройства виртуальной, дополненной и смешанной реальностей нашли широкое применение в различных областях человеческой жизни. Например, они используются в виде нашлемных индикаторов и очков виртуальной, дополненной и смешанной реальностей, в виде индикаторов на лобовом стекле в автомобильной и авиа-индустрии, в системах тренажеров. При создании тренажеров они используются для создания интуитивно-понятных интерфейсов взаимодействия виртуального и реального миров.

Высококачественная оптика, реализм изображения виртуального мира и высокое разрешение ЖК экранов являются только необходимым условием для представления естественного восприятия виртуального мира. Достаточным условием естественного восприятия виртуального изображения является согласованность вергенции и аккомодации зрения наблюдателя. Данная проблема особенно остро встает в стереоскопических системах видео 360 [4, 7]. При съемке видео 360 каждая часть изображения снимается парой со-направленных камер, расположенных друг относительно друга на среднем межцентровом расстоянии глаз человека. Данная схема соответствует наблюдению бесконечно удаленного объекта. При просмотре такого видео, пользователь может комфортно наблюдать объекты, расположенные на удаленном расстоянии, однако при попытке рассмотреть объекты вблизи, изображение начинает двоиться, так как ожидаемой дистанцией наблюдения является бесконечность. Данная проблема не может быть решена в рамках видео 360 и единственным решением на настоящий момент является рекомендация избегать съемки близких объектов.

В системах виртуальной реальности данная проблема может быть частично решена за счет отслеживания положения зрачка пользователя. Когда направление взгляда изменяется, необходимо создать новые изображения для левого и правого глаза, соответствующие новому направлению взгляда. Однако правильная ориентация виртуальных камер, соответствующих левому и правому глазам, не является достаточным условием для естественного восприятия виртуального мира. Необходимо также соответствие аккомодации вергенции зрения [2, 3, 6], поскольку при конвергенции человеческих глаз на объекте происходит фокусировка (аккомодация) на этом объекте. В результате наблюдаемый объект виден четко, а объекты, которые находятся ближе или дальше, размыто. Эффект

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

размытия зависит от расстояния от плоскости фокусировки зрения до объекта и от размера зрачка.

Чаще всего для синтеза изображения в системах виртуальной реальности используется идеальная камера с бесконечно малым зрачком, что дает полностью четкую картинку. Если система виртуальной реальности позволяет отследить положения зрачка, то она может синтезировать четкое изображение для объектов в плоскости фокусировки и размытое для удаленных от нее.

С другой стороны, для того, чтобы пользователь мог корректно сфокусировать зрение на объектах виртуальной сцены, необходимо произвести диоптрическую коррекцию проекционной оптики системы виртуальной реальности, то есть согласовать ее с реальной аккомодацией глаза. В таком случае объекты, на которые он смотрит, будут видны четко. Однако и остальные объекты будут видны четко, что приведет к дискомфорту. Рисунок 1 иллюстрирует данный эффект.

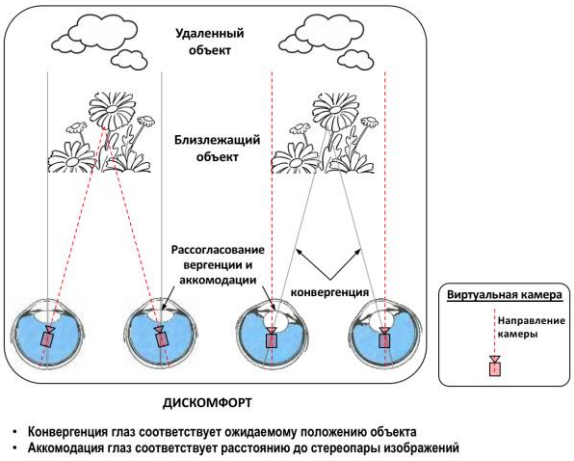


Рис. 1. Эффект конфликта вергенции и аккомодации зрения.

С другой стороны, отсутствие диоптрической коррекции приведет к размытию всего изображения целиком, что также вызовет дискомфорт. Таким образом, задача проектирования системы виртуальной, дополненной или смешанной реальности требует одновременно синтеза естественно размытого изображения (в соответствии с аккомодацией глаз) с диоптрической коррекцией проекционной оптической системы. На рисунке 2 показано условие формирования естественного восприятия виртуального мира без конфликта вергенции и аккомодации зрения.

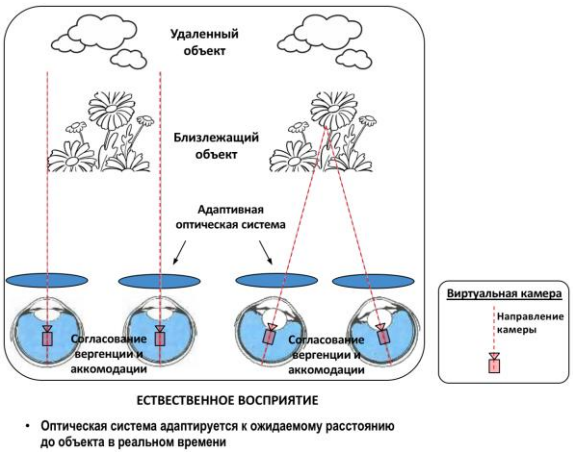


Рис. 2. Учет вергенции и аккомодации зрения при синтезе изображений для левого и правого глаз.

В настоящее время существует несколько методов для решения проблемы соответствия аккомодации и вергенции [3, 5]. В основном они базируются или на синтезе правильно-размытого изображения с изменяемым фокусным расстоянием оптической системы, или на преобразовании плоского изображения ЖК экрана в псевдотрехмерное, соответствующее глубинам сцены.

В случае использования адаптивной панкратической системы, система синтеза изображения должна моделировать эффект дефокусировки, вызванный естественной аккомодацией зрения. Данный эффект является аберрацией первого порядка и может легко моделироваться современными системами компьютерной графики. Однако для естественного восприятия дефокусированного изображения панкратическая система должна фокусироваться на плоскости в виртуальной сцене или, другими словами, адаптироваться к данным, полученным от системы отслеживания положения зрачков пользователя. Таким образом изменение направления взгляда пользователя будет приводить к синтезу правильно дефокусированного изображения совместно с адаптацией фокусного расстояния, что в сумме приведет к естественному восприятию изображения. К сожалению, задача проектирования адаптивной панкратической системы технически сложна и на настоящий момент еще не решена.

2. Построение изображения

Изображение, отображаемое на ЖК экранах системы виртуальной реальности, должно быть согласовано, во-первых, с оптической системой проецирования изображения на зрачок глаза и, во-вторых, с условиями наблюдения, то есть вергенцией и аккомодацией зрения, то есть того, на что смотрит пользователь. Естественно, что оптическая система проецирования изображения не может быть безаберрационной по причине наложенных на нее жестких ограничений, таких как широкий угол зрения и малый физический размер. К счастью, размер зрачка глаза существенно меньше, чем область формирования изображения оптической системы, и, следовательно, аберрации будут ограничены главным образом дисторсией изображения. Современные системы синтеза изображений могут учитывать дисторсию, и, соответственно, строить изображение с обратной дисторсией, делая данный эффект невидимым пользователю. Остается проблема несоответствия между вергенцией и аккомодацией зрения. Современные системы виртуальной реальности отслеживают положение зрачков и могут определить направление взгляда пользователя. На основе этой информации можно вычислить точку сцены, на которую направлен взгляд. В случае систем виртуальной реальности можно также точно определить объект, на который именно направлен взгляд.

С помощью современных систем синтеза изображений можно построить изображения для левого и правого глаз, используя направления взгляда при задании камер. Однако остается проблема аккомодации. Когда глаз фокусируется на каком-либо объекте, то другие объекты, находящиеся вне плоскости фокусировки, размываются. Данный эффект корректной фокусировки (и дефокусировки) может быть достигнут двумя методами.

Первый – это преобразование плоского изображения, отображенного на ЖК экране в псевдотрехмерное: в зависимости от расстояний до объектов, соответствующих точкам экрана, пространственный модулятор света может задать фазовое запаздывание отделенных пикселей ЖК матрицы, тем самым преобразуя 2-мерное изображение в 2.5-мерное [5]. От системы синтеза изображений данный

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

подход требует только вычисления карты глубин, что не является технической проблемой.

Другой метод – это синтез 2-мерного изображения с дефокусировкой, соответствующей естественному восприятию глаза. В данном случае дефокусировка будет основной среди всех возможных аберраций глаза. В случае дефокусировки, функцией рассеивания точки будет цилиндр, чей радиус зависит от фокусного расстояния оптической системы, размера зрачка и смещения от плоскости фокусировки. Учитывая, что фокусное расстояние и размер зрачка постоянны для всей сцены, в результате радиус зависит только от смещения от объекта до плоскости фокусировки. Таким образом, синтезировав картинку «идеальной» точечной камеры с картой глубин, достаточно отфильтровать ее данной цилиндрической функцией.

Использование этого метода должно базироваться на физической корректности вычислений. Во-первых, требуется сформировать карту расстояний до объектов, а не до первых видимых объектов. Например, если какой-либо объект виден отраженным в зеркале, дефокусировка должна быть применена к этому объекту, а не к зеркалу. Во-вторых, может возникнуть проблема неестественного размытия общей границы объектов, находящихся на разном расстоянии от наблюдателя: поскольку при синтезе изображений используется «идеальная» камера с бесконечно малым размером зрачка, то размытие границы может выглядеть неестественным образом.

Целью исследования был анализ и оценка возможного несоответствия между аккомодацией и вергенцией зрения в системах виртуальной, дополненной и смешанной реальностей. По этой причине для синтеза изображений были использованы методы синтеза физически корректных изображений, отличные от тех, которые используются в системах виртуальной реальности. Эти методы основаны на физически точном расчете переноса светового излучения, использованные в программном продукте Lumicept [8]. Они позволяют синтезировать как реалистичные изображения сцены, видимые левым и правым глазами, так и то, что пользователь увидит при использовании различных устройств виртуальной, дополненной или смешанной реальности [1]. Сравнивая полученные изображения, можно оценить возможный дискомфорт визуального восприятия.

3. Модель конфликта вергенции и аккомодации зрения

Целью данной работы была разработка быстрого и точного метода прототипирования различных подходов, нацеленных на избавление от эффекта конфликта вергенции и аккомодации зрения. На настоящий момент существует три основных подхода к отображению объектов виртуального мира, это виртуальная (VR), дополненная (AR) и смешанная (MR) реальности. Каждый из этих подходов имеет свои особенности, которые должны быть учтены при решении задач снижения дискомфорта восприятия пользователем виртуального мира. В случае виртуальной реальности, весь видимый пользователем мир виртуален, и он видит только объекты, созданные системой отображения. Следовательно, если система снабжена устройством отслеживания положения глаз, то можно синтезировать изображение, корректно учитывающее все эффекты, в том числе и то, как решается проблема конфликта вергенции и аккомодации зрения оптической системой. В случае смешанной реальности проблема становится сложнее, так как в ней существует два мира: реальный и виртуальный. Система синтеза изображений имеет полную информацию о виртуальном мире, но при этом только частичную о реальном, что, в свою очередь, не позволяет вычислить точку конвергенции в реальном мире по причине возможного отсутствия информации о ней. Это может привести к дополнительному дискомфорту восприятия. Системы дополненной реальности являются частным случаем систем смешанной реальности. Более того, в большинстве случаев они не являются стереоскопическими и, следовательно, не подвержены эффекту конфликта вергенции и аккомодации зрения.

Учитывая проблемы и сложность прототипирования систем виртуальной, дополненной и смешанной реальностей, авторы предлагают специально разработанную для данной цели систему виртуального прототипирования. Система состоит из шести основных компонент, а именно:

- виртуальный прототип глаза (как изображение формируется на сетчатке глаза);
- виртуальный прототип оптической системы, проецирующей изображение с ЖК экрана на зрачок глаза;
- виртуальный прототип реального мира (актуально только для систем дополненной и смешанной реальности), соответствующий изображению наблюдаемого реального мира, формируемому на сетчатке глаза;
- виртуальный прототип оптической системы между пользователем и реальным миром (актуально только для систем дополненной и смешанной реальности, например, очки);
- виртуальный прототип виртуального мира, соответствующий изображению виртуального мира, формируемого на сетчатке пользователя, если бы он наблюдал его как реальный;
- виртуальный прототип системы синтеза изображений, отображаемых на ЖК экране.

В результате решаемая задача сводится к:

- созданию виртуального устройства виртуальной, дополненной или смешанной реальности, состоящего из перечисленных выше компонент;
- созданию реального (ожидаемого) изображения сцены (реального или виртуального мира), формируемого на сетчатке глаза (данное изображение не будет иметь конфликта вергенции и аккомодации зрения);
- построению изображения, отображаемого на ЖК экранах системы виртуальной, дополненной или смешанной реальности при тех же условиях наблюдения и проецированию их через оптическую систему на сетчатку глаза;
- сравнению полученных изображений (первое будет свободно от конфликта вергенции и аккомодации зрения, тогда как второе будет тем, что пользователь реально увидит через проектируемую систему).

В данной статье приведены результаты виртуального прототипирования системы виртуальной реальности как подверженной конфликту вергенции и аккомодации зрения, так и избавленной от нее. Был проанализирован источник возникновения конфликта и предложен метод его исключения за счет использования проекционной системы с переменным фокусным расстоянием. Таким образом, изображения для левого и правого глаз строятся в соответствии с конвергенцией глаз пользователя и фокусное расстояние оптической системы адаптируется к той же конвергенции, полностью исключая конфликт.

4. Результаты моделирования

Для анализа возможного дискомфорта использования системы ВР по причине несоответствия вергенции и аккомодации зрения была создана простая сцена виртуального мира, представленная на рисунке 3. Сцена построена таким образом, что пользователь видит серию находящихся на полу объектов, расположенных на различном расстоянии от него и попадающих в угол зрения

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

в 10°. Расстояние до ближайшего объекта 0.5 метра, а до дальнего 32 метра. В результате возможное несоответствие между вергенцией и аккомодацией будет велико, что может вызвать существенный дискомфорт у наблюдателя.

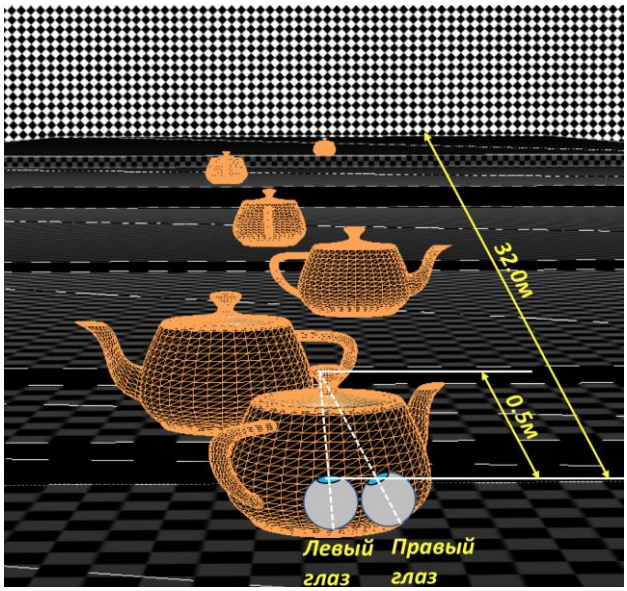


Рис. 3. Сцена виртуального мира.

Было использовано две модели наблюдения сцены. Первая – это модель зрения человека. Изображение, видимое каждым глазом по отделенности, синтезировалось с учетом вергенции и аккомодации глаза. Для построения изображения, формируемого на сетчатке глаза, была использована система фотореалистичной визуализации [7, 8].

Во второй модели строилось изображение, видимое каждым глазом по отдельности, но через систему виртуальной реальности, отображающую ту же самую сцену. На рисунке 4 представлены две модели устройств виртуальной реальности (схема для одного глаза), рассмотренные в рамках данного исследования:

- Первая модель представляет собой идеальную линзу, строящую безаберрационное изображение. Данная модель использовалась как эталон при проектировании более сложных систем. В описываемой модели ЖК матрица и линза смещаются в зависимости от точки наблюдения, тем самым изменяя расстояние фокусировки.
- Вторая модель представляет собой специально спроектированный окуляр системы виртуальной реальности. Он предоставляет высокое качество изображения в угле зрения в 10° и может также считаться безаберрационным.
-

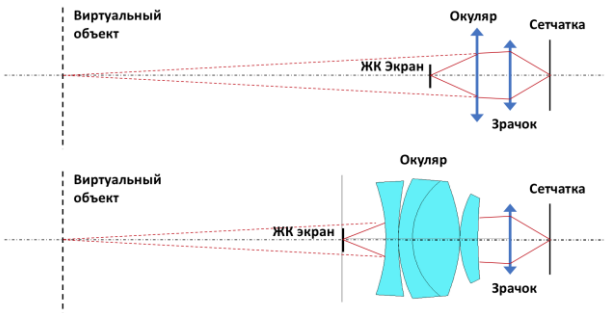


Рис. 4. Модели систем виртуальной реальности.

На рисунке 5 представлены результаты синтеза видимого изображения для его левого и правого глаза при условии естественного наблюдения пользователем данной сцены. В верхней части показано изображение, смоделированное для поля зрения 40°, а в нижней – центральная область поля зрения 10°. Глаза при этом аккомодировались и конвергировались на расстояние в 565мм. Хорошо видно, что в точке фокусировки картинка четкая, тогда как вне фокуса она размыта.

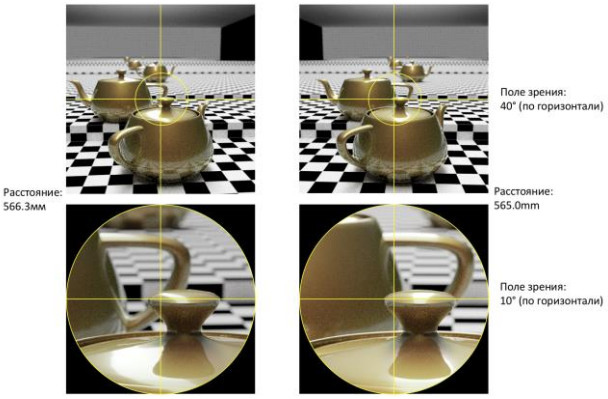


Рис. 5. Реальное наблюдение виртуальной сцены.

На рисунке 6 представлены результаты синтеза изображения сцены, наблюдаемого через систему виртуальной реальности, отслеживающей положение зрачка и, соответственно, строящей изображения для левого глаза с учетом конвергенции зрения. На данных изображениях хорошо прослеживается возникновение эффекта рассогласования вергенции и аккомодации глаз. В обоих представленных на рисунке случаях система рендеринга создает изображение с помощью модели камеры-обскуры. Также в обоих случаях положение ЖК-дисплея, размещенного в фокусе окуляра, не изменяется. В первом случае глаза наблюдателя аккомодировались к бесконечности. В результате все объекты в фокусе, что очевидно приводит к рассогласованию вергенции - аккомодации. Во втором случае глаза наблюдателя аккомодировались к конечному расстоянию наблюдения, равному 565мм, что приводит к расфокусировке изображения по причине неподвижности ЖК-дисплея. Таким образом, в обоих приведенных случаях мы наблюдаем рассогласование вергенции - аккомодации.

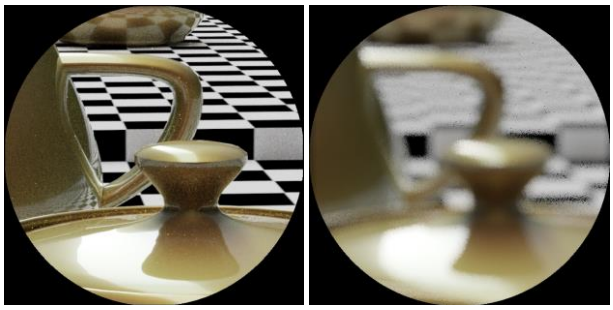


Рис. 6. Моделирование изображений виртуальной сцены наблюдаемой левым глазом через оптическую систему виртуальной реальности. ЖК экран расположен в фокусе окуляра. Слева: глаз сфокусирован на бесконечность. Справа: глаз сфокусирован на конечное расстояние 565мм.

Для исключения рассогласования вергенции и аккомодации требуется строить изображения для естественного значения аккомодации (в данном случае 565мм) одновременно с изменением параметров оптической системы формирования изображения (фокусного

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

расстояния оптической системы и расстояния до ЖК экрана). На рисунке 7 представлен результат синтеза изображения сцены, наблюдаемый в системе виртуальной реальности при условии, что изображение, сформированное на ЖК экране, фокусное расстояние оптической системы и расстояние до ЖК экрана выставлены для соблюдения условий естественного восприятия. Видно, что эффект конфликта вергенции и аккомодации зрения отсутствует.

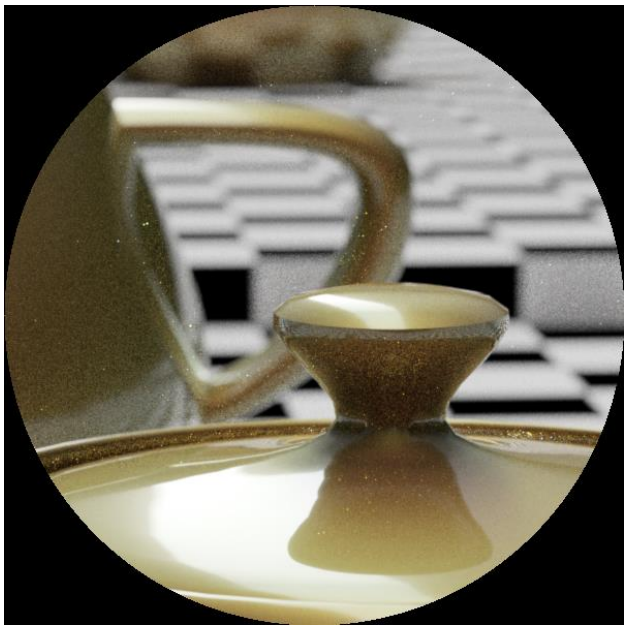


Рис. 7. Наблюдение виртуальной сцены левым глазом через систему виртуальной реальности. Фокусное расстояние оптический системы и расстояние до ЖК экрана скорректированы. Глаз сфокусирован на расстоянии в 565мм.

5. Заключение

В работе был представлен метод использования виртуального прототипирования для проектирования систем виртуальной, дополненной и смешанной реальностей. Данный метод позволяет как выявить и оценить конфликт вергенции и аккомодации, так и разработать устройство, не подверженное рассматриваемому эффекту. Был представлен пример схемы устройства, не подверженного эффекту конфликта вергенции и аккомодации зрения. Дальнейшие исследования будут направлены на методы количественного анализа присутствия конфликта вергенции и аккомодации в виртуальной, дополненной и смешанной реальности.

6. Благодарности

Данная работа была частично профинансирована грантами РФФИ № 16-01-00552 и 18-08-01484.

7. Литература

[1] Жданов Д.Д., Галактионов В.А., Волобой А.Г., Жданов А.Д., Гарбуль А.А., Потемин И.С., Соколов В.Г., “Методы фотореалистичной визуализации изображений, формируемых оптическими системами дополненной реальности”, Программирование, 2018, №4.

[2] Konrad R., Cooper E.A., Wetzstein G., “Novel Optical Configurations for Virtual Reality: Evaluating User Preference and Performance with Focus-tunable and Monovision Near-eye Displays” ACM CHI, 2016.

[3] Kramida G., “Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays” IEEE TVCG 22, 7, pp. 1912–1931, 2016.

[4] LaValle S.M., “VIRTUAL REALITY” University of Illinois, Cambridge University Press, 2017.

[5] Matsuda N., Fix A., Lanman D., “Focal Surface Displays” ACM Trans. Graph. 36, 4, Article 86, July 2017.

[6] Peli E., “Optometric and perceptual issues with head-mounted displays. In Visual Instrumentation” Optical Design and Principles, P. Mouroulis (Ed.). McGraw-Hill, pp. 205-275, 1999.

[7] Wetzstein G., “A personalized VR/AR system that adapts to the user is crucial to deliver the best possible experience” The BRIDGE, Vol. 46, No. 4, Winter 2016.

[8] "Lumicept – Hybrid Light Simulation Software," Integra Inc., 2018, <http://www.integra.jp/en/products/lumicept>.

Об авторах

Жданов Дмитрий Дмитриевич, к.ф.-м.н., заведующий кафедрой технологий визуализации Университета ИТМО. Его e-mail ddzhdanov@mail.ru.

Потемин Игорь Станиславович, к.т.н., доцент кафедры технологий визуализации Университета ИТМО. Его e-mail ipotemin@yandex.ru.

Богданов Николай Николаевич, аспирант кафедры технологий визуализации Университета ИТМО. Его e-mail nnbogdanov@corp.ifmo.ru.

Жданов Андрей Дмитриевич, аспирант кафедры технологий визуализации Университета ИТМО. Его e-mail adzhdanov@corp.ifmo.ru.

Галактионов Владимир Александрович, профессор, доктор физ.-мат. наук, заведующий отделом компьютерной графики и вычислительной оптики ИПМ им. М.В.Келдыша РАН. Его e-mail vlgal@ginkeldysh.ru.

Лившиц Ирина Леонидовна, старший научный сотрудник, к.т.н., Заведующий научно-технической лабораторией автоматизированного проектирования оптико-информационных и энергосберегающих систем Университета ИТМО. Ее e-mail ecenter-optica10@yandex.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

ISO 12233 и программные средства измерения частотно-контрастных характеристик изображений

Рудникович А.С., Шипунова К.В.
ra@diviline.ru | shipunovakseniya@mail.ru
Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

В работе рассмотрен метод измерения частотно-контрастных характеристик (ЧКХ) на основе наклонной границы объекта, выполнен краткий обзор соответствующего программного обеспечения, представлены результаты измерения ЧКХ с использованием специальных программ (Imatest, MTF Mapper, Slanted Edge MTF, Quick MTF, DiViLab-MTF). Для измерений использованы изображения в соответствии с ISO 12233, а также изображения со специально введенными отклонениями (шум, угол наклона).

Ключевые слова: *параметры изображений, частотно-контрастная характеристика, ЧКХ, ФРП, программные средства измерений, стандартизация измерений, ISO 12233, точность измерений.*

ISO 12233 and software tools for measuring of images modulation transfer function

Rudnikovich A, Shipunova K.
ra@diviline.ru | shipunovakseniya@mail.ru
Tomsk state university of control systems and radio electronics, Tomsk, Russia

The paper describes the main ESF based method of modulation transfer function (MTF) measuring, available a brief overview of the relevant software and the results of MTF measuring by such tools (Imatest, MTF Mapper, Slanted Edge MTF, Quick MTF, DiViLab-MTF). Images in accordance with ISO 12233, as well as images with specially introduced deviations (noise, angle) have been used for the experiment.

Keywords: *image quality, modulation transfer function, MTF, ESF, ISO 12233, measurement software, measurement accuracy.*

1. Введение

Разрешающая способность, функция рассеяния перепада (ФРП) и частотно-контрастная характеристика (ЧКХ) это набор связанных и наиболее важных характеристик для оценки качества изображений современной фото- видеоаппаратуры.

Существует несколько методов измерений ЧКХ. Наибольшее распространение получил метод, который основан на использовании наклонной границы объекта, именно этот метод описан в стандарте ISO 12233 [1-3]. Стандартизация алгоритма измерений, условий освещения и тестовых таблиц должна была обеспечить приемлемую точность измерений различных программных инструментов для измерения ЧКХ современной фото-видеоаппаратуры, однако, на сколько разработчики таких инструментов следуют стандарту – вопрос открытый.

При работе с видеоданными существующие программы применимы, но при этом требуются дополнительные манипуляции с видеофайлами для извлечения изображений. Практика измерений ЧКХ большого количества видеоустройств показала, что выбор кадра для измерений может существенно влиять на результат даже в лабораторных условиях с соблюдением всех требований ISO 12233. Ряд устройств видеозаписи в процессе записи продолжают выполнять подстройку параметров, кодеки используют различные типы кадров влияют на размытие элементов изображений, действия оператора (включение и выключение записи) приводят к смазам, мешающим получить истинную ЧКХ.

Одним из применений измерения ЧХК является видеотехническая экспертиза. Анализ профиля перепада яркости сигнала на границах объектов может служить инструментом при выявлении фактов внутрикадрового монтажа или других манипуляций с изображениями. При этом, интерес представляет не только частота при заданном уровне контраста (например, 30% или 50%) а также форма ЧКХ и профиля перепада уровня сигнала в различных направлениях.

Несовершенство оптики современных фото-видеоустройств порождает неоднородность характеристик в плоскости изображений. Как правило, разрешение выше в центральной части изображений и уменьшается при удалении от центра в разной степени для различных устройств. Нередко наблюдается выраженное смещение оптического центра. Поэтому для решения идентификационных задач видеотехнической экспертизы необходим инструмент способный осуществлять измерение пространственного распределения характеристик по всей плоскости изображений.

Зачастую оценка ЧКХ требует многократных повторов, при этом критически важной является возможность точно определить область измерений. Данные таблиц этой работы могли бы быть более точны, если бы все программы позволяли задать одинаковую область измерений (ROI). При сравнении ЧКХ различных фото- видеоустройств отсутствие возможности задать одинаковую область измерений приводит к увеличению погрешности измерений и дополнительным ошибкам оператора, которые практически неизбежны, когда речь идет о сотнях измерений.

Набор описанных факторов, требований к ПО и отсутствие специализированных инструментов измерения ЧКХ по видеоданным мотивировали разработку нового программного модуля в составе программного обеспечения DiViLab [4]. Одним из авторов данной работы разработан программный модуль, который далее обозначен как DiViLab-MTF. Данный модуль, также, как и другие программы, ориентирован на измерение ЧКХ по перепаду уровня сигнала, при этом разрабатывался он с учетом перспектив измерения ЧКХ по реальным изображениям.

Целью работы являлось сопоставление результатов измерений разными программами, оценка эффективности текущей версии DiViLab-MTF, актуализация вопроса эффективности стандартизации измерений ЧКХ, а также

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

оценка пригодности существующих программ с точки зрения видеотехнической экспертизы.

2. Измерения ЧКХ с использованием наклонной границы объекта

ЧКХ связана с разрешающей способностью, и также характеризует способность передавать мелкие близко расположенные фрагменты изображений. ЧКХ показывает зависимость коэффициента передачи контраста от пространственной частоты. Обычно количественная оценка ЧКХ это значение пространственной частоты на половине максимального значения контраста (ЧКХ50).

График ЧКХ описывает параметры оптико-электронной системы более полно, чем величина разрешающей способности при определенном уровне контраста. Наибольшее распространение получил метод измерения ЧКХ на основе оценки параметров наклонной границы объекта (перепада уровня сигнала), описанный в стандарте ISO 12233.

Функция рассеяния перепада (ФРП) или edge spread function (ESF), является основой оценки частотных свойств системы по изображениям. Функция показывает, как изменяется уровень сигнала в месте перехода от темного к светлому и наоборот. Зона перехода называется зоной нерезкости и по её продолжительности, а также по крутизне перехода судят о частотных свойствах изображения и соответственно о свойствах системы, которой оно получено.

Первое издание стандарта ISO 12233, выпущенное в 2000 году, предоставило научно-обоснованный и легко реализуемый способ измерения ЧКХ систем, включающих дискретный датчик изображения, например, цифровых фотокамер.

В стандартах ISO:12233 описан алгоритм измерения ЧКХ по изображениям с контрастным наклонным объектом на светлом фоне. Алгоритм основан на автоматическом измерении пространственно-частотного отклика в области интереса (ROI), которая представляет собой наклоненную вертикально кромку с переходом между уровнями сигнала.

В использованной области интереса, происходит перебор строк изображения и вычисляется производная для получения функции размытия линии (ФРЛ). ФРП как обобщение ФРЛ, получается путем усреднения данных для нескольких строк изображения (рис. 1). Это позволяет увеличить частоту дискретизации ФРП и соответственно разрешение итоговой ЧКХ. Для уменьшения влияния шумов стандарт предусматривает применение к ФРП функции окна Хэмминга. Применение дискретного преобразования Фурье к ФРП, позволяет получить частотно-контрастную характеристику или функцию передачи модуляции (MTF).

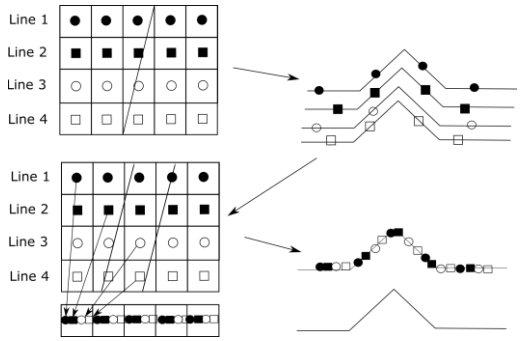


Рис. 1. Алгоритм вычисления ЧКХ, ISO:12233 2000 [1].

В последующих версиях стандарта 2014 и 2017 года в отношении этого алгоритма наиболее существенное изменение коснулось контраста тестовой таблицы, т.к. высокий контраст таблицы первой редакции стандарта приводил к значимым нелинейным искажениям.

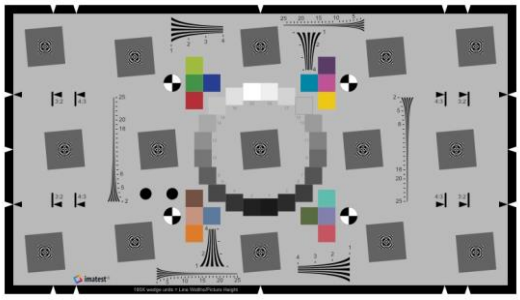


Рис. 2. Тестовая таблица ISO:12233 2014 [6].

Одним из недостатков описанного метода является сложность точного совмещения центров строк изображений при расчете ФРП. При этом методы на основе ФРТ, ФРЛ и ФРП представляется единственно возможными для измерений ЧКХ по изображениям реальных сцен.

3. Обзор и экспериментальное сравнение ПО для измерения ЧКХ

При выборе программного обеспечения измерения ЧКХ логичными представляются следующие специальные требования: возможность выбора заданной области интереса на изображении (повторяемость эксперимента); устойчивость измерений к шумам на изображении; возможность работы с видеоданными.

В результате обзора программных инструментов измерения ЧКХ для проведения эксперимента и сравнения функциональных особенностей были выбраны пять наиболее известных программных продуктов:

1. Imatest 5.0.16. Компания производитель предоставляет клиентам программное обеспечение, тестовые таблицы и оборудование для определения качества изображения и их соответствия стандартам [5].
2. MTF Mapper 0.6.18 - это набор инструментов для измерения разрешения на основе значений MTF50. Основной инструмент может автоматически обнаруживать темные прямоугольные объекты на светлом фоне и извлекать значения MTF50 по каждому краю этих объектов [6].
3. Slanted Edge MTF 2.0 (плагин для ImageJ) - выполняет расчеты ЧКХ до удвоенной частоты Найквиста над наклонными объектами в изображении с вертикальным краем с темной стороной слева [7].
4. Quick MTF 2.11 – это программа для быстрой оценки MTF по отдельным изображениям [8].
5. DiViLab-MTF 2.0.0.1684 – инструмент для измерения ЧКХ заданной области изображения, позволяет работать с видеоданными [4].

Все перечисленные программы выполняют измерение ЧКХ по изображению наклонного края объекта.

Выполнение экспериментальных измерений позволило сделать в таблице 1 некоторые обобщения функциональных возможностей программ. Номера в названиях столбцов приведенных далее таблиц соответствуют порядку перечисления программ:

1. Imatest.
2. MTF Mapper.
3. Slanted Edge MTF.
4. Quick MTF (настройки по умолчанию).
5. Quick MTF (настройки по ISO 12233).
6. DiViLab - MTF.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Таблица 1.

Функциональные характеристики ПО	1	2	3	4, 5	6
Использование ФРП для измерения ЧКХ	+	+	+	+	+
Вывод графиков ЧКХ и функции перепада яркости	+	+	+	+	+
Возможность выбора области изображения для измерений	+	-	-	+	+
Возможность точно задать ROI (повторяемость измерений)	+	-	-	-	+
Пригодность для измерения отдельно в вертикальном или горизонтальном направлении	+	-	+	+	+
Возможность работы с видеопоследовательностью	-	-	-	-	+
Поддержка различных форматов изображений	+	+	+	+	-
Устойчивость к шумам	+	+	-	-	+

Обзор доступной информации об этих программах показал, что производители этих программ стараются внести улучшения в алгоритм измерений по отношению к ISO 12233.

В качестве исходных изображений при проведении эксперимента использовались изображения с разрешением 1920x1080 (рис. 3), полученные из файлов с векторной графикой. В соответствии с ISO 12233 угол наклона объектов базового изображения составил 5°, контраст 4:1.

Рис. 3. Исходное изображение

Результаты измерения ЧКХ50 для размытых фильтром Гаусса изображений с разными радиусами размытия, представлены в таблице 2.

Таблица 2.

Радиус размытия	**1**	**2**	**3**	**4**	**5**	**6**
0,5	0,374	0,384	0,258	0,381	0,287	0,285
1	0,183	0,188	0,161	0,187	0,17	0,157
1,5	0,125	0,127	0,112	0,131	0,12	0,098

Измеренные значения ЧКХ50 существенно различаются, при этом, отклонение результатов измерений достигает 33% при значениях ЧКХ50 близких к 0,3 частоты дискретизации (радиус размытия 0,5, таблица 2). Можно предположить, что это связано с использованием различных алгоритмов интерполяции области перепада сигнала на этапе передискретизации. На рисунках 4 и 5 приведены результаты, полученные соответственно в программах Imatest и Quick MTF при одинаковых условиях измерений. Стоит отметить, что при получении данных таблицы 2 изменения в расположении области интереса (ROI) и ее размеров в пределах ±20%, при размере 200x300 элементов, не оказывают существенного влияния на результаты измерений.

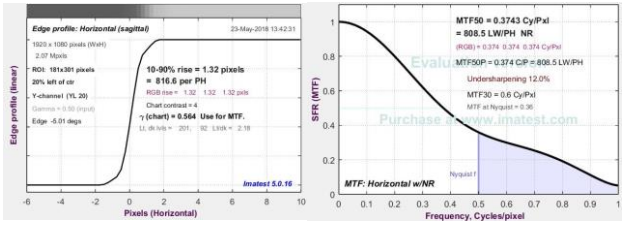


Рис. 4. Пример результата ПО Imatest (MTF50 = 0,374)**.**

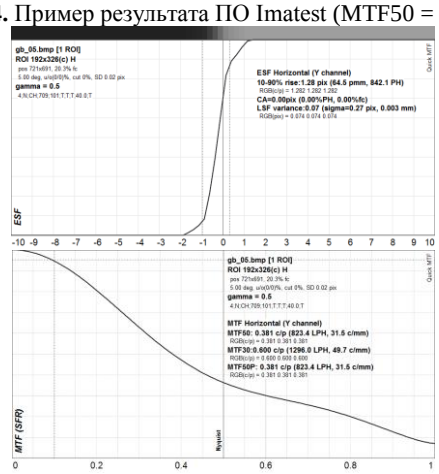


Рис. 5. Пример результата ПО Quick MTF (MTF50 = 0,381).

Отдельные применения для измерений ЧКХ требуют устойчивости к высокому уровню шумов. Исследователи постоянно возвращаются к вопросам устойчивости методов измерения ЧКХ к шумам [9]. Особенно это актуально при выполнении измерений по изображениям реальных сцен, медицинских и спутниковых изображений. Если исходить из предположения, что в изображении присутствуют границы объектов, соответствующие требованиям к контрасту и углу наклона, то определяющим фактором применимости ПО становится устойчивость к шумам. В таблице 3 проведены результаты измерений ЧКХ50 для изображения с радиусом размытия 0,5 с добавлением нормального шума. Параметры добавленного шума приведены в столбцах СКО и PSNR.

Таблица 3.

СКО	**PSNR**	**1**	**2**	**3**	**4**	**5**	**6**
5	34,098	0,381	0,379	0,274	0,381	0,277	0,277
10	28,121	0,371	0,391	0,25	0,295	0,295	0,289
15	24,604	0,34	0,342	0,274	0,27	0,27	0,285
20	21,703	0,347	0,432	NaN	NaN	NaN	0,301
25	19,911	0,339	0,35	NaN	NaN	NaN	0,295
30	18,476	0,415	0,416	NaN	NaN	NaN	0,312
35	17,086	0,357	0,473	NaN	NaN	NaN	0,271
40	15,942	0,347	0,448	NaN	NaN	NaN	0,244
45	14,984	0,204	NaN	NaN	NaN	NaN	0,148
50	14,033	0,143	0,41	NaN	NaN	NaN	0,078
	MAD:	**0,092**	**0,073**	**0,194**	**0,287**	**0,214**	**0,056**

Как видно из таблицы 3, при существенных уровнях шумов на изображениях работоспособность сохраняют не все программы. В качестве меры устойчивости к шуму использовано среднее абсолютное отклонение (MAD) относительно значений, полученных при отсутствии шума (строка 1 таблицы 2). Плагин для ImageJ и Quick MTF не позволяют измерять ЧКХ при СКО от 20 и более. Существенная разница результатов измерений программ, сохранивших работоспособность при высоком уровне шумов, говорит, как минимум, об использовании

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

различных алгоритмов интегрирования данных строк изображения (ROI).

Отдельный интерес представляет возможность использования существующего программного обеспечения для оценки ЧКХ по изображениям реальных сцен. Очевидно, что при этом критичным будет чувствительность ПО к форме и наклону перепада яркости в плоскости изображения.

Известны работы, в которых дается оценка устойчивости алгоритмов измерений к углу наклона объекта, предложены различные методы компенсации, позволяющие сохранить погрешность измерений в пределах 1% [10], при этом не удалось найти результаты таких измерений для большей части доступного ПО. В таблице 4 приведены результаты измерений ЧКХ50 при разных углах наклона объектов. Среднее абсолютное отклонение рассчитано относительно значений, полученных при наклоне сторон объектов на 5° (строка 2 таблицы 4).

Таблица 4.

Угол, град.	**1**	**2**	**3**	**4**	**5**	**6**
0	0,256	NaN	0,258	0,262	0,262	0,263
5	0,374	0,384	0,258	0,381	0,287	0,285
10	0,372	0,382	0,274	0,375	0,290	0,283
15	0,371	0,380	0,274	0,362	0,290	0,279
20	0,370	0,377	0,258	0,350	0,292	0,274
25	0,367	0,373	0,241	0,337	0,297	0,264
30	0,363	0,370	0,241	0,300	0,319	0,255
35	0,362	0,368	0,225	0,300	0,305	0,245
40	0,360	0,366	0,230	0,281	0,310	0,233
45	0,324	0,338	NaN	0,244	0,317	0,222
MAD:	**0,025**	**0,056**	**0,043**	**0,069**	**0,017**	**0,027**

Из таблицы 4 можно сделать следующие выводы:

1. ПО MTF Mapper неспособно работать с вертикальной границей (угол наклона объекта 0 град.).

2. ПО Imatest и Quick MTF (настройки по умолчанию) демонстрирует существенное отклонение результата, полученного на строго вертикальной границе, от результатов на наклонных перепадах. Можно предположить, что на вертикальной границе при передискретизации эти программы начинают использовать линейную интерполяцию.

3. Значения ЧКХ50, полученные в Quick MTF, существенно зависят от угла наклона объекта, при этом, разнонаправленно, в зависимости от настроек. Quick MTF (настройки по умолчанию) демонстрирует уменьшение ЧКХ50 на 36% с увеличением угла наклона. Quick MTF (настройки по ISO 12233) демонстрирует увеличение ЧКХ50 на 10% с увеличением угла наклона. Такие результаты представляются странными, особенно при сравнении с результатами DiViLab-MTF (уменьшение ЧКХ50 на 22% при отсутствии компенсация угла наклона границ объектов), которые в большей мере согласуются с другими исследованиями [10].

4. Заключение

В таблицах продемонстрированы заметные отклонения результатов работы программ для измерения ЧКХ. В целом по результатам обзора и эксперимента можно сделать вывод о том, что разработчики ПО для измерения ЧКХ стараются усовершенствовать свое ПО и вносят изменения в алгоритм, предусмотренный стандартом ISO 12233.

В условиях существования международного стандарта, полученные отклонения результатов измерений ЧКХ представляются значительными. Наибольшие расхождения наблюдаются при более резких перепадах уровня сигнала. Такие расхождения объясняются применением различных алгоритмов интерполяции при передискретизации сигнала и различными методами поиска центров ФРЛ при их интегрировании для получения ФРП.

Полученные результаты актуализируют вопрос о точности абсолютных значений результатов измерений современных мерительных инструментов и эффективности международной стандартизации измерений ЧКХ стандартом ISO 12233.

Не смотря на относительную простоту базовой части алгоритма измерений ЧКХ и множества публикаций, в которых данный алгоритм воспроизводится, отсутствует доступный программный инструмент, позволяющий эффективно работать с видеоданными.

5. Литература

[1] ISO 12233:2000. [Электронный ресурс]. Режим доступа:https://www.iso.org/obp/ui/#iso:std:iso:12233:ed-1:v1:en (дата обращения: 10.05.2018).

[2] ISO 12233:2014. [Электронный ресурс]. Режим доступа: https://www.iso.org/standard/59419.html (дата обращения: 28.03.2018).

[3] ISO 12233:2017. [Электронный ресурс]. Режим доступа: https://www.iso.org/obp/ui/#iso: std: iso:12233: ed-3: v1: en (дата обращения: 28.03.2018).

[4] Рудникович А.С., Крутовская Н.И.: «Подготовка специалистов в области видеоинформационных технологий с использованием программного обеспечения DiViLab» // Современное образование: развитие технологий и содержания высшего профессионального образования как условие повышения качества подготовки выпускников: материалы междунар. науч. метод. конф., 26-27 января 2017 г. Россия Томск. – Томск: Изд-во Томск. гос. ун-та систем упр. и радиоэлектроники 2017. – С. 184-186.

[5] Imatest. [Электронный ресурс]. Режим доступа: http://www.imatest.com/ (дата обращения: 28.03.2018).

[6] MTF mapper. [Электронный ресурс]. Режим доступа: https://sourceforge.net/projects/mtfmapper/ (дата обращения: 28.03.2018).

[7] Slanted Edge MTF. [Электронный ресурс]. Режим доступа: https://imagej.nih.gov/ij/plugins/se-mtf/index.html (дата обращения: 28.03.2018).

[8] Quick MTF, an image quality testing application. [Электронный ресурс]. Режим доступа: http://www.quickmtf.com/ (дата обращения: 28.03.2018).

[9] Antonio González-López, “Effect of noise on MTF calculations using different phantoms”, Medical PhysicsVolume 45 (03 March 2018); https://doi.org/10.1002/mp.12847.

[10] Jackson K. M. Roland, "A study of slanted-edge MTF stability and repeatability", Proc. SPIE 9396, Image Quality and System Performance XII, 93960L (8 February 2015); https://doi.org/10.1117/12.2077755.

Об авторах

Шипунова Ксения Владимировна, магистрант кафедры телевидения и управления Томского государственного университета систем управления и радиоэлектроники, e-mail: shipunovakseniya@mail.ru.

Рудникович Андрей Сергеевич, аспирант кафедры телевидения и управления Томского государственного университета систем управления и радиоэлектроники, e-mail: ra@diviline.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Съемка сферических панорам с каналом освещенности

М.И. Свириденко
twistms@mail.ru
Санкт-Петербургский Политехнический Университет, Санкт-Петербург, Россия

Данная работа описывает концепцию нового метода съёмки сферических панорам с информацией об освещенности. Информация об освещенности получается путем сьемки сферических панорам массивом фотоприёмных устройств. Получаемые сферические панорамы могут использоваться при проведении трёхмерной визуализации, при этом данные освещенности используются как информация для освещения в виртуальном пространстве, а сферические панорамы, полученные при помощи мини-камер устройства, для построения верных отражений на трехмерной модели.

Ключевые слова: *сферическая панорама, визуализация, рендеринг, стабилизация видеоизображения, HDRI.*

Shooting spherical panoramas with the channel of illumination

M.I. Sviridenko
twistms@mail.ru
St. Petersburg Polytechnic University, Saint Petersburg, Russia

This work provides the new concept disclosing a method of shooting panoramic videos with an additional illumination source. The information about illumination is obtained from 3D video-footage taken by an array with a photodiodes. The resulting panoramic video can be used for creating three-dimensional visualization, wherein the information about illumination can be used for illuminating virtual space, as well as channel image data can be used for constructing correct reflections on a three-dimensional model.

Keywords: *spherical panorama, visualization, rendering, video image stabilization, HDRI.*

1. Введение

Разрабатываемая концепция съемки с каналом освещенности позволяет упростить процесс освещения виртуальных сцен путем съемки двух типов сферических панорам: сферическая панорама для использования в качестве освещения в 3D сценах и сферическая панорама для использования в качестве источника отражений в 3D. Причем съёмка должна вестись относительно объекта интереса и с требуемой частотой кадров.

В настоящее время в кинопроизводстве при освещении трёхмерных объектов для интеграции в видеоизображение в основном пользуются двумя способами: используют для освещения сферическую панораму с расширенным динамическим диапазоном [5], либо используют виртуальные источники света. В первом случае из фотоснимков создается HDR-панорама (High Dynamic Range), далее она используется как источник света и отражений в виртуальных сценах (IBL – Image-Based Lighting) [4]. Чем выше разрешение и динамический диапазон яркости в такой панораме – тем лучше качество интеграции трёхмерных моделей в видеоизображение. Для создания качественных сферических HDR-панорам требуется много времени, это трудоёмкий процесс, итоговое изображение должно пройти несколько этапов создания, включающих вмешательство специалиста:

1. Съемка фотокамерой во всех направлениях для создания полноценной сферической панорамы с учетом особенностей съемки для получения расширенного диапазона яркости.
2. Обработка изображений (компенсация недостатков оптической системы камеры, получение расширенного диапазона яркости из серии снимков с разной экспозицией и др.); даже автоматические режимы обработки потребуют вмешательство специалиста. [7].
3. Сшивание снимков в сферическую HDR-панораму [8].

Во втором случае снимается обычная сферическая панорама, на основе которой специалисты по виртуальному освещению выстраивают источники света, настраивая интенсивность данных источников, пользуясь различными вспомогательными устройствами. В этом варианте для отражений пространства на трёхмерной модели может потребоваться моделирование окружения, таким образом, этот метод также является достаточно трудоёмким.

Основными задачами являются получение цепочки сферических панорам с информацией для освещения, получение цепочки сферических панорам с информацией для отражений и автоориентирование этих панорам (для дальнейшего использования в трёхмерных сценах потребуются автоориентирование и стабилизация всего отснятого материала относительно одной системы координат) [3].

2. Реализация

Реализация концепции представляется устройством являющимся системой камер расположенных на сферическом корпусе во всех направлениях для съемки панорам. Помимо камер, на устройстве планируется равномерно расположить фотоприёмные устройства для съемки канала освещенности. Предлагается использовать фотодиоды, поскольку они дают меньше шума в данных при работе в фотоприёмном режиме, нежели другие типы доступных фотоприёмных устройств [1]. И последним важным элементом будет являться наличие датчиков ориентации в пространстве: цифровые гироскоп, акселерометр и магнитометр [2].

Одной из важных задач реализации является настройка синхронизации съемки видеокамерами и записи данных со всех датчиков. Для решения этой задачи предлагается использовать цифровые счетчики, аналого-цифровые преобразователи и микропроцессор. После съемки потребуется обработка всех данных и для этого

планируется создание отдельной компьютерной программы, которая будет выполнять в автоматическом режиме основные задачи:

1. Склейку видеоизображений с камер в единую сферическую видеопанораму.
2. Создание цепочки сферических панорам канала освещенности.
3. Обработку данных с фотодиодов и данных с датчиков ориентации в пространстве.
4. Ориентирование и стабилизацию обработанных материалов.

В силу того, что отражения на трехмерной модели будут искажены физическими свойствами материала, применение панорам в качестве информации для отражений в виртуальном пространстве не требует больших размеров исходных изображений, из чего следует, что для устройства не требуются дорогостоящие и массивные камеры для съемки. Использование мини-камер имеет некоторые преимущества: масса устройства будет меньше, данные будут занимать меньший объем, мини-камеры энергоэффективнее, изображения небольшого размера гораздо быстрее обрабатываются. Все собранные данные будут записываться на карту памяти устройства.

Равномерно расположенные на устройстве (рис. 1) фотодиоды будут синхронно снимать данные освещенности с частотой от 25 раз в секунду (аналогично частоте кадров в секунду у камер). Обработанные данные будут представлять собой сферические панорамы с расширенным диапазоном яркости. Полученные панорамы будут иметь только информацию об освещенности, использование этой информации в качестве данных для корректного освещения в виртуальном пространстве невозможно, поскольку снимаемый свет должен иметь информацию о цвете. Для получения информации о цвете предлагается использовать сферические цепочи панорам из мини-камер устройства.

Съемка камерами должна вестись параллельно со сбором данных с фотоприёмных устройств. Поскольку количество фотодиодов на устройстве не позволит получить точную информацию об освещенности для каждого пикселя сферической панорамы (рис. 2), то для получения корректной карты освещенности потребуется добавить информацию о цвете и границах обьектов пространства (рис. 3). Эту информацию предлагается брать из сферических панорам, снятых мини-камерами устройства (рис. 4). Такой метод позволит сохранить резкость краев источника света и избежать избыточной информации об освещенности в тех областях изображения, где на самом деле нету источников света. Также это решит задачу с добавлением информации цвета в итоговые сферические панорамы для освещения 3D сцен. [9]. Предлагается использование 62-х фотодиодов, которые потребуется равномерно расположить на сферическом корпусе устройства. В дальнейшем, для более точного сбора данных, это количество можно будет увеличить и прийти к оптимальному соотношению цены и качества.

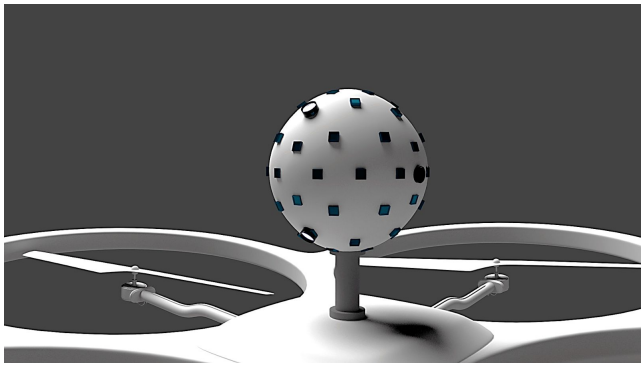


Рис. 1. Концепт устройства.

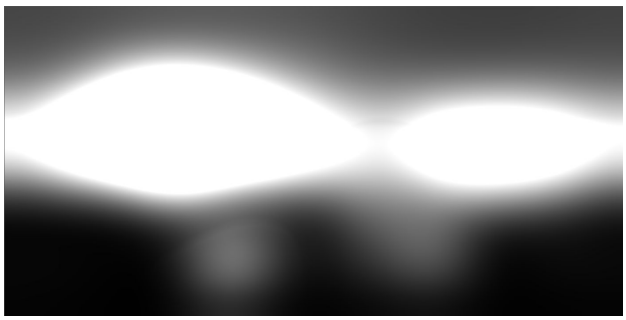


Рис. 2. Пример необработанной сферической панорамы с информацией об освещенности, полученной при помощи фотодиодов.

Рис. 3. Пример итоговой сферической панорамы с информацией об освещенности, с границами объектов и с цветом.

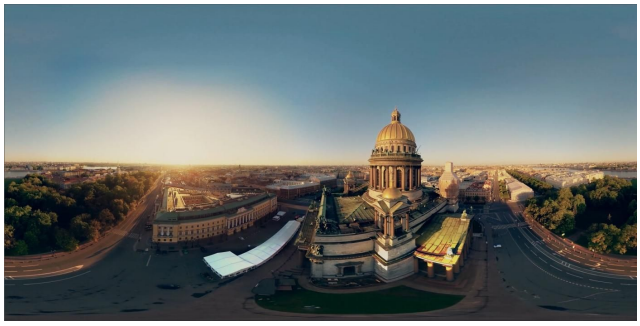


Рис. 4. Пример итоговой сферической панорамы, полученной при помощи мини-камер, для использования в качестве отражений в 3D сценах.

Для использования полученных видеопанорам в виртуальных сценах потребуется их ориентация в пространстве. Для достижения поставленной задачи предлагается разместить в устройстве цифровые гироскоп, акселерометр и магнитометр, данные с которых будут записываться синхронно со сбором данных с камер и фотодиодов. При помощи полученных данных мы сможем автоматически ориентировать цепочки полученных сферических панорам.

Поскольку устройство планируется устанавливать на различные объекты, то при съемке будет перекрываться часть обзора камер самим объектом. Эту проблему можно решить только генерацией части перекрытого изображения путем отслеживания контрастных точек в видеоизображении.

Управление устройством предлагается по таким беспроводным каналам связи как Wi-Fi или Bluetooth, что подразумевает разработку необходимого ПО, как для мобильных устройств, так и для персональных компьютеров.

3. Применение

Алгоритм действий специалиста при использовании предлагаемого устройства:

1. Установка устройства на объект интереса (часть окружения сцены или любой объект, заменяемый в последствии трехмерной моделью)
2. Соединение с устройством по беспроводному каналу связи (Wi-Fi или Bluetooth), через мобильные устройства или через персональные компьютеры.
3. Настройка устройства перед съёмкой (частота кадров, размер изображений и т.д.)
4. Управление устройством посредством отправки простейших команд, таких как “запись”, “стоп”, “выключение” и т.д.
5. Копирование всех собранных устройством данных на персональный компьютер, используя либо беспроводную связь с устройством, либо копируя данные напрямую с карты памяти устройства.
6. Выбор нужных отрезков времени съёмки и автоматическая обработка полученных данных при помощи ПО.
7. Передача двух итоговых цепочек сферических панорам (с информацией для освещения и с информацией для отражений в 3D сценах) специалистам по компьютерной графике для дальнейшего выполнения поставленных задач. Данные цепочки сферических панорам можно будет применять во всех современных редакторах трехмерной компьютерной графики.

Планируемая масса устройства не более 300 г., это позволит устанавливать его на малые беспилотные летательные аппараты и не только. Установив устройство на беспилотный аппарат, будет возможным вести съемку как в павильоне, так и за его пределами, и в дальнейшем качественно интегрировать трёхмерную модель в видеоизображение, например, заменив на этапе монтажа и постобработки малый БПЛА на требуемый объект трёхмерной графики.

После обработки всех собранных с устройства данных, мы получим цепочку сферических панорам, для использования в качестве информации для отражений в виртуальном пространстве и цепочку сферических панорам с расширенным диапазоном яркости для использования в качестве информации для освещения в виртуальном пространстве.

В качестве примера можно рассмотреть определенную ситуацию на съемочной площадке - предположим, что в павильоне для съемок расставлены источники света, декорации и присутствуют актеры. По сценарию требуется снять полет летательного аппарата по освещенному туннелю, полет планируется от начала и до конца туннеля. Для этой задачи будет выгодно использование разрабатываемого устройства: установив его на малый БПЛА, мы сможем синхронно снимать карту освещенности и карту отражений относительно объекта который в дальнейшем потребуется заменить трёхмерной моделью. При помощи полученных сферических панорам будет возможно качественно визуализировать интегрированную в видеоизображение трехмерную модель, так как в каждый момент времени на объекте будут корректные отражения окружения и корректное освещение относительно нужной точки пространства.

4. Заключение

Прототип устройства имеет свои недостатки и преимущества.

Преимущества:

1. Малый вес устройства.
2. Простота использования.
3. Дешевизна производства.
4. Новизна, заключающаяся в гибридной синхронной съёмке фотодиодами и камерами.

Недостатки:

1. Недостаточное разрешение мини-камер.
2. Относительно малое количество фотодиодов для создания более точных сферических панорам с информацией об освещенности.
3. Частота кадров ограничена возможностями микроконтроллера и мини-камер.
4. Невозможность просмотра полученного результата в режиме реального времени.
5. Использование двух типов сферических панорам может вызывать затруднения при работе с ними в трехмерных редакторах.

В целом, разрабатываемое устройство дает возможность сократить время и бюджет, затраченные на создание определенных сцен в кино, музыкальных клипах или любительских видеороликах, а концепция съемки канала освещенности может применяться не только в киноиндустрии, но и в технологиях виртуальной реальности (VR), компьютерных играх, военных разработках. Автоматическая ориентация сферических видеоизображений при помощи данных с датчиков ориентации может найти применение в различных типах съемки, когда съемка ведется с движущегося (вращающегося) объекта, а внимание зрителя требуется в определенном направлении. На данный момент при просмотре таких видеороликов приходится самостоятельно вращать сферическое видеоизображение до точки интереса, что может вызывать дискомфорт или головокружение [6].

5. Благодарности

Автор выражает благодарность управляющему директору студии ЛенVR Никонову Максиму за предоставленное разрешение на использование сферической панорамы.

6. Литература

[1] Аксененко М.Д., Бараночников М.Л., Смолин О.В. Микроэлектронные фотоприемные устройства. М.: Энергоатомиздат. 1984.

[2] Игнатьев А.А., Проскуряков Г.М., Васильев А.В. Алгоритмы работы миниатюрной системы ориентации вращающихся подвижных объектов//Гетеромагнитная микроэлектроника: сб. науч. тр. Саратов: Изд-во Сарат. ун-та, 2014. Вып. 17: Гетеромагнитная микро-и наноэлектроника. Методические аспекты физического образования. Экономика в промышленности. С. 78-91.

[3] Свириденко М.И. Использование цепочки сферических панорам для качественной интеграции трехмерных моделей в видеоизображение. Сборник докладов недели науки. Санкт-Петербург: СПбГПУ, 2015, 121с.

[4] Debevec P. 2005. Image-based lighting. In ACM SIGGRAPH 2005 Courses (SIGGRAPH '05), John Fujii (Ed.). ACM, New York, NY, USA, Article 3.

[5] Heckenberg D., Agland S., leBlanc J.P., Barth R. 2017. Automated light probes from capture to render for Peter Rabbit. In ACM SIGGRAPH 2017 Talks (SIGGRAPH '17). ACM, New York, NY, USA, Article 17, 2 pages.

[6] Hong S., Kim G.J., 2016. Accelerated viewpoint panning with rotational gain in 360 degree videos. In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST '16). ACM, New York, NY, USA, 303-304.

[7] Jinno T., Okuda M., "Multiple Exposure Fusion for High Dynamic Range Image Acquisition," in IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 358-365, Jan. 2012.

[8] Peterson J. Pat. 7386188 USA. Merging images to form a panoramic image / 10.06.2008.

[9] Yoo H.G., Kim K.Y., Kim H.M., Park K.S., Lee S.J., Ko K.H., Lee K.H. 2007. Color correction of high dynamic range images at HDR-level. In ACM SIGGRAPH 2007 posters (SIGGRAPH '07). ACM, New York, NY, USA, Article 67.

7. Об авторах

Свириденко Максим Игоревич, аспирант Санкт-Петербургского Государственного Политехнического Университета, Институт Металлургии, Машиностроения и Транспорта, Кафедра «Инженерная графика и дизайн».
Его e-mail twistms@mail.ru.

Восстановление двунаправленной функции рассеяния света для среды с объемным рассеянием

В.Г. Соколов[2], Д.Д.Жданов[1], И.С. Потемин[1], Б.Х.Барладян[2], Н.Н. Богданов[1], А.Д.Жданов[1]
sokolovv1969@gmail.com|ddzhdanov@mail.ru|ipotemin@yandex.ru|bbarladian@gmail.com
|nnbogdanov@corp.ifmo.ru|adzhdanov@corp.ifmo.ru
[1]Университет ИТМО, Санкт-Петербург, Россия;
[2] ИПМ им. М.В.Келдыша РАН, Москва, Россия.

Для проектирования и моделирования изображающих или неизображающих оптических систем требуется точное описание оптических свойств рассеивающих материалов. Одним из способов описания таких свойств является двунаправленная функция рассеяния света (ДФР). Не смотря на то, что получить информацию об оптических свойствах образца можно путем измерения ДФР, довольно сложно выделить свойства для отдельных компонентов образца (ДФР для поверхности образца, параметры объемного рассеяния и др.). В таких случаях необходимо восстановить эти свойства. Для такой операции существует множество методов, таких, как восстановление ДФР для микрорельефа, однако они не применимы для объемного рассеяния. В данной работе предложен метод восстановления ДФР путем оптимизации параметров объемно рассеивающей среды. Полученные результаты хорошо согласуются с результатами измерений ДФР образца.

Ключевые слова: *ДФР, измерение ДФР, восстановление ДФР, объемное рассеяние, рассеяние Ми, рассеяние Рэлея.*

Reconstruction of BSDF for the volume scattering medium

Vadim G. Sokolov[2], Dmitry D. Zhdanov[1], Igor S. Potemin[1], Boris Barladian[2], Nikolai N. Bogdanov[1], Andrey D. Zhdanov[1]
sokolovv1969@gmail.com|ddzhdanov@mail.ru|ipotemin@yandex.ru|bbarladian@gmail.com
|nnbogdanov@corp.ifmo.ru|adzhdanov@corp.ifmo.ru
[1] ITMO University, St. Petersburg, Russia;
[2] Keldysh Institute of Applied Mathematics RAS, Moscow, Russia.

Lighting design or stray light simulation of imaging or non-imaging optical systems requires a precise specification of the optical properties of the scattering materials and one of the ways of the proper specification is the Bidirectional Scattering Distribution Function (BSDF). Although, it is possible to obtain data about the optical properties of the sample for example by measurement of the BSDF, but it is difficult to extract the properties of the sample components (BSDF of the sample boundary, parameters of volume scattering, etc.). In such cases, it is required to reconstruct these properties. For this operation, there are many methods, like the reconstruction of the BSDF of the microrelief, but they not applicable in cases when the volume scattering is used. Authors have developed a method for optimization of the volume scattering media parameters, which shows good agreement with the measurев BSDF of the sample.

Keywords: *BSDF, BSDF measurements, BSDF reconstruction, volume scattering, Mie scattering, Rayleigh scattering.*

1. Введение

В настоящее время существует множество электронных приборов, в конструкции которых используются оптические системы с элементами, создающими сложное пространственное распределение интенсивности света. К таким электронным системам можно отнести, например, светопроводящие системы для жидкокристаллических экранов, автомобильные приборные панели, светодиодные светильники и др. Качество и технические характеристики устройств совершенствуются с каждым днем, в связи с чем повышаются сложность и технические требования к данным изделиям, что создает ряд проблем для разработчиков и проектировщиков таких устройств. Одной из серьезных проблем является физически корректное моделирование оптических систем со светорассеивающими элементами. Ошибки в моделировании света, приводят к некорректной оценке работы проектируемого электронного устройства в целом, что, как правило, приводит к отклонению от технических требований. Также важна скорость расчетов, так как в большинстве случаев разработчикам отведено ограниченное время, в течение которого им необходимо рассмотреть множество вариантов и провести оптимизацию всей системы. Именно поэтому скорость и корректность вычислений имеют большое значение.

В целом для корректного моделирования важно иметь как можно более точные данные об оптических свойствах рассеивающих материалов. В настоящий момент, одним из способов описания оптических свойств рассеивающих материалов является использование двунаправленной функции рассеяния света (ДФР). Эта функция определяет угловую зависимость диффузного рассеяния для поверхности. Двунаправленная функция рассеяния света имеет довольно сложное выражение. Если говорить кратко, это суперпозиция и обобщение двух функций: двунаправленной функции отражения (ДФО, eng. BRDF) и двунаправленной функции пропускания (ДФП, eng. BTDF). Ключевая особенность всех этих функций заключается в том, что они могут быть представлены в виде «черного ящика», на входе которого подается информация о падающем луче, а на выходе получаем отраженный или пропущенный луч в заданной точке поверхности.

Рассеяние света возникает, когда луч падает на шероховатую поверхность и отражается не зеркально, а диффузно, т.е. по различным направлениям. В настоящее время существует множество разнообразных диффузных рассеивающих материалов, которые включают в себя

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

шероховатую поверхность или среду с объемным рассеянием или их комбинации, представленные на Рис. 1. [3].

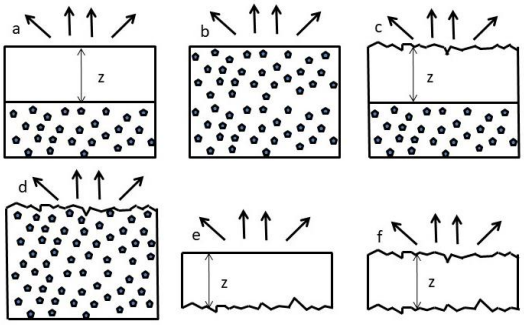


Рис. 1. Схематическое представление 6 типов диффузных образцов. Распределение света измеряется над верхними поверхностями. Области с объемным рассеянием представлены в виде точек. Области без обозначения подразумеваются как нерассеивающие. Шероховатые поверхности представлены на c – f.

В большинстве случаев, у разработчиков и проектировщиков есть образец материала, например, в форме плоскопараллельной пластины, но отсутствуют данные о характеристиках отдельных компонентов, таких как рассеивающая среда и шероховатая поверхность. В таких случаях решение задач, связанных с моделированием отдельных компонентов диффузного образца, к сожалению, становится невозможным.

Главная причина заключается в том, что чрезвычайно сложно определить свойства шероховатой поверхности и рассеивающего слоя по отдельности, т.е. провести измерения ДФР для отдельных компонентов диффузной пластины. Провести измерение ДФР возможно только для всего образца в целом. Результаты таких измерений весьма полезны для моделирования образцов, толщиной которых можно пренебречь. Однако, возникают ситуации, когда форма рассеивающего образца отличается от формы измеренного образца и, таким образом, расчетная модель образца не может быть сведена к модели поверхности. Например, модель с единственной ДФР (для всего образца) не применима, когда этот образец является элементом светопроводящей системы или объемного диффузора ввиду того, что возникают множественные переотражения либо между поверхностями образца, либо между рассеивающими частицами материала образца. По этой причине, важно определять оптические свойства поверхностей и параметров рассеивающего слоя по отдельности. Процесс определения характеристик рассеивающего слоя и параметров рассеивания граничной шероховатой поверхности является процессом восстановления полной ДФР образца.

Существует множество методов восстановления одной из частей ДФР – ДФО. Все эти методы применяются в компьютерной графике и в задачах по моделированию распространения света, однако, они не применимы для решения задач, связанных с диффузными и прозрачными материалами, при условии физически корректного моделирования. Тем не менее, существуют эффективные методы восстановления ДФР в рассеивающих материалах с шероховатой поверхностью [7, 8]. Однако, все они не способны предоставить верные результаты в случаях, когда у оптического материала есть не только одна единственная шероховатая поверхность, но и рассеивающий слой, в котором возникает объемное рассеяние. Совершенно точно можно сказать о том, что существуют эффективные методы для случаев с объемным рассеянием света, например, в работе [4]. Этот метод показывает очень хорошие результаты, но к сожалению, отсутствует детальная информация о том, как он на самом деле работает. Также, стоит отметить, что он был разработан только для программного комплекса LightTools. В представленной работе, авторы предлагают собственный метод восстановления ДФР для среды с объемным рассеянием.

2. Методы и модели

Существует несколько подходов для описания рассеяния света. В случаях, когда размер рассеивающих частиц сопоставим с длиной волны, причем ненамного больше или меньше длины волны излучения, то применяют теорию рассеяния Ми.

Другой моделью является Рэлеевское рассеяние. Используется в случаях упругого рассеяния света на частицах, размер которых намного меньше длины волны. Такими частицами могут быть отдельные атомы или молекулы. Рассеяние возникает при распространении света через прозрачную среду и лучше всего наблюдается в газах. Размер рассеивающей частицы определяется выражением (1):

$$x=\frac{2\pi r}{\lambda} \qquad (1)$$

где r – характерная длина, λ – длина волны.

Существует фазовая функция Хени - Гринштейна, широко используемая в расчетах рассеяния света. Функция удобна для некоторых численных вычислений благодаря тому, что имеет аналитическое выражение и близка к полиномам Лежандра. Более того, она может обеспечить простые вычисления приближенного распределения светового излучения [1, 4].

Функция Хени – Гринштейна представлена следующим выражением (2):

$$p_{HG}(\mu,g)=\frac{1-g^2}{(1-2g\mu+g^2)^{3/2}},\ \mu=\cos\nu \qquad (2)$$

Где ν – угол рассеяния, g – ассиметричный параметр, равный среднему значению косинуса, cos ν – полное угловое распределение, определяемое как:

$$\cos v=\frac{1}{2}\int_{-1}^{1}p(\mu)\mu d\mu \qquad (3)$$
$$\frac{1}{2}\int_{-1}^{1}p(\mu)\mu d\mu=1.$$

Таким образом, в различных ситуациях мы можем использовать соответствующий метод для расчёта объемного рассеяния. Разработанный авторами алгоритм включает все перечисленные методы, и проектировщик может выбрать один из них. В качестве характеристик и параметров, которые могут быть оптимизированы, используются размеры частиц, показатель преломления, концентрацию рассеивающих частиц при рассеянии Ми/Рэлея, или параметры фазовой функции Хени – Гринштейна.

Ранее уже описывался метод восстановления ДФР по параметрам микрорельефа [10]. Сейчас мы предлагаем комбинированный подход. Действия алгоритма включают в себя следующие операции, представленные на Рис.2.

Во-первых, необходимо получить информацию, требуемую для моделирования и оптимизации. Эта информация должна содержать данные о параметрах рельефа (распределение высот) поверхности образца, которые могут быть получены с помощью профилометра или атомно-силового микроскопа. Этот набор данных используется на следующем этапе для оптимизации ДФР образца. Обычно для измерения ДФР используется гониоспектрофотометр. Для этой задачи мы предлагаем использовать GCSM-4 [12] или Gonio-Photometer II производства “Pab Ltd” [9]. Эти измерительные системы сертифицированы, а получаемые данные достоверны.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Рис.2. Обобщенная схема восстановления параметров объемного рассеяния и ДФР оптической среды.

Следующие 4 шага отображают процесс оптимизации:

1. Подготовка модели в соответствии с измеренным образцом. Создать микрорельеф по измеренному распределению высот с помощью инструмента OPTOS MicroRelief [11, 2] используемый для моделирования распространения света. Выбрать модель объемного рассеяния (Ми, Рэлея, функцию Хени – Гринштейна) с соответствующими параметрами. Начальные параметры необходимо принять из измерений или общепринятой величины (параметры объемного рассеяния). Затем оптимизатор изменяет параметры таким образом, чтобы достичь заданной цели.

2. Компьютерное моделирование процесса измерения ДФР виртуального образца, измеренного ранее с использованием гониоспектрофотометра GCSM-4 или Gonio-Photometer II. Результатами измерений легко сравниваются друг с другом, так как алгоритм моделирования полностью соответствует процедуре измерений.

3. Сравнение результатов моделирования и измерений. Обычно результатом сравнения является значение среднеквадратического отклонения (RMS). Принимая во внимание тот факт, что результаты измерений (и соответствующего моделирования) могут содержать от сотен до тысяч значений, мы можем назначить специальный «вес» для каждого значения. Например, мы можем повлиять на результаты измерений, полученных при ортогональном освещении образца таким образом, чтобы эти данные имели большее значение по сравнению с результатами, полученными при скользящих углах.

4. Анализ результатов моделирования. Если отклонение достигает заданного уровня, то процесс оптимизации завершается и выводятся результаты моделирования. В противном случае, процесс оптимизации продолжается с новыми параметрами. Оптимизатор определяет новые параметры объемного рассеяния и параметры изменения микрорельефа (распределение высот). Учитывая то, что распределения высот микрорельефа было измерено, можно заключить, что ДФР микрорельефа изначально корректна, поэтому для оптимизационного процесса «вес» параметров объемного рассеяния назначаем намного больше «веса» параметров модификации микрорельефа. После изменения параметров процедура оптимизации проходит в цикле.

По завершению процесса оптимизации, данные приводятся к окончательному виду. Оформление результатов состоит из двух шагов. На первом этапе данные плагина OPTOS MicroRelief конвертируются в ДФР границы двух сред. Поскольку мы отделили поверхность от слоя с объемным рассеянием, то преобразование оптимизированного микрорельефа в ДФР проводится для прозрачной среды.

В заключении, проводится проверка оптимизированной модели с ДФР на поверхности образца и рассеивающим слоем. Основное отличие проверочной модели от модели, используемой в оптимизационном процессе, заключается в модели рассеивания света на поверхности. Оптимизатор использует плагин OPTOS MicroRelief, в то время как в финальной сцене используется ДФР, сконвертированная с помощью плагина OPTOS MicroRelief. Учитывая, что плагин OPTOS MicroRelief применяет лучевой подход, в то время как в процедуре конвертации может использовать как лучевой подход, так и волновое решение, то возможны отклонения между оптимизированной и финальной ДФР.

3. Эксперименты и результаты

Как было отмечено ранее, основной целью рассматриваемого процесса является восстановление свойств рассеивающей среды таким образом, чтобы они были близки к свойствам реального образца и сведены к модели бесконечно тонкого листа.

Нами были исследованы несколько проблемных образцов. Прежде всего, было необходимо продемонстрировать, что методы, используемые для оптимизации рельефа поверхности рассеивающего материала не применимы при моделировании объемного рассеяния. На рисунке 3 показаны результаты исследования случая, указанного выше. Сплошная кривая на графике соответствует измерениям, пунктирной линией обозначен результат моделирования, и как видно из графика, результаты моделирования существенно отличаются от результатов измерений. Это доказывает, что метод оптимизации микрорельефа не следует использовать в случае объемного рассеяния.

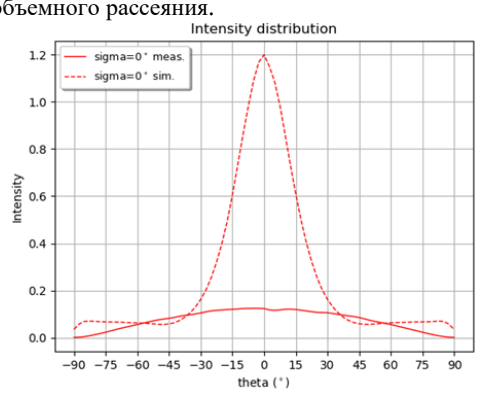


Рис. 3. Результаты восстановления ДФР, полученные на основе оптимизации микропрофиля для материалов с объемным рассеянием. Сплошная линия соответствует измеренной интенсивности, пунктирной линией обозначена интенсивность, полученная путем моделирования.

Для следующего эксперимента требуется подробное описание используемого метода. Исследуемый образец (пластина толщиной 1 мм) имеет размеры 100 x 100 мм, толщина слоя материала с объемным рассеянием, который располагался в верхней части образца, составляет 0.15 мм, показатель преломления 1.5 (Рисунок 5).

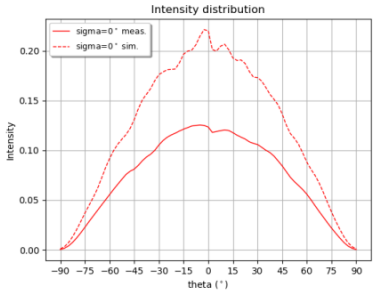
The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Также для эксперимента требуются данные ДФР исследуемого образца, и в связи с этим, необходимы ДФП и ДФО измеренные с обеих сторон исследуемой пластины. После определения указанных данных осуществляется подготовка компьютерной модели образца.

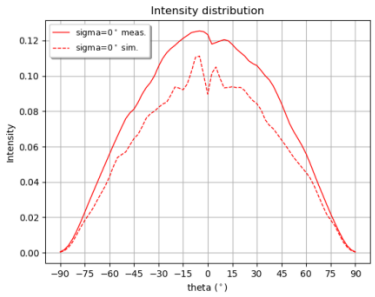
Параметры оптимизации включают в себя параметры рассеивающего компонента: объемную концентрацию частиц, радиус частиц, мнимая и действительная части показателя преломления частиц, а также параметры слоя связующей среды: размер слоя, показатель преломления. Можно выбрать подходящие параметры, а также задать диапазон от минимального до максимального значения и шаг изменения. На этом этапе, также можно отрегулировать необходимый диапазон длин волн и количество итераций.

На рисунке 4 представлен процесс оптимизации параметров объемного рассеивания.

a



b



c

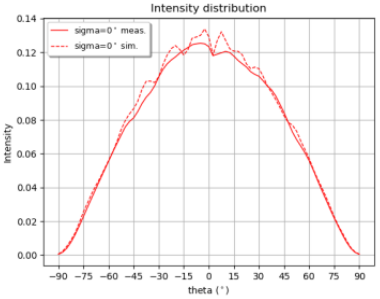


Рис. 4. Результаты ДФО восстановления, основанные на оптимизации параметров материала с объемным рассеянием.

Мы предположили, что для исследуемого образца (Рисунок 5) наиболее подходящей является модель рассеяния Ми и, соответственно, в качестве параметров оптимизации использовались параметры модели Ми. Для графика Рис. 4a исходное значение объемной концентрации частиц составляет 3%, минимальное значение – 2% и максимальное – 4 %.

Следующий параметр - радиус частиц: исходное значение составляло 500нм, минимальное значение - 100нм и максимальное – 1000нм. Кроме того, мы попытались оптимизировать показатель преломления образца: при его исходном значении 1.7 минимальное значение составило 1.5, а максимальное – 2.0.

Последний параметр оптимизации – размер слоя. Его исходная величина составляла 0.2 мм, минимальное значение 0.1 мм и максимальное – 0.3 мм. Стандартное отклонение для первой итерации составляло 37.4 %. На пятнадцатом шаге оптимизации (Рисунок 4b) параметры изменились следующим образом: значение концентрации частиц возросло до 4.09%, радиус частиц - 521.07 нм, показатель преломления увеличился до 1.9, а размер слоя – до 0.203 мм.

Среднеквадратическое отклонение для пятнадцатой итерации составило 12.7 %. Фактически полный процесс оптимизации составил тридцать четыре итерации, однако наилучший результат оптимизации был достигнут на двадцать девятом шаге (Рисунок 4c). На данном этапе оптимизации наблюдались следующие параметры: концентрация частиц составила 3.17 %, радиус частиц – 547.93 нм, показатель преломления – 1.93, и толщина слоя – 0.18 мм. Среднеквадратическое отклонение для двадцать девятой итерации составило 4.5 %.

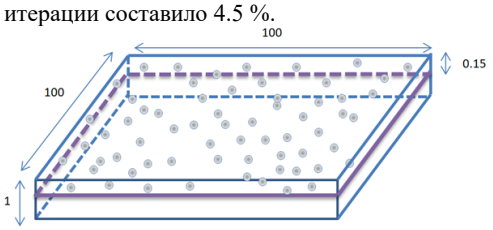


Рис. 5. Схема модели исследуемого образца

Этот пример демонстрирует хорошую эффективность предлагаемого метода, поэтому мы применили его для более сложной ситуации, которая тесно связана с реальной ДФР.

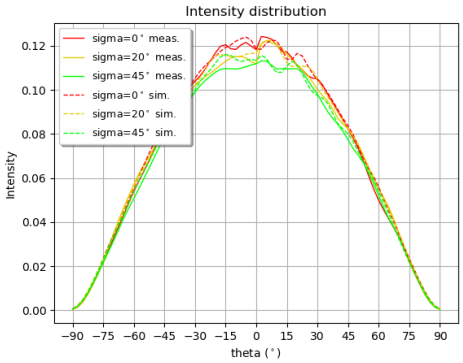


Рис. 6. Результаты восстановления ДФР, основанное на оптимизации параметров материала с объемным рассеиванием для конкретных углов луча: 0◦, 20◦, 45◦.

На Рисунке 6 приведены результаты для нескольких углов падения. Процесс оптимизации состоял из тридцати пяти итераций, а наилучший результат был достигнут при значении стандартного отклонения 3.5 %. Процесс оптимизации достаточно быстрый, время прохождения всех шагов - 7 минут.

4. Заключение

Предложенный метод восстановления оптических характеристик образцов материалов с объемным рассеиванием и шероховатыми границами хорошо согласуются с результатами измерений ДФР соответствующих образцов материала.

Применение данного метода обеспечивает быстрое и физически корректное восстановление отдельных оптических параметров (ДФР граничной поверхности и характеристик среды объемного рассеивания) для образцов со сложными оптическими свойствами.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

5. Благодарности

Работа выполнена при частичной государственной финансовой поддержке грантами РФФИ № 17-01-00363, № 16-01-00552 и № 18-01-00569, а также компанией Integra Inc.

6. Литература

[1] Calabro, K.W. and Cassarly, W., Modeling scattering in turbid media using the Gegenbauer phase function // Proc. SPIE 9333, Biomedical Applications of Light Scattering IX, 93330F (2015). doi: 10.1117/12.2079695

[2] Ershov, S.V. and Zhdanov, D.D., Efficient application of Optical Objects in light simulation software // Proc. 15th International Conference on Computer Graphics and Applications - GraphiCon-2005, Russia, Novosibirsk, 2005, p. 284-288.

[3] Germer, T.A., Bidirectional scattering distribution function measurements from volume diffusers: correction factors and associated uncertainties // Applied optics. – 2016. Vol.55, No. 25 – P. 75-85. http://dx.doi.org/10.1364/AO.55.006978

[4] Kuperman-Le Bihan, Q., Volume scattering characterization for illumination design // Proc. SPIE 8550, Optical Systems Design – 2012. 85502K; doi: 10.1117/12.981995

[5] Magarill, S., Cassarly, W.J.,. Jenkins, D.R. and others. Parametric Model of Volumetric Scattering // Design and Fabrication Congress 2017 (IODC, Freeform, OFT), OSA – 2017.

[6] Официальный сайт компании Muracami Color Research Laboratory. URL: http://www.mcrl.co.jp/english/index.html (дата обращения 14.05.2018)

[7] Nielsen, J.B., Jensen, H.W. and Ramamoorthi, R., On Optimal, Minimal BRDF Sampling for Reflectance Acquisition // ACM TOG. – 2015. № 34(6). – P.1-11.

[8] Nurcan, S., Serkan, E. and Aydın, Ö., BRDF Reconstruction Using Compressive Sensing // 21st International Conference on Computer Graphics, Visualization and Computer Vision. – 2013, – P. 88-94.

[9] Официальный сайт компании Pab advanced technologies Ltd, URL: http://www.pab.eu/?d=gonio-photometer (дата обращения 14.05.2018)

[10] Sokolov, V.G., Zhdanov, D.D., Potemin, I.S., Bogdanov, N.N., Zhdanov, A.D. and Denisov, E.U., Reconstruction of BSDF based on optimization of Micro-Relief normal distribution // Proc. of 27th International Conference on Computer Graphics and Vision. – 2017. –P. 37-41.

[11] Sokolov, V.G., Zhdanov, D.D., Potemin, I.S., Garbul, A.A., Voloboy, A.G., Galaktionov, V.A. and Kirilov. N., Reconstruction of scattering properties of rough air-dielectric boundary // Optical Review. – 2016. Vol.23(5), – P. 834-841. doi: 10.1007/s10043-016-0250-6.

[12] Zhdanov, D.D., Garbul, A.A., Potemin, I.S., Voloboy, A.G., Galaktionov, V.A., Ershov, S.V. and Sokolov, V.G., Photorealistic Volume Scattering Model in the Bidirectional Stochastic Ray Tracing Problem // Programming and Computer Software. – 2015. Vol. 41, No. 5. – P. 295–301. doi: 10.1134/S0361768815050102.

[13] Zhdanov, D.D., Potemin, I.S., Sokolov, V.G., Garbul, A.A., Voloboy, A.G. and Galaktionov, V.A., Simulation of the BSDF measurement capabilities for various materials with GCMS-4 goniospectrophotometer // Proceedings of 12th International Conference "Applied Optics 2016". – 2016. Vol. 2. –P. 219-223.

Об авторах

Соколов Вадим Геннадиевич, научный сотрудник информационно вычислительного центра ИПМ.

Его e-mail sokolovv1969@gmail.com.

Жданов Дмитрий Дмитриевич, к.ф.-м.н., заведующий кафедрой технологий визуализации университета ИТМО. Его e-mail ddzhdanov@mail.ru.

Потемин Игорь Станиславович, к.т.н., доцент кафедры технологий визуализации университета ИТМО. Его e-mail ipotemin@yandex.ru.

Борис Хаимович Барладян, к.ф.-м.н., старший научный сотрудник отдела компьютерной графики и вычислительной оптики ИПМ им.М.В.Келдыша РАН. Его e-mail bbarladian@gmail.com.

Богданов Николай Николаевич, аспирант кафедры технологий визуализации университета ИТМО. Его e-mail nnbogdanov@corp.ifmo.ru.

Жданов Андрей Дмитриевич, аспирант кафедры технологий визуализации университета ИТМО. Его e-mail adzhdanov@corp.ifmo.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Использование свёрточных нейронных сетей для классификации видов искусственного освещения

М.И. Сорокин[1], Д.Д. Жданов[1]
vergotten@gmail.com|ddzhdanov@mail.ru
[1]Университет ИТМО, Санкт-Петербург, Россия.

Актуальность данной темы обусловлена бурным развитием систем виртуальной, смешанной и дополненной реальностей. Проблема заключается в формировании естественных условий освещения объектов виртуального мира в реальном пространстве. Реально наблюдаемая картина всегда содержит источники освещения, и когда в реальный мир добавляется виртуальный объект, он должен быть освещен естественным образом, т.е. со стороны источников света и отбрасывать тень в противоположную сторону, другими словами соответствовать освещению реального мира. Задача заключается в анализе и классификации источников освещения для дальнейшего построения соответствующих оптических параметров среды с использованием технологии свёрточных нейронных сетей, которые себя хорошо зарекомендовали в области классификации. В данной работе предложен подход использования свёрточных нейронных сетей для классификации типа освещения с использованием библиотеки машинного обучения Keras.

Ключевые слова: *классификация, освещение, свёрточные нейронные сети.*

Using convolutional neural networks to classify types of artificial lighting

M.I. Sorokin[1], D.D. Zhdanov[1]
vergotten@gmail.com|ddzhdanov@mail.ru
1ITMO University, St. Petersburg, Russia.

The relevance of this topic is due to the rapid development of virtual, mixed and augmented reality systems. The problem lies in the formation of natural conditions for lighting objects of the virtual world in real space. The actual picture, that is observed always contains lighting sources, and when a virtual object is added to the real world, it should be illuminated in a natural way, i.e. from the side of light sources and cast the shadow in the opposite direction, in other words, correspond to the illumination of the real world. The task is to analyze and classify the light sources for further construction of the corresponding optical parameters of the medium using the technology of convolutional neural networks that have already proved themselves in the field of classification. In this paper, an approach is proposed for the use of convolutional neural networks to classify the type of illumination using the Keras machine learning library.

Keywords: *classification, illumination, convolutional neural networks.*

1. Введение

Главное отличие систем дополненной реальности от систем виртуальной – это то, что в системах дополненной реальности слои изображения не просто симулируются, а накладываются на реальный мир, и, следовательно, основная проблема адаптации синтезированных объектов – освещение.

Освещение играет очень важную роль в представлении систем дополненной реальности, ведь неправильное освещение негативно влияет на восприятие, и неправильно освещенный объект может ввести наблюдателя в заблуждение, например, из-за отсутствия теней, ориентированных естественным образом.

Проблема заключается в дискомфорте восприятия синтезированных изображений виртуальной реальности с реальными объектами. Человеческий мозг находит несоответствие с тем, что он видит, и тем, что он ожидает увидеть. Например, некорректно воспринимаются тени от объектов: они излишне контрастны (без полутонов) или имеют неверную ориентацию. В работе [1] предложен метод вставки синтезированных изображений в уже существующие фотографии путем не только анализа освещения, но и геометрии помещения.

Как известно, освещение бывает двух видов: естественное и искусственное. Искусственное же в свою очередь можно разделить по его расположению на следующие виды: боковое, потолочное, напольное и комбинированное. Каждый вид имеет свои оптические параметры и свойства, такие как угол падения и интенсивность, что является ключевым моментом в образовании теней. Поэтому очень важно правильно определить тип освещения.

Актуальность анализа освещения определяется бурным развитием систем виртуальной, дополненной и смешанной реальностей. В данной работе предлагается подход классификации видов искусственного освещения с использованием свёрточных нейронных сетей.

Стоит заметить, что впервые свёрточная нейронная сеть была представлена Алексом Крижевски [2] на соревнованиях по распознаванию изображений ImageNet в 2012 году, где с огромным отрывом опередила своих конкурентов. В дальнейшем, свёрточная нейронная сеть уже использовалась повсеместно, лишь меняя свою архитектуру для достижения более лучших результатов.

2. Состояние проблемы

На данный момент свёрточные нейронные сети активно используются для различных задач, связанных непосредственно с анализом и обработкой изображений, будь то классификация или распознавание каких-либо отдельных участков. К основным преимуществам свёрточных нейронных сетей (CNN) следует отнести удобное распараллеливание вычислений, устойчивость к сдвигу изображения и обучению при помощи

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

классическому методу обратного распространения ошибки. Из недостатков - большое количество настраиваемых параметров, т.е. для решения какой-либо задачи следует испробовать различные слои и параметры для выбора лучшего решения, к таким параметрам относятся: размерность ядра свёртки, степень уменьшения размерности, использование слоёв субдискретизации, выбор функции активации и т.д. Более подробная информация представлена в официальной документации [3].

Естественно, что свёрточная нейронная сеть хорошо подходит для сегментации изображений. Например, в работе [4] используется свёрточная нейронная сеть для сегментации биомедицинских изображений.

В работе [5] используются свёрточные нейронные сети для оценки состояния окружения на открытом воздухе. Обучаясь на панорамных изображениях и анализируя такие параметры, как природные условия, расположение и яркость солнца и неба, сеть строит предсказание, как должна располагаться тень объекта при заданных условиях.

Аналогичным образом в работе [6] также идет анализ объектов на открытом воздухе. Здесь сеть не только анализирует положение неба, но и делит изображение пополам и анализирует нижнюю часть на предмет теней, анализируя которые, обучается и строит свои собственные предсказания.

В работе [7] происходит анализ небольшого участка помещения, по ко которому может быть корректно сформировано освещение всей видимой части сцены. В данном алгоритме сеть на вход получает изображение, по которому строится маска освещения, и на выходе получается панорама изображения с прогнозируемым освещением.

3. Архитектура CNN

В качестве архитектуры свёрточной нейронной сети в данной работе было решено взять архитектуру VGG 16 NET, которая успешно использовалась в работах [8,9], состоящая из трех слоев свертки размерности 3x3 и размером ядер подвыборки 2x2. Эти слои необходимы для составления карт признаков. После свёрточных слоев следуют полносвязные слои, которые используются для классификации. Далее идет слой исключения (dropout) с параметром 0.5, который необходим для предотвращения обучения, путем исключения из работы определенного процента нейронов. Так как в данной работе классифицируются изображения на два типа освещения, потолочное и боковое (wall, ceiling), то используется "бинарная функция ошибки" (binary crossentropy), оптимизатор - "адам", выходная функция активации "sigmoid" - для плавного изменения значений от 0 до 1, а в качестве метрики обучения - "аккуратность" (accuracy).

В качестве функции активации между слоями выбрана функция "RelU", которая хорошо себя зарекомендовала в работе [10]. Функцию RelU можно представить в виде формулы:

$$f(x) = max(0, x),$$

которая реализует простой пороговый переход в нуле.

Слои дискретизации работают следующим образом: на его вход подается некий массив с информацией, и нужно процедить это информацию, т.е. уменьшить размерность. На рисунке 1 изображен пример дискретизации.

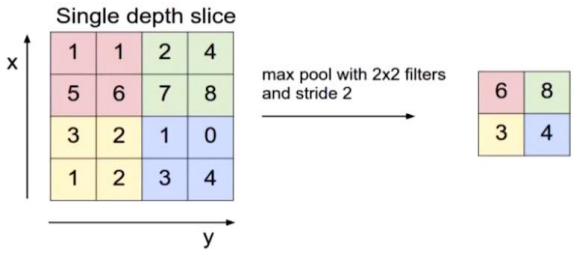


Рис. 1. Операция субдискретизации.

Слои дискретизации необходимы, т.к. использование этих слоёв имеет свои плюсы. Во-первых, уменьшается размерность, а чем меньше размерность, тем больше скорость обучения. Во-вторых, побеждают наиболее активные нейроны. И в-третьих, вырабатывается инвариантность к небольшим сдвигам.

4. Обучение

Так как обучение происходит "с учителем", необходимо для начала сформировать набор изображений каждого вида освещения, в данном случае нейронная сеть обучена определять два вида освещения - потолочное (ceiling) и боковое (wall). Всего набор данных состоит из 1560 изображений каждого класса размерности 200 на 300 пикселей, из которых 15% выделены на тестовый набор изображений и 15% на проверочный набор. Итого, каждый класс состоит из 1092 тренировочных изображений и по 234 тестовых и проверочных. Структура каталогов представлена на рисунке 2, где "train" - набор данных для тренировки, "validation" - для проверки, и "test", соответственно набор данных для тестирования. При составлении набора данных, изображения были зеркально отражены и немного наклонены в разные стороны, что необходимо для увеличения размерности набора данных и выработки инвариантности к наклону.

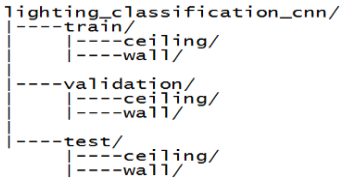


Рис. 2. Структура каталогов наборов данных

Изображения из набора данных представлены на рисунке 3.

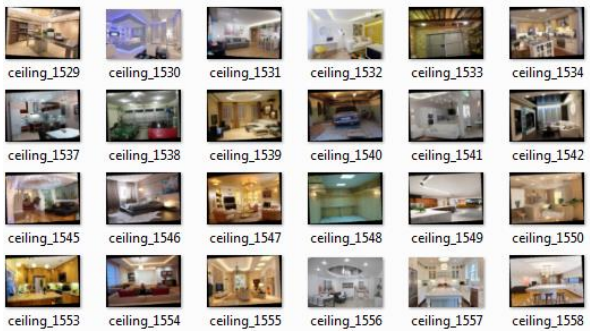


Рис. 3. Набор данных потолочного освещения

Обучение нейронной сети происходило при помощи библиотеки машинного обучения "keras", в качестве

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

бэкенда использовася "tensorflow". При обучении использовался метод "model.fit_generator", которому передавалось на вход два параметра "train_generator" - набор данных с изображениями для обучения и "val_generator" - набор данных для проверки. Обучение выполнялось в течении 30 эпох. На рисунке 4 изображена история обучения, где "acc" - точность на обучающем наборе данных, а "val_acc" - точность на проверочном наборе данных.

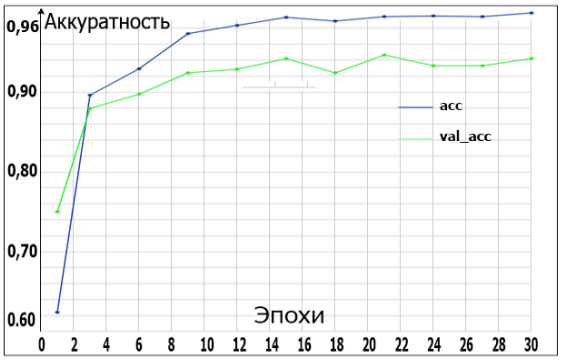


Рис.4. История обучения нейронной сети по эпохам

Как видно из графика, составленного по значениям аккуратности после каждой итерации эпохи, аккуратность сети на тестовых данных равняется 96.43%. Об этом так же говорит функция model.evaluate_generator, представленная на рисунке 5.

```
In [14]: scores = model.evaluate_generator(test_generator,

In [15]: print("Аккуратность на тестовых данных: %.2f%%" %

         Аккуратность на тестовых данных: 96.43%
```

Рис.5. Функция оценки аккуратности нейронной сети

5. Пример работы

После того, как сеть обучилась, можно для проверки ее работоспособности запустить на обработку изображение, которое не участвовало в обучении и тестировании. На рисунке 6 изображен процесс подачи изображения на вход нейронной сети, а на рисунке 7 его классификация к классу wall (боковое).

```
In [32]: img = image.load_img('3.jpg', target_size=(
         plt.imshow(img)
         plt.show()
```

Рис. 6. Подача изображения нейронной сети

```
In [31]: print(prediction)
         print(classes[np.argmax(prediction)])

         [[0.]]
         wall
```

Рис 7. Классификация изображения к классу wall

6. Заключение

В последнее время, свёрточные нейронные сети набирают огромную популярность в решении различных задач, таких как поиск, классификация или обработка изображений. Задачи по анализу освещения не являются исключением. Свёрточные нейронные сети отлично себя показали в этих и смежных задачах. С каждым годом разрабатываются новые техники и рекомендации по более эффективному обучению, появляются новые архитектуры и новые наборы данных, что позволяет надеяться на возможность использования нейронных сетей не только для качественного, но и количественного анализа освещения.

В данной работе рассмотрен подход классификации видов освещения с помощью свёрточных нейронных сетей, который может использоваться для определения типа освещения с дальнейшей его обработкой и восстановления оптических параметров сцены. Хотя нейронная сеть в большинстве случаев работает корректно, однако еще многое предстоит усовершенствовать, так, например, возможно улучшить алгоритм определения источников освещения или разработать алгоритм для определения диффузных поверхностей. Помимо этого, следует дополнить набор данных тестовых изображений и изображений для тестирования.

7. Литература

[1] Kevin Karsch, Varsha Hedau, David Forsyth, Derek Hoiem, Rendering synthetic objects into legacy photographs // Proceedings of the 2011 SIGGRAPH Asia Conference Article No. 157

[2] Alex Krizhevsky, Geoffrey E. Hinton, Ilya Sutskever, ImageNet Classification with Deep Convolutional Neural Networks // Advances in neural information processing systems 2012, 1097-1105.

[3] Convolutional Neural Networks (LeNet) - DeepLearning 0.1 documentation. DeepLearning 0.1. LISA Lab.

[4] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pages 234–241, 2015.

[5] Y. Hold-Geoffroy, K. Sunkavalli, S. Hadap, E. Gambaretto, and J.-F. Lalonde. Deep outdoor illumination estimation. In Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[6] J.-F. Lalonde, A. A. Efros, and S. G. Narasimhan. Estimating the natural illumination conditions from a single outdoor image. International Journal of Computer Vision, 98(2): 123–145, 2012.

[7] M.-A. Gardner, K. Sunkavalli, E. Yumer, X. Shen, E. Gambaretto, C. Gagné, and J.-F. Lalonde. Learning to predict indoor illumination from a single image. ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia), preprints, 2017.

[8] Song Han, Huizi Mao, William J. Dally, Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Coding // Computer Vision and Pattern Recognition (cs.CV); Neural and Evolutionary Computing (cs.NE)
[9] Jonathan Long, Evan Shelhamer, Trevor Darrell, Fully Convolutional Networks for Semantic Segmentation // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431-3440
[10] GE Hinton, V Nair, Rectified linear units improve restricted boltzmann machines // In Proceedings of the 27th International Conference on Machine Learning. (ICML-10), 807–814.

Об авторах

Сорокин Максим Игоревич, аспирант кафедры визуализации (базовой) факультета программной инженерии и компьютерной техники Университета ИТМО. Его e-mail vergotten@gmail.com.

Жданов Дмитрий Дмитриевич, к.ф.-м.н., заведующий кафедрой технологий визуализации Университета ИТМО. Его e-mail ddzhdanov@mail.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Научная визуализация и визуальная аналитика

Scientific Visualization and Visual Analytics

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Edge detection using neural network committee

Diego Oliva[1], Demin Anton[2], Demeshko Maria[2]
diego.oliva@cucei.udg.mx | ad@tpu.ru | mvd7@tpu.ru
[1]Depto. de Ciencias Computacionales, Universidad de Guadalajara, CUCEI, Av. Revolucion 1500, Guadalajara, Jal, Mexico
[2]National Research Tomsk Polytechnic University, pr. Lenina 30, Tomsk, Russia

This article considers edge detection using neural network approach, and describes the integration of neural networks into a committee. The novelty of the proposed technique comes from two ideas: firstly, since the training set describes a solution, the proposed algorithm is able to detect multidirectional edges on noised image. Secondly, to increase the accuracy of detection, neural networks were combined into a committee.

Keywords: *edge detection, neural network, neural network committee, committee classification, digital image.*

1. INTRODUCTION

The necessity to detect edges is one of the crucial task in computer vision and image processing systems. Commonly it is used for applications such as segmentation or object recognition [1]. Edges are features that define the area of an object in digital images. Basically, an edge is a boundary between two homogeneous regions [2]. The gray level properties of the two regions on either side of an edge are distinct and exhibit some local uniformity or homogeneity among themselves.

There are different approaches used for edge detection and it is important to emphasize that each one of them can be useful in specific cases defined by the scene. This situation exists since edges can be defined in different ways for different situations. One of the definitions, which have many applications, is the local variation of the brightness in the image [3].

On the other hand, several classic methods of edge detection use the gradient evaluation. One of the first methods was proposed by L. Roberts [5], it is based on cross matrix operator, which contains differences between neighbor elements. In this sense, J. Prewitt introduced an operator based on central difference [5], that is well-known and extensively used for edge detection. Another interesting technique to detect edges in digital images is the discrete Laplace operator, its kernel is evaluated with discrete partial derivatives [6].

Some other approaches consider filters with weighted kernels used for decreasing the noise [5]. An example of this kind of technique was developed by J. Canny [4]. It is an edge detection method which includes gradient evaluation but in addition it uses a preprocessing step (blurring) and non-maximum suppression. Morphological operations can also be applied for the described problem [7]. This approach employs the difference that exists between the original image and its erosion output. The main disadvantage of this method is that it can be applied only for binary images. Meanwhile, several studies demonstrate that the usage of wavelet transform for edge detection is an interesting tool and provide competitive results [8, 9].

All the methods previously described represent interesting alternatives for edge detection. However, the main problems of all of them are the computational effort required and the lack of accuracy presented on the detection [10]. This article considers edge detection using neural network approach, and describes the integration of neural networks into a committee. The novelty of the proposed technique comes from two ideas: firstly, since the training set describes a solution, the proposed algorithm is able to detect multidirectional edges on noised image. Secondly, to increase the accuracy of detection, neural networks were combined into a committee.

The remainder paper is organized as follows: Section 2 presents the architecture of the Neural Network. In Section 3 the training process is explained. Section 4 explains the bagging procedure used for networks committee. In Section 5 the process to form the inputs of the Neural Network is presented. The experimental results are presented in Section 6. Meanwhile, Section 7 shows a comparative study of the proposed approach. Finally, in Section 8 some conclusions are discussed.

2. NEURAL NETWORK STRUCTURE

In this article multilayer feedforward neural networks are used. In this kind of networks the information moves only in one direction from the input nodes through the hidden nodes and to the output nodes. Signal transmission occurs as follows: the sum of the products of the weights and the inputs is calculated in each node. Values calculated in output neurons become network outputs and can represent significant information about input vector.

In order to choose optimal network configuration several test networks were trained. These networks had different configurations:

1. 25 input neurons, 10 hidden neurons, 1 output neuron;
2. 25 input neurons, 25 hidden neurons, 1 output neuron;
3. 25 input neurons, 30 hidden neurons, 1 output neuron;
4. 25 input neurons, 50 hidden neurons, 1 output neuron.

Number of inputs is 25 for all of the experiments because of chosen neighborhood. Every input vector represents the set of values corresponding to pixels from neighborhood with size 5×5. Way of formation of these values is described in Section 5.

Experiments with these networks included training with the study images described in Section 3 and detecting edges for one image. Results were analyzed visually. Network with 30 hidden neurons detected edges better than others.

Selected structure is shown in the Fig. 1.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

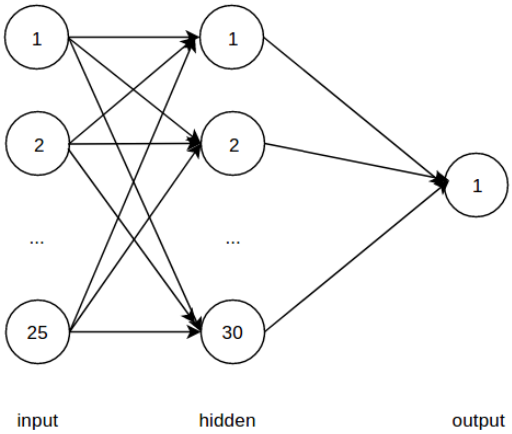


Figure 1. Neural network structure.

In accordance with experiment neural network includes 25 input neurons, 30 hidden neurons and 1 output neuron. The value of output neuron will be set to zero if current pixel belongs to background or object, and will be set to one if current pixel belongs to the edge.

Described detection system includes three networks with such structure.

3. NETWORK TRAINING

In order to properly train the neural network, a dataset is generated that contains different samples of edges. The images on the training set were prepared in such a way that they have edges of different “steepness”. Also, they are contaminated with noise to achieve appropriate noise resistance. All these operations were performed with the help of GIMP editor.

The samples included in the training dataset are presented in Fig. 2–4.

Figure 2. Couple of study images: original and marked (horizontal edges).

Figure 3. Couple of study images: original and marked (vertical edges).

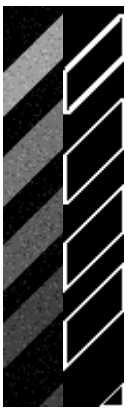


Figure 4. Couple of study images: original and marked (diagonal edges).

In Fig. 2–4 the marked images were prepared with the help of the Canny detector and after that were modified manually to achieve necessary accuracy. In this way, each image is a multiple testing set: the sets formed from the images in Fig. 2–3 both contains 1500 sub-images, the set formed from the images in Fig. 4 contains 6000 sub-images. Moreover, some of these sets contain edges and some do not: training with these images takes into account positive and negative examples.

4. BAGGING

Bagging is one of the committee classification methods. These kinds of methods are used in cases when it is necessary to define to which class the object belongs. In terms of edge detection committee classification approaches can determine whether current pixel belongs to the edge or to the object and background. The application of committee techniques theoretically is not worse than methods which use one classifier [11]. This rule is often observed in practice, but there are cases when the committee classification works worse than one classifier. Therefore, the usefulness of the committee classification for the solution of a particular problem is determined experimentally.

Bagging is one of the fundamental committee classification algorithms, in which the decision is made on the basis of averaging decisions of separate classifiers. Described solution uses three classifiers trained on various images (explained in the previous section). Their results are summarized and multiplied by empirical coefficient. If result is less than fixed threshold it means that current pixel belongs to background or object. Otherwise, if result is greater than threshold, current pixel belongs to the edge.

5. FORMATION OF INPUTS

In order to properly identify the optimal way of forming input vector, several experiments were conducted. Each experiment consists of using particular way of forming inputs and submitting this data to neural network for training with the back-propagation algorithm. The resulting network was used for edge detection on different images and results of the detection were visually analyzed. The neighborhood had the same size (5×5) for all experiments. In first experiment input vector corresponded to the normalized values from the neighborhood of the point. This experiment gave torn thin edges which lead to three other experiments. Assuming that using differences will give better result, those experiments were conducted with usage of three types of differences: difference between central element and current element, difference between current element and its row neighbor, difference between current element and its column neighbor. Such differences are explained in Fig. 5–7.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

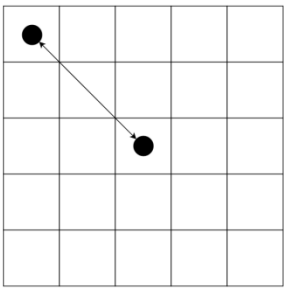


Figure 5. Difference between central and current element.

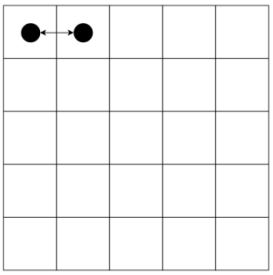


Figure 6. Difference between current element and its row neighbor.

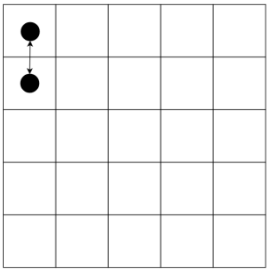


Figure 7. Difference between current element and its column neighbor.

It is interesting that using difference with column neighbor gave no edges. Using difference with central element gave best results (this method is in a sense similar to filter weighing), so it was the way of formatting inputs that was chosen.

6. RESULTS

This section presents some interesting results selected from the experiments. Results of edges extraction using the proposed method after the training and testing steps are presented in Table 1.

First row in Table 1 shows edge detection on the image of Rubik's Cube. Neural network detected most of the squares on the sides of the cube but did not detect blue ones. Presumably it happened because of less abrupt brightness transition.

Second row in Table 1 shows edge detection on the photo of pawn. It can be noticed that right corner of the pawn is not detected because of lightning, but overall outline of the chess figure is obvious.

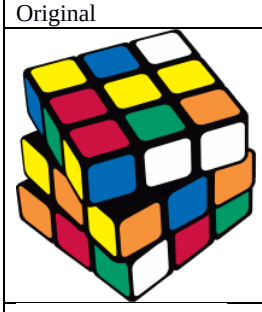
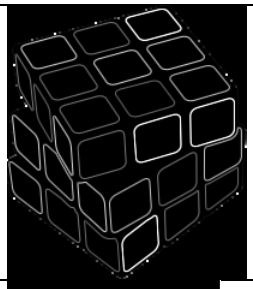
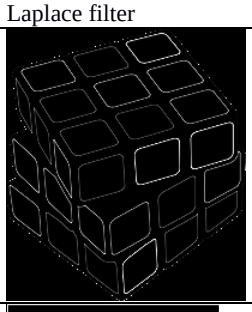
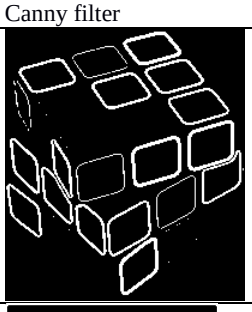
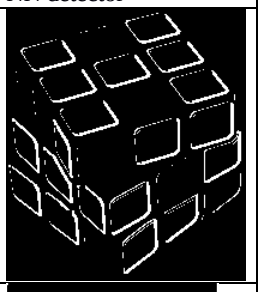
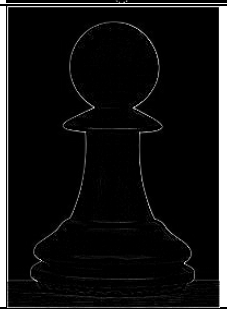
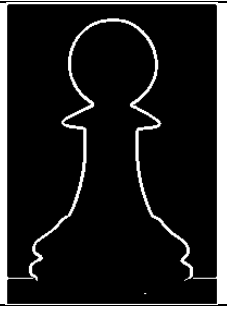
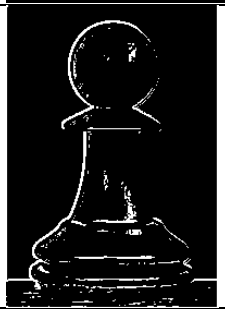
Last row in Table 1 shows result of edge detection on the image of sea wave. It is evident that detector identified many details and found direction of the waves.

It is obvious that the classifier detects edges of the objects. Sometimes it reacts to noise more than necessary (image with the pawn), but in other cases it can be useful (image with the wave).

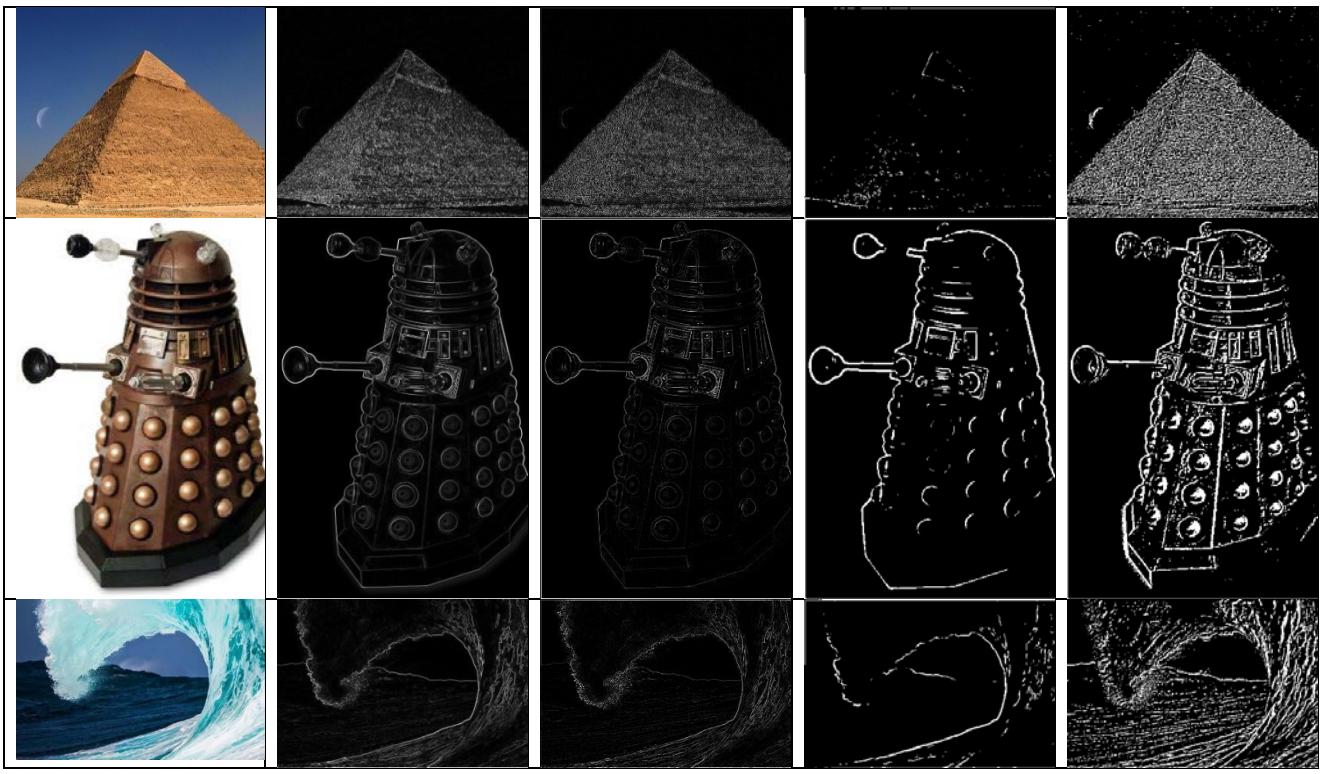
7. COMPARISON WITH OTHER METHODS

For a comparative study the Table 1 shows the results of edge detection considering different methods (explained in Section 1): the Roberts filter, the Laplace filter and the Canny filter.

Table 1. Comparison with other edge detection methods

Original	Roberts filter	Laplace filter	Canny filter	NN detector
				
				

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18



From Table 1 it can be noted that developed classifier is comparable with Canny detector in terms of thickness and accuracy. Also the last example shows that developed detector is sensitive to directed lines. This feature can be useful in detecting dynamic changes in image series. In this context, the image with Rubik's cube shows that sensitivity of developed detector is higher than Canny's but lower than Roberts' and Laplace's. On the other hand intensity of edges detected with described committee is higher than intensity of edges given by Roberts and Laplace filters.

8. CONCLUSION

Proposed classifier can be used as a preprocessing step in image segmentation and objects recognition algorithms. It detects edges which are as wide as edges detected by Canny filter. In cases where it is necessary to detect thin boundaries skeletonization methods can be applied. The application of the neural network committee allows usage of weak classifiers. Moreover, it is possible to replace sub-classifier for cases when it is necessary to detect edges with different characteristics (sub-classifier will be learned on different samples). Also edge detection can be applied to scientific visualization as a step of visualization quality assessment. Opportunity to replace sub-classifiers allows forming different requirements to the result. For example, in case of visualization of air flows the requirement for edge clarity between flows with different temperature can be brought. Committee composed of classifiers trained to detect edges of such clarity can find them and give the results to sub-system which makes the decision about the correspondence of the expected and actual result.

On the one hand, disadvantage of this classifier is its noise sensitivity, which is higher than Canny detector's sensitivity. But on the other hand, it can be useful in solving certain problems. For example, third image in Table 1 lost all features after Canny filtration, but described classifier detected all of the local changes in pyramid texture. It means that it is necessary to select edge detection tool in accordance with the task being solved.

The work was carried out within the TPU Competitiveness Enhancement Program.

9. REFERENCES

1. Belim S.V., Kutlunin P.E. Boundary extraction from images using clustering algorithm // Computer Optics. 2015. 39 (1). P. 119–124.
2. R.C. Gonzalez, R.E. Woods, Digital Image Processing, Pearson, Prentice-Hall, New Jersey, 1992.
3. Demin A.Yu., Dorofeev V.A. Parallelization research of Algorithm for detecting borders on the basis of graph representation // 12th International conference on actual problems of electronic instrument engineering (APEIE), 2014. // Novosibirsk, October 2-4 , 2014. Vol. 6. P. 232–236.
4. John Canny. A computational approach to edge detection // Pattern Analysis and Machine Intelligence, IEEE Transactions on, PAMI-8(6). P. 679–698.
5. Fisenko V.T. Computer processing and image recognition // ITMO University, 2008. P. 192.
6. Bui T.T.T., Spitsyn V.G. Analysis of methods of digital images edge detection // Reports of Tomsk State University of Control Systems and Radioelectronics, (2-2 (22)). P. 221–223.
7. Shapiro L., Sockman G. Computer vision // BKL Publishers, 2006. P. 752.
8. Antoshuk S.G. The edge selection of object images by double hyperbolic wavelet transform method // Electrical Machine-building and Electrical Equipment (65), 2005. P. 65–69.
9. Zhiznjakov A.L. Detection and analysis of shapes attributes of halftone images using multiresoluation representation // Methods and devices of information transmission and processing (7), 2006. P. 143–148
10. O.P. Verma, R. Sharma, M. Kumar, N. Agrawal, An Optimal Edge Detection Using Gravitational Search Algorithm, Lect. Notes Softw. 1, 2013. P. 148–152.
11. Bishop C. Pattern recognition and machine learning // Springer, 2006.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Using IoT Devices Powered by Scientific Visualization Tools to Create Interactive Paleontological Museum Exhibitions

K. Ryabinin[1], M. Kolesnik[2]
kostya.ryabinin@gmail.com|kolesnik.ma@outlook.com
[1]Perm State University, Perm, Russia;
[2]Perm Regional Museum / branch Museum of Permian Antiquities, Perm, Russia

In this paper, we propose to create interactive paleontological exhibitions using scientific visualization software tools and modern IoT capabilities. We implement this concept with a help of the multiplatform adaptive scientific visualization system SciVi. This systems' distinctive feature is knowledge-driven functioning. It enables extending SciVi capabilities by replenishing underlying knowledge base that contains ontologies describing supported visual objects, graphical scenes, data transformers and filters, etc. Thereby it becomes easy to add to SciVi new rendering algorithms needed for particular exhibitions. Another important SciVi feature is the built-in firmware generator that automates the installation of this system on the IoT electronic gadgets. SciVi provides the user with high-level data flow diagram editor that enables composing data transformation and rendering pipelines by chaining supported data filters, visual objects and other elements described in the knowledge base. This editor turns SciVi into high-level self-service visual programming tool for IoT devices that is easy enough to be used by non-professional or inexperienced programmers, for example by museum research workers. We demonstrate the proposed approach by an example of interactive Dimetrodon grandis exhibit that includes scientifically accurate animated 3D reconstruction of dimetrodon based on the latest paleontological research. The movements of the reconstructed model are controlled by an IoT-based light direction sensor that makes dimetrodon to rotate its sail against the incoming light. The rendering is performed using Peppers' Ghost technique to reach volumetric illusion. The exhibit is created in the Museum of Permian Antiquities making the first step towards the smart paleontological museum in Perm.

Keywords: *Scientific Visualization, IoT, Smart Museum, Paleontology, Dimetrodon Reconstruction.*

1. Introduction

The reconstruction of prehistoric life forms is a complicated problem in paleontology. When it comes to the creatures lived millions of years ago, fossils and prints are almost the only evidence of their existence. These data are very uncertain and ambiguous, so it seems to be almost impossible to uniquely reconstruct outfit and behavior of extinct animals. But the new discoveries in biology and paleontology enable the look at the prehistoric fauna to evolve. For example, since the 1960s the new interpretations of paleontological facts and fossils built up the completely new point of view at the dinosaurs. This scientific revolution is called "dinosaur renaissance" [1]. It is already proven, that some dinosaurs were feather-covered, warm-blooded and had complicated behavior.

However, the media image of prehistoric life is still based on outdated concepts and hypotheses. The modern films, cartoons, and games often present prehistoric animals as bloodthirsty monsters, indicating the ancient times as eon of harmless predators and doomed victims. Such dramatization has desired artistic effect but forms a kind of "media trap" for the casual audience. People who are far from biology often imagine the prehistoric world as some fabulous place and do not distinguish the real but extinct animals like dinosaurs from the mythical creatures like dragons. This consequently leads to some more heavy misconception, when modern people despite their education cannot truly distinguish evidence-based science from fairy tales, which in turn brings serious world-view problems.

The possible way out of this "media trap" is high-quality multimedia brightening content that can be shown in museums, thematic exhibitions, etc. This content should incorporate the latest scientific data about prehistoric life and present them in visual and memorable from. While professional paleoartists already go this way [12], scientifically accurate artistic content is still badly needed.

We propose using scientific visualization techniques combined with the Internet of Things (IoT) [8] to develop interactive museum exhibitions, which can attract the visitors and help them to discover state-of-the-art paleontological concepts. IoT principles ensure high interactivity of exhibitions, while scientific visualization enables observable and attractive way for scientific data presentation.

2. Related Work

2.1. Smart Museum Concept

In the present time the museums are not just "storages for old things", but fully featured brightening places equipped with modern tools to represent information. The possibilities of todays microelectronics together with state-of-the-art IoT technologies enable so-called Smart Museums [7]. This concept assumes placing the IoT devices inside the museum rooms to allow advanced navigation and guides for visitors having smartphones or tablets with WiFi and Bluetooth support. Another application of IoT technologies in

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

museums is monitoring the visitors' activity inside the museum, which may be useful to find out which parts of exhibition attract more attention [4].

While state-of-the-art Smart Museum solutions focus on the indoor navigation, IoT technologies also provide a wide range of interaction possibilities. We propose to use sensors, actuators, and displays assembled into IoT devices to make the museum exhibitions much more vivid, ensuring different reactions to the visitors' actions performed both by mobile devices and by the people themselves.

2.2. Scientific Visualization Tools

In our previous research, we analyzed the state-of-the-art scientific visualization systems and found out that there is a lack of high-level tools to efficiently solve non-standard scientific visualization problems [10]. This was the motivation to create original scientific visualization system SciVi based on ontology engineering methods. The behavior of SciVi is completely governed by its ontological knowledge base. Hence it can be easily extended by replenishment of underplayed ontologies and thereby adapted to visualization tasks in almost any application domain with arbitrary data sources as input.

As the next step of SciVi development, we implemented the module enabling parts of SciVi to be incorporated in the firmware of IoT devices. This allows to automatically create lightweight visualization servers inside the IoT ecosystem, which ensure monitoring of IoT generated data as well as calibration of related electronic gadgets [9]. Consequently, we propose using SciVi as a visualization core for the interactive IoT-based museum exhibitions.

2.3. Volumetric Displays

There are many attempts to create volumetric displays [2], which enable showing 3D models with true volume, not just flat projections. Most of these technologies involve relative expensive devices and are quite complicated to assemble.

However, there are also some "compromise" solutions, which do not reach real volumetric images but provide optical illusions ensuring the perception of volume. First of all, those are technologies based on stereoscopy. Their drawback is the need of separation glasses or parallax-barrier displays.

The alternative solution is so-called "Pepper's Ghost" [3]. Its drawback is semitransparency of the resulting image, but sometimes this effect could be desired. Being invented in 1862, this technology was reborn since LCD displays appear [11]. In the technical and marketing slang it is currently called "3D Hologram", however, this name is not correct in terms of physics: no holography is used, just a light reflection on a slant transparent surface.

We propose using Pepper's Ghost as a way to display 3D models in museum exhibitions because this technology enables easy and cheap installation and ensures the possibility to combine shown 3D models with real exhibits being a kind of augmented reality.

2.4. *Dimetrodon grandis* Reconstruction

The dinosaurs are the most popular prehistoric animals attracting much more public attention than any other extinct creatures. Probably this is why the look at them evolves faster. However, the new paleontological knowledge allows not only dinosaur renaissance but also a renaissance of synapsids. Traditionally thought to be primitive, most likely synapsids were been much more progressive.

State-of-the-art work of Scott Hartman [6] describes the new look at *Dimetrodon grandis* Romer and Price, 1940, early Permian synapsid lived more than 290 million years ago. The distinctive features of Hartman's dimetrodon reconstruction are new sail shape, new back curvature, and high walk pose. While this reconstruction is very plausible, there is still lack of its visual representations across museums and media environment. It especially relates to 3D models. Almost the only high-quality scientifically accurate 3D model of *Dimetrodon grandis* found in the Internet[1] is not available by any license, not even commercial one.

To fill this media content gap we created our own 3D reconstruction of *Dimetrodon grandis* according to the S. Hartman's concept. The outfit of this reconstruction was inspired by the book "All Yesterdays", J. Conway and C.M. Kosemen [5] as well as by J. Conway drawing[2]. This model was used as a visual basis for the IoT-powered interactive museum exhibit.

3. IoT-Powered Paleontological Exhibit

We developed an exhibit of *Dimetrodon grandis* demonstrating two aspects of modern paleontological knowledge about this synapsid:

1. Anatomical features according to S. Hartman's reconstruction.
2. The concept that the dimetrodon's sail served to effectively heat the creature's body by the sun rays.

The first aspect is demonstrated by animated 3D model created in Blender 3D editor and displayed by the scientific visualization system SciVi using Pepper's Ghost technique.

The demonstration of the second aspect involves interactivity powered by IoT device, that includes the light direction sensor and WiFi connection module. The firmware for this device is generated by SciVi and includes visualization HTTP-server code in C++ and client code in JavaScript. Arbitrary third-party mobile device or computer can connect to this server via WiFi, obtain visualization client over HTTP and dis-

[1] https://sketchfab.com/models/21c7948d4d1e4c219ff6af19f71c8088
[2] http://johnconway.co/high-walkin

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

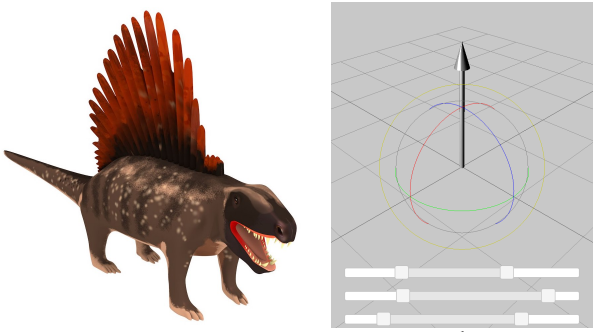
play the animated dimetrodon model using WebGL. The visualization server hosted on the IoT device provides WebSocket connection and transmits direction of light detected by the sensor. According to this data, the virtual light source is located on the graphical scene and the rendered dimetrodon automatically rotates its body to make its sail nearly perpendicular to the light rays.

The museum visitors can illuminate the sensor with a flashlight and see the dimetrodon reaction either on their smartphones/tablets connected to the IoT device or on the specially installed Pepper's Ghost display that provides the volumetric illusion.

3.1. 3D Model

The 3D model of *Dimetrodon grandis* is created in Blender using S. Hartman's reconstruction as the main anatomy reference, J. Conway's art as a texture reference and the skeleton anatomical casting from the Museum of Permian Antiquities as an auxiliary anatomy reference. Although this anatomical casting is based on "old school" concepts of dimetrodon anatomy, it still was useful to clarify some details not changed by S. Hartman, like, for example, the structure of fingers.

The created model is shown in Fig. 1 a. It includes ca. 3 000 vertices connected in ca. 5 000 triangles and contains 33 bones for skeletal animation. The surface material includes color textures and normal maps. For compatibility with rendering mechanisms, the model was exported to COLLADA format.



a b

Figure 1. Dimetrodon 3D model rendering result (a) and calibration scene of light direction sensor (b).

3.2. Light Direction Sensor

The detailed description of a light direction sensor used in the mentioned exhibit can be found in [9]. This device is based on the ESP8266 microcontroller with onboard WiFi communication module and contains 3 photoresistors VT90N2 placed at the vertices of an equilateral triangle with their surfaces oriented 45° to the ground. Each photoresistor detects the amount of light in the certain direction, collinear to its surface normal. The 3 surface normals $\vec{n}_1$, $\vec{n}_2$, $\vec{n}_3$ of photoresistors build the basis of vector space where the light incidence direction can be expressed as $l = a_1\vec{n}_1 + a_2\vec{n}_2 + a_3\vec{n}_3$. The expansion coefficients a_1, a_2, a_3 of the light incidence vector are evaluated based on the voltages on the photoresistors.

The firmware for this device was generated by scientific visualization system SciVi according to the data flow diagram composed in the special built-in high-level graphical editor. This diagram describes the entire software logic of the device including the photoresistors polling, calibration mechanisms, and graphical scene.

The functioning cycle of light direction sensor consists of two main stages: calibration and 3D model rendering. During the calibration stage, the user (or the museum hall administrator) sets up the basis vectors of the virtual scene space to match the real vectors of photoresistors' surface normals. This ensures the matching of light direction detected by the sensor with the direction of simulated virtual light. For that, a special graphical scene is used with the feedback to the sensor. This scene is demonstrated in the Fig. 1 b. The big arrow represents the direction to the light source, which corresponds to $l = M(a_1\vec{n}_1 + a_2\vec{n}_2 + a_3\vec{n}_3)$, where M is transformation matrix for the basis. The user can rotate the rendered arrow on the scene using drag'n'drop and thereby modify the matrix M. Three range sliders on the bottom in Fig. 1 b represent the mapping the resistor's voltages to the expansion coefficients $a_i = \frac{u_i}{1023}(R_i - r_i) + r_i$, where r_i and R_i are minimal and maximal values of the slider respectively, u_i is measured voltage on the resistor, $i = \overline{1,3}$, 1023 stands from maximal value of analog-to-digital converter of the device. The user can alter this mapping to compensate possible deviations in voltage measurements and thereby increase the sensor's precision. After the calibration is done, rendering stage can begin.

3.3. Rendering

The renderer is based on Three.js[3] engine and uses WebGL API to ensure high-quality real-time 3D model visualization. Two rendering modes are supported: regular rendering suitable for user's smartphone or tablet and special four-sided view rendering suitable for Pepper's Ghost illusion. In the first case, the dimetrodon model is visualized to the full screen. The rendering schema for the second case is shown in the Fig. 2 a. The letter "R" indicates orientation of the image.

The idea is to render the 3D model to texture and then map this texture to 4 trapezium-shaped sprites positioned as shown in the Fig. 2 a. The screen that displays the rendered image is located horizontally with a plexiglass frustum placed on top of it as shown in Fig. 2 b. The sides of the frustum are aligned to the

[3] http://threejs.org/

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

sprites. The angle between each side and screen surface is 45°. The texture content reflects on each side and the reflections build an illusion as if the rendered 3D model was inside the frustum.

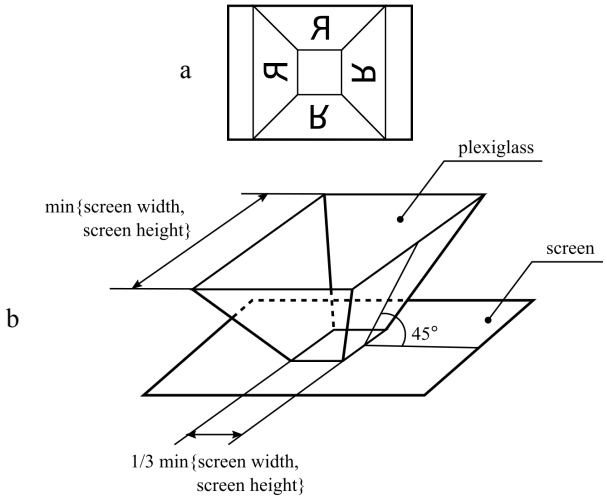


Figure 2. Rendering schema (a) and reflective frustum schema (b) for Pepper's Ghost illusion.

4. Conclusion

While IoT enables great capabilities for creating smart museums, one still needs high-level software to utilize different kinds of devices assembling interactive museum environment. We propose using adaptive scientific visualization system SciVi as a self-service software tool for creating IoT-powered interactive paleontological exhibitions. It allows generating the firmware for IoT-devices according to the user-defined data flow diagrams combining the calibration, interaction and visualization logic. Thanks to high-level graphical user interface SciVi can be used as a visual programming tool suitable for non-experienced programmers and device makers, for example, for museum research workers who create interactive exhibitions.

We demonstrate the concept of IoT-powered smart paleontological museum by creating the interactive reconstruction of *Dimetrodon grandis*. This exhibit includes scientifically accurate 3D reconstruction of dimetrodon that rotates its sail against the incoming light detected by the light direction sensor based on ESP8266 microcontroller. This exhibit is made in the museum of Permian antiquities.

For the next project, we plan to assemble the interactive scale of geological periods based on ultrasonic proximity sensor to enable museum visitors to navigate through the geological periods by moving hand along the scale and get the feedback from the exhibit.

5. References

[1] Bakker, R.T. Dinosaur Renaissance // Scientific American, Inc. – 1975. – PP. 58–78.

[2] Blundell, B.G. About 3D Volumetric Displays // Walker & Wood. – 2011. – 158 p.

[3] Blundell, B.G. Displays and Spatial Interaction: From Perception to Technology Volume I Exploring the Science, Art, Evolution and Use of 3D Technologies // Walker & Wood. – 2011. – 419 p.

[4] Chianese, A., Piccialli, F. Designing a Smart Museum: when Cultural Heritage Joins IoT // Third International Conference on Technologies and Applications for Smart Cities (I-TASC'14). – 2014. – 7 p. DOI: 10.1109/NGMAST.2014.21.

[5] Conway, J., Kosemen, C. M., Naish, D. All Yesterdays: Unique and Speculative Views of Dinosaurs and Other Prehistoric Animals // Irregular Books. – 2012. – 100 p.

[6] Hartman, S. Taking a 21st Century Look at Dimetrodon // Scott Hartman's Skeletal Drawing.com. – 2016 [Electronic Source]. URL: http://www.skeletaldrawing.com/home/21stcenturydimetrodon (Last accessed 05.05.2018).

[7] Kale, N., Itkarkar, R. Smart Museum Based on IoT // International Journal of Innovations & Advancement in Computer Science. – 2017. – Vol. 6, I. 8. – PP. 241–247.

[8] Rose, K., Eldridge, S., Chapin, L. The Internet of Things: an Overview // The Internet Society (ISOC). – 2015 [Electronic Source]. URL: https://www.internetsociety.org/resources/doc/2015/iot-overview (Last accessed 05.05.2018).

[9] Ryabinin, K., Chuprina, S., Kolesnik, M. Calibration and Monitoring of IoT Devices by Means of Embedded Scientific Visualization Tools // Lecture Notes in Computer Science. – Springer, 2018. – Vol. 10861. – PP. 655–668. DOI: 10.1007/978-3-319-93701-4_52.

[10] Ryabinin, K.V. Methods and Means for the Development of Adaptive Multiplatform Systems to Visualize Scientific Experiments / PhD Thesis: 05.13.11. – M., 2015. – 207 p. URL: http://library.keldysh.ru/diss.asp?id=2015-ryabinin (Last accessed 05.05.2018).

[11] Vasanth, N. How to Make a 3D Hologram Pyramid for Your Smartphone // Maker Pro. – 2015 [Electronic Source]. URL: https://maker.pro/custom/projects/diy-hologram (Last accessed 05.05.2018).

[12] Witton, M.P., Naish, D., Conway, J. State of the Palaeoart // Palaeontologia Electronica. – 2014. – Vol. 17, I. 3. – 10 p.

About the authors

Konstantin Ryabinin, PhD, assistant professor of computer science department, Perm State University.

Mariia Kolesnik, leading researcher of Perm Regional Museum / branch Museum of Permian Antiquities.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

К вопросу о феномене присутствия в системах компьютерной визуализации на базе виртуальной реальности

Н.В. Авербух[1]
natalya_averbukh@mail.ru
[1]УрФУ, Екатеринбург

В данной статье рассматриваются проблемы исследования феномена присутствия в связи разработкой систем компьютерной визуализации на базе виртуальной реальности. Ставятся задачи изучения феномена присутствия в различных условиях.

Ключевые слова: *виртуальная реальность, феномен присутствия, визуализация.*

On the phenomenon of presence in computer visualization systems based on virtual reality

N. Averbukh[1]
natalya_averbukh@mail.ru
[1]UrFU, Ekaterinburg

This article is about the problems of investigating the phenomenon of presence in connection with the development of computer visualization systems based on virtual reality. The problems of studying the phenomenon of presence under various conditions are set.

Keywords: *virtual reality, phenomenon of presence, visualization.*

1. Введение

Виртуальная реальность широко используется для игровых целей, в качестве тренажёра, при лечении фобий и для компьютерной визуализации. Однако при использовании сред виртуальной реальности для задач компьютерной визуализации встаёт ряд вопросов, связанных с целесообразностью использования. Виртуальная реальность требует специальной техники (очки виртуальной реальности или CAVE-системы), кроме того, может вызвать побочные эффекты, например, киберболезнь, и пребывание в ней в течение долгого времени не рекомендуется. В связи с этим можно подвергнуть сомнению использование технологий виртуальной реальности для задач, могущих быть реализованных на обычном экране. Однако виртуальная реальность может дать новое качество визуализации, и это не только трёхмерная графика, но и возможность рассмотреть объект виртуальной среды с разных ракурсов, заглянуть изнутри, обойти или облететь вокруг и т.п. [1]

Если в таких действиях возникает необходимость, то виртуальная реальность, способная погрузить пользователя в изучаемую среду, будет наилучшим решением.

В этой работе ставятся вопросы исследования центрального для виртуальной реальности психологического феномена – феномена присутствия.

2. Феномен присутствия, его определение и методы исследования

Определяющим понятием виртуальной реальности является переживание присутствия – состояния, при котором пользователь воспринимает себя находящимся «там», в среде, предложенной ему виртуальной реальностью. Как правило, это понятие используется для более или менее реалистичных сред, в которых пользователю предлагается передвигаться, желательно максимально естественным образом. Исследователи, принадлежащие к так называемой школе внутреннего присутствия [10–13] предлагают определять присутствие как явление, не связанное с технологиями виртуальной реальности и говорят о присутствии в реальном мире, которое в их понимании регулируется теми же механизмами, что и присутствие в виртуальной реальности.

Однако особенный интерес представляет не просто присутствие, а присутствие в виртуальной реальности при решении интеллектуальных задач, ведь именно они являются целью визуализации.

Согласно работам школы внутреннего присутствия [11, 12] присутствие тесно связано с возможностью осуществлять свои намерения, пользуясь предоставляемыми средой возможностями. Согласно [5, 6, 10] присутствие тесно связано с эмоциями: среда, вызывающая больше присутствия, вызовет больше эмоций и обратно, более эмоционально нагруженная среда вызывает более высокое переживание присутствия.

Методы исследования присутствия подробно описаны в обзоре [3]. К ним относятся объективные, измеряющие эмоциональную реакцию (как правило, страх) во время пребывания в потенциально пугающей среде (например, у края пропасти), поведенческие, фиксирующие реакцию человека, например, если человек заслоняется от летящего в него в виртуальной реальности мяча, то говорят о переживании им присутствия, и субъективные. Традиционно субъективные методы присутствия используют опросники, в которых задаются вопросы относительно реалистичности происходящего в среде, того, насколько пользователь был ею поглощён, обращал ли внимания на средство вывода информации и контроллеры, было ли ему удобно, замечал ли он события реального мира, следил ли за временем и т.п. Предполагается, что, при переживании присутствия человек забывает о реальном мире, забывает о времени и вовлекается в события виртуальной среды, не помня при этом, что взаимодействие с ней опосредуется технологией. Так, например, в опроснике Витмера и Сингера (Witmer B.G., Singer M.J.), [14].

Надо отметить, что понятие присутствия относится к тем понятием, которые до сих пор не имеют однозначного определения [7]. Присутствие является субъективным переживанием и во многом зависит от внутреннего состояния пользователя [6], от того, был ли он достаточно

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

вовлечён в среду, какие эмоции испытывал и т.п. Поэтому представляется необходимым проводить качественное исследование, в котором пользователям будет предоставлена возможность описать свои ощущения во время пребывании в виртуальной реальности. Предполагается, что в результате будет получено подробное описание того, что именно входит в переживание человеком феномена присутствия, описание, основанное не на теоретических выкладках, а на реальных случаях взаимодействия с виртуальной средой. Такая работа была проведена [2, 4], в ней испытуемым предлагался полёт на виртуальном вертолёте в рамках игры ГТА, приспособленной для нужд эксперимента. Испытуемым предлагалось структурированное интервью, вопросы в котором были связаны с событиями, произошедшими во время полёта. Ответы испытуемых показали яркие переживания, связанные с полётом, а также интересные случаи противоречивых ощущений, когда человек ощущал своё местонахождение одновременно в двух мирах: реальном и виртуальном.

3. Исследование присутствия в пассивных средах

Встаёт вопрос об исследовании присутствия в рамках работы пользователя с системами визуализации. Отличительной особенностью многих (хотя и, разумеется, не всех) систем является относительная неподвижность пользователя относительно среды и произведение мыслительных операций с представленной информацией. Поэтому за основу была взята та же задача, что и в первой части статьи [2], на базе которой исследовалась эффективность выполнения интеллектуальных задач в виртуальной реальности и в частности – при переживании присутствия. Данная задача представляет собой компьютерную версию теста «Кубики коса», в которой необходимо из верхних граней сначала четырёх, а потом девяти кубиков складывать такой же узор, как представлен на лежащей рядом с испытуемым карточке с образцом.

В работе [2] присутствие при решении интеллектуальной задачи исследовалось посредством опросника присутствия (Presence Questionnaire), разработанного Витмером и Сингером (Witmer B.G., Singer M.J.), [14], а также прямым опросом («Чувствовали ли Вы себя так, будто находитесь в представленной перед Вами среде?»). Однако такой метод вызвал вопросы, так как опросник не учитывает реальной ситуации, в которой находится испытуемый при взаимодействием с виртуальной реальностью, а прямой вопрос не позволяет выявить нюансы. После разработки метода интервью применительно к виртуальной реальности [2, 4] было решено вернуться к изучению присутствия при решении интеллектуальной задачи.

При разработке данного исследования встал вопрос об опорных точках интервью. При опросе испытуемых после полёта над виртуальным городом вопросы касались событий, происходящих в среде в результате умелого или неумелого управления виртуальным вертолётом. Однако при решении головоломки действия испытуемых не вызывали событий и среда не предполагала возможность для их появления.

В связи с этим необходимо поставить вопрос о возможностях изучения присутствия в пассивных средах, то есть в таких, в которых любые изменения строго контролируются действиями пользователя, и, более того, вопрос о том, в какой форме будет переживаться присутствие в такой среде. Очевидно, что в активной среде, то есть в той, в которой происходят неконтролируемые пользователем события, если она также учитывает действия пользователя и откликается на них, присутствие будет переживаться как ощущение реальности происходящего, как ожидание естественного для изображаемой ситуации взаимодействия и воздействия элементов среды. В такой среде возможен страх перед пугающими объектами или, напротив, поисковый интерес, любопытство по отношению к объектам и закономерностям, желание выяснить, какие возможности предоставлены данной средой. В среде же, созданной для интеллектуальной деятельности, не может быть пугающих объектов, так как они будут отвлекать от умственной деятельности, для которой разработана данная система. Точно также менее вероятен поисковый интерес и простое любопытство, поскольку их вытесняет заинтересованность в решении поставленной задаче. Таким образом, ожидается сосредоточенность пользователей на поставленной задаче и, возможно, на управлении объектами (что подтверждается предварительными результатами), а не на взаимодействии со средой. Однако предварительные исследования показывают, что даже пассивная среда может вызывать намерения и ожидания, которые можно связать с переживанием феномена присутствия: так, например, пользователь может пытаться или хотя бы хотеть раздвинуть руками неудобно стоящие виртуальные объекты вместо того, чтобы стремиться сделать это с помощью предложенного контроллера.

4. Проблематика, связанная с геймификацией

Задача изучения присутствия связана также с популярным подходом – геймификацией сред на базе виртуальной реальности, используемых для компьютерной визуализации [8, 9]. Предполагается, что пользователь, находящийся в привычном для себя игровом окружении и используя привычные методы игрового взаимодействия, сможет легко освоить среду визуализации. Однако это решение требует дополнительного изучения и экспериментов, так как неясно, что даст присутствие в игровой среде, не будет ли оно отвлекать от интеллектуальной деятельности. Кроме того, геймификация порождает вопросы, не будет ли конструктивное требование выполняемой деятельности вступать в противоречие с наработанными привычками игрока. Проще говоря, у специалиста-игрока может оказаться определённое ожидание от того, какие действия должны быть вызваны определённым привычным ему сочетанием клавиш, тогда как по причинам, связанным с особенностями разработки или поставленной задачи, это сочетание клавиш инициирует совершенно другое действие, неожиданное для него, на которое ему придётся заново переучиваться. Кроме того, встаёт вопрос, не окажется ли специалист, не имеющий опыта компьютерных игр, в проигрышном положении по сравнению с опытным геймером и не потребуется ли много усилий, чтобы освоить механизмы управления, понятные играющему в компьютерные игры специалисту за счёт привычки, но неудобные или непривычные сами по себе.

5. Заключение

Задача исследования присутствия в пассивных средах имеет теоретическое значение, так как может предложить новые данные, необходимые для описания этого явления и отделения его от близких понятий погружения, вовлечённости, интеллектуальной и эмоциональной, и т.п. Также нельзя забывать о практическом значении, так как, изучив вопрос порождения переживания присутствия, мы

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

можем вызывать его там, где оно будет способствовать решению задачи и по возможности предотвращать там, где оно будет вредить.

6. Литература

[1] Авербух В.Л., Анненкова О.Г., Артемова Н.А., Бахтерев М.О., Васёв П.А., Пестова М.С. Стародубцев И.С., Ушакова О.В., Разработка специализированной системы визуализации вычислительных сеток // Тезисы XVI Международной конференции Супервычисления и Математическое Моделирование». Тезисы. ФГУП «РФЯЦ ВНИИЭФ». Саров. 2016, стр. 15-16.

[2] Авербух Н.В. Особенности переживания феномена присутствия Материалы международной объединенной конференции "Интернет и современное общество". В печати. 2018

[3] Авербух Н.В. Психологические аспекты феномена присутствия в виртуальной среде // Вопросы психологии. 2010. № 5. С. 105-113.

[4] Авербух Н.В., Щербинин А.А. Феномен присутствия и его влияние на эффективность решения интеллектуальных задач в средах виртуальной реальности // Психология. Журнал Высшей школы экономики. Том 8, N 4, 2011, стр. 102-119.

[5] Baños R.M., Botella C., Alcañiz M., Liaño V., Guerrero B., Rey B. Immersion and Emotion: Their impact on the sense of presence. Cyberpsychology and Behavior, 7, 734-741. 2004.

[6] Bouvier P. The five pillars of presence: guidelines to reach presence. Presence 2008: The 11th Annual International Workshop on Presence, Pp. 246-249

[7] Lombard, M., & Jones, M. T. (2015). Defining presence. In M. Lombard, F. Biocca, W.A. Ijsselsteijn, J. Freeman, & R. Schaevitz (Editors), Immersed in Media: Telepresence Theory, Measurement and Technology. London: Springer. Pp. 13-34.

[8] Merino L., Ghafari M., Anslow C., Nierstrasz O. CityVR: Gameful Software Visualization. IEEE International Conference on Software Maintenance and Evolution (ICSME TD Track). 2017. Pp.633-637

[9] Oberhauser R., Lecon C. Gamified Virtual Reality for Program Code Structure Comprehension. The International Journal of Virtual Reality, 2017, 17 (02). Pp. 79-88

[10] Riva G., Mantovani F., Capideville C. S., Preziosa A., Morganti F., Villani D., Gaggioli A., Botella C., Alcaniz M. Affective interactions using virtual reality: the link between presence and emotions. Cyberpsychology and Behavior, 10, 45-56. 2007.

[11] Riva, G., Waterworth, J. A., Waterworth, E. L., and Mantovani, F. From intention to action: the role of presence. New Ideas Psychol. 29, 24–37. 2011.

[12] Triberti S., Riva, G. Being present in action: a theoretical model about the “interlocking” between intentions and environmental affordances Frontiers in Psychology (2016) 6:2052, Pp. 21-28.

[13] Waterworth, J., and Riva, G. Feeling Present in the Physical World and in Computer-Mediated Environments. London: Palgrave Macmillan. 2014

[14] Witmer B.G., Singer M.J. Measuring Presence in Virtual Environments: A Presence Questionnaire // Presence. Vol. 7. No 3. 1998. P. 225–240.

Об авторах

Авербух Наталья Владимировна, младший научный сотрудник лаборатории мозга и нейрокогнитивного развития кафедры клинической психологии и психофизиологии департамента психологии уральского гуманитарного института Уральского Федерального Университета, г. Екатеринбург. Её e-mail natalya_averbukh@mail.ru

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Визуализация программного обеспечения с использованием виртуальной реальности

В.Л. Авербух [1,2], И.Л. Гвоздарев [2], Г.И. Левчук[2]
averbukh@imm.uran.ru, ilyagvz@gmail.com, levchukgeorgy1@gmail.com
[1] ИММ УрО РАН, [2]УрФУ, г. Екатеринбург, Российская Федерация

Работа посвящена системам визуализации программного обеспечения на базе виртуальной реальности. Делается обзор разработки таких систем. Описываются проекты систем визуального программирования и визуализации программного обеспечения на базе виртуальной реальности с использованием метафор солнечной системы и расширенной метафоры города.

Ключевые слова: *визуализация программного обеспечения, виртуальная реальность, метафоры визуализации.*

Software Visualization based on Virtual Reality

V. L. Averbukh, I. L. Gvozdarev, G. I. Levchuk

The paper is devoted to Software Visualization systems based on Virtual Reality. The overview of design and development of such systems is presented. The projects of visual programming and software visualization systems based on virtual reality using Solar System metaphor and the extended City metaphor are described.

Keywords: *software visualization, virtual reality, visualization metaphors.*

1. Введение

Работа посвящена проблематике визуализации программного обеспечения и современному состоянию этой дисциплины в связи с использованием сред виртуальной реальности. Визуализация программного обеспечения (Software Visualization), которая активно развивается уже более тридцати лет, является подобластью дисциплины Компьютерная визуализация. Этой теме посвящено большое количество публикаций и выступлений на различных конференциях. Ежегодно проводятся конференции по визуализации программного обеспечения ***VISSOFT.*** В последнее время вновь проявился интерес к использованию сред виртуальной реальности в системах визуализации программного обеспечения. Ниже будут описаны некоторые положения этой дисциплины. На примере метафоры города рассматривается понятие ***свойство метафор.*** Затем приводится обзор систем визуализации программного обеспечения на базе сред виртуальной реальности. Анализ этих систем показывает необходимость проектирования и разработки собственных прототипов и проведения исследований возникающих возможностей и проблем.

2. Визуализация программного обеспечения

Под визуализацией программного обеспечения понимается совокупность методик использования графики и средств человеко-машинного взаимодействия, применяемых для лучшего уяснения понятий и эффективной эксплуатации программного обеспечения, а также для спецификации и представления программных объектов в процессе создания программ. Первые системы визуального программирования и визуализации программ, которые можно отнести к визуализации программного обеспечения, появились в конце 70-ых — начале 80-ых годов. Появление этой дисциплины связано с практикой и идеями графического представления программ, основанными на активном использовании программистами блок-схем (графов потоков управления) и графов потоков данных. В число систем визуализации программного обеспечения обычно включаются системы визуального программирования, системы визуализации программирования. Выдвигалась идея, что и разработка программ, и и их отладка, и их сопровождение должны осуществляться в рамках единой системы с одинаковым графическим представлением программных сущностей. Описание новой дисциплины было осуществлено в начале 90-ых. Несмотря на использование примитивной (по нынешним временам) графики в 80-ых и 90-ых годах были разработаны интересные системы визуального программирования и полноценные визуальные отладчики. Однако использование традиционных методов двумерной графики очень быстро стало ограничением для систем визуализации программного обеспечения. Даже сравнительно небольшие отображения, например, сотни связанных между собой объектов были сложны как для восприятия, так и для интерпретации. Появились публикация, в которых указывалось на кризис в области визуализации программного обеспечения. Выходом в этом случае стало использование метафор визуализации, трехмерной графики и средств виртуальной реальности.

3. Метафоры визуализации

Неформально под метафорой визуализации понимается главная идея при отображении прикладной области на визуальный мир. Метафоры используются для определения деятельности пользователя программной системы и его восприятия объектов и операций над ними. Среди метафор, используемых в системах визуализации программного обеспечения, большой популярностью пользуются пространственные метафоры, в частности, метафора города. На этом примере мы рассмотрим такое понятие как «свойство метафоры». Метафора города дает возможность получить представление структуры программного проекта в виде карты некоторой местности, реалистичного трёхмерного представления городских кварталов, завода или обобщенных образов, напоминающих фантастический пейзаж. Возможности этой метафоры позволяют отображать *иерархию элементов системы* (отдельные строения в рамках городского квартала), *потоки управления, потоки данных, иные связи между элементами программ, программными конструкциями или частями программного комплекса* (улицы и дороги, связывающие цеха фабрики). Эта метафора может служить основой для построения видов отображения в системах отладки, тестирования и мониторинга работы параллельных и распределенных программ, также как и программ, обрабатывающих события и обеспечивающих реакции на

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

них. Среди свойств метафоры города можно выделить ***неограниченный контекст, естественность, организация внутренней структуры, устойчивость к масштабированию***. Метафора города предполагает наличие внутренней структуры и допускает представление достаточно большого объема однородной в визуальном смысле информации. При этом интерпретация полученных графических выводов представляется простой. Возможность пролета над городом позволяет осуществление удобной и легкой навигации. Полет с изменением высоты дает возможность реализовать масштабирование и зуминг. Это позволяет использовать метафору города при разработке систем визуализации программного обеспечения на базе ***сред виртуальной реальности***.

4. Визуализация программного обеспечения на базе сред виртуальной реальности

В первой половине 90-ых годов появилась первая система визуализации программного обеспечения, активно использующая средства виртуальной реальности и метафору комнаты-здания в трехмерном варианте и функционировала на базе среды виртуальной реальности CAVE. Визуальная среда Avatar [12] была предназначена для представления больших объемов данных о производительности параллельных систем, получаемых непосредственно в ходе их работы, что использовалось для отладки производительности параллельных программ. По ходу работы пользователь как бы оказывается внутри трехмерного помещения, где на стены проецируется видеоизображение. На внутренних гранях куба - полу и стенах помещения выводятся оси и кривые, описывающие метрики производительности параллельной программы. Данные о производительности параллельных программ отображались в виде традиционных двумерных графиков. Реализовано объединение отдельных элементов, которое напоминает стеклянный небоскреб. Вводится визуальное отображение - аналог фазового портрета - "лента истории". Путешествие (“виртуальный полет”) по небоскребу на основе этой ленты давал возможность исследования данных о производительности параллельной программы. Обратим внимание на то, что несмотря на очень интересную реализацию системы Avatar сведений о ее использовании или продолжении разработок найти не удалось, что возможно связано как с недостаточной информативностью двумерных графиков, использованных при визуализации данных о производительности процессоров, так и с возможными неприятными и даже болезненными ощущения проявлениями, вызванными у пользователя при “виртуальном полете”, который проходит согласно ленте истории работы многопроцессорной системы. Проявления ***киберболезни*** (*cyberthickness*) часто связываются с отсутствием возможности активного управления событиями виртуальной реальности.

Новый интерес к использованию виртуальной реальности в системах визуализации программного обеспечения проявился в самом начале 2000-ых годов [4], [5], хотя дальше интересных идей первые публикации нового тысячелетия не пошли. Вновь работы в этом направлении начали появляться несколько лет назад.

В работе [3] было описан прототип системы визуализации программного обеспечения на основе метафоры города, в которой использовался шлем виртуальной реальности и интерфейс на основе языка жестов. Для первичной оценки системы авторы провели структурированные интервью, где участники должны были решить три задачи на понимание программы и оценить удобство использования жестов, а также указать свои общие впечатления от использования средств виртуальной реальности для понимания смысла программы. Также эта исследовательская команда опубликовала работу по использованию физического моделирования на базе 3D-принтинга для задач визуализации программного обеспечения. [2] Визуализация программных комплексов проводится на основе метафоры города. В результате пользователи получают физические модели объектов программы, что повышает возможности их представления и понимания результата работы. Наличие физических моделей способствует лучшему взаимопониманию участников разработки. Предполагается возможность использования этих моделей в процессе образования.

Метафора города является весьма популярной в системах визуализации программного обеспечения, использующих средства виртуальной реальности. Две системы, известные из публикаций 2017 года, носят весьма схожие имена - VR City [15] и CityVR [7]. В системе VR City используется модифицированная метафора города для представления программных систем и связанных с ними аналитических данных. Визуализируются статические (метрики) и динамические (трассы) аспекты программ. Пользователи могут наблюдать и взаимодействовать с объектами «города» в среде иммерсивной виртуальной реальности. Предоставляется функция просмотра исходного кода. В публикации [7] рассматривается не только использование метафоры города для визуализации программного обеспечения на базе виртуальной реальности, но и подход, который называется ***геймификация*** задач разработки программного обеспечения. Этот подход подразумевает создание инструментов, которые предоставляют разработчикам программного обеспечения интерфейс, аналогичный компьютерным играм. Проводился анализ того, как разработчики взаимодействуют с системой визуализации CityVR. Разработчики были взволнованы, они почувствовали не только интерес к работе, но и определенный вызов. Они испытывали погружение в виртуальный мир при сохранении управляемости системы. При взаимодействии для навигации в виртуальном мире и при выборе нужных элементов программы было затрачено значительное время, при этом пользователи считали, что время прошло быстрее, чем в реальности. Также на ***геймификацию*** при разработке систем визуализации программного обеспечения на базе виртуальной реальности указывается в [11]. Разрабатываемая среда на базе виртуальной реальности должна обеспечить работу со структурами программного кода с использованием нескольких метафор (включая метафору города) для визуализации, навигации и передачи информации о программном коде в интерактивном режиме. Были реализованы игры, продемонстрировавшие потенциал геймификации для улучшения понимания структурных зависимостей и модуляризации кода. Эта команда исследователей опубликовала несколько работ, посвященных разработке систем визуализации программного обеспечения на базе виртуальной реальности [9],[10] (включая работу по программированию с использованием виртуальной реальности [8]). В работе [9] рассматривается перспективы использования погружения (immersion) разработчиков в структуры программ, а в [10] виртуальный полет над программными объектами в рамках различных метафор визуализации. Работа [14] посвящена визуализации программных систем на базе виртуальной реальности. Инструмент позволяет программистам, менеджерам проектов или заказчикам исследовать архитектуру и получить первое впечатление о размерах компонентов и их зависимостях. В виртуальных сценах используется простая графика — кубы различных размеров

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

и цветов для представления программных объектов с показом зависимостей в виде соединительных линий. В работе [13] рассматривается использование смешанной реальности для отладки кода. Предлагается интерактивная 3D визуализация потока сообщений, сочетающая традиционные методики отображения с дополненной реальностью на базе шлема виртуальной реальности. В [8] представлена система на базе виртуальной реальности для программистов. Эта система должна обеспечить визуализацию структуры программного кода в нескольких метафорах. Реализован встроенный просмотр работы с реальными клавиатурой и мышью и взаимодействие в смешанной реальности. Это обеспечивает поддержку основных задач программирования, не выходя из среды виртуальной реальности. Реализован прототип системы. Проведено исследование, демонстрирующее осуществимость идей. Получены результаты эмпирической оценки, демонстрирующие потенциал системы.

В работе [1] описан инструмент визуального программирования с использованием иммерсивной виртуальной реальности. Предполагается использование данной системы для создания встроенных цифровых комплексов. Система позволяет описывать сложные объекты и их связи, задавать логические конструкции визуализировать в реальном масштабе времени потоки данных между реальными объектами разрабатываемых комплексов. Используется естественная для приложений образность. Первичный анализ экспериментальных результатов показал, что участники различных уровней квалификации могут успешно создавать и отлаживать программы в рамках заданных сценариев.

Наконец отметим работу, посвященную оценке эффективности систем трехмерной визуализации программного обеспечения [6]. Для эксперимента была реализована трехмерная визуализация на базе метафоры города на стандартном экране компьютера, в среде виртуальной реальности и на физической модели с использованием 3D-печати. Участники (которые были разделены на три группы для каждого носителя) провели визуализацию программных систем различных размеров, решили набор задач на понимание результатов и заполнили анкету. Эффективность визуализации была оценка с точки зрения производительности, набора впечатлений и пользовательского опыта. Хотя разработчики, использовавшие физическую визуализацию, затратили наименьшее времени для идентификации резко отклоняющихся значений, наименьшие сложности были достигнуты при визуализации систем на основе стандартных экранов. При этом разработчики, использовавшие средства виртуальной реальности, получили самые большие впечатления.

5. Разработка прототипов систем

Анализ примеров систем визуализации программного обеспечения на базе сред виртуальной реальности показывает, что несмотря на 25 лет истории, остается еще много вопросов по применимости этих сред на практике. Все это требует дополнительных исследований и разработок систем с учетом конкретных задач программного обеспечения. Ниже описываются проекты и реализации прототипов систем визуализации программного обеспечения на базе сред виртуальной реальности, которые послужат для этих нужд.

Разрабатывается визуальный язык на основе объектов, когда визуальная среда программирования представляет графические или символьные элементы, которыми можно манипулировать в соответствии с некоторыми правилами. Целью проекта является создание среды визуального программирования с возможностью абстрагироваться от правил и спецификаций конкретного текстового языка, чтобы сосредоточиться на решаемой проблеме при написании программы. Таким образом, программист, манипулируя графическими объектами на разных уровнях абстракции, сможет лучше понимать структуру программы и минимизировать число связанных ошибок (несоответствие типов, нарушение синтаксиса и т.п.).

В качестве идеи среды визуального программирования выбрана метафора солнечной системы, среди свойств которой можно указать на ***естественность, организация внутренней структуры, устойчивость к масштабированию***. Данную метафору можно использовать при разработке систем визуализации программного обеспечения на базе ***сред виртуальной реальности***. Все пользовательские классы представляются планетами. (Рис. 1). У каждой планеты (класса) есть два вида отображения: **свободный** (привычный вид планеты со спутниками и кольцами) и **активный** (когда класс выбран, выглядит как развертка планеты). Активный вид представляет собой круг с сектором колец слева и упорядоченными спутниками справа. **Сектор колец** - методы класса. Каждое кольцо – отдельный метод. **Внешние кольца** – public-методы, внутренние – private и protected. **Спутники** – поля класса. Сверху вниз сначала идут спутники, более удаленные от планеты – public-поля, менее удаленные – private и protected. Внутри планеты также есть сектор колец, упорядоченные спутники (static-методы и поля), принадлежащие самому классу и в центре ядро хранящее все конструкторы класса. Использование виртуальной реальности дает преимущества в навигации. Появляется возможность путешествия по структуре классов, что поможет устранить неоднозначности сложных абстрактных представлений, позволит пользователю оценить относительный размер объектов и дистанции между ними. Предполагается осуществлять взаимодействие с системой с помощью используемых в средах виртуальной реальности устройств, а не традиционной клавиатуры. Реализована возможность создания примитивных классов, внутри которых нет зависимости от других классов (поля-примитивы), с помощью манипулирования графическими объектами. Также реализована возможность создания циклов и условных операторов с помощью манипулирования графическими объектами. Увеличение сложности ПО не должно сильно увеличивать сложность графического представления. В дальнейшем возможно расширение среды визуальной разработки ПО.

Кроме того, разрабатывается проект визуализации программного обеспечения для нужд разработчиков и тестеровщиков на базе виртуальной реальности. Предполагается использование двух форм представления — графовое и на основе метафоры города. В результате обеспечивается взаимодействие с пользователем при формировании визуализации выбранного проекта с выведением информации о всех составляющих единицах кода. Представляется, что для тестировщиков удобнее визуализация на базе метафоры города. Разработчик, уяснив проблему, открывают визуализацию кода в метафоре графа. Графовое представление используется для исследования кода и получения информации о его структуре на данный момент, а метафора города позволяет получить информацию для коррекции кода. (Рис. 2). Разрабатывается функционал для добавления класса в виде трехмерной модели здания, что позволит использовать внутренние элементы здания для визуализации класса. Предполагается расширение метафоры города за счет добавления активных агентов путем ввода параметров в определенные функции и методы. Агенты могут передвигаться внутри города, тем самым определяя в каких

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

местах они используются, где изменяются и как проходит процесс работы алгоритмов. Рассматривается возможность обучения системы за счет сравнения кусков кода, поиска примерно одинаковых мест. Это позволит предсказывать влияние изменений кода на его работу

При разработке программного обеспечения, контролирующего работу реальных изделий и механизмов, целесообразно визуализировать их работу, используя естественную графику. Виртуальная реальность позволяет построить отображения изнутри объектов, размещать в трехмерной модели различные датчики и увидеть, как функционирует устройство под управлением программы.

Рис. 1. Общий вид среды визуального программирования на основе метафоры солнечной системы.

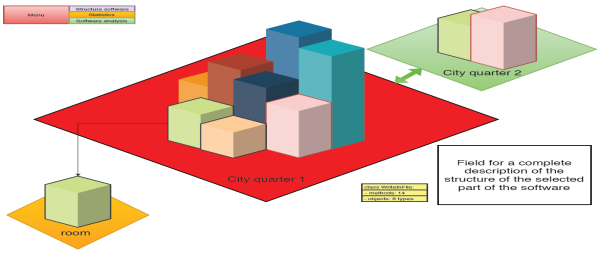


Рис. 2. Общий вид интерфейса системы на базе метафоры города

6. Заключение

Разработка прототипов систем на базе виртуальной реальности — лишь первый этап исследований. На базе этих систем можно прояснить те вопросы, на которые нельзя найти ответы в имеющихся публикациях. Необходим поиск критериев выбора метафор визуализации и подходов к проектированию видов отображения. Не ясно насколько эффективны среды виртуальной реальности для использования в задачах визуализации программного обеспечения. Остается множество вопросов к адаптации таких систем с учетом особенностей пользователей, включая индивидуальные ограничения по восприятию и длительному нахождению в средах виртуальной реальности. Идея геймификации весьма интересна, но не всегда может быть использована в серьезных разработках. Отсюда вытекает проблема описания деятельности программиста в качестве пользователя систем визуализации программного обеспечения. Мы занимаемся изучением деятельности пользователей систем научной визуализации. Причем эта задача требует большого числа экспериментов и соответствующей программной поддержки. Отметим, что упомянутые в приведенном выше обзоре физическая визуализация и жестовые интерфейсы использовались нами в случае специализированных систем научной визуализации. Причем физическая визуализация моделей оказалась весьма интересной и полезной для конечных пользователей, а простой язык жестов оказался достаточно эффективным. Для систем визуализации программного обеспечения необходимы новые исследования и экспериментальные разработки. Наконец, при проектировании систем на базе виртуальной реальности следует учитывать такой фактор как феномен присутствия (*presence*), а также проявления киберболезни, тем более, что до 30% участников исследований могут почувствовать ее проявление [9].

Литература

[1] Ens B., Anderson F., Grossman T., Annett M., Irani P., Fitzmaurice G. Ivy: Exploring Spatially Situated Visual Programming for Authoring and Understanding Intelligent Environments // Proceedings of the 43rd Graphics Interface Conference GI '17. 2017. Pp. 156-162.

[2] Fittkau F., Koppenhagen E., Hasselbring W. Research Perspective on Supporting Software Engineering via Physical 3D Models // 2015 IEEE 3rd Working Conference on Software Visualization (VISSOFT). Pp. 125-129.

[3] Fittkau F., Krause A., Hasselbring W. Exploring Software Cities in Virtual Reality // 2015 IEEE 3rd Working Conference on Software Visualization. Pp. 130-134.

[4] Knight C., Munro M. Virtual but Visible Software // IV 2000 Information Visualization. IEEE International Conference on Information Visualization, 2000. Proceedings. Pp. 198-205.

[5] Maletic J., Leigh J., Marcus A. Visualizing Software in an Immersive Virtual Reality Environment // Proceedings of the Workshop on Software Visualization. 2001. Pp. 49-54.

[6] Merino L., Fuchs J., Blumenschein M., Anslow C., Ghafari M., Nierstrasz O., Behrisch M., Keim D.A. On the Impact of the Medium in the Effectiveness of 3D Software Visualizations // Proceedings on 2017 IEEE Working Conference on Software Visualization, 2017. Pp. 11-21.

[7] Merino L., Ghafari M., Anslow C., Nierstrasz O. CityVR: Gameful Software Visualization // IEEE International Conference on Software Maintenance and Evolution (ICSME TD Track). 2017. Pp. 633-637.

[8] Oberhauser R. Immersive Coding: A Virtual and Mixed Reality Environment for Programmers // Proceedings of The Twelfth International Conference on Software Engineering Advances (ICSEA 2017). Pp. 250-255.

[9] Oberhauser R. Carsten Lecon Immersed in Software Structures: A Virtual Reality Approach // ACHI 2017. The Tenth International Conference on Advances in Computer-Human Interactions. Pp. 181-186.

[10] Oberhauser R. Carsten L. Virtual Reality Flythrough of Program Code Structures // Proc. of the 19th ACM Virtual Reality International Conference. ACM, 2017. (4 pp.)

[11] Oberhauser R. Carsten L. Gamified Virtual Reality for Program Code Structure Comprehension // The International Journal of Virtual Reality, 2017, 17 (02). Pp. 79-88.

[12] Reed D., Scullin W., Tavera L., Shields K., Elford Ch. Virtual Reality and Parallel Systems Performance Analysis // IEEE Computer, V.28, N 11, (1995). Pp. 57-67.

[13] Reipschläger P., Gumhold S., Ozkan B.K., Majumdar R., Mathur A.Sh., Dachselt R. DebugAR: Mixed Dimensional Displays for Immersive Debugging of Distributed Systems // Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. (6 pp.)

[14] Schreiber A., Bruggemann M. Interactive Visualization of Software Components with Virtual Reality Headsets // 2017 IEEE Working Conference on Software Visualization (VISSOFT), 2017, Pp. 119-123.

[15] Vincur J., Navrat P., Polasek I. VR City: Software Analysis in Virtual Reality Environment // 2017 IEEE International Conference on Software Quality, Reliability and Security Companion. Pp. 509 – 516.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Об авторах

Авербух Владимир Лазаревич, заведующий сектором компьютерной визуализации ИММ УрО РАН; доцент кафедры вычислительной математики и компьютерных наук УрФУ. E-mail – averbukh@imm.uran.ru.

Гвоздарев Илья Леонидович, аспирант кафедры вычислительной математики и компьютерных наук УрФУ E-mail – ilyagvz@gmail.com.

Левчук Георгий Иванович, аспирант кафедры вычислительной математики и компьютерных наук УрФУ E-mail – levchukgeorgy1@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Построение стереоанимаций на современных стереоустановках

Андреев С.В., Бондарев А.Е.
esa@keldysh.ru | bond@keldysh.ru
ИПМ им. М.В. Келдыша РАН, Москва, Россия

Данная работа посвящена задачам построения стереоизображений и стереоанимаций при помощи современных стереоустановок активного и пассивного типов. Рассматриваются различные способы построения стереоизображений с учетом приобретённого практического опыта. Приводятся иллюстрации для тестовых изображений.

Ключевые слова: *стереоанимация, карта глубины, многовидовый кадр.*

The construction of stereo animations on modern stereo devices

Andreev S.V., Bondarev A.E.
esa@keldysh.ru | bond@keldysh.ru
Keldysh Institute of Applied Mathematics RAS, Moscow, Russia

This work is devoted to the tasks of constructing stereo images and stereo animations using modern stereosets of active and passive types. Various methods for constructing stereo images are considered, taking into account the acquired practical experience. Illustrations for test images are given.

Keywords: *stereo animation, depth map, multi-view frame.*

1. Введение

Человеку по своей природе свойственно восприятие окружающего мира в объемном, стереоскопическом виде, то есть способность одновременно четко видеть объекты нашего окружения двумя глазами одновременно. И, соответственно, если мы хотим добиться полной реалистичности визуализации объекта, применительно к результатам расчетов, необходимо учитывать особенность бинокулярного стереоскопического восприятия человеком визуальной информации. При представлении визуальной информации в стереоскопическом, объемном виде, естественном для восприятия человека, особенно результатов расчетов объекта, не существующего в реальности, данная информация будет восприниматься и усваиваться зрителем намного быстрее и продуктивнее, чем визуализация в плоском, двумерном виде на экранах, так как восприятие будет происходить на уровне подсознания, с автоматической оценкой мозгом формы объекта визуализации и расстояния до него на фоне других объектов, всех преимуществ и недостатков сгенерированного объекта, что позволяет почти автоматически найти возможные ошибки расчетов, проблемы и пути для их решения.

2. Современные системы демонстрации изображений в объемном представлении

В целом, системы демонстрации изображений в объемном представлении (стереоустановки) можно разделить на два основных вида: пассивные и активные. Пассивные комплексы предназначены для демонстрации материала по заранее намеченному сценарию, когда зритель лишен возможности повлиять на процесс показа; активные (или интерактивные) комплексы отличаются тем, что зритель сам каким-либо образом влияет на процесс показа демонстрируемого материала.

Общим для всех этих систем является наличие большого экрана (и даже нескольких экранов), таким образом, типичного разрешения монитора чаще всего не хватает для генерации кадра, выводимого на экран стереоустановки. Кроме того, такие системы должны обеспечивать показ в объемном представлении, то есть в режиме стерео, когда каждый кадр генерируется отдельно для левого и правого глаза (правый и левый стереоканал). Таким образом, требования к вычислительным ресурсам для генерации кадров возрастают вдвое. Часто такие активные стереоустановки управляются не отдельным компьютером, мощности которого не хватает для генерации кадров в режиме реального времени, а системой компьютеров, объединенных в локальную сеть и обеспечивающих вывод составного стереокадра на несколько экранов. Специфические проблемы, возникающие при использовании системы компьютеров для генерации и визуализации составного мультиэкранного стереокадра, и методы решений таких проблем достаточно подробно описаны в [1,2].

Постепенное распространение современных стереоустановок, как пассивного, так и активного типов, породило большой интерес к разработке практических алгоритмов стереопредставления статических и анимированных изображений, что отражено в работах [3,5 - 7]. Данная работа относится к общей тематике построения стереоизображений и стереоанимаций результатов математического моделирования сложных технических объектов и физических процессов в сплошных средах и представляет собой практический опыт построения стереоанимаций для конкретных устройств, которыми располагает ИПМ им. М.В. Келдыша РАН. Первый тип устройства представляет собой 3D проекционную стереосистему для показа стереопрезентаций, обучающих приложений, графики и фильмов. Включает графическую станцию, проекционную подсистему из двух проекторов (со специализированным экраном), акустическую систему. Данная система является классической стереографической системой с использованием двух проекторов, экрана и очков линейной поляризации.

Второй тип устройства представляет собой автостереоскопический монитор Dimenco DM654MAS. Автостереоскопические мониторы обеспечивают показ стереоизображений без необходимости отслеживать позицию наблюдателя, то есть, нет необходимости подстраивать оборудование в зависимости от положения наблюдателя – как правило, такие мониторы позволяют наблюдать стереоизображения, обеспечивая несколько фиксированных сегментов в пространстве для наблюдения,

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

причем зритель может перемещаться из одного сегмента в другой, получая возможность рассматривать демонстрируемый объект в 3D с различных углов зрения. Принцип работы автостереоскопического монитора – это использование параллаксных перегородок или линз Френеля, установленных за защитным стеклом экрана (рис. 1).

Когда голова наблюдателя находится в определенной позиции перед автостереоскопическим монитором, его правый и левый глаз получают различные изображения (стереопара). Таким образом, создается убедительная иллюзия 3D глубины.

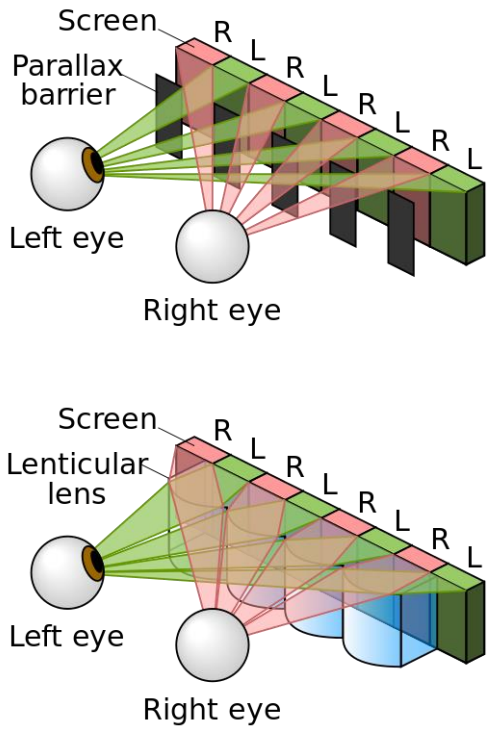


Рис. 1. Сравнение технологий параллаксных перегородок и линз Френеля для показа 3D изображений в автостереоскопических мониторах

Для визуализации расчетов в объемном виде на данных устройствах нами были использованы несколько методов.

3. Метод «облета камеры» при использовании угловой стереобазы

При создании стереопар, требуемых для кадров стереоанимации, используется либо линейное, либо угловое смещение камеры, то есть линейная или угловая стереобаза. В большинстве случаев для визуализации результатов расчетов необходимо рассмотреть объект визуализации с разных сторон, поэтому для этих целей наиболее подходящим является метод «облета» камеры вокруг объекта, то есть угловая стереобаза.

Практически любая CAD-система позволяет визуализировать трехмерный объект с заданными координатами положения по отношению к пользователю. Поэтому задача сводится к построению последовательности изображений CAD модели с зафиксированной осью Z, проходящей через ее геометрический центр, и смещением на определенный угол. После каждого такого смещения проводится рендеринг модели, позволяющей достичь фотореалистичности, и сохранение результата в виде графического файла в высоком разрешении. Для левого стереоканала устанавливалась начальная позиция, отличающаяся от правого стереоканала на угол, определяющий параллакс, и позволяющая добиться объемности объекта в результирующем стереофильме. Данный метод привлекателен еще и тем, что для статической модели, то есть с неизменной геометрией и цветовым распределением, возможно повторное использование уже рассчитанных кадров, например, если для левого стереоканала мы используем 1,2,3,4, … кадры, то для правого стереоканала достаточно начать последовательность со второго или с третьего кадра (2,3,4,5, …), автоматически обеспечивая сдвиг объекта на необходимый угол стереобазы.

Дополнительный выигрыш такого «облета камеры» состоит также в том, что имея всего 360 монокадров можно организовать зацикливание вращения объекта визуализации в объемном представлении, то есть организовать его непрерывное, сколь угодно долгое вращение, повторно демонстрируя зрителю все стороны объекта (рис. 2).

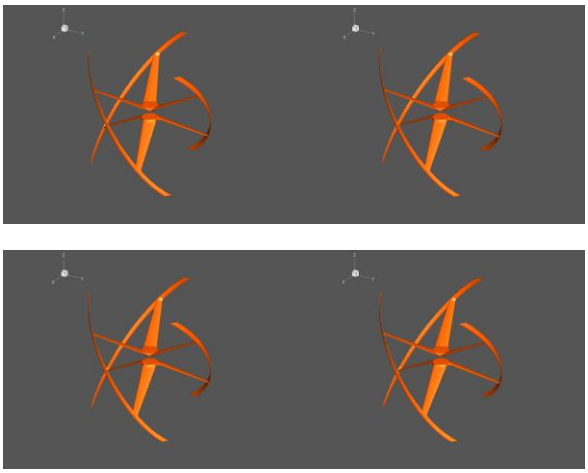


Рис. 2. Стереопары с повторным использованием рассчитанных кадров при угловом смещении (угловая стереобаза).

Данный метод был успешно реализован при визуализации моделирования работы энергетической установки сложной формы в режиме стереоанимации как для классической стереоустановки проекционного типа, так и для визуализации на автостереоскопическом мониторе [4].

4. Использование карт глубины

При создании многоракурсного видео для автостереоскопических мониторов возникает потребность хранения большого количества данных – видеопотоков для каждого из ракурсов. Даже с учетом того, что современные методы цифрового видеосжатия позволяют эффективно учитывать временную и пространственную избыточность, объем данных при многоракурсном видео возрастет многократно. Особенно это критично для автостереоскопических мониторов.

Один из эффективных способов решения проблемы большого объема данных состоит в использовании так называемого формата 2D+Z. Любому обычному (2D) изображению можно сопоставить информацию об удаленности каждого пиксела от наблюдателя (Z-координату). Такое представление изображения называют "формат 2D+Z", а плоскость координат Z – "картой

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

глубины". Ее можно представить в виде монохромного изображения. В карте глубины градациями серого обозначается удалённость точек изображения от наблюдателя. То есть самая ближняя точка к зрителю станет белой, а самая дальняя – черной. На рисунке 3 ниже показан пример оригинального изображения и его карты глубины.

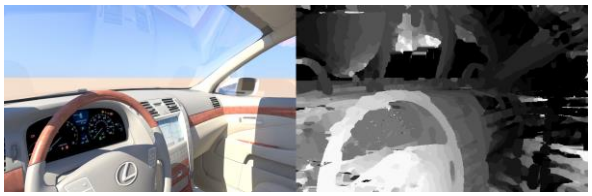


Рис. 3. Оригинальное изображение и его карта глубины.

Формат 2D+Z является дальнейшим развитием концепции представления информации об изображении по компонентам. Широко известно, что и в аналоговом, и в цифровом телевидении изображение формируется из яркости и двух цветовых составляющих. Добавление карты глубины к этим составляющим, характеризующей объемность изображения, является вполне логичным развитием и вполне согласуется с принципами совместимости.

Идея, лежащая в основе построения карты глубины по стереопаре, достаточно очевидна. Для каждой точки на одном изображении выполняется поиск парной ей точки на другом изображении. А по паре соответствующих точек можно определить координаты их прообраза в трехмерном пространстве. Имея же трехмерные координаты прообраза, глубина вычисляется, как расстояние до плоскости камеры.

Парную точку нужно искать на эпиполярной линии [8]. Соответственно, для упрощения поиска, изображения выравнивают так, чтобы все эпиполярные линии были параллельны сторонам изображения (обычно горизонтальны).

Более того, изображения выравнивают так, чтобы для точки с координатами ($x0$, $y0$) соответствующая ей эпиполярная линия задавалась уравнением $x = x0$, тогда для каждой точки соответствующую ей парную точку нужно искать в той же строчке на изображении со второй камеры. Такой процесс выравнивания изображений называют ректификацией (rectification).

После того как изображения ректифицированы, выполняют поиск соответствующих пар точек. Наиболее простой метод состоит в следующем: для каждого пиксела левой картинки с координатами ($x0$, $y0$) выполняется поиск пиксела на правой картинке. При этом предполагается, что пиксел на правой картинке должен иметь координаты ($x0 – d$, $y0$), где d – величина называемая несоответствие/смещение (disparity). Поиск соответствующего пиксела выполняется путем вычисления максимума функции отклика, в качестве которой может выступать, например, корреляция окрестностей пикселей. В результате получается карта смещений (disparity map).

Значения глубины обратно пропорциональны величине смещения пикселей. Если использовать обозначения с левой половины рисунка выше, то зависимость между disparity и глубиной можно выразить следующим способом:

$$\frac{T-d}{Z-f} = \frac{T}{Z} \rightarrow Z = \frac{fT}{d}.$$

Это один из методов построения карты глубины.

Необходимо отметить, что данный метод при использовании его на автостереоскопическом мониторе предоставляет больше возможностей для зрителя по сравнению с обычной стереоанимацией, так как наблюдатель получает возможность рассмотреть изображение объекта под разными углами, меняя свое местоположение.

5. Составные многовидовые кадры

Автостереоскопический монитор также обладает возможностью демонстрации объекта визуализации с использованием составного кадра, содержащего виды объекта визуализации под различными углами, образующими определенный сектор обзора. Обычно, это девять видов (рис. 4).

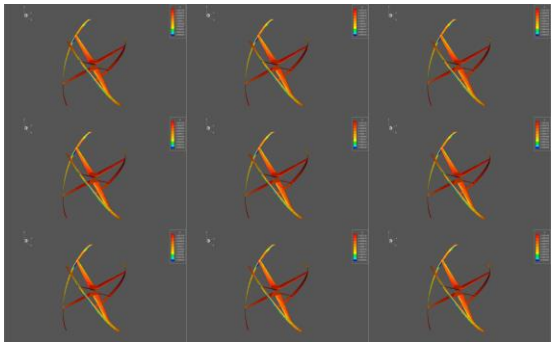


Рис. 4. Составной кадр автостереоскопического монитора.

При этом эти девять видов образуют восемь стереопар ([1|2], [2|3], [3|4], …, [8|9]), и наблюдатель может обозревать только одну из стереопар в зависимости от его положения в том или ином угловом секторе наблюдения. Перемещаясь из сектора в сектор, наблюдатель получает объемную информацию об объекте, используя все девять ракурсов, то есть, как бы оглядывая объект визуализации с разных сторон. Заметим при этом, что фактически происходит «облет камеры» вокруг объекта, и, следовательно, тот же метод уменьшения необходимого для визуализации числа монокадров вполне применим.

Действительно, в случае с величиной угловой стереобазы в один градус и с последовательным углом смещения вокруг оси *OZ*, мы получаем, что первый кадр состоит из последовательности ракурсов с последовательным изменением угла облета камеры:

1	2	3
4	5	6
7	8	9

второй кадр будет иметь вид:

2	3	4
5	6	7
8	9	10

соответственно, третий кадр будет иметь вид:

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

3	4	5
6	7	8
9	10	11

и так далее. Последний кадр, имеющий вид:

360	1	2
3	4	5
6	7	8

позволит зациклить последовательность ракурсов, и имея всего 360 монокадров, можно получить сколь угодно долгое вращение объекта визуализации в многоракурсном представлении на экране автостереоскопического монитора.

Необходимо также отметить, что данный вид представления объекта визуализации в объемном виде на экране автостереоскопического монитора обеспечивает максимальное качество по сравнению с обычным стереоскопическим представлением и даже методом с использованием карт глубины, позволяя рассматривать объект в достаточно широком секторе обзора.

Данный метод визуализации был успешно реализован при моделировании работы узла лопастей сложной формы в энергетической установке [4].

6. Заключение

Данная работа представляет часть результатов общего проекта организации стереоанимированного представления результатов численного моделирования задач математической физики. Реализация подобных анимаций является заключительным этапом вычислительной технологии получения и представления результатов моделирования сложных физических процессов и работы сложных технических установок. Представлены результаты анимаций для классической пассивной установки и для автостереоскопических мониторов. В последнем случае использовались такие методы, как построение карт глубины и многовидовых кадров.

.

7. Литература

[1] Синтез фотореалистичных трехмерных изображений в современных системах презентаций / Андреев С.В. [и др.] // Программные продукты и системы. 2007. № 3. С. 37-40.

[2] Организация стереопредставлений в задачах синтеза фотореалистичных изображений и научной визуализации /Андреев С.В. [и др.] // Препринты ИПМ им. М.В.Келдыша. 2010. № 61. 14 с. URL: http://library.keldysh.ru/preprint.asp?id=2010-61.

[3] Андреев С., Филина А. Применение стереоизображений для визуализации результатов научных вычислений / Научная визуализация. 2012. Т.4. № 1. С.12-21.

[4] Моделирование и визуализация работы узла лопастей сложной формы в энергетической установке / Андреев С.В. [и др.] // Научная визуализация. 2015. Т.7. № 4. С.1-12.

[5] Программно-аппаратный комплекс 3D презентаций на основе виртуальной студии и виртуального окружения / Ванданов В.Г. [и др.] // Труды Первой международной конференции «Трехмерная визуализация научной, технической и социальной реальности. Кластерные технологии моделирования». Ижевск. 2009. С.73-77.

[6] Меженин А.В., Тозик В.Т. 3D Визуализация с использованием эффекта стереоизображений // Труды Второй международной конференции «Трехмерная визуализация научной, технической и социальной реальности. Кластерные технологии моделирования». Ижевск. 2010.

[7] Михайлюк М.В., Хураськин И.А. Синтез стереоизображения для систем виртуальной реальности с использованием оптической трекинговой системы / Программные продукты и системы. 2006. № 3. С.10-14.

[8] Калиниченко А.В., Свешникова Н.В., Юрин Д.В. Эпиполярная геометрия и оценка ее достоверности по результатам восстановления трехмерной сцены алгоритмами факторизации // Труды 16-й Международной конференции по компьютерной графике и зрению ГрафиКон'2006.

Об авторах

Андреев Сергей Валерьевич, научный сотрудник Института прикладной математики им. М.В. Келдыша РАН. Его e-mail esa@keldysh.ru.

Бондарев Александр Евгеньевич, к.ф.-м.н., старший научный сотрудник Института прикладной математики им. М.В. Келдыша РАН. Его e-mail bond@keldysh.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Построение отображения текстовой информации в стереопрезентациях

Андреев С.В.[1], Бондарева Н.А.[2]
esa@keldysh.ru|niki.99@mail.ru
[1]ИПМ им. М.В. Келдыша РАН, Москва, Россия;
[2] Российский технологический университет (МИРЭА), Москва, Россия

Данная работа посвящена практическим вопросам построения стереопредставлений текстов и формул в стереопрезентациях, предназначенных для отображения результатов численного моделирования. Задача построения стереоизображений текстов и формул является структурной подзадачей общего исследования, посвященного разработке способов и алгоритмов построения стереопрезентаций результатов научных исследований на современных стереоустановках как активного, так и пассивного типов. Рассматриваются различные подходы, обеспечивающие наибольшую выразительность и стереоэффект. Представлены иллюстрации практического применения.

Ключевые слова: *стереопрезентация, текстовая информация, выразительность, стереоэффект.*

Constructing a representation of textual information in stereo presentations

Andreev S.V.[1], Bondareva N.A.[2]
esa@keldysh.ru|niki.99@mail.ru
[1]Keldysh Institute of Applied Mathematics RAS, Moscow, Russia;
[2]Russian Technological University MIREA, Moscow, Russia

This paper is devoted to practical questions of constructing stereo images of texts and formulas in stereo presentations intended for displaying the results of numerical simulation. The task of constructing stereo images of texts and formulas is a structural subtask of a general study devoted to the development of methods and algorithms for constructing stereo presentations of scientific research results on modern stereo devices of both active and passive types. Various approaches are considered that provide the greatest expressiveness and stereo effect. Illustrations of practical application are presented.

Keywords: *stereo presentation, textual information, expressiveness, stereo effect.*

1. Введение

Появление, развитие и совершенствование современных стереоустановок, как пассивного, так и активного типов, сделало возможным построение стереоанимационных представлений результатов научных исследований Значимость и актуальность новых возможностей, открываемых использованием стереоустановок, трудно переоценить. Представление результатов численного моделирования, отображающих протекание во времени сложных физических процессов или показывающих в трехмерном представлении работу сложных технических установок, позволяет очень многое. Во-первых, трехмерное представление развития процесса во времени помогает верифицировать вычислительную модель и применяемые алгоритмы. Во-вторых, обеспечивает наблюдателю полноценное понимание моделируемого явления. И, наконец, популяризирует проведенное исследование и его результаты для социума, в том числе и для лиц, принимающих решения. Все перечисленное делает теоретические и практические разработки в области построения стереопредставлений весьма важными.

По мере появления и развития стереоустановок стало появляться все больше работ, посвященных вопросам построения стереоизображений. К таким работам следует отнести работы [1,6-8]. Поначалу большинство исследований в области стереоизображений относились к задачам построения виртуальной реальности, симуляторов и тренажеров, а также обучающим системам. Однако позднее стал весьма актуальным вопрос построения презентационных комплексов, которые в том числе позволяли демонстрировать в стереоскопическом режиме результаты научных исследований. К таким работам можно отнести работы [2-4]. В работе [3] представлены результаты отображения взрыва Сверхновой в стереорежиме. Работа [4] посвящена созданию вычислительной технологии моделирования работы трехмерного узла лопастей энергоустановки при обтекании потоком вязкого сжимаемого теплопроводного газа. Построение анимированных стереопредставлений результатов расчетов являлось одной из важных компонент разработанной технологии.

Однако при построении стереоанимации возникает достаточно важная частная структурная подзадача. Стандартное представление научного доклада, как правило, включает в себя не только визуализацию сложной геометрии или полей физических величин, но и пояснения, обычно представляемые текстом и формулами. Если для геометрий и физических величин вопросы построения стереопредставлений разработаны достаточно, то задачи построения стереоотображений текстов и формул с достаточной выразительностью и необходимым стереоэффектом остаются открытыми. Вопросам разработки конкретных практических подходов в этой области посвящено довольно мало работ. Здесь можно упомянуть только работу [10], описывающую проведенное в Японии исследование воспринимаемости стереоизображений шрифтов на экранах стереоскопических мобильных устройств. Однако это исследование носило более социальный характер, нежели давало рецепты построения стереоизображений текстов и формул.

Данная работа призвана заполнить пробел в этой области и представить практический опыт построения стереопредставлений текстовых надписей и формул в стереоанимациях.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Данная работа относится к общей тематике построения стереоизображений и стереоанимаций результатов математического моделирования сложных технических объектов и физических процессов в сплошных средах и представляет собой практический опыт построения стереоанимаций для конкретных устройств, которыми располагает ИПМ им. М.В. Келдыша РАН. Первый тип устройства представляет собой 3D проекционную стереосистему для показа стереопрезентаций, обучающих приложений, графики и фильмов. Данная система является классической стереографической системой с использованием двух проекторов, экрана и очков линейной поляризации.

Второй тип устройства представляет собой автостереоскопический монитор Dimenco DM654MAS. Принцип работы автостереоскопического монитора – это использование параллаксных перегородок или линз Френеля, установленных за защитным стеклом экрана. Когда голова наблюдателя находится в определенной позиции перед автостереоскопическим монитором, его правый и левый глаз получают различные изображения (стереопара). Таким образом, создается убедительная иллюзия 3D глубины.

Исследования по построению стереопредставлений текстов и формул на сегодняшний день проведены только для систем пассивного типа, то есть для классической 3D проекционной стереосистемы.

2. Основные понятия построения стереоизображений

При создании стереопар, требуемых для кадров стереофильма, используется либо линейное, либо угловое смещение камеры, то есть линейная или угловая *стереобаза*. В большинстве случаев линейная стереобаза применяется при визуализации объектов сцены со значительно отдаленным задним фоном. При необходимости создания стереофильма для визуализации результатов расчетов модели объекта, то есть набора последовательно меняющихся кадров, линейное смещение применяется довольно редко, так как чаще всего необходимо получить представление об объекте визуализации со всех сторон. При линейном же смещении область видимой части объекта ограничена.

В большинстве случаев результатом расчетов является достаточно инкапсулированный объект, не привязанный к окружающей среде, такой как деталь самолета, автомобиля, то есть не привязанный к местности. И в первую очередь зрителям важно осмотреть этот объект со всех сторон. В таких случаях применяется угловое смещение, то есть угловая стереобаза. Чаще всего определяется центр пространственных координат, находящийся внутри объекта, и происходит угловое смещение камеры.

При этом для показа объекта визуализации со всех сторон достаточно зафиксировать расстояние от камеры до этого центра координат и произвести полный оборот в 360 градусов в горизонтальной плоскости вокруг оси *OZ*, последовательно меняя только угол. В результате мы получим стереоэффект «висящего» в воздухе объекта визуализации, центр координат которого находится на расстоянии от глаз зрителя до плоскости экрана.

При оценке величины стереобазы необходимо учитывать такое явление как *параллакс*.

Параллакс (греч. Parallaxis – уклонение) – это кажущееся смещение рассматриваемого объекта, вызванное изменением точки наблюдения. Если смотреть на объект одним глазом, затем другим, то можно увидеть, что объекты, расположенные на расстоянии, смещаются относительно друг друга. Это смещение и называется параллаксом. Благодаря этому параллаксу человек получает представление о взаимном расположении объектов и объемности сцены по отношению к плоскости экрана (см. [2]).

Существуют три вида параллакса: позитивный, нулевой и негативный.

Когда объект смещается вправо в левой части стереопары, и наоборот, смещается влево в правой части стереопары, бинокулярное зрение оценивает этот объект, как находящийся позади плоскости экрана. Такое явление получило название *положительный параллакс*.

Когда изображение объекта на правой и левой частях стереопары накладываются одно на другое (то есть смещение близко к нулю), бинокулярное зрение подразумевает, что объект находится на плоскости экрана. Такой эффект получил название *нулевой параллакс*.

Когда объект смещается влево в левой части изображения и смещается вправо в правой части изображения, бинокулярное зрение расценивает этот объект, как находящийся между зрителем и плоскостью экрана. Данный эффект получил название *отрицательный параллакс*.

Соответственно, в зависимости от того, каким является параллакс (положительным, нулевым или отрицательным), точка конвергенции (точка пересечения осей взгляда глаз наблюдателя) располагается либо за плоскостью экрана, либо на плоскости экрана, либо перед плоскостью экрана. При создании конкретного стереоизображения оценки стереобазы и параллакса переднего плана можно получить с помощью достаточно известных соотношений стереофотографии, что было использовано в процессе практических тестов. Эти соотношения представлены ниже для случая линейной стереобазы и для случая угловой стереобазы

$$Base = \frac{Parallax_{fore} * (L_{fore} / f - L_{fore} / L)}{M}$$

$$AngularBase_{fore} = 2 * arctg\left(\frac{Parallax_{fore} * (L_{fore} / f - L_{fore} / L)}{2 * M * (P - L_{fore})}\right).$$

3. Построение стереоизображений текста

Построения стереоизображений текстов и формул в данной работе проводились для классической 3D проекционной системы, то есть для установки пассивного типа, предполагающей использование очков линейной поляризации. В качестве основы построения стереоизображения была выбрана линейная стереобаза. Таким образом, построение предполагало создание левого и правого кадра для каждой текстовой надписи.

При построении основной целью было следование ряду достаточно простых принципов, предполагавшихся изначально. К этим принципам относятся следующие:

- Надпись должна быть трехмерной, т.е. иметь объем, согласно выводам работы [2], где было указано, что наилучший стереоэффект получается для объемных фигур. В силу этого рекомендовалось все линии представлять в объемном виде. Следовательно, для построения надписей и формул необходимо использовать объемные шрифты или эффекты, создающие объем.

- Надписи должны четко читаться, в силу чего рекомендуется использовать «строгие» шрифты.

- Надписи и формулы должны четко выделяться на фоне кадра, следовательно, фон не может быть слишком ярким.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

- В стереоанимационном фильме, представляющем результаты научных исследований, основным является представление результатов исследований. Пояснения в виде текстовых надписей и математических формул играют вторичную роль. Поэтому кадр, содержащий сопроводительную информацию в виде надписей и формул, не должен быть ярче кадров, несущих основную информацию.

Для построения тестовых вариантов надписей была выбрана хорошо известная программа Microsoft Power Point, обладающая достаточным набором встроенных шрифтов, а также широким диапазоном опций и спецэффектов, позволяющих обеспечить объемность надписей и проведение действий с ними. Для вариации стереобазы использовалась сетка, которая позволяла регулировать расстояние сдвига по стереобазе для организации левого и правого кадра.

Подобный подход позволил провести ряд экспериментов по построению наиболее подходящих кадров с текстовыми надписями и формулами. Было построено около 100 вариантов различных представлений надписей и формул, где варьировались:

- лежащие в основе надписей шрифты;
- рельефность шрифтов;
- поворот надписи целиком на определенный малый угол для обеспечения трехмерности представления;
- расстояние сдвига по стереобазе для получения левого и правого кадров;
- цвет шрифта;
- цвет фона.

Некоторые из построенных вариантов представлены на рисунках ниже. На рисунке 1 представлен вариант построения надписи с помощью рельефного шрифта Arial Black с поворотом всей надписи. Рисунок 2 представляет аналогичное построение с помощью рельефного шрифта Cambria с поворотом всей надписи. Левый и правый кадры отмечены на рисунках ниже соответственно синим и зеленым маркерами.

Рис. 1. Пример построения надписи на основе рельефного шрифта Arial Black с поворотом надписи.

Рис. 2. Пример построения надписи на основе рельефного шрифта Cambria с поворотом надписи.

Ниже на рисунках 3 и 4 представлены сравнительные результаты для рельефного шрифта Arial Narrow с поворотом надписи на больший угол (рис. 3) и на меньший угол (рис. 4).

Рис. 3. Построение надписи на основе рельефного шрифта Arial Narrow с поворотом надписи на больший угол.

Рис. 4. Построение надписи на основе рельефного шрифта Arial Narrow с поворотом надписи на меньший угол.

Для большинства тестовых надписей проводилась вариация сдвига по стереобазе. Общая методика построения тестового стереоизображения для выбранной надписи может быть описана следующим образом:

- выбор основного шрифта;
- выбор цвета шрифта и фона;
- выбор степени рельефности;
- выбор поворота надписи целиком.

Далее из созданных образцов отбирались те, что в наибольшей степени соответствовали принципам, изложенным в начале раздела. Для отобранных образцов проводилось исследование вариации по расстоянию сдвига по сетке, то есть, по линейной стереобазе. Образцы стереокадров проверялись на стереоустановке с целью оценки стереоэффекта.

Следует заметить, что в силу отсутствия возможности организации полноценной экспертной оценки, отбор тестовых вариантов проходил на основе изначально выбранных принципов и оценки зрительного восприятия удаленности тестовой надписи от плоскости экрана. При сдвиге тестовой надписи на 3.5 деления сетки вправо и влево зрители, находящиеся в 4-х метрах от экрана, воспринимали надпись, как удаленную от плоскости экрана приблизительно на 1 метр, что было признано достаточным.

В итоге было решено, что наиболее оптимальным способом представления надписей является представление на основе рельефного шрифта Arial Narrow на светлом фоне, затененном в нижней части, без поворота надписи. Пример такого представления надписи приведен на рисунке 5.

Рис. 5. Построение надписи на основе рельефного шрифта Arial Narrow с поворотом надписи на меньший угол.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Результаты экспериментов нашли практическое применение. При разработке вычислительной технологии, обеспечивающей построение оптимальной с точки зрения силовых нагрузок формы узла лопастей энергоустановок [4,5,9] предполагался полный цикл моделирования энергоустановки от построения исходной CAD модели до организации визуального представления результатов в стереорежиме. Результаты экспериментов были использованы при построении стереофильмов в разрабатываемой вычислительной технологии [9]. Все титры, надписи и формулы в данном стереофильме были выполнены на основе экспериментов по представлению надписей и формул, представленных в настоящей работе. Версия стереофильма доступна по ссылке [9] и может быть просмотрена на классической 3D проекционной стереоустановке пассивного типа с очками линейной поляризации.

На рисунке 6 представлен кадр из данного стереофильма [9], показывающий вращение узла лопастей энергоустановки. На рисунке 7 представлен кадр из стереофильма [9], содержащий надпись, построенную в результате описанных экспериментов.

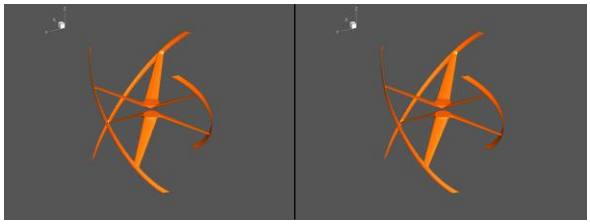


Рис. 6. Кадр из стереофильма [9], показывающий вращение узла лопастей энергоустановки.

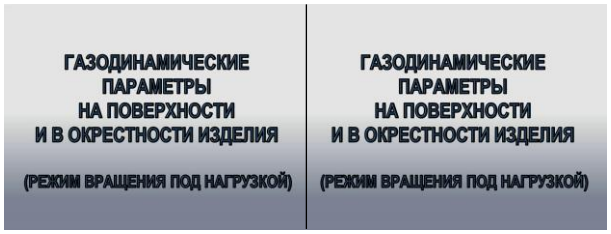


Рис. 7. Кадр из стереофильма [9] с типичным представлением пояснительного текста.

Данная работа не является завершенной. Предполагается проведение дальнейших экспериментов на стереоустановке пассивного типа, но уже для угловой стереобазы. Также предполагается проведение исследований по разработке построения надписей и формул для установок активного типа, не требующих применения очков – автостереоскопических мониторов. Это потребует построения карты глубины для тестовой надписи или построения нескольких видов надписи с разных углов зрения.

4. Заключение

Данная работа представляет опыт построения текстовой информации, а именно, текстов и формул, в стереопрезентациях для стереоустановки пассивного типа. Работа является частью общего проекта, посвященного построению стереопрезентаций для результатов решения задач математического моделирования. Представлены результаты использования линейной стереобазы. В дальнейшем планируется исследовать возможности применения угловой стереобазы и проведение аналогичного исследования для установок активного типа – автостереоскопических мониторов.

5. Литература

[1] Синтез фотореалистичных трехмерных изображений в современных системах презентаций / Андреев С.В. [и др.] // Программные продукты и системы. 2007. № 3. С. 37-40.

[2] Организация стереопредставлений в задачах синтеза фотореалистичных изображений и научной визуализации /Андреев С.В. [и др.] // Препринты ИПМ им. М.В.Келдыша. 2010. № 61. 14 с. URL: http://library.keldysh.ru/preprint.asp?id=2010-61м.

[3] Андреев С., Филина А. Применение стереоизображений для визуализации результатов научных вычислений / Научная визуализация. 2012. Т.4. № 1. С.12-21.

[4] Моделирование и визуализация работы узла лопастей сложной формы в энергетической установке / Андреев С.В. [и др.] // Научная визуализация. 2015. Т.7. № 4. С.1-12.

[5] Моделирование и визуализация работы энергетической установки сложной формы в режиме стереоанимации / Андреев С.В. [и др.] // Труды 25-й Международной Конференции по Компьютерной Графике и Зрению Гра-фиКон'2015, Протвино, Россия, 22-25 сентября 2015 г., с. 183-187.

[6] Программно-аппаратный комплекс 3D презентаций на основе виртуальной студии и виртуального окружения / Ванданов В.Г. [и др.] // Труды Первой международной конференции «Трехмерная визуализация научной, технической и социальной реальности. Кластерные технологии моделирования». Ижевск. 2009. С.73-77.

[7] Меженин А.В., Тозик В.Т. 3D Визуализация с использованием эффекта стереоизображений // Труды Второй международной конференции «Трехмерная визуализация научной, технической и социальной реальности. Кластерные технологии моделирования». Ижевск. 2010.

[8] Михайлюк М.В., Хураськин И.А. Синтез стереоизображения для систем виртуальной реальности с использованием оптической трекинговой системы / Программные продукты и системы. 2006. № 3. С.10-14.

[9] Стереофильм «Моделирование и визуализация работы энергетической установки сложной формы в режиме стереоанимации» https://yadi.sk/i/3CcYhlMB3TCH4r (дата обращения 01.03.2018).

[10] Visibility Experiment and Evaluation of 3D Character Representation on Mobile Displays / Hiromu Ishio [et al.] // C. Stephanidis (Ed.): Posters, Part II, HCII 2011, CCIS 174, pp. 46–51, 2011.

Об авторах

Андреев Сергей Валерьевич, научный сотрудник Института прикладной математики им. М.В. Келдыша РАН. Его e-mail esa@keldysh.ru.

Бондарева Николь Александровна, студентка Российского технологического университета (МИРЭА). Ее e-mail niki.99@mail.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Формирование пространства признаков для определения границ зерен структурных компонентов сплава по полихроматическому растровому изображению его микроструктуры

К.Ю. Андросов[1], А.Н. Горбунов[1], С.В. Давыдов[1]
androkirl@yandex.ru|angor47@yandex.ru|davidov69@tu-bryansk.ru

[1]Брянский государственный технический университет, Брянск, Россия

Поиск дефектов в сплаве может быть значительно облегчен, если имеется возможность определения формы зерен компонентов сплава путем анализа его микроструктуры по имеющемуся полихроматическому растровому изображению. В данной работе рассматриваются вопросы определения необходимого набора параметров для формирования пространства признаков по определению границ зерен, алгоритма поиска этих границ и выполнения операции по их сегментации.

Ключевые слова: *сплав, границы зерен, микроструктура, полихроматическое растровое изображение, пространство признаков.*

Formation of the feature space to determine the grain boundaries of the structural components of the alloy by the polychromatic raster image of its microstructure

K.Yu. Androsov[1], A.N. Gorbunov[1], S.V. Davydov[1]
androkirl@yandex.ru|angor47@yandex.ru |davidov69@tu-bryansk.ru

[1]Bryansk state technical University, Bryansk, Russia

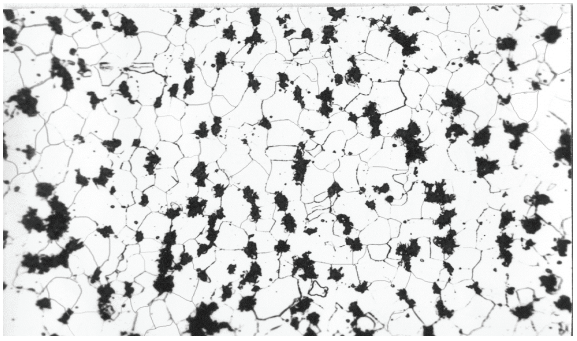
Search of defects in the alloy can be greatly facilitated if there is a possibility of determining the shape of the grains of the alloy components by the analysis of the microstructure according to available polychromatic raster image. In this paper, we consider the issues of determining the necessary set of parameters for the formation of the feature space to determine grain boundaries, the algorithm for finding these boundaries and perform operations on their segmentation.

Keywords: *alloy, grain boundaries, microstructure, polychromatic raster image, feature space.*

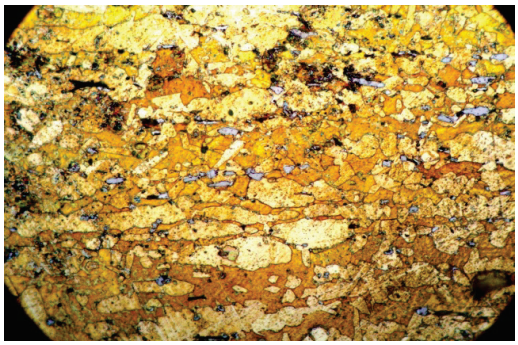
1. Введение

Наличие структурных дефектов в сплаве, таких как газовые поры, неметаллические включения, зоны усадки, микротрещины, можно определить, осуществляя анализ его микроструктуры по имеющемуся полихроматическому растровому изображению путем определения формы зерен фаз и структурных элементов, формирующих структурный состав сплава. Для этого необходимо сформировать пространство признаков, характеризующих зерно по некоторому ряду параметров, определить границы зерна и его конфигурацию. Схожие по конфигурации зерна объединяются в сегменты, формирующие различные области в структуре сплава. Отличия в конфигурации сегментов друг от друга дают возможность определять наличие и место расположения дефектов в структуре сплава.

В зависимости от типа сплава металлографическое изображение является ключевым носителем информации либо в виде черно-белой структуры (рис. 1а), либо в виде цветового изображения структурных компонентов сплава (рис. 1б).



(а)



(б)

Рис. 1. Структурное (а) и цветовое (б) изображение сплава (а – ковкий ферритный чугун, х100, б – бронза, х100).

Приведенные изображения получены при фотографировании структуры сплавов с помощью стационарного металлографического инвертированного микроскопа Leica DMI3000M, который позволяет рассматривать непрозрачные тела в отраженном свете. Таким образом, возникает проблема выбора алгоритма сегментации в равной степени применимого для обоих типов изображений. В связи с этим необходимо построить объединенное структурно-цветовое пространство, являющееся объединением двух более простых пространств – цветовых и структурных признаков [1].

2. Построение структурного пространства признаков

Исходные металлографические изображения, с которыми предстоит работать, представлены в формате BMP, и цветовые характеристики каждого пикселя представлены в обычном RGB пространстве. Для

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

качественного описания характеристик цвета необходимо учитывать не только характеристики основных цветов (Red, Green, Blue – красный, зелёный, синий), но и характеристики яркости (светлоты) и цветности (тона и насыщенности). Поэтому в работе будет использовано цветовое пространство CIE Lab. Преобразование координат из пространства RGB в CIE Lab организуется путем цепочки последовательных преобразований цветовых пространств: из RGB в CIEXYZ, из CIEXYZ в CIE Lab [2].

Для перевода из RGB в CIEXYZ воспользуемся следующими уравнениями:

$$\left.\begin{array}{l} X = 0.412453 * r + 0.357580 * g + 0.180423 * b, \\ Y = 0.212671 * r + 0.715160 * g + 0.072169 * b, \\ Z = 0.019334 * r + 0.119193 * g + 0.950227 * b \end{array}\right\} \quad (1),$$

где r, g, b координаты характеристик для отдельно взятого пикселя в цветовом пространстве RGB, расчет проводится для каждого пикселя изображения. Далее рассчитываются непосредственно яркостная – L и хроматические составляющие – a, b по формулам:

$$L = \begin{cases} \sqrt[3]{116 * Y/Y_n} - 16, & Y/Y_n > 0.008856; \\ 903.3 * \dfrac{Y}{Y_n}, & \dfrac{Y}{Y_n} \le 0.008856; \end{cases} \quad (2)$$

$$a = \begin{cases} 500 * \left(\sqrt[3]{X/X_n} - \sqrt[3]{Y/Y_n}\right), & \frac{Y}{Y_n} > 0.008856 \wedge \frac{X}{X_n} > 0.008856; \\ 500 * [f(X/X_n) - f(Y/Y_n)], & \frac{Y}{Y_n} \le 0.008856 \vee \frac{Z}{Z_n} \le 0.008856; \end{cases} \quad (3)$$

$$b = \begin{cases} 200 * \left(\sqrt[3]{Y/Y_n} - \sqrt[3]{Z/Z_n}\right), & \frac{Y}{Y_n} > 0.008856 \wedge \frac{X}{X_n} > 0.008856; \\ 200 * [f(Y/Y_n) - f(Z/Z_n)], & \frac{Y}{Y_n} \le 0.008856 \vee \frac{Z}{Z_n} \le 0.008856; \end{cases} \quad (4)$$

$$f\left(\frac{Y}{Y_n}\right) = 7.87 * \frac{Y}{Y_n} - \frac{16}{116}; \quad (5)$$

$$f\left(\frac{Y}{Y_n}\right) = 7.87 * \frac{Y}{Y_n} - \frac{16}{116}; \quad (6)$$

$$f\left(\frac{Z}{Z_n}\right) = 7.87 * \frac{Z}{Z_n} - \frac{16}{116}, \quad (7)$$

где: X_n, Y_n, Z_n показатели трех основных цветов в CIEXYZ пространстве, взятые как белый эталон. В работе будет использоваться эталон белого цвета D_{65} [2].

Для построения структурного пространства признаков, определяющих границы зерен сплава, будем использовать оценку локальной изменчивости изображения по нескольким направлениям для разных локальных масштабов, вычисляемую для каждой точки изображения на основе локальной вариации v и локального изменения градиента g вдоль выбранных направлений и масштабов. В качестве анализируемой информации будут использованы значения яркости L, вычисляемые по формуле (2). Рассмотрим непосредственно вычисление признаков принадлежности точки изображения (пикселя) к соответствующей структуре.

Сначала рассмотрим вычисление признака принадлежности пикселя к соответствующей структуре для единичного масштаба (расстояние между двумя соседними пикселями). Рассмотрим отрезок длиной $2L$, центрированный в точке (m, n) под углом $\alpha = k\pi/K$, где $k \in \overline{0, K-1}$. На этом отрезке равномерно отложим $2L+1$ точек, и каждой из этих точек $i \in \overline{-L, L}$ присвоим значение x_i, равное значению ближайшей точки изображения.

Пусть $d_i = x_{i+1} - x_i$ есть разница между двумя соседними точками разбиения на отрезке, т.е. является градиентом в точке i по направлению отрезка. Далее вычисляются взвешенные суммы подъемов v^+ и спусков v^- яркости вдоль направления отрезка по формулам (8), (9). Для этого в качестве весовой функции для каждого значения d_i будет рассчитываться функция q_i, вычисляемая по формуле (10).

$$v^+ = \sum_{i=-L}^{L-1} q_i d_i, \text{если} d_i \ge 0; \quad (8)$$

$$v^+ = \sum_{i=-L}^{L-1} q_i(-d_i), \text{если} d_i < 0; \quad (9)$$

$$q_i = \cos\left(\frac{i\pi}{2L}\right) / \sum_{i=-L}^{L} \cos\left(\frac{i\pi}{2L}\right). \quad (10)$$

Для подавления влияния контурных перепадов на значение локальной вариации по направлению из двух оценок взвешенных значений подъемов и спусков яркости выбирается наименьшее:

$$v_k = \min(v^+, v^-). \quad (11)$$

Значение изменения локального градиента g_k определяется как число перемен знака градиента d_i для $i \in \overline{-L, L}$ на выбранном отрезке, при условии что $d_i > \varepsilon$, где ε – выбранный порог чувствительности.

Произведение изменения локального градиента на значение локальной вариации по направлению k дает локальное значение структурного признака пикселя по выбранному направлению:

$$t'_k = g_k v_k. \quad (12)$$

Для выравнивания значений признаков высоко и мало контрастных структур можно использовать функцию гиперболического тангенса

$$t_k = \tanh(t'_k), \quad (13)$$

значение t_k будем считать структурным признаком по направлению k в соответствующей точке изображения. Такие значения рассчитываются для каждого направления $k \in \overline{0, K-1}$, а так как структура, очевидно, является масштабно зависимой характеристикой, то и для всех масштабных уровней S.

Расчет структурных признаков для масштабов различного уровня необходим с целью искусственного выделения микро- и макроструктурных компонентов сплава. Для решения указанной задачи требуется установление структурных признаков компонентов сплава, которые соответствуют дифференцированным пространственным частотам. Для расчета значений структурных признаков для разных масштабов можно воспользоваться формулами (4)-(9), для этого необходимо взять отсчеты с другим шагом, т.е. при сохранении числа точек, равном $2L-1$, необходимо соответствующим образом изменить размер самого отрезка. Теоретически, увеличив размер отрезка в 2^n раз, можно учесть в 2^n раз больше низких частот, влияние которых можно ослабить применением сглаживающих фильтров.

Таким образом, по формулам (8)-(13) для каждой точки изображения по всем выбранным направлениям $k \epsilon \overline{0, K-1}$ и масштабам $1, 2^1, \ldots, 2^{S-1}$ можно вычислить значения структурного признака, где K и S общее число направлений и масштабов соответственно, т.е. структурный признак будет иметь значение:

$$t'^S_k = v^S_k g^S_k; \quad (14)$$

$$t^S_k = \tanh(\alpha v^S_k g^S_k), \quad (15)$$

где v^S_k и g^S_k значения локальной вариации и изменения локального градиента масштаба s по направлению k, а общее число рассчитываемых структурных признаков для точки изображения будет равным $K * S$.

Если в качестве K и S использовать значения $S = 4$ и $K = 4$, то для каждой точки будет рассчитываться 16 значений структурных признаков. Дальнейшее увеличение числа масштабов и направлений приведет к неоправданному увеличению вычислительной нагрузки при реализации.

Объединив полученные значения структурных признаков и признаков цвета для точки изображения,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

получим вектор признаков изображения, а в совокупности эти векторы образуют объединенное структурно-цветовое пространство признаков.

3. Метрика в объединенном структурно-цветовом пространстве признаков

Важнейшим аспектом задачи сегментации является вопрос о близости соседних элементов, т. е вопрос о метрике в пространстве признаков. Поскольку используемое в работе пространство является объединением пространства цветовых и структурных признаков, то для каждого из них будет выбрана своя метрика, которая затем объединяется в единую структурно-цветовую метрику, как это сделано в работе [3].

Структурная метрика будет задаваться формулой:

$$\rho(x,y)=\sum_{s=1}^{S}\omega^s\sum_{k=0}^{K-1}|t_k^s(x)-t_k^s(y)|, \tag{16}$$

где x,y – две точки структурного пространства, $t_k^s(x)$ и $t_k^s(y)$ значения структурных признаков в этих точках по направлениям и масштабам k,s, а ω^s весовой коэффициент соответствующего масштаба S.

В качестве цветовой метрики будет использована метрика CIEDE2000. Данная метрика предполагает использование цветового пространства CIE Lab.

Рассмотрим подробнее эту метрику. Возьмем две точки $L_1^*a_1^*b_1^*$ и $L_2^*a_2^*b_2^*$ (рис. 2) в пространстве цветовых признаков, описанном выше. Цветовое различие (метрика) между этими двумя точками обозначим как:

$$\Delta E_{00}(L_1^*a_1^*b_1^*, L_2^*a_2^*b_2^*)=\Delta E_{00}^{12}=\Delta E_{00}. \tag{17}$$

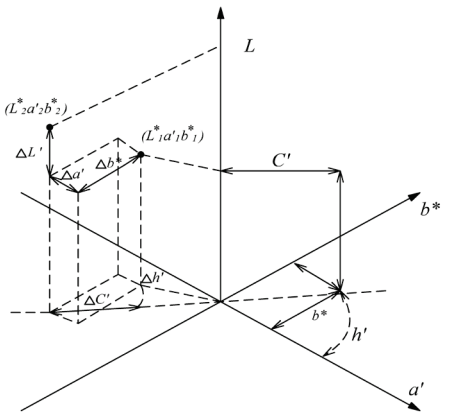


Рис. 2. Геометрическое представление значений, вычисляемых в CIEDE2000.

Вычисление ΔE_{00} можно разделить на 3 основных шага:

1. Вычисление значений C_i' и h_i', где модифицированный оттенок h_i' вычисляется по формуле (23) и представляет собой угловую координату точки (a_i', b_i^*) в двумерной плоскости $a'-b^*$, взятую от положительной оси a' (рис. 2). Так как большинство языков программирования при вычисления арктангенса возвращают значение в радианах, а для дальнейших расчетом нам необходимы значения в градусах, то в приведенных формулах учитывается перевод значений.

$$C_{i,ab}^*=\sqrt{(a_i^*)^2+(b_i^*)^2}\quad i=1,2; \tag{18}$$

$$\bar{C}_{i,ab}^*=\frac{C_{1,ab}^*+C_{2,ab}^*}{2}; \tag{19}$$

$$G=0.5\left(1-\sqrt{\frac{\bar{C}_{ab}^{*7}}{\bar{C}_{ab}^{*7}+25^7}}\right); \tag{20}$$

$$a_i'=(1+G)a_i^*,\qquad i=1,2; \tag{21}$$

$$C_i'=\sqrt{(a_i')^2+(b_i^*)^2},\qquad i=1,2; \tag{22}$$

$$h_i'=\begin{cases}0, & a_i'=b_i^*=0\\ tan^{-1}(b_i^*,a_i'), & a_i'\neq 0, b_i^*\neq 0\end{cases}\quad i=1,2. \tag{23}$$

2. Далее вычисляются значения $\Delta L', \Delta C', \Delta H'$ по формулам (24), (25), (27) соответственно. Геометрическая интерпретация $\Delta L', \Delta C'$ представлена на рис. 2.

$$\Delta L'=L_2^*-L_1^*; \tag{24}$$

$$\Delta C'=C_2'-C_1'; \tag{25}$$

$$\Delta h'=\begin{cases}0 & C_1'C_2'=0\\ h_2'-h_1' & C_1'C_2'=0;|h_2'-h_1'|\le 180^\circ\\ h_2'-h_1'-360 & C_1'C_2'=0;(h_2'-h_1')>180^\circ\\ h_2'-h_1'+360 & C_1'C_2'=0;(h_2'-h_1')<-180^\circ\end{cases} \tag{26}$$

$$\Delta H'=2\sqrt{C_1'C_2'}\sin\left(\frac{\Delta h'}{2}\right). \tag{27}$$

3. На третьем шаге вычисляется непосредственно значение цветового различия ΔE_{00}. Для этого вычисляются поправки на яркость S_L, тон S_H, и насыщенность S_C по формулам (34)-(36). Значение цветового различия двух точек $(L_1^*a_1^*b_1^*)$ и $(L_2^*a_2^*b_2^*)$ будет вычисляться по формуле (38), где k_L, k_C, k_H весовые коэффициенты, позволяющие усилить или ослабить влияние нужных цветовых компонент на значение цветового отличия.

$$\bar{h}'=\begin{cases}h_2'+h_1' & C_1'C_2'=0\\ \frac{(h_2'+h_1')}{2} & C_1'C_2'\neq 0;|h_1'-h_2'|\le 180^\circ;\\ h_2'+h_1'+360 & C_1'C_2'\neq 0;|h_1'-h_2'|>180^\circ;h_1'+h_2'<360^\circ\\ h_2'-h_1'-360 & C_1'C_2'\neq 0;|h_1'-h_2'|>180^\circ;h_1'+h_2'\ge 360^\circ\end{cases} \tag{28}$$

$$\overline{L'}=\frac{L_1^*+L_2^*}{2}; \tag{29}$$

$$\overline{C'}=\frac{C_1^*+C_2^*}{2}; \tag{30}$$

$$T=1-0.17\cos\left(\bar{h}'-30^\circ\right)+0.24\cos\left(2\bar{h}'\right)+$$
$$+0.32\cos\left(3\bar{h}'+6^\circ\right)-0.20\left(4\bar{h}'+63^\circ\right); \tag{31}$$

$$\Delta\theta=30exp\left\{-\left[\frac{\bar{h}'+275^\circ}{25}\right]^2\right\}; \tag{32}$$

$$R_c=2\sqrt{\frac{\overline{C'}^7}{\overline{C'}^7+25^7}}; \tag{33}$$

$$S_L=1+\frac{0.015(\overline{L'}-50)^2}{\sqrt{20+(\overline{L'}-50)^2}}; \tag{34}$$

$$S_C=1+0.045\overline{C'}; \tag{35}$$

$$S_H=1+0.015\overline{C'}T; \tag{36}$$

$$R_T=-\sin(2\Delta\theta)\,R_C; \tag{37}$$

$$\Delta E_{00}=\sqrt{\left(\frac{\Delta L'}{k_LS_L}\right)^2+\left(\frac{\Delta C'}{k_CS_C}\right)^2+\left(\frac{\Delta H'}{k_HS_H}\right)^2+R_T\left(\frac{\Delta C'}{k_CS_C}\right)\left(\frac{\Delta H'}{k_HS_H}\right)}. \tag{38}$$

Таким образом, получив формулы структурной и цветовой метрик, можно задать формулу объединенной структурно-цветовой метрики. В работе она будет определяться как взвешенная сумма расстояний (метрик) в структурном и цветовом пространстве признаков и вычисляться по формуле:

$$D(x,y)=\omega_c\rho(x,y)+\omega_t\Delta E_{00}(x,y) \tag{39},$$

где ω_c,ω_t значения, определяющие весовые коэффициенты для цветовой и структурной метрик.

На основании полученных признаков строятся алгоритмы поиска границ зерен и выполнения операции по их сегментации для получения структуры сплава и дифференциации его структурных компонентов на предмет выявления областей с дефектами. Поиск дефектов производится путем сравнения конфигурации выделенных в процессе обработки зерен, образующих отдельные сегменты. На рис. 1б в центре рисунка отчетливо виден слой зерен вытянутой и уплощенной конфигурации, которые резко отличаются от остальных.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4. Результаты

Практическим результатом данной работы стало программное средство, реализующее метод сегментации, построенный на основании рассмотренного алгоритма. В качестве среды разработки используется Microsoft Visual Studio 2015. В качестве языка программирования для данной разработки был выбран язык C#. Тестирование проводилось на тестовых исходных металлографических изображениях (рис. 3) разрешением 1980×960 пикселей. Так как предложенный метод предполагает работу с большим количеством параметров, что требует значительных вычислительных мощностей, и, в частности, использования параллельных вычислений, предложенный метод сегментации был протестирован в усеченном варианте, т.е. с использованием несколько ограниченного пространства признаков и упрощенной метрики. Тем не менее, в таком варианте программное средство показало удовлетворительные результаты с точки зрения металлографии, (рис. 3 и рис. 4).

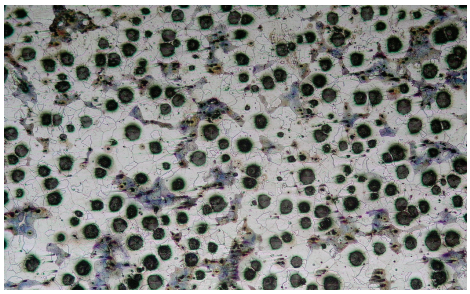


Рис. 3. Исходное тестовое металлографическое изображение (высокопрочный феррито-перлитный чугун, x100).

На рис. 4 представлены результаты сегментации изображения рис. 3 разработанным в данной работе методом.

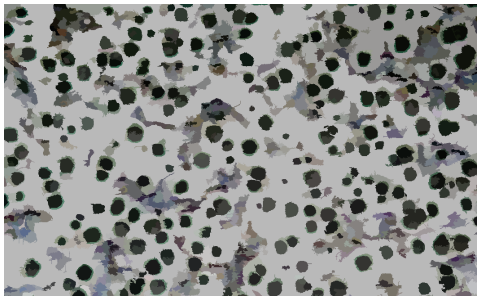


Рис. 4. Изображение, сегментированное разработанным методом.

Сравнение изображений показывает, что наблюдается общая тенденция укрупнения сегментов, вызванная слиянием более мелких областей с более крупными соседями.

5. Заключение

Предлагаемый алгоритм сегментации показал хорошие результаты на металлографических изображениях со сложной структурой. Это стало возможным благодаря использованию комбинированного структурно-цветового построения пространства признаков. Однако во время экспериментальной проверки метода были выявлены и недостатки этого подхода. Наиболее существенным из них стала большая вычислительная сложность, как самих признаков, так и меры в пространстве признаков.

Обработка рассмотренного в работе изображения занимала порядка 40 секунд на 4-х ядерном комьютере с тактовой частотой 2,6 ГГц. В дальнейшем для решения задачи в полном объеме планируется использовать кластер параллельных вычислений, созданный на базе учебной лаборатории кафедры И и ПО БГТУ.

В результате комплексного исследования данных проблем были выявлены следующие узкие места алгоритма, приводящие к указанным выше проблемам:

1. Сложность расчета признаков структуры для каждого пикселя. Так как для каждого пикселя вычисляются не менее шестнадцати таких признаков, то в масштабах небольшого изображения разрешением 1980×960 их количество достигает 30412800.
2. Сложность используемой цветовой меры близости. Метрика CIEDE2000 (используемая в алгоритме) является наиболее совершенной на текущей момент мерой близости для цветового пространства CIE Lab, но и одновременно является наиболее сложной для вычислений.
3. Необходимость разработки процедуры выделения на сегментированной структуре дефектных зон, отличающихся по своим пространственным признакам от структурных бездефектных зон компонентов сплава.

В дальнейшем представляются следующие пути развития описанного метода сегментации, при сохранении общей идеи:

1. Для упрощения расчетов признаков структуры предполагается несколько путей развития, такие как: применение более рациональных алгоритмов для расчета признаков масштаба большего, чем единичный; введение усредненного признака структуры.
2. Для решения проблем сложности используемой меры близости предполагается два пути. Первый – использование более простой меры близости в том же цветовом пространстве. Второй – отказ от пространства CIE Lab и поиск иного пространства цветовых признаков, и меры в нем.

6. Литература

[1] Fukunaga, K. The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition / Fukunaga K., Hostetler L.D. // IEEE Trans. Information Theory. – 1975. – № 21. – С. 32-40.

[2] Андросов К.Ю., Горбунов А.Н. Обработка цифрового изображения микроструктуры сплава для идентификации его структурных составляющих / Вестник образовательного консорциума Среднерусский университет. Серия: Информационные технологии. – 2015. – № (6) – С. 11-14.

[3] Чочиа, П.А. Пирамидальный алгоритм сегментации изображений / Чочиа П.А. // Информационные процессы. – 2010. – №1. – С. 23-30.

Об авторах

Андросов Кирилл Юрьевич, аспирант кафедры «Информатика и программное обеспечение» Брянского государственного технического университета. Его e-mail androkirl@yandex.ru.

Горбунов Александр Николаевич, к.т.н. доцент кафедры «Информатика и программное обеспечение» Брянского государственного технического университета. Его e-mail angor47@yandex.ru.

Давыдов Сергей Васильевич, д.т.н., профессор кафедры «Триботехническое материаловедение и технология материалов» Брянского государственного технического университета. Его e-mail davidov69@tu-bryansk.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Редуктивно-аккомодативный принцип обработки гетерогенных данных мультисенсорной системы

Р.А. Багутдинов[1], С.Г. Небаба[1], А.А. Захарова[1]
ravil_bagutdinov@yahoo.com|stepanlfx@tpu.ru|zaa@tpu.ru
[1]Томский Политехнический Университет, Томск, Россия

В данной работе представлена классификационная характеристика гетерогенных данных и подбор соответствующих методов, рекомендуемых для их обработки. В работе предложен редуктивно-аккомодативный принцип обработки гетерогенных данных мультисенсорной системы. Изучена зависимость одной выборки данных от других. Основываясь на теоретических исследованиях, в том числе в области системного анализа, проведен классификационный анализ разнородных и разномасштабных данных и соответствующих методов их обработки, в том числе с использованием методов математической статистики.

Ключевые слова: *разнородные данные, разномасштабные данные, датчики, робототехнические комплексы, гетерогенные данные, обработка данных, мультисенсорные системы.*

Reductive-accommodative principle of processing heterogeneous data of a multi-sensor system

R.A. Bagutdinov[1], S.G. Nebaba[1], A.A. Zakharova[1]
ravil_bagutdinov@yahoo.com|stepanlfx@tpu.ru|zaa@tpu.ru
[1]Tomsk Polytechnic University, Tomsk, Russia

In this paper, a classification characteristic of heterogeneous data is presented and the selection of appropriate methods recommended for their processing. The reductive-accommodative principle of processing heterogeneous data of a multi-sensor system is proposed in the work. The dependence of one sample of data on others was studied. Based on theoretical studies, including in the field of system analysis, a classification analysis of heterogeneous and different-scale data and corresponding methods of their processing, including using mathematical statistics methods, was carried out.

Keywords: *heterogeneous data, multiscale data, sensors, robotic complexes, heterogeneous data, data processing, multisensor systems.*

1. Введение

Наиболее активно развивающимися областями науки являются направления, связанные с обработкой больших массивов данных: компьютерное зрение, робототехника, физика элементарных частиц и другие. Проблема обработки, анализа и хранения больших объемов данных, получаемых от различных сенсоров, является актуальной задачей [13]. Сенсор в контексте задач получения и обработки данных может рассматриваться достаточно широко, включая в себя любой источник цифровой информации. Следовательно, совокупность таких источников представляет собой мультисенсорную систему (МС). При этом характерной особенностью становится не только рост объема данных, но и увеличение их неоднородности. В связи с этим происходит отход от традиционных подходов обработки данных [9, 12].

С развитием информационных технологий, наблюдается повышенный интерес к решению задач обработки больших объемов данных (Big Data). Однако все еще не существует эффективного решения проблемы создания универсальных моделей, способов, алгоритмов и методов для разнородных, неформализованных и неструктурированных данных, имеющих различные типы и источники происхождения. Успешное решение этой проблемы приведет к существенному, прогрессу решения прикладных задач за счет повышения эффективности, скорости работы таких систем и принятия решений на основе обработки большого объема разнородных данных.

Под термином «редукция данных» понимается совокупность аналитических методов для уменьшения размерности больших данных.

2. Редуктивно-аккомодативный принцип обработки гетерогенных данных МС.

Быстрый выбор наиболее оптимального метода для обработки больших данных имеет большое значение для многих сфер науки и техники, и являются основной частью многих интеллектуальных систем, основанных на идее работы с большими объемами гетерогенных данных. Сложность задач, решаемых такими системами, постоянно увеличивается, а требования к их техническим характеристикам повышаются. При этом все острее встают вопросы надежности и точности подобных систем, особенно для таких критичных направлений как автономные транспортные средства, системы безопасности, медицинские технологии, моделирование и прогнозирование природных и социальных событий [1].

Учитывая текущий прогресс в области вычислительной техники и сбора данных, а также связанный с этим экспоненциальный рост объемов информации, поступающей из различных источников, актуальными становятся такие проблемы, как

1. структуризация и классификация типов данных и существующего разнообразия методов обработки, определение их зависимости для выбора оптимального решения;
2. разработка новых подходов к созданию универсальных технологий и систем обработки разнородных больших данных.

Предлагается редуктивно-аккомодативный принцип обработки гетерогенных данных МС, а также классификационная характеристика данных (КХД) и соответствующая классификация методов обработки

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

данных в зависимости от требуемого приоритета решения задачи (определение качественных или количественных показателей). Под предлагаемым редуктивно-аккомодативным принципом (от лат. reduction – сокращение, accommodatio – приспособление) понимается сокращение размерности больших гетерогенных данных за счет выделения ключевых параметров, объясняющих в той или иной степени изменение исходных данных. Для решения поставленной задачи используются методы структурирования, классификации, обработки гетерогенных данных МС.

Существует множество исследований и практических реализаций систем обработки разнородных данных для различных специализированных задач: мониторинг технического состояния различных объектов, поиск аномальных и ложных данных среди источников информации, прогнозирование погодных явлений [6, 10, 2]. Несмотря на это, недостаточно внимания уделяется разработке универсального, комплексного и системного подхода к процессу разработки таких систем.

Зачастую данные синхронизированы по времени, но также встречаются и другие данные, которые не могут быть синхронизированы по времени, отсчетными точками для таких данных могут быть схожие параметры [11].

При создании такой модели КХД можно учитывать выбор данных по определенным критериям, относительно времени, часть данных могут переходить из одного типа данных в другой или быть одновременно частью этих типов. КХД вводится для упрощения обработки данных и последующего решения конкретной практической задачи. В процессе классификации выбираются схожие параметры для того или иного типа данных. Т.е. предлагается универсальный способ описания данных и представление характеристики данных в качестве некоего комплекса знаний, в соответствии с которым можно будет предложить именно те методы обработки данных, которые наиболее подходят для решения конкретной задачи.

Данные также могут быть как регулярные, так и не регулярные, т.е. не только разнородные, но и разномасштабные по времени [14]. Для обработки нерегулярных данных, как можно заключить, требуется наиболее сложные методы обработки или комплекс методов [3]. Для того чтобы лучше понять способ применения предлагаемых классификаций на рисунке 1 приведен пример реализации обработки разнородных данных, полученных с газоанализатора и тепловизора [8]. Данные поступают с датчиков, проводится их первичная обработка, в соответствии с каждым набором данных выводится их КХД, в соответствии с которой впоследствии предлагается наиболее подходящий метод или методы обработки данных, которые приведут к получению наиболее полной ценной информации обо всей совокупности разнородных данных.

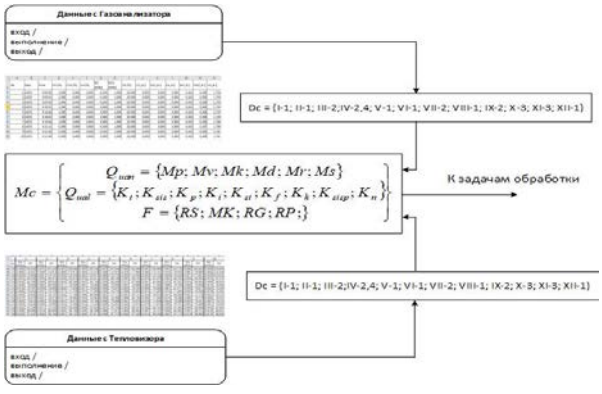


Рис. 1. Модель применения КХД.

Пример расшифровки описания КХД:

$D_1 = \{I1; II1; III1; IV1; V1; VI1; VII1; VIII1; IX1; X1; XI1; XII1\}$

Расшифровка: По способу ввода-вывода информации в системе данные поступают параллельно, источник данных – датчики (конкретизация типа датчиков описывается дополнительно), данные имеют статический вид, данные открыты аппаратно, имеют открытый доступ, данные цифровые, универсальные, по форме представления двумерные, глобальные, неструктурированные, модель данных иерархическая, хранятся локально.

Проблема обработки больших разнородных данных зависит от постановки задач, методов обработки, и подходов к решению. Требуется объединение данных, полученных с различных источников, приведение их к общей понятной форме, для проведения последующих операций обработки. Решение такой проблемы обработки данных позволяет уйти от необходимости самостоятельно отбирать источники с наиболее важной информацией.

В научных исследованиях сформулировано множество условий, определений и методов, обработки больших данных. Но даже сейчас нет четкого соглашения о том, как ускорить этот процесс, упростить работу пользователя с данными, выявить наиболее ценную информацию из огромного потока данных, не потеряв при этом часть информации, которая может быть важна.

3. Математический аппарат обработки гетерогенных данных мультисенсорной системы, проведение эксперимента и анализ результатов.

Для проведения эксперимента используется МС на базе чипа Atmega 2560, на которой установлены различные датчики, данные с которых считываются через авторский скетч и поступают либо в таблицу Excel с помощью соответствующего скрипта, либо в БД для последующей обработки.

Упрощенно весь процесс обработки данных можно разделить на несколько этапов, приведенных ниже.

1. Сырые данные приводим к единой форме.
2. Проводим синхронизацию по времени.
3. Проводим сортировку данных, чтобы поток данных делился на интервалы по 100 замеров каждый для удобства расчетов.
4. Проводим комплексный корреляционный анализ.
5. Проводим комплексный регрессионный анализ.

Т.к. регрессионный анализ данных проводят по двум выборкам, то можем сравнить последовательно каждую выборку данных.

Для решения задачи регрессионного анализа отбросим те данные, которые не влияют на конечный результат анализа (всего 12 столбцов): последний столбец (т.к. там данные порядкового характера), первый столбец (т.к. там данные времени), второй столбец, т.к. там данные температуры, которые в нашем случае не изменялись). Итого 9 столбцов.

Сравниваем последовательно каждую выборку:

Для удобства обозначим каждый столбец буквой латинского алфавита (таблица 1).

Humidity (%)	PhotoPin (lux)	MQ2 (ppm)	Ratio (Om)	LPG (ppm)	Methane (ppm)	Smoke (ppm)	Hydrogen (ppm)	PIR [0;1]
A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**

Таблица 1. Условное обозначение каждой выборки данных МС.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Для того чтобы провести регрессионный анализ со всеми данными, необходимо сравнить значение каждого столбца последовательно со значениями каждого последующего столбца (таблица 2):

AB	-	-	-	-	-	-	-
AC	**BC**	-	-	-	-	-	-
AD	**BD**	**CD**	-	-	-	-	-
AE	**BE**	**CE**	**DE**	-	-	-	-
AF	**BF**	**CF**	**DF**	**EF**	-	-	-
AG	**BG**	**CG**	**DG**	**EG**	**FG**	-	-
AH	**BH**	**CH**	**DH**	**EH**	**FH**	**GH**	-
AI	**BI**	**CI**	**DI**	**EI**	**FI**	**GI**	**HI**

Таблица 2. Матрица условных обозначений каждой выборки данных МС для последующей обработки.

Уравнение регрессии имеет вид $y = bx + a$. Оценочное уравнение регрессии по выборочным данным будет иметь вид $y = bx + a + \varepsilon$,
где ε_i – случайная ошибка (отклонение), a и b соответственно оценки параметров α и β регрессионной модели, которые следует найти.

Причины существования случайной ошибки: отброс значимых данных, аппроксимация данных с различными параметрами, неверное описание модели, ошибки измерения [7].

Так как ε_i для наблюдения i – случайны и их значения в выборке неизвестны, то по наблюдениям x_i и y_i можно получить оценки параметров α и β регрессионной модели, которыми являются случайные величины a и b, т. к. соответствуют случайной выборке.

Для оценки параметров α и β - используют МНК (метод наименьших квадратов). МНК дает наилучшие результаты, но только в том случае, если выполняются определенные предпосылки относительно случайного члена (ε) и независимой переменной (x).

Формально критерий МНК можно записать так:

$$S = \sum(y_i - yi)2 \to min.$$

Система нормальных уравнений:

$$\begin{cases} an + b\sum x = \sum y \\ a\sum x + b\sum x2 = \sum xy \end{cases}.$$

Следующим этапом строится расчетная таблица для расчета параметров регрессии, далее получаем эмпирическое уравнение регрессии для каждой из выборок:

1) y=0.3026x+781.64 (PhotoPin);
2) y=0.408x+153.99 (MQ2 (вся газовая смесь));
3) y=0.0008x+0.4086 (Ratio);
4) y=13.376x+3645.2 (LPG);
5) y=140.87x+33264 (Methane);
6) y=104.03x+26952 (Smoke);
7) y=16.108x+4747.3 (Hydrogen);
8) y=0.014x-0.3294 (PIR).

Прогнозируемый уровень показывает коэффициент a, но только в тех случаях, когда x находится близко с выборочными значениями, если нет, то буквальная интерпретация данных может привести к неверным результатам, и даже если линия регрессии точно описывает значения наблюдаемой выборки, то это не значит, что аналогично будет при экстраполяции влево или вправо. Подставив в уравнение регрессии соответствующие значения x, можно определить выровненные (предсказанные) значения результативного показателя $y(x)$ для каждого наблюдения [4, 5].

Связь между y и x определяет знак коэффициента регрессии b (если > 0 – прямая связь, иначе - обратная). В нашем примере связь прямая. Далее находим числовое значение регрессии для каждой из пар выборки, а также корреляцию (соответствующие матрицы корреляционных и регрессионных данных представлены в таблице 3, а карты данных корреляции и регрессии представлены на рисунке 2). Карты корреляции и регрессии в данном случае строятся для быстрого визуального анализа применения предлагаемого принципа. На карте корреляции видно, что наибольшее сгущение данных происходит в области водорода, дыма, метана и LPG, карта регрессии подтверждает, что наибольшее значение регрессии наблюдается именно в области углеводородов (пик значений прослеживается в области дыма и метана).

В результате получаем:

Корреляция данных

	Humidity (%)	PhotoPin	MQ-2	Ratio	LPG	Methane	Smoke	Hydrogen	PIR	
Humidity (%)	1									
PhotoPin	0,095	1								
MQ-2	0,864	-0,152	1							
Ratio	-0,099	-0,996	0,143	1						
LPG	0,155	0,984	-0,054	-0,984	1					
Methane	0,153	0,961	-0,031	-0,964	0,992	1				
Smoke	0,179	0,924	0,008	-0,932	0,967	0,989	1			
Hydrogen	0,199	0,892	0,038	-0,903	0,942	0,972	0,995	1		Среднее
PIR	0,061	-0,345	0,036	0,328	-0,373	-0,373	-0,358	-0,331	1	0,310646

Регрессия данных

	Humidity (%)	PhotoPin	MQ-2	Ratio	LPG	Methane	Smoke	Hydrogen	PIR	
Humidity (%)	1									
PhotoPin	7,081	1								
MQ-2	-9,547	-304,682	1							
Ratio	-0,018	0,599	0,101	1						
LPG	-313,008	9988,786	1663,881	-662,824	1					
Methane	-3296,218	105197,630	17522,959	-6980,321	6000671,229	1				
Smoke	-2434,462	77886,323	12940,132	-5155,107	443584,800	4187973,001	1			
Hydrogen	-376,917	12028,981	2003,684	-798,182	68984,683	648465,565	517125,765	1		Среднее
PIR	-0,327	10,455	1,742	-0,693	59,697	563,604	449,452	77,038	1	281078,6

Таблица 3. Матрица корреляционных и регрессионных данных.

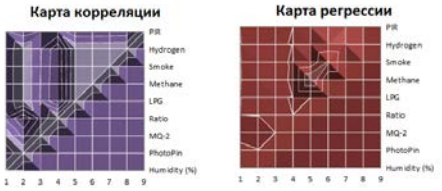


Рис. 2. Карты корреляции и регрессии.

Линейный коэффициент корреляции принимает значения от -1 до +1. Связи между признаками могут быть слабыми и сильными (тесными). Их критерии оцениваются по шкале Чеддока. В данном случае среднее значение корреляции составляет 0,311, т.е. связь между данными слабая, но прямая.

Коэффициент линейной парной корреляции может быть вычислен через коэффициент регрессии b и в среднем составит:

$$r_{x,y} = b\frac{S(x)}{S(y)} = 0.311,$$

где S – среднеквадратичное отклонение.

Общий коэффициент детерминации по всем выборкам в этом случае будет равен:

$$R^2 = 1 - \frac{\sum(y_i - y_x)^2}{\sum(y_i - \bar{y})^2} = 0.431,$$

т.е. на основании выше сказанного можно заключить, что связь между разнородными данными слабая, но прямая, в 43.1% случаев изменения одних данных приводят к изменению других данных. Точность подбора уравнения регрессии - высокая. Остальные 56.9% изменения данных объясняются факторами, не учтенными в модели (а также ошибками спецификации).

Наибольшее влияние в ходе эксперимента прослеживается между значениями углеводородной смеси, метана, дыма и водорода (в общем, порядка 60-70%). Наименьшее значение между данными влажности, сопротивлением МС и значениями инфракрасного датчика присутствия (ИКД).

Исходя из рассмотренного примера, можно выдвинуть следующую гипотезу: во многих прикладных задачах, требующих больших объемов разнородных данных для нахождения решения приемлемой точности, есть возможность получить то же решение даже по информации от существенно ограниченного набора датчиков МС при условии правильной интерпретации имеющихся данных и

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

нахождении новых, неявных зависимостей между данными. Другими словами, если с помощью имеющейся МС необходимо решить практическую задачу на основании лишь двух-трех выборок значений данных с газоанализатора, то эту задачу можно решить без установки дорогостоящего дополнительного оборудования. Решение было получено по ограниченному набору датчиков, в данном случае удалось обойтись без датчика температуры, а также доказано, что значения ИКД и влажности незначительно влияют на результат.

Полученные в ходе эксперимента теоретические результаты и формулируемые на их основе выводы подтверждаются строгостью математических выкладок, базирующихся на аппарате интегрального и дифференциального исчисления, теории вероятностей и математической статистики. Справедливость выводов относительно эффективности предложенной системы подтверждена статистическим моделированием и опытно-методической обработкой реальных результатов.

4. Заключение

В работе предложен редуктивно-аккомодативный принцип обработки гетерогенных данных мультисенсорной системы. Другими словами предлагается некий унифицированный подход к проблеме разработки высокопроизводительных систем обработки многомерных разнородных данных под конкретную прикладную задачу и заданные требования. В дальнейшем разработанная система может легко масштабироваться различными наборами данных для применения в конкретных областях технических задач.

Изучена зависимость одной выборки данных от других. Для расчетов была выбрана парная линейная регрессия. Оценены её параметры методом наименьших квадратов. Статистическая значимость уравнения проверена с помощью коэффициента детерминации. Установлено, что в исследуемой ситуации 0.9% общей вариабельности одних данных объясняется изменением других данных. Установлено также, что параметры модели статистически не значимы.

Из-за постоянной компликации задач обработки, поиска, сбора и распределения большого объема данных возникает необходимость универсальной системы классификации и их взаимосвязи, а также хорошо проработанных сценариев их обработки. Предлагаемая классификация способствует быстрому оптимальному поиску решения задач, т.к. позволяет единовременно увидеть всю картину существующих связей.

Результаты исследования применимы в сфере мониторинга и обработки разнородных данных для получения быстрой информации о согласованности данных с целью получения значимой информации, на основании которой можно быстро принять правильное решение. Данные исследования могут быть использованы при моделировании ситуаций, требующих быстрого реагирования, таких как: осуществление ликвидации аварий, моделирование эвакуации людей из зданий в чрезвычайных ситуациях, моделирование ситуаций при террористических атаках.

5. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-41-700001 р_а.

1. Литература

[1] Bagutdinov R.A., Zaharova A.A. The task adaptation method for determining the optical flow problem of interactive objects recognition in real time. Journal of Physics: Conference Series. 2017; 803(1): 012014.https://doi.org/10.1088/1742-6596/803/1/012014

[2] Kashnikov A., Lyadova L. Integration of heterogeneous sources of data based on recurrent decomposition / International Journal "Information Technologies & Knowledge" Vol.5, Number 3, 2011, P.274-284.

[3] Hastie, T., Tibshirani, R., & Friedman, J. H. (2001). The elements of statistical learning : Data mining, inference, and prediction. New York: Springer.

[4] Witten, I. H., & Frank, E. (2000). Data mining. New York: Morgan-Kaufmann.

[5] Witten, I. H., & Frank, E. (2011). Data Mining: Practical Machine Learning Tools and Techniques.

[6] Багутдинов Р.А. Гносеологические аспекты к определению назначения и состава СТЗ в задачах проектирования и разработки робототехнических комплексов. Программные системы и вычислительные методы. 2017;1:39-45. https://doi.org/10.7256/2454-0714.2017.1.20372

[7] Буре В.М. Евсеев Е.А. Основы эконометрики: Учеб. Пособие. - СПб.: Изд-во С.-Петерб. ун-та, 2004.- 72 с.

[8] Галямов А.Ф., Ризванов Д.А., Сметанина О.Н., Юсупова Н.И. Модели и алгоритмы глобально распределённой обработки слабоструктурированных данных на основе микроразметки для поддержки принятия решений // Фундаментальные исследования. – 2017. – № 1. – С. 27-35

[9] Клеменков П.А., Кузнецов С.Д. Большие данные: современные подходы к хранению и обработке. //Труды ИСП РАН. Том: 23. 2012. - С. 143-158

[10] Лысенко Н.В., Мончак А.М. Анализ эффективности гетерогенных видеоинформационных систем на основе критерия доминирования. / Радиоэлектроника. 2018. С. 57-62.

[11] Сибиряков М.А., Васяева Е.С. Модификация и моделирование алгоритмов обработки данных в кэш-памяти систем хранения данных / Кибернетика и программирование. — 2016. - № 4. - С.44-57. doi: 10.7256/2306-4196.2016.4.18058.

[12] Скобло Т.С., Клочко О.Ю., Белкин Е.Л., Сидашенко А.И. Новые подходы в исследовании неоднородности гетерогенных структур / Металлофизика и новейшие технологии, 2018.Т. 40, № 2. — С. 255-280.

[13] Финогеев А.А., Финогеев А.Г., Нефедова И.С. Технология конвергентной обработки данных в защищенной сети системы мониторинга/ Фундаментальные исследования. – 2015. – № 11-5. – С. 923-927

[14] Юревич Е.И. Сенсорные системы в робототехнике: учеб. пособие / СПб.: Изд-во Политехн. ун-та, 2013. - 100 с.

2. Об авторах

Багутдинов Равиль Анатольевич – аспирант, программист ОАР ИШИТР НИ ТПУ. E-mail: ravil_bagutdinov@yahoo.com.

Небаба Степан Геннадьевич – к.т.н., инженер Научно-образовательной лаборатории 3DM и ПД ОАР ИШИТР НИ ТПУ. E-mail: stepanlfx@tpu.ru.

Захарова Алена Александровна – профессор, д.т.н., зав. Научно-образовательной лаборатории 3D-моделирования и промышленного дизайна ОАР ИШИТР НИ ТПУ. E-mail: zaa@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Система процедур визуального анализа многомерных данных

Бондарев А.Е., Галактионов В.А., Клышинский Э.С., Шапиро Л.З.
bond@keldysh.ru|vlgal@gin.keldysh.ru|klyshinsky@mail.ru|pls@gin.keldysh.ru
ИПМ им. М.В. Келдыша РАН

В работе рассматриваются задачи визуального анализа многомерных наборов данных. Для визуального анализа применяется подход построения упругих карт. В работе приведены результаты применения упругих карт для визуального анализа многомерных наборов данных различного происхождения. В частности, рассматривается задача анализа текстовой информации, представленной в виде многомерного массива частот совместного употребления глаголов и существительных. Описан ряд процедур обработки данных, позволяющих улучшить полученные результаты.

Ключевые слова: *многомерные данные, визуальный анализ, упругие карты.*

System of procedures for visual analysis of multidimensional data

Bondarev A.E., Galaktionov V.A., Klyshinsky E.S., Shapiro L.Z.
bond@keldysh.ru|vlgal@gin.keldysh.ru|klyshinsky@mail.ru|pls@gin.keldysh.ru
Keldysh Institute of Applied Mathematics RAS, Moscow, Russia

The paper considers the problems of visual analysis of multidimensional data. For visual analysis, the elastic maps approach is used. The paper presents the results of applying elastic maps for the visual analysis of multidimensional datasets having various origins. In particular, the problem of the analysis of textual information represented in the form of a multidimensional array of frequencies of joint use of verbs and nouns is considered. A number of data processing procedures are described that allow improving the results obtained.

Keywords: *multidimensional data, visual analysis, elastic maps.*

1. Введение

В анализе многомерных данных особое место занимают задачи классификации. При решении задач классификации весьма полезными оказываются подходы визуальной аналитики, являющиеся синтезом нескольких алгоритмов понижения размерности и визуального представления многомерных данных во вложенных в исходный объем многообразиях меньшей размерности.

К таким алгоритмам можно отнести отображение исходного многомерного объема в упругих картах [2,6,7] с разными свойствами упругости или эластичности. Эти методы позволяют тем или иным образом выделить из исходного многомерного объема данных содержащуюся в нем кластерную структуру. Авторами подхода [2,6,7] разработан программный комплекс ViDaExpert [3], позволяющий проводить построение и визуальное представление упругих карт. Основные функциональные особенности данного программного комплекса подробно описаны в [2].

Интерес к упругим картам появился у нас в процессе реализации проекта по разработке вычислительной технологии для построения, обработки, анализа и визуального представления многомерных параметрических решений задач газовой динамики. Вычислительная технология реализована как единая технологическая цепочка алгоритмов производства, обработки, визуализации и анализа многомерных данных. Такая технологическая цепочка может рассматриваться как прототип обобщенного вычислительного эксперимента для нестационарных задач вычислительной газовой динамики. В итоге подобный обобщенный вычислительный эксперимент позволит получать решение не одной отдельно взятой задачи, а решение для целого класса задач, задаваемого диапазонами изменения определяющих параметров. Также следует отметить универсальность подобного обобщенного вычислительного эксперимента. Он может быть применен к широкому кругу задач математического моделирования нестационарных процессов. Практическая реализация подобного обобщенного эксперимента может обеспечивать организацию крупномасштабных промышленных расчетов. Описание элементов реализованной вычислительной технологии приведено в работах [4,5].

На практике упругие карты оказались полезным и достаточно универсальным инструментом, что позволило применять их к многомерным объемам данных разного типа. Данный подход был применен к задачам анализа текстовой информации, где в качестве числовых характеристик выступали частоты употребления слов [1].

2. Упругие карты

Идеология и алгоритмы реализации построения упругих карт подробно представлены в работах [2,6]. Подобная карта представляет собой систему упругих пружин, вложенную в многомерное пространство данных. Этот подход основывается на аналогии с задачами механики: главное многообразие, проходящее через «середину» данных, может быть представлено как упругая мембрана или пластинка. Метод упругих карт формулируется как оптимизационная задача, предполагающая оптимизацию заданного функционала от взаимного расположения карты и данных.

Согласно [2], основой для построения упругой карты является двумерная прямоугольная сетка G, вложенная в многомерное пространство, которая аппроксимирует данные и обладает регулируемыми свойствами упругости по отношению к растяжению и изгибу.

Варьирование параметров упругости заключается в построении упругих карт с последовательным уменьшением коэффициентов упругости, в силу чего карта становится более мягкой и гибкой, наиболее оптимальным образом подстраиваясь к точкам исходного многомерного объема данных. После построения упругую карту можно развернуть в плоскость для наблюдения кластерной структуры в изучаемом объеме данных. Применение упругих карт позволяет более точно и четко определять кластерную структуру изучаемых многомерных объемов данных.

Следует отметить, что при построении упругих карт в многомерном облаке данных, состоящем из сгущений и

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

отдельных отдаленных точек, возникает проблема масштабируемости. Упругая карта будет пытаться подстроиться под рассматриваемый объем в целом – как к отдаленным точкам, так и к областям сгущения, что, естественно, не может получиться одинаково хорошо. Для того чтобы решить эту проблему и обеспечить четкое представление о данных в области сгущений в работе [1], был предложен подход, названный quasi-Zoom, заключающийся в вырезании области сгущения из рассматриваемого облака многомерных данных и построения для вырезанной области упругой карты заново.

3. Процедуры обработки многомерных данных при построении упругих карт

Рассмотрим пример построения упругих карт для объема многомерных данных, представляющих собой описание характеристик полезных ископаемых, а именно, трех сортов угля из месторождений Польши [8].

Рассматриваются многомерные данные, представляющие собой точки в многомерном пространстве признаков (характеристик образцов угля). Пространство признаков состоит из следующих характеристик образцов угля – плотность, масса, удельная теплота сгорания, зольность, содержание серы, содержание летучих компонент, содержание влаги.

Таким образом, мы имеем набор точек в 7-мерном пространстве, соответствующих различным образцам угля. В наборе данных отображены три сорта угля. Рассматривается визуальный анализ с помощью применения упругих карт и главных компонент с целью изучения кластеризации многомерного облака данных и разделения сортов угля. Здесь и далее построение и визуальное представление упругих карт реализовано с помощью программного комплекса ViDaExpert [3], подробно описанного в [2].

Для исходного объема строится «мягкая» упругая карта, отображаемая в пространстве, образованном первыми тремя главными компонентами. Красные, зеленые и синие точки соответствуют трем типам угля (см. рис. 1). Далее представляем развертку построенной карты (см. рис. 2) на плоскость, образованную двумя первыми главными компонентами.

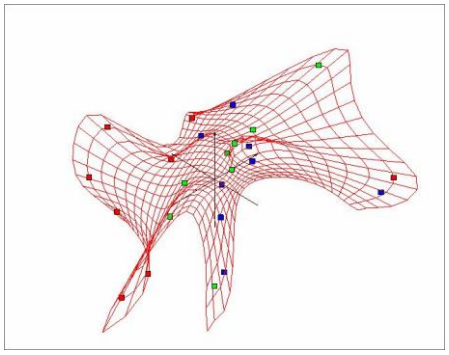


Рис. 1. Построение «мягкой» упругой карты, представляющей три сорта угля.

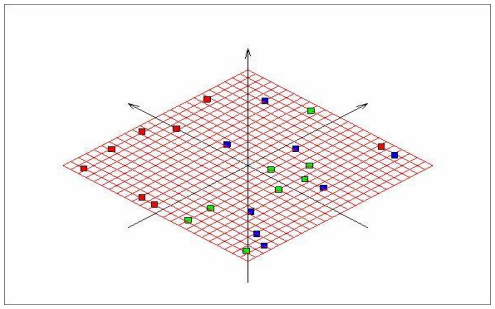


Рис. 2. Развертка «мягкой» упругой карты, представленной на предыдущем рисунке.

По развертке видно, что красные точки образуют отделившийся кластер, за исключением точки в правой части рисунка. Зеленые и синие точки перемешаны.

Для дальнейшего улучшения разделения применим фильтрацию исходного объема данных. Следует отметить, что для некоторых точек в исходном массиве представлены неполные данные, то есть для некоторых образцов информация по ряду характеристик отсутствует или находится в широком диапазоне вариации, а не представлена точно. В частности, данные по размерам образцов представлены неопределенной величиной меньше некоторого или больше некоторого предела (огромные куски или пыль). Попробуем провести фильтрацию данных, то есть убрать все точки, данные по которым представлены нечетким или неполным образом.

Удаление подобных точек из исходного объема приводит к следующим результатам (см. рис. 3).

На рисунке 3 представлена развертка «мягкой» упругой карты для измененного объема данных.

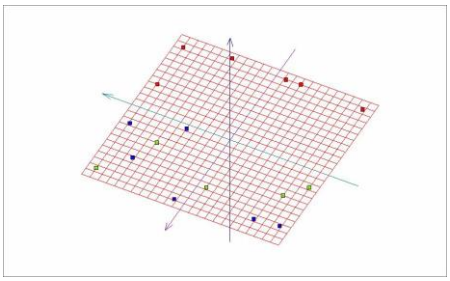


Рис. 3. Развертка «мягкой» упругой карты после фильтрации данных.

Сорт угля, представленный точками красного цвета, полностью отделился после процедуры удаления из исходного объема данных точек с нечетко определенными координатами в 7-мерном пространстве.

Сорта, представленные синими и зелеными точками, остались смешанными. Попробуем еще раз провести ту же процедуру удаления точек из данных. Однако на этот раз исключим из рассматриваемого объема красные точки целиком. Назовем эту процедуру флотацией (от английского термина flotation) аналогично термину, применяющемуся при очистке горных пород, когда более легкие фракции всплывают на поверхность и удаляются.

Теперь для четкого разделения двух оставшихся сортов достаточно отобразить точки нового объема данных в пространстве трех первых главных компонент (см. рис. 4).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

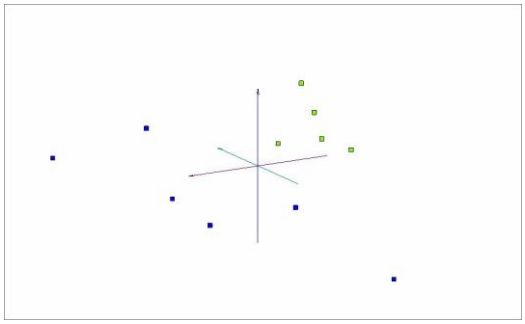


Рис. 4. Оставшиеся точки в пространстве трех первых главных компонент после флотации (удаления отделившегося полностью класса).

Итак, комбинируя построение упругих карт, отображение на вложенные в исходное пространство главные 2D и 3D многообразия и операции удаления нечетких точек и отделившихся классов (фильтрация и флотация данных), можно полностью провести разделение заданных в исходном объеме образцов на три класса, соответствующих трем сортам угля.

Эти же приемы обработки исходных данных послужили основой для разработки приема quasi-Zoom при анализе текстовой информации в работе [1]. В этой работе основное внимание уделялось изучению возможности применения методов упругих карт и t-SNE (метод стохастического распределения соседей с использованием t-распределения Стьюдента) для анализа тематической близости слов русского языка.

Основой предлагаемого метода является анализ непосредственного окружения слов. Основная гипотеза состоит в том, что близкие по смыслу слова должны встречаться в примерно одинаковом контексте. В связи с этим в пространстве признаков они будут находиться на относительно близком расстоянии друг от друга, тогда как отличающиеся слова будут находиться на более удаленном друг от друга расстоянии.

Рассмотрим результаты построения упругих карт для тестового объема [1]. В качестве тестовых данных использовались текстовые корпуса из новостных источников (ленты новостей за определенный период). Процедуры извлечения синтаксически связанных пар слов подробно описаны в [1]. Для первичных тестов было отобрано около 100 глаголов со 155 наиболее связанными с ними существительными. Полученные таким образом данные далее рассматривались как многомерный объем данных, представляющий собой 100 точек в 155-мерном пространстве. Числовые значения получающейся в результате матрицы определяются как частоты совместного употребления. Рассмотрим применение упругих карт к изображению данного объема. На рисунке 5 представлена развертка «мягкой» упругой карты для данного объема.

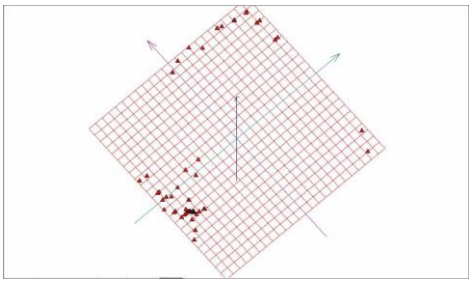


Рис. 5. Построение развертки «мягкой» упругой карты на плоскость первых двух главных компонент [1].

На рисунке видно, что изучаемый объем данных содержит область высокой плотности данных и точки, достаточно далеко отстоящие от этой области.

Именно в таких случаях возникает проблема масштабируемости, описанная ранее. Для решения этой проблемы был разработан подход quasi-Zoom, представленный в работе [1]. При исследовании объема частот совместного употребления глаголов и существительных в работе [1] практическая задача ставилась следующим образом. Нужно было максимально разделить «слипшиеся» точки. Для этой цели был разработан подход quasi-Zoom, который позволил решить эту задачу. Суть этого технологического приема заключается в том, что для более тонкой подстройки необходимо выделять большие кластеры в исследуемом объеме многомерных данных и проводить построение упругих карт для выделенных кластеров отдельно, организуя тем самым эффект, подобный функции «zoom» в современной фототехнике. Это позволит избежать проблем с масштабируемостью, когда упругая карта должна описывать как области сгущения, так и сильно удаленные отдельные точки. Этот прием сродни проиллюстрированным ранее фильтрации и флотации, однако он отличается от них тем, что имеет четкую последовательность применения.

Для улучшения разделения в области сгущения применим подход quasi-Zoom. Из исходного многомерного объема данных вырежем отделившиеся точки верхней части развертки. К получившемуся в результате этой процедуры новому объему многомерных данных заново применим построение упругой карты. При этом упругую карту будем строить сразу максимально мягкой.

После первого применения подхода quasi-Zoom разделение точек удалось значительно улучшить. Однако под точкой, соответствующей наивысшей плотности данных скрываются еще 37 «слипшихся» слов – глаголов, приведенных на рисунке 6.

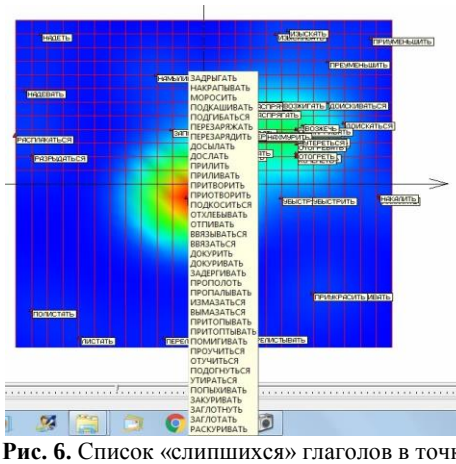


Рис. 6. Список «слипшихся» глаголов в точке наивысшей плотности данных после первого применения подхода quasi-Zoom с раскраской по плотности.

Для того, чтобы обеспечить разделение оставшихся 37 глаголов в точке наивысшей плотности данных, применим прием quasi-Zoom еще раз. Снова вырезаются отделившиеся точки таким образом, чтобы в результирующем многомерном объеме данных остались только 37 «слипшихся точек».

Повторно проводим для получившегося нового многомерного объема данных из 37 точек построение максимально мягкой упругой карты из сетки с таким же

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

количеством узлов, как и на предыдущем этапе. Большую часть точек удалось разделить, однако в зоне наивысшей плотности данных осталось еще 17 «слипшихся» точек.

Проводим процедуру quasi-Zoom в третий раз для оставшихся 17 точек. Результаты представлены на рисунке 7. В результате третьего применения процедуры удалось отделить еще 4 точки. Непосредственный анализ частот встречаемости для оставшихся 13 точек показал, что все они имеют одинаковые координаты по всем измерениям равные нулю. Следовательно, их разделение невозможно в принципе.

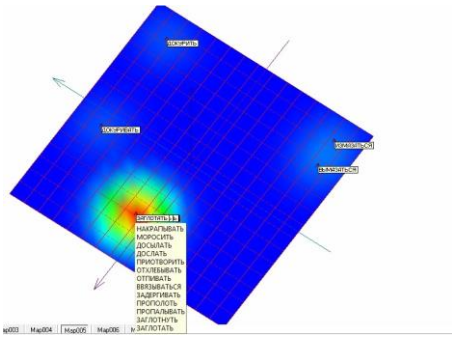


Рис. 7. Развертка упругой карты с раскраской по плотности после третьего применения процедуры quasi-Zoom [1].

Таким образом, процесс разделения точек в тестовом многомерном объеме данных с помощью построения упругих карт и применения процедуры quasi-Zoom завершен полностью и успешно. Удалось разделить все точки многомерного объема данных, имеющие различные и отличающиеся от нуля координаты. Апробация приема quasi-Zoom, проведенная в работе [1] показала эффективность разработанного подхода. При этом результаты показали необходимость проведения фильтрации данных на предварительном этапе с целью удаления точек с полностью нулевыми или полностью одинаковыми координатами. Для этой цели был построен автоматический фильтр в виде программного модуля, позволяющий проводить фильтрацию по указанному признаку. Для полного разделения точек рассматриваемого объема потребовалось только два раза последовательно применить quasi-Zoom. Повторное применение quasi-Zoom уже не оставляет слипшихся точек. Таким образом, можно сделать вывод, что предварительная очистка данных путем фильтрации существенно ускоряет обработку многомерного массива и улучшает качество обработки.

Следующим интересующим нас вопросом являлся вопрос влияния на результат транспонирования рассматриваемого многомерного массива данных. Что будет, если поменять существительные и глаголы местами? Теперь рассмотрим 155 точек, характеризующих существительные, а в качестве измерений у нас будут служить 100 глаголов (87 после фильтрации). Соответственно многомерный массив будет также состоять из частот совместного употребления существительных и глаголов, и мы сможем посмотреть и оценить, насколько схожие в понятийном смысле слова будут близки на развертке упругой карты. На рисунках 8 и 9 представлены верхняя правая и нижняя правая части развертки упругой карты. На рисунке 8 можно видеть достаточно четко выделенный понятийный кластер – ЧЕЛОВЕК, ЖЕНЩИНА, ДЕВОЧКА, ДЕВУШКА, МАТЬ, РЕБЕНОК. Чуть выше менее четко просматривается смысловой кластер – РУКА, ПАЛЕЦ, ЛИЦО, ТЕЛО.

На рисунке 9 также просматривается кластер понятий – СМЫСЛ, ДОКУМЕНТ, БУМАГА, ГАЗЕТА, КНИГА, РОМАН, ОТВЕТ.

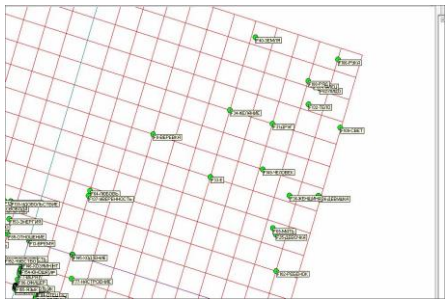


Рис. 8. Верхняя правая часть развертки упругой карты.

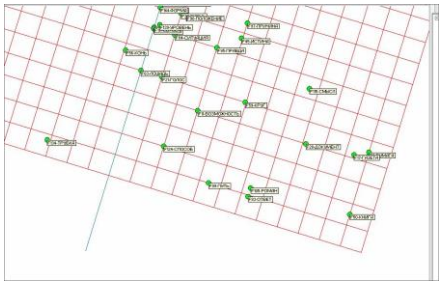


Рис. 9. Нижняя правая часть развертки упругой карты.

Это открывает широкие возможности по анализу и трактовке смысловых группировок для специалистов в этой области. Для более тщательного анализа и трактовки смысловой близости следует развивать данные подходы с точки зрения анализа расстояний между точками многомерного массива для различного задания метрики. Также интересной и важной задачей в этом плане должно стать выяснение влияния выбора метрики.

Для улучшения разделения необходимо в дальнейшем усовершенствовать алгоритм построения упругих карт за счет возможности сгущения исходной сетки карты в областях повышенной плотности данных. Подобная возможность широко используется в задачах математического моделирования сплошных сред.

4. Заключение

Для анализа кластерных структур в многомерных объемах данных использованы технологии построения упругих карт, представляющие собой методы отображения точек исходного многомерного пространства на вложенные многообразия меньшей размерности. Рассмотрен ряд приемов обработки данных, позволяющих улучшить результаты – предварительная фильтрация данных, удаление отделившихся кластеров (флотация). Для решения проблемы масштабируемости, когда упругая карта подстраивается как под область сгущения точек данных, так и к отдельно расположенным точкам облака данных, применяется подход quasi-Zoom. Суть подхода заключается в том, что для более тонкой подстройки в многомерном облаке данных выделяются большие кластеры, после чего проводится отдельное построение упругих карт для выделенных кластеров. Приведены примеры построения упругих карт для многомерных данных различного происхождения.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

5. Благодарности

Данная работа выполнена при поддержке Российского фонда фундаментальных исследований (проекты 16-01-00553а и 17-01-00444а).

6. Литература

[1] Бондарев А.Е., Бондаренко А.В., Галактионов В.А., Клышинский Э.К. Визуальный анализ кластерных структур в многомерных объемах текстовой информации / Научная визуализация. 2016. Т.8. № 3. с. 1-24.

[2] Зиновьев А.Ю. Визуализация многомерных данных - Красноярск: Изд-во КГТУ, 2000 - 168 с.

[3] Программный пакет ViDaExpert http://bioinfo.curie.fr/projects/vidaexpert/ (дата обращения 01.02.2018).

[4] Bondarev A.E., Galaktionov V.A. Analysis of Space-Time Structures Appearance for Non-Stationary CFD Problems / Proceedings of 15-th International Conference On Computational Science ICCS 2015 Rejkjavik, Iceland, June 01-03 2015, Procedia Computer Science. Vol. 51. P. 1801–1810.

[5] Bondarev A.E., Galaktionov V.A. Multidimensional data analysis and visualization for time-dependent CFD problems / Programming and Computer Software. 2015. Vol. 41. №. 5. P. 247–252. DOI: 10.1134/S0361768815050023

[6] Gorban A., Kegl B., Wunsch D., Zinovyev A. (Eds.), Principal Manifolds for Data Visualisation and Dimension Reduction, LNCSE 58, Springer, Berlin – Heidelberg – New York, 2007.

[7] Gorban A.N., Zinovyev A. Principal manifolds and graphs in practice: from molecular biology to dynamical systems / International Journal of Neural Systems. 2010. Vol. 20. №. 3. P. 219–232. DOI: 10.1142/S0129065710002383

[8] Niedoba T. Multi-parameter data visualization by means of principal component analysis (PCA) in qualitative evaluation of various coal types / Physicochemical Problems of Mineral Processing. 2014. Vol. 50. № 2. P. 575-589.

Об авторах

Бондарев Александр Евгеньевич, к.ф.-м.н., старший научный сотрудник, ИПМ им. М.В. Келдыша РАН.
Его e-mail bond@keldysh.ru.

Галактионов Владимир Александрович, д.ф.-м.н., профессор, зав. отделом, ИПМ им. М.В. Келдыша РАН. Его e-mail vlgal@gin.keldysh.ru.

Клышинский Эдуард Станиславович, к.т.н, доцент, НИУ ВШЭ МИЭМ. Его e-mail klyshinsky@mail.ru.

Шапиро Лев Залманович, к.т.н, старший научный сотрудник, ИПМ им. М.В. Келдыша РАН. Его e-mail pls@gin.keldysh.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Сцены визуализации в вычислительных блокнотах

П. А. Васёв[1], В. Л. Авербух[1,2], М. О. Бахтерев[1,2],
Д. В. Манаков[1], Д. Ю. Филоненко[3], М. А. Форгани[2]
vasev@imm.uran.ru
[1]Институт математики и механики им. Н.Н. Красовского УрО РАН, г. Екатеринбург
[2] Уральский федеральный университет им. Б.Н. Ельцина, г. Екатеринбург
[3] Уральский государственный архитектурно-художественный университет, г. Екатеринбург

В работе описан феномен вычислительных блокнотов и современное состояние визуализации в них. Проведен анализ сложившегося состояния в контексте визуального программирования. Представлена реализация визуального программирования сцен визуализации для блокнотов.

Ключевые слова: *визуализация, вычислительные блокноты, визуальное программирование.*

1. Введение

К настоящему времени стал популярным способ численного моделирования с помощью вычислительных блокнотов. Пользователь через веб-браузер вводит команды и программы, а блокнот их интерпретирует, производит вычисления, и показывает результаты.

Однако современные блокноты имеют не всегда удобные средства визуализации данных. По нашему мнению, это связано с тем, что описание видов отображения (сцен визуализации) проводится на том же языке, что и вычисления. Однако языки вычислений не приспособлены для такой задачи, что влечет громоздкость описаний и лишние затраты сил.

Мы предлагаем формировать описание визуализации с помощью графического интерфейса с небольшим набором основных действий: добавить визуализацию, выбрать способ отображения, указать необходимые переменные блокнота — источники данных для построения графического объекта.

Предполагаем, что этот способ поможет инженеру-ученому легче и быстрее конструировать визуальное представление промежуточных и конечных результатов счёта.

2. Вычислительные блокноты

Вычислительный блокнот (computational notebook, [2]) – это человеко-машинный интерфейс, позволяющий создавать программы из последовательности т. н. клеток. В клетках может содержаться текст, программный код, другие данные. Текст отображается как есть, а программный код — выполняется, и результат его работы выводится на экран блокнота рядом с исходной клеткой.

Пример блокнота показан на рис 1. Заголовки, текст и формулы — это клетка с текстом. Клетки «in[2]» и «in[5]» — с программным кодом. Изображения — результат выполнения кода клетки «in[5]».

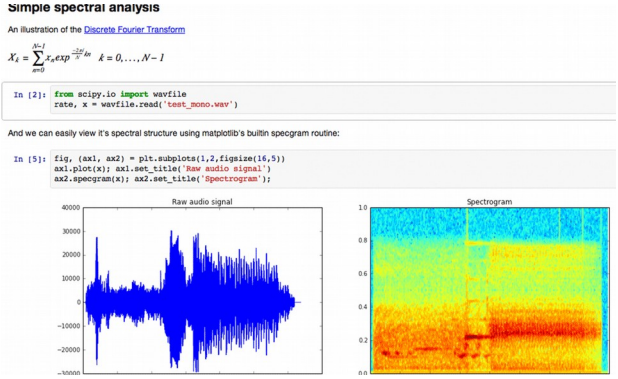


Рис. 1. Вычислительный блокнот.

Важной особенностью блокнотов является то, что обычный текст и программный код поставлены в равнозначные позиции. Технически и то и другое является входом для интерпретации; текст интерпретируется как данные в формате markdown, а код — интерпретатором программ.

Происхождение вычислительных блокнотов авторы связывают со следующими идеями программирования:

- Цикл «чтение-вычисление-вывод» (REPL). Эта идея развивается блокнотами: очередная команда пользователя не просто выполняется, а сохраняется в отдельной клетке, которую можно в последствии редактировать и выполнять повторно;
- Грамотное программирование (literate programming). Идея развивается блокнотами до состояния, когда и текст, и программный код считаются равноправными участниками программы.

Широко известные программы, использующие идею вычислительных блокнотов: Mathematica, Maple, Mathcad.

В связи с развитием веб-технологий появились новые блокноты, например:

- Observable Notebooks
- Apache Zeppelin
- Apache Spark Notebook
- JupyterLab (Jupyter Notebook, IPython)
- Nteract
- R Notebooks
- Google Colab

Одни блокноты поддерживают простой режим вычисления клеток — сверху вниз.

Другие блокноты поддерживают «реактивность»: определяют зависимости и автоматически пересчитывают клетки, если данные, от которых клетки зависят, изменились. Это может касаться как данных, рассчитанных в других клетках, так и данных из внешних источников.

Некоторые блокноты поддерживают потоковый режим поступления данных. Например, Spark Notebook умеет обрабатывать потоки из внешних источников (см. видеозапись на странице [3]). Также Spark и Observable умеют порождать в клетках объекты, являющиеся источниками потоков (их термины - streams, generators).

Открытый вопрос: могут ли блокноты и сами являться источником потоковых данных, при подключении их в другие блокноты в качестве модулей?

3. Визуализация в блокнотах

Блокноты предлагают различные варианты визуализации данных. Во-первых, они имеют встроенные средства визуализации. Например, они могут представить результат выполнения кода клетки в виде таблицы,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

изображения, графика или диаграммы. Эти средства просты и достаточны для широкого класса задач.

Кроме того, многие блокноты предлагают возможность внедрения дополнительных средств визуализации.

Анализ таких средств показал, что они разделяются на два вида:

1. Специализированные. Обеспечивают визуализацию заданного типа объектов с помощью определённого заранее вида отображения.

2. Универсальные. Предоставляют пользователю возможность самому сформировать вид отображения его данных.

Наблюдения за блокнотами показали, что обычный режим использования универсальных средств следующий.

Пользователь подключает специальную библиотеку визуализации (например matplotlib или plotly). Вводит программный код в клетку, который через программный интерфейс библиотеки визуализации формирует вид отображения.

В результате выполнения кода этой клетки формируется статичный или интерактивный графический вывод.

На рис. 2 показан пример работы в таком режиме.

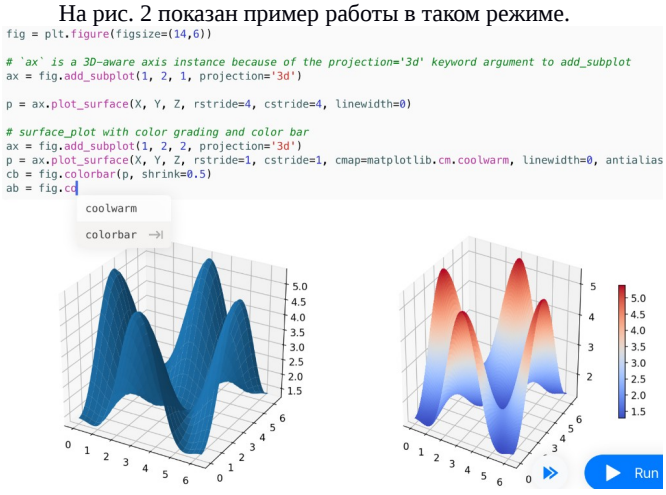


Рис. 2. Код визуализации и результат его работы.

4. Идея внедрения визуального языка

Идея состоит в том, чтобы предоставить пользователю возможность настраивать получаемую визуализацию в блокнотах с помощью визуальных средств. Например, с помощью мышки менять тип графиков, добавлять новые представления (возможно не только графики), настраивать их цвета и другие атрибуты.

То есть вместо того, чтобы писать программные коды, формирующие вид отображения – задавать его визуально.

Это могло бы облегчить работу пользователя. Ведь для написания кодов на программном языке необходимо переключать контекст мышления, так как программные интерфейсы требуют удержания их в памяти пользователя.

В свою очередь визуальные интерфейсы по своей природе построены на системе подсказок. С их помощью снижаются ментальные затраты пользователя.

Применение подобных идей в блокнотах уже наблюдается, но в ограниченном виде. Например, с помощью визуальных команд можно настроить некоторые атрибуты графиков в блокнотах Zeppelin.

Однако к настоящему времени мы не обнаружили ни одного полноценного примера визуального языка для создания видов отображения в блокнотах.

Мы предполагаем, что скорее всего такие попытки были, но они по каким-то причинам не прижились. Далее мы попытаемся выяснить эти причины.

5. О визуальных языках в блокнотах

Классически системы визуализации разделяются на специализированные и универсальные [1].

Специализированные системы предназначены для решения конкретной известной задачи визуализации. При этом они учитывают специфику не только по объектам визуализации, но и по использованию методик компьютерной графики, по научным направлениям.

Универсальные системы визуализации это среды, в которых пользователю предоставлена возможность создать требуемый ему вид отображения самостоятельно.

Можно считать, что это среды программирования, в которых пользователь описывает (программирует), какой вид отображения в данный момент требуется.

Наше обычное понимание универсальных систем визуализации строится из того, что они предлагают визуальный язык программирования. Он может быть простой, например построенный на кнопках, меню и т. д., или более развитый, например, позволяющий создавать графы потоков данных с помощью мышки.

Однако формально система визуализации не обязана предоставлять только визуальный язык. Она может обеспечивать возможность общения и через программный интерфейс. Это хорошо видно на примере макросов в универсальных системах визуализации (например в Excel или Paraview). Более того, эти системы умеют не только потреблять, но и генерировать макросы - исходя из визуальных команд, подаваемых пользователем.

Проведённый авторами анализ показал, что для получения визуализации в блокнотах в подавляющем большинстве случаев используется именно программный язык, а не визуальный графический интерфейс.

Визуальные методы, как принято считать, обеспечивают пользователю интуитивно-понятную возможность формирования необходимых описаний.

Почему же при формировании видов отображения в блокнотах пользователи предпочитают программный язык взамен визуального? И это при условии, что в программных описаниях, во всех случаях, которые нам удалось найти, мы обнаружили большую громоздкость кода. Рис. 2 иллюстрирует это — потрачено 7 строк кода для получения двух графиков. В среде Интернет мы наблюдали примеры и в 20, и в 100 строк кода для получения достаточно тривиальных визуальных сцен.

Возможные причины, почему пользователи предпочитают программные коды в случае блокнотов:

- программирование визуализации ведется на том же языке, что и вычисления. В этом случае пользователю надо преодолеть только один когнитивный барьер — освоить модель библиотеки визуализации. Мы считаем, что это серьезный барьер, так как он несёт дополнительную деталь, которую требуется удерживать в памяти пользователя помимо самих вычислений в прикладной области. Но практика показывает, что пользователей это не отталкивает. Также их не отталкивает, что язык вычислений может не подходить для описания сцен и нести излишние детали.

- возможность переносить программные коды между разными блокнотами. Пользователь формирует программный код, затрачивая на это массу усилий, но в последующей работе он тратит гораздо меньше сил, копируя и адаптируя предыдущие коды.

- возможность комментирования, встроенная в языки программирования - снижает когнитивные затраты при переносе и последующем восприятии кодов визуализации.

- изобилие примеров кодов в среде Интернет, которые дают пользователю возможность скопировать и адаптировать этот код под свои потребности.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

В результате у пользователя формируется представление, что специализированная визуализация в блокноте с помощью программных кодов – это хоть и затратная, но доступная и предсказуемая процедура.

6. О совместимости визуальных языков с концепцией блокнотов

Возможно также, что препятствие на пути использования визуальных языков программирования для получения визуализации в блокнотах лежит в плоскости самих блокнотов. А именно, в концепции клетки как входного объекта для вычислителя, в результате вычисления которого формируется выходной объект.

В этой концепции логично создавать описание вида отображения именно в клетке, а в выводе клетки наблюдать результат работы этого описания.

Получается, например, что настройка вида отображения на визуальном языке в самой выходной области клетки противоречит этой концепции. Вместе с тем многие визуальные языки построены на парадигме непосредственного взаимодействия с объектами [4].

Среди возможных решений этого противоречия мы предполагаем следующие:

1. Смена концепции вычислительных блокнотов на что-то принципиально новое.
2. Сохранение описания, визуально формируемого пользователем, в коде клетки (по аналогии с формированием макросов в системах визуализации).
3. Сохранение этого описания в мета-информации выхода клетки (поддерживается только в Zeppelin). Это может быть полезным и удобным для пользователя, хоть и является необычным с точки зрения концепции блокнотов.
4. Внедрение визуального языка для клеток, когда код клетки формируется не текстом, а визуально.
5. Визуально-текстовый язык, когда код клетки можно и представлять, и редактировать как в виде текста, так и визуально.
6. Создание типа клетки «визуализация» (в дополнение к используемым сейчас типам «код» и «текст»).

Возможно, существуют другие варианты или точки зрения; авторы будут рады их услышать.

7. Тестовая реализация

Для воплощения идеи визуального программирования сцен визуализации в блокнотах авторами разработан модуль для Jupyter Notebook.

Пример использования модуля показан на рис. 3:

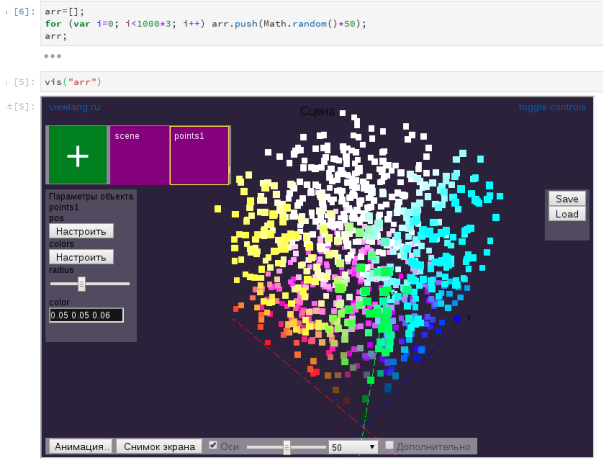


Рис. 3. Пример визуализации в блокноте, созданной с помощью визуального языка.

В этом примере пользователь рассчитал массив данных и желает визуализировать его в виде облака точек (мы посчитали это достаточно типовой задачей визуализации).

Для получения визуализации пользователь в блокноте написал вызов специальной функции. В результате её работы появилась специальная область для визуализации.

В этой области пользователь с помощью мышки добавил визуальный объект - «точки», и настроил связь между их положением в пространстве и цветом с рассчитанным в блокноте массивом. Таким образом, пользователь визуально создал требуемый ему вид отображения.

В текущей реализации пользователь может указывать не один, а несколько объектов данных для передачи системе визуализации. Также он может создавать разные визуальные объекты — точки, сферы, отрезки, стрелки, треугольники и т. д.

Например, на рис. 4 пользователь рассчитал дополнительный массив, добавил клетку с вызовом функции визуализации и мышкой создал вид отображения, состоящий из точек и отрезков. Положение точек и отрезков он связал с первым массивом, а цвет точек — со вторым. Дополнительно пользователь мышкой задал цвет отрезков.

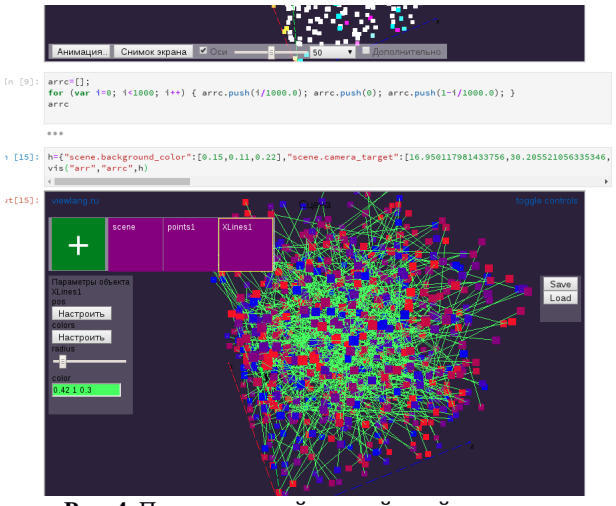


Рис. 4. Пример дальнейших действий пользователя.

Для сохранения состояния сцены применяется следующий подход. Если пользователя удовлетворяет полученная сцена, он нажимает кнопку «Сохранить» и получает текстовую запись с описанием настроенного вида отображения в формате json. Далее пользователь сохраняет его в программном коде клетки блокнота, передавая в качестве аргумента в функцию вызова визуализации. При последующем запуске вид отображения воспроизводится по этому описанию.

8. Заключение

Дальнейшее развитие представленной работы авторы видят в следующем:

1. Проверка тестовой реализации с целью понимания меры удобства использования в реальных задачах.
2. Выявление набора видов отображений, комбинации которых эффективно решают задачи визуализации в прикладных областях.
3. Создание удобного визуального языка формирования специализированных видов отображения.

Результаты программных разработок публикуются на сайте Сектора компьютерной визуализации ИММ УрО РАН www.cv.imm.uran.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Благодарности

Авторы выражают благодарность О. Г. Анненковой за экспертизу в области вычислительных блокнотов и критику настоящей работы.

Литература

[1] Авербух В.Л., Байдалин А.Ю., Бахтерев М.О., Васёв П.А., Казанцев А.Ю., Манаков Д.В., Опыт разработки специализированных систем научной визуализации // Научная визуализация. Квартал 4. Том 2. Номер 4. 2010. Стр. 27-39.

[2] https://en.wikipedia.org/wiki/Notebook_interface

[3] https://www.lightbend.com/blog/scala-and-spark-notebook-the-next-generation-data-science-toolkit

[4] https://en.wikipedia.org/wiki/Direct_manipulation_interface

Об авторах

Авербух Владимир Лазаревич, к.т.н., начальник Сектора компьютерной визуализации ИММ УрО РАН, averbukh@imm.uran.ru.

Васёв Павел Александрович, научный сотрудник Сектора компьютерной визуализации ИММ УрО РАН, vasev@imm.uran.ru.

Бахтерев Михаил Олегович, научный сотрудник Сектора компьютерной визуализации ИММ УрО РАН, m.bakhterev@imm.uran.ru.

Манаков Дмитрий Валерьявнович, программист Сектора компьютерной визуализации ИММ УрО РАН, manakov@imm.uran.ru.

Филоненко Дмитрий Юрьевич, канд. культурологии, доцент кафедры Графического дизайна УрГАХУ, philonenko@gmail.com.

Форгани Маджид Али, аспирант Института математики и компьютерных наук УрФУ, majid.forqani@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Количественная визуализация формы физических тел с помощью точечно структурированного оптического излучения

А.М. Вин[1], Б.С. Ринкевичюс[2], А.Ю. Поройков[3]
humburger49@gmail.com| rinkevbs@mail.ru | poroykovay@gmail.com
[1]Национальный исследовательский университет «МЭИ», Москва, Россия;
[2]Национальный исследовательский университет «МЭИ», Москва, Россия;
[3]Национальный исследовательский университет «МЭИ», Москва, Россия

В статье предложен и разработан новый оптико-электронный метод количественной визуализации геометрических параметров физических тел с помощью точечно структурированного оптического излучения (ТСИ), получаемого от источников когерентного излучения с использованием дифракционных оптических элементов (ДОЭ). В данной статье используется ТСИ в виде прямоугольной сетки точечных равноотстоящих источников излучения с заданными расстояниями между ними. Исследуемые геометрические фигуры устанавливаются на вертикальном экране и освещаются ТСИ. Получаемые изображения регистрируются цифровой фотокамерой с целью определения геометрических параметров исследуемых тел с помощью стандартных алгоритмов обработки. Приведено подробное описание элементов измерительного комплекса, который состоит из источника непрерывного когерентного излучения с длиной волны 0,552 мкм, ДОЭ, формирующего ТСИ в виде квадратной матрицы, соответствующей набору тонких пучков излучения с расстояниями между ними 3,4 мм на дистанции 750 мм до экрана. Также приведено подробное описание методики обработки плоских картин визуализации. Данный метод получения изображения геометрической фигуры отличается от прямого фотографирования тем, что наряду с получением изображения объекта регистрируется информация о расположении точечных источников излучения в составе ТСИ, которая в дальнейшем используется для определения геометрических параметров исследуемой фигуры или проекции физического тела. Основным преимуществом предлагаемого метода является то, что сопоставление с измерительной матрицей происходит на физическом уровне и при обработке изображений не требуется дополнительная калибровка для определения абсолютных геометрических размеров исследуемого тела.

Ключевые слова: *точечно структурированное оптическое излучение, дифракционный оптический элемент, диффузно рассеивающая поверхность, фотографирование, компьютерная обработка картин визуализации.*

Quantitative visualization of physical shapes using point-structured optical radiation

A.M. Win[1], B.S. Rinkevichyus[2], A. Yu. Poroikov[3]
humburger49@gmail.com| rinkevbs@mail.ru | poroykovay@gmail.com
[1]National Research University “MPEI”, Moscow, Russia;
[2]National Research University “MPEI”, Moscow, Russia;
[3]National Research University “MPEI”, Moscow, Russia

This article shows a new electro-optical method for quantitative imaging parameters of physical shapes with the use of point-structured optical radiation (PSR) obtained from the coherent radiation source and diffractive optical elements (DOE). PSR is used in the form of a rectangular grid of point equidistant radiation sources with given distances between them. The physical figures under consideration are mounted on a vertical screen and illuminated by the PSR. The parameters of the figures can be determined from the obtained images using standard processing algorithms. The images of the physical shapes can be photographed with a digital camera and then reproduced on a computer screen with further processing by modern digital methods. There is a detailed description of the measurement system elements, which consists of a continuous coherent radiation source with a wavelength of 0.552 μm, the DOE and a large aperture lens, which enable to receive PSR as a square matrix of thin beams with distances between them of 3.4 mm at a distance of 750 mm up to the screen. Also it is given detailed description of the processing technique of flat visualization picture. This method for obtaining a photographic image of physical shape differs from the standard photographing in the fact that in addition to obtaining an object image is recorded information on the parameters of point radiation sources in the composition of PSR, which is later used to determine the geometric parameters of the physical shapes.

Keywords: *point structured optical radiation, diffraction optical element, diffusely scattering surface, photography, computer processing of visualization pictures.*

1. Введение

Для качественной и количественной визуализации и определения параметров диффузно рассеивающих исследуемых физических фигур на практике широко используются оптические методы: прямого фотографирования, голографические, интерференционные и ряд других [1-3]. Каждый из этих методов обладает своими особенностями и областью применения.

Методы прямого фотографирования наиболее просты и широко распространены, однако требуют дополнительной калибровки для получения численных значений геометрических параметров фотографируемых исследуемых физических фигур, интерференционные и голографические методы позволяют получать трехмерную визуализацию, однако сложны в реализации [3].

Для количественной визуализации геометрических параметров диффузно рассеивающих плоских объектов в работах [4, 5] было предложено использовать точечно структурированное оптическое излучение (ТСИ), получаемое с помощью дифракционных оптических элементов (ДОЭ). Дифракционные оптические элементы широко используются в лазерной рефрактографии –

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

современной измерительной технологии, предназначенной для количественной визуальной диагностики процессов в объеме и пограничных слоях жидкостей [6].

Реализованный принцип измерения параметров диффузно рассеивающих объектов адаптирован для компьютерных методов обработки оптических изображений и позволяет регистрировать параметры объектов с линейными размерами от 1 до 30 см (ограничения связаны с разрешающей способностью и апертурой использованных объективов). Обработка оптических изображений проводится стандартными компьютерными алгоритмами [7].

2. Оптическая система визуализационного комплекса и методы исследования

Разработанный оптико-электронный комплекс для количественной визуализации геометрических параметров диффузно рассеивающих исследуемых физических тел с помощью ТСИ показан на рис. 1.

Источник оптического излучения 1 – монохроматический, работающий на длине волны 0,552 мкм, мощность излучения 20 мВт, исходный диаметр пучка составляет 2 мм, установлен на неподвижном штативе. Дифракционный оптический элемент (ДОЭ) 2 представляет собой круглую пластинку толщиной 1 мм, диаметром 5 мм, укрепленную на регулируемой по высоте стойке. После ДОЭ в результате дифракции когерентного пучка получается 51 × 51 тонких равноотстоящих микропучков, диаметр которых составляет не более 1 мм.

Оптическая система представляет собой многолинзовый объектив 3 диаметром 75 мм и фокусным расстоянием 750 мм. Объектив установлен на штативе, регулируемом по высоте, углу наклона и положению. Экран представляет собой пластиковую пластину, окрашенную в белый цвет. Регистрация изображений предметов на экране осуществляется цифровым фотоаппаратом с разрешением 0,1 мм.

Для исследования применяются оптико-электронные методы, т. е. методы, в основе которых лежит фотоэлектронное преобразование оптического излучения с последующим измерением параметров электрического сигнала, несущего информацию об оптических характеристиках исследуемых физических тел. Тонкие пучки формируются с помощью ДОЭ и при распространении сохраняют свою заранее заданную поперечную структуру с высокой точностью. Использование различных ДОЭ позволяет гибко менять форму пучков при сохранении структурной устойчивости и обеспечить выделение размеров и дефектов поверхности с заранее заданными характеристиками (микроструктуру поверхности).

Таким образом, данный метод основан на зондировании исследуемого объекта ТСИ, регистрации рассеянного на физической фигуре излучения с помощью цифрового фотоаппарата и компьютерной обработке полученных изображений. Формирование ТСИ позволяет сохранить его высокую когерентность и обеспечить малую расходимость пучков. Это дает возможность использовать для описания ТСИ представления геометрической оптики.

3. Экспериментальная установка

Схема созданной установки для количественной визуализации параматров исследуемых физических фигур показана на рис. 1. Излучение от оптического модуля 1, состоящего из источника когерентного излучения и блок питания лазера 2 и ДОЭ 3, проходит через объектив 4 и освещает экран 5 с закрепленной на нем исследуемой геометрической фигурой 6. Изображение регистрируется с помощью цифрового фотоаппарата 7 и отображается на экране персонального компьютера 8.

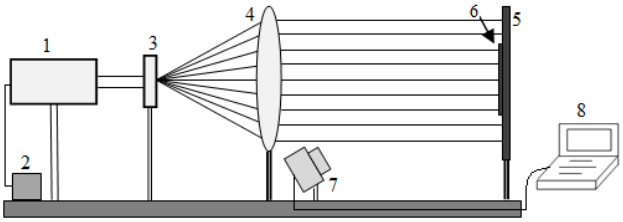
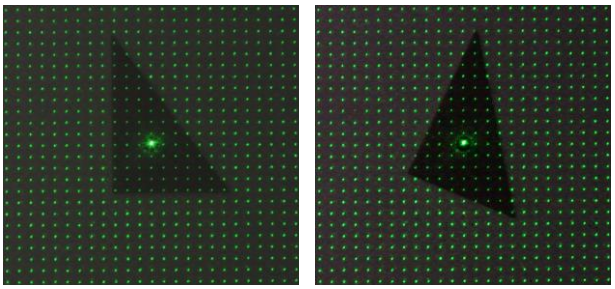


Рис.1. Схема экспериментальной установки: 1 – источник оптического когерентного излучения, 2 – блок питания лазера, 3 – дифракционный оптический элемент (ДОЭ), 4 – объектив, 5 – экран, 6 – геометрическая фигура, 7 – цифровой фотоаппарат, 8 – персональный компьютер.

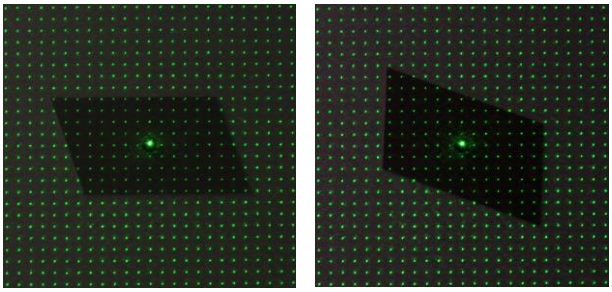
Источник оптического излучения создает узкий поляризованный пучок, который направляется на ДОЭ. При юстировке он вращается вокруг оптической оси для получения максимальной резкости дифракционной картины. После ДОЭ получается много дифрагированных узких пучков, которые направляются на коллимирующий объектив 3, создающий матрицу параллельных пучков, распространяющихся вдоль оптической оси. На экране 4 получается изображение дифракционной картины, которое регистрируется с помощью цифрового фотоаппарата 6.

На экране размещаются исследуемые физические фигуры, параметры которых необходимо определить. На рис. 2 и 3 показаны изображения двух различных исследуемых физических фигур: треугольника и параллелограмма при различных углах их поворота. Центр изучаемых фигур совпадает с центром экрана.



(а) (б)

Рис. 2. Изображения треугольной фигуры: (а) исходное, (б) повернутое.



(а) (б)

Рис. 3. Изображение параллелограммной фигуры: (а) исходное, (б) повернутое.

4. Алгоритм цифровой обработки физических фигур

Основной целью разработки алгоритма является возможность определения с его помощью геометрических

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

параметров исследуемых фигур. Особенностью алгоритма является использование ТСИ, присутствующего на экспериментальных изображениях, для определения масштабных коэффициентов. Такой подход позволяет проводить бесконтактные измерения геометрических параметров плоских фигур без предварительного проведения процедуры калибровки цифровой видеокамеры.

Точечное структурированное излучение на изображениях представляет собой периодическую структуру, состоящую из эквидистантно расположенных пятен (точек) лазерного излучения. Расстояния между пятнами позволяет оценить вертикальный и горизонтальный масштабный коэффициент. Также при необходимости можно провести калибровку видеокамеры [8] по известным расположениям точек структурированного излучения, в том числе для компенсации оптических аберраций и уменьшения погрешности измерения.

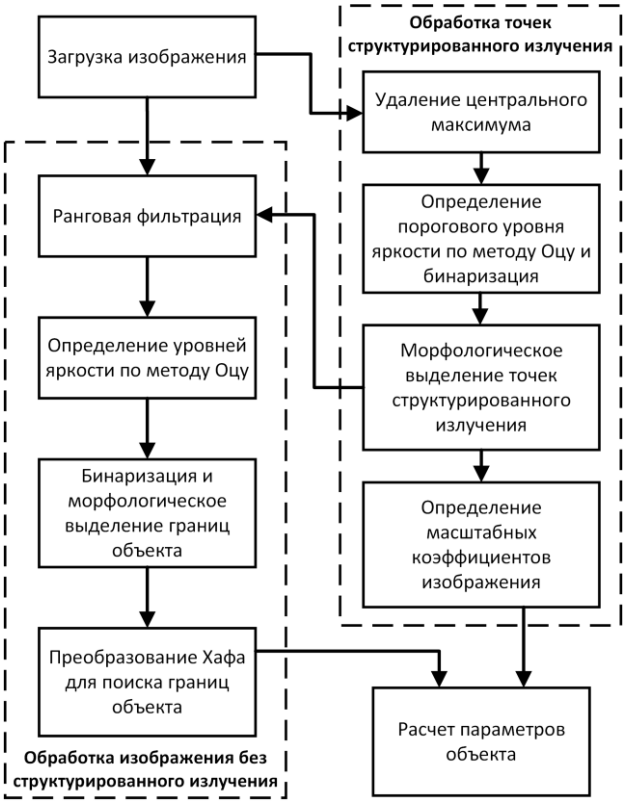


Рис. 4. Алгоритм цифровой обработки изображений для определения геометрических параметров объектов.

Предлагаемый алгоритм определения геометрических параметров плоских фигур представлен на рис. 4. Из рисунка видно, что он разделен на две части: обработка пятен или точек ТСИ на изображении и обработка дифракционных картин без точек ТСИ. Информация, полученная в первой части, используется во второй для определения размера маски, для ранговой фильтрации и в итоговом расчете параметров фигур в виде масштабных коэффициентов изображения.

Рассмотрим детально все шаги данного алгоритма. После загрузки изображения происходит удаление с него центрального максимума, который является паразитным элементом и мешает дальнейшей обработке. Удаление происходит по признаку наиболее яркого и большого по площади объекта на изображении. Далее для выделения оставшихся пятен структурированного излучения используется бинаризация изображения (пороговая фильтрация). Для проведения бинаризации автоматически определяется оптимальный уровень порога по методу Оцу [9].

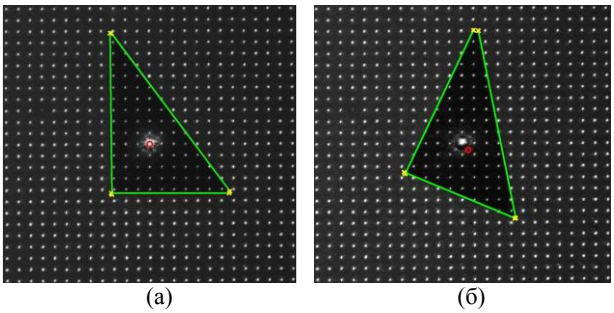
Для выделения точек используется морфологическое выделение замкнутых отдельно стоящих фигур [10]. Оценка параметров выделенных фигур позволяет определить средний размер точек. Этот размер в дальнейшем используется в ранговой фильтрации. По расположению выделенных объектов также рассчитываются масштабные коэффициенты изображения, которые в последствии учитываются при расчете геометрических параметров объектов. На этом использование точек структурированного излучения на изображении заканчивается.

В следующей части алгоритма точки становятся нежелательными объектами, которые необходимо удалить. Для этого применяется ранговая фильтрация. Размер фильтра выбирается исходя из среднего размера точек, полученного в предыдущих действиях. Для приведенного далее примера размер маски составил 15×15 пикселей при размере точек порядка 10-12 пикселей, для фильтрации выбирался 20-й по порядку элемент. Ранговая фильтрация позволяет качественно удалить точки с изображения, оставив только исследуемую фигуру и фон.

Далее происходит расчет нескольких уровней яркости по методу Оцу для выделения исследуемой фигуры. Более одного уровня требуется из-за неравномерного освещения заднего фона изображения. Выделение области, лежащей в центре изображения, и бинаризация ее по уровням яркости позволяет надежно отделить искомую фигуру от фона.

На следующем шаге алгоритма морфологические операции закрытия и выделения границ используются для точного выделения, как самой фигуры, так и ее границ. Выделенная граница подвергается преобразованию Хафа для поиска прямых [11], описывающих периметр фигуры. Для определения количества локальных максимумов в преобразовании используется априорная информация об исследуемой фигуре (количество сторон многоугольника).

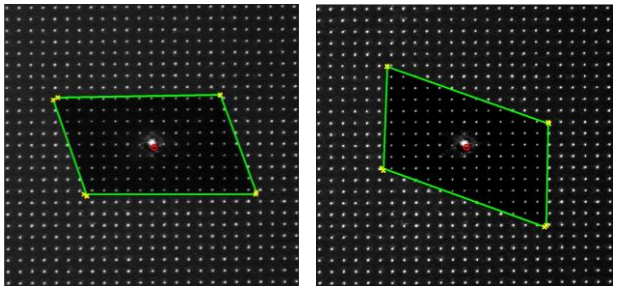
На конечном этапе алгоритма происходит расчет геометрических параметров плоской фигуры: площади, периметра, положение центра, ориентация (по главной оси описывающего объект эллипса), и др. При расчете используются масштабные коэффициенты изображения, полученные на предыдущих шагах при обработке изображения. Пример результата обработки дифракционной картины с плоским многоугольником с помощью разработанного алгоритма приведены на рис.5- 6.



(а) (б)

Рис. 5. Результаты обработки изображения треугольной фигуры на основе разработанного алгоритма: площадь 17755 мм^2, периметр 1896 мм, ориентация -62,0°, (а) исходное, (б) повернутое.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18



(а) (б)

Рис. 6. Результаты обработки изображения фигуры в виде параллелограмма: площадь 19237 мм^2, периметр 2063 мм, ориентация -12,6°, (а) исходное, (б) повернутое.

5. Заключение

Разработан бесконтактный метод визуализации проекции диффузно рассеивающего физического тела и определения его геометрических параметров на основе сопоставления изображения проекции с матрицей точек, расположенных на фиксированном расстоянии друг от друга. Сопоставление происходит на физическом уровне, что позволяет избежать возможных погрешностей, связанных с искажением объекта при фотографировании, т.е. предлагаемый метод не требует калибровки. В данном случае, матрица точек, создаваемая ДОЭ, играет роль «измерительной линейки», поэтому погрешность прямых измерений геометрических параметров определяется погрешностью расстояния между точками и для данной установки составляет 1мм.

Для реализации метода создана экспериментальная установка, содержащая источник когерентного оптического излучения, дифракционный элемент для получения ТСИ, экран, цифровую фотокамеру. Приведены результаты экспериментов по измерению параметров ряда исследуемых физических фигур при помощи ТСИ.

Дополнительное преимущество данного метода измерения геометрических параметров физических тел перед другими оптическими методами обусловлено дискретным характером регистрируемых изображений, что даёт возможность применять хорошо разработанные цифровые методы обработки информации [6].

Данный метод бесконтактного определения геометрических параметров физических тел на основе использования точечно структурированного излучения может найти широкое применение в различных областях современной техники, в частности, при создании автоматизированной системы бесконтактной классификации объектов по заданным параметрам и при восстановлении формы и размеров тела по его проекциям.

6. Литература

[1] Блейкер А. Применение фотографии в науке. М.: Мир, 1980. – 248 с.

[2] Сойфер В.А., Безус Е.А., Быков Д.А., Досколович Л.Л., Ковалев А.А. и др. Дифракционная оптика и нанофотоника. М.: Физматлит, 2014. – 608 с.

[3] Джоунс Р., Уайкс К. Голографическая и спекл-интерферометрия. Пер. с англ. – М.: Мир, 1986. – 328 с.

[4] Вин А.М., Павлов И.Н., Ринкевичюс Б.С. Визуализация геометрических параметров диффузно рассеивающих плоских тел с помощью структурированного оптического излучения: Научная визуализация, 2017, Том 9, № 2. с. 43 – 48.

[5] Вин А.М., Павлов И.Н., Ринкевичюс Б.С. Оптико-электронный комплекс для визуализации формы шероховатой поверхности с помощью структурированного оптического излучения: Труды XIV Международной научно-технической конференции. [Электронный ресурс] – М. Издательство Перо, 2017. с. 389 – 393.

[6] Rinkevichyus B.S., Evtikhieva O.A., Raskovskaya I.L. Laser Refractography. New York, Springer, 2011. – 189 p.

[7] Методы компьютерной обработки изображений/ Под ред. В.А.Сойфера - М.: Физматлит. 2001.-784 с.-ISBN 5-9221-0180-3.

[8] Salvi J., Armangué X., Batlle J. A comparative review of camera calibrating methods with accuracy evaluation // Pattern recognition. – 2002. – Vol. 35. – №. 7. – Pp. 1617-1635.

[9] Otsu N. A threshold selection method from gray-level histograms // IEEE transactions on systems, man, and cybernetics. –1979. –Vol. 9. –№. 1. – Pp. 62-66.

[10] Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: ТЕХНОСФЕРА, 2012. 1104 с.

[11] Шапиро Л., Стокман Д. Компьютерное зрение. БИНОМ. Лаб. знаний, 2006. 752 с.

Об авторах

Аунг Мьо Вин, аспирант кафедры физики им. В.А. Фабриканта Национального исследовательского университета «МЭИ», humburger49@gmail.com.

Ринкевичюс Бронюс Симович, д.ф.-м.н., профессор кафедры физики им. В.А. Фабриканта Национального исследовательского университета «МЭИ»,rinkevbs@mail.ru.

Поройков Антон Юрьевич, к.т.н., доцент кафедры физики им. В.А. Фабриканта Национального исследовательского университета «МЭИ», poroykovay@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Кластерный анализ вычислительных задач системы управления загрузкой эксперимента ATLAS на LHC с использованием визуальной аналитики

Т.П.Галкин[1], М.А.Григорьева[2,3], А.А.Климентов[2], Т.А.Корчуганова[3], И.Е.Мильман, В.В.Пилюгин[1], М.А.Титов[2]
z@wqc.me | magsend@gmail.com | alexei.klimentov@cern.ch | tatiana.korchuganova@cern.ch |
igal.milman@gmail.com | VVPilyugin@mephi.ru | mikhail.titov@cern.ch
[1]Национальный исследовательский ядерный университет «МИФИ», Москва, Россия;
[2]Национальный исследовательский центр «Курчатовский институт», Москва, Россия;
[3]Национальный исследовательский Томский политехнический университет, Томск, Россия.

При проведении экспериментов на научных установках, таких как LHC, RHIC, KEK, для решения задач в области физики высоких энергий (ФВЭ) и ядерной физики (ЯФ) получены сотни петабайт экспериментальных данных. По мере модернизации ускорителей (повышение энергии и светимости), объемы данных стремительно растут и достигли эксабайтной отметки, что также способствует увеличению количества выполняемых задач анализа и обработки данных, непрерывно конкурирующих между собой за вычислительные ресурсы. Последнее обуславливает повышение производительности вычислительной инфраструктуры привлечением высокопроизводительных вычислительных ресурсов, образуя гетерогенную распределённую вычислительную среду (сотни распределённых вычислительных центров). При распределённой модели обработки и анализа данных, оптимизация работы систем управления данными и загрузкой становится фундаментальной задачей, и отсутствие её своевременного решения приводит к экономическим, функциональным и временным потерям. Данная работа описывает первый этап исследований, направленных на решение задачи повышения стабильности и эффективности функционирования распределённых систем обработки данных экспериментов класса мега-сайенс с использованием методов визуальной аналитики. На примере обработки данных эксабайтного диапазона в эксперименте ATLAS на LHC будет продемонстрирована возможность использования визуальных методов для кластерного анализа вычислительных задач/заданий системы управления загрузкой PanDA. При этом будут исследованы и графически интерпретированы взаимозависимости и корреляции между различными параметрами упомянутых задач/заданий в N-мерном пространстве с использованием трёхмерных проекций. Визуальный анализ позволит выявлять схожие (подобные) задачи, а также аномальные задачи, при этом определять, чем обусловлена аномальность. Дальнейшее развитие работ в данном направлении будет основано на увеличении количества анализируемых вычислительных задач и разработке соответствующего программного инструментария.

Ключевые слова: *визуальная аналитика, физика высоких энергий, ядерная физика, эксперимент ATLAS, кластерный анализ.*

Visual Cluster Analysis for Computing Tasks at Workflow Management System of the ATLAS Experiment at the LHC

T.P.Galkin[1], M.A.Grigorieva[2,3], A.A.Klimentov[2], T.A.Korchuganova[3], I.E.Milman, V.V.Pilyugin[1], M.A.Titov[2]
on behalf of the ATLAS Collaboration
z@wqc.me | magsend@gmail.com | alexei.klimentov@cern.ch | tatiana.korchuganova@cern.ch |
igal.milman@gmail.com | VVPilyugin@mephi.ru | mikhail.titov@cern.ch
[1]National Research Nuclear University “MEPHI”, Moscow, Russian Federation;
[2]National Research Center “Kurchatov Institute”, Moscow, Russian Federation;
[3]National Research Tomsk Polytechnic University, Tomsk, Russian Federation.

Hundreds of petabytes of experimental data in high energy and nuclear physics (HENP) have already been obtained by unique scientific facilities, such as LHC, RHIC, KEK. As the accelerators are being modernized (energy and luminosity are increasing), data volumes are rapidly growing and have reached the exabyte scale, that also leads to an increasing the number of analysis and data processing tasks, that are competing continuously for computational resources. The increase in processing tasks causes an increase in the performance of the computing infrastructure by the involvement of high-performance computing resources and forming a heterogeneous distributed computing environment (hundreds of distributed computing centers). With a distributed model of data processing and analysis, the optimization of data management and workload systems becomes a critical task, and the absence of an adequate solution leads to economic, functional and time losses. This work describes the first stage of the study aiming at solving the task of increasing the stability and efficiency of workflow management systems for mega-science experiments by using visual analytics methods. Using the case of the ATLAS experiment at LHC the visual methods for cluster analysis of the workload management system computing tasks/jobs will be applied. The interdependencies and correlations between various tasks/jobs parameters will be investigated and graphically interpreted in N-dimensional space using 3D projections. Visual analysis allows to identify similar jobs, as well as anomalous jobs, and to determine what causes such anomaly. A further evolvement of the work in this direction will be focused on the increasing the amount of analysed computing jobs and the development of the appropriate infrastructure.

Keywords: *visual analytics, high-energy physics, nuclear physics, ATLAS experiment, cluster analysis.*

1. Введение

Научные установки, используемые для решения задач в области физики высоких энергий (ФВЭ) и ядерной физики (ЯФ), генерируют огромные объёмы данных. Эти научные области одними из первых столкнулись с необходимостью анализа и обработки эксабайтных объёмов данных и сопутствующих метаданных.

Стремительное увеличение и усложнение распределённой вычислительной инфраструктуры современных научных экспериментов в области ФВЭ и ЯФ и экспоненциальный рост объёмов обрабатываемых данных обусловили появление новых задач, решение которых

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

затруднено или невозможно без визуальной аналитики. Исторически, научные эксперименты в области ФВЭ и ЯФ использовали различные пакеты визуализации для представления данных и решения различных классов задач: моделирование работы детектора, анализ событий, представление результатов исследований для обмена информацией в научном сообществе (HBOOK[6], PAW[5], ROOT[3], MINUIT[7], Ganglia[9], GEANT[2]).

В представленной работе исследуются данные эксперимента ATLAS [1] на LHC - крупнейшего эксперимента в области ФВЭ и ЯФ. Информация, накопленная за многие годы работы системы обработки данных эксперимента ATLAS (ProdSys2 / PanDA [4,8]), содержит данные о ходе выполнения более чем 10 миллионов заданий и порядка 3000 миллионов задач (см. подробнее о заданиях и задачах в разделах 2 и 3). Существующие программно-аппаратные средства позволяют осуществлять контроль, мониторинг и оценку многих параметров и метрик в реальном времени. Однако текущая инфраструктура мониторинга не имеет инструментов оценки корреляций между многочисленными свойствами объектов, в том числе для анализа временных задержек при выполнении вычислительных задач в распределённой компьютерной среде.

Для решения этих проблем будут применены методы визуальной аналитики для поиска новых (неявных) знаний об объектах и обеспечения эффективного взаимодействия с данными, соответствующего человеческим когнитивным системам при обработке сложной информации.

В данной статье описывается первый этап работ по применению визуальной аналитики к исследованию функционирования систем управления загрузкой эксперимента ATLAS: кластерный анализ вычислительных задач с использованием визуально-аналитических методов. Этот анализ позволит пользователю визуально интерпретировать наиболее близкие по параметрам задачи с использованием 3х-мерных проекций, отслеживая при этом корреляции различных комбинаций параметров.

2. Система управления загрузкой эксперимента ATLAS

Для обработки больших объемов данных, эксперименты на LHC используют вычислительную Грид инфраструктуру Worldwide LHC Computing Grid (WLCG)[1], а также ресурсы облачных вычислений и суперкомпьютеры. Система обработки данных эксперимента ATLAS второго поколения (Production System - ProdSys2) предназначена для выполнения комплексных вычислительных приложений, и характеризуется следующими показателями: в среднем выполняется 350К задач в день в более 200 вычислительных центрах (более 250К узлов) тысячами пользователей [4]. Основные составляющие ProdSys2:

- Интерфейс запросов (Request Interface). Отвечает за определение параметров запуска вычислительных заданий и/или группы заданий в форме соответствующего запроса.
- DEfT (Database Engine for Tasks). Данный компонент формирует отдельные задания (tasks), цепочки заданий (пример этапов выполнения: генерация, симуляция, реконструкция и т.д.) или группы заданий на основе установленных параметров в запросе.
- JEDI (Job Execution and Definition Interface). Компонент отвечает за управление полезной нагрузкой на уровне заданий; в основе данного решения лежит динамическое создания задач (jobs) на основе сформированных заданий и управление их выполнением (позволяет оптимизировать загруженность ресурсов), включает в себя механизм принятия решений. Данный компонент также является частью системы PanDA.
- PanDA - система управления загрузкой гетерогенной вычислительной инфраструктуры, отвечает за выполнение задач (внутренний планировщик определяет порядок запуска задач и распределяет по ресурсам). Данная система разработана на основе концепции пилотных задач (pilots / pilot jobs), обеспечивающих контроль состояния ресурсов и назначение рабочей нагрузки (т.н. “поздняя привязка” полезной нагрузки).

3. Исследуемые метрики вычислительных задач системы управления загрузкой

Для запуска цикла обработки данных, который представляет собой конвейер заданий по трансформации и анализу данных, менеджер группы формирует запрос к системе ProdSys2. Запрос преобразуется в набор заданий (tasks) и задач (jobs), которые распределяются системой управления загрузкой PanDA на ресурсах Грид инфраструктуры. Каждое задание разбивается на вычислительные задачи (от нескольких единиц до нескольких тысяч). На первом этапе исследований анализируются визуальные трёхмерные проекции задач, принадлежащих определенным заданиям. Для описания вычислительных задач были выбраны следующие пробные параметры, характеризующие потребление ресурсов:

- Идентификатор задания в системе ProdSys2 / PanDA: ID (integer)
- Время выполнения задачи: duration = endtime - starttime (integer)
- Объём входных данных для задачи: inputFileBytes (integer)
- Объём выходных данных задачи: outputFileBytes (integer)
- Эффективность процессора (отношение общего процессорного времени к произведению времени выполнения запроса на количество ядер): CPU eff per core (integer)
- Потребление процессорного времени: CPU consumption (integer)
- Средний объём общей памяти, который хранится в ОЗУ для процесса: avgRSS (integer)
- Средний размер потребляемой оперативной памяти, в котором учитывается, что страницы памяти могут быть разделены между несколькими процессами: avgPSS (integer)
- Средний размер выделенной виртуальной памяти: avgVMEM (integer)

4. Метод визуальной аналитики для кластерного анализа

Параметры вычислительных задач (jobs) системы ProdSys2 / PanDA можно представить в виде многомерных табличных данных. Тогда каждую строку этой таблицы можно поставить в соответствие с точками в многомерном пространстве E_n, координатами которых являются нормированные значения параметров: $p_i = (p_i^1, p_i^2, \dots, p_i^n) \in E_n$. В данной работе в качестве метрики различия задач выбрано Евклидово расстояние $d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$

[1] The Worldwide LHC Computing Grid, http://wlcg.web.cern.ch

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

между точками $p(p_1, \dots p_n)$ и $q(q_1, \dots q_n)$. Для анализа расстояния между точками n-мерного пространства предлагается использовать визуальное отображение этих точек. Нормализация всех исследуемых значений параметров выполняется по следующей формуле: $x_n = (x - x_{min})/(x_{max} - x_{min})$, где x_{min}, x_{max} - минимальное и максимальное значения параметра x соответственно.

Далее осуществляется проецирование исходного множества точек на одно из трёхмерных пространств. Многомерная точка p_i проецируется в сферу S_i (при этом все координаты, кроме 3-х используемых, приравниваются к нулю). Для визуального представления связных объектов вводится некоторое пороговое расстояние d, задаваемое в интерактивном режиме. Если расстояние между точками n-мерного пространства меньше d, то эти точки соединяются цилиндром, цвет которого меняется от красного (малое расстояние) до синего (расстояние, близкое к d).

На следующем шаге выполняется графическое проецирование сфер и цилиндров на картинную плоскость с последующим их визуальным анализом. Результирующая совокупность сфер и цилиндров образует пространственную сцену с заданной геометрией и оптическими (цветовыми) характеристиками.

5. Программная реализация метода визуальной аналитики для кластерного анализа

Для визуализации и решения задачи анализа используется программа IVAMD [10]. Данная программа реализована на скриптовом языке MAXScript с дополнительным модулем, написанным на языке программирования C#. Основной функционал программного средства включает в себя: отображение пространственной сцены с использованием пользовательских параметров визуализации (порогового расстояния d, радиусов сфер и цилиндров, трёхмерного пространства для проекции), проведение аффинных преобразований трёхмерного пространства, расчёт расстояния в исходном n-мерном пространстве, разбиение на кластеры (обозначаются с использованием различных цветов), проведение микроанализа пространств. При микроанализе, а именно, анализе удаленных точек, важным является то, какие именно координаты вносят больший вклад в расстояние — происходит ли это за счёт всех координат или за счёт большого отличия только нескольких координат. Для определения этого строятся графические проекции исходного множества на плоскости (x_i, x_i) и затем просматриваются все эти проекции при различных i. Результаты кластерного анализа можно получить в виде таблицы, где строки отмечены цветом в зависимости от заданного цвета кластера.

В силу особенностей метода, необходимое для решения задачи (n строк, m столбцов) количество объектов для визуализации: n сфер (равно количеству строк), количество цилиндров равно $n \times (n-1)/2$. Итого $n \times (n+1)/2$, т.е., для отображения 100 строк необходимо отобразить около 5000 объектов. Описываемое программное средство является прототипом реализации метода и имеет ограничение на количество обрабатываемых объектов. Дальнейшая разработка и усовершенствование прототипа, и оптимизация в рамках высокопроизводительной программно-аппаратной инфраструктуры позволит устранить текущие ограничения.

6. Применение метода визуальной аналитики

В качестве тестового испытания был выполнен визуальный анализ вычислительных заданий, состоящих из множества задач. Например, задание №12196428 из данных за 2017-10-02 состоит из 74 задач. Первым шагом был проведен макроанализ всего подпространства, т.е. всё исходное 9-и мерное пространство значений было кластеризовано. Трёхмерная модель выполнена на проекции трех параметров: avgPSS, duration, outputFileBytes (представлено на рисунке 1). В результате выделено 6 кластеров (мощностью 31, 17, 8, 5, 3 и 2), а также 2 одиночные точки (кластеры мощностью 1).

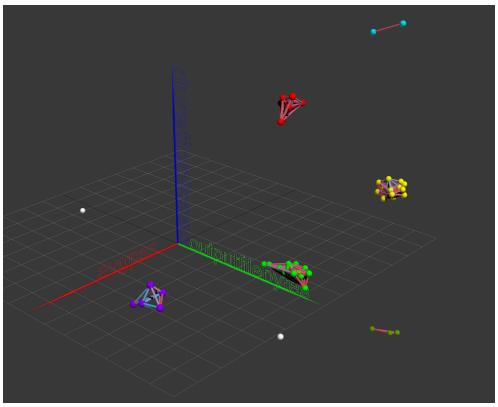


Рис. 1. Разбиение исходных точек на кластеры в трёхмерном пространстве (красная ось - avgPSS, зелёная ось - outputFileBytes, синяя ось - duration in sec).

Далее был проведен микроанализ для определения вклада различных параметров в разбиение на кластеры, а также оценки влияния различных параметров объектов на продолжительность выполнения вычислительных задач (duration).

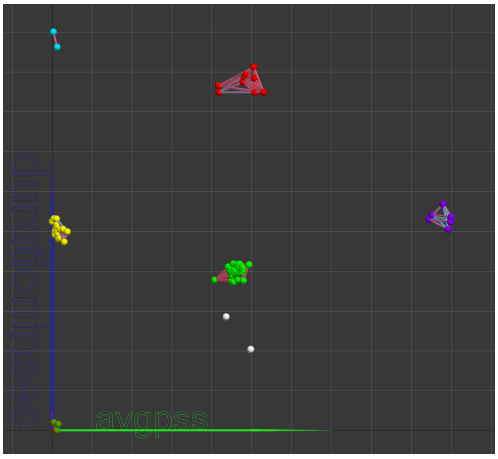


Рис. 2. Графическая проекция на плоскость (avgPSS, duration).

На рисунке 2 можно увидеть, что при одинаковом avgPSS, различие в duration у красного и зелёного кластера значительны. Аналогичная картина наблюдается для параметров avgRSS и avgVMEM, что позволяет сделать вывод о том, что перечисленные параметры вносят вклад в длительность выполнения задач, но только при средних значениях. Данная зависимость требует дополнительного исследования на большем количестве точек.

Рассмотрим кластеры в другом подпространстве. На рисунке 3 видно, что различие в потреблении процессорного времени (CPU Consumption) прямо пропорционально влияет на duration, как и ожидалось.

На рисунке 4 видно, что степень влияния объёма входных файлов (inputFileBytes) на длительность выполнения задач обработки и анализа данных не является решающей при распределенной обработке. Это также

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

применимо и к выходным данным/файлам (outputFileBytes). Для количественной оценки влияния объёма данных на длительность выполнения задач требуются дополнительные исследования на большей статистике, которые будут проведены в дальнейшем.

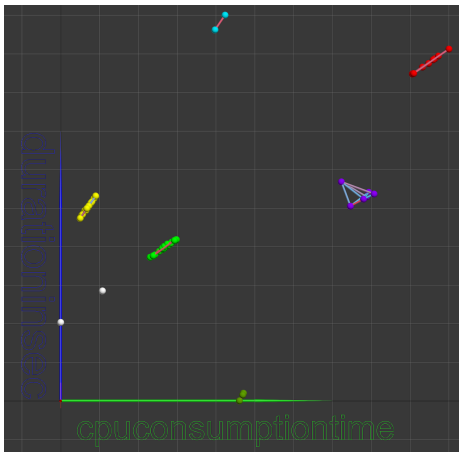


Рис. 3. Графическая проекция на плоскость (cpuConsumptionTime, duration).

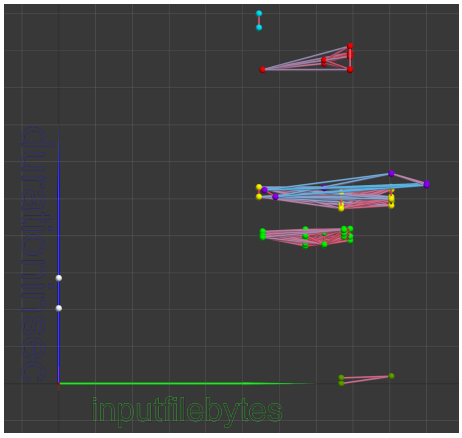


Рис. 4. Графическая проекция на плоскость (inputFileBytes, duration).

7. Заключение

Полученные результаты кластеризации задач с использованием разработанного метода и инструментария являются первым этапом работ по визуальной аналитике данных системы управления загрузкой эксперимента ATLAS. Стоит отметить, что данная работа является новаторской в рамках анализа данных системы ProdSys2 / PanDA. В ходе дальнейших работ планируется построение модели многоуровневой интерактивной визуальной кластеризации. При этом будут анализироваться целые классы задач и заданий анализа и обработки данных эксперимента, на которых будут опробованы различные методы кластеризации, включая методы машинного обучения. Разработанный инструментарий для визуализации будет усовершенствован в сторону обеспечения интерактивного переключения между кластерами различных уровней, обеспечив тем самым удобный для исследователя метод интерпретации результатов анализа разной степени детализации данных.

Предполагается, что полученные результаты могут быть использованы для визуального мониторинга системы управления загрузкой эксперимента ATLAS, для разработки рекомендаций по оптимизации времени выполнения задач анализа и обработки данных.

8. Благодарности

Работа выполнена при поддержке гранта РНФ №18-71-10003 от 02.08.2018г.

9. Литература

[1] The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider // Journal of Instrumentation, vol. 3, S08003, 2008.

[2] S Agostinelli et al. Geant4 - a simulation toolkit // Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 506, no. 3, pp. 250--303, 2003.

[3] I Antcheva et al. ROOT - A C++ framework for petabyte data storage, statistical analysis and visualization // Computer Physics Communications, vol. 180, no. 12, pp. 2499--2512, 2009.

[4] F H Barreiro et al. The ATLAS Production System Evolution: New Data Processing and Analysis Paradigm for the LHC Run2 and High-Luminosity // Journal of Physics: Conference Series, vol. 898, no. 5, 2017.

[5] R Brun, O Couet, C Vandoni and P Zanarini. PAW, a general-purpose portable software tool for data analysis and presentation // Computer Physics Communications, vol. 57, no. 1, pp. 432--437, 1989.

[6] R Brun and P Palazzi. Graphical Presentation for Data Analysis in Particle Physics Experiments: The HBOOK/HPLOT Package // Proceedings Eurographics '80, pp. 93--104, 1980.

[7] F James and M Roos. Minuit - a system for function minimization and analysis of the parameter errors and correlations // Computer Physics Communications, vol. 10, no. 6, pp. 343--367, 1975.

[8] A Klimentov et al. Migration of ATLAS PanDA to CERN // Journal of Physics: Conference Series, vol. 219, no. 6, 2010.

[9] M Massie, B Chun and D Culler. The ganglia distributed monitoring system: design, implementation, and experience // Parallel Computing, vol. 30, no. 7, pp. 817--840, 2004.

[10] D Popov, I Milman, V Pilyugin and A Pasko. A solution to a multidimensional dynamic data analysis problem by the visualization method // Scientific Visualization, vol. 8, no. 1, pp. 45--47, 2016.

Об авторах

Галкин Тимофей Петрович, инженер, Национальный исследовательский ядерный университет "МИФИ", email: z@wqc.me

Григорьева Мария Александровна, к.т.н., старший научный сотрудник, Национальный исследовательский центр "Курчатовский институт", email: magsend@gmail.com

Климентов Алексей Анатольевич, к.ф-м.н., начальник лаборатории, Национальный исследовательский центр "Курчатовский институт", email: alexei.klimentov@cern.ch

Корчуганова Татьяна Александровна, инженер, Национальный исследовательский Томский политехнический университет, email: tatiana.korchuganova@cern.ch

Мильман Игаль Евгеньевич, email: igal.milman@gmail.com

Пилюгин Виктор Васильевич, к.т.н., профессор, Национальный исследовательский ядерный университет "МИФИ", email: VVPilyugin@mephi.ru

Титов Михаил Анатольевич, PhD, научный сотрудник, Национальный исследовательский центр "Курчатовский институт", email: mikhail.titov@cern.ch

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Модель визуализации изменений в графе внутреннего представления программ языка Cloud Sisal

Д.С. Гордеев
gds@iis.nsk.su
Институт систем информатики СО РАН, Новосибирск, Россия

В работе описывается решение задачи визуализации изменений в графовой модели внутреннего представления программ с целью отражения процессов, происходящих с графовой моделью в процессы вычислений или обработки с помощью графовых алгоритмов. Для решения задачи применяется модель, описывающая изменения в графовой модели как графические анимации, отображающие изменения в визуальных стилях, которые соответствуют изменениям в атрибутах графовой модели. Каждое изменение атрибута, добавление вершины или удаление ребра отражается на визуальном отображении анимацией графических стилей. Ключевым моментом модели является, выделение понятия контекста графической анимации, как визуального отражение внешнего окружения по отношению к происходящему изменению.

Ключевые слова: *визуализация алгоритмов, анимация алгоритмов, динамический граф.*

Visualization of changes in Cloud Sisal program internal representation graph

D.S. Gordeev
gds@iis.nsk.su
Institute of Informatics Systems SB RAS, Novosibirsk, Russia

This paper describes the solution of problem of visualization of changes in graph model of internal representation of programs in order to reflect processes which happens during calculation of programs or processing with graph algorithms. We use model which interprets every change of graph model as corresponding change of visual graphical styles, which corresponds to changes of attributes of graph model elements or to structure of graph model. Each single changing of an attribute value, adding of node to graph model or removing of edge from graph model is reflected over visual representation as animation of graphical styles of pieces of image. The key point of described model is an idea of context of visual animation as surrounding visual information in relation to happening change in graph model.

Keywords: *algorithm visualization, algorithm animation, dynamic graph.*

1. Введение

В рамках развития проекта облачной системы параллельного программирования CPPS [8], целью которой является предоставление возможности решения вычислительно-ёмких задач с помощью доступных по сети супервычислителей, поддерживаемых системой CPPS, широкому кругу пользователей без прямого доступа к вычислителям большой мощности, но имеющих выход в Интернет, где пользователи создают и отлаживают программы на языке Cloud Sisal [4] с помощью вебприложения в одном из стандартных браузеров, а решение вычислительных задач осуществляется на поддерживаемом системой супервычислителе после предварительной адаптации с помощью настраиваемого оптимизирующего компилятора из комплекта системы CPPS. При этом возникает задача поддержки процесса программирования и отладки исполнения Cloud Sisal программы на конкретном вычислительном оборудовании, при условии, что код программы инструментирован для получения соответствующей информации о ходе вычислений.

При трансляции программы на языке Cloud Sisal генерируется граф внутреннего представления IR входной программы, который представим в виде иерархического ориентированного графа [3]. При отладке программы полезно иметь визуальное отображение исполнения для понимания движения данных по графу, и на каких узлах вычислителя осуществляются вычисления тех или иных функций. При этом возникает задача отображения графа, и свойств вершин и дуг этого графа с помощью графических примитивов, таких как многоугольники, ломаные линии, различных цветовых стилей, и визуального отображения изменений атрибутов вершин или дуг. Данная задача является нетривиальной, так как основным пользователем системы является человек, и его ожидания должны совпадать с наблюдением за отображением изменений, чтобы возникло корректное представление о том, как происходило вычисление в заданной программе.

Методы визуализации процесса вычислений зависят от решаемой задачи и её предметной области, в данном случае это предметная область графов потока данных, где результатом визуализации в том или ином виде является последовательность изображений состояния графовой модели со всеми графически отображаемыми атрибутами и структурой. Другими словами результатом являет динамический граф, и для визуализации динамических графов в основном известно два семейства техник визуализации: автономная (offline) и интерактивная (online) [9]. В первом случае строится последовательность изображений графа, при визуализации которой между двумя соседними изображениями используются техники интерполяции для получения промежуточного изображения, и набор таких промежуточных изображений описывает трансформацию между двумя соседними автономными изображениями графа. Во втором случае используется динамическое вычисление очередного изображения исходя из текущего состояния графовой модели и следующего события, которое следует графически отобразить.

В данной работе обсуждается реализация модели визуализации динамических графов для решение задачи отображения процесса вычислений программ на языке

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Cloud Sisal в рамках проекта развития облачной системы параллельного программирования CPPS.

2. Внутреннее представление IR

Текст Cloud Sisal программы транслируется внутреннее представление IR, которое описывается как иерархический ориентированный граф с портами. Формально такие графы можно описывать следующим образом. Пусть задан граф

$$\mathrm{G} = (\mathrm{V},\ \mathrm{P},\ F^v_{p,i},\ F^v_{p,o},\ \mathrm{E}), \quad (1)$$

где V это непустое конечное множество вершин, P это непустое конечное упорядоченное множество портов, представимое в виде объединения двух непересекающихся множеств P_i и P_o. $F^v_{p,i}$ это функция $F^v_{p,i} : P_i \to V$, $F^v_{p,o}$ это функция $F^v_{p,o} : P_o \to V$,. E это подмножество декартового произведения $< P_o \times P_i >$.

Множество P_i называется множеством входных портов, и с помощью функции $F^v_{p,i}$ позволяет говорить об упорядоченном множестве входных портов для каждой вершины из графа G. Аналогично множество P_o называется множеством выходных портов, и с помощью функции $F^v_{p,o}$ позволяет говорить об упорядоченном множестве выходных портов для каждой вершины из графа G.

В случае если фунции $F^v_{p,i}$ и $F^v_{p,o}$ являются тождественными, то каждая вершина имеет ровно один входной порт и ровно один выходной порт, и граф G эквивалентен классическому определению графа G′ = (V', V'), где E' есть подмножество декартового произведения $< V' \ x \ V' >$.

Для вершины v из V и ребра (p_o, p_i) из E будем говорить о инцидентности, если и только если $F^v_{p,i}(p_i) = v$ или $F^v_{p,o}(p_o) = v$. Для вершин v_1 и v_2 из V будем говорить о смежности вершин, если и только если существует ребро (p_o, p_i) из E такое, что ($F^v_{p,o}(p_o) = v_1$ и $F^v_{p,i}(p_i) = v_2$ или ($F^v_{p,o}(p_o) = v_2$ и $F^v_{p,i}(p_i) = v_1$).

Каждая вершина графа G внутреннего представления IR соответствует функции Cloud Sisal программы, по которой построен граф внутреннего представления. Входные порты P_i вершины v отвечают аргументам соответствующей функции, а выходные порты P_o вершины v отвечают результатам вычислений соответствующей функции. Дуга e в графе G, соединяющая выходной порт p^1_o и p^2_i, обозначает передачу данных из выходного порта p^1_o во входной порт p^2_i.

3. Графовая модель визуального представления внутреннего представления IR

При визуализации вершины графов обычно представляются геометрическими фигурами, такими как прямоугольники, окружности или более сложные многоугольники. Дуги часто представлены в виде ломаных или гладких кривых. Порты, в зависимости от применения изображения графа, могут отображаться с помощью кругов или более мелких квадратов, относительно вершины. Причём обычно порты располагаются на границах многоугольников, представляющих вершины. При записи решения задачи с помощью языка программирования часто функции вызывают другие функции. На графе внутреннего представления это отражается наличием вложенных графов, являющихся метками вершин [3]. Таким образом задаётся иерархия в графе внутреннего представления.

Также для определения стилей визуального отображения геометрических фигур требуется хранить информацию об атрибутах вершин и их значениях. Далее будем рассматривать графовую модель G', расширяющую определение графа G (1) атрибутами. Пусть

$$\mathrm{G}' = (\mathrm{V},\ \mathrm{P},\ F^v_{p,i},\ F^v_{p,o},\ \mathrm{E}, \mathrm{A}_v, \mathrm{A}_p, \mathrm{A}_e), \quad (2)$$

где $A_v : V \to \{(n_v,\ u_{n_v})\}$, где n_v это уникальное имя атрибута, а u_{n_v} это значение атрибута с именем n_v. Аналогично $A_p : P \to \{(n_p,\ u_{n_p})\}$, и $A_e : E \to \{(n_e,\ u_{n_e})\}$, . Для вершин, например, именем атрибута может быть “форма” со значениями “прямоугольник” или “окружность”, “цвет границы” со значениями цвета соответственно, “цвет фона” также со значениями цвета или “подграф” со значением графа, удовлетворяющего определению (2).

Визуальное изображение графа G' строится в зависимости от структуры графа, а также атрибутов и их значений. При наличии такой зависимости отражать ход вычислений на графе IR можно с помощью изменения структуры графа или изменения значений атрибутов. В этом случае можно говорить о динамическом графе [5] DG', где DG' это последовательность графов G'_i, где каждый G'_{i+1} граф получается из G'_i через преобразование $G'_{i+1} = \varphi(G'_i)$, где φ осуществляет теорерико-графовую операцию, например, удаление или добавление ребра или дуги, добавление или удаление вершины, а также изменение значения атрибута вершины, дуги или порта.

4. Поток изменений в графовой модели

Каждое изменение в графе DG', соответствующее преобразованию φ, можно описать как событие изменения [2] в графовой модели G'. Определим каждое такое событие ev как вектор $(el, n, v_1, v_2, t, ctx)$, где el это вершина, порт или дуга графа G', n это имя атрибута элемента el, v_1 это предыдущее значение атрибута с именем n, v_2 это новое значение атрибута с именем n, t это момент времени, когда произошло изменение значения атрибута с именем n, а ctx это вектор контекстной информации, представляющий собой пару (a, b), где a обозначает признак либо входа либо выхода из контекста, b это контекстная информация, например, множество вершин, множество дуг, или целый подграф входного графа.

Рассматриваются следующие события изменений:

1. Добавление вершины в граф. Например, это может соответствовать добавлению функции или выражения в текст программы или добавлению вершины во вспомогательные структуры данных, например стек или очередь.
2. Добавление порта к вершине. Например, это может соответствовать добавлению аргумента в список аргументов функции или добавлению дуги во вспомогательные структура данных.
3. Добавление дуги. Например, это может соответствовать использованию результатов функции в выражении.
4. Удаление вершины.
5. Удаление порта.
6. Удаление дуги.
7. Изменение значения атрибута вершины, порта или дуги.

Кроме того рассматриваются следующие события:

1. Вычисление предиката на вершине.
2. Вычисление предиката на порте.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

3. Вычисление предиката на дуге.
4. Перебор смежных вершин с предикатом или без.
5. Перебор смежных дуг с предикатом или без.
6. Получение предка по ориентированной дуге.
7. Получение смежные дуг с предикатом или без.
8. Получение смежные вершин с предикатом или без.
9. Получение элемента из множества по предикату.
10. Чтение значения атрибута вершины, дуги или порта.
11. Вход или выход из контекста.

Поток событий можно получить разными способами, например, это может быть информация о том, какая вершины вычислялась в какой момент времени, и по каким дугам передавались данные между портами. Также можно получать методом выполнения какого-то графового алгоритма, не теряя общности, когда алгоритм запускается на входном графе, в таком случае граф должен быть иерархическим ориентированным графов с портами, возможно несвязным, если в изображении используются вспомогательные структуры данных из алгоритма. Например, стек из алгоритма обхода в глубину.

Модель обрабатывает полученный поток событий изменений, сверяясь со своим внутренним состоянием, и преобразовывая каждое изменение в серию изменений графического состояния подсистемы вывода, о которой речь пойдёт в следующем разделе.

5. Визуальное представление графовой модели

В каждом программном проекте существует подсистема вывода информации. В рамках данной работы подсистема вывода является графическим интерфейсом, основанным на технологии SVG, которая поддерживается большинством современных браузеров. Обычно для изображения элементов графов используются различные геометрические фигуры. Для описания двумерной векторной графики существует язык SVG, который также является расширением языка XML [6]. Язык SVG поддерживает такие фигуры как эллипсы, прямоугольники и более сложные фигуры, задаваемые с помощью последовательностей точек. Данный язык поддерживается основными браузерами, что делает его подходящим для представления графов при визуализации в контексте реализации облачного сервиса. Будем строить изображение графов по следующим правилам. Вершины будем представлять с помощью прямоугольников с границей. Порты будем представлять с помощью точек с координатами, заданными относительно содержащих их вершин. Порты вершин будем располагать только на верхних или нижних гранях прямоугольников, изображающих вершины. Координаты вершин также задаются относительно содержащих их вершин-графов. Дуги будем представлять с помощью ломаных линий или кривых Безье. Последние поддерживаются форматом SVG естественным образом [6]. Кроме того, следует отметить, что при использовании SVG изображений в браузерах, изображения естественным образом включаются в DOM модель документа [7]. Данное обстоятельство позволяет использовать стандартные методы для обеспечения интерактивного взаимодействия с построенным изображением. Например, это позволяет с помощью стандартных браузерных событий реализовать сворачивание и разворачивание вершин графа, что облегчает навигацию по изображению. Также естественное включение в DOM позволяет использовать существующие средства для поиска вершин в графе по заданным значениям атрибутов, например, по имени функции.

На рис. 1 представлен пример изображения внутреннего представление IR программы языка Cloud Sisal с помощью технологии SVG [1].

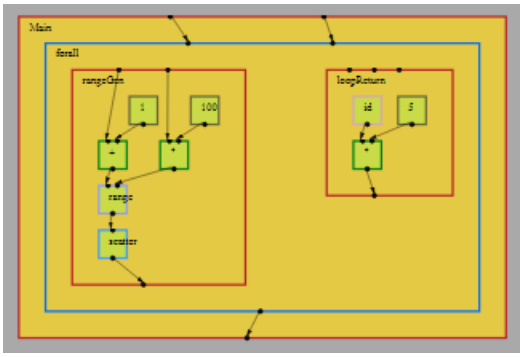


Рис. 1. Граф внутреннего представления программы, содержащей цикл.

6. Визуальные эффекты и отображение контекстной информации

Визуальные эффекты предназначены для моделирования графического отображения изменений в графовой модели. Простой визуальный эффект можно описать как непрерывную функцию F: $[0,1] \to S$, где S это множество векторов (l, w, c), описывающих визуальное состояние геометрических линий, где l это вектор точек, задающий форму линии, w это толщина линии, а c это цвет линии.

Составные визуальные эффекты представляют собой последовательности простых эффектов, где каждый следующий эффект запускается после предыдущего. Таким образом составной эффект можно описать как функцию:

$$F^c(t) = \begin{cases} F_1^S(t * 2), & 0 \le t \le \frac{1}{2} \\ F_2^S(t * 2), & \frac{1}{2} < t \le 1 \end{cases}, \quad (1)$$

в случае, если составной эффект состоит из двух эффектов, и первый из них действует половину времени, а второй оставшуюся половину времени.

В более общем случае эффект может быть задан следующим образом:

$$F^c(t) = \begin{cases} F_1(t * |d_1|), & t \in d_1 \\ F_2(t * |d_2|), & t \in d_2 \\ \dots \\ F_n(t * |d_n|), & t \in d_n \end{cases}, \quad (2)$$

где $d_1, d_2, \dots, d_n$ это интервалы, в объединении покрывающие отрезок $[0.1]$.

Таким образом, если требуется отображать изменение значения атрибута N вершин от значения A до значения B в течение, например, двух секунд, то для этого можно использовать функцию $F_N(t)$, где $F_N(0) = A$, а $F_N(1) = B$, причём для получения промежуточных значений использовать выражение $F_N(x/2000)$, где x меняется от 0 до 2000 миллисекунд.

При отображении изменений значений атрибутов элементов графовой модели, для которого используются функции визуальных эффектов, также полезно отображать контекстную информацию, соответствующую контексту, где происходит изменение значения атрибута. Так, например, при перечислении смежных вершин для заданной вершин происходит последовательный перебор элементов из множества смежных вершин, причём для

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

каждой отдельно итерации перебора, множество смежных вершин является контекстной информацией.

Для отображения контекстной информации также применимы визуальные эффекты с тем замечанием, что контекстная информация отображается графически в течение всего времени жизни контекста. Например, множество смежных вершин графически выделено всё время, пока производится перебор смежных вершин.

Другими примерами контекстной информации могут служить отображение смежной вершины из множества в рамках итерации цикла, который осуществляет перебор, или отображение подграфа, на котором производится вычисление внутри подпрограммы, если программа использует подпрограммы для осуществления вычислений.

7. Реализация модели

Реализация предложенной модели визуализации предназначена для использования в качестве компонента подсистемы отладки в системе CPPS, где, при интеграции с компонентом, интерпретирующем отлаживаемую Cloud Sisal программу, компонент визуализации графовой модели использует интерпретатор в качестве источника событий, отражающих изменения в графовой модели внутреннего представления Cloud Sisal программы. Предложенная модель позволяет отображать поток событий, отражающих изменения графовой модели, в множество взаимозависимых анимаций над визуальными образами элементов графа, а также осуществлять приостановку анимаций, и запуск анимаций в обратном направлении. На момент написания реализация модели находится на стадии технического демо, представляющего собой автономное вебприложение, реализованного на языке JavaScript, где вместо потока событий от интерпретатора Cloud Sisal программы используется симуляция ожидаемых событий.

8. Заключение

Целью данной работы является описать решение задачи визуализации графовых алгоритмов с помощью модели, описывающей каждое изменение в графовой модели под воздействием алгоритма, как графический эффект, отображающий изменения в визуальных стилях, которые соответствуют изменениям в атрибутах графовой модели. Каждое изменение атрибута, добавление вершины или удаление ребра отражается на визуальном отображении анимацией графических стилей. При графическом отображении графов обычно требуется тем или иным способом решить задачу определения размеров вершин графа и укладки графа на плоскость, представляя вершин геометрическими фигурами, а дуги ломаными или гладкими кривыми. В случае, если при графическом отображении изменений в графовой модели изменения соответствуют добавлению вершин, портов или дуг, то задачу укладки приходится решать на лету. Существенной трудностью здесь является сохранение относительного расположения вершин после удаления, так как алгоритм укладки может существенно изменить финальное графическое изображение, передвинув оставшиеся вершины в неожиданные места. Для практического применения существует метод, который существенно упрощает реализацию модели.

Графовая модель внутреннего представления модифицируется добавлением служебного атрибута ко всем элементам графа, обозначающим «видимость». Если значение атрибута задано, то вершина должна быть отображена графически, если не задано, то соответственно вершина должна быть невидима. Также меняется логика вычислений над множествами вершин, портов и дуг графа так, что, если значение атрибута видимости задано, то вершина при операциях перечисления элементов в множестве должна быть доступна для перечисления, а если значение не задано, соответственно, вершин должна быть недоступна для перечисления. Также меняется алгоритм укладки графой модели на плоскость, и вычисление будет проходить по следующей схеме:

1. Потребуется взять всю последовательность графов из динамического графа DG', объединить множества вершин всех элементов последовательности, а также множества портов и дуг.
2. Применить алгоритм укладки.
3. Пометить все элементы первого графа из последовательности DG' как видимые с помощью атрибута «видимости».

Соответственно, при отображении событий добавления или удаления вершин, вместо удаления проставляется или очищается значение атрибута видимости.

Данный метод позволяет решить задачу укладки более простым способом, чем перевычислять укладку при очередном удалении или добавлении вершин в графовую модель в случае, если операций удаления или добавления значительное количество, относительно общего числа изменений графовой модели.

9. Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект 18-11-00118).

10. Литература

[1] Гордеев Д.С. визуализация внутреннего представления программ на языке Cloud SISAL // Научная визуализация. - 2016.- Том. 8, N 2.- С.98 - 106.

[2] Demetrescu C., Finocchi I., Stasko J. T., Specifying Algorithm Visualizations: Interesting Events or State Mapping? // In Proc. of Dagstuhl Seminar on Software Visualization – Lect. Notes in Comput. Sci. – 2001. – P. 16–30.

[3] Касьянов, В. Н., Евстигнеев, В. А. Графы в программировании: обработка, визуализация и применение. – СПб.: БХВ-Петербург, 2003. – 1104 с.– ISBN 5-94157-184-4.

[4] Касьянов В. Н., Касьянова Е. В. Язык программирования Cloud Sisal. – Новосибирск, 2018. – 42 с. – (Препринт/ РАН, Сиб. отд-ние, ИСИ; N181).

[5] Zaki A. Comprehensive survey on dynamic graph models // IJACSA – 7(2). – 2016.

[6] http://www.w3.org/TR/SVG/

[7] http://www.w3.org/DOM/

[8] Касьянов В. Н., Касьянова Е. В. Методы и система облачного параллельного программирования // Проблемы оптимизации сложных систем: Материалы XIV Международной Азиатской школы-семинара - Алматы, 2018. - Часть 1. - С. 298-307.

[9] Federico P., Aigner W., Miksch S., Windhager F., Zenk L.: A visual analytics approach to dynamic social networks. In Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies - 2011. - p. 47.

Об авторе

Гордеев Дмитрий Станиславович, м.н.с лаборатории конструирования и оптимизации программ института систем информатики СО РАН.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Оптическая диагностика конвективных структур в вертикальном слое воды, ограниченном теплообменными поверхностями при нестационарных граничных условиях

Ю.Н. Дубнищев[1, 3], В.А. Арбузов[1, 3], Э.В. Арбузов[1, 2, 3], В.С. Бердников[1, 3], С.А. Кислицын[1, 3], О.С. Мелёхина[3]
dubnistchev@itp.nsc.ru | arbuzov@itp.nsc.ru | arbuzov@math.nsc.ru | berdnikov@itp.nsc.ru | 100pch@mail.ru [1, 3] | melexina-olga17@yandex.ru [3]
[1] Институт теплофизики им. С.С. Кутателадзе
Сибирского отделения Российской академии наук (ИТ СО РАН), Новосибирск, Россия;
[2]Институт математики Сибирского отделения Российской академии наук (ИМ СО РАН), Новосибирск, Россия;
[3] Новосибирский государственный технический университет (НГТУ), Новосибирск, Россия

Методами гильберт-оптики и сдвиговой интерферометрии визуализирован процесс эволюции конвективных структур в вертикальном слое воды, ограниченном плоскими теплообменными поверхностями при нестационарных граничных условиях. Восстановлено температурное поле. Решена обратная задача реконструкции интерферограмм и гильбертограмм из численной модели теплового поля в конвективном течении.

Ключевые слова: *оптическая диагностика, гильберт-оптика, сдвиговая интерферометрия, конвективные течения.*

Optical diagnostics of convective structures in a vertical water layer, limited by heat exchange surfaces under non-stationary boundary conditions

Yu.N. Dubnishchev[1, 3], V.A. Arbuzov[1, 3], E.V. Arbuzov[1, 2, 3], V.S. Berdnikov[1, 3], S.A. Kislitsin[1, 3], O.S. Melekhina[3]
dubnistchev@itp.nsc.ru | arbuzov@itp.nsc.ru | arbuzov@math.nsc.ru | berdnikov@itp.nsc.ru | 100pch@mail.ru[1, 3]| melexina-olga17@yandex.ru [3]
[1] Kutateladze Institute of Thermophysics
Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia;
[2] Sobolev Institute of Mathematics Siberian Branch of the Russian Academy of Sciences, *Novosibirsk, Russia*
[3] Novosibirsk State Technical University, Novosibirsk, Russia

Using the methods of Hilbert optics and shear interferometry, the evolution of convective structures in a vertical layer of water, limited by flat heat exchange surfaces under non-stationary boundary conditions, is visualized. The temperature field is restored. The inverse problem of reconstructing interferograms and Hilbertograms from the numerical model of the thermal field in a convective flow is solved.

Keywords: *optical diagnostics, Hilbert-optics, shear interferometry, convective currents.*

1. Введение

Методы оптической диагностики в значительной мере определяют уровень экспериментальных исследований в современной гидро- и газодинамике. Выбор конвективных течений в качестве объекта приложений оптической диагностики объясняется особой важностью конвекции в геодинамике, физике атмосферы и океана, в гидродинамических и теплофизических процессах с фазовыми переходами [1, 2].

Интерес к этим проблемам обострился в последнее время в связи с интенсификацией циклических процессов таяния и образования льдов в арктических и антарктических районах, оказывающих огромное влияние на мировой океан и климат. Экспериментальные и теоретические исследования конвективных процессов и динамики фазового перехода вода–лёд с учётом аномалии плотности воды в области 0÷4°С необходимы для многочисленных технических и технологических приложений [3]. Изучение динамики фазовых переходов в переохлаждённой жидкости и эволюции конвективных структур в системе вода–лёд требует применения невозмущающих методов диагностики, основанных на современных достижениях оптики, лазерной техники и информационных технологий. Необходимостью таких исследований мотивирована настоящая работа.

2. Метод и экспериментальная установка

Упрощённая схема исследовательского комплекса, содержащего систему оптической диагностики и экспериментальный стенд, показана на рисунке 1. Схема содержит осветительный модуль, состоящий из источника света *1* (ртутная лампа ДРШ–250), конденсорной линзы *2* и щелевой диафрагмы *3*, помещённой в передней фокальной плоскости объектива *4*, формирующего зондирующее поле в исследуемой среде. Такой средой является вода, заполняющая прямоугольную полость *5* с прозрачными стёклами оптического качества (рис. 2). Внутренние размеры полости 136×86×30 мм, боковые стенки образованы охлаждаемыми плоскими параллельными медными пластинами высокого качества *6* и *7*. Температура одной из пластин (T_1) регулируется и поддерживается термостатом в пределах до –28°С. Температура другой пластины (T_2) – в диапазоне от комнатной температуры до ~ 8,0°С. Внутри пластин имеются полости, в которые подаётся охлаждающая жидкость из термостатов *8* и *9*.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Термостаты управляются компьютером *10* по заданной программе. Трубки, подающие хладагент в теплообменные боковые пластины, обозначены цифрами *11–14*, трубки оттока – *15, 16*. Расположение подающих трубок и трубок оттока может меняться в соответствии с условиями эксперимента. Расстояния между трубками – 80 мм.

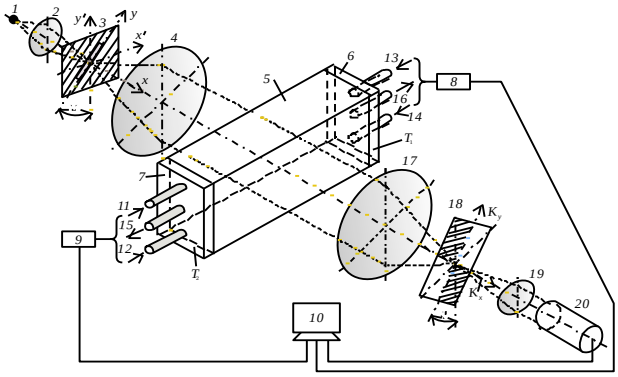


Рис. 1. Схема экспериментальной установки.

Объектив *17* формирует в частотной плоскости фурье-спектр возмущённого исследуемой средой зондирующего светового поля. Квадрантный гильберт-фильтр *18* помещён в фурье-плоскость объектива *17*. Частотная ось K_x гильберт-фильтра ортогональна направлению изображения щелевой диафрагмы светового источника (ψ=45°). Объектив *19* выполняет обратное фурье-преобразование фильтрован-ного пространственно-фильтрованного спектра фазовой структуры светового поля. Восстановленное объективом *19* изображение гильберт-образа фазовой структуры светового поля, возмущённого исследуемой средой, регистрируется цифровой видеокамерой *20*, подключенной к компьютеру *16*.

Рис. 2. Исследуемая прямоугольная полость с водой. Видны оптические трубки, подающие хладогент в теплообменные боковые пластины.

Оптическая система диагностики реализована на базе серийного теневого прибора ИАБ–463М [3] с модифицированными под задачи эксперимента модулями светового источника, пространственно-частотной фильтрации оптического сигнала, а также модулем регистрации изображения визуализированных фазовых возмущений светового поля, индуцированных исследуемой средой.

Зондирующее поле, сформированное объективом *4* от щелевого светового источника, проходит через исследуемую среду (воду), в которой боковыми верхней термостатированными поверхностями *6* и *7* задаются по заданной программе температурные граничные условия: T_1 и T_2, $T_1 < T_2$. В пространстве между термостатированными поверхностями возникают конвективные структуры, которые проявляются как возмущения полей оптической фазовой плотности. Эти структуры индуцируются рэлей-бенаровской конвекцией, осложнённой наличием фазового перехода и аномалией плотности воды в температурном интервале 0÷4°С.

Вблизи боковой термостатированной поверхности при росте градиента температуры вода становится переохлаждённой жидкостью, переходя в состояние неустойчивого равновесия. В такой среде возникает переход из жидкой фазы в твёрдокристаллическое состояние. Это – фазовый переход первого рода. Он проявляется возникновением волны кристаллизации и сопровождается энерговыделением, что в свою очередь влияет на динамическое распределение градиента оптической фазовой плотности в переохлаждённой воде и индуцирует фазовые возмущения в зондирующем световом поле, фурье-спектр которых $s(K_x, K_y)$ формируется в частотной плоскости объектива *17*.

Когерентная передаточная функция пространственно-частотного фильтра *18*, выполняющего одномерное преобразование Гильберта, описывается выражением:

$$H(K_x, K_y) = \left[e^{-i\varphi}\sigma(K_x) + e^{i\varphi}\sigma(-K_x)\right]\sigma(K_y) + \\ +\left[e^{i\varphi}\sigma(K_x) + e^{-i\varphi}\sigma(-K_x)\right]\sigma(-K_y) = \\ = \cos\varphi - i\sin\varphi\,\mathrm{sgn}\,K_x\,. \qquad (1)$$

где K_x и K_y – пространственные частоты; $\sigma(\pm K_x)$ и $\sigma(\pm K_y)$ – функции Хевисайда; φ – фазовый сдвиг, задаваемый соответствующим квадрантом пространственно-частотного фильтра. Фильтр с когерентно-передаточной функцией (1) выполняет одномерное преобразование Фуко–Гильберта.

В фурье-плоскости $H(K_x, K_y)$ пространственно-частотная ось K_x ортогональна изображению щелевого источника. Фурье-спектр светового поля непосредственно после фильтра:

$$s(K_x, K_y)H(K_x, K_y) = s(K_x, K_y)\,[(\cos\varphi - i\sin\varphi\,\mathrm{sgn}K_x)] = \\ = s(K_x, K_y)\,\cos\varphi + \hat{s}_x\,(K_x, K_y)\sin\varphi\,. \qquad (2)$$

Здесь $s(K_x, K_y)$ – пространственно-частотный фурье-спектр светового поля, возмущённого исследуемой средой; $\hat{s}_x(K_x, K_y) = -i\mathrm{sgn}(K_x)s(K_x, K_y)$ – фурье-спектр светового поля, подвергнутого одномерному преобразованию Гильберта по оси K_x.

Фазовый сдвиг φ является функцией длины волны λ зондирующего светового поля: φ=φ(λ).

При длине волны $\lambda=\lambda_0$, удовлетворяющей условию $\varphi(\lambda) = \pi/2$, фурье-фильтр 18 выполняет одномерное преобразование Гильберта:

$$\hat{s}_x(K_x, K_y) = -i\mathrm{sgn}(K_x)s(K_x, K_y) \qquad (3)$$

Если исследуемая среда индуцирует исключительно фазовые возмущения зондирующего поля, фурье-спектр возмущённого поля описывается выражением

$$s(K_x, K_y) = e^{i\psi(K_x,K_y)}.$$

Фильтрованный фурье-спектр фазовых возмущений:

$$s(K_x, K_y)H(K_x, K_y) = \\ = \left\{\sum_m \left[1 + i\tilde{\Psi}_{m_x}(K_x, K_y)\right]\right\}\,[-i\mathrm{sgn}K_x] = \sum_m \hat{\Psi}_{m_x}(K_x, K_y), \qquad (4)$$

где $\hat{\Psi}_{m_x}(K_x, K_y)$ – фурье-спектр фазовых возмущений, соответствующий гильберт-полосе *m*-го порядка. Здесь учтено, что гильберт-преобразование обладает квазидифференцирующими свойствами. Поэтому гильберт-образ отображает структуру градиентов поля фазовых возмущений.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Объектив *19* камкордера выполняет обратное фурье-преобразование фильтрованного фурье-спектра фазовых возмущений:

$$\sum_m \widehat{\Psi}_{m_x}(K_x, K_y) \leftrightarrow \sum_m \widehat{\Psi}_{m_x}(x, y) \quad (5)$$

Визуализированные посредством одномерного гильберт-преобразования фазовые структуры (4) регистрируются фотоматрицей камкордера.

Изображение одномерного гильберт-образа фазовых возмущений

$$\sum_m \left|\widehat{\Psi}_{m_x}(x, y)\right|^2$$

анализируется в системе координат (x', y'), повёрнутой относительно системы координат (x, y) на угол α=45°:

$$\sum_m \left|\widehat{\Psi}_{m_x}(x', y')\right|^2, \quad (6)$$

где

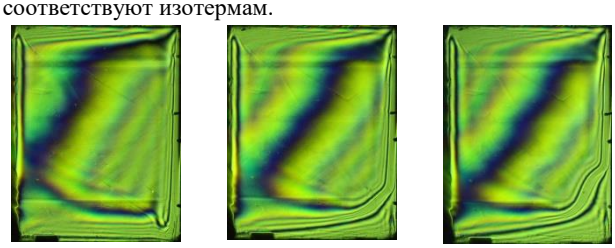
$$x' = \frac{1}{2}(x + y);\ y' = \frac{1}{2}(x - y)\,.$$

В случае широкополосного источника выполняется одномерное преобразование Фуко–Гильберта с когерентной передаточной функцией фильтра (1). Результатом фильтрации является трансформация поля фазовых возмущений в аналитический сигнал, являющийся суперпозицией фильтрованного сигнала и его гильберт-образа. Соответственно, сигнал, зарегистрированный фотоматрицей, представляет собой суперпозицию изображений фазового возмущения и его гильберт-образа. Регистрируемое изображение представляет собой структуру, состоящую из гильберт-полос, отображающих возмущения градиентов оптической фазовой плотности в исследуемой среде.

3. Результаты

На рисунке 3 показана выборка кадров из видеофильма, иллюстрирующих гильберт-визуализацию конвективных структур в слое пресной воды, ограниченном вертикальными плоскими теплообменными поверхностями.

На рисунке 4 представлен пример конвективных структур и фазового перехода в том же слое, визуализированных методами сдвиговой интерферометрии. Гильбертограммы и сдвиговые интерферограммы отображают экстремумы фазовых градиентов и соответствуют изотермам.



IMG–0503.jpg IMG–0510.jpg IMG–0518.jpg

Рис. 3. Кадры видеофильма, иллюстрирующие гильберт-визуализация конвективных структур и фазового перехода в воде.

Использование для аппроксимации полиномов Бернштейна позволяет построить сетку, адаптирующуюся к деформациям интерференционных линий. По полученной сетке проводится двумерная сплайн-интерполяция и восстанавливается температурное поле (рис. 6, 7). Рисунок 5 иллюстрирует аппроксимацию интерференционных полос кривыми Безье.

22.jpg 25.jpg 28.jpg

Рис. 4. Сдвиговые интерферограммы, визуализирующие конвективные структуры и фазовый переход в вертикальном слое пресной воды.

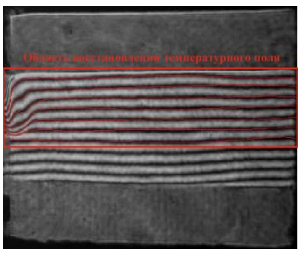


Рис. 5. Аппроксимация кривыми Безье интерференционных полос на сдвиговой интерферограмме.

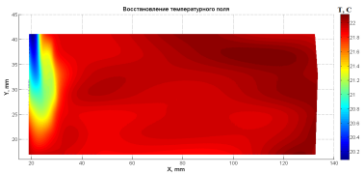


Рис. 6. Восстановленное температурное поле по сдвиговой интерферограмме (оси *x*–*y*).

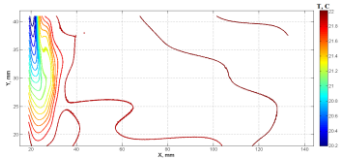


Рис. 7. Изотермы восстановленного температурного поля (оси *x*–*y*).

На рисунке 7 показан пример изотерм восстановленного температурного поля.

Аппроксимация интерференционных и гильберт-полос полиномами Бернштейна позволяют строить равномерную сетку, по которой производится двумерная сплайн-интерполяция и восстанавливается температурное поле.

В работе [4], в качестве верификации численной модели температурного поля, выполнено численное моделирование сдвиговых интерферограмм и гильберт-изображений по численной модели изотерм.

Гильбертограммы и сдвиговые интерферограммы, синтезированные из численных моделей изотерм, относящимся к моментам времени: a) – 122 с; b) – 305 с; c) – 710 с. после включения термостатов, управляющих температурой теплообменных поверхностей, представлены на рисунке 8.

Изотерма, соответствующая инверсии плотности воды (+4°C), выделена красным цветом. Она делит пространство на две области. Над изотермой (+4°C) возникает вихревое движение конвективных структур, направленное по часовой стрелке.

Под этой изотермой вихревое движение конвективных структур происходит против часовой стрелки. Сопряжённые вихри, сформированные над и под изотермой

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

(+4°C), переносят тёплую воду вниз, а охлаждённую воду – наверх.

На рисунке 9 показаны сдвиговые интерферограммы: a – экспериментально полученная; b – численно смоделированная.

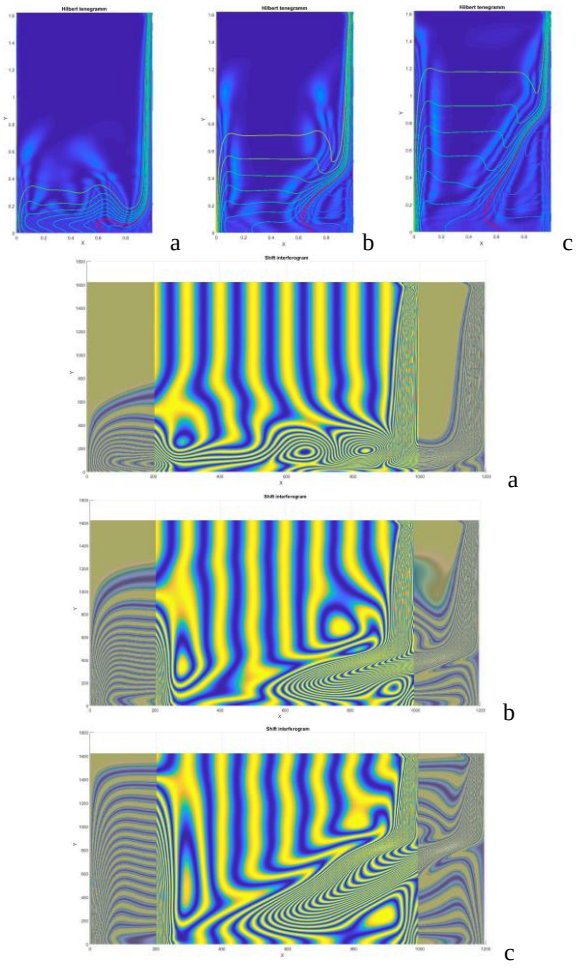


Рис. 8. Смоделированное поле изотерм, соответствующие ему гильбертограммы и сдвиговые интерферограммы (красная линия-изотерма +4°C).

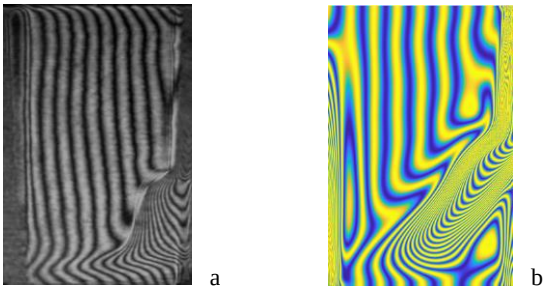


Рис. 9. Сдвиговые интерферограммы: a – экспериментальная; b –смоделированная.

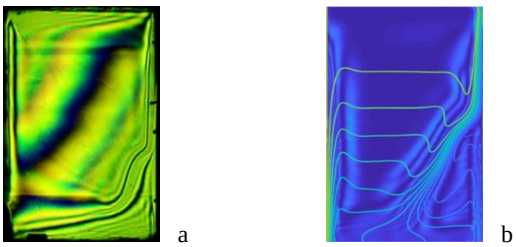


Рис. 10. Гильбертограммы: a – экспериментальная; b – смоделированная.

На рисунке 10 приведены гильбертограммы: a – полученная экспериментально; b – синтезированная из структуры численно смоделированных изотерм.

4. Заключение

Исследования индуцированных нестационарными граничными условиями конвективных течений и их влияния на структуру и динамику фронтов кристаллизации актуальны для развития новых технологий выращивания кристаллов, понимания аномалий процессов формирования и таяния льда в арктических и антарктических районах. Исследование этой проблемы к настоящему времени выполнено только в узком диапазоне параметров и пока не имеет системного характера. Похожими зависимостями плотности от температуры обладают растворы галлия, эвтектики кадмий–ртуть–теллур и некоторых других веществ и материалов. Поэтому вода может использоваться в качестве жидкости-имитатора расплава для отработки технологий получения монокристаллов.

Область приложений не ограничивается течениями, имеющими конвективную природу. Она может быть расширена на решение задач фундаментального и прикладного характера в экспериментальной гидро- и газодинамике, теплофизике, биологии, медицине.

5. Благодарности

Работа выполнена при частичной поддержке РФФИ (18-38-00790 мол. а).

6. Литература

[1] Smits A.J., Lim N.N. Flow Visualization. Imperial College Press. 2010.

[2] Lappa M. Thermal Collection: Patterns, Evolution and Stability. Chichester: John Willey Sons. 2010.

[3] Арбузов В.А., Арбузов Э.В., Дубнищев Ю.Н., Бердников В.С., Мелёхина О.С. Оптическая диагностика фронта кристаллизации, индуцированного температурным градиентом на верхней границе горизонтального слоя жидкости. Автометрия. Т. 53, № 2. С. 39–44.

[4] Бердников В.С., Винокуров В.А., Винокуров В.В. Влияние нестационарных режимов свободной и смешанной конвекции расплавов на теплообмен и формы фронтов кристаллизации в методе Чохральского // Известия РАН. Серия Физическая. 2017. Т. 81, № 10. С. 1395–1402.

Об авторах

Дубнищев Юрий Николаевич, д.т.н., профессор, г.н.с. Института теплофизики СО РАН, dubnistchev@itp.nsc.ru.

Арбузов Виталий Анисифорович, д.т.н., в.н.с. Института теплофизики СО РАН, arbuzov@itp.nsc.ru.

Бердников Владимир Степанович, д.ф.-м.н., г.н.с. Института теплофизики СО РАН, berdnikov@itp.nsc.ru.

Арбузов Эдуард Витальевич, к.ф.-м.н., с.н.с. Института математики СО РАН, arbuzov@math.nsc.ru.

Кислицын Степан Александрович, аспирант НГТУ, инженер-исследователь Института теплофизики СО РАН, 100pch@mail.ru.

Мелёхина Ольга Сергеевна, аспирант НГТУ, melexina-olga17@yandex.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Гильберт-диагностика пламён

Ю.Н. Дубнищев[1, 2], В.А. Арбузов,[1, 2], В.В. Лукашов[1], В.В. Леманов[1]
dubnistchev@itp.nsc.ru | arbuzov@itp.nsc.ru | luka@itp.nsc | lemanov@itp.nsc.ru
[1] Институт теплофизики им. С.С. Кутателадзе
Сибирского отделения Российской академии наук (ИТ СО РАН), Новосибирск, Россия;
[2] Новосибирский государственный технический университет (НГТУ), Новосибирск, Россия

Обсуждается исследование пламени методами гильберт-оптики. Адаптированная к изучению проблем горения диагностика основана на визуализации фазовых возмущений, индуцированных в зондирующем световом поле исследуемой средой, путём использования оптических преобразований Гильберта и Фуко–Гильберта. Диагностический комплекс реализован на основе серийного прибора ИАБ–462М, в котором применяется специально разработанный модуль гильберт-фильтрации, сопряжённый с источником освещения. Исследовалось диффузионное горение струй пропана и водорода. Выявлено влияние на динамическую структуру и эволюцию факела со стороны локальных турбулентных возмущений (puff, slug).Эти возмущения случайно возникают в трубке, формирующей струю, режимах течения соответствующих ламинарно-турбулентному переходу в трубке.

Ключевые слова: *оптическая диагностика потоков, струйное пламя, контроль горения, вихревые структуры*.

Hilbert-diagnostics of flames

Yu.N. Dubnishchev [1, 2], V.A. Arbuzov [1, 2], V.V. Lukashov [1], V.V. Lemanov

dubnistchev@itp.nsc.ru | arbuzov@itp.nsc.ru | luka@itp.nsc.ru | lemanov@itp.nsc.ru
[1]Kutateladze Institute of Thermophysics
Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia;
[2] Novosibirsk State Technical University, Novosibirsk, Russia

The flame investigation by Hilbert-optics methods is discussed. The diagnostics adapted to the study of combustion problems is based on visualization of the phase perturbations induced in the sounding light field by the medium under study, using the optical Hilbert and Foucault-Hilbert transformations. The diagnostic complex is implemented on the basis of the device IAB-462M, which uses a specially developed Hilbert filtering module, conjugated with the light source. The combustion of propane and hydrogen jets was investigated. The influence on the dynamic structure and evolution of the torch from the side of local turbulent perturbations (spots, puff) is revealed. These perturbations occur accidentally in the tube forming the jet, under different burning conditions, depending on the dilution of the burning gas with an inert gas.

Keywords: *optical flow diagnostics, jet flame, combustion control, vortex structures.*

1. Введение

Оптическая диагностика потоков давно и успешно применяется в экспериментальной гидро- и газодинамике. Она отличается многообразием методов и средств, которые составляют современный арсенал исследователя [1]. Большой потенциал развития и применения имеет гильберт-диагностика потоков (HFV – Hilbert Flow Visualization) [2, 3]. Она основана на визуализации пространственно-временных фазовых возмущений, индуцированных в световом поле исследуемой средой путём формирования гильберт-сопряжённого оптического сигнала.

Целью работы является адаптация методов гильберт-оптики к задачам диагностики пламён. Мотивация следует из научной и практической значимости этой проблемы, которая заключается в поиске методов тонкого управления пространственно-временными и термодинамическими параметрами струйного горения [4]. Струйное диффузное горение в режимах контролируемого развития неустойчивости внутри и вне источника формирования струи является одним из таких перспективных методов. Сложность эксперимента заключается в необходимости обеспечения высокочувствительной визуализации полей фазовой оптической плотности газовой среды, возмущённой вихревыми структурами и температурными градиентами.

2. Описание экспериментальной установки

Комплекс оптической диагностики был создан на основе серийного теневого прибора ИАБ–463М [www.skb–photon], в котором использовался специально разработанный для целей эксперимента модуль оптической гильберт-фильтрации и источник освещения.

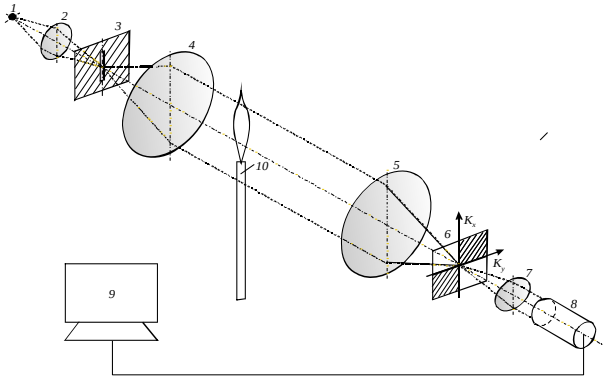


Рис. 1. Схема гильберт-визуализатора.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

На рис. 1 показана упрощённая схема гильберт-визуализатора. Схема содержит осветительный модуль, состоящий из источника света 1, коллиматорной линзы 2 и щелевой диафрагмы 3. Щелевая диафрагма расположена в передней фокальной плоскости объектива 4, формирующего зондирующее поле. Фурье-спектр фазовых возмущений, индуцируемых в исследуемой среде, формируется объективом 5 в частотной плоскости, где помещён квадрантный гильберт-фильтр 6. Объектив 7 выполняет обратное фурье-преобразование фильтрованного светового поля, визуализируя его фазовые возмущения, которые регистрируются цифровой видеокамерой 8, подключенной к компьютеру 9. Этот оптический комплекс реализован на базе теневого прибора ИАБ–463М с модифицированным световым источником, модулями гильберт-фильтрации оптического сигнала и регистрации изображения. Для интерферометрических исследований используется модифицированный интерферометр сдвига.

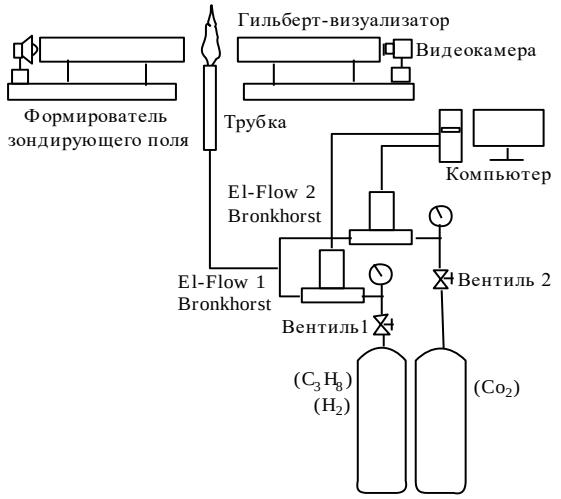


Рис. 2. Схема экспериментального стенда.

Когерентная передаточная функция пространственно-частотного фильтра 6, выполняющего одномерное преобразование Фуко–Гильберта [3], описывается выражением

$$H(K_x, K_y) = \cos\varphi - i\sin\varphi\,\mathrm{sgn}K_x. \quad (1)$$

Фазовый сдвиг φ является функцией длины волны λ зондирующего светового поля: $\varphi = \varphi(\lambda)$. В частотной плоскости для фурье-спектра фазовой оптической плотности светового поля, возмущённого исследуемой средой (пламенем) имеем, с учётом (1):

$$H(K_x, K_y)s(K_x, K_y) = s(K_x, K_y)\cos\varphi + \hat{s}_x(K_x, K_y)\sin\varphi, \quad (2)$$

где

$$\hat{s}_x(K_x, K_y) = -i\mathrm{sgn}K_x s(K_x, K_x). \quad (3)$$

Энергетический спектр оптического сигнала непосредственно за фильтром 6, с учётом (1)

$$\left|H(K_x, K_y)s(K_x, K_y)\right|^2 = \cos^2_\varphi\left|s(K_x, K_y)\right|^2 + \\ +\cos^2_\varphi\left|\hat{s}_x(K_x, K_y)\right|^2. \quad (4)$$

Объектив 7 выполняет обратное фурье-преобразование фильтрованного сигнала (2):

$$E(x, y) = s(x, y)\cos\varphi + \hat{s}_x(x, y)\sin\varphi \leftrightarrow \\ \leftrightarrow s(K_x, K_x)\cos\varphi + \hat{s}_x(K_x, K_x)\sin\varphi. \quad (5)$$

Интенсивность сигнала (5) регистрируется фотоматрицей камеры 8:

$$I(x, y) = \rho\{|s(x, y)|^2\cos^2\varphi + |\hat{s}_x(x, y)|^2\sin^2\varphi\}, \quad (6)$$

где ρ – коэффициент, учитывающий чувствительность фотоматрицы. На длине волны $\lambda=\lambda_0$, удовлетворяющей условию $\varphi(\lambda_0) = \pi/2$, фурье-фильтр выполняет одномерное преобразование Гильберта. Гильберт-преобразование обладает свойствами квазидифференцирования. Поэтому экстремумы и градиенты фазовой оптической плотности исследуемой среды трансформируются в визуализированные структуры гильберт-полос. Пространственное распределение гильберт-полос несёт информацию о возмущениях фазовой оптической плотности, индуцированных вихревыми структурами в реагирующих потоках.

На рисунке 2 показана упрощённая схема экспериментального стенда. Эксперименты проводились при атмосферном давлении и начальной комнатной температуре потока. Экспериментальная установка состояла из баллонов с газом, газовых редукторов и двух расходомеров. Общий расход газов и состав топливной смеси задавался с помощью цифровых расходомеров El-Flow Bronkhorst. В качестве источника струи использовалась трубка из кварца внутреннего диаметра d = 3.2мм, длиной l = 550 мм с толщиной стенки 1.2 мм.

Истечение струи смеси газов осуществлялось в воздушное пространство. Газ подводился к рабочему участку с помощью гибких соединительных шлангов.

3. Результаты

На рис. 3 представлена временная последовательность кадров гильберт-визуализации горения смеси C_3H_8/CO_2, истекающего в неподвижную атмосферу. Время между ближайшими номерами кадров составляет 1/50 с. При числе Рейнольдса Re=3476 внутри источника факела (в трубке) в результате процесса ламинарно-турбулентного перехода в режиме перемежаемости формируются турбулентные пятна (puff) [5].

На кадрах №703–704 представлены изображения отсоединённого ламинарного пламени (в этом интервале времени турбулентные пятна отсутствовали). Начиная с кадра 705 наблюдается развитие турбулентных пятен в потоке факела. На рисунке зарегистрирован случай прохождения puff, который привёл к полному погасанию пламени (кадр 711). Для условий представленного опыта, примерно в 3–5 процентах случаев возникает ситуация, приводящая к распаду струи топлива и полному погасанию факела. По-видимому, режим погасания связан с параметром перемежаемости и уровнем флуктуаций скорости в турбулентных пятнах. Дальнейшее увеличение Re приводит к стабилизации горения.

Как установлено в [5], разбавление пропана инертным газом (в нашем случае рассматривается C_3H_8/CO_2 или C_3H_8/He) может изменять режим горения. При одном и том же числе Рейнольдса и диаметре трубки уменьшение доли пропана может привести к переходу от присоединенного к приподнятому ламинарному или турбулентному пламени.

Примеры визуализированных структур приподнятого (отсоединённого) пламени, приведены на рисунке 4. Слева направо показана временная последовательность изображений пламени, изменяющегося под воздействием ТП. Момент воздействия турбулентного пятна зафиксирован на втором кадре.

На рисунке 5 приведена визуализация факела, иллюстрирующая горение чистого пропана (Re=2 700). Это два изображения (а, б), полученные при гильберт-визуализации пропано-воздушного диффузионного факела в переходном режиме истечения топлива в подводящей трубке (течение с перемежаемостью). На рисунке 5а представлен кадр, соответствующий ламинарному участку течения, на рисунке 5б отображён момент прохождения турбулентного пятна (течение снизу вверх).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Также были проведены опыты для горения смеси H_2/CO_2, истекающей из кварцевой трубки в атмосферу воздуха. Как известно, свойства водорода сильно отличаются от теплофизических и химических свойств углеводородов – H_2 имеет низкую плотность, высокие значения коэффициента диффузии и, как следствие, высокую скорость распространения пламени.

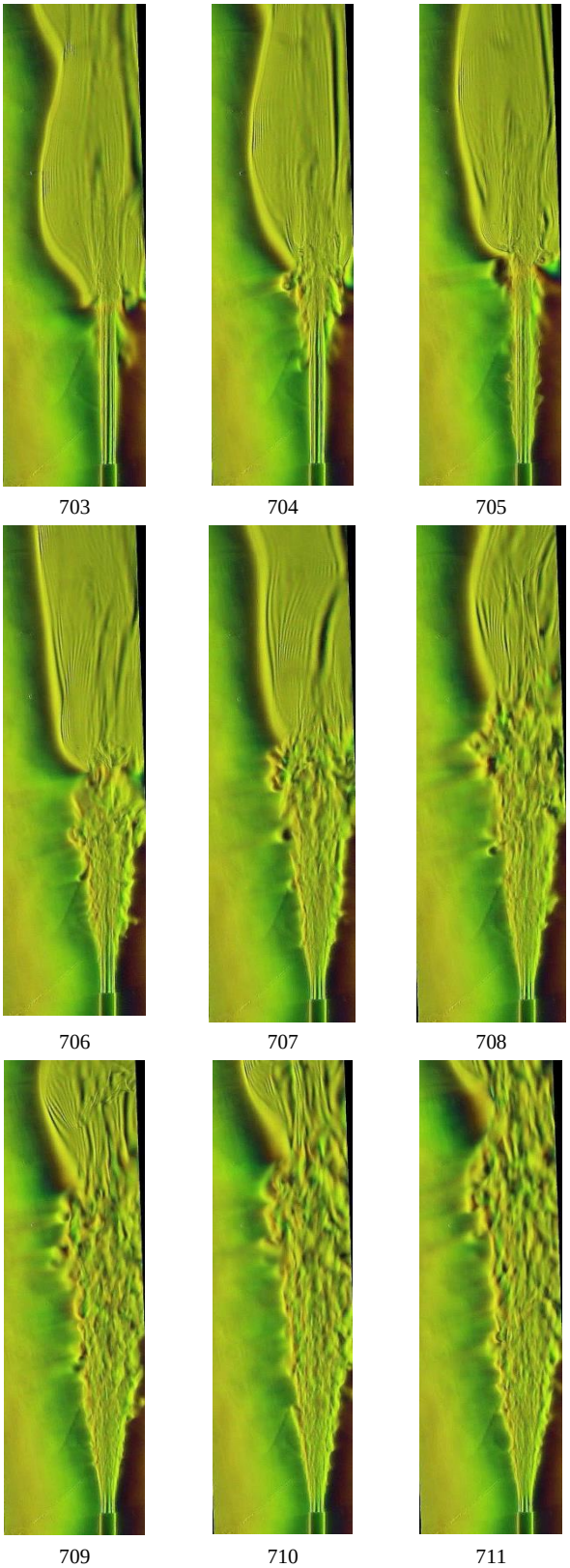


Рис. 3. C_3H_8/CO_2, XC_3H_8=39%, Re=3476, d=3.2 мм, струя снизу вверх

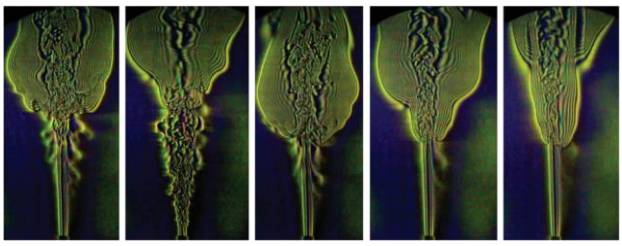
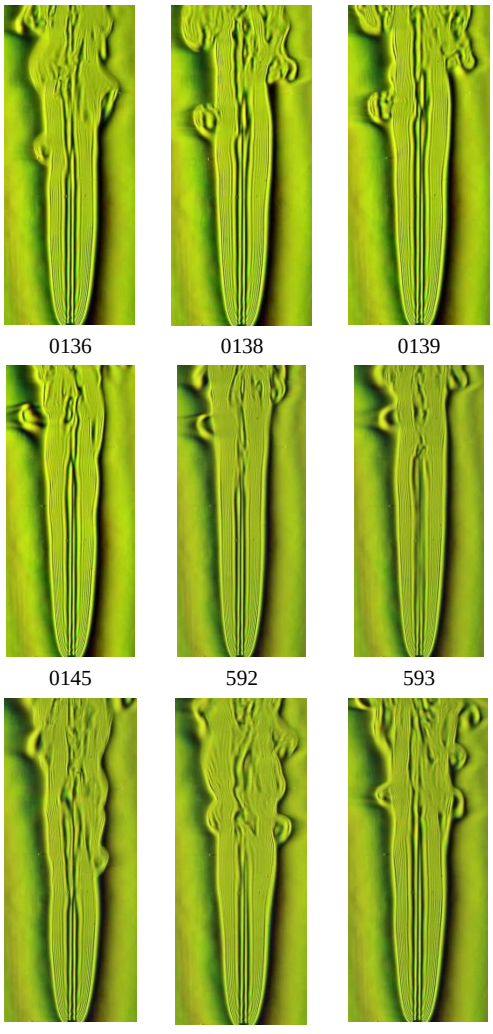


Рис. 4. Визуализированные структуры приподнятого (отсоединённого) пламени.

Рис. 5. Горение чистого пропана.

Это не может не сказываться на различиях в реакции пламени на появление в топливной струе турбулентных пятен в переходном режиме течения в трубке.

Опыты проведены для горения смеси (H_2/CO_2), истекающей из кварцевой трубки диаметром d=3.2 мм и длиной l=550 мм, в атмосферу воздуха. Течение направлено снизу вверх. На рисунке 6 представлена гильберт-визуализация течения с частотой 50 Гц. Цифры под рисунками обозначают номер кадра видеофильма при постоянном числе Рейнольдса Re=2971.



Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Рис. 6. H_2/CO_2, X_1H_2=57%, Re=2971, d=3.2 mm

Факел имеет две характерных области: а) струйное ядро с поперечным размером порядка диаметра трубки; б) внешнюю часть – с поперечным размером существенно больше d. Внутри источника факела (в трубке) случайным образом формируются турбулентные пятна (puff) как следствие ламинарно-турбулентного перехода. Горение водородного топлива сопровождается наличием возмущений как в ядре, так и на периферии факела. На внешней границе факела, в условиях опыта, практически подавляются низкочастотные фликер-колебания. Вместо них наблюдаются уединённые возмущения, которые возникают при прохождении вихревых структур (puff) вдоль ядра факела. В данном случае в ламинарном пламени возможно появление отдельных микропламён, отрывающихся от основного тела факела (кадры 0139 и 0145).

4. Заключение

В работе представлено исследование диффузионного пламени методами гильберт-оптики. Основным инструментарием являлось использование оптических преобразований Гильберта и Фуко–Гильберта. В плане исследования физических процессов главное внимание уделено механизму взаимодействия трубной и струйной неустойчивости. Струйное диффузное горение на режимах контролируемого развития неустойчивости внутри и вне источника формирования струи является одним из перспективных методов управления пламенем. Для изучения диффузионного горения использовались пропан и водород в смеси с инертным разбавителем (CO_2). Опыты проведены при истечении дозвуковой газовой струи в воздушное пространство из длинной трубки в диапазоне чисел Рейнольдса 2000–4000, характерных для ламинарно-турбулентного перехода в трубах. В этом интервале чисел Рейнольдса для диффузионного факела обнаружен механизм двухкаскадной неустойчивости, вызванной формированием турбулентных пятен (puff) внутри трубы (1-и каскад) [5] и генерацией вихревых структур в слое смешения струи (2-и каскад) [6]. Эти организованные структуры (puff) оказывают сильное воздействие как на изотермическую струю [7-9], так и на пламя. С помощью puff можно управлять как низкочастотными фликер-колебаниями, так и высокочастотными модами в приосевой зоне факела. Показана возможность перевода присоединенного пламени в отсоединенное и обратно. Puff могут как инициировать переход к турбулентному факелу, так и приводить к его ламинаризации. Зарегистрирована динамика погасания отсоединённого пламени. Сам механизм погасания, по-видимому, связан с параметром перемежаемости и уровнем флуктуаций скорости в турбулентных пятнах

5. Благодарности

Работа выполнена при частичной поддержке РФФИ (грант № 17-08-00958), в части измерений условий возникновения puffна выходе из трубы выполнены за счет средств, полученных от ФАНО России. Авторы благодарны Абдрахманову Р.Х., Шарову К.А. за помощь при проведении экспериментов.

6. Литература

[1] Smits A, Lim T. Flow Visualization. Techniques and Examples. Imperial College Press. 2000 London: 306 p.

[2] Арбузов В.А., Дубнищев Ю.Н. Методы гильберт-оптики в измерительных технологиях. Новосибирск: Изд-во НГТУ, 2007. 316 с.

[3] Arbuzov V.A, Dvornikov N.A, Dubnishchev Yu.N, Nechaev V.G, Novosyolova O.N, Shlapakova E.O (2016) Mint: Hilbert Diagnostics of Vortex Rings Indused in Air by a Pressure Pulse on a Hole. International Journal of Spray and Spray and Combustion Dynamics 2016;8(3): 197–204. DOI: 10.116].

[4] Литвиненко Ю. А. Устойчивость дозвуковых макро- и микроструйных течений и микроструйное горение (обзор) // Сибирский физический журнал. 2017. Т. 12, № 3. С. 83–99. DOI 10.25205/2541-9447-2017-12-3-83-99.

[5] Mullin T. 2011 Experimental studies of transition to turbulence in a pipe. Annu. Rev. Fluid Mech. 42 1–24

[6] Lee B. J., Kim J. S., Chung S. H. Effect of dilution on the liftoff of nonpremixed jet flames // Proc. Combust. Inst. 1994. V. 25. P. 1175–1181.

[7] Ho C.M and Huerre P. 1984. Perturbed free shear layers. Annu. Rev. Fluid Mech. 16 365-424.

[8] Lemanov V.V, Terekhov V.I, Sharov K.A and Shumeiko A.A. 2013 An experimental study of submerged jets at low Reynolds numbers Tech. Phys. Lett. 39 421–423.

[9] Lemanov V.V, Lukashov V.V, Abdrakhmanov R, Arbuzov V.A, Dubnishchev Yu.N and Sharov K.A. 2018 Regimes of Unsteady Exhaustion and Diffusion Combustion of a Hydrocarbon Fuel Jet. Combustion Explosion and Shock Waves 54 255-263.

Об авторах

Арбузов Виталий Анисифорович, д.т.н., в.н.с. Института теплофизики СО РАН, arbuzov@itp.nsc.ru.

Дубнищев Юрий Николаевич, д.т.н., профессор, г.н.с. Института теплофизики СО РАН, dubnistchev@itp.nsc.ru.

Леманов Вадим Владимирович, к.т.н., с.н.с. Института теплофизики СО РАН, lemanov@itp.nsc.ru.

Лукашов Владимир Владимирович, к.т.н., в.н.с., зав. лаб. термохимической аэродинамики ИТ СО РАН, luka@itp.nsc.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Когнитивная интерпретация визуализации

А.А. Захарова[1], Е.В. Вехтер[2], А.В. Шкляр[1]
zaa@tpu.ru|vehter@tpu.ru|shklyarav@tpu.ru
[1]ИПМ им. М.В.Келдыша РАН, [2]Томский политехнический университет

В работе показаны результаты исследований когнитивной интерпретируемости визуализации и предложены подходы к получению эмпирических оценок практической результативности средств визуализации. Полученные результаты использованы при разработке технологии когнитивной интерпретации гетерогенных данных и необходимых для этого средств визуальной аналитики.

Ключевые слова: *визуальная аналитика, визуальная модель, анализ данных, интерпретация, визуальное восприятие.*

Cognitive interpretation of visualization

A.A. Zakharova[1], E.V. Vekhter[2], A.V. Shklyar[1]
zaa@tpu.ru|vehter@tpu.ru|shklyarav@tpu.ru
[1]Keldysh Institute of Applied Mathematics RAS, [2]Tomsk Polytechnic University

The paper shows some results of studies of cognitive interpretability of visualization and suggests approaches to obtaining empirical estimates of the practical effectiveness of visualization tools. The obtained results were used in developing the technology of cognitive interpretation of heterogeneous data and the necessary tools for visual analytics.

Keywords: *visual analytics, visual model, data analysis, visual interpretation, visual perception.*

1. Введение

На современном этапе развития средств информационной коммуникации визуализация рассматривается как доминирующая форма информационного взаимодействия между пользователями и источниками информации [12]. Однако, к настоящему моменту по-прежнему не сформирован единый обоснованный подход к изучению и использованию возможностей визуализации в качестве инструмента научного исследования и передачи информации. Одним из препятствий для движения в этом направлении является необходимость проведения объемных междисциплинарных исследований, для которых характерна проблема общего определения объекта исследований и применимости специализированных методов [3].

Специфичным эффектом использования визуализации, согласно [11], является влияние на сознание наблюдателя через формирование новой информационной реальности. Объяснение этого результата основано на изучении влияния иллюзии объективности визуализации на результат ее интерпретации. Для существующих систем визуализации использование указанного эффекта является резервом, привлечение которого затруднено отсутствием систематизированного подхода к визуальной аналитике.

Переход от исключительно иллюстративной визуализации к использованию когнитивного потенциала визуального восприятия происходит медленнее, чем развиваются технологии визуализации. Когнитивная интерпретация визуальных образов предъявляет к создателям подобных образов требования, выполнение которых затруднено необходимостью направленного использования особенностей восприятия и мышления, многие из которых недостаточно изучены. В работе представлены результаты, полученные в результате исследования особенностей визуального восприятия и когнитивной интерпретации абстрактных образов.

2. Комплексность

Одним из перспективных направлений развития теории визуализации следует считать исследование соотношения в интерпретации информативного образа объективной и субъективной составляющих. В данном случае, к субъективным факторам, влияющим на общий результат когнитивной интерпретации, могут быть отнесены: условия создания визуального образа, специфика восприятия, культурные традиции и другие данные. Следовательно, в визуализации возникает необходимость согласованного использования методов семиотики, лингвистики, психоанализа, феноменологии, герменевтики, системного анализа и т.д.

В исследованиях [1] сформулированы положения, указывающие на необходимость учета при анализе визуальных образов нескольких уровней дополнительной информации: технологического, лингвистического, жанрового, эстетического, идеологического. В работе [13] предложено выделение трех информационных составляющих: функциональной, семиотической и психоэмоциональной, — управление которыми позволяет контролировать результаты когнитивной интерпретации. Выбор составляющих, вносящих вклад в результат процесса когнитивной интерпретации, является основой создания общей схемы визуализации, а также позволяет определить способы ее улучшения.

3. Проблематика визуализации

Исследователями Drexel University's College of Computing & Informatics (CCI) [2] приведен рейтинг нерешенных проблем визуализации. К настоящему времени, этот список может быть уточнен с учетом актуальных задач, к решению которых может быть привлечена визуализация, в т.ч. научная и когнитивная:

1. Юзабилити (Достаточность). Проблема достижения средствами визуального исследования пользовательских характеристик, обеспечивающих возможность получения достоверного решения поставленной задачи исследования лишь при использовании визуальной интерпретации.
2. Перцепция. Задача увеличения информативности и интерпретируемости средств визуальной аналитики в результате направленного использования законов восприятия в схеме визуализации.
3. Предварительная информированность. Проблема избыточного информирования пользователя,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

снижающего когнитивную результативность визуализации.

4. Обучение. Увеличение сложности задач визуального исследования, создаваемых для этого средств визуальной аналитики и необходимость предварительного обучения пользователей.
5. Качественная мера. Проблема сравнения и выбора средств визуализации из-за отсутствия принятой схемы визуализации и соответствующей ей системы критериев оценки.
6. Масштабируемость. Задача поиска способов визуализации информации, сохраняющих высокую когнитивную интерпретируемость при любых изменениях детализации или объема данных.
7. Интерпретируемая эстетика. Исключение или использование субъективного критерия визуальной эстетики для достижения цели исследования.
8. Динамика. Задача использования восприятия изменяющихся образов для усиления когнитивной интерпретируемости визуализации.
9. Внутрисистемная логика. Формализация использования визуализации в качестве знаковой языковой системы для получения новых знаний при использовании лишь внутрисистемных операций.
10. Представление знаний. Решение комплексной задачи эффективного хранения, передачи, получения и использования информации в результате создания ее визуального представления.

4. Методика исследования

Получение ответов на сформулированные вопросы сталкивается с проблемой получения объективных данных, подтверждающих существование процессов, включенных в любую из разработанных схем визуализации [4, 6, 10]. Наличие и корректная интерпретация подобных эмпирических данных создадут условия для построения средств визуализации, обладающих обоснованным и направляемым когнитивным эффектом. Для получения необходимых данных разработана методика исследования возможностей визуального представления, апробированная на нескольких вариантах средств визуального анализа многомерных гетерогенных данных (Рис. 1).

Суть предложенной методики исследования заключается в проведении серий тестовых решений задач визуального анализа, характеризуемых рядом контролируемых ограничений, а также в измерении интервалов времени, соответствующих различным этапам в схеме визуализации. На каждом этапе происходит взаимодействие пользователя с визуальным образом, которое подразумевает любые доступные пользователю операции, за исключением изменения набора анализируемых данных.

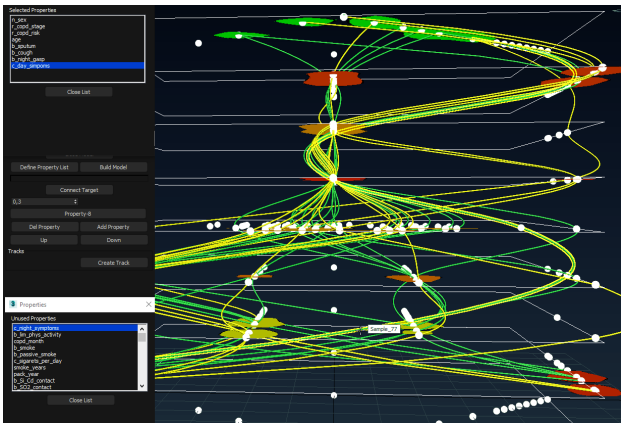


Рис. 1. Применение методики тестирования для одной из визуальных моделей.

Тестовое решение предполагает создание образа исследуемых данных с использованием предопределенного способа визуального представления. Выбор первоначального способа визуализации соответствует типу задачи исследования (задача обучения, информирования, принятия решения) и не зависит от предпочтений исследователя. Таким образом, на первом шаге пользователь располагает визуальной моделью данных и знает цель (вопрос) решаемой им задачи [7].

В ходе тестового решения, исследователю, анализирующему образ данных, позволяется формулировать неограниченное число гипотез ответа на вопрос задачи, каждая гипотеза считается шагом анализа. Формулирование верной гипотезы означает завершение тестового решения. Цикличность процедуры решения позволяет автоматически фиксировать продолжительность каждого этапа для последующего анализа результативности визуализации.

Полученные результаты позволили сделать ряд заключений о характере взаимодействия пользователя и визуальной модели в процессе решения задачи анализа. Прежде всего, предложенная методика изучения когнитивных особенностей визуального анализа позволила оценить общее время и влияние ряда факторов на длительность решения. Например, в случае проведения серии измерений с участием одного пользователя это позволило оценить индивидуальные особенности когнитивных процессов. Типичный пример результатов таких измерений (Рис. 2) показывает существенное изменение скорости мышления исследователя в ходе решения однотипных задач.

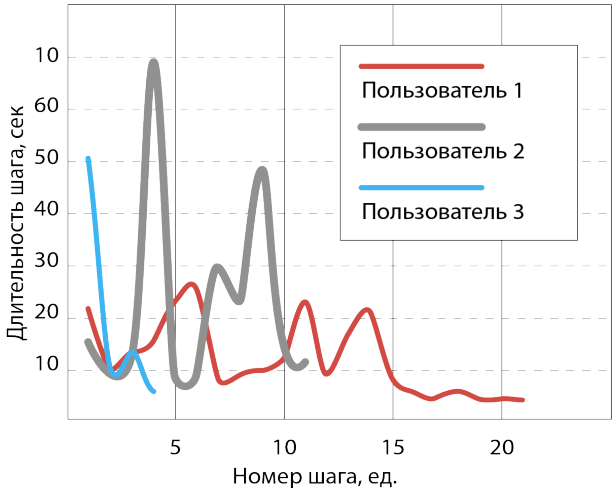


Рис. 2. Длительность шагов в процессе когнитивной интерпретации визуализации.

5. Обучение пользователя

Для получения экспериментальных результатов, позволяющих делать выводы о когнитивных свойствах визуализации, необходима постановка тестовых задач, минимизирующая влияние факторов, относящихся к формальным логическим рассуждениям, опирающимся на существующие у пользователя знания. В качестве тестовых задач пользователю предлагалось исследование наборов многомерных данных, описывающих реальные системы. В соответствии с целью проведения тестовых решений, пользователю было неизвестно происхождение данных и их номинальные значения. Набор визуализированных данных дополнялся искусственно сгенерированными данными,

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

которые считались аномальным отклонением от реальных значений. Обнаружение такого аномального объекта в наборе исходных данных являлось целью тестового визуального исследования.

Получение с помощью предложенной методики повторяющихся результатов измерений, позволило сделать предположения о некоторых общих закономерностях использования визуализации в качестве инструмента аналитического исследования и получить представление о способах влияния на результаты такого использования. По результатам проведенных измерений, в процессе взаимодействия пользователя и визуальной модели выделены интервалы, на протяжении которых время, затрачиваемое пользователем на построение очередной гипотезы, непрерывно сокращается. Во многих случаях, такие интервалы расположены в начале процесса визуального анализа, однако, могут повторяться и на более поздних шагах решения. Высокая повторяемость полученных результатов и их зависимость от пользователя, принимающего участие в тестовых измерениях, являются основанием для сопоставления с когнитивными аспектами визуализации.

С этой точки зрения, интервалы ускоряющегося формирования новых гипотез, возникающие на начальных этапах схемы визуализации, могут быть интерпретированы как стадии анализа, соответствующие ознакомлению пользователя с инструментом визуального анализа, его особенностями в решаемой задаче, а также накоплению дополнительных сведений, возникающих в результате принятия неверных решений. Следует обратить внимание, на то, что формирование очередной гипотезы ответа, даже в случае ошибочности прежней, многократно повторяется вместе с сокращением затрачиваемого времени. Таким образом, в схеме визуализации, содержащей многократное повторение циклов приема и передачи информации [9], может быть выделена стадия накопления пользователем информации, т.е. этап обучения.

На основании проведенных измерений, показана зависимость продолжительности этапа обучения от способа визуального представления исследуемых данных, а также от собственных характеристик пользователя. Выделение и использование факторов, сокращающих длительность когнитивных процессов, является способом решения многих задач, указанных в качестве проблематики визуализации.

6. Сомнения и рефлексия

Изменение скорости построения новых гипотез, наблюдаемое на протяжении процесса визуального анализа, носит периодический характер. Во время решения тестовых задач пользователь, в рамках проводившегося исследования, не получал сведений из внешних источников, поэтому изменение времени, затрачиваемого на каждом шаге анализа, может быть объяснено изменением соотношения между прямыми и обратными процессами [5]. С точки зрения изучения когнитивных возможностей визуализации, изменяется скорость процедур поиска информации и проверки результатов поиска.

Незначительное снижение скорости решения, с этой точки зрения, свидетельствует об увеличении времени, затрачиваемом на построение и проверку очередной гипотезы, и может интерпретироваться как сомнение в правильности совершаемых действий, вызванное ошибками, допущенными на предыдущих шагах (Рис. 3). Возникновение сомнений приводит к повторному обращению к накопленным знаниям или повторению логических операций с привлечением внутрисистемной логики. Улучшение когнитивной результативности в этом случае может быть достигнуто в результате использования средств визуализации, предоставляющих оперативный доступ к необходимым сведениям.

В большинстве измерений, резкое снижение скорости формирования пользователем новой гипотезы наблюдалось после серии быстрых ошибочных предположений. Это говорит об отсутствии у пользователя понимания результатов визуальной интерпретации информации, т.е. о противоречивых результатах когнитивных процессов. Значительное снижение скорости анализа возникает с периодичностью, отличающейся у разных пользователей и может быть использовано в качестве одной из характеристик визуализации. Длительность подобного этапа рефлексии, используемого для переосмысления доступной информации и совершенных ошибок, зависит от индивидуальных особенностей исследователя. Особенностью этапа рефлексии является преобладание операций обращения к сведениям, хранящимся в памяти пользователя. Таким образом, сокращение продолжительности этапа рефлексии возможно в случае устранения многократного перекодирования визуальной информации и знаний пользователя в результате выполнения операций поиска и проверки необходимой информации. Одним из возможных решений этой задачи является сохранение в визуальном образе исследуемых данных сведений об ошибочных решениях, сделанных ранее. Подобный прием является способом накопления визуальной информации и его результативность должна исследоваться дополнительно, т.к. не соблюдается граничное требование используемой методики, связанное с неизменностью исходных данных.

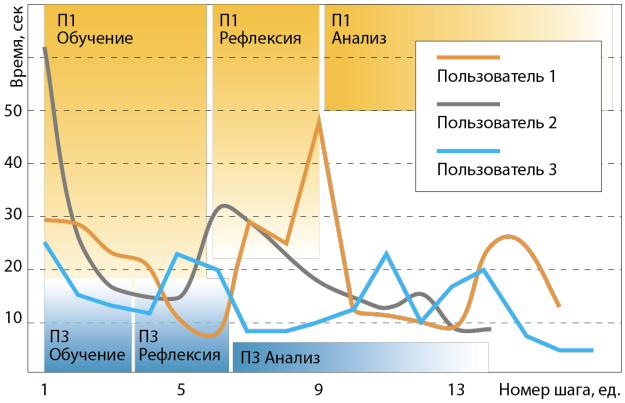


Рис. 3. Стадии обучения и рефлексии.

7. Форма и цвет

На основании измерения времени решения тестовых задач с помощью средств визуального анализа, использующих различные способы визуального представления, проведена оценка некоторых аспектов визуального образа с точки зрения их влияния на скорость формирования новой гипотезы. При проведении серийных измерений предполагалось, что все прочие факторы, кроме исследуемых в соответствующей серии, оказывают на результаты измерения известное и ограниченное влияние.

При сравнении результатов использования визуальных моделей, отличающихся способами цветового кодирования анализируемых данных (Рис. 4), были получены результаты, необходимые для уточнения существующих способов визуализации данных. Полученные результаты интерпретированы как вероятное негативное влияние использование цвета при создании визуального представления данных.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

В соответствии с результатами тестовых измерений, цветовая составляющая, во-первых, воспринимается пользователем как дополнительные данные, участвующие в анализе. Это приводит к снижению когнитивной интерпретируемости визуального образа в случае дублирования уже представленных данных или интерпретации пользователем несущественных данных. Во-вторых, пользователь, по результатам опросов участников тестовых решений, во многих случаях отдавал предпочтение информативности формы визуальных объектов, составляющих образ данных, и игнорировал цветовую составляющую. Таким образом, сопоставление результативности использования визуальных моделей в решении различных задач анализа, показало необходимость оправданного использования цветового кодирования при создании визуальных моделей, рассчитанных на когнитивную интерпретацию.

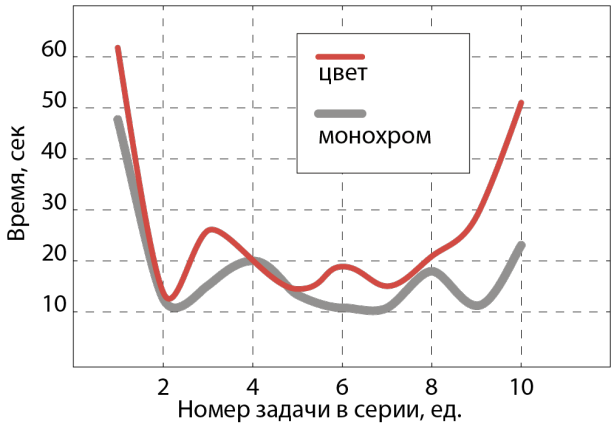


Рис. 4. Влияние цветового кодирования на время решения.

8. Характеристики средств визуализации

На основании разработанной авторами методики проведения визуального исследования данных и созданного обоснования применимости различных средств визуальной аналитики с целью увеличения общей результативности визуального исследования, сделан вывод о необходимости введения собственных характеристик средств визуализации. Предложны определения двух величин, каждая из которых может быть использована для упрощения процессов разработки, выбора и использования средств визуализации.

Для решения вопросов, связанных с выбором средств визуализации, введена комплексная характеристика, объединяющая длительности интервалов времени, затрачиваемых на построение визуальной модели, ознакомление пользователя с ее свойствами и получение ответа на вопрос исследования (Рис. 5). Разработанная методика позволила провести необходимые измерения и получить комплексные характеристики исследованных средств визуализации.

В качестве примера приведены комплексные характеристики трех визуальных моделей, примененных для решения одной задачи анализа данных. Каждый столбец диаграммы (1-3) соответствует одному из тестируемых способов визуализации, комплексная характеристика модели (Рис. 1) соответствует характеристике 2. Характеристика 1 относится к визуальной модели, представляющей облако точек в пространстве с определенной системой шкал, характеристика 3 — относится к визуальной модели, дополненной цветовым кодированием. Нижний сегмент (синий) — этап построения модели, средний (оранжевый) — этап обучения, верхний (серый) — этап решения. Численные характеристики соответствуют средней продолжительности соответствующих этапов (секунды), полученной с использованием предложенной методики измерений.

Показан вариант нормированных характеристик, позволяющий анализировать когнитивные характеристики визуализации, достигаемые при использовании определенных ресурсов (временных и вычислительных). Одним из практических выводов, которые могут быть сделаны на основании комплексной характеристики, является соотношение между когнитивной интерпретируемостью средств визуализации и их ресурсоемкостью.

Визуальная модель данных является сложным многофакторным объектом и может быть охарактеризована с точки зрения соответствия своему назначению, т.е. процессу когнитивной интерпретации. Для функции визуального представления, формирующей визуальный образ исследуемых данных, предложено введение декомпозиционной оценки восприятия (ДОВ), подтверждающей реализацию в процессе интерпретации образа данных соответствующих принципов восприятия. ДОВ представляет собой многомерный бинарный вектор, содержащий экспертную оценку используемых функцией визуального представления принципов восприятия. В Таб.1 приведен вариант подобной оценки для тестируемой модели (Рис. 1).

Таблица 1

Декомпозиционная оценка визуальной модели

Принцип восприятия	Реализация
Константность	да
Целостность	нет
Обобщенность	нет
Предметность	да
Движение	нет
Цвет	да

Введение этой оценки (ДОВ) для разработанных функций визуального представления позволяет алгоритмизировать процесс решения задач визуального исследования, сократить затраты временных и вычислительных ресурсов и, следовательно, увеличить его результативность.

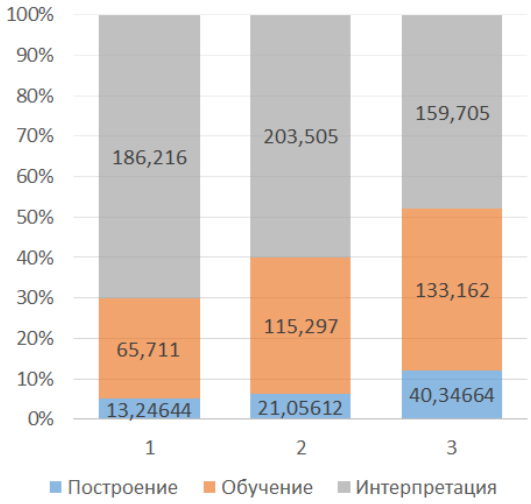


Рис. 5 Примеры комплексных характеристик.

9. Практическое использование

Предложенный подход к исследованию когнитивной интерпретации визуализации позволил разработать средства визуального исследования для ряда практических задач.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Примерами таких решений являются средства визуального анализа гетерогенных данных, предназначенных для оперативного исследования медико-биологических данных [7] (Рис. 6), а также предложенное решение для задачи определения степени изученности заданной предметной области с использованием средств визуализации [8].

В результате применения предложенного подхода к разработке и исследованию возможностей средств визуализации, увеличена результативность решения практических задач. Оценка результативности в этом случае происходила на основании измерения средних интервалов времени, затрачиваемых пользователем на формулирование гипотезы ответа (Таб. 2). Для получения численной оценки результативности предложено сравнение общего времени решения задачи исследования и суммарного интервала обучения и рефлексии. Предполагается, что использование когнитивной интерпретации сокращает время (в проведенных измерениях, на 40-45%), необходимое для общего решения, а завершение цикла обучение/рефлексия соответствует переходу к новой мысленной формулировке цели исследования.

Таблица 2
Оценка результативности решения

Общее время решения, сек	Цикл обучения, сек	Эффект, %
250,9	148,4	40,8
153,2	80,3	47,6
79,5	37,7	52,5
157,6	97,7	38,0
244,9	110,6	54,8
160,4	100,8	37,2

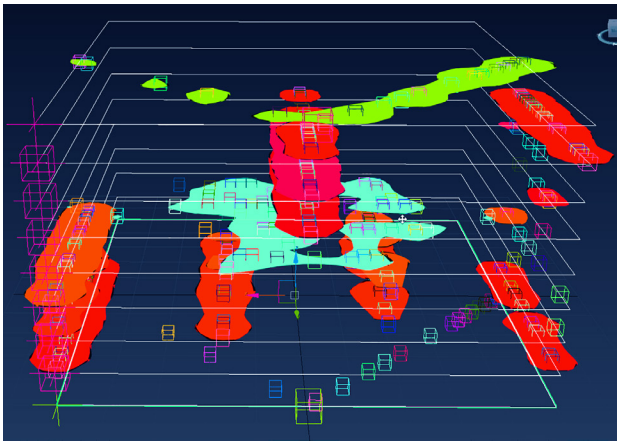


Рис. 6. Модель медико-биологических данных.

10. Заключение

В работе показаны подходы к оценке когнитивной интерпретируемости средств визуализации, используемых при решении практических задач исследования данных. Приведены результаты изучения результативности средств визуализации и показаны некоторые возможные пути управления ее значением. Предложенные характеристики средств визуализации позволяют систематизировать использование визуализации при решении широкого круга задач, имеющих научно-исследовательскую и прикладную направленность. Показана возможность увеличения общей результативности визуального исследования до 35-45% на основании целенаправленного использования возможностей когнитивной интерпретации визуальных образов данных. Перспективным направлением исследований в рассматриваемой области является согласование когнитивной интерпретируемости визуализации и закономерностей восприятия. Средствами достижения этой цели является определение предложенных характеристик визуализации и факторов, оказывающих ни них влияние.

Работа выполнена при поддержке Российского научного фонда, проект 18-11-00215.

11. Литература

1. Berger A.A. Seeing is believing : an introduction to visual communication / A.A. Berger, McGraw-Hill Education, 2011. 267 с.

2. Chen C. Top 10 unsolved information visualization problems // IEEE Computer Graphics and Applications. 2005. № 4 (25). С. 12–16.

3. Elkins J. Interpreting Non-art Images 2001.

4. Green T.M., Ribarsky W., Fisher B. Visual analytics for complex concepts using a human cognition model 2008. 91–98 с.

5. Pirolli P. Information Foraging Theory: Adaptive Interaction with Information // Information Foraging Theory. 2007. 1–28 с.

6. Pirolli P., Card S. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis // Proceedings of International Conference on Intelligence Analysis. 2005. (2005). С. 2–4.

7. Shklyar A. [и др.]. Visual detection of internal patterns in the empirical data Springer Verlag, 2017. 215–230 с.

8. Shklyar A. [и др.]. Visual modeling in an analysis of multidimensional data // Journal of Physics: Conference Series. 2018. № 1 (5). С. 125–128.

9. Spivey M. The Continuity of Mind / M. Spivey, 2007. 1-443 с.

10. Ward M.O., Theroux K.J. Perceptual Benchmarking for Multivariate Data Visualization 1997. 314–321 с.

11. Барт Р. Риторика образа // Избранные работы. Семиотика. Поэтика. / Р. Барт, М.:, 1989. 616 с.

12. Розин В.М. Визуальная культура и восприятие. Как человек видит и понимает мир / В.М. Розин, 6-е изд., URSS, 2015. 272 с.

13. Шкляр А., Захарова А. Структурный подход к визуализации данных Издательство МАИ-Принт (Москва), 2017. 609–611 с.

Об авторах

Захарова Алена Александровна, Институт прикладной математики им. М.В. Келдыша РАН, zaa@tpu.ru.

Вехтер Евгения Викторовна, Томский политехнический университет, Томск, Россия, vehter@tpu.ru.

Шкляр Алексей Викторович, Институт прикладной математики им. М.В. Келдыша РАН, shklyarav@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Кросс-корреляционные методы анализа изображений в экспериментальной визуализации потоков

И.А. Знаменская, Е.Ю. Коротеева, Ф.Н. Глазырин
znamen@phys.msu.ru|koroteeva@physics.msu.ru|glazyrin@physics.msu.ru
МГУ им. Ломоносова, Москва, Россия

Обсуждается применение кросс-корреляционных алгоритмов для анализа высокоскоростных газоплазменных и жидких потоков. В приведенных экспериментах исследовались поля плотности прозрачных сред и векторные поля скорости, полученные с использованием как засевного, так и беззасевного трассирования. Для цифровой обработки полученных экспериментальных изображений потоков использовался кросс-скорреляционный алгоритм, реализованный в программе Davis фирмы LaVision. При схожем алгоритме обработки и анализа физические принципы регистрации физических полей различны.

Ключевые слова: *кросс-корреляционный анализ, цифровая регистрация, рефракция, трассирование, разрывы.*

1. Введение

Методы визуализации течений газа, жидкости, плазмы, а также многофазных сред описывались в 20 веке в большом количестве монографий, обзоров и статей. Физические основы методов мало меняются; в них используются преимущественно свойства света - рефракция, рассеяние, поглощение - и такие явления, как интерференция, дифракция, флюоресценция. Однако методы визуализации сегодня получили мощнейший толчок в связи с цифровой (компьютерной) революцией, которая привела к многократному увеличению количества исследований в области динамики течений и объема собранных данных.

Одно из основных преимуществ цифровых технологий – возможность компьютерной обработки цифровых изображений с целью получения количественной информации о параметрах потока.

2. Области применения кросс-корреляционной обработки в эксперименте

Кросс-корреляционный алгоритм сегодня – один из главных инструментов цифровой обработки изображений. На начальном этапе развития этих инструментов он использовался преимущественно для анализа полей скоростей в методе трассирования с засевом частиц (слежение за частицами в потоке, particle image velocimetry или просто PIV). Сегодня можно назвать ряд методов визуализации потоков, где кросс-корреляционный алгоритм является основным: теневой фоновый метод (ТФМ), беззасевные теневые методы [9], термографическое PIV [4], измерение приповерхностного поля скорости в вязких покрытиях, микро, стерео, томографические модификации PIV [3] и др.

Первоначально для определения смещений частиц в засеянном потоке жидкости или газа широко использовался базовый кросс-корреляционный алгоритм, который в дальнейшем был назван стандартным. Стандартный алгоритм состоит из следующих основных операций:

а) разбиение пары изображений на элементарные расчетные области равного размера;

б) расчет кросс-корреляционной функции;

в) нахождение максимума на корреляционной функции;

г) подпиксельная интерполяция максимума корреляционной функции; используется преимущественно при исследованиях несжимаемой жидкости.

Сегодня разработано множество коммерческих и специальных алгоритмов, адаптированных к конкретным течениям. В настоящей работе обсуждается применение кросс-корреляционных алгоритмов для анализа высокоскоростных газоплазменных и жидких потоков. Исследовались

1) поля плотности (теневой фоновый метод, ТФМ)

2) векторные поля скорости (цифровое трассирование, PIV).

Кроме того, рассматривается опыт применения кросс-корреляционной обработки в высокоскоростной беззасевной термографии пограничного слоя жидкости.

Во всех случаях использован метод кросс-корреляции, реализованный в программном обеспечении Davis фирмы LaVision. При схожем алгоритме обработки - определение смещения характерных элементов на изображении - физические принципы регистрации физических полей различны.

В первом случае - при визуализации полей показателя преломления прозрачной неоднородной среды - определяются смещения разделенных в пространстве изображений точек фона, помещенного за исследуемым объектом. Смещение точек на изображении фона происходит за счет отклонения света в неоднородности плотности среды, через которую происходит регистрация фона. Визуализация оптической неоднородности происходит за счет физического явления *рефракции* света.

Во втором случае – при регистрации мгновенных полей скорости среды в сечении потока (например, лазерным листом) - регистрируются разделенные во времени изображения трассеров, зафиксированные с малой экспозицией (равной, как правило, времени подсветки лазерным листом). Регистрируя с интервалом δt движение трассирующих частиц в потоке, можно получить изображения отрезков пути δs, пройденные частицами за δt. Величина $v=\delta s/\delta t$ представляет собой среднюю скорость частицы на этом отрезке. Визуализация трассеров в классическом методе PIV происходит за счет *рассеяния* лазерного излучения на трассере.

При интегральной регистрации следа частицы в газе программа обработки данных метода цифрового трассирования позволяет визуализировать и наблюдать траектории частиц, поля завихренности, поля скоростей в векторном и скалярном представлении. На рис. 1 представлено изображение линий тока частиц, совмещенное с полем мгновенных скоростей, обозначенных цветом; приведена соответствующая шкала. Визуализирована область нестационарного течения на выходе из прямоугольного канала ударной трубы.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

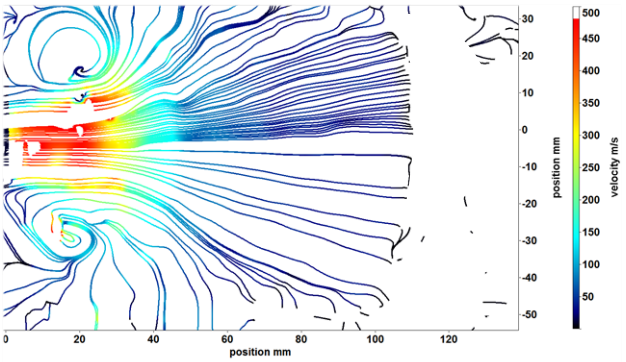


Рис. 1. Скалярное представление результатов цифрового трассирования.

Правая граница линий траекторий частиц соответствует положения фронта ударной волны, скорость потока за ним около 50 м/с. На сверхзвуковом участке струи скорость газа достигает 400-450 м/с.

3. Коррекция результатов

Особую сложность представляет использование кросс-корреляционного алгоритма анализа данных в высокоскоростных газоплазменных течениях с ударными волнами - вследствие скачкообразного изменения параметров (плотности и скорости) на фронте разрыва. Так, обнаружены значительные погрешности при использование классической схемы ТФМ для количественного определения скачка плотности на фронте ударной волны [6]. Это обусловлено выходом детектируемой величины рефракции за рамки чувствительности метода — из-за сильного преломления света на фронте ударной волны отклоненный луч может выйти за пределы оптической схемы и не быть зарегистрированным [2].

Несмотря на то, что в последние годы метод PIV все более активно применяется для визуализации транс- и даже сверхзвуковых течений, наблюдение нестационарных потоков с сильными разрывами этим методом по-прежнему затруднено. Необходимо, чтобы трассирующие частицы следовали движению частиц газа, являясь элементом сплошной среды. Из-за запаздывания частиц, немгновенного характера измерений в методе PIV, влияния алгоритмов цифровой обработки и т.д, данные PIV в таких течениях могут существенно отличаться от истинных. Основной причиной расхождения результатов PIV и реальной скорости потоков является инерционное запаздывание трассирующих частиц. Трассирующие частицы, жидкие или твердые, никогда не следуют потоку абсолютно точно. Сила вязкого трения, действующая на частицы со стороны газа, стремится уравнять их скорость со скоростью потока, ускоряя или замедляя их в зависимости от их относительной скорости. В результате количественное значение скачка скорости оказывается заниженным, а сам фронт существенно «размывается» [5].

В то же время возможно восстановить исходное течение с помощью численного моделирования на основе уравнений газодинамики. В этом случае моделируется динамика трассирующих частиц в рассчитанном поле течения [7]. Если предположить, что численное моделирование на основе уравнений газодинамики (CFD) точно предсказывает фактическую скорость потока, мы можем количественно оценить ошибку данных PIV. На рис. 2 приведено сравнение экспериментального профиля скорости на фронте взрывной волны от импульсного цилиндрического разряда с расчетом на основе численного моделирования.

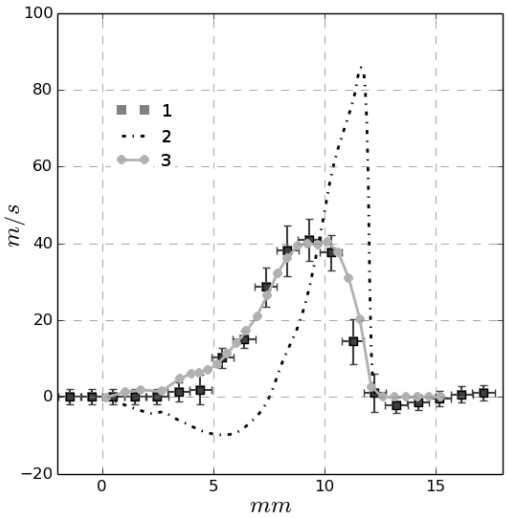


Рис. 2. Учет погрешности экспериментальных данных на основе численного моделирования: 1 – экспериментальный профиль (PIV); 2 – результат CFD моделирования; 3 – профиль «виртуальных» трассирующих частиц в рассчитанном поле течения.

4. Кросс-корреляционный анализ газоплазменного течения с разрывами.

На рис. 3 и 4 приведены результаты экспериментов по исследованиям полей плотности и скорости в нестационарном разрывном газоплазменном течении, возникающем при инициировании наносекундного поверхностного скользящего разряда (плазменного листа) в канале ударной трубы. Разряды площадью 100x30 мм^2 поджигались одновременно на верхней и нижней стенках канала.

На рис. 3 представлены результаты, полученные на основе теневого фонового метода. На рис. 3(а) – исходный снимок разрядной области с фоном, репером и свечением яркого канала разряда. На рис. 3(б) – результат кросс-корреляционной обработки – поле плотности. Наличие разрывов приводит к большой погрешности измерений [1]. Структура течения отражена достаточно правильно.

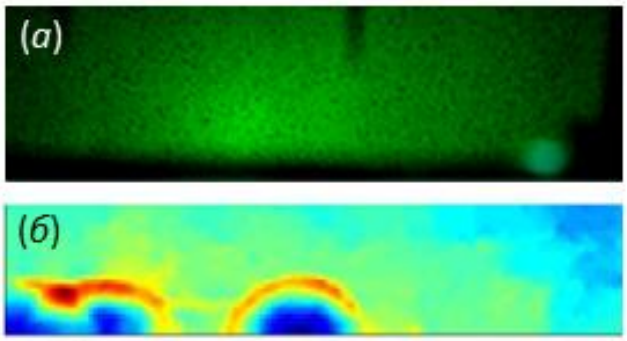


Рис. 3. Теневой фоновый метод. Исходный снимок (а) и результат обработки (б) поля течения, индуцированного импульсными поверхностными разрядами.

На рис. 4 приведено поле того же течения, но полученное на основе метода цифрового трассирования (PIV). На рис. 4(а) – исходное изображение без засева, демонстрирующее область пересечения области поверхностного разряда с лазерным листом. На рис. 4(б) – мгновенное изображение поля течения после разряда с засевом. Видно сгущение трассеров за фронтом взрывной волны и их практическое отсутствие в области разрежения.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

На рис. 4(в) – результат кросс-корреляционной обработки пары изображений потока - поле скоростей; время после разряда - 21 мкс.

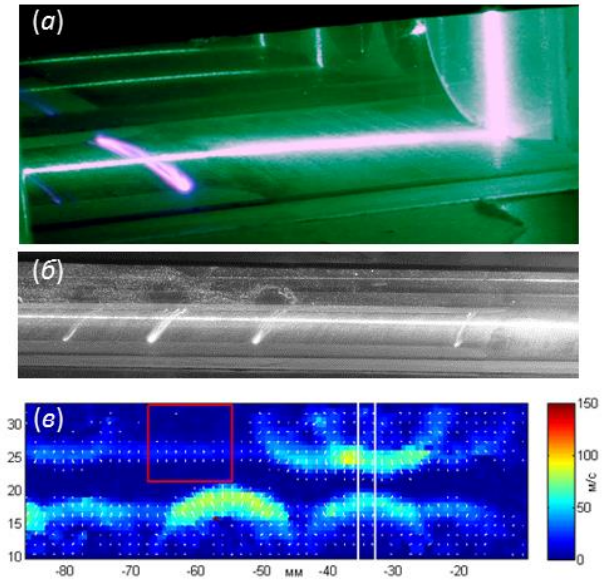


Рис. 4. Метод цифрового трассирования. Вид области регистрации (а-б) и результат обработки (в) поля течения через 21 мкс после инициирования импульсных поверхностных разрядов.

5. Беззасевные методы трассирования

В последние годы благодаря распространению кросс-корреляционных алгоритмов обработки изображений появилось значительное количество работ по использованию беззасевного трассирования – слежение за неоднородностями, структурными элементами, маркерами, присутствующими в самом потоке.

Одним из таких методов является предложенный авторами метод трассирования тепловыми точками (ТТТ), основанный на термографической визуализации пограничного слоя жидкости. Программой кросс-корреляции измеряется смещение точек равной температуры (градации серого) на двух соседних термограммах. На рис. 5 приведено усредненное по 100 кадрам поле скоростей потока в пограничном слое воды на основе нового метода ТТТ [8]. Использовалась высокоскоростная съемка тепловизором импактной затопленной неизотермической струи через окно, прозрачное для инфракрасного излучения.

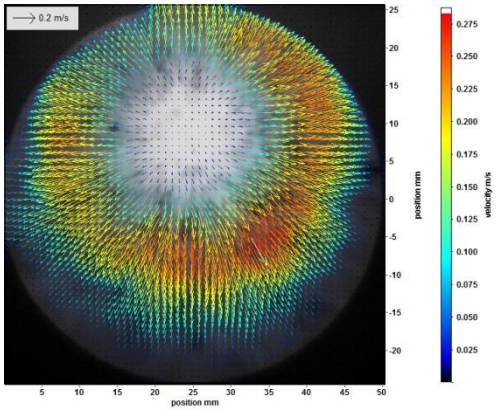


Рис. 5. Метод беззасевного трассирования тепловыми точками. Поле скоростей в погранслое импактной затопленной струи.

6. Заключение

Рассмотрены примеры применения кросс-корреляционных алгоритмов для анализа высокоскоростных потоков. Исследовались поля плотности прозрачных сред и поля параметров газоплазменных и жидких потоков, полученные с использованием как засевного, так и беззасевного трассирования. Показано, что для различных физических принципов регистрации могут быть использованы аналогичные методы обработки. Рассмотрены проблемы, связанные регистрацией параметров в области ударных и взрывных волн.

7. Благодарности

Работа выполнена при поддержке гранта РНФ № 18-19-00672.

8. Литература

[1] Ф. Глазырин, И. Знаменская, Е. Коротеева, И. Мурсенкова, Н. Сысоев. Использование теневого фонового метода для исследования нестационарного потока с ударной волной. Научная визуализация. 2013. Т. 5(3). С. 65–74.

[2] Ф.Н. Глазырин и др. Исследования ударно-волнового течения в канале теневым и теневым фоновым методами. Автометрия. 2012. Т. 48 (3). С. 101–110.

[3] R.J. Adrian. Particle-Imaging Techniques for Experimental Fluid Mechanics. Annu. Rev. Fluid Mech. 1991. Vol. 23 (1). pp. 261–304.

[4] Charogiannis, I. Zadrazil, and C. N. Markides, Thermographic particle velocimetry (TPV) for simultaneous interfacial temperature and velocity measurements. Int. J. Heat Mass Transfer. 2016. Vol. 97. pp. 589–595.

[5] F. N. Glazyrin, I. V. Mursenkova, I. A. Znamenskaya. PIV tracer behavior on propagating shock fronts. Meas. Sci. Technol. 2016. V. 27(1). 015302.

[6] Goldhahn E. Quantitative Measurements of Three-Dimensional Density Fields Using the Background Oriented Schlieren Technique. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[7] Koroteeva E., Mursenkova I., Liao Y., and Znamenskaya I. Simulating particle inertia for velocimetry measurements of a flow behind an expanding shock wave**.** Physics of Fluids. 2018. V. 30. 11702.

[8] E. Koroteeva, I. Znamenskaya, P. Ryazanov, and A. Novinskaya. Velocimetry of water boundary layer flows by thermal imaging. In Proceedings of 11 Pacific Symposium on Flow Visualization and Image Processing (PSFVIP-11)**.** 2017. ID 087.

[9] G. S. Settles, Schlieren and Shadowgraph Techniques. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001.

9. Об авторах

Знаменская Ирина Александровна, д.ф.-м.н., профессор кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета. Ее e-mail znamen@phys.msu.ru.

Коротеева Екатерина Юрьевна, к.ф.-м.н., с.н.с. кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета. Ее e-mail koroteeva@physics.msu.ru

Глазырин Федор Николаевич, аспирант кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Цифровые технологии в анализе теневых изображений газоплазменных потоков

Знаменская Ирина[1], Наумов Дмитрий[1], Дорощенко Игорь[1], Кузнецов Александр[1], Луцкий Александр[2]
znamen@phys.msu.ru|n_ds_n@inbox.ru|doroshchenko93@gmail.com|sasha15555@mail.ru|allutsky@yandex.ru
[1]Московский государственный университет имени М.В. Ломоносова, Москва, Россия;
[2]ИПМ им. М.В.Келдыша РАН

Проанализированы принципиально новые возможности, обеспечиваемые современной цифровой высокоскоростной регистрацией течений с разрывами, в частности, при прямом сравнении расчетной и экспериментальной визуализации течений. Рассмотрены примеры высокоскоростной теневой съемки и анализа газоплазменных высокоскоростных течений с разрывами при различных параметрах регистрации. Анализируются новые проблемы, сопровождающие переход на цифровые технологии и связанные с большими объемами получаемых экспериментальных данных в виде цифровых видеофильмов.

Ключевые слова: *теневые методы, цифровая регистрация, фильтрация шума, быстропротекающие процессы, разрывы, сравнение с численными моделированием.*

1. Введение

Развитие панорамных методов регистрации обработки, количественного анализа изображений применительно к задачам динамика разрывов и газоплазменных образований находится в прямой зависимости от прогресса в области создания цифровой регистрирующей аппаратуры. Развитие методов и приборов для регистрации быстропротекающих процессов в 50-е годы 20 века было связано с прикладными задачами в области атомной физики и ракетно – космической отрасли. Наиболее быстрыми регистрируемыми процессами были - динамика ионизирующих ударных и взрывных волн, детонация. Революционный переход от пленочных технологий к цифровой высокоскоростной регистрации быстропротекающих процессов в газах, плазме, жидкости привел, однако, к удорожанию эксперимента, и в России зачастую – практически к недоступности съемки процессов микросекундного диапазона.

Барабанные пленочные камеры: СФР, ЖЛВ, ВСК позволяли получить до 48 кадров за 1 эксперимент при хорошем пространственном разрешении при съемке до 1 миллиона кадров в секунду либо развертки процесса. Цифровые современные приборы на такой же скорости могут снимать фильмы объемом до 10 Гб и длительностью несколько секунд. Это позволяет регистрировать не отдельные фрагменты процессов миллисекундной длительности – а весь процесс от начала до конца - например, течение в ударной трубе; течение при запуске газодинамического сопла и др. Другим преимуществом цифровой съемки является возможность сравнения любой стадии процесса с анимацией, полученной при визуализации результатов численного моделирования процесса. В последнее десятилетие стало возможным регистрировать соответствующей аппаратурой и субмикросекундные процессы – микровзрывы, эволюцию импульсных плазменных образований различного происхождения.

При этом возникает проблема сохранения, обработки, анализа больших объемов данных – как эксперимента, так и численных данных и осмысление, интерпретация физического процесса на основе их сравнения.

Данная работа рассматривает возможности исследования формирующихся при инициировании сильноточных локализованных в сверхзвуковом потоке разрядов и динамики возникающих при этом разрывов. Использовались различные модификации теневых методов. Для регистрации использовалась цифровая съемка камерой Photron FastCam SA5. Для визуализации свечения при регистрации плазмодинамических процессов субмикросекундного диапазона также использовались высокоскоростные камеры Nanogate и KOIR 010 KOIR 008. Синхронизация, регистрация и анализ сигналов - осциллограмм тока, данных с пъезодатчиков – осуществлялась на базе цифровых комплексов.

2. Описание эксперимента и аппаратуры.

Традиционные оптические методы визуализации течений, основанные на рефракции, развивались и совершенствовались в течении 19 и 20 столетий. Теневой фоновый метод визуализации ТФМ стал возможным только в последние 2 десятилетия с производством цифровых камер высокого разрешения, разработкой программ цифровой обработки изображений, доступностью компьютеров с большим объемом памяти для обработки и хранения изображений и фильмов [4]. Суть теневого фонового метода заключается в сравнении двух изображений одного и того же фона, снятых при отсутствии и при наличии между фотоаппаратом и фоном исследуемого прозрачного объекта с неоднородностями [3]. Изменение показателя преломления вдоль линии наблюдения в случае съемки фона через течение приводит к несовпадению исходного (реперного) и «рабочего» изображений. Проанализировав смещение характерных элементов фона на снимках, можно получить количественную информацию об интегральных показателях преломления исследуемой среды вдоль оптического луча.

Рассмотрим свойства течений, формирующихся в плоском канале при взаимодействии плоской ударной волны с областью вложения энергии конечных размеров. Схема экспериментальной реализации импульсного вложения энергии в сверхзвуковой поток представлена на рис. 1.

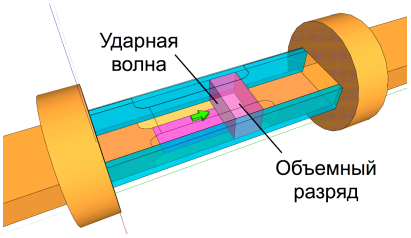


Рис. 1. Схема экспериментальной установки.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

В работе [7] на основе систематического сравнения результатов численной и теневой визуализации проведено исследование квазидвумерного нестационарного разрывного газодинамического течения, возникающего ври взаимодействии плоской ударной волны с импульсным комбинированным разрядом.

Рассмотрены конфигурации и эволюция разрывов. В экспериментах использовалась установка [5-6], основной частью которой являлась ударная труба сечением 48х24мм, состоящую из камеры высокого давления и камеры низкого давления со встроенной разрядной секцией. Рабочий газ - воздух, толкающий газ – гелий. В разрядной камере создается импульсный объемный разряд с предыонизацией ультрафиолетовым излучением от плазменных электродов на участке потока длиной 10 см. Межэлектродное расстояние 24 мм. Разряд запускается в заданный момент движения плоской ударной волны по каналу ударной трубы (разрядной секции) и протекает в два этапа: сначала на верхней и нижней поверхностях камеры горят плазменные листы – сильноточные разряды, скользящие по поверхности диэлектрика (60 – 100 нс). Область плазменного листа является однородной зоной приповерхностного импульсного энерговыделения на противоположных горизонтальных стенках прямоугольного канала. Скользящий разряд создаёт ультрафиолетовое излучение высокой интенсивности и производит ионизацию воздуха в рабочей секции, подготавливая таким образом рабочий объём к созданию однородного импульсного объёмного разряда. Поверхностный разряд создает ударные волны, распространяющиеся в направлении, перпендикулярном направлению потока за падающей ударной волной в канале. На втором этапе происходит замыкание основного разрядного промежутка, возникает объемный разряд и энергия вкладывается в межэлектродную область (100 – 300 нс). Наличие неоднородностей плотности газа (ударной волны) в газодинамическом потоке приводит к пространственному перераспределению плазмы объемного разряда. При наличии ударной волны в межэлектродном пространстве в момент инициирования разряда, область свечения и энерговклад локализуются в зоне низкого давления перед её фронтом. Мгновенное, по сравнению с длительностью характерных процессов в газодинамике, повышение температуры и давления перед фронтом исходной ударной волны при интенсивном импульсном энергоподводе приводит к нарушению условий Ренкина-Гюгонио. Фронт исходной ударной волны в момент времени $t = 0$ становится плоской границей раздела двух областей газа с различными, не связанными друг с другом, параметрами. Таким образом, возникают условия для распада разрыва. Начальный плоский разрыв распадается (в основном течении) на две ударные волны и контактную поверхность между ними.

Посредством анализа динамики двумерных ударно-волновых течений восстановлено пространственное распределение энергии разряда. В каждом эксперименте получали 1 изображение на цифровом фотоаппарате при подсветке лазерным излучением. Длительность экспозиции определялась длительностью лазерного импульса и составляла 6 нс. Для анализа эволюции течения необходима хорошая повторяемость экспериментов и малый разброс параметров. Разброс задержки в системе синхронизации оборудования приводил к тому, что не в каждом эксперименте была получена информация, представляющая научный интерес.

Проведено численное 2D моделирование задачи (рис. 2,а). Целью численного моделирования было уточнение распределения энергии разряда в момент $t = 0$ на основе сравнения экспериментальных и расчетных изображений течения в различные моменты времени. Численное моделирование проводилось в рамках математической модели нестационарных 2D уравнений Эйлера. Область между электродами: $0 \le x \le 10$ см. В начальный момент времени плоская ударная волна с числом Маха $M = 2.3$ для первого варианта задачи расположена в сечении $x_0 = 6$ см, для второй – $x_0 = 3$ см. В эксперименте в объемную часть разряда $x_0 < x < 10$ см, $|y| < 1.1$ см вкладывалось 0.4 Дж, в плазменные листы около стенок $1.1 < |y| < 1.2$ см – по 0.1 Дж. На основании ранее проведенных исследований полагалось, что в поступательные степени свободы мгновенно переходит 50% вложенной энергии.

Наличие цифровой высокоскоростной съемки позволило получать в 1 эксперименте ряд изображений последовательных стадий процесса, включая динамику исходной ударной волны (это дало возможность уточнить скорость падающей волны). На рис. 2.б представлены соответствующие теневые изображения, полученные в одном эксперименте на частоте съемки 300000 кадров в секунду для числа Маха падающей ударной волны M=2.

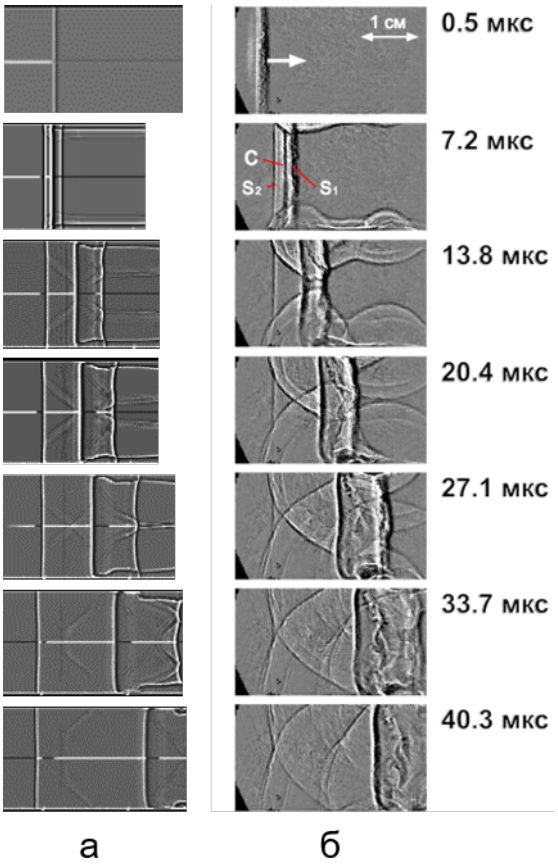


Рис. 2. Численное 2D моделирование (а) и кадры высокоскоростной теневой съемки (б) ударно-волнового течения, возникающего при распаде разрыва на фронте плоской ударной волны. S_1, S_2 – ударные волны, C – контактная поверхность.

Результаты экспериментальной анимации сравниваются напрямую с соответствующими кадрами численной визуализации (Рис. 2.а). Результаты двумерного моделирования представлены в виде синтетических теневых изображений – визуализирована вторая производная плотности газа. [2, 4] Такое представление результатов за последнее десятилетие укоренилось в мировом научном сообществе [3].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Одна из целей сравнения - нахождение энергии локализованного энерговклада путем сравнения численных и экспериментальных изображений с известными контролируемыми параметрами. При экспериментальной цифровой многокадровой видеорегистрации получена возможность проводить расчет и сравнение экспериментальных видеофильмов для каждого эксперимента. При скорости съемки 150000 кадров в секунду сравнение с двумерным численным моделирование возможно для 9-10 последовательных изображений потока, развивающегося после инициирования разряда. При скорости 420 000 кадров/с и 525 000 сравнение с одномерным моделированием возможно на 13-15 изображениях. При такой скорости регистрация ведется из приосевой области, и сравнение идет с одномерным расчетом задачи о распаде разрыва на границе газ – плазма [1].

Для каждой из экспериментальных съёмок течения проводится реперная съёмка в отсутствие течения, с идентичными параметрами. Для улучшения качества полученных цифровых теневых изображений применяется их компьютерная постобработка. Поле интенсивности каждого экспериментального кадра подвергается цифровой обработке: из него вычитается поле интенсивности соответствующего кадра реперной съёмки, и результат сохраняется с линейным преобразованием контраста. Такая обработка позволяет существенно снизить влияние неравномерности освещения поля и дефектов оптических поверхностей, улучшить восприятие газодинамических структур, регистрируемых теневым методом (рис. 3).

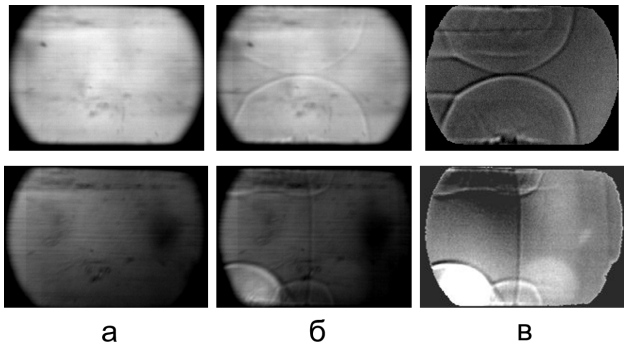


Рис. 3. Пример постобработки теневых изображений: а – реперный кадр, б – рабочий кадр, в – рабочий кадр после вычитания реперного и линейного изменения контраста. Скорость съемки 124 000 кадра/с, экспозиция кадра 1 мкс, разрешение 256 х 176 пикселей.

3. Заключение

Таким образом, современная цифровая высокоскоростная теневая регистрация течений с разрывами существенно повышает возможности эксперимента и точность результатов сравнения расчетной и экспериментальной визуализации течений. При этом возникают новые проблемы, сопровождающие повсеместно переход на цифровые технологии – проблемы, связанные с большими объемами данных: необходимости их систематизации, интерпретации, анализа, построения физической модели. В связи с этим возрастает роль степени подготовленности специалистов – не только в области цифровых технологий, но прежде всего – в области физики, механики, оптики.

4. Благодарности

Работа выполнена при поддержке гранта РНФ 18-19-00672.

5. Литература

[1] Igor Doroshchenko, Irina Znamenskaya, Dmitry Koroteev, and Tahir Kuli-zade., When shock is shocked: Riemann problem dynamics at pulse ionization of a shock wave., Physics of Fluids, Vol. 29, № 10, 2017.

[2] A. A. Fursenko, D. M. Sharov, E. V. Timofeev and P. A. Voinovich. Numerical Simulation of Shock Wave Interactions with Channel Bends and Gas Nonunoformities. Computers Fluids Vol. 21, No. 3, pp. 377-396, 1992.

[3] Gary S. Settles, Michael J. Hargather. A review of recent developments in schlieren and shadowgraph techniques. Meas. Sci. Technol., Vol. 28(4), 2017.

[4] Y. Tamura and K. Fujii. Visualization for computational fluid dynamics and the comparison with experiments, Paper AIAA-90-3031 (1990).

[5] Znamenskaya I.A., Koroteev D.A., Lutsky A.E., Discontinuity breakdown on shock wave interaction with nanosecond discharge., Physics of Fluids, Vol. 20, № 5, 2008.

[6] И. Знаменская, А. Луцкий, Я. Ханхасаева, Ц. Цзинь. Взаимодействие численной и экспериментальной визуализации при исследовании свойств разрядов методом анализа ударно-волновых конфигураций., Научная визуализация. Т. 5, № 3, С. 40-51, 2013.

Об авторах

Знаменская Ирина Александровна, д.ф.-м.н., профессор кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета. Ее e-mail znamen@phys.msu.ru.

Дорощенко Игорь Александрович, аспирант кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета. Его e-mail doroshchenko93@gmail.com.

Наумов Дмитрий Сергеевич, аспирант кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета. Его e-mail n_ds_n@inbox.ru.

Кузнецов Александр Юрьевич, аспирант кафедры молекулярных процессов и экстремальных состояний вещества Московского государственного университета. Его e-mail sasha15555@mail.ru.

Луцкий Александр Евгеньевич, д.ф.-м.н., главный научный сотрудник ИПМ им. М.В. Келдыша РАН, e-mail lutsky@kiam.ru

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Визуализация графовых представлений потоковых программ

В.Н. Касьянов[1,2], Т.А. Золотухин[1,2]
kvn@iis.nsk.su|tzolotuhin@gmail.com
[1] Институт систем информатики, Новосибирск, Россия;
[2] Новосибирский государственный университет, Новосибирск, Россия

В докладе рассматривается задача визуализации графовых представлений потоковых программ. Описывается эффективный алгоритм построения наглядных изображений иерархических графов с портами, а также его эффективная реализация в рамках системы Visual Graph.

Ключевые слова: *графовые представления потоковых программ, графы с портами, визуализация графов, иерархические графы, системы визуализации графов.*

Visualization of graph presentations of data-flow programs

V.N. Kasyanov[1,2], T.A. Zolotuhin[1,2]
kvn@iis.nsk.su|tzolotuhin@gmail.com
[1]Institute of Informatics Systems, Novosibirsk, Russia;
[2]Novosibirsk State University, Novosibirsk, Russia

In this paper, the problem of visualization of graph representations of data-flow programs is considered. An effective algorithm for constructing visual images of hierarchical graphs with ports is described, as well as its effective implementation within the framework of the Visual Graph system.

Keywords: *graph presentations of data-flow programs, graphs with ports, visualization of graph, hierarchical graphs, graph visualization systems.*

1. Введение

Используя традиционные языки и методы, очень трудно разработать высококачественное, переносимое программное обеспечение для параллельных компьютеров. В частности, параллельное программное обеспечение не может быть разработано с малыми затратами на последовательных компьютерах и потом перенесено на параллельные вычислительные системы без существенного переписывания и отладки. Поэтому высококачественное параллельное программное обеспечение может разрабатываться только небольшим кругом специалистов, имеющих прямой доступ к дорогостоящему оборудованию. Однако, используя языки программирования с неявным параллелизмом, такие как потоковый язык Cloud Sisal [8], можно преодолеть этот барьер и предоставить широкому кругу специалистов, которые не имеют доступа к параллельным вычислительным системам, но могут многое сделать в своих прикладных областях исследований, возможность быстрой разработки переносимых параллельных алгоритмов на своем рабочем месте.

Семантика языков программирования с неявным параллелизмом гарантирует детерминированные результаты для параллельной и последовательной реализаций — то, что невозможно гарантировать для традиционных языков, подобных языку Фортран. Более того, неявный параллелизм языка снимает необходимость переписывания исходного кода при переносе его с одного компьютера на другой. Гарантировано, что программа с неявным параллелизмом, правильно исполняющаяся на персональном компьютере, будет также давать правильные результаты на высокоскоростном параллельном или распределенном вычислителе.

В докладе рассматривается методы и алгоритмы визуализации графовых представлений программ на языке Cloud Sisal в рамках облачной системы параллельного программирования CPPS (Cloud Parallel Programming System). CPPS – это система, которая разрабатывается как интегрированная облачная среда программирования на языке Cloud Sisal, которая содержит как интерпретатор, поддерживающий диалоговое взаимодействие с пользователем при построении и отладке программы, так и оптимизирующий кросс-компилятор, осуществляющий построение параллельной программы по ее функциональной спецификации. [9]

Цель создания системы – дать возможность широкому кругу лиц, находящихся в удаленных населенных пунктах или в местах с недостаточными вычислительными средствами, но имеющих выход в Интернет, дистанционно и без установки дополнительного программного обеспечения на своих недорогих вычислительных устройствах в визуальном стиле создавать и отлаживать переносимые параллельные программы на языке Cloud Sisal, а также затем дистанционно (в облаке) осуществлять эффективное решение своих задач, исполняя на некоторых супервычислителях, доступным им по сети, созданные и отлаженные переносимые Cloud-Sisal-программы, предварительно адаптировав их под используемые супервычислители с помощью облачного оптимизирующего кросс-компилятора, предоставляемого системой.

На рынке широко представлены наукоемкие программные продукты, использующие методы визуализации информации на основе графов. В основном это многочисленные специализированные системы или встроенные компоненты систем, ориентированные на визуализацию определенных подклассов графов, но есть и универсальные системы, предназначенные для визуализации графов общего вида и широкого назначения, такие как Higres, aiSee, yEd и Cytoscape [6].

2. Графовое представление Cloud Sisal программ

Система CPPS использует графовое представление Cloud Sisal программ, которое ориентировано на их семантическую и визуальную обработку и основано на атрибутированных иерархических графах [4]. Это представление, получившее название IR (Intermediate

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Representation), в отличие управляющего графа [5], обычно используемого в оптимизирующих компиляторах для императивных языков (таких как C или Фортран), выражает не поток управления, а поток данных в программе.

Вершины IR графа соответствуют выражениям программы, а дуги отражают передачи данных между портами вершин, упорядоченные множества которых приписаны вершинам в качестве их аргументов (входных портов или входов) и результатов (выходных портов или выходов). В силу свойства языка Cloud Sisal этот граф является ациклическим и не содержит двух дуг, заходящих в один и тот же вход.

Вершины обозначают операции над своими входами (аргументами), результаты которых находятся на выходах вершин. Существует, однако, специальный вид вершин, обозначающих литералы (константы) любого типа, каждая из которых имеет один выход и пустое множество входов.

Вершины бывают простыми и составными. Простые вершины (или просто вершины) не имеют внутренней структуры помимо ассоциированной с ними операции. Составные вершины (или фрагменты) соответствуют составным выражениям Cloud Sisal программы, таким, как, например, циклическое выражение или тело функции, и дополнительно непосредственно содержат упорядоченные множества вершин, соответствующих выражениям, из которых они состоят. Для каждого фрагмента F количество и типы того множества вершин M, которые непосредственно содержатся в нем, а также того множества дуг (p, q), которые существуют между его портами и портами этих вершин, задаются типом (или семантикой) составного выражения, но удовлетворяют следующему свойству: p – выход некоторой вершины из M или вход фрагмента F, а q - вход некоторой вершины из M или выход фрагмента F. Понятно, что фрагмент F помимо указанных выше вершин M и инцидентных их портам дуг, которые непосредственно содержатся в F, будет содержать и другие вершины и дуги графа, если среди вершин из M есть фрагменты, но в силу свойств языка Cloud Sisal среди них нет дуг, которые инцидентны портам фрагмента F и не являются непосредственно вложенными в F.

3. Изображения IR графов

Предполагается, что IR представления Cloud Sisal программ демонстрируются пользователям системы наряду с их текстовыми представлениями и используются пользователями для целей визуальной отладки Cloud Sisal программ и управляемой их оптимизации. Возникает необходимость автоматического построения наглядного изображения IR графов.

Предполагается следующие правила изображения IR графов:

1. Простые вершины без входов изображаются в виде кругов, содержащих обозначения констант.

2. Простые вершины с входами изображаются в виде прямоугольников с полукруговыми выступами сверху, изображающими входы вершины в их упорядоченности слева направо, и полукруговыми выступами снизу, изображающими выходы вершины. Внутри прямоугольников находятся обозначения операций.

3. Составные вершины изображаются в виде прямоугольников с прямоугольным выступом сверху слева для указания типа составной вершины, а также с кругами сверху справа и снизу для изображения входов и выходов этой вершины в их упорядоченности слева направо. Каждый круг, изображающий некоторый полюс фрагмента, состоит полукругового выступа наружу и полукругового выступа внутрь этого прямоугольника. Внутри прямоугольника, изображающего составную вершину, находятся изображения всех вершин и всех дуг, в ней содержащихся, и только они.

4. Изображения двух разных вершин либо не пересекаются, либо одно из них целиком лежит в другом.

5. Дуги изображаются в виде кривых (сплайнов) со стрелками, которые соединяют соответствующие порты и не пересекают изображения вершин по своим внутренним точкам.

Основные трудности решения построения наглядного изображения IR графов связаны с тем, что в отличие от стандартной задачи укладки графа на плоскости [2-7] в IR графе вершины графа соединяются дугами через свои порты и могут иметь разный размер, который определяется характеристиками изображения тех графов, которые вложены в данную вершину.

4. Общая схема алгоритма

При описании алгоритма предполагается, что исходный граф является дэгом, т. е. ациклическим ориентированным графом. Если исходный граф является неориентированным и/или содержит циклы, то он сначала может быть приведен к дэгу путем задания и/или смены ориентации у части его ребер, а затем по построенной укладке дэга обратным преобразованием, примененным к изображениям дуг, получаем укладку исходного графа [3, 5].

Алгоритм состоит в последовательном выполнении шагов построения изображений содержимого фрагментов исходного графа, начиная с самого внутреннего, на каждом из которых происходит укладка некоторого фрагмента с использованием размеров и расположений непосредственно вложенных в него вершин.

Само построение изображения одного текущего фрагмента базируется на методике так называемого поуровневого подхода к построению укладки ациклического ориентированного графа, которая была предложена К. Сугиямой (K. Sugiyama), и состоит из следующих трех основных этапов [3, 5]:

1. Распределение вершин по уровням так, чтобы дуги следовали одному направлению, т. е. определение y-координат для вершин.
2. Выбор порядка вершин на уровне с целью минимизации пересечений дуг.
3. Определение x-координат вершин на уровне с целью минимизации общей длины дуг и количества сгибов.

Таким образом, данный подход позволяет разбить задачу нахождения расположения ациклического графа на три достаточно независимых шага, реализация каждого из которых может опираться на свои методы и улучшать изображение по тем или иным эстетическим критериям.

При размещении ациклических графов обычно во внимание принимается следующие эстетические критерии:

- совпадающее направление дуг;
- минимизация площади изображения;
- равномерность распределения вершин по площади всего изображения;
- уменьшение числа пересечений дуг;
- отсутствие слишком длинных дуг;
- уменьшение числа сгибов дуг.

Одновременная оптимизация изображения по всем перечисленным критериям является невозможной, т. к. критерии зачастую оказываются противоречащими друг другу. Например, уменьшение числа сгибов дуг может нарушать равномерность распределения вершин и приводить к увеличению общей площади изображения.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

5. Распределение вершин по уровням

Задачей данного шага является присваивание каждой вершине ее конечной вертикальной *y*-координаты. Для этого исходный ациклический граф $G = (V, E)$ должен быть приведен к *поуровневому* представлению, которое есть разбиение V на подмножества $L_1, L_2, \ldots, L_h$, так что для каждой дуги $(u,v) \in E$, где $u \in L_i$ и $v \in L_j$, верно то, что $i > j$. Подразумевается, что все вершины одного уровня будут расположены на одной горизонтальной прямой. *Высотой* представления называется количество уровней h, а *шириной* w_i – наибольшее число вершин на одном уровне L_i. *Зазором* дуги (u,v), где $u \in L_i$ и $v \in L_j$, называется разность $i - j$. Представление называется *сжатым*, если не существует дуги с зазором, большим единицы. После разбиения вершин по уровням всем вершинам одного уровня `присваивается одна и та же вертикальная *y*-координата, т. е. $v_y = i$ для всех вершин с уровня L_i.

Существование поуровневого представления для любого ациклического графа очевидно. Однако не для всякого ациклического графа существует его сжатое представление — это может быть продемонстрировано уже для графа, состоящего всего из трех вершин (см. Рис. 1). Необходимость же построения сжатого разбиения диктуется следующими двумя причинами. Во-первых, сжатое разбиение минимизирует горизонтальную длину дуг и количество их сгибов. При сжатом разбиении дуги связывают вершины только соседних уровней и могут быть изображены прямолинейными отрезками. Во-вторых, сжатое разбиение существенно упрощает задачу минимизации пересечений дуг графа, которая решается на втором шаге алгоритма. Сжатое разбиение сводит эту глобальную задачу к задаче минимизации пересечений для двух соседних уровней.

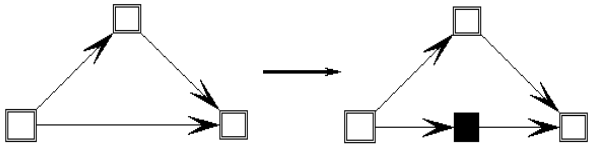


Рис. 1. Пример укладки графа, в которой вершины одного уровня расположены на одной вертикальной линии. Заштрихованная вершина является фиктивной.

Для построения сжатого разбиения произвольного ациклического графа применяется техника вставки *фиктивных* вершин. Фиктивные вершины добавляются в граф вдоль длинных дуг, т. е. дуг с зазором, большим единицы. Для каждой такой дуги (u, v), $u \in L_i$ и $v \in L_j$, с зазором $k = i - j > 1$ добавляется $k - 1$ вершин $u_1, \ldots u_{k-1}$, таких, что $u_m \in L_{i-m}$, и дуг (u_l, u_{l+1}) для всех $l < k - 1$. Такое построение приводит к тому, что в графе остаются только дуги, соединяющие вершины соседних уровней, и задача минимизации пересечения дуг сводится к соответствующей задаче для двухуровневого случая. Данное построение заодно решает и задачу проведения дуг в конечном изображении. Дуга изображается в виде сплайна, концами которого являются сами вершины, а другими опорными точками (точками сгиба) — вставленные на первом этапе фиктивные вершины. Таким образом, исчезает возможность пересечения дуг с вершинами.

Выполнение данного этапа связано с обработкой графа $G = (V, E)$, представляющего некоторый фрагмент F, в котором V состоит из полюсов фрагмента F и вершин, непосредственно содержащихся в F. Пусть t — длина самого длинного пути по G. Тогда будет построено поуровневое представление $L_1, L_2, \ldots, L_{t+1}$, в котором L_1 — все входы F, а L_{t+1} — все выходы F.

Сначала из полюсов фрагмента строятся L_1 и L_{t+1} и эти полюса вместе с инцидентными дугами удаляются из G. Процесс построения продолжается по шагам, на каждом из которых одна из вершин, не имеющая заходящих дуг в текущем состоянии G, включается в множество L_{i+1}, где i — максимальный номер множества L_i, содержащего ее предшественника в исходном графе G, с одновременным удалением этой вершины из G вместе со всеми исходящими из нее дугами. Причем, среди вершин, не имеющих заходящих дуг в текущем состоянии G, выбирается и включается в соответствующее множество та вершина, у которой было наименьшее число входящих дуг и наибольшее число исходящих дуг в исходном графе G (эти числа предварительно подсчитываются для всех вершин исходного графа). Этот процесс продолжается пока текущий граф G не станет пустым. После этого проводится приведение построенного представления графа G к сжатому виду, используя технику фиктивных вершин.

6. Выбор порядка вершин на уровне

Задачей данного этапа является нахождение порядка вершин на каждом уровне, с целью минимизации количества пересечений дуг.

Следует отметить, что количество пересечений дуг в поуровневом представлении графе не зависит от конечных горизонтальных координат вершин, а зависит только от их относительного положения внутри каждого уровня (их порядкового номера на данном уровне). Таким образом, задача данного этапа является не геометрической, а всего лишь комбинаторной. Однако эта задача является NP-полной уже для графа, имеющего всего лишь два уровня.

Выполнение данного этапа для заданного фрагмента осуществляется следующим образом:

1. Если у фрагмента есть входные и/или выходные порты, размещенные на уровне L_1 и/или L_h соответственно, то следует приписать этим портам соответствующие им порядковые номера.
2. Рассматриваем вершины уровня L_1 в их упорядоченности, если у фрагмента нет портов, или L_2, если порты у него есть, и осуществляем обход в глубину содержимого фрагмента, начиная с этих вершин, с использованием стека.
3. Если стек пуст, то либо продолжаем шаг 2, либо данный этап завершен. В противном случае рассматриваем вершину, находящуюся на верху стека, но не удаляем ее из стека. Если не получивших номеров преемников у рассматриваемой вершины нет, то присваиваем вершине текущий номер порядка, удаляем ее из стека и переходим на шаг 3. Если же такие преемники у вершины есть, то добавляем одну из них в стек и переходим на шаг 3; при выборе преемника для размещения в стек учитываем следующее:
   - Если все преемники данной вершины соединены дугами, исходящими из разных портов, то порядок включения этих вершин не противоречит порядку портов.
   - Если среди преемников есть вершины, связанные с вершинами, которые уже получили порядковые номера, то они включаются раньше других.
   - Если среди преемников есть фиктивные вершины, то они выбираются в таком порядке, чтобы они разместились на уровне посередине, а реальные вершины на краях уровня.

7. Определение координат вершин на уровне

После определения порядка вершин на уровне необходимо определить их реальные координаты. Дуги

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

графа изображаются в виде ломанных с точками излома, находящимися в фиктивных вершинах, поэтому задача определения окончательных координат всех вершин одновременно является и задачей проведения дуг. Если же дуги графа имеют какую-либо другую форму и / или способ проведения, то соответствующие критерии должны быть рассмотрены при решении задачи об определении координат вершин.

На входе данного этапа имеем разбиение вершин по уровням L_1, L_2, … , L_h, вершины на каждом уровне имеют порядок от 1 до w_i, где i изменяется от 1 до h. Причем самое большое количество вершин находится на уровне L_1 (или на уровне L_{h-1}, если выходных портов нет).

Начиная с последнего уровня, используя метод барицентров в сочетании с ограничениями на порядок вершин, полученными на предыдущем шаге, определяем x-координату на предыдущем уровне.

Вершины последнего уровня распределяются равномерно на некотором отрезке горизонтальной прямой, выделенной под вершины этого уровня. Затем для каждого следующего уровня координаты его вершин последовательно определяются как среднее арифметическое координат их соседей из уже поставленных уровней. При этом, не допускается нарушение изначального порядка вершин на уровне.

Если на каком-то шаге происходит наслоение, т. к. двум вершинам присвоится одна x-координата, то следует немного раздвинуть последний уровень. Например, предположим, что изначально вершины на этом уровне имеют x-координаты 0, 1, 2, 3, 4 и т. д. При следующем шаге они будут иметь x-координаты 0, 2, 4, 6, 8 и т.д. А если снова не получится уложить граф, то можно еще кратно увеличить основание: 0, 3, 6, 9 и т. д.

8. Реализация алгоритма

При описании алгоритма предполагалось, что исходный граф является ациклическим ориентированным графом. Однако его реализация в рамках системы Visual Graph [6] выполнена для более общего случая и включает также шаги предварительной и окончательной обработки графа, позволяющие системе работать не только с дэгами. Суть и цель этих преобразований заключается в обратимом преобразовании структуры исходного графа, так чтобы после размещения дэга, возможно было безболезненно для качества получаемого изображения восстановить структуру исходного графа и, тем самым, построить его изображение.

Если исходный граф содержит неориентированные ребра, то в качестве предварительного шага реализовано задание им направления в соответствии с обходом графа в глубину с удалением ориентации при окончательной обработке.

Если исходный граф содержит циклы, то происходит изменение ориентации у части дуг, входящих в цикл, а затем в построенной укладке дэга обратным преобразованием строится изображение исходного графа.

Поскольку задача нахождения минимального числа дуг, изменение ориентации которых переводит граф в дэг, является NP-полной, нами используется простой метод нахождения таких дуг: в процессе обхода в глубину выделяются все обратные дуги, у которых и происходит смена ориентации. И хотя у произвольного графа $G=(V,E)$ обратных дуг может быть достаточно много (порядка $|E| - |V| - 1$), для графов программ, на которые ориентирована система, это не имеет значения, поскольку у таких графов (в частности, для управляющих графов программ) есть четко выраженное направление и, как правило, если и есть, то незначительное количество обратных дуг, образующих циклы.

Если граф содержит вершины с более чем одним предшественником, то у каждой из них оставляем только одного по следующему правилу. Если среди предшественников вершины есть только одна фиктивная или только одна не фиктивная вершина, оставляем ее одну. Иначе ищем общего предшественника для предшественников данной вершины (в случае необходимости создаем фиктивного предшественника) и прокладываем путь от него до нашей вершины.

Предполагается, что система Visual Graph взаимодействует с интерпретатором и кросс-компилятором системы CPPS через представление IR графа Cloud Sisal программы на языке GraphML [1], который позволяет специфицировать произвольного вида атрибутированные иерархические графы с полюсами.

9. Заключение

Рассмотренный в докладе алгоритм укладки графов с портами имеет квадратичную временную сложность и позволяет строить наглядные изображения графовых представлений потоковых Cloud Sisal программ.

Реализация алгоритма в рамках системы Visual Graph также достаточно эффективна, поскольку она позволяет на обычном персональном компьютере за реальное время (без видимых задержек) визуализировать произвольного вида атрибутированные иерархические графы с портами, содержащие порядка 10000 элементов (вершин и дуг).

10. Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект 18-11-00118).

11. Литература

[1] Di Battista G., Eades P., Tamassia R., Tollis I. G. Graph Drawing: Algorithms for Vizualization of Graphs. - Prentice Hall, 1999.

[2] Drawing Graphs. Methods and Models. - Berlin: Springer, 2001.

[3] Sugiyama K. Graph drawing and applications. For software and knowledge engineers. - World Scientific. 2002.

[4] Касьянов В.Н. Иерархические графы и графовые модели: вопросы визуальной обработки // Проблемы систем информатики и программирования. - Новосибирск, 1999. - С. 7-32.

[5] Касьянов В.Н., Евстигнеев В.А. Графы в программировании: обработка, визуализация и применение. - СПб.: БХВ-Петербург, 2003.

[6] Касьянов В.Н., Золотухин Т.А. Visual Graph — система для визуализации сложно структурированной информации большого объема на основе графовых моделей // Научная визуализация. – 2015. - Том. 7, N 4. - С. 44 – 59.

[7] Касьянов В. Н., Касьянова Е. В. Визуализация информации на основе графовых моделей // Научная визуализация – 2014. – Т. 6, N 1. – С. 31 - 50.

[8] Касьянов В. Н., Касьянова Е. В. Язык программирования Cloud Sisal. – Новосибирск, 2018. – 42 с. – (Препринт/ РАН, Сиб. отд-ние, ИСИ; N181).

[9] Касьянов В.Н., Касьянова Е.В. Методы и система облачного параллельного программирования // Проблемы оптимизации сложных систем: Материалы XIV Международной Азиатской школы-семинара. – Алматы, 2018. – Часть 1, С. 298-307.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Визуализация препятствий на индикаторе на лобовом стекле с помощью оценки оптического потока

В.В. Князь[1,2], С.Ю. Данилов[1,2], О.В. Выголов[1], А.Н. Бордодымов[1]
vl.kniaz@gosniias.ru|danilov@gosniias.ru|o.vygolov@gosniias.ru|bordodymov@gmail.com
[1]125319, Москва, ул. Викторенко, 7;
[2]Московский физико-технический институт (государственный университет), Москва, Россия

Индикатор на лобовом стекле (ИЛС) позволяет снизить нагрузку на экипаж самолёта во время сложных этапов полёта, таких как взлёт, посадка и рулёжка. Перспективным средством повышения ситуационной осведомлённости экипажа является отображение на ИЛС видеопоследовательности от датчика дальнего инфракрасного диапазона. Однако для практической визуализации кадров ИК-камеры на ИЛС требуется предварительная обработка видеопоследовательности для снижения яркости фоновых объектов и повышения яркости препятствий. В данной статье рассматривается алгоритм фильтрации кадров системы улучшенного видения для эффективной визуализации на ИЛС на основе оценки оптического потока.

Ключевые слова: *визуализация, оптический поток, индикатор на лобовом стекле, визуальная аналитика.*

Obstacle visualization on a head up display using optical flow estimation

V.V. Kniaz[1,2], S.Yu. Danilov[1,2], O.V. Vygolov[1], A.N. Bordodymov[1]
vl.kniaz@gosniias.ru|danilov@gosniias.ru|o.vygolov@gosniias.ru|bordodymov@gmail.com
[1]State Res. Institute of Aviation Systems (GosNIIAS), 125319, 7, Victorenko str., Moscow, Russia;
[2]Moscow Institute of Physics and Technology (MIPT), Russia

A Head up Display (HUD) reduces cabin crew payload during complicated stages of the flight such take-off, landing and taxi. Projection of an infrared sensor video on a HUD provides a perspective means for increasing of crew situational awareness. However the visualization of an infrared video on a HUD in practice requires video preprocessing for reduction of brightness of background objects and increase of brightness of obstacles. This paper is focused on the algorithm for effective visualization of EVS video on a HUD using optical flow estimation.

Keywords: *visualization, optical flow, head up display, visual analytics.*

1. Введение

Использование в составе информационного поля кабины экипажа ИЛС позволяет обеспечить значительное улучшение ситуационной осведомленности экипажа за счёт компактного и информативного представления данных и их удобного отображения. Визуализация необходимой полётной информации на ИЛС в режиме полупрозрачного «наложения» на реальную сцену, наблюдаемую пилотом в процессе управления, позволяет обеспечить концентрацию внимания. Такой подход позволяет предоставить пилоту критически важную информацию в ходе сложных этапов полета, таких как взлёт, посадка и рулёжка.

Целью данной работы является разработка алгоритма, производящего снижение яркости малоинформативных областей и повышение яркости динамических препятствий на видеопоследовательности системы улучшенного видения для эффективного отображения на индикаторе на лобовом стекле. Рассматриваемый подход основывается на гипотезе о том, что комплексирование данных систем синтезированного (ССВ) и улучшенного видения (СУВ) позволяет увеличить контрастность кадра СУВ при отображении на ИЛС. Такое комплексирование применяется в системе комбинированного видения [1, 8, 9], где производится смешивание значений интенсивности кадров двух систем для включения в состав виртуальной сцены динамических препятствий. Обзор современной литературы показал, что до настоящего времени не проводилось исследований алгоритмов комплексирования кадров систем синтезированного и улучшенного видения, обеспечивающих фильтрацию кадра датчика инфракрасного диапазона для эффективного отображения на ИЛС.

(a) ССВ **(b) СУВ** **(c) ИЛС**

Рис. 1. Системы воздушных судов, используемые в предлагаемом методе: (а) Система синтезированного видения, (б) Система улучшенного видения, (с) Индикатор на лобовом стекле с кадром системы синтезированного видения.

2. Постановка задачи

Рассматривается задача фильтрации данных, поступающих от датчика инфракрасного диапазона, для эффективного представления на индикаторе на лобовом стекле. Дан комплекс программного обеспечения, моделирующего бортовые системы синтезированного и улучшенного видения, трёхмерная модель рельефа и

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

объектового состава аэропорта Сочи (рис. 1). Требуется разработать алгоритм, производящий снижение яркости малоинформативных областей и повышение яркости динамических препятствий.

3. Алгоритм фильтрации ClearHUD

Входными данными для алгоритма являются последовательности кадров систем синтезированного и улучшенного видения. На выходе алгоритм выдаёт монохромную видеопоследовательность, представляющую собой кадр улучшенного видения, подготовленный для проекции на ИЛС. Метод комплексирования данных основан на предположении о том, что реальный оптический поток, построенный по кадрам системы улучшенного видения, и модельный оптический поток, построенный по данным системы синтезированного видения, должны быть равны друг другу, если в поле зрения датчика отсутствуют динамические препятствия. Следовательно, разница модельного и реального оптических потоков позволяет выделить все динамические препятствия, присутствующие в поле зрения датчика системы улучшенного видения. Словесное описание разработанного алгоритма ClearHUD представлено в листинге 1.

Algorithm 1: ClearHUD

Input: Положение камеры $\mathbf{X}_0^i,\ \mathbf{X}_0^{i-1}$, текущий кадр $\mathbf{I}_r^i$ и предыдущий кадр $\mathbf{I}_r^{i-1}$ поступающие от СУВ, текущий кадр от ССВ $\mathbf{I}_s^i$

Output: Кадр СУВ $\mathbf{I}_{HUD}$, обработанный для эффективного отображения на ИЛС (HUD).

1 Оценить оптический поток СУВ $\mathbf{U}_r^i = \texttt{FlowNet2-SD}(\mathbf{I}_r^{i-1}, \mathbf{I}_r^i)$;

2 Оценить оптический поток ССВ $\mathbf{U}_s^i = \texttt{RayTracing}(\mathbf{X}_0^i,\ \mathbf{X}_0^{i-1})$, eq. (1);

3 Вычислить разность оптических потоков $\mathbf{U}_d = \mathbf{U}_r - \mathbf{U}_s$;

4 Отрисовать кадр ССВ $\mathbf{I}_s^i$ с использованием монохромной палитры;

5 Сформировать улучшенный кадр СУВ с использованием алгоритма «Retinex» $\mathbf{I}_{re}^i = \texttt{Retinex}(\mathbf{I}_r^i)$;

6 Найти сумму кадра ССВ и разности оптических потоков для формирования канала прозрачности $\mathbf{I}_a = \frac{\mathbf{I}_s^i}{255} + \texttt{clip}(\mathbf{U}_d, 0, 1)$;

7 Сформировать результирующий кадр путём умножения кадра СУВ на альфа-канал (канал прозрачности) $\mathbf{I}_{HUD} = \mathbf{I}_{re}^i \cdot \mathbf{I}_a$;

Алгоритм ClearHUD принимает следующий набор входных данных: текущий кадр $\mathbf{I}_r^i$ и предыдущий кадр $\mathbf{I}_r^{i-1}$ системы улучшенного видения, текущий кадр системы синтезированного видения $\mathbf{I}_s^i$, положение виртуальной камеры $\mathbf{X}_0^i,\ \mathbf{X}_0^{i-1}$ для текущего и предыдущего кадров, а также трёхмерную модель сцены, заданную системой синтезированного видения. На первом этапе производится оценка реального оптического потока $\mathbf{U}_r$ с использованием глубокой свёрточной нейронной сети FlowNet2-SD. Параллельно с оценкой реального оптического потока производится расчёт модельного оптического потока $\mathbf{U}_s$ с использованием метода трассировки лучей. Практическая применимость данного метода подтверждена в работе [7]. Затем вычисляется разность оптических потоков $\mathbf{U}_d = \mathbf{U}_r - \mathbf{U}_s$.

Следующим этапом является предварительная обработка кадров системы улучшенного и синтезированного видения. Предварительная обработка кадра синтезированного видения заключается в отрисовке трёхмерной модели сцены с использованием специальной монохромной палитры. В данной палитре наиболее высокая яркость присвоена наиболее важным визуальным ориентирам (разметка ВПП, горы). Предварительная обработка кадров системы синтезированного видения выполняется с использованием модифицированного алгоритма «Retinex» [6, 8].

На следующем этапе формируется канал прозрачности, заданный разностью модельного и реального оптических потоков $\mathbf{U}_d$ с и предварительно обработанным кадром системы синтезированного видения. Изображение $\mathbf{I}_a$ представляет собой канал прозрачности кадра, который будет отображён на ИЛС. Для построения окончательного изображения предварительно обработанный кадр системы синтезированного видения умножается на канал прозрачности $\mathbf{I}_a$. Структурна схема алгоритма представлена на рисунке 2.

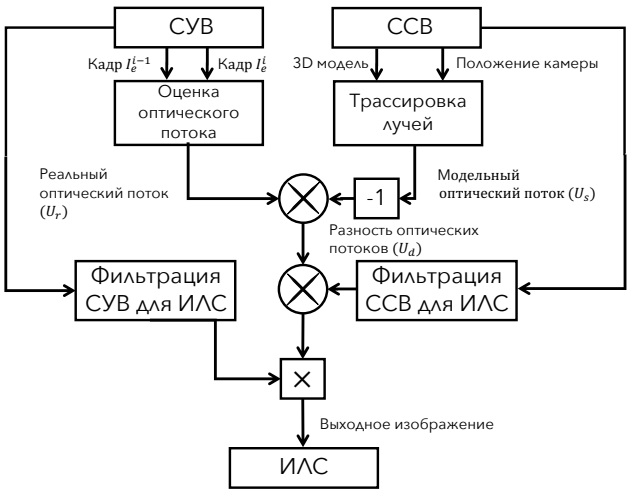


Рис. 2. Структурна схема разработанного алгоритма.

Оценка модельного оптического потока. Для построения модельного оптического потока использовался метод трассировки лучей. Пусть $\mathbf{x}_p^i$ проекция точки сцены в i-ом кадре. Тогда модельный оптический поток задан [4]:

$$\mathbf{u}_s^i = \mathbf{x}_p^i - \mathbf{x}_p^{i-1}, \tag{1}$$

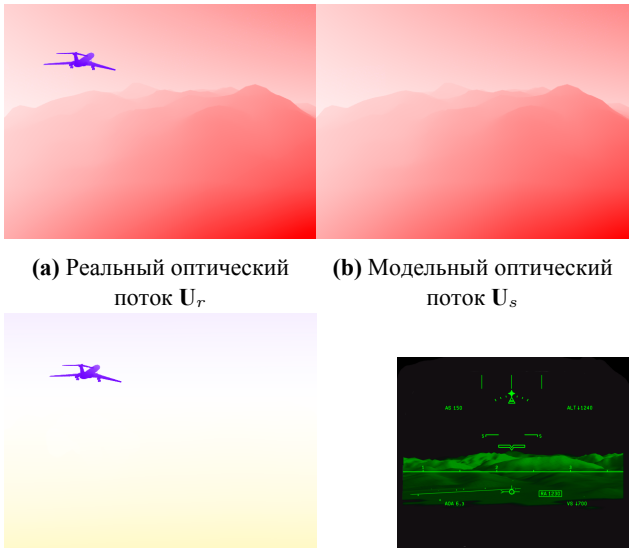
где $\mathbf{x}_p^i$ задано формулами:

$$\mathbf{x}_p^{i-1} = \mu^{i-1} \cdot \mathbf{R}_{oc}^{i-1} \cdot (\mathbf{X}_p - \mathbf{X}_0^{i-1}), \tag{2}$$

$$\mathbf{x}_p^i = \mu^i \cdot \mathbf{R}_{oc}^i \cdot (\mathbf{X}_p - \mathbf{X}_0^i), \tag{3}$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

где μ^i – масштабный коэффициент, $\mathbf{R}^i_{oc}$ – матрица перехода из системы координат сцены в систему координат камеры, $\mathbf{X}_p$ – трехмерная координата точки сцены, $\mathbf{X}^i_0$ – положение камеры. Разность двух последних уравнений даёт окончательное уравнение для модельного оптического потока $\mathbf{u}^i_m$. Примеры реального и модельного оптических потоков представлены на рисунке 3.



(a) Реальный оптический поток $\mathbf{U}_r$

(b) Модельный оптический поток $\mathbf{U}_s$

(c) Разность оптических потоков $\mathbf{U}_d$

(d) Кадр системы синтезированного видения обработанный для отображения на ИЛС

Рис. 3. Пример оценки реального и модельного оптического потока.

Оценка реального оптического потока. Проведённый анализ современных алгоритмов оценки оптического потока показал, что наиболее перспективным подходом является глубокая свёрточная нейронной сеть FlowNet 2.0 [2, 3]. В работах [2, 3] представлено несколько вариантов архитектуры сети: неглубокие сети, обеспечивающие высокое быстродействие, и глубокие сети, обеспечивающие высокую точность оценки оптического потока. Наиболее подходящим вариантом архитектуры сети по критерию точность-быстродействие представляется архитектура FlowNet2-SD. Данная архитектура включает 13 свёрточных (конволюционных) слоёв, 1 полносвязный слой и 12 деконволюционных слоёв. Архитектура сети представлена на рисунке 4.

Обучение сети. Обучение разработанной реализации сети FlowNet2-SD производилось с использованием библиотеки PyTorch. Вычисления производились на графическом процессоре NVIDIA 1080 Ti. Полный процесс обучения занял 40 часов. В качестве обучающей выборки использовалось объединение трёх выборок: MPI-Sintel Flow [5], многоспектральной обучающей выборки ZEFIRUS [4] и ClearHUD, предназначенной для решения задач обнаружения динамических препятствий. Выборка ClearHUD, моделирующая заход на посадку, подготовлена с использованием средств 3D моделирования. Ошибка обученной сети составила 5,1 пиксел.

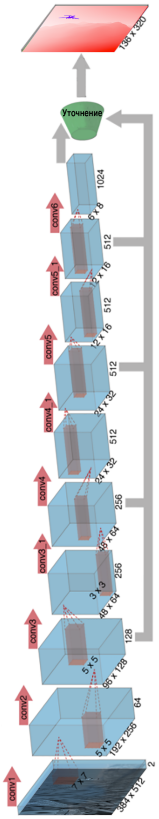


Рис. 4. Архитектура глубокой свёрточной нейронной сети FlowNet2-SD [2].

4. Анализ результатов

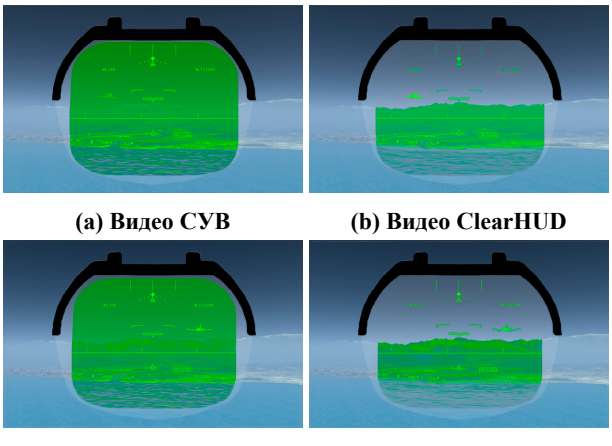
Для тестирования алгоритма разработан программный имитатор индикатора на лобовом стекле, позволяющий производить отображение кадров систем улучшенного и синтезированного видения, а также результатов работы алгоритма ClearHUD. Произведено тестирование алгоритма ClearHUD с использованием истинной разметки областей динамических препятствий. Результаты приведены на рисунке 5. Алгоритм продемонстрировал высокую точность обнаружения динамических препятствий, значение коэффициента Жаккара составило 0,88для дневных последовательностей и 0,75для ночных последовательностей. Выполнено тестирование работы алгоритма на видеопоследовательностях, моделирующих различные этапы полёта: взлёт, заход на посадку, рулёжка.

5. Заключение

Разработан алгоритм ClearHUD, производящий снижение яркости малоинформативных областей и повышение яркости динамических препятствий. Алгоритм использует видеопоследовательности от систем

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

улучшенного и синтезированного видения в качестве входных данных и производит комплексирование кадров двух систем для снижения яркости малоинформативных областей в кадре системы улучшенного видения и увеличения яркости динамических препятствий.



(a) Видео СУВ **(b) Видео ClearHUD**

(c) Видео СУВ **(d) Видео ClearHUD**

Рис. 5. Примеры обработки изображений с помощью алгоритма ClearHUD.

Уменьшение яркости выполняется на основе сегментации областей из кадра системы синтезированного видения. Для повышения яркости динамических препятствий производится их обнаружение на основе вычисления разности оптических потоков кадров систем улучшенного и синтезированного видения. Реальный оптический поток оценивается по видеопоследовательности системы улучшенного видения. Модельный оптический поток рассчитывается с помощью трассировки лучей и 3D модели сцены системы синтезированного видения.

Разработанный алгоритм производит параллельную обработку данных и может быть реализован на платформе современной интегрированной модульной авионики. Вычислительная сложность алгоритма позволяет реализовать его для обработки кадров системы улучшенного видения с разрешением 640×480 пикселей, с использованием 3 процессоров с тактовой частотой 1,8 ГГц и одного графического процессора NVIDIA Tegra. Дальнейшая работа будет посвящена реализации и тестированию алгоритма ClearHUD с использованием бортового оборудования.

6. Благодарности

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант РФФИ 17-29-03185.

7. Литература

[1] Design and Flight Test of a Primary Flight Display Combined Vision System / Patricia May Ververs, Gang He, John Suddreth et al. // SAE International Journal of Aerospace. — 2011. — Vol. 4, no. 2011-01-2525. — P. 738–750.

[2] FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks / E. Ilg, N. Mayer, T. Saikia et al. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — Jul. — URL: http://lmb.informatik.uni-freiburg.de//Publications/2017/IMKDB17.

[3] FlowNet: Learning Optical Flow with Convolutional Networks / Alexey Dosovitskiy, Philipp Fischery, Eddy Ilg et al. // 2015 IEEE International Conference on Computer Vision (ICCV). — 2015. — P. 2758–2766.

[4] Kniaz V V. A photogrammetric technique for generation of an accurate multispectral optical flow dataset // SPIE Optical Metrology / Ed. by Fabio Remondino, Mark R Shortis. — SPIE, 2017. — Jun. — P. 103320G–12.

[5] Lessons and Insights from Creating a Synthetic Optical Flow Benchmark / Jonas Wulff, Daniel J Butler, Garrett B Stanley, Michael J Black // Computer Vision – ECCV 2012. Workshops and Demonstrations. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2012. — Oct. — P. 168–177.

[6] Petro Ana Belén, Sbert Catalina, Morel Jean-Michel. Multiscale retinex // Image Processing On Line. — 2014. — P. 71–88.

[7] Stochastic ray tracing methods in problems of photorealistic image synthesis for augmented reality systems / DD Zhdanov, IS Potemin, AA Kishalov et al. // Proc. 27th International Conference on Computer Graphics and Vision, Perm, PSU. — 2017. — P. 24–28.

[8] Vygolov Oleg, Zheltov Sergey. Enhanced, synthetic and combined vision technologies for civil aviation // Computer Vision in Control Systems-2. — Springer, 2015. — P. 201–230.

[9] A real-time photogrammetric algorithm for sensor and synthetic image fusion with application to aviation combined vision / MA Lebedev, DG Stepaniants, DV Komarov et al. // The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2014. — Vol. 40, no. 3. — P. 171.

Об авторах

Князь Владимир Владимирович, начальник группы ГосНИИАС, магистр ФУПМ МФТИ. Его e-mail vl.kniaz@gosniias.ru.

Данилов Сергей Юрьевич, инженер ГосНИИАС, магистрант ФУПМ МФТИ. Его e-mail danilov@gosniias.ru.

Выголов Олег Вячеславович, к.т.н., начальник лаборатории ГосНИИАС. Его e-mail o.vygolov@gosniias.ru.

Бордодымов Артём Николаевич, инженер ГосНИИАС. Его e-mail bordodymov@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Визуальная аналитика и модели абстракции данных

Д.В. Манаков
manakov@imm.uran.ru
Институт математики и механики им. Н.Н. Красовского Уральского отделения РАН, Екатеринбург, Россия

Поскольку визуализация становится средой автоматизированного аналитического процесса, для визуальной аналитики представляют интерес направления, связанные с самоорганизацией, например, диссипативные систем. С этих позиций можно уточнить понятие структурной единицы визуального анализа, включая модели абстракции данных.

Ключевые слова: *логическое пространство, диссипативные системы, предельная неопределенность.*

Visual analytics and data abstraction models

D.V. Manakov
manakov@imm.uran.ru
Institute of Mathematics and Mechanics im. N.N. Krasovskogo Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russia

Since visualization becomes the environment of an automated analytical process, the directions related to self-organization, for example, dissipative systems, are of interest for visual analytics. From these positions, it is possible to clarify the notion of a structural unit of visual analysis, including data abstraction models.

Keywords: *logical space, dissipative systems, limited uncertainty.*

1. Введение

При рассмотрении компьютерной визуализации как самостоятельной дисциплины, необходимо построение ее ментального пространства со своей семантикой, прагматикой и базисом. Тогда любые два специалиста по визуализации смогут говорить на одном языке. Этот базис выбирается из достаточно широкой междисциплинарной области знаний. Верификацию визуализации в духе нечетких множеств определим через отношение двух базисных функций точности и полноты визуализации. Она должна гарантировать, что конечному пользователю предложена формально правильная модель визуализации, или, другими словами, что разработчики систем визуализации решили поставленную задачу. На современном этапе развития компьютерной визуализации критерий полноты является более важным. Сначала необходимо сформировать ментальное пространство, а затем, уточняя семантику, прагматику и базис, заменяя ментальное пространство логическим пространством, перейти к верификации визуализации. Построение монотонно возрастающих базисных функций, например, точность визуализации: постановка задачи, прототип, приложение, сервис, позволяет рассматривать классификацию как непрерывный процесс. Возможные постановки задач рассматриваются как вызовы и определяют не только перспективные направления развития визуализации, но и их множество продуцирует функцию полноты.

2. К теории компьютерной визуализации и визуальной аналитики

Визуальная аналитика должна предоставить технологии, которые сочетают сильные стороны человеческой и электронной обработки данных. Визуализация становится средой полуавтоматического (автоматизированного) аналитического процесса, где люди и машины сотрудничают для достижения наиболее эффективных результатов [15]. Цель визуальной аналитики - сделать способы обработки данных и информации прозрачными для аналитического дискурса. И так, визуальная аналитика наука об аналитических рассуждениях, упрощаемых интерактивными визуальными интерфейсами. Частный случай визуальной аналитики - бизнес аналитика.

Модели абстракции данных - обобщения, позволяющие абстрагироваться от источника и происхождения (онтологии) данных во время визуального анализа. Стандартный цикл компьютерного моделирования [7], включающий физическую, математическую, алгоритмическую, программную, визуальную модели и модели визуальной аналитики, можно рассматривать как последовательность отображения данных. Очевидно, что любая формализация повышает уровень абстракции данных, в то время как интерпретация (связанная с визуальной аналитикой) нацелена на снижение этого уровня. В тоже время интерпретация невозможна без учета контекста, то есть некоторого ментального представления о формальной модели. Модели абстракции данных ограничены диапазоном от математической модели до визуальной модели данных. Их основные структурные единицы, в том числе и по частоте встречаемости - фильтрация, кластеризация, выборка (выбор) не имеют естественной образности.

Датацентрическая теория [17] ориентирована на шаблоны (типы, структуры) данных, используемые в программирование. Такие структуры данных, как *k*-деревья, *R*-деревья, ассоциативные массивы, стеки (итераторы) не только могут быть связаны с оптимизацией вычислительного алгоритма, но и отвечать за способ распараллеливания. Например, k-деревья (в частности, окта-деревья) применяются для решения задач с различной математической постановкой: параллельные вычисления (алгоритмы с внешней памятью), визуализация (объемный рендеринг с уровнем детализации), дискретная оптимизация (геометрическая аппроксимация транспортной задачи). Таким образом, *k*-деревья являются структурной единицей моделей алгоритмического уровня абстракции данных.

Для того, чтобы компьютерная визуализация и тем более визуальная аналитика считались самостоятельными дисциплинами, необходимо сформировать (например, с точки зрения когнитивной психологии) их ментальное пространство со своей семантикой, прагматикой и базисом. Тогда любые два специалиста по визуализации смогут

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

говорить на одном языке. Этот базис выбирается из достаточно широкой области знаний [3] и должен определять размерность пространства (визуализации). Термин когнитивная размерность, используемый, например, в “usability”, можно считать удачным в качестве дополнительной размерности, например, к декартовым координатам. Когнитивная размерность определяется количеством эвристик или термов, которые изначально считаются ортогональными или независимыми [13]. В тоже время подобные определения базиса очевидно некорректны в математическом смысле.

Альтернативным решением является понятие структурной (семантической) единицы или просто юнита, подходящие для моделей абстракции данных. С точки зрения теории множеств, структурную единицу можно рассматривать как непрерывное отображение класса подмножеств данных на логическое пространство. В рамках датацентрической теории подобное отображение можно реализовать с помощью ассоциативных массивов с множественным ключом и с множественным значением. Обычное логическое пространство [8] определяется как пара $LS = (V, 2^V)$, где фиксированное непустое множество значений истинности - V рассматривается вместе с множеством своих подмножеств - 2^V. Иными словами, множество истинностных значений называется логическим пространством, если на нем выделены определенные подмножества. Поскольку визуализация по своей природе многозначная и нечеткая, для визуальной аналитики представляют интерес истинностные значения k-логики и нечеткой логики.

Построение логического пространства для системы визуализации обеспечивает автоматическую верификацию деятельности пользователя или агента программирования. Распространенными построениями логического пространства являются: отображение симметрической группы, n-линейное и ограниченное отображение, к которому относятся тестирование и алгоритм MapReduce. В случае криволинейных координат (трилинейное ограниченное отображение) значение определителя (Якобиана) формирует логическое пространство (если определитель больше нуля, то отображение является взаимно-однозначным, следовательно, решение существует и единственное, если определитель равен нулю, то решение неединственное, если определитель меньше нуля образуются восемь видов вырожденных ячеек).

Семантические (различимые по смыслу) единицы - непересекающиеся (частично пересекающиеся) визуальные тексты [4] с истинностным значением близким к единице участвуют в формирование базиса. В случае n-линейного и ограниченного отображения размерность пространства равна n. Для n-линейного ограниченного отображения существует взаимно-однозначное отображение на числовую прямую (сюръекция, действие). В этом случае применение теории вероятности или нечетких множеств оправданно (логическое пространство как вероятностное пространство).

Возможные структурные единицы визуализации: визуальные переменные, знаки, образы, тексты, виды отображения, метафоры визуализации и взаимодействия. Например, визуальные переменные используются в визуальной семиотике. Аналогично лингвистическому направлению визуализацию можно рассматривать как представление и как процесс. Выше перечислены структурные единицы представления. Понятие визуальной парадигмы определяется как цель визуализации (целенаправленность, интенциональность). Достижение цели можно рассматривать, как процесс

Цикличность процедуры визуального анализа приводит к появлению понятия об элементарном структурном объекте, повторение которого формирует ход решения. Структурной единицей визуального анализа (визуального процесса) предлагается считать [2] состояние визуальной модели, интерпретация которого предоставляет наблюдателю объем информации, необходимый для последующего получения общего результата проводимого анализа. Структурная единица может быть самостоятельной визуальной моделью, либо входит в состав более сложного объединения, являющегося средством визуального анализа применительно к поставленной задаче.

Структурная единица визуального анализа является визуально воспринимаемым образом (или другой структурной единицей визуального представления), интерпретируемым как истинностный ответ (в первоисточнике [2] однозначный ответ) на один из промежуточных вопросов. Сложность и структура вопроса определяется уровнем абстракции модели с ограничениями по времени, вычислительным ресурсам и на модель конечного пользователя.

В терминах системного анализа, структурная единица визуального анализа представляет собой управляемую систему S с обратной связью $R(t)$. Объем исследуемых данных поступает на неуправляемый вход $V(t)$ структурной единицы рис. 1. Результатом является информация, изменяющая состояние выхода $Y(t)$. Особенностью такой системы является возможность изменения состояния управляемого входа системы $U(t)$ в зависимости от полученных результатов и благодаря наличию обратной связи $R(t)$. Пользователь P является, с точки зрения комплексного подхода к визуальному анализу, обязательным участником структурной единицы. Влияние пользователя проявляется в изменении состояний управляемых входов и выходов, а также в регулировании работы системы S и принятии решения о завершении ее функционирования.

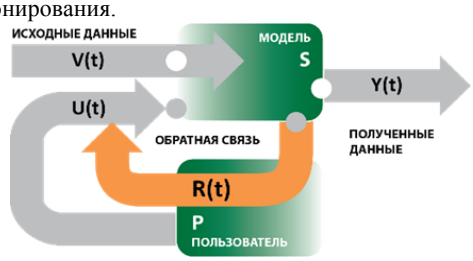


Рис. 1. Структурная единица визуального анализа [2]**.**

Поскольку визуализация становится средой **автоматизированного** аналитического процесса, для визуальной аналитики представляют интерес направления, связанные с **самоорганизацией:** диссипативные системы, автономные вычисления и синергетика. Например, можно уточнить понятие структурной единицы визуального анализа рис. 1 с позиции диссипативных систем

Диссипативная система - квазистационарная открытая система, характерной особенностью которой является процесс самоорганизации, происходящий в результате действия отрицательного вектора, например, силы трения. Поясним отличие диссипативной системы от задачи оптимального управления в случае линейных систем:

$$\dot{y} = Ay - Bu,$$

где u – оптимальное управление, которое надо найти.

Поскольку диссипативная система является открытой системой управление $V(t)$ приходит извне, оно изначально задано. Введение в рассмотрение отрицательного вектора или отрицательной обратной связи $R(t)$ сужает круг задач до моделей с насыщением. Возникновение обратной связи стоит рассматривать не только как результат целенаправленных, но и хаотических, случайных действий пользователя (например, аналогично броуновскому

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

движению), поскольку процесс самоорганизации происходит гораздо быстрее при наличии в системе внешних и внутренних шумов. Управление пользователя *U(t)* желательно, чтобы было оптимальным *u=u(U(t),V(t))*, по крайней мере, оно должно существовать. Например, можно предложить пользователю y_i - монотонно сходящеюся последовательность ращений [3].

В принципе, рассмотрение восприятия-понимания человека [5] может базироваться на частном случае диссипативных систем теории функциональных систем П. К. Анохина. Человек находится в состояние гомеостаза (почти статического равновесия). Выделяют два типа функциональных систем. Системы первого типа обеспечивают гомеостаз за счёт внутренних (уже имеющихся) ресурсов организма, не выходя за его пределы – замкнутые системы. Системы второго типа поддерживают гомеостаз за счёт изменения поведения, которое формируется в результате отрицательной обратной связи – открытые (диссипативные) системы. Возможно слияние компьютерных моделей и моделей восприятия-понимания через ввод понятия интеллектуального агента. Под агентом понимают открытую, активную, целенаправленную систему, которая способна сама формировать собственное поведение в не полностью определенной среде [9]. Интеллектуальный агент – агент способный к самопознанию.

Ключевым понятием диссипативных систем является энтропия. Например, термодинамическая энтропия - мера необратимой диссипации энергии. Поскольку понятия термодинамической энтропии и информационной энтропии (информативности) эквивалентны, применение теории диссипативных систем в области информационных технологий и визуализации обоснованно. Так как данная работа ориентирована на модели абстракции данных ,рассмотрим иллюстративный пример динамики диссипативных систем в области информационных технологий (задача кластеризации). Рис. 2 представляет визуальный анализ деятельности кредитных организаций. Он позволяет строить суждения об образовании кластеров (диссипативных структур) и находить объекты (выделено красным цветом), стремящиеся оказаться в кластере [6] (аналогично, модели с насыщением - движения амеб в сторону вещества с наибольшей концентрацией). С точки зрения диссипативных систем, на следующем шаге кластер должен разрушиться (обычно по превышению некоторого порогового значения), и на его периферии должен сформироваться новый кластер.

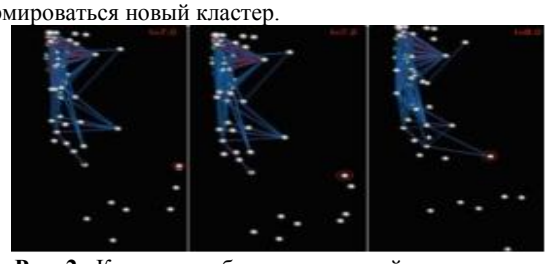


Рис. 2. Кластер и объект, стремящейся оказаться в кластере [6].

Негосударственные интернет-деньги, включая криптовалюту, формируют новый кластер. Перспективным направлением в области параллельных и распределенных вычислений является модель потока данных. Если в области финансов примером модели потока данных является Blockchain (способ последовательно записать такой поток), то в области визуализации Vistrail [10]. Vistrail – классический конструктор модели потока данных, например, параметры модели можно не только менять, но и связывать (графы, таблицы). Он использует графическую библиотеку VTK. При переносе объектно-ориентированных библиотек в область параллельных вычислений существуют определенный ограничения на реализацию непосредственного взаимодействия, в частности, параллельный VTK ориентирован на параллельный рендеринг, поэтому ввод вида отображения multiple-view (рис. 3 - множеству фиксированных параметров ставится в соответствие множество изображений) является вынужденным.

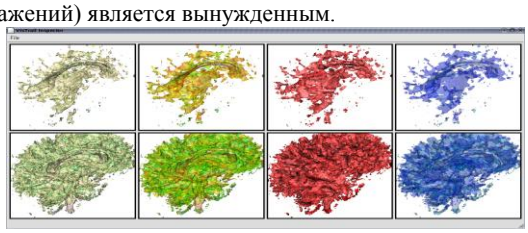


Рис. 3. multiple-view [10]

Формализация модели потока данных может основываться на денотационной семантике Скотта для λ-исчисления. Для построения непрерывного отображения топологический подход, основанный на построении замыкания с такими определяющими свойствами, как частичный порядок и существование супремума, является конструктивным. В качестве примера в области визуализации можно привести семиотическое определение метафоры визуализации, рассматриваемой как непрерывное отображение исходного множества на целевое множество [3] (или любой другой структурной единицы визуализации). В стандартное определение метафоры по Лакоффу добавлено только свойство непрерывности. Хотя, монотонность является следствием частичного порядка и ограниченности, в [3] предлагается рассматривать монотонное отображение с целью расширения класса решаемых задач. Визуализация по своей природе многозначная и нечеткая, можно вести понятие *k*-монотонности и α - монотонности соответственно для *k*-логики и нечеткой логики так, чтобы отображение по-прежнему оставалось непрерывным. Для непрерывной модели разделение визуализации на представление и процесс не существенно. Наравне с топологическим подходом (денотационная семантика) непрерывную модель можно определить через малое изменение параметров модели (операционная семантика). Например, модель процесса исследования визуализации (A Model for the Visualization Exploration Process) [14] основана на исчисление параметров. Если существует визуальная парадигма, то процесс визуализации можно рассматривать, как направленное действие. Для любого направленного действия существует конечный предел.

Непрерывность модели – непрерывность цикла прохождения задания и непрерывность познания, включая визуальный анализ результатов вычислений. Таким образом, для систем визуализации возможна постановка таких стандартных в математическом плане задач:

1. Определение информативных признаков, то есть какими свойствами должна обладать система визуализации, чтобы работа с ней была эффективной, и построение по ним логического пространства.
2. Задача визуального анализа чувствительности решения в зависимости не только от параметров прикладной задачи (модели), но и от параметров параллельной программы, и от параметров визуализации [4].

Анализируя визуальные модели, скорее стоит говорить о ментальной модели, чем о математической модели. Хотя формальное понимание может помочь в разработке новых систем визуализации и в уточнении текущих[14]. В результате сформировалось направление по верификации визуализации, которое должно предложить пользователям формально правильные системы визуализации. Начиная с

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2012 года, проходят Европейские семинары по проблеме воспроизводимости, верификации и валидации в визуализации (EuroRV3).

Можно ввести две базисные функции или меры верификации: полнота и точность верификации. В работе [16] вводится понятие верифицируемой визуализации, которая отслеживает, как распространяется погрешность (неопределенность) на всем этапе вычислительного конвейера, включая визуализацию. Математические направления, используемые для формализации визуализации, можно рассматривать как возможный способ структурирования.

При исследовании данных большого объема и сложной структуры возможностей одной модели недостаточно для принятия решения. Комбинированная (нелинейная) модель - построение из набора моделей, каждая из которых соответствует некоторому фрагменту исходного объема данных, либо отвечает на вопрос анализа лишь частично [2] (ответ как истинностное значение), имеет когнитивное значение, превосходящее сумму значений частных моделей.

После постановки проблемы “больших данных” и задач их визуализации по этим вопросам опубликовано большое количество работ. Есть понимание того, что формируется новая предметная область. Сформировалась структура новой дисциплины, хотя и требующая дальнейшей детализации. Сравнивая структуру обработки “больших данных” (рис. 4) и структурную единицу визуального анализа (рис. 1) стоит подчеркнуть наличие дополнительного элемента – моделирования, и еще одной обратной связи – рефакторинг модели. Визуальный анализ больших данных не только невозможен без некоторого ментального представления о формальной модели, но и должен уточнять эту модель или создавать новую. Рефакторинг модели (программы, базы данных, визуализации) – уточнение параметров модели, в результате обратной связи и постановки задачи минимума или верификации и валидации модели [3]. Переопределение понятия рефакторинга связано с необходимостью рассмотрения программ визуализации, не только как замкнутых, но и как открытых детерминированных систем, например, визуальный агент может иметь адаптивное поведение. Частный случай рефакторинга – калибровка параметров модели. Рефакторинг как метод визуальной аналитики сочетает сильные стороны человеческой (обратная связь) и электронной (критерий оптимальности) обработки данных

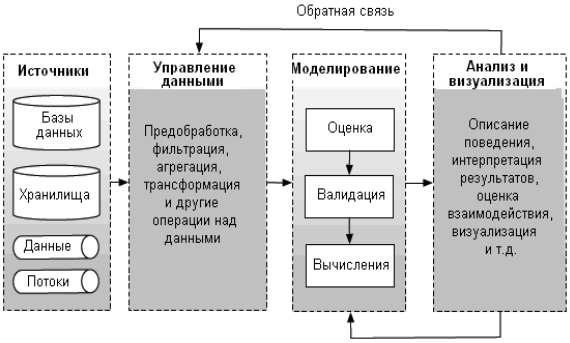


Рис. 4. Структура обработки “больших данных”.

Большие данные - предельный (на данный момент) случай обработки данных, при котором универсальные подходы к анализу и визуализации не работают или неэффективны. Тогда в качестве больших данных могут рассматриваться многомерные и многокатегориальные данные, данные большого объема, данные с неполной информацией (модель с неопределенностью). Предельный случай формирует вызовы, на которые необходимо ответить, чтобы двигаться дальше. Решение возникающих проблем приводит к тому, что сегодняшние “большие данные” завтра становятся нормой [1]. При анализе и визуализации больших данных рассмотрение предельной неопределенности - неопределенности, которая имеет конечный предел в конкретной метризуемой топологии, является вынужденным.

Модели абстракции данных: фильтрация, кластеризация, выборка должны быть ориентированы на обработку больших данных. В данной работе они верифицируются в рамках экономической модели, которая допускает стандартную оценку эффективности. Используемое в экономике определение эффективности (продуктивности) через отношение продукта к источнику продукта (ресурсам) достаточно адекватно (логическое значение близко к единице). Формально, это скорость или полный дифференциал:

$$\mathrm{V} = \frac{\mathrm{df}}{\mathrm{dt}},\ \mathrm{df} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\partial \mathrm{f}}{\partial \mathrm{x_i}} \mathrm{dx_i},$$

где f- продукт, x_i- i источник продукта.

Также возможно рассмотрение данного отношения и с позиций нечетких множеств (нечеткая экономическая модель), так как экономическая эффективность всегда меньше или равна единице, например, вследствие закона сохранения массы. Поскольку значение эффективности должно быть оптимальным ищется решение оптимизационной задачи, например, на нахождение минимума функционала:

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\mathrm{f_i}}{\mathrm{x_i}} \to \min$$

С целью валидации моделей стандартно используют измерение потерь – L. Очевидно, что мера потерь является двойственной мерой эффективности $|V|=1-|L|$, а задачи $V \to max$ и $L \to min$ эквивалентными. В качестве потерь рассматривают относительную эффективность в области параллельных вычислений, например, закон Амдала, а в области визуализации, например, через отношение расстояний, площадей, объемов. В работе [11] рассматривается мера разницы гистограмм (агрегированная дисперсия), частным случаем которой является мера ближайший соседний, используемой для поиска похожих изображений. А также приведены другие примеры отношений и соответствующие ссылки. Стандартно относительную эффективность определяют через отношение двух базисных функций, таких как точность и полнота визуализации, что соотносится с определением верификации визуализации. Аналогично мера разницы гистограмм определяется через “уровень абстракции - соотношение между размерами абстрагированного и исходного набора данных (точность визуализации), и качество абстракции данных (полнота визуализации) - степень, в которой абстрагированный набор данных представляет собой исходный набор данных”.

Относительную эффективность можно рассматривать и как условную вероятность. Если в критерий оптимальности добавить коэффициенты, то возможна калибровка параметров модели. Например, в случае трилинейного вероятностного пространства: $k_1 p_{1,2} + k_2 p_{1,3} + k_3 p_{2,3} = 1$, где $p_{i,j}$ - условные вероятности. Стратегия визуальной калибровки достаточно простая: $k_1 = 1$, k_2 непрерывно изменяется, например, используя слайдер, k_3 вычисляется, например, через дисперсию. Исследователь, выбирая наилучшее визуальное представление, калибрует модель.

Нечеткие модели достаточно распространены (логическое значение близко к единице), в том числе, и в области компьютерной визуализации [12] и семинар EuroRV3 имеет подзаголовок на 2018 год визуализация с неопределенностью, частный случай которых частота встречаемости слов, информативных признаков.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

В работе [12] для верификации визуализации предлагается комбинированная модель с неопределенностью, включающая теорию монотонной меры и расширение понятия треугольного нечеткого числа в рамках аффинной арифметики (первой степенью многочлена, заданной через отношение, что соответствует нечеткой экономической модели нашей классификации). Число диапазон (интервальное число (range number)) - гибридное число, которое явно фиксирует все соответствующие типы неопределенности в одной величине и сопоставляет с каждой величиной степень достоверности.

Мы предлагаем [3] другое расширение и аннотацию нечеткого числа – параметрическое (темпоральное) или локально компактное нечеткое число. Продемонстрируем это понятие на примере параллельных вычислений. Сложность алгоритма задается иногда, как дробь, такая как, например, 2/3, где 3 - сложность последовательного алгоритма, а 2 - сложность передачи данных. Вообще говоря, это могут быть не числа, а функции, зависящие от количества процессоров: ***[2,2]/[3,3](p)=L(p)***.

Темпоральное нечеткое число допускает рассмотрение предельной неопределенности (нечетко-темпоральная модель), например, критерий оптимальности в данном случае будет иметь вид: $\lim_{p\to\infty} L(p) \to min$.

Для диссипативных систем в качестве параметра предпочтительнее рассматривать информационную энтропию, в данном случае, *p/N*, где *N* – количество данных. В рамках визуальной аналитики вышеприведённый критерий оптимальности эквивалентен нахождению минимума таких абстрактных метрик, как когнитивное расстояние, информационный разрыв, инсайт. Как уже отмечалось, относительную эффективность определяют через отношение двух базисных функций, таких как точность и полнота визуализации. Если в качестве параметра определить количество данных, то с помощью нечетко-темпоральной модели можно описать визуализацию с уровнем детализации.

3. Заключение

В секторе компьютерной визуализации ИММ УрО РАН рассматривается возможность разработки он-лайн сервисов параллельных вычислений. На базе конструктора веб-визуализации можно реализовать автономную поддержку стандартных моделей абстракции данных. С целью выделения общих подходов мы разрабатываем нечеткую верифицированную классификацию, которая учитывает как частоту встречаемости моделей, структурных единиц, информативных признаков, так и математический уровень абстракции данных, К структурным единицам визуального процесса относятся визуальная парадигма, анализ чувствительности, рефакторинг, калибровка, предельная неопределенность, веб-визуализация. Построение логического пространства обеспечивает автоматическую верификацию. Мы предлагаем рассматривать структурную единицу как **непрерывное** отображение класса подмножеств данных на **логическое пространство**.

4. Литература

[1] Авербух В.Л., Манаков Д.В**.** Анализ и визуализация “больших данных” // Труды международной научной конференции “Параллельные Вычислительные Технологии” (ПаВТ'2015). Екатеринбург, 31 марта - 2 апреля 2015. Челябинск, Издательский центр ЮУрГУ. 2015. Стр.332-340.

[2] Захарова А.А., Вехтер Е.В., Шкляр А.В. Методика решения задач анализа данных при использовании аналитических визуальных моделей // Научная визуализация 2017. Кв.4. Том 9. N: 4. Стр. 78 - 88.

[3] Манаков Д., Авербух В. Верификация визуализации // Научная визуализация 2016. Кв.1. Том 8. N: 1. Стр. 58 -94.

[4] Манаков Д. В., Авербух В. Л., Васёв П. А. Визуальный текст как истинностное подмножество универсального пространства // Научная визуализация 2016. Кв.4. Том 8. N: 4. Стр. 38 - 49.

[5] Манаков Д.В. Модели восприятия визуальной информации // GraphiCon 2017. 27-я Международная конференция по компьютерной графике и машинному зрению. Труды конференции. ПГНИУ. Пермь. 2017. Стр. 129-132.

[6] Попов Д.Д., Мильман И.Е., Пилюгин В.В., Пасько А.А. Решение задачи анализа многомерных динамических данных методом визуализации // Научная визуализация 2016. Кв.1. Том 8. N: 1. Стр. 55 – 57.

[7] Самарский А.А. Математическое моделирование и вычислительный эксперимент // Вестник АН СССР 1979, N 5. Стр. 38--49.

[8] Тарасов В.Б. Универсальная логика, грануляция информации и искусственный интеллект. http:// www.raai.org/news/pii/ppt/2015/tarasov2015.ppt

[9] Тарасов В.Б., Калуцкая А.П., Святкина М.Н. Гранулярные, нечеткие и лингвистические онтологии для Обеспечения взаимопонимания между когнитивными агентами // Материалы II Междунар. науч.-техн. конф. «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS–2012). Минск: БГУИР. 2012

[10] Bavoil L., Callahan S. P., Crossno P. I., Freire I., Scheidegger C. E., Silva C. T., Vo H. T. VisTrails: Enabling interactive multiple-view visualizations. IEEE Visualization, 2005.

[11] Cui Q., Ward M., Rundensteiner E., Yang J. Measuring data abstraction quality in multiresolution visualizations. IEEE TVCG, 12(5):709–716, 2006.

[12] Fout N., Ma K.-l. Reliable Visualization: Verification of Visualization based on Uncertainty Analysis/ Tech. rep., University of California, Davis, 2012.

[13] Green T.R.G., Petre M. Usability analysis of visual programming environments: a “cognitive dimensions” framework // J. Visual Languages and Computing, 7, 1996. Pp. 131-174.

[14] Jankun-Kelly T., Ma K., Gertz M.. A model for the visualization exploration process. In IEEE Visualization, 2002.

[15] Keim D., Andrienko G., Fekete J-D., Görg C., Kohlhammer J., Melançon G. Visual Analytics: Definition, Process, and Challenges. In: Kerren A., Stasko J.T., Fekete J-D., North C., (Eds). Information Visualization - Human-Centered Issues and Perspectives, Vol. 4950, LNCS State-of-the-Art Survey. Springer: Berlin, 2008. Pp. 154-175.

[16] Kirby R., Silva C. The need for verifiable visualization // IEEE Computer Graphics and Applications, 28(5) Sep 2008. Pp. 78–83.

[17] Purchase H. C., Andrienko N., Jankun-Kelly T.J., Ward M. Theoretical foundations of information visualization. In: Kerren A., Stasko J.T., Fekete J-D., North C., (Eds). Information Visualization - Human-Centered Issues and Perspectives, Vol. 4950, LNCS State-of-the-Art Survey. Springer: Berlin, 2008. Pp. 49–64.

Об авторах

Манаков Дмитрий Валерьянович ведущий программист ИММ УрО РАН. Его e-mail manakov@imm.uran.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Методы визуализации и анализа больших данных в облачных информационных системах

С.В. Мещеряков, Д.А. Щемелинин
serg-phd@mail.ru|dshchmel@gmail.com
Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

Эффективная эксплуатация глобально распределенных вычислительных комплексов с многофункциональными информационными сервисными системами и большими потоками данных невозможна без интеграции процессов непрерывного мониторинга и визуализации текущего состояния всей облачной инфраструктуры по ключевым критериям производительности. В цели и задачи современного мониторинга входит не только визуализация и оповещение обслуживающего персонала и пользователей о временных перерывах работоспособности всей инфраструктуры или ее отдельных частей, но и по возможности автоматическое восстановление работоспособности информационных систем и сервисов. Все эти вопросы рассмотрены в данной статье на конкретном примере большой глобально распределенной интернациональной IT компании.

Ключевые слова: *информационные системы, глобально распределенная инфраструктура, большие данные, интеграция, облачные вычисления, мониторинг, визуализация.*

Methods of big data visualization and analysis in cloud computing information systems

S.V. Mescheryakov, D.A. Shchemelinin
serg-phd@mail.ru|dshchmel@gmail.com
Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

Effective operations of a globally distributed cloud computing infrastructure with multi-service information systems and big data flows is impossible without integration of continuous monitoring processes and visualization of the current status of the entire infrastructure using key performance indicators (KPIs). Modern goals and initiatives of the monitoring system include not only visualization and alerting to technical support and to customers about temporary service outage in either entire or partial production infrastructure but also auto-restoration of information services as soon as possible. In this paper, all mentioned problems are discussed using particular example of big globally distributed International IT Company.

Keywords: *information systems, globally distributed computing infrastructure, big data, integration, cloud computing, monitoring, visualization.*

1. Введение

Исследования проведены в крупной интернациональной IT компании RingCentral (RC), США [5], предоставляющей следующие телекоммуникационные сервисы в разных регионах мира – США и Канаде, Северной и Западной Европе, Юго-Восточной Азии и Океании:

1. Сервисные услуги виртуальной автоматической телефонной станции (АТС) посредством голосового Интернет протокола (VoIP), включая факсимильные и короткие текстовые сообщения (SMS).
2. Облачные программные решения для организации центра обработки звонков, так называемых кол-центров.
3. Видео и телеконференции, включая селекторные совещания между тремя и более участниками.
4. Службы мгновенных сообщений, организации групповых текстовых чатов, звуковых сигналов, файлов, изображений, видео, в том числе на мобильные устройства.

База данных клиентов RC насчитывает более 350000 юридических предприятий, каждое из которых может обслуживать тысячи физических лиц, в разных регионах мира.

На рис. 1 приведена облачная архитектура RC, размещенная в 9 центрах обработки данных на западном и восточном побережье США, в странах Западной Европы и Юго-Восточной Азии. Масштабируемость инфраструктуры RC обеспечивает ежегодный рост на 40% клиентской базы, программного обеспечения и компьютерного оборудования для тестирования. Общий объем физических и виртуальных серверов на сегодняшний день составляет более 20000 вычислительных машин [3]. Все серверы распределены на компоненты и группы, находящиеся под балансировкой нагрузки в зависимости от активности со стороны пользователей и сторонних провайдеров информационных услуг.

2. Система мониторинга ИС

Для мониторинга глобально распределенной облачной инфраструктуры RC использована программная система с открытым исходным кодом Zabbix [6], которая входит в 5 наиболее популярных мониторинговых решений в мире и эффективно используется для контроля системных метрик, таких как доступность сетевых сервисов, использование процессора, наличие памяти, свободное место на внешних дисках и др. Реагирование на возможные неполадки происходит в реальном режиме времени.

Система мониторинга Zabbix построена по стандартной архитектуре «клиент-сервер», состоящей из центральной базы данных MySQL реляционного типа и удаленных серверов (рис. 2). Дополнительные прокси серверы установлены с целью уменьшения нагрузки больших потоков данных со стороны удаленных серверов на базу данных. Интервал времени опроса удаленных серверов составляет для разных системных метрик от 1 до 30 минут, а объем хранимых исторических данных варьируется от месяца до года. Эти параметры заранее определяются конфигурацией Zabbix, которая включает [4]:

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

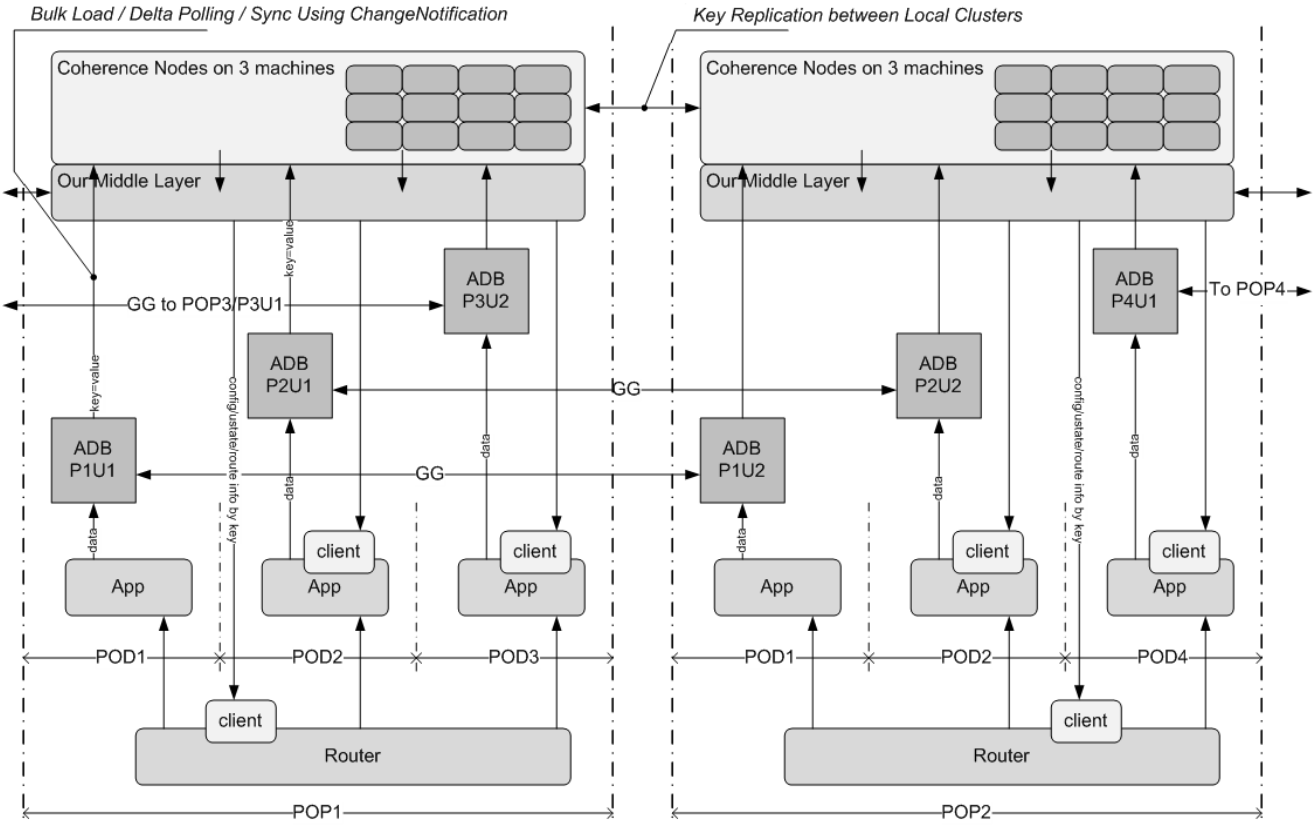


Рис. 1. Масштабируемая архитектура RC.

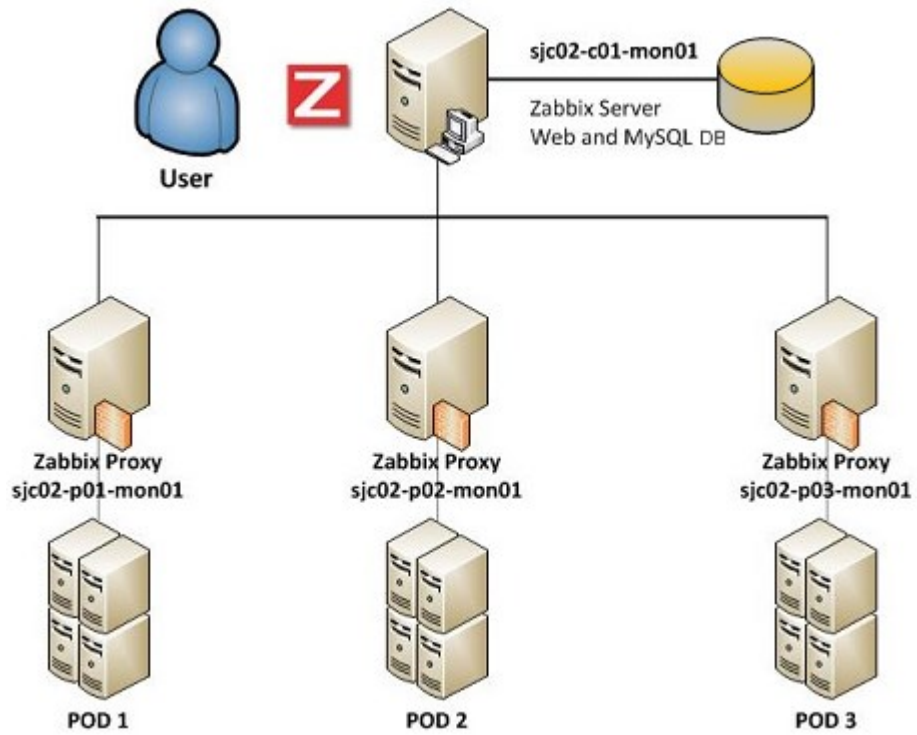


Рис. 2. Архитектура системы мониторинга Zabbix.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

1) список всех серверов в системе мониторинга;
2) описание метрик для мониторинга доступности и производительности серверов, системных служб и приложений, работающих на серверах;
3) триггеры различного приоритета (информационный, предупреждающий, критический, аварийный), которые срабатывают на события, когда текущее значение метрики выходит за пределы заранее предопределенной пороговой величины;
4) графики анализа данных о производительности и исторических тенденций поведения как отдельных метрик серверов, так и ИС в целом.

3. Интерактивный графический интерфейс мониторинга ИС

Система мониторинга Zabbix не обеспечивает реализацию графического интерфейса визуализации состояния всей облачной инфраструктуры с алгоритмом подавления информационного шума и группировки событий, происходящих в RC. С увеличением количества обслуживаемого оборудования возникла необходимость решения следующих задач:

1) уменьшить общее число ложных срабатываний триггеров событий во время отказов;
2) улучшить качество средств мониторинга с точки зрения визуализации корневой причины отказа ИС;
3) сократить время, затраченное на устранение поломок и автоматическое восстановление сервисов.

Для решения поставленных задач реализована новая технология построения среды мониторинга, основанная на интеграции используемых программно-инструментальных средств, – Zabbix API, Attlassian Jira, vCenter, MySQL, Laravel и др., а также собственного программного обеспечения. Разработан интегрированный графический интерфейс (рис. 3), который формируется как сводный отчет в формате HTML, где весь информационный шум (как, например, триггеры C04 на рис. 3) представлен методами [1, 2] в виде иерархического дерева триггеров, сгруппированных под одним корневым событием.

4. Прогнозирование отказов в ИС

Общий подход к мониторингу состоит в сборе данных с удаленных серверов с заданным интервалом времени опроса, сравнении последних значений с указанным пороговым значением и активации триггера в случае превышения. Цель нового решения – создать процесс мониторинга с прогнозированием и заблаговременным предотвращением аномалий, не допуская отказов в сервисном обслуживании. Для этого предложено создать модели прогнозирования отказов, которые оценивают будущие значения метрик и тенденции поведения компонент ИС на основе исторической статистики.

Для прогнозирования отказов в системах с большими данными предложен следующий алгоритм работы:

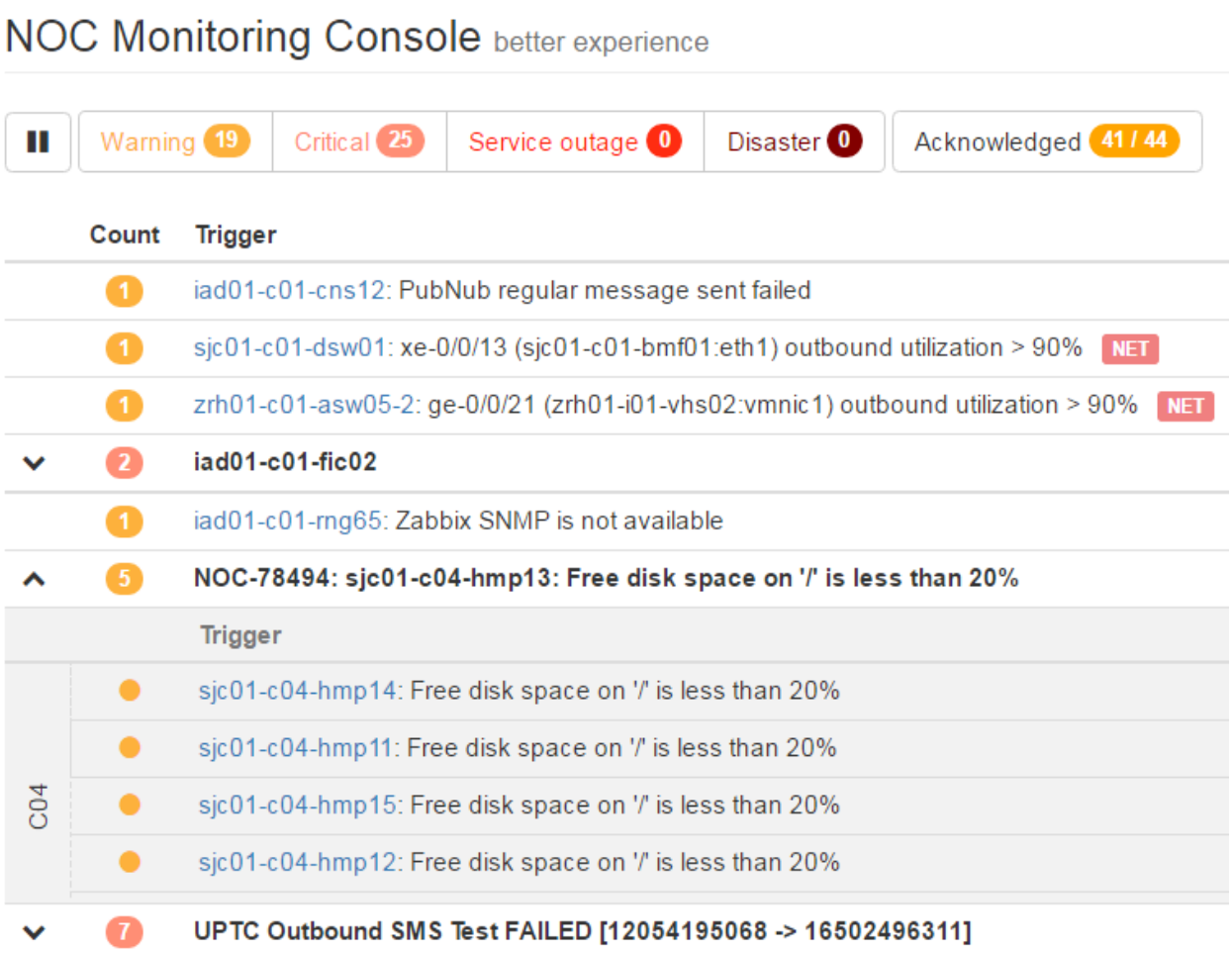


Рис. 3. Интерактивный интерфейс визуализации состояния ИС.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

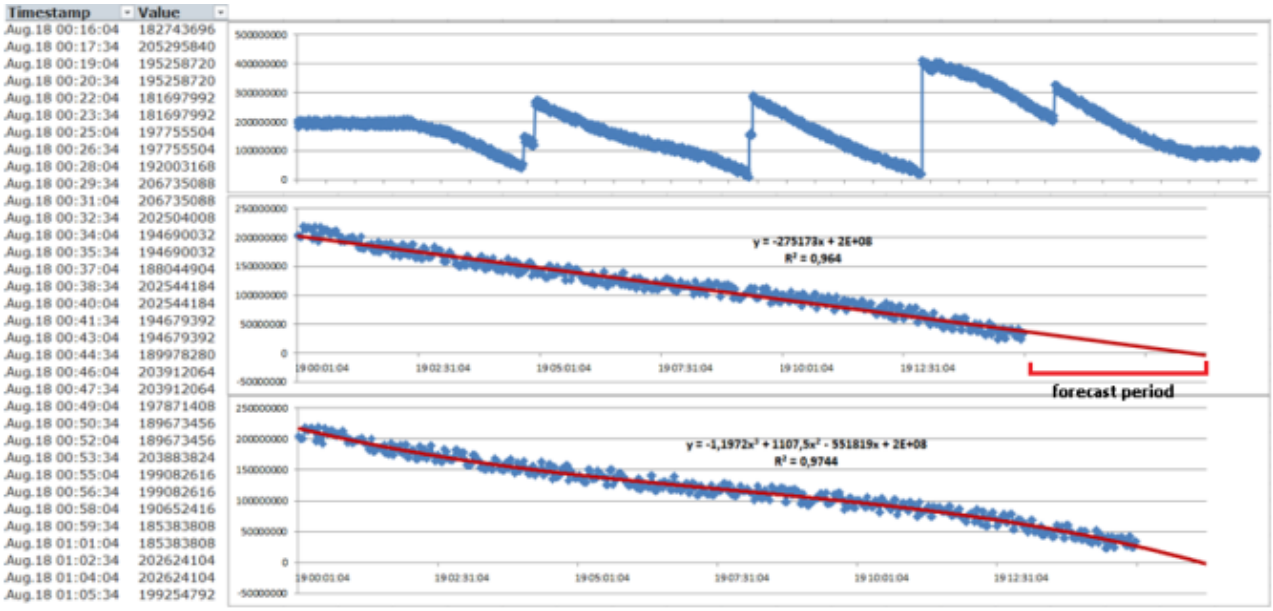


Рис. 4. Пример определения функции прогнозирования поведения элемента ИС.

1) экспортировать статистические значения из базы данных мониторинга в одну из математических систем анализа – Mathematica, R, Mathlab (рис. 4);
2) задать исторический период изменения показателей;
3) настроить тип модели данных и горизонт прогноза;
4) полученное регрессионное выражение использовать в системе мониторинга.

5. Выводы

Практическая реализация предложенной технологии мониторинга в международной IT компании RingCentral позволили повысить доступность облачных сервисов в режиме 24/7 до общепринятого в мире уровня 99,999%.

Улучшение визуализации текущей рабочей обстановки достигнуто за счет разработанного интерактивного графического интерфейса, позволяющего отображать только наиболее значимые события, являющиеся первопричиной аномалий в глобально распределенной системе, а также управлять этими событиями в реальном режиме времени.

Эффективность регламентного обслуживания улучшена за счет практической реализации и внедрения моделей прогнозирования в систему мониторинга. Это позволило получать уведомления о неполадках заблаговременно и выполнять превентивные мероприятия по их устранению в часы плановой профилактики, а не во время пиковой нагрузки со стороны пользователей.

6. Литература

[1] Мещеряков С.В. Сравнительный анализ вариантов организации иерархии в базах данных // Системы управления и информационные технологии. – 2009. – № 1 (35). – С. 34-37.

[2] Мещеряков С.В., Иванов В.М. Реализация модели данных для описания иерархических объектов с произвольными атрибутами // Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление. – 2009. – № 1 (72). – С. 139-144.

[3] Bortyakov D.E., Mescheryakov S.V., Shchemelinin D.A. Integrated Management of Big Data Traffic Systems in Distributed Production Environments // St. Petersburg State Polytechnic University Journal. Computer Science. Telecommunications and Control Systems. – 2014. – № 1 (188). – pp. 105-113.

[4] Hegedus J.C. Perrobix + Zabbix DB Monitoring. Zabbix Annual International Conference, 2013. Available: http://youtu.be/pJCV_ui0orQ. Published: Oct 14, 2013. Last access: 2018.

[5] RingCentral Inc. Official web site. http://www.ringcentral.com/. Last access: 2018.

[6] Zabbix Enterprise-class Monitoring Solution. Official web site. Available: http://www.zabbix.com/product.php. Last access: 2018.

Об авторах

Мещеряков Сергей Владимирович, д.т.н., профессор кафедры «Инженерная графика и дизайн» Санкт-Петербургского политехнического университета Петра Великого. Его e-mail serg-phd@mail.ru.

Щемелинин Дмитрий Александрович, к.т.н., докторант кафедры «Инженерная графика и дизайн» Санкт-Петербургского политехнического университета Петра Великого. Его e-mail dshchmel@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Анализ и визуализация данных энергоаудита методом агрегирования предпочтений

С.В. Муравьев, М.А. Борисова
muravyov@tpu.ru|marits@tpu.ru

Национальный исследовательский Томский политехнический университет, Томск, Россия

Систематическое проведение энергетических обследований электрических сетей с целью выявления потерь электроэнергии – основа энергосбережения и энергоэффективности. В статье предложен основанный на агрегировании предпочтений метод, позволяющий представить большой объем данных инструментальных обследований расхода энергии на собственные нужды подстанциями в форме компактной интегральной оценки в порядковой шкале, удобной для принятия решений и визуализации. Обсуждается применение предложенного метода для анализа реальных данных энергетического обследования магистральных электрических сетей России.

Ключевые слова: *агрегирование, энергосбережение, обработка данных, визуализация, энергоаудит.*

1. Введение

Основной задачей проведения энергетических обследований является расчет потерь электроэнергии на электрических подстанциях. Международные рекомендации в области энергетики требуют, чтобы в процессе передачи электроэнергии в электрических сетях потери не превышали 4 % [9], однако фактические потери магистральных электрических сетей, как правило, составляют не менее 10 %. Значимой составляющей фактических потерь является *расход электрическими подстанциями энергии на собственные нужды* [1].

Традиционно обработка результатов энергетических обследований представляет собой работу с большим объемом неструктурированных данных [3], которые плохо поддаются полному учету. В таких случаях традиционно используются методы многокритериального выбора, такие, например, как метод обобщенного показателя качества или метод анализа иерархий [7,12], которые реализуют сравнение анализируемых объектов на основе взвешенной суммы частных показателей. Эти методы не поддаются серьезной аксиоматизации, и могут приводить к ошибочным решениям [8,10]. Этих недостатков удается избежать при применении агрегирования предпочтений [2], основанного на обработке исключительно бинарных отношений, не требующей вычисления весов, и дающего компактную интегральную оценку объектов в порядковой шкале. Применительно к энергоаудиту, агрегирование предпочтений обеспечивает возможность эффективного сжатия данных без потери информации и их наглядную визуализацию.

2. Агрегирование предпочтений

Пусть задано множество m ранжирований $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_m\}$ n объектов (альтернатив) из множества $A = \{a_1, a_2, ..., a_n\}$. Каждое ранжирование имеет вид цепочки и задает отношение предпочтения $\lambda_k = (a_1 \succ a_2 ... \sim a_s \sim a_t \succ ... \sim a_n)$ на множестве A. Отношение предпочтения λ является объединением двух отношений: отношения строгого предпочтения ρ, т. е. $a_i \succ a_j$, и отношения толерантности τ, т. е. $a_i \sim a_j$, т.е. $\lambda = \rho \cup \tau$. Множество ранжирований Λ будем называть профилем предпочтения для заданных m и n.

Агрегировать m предпочтений, заданных на множестве n альтернатив, означает определить единственное отношение предпочтения β, называемое *ранжированием консенсуса*, которое обеспечивает между ранжированиями исходного профиля наилучший компромисс. Трактовка понятия "наилучший компромисс" определяется используемым при нахождении ранжирования консенсуса *правилом* агрегирования предпочтений. В данной работе будем использовать *правило Кемени* [4], состоящее в нахождении такого линейного порядка (ранжирования Кемени) β альтернатив, что расстояние $D(\beta, \Lambda)$ (определенное в терминах числа парных несоответствий между ранжированиями) от β до ранжирований исходного профиля Λ минимально для всех возможных линейных порядков (перестановок) альтернатив.

Правило Кемени допускает существование неединственного ранжирования консенсуса: число N найденных по этому правилу оптимальных решений может достигать значений более 10^7 даже для небольших $m = 4$ и $n = 15$ [6]. Для свертки множества оптимальных решений $B = \{\beta_1, \beta_2, \ldots, \beta_N\}$ к единственному итоговому ранжированию $\beta_и$ будем использовать следующее условие: если число пар $a_i \succ a_j$ равно числу пар $a_i \prec a_j$ во всех ранжированиях консенсуса, то $a_i \sim a_j$ в итоговом ранжировании $\beta_и$; в противном случае в $\beta_и$ включается та из пар $a_i \succ a_j$ и $a_i \prec a_j$, которая встречается среди оптимальных решений большее число раз.

Для нахождения ранжирования Кемени будем использовать рекурсивный алгоритм собственной разработки RECURSALL, реализующий метод ветвей и границ, позволяющий находить все возможные ранжирования Кемени для заданного исходного профиля предпочтений [5].

При применении правила Кемени следует учитывать, что задача нахождения отношения консенсуса является *NP*-полной, т.е. характеризующейся экспоненциальным ростом времени решения в зависимости от размерности $n = |A|$ задачи [5]. Заметим, что при подходящей для практического применения размерности задачи $n \leq 20$ алгоритм RECURSALL позволяет находить все точные решения за приемлемое время – порядка нескольких миллисекунд. В ситуациях, где $n > 20$, следует прибегать к разбиению множества A на непересекающиеся подмножества A_i, т.е. $A = A_1 \cup A_2 \cup \ldots \cup A_k$, $\bigcap_{i=1}^{k} A_i \neq \varnothing$, где $|A_i| \leq 20$, $i = 1, \ldots, k$.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Операция разбиения множества A приводит к декомпозиции профиля предпочтений Λ, тем самым преобразуя его структуру из линейной в двухуровневую, а при неоднократном повторении этой операции – в иерархическую. Отношения консенсуса β_i, найденные на множестве A_i, войдут в состав профиля вышестоящего уровня иерархии, для которого также может быть найдено отношение консенсуса. Процесс продолжается до достижения наивысшего уровня иерархии.

3. Обработка данных реальных энергетических обследований

Правило Кемени и алгоритм RECURSALL были положены в основу метода анализа данных энергетических обследований магистральных электрических сетей (МЭС) России с помощью агрегирования предпочтений. Единая национальная электрическая сеть России включает в себя восемь МЭС, которые в свою очередь состоят из предприятий магистральных электрических сетей (ПМЭС), каждое из которых объединяет большое количество *подстанций* (ПС).

Исходные данные для метода содержались в обширных таблицах значений *нормативных расходов на собственные нужды* (НРСН) *подстанций* [11]. Данные были получены в ходе проведения энергетического обследования МЭС России и предоставлены ПАО «ФСК ЕЭС». Структура НРПСН содержит 9 основных составляющих: обогрев оборудования (λ_1), освещение зданий (λ_2), освещение территории (λ_3), охлаждение трансформаторов (λ_4), обогрев имеющегося оборудования (λ_5), зарядные и подзарядные устройства (λ_6), аппаратура связи и телемеханики (λ_7), вентиляция и кондиционирование зданий (λ_8) и прочие расходы (λ_9).

Поскольку во многих случаях число подстанций n превышало верхнюю допустимую границу, равную 20 (см. раздел 2), нами была произведено разбиение множеств подстанций на подмножества (кластеры) по признаку их близкого географического расположения. Это было оправдано необходимостью анализа потребления ресурсов подстанциями, функционирующими в подобных климатических условиях.

Метод состоит из следующих основных этапов:

1) формирование множества подстанций $A = \{a_1, a_2, \ldots, a_n\}$, нормативное потребление энергии на собственные нужды которых необходимо проанализировать;
2) построение профиля предпочтений $\Lambda = \{\lambda_1, \ldots, \lambda_m\}$, состоящего из m ранжирований n подстанций по каждой из составляющих (признаку) НРСН; пара подстанций a_i и a_j будет находиться в бинарном отношении $a_i \succ a_j$ или $a_i \prec a_j$, или $a_i \sim a_j$ по признаку λ_k, если соответствующие пары значений НРСН v_i и v_j состоят в бинарном отношении $v_i < v_j$ или $v_i > v_j$, или $v_i = v_j$ по признаку λ_k соответственно;
3) нахождение для профиля Λ ранжирований консенсуса по правилу Кемени $B = \{\beta_1, \ldots, \beta_N\}$;
4) определение с помощью свертки (см. раздел 2) итогового ранжирования консенсуса β_{fin}.

Для того, чтобы продемонстрировать применение предложенного метода на реальных данных, рассмотрим множество подстанций, расположенных возле г. Арзамас: Арзамасская (a_1), Бобылская (a_2), Лукьяновская (a_3) и Луч 500 (a_4). Это множество подстанций относится к ПМЭС «Нижегородское» наряду с другими ареалами, расположенными вокруг населенных пунктов Нижний Новгород, Порецкое и Саранск. Близ городов Арзамас и Саранск расположены по 4 подстанции, близ Нижнего Новгорода – 11 подстанций, около Порецкого – 5 подстанций. Значения v НРСН по всем признакам четырех подстанций в ареале г. Арзамас приведены в таблице 1.

Таблица 1. НРСН четырех подстанций в ареале г. Арзамас, тыс. кВт·ч.

A	Профиль Λ								
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9
a_1	449,3	9,7	95,9	1689,2	777,6	132,8	52,5	3,3	4,2
a_2	351,1	8,2	10,0	137,4	104,5	44,2	106,2	2,7	4,4
a_3	20,5	5,5	6,0	225,2	26,1	44,2	0	14,1	20,9
a_4	915,5	21,2	24,0	1159,8	370,8	132,8	52,5	65,3	329,8

В таблицах 2 и 3 представлены исходные профили предпочтений Λ, сформированные для подстанций расположенных близ г. Арзамас из данных таблицы 1 методом агрегирования предпочтений, в аналитической (символьной) и графической формах, приведенных соответственно в левой и правой частях таблицы, а также рассчитанные для этих профилей алгоритмом RECURSALL ранжирования консенсуса β_{fin} с учетом правила свертки (см. раздел 2).

Аналитическое представление использует вертикальную форму задания ранжирований, в которой более предпочтительный элемент расположен ниже менее предпочтительного, толерантные элементы занимают одну и ту же позицию, например, элементы a_2 и a_3 в ранжировании λ_6. К сожалению, аналитическое представление не обеспечивает достаточного уровня наглядности данных о предпочтениях.

Таблица 2. Аналитическое представление результатов обработки данных энергетического обследования подстанций близ г. Арзамас

Профиль Λ									
λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	$\boldsymbol{\beta_{fin}}$
a_3	a_3	a_3	a_2	a_3	a_2	a_3	a_2	a_1	a_3
a_2	a_2	a_2	a_3	a_2	a_3	a_1	a_1	a_2	a_2
a_1	a_1	a_4	a_4	a_4	a_3	a_4	a_3	a_3	a_1
a_4	a_4	a_1	a_1	a_1	a_1	a_2	a_4	a_4	a_4

Значительно большей наглядности можно достичь графическим представлением профилей предпочтения, если воспользоваться соответствием между определенной интенсивностью цвета и позицией r_i (рангом) элемента a_i в ранжировании λ_j показана соответствующим цветом. Поскольку числа элементов в профилях могут быть различными, необходимо иметь цветовую шкалу для всех вариантов значений n, обусловленных исходными анализируемыми данными. На рис. 1 приведена цветовая шкала, которая была использована для визуализации результатов обработки данных энергетического обследования подстанций близ г. Арзамас (таблица 3).

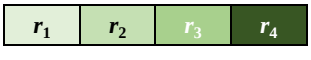


Рис. 1. Цветовая шкала для представления ранжирований (ранг r_i элемента a_i в ранжировании λ_j показан соответствующим цветом)

Результатом применения метода агрегирования предпочтений является выявление двух подстанций a_1 (Арзамасская) и a_4 (Луч 500) в ареале г. Арзамас, для которых необходимо провести мероприятия по снижению НРСН.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Таблица 3. Графическое представление результатов обработки данных энергетического обследования подстанций близ г. Арзамас

Профиль Λ									
λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	β_{fin}
									a_3
									a_2
									a_1
									a_4

4. Заключение

В статье предложен основанный на агрегировании предпочтений метод анализа и визуализации данных энергетических обследований магистральных электрических сетей. Метод позволяет выявить источники экономически нерациональных затрат энергетических ресурсов и неоправданных потерь энергии, обеспечить сжатие больших объемов данных энергообследований без потери существенной информации. Предложенный метод может стать удобным перспективным инструментом для организаций, занимающихся энергоконсалтингом.

5. Благодарности

Работа поддержана Российским научным фондом, проект № 18-19-00203, в части разработки метода агрегирования предпочтений, и Министерством образования и науки РФ, базовая часть госзадания «Наука», проект № 2.5760.2017/8.9 в части анализа данных реальных энергообследований.

6. Литература

[1] Abeysinghe S., Wu J., Sooriyabandara M., Abeysekera M., Xu T., Wang Ch. Topological properties of medium voltage electricity distribution networks // Applied Energy. – 2018. – V. 210. – P. 1101-1112.

[2] Artemenko V., Martyusheva P., Muravyov S.V., Znamenschikova N. Methods of rating competitors for quality awards: tentative comparative analysis // Proceedings of the 10th IMEKO TC7 International Symposium on Advances of Measurement Science (June 30-July 2, 2004, Saint-Petersburg, Russia). – V. 2. – P. 491-496.

[3] Eyupoglu C., Aydi M.A. Energy efficiency in backbone networks // Procedia – Social and Behavioral Sciences. – 2015. – V. 195.– P. 1966–1970.

[4] Kemeny J.G., Snell J.L. Mathematical models in the social sciences. – Cambridge: MIT Press. – 1972. – P. 145.

[5] Muravyov S.V. Ordinal measurement, preference aggregation and interlaboratory comparisons // Measurement. – 2013. – V. 46. – P. 2927-2935.

[6] Muravyov S.V. Dealing with chaotic results of Kemeny ranking determination Measurement. – 2014. – V. 51. – P. 328-334.

[7] Saaty T.L., Vargas L.G. Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks. New York: Springer Science & Business Media. – 2013. – P. 363.

[8] Вентцель Е.С. Исследование операций: задачи, принципы, методология. – М.: КНОРУС, 2013. – 192 с.

[9] Перспективы энергетических технологий Международного энергетического агентства <http://www.iea.org/media/ translations/eer_ru.pdf > (последнее обращение 25.02.18)

[10] Подиновский В.В., Подиновская О.В. О некорректности метода анализа иерархий // Математические проблемы управления, 2011. – 8 - 13 с.

[11] Федеральная Сетевая Компания Единой Энергетической Системы <http://www.fsk-ees.ru/eng> (последнее обращение 05.02.18)

[12] Хвастунов Р.М., Ягелло О.И., Корнеева В.М., Поликарпов М.П. Экспертные оценки в квалиметрии машиностроения. – М.: Технонефтегаз, 2002. – 140 с.

Об авторах

Муравьев Сергей Васильевич, д.т.н., профессор отделения автоматизации и робототехники инженерной школы информационных технологий и робототехники Томского политехнического университета. Его e-mail muravyov@tpu.ru.

Борисова Мария Андреевна, аспирант отделения автоматизации и робототехники инженерной школы информационных технологий и робототехники Томского политехнического университета. Ее e-mail marits@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Критерии когнитивной ясности как основа для построения метафоры визуализации нечетких когнитивных карт

А.Г. Подвесовский[1], Р.А. Исаев[1]
apodv@tu-bryansk.ru|ruslan-isaev-32@yandex.ru
[1]Брянский государственный технический университет, Брянск, Россия

Целью работы является построение метафоры визуализации когнитивных моделей, основанных на нечетких когнитивных картах. Для данного класса задач определены две составляющие метафоры визуализации – пространственная метафора и метафора представления. Рассмотрены критерии когнитивной ясности, которые предложено использовать в качестве основы для построения метафоры визуализации. Приведены примеры использования метафоры визуализации и интерактивного управления визуальным образом нечеткой когнитивной карты в системе программной поддержки когнитивного моделирования «ИГЛА».

Ключевые слова: *нечеткая когнитивная карта, визуализация графов, когнитивная ясность, метафора визуализации.*

Cognitive clarity criteria as a basis for constructing a visualization metaphor for fuzzy cognitive maps

A.G. Podvesovskii[1], R.A. Isaev[1]
apodv@tu-bryansk.ru|ruslan-isaev-32@yandex.ru
[1]Bryansk State Technical University, Bryansk, Russia

The research is aimed to construct a visualization metaphor for cognitive models based on fuzzy cognitive maps. Two basic components of a visualization metaphor, namely spatial metaphor and representation metaphor, are defined for that kind of tasks. The cognitive clarity criteria, which are proposed to use as a basic for constructing a visualization metaphor, are described. Some examples of using a visualization metaphor and interactive control of visual appearance of fuzzy cognitive map in the cognitive modeling support system IGLA are given.

Keywords: *fuzzy cognitive map, graph visualization, cognitive clarity, visualization metaphor.*

1. Введение

Одним из подходов к исследованию слабоструктурированных систем, широко применяемым в настоящее время, является когнитивный подход. В соответствии с определением, приведенным в [2], данный подход ориентирован на разработку формальных моделей и методов, поддерживающих интеллектуальный процесс решения проблем благодаря учету в этих моделях и методах когнитивных возможностей человека (восприятие, представление, познание, понимание, объяснение) при решении управленческих задач. В общем виде под когнитивным моделированием понимается исследование структуры системы и процессов ее функционирования и развития путем анализа ее когнитивной модели.

Процесс когнитивного моделирования начинается с построения когнитивной карты исследуемой системы на основе информации, получаемой от экспертов или путем анализа имеющихся данных о системе. На следующем этапе происходит непосредственно моделирование, основными целями которого являются формирование и проверка гипотез о структуре исследуемой системы, позволяющих объяснить ее поведение, а также выработка стратегий поведения в той или иной ситуации с целью достижения заданных целевых состояний.

В качестве математического аппарата, применяемого для представления и анализа когнитивных моделей, чаще всего используется нечеткая логика. Существует целый класс когнитивных моделей, основанный на различных типах нечетких когнитивных карт (НКК) – достаточно подробный обзор таких моделей можно найти, например, в монографии [3]. Одной из разновидностей НКК, хорошо зарекомендовавшей себя в практических задачах анализа и моделирования слабоструктурированных социальных, организационных, и экономических систем являются нечеткие когнитивные карты Силова [10, 13]. НКК Силова отражает субъективное представление исследователя о ней (индивидуальное или коллективное) в виде множества семантических категорий (называемых факторами или концептами) и набора причинно-следственных связей между ними. Таким образом, данный тип НКК может быть наглядно представлен в виде взвешенного ориентированного графа, вершины которого соответствуют концептам, а дуги – причинно-следственным связям.

Одним из условий эффективной работы с когнитивной моделью является обеспечение ее визуального представления. В работе предлагается метафора визуализации нечетких когнитивных карт, в основе которой лежат алгоритмы визуализации графов и понятие когнитивной ясности, для формализации которого вводится набор критериев. Рассматриваются примеры применения различных метафор визуализации и возможности интерактивного управления визуальным образом нечеткой когнитивной карты в разработанной с участием авторов системе поддержки принятия решений «ИГЛА». Более подробное описание системы «ИГЛА» можно найти в [9], а демо-версию – по адресу http://iipo.tu-bryansk.ru/quill. Система представляет собой Windows-приложение на платформе Microsoft .NET Framework, и в настоящее время ведется разработка сетевой многопользовательской версии.

2. Задача визуализации нечетких когнитивных карт

Визуальный анализ является неотъемлемой составляющей процесса когнитивного моделирования, как на этапе построения когнитивной карты, так и на этапах ее верификации и исследования. Важность обеспечения наглядного и удобного для восприятия визуального представления когнитивной карты обусловлено следующими обстоятельствами.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

1) Визуальное представление когнитивной карты предоставляет исследователю (эксперту, аналитику) зачастую единственную возможность «охватить модель одним взглядом», в результате чего тот способен быстро заметить ошибки (например, пробелы или избыточность), допущенные при ее построении.
2) Наличие визуального представления является необходимым условием для обеспечения возможности отображения результатов верификации когнитивной карты (в простейшем случае верификация представляет собой поиск избыточных транзитивных путей передачи влияния между концептами и замкнутых бесконечно усиливающих или ослабляющих циклов влияния).
3) Визуальное представление может быть использовано для отображения результатов структурно-целевого и сценарного анализа когнитивной модели.

Тем не менее, в публикациях по когнитивному моделированию проблеме визуализации когнитивных карт уделяется достаточно мало внимания. Отдельные аспекты данной проблемы рассматривались в работах [1, 11]. В опубликованной авторами ранее работе [6] был предложен подход, подразумевающий сведение задачи визуализации НКК Силова к задаче визуализации графов. Исследовался ряд алгоритмов визуализации графов. В соответствии с идеей, лежащей в основе построения и работы этих алгоритмов, среди них можно выделить алгоритмы, основанные на физических аналогиях (силовые алгоритмы), алгоритмы, основанные на принципах самоорганизации, и алгоритмы поуровневого изображения ориентированных графов.

По результатам экспериментальной проверки рассмотренных алгоритмов на реальных НКК, наиболее подходящими для реализации в системе «ИГЛА» были признаны алгоритм LinLog [15], относящийся к классу силовых алгоритмов, и алгоритм ISOM [14], основанный на принципах самоорганизации. Дальнейшим развитием данного подхода является его расширение до метафоры визуализации, с учетом требований обеспечения когнитивной ясности визуального представления.

3. Построение метафоры визуализации

В работе [5] метафора визуализации определяется как отображение визуализируемого объекта из пространства данных исходной задачи в объект пространства представления, происходящее посредством условного переноса признаков элементов одного множества на элементы другого множества.

При этом, согласно [5], метафора визуализации может использоваться на двух стадиях решения задач визуализации. Первый раз – при переходе от исходных данных к пространству визуальной модели. Для этого случая используется термин «пространственная метафора». Второй раз необходимость в применении этого приема возникает на стадии уточнения результатов визуализации, усиления необходимых для решения составляющих, фильтрации избыточных компонент созданного образа и т.д. Для данного случая используется термин «метафора представления». Фактическая роль пространственной метафоры состоит в преобразовании данных решаемой задачи в данные, описывающие элементы пространства визуальной модели. Метафора представления используется как средство, раскрывающее потенциальные преимущества уже примененной пространственной метафоры.

Рассмотрим предлагаемый способ построения двух составляющих метафоры визуализации НКК – пространственной метафоры и метафоры представления.

Поскольку пространственная метафора подразумевает переход к пространству визуальной модели, необходимо определить в первую очередь вид этого пространства – в частности, его размерность.

В работе [11] приводится модель «когнитивное облако» (фактически это пространственная метафора), подразумевающая расположение когнитивной карты в трехмерном пространстве. Выдвигается гипотеза о том, что такое визуальное представление в большей степени (по сравнению с расположением карты на плоскости) будет способствовать пониманию когнитивной модели при ее чтении. В подтверждение выдвинутой гипотезы приводится пример применения предложенного подхода.

Вместе с тем, метафора, основанная на модели «когнитивное облако», не является универсальной, поскольку она ориентирована на визуализацию когнитивных карт специфичной структуры (а именно, с наличием ярко выраженных «факторов-ядер»), в то время как потенциальная эффективность применения данной метафоры для визуализации карт произвольной структуры вызывает сомнения. Помимо этого, визуализация в трехмерном пространстве, очевидно, является более ресурсоемким процессом, чем в двумерном, что может негативно сказаться на скорости отрисовки модели и времени ее отклика на действия пользователя, особенно при большом количестве концептов и причинно-следственных связей. Таким образом представляется целесообразной разработка такой метафоры в двумерном пространстве (на плоскости), которая, с одной стороны, свободна от недостатков и усложнений трехмерной метафоры, а с другой, является более универсальной и подходящей для НКК произвольной структуры.

Как упоминалось ранее, задача визуализации НКК в целом сводится к задаче визуализации графа, которая, в двумерном случае, может быть решена при помощи обширного класса алгоритмов укладки графов. Таким образом, в основе пространственной метафоры визуализации НКК должны лежать указанные алгоритмы. Однако при этом также должны быть учтены проблемы ограниченности когнитивных возможностей человека при чтении графов (подробный анализ данной проблемы можно найти, например, в [12]). Предложенный в упомянутой выше работе [6] подход позволяет получить удовлетворительные результаты визуализации, однако при этом не учитывается важный аспект качества получаемого изображения, а именно, достижение его когнитивной ясности. Понятие когнитивной ясности и связанные с ним критерии будут рассмотрены далее.

В результате применения к исходным данным (структуре когнитивного графа) пространственной метафоры будет получено расположение (т.е. координаты на плоскости) его вершин и дуг, являющееся оптимальным в некотором отношении. Разработка и описание конкретных критериев оптимального расположения выходит за рамки настоящей работы и является направлением дальнейших исследований.

Метафора представления, как уже отмечалось, призвана «раскрыть преимущества» используемой пространственной метафоры, в результате чего совершается переход от визуальной модели исследуемого объекта к его визуальному образу. Исходя из многоэтапности процесса когнитивного моделирования (включающего в себя как построение модели и ее верификацию, так и различные виды ее анализа), возникает потребность в разработке нескольких различных метафор представления. При этом каждая из этих метафор должна соответствовать определенному этапу моделирования и способствовать достижению исследователем поставленной на данном этапе цели. Таким образом, метафоры представления должны строиться, исходя из акцентов восприятия исследователя на том или ином этапе когнитивного моделирования.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Перечислим с примерами основные метафоры представления НКК, реализованные в системе «ИГЛА». Для примеров будем использовать разработанную при непосредственном участии авторов когнитивную модель поддержки принятия решений в области автоматизации проектирования технологической оснастки [8].

1) Метафора, используемая «по умолчанию» (рис. 1), позволяющая исследователю акцентировать внимание на структуре когнитивной карты в целом, не отвлекаясь на особенности отдельных ее концептов. В рамках данной метафоры для представления всех концептов используется единый цвет (серо-голубой). Цвет дуги задает знак соответствующей причинно-следственной связи между концептами: красный цвет означает положительную связь («усиление»), синий цвет – отрицательную («торможение»). Толщина дуги определяет интенсивность связи (более подробное описание структуры и параметров НКК Силова можно найти, например, в работах [9, 10, 13]).
2) Метафоры, подразумевающие визуальное разделение концептов в зависимости от их типов либо их принадлежности к тем или иным смысловым группам, которые задаются самими экспертами при построении когнитивной модели. На рис. 2 показан пример визуального разделения с целью акцентировать внимание исследователя на взаимном расположении управляемых (желтый цвет), неуправляемых (зеленый цвет) и целевых (серо-голубой цвет) концептов.

Рис. 1. Основная метафора

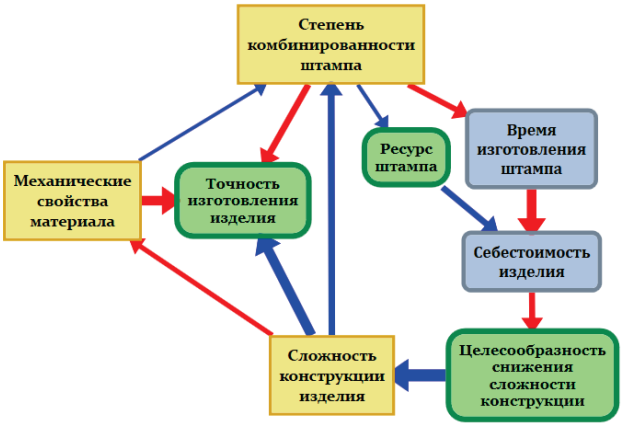


Рис. 2. Метафора разделения концептов по типам

3) Группа метафор, позволяющих визуализировать в наглядной форме системные показатели НКК, рассчитанные в результате ее структурно-целевого анализа (подробное описание системных показателей НКК можно найти, например, в работах [8-10]). В примере на рис. 3 объектами визуализации являются сравнительная интенсивность и направление влияния всех факторов на систему. По аналогии с цветами дуг, для представления концептов, положительно влияющих на систему, используются оттенки красного цвета, а отрицательно влияющих – оттенки синего цвета. Насыщенность цвета определяет интенсивность влияния (при этом белый цвет означает, что данный концепт не оказывает существенного влияния на систему, что характерно для целевых концептов, являющихся «стоковыми»).
4) Метафора «альфа-срез», позволяющая выделить классы взаимосвязанных концептов, характеризуемых заданным уровнем среза относительно выбранного системного показателя (например, взаимного отрицательного влияния концептов – рис. 4). Отметим, что для альфа-среза не учитываются направленность знак и интенсивность причинно-следственной связи между концептами, поэтому в данном случае для представления всех связей используются ненаправленные линии единого цвета и толщины.

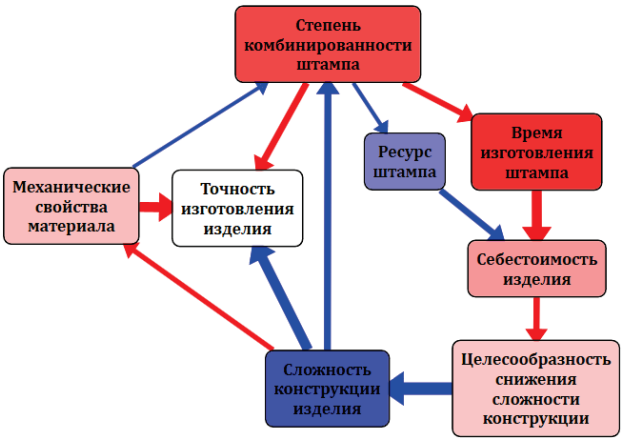


Рис. 3. Метафора влияния концептов на систему

Рис. 4. Метафора альфа-среза по взаимовлиянию

4. Критерии когнитивной ясности

Понятие когнитивной ясности некоторой информации в основном характеризуется легкостью интуитивного понимания соответствующих описаний, сообщений и т.п. Отсутствие когнитивной ясности проявляется, в том, что человек задумывается, чувствует затруднения, пытаясь понять сказанное или написанное. Это может выражаться в наблюдаемом замедлении процесса понимания. Другим последствием может быть упущение значимой информации, которая остается незамеченной [1].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

С другой стороны, в работах, затрагивающих тематику визуализации графов, зачастую приводятся так называемые «эстетические критерии», связанные с повышением наглядности получаемого изображения графа. Достаточно подробный обзор таких критериев приведен в работе [7]. Поскольку НКК, как упоминалось выше, является взвешенным ориентированным графом, то вполне логично, что эстетические критерии будут вносить существенный вклад в повышение наглядности ее изображения.

При этом можно заметить, что многие эстетические критерии неявно направлены на повышение когнитивной ясности изображения графа, и, напротив, изображение, в полной мере обладающее свойствами когнитивной ясности, вероятнее всего, будет привлекательным в эстетическом отношении. Таким образом, эстетические критерии будет вполне целесообразным также относить к критериям когнитивной ясности, и в дальнейшем можно говорить только об этой группе критериев.

С учетом сказанного, в качестве основы для построения метафоры визуализации НКК предлагаются следующие критерии когнитивной ясности:

1) направленность дуг: более удобными (здесь и далее – при прочих равных условиях) для «прочтения» НКК являются направления дуг «сверху-вниз» и «слева-направо» (именно эти направления совпадают с привычными для многих направлениями чтения);
2) однонаправленность изображения последовательных дуг: если не требуется постоянно менять направление взгляда, то зрительный обхват путей и циклов графа будет совершаться быстрее;
3) минимизация пересечений дуг: в идеале они должны отсутствовать, а при невозможности этого (для непланарного графа) их число следует минимизировать;
4) минимизация количества искривленных дуг: более удобными для восприятия являются изображения с прямыми дугами;
5) минимизация длины дуг (как совокупной длины, так и наибольшей из длин): чем короче дуги, тем проще увидеть, какие концепты связаны друг с другом, и тем больше связей можно увидеть одновременно;
6) минимизация разброса длины дуг: более удобными для восприятия являются изображения, в которых все дуги имеют примерно одинаковую длину;
7) максимизация углов между дугами, инцидентными одной вершине: при небольших углах между такими дугами они будут «сливаться» друг с другом вблизи вершин, что может существенно затруднить визуальное определение их направлений;
8) оптимизация области размещения: для эффективного использования пространства граф должен быть размещен в прямоугольной области, формат которой (т.е. соотношение ширины и высоты) соответствует текущему формату графической области подсистемы визуализации;
9) подчеркивание симметрии графа: изображения, симметричные относительно некоторой оси или центра, более удобны для восприятия и анализа.

Возвращаясь к описанной ранее метафоре визуализации НКК, можно теперь сказать, что вторым по важности ее компонентом (после алгоритмов визуализации графов) должны стать введенные критерии когнитивной ясности. Анализируя эти критерии, можно прийти к выводу, что многие из них противоречат друг другу, и обеспечить соответствие изображения одновременно всем критериям с алгоритмической точки зрения, как правило, невозможно. Таким образом, независимо от конкретных особенностей реализации метафоры визуализации НКК, необходима разработка решающих правил, моделирующих различные формы компромисса между критериями.

5. Интерактивное управление визуальным представлением нечеткой когнитивной карты

В работе [4] отмечается, что наличие интерактивного управления визуальным образом обеспечивает прямое участие пользователя в манипулировании образом и является основанием для глубокого анализа данных. Кроме того, система интерактивного управления моделью является одним из способов верификации полученных решений и потому обеспечивает ускоренное достижение цели анализа в случае, когда этот способ удобен для пользователя. Таким образом, интерактивность визуальной модели становится условием ее высокой результативности.

Механизм интерактивного управления визуальным представлением НКК когнитивной моделью в системе «ИГЛА» обеспечивает следующие возможности:

1) редактирование структуры когнитивной модели посредством редактирования ее визуального образа;
2) перестройка пространственной метафоры с учетом выбранного алгоритма визуализации графов;
3) трансформация изображения когнитивной модели;
4) переключение между метафорами представления, а также настройки отдельных метафор.

Редактирование структуры модели подразумевает возможность добавления, удаления и изменения концептов и связей путем манипулирования элементами графического интерфейса и визуального образа самой модели.

Трансформация изображения не затрагивает ни структуру модели, ни используемую в текущий момент метафору представления, однако позволяет гибко изменять свойства конечного изображения, выводимого на экран (таким образом, выполняя своего рода постобработку). К возможностям трансформации относятся: плавное масштабирование изображения; поворот изображения на заданный угол по часовой стрелке или против нее; зеркальное отражение изображения относительно горизонтальной или вертикальной оси; сжатие и растяжение изображения вдоль выбранных направлений.

Наконец, благодаря возможности переключаться между метафорами представления исследователь в любой момент может направить внимание на те аспекты модели, которые представляют для него наибольший интерес на текущем этапе анализа. При этом функции настройки визуальных характеристик метафор представления также способствуют повышению гибкости этого инструмента и его удобства для исследователя (рис. 5).

6. Заключение

В работе представлена метафора визуализации нечетких когнитивных карт. Определены две составляющие метафоры визуализации – пространственная метафора и метафора представления. Рассмотрено понятие когнитивной ясности, в результате анализа которого сделан вывод о наличии связи между качеством применяемой метафоры визуализации НКК и уровнем когнитивной ясности полученного визуального образа: чем более высокий уровень когнитивной ясности обеспечивает метафора визуализации, тем более простым является процесс экспертного понимания когнитивной модели при ее визуальном анализе. Для оценки уровня когнитивной ясности предложен набор критериев.

Таким образом, наряду с алгоритмами визуализации графов, показатели когнитивной ясности составляют основу для построения метафоры визуализации НКК и одновременно с этим являются средством наиболее естественной оценки качества построенной метафоры.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

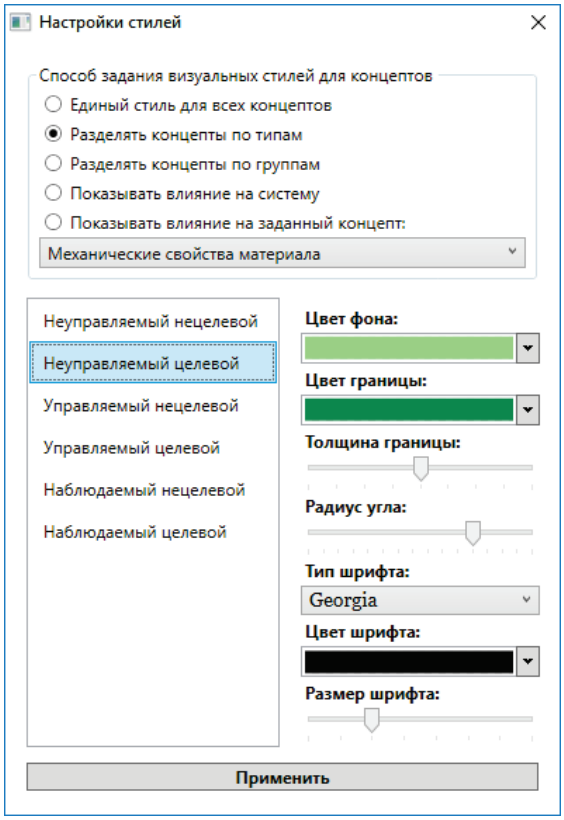


Рис. 5. Пользовательский интерфейс выбора и настройки метафор представления в системе «ИГЛА»

Также рассмотрены возможности интерактивного управления визуальным представлением НКК в системе «ИГЛА» обеспечивающие гибкую настройку визуальных характеристик метафор представления, что позволяет исследователю в ходе анализа когнитивной модели фокусировать внимание на тех ее аспектах, которые представляют для него наибольший интерес на том или ином этапе анализа.

Возможности использования метафор визуализации и интерактивного управления визуальным представлением НКК в системе «ИГЛА» проиллюстрированы на примере когнитивной модели поддержки принятия решений в области проектирования технологической оснастки [8].

Укажем направления дальнейших исследований.

Во-первых, это формализация описанных выше критериев когнитивной ясности и разработка метода оценки качества метафоры визуализации НКК на основе формализованных критериев, с реализацией этого метода в подсистеме визуализации системы «ИГЛА».

Во-вторых, совершенствование механизма интерактивного управления визуальным образом НКК в системе «ИГЛА», в частности, реализация функции автоматического подбора оптимальной пространственной метафоры с учетом указанных пользователем приоритетов по критериям когнитивной ясности.

7. Литература

[1] Абрамова Н.А., Воронина Т.А., Порцев Р.Ю. О методах поддержки построения и верификации когнитивных карт с применением идей когнитивной графики // Управление большими системами. Специальный выпуск 30.1 «Сетевые модели в управлении». – М.: ИПУ РАН, 2010. – С. 411-430.

[2] Авдеева З.К., Коврига С.В., Макаренко Д.И. Когнитивное моделирование для решения задач управления слабоструктурированными системами (ситуациями) // Управление большими системами. – 2007. – Вып. 16. – С. 26-39.

[3] Борисов В.В., Круглов В.В., Федулов А.С. Нечеткие модели и сети. – М.: Горячая линия – Телеком, 2012. – 284 с.

[4] Захарова А.А., Вехтер Е.В., Шкляр А.В. Методика решения задач анализа данных при использовании аналитических визуальных моделей // Научная визуализация. – 2017. – Т. 9. – № 4. – С. 78-88.

[5] Захарова А.А., Шкляр А.В. Метафоры визуализации // Научная визуализация. – 2013. – Т. 5. – № 2. – С. 16-24.

[6] Исаев Р.А., Подвесовский А.Г. Совершенствование механизма визуализации нечетких когнитивных моделей в системе поддержки принятия решений «ИГЛА» // Современные технологии в науке и образовании – СТНО-2017: сб. тр. междунар. науч.-техн. и науч.-метод. конф.: в 9 т. Т.1. – Рязань: Рязан. гос. радиотехн. ун-т, 2017. – С. 138-142.

[7] Касьянов В.Н., Евстигнеев В.А. Графы в программировании: обработка, визуализация и применение. – СПб.: БХВ-Петербург, 2003. – 1104 с.

[8] Копелиович Д.И., Подвесовский А.Г., Сафонов А.Л., Вилюха А.В., Исаев Р.А. Применение нечетких когнитивных моделей в автоматизации проектирования технологической оснастки. // Вестник компьютерных и информационных технологий. – 2018. – № 3. – С. 20-35.

[9] Подвесовский А.Г., Лагерев Д.Г., Коростелев Д.А. Применение нечетких когнитивных моделей для формирования множества альтернатив в задачах принятия решений // Вестник Брянского государственного технического университета. – 2009. – № 4 (24). – С. 77-84.

[10] Силов, В.Б. Принятие стратегических решений в нечеткой обстановке / В.Б. Силов – М.: ИНПРО-РЕС, 1995. – 228 с.

[11] Феррейра Опасо Е.В., Терелянский П.В. Представление когнитивных карт в трехмерном пространстве // XII Всероссийское совещание по проблемам управления ВСПУ-2014: Труды совещания. М.: ИПУ РАН, 2014. – С. 6149-6154.

[12] Huang W., Hong S.H., Eades P. Predicting Graph Reading Performance: A Cognitive Approach // Proc. Asia Pacific Symposium on Information Visualization (APVIS2006). – Tokyo, Japan, 2006. – P. 207–216.

[13] Isaev R.A., Podvesovskii A.G. Generalized Model of Pulse Process for Dynamic Analysis of Sylov's Fuzzy Cognitive Maps // CEUR Workshop Proceedings of the Mathematical Modeling Session at the International Conference Information Technology and Nanotechnology (MM-ITNT 2017), Vol. 1904. – P. 57-63.

[14] Meyer B. Self-Organizing Graphs – A Neural Network Perspective of Graph Layout // Whitesides S.H. (eds) Graph Drawing. GD 1998. Lecture Notes in Computer Science, vol 1547. Springer, Berlin, Heidelberg.

[15] Noack A. An energy model for visual graph clustering // Proceedings of the 11th International Symposium on Graph Drawing. Springer-Verlag, 2004. – P. 425-436.

Об авторах

Подвесовский Александр Георгиевич, к.т.н., доцент, заведующий кафедрой «Информатика и программное обеспечение» Брянского государственного технического университета. Его e-mail apodv@tu-bryansk.ru.

Исаев Руслан Александрович, аспирант Брянского государственного технического университета. Его e-mail ruslan-isaev-32@yandex.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Визуализация траекторий пузырьков газа в акустическом поле

И.Л.Расковская [1], Б.С. Ринкевичюс[2], А.В. Толкачев[3], А.В. Тихомирова[4]
raskovskaIL@mail.ru | rinkevbs@mail.ru |avtolk2008@yandex.ru | a.v.tikhomirova77@gmail.com
[1]ФГБОУ НИУ «МЭИ», Москва, Россия;
[2] ФГБОУ НИУ «МЭИ», Москва, Россия;
[3] ФГБОУ НИУ «МЭИ», Москва, Россия;
[4] ФГБОУ НИУ «МЭИ», Москва, Россия;

В данной работе проводились исследования сложного поведения газового кавитационного пузырька в жидкости. Исследование проводилось оптическим методом, путём видеосъемки подсвеченного лазерной плоскостью потока пузырьков.

Дано описание экспериментальной установки, получены видеокадры сложного поведения пузырька в акустическом поле. Обработка полученных изображений осуществляется с помощью разработанной программы слежения за одиночными пузырьками.

Ключевые слова: *акустическое поле, исследование потоков кавитационных пузырьков, алгоритм слежения, определение траектории.*

Visualization of the trajectory of gas bubbles in acoustic field

I.L. Raskovkaya[1], B.S. Rinkevichyus[2], A.V. Tolkachev[3], A.V. Tikhomirova[4]
raskovskaIL@mail.ru | rinkevbs@mail.ru |avtolk2008@yandex.ru | a.v.tikhomirova77@gmail.com
[1]National Research University “MPEI”, Moscow, Russia;
[2] National Research University “MPEI”, Moscow, Russia;
[3] National Research University “MPEI”, Moscow, Russia;
[4] National Research University “MPEI”, Moscow, Russia;

The paper is dedicated to the study of complex behavior of a gas cavitation bubble in liquid. The study was carried out with the usage of optical method through videotaping of particle flux illuminated by laser plane and being affected by acoustic field with a frequency from 120 to 150 kHz.

A description of an experimental setup is given, video frames of the complex bubble behavior in acoustic field are obtained. Processing of the received photos was carried out by means of the created program of tracking of single objects.

Keywords: *acoustic field, flow study, track algorithm, determination of trajectory.*

1. Введение

Исследование траектории движения газового пузырька, возникающего вследствие явления акустической кавитации (газообразования при локальном понижении давления в жидкости), имеет значение как для решения теоретических, так и практических задач [1].

С проблемами кавитации сталкиваются при рассмотрении широкого круга вопросов, связанных с течениями жидкостей, от исследования тока крови в сосудах до проектирования турбин и корабельных винтов [2]. Её возникновение зависит от физических свойств жидкости и параметров течения. В технике кавитация приводит к значительному снижению эффективности машин и к их эрозии, а кавитация в крови может вызвать заболевание сердца и артерий. В ряде случаев кавитация играет положительную роль, поэтому в настоящее время делаются попытки использовать это явление в некоторых областях техники, химической технологии и медицине, в связи с чем требуется более подробное исследование процессов, происходящих в акустическом поле.

Данные процессы отражаются в сложном поведении газового пузырька, движение которого можно зарегистрировать оптическими методами.. Однако непредсказуемость траектории, изменение формы и размера газового пузырька в акустическом поле усложняет анализ его движения. Существующие программы определения скорости потока достаточно сложны для пользователя в настройке параметров, ресурсоемки, требуют значительного времени обработки для ряда изображений и локально усредняют определяемую скорость [3].

В данной работе проведены исследования сложного движения пузырьков в акустическом поле в диапазоне частот от 120 до 150 кГц. Дано описание экспериментальной установки и приведён пример результата обработки полученных видео снимков с помощью специально разработанной программы по отслеживанию траектории одиночного пузырька.

2. Экспериментальная установка

Для исследования закономерностей движения кавитационных пузырьков в акустическом поле была собрана установка, структурная схема которой представлена на рисунке 1. Здесь, 1 – лазер; 2 – оптическая система; 3 – лазерная плоскость; 4 – кювета; 5 – цифровая видеокамера; 6 – акустическое поле в жидкой среде; 7 – область регистрации; 8 – пъезоэлемент, рабочая частота 120 – 150 кГц; 9 – низкочастотный генератор Г-109 с рабочим диапазоном частот 17,7 Гц – 200 кГц;

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

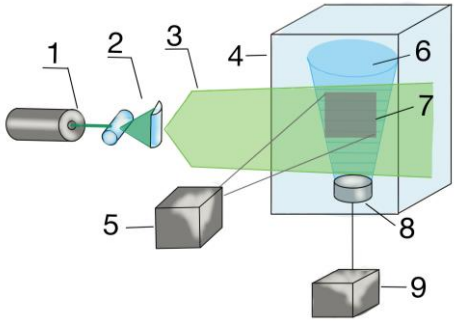


Рис. 1. Схема экспериментальной установки.

Лазер – полупроводниковый лазерный модуль «G80» (длина волны излучения – 0,532 мкм; диаметр пучка 3 мм; максимальная мощность излучения 80 мВт .

Блок 2 формирования лазерной плоскости состоит из одной цилиндрической линзы из стекла К8 диаметром 5 мм и показателем преломления n=1,51916, обеспечивающей получение расходящейся в пространстве лазерной плоскости. Так же для некоторых видов съёмок добавлялась плосковыпуклая (цилиндрическая) линза с фокусным расстоянием 12 мм для расширения лазерной плоскости в направлении перпендикулярном плоскости наблюдения, для обеспечения возможности визуализации двумерного движения газового пузырька .

Акустические поля, создавались пьезоэлектрическим ультразвуковым излучателем в кювете с внутренними размерами 190×181×292 мм^3 и прозрачными стенками. Объем заливаемой воды был равен 5 литрам. Ультразвуковой излучатель диаметром 9 мм работал в диапазоне частот от 120 до 150 кГц и располагался на дне кюветы

На данной установке могут проводиться эксперименты по наблюдению движения пузырьков (рис.2-4), возникающих под действием акустического поля, с последующей цифровой обработкой картин визуализации).

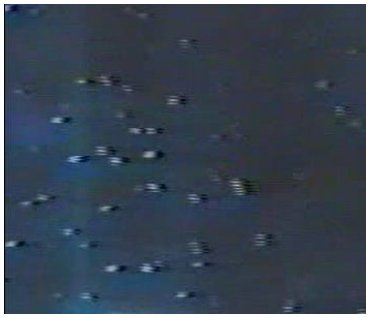


Рис. 2. Визуализация пузырьков в рассеянном излучении.

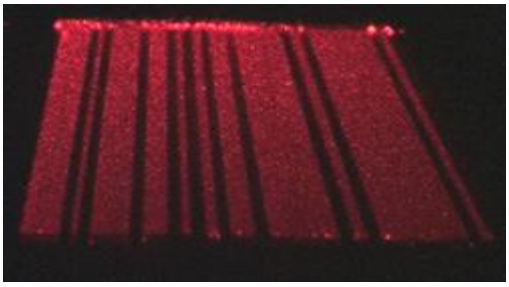


Рис. 3. Визуализация «тени» от пузырьков в плоском пучке.

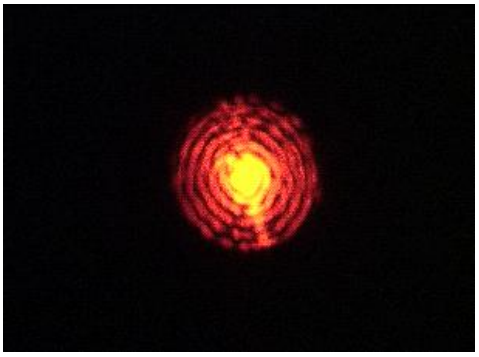


Рис. 4. Дифракционное изображение пузырька.

Результаты экспериментов по визуализации тени от пузырьков и интерференционных изображений были подробно исследованы в [4-5], однако в указанных работах акцент был сделан в основном на исследовании размеров пузырьков, скорость их регистрировалась только в момент пересечения пузырьками лазерного пучка. В [5-8] исследовалась возможность регистрации параметров кавитационного шума на основе рефракционных искажений лазерного пучка.

В представленном докладе визуализируются траектории пузырьков и выявляются характерные закономерности их движения в условиях кавитации.

3. Цифровая обработка изображений

Для улучшения качества изображения, а точнее для подавления шумов, сглаживания контуров и выделения границ с последующей сегментацией используются цифровые методы обработки изображений, которые можно разделить на пространственные и частотные методы [9].

Термин пространственная область относится к плоскости изображения как таковой, и данная категория объединяет подходы, основанные на прямом манипулировании пикселями изображения. Фильтрация изображения f , имеющего размеры $M \times N$, с помощью фильтра w размерами $m \times n$ задается выражением общего вида:

$$g(x,y) = \sum_{S=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t), \qquad (1)$$

где $a = ((m-1))/2$ и $b = ((n-1))/2$. При фильтрации всего изображения данная формула должна быть вычислена для всех сочетаний $x = 0, 1, 2, \ldots, M-1$, $y = 0, 1, 2, \ldots, N-1$. Это означает, что все элементы изображения будут обработаны по заданной маске. Данное выражение принято называть операцией свертки.

Обычно если нужно выделить границы изображения применяют следующую последовательность фильтров:

1) сглаживающий (чаще всего фильтр Гаусса);
2) бинаризацию;
3) выделения границ (чаще всего фильтр Собеля).

Данная последовательность заставляет выполнять операцию свертки по крайней мере два раза, что уменьшает скорость работы алгоритмов. Однако, если воспользоваться фильтром, полученным взятием первых частных производных от функции Гаусса:

$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \qquad (2)$$

где σ – коэффициент, определяющий радиус ядра Гауссианы; x, y – координаты внутри фильтра, то можно

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

получить фильтр выделения контура изображения, который требует провести операцию свертки только один раз, что заметно сокращает время обработки [10].

Такой фильтр часто называют ядром Канни (в некоторых источниках Кэнни) и используют для дальнейшей обработки, целью которой является выделение тонких контуров [11].

В данной работе по контуру объекта определяется его центр с использованием формул:

$$X_c = \left(\sum_{i=x_s}^{x_e}\sum_{j=y_s}^{y_e} j \cdot f(x,y)\right) \Big/ \left(\sum_{i=x_s}^{x_e}\sum_{j=y_s}^{y_e} f(x,y)\right), \quad (3)$$

$$Y_c = \left(\sum_{i=x_s}^{x_e}\sum_{j=y_s}^{y_e} i \cdot f(x,y)\right) \Big/ \left(\sum_{i=x_s}^{x_e}\sum_{j=y_s}^{y_e} f(x,y)\right). \quad (4)$$

где x_s, y_s, x_e, y_e – координаты начала и конца области изображения $f(x,y)$, содержащей одиночный объект.

Пример обработки такой области изображения с центром объекта, показанным красной точкой, представлен на рисунке 5.

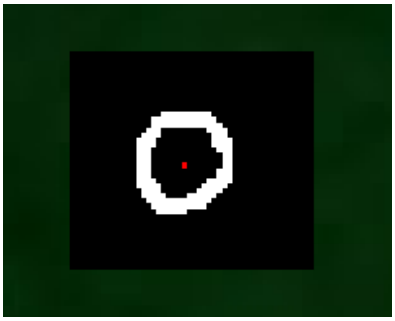


Рис. 5. Выделение контуров одиночного пузырька и определение его центра

4. Программа определения траектории выделенного объекта

В данной работе в интегрированной среде разработки программного обеспечения Microsoft Visual Studio на языке C# создавалось приложение Windows Forms (оконное приложение) для быстрой обработки ряда фотографий с целью определения траектории движения отдельно взятого пузырька. Вид созданного оконного приложения показан на рисунке 6.

Рис. 6. Окно программы до начала работы

В данном приложении пользователь может выбрать первую фотографию, мышкой выделить область на изображении, которая содержит интересующий объект, затем выбрать ряд последовательных фотографий, на которых требуется определить положение данного объекта.

В результате обработки приложение определит центр объекта на каждой фотографии, занесёт его в массив данных, как координаты пикселей на фотографии, который будет затем сохранен в текстовом документе под названием matrix.txt в заранее выбранной пользователем папке, так же происходит визуализация движения объекта на первой фотографии линией, которая меняет свой цвет от чёрного к красному от начальной до конечной точки как показано на рисунке 7.

5. Результаты обработки изображений

Пример результата работы программы в виде визуализации траектории газового пузырька, движущегося при воздействии акустического поля с частотой 150 кГц представлен на рисунке 7.

Рис. 7. Движение пузырька по эллипсоидальной траектории

По массиву полученных координат центра пузырька в пикселях по каждому кадру построены графики, представленные на рисунках 8 и 9, а на рисунке 10 вычисленная по этим координатам скорость.

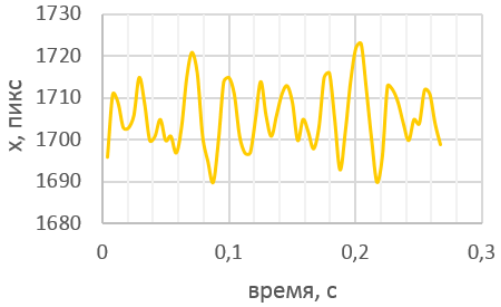


Рис. 8. Значение координаты *x* на каждом кадре при движении по эллипсоиду

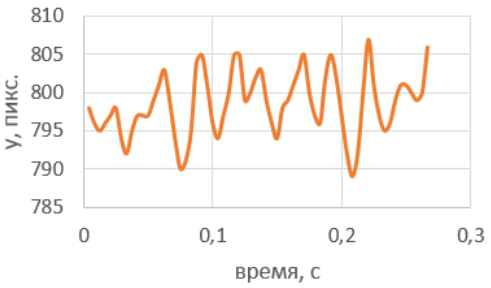


Рис. 9. Значение координаты y на каждом кадре при движении по эллипсоиду

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

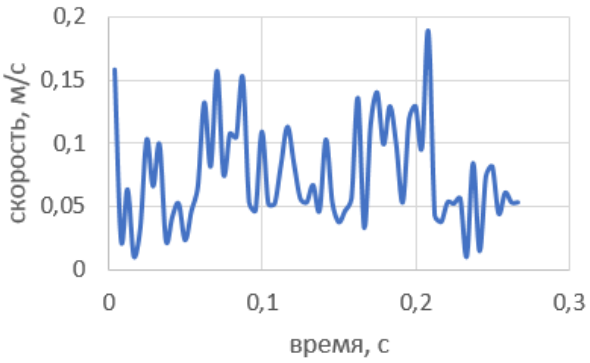


Рис. 10. Значение скорости между двумя соседними кадрами при движении пузырька по эллипсоиду

Если посчитать среднее арифметическое по всем кадрам, то скорость будет равна $u_{\text{ср.ар}} = 0{,}075$ м/с. Сделав преобразование Фурье по одной из координат, можно узнать с какой частотой пузырёк занимает изначальную позицию, рисунок 11.

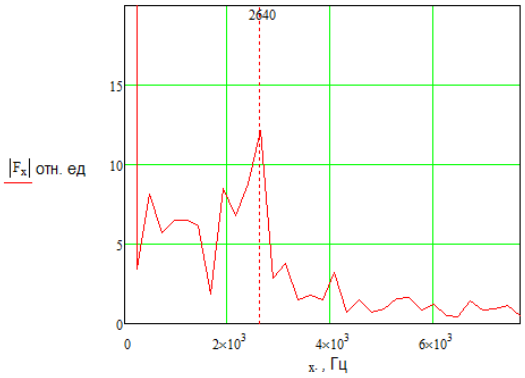


Рис. 11. Частота движения пузырька по эллипсоидальной траектории, построенная по координате y

Таким образом с помощью составленной программы можно проанализировать сложное движение объекта и выявить характерные закономерности данного движения.

6. Заключение

В данной работе было изучено сложное поведение пузырьков газа в акустическом поле. Исследование проводилось на созданной экспериментальной установке посредством видеофиксации на камеру с частотой съемки 240 кадров в секунду. Обработка полученных видеокадров проводилась специально разработанной программой. Из преимуществ программы можно выделить:

а) возможность обработки большого числа изображений;
б) скорость обработки;
в) относительно не сложную настройку параметров обработки;
г) работа программы не зависит от характера движения объекта, за которым ведется наблюдение;
д) возможность получить данные о скорости движения объекта по каждому кадру;
е) возможность сохранения промежуточных результатов обработки с заданной периодичностью;
ж) визуализация траектории движения объекта.

Из недостатков:

а) возможность слежения только за одним объектом за раз;
б) сильная погрешность или неадекватность работы программы при перекрытии объектов;
в) разрешающая способность сильно зависит от камеры, на которой велась съёмка;
г) необходимость подбора настроек, чтобы объект не исчез из поля наблюдения, и чтобы на результат не повлияли другие объекты.

Таким образом, на основе проведенного исследования разработана методика визуализации траектории пузырьков и анализа их движения в условиях акустической кавитации.

7. Литература

[1] Пирсол, И.С. Кавитация [Текст] / пер. с англ. Ю. Ф. Журавлева, под редакцией Л. А. Эпштейна - М.: «МИР», 1975. 95 с

[2] Сиротюк, М. Г. Акустическая кавитация [Текст] / М. Г. Сиротюк; отв. ред. В. А. Акуличев, Л. Р. Гаврилов. – Тихоокеан. океанолог. ин-т им. В.И. Ильичева ДВО РАН. – М.: Наука, 2008. – 275 с.

[3] Крачевский, М.Н. Алгоритмы распознавания и слежения за пузырями для измерения параметров кавитации на гидрокрыле / М. Н. Крачевский, И. Е. Полетаев, Г. С. Сухоруков // Вестн. Новосиб. гос. ун-та. Серия: Информационные технологии. – 2016.– Т. 14, №1. С. 23-38.

[4] Расковская, И.Л. Теоретическое и экспериментальное обоснование лазерных методов диагностики акустического поля в жидкостях и газах/ диссертация на соискание ученой степени кандидата физико-математических наук. Москва, 2005.

[5] Rinkevichyus, B.S., Raskovskaya, I.L., Tolkachev, A.V. Diffraction-shadow technique of single bubble size and velocity measurements/ В сборнике: Proceedings of SPIE - The International Society for Optical Engineering. Optical Methods of Flow Investigation. Сер. "Optical Methods of Flow Investigation", Russia. Moscow, 2006. С. 62620H.

[6] Расковская, И.Л. Распространение лазерного пучка в среде с акустической волной / Радиотехника и электроника. 2004. Т. 49. № 11. С. 1382-1389.

[7] Grechikhin, V.A., Raskovskaya, I.L., Runkevichyus, B.S. Influence of the acousto-optic effect on measurement error of particle vibration velocity in using the laser doppler anemometer / Автометрия. 2000. № 5. С. 92-101.

[8] Raskovskaya, I.L., Pavlov, I.N. Rinkevichyus, B.S. Tolkachev, A.V. Vedyashkina A.V. Laser techniques for monitoring physical processes in water under substantial refraction conditions / Proceedings of XVII International Conference «Laser Optics 2016», St. Petersburg, June 27-July 1, 2016, pp. R7-5.

[9] Гонсалес, Р., Цифровая обработка изображений [Текст] / Р. Гонсалес, Р. Вудс Москва: Техносфера, 2005. – 1027 с. – ISBN 5-94836-028-8

[10] Sohaib Khan Canny's edge detector: implementation / Sohaib Khan. – October, 2002.

[11] Есин, М.В., Расковская, И.Л., Ринкевичюс, Б.С., Толкачев, А.В. Трехмерные рефрактограммы и их применение для диагностики градиентных неоднородностей /Радиотехника и электроника. 2012. Т. 57. № 4. С. 485-491.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Методы визуальной аналитики вариативности речевого поведения пользователей социальных сетей в зависимости от психологических черт личности

К.В. Рябинин[1], С.И. Чуприна[1], К.И. Белоусов[1], С.С. Пермяков[1]
kostya.ryabinin@gmail.com|chuprinas@inbox.ru|belousovki@gmail.com|rewmad@gmail.com
[1]Пермский государственный национальный исследовательский университет, Пермь, Россия

Работа посвящена вопросам создания методов и средств визуального анализа, направленного на выявление зависимостей между параметрами речевого поведения пользователей социальных сетей и психологическими характеристиками личности. Данные, подлежащие анализу, обладают высокой связностью, поэтому в качестве основного визуального средства отображения предлагается использовать различные типы графов. В работе представлены новые способы визуализации кругового графа с кольцевой иерархической шкалой и шкалой срезов данных, а также различные варианты укладки на плоскости графа свободной структуры. Оба средства визуализации имеют унифицированный интерфейс доступа к данным, что позволяет подбирать наиболее подходящий для анализа способ отображения данных и тем самым легко адаптировать аналитическую систему к специфике решаемых задач. Для увеличения когнитивной мощности разработанных средств предусмотрены возможности интерактивного взаимодействия с пользователем и механизмы настраиваемой семантической фильтрации данных. Разработанные средства включены в состав управляемой онтологиями адаптивной мультиплатформенной системы научной визуализации SciVi, которая интегрирована в систему лингвистического анализа Семограф в качестве основного инструмента визуальной аналитики.

Ключевые слова: *визуальная аналитика, онтологический инжиниринг, графы, языковые параметры, психологические характеристики, BFI, пользователи социальных сетей.*

Visual analytics methods of the verbal behavior variability of social networks users depending on their individual psychological features

K.V. Ryabinin[1], S.I. Chuprina[1], K.I. Belousov[1], S.S. Permyakov[1]
kostya.ryabinin@gmail.com|chuprinas@inbox.ru|belousovki@gmail.com|rewmad@gmail.com
[1]Perm State University, Perm, Russia

The paper is devoted to the creation of visual analytics methods and means for identifying the dependencies between the parameters of verbal behavior of social networks users and their psychological characteristics. The data, which are to be analyzed, have a high connectivity, therefore we suggest to display them using different types of graphs. The paper presents new ways to visualize a circular graph with a hierarchical ring scale and a data states scale, as well as various options for laying out a free-structure graph on the plane. Both visualization tools have a unified data access interface, which allows to choose the most suitable depicting way for particular data analysis and thereby to adapt the analytical system to the specifics of the tasks being solved. To increase the cognitive power of the developed tools advanced user interactions and custom semantic data filtering mechanisms are supported. The developed tools are included in the SciVi ontology driven adaptive multiplatform scientific visualization system. This system is integrated into the Semograph linguistic analysis system as the main visual analytics engine.

Keywords: *visual analytics, ontology engineering, graphs, language parameters, psychological characteristics, BFI, social networks users.*

1. Введение

В современной науке существует запрос на создание фундаментальной концепции личности, которая бы позволила описывать, объяснять и прогнозировать речевое и неречевое поведение человека и социальных групп, включая группы пользователей социальных интернет-сервисов (англ. Social Network Services, SNS). Несмотря на широкий спектр задач, решаемых в области исследования SNS, в открытых источниках пока не встречаются концепции комплексного анализа типов их пользователей, взаимосвязей между ними и моделей их поведения.

Комплексное описание пользователей SNS должно основываться на моделях интеграции социального, поведенческого, психологического и языкового (и, более широко, мультимодального) профилей цифровых проекций личностей. В качестве социальных параметров рассматривается информация из профиля пользователя (пол, возраст, образование, сфера интересов, социальное окружение и др.); в качестве поведенческих – предпочтения (например, отмеченные как понравившиеся публикации и др. материалы, размещаемые в сети) и т. п. Психологические параметры выявляются в результате психологических опросов, а языковые – на основе анализа комментариев пользователей и тегов к размещаемым материалам.

Для автоматизации анализа и интерпретации данных из социальных сетей предлагается использовать аппарат визуальной аналитики, основанный на когни-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

тивной графике, а также на методах и средствах научной визуализации. Данные из социальных сетей зачастую обладают высокой связностью, и именно наличие и характер связей обычно выступают предметом анализа. Соответственно, в качестве базового способа визуализации предлагается использовать интерактивные графы различной структуры, расширенные настраиваемыми средствами предобработки (в частности, семантической фильтрации) отображаемых данных.

В качестве аналитической платформы предлагается использовать систему лингвистического анализа Семограф [2]. Целью данной работы является внедрение в Семограф новых средств визуального анализа вариативности речевого поведения пользователей социальных сетей в зависимости от психологических черт их личности. За визуализацию данных, обрабатываемых системой Семограф, отвечает интегрированная с ней система научной визуализации SciVi [10]. Для достижения поставленной цели разработаны и включены в состав SciVi два модуля визуализации графов, обладающие унифицированным интерфейсом доступа к данным: модуль визуализации круговых графов с настраиваемой кольцевой иерархической шкалой и шкалой срезов данных, а также модуль визуализации графов со свободной структурой.

Для тестирования и отладки созданных средств визуализации в работе использовались данные 821 пользователя социальной сети ВКонтакте, участвовавших в психологическом опросе (Вопросник Большой Пятерки – BFI [8]). Языковые параметры выделялись на материале многоуровневого лингвистического анализа 18000 автоматизировано собранных реплик информантов. Анализ выполнялся тремя экспертами в системе Семограф. В данной работе рассмотрены некоторые из анализируемых речевых параметров, относящиеся к семантике (дейксис, модальность), стилистические параметры (бранная лексика, аугментативы и др.) и параметры, касающиеся использования в диалогах пользователей графических средств (эмотикон) в качестве полноценных реплик.

2. Онтологический инжиниринг в визуальной аналитике

В ходе предыдущих исследований были проанализированы наиболее популярные системы научной визуализации (например, TecPlot, ParaView, Avizo, VizIt и др.) и установлено, что самым серьёзным недостатком большинства из них является отсутствие высокоуровневых средств адаптации к нестандартным задачам [6]. Нестандартность задач визуализации может выражаться либо в особенностях источника данных (что требует использования специальных алгоритмов доступа или представления данных в некотором нестандартном формате), либо в специфических требованиях к отображаемым графическим объектам и сценам.

Один из возможных путей для обеспечения высокой гибкости программных средств научной визуализации – использование при их создании модельно-ориентированного подхода. Система, поведение которой полностью или хотя бы частично управляется некоторой декларативной формальной моделью, может быть быстро перенастроена для наилучшего соответствия требованиям решаемых задач. В роли такой модели предлагается использовать онтологии, так как они обеспечивают человекочитаемость и самодокументируемость. В состав базы знаний системы визуализации предлагается включить онтологию визуальных объектов и графических сцен, описывающую поддерживаемые системой средства визуализации, а также онтологию семантических фильтров, описывающую допустимые способы трансформации входных данных.

Для решения специализированных задач предлагается использовать дополнительные онтологии. Например, если источником подлежащих визуализации данных выступает некоторый решатель (англ. Solver – расчётная программа), в состав системы визуализации предлагается включить онтологию синтаксических конструкций ввода/вывода языка программирования, на котором написан этот решатель. Такая онтология служит для целей автоматической генерации синтаксического анализатора, задачей которого является извлечение из исходного кода решателя структуры его выходных данных и управляющих параметров. Это, в свою очередь, позволяет автоматизировать процесс настройки системы визуализации на взаимодействие с решателем. В случае необходимости встраивания системы визуализации в некоторое аппаратное устройство, например, в какой-либо элемент экосистемы Интернета вещей, предлагается использовать онтологию электронных компонентов [7]. Она служит для автоматической генерации прошивки электронного устройства и интеграции в эту прошивку кода визуализации.

Предложенные принципы построения систем научной визуализации были реализованы при разработке упомянутой выше мультиплатформенной системы SciVi. Методы и средства онтологического инжиниринга, использованные при создании SciVi, хорошо зарекомендовали себя на практике в контексте обеспечения высокой настраиваемости и адаптивности системы визуализации к специфике решаемых задач и индивидуальным предпочтениям пользователей, а также унификации процесса пополнения функциональности системы [7]. Так, например, добавление поддержки новых механизмов рендеринга сводится к пополнению онтологии визуальных объектов и графических сцен описанием новой функциональности и ссылками на внешние модули, реализующие эту функциональность. При этом нет необходимости в модификации исходного кода ранее отлаженных функций и ядра системы.

В тех случаях, когда помимо наглядного представления научных данных требуется также их глубинный анализ, зачастую необходимо выполнить над ними некоторые преобразования. Например, фильтрацию данных в соответствии с заданными критериями (поро-

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

говыми функциями и т. п.), математические преобразования (масштабирование, нормализация и т. п.), классификацию и кластеризацию, статистический анализ и т. д. Для поддержки и унификации механизма такого рода трансформаций предлагается использовать т. н. семантические фильтры – операторы преобразования подлежащих визуализации данных. Описание алгоритма работы этих операторов, их входов, выходов и настроечных параметров хранится в онтологии семантических фильтров. Это позволяет пополнять их набор путём внесения изменений только в базу знаний системы научной визуализации с помощью средств высокоуровневого пользовательского интерфейса.

Семантические фильтры допускают суперпозицию, задаваемую диаграммой потока данных [5]. В системе SciVi присутствует специальный высокоуровневый графический редактор для составления таких диаграмм. Вершинами в них выступают источники данных, семантические фильтры, визуальные объекты и графические сцены, а связи выражают пути передачи данных. Палитра допустимых вершин формируется автоматически на основе соответствующих онтологий.

Наличие расширяемого набора семантических фильтров, а также удобных средств для их комбинирования, наряду с другими описанными в данной статье возможностями превращает систему научной визуализации SciVi в полноценное средство визуальной аналитики [6].

3. Онтологический профиль анализируемых данных

Для того чтобы представить алгоритм обработки и визуализации данных в виде диаграммы потока данных, необходимо в качестве начальной вершины задать их источник. Для этого источник данных должен быть описан онтологически при помощи тех же концептов, которые используются для описания семантических фильтров и визуальных объектов (выходные данные с указанием их типов, настроечные параметры и т. п.). Такое описание составляет т. н. онтологический профиль.

На рис. 1 в обобщенном виде приведен онтологический профиль данных, извлекаемых из системы Семограф, которые предназначены для визуализации в виде графов различной структуры.

Основными составляющими частями такого профиля являются вершины (англ. Nodes) и связи (англ. Edges), имеющие расширяемый набор атрибутов, таких, как название (англ. Label), вес (англ. Weight) и др. В зависимости от конкретных решаемых задач, набор атрибутов может пополняться. Влияние значений тех или иных атрибутов на результат визуализации задаётся пользователем посредством диаграммы потока данных.

Набор данных для визуализации может включать несколько срезов по какому-либо показателю, например, по времени, месту или персоне. В этом случае отображаемый граф имеет несколько состояний (англ. Graph States), переключение между которыми осуществляется при помощи специальной шкалы.

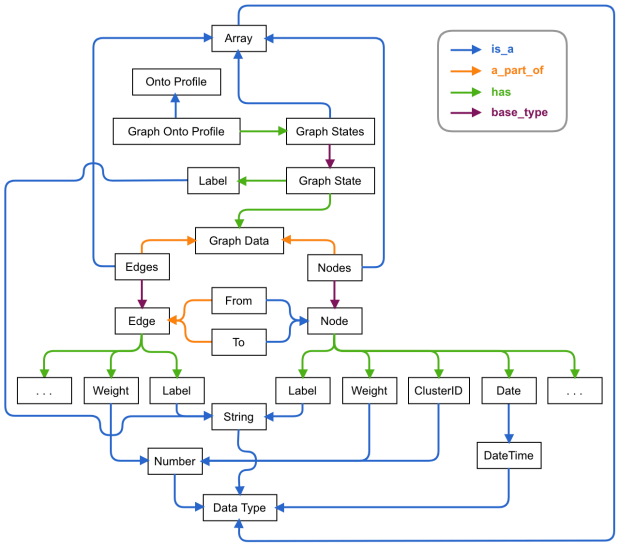


Рис. 1. Обобщённый вид онтологического профиля данных, предназначенных для представления в виде графа.

Для целей анализа зачастую может потребоваться кластеризация данных. Она может быть выполнена заранее (на стороне системы Семограф, различные подсистемы которой выступают здесь в качестве решателей), и в этом случае в число атрибутов вершины будет входить идентификатор кластера (англ. ClusterID). Кроме этого, можно задать своего рода «кластеризацию на лету», если того требуют цели визуального анализа данных. В этом случае данные разбиваются на кластеры не решателем, а самой системой визуализации, благодаря наличию в её составе соответствующих семантических фильтров. На данный момент для кластеризации используется Лёвенский алгоритм [3].

4. Круговой граф с настраиваемой иерархической кольцевой шкалой

Для структурированного отображения данных высокой связности используется круговой граф [1]. Его вершины расположены по окружности на равном расстоянии друг от друга. Вес вершин отображается гистограммой, столбцы которой рисуются как фон для названий вершин. Принадлежность вершины к тому или иному кластеру отображается при помощи цвета.

Дуги графа представляют собой квадратичные параболы, построенные по трём контрольным точкам. Первая и третья контрольные точки находятся в соединяемых вершинах, а вторая лежит в центре окружности. Толщина дуг отражает их вес.

В зависимости от настроек, сделанных пользователем, к вершинам может быть применена многоуровневая группировка по ряду связанных с ними показателей. Принадлежность вершин к группам, сформированным по этим показателям, отображается при помощи иерархической кольцевой шкалы.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Для обеспечения необходимой аналитической функциональности в круговом графе реализована поддержка интерактивности, включающая возможности масштабирования, фильтрации вершин и дуг по весу, выделения, переноса и переименования отдельных вершин, изменения цвета, выделения дуг, перехода к отображению одиночных кластеров вершин (т. н. квази-зум [9]), а также изменения порядка следования уровней иерархии кольцевой шкалы (с соответствующей перегруппировкой вершин). Модуль визуализации кругового графа функционирует на основе библиотеки графического расширения PixiJS[1] и оптимизирован для работы в WebGL-совместимом браузере.

На рис. 2 приведён круговой граф, построенный по результатам лингвистического анализа 18000 реплик информантов – пользователей социальной сети ВКонтакте, участвовавших в психологическом опросе BFI [8]. Часть из выделенных языковых параметров представлена в нижнем полукружии, психологические параметры даны в верхнем полукружии. Языковые параметры выделялись при анализе отдельных реплик, психологические параметры информанта приписывались всем его репликам. Таким образом, «психологические» вершины графа, которые представляют определенное количество информантов-носителей данных черт личности, соединялись с «языковыми» вершинами графа, отражающими языковые параметры реплик, принадлежащих информантам данного типа. Кроме того, результаты, представленные на графе, можно дополнительно отфильтровать по гендерной принадлежности информантов.

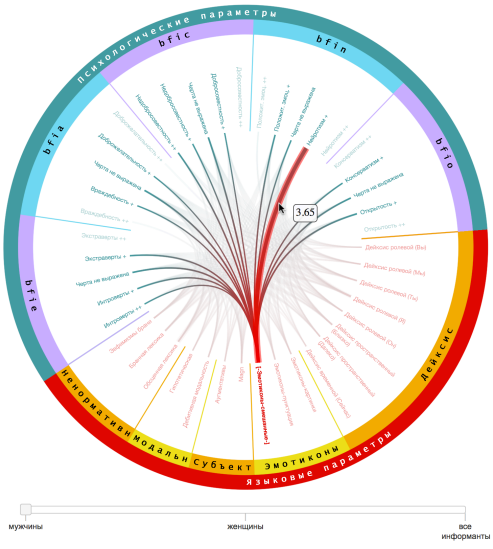


Рис. 2. Круговой граф с иерархической кольцевой шкалой и шкалой срезов данных.

5. Граф свободной структуры

Для отображения произвольных связных данных предлагается использовать граф свободной структуры. Вершины в нём располагаются согласно алгоритму укладки на основе квази-физических аналогий [4] в режиме реального времени.

Вершины представляют собой различные геометрические примитивы, такие как квадрат или окружность. Поддерживается ранжирование вершин, при этом вершинам разного ранга назначается различный внешний вид. Ранжирование осуществляется на основе кластеризации данных по различным критериям. Размеры вершин отражают их веса. Дуги графа отображаются при помощи отрезков прямых, их веса отражаются толщиной отрезков. Модуль визуализации графа свободной структуры разработан на основе библиотеки VivaGraphJS[2] и оптимизирован для работы в WebGL-совместимом браузере.

На рис. 3 приведён пример графа свободной структуры, построенного по тем же данным, что и граф на рис. 2.

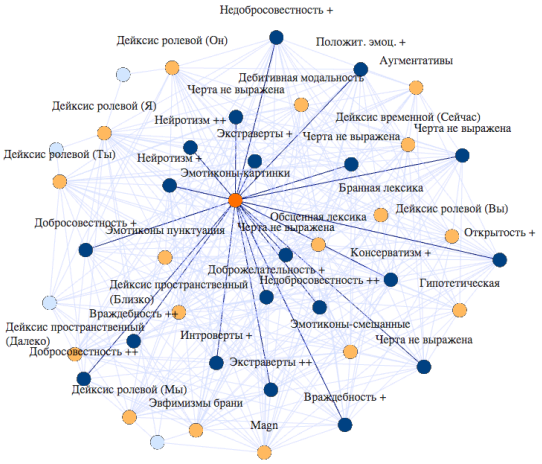


Рис. 3. Двудольный граф свободной структуры.

Поддерживаются различные средства интерактивности: перемещение, масштабирование и поворот всего изображения; перемещение отдельных вершин; задание/изменение цвета, приписываемого вершин одного ранга; возможность показывать/скрывать отдельные вершины. Реализован механизм подсветки выбранной вершины в зависимости от её ранга: при выборе вершины автоматически подсвечиваются все связанные с ней вершины более низкого ранга.

Для уменьшения загруженности визуального образа по умолчанию скрываются названия всех вершин, кроме вершин наивысшего ранга. Пользователь может изменять настройки, связанные с укладкой графа на плоскости. Поддерживается фильтрация отображаемых вершин и дуг по весу, при этом в случае работы с двудольным графом фильтрацию вершин можно осуществлять отдельно для каждой доли.

6. Заключение

В работе представлены новые возможности системы научной визуализации SciVi: круговой граф с кольцевой иерархической шкалой и шкалой срезов данных,

[1] http://www.pixijs.com/
[2] https://github.com/anvaka/VivaGraphJS

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

а также граф свободной структуры, поддерживающий ранжирование вершин и различные варианты укладки на плоскости. Эти средства были использованы для наглядного представления и визуальной аналитики извлекаемых из социальных сетей данных о речевом поведении пользователей. Первичная обработка данных производилась средствами системы лингвистического анализа Семограф.

Интеграция SciVi в систему Семограф позволила выявить соответствия между установленными в результате психологического опроса, в котором принимали участие более 800 пользователей социальной сети ВКонтакте, чертами личности пользователей и особенностями их речевого поведения в социальной сети, выявленными при помощи средств лингвистического анализа.

Разработанные средства визуализации позволили на основе анализа имеющихся и отсутствующих связей между отдельными языковыми и психологическими параметрами получить значимую для предметной области информацию. Так, в частности, были выявлены различия в использовании средств ролевого (в том числе социально маркированного) и пространственного дейксиса у пользователей-экстравертов и интровертов. Интересны различия в использовании бранной и обсценной лексики в письменной речи пользователей, имеющих ярко выраженные черты консерватизма и открытости и мн. др. Кроме того, речевая вариативность может объясняться не только психологическими различиями, но и гендерными (в частности, публичное использование обсценной лексики имеет связи с разными психологическими характеристиками в мужской и женской группах пользователей). Визуальная модель языковых и психологических соответствий, построенная с учетом гендерных характеристик, позволяет извлекать релевантную информацию об организации данной предметной области; строить интерпретации с опорой на визуальную основу (это особенно удобно, так как модель дает возможность перебора самых разных вариантов); детализировать направления научного поиска, которые дают нетривиальную информацию о связях психологической, языковой и социальной природы человека в процессе осуществления своей жизнедеятельности.

Благодаря управляемым онтологиями адаптационным возможностям системы SciVi разработанные средства могут быть использованы для визуального анализа любых многомерных данных, генерируемых различными решателями, включая устройства в составе экосистемы Интернета вещей.

В ближайших планах предусмотрено использование SciVi для комплексного анализа речевого и неречевого поведения пользователей социальных сетей. В дальнейшем в составе SciVi планируется реализовать модули визуализации данных с картографической привязкой, а также модули построения настраиваемых многоуровневых трёхмерных иерархических графов.

7. Благодарности

Работа выполнена в рамках государственного задания Минобрнауки России (проект 34.1505.2017/4.6).

8. Литература

[1] Ageev A. A Triangle-free Circle Graph with Chromatic Number 5 // Discrete Mathematics. – 1996. – Vol. 152. – PP. 295–298. DOI: 10.1016/0012-365X(95)00349-2.

[2] Belousov K., Erofeeva E., Leshchenko Y., Baranov D. “Semograph” Information System as a Framework for Network-Based Science and Education // Smart Innovation, Systems and Technologies. Smart Education and e-Learning. – Springer, 2017. – PP. 263–272. DOI: 10.1007/978-3-319-59451-4_26.

[3] Blondel V.D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks // Journal of Statistical Mechanics: Theory and Experiment. – 2008. – No. 10. – 12 P. DOI: 10.1088/1742-5468/2008/10/P10008.

[4] Jacomy M., Venturini T., Heymann S., Bastian M. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software // PLoS ONE. – 2014. – No. 9, I. 6. DOI: 10.1371/journal.pone.0098679.

[5] Lee B., Hurson A. Issues in Dataflow Computing // Advances in Computers. – Elsevier, 1993. – Vol. 37.– PP.285–333.DOI: 10.1016/S0065-2458(08)60407-6.

[6] Ryabinin K., Chuprina S. High-Level Toolset For Comprehensive Visual Data Analysis and Model Validation // Procedia Computer Science. – Elsevier, 2017. – Vol. 108. – PP. 2090–2099. DOI: 10.1016/j.procs.2017.05.050.

[7] Ryabinin K., Chuprina S., Kolesnik M. Calibration and Monitoring of IoT Devices by Means of Embedded Scientific Visualization Tools // Lecture Notes in Computer Science. – Springer, 2018. – Vol. 10861. – PP. 655–668. DOI: 10.1007/978-3-319-93701-4_52.

[8] Shchebetenko S. Reflexive characteristic adaptations explain sex differences in the Big Five: But not in neuroticism // Personality and Individual Differences. – Elsevier, 2017. – Vol. 111. – PP. 153–156. DOI: 10.1016/j.paid.2017.02.013.

[9] Бондарев А.Е., Галактионов В.А., Шапиро Л.З. Обработка и визуальный анализ многомерных данных // Научная визуализация. – М.: Национальный исследовательский ядерный университет МИФИ, 2017. – К. 4, Т. 9, №5. – С. 86-104. DOI: 10.26583/sv.9.5.08.

[10] Рябинин К.В., Баранов Д.А., Белоусов К.И. Интеграция информационной системы Семограф и визуализатора SciVi для решения задач экспертного анализа языкового контента // Научная визуализация. – М.: Национальный исследовательский ядерный университет МИФИ, 2017. – К. 4, Т. 9, №4. – С. 67–77. DOI: 10.26583/sv.9.4.07.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Визуализация режимов колебаний балок с учетом их контактного взаимодействия

Салтыкова О.А.[1,2], Папкова И.В.[1], Афонин О.А. [1], Крысько А.В.[1,2]
olga_a_saltykova@mail.ru|Ikravzova@mail.ru|anton.krysko@gmail.com
[1]Саратовский государственный технический университет имени Гагарина Ю.А., Саратов, Россия;
[2]Томский политехнический университет, Томск, Россия

Целью работы является анализ режимов колебаний балочных структур с помощью визуализации. В работе представлены результаты работы программного продукта для комплексного анализа нелинейной динамики контактного взаимодействия балочных структур. Особенность разработанного программного продукта заключается в том, что реализована возможность комплексного исследования задач механики с помощью методов нелинейной динамики и качественной теории дифференциальных уравнений. Демонстрация возможностей программного комплекса представлена на примере анализа нелинейной динамики контактного взаимодействия балочных структур, находящихся под действием внешней знакопеременной нагрузки.

Ключевые слова: *балочные структуры, визуализация режимов колебаний, метод конечных разностей.*

Visualization of beam's vibrations modes taking into account their contact interaction

Saltykova O.A.[1,2], Papkova I.V.[1], Афонин О.А. [1], Krysko A.V.[1,2]
olga_a_saltykova@mail.ru|Ikravzova@mail.ru|anton.krysko@gmail.com
[1]Yuri Gagarin Saratov State Technical University, Saratov, Russia;
[2]Tomsk Polytechnic University, Tomsk, Russia.

The aim of the work is to analyse the modes of beam structures vibration with the help of visualization. The paper presents the results of the software product for a complex analysis of the nonlinear dynamics of the contact interaction of beam structures. The possibility of a comprehensive study of the problems of mechanics using the methods of nonlinear dynamics and the qualitative theory of differential equations is the peculiarity of the developed software product. Demonstration of the software complex capabilities is presented on the example of analysis of nonlinear dynamics of the contact interaction of beam structures under the action of an external alternating load.

Keywords: *beam structures, visualization of vibration modes, finite difference method.*

1. Введение

В современных условиях быстро развивающейся промышленности актуальным является исследование динамического поведения элементов механических структур, таких как балки, пластинки, оболочки. Проведение натурного эксперимента является ресурсозатратным и не всегда возможно. Поэтому необходима разработка программного комплекса для проведения таких исследований.

Изучение уже существующих на рынке программных комплексов, таких как Flow HPC [10], SOLVCON [1], DIASTARS [8] и др. показало их узкую специализацию.

Если говорить о программных комплексах конечно-элементного моделирования, таких как ANSYS, ABAQUS, COMSOL, способных решать широкий класс механических и физических задач, то помимо высокой стоимости к их недостаткам можно отнести закрытость программного кода. Эти программные продукты направлены на решение задач механики и физики, в то время как в настоящее время актуальным вопросом является возможность визуализации больших объемов данных для получения зрительных образов изучаемых систем. Многоузловой системный подход широко используется для анализа динамического поведения систем, и для этой задачи доступно несколько программ. Но по-прежнему отсутствует инструментарий для систематического синтеза нелинейных динамических систем. Синтез требует оптимального выбора параметров системы по отношению к данному показателю эффективности.

Целью данной работы является разработка программного комплекса для визуализации решений задач нелинейной динамики. В качестве языка программирования был выбран C# в силу его гибкости и эффективности.

2. Описание программного комплекса

Разработанный программный комплекс является универсальным и позволяет решать системы нелинейных дифференциальных уравнений, описывающих колебания элементов механических структур, с помощью нескольких принципиально разных методов: по пространственной координате это метод конечных элементов в форме Бубнова-Галеркина, метод конечных разностей с аппроксимацией $O(h^2)$, $O(h^4)$, $O(h^6)$, метод Фаэдо-Галеркина в высших приближениях; по времени это метод конечных разностей, методы типа Рунге-Кутта, такие как метод Рунге-Кутта 4-го, 2-го порядков, метод Рунге-Кутта-Фелберга 4-го порядка, метод Кеш-Карпа 4-го порядка, Рунге-Кутта Принса-Дорманда 8-го порядка, неявные методы Рунге-Кутта 2-го и 4-го порядка [7]. Для каждой задачи предусмотрена возможность построения сигналов, фазовых портретов 2D и 3D, спектров мощности Фурье, вейвлет спектров, эпюр прогибов, сечений Пуанкаре. В качестве анализирующих вейвлетов выбирались: вейвлет Хаара, Шеннона-Котельникова, Мейера, вейвлеты Добеши, кауфлеты, симлеты, а также вейвлет Морле, комплексный Морле и вейвлеты на основе производной функции Гаусса старше 8 порядка [2]. Предусмотрена возможность вычисления ляпуновских показателей по алгоритмам: Канца [3], Вольфа [11], Розенштейна [6]. Применение нескольких методов для решения одной и той же задачи

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

позволяет обеспечить истинность получаемых решений, что показано авторами настоящей работы ранее [4], [5].

Для принятия решений по управлению механическими системами с помощью внешних воздействий, реализована возможность построения карт динамических режимов и карт Ляпуновских показателей в зависимости от управляющих параметров. Карты представляют собой двумерные изображения характера колебаний изучаемой системы в зависимости от изменения управляющих параметров. Разрешение карт может быть различным, в зависимости от цели. Для наших исследований определяющей является зависимость характера колебаний от амплитуды и частоты внешней силы. С точки зрения прикладной механики важно четко выделить зоны изменения амплитуды внешней силы, при которых колебания носят хаотический характер. “Карта” — есть цветовое представление характера колебания в зависимости от амплитуды и частоты внешней нагрузки. “Карты” строятся на основе анализа спектров мощности колебаний центральной точки балки. То есть, для каждой пары управляющих параметров $\{q_0, \omega_p\}$, где q_0 — амплитуда, ω_p — частота вынуждающей нагрузки, производилось численное моделирование колебаний балки, далее рассчитывался спектр мощности для колебаний центральной точки и на основе анализа полученного спектра определялся характер колебаний всей балки. Каждый режим обозначен некоторым цветом. Отметим выделенные нами характерные режимы:

1. Затухание колебаний — случай, когда энергии внешней силы недостаточно для преодоления сил диссипации и поддержания колебаний (черный цвет).
2. Гармонические колебания на частоте ω_p — случай синхронизации системы и внешней силы (синий цвет).
3. Гармонические колебания на частоте $\omega_p / 2$ (красный цвет).
4. Колебания на частоте $\omega_p / 2^n, n \in N$ (желтый цвет).
5. Хаотические колебания — случай, когда устанавливаются негармонические колебания (белый цвет).
6. Двухчастотные колебания на частоте вынуждающей нагрузки и на линейно независимой от нее частоте (голубой цвет).
7. Колебания на частотах вынуждающей нагрузки, линейно независимой от нее частоте и их линейных комбинациях (сиреневый цвет).
8. Колебания на частоте ω_p и $n\omega_p / 3, n \in N$ (зеленый цвет).

3. Постановка задачи

Рассматриваемая структура из двух балок представляет собой двумерную область пространства R^2 с декартовой системой координат, введенной следующим образом: в теле балки 2 фиксируется линия приведения, называемая срединной линией $z = 0$, ось *OX* направлена слева направо вдоль срединной линии, ось *OZ* – вниз, перпендикулярно *OX*. В указанной системе координат структура из двух балок, как двумерная область Ω определяется следующим образом (рис. 1):

$$\Omega = \{x \in [0, a], -h \le z \le h_k + 3h\},\ 0 \le t \le \infty .$$

Уравнение движения балок, а также граничные и начальные условия получены из энергетического принципа Гамильтона-Остроградского.

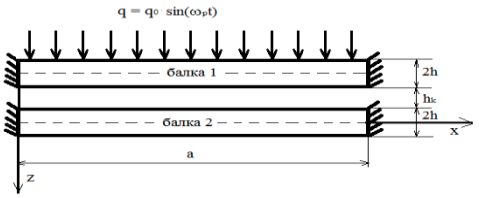


Рис. 1. Расчетная схема

При построении математической модели балок были приняты следующие гипотезы.

1. Ось балок является прямой линией.
2. Нагрузка действует в направлении оси *OZ* и внешние силы не меняют своего направления при деформации балки.
3. Контактное давление учитывается по модели Кантора Б.Я. [12].
4. Нормальные напряжения на площадках, параллельных оси, пренебрежимо малы.
5. Геометрическая нелинейность учитывается в форме Т. Кармана.

В настоящей работе рассматривается случай, когда балка 1 описывается кинематической гипотезой второго приближения, а балка 2 – кинематической гипотезой третьего приближения.

Гипотеза третьего приближения (Рэдди-Шереметьева-Пелеха) [13] подразумевает, что поперечные сечения не остаются плоскими, и не перпендикулярны деформированной оси балки, а также учитываются инерционные составляющие, связанные с поворотом сечений.

Гипотеза второго приближения [9] – гипотеза Тимошенко С.П., заключается в том, что тангенциальные перемещения распределены по толщине балки согласно линейному закону.

Для моделирования контактного взаимодействия балок по модели Кантора Б.Я. в уравнения балок необходимо ввести слагаемое $(-1)^i K(w_1 - w_2 - h_k)\Psi$, $i = 1, 2$ — номер балки, функция Ψ определена формулой $\Psi = \frac{1}{2}\left[1 + sign(w_1 - h_k - w_2)\right]$, то есть $\Psi = 1$, если есть контакт между балками — $w_1 > w_2 + h_k$, иначе контакта нет [12], K — коэффициент жесткости трансверсального обжатия структуры в зоне контакта, h_k — зазор между балками.

Для удобства изложения материала, будем под балкой 1 понимать балку, на которую действует нагрузка, а под балкой 2 – не нагруженную балку.

Приведем исходные уравнения движения пакета балок:

$$\begin{aligned}
&\frac{1}{3}\left(\frac{\partial^2 w_1}{\partial x^2} + \frac{\partial \gamma_{x1}}{\partial x}\right) + \\
&+ \frac{1}{\lambda^2}\left(L_1(w_1, u_1) + \frac{3}{2}L_2(w_1, w_1) + L_3(w_1, u_1)\right) + \\
&+ K(w_1 - w_2 - h_k)\Psi + q(t) - \frac{\partial^2 w_1}{\partial t^2} - \varepsilon_1 \frac{\partial w_1}{\partial t} = 0; \\
&\frac{\partial^2 u_1}{\partial x^2} + L_4(w_1, w_1) - \frac{\partial^2 u_1}{\partial t^2} = 0;
\end{aligned} \qquad (1)$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

$$\frac{\partial^2\gamma_{x1}}{\partial x^2}-8\lambda^2\left(\frac{\partial w_1}{\partial x}+\gamma_{x1}\right)-\frac{\partial^2\gamma_{x1}}{\partial t^2}=0;$$

$$\frac{1}{\lambda^2 63}\left[\frac{4}{5}\frac{\partial^3\gamma_{x2}}{\partial x^3}-\frac{1}{4}\frac{\partial^4 w_2}{\partial x^4}\right]+k^2\frac{G_{13}}{E_1}\cdot$$

$$\cdot\left[\frac{\partial\gamma_{x2}}{\partial x}+\frac{\partial^2 w_2}{\partial x^2}\right]+\frac{1}{\lambda^2}\left[L_3(w_2,u_2)+L_1(w_2,u_2)+\right.$$

$$\left.+\frac{3}{2}L_2(w_2,w_2)\right]+K(w_1-w_2-h_k)-$$

$$-\frac{\partial^2 w_2}{\partial t^2}-\varepsilon_1\frac{\partial w_2}{\partial t}=0;$$

$$\frac{\partial^2 u_2}{\partial x^2}+L_4(w_2,w_2)-\frac{\partial^2 u_2}{\partial t^2}=0;$$

$$\frac{204}{315}\frac{\partial^2\gamma_{x2}}{\partial x^2}-\frac{48}{315}\frac{\partial^3 w_2}{\partial x^3}-$$

$$-12\lambda^2k^2\frac{G_{13}}{E_1}\left[\gamma_{x2}+\frac{\partial w_2}{\partial x}\right]-\frac{\partial^2\gamma_{x2}}{\partial t^2}=0.$$

Здесь $L_1(w_i,u_i)=\frac{\partial^2 w_i}{\partial x^2}\frac{\partial u_i}{\partial x}$, $L_2(w_i,w_i)=\frac{\partial^2 w_i}{\partial x^2}\left(\frac{\partial w_i}{\partial x}\right)^2$, $L_3(w_i,u_i)=\frac{\partial w_i}{\partial x}\frac{\partial^2 u_i}{\partial x^2}$, $L_4(w_i,w_i)=\frac{\partial w_i}{\partial x}\frac{\partial^2 w_i}{\partial x^2}$ — нелинейные операторы, γ_{xi} — функция поперечного сдвига, w_i, u_i — функции прогибов и перемещений балок соответственно. К системе уравнений (1) следует присоединить граничные и начальные условия.

Граничные условия для заделки обоих концов балки 1:

$$w_1(0,t)=w_1(1,t)=0;\ u_1(0,t)=u_1(1,t)=0;$$
$$\gamma_{x1}(0,t)=\gamma_{x1}(1,t)=0. \qquad (2)$$

Начальные условия для балок 1 и 2:

$$w_1(x,t)_{|t=0}=u_1(x,t)_{|t=0}=\gamma_{x1}(x,t)_{|t=0}=0,$$
$$\frac{\partial w_1(x,t)}{\partial t}_{|t=0}=\frac{\partial u_1(x,t)}{\partial t}_{|t=0}=\frac{\partial\gamma_{x1}(x,t)}{\partial t}_{|t=0}=0. \qquad (3)$$

Граничные условия для заделки обоих концов балки 2:

$$w_i(0,t)=w_i(1,t)=0;\ u_i(0,t)=u_i(1,t)=0;$$
$$\gamma_{xi}(0,t)=\gamma_{xi}(1,t)=0. \qquad (4)$$

Система уравнений (1) – (4), приведена к безразмерному виду с помощью переменных (5).

$$\bar{w}=\frac{w}{2h},\ \bar{u}=\frac{ua}{(2h)^2},\ \bar{x}=\frac{x}{a},\ \lambda=\frac{a}{(2h)},\bar{q}=q\frac{a^4}{(2h)^4E},$$
$$\bar{t}=\frac{t}{\tau},\ \tau=\frac{a}{c},\ c=\sqrt{\frac{Eg}{\gamma}},\ \bar{\varepsilon}_1=\varepsilon_1\frac{a}{c},\ \bar{\gamma}_x=\frac{\gamma_x a}{(2h)}. \qquad (5)$$

Здесь: E — модуль Юнга, g — ускорение свободного падения, γ — удельный вес материала балок, ε_1 — коэффициент демпфирования.

На балку 1 действует поперечная распределенная по поверхности знакопеременная нагрузка вида:

$$q(t)=q_0\sin(\omega_p t). \qquad (6)$$

Полученная система нелинейных дифференциальных уравнений в частных производных (1) — (4) сводится к системе обыкновенных дифференциальных уравнений методом конечных разностей с аппроксимацией $O(c^2)$, где c — шаг по пространственной координате.

4. Результаты численного эксперимента

Приведем карты динамических режимов исследуемой балочной структуры при следующих параметрах: разрешение карт 400×400, $h_k=0.25;0.1;0.05$ — зазор между балками, $\omega_p\in[0.5;10.35]$, $q_0\in[0;15000]$, $\varepsilon_1=1$, $\lambda=50$, $\Delta t=9.76562\cdot10^{-4}$. Предварительно исследовался вопрос о разрешении при построении карт. При используемом разрешении было решено $16\cdot10^4$ вариантов задач для каждой балки.

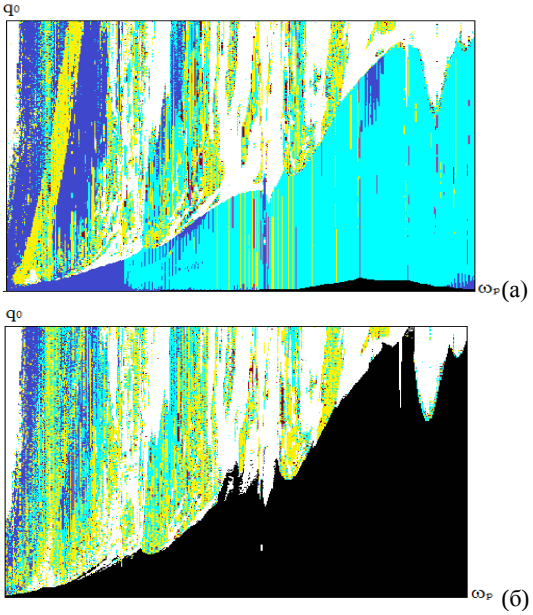


Рис. 2. Карты динамических режимов при $h_k=0.25$ а) карта для балки 1; б) карта для балки 2.

Исходя из полученного графического представления режимов колебаний балочной структуры при $h_k=0.25$ видно, что балка 2 переходит от затухающих колебаний к хаотическим по тому же контуру, по которому происходит переход балки 1. Балка 1 переходит к хаотическим колебаниям в основном чрез колебания на частоте вынуждающих колебаний и на линейно независимой от нее частоте. Площадь области гармонических колебаний балки 1 (рис. 2 а) в три раза больше, чем балки 2 (рис. 2 б), где она составляет 5%. Площадь областей хаотических колебаний балок почти одинакова и составляет 21% и 22% соответственно.

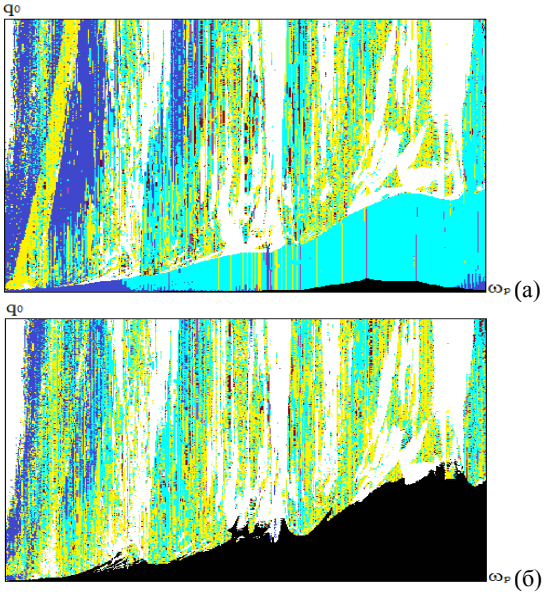


Рис. 3. Карты динамических режимов при $h_k=0.1$ а) карта для балки 1; б) карта для балки 2.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Уменьшение величины зазора (рис. 3) приводит к увеличению площади хаотических колебания для обеих балок до 25.5% и 27.8%. Для балки 2 значительно сокращается зона затухающих колебаний, однако ее контур по-прежнему совпадает с контуром перехода к хаотическим колебаниям балки 1.

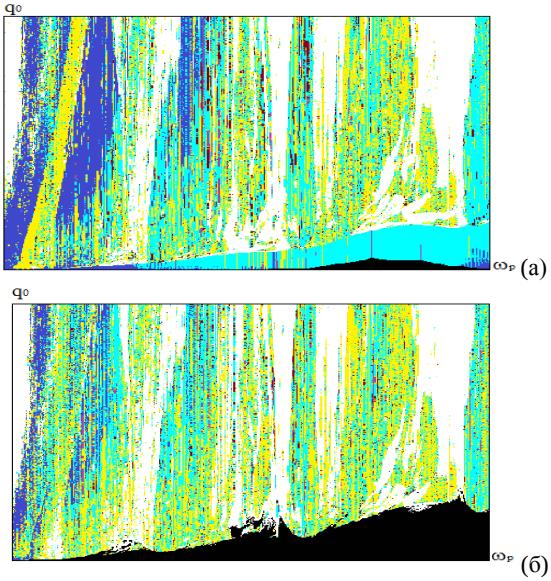


Рис. 4. Карты динамических режимов при $h_k = 0.05$ а) карта для балки 1; б) карта для балки 2.

При зазоре $h_k = 0.05$ доля зон хаотических колебаний максимальна и составляет 26% и 29%. С уменьшением величины зазора между балками увеличилась и зона бифуркационных режимов колебаний. Приведенные на рисунках 2 — 4 карты динамических режимов построены в векторной форме (не растровая графика) и позволяют с точностью до разрешения карты определять значения управляющих параметров, характерных для той или иной зоны колебаний.

5. Заключение

Разработанный программный комплекс позволяет проводить всесторонний анализ нелинейной динамики механических структур в зависимости от управляющих параметров. В качестве управляющих параметров могут рассматриваться граничные и начальные условия, амплитуда и частота вынуждающей нагрузки, тип нагрузки, величина зазора между элементами структуры.

6. Благодарности

Работа выполнена при финансовой поддержке грантов РФФИ № 18-41-700001 р_а, 18-01-00351 а.

7. Литература

[1] Chen Y.-Y., Yu S.-T.J. Constructing a supercomputing framework by using python for hybrid parallelism and GPU cluster // 20th AIAA Computational Fluid Dynamics Conference. 2011.

[2] Grossman A., Morlet S. Decomposition of Hardy functions into square separable wavelets of constant shape // SIAM J. Math. Anal, 1984, vol. 15, № 4, p. 723.

[3] Kantz H. A robust method to estimate the maximum Lyapunov exponent of a time series // Phys. Lett. A, 185, 1994, p. 77-87.

[4] Krysko V.A., Awrejcewicz J., Papkova I.V., Saltykova O.A., Krysko A.V. On reliability of chaotic dynamics of two Euler–Bernoulli beams with a small clearance. // International Journal of Non-Linear Mechanics, 2017.

[5] Papkova I.V., Krysko A.V., Saltykova O.A., Zakharova A.A., Krysko V.A. Contact interaction of flexible Timoshenko beams with small deflections. // Journal of Physics: Conference Series, Vol. 944, No. 1, January 2018, p. 012087.

[6] Rosenstein M.T., Collins J.J., Carlo J. De Luca A practical method for calculating largest Lyapunov exponents from small data sets // Neuro Muscular Research Center and Department of Biomedical Engineering, Boston University, November 20, 1992.

[7] Süli E., Mayers D. An Introduction to Numerical Analysis // Cambridge: Cambridge University Press, 2003.

[8] Tanabe M., Wakui H., Matsumoto N., (...), Sogabe M., Komiya S. Computational model of a Shinkansen train running on the railway structure and the industrial applications // Journal of Materials Processing Technology 140(1-3 SPEC.), 2003, p. 705-710.

[9] Timoshenko S.P. On the correction for shear of differential equation for transverse vibration of prismatic bar // Philosophical Magazin. 41, 36, 1921, p. 744-746.

[10] Wilson J.R., Gramoll K.C. Viscous fluid dynamics app for mobile devices using a remote high performance cluster // ASEE Annual Conference and Exposition, Conference Proceedings 122nd ASEE Annual Conference and Exposition: Making Value for Society, 2015.

[11] Wolf A., Swift J.B., Swinney H.L., Vastano J.A. Determining Lyapunov Exponents from a time series // Physica 16 D, 1985, p. 285-317.

[12] Кантор Б.Я. Контактные задачи нелинейной теории оболочек вращения, Киев, Наукова Думка, 1990.

[13] Шереметьев М.П., Пелех Б.Л. К построению уточненной теории пластин // Инж.журн., т. 4, вып. 3., 1964, с. 34–41.

Об авторах

Салтыкова Ольга Александровна, к.ф.-м.н., доцент кафедры Математика и моделирование Саратовского государственного технического университета; с.н.с. лаборатории 3D-моделирования и промышленного дизайна кафедры инженерной графики и промышленного дизайна Института кибернетики Национального исследовательского Томского политехнического университета. Ее e-mail olga_a_saltykova@mail.ru

Папкова Ирина Владиславовна, к.ф.-м.н., доцент кафедры Математика и моделирование Саратовского государственного технического университета. Ее e-mail Ikravzova@mail.ru.

Афонин Олег Александрович, к.ф.-м.н., доцент кафедры Математика и моделирование Саратовского государственного технического университета имени Гагарина Ю.А. Его e-mail: oaafonin@gmail.ru

Крысько Антон Вадимович, д.ф.-м.н., профессор кафедры Прикладной математики и системного анализа Саратовского государственного технического университета, инженер лаборатории 3D-моделирования и промышленного дизайна кафедры инженерной графики и промышленного дизайна Института кибернетики Национального исследовательского Томского политехнического университета Его e-mail anton.krysko@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Методы реставрации потерянных и искаженных участков изображений

А.О. Трубаков[1], Т.Д. Праздникова[1]
trubakovao@gmail.com|tanya.dmitrievna.2016@bk.ru
[1] Брянский государственный технический университете, Брянск, Россия

В данной статье демонстрируется сравнение методов восстановления искаженных участков изображения, проводится оценка эффективности данных методов, а также описывается модифицированный метод на основе минимизации энтропии ДКП и сравнение его результатов с остальными методами.

Ключевые слова: *реставрация участков изображения, метод минимизации энтропии, ДКП, ПССШ, изотропная диффузия, косинусное преобразование, восстановление, реконструкция.*

Restoration of distorted image areas

A.O. Trubakov[1], T.D. Prazdnikova[1]
trubakovao@mail.ru|tanya.dmitrievna.2016@bk.ru
[1]Bryansk State Technical University, Bryansk, Russia

This article demonstrates comparison of methods for recovered image areas that were distorted, and spends the evaluation of the effectiveness these methods, and describes the modified method that based on minimization the entropy of DCT and compare its results with other methods.

Keywords: *restoration of image areas, the method of minimization of entropy, DCT, isotropic diffusion, cosine transformation.*

1. Введение

В последнее время заметно вырос интерес к корректировке изображений, их ретушированию и реставрации. То, что недавно казалось невозможным, набирает популярность с довольно высокой скоростью. Часто в кадр попадает объект, который сразу бросается в глаза и портит картину, например, морщинки на прекрасном женском лице, выдающие ее возраст или, к примеру, мусор на красивом морском побережье или частицы пыли, появившиеся в результате сканирования изображения, или еще более серьезные проблемы, такие как восстановление рваных и мятых фотографий с потерянными участками. Такие объекты стараются убрать, заретушировать, а для этого необходимы качественные математические расчеты, для определения наиболее подходящих пикселей на основе доступных элементов изображения. На данный момент идеально восстановить пиксели не может ни один из методов, но со временем каждый из них модифицируется, добавляются новые подходы, обладающие большим качеством по сравнению с предыдущими.

2. Существующие подходы к восстановлению искаженных участков изображений

В данной работе для анализа были выбраны три наиболее популярных метода реставрации изображений:

- метод на основе минимизации энтропии ДКП [1];
- метод на основе синтеза текстуры [2];
- метод на основе изотропной диффузионной модели [3].

Для того чтобы оценить работу методов было выбрано полутоновое изображение Barbara (рис. 1.а), которое является одним из стандартных снимков, используемых для тестирования алгоритмов обработки и анализа изображений. Данная фотография впервые была использована в базе «Marco Schmidt's standard test images database». Привлекательность этого изображения состоит во множественных линиях, которые проверяют свойства сглаживающих алгоритмов, изменения размера, избавления от шума и алгоритмов сверхвысокого разрешения.

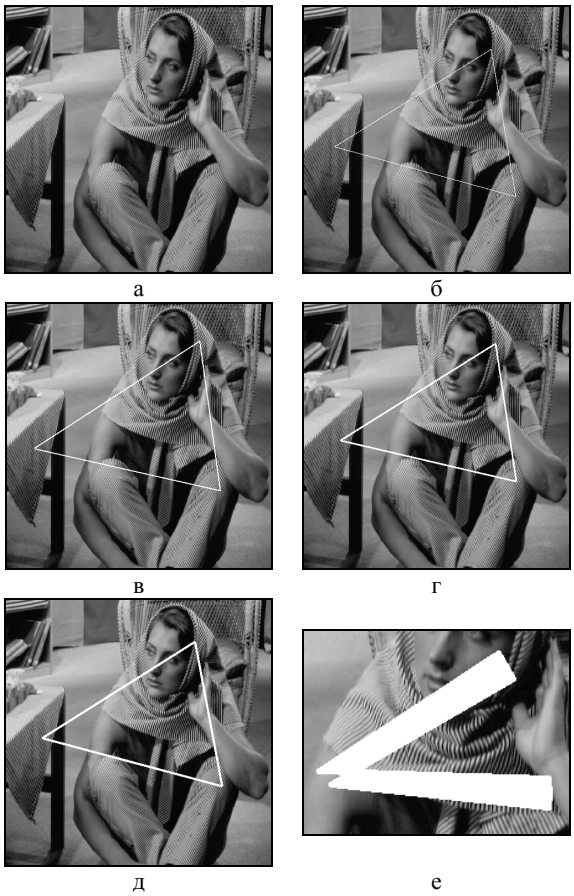


Рис. 1. Тестовое изображение Barbara:
а – исходное изображение; б – 0,4% повреждения; в – 0,8% повреждения; г – 1,2% повреждения; д – 1,6% повреждения; е - 46% повреждения.

Для проведения экспериментов на фотографию были наложены повреждения разной степени сложности в виде белых линий. Данные линии располагаются таким образом, чтобы зона повреждения включала в себя однородные,

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

текстурные, шумоподобные участки, районы перепадов, а также мелкие детали изображения.

Далее полученные тестовые изображения были восстановлены каждым из трех методов.

3. Оценка эффективности методов восстановления поврежденных участков изображения

Сравнение методов проводилось на основе визуальных показателей, а также показателя пикового соотношения сигнал/шум (ПССШ) между истинными и восстановленными значениями пикселей изображений [4,5].

Данная величина определяется через нахождение среднеквадратичной ошибки. Это показатель для двух монохромных изображений, одно из которых считается зашумленным приближением другого. Чем больше значение ПССШ, тем лучше качество восстановления и эффективнее метод.

Для начала рассматривалась визуальная оценка восстановления поврежденных участков. С помощью неё можно сделать вывод о том, что для первых двух поврежденных изображений (малые повреждения) метод минимизации энтропии достаточно точно восстановил искаженные пиксели, а после реконструкции остальными методами зона повреждения осталась слегка заметной. Относительно следующих двух изображений (средние повреждения), наблюдается значительное ухудшение восстановления методом минимизации энтропии, в то время как остальные методы справляются примерно так же, как и ранее. Процесс реконструкции последнего поврежденного изображения (очень большие повреждения) происходит последовательно от краев потерянной области к центру. По результатам восстановления можно сделать вывод о том, что поврежденный участок слабо восстановился, следовательно данные методы совершенно не подходят для восстановления особо крупных областей.

Результат вычислений оценки с помощью ПССШ представлен в таблице 1.

Таблица 1. Результаты реконструкции поврежденных изображений различными методами

Изображение	Процент повреждения, %	ПССШ, дБ		
		Метод минимизации энтропии	Метод синтеза текстуры	Метод диффузионной модели
б	0.4	29.5806	21.1400	21.9957
В	0,8	27.4575	22.1025	21.8328
Г	1,2	22.1818	21.5498	23.2900
д	1,6	18.2451	21.3076	22.0731
е	46	5.1990	6.8912	8.1896

По данным таблицы заметно, что для первых двух изображений наивысший показатель у метода минимизации энтропии, далее он значительно ухудшается, и лидирующим становится метод диффузионной модели.

4. Модифицированный метод на основе минимизации энтропии ДКП

1.1. Описание метода

Данный метод представляет собой усовершенствованный метод на основе минимизации энтропии ДКП. Он состоит в выполнении следующих шагов.

1. На первом этапе определяется маска изображения. Далее выбирается размер блока подсчета, внутри которого будет осуществляться поиск истинных значений для каждого пикселя. В данном случае блок будет равен 8x8. После этого по маске производится поиск поврежденных областей.

2. На основе окрестностей, осуществляется изотропная диффузия [9] внутри поврежденного участка. Сперва изображение фильтруется блоком A, затем блоком D (4).

$$A = \begin{bmatrix} a & b & a \\ b & 0 & b \\ a & b & a \end{bmatrix} \quad D = \begin{bmatrix} c & c & c \\ c & 0 & c \\ c & c & c \end{bmatrix}, \tag{4}$$

где $a = 0.073235$, $b = 0.176765$, $c = 0.125$.

Для определения оптимального количества повторений метода на основе диффузионной модели, был проведен ряд испытаний с коллекцией тестовых изображений, которые содержали от 0.4% до 46% поврежденных пикселей. На рисунке 2.9 изображен график зависимости показателя ПССШ от количества повторений метода.

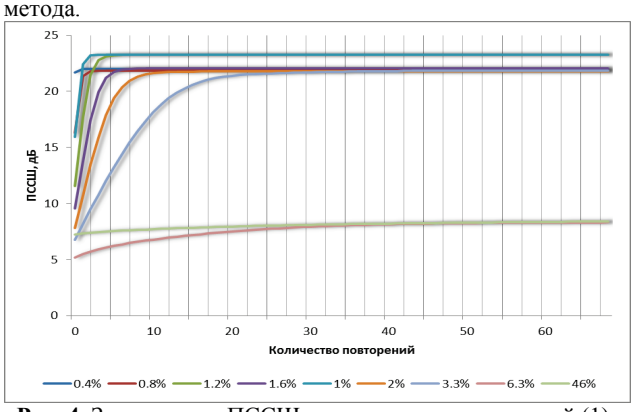


Рис. 4. Зависимость ПССШ от количества повторений (1)

По результатам испытаний было выявлено, что оптимальное количество повторений находится в диапазоне от 10 до 40 раз в зависимости от размера зоны повреждения. Большее количество повторений практически не улучшает результат.

3. На третьем шаге по маске последовательно выбирается каждый искаженный пиксель изображения и область поиска вокруг него (рис. 2.7).

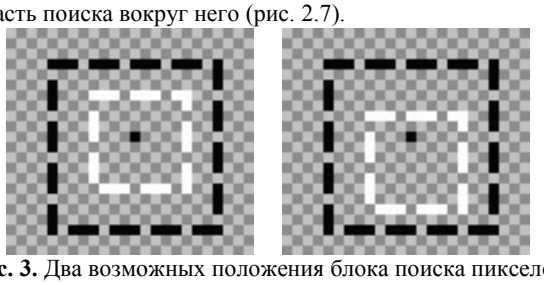


Рис. 3. Два возможных положения блока поиска пикселей (белая рамка) внутри области поиска (черная рамка) для оценки значения утерянного пикселя (черная точка)

4. Далее искомому пикселю присваивается минимально возможное значение.

5. Затем для выбранного блока по формуле (5) вычисляются коэффициенты ДКП [2]. В таблице 2 представлена матрица коэффициентов косинусного преобразования размером 8×8 [1].

$$DCT = D * A * D^T, \tag{5}$$

$$D_{pq} = \begin{Bmatrix} \frac{1}{\sqrt{M}}, & i = 0; \\ \sqrt{\frac{2}{M}} \cos\left(\frac{\pi i(2j+1)}{2M}\right), & i > 0. \end{Bmatrix}, \tag{6}$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

где D – матрица косинусного преобразования размером $M \times M$;
D^T – транспонированная матрица преобразования;
A – матрица пикселей изображения.

Таблица 2. Матрица коэффициентов косинусного преобразования размером 8x8

0.3535	0.3535	0.3535	0.3535	0.3535	0.3535	0.3535	0.3535
0.4903	0.4158	0.2779	0.0978	-0.0971	-0.2773	-0.4153	-0.4902
0.4619	0.1916	-0.1908	-0.4616	-0.4622	-0.1923	0.1901	0.4613
0.4148	-0.0971	-0.4902	-0.2786	0.2766	0.4907	0.0994	-0.4144
0.3536	-0.3531	-0.3542	0.3525	0.3548	-0.3520	-0.3553	0.3514
0.2779	-0.4902	0.0963	0.4167	-0.4144	-0.1002	0.4910	-0.2746
0.1916	-0.4622	0.4613	-0.1894	-0.1938	0.4631	-0.4604	0.1871
0.0978	-0.2786	0.4167	-0.4908	0.4897	-0.4135	0.2740	-0.0924

6. На шестом шаге по формуле (7) вычисляется значение энтропии для каждого возможного значения пикселя [3].

$$E = 3NM + \sum_{i=1}^{N}\sum_{j=1}^{M}\sqrt{|X_{ij}|}, \quad (7)$$

где X_{ij} – значение коэффициента ДКТ блока;
N и M – размеры блока изображения.

После этого значение пикселя увеличивается на единицу и шаг 6 повторяется до тех пор, пока не будет вычислена энтропия для всех возможных значений пикселя. Далее из найденных значений выбирается то, с которым энтропия минимальна.

7. Шаги 5-6 повторяются для каждого положения блока поиска. Затем результаты усредняются и полученное значение присваивается искомому пикселю. Процесс повторяется для каждого поврежденного пикселя.

8. Для определения оптимального количества повторений шагов 3-7, был проведен ряд испытаний с коллекцией тестовых изображений, которые содержали от 0.4% до 46% поврежденных пикселей. На рисунке 2.10 изображен график зависимости показателя ПССШ от количества повторений метода.

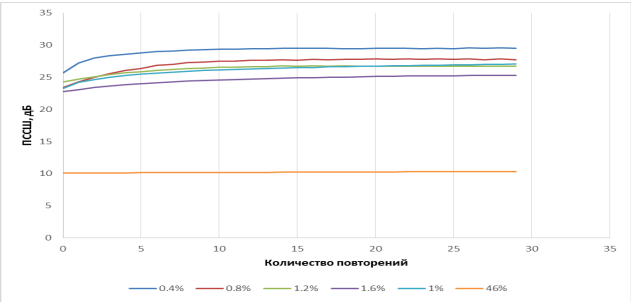


Рис. 4. Зависимость ПССШ от количества повторений (2)

По результатам испытаний было выявлено, что количество повторений зависит от размера области повреждения и вычисляется по формуле (8).

$$P = N * 10, \quad (8)$$

где N – толщина поврежденного участка.

В среднем достаточно 30 повторений для получения максимально возможного результата или близкого к нему.

1.2. Демонстрация

Для демонстрации метода была выполнена обработка всех тестовых изображений с разным процентом повреждения предложенным методом. Итоги обработки представлены на рис. 5.

Как и предполагалось, данный метод отлично справился с восстановлением потерянных участков и показал лучшие визуальные результаты по сравнению с остальными методами, проанализированными ранее.

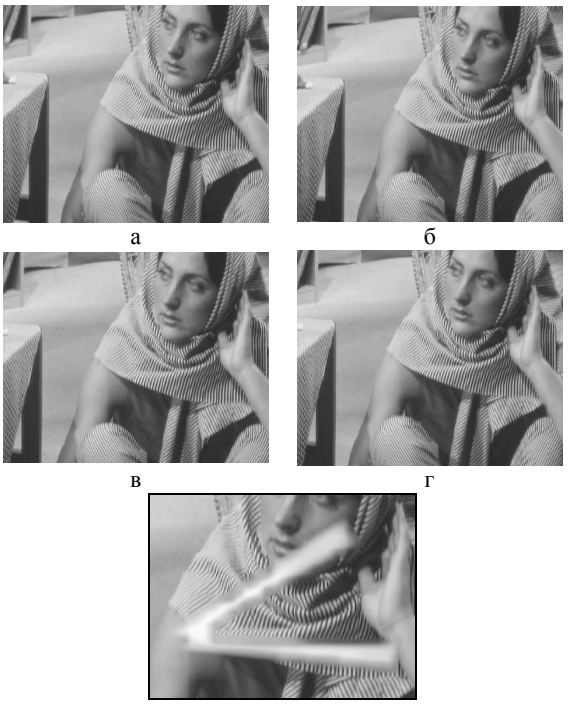


Рис. 5. Восстановленное изображение Barbara:
а – 0,4% повреждения; б – 0,8% повреждения; в – 1,2% повреждения; г – 1,6% повреждения; д - 46% повреждения.

1.3. Сравнение эффективности модифицированного метода со стандартными методами восстановления

В завершении работы было проведено сравнение предложенного метода со стандартными методами восстановления на основе значения ПССШ. Результат вычислений представлен в таблице 3.

Таблица 3. Результаты реконструкции поврежденных изображений различными методами с добавлением предложенного метода

Изображение	Процент повреждения, %	ПССШ, дБ			
		Метод минимизации энтропии	Метод синтеза текстуры	Метод диффузной модели	Предложенный метод
б	0.4	29.5806	21.1400	21.9957	29.5981
в	0.8	27.4575	22.1025	21.8328	27.8876
г	1.2	22.1818	21.5498	23.2900	26.8148
д	1.6	18.2451	21.3076	22.0731	25.2899
е	46	5.1990	6.8912	8.1896	9.8673

Как и следовало ожидать, при незначительных поврежденных областях предложенный метод имеет практически такие же результаты, как и оригинальный вариант на основе минимизации энтропии ДКП. Это объяснимо, потому что основные алгоритмические приемы предложенного метода базируются на методе минимизации энтропии ДКП. Как и свой предшественник, он достаточно хорошо справился с тестовыми изображениями и показал наилучшие результаты (по сравнению с методом синтеза текстур и методом диффузионной модели).

Однако стоит заметить, что при увеличении объема поврежденных участков изображения, оригинальный метод минимизации энтропии ДКП очень сильно теряет в

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

качестве восстановления, что ограничивает его применение на практике в подобных ситуациях. Уже при повреждении в 1.5-2% процента качества восстановления становится сравнимым с двумя другими подходами. Предложенный же метод характеризуется тем, что качество его восстановления не так сильно падает с увеличением искаженных областей. При этом оно является максимальным среди всех исследованных алгоритмов, что подтверждает эффективность данного метода. В отличии от своего предшественника, предложенный метод может применяться в более широких диапазонах искажения, что расширяет рамки его использования на практике.

Из приведенной выше таблицы видно, что в действительности модифицированный метод минимизации энтропии ДКП обладает наивысшими показателями для всех протестированных изображений.

5. Заключение

В данной работе были продемонстрированы результаты восстановления искаженных пикселей для пяти тестовых изображений каждым из трех исследуемых методов. Также было проведено сравнение показателей ПССШ для реконструированных изображений. Полученные результаты показали, что для близкого к идеальному восстановлению мелких повреждений толщиной в 1-2 пикселя отлично подходит метод на основе минимизации энтропии ДКП. Что касается более серьезных искажений, для них лучшие результаты демонстрирует метод на основе изотропной диффузионной модели. Повреждения остаются заметны, но в меньшей степени.

Исходя из этого, был разработан новый модифицированный метод, который показывает отличные результаты.

6. Литература

[1] Марчу, В.И. Методы цифровой обработки сигналов для решения прикладных задач / В.И. Марчук, 2012. – М.: Радиотехника, 2012. – 128 с.

[2] Воронин, В.В. Восстановление искаженных пикселей изображений на основе текстурно-геометрической модели / В.В. Воронин, Е.А. Семенищев, В.А. Франц, И.С. Свирин // Фундаментальные исследования. – 2015. – № 12-1. – С. 17-21

[3] Perona, P. Scale-space and edge detection using anisotropic diffusion / P. Perona, J. Malik // IEEE Transactions on Pattern Analysis and Machine Intelligence 12(7), 1990. – PP. 629-639.

[4] Гаврилов, Н.И. Оценка качества визуализации медицинских и научных данных в величинах отношения пикового сигнала к шуму / Н.И. Гаврилов, В.Е. Турлапов // Вестник Нижегородского университета им. Н.И.Лобачевского, 2012. – №5. – С.266-275.

[5] Wang Z., Bovik A.C. Modern image quality assessment // Synthesis lectures on image, video & multimedia processing. Morgan & Claypool, 2006.

Об авторах

Трубаков Андрей Олегович, к.т.н., доцент кафедры «Информатика и программное обеспечение» Брянского государственного технического университета. Его e-mail trubakovao@gmail.com.

Праздникова Татьяна Дмитриевна, магистрант кафедры «Информатика и программное обеспечение» Брянского государственного технического университета. Ее e-mail tanya.dmitrievna.2016@bk.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Об СКО, методе *K*-средних и деревьях в приложении к обработке изображений

М.В. Харинов
khar@iias.spb.su
Санкт-Петербургский Институт информатики и автоматизации Российской академии наук (СПИИРАН),
Санкт-Петербург, Россия

В статье обсуждаются сложившиеся в обработке изображений стереотипы, которые препятствуют активному применению кластерного анализа в задачах первоначальной локализации и детектирования объектов. Рассматриваются методы получения и оптимизации по СКО иерархической последовательности кусочно-постоянных приближений изображения. Предлагается новый способ кодирования и визуализации иерархии разбиений изображения посредством «стереопары рейтинга объектов». Описываются особенности скоростных вычислений в реальном времени.

Ключевые слова: *среднеквадратичное отклонение, метод Уорда, метод К-средних, динамические деревья, пиксели, цветовое изображение.*

On utilization of standard deviation, K-means method and trees in image processing

M.V. Kharinov
khar@iias.spb.su
St.Peterburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS),
St.Petersburg, Russia

The article discusses the stereotypes of image processing, which hinder the active application of cluster analysis for preliminary object localization and detection. Generating and optimizing of a hierarchical sequence of piecewise constant image approximations are considered. The methods of generating and optimizing by standard deviation of hierarchical sequence of piecewise constant image approximations are treated. A new coding and visualizing way of a hierarchy of image partitions by means of so called «object rating stereopair» is proposed. The peculiarities of high-speed computing in real-time are discussed.

Keywords: *standard deviation, Ward's clustering, K-means method, dynamic trees, pixels, color image.*

1. Введение

В недавнем прошлом линейка, циркуль и арифмометр являлись рабочими инструментами практических вычислений. В расчете на применение этих инструментов был разработан метод *K*-средних [8], который в настоящее время с переменным успехом активно применяют для кластеризации пикселей изображения с помощью компьютера [5].

Метод сводится к *реклассификации* (перенесению) пикселей или, лучше, множеств пикселей из одного кластера в другой до тех пор, пока их средние яркостные значения не окажутся ближайшими к трехкомпонентным средним яркостям k кластеров пикселей. Декларируется, что при этом обеспечивается минимизация СКО (среднеквадратичного отклонения) σ или ошибки аппроксимации (суммарной квадратичной ошибки) $E = 3N\sigma^2$ изображения из N пикселей своим кусочно-постоянным приближением средними значениями пикселей, усредненными внутри k кластеров. Утверждение о минимизации E, если не ошибочно, то, по крайней мере, не точно, т.к. не следует из условия минимизации декларируемого функционала. Если при реклассификации множеств пикселей в методе *K*-средних критерий максимальной близости к центру кластера заменить условием минимизации E или σ, то получим более сильный метод минимизации E, названный методом «*K*-средних без средних» или *K*-meanless методом [3], который улучшает результаты метода *K*-средних без увеличения вычислительной сложности.

По поводу СКО в современных учебниках по обработке изображений часто утверждается, что оно “плохо согласуется с критерием субъективного восприятия” [6]. Однако следует иметь в виду, что это утверждение касается сравнения пары изображений, значения пикселей которых не связаны строгой зависимостью. Для сравнения изображения со своим кусочно-постоянным *приближением*, значения пикселей которого получаются усреднением пикселей изображения, СКО σ или ошибка аппроксимации E является вполне адекватной оценкой при условии достаточного приближения $E = 3N\sigma^2$ к *оптимальному* значению, минимально возможному при данном значении k. А неадекватная кластеризация пикселей, скорей всего, получается из-за недостаточной минимизации E:

- в устаревшей версии метода *K*-средних, моделирующей на компьютере применение линейки, циркуля и арифмометра;
- ограниченным рассмотрением реклассификации только одиночных пикселей;
- заданием начального разбиения N пикселей изображения на k кластеров с *заведомо большой ошибкой* аппроксимации E, превышающей некоторый установленный порог.

В докладе обсуждается аппроксимация изображения своими *квазиоптимальными приближениями*, генерируемыми с эффективной минимизацией ошибки аппроксимации E для последующей автоматической локализации, детектирования и распознавания объектов сцены.

2. Постановка задачи

Постановка задачи локализации и детектирования объектов сцены подобна постановке задачи вычисления на изображении «наиболее заметных областей» (salient region detection [2, 9]), которая отличается тем, что не использует априорных предположений о конкретном содержании изображения. При этом для изображения строится, так называемая, карта заметных областей, которая обычно преобразуется в черно-белую маску «объект-фон» пороговым преобразованием.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

В нашей модели кластеризации пикселей изображения обобщается и развивается постановка задачи [2,9].

На выходе кластеризации поддерживается бинарная иерархия кластеров, сформированная в алгоритме итеративного слияния множеств пикселей. Выходная иерархия кластеров содержит N приближений изображения. Приближения содержат всего $2N-1$ различных кластеров пикселей. Из них N кластеров неделимы, т к. состоят из отдельных пикселей, а для каждого из $N-1$ остальных кластеров поддерживается операция *разделения надвое*, при которой запоминается и, по мере необходимости, восстанавливается пара кластеров пикселей 1 и 2, слиянием которых данный кластер 3 получен: $3 \to 1,2 : 3 = 1 \cup 2$.

Иерархическая кластеризация пикселей, в частности, сегментация изображения, выполняется для последующей фильтрации и формирования «объектов интереса» из множества «объектов», детектируемых компьютером в виде кластеров пикселей приближений изображения. Пиксели приближения, относимые к различным кластерам, при актуальных, достаточно малых значениях числа кластеров различаются по цвету.

На этапе кластеризации пикселей, выполняемой для последующей фильтрации объектов, возникает проблема упорядочения кластеров пикселей (цветов) по некоторому признаку «заметности», «контрастности» [2,9] или, в терминологии [10], «сложности» фрагмента изображения. Фактически, требуется ввести своего рода *меру яркостной неоднородности* пикселей на участках изображения, обеспечивающую локализацию и детектирование иерархически структурированных объектов и при этом являющуюся таким же предсказуемым параметром, как число n пикселей в кластере, которое монотонно уменьшается по мере разделения данного кластера $1 \cup 2$ на составные части 1 и 2: $n(1 \cup 2) > n(1), n(2)$.

3. Минимизация СКО

Для решения сформулированной задачи необходимо эффективно вычислять *квазиоптимальные* приближения изображения, которые аппроксимируют оптимальные приближения с установленной точностью.

Известно, что последовательность оптимальных приближений в зависимости от числа g кластеров пикселей в приближении описывается монотонно возрастающей последовательностью $\Delta E_2 \le \Delta E_3 \le ... \le \Delta E_{N-1} \le 0$ неположительных приращений E или *выпуклой* последовательностью самих значений E:

$$E_g \le \frac{E_{g-1} + E_{g+1}}{2}, \quad g = 2, 3, ..., N-1,$$

где g — число кластеров, или цветов в приближении изображения.

В общем случае последовательность оптимальных приближений не является иерархической, в чем нетрудно убедиться на примере изображения с линейным возрастанием яркости по одной из координат. Для вычисления оптимальных приближений цветового изображения с разумной вычислительной сложностью, а также для их запоминания и преобразования пока не создано эффективных алгоритмов. Поэтому в развиваемой модели квазиоптимальной кластеризации пикселей [17] решается задача аппроксимации последовательности оптимальных приближений иерархической последовательностью приближений с числом g кластеров пикселей от 1 до N. Полагается что:

1. В целевой иерархии квазиоптимальных приближений любая тройка вложенных друг в друга кластеров описывается выпуклой последовательностью трех значений ошибки аппроксимации — так же, как это имеет место в случае неиерархической последовательности оптимальных приближений.
2. При некотором числе кластеров g_0, задаваемом в диапазоне от 2 до $N-1$, иерархия квазиоптимальных приближений максимально приближается по E к последовательности оптимальных приближений.

Здесь число кластеров g_0 вводится как модельный параметр. Полагается, что при $g = g_0$ изображение разделяется на *целостные* объекты, части которых не встречаются при $g < g_0$, а объединения не встречаются при $g > g_0$. Эти соображения позволяют выбирать параметр g_0 для наилучшего детектирования конкретных «объектов интереса».

Формально, требования к минимизации E при аппроксимации изображения приближением с g_0 кластерами пикселей выражаются тремя утверждениями 1)-3):

1) итеративное разделение каждого кластера надвое в зависимости от числа кластеров g описывается выпуклой последовательностью значений E;
2) ошибку аппроксимации E при данном числе g_0 кластеров пикселей (цветов в приближении изображения) нельзя уменьшить посредством разделения одного кластера надвое и слияния двух других кластеров;
3) при установленном параметре $g = g_0$ ошибку аппроксимации E нельзя уменьшить посредством реклассификации (переноса) той или иной предусмотренной части из одного кластера в другой кластер пикселей.

Здесь условие 1) непосредственно следует из условия 1 предыдущего списка и в расчетах обеспечивается итеративной кластеризацией пикселей по методу Уорда [1,7]. Условие 2) обеспечивается применением версии *SI* (Segmentation Improvement)-метода [12] снижения ошибки аппроксимации E. Условие 3) обеспечивается минимизацией E или σ методом «*K*-средних без средних» [3,14].

Выполнение условий 1) и 2) обеспечивается применением оригинальной версии Уорда [7] или, в ускоренном варианте, поочередным применением *SI*-метода в сочетании с обработкой методом Уорда изображения *по частям*, т.е. по кластерам пикселей, обрабатываемым как самостоятельные изображения. При выполненных условиях 1) и 2) иерархия приближений с числом кластеров g от g_0 до N методом Уорда достраивается до полной иерархии приближений с g от 1 до N, описываемой выпуклой последовательностью N значений ошибки аппроксимации E. При этом благодаря условию 2) не возникает нарушения выпуклости при g_0 кластерах, и при $g = g_0$ достигается значение ошибки аппроксимации E_g, которое не превышает определенного порога:

$$E_{g_0} \le E_1 \left(1 - \frac{g_0}{N}\right),$$

где E_1 ошибка аппроксимации изображения одинаковыми пикселями средней яркости. Таким образом гарантируется исправление грубых начальных приближений изображения. Условие 3) при выполненных условиях 1) и 2) обеспечивает дальнейшую минимизацию E_g при $g = g_0$ кластерах.

В целом, условия 1)-3) обеспечивают аппроксимацию оптимальных приближений *квазиоптимальными*, и аппроксимацию выпуклой последовательности оптимальных значений $E_g^{\min}$ такой же выпуклой последовательностью значений E_g, с минимизированной разницей $E_g - E_g^{\min} \approx 0$ при $g = g_0$ кластерах.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

4. Метод Уорда

Апостериори любую выпуклую последовательность $E_{g=1}$, $E_{g=2}$, ..., $E_{g=N}$, можно получить укрупнением пикселей методом Уорда, в котором вначале каждый пиксель составляет самостоятельный кластер. Затем на каждом шаге сливается пара кластеров i, j, отвечающая минимальному приращению ошибки аппроксимации $\Delta E_{merge}(i, j)$:

$$i, j \to i \cup j: \quad i, j = \underset{i,j=0,2,...g-1}{\arg\min} \ \Delta E_{merge}(i, j),$$

где число кластеров снижается от N до 1, а приращение ошибки аппроксимации $\Delta E_{merge}(i, j)$ выражается через число писелей n_i, n_j и трехкомпонентные средние значения I_i, I_j пикселей внутри кластеров i, j в виде:

$$\Delta E_{merge}(i, j) = -\Delta E_{split}(i \cup j) = \frac{n_i n_j}{n_i + n_j} \left\| I_i - I_j \right\|^2 \geq 0,$$

где $\Delta E_{split}(i \cup j) \leq 0$ неположительная величина приращения ошибки аппроксимации при разделении надвое кластера $i \cup j$.

Если каждому кластеру $i \cup j$, из двух и более пикселей сопоставляется падение $\left|\Delta E_{split}(i \cup j)\right|$ ошибки аппроксимации E при его *разделении надвое*, т.е. на пару кластеров i, j, слиянием которых он получен, то величина $\left|\Delta E_{split}(i \cup j)\right|$ монотонно убывает по мере разделения кластеров надвое:

$$\left|\Delta E_{split}(i \cup j)\right| \geq \left|\Delta E_{split}(i)\right|, \left|\Delta E_{split}(j)\right|.$$

Поэтому величина $\left|\Delta E_{split}(i \cup j)\right|$ может рассматриваться в качестве искомого параметра, характеризующего меру яркостной неоднородности или сложности фрагмента изображения, который составляют пиксели данного кластера $i \cup j$.

Рис. 1 иллюстрирует результаты обработки изображения «Лена» методом Уорда. Вместо исходного изображения в левом верхнем углу демонстрируется кусочно-постоянное приближение изображения в 16 цветах.

Рис. 1. Верхний ряд: приближение изображения в 16 цветах (слева) и пара рейтинговых представлений объектов (справа). Нижний ряд: начальные составные кластеры пикселей №1-3 (слева-направо).

Справа от изображения на рис. 1 показана стереопара полутоновых представлений изображения, которая задает кластеризацию пикселей изображения на 1, 2,...16 кластеров и называется «стереопарой рейтинга объектов». В нижнем ряду демонстрируются сами кластеры изображения с номерами 1, 2, 3 при общей нумерации # составных кластеров от 0 до $N-2$ в порядке убывания $\left|\Delta E_{split}(\#)\right|$. Одна из двух «половин», на которые разделяется каждый кластер, помечена узором в виде квадратной решетки. Под представлениями изображения приводится сопутствующая информация.

Кодирование иерархии кластеров пикселей в стереопаре рейтинга объектов выполняется в порядке последовательного разделения кластеров надвое и сводится к записи номера итерации по координатам одной «половины» кластера в первом представлении стереопары и другой «половины» во втором представлении. При этом альтернативные

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

«половины» кластеров в каждом представлении стереопары сохраняют обозначения материнского кластера.

Для считывания координат пикселей минимального кластера с максимальным номером # достаточно объединить списки координат максимальной яркости для каждого представления стереопары. При этом замена максимальных значений пикселей на номера материнского кластера в каждом представлении обеспечивают редукцию стереопары рейтинга объектов и извлечение кластера с номером #–1 .

Стереопара рейтинга объектов позволяет задать иерархию $N-1$ составных кластеров пикселей в виде двух полутоновых изображений с одинаковым числом $N-1$ тонов яркости. Наглядную визуализацию иерархической кластеризации рис. 1 полезно использовать при настройке локализации и детектирования объектов, например, при установлении диапазона рейтинговых номеров # , обеспечивающих проявление на изображении «объектов интереса» при достаточно малой доле случайных искажений.

5. Метод *K*-средних без средних

В силу выраженной повторяемости пар кластеров i, j, отвечающих минимальным значениям $\Delta E_{merge}(i, j)$, на начальных шагах слияния кластеров, результаты итеративного слияния кластеров методом Уорда зависят от выбора порядка слияния и описывается множеством переплетающихся выпуклых последовательностей значений ошибки аппроксимации $E_{g=1}$, $E_{g=2}$, ..., $E_{g=N}$ [11] (рис.2).

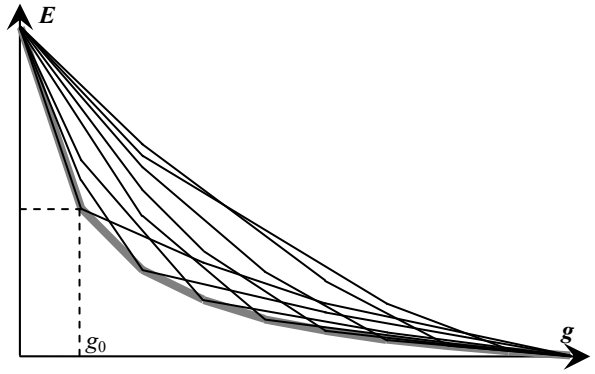


Рис.2. Неоднозначная аппроксимация последовательности оптимальных приближений (граничная серая кривая) выпуклыми последовательностями квазиоптимальных приближений (верхние кривые).

Выбор иерархии и отвечающей ей кривой $E_{g=1}$, $E_{g=2}$, ..., $E_{g=N}$, в точке $g = g_0$ приближающейся к кривой $E^{\min}_{g=1}$, $E^{\min}_{g=2}$, ..., $E^{\min}_{g=N}$, для оптимальных значений, обеспечивается методом *K*-средних без средних, в котором для иерархии частей j кластеров i минимизируется (если достигается) отрицательное приращение $\Delta E_{\text{correct}}(i, j, k)$ ошибки аппроксимации:

$$i, j, k = \underset{i,j=1,2,\ldots,g_0,\, k=1,\ldots,2n_i-1}{\arg\min} \Delta E_{correct}(i, j, k) < 0 ,$$

где приращение $\Delta E_{correct}(i, j, k)$ ошибки аппроксимации при реклассификации k –й части кластера i в кластер j выражается в виде:

$$\Delta E_{correct}(i, j, k) = \frac{n_j n_k}{n_j + n_k} \left\| I_j - I_k \right\|^2 - \frac{n_i n_k}{n_i - n_k} \left\| I_i - I_k \right\|^2 ,$$

где n_i, n_j, $n_k < n_i$ и I_i, I_j, I_k — значения числа пикселей и трехмерных средних яркостей в кластерах i, j, k, соответственно.

Выписанная формула следует из условия минимизации E или σ. Если в ней опустить дробные коэффициенты, учитывающие количества пикселей n_i, n_j, n_k в кластерах i, j, k, то получится критерий минимизации метода *K*-средних. При расчетах с помощью линейки, циркуля и арифмометра метод *K*-средних позволяет избежать операций с чрезмерно большими числами и трудоемкой операции извлечения квадратного корня на арифмометре. Но в приложении к кластеризации пикселей этот метод приводит к большому числу ложных решений, главным образом из-за нарушения порядка слияния/разделения кластеров, выбираемого из условия минимизации приращения ошибки аппроксимации ΔE на каждом шаге формирования бинарной иерархии приближений изображения.

6. Динамические деревья

Скоростные операции с иерархически структурированными кластерами пикселей изображения поддерживаются в терминах деревьев. При этом удобней вместо традиционных деревьев, дендрограмм и пр. использовать динамические деревья Слейтора-Тарьяна [4,16]. Если в традиционной интерпретации дерева при слиянии множеств пикселей порождается новый узел, то при интерпретации по Слейтору-Тарьяну множества пикселей упорядочиваются в структуре дерева и слияние множеств пикселей описывается установлением дуги между корневыми узлами деревьев.

Рис. 3 поясняет различие в интерпретации деревьев на примере изображения из четырех пикселей.

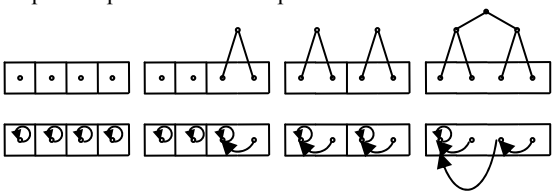


Рис.3. Формирование бинарной иерархии кластеров пикселей в терминах традиционных деревьев (вверху) и динамических деревьев Слейтора-Тарьяна (внизу).

В отличие от традиционных деревьев при использовании динамических деревьев Слейтора-Тарьяна:

1. Метаданные, описывающие иерархию кластеров пикселей, и сами кластеры задаются на одном и том же множестве координат.
2. Для минимизации E или σ реализуется обобщенный механизм обратимых вычислений, при котором состояние вычислительной системы в любой момент времени восстанавливается не обязательно прежним, и реализуется минимизация E при обратном ходе вычислений.
3. Бинарная иерархия кластеров пикселей, описываемая бинарными традиционными деревьями, задается, вообще говоря, нерегулярными динамическими деревьями Слейтора-Тарьяна.

Динамические деревья Слейтора-Тарьяна позволяют реализовать все возможности традиционных деревьев при минимальных затратах памяти и благодаря первому из перечисленных свойств обеспечивают простейшую реализацию обратимых вычислений. В разработанной структуре данных динамические деревья (ациклические графы) строятся в нескольких видах и дополняются *циклами* (циклическими графами), в сочетании с которыми они образуют *динамическую сеть*, поддерживающую скоростное вычисление, запоминание и преобразование миллионов множеств пикселей в оперативной памяти компьютера.

7. Применение

Реализованный в терминах динамической сети аппарат генерации и оптимизации кусочно-постоянных приближе-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

ний изображения по ошибке аппроксимации E или среднеквадратичному отклонению σ (СКО) обеспечивает:

1. Улучшение любой традиционной сегментации изображения, причем, как по E (или σ), так и по визуальному восприятию [13].
2. Улучшение любой иерархической сегментации изображения по тем же параметрам [15].
3. Синхронное выделение объектов на двух изображениях стереопары или нескольких изображениях, снятых в различном ракурсе и совмещенных в одно изображение [11].

Правда, имеющееся в настоящее время программное обеспечение не позволяет единообразно выделять объекты на стереопарах, обрабатываемых как раздельные изображения [11]. По всей видимости, это связано с разбросом оптимизированных приближений по E (или σ) при начальных значениях g_0 (см. рис. 2), для которых пока программно не реализован метод *K*-средних без средних.

8. Заключение

Типичным оправданием недостаточной минимизации СКО при вычислении кусочно-постоянных приближений изображения служит утверждение об NP-полноте задачи минимизации, которое не вполне уместно при обработке цветовых изображений, т.к., речь обычно идет не о точном, а о приближенном решении оптимизационной задачи. Точное решение задачи обеспечивается мультипороговым методом Оцу только для полутоновых изображений. Но и при этом требуется разбить изображение на ограниченное число 2, 3, 4 и. т.д., скажем, до 10 кластеров, что вполне доступно для современных компьютеров.

Другим возражением при применении методов классического кластерного анализа в обработке изображений является аргумент, что сложность современных изображений многократно возросла и продолжает возрастать. Однако и это возражение недостаточно актуально. С ростом разрешения и возможностей цветопередачи возросло только количество пикселей в изображении. Но до тех пор, пока средства машинного зрения уступают возможностям человека, сложность изображений по количеству объектов меняется незначительно. Для того, чтобы снять проблему возрастания объема данных, достаточно решить задачу адекватного начального приближения изображения ограниченным числом, скажем, десятком или сотней тысяч сегментов с уникальными значениями средней яркости. Эту задачу несложно решить, например, применением метод Уорда по частям изображения или другими доступными методами.

Для эффективного использования мощности современных вычислительных средств обработки цифровых изображения пока не хватает адекватных моделей изображения произвольного содержания. Возможно для создания таких моделей необходимо пересмотреть некоторые стереотипы применения классических методов, что обсуждается в докладе.

9. Литература

[1] Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерности. М.: Финансы и статистика, 1989. — 607 с.

[2] Achanta R., Hemami S., Estrada F., Susstrunk S. Frequency-tuned salient region detection // Computer vision and pattern recognition (CVPR), IEEE conference, 2009. — Pp. 1597-1604.

[3] Dvoenko S. D., 2014. Meanless *k*-means as *k*-meanless clustering with the bi-partial approach. // Proc. of the 12th International Conference on Pattern Recognition and Information (PRIP'2014), Minsk, Belarus, Pp. 50–54.

[4] Nock, R., Nielsen F, Statistical Region Merging. IEEE Trans. Pattern Anal. Mach. Intell. 2004. 26(11), — Pp. 1452–1458.

[5] Jain A.K. Data Clustering: 50 Years Beyond K–Means // Pattern Recognition Letters, Vol. 31, No. 8, 2010. — Pp. 651-666.

[6] Сойфер В.А. (Ред.) Методы компьютерной обработки изображений. М.:Физматлит, 2001. — 784 с. – ISBN 5–9221–0180–3.

[7] Ward J.H., Jr. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 1963. Vol. 58, Issue 301,— Pp. 236-244.

[8] Steinhaus H. Sur la division des corps materiels en parties // Bull. Acad. Polon. Sci., C1. III Vol. IV, 1956. — Pp. 801–804.

[9] Cheng M.M., Mitra N.J., Huang X., Torr P.H., Hu S.M. Global contrast based salient region detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. Т. 37. №. 3. — Pp. 69-582.

[10] Чочиа П.А. Теория и методы обработки видеоинформации на основе двухмасштабной модели изображения / Дис. докт. технич. наук. М.:ИППИ РАН, 2016. 302 с.

[11] Ханыков И.Г., Харинов М.В. Кластеризация пикселей цветового изображения методом Уорда // Компьютерная графика и зрение (Графикон'2017) / Труды 27-й международной конференции по компьютерной графике и зрению: Москва, МГУ им. М. В. Ломоносова, 24 – 28 сентября 2017 г., Пермь, — С. 232–235.

URL: http://graphicon.ru/html/2017/papers/pp232-235.pdf

[12] Харинов М.В., Ханыков И.Г. Оптимизация кусочно-постоянного приближения сегментированного изображения. // Труды СПИИРАН. 2015. Вып. 3(40). — С. 183–202. — ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн).

URL: http://www.proceedings.spiiras.nw.ru/ojs/index.php/sp/article/view/3143/1804

[13] Khanykov, M. V. Kharinov and C. Patel Image segmentation improvement by reversible segment merging // International Conference on Soft Computing and its Engineering Applications (icSoftComp), Changa, Anand, India, IEEE Publisher, 2017, — Pp. 1-8.

[14] Kharinov M.V. Reclassification formula that provides to surpass K–means method // arXiv preprint, arXiv:1209.6204, 28 Sep 2012. — 10 p.

URL: https://arxiv.org/ftp/arxiv/papers/1209/1209.6204.pdf

[15] Харинов М.В. Обратимое слияние структурированных кластеров пикселей // Компьютерная графика и зрение (Графикон'2016) / Труды 26-й международной конференции по компьютерной графике и зрению: Москва, МГУ им. М. В. Ломоносова, 19 – 23 сентября 2016 г., Нижний Новгород, — С. 298–302.

URL: http://graphicon.ru/html/2016/papers/Pages_298-302.pdf

[16] Харинов М.В. Разработка динамических структур данных системы автоматизированного распознавания изображений / руков. В.В. Александров / Автореф. Дис. канд. технич. наук. — С.П. 1993. — 20с.

[17] Харинов М. В. Модель квазиоптимальной иерархической сегментации цветового изображения // Оптический журнал. Т. 82, № 7, 2015. — С. 37-43. — ISSN 0030-4042.

Об авторах

Михаил Харинов – к.т.н., доцент, с.н.с. лаборатории прикладной информатики СПИИРАН.
Его адрес: khar@iias.spb.su

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Встраивание информации в сжатые JPEG-изображения с минимизацией проявления демаскирующих признаков

О.О. Шумская, О.О. Евсютин
shumskaya.oo@gmail.com|eoo@keva.tusur.ru
Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

Представлено исследование информативности признаков в пространственной и частотной областях сжатых JPEG-изображений при стегоанализе. Проведен отбор признаков с применением жадного алгоритма с исключением и получен набор информативных признаков, на основе которых определена целевая функция стеганографического встраивания. Встраивание осуществляется по алгоритму на основе операции замены. Зная, какие признаки несут информацию о наличии вложения в цифровом изображении, можно стремиться к минимизации искажений этих параметров с целью максимизации незаметности встраивания как визуально, так и для стегоанализаторов. Эксперименты показали, что такой подход позволяет при встраивании более чем 4000 бит в полутоновое изображение с разрешением 256×256 пикселей достичь минимальных искажений изображения, держа метрику визуального качества PSNR не ниже 52дБ, и обеспечивая устойчивость перед стегоанализатором.

Ключевые слова: *стеганография, стегоанализ, цифровые изображения, сжатые JPEG-изображения, демаскирующие признаки.*

Information embedding into compressed JPEG-images with minimization of manifestation of the unmasking features

O.O. Shumskaya, O.O. Evsutin
shumskaya.oo@gmail.com|eoo@keva.tusur.ru
Tomsk state university of control systems and radioelectronics, Tomsk, Russia

The research of informational content of features in the spatial domain and the frequency domain of compressed JPEG-images at the stegoanalisys is presented. Selection of features with application of the greedy algorithm with an exception is made and the set of informative features on the basis of which criterion function of steganographic embedding is defined is received. Embedding is carried out by means of the algorithm on the basis of replacement operation. Knowing what features contain information about existence of the secret message in the digital image, it is possible to seek for minimization of distortions of these parameters for the purpose of maximizing obscurity of embedding as visually, and for stegoanalyzer. Experiments have shown that such approach allows when embedding more than 4000 bits into the grayscale image with size of 256×256 pixels to reach the minimum distortions of the image, holding a metrics of visual quality PSNR not lower than 52 dB, and stability before the stegoanalyzer.

Keywords: *steganography, steganalisys, digital images, compressed JPEG-images, unmasking features.*

1. Введение

Сегодня трудно представить какую-либо сферу деятельности человека без использования цифровых изображений. Ежедневно в Интернет загружают миллионы изображений различного характера и содержания. Этот факт позволяет пользователям обмениваться информацией с использованием методов цифровой стеганографии, предназначенных для незаметного встраивания дополнительной информации в цифровые объекты.

Эффективность стеганографического встраивания информации в цифровые изображения оценивается с помощью различных показателей. Основным требованием является обеспечение устойчивости перед стегоанализом, что выражается в статистической неразличимости стегоизображений и изображений, не содержащих вложений. Для обеспечения устойчивости перед стегоанализом при встраивании информации в цифровое изображение необходимо, чтобы встраивание не приводило к проявлению демаскирующих признаков.

Выявление демаскирующих признаков может быть произведено посредством анализа массивов изображений (стегоизображений и изображений, не содержащих вложений). В свою очередь, для построения надежного стеганографического алгоритма необходимо проведение анализа выявленных признаков и выбор наиболее информативных из них.

2. Признаки цифровых изображений, используемые в стегоанализе

В качестве признаков в задаче стегоанализа используются разнообразные статистические характеристики цифровых изображений. Известны методы стегоанализа, основанные на анализе признаков изображения в пространственной области, и методы, основанные на анализе признаков в частотной области. Также встречаются комбинированные методы, использующие признаки и в пространственной, и в частотной областях. В каждом методе стегоанализа значения некоторого набора признаков объединяются в один вектор, с которым далее работает классификатор.

Ключевой задачей является выбор информативных признаков, анализ которых позволяет отделять стегоизображения от чистых изображений.

Кратко рассмотрим наборы признаков, представленные в некоторых современных исследованиях, посвященных стегоанализу JPEG-изображений.

В работе [7] предлагается набор признаков для стегоанализа, основанных на соотношениях между энергией, собранной в отдельных коэффициентах спектра дискретного косинусного преобразования (ДКП):

$$F_1 = \frac{E(f_0)}{E(f_{|\eta|=1})},$$

где $E(f_0)$ – среднее значение частот нулевых АС-коэффициентов изображения по блокам.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

$$F_2 = \frac{\sum_{|\eta|>1} E(f_\eta)}{E(f_{|\eta|=1})},$$

где $E(f_{|\eta|=1})$ – среднее значение частот АС-коэффициентов изображения, по абсолютной величине равных 1.

$$F_3 = \frac{En_{|\eta|>1}}{En_{|\eta|\leq 1})},$$

где $En_{|\eta|>1}$ – энергия АС-коэффициентов изображения, по абсолютной величине больших 1.

Для изображений характерна межблочная корреляция. Во время встраивания вносятся изменения в блоки изображения, что может привести к нарушению связи между блоками:

$$F_4 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

где $\bar{x}$, $\bar{y}$ – средние значения АС-коэффициентов соседних блоков;

x_i, y_i – i-й АС-коэффициент соседних блоков.

Авторы работ [2, 4] используют 23 признака, как в частотной области, так и в пространственной:

1) общая гистограмма ДКП-коэффициентов изображения;
2) гистограммы первых пяти АС-коэффициентов, для каждого отдельно;
3) межблочные зависимости в разных направлениях (признаки в пространственной области);
4) двойные гистограммы АС-коэффициентов, значение которых находится в диапазоне [–5, 5]:

$$F_5, \dots, F_{15} = \frac{\sum_{k=1}^{B} \delta(d, d_k(i,j))}{\left\|\sum_{k=1}^{B} \delta(d, d_k(i,j))\right\|_{L_1}},$$

где d – фиксированное значение коэффициента, $d \in [-5,5]$;

B – количество блоков в изображении;

i, j – координаты положения коэффициента в блоке;

$\delta(u,v) = \begin{cases} 1, u = v, \\ 0, \text{иначе}; \end{cases}$

L_1 норма – максимальная из сумм элементов по столбцам.

Каждая величина рассчитывается дважды: для исследуемого изображения (J_1) и для изображения, которое получают путем обрезания исследуемого изображения сверху и слева на 4 пикселя (J_2). Подобное действие объясняется следующим образом: при обрезании изображения слева и сверху его разделение на блоки сдвигается, и повторное квантование ДКП-коэффициентов уничтожает имеющееся вложение. В результате характеристики нового изображения будут значительно отличаться от характеристик исходного изображения, если последнее содержало вложение. В противном же случае отличия будут незначительны.

Поэтому в качестве конечного значения каждого признака принимается значение функционала:

$$F = \|f(J_1) - f(J_2)\|_{L_1},$$

где L_1 норма – максимальная из сумм элементов по столбцам.

Для стегоанализа можно использовать не только специально формируемые признаки, но и произвольные признаки, используемые в иных задачах обработки цифровых изображений. В связи с этим в исследуемый набор признаков было включено 10 текстурных признаков $F_{16}, \dots, F_{25}$, приведенных в статье [12].

Метод, представленный авторами [1], основан на законе Бенфорда: вероятность появления цифры на первом месте в числе тем выше, чем меньше эта цифра. Основываясь на выводах работы [5], посвященной исследованию справедливости закона Бенфорда в отношении ДКП-коэффициентов JPEG-изображений до и после квантования, авторы предложили частный случай закона Бенфорда, так как квантованные ДКП-коэффициенты не подчиняются строго закону Бенфорда.

В набор признаков были включены признаки на основе закона Бенфорда для каждой из 9 цифр (1), а также признаки, объединяющие суть закона Бенфорда и идею смещения изображения на 4 пикселя из [2, 4] (2):

$$F_{26}, \dots, F_{34} = N \log_{10}\left(1 + \frac{1}{s + x^q}\right), \quad (1)$$

где $x = 1, \dots, 9$;

N, s, q – параметры, зависящие от качества JPEG-сжатия.

$$F_{35}, \dots, F_{43} = \|F_{26}, \dots, F_{34}(J_1) - F_{26}, \dots, F_{34}(J_2)\|_{L_1}. \quad (2)$$

3. Отбор информативных признаков

3.1 Методы исследования

Существует много алгоритмов отбора информативных признаков, основанных на выявлении статистических зависимостей (между элементами набора, между элементами набора и выходным значением), сравнительных экспериментах, сложных вычислениях. Например, жадные алгоритмы поиска, алгоритмы поочередного перебора, генетический алгоритм (ГА).

Жадные алгоритмы поиска часто используются, так как достаточно быстры и дают хороший результат во многих задачах. Группа алгоритмов получила такое название из-за того, что если один из признаков был выбран в поднабор (или исключен), то в дальнейшем он остается в наборе (в случае жадного включения) или навсегда будет отсутствовать (в случае жадного исключения) [3].

В рамках настоящего исследования выбран жадный алгоритм с исключением, так как он является достаточно быстрым и эффективным согласно исследованиям [6, 9], учитывает все элементы набора признаков в совокупности, а не как отдельные составляющие. Применение исключения позволяет сравнить обновленный набор с тем, что был ранее и понять, показывает ли новый набор лучший результат при проведении стегоанализа.

Для классификации изображений использовался наивный байесовский классификатор.

3.2 Построение набора информативных признаков

Для формирования обучающей и тестовой выборок были взяты стандартные изображения (Lena, Peppers, Baboon и т.д.), а также изображения из баз UCID [8] и USC-SIPI ID [10] размером 256×256 пикселей. Обучающая (660 экземпляров) и тестовая (440 экземпляров) выборки содержат изображения:

1) без вложения;
2) с вложением по классическим алгоритмам F5, JSteg, PM1.

Так как в работе применен жадный алгоритм с исключением, то изначально рассматривался полный набор из 43 признаков, описанный ранее.

В табл. 1 частично представлены результаты экспериментов. Так как набор признаков достаточно объемный, то данная таблица включает три наибольших вероятности обнаружения на каждом шаге.

Шаг	Место		
	1	2	3
1	F_{25} (корреляция) 76,32	F_{33} (Бенфорд 8) 76,19	F_{24} (ковариация) 74,6
2	F_{17} (энтропия) 77,78	F_4 (корреляция межблочная) 76,19	F_{18} (однородность) 76,19
3	F_{30} (Бенфорд 5) 77,78	F_{27} (Бенфорд 2) 73,02	F_{19} (контраст) 71,43

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4	F_{19} (контраст) 77,78	F_{26} (Бенфорд 1) 77,78	F_{35} (Бенфорд+ смещение 1) 74,6
5	F_{26} (Бенфорд 1) 79,37	F_{11} (двойная гист-ма 1) 77,78	F_{28} (Бенфорд 3) 76,19
6	F_{24} (ковариация) 79,37	F_5 (двойная гист-ма –5) 74,6	F_{37} (Бенфорд+ смещение 3) 74,6
7	F_{27} (Бенфорд 2) 80,95	F_{14} (двойная гист-ма 4) 79,37	F_{32} (Бенфорд 7) 79,37
8	F_{31} (Бенфорд 6) 84,13	F_{36} (Бенфорд+ смещение 2) 84,13	F_{15} (двойная гист-ма 5) 82,54
9	F_7 (двойная гист-ма –3) 84,13	F_{10} (двойная гист-ма 0) 84,13	F_{13} (двойная гист-ма 3) 84,13
10	F_{35} (Бенфорд+ смещение 1) 84,13	F_8 (двойная гист-ма –5) 82,56	F_{14} (двойная гист-ма 4) 82,56
11	F_{15} (двойная гист-ма 5) 84,13	F_{32} (Бенфорд+ смещение 2) 82,56	F_1 (энергия част. 1) 82,56
12	F_{32} (Бенфорд 7) 84,13	F_{32} (Бенфорд+ смещение 2) 84,13	F_1 (энергия част. 1) 82,54
13	F_{13} (двойная гист-ма 3) 84,13	F_{10} (двойная гист-ма 0) 82,56	F_2 (энергия част. 2) 82,54
14	F_{14} (двойная гист-ма 4) 84,13	F_{40} (Бенфорд+ смещение 6) 84,13	F_{10} (двойная гист-ма 0) 82,54
15	F_6 (двойная гист-ма –4) 84,13	F_1 (энергия част. 1) 82,54	F_{12} (двойная гист-ма 2) 82,54
16	F_{33} (Бенфорд 8) 85,71	F_{40} (Бенфорд+ смещение 6) 85,71	F_3 (энергия част. 3) 84,13
17	F_{34} (Бенфорд 9) 87,3	F_{38} (Бенфорд+ смещение 4) 84,13	F_4 (корреляция межблочная) 82,54
18	F_{42} (Бенфорд+ смещение 8) 85,71	F_5 (двойная гист-ма –5) 84,13	F_{23} (дисперсия по *j*) 84,13

Табл. 1. Отбор информативных признаков

В результате экспериментов был построен следующий набор из 26 информативных признаков:

$\{F_1 - F_5, F_8 - F_{12}, F_{16}, F_{18}, F_{20} - F_{23}, F_{28}, F_{29}, F_{36} - F_{43}\}$.

3.3 Использование построенного набора признаков для повышения незаметности стеганографического встраивания

Построенный набор признаков был использован для улучшения стеганографического алгоритма, полученного в [11] и работающего с JPEG-изображениями. Данный алгоритм встраивает один бит сообщения в один ДКП-коэффициент с помощью операции замены: коэффициент заменяется значением *x*, если встраиваемый бит равен 1, и значением *–x* — в противном случае. Коэффициенты, по абсолютной величине равные *x* и не используемые для встраивания, корректируются, чтобы извлечение было однозначным.

Выбор оптимальных позиций для размещения *n*-битового фрагмента сообщения в ДКП-блоке осуществляется с помощью ГА. За целевую функцию принимается значение метрики PSNR.

В настоящем исследовании в основу целевой функции был положен выход классификатора, осуществляющего стегоанализ изображений–контейнеров:

$$f(x^*) = \max_{n \in N, x \in X}(1 - \mathrm{acc}_{n,x}),$$

где $\mathrm{acc}_{n,x}$ — точность классификации стегоизображений с разрешением *N*×*N* пикселей при встраивании *n*-битовых фрагментов сообщения в ДКП-блоки с величиной замены *x*.

Таким образом, повышение незаметности стеганографического встраивания заключалось в минимизации проявления демаскирующих признаков, которая, в свою, очередь реализовывалась посредством максимизации ошибки классификатора, обученного с использованием данных признаков.

Результаты экспериментов по встраиванию сообщений в JPEG-изображения с минимизацией проявления демаскирующих признаков приведены на рис. 1. В качестве алгоритма оптимизации, как и в [11], был выбран ГА. Параметры ГА были заданы следующим образом: 30 особей, 30 поколений, 40 итераций.

Для оценки визуального качества стегоизображений использовалась стандартная метрика PSNR.

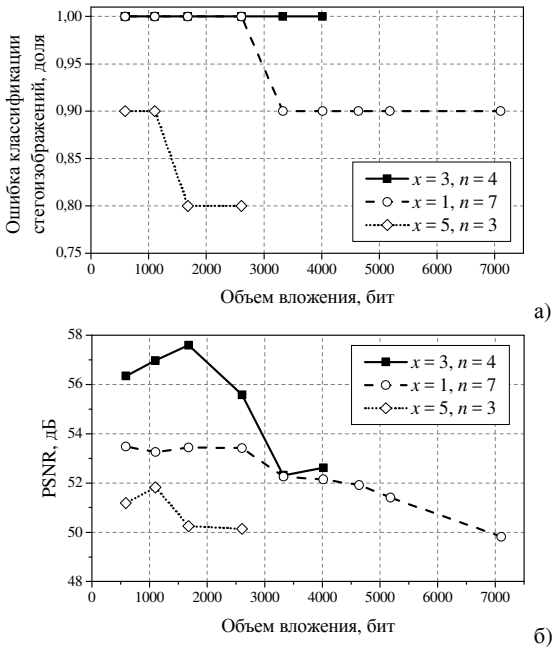


Рис. 1. Результаты встраивания: а) зависимость ошибки классификатора от объема вложения; б) зависимость значения PSNR от объема вложения

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

4. Обсуждение результатов

Эксперименты с жадным алгоритмом показали, что набор из 43 признаков в пространственной и частотной областях можно сократить на 17 элементов, при этом повысив общую точность классификации на 19%.

Анализируя результаты экспериментов, стоит отметить, что среди 17 удаленных признаков находится 7 признаков на основе закона Бенфорда (остались лишь для значений 3 и 4). Признаки на основе совмещения закона Бенфорда с идеей смещения изображения оказались более устойчивыми, из них выбыл всего 1 (для значения 1), что позволяет сделать вывод о большей информативности и большем вкладе модифицированной версии признаков.

Также в течение нескольких первых итераций жадного алгоритма большую точность обнаружения показывали наборы без структурных признаков в пространственной области (корреляция, энтропия, контраст, ковариация). Однако после удаления этих 4 признаков в табл. 1 практически не было замечено присутствие других элементов из данной группы.

Оставшиеся 5 выбывших признаков – двойные гистограммы. Тут стоит обратить внимание на то, для каких значений коэффициентов были эти признаки: –4, –3, 3, 4, 5. Известно, что после ДКП и квантования наибольшее количество элементов принимают значения в диапазоне [–1, 1], и чем дальше значение от границ этого диапазона, тем меньше вероятность встретить его в матрице ДКП-коэффициентов. Поэтому двойные гистограммы для данных значений выбыли из общего набора признаков.

Наборы с исключением первых 4 признаков: 3 энергетических признака в частотной области и межблочная корреляция между блоками с близостью DC-коэффициентов не более 3, – всегда показывали результат ниже среднего, что говорит о важном вкладе этих признаков в информацию о содержании изображения.

Применение результирующего набора из 26 информативных признаков позволило повысить незаметность стеганографического встраивания информации в цифровые изображения с помощью алгоритма, основанного на операции замены квантованных ДКП-коэффициентов.

Как можно увидеть из графика на рис. 1а, наилучший результат в части минимизации проявления демаскирующих признаков достигается при встраивании в один ДКП-блок 4 бит сообщения с использованием величины замены $x = 3$. В этом случае классификатор, обученный на наборе информативных признаков, пропускает все стегоизображения независимо от объема вложения. Уменьшение значения x с одновременным увеличением значения n, либо увеличение x с уменьшением n показывает худший результат: при повышении объема вложения 10–20 % стегоизображений детектируются верно.

Визуальное качество получаемых стегоизображений в сравнении с исходными изображениями было оценено с помощью метрики PSNR. Повышение визуального качества встраивания не было целью оптимизации, поэтому соответствующий график на рис. 1б не показывает явно выраженную зависимость значения PSNR от объема вложения. Однако даже при максимальном объеме вложения PSNR остается на уровне 50 дБ, что говорит об отсутствии заметных искажений на стегоизображении. При этом наилучший результат по качеству встраивания также обеспечивают значения параметров встраивания $x = 3$, $n = 4$.

Дальнейшего улучшения можно добиться, если усложнить задачу оптимизации и встраивать сообщение в ДКП-блоки изображения-контейнера фрагментами переменного размера.

5. Заключение

В настоящей работе проведено исследование информативности признаков сжатых JPEG-изображений, используемых в задаче стегоанализа. Сформирован набор, включающий признаки, представленные в литературе, и их модификации. Для проведения стегоанализа использован наивный байесовский классификатор. С помощью жадного алгоритма мощность набора уменьшена с 43 признаков до 26 с одновременным повышением точности классификации.

Сформированный набор информативных признаков использован для решения обратной задачи: повышение устойчивости стеганографического встраивания перед стегоанализом за счет минимизации проявления демаскирующих признаков. В основу целевой функции положен выход классификатора, осуществляющего стегоанализ изображений–контейнеров с использованием выбранных признаков.

Для экспериментов взят алгоритм встраивания, основанный на операции замены квантованных ДКП-коэффициентов. Минимизация проявления демаскирующих признаков проведена с помощью генетического алгоритма. В результате при встраивании более чем 4000 бит в полутоновое изображение с разрешением 256×256 пикселей удалось достичь минимальных искажений изображения, держа метрику визуального качества PSNR не ниже 52дБ.

Полученные результаты могут быть использованы для построения каналов скрытой передачи конфиденциальной информации с целью ее защиты от несанкционированного доступа. Уменьшение размерности признакового пространства приводит к уменьшению времени оптимизации, поэтому предлагаемое решение целесообразно использовать в приложениях, требовательных к ресурсам, например, в мобильных устройствах или в устройствах «интернета вещей».

6. Благодарности

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации по государственной поддержке ведущих научных школ Российской Федерации № НШ-3070.2018.8

7. Литература

[1] Andriotis P. JPEG steganography detection with Benford's Law / P. Andriotis, G. Oikonomou, T. Tryfonas // Digital Investigation. – 2013. – Vol. 9. – P. 246–257.

[2] Chen M.-C. Alpha-trimmed Image Estimation for JPEG Steganography Detection / M.-C. Chen // Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics. – San Antonio, Texas, USA. – 2009. – P. 4581–4585.

[3] Cormen Th.H. Introduction to algorithms. Third edition / Th.H. Cormen, Ch.E. Leiserson, R.L. Rivest, C. Stein. – Third edition. – London: The MIT Press, 2013. – 1324 p.

[4] Fridrich J. Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes / J. Fridrich // Proceedings of the Sixth International Workshop on Information Hiding, Lecture Notes in Computer Science. – 2014. – Vol. 3200. – P. 67–81.

[5] Fu D. A generalized Benford's law for JPEG coefficients and its applications in image forensics / D. Fu, Y.Q. Shi, W. Su // Proceedings of SPIE 6505, Security, Steganography, and Watermarking of Multimedia Contents IX. – USA, San Jose. – 2007. – P. 1L1–1L11.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

[6] Guyon I. An introduction to variable and feature selection / I. Guyon, A. Elisseeff // Journal of Machine Learning Research. – 2003. – Vol. 3. – P. 1157–1182.
[7] Jia-Fa M. A steganalysis method in the DCT domain / M. Jia-Fa, N. Xin-Xin, X. Gang, Sh. Wei-Guo, Zh. Na-Na // Multimedia Tools and Applications. – 2016. – №75. – P. 5999–6019.
[8] Image Databases [Electronic recourse] // ImageProcessingPlace.com. – Electronic data. – URL: http://www.imageprocessingplace.com/root_files_V3/image_databases.htm.
[9] Molina L.C. Feature Selection Algorithms: A Survey and Experimental Evaluation / L.C. Molina, L. Belanche, A. Nebot // Proceedings of the 2002 IEEE International Conference on Data Mining, IEEE Computer Society. – 2002. – P. 306–313.
[10] SIPI Image Database [Electronic recourse] // USC University of Southern California. – Electronic data. – Los Angeles, CA, 2017. – URL: http://sipi.usc.edu/database/.
[11] Евсютин О.О. Алгоритм встраивания информации в сжатые цифровые изображения на основе операции замены с применением оптимизации / О.О. Евсютин, А.А. Шелупанов, Р.В. Мещеряков, Д.О. Бондаренко // Компьютерная оптика. – 2017. – Т. 41, № 3. – С. 412–421.
[12] Мицель А.А. Непараметрический алгоритм текстурного анализа аэрокосмических снимков / А.А. Мицель, Н.В. Колодникова, К.Т. Протасов // Известия Томского политехнического университета. – 2005. – Т. 308(1). – С. 65–70.

Об авторах

Шумская Ольга Олеговна, техник кафедры безопасности информационных систем факультета безопасности Томского государственного университета систем управления и радиоэлектроники. Ее e-mail shumskaya.oo@gmail.com.

Евсютин Олег Олегович, канд. техн. наук, доцент кафедры безопасности информационных систем факультета безопасности Томского государственного университета систем управления и радиоэлектроники. Его e-mail eoo@keva.tusur.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Компьютерное зрение

Computer Vision

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Контурная глубина плоской фигуры: применение для определения симметрии и классификации форм

Д.С. Белобородов[1]
dmitribeloborodov@yandex.ru
[1]МГУ им. М.В. Ломоносова, Москва, Россия

В статье предлагается новый подход к выделению признаков формы из плоских фигур и бинарных изображений — контурной глубины силуэта. Эти признаки вычисляются при помощи непрерывного скелета плоской фигуры и позволяют описать произвольный силуэт без отверстий в виде вектора конечной длины. Рассматривается применение контурной глубины для выявления симметрии объектов, а также для классификации листьев растений из набора данных Flavia.

Ключевые слова: *распознавание формы, выделение признаков, непрерывный скелет, поиск симметрии.*

Contour depth of a flat figure: application for symmetry detection and shape classification

D.S. Beloborodov[1]
dmitribeloborodov@yandex.ru
[1]Lomonosov MSU, Moscow, Russia

In this paper I propose a novel approach to feature extraction from flat figures and binary images — contour depth of a silhouette. These features are constructed from continuous skeleton of a flat figure, they allow to represent any silhouette without holes as a finite-sized vector. These features are used for symmetry detection, as well as for plant leaves classification in Flavia dataset.

Keywords: *shape recognition, feature extraction, continuous skeleton, symmetry detection.*

1. Введение

Задача распознавания формы является одной из важных задач в области компьютерного зрения. Как правило, эта задача предполагает определение класса различных форм или силуэтов, а также извлечение другой полезной информации. Роль исходных данных играют силуэты объектов, чаще всего представленные бинарными изображениями.

Вначале фигуру необходимо представить в виде, допускающем некоторый способ сравнения с другими фигурами, например, выделить из нее числовые признаки. Для задач анализа формы к этим признакам обычно предъявляются требования инвариантности относительно поворотов и сдвигов, а иногда и относительно деформаций. Наиболее распространенный подход заключается в том, чтобы вычислить некоторые статистики с помощью самой фигуры или ее скелета. Такие признаки инвариантны к поворотам и сдвигам, однако они не учитывают важную информацию о структуре фигуры. Существующие подходы можно разделить на три группы: основанные на данных контура, основанные на данных скелета, а также использующие одновременно контур и скелет.

Подходы, использующие контур, часто применяют алгоритмы выравнивания к последовательностям контура, также существуют алгоритмы на основе преобразования (изгиба) контуров, либо подсчета каких-либо статистик контура. К данным работам относятся [5, 6, 9, 10, 12]. Во втором подходе чаще всего применяются алгоритмы выравнивания скелетов фигур: [3, 4, 7, 11, 15]. В работах [2, 16] фигуры сравниваются по морфологическим спектрам (распределение ширины в фигуре). В [13, 14] для подсчета статистических признаков фигуры используются и скелет, и граница. В работе [8] производится выравнивание последовательностей, построенных при обходе скелета и учитывающих также информацию из контура. Предложенный в данной работе метод можно также отнести к последней группе.

Предлагается новый дескриптор плоской фигуры с гладкой границей, «функция контурной глубины», имеющая смысл распределения площади фигуры по ее границе. Этот дескриптор обобщается на случай многоугольной фигуры, граница которой уже не является гладкой. Предлагается способ выделения из силуэта вектора признаков конечной длины, «вектора контурной глубины», при помощи данного дескриптора. С помощью этого вектора можно, например, устанавливать симметрию объектов, а также сравнивать их напрямую в некоторых случаях. В следующих разделах будет подробно описан алгоритм вычисления вектора контурной глубины.

2. Функция контурной глубины

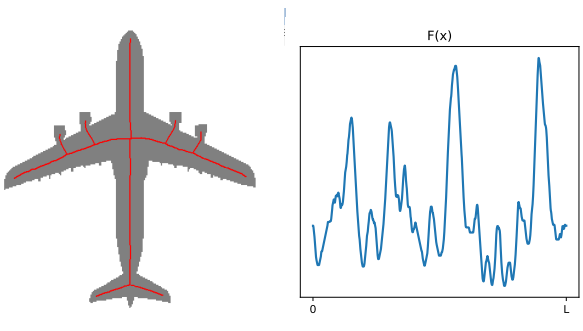
Рассмотрим плоскую фигуру с гладкой границей, не имеющую отверстий. Тогда для каждой точки границы мы можем рассмотреть множество точек внутри фигуры, для которых данная точка границы будет ближайшей. Введем отображение $S(Y)$: любому подмножеству Y границы фигуры оно ставит в соответствие множество точек внутри фигуры, для которых расстояние до множества Y меньше, чем до остальных

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

ных точек границы (локус). Можно выбрать на границе фигуры произвольную точку отсчета и параметризовать границу фигуры некоторой переменной $x \in [0, L)$, где L — длина границы фигуры. Пусть $\partial[x_1, x_2]$ — участок границы, параметризованный значениями $x \in [x_1, x_2]$. Тогда мы можем рассмотреть функцию $G(x) = |S(\partial[0, x])|$, $x \in [0, L)$ — она имеет смысл площади фигуры, заметаемой при проходе вдоль границы от начальной точки до x. Если у этой функции существует производная $G'(x) = \widetilde{F}(x)$ (для фигур с гладкой границей она будет существовать), будем называть ее функцией контурной глубины. Эта функция имеет смысл распределения площади фигуры относительно ее границы. Поскольку $x = 0$ и $x = L$ параметризуют одну точку на контуре, а $\widetilde{F}(x)$ задает прирост площади в данной точке, можно доопределить функцию в L: $\widetilde{F}(0) = \widetilde{F}(L)$. Из тех же соображений можно считать функцию $\widetilde{F}(x)$ периодической с периодом L (что имеет смысл при обходе контура более одного раза, а также обходе в обратном направлении). Стоит заметить, что некоторые точки внутри фигуры являются ближайшими для нескольких точек границы, и эти точки можно не учитывать при подсчете $G(x)$: множество всех таких точек является множеством центров максимальных вписанных кругов (скелетом фигуры) и имеет нулевую площадь. Поэтому $G(L) = \int_0^L \widetilde{F}(x) = S$.

При смене точки начала отсчета на границе сама функция $\widetilde{F}(x)$ принципиально меняться не будет, произойдет циклический сдвиг: $\widetilde{F}(y) = \widetilde{F}(x - l)$, где x, y — параметризации границы фигуры с точками начала отсчета x_0, y_0, а l — длина участка границы, заключенного между ними (при обходе от x_0 к y_0). Таким образом, данная функция обладает некоторой инвариантностью к поворотам фигуры: при смене точки начала обхода она сдвигается циклически. Также очевидна инвариантность относительно перемещения фигуры на плоскости.

Функция $F(x) = \widetilde{F}(x\sqrt{S})/\sqrt{S}$, $x \in [0, L/\sqrt{S}]$ должна быть инвариантна относительно масштаба фигуры, это функция контурной глубины аналогичной фигуры с единичной площадью. Функцию $F(x)$ будем называть нормированной контурной глубиной.



(a) **(b)**

Рис. 1. Плоская фигура и ее скелет (a) и функция контурной глубины (b)

Пример функции контурной глубины можно видеть на рисунке 1.

3. Контурная глубина многоугольных фигур

Для многоугольной фигуры определить функцию контурной глубины будет сложнее: вершины при вогнутых углах обладают множествами ближайших точек ненулевой площади, из-за чего функция $G(x)$ будет разрывной. В это же время вершины при выпуклых углах будут всегда вносить нулевой вклад в площадь. Таким образом, любые углы в контуре будут очень сильно влиять на значения функции.

Очевидным решением будет как-либо сгладить границы фигуры, и в таком случае будет гораздо удобнее рассматривать $F(x)$ как распределение площади по границе, и саму эту функцию подсчитывать с помощью непрерывного скелета фигуры и площадей ячеек Вороного.

Для описания алгоритма построения необходимо использовать уже введенное понятие непрерывного скелета: непрерывным скелетом плоской фигуры называется множество центров максимальных кругов, вписанных в данную фигуру. Для многоугольной фигуры ее скелет может быть представлен в виде плоского геометрического графа, узлы которого являются точками на плоскости, а ребра — отрезками прямых или парабол. При этом каждая точка на графе соответствует некоторому вписанному в эту фигуру максимальному кругу. Чтобы построить непрерывный скелет бинарного изображения, необходимо сначала приблизить его границы некоторой многоугольной фигурой, для которой впоследствии и строится скелет.

Существуют эффективные алгоритмы построения непрерывного скелета для бинарных изображений, они подробно описаны в [1].

Понятие непрерывного скелета тесно связано с понятием диаграммы Вороного: граф скелета является подмножеством границ ячеек Вороного. Диаграммой Вороного для некоторого набора элементов (сайтов) называется разбиение плоскости на области (ячейки Вороного), каждая из которых состоит из точек, ближайших к конкретному сайту. В терминах многоугольных фигур сайтами являются вершины и стороны фигуры. При этом каждая ячейка Вороного сопоставлена некоторому сайту, а каждый участок границы между ячейками Вороного представляет собой линию, в каждой точке которой расстояния до некоторых двух сайтов равны. Если сайты — это две точки или два отрезка прямых, разделяющая кривая будет отрезком прямой, если же эти два сайта являются точкой и отрезком, то кривая будет параболической. Несложно заметить, что непрерывный скелет фигуры будет подмножеством границ ячеек Вороного (все, лежащие внутри фигуры, за исключением границ между ребрами, соединяющимися под углом более 180°). Зная скелет, можно также восстановить границы ячеек Вороного.

Таким образом, при построении диаграммы Вороного внутренняя часть фигуры разбивается на ячейки,

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

соответствующие точкам и сегментам границы. Границы этих ячеек будут состоять из отрезков прямых и парабол. При этом, поскольку сайты границы неизбежно соприкасаются со своими соседями, сам сайт также обязательно будет частью границы ячейки. Зная непрерывный скелет фигуры, можно вычислить площади всех ячеек и установить их принадлежность к сайтам границы. Другими словами, мы знаем некоторый закон распределения площади фигуры по ее границе.

Вспомним, что на границе фигуры есть два типа вершин: выпуклые (внутренний угол при вершине не более 180°) и вогнутые (внутренний угол при вершине больше 180°). Ячейка Вороного для выпуклой вершины внутри фигуры всегда имеет нулевую площадь, ячейка для вогнутой вершины имеет ненулевую площадь и содержит саму эту вершину как часть своей границы (но не содержит никаких других точек из границы фигуры). В этом случае получается, что ячейка с ненулевой площадью отвечает участку границы нулевой длины.

Пусть для некоторой многоугольной фигуры известна диаграмма Вороного. Тогда можно составить список последовательных сегментов границы и площадей ячеек Вороного, соответствующих им. Такой же список можно составить для вершин границы и площадей их ячеек (для выпуклых вершин эти площади нулевые). Интуитивно это можно представить как разрезание фигуры на «дольки» и разворачивание в линию. Параметр $x \in [0, L]$ параметризует линию границы начиная с выбранной точки, где L — длина границы фигуры.

Будем считать, что площадь ячеек вершин не сконцентрирована в одной точке границы, а разделяется соседними точками границы некоторым образом. Пусть x_i — координата данной вершины на параметризованной линии границы, а S_i — площадь ячейки. Положим, что площадь ячейки распределена по точкам границы следующим образом: основная часть сконцентрирована в x_i, а при удалении от нее доля площади, приходящаяся на каждую точку, убывает линейно. Это можно формализовать в виде симметричной треугольной положительной функции с центром в x_i, определенной на $[0, L]$, интеграл от которой равен S_i :

$$f_i(x) = \begin{cases} k(x - x_i) + \sqrt{kS_i}, & x \in \left[x_i - \sqrt{\frac{S_i}{k}}, x_i\right] \\ k(x_i - x) + \sqrt{kS_i}, & x \in \left[x_i, x_i + \sqrt{\frac{S_i}{k}}\right] \\ 0, & else \end{cases}$$

где $k > 0$ — параметр размытости площади вокруг точки (наклон сторон треугольника), одинаковый для всех сайтов фигуры.

Теперь построим аналогичную функцию для сегментов границы. Вместо треугольника используем трапецию с тем же наклоном сторон k : это значит, что площадь ячейки равномерно распределена по ребру в центре, а по краям размывается. Пусть x_j^b, x_j^e — координаты начала и конца сегмента, $x_j^c = \frac{x_j^b + x_j^e}{2}$ — координата центра ребра. Также S_j — площадь ребра, l_j — его длина. Если $l_j \geqslant \frac{S_j}{kl_j}$ (ребро достаточно длинное, чтобы построить трапецию), зададим распределение площади следующей функцией:

$$g_j(x) = \begin{cases} k(x - x_j^b) + \frac{S_j}{2l_j}, & x \in \left[x_j^b - \frac{S_j}{2kl_j}, x_j^b + \frac{S_j}{2kl_j}\right] \\ \frac{S_j}{l_j}, & x \in \left[x_j^b + \frac{S_j}{2kl_j}, x_j^e - \frac{S_j}{2kl_j}\right] \\ k(x_j^e - x) + \frac{S_j}{2l_j}, & x \in \left[x_j^e - \frac{S_j}{2kl_j}, x_j^e + \frac{S_j}{2kl_j}\right] \\ 0, & else \end{cases}$$

где $k > 0$ — тот же параметр размытости.

Заметим, что площадь этой трапеции равна S_j, центры наклонных сторон имеют координаты x_j^b, x_j^e, а центр трапеции находится в x_j^c.

Если же условие $l_j \geqslant \frac{S_j}{kl_j}$ не выполняется, вместо трапеции используется треугольник:

$$g_j(x) = \begin{cases} k(x - x_j^c) + \sqrt{kS_j}, & x \in \left[x_j^c - \sqrt{\frac{S_j}{k}}, x_j^c\right] \\ k(x_j^c - x) + \sqrt{kS_j}, & x \in \left[x_j^c, x_j^c + \sqrt{\frac{S_j}{k}}\right] \\ 0, & else \end{cases}$$

Такой треугольник центрирован относительно x_j^c и имеет площадь S_j.

Эти формулы выписаны для случая, когда сайт границы находится достаточно далеко от начальной точки. Для сайтов, находящихся рядом с точкой старта, эти функции должны циклически продолжаться между концом и началом отрезка $[0, L]$ (поскольку граница является циклическим объектом, и любая точка может быть начальной).

Теперь функция контурной глубины может быть выражена через сумму всех полученных функций:

$$\widetilde{F}(x) = \sum_i f_i(x) + \sum_j g_j(x),$$

суммирование ведется по всем вершинам и сегментам границы соответственно.

Ее можно нормировать: $F(x) = \widetilde{F}(x\sqrt{S})/\sqrt{S}$, в терминах введенных выше функций это означает, что значения $x_i, x_j^b, x_j^e, x_j^c, l_i$ необходимо масштабировать на значение $1/\sqrt{S}$, а S_i — на значение $1/S$.

4. Вектор контурной глубины

Полученная нормированная функция контурной глубины является информативным дескриптором формы, однако саму эту функцию сложно использовать в задачах распознавания форм. Вместо этого предлагается использовать вектор фиксированной длины n. Для этого достаточно просто взять n точек t_n, расположенных равномерно на отрезке $[0, L/\sqrt{S}]$ и посчитать в них функцию $F(x)$. При этом $\sum_n F(t_n)\frac{L}{n} \approx 1$. Параметр n регулирует соотношение между информативностью контурной глубины и длиной вектора (сложностью вычислений). Также, поскольку длины контуров

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

нормированных фигур могут быть разными, правильнее будет зафиксировать шаг вычисления d_t, на который отстоят значения t_i и t_{i+1} — в этом случае для разных фигур длины векторов будут отличаться.

Полученные признаки инвариантны относительно сдвигов и масштабирования изображения, а также очевидно меняются при повороте: можно считать, что при этом просто происходит циклический сдвиг вектора. Они также устойчивы к упругим деформациям объектов. Однако, увеличение длины контура за счет появления дополнительных элементов может привести к заметному изменению признаков.

Параметр k отвечает за степень размытости площади по границе: чем больше k, тем больше наклон у сторон треугольников и трапеций, тем более детальной получается функция F. Наоборот, маленькие значения k позволяют сгладить эту функцию. Параметр n отвечает за длину вектора, в который преобразуется фигура, и также влияет на детализацию. Эти параметры связаны: при маленьких значениях k параметр n также можно уменьшить.

Теперь, если для контуров двух фигур фиксированы начальные точки, для них можно вычислить вектора контурной глубины a, b, которые можно сравнивать напрямую. Поскольку эти вектора отражают распределение площади по границе, для их сравнения можно использовать функции близости для множеств, например аналог коэффициента Жаккара: $J(a,b) = \frac{\sum_i min(a_i,b_i)}{\sum_i max(a_i,b_i)}$. Это значение имеет смысл отношения пересечения к объединению фигур (поскольку вычисленные признаки представляют площади).

Другим важным вопросом является поиск правильных начальных точек, с которых можно начинать сравнение. Помимо очевидного способа перебрать все n точек в каждом контуре или их подмножество, в некоторых случаях можно использовать специфичную структуру объектов, например, их симметрию.

5. Определение симметрии

Вектор контурной глубины можно использовать для простого поиска симметрии: для этого достаточно перебрать все циклические сдвиги вектора и найти сдвиг, максимизирующий коэффициент Жаккара для первой половины вектора и второй половины в обратном порядке, это займет $O(n^2)$. Точка сдвига определяет одну из точек на контуре, через которую проходит ось симметрии. В качестве второй точки оси симметрии берется точка, отстоящая от первой на $n/2$. По индексам компонент вектора можно восстановить и саму ось симметрии.

При нарушении симметрии фигуры (наличии дефектов контура) такой наивный алгоритм приведет к неточному результату. Для повышения точности можно воспользоваться описанной далее процедурой. Сначала необходимо циклически сдвинуть вектор на значение, найденное наивным алгоритмом. Затем производится симметрическое выравнивание вектора: выравнивание самого вектора контурной глубины с этим же вектором, записанным в обратном порядке, с помощью Алгоритма Нидлмана-Вунша выравнивания последовательностей. (Этот алгоритм изначально использовался для выравнивания символьных последовательностей с табличными функциями стоимости, однако он легко обобщается и на вещественные последовательности. В данной работе используются функции стоимости $f(a,b) = \min\{a,b\}/\max\{a,b,\epsilon\}$, $g(a) = -0.1$, где ϵ имеет маленькое значение для численной устойчивости). Тогда вторая точка оси симметрии будет задаваться элементом вектора, который был выравнен с самим собой (либо с соседним элементом). Этот процесс изображен на рисунке 2.

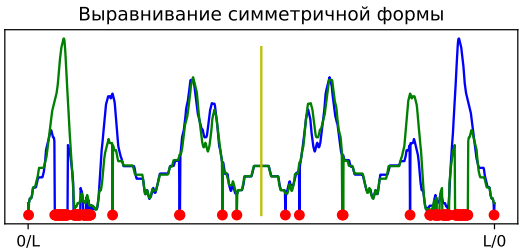


Рис. 2. Симметрическое выравнивание вектора. Красные точки — пропуски выравнивания, желтая линия — искомая точка оси симметрии.

Поскольку наивный алгоритм мог неверно определить первую точку оси симметрии, этот процесс можно повторить, использовав в качестве первой точку, найденную с помощью выравнивания. Результат выравнивания можно видеть на примере сложного контура на рисунке 3: в качестве контура выступает изображение самолета, один из двигателей которого соединился с корпусом из-за ошибки бинаризации. Это сильно влияет на симметрию (контур делится осью симметрии на неравные части), однако с помощью выравнивания можно точно восстановить ось симметрии. Сложность такого подхода составляет $O(n^2)$, где n — число компонент вектора контурной глубины.

(a) **(b)**

Рис. 3. Оси симметрии наивного метода (a) и с использованием выравнивания (b).

6. Классификация

Векторы контурной глубины были также применены в задаче классификации листьев растений в наборе данных Flavia [17], который содержит 1907 изображений листьев растений 32 разных видов.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Предварительно изображения были бинаризованы по пороговому значению. Для подсчета векторов контурной глубины использовалось значение $k = 0.1$ и шаг пересчета компонент вектора $d_t = 0.015$. Каждый вектор циклически сдвигался таким образом, чтобы наименьшая компонента (отвечающая одному из заостренных участков листа) оказалась в начале вектора. Также из каждого вектора вычиталось его минимальное значение — это позволяет перенести акцент с формы в целом на более тонкие детали. Затем для каждых двух векторов a, b считался аналог меры сходства Жаккара: $J_s = J(a,b) = \frac{\sum_i min(a_i,b_i)}{\sum_i max(a_i,b_i)}$. При этом пары векторов выравнивались по центру, а кратчайший вектор дополнялся нулями с двух сторон. Поскольку данные вектора оценивают только форму листа в целом, дополнительно была использована информация о цвете листьев: для этого были использованы гистограммы цвета (26 ячеек на каждый канал), для которых также использовалась мера сходства $J_c = J(u,v)$, где u, v — распределения цветов по ячейкам гистограммы. Для классификации был использован метод ближайших соседей с 10 соседями и мерой сходства $0.85J_s + 0.15J_c$. Полученная точность (accuracy) составила 0.866.

7. Заключение

Предложенные в данной статье признаки контурной глубины описывают распределение площади фигуры по ее границе. Эти признаки легко интерпретируемы, а гиперпараметры позволяют подбирать как детализацию, так и размер вектора. С помощью этих векторов можно сравнивать формы фигур, а также точно находить ось симметрии даже при нарушении контура.

8. Благодарности

Данная работа выполнена при поддержке гранта РФФИ 17–01–00917.

9. Литература

[1] Местецкий, Л. М., 2009. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. Москва, Физматлит.

[2] Местецкий, Л. М., 2014. Медиальная ширина фигуры — дескриптор формы изображений. Машинное обучение и анализ данных, 1(9), 1291-1318.

[3] Aslan, C., Erdem, A., Erdem, E., Tari, S., 2008. Disconnected skeleton: shape at its absolute scale. IEEE Trans. Pattern Analysis and Machine Intelligence 30, 2188–2203.

[4] Bai, X., Latecki, L., 2008. Path similarity skeleton graph matching. IEEE Trans. Pattern Analysis and Machine Intelligence 30, 1282–1292.

[5] Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Analysis and Machine Intelligence 24, 509–522.

[6] Felzenszwalb, P.F., Schwartz, J., 2007. Hierarchical matching of deformable shapes, in: CVPR.

[7] Klein, P., Tirthapura, S., Sharvit, D., Kimia, B., 2000. A tree-edit-distance algorithm for comparing simple, closed shapes. Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp. 696–704.

[8] Kushnir, O., Seredin, O., 2015. Shape matching based on skeletonization and alignment of primitive chains. International Conference on Analysis of Images, Social Networks and Texts, 123-136, Springer, Cham.

[9] Ling, H., Jacobs, D.W., 2007. Shape classification using the innerdistance. IEEE Trans. Pattern Analysis and Machine Intelligence 29, 286–299.

[10] Neuhaus, M., Bunke, H., 2006. Edit distance-based kernel functions for structural pattern classification. Pattern Recognition 39.10, pp. 1852–1863.

[11] Sebastian, T., Klein, P., Kimia, B., 2004. Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Analysis and Machine Intelligence 26, 550–571.

[12] Sederberg, T.W., Greenwood, E., 1992. A Physically Based Approach to 2-D Shape Blending, Comput. Graphics, no. 26 (2), pp. 25–34.

[13] Shen, W., Wang, X., Yao, C., Bai, X., 2014. Shape Recognition by Combining Contour and Skeleton into a Mid-Level Representation. Pattern Recognition, Springer Berlin Heidelberg, pp. 391–400.

[14] Shen, W., Jiang, Y., Gao, W., Zeng, D., Wang, X., 2016. Shape recognition by bag of skeleton-associated contour parts. Pattern Recognition Letters, 83, 321-329.

[15] Siddiqi, K., Shokoufandeh, A., Dickinson, S., Zucker, S., 1999. Shock graphs and shape matching. Int'l J. Computer Vision 35, 13–32.

[16] Vizilter, Yu.V., Sidyakin, S.V., Rubis, A.Yu., Gorbatsevich, V.S., 2012. Morphological shape comparison based on skeleton representations. Pattern Recognition and Image Analysis 22.3, pp. 412–418.

[17] Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.-X.,Chang, Y.-F., Xiang, Q.-L., 2007. A leaf recognition algorithm for plant classification using probabilistic neural network. Signal Processing and Information Technology, 2007 IEEE International Symposium, 11–16.

Об авторах

Белобородов Дмитрий Сергеевич, студент магистратуры кафедры Математических методов прогнозирования факультета Вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова. Его e-mail dmitribeloborodov@yandex.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Исследование свёрточных нейронных сетей класса YOLO для мобильных систем детектирования объектов на изображениях

А.П. Береснев[1], И.В. Зоев[1], Н.Г. Марков[1]
apb3@tpu.ru|ivz3@tpu.ru|markovng@tpu.ru
[1]Томский политехнический университет, г. Томск, Российская Федерация

Предлагаются свёрточные нейронные сети класса YOLO новой архитектуры для решения задачи детектирования объектов на изображениях с помощью аппаратно-реализованной системы компьютерного зрения. Приводятся результаты анализа эффективности таких сетей в сравнении со свёрточными сетями класса YOLO с другими архитектурами.

Ключевые слова: *свёрточные нейронные сети, класс нейронных сетей YOLO, детектирование объектов на изображениях, мобильные системы компьютерного зрения.*

Research on convolutional neural networks of YOLO class for mobile object detection system

A.P. Beresnev[1], I.V. Zoev[1], N.G. Markov[1]
apb3@tpu.ru|ivz3@tpu.ru |markovng@tpu.ru
[1]Tomsk Polytechnic University, Tomsk, Russian Federation

This article presents new convolutional network architectures of the YOLO class to solve the problem of detecting objects in images using a hardware-based computer vision system. The efficiency analysis results of proposed networks in comparison with convolutional networks of YOLO class with other architectures are presented.

Keywords: *convolutional neural network, neural networks of YOLO class, object detection, mobile computer vision systems.*

1. Введение

В последние годы при решении задачи детектирования объектов на изображениях наметилась тенденция к созданию мобильных систем компьютерного зрения (СКЗ) на основе аппаратно-реализованных свёрточных нейронных сетей (СНС) [1]. При этом весьма перспективными являются СНС класса YOLO [7]. Такие СНС обеспечивают довольно высокую точность и скорость детектирования объектов на изображениях, но сложны в аппаратной реализации.

В статье описываются СНС новой архитектуры из класса YOLO, более простые в аппаратной реализации, чем другие сети из этого класса. Приводятся результаты исследования эффективности таких сетей. Дается их сравнение с результатами исследований эффективности СНС других архитектур этого же класса.

2. Задача аппаратной реализации СНС для детектирования объектов на изображениях

В настоящее время во многих областях человеческой деятельности востребованы мобильные (СКЗ), устанавливаемые на беспилотные летательные аппараты (БПЛА). Такие СКЗ должны решать задачи детектирования объектов различной физической природы на изображениях земной поверхности. Под задачей детектирования обычно понимается определение ограничивающих областей и класса для каждого объекта на анализируемых изображениях. При решении задачи детектирования наметилась тенденция к использованию аппаратно-реализованных СНС с целью повышения скорости детектирования объектов. При этом, создание мобильных СКЗ значительно усложняется тем, что при их разработке необходимо искать баланс между точностью детектирования, энергопотреблением и быстродействием бортового вычислительного устройства, реализующего СНС, из-за ограниченности используемых вычислительных ресурсов СКЗ. Для этого требуются дополнительные исследования эффективности разрабатываемых мобильных СКЗ при использовании в них различных классов и архитектур СНС.

Цель данной работы – исследование эффективности различных архитектур СНС для применения в мобильных СКЗ, которые наиболее полно отвечают требованиям по точности детектирования объектов на изображениях, скорости выполнения каждой из архитектур СНС и по сложности её аппаратной реализации. Последние два требования, по сути, определяют требование к быстродействию бортового вычислительного устройства в составе СКЗ.

Анализ показывает, что среди наиболее подходящих для решения задачи детектирования СНС предпочтение сегодня следует отдать сетям класса YOLO [7]. Эти сети обеспечивают большие точность и скорость детектирования объектов на изображениях, чем сети других классов. Проанализируем наиболее известные архитектуры сетей класса YOLO.

Наиболее исследованной и перспективной считается архитектура YOLOv2, включающая 22 свёрточных слоя. В качестве функции активации сетей этого класса используется функция leaky ReLU, определяемая по формуле $max(0{,}1x, x)$. Сеть этой архитектуры имеет ряд отличий от исходной. YOLOv2 – полностью свёрточная сеть (англ. fully convolutional network), из архитектуры этой сети убрано два полносвязных слоя на выходе, что позволяет не использовать метод dropout на этапе обучения для предотвращения переобучения. Чтобы детектировать объекты разного масштаба карты признаков перед последним слоем подвыборки (англ. pooling layer) объединяются с выходными картами признаков выходного слоя. Для определения ограничивающих областей объектов используются смещения относительно якорных значений (англ. anchor boxes), которые вычисляются на из обучащей выборки с использованием метода K средних. Точность

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

детектирования по методике mAP составляет 76,8% на выборке Pascal VOC 2007, а скорость выполнения этой сети при такой точности достигает 67 кадров в секунду на видеокарте NVIDIA Titan X. В силу ограниченности вычислительных ресурсов в мобильных СКЗ перспективной для аппаратной реализации является более простая архитектура tiny-YOLO, имеющая только 9 свёрточных слоёв. В этой архитектуре не используется добавление карт признаков из предыдущих слоев в последующие. Однако точность детектирования и скорость выполнения сети близки к соответствующим значениям в случае использования архитектуры YOLOv2.

Отметим, что во всех архитектурах сетей класса YOLO используется операция батч-нормализации (англ. batch-normalization) после каждого свёрточного слоя. Это усложняет процесс вычислений в сети. Естественно, что при аппаратной реализации СНС для многократного выполнения такой операции должны использоваться дополнительные вычислительные ресурсы и возрастет энергопотребление СКЗ. В этой связи необходимо разрабатывать СНС новой архитектуры без операции батч-нормализации.

3. СНС класса YOLO новой архитектуры

Глубокие нейронные сети были предметом повышенного интереса до недавнего времени. Однако сегодня мобильные устройства позволяют проводить эксперименты по использованию нейронных сетей для решения различного рода задач на самом устройстве. Различные методы позволяют сократить затраты на вычисления и используемые ресурсы, при этом сохраняя точность работы сети в сравнении с более глубокими вариантами сетей. Например, на выборке ImageNet хорошо себя зарекомендовал подход с использованием Inception-модулей, которые применялись в СНС GoogLeNet [4].

В статье [5] приведены результаты исследования, которое показало, что увеличение количество слоёв сети не ведет к увеличению точности детектирования. Вместе с этим, была предложена новая архитектура с использованием ResNet-модулей, где данные на входе слоя подаются также и на вход следующего за ним через два слоя. Это позволяет использовать меньшее количество слоёв сети при более высокой точности детектирования.

Возможна комбинация модулей, когда в Inception-модуль встраивается ResNet-модуль. Такой подход и различные комбинации модулей представлены в статье [5]. В результате сокращается число вычислений в сравнении с более глубокими сетями, но сохраняется точность детектирования. Используем данный подход для развития архитектуры СНС tiny-YOLO, заменив часть слоев на InceptionResNet-модули в различных комбинациях. На рис. 1 представлена детализированная схема Inception-ResNet-модуля.

В табл. 1 представлена исходная архитектура СНС tiny-YOLO. Так как выход InceptionResNet-модуля имеет размер в 256 карт признаков и объединяется со входом, то изменим количество карт признаков четвертого свёрточного слоя с 128 на 256 карт признаков. Также уменьшим количество карт признаков в последующих слоях свёртки сети tiny-YOLO.

Другим важным отличием предлагаемой архитектуры от исходной является отказ от использования операции батч-нормализации.

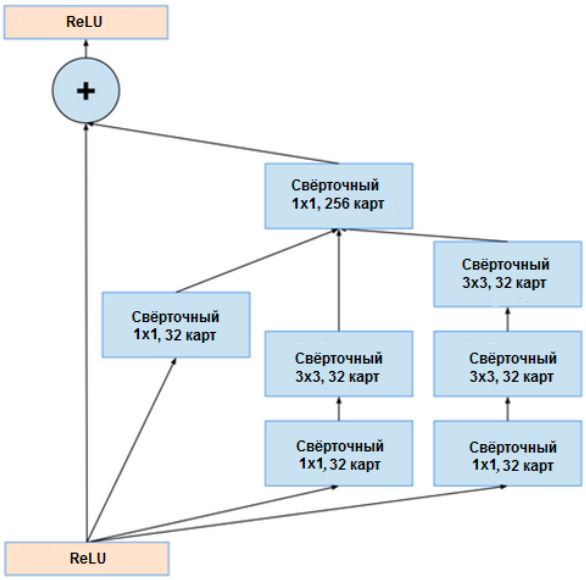


Рисунок 1. Inception-ResNet-модуль

Таблица 1. Архитектура СНС tiny-YOLO

№	Тип слоя	Ядра	Размеры/шаг	Размеры входа
1	Свёрточный	16	3x3/1	416x416x3
2	Подвыборки	-	2x2/2	416x416x16
3	Свёрточный	32	3x3/1	208x208x16
4	Подвыборки	-	2x2/2	208x208x32
5	Свёрточный	64	3x3/1	104x104x32
6	Подвыборки	-	2x2/2	104x104x64
7	Свёрточный	128	3x3/1	52x52x64
8	Подвыборки	-	2x2/2	52x52x128
9	Свёрточный	256	3x3/1	26x26x128
10	Подвыборки	-	2x2/2	26x26x256
11	Свёрточный	512	3x3/1	13x13x256
12	Подвыборки	-	2x2/1	13x13x512
13	Свёрточный	1024	3x3/1	13x13x512
14	Свёрточный	1024	3x3/1	13x13x1024
15	Свёрточный	125	1x1/1	13x13x1024
16	Детектор	-	-	

В итоге в табл. 2 представлена новая архитектура tiny-YOLO-InceptionResnet.

Таблица 2. Архитектура СНС tiny-YOLO-InceptionResNet

№	Тип слоя	Ядра	Размеры/шаг	Размеры входа
1	Свёрточный	16	3x3/1	416x416x3
2	Подвыборки	-	2x2/2	416x416x16
3	Свёрточный	32	3x3/1	208x208x16
4	Подвыборки	-	2x2/2	208x208x32
5	Свёрточный	64	3x3/1	104x104x32

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

6	Подвыборки	-	2x2/2	104x104x64
7	Свёрточный	256	3x3/1	52x52x64
8	Подвыборки	-	2x2/2	52x52x256
9	Inception ResNet	-	-	26x26x256
10	Свёрточный	128	3x3/1	26x26x256
11	Подвыборки	-	2x2/2	26x26x128
12	Свёрточный	256	3x3/1	13x13x128
13	Подвыборки	-	2x2/1	13x13x256
14	Свёрточный	128	3x3/1	13x13x256
15	Свёрточный	125	1x1/1	13x13x128
16	Детектор	-	-	

Предложим еще несколько новых архитектур СНС на базе сети tiny-YOLO. Так используем последовательно по два и три InceptionResNet-модуля, которые следуют друг за другом и получим архитектуры tiny-YOLO-InceptionResNet2 и tiny-YOLO-InceptionResNet3. Также параллельно используем два и три таких же модуля, имеющих один вход и конкатенацию их выходных карт признаков, что дает новые архитектуры tiny-YOLO-InceptionResNet-x2 и tiny-YOLO-InceptionResNet-x3.

4. Результаты исследований эффективности сетей класса YOLO

Исследования эффективности СНС YOLOv2, tiny-YOLO и предложенных новых архитектур сетей проводились по скорости выполнения каждой из архитектур, по сложности их для аппаратной реализации и по точности детектирования объектов на тестовой выборке PascalVOC2007[6].

Скорость выполнения каждой из архитектур определялась как среднее время выполнения сети при детектировании объектов одного изображения из тестовой выборки, состоящей из 4952 изображений. В табл. 3 представлены значения среднего времени выполнения для всех исследуемых архитектур СНС. Видим, что наилучшей скоростью обладает архитектура tiny-YOLO-Inception-ResNet , а наихудшей – сеть YOLOv2.

Таблица 3. Среднее время детектирования объектов на одном изображении

Архитектура СНС	Среднее время выполнения СНС, мс
YOLOv2	35,016
tiny-YOLO	12,124
tiny-YOLO-Inception-ResNet	6,943
tiny-YOLO-Inception-ResNet2	7,546
tiny-YOLO-Inception-ResNet3	8,337
tiny-YOLO-Inception-ResNet_x2	7,948
tiny-YOLO-Inception-ResNet_x3	9,047

Сложность при аппаратной реализации СНС определяется числом операций, реализуемых сетью в единицу времени, и объемом требуемой памяти, в первую очередь, для весовых коэффициентов сети. Основные операции, используемые в сетях класса YOLO: умножение с накоплением, сравнение, операции активации, сложение, деление. Причем последние две необходимы при реализации батч-нормализации, которая отсутствует в предложенных новых архитектурах СНС. В табл. 4 в среднем столбце представлено количество всех операций, реализуемых каждой из архитектур. Чем меньше таких операций, тем проще аппаратная реализация архитектуры СНС. В правом столбце приведены значения объема требуемой памяти для весовых коэффициентов СНС.

Таблица 4. Требуемые ресурсы для аппаратной реализации исследуемых архитектур СНС

Архитектура СНС	Число операций СНС, GFlops	Объем весовых коэффициентов СНС, МБ
YOLOv2	29,36	202,7
tiny-YOLO	6,973	63,5
tiny-YOLO-Inception-ResNet	2,451	4,6
tiny-YOLO-Inception-ResNet2	3,136	6,6
tiny-YOLO-Inception-ResNet3	3,622	8,1
tiny-YOLO-Inception-ResNet_x2	3,129	6,6
tiny-YOLO-Inception-ResNet_x3	3,719	8,4

Из табл. 4 следует, что минимальные ресурсы требуются для аппаратной реализации архитектуры tiny-YOLO-Inception-ResNet. Для архитектур YOLOv2 и tiny-YOLO требуется максимальное количество ресурсов, в том числе из-за использования операции батч-нормализации.

Исследование эффективности СНС в части точности детектирования объектов осуществлялось с использованием метрики, применяемой в PascalVOC2007. В PascalVOC2007 точность оценивается интегральным параметром mean Average Precision (mAP). Данный параметр является средним значением от Average Precision (AP) для каждого класса из обучающей выборки. AP рассчитывается по формуле:

$$AP = \frac{1}{11} \sum_{r \in \{0, 0.1, \ldots, 1\}} P(r),$$

$$P = \frac{TP}{TP + FP},$$

$$r = \frac{TP}{TP + TN},$$

где P – precision r – recall, TP – true positive (правильно распознанный объект), FP – false positive (неправильно распознанный объект), FN – false negative (ложное срабатывание детектора на фоне).

Таблица 5. Точность детектирования СНС

Архитектура СНС	mAP, %
YOLOv2	76,80
tiny-YOLO	57,1
tiny-YOLO-Inception-ResNet	46,22
tiny-YOLO-Inception-ResNet2	53,68
tiny-YOLO-Inception-ResNet3	55,45
tiny-YOLO-Inception-ResNet_x2	49,54
tiny-YOLO-Inception-ResNet_x3	46,52

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

В табл. 5 значения mAP для архитектур YOLOv2 и tiny-YOLO взяты из [7].

Нами для обучения СНС с новыми архитектурами использовался open source фреймворк, названный darknet [2]. Он поддерживает декларативное описание архитектуры сетей в виде конфигурационного файла с расширением cfg. При обучении использовались параметры, предлагаемые авторами архитектуры YOLOv2 и tiny-YOLO, однако размер батчей был ограничен 16 элементами обучающей выборки. Обучение предлагаемых архитектур проводилось на видеокарте NVIDIA GTX 1060 на фреймворке darknet с использованием библиотек CUDA 9.1 и cuDNN 7.

Из табл. 5 видно, что наилучшие результаты по точности детектирования даёт СНС YOLOv2, однако использование этой сети, как следует из табл. 3 и табл. 4, требует значительных аппаратных ресурсов и уступает по скорости другим архитектурам. К сожалению, архитектура tiny-YOLO-Inception-ResNet имеет невысокую точность детектирования. Разработчикам мобильных СКЗ следует обратить внимание на архитектуры tiny-YOLO-Inception-ResNet2 и tiny-YOLO-Inception-ResNet3, поскольку именно для них имеет место баланс между точностью детектирования объектов и скоростью выполнения этих СНС.

5. Заключение

В последнее время при решении задачи детектирования объектов на изображениях наметилась тенденция к созданию мобильных СКЗ. В составе таких СКЗ используются аппаратно-реализованные СНС. В данной работе ставилась задача разработки и исследования эффективности различных архитектур СНС класса YOLO для мобильных СКЗ.

Предложены новые архитектуры СНС на основе архитектуры tiny-YOLO без использования батч-нормализации и с включением в архитектуру tiny-YOLO Inception-ResNet-модулей в различных конфигурациях.

Проведенные исследования эффективности известных и предложенных архитектур СНС класса YOLO позволили считать, что наиболее перспективными для аппаратной реализации СНС в составе СКЗ являются предложенные архитектуры tiny-YOLO-Inception-ResNet2 и tiny-YOLO-Inception-ResNet3.

Именно с использованием таких архитектур СНС разработчикам мобильных СКЗ можно найти баланс между точностью и скоростью детектирования объектов на изображениях.

6. Благодарности

Исследования были поддержаны грантом РФФИ № 18-47-700010 р_а.

7. Литература

[1] Зоев, И.В. Устройство на основе ПЛИС для распознавания рукописных цифр на изображениях/ И.В. Зоев, А.П. Береснев, Н.Г. Марков, А.Н. Мальчуков // Компьютерная оптика. – 2017. – Т. 41, № 6. – С. 938-949. – DOI: 10.18287/2412-6179-2017-41-6-938-949.

[2] Darknet framework. [Электронный ресурс] – URL: http://github.com/pjreddie/darknet (дата обращения 14.05.2018)

[3] Deep Residual Learning for Image Recognition [Электронный ресурс] – URL: https://arxiv.org/abs/1512.03385 (дата обращения 14.05.2018)

[4] Going Deeper with Convolutions. [Электронный ресурс] – URL: https://arxiv.org/abs/1409.4842 (дата обращения 14.05.2018)

[5] Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. [Электронный ресурс] – URL: https://arxiv.org/abs/1602.07261 (дата обращения 14.05.2018)

[6] Pascal VOC database. [Электронный ресурс] – URL: https://host.robots.ox.ac.uk/pascal/VOC (дата обращения 14.05.2018)

[7] YOLO9000: Better, Faster, Stronger. [Электронный ресурс] – URL: https://arxiv.org/pdf/1612.08242.pdf (дата обращения 14.05.2018)

Об авторах

Береснев Алексей Павлович, магистрант отделения информационных технологий инженерной школы информационных технологий и робототехники Томского политехнического университета, e-mail: apb3@tpu.ru.

Зоев Иван Владимирович, аспирант отделения информационных технологий инженерной школы информационных технологий и робототехники Томского политехнического университета., e-mail: ivz3@tpu.ru.

Марков Николай Григорьевич, д.т.н., профессор отделения информационных технологий инженерной школы информационных технологий и робототехники Томского политехнического университета, e-mail: markovng@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Аппаратная реализация на ПЛИС свёрточных нейронных сетей для распознавания объектов на изображениях

И.В. Зоев, Н.Г. Марков,А.П. Береснев,Т.А. Ягунов
ivz3@tpu.ru | markovng@tpu.ru | apb3@tpu.ru | tay1@tpu.ru
Томский политехнический университет, г. Томск, Российская Федерация

Рассмотрены особенности аппаратной реализации на ПЛИС, входящей в систему на кристалле Cyclone V SX, свёрточных нейронных сетей (СНС) класса LeNet 5 для распознавания объектов на изображениях. В аппаратной СНС используются вычислительные блоки двух типов: свёртки и подвыборки. Разработан метод организации вычислений в такой СНС. Приведены результаты исследования его эффективности при решении задачи распознавания рукописных цифр на изображениях.

Ключевые слова: *аппаратные свёрточные нейронные сети, программируемые логические интегральные схемы, системы компьютерного зрения, распознавание объектов на изображениях.*

FPGA-based hardware implementation of convolution neural networks for images recognition

I.V. Zoev, N.G. Markov, A.P. Beresnev, T.A. Yagunov
ivz3@tpu.ru | markovng@tpu.ru | apb3@tpu.ru | tay1@tpu.ru
Tomsk Polytechnic University, Tomsk, Russian Federation

This text presents specific of hardware implementation of convolution neural networks like LeNet 5 using system on chip Altera Cyclone V SX with FPGA. Hardware implementation has two types of calculating units: convolution and pooling. We developed a method of calculation organization in hardware implementation CNN. Also this text consist the results of efficiency for handwritten digit recognition task for developed method.

Keywords: *hardware convolution neural networks, field-programmable gate array, computer vision, images recognition.*

1. Введение

В последние годы для распознавания объектов на изображениях часто используются свёрточные нейронные сети (СНС). При создании мобильных систем компьютерного зрения (СКЗ) применяются аппаратно реализованные СНС. Учитывая высокую вычислительную сложность СНС, разработчикам таких СКЗ приходится искать баланс между точностью распознавание объектов на изображениях, производительностью и энергопотреблением таких систем. Программируемые логические интегральные схемы (ПЛИС) современных систем на кристалле имеют низкое энергопотребление, а также позволяют реализовывать параллельные вычисления. В этой связи для аппаратной реализации СНС в составе СКЗ все чаще применяют современные ПЛИС.

В работе [3] показаны результаты ускорения вычислений при реализации СНС весьма сложной архитектуры на ПЛИС с большими вычислительными ресурсами, но и со значительным энергопотреблением. На наш взгляд, при создании мобильных СКЗ можно использовать широко распространённые ПЛИС с ограниченными вычислительными ресурсами и, соответственно, с малым энергопотреблением. Так, в работе [4] приведены результаты исследования эффективности аппаратной СНС класса LeNet 5 для распознавания на изображениях рукописных цифр, при создании которой использовалась именно такая ПЛИС системы на кристалле Altera Cyclone V SX. Авторами предложена и аппаратно реализована в виде устройства архитектура СНС, основанная на хорошо известной архитектуре LeNet 5. В предложенной архитектуре 4 свёрточных слоя и 2 слоя подвыборки, а в качестве функции активации применяется оператор ReLU. В аппаратной СНС используются созданные на ПЛИС вычислительные блоки двух типов: свёртки и подвыборки. Блоки соответствующего типа выполняются параллельно внутри слоя свёртки или подвыборки. Из-за использования процедур свёртки и подвыборки с разными значениями параметров, число одновременно выполняемых блоков одного типа в разных слоях отличается и определяется конкретным слоем. Нетрудно видеть, что такой простой способ организации вычислений ведёт к простою ряда вычислительных блоков и, в итоге, к невысокой производительности устройства.

В обзоре [2] приведена схема конвейерной архитектуры вычислителя для СНС. В этой архитектуре кроме ПЛИС используется центральный процессор системы на кристалле и вычисления ведутся над числами в различных форматах. Однако в [2] отсутствуют результаты по энергопотреблению вычислителя такой архитектуры, что не позволяет оценить его пригодность в мобильных СКЗ.

Анализ этих и других подобных им работ указывает на актуальность дальнейших исследований при создании мобильных СКЗ в части используемых моделей вычислений и методов и способов реализации СНС на ПЛИС.

В данной работе предлагается метод организации вычислений в аппаратной СНС класса LeNet 5 с использованием унифицированных блоков свёртки и подвыборки. Приведены особенности реализации метода и результаты исследования его эффективности при решении задачи распознавания рукописных цифр на изображениях.

2. Метод организации вычислений в аппаратной СНС и особенности его реализации

Нами предлагается при организации вычислений в аппаратной СНС с архитектурой из [4], подобной LeNet 5, использовать вычислительные возможности не только ПЛИС, но и других компонентов современных систем на кристалле. Так, в системе на кристалле Altera Cyclone V SX

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

кроме ПЛИС имеется двухъядерный процессор ARM Cortex A9, который имеет прямой доступ к внешней памяти. Архитектура рассматриваемой системы на кристалле позволяет организовать прямой доступ аппаратно реализованной СНС к этой внешней памяти. Реализация такого способа взаимодействия ПЛИС с внешней памятью процессора позволит выполнять некоторые операции, отличные от процедур свёртки и подвыборки, не на ПЛИС, а на самом процессоре.

Опираясь на этот способ, можно предложить метод организации вычислений в аппаратной СНС на ПЛИС, отличающийся от известных методов использованием унифицированных вычислительных блоков свёртки и подвыборки. Унификация блоков свёртки/подвыборки достигается путем извлечения параметров блоков, обычно задающихся на этапе их синтеза, и размещения их в отдельную изменяемую область памяти ПЛИС, называемую конфигурационной областью. Это позволит использовать блоки в слоях СНС, которые имеют разные параметры, такие как высота и ширина ядра свёртки/подвыборки, так и количество входных/выходных карт признаков.

Другой особенностью метода является то, что число задействованных в аппаратной СНС вычислительных блоков может быть переменным, причем масштабирование ведется как в целом для СНС, так и для её отдельных слоёв. Это позволит значительно сократить используемые вычислительные ресурсы ПЛИС и реализовать различные архитектуры СНС без реконфигурации ПЛИС.

При проведении унификации также необходимо:

- реализовать конфигурационную область памяти на ресурсах ПЛИС, которая будет хранить конфигурируемые параметры, необходимые не только для вычислительных блоков, но и для вспомогательных блоков аппаратной СНС;
- организовать доступ к этой области памяти процессору ARM, который в дальнейшем и будет осуществлять конфигурирование аппаратной СНС.
- создать дополнительное программное обеспечение в виде драйвера аппаратной СНС, исполняемого на процессоре ARM под управлением ядра ОС Linux, что позволит обеспечить значительную гибкость аппаратной реализации СНС.

Стоит отметить такой параметр конфигурационной области памяти ПЛИС как флаг прерывания слоя, благодаря которому по завершению вычислений в слое аппаратная СНС приостанавливает работу и даёт возможность процессору реализовать преобразование данных для выполнения операций в следующем слое сети.

Обработку таких прерываний осуществляет драйвер аппаратной СНС. Также драйвер выделает область во внешней памяти для весовых коэффициентов, входных значений СНС и промежуточных результатов её работы. Это происходит на основе передаваемых драйверу конфигурационных данных СНС, которые он в последующем записывает в конфигурационную область памяти ПЛИС. Драйвер осуществляет доступ к выделенным областям внешней памяти для аппаратной СНС и для процессора, что позволяет последнему быстро менять как входные значения сети – изображения, которые необходимо распознать, так и весовые коэффициенты СНС. Благодаря этому на ПЛИС, в принципе, можно сконфигурировать и другие классы СНС, отличные от класса LeNet 5.

Нетрудно видеть, что процессы обмена между ПЛИС и внешней памятью могут негативно сказаться на производительности разрабатываемой СКЗ из-за задержек по времени на реализацию операций ввода/вывода.

3. Результаты исследований эффективности метода

Для проведения исследований эффективности предлагаемого метода был создан макет СКЗ на отладочной плате Terasic SoCkit с системой на кристалле Altera Cyclone V. Архитектура СНС, аппаратно реализованная в макете, взята из работы [4]. Эта СНС относится к классу LeNet 5 и имеет следующие параметры. Первый слой - свёрточный, количество входных карт признаков 3, выходных 6, ядро свёртки 7x7, шаг 1. Второй слой - подвыборки, количество входных карт признаков 6, выходных 6, ядро подвыборки 2x2, шаг 2. Третий слой - свёрточный, количество входных карт признаков 6, выходных 32, ядро свёртки 5x5, шаг 1. Четвертый слой - подвыборки, количество входных карт признаков 32, выходных 32, ядро подвыборки 2x2, шаг 2. Пятый слой - свёрточный, количество входных карт признаков 32, выходных 100, ядро свёртки 5x5, шаг 1. Шестой (полносвязный) слой – свёртки, количество входных карт признаков 100, выходных 10, ядро свёртки 1x1, шаг 1. Весовые коэффициенты для аппаратной СНС были получены при программной реализации этой архитектуры СНС (использовалось 32-разрядное представление формата чисел с плавающей запятой), которая обучалась на выборке изображений рукописных цифр из базы MNIST размером 32x32 пикселя [1]. Точность распознавания рукописных цифр для программной реализации СНС составляет 98,71% и принята нами за эталонную точность. В макете СКЗ для сокращения необходимых вычислительных ресурсов ПЛИС не использовалась нормализация выходных данных слоёв свёртки. Это привело к необходимости применения модели вычислений в виде операций над числами с плавающей запятой.

В табл. 1 представлены результаты исследования точности распознавания рукописных цифр на изображениях из базы MNIST с помощью макета СКЗ. Разработаны несколько вариантов макета СКЗ, в которых на ПЛИС реализованы базовые операции над числами в формате с плавающей запятой различной разрядности. Результаты по точности распознавания рукописных цифр на изображениях в случаях использования 26 и 32 разрядных чисел близки к значению эталонной точности и в табл. 1 не приведены. Для сравнения второй строкой в этой таблице приведены результаты точности распознавания этих же объектов, полученные с помощью устройства из работы [4].

Таблица 1. Зависимость точности распознавания объектов от разрядности формата чисел с плавающей запятой.

Разрядность / Устройство	12 бит	14 бит	16 бит
Макет СКЗ	97,03 %	98,51 %	98,66%
Устройство из [4]	93,48 %	98,27 %	98,34%

Видим, что макет СКЗ показал лучшую точность распознавания, чем устройство из [4] при всех разрядностях чисел формата с плавающей запятой, используемых в макете. Это связано с изменением методики переноса весовых коэффициентов СНС на ПЛИС. В работе [4] перенос осуществлялся на этапе синтеза аппаратной СНС с помощью библиотеки numpy, а в случае макета СКЗ – путем записи весов в память процессора ARM Cortex A9 при помощи библиотеки Half-precision floating point library.

Второй эксперимент был направлен на исследование производительности макета СКЗ. Далее в качестве примера приведены результаты эксперимента для макета, который

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

работает с 16 разрядным форматом чисел с плавающей запятой.

Для последующего сравнения производительности макета СКЗ и устройства из [4], в табл. 2 приведены результаты теоретических расчетов производительности при аппаратной реализации СНС в устройстве из [4]. Расчеты проводились на основе того, что каждый вычислительный блок работает на частоте 50 МГц, что позволяет ему выполнять $50*10^6$ операций в секунду.

Таблица 2. Результаты расчетов производительности устройства из [4] при распознавании одного изображения.

№ слоя	Кол-во и тип используемых вычислительных блоков	Кол-во операций	Расчетное время выполнения, мс
1	6 свёрточных	596232	1,987
2	6 блоков подвыборки	4056	0,014
3	32 свёрточных	388800	0,243
4	32 блока подвыборки	3200	0,002
5	100 свёрточных	80000	0,016
6	10 свёрточных	1000	0,002
		Итого	2,264

Из табл. 2 следует неравномерность распределения количества операций по слоям СНС, поэтому вычисления в слоях выполняются с достаточно большой разницей по времени. Теоретическое значение времени работы в целом устройства при распознавании цифр на одном изображении равно 2,264 мс и близко к полученному экспериментально значению из [4], которое составляет 2,417 мс. Общее требующееся в устройстве количество вычислительных блоков свертки — 148, а подвыборки — 38. Такое количество блоков занимает практически все ресурсы ПЛИС при использовании 16 разрядного формата чисел с плавающей запятой. Очевидно, что такая конфигурация блоков для способа организации вычислений, принятого в [4], является для ПЛИС системы на кристалле Altera Cyclone V SX максимально возможной.

Для макета СКЗ также получены расчетным путем результаты его теоретической производительности. Расчет вёлся таким же способом, что и для устройства из [4], с учётом того, что вычислительные блоки макета СКЗ работают на той же частоте 50 МГц. Теоретические и экспериментальные значения производительности макета, полученные в ходе исследования при распознавании рукописных цифр на изображениях из базы MNIST, представлены в табл. 3-6 для различного числа вычислительных блоков СНС. Число блоков в эксперименте задается (масштабируется), исходя из количества выходов слоев реализуемой архитектуры СНС.

Таблица 3. Результаты исследования производительности макета СКЗ при распознавании одного изображения (количество блоков свёртки — 6, подвыборки — 6).

№ слоя	Кол-во и тип используемых вычислительных блоков	Кол-во операций	Время выполнения, мс	
			расчетное	эксперимен.
1	6 свёрточных	596232	1,987	4,176
2	6 блоков подвыборки	4056	0,014	0,381
3	6 свёрточных	388800	1,296	9,221
4	6 блоков подвыборки	3200	0,011	0,362
5	6 свёрточных	80000	0,267	1,981
6	6 свёрточных	1000	0,003	0,343
Итого			3,578	16,464

Таблица 4. Результаты исследования производительности макета СКЗ при распознавании одного изображения (количество блоков свёртки — 10, подвыборки — 10).

№ слоя	Кол-во и тип используемых вычислительных блоков	Кол-во операций	Время выполнения, мс	
			расчетное	эксперимен.
1	6 свёрточных	596232	1,987	4,113
2	6 блоков подвыборки	4056	0,014	0,344
3	10 свёрточных	388800	0,778	7,633
4	10 блоков подвыборки	3200	0,006	0,295
5	10 свёрточных	80000	0,160	1,485
6	10 свертки	1000	0,002	0,115
		Итого	2,947	13,985

Таблица 5. Результаты исследования производительности макета СКЗ при распознавании одного изображения (количеством блоков свёртки — 32, подвыборки — 32).

№ слоя	Кол-во и тип используемых вычислительных блоков	Кол-во операций	Время выполнения, мс	
			расчетное	эксперимен.
1	6 свёрточных	596232	1,987	4,171
2	6 блоков подвыборки	4056	0,014	0,369
3	32 свёрточных	388800	0,243	0,702
4	32 блока подвыборки	3200	0,002	0,262
5	32 свёрточных	80000	0,050	1,311
6	10 свёрточных	1000	0,002	0,118
		Итого	2,298	6,933

Таблица 6. Результаты исследования производительности макета СКЗ при распознавании одного изображения (количеством блоков свёртки — 100, подвыборки — 32).

№ слоя	Кол-во и тип используемых вычислительных блоков	Кол-во операций	Время выполнения, мс	
			расчетное	эксперимен.
1	6 свёрточных	596232	1,987	4,31
2	6 блоков подвыборки	4056	0,014	0,391
3	32 свёрточных	388800	0,243	0,726
4	32 блока подвыборки	3200	0,002	0,274
5	100 свёрточных	80000	0,016	1,281
6	10 свёрточных	1000	0,002	0,221
		Итого	2,264	7,202

Из табл. 3-6 видим, что макет СКЗ по производительности уступает устройству из [4]. Однако его расчетная производительность не сильно отличается от расчетной производительности устройства из табл. 2.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Основная причина различия расчетной и найденной в результате эксперимента производительности макета СКЗ заключается в наличии большого числа операций чтения данных из внешней памяти в ПЛИС. Для уменьшения таких издержек в дальнейшем следует организовать кэширование промежуточных данных и весовых коэффициентов СНС, а также организовать пакетное чтение этих данных. Это, по-видимому, позволит увеличить производительность макета СКЗ и приблизить её к расчетной производительности.

Следует также отметить возрастание как расчетной, так и экспериментально полученной производительности макета СКЗ при увеличении количества вычислительных блоков. Более того, в табл. 5 и табл. 6 значения расчетного времени выполнения распознавания объекта на одном изображении имеют небольшое различие, а также очень близки к расчетному времени вычислений для устройства из табл 2. Из-за близости расчетного времени работы макета СКЗ в табл. 5 и в табл. 6 должны быть близкие значения времени вычислений, полученные экспериментальным путем. Однако этого не происходит из-за использования динамической памяти типа DDR3 (имеет циклы обновления памяти, дающие определенную погрешность замеров).

Все варианты макета СКЗ используют меньшее количество вычислительных блоков свёртки и подвыборки по сравниванию с устройством из [4] и, соответственно, требуют меньшее количество вычислительных ресурсов ПЛИС. Анализ данных табл. 3-6 по критериям минимального времени распознавания одного изображения и минимального потребления ресурсов ПЛИС позволяет считать оптимальным для исследуемой архитектуры СНС вариант макета СКЗ с 32 блоками свертки и 32 блоками подвыборки. Кроме того, можно предположить, что после сокращения издержек при обменах ПЛИС с внешней памятью варианты макета СКЗ с числом блоков из табл. 5 (32 блока свёртки и 32 подвыборки) и с числом блоков из табл. 6 (100 блоков свёртки и 32 блока подвыборки) будут иметь производительность, близкую к производительности устройства из [4].

Полное отсутствие простоя вычислительных блоков характерно только для варианта макета СКЗ, характеристики которого приведены в табл. 3. Однако он показывает невысокую производительность из-за малого числа используемых вычислительных блоков. Для остальных вариантов макета СКЗ количество простаивающих вычислительных блоков СНС сократилось по сравнению с устройством из [4]. Чтобы решить проблему оставшихся простаивающих вычислительных блоков необходимо разработать новые алгоритмы распараллеливания вычислений.

Экспериментально установлено, что макет СКЗ и устройство из [4] потребляют одинаковое количество электроэнергии - 5,1 Вт. Невысокое энергопотребление важно при создании мобильных СКЗ.

4. Заключение

В настоящее время при создании мобильных СКЗ все чаще применяют аппаратно реализованные СНС. Причем реализация СНС выполняется на современных ПЛИС, имеющих низкое энергопотребление и позволяющих проводить параллельные вычисления, что очень важно при удовлетворении требований к таким СКЗ по энергопотреблению и производительности.

Предложен метод организации вычислений в аппаратной СНС на ПЛИС, использующий унифицированные вычислительные блоки свёртки и подвыборки. Метод реализован в макете СКЗ, включающем аппаратную СНС из класса LeNet 5 на основе ПЛИС системы на кристалле Altera Cyclone V SX.

Проведены исследования точности распознавания рукописных цифр на изображениях с помощью этого макета. Показано, что он дает большую точность распознавания, чем устройство из работы [4] при одинаковой разрядности форматов чисел с плавающей запятой. Однако по производительности макет уступает этому устройству. Основной причиной этого является большое число обращений ПЛИС к внешней памяти. Для решения этой проблемы указаны возможные подходы по снижению числа обращений к такой памяти. Выявлена тенденция возрастания производительности макета СКЗ при увеличении количества вычислительных блоков. Использование наряду с критерием максимальной производительности макета критерия минимального потребления ресурсов ПЛИС позволяет считать оптимальным вариант макета СКЗ с 32 блоками свертки и 32 блоками подвыборки.

Исходя из полученных результатов исследований, можно считать, что предложенный метод организации вычислений в аппаратной СНС позволяет эффективно использовать вычислительные ресурсы ПЛИС. На наш взгляд, учитывая масштабируемость унифицированных вычислительных блоков, метод может быть перспективен при организации вычислений в аппаратно реализованных СНС более сложной архитектуры, чем архитектуры нейросетей класса LeNet 5.

5. Благодарности

Исследования были поддержаны грантом РФФИ № 18-47-700010 р_а.

6. Литература

[1] The MNIST database of handwritten digits [Электронный ресурс]. – URL: http://yann.lecun.com/exdb/mnist (дата обращения 26.07.2017)

[2] Toolflows for Mapping Convolutional Neural Networks on FPGAs: A Survey and Future Directions [Электронный ресурс]. – URL: https://arxiv.org/pdf/1803.05900.pdf (дата обращения 14.08.2018)

[3] Zhang, C. Optimizing fpga-based accelerator design for deep convolutional neural networks / C. Zhang, P. Li, G. Sun, Y. Guan, B. Xiao, J. Cong //Proceedings of the 2015 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays. – ACM, 2015. – P. 161-170. – DOI: 10.1145/2684746.2689060

[4] Зоев, И.В. Устройство на основе ПЛИС для распознавания рукописных цифр на изображениях/ И.В. Зоев, А.П. Береснев, Н.Г. Марков, А.Н. Мальчуков // Компьютерная оптика. – 2017. – Т. 41, № 6. – С. 938-949. – DOI: 10.18287/2412-6179-2017-41-6-938-949.

Об авторах

Зоев Иван Владимирович, аспирант отделения информационных технологий Томского политехнического университета, e-mail: ivz3@tpu.ru.

Марков Николай Григорьевич, д.т.н., профессор отделения информационных Томского политехнического университета, e-mail: markovng@tpu.ru

Береснев Алексей Павлович, магистрант отделения информационных технологий Томского политехнического университета, e-mail: apb3@tpu.ru.

Ягунов Тимофей Антонович, магистрант отделения информационных технологий Томского политехнического университета, e-mail: tay1@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Обнаружение спуфинг-атак на систему лицевой биометрии

И.А. Калиновский[1,2], Г.М. Лаврентьева[2,3]
kalinovskiy@speechpro.com|lavrentyeva@speechpro.com
[1]ООО «ЦРТ», Санкт-Петербург, Российская Федерация;
[2]ООО «ЦРТ-инновации», Санкт-Петербург, Российская Федерация;
[3]Университет ИТМО, Санкт-Петербург, Российская Федерация

В статье предлагается способ обнаружения спуфинг-атак на систему лицевой биометрии, основанный на глубоких нейронных сетях. Разработанная мультимодальная архитектура сверточной нейронной сети с механизмом внимания, позволяет достичь лучших результатов на бенчмарке CASIA-FASD, превосходя другие нейросетевые модели по качеству детектирования подложных образцов. Приводятся результаты ряда экспериментов по исследованию различных алгоритмов предобработки изображения для улучшения качества детектирования спуфинга. Показано, что использование метода нормализации Tan-Triggs позволяет значительно улучшить качество работы нейросетевого классификатора для рассматриваемой задачи.

Ключевые слова: *спуфинг-атака, биометрические системы, сверточные нейронные сети.*

Face anti-spoofing for biometric systems

I.A. Kalinovskiy[1,2], G.M. Lavrentyeva[2,3]
kalinovskiy@speechpro.com|lavrentyeva@speechpro.com
[1]"STC" Ltd., Saint Petersburg, Russian Federation;
[2]"STC-Innovations" Ltd., Saint Petersburg, Russian Federation;
[3]ITMO University, Saint Petersburg, Russian Federation

This paper proposes spoofing detection approach for face recognition systems, based on deep neural networks. The suggested architecture of multimodal convolutional neural network with attention mechanism allows to achieve state-of-the-art results for the CASIA-FASD benchmark database, exceeding other neural network models in terms of spoofing-detection quality. This research presents experimental results of the investigation of different image preprocessing techniques implemented to increase spoofing detection quality. Results confirm that Tan-Triggs normalization method leads to significant quality improvement of neural network classifier for the considered task.

Keywords: *spoofing attack, biometric systems, convolutional neural networks.*

1. Введение

Биометрические системы распознавания личности в последнее время получили широкое распространение. В основе подобных систем лежат технологии преобразования различных типов индивидуальных биометрических характеристик (отпечатки пальцев, радужная оболочка глаза, голос, лицо и др.) в цифровой код, который используется для решения задач верификации и идентификации профиля человека среди миллионов записей в базе данных. Наиболее простой с точки зрения считывания биометрических дынных является технология лицевой биометрии [11]. Современные системы распознавания личности по лицу способны с высокой точностью работать даже в неконтролируемых условиях. Однако это приводит к повышенному риску взлома, так как для прохождения верификации, системе достаточно предоставить фотографию, сделанную на обычную камеру или взятую из открытых источников. В связи с этим возникает ряд задач по предотвращению попыток подмены биометрических данных, обычно называемых спуфинг-атаками.

Спуфинг-атака на лицевую биометрическую систему может быть осуществлена разными способами. Наиболее эффективным из них является прямая загрузка фотографии в систему, однако для этого злоумышленникам требуется получить доступ к программному обеспечению, что значительное усложняет атаку. В случае доступности ответа системы в виде степени схожести или вероятности верификации, возможно осуществление атаки непосредственно на алгоритмы построения биометрического шаблона, которые, как правило, основаны на глубоких сверточных нейронных сетях (СНС) [11]. Как известно СНС уязвимы к так называемым «состязательным атакам» (adversarial attack) [21], что позволяет злоумышленникам использовать специальным образом нанесенный макияж для существенного увеличения вероятности ложного срабатывания системы. Наконец, наиболее простым способом осуществления атаки является атака на уровне сенсора с использованием фотографии зарегистрированного в системе пользователя.

В данной работе предлагается способ защиты от последнего вида спуфинг-атак на системы контроля и управления доступом (СКУД). Ввиду большой популярности социальных сетей весьма просто заполучить фотографию какого-либо человека, работающего на предприятии. При этом она может демонстрироваться сенсору как в распечатанном виде, так и на смартфоне или планшете (рис. 1). Далее будет рассмотрена единая модель, позволяющая детектировать одновременно все варианты спуфинга на уровне сенсора используя только одно изображение при малом количестве ложных отказов.

Рис 1. Примеры спуфинг-атак из базы CASIA-FASD [22]

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2. Обзор методов обнаружения спуфинга

Проблема детектирования спуфинг-атак или обнаружения витальности (liveness detection) применительно к системам лицевой биометрии, приобрела популярность у исследователей в середине 2000-х готов [15]. К настоящему времени предложено множество методов ее решения, которые условно разделяются на две группы: активные и пассивные. Активные методы запрашивают от пользователя совершения определенного действия, например: улыбнуться, моргнуть, наклонить или повернуть голову и др. [2, 18]. Пассивные методы обычно используют для анализа только одного изображения, по которому непосредственно строится биометрический шаблон [17]. В связи с этим они удобнее в использовании, а также позволяют исключить ситуацию демонстрации фотографии в промежуток между процессами определения витальности лица и верификации.

Большое количество работ, посвященных алгоритмам пассивного обнаружения спуфинг-атак, основаны на анализе текстуры области лица. Для описания текстуры обычно используются различные разновидности LBP [5], SURF [3], HOG-LPQ [12] и других дескрипторов, а в качестве классификатора – SVM. При этом исследуется влияние различных цветовых пространств на точность обнаружения, а также применяется анализ в частотной области [8]. Несмотря на хорошее качество подобных алгоритмов, достигаемое на стандартных бенчмарках [22], оно существенно ухудшается при изменении условий регистрации изображений.

В ряде работ изучается возможность использования оптических эффектов, создаваемых камерой (муар [13], расфокусировка [7], дисторсия [6]). Недостатком подобных методов является необходимость в точной калибровке алгоритмов для конкретного оборудования. В тоже время использование дополнительного специализированного оборудования, например камер светового поля или 3D-сенсоров, позволяет добиться наиболее высокой степени защиты от спуфинк-атак, но ценой удорожания стоимости СКУД для конечного пользователя.

В последние годы для решения рассматриваемой задачи стали активно применяться методы глубокого обучения. В [20] использовалась СНС типа AlexNet для бинарной классификации реальное лицо/спуфинг-атака, а также исследовалось влияние размера контекста, захватываемого при выравнивании изображения лица, на качество такой классификации. Авторы [9] использовали SVM для классификации высокоуровневых признаков, полученных с последних слоев дообученной сети VGG-Face. В [19] использовалась комбинация СНС и рекуррентной сети типа LSTM для классификации последовательности кадров. В [1] была предложена оригинальная комбинация «patched-based» СНС, классифицирующей группу локальных участков изображения, и «depth-based» СНС, которая генерировала плотную карту глубины для каждого пикселя. Для обучения «depth-based» модели карта глубины реальных изображений лиц генерировалась с помощью нежесткой 3D-модели, а для поддельных задавалась как нулевая. В [10] авторы распространили предложенный подход на обработку последовательности кадров [19].

Несмотря на значительный прогресс в обнаружении витальности лица, эта задача по-прежнему остается не решенной в общем случае. Сверхвысокая плотность пикселей и естественная цветопередача современных дисплеев делают отображаемое изображение лица почти не отличимым от реального. В связи с этим, существует потребность в совершенствовании предложенных алгоритмов обнаружения спуфинга и в разработке новых подходов.

3. Мультимодальная сверточная нейронная сеть с механизмом внимания

В настоящее время глубокие сверточные нейронные сети являются стандартным блоком практически во всех задачах, связанных с обработкой изображений. Современные библиотеки для машинного обучения значительно ускоряют разработку новых нейросетевых архитектур, что приводит к постепенному усложнению их вычислительного графа. Недавние работы в области глубокого обучения продемонстрировали эффективность механизма внимания для улучшения работы СНС в задачах распознавания образов, генерации подписей к изображениям и др. [14]. Блоки внимания позволяют выделять наиболее информативные участки карт признаков на разных уровнях нейронной сети, специфичных для каждого конкретного класса. В основе предлагаемого подхода к обнаружению спуфинг-атак лежит сверточная нейронная сеть с модулями внимания, имеющая несколько ветвей для обработки различных входных данных (таблица 1).

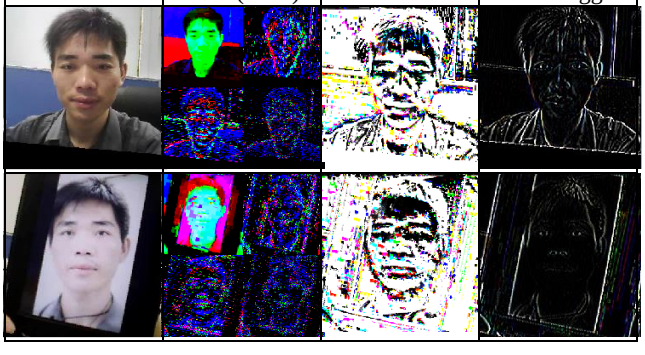
3.1 Мультимодальная архитектура

Для обучения сверточных нейронных сетей обычно используются непосредственно RGB-каналы изображения. В тоже время во множестве предшествующих работ по обнаружению витальности лица была продемонстрирована эффективность использования различных текстурных дескрипторов для кодирования лицевой области. Необходимость обобщать различия в освещенности и позе лиц для двух визуально похожих классов, одновременно с поиском мелкомасштабных текстурных отличий, усложняет задачу оптимизации нейронной сети в контексте рассматриваемой задачи. В связи этим, в настоящей работе предлагается использовать ряд процедур предобработки исходных изображений с целью извлечения наиболее информативных текстурных характеристик.

Для совместной обработки полученных таким образом различных карт-признаков, предлагается архитектура СНС, имеющая несколько идентичных входов (табл. 2). Обработка сетью информации, поступающей на каждый вход, осуществляется независимо вплоть до определенного уровня глубины, после чего полученные карты конкатенируется в единый пул. При этом каждая ветка может включать один или несколько модулей внимания.

В ходе работы были проведены эксперименты с различными алгоритмами «ручного» выделения признаков: цветовые пространства HSV и YCbCr, LBP-дескрипторы, вельветы Хаара и Добеши, алгоритм нормализации Tan-Triggs [16] (табл. 1). Результаты исследования зависимости качества обнаружения спуфинг-атак от алгоритма предобработки входных изображений обсуждаются в разделе 4.

Таблица 1. Иллюстрация вариантов предобработки данных

RGB	Haar (HSV)	LBP	Tan-Triggs
			

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Таблица 2. Архитектура СНС для обнаружения спуфинга

Слой	Фильтр / шаг	Размер выхода
Input		192×192×3 (*N*)
N веток для обработки каждого входа		
Conv-1.1 (SELU, BN)	3×3/1	192×192×32
Conv-1.2	3×3/2	96×96×32
Conv-2.1 (SELU, BN)	3×3/1	96×96×64
Conv-2.2	3×3/2	48×48×64
Conv-3.1 (SELU, BN)	3×3/1	48×48×128
Attention-1 (*K*=4)		1×1×2
Conv-3.2	3×3/2	24×24×128
Основная ветка графа		
Concatenation		24×24×(128×*N*)
Conv-4 (SELU, BN)	3×3/1	24×24×256
Attention-2 (*K*=4)		1×1×2
MaxPooling-1	3×3/2	12×12×256
Conv-5 (SELU, BN)	3×3/1	12×12×512
Attention-3 (*K*=4)		1×1×2
MaxPooling-2	3×3/2	6×6×512
Conv-6 (SELU, BN)	3×3/1	6×6×1024
Dropout-1		
FC-1		1×1×2048
Dropout-2		
FC-2		1×1×2
Add (attentions layers)		1×1×2

3.2 Механизм внимания

В данной работе применятся простой механизм внимания, описанный в [14]. Авторы предложили универсальную архитектуру модуля внимания, который может быть подключен к любому уровню сети. При этом для его обучения не требуются дополнительные метки. В состав модуля входят три основных блока:

1) слой внимания, состоящий из сверточного слоя с одним выходом и ядром 1×1, а также softmax-слой, который производит вероятностную карту;
2) выходной полносвязный слой, генерирующий гипотезу о классе на основе взвешенных с выходом softmax-слоя входных карт признаков;
3) ворота (gates), контролирующие уровень выходного сигнала от данного модуля.

В данной работе используется модифицированная архитектура модуля, представленная в таблице 3. К одному участку нейронной сети может быть подключено одновременно *K* модулей внимания, при этом их выходы суммируются.

Таблица 3. Архитектура модуля внимания

Слой	Фильтр / шаг	Размер выхода
Input		*S*×*S*×*D*
Conv-1	3×3/1	*S*×*S*×1
Softmax-1		*S*×*S*×1
Tile		*S*×*S*×*D*
Input×Tile		*S*×*S*×*D*
GlobalAvegraePooling		1×1×*D*
Ветка для вычисления значений ворот		
FC-1		1×1×256
FC-2		1×1×2
Tanh-1		1×1×2
Softmax-2		1×1×2
Ветка для вычисления входа		
FC-1		1×1×256
FC-2		1×1×2
FC-2×Softmax-2		1×1×2

3.3 Выравнивание и аугментация изображений

Во всех проведенных экспериментах изображения лиц были предварительно выровнены по положению глаз и обрезаны до размера 192×192 пикселя (соотношение между размером лица и шириной «кропа» определяется коэффициентом α). Для расширения обучающей выборки спуфинга применялся ряд процедур аугментации, имитирующих различный эффекты: размытие гауссовским фильтром с размером из диапазона [11; 21], изменение разрешения с коэффициентом из диапазона [0,1; 0,2] и последующим восстановлением до исходного, добавление шума с нормальным распределением (μ=25, σ=25).

4. Результаты экспериментов

В этом разделе приводятся результаты экспериментального анализа предложенного способа детектирования спуфинг-атак на наборе данных CASIA-FASD [22]. В записи этой базы принимали участие 50 человек, для каждого из которых были сделаны 12 коротких видеороликов (3 реальных и 9 поддельных). Спуфинг-атаки включают демонстрацию фотографии на бумаге и планшете. Обучающее подмножество включает 240 записей (20 человек), тестовое – 360 записей (30 человек). В соответствии с протоколом [1] качество моделей оценивается с помощью метрики HTER=0,5×(FAR+FRR), при этом классификация видео осуществляется по 30 случайно выбранным кадрам.

Эксперименты были проведены на фреймворке TensorFlow: оптимизатор – Adadelta; начальная скорость обучения – 0,3; уменьшение скорости – экспоненциальное с шагом 500 и коэффициентом 0,01; размер партии – 32; количество итераций – 20000.

В таблице 4 приведены результаты тестирования предложенной модели СНС с одним входом. Классификация изображений при α = 1 является наиболее сложной задачей, т.к. доступна информация только о текстуре лицевой области. Лучшее качество достигается при α = 3, когда границы фотографии и рамки планшета попадают в область интереса. При этом включение модулей внимания приводит к повышению HTER только в этом случае, что подтверждает высокую информативность контекста вокруг лица для рассматриваемой задачи. В тоже время, даже если изображение лица занимает весь угол обзора камеры (рис. 1), спуфинг-атака по-прежнему может быть идентифицирована. Дальнейшие эксперименты были проведенны с α = 3.

Таблица 4. Тестирование на базе CASIA-FASD (HTER, %)

α	1	1,5	2	3
off	12,34	11,48	9,82	4,26
on	12,96	12,41	9,82	3,33

***** off/on attention mechanism

В таблице 5 приведены результаты тестирования алгоритмов предобработки исходных изображений и их комбинаций. В соответствии с полученными данными можно сделать следующие выводы:

1) Среди различных вариантов цветовых пространств лучшим является HSV. Это объясняется свойством данного пространства, в котором тон и яркость пикселей разделяются на разные каналы, поэтому оно часто используется для выделения участков кожи.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2) Популярные LBP-дескрипторы плохо поддаются классификации сверточной нейронной сетью.
3) Лучшими по соотношению качество/производительность являются вельветы Хаара.
4) Наиболее высокое качество для СНС с одним входом достигается при применении нормализации Tan-Triggs, которое устраняет вариации в освещенности и усиливает градиенты на границах. Однако вычислительная сложность такого преобразования значительно выше, чем у остальных.
5) Наилучшим решением является сочетание нормализации Tan-Triggs с вельветами Хаара. При этом дальнейшее увеличение числа входов не приводит к уменьшению HTER.

Таблица 5. Сравнение эффективности различных алгоритмов предобработки на базе CASIA-FASD (α = 3)

Преобразование	HTER. %
RGB	8,89
HSV	7,96
YCbCr	13,33
LBP	10,19
Haar	6,3
Tan-Triggs	1,85
Tan-Triggs, Haar	**1,67**
Tan-Triggs, D4	3,15
Tan-Triggs, LBP	5,0
Tan-Triggs, LBP, Haar	4,82
Tan-Triggs, RGB, LBP, Haar	7,59

Сравнение предложенного подхода к решению задачи обнаружения витальности лица с другими современными алгоритмами, основанными на нейронных сетях, приведено в таблице 6. Разработанная мультимодальная сверточная нейронная сеть с механизмом внимания в сочетании с нормализацией Tan-Triggs и вельветами Хаара достигает state-of-the-art результатов на бенчмарке CASIA-FASD, превосходя другие алгоритмы.

Таблица 6. Сравнение нейросетевых методов обнаружения витальности лица на базе CASIA-FASD

Алгоритм	EER, %	HTER, %
Fine-tuned VGG-Face [9]	5,20	–
DPCNN [9]	4,50	–
Yang et al. [20]	4,92	–
LSTM-CNN [19]	5,17	5,93
Patch-based CNN [1]	4,44	3,78
Depth-based CNN [1]	2,85	2,52
Fusion [1]	2,67	2,27
Волкова и др.[23]	–	8,71
Предложенный	–	**1,67**

5. Заключение

В данной работе предложен эффективный способ обнаружения спуфинг-атак на системы лицевой биометрии по одной фотографии, превосходящий качество современных нейросетевых алгоритмов на бенчмарке CASIA-FASD. Однако более сложные варианты спуфинга, основанные на 3D-реконструкции лица, не были рассмотрены. Идентификация подложной 3D-маски высокого качества явятся трудной задачей и, возможно, может быть решена только с использованием видеопоследовательности, поэтому предложенный подход требует дальнейшего исследования.

Исследование проводится в рамках соглашения о предоставлении субсидии № 14.578.21.0189 (RFMEFI57816X0189)

6. Литература

[1] Atoum, Y. Face anti-spoofing using patch and depth-based CNNs, Biometrics, 2017
[2] Bharadwaj, S. Computationally efficient face spoofing detection with motion magnification, CVPRW, 2013
[3] Boulkenafet, Z. Face anti spoofing using speeded up robust features and fisher vector encoding, IEEE Signal Processing Letters, vol. 24, issue: 2, 2017
[4] Erdogmus, N. Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect, in Proc. IEEE Biometrics, Theory, 2013, pp. 1-6
[5] Kim, G. Face liveness detection based on texture and frequency analyses. In International Conference on Biometrics, 2012, pp. 67–72
[6] Kim, I. Face spoofing detection with highlight removal effect and distortions. systems, SMC, 2016
[7] Kim, S. Face Liveness Detection Using Defocus, Sensors, vol. 15, 2015, pp. 1537–1563
[8] Li, J. Live face detection based on the analysis of fourier spectra, Biometric technology for human identification, Proc. SPIE, vol. 5404, 2004, pp. 296-303
[9] Li, L. An original face anti-spoofing approach using partial convolutional neural network. In IPTA, 2016.
[10] Liu Y. Learning deep models for face anti-spoofing: binary or auxiliary supervision, CVPR, 2018
[11] Mahmood, Z. A review on state-of-the-art face recognition approaches.. Fractals, 2017
[12] Mohan, K. Object face liveness detection with combined HOGlocal phase quantization using fuzzy based SVM classifier, 2017
[13] Patel, K. Live face video vs. spoof face video: use of moire patterns to detect replay video attacks. ICB, 2015, pp. 8-15
[14] Rodriguez, P. A painless attention mechanism for convolutional neural networks, ICLR, 2018, pp. 1-8
[15] Souza, L. How far did we get in face spoofing detection?. arxiv:1710.09868, 2017, pp. 1-48
[16] Tan X. and Triggs B., Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process, 2010, pp. 1635–1650
[17] Wang S.-Y. Face liveness detection based on skin blood flow analysis, 2017
[18] Wang, T. Face liveness detection using 3D structure recovered from a single camera, international conference on biometrics, 2013.
[19] Xu, Z. Learning temporal features using LSTM-CNN architecture for face anti-spoofing. ACPR, 2015, pp. 41-45
[20] Yang, J. Learn convolutional neural network for face anti-spoofing. arXiv preprint arXiv:1408.5601, 2014
[21] Yuan, X. Adversarial examples: attacks and defenses for deep learning, arXiv:1712.07107, 2017, pp. 1-22
[22] Zhang, Z. A face antispoofing database with diverse attacks. Proceedings of 5th IAPR International Conference on Biometrics, IEEE (2012), pp. 26-31
[23] Волкова, С. Применение сверточных нейронных сетей для решения задачи противодействия атаке спуфинга в системах лицевой биометрии / С. Волкова, Ю. Н. Матвеев // Научно-технический вестник информационных технологий, механики и оптики, Т.17, № 4, с. 702-710, 2017.

Об авторах

Калиновский Илья Андреевич, к.т.н., научный сотрудник ООО «ЦРТ-инновации». Его e-mail kalinovskiy@speechpro.com.

Лаврентьева Галина Михайловна, аспирантка каф. РИС Университета ИТМО, научный сотрудник ООО «ЦРТ-инновации». Ее e-mail lavrentyeva@speechpro.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Распределённый алгоритм сопровождения людей в видео

Купляков Д.А.[1], Шальнов Е.В.[2], Конушин В.С.[3], Конушин А.С.[4]
denis.kuplyakov@graphics.cs.msu.ru|eshalnov@graphics.cs.msu.ru|vadim@tevian.ru|ktosh@graphics.cs.msu.ru
[1,2,4]Факультет вычислительной математики и кибернетики,
Московский государственный университет им. М.В. Ломоносова, Москва, Россия;
[1,3]ООО «Технологии видеоанализа», Москва, Россия

В работе рассматривается задача сопровождения людей в видеопотоке с целью их подсчёта. Современные системы видеонаблюдения такие, как система видеонаблюдения Москвы, содержат сотни тысяч камер. Использование современных методов, разрабатываемых для работы на одном компьютере с дорогим графическим ускорителем, не является экономически целесообразным для систем подобного масштаба. В работе предлагается распределённый алгоритм сопровождения. Он позволяет за счёт детекции на разреженном множестве кадров сократить число необходимых вычислительных ресурсов. Детекция производится на серверах в центре обработки данных, а видеопоток обрабатывается локальными для камер узлами. Представлена методика оценки качества подсчёта людей. Экспериментальная оценка показала, что предложенный алгоритм позволяет достичь приемлемого качества при частоте детекции $4/3$ Гц.

Ключевые слова: *компьютерное зрение, сопровождение в видео, подсчёт людей.*

Distributed pedestrian video tracking algorithm

Kuplyakov D.[1], Shalnov E.[2], Konushin V.[3], Konushin A.[4]
denis.kuplyakov@graphics.cs.msu.ru|eshalnov@graphics.cs.msu.ru|vadim@tevian.ru|ktosh@graphics.cs.msu.ru
[1,2,4]Department of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia;
[1,3]«Video Analysis Technologies» LLC, Moscow, Russia

Problem of tracking people on video with the purpose of their counting is considered. Modern video surveillance systems, like Moscow's video surveillance system, consist of hundreds of thousands of cameras. The use of modern tracking methods, developed to work on a single computer with expensive graphics accelerator, is not economically feasible for systems of such scale. The paper proposes a distributed tracking algorithm. It allows to reduce amount of neccessary computational resources by detection on rare frames. Detection is performed on servers in a data center, and the video stream is processed by local to cameras nodes. A methodology for assessing the quality of people counting is presented. The experimental evaluation shows that the proposed algorithm allows to reach an acceptable quality with a detection frequency $4/3$ Hz.

Keywords: *computer vision, video tracking, people counting.*

1. ВВЕДЕНИЕ

Практически важной задачей является подсчёт числа людей, прошедших через определённые зоны общественной инфраструктуры, такие как пешеходные переходы, тротуары, площади и т.д. Одним из возможных путей её решения является сопровождение людей в видеопотоке, то есть построение траекторий их движения. *Траектория* однозначно соответствует человеку и содержит его положение на всех кадрах, где он представлен. Для подсчёта людей обычно на кадре фиксируется *сигнальный отрезок*. При пересечении траекторией человека данного отрезка фиксируется один проход, что позволяет подсчитывать людей.

В работе рассматривается задача полностью автоматического подсчёта людей в видео, снятого неподвижной камерой. На вход алгоритма поступает видеопоток $\{F_i\}_{i=1}$ кадров, полученных с неподвижной камеры и сигнальный отрезок, заданный упорядоченной парой точек (L_a, L_b) в кадре. На выходе — поток событий пересечения сигнального отрезка $\{E_i\}_{i=1}$. Событие $E_i = (k_i,\, r_i,\, d_i)$ описывается номером кадра k_i, на котором произошло пересечение сигнального отрезка, прямоугольником d_i, ограничивающим изображение человека на кадре в момент пересечения, и направлением пересечения сигнального отрезка r_i. Данная задача рассматривалась также в работе [16].

Современные методы сопровождения людей используют подход к сопровождению через обнаружение. Будем рассматривать только алгоритмы способные обрабатывать видеопоток. Их можно классифицировать на 2 группы по временному контексту, используемому для построения результатов на определённом кадре. Первые используют только предыдущие кадры [2], [12], [14], [15], вторые учитывают информацию со следующих кадров, используя метод скользящего окна [1], [3], [6], [11]. Последние позволяют добиться лучшего качества, но результат генерируется с некоторой задержкой.

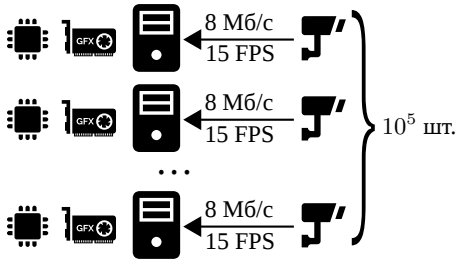
Вторым ключевым аспектом является детектируемый объект. Наиболее популярным является обнаружение всего человека в кадре [2], [3], [14], [15]. В этом случае перекрытия могут существенно снизить качество результата. Обнаружение голов людей [1], [6], [11] является более стойким к перекрытиями, но детекторы голов обычно производят много ложных обнаружений. Третья группа методов использует детекторы на осно-

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

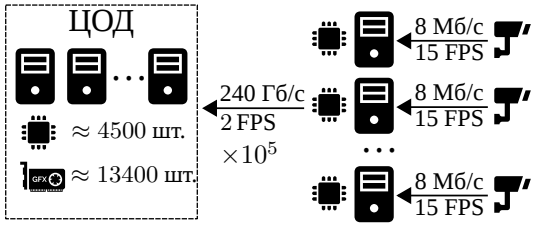
ве деформируемых моделей частей [12] и одновременно учитывает информацию о их положении.

Третьим аспектом является выбор алгоритма объединения обнаружений в траектории. Данную задачу сводят к минимизации функции энергии модели объединения обнаружений. Алгоритмы, обрабатывающие видеопоток, обычно учитывают в функции энергии связи только между соседними обнаружениями траектории. Задачу тогда можно свести к наращиванию траекторий с помощью их сопоставления с обнаружениями со следующего кадра Венгерским алгоритмом [5] (работы [2], [14], [15]), либо приближённо жадными алгоритмами [12]. Для обработки с помощью скользящего окна возможно использование алгоритма на основе модифицированного поиска потока минимальной стоимости [7]. Модель может быть усложнена добавлением типа траектории для выявления ложных обнаружений. В этом случае оптимизация выполняется с помощью метода Монте-Карло [1], [6], [11].

Члены функции энергии могут использовать различные признаки, извлекаемые из видео: нейросетевая реидентификация [14], [15], визуальное сопровождение [1], [6], [11], оптический поток [3]. Некоторые алгоритмы [1], [3], [6], [11] строят *треклеты*, части траектории в коротком отрезке видео, что позволяет строить более точные члены функции энергии для модели. В некоторых работах используется калибровка камеры, что позволяет находить координаты людей в пространстве. В статье [10] представлен метод автоматического получения матрицы калибровки по обнаружениям. Калибровка может быть использована для повышения качества детекции [17].



(*а*) Обработка каждого видеопотока своим компьютером.



(*б*) Обработка видеопотоков локально с размещением детекторов в ЦОД.

Рис. 1. Оценочное сравнение конфигураций систем для подсчёта людей.

Современные алгоритмы сопровождения разработаны для работы на одном компьютере с мощным графическим ускорителем. Он требуется для осуществления детекции с помощью нейронных сетей, что позволяет добиться высокого качества обнаружений. Так как обнаружение необходимо производить для всех кадров видео, а качественные алгоритмы детекции вычислительно ёмкие, потребуется число графических ускорителей равное числу обрабатываемых видеопотоков. Данный подход делает применение системы экономически необоснованным (рис. 1, *а*)

Учитывая перечисленные выше недостатки, в данной работе предлагается распределённый алгоритм сопровождения для системы, состоящей из подключённых к камерам видеонаблюдения локальных вычислительных узлов и ЦОД с серверами для обнаружения людей (рис. 1, *б*). На ресурсах ЦОДа выполняется обнаружение людей, но лишь на разреженном множестве кадров. Это позволит уменьшить число используемых графических ускорителей и снизить требование на пропускную способность канала. Чтобы компенсировать отсутствие информации от детектора на остальных кадрах видео, выполняющаяся на локальных узлах часть алгоритма сопровождения будет производить дополнительный анализ видеопотока. Эта часть алгоритма сопровождения должна быть вычислительно простой, чтобы снизить стоимость локальных узлов.

Предположения, используемые для сравнения на рис. 1: один графический ускоритель производит детекцию с частотой $15\,\text{Гц}$; предложенный алгоритм для варианта (б) будет производить детекцию с частотой $2\,\text{Гц}$; объём информации одного пересылаемого детектору кадра — 150 КБайт; один центральный процессор может обеспечить работу детекторов на трёх графических ускорителях.

2. ПРЕДЛОЖЕННЫЙ АЛГОРИТМ

Предложенный алгоритм представляет собой модификацию алгоритма [2]. Он состоит из 4-х этапов: 1) детекция людей на ключевых кадрах; 2) построение матрицы стоимостей (энергий) связывания текущего набора траекторий с новыми обнаружениями; 3) поиск оптимального связывания: каждое обнаружение присоединяется к текущей траектории, либо даёт начало новой траектории; 4) поиск пересечений сигнального отрезка. Ниже подробней описаны этапы, претерпевшие модификации.

2.1. Детекция

Детекция производится на *ключевых кадрах* видеопотока: F_1, F_{1+step}, $F_{1+2\,step}$, ... Результатом детекции является набор ограничивающих изображения людей в кадре прямоугольников. Для решения этой подзадачи используется детектор на основе подхода FasterRCNN [9] с использованием свёрточных нейронных сетей. Данный этап является единственным, выполняющимся на удалённом сервере.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

2.2. Построение матрицы стоимостей

Построение матрицы стоимостей в базовом алгоритме производится на основе предсказания положения человека фильтром Калмана и вычислением IOU для предсказанного положения и нового обнаружения. Пары траектории и обнаружения, для которых получившееся значение ниже заданного порога не могут быть сопоставлены. IOU (intersection over union) — функция, вычисляющая отношение площади пересечения прямоугольников к площади их объединения.

В данной работе предлагается внести описанные в следующих подразделах модификации в процедуру вычисления стоимости связывания.

2.2.1. Оценка скорости с помощью визуального сопровождения

Для уменьшения требуемого объема вычислительных ресурсов необходимо уменьшать частоту детекции. При больших значениях $step$ качество работы базового алгоритма сильно ухудшается (см. рис. 3). Это объясняется тем, что адекватную оценку скорости базовый алгоритм получает из фильтра Калмана только при накоплении достаточной статистики. Несколько первых обнаружений траектории должны сопоставиться друг с другом по значению IOU. Увеличение $step$ приводит к тому, что изображение человека в кадре перемещается между ключевыми кадрами сильно. Следовательно, значение IOU становится околонулевым и сопоставления не происходит.

Предлагается решить данную проблему, оценив скорость человека v_x, v_y и включив её в видимое состояние фильтра Калмана. Для измерения скорости воспользуемся алгоритмом визуального сопровождения. Он позволит найти изображение человека с кадра $1 + k\,step$ на кадре $1 + k\,step + \Delta$, где $\Delta \in \mathrm{Z}_+$ — параметр алгоритма. Пусть положение человека, заданное координатами левого верхнего угла и размерами ограничивающего прямоугольника, на первом кадре (x, y, w, h), а на втором $(\tilde{x}, \tilde{y}, \tilde{w}, \tilde{h})$. Оценим скорости через перемещение центра нижней границы ограничивающего прямоугольника:

$$\begin{aligned} v_x &= \left((\tilde{x} + \frac{\tilde{w}}{2}) - (x + \frac{w}{2}) \right) \frac{step}{\Delta}, \\ v_y &= \left((\tilde{y} + \tilde{h}) - (y + h) \right) \frac{step}{\Delta}. \end{aligned} \tag{1}$$

Среди современных методов визуального сопровождения, проанализированных по результатам конкурса VOT2016 [4] и имеющих открытую реализацию, был выбран метод ASMS [13]. Он обладает небольшой вычислительной сложностью, что позволяет осуществлять визуальное сопровождение в режиме реального времени.

Выбранный алгоритм визуального сопровождения локализует изображение человека в кадре не так, как это делает детектор. Это приводит к наличию постоянной ошибки в определении скорости. Предлагается решить данную проблему, применив визуальное сопровождение к паре кадров с номерами $1 + k\,step$ и $1 + k\,step + \varepsilon$, где $\varepsilon \in \mathbb{Z},\ \varepsilon < \Delta$ — параметр алгоритма. Соответствующее предсказание обозначим через $(\bar{x}, \bar{y}, \bar{w}, \bar{h})$. Тогда оценим скорость через:

$$\begin{aligned} v_x &= \left((\tilde{x} + \frac{\tilde{w}}{2}) - (\bar{x} + \frac{\bar{w}}{2}) \right) \frac{step}{\Delta - \varepsilon}, \\ v_y &= \left((\tilde{y} + \tilde{h}) - (\bar{y} + \bar{h}) \right) \frac{step}{\Delta - \varepsilon}. \end{aligned} \tag{2}$$

Оба предсказания $(\tilde{x}, \tilde{y}, \tilde{w}, \tilde{h})$ и $(\bar{x}, \bar{y}, \bar{w}, \bar{h})$ получены одним алгоритмом визуального сопровождения. Возникающая из-за различий в локализации ошибка таким образом будет устранена.

2.2.2. Использование оценки скорости для новых обнаружений

Заметим, что при оценке скорости оценка между кадрами с разницей в номерах $\Delta - \varepsilon$ экстраполируется на разницу $step$, исходя из модели равномерного прямолинейного движения. Траектории людей в кадре могут быть с достаточной точностью описаны данной моделью только на коротких интервалах времени. Также увеличение Δ приводит к ошибкам в визуальном сопровождении. Это приводит к тому, что последующее увеличение параметра $step$ ухудшает качество алгоритма (см. рис. 3).

Данную проблему возможно частично устранить. Для этого будем предсказывать положение человека не на $step$ кадров вперед, а на $step/2$. Оценки скоростей через визуальное сопровождение (2) изменятся на:

$$\begin{aligned} v_x &= \left((\tilde{x} + \frac{\tilde{w}}{2}) - (\bar{x} + \frac{\bar{w}}{2}) \right) \frac{step/2}{\Delta - \varepsilon}, \\ v_y &= \left((\tilde{y} + \tilde{h}) - (\bar{y} + \bar{h}) \right) \frac{step/2}{\Delta - \varepsilon}. \end{aligned} \tag{3}$$

Для новых обнаружений на очередной итерации алгоритма будем предсказывать положение на $step/2$ кадров назад. Для этого достаточно применить визуальное сопровождение к одному из предыдущих кадров. Будем вычислять новую функцию стоимости как IOU между предсказанным фильтром Калмана на $step/2$ кадров вперед положением траектории и предсказанным на $step/2$ кадров назад положением обнаружения.

Данный подход позволяет сократить в два раза интервал, на который экстраполируется оценка скорости. Однако теперь алгоритм визуального сопровождения необходимо запускать четыре раза для каждого обнаружения, вместо двух. Число запусков можно сократить до трёх. Для этого нужно использовать 0 в качестве значения параметра ε.

2.3. Поиск событий

Данный этап не присутствует в базовом алгоритме, так как он решает задачу сопровождения, а не подсчёта людей. Чтобы получить искомые события, необходимо для каждой траектории найти её пересечения с сигнальным отрезком. Для этого рассматривается ломанная линия, соответствующая нижней середине обнаружений (рис. 2).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Чтобы подавить ложные события вызванные незначительными колебаниями обнаружений траектории, для сигнального отрезка вводится ширина W_L. Пересечение траекторией сигнального отрезка приводит к генерации события, только если с момента предыдущего события для этой траектории (либо с начала траектории, если не было событий) существует обнаружение, находящиеся на отдалении не менее $W_L/2$ с соответствующей стороны от прямой, проходящей через сигнальный отрезок.

2.3.1. Экстраполяция траекторий

Во время работы алгоритма из-за перекрытий, неточностей детекции или каких-то других ошибок человек может начать сопровождаться после момента пересечения сигнального отрезка. Если траектория началась около сигнального отрезка и продолжилась в сторону от него, разумно предположить, что ранее произошло пересечение сигнального отрезка. Чтобы учесть описанную ситуацию, предлагается произвести предобработку каждой траектории путем её экстраполяции на одно обнаружение в обе стороны. Оценив скорость от второго к первому обнаружению траектории, добавим в начало траектории обнаружение полученное сдвигом первого обнаружения на вектор, соответствующий перемещению с полученной скоростью за число кадров Δ_{ex}. Аналогично поступим для последних двух обнаружений траектории, добавив новое обнаружение в конец. Если траектория началась не позднее, чем Δ_{ex} кадров с момента пересечения сигнального отрезка человеком, и он двигался с постоянной скоростью, то пропущенное событие будет восстановлено.

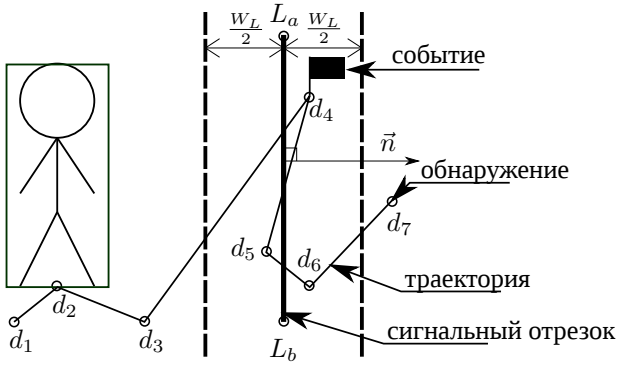


Рис. 2. Пример результатов работы алгоритма поиска событий. Использование значения ширины сигнального отрезка позволило избежать ложных событий для обнаружений d_5 и d_6.

3. ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА

3.1. Методика исследования качества

3.1.1. Наборы данных

Для тестирования базового алгоритма и его модификации были использованы наборы с ручной разметкой траекторий людей для задачи сопровождения. Всего использовалось 22 видеопоследовательности общей продолжительностью 6 часов из коллекции конкурса MOTChallenge [8], собранной ООО «Технологии видеоанализа» коллекции и набора Towncentre [1]. Для каждого набора вручную было установлено положение сигнального отрезка. Далее с помощью алгоритма поиска событий (разд. 2.3) была получена эталонная разметка $\widehat{E}$ для задачи подсчёта людей. Получившееся число эталонных событий во всех видеопоследовательностях — 2657.

3.1.2. Сопоставление событий

Среди результирующих событий могут быть верные и ложные. Ложные события не имеют соответствия в эталонной разметке. Будем считать, что событие из выходных данных E_i соответствует событию в эталонной разметке $\widehat{E}_j$, если они соответствуют переходу через сигнальный отрезок одного и того же человека в один момент времени. Заметим, что эти события могут быть на разных кадрах из-за разницы локализации изображения человека ограничивающим прямоугольником и из-за того, что предложенный алгоритм может выдать событие только для ключевого кадра.

Для сопоставления событий из выходных данных и эталонной разметки предлагается воспользоваться жадным алгоритмом с отсечением сопоставлений по двум порогам: Δ_f и IOU_e. Первый задаёт допустимую разницу во времени между событиями, второй задаёт допустимую погрешность в определении позиции человека в кадре. Жадный алгоритм также получает на вход разметку для задачи сопровождения с информацией о сопоставлении эталонных событий и траекторий, т.к. событие E_i по положению необходимо сопоставлять не с положением $\widehat{E}_j$, а с положением соответствующей $\widehat{E}_j$ траектории в момент времени события E_i.

3.1.3. Оценка качества по числу событий

На практике результаты подсчёта числа людей часто рассматривают в виде числа проходов за определённый отрезок времени, не учитывая пространственное расположение событий в кадре. Предлагается ввести ещё одну характеристику качества, соответствующую такому подходу.

Рассмотрим отрезки видеопоследовательности, содержащие хотя бы P эталонных событий. Из них рассмотрим те, которые нельзя сузить, не нарушив условие на число эталонных событий. Согласно разд. 3.1.2 сопоставим множества событий E и $\widehat{E}$. Для каждого отрезка вычислим: число ложных событий FP_i как число несопоставленных событий из результата, число пропущенных событий FN_i как число несопоставленных событий из эталона; общее число эталонных событий на отрезке GT_i. Вычислим относительную ошибку подсчёта на отрезках:

$$\mathbb{E}_i = \frac{|FP_i - FN_i|}{GT_i}. \quad (4)$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Введем характеристику качества $\mathbb{E}$, как среднюю относительную ошибку на всех рассматриваемых отрезках.

3.2. Качество подсчёта людей

Для сопоставления событий (разд. 3.1.2) использовались соответствующее двум секундам видео значение Δ_f и $IOU_e = 0{,}4$. Для оценки качества по числу событий (разд. 3.1.3) $P = 10$. Относительная ошибка подсчёта $\mathbb{E}$ вычислялась после объединения всех отрезков всех наборов. Модификации добавлялись к базовому алгоритму одна за другой, чтобы оценить их вклад в полученный результат. Полученные результаты приведены на рис. 3.

Можно сделать вывод, что каждая из предложенных модификаций улучшает качество алгоритма. Предложенный алгоритм при частоте детекции $4/3$ Гц показывает приемлемое качество и позволяет сократить число используемых графических ускорителей по сравнению с базовым алгоритмом для изображенной на рис. 1, *а* системы (15 Гц) в $11{,}25$ раза.

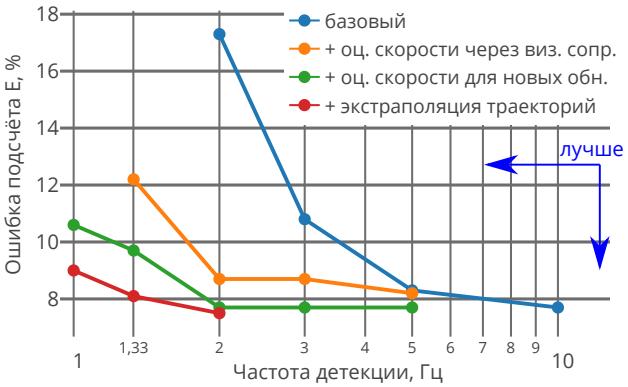


Рис. 3. Результаты экспериментальной оценки.

4. ЗАКЛЮЧЕНИЕ

Предложен алгоритм, являющийся модификацией работы [2], способный работать в распределённой системе с детекцией на разреженном видеопотоке и локальной обработкой, заключающейся в визуальном сопровождении найденных обнаружений. Метод позволяет значительно снизить необходимый объём вычислительных ресурсов для задачи подсчёта людей, не уступая в качестве базовому алгоритму.

5. БЛАГОДАРНОСТИ

Работа выполнена при поддержке гранта РФФИ 16-29-09612 офи_м «Исследование и разработка методов биометрической идентификации человека по походке, жестам и комплекции в данных видеонаблюдения».

6. ЛИТЕРАТУРА

[1] Ben Benfold и Ian Reid. «Stable Multi-Target Tracking in Real-Time Surveillance Video». В: *CVPR*. Июнь 2011, с. 3457—3464.

[2] Alex Bewley, ZongYuan Ge, Lionel Ott и др. «Simple Online and Realtime Tracking.» В: *CoRR* abs/1602.00763 (2016).

[3] Wongun Choi. «Near-Online Multi-target Tracking with Aggregated Local Flow Descriptor.» В: *CoRR* abs/1504.02340 (2015).

[4] Matej Kristan, Aleš Leonardis, Jiri Matas и др. *The Visual Object Tracking VOT2016 challenge results*. Springer. Окт. 2016. URL: `http://www.springer.com/gp/book/9783319488806`.

[5] Harold W Kuhn. «The Hungarian method for the assignment problem». В: *Naval Research Logistics (NRL)* 2.1-2 (1955), с. 83—97.

[6] D. Kuplyakov, E. Shalnov и A. Konushin. «Markov chain Monte Carlo based video tracking algorithm». Англ. В: *Programming and Computer Software* 43.4 (2017), с. 224—229. ISSN: 1608-3261; 0361-7688. DOI: `10.1134/S0361768817040053`.

[7] Philip Lenz, Andreas Geiger и Raquel Urtasun. «FollowMe: Efficient Online Min-Cost Flow Tracking With Bounded Memory and Computation». В: *The IEEE International Conference on Computer Vision (ICCV)*. Дек. 2015.

[8] A. Milan, L. Leal-Taixé, I. Reid и др. «MOT16: A Benchmark for Multi-Object Tracking». В: *arXiv:1603.00831 [cs]* (март 2016). URL: `http://arxiv.org/abs/1603.00831`.

[9] Shaoqing Ren, Kaiming He, Ross Girshick и др. «Faster r-cnn: Towards real-time object detection with region proposal networks». В: *Advances in neural information processing systems*. 2015, с. 91—99.

[10] E. V. Shalnov, A. S. Konushin и V. S. Konushin. «Convolutional neural network for camera pose estimation from object detections». Англ. В: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42.2-W4 (2017), с. 1—6. DOI: `10.5194/isprs-archives-XLII-2-W4-1-2017`.

[11] Eugeniy Shalnov, Vadim Konushin и Anton Konushin. «An improvement on an MCMC-based video tracking algorithm». В: *Pattern Recognition and Image Analysis* 25 (2015), с. 532—540.

[12] Guang Shu, Afshin Dehghan, Omar Oreifej и др. «Part-based multiple-person tracking with partial occlusion handling.» В: *CVPR*. IEEE Computer Society, 2012, с. 1815—1821. ISBN: 978-1-4673-1226-4.

[13] Tomas Vojir, Jana Noskova и Jiri Matas. «Robust scale-adaptive mean-shift for tracking». В: *Scandinavian Conference on Image Analysis*. Springer. 2013, с. 652—663.

[14] N. Wojke, A. Bewley и D. Paulus. «Simple Online and Realtime Tracking with a Deep Association Metric». В: *ArXiv e-prints* (март 2017).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

[15] Fengwei Yu, Wenbo Li, Quanquan Li и др. «Poi: Multiple object tracking with high performance detection and appearance feature». В: *European Conference on Computer Vision*. Springer. 2016, с. 36—42.

[16] И. В. Филиппов, В. А. Кононов, В. С. Конушин и др. «Подсчет количества людей в видеопоследовательности на основе детектора головы человека». В: *Программные продукты и системы* 1 (2015), с. 121—126. ISSN: 0236-235X.

[17] Е. В. Шальнов и А. С. Конушин. «Использование геометрии сцены для увеличения точности детекторов». В: *Программные продукты и системы* 30.1 (2017), с. 106—111. ISSN: 0236-235X. DOI: 10.15827/0236-235X.030.1.106-111.

ОБ АВТОРАХ

Купляков Денис Анатольевич, магистр факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова. Разработчик программного обеспечения ООО «Технологии видеоанализа»>. Его e-mail denis.kuplyakov@graphics.cs.msu.ru

Шальнов Евгений Вадимович, к.ф.-м.н., сотрудник факультета вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова. Его e-mail eshalnov@graphics.cs.msu.ru

Конушин Вадим Сергеевич, генеральный директор ООО «Технологии видеоанализа». Его e-mail vadim@tevian.ru

Конушин Антон Сергеевич, к.ф.-м.н., доцент Московского государственного университета им. М.В. Ломоносова и Национального исследовательского университета «Высшая школа экономики». Его e-mail ktosh@graphics.cs.msu.ru

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Метод быстрого отбора идентичных контрольных точек на изображениях отпечатка пальца

Д.Н. Лепихова, В.Ю. Гудков
lepikhova@sonda.ru|diana@sonda.ru
ООО «Сонда Про», Челябинская обл., г. Миасс, Россия
Челябинский государственный университет, Миасский филиал, г. Миасс, Россия

Изображение отпечатка пальца представляется в виде набора гнезд, которые строятся для каждой контрольной точки. При сравнении формируется список отобранных пар гнезд, от которых производится развитие фрагментов изображений. В статье рассматривается метод, ускоряющий отбор идентичных контрольных точек с помощью сравнения гнезд.

Ключевые слова: *отпечаток пальца, гребневый счет, контрольная точка*

Method of Fast Identical Minutiae selection

D.N. Lepikhova, V.U. Gudkov
lepikhova@sonda.ru|diana@sonda.ru
Sonda Pro Ltd, Miass, Chelyabinsk reg., Russia
Chelyabinsk State University, Miass, Chelyabinsk reg., Russia

Fingerprint image is represented as a set of local minutiae structures, which are formed for each minutia. While matching the list of selected pairs of nests is formed. For this pairs image fragment developing is produced. The article considers a method to accelerate the selection of identical pairs of minutiae when comparing nests.

Keywords: *fingerprint, ridge count, minutiae*

1. Введение

Большинство известных алгоритмов идентификации по отпечаткам пальцев основывается на сравнении шаблонов дактилоскопических изображений. Шаблоны могут содержать данные о контрольных точках (КТ) изображения [1], папиллярных линиях, а также дополнительную информацию о структуре изображения.

Локальные структуры для контрольных точек, устойчивые к шуму и искажениям, были предложены еще в 2000 году [13]. Позднее Feng предложил объединять такие структуры в пары [12], а Cao – в звездчатые множества [8]. В [15] рассмотрены варианты формирования таких множеств для двух, трех, k контрольных точек.

В [10] структура, построенная для каждой КТ, именуется k-плет. Для идентификации предложен алгоритм предварительного поиска в ширину (Coupled Breadth First Search), выполняющий сопоставление двух графов, построенных на основе КТ и скелета ДИ путем их обхода и сопоставления k-плетов на каждом этапе идентификации. В дальнейшем концепция k-плетов получила развитие в [9], в которой к структуре КТ добавляется дополнительная ассоциативная точка (SAP), а отпечаток представляется в виде ориентированного графа.

В [15] предлагается формировать для ДИ граф смежности контрольных точек (Minutiae Adjacency Graph). Чтобы построить MAG для некоторой центральной КТ, выбираются все такие КТ, для которых евклидово расстояние до центра меньше заданного порога. Каждое ребро графа записывается в виде $\{i, j, d_{ij}, r_{ij}, \varphi_{ij}\}$, где i, j – номера КТ, образующих ребро, d_{ij} - евклидово расстояние между КТ i и j, r_{ij} – значение гребневого счета между КТ i и j, φ_{ij} – ориентация ребра.

В [14] предложено строить подобные структуры не только для частных признаков ДИ, но и для общих признаков, например, для центра узора.

Самое большое влияние на скорость сравнения оказывает первичный этап отбора идентичных пар КТ. Список этих пар КТ можно сформировать различными способами, например использованием гребневого счета [4]. Значение гребневого счета для КТ определяется по числу линий, пересекаемых отрезками от нее до соседних КТ (в пределах некоторого заданного радиуса) [2].

Рассмотрим модель классического гребневого счета.

Для каждой КТ m_i выделяется некоторая окрестность, в которой измеряется гребневый счет r_{ik} от нее до КТ m_k как число папиллярных линий, пересекаемых прямой, проводимой между этими КТ. Затем для КТ m_i строится гнездо в виде

$N_i = \{m_k, r_{ik}, \gamma_{ik}\},\ m_k = \{x_k, y_k, \alpha_k\},\ i, k \in K,$

где γ_{ik} – азимут как угол доворота вектора направления α_i для m_i до отрезка, соединяющего m_i и m_k; K – множество всех КТ изображения (рис. 1).

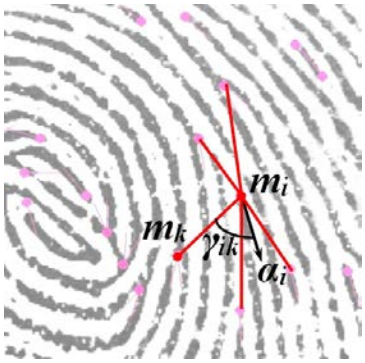


Рис. 1. Гнездо для контрольной точки m_i

Сравнение запросного и хранимого отпечатков сводится к сравнению гнезд из шаблонов этих отпечатков [4].

Вышеперечисленные алгоритмы демонстрируют достаточно хорошую производительность, однако дополнительное их ускорение не представляется возможным, поскольку на этапе обработки изображения не определено разбиение формируемых структур по классам. Следовательно, введение новых признаков повлечет за собой резкое увеличение количества вычислений и, как следствие, снижение скорости идентификации. Определение классов признаков и их упорядочивание в соответствии с классами

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

позволяет дополнительно повысить производительность алгоритмов.

Таким образом, актуальной задачей является ускорение отбора пар КТ, которые потенциально являются идентичными на изображениях одного и того же пальца. Для этого необходимо определить дополнительные признаки, которые позволяют заранее сформировать множество классов и определить принадлежность структуры КТ к какому-либо классу.

2. Описание алгоритма

При сравнении гнезд запросного и архивного отпечатков с целью выделения списка идентичных пар КТ каждое гнездо запросного отпечатка сравнивается с каждым гнездом архивного отпечатка пальца. Сравнение каждой пары гнезд предполагает построение множества геометрических и топологических оценок (количество оценок обычно превышает количество особенностей в гнезде) с последующим вычислением единой оценки степени подобия этой пары гнезд, что в случае полного перебора всех гнезд запросного и архивного отпечатков значительно увеличивает время сравнения [5, 11, 12, 15]. Таким образом, актуальным является алгоритм, позволяющий ускорить первичный отбор гнезд для их последующего сравнения.

Проведем сечение через каждую КТ вправо и влево на глубину нескольких линий перпендикулярно папиллярным линиям и пронумеруем пересеченные линии (как связи) по часовой стрелке, начиная от центра. По каждой линии проследим путь до первой КТ, встреченной на связи, либо до ее проекции на линию [3-5] (рис. 2).

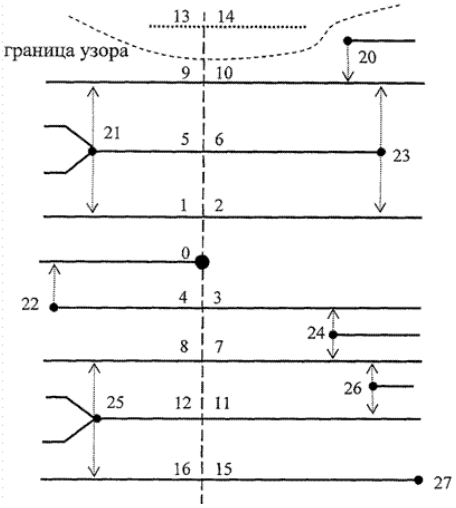


Рис. 2. Пример построения сечения

Пусть каждое гнездо описывается матрицей, в которой на основе i-й связи синтезируется строка вида $|d_i, \gamma_i, e_i, v_i, n_i|$, где d_i – расстояние от центра сечения до КТ на i-й связи; γ_i – угол доворота по часовой стрелке направления α_i для КТ, образующей гнездо, до луча, ориентированного на КТ по -й связи; e_i - событие на i-й связи; $v_i \in \{0; 1\}$ – состояние i-й связи; n_i - номер КТ на i-й связи (см. рис. 2) [5].

Заметим, что параметры связи не зависят от ориентации и выбора центра координат.

Для КТ m_i выберем несколько ее ближайших соседей, например, два – m_1 и m_2 (число выбираемых КТ можно варьировать, повышая тем самым точность отбора, но снижая скорость). Значение гребневого счета до выбранных соседних КТ равно соответственно r_{i1} и r_{i2}. Тогда m_i образует два диполя с выбранными КТ m_1 и m_2 (рис. 3).

Для выбранных КТ определяют значение азимута $\gamma_{ik} \in \{0; 359\}, k \in \{1,2\}$. С помощью квантования значение азимута сводят к значению $\gamma_{ik} \in \{0; 63\}$. Тем самым получают 64 возможных класса для идентификации диполей.

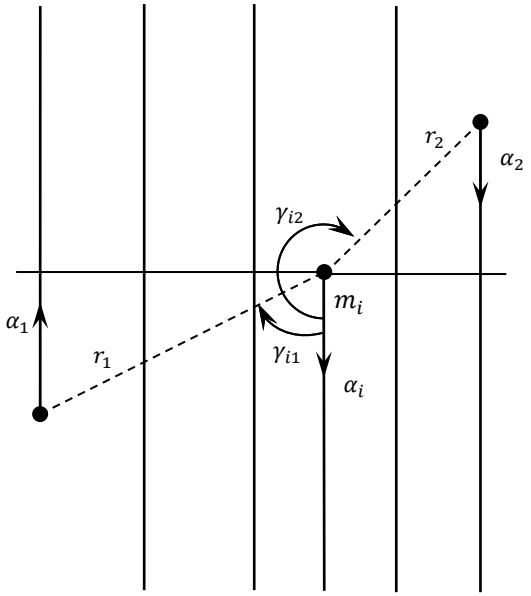


Рис. 3. Построение диполя

Для каждого диполя определены его длина d_{ik}, равная расстоянию между КТ m_i и m_k, значение гребневого счета r_{ik} между КТ m_i и m_k, угол φ_{ik} разности векторов α_i и α_k. Таким образом, диполь имеет вид $\{\gamma_{ik}, d_{ik}, r_{ik}, \varphi_{ik}\}$.

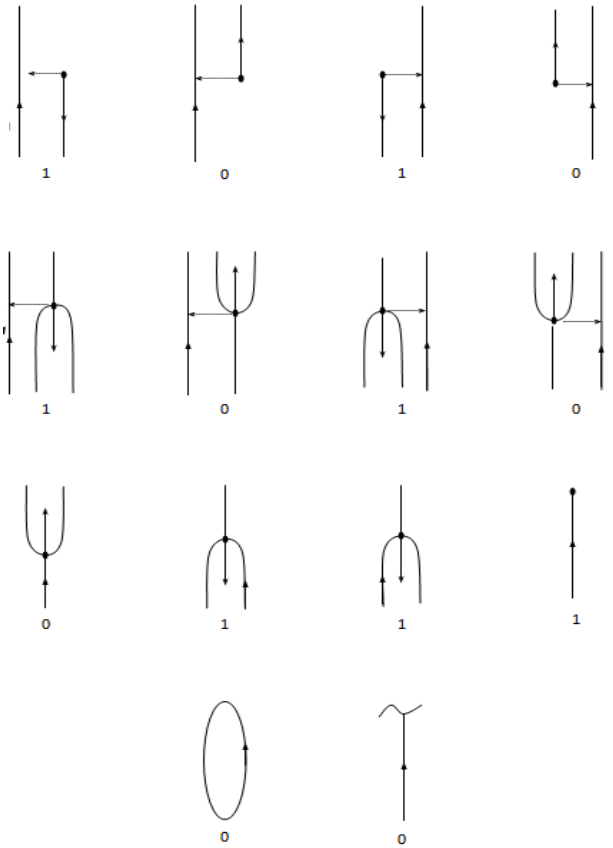


Рис. 4. Значение бита в зависимости от направления КТ на связи

В силу малого числа классов высока вероятность ошибочной идентификации двух различных диполей. Чтобы

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

этого избежать, добавим к описанию диполя дополнительный бит, удваивающий число классов. Значение дополнительного бита характеризует направление α_k события, которое образует КТ m_k, на связи с КТ m_i, причем 1 означает направление КТ по ходу связи, 0 – против хода (рис. 4).

При добавлении бита число возможных классов, которым может принадлежать диполь, удваивается до 128. Этот бит добавляется старшим битом к азимуту γ_{ik}. При этом получаются те самые 128 классов c_{ik}.

Для сравнения диполи сортируют по классам в зависимости от значения c_{ik}. Классы позволяют при упорядочивании диполей быстро найти пары потенциально идентичных КТ. Затем для диполей из запросного и архивного отпечатков вычисляется пересечение списка классов слиянием [7]. Для равных классов дополнительно проверяют разность соответствующих углов φ_{ik}, расстояний d_{ik}, значений гребневого счета r_{ik} для запросного и архивного отпечатков. Значения этих разностей должны быть в пределах заданных допусков. Если это условие выполняется, то КТ m_i со стороны запросного отпечатка и соответствующая КТ со стороны архивного отпечатка добавляются в список L_0 возможных кандидатов идентичных пар КТ.

За счет того, что формальная отсечка непохожих пар КТ выполняется на основе упорядоченного списка классов, а остальные проверки разностей проводятся редко, этот подход позволяет значительно увеличить скорость. В результате поиска по классам очень быстро формируются возможные пары идентичных КТ.

На основе списка L_0 для каждого гнезда определяется 32-битное слово, каждый бит которого характеризует ориентацию КТ по ходу или против хода при прослеживании связи (см. рис 4). Полученные два слова объемом 32 бита сравниваются при помощи логической операции исключающего «или», причем для каждого сравнения подсчитывается число совпадающих бит. Отбирая из списка L_0 пары КТ по наибольшему числу совпадающих бит, формируют список L_1.

Из списка L_1 выбираются 1-3 возможных пары, которые, как правило, являются идентичными на отпечатках, принадлежащих одному и тому же пальцу. В дальнейшем выбранные пары связей используются как затравки для развития похожих фрагментов согласно методу, описанному в [6]. Собственно развитие фрагментов путем наращивания подобных областей формирует высококачественный, но медленный поиск. Ускорение достигается за счет быстрого первичного отбора идентичных пар – затравок.

3. Результаты экспериментов

Для тестирования предложенного алгоритма из баз FVC 2000, 2002, 2004 и 2006 г. были отобраны примерно 6000 изображений, полученных с оптического сенсора. При этом выбирались изображения, наиболее сложные при идентификации. Для отобранных отпечатков исследована зависимость точности сравнения от скорости идентификации.

График на рис. 5 иллюстрирует оценку величины ошибки, определяемой по DET-кривой, в зависимости от скорости сравнения. По оси абсцисс отражены данные о количестве сравниваемых миллионов пар отпечатков в секунду в режиме идентификации, по оси ординат – значение интегральной ошибки в процентилях. Левый край графика соответствует сравнению отпечатков без ускорителей, а правый – с ускорителями с минимальными допусками на разности параметров диполей.

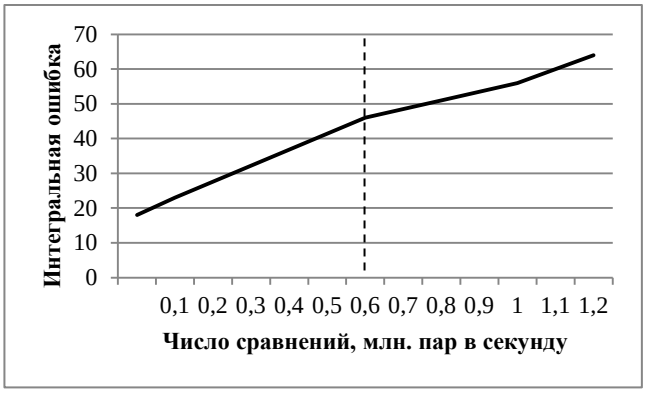


Рис. 5. Результаты экспериментов

4. Заключение

В работе рассматривается метод, позволяющий ускорить отбор пар КТ для последующего сравнения изображений отпечатков пальцев. Ускорение обеспечивается за счет сокращения числа пар КТ, используемых в качестве затравок для развития похожих фрагментов и дальнейшего сравнения.

Дополнительная информация, характеризующая каждое гнездо, незначительно увеличивает размер шаблона. Однако за счет того, что формирование списка возможных кандидатов L_0 и списка затравок L_1 производится практически мгновенно, общее время сравнения значительно сокращается.

Таблица. Сравнительный анализ алгоритмов

Алгоритм	Время сравнения, мс	Размер шаблона, байт	FMR_{10000}, %
PPM	46	36922	6,179
M3gl	667	88958	9,643
MinutiaeClusterFull	96	3344160	10,214
MCC-Based (Baseline)	242	214629	20,107

В таблице приводятся результаты тестирования FVC-onGoing [11] по состоянию на 15.05.2018 г. Рассматривались алгоритмы идентификации по ладоням на основе полных отпечатков. Метод, представленный в работе, применяется в алгоритме PPM. В таблице видно, что при высокой производительности качественные характеристики алгоритма высоки, что демонстрирует доля ошибочно идентифицированных идентичных отпечатков на 10000 сравнений (значение FMR_{10000}).

5. Литература

[1] ГОСТ Р ИСО/МЭК 19494-2 - 2005 Форматы обмена биометрическими данными. Часть 2: Данные изображения отпечатка пальца - контрольные точки. Biometric Data Interchange Format – Part 2: Finger Image – Finger Minutiae.

[2] Гудков В.Ю. Модель гребневого счета на основе топологии дактилоскопического изображения // Вестник ЧелГУ, 2011. Вып. 13. С. 99-108.

[3] Гудков В.Ю., Аркабаев Д.И. Пат. 2444058, МПК G06K 9/62. Способ гребневого счета на основе топологии дактилоскопического узора. № 2010110115/08, заявл. 17.03.2010; опубл. 27.02.2012 Бюл. № 6.

[4] Гудков В.Ю. Способы математического описания и идентификации отпечатков пальцев / под ред. член-

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

корр. РАН В.Л. Арлазарова и д.т.н. проф. Н.Е. Емельянова // Обработка изображений и анализ данных: Труды ИСА РАН. Т. 38. С. 336--356. 2008.
[5] Гудков В.Ю., Аркабаев Д.И.. Пат. 2331108, МПК G06K 9/62. Способ сравнения отпечатков папиллярных узоров. № 2006140939/09, заявл. 20.11.2006; опубл. 10.08.2008; бюл. № 22.
[6] Гудков В.Ю., Боков А.С. Пат. 2185661, МПК G06K 9/62. Способ сравнения папиллярных узоров пальцев. № 2000121735/09, заявл. 14.08.2000; опубл. 20.07.2002; бюл. № 20.
[7] Новиков Ф.А. Дискретная математика: Учебник для вузов. 2-е изд. Стандарт третьего поколения. Издательский дом "Питер", 2012 г. 400 с.
[8] Cao J., Feng J., A Robust Fingerprint Matching Algorithm Based on Compatibility of Star Structures. Proceedings of the Sixth International Symposium on Multispectral Image Processing and Pattern Recognition (MIPPR 2009), vol. 7498, Remote Sensing and GIS Data Processing and Other Applications, pp.74983X, 2009
[9] Chen X., Wang L., Li M. An Efficient Graph-Based Algorithm for Fingerprint Representation and Matching, Proceedings of the 3rd International Conference on Multimedia Technology（ICMT 2013), pp. 1019-1029, 2013
[10] Chikkerur S., Cartwright A., Govindaraju V. K-plet and cbfs: A graph based fingerprint representation. International Conference on Biometrics ICB 2006: Advances in Biometrics, p. 309-315.
[11] Dorizzi B., Cappelli R., Ferrara M., et al. Fingerprint and On-Line Signature Verification Competitions at ICB 2009 //International Conference on Biometrics ICB 2009 (Alghero, Italy, June 2-5, 2009). P. 725-732.
[12] Feng Y., Feng J., Chen X., et al. A Novel Fingerprint Matching Scheme Based on Local Structure Compatibility, International Conference on Pattern Recognition, p. 374-377, 2006.
[13] Jiang X., Yau W.Y. Fingerprint Minutiae Matching Based on the Local and Global Structures. International Conference on Pattern Recognition, vol. 2(1): p. 1038–1041, 2000.
[14] Leslie S., Sumathi C.P. A Robust Hierarchical approach to Fingerprint matching based on Global and Local Structures. International Journal of Applied Engineering Research ISSN 0973-4562 Vol. 13, N. 7, p. 4730-4739, 2018
[15] Ratha N. K., Pandit V. D., Bolle R. M., et al. Robust fingerprint authentication using local structure similarity. In Workshop on applications of Computer Vision, p. 29–34, 2000.

Об авторах

Лепихова Дарья Николаевна, инженер-программист ООО «Сонда ПРО». Её e-mail lepikhova@sonda.ru.
Гудков Владимир Юльевич, д.ф.-м.н, доцент кафедры прикладной математики Миасского филиала Челябинского Государственного университета. Его e-mail diana@sonda.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Нейронные сети для распознавания формы по медиальному представлению

Н.А. Ломов[1], С.П. Арсеев[1]
nikita-lomov@mail.ru | 9413serg@gmail.com
[1]Московский государственный университет имени М.В. Ломоносова, Москва, Россия

Статья посвящена разработке архитектур для нейронных сетей, оперирующих данными особого вида — медиальным представлением формы, которое рассматривается как частный случай неориентированного графа. Предложены способы решения проблем, затрудняющих обработку данных такого типа традиционными нейросетями: различная длина входных данных, неоднородность их структуры, неупорядоченность составных элементов. Сформулированы скелетные аналоги стандартных операций, используемых в свёрточных нейронных сетях. Проведены эксперименты по классификации рукописных цифр.

Ключевые слова: *глубинное обучение, свёрточные нейронные сети, медиальное представление, скелет.*

Neural Networks for Shape Recognition by Medial Representation

N.A. Lomov[1], S.P. Arseev[1]
nikita-lomov@mail.ru | 9413serg@gmail.com
[1]Lomonosov Moscow State University, Moscow, Russia

The article is dedicated to the development of neural networks that process data of a special kind — a medial representation of the shape, which is considered as a special case of an undirected graph. Methods for solving problems that complicate the processing of data of this type by traditional neural networks — different length of input data, heterogeneity of their structure, unordered constituent elements — are proposed. Skeletal counterparts of standard operations used in convolutional neural networks are formulated. Experiments on the classification of handwritten digits are carried out.

Keywords: *deep learning, convolutional neural networks, medial representation, skeleton.*

1. Введение

Стремительный рост популярности методов глубинного обучения несколько лет назад начался с успехов в ряде задач распознавания изображений. Изображения обычно задаются матрицами точек, поэтому их можно считать примером организованных, стандартизированных данных. В дальнейшем методы глубинного обучения также добились серьёзного прогресса в задачах с данными более сложной природы, которые можно представить последовательностями произвольной длины: распознавание речи, распознавание рукописных текстов, машинный перевод и т. д.

Предложенные архитектуры используют структурированные (по времени или положению) данные. Вместе с тем проблема обработки данных менее «удобной» для нейросетей природы остаётся сравнительно мало изученной. Характерным примером данных такого типа являются графы, а в целом подобные данные часто возникают в задачах вычислительной геометрии: это облака точек, полигональные модели, триангуляции, контурно-скелетные представления. Трудности заключаются в том, что объекты, подлежащие обработке, не упорядочены и могут иметь произвольную размерность и структуру связей. Обстоятельное исследование этой проблемы и возможных подходов к её решению можно найти в статье [3].

Цель данной работы — применить нейросетевые методы к графам особого класса — скелетам, которые являются описанием формы в виде её некоторой основы, «каркаса». Хотя мы исходим из того, что скелет является частным случаем неориентированного графа, некоторые его особенности, в частности геометрическая природа и сильная разреженность, позволяют извлечь из них дополнительную пользу.

Непосредственно скелеты используются в нейросетевых методах в очень ограниченном контексте: как правило, это задачи распознавания действий и жестов [5, 7], а сами скелеты получены с помощью сенсорных устройств (например, Kinect). Такие скелеты предназначены исключительно для описания фигуры человека и имеют строго заданную структуру. Примеры использования в качестве входных данных скелетов для описания произвольных форм авторам не известны. В статье предпринята попытка начать исследования в данной области.

2. Обработка скелетов при помощи нейросети

2.1. Базовые понятия

Скелетом называется множество центров всех вписанных пустых кругов фигуры.

Радиальной функцией — функция, которая каждой точке скелета ставит в соответствие радиус вписанного круга.

Скелет и радиальная функция в совокупности образуют *медиальное представление формы*.

Стрижкой (регуляризацией) скелета называется процесс отсечения ветвей, вносящих незначительный вклад в образование формы. Способ стрижки, который будет использоваться далее, описан в [1].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Скелет можно рассматривать как геометрический граф, рёбрами которого являются отрезки прямых линий и парабол. К такому представлению ведёт получение скелета из диаграммы Вороного [2]. В рамках данной задачи мы будем рассматривать скелет в виде $S = (V, E)$, где каждая вершина $v \in V$ представляется в виде тройки (x, y, r) — координат и значения радиальной функции, а каждое ребро $e \in E$ является неупорядоченной парой вершин.

2.2. Упорядочение вершин

Наша первоначальная цель — добиться принципиальной возможности подавать скелеты на вход нейронной сети. На этом пути мы видим два препятствия: 1) скелеты имеют различную размерность в силу различного числа рёбер и вершин, 2) элементы скелета не имеют «естественного» порядка. В работе [10] предложено решение этой задачи: вершины скелета должны быть упорядочены согласно их центральности по посредничеству (betweenness centrality). Для вершины v эта величина определяется следующим образом:

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

где σ_{st} — число кратчайших путей из вершины s в вершину t, а $\sigma_{st}(v)$ — число таких путей, проходящих через v. Таким образом, в начало списка попадут вершины, которые имеют больше соседей по сравнению с остальными и достаточно удалены от терминальных. В конец же списка попадут изолированные и терминальные вершины. В качестве дополнительной характеристики мы используем центральность по близости (closeness centrality, $h(v) = \frac{1}{\sum_u d(u,v)}$), а в случае совпадения будем сравнивать расстояния ($d(u, v)$) вершин-кандидатов до вершин, уже добавленных в список.

Algorithm 1 Упорядочение вершин скелета

Дано: множество вершин $\{v_i\}_{i=1}^n$
for $i \in 1, \ldots, n$ **do**
  $F_i = [g(v_i), h(v_i)]$
$I = [1, \ldots, n]$
$L = [\,]$
while $I \neq [\,]$ **do**
  Упорядочить список I лексикографически по убыванию F_i
  $k = \text{pop}(I)$ // Извлечь первый элемент из списка
  Приписать k в конец L
  for $i \in I$ **do**
    Приписать $-d(v_i, v_k)$ в конец F_i

Отметим, что упорядочение может и не понадобиться, если на последнем слое мы будем применять ко всей карте признаков симметричную pooling-операцию: например, максимум или среднее значение.

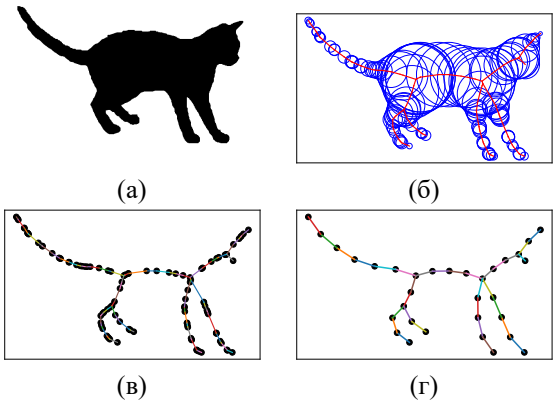
2.3. Унификация скелета

Вновь обратимся к работе [10]: чтобы оставить в графе заданное число вершин n, предлагается взять из полученного на предыдущем списка первые n вершин с сохранением порядка, а если вершин меньше n — дополнить список фиктивными вершинами. В случае скелета это исказит сам геометрический облик графа: если рёбер было слишком много, концы ветвей будут отсечены. Возникает идея использовать скелет, сохраняющий очертания исходного, но с фиксированным числом вершин и рёбрами примерно одинаковой длины.

Algorithm 2 Унификация скелета

Дано: скелет S, число вершин n, минимальная относительная длина ε.
Если в скелете больше одной компоненты связности, сделать скелет связным с помощью дополнительных рёбер минимальной длины.
Разбить скелет на ветви — линии, концами которых являются точки степени 1 и 3, и которые не содержат промежуточных точек такого типа. Число ветвей обозначим за n_e, число их концов — за n_v.
if $n_v > n$ **then**
  Для $n_v - n$ кратчайших ветвей стянуть концы в одну точку посередине ветви
else
  Определить длину каждой ветви l_i, общую длину скелета L и среднюю длину $l = \frac{L}{n_e + n - n_v}$
  Для ветвей, длина которых меньше εl, стянуть концы
  Пересчитать n_e и n_v
  for $i \in 1, \ldots, n_e$ **do**
    $n_i = \lceil \frac{l_i}{l} \rceil - 1$
  while $n_v + \sum_{i=1}^{n_e} n_i \geq n$ **do**
    Выбрать вершину с наименьшим $\frac{l_i}{n_i - 1}$ и уменьшить n_i на 1
  На каждую ветвь равномерно добавить n_i вершин

Процесс подготовки входных данных проиллюстрирован рис. 1.



(а) (б) (в) (г)

Рис. 1. Построение скелета с рёбрами равной длины. а) Изображение б) медиальное представление в) скелет г) унифицированый скелет.

2.4. Свёрточный слой

В случае изображений операция свёртки — это линейная комбинация признаков пикселей, находящих-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

ся в окрестности выбранного. Действуя по аналогии, сформируем для выбранной вершины окрестность из r ближайших вершин в графе, а если достижимых вершин меньше r, дополним окрестность фиктивными вершинами с нулевыми признаками. Чтобы избежать элемента случайности, используем дополнительные критерии:

1. Список длин кратчайших путей до уже выбранных для окрестности точек.
2. Центральность по посредничеству.
3. Центральность по близости.

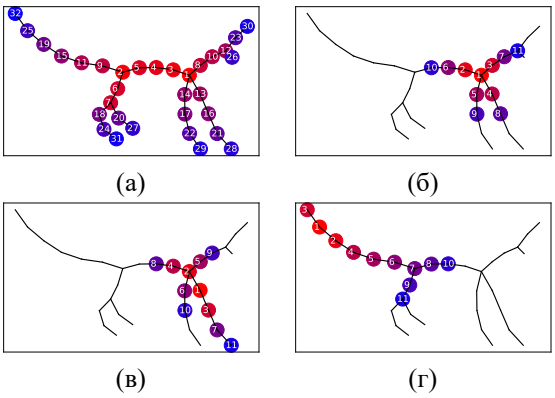
Общий вид процедуры аналогичен случаю упорядочения вершин в скелете.

Можно формально определить процедуру графовой свёртки окрестностей следующим образом. Пусть $F \in \mathbb{R}^{n\times m}$ — матрица вершин-признаков, $P \in \mathbb{N}_0^{n\times r}$ — матрица индексов элементов, образующих окрестности. Тогда $E(F,P) \in \mathbb{R}^{n\times r\times m}$, и

$$e_{ijk} = \begin{cases} f_{p_{ij}k}, & \text{если } p_{ij} \geq 0, \\ 0, & \text{если } p_{ij} = 0. \end{cases}$$

Теперь, если $W \in \mathbb{R}^{r\times m\times q}$, результат свёртки $C(E,W) \in \mathbb{R}^{n\times q}$ определяется как:

$$c_{ij} = \sum_{1\leq t\leq r, 1\leq s\leq m} e_{its} w_{tsj}.$$



(а) (б) (в) (г)

Рис. 2. Упорядочение всех вершин (а) и построение окрестностей длины 11 для отдельных вершин (б,в,г). При проходе по списку цвета меняются от красного к синему.

Примеры построения окрестностей, а также результат упорядочения всех вершин отражены на рис. 2.

2.5. Слой субдискретизации

Целью слоя субдискретизации является снижение размерности данных. Будем разбивать вершины графа на группы с тем расчётом, чтобы подграфы, образованные вершинами в группе, были связными. Можно предложить метод, который разбивает вершины по парам, убирая из графа пару из какой-то терминальной вершины и инцидентной ей, если её степень не больше 2. В том случае, если такой шаг невозможен, используется ряд эвристических правил для разрешения коллизий. Результатом применения такой процедуры будет новый граф, в котором пары считаются смежными если есть хотя бы одно ребро, соединяющее вершины из разных пар. Повторяя такую процедуру k раз, можно добиться разбиения множества вершин на группы мощностью 2^k (рис. 3).

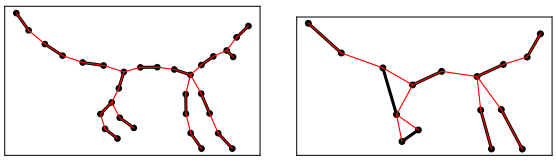


Рис. 3. Объединение вершин в пары и симплификация графа путём замены пары вершин на одну новую

Теперь можно определять архитектуру «скелетной» сети с помощью тех же обозначений, которые традиционно используются для свёрточных сетей.

3. Архитектура без унификации

Альтернативным способом применения свёрточных нейронных сетей для обработки скелетов является использование графовой свёрточной архитектуры, не требующей унификации графа. Для сравнения подходов был реализован алгоритм на основе архитектуры, описанной в [9]. Сеть в этой архитектуре является полностью свёрточной и состоит из последовательности слоёв, описываемых формулой:

$$H^{(l)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l-1)}W^{(l)}\right).$$

Здесь l — номер слоя, $H^{(l)} \in \mathbb{R}^{N\times D}$ — значения активаций на слое l, $\widetilde{A} = A + I_N$ — матрица смежности графа с добавленными петлями на каждой вершине: A — матрица смежности, а I_N — единичная матрица. $\widetilde{D}_{ii} = \sum_j \widetilde{A}_{ij}$, $W^{(l)}$ — обучаемая матрица весов слоя l, а σ — функция активации. На каждом слое выход предыдущего слоя (на первом слое используется матрица признаков исходного графа) умножается слева на симметрическую нормализованную матрицу Кирхгофа для данного графа, а справа — на матрицу весов слоя.

Поскольку при такой операции информация распространяется только по непосредственным соседям узла графа, в граф добавляется фиктивная вершина, соединённая со всеми остальными. Также, метка графа при классификации определяется выходом последнего свёрточного слоя для этой вершины.

4. Эксперименты

Эксперименты проводились на популярной базе изображений рукописных цифр MNIST. Сначала изображения были бинаризованы методом Оцу, далее для них были вычислены скелеты, которые подверглись процедуре стрижки с параметром в 1 пиксель. Унифицированные скелеты содержали 32 вершины. Архитектура сети состояла из следующих слоёв:

1. Свёрточный слой размера $7 \times 3 \times 16$ (окрестность мощности 7, 3 признака — x, y, r и 16 фильтров) + функция активации ReLU ($f(x) = x^+ = \max(0, x)$).
2. Свёрточный слой размера $9 \times 7 \times 32$ + ReLU.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

3. Полносвязный слой размера 256.
4. Свёрточный слой с размерностью признаков на выходе 10.
5. Слой softmax-функции ($\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}}$) для определения итоговой метки класса.

В свёрточных слоях применялась регуляризация по методу dropout, нейроны исключались из сети с вероятностью 50%. Точность классификации составила 98,36%. Это несколько уступает результату, который показывают наиболее эффективные из графовых сетей ChebNet[4] и MoNet[8], когда вершины в графе соответствуют пикселям изображения (таблица 1). Следует, однако, отметить, что все графы при этом имеют одинаковую структуру, поэтому архитектура сети оказывается близкой к традиционным свёрточным сетям, таким как [6]. Неоднородность возникает при выделении на изображениях суперпикселей, особым способом описания которых можно считать рёбра скелета. Таким образом, в случае произвольных графов предложенный метод показывает результат, превосходящий существующие аналоги. Отметим также компактность скелетного описания: каждый граф имеет лишь 32 вершины с 3 признаками; если потребовать от других графовых сетей подобной длины описания, превосходство нашего метода станет ещё более ощутимым.

Метод	LeNet5[6]	ChebNet[4]	MoNet[8]
Полная сетка	99,33%	99,14%	99,19%
$\frac{1}{4}$ сетки	98,59%	97,70%	98,16%
300 суперпикс.	–	88,05%	97,30%
150 суперпикс.	–	80,94%	96,75%
75 суперпикс.	–	75,62%	91,11%

Таблица 1. Точность классификации для различных форматов входных данных

Для альтернативной архитектуры без унификации графа сеть состояла из следующих слоёв:

1. Два свёрточных слоя с размерностью признаков на выходе 32 + ReLU.
2. Четыре свёрточных слоя с размерностью 64 + ReLU.
3. Свёрточный слой с размерностью 10.
4. Слой softmax-функции для определения итоговой метки класса.

Несмотря на более сложную архитектуру, точность классификации составила всего 90%. Это свидетельствует о том, что выбранный подход, успешно применявшийся в [9] для задачи классификации научных работ, плохо поддаётся переносу на задачу классификации изображений. Вероятно, это связано с оператором графовой свёртки, в результате применения которого может теряться информация о взаимном расположении узлов. Ещё одним источником трудностей является отсутствие аналога полносвязного слоя для таких данных, что влечёт за собой усложнение архитектуры.

5. Заключение

В работе были рассмотрены два метода глубинного обучения с использованием подхода графовой свёртки на примере задачи классификации изображений. Были предложены два подхода, позволяющих использовать медиальное представление, заданное графом произвольной структуры, для решения задачи классификации. Первый подход, основанный на предобработке и унификации медиального представления, показал на наборе данных MNIST качество работы, сопоставимое с традиционными нейросетевыми алгоритмами. Попытка исключить этап унификации, предпринятая во втором подходе, показала снижение точности работы алгоритма.

6. Благодарности

Работа выполнена при поддержке РФФИ грант 17-01-00917.

7. Литература

[1] *Местецкий Л.М.* Непрерывная морфология бинарных изображений: фигуры, скелеты циркуляры. М.: ФИЗМАТЛИТ, 2009.

[2] *Местецкий Л. М.* Скелет многоугольной фигуры — представление плоским прямолинейным графом. — Труды 20-й международной конференции «Графикон-2010». СПб, ИТМО, с. 222–229, 2010.

[3] *M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, P. Vandergheynst.* Geometric deep learning: going beyond Euclidean data // IEEE Signal Processing Magazine, Vol. 34(4), pp. 18–42, 2017.

[4] *M. Defferrard, X. Bresson, P. Vandergheynst.* Convolutional neural networks on graphs with fast localized spectral filtering // Proceedings of the 30th International Conference on Neural Information Processing Systems, pp. 3844–3852, 2016.

[5] *Y. Du, W. Wang, L. Wang.* Hierarchical recurrent neural network for skeleton based action recognition // 2015 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1110–1118, 2015.

[6] *Y. LeCun, L. Bottou, Y. Bengio, P. Haffner.* Gradient-based learning applied to document recognition // Proc. IEEE, 86(11):2278–2324, 1998.

[7] *C. Li, P. Wang, S. Wang, Y. Hou, W. Li.* Skeleton-based action recognition using LSTM and CNN // 2017 IEEE International Conference on Multimedia & Expo Workshops, pp. 585–590, 2017.

[8] *F. Monti, D. Boscaini, J. Masci, E. Rodolà, J. Svoboda, M. Bronstein.* Geometric Deep Learning on Graphs and Manifolds Using Mixture Model CNNs // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5425–5434, 2017.

[9] *T. N. Kipf, M. Welling.* Semi-Supervised Classification with Graph Convolutional Networks // arXiv preprint, arXiv:1609.02907, 2016.

[10] *M. Niepert, M. Ahmed, K. Kutzkov.* Learning Convolutional Neural Networks for Graphs // Proceedings of the 33rd International Conference on International Conference on Machine Learning, Vol. 48, pp. 2014–2023, 2016.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Методы и средства томографического видения пространства активно-импульсными телевизионными измерительными системами

А.К. Мовчан[1], В.В. Капустин[1], М.И. Курячий[1]
mr.movchann@mail.ru|peregnun@mail.ru|kur@tu.tusur.ru
[1]ТУСУР, Томск, Россия

Представлены результаты измерения освещенности объектов наблюдения в зависимости от задержки стробирования фотоприемного устройства активно-импульсной телевизионной измерительной системы (АИ ТИС) в чистой и малопрозрачных средах. Построены зависимости абсолютного и относительного контрастов объектов наблюдения в зависимости от задержки стробирования фотоприемного устройства и среды распространения. Определена аппаратурная погрешность АИ ТИС по задержке стробирования и оценена дальность до наблюдаемых объектов.

Ключевые слова: *контраст, активная зона видения, томографическое видение, активно-импульсная телевизионная измерительная система.*

Methods and means of tomographic vision of space by active-pulse television measurement systems

A.K. Movchan[1], V.V. Kapustin[1], M.I. Kuryachiy[1]
mr.movchann@mail.ru|peregnun@mail.ru|kur@tu.tusur.ru
[1]TUSUR, Tomsk, Russia

The results of measurement of the illumination of the objects of observation are presented depending on the delay in the gating of the photoreceiver of the active-pulse television measuring system (AP TMS) in pure and low transparent space. Dependences of the absolute and relative contrasts of the objects of observation are constructed depending on the delay in the gating of the photodetector device and the propagation medium. The instrumental error of the AP TMS on gating delay was determined and the range to the observed objects was estimated.

Keywords: *contrast, active vision area, tomographic vision, active pulse television measuring system.*

1. Введение

Качество изображения обычных систем видеонаблюдения, работающих в сложных условиях видимости (туман, дымка, пыль, снегопад и т.д.), значительно ухудшается, так как источник излучения подсвечивает как объекты наблюдения, так и фон, создающий обратное рассеянное в окружающей среде излучение и встречную засветку, снижающую контраст изображения и возможность достоверного визуального обнаружения, распознавания и идентификации изображений объектов.

Для наблюдения в сложных условиях видимости применяются активно-импульсные телевизионные измерительные системы (АИ ТИС) [1].

Основными достоинствами АИ ТИС являются устранение помехи обратного рассеивания и значительное ослабление световых помех естественной и искусственной природы. Принцип работы активно-импульсных систем основан на импульсном подсвете пространства и стробировании по времени фотоприемного устройства, оснащённого быстродействующим затвором.

Сущность метода сводится к следующему. Объект наблюдения освещается световыми импульсами, длительность которых значительно меньше времени распространения света до объекта и обратно. В том случае, когда временная задержка между моментом излучения импульса и моментом открывания затвора равна удвоенному времени, необходимому для прохождения светом расстояния до объекта и обратно, наблюдатель будет видеть только сам объект и участок пространства, непосредственно его окружающий из чего следует, что импульсный режим работы системы позволяет сформировать активную зону видения и исключить все, что находится за ее пределами.

Глубина активной зоны видения определяется как временем открытого состояния затвора, так и длительностью светового импульса [2,5].

Таким образом, средствами АИ ТИС возможно осуществить томографическое видение окружающего пространства. Под томографическом видением понимается получение послойного изображения окружающего пространства с целью наблюдения за объектами в выбранном участке пространства (слое), а так же получения информации как о самом объекте (линейные размеры, расстояние), так и об объектах находящихся в непосредственной близости от объекта наблюдения и таким образом осуществлять классификацию объектов как по дальности так и по линейным размерам [4].

На рис. 1 представлен принцип томографического видения пространства с помощью АИ ТИС.

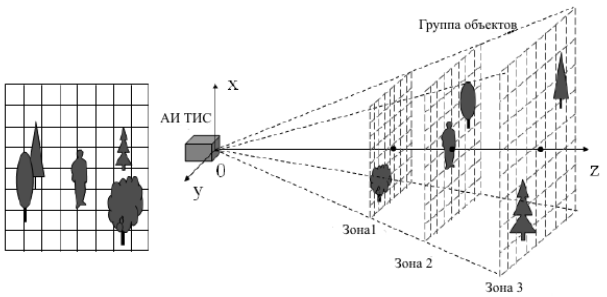
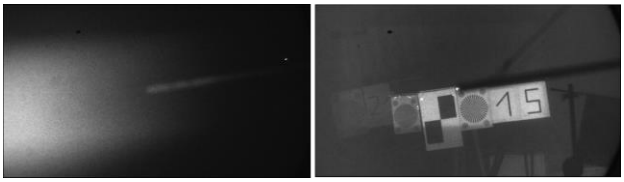


Рис. 1. Принцип томографического видения пространства с помощью АИ ТИС.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Особенно очевидны преимущества АИ ТИС при томографическом видении пространства в сложных условиях видимости. На рис. 2 представлены сравнительные изображения, полученные с помощью АИ ТИС в режиме наблюдения обычных систем видеонаблюдения, а так же в активно-импульсном режиме основным для АИ ТИС.



а *б*

Рис. 2. Сравнительные изображения, полученные в условиях плотного тумана: *а* – режим обычной системы видеонаблюдения; *б* – активно-импульсный режим работы АИ ТИС.

2. Томография в чистой среде

В целях исследования характеристик изображений получаемых с помощью АИ ТИС, были проведены эксперименты в большой аэрозольной камере (БАК) Института оптики и атмосферы имени В.Е. Зуева СО РАН [3].

Для фиксации изображений внутри аэрозольной камеры АИ ТИС была установлена в смотровом иллюминаторе БАК.

В качестве объектов наблюдения использовались измерительные таблицы установленные на определенных расстояниях от расположения АИ ТИС.

Эксперименты проводились как в чистой среде так и в средах с пониженной оптической прозрачностью, имитирующих условия туманов и дымов.

Эксперимент томографического видения в чистой среде осуществлялся путем фиксации изображения объекта наблюдения представляющего собой измерительную таблицу типа «шахматное поле», находящейся на расстоянии 15 м от расположения АИ ТИС (рис. 3). Фиксация изображений объекта наблюдения производилась при каждом изменении задержки стробирования фотоприемного устройства АИ ТИС на фиксированный шаг 5–10 нс, что в свою очередь позволило сформировать определенное количество пересекающихся активных зон видения. Известно, что формы импульсов излучения источника подсвета и импульса стробирования фотоприемного устройства определяют форму активной зоны видения АИ ТИС [2,3]. На рис. 4 приведена осциллограмма тока импульса излучения источника подсвета и напряжения на фотокатоде электронно-оптического преобразователя.

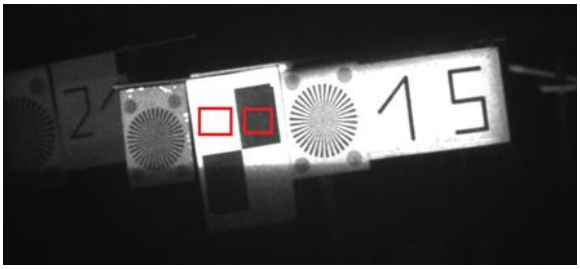


Рис. 3. Изображение объекта наблюдения в чистой среде.

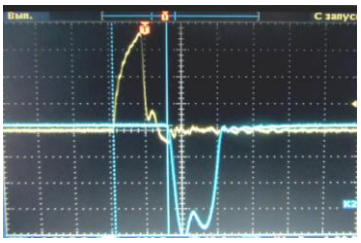


Рис. 4. Осциллограммы импульса источника излучения (верх) и импульса стробирования ЭОП (низ).

Длительность импульса подсвета равна 40 нс, а импульса стробирования 60 нс. Количество импульсов подсвета за время одного кадра соответствовало 5 импульсам.

По полученным изображениям объекта наблюдения в каждой активной зоне видения фиксировались усредненные уровни белого U_{max} и черного U_{min}, в областях изображений, отмеченных «рамками» на рис. 3 по которым производилось вычисление абсолютного K_1 и относительного K_2 контрастов

$$K_1 = \frac{U_{max}}{U_{min}}, \; K_2 = \frac{U_{max} - U_{min}}{U_{max}}.$$

Полученные данные измерений и вычислений сведены в таблицу 1.

Данные из таблицы 1 представлены на рис. 5 -7.

Рис. 5. Графики зависимостей уровней белого (пунктирная) и черного (сплошная) от задержки стробирования (чистая среда).

Рис. 6. График зависимости относительного контраста от задержки стробирования (чистая среда).

Рис. 7. График зависимости абсолютного контраста от задержки стробирования (чистая среда).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Таблица 1. Результаты измерений уровней белого и черного, относительного и абсолютного контрастов изображений объекта наблюдения в активных зонах видения (чистая среда).

Задержка стробирования, нс	Дальность, м	Уровень белого	Уровень черного	Относительный контраст	Абсолютный контраст
15	2,3	35,8	17,3	0,51	2
25	3,8	57,2	19,4	0,66	2,9
35	5,3	120,5	28,6	0,76	4,2
40	6	183	37,4	0,78	4,9
45	6,8	200	39	0,8	5
50	7,5	232,5	45,2	0,8	5,1
60	9	255	60	0,77	4,2
65	9,7	255	68,8	0,73	3,7
70	10,5	255	71	0,72	3,6
80	12	255	69,7	0,73	3,6
90	13,5	255	59,3	0,76	4,3
100	15	230	46	0,8	5
110	16,5	120,3	27	0,77	4,3
120	18	36,7	14,7	0,6	2,5
130	19,5	20,6	12	0,4	1,7

В виду того, что фотоприемное устройство АИ ТИС в диапазоне задержек стробирования от 60 нс до 90 нс находилось в насыщении по уровню белого, следовательно доверять значениям абсолютного и относительного контрастов на рис. 6, 7 в данном диапазоне некорректно, поэтому эти значения обозначены пунктирной линией на графиках контрастов.

Не смотря на ограничения по уровню белого можно сделать выводы по кривой уровня черного, которая симметрична и имеет явно выраженный максимум при задержке стробирования равной 70 нс что соответствует дальности согласно таблице 1 в 10,5 м. Объект наблюдения находится на расстоянии 15 м, что отстает от рассчитанной на 4,5 м или в пересчете в задержку стробирования на 30 нс что объясняется задержкой сигналов в АИ ТИС.

Графики контрастов (рис. 6, 7) в основном повторяют симметрию уровней белого и черного.

3. Томография в малопрозрачной среде

В целях исследования характеристик изображений получаемых с помощью АИ ТИС в малопрозрачной среде, объём БАК был заполнен дымами полученных путем пиролиза сосновых брусков.

В качестве объектов наблюдения использовались измерительные миры установленные на определенных расстояниях от АИ ТИС.

Эксперимент проводился таким же образом как и в условии чистой среды за исключением того, что объект наблюдения в данном случае располагается на расстоянии 21 м от АИ ТИС (рис. 8).

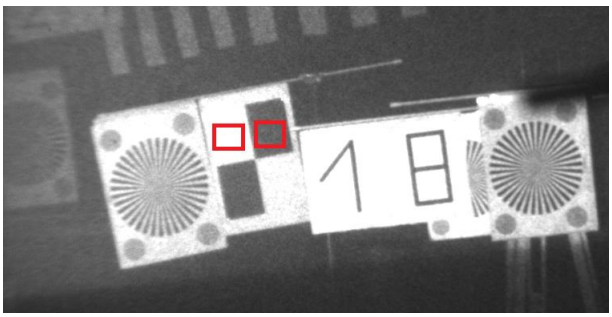


Рис. 8. Изображение объекта наблюдения в малопрозрачной среде.

По полученным изображениям объекта наблюдения в каждой активной зоне видения фиксировались усредненные уровни белого и черного, в областях изображений, отмеченных «рамками» на рис. 8 по которым производилось вычисление относительного и абсолютного контрастов. Полученные данные измерений и вычислений сведены в таблицу 2. При проведении эксперимента в малопрозрачной среде в целях предотвращения «пересвета» и дальнейшего повреждения фотоприемного устройства в АИ ТИС корректировалось количество импульсов подсвета за кадр: 29 импульсов при задержках стробирования 65 нс и 70 нс; 40 импульсов при задержке стробирования 80 нс; 59 импульсов при задержке стробирования 90 нс; 79 импульсов при задержке стробирования 100 нс; 99 импульсов при задержках стробирования от 110 до 160 нс.

Данные из таблицы 2 представлены на графиках рис. 9 – 11.

Рис. 9. Графики зависимостей уровней белого (пунктирная) и черного (сплошная) от задержки стробирования (малопрозрачная среда).

Рис. 10. График зависимости относительного контраста от задержки стробирования (малопрозрачная среда).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Таблица 2. Результаты измерений уровней белого и черного, относительного и абсолютного контрастов изображений объекта наблюдения в активных зонах видения (малопрозрачная среда).

Задержка стробирования, нс	Дальность, м	Уровень белого	Уровень черного	Относительный контраст	Абсолютный контраст
65	9,7	104,6	86,6	0,17	1,2
70	10,7	100	80,5	0,19	1,24
80	12	107,5	74,6	0,31	1,44
90	13,5	158,5	84,4	0,47	1,88
100	15	216,7	88,4	0,6	2,45
110	16,5	254,8	103,6	0,6	2,46
120	18	246,3	75,1	0,7	3,28
133	20	180	51,9	0,71	3,46
140	21	148,6	42	0,72	3,54
150	22,5	102,2	31,5	0,7	3,2
160	24	40,6	19,1	0,53	2,1

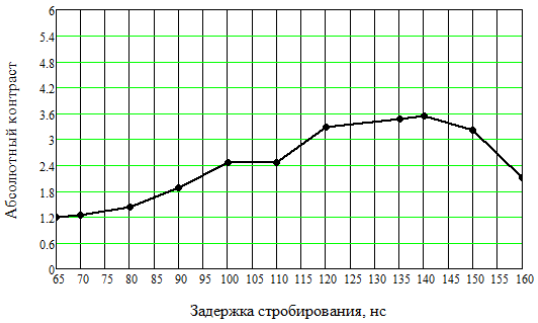


Рис. 11. График зависимости абсолютного контраста от задержки стробирования (малопрозрачная среда).

Анализируя графики уровней белого и уровня черного для малопрозрачной среды можно прийти к следующим выводам: графики (рис. 9) менее симметричны по сравнению с аналогичными графиками для чистой среды, имеют подъем обоих уровней при малых задержках стробирования, а при задержке стробирования в 110 нс явно выраженные максимумы уровней белого и черного, что согласно таблице 2 соответствует дальности в 16,5 м; разница с истинным расстоянием до объекта наблюдения составляет 4,5 м, при этом он находится на расстоянии 21 м, что в свою очередь при пересчете в задержку стробирования составляет 30 нс, как и в случае с чистой средой.

Следовательно в обоих экспериментах присутствует некоторая одинаковая аппаратурная погрешность в определении задержки стробирования в 30 нс.

4. Заключение

В виду особенностей режимов работы АИ ТИС по сравнению с обычными системами видеонаблюдения АИ ТИС позволяет осуществить томографическое (послойное) видение окружающего пространства, как в обычных, так и в сложных условиях видения, что позволяет производить накопление информации об конкретных объектах, находящихся в активной зоне видения системы и с учетом заранее известных погрешностей получать достаточно точную оценку характеристик наблюдаемых объектов, например, расстояние до них.

По результатам экспериментальных исследований АИ ТИС, применяя метод томографического видения окружающего пространства, получены зависимости абсолютного и относительного контрастов, которые для малопрозрачной среды имеют явную асимметрию из–за воздействия помехи обратного рассеяния. Показана возможность определения расстояния до конкретного объекта наблюдения по максимальному сигналу фотоприемного устройства при соответствующей задержке стробирования.

Благодарности

Работа выполнена при финансовой поддержке Минобрнауки России по проекту № 8.9562.2017/8.9 и РФФИ в рамках научного проекта № 16-47-700939.

5. Литература

[1] Горелик С.Л. Телевизионные измерительные системы / С.Л. Горелик, Б.М. Кац, В.И. Киврин. – М.: Связь,1980. – 168 с.

[2] Kapustin V.V. Vision area parameters analysis for active-pulse television-computing systems / V.V. Kapustin, A.K. Movchan, M.I. Kuryachiy // 2017 International Siberian Conference on Control and Communications (SIBCON). - 2017. –P. 1 –4.

[3] Kapustin V.V., Movchan A.K., Zaytseva E.V., Kuryachy M. I. / Active pulse television measuring systems for ensuring navigation of transport means in heavy weather conditions // Transportation systems and technology. – 2018. – Т. 4. – №. 1. – С. 068-083.

[4] Левин Г.Г., Вишняков Г.Н. Оптическая томография. М.: Радио и связь, 1989.-224 с.

[5] Wang Xinwei, Zhou Yan, Fan Songtao, Liu Yuliang, Liu Hongjun. «Echo Broadening Effect in Range-Gated Active Imaging Technique» Proc. of SPIE Vol. 7382 738211-9 (2009).

Об авторах

Мовчан Андрей Кириллович, аспирант кафедры телевидения и управления (ТУ) ТУСУР. E–mail: mr.movchann@mail.ru.

Капустин Вячеслав Валериевич, к.т.н., младший научный сотрудник кафедры ТУ ТУСУР. E–mail: peregnun@mail.ru.

Курячий Михаил Иванович, к.т.н., с.н.с., доцент кафедры ТУ ТУСУР. E–mail: kur@tu.tusur.ru.

.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Ускорение метода перебора плоскостей при трехмерной реконструкции строений по аэроснимкам

Л.В. Новоторцев[1], А.Г. Волобой[2]
torets13@gmail.com|voloboy@gin.keldysh.ru
[1]1Институт Прикладной Математики им. Келдыша РАН, Москва, Россия;
[2]АО "Фирма Ракурс", Москва, Россия;

В фотограмметрии нередко встречается задача трехмерной реконструкции строений по набору аэроснимков. Многие алгоритмы, решающие данную задачу, либо имеют невысокую точность, либо низкую скорость работы. В данной статье предлагается использовать модификацию метода перебора плоскостей для достижения высокой скорости работы при сохранении точности реконструкции.

Ключевые слова: *аэроснимки, трехмерная реконструкция, компьютерное зрение.*

Plane-sweep stratagy acceleration in the task of 3d building reconstruction from aerophoto

L.V. Novotortsev[1], A.G. Voloboy[2]
torets13@gmail.com|voloboy@gin.keldysh.ru
[1]Keldysh Institute of Applied Mathematics RAS, Moscow, Russia;
[2]Racurs, Moscow, Russia;

One of the common tasks of phogrammetry is three-dimentional bulding reconstruction. Most of methods which solve that problem are either unaccurate or slow. In this paper we propose to use modification of plane-sweep stratagy method in ortder to achive high computation speed without loosing accuracy.

Keywords: *aerophoto, 3d reconstruction, computer vision.*

1. Введение

Задача распознавания и реконструкции строений на снимках, полученных при помощи аэро- или космосъемки, является распространенной в рамках фотограмметрии.

Данная задача часто решается в рамках автоматического создания или обновления электронных карт, где наличие трехмерных моделей значительно увеличивает наглядность продукта и упрощает чтение карты.

Другая сфера, где рассматриваемая задача активно используется, — анализ застройки. К примеру, это необходимо для поиска выгодных мест для новой застройки и дальнейшей ее продажи. Также анализ застройки может использоваться в природоохранных целях: обнаружение нелегальной или опасной застройки

Характерной особенностью рассматриваемой задачи является большой объем исходных данных. В некоторых случаях размер изображений достигает нескольких гигабайтов. Ручная обработка данных такого размера представляет из себя дорогую, а в некоторых случаях невыполнимую задачу. Более того, развитие способов съемки (БПЛА), скорости передачи данных при быстрой потери их актуальности, приводит к острой необходимости использовать автоматические методы построения. Однако точность автоматических методов уступает ручному сопоставлению, что вызывает необходимость совершенствования, модификации и разработки новых алгоритмов. Исследованию этой задачи посвящена данная работа.

2. Обзор существующих методов

Основные распространенные методы трехмерной реконструкции можно разделить на следующие группы:

1. построение на основе данных лидар'а;
2. построение на основе набора растровых снимков
   (a) по облаку точек;
   (b) по отрезкам.

Построение на основе данных лидар'а. Лидар (LIDAR англ. Light Identification Detection and Ranging — обнаружение, идентификация и определение дальности с помощью света) — технология получения и обработки информации об удаленных объектах с помощью активных оптических систем, использующих явления поглощения и рассеяния света в оптически прозрачных средах [8, 13].

Реконструкция сцены с использованием этих данных значительно облегчена, поскольку сами данные уже содержат информацию о высоте в каждой точке. Недостатком использования данного типа данных для реконструкции строений является большая стоимость оборудования и полета. Несмотря на большую вариативность методов опирающихся на использование данных лидар'а [10] (в том числе и построение облака точек), их общий недостаток - стоимость, сильно ограничивает их применение. По этой причине их более подробное рассмотрение выходит за рамки представленной работы.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Построение по облаку точек. Создание облака точек является довольно затратной задачей, требующей больших вычислительных затрат, которые значительно увеличиваются при отсутствии хорошей калибровки изображений. Результат, получаемый этим методом представляет собой высокоточную реконструкцию сцены [5, 7].

К недостаткам представленного метода можно отнести очень медленную скорость работы и необходимость дополнительного анализа полученной трехмерной сцены для выделения из нее строений, что еще больше увеличивает вычислительные затраты методов, основанных на подобном подходе.

Сочетание представленных недостатков с тем фактом, что задача выделения строений по уже построенной трехмерной модели является нетривиальной, делает данный подход не оптимальным в рамках рассматриваемой задачи, несмотря на высокую распространенность и эффективность решения задач, связанных с реконструкцией сцен по нескольким снимкам.

Построение по выделенным отрезкам. Данный метод подразумевает предварительный анализ каждого из снимков, выделение отрезков и последующее их сопоставление [3]. Этот подход обладает большой гибкостью и скоростью. Более того, поскольку строения в своем большинстве имеют линейные черты, то объекты, выделенные данным подходом, в основном будут являться строениями. Дополнительный фильтр с использованием таких простых метрик, как размер объектов, почти полностью исключает ложные срабатывания. Это существенно облегчает одну из частей поставленной задачи: обнаружение строений.

Также стоит отметить, что в большинстве случаев аэроснимки поставляются с данными элементов внешнего ориентирования (ЭВО), которые используется построения первого приближения сопоставления снимков. Далее по тексту предполагается, что ЭВО есть и используется.

Методы, основанные на данном походе, уступают по точности реконструкции тем, что основываются на построении облака точек, но выигрывают по скорости работы (до 30 раз быстрее) [4].Переставленные преимущества являются крайне важными в рамках задачи реконструкции и обнаружения строений по аэроснимкам.

Одним из наиболее распространенных подходов, относящихся к представленной группе, является метод перебора плоскостей [3] и его модификация [12]. Данный алгоритм дает высокую точность обнаружения, но работает значительно медленнее подобных алгоритмов. Представленная работа посвящена разработке метода ускорения рассматриваемого подхода без значительной потери точности.

3. Описание метода

Как уже упоминалось выше данный подход состоит из следующих этапов:

1. выделение отрезков на каждом изображении;
2. связывание соседних отрезков;
3. сопоставление отрезков;
4. предварительное построение плоскостей;
5. валидация и уточнение плоскостей;
6. построение результата.

Из приведенных этапов, пункты 3 и 5 являются новыми шагами, представленными рассматриваемым методом, этап 4 является существенно модифицированным шагом алгоритма перебора плоскостей. Остальные этапы либо являются темой отдельного исследования (пункты 1 и 2), либо существенно не отличаются от аналогичных этапов метода перебора плоскостей (пункт 6).

3.1. Выделение отрезков

Рассматриваемый алгоритм сильно зависит от результата, полученного в результате операции выделения отрезков. Несмотря на то, что в рамках представленного метода пропущенные и разорванные отрезки частично компенсируются при реконструкции, пропуск слишком большого числа отрезков может привести к существенному ухудшению результата. По этой причине предлагается использовать алгоритм LSD [6] или алгоритмы, основанные на нем.

3.2. Связывание соседних отрезков

После того, как были выделены отрезки со снимков, в данной работе предлагается выполнить связывание соседних отрезков. Иными словами - соединить отрезки в ломанные на каждом из снимков (Рис. 1).

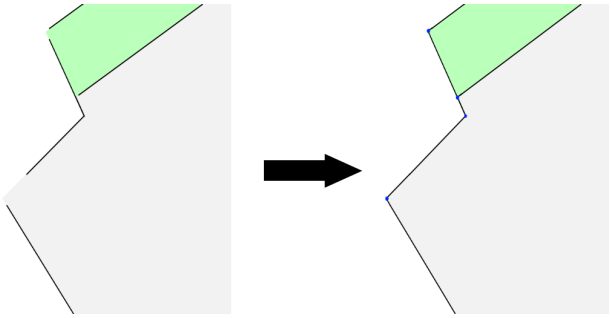


Рис. 1. Объединение отрезков в ломаную.

Для выполнения этой операции предлагается продолжить проверяемые отрезки прямыми и вычислить точку пересечения прямых. Далее выполняется проверка расстояния, на которое необходимо продолжить отрезки до пересечения. Для этого определяются входные параметры: нижний и верхний порог для отбраковки. Величина этих параметров определяется масштабом и качеством снимков, а также характером объектов на снимках (плотность и размер строений). Если расстояние от концов отрезков до точки пересечения меньше нижнего порога, то отрезки продлеваются без дополнительных проверок, если больше верхнего порога,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

то никаких операций не требуется. В противном случае, если расстояние больше нижнего порога, но меньше верхнего, то выполняется дополнительный анализ поля градиентов в той области, в которой должно проходить продление отрезка до точки пересечения [1].

Также важно отметить, что в случае, если точка пересечения продленных отрезков находится на одном из них, то отрезок разбивается на два и в дальнейших вычислениях воспринимается как два отрезка.

3.3. Сопоставление отрезков

После того, как отрезки были выделены на каждом из изображений, необходимо выполнить их сопоставления: связывание отрезков на разных снимках, соответствующих одним и тем же граням объектов на сцене. Данный этап оказывает существенное влияние на итоговый результат. По этой причине необходимо использовать алгоритмы, обеспечивающие высокую точность сопоставления, к примеру SIFT [11] или ORB [9].

Данные о связывании соседних отрезков, полученные на предыдущем этапе, также могут использоваться для ускорения сопоставления. Связи между отрезками можно использовать в качестве подсказок: связанные (пересекающиеся) отрезки имеет смысл сопоставлять со связанными отрезками на другом снимке [2].

3.4. Предварительное построение плоскостей

На данном этапе в методе перебора плоскостей предлагается для каждого отрезка перебирать все плоскости, образуемые им (все углы поворота вокруг отрезка). Далее, на проекции плоскости на снимок выбираются точки интереса, и сопоставляются с соответствующими точками на других снимках (соответствующие точки находятся за счет преобразования проекции плоскости с одного снимка на другой при использовании информации об ориентировании снимков) [3].

Данная операция дает точные результаты, но полный перебор плоскостей занимает слишком много времени (минуты), особенно если использовать маленький шаг дискретизации угла [3] и является самым длительным этапом алгоритма (до 50% вычислительного времени). Далее считаем, что отрезки перебираются от самого длинного в порядке убывания.

В данной работе предлагается решать эту проблему при помощи информации о связанных отрезках из раздела 3.2. Далее для определенности будем называть отрезок, относительно которого вращаем полуплоскость, — отрезком-осью. Связанный с ним — отрезком-предположением. Используя данные о связи сопоставленных отрезков можно сделать оценку угла наклона полуплоскости относительно отрезка-оси и применять метод перебора плоскостей в окрестности угла, определяемым связанным отрезком (два отрезка однозначно определяют полуплоскость, но из-за погрешностей и ошибок сопоставления исключительно на это полагаться нельзя). Диапазон перебора углов относительно предположительного угла и шаг перебора являются входными параметрами алгоритма и определяются качеством снимков и характером объектов изображенных на них.

Если брать отрезок-подсказку как есть, то велика вероятность, что из-за ошибок сопоставления и 3d-реконструкции полуплоскость на рассматриваемом отрезке-оси и отрезке-подсказки не будет найдена. По этой причине для нахождения начального приближения полуплоскости используются трехмерные координаты точки пересечения отрезков и дальнего конца отрезка-подсказки. Также стоит учитывать, что подсказка оказалась ложной и искомой полуплоскости не существует. Это определяется слабой корреляцией точек интереса при использовании алгоритма перебора плоскостей.

Если у рассматриваемого отрезка есть несколько связанных с ним отрезков, то описанная выше процедура применяется для каждого из них.

3.5. Валидация и уточнение плоскостей

На предыдущем этапе были получены полуплоскости с большой точностью, однако есть возможность ее увеличить, если есть такая необходимость. Для этого рассматриваются отрезки, находящиеся на проекции плоскости на изображения. Из этих отрезков отфильтровываются короткие для исключения шума и деталей крыши (обычно достаточно фильтровать все отрезки короче трех метров в трехмерном пространстве, но конкретная величина зависит от качества и вида снимков). Далее отфильтровываются отрезки, расстояние концов которых до плоскости превышает погрешность при построении. Данная величина высчитывается на основе шага угла при переборе плоскостей и размера рассматриваемого отрезка: расстояние от одного из концов отрезка до плоскости при его отклонении от нее на угол, равный шагу перебора. Также не рассматриваются отрезки, находящиеся дальше удвоенного размера отрезка-оси.

Из оставшихся отрезков выбираются не более десяти имеющих наименьшее отклонения. И по трехмерным координатам концов отрезков проводится линейная регрессия для уточнения плоскости. Данная операция сходится быстро благодаря точному начальному приближению, полученному на предыдущем этапе.

3.6. Построение результата

На данном этапе предлагается использовать аналогичный этап метода перебора плоскостей без существенных изменений.

Суть этого этапа заключается в ограничивании полуплоскостей отрезками, лежащими на них, и выделении новых граней за счет пересечения полуплоскостей.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4. Результаты и выводы

На рис. 2 представлен пример результата работы представленного метода.

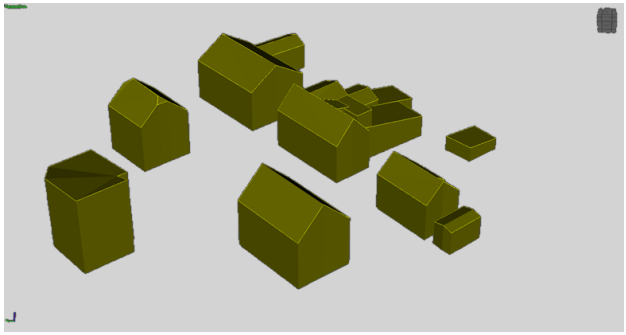


Рис. 2. Объединение отрезков в ломаную.

В таблице 1 приведены данные сравнения представленного метода с методом перебора плоскостей и с методом [5], основанном на использовании облака точек. Представленные методы также были реализованы автором данной работы. Данные приведены для набора из 6 фрагментов снимков размером 1028x891x24. Сцена, представленная на изображениях содержит 8 строений.

Метрика	Метод Baillard'00	Метод Dahlke'15	Данный метод
Точность по плану (пикс)	0.89	0.68	0.81
Точность по высоте (пикс)	2.98	2.01	2.76
Время (сек)	68	236	26

Таблица 1. Таблица 1. Сравнение методов реконструкции

Из результатов, приведенных в таблице 1, следует что представленный метод работает быстрее, чем метод перебора плоскостей и предоставляет точность, не уступающую этому методу. Однако описанный алгоритм уступает более точному алгоритму, использующему облако точек. Данная разница объясняется компромиссом между точностью и скоростью работы.

5. Литература

[1] Новоторцев Л.В., Волобой А.Г. Автоматическое распознавание зданий на аэрофотоснимках // Труды 16-ой международной конференции CAD/CAM/PDM-2016 "Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта". — 2016. — С. 68-72.

[2] Волобой А. Г., Новоторцев Л. В. Сопоставление областей на аэрофотоснимках в задаче распознавания строений // Новые информационные технологии в автоматизированных системах.Материалы двадцать первого научно-практического семинара. — 2018. — С. 17-23.

[3] Baillard C., Zisserman A. A plane-sweep strategy for the 3D reconstruction of buildings from multiple images // International Archives of Photogrammetry and Remote Sensing. — 2000. — Vol. 33(B2). — P. 56-62.

[4] Dahlke D., Linkiewicz M. Comparison between two generic 3d building reconstruction approaches — point cloud based vs. image processing based // International Archives of Photogrammetry and Remote Sensing. — 2016. — Vol. 41(B3) — P. 599-604.

[5] Dahlke D., Linkiewicz M., Meissner H. True 3D building reconstruction: façade, roof and overhang modelling from oblique and vertical aerial imagery // International Journal of Image and Data Fusion. — 2015. — Vol. 6(4). — P. 314-329.

[6] von Gioi R. G. et al. LSD: A fast line segment detector with a false detection control // IEEE transactions on pattern analysis and machine intelligence. — 2010. — Vol. 32(4). — P. 722-732.

[7] Malihi S. et al. 3D Building Reconstruction Using Dense Photogrammetric Point Cloud //International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Prague, Czech Republic. — 2016. — Vol. 3. — P. 71-74.

[8] Middleton W. E. K, Spilhaus A. F., Meteorological instruments, University of Toronto, 3rd ed. 1953.

[9] Rublee E. et al. ORB: An efficient alternative to SIFT or SURF // 2011 IEEE international conference on Computer Vision (ICCV). —2011. — P. 2564-2571.

[10] Wang R., Peethambaran J., Chen D. LiDAR Point Clouds to 3-D Urban Models : A Review // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. — 2018. — Vol. 11(2). — P. 606-627.

[11] Wang X., Fu W. Optimized SIFT image matching algorithm // IEEE International Conference on Automation and Logistics. — 2008. — P. 843-847.

[12] Wang Y. C. et al. 3D Reconstruction of Piecewise Planar Models from Multiple Views Utilizing Coplanar and Region Constraints // Journal of Information Science and Engineering. — 2013. — Vol. 29(2). — P. 361-378.

[13] Wu B. et al. A graph-based approach for 3D building model reconstruction from airborne LiDAR point clouds //Remote Sensing. — 2017. — Vol. 9(1). — P. 92.

Об авторах

Новоторцев Леонид Владимирович, Программист-разработчик АО "Фирма Ракурс". Его e-mail torets13@gmail.com.

Волобой Алексей Геннадьевич, доктор физ.-мат. наук, ведущий научный сотрудник; Институт прикладной математики им. М.В. Келдыша РАН. Его e-mail voloboy@gin.keldysh.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Программа и алгоритм сегментации и распознавания рукопечатных символов с помощью сверточных нейронных сетей

Е.С. Попова, В.Г. Спицын, Ю.А. Болотова
esp9@tpu.ru|spvg@tpu.ru|jbolotova@tpu.ru
Национальный Исследовательский Томский Политехнический Университет, Томск, Россия

Статья посвящена разработке алгоритма сегментации и распознавания рукопечатных символов на изображениях. Излагается обобщенный алгоритм работы системы распознавания текстов. Приводится описание методов сегментации текстовых документов. Предлагается применение нейросетевого подхода, основанного на архитектуре сверточных нейронных сетей, для решения задачи распознавания рукопечатных символов. Результаты численных экспериментов по распознаванию бланков ЕГЭ на основе предложенного подхода показали точность распознавания 94.1%, что превосходит точность распознавания 82%, полученную с помощью системы ABBYY FineReader.

Ключевые слова: *компьютерное зрение, распознавание образов, сверточные нейронные сети, сегментация текстовых изображений.*

Program and algorithm for segmentation and recognition of hand-printed characters using convolutional neural networks

E.S. Popova, V.G. Spitsyn, Yu.A. Bolotova
esp9@tpu.ru|spvg@tpu.ru|jbolotova@tpu.ru
National Research Tomsk Polytechnic University, Tomsk, Russia

The article is devoted to the development of the algorithm for segmentation and recognition of hand-printed symbols on images. A generalized algorithm for the operation of the text recognition system is presented. The paper describes the methods of segmentation of text documents. The application of a neural network approach based on the architecture of convolutional neural networks is proposed to solve the problem of recognizing hand-printed characters. Results of numerical experiments on the recognition of USE forms on the basis of the proposed approach showed a recognition accuracy of 94.1%, which exceeds the recognition accuracy of 82% obtained with the help of ABBYY FineReader.

Keywords: *computer vision, pattern recognition, convolutional neural networks, segmentation of text images.*

1. Введение

На сегодняшний день существует множество направлений науки и техники, которые в значительной степени ориентированы на развитие систем, анализирующих информацию, представленную в виде изображений. Задача обработки и распознавания изображений относится к разряду трудно формализуемых задач, и является одной из наиболее важных на сегодняшний день.

Задача сегментации и распознавания текстовых областей, содержащих рукопечатные символы, является актуальной, так как в мире существует большое количество документов, нуждающихся в оцифровке.

Рукописные цифры или буквы, очевидны для человека, но для компьютеров идентификация таких символов — сложная задача, для решения которой часто применяются сверточные нейронные сети, являющиеся наиболее подходящими для анализа изображений.

Свёрточная нейронная сеть впервые была представлена в 1998 году французским исследователем Яном Лекуном, как развитие модели неокогнитрона, предназначенного для эффективного распознавания изображений.

Процесс распознавание текстовых областей на изображениях условно можно разделить на два этапа: этап сегментации изображения на строки, слова и символы, и этап распознавания символов.

Целью данной работы является разработка алгоритма сегментации и распознавания бланков ЕГЭ, которые на сегодняшний день могут обрабатываться как вручную, так и с помощью специализированного ПО. Однако, используемое программное обеспечение не всегда дает достаточную точность распознавания, необходимую для полной автоматизации процесса.

В данной работе проводится сравнение различных архитектур нейронных сетей и алгоритмов обучения, основанных на алгоритме градиентного спуска для решения поставленной задачи. Так же приводится описание алгоритма для сегментации текстового изображения, основанного на комбинации метода связанных компонент и порогового метода анализа гистограмм.

Далее в статье будут рассмотрены методы и алгоритмы, используемые для решения поставленной задачи.

2. Сегментация текстового изображения

2.1 Бинаризация изображения

Прежде чем произвести бинаризацию изображения, его необходимо перевести в градации серого, яркость пикселя вычисляется по формуле [1-3]:

$$I = R * 0.299 + G * 0.587 + B * 0.114,$$

где R, G, B — красный, зеленый и синий канал соответственно.

Для последующей бинаризации изображения яркость каждого пикселя $I(x,y)$ сравнивается с некоторым пороговым значением P. Если яркость пикселя больше порогового значения, то цвет пикселя принимается равным 1, иначе 0.

Глобальный порог бинаризации вычисляется по формуле [8, 9]:

$$P = \frac{I_{max}+I_{min}}{2},$$

где I_{max} – максимальное значение яркости изображения, I_{min} – минимальное значение яркости изображения.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2.2 Сегментация изображения на строки и слова

Для сегментации фрагмента текста на строки и слова используется метод гистограмм.

Метод предполагает построение гистограммы для черных пикселей бинарного изображения. По оси *Y* располагается шкала распределения пикселей по количеству, а на оси *X* размещены номера строк, в случае сегментирования изображения на строки.

Алгоритм основан на предположении, что количество черных пикселей в межстрочных интервалах существенно меньше, чем в текстовых строках. Основываясь на этом предположении, определим каким должно быть наименьшее количество черных пикселей в строке, чтобы отнести ее к текстовой строке. Рассчитаем значение по формуле:

$$I = 0.1 * Str_{max},$$

где Str_{max} – максимальное число черных пикселей в строке изображения. Следовательно, используя найденный порог разделим все изображение на строки.

Работа алгоритма сегментации строк заключается в последовательном просмотре массива, содержащего количество черных пикселей для каждой строки и сравнение их с минимальным количеством N, затем выявлении множества пар индексов (s^1_i, s^2_i) строк, соответствующих границам печатных строк.

Аналогично осуществляется сегментация строк на слова, только в данном случае условием выделения пробелов является последовательность из *K* белых пикселей в строке изображении.

2.3 Метод связанных компонент

Для сегментации слов на символы в работе используется двухпроходной метод связанных компонент (МСК) [3,4]. Под выделением связных компонент понимают присвоение уникальной метки каждому объекту изображения. При следующем анализе данные метки служат в качестве идентификаторов при обращении к объектам.

Для описания алгоритма введем некоторые понятия. Обозначим через *I* матрицу изображения. Если $I(i, j) = 0$, то пиксель является фоновым. Если $I(i, j) = 1$, то пиксель принадлежит объекту. Через *L* обозначим двумерную матрицу (скан-маску), которая используется для хранения информации о метках и имеющую размеры, равные размерам изображения.

Первый проход по изображению осуществляется из верхнего левого угла, слева на право и сверху вниз. На внешнем цикле – перебор строк, на внутреннем – перебор столбцов строки, анализируются только соседи сверху и слева.

Каждый раз при обнаружении черного пикселя его соседи, принадлежащие скан-маске, исследуются для определения метки, которая будет присвоена рассматриваемому пикселю. Если в скан-маске содержатся только фоновые пиксели, то рассматриваемый пиксель получает новую промежуточную метку, если скан-маска содержит только один пиксель интереса, то рассматриваемый пиксель получает метку соседа. Если скан-маска содержит несколько точек интереса, то их промежуточные метки являются эквивалентными, среди них выбирается метка с наименьшим значением и пикселю присваивается значение выбранной метки-представителя.

После окончания первого прохода каждый объектный пиксель получает временную метку, на втором проходе значение метки уточняется.

Второй проход осуществляется в противоположенном направлении, снизу-вверх и справа налево, на втором проходе исследуются все соседние пиксели.

Следовательно, на втором проходе осуществляется поиск связей между маркированными пикселями различных категорий. Если связь найдена, то все «старшие» метки заменяются на «младшие».

3. Распознавание символов

3.1. Обоснования выбора сверточных нейронных сетей

Свёрточные нейронные сети, в отличие от других нейросетевых архитектур, обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям. Они объединяют три архитектурные идеи, для обеспечения инвариантности к изменению масштаба, повороту сдвигу и пространственным искажениям [6]:

- локальные рецепторные поля (обеспечивают локальную двумерную связность нейронов);
- общие синаптические коэффициенты (обеспечивают детектирование некоторых черт в любом месте изображения и уменьшают общее число весовых коэффициентов);
- иерархическая организация с пространственными подвыборками.

3.2. Обучающая выборка

В качестве наборов данных для обучения и тестирования будет использоваться обучающая выборка, содержащая 146324 изображений рукописных символов, включающая цифры и буквы русского алфавита.

Изображения в наборе данных имеют разрешение 32×32 пикселей и хранятся в формате оттенков серого, следовательно, каждое значение пикселя лежит в диапазоне от 0 (представляет белый цвет) до 255 (представляет черный цвет).

Для ускорения работы сети необходимо инвертировать значения пикселей следующим образом 0 – белый цвет, 255 – черный цвет, по следующей формуле:

$$I = 255 - x_i.$$

Для ускорения сходимости обучения сети значения входных пикселей изображений нормализуются по формуле:

$$y_i = \frac{x_i}{255},$$

где x_i– значение *i*-го пикселя изображения из базы, y_i – значение, подаваемое на вход сети.

3.3. Алгоритм обучения

В данной работе используется наиболее распространенный алгоритм обучения нейронных сетей, основанный на методе градиентного спуска (метод обратного распространения ошибки) и его модификации: SGD, Adam, AdaGrad, AdaDelta. [7, 10-13].

Алгоритм обратного распространения ошибки имеет несколько режимов обучения [7]. В данной работе используется подход с Mini-batch, как компромисс между последовательным и пакетным режимами, в этом случае корректировка синаптических весов сети происходит после небольшого количества обучающих образцов.

3.4. Функция активации

В данной работе используются функции активации ReLU и для выходного слоя Softmax.

ReLU является выпрямленной линейной функцией и на данным момент считается гораздо более простым и эффективным с точки зрения вычислительной сложности вариантом передаточной функции [5]:

$$f(S) = \max(0, S) = \begin{cases} 0, \text{при } S < 0 \\ S, \text{при } S \geq 0 \end{cases}.$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

На сегодняшний день существует семейство различных модификаций функции ReLU, решающих проблемы надёжности этой передаточной функции при прохождении через нейрон больших градиентов: Leaky ReLU, Parametric ReLU, Randomized ReLU.

Softmax функция активации разработана, чтобы превратить любой вектор с реальными значениями в вектор распределения вероятностей и определяется для *i*-ого нейрона следующим образом [7]:

$$z_i = \frac{\exp(y_i)}{\sum_{j=1}^{n} \exp(y_j)},$$

где z_i – искомое значение выхода *i*-ого нейрона, y_i – исходное значение выхода *i*-ого нейрона.

3.5. Функция потерь

Так как в качестве функции активации в выходном слое будет использоваться функция активации softmax, которая преобразует любой входной вектор в вектор вероятностей, то для сравнения двух вероятностных распределений необходимо выбрать корректную меру. В качестве такой меры будет использоваться перекрестная энтропия [3, 7]:

$$\mathrm{C} = -\sum_{j=1}^{n} t_j \log(y_j),$$

где t_j – требуемый выход для текущего обучающего примера, y_j – реальный выход нейронной сети.

3.6. Выбор библиотеки машинного обучения

Для подбора оптимальных параметров разрабатываемой сети, был произведен обзор существующих библиотек машинного обучения. На основе выбранной библиотеки спроектирована нейронная сеть с различными параметрами, для выявления наиболее подходящей конфигурации сверточной сети.

В качестве библиотеки машинного обучения для проектирования и тестирования сети была выбрана открытая нейросетевая библиотека Keras, которая позволяет на более высоком уровне работать с нейросетями. В качестве базовой библиотеки для вычислений Keras может использовать Theano и Tenzorfow, в данной работе использовалась библиотека машинного обучения Theano.

3.7. Обучение сети

Для решения поставленной задачи была спроектирована сверточная нейронная сеть с одним сверточным и подвыборочным слоем, размер матрицы свертки 5×5, размер окна подвыборки 2×2. Конфигурация сети представлена в таблице 1.

Таблица 1. Конфигурация сверточной нейронной сети

Тип слоя	Функция активации	Кол-во настраиваемых параметров
Слой свертки, Кол-во карт признаков: 32	ReLU	832
Слой подвыборки Кол-во карт признаков: 32	–	–
Полносвязный слой, кол-во нейронов:128	ReLU	802944
Полносвязный слой, кол-во нейронов:32	Softmax	4128
Общее кол-во настраиваемых параметров		807 904

В таблицах 2 и 3 приведен анализ влияния количества эпох обучения и размера минивыборки на точность распознавания сети для различных модификаций метода градиентного спуска.

Таблица 2.Влияние количества эпох на точность распознавания.

Размер мини выборки = 100				
Кол-во эпох	Алгоритмы оптимизации			
	SGD	Adam	AdaGrad	AdaDelta
5	81.03%	94.53%	96.14%	93.09%
10	84.41%	95.18%	96.30%	94.21%
20	88.42%	95.98%	96.46%	95.82%

Таблица 3. Влияние размера минивыборки на точность распознавания.

Количество эпох = 10				
Размер мини выборки	Алгоритмы оптимизации			
	SGD	Adam	AdaGrad	AdaDelta
50	89.23%	95.98%	96.46%	95.18%
100	84.41%	95.18%	96.30%	94.21%
200	83.92%	94.53%	96.46%	94.21%

Так же было проанализировано влияние количества карт признаков на точность распознавания, результат представлен в таблице 4.

Таблица 4. Влияние количества карт признаков на точность распознавания

Количество эпох = 10, размер минивыборки = 100				
Количество карт признаков выборки	Алгоритмы оптимизации			
	SGD	Adam	AdaGrad	AdaDelta
32	84.41%	95.18%	96.30%	94.21%
64	85.21%	95.34%	96.78%	95.98%

Исходя из полученных результатов были выбраны оптимальные параметры обучения и конфигурации сети:

- Алгоритм оптимизации: AdaGrad
- Количество эпох обучения: 10
- Размер минивыборки: 200
- Количество карт признаков: 32

4. Результаты тестирования

Для оценивания результатов классификации была рассчитана доля верно классифицированных объектов к общему количеству объектов. На данный момент лучшая точность распознавания на тестовой выборке, которой удалось добиться при реализации выбранной модели составляет 94.1%. Значение точности вычислялось по формуле:

$$R = \frac{n}{N} = \frac{623}{662} = 0.941\ ,$$

где R – точность распознавания по всему набору тестовой выборки, n – количество правильно распознанных символов из тестовой выборки, N – количество элементов в тестовой выборке.

На рисунке 1 представлены некоторые примеры неверного распознавания символов различных категорий.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

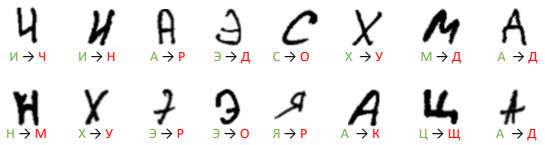


Рис. 1. Примеры не верно распознанных символов

Были проведены исследования по распознаванию бланков ЕГЭ с помощью системы ABBYY FineReader, которая показала точность распознавания 82%. В связи с этим, можно сделать вывод о том, что разрабатываемая система является перспективной, а дальнейшее ее улучшение – востребованным.

5. Заключение

В ходе работы опытным путем была выявлена наиболее подходящая архитектура сверточной нейронной сети для решения поставленной задачи, исследованы алгоритмы оптимизации на основе метода градиентного спуска и виды функции активации нейронов. Все тесты проводились с использованием базы данных рукописных символов русского алфавита.

Были исследованы алгоритмы сегментации текстовых изображений, на основе пороговых алгоритмов и метода связанных компонент.

6. Благодарности

Работа выполнена в рамках Программы повышения конкурентоспособности ТПУ при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А.

7. Литература

[1] Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. Цифровая обработка изображений в информационных системах: Учеб. пособие. – Новосибирск.: Изд-во НГТУ, 2003. – 352 с.

[2] Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.

[3] Маркелов А.А. Алгоритмы и программная система классификации полутоновых изображений на основе нейронных сетей: Диссертация на соискание ученой степени кандидата технических наук / А. А. Маркелов. – Томск, 2007.

[4] Поршнев С.В., Левашкина А.О., Универсальная классификация алгоритмов сегментации изображений // Журнал научных публикаций аспирантов и докторантов [Электронный ресурс] — Электронный научный журнал – 2006. –Режим доступа: http://jurnal.org/articles/2008/inf23.html

[5] Спицын В.Г., Интеллектуальные системы: учебное пособие / В.Г. Спицын, Ю.Р. Цой; Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2012. – 176 с.

[6] Солдатова О.П., Гаршин А.А. Применение сверточной нейронной сети для распознавания рукописных цифр. Компьютерная оптика. – 2010. – Том 34,№2. – с. 252-260 – ISSN 0134-2452.

[7] Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2006. - 1104 с.

[8] Шапиро Л. Компьютерное зрение / Л. Шапиро, Дж. Стакан, 206. – 53 с.

[9] Введение в цифровую обработку изображений: лекция 3. [Электронный ресурс] – 2011. – Режим доступа: http://cvbeginner.blogspot.ru/2011/09/3.html

[10] LeCun, Y. Efficient BackProp in Neural Networks: Tricks of the trade / Y.LeCun, L. Bottou, G. Orr, K. Muller – Springer, 1998.

[11] LeCun, Y. Scaling learning algorithms towards AI / Y.LeCun, Y. Bengio – MIT Press, 2007.

[12] Mike O'Neill. Neural Network for Recognition of Handwritten Digits. [Электронный ресурс] Точка доступа: https://www.codeproject.com/Articles/16650/Neural-Network-for-Recognition-of-Handwritten-Digi

Об авторах

Спицын Владимир Григорьевич, д.т.н., профессор инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail spvg@tpu.ru.

Болотова Юлия Александровна, к.т.н., доцент инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail: jbolotova@tpu.ru.

Попова Екатерина Сергеевна, аспирант инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail: esp9@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Обзор методов распознавания человека по походке в видео

А.И. Соколова[1], А.С. Конушин[1,2]
ale4kasokolova@gmail.com|anton.konushin@graphics.cs.msu.ru
[1]Национальный Исследовательский Университет Высшая Школа Экономики, Москва, Россия;
[2]Московский Государственный Университет им. Ломоносова, Москва, Россия

Походка - важный биометрический показатель, позволяющий идентифицировать человека без личного контакта. В данной работе проводится обзор современных методов распознавания человека по походке, их анализ и сравнение, а также выявляются проблемы, препятствующие окончательному решению задачи распознавания по походке.

Ключевые слова: *походка, биометрия, силуэт, нейронные сети, идентификация.*

Review of video gait recognition methods

A. Sokolova[1], A. Konushin[1,2]
ale4kasokolova@gmail.com|anton.konushin@grapics.cs.msu.ru
[1]National Research University Higher School of Economics, Moscow, Russia;
[2]Lomonosov Moscow State University, Moscow, Russia

Human gait is an important biometrical index that allows to identify the human without any contact. In this paper, we provide a survey of state-of-the-art methods of gait recogntion, their analysis and comparison. We additionally reveal the problems that prevent the final solution of gait recognition challenge.

Keywords: *gait, biometrics, silhouette, neural networks, identification.*

1. Введение

Задача распознавания человека в видео по походке особенно актуальна в современном мире. Как показывают биометрические исследования, каждый человек обладает своей манерой движения, которую практически невозможно подделать, вследствие чего походка оказывается уникальным идентификатором, подобным сетчатке глаза или отпечаткам пальцев. С развитием систем видеонаблюдения походка становится наиболее удобной характеристикой для распознавания благодаря возможности наблюдать этот показатель без ведома человека. В связи с этим идентификация человека по походке может быть применима, например, в сфере безопасности для распознавания преступников или ограничения доступа на закрытые территории. Задача распознавания человека в видео усложняется множеством факторов, влияющих как на представление походки, так и на само движение: ракурс, разная одежда и переносимые предметы и изменение походки человека со временем.

В статье представлен обзор методов распознавания человека по походке в видео, а также их сравнение на популярных наборах данных.

На сегодняшний день существует два основных подхода к получению признаков походки и их классификации: построение признаков вручную и обучение признаков. Первый способ более традиционен и, как правило, основывается на вычислении различных свойств бинарных масок силуэта человека или на исследовании взаимного расположения суставов, относительных расстояний и скоростей и других кинетических показателей. Обучение признаков характерно для искусственных нейронных сетей, набравших популярность в последние годы благодаря выдающимся результатам в решении многих задач компьютерного зрения, таким как классификация видео и изображений, сегментация изображений, детекция объектов, визуальный трекинг и другие. Признаки, обучаемые с помощью нейронных сетей, часто обладают более высоким уровнем абстракции, необходимым для качественного распознавания. Кроме того, высокое качество идентификации достигается методами, комбинирующими два описанных подхода. На начальном этапе вручную вычисляются базовые характеристики походки, а на их основе обучается нейронная сеть, выделяющая более абстрактные признаки. Несмотря на успешность методов глубинного обучения, на данный момент наилучшего результата на некоторых наборах данных достигают неглубокие алгоритмы, поэтому оба глобальных подхода достойны внимания.

2. Базовые признаки походки

Рассмотрим сначала некоторые базовые подходы, в которых признаки походки извлекаются вручную из естественных соображений.

2.1 Бинарные силуэты человека

Наиболее распространенной характеристикой походки является изображения энергии походки (Gait Energy Image, GEI [5]). Такие изображения − усредненные по одному циклу походки бинарные маски силуэта движущегося человека. Изображения энергии походки характеризуют частоты нахождения человека в той или иной позе во время движения. Этот подход получил широкое распространение и лег в основу множества других методов распознавания походки. Кроме того, многие подходы предлагают похожую агрегацию других базовых признаков. Например, распознавание возможно по изображениям энтропии походки [1], где вместо усреднения силуэтов вычисляется энтропия каждого пикселя, или по дискретному преобразованию Фурье набора силуэтов [11].

Несмотря на простоту всех этих методов, именно изображения энергии используются и развиваются до сих пор. По таким изображениям можно вычислять дальнейшие признаки, такие как гистограммы ориентированных градиентов, или строить более сложные алгоритмы классификации, использующие специфику задачи распознавания походки.

Так, одними из наиболее успешных многоракурсных подходов являются два неглубоких метода, использующих в качестве базовых признаков изображения энергии походки. Первый из них − байесовский подход, предложенный Ли в [9]. Авторы предлагают считать

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

изображения энергии походки случайными матрицами, получающимися из собственно походки и независимого от нее шума, соответствующего различным изменяющимся условиям, причем предполагается, что оба слагаемых − нормальные случайные величины. Рассмотрение совместного распределения двух представлений походки в предположении совпадения классов или их различия сводит проблему к задаче оптимизации ковариационных матриц, решаемую с помощью EM-алгоритма. Во втором подходе [12] предлагается проводить многоракурсный дискриминантный анализ. Для признаков походки, посчитанных для каждого угла обзора, обучается отдельное вложение, так чтобы внутриклассовый разброс был минимален, а межклассовый − максимален.

Оба эти подхода интуитивно понятны и математически просты, что в дополнение к высоким результатам распознавания дает им преимущество перед многими другими методами. Общим недостатком методов, использующих GEI для многоракурсного распознавания, является необходимость вычислять изображение энергии для каждого ракурса, присутствующего в выборке. Поэтому для каждого кадра нужно знать, под каким углом он был снят, что не всегда возможно в реальных данных.

2.2 Поза человека

Помимо бинарных масок силуэта и всевозможных способов их агрегации многие исследователи предлагают обратить внимание на позу человека, а именно положение ключевых точек фигуры (основных частей тела и суставов) в каждом кадре. Так, например, в работе [22] в каждом кадре оценивается скелет человека и исследуется движение ключевых точек: период походки, скорость, траектории бедер, коленей и лодыжек, а также относительные углы между точками тела. Лу [10] в своем подходе предлагает многослойную деформируемую модель, характеризующую форму и динамику частей тела человека: их относительные размеры, позицию и ориентацию. Параметры такой модели, определяющие позу, восстанавливаются по имеющемуся (или оцененному каким-либо методом) силуэту, а по ним, в свою очередь, производится распознавание человека.

2.3 Траектории точек фигуры

Еще один неглубокий подход, показывающий высокое качество распознавания, предложен в [3], где рассматриваются траектории движения точек фигуры человека и по этим траекториям строятся дескрипторы движения Фишера, которые классифицируются методом опорных векторов.

3. Нейросетевые подходы

Несмотря на обилие структурных неглубоких подходов, сверточные нейронные сети (CNN) занимают прочную позицию во всех задачах компьютерного зрения и в том числе распознавании походки. За последние несколько лет было предложено множество нейросетевых методов идентификации по походке, отличающихся как технически (выбором архитектур сетей, функции потерь, способов обучения), так и идейно - методом обработки данных и извлечения первичных признаков, подаваемых на вход сети. В некоторых работах основным источником информации считаются сами кадры видео, и именно они подаются на вход сети. Как и для смежной задачи распознавания действий [8], классификация по отдельным кадрам дает удовлетворительный результат. Однако во многих подходах основное внимание уделяется именно движению фигуры человека, поэтому мы остановимся подробнее на нескольких более сложных моделях.

Одним из способов предобработки данных перед подачей в сеть является извлечение оптического потока − векторного поля видимого движения точек сцены. Преимущество этого подхода состоит в том, что обученная на таких данных модель не обращает внимания на цвет, яркость или контрастность кадров видео. Влияние на распознавание оказывает только движение отдельных точек фигуры, а именно это и составляет походку человека.

В нескольких работах [2, 16], появившихся практически одновременно, предлагается рассматривать блоки карт оптического потока аналогично временной составляющей классической двухпоточной модели [15] для распознавания действий. Для нескольких идущих подряд пар соседних кадров вычисляется оптический поток и строится тензор − блок из нескольких карт потока. Для большей точности из этого блока вырезается часть, во всех кадрах содержащая фигуру человека, и на таких блоках обучается нейронная сеть. На этапе тестирования сеть используется для извлечения признаков, которые потом можно классифицировать любым методом машинного обучения, например, методом опорных векторов (SVM) или методом ближайшего соседа (kNN).

Второй и наиболее популярный источник информации, на основе которого происходит обучение сетей, − бинарные маски силуэтов, о которых уже шла речь при рассмотрении неглубоких методов. В простейшем случае [24] сверточная архитектура обучается по отдельным силуэтам предсказывать человека, которому этот силуэт принадлежит. Как и предыдущих методах, сеть в дальнейшем используется для извлечения признаков, причем переход от дескрипторов отдельных кадров ко всему видео происходит путем выбора максимального отклика по циклу походки. Этот метод наиболее простой из всех глубоких подходов, т.к. при наличии масок силуэтов людей не требует практически никакой дополнительной предобработки.

Еще один метод, использующий сами силуэты, предложен в [18]. Двухэтапный алгоритм сначала определяет угол съемки видео, а потом по изначальным данным и найденному углу предсказывает человека. Для первой задачи случайным образом выбирается набор кадров видео и поступает на вход сети, предобученной для задачи классификации видео. В зависимости от предсказанного угла, данные подаются в новую сеть, свою для каждого ракурса. Чтобы учитывать не только пространственные, но и временные характеристики движения, на вход сети подаются не отдельные маски, а блоки из нескольких силуэтов, идущих подряд. Кроме того, изменчивость между кадрами учитывается благодаря архитектурам сетей в обеих подзадачах: авторы используют трехмерные сети, в которых свертки производятся не только в пространстве, но и во времени.

Трехмерные свертки применяются также в [20], где источником информации служит комбинированное трехканальное изображение: в качестве первого канала берется сам кадр, приведенный к черно-белому виду, а два других − компоненты карт оптического потока. Такая модель, благодаря своей структуре, также учитывает и пространственные характеристики видео, и временные.

Отдельно стоит выделить множество методов, совмещающих “ручное” выделение признаков, и глубинное обучение. В качестве входных данных для сетей часто используют изображения энергии походки, упоминавшиеся выше. Основанные на этой идее модели варьируются от простейших [14], где неглубокая сеть предсказывает человека по получаемым изображениям энергии, вычисленным для различных ракурсов, до более сложных, как, например, [21], где определяется степень похожести пары GEI изображений и исследуются различные методы

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

сравнения нейросетевых признаков походки. В [17, 23] также с помощью двух- и трехпоточных сиамских архитектур определяется, какие изображения походки близки и принадлежат одному человеку, а какие − разным.

Наряду с классическими сверточными сетями прямого распространения для распознавания по походке, как и для других задач анализа видео, используются рекуррентные нейросети. В них также можно подавать всевозможные данные: сами бинарные силуэты, как это делают в [19], или более сложные преобразованные данные, например, информацию о позе человека в каждом кадре [4].

Многие из обсуждаемых подходов оцениваются на одних и тех же данных при одних и тех же условиях, в этом обзоре мы приводим сравнение некоторых описанных методов и выделяем наиболее успешные решения.

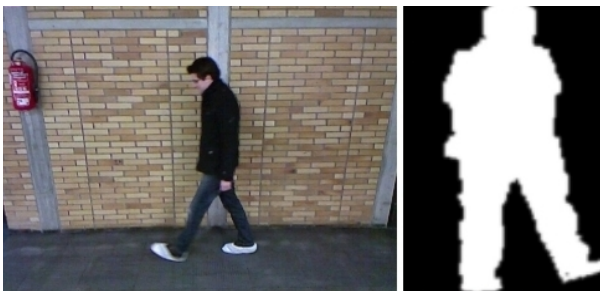


Рис. 1. Изображения кадров из базы TUM-GAID (слева) и из базы OU-ISIR (справа).

4. Базы данных для распознавания человека по походке

Наиболее широко используемыми в наше время сложными наборами данных для распознавания человека по походке являются базы OU-ISIR Large Population Dataset [7] и TUM-GAID [6]. Первый набор применим для многоракурсного распознавания, так как состоит из видеопоследовательностей для более, чем 4000 человек, снятых двумя камерами, причем ракурс съемки плавно меняется от 55° до 85°. Данные из этой коллекции распространяются в виде масок силуэтов, поэтому не все описанные методы применимы к этой базе данных.

База TUM предназначена для распознавания сбоку, все видео в ней сняты под углом 90°, она гораздо меньше предыдущей (по 10 видео для 305 человек), однако состоит из полноценных цветных видео, что делает ее применимой для большого количества подходов. Кроме того, в этой базе присутствуют данные, снятые с разницей в полгода, что дает возможность проверить устойчивость алгоритмов к изменениям походки со временем. Примеры кадров из обеих баз изображены на рис. 1.

Многие из описанных методов оценены на этих наборах данных, поэтому мы приведем сравнение подходов на них.

5. Результаты работы методов

На описанных базах данных оценка качества алгоритмов происходит следующим образом. Модель настраивается на части данных (как правило, это все видео для некоторого подмножества людей), а затем тестируется на новом наборе данных, состоящем из видео для других людей. Для баз TUM и OU-ISIR разделение на обучающую и тестовую выборки предоставлено авторами коллекций. В качестве метрики качества обычно рассматривается точность распознавания.

Метод	Точность
Sokolova [16]	97.5%
Castro, SNN+ SVM [2]	98.0%
Marín. Jiménez [13]	98.9%
Castro, дескрипторы Фишера [3]	99.2%
Zhang [24]	97.7%

Таб. 1. Результаты распознавания на базе TUM-GAID.

В таблице 1 приведено сравнение результатов распознавания на базе TUM-GAID. Наилучшего качества достигает единственный метод [3], не использующий нейронные сети, что показывает, что нейросетевые подходы пока не могут полностью вытеснить неглубокие.

Интересно также рассмотреть, насколько алгоритмы устойчивы ко времени съемки видео. Оказывается, что идентификация человека в случае, когда между несколькими съемками проходит длительный период времени, гораздо сложнее, и качество таких экспериментов оказывается невысоким. Таблица 2 отражает результаты соответствующих экспериментов. Точность каждого из представленных алгоритмов падает примерно на 30%, что говорит о том, что на сегодняшний день алгоритмы плохо переносимы во времени.

Метод	Точность
Castro, CNN + SVM [2]	59.4%
Marín. Jiménez [13]	63.6%
Castro, дескрипторы Фишера [3]	60.4%

Таб. 2. Сравнение результатов распознавания видео, снятых в разные дни, на базе TUM-GAID.

Для многоракурсных баз сравнение, как правило, производится для всевозможных пар ракурсов: данные, снятые под некоторым “тестовым” углом классифицируются алгоритмом, при настройке которого используется другой, “обучающий” угол съемки.

Метод	0°	10°	20°	30°
Zhang [24]	94.1%	71.6%	21.8%	2.9%
Shiraga [14]	94.9%	93.9%	90.5%	80.65%
Li [9]	98.3%	98.2%	97.3%	94.6%
Mansur [12]	96.8%	96.3%	94.2%	90.3%

Таб. 3. Результаты распознавания на базе OU-ISIR.

Для базы OU-ISIR есть два популярных протокола тестирования. Один из них предоставлен авторами: из коллекции выбрано 1912 человек, которые пятью способами разделяются пополам на обучающую и тестовую выборки, после чего качество моделей усредняется. Второй протокол реализует кросс-валидацию, причем модели строятся на данных для 3844 людей, для которых в базе присутствуют данные с обеих камер. Для удобства результаты сравнения агрегируют, рассматривая разности между “обучающим” и “тестовым” углами. В таблице 3 показана средняя точность методов для каждого из 4 возможных значений разности углов. Самый простой метод [24] оказывается несостоятелен, когда углы съемки сильно отличаются, однако остальные подходы показывают достаточно высокое качество распознавания. Наилучших результатов при таких экспериментах тоже достигает метод, не использующий нейронные сети.

Метод	0°	10°	20°	30°
Shiraga [14]	96.5%	95.8%	92.5%	84.9%
Wu [21]	98.9%	95.5%	92.4%	85.3%
Takemura [17]	99.3%	99.2%	98.6%	96.9%

Таб. 4. Результаты кросс-валидации на базе OU-ISIR.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Однако при использовании большего количества данных для обучения и тестирования нейросетевые методы оказываются очень успешны. В таблице 4 приведены результаты сравнения алгоритмов при использовании данных для почти 4 тысяч людей.

Благодаря большему размеру обучающей выборки методы, использующие такой протокол тестирования, достигают более высокой точности. Отсутствие открытых реализаций и общего протокола тестирования не дает возможности сравнить все методы и найти оптимальный, однако имеющиеся результаты показывают, что на сегодняшний день глубокие и неглубокие методы развиваются и показывают практически равное качество.

6. Заключение

Несмотря на все многообразие предлагаемых подходов, задача распознавания походки все еще не потеряла актуальности: существующие решения пока не достигли идеальной точности идентификации. Множество различных условий, влияющих на представление движения, и отсутствие достаточно больших наборов данных, способных учесть все возможные вариации походки, препятствует созданию совершенной модели.

7. Благодарности

Работа выполнена при поддержке гранта #16-29-09612 Российского фонда фундаментальных исследований.

8. Литература

[1] Bashir, K., Xiang, T., S, G. Gait recognition using gait entropy image // Proceedings of 3rd international conference on crime detection and prevention, 2009

[2] Castro, F.M., Marín.Jiménez, M.J., Guil, N., Pérez de la Blanca, N. Automatic learning of gait signatures for people identification // Advances in Computational Intelligence. 2017. pp. 257−270

[3] Castro, F.M., Marín.Jiménez, M., Medina Carnicer, R. Pyramidal Fisher Motion for multiview gait recognition // 22nd International Conference on Pattern Recognition, 2014, pp. 692−1697

[4] Dan.Liu, X.L.F.Z. Mao~Ye, Lin, L. Memory-based gait recognition // Proceedings of the British Machine Vision Conference (BMVC), 2016, pp. 82.1−82.12

[5] Han, J., Bhanu, B. Individual recognition using gait energy image // IEEE Trans Pattern Anal Mach Intell, 2006, 28, pp. 316−322

[6] Hofmann, M., Geiger, J., Bachmann, S., Schuller, B., Rigoll, G.: The TUM Gait from Audio, Image and Depth (GAID) database: Multimodal recognition of subjects and traits // J of Visual Com and Image Repres, 2014, 25(1), pp.195 − 206

[7] Iwama, H., Okumura, M., Makihara, Y., Yagi, Y.: The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition // IEEE Trans on Information Forensics and Security, 2012, 7, Issue 5, pp.1511−1521

[8] Karpathy, A. Toderici, G. Shetty, S. Leung, T. Sukthankar, R and Fei-Fei, L. Large-Scale Video Classification with Convolutional Neural Networks // Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition.

[9] Li, C., Sun, S., Chen, X., Min, X. Cross-view gait recognition using joint Bayesian // Proc. SPIE 10420, Ninth International Conference on Digital Image Processing, 2017

[10] Lu, H. Plataniotis, K. N. Venetsanopoulos, A. N. A Full-Body Layered Deformable Model for Automatic Model-Based Gait Recognition // EURASIP Journal on Advances in Signal Processing, 2007, pp 261−317

[11] Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., Yagi, Y. Gait recognition using a view transformation model in the frequency domain // Computer Vision − ECCV 2006, 2006, pp. 151−163

[12] Mansur, A., Makihara, Y., Muramatsu, D., Yagi, Y. Cross-view gait recognition using view-dependent discriminative analysis // IEEE/IAPR International Joint Conference on Biometrics, 2014

[13] Marín. Jiménez, M., Castro, F., Guil, N., de la Torre, F., Medina.Carnicer, R. 'Deep multi-task learning for gait-based biometrics'.// IEEE International Conference on Image Processing (ICIP), 2017

[14] Shiraga, K., Makihara, Y., Muramatsu, D., Echigo, T., Yagi, Y. GEINet: View-invariant gait recognition using a convolutional neural network // International Conference on Biometrics (ICB)}, 2016

[15] Simonyan, K., Zisserman, A. Two-stream convolutional networks for action recognition in videos // Proceedings of the 27th International Conference on Neural Information Processing Systems. 2014, pp. 568−576

[16] Sokolova, A., Konushin, A. Gait recognition based on convolutional neural networks // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. vol. XLII-2/W4. 2017. pp. 207−212

[17] Takemura, N., Makihara, Y., Muramatsu, D. On input/output architectures for convolutional neural network-based cross-view gait recognition // IEEE Trans Circuits Syst Video Technol 1-1. 2017

[18] Thapar, D., Nigam, A., Aggarwal, D., Agarwal, P. VGR-net: A view invariant gait recognition network // IEEE 4th International Conference on Identity, Security, and Behavior Analysis, 2018, 1-8.

[19] Tong, S., Fu, Y., Ling, H., Zhang, E. Gait identification by joint spatial-temporal feature // Biometric Recognition, 2017, pp. 457−465

[20] Wolf, T., Babaee, M., Rigoll, G. Multi-view gait recognition using 3D convolutional neural networks // Proceedings of the IEEE International Conference on Image Processing, 2016, pp. 4165-4169.

[21] Wu, Z., Huang, Y., Wang, L., Wang, X., Tan, T. A comprehensive study on cross-view gait based human identification with deep cnns // IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39

[22] Yoo, J. and Nixon, M. S. Automated Markerless Analysis of Human Gait Motion for Recognition and Classification // ETRI Journal, 2011, 33: 259−266

[23] Zhang,C. Liu, W., Ma, H., Fu, H. Siamese neural network based gait recognition for human identification // IEEE International Conference on Acoustics, Speech and Signal Processing, 2016

[24] Zhang, X., Sun, S., Li, C., Zhao, X., Hu, Y. Deepgait: A learning deep convolutional representation for gait recognition // Biometric Recognition, 2017

Об авторах

Соколова Анна Ильинична − аспирантка департамента больших данных и информационного поиска факультета компьютерных наук НИУ ВШЭ. Email: ale4kasokolova@gmail.com.

Конушин Антон Сергеевич − к.ф.-м.н., доцент, заведующий лабораторией компьютерной графики и мультимедиа факультета ВМК МГУ имени М.В. Ломоносова, доцент департамента больших данных и информационного поиска факультета компьютерных наук НИУ ВШЭ. Email: anton.konushin@graphics.cs.msu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Нейросетевой метод обнаружения размытых изображений лиц

В.Г. Спицын[1], Ю.В. Савицкий[1], А.Б. Казиев[1], Ю.А. Болотова[1]
spvg@tpu.ru|yvs18@tpu.ru|kaziev_ab@mail.ru|jbolotova@tpu.ru
[1]Национальный исследовательский Томский политехнический университет, Томск, Россия

Предложен метод оценки степени размытия изображения лица на основе аппарата свёрточных нейронных сетей. исследована возможность использования в качестве входных данных для нейронной сети исходного изображения, модульной составляющей частотного спектра и фазовой составляющей частотного спектра исходного изображения. Тестирование метода проводилось на собственной базе изображений лиц, полученных с ip-камеры в реальных условиях. Предложенный метод сравнивается с известным методом оценки размытия на основе количественного анализа модульных составляющих частотного спектра изображения. На собранной тестовой выборке предложенный метод показал точность распознавания 98,57%, метод на основе количественного анализа модульных составляющих частотного спектра показал точность 77,12%.

Ключевые слова: *свёрточные нейронные сети, оценка качества изображения, преобразование Фурье.*

Neural network method for detecting fuzzy images of faces

V.G. Spitsyn[1], Yu.V. Savitsky[1], A.B. Kaziev[1], Yu.A. Bolotova[1]
spvg@tpu.ru|yvs18@tpu.ru|kaziev_ab@mail.ru|jbolotova@tpu.ru
[1]National Research Tomsk Polytechnic University, Tomsk, Russia

A method for estimating the degree of blurring of an image of a person based on the apparatus of convolutional neural networks is proposed. The possibility of using the original image, the modular component of the frequency spectrum and the phase component of the frequency spectrum of the original image as input to the neural network is investigated. The method was tested on an own image base of faces obtained from an ip camera in real conditions. The proposed method is compared with the known method of estimating the blur based on a quantitative analysis of the modular components of the frequency spectrum of the image. On the collected test sample, the proposed method showed the recognition accuracy of 98.57%, the method based on the quantitative analysis of the modular components of the frequency spectrum showed an accuracy of 77.12%.

Keywords: *convolutional neural networks, image quality estimation, Fourier transform.*

1. Введение

В настоящее время повсеместное распространение получают системы биометрической идентификации человека по изображению лица. Условия, в которых приходится работать этим системам, сильно отличаются от условий, в которых набираются изображения для бенчмарков. В данной работе рассматриваются изображения, взятые с камер наблюдения внутри помещения, например, на проходной предприятия.

В условиях получения изображений в реальном времени, входные данные могут подвергаться искажениям. Одним из видов искажений, встречаемых наиболее часто в ходе работы в реальных условиях, является искажение «размытия». К данному виду могут быть отнесены искажения, получающиеся в результате смазывания в движении или расфокусировки камеры. Данный вид искажений приводит к тому, что изображение становится непригодным для распознавания. Отбраковка таких изображений позволит снизить нагрузку на систему распознавания, сократив тем самым вычислительные ресурсы, необходимые для работы биометрической системы.

При оценке качества изображения, в смысле оценки степени его размытия, алгоритмы оценки можно разделить на следующие категории [14, 18, 19]:

1. Эталонные методы – предполагают наличие эталонного изображения, которое изначально не содержит искажений. В таком случае оценка качества изображения может основываться на оценке схожести изображений, в том или ином смысле.
2. Не эталонные методы – предполагают отсутствие эталонного изображения.

Псевдоэталонные методы – предполагают наличие ограниченной дополнительной информации о том, какими характеристиками должно обладать качественное изображение.

В данной работе предложен метод оценки качества изображений на основе использования аппарата свёрточных нейронных сетей (СНС). Предложенный метод может быть отнесён к неэталонным методам оценки качества изображений.

2. Обоснование выбора СНС для оценки качества изображений

Свёрточная нейронная сеть - особый вид архитектуры нейронной сети прямого распространения, специально разработанный для анализа изображений. Основной особенностью данного вида архитектуры нейронной сети является использование рецептивных полей, т.е. фильтров, чувствительных к определённому визуальному образу.

Предшественниками свёрточных нейронных сетей были модели когнитрона и неокогнитрона [5]. Свёрточные нейронные сети в современном виде были представлены в работах Ле Куна [11-13] и A. Krizhevsky, I. Sutskever, G.E. Hinton [10]. Основными слоями, используемыми в свёрточных сетях, являются свёрточный слой, слой субдискретизации и полносвязный слой.

Сочетание функции классификации с функцией выделения признаков с помощью ядер свёртки, получаемых в процессе обучения, позволяет выделять оптимальный набор признаков. Получить данный набор признаков, подбирая метод извлечения признаков вручную, является практически невозможной задачей.

Исходя из вышесказанного, было выдвинуто предположение, что аппарат свёрточных нейронных сетей может показать высокую эффективность в решении задачи оценки степени размытия изображения. Для проверки выдвинутого предположения было принято решение об

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

исследовании различных вариантов архитектур, с целью определения наиболее эффективной для решения поставленной задачи.

Полученную модель необходимо сравнить с подходом, не использующим нейронные сети по характеристике эффективности решения поставленной задачи. Так как данная работа нацелена на использование в конечной системе распознавания лиц, реализуемая модель должна быть достаточно компактной и обеспечивать возможность работы в реальном времени.

В качестве базовых видов архитектур, на основе которых можно построить архитектуру сети, были рассмотрены LeNet [10-13], Inception [1, 16] и ResNet [6] . Данные виды архитектур являются наиболее распространёнными на текущий момент.

На основе анализа работ, использующих сети с указанными архитектурами, и предварительном тестировании сложности вычислений, был сделан вывод, что для данной работы будет целесообразно использовать сеть, построенную на основе архитектуры LeNet. Данный вариант архитектуры используется в качестве базового для подбора основных параметров обучения сети.

Модификации архитектуры LeNet производятся путём:

- изменения количества слоёв свёртки и субдискретизации;
- изменения количества и размера фильтров слоёв свёртки;
- изменения размера и шага областей субдискретизации;
- изменения количества и размера полносвязных слоёв;
- добавления слоёв нормализации;
- изменения нелинейной функции активации нейронов.

3. Формирование обучающей и тестовой выборок

Отбор обучающего материала является одним из наиболее значимых факторов в машинном обучении и в обучении нейронных сетей в частности. Для формирования обучающей выборки были записаны видеоролики проходов людей по направлению к камере. В ходе записи менялись условия освещения, высота установки и угол наклона камеры относительно горизонта. Для выделения изображений лиц на видео была использована реализация детектора лиц на основе метода Виолы-Джонса открытой библиотеки OpenCV. После обработки полученных роликов алгоритмом детекции лиц, были получены 200 000 изображений.

Сортировка полученных изображений была проведена вручную, так как если бы имелось какое-либо средство автоматической сортировки по заданному критерию, то не было бы необходимости в данной работе. Весь процесс сортировки занял 7 дней, в результате изображения были распределены по 4 группам:

1. Изображения с высокой степенью размытия
2. Изображения без размытия
3. Изображения с промежуточной степенью размытия
4. Бракованные изображения (пересвеченные, затемнённые, с высокой степенью зашумления).

В результате сортировки к первой и второй группе были отнесены 130 000 изображений. Итоговое соотношение чётких изображений и размытых изображений оказалось 3 к 1.

В качестве входных данных были рассмотрены три варианта, примеры изображений приведены на рис. 1:

1. Исходное изображение (левая колонка).
2. Изображение модульных составляющих частотного спектра (по центру).
3. Изображение фазовых составляющих частотного спектра (правая колонка).

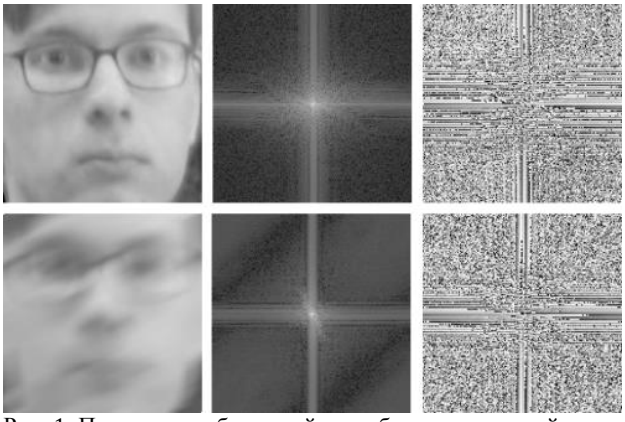


Рис. 1. Примеры изображений для обучения: верхний ряд – чёткое изображение и его спектральные изображения, нижний ряд – размытое изображение и его спектральные изображения.

Для обучения нейронных сетей была выбрана библиотека Dlib [9]. Данная библиотека ориентирована на применение в коммерческой разработке. В отличие от большинства библиотек машинного обучения она ориентирована на использование api на языке C++. Активно использует стандарт C++11. По сравнению с библиотекой Caffe [2] Dlib обеспечивает большую скорость работы, за счёт формирования архитектуры сети с помощью шаблонов.

При обучении сети задача обучения была сформулирована как задача бинарной классификации. В данном случае первым классом являлись не размытые изображения, а вторым классом – размытые изображения.

В работе проведен анализ особенностей модификаций метода градиентного спуска: стохастический градиентный спуск, ускоренный градиентный спуск Нестерова [15], Adagrad [4], RMSProp/AdaDelta [17, 20], Adam [3]. На основе проведенного сопоставления перечисленных модификаций метода градиентного спуска для обучения сети был выбран алгоритм Adam. Данный метод был выбран в первую очередь потому, что при его использовании процесс обучения сходится быстрее и занимает меньше времени, что позволяет провести большое количество запусков обучения конкретного варианта архитектуры и собрать достоверную статистику.

Для регуляризации весов использовался приём batch normalization [7, 13]. Использовался приём уменьшения коэффициента обучения на порядок по достижению сетью плато на поверхности ошибок.

Для поиска оптимальных значений параметров обучения сети были выбраны три основных параметра, оказывающих наиболее значимое влияние на процесс обучения нейронной сети: начальная скорость обучения или коэффициент обучения, коэффициент регуляризации и число итераций без снижения ошибки, после которого происходит снижение коэффициента обучения. Общий алгоритм подбора параметров выглядит следующим образом:

1. Подбор начальной скорости обучения. Данный параметр признаётся как наиболее значимый параметр обучения, поэтому его рекомендуется подбирать в первую очередь.
2. Подбор параметра регуляризации. На предыдущем этапе данный параметр устанавливался равным 0.
3. Подбор предельного числа итераций без прогресса. Увеличение данного параметра увеличивает продолжительность обучения, давая алгоритму обучения больше времени на поиск оптимального решения. Однако слишком большое число итераций без прогресса может привести к переобучению модели.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

При каждом изменении какого-либо параметра обучалось по 10 вариантов весов для базовой архитектуры сети. Каждый полученный вариант сети тестировался на 10 наборах тестовой выборки. Итоговые величины среднего значения и стандартного отклонения точности определялись по результатам всех тестовых проходов.

На основе проведенных экспериментов установлено, что наилучших результатов при тестировании обучаемой модели удаётся добиться при использовании значения начальной скорости обучения равном $0.5*10^{-2}$, для выборки изображений модульных составляющих частотного спектра – $0.5*10^{-2}$ и для выборки фазовых составляющих частотного спектра – $0.5*10^{-3}$.

На следующем этапе было определено оптимальное значение параметра регуляризации. Значение параметра регуляризации подбиралось при использовании значения параметра начальной скорости обучения, определенного в ходе предыдущего этапа, значение параметра числа итераций без прогресса как и на предыдущем этапе равнялось 300. В результате проведенных экспериментов установлено, что наилучших результатов при тестировании обучаемой модели удаётся добиться при использовании значения параметра регуляризации для выборки исходных изображений равном $0{,}5*10^{-5}$, для выборки модульных составляющих частотного спектра – $0{,}5*10^{-4}$, для фазовых составляющих частотного спектра – $0{,}5*10^{-3}$.

На заключительном этапе опренделения параметров обучения был подобран параметр “число итераций без прогресса”. Данный параметр напрямую влияет на продолжительность обучения сети. Проведенные эксперименты показали, что оптимальным значением параметра для всех выборок является значение 1000.

По результатам обучения и тестирования наибольшую точность продемонстрировали сети, обученные на выборке исходных изображений. На втором месте по точности оказались сети, обученные на выборке модульных составляющих частотного спектра. С большим отставанием на третьем месте оказались сети, обученные на фазовых составляющих частотного спектра.

4. Выбор архитектуры сети

Как уже говорилось ранее, одним из требований, предъявляемых к разрабатываемому методу, является его быстродействие. Быстродействие нейронных сетей напрямую зависит от числа параметров, которое использует сеть. В таблице 1 приведено описание предложенных вариантов архитектуры LeNet: BN1-BN4.

В таблице 1 используются следующие обозначения: Input – размер входного изображения, Con – слой свёртки, MaxPool – слой субдискретизации, FC – полносвязный слой, N – число ядер свёртки в слое, K – размер окна свёртки или субдискретизации в пикселях, S – шаг смещения в пикселях, Fun – нелинейная функция, Sigmoid – функция логистической сигмоиды. Входной размер изображения 183x183 обусловлен особенностями работы целевой системы распознавания лиц, в составе которой планируется использовать обученную модель. Полученную модель можно использовать не только как бинарный классификатор размытое/чёткое. На выходе последнего слоя будет число из диапазона (0;1), которое укажет «степень размытия», позволяя конечной системе самой решать, с каким порогом производить фильтрацию входных изображений.

Тестирование предложенных вариантов архитектуры LeNet: BN1-BN4 проходило по тому же принципу, что и во время подбора параметров обучения. На каждый вариант архитектуры было обучено по 10 вариантов весов, каждый из которых был протестирован на 10 тестовых выборках для получения среднего значения точности и его стандартного отклонения.

Таблица 1. Описание вариантов архитектур

	BN1	BN2	BN3	BN4
Input	183x183	183x183	183x183	183x183
Con1	N=16 K=5x5 S=1x1	N=8 K=5x5 S=1x1	N=8 K=5x5 S=1x1	N=16 K=5x5 S=1x1
MaxPool1	K=3x3 S=2x2	K=3x3 S=2x2	K=3x3 S=2x2	K=3x3 S=2x2
Con2	N=32 K=3x3 S=1x1	N=16 K=3x3 S=1x1	N=16 K=3x3 S=1x1	N=32 K=3x3 S=1x1
MaxPool2	K=3x3 S=2x2	K=3x3 S=2x2	K=3x3 S=2x2	K=3x3 S=2x2
Con3	N=64 K=3x3 S=1x1	N=32 K=3x3 S=1x1	N=16 K=3x3 S=1x1	-
MaxPool3	K=3x3 S=2x2	K=3x3 S=2x2	K=3x3 S=2x2	-
Con4	-	-	N=32 K=3x3 S=1x1	-
MaxPool4	-	-	K=3x3 S=2x2	-
FC	160	120	80	160
Fun	Sigmoid	Sigmoid	Sigmoid	Sigmoid
FC	1	1	1	1
Fun	Sigmoid	Sigmoid	Sigmoid	Sigmoid

Результаты тестирования представлены на рис. 2.

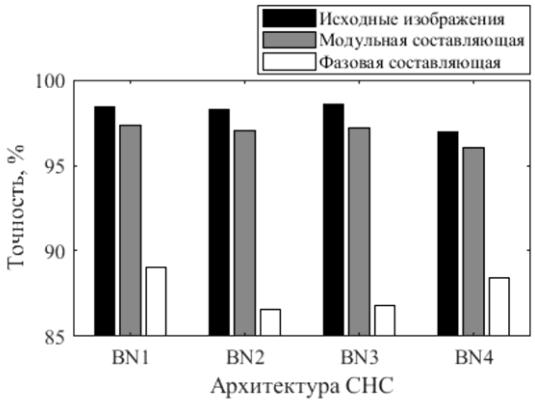


Рис. 2. Результаты тестирования архитектур СНС.

Наиболее эффективным оказался подход, при котором в качестве вектора входных признаков использовалось исходное изображение лица. Это говорит о высокой информативности признаков, содержащихся в данных специфичных изображениях, для СНС с точки зрения оценки размытости и последующей бинарной классификации. Данный подход показывает как наибольшую точность, так и наибольшую стабильность, то есть вероятность высоких результатов обучения при использовании данного вектора признаков и подобранных параметров обучения.

Подход, при котором в качестве входного вектора признаков использовались модульные составляющие гармоник частотного спектра, показал меньшую эффективность относительно первого подхода. Наименее эффективным оказался третий подход, при котором использовались фазовые составляющие гармоник частотного спектра.

В ходе тестирования были произведены замеры времени, которое необходимо для обработки изображения каждой архитектурой. Тестирование проводилось на восьмиядерном CPU AMD FX-8320 3.5 ГГц, размер ОЗУ 8

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Гб. Библиотека Dlib была собрана с поддержкой пакета OpenBLAS из репозитория Ubuntu 14.04 и без поддержки GPU. При использовании данной конфигурации среднее время обработки одного изображения для BN1 составило 15.0 миллисекунды (мс), для BN2 – 9.3 мс, для BN3 – 8.5 мс, для BN4 – 18.4 мс.

Следует отметить, что все 4 архитектуры показали сопоставимый результат на каждой из выборок исходных данных. Наилучший итоговый результат показали сети, обученные на выборке исходных изображений, наихудший – на выборке фазовой составляющей частотного спектра. Исходя из полученных данных следует, что для использования предложенного метода в конечном приложении следует выбрать вариант BN3 архитектуры LeNet, так как данная архитектура является самой компактной.

5. Анализ полученных результатов

На заключительном этапе было проведено сравнение результатов предложенного метода с методом, описанном в работе [8].

Метод из [8] был реализован и протестирован на описанном выше наборе тестовых данных изображений лиц. Полученные результаты показывают точность классификации 77,12 % , что значительно уступает значению 98,57%, приведенному на рис. 2.

Таким образом, предложенный в данной работе метод обнаружения размытия на изображениях лиц на основе сверточных нейронных сетей существенно превосходит метод, основанный на подсчёте модульной составляющей спектра [8].

6. Заключение

Предложен и протестирован нейросетевой метод обнаружения размытых изображений лиц. Результаты сопоставления предложенного метода с известным методом, основанным на подсчёте модульных составляющих спектра изображения, показали точность распознавания размытых изображений лиц превосходящую на 21,45 % результаты тестирования метода [8] на той же базе изображений. Исходя из распределения чётких и размытых изображений, на основе применения разработанного метода, можно сократить количество обрабатываемой информации на 30%, что приведет к существенному снижению нагрузки на систему распознавания лиц.

7. Благодарности

Работа выполнена в рамках Программы повышения конкурентоспособности ТПУ при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А.

8. Литература

[1] Arora S., Bhaskara A., Ge R., Ma T. Provable Bounds for Learning Some Deep Representations // arXiv. 2013. №1310.6343v1.

[2] Caffe deep learning framework. Available at: http://caffe.berkeleyvision.org (accessed 20 October 2017).

[3] Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization // Source: https://arxiv.org/pdf/1412.6980.pdf.

[4] Duchi J., Hazan E., Singer Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. Journal of Machine Learning Research. 2011, Vol. 12, pp. 2121-2159.

[5] Fukushima K. Neocognitron: A hierarchical neural network capable of visual pattern recognition // Neural networks. 1988, Vol. 1(2), pp. 119–130.

[6] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // arXiv. 2015. №1512.03385, http://www.ms.uky.edu/~qye/MA721/JinpengLiu_ResNet.pdf.

[7] Ioffe S., Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift // arXiv. 2015. Source: https://arxiv.org/pdf/1502.03167.pdf.

[8] Kanjar De., Masilamani V. Image Sharpness Measure for Blurred Images in Frequency Domain // Procedia Engineering. 2013, Vol. 64, pp. 149-158.

[9] King, D. E. Dlib-ml: A Machine Learning Toolkit // Journal of Machine Learning Research. 2009, Vol. 10, pp. 1755-1758.

[10] Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems. 2012, pp. 1097–1105.

[11] LeCun Y. Backpropagation applied to handwritten zip code recognition // Neural computation. 1989, Vol. 1(4), pp. 541–551.

[12] LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE. 1998, Vol. 86(11), pp. 2278-2324.

[13] LeCun Y. Efficient backprop // Neural Networks: Tricks of the Trade: Lecture Notes in Computer Science / G. Montavon, G. B. Orr, K.-R. Muller (Eds.) – Springer, 2012, pp. 9-48.

[14] Marzialono P., Dufaux F., Winkler S., Ebrahimi T. A no-reference perceptual blur metric // Proceedings of IEEE international conference on Image Processing. 2002, Vol. 3., pp. 57-60.

[15] Sutskever I., Martens J., Dahl G., Hinton G. On the Importance of Initialization and Momentum in Deep Learning // Proceedings of the 30th International Conference on Machine Learning. 2013, Vol. 28., pp. III-1139–III-1147.

[16] Szegedy C. et al. Going deeper with convolutions // arXiv. 2014. Source: https://arxiv.org/pdf/1409.4842.pdf.

[17] Tieleman T., Hinton G. RMSProp: Divide the gradient by a running average of its recent magnitude // Coursera: Neural Networks for Machine Learning: Technical report. 2012, pp. 26-31.

[18] Tong H., Li M., Zhang H., Zhang C. Blur detection for digital images using wavelet transform // Proceedings of IEEE international conference on Multimedia and Expo. 2004, Vol. 1, pp. 17-20.

[19] Wang Z., Bovik A. C. Modern image quality assessment: Synthesis lectures on image, video & multimedia processing.– San Rafael: Morgan & Claypool Publishers, 2006, 156 p.

[20] Zeiler M. Adadelta: An adaptive learning rate method // arXiv. 2012. Source: https://arxiv.org/pdf/1212.5701.pdf.

Об авторах

Спицын Владимир Григорьевич, д.т.н., профессор отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: spvg@tpu.ru.

Савицкий Юрий Васильевич, аспирант отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: yvs18@tpu.ru.

Казиев Александр Борисович, магистрант отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: kaziev_ab@mail.ru.

Болотова Юлия Александровна, к.т.н., доцент отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: jbolotova@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Синтез изображений дорожных знаков с помощью условных порождающих противоборствующих нейросетей

П.В. Хрушков[1], В.И. Шахуро[1,2], А.С. Конушин[1,2]
pvkhrushkov@edu.hse.ru|vlad.shakhuro@hse.ru|anton.konushin@graphics.msu.ru
[1]НИУ Высшая Школа Экономики, Москва, Россия;
[2]МГУ имени М. В. Ломоносова, Москва, Россия

В работе рассматривается метод генерации синтетических обучающих выборок для задачи классификации дорожных знаков. Метод основан на использовании порождающих конкурирующих нейросетей и метрики Васерштейна. Исследуется метод условной генерации изображений, когда на вход порождающей нейросети подается случайный шум и метка класса изображения, которое нужно сгенерировать. Для обучения такой нейросети предлагается использовать перекрестную энтропию в добавление к метрике Васерштейна. Для стабилизации процесса обучения используются веса для обучающей выборки. Экспериментальная оценка метода показывает, что условная порождающая сеть работает лучше, чем простая генерация дорожных знаков по иконке, однако не дотягивает до метода, в котором для каждого класса обучается отдельная порождающая нейросеть.

Ключевые слова: классификация дорожных знаков, синтетические выборки, условные порождающие противоборствующие нейросети

Generation of synthetic traffic sign images using conditional generative adversarial networks

P.V. Khrushkov[1], V.I. Shakhuro[1,2], A.S. Konushin[1,2]
pvkhrushkov@edu.hse.ru|vlad.shakhuro@hse.ru|anton.konushin@graphics.msu.ru
[1]NRU Higher School of Economics, Moscow, Russia;
[2]Lomonosov Moscow State University, Moscow, Russia

In this work we research method for synthesis of training samples for classification of traffic signs. Method is based on generative adversarial networks and Wasserstein metric. We consider conditional generative neural network, which takes random noise and class label as input and outputs image of object of needed class. To train such generative network, we use cross entropy in addition to Wasserstein metric. To stabilize training, we use weighting of samples. Experimental evaluation of method shows that conditional neural network outputs samples which are better than naive samples generated using icons, but worse than set of networks trained separately for each traffic sign class.

Keywords: traffic sign classification, synthetic data, conditional generative adversarial networks.

1. Введение

В настоящее время прогресс в компьютерном зрении в значительной степени обусловлен появлением больших наборов размеченных данных. В качестве примера таких наборов данных можно привести [3, 5, 11, 12] для задач классификации, сегментации и детектирования объектов на изображениях. В зависимости от сложности разметки такие наборы данных насчитывают от 5000 (Cityscapes, семантическая сегментация) до 9 миллионов (OpenImages, детектирование объектов) и 14 миллионов (ImageNet, классификация) изображений. Разметка таких выборок — дорогой и трудозатратный процесс, который можно провести силами большого количества наемных либо добровольных рабочих. При этом даже после составления размеченной выборки изображений могут возникнуть проблемы:

1. Существуют редкие классы объектов. Даже несмотря на большие размеры выборок, обучающих примеров изображений таких объектов недостаточно для обучения современных методов машинного обучения (как правило, нейросетевых).
2. При появлении новых классов объектов требуется обновление выборки.

Описываемую проблему получения качественных обучающих выборок можно решить с помощью генерации синтетических изображений. В последнее время стали активно исследоваться модели порождающих конкурирующих нейросетей для генерации реалистичных изображений. В этих моделях используются две нейросети: генератор и критик (используется подход и терминология из [1]). Генератор принимает на вход случайный шум (например, нормальный) и выдает изображение, а критик выдает меру реалистичности поданного ему на вход изображения. В данной работе исследуется генератор, который помимо случайного шума на вход принимают также метку класса изображения, которое нужно сгенерировать. Чтобы стабильно обучать такой генератор, в функцию потерь критика добавляется дополнительное слагаемое. Помимо меры реалистичности изображения он должен выдавать вектор вероятностей меток класса. Кроме этого, чтобы бороться с несбалансированностью обучающей выборки, задаются веса для обучающих приме-

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

ров. Два перечисленных метода стабилизации обучения условной порождающей нейросети являются основным вкладом данной работы. Качество работы метода генерации синтетических изображений оценивается на задаче классификации дорожных знаков.

2. Обзор методов генерации синтетических изображений

Синтетические обучающие выборки применяются для обучения алгоритмов распознавания изображений в случаях, когда получить и разметить реальные данные невозможно или требуется слишком много ресурсов. Один из примеров такой ситуации — анализ медицинских изображений. Данные могут быть труднодоступны (снимки пациентов с редким заболеванием) и трудоемки в разметке (для сегментации снимков, например, может понадобиться большое количество времени высококвалифицированного специалиста). Генерация реалистичных изображений — плохо определенная задача, т.к. в обработке изображений и компьютерном зрении на сегодняшний день не существует метрики фотореалистичности изображения, а в каждой задаче обучающая выборка имеет особенности. Кроме этого, пока не придумано надежного способа сравнения между собой различных генераторов синтетических изображений [18]. Поэтому для генерации обучающих примеров сейчас используются два метода: размножение реальных данных и трехмерное моделирование.

Размножение изображений активно используется для обучения нейронных сетей. В [13] для обучения классификатора изображений обучающая выборка увеличивается на порядки за счет случайной обрезки изображения и зеркальных отражений относительно вертикальной оси изображения. В [2] изображения дорожных знаков размножаются за счет случайных поворотов, сдвигов и масштабирований.

Трехмерное моделирование для создания обучающих выборок успешно используется в задачах, где не требуется высокая степень фотореалистичности. Например, в [16] обучается регрессор позы человека по карте глубины. С помощью трехмерного моделирования генерируются зашумленные карты глубины, которые, в отличие от реалистичных RGB-изображений, сгенерировать достаточно просто. Для задачи вычисления оптического потока с помощью сверточных нейронных сетей [6] используется нереалистичная выборка «Летающие стулья». Для вычисления оптического потока нейросеть должна научиться сопоставлять области двух кадров, поэтому правдоподобность данных не требуется.

В [8, 15] игровые движки используются для генерации размеченных городских сцен. Экспериментальная оценка показывает, что только синтетических данных недостаточно для качественного обучения алгоритма детектирования объектов и сегментирования изображений. Однако использование синтетических данных вместе с реальными позволяет улучшить качество итогового алгоритма.

В работах [4, 14] рассматривается простой метод генерации изображений дорожных знаков. Знак — стандартизированный объект, поэтому для него в качестве модели можно взять иконку. Затем с помощью случайных преобразований (гауссово размытие, размытие движения, поворот, сдвиг, масштабирование, наложение на фон, изменение контраста и цветности) из иконки получается изображение дорожного знака. Метод требует априорного задания преобразований, применяемых к модели.

В последнее время появился и активно развивается метод генерации фотореалистичных изображений на основе порождающих конкурирующих нейронных сетей [9]. В этом подходе попеременно обучаются две нейросети: генератор и дискриминатор (критик). Генератор преобразует входной случайный шум в реалистичное изображение, а дискриминатор пытается отличить реальное изображение от сгенерированного. Генератор и дискриминатор обучаются попеременно. В качстве функции потерь используется бинарная перекрестная энтропия.

Конкурирующие порождающие нейронные сети также используются для генерации обучающих выборок. В [19] синтетические данные в добавление к реальным данным используются для повышения качества повторной идентификации людей в видео. В [7] синтетические данные добавляются в обучающую выборку для улучшения качества классификации поражений печени. В [20] порождающие нейронные сети используются для генерации изображений дорожных знаков, при этом для каждого класса обучалась собственная нейросеть-генератор. В настоящей работе ставится следующая цель: обучить одну условную нейросеть-генератор дорожных знаков, т.е. нейросеть, принимающую на вход метку класса, и синтезирующую изображение этого класса.

3. Описание метода условной генерации дорожных знаков

В данной работе используется метод обучения порождающих нейросетей с помощью метрики Васерштейна. В процессе обучения участвуют две нейросети: генератор g_θ и критик f_w. Сеть-генератор получает на вход нормальный шум и метку класса и генерирует изображение. Сеть-критик получает на вход изображение и пытается отличить реальное изображение от сгенерированного. Для обучения генератора и критика используется метрика Васерштейна:

$$W(\mathsf{P}_r, \mathsf{P}_g) = \max_{w \in \mathcal{W}} \mathsf{E}_{x \sim \mathsf{P}_r}[f_w(x)] - \mathsf{E}_{z \sim p(z)}[f_w(g_\theta(z))].$$

Здесь $z \sim p(z)$ — случайный шум, подаваемый на вход нейросети-генератору, P_r и P_g — распределения реальных и синтетических изображений соответственно.

Для того, чтобы метрику можно было использовать для обучения нейросетей, нужно добиться того, чтобы функция f_w была липшицевой с константой 1 (подроб-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

нее об этом условии см. [1]). В работе [10] это достигается с помощью использования дополнительного слагаемого в функции потерь:

$$L_R = -\lambda(||\nabla_{\hat{x}} f_w(\hat{x})|| - 1)^2.$$

Здесь $\hat{x} = tx + (1-t)g_\theta(z)$, $z \sim p(z)$, $t \sim U[0;1]$ — выпуклая комбинация реального и синтезированного изображения.

Для того, чтобы генератор учитывал метку класса, критик дополнительно учится классифицировать сгенерированные изображения на заданное количество классов. Для этого используется перекрестная энтропия:

$$L_C = \mathsf{E}[\log P(C = c|I_{real})] + \mathsf{E}[\log P(C = c|I_{fake})]$$

Здесь c — метка класса реального или синтезированного изображения.

Итоговая функция потерь L — сумма $W(\mathsf{P}_r, \mathsf{P}_g)$, L_C и L_R с некоторыми весами.

4. Обучение на несбалансированной выборке

Напомним, что в функцию потерь входит слагаемое L_R, в котором используется выпуклая комбинация $\hat{x}$ синтетического и реального изображений. Заметим, что не имеет смысла находит выпуклую комбинацию изображений двух различных классов. Процесс обучения устроен следующим образом:

1. Сэмплируется мини-батч реальных изображений.
2. Генерируем синтетические изображения с теми же метками, что и изображения в мини-батче. Таким образом, для каждого реального изображения имеем синтетическое с тем же классом.

Отметим также, что на шаге обновления генератора нужно сэмплировать метки классов для генерации изображений. Важно здесь то, что необходимо сэмплировать метку класса с той вероятностью, с которой она встречается в обучающей выборке. Это видно из формулы полной вероятности:

$$\mathsf{P}_r = p(y_1)\mathsf{P}_r(x|y_1) + \cdots + p(y_k)\mathsf{P}_r(x|y_k).$$

Даже если генератор умеет хорошо синтезировать объекты по метке (т.е. хорошо оценил условные распределения вида $\mathsf{P}_r(x|y)$), но метки сэмплируются не с вероятностями $p(y_1), \ldots, p(y_k)$, то получается совершенно другое распределение, отличное от P_r. Это, в свою очередь, означает, что функция потерь штрафует генератор за «неправильно» выученное распределение.

Предположим теперь, что какая-то метка y_i встречается намного чаще другой метки y_j. Это означает, что вероятностная мера, соответствующая y_i, войдёт в P_r с большим весом, нежели вероятностная мера, соответствующая y_i. Из этого следует, что генератору выгоднее генерировать объекты с меткой y_i лучше, чем объекты с меткой y_j. Теоретически WGAN способен идеально оценить реальное распределение, однако наша генерирующая модель ограничена в своих возможностях и не может сохранить в себе слишком много информации. Поэтому если WGAN и будет стоять перед выбором: качеством генерации какого класса пожертвовать, этим классом будет скорее y_j, а не y_i.

Чтобы достичь равного качества генерации изображений разных классов, добьемся равномерности распределения классов. Естественно, можно было бы применить undersampling или oversampling и непосредственно сравнять число объектов каждого класса, однако можно поступить по-другому. Как уже было сказано, будем учить нейросеть оценивать не исходное распределение P_r, а его «равномерную версию»

$$\mathsf{P}'_r = \frac{\mathsf{P}_r(x|y_1) + \cdots + \mathsf{P}_r(x|y_k)}{k}.$$

Заметим, что

$$\mathsf{E}_{x\sim\mathsf{P}'_r}[f(x)] = \int_{\mathcal{X}} f(x)\mathsf{P}'_r(dx) = \int_{\mathcal{X}} \frac{f(x)}{k} \sum_{i=1}^{k} \mathsf{P}_r(dx|y_i)$$

$$= \int_{\mathcal{X}} \frac{f(x)}{k} \sum_{i=1}^{k} \frac{p(y_i)}{p(y_i)} \mathsf{P}_r(dx|y_i)$$

$$= \int_{\mathcal{X}} \sum_{i=1}^{k} \frac{f(x)}{kp(y_i)} p(y_i)\mathsf{P}_r(dx|y_i)$$

$$= \mathsf{E}_{x,y\sim\mathsf{P}_r}\left[\frac{f(x)}{kp(y)}\right],$$

что соответствует введению веса $\frac{1}{kp(y)}$ для объектов класса y. Таким образом, можно пересчитывать все математические ожидания в задаче оптимизации WGAN и заставить модель оценивать «равномерную версию» реального распределения.

5. Экспериментальная оценка

Для экспериментальной оценки предложенного метода использовалась база автодорожных знаков gtsrb [17]. Эта выборка содержит 52 тысячи изображений 43 классов дорожных знаков.

На рис. 1 показаны примеры изображений, генерируемых с помощью условной порождающей нейросети, обученной с помощью метрики Васерштейна и вспомогательного классификатора. Многие изображения получаются визуально неотличимыми от реальных, однако некоторые классы путаются друг с другом (например, ограничение скорости 20 и класс «осторожно, дети», ограничения скорости 100 и 120).

Для получения количественных оценок был обучен нейросетевой классификатор дорожных знаков. Архитектура нейросети взята из [2] и описана в табл. 1.

В табл. 2 показаны результаты классификатора (точность многоклассовой классификации в процентах), обученного на различных тренировочных выборках. Метод условной генерации дорожных знаков, предложенный в данной работе (строчки «условная WGAN» в табл. 2) позволяет генерировать более качественные изображения, чем метод генерации по иконке дорожного знака, однако проигрывает по качеству 43 нейросетям, отдельно обученным на разных классах (строчки «WGAN синтетика» в табл. 2).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

#	Тип	Кол-во карт и нейронов в слое	Ядро
0	Input	3 карты по 48×48 нейронов	
1	Conv	100 карт по 100×100 нейронов	7×7
2	Maxpool	100 карт по 21×21 нейронов	2×2
3	Conv	150 карт по 18×18 нейронов	4×4
4	Maxpool	150 карт по 9×9 нейронов	2×2
5	Conv	250 карт по 6×6 нейронов	4×4
6	Maxpool	250 карт по 3×3 нейронов	2×2
7	FC	300 нейронов	1×1
8	FC	43 нейрона	1×1

Таблица 1. Архитектура свёрточной нейросети, использовавшейся для классификации знаков.

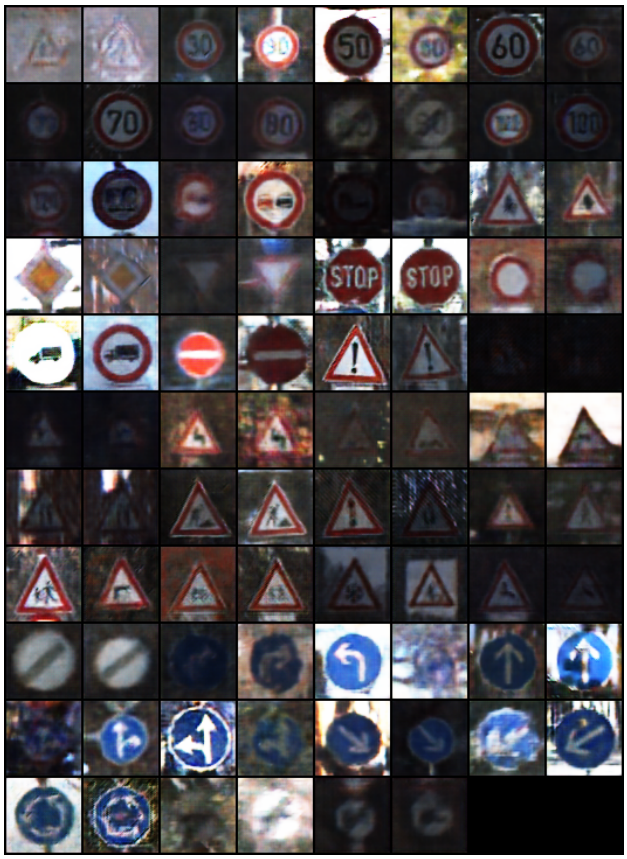


Рис. 1. Примеры изображений, сгенерированных с помощью условного Wasserstein GAN.

6. Заключение

В данной работе предложен метод обучения условной нейросети с помощью метрики Васерштейна. Обученная нейросеть получает на вход случайный шум и метку класса, а на выход выдает изображение заданного класса. Экспериментальная оценка метода на задаче классификации показала, что условная нейросеть работает лучше наивной генерации синтетических знаков по иконке, однако проигрывает по качеству поклассово обученным порождающим нейросетям.

7. Благодарность

Работа выполнена при поддержке гранта РНФ 17-71-20072 «Нейробайесовские методы в задачах машинного обучения, масштабируемой оптимизации и компьютерного зрения».

8. Литература

[1] Arjovsky M., Chintala S., Bottou L. Wasserstein gan //arXiv preprint arXiv:1701.07875. – 2017.

[2] Cireşan D. et al. Multi-column deep neural network for traffic sign classification //Neural networks. – 2012. – T. 32. – C. 333-338.

[3] Cordts M. et al. The cityscapes dataset for semantic urban scene understanding //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – C. 3213-3223.

[4] Chigorin A., Konushin A. A system for large-scale automatic traffic sign recognition and mapping //CMRT13–City Models, Roads and Traffic. – 2013. – T. 2013. – C. 13-17.

[5] Deng J. et al. Imagenet: A large-scale hierarchical image database //Computer Vision and Pattern Recognition, 2009. - C. 248-255.

[6] Dosovitskiy A. et al. Flownet: Learning optical flow with convolutional networks //Proceedings of the IEEE International Conference on Computer Vision. – 2015. – C. 2758-2766.

[7] Frid-Adar M. et al. Synthetic data augmentation using GAN for improved liver lesion classification //Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on. – IEEE, 2018. – C. 289-293.

[8] Gaidon A. et al. Virtual worlds as proxy for multi-object tracking analysis //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – C. 4340-4349.

[9] Goodfellow I. et al. Generative adversarial nets //Advances in neural information processing systems. – 2014. – C. 2672-2680.

[10] Gulrajani I. et al. Improved training of wasserstein gans //Advances in Neural Information Processing Systems. – 2017. – C. 5767-5777.

[11] Krasin I. et al. OpenImages: A public dataset for large-scale multi-label and multi-class image classification, 2017.

[12] Lin T. Y. et al. Microsoft coco: Common objects in context //European conference on computer vision. – Springer, Cham, 2014. – C. 740-755.

[13] Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks //Advances in neural information processing systems. – 2012. – C. 1097-1105.

[14] Moiseev B. et al. Evaluation of traffic sign recognition methods trained on synthetically generated data //International Conference on Advanced Concepts for Intelligent Vision Systems. – Springer, Cham, 2013. – C. 576-583.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Тренировочная выборка	39 тыс. без размножения	215 тыс. без размножения	39 тыс. / 215 тыс. с размножением
Реальные данные	96.6	—	98.4 / —
WGAN синтетика	95.3	96.1	97.6 / 98.1
Реальные данные + WGAN синтетика	—	97.7	— / 98.4
Условная WGAN	79.2	83.7	81.3 / 81.5
Реальные данные + условная WGAN	—	95.2	— / 95.5
Синтетика по иконке	46.5	53.7	67.8 / 69.7
Реальные данные + синтетика по иконке	—	96.5	— / 97.9

Таблица 2. Результаты тестирования классификатора на различных выборках знаках

[15] Richter S. R. et al. Playing for data: Ground truth from computer games //European Conference on Computer Vision. – Springer, Cham, 2016. – С. 102-118.

[16] Shotton J. et al. Real-time human pose recognition in parts from single depth images //Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. – Ieee, 2011. – С. 1297-1304.

[17] Stallkamp J. et al. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition //Neural networks. – 2012. – Т. 32. – С. 323-332.

[18] Theis L., Oord A., Bethge M. A note on the evaluation of generative models //arXiv preprint arXiv:1511.01844. – 2015.

[19] Zheng Z., Zheng L., Yang Y. Unlabeled samples generated by gan improve the person re-identification baseline in vitro //arXiv preprint arXiv:1701.07717. – 2017. – Т. 3.

[20] Шахуро В. И., Конушин А. С. Синтез обучающих выборок для классификации дорожных знаков с помощью нейросетей //Компьютерная оптика. – 2018. – Т. 42. – №. 1.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Классификация изображений таблиц на основе структурного подхода

Т.Н. Шолохова[1], Л.М. Местецкий[1]
tanja200596@mail.ru|mestlm@mail.ru
[1]МГУ им. М.В. Ломоносова, Москва, Россия

В работе предложен новый алгоритм преобразования документа с таблицами в электронный вид. Целью алгоритма является поиск таблиц на изображении документа, их классификация и определение их содержимого на основе обучающего набора таблиц.

Алгоритм основан на непрерывно-морфологическом подходе. Непрерывно-морфологический подход позволяет уйти от необходимости оперировать растровыми терминами и перейти к векторному описанию - линиями, формами и фигурами. В основе алгоритма лежит метод скелетизации.

Ключевые слова: *изображение таблицы, распознавание структуры, классификация, непрерывно-морфологический подход.*

Table images classification based on structure approach

tanja200596@mail.ru|mestlm@mail.ru
[1]Lomonosov Moscow State University, Moscow, Russia

In this work, we propose an algorithm of digital representation of printed table documents. The key steps of the algorithm are detection of tables on an image of a document and matching table structure to one of the predefined table templates.

The algorithm is based on the structure approach. This approach allows describing an image of a table in terms of its shape, morphological structure, and geometrical properties rather than in terms of rasterized image. The core routine of the proposed algorithm is skeleton method.

Keywords: *table image, structure recognition, classification, skeleton approach.*

1. Введение

Одной из важных функций систем электронного документооборота является перевод документов на бумажных носителях в электронную форму. Современные средства - фотокамеры, сканеры - обеспечивают ввод изображений документов в компьютер. Далее необходимо преобразовать изображение в форму электронного документа. Важной частью этого процесса является преобразование таблиц, так как таблицы наиболее часто используются для представления структурированных данных. Целью данной работы является получение эффективного решения описанной задачи.

Для задачи распознавания таблиц необходимо сгенерировать признаковое описание таблицы. Распознавание структуры таблиц на основе поиска горизонтальных и вертикальных линий описано в работах [3, 5]. В статье [6] предложен метод поиска горизонтальных и вертикальных линий на основе гистограмм проекций насыщенности пикселей, обобщение этого метода (с использованием алгоритма Хафа для обнаружения произвольных кривых на изображениях) описано в статьях [4, 7]. В рассмотренных статьях решается задача распознавания качественных отсканированных документов, поэтому получение качественного изображения для исходной, возможно искаженной, фотографии является отдельной важной частью этой работы.

Предложенный подход основан на анализе изображений таблиц с использованием непрерывных скелетов - внутреннего и внешнего [1, 2]. Внутренний скелет представляет собой множеств срединных осей бинарного изображения, в котором чёрные пиксели соответствуют линиям и текстам, а белые - бумажному фону. Внешний скелет - наоборот, есть множество срединных осей фоновой части изображения. На рис. 1 приведены примеры внутреннего (зелёные линии) и внешнего (красные линии) скелетов.

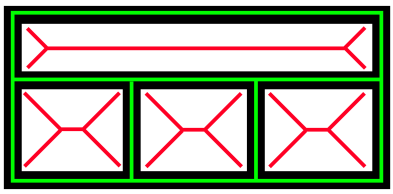
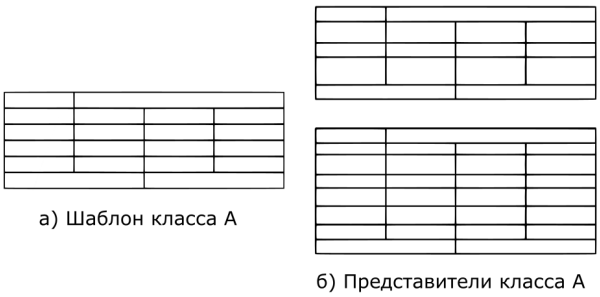


Рис. 1. Иллюстрация скелета

Основные идеи предложенного подхода:

1. Нахождение протяжённых вертикальных и горизонтальных отрезков на основе внутреннего скелета (замена алгоритма Хафа).
2. Отрисовка найденных отрезков на “чистой” странице в растровом формате (удаление лишних элементов - букв).
3. Нахождение ячеек на полученном изображении на основе внешнего скелета (компоненты связности и длинные срединные оси - ячейки).
4. Построение графа смежности ячеек. Граф смежности с информацией о вершинах ячейки - основа для оценки сходства структур таблиц.
5. Найденная структура используется для нахождения ячеек исходной таблицы. Найденные ячейки могут быть отправлены на распознавание в OCR.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18



а) Шаблон класса А

б) Представители класса А

Рис. 2

2. Постановка задачи

Процесс идентификации таблиц состоит в сравнении исследуемого изображения таблицы с описаниями, заданными базой эталонов. Эталоны описываются заранее в процессе разметки.

Размеченные данные представляют собой качественные изображения, каждое изображение содержит ровно одну таблицу, в дополнение к изображению хранится файл разметки этой таблицы, который содержит информацию о местоположениях и типах ячеек.

Внутри одного класса таблиц может изменятся: наполнение таблицы, количество строк в таблице, высота строк в таблице. Не может изменяться: количество столбцов, ширина столбцов.

Размеченные данные будем называть шаблонными данными или шаблонными таблицами. На рис. 2 изображен пример шаблонной таблицы (слева) и её возможных вариаций (справа).

В процессе тестирования необходимо по данному изображению документа (который может содержать несколько таблиц) определить положения всех таблиц, классы найденных таблиц, координаты и типы ячеек.

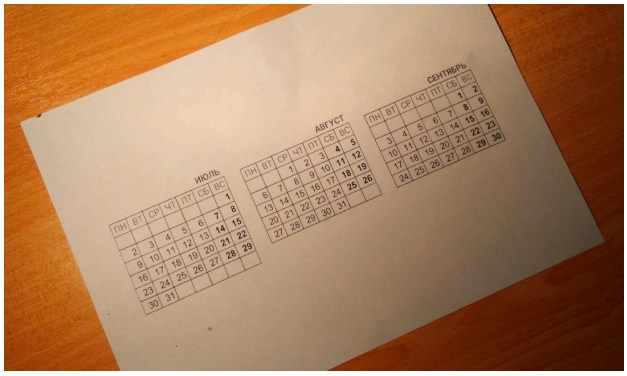


Рис. 3. Пример исходного изображения

На рис. 4 сверху изображена шаблонная таблица, ниже изображен бинаризованный и верно-ориентированный документ из рис. 3, содержащий три таблицы того же типа; на документе цветом отмечены найденные таблицы и ячейки, каждая ячейка соответствует некоторой ячейке шаблонного документа.

ЯНВАРЬ

ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

а) Шаблон

ИЮЛЬ

ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

АВГУСТ

ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

СЕНТЯБРЬ

ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

б) Результат работы

Рис. 4

3. Предобработка входного изображения

Первым шагом необходимо качественно бинаризовать исходную фотографию. Она может быть зашумлена и содержать искажения.

Для бинаризации из исходной фотографии вычитается фон, далее используется адаптивная бинаризация. Для получения фона изображения используется медианное сглаживание с большим радиусом окна (равным четверти высоты изображения).

Для получения верной ориентации используется следующий метод: на бинаризованном изображении строится внутренний скелет, каждое ребро которого имеет некоторую длину и некоторый наклон. На множестве полученных рёбер строится гистограмма углов наклона, каждое ребро имеет вес, равный его длине. На исходных изображениях преобладание по длине рёбер у линий-границ таблиц и максимум гистограммы соответствует преобладающему углу, следовательно при повороте изображения на преобладающий угол линии-границы таблиц станут параллельны осям координат.

На рис. 5 представлен пример предобработки документа, изображенного на рис. 3.

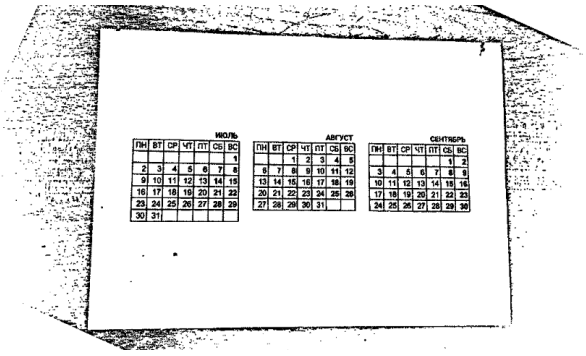


Рис. 5. Результат предобработки

4. Поиск линий и ячеек

Для поиска ячеек необходимо найти на изображении прямые линии, далее на основе линий найти связные ячейки таблицы.

Для поиска линий используется внутренний скелет изображения. Рёбра внутреннего скелета разделяются

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

на горизонтальные и вертикальные. Далее рёбра объединяются в линии в соответствии со следующим алгоритмом:

```
СОРТИРОВКА горизонтальных рёбер
           по координате левого конца
горизонтальные линии = {}
ЦИКЛ по горизонтальным рёбрам:
ЕСЛИ ребро по высоте соответствует одной из линий,
ТО добавить ребро к соответствующей линии
ИНАЧЕ создать новую линию
```

Аналогичный алгоритм используется для объединения вертикальных линий.

Для поиска ячеек на изображении остаются только линии, после этого выполняется построение внешнего скелета.

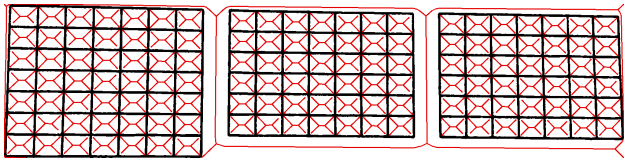


Рис. 6. Внешний скелет

Пример построенного внешнего скелета представлен на рис. 6. Все компоненты связности внешнего скелета, считаются найденными ячейками.

5. Описание структуры таблицы

После определения местоположений всех ячеек таблицы, необходимо классифицировать данную таблицу, определить к какой именно шаблонной таблице она относится. Для этого необходимо определить структуру таблицы.

Структурой таблицы, считается граф, вершинами которого являются ячейки, а рёбра соответствуют смежности ячеек. При этом определяются два типа рёбер: горизонтальная связь и вертикальная связь.

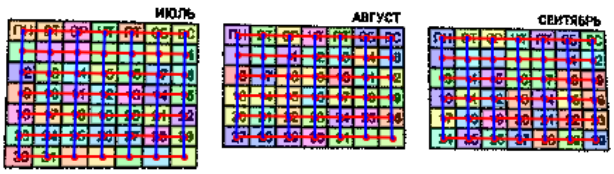


Рис. 7. Структура документа

На рис. 7 представлена полученная структура. Вершины, отмеченные красными точками, расположены в центрах найденных ячеек, горизонтальные рёбра изображены красным цветом, вертикальные — синим.

Дополнительными структурными признаками являются количество столбцов и их ширина. Для определения этих признаков все горизонтальные компоненты графа (все строки таблицы) преобразуются в численные вектора: $col_i = (w_1, w_2, \dots,)$, в которых w_k — соответствует длине очередного ребра (ширине очередного столбца). Далее, из всех векторов $col_1, col_2, \dots,$ находится вектор, который встречается в наборе наибольшее количество раз, назовём его повторяющимся вектором либо повторяющейся строкой. Повторяющийся вектор сохраняется и включается в описание таблицы. Количество столбцов определяется длиной такого вектора.

6. Классификация

Для определения к какому шаблонному типу принадлежит каждая таблица в документе для всех шаблонных таблиц подсчитываются структуры. Так как размерности признаковых пространств для разных таблиц могут быть разными (в качестве признака хранится ширина столбцов, разное количество столбцов означает разную размерность вектора признаков), используемый алгоритм классификации — метод ближайших соседей со специально введённой метрикой для таблиц.

Так, после вычисления структуры таблица T описывается следующим набором:

$$\begin{aligned} describe(T) = [& \\ comp_h(T) &= \{K_h^1, K_h^2, \dots, K^{|T_h|}\}, \\ comp_v(T) &= \{K_v^1, K_v^2, \dots, K^{|T_v|}\}, \\ cols(T) &= [w_1, w_2, \dots, w_{|T_{cols}|}] \quad] \end{aligned}$$

где K – компонента, которая описывается парой чисел: количество вершин и количество рёбер в ней;

$comp_h$ – множество горизонтальных компонент,

$comp_v$ – множество вертикальных компонент,

$cols$ – список размеров столбцов, упорядоченный в порядке столбцов и нормированный на ширину всей таблицы.

В метрику включены величины, описывающие близость таблиц. Например, мощность пересечения множеств горизонтальных компонент одной таблицы с множеством горизонтальных компонент другой таблицы:

$$\frac{|comp_h(T_1) \cap comp_h(T_2)|}{\min(|comp_h(T_1)|, |comp_h(T_2)|)} \cdot \gamma_1$$

То же для вертикальных компонент:

$$\frac{|comp_v(T_1) \cap comp_v(T_2)|}{\min(|comp_v(T_1)|, |comp_v(T_2)|)} \cdot \gamma_2$$

При равенстве количеств столбцов, евклидово расстояние между наборами размеров столбцов:

$$\sqrt{(cols(T_1) - cols(T_2))^2} \cdot \gamma_3$$

. Так как ширина столбцов нормирована на ширину всей таблицы, все вектора будут расположены внутри шара радиуса 1. Значит евклидово расстояние между любой парой векторов не будет превышать 2, то есть в случае разного числа столбцов можно использовать константу 2 в качестве большого расстояния между таблицами. γ_i — параметры классификации настраиваются по валидационной выборке.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

7. Определение типов ячеек

При выделении ячеек таблиц на тестовых документах необходимо восстановить положения ячеек тестовых таблиц и попытаться соотнести их с ячейками шаблонных. Могут возникнуть ошибки двух типов: некоторая ячейка не распознана совсем, либо несколько ячеек объединились в одну.

В структуре тестовой таблицы подсчитывается количество строк и высота каждой строки. Считается, что таблицы могут иметь несколько строк-заголовков и несколько строк-окончаний, отличающихся от повторяющейся строки, именно количество повторяющейся строки может изменяться в тестовой таблице по сравнению с шаблонной. (Строками здесь называются горизонтальные компоненты).

Для восстановления положения ячеек предложен следующий подход: шаблонная таблица преобразуется в соответствии с подсчитанными в тестовой таблице параметрами, после этого в явном виде помещается на место тестовой таблицы. После этого все размеченные ячейки шаблонной таблицы переходят в ячейки тестовой таблицы.

8. Эксперимент

Эксперимент проводится на изображениях финансовых документов.

Эксперимент состоит из двух частей: оценка качества классификации таблиц и оценка качества метода определения положений ячеек.

Качество данного классификатора оценивается на выборке таблиц, часть из которых является качественными отсканированными документами, вторая часть является фотографиями с возможным шумом и искажениями. Метрика качества — точность (accuracy).

Количество эталонных размеченных таблиц равно 10. Для каждой шаблонной таблицы есть ≈ 5 качественных изображений с различными вариациями структуры. Итого ≈ 40 качественных изображений. Каждое качественное изображение было распечатано и сфотографировано ≈ 5 раз. Итого ≈ 200 зашумленных изображений.

Полученное качество классификации таблиц $\approx 93\%$, на отсканированных документах качество достигло 98%, на фотографиях 87%. Ошибки связаны с недостаточным качеством предобработки, либо с похожими таблицами в шаблонной выборке и невозможностью различия их исключительно по структуре.

Для оценки качества метода определения положений ячеек используются только те документы, для которых верно предсказаны типы всех таблиц. Ячейка считается распознанной верно, если отношение совместной площади к площади пересечения с найденной ячейкой $> 90\%$ и совпадает тип. Метрика качества — точность (accuracy).

Полученное качество определения ячеек 82%, на отсканированных документах качество достигло 96%, на фотографиях 69%. Ошибки связаны с недостаточным качеством предобработки, так как фотографии могут содержать проективные искажения, либо искажения второго порядка, в таком случае наиболее частая ошибка — это неверное определения числа строк в таблице.

Основной проблемой в задаче осталась проблема правильной предобработки изображений.

9. Заключение

В ходе работы предложены и реализованы следующие методы:

- метод описания структуры таблицы;
- метод классификации таблиц на основе структуры (качество 93%);
- метод определения положений ячеек таблиц (качество 82%);

10. Благодарности

Работа выполнена при поддержке РФФИ грант 17-01-00917.

11. Литература

[1] Местецкий Л.М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. Москва, Физматлит, 2009.

[2] Местецкий Л.М. Скелет многоугольной фигуры — представление плоским прямолинейным графом. Санкт-Петербург, ГРАФИКОН-2010.

[3] Fan K-C, Wang Y-K, Chang M-L. Form document identification using line structure based features. Guilin, Chinese control and decision conference, CCDC'09, 2001.

[4] Liolios N, Fakotakis N, Kokkinakis G. On the generalization of the form identification and skew detection problem. Pattern Recognit, 2002.

[5] Liu J, Jain AK Image-based form document retrieval. Pattern Recognit, 2000.

[6] Mandal S, Chowdhury S, Das A, Chanda B. A hierarchical method for automated identification and segmentation of forms. Guilin, Chinese control and decision conference, CCDC'09, 2005.

[7] Ohtera R, Horiuchi T. (2004) Faxed form identification using histogram of the Hough-space. Guilin, Chinese control and decision conference, CCDC'09, 2004.

Об авторах

Местецкий Леонид Моисеевич, д.т.н., профессор кафедры математический методов прогнозирования факультета вычислительной математики и кибернетики Московского государственного университета. e-mail: mestlm@mail.ru.

Шолохова Татьяна Николаевна, студентка магистратуры кафедры математический методов прогнозирования факультета вычислительной математики и кибернетики Московского государственного университета. e-mail: tanja200596@mail.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Обработка и анализ биомедицинских изображений

Biomedical Image Processing and Analysis

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Image-based algorithm for the catheter recognition based on GLCM

V.V. Danilov[1,2], I.P. Skirnevskiy[1], R.A. Manakov[1], O.M. Gerget[1], F. Melgani[2]
viacheslav.v.danilov@gmail.com|skirnevskiy@tpu.ru|ram290495@gmail.com|gerget@tpu.ru|farid.melgani@unitn.it
[1]Medical Devices Design Laboratory, Tomsk Polytechnic University, Tomsk, Russia
[2]Signal Processing and Recognition Laboratory, University of Trento, Trento, Italy

The focus of this study was to develop an image-based algorithm for the catheter detection. Epicardial full-volume 3D echocardiography datasets were used as an input data of the algorithm. In total, 9 datasets consisted of 15 three-dimensional timeframes were processed. To correctly detect the catheter, the feature-based approach was applied to recognition the catheter within the 3D echocardiography datasets. The algorithm based on gray-level co-occurrence matrix (GLCM) was applied as a feature extraction technique. Once the GLCM was computed, we obtained correlation, contrast, homogeneity and energy features. Then we applied feature thresholds to the catheter detection. These thresholds were obtained using Support Vector Machine (SVM) with the linear kernel function and standardization the predictor data. The average segmentation and recognition accuracies of the algorithm equal 94.16% and 87.2% respectively. The processing time for one 2D slice and one 3D dataset are equal to 9±0.2 milliseconds and 1.96±0.045 seconds, respectively. Though the algorithm is not time-consuming for 2D mode, it is still complicated to apply it for real-time surgery in 3D mode.

Keywords: *catheter recognition, tracking, texture analysis, GLCM, echocardiography.*

1. Introduction

Image processing and visualization tools are an indispensable link during performing microsurgical endovascular procedures with catheters on the heart. Existing visualization and image processing methods limit the range of complex manipulations performed on the beating heart. In its turn, classical visualization techniques (echocardiography, fluoroscopy, and angiography) [3, 8, 14] possess several crucial disadvantages such as insufficient quality of the output data, high level of noise and limited scope (field of view). These drawbacks significantly reduce the range of possible transcatheter procedures on the heart by means of endovascular technology and also hinder the development of this area. Two-dimensional echocardiography (EchoCG) does not allow adequately displaying the spatial position of the medical instrument (catheter) and visualize the area of interest. Additionally, 3D EchoCG displays a limited area with a significant level of digital noise, shadows, and artifacts. Thus, for today, a wide range of cardiosurgical operations are performed only on an open heart using artificial circulation [10, 16]. Such methods of pathology correction have a number of shortcomings associated with severe postoperative complications and long-term rehabilitation, which is an important problem in cardiology. The availability of new methods of visualization, image processing and delivering catheter devices allows performing a number of interventions on the working heart, including stenting of vessels, radiofrequency ablation in cases of cardiac arrhythmias, closing of atrial and interventricular septal defects. Currently, there are a few numbers of research studies that are devoted to the application of recognition and tracking of medical devices.

Nowadays biplane fluoroscopy that uses X-ray image intensifier is one of the most frequently used methods of visualization (up to 60%) [17]. In the studies of Matthias Hoffman, a semi-automatic reconstruction of the catheter is performed based on two projections [6, 7]. The authors minimized the interaction with the user in their algorithm; nevertheless, recognition still requires the assigning of at least one point belonging to the catheter area. The accuracy of the methods is at a high level, but the total processing time takes about 8 seconds, which confirms the impossibility of using this method in real time. Moreover, the X-ray based methods are not applicable for heart imaging, because they cannot provide soft tissue information.

Research study on the skeletonization and visualization of the catheter in the three-dimensional space of the vasculature were carried out by Baert Shirley, Van de Kraats Everine and others [2]. Another research devoted to the detection and tracking of different types of catheters was performed by Ying Liang Ma et al. [9]. In their study, scientists did not reconstruct the catheter but only used the tracking algorithm for three different types of catheters.

In the other work, P. Ambrosini et al. proposed a method, based on a hidden Markov model, for 3D catheter tip tracking with 2D X-ray catheterization sequence and 3D rotational angiography [1]. But in the research intraoperative images are enhanced using contrast agent for visualizing the vasculature.

In paper [18] the authors describe a method based on registration X-rays and echocardiography. The presented algorithm had an average error of less than 1 mm and the speed could reach 1.5 fps. However, the combination of X-rays and ultrasound signals impose restrictions on the structure of input data end equipment. In this regard, the main scientific direction of this study is the research and development of the algorithm for intraoperative imaging and catheter visualization using the 3D echocardiography data.

2. Materials

2.1 Data description

Epicardial full-volume 3D echocardiography datasets were acquired on the porcine heart using an X7-2t transducer on a Philips iE33 machine and PMS5.1 Ultrasound software (Philips Healthcare, Andover MA). During the data acquisition step, 9 datasets were acquired. Each dataset consists of 15-17 timeframes and each of them includes 208 slices of 176*176-pixel size. Example of the data acquiring using Philips iE33 echocardiographer is shown in Fig. 1.

The experimental protocols were approved by the Boston Children's Hospital Institutional Animal Care and Use Committee (IACUC).

The datasets were processed offline on the computer equipped with Intel Core i7-4790K 4.0 GHz CPU and NVIDIA GeForce 960 GT using MATLAB (MathWorks, Natick MA).

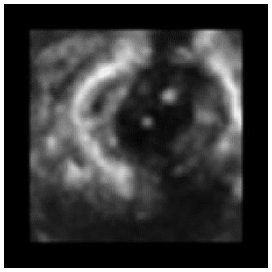


Fig. 1. An exemplary slice of the data.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2.2 Methods

2.2.1 Workflow

To efficiently recognize and segment the catheter we used sequential steps with particular options and setting for each of them. Features selected for the algorithm are the most relevant properties that can be used for the catheter recognition. The main steps of the algorithm are presented as cascade-based workflow shown in Fig. 2.

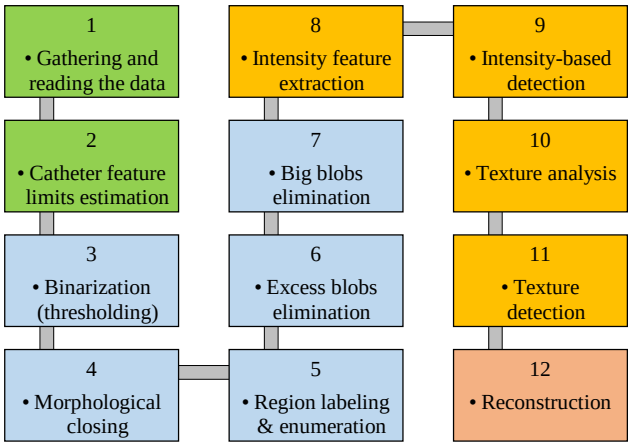


Fig. 2. Basic steps for recognition and segmentation algorithm of the catheter: green blocks – data acquisition and feature studying, blue blocks – coarse detection and segmentation, orange blocks – delicate detection and segmentation, red block – smoothed 3D reconstruction.

To gather the whole data, we used real-time streaming of Philips iE33 machine. After getting all the medical data, we transfer it from DICOM Philips format into Nearly Raw Raster Data (NRRD) format.

2.2.2 Morphology

In regard to the initial processing methods, we applied automatic Otsu's thresholding for binarization an image [12] and morphological closing based on morphological reconstruction for filling the holes [15]. The disk-shaped element was used as the main structuring element for this morphological procedure. Using a disk structuring element allows preserving the circular nature of the object. In our case, the catheter has the circular or elliptical shape depending on its position and the position of the ultrasound transducer.

Small artifacts and noise emissions were removed by the imposing area restrictions. The output of the previous steps was a set of labeled regions including the region of interest (catheter). However, among these regions, there are false ones (see Fig. 4b).

2.2.3 Feature selection

We put forward a hypothesis that, by imposing different feature constraints, the catheter can be accurately found in the image. We selected several high-impact features such as ROI area, mean and standard deviation of intensity within a region and four textural features such as contrast, correlation, energy, and homogeneity. In order to limit the number of obtained regions, we calculated the values of all features and entered the limits for them, which are equal to the ranges, corresponding to the region of the catheter. Textural features were calculated using gray-level co-occurrence matrix (GLCM) [4, 5]. The GLCM for an exemplary slice is shown in Fig. 3.

Fig. 3. GLCM for an exemplary slice.

2.2.4 GLCM

In order to compute textural features, GLCM should be normalized, so that the sum of its elements is equal to 1. Each element *(r,c)* in the normalized GLCM is the joint probability occurrence of pixel pairs with a defined spatial relationship having gray level values *r* and *c* in the image.

Statistical properties of the image derived from GLCM are following:

1. Contrast is a measure of the intensity variance or inertia between a pixel and its neighbor over the whole image. Contrast range for GLCMs (255*255 size) is from 0 to 64516. For the constant image, contrast is equal to 0. Contrast is calculated as follows:

$$Contrast = \sum_{i,j=0}^{N-1} |i-j|^2 p(i,j)$$

2. Correlation is a measure of how correlated a pixel is to its neighbor over the whole image. Correlation varies from 1 to -1. Perfectly positively or negatively correlated image corresponds to 1 or -1 for correlation. Correlation is calculated as follows:

$$Correlation = \sum_{i,j=0}^{N-1} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$$

3. Energy is the sum of squared elements in the GLCM and it ranges from 0 to 1. Energy is calculated as follows:

$$Energy = \sum_{i,j=0}^{N-1} p(i,j)^2$$

4. Homogeneity is a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$Homogeneity = \sum_{i,j=0}^{N-1} \frac{p(i,j)}{1+|i-j|}$$

where:

p(i, j) is an element *(i, j)* of the normalized symmetrical GLCM;

N is the number of gray levels;

μ is the GLCM mean (being an estimate of the intensity of all pixels in the relationships that contributed to the GLCM), calculated as follows:

$$\mu = \sum_{i,j=0}^{N-1} ip(i,j)$$

σ is the variance of the intensities of all reference pixels in the relationships that contributed to the GLCM, calculated as:

$$\sigma = \sum_{i,j=0}^{N-1} p(i,j)(i-\mu)^2$$

The most regions similar to the region of the catheter were excluded after application delicate detection based on texture features and Gray-Level Co-Occurrence Matrix. It should also be

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

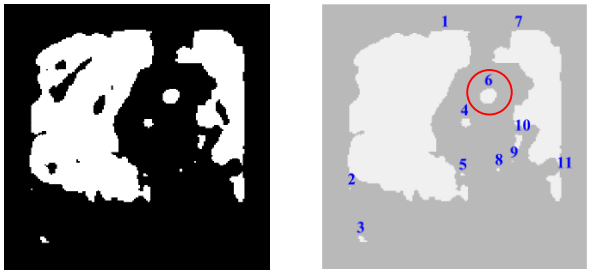
noted, that threshold levels for textural and intensity features were empirically found.

GLCM is a widely used technique, which is applied not only to feature extraction but also for segmentation tasks. For instance, A. Rampun, H. Strange et al. proposed a segmentation method based on consideration each feature at two different configurations in paper [13]. Mahesh B. Nagarajan, Markus B. Huber et al. used GLCM for lesion segmentation in the work [11].

2.2.5 Image processing

After acquiring the data and figuring out the catheter feature distribution, each slice is processed by automatic thresholding algorithm. The result of the application of Otsu's thresholding method is shown in Fig. 4a.

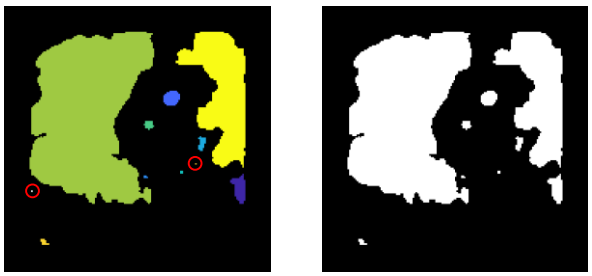
The morphological closing follows the thresholding step and performs a flood-fill operation on background pixels (see Fig. 4b). This operation fills the holes and determines the connected area. At this step, we used 4-connected neighborhoods connectivity. As it can be seen in Fig. 4b, the number of objects of these steps is 11, where the catheter is represented by the region 6.



(a) (b)

Fig. 4. Automatic thresholding (a) and flood-fill segmentation (b).

As we described above, one of the workflow steps is an intermediate procedure connected with excess blobs elimination. This step assumes that the region of the catheter cannot be smaller than 5 pixels even for an edged slice of the distal end of the catheter. This is why regions 2 and 9 reflected in Fig. 4b and shown in Fig. 5a in red circles were removed. The output of this step is a black and white (BW) mask with 9 objects.



(a) (b)

Fig. 5. Excess blobs elimination: colorized input (a) and BW output (b).

As illustrated in Fig. 6 most regions are removed after the step 7 of the workflow but there are 4 excess regions that meet the area criteria as well as the region of the catheter.

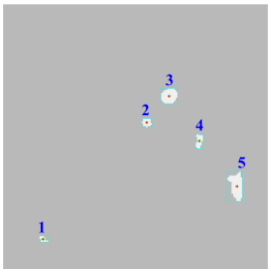


Fig. 6. Exclusion of the regions by area criteria.

After finishing the coarse detection and segmentation, we apply intensity analysis with further segmentation using the features. The first two features related to this step are the intensity mean and standard deviation of a certain region. To detect the region of the catheter, the mean and standard deviation were computed for each region on each slice of the dataset. The distribution of these two features for the current slice is presented in Fig. 7, where the average intensity and standard deviation for the catheter are 112 and 27, respectively. In its turn, several regions significantly differ from the region of the catheter, for example, region 1 and 4. In this regard, the obvious difference between the two classes (tissue and catheter) allowed linear SVM to set a threshold value for the features.

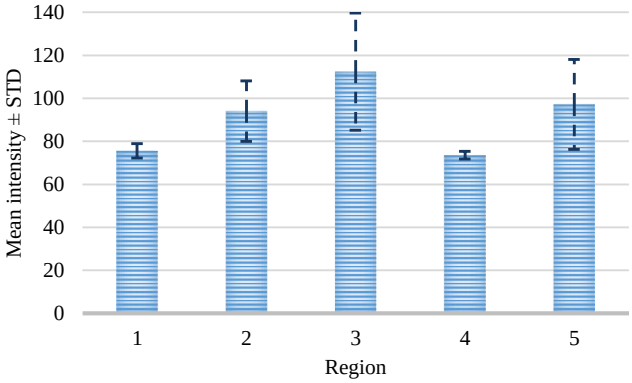
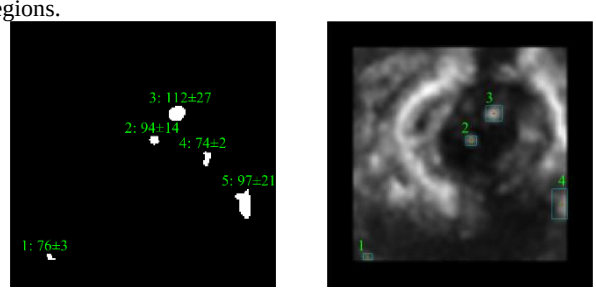


Fig. 7. Intensity analysis for the regions.

The values of these features for each region are compared to the limits and are defined whether they meet the criteria or not. Fig. 8 displays how these two features excluded one of the redundant regions.



(a) (b)

Fig. 8. Intensity analysis (a) with the following intensity-based detection using mean and standard deviation (b).

The most precise step of the workflow is texture analysis with GLCM. During this stage, four parameters such as contrast, homogeneity, correlation, and energy are computed using GLCMs of each region.

Four GLCMs for the remained regions are shown in Fig. 9. Afterward, textural features were compared with the established limits. Computed feature values of the texture analysis shown below in Fig. 10 were compared to the reference ones shown in Fig. 12.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

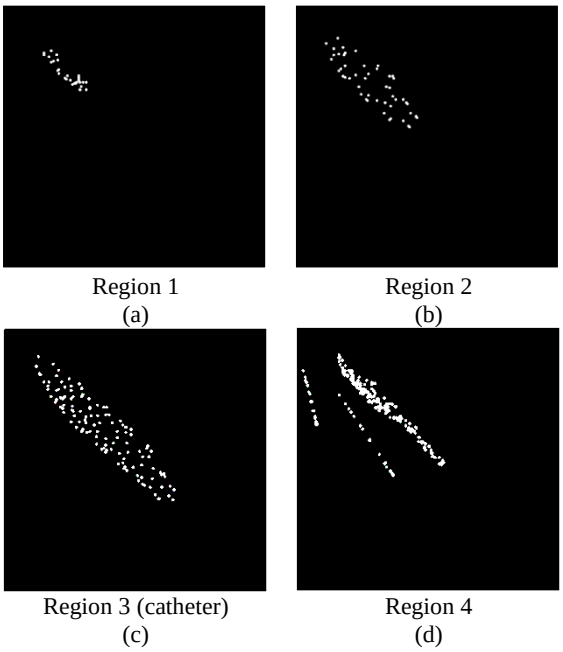


Fig. 9. Visual representation of GLCMs for remained regions.

The catheter (region 3) and two other regions (region 1 and 2) in Fig. 10 have sufficient uniformity, as indicated by the arrangement of the matrix elements on the central diagonal. In turn, region 4 has internal brightness differences. Since the GLCM uniformity of the region cannot be used as a feature of the catheter. Correlation, contrast, homogeneity, and energy were computed based on the obtained data. Further processing was performed on the obtained features.

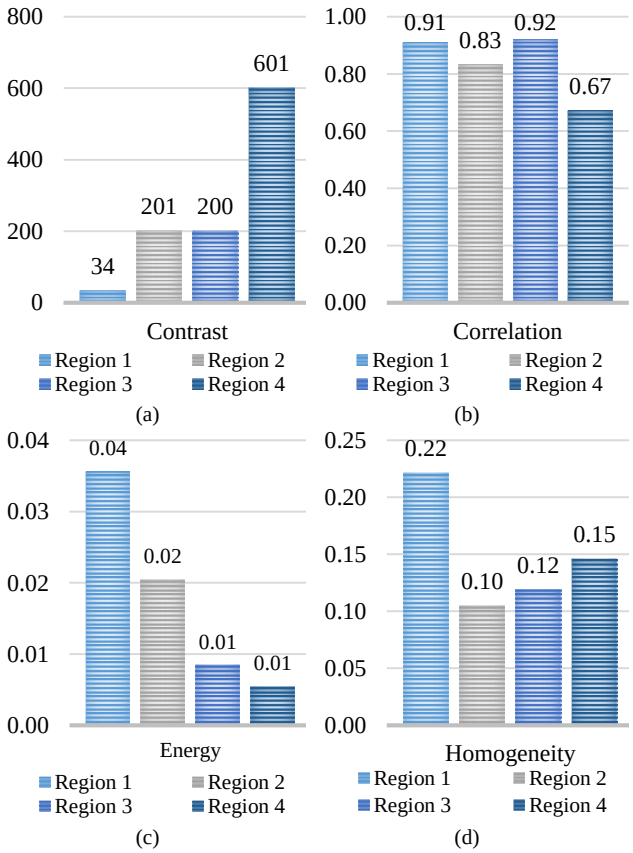


Fig. 10. Texture analysis: contrast (a), correlation (b), energy (c) and homogeneity (d).

The final procedure of the 2D stage found the desired ROI with certain features (see Fig. 11). As it can be seen in Fig. 11b, this step leaves only the region of the catheter and gives the output as a black and white mask which can be used for further 3D reconstruction.

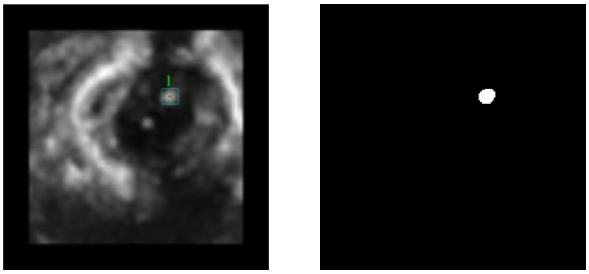


Fig. 11. A detected catheter (a) and its BW mask (b).

2.2.6 *Limitations*

In order to eliminate blobs and leave the desired region of interest (ROI), we imposed general area limitations performed in the beginning: the lowest and highest area limits are 5 and 200 pixels respectively. We also performed an estimation of feature limits within one timeframe consisting of 208 slices in order to find out the distribution of lowest and highest values for further restriction. This number of slices is sufficient to figure out the feature distribution because all settings for data acquisition were not changed within the study. In this regard, other 3D timeframes have relatively the same feature distribution. Feature limits obtained with 208 slices are shown in Fig. 12.

One of the main limitations of the algorithm is area restriction. Many excess binary large objects (blobs) frequently remain after application thresholding.

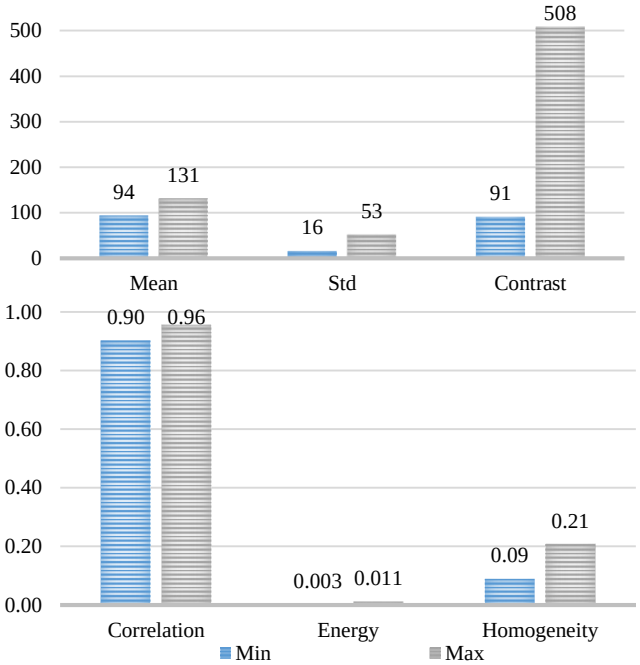


Fig. 12. Feature limits for the catheter.

3. Results

3.1 Accuracy

To access the accuracy of the algorithm, we compared two diameters: reference diameter and the diameter received after the application of the proposed algorithm. The catheter used for

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

performing medical procedures in our case had a diameter equal to 7 Fr (French Gauge) or 2.333 mm.

The diameter of the catheter was estimated using 2 ways: manual and automatic. For manual measurements, we used the short-axis view. Using this view gives a bigger sample of measurements what makes calculations of the diameter with higher precision. However, these measurements can be taken in the long-axis view but with lower accuracy. To assess the accuracy, the diameter was measured in 504 slices. To gauge the diameter automatically, we described the region of the catheter by the corresponding ellipse that has the same second moment (see Fig. 13).

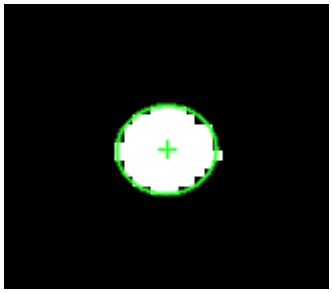


Fig. 13. Zoomed mask of the catheter region with an ellipse replacement.

As it can be seen in Table 1, the average diameter of the catheter after executing all processing steps equals to 2.47 mm what is 5.84% more than the reference diameter. However, such error does not have a strong influence on the visualization because 0.14 mm error lets a surgeon perform medical procedures without any confusions. It was observed that Otsu's thresholding method incrassates the region of the catheter by 5-10% this is why the error is increased as well. To decrease the error, modified Otsu's method and/or more delicate threshold level should be used. It's important to note that accuracy measurement performed by the algorithm took into account position and inclination of the catheter, which allowed better orientation along the correct axes.

Table 1. Comparison of the reference catheter diameter with the diameter obtained by the proposed algorithm.

	Reference diameter	Diameter measured by the algorithm
Mean, mm	2.33	2.47
Std, mm	0	0.16
Error, %	0	5.84

A reduction tendency of a number of regions over the different steps of the proposed algorithm is shown in Table 2. According to the results, there is a trend of decreasing the number of regions. The intensity-based and area-based features reduced the number of regions approximately by half. In turn, texture feature based on GLCM allowed excluding the false regions more accurately. Speaking of recognition accuracy and several mathematical statistics, they are shown in Table 3. The average recognition accuracy of the catheter is 87.2%. The confidence level for the sample including 504 cases equals 2.5%. The latter means that the recognition accuracy varies from 84.6% to 89.7% with a 5% significance level.

Table 2. A number of regions within different steps of the proposed algorithm.

Number of regions	Mean±STD
Initial number of regions	10.1±3.4
Number of false regions	9.1±3.4
Number of regions after area restriction (low limit)	8.9±3.2
Number of regions after area restriction (high limit)	6.7±2.9
Number of regions after intensity restriction	5.5±2.5
Number of regions after texture restriction	1.2±0.3

Table 3. Mathematical statistics and recognition accuracy.

Parameter	Value
Significance level	0.05
STD	28.8%
Size	504
Confidence interval	2.5%
Recognition accuracy	87.2%

3.2 Processing time

The desktop computer with equipment described in section 2.1 was used for time assessment. We did not isolate the testing process from the influence of other processes and did not allocate a separate thread. To find the mean and standard deviation of processing time we performed 150-iteration assessment. Processing time for one timeframe is 1.96±0.045 seconds. Each timeframe includes 208 slices what means that average processing time for 176*176 slice is 9±0.2 ms.

4. Conclusion

In this research study, we developed an image-based algorithm detecting and tracking the distal end of the catheter. To correctly detect and segment the catheter, we applied a feature-based approach that can recognize the catheter along the whole 3D dataset. However, the algorithm works in pseudo-3D what means that it processes the data in slice-by-slice mode. It worth noticing that the accuracy of the algorithm is at the relatively high level and equal to 94.16%. This means that the algorithm is in error by 140 micrometers. However, this error is acceptable for performing minimally-invasive cardiac surgery. Another vital feature of the algorithm is its processing time. Average processing time for one 3D timeframe is equal to 1.96 seconds and approximately 9 ms per slice. Though the algorithm is not time-consuming, it is still complicated to apply it to real-time surgery because of the huge amount of data obtained by echocardiography.

5. Acknowledgments

This work was supported in part by the Russian Federation Governmental Program "Nauka" № 12.8205.2017/БЧ (addition number: 4.1769.ГЗБ.2017). The experimental calculations are carried out at Tomsk Polytechnic University within the framework of Tomsk Polytechnic University Competitiveness Enhancement Program grant.

6. References

[1] Ambrosini, P., Smal, I., Ruijters, D., Niessen, W.J., Moelker, A. and Van Walsum, T. 2017. A Hidden Markov Model for 3D Catheter Tip Tracking with 2D X-ray Catheterization Sequence and 3D Rotational Angiography. *IEEE Transactions on Medical Imaging*. 36, 3 (2017), 757–768. DOI:https://doi.org/10.1109/TMI.2016.2625811.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

[2] Baert, S.A.M., Van de Kraats, E.B., Van Walsum, T., Viergever, M.A. and Niessen, W.J. 2003. Three-Dimensional Guide-Wire Reconstruction from Biplane Image Sequences for Integrated Display in 3-D Vasculature. *IEEE Transactions on Medical Imaging*. 22, 10 (2003), 1252–1258. DOI:https://doi.org/10.1109/TMI.2003.817791.

[3] Cheung, N.K., Boutchard, M., Carr, M.W. and Froelich, J.J. 2018. Radiation exposure, and procedure and fluoroscopy times in endovascular treatment of intracranial aneurysms: a methodological comparison. *Journal of NeuroInterventional Surgery*. (Jan. 2018), neurintsurg-2017-013596. DOI:https://doi.org/10.1136/neurintsurg-2017-013596.

[4] Haralick, R.M., Shanmugam, K. and Dinstein, I. 1973. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*. SMC-3, 6 (Nov. 1973), 610–621. DOI:https://doi.org/10.1109/TSMC.1973.4309314.

[5] Haralick, R.M. and Shapiro, L.G. 1992. *Computer and Robot Vision*.

[6] Hoffmann, M., Brost, A., Jakob, C., Bourier, F., Koch, M., Kurzidim, K., Hornegger, J. and Strobel, N. 2012. Semi-automatic catheter reconstruction from two views. *Medical image computing and computer-assisted intervention (MICCAI)*. 15, Pt 2 (2012), 584–91. DOI:https://doi.org/10.1007/978-3-642-33418-4_72.

[7] Hoffmann, M., Brost, A., Koch, M., Bourier, F., Maier, A., Kurzidim, K., Strobel, N. and Hornegger, J. 2016. Electrophysiology catheter detection and reconstruction from two views in fluoroscopic images. *IEEE Transactions on Medical Imaging*. 35, 2 (2016), 567–579. DOI:https://doi.org/10.1109/TMI.2015.2482539.

[8] Luani, B., Zrenner, B., Basho, M., Genz, C., Rauwolf, T., Tanev, I., Schmeisser, A. and Braun-Dullaeus, R.C. 2018. Zero-fluoroscopy cryothermal ablation of atrioventricular nodal re-entry tachycardia guided by endovascular and endocardial catheter visualization using intracardiac echocardiography. *Journal of Cardiovascular Electrophysiology*. 29, 1 (Jan. 2018), 160–166. DOI:https://doi.org/10.1111/jce.13354.

[9] Ma, Y., Gogin, N., Cathier, P., Housden, R.J., Gijsbers, G., Cooklin, M., O'Neill, M., Gill, J., Rinaldi, C.A., Razavi, R. and Rhode, K.S. 2013. Real-time x-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions. *Medical physics*. 40, 7 (2013), 071902. DOI:https://doi.org/10.1118/1.4808114.

[10] Muraru, D., Niero, A., Rodriguez-Zanella, H., Cherata, D. and Badano, L. 2018. Three-dimensional speckle-tracking echocardiography: benefits and limitations of integrating myocardial mechanics with three-dimensional imaging. *Cardiovascular Diagnosis and Therapy*. 8, 1 (Feb. 2018), 101–117. DOI:https://doi.org/10.21037/cdt.2017.06.01.

[11] Nagarajan, M.B., Huber, M.B., Schlossbauer, T., Leinsinger, G., Krol, A. and Wismüller, A. 2013. Classification of small lesions in dynamic breast MRI: eliminating the need for precise lesion segmentation through spatio-temporal analysis of contrast enhancement. *Machine Vision and Applications*. 24, 7 (Oct. 2013), 1371–1381. DOI:https://doi.org/10.1007/s00138-012-0456-y.

[12] Otsu, N. 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*. 9, 1 (Jan. 1979), 62–66. DOI:https://doi.org/10.1109/TSMC.1979.4310076.

[13] Rampun, A., Strange, H. and Zwiggelaar, R. 2013. Texture segmentation using different orientations of GLCM features. *Proceedings of the 6th International Conference on Computer Vision / Computer Graphics Collaboration Techniques and Applications - MIRAGE '13* (2013).

[14] Schwein, A., Chinnadurai, P., Behler, G., Lumsden, A.B., Bismuth, J. and Bechara, C.F. 2018. Computed tomography angiography-fluoroscopy image fusion allows visceral vessel cannulation without angiography during fenestrated endovascular aneurysm repair. *Journal of Vascular Surgery*. (Feb. 2018). DOI:https://doi.org/10.1016/j.jvs.2017.11.062.

[15] Soille, P. 2004. *Morphological Image Analysis*. Springer Berlin Heidelberg.

[16] Spitzer, E., Ren, B., Zijlstra, F., Mieghem, N.M. Van and Geleijnse, M.L. 2017. The Role of Automated 3D Echocardiography for Left Ventricular Ejection Fraction Assessment. *Cardiac Failure Review*. 3, 2 (Nov. 2017), 97. DOI:https://doi.org/10.15420/cfr.2017:14.1.

[17] Tepeler, A., Binbay, M., Yuruk, E., Sari, E., Kaba, M., Muslumanoglu, A.Y. and Tefekli, A. 2009. Factors affecting the fluoroscopic screening time during percutaneous nephrolithotomy. *J Endourol*. 23, 11 (2009), 1825–1829. DOI:https://doi.org/10.1089/end.2009.0256.

[18] Wu, X., Housden, J., Varma, N., Ma, Y., Rueckert, D. and Rhode, K. 2013. Catheter tracking in 3D echocardiographic sequences based on tracking in 2D X-ray sequences for cardiac catheterization interventions. *Proceedings - International Symposium on Biomedical Imaging* (2013).

About the authors

Viacheslav V. Danilov is a Ph.D. student at Tomsk Polytechnic University and a research scientist at the Medical Devices Design Laboratory. His e-mail is viacheslav.v.danilov@gmail.com.

Igor P. Skirnevskiy is a lecturer at Tomsk Polytechnic University and a research scientist at the Medical Devices Design Laboratory. His e-mail is skirnevskiy@tpu.ru.

Roman A. Manakov is a Master's student at Tomsk Polytechnic University. His e-mail is ram290495@gmail.com.

Olga M. Gerget is an Assistant Professor at Tomsk Polytechnic University and a research scientist at the Medical Devices Design Laboratory. Her e-mail is gerget@tpu.ru.

Farid Melgani is a Professor at the University of Trento and a Head of the Signal Processing and Recognition Lab. His e-mail is farid.melgani@unitn.it.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

CNN-based histological images segmentation of mucous glands

A. Khvostikov[1], A. Krylov[1], O. Kharlova[2], N. Oleynikova[2], I. Mikhailov[2], P. Malkov[2]
xubiker@gmail.com | kryl@cs.msu.ru
olga.arsenteva@gmail.com | noleynikova@mc.msu.ru | imihailov@mc.msu.ru | pmalkov@mc.msu.ru
[1]Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, Russia;
[2]University Medical Center, Lomonosov Moscow State University, Moscow, Russia

Mucous glands lesions analysis and assessing of malignant potential of colon polyps are very important tasks of surgical pathology. However, differential diagnosis of colon polyps often seems impossible by classical methods and it is necessary to involve computer methods capable of assessing minimal differences to extend the capabilities of the classical pathology examination. Accurate segmentation of mucous glands from histology images is a crucial step to obtain reliable morphometric criteria for quantitative diagnostic methods. We review major trends in histological images segmentation and design a new convolutional neural network for mucous gland segmentation.

Keywords: *Image segmentation, Convolutional Neural Networks, Histology, Pathology, Mucous glands*

1. Introduction

Mucous glands are important histological structures presented in certain organ systems as the main mechanism for secreting proteins and carbohydrates. Malignant tumors arising from glandular epithelium, also known as adenocarcinomas, are the most prevalent form of cancer. However for the mucous glands the sequence of precancerous processes leading to the development of a malignant tumor has been studied most well: metaplasia-dysplasia-cancer. The detection of metaplasia and dysplasia in the epithelium of the mucous glands may be an extremely difficult task for surgical pathology. Minimal differences between these processes are often not available for the classical pathology examination. An important practical task is to find analytical instrument for objective diagnosis. Computational analysis is an additional analytical instrument for the segmentation of mucous glands on a microscopic images, shape analysis of the whole gland and it's lumen (Fig.1), calculating nuclear-cytoplasmic ratio in mucus-forming cells, and localization of the expression of certain immunohistochemical markers. All these parameters together can potentially significantly increase the accuracy of diagnosis of pathological processes associated with mucous glands in colon, stomach, prostate, breast, bronchus, etc.

Accurate segmentation of glands is often a crucial step to obtain reliable morphological statistics. Nonetheless, the task by nature is very challenging due to the great variation of glandular morphology in different histological grades.

Today the standard dataset for the task of histological images segmentation is Warwick-QU image dataset [1], first used in [2]. The dataset provides 165 Hematoxylin and Eosin (H&E) stained slides, consisting of a variety of histological grades. The dataset is provided together with ground truth annotations made by expert pathologists. Every image contains a number of colon glands, each of them consists of lumen, cytoplasm and nuclei (Fig.1). Warwick-QU dataset was also used in the Gland Segmentation in Colon Histology Images (GlaS) contest [3].

The goal of this paper is to give a short overview of recent trends in histological images segmentation and propose a new algorithm for mucous glands segmentation. The target aim of our research is the computer-aided diagnostics of colon polyps – adenomas (ICD-O 8140/0) and serrated lesions (ICD-O 8213/0).

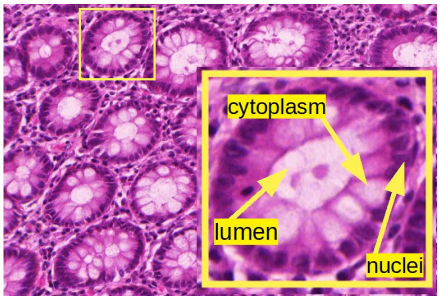


Figure 1. A histological image of colon mucous glands.

2. Review of gland segmentation methods

All major trends in gland segmentation can be roughly split into two groups (those which are based on the use of conventional image processing techniques and more modern methods based on convolutional neural networks), which are overviewed in Section 2.1 and Section 2.2 accordingly.

2.1 Conventional methods of segmentation

The first methods of image segmentation ever applied to histological images were methods based on K-means clustering, Chan-Vese algorithm, histogram analysis, edge detection, and watershed algorithm. Although these methods of segmentation are very simple, they are not effective enough, and can be used only for semiautomatic segmentation [4].

The next generation of algorithms was based on a more complex feature-based approaches and could perform automatic image segmentation in most cases. E.g. a region growing algorithm is presented in [5].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

The initial seeding regions are identified based on the lumen inside the glands by fitting with a large moving window. The seeding regions are then expanded by repetitive application of a morphological operation. False gland regions are removed based on either their excessive ages of active growth or inadequate thickness of dams formed by goblet cell nuclei outside the grown regions.

In [6] the tissue image is segmented using the definition of object-graphs instead of using the pixel-based approach. The image is decomposed into a set of circular objects, and depending on the organizational properties of these objects initial seeds are determined. After that the inner regions of glands are found by growing the seeds in accordance with nuclei location.

A more complex approach was proposed in [2]. Each gland in the image is treated as a polygon of random number of vertices, while the vertices represent approximate locations of epithelial nuclei. The problem of constructing such a graph is formulated as a Bayesian inference problem by defining a prior for spatial connectivity and arrangement of neighboring epithelial nuclei and a likelihood for the presence of a gland structure.

2.2 Segmentation with neural networks

Recently, image segmentation methods using convolutional neural networks are gaining in popularity. Due to their good generalization capacity and versatility these methods demonstrate the state-of-the-art level of performance. Almost all CNN-based segmentation methods use the same idea of convolutional autoencoder (CAE) [7]. With minor changes these CNN-based segmentation methods can be also applied to histological images.

The first approach of semantic image segmentation using CNN was Fully Convolutional Network (FCN) proposed by Long et al. in [8]. The key ideas of the approach are to replace the fully-connected layers with convolutions, perform upsampling with transposed convolutions (or so called upconvolutions) and use skip connections. All in all FCN re-purposes imagenet pretrained networks for semantic segmentation problem.

Despite it's upconvolutional layers and a few shortcut connections FCN produces too coarse segmentation maps. SegNet [9] proposes another way of upsampling. The decoder part of the SegNet uses pooling indices computed in the max-pooling step of the corresponding encoder part to perform non-linear upsampling. This leads to more accurate segmentation and makes SegNet more memory efficient than FCN.

The main contribution of U-Net [10] compared to other fully convolutional segmentation networks is that while upsampling and going deeper in the network the upsampled features are concatenated with the higher resolution features from down part for better localization. Furthermore, apart from proposing the new architecture and a special way of data augmentation U-Net was the first CNN-based method applied for biomedical image segmentation.

The main problem of the mentioned above approaches of segmentation is the impossibility of the algorithms to separate close or contiguous objects. Various ideas were proposed to solve this problem.

For example in [11] the authors introduced a new high-level loss function, that takes into account high level shape priors, such as smoothness and preservation of complex interactions between object regions. This loss was used in the FCN network trained on the the Warwick-QU dataset and demonstrated better performance compared to the FCN trained with conventional per-pixel loss.

In [12] Kainz et al. trained two FCN networks on the Warwick-QU dataset. The first network performs a 4-class object segmentation (benign background, benign gland, malignant background and malignant gland), while the second one is trained to separate close glands. The CNN predictions are then regularized using weighted total variation to produce the final segmentation result.

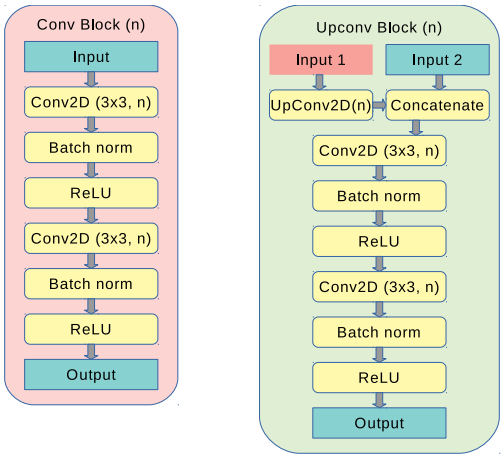
Chen et al. in [13] proposed a DCAN architecture with similar idea of object detection and separation, but unlike [12] these two steps are performed simultaneously with one FCN-based network that has two outputs. First output predicts probabilities of gland object, while the second predicts the probability map of contours separating glands. The final segmentation masks are calculated using the threshold rule. To strengthen the training process DCAN uses 3 weighted auxiliary classifiers in the 3 deepest layers of the network.

The idea of splitting segmented glands got a further development in [14]. The authors introduce a CNN with 3 pipelines: a FCN for the foreground segmentation, Faster R-CNN [15] for the object detection and HED [16] for edge detection. All three pipelines a fused into one, and are followed with several convolution layers to predict the final instance segmentation map. This approach leads to the state-of-the-art level of segmentation accuracy.

3. Proposed method

In this paper we consider the problem of histological mucous gland segmentation. For this problem we suggest a rather simple yet effective CNN architecture. The approach is based on the U-Net [10] with batch normalization [17] layers added after each convolution. The main building blocks of the network are Conv block (Fig.2(a)) and Upconv block (Fig.2(b)).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18



(a) Convolution block (b) Upconvolution block

Figure 2. Base building blocks of the proposed network.

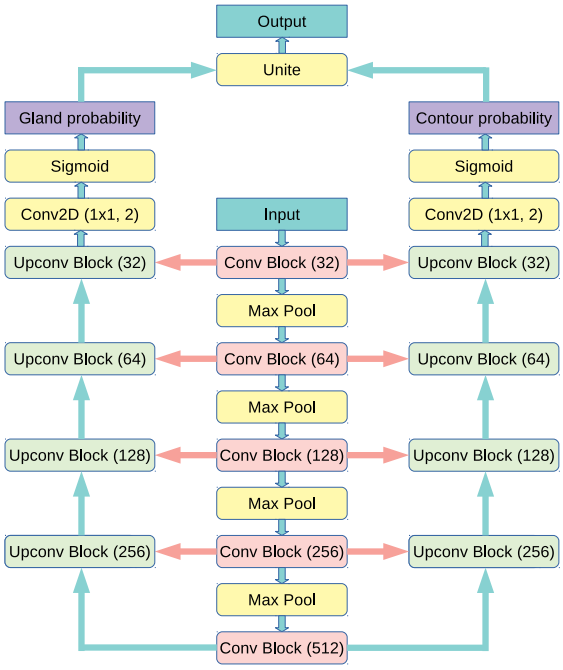


Figure 3. Proposed network architecture for mucous glands segmentation.

Conv blocks form the contracting path of the network, while Upconv blocks form the expansive path. Here Conv2D (SxS, n) operation stands for the 2D convolution with n filters of size $S \times S$, while UpConv2D (n) operation stands for the double upsampling of the feature map followed by 2D convolution with n filters of size 2×2. In contrast to the U-Net we use two different expansion paths in order to detect glands and separating contours. Each expansion paths ends with a 1×1 2D convolution, which reduces the number of features to 2 (background/foreground) and a sigmoid activation function to get the output probability map. Thus, the first expansion path outputs the probability of gland and the second - the probability of gland contour. Uniting these two probabilities by threshold we get the final result of segmentation (Fig.3). The network is learned using RMSprop optimizer [18], dividing the gradient by a running average of its recent magnitude. We use a sum of a binary cross entropy value and a dice coefficient as the loss for both of the network outputs.

4. Experiments and results

The algorithm was evaluated on the Warwick-QU dataset [1], which consists of a wide range of histological grades from benign to malignant subjects. The dataset contains 85 train and 80 test images. We enlarged the dataset by 10 times using augmentation process with random shift, rotation, zoom and flip operations.

The proposed algorithm was implemented using open source neural network library Keras [19] with TensorFlow backend. The training was performed on a platform with Intel(R) Core(R) i7-6700HQ CPU and NVIDIA GeForce GTX 960M GPU. Due to the relatively high resolution of the images from the used dataset, every processed image is first split into non-overlapping patches of 256×256 size, each of the patches is fed to the network and after that the output image is merged from the processed patches.

The segmentation results after training the proposed network for 12 epochs are shown in Fig.4. The sum of the binary cross entropy and the dice coefficient was 0.872 on the validation set. The evaluation on the test set demonstrates sufficiently good results in both gland and contour segmentation.

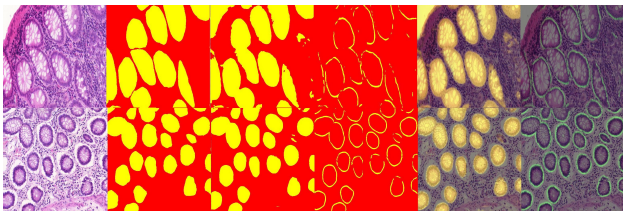


Figure 4. Segmentation results on the test images from dataset [1]. The columns: source image, ground truth mask, network outputs (gland and contour detection), superpositions of source image and the outputs.

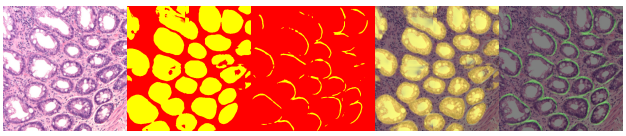


Figure 5. Segmentation results of an image fragment of mucous glands in hyperplastic colon polyp in real biopsy diagnostic material. The columns: source image; network outputs (gland and contour detection); superpositions of source image and the outputs.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Moreover, Fig.5 demonstrates the results of applying the proposed algorithm for the classification of real biopsy diagnostics material. It should be noticed, that for this example the network was trained on the Warwick-QU dataset [1] and was not fine-tuned on the used data due to the absence of the ground-truth markup. Even though, the obtained results look reasonable. Creating the markup for this data and fine-tuning the network are the further work directions.

5. Conclusion and further directions

Further directions of the our work include performing a more complex segmentation to make also an inner-gland segmentation (detect nuclei, lumen and cytoplasm), which can be used for the ensuing analysis. In particular, analyzing the histological images of mucous glands helps to detect changes in its lumen shape (serration), in the nuclear-cytoplasmic ratio inside mucus-forming cells, and in the character of the expression of immunohistochemical markers [20].

6. Acknowledgements

The work was supported by Russian Science Foundation grant 17-11-01279. In our study, the review of segmentation methods and designing the CNN were performed by the authors from the Faculty of Computational Mathematics and Cybernetics, the medical problem statements and experimental work - by the authors from the University Medical Center.

7. References

[1] Warwick-QU image dataset description. https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/about/.

[2] Korsuk Sirinukunwattana, David RJ Snead, and Nasir M Rajpoot. A stochastic polygons model for glandular structures in colon histology images. *IEEE transactions on medical imaging*, 34 (11):2366–2378, 2015.

[3] Korsuk Sirinukunwattana and et al. Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis*, 35: 489–502, 2017.

[4] Rodrigo Fernandez-Gonzalez and et al. Automatic segmentation of histological structures in mammary gland tissue sections. *Journal of biomedical optics*, 9(3):444–454, 2004.

[5] H-S WU, R Xu, N Harpaz, D Burstein, and J Gil. Segmentation of intestinal gland images with iterative region growing. *Journal of Microscopy*, 220(3):190–204, 2005.

[6] Cigdem Gunduz-Demir, Melih Kandemir, and et al. Automatic segmentation of colon glands using object-graphs. *Medical image analysis*, 14 (1):1–12, 2010.

[7] Jonathan Masci, Ueli Meier, Dan Cireşan, and Jürgen Schmidhuber. Stacked convolutional auto-encoders for hierarchical feature extraction. In *ICANN*, pages 52–59. Springer, 2011.

[8] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

[9] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on PAMI*, 39(12): 2481–2495, 2017.

[10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[11] Aïcha BenTaieb and Ghassan Hamarneh. Topology aware fully convolutional networks for histology gland segmentation. In *MICCAI*, pages 460–468. Springer, 2016.

[12] Philipp Kainz, Michael Pfeiffer, and Martin Urschler. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. *PeerJ*, 5: e3874, 2017.

[13] Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical image analysis*, 36:135–146, 2017.

[14] Yan Xu and et al. Gland instance segmentation using deep multichannel neural networks. *IEEE Transactions on Biomedical Engineering*, 64(12): 2901–2912, 2017.

[15] Shaoqing Ren and et al. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.

[16] Saining Xie and Zhuowen Tu. Holistically-nested edge detection. In *Proceedings of the IEEE international conference on computer vision*, pages 1395–1403, 2015.

[17] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.

[18] RMSprop optimizer. https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.

[19] François Chollet et al. Keras. https://keras.io, 2015.

[20] OA Kharlova, NV Danilova, PG Malkov, NV Ageikina, and MV Knyazev. Serrated lesions of the large intestine: Prevalence analysis and clinical and morphologic characteristics. *Arkhiv patologii*, 77(4):24–32, 2015.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Computer analysis and restoration of motion blurred images

A.A. Melnikova, A.S. Krylov
alexandra.melnikova.cmc.msu@gmail.com|kryl@cs.msu.ru
Faculty of Computational Mathematics and Cybernartics
Lomonosov Moscow State University, Moscow, Russia

Histogram based and correlation based algorithms for computer analysis and restoration of motion blurred images have been proposed. The histogram algorithm uses ridge detection method and direction distribution histogram analysis. The correlation algorithm finds a detection function, that when superimposed on the Fourier spectrum of the motion blurred image, gives a minimum. The results of comparison between the histogram method and the correlation method are presented. The tests show practically good results for the realistic motion blur lengths.

Keywords: *image blur, motion bur, blur kernel estimation, ridge detection, histogram method.*

1. Introduction

Motion blur on the image occurs when the relative motion between the scene and/or the sensor occurs during exposure. Such an effect makes the recognition and perception of images difficult. To apply algorithms to restore blurred images, it is required to evaluate the blur kernel. Thus, success of blurred image restoration significantly depends on an accuracy of blur parameters estimation.

This paper is the continuation of [1], where an algorithm for image blur kernel estimation for the case of uniform motion blur was proposed. In this work the proposed histogram method was adapted for the blur kernel, which uses a third, additional, Gaussian blur parameter. A comparison between the histogram method and correlation method based on the idea suggested in [4] has been also performed.

For motion blur parameter estimation various techniques are being used. A method based on the analysis of power cepstrum of the image has been proposed in [2]. In [3] motion blur estimation is based on the analysis of shape, the homogeneity and the smoothness. It is proposed to analyze the correlation between the Fourier spectrum of blurred image and a periodic detecting function in [4], and the results of comparison between method [4], [2] and [3] were presented. In [5] a spectral analysis of image gradients had been performed to identify the blurring kernel, in [6] it was proposed to analyze the Hough transform. In [7] the authors suggest using modifications of the Radon transform for the identification of the blur spectrum pattern. In [8] a novel PSF estimation scheme based on frequency spectrum analysis has been proposed.

The motion blur parameter estimation is used in various areas. As an example, in [9] motion blurred images analysis was used for vehicle speed detection. Also, motion blur estimation is used for document image deblurring [10].

Formulation of the problem. In this paper we consider the uniform linear motion blur. Blurred image can be represented as:

$$g(x, y) = f(x, y) * h(x, y),$$

where $h(x, y)$ - is a blur function, or PSF, $f(x, y)$ - uncorrupted image, $g(x, y)$ - blurred image.

The PSF (Point Spread Function), or blur kernel, in the case of uniform blur, can be defined as [11]:

$$h(x, y) = \begin{cases} \frac{1}{L}, & \sqrt{x^2 + y^2} \le \frac{L}{2}, y = x \tan(\theta) \\ 0, & else \end{cases}, \quad (1)$$

where L, θ are the length and the angle, which defines blur direction, respectively. Figure 1 shows an example of the PSF (1).

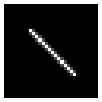


Figure 1. An illustration of the PSF function (1), corresponded to $\alpha = 45, L = 15$). The contrast of the image was increased.

Also, we used the Gaussian function for motion blur simulation:

$$h(x', y', \alpha) = e^{\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)}, \quad (2)$$

where

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (3)$$

In this model we use 3 parameters: $\sigma_x, \sigma_y, \alpha$ - length, width and direction respectively. Adding a width parameter and a Gaussian blur makes the model more appropriate to real images.

Thus, in this paper the definition of motion blur parameters is the determination of length and direction or length, direction and width. The ridge detection method [12] and direction distribution histogram analysis are used as the main techniques.

This paper is organized as follows. Section 2 includes the histogram algorithm description. Correlation method for motion blur parameters estimation is described in Section 3. The testing results and their discussion are given in Section 4.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2. Histogram algorithm for motion blur parameters estimation

Algorithm for determining the direction of the motion blur vector. To determine blur direction we use an approach based on ridge detection.

The ridges (valleys) of a smooth function of two variables are a set of curves whose points are local maxima (minima) of the function in at least one dimension.

The curves whose points correspond to the maxima are called ridges, and whose points correspond to the minima are called valleys. We use ridges for motion blur analysis.

Let us consider an image as a function of two variables $I(x, y)$. For each pixel we compose the Hessian:

$$\begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix}.$$

Let λ_1 and λ_2 be the eigenvalues of this matrix, then ridges can be defined as $\lambda_{max} = MAX(\lambda_1, \lambda_2)$, where the maximum denotes the eigenvalue with the greatest magnitude and where $\lambda_1 \gg \lambda_2$. The direction of the eigenvector, which corresponds to λ_{max} define direction of the ridge.

To extract the skeletons corresponded to ridges we apply non-maximum suppression algorithm. This step significantly increases the accuracy of determining blur direction.

Thus, for each point of the selected ridge we can define the angle of inclination of the tangent to the ridge by following formula:

$$\theta = tan^{-1}\left(I_{xy}/(\lambda_{max} - I_{xx})\right). \quad (4)$$

When the motion blur occurs, the contours of the image are stretched in the corresponding direction, which we take as the desired parameter. To determine it, we construct a direction distribution histogram. Note that the angle θ is rounded to integer value when computing histogram.

The horizontal axis in the histogram corresponds to the value of the angle from 0^o to 180^o, and the vertical axis shows the number of pixels with the corresponding angle value calculated from the formula (4). To increase the accuracy, we use for histogram construction not only the pixels selected by non-maximum suppression algorithm, but also their 1 pixel radius neighborhood. We take the average maximum obtained from the peak analysis on the histogram as the desired direction parameter.

Algorithm for determining the length of the motion blur vector. To determine the length of the motion blur vector, we suggest gradually blurring the image in a direction that is perpendicular to the existing blur direction, gradually increasing the blur length. The initial length of the perpendicular blur and the step inside the loop can be chosen depending on the considering task. With each iteration the image becomes more and more defocused, and the direction distribution histogram changes.

We denote the value $S_L = \sum_{i=0}^{90} hist(i)$ $\left(S_R = \sum_{i=90}^{180} hist(i)\right)$ as *the left sum (the right sum) of the direction distribution histogram.*

Thus, with each iteration the histogram tends to the moment when the sum of the left part of the histogram is approximately equal to the sum of the right part of the histogram. It is proposed to leave the loop at the moment when the corresponding sum of the histogram becomes greater than the opposite sum. We take the value of the iteration parameter, at which the sign of the difference between the left and right histogram sums has changed, as desired parameter of length.

Algorithm for determining the "width" of the motion blur kernel. The ridge detection method depends on the parameter σ. Changing it we can find ridges at different scales. Therefore, for different values of σ, we can get different values of the direction of the blur vector.

Thus, the accuracy of determining the direction of motion blur depends on the accuracy of finding the parameter σ, which corresponds to the width of the motion blur kernel (2). Finding a good approximation of σ makes possible to increase the method accuracy and makes the method automatic. For these purposes we used a multiscale ridge detection method.

We build a grid of the following values $\sigma \in \{\sigma_1, ..., \sigma_n\}, \sigma_i = \sigma_0 \nu^{i-1}$, where $\sigma_0 = 1, \nu = \sqrt{2}$. We used $n = 6$. Consequently, the corresponding values are equal to $\sigma = \{1, \sqrt{2}, 2, 2\sqrt{2}, 4, 4\sqrt{2}\}$. For each $sigma_i$ we perform ridge detection and get the image $H^{\sigma_i}(x, y)$. To equalize the values between Hessians, obtained for different σ_i, we multiply the derivatives of the image by σ^2 at the construction stage. To suppress false filter responses we use threshold value of 0.1 for small σ.

It is assumed that if there are point sources on the image, than the number of pixels on the skeletons corresponded to ridges will predominate in some direction. Basing on this consideration we choose σ_i corresponding to the image with the largest number of ridges.

3. Correlation method for motion blur parameters estimation

To compare the results obtained by method described at section 2, we developed a correlation algorithm based on the idea from [4]. Such approach can be applied to the images corrupted with motion blur, which corresponds to (1) model. Let us denote correlation algorithm as "algorithm T".

The main idea of the algorithm T is to choose such a detection function, which, when superimposed on

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

the Fourier spectrum of the motion blurred image, gives a minimum.

We consider a blurry image $g(x,y) = h(x,y) * f(x,y)$, where $h(x,y)$ is PSF, $f(x,y)$ is uncorrupted image and $g(x,y)$ is blurred image. Then the Fourier transform of image (hereinafter referred to as FT) can be represented as $G(\xi,\eta) = H(\xi,\eta)F(\xi,\eta)$, where $G(\xi,\eta), H(\xi,\eta), F(\xi,\eta)$ - are FT of $g(x,y), h(x,y), f(x,y)$ respectively. The main idea of the method is to represent the function $H(\xi,\eta)$ using *sinc*-function. On images that have a blur kernel close to the model (1) we can observe zero-patterns on Fourier spectrum. It is required to identify zero-patterns on the Fourier spectrum of the blurred image by estimating the correlation between the detection function $\rho(\mu)$ and the Fourier spectrum of the blurred image.

We assume that the detection function $\rho(\mu)$ is periodic. Consequently, changing the period we can achieve the smallest value of the correlation function.

The correlation function based minimization method can be presented as

$$I(L,\theta) = \int\int w(\xi,\eta)\rho(\mu(\xi,\eta;L,\theta)\log|G(\xi,\eta)|)d\xi d\eta,$$

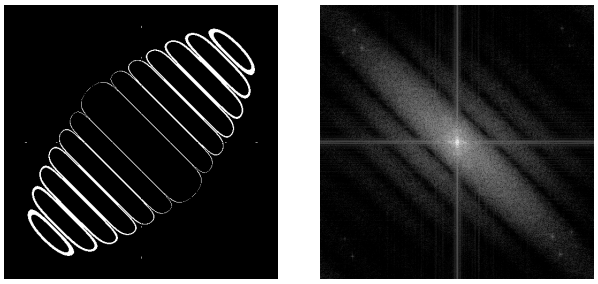
$$\mathrm{I(L,}\ \theta) \to min,$$

where $w(\xi,\eta)$ - weighting function, $\mu(\xi,\eta;L,\theta) = L(\xi\cos\theta + \eta\sin\theta)$. To reduce the influence of noise, the weight function is built according to the rule: the higher the frequency, the less weight.

The detection function was built according to the following rule:

$$k(\mu) = |\frac{\sin(\mu)}{\mu}| * 500 - \frac{1}{100 * G(x,y,\sigma_x,\sigma_y)},$$
$$\mu = \frac{1}{l}(\xi\cos\theta + \eta\sin\theta),$$
$$z(\mu) = \begin{cases} \frac{k(\mu)}{2}, & |\mu| < \frac{l}{2} \\ 0, & z(\mu) < 0 \\ k(\mu) & else \end{cases},$$
$$\rho(\mu) = \begin{cases} 255, & z(\mu) > 0, |z(\mu)| < 10 \\ 0, & else \end{cases},$$

where $G(x,y,\sigma_x,\sigma_y)$ - Gaussian function, $\sigma_x = 30$, $\sigma_y = 70$. The parameter α corresponds to the direction of the blur, and l corresponds to the blur length. To perform the rotation of this function, the coordinates are transformed using the formula (3). Figure 2 shows an example of the detection function and the Fourier spectrum of the blurred image.



a) b)

Figure 2. Detection function (a) corresponded to Fourier spectrum of baboon image ($\alpha = 45, L = 5$) (b).

4. Testing and results

Comparison of the results of histogram and correlation algorithms. The comparison results (just an example) of the histogram algorithm and the correlation algorithm are given in the Table. 1. Note that here we use half of the length.

$L/2$-ϕ	$L/2$	$L/2$(method T)	ϕ	ϕ (method T)
5-45	5	6	43^o	42^o
7-45	7	6	43.5^o	42^o
8-45	7	11	43.5^o	45^o
9-45	8	11	44^o	42^o
10-45	8	11	44.5^o	45^o

Table 1. Comparison of the results with the correlation algorithm. Test images were blurred with the kernel (1).

It can be concluded that the histogram algorithm is more accurate for a small value of the parameter L than the correlation algorithm. With the increase of the blur length, both algorithms give approximately the same accuracy.

The results of the histogram algorithm. The results of the histogram algorithm applied to images blurred with a kernel (2) presented in Table.2.

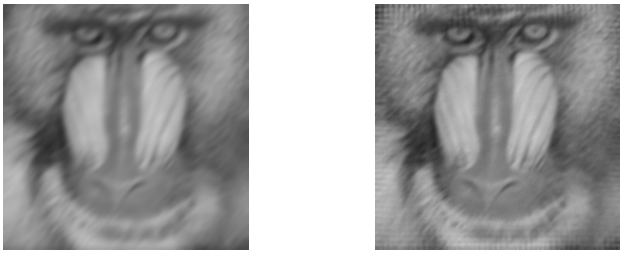
L-α-W	α (estimated)	L(estimated)	W(estimated)
4-30-1	4	29.5^o	1.4
5-30-3	5	30.5^o	2
7-30-3	7	29.5^o	2
8-30-3	8	29.5^o	2
10-30-4	9	28^o	4
15-45-3	15	45.5^o	4

Table 2. The results of the histogram algorithm applied to images blurred with a kernel (2).

Restoration of motion blurred images. Using the proposed algorithms it is possible to determine the desired parameters of the PSF, and now it becomes possible to apply an inverse regularizing method to find an unblurred image. Since this task is ill-posed [13],

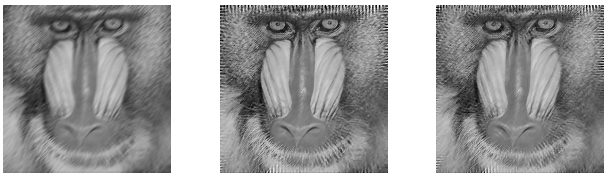
The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

even knowing the blur kernel it is difficult to restore the original image. To restore the blurred images we used a total variation based method [14]. After applying the restoration algorithm, the sharpness of the images is increased (Fig.3, Fig.4).



a) b)

Figure 3. Blurred image baboon ($\alpha = 45, L = 8, W = 3$;) (a); image restored using total variation deblurring [14] (b).



a) b) c)

Figure 4. Image blurred with $\alpha = 45$, $L = 7$(a); image (a) restored using parameters acquired using histogram method (b) and correlation method (c).

5. Conclusion

In this paper, histogram and correlation method for analyzing and reconstructing blurred images have been developed, the tests show practically good results. The histogram algorithm is more accurate for a small value of the parameter L than the correlation algorithm. With the increase of the blur length, both algorithms give approximately the same accuracy.

6. Gratitude

The work was supported by the Russian Foundation for Basic Research, grant 16-07-01207.

7. References

[1] *Melnikova A.A., Krylov A.S.* "Image motion blur kernel estimation using ridge analysis." GraphiCon 2017: proceedings of the 27th GraphiCon international conference, 2017. - P. 279–283.

[2] *Cannon M.* "Blind deconvolution of spatially invariant image blurs with phase." IEEE Transactions on Acoustics, Speech, and Signal Processing 24, no. 1 (1976): 58-63.

[3] *Yitzhaky Y., Kopeika N. S.* "Identification of blur parameters from motion blurred images." Graphical models and image processing 59, no. 5 (1997): 310-320.

[4] *Tanaka M., Yoneji K., Okutomi M.* "Motion blur parameter identification from a linearly blurred image." In Consumer Electronics, 2007. ICCE 2007. Digest of Technical Papers. International Conference on, pp. 1-2. IEEE, 2007.

[5] *Ji H., Liu C.* "Motion blur identification from image gradients." In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pp. 1-8. IEEE, 2008.

[6] *Phansalkar N.* "Determination of linear motion point spread function using hough transform for image restoration." In Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference on, pp. 1-4. IEEE, 2010.

[7] *Oliveira J. P., Figueiredo M. A. T., Bioucas-Dias J. M.* "Parametric blur estimation for blind restoration of natural images: Linear motion and out-of-focus." IEEE Transactions on Image Processing 23, no. 1 (2014): 466-477.

[8] *Wang Z., Yao Z., Wang Q.* "Improved scheme of estimating motion blur parameters for image restoration." Digital Signal Processing 65 (2017): 11-18.

[9] *Lin H. Y., Li K. J.* "Motion blur removal and its application to vehicle speed detection." In ICIP, pp. 3407-3410. 2004.

[10] *Qi X. Y., Zhang L., Tan C. L* "Motion deblurring for optical character recognition." In Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on, pp. 389-393. IEEE, 2005.

[11] *Aizenberg I., Paliy D.V., Zurada J.M. and Astola J.T.* "Blur identification by multilayer neural network based on multivalued neurons." IEEE Transactions on Neural Networks 19, no. 5 (2008): 883-898.

[12] *Serrat J., López A., Lloret D.* "On ridges and valleys [image registration]." In Pattern Recognition, 2000. Proceedings. 15th International Conference on, vol. 4, pp. 59-66. IEEE, 2000.

[13] *Tikhonov A.N. and Arsenin V.Y.* Methods for solving ill-posed problems. John Wiley and Sons, Inc, 1977.

[14] *Beck A., Teboulle M.* "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems." IEEE Transactions on Image Processing 18, no. 11 (2009): 2419-2434.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Диагностика меланомы с использованием нейронной сети Inception v3

Гаврилов Д.А., Мелерзанов А.В, Щелкунов Н.Н., Закиров Э.И.
gavrilou@gmail.com, m83071@gmail.com, shchelkunov.nn@mipt.ru, bonlimezak@gmail.com
Московский Физико-Технический институт (государственный университет)
Кафедра радиоэлектроники и прикладной информатики
Факультет биомедицинской физики

Меланома — один из самых опасных видов рака. Точность визуальной диагностики данного заболевания напрямую зависит от квалификации и специализации врача. Современное развитие технологий в области обработки изображений и машинного обучения позволяет создать системы на основе искусственных нейронных сверточных сетей, превосходящие человека в задачах классификации объектов, в том числе в случае диагностики злокачественных новообразований кожи. В настоящей работе представлен алгоритм ранней диагностики меланомы, основанный на использовании сверточной нейронной сети с архитектурой Inception.v.3. Предобученная нейронная сеть настраивалась на классификацию кожных заболеваний путем удаления верхних классифицирующих слоев и добавления новых нейронов для определения кожных заболеваний. В результате обеспечивается классификация доброкачественных и злокачественных новообразований кожи с точностью не менее 91% при просмотре дерматоскопических изображений.

Ключевые слова: *глубокие сверточные нейросети, диагностика меланомы, техническое зрение, телемедицина.*

Melanoma diagnosis by means transfer learning method for Inception v3 CNN

Gavrilov D.A., Melerzanov A.V., Shchelkunov N.N., Zakirov E.I.
gavrilou@gmail.com, m83071@gmail.com, shchelkunov.nn@mipt.ru, bonlimezak@gmail.com
Moscow Institute of Physics and Technology (State University)

Melanoma is one of the most virulent lesions of human's skin. The visual diagnosis accuracy of melanoma directly depends on the doctor's qualification and specialization. State-of-the-art solutions in the field of image processing and machine learning allows to create intelligent systems based on artificial convolutional neural network exceeding human's rates in the field of object classification, including the case of malignant skin lesions. This paper presents an algorithm for the early melanoma diagnosis based on artificial deep convolutional neural networks. The algorithm proposed allows to reach the classification accuracy of melanoma at least 91%.

Key words: *deep convolutional neural networks, melanoma diagnosis, computer vision, telemedicine.*

1. Ведение

Меланома — злокачественная опухоль преимущественно кожной локализации, являющаяся на сегодняшний день одним из самых опасных видов рака. По данным Всемирной организации здравоохранения (ВОЗ) частота возникновения этого заболевания неуклонно растет из года в год [12]. Меланома вносит наибольший вклад в смертность от всех типов рака кожи и является одним из наиболее быстро количественно растущих типов рака среди всех видов онкологических заболеваний. Высокий уровень смертности при заболевании меланомой, как правило, объясняется поздней диагностикой заболевания. В то же время, при обнаружении на ранних стадиях, хирургическое лечение дает хорошие прогностические результаты и может обеспечить практически 100 % -ю выживаемость [13].

В борьбе со смертностью от меланомы на первый план выходит вторичная профилактика или превентивная медицина. Своевременные осмотры, терапия и контроль распространения рецидивов. По данным некоторых авторов, своевременное обнаружение и начало лечения меланомы при регулярном самостоятельном или дерматологическом обследовании снижает смертность пациентов на 63% [1]. Особенно важно, что процент обнаружения меланомы, при использовании современных подходов и методик, зачастую незначительно отличается в случае самостоятельного обследования пациента (33%) от диагностики при осмотре профессиональным врачом в рамках общего обследования (36%) [5]. Таким образом, ответственный подход к самообследованию на предмет появления атипичных пигментных кожных образований может обеспечить высокие результаты первичной диагностики и, в случае необходимости, позволит пациенту своевременно обратиться за медицинской помощью.

Первичная диагностика меланомы, как правило, выполняется врачом визуально, в дальнейшем диагноз уточняется с помощью биопсии и морфологических исследований.

Современное развитие технологий в области обработки изображений и машинного обучения позволяет создавать системы на основе искусственных нейронных сверточных сетей, превосходящие человека в задачах

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

классификации объектов, в том числе при диагностике злокачественных новообразований кожи [6, 10]. Существуют алгоритмы автоматизированного компьютерного анализа дерматологических изображений, позволяющие определять границу, яркость, диаметр и симметрию пигментации [4] и, тем самым, обеспечивающие помощь врачам для повышения точности диагностики [3]. В последнее время активно ведется разработка систем, основанных на технологиях искусственного интеллекта, в частности сверточных нейросетях, и обеспечивающих возможность самостоятельного обнаружения злокачественных образований на ранней стадии [2]. Исследования показывают высокий потенциал применения нейросетевых технологий для диагностики новообразований кожи [6].

В настоящей работе представлена система ранней диагностики меланомы на основе искусственных глубоких сверточных нейронных сетей, позволяющая автоматически определять с точностью не менее 90% наличие злокачественных поражений кожи с учетом получения фотографий вне лабораторных условий: при различных вариантах освещения, угла съемки и разрешения средства для фотосъемки. При этом точность определения кожной меланомы сравнима с точностью диагностики данного заболевания врачами.

2. Методика

Понятие нейронной сети – математической модели, функционирующей в некоторой степени схожим образом с клетками мозга человека, известно с середины 20 века. Нейронные сети уже находили свое применение при решении различных задач, имели своих сторонников и противников, однако в последние 5 лет данная технология совершила революцию в области задач классификации изображений, превысив не только результаты классических алгоритмов, но и качество классификации, демонстрируемое человеком.

Идея использования компьютерного зрения для решения задачи определения рака кожи возникла достаточно давно. Технологии компьютерного или технического зрения представляют собой комплекс методов, позволяющих производить обнаружение и классификацию различных объектов. Научные исследования, целью которых было улучшение дифференциальной диагностики меланомы с помощью применения компьютерных технологий и экспертных систем, проводятся во многих крупных центрах Германии, Австрии и других странах начиная с 1987 г. по предложению N. Cascinelli, президента Меланомной Программы ВОЗ (WHO Melanoma Program), в настоящее время преобразованной в World Melanoma Society. Продолжительное время результаты не были достаточно точными для применения на практике.

Основная проблема построения глубоких сверточных нейронных сетей заключается в том, что на данный момент не существует достаточной публичной выборки изображений, которые возможно было бы использовать для обучения системы и настройки моделей. Одним из самых крупных архивов изображений кожных новообразований является International Skin Imaging Collaboration (ISIC) [14]. Однако, представленные в архиве наборы данных содержат недостаточное количество требуемых изображений, представленные изображения получены при проведении реальных клинических исследований, поэтому могут быть частично зашумлены или содержать различные помехи (Рис. 1).

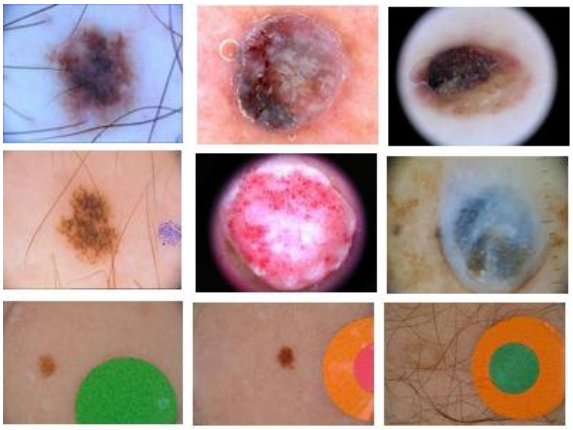


Рис. 1. Примеры помех на изображениях

Авторами были разработаны алгоритмы, позволяющие проводить анализ искаженных помехами изображений в условиях ограниченной выборки, вариабельности освещения и разброса в условиях съемки.

Для решения поставленной задачи использовался подход, известный как «перенос обучения» или transfer learning [11]. В данном случае не предусматривалось решение задачи обучения нейросети классификации кожных заболеваний с нуля.

В качестве модели была выбрана нейронная сеть с архитектурой Inception.v.3 (Рис. 2) [8, 9]. В сети Inception.v.3 использована техника пакетной нормализации Batch normalization. В архитектуре каждый блок из двух свёрток включает слой Batch Normalization между ними. Batch Normalization позволяет существенно ускорить процесс обучения нейронной сети за счет использования более высокого значения параметра скорости обучения (learning rate), а также снизить переобучение.

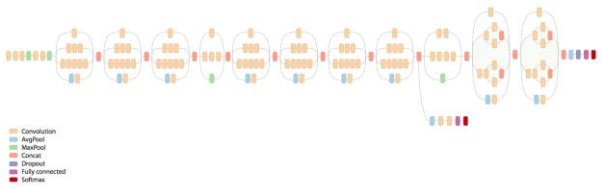


Рис. 2. Архитектура нейронной сети «Inception v.3» [17]

Inception.v.3 демонстрирует высокое качество классификации различных изображений в соревновании ILSVRC (ImageNet Large Scale Visual Recognition Challenge). Выбранная предобученная нейронная сеть обладает возможностью переучивания для последующего ее использования в качестве компонента более крупной сети. С помощью сети Inception.v.3 осуществлялась подготовка модели классификации изображений на данных архива ImageNet Challenge [7, 15].

Предобученная нейронная сеть настраивалась на классификацию кожных заболеваний путем удаления верхних классифицирующих слоев и добавления новых нейронов для определения кожных заболеваний. Использование предобученной на изображениях ImageNet сети обусловлено отсутствием достаточно большой обучающей выборки. Использование «переноса обучения» позволяет сократить объем обучающей выборки и время обучения сети. Обучение с нуля на недостаточно большой выборке могло привести к переобучению и невозможности дальнейшей качественной классификации на новых данных.

Впоследствии производилась перенастройка

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

полученной модели на обучающей выборке, включающей 10000 фотографий кожных образований, для обеспечения нейронной сети возможности различать тип кожного образования. С целью увеличения объема обучающей выборки количество исходных изображений было искусственно расширено до 1000000 путем аугментации данных с помощью различных искажений. Вносимые искажения включали в себя: случайные повороты в пределах 360^0, зеркальные отражения, вырезание лишь части изображения, растяжение и сжатие по осям абсцисс и ординат с коэффициентом 0.1, масштабирование с коэффициентом 0.1, а также варианты изменения освещения.

Зачастую в выборке присутствуют изображения с интерференцией, кроме того искусственные искажения вносимые при аугментации исходных изображений включают, в том числе и интерференцию. Факторы интерференции и прочие искажения оказывают влияние на точность распознавания, однако, данное влияние не всегда имеет отрицательный характер. Следует обратить внимание, что точность работы сети возрастает с ростом обучающей выборки. Таким образом, сеть учится распознавать и искаженные изображения.

Неравномерность яркости изображений, разные расстояния камеры от объекта при съемке, в также различие размеров объектов на изображениях учитывались в обучающей выборке методом наращивания возможных вариаций.

Для контроля обучения использовалась тестовая выборка, изображения из которой не использовались при обучении алгоритма.

В общей сложности, для повышения качества распознания и классификации, было обучено 5 сетей с одинаковой архитектурой, но имеющих разные веса. В дальнейшем все обученные нейросети были объединены в ансамбль моделей, обеспечивающих принятие решений мажоритарным голосованием.

Исходная выбранная модель Inceptionv.v.3 работала исключительно с изображениями разрешением не более 300x300 пикселей, что существенно меньше обычного размера дерматологических снимков. С учетом того, что для достоверной диагностики заболеваний крайне важными являются мелкие детали изображения, неразличимые при таком разрешении, в существующую модель были добавлены сверточные слои, позволяющие осуществлять классификацию на изображениях большего разрешения.

Доработанная модель позволяет проводить исследования дерматоскопических изображений разрешением до 700x700 пикселей.

В итоговой тестовой модели на вход алгоритма поступают изображения новообразований кожи, выполненные с помощью мобильного телефона в условиях дневного света. На выходе система выдает вероятность соответствия полученного изображения одному из 4 классов: меланома, невус, себорейный кератоз или другое заболевание.

3. Результаты

Основными критериями качества результата работы сети являлись площадь под кривой точности-полноты и итоговая вероятность обнаружения. В результате работы обеспечивается точность распознавания меланомы кожи более 91% с AUC-ROC 0.96 (ROC-кривая — график, позволяющий оценить качество бинарной классификации), что сопоставимо с результатами диагностики высококвалифицированных врачей. График AUC-ROC распознавания меланомы-невуса для ансамбля моделей представлен на Рис. 3.

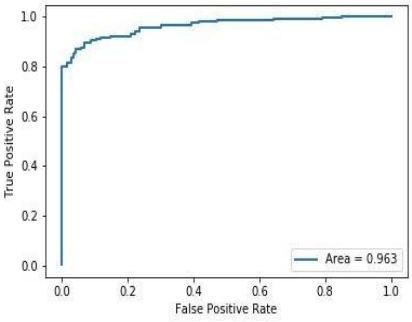


Рис. 3. График AUC-ROC распознавания меланомы-невуса для ансамбля моделей

Многообразие кожных новообразований, сложность их строения, сходство клинической картины различных форм поражений кожи обуславливает трудности визуального осмотра и диагностики даже у специалистов. С учетом того, что точность визуальной диагностики меланомы напрямую зависит от квалификации и специализации врача, а также от частоты встречаемости заболевания в его повседневной практике, в качестве дополнительной задачи разработанная модель позволяет отделять злокачественные новообразования от доброкачественных. Так, например, точность определения такого доброкачественного образования как себорейный кератозис по дерматологической фотографии составляет 97% и AUC-ROC 0.99.

Таким образом, предлагаемый алгоритм позволяет с уверенностью отличить меланому от себорейного кератозиса, а также меланому от доброкачественного невуса. Пример классификации представлен на Рис. 4.

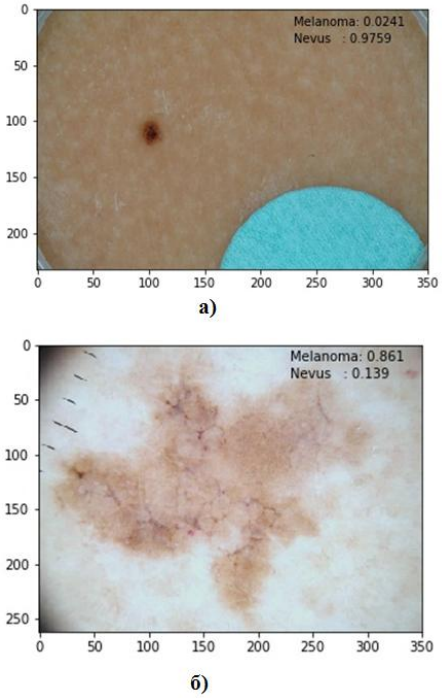


Рис. 4. Пример классификации а) невуса, б) меланомы

4. Заключение

Проведенная работа позволила создать сверточную нейронную сеть глубокого обучения для проведения автоматизированной диагностики кожных новообразований.

В ходе работ сформирована обучающая выборка,

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

включающая реальные фотоснимки кожных новообразований и изображения, полученные из исходных с помощью различных искажений. Уникальные алгоритмы аугментации вкупе с функциями потерь, являются основной авторской разработкой, позволяющей достичь высокого качества классификации кожных заболеваний при ограниченной обучающей выборке.

Разработанная модель позволяет проводить качественную диагностику меланомы кожи с точностью не менее 91%, что сравнимо с возможностями диагностики высококвалифицированного врача-дерматолога. Использование интеллектуальных систем такого типа для идентификации кожных заболеваний окажет существенную поддержку в постановке диагноза как дерматологам, так и врачам общей практики.

Система доступна в тестовом варианте по адресу **skincheckup.online.** Общедоступность данного проекта позволит широким массам населения осуществлять предварительную самодиагностику с помощью собственных фотоснимков.

Дальнейшее развитие предложенной модели может осуществляться в направлении создания встроенных систем, WEB- и мобильных приложений.

Разработка имеет большие перспективы применения в составе телемедицинских систем комплексного анализа здоровья человека.

Авторы выражают большой интерес к сотрудничеству в области создания прикладных систем технического зрения.

5. Литература

1. Berwick M., Begg C.B., Fine J.A., Roush G.C., Barnhill R.L. Screening for cutaneous melanoma by skin self-examination // J Natl Cancer Inst. 1996. № 1 (88). С. 17–23.
2. Esteva A. [и др.]. Dermatologist-level classification of skin cancer with deep neural networks // Nature. 2017. № 7639 (542). С. 115–118.
3. Fink C., Jaeger C., Jaeger K., Haenssle H.A. Diagnostic performance of the MelaFind device in a real-life clinical setting // J Dtsch Dermatol Ges. 2017. № 4 (15). С. 414–419.
4. Gareau D.S., da Rosa J.C., Yagerman S., Carucci J.A., Gulati N., DeFazio J. L., Suárez- Fariñas M., Marghoob A., Krueger J.G. Digital imaging biomarkers feed machine learning for melanoma screening // Experimental Dermatology. 2017. (26). С. 615–618.
5. Giorgi V., Grazzini M., Rossari S., Gori A., Papi F., Scarfi F. De Is skin self-examination for cutaneous melanoma detection still adequate? A retrospective study // Dermatology. 2012. № 1 (225). С. 1–6.
6. Haenssle H.A., Fink C., Schneiderbauer R., Toberer F., Buhl T., Blum A., KallooA A., Hassen A. B. H., Thomas L., Enk A., Uhlmann L. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists // Annals of Oncology. 2018. № 8 (29). С. 1836–1842.
7. Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A.C., Fei-Fei L. ImageNet Large Scale Visual Recognition Challenge // International Journal of Computer Vision. 2015. № 3 (115). С. 211–252.
8. Shlens J. Train your own image classifier with Inception in TensorFlow [Электронный ресурс]. URL: https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html (дата обращения: 24.01.2018).
9. Szegedy C. [и др.]. Rethinking the Inception Architecture for Computer Vision [Электронный ресурс]. URL: http://arxiv.org/abs/1512.00567 (дата обращения: 25.01.2018).
10. Talha Khan B. Machine learning model in melanoma // The Lancet Oncology. 2018. № 7 (19). С. 340.
11. Torrey L., Shavlik J. Transfer Learning / L. Torrey, J. Shavlik, IGI Global, 2009. 242-264 с.
12. Гидранович А.В., Луд Н.Г., Шляхтунов Е.А., Луд Л.Н., Жуковец А.Г., Прокошин А.В., Кожар В.Л. Перспектива использования экспертных систем в ранней диагностике меланомы кожи // Вестник ВГМУ. 2015. № 2 (14). С. 46–55.
13. Козлов С.В., Неретин Е.Ю., Куколкина В.В. Перспективы применения «машинного зрения» в цифровой дерматоскопии при подозрении на меланому кожи // Саратовский научно-медицинский журнал. 2014. № 2 (10). С. 6–10.
14. ISIS Archiv // Kitware, Inc. [Электронный ресурс]. URL: https://isic-archive.com/ (дата обращения: 24.01.2018).
15. ImageNet Large Scale Visual Recognition Competition (ILSVRC) [Электронный ресурс]. URL: http://www.image-net.org/challenges/LSVRC/ (дата обращения: 13.07.2018).

Об авторах

Гаврилов Дмитрий Александрович, к.т.н., доцент кафедры радиоэлектроники и прикладной информатики, руководитель лаборатории цифровых систем специального назначения, Московский физико-технический институт (государственный университет), тел. 8-925-642-22-71, gavrilou@gmail.com

Мелерзанов Александр Викторович, к.м.н., Декан факультета биологической и медицинской физики МФТИ, Московский физико-технический институт (государственный университет), m83071@gmail.com

Щелкунов Николай Николаевич, к.т.н., зам. зав. кафедры радиоэлектроники и прикладной информатики МФТИ, Московский физико-технический институт (государственный университет), тел. 8-916-654-42-54, shchelkunov.nn@mipt.ru

Закиров Эмиль Ильдарович, студент, стажер лаборатории цифровых систем специального назначения МФТИ, Московский физико-технический институт (государственный университет), bonlimezak@gmail.com

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Алгоритм нелокального среднего основанный на модифицированном индексе структурного сходства

Довганич А.А.[1], Крылов А.С.[1], Юрин Д.В.[1]
a.dovganich@yandex.ru|kryl@cs.msu.ru|yurin@cs.msu.ru

[1]Факультет Вычислительной математики и кибернетики

МГУ имени М.В.Ломоносова, Москва, Россия

Предлагается новый алгоритм шумоподавления на изображениях. Он является модификацией алгоритма нелокального среднего с использованием метрики основанной на индексе структурного сходства(SSIM). Производится анализ этой весовой функции с точки зрения статистических характеристик изображений, физиологии восприятия изображений и структуры наблюдаемой сцены. Индекс SSIM разбирается на отдельные составляющие и для каждой компоненты задается физически обоснованная весовая функцию. Результаты модифицированного метода сравниваются с алгоритмом нелокального среднего, где веса вычисляются по метрике L2 и SSIM.

Ключевые слова: *фильтрация изображений, подавление шума, нелокальное среднее, SSIM .*

Nonlocal mean algorithm based on a modified structural similarity index

A.A. Dovganich[1], A.S. Krylov[1], D.V. Yurin[1]
a.dovganich@yandex.ru|kryl@cs.msu.ru|yurin@cs.msu.ru
[1]Laboratory of Mathematical Method of Image Processing
Faculty of Computational Mathematics and Cybernetics Lomonosov Moscow State University, Moscow, Russia

In this paper, we propose new image denoising algorithm. It is nonlocal mean algorithm based on the modified structural similarity index (SSIM). The weight function is analysed with the statistical characteristics of images, human vision perception and the structure of the scene. The SSIM index is divided into the separate components and a physically valid weight function is assigned to each component. The results of the modified method are compared with the standard nonlocal mean algorithm with the weights calculated using L2 metric and non-modified SSIM.

Keywords: *image filtering, image denoising, nonlocal mean, SSIM.*

1. Введение

Задача подавления шума на изображениях является одной из самых старых но, по-прежнему, актуальных [1,2,3,4]. Одновременно с ростом качества аппаратных средств съемки наблюдается не менее существенный рост и запросов по съемке в гораздо более сложных условиях. Можно привести следующие области применения, где необходимость подавления шума будет и в дальнейшем сохранять актуальность:

1) Съемка бытовыми камерами и смартфонами в условиях плохой освещенности или искусственного освещения. Запросы постоянно растут – сумерки, приглушенное освещение, неудачный спектр освещения.

2) Аэрокосмическая съемка Земли - растет желаемое пространственное разрешение. Сейчас вполне типичным является разрешение свыше 1 метра, тогда как типичная высота космического аппарата составляет около 1000 км, т.е. 1 пиксель видим под ничтожным углом 10^{-6} радиан. Более того, т.к. космический аппарат движется со скоростью около 7000 м/с, то время экспозиции при съемке с таким разрешением не должно превышать 1/7000 секунды, что обуславливает крайне низкий световой поток.

3) В медицине изображения получаются, как правило, с помощью излучений вредных для здоровья – рентген, гамма излучение, микроволновое излучение. Поэтому задача максимального снижения дозы облучения ведет опять к съемке в условиях слабого освещения и высоких шумов.

В принципе, подавляющее большинство алгоритмов представляет собой какой-либо вариант усреднения некоторого подмножества пикселей в некоторой окрестности рассматриваемой точки. Классический алгоритм среднего по окрестности – это усреднение всех пикселей из прямоугольной или круглой окрестности точки [2,3]. Линейные алгоритмы усреднения – это усреднение пикселей локальной окрестности с некоторым весом, который зависит только от расстояния от этих пикселей до рассматриваемой точки. Весовая функция является ядром оператора свертки [2,3]. Алгоритмы на основе анизотропной и нелинейной диффузии [5,6] можно рассматривать как аналог свёрточных линейных методов, но с ядром свертки, меняющимся по плоскости изображения, например, вблизи границ объектов, вытягивающимся вдоль границы. При этом, размытие происходит вдоль границ, а не поперек, эффективный размер ядра меняется в зависимости от сложности изображения. В однородных областях обеспечивается большее размытие и поэтому сглаживание шума; в информационно насыщенных областях степень размытия снижается, поэтому сохраняется больше как шума, так и деталей изображения.

В ранговых алгоритмах [2,7,8], которые относятся уже к нелинейным фильтрам, помимо отбора по расстоянию (форма окрестности), производится отбор и по яркости пикселей, в усреднении участвуют только пиксели, удовлетворяющие некоторым ограничениям, например, K ближайших по яркости к рассматриваемому пикселю (среднее по KNV окрестности), или пиксели окрестности, яркость которого отклоняется не более чем на ε от яркости центрального пикселя (среднее по ε_V окрестности). В отличие от линейных алгоритмов, ранговые алгоритмы уже не обязательно приводят к размытию границ объектов и линий. Так, для упомянутых двух алгоритмов характерно снижение контраста границы объектов при сохранении резкости профиля перепада яркости. Аналогом алгоритма усреднения по ε_V окрестности является алгоритм билатеральной фильтрации [9]. Он отличается тем, что весовая функция в зависимости от расстояния от центра окрестности и весовая функция в зависимости от расстояния от ведущего пикселя - Гауссовы, а не прямоугольные ступеньки.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Следующим шагом развития подобных подходов являются методы самоподобия, как алгоритм нелокального среднего [10] (NLM) и ряд его разновидностей [11,12,13]. В усреднении участвуют только точки, у которых похожи малые локальные окрестности. Т.е. отбор ведется не только по пространственному признаку или яркости, но и по паттерну, текстуре вокруг точки. Фактически используется предположения самоподобия (self-similarity) и усредняются похожие фрагменты. Дальнейшим развитием этого направления является алгоритм BM3D [14], который после отбора похожих фрагментов собирает их в трехмерную структуру и подвергает фильтрации Винера. Этот алгоритм на сегодняшний день считается наилучшим по качеству подавления шума, внося минимальные повреждения в изображение.

Для случая аддитивного Гауссового шума, за достаточно короткий период, одними из наиболее популярных методов подавления шума стали методы, использующие глубокое обучение. Они показывают достаточно конкурентные результаты в сравнении с методами усреднения [15,16]. Однако необходимо отметить, что методы, основанные на самоподобии, более устойчивы к изменению типа шума и его однородности на изображении. В связи с этим необходимо отметить работу [4], в которой предлагают метод итерационного подавления шума как с использованием CNN, так и методов самоподобия.

2. Алгоритм нелокального среднего

Рассмотрим процедуру отбора похожих блоков в алгоритме нелокального среднего [10]:

$$I_{NLM}(x,y)=\frac{1}{W(x,y)}\sum_{x',y'\in\Omega(x,y)} w(x,y,x',y')I(x',y') \quad \text{где} \quad W(x,y)=\sum_{x',y'\in\Omega(x,y)} w(x,y,x',y') \tag{1}$$

Размер окрестности Ω здесь может быть произвольным, в том числе и всем изображением, откуда и идет название метода ***нелокальное*** среднее. Здесь $w(x,y,x',y')$ - весовая функция, зависящая от схожести блоков с центрами в точках (x,y) и (x',y'), определенная как [10]:

$$w(x,y,x',y')=\exp\left(-\frac{\sum_{\xi,\eta\in\Omega_{patch}}\left(I(x+\xi,y+\eta)-I(x'+\xi,y'+\eta)\right)^2}{2\rho^2}\right) \tag{2}$$

Здесь окрестность Ω_{patch} - это фрагмент изображения вокруг точки, по которому и выполняется сравнение похожести двух точек, т.е. анализ схожести их текстур или блоков. Числитель в экспоненте (2) является, по сути, нормой L_2 разности двух векторов-параметров Ω_{patch}, построенных как построчные выборки пикселей этих окрестностей.

Из анализа формул (1)-(2) следует, что они вычисляют немного не то, что декларировалось: в силу использования нормы L_2, блоки, содержащие совсем разные картинки, и блоки, содержащую одинаковую картинку, отличающуюся только яркостью или контрастом одинаково различны. При этом, мера отличия определяется только суммой квадратов разностей соответствующих пикселей и не зависит от графического образа картинки или структурного сходства блоков, что, по-видимому, не вполне то, что было бы желательно.

3. Структурный индекс подобия

Естественный способ преодолеть подобный недостаток – это подобрать более адекватную норму. В работе [17] предложен структурный индекс подобия SSIM, его достоинства по сравнению с L_2 рассмотрены в [18]. Он представляет собой произведение трех компонент:

$$SSIM(x,y,x',y')=l(x,y,x',y')c(x,y,x',y')s(x,y,x',y') \tag{3}$$

где,

$$l(x,y,x',y')=\frac{2\mu(x,y)\mu(x',y')+C_1}{\mu^2(x,y)+\mu^2(x',y')+C_1} \tag{4}$$

- функция подобия яркости,

$$c(x,y,x',y')=\frac{2\sigma(x,y)\sigma(x',y')+C_2}{\sigma^2(x,y)+\sigma^2(x',y')+C_2} \tag{5}$$

-функция подобия контраста,

$$s(x,y,x',y')=\frac{\Gamma(x,y,x',y')+C_3}{\sigma(x,y)\sigma(x',y')+C_3} \tag{6}$$

- функция подобия структуры.

Константы C_1, C_2, C_3 введены исключительно для предотвращения деления на нуль и, в принципе, не обязательны, так как при численных расчетах подобные предосторожности могут быть выполнены и иными средствами исходя из желаемого поведения соответствующих компонент при предельном стремлении к нулю величин из знаменателя. В формулах (4)-(6) использованы следующие величины $\mu(x,y)=\langle I(x,y)\rangle$ - средняя яркость, $\sigma^2(x,y)=\langle I^2(x,y)\rangle-\langle I(x,y)\rangle^2$ - стандартное отклонение (корень из дисперсии), $\Gamma(x,y,x',y')=\langle I(x,y)I(x',y')\rangle-\langle I(x,y)\rangle\langle I(x',y')\rangle$ - корреляционная функция. Усреднение $\langle\cdot\rangle$ вычисляется по блоку Ω_{patch} с центром в точке (x,y). Выражение (6) зависит от корреляционной функции между двумя блоками с центрами в (x,y) и (x',y'), и представляет собой коэффициент корреляции между этими блоками. Отсюда видно, что:

$$\begin{aligned} 0\le l(x,y,x',y')\le 1\\ 0\le c(x,y,x',y')\le 1\\ -1\le s(x,y,x',y')\le 1\end{aligned} \tag{7}$$

В [19] предложена модификация алгоритма нелокального среднего, использующая в качестве весовой функции в (1) индекс структурного сходства:

$$w(x,y,x',y')=SSIM(x,y,x',y') \tag{8}$$

Кроме того, формула (1) также претерпела изменения. Поскольку теперь каждый блок характеризуется раздельно яркостью, контрастом и структурой, логичным представляется перед выполнением усреднения центральных точек привести яркости и контрасты соответствующих блоков к целевому блоку. Поэтому

$$I_{SSIM}(x,y)=\frac{1}{W(x,y)}\sum_{x',y'\in\Omega(x,y)} w(x,y,x',y')J(x',y') \tag{9}$$

где

$$J(x',y')=\frac{\sigma(x,y)}{\sigma(x',y')}\left(I(x',y')-\mu(x',y')\right)+\mu(x,y) \tag{10}$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

4. Предлагаемый метод

Предложенная в [19] модификация метода нелокального среднего на наш взгляд представляется незаконченной и недостаточной. Согласно (3)-(8) высокие веса соответствуют блокам, которые имеют ***И*** одинаковую структуру (паттерн), ***И*** одинаковую яркость, ***И*** одинаковый контраст. В реальности же, очень часто фрагменты объектов реального мира имеют одинаковую структуру, но наблюдаются в различных условиях освещенности, т.е. отличаются яркостью и контрастом. Более того, человеческое зрение к таким различиям часто не чувствительно и как бы автоматически их игнорирует. Действительно, с точки зрения компьютерной обработки и вычисления величины перепадов яркости, часто наиболее значимыми на изображении являются границы света и тени, в то время как человек обычно их игнорирует и акцентируется на структуре сцены и объектов её составляющих, а не условиях освещения. В этой связи наиболее важным представляется третий сомножитель (6), а первые два, (4) и (5) - существенно менее важными. Кроме того, (6) а, следовательно, и (3) и (8), могут принимать отрицательные значения. Теоретически это могло бы позволить в процессе усреднения использовать не только похожие паттерны, но и их негативы. Однако, реальная сцена, состоящая из изображений каких либо объектов и их негативов представляется чем-то весьма экзотичным. Как минимум необходим специальный анализ причин, почему в конкретном случае реальные объекты выглядят как позитив и негатив. Наличие же отрицательных весов может сделать схему фильтрации крайне нестабильной из-за возможности нулевого знаменателя в (1). Поэтому мы предлагаем полагать третий сомножитель равным нулю при отрицательном значении коэффициента корреляции.

Рассмотрим теперь множители, зависящие от яркости (4) и контраста (5). Казалось бы, при наличии процедуры нормализации яркости (10), эти множители можно было бы просто игнорировать. Однако, в реальности изображение содержит шум и, даже при идеальных условиях съемки, есть шум обусловленный оцифровкой на определенное число градаций яркости. Поэтому, при попытке использовать менее контрастный блок для вычисления средней яркости для центральной точки более ярких блоков, согласно формуле (10) яркость этого блока будет растягиваться, что приведет к увеличению шума в такое же число раз. Одновременно точность данных в такое же число раз снизится, т.к. данные заданы с определенным шагом по яркости, который также в соответствующее число раз возрастет. Таким образом, представляется очевидными, что в формуле (10) соотношение $\sigma(x,y)/\sigma(x',y')$ следует контролировать и в случае его слишком большого значения, резко уменьшать вес для таких точек. Наоборот, малые значения этого коэффициента весьма желательны, так как ведут ко вкладу данных с ***пониженным шумом и повышенной точностью оцифровки***.

Зависимость от яркости блоков менее очевидна. Так как фоторегистрирующая аппаратура обычно работает в линейном режиме, то постоянная добавка к яркости вроде бы не должна ни на что влиять. Тем не менее возможны некоторые косвенные эффекты. Например, при отношении $\sigma(x,y)/\sigma(x',y')$ меньшем единицы, в (10) более точно учитывается среднее значение яркости блока, который приводится к целевому. С другой стороны, изменение средней яркости на единицу при высокой средней яркости представляется менее достоверной, измеренной с меньшей точностью, чем то же изменение в условиях малой средней яркости – измерение изменений в разы представляется более достоверным, нежели чем на доли процента. Поэтому для яркостной компоненты предлагается просто ввести ограничение по допустимому различию яркости блоков $\mu(x,y)/\mu(x',y')$ симметричным образом в обе стороны.

Суммируя выше сказанное, предлагается следующий вид весовой функции для модифицированного алгоритма нелокального среднего

$$\begin{aligned} w(x,y,x',y') = \\ = \theta\left(T_1 \cdot \mu(x,y)\mu(x',y') - \left(\mu^2(x,y) + \mu^2(x',y')\right)\right) \\ \cdot \theta\left(T_2 \cdot \sigma(x',y') - \sigma(x,y)\right) \\ \cdot \theta\left(s(x,y,x',y')\right) \cdot f\left(s(x,y,x',y')\right) \end{aligned} \quad (11)$$

которая подставляется в (9) и (10). Здесь $\theta(\cdot)$ - ступенчатая единичная функция Хэвисайда. Функция $f(\cdot)$ служит для регулирования крутизны учета влияния слабо коррелированных блоков, в настоящей работе эта функция полагается линейной, $f(x)=x$. Пороги T_1 и T_2 определяют допустимые пределы по отклонению яркости (симметрично) и контраста (несимметрично) и задаются из априорных соображений о допустимом повышении шума. Например, для пятикратного разброса яркостей T_1=2.6, и допустимого растяжения контраста вдвое T_2=2. Возможно, более целесообразным было бы использование менее жестких, плавных регулировок, нежели функция Хэвисайда, однако для простоты и ясности изложения в настоящей работе мы ограничились таким простейшим случаем, который наиболее наглядно демонстрирует свойства алгоритма.

Практически, внесённые изменения должны привести к следующим отличиям предлагаемого алгоритма от классического нелокального среднего [10] и основанного на SSIM [19].

1) В указанных алгоритмах в усреднении эффективно участвуют только точки блоков, похожих как по структуре, так и по яркости и контрасту. В предложенном алгоритме – от блоков с близкой структурой, при широком разбросе яркостей и контрастов. Усредняются все похожие фрагменты сцены, независимо от того на свету они или затенены;
2) Несимметричное ограничение на допустимый разброс контрастов ведет к дополнительному снижению шума;
3) Отсутствие отрицательных весов в отличие от [19].

Заметим, что проблемы связанной с порчей блоков из-за зашкаливания яркости как в сторону засвета, так и обрезки темных пикселей не возникает, так как в этом случае будет уменьшаться множитель, зависящий от коэффициента корреляции.

5. Результаты

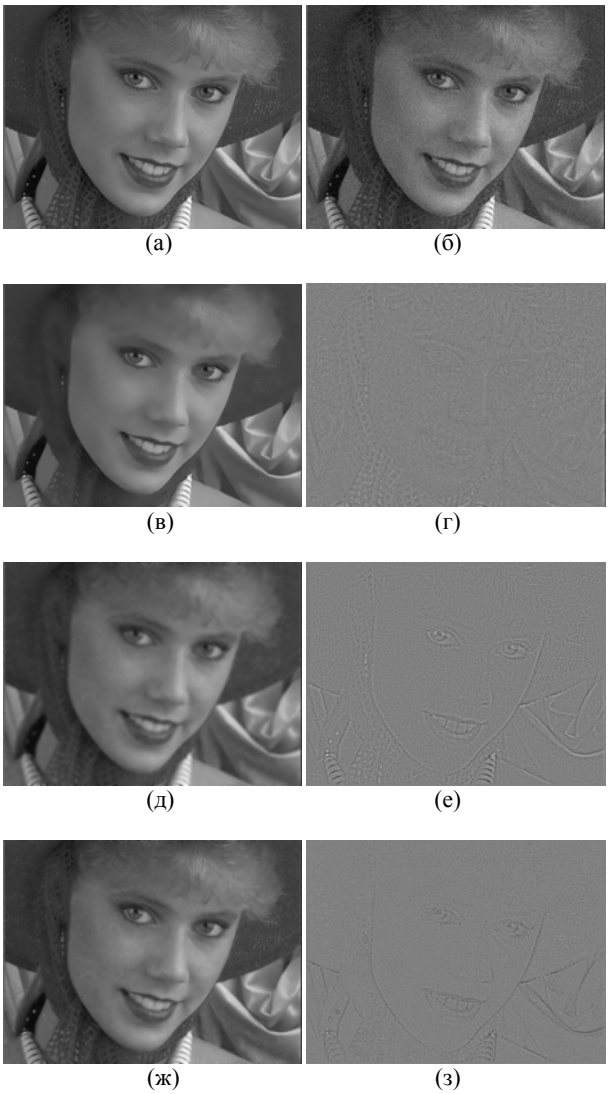
Тестирование производилось на изображениях базы TID2013[19] с различными типами шума. Результаты для нескольких изображений приведены в таблице 1. Значение MOS(Mean Opinion Score) для зашумленных изображений было следующее: I04/Additive Gaussian Noise - 4.27, I22/ High Frequency Noise - 4.18, I23/Spatially Correlated Noise - 3.26. Алгоритм нелокального среднего на основе L_2 обозначается как NLM-L2; алгоритм, использующий индекс структурной схожести, обозначается как NLM-SSIM, предлагаемый метод обозначен как NLM-MSSIM. Блоки ищутся в окрестности с радиусом 7 пикселей, радиус каждого блока равен 3, то есть с учетом центрального пикселя диаметр блоков 3+1+3=7 пикселей, а размер7x7 пикселей. При поиске похожих блоков их центры ищутся по окну 15x15 пикселей. Параметры C_1, C_2, C_3 заданы следующим образом: $C_1 = (k_1 * L)$, $C_2 = (k_2 * L)$, $C_3 = C_2 / 2$, где k_1=0.01, k_2=0.03, L - динамический диапазон пикселей. Для NLM-L2 использовался параметр ρ = 8.3. Он выбирался исходя из

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

параметров шума на изображении (стандартного отклонения). #

Изображение и тип шума	I04/Additive Gaussian Noise	I22/ High Frequency Noise	I23/Spatially Correlated Noise
PSNR			
Noisy image	24.4104	24.3689	24.6671
NLM-L2	31.0202	28.1615	34.5181
NLM-SSIM	30.4254	29.8546	31.1828
NLM-MSSIM	32.3352	30.3894	34.6982
SSIM			
Noisy image	0.7691	0.6164	0.5892
NLM-L2	0.9212	0.8897	0.8893
NLM-SSIM	0.9446	0.9245	0.8964
NLM-MSSIM	0.9737	0.9358	0.9156

Таблица 1. Результаты сравнения PSNR и SSIM.



(а) (б)

(в) (г)

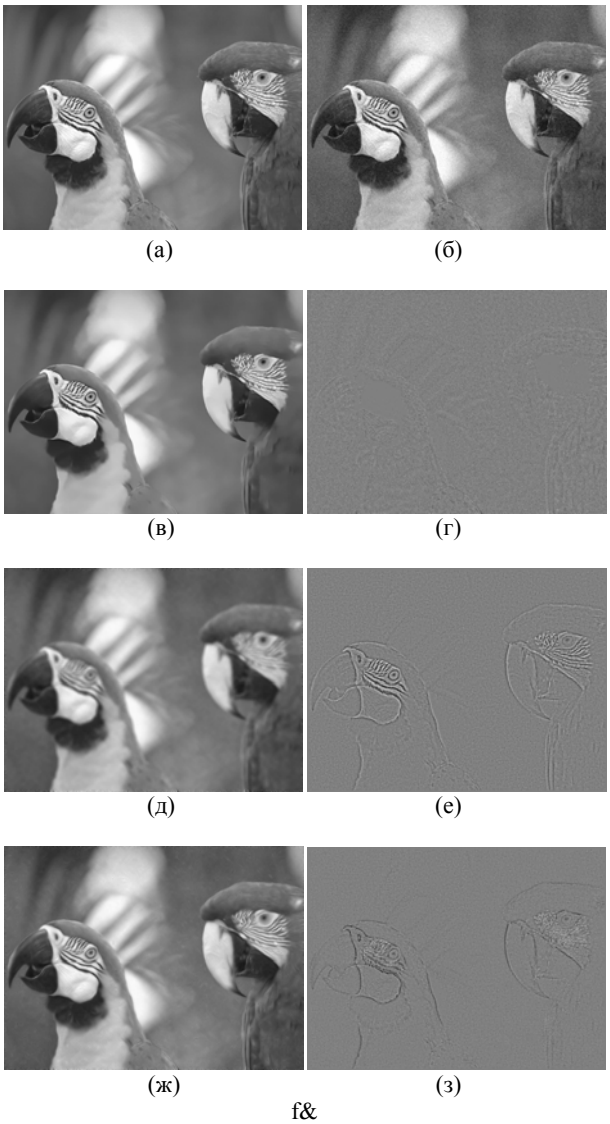
(д) (е)

(ж) (з)

Рис. 1. ***I04.***

(а) – Исходное изображение,
(б) – Зашумленное изображение
(в) – NLM-L2,
(г) – **Δ** между NLM-L2 и зашумленным изображением
(д) – NLM-SSIM,
(е) – **Δ** между NLM-SSIM и зашумленным изображением,
(д) – NLM-MSSIM,
(е) – **Δ** между NLM-MSSIM и зашумленным изображением

Подбор параметров производился при помощи прогона по сетке параметров и выбора наилучшего результата по метрике PSNR. Для NLM-L2 варьировался параметр ρ, для NLM-SSIM и NLM-MSSIM изменялись параметры C_1, C_2, C_3. Также для всех методов менялись размеры блока и область поиска. Методы NLM-SSIM и NLM-MSSIM имеют лучшую производительность, чем NLM-L2. Разрыв производительности между методами ещё больше возрастает при увеличении размеров блоков и диапазона поиска



(а) (б)

(в) (г)

(д) (е)

(ж) (з)

f&

Рис. 2. ***I23.***

(а) – Исходное изображение,
(б) – Зашумленное изображение
(в) – NLM-L2,
(г) – **Δ** между NLM-L2 и зашумленным изображением
(д) – NLM-SSIM,
(е) – **Δ** между NLM-SSIM и зашумленным изображением,
(д) – NLM-MSSIM,
(е) – **Δ** между NLM-MSSIM и зашумленным изображением

Результаты в таблице выбраны неслучайно. Изображение из первого примера представлено на рисунке 1. Можно заметить, что эффективность NLM-MSSIM по сравнению NLM-L2 возрастает на изображениях, которые имеют мелкие структурные элементы, которые не портятся в результате применения более эффективной метрики. Это можно

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

видеть в области прически на изображении 1. Во втором примере из таблицы мелкие детали более разрознены, что почти уравнивает NLM-SSIM и NLM-MSSIM. В последнем примере из таблицы, он изображен на рисунке 2, наоборот мелких деталей мало. Что делает это изображение благоприятным для применения классического NLM-L2, но благодаря участию блоков с близкой структурой при разбросе яркостей и контрастов NLM-MSSIM не проигрывает NLM-L2 в отличие от NLM-SSIM.

6. Заключение

Проанализированы классический метод нелокального среднего и его модификация на основе метрики SSIM, рассмотрены достоинства и недостатки этих методов. Предложен новый метод шумоподавления (9)-(11) на основе алгоритма нелокального среднего использующий новую весовую функцию (11) представляющую собой модификацию индекса структурного сходства. Для базы изображений TID2013 проведены его сравнения с классическим алгоритмом нелокального среднего [2] и алгоритмом [19], использующим стандартный индекс структурной схожести. Предложенный алгоритм показал себя более универсальным методом и выдал лучшие или сопоставимые результаты на различных типах изображений.

7. Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект № 17-11-01279).

8. Литература

[1] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one //Multiscale Modeling and Simulation, vol. 4, no. 2, pp. 490 530, 2005.

[2] Р.Гонсалес, Р.Вудс. Цифровая обработка изображений. М.Техносфера, 2005, 1072с.

[3] Л.П. Ярославский. Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. – М.:Радио и связь, 1987. -296 с.: ил.

[4] Cruz C. et al. Nonlocality-Reinforced Convolutional Neural Networks for Image Denoising //arXiv preprint arXiv:1803.02112. – 2018.

[5] Joachim Weickert, Anisotropic Diffusion in Image Processing, ECMI Series, Teubner-Verlag, Stuttgart, Germany, 1998.

[6] Perona P., Malik J. Scale-space and edge detection using anisotropic diffusion // Proceedings of IEEE Computer Society Workshop on Computer Vision. pp. 16–22 (1987).

[7] Storozhilova M.V., Yurin D.V. “Fast Rank Algorithms Based on Multiscale Histograms” // In: 21-th International Conference on Computer Graphics GraphiCon'2011. Moscow, Russia, 2011, pp. 132–135.

[8] Storozhilova M.V., Yurin D.V. “Fast Rank Algorithms with Multiscale Histograms Lazy Updating” // 8th Open German-Russian Workshop “Pattern Recognition and Image Understanding” (OGRW-8-2011), November 2011, pp. 380–383.

[9] Tomasi C., Manduchi R. “Bilateral Filtering for Gray and Color Images” // Proceedings of the IEEE Sixth International Conference on Computer Vision (ICCV'98), January 1998, pp. 839–846.

[10] Buades A., Morel J.M. “A Non-Local Algorithm for Image Denoising” // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, vol. 2, 20–26 June 2005, pp. 60–65.

[11] Wang S., Xia Y., Liu Q., Luo J., Zhu Y., and Feng D. Gabor feature based nonlocal means filter for textured image denoising // Journal of Visual Communication and Image Representation, Oct. 2012, vol. 23, issue 7, pp. 1008-1018.

[12] N. V. Mamaev, A. S. Lukin, D. V. Yurin. “HeNLM–LA: a Locally Adaptive Nonlocal Means Algorithm Based on Hermite Functions Expansion” // Programming and Computer Software, Vol. 40, No. 4, Pleiades Publishing, Ltd., 2014, pp. 199−207.

[13] Manzanera A. Local Jet based similarity for NL-Means ltering // 20th International Conference on Computer Vision and Pattern Recognition (ICPR), 2010, pp. 2668-2671.

[14] Dabov K., Foi A., Katkovnik V., Egiazarian K., “Image denoising by sparse 3D transform-domain collaborative filtering,” IEEE Trans. Image Process., vol. 16, no. 8, pp. 2080–2095, August 2007.

[15] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, “Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising”//IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142–3155, 2017.

[16] K. H. Jin, M. T. McCann, E. Froustey, M. Unser, “Deep convolutional neural network for inverse problems in imaging”//IEEE Transactions on Image Processing, vol. 26, no. 9, pp. 4509–4522, 2017.

[17] Z. Wang, A. Bovik, H. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity // IEEE Trans. on Image Processing, vol. 13, pp. 600-612, Apr. 2004.

[18] Z.Wang, A.C. Bovik Mean squared error: love it or leave it? - a new look at signal fidelity measures //IEEE Signal Processing Magazine, vol. 26, pp. 98-117, Jan. 2009.

[19] Rehman A., Wang Z. SSIM-based non-local means image denoising // Image Processing (ICIP), 2011 18th IEEE International Conference on. -IEEE, 2011. -P. 217-220.

[20] Ponomarenko N. et al. Color image database TID2013: Peculiarities and preliminary results // Visual Information Processing (EUVIP), 2013 4th European Workshop on. -IEEE, 2013. -P. 106-111.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Лучевой метод реконструкции поверхности сердца человека по данным компьютерной томографии

М.В. Дубровская[1], М.М. Новожилов[1]
maria.v.dubrovskaya@gmail.com|michael.novogilov@gmail.com
[1] Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия

Выбор КТ в качестве исходных данных для персонализации модели сердца обусловлен сравнительно высоким пространственным разрешением данных во всем объеме моделирования и доступностью КТ. Реконструируется внешняя поверхность сердца, поверхность миокарда желудочков. Применены методы: двумерный и трёхмерный Ray Casting, триангуляция внешней поверхности миокарда, заданной точками контуров, полученных лучевым методом. Проведён предварительный сбор статистики плотностей для тканей грудной клетки на исследованных томограммах для более точного определения расположения точек поверхности. Реализованы масштабирование, угловая и параметризация по узлам сетки окружающего параллелепипеда. Результат реконструкции является достаточно качественным для использования в качестве ограничивающей поверхности данных томографии сердца пациента для дальнейшего анализа аномалий стенок желудочков и межжелудочковой перегородки.

Ключевые слова: *реконструкция поверхности, компьютерная томография, Ray Casting, сегментация медицинских изображений, параметризация, полигональная модель.*

Ray Casting for human heart surface reconstruction, based on CT

M.V. Dubrovskaya[1], M.M. Novozhilov[1]
maria.v.dubrovskaya@gmail.com|michael.novogilov@gmail.com
[1] National Research Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation

The choice of CT as source data for heart model personalization is caused by it`s relatively high spatial data resolution of the whole modelling volume, and by accessibility of CT. We reconstruct external heart surface and ventricular myocardium surface. The next methods are applied: 2D and 3D Ray Casting, and exterior myocardium surface triangulation, specified by points of contours, acquired by Ray Casting. The preliminary densities` statistics for chest tissues is accumulated for investigated CT-scans in order to detect surface points` positions more precisely. We implemented scaling, angle parameterization and parameterization along bounding box grid. The result of the reconstruction is qualitative enough to be used as a bounding surface of the patient's heart tomography data for further analysis of abnormalities of the ventricular wall and interventricular septum.

Keywords: *surface reconstruction, CT-scans, Ray Casting, medical image segmentation, parameterization, polygonal model.*

1. ВВЕДЕНИЕ

Целью работы является расширение возможностей неинвазивной диагностики сердечной мышцы пациентов. Необходимым шагом на этом пути является отделение тканей сердца от всех внешних по отношению к сердцу объектов. Такой границей и является внешняя поверхность сердца. Внутри этой оболочки уже можно решать задачу детектирования и количественной оценки аномалий остро необходимую в практике цифровой диагностики. Для кардиохирургов это также возможность «заглянуть» внутрь сердца пациента для детальной подготовки операции. Результаты могут быть использованы также в медицинских институтах для исследования и демонстрации сердечных патологий. Для обеспечения возможности управления формой реконструированной поверхности во времени и пространстве важно, чтобы модель была параметризована в анатомически адекватных системах и координатах. Такая параметризация позволила бы также успешно решать задачу совмещения моделей двух состояний одного сердца или даже двух разных сердец.

2. ОБЗОР

Aubin [1] (1995) основал на компьютерной томографии построение геометрической модели ряда тканей с использованием конечно-элементного метода. В 2005 Bean [2] рассмотрел свойства 3D изображений КТ как базы для исследований. Zheng [17] (2008) осуществил автоматическую сегментацию всех камер сердца на основе КТ. Сегментация формулируется как задача обучения из двух этапов: локализация анатомической структуры и очерчивание границ. Общее расположение основано на обучении в рамках окружающего пространства. Построенная модель представлена в форме закрытой сетки. Для обучения определены «управляемые признаки», такие, что информация об ориентации и масштабе вкладывается в распределение точек выборки, в то время как каждый конкретный признак определён локально. Ключевые точки модели определяются как контрольные точки. В 2009 Heismann [6] решил проблему повышения точности реконструкции на основе КТ. Он разработал механизм спектральной реконструкции, где данные КТ оцениваются как спектральная интеграция коэффициента спектрального затухания, которым определяется истинное положение (ground truth) сканируемого объекта. Симулятор сердца человека, способный воспроизводить сокращения миокарда и движение кровеносного потока был создан Seiryo Sugiura [13] (2013). Модель построена на основе молекулярного механизма возбуждения-сжатия волокон миокарда, а также основана на конечно-элементном методе и решает уравнения, описывающие электрическую активность, твердотельную механику и динамику жидкости. Использует данные КТ.

Norouzi [10] (2014) рассмотрел методы классификации и кластеризации медицинских изображений, а также методы сегментации, основанные на разрастании регионов. Сегментации посвящено исследование Xu [16] (2015), где рассмотрены различные техники многоатласной сегментации. Предложенный метод основан на технике SIMPLE, реализующей статистическое изучение контекста, и методе JLF (joint label fusion), позволяющему уменьшить

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

вклад ошибок корреляции среди выбираемых атласов. Для упорядочивания объединённой сегментации используется техника усечения графа. Liu [8] (2015) сформулировал задачу реконструкции как проблему оптимизации. Предложена параметризация дыхательного движения, что позволило создать его точную кинематическую модель. Реконструкция формы объекта формулируется как задача оптимизации с регуляризацией полной вариации с использованием референсного изображения и векторных полей. Данная проблема решается с помощью метода «proximal alternating minimization algorithm».

Mansi [9] (2015) создал твердотельную модель желудочков сердца, ориентированную на конкретных пациентов. Модель генерируется на основе анатомической модели конкретного пациента для левого и правого желудочков и клинических данных пациентов, и включает биомеханику, электрофизиологию и гемодинамику, благодаря чему уточняются первоначальные параметры модели. Динамика жидкости, основанная на потоках крови сердца, рассматривается для оценки внутренних параметров сердца. Результирующая закрытая сетка используется для создания объёмной тетраэдральной сетки. Присутствует модель ориентации волокон.

Лидером в создании параметризованных моделей сердца, восстановленных из КТ, является фирма SIEMENS [12] (2015). Во время зондирования специальным зондом, снимающим координаты и электрическую активность сразу в нескольких точках на внутренней поверхности левого предсердия (ЛП) собираются данные о поверхности. Параметрическая модель представлена во внутренних координатах и приводится в соответствие электрическим данным и предварительной статистике. При построении модели происходит трансформация точек данных, которые соответствуют трубкам и эллипсоидам в параметрической модели. Соотнесение предварительной статистики с параметрической моделью включает подготовку сегментированных сеток данных из КТ и вычисление анатомических признаков из сеток данных, а также вычисление коэффициентов корреляции между различными экземплярами анатомических признаков. Использование электрических данных в соотнесённой параметрической модели включает минимизацию целевой функции, которая описывает оцененную ошибку параметрической модели в соответствии с электрическими данными. Минимизация выполняется через назначение соответствующих весов параметрам модели. Заполнение параметрической модели производится взвешиванием искривления. Модель поверхности соотносится с каждой сеткой набора данных, и к результирующей поверхности и/или её параметрам применяется статистический анализ. Вначале выполняется «грубое» соотнесение, так что исходные оценки должны давать приемлемый результат.

Panthee, N. [11] (2016) в своей работе с ресинхронизирующей терапией сердца (CRT, cardiac resynchronization therapy) на данных КТ создал модель желудочков сердца собаки частично с использованием коммерческого ПО. Получены конечно-элементные сетки желудочков, а также воксельные сетки сердца для электрофизиологии и тетраэдральные сетки для анализа механики.

Jinwoo Choi [3] (2016), также работавший с CRT, осуществил трёхмерное моделирование точки воздействия терапии на левый желудочек (ЛЖ). Для этого предварительно полуавтоматически сегментируется всё сердце с использованием приложения ITK-Snap.

Одним из способов 3D реконструкции из данных томографии является уточнение методом активного контура, которое рассмотрел Vasil'ev [15] (2016). Начальное приближение контура определяется оценкой локализации на изображении плотностей, характерных для объекта, методом интегральных проекций этих плотностей на оси координат.

Hammon [5] (2017) утверждает, что положительный опыт реконструкции из КТ очень скромен. В своей работе он предлагает пороговую сегментацию, основанную на статистическом анализе, и показавшую результат с нулевой ошибкой.

Ряд работ посвящается созданию моделей для 3D принтеров. Kiraly [7] (2018) для реконструкции виртуальной модели из данных томограмм использует полуавтоматическую сегментацию. Uccheddu [14] (2018) даёт обзор последних достижений по сегментации, в числе которых интерактивные методы сегментации и реконструкции медицинских изображений.

3. МЕТОД РЕКОНСТРУКЦИИ ПОВЕРХНОСТИ

Для реконструкции внешней поверхности сердца разработана методология использования лучевого метода обратной трассировки (Ray Casting). Результатом работы метода являются контуры сердца для каждого из последовательности слоев томограммы.

При своей относительной простоте по сравнению с рассмотренными альтернативами алгоритм показывает хорошие результаты для области внешней поверхности желудочков сердца. Метод является малозатратным (порядка 10 секунд), но позволяет менять разрешение и масштабы модели благодаря параметризации.

Методология, в отличии от Zheng [17], Mansi [9] и Safran [12], основывается на испускании двумерных и трёхмерных лучей в пределах исходного объёма данных томограммы. Два данных способа реализованы независимо. Метод является полуавтоматическим. На этапе алгоритма реконструкции, выполняемом вручную, берётся предварительная разметка центров испускания лучей. Центры устанавливаются внутри сердца, независимо от попадания их в полости камер или на ткани миокарда и сопутствующие ткани, а также от наличия или отсутствия контраста в крови внутри камер.

Исходный объём данных представляет собой трёхмерный набор плотностей, снятых томографом с регулярным шагом, поэтому в методах реконструкции для получения плотностей в точках с произвольными координатами используется линейная интерполяция.

3.1 Двумерный Ray Casting

В данной вариации реконструкции двумерные лучи выпускаются послойно и радиально в пределах слоя. Перед работой основной части алгоритма строится локальная система координат, которая основывается на точках, полученных при ручной разметке. Данная система координат является радиальной (круговой). В основании системы координат лежит (в общем случае наклонный) цилиндр. Точки, центры испускания лучей, выбираются экспертным методом.

Лучи для каждого слоя томограммы (слои расположены вертикально, вдоль оси Oz) радиально испускаются по заданному углу в двумерном пространстве в пределах каждого слоя. Третьей координатой является номер слоя, умноженный на размер шага, с которым томограф сканировал сердце пациента по оси Oz. Количество лучей в одном ярусе (в данном случае, слое) аналогично соответствующему значению для метода трёхмерного Ray Casting`a (раздел 3.2).

3.1.1 Статистический анализ плотностей

Далее вдоль лучей решается задача обработки плотностей. Предварительно перед выбором точек

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

поверхности вдоль лучей был проведён статистический анализ плотностей тканей грудной клетки для рассмотренных томограмм. Диапазоны плотностей тканей в КТ стандартизованы таким образом, что они могут принимать значения от -1000 HU (Hounsfield unit, специальная единица измерения плотностей по шкале Хаунсфилда для КТ) до 3000 HU, а также установлены соответствия значения данной шкалы для воды (0 HU) и воздуха (-1000 HU). Тем не менее, для томограмм, полученных различными томографами, величины плотностей в HU могут варьироваться для разных тканей, поэтому для наиболее корректного выбора пороговых значений был проведён статистический анализ плотностей исследуемых томограмм (рис. 1). Сбор статистики производился в приложении RadiAnt. Статистика подтверждает, что плотности всех элементов сердца положительны, а жировая ткань отрицательна в HU.

Ткань / вещество	Плотность, HU
Вода	0
Воздух	-1000
Воздух (вне тела)	[-1024; -956]
Воздух (в трахее и лёгочных путях)	[-979; -822]
Лёгкие	[-877; -297]
Кости	[215; 2976]
Жир	[-224; -11]
Печень	[36; 143]
Кровь (контраст, предсердия)	[82; 173]
Кровь (контраст, желудочки)	[82; 144]
Миокард (предсердия)	[7; 67]
Миокард (желудочки)	[12; 105]
Перегородка между предсердиями	[68; 115]
Перегородка между желудочками	[33; 113]

Рис. 1. Таблица Хаунсфилда, составленная по данным исследованных томограмм

Двигаться вдоль лучей нужно с достаточно мелким шагом, чтобы не пропускать переходы от одних тканей к другим. Для получения значений плотности в точках, попадающих между снятыми томографом значениями, используется билинейная интерполяция. Использовался шаг параметризации по углу, равному 5 градусов. Например, для объёма данных, в котором это расстояние составляет 3мм по оси Oz и ~1.7мм, по осям Ox и Oy использовался шаг длиной 0.125мм. Для этого из координат всех векторов-направлений (для испускания лучей) были вычтены координаты выбранных центров: собственный центр для каждого среза при послойной обработке и единственный центр при 3D обработке. Затем векторы направлений лучей нормировались и их длины делились на 8, поскольку экспериментально установлено [4] (Gavrilov, 2014), что такая длина шага достаточно мала, чтобы не пропускать переходы из одних тканей в другие.

3.1.2 Устранение шумов вдоль лучей

После получения плотностей, найденных с выбранным шагом вдоль каждого луча, проведён их анализ. Для внешней поверхности сердца нужно определять, когда луч покидает зону контраста (или не контрастированной крови) и миокарда.

Для того чтобы убедиться, что переходы между группами соседних плотностей, соответствующих определённым тканям, хорошо выражены, были построены графики плотностей вдоль лучей (рис. 2).

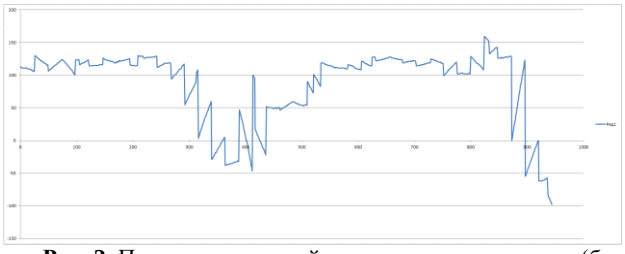


Рис. 2. Пример плотностей, полученных вдоль луча (без постобработки). По вертикальной оси: значения плотностей; по горизонтальной: порядковые номера отсчётов вдоль луча, в которых были получены плотности

Как видно из рисунка (рис. 2), плотности, снятые вдоль лучей, непригодны для требуемого принятия решений, поэтому значения плотностей вдоль лучей были сглажены (рис. 3), так чтобы исчез эффект изрезанности, но остались видны перепады плотностей, обозначающие границы между тканями. Экспериментально выбрано ядро размера 5мм (шаг 1/8мм).

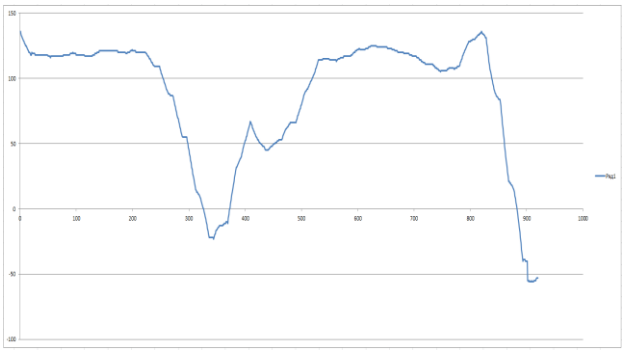


Рис. 3. Пример плотностей, полученных вдоль луча, сглаженных с окном сглаживания 5мм. По вертикальной оси: значения плотностей; по горизонтальной: порядковые номера отсчётов вдоль луча, в которых были получены плотности

Сглаживание плотностей вдоль лучей позволило в необходимой степени уменьшить шум томограммы.

3.1.3 Выбор точек поверхности на лучах

За решающее правило принят тот факт анатомии сердца человека, что оно окружено перикардом, сплошным жировым мешочком толщиной до нескольких миллиметров. Перикард фактически обволакивает внешнюю поверхность сердца и основания крупных сосудов. Его плотность в Хаунсфилдах отрицательна, в отличие от плотностей всех областей сердца. При переходе от миокарда к ткани перикарда берётся ближайшая плотность перикарда. Таким образом, фактическим условием принятия точки вдоль луча в качестве точки внешней поверхности миокарда выбран переход от положительных значений плотностей к отрицательным.

Поверхность, изображённая на рисунке (рис. 4), является внешней поверхностью желудочков сердца, построенной методом двумерного Ray Casting`а. Решение рассматривать поверхность этой части сердца отдельно было принято по следующим причинам:

1. На данной части поверхности сердца удаётся получить наилучшее качество, так как для нее алгоритм сегментации работает устойчиво.
2. Части сердца, включающие в себя места присоединения крупных сосудов, а также сами сосуды сложно детектировать и анализировать их расположение, в результате чего реконструированные поверхности имеют проблемные зоны в этих областях.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

3. Среди всех известных попыток корректной реконструкции поверхности сердца из томограммы лучшая принадлежит крупнейшей корпорации SIEMENS [12] в 2015 году, причём запатентована была восстановленная параметризованная модель только участка поверхности сердца, а именно внутренняя поверхность левого предсердия.

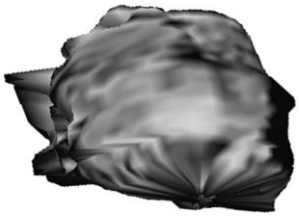


Рис. 4. Поверхность, полученная двумерным лучевым методом (внешняя поверхность миокарда желудочков); вид на верхушку сердца (Apex)

Рисунок (рис. 5) демонстрирует пример расположения детектированных граничных точек на срезе исходных данных для внешней поверхности сердца, построенной методом двумерного Ray Casting'а. Видно, что большинство точек поверхности детектируется верно, однако есть «выбросы», происходящие из-за недостаточности одного критерия, равенства нулю плотности на границе поверхности, (рис.5, желтая обводка, оранжевая обводка около точки *C*), обработка которых требует усложнения алгоритма.

Рис. 5. Пример расположения детектированных граничных точек на срезе исходных данных для внешней поверхности сердца, построенной методом двумерного Ray Casting`а; *C* – центр испускания лучей

3.2 Трёхмерный Ray Casting

Для испускания трёхмерных лучей также строится локальная система координат. В данном случае она построена в форме регулярной сетки по окружающему параллелепипеду. Сглаживание плотностей вдоль лучей проводится аналогично двумерному Ray Casting`у.

В трёхмерной вариации метода все лучи с определённым шагом испускаются из одной точки исходного массива данных. Эта точка располагается в области соединения всех четырёх камер сердца. Лучи идут по регулярной сетке, организованной следующим образом. За каркас сетки принят окружающий параллелепипед области сердца. На каждой из шести граней параллелепипеда строится своя регулярная сетка с аналогичными количествами делений для каждой грани. Каждая грань и по высоте, и по ширине равномерно разбивается на 16 сегментов (по каждой стороне). Все испускаемые лучи идут из выбранного заранее центра по шести сеткам размера 17x17.

Для получения плотностей во всех точках, выбранных с заданным шагом вдоль лучей, используется трилинейная интерполяция.

На рисунке (рис. 6) представлен результат работы алгоритма. Данная реконструкция хорошо строит боковую поверхность сердца, однако в местах перехода от одних подструктур сердца к другим наблюдаются разрывы и неоднозначности.

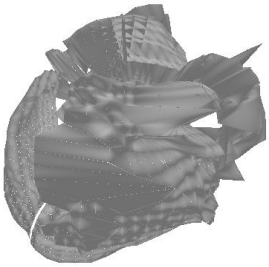


Рис. 6. Поверхность, полученная трёхмерным лучевым методом; вид сверху на верхушку сердца (Apex), ЛЖ и ПЖ

Основной причиной перечисленных проблем является трудность сегментации такого сложного объекта как сердце на лучах, выходящих из одной точки. Действительно, множество тканей неоднократно перекрываются на пути лучей, пускаемых в трёхмерном пространстве.

3.3 Сравнение лучевых методов

На внешних поверхностях сердца, построенных методом двумерного Ray Casting`а, неоднозначностей значительно меньше, чем на поверхностях, построенных методом трёхмерного Ray Casting`а из-за искажений линейной интерполяции. Томограммы строятся послойно с фиксированным разрешением по осям X, Y, но шаг по оси Z часто кратно больше, чем по X, Y, что увеличивает ошибку линейной интерполяции.

4. ПАРАМЕТРИЗАЦИЯ

На данный момент реализованы следующие виды параметризации:

1. Угловая: регулярная параметризация по испускаемым двумерным лучам в пределах слоя.
2. По узлам сетки окружающего параллелепипеда (для 3D лучей). Параметром является уровень разбиения рёбер граней регулярной сетки, покрывающей окружающий параллелепипед, через точки которой происходит испускание лучей.
3. Масштабирование. Позволяет как масштабировать полученные на лучах точки (то есть, всю поверхность), так и растягивать все точки поверхности по осям: Oz, Ox, Oy, Ox и Oy.

Параметризация определяется в том числе типом системы координат, использованной в алгоритме.

Угловая параметризация, а также параметризация лучей по регулярной сетке дают базовые возможности для совмещения атлас-моделей с данными томограммы конкретного пациента.

Масштабирование реализует изменения масштаба по различным ортогональным осям, для двумерного и трёхмерного Ray Casting`а соответственно. Получаемые внешние поверхности сердца строятся описанными выше методами, для которых сетки окружающего сердце параллелепипеда после получения плотностей на лучах, выпускаемых в направлениях их узлов, равномерно масштабируются в выбранных направлениях. Таким образом, есть возможность как полного масштабирования поверхности, так и растяжения всех точек поверхности вдоль осей: Oz, Ox, Oy, Ox и Oy, что является альтернативой работы на основе полярной или цилиндрической систем.

Выбранные виды параметризации следуют непосредственно из принципов лучевого метода, и обеспечивают гибкость его использования.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

5. ВИЗУАЛИЗАЦИЯ ПОВЕРХНОСТЕЙ

Визуализация поверхностей осуществляется с использованием библиотеки OpenGL 4.x.

Для двумерного послойного алгоритма поверхность является результатом триангуляции последовательности участков поверхности, натянутых на пару контуров детектированных на соседних слоях томограммы. В таком виде она и подается на визуализацию, а лента из треугольников между контурами подается на визуализацию как GL_TRIANGLE_STRIP.

Для трехмерного лучевого алгоритма, формирующего вершины модели поверхности в виде пар (точка, нормаль) поверхность считается реконструированной в виде сетки треугольных пачей Безье. В этом случае мы можем воспользоваться примитивами OpenGL 4.0 GL_PATCHES или визуализировать поверхность как описано выше в виде ленты опорных треугольников (треугольных оснований патчей).

6. ЗАКЛЮЧЕНИЕ

В работе предложен и исследован лучевой метод реконструкции поверхности сердца человека по данным компьютерной томографии. Реконструируется внешняя поверхность сердца, поверхность миокарда желудочков. Выполнен обзор по тематике исследования, подтверждающий её актуальность. Применены методы двумерного и трёхмерного бросания лучей (Ray Casting), алгоритмизована триангуляция внешней поверхности миокарда, заданной точками контуров, полученных лучевым методом. Реализованы различные способы параметризации реконструированной поверхности.

Время реконструкции внешней поверхности сердца в среднем составляет около 10 секунд.

По сравнению с рассмотренными аналогами метод требует минимального участия пользователя на этапе ручной настройки. Результат реконструкции является достаточно качественным для его использования в качестве ограничивающей поверхности данных томографии сердца пациента для дальнейшего анализа аномалий стенок желудочков и межжелудочковой перегородки.

7. ЛИТЕРАТУРА

[1] Aubin CE1, Descrimes JL et al. Geometrical modeling of the spine and the thorax for the biomechanical analysis of scoliotic deformities using the finite element method. Annales de Chirurgie 49(8):749-61, February 1995

[2] Bean MJ, Pannu H et al. Three-dimensional computed tomographic imaging of complex congenital cardiovascular abnormalities. J Comput Assist Tomogr2005; 29: 721–724.

[3] Choi, J., Radau, P. et al. G. X-ray and Magnetic Resonance Imaging Fusion for Cardiac Resynchronization Therapy, Medical Image Analysis (2016), doi: 10.1016/j.media.2016.03.004

[4] Gavrilov, N.I., Turlapov, V.E. Novel approach to development of direct volume rendering algorithms based on visualization quality assessment // Programming and Computer Software, (2014) 40 (4), pp. 174-184. DOI: 10.1134/S0361768814040045.

[5] Hammon, M., Rompel, O. et al. Accuracy and Specific Value of Cardiovascular 3D-Models in Pediatric CT-Angiography. Pediatr Cardiol (2017) 38: 1540. https://doi.org/10.1007/s00246-017-1693-7

[6] B. Heismann and M. Balda. Quantitative image-based spectral reconstruction for computed tomography. Medical Physics 36, 4471 (2009); doi: 10.1118/1.3213534

[7] Kiraly, Laszlo. Three-Dimensional Modelling and Three-Dimensional Printing in Pediatric and Congenital Cardiac Surgery. *Translational Pediatrics* 7.2 (2018): 129–138. *PMC*. Web. 19 Aug. 2018.

[8] J. Liu et al., "5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography", Inverse Prob., vol. 31, no. 11, pp. 115007, 2015.

[9] Mansi; Tommaso; et al., Siemens Corporation. Method and system for advanced measurements computation and therapy planning from medical data and images using a multi-physics fluid-solid heart model. Patent: US 9129053 B2. Sep. 8, 2015.

[10] Alireza Norouzi, Mohd Shafry et al. Medical Image Segmentation Methods, Algorithms, and Applications, IETE Technical Review, 31:3, 199-213, DOI: 10.1080/02564602.2014.906861, 2014

[11] Panthee, N., Okada, J. et al. (2016). Tailor-made heart simulation predicts the effect of cardiac resynchronization therapy in a canine model of heart failure. Med. Image Anal. 31, 46–62. doi: 10.1016/j.media.2016.02.003

[12] Moshe Safran, Siemens Corporation. Model based reconstruction of the heart from sparse samples. Patent: US 20150018698 A1. Jan. 15, 2015.

[13] Sugiura, Seiryo & Washio, Takumi & Okada, et al. (2013). Tailor-made Medicine Using the Multi-scale Heart Simulator "UT-Heart". Journal of Cardiac Failure. 19. S107. 10.1016/j.cardfail.2013.08.023.

[14] F. Uccheddu, M. Carfagni et al. 3D printing of cardiac structures from medical images: an overview of methods and interactive tools. International Journal on Interactive Design and Manufacturing, 12 (2) (2018), pp. 597-609

[15] Vasil'ev, E.P., Belokamenskaja, A.A. et al. A parallel algorithm for 3D reconstruction of internal organs according to imaging based on the active contour model / CEUR Workshop Proceedings (2016) 1576, c. 482-489

[16] Zhoubing Xu, Ryan P. Burke et al. Efficient multi-atlas abdominal segmentation on clinically acquired CT with SIMPLE context learning, In Medical Image Analysis, Volume 24, Issue 1, 2015, Pages 18-27, ISSN 1361-8415, https://doi.org/10.1016/j.media.2015.05.009.

[17] Y. Zheng, A. Barbu, et al. "Four-Chamber Heart Modeling and Automatic Segmentation for 3-D Cardiac CT Volumes Using Marginal Space Learning and Steerable Features," in *IEEE Transactions on Medical Imaging*, vol. 27, no. 11, pp. 1668-1681, Nov. 2008. doi: 10.1109/TMI.2008.2004421

Об авторах

Дубровская Мария Владимировна, аспирант кафедры математического обеспечения и суперкомпьютерных технологий института информационных технологий, математики и механики Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского (ННГУ). Его e-mail maria.v.dubrovskaya@gmail.com.

Новожилов Михаил Михайлович, аспирант кафедры математического обеспечения и суперкомпьютерных технологий института информационных технологий, математики и механики Национального исследовательского Нижегородского государственного университета им. Н.И. Лобачевского (ННГУ). Его e-mail michael.novogilov@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Нахождение параметров метода шумоподавления с использованием вейвлет-преобразования на медицинских изображениях на основе анализа хребтовых структур

Мамаев Н.В.[1], Юрин Д.В.[1], Крылов А.С.[1]
mamaev.nikolay93@mail.ru|kryl@cs.msu.ru|yurin@cs.msu.ru

[1]Факультет Вычислительной математики и кибернетики
МГУ имени М.В.Ломоносова, Москва, Россия

Методы подавления шума зависят от параметров, регулирующих силу фильтрации. На практике незашумленное изображение недоступно, и возникает задача выбора оптимальных параметров, использующая только исходное зашумленное и отфильтрованное изображения. Качество можно измерить в областях, где присутствуют хребтовые структуры (хребты и долины) с помощью анализа разностного кадра. Предложен метод, позволяющий оценить качество фильтрации на основе вычисления взаимной информации между значениями разностного кадра в точках, где присутствуют хребтовые структуры. Для нахождения хребтовых структур используется анализ матрицы Гессе для нахождения направлений и характерной ширины хребтов и долин. Тестирование метода производилось для шумоподавления с использованием вейвлет-преборазования на зашумленных изображениях из базы изображений глазного дна DRIVE.

Ключевые слова: *оценка качества шумоподавления, вейвлет-преобразование, хребтовые структуры, взаимная информация.*

Choice of the Parameter for Wavelet Transform Based Denoising Method for Medical Images Using Ridge Analysis

Mamaev N.[1], Yurin D.[1], Krylov A.[1]
mamaev.nikolay93@mail.ru|kryl@cs.msu.ru|yurin@cs.msu.ru

[1]Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University

Denoising methods depend on parameters that control filtering strength. In practice image without noise is unavailable and the problem of filter parameters choice arises using only source noisy and image. A quality of filtering can be measured in areas there ridges are present using method noise analysis. The method that measures filtering quality based on calculation of mutual information between method noise values in points there ridges are present is proposed. A Hessian matrix analysis is used for detection of ridges direction and scale. The method is tested for wavelet transform based denoising method on noised images from DRIVE retinal image database.

Keywords: *denoising quality assessment, wavelet transform, ridges, mutual information.*

1. Введение

Задача подавления шума является одной из классических задач обработки изображений, так как шум присутствует на любых цифровых изображениях. Зачастую подавление шума является подготовительным этапом перед работой другого метода, например, перед сегментацией. При этом в процессе шумоподавления важно сохранить границы объектов и мелкие объекты, присутствовавшие на исходном изображении. В подавляющем числе случаев шум на цифровых изображениях является аддитивным гауссовым. Для данного типа шума разработано множество эффективных методов шумоподавления, сохраняющих границы. Одним из наиболее простых методов является семейство ранговых фильтров (среднее по KNV и ε_v-окрестности) [13]. В этих методах, интенсивность выходного пикселя есть среднее интенсивностей некоторых пикселей исходного изображения из его окрестности. Вхождение каждого пикселя определяется его позицией по отношению к усредняемому пикселю в ранговом ряду интенсивностей. В методе среднего по ε_v-окрестности в усреднении участвуют только пиксели, отличающиеся по интенсивности от усредняемого меньше, чем на ε, а среднего по KNV – только K ближайших по интенсивности (рангу) пикселей. Обобщением усреднения по ε_v-окрестности является билатеральный фильтр [12], в котором отбор пикселей для усреднения по расстоянию и по яркости выполняется не на пороговой основе, а введением Гауссова веса. Метод нелокального среднего (NLM) [1] можно рассматривать как дальнейшую модификацию билатерального фильтра. Если в билатеральном фильтре вес пикселя в среднем значении стал зависеть от его похожести по яркости на центральный пиксель окрестности, то здесь веса зависят от похожести между небольшими окрестностями (патчами) вокруг пикселей. Эта модификация дает большой прирост качества в областях, где присутствуют границы и текстуры, но также имеет и недостатки: высокую вычислительную сложность и отсутствие инвариантности к повороту окрестности. Для устранения этих недостатков был предложен алгоритм LjNLM [7]. В этом алгоритме расстояние между окрестностями заменяется расстоянием между векторами признаков, характеризующими окрестность. Размерность этих векторов ниже, чем размерность количества пикселей в окрестности, поэтому вычислительная сложность алгоритма ниже. Признаки здесь – значения сверток с производными функции Гаусса до второго порядка. Эти признаки легко пересчитываются в системе координат, связанной с направлением градиента путем преобразования базиса, что обеспечивает инвариантность сравнений к ориентации окрестности. Еще одним методом шумоподавления является метод на основе обрезки коэффициентов вейвлет-преобразования [6]. Этот метод является вычислительно эффективным, так как состоит только из вычисления

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

вейвлет-преобразования, обрезки коэффициентов и вычисления обратного преобразования.

Все методы подавления шума зависят от параметров, регулирующих силу фильтрации, поэтому возникает задача выбора этих параметров. Идеалом является такое удаление шума, что после него изображение становится таким, каким оно было бы, если бы шума на нем изначально не было. Но реальной жизни незашумленные изображение недоступно. Поэтому необходимы методы, которые оценивают качество фильтрации, не использующие незашумленное изображение. Существуют методы, оценивающие качество изображения в целом, на основе локальных статистик [8, 9] или на основе анализа частот [10]. Однако эти методы не учитывают то, что изображение было получено после применения метода шумоподавления. Для анализа качества фильтра целесообразно использовать разностный кадр (разность между исходным зашумленным изображением и изображением после фильтрации). Разностный кадр в случае идеальной фильтрации содержит чистый шум, и значения между соседними пикселями разностного кадра будут некоррелированными. Если в процессе фильтрации были размыты или стерты некоторые объекты или границы, то на разностном кадре будет виден этот объект или полоска вдоль границы. Отметим также, что проверка наличия регулярных структур необходима в областях, где присутствуют объекты, которые при фильтрации необходимо сохранить, такие как хребтовые структуры. Поэтому для эффективной оценки качества фильтрации необходимо нахождение таких областей.

2. Шумоподавление на основе вейвлет-преобразования

Одним из вычислительно эффективных методов подавления шума, сохраняющего границы, является метод, использующий вейвлет-преобразование [6]. Сначала вычисляются коэффициенты вейвлет-преобразования, затем к полученным коэффициентам применяется пороговая функция:

$$\psi(x) = \begin{cases} x, |\, x \,| \geq \kappa \\ 0, |\, x \,| < \kappa \end{cases}. \quad (1)$$

Отфильтрованное изображение получается путем обратного преобразования отфильтрованных коэффициентов.

Как видно из (1), сила фильтрации зависит от значения порога κ: чем больше κ, тем большее количество коэффициентов удаляется, поэтому удаляется и больше шума, но вместе с этим и большее количество полезной информации также удаляется.

3. Многомасштабное выделение хребтовых структур

Фильтром, определяющим наличие хребтовых структур (ridges) [2], может являться лапласиан $\Delta L^{\sigma}(x,y)$, где:

$$L^{\sigma}(x,y) = \sigma^2 \cdot I(x,y) * G_{\sigma}(x,y), \quad (2)$$

$$G_{\sigma}(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{x^2+y^2}{2\sigma^2}}. \quad (3)$$

Множитель σ^2 присутствует в (2) для того, чтобы оклик лапласиана на различных масштабах был одинаков. Иными словами, если увеличить и изображение, и σ в одинаковое число раз, то значение $L^{\sigma}(x,y)$ в соответствующих точках будет одинаково. Также отметим, что дифференцирование $L^{\sigma}(x,y)$ эквивалентно свертке изображения $I(x,y)$ с соответствующей производной функции Гаусса. Пусть $L^{\sigma}_{xx}(x,y)$, $L^{\sigma}_{xy}(x,y)$ и $L^{\sigma}_{yy}(x,y)$ – производные $L^{\sigma}(x,y)$ по соответствующим переменным. Тогда лапласиан $\Delta L^{\sigma}(x,y)$ будет иметь максимум (минимум) в центральной точке долины (хребта) ширины 2σ. Для нахождения направления хребта или долины рассмотрим матрицу Гессе:

$$H^{\sigma}(x,y) = \begin{pmatrix} L^{\sigma}_{xx}(x,y) & L^{\sigma}_{xy}(x,y) \\ L^{\sigma}_{yx}(x,y) & L^{\sigma}_{yy}(x,y) \end{pmatrix}. \quad (4)$$

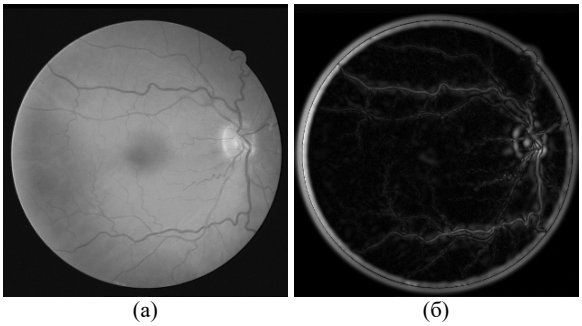
Тогда собственный вектор $\overrightarrow{v^{\sigma}}(x,y)$ матрицы (4), соответствующий наименьшему по модулю собственному значению будет направлен вдоль хребта (долины) [10].

Для многомасштабного нахождения хребтовых структур используются различные значения $\sigma \in \{\sigma_1,...,\sigma_n\}$, $\sigma_i = \sigma_0 \cdot \nu^{i-1}$. Вычислив соответствующие $\Delta L^{\sigma_i}(x,y)$ и $\overrightarrow{v^{\sigma_i}}(x,y)$, мы можем определить значения характерного масштаба и направления хребтовой структуры как:

$$s(x,y) = \arg\max_{\sigma_i}\left(\left|\Delta L^{\sigma_i}(x,y)\right|\right), \quad (5)$$

$$\vec{v}(x,y) = \overrightarrow{v^{s(x,y)}}(x,y). \quad (6)$$

Пример изображения из базы DRIVE и изображение значения модуля лапласиана $\left|\Delta L^{s(x,y)}(x,y)\right|$ показаны на Рис. 1.



(а) (б)

Рис. 1. (а) изображение из базы DRIVE, (б) значение модуля лапласиана $\left|\Delta L^{s(x,y)}(x,y)\right|$.

4. Оценка качества фильтрации на основе анализа хребтовых структур

Мы используем разностный кадр для анализа качества фильтрации. Разностный кадр $I_d(x,y)$ для изображения $I(x,y)$ определяется как разность между исходным зашумленным изображением $I(x,y)$ и изображением $I_f(x,y)$, полученным в результате применения метода шумоподавления:

$$I_d(x,y) = I(x,y) - I_f(x,y). \quad (7)$$

В идеальном случае разностный кадр содержит только шум. Присутствие деталей, которые также присутствовали на исходном изображении, означает, что эти детали были размыты или полностью стерты после применения метода шумоподавления, хотя мы хотим их сохранить. Главной характеристикой, отображающей, что разностный кадр содержит только лишь шум, является отсутствие корреляции между соседними пикселями. Присутствие регулярных структур на разностном кадре влечет за собой наличие корреляции между соседними пикселями в области, где эта

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

структура присутствует. В основном, эти структуры находятся в области, где находятся хребтовые структуры исходного изображения. Если после шумоподавления некоторые из хребтовых структур были размыты, разностный кадр будет содержать контуры этих структур, что выливается в наличие корреляции между значениями пикселей вдоль них, или, иными словами, возникновение в хаосе упорядоченности, мерой которой может быть высокое значение взаимной информации.

Для вычисления взаимной информации интенсивностей на разностном кадре построим совместное распределение $p(k,m)$ случайных величин K – значения разностного кадра в точке (x,y), и M – значения в точке $\overrightarrow{(x,y)} + s(x,y)\cdot\vec{v}(x,y)$. Для того, чтобы принимать во внимание только точки, принадлежащие хребтовым структурам, мы использует только точки, для которых $\left|\Delta L^{s(x,y)}(x,y)\right| > T$. Также значения разностного кадра квантуются на N значений так, что близкие значения интерпретируются как одно событие. Мы использует бикубическую интерполяцию для нахождения значений разностного кадра вне пиксельной сетки. Таким образом, $p(i,j)$ определяется следующим образом:

$$p(k,m) = \frac{1}{P}\cdot\#\left\{ \quad (x,y) \quad : \quad \left\lfloor \frac{I_d(x,y)\cdot N}{I_{\max}} \right\rfloor = k, \right.$$
$$\left. \left\lfloor \frac{I_d(\tilde{x},\tilde{y})\cdot N}{I_{\max}} \right\rfloor = m, \quad \left|\Delta L^{s(x,y)}(x,y)\right| \geq T \right\}, \tag{8}$$

$$\tilde{x} = x + s(x,y)\cdot\vec{v}(x,y)_x, \tilde{y} = y + s(x,y)\cdot\vec{v}(x,y)_y \tag{9}$$

где $\#\{...\}$ – мощность множества, а P – нормировочная константа, такая что $\sum_{k=1}^{N}\sum_{m=1}^{N} p(k,m) = 1$. Отметим, что в случае $s(x,y) = s \equiv const$, $\vec{v}(x,y) = \vec{v} \equiv const$ и $T = 0$ совместное распределение представляет собой матрицу смежности (GLCM) [3], построенную по направлению $s\cdot\vec{v}$.

Взаимная информация может быть использована как мера независимости случайных величин в совместном распределении. Для случайных величин K и M она определяется как:

$$\mu(K;M) = \sum_{k=1}^{N}\sum_{m=1}^{N} p(k,m)\log\left(\frac{p(k,m)}{p(k)p(m)}\right) \tag{10}$$

где $p(k)$ и $p(m)$ – функции маргинального распределения:

$$p(k) = \sum_{m=1}^{N} p(k,m), \quad p(m) = \sum_{k=1}^{N} p(k,m). \tag{11}$$

Чем меньше значение μ, тем меньше значения интенсивностей разностного кадра коррелированны, тем более беспорядочен (шумовой) разностный кадр. Отметим также, что лапласиан $\Delta L^{s(x,y)}(x,y)$ будет иметь большой отклик не только на хребтовых структурах, но и на границах и пятнах. Это означает, что значение μ также зависит от значений пар пикселей, принадлежащим этим структурам. Это не противоречит условию сохранения хребтовых структур после фильтрации, так как границы и пятна так же размываются после шумоподавления.

Таким образом, оценка качества фильтрации дается значением μ, подсчитанным по разностному кадру $I_d(x,y)$. Значения лапласиана, характерных масштабов и направлений хребтовых структур вычисляются по исходному зашумленному изображению $I(x,y)$. Минимальное значение μ будет соответствовать оптимальным параметрам фильтрации.

5. Результаты

Для тестирования предложенного метода использовалась база изображений глазного дна DRIVE [11], содержащая 40 изображений. Эти изображения содержат большое число хребтовых структур различных масштабов и направлений, поэтому эта база является хорошим тестовым набором для предложенного метода. Значения интенсивностей для изображений находятся в диапазоне от 0 до 255. Процедура тестирования происходила следующим образом:

1. К каждому изображению из базы добавлялся гауссов шум со стандартным отклонением σ_{noise}. Так, для каждого изображения I мы имеем изображение $I_{\sigma_{noise}}$. Мы использовали значения $\sigma_{noise} \in \{2i,\ i = 1,...,10\}$.
2. К изображению $I_{\sigma_{noise}}$ применялся метод шумоподавления на основе обрезки коэффициентов вейвлет-преобразования с различными значениями порога $\kappa \in \{\kappa_1,...,\kappa_L\}$. В качестве материнского вейвлета использовался биортогональный вейвлет bior3.5. Таким образом, мы имеем набор изображений после применения метода шумоподавления I_{f_κ}.
3. Вычислялись значения μ для изображений I_{f_κ}. Так как мы знаем, что шум является аддитивным гауссовым, то значения метрики PSNR между исходным изображением I и изображениями I_{f_κ} будут отображать качество фильтрации, и максимум PSNR будет соответствовать оптимальным параметрам. Мы использовали следующие параметры детектора хребтовых структур: $\sigma_0 = 1$, $v = \sqrt{2}$, и $n = 8$. Были протестированы различные значения количества уровней квантования $N \in \{2^i,\ i = 2,...,7\}$, и было обнаружено, что лучшие результаты дает значение $N = 16$.
4. Находились оптимальное значение параметров фильтрации $\kappa_{opt} = \arg\max_{\kappa}(PSNR(I, I_{f_\kappa}))$ и оцененное $\kappa_{est} = \arg\min_{\kappa}(\mu(I_{f_\kappa}))$. Вычислялась относительная погрешность оцененных параметров и оптимальных:

$$e = \frac{|\kappa_{est} - \kappa_{opt}|}{\kappa_{opt}}. \tag{12}$$

Описанная процедура так же применялась к методам оценки качества изображения BRISQUE [8], BLIDNS-2 [10] и DIVINE [9] путем замены вычисления μ на вычисление соответствующей оценки в пункте 3. Значения оценок максимизировались вместо минимизации взаимной информации. График зависимости среднего значения e по изображениям из базы DRIVE [11] для предложенного метода и методов [8-10] от уровня добавляемого шума показан на Рис. 2. Видно, что для всех уровней шума предлагаемый метод находит оптимальный параметр точнее, чем другие

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

методы. Графики зависимостей значений μ и PSNR между исходным незашумленным изображением и отфильтрованным для изображения, изображенного на Рис. 1(а) показаны на Рис. 3. Видно, что значение параметра, минимизирующего μ совпадает со значением параметра, максимизирующего PSNR. Также видно, что кривая достаточно гладкая для использования метода численного нахождения минимума для автоматического выбора параметра фильтрации. Результаты шумоподавления с различными значениями κ и соответствующие им разностные кадры показаны на Рис. 4.

6. Заключение

Предложен метод оценки качества фильтрации на основе анализа разностного кадра в областях, где присутствуют хребтовые структуры. Предлагаемый метод находит оптимальные параметры метода шумоподавления на основе вейвлет-преобразования лучше, чем существующие методы оценки качества изображения BRISQUE, BLIDNS-2 и DIVINE на изображениях глазного дна. На практике метод может быть использован для автоматического выбора параметров фильтрации изображений. Направлением дальнейших исследований является, в том числе, анализ предложенного алгоритма в случае мультипликативного шума.

7. Благодарности

Работа выполнена при поддержке гранта РФФИ 16-07-01207.

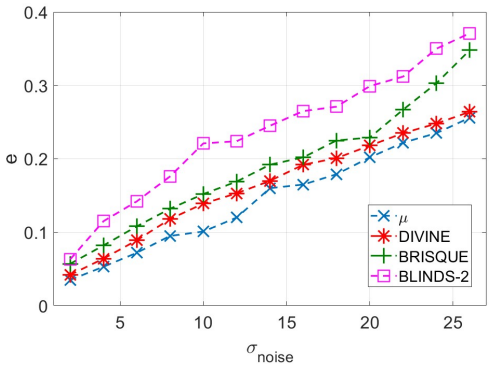


Рис. 2. Графики зависимости значений относительной ошибки (12) от уровня добавленного шума для методов [8-10] и предложенного метода.

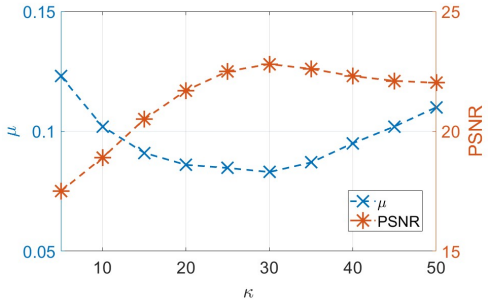


Рис. 3. Графики зависимости взаимной информации μ (10) и PSNR между отфильтрованным и исходным изображением без шума от значения порога κ при уровне шума $\sigma_{noise} = 16$. В отличие от PSNR, вычисление μ не требует знания незашумленного сигнала, который при обработке реальных данных недоступен и, собственно, является искомым. Видно, что экстремумы обеих кривых совпадают.

8. Литература

[1] Buades A., Coll B., Morel J. M. A review of image denoising algorithms, with a new one //Multiscale Modeling & Simulation. – 2005. – V. 4. – no. 2. – P. 490-530.

[2] Eberly D. Ridges in image and data analysis. – Springer Science & Business Media, 1996. – V. 7.

[3] Haralick R. M. et al. Textural features for image classification //IEEE Transactions on systems, man, and cybernetics. – 1973. – no. 6. – P. 610-621.

[4] Kim V., Iaroslavskii L. Rank algorithms for picture processing //Computer Vision graphics and image processing. – 1986. – V. 35. – P. 234-258.

[5] Lindeberg T. Edge detection and ridge detection with automatic scale selection //International Journal of Computer Vision. – 1998. – V. 30. – no. 2. – P. 117-156.

[6] Mallat S. G. A theory for multiresolution signal decomposition: the wavelet representation //IEEE transactions on pattern analysis and machine intelligence. – 1989. – V. 11. – no. 7. – P. 674-693.

[7] Manzanera A. Local jet based similarity for NL-means filtering //Pattern Recognition (ICPR), 2010 20th International Conference on. – IEEE, 2010. – P. 2668-2671.

[8] Mittal A., Moorthy A. K., Bovik A. C. No-reference image quality assessment in the spatial domain //IEEE Transactions on Image Processing. – 2012. – V. 21. – no. 12. – P. 4695-4708.

[9] Moorthy A. K., Bovik A. C. Blind image quality assessment: From natural scene statistics to perceptual quality //IEEE transactions on Image Processing. – 2011. – V. 20. – no. 12. – P. 3350-3364.

[10] Saad M. A., Bovik A. C., Charrier C. Blind image quality assessment: A natural scene statistics approach in the DCT domain //IEEE transactions on Image Processing. – 2012. – V. 21. – no. 8. – P. 3339-3352.

[11] Staal J. et al. Ridge-based vessel segmentation in color images of the retina //IEEE transactions on medical imaging. – 2004. – V. 23. – no. 4. – P. 501-509.

[12] Tomasi C., Manduchi R. Bilateral filtering for gray and color images //Computer Vision, 1998. Sixth International Conference on. – IEEE, 1998. – P. 839-846.

[13] Kim V., Iaroslavskii L. Rank algorithms for picture processing //Computer Vision graphics and image processing. – 1986. – V. 35. – P. 234-258.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

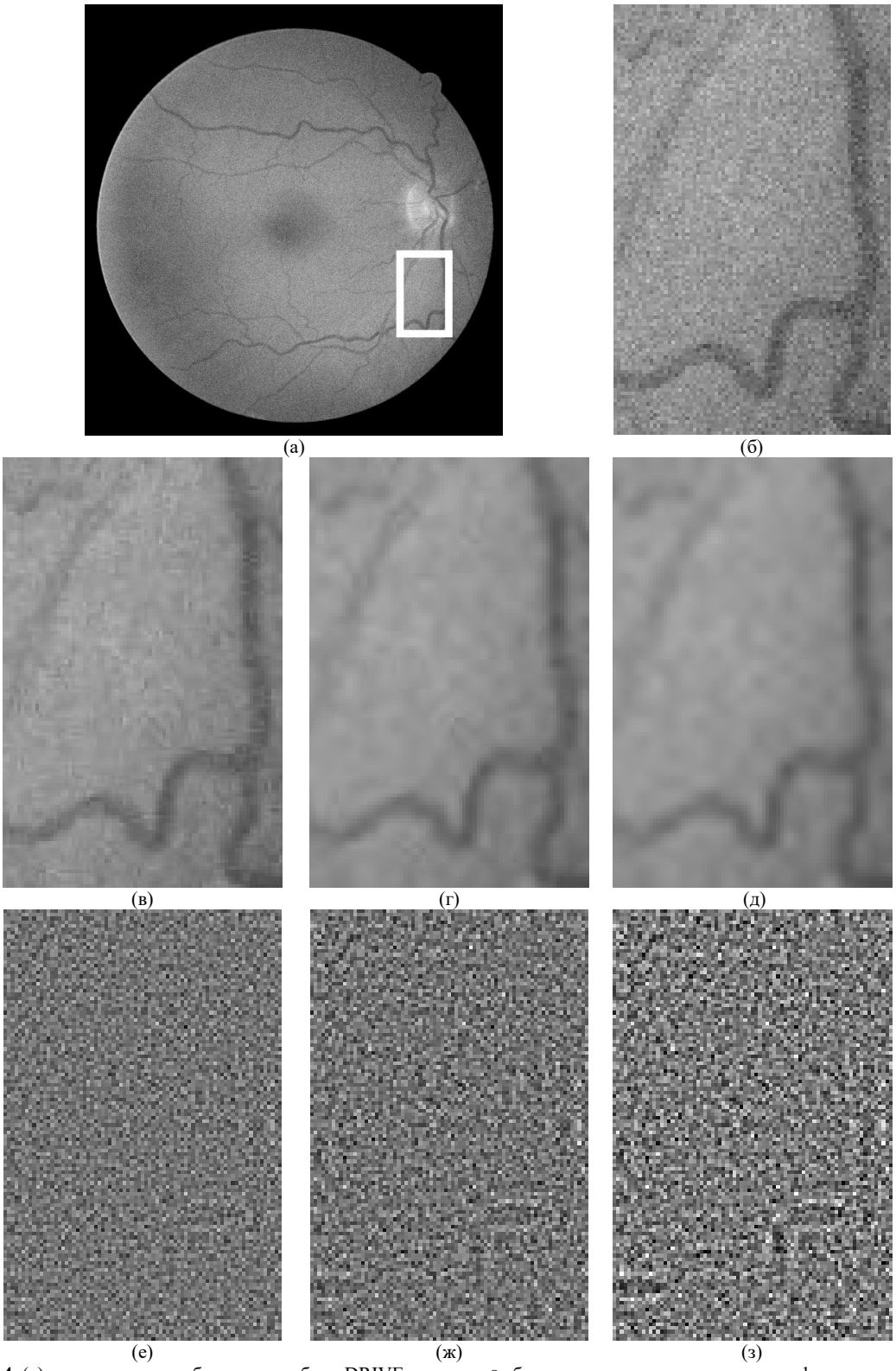


Рис. 4. (а) зашумленное изображение из базы DRIVE с $\sigma_{noise} = 8$, белым прямоугольником выделен фрагмент, показанный на (б), (в-д) фрагмент после применения метода шумоподавления с заниженным, оптимальным и завышенным параметром фильтрации, (е-з) соответствующие им разностные кадры.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Повышение качества изображений оптической когерентной томографии сетчатки глаза с использованием метода полной обобщённой вариации

А.В. Насонов[1], И.С. Лузин[1], А.А. Насонова[1], А.С. Крылов[1], Н.С. Семенова[2], В.С. Акопян[2]
kryl@cs.msu.ru
[1]Лаборатория математических методов обработки изображений
Факультет вычислительной математики и кибернетики
МГУ имени М.В.Ломоносова, Москва, Россия;
[2]Кафедра офтальмологии, факультет фундаментальной медицины
МГУ имени М.В.Ломоносова, Москва, Россия.

В работе предложен вариационный метод повышения резкости и подавления шума на трёхмерных изображениях оптической когерентной томографии сетчатки глаза. Метод основан на использовании функционала полной обобщённой вариации второго порядка в качестве стабилизатора в регуляризирующем функционале. По сравнению с обычным функционалом полной вариации, использование данного функционала позволяет устранить эффект ступенчатости. В работе также предложен численный метод минимизации регуляризирующего функционала.

Ключевые слова: *полная обобщённая вариация, повышение резкости, подавление шума, изображения оптической когерентной томографии.*

An enhancement of optical coherence tomography retinal images using total generalized variation

A.V. Nasonov[1], I.S. Luzin[1], A.A. Nasonova[1], A.S. Krylov[1], N.S. Semenova[2], V.S. Akopyan[3]
kryl@cs.msu.ru
[1]Laboratory of Mathematical Methods of Image Processing
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University, Moscow, Russia;
[2]Chair of Ophtalmology, Faculty of Fundamental Medicine
Lomonosov Moscow State University, Moscow, Russia.

A variational method for deblurring and denoising of optical coherence tomography retinal images is proposed. The method is based on using second order total generalized variation as a stabilizer term in a regularization functional. Compared to simple total variation, the use of total generalized variation can suppress staircase effect. The work also proposes a numerical scheme for regularization functional minimization.

Keywords: *total generalized variation, deblurring, denoising, optical coherence tomography images.*

1. Введение

Восстановление размытых и зашумлённых изображений является одной из основных обратных задач в области обработки изображений и компьютерного зрения.

Математически задача восстановления изображений ставится в виде обратной задачи нахождения исходного изображения v по известному размытому и зашумлённому изображению u:

$$u = Av + n, \tag{1}$$

где A — оператор размытия, обычно представляемый в виде операции свёртки с ядром размытия H: $v * H$, n — аддитивный шум.

Задача (1) является некорректно поставленной. Для нахождения её приближённого решения высокую эффективность показали вариационные методы, представимые в виде минимизации регуляризирующего функционала:

$$v_R = \arg\min_v J(v), \quad J(v) = F(v) + \alpha R(v), \tag{2}$$

где $F(v)$ — функционал, определяющий близость восстанавливаемого изображения к исходным данным, $R(v)$ — стабилизатор, накладывающие дополнительные ограничения на восстанавливаемое изображение, α — регуляризирующий параметр. Функционал $F(v)$ обычно задаётся в виде невязки

$$F(v) = \|Av - u\|^2 \tag{3}$$

с соответствующим выбором нормы $\|\cdot\|$. Как правило, используется обычная квадратичная норма.

Для задач восстановления изображений высокую эффективность показало использование функционала полной вариации, предложенное в работе [1]:

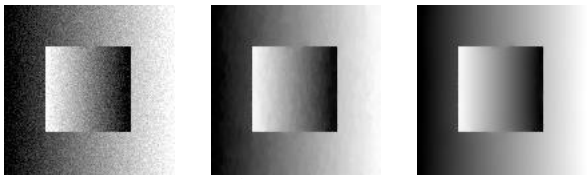
$$R(v) = \|\nabla v\|_1. \tag{4}$$

Преимуществом использования данного стабилизатора является сохранение контуров, а недостатком — потеря мелких деталей и приведение изображения к кусочно-постоянному виду при большом параметре регуляризации.

Несмотря на почти 30-летнюю историю, использование полной вариации остаётся актуальным и сейчас. Полная вариация используется, например, как функция потерь при обучении свёрточных нейронных сетей [2], в ретушировании [3], в восстановлении с помощью разреженных представлений [4]. В современных методах восстановления размытых изображений, основанных на использовании свёрточных нейронных сетей, машинное обучение используется для нахождения ядра размытия, тогда как непосредственно восстановление осуществляется с помощью регуляризирующих методов.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Идея использования производных более высокого порядка была предложена с целью устранения эффекта ступенчатости, присущего методам, основанным на использовании функционала полной вариации (см. Рис. 1).



а) зашумлённое изображение

б) подавление шума с использованием полной вариации

в) подавление шума с использованием полной обобщённой вариации

Рис. 1. Сравнение подавления шума с помощью обычной и обобщённой полной вариации.

Для исключения данного эффекта для решения задачи шумоподавления изображений, в работе [5] был предложен функционал со второй производной. Применение этой идеи в модифицированной форме было осуществлено в работе [6], где производная второго порядка заменена оператором Лапласа. Ещё один функционал был предложен в [7]. В работе [8] проведён анализ применения функционала полной обобщённой вариации для устранения эффекта ступенчатости. Регуляризирующие функционалы более высокого порядка для дискретного случая, непосредственно связанные с вычислениями производных второго порядка, были также рассмотрены Setzer S. and Steidl G. [9].

Значительный интерес представляет разработка методов обработки трёхмерных изображений. Подавляющее большинство существующих программных пакетов осуществляют обработку исключительно двумерных изображений. Адаптация двумерных алгоритмов к трёхмерным изображениям представляет большой теоретический и практический интерес.

В данной работе рассматривается задача подавления шума и повышения резкости трёхмерных изображений оптической когерентной томографии сетчатки глаза.

2. Полная обобщённая вариация

Функционал полной обобщённой вариации второго порядка представляется в виде [10]:

$$TGV_\alpha^2(v) = \min_w(\alpha_1\|\nabla v - w\|_1 + \alpha_2\|\nabla w\|_1). \qquad (5)$$

Здесь минимум берётся по всему векторному полю w, а положительные веса α_1 и α_2 являются параметрами функционала.

2.1 Устранение вспомогательной переменной

Одной из проблем, связанных с использованием стабилизатора TGV (5), является разработка эффективных алгоритмов минимизации регуляризирующего функционала (2). Для упрощения процедуры минимизации, мы устраняем вспомогательную переменную w и переходим к прямому вычислению вариации второй производной:

$$TGV_\alpha^2(v) = \alpha_1\|\nabla v\|_1 + \alpha_2\|\nabla(\nabla v)\|_1. \qquad (6)$$

2.2 Вычисление модуля первой производной

Ещё одной из проблем, связанных с использованием функционалов полной вариации и полной обобщённой вариации, является вычисление производных. Выбор способа численного дифференцирования оказывает большое внимание на результат. Например, производная в виде

$$|\nabla v|_{i,j,k} = \sqrt{v_x^2 + v_y^2 + v_z^2} =$$
$$= \sqrt{\left(v_{i+1,j,k} - v_{i,j,k}\right)^2 + \left(v_{i,j+1,k} - v_{i,j,k}\right)^2 + \left(v_{i,j,k+1} - v_{i,j,k}\right)^2}$$

приводит к несимметричной обработке симметричных объектов, тогда как симметричная производная

$$|v|_{i,j,k} = \sqrt{\left(\frac{v_{i+1,j,k} - v_{i-1,j,k}}{2}\right)^2 + \cdots + \left(\frac{v_{i,j,k+1} - v_{i,j,k-1}}{2}\right)^2}$$

не использует значение центрального пикселя $z_{i,j,k}$, что приводит на практике к появлению шумового паттерна в виде сетки.

Пользуясь тем, что нам важен суммарный модуль градиента, но не его компоненты, можно использовать следующее представление для модуля градиента:

$$|\nabla v|_{i,j,k} = [\frac{1}{2}\left(\left(v_{i+1,j,k} - v_{i,j,k}\right)^2 + \left(v_{i-1,j,k} - v_{i,j,k}\right)^2\right) +$$
$$+\frac{1}{2}\left(\left(v_{i,j+1,k} - v_{i,j,k}\right)^2 + \left(v_{i,j-1,k} - v_{i,j,k}\right)^2\right) +$$
$$+\frac{1}{2}\left(\left(v_{i,j,k+1} - v_{i,j,k}\right)^2 + \left(v_{i,j,k-1} - v_{i,j,k}\right)^2\right)]^{1/2}.$$

Использование полной вариации в таком виде вычислительно неэффективно из-за вычисления корня и операции деления при вычислении производной регуляризирующего функционала. Вычислительно более эффективным является следующий способ вычисления вариации:

$$|\nabla v|_{i,j,k} = \frac{1}{2}\left(\left|v_{i+1,j,k} - v_{i,j,k}\right| + \left|v_{i-1,j,k} - v_{i,j,k}\right| + \cdots\right),$$

однако он приводит к неоднородной обработке в зависимости от направления. Для устранения этого эффекта в работе [11] было предложено использовать функционал билатеральной вариации, представляющий собой взвешенную сумму модулей производных по направлению. В трёхмерном случае данный функционал записывается в виде:

$$|\nabla v|_{i,j,k} = \sum_{x,y,z=-p}^{p} \frac{1}{\sqrt{x^2 + y^2 + z^2}}\left|z_{i+x,j+y,k+z} - z_{i,j,k}\right|,$$
$$x^2 + y^2 + z^2 > 0.$$

Значение p задаёт размер окрестности, по которой вычисляются разностные производные. Обычно используются значения $p = 1$ (выше скорость) и $p = 2$ (выше качество). Использование значений, больших, чем 2, не приводит к дальнейшему повышению качества. Мы используем $p = 1$ из соображений вычислительной эффективности.

2.3 Вычисление модуля второй производной

Вычисление модуля второй производной предлагается осуществлять аналогично первой производной:

$$|\nabla(\nabla v)|_{i,j,k} = \sum_{x,y,z=-p}^{p} \frac{1}{\sqrt{x^2 + y^2 + z^2}}|v_{i+x,j+y,k+z} -$$
$$- 2v_{i,j,k} + v_{i-x,j-y,k-z}|, \quad x^2 + y^2 + z^2 > 0.$$

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

3. Метод минимизации

Мы используем метод Суцкевера-Нестерова для быстрой градиентной минимизации регуляризирующего функционала (2):

$$z^{(0)} = 0,\ v^{(0)} = 0$$
$$g^{(k+1)} = \nabla J\left(z^{(k)} + \mu v^{(k)}\right)$$
$$v^{(k+1)} = \mu v^{(k)} - \beta^{(k+1)} g^{(k+1)}$$
$$z^{(k+1)} = z^{(k)} + v^{(k+1)}$$

Число итераций N и шаг $\beta^{(k)}$ выбираются согласованно следующим образом:

$$\beta^{(k+1)} = 25 \cdot 0{,}01^{k/N} \cdot P / \left\| g^{(k+1)} \right\|_1,$$

где P — число пикселей (размер) изображения, а диапазон значений пикселей лежит в интервале $[0, 255]$.

Точность алгоритма зависит от числа итераций N и величине момента μ. Эксперименты показали, что минимально возможное число итераций N, при которых метод выдаёт результаты, визуально неотличимые от точной минимизации, равно 30, при этом $\mu = 0{,}8$.

4. Выбор регуляризирующих параметров

Соотношение между весами α_1 и α_2 в (6) определяет баланс весов первой и второй производных. При ограничении $\alpha_2 = 0$ функционал (6) представляет собой функционал полной вариации. При повышении α_2 с одновременным уменьшением α_1 происходит снижение эффекта ступенчатости. Начиная с некоторого момента, при больших значениях α_2 и малых значениях α_1, добавляется эффект размытия, и качество изображения снижается.

В связи с отсутствием трёхмерных референсных изображений оптической когерентной томографии сетчатки глаза, для их генерации применялся следующий подход:

1. К двумерным референсным изображениям из базы TID [13] были применены соответствующие реальным размытие и шум из рассматриваемого класса изображений. На полученных изображениях была произведена оптимизация параметров α_1 и α_2, причём оказалось, что наилучшие результаты достигаются при $\alpha_1 \approx \alpha_2$.

2. Предполагая, что параметры α_1 и α_2 остаются неименными при постоянных характеристиках размытия и шума, двумерный регуляризирующий алгоритм был применён к изображениями оптической когерентной томографии глаза по слоям.

3. Далее проводилась оптимизация масштабирующего параметра — нормировки для коэффициентов α_1 и α_2 при переходе к трёхмерному алгоритму, используя полученные на предыдущем шаге изображения как референсные.

5. Результаты

Для тестирования предложенного алгоритма были взяты трёхмерные изображения оптической когерентной томографии сетчатки глаза из базы [12].

Сравнение проводилось между описанным в работе трёхмерным алгоритмом и применением аналогичного двумерного алгоритма по срезам.

Примеры применения рассмотренных алгоритмов приведены на Рис. 2 и Рис. 3. На изображениях видно как преимущество полной обобщённой вариации перед простой полной вариацией, так и преимущество трёхмерного алгоритма перед двумерным. Использование обобщённой полной вариации позволяет избавиться от эффекта ступенчатости, присущего использованию функционала простой полной вариации. Трёхмерный алгоритм позволяет достичь лучшего шумоподавления за счёт того, что объёмный функционал полной обобщённой вариации захватывает больше пикселей, чем двумерных.

6. Заключение

Рассмотрена задача подавления шума и повышения резкости трёхмерных изображений оптической когерентной томографии сетчатки глаза. Для решения задачи был разработан трёхмерный регуляризирующий алгоритм со стабилизатором в виде функционала полной обобщённой вариации. Предложен эффективный численный метод минимизации регуляризирующего функционала. Применение алгоритма к реальным изображениям показало его эффективность и преимущество перед применением аналогичного двумерного алгоритма по срезам.

7. Благодарности

Работа выполнения при поддержке гранта РФФИ 15-29-03896.

8. Литература

[1] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms” // Physica D: nonlinear phenomena, vol. 60, no. 1-4, pp. 259–268, 1992.

[2] T. Wang, Z. Qin, M. Zhu, “An ELU network with total variation for image denoising” // International Conference on Neural Information Processing, 2017, pp. 227–237.

[3] M. V. Afonso, J. M. R. Sanches, “Blind inpainting using l0 and total variation regularization” // IEEE Transactions on Image Processing, vol. 24, no. 7, pp. 2239–2253, 2015.

[4] J. Liu, T.-Z. Huang, I. W. Selesnick, X.-G. Lv, P.-Y. Chen, “Image restoration using total variation with overlapping group sparsity” // Information Sciences, vol. 295, pp. 232–246, 2015.

[5] Chambolle A. and Lions P.-L. “Image recovery via total variation minimization and related problems” // Numerische Mathematik, 1997. Vol. 76, No 2, pp. 167–188.

[6] Chan T.F., Esedoglu S. and Park F.E. “A fourth order dual method for staircase reduction in texture extraction and image restoration problems” // International Conference on Image Processing, 2010, pp. 4137–4140.

[7] Chan T.F., Marquina A. and Mulet P. “Higher order total variation-based image restoration” // SIAM Journal on Scientific Computing, 2000. Vol. 22, No 2. pp. 503–516.

[8] Dal Maso G., Fonseca I., Leoni G. and Morini M. “A higher order model for image restoration: The one-dimensional case” // SIAM Journal on Mathematical Analysis, 2009. Vol. 40, No 6, pp. 2351–2391.

[9] Setzer S. and Steidl G. “Variational methods with higher order derivatives in image processing” // Approximation Theory XII, ed. by Neamtu M. and Schumaker L.L., Nashboro Press, Brentwood, TN, 2008. pp. 360–386.

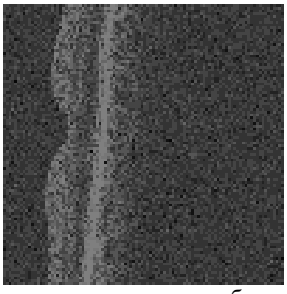
[10] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, “Second order total generalized variation (TGV) for MRI” // Magnetic resonance in medicine, vol. 65, no. 2, pp. 480–491, 2011.

[11] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super resolution” // IEEE Transactions on Image Processing, vol. 13, no. 10, pp. 1327–1344, 2004.

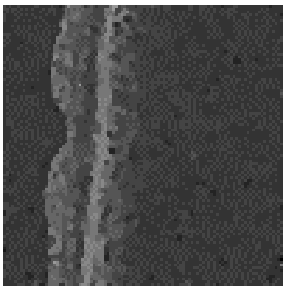
[12] M. Golabbakhsh, H. Rabbani, “Vessel-based registration of fundus and optical coherence tomography projection images of retina using a quadratic registration model” // IET Image Processing, 7(8), 768–776, 2013.

[13] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, C.-C. Jay Kuo, “Image database TID2013: Peculiarities, results and perspectives” // Signal Processing: Image Communication, vol. 30, 2015, pp. 57–77.

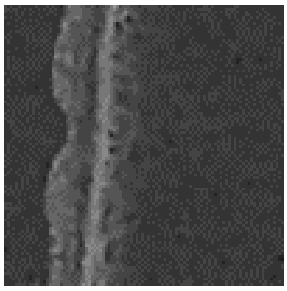
Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18



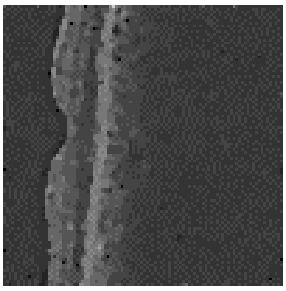
а) Фрагмент исходного изображения.



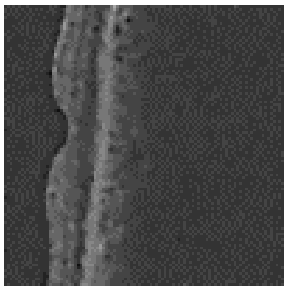
б) Результат применения двумерного алгоритма со стабилизатором в виде функционала полной вариации.
$\alpha = 1$



в) Результат применения двумерного алгоритма со стабилизатором в виде функционала полной обобщённой вариации.
$\alpha_1 = 0{,}5, \alpha_2 = 0{,}5$

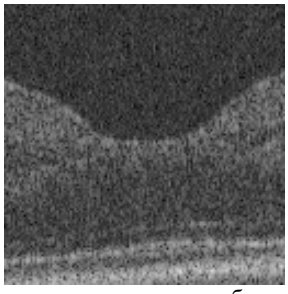


г) Результат применения трёхмерного алгоритма со стабилизатором в виде функционала полной вариации.
$\alpha = 0{,}2$

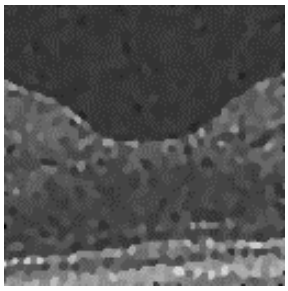


д) Результат применения предложенного трёхмерного алгоритма со стабилизатором в виде функционала полной обобщённой вариации.
$\alpha_1 = 0{,}1, \alpha_2 = 0{,}1$

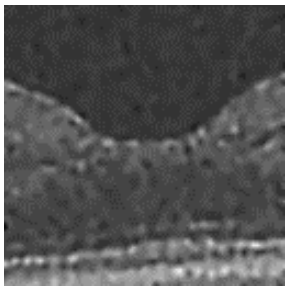
Рис. 2. Сравнение регуляризирующих алгоритмов подавления шума и повышения резкости для изображений оптической когерентной томографии сетчатки глаза. Ядро размытия — фильтр Гаусса с параметром $\boldsymbol{\sigma = 1}$.



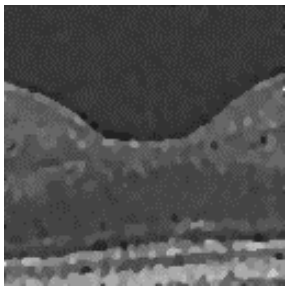
а) Фрагмент исходного изображения.



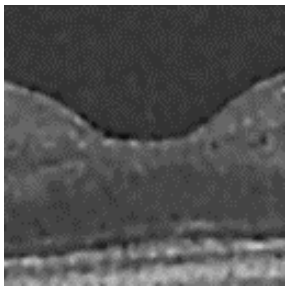
б) Результат применения двумерного алгоритма со стабилизатором в виде функционала полной вариации.
$\alpha = 1$



в) Результат применения двумерного алгоритма со стабилизатором в виде функционала полной обобщённой вариации.
$\alpha_1 = 0{,}5, \alpha_2 = 0{,}5$



г) Результат применения трёхмерного алгоритма со стабилизатором в виде функционала полной вариации.
$\alpha = 0{,}2$



д) Результат применения предложенного трёхмерного алгоритма со стабилизатором в виде функционала полной обобщённой вариации.
$\alpha_1 = 0{,}1, \alpha_2 = 0{,}1$

Рис. 3. Сравнение регуляризирующих алгоритмов подавления шума и повышения резкости для изображений оптической когерентной томографии сетчатки глаза. Ядро размытия — фильтр Гаусса с параметром $\boldsymbol{\sigma = 1, 5}$.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Реконструкция поверхности по точкам контуров в ортогональных сечениях томограммы

М.М. Новожилов[1], М.В. Дубровская[1]
michael.novogilov@gmail.com|maria.v.dubrovskaya@gmail.com
1 Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия

*В настоящее время существует много средств и методов решения задачи реконструкции полигональной поверхности по контурам, образованным при сечении поверхности параллельными плоскостями. Использование этого арсенала для реконструкции анатомических изоповерхностей по данным томографии предполагает послойное формирование последовательности параллельных контуров. Последующие этапы: детектирование когерентных контуров и триангуляция поверхности. Для оценки когерентности предложено строить контуры в плоскостях, ортогональных первоначальным контурам, если исходные данные позволяют выполнить такое построение. Задача поиска точек пересечения контуров, в общем случае, имеет сложность O(N*N), где N - количество точек. Показано, как можно решить данную задачу за линейное время, предложен эффективный алгоритм триангуляции полученного графа поверхности. Работоспособность алгоритма демонстрируется на примере реконструкции поверхности сердца.*

Ключевые слова: *обработка данных томографии, реконструкция изоповерхности, контур, граф.*

Surface reconstruction by contour points in orthogonal cross-sections of tomogram

M.M. Novozhilov[1], M.V. Dubrovskaya[1]
michael.novogilov@gmail.com|maria.v.dubrovskaya@gmail.com
[1] Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation

*Nowadays there are a lot of tools and methods for solving the problem of polygonal surface reconstruction from contours, created by surface cut using parallel planes. The use of this arsenal for the reconstruction of anatomical isosurfaces from tomography data suggests the layerwise formation of a sequence of parallel contours. Subsequent steps: detection of coherent circuits and triangulation of the surface. To estimate the coherence, it is proposed to construct contours in planes orthogonal to the original contours, if the initial data allow such a construction. The problem of intersection points search in common case has complexity O(N*N), where N is a number of points. It is shown how it is possible to solve this problem in linear time, an effective algorithm for triangulation of the obtained surface graph is proposed. The efficiency of the algorithm is demonstrated by the example of reconstruction of the heart surface.*

Keywords: *Surface reconstruction, contour, graph.*

1. Введение

Задача реконструкции поверхности по контурам является широко распространённой в областях, связанных с геометрическим моделированием и обработкой трехмерных медицинских данных, прежде всего данных томографии. Исходные контуры должны быть заданы набором точек, или получены в результате их сегментации на изображениях слоев томограммы. В работе [1] приводится обзор базовых подходов к сегментации медицинских изображений, в результате применения которых исследуемый орган помечается определённым цветом, что позволяет реконструировать на каждом срезе контур по границе органа.

Существует много публикаций, предлагающих общее решение задачи реконструкции поверхности по контурам. В работе [2] исходный набор контуров представлен в виде последовательности бинарных изображений, которые преобразуются в набор точек в R^3. Далее, для вычисления нормали в каждой точке создаётся бинарный объём, ограниченный исходными контурами, применяется фильтр Гаусса и в точках исходных контуров вычисляются градиенты. Таким образом, задача реконструкции поверхности по контурам сводится к более общей задаче реконструкции поверхности по точкам.

В работе [3] предлагается гладкое решение проблемы сшивки когерентных контуров, заданных последовательностью их точек. Также, приводится подробный обзор решений проблем: соответствия точек когерентных контуров; триангуляции (tiling) поверхности на когерентных контурах; ветвления контуров.

Алгоритм [4], подобно [2], также принимает на вход последовательность бинарных изображений. В данном методе, для улучшения качества поверхности выполняется построение промежуточных контуров посредством адаптируемой схемы уточнения. При этом, на каждой стадии реконструкции используется большее двух сечений, что отличает данный метод от большинства других. В результате, как и в методе [2], получается гладкая поверхность.

В работе [5] приводится алгоритм разрешения неоднозначности в контуре (образование "замочных скважин"), а так же шаблон для триангуляции быстрых переходов.

В случае, когда медицинские изображения образованы данными компьютерной томограммы достаточного разрешения, реконструкцию поверхности можно выполнить методами, базирующимися на марширующих кубах [6, 7]. Данные методы разделяют пространство на кубические ячейки и реконструируют поверхность отдельно в каждой из них. Однако в данных работах не рассматривается проблема объединения отдельных поверхностей в единую связную поверхность, что может привести к существенным вычислительным затратам. Данная проблема решается алгоритмом, предложенным в работе [8]. В данном методе рассматриваются регулярные сетки томограммы, ортогональные направлениям X, Y и Z. Для каждого ребра сетки проверяется условие его пересечения с поверхностью, и в случае его выполнения

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

фиксируется точка пересечения. Затем, на построенных точках формируется контур и объединяется с другими, находящимися в ортогональных плоскостях. Таким образом, формируется граф поверхности (рис. 1).

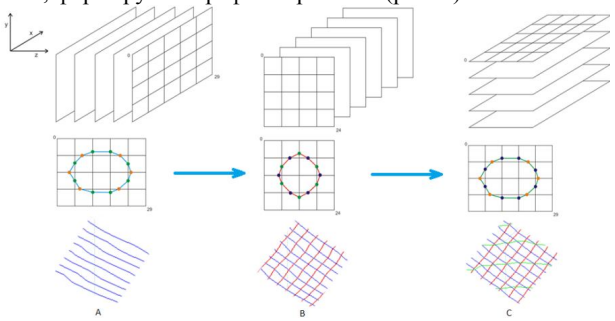


Рис. 1. Построение графа поверхности.

Преимущество данного подхода в том, что он эффективно решает проблему объединения поверхностей в соседних кубических ячейках, и при этом имеет линейное ускорение при распараллеливании. Однако недостатком приведённого алгоритма, а также базирующихся на марширующих кубах методах является наличие неоднозначностей при реконструкции поверхности, поскольку поверхность восстанавливается локально в каждой области.

Один из способов разрешения неоднозначностей - реконструкция контура поверхности целиком (глобально) на срезе, исходя, например, из его геометрических свойств, особенностей изображения или заданной модели.

В данной работе предлагается подход, позволяющий глобально восстановить контуры поверхности в ортогональных плоскостях, и в то же время обладающий линейным ускорением при построении поверхности в параллельной реализации. Кроме того, данный алгоритм может использоваться в сочетании с подходами [2,4] в ортогональных плоскостях. При этом решается проблема соответствия точек и триангуляции.

2. Описание метода получения контуров

Для сегментации и выделения контуров в данной работе применяется метод Ray Casting [9] на каждом слое данных компьютерной томограммы. Одним из возможных решений здесь является испускание лучей радиально из области внутри сердца, близкой к его центру. Критерием выбора точек контуров является переход в этой точке, при движении вдоль луча, к отрицательным значениям (к перикарду, жировому мешочку толщиной порядка нескольких миллиметров, плотно окружающему сердце и основание прилегающих крупных сосудов).

Лучевой метод в качестве основы алгоритма построения контуров выбран благодаря его ориентации на параллельные вычисления и, как следствие, приемлемому времени работы по сравнению с аналогами. Более того, сегментация лучевым методом является в значительной степени автоматической и, следовательно, почти не требует сопровождения. При всей простоте, Ray Casting показывает результаты приемлемые для дальнейшей обработки.

3. Описание метода реконструкции поверхности

На вход алгоритма поступают контуры и кривые заданные последовательностью своих точек в R^3, и как бы образованные при сечении реконструируемой поверхности ортогональными плоскостями. Предлагаемый подход основан на алгоритме [8], который принимает на вход данные компьютерной томограммы, строит контуры изоповерхности в ортогональных плоскостях и за линейное время выполняет поиск точек их пересечений. Однако для применения этого алгоритма в задаче реконструкции поверхности по контурам не обязательно выполнять построение скалярного поля. Достаточно спроектировать каждый контур на регулярную сетку необходимого разрешения, как показано на рис. 2.

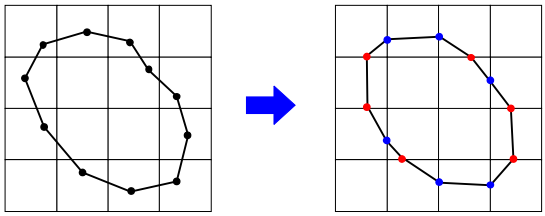


Рис. 2. Проекция контура на регулярную сетку.

При этом, в процессе обхода вершин текущего контура, для каждого его ребра, вычисляются точки пересечения с горизонтальными и вертикальными линиями сетки. Вычисление исходного контура происходит на одном этапе с обработкой соответствующей плоскости, что не нарушает масштабируемости алгоритма при параллельном исполнении.

Однако исходные контуры, вычисляемые в ортогональном направлении, могут не пересекаться с первоначальными контурами на одном ребре регулярной сетки, что необходимо для идентификации точки пересечения. Поэтому процедура дополняется возможностью, для соответствующего первоначального контура, рассматривать также соседние рёбра регулярной сетки, на предмет наличия точки контура в ортогональном направлении. Затем, точки полученного контура смещаются в целях его сглаживания и соответствия рёбрам для первоначальных контуров. Сглаженный контур снова проектируется на регулярную сетку, для возможности дальнейшего эффективного получения графа поверхности.

После получения графа, выполняется перебор его граней и их триангуляция. Алгоритм состоит из следующих шагов:

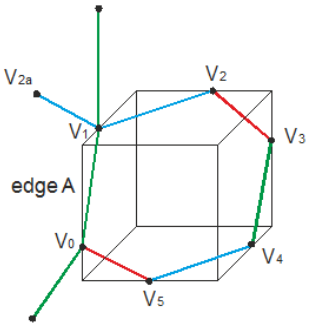


Рис. 3. Триангуляция графа поверхности.

1. Рассматриваются контуры поверхности, принадлежащие одной плоскости;
2. На текущем контуре, рассматриваются его рёбра (ребро *A* на рис. 3);
3. Вершины графа поверхности добавляются в список вершин грани. Вершины V_0 и V_1 можно добавить единственным образом;
4. Далее, мы рассматриваем две грани, вершины которых начинаются с $\{V_0, V_1, V_2\}$ и $\{V_0, V_1, V_{2a}\}$. На всех стадиях формирования грани мы добавляем в список вершину, ребро которой лежит на плоскости, ортогональной к плоскости рассматриваемого ребра;
5. Необходимо проверить существование грани, инцидентной первым трём добавленным вершинам

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

(например, посредством проверки списка граней для каждой вершины). В случае существования, обход грани завершается.

6. Добавляем следующую вершину V_n, лежащую на той же стороне от плоскости ребра (V_{n-1}, V_{n-2}) что и вершина V_{n-3};
7. Если V_n совпадает с V_0, обход грани успешно завершён. Иначе, алгоритм завершается, если количество итераций больше шести.
8. В результате, получаем грань степени от 3 до 6, которая может быть легко триангулирована.

4. Заключение

Для получения исходного набора контуров мы воспользовались методом трассировки луча, применяя его к данным компьютерной томограммы сердца. В результате работы получена связная сетка поверхности (рис. 4), и ускорение в параллельной реализации на процессоре Intel®Core™i5 750 2.66 GHz, для поверхности из 477057 примитивов, оказалось близким к линейному:

Стадия	Время (сек.)		
	1 ядро	2 ядра	4 ядра
Сегментация	20,96	10,80	5,72
Построение графа	5,64	2,37	1,32
Триангуляция	1,01	0,89	0,77

Некоторые области оказались не сегментированными, из-за чего в поверхности возникли отверстия. Тем не менее, данный подход, в отличие от упомянутых ранее, позволяет разрешить неоднозначности, например, при отделении поверхности сердца от диафрагмы. Планируется продолжить развитие данного метода сегментации.

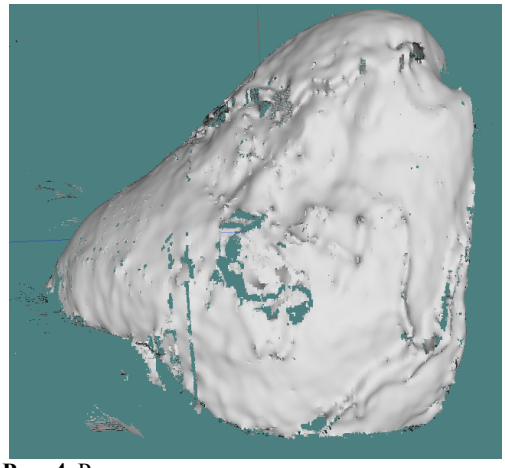


Рис. 4. Реконструированная поверхность сердца.

5. Литература

[1] Alireza Norouzi, Mohd Shafry Mohd Rahim, Ayman Altameem, Tanzila Saba, Abdolvahab Ehsani Rad,Amjad Rehman & Mueen Uddin (2014) Medical Image Segmentation Methods, Algorithms, and Applications, IETE Technical Review, 31:3, 199-213.

[2] Braude, I., Marker, J., Museth, K., Nissanov, J., Breen, D.: Contour-based surface reconstruction using MPU implicit models. GRAPHICAL MODELS 69, 2007.

[3] Nurzyńska K. (2009) 3D Object Reconstruction from Parallel Cross-Sections. In: Bolc L., Kulikowski J.L., Wojciechowski K. (eds) Computer Vision and Graphics. ICCVG 2008. Lecture Notes in Computer Science, vol 5337. Springer, Berlin, Heidelberg.

[4] Shay Kels, Nira Dyn, Reconstruction of 3D objects from 2D cross-sections with the 4-point subdivision scheme adapted to sets, Computers & Graphics, Volume 35, Issue 3, 2011, Pages 741-746.

[5] Sunderland, Kyle; Woo, Boyeong; Pinter, Csaba; Fichtinger, Gabor. / Reconstruction of surfaces from planar contours through contour interpolation. Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling. Vol. 9415 SPIE, 2015.

[6] Lewiner T. et al, 2003. Efficient Implementation of Marching Cubes' Cases with Topological Guarantees. *Graphics Tools*. Vol. 8, No. 2. pp. 1-15.

[7] Sundaresan Raman and Rephael Wenger, May 26-28, 2008. Quality Isosurface Mesh Generation Using an Extended Marching Cubes Lookup Table. *Proceedings of the 10th Joint Eurographics / IEEE - VGTC conference on Visualization.* Eindhoven, The Netherlands, pp. 791-798.

[8] Новожилов М.М. Реконструкция структурированных изоповерхностей по 3D данным компьютерной томографии. В сборнике: ГРАФИКОН'2016 Труды 26-й Международной научной конференции. 2016. С. 185-187.

[9] Gavrilov, N. I., Turlapov, V .E. Novel approach to development of direct volume rendering algorithms based on visualization quality assessment // Programming and Computer Software, (2014) 40 (4), pp. 174-184. DOI: 10.1134/S0361768814040045.

Об авторах

Михаил Михайлович Новожилов – аспирант ИТММ ННГУ, michael.novogilov@gmail.com.

Мария Владимировна Дубровская – аспирант ИТММ ННГУ, maria.v.dubrovskaya@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Машинное обучение по текстурным признакам в задаче детектирования нейронов на срезах мозга, окрашенных по Нисслю

С.А. Носова, В.Е. Турлапов
svetlana.nosova@itmm.unn.ru | vadim.turlapov@itmm.unn.ru
Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия

Рассматривается задача детектирования и классификации нейронов на срезах мозга мыши окрашенных по Нисслю в оптической микроскопии. Исследуется применение в этой задаче алгоритмов машинного обучения в сочетании с использованием текстурных признаков объекта. Предложенный метод включает в себя два основных шага: классификацию пикселей и детектирование центров нейронов. На этапе классификации при помощи алгоритма k ближайших соседей (kNN) определяется принадлежность каждого пикселя телу нейрона. В работе исследованы 3 типа текстурных признаков, построенных на основе: нормализованной гистограммы; структурного тензора; их комбинации. Для нахождения центров нейронов все пиксели, принадлежащие нейронам, кластеризуются при помощи алгоритма сдвига среднего (mean shift). Эксперименты показали, что наилучшее качество детектирования достигнуто при использовании комбинированных текстурных признаков (по гистограмме+структурный тензор) с радиусом окрестности R=7 со следующими показателями качества: recall = 0,77; precision =0,79; F1 score = 0,78. На рассматриваемом наборе данных выбор параметра k>15 не дает существенного повышения качества детектирования. Показано, что kNN является наилучшим классификатором для предложенного алгоритма детектирования по сравнению с байесовским классификатором (NBC) и машиной опорных векторов (SVM).

Ключевые слова: *обработка изображений, срез мозга, оптическая микроскопия, окрашивание по Нисслю, нейроны мозга, детектирование, классификация, кластеризация, машинное обучение, kNN, текстурный анализ, гистограмма, структурный тензор*

Machine learning by textural features in the task of detecting neurons on brain slices stained by Nissl

S.A. Nosova, V.E. Turlapov
svetlana.nosova@itmm.unn.ru | vadim.turlapov@itmm.unn.ru
Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation

The problem of detecting and classifying neurons on mouse brain sections stained by Nissl in optical microscopy is considered. The application in this problem of machine learning algorithms in combination with the use of textural features of the object is investigated. The proposed method includes two main steps: the classification of pixels and the detection of neuron centers. At the stage of classification using the k nearest neighbors (kNN) algorithm determines the belonging of each pixel to the neuron's body. In the paper, 3 types of textural features were constructed, based on: a normalized histogram; structural tensor; their combination. To find the centers of neurons, all the pixels belonging to the neurons are clustered using the mean shift algorithm. The experiments showed that the best detection quality was achieved using combined textural features (histogram + structural tensor) with a radius of neighborhood R = 7 with the following quality indicators: recall = 0.77; precision = 0.79; F1 score = 0.78. On the data set under consideration, the choice of the parameter k> 15 does not significantly improve the detection quality. It is shown that kNN is the best classifier for the proposed detection algorithm in comparison with the Bayesian classifier (NBC) and the support vector machine (SVM).

Keywords: *image processing, brain slice, optical microscopy, Nissl staining, brain neurons, detection, classification, clustering, machine learning, kNN, textural analysis, histogram, structural tensor.*

1. Введение и обзор литературы

Первыми задачами обработки изображений срезов коры мозга, окрашенных по Нисслю, являются: определение границ слоев коры мозга; детектирование и классификация нейронов, сбор статистики о расположении нейронов. Стоит отметить высокую техническую сложность процедуры окрашивания по Нисслю и, как следствие, отсутствие для него формализованной процедуры классификации нейронов и астроцитов.

В данной работе рассматривается задача автоматизации детектирования нейронов. Выбор методов решения задачи зависит от рассмотрения двух вопросов: особенности нейронов как объектов на изображении и качества тренировочных данных (знаний), полученных от специалиста.

Задача детектирования нейронов осложняется следующими особенностями:

1. Нейроны различны по форме.
2. Размеры нейронов на одном изображении могут различаться более, чем в 2 раза.
3. Изображения тел нейронов часто перекрываются
4. Изображения тел нейронов могут иметь низкую контрастность с фоном.
5. Изображения нейронов одного типа могут иметь значительные различия в гистограмме.
6. Изображения нейронов могут иметь “пустоты” внутри своих тел.

С точки зрения качества решения задач детектирования и классификации, наилучшим вариантом является большой размер тренировочной базы, когда потрачены значительные ресурсы на качественную разметку объектов (десятки-сотни сложных объектов, тысячи простых). В этом случае для решения задач детектирования сегодня часто используются различные варианты глубокого обучения [1] или комплексные методы машинного обучения [5].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

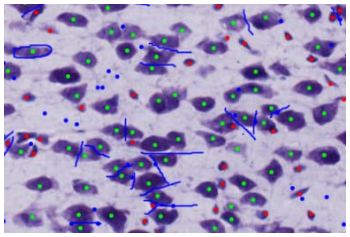


Рис. 1. Пример разметки специалистом. Зеленый цвет – центр нейрона, красный цвет – другие объекты, синий цвет – фон и точки, не являющиеся центрами нейронов

В случае меньших ресурсов на разметку и, соответственно при меньшем объеме тренировочной базы и задача традиционно решается методами машинного обучения, в том числе, k ближайших соседей (kNN) [10], машина опорных векторов (SVM) [7].

Когда образцы разметки единичны или отсутствуют, приходится производить довольно обширное исследование признаков, обеспечивающих приемлемую надежность детектирования и классификации объектов. В этом случае для исследования применяются классические методы обработки изображений, например, методы водораздела [9], математической морфологии, поиск по шаблону и др. [8].

Исследование в данной работе относится ко второму (среднему) случаю. Пример разметки приведен на Рис. 1.

Предложенный метод состоит из двух этапов: классификация пикселей, принадлежащих телам нейронов и нахождение центров кластеров пикселей. На этапе классификации пикселей используется метод k ближайших соседей (kNN) [3]. Каждый пиксель описывается вместе с его окрестностью радиуса R набором текстурных характеристик, которые вычисляются: по гистограмме [13], на основе структурного тензора [11] и комбинации этих двух характеристик.

Для нахождения центров кластеров пикселей используется алгоритм сдвига среднего (mean shift) [4]. Для оценки качества детектирования использованы оценки, описанные в [6].

Для тестирования алгоритма использован набор данных “Mus Musculus” портала BrainMaps [2]. В работе использованы реализации классификаторов библиотеки OpenCV [12].

2. Текстурные признаки

Мы используем текстурные признаки окрестности каждого пикселя для его классификации как пикселя фона или пикселя нейрона. Удобно задавать окрестность пикселя радиусом, но можно также описывать ее как прямоугольную область I изображения размером $N\times N$. Здесь и далее $N=2*R+1$. Вычисляются следующие текстурные характеристики:

1. На основе *гистограммы*. Строится гистограмма $p(I)$ - распределение интенсивности изображения размером $N\times N$ с центром в текущем пикселе. Вычисляются следующие характеристики гистограммы:

$$mean=\sum_{i=0}^{255} i*p_i(I); \qquad m2=\sum_{i=0}^{255}(i-mean)^2 p_i(I);$$

$$m3=\sum_{i=0}^{255}(i-mean)^3 p_i(I); \quad m4=\sum_{i=0}^{255}(i-mean)^4 p_i(I);$$

$$stddev=\sqrt{m2}; \quad skewness=\frac{m3}{stddev^3}; \quad kurtosis=\frac{m4}{m2^2}$$

Вектор <*stddev, skewness, kurtosis*> образует текстурную классификационную характеристику пикселя.

2. На основе *структурного тензора*. Для вычисления текстурной характеристики необходимо:
   a. вычислить значения производных I_x и I_y для исходного изображения:
   $$I_x(i,j)=I(i,j+1)-I(i,j-1),\ 0\le i\le N,\ 0\le j\le N$$
   $$I_y(i,j)=I(i+1,j)-I(i-1,j),\ 0\le i\le N,\ 0\le j\le N$$
   b. найти средние значения
   $$\overline{I_xI_x}=\frac{1}{N\cdot N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}I_x(i,j)*I_x(i,j)$$
   $$\overline{I_yI_y}=\frac{1}{N\cdot N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}I_y(i,j)*I_y(i,j)$$
   $$\overline{I_xI_y}=\frac{1}{N\cdot N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}I_x(i,j)*I_y(i,j)$$

   Вектор $\left\langle \overline{I_xI_x}, \overline{I_xI_y}, \overline{I_yI_y} \right\rangle$ образует текстурную характеристику.

3. *Составная* текстурная характеристика (вектор признаков) строится из двух предшествующих:
   $$\left\langle stddev, skewness, kurtosis, \overline{I_xI_x}, \overline{I_xI_y}, \overline{I_yI_y} \right\rangle$$

Для все текстурных характеристик выполняется нормировка на основе значений, полученных для объектов из тренировочной базы.

3. Шаг 1. Метод k ближайших соседей

Для предварительной классификации изображения, отделения пикселей, относящихся к нейронам, от всех остальных пикселей, использован метод машинного обучения *k* ближайших соседей (*kNN, k Nearest Neighbor*). Этот метод является одним из самых интуитивно понятных алгоритмов классификации. Так как мы используем вектор признаков, то в процедуре *kNN* применяется следующее правило: для каждого используемого набора признаков в качества класса выбирается тот класс, за который «проголосовало» наибольшее количество тестовых признаков из *k* ближайших соседей. В качестве метрики расстояния использована Евклидова метрика.

Пример предварительной разметки приведен на Рис.1. Для каждого пикселя (x,y) разметки вычисляется набор текстурных признаков для области изображения с центром в (x,y) и радиусом R. Результат заносится в тренировочную базу.

4. Шаг 2. Метод сдвига среднего

На данном этапе происходит непосредственное вычисление центров кластеров как центра тяжести пикселов, отнесенных к классу «нейрон» в окрестности радиуса R пикселя x. Процедура поиска центра кластера, к которому относится пиксель x, заключена в следующем:

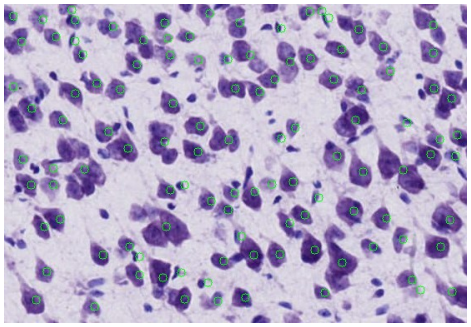
1. Положить, что x – центр кластера.
2. Вычислить положение $m(x)$ средней плотности пикселей, детектированных как нейрон, из окрестности R текущего пикселя по формуле ниже.

$$K(x)=\begin{cases}1, if\ \|x\|\le R\\ 0, if\ \|x\|>R\end{cases};\quad m(x)=\frac{\sum_{s\in S}K(s-x)s}{\sum_{s\in S}K(s-x)}$$

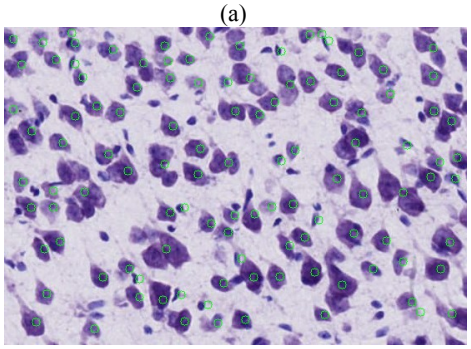
где s – пиксель из окрестности x, принадлежащий нейрону; S – множество пикселей, детектированных как нейроны, из окрестности радиуса R пикселя x.

3. Вычислить сдвиг среднего $m(x) - x$. Если $|m(x) - x|>\varepsilon$, *п*ереместить x в точку $m(x)$ и повторить процедуру, начиная с п.2. В противном случае перейти к следующему пикселю.

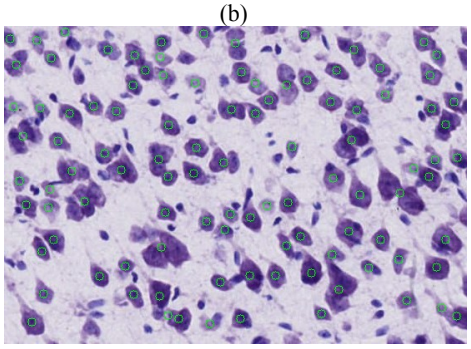
Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18



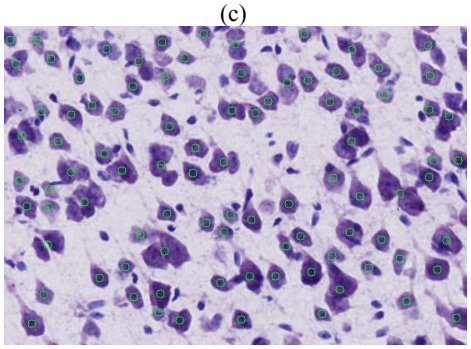
(a)



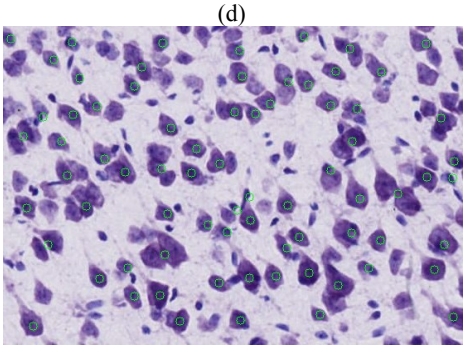
(b)



(c)

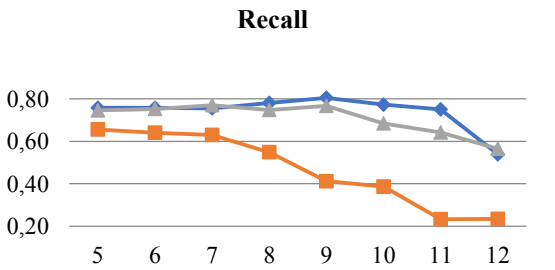


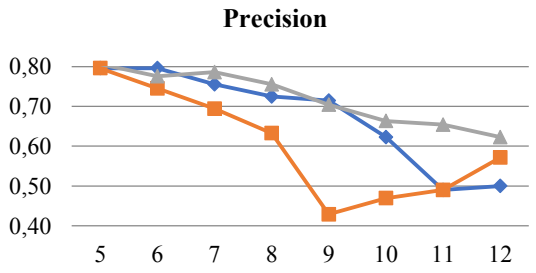
(d)

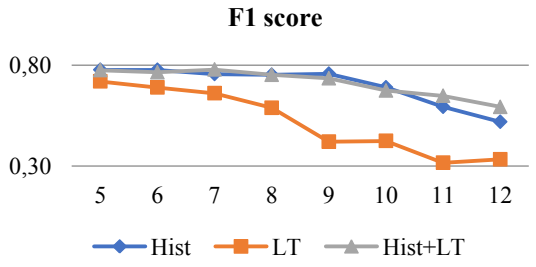


(e)

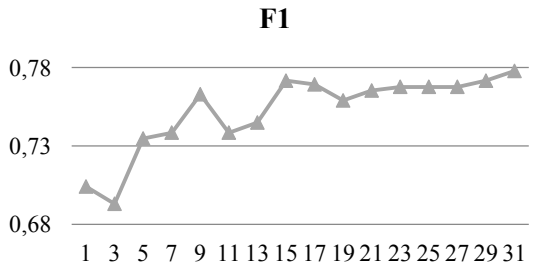
Рис. 3. Результаты детектирования центров нейронов алгоритмом kNN при использовании текстурных характеристик: (a) - по гистограмме, (b) - структурный тензор, (c) – комбинированная; результаты детектирования при использовании комбинированной текстурной характеристики и алгоритма классификации пикселей: (d)–NBC, (e) - SVM



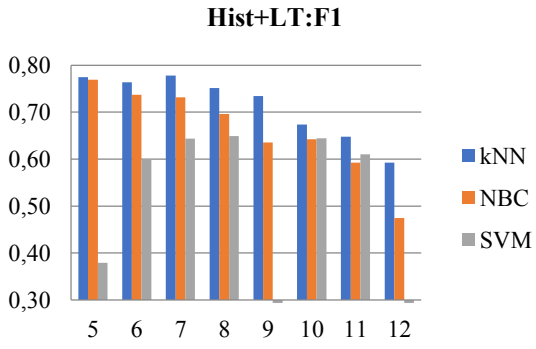




(a)



(b)



(c)

Рис. 2. (a) - показатели качества алгоритма детектирования нейронов; (b) – зависимость качества детектирования от выбора параметра k в алгоритме kNN; (c) - характеристика F1 score результатов детектирования методом kNN, NBC, SVM (с радиальным ядром и максимальным количеством итераций 20)

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

5. Оценка качества детектирования

Для оценки качества детектирования необходимо вычислить следующие характеристики: количество верных детектирований центров нейронов (*TP, true positive*), количество ложных детектирований центров нейронов (*FP, false positive*) и количество недетектированных центров нейронов (*FN, false negative*) из тестовой базы. Далее вычисляются следующие оценки качества детектирования:

$$\mathrm{Re}call = \frac{TP}{TP + FP}$$

$$\mathrm{Pr}ecision = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * \mathrm{Re}call * \mathrm{Pr}ecision}{\mathrm{Re}call + \mathrm{Pr}ecision}$$

Для того, чтобы определить, к какому классу (*TP,FP,FN*) относится центр, применяется следующая процедура для всех детектированных центров нейронов c_i и размеченных специалистом центров нейронов m_j:

1. Положить *TP = 0, FP = 0, FN = 0.*
2. Для каждого детектированного центра c_i:
   a. Найти ближайший m_j
   b. Если $|c_i - m_j| < R_{acc}$, где R_{acc} – средний радиус нейрона; то увеличить *TP* на 1 и удалить m_j и c_i из рассмотрения. В противном случае, увеличить *FP* на 1.
3. Для каждого центра m_i:, для которого не был найден c_i:
   a. Увеличить *FN* на 1.

6. Результаты

Результаты работы алгоритма для разных наборов текстурных признаков представлены на Рис. 3. Зависимости качества детектирования от выбранного радиуса *R* оценки текстурных характеристик представлены на Рис. 2(a).

Зависимость качества детектирования от выбора параметра k представлена на Рис. 2(b). Результаты измерения качества детектирования для *R=7* при использовании классификтора *kNN* при *k=31*, байесовского классификатора для величин, распределенных по нормальному закону (NBC, [14]) и машиной опорных векторов (SVM, [14]) с радиальным ядром и остановкой по максимальному количеству итераций 20 представлены на Рис. 2(c).

7. Заключение

Предложен алгоритм для детектирования центров нейронов на основании принадлежности пиксела к изображению нейрона по текстурным признакам. Алгоритм состоит из двух основных шагов: классификации пикселей при помощи алгоритма kNN (препроцессинг) и применение алгоритма mean shift для кластеризации полученных пикселей (непосредственно фаза детектирования нейрона).

Эксперименты показали, что наилучшее качество детектирования достигается при использовании комбинированной текстурной характеристики (по гистограмме+структурный тензор) с использованием радиуса рассматриваемой окрестности R=7 со следующими показателями качества: recall = 0,77; precision =0,79; F1 score = 0,78. Сравнение результатов использования алгоритмов kNN, NCB, SVM для классификации пикселов показало, что наилучшее качество детектирования нейронов достигается при использовании kNN. На рассматриваемом наборе данных выбор параметра k>15 не дает существенного повышения качества детектирования.

Качество детектирования является приемлемым для задач подсчета количества нейронов в различных участках мозга, для особей различного возраста. Тем не менее, вопрос оптимальности выбора текстурных признаков, по критерию качества детектирования, остается открытым. Планируется исследование текстурных признаков Лавса и текстурных признаков, построенных на матрице взаимной встречаемости.

8. Литература

[1] Akram S.U., Kannala J., Eklund L., Heikkilä J. , 2016. Cell Segmentation Proposal Network for Microscopy Image Analysis. In: Carneiro G. et al. (eds) Deep Learning and Data Labeling for Medical Applications. LABELS 2016, DLMIA 2016. Lecture Notes in Computer Science, vol 10008. Springer, Cham. doi: https://doi.org/10.1007/978-3-319-46976-8_3

[2] BrainMaps: An Interactive Multiresolution Brain Atlas; http://brainmaps.org [retrieved on 15.05.2018]

[3] Bo Yang Yu, Caglar Elbuken, Carolyn L. Ren, Jan Paul Huissoon, "Image processing and classification algorithm for yeast cell morphology in a microfluidic chip," Journal of Biomedical Optics 16(6), 066008 (1 June 2011). https://doi.org/10.1117/1.3589100

[4] Cheng Y. Mean Shift, Mode Seeking, and Clustering, IEEE Trans. Pattern Anal. Machine Intel l., vol. 17, 790-799, 1995.

[5] Inglis A., et al., 2008. Automated identification of neurons and their locations. Journal of Microscopy. Jun;230(Pt 3):339-52. doi: 10.1111/j.1365-2818.2008.01992.x.

[6] Jiao, Y. & Du, P. Quant Biol (2016) 4: 320. https://doi.org/10.1007/s40484-016-0081-2

[7] Han, J.W., Breckon, T.P., Randell, D.A. et al. Machine Vision and Applications (2012) 23: 15. https://doi.org/10.1007/s00138-010-0275-y

[8] He, Y. et al., 2015.iCut: an Integrative Cut Algorithm Enables Accurate Segmentation of Touching Cells. Sci. Rep. **5**, 12089; doi: 10.1038/srep12089.

[9] Li S., Wu L., Sun Y. Cell Image Segmentation Based on an Improved Watershed Transformation. 2010 International Conference on Computational Aspects of Social Networks. doi : 10.1109/CASoN.2010.28

[10] Li J, Tseng K-K, Hsieh ZY, Yang CW, Huang H-N (2014) Staining Pattern Classification of Antinuclear Autoantibodies Based on Block Segmentation in Indirect Immunofluorescence Images. PLoS ONE 9(12): e113132. doi:10.1371/ journal.pone.0113132

[11] Luo F, Sun J, Hou S. Medical Images Fusion with Patch Based Structure Tensor. The Open Biomedical Engineering Journal. 2015;9:199-203. doi:10.2174/1874120701509010199.

[12] OpenCV(Open Source Computer Vision Library), 2018. https://opencv.org/

[13] Park BE, Jang WS, Yoo SK. Texture Analysis of Supraspinatus Ultrasound Image for Computer Aided Diagnostic System. Healthcare Informatics Research. 2016;22(4):299-304. doi:10.4258/hir.2016.22.4.299.

[14] Джеймс Г., Уиттон Д., Хасти Т., Тибширани Р. Введение в статистическое обучением с примерами на языке R. Изд. второе, испр. - М.: ДМК Пресс, 2017.- 456с.

Об авторах

Турлапов Вадим Евгеньевич, д.т.н., профессор кафедры математического обеспечения и суперкомпьютерных технологий ННГУ им.Н.И.Лобачевского,
email:vadim.turlapov@itmm.unn.ru

Носова Светлана Александровна, аспирант 2 года,
ИИТММ ННГУ им.Н.И.Лобачевского,
email: svetlana.nosova@itmm.unn.ru

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Цифровая Земля и Большие Данные

Digital Earth and Big Data

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Улучшенный алгоритм встраивания информации в область ДКП цифровых изображений на основе метода QIM с переменным шагом квантования

О.О. Евсютин, А.С. Кокурина, Р.В. Мещеряков
eoo@keva.tusur.ru | annakokurina94@yandex.ru | mrv@tusur.ru
Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

Стеганографическое сокрытие информации в частотной области цифровых изображений сопряжено с такими проблемами, как потеря встроенной информации и искажение естественной модели изображения. Ранее авторами настоящего исследования был предложен подход к решению первой проблемы, заключающийся в итеративном исправлении ошибок встраивания, и получен алгоритм, основанный на методе QIM и обеспечивающий безошибочность извлечения встроенной информации за счет предложенного подхода. В настоящей работе предлагается новый подход к минимизации искажений естественной модели цифрового изображения в области ДКП, основанный на изменении шага квантования в зависимости от локальных характеристик изображения в частотной области. Результатом работы является улучшенный алгоритм встраивания информации в область ДКП цифровых изображений, объединяющий два предложенных подхода. Результаты экспериментов показывают, что разработанный алгоритм наряду с обеспечением безошибочности извлечения позволяет существенно уменьшить искажения, вносимые в коэффициенты ДКП, и обеспечить статистическую неразличимость исходных изображений и стегоизображений в 75% случаев.

Ключевые слова: *цифровые изображения, стеганография, дискретное косинусное преобразование, QIM.*

An improved algorithm for embedding information into the digital images DCT domain based on the QIM method with a variable quantization step

O.O. Evsutin, A.S. Kokurina, R.V. Meshcheryakov
eoo@keva.tusur.ru | annakokurina94@yandex.ru | mrv@tusur.ru
Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia

Steganographic concealment of information in digital images frequency domain involves such problems as loss of embedded information and distortion of the natural model of digital images. Previously, the authors of this study proposed an approach to solving the first problem, consisting in iterative correction of embedding errors, and obtained an algorithm based on the QIM method and ensuring the error-free extracting of embedded information through the proposed approach. In this paper, we propose an approach to minimizing the distortions of the natural digital image model in the DCT domain, based on changing the quantization step depending on the image local characteristics in the frequency domain. The result of the work is an improved algorithm for embedding information in the DCT domain of digital images, combining the two proposed approaches. The results of the experiments show that the developed algorithm along with ensuring the error-free extraction makes it possible to reduce the distortions introduced into the DCT coefficients and to ensure the statistical indistinguishability of the cover images and stego images in 75% of cases.

Keywords: *digital images, steganography, discrete cosine transform, QIM.*

1. Введение

Цифровая стеганография предназначена для обеспечения конфиденциальности информации посредством ее скрытой передачи внутри некоторых цифровых объектов. Наибольшей популярностью пользуется сокрытие информации в цифровых изображениях по причине их широкой распространенности, простоты обработки и последующей передачи.

Одним из основных требований к методам встраивания информации является возможность ее последующего безошибочного извлечения. Это особенно актуально для частотного встраивания, когда биты сообщения записываются в частотные коэффициенты, полученные из пикселей изображения с помощью некоторого частотного преобразования.

Ранее авторами настоящего исследования был предложен новый подход к обеспечению безошибочности, заключающийся в итеративном исправлении ошибок встраивания, и получен алгоритм, основанный на методе модуляции индекса квантования (QIM) в области дискретного косинусного преобразования (ДКП) и обеспечивающий безошибочность извлечения встроенной информации за счет предложенного подхода [2].

Другим важным требованием к стеганографическим методам является незаметность встраивания. Внедрение дополнительной информации в элементы данных, составляющие цифровое изображение, нарушает естественный вид модели цифрового изображения и раскрывает факт передачи в нем дополнительной информации. В результате изображения–контейнеры и стегоизображения становятся статистически различимы. Вид искажений зависит от операции встраивания.

Данная работа развивает исследование [2] и предлагает новый подход к минимизации искажений естественной модели цифрового изображения в области ДКП при встраивании по методу QIM, основанный на изменении шага квантования в зависимости от локальных характеристик изображения в частотной области.

Основным результатом работы является новый улучшенный алгоритм встраивания информации в область ДКП цифровых изображений, одновременно обеспечивающий безошибочность встраивания и статистическую неразличимость изображений–контейнеров и стегоизображений.

Описание и обоснование решений, лежащих в основе разработанного алгоритма, представлено в следующих разделах данной статьи.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

2. Обеспечение безошибочного извлечения информации, встроенной по методу QIM

Основная идея метода QIM [4] заключается в модуляции яркости пикселей или значений частотных коэффициентов в зависимости от значений встраиваемых битов. Существует большое количество вариаций классического метода QIM. Например, в работе [3] представлена основанная на QIM техника внедрения цифровых водяных знаков для трёхмерных сеток. Авторы работы [10] предлагают улучшенную версию QIM, обеспечивающую устойчивость к масштабированию и ряду других искажений. В [7] представлен вариант метода QIM для JPEG-изображений, направленный на повышение устойчивости к атакам, основанным на гистограммах. В статье [8] описан метод стеганографического встраивания информации в область ДВП для изображений в формате JPEG2000. Авторами работы [9] предлагается использовать двойное вложение в рамках QIM, где один коэффициент модифицируется дважды для встраивания двух битов.

В основе алгоритма, представленного в настоящей статье, лежит формула, предложенная в статье [1]. Исходная формула описывает встраивание информации в пиксели изображения. В настоящей работе она модифицирована для применения к коэффициентам ДКП несжатых цифровых изображений и имеет следующий вид:

$$C'_{x,y} = \begin{cases} \operatorname{sgn} C_{x,y} \cdot \left(\left[\frac{|C_{x,y}|}{q} \right] + \left\{ \frac{|C_{x,y}|}{q/2} \right\} \right), & \text{если } b_i = 0; \\ \operatorname{sgn} C_{x,y} \cdot \left(\left[\frac{|C_{x,y}|}{q} \right] + \left\{ \frac{|C_{x,y}|}{q/2} \right\} + \frac{q}{2} \right), & \text{если } b_i = 1; \end{cases} \quad (1)$$

где $C_{x,y}$ – коэффициент ДКП до встраивания, $C'_{x,y}$ – коэффициент ДКП после встраивания, q – шаг квантования, b_i – бит сообщения, $[a]$ – целая часть a, $\left\{\frac{a}{b}\right\}$ – остаток от деления a на b.

Извлечение информации производится по формуле

$$b_i = \begin{cases} 0, & \text{если } \left\{ \frac{|C'x,y|}{q} \right\} < \frac{q}{2}, \\ 1, & \text{если } \left\{ \frac{|C'x,y|}{q} \right\} \geq \frac{q}{2}. \end{cases} \quad (2)$$

Алгоритм, основанный на формулах (1) и (2), сформулирован в статье [2]. Его основной особенностью является безошибочное извлечение информации из стегоизображения за счёт итеративной процедуры встраивания, впервые предложенной авторами настоящей работы в статье [5]. Итеративная процедура встраивания заключается в следующем. На этапе встраивания выполняется извлечение информации с проверкой на наличие ошибок. В случае, если при извлечении возникли ошибки, они исправляются, а затем выполняется повторное встраивание информации в уже изменённый блок. Если безошибочного извлечения не удалось достичь за определённое число шагов, блок обрабатывается таким образом, чтобы при извлечении ни один из его элементов не мог быть интерпретирован как содержащий встроенный бит. В данном алгоритме для этих целей применяется встраивание бита-флага в один из заранее выбранных коэффициентов. Областью встраивания является 21 высокочастотный коэффициент блока (с номерами 43–63 при зигзаг-сканировании).

Эксперименты, проведенные с данным алгоритмом, демонстрируют хорошие показатели ёмкости и визуального качества встраивания, оцениваемого по метрике PSNR [2].

3. Искажение гистограммы изображения в области ДКП при встраивании по методу QIM

Несмотря на высокое качество встраивания алгоритм [2] приводит к существенному искажению естественной модели цифрового изображения в частотной области.

Для иллюстрации этой проблемы обратимся к рис. 1, на котором изображены фрагменты гистограмм распределения значений коэффициентов ДКП классического тестового изображения «Lena» до и после встраивания информации согласно указанному алгоритму (шаг квантования $q = 10$). Вид гистограммы стегоизображения значительно отличается от вида гистограммы оригинального изображения. На гистограмме из рис. 1 (б) отчётливо заметно возникновение пиков в районе ± 5, что соответствует величине $\pm q/2$.

Эта проблема [6] является типичной для классического алгоритма QIM. Она связана с тем, что QIM сужает число возможных вариантов значений элементов данных цифрового изображения (пикселей или коэффициентов ДКП), в которые производится встраивание сообщения. В случае конкретного алгоритма, возникновение данной проблемы напрямую следует из формулы (1). Например, при встраивании единичного бита в коэффициенты, значения которых близки к нулю, формула преобразуется к виду $\operatorname{sgn} C_{x,y} \cdot \left(|C_{x,y}| + \frac{q}{2} \right)$, что в итоге трансформирует коэффициенты ДКП в значения, близкие к $\pm q/2$.

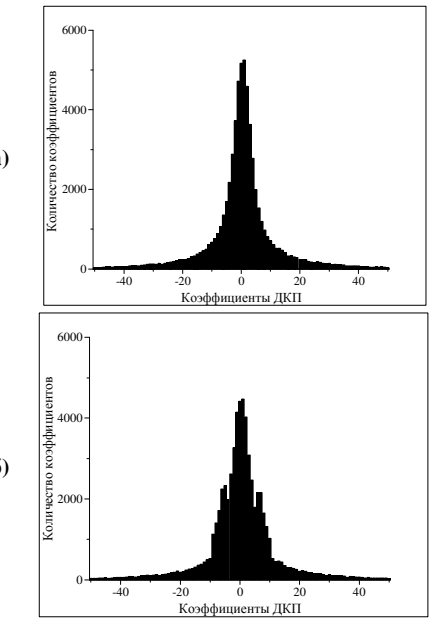


Рис. 1. Гистограммы изображения «Lena»: а) до встраивания; б) после встраивания

Описанная проблема напрямую связана с уязвимостью алгоритма перед стегоанализом, поскольку существенное искажение естественной модели в частотной области цифрового изображения может служить явным признаком наличия стеганографического вложения.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4. Восстановление естественной модели цифрового изображения за счет переменного шага квантования

Чтобы решить указанную проблему алгоритм [2] был модифицирован таким образом, чтобы шаг квантования для каждого блока коэффициентов ДКП зависел от характеристик этого блока. Для этого была введена формула, определяющая величину шага квантования:

$$q_n = \begin{cases} \lfloor F / E \rfloor \cdot K, \text{если } F \geq E, \\ \lfloor E / F \rfloor \cdot K, \text{если } F < E. \end{cases} \quad (3)$$

где E – среднее значение модулей АС-коэффициентов ДКП по области встраивания, F – среднее значение модулей АС-коэффициентов ДКП вне области встраивания, K – целочисленный коэффициент, n – номер блока коэффициентов ДКП изображения контейнера.

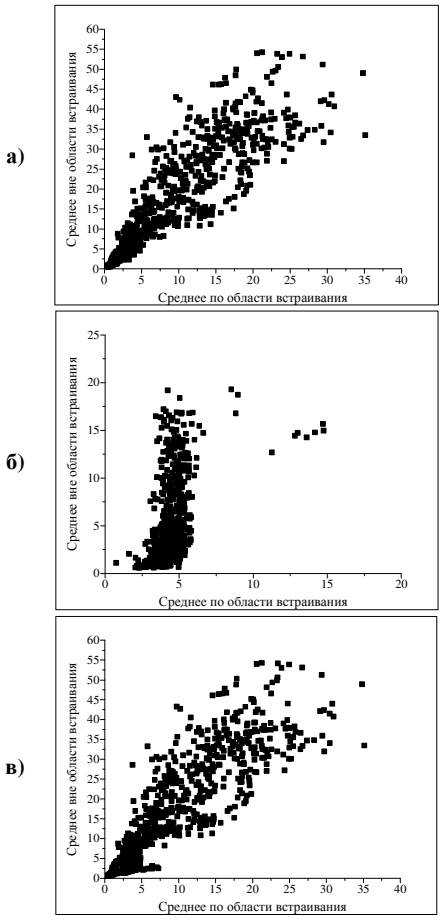


Рис. 2. Соотношение между средними значениями коэффициентов ДКП для: а) изображения контейнера; б) стегоизображения, полученного по алгоритму [2]; в) стегоизображения, полученного по улучшенному алгоритму [2] с переменным шагом квантования

Идея подобного способа вычисления шага квантования возникла из следующего наблюдения: соотношение между средними значениями коэффициентов ДКП по области встраивания и вне её имеет схожий характер для различных изображений. На рис. 2 (а) представлен пример, иллюстрирующий данное соотношение для тестового изображения «Airplane». График на рис. 2 (б) существенно отличается от графика на рис. 2 (а) по причине искажения естественной модели цифрового изображения в частотной области за счёт встраивания информации с постоянным шагом квантования (шаг квантования $q = 10$). Однако график, представленный на рис. 2 (в), демонстрирует, что встраивание информации с переменным шагом квантования, рассчитанным по формуле (3), позволяет значительно уменьшить изменение данного соотношения.

Рассмотрим влияние произведённой модификации алгоритма на распределение значений коэффициентов ДКП после встраивания. На рис. 3 представлены фрагменты гистограмм изображения «Airplane». Гистограмма на рис. 3 (б) содержит явные признаки наличия вложения. Аналогичная ситуация уже была проиллюстрирована на рис. 1 для изображения «Lena», что подтверждает тот факт, что искажение гистограммы при встраивании по алгоритму [2] возникает независимо от исходного изображения-контейнера. Рис. 3 (в) демонстрирует влияние переменного шага квантования на степень искажений гистограммы: пики сглажены, гистограмма визуально более похожа на гистограмму исходного изображения.

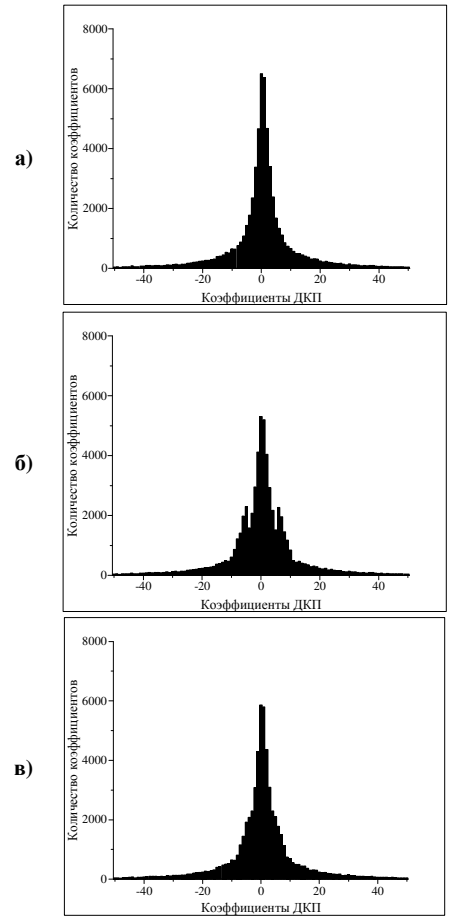


Рис. 3. Гистограммы изображения «Airplane»: а) до встраивания; б) после встраивания с постоянным шагом; в) после встраивания с переменным шагом

5. Результаты экспериментов и обсуждение

Для проверки эффективности предложенного подхода к корректировке искажений гистограммы была выполнена проверка статистической значимости различий между гистограммой контейнера и стегоизображения с помощью непараметрического критерия Уилкоксона. В табл. 1 представлены значения p-уровня значимости для данного критерия. Чтобы две гистограммы считались статистически неразличимыми, должно выполняться условие $p > 0{,}05$.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Изображение	*p*-уровень значимости	
	Постоянное $q = 5$	Переменное q
«Aerial»	$2{,}65 \cdot 10^{-5}$	0,099
«Airplane»	$6{,}21 \cdot 10^{-6}$	0,663
«Aps»	$2{,}03 \cdot 10^{-4}$	0,042
«Baboon»	$1{,}68 \cdot 10^{-7}$	0,003
«Barbara»	$2{,}75 \cdot 10^{-10}$	0,053
«Boat»	$2{,}06 \cdot 10^{-5}$	0,237
«Cameraman»	$1{,}46 \cdot 10^{-6}$	0,282
«Couple»	$1{,}87 \cdot 10^{-9}$	0,105
«Elaine»	$6{,}41 \cdot 10^{-6}$	0,183
«Goldhill»	$3{,}63 \cdot 10^{-5}$	0,512
«House»	$2{,}54 \cdot 10^{-8}$	0,624
«Lake»	$3{,}90 \cdot 10^{-5}$	0,178
«Lena»	$3{,}83 \cdot 10^{-10}$	0,022
«Man»	$1{,}57 \cdot 10^{-8}$	0,060
«Peppers»	$1{,}37 \cdot 10^{-5}$	0,463
«Splash»	$1{,}50 \cdot 10^{-4}$	0,762
«Stream and bridge»	$2{,}67 \cdot 10^{-7}$	0,414
«Tank»	$1{,}63 \cdot 10^{-5}$	0,008
«Tiffany»	$1{,}23 \cdot 10^{-10}$	0,000
«Truck»	$9{,}63 \cdot 10^{-5}$	0,550

Табл. 1. Проверка статистической различимости гистограмм по критерию Уилкоксона

Как видно из табл. 1, при встраивании информации с постоянным шагом квантования гистограммы статистически различимы во всех случаях. Для переменного шага квантования гистограммы до и после встраивания информации статистически неразличимы в 75% случаев. Данный результат наглядно демонстрирует положительное влияние предложенного подхода на уменьшение искажений естественной модели цифрового изображения в частотной области. При этом качество встраивания улучшилось: при сравнимой ёмкости порядка 7 500 бит среднее значение PSNR составило 40,7 дБ и 41,8 дБ для алгоритмов с постоянным и переменным шагом квантования соответственно.

6. Заключение

В работе предложен новый подход к минимизации искажений естественной модели цифрового изображения в области ДКП при встраивании по методу QIM, основанный на изменении шага квантования в зависимости от локальных характеристик изображения в частотной области. С помощью данного подхода разработан улучшенный алгоритм встраивания информации в область ДКП, который обеспечивает не только безошибочное извлечение информации, но и статистическую неразличимость изображений–контейнеров и стегоизображений. Наличие эффекта от применения предложенного подхода подтверждается результатами проведённых экспериментов.

7. Благодарности

Работа выполнена при финансовой поддержке гранта Президента Российской Федерации по государственной поддержке ведущих научных школ Российской Федерации № НШ-3070.2018.8.

8. Литература

[1] Глумов Н.И. Алгоритм встраивания полухрупких цифровых водяных знаков для задач аутентификации изображений и скрытой передачи информации / Н.И. Глумов, В.А. Митекин // Компьютерная оптика. – 2011. – Том. 35, №2. – С. 262 – 267.

[2] Евсютин О.О. Алгоритм встраивания информации в цифровые изображения с безошибочным извлечением, основанный на методе QIM / О.О. Евсютин, А.С. Кокурина, Р.В. Мещеряков // Труды XVI Всероссийской конференция «Распределенные информационно - вычислительные ресурсы. Наука – цифровой экономике» (DICR-2017). — 2017. — Россия, Новосибирск. – С. 314-322.

[3] Borah S. Quantization index modulation (QIM) based watermarking techniques for 3D meshes / S. Borah, B. Borah // Proceedings of the Fourth International Conference on Image Information Processing (ICIIP). — 2017. — India, Shimla. — P. 1–6.

[4] Chen B. Quantization index modulation: a class of provably good methods for digital watermarking and information embedding / B. Chen, G.W. Wornell // IEEE Transactions on Information Theory. – 2001. – Vol. 47, №4. – P. 1423 – 1443.

[5] Evsutin O.O. An adaptive algorithm for the steganographic embedding information into the discrete Fourier transform phase spectrum / O.O. Evsutin, A.S. Kokurina, R.V. Mescheryakov, O.O. Shumskaya // Proceedings of the First International Scientific Conference «Intelligent Information Technologies for Industry» (IITI'16). — 2016. — Russia, Sochi. — P. 47–56.

[6] Mitekin V. A new QIM-based watermarking algorithm robust against multi-image histogram attack / V. Mitekin, V. Fedoseev // Procedia Engineering. – 2017. – Vol. 201. – P. 453 – 462.

[7] Noda H. Application of QIM with dead zone for histogram preserving JPEG steganography / H. Noda, M. Niimi, E. Kawaguchi // Proceedings of the IEEE International Conference on Image Processing 2005. — 2005 (vol. 2). — Italy, Genova. — P. II-1082-5.

[8] Noda H. JPEG2000 steganography possibly secure against histogram-based attack / H. Noda, Y. Tsukamizu, M. Niimi // Proceedings of the 7th Pacific Rim Conference on Multimedia. — 2006. — China, Hangzhou. — P. 80-87.

[9] Sarkar A. Double embedding in the quantization index modulation framework / A. Sarkar, B.S. Manjunath // Proceedings of the 16th IEEE International Conference on Image Processing (ICIP). — 2009. — Egypt, Cairo. — P. 3653-3656.

[10] Zhu X. Improved quantization index modulation watermarking robust against amplitude scaling and constant change distortions / X. Zhu, Z. Tang // Proceedings of the 15th IEEE International Conference on Image Processing. — 2008. — USA, San Diego. — P. 433–436.

Об авторах

Евсютин Олег Олегович, канд. техн. наук, доцент кафедры безопасности информационных систем факультета безопасности Томского государственного университета систем управления и радиоэлектроники. Его e-mail eoo@keva.tusur.ru.

Кокурина Анна Сергеевна, младший научный сотрудник кафедры безопасности информационных систем факультета безопасности Томского государственного университета систем управления и радиоэлектроники. Её e-mail annakokurina94@yandex.ru.

Мещеряков Роман Валерьевич, д-р техн. наук, заведующий кафедрой безопасности информационных систем факультета безопасности Томского государственного университета систем управления и радиоэлектроники. Его e-mail mrv@keva.tusur.ru

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Визуализация и новое определение знака

Е.Н. Ерёмченко[1]
eugene.eremchenko@gmail.com
[1]МГУ, Москва, Россия

Предлагается к обсуждению новое определение знака, открывающее новые возможности для исследования семиотики визуализации.

Ключевые слова: *Цифровая Земля, визуализация, семиотика, знак.*

Visualization and New Definition of Sign

E.N. Eremchenko[1]
eugene.eremchenko@gmail.com
[1] Lomonosov MSU, Moscow, Russia

New definition of sign, that opens new possibilities for studying semiotics of visualization, is proposed for discussions.

Keywords: *Digital Earth, visualization, semiotics, sign.*

1. Введение

Знак занимает центральное место в человеческой цивилизации. Знаки используются для идентификации человека, о знаках как источнике всего сущего прямо указывается в первом же стихе одной из книг Нового Завета[1]. Знаки воспринимаются как единственно возможные носители информации и отождествляются с ней до степени смешения. Слово «знак» в русском и во многих иных языках непосредственно соотносится со знанием и является однокоренным с глаголом «знать» и с идентичным существительным, которым когда-то обозначалось то, что нынче стало просто «элитой» [4]. Представления о различных частных знаковых системах – например, математике, картографии и грамматике – были осознаны и доведены до совершенства ещё в глубокой древности [3]. Попытки осмыслить природу знаков как таковых безотносительно к их конкретной тематической области также предпринимались многие столетия назад.

Вместе с тем, как это ни парадоксально, семиотика как самостоятельная научная дисциплина, изучающая знаки во всём их многообразии, является одной из самых молодых наук и насчитывает менее двух столетий. К настоящему времени семиотика уже достигла известных успехов. Был осознан её всеобъемлющий характер, был сформулирован общий подход к изучению знака, предложены его определения. Сложилось общее представление об искусственности знаков, предложен ряд их классификаций. Семиотические исследования обрели огромную популярность и фактически стали метафорой, объединяющей попытки вскрыть истинный, неочевидный смысл объектов, процессов, явлений.

Однако на ряд вопросов удовлетворительные ответы до сих пор не найдены. Непонятен характер генетической связи между реальным и знаковым мирами, и в наше время сосуществуют как минимум две авторитетные цивилизационные парадигмы, имеющие на этот счёт полярные и несовместимые воззрения. Неясно, где находятся пределы семиотики, и что лежит за ними. Существенной проблемой продолжают оставаться так называемый «иконический» знак и в целом семиотика визуализации. В ряде аспектов существующие представления о природе визуальных знаков носят парадоксальный характер; это странно, поскольку высокая эффективность восприятия визуальных образов очевидна и безусловна, визуализация всё шире используется на практике и в отдельных областях – например, в геовизуализации – уже стала основным средством передачи образа обстановки.

В этой ситуации видится уместным критически рассмотреть существующие определения знака.

[1]«В начале было Слово, и Слово было у Бога, и Слово было Бог» [6].

2. Определение знака

К настоящему времени предложен целый ряд определений знака (см., например, [2]), в своей основе схематически отображаемых с помощью так называемого треугольника Фреге. Существуют разнообразные вариации этого треугольника, но все они сводятся в конечном итоге к формуле «знак – обозначает – обозначаемое» (рис. 1).

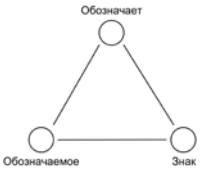


Рис. 1. Классическое представление взаимосвязей между знаком, обозначаемым и обозначением (треугольник Фреге).

Треугольник Фреге интересен своей самоочевидной бесспорностью и плодотворностью подхода, о котором свидетельствует бурное развитие семиотики, богатство и разнообразие тем, порождённых этим развитием на текущем этапе[2].

Вместе с тем, задаваемая этим подходом концепция знака имеет и свои естественные ограничения. В частности, она не позволяет выйти за пределы уже сформировавшихся отношений знака и обозначаемого и тем самым выйти за пределы реальности, декларируемой как реальность семиотическая и даже поставить вопрос о существовании чего-либо, что знаком не является. Существенной проблемой становятся «иконические» знаки. Очевидно, что в наиболее законченной форме отождествление объекта

[2]На XII Всемирном семиотическом конгрессе, прошедшем в Софии (Болгария) в 2014 году, доклады были распределены по 57 тематическим секциям, среди которых были секции «семиотика религии», «семиотика еды» и «музыкальная семиотика», «семиотика и кино» и «семиотика спектакля», «семиотика веб-сёрфинга и его пользователей», и другие аналогичные [1].

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

(процесса, явления и т.д.) с его собственным знаком полностью выхолащивает саму идею знака и семиотики, поскольку отождествляет любое чувственное восприятие с его собственным знаком и тем самым превращает знак во вторичную, искусственную и надуманную сущность. Совершенно необъяснимым видится феномен Цифровой Земли, в которой произошёл намеренный отказ от, казалось бы, досконально разработанных за долгие тысячелетия и доведенных до мыслимых пределов совершенства геопространственных условных знаков в пользу простых и примитивных с точки зрения семиотики не опосредованных никакими условностями изображений.

Особой проблемой существующая концепция знака становится для семиотики визуализации вообще. Можно сформулировать ряд вопросов:

1. Какова роль и место знака в восприятии субъектом внешней реальности? Возможно ли восприятие, не опосредованное знаками, или знак является исключительным средством, с помощью которого субъект воспринимает внешнюю реальность?
2. Что является продуктом визуализации с точки зрения семиотики? Только лишь знаки? Беззнаковые визуальные образы? Или то и другое вместе?
3. Как происходит или происходил генезис знаков вообще и генезис знаков иконических – в частности?
4. Как можно определить конечную цель и задачу визуализации как научного метода с точки зрения семиотики? К чему должна стремиться визуализация как наука?
5. Как происходит восприятие обстановки субъектом, какую роль играют в нём знаки и какую – гипотетически возможные иные средства восприятия обстановки?
6. Какой режим восприятия обстановки субъектом более эффективен – знаковый или беззнаковый?

Уже было показано, что кризис в семиотике можно разрешить посредством введения в семиотику представления о «нулевом знаке» [5] по аналогии с математикой.

Понятие о нулевом знаке позволяет очертить границы семиотики и исследовать генезис знаков из незнаковой реальности. Вместе с тем, оно нуждается в определении знака, которое бы не ограничивалось внутренними связями между знаком, означаемым и значением, задаваемым текущей семиотической парадигмой, и фиксировало бы действительно уникальное свойство (или набор свойств), имеющееся у любого знака и отличающее его от любой иной сущности.

Для выработки такого определения можно воспользоваться уникальной особенностью знака – его неразрывной связью с категориальным мышлением. Знак может соотноситься исключительно с категорией, т.е. с ясно осознаваемой субъектом сущностью, и при этом любая категория может быть обозначена различными по своей природе знаками, идентичными друг другу и обозначающими одно и то же понятие. Невозможны категории, которые вообще не могут быть обозначены каким-либо знаком, либо которые могут быть обозначены одним и только одним знаком. Знаком можно назвать лишь то, что имеет множество идентичных ему аналогов, так же исчерпывающе полно характеризующих ту же самую категорию, тождественных друг другу и неразличимых по своему «значению» вследствие своей тождественности (рис. 2). Знак в любом контексте может быть заменён иным, тождественным ему знаком. Соответственно можно предложить следующее определение знака:

Знак – элемент восприятия, обладающий свойством тождественной неразличимости.

Насколько целесообразно вводить в рассмотрение такое определение, в какой мере оно обосновано и насколько совместимо с уже имеющимися определениями? Очевидно, что оно не противоречит массиву определений, основанных на взаимоотношениях между знаком, означаемым и значением, и, наоборот, является комплементарным к ним, позволяя посмотреть на проблему знака с иной стороны. Новое определение фокусируется на отношениях между знаками и категориями, посредством которых осуществляется мышление, и фиксирует вторичность знака по отношению к обозначаемой им категории, проявляющуюся во множественном характере знаков. При этом данное определение не ограничивает восприятие только лишь знаками.

Рис. 2. Одно понятие может быть выражено различными знаками, тождественными друг другу по своему значению.

Акцент на свойстве тождественной неразличимости может оказаться полезным при исследовании взаимосвязи семиотики с уже сформировавшимися научными направлениями. Принцип неразличимости тождественных частиц известен из физики, однако нетрудно заметить, что он является основой мышления вообще. Мы можем мыслить лишь дискретными категориями, и все сущности, относящиеся к определённой категории, тождественны друг другу в рамках заданной классификации и, соответственно, неразличимы. Точно так же любая категория может быть обозначена множеством знаков, тождественных друг другу по своему значению. В какой мере такая дискретизация мышления ограничивает когнитивный потенциал науки – тема, нуждающаяся в дополнительной разработке.

Рис. 3. Условное обозначение взаимосвязей между элементом восприятия, субъектом и тождественными знаками.

Способно ли предлагаемое определение помочь в разрешении проблемы знака? Ответ на этот вопрос видится утвердительным. Во-первых, оно позволяет исследовать феномен беззнаковых носителей информации, а также генезиса знаков. Во-вторых, оно позволяет подойти к изучению феномена визуализации и границ применимости понятия об «иконическом знаке» в области визуализации, которая на наших глазах смешивается с восприятием мира вообще. В-третьих, оно выделяет интересную особенность любых знаковых конструкций, которая, по всей видимости, является уникальным признаком знака – знак всегда может быть заменён иным знаком, идентичным ему по своему значению.

3. Заключение

Предложенное определение знака позволяет по-новому выявить характерные особенности как знаков, так и иных, незнаковых потенциальных носителей информации, и нуждается в критическом рассмотрении.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4. Литература

[1] New Semiotics: Between Tradition and Innovation. Program. URL: http://semio2014.org/en/program Accessed: 01.09.2018.
[2] Sign. English Oxford Living Dictionaries URL: https://en.oxforddictionaries.com/definition/sign Accessed: 01.09.2018.
[3] Аверинцев С.С. Античная риторика и судьбы античного рационализма. - В сб. Образ античности. - СПб.: «Азбука-классика». 2004. с. 7.
[4] Знак. Этимологический словарь Фасмера. URL: https://vasmer.lexicography.online/з/знак Accessed: 01.09.2018.
[5] Ерёмченко, Е.Н. Концепция знака в контексте неогеографии. Информационные и математические технологии в науке и управлении. 2016. №1. Стр. 49-54.
[6] Иоанн. 1:1.

Об авторах

Ерёмченко Евгений Николаевич, научный сотрудник географического факультета Московского государственного университета имени Ломоносова, руководитель группы «Неогеография», член совета ISDE. Его e-mail eugene.eremchenko@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Цифровая Земля и «цифровые» инициативы современности

Е.Н. Ерёмченко[1], В.С. Тикунов[1], А.А. Захарова[2]
eugene.eremchenko@gmail.com|tikunov@geogr.msu.su|zaa@tpu.ru
[1]МГУ, Москва, Россия;
[2]ТПУ, Томск, Россия

В работе Цифровая Земля рассматривается в качестве особого режима геовизуализации среди разнообразных «цифровых» инициатив современности и, в частности, в связи с так называемой «цифровой экономикой». Обсуждается представление о беззнаковых носителях информации, вводимых в семиотику в полной аналогии с понятием о нулевом знаке в математике. Рассматривается возможность разрешения «парадокса береговой линии» и связанных с ним расходимостей кадастровых данных с помощью Цифровой Земли.

Ключевые слова: *Цифровая Земля, визуализация, типология, семиотика, экономика, устойчивое развитие.*

Digital Earth and «Digital» Initiatives of Modern Era

E.N. Eremchenko[1], V.S.Tikunov[1], A.A.Zakharova[2]
eugene.eremchenko@gmail.com|tikunov@geogr.msu.su|zaa@tpu.ru
[1]Lomonosov MSU, Moscow, Russia;
[2]TPU, Tomsk, Russia

The paper considers the place of Digital Earth as a special mode of geovisualization among various "digital" initiatives of our time, and, in particular, its connection with the so-called "digital economy". The concept of unsigned carriers of information introduced into semiotics in complete analogy with the concept of a zero sign into mathematics, is discussed. The possibility of resolving the "coastline paradox" and associated discrepancies of cadastral data with the help of the Digital Earth is considered.

Keywords: *Digital Earth, visualization, typology, semiotics, economics, Sustainable Development.*

Памяти С.В. Клименко посвящается

1. Введение

Понятия «цифровое» и «визуальное», «визуализация» широко используются в наши дни, всё чаще воспринимаясь в качестве панацеи для основных проблем сегодняшнего дня, с ними увязываются надежды на обеспечение так называемого устойчивого, или бескризисного, развития человечества, заявленного в качестве одного из пяти официальных направлений работы ООН [13]. Огромные ожидания связываются с цифровой экономикой. За рубежом и в России она уже декларирована приоритетным направлением развития экономики вообще [9]. Во всём мире растёт интерес к концепции Цифровой Земли (Digital Earth), впервые озвученной в США два десятилетия назад вице-президентом Альбертом Гором [7] и официально признанной в 2017 году в Австралии [4] и в России [17]. Однако содержание «цифровых» понятий продолжает оставаться туманным, поэтому их анализ видится актуальной задачей.

2. «Цифровые» инициативы с точки зрения семиотики

Использование «данных», т.е. представленных в цифровом виде параметров, зачастую считается отличительным признаком цифровой экономики, отличающим её от экономики нецифровой [10]. Однако специфической особенностью цифровой экономики как режима управления не может быть использование «данных». Наоборот, цифра является неспецифическим признаком экономического способа управления [15]. Цифра, т.е. знак, выражающий скалярное значение, есть необходимый признак любых экономических отношений, любая экономика может быть только цифровой и не может быть нецифровой[1]. Цифровая передача информации даже с помощью электромагнитных волн широко известна уже как минимум два столетия.

Воспользуемся системой семиотических определений экономики, цифровой экономики, цифровой революции и Цифровой Земли, предложенных и обоснованных в работе [15]:

Экономика – режим управления, основывающийся на использовании в качестве управляющего параметра единственной скалярной величины – стоимости.

Цифровой поворот (Digital Turn), или цифровая революция (Digital Revolution) – процесс диссоциации материи и знаков, проявляющийся в отчуждении знаков от их специфических материальных носителей.

Цифровая экономика – режим управления, в котором, наряду со стоимостью, применяются и другие опосредованные знаками управляющие параметры.

Цифровая Земля – среда, объединяющая знаковую и истинную реальность в едином виртуальном геоцентрическом континууме, подобном реальному.

Процесс отчуждения знаков от их материальных носителей, или диссоциация реального мира и мира знаков, или виртуальной реальности, является главным смыслом «цифровой революции». Новая виртуальная реальность, несмотря на свой «искусственный» и «вторичный» характер, сущностно тождественна миру реальному и противопоставлена ему, что делает их взаимодействие изначально конфликтным. Для иллюстрации масштаба ещё только начинающего проявляться конфликта достаточно

[1]Из классического определения экономики как отрасли знания, изучающей производство, потребление и взаимообмен богатства (economics is «the branch of knowledge concerned with the production, consumption, and transfer of wealth", [5]) следует, что предметом экономики является количественный аспект, представимый в виде скалярного значения.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

отметить, что виртуальный мир, едва появившись, стал потреблять энергию в аномально значительных количествах. Так, впервые в истории человечества энергия стала безвозвратно расходоваться уже не на создание материальных продуктов, имеющих стоимость, но непосредственно на производство нематериальных знаков стоимости, не имеющих каких-либо материальных носителей [1].

Предложенные выше определения характеризуют экономику как предельно элементарный режим управления, использующий в управлении единственный скалярный параметр – стоимость. Естественно, возможности «обычной» экономики ограничены, и цифровая экономика преодолевает эту ограниченность за счёт привлечения к управлению других аналогичных параметров. Однако в ситуации исчезновения связи между реальным и виртуальным мирами этого уже недостаточно. Необходимо восстановить непосредственную связь с геопространством, существовавшую в период «вмороженности» знаков в вещество и исчезнувшую с диссоциацией реального и виртуального миров. Восстановить эту связь позволяет Цифровая Земля – виртуальный образ реального геопространства, помещённый в виртуальное пространство и способный восстановить или имитировать связь знаков с геопространством.

Цифровая Земля как визуальный феномен интересна с точки зрения семиотики, поскольку существенно отличается от классического геопространственного продукта – карт и ГИС. Поэтому необходимо рассмотреть, что общего между картами и Цифровой Землёй и в чём состоят различия между ними.

3. Семиотические особенности Цифровой Земли

Цифровая Земля, как карты или ГИС, является геоинформационной системой [3,6,7]. Отличием Цифровой Земли является возможность обеспечения одновременно и внемасштабности, и всеракурсности, т.е. произвольного и интерактивного изменения пользователем эффективного ракурса и дистанции просмотра. Такой режим недостижим в классической картографии, поскольку карты одноракурсны и одномасштабны в силу принципиальных особенностей картографического метода. В Цифровой Земле внемасштабность и всеракурсность достигаются благодаря использованию вместо картографических знаков изображений, не опосредованных знаками.

С точки зрения классической семиотики, изображения – это иконические знаки, характерным и наиболее наглядным примером которых может служить рисунок. Иконические знаки отображают изображаемое посредством внешнего визуального сходства с ним. Но можно ли распространить это определение на такой вид изображений, как данные дистанционного зондирования и фотографические изображения вообще?

Современные технические средства позволяют получать изображения очень высокого качества, способные подменить собой непосредственное визуальное восприятие, в них последовательно устраняются последние остатки субъективизма, связанные с выбором композиции – ракурса, поля зрения и дистанции. Грань между прямым визуальным восприятием и изображениями, полученными техническими средствами, размывается. Поэтому закономерен вопрос – следует ли относить к иконическим знакам (и знакам вообще) все виды визуальных восприятий, зафиксированные техническими средствами?

Можно показать, что утверждение о тождественности знаков и визуальных восприятий внутренне противоречиво и потому логически несостоятельно. Во-первых, оно обессмысливает идею знака как таковую. Знак становится вторичной сущностью, эквивалентной чувственному восприятию, и превращается в дубликат восприятия вообще.

Во-вторых, оно несовместимо с принципом искусственности знаков, т.е. осознанного создания их субъектом. Проблему можно разрешить лишь признанием наличия иной переносящей информацию сущности, которая сама по себе ещё знаком не является, но которая способна инициировать создание знаков. Такую сущность можно назвать «не-знаками», или «нуль-знаками» [14]. Введение в семиотику понятия о «нулевых» знаках позволяет избавить её от внутренней противоречивости, неполноты и открывает возможность более широкого толкования информации.

Нулевой знак в семиотике является полным аналогом нулевого знака в математике.

4. Цифровая Земля и парадокс геовизуализации

Проблемы, возникающие при подмене реальности знаками, можно наглядно продемонстрировать на примере известного картографического парадокса. Широко распространено мнение о том, что геопространственные объекты (например, Австралию или Томскую область) можно измерить, точно определив их существенные метрические характеристики – площади территорий и акваторий, длины береговых линий, и т.д. Тем не менее, это не так. Континент, остров и т.д. в принципе не имеют и не могут иметь какой-либо определённой площади, определённой длины береговой линии и других аналогичных атрибутов, которые можно было бы выразить скалярными значениями, поскольку все эти параметры всегда являются производными не континента как такового, но его карты. В некоторых случаях скалярные характеристики географических объектов не имеют не только определённого значения, но и разумного предела, к которому они могли бы стремиться. Какое-либо «правильное» численное значение, представимое скалярной величиной, у этих параметров отсутствует в принципе[2], поскольку отсутствует единственно «правильный» масштаб – все масштабы равноправны. Если каждая карта – всегда и обязательно масштабна, а отображаемые ею объекты всегда принципиально внемасштабны. В отличие от своего собственного картографического представления, сам отображаемый географический объект существует всегда во всех масштабах одновременно, независимо от нашего восприятия, и не сводим, не может быть отождествлён ни с одной из опосредующих его карт.

Отмеченный выше факт принципиального отсутствия у географических объектов каких-либо точных измеряемых в виде скалярных значений характеристик новым не является и известен как «парадокс береговой линии» (coastline paradox) [2]. Именно этот парадокс, отмеченный впервые при попытке определения длины береговой линии Великобритании, стимулировал разработку концепции фрактальности [12] и носит для островного государства настолько фундаментальный характер, что Великобритания до сих пор не имеет собственного земельного кадастра [8]. Действие указанного парадокса распространяется не только на береговые линии, но и на любые территориальные границы, в том числе – границы кадастровых участков.

[2] “Geographical curves are so involved in their detail that their lengths are often infinite or, rather, undefinable” (Benoit Mandelbrot, 1966) [11].

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Парадоксальная на первый взгляд ситуация в картографии в точности соответствует логике парадокса, известного в математике как квадратура круга [16]. В геометрии она была успешно решена более двух тысячелетий тому назад введением понятия о так называемых числах иррациональных. Как невозможно определить отношение между длиной окружности и её радиусом в виде рациональной дроби, так же невозможно точно отобразить реальные географические образования при помощи векторного слоя и тем более представить их с помощью скалярных значений.

Отмеченный парадокс, в частности, ответствен за проблемы в кадастровом учёте, неустранимую некорректность кадастровых слоёв [18] и в целом за принципиальную невозможность ведения и поддержания кадастра в векторном виде. Только изображения являются корректным инструментом представления геопространства, точно так же, как только иррациональным числом может быть представлено отношение длины окружности к её радиусу.

5. Заключение

Цифровая Земля является репликой реального геопространства, восстанавливающей связь между реальным и цифровым мирами в эпоху их диссоциации. Её уникальными особенностями являются внемасштабность и всеракурсность, достигаемые за счёт использования беззнаковых носителей информации. Они позволяют разрешить парадокс, связанный с принципиальной неспособностью представления геопространства с помощью векторных или скалярных знаков из-за невозможности обеспечения межмасштабной топологии классическими картографическими средствами.

6. Благодарности

Неоценимую помощь в постановке проблемы, в критическом обсуждении полученных результатов и в целом в выработке и развитии семиотического подхода к визуализации и поддержке работ в этом направлении оказал профессор Станислав Владимирович Клименко (1941-2018), с которым неоднократно обсуждались затрагиваемые в этой работе темы.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-41-700001 р_а.

7. Литература

[1] Bitcoin Energy Consumption Index. URL: https://digiconomist.net/bitcoin-energy-consumption Accessed: 03.08.2018.

[2] Coastline paradox URL: https://en.wikipedia.org/wiki/Coastline_paradox Accessed: 07.07.2018.

[3] Craglia M., Kees de Bie, Jackson D., Pesaresi M., Remetey-Fülöpp G., Wang C., Annoni A., Bian L., Campbell F., Ehlers M., van Genderen J., Goodchild M., Guo H., Lewis A., Simpson R., Skidmore A. & Woodgate P. (2012). Digital Earth 2020: towards the vision for the next decade, International Journal of Digital Earth, 5:1, 4-21, DOI: 10.1080/17538947.2011.638500.

[4] Digitalearth-isde.org, 2017. Digital Earth Australia URL: http://www.digitalearth-isde.org/news/805 [Accessed: 27 December 2017].

[5] Economics (definition). URL: https://en.oxforddictionaries.com/definition/economics Accessed: 19.08.2018.

[6] Eremchenko E., Tikunov V., Ivanov R., Massel L., Strobl J., 2015. Digital Earth and evolution of cartography. Procedia computer science. – 66(C). – pp. 235–238.

[7] Gore A. The Digital Earth: Understanding our planet in the 21st Century. Al Gore speech at California Science Center, Los Angeles, California, on January 31, 1998.

[8] Grover R. Why the United Kingdom does not have a cadastre – and does it matter? URL: https://www.fig.net/resources/proceedings/2008/verona_am_2008_comm7/papers/12_sept/7_2_grover.pdf Accessed: 03.08.2018.

[9] Minsvyaz.ru, 2017a. Цифровая экономика. URL: http://minsvyaz.ru/ru/activity/directions/779/ [Accessed: 27 December 2017].

[10] Minsvyaz.ru, 2017b. Интервью | Николай Никифоров: цифровая экономика — новая веха в экономической жизни страны. URL: http://minsvyaz.ru/ru/events/36951/ [Accessed: 27 December 2017.

[11] Mandelbrot B. How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. Science, New Series. 1967. Vol. 156, N.3775, pp.636-638.

[12] Mandelbrot B. The Fractal Geometry of Nature. W.H. Freedman and Company. New York. P. 468. Updated and Augmented Edition, 1983.

[13] UN.org, 2017. Содействие устойчивому развитию. URL: http://www.un.org/ru/sections/what-we-do/promote-sustainable-development/index.html [Accessed: 27 December 2017].

[14] Ерёмченко Е.Н. Концепция знака в контексте неогеографии. Информационные и математические технологии в науке и управлении. 2016. №1. Стр. 49-54.

[15] Ерёмченко Е., Тикунов В., Никонов О., Мороз В., Массель Л., Захарова А., Дмитриева В., Панин А. 2018. Цифровая Земля и цифровая экономика. Геоконтекст №5(5), стр. 40-54. URL: http://www.geo-context.org/index.php/geocontext/article/view/33 [Accessed: 14 May 2018].

[16] Квадратура круга. URL: https://ru.wikipedia.org/wiki/Квадратура_круга Accessed: 03.05.2018.

[17] Роскосмос. Проект «Цифровая Земля» стимулирует рост российской экономики. https://www.roscosmos.ru/23571/ Accessed June 24, 2017.

[18] Россия в два раза больше, чем написано в учебниках. URL: http://ulgrad.ru/?p=144526 Accessed: 03.05.2018.

Об авторах

Ерёмченко Евгений Николаевич, научный сотрудник географического факультета Московского государственного университета имени Ломоносова, руководитель группы «Неогеография», член совета ISDE. Его e-mail eugene.eremchenko@gmail.com.

Тикунов Владимир Сергеевич, доктор географических наук, профессор, заведующий Лабораторией комплексного картографирования и Центр Мировой системы данных по географии географического факультета Московского государственного университета им. М.В. Ломоносова. Экс-вице-президент Международной картографической ассоциации (ICA), председатель комиссии по устойчивому развитию ICA. Его e-mail tikunov@geogr.msu.su.

Захарова Алёна Александровна, доктор технических наук, профессор Томского политехнического университета. Её e-mail zaa@tpu.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Анализ многомерных данных на основе комплексной модели нефтегазового месторождения

Д.А.Завьялов, А.А.Захарова
zda@tpu.ru | zaa@tpu.ru
ИПМ им. М.В.Келдыша РАН

В работе предложена комплексная модель месторождения углеводородов, которая позволяет повысить эффективность как самого процесса проектирования разработки, так и точность и достоверность проектного решения. Сформулированы условия реализации комплексной модели месторождения и эффективной обработки многомерных данных на ее основе.

Ключевые слова: *многомерные данные, месторождение, моделирование, визуализация, модель месторождения, комплексная модель.*

Multidimensional data analysis based on the integrated oilfield model

D.Zavyalov, A.Zakharova
zda@tpu.ru | zaa@tpu.ru
Keldysh Institute of Applied Mathematics RAS

The paper suggests an integrated hydrocarbon field model, which makes it possible to improve the efficiency of both the development process itself and the accuracy and reliability of its result. The conditions for the implementation of integrated field model and the effective processing of multidimensional data based on it are formulated.

Keywords: *multidimensional data, field, modelling, visualization, oilfield model, integrated model.*

1. Введение

Технология моделирования разработки нефтегазовых месторождений предполагает построение на основе многомерных исходных данных ряда трехмерных цифровых моделей для оценки объемов запасов углеводородов и планирования уровней их добычи на полное развитие месторождения.

Главная проблема в проектировании разработки месторождений – это высокая степень неопределенности. Модель месторождения представляет собой интерполяцию характеристик реального пласта, полученных в отдельно взятых скважинах, на расстояния в несколько километров, поэтому крайне важно для повышения достоверности моделей использовать все доступные данные. К таким данным относятся фактические показатели эксплуатации скважин, используя которые можно вычислить характеристики пласта для уточнения областей, не затронутых специальными исследованиями.

Основными и самыми значимыми этапами сложившейся и регламентированной [6, 7] технологии проектирования разработки месторождений являются:

- геологическое моделирование,
- гидродинамическое моделирование,
- прогноз разработки,
- экономическое моделирование,
- управление и принятие решений.

Каждому этапу предшествует анализ многомерных массивов исходных данных.

Целью создания геологической модели месторождения является оценка строения залежей, их свойств и объемов запасов углеводородов, в то время как гидродинамическая модель представляет собой уточнённую и дополненную новыми данными геологическую. При наличии истории разработки месторождения гидродинамическая модель может отличаться от геологической из-за настройки с учетом фактических режимов работы эксплуатационных скважин – адаптации модели. Кроме того, отличие может быть вызвано некорректной процедурой апскейлинга (загрубления) – уменьшения размерности модели, что в современных условиях развития технических средств становится все менее актуальным. При этом максимально допустимые расхождения моделей строго регламентированы [6, 7] и выявляются по результатам обязательной экспертизы.

При существующей технологии проектирования [8, 9] отсутствует возможность выполнения обратного перехода между этапами геологического и гидродинамического моделирования – нельзя скорректировать утвержденную геологическую модель на основе имеющихся фактических данных о режимах работы скважин. Такая особенность приводит к ряду проблем в процессе проектирования разработки: к росту несогласованности моделей, дополнительным финансовым и трудозатратам для выполнения нового подсчета запасов (для актуализации геологической модели).

Причиной сложившейся ситуации является то, что контроль за проектами разных типов осуществляется различными государственными структурами, кроме того геологическое статическое моделирование менее ресурсозатратно и более предпочтительно для выполнения грубого подсчета запасов объемным методом.

Решением проблемы может быть объединение этапов геологического и гидродинамического моделирования и переход к комплексной модели месторождения.

2. Условия реализации комплексной модели месторождения

На протяжении всего жизненного цикла месторождения углеводородов, который можно представить в виде спирали (рис. 1), происходит накопление информации и уточнение знаний об объекте разработки.

Сам процесс проектирования разработки месторождений связан с обработкой больших объемов многомерных данных, зачастую противоречивых или некорректных.

На ранних стадиях жизненного цикла для оценки запасов и строения залежей применяется геологическое моделирование, объемы исследований невелики, модели характеризуются низкой достоверностью, имеет место крайне высокая степень геологической неопределенности. В то время как на стадиях с более высокой степенью

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

изученности, при наличии истории разработки месторождения, появляется возможность применения комплексного подхода к моделированию и уточнения строения залежей и объемов запасов путем решения обратных задач на основе данных об эксплуатации скважин.

Предлагаемая концепция предполагает объединение всех моделей (геологическая, гидродинамическая, экономическая, технологическая) в единую комплексную модель месторождения на основе комплексного подхода, а также использование иерархии моделей разной степени детализации в зависимости от задачи.

Несмотря на высокую степень развития средств моделирования остается противоречие между геологической и гидродинамической моделями месторождений. В настоящее время можно утверждать об устаревании такого разделения. Следует говорить о модели месторождения вообще, которая на ранних стадиях жизненного цикла строится на основе результатов исследований единичных разведочных скважин, а на более поздних стадиях должна уточняться по результатам фактических режимов работы эксплуатационных скважин. Однако реализации такого подхода препятствуют административные барьеры.

Рис. 1. Схема этапов (обозначены цифрами 1-5) жизненного цикла и этапов проектирования разработки месторождения углеводородов.

Таким образом, реализация предлагаемого подхода возможна в рамках оперативной модели месторождения, используемой недропользователем для управления промыслом и принятия решений, либо при выполнении комплекса проектов: подсчет запасов и проектный технологический документ. Так как при адаптации не допускается рассогласованность моделей более 5%, выполнение таких проектных документов, как дополнения к технологическим схемам разработки, в случае существенного рассогласования моделей требует дополнительно новых проектов подсчетов запасов.

При создании модели решается прямая задача – структурное моделирование и интерполяция свойств пласта. Адаптация в свою очередь позволяет решить обратную задачу – актуализацию и уточнение модели на основе данных об эксплуатации скважин.

Процедура адаптации гидродинамической модели месторождения предполагает корректировку следующих трехмерных кубов параметров:

- проницаемость,
- относительные фазовые проницаемости,
- давление.

Данные параметры позволяют уточнить построенную по результатам геофизических исследований скважин модель на основе фактических режимов работы эксплуатационных скважин (дебиты, объемы добычи, давления). Однако в проектировании не учитывается возможность уточнения на основе данных показателей объемов запасов и коэффициентов извлечения.

Для постановки на государственный баланс запасы принято считать объемным методом:

$$Vo = Vb * Poro * So * \ \rho o * Bo \ ,$$

где Vo – объем запасов нефти, Vb – эффективный нефтенасыщенный объем породы, $Poro$ – коэффициент пористости, So – коэффициент нефтенасыщенности, ρo – плотность нефти, Bo – объемный коэффициент нефти. При этом So в системе двух фаз рассчитывается по формуле:

$$So = 1 - \left(\frac{3{,}183 * (\rho w - \rho o) * g * h}{A * \gamma * \cos(\theta)} * \sqrt{\frac{Perm}{Poro}} \right)^{-\frac{1}{B}} ,$$

где ρw – плотность воды, ρo – плотность нефти, g – ускорение свободного падения, h – высота относительно уровня свободной воды, A, B – коэффициенты степенной зависимости J-функции Леверетта $J(Sw) = A * Sw^{-B}$, расчет которых производится авторским программным обеспечением J-Function на основе результатов исследований керна, γ – поверхностное натяжение между нефтью и водой, θ – угол смачиваемости.

Часть параметров определяются по результатам порометрических исследований керна, однако из формулы видно, что уточнение распределения проницаемости позволяет получить более точную оценку нефтенасыщенности, а следовательно объема запасов.

Скорректированные по результатам адаптации относительные фазовые проницаемости и распределение остаточной нефтенасыщенности позволяют уточнить значение коэффициента извлечения углеводородов, которое, как правило, принимается экспертно или по аналогии.

Анализ истории проектирования и разработки (показатели эксплуатации скважин и состояния залежи с шагом по времени) ряда месторождений Томской области показал, что основными наиболее частыми причинами недостоверности моделей и, как следствие, невыполнения утвержденных проектом показателей являются:

- отсутствие полного комплекса необходимых данных;
- низкое качество исследований, некорректность исходных данных;
- конфликт интересов управляющих субъектов процесса проектирования разработки, к которым относятся государство, недропользователь и проектный институт;
- несогласованность моделей из-за невозможности корректировки утвержденной геологической модели при адаптации гидродинамической;
- междисциплинарность процесса, которая приводит к разобщенности вовлеченных специалистов и несогласованности их работы;
- человеческий фактор: ошибки при обработке данных и моделировании.

В результате были сформулированы условия реализации комплексной модели месторождения и эффективного проектирования на основе комплексного подхода к моделированию, которые представлены ниже.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Переход к визуальным моделям данных (переход от численного моделирования к визуальному, использование когнитивной визуализации). Построение визуальных моделей данных и использование их в качестве когнитивного инструмента при решении задач, связанных с анализом многомерных данных, способно значительно сократить время, необходимое для выработки необходимых решений.

Визуальная модель, активно использующая ресурсы зрительного восприятия, трансформируя данные в объект, воспринимаемый с помощью этих когнитивных инструментов, существенно упрощает процесс осмысления. [4, 5]

Кроме того, снижение количества численного моделирования в пользу визуальных моделей данных ведет к уменьшению ресурсоемкости процесса проектирования.

Объединение этапов геологического и гидродинамического моделирования – позволяет получать более точное описание геологического строения месторождения с учетом истории разработки. В процессе адаптации модели решается обратная задача нахождение геофизических свойств пласта, позволяющих достичь фактических показателей при моделировании.

Централизованное хранение, формализация и интеграция данных обеспечивает оперативный доступ к информации, а также верификацию и контроль целостности данных на всех этапах проектирования. Единое хранилище исходных данных и моделей данных обеспечивает доступ всех вовлеченных специалистов, в том числе в распределенных проектах, что решает проблему междисциплинарности процесса. Использование баз данных и технологии OLAP для аналитической обработки многомерных данных позволяет получать быстрый доступ к массивам информации об истории проектирования и разработки группы близких по строению месторождений. [10]

Постоянная верификация исходных данных – раннее и своевременное обнаружение ошибок позволяет повысить достоверность моделей и точность прогноза. Для выполнения этой задачи предлагается использовать визуальные модели многомерных данных. [3]

Соблюдение баланса интересов субъектов управления процессом проектирования – неотъемлемое условие успешного и эффективного проектного решения. К основным субъектам управления процессом проектирования относятся государство, как владелец недр и регулятор, недропользователь, как выгодоприобретатель, и проектный институт, выполняющий работы по проектированию. Интересы этих субъектов, как правило, пересекаются лишь частично.

Минимизация избыточности решения задачи предполагает использование модели как можно более низкого уровня, обеспечивающего достаточную точность получаемого решения. [1]

Соответствие показателям эффективности и требованиям к качеству. Выполнение регламентированных требований к качеству является обязательными при проектировании разработки. Показатели эффективности характеризуют насколько эффективно выполняет работу проектный институт и насколько эффективно представленное им проектное решение с точки зрения его практической реализации.

3. Заключение

Комплексная модель позволяет решать как прямые, так и обратные задачи, связанные с обработкой и анализом многомерных данных при моделировании разработки месторождений.

Применение такого подхода позволяет нивелировать переход между геологической и гидродинамической моделями, а также повысить точность подсчета запасов.

Учет истории разработки при создании комплексной модели месторождения позволяет избавиться от главного недостатка существующей технологии – отсутствия возможности обратного перехода к этапу геологического моделирования и несогласованности моделей.

Комплексный подход позволяет расширить области решения задач на смежные отрасли (экология, модели технологических процессов и пр.). К примеру, такой подход авторами работы был применен при моделировании зон закачки отходов в пласт, и позволил спрогнозировать безопасные объемы и места осуществления закачки. [2]

4. Благодарности

Работа выполнена при поддержке Российского научного фонда, проект 18-11-00215.

5. Литература

[1] Антоненко Д.А. 2010, Москва. Системные подходы к снижению рисков при моделировании разработки нефтегазовых месторождений

[2] Завьялов Д. А., Комплексный подход к моделированию полигона утилизации жидких нефтяных отходов // XIV Международная научно-практическая конференция студентов, аспирантов и молодых учёных «Молодёжь и современные информационные технологии» - Томск, 7-11 ноября 2016

[3] Завьялов Д.А., Шкляр А.В., Захарова А.А. Визуальные методы оценки и модели данных в проектировании разработки месторождений углеводородов // Графикон, 2017

[4] Захарова А.А., Шкляр А.В. Основные принципы построения визуальных моделей данных на примере интерактивных систем трехмерной визуализации // Научная визуализация. - 2014. - №2. - С. 62-73

[5] Захарова А.А., Шкляр А.В., Ризен Ю.С. Измеряемые характеристики задач визуализации // Научная визуализация. - 2016. - №1. - С. 95-107

[6] Методические рекомендации по подготовке технических проектов разработки месторождений углеводородного сырья // Минприроды России. Утв. 18.05.2016

[7] Методические рекомендации по применению классификации запасов и ресурсов нефти и горючих газов // Минприроды России. Утв. 01.02.2016

[8] Силич В.А., Комагоров В.П., Савельев А.О. Принципы разработки системы мониторинга и адаптивного управления разработкой «интеллектуального» месторождения на основе постоянно действующей геолого-технологической модели // Известия Томского политехнического университета. - 2013 - Т. 323 - №. 5. - С. 94-100

[9] Ямпольский В.З., Заикин И.А. Онтология «Интеллектуальное месторождение» // Известия Томского политехнического университета. - 2013 - Т. 323 - № 5 - С. 112-117

[10] Zavyalov D.A. The application of OLAP technology in automated risk assessment system for oil and gas fields // Journal of Physics: Conference Series (JPCS) (Scopus). — 2016

Об авторах

Завьялов Дмитрий Алексеевич, Институт прикладной математики им. М.В. Келдыша РАН. Его e-mail zda@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Захарова Алёна Александровна, Институт прикладной математики им. М.В. Келдыша РАН. Ее e-mail zaa@tpu.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Визуализация больших данных с применением методов корреляции событий информационной безопасности в инфраструктуре интернета вещей

С.Ю. Исхаков[1], А.Ю. Исхаков[1], Р.В. Мещеряков[1]
iskhakov.sy@gmail.com | iskhakovandrey@gmail.com | mrv@ieee.org
[1]Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

Статья посвящена вопросам визуализации больших объемов данных, обрабатываемых в процессе обеспечения безопасности инфраструктуры интернета вещей. Рассмотрены методы анализа и обработки больших потоков данных. Проведен обзор и сравнение методов нормализации и корреляции событий в системах управления информационной безопасностью. Предложена методика визуализации как средство совершенствования методов управления инцидентами и представлены результаты ее апробации. Рассмотрены варианты использования современных инструментов для поиска и визуализации больших данных.

Ключевые слова: *интернет вещей, визуализация больших данных, нормализация, корреляция, инцидент.*

Big data visualization with application of correlation methods of information security events in Internet of things infrastructure

S.Y. Iskhakov[1], A.Y. Iskhakov[1], R.V. Mescheryakov[1]
iskhakov.sy@gmail.com | iskhakovandrey@gmail.com | mrv@ieee.org
[1]Tomsk state university of control system and radioelectronics

Article is devoted to questions of visualization of big data processed in the Internet of things infrastructure. Methods of the analysis and processing of big data flows are considered. The review and comparison of normalization and correlation methods in security information management systems is carried out. The visualization technique as means of incident management methods improvement is offered and approbation results are presented. Modern tools for search and visualization of big data are considered.

Keywords: *Internet of things, visualization of big data, normalization, correlation, incident.*

1. Введение

Широкое распространение беспроводных технологий передачи данных, развитие сервисов облачных вычислений и увеличение адресного пространства в сети Интернет посредством внедрения IPv6 [1] стали основными предпосылками возникновения новой парадигмы – Интернета вещей (Internet of Things, IoT) [1], что можно определить как динамическую глобальную сетевую инфраструктуру на основе стандартных и совместимых протоколов связи, где физические и виртуальные «вещи» имеют идентификаторы, используют интеллектуальные интерфейсы и интегрируются в информационную сеть. Обобщение результатов научных исследований, проведенных за последние несколько лет авторским коллективом научной школы, позволяет говорить о глубокой интеграции современных робототехнических систем и IoT-инфраструктуры.

В отличие от компьютеров и смартфонов, значительный процент IoT-устройств не способен применять средства защиты от вредоносного программного обеспечения из-за отсутствия инфраструктуры для запуска подобных приложений. Однако, подавляющее большинство подобных устройств способны генерировать события для ведения журналов. Эта особенность позволяет реализовывать решения по их защите путем анализа лог-файлов. Поскольку в состав IoT-инфраструктуры входит огромное количество устройств с высоким уровнем неоднородности, то генерируемые журналы событий образуют большие объемы данных различного формата.

Настоящая статья нацелена на обобщение методических и практических наработок авторов в части организации сбора и обработки подобных данных для выявления инцидентов информационной безопасности. В качестве средства совершенствования методов управления инцидентами предложена методика визуализации.

2. Сбор и нормализация данных

Общих требований к структуре журналов событий на рынке IoT в настоящее время нет, поэтому разработчики аппаратных и программных продуктов создают системы логирования, исходя из собственных принципов. В одних случаях сообщения записываются в текстовый файл, в других отправляются на syslog-сервер [4,6] или помещаются в базу данных. Форматы хранения могут быть различными, даже если основной механизм транспорта лог-файлов совпадает. Несмотря на это наблюдается переход от средств управления журналами регистрации (Log Management) к SIEM-системам [6], обеспечивающим возможность анализа регистрируемых событий с точки зрения информационной безопасности.

Для событий, полученных от разных источников, могут отличаться формат (TXT, XML, JSON) [2] и стандарт записи. Нормализация – процедура приведения необработанного события к нормализованному виду в соответствии с заранее заданной для источника и типа события формулой нормализации. Два основных этапа нормализации это «разбор событий» (парсинг) и «сопоставление полей» (маппинг). Учитывая большое количество источников данных в инфраструктуре интернета вещей, необходимо осуществлять фильтрацию и укрупнение выборок событий, подлежащих анализу. Для сокращения таких выборок выполняется агрегация – процесс отбора событий, удовлетворяющих условию заранее настроенного правила агрегации, и объединения их в одно агрегированное событие. Ниже представлен пример нормализованного события в MaxPatrol SIEM [2]. В данной системе используется стандарт, разработанный на основе Common Event Expression (CVE) [2,7].

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

"action": "allow",
"object": "connection",
"status": "success",
"datafield1": "Access_Control_Policy_DMZ",
"datafield2": "HTTP",
"dst.ip": "192.168.7.16",
"dst.port": 80,
"event_src.category": "IDS/IPS",
"event_src.hostname": "NGIPSV-5555X",
"event_src.title": "asa_firepower",
"event_src.vendor": "cisco",
"importance": "info",
"protocol": "TCP",
"src.ip": "195.38.41.42",
"src.port": 26449,
"time": "2018-04-23T11:20:26Z"

Имея в распоряжении нормализованные и агрегированные данные о состоянии информационной безопасности ИТ-инфраструктуры, можно осуществлять действия, направленные на анализ событий и выявление инцидентов. Основным механизмом в решении данной задачи является применение методов корреляции.

3. Корреляция событий

В рамках данного исследования под инцидентом авторами понимается наличие на отрезке наблюдений зафиксированного факта нехарактерного изменения в сценарии работы объекта (устройства), подтвержденного результатом выполнения условия правила корреляции.

Набор данных, подлежащий обработке правилами корреляции, формируется экспертами на этапе нормализации событий и его описание выходит за рамки данной статьи. В терминах настоящего исследования корреляция событий – это процесс обнаружения инцидентов информационной безопасности путем анализа потока нормализованных данных [2,5]. При обнаружении в потоке событий такой их последовательности, которая указана в условии одного из заранее настроенных правил корреляции, регистрируется корреляционное событие.

В [3,5,7] представлена классификация, которая различает сигнатурные и бессигнатурные методы корреляции. Сигнатурные методы (англ. термин rule based) подразумевают создание человеком неких правил определения инцидентов. Бессигнатурные основаны на обнаружении аномалий по принципу черного ящика, среди которых выделяются подходы, основанные на спецификации и базирующиеся на интеллектуальном анализе данных. Анализ рынка [5-8] SIEM-систем, позволил сформировать список наиболее распространенных к применению на практике методов:

Statistical – бессигнатурный метод корреляции событий, основанный на измерении двух или более переменных и вычислении степени статистической связи между ними.

RBR Rule-based (pattern based) – метод, в котором взаимосвязи между событиями определяются аналитиками в заранее заданных специфических правилах.

CBR Codebook (case based). Корреляция производится по подходящим векторам из предварительно заданной матрицы событий.

MBR (model based reasoning) — метод основан на абстракции объектов и наблюдения за ними в рамках модели.

Graph based. Корреляция определяется на основе поиска зависимостей между системными компонентами в графическом представлении и построении на их основе графа (рис. 1). В случае обнаружения корреляции граф используется для поиска причины инцидента.

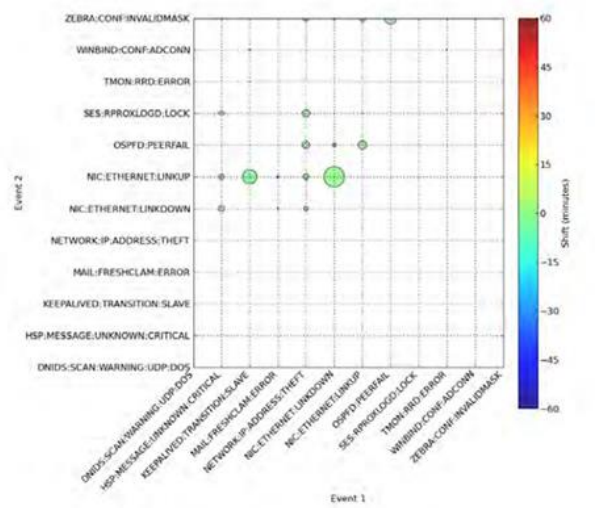


Рис. 1. Пример использования метода Grpah based

В большинстве случаев выявление причины необходимо для адекватной реакции на инцидент или осуществляется в контексте инцидента (например, пользователь изменил параметры запуска служб, необходимо выяснить причину). Наступлению каждого инцидента предшествуют различные события S (рис. 2): сканирование сетевых ресурсов, попытки установить соединение, подозрительные вложения в почтовом трафике. Корреляция позволяет объединить и группировать их для определения момента возникновения инцидента.

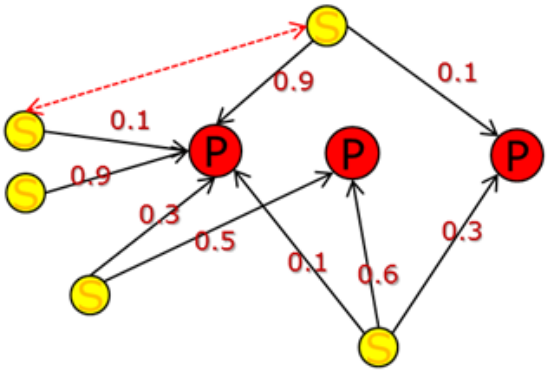


Рис. 2. Использование вероятностной модели

Правило имеет триггер, срабатывающий по условию, счетчик и сценарий реакции. Часть систем включают интуитивно понятные правила в графическом режиме. Счетчики предназначены для подсчета количества совпадений по одному и тому же правилу. Правила корреляции могут включать условия различной сложности.

4. Методика визуализации инцидентов безопасности

Визуализация информации о событиях и инцидентах в инфраструктуре интернета вещей предназначена для интерпретации и принятия решений по оперативной корректировке деятельности при реагировании на инциденты. В [6,7] представлены подходы к формализации данной задачи, в [5] предложена методика для визуализации данных о состоянии топологии сетей различного масштаба. Недостатком данного методического обеспечения является его направленность на визуальное представление топологии сети и состояния сетевых объектов (хостов). Ниже представлена методика визуализации данных, обрабатываемых SIEM-системами с учетом методов нормализации, агрегации и корреляции событий.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Шаг 1. Определение и выбор ИТ-активов, являющихся источниками данных в рамках конкретной решаемой задачи. Для каждого ИТ-актива, используемого в конкретной задаче необходимо определить механизмы транспорта лог-сообщений (интерфейсы и протоколы взаимодействия) и организовать получение событий SIEM-системой.

Шаг 2. Провести анализ получаемых событий с целью выделения тех сообщений, которые имеют ценность для обнаружения инцидентов и подлежат нормализации в ходе обработки.

Шаг 3. В соответствии с рассмотренными выше методами разработать правила нормализации событий, определенных на шаге 2.

Шаг 4. В соответствии с рассмотренными выше методами определить необходимость и разработать способы агрегации сообщений, определенных на шаге 3. В случае отсутствия необходимости агрегации для конкретного типа сообщений данный шаг может быть пропущен.

Шаг 5. Основываясь на рассмотренных выше методах, разработать алгоритмы и правила корреляций нормализуемых событий. Поскольку именно в результате обнаружения корреляций создаются записи об инцидентах, то именно они и будут являться данными, подлежащими визуализации. При этом необходимо определить принципиальные схемы событий. Методика основывается на положении, что в каждом событии можно выделить следующие элементы:

– субъект: инициатор взаимодействия, о котором повествует событие;

– объект: основная сущность, описываемая событием;

– источник: наблюдатель события, фактически ИТ-актив, сформировавший и приславший это сообщение.

Исходя из этого можно выделить следующие схемы событий с точки зрения взаимодействия субъекта и объекта.

1) В событии не содержится информации о каком-либо взаимодействии – источник передает информацию о состоянии объекта (рис. 3а).

2) В событии есть информации только об объекте, но субъект подразумевается контекстом события (рис. 3б).

3) В событии объект выполняет действие над самим собой (рис. 3в).

4) В событии имеется информация о субъекте и объекте, а также их взаимодействии (рис. 3г).

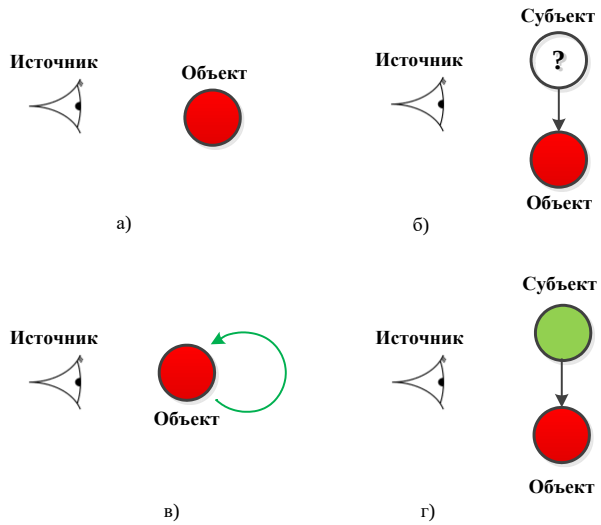


Рис. 3. Схемы событий с точки зрения взаимодействия субъекта и объекта

Шаг 6. Обеспечить накопление событий и записей об инцидентах в хранилище, обеспечивающем возможность работы с большими данными и предоставляющим функции поиска, в том числе поиска с нечеткими условиями.

Шаг 7. Определить возможные для использования графические модели [5] и скорректировать их с учетом сценария работы SIEM-системы. При определении моделей возможно использование критериев эффективности восприятия подсистемы визуализации [5].

Шаг 8. В соответствии с определенными на шаге 7 моделями реализовать визуализацию данных об инцидентах, извлекаемых из хранилища с помощью средств поиска, посредством использования программных компонентов или отдельных продуктов.

5. Апробация методики

Провести сравнение предложенного методического аппарата с исследованиями других ученых [3-5] на основе количественных характеристик не представляется возможным ввиду их различия в части основных положений методик. В связи с этим был проведен эксперимент по применению предложенной методики для решения задачи визуализации событий информационной безопасности в сети лаборатории безопасности интернета вещей ТУСУР.

Ранее было установлено, что в результате логирования событий, происходящих на IoT-устройствах, могут формироваться большие объемы данных. Для практической реализации предложенной авторами методики необходимо решить проблемы обработки больших данных. В [5,8] представлены обзоры рынка современных решений в области визуализации данных, среди которых стоит отметить продукт Elasticsearch. Это свободно распространяемый движок, предоставляющий распределенное аналитическое ядро поисковой системы. Используется в составе с Logstash [5] и Kibana [5]. Для взаимосвязи компонентов используется платформа обмена сообщениями RabbitMQ [8]. На рис. 4 представлена архитектура взаимодействия компонентов для организации возможности поиска, анализа и визуализации больших данных.

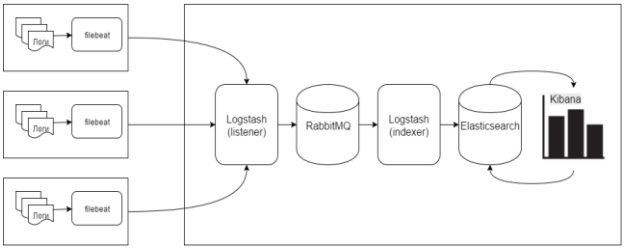


Рис. 4. Архитектура взаимосвязи компонентов

На базе стационарного компьютера был развернут стенд, имеющий следующие характеристики: 4-х ядерный процессор с тактовой частотой 3,1 ГГц; 16 Гб ОЗУ, на котором установлен стек ELK в составе: Elasticsearch 6.2.0, Logstash 6.2.0, Kibana 6.2.0. Стенд был развернут на базе операционной системы CentOS 7. Средняя скорость индексирования в Elasticsearch составила 15000 пакетов в секунду при средней загрузке каждого из ядер процессора 50%. Однако, поскольку основной целью эксперимента была оценка времени поиска среди индексированных сообщений, то были предприняты попытки отправки запросов в Elasticsearch для поиска среди сообщений слова из 6 букв. Среднее время поиска составило 1,02 мс.

Несмотря на достигнутые результаты в части поиска данных, были выявлены следующие проблемы при обработке больших данных, формирующихся в ходе логирования событий IoT-устройств. Во-первых, стек ELK критичен к ошибкам типа «OutOfMemory», что приводит к частым перебоям в ходе его использования. Кроме того,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

высокая скорость поиска в больших объемах данных, достигаемая в первую очередь за счет документоориентированности сопровождается низким коэффициентом восстановления работоспособности в случае перебоев, что зачастую приводит к безвозвратной потере данных. С точки зрения безопасности ElasticSearch не имеет возможности для авторизации.

Одним из наиболее практичных подходов является применение правил, ограничивающих глубину корреляции и разделение базы событий на онлайн и архивную части. Например, события, произошедшие за последние сутки, хранятся в онлайн базе, по истечению таймера помещаются в архивную часть. Для работы с большим объемом данных применяются различные специализированные поисковые движки и инструменты визуализации. В следующих этапах исследования будут проведены эксперименты по использованию данного продукта в качестве поискового инструмента к различным реляционным СУБД для повышения стабильности работы и сохранности обрабатываемых данных. В некоторых публикациях [3-5] представлены данные смежных экспериментов. Ниже рассмотрены попытки сравнить полученные данные.

1. Аппаратные ресурсы. В большинстве рассмотренных примеров [3-5] используется распределение нагрузки на потоки в кластере из нескольких узлов (серверов). В данном случае использовались виртуальные машины на базе одного физического стенда, характеристики которого указаны выше.

2. Скорость индексирования и анализа используемых данных. В связи с невозможностью проведения экспериментов при равных условиях на идентичном оборудовании [3-5] проведение какой-либо количественной оценки по данному параметру не представляется возможным. Однако, такие сравнения вероятно будут проведены на следующих этапах исследования.

3. Механизм обработки данных. В эксперименте, как и в большинстве смежных исследований [3-5], применялась потоковая обработка.

4. Пул задач. Все рассмотренные задачи в смежных исследованиях имеют конкретную постановку и получить результаты их сравнения в численном виде не представляется возможным. Однако, на следующем этапе исследования будет возможно применение предлагаемого методического обеспечения для решения различных задач и получения количественных оценок будущих результатов.

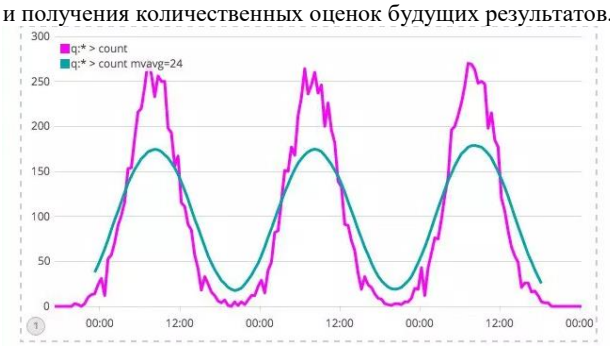


Рис. 5. Пример визуализации с помощью Kibana

На рисунке 5 представлен пример визуализации данных с помощью Kibana.

6. Заключение

Обобщение методических и практических наработок авторов в части организации сбора и обработки данных представлено в виде методики визуализации как средства совершенствования методов управления инцидентами. Информация о состоянии защищенности ИТ-активов в инфраструктуре интернета вещей может быть использована для непосредственного анализа, обнаружения инцидентов, а также их расследования и принятия решений.

Проведенные эксперименты показали, что применение методов корреляции является одним из инструментов совершенствования методов управления инцидентами в инфраструктуре Интернета вещей.

7. Благодарности

Работа выполнена при финансовой поддержке РФФИ (проект № 16-47-700350 р_а).

8. Литература

[1] *Abomhara M., Kien G.M.* Cyber security and the internet of things: vulnerabilities, threats, intruders and attacks // Journal of Cyber Security. – 2015. – Vol. 4. – Pp. 65–88.

[2] MaxPatrol SIEM [Электронный ресурс]. – Режим доступа: https://www.ptsecurity.com/ru-ru/products/mpsiem/?utm_source=slider (дата обращения: 23.06.2018).

[3] Dumitras T., Shou D.Toward a Standard Benchmark for Computer Security Research: the Worldwide Intelligence NetworkEnvironment (WINE) // Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS'11). 2011. pp. 89–96.

[4] Shu X., Smiy J., Yao D., Lin H. Massive Distributed and Parallel Log Analysis For Organizational Security // IEEE Globecom Workshops. December 2013. pp. 194–199

[5] Котенко И.В., Кулешов А.А., Ушаков И.А. Система сбора, хранения и обработки информации и событий безопасности на основе средств Elastic Stack/ [Электронный ресурс]. – Режим доступа http://proceedings.spiiras.nw.ru/ojs/index.php/sp/article/view/3590/2090 (дата обращения: 23.06.2018)/

[6] Милославская Н., Толстой А., Бирюков А. Визуализация информации при управлении информационной безопасностью информационной инфраструктуры организации //Научная визуализация. 2014. №2. [Электронный ресурс]. – Режим доступа http://sv-journal.org/2014-2/06/ru/index.php?lang=ru (дата обращения: 23.06.2018).

[7] Семёнов Д.П. Визуализация процессов информационной безопасности // Актуальные проблемы авиации и космонавтики. 2017. №13. [Электронный ресурс]. – Режим доступа: https://cyberleninka.ru/article/n/vizualizatsiya-protsessov-informatsionnoy-bezopasnosti (дата обращения: 23.06.2018).

[8] Корреляция SIEM. [Электронный ресурс]. – Режим доступа: https://www.securitylab.ru/analytics/431459.php (дата обращения: 23.06.2018)

Об авторах

Исхаков Сергей Юнусович, к.техн.н., доцент кафедры безопасности информационных систем Томского государственного университета систем управления и радиоэлектроники iskhakov.sy@gmail.com.

Исхаков Андрей Юнусович, к.техн.н., доцент кафедры безопасности информационных систем Томского государственного университета систем управления и радиоэлектроники iskhakovandrey@gmail.com.

Мещеряков Роман Валерьевич, д.техн.н., профессор кафедры безопасности информационных систем Томского государственного университета систем управления и радиоэлектроники mrv@ieee.org.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Пространственный анализ гиперспектральных изображений

Л.И. Лебедев[1], А.О. Шахлан[2]
lebedev@pmk.unn.ru|amega123@mail.ru
[1]ННГУ им. Н.И. Лобачевского, ул. Ульянова, 10, Нижний Новгород, 603005, Россия;
[2]МГТУ им. Н.Э. Баумана, 2-я Бауманская ул., д.5, стр.1, Москва, 105005, Россия

В работе предлагается методы и технология пространственного анализа гиперспектральных изображений. Сжатие, распознавание и кластерный анализ гиперспектральных изображений основаны на получении подпространств, с требуемыми для решения этих задач характеристиками. Сжатое описание изображений и решающие правила классификации базируются на информативных признаках. Методами прослеживания контуров формируется описание границ кластеров и классов на изображении. Предлагаемые методы и технология анализа демонстрируется на гиперспектральном изображении Deepwater Horizon Oil Spill спектрометра Spectir.

Ключевые слова: *гиперспектральное изображение, сжатие, распознавание, кластерный анализ, подпространство, собственные вектора, обучающие последовательности, информативные признаки, контур.*

Spatial analysis of hyperspectral images

L.I. Lebedev[1], A.O. Shakhlan[2]
lebedev@pmk.unn.ru|amega123@mail.ru
[1] Lobachevsky State University of Nizhni Novgorod, ul Ul'janova, 10, Nizhny Novgorod, 603005, Russia;
[2] Bauman Moscow State Technical University, 2-ja Baumanskaja ul., d.5, str.1, Moscow, 105005, Russia

In the paper, methods and technology of spatial analysis of hyperspectral images are proposed. Compression, recognition and cluster analysis of hyperspectral images are based on obtaining subspaces, with the characteristics required for solving these problems. The compressed description of images and the decisive rules of classification are based on informative signs. The methods of tracing outlines form a description of the boundaries of clusters and classes in the image. The proposed methods and analysis technology is demonstrated on the hyperspectral image «Deepwater Horizon Oil Spill» obtained by the Spectir spectrometer.

Keywords: *hyperspectral image, compression, recognition, cluster analysis, subspace, eigenvectors, training sequences, informative feature, contour.*

1. Введение

На протяжении двух десятилетий ведутся исследования использования снимков высокого спектрального разрешения для решения различных задач в области разведки земных ресурсов, состояния окружающей среды, мониторинга различного рода рисков. Области использования гиперспектральных изображений (ГСИ) обширны. Анализ гиперспектральных снимков стал одним из самых эффективным и актуальным применением решения задач дистанционного зондирования. Сегодня гиперспектральные снимки, в отличие от других ДДЗ (дистанционное зондирование Земли), позволяют извлекать более точную и детальную информацию. Результаты обработки гиперспектральных данных ДЗЗ находят все большее применение в агропромышленном комплексе, геологической разведке, экологическом мониторинге, что подтверждается возрастающим количеством публикаций по анализу ГИ в этих тематических областях [2,6-8].

Основными направлениями в анализе ГИ по праву считаются задачи сжатия (фильтрации) изображений, распознавания (идентификации) объектов на снимках, вопросы кластерного анализа и получения (уточнения) границ областей интереса. Решение этих задач нашло отражение в работах и докладах IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2018). В данной работе задачи сжатия, распознавания, кластеризации и выделения границ объектов предлагается осуществлять на основе получения и анализа соответствующих подпространств с использованием единой технологии [5], получившей дальнейшее развитие.

Пусть гиперспектральное изображение задано моделью, в которой каждый пиксель описывается набором признаков, значения которых соответствуют интенсивностям отражения при различных частотах одного и того же исследуемого элементарного фрагмента. В этом контексте, описание пикселя с одной стороны соответствует амплитудно-фазовому преставлению, как это отображено на рис. 1, а с другой стороны является точкой в линейном пространстве R_n, $x \in R_n$. Если в прямоугольной таблице в ячейках по строкам теперь записать значения интенсивностей пикселя, то в итоге получим классическую таблицу «объект - свойство» (ТОС) T_{nM}, свойственную для задач распознавания образов, в которой параметром M задано количество пикселей.

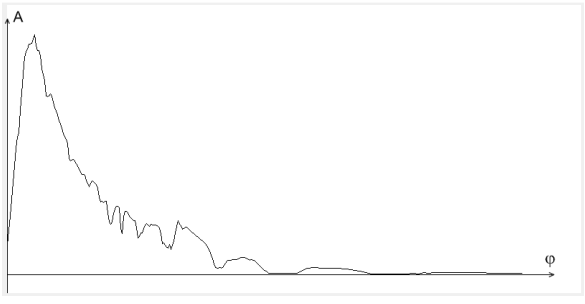


Рис.1. Амплитудно-частотное представление пикселя

2. Методы анализа

Пусть в пространстве R_n задано подпространство R_m, определяемое базисом $U = (u^1, u^2, ..., u^m)$. Тогда любая точка $x \in R_m$ может быть представлена своими

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

коэффициентами сверток $c_j = x^T \cdot u^j$, $j = 1,2,...,m$, причем, $x = \sum_{j=1}^{m} c_j \cdot u^j$. Отсюда,

$$x_k = \sum_{j=1}^{m} c_j \cdot u_k^j = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} x_i \cdot u_i^j \right) \cdot u_k^j = \\ = \sum_{i=1}^{n} x_i \cdot \left(\sum_{j=1}^{m} u_k^j \cdot u_i^j \right) = \sum_{i=1}^{n} x_i \cdot \gamma_i^k = x^T \cdot \gamma^k . \quad (1)$$

Имеем, $\Gamma = (\gamma^1, \gamma^2, ..., \gamma^n)$. В матричном виде $\Gamma = U \cdot U^T$. Среди $(\gamma^1, \gamma^2, ..., \gamma^n)$ линейно независимых точно равно m. Действительно, строки матрицы U можно рассматривать как коэффициенты разложения векторов Γ в базисе U, ранг матрицы которого равен m. Пусть это вектора $G = (g^1, g^2, ..., g^m)$. Пусть номера линейно-независимых строк матрицы U будут $\nu_1, \nu_2, ..., \nu_m$, тогда $g^j = \gamma^{\nu_j}$, $j = 1,2,...,m$. Отсюда, в матричном представлении $G = U \cdot V^T$, где V, матрица составленная из линейно-независимых строк матрицы U. Пусть компоненты матрицы $K = (k^1, k^2, ..., k^m)$ являются коэффициентами разложения векторов из Γ по базису G, то есть $\Gamma = K \cdot G^T$. Отсюда, $U \cdot U^T = K \cdot V \cdot U^T$, и $K = U \cdot V^{-1}$. Из (1) следует

$$x = (x_1, x_2, ..., x_n)^T = (x^T \cdot \gamma^1, x^T \cdot \gamma^2, ..., x^T \cdot \gamma^n)^T = \\ = \left(\sum_{j=1}^{m} k_1^j \cdot (x^T \cdot \gamma^{\nu_j}), \sum_{j=1}^{m} k_2^j \cdot (x^T \cdot \gamma^{\nu_j}), ..., \sum_{j=1}^{m} k_n^j \cdot (x^T \cdot \gamma^{\nu_j}) \right).$$

Окончательно, $x = \sum_{j=1}^{m} x_{\nu_j} \cdot k^j$, где x_{ν_j}, $j = 1,2,...,m$ линейно-независимые признаки. Если точка $x \notin R_m$, то в целях согласования с коэффициентами сверток для ее восстановления надо брать сглаженные признаки $y_j^* = x_{\nu_j}^* = x^T \cdot g^j$. Тогда по значениям коэффициентов сверток $c^T = (c_1, c_2, ..., c_m)$ однозначно можно получить набор информативных признаков $(y^*)^T = (y_1^*, y_2^*, ..., y_m^*)$ и наоборот: $c = V^{-1} \cdot y^*$, $y^* = V \cdot c$.

3. Получение подпространств.

В предлагаемых методах сжатия, распознавания и кластеризации базис подпространства формируются на основе собственных векторов, соответствующих наибольшим собственным значениям матрицы D специального вида:

$$D = \sum_{x \in X} \sum_{y \in Y} \frac{(x-y) \cdot (x-y)^T}{\|x-y\|^{\beta}} , \quad (2)$$

где X и Y - обучающие последовательности.

При сжатии ГИ в качестве множества X первоначально берутся все пиксели изображения, а множество Y представляет либо нулевая точка, либо средняя точка $\bar{x}$, полученная по всем пикселям ГИ. Значение β устанавливается равным нулю. Количество векторов m, базиса подпространства R_m берется с тем расчетом, чтобы отношение разброса по первому вектору к разбросу по m–ному направлению удовлетворяло заданному порогу, то есть $\sqrt{\lambda_1 / \lambda_m} > Thr$. Количество взятых векторов базиса U обеспечивает не только уровень среднеквадратической ошибки и качество аппроксимации пикселей, но и определяет набор информативных признаков нового описания пикселей и коэффициент сжатия k_r. В случае линейного подпространства R_m, когда $Y = \{0\}$, коэффициент сжатия тогда равен $k_r = n \cdot M / [m \cdot (M+n)]$, а случае линейного многообразия $\mathrm{M} = \bar{x} + R_m$, когда $Y = \{\bar{x}\}$, его значение равно $k_r = n \cdot M / [m \cdot (M+n+1)]$. Отметим, что значения информативных признаков $(x_{\nu_1}, x_{\nu_2}, ..., x_{\nu_m})$ остаются в том же диапазоне, что и для исходного набора и поэтому, их кодировка не приводит к увеличению разрядности при передаче, что может нарушаться при кодировании интегральными признаками $(c_1, c_2, ..., c_m)$. Для дальнейшего обеспечения качества аппроксимации требуется формирование новой обучающей последовательности X^1. Использование других полученных собственных векторов нецелесообразно, так как они аппроксимируют случайные данные. Новая последовательность должна формироваться на основе пикселей не удовлетворяющих качеству аппроксимации. Формирование может осуществляться при использовании различных критериев (Евклидова метрика, метрика Чебышёва и т.д.) на базе ошибок аппроксимации точек множества X. Далее аналогичным образом для множества X^1 находится базис и определяется дополнительный набор информативных признаков, который включается в общий список. Эта итеративная процедура завешается при выполнении критерия качества аппроксимации. Следует отметить, что алгоритмом сжатия предусмотрена смешанная аппроксимация пикселей как наборами $(x_{\nu_1}^*, x_{\nu_2}^*, ..., x_{\nu_m}^*)$, так и $(x_{\nu_1}, x_{\nu_2}, ..., x_{\nu_m})$ в зависимости от полученного качества их восстановления.

При распознавании значение β берется положительным, чтобы обеспечить подавление наиболее удаленных точек обучающих множеств при построении ковариационной матрицы межклассового разброса и, наоборот, увеличить вклад ближайших точек, от разделения которых в основном зависит качество распознавания. Это позволяет получить лучший, в плане построения решающих правил, базис разделяющего подпространства множеств. При решении практических задач распознавания объектов на гиперспектральных изображениях задание обучающих последовательностей может сводиться к заданию соответствующих фрагментов. В этих ситуациях после верификации всех пикселей фрагментов все оставшееся множество точек разбивается на два подмножества X и Y, которые не должны содержать пиксели одноименных классов. В случае задания фрагмента только для одного класса (множество X) множество Y формируется на основе всех пикселей изображения, не имеющих в описании сходства с точками распознаваемого класса, которое можно установить корреляционно-экстремальными контурными методами [4].

При решении задач кластерного анализа из тех же соображений берется значение $\beta < 0$. Здесь все множество пикселей изображения разбиваются на два подмножества X и Y произвольным образом.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Построение решающих правил осуществляется в подпространстве R_2, собственных векторов, соответствующих наибольшим собственным значениям. Используется иерархическая структура классификатора. На каждом уровне формируются новые обучающие последовательности из точек зоны отказов предыдущего уровня.

Построение классификатора осуществляется на базе информативных признаков. Для этого используются только кусочно-линейные решающие правила. Если решающее правило задано в виде $F(z) = u^T \cdot z - d = 0$, $\|u\| = 1$, то информативные признаки поэтапно выбираются в порядке убывания $\left|u_{\nu_j}\right|$, $j = 1,2,...,m$. При этом на каждом шаге решающее правило будет иметь вид $F(\hat{z}) = \hat{u}^T \cdot \hat{z} - \hat{d} = 0$, где $\hat{u} = (u_{\nu_1}, u_{\nu_2}, ..., u_{\nu_{m-1}}, \Phi_{m-1}/u_{\nu_m})$, а $\Phi_{m-1} = 1 - \sum_{j=1}^{m-1} (u_{\nu_j})^2$.

Для получения границ кластеров или распознанных объектов строится матрица с полутоновым описанием ячеек, размеры которой совпадают с исходным снимком. Всем пикселям изображения, удовлетворяющим условию принадлежности данному классу или кластеру, в соответствующие ячейки матрицы заносится идентификационный код. Далее методами прослеживания контуров формируется их описание [1,3].

4. Экспериментальные исследования

В качестве исходного изображения возьмем ГСИ Deepwater Horizon Oil Spill, полученное с использованием спектрометра Spectir.. Данное ГСИ содержит 371200 пикселей (матрица из 320 столбцов и 1160 строк), размерностью $n = 360$ каждый (число частот - каналов в спектрометре). Амплитуда сигнала колеблется в диапазоне от 1 до 32766 и была описана целой двухбайтовой переменной. Общий объем файла составил 267264000 байт. В связи с большим количеством пикселей (сотни тысяч), высокой размерностью их описания (сотни признаков) и большой вариабельностью значений по каждой частоте (десятки тысяч) алгоритмы и программы были ориентированы на использование 64-разрядных переменных целого типа и 80-разрядных переменных действительного типа, что естественно отразилось на их быстродействии.

Для сжатия ГСИ были построены за один проход две матрицы D при $Y = \{0\}$ и $Y = \{\bar{x}\}$ и найдены их собственные вектора. В результате анализа (автоматического, на основании отношений собственных значений) был в качестве базовой аппроксимации взят вариант с использованием линейного многообразия на основе пяти собственных векторов ($\lambda_1/\lambda_6 = 270$). Далее были определены номера пяти линейно-независимых признаков (каналов): 120, 13, 131, 101, 48. Таким образом, сжатое базовое описание определяется средней точкой $\bar{x}$ и набором признаков $(x^*_{120}, x^*_{13}, x^*_{131}, x^*_{101}, x^*_{48})$. Это вариант сжатия набора пикселей с использованием покоординатного среднеквадратического критерия по ансамблю точек. В нашем случае полученный набор признаков не смог обеспечить необходимое качество аппроксимации всех пикселей, которое оценивалось на основе среднеквадратической ошибки (СКО). На основе пикселей, у которых СКО превышало пороговую величину было сформирован новое множество. Описаниями точек этого множества E служили ошибки аппроксимации соответствующих пикселей. На рис.2 отображен этот набор пикселей, из которого видно разбиение этого множества на кластеры, для которых получение контурной модели представления является достаточно простой задачей.

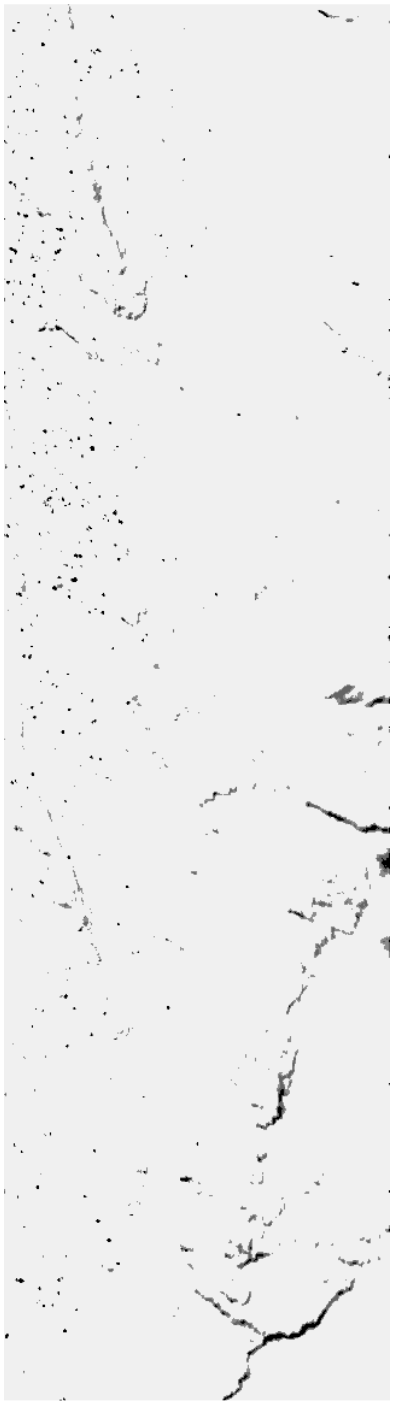
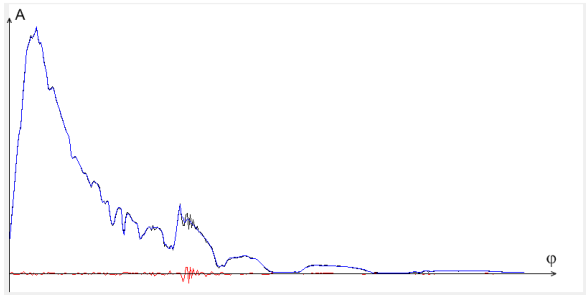


Рис.2. Выделение кластеров ГСИ

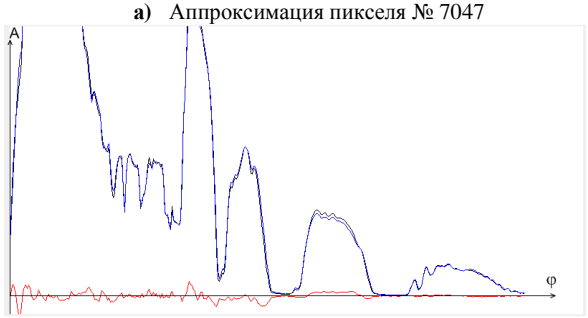
Аналогичным образом по множеству E была найдена ковариационная матрица и получены ее собственные вектора. В результате был получен набор из 5 признаков для аппроксимации точек из E: 120, 6, 12, 88, 132. В итоге,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

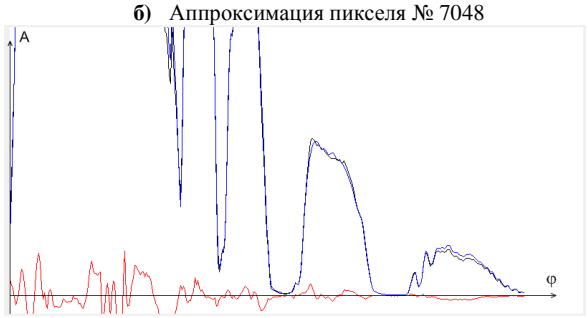
стала возможным аппроксимация с хорошим качеством даже «возмущенных» пикселей, представленных на рис. 3.



а) Аппроксимация пикселя № 7047



б) Аппроксимация пикселя № 7048



в) Аппроксимация пикселя № 7049

Рис.3. Динамика изменения характеристик пикселей ГСИ

Неудовлетворительная аппроксимация ошибок этого уровня была получена у 779 пикселей. Так как передача напрямую 779 ошибок будет по объему занимать меньше байт, чем введение в список информативных даже одного нового признака, то процедура сжатия бала завершена на этом этапе. Коэффициент сжатия составил более 33.

Методами построения простых решающих правил [5] с учетом специфики данных была решена задача распознавания помеченных выше 779 пикселей. Классификатор для выделения точек данного класса был построен на трех информативных признаках.

5. Заключение

Полученный результат применения предложенных методов сжатия, распознавания и кластеризации показал эффективность описанной технологии анализа ГСИ. Характерной чертой этих методов является возможность управляемого воздействия на решение задач анализа ГСИ (в том числе и интерактивного). Наибольшая трудность этой технологии заключена во временной сложности построения автокорреляционных матриц при распознавании и кластеризации ГСИ.

6. Благодарности

Работа выполнена при поддержке гранта РНФ № 16-11-00068.

7. Литература

[1] Васин, Ю.Г. Математические модели структурированного описания графических изображений./Ю.Г. Васин, О.А. Башкиров, С.Б. Рудометова//Межвузовский тематический сборник научных трудов «Автоматизация обработки сложной графической информации. – Изд-во ГГУ, Горький, 1984. – С. 92-117.

[2] Гашников, М.В. Иерархическая сеточная интерполяция при сжатии гиперспектральных изображений./ М.В. Гашников, Н.И. Глумов // Компьютерная оптика. – 2014. – Т. 38, – № 1. – С. 87-93.

[3] Дуда, Р. Распознавание образов и анализ сцен./ Р. Дуда, П. Харт// М.: Издательство «Мир». – 1976. – 512 с.

[4] Лебедев, Л.И. Корреляционно экстремальные контурные методы распознавания. Теоретические основы: Учебное пособие./ Л.И. Лебедев// Изд-во Нижегородского государственного университета, Нижний Новгород - 2013. -113с. - ISBN 978-5-91326-308-7.

[5] Лебедев, Л.И. Простые решающие правила в задачах распознавания образов и обработки графической информации// Научная библиотека диссертаций и авторефератов disserCat - http://www.dissercat.com/content/prostye-reshayushchie-pravila-v-zadachakh-raspoznavaniya-obrazov-i-obrabotki-grafricheskoi-in, 1995.

[6] Пахомов, П.А. Исследование сигнатур пикселов гиперспектрального изображения методом эмпирических мод/ П.А. Пахомов, А.В. Борусяк, В.Е. Турлапов //Сборник трудов IVмеждународной конференции и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2018). – Самара, 2018.

[7] Федосеев, В.А. Классификация гиперспектральных спутниковых изображений по эталонным участкам территории /В.А. Федосеев //Сборник трудов IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2018). – Самара, 2018.

[8] Wang, J. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis / J. Wang, C.-I. Chang // IEEE Trans. Geosci. Remote Sens. –2006. – Vol. 44 (6). – P. 1586-1600.

Об авторах

Лебедев Леонид Иванович, к.ф.-м.н., ведущий научный сотрудник Института информационных технологий, математики и механики Нижегородского государственного университета. Его e-mail lebedev@pmk.unn.ru.

Шахлан Анастасия Олеговна, магистрант кафедры компьютерные системы и сети факультета информатики и систем управления Московского государственного технического университета. Её e-mail amega123@mail.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Обнаружение термических аномалий на инфракрасных изображениях вулканов

И.П. Урманов[1], А.Н. Камаев[1], А.А. Сорокин[1]
[1]Вычислительный центр ДВО РАН, Хабаровск, Россия
urmanov@ccfebras.ru|kamaev_an@mail.ru|alsor@febras.net

В публикации представлен разработанный алгоритм поиска термических аномалий на фотоснимках вулканов, полученных в ближнем инфракрасном диапазоне света. Дано описание выделяемых признаков аномалий и используемого классификатора. Проведена апробация предложенного алгоритма на примере изображений вулканов Камчатки и рассмотрены перспективы его совершенствования.

Ключевые слова: *вулкан, алгоритм, термическая аномалия, инфракрасные изображения, программное обеспечение*

Thermal anomalies detection in infrared images of volcanoes

I.P.Urmanov[1], A.N.Kamaev[1], A.A.Sorokin[1]
[1]Computing Center FEB RAS, Khabarovsk, Russia
urmanov@ccfebras.ru|kamaev_an@mail.ru|alsor@febras.net

The publication presents a developed algorithm for the search for thermal anomalies on photographs of volcanoes obtained in the near infrared range of light. The characteristics of the anomalies and the classifier used are described. Approbation of the proposed algorithm on the example of images of volcanoes of Kamchatka is carried out and prospects of its improvement are considered.

Keywords: *volcano, algorithm, thermal anomaly, infrared images, software.*

1. Введение

В рамках работ по развитию программных средств для обработки изображений вулканов Камчатки [6], авторами ранее был разработан ряд специализированных алгоритмов [2, 7], обеспечивающих фильтрацию и автоматизированный поиск фотоснимков с признаками активности указанных природных объектов. В ночное время суток камеры системы видеонаблюдения ведут съемку в ближнем инфракрасном диапазоне, что при наличии соответствующих решений, позволяет выявлять возникновение и развитие термических аномалий. Как правило задача поиска термальных аномалий решается для термальных снимков [5]. А при анализе снимков ближнего инфракрасного диапазона такие аномалии возможно найти с помощью алгоритма, представленного в работе [4], но он имеет ряд ограничений, связанных с периодической посторонней засветкой и переменной видимостью вулкана. Кроме того, анализ данных снимков осложняется наличием на изображениях спекл-шума.

Статья посвящена исследованиям данного вопроса и описанию разработанного алгоритма для выявления термических аномалий на ночных снимках вулканов, полученных в ближнем инфракрасном диапазоне.

2. Обнаружение термических аномалий

Термические аномалии на снимках представляют собой пятна, яркость которых превышает яркость окрестности и спадает от центра к краям аномалии (рис. 1).

Размер аномалий может быть различным: от нескольких единиц до нескольких сотен пикселей. Термическим аномалиям соответствует широкоизвестная в компьютерном зрении визуальная особенность «blob», использующаяся, например, в алгоритме SIFT [3] для определения положения особых точек.

Рис. 1. Пример снимка вулкана с термической аномалией.

Центры потенциальных термических аномалий соответствуют точечным особенностям [3], обнаруженным в максимумах на DoG слоях. Для описания процедуры выделения областей, занимаемых аномалиями на i-м DoG слое $\mathrm{D}_i(x, y)$, где i – номер слоя введём множество соседей точки $\mathbf{p} = (p_x, p_y)$ на слое $\mathrm{D}_i(x, y)$

$M(\mathbf{p}) =$
$\{\mathbf{p}' = (p'_x, p'_y) \mid |p_x - p'_x| \leq 1 \wedge |p_y - p'_y| \leq 1 \wedge \mathbf{p} \neq \mathbf{p}'\}$.

Пусть обнаружено K центров аномалий на слое i: $\mathbf{c}_k = (c_{kx}, c_{ky})$, где $k = 1,2,\dots,K$, тогда алгоритм разметки областей, принадлежащих данным аномалиям можно записать следующим образом:

Шаг 1: Присвоить $\mathbf{A}_{x,y} = 0$, где $x = 0,1,\dots,w-1$, $y = 0,1,\dots,h-1$ – ширина и высота исследуемого изображения.

Шаг 2: Для $k = 1,2,\dots,K$: присвоить $\mathbf{A}_{c_{kx},c_{ky}} = k$, поместить точку (c_{kx}, c_{ky}) в очередь Q.

Шаг 3: Пока $Q \neq \emptyset$ выполнять шаги 3а – 3б.

Шаг 3а: Извлечь точку из очереди: $\mathbf{p} \leftarrow Q$

Шаг 3б: Если $\mathrm{D}_i(p_x, p_y) \geq 0.1\mathrm{D}_i\left(c_{A_\mathbf{p}x}, c_{A_\mathbf{p}y}\right)$ то $\forall \mathbf{p}' \in M(\mathbf{p}) \wedge \mathbf{A}_{\mathbf{p}'} = 0$: присвоить $\mathbf{A}_{\mathbf{p}'} = \mathbf{A}_\mathbf{p}$, поместить $\mathbf{p}'$ в очередь $Q \leftarrow \mathbf{p}'$.

После выполнения алгоритма разметки областей, для каждой точки i-го изображения маска $\mathbf{A}$ будет содержать

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

либо индекс аномалии, которой точка принадлежит, либо ноль, если точка не принадлежит какой-либо аномалии.

3. Признаки термических аномалий

Лишь небольшое число найденных потенциальных термических аномалий в действительности являются реальными аномалиями. Можно выделить ряд признаков, позволяющих произвести классификацию выделенных областей на реальные аномалии и не аномалии.

К признакам, которые могут быть рассчитаны для потенциальной термической аномалии можно отнести величину максимума, вытянутость и яркость.

$D_i(c_{kx}, c_{ky})$ – величина максимума k-й аномалии. Как правило, для термических аномалий характерны значения выше определённого порога, тогда как максимумы, вызванные шумом на изображении, имеют меньшие величины.

Признак вытянутости характеризует, насколько различается скорость убывания функции вокруг максимума в разных направлениях. Для термических аномалий, сгенерированных вулканами, различия в скорости невелики, тогда как для ярких пятен, образующихся у контура вулкана в лучах восходящего солнца, скорость убывания функции в одном направление существенно превышает скорость убывания в другом – рис. 2.

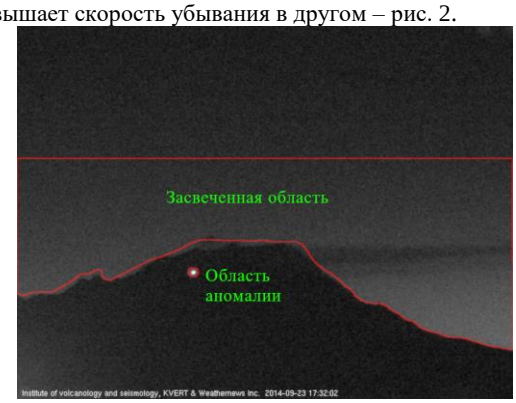


Рис. 2. Примеры областей.

Последний признак центра аномалии – его яркость на соответствующем масштабном слое изображения. Для реальных термических аномалий это значение, как правило, превышает некоторый порог.

В качестве признаков области аномалии предлагается использовать коэффициенты периметра, асимметричности границ и пика. Рассмотрим эти коэффициенты.

1. Коэффициент периметра показывает, во сколько раз периметр k-й аномалии на i-ом слое больше, чем минимально возможный периметр аномалии такой же площади.
2. Коэффициент асимметричности границ показывает, насколько равномерно спадает яркость аномалии к её границе.
3. Коэффициент пика показывает, во сколько раз перепад яркости аномалии превосходит среднюю яркость её границы.

4. Классификация

Возьмем размеченный вручную набор изображений с наличием на них реальных аномалий и построим гистограммы (рис. 3) статистических рядов каждого признака двух рассматриваемых классов аномалий: реальные аномалии и ложные.

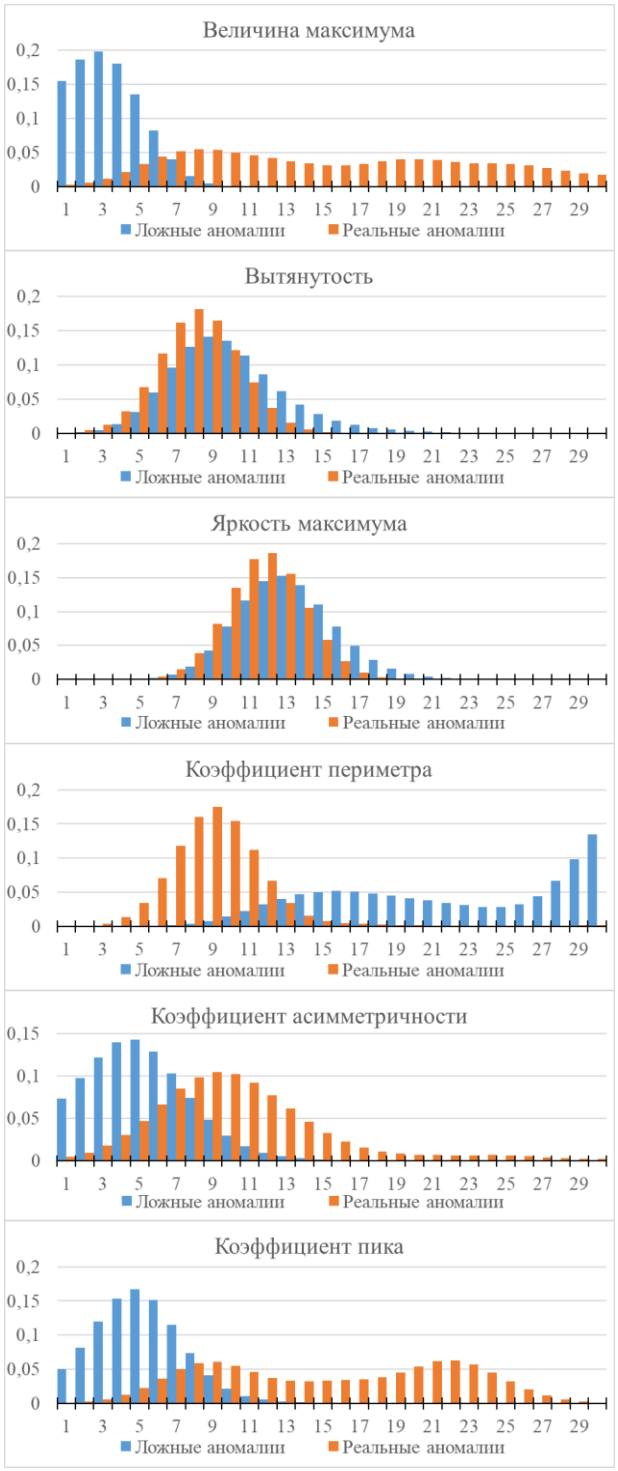


Рис. 3. Гистограммы статистических рядов каждого признака аномалий для 30 интервалов.

Заметим, что ни один признак не позволяет однозначно определить принадлежность аномалии к какому-либо классу. В связи с этим для поиска термальных аномалий на снимке, характеризующих состояние вулкана, используем наивный байесовский классификатор:

$$a = \underset{y \in Y}{\operatorname{argmax}}\, p(x) P_y,$$

где Y – множество классов,

x – исследуемый объект–аномалия,

$p(x)$ – функция правдоподобия,

P_y – априорная вероятность класса y.

При этом функцию плотности распределения значений j-го признака будем определять по формуле:

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

$$p_{yj}(\xi_j) = \frac{A(x)}{\int A(x)dx},$$

где $A(x)$ – кусочно-линейная функция, построенная по статистическому ряду, при этом значения частот сглаживаются одномерной функцией Гаусса.

5. Результаты

Для апробации созданного алгоритма была разработана компьютерная программа с применением которой было проведено обучение классификатора на 83 изображениях с размеченными зонами реальных аномалий. Тестирование алгоритма было проведено на 5000 снимков вулкана Шивелуч, сделанных в период с 12.10.2015 по 19.10.2015. Примеры найденных реальных термических аномалий на снимках представлены на рис. 4.

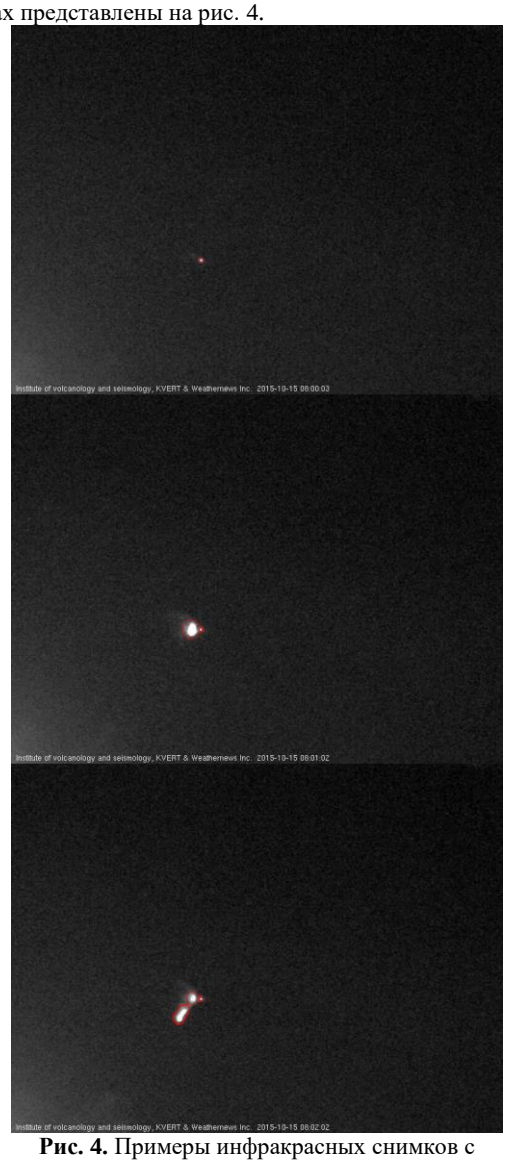


Рис. 4. Примеры инфракрасных снимков с выделенными реальными термическими аномалиями.

Результаты тестирования были оценены вручную, Количество снимков с ошибочно классифицируемыми аномалиями относительно общего числа обработанных изображений не превышает 1%. Анализ таких снимков показал, что ложно обнаруженные реальные аномалии (рис. 5) находятся на значительном расстоянии от вершины вулкана. Данную проблему можно решить, ограничив область поиска аномалий активной частью наблюдаемого вулкана.

Рис. 5. Пример ложно обнаруженных аномалий вне активной области вулкана.

Не все реальные аномалии были обнаружены (рис. 6) в ходе тестирования, решениями проблемы представляются расширение обучающей выборки и использование другого классификатора.

Рис. 6. Пример не обнаруженных реальных аномалий (выделено красным).

6. Заключение

В результате исследований был разработан алгоритм обнаружения и классификации термических аномалий на ночных снимках вулканов, снятых в инфракрасном диапазоне света. Алгоритм протестирован на тестовой выборке изображений системы видеонаблюдения за вулканами Камчатки, а полученные результаты позволяют говорить о возможности использования алгоритма в задачах оперативного мониторинга вулканической активности на Камчатке [1].

7. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 16-07-00156.
При выполнении расчетов были использованы информационно-вычислительные ресурсы ЦКП “Центр данных ДВО РАН”.

8. Литература

[1] Гордеев Е.И., Гирина О.А. Вулканы и их опасность для авиации // Вестник Российской академии наук. – 2014. – Т. 84, № 2. – С. 134-142. – DOI: 10.7868/S0869587314020121.
[2] Камаев А.Н., Урманов И.П., Сорокин А.А., Карманов Д.А., Королёв С.П. Анализ изображений для определения видимости вулканов // Компьютерная оптика. – 2018. – Т. 42, № 1. – С. 128-140. – DOI: 10.18287/2412-6179-2018-42-1-128-140.
[3] Lowe D. G. Distinctive image features from scale-invariant keypoints // International Journal of Computer Vision. – 2004. – Vol. 60(2). – P. 91-110.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

[4] Мельников Д.В., Маневич А.Г., Гирина О.А. Количественные характеристики активности вулканов Камчатки по данным веб-камер // Тезисы докладов XVIII региональной конференции, посвященной Дню вулканолога, 30 марта - 1 апреля 2015 г. Петропавловск-Камчатский: ИВиС ДВО РАН. – С. 92-94.
[5] Spampinato L., Calvari S., Oppenheimer C., Boschi E. Volcano surveillance using infrared cameras // Earth-Science Reviews. – 2011. – Vol. 106(1-2). – P. 63–91
[6] Sorokin A., Korolev S., Romanova I. et al. The Kamchatka Volcano Video Monitoring System // Proceedings of 2016 6th International Workshop on Computer Science and Engineering (WCSE 2016). – Tokyo, Japan, 17–19 June 2016. – P. 734–737. ISBN 978-981-11-0008-6.
[7] Урманов И.П., Камаев А.Н., Сорокин А.А., Королев С.П. Оценка видимости и состояния вулканов по последовательностям изображений стационарных камер наблюдения // Вычислительные технологии. – Т. 21 № 3. – 2016. – С. 80-90

Об авторах

Урманов Игорь Павлович, младший научный сотрудник Вычислительного центра ДВО РАН, лаборатория информационно-телекоммуникационных систем, Хабаровск, Россия. Контактный адрес электронной почты urmanov@ccfebras.ru.

Камаев Александр Николаевич, кандидат технических наук, научный сотрудник Вычислительного центра ДВО РАН, лаборатория численных методов в математической физике, Хабаровск, Россия. Контактный адрес электронной почты kamaev_an@mail.ru.

Сорокин Алексей Анатольевич, кандидат технических наук, заведующий лабораторией информационно-телекоммуникационных систем Вычислительного центра ДВО РАН, Хабаровск, Россия. Контактный адрес электронной почты alsor@febras.net

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Геометрическое моделирование. Компьютерная графика в образовании

Geometric Modeling. Computer Graphics in Education

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Techniques for modeling a high-quality *B*-spline curves by *S*-polygons in a float format

Rushan Ziatdinov[1], Valerijan Muftejev[2], Rifkat Nabiyev[2], Albert Mardanov[2], Rustam Akhmetshin[2]
ziatdinov@kmu.ac.kr|muftejev@mail.ru|dizain55@yandex.ru|systems@list.ru|ahmetshin@bk.ru
[1]Department of Industrial and Management Engineering, Keimyung University, Daegu, Republic of Korea;
[2]Department of the Fundamentals of Mechanisms and Machines Design, Ufa State Aviation Technical University, Ufa, Russian Federation

This article proposes a technique for the geometrically stable modeling of high-degree B-spline curves based on S-polygon in a float format, which will allow the accurate positioning of the end points of curves and the direction of the tangent vectors. The method of shape approximation is described with the purpose of providing geometrical proximity between the original and approximating curve. The content of the notion of a "harmonious, regular form" of B-spline curve's S-polygon in a float format is revealed as a factor in achieving a high-quality of fit for the generated curve. The expediency of the shape modeling method based on S-polygon in a float format at the end portions of the curve for quality control of curve modeling and editing is substantiated. The results of a comparative test are presented, demonstrating the superlative efficacy of using the Mineur-Farin configuration for constructing constant and monotone curvature curves based on an S-polygon in a float format. The findings presented in this article confirm that it is preferable to employ the principle of "constructing a control polygon of a harmonious form (or the Mineur-Farin configuration) of a parametric polynomial" to a B-spline curve's S-polygon in a float format, and not to a B-polygon of the Bézier curve. Recommendations are given for prospective studies in the field of applying the technique of constructing a high-quality B-spline curves to the approximation of log-aesthetic curves, Ziatdinov's superspirals, etc. The authors of the article developed a technique for constructing smooth connections of B-spline curves with ensuring a high order of smoothness of the composite curve. The proposed techniques are implemented in the FairCurveModeler program as a plug-in to engineering CAD systems.

Keywords: *S-polygon, NURBS, fair curve, typical curve, transition curve.*

"The Golden Age will include people who will learn to unite."
Konstantin Tsiolkovsky

1. Introduction

Mineur et al. [1] and Farin [2] proposed a method for constructing Bézier curves with a monotone change in curvature, based on the construction of the Bézier polygon (also known as *B*-polygon) with a monotonic change in the control polygon legs at fixed angles between the legs. The error in this method was recently demonstrated [18], although this fact was known a decade ago to one of our Japanese colleagues, Norimasa Yoshida[1].

In this paper, the aforementioned method adapts to the construction of *B*-spline curves with monotone curvature and the application of the *S-polygon* with the *Mineur-Farin configuration*. A technique for modeling a curve with an *S*-polygon in a float format is proposed, which allows for the precise positioning of the end points of the curve and the direction of the tangents. More detailed equivalence parameters for the two curves, such as the similarity of the *m*-level curvature and the similarity of the *m*-level curvature of torsion, are introduced. In addition, a method for approximating analytic curves with the equivalence of curvature and curvature of torsion to the *m*th level is proposed.

Modern CAD systems use a method of modeling *class A curves* using *S*-polygons of high-degree NURBS curves. To model closed curves, the *S*-polygon is used in a float format, and the closed format of the *S*-polygon is used for non-closed curves. Based on the configuration of the *S*-polygon in a float format, the quality of the modeled curve is easily judged. The harmonious, regular form of the *S*-polygon in a float format of the *B*-spline curve allows us to ensure high quality in the generated curve (hereinafter in this article we will clarify the concept of harmony of the *S*-polygon shape). Using a closed *S*-polygon is dictated by the need for precise positioning of the end points of the curve and the directions of the tangent vectors at the end points. The end portions of the closed *S*-polygon coincide with the initial portions of the *B*-polygons of the end segments of the curve, which allows precise positioning of the end points and directions of the tangent vectors at these points. However, using an *S*-polygon in a closed format causes instability in the curve modeling at the end portions of the curve.

2. Techniques for modeling B-spline curves

The Cox-de Boor algorithm [4, 5] using the format of the representation of the *B*-spline curve with an *S*-polygon in a float format is numerically stable, since it uses a sequence of operations dividing the segment into a given ratio. In addition, the operations of calculating points and differential characteristics of the curve are uniform on any segment. Obviously, for the best approximation of the circular arc, the segment of the *B*-spline curve of the vertex of the *S*-polygon in a float format must be uniformly chosen on the circle. We construct an *S*-polygon in a float format of a segment of a *B*-spline of the ninth degree along the vertices of a regular dodecagon inscribed in a circle with a radius of $R = 10$. The point of the largest deviation of the evolute of the curve from the center (0, 0) of the circle is the point with the coordinates (0.000000015321, 0.000000057206).

Transformations of the control polygon formats of *B*-spline curves and Bézier splines are performed using BS and SB algorithms [6-8]. In [9] NBS and NSB algorithms of transformations for rational splines are proposed. SB (NSB) algorithms for the transition from the format of the *S*-polygon *B*-spline to the format of the generalized *B*-polygon of the Bézier spline are stable, since they contain only operations of dividing the segments into a given ratio. BS (NBS) algorithms are less stable, so they contain extrapolation operations of a segment in a given ratio. The algorithms for converting *S*-polygon formats of multi-segment splines partially contain the operations of BS (NBS) and SB (NSB) algorithms.

[1] Private communication.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Fig. 1 shows the segment of the *B*-spline curve of the ninth degree associated with the *S*-polygon in a float format and the closed *S*-polygon coinciding with the *B*-polygon of the Bézier curve of the ninth degree.

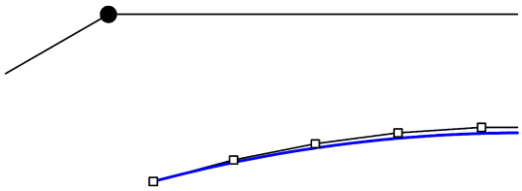


Fig. 1. An *S*-polygon in a float format and a closed *S*-polygon of a segment of a *B*-spline curve of ninth degree.

The vertices of the closed *S*-polygon (i.e., the *B*-polygon) do not lie on the circle. The legs of the *S*-polygon do not have the same length (the lengths of the legs are 0.518838, 0.519949, 0.520746, 0.521227, 0.521387, ...). The angles between the legs are also not equal: in radians they are 0.0653566, 0.0654361, 0.0654898, 0.0655169, ...). At the same time, it is this configuration that provides a high-precision approximation of the circular arc with an accuracy of 5.92221×10^{-8}.

In modern CAD systems for modeling class A curves, *B*-splines with *S*-polygons in closed format are used. It is impossible to draw a closed *S*-polygon with the same configuration as in Fig. 2 to approximate a circular arc with the same level of accuracy. The difficulty of ensuring proper quality at the end portions is emphasized in the guidelines for modeling class A curves in CAD systems.

3. Technique of geometrically stable modeling of high-degree B-spline curves using an S-polygon in a float format

An *S*-polygon in a float format of a *B*-spline curve of degree *m* can be used for quality control of modeling and for editing a curve [10].

Differential characteristics of the spline are determined by the divided differences of the *S*-polygon in a float format [5, 11]. A harmonious, regular form of an *S*-polygon in a float format of a *B*-spline curve of degree *m* must satisfy the second order isogeometry and regularity of an *S*-polygon of order *m*-1 [9]. Such a configuration of an *S*-polygon can be achieved by providing the required law of variation of the discrete approximation of the tangent vectors and curvature [9], as computed from the central divided differences of the *S*-polygon of the first and second order. More stringent requirements to the harmonicity of the polyline configuration can be introduced if one takes into account the condition that portions of a modeled or approximated curve with a monotone change in curvature are determined by segments of a polygonal line with a monotonous variation of the discrete curvature, as, for example, segments of a polygonal line with a Mineur-Farin configuration.

Modeling and editing of the high-degree *B*-spline curve at the end portions of the curve by means of a closed *S*-polygon is extremely unstable. Even simple rounding of the coordinates of a closed *S*-polygon's vertices can cause distortion of the shape. For example, let the arc of a *B*-spline curve of the ninth degree be given by a convex *S*-polygon in a float format (see Figure 1), built on the vertices of a dodecagon inscribed in a circle of radius $R = 10$. Rounding the coordinates of the second vertex of the closed *S*-polygon (see Fig. 1) from the value (-1.807034341228000, 8.748581800695999) to the value (-1.807, 8.748) causes a significant change in the shape of the curve and an unpredictable change in the shape of the *S*-polygon in a float format. The vertices of the original *S*-polygon (-7.071, -7.071), (-9.659, -2.588), (-9.659, 2.588), etc. have changed to (-83.972, 1295.780), (-8.274, -26.046), (-9.832, 5.520), and so on. The error of circular arc approximation increases from 5.92221×10^{-8} to a value of 3.06354×10^{-1}.

It is the stability of shaping with an *S*-polygon in a float format at the end portions of the curve that makes this method preferable.

The uncertainty of positioning the end points of the curve and the values of the tangent vectors is eliminated by employing a technique which consists of the following steps:

1. Construct the *S*-polygon of a curve in a float format of the proper shape.
2. Control the positions of the end points and the tangent directions when formatting the closed *S*-polygon.
3. Define the position difference vector (the vector determined by the difference of the actual end point versus the desired end point) and the difference angle of the direction (the angle between the actual tangent direction and the required one).
4. Switch the current formatting to that of *S*-polygon in a float format.
5. Eliminate misalignment in the position of the end points, requiring translation of the *m* terminal vertices of the *S*-polygon by the magnitude of the difference vector in the direction of decreasing the error value.
6. Eliminate the difference in the direction of the tangent vectors, requiring rotation of *m* end vertices of the *S*-polygon by the amount of the difference angle in the direction of decreasing the difference value.
7. Edit the changed configuration of the *S*-polygon in a float format until it reaches a harmonious, regular shape. Repeat the position control on the closed *S*-polygon and edit the *S*-polygon in a float format until a harmonious, regular form of the *S*-polygon in a float format and the accuracy of the closed *S*-polygon position are obtained.

4. Typical curves

In the works of Mineur et al. [1] and Farin [2] algorithms for generation of the control points for the formation of *typical Bézier curves* with a monotone change in curvature were developed. In these algorithms, a *B*-polygon of the Bézier curve is generated as the control polygon. The essence of the algorithms lies in the construction of a *B*-polygon with a monotonous change in the length of the leg, while maintaining a constant value of the angle between the legs and the angle of torsion. Such a configuration of a *B*-polygon is called the *Mineur-Farin configuration.*

The principle of constructing a control polygon by a Mineur-Farin configuration is correct in itself, however, as will be shown below, it is more appropriate to apply it to an *S*-polygon in a float format of the *B*-spline curve, and not to the *B*-polygon of the Bézier curve.

We will perform a comparative testing of these two different approaches for the limiting case, namely, modeling the curve of constant curvature. Obviously, the elongation coefficient of the leg in the Mineur-Farin configuration must be equal to 1 for generating the curve of constant curvature. It is also obvious that the Bézier curve and polynomial splines do not accurately approximate the geometry of the circular arc. However, a comparison of the approximation accuracy can demonstrate the advantage of using one method over another.

To generate the vertices of an *S*-polygon with a *B*-spline curve, we use an equilateral polygon inscribed in a circle of radius $R = 10$. For a degree of $n = 3$, a hexagon is used, and for degrees of $n = \{5, 7, 9\}$ a dodecagon is used.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

The construction of the vertices of a *B*-polygon is performed as follows: The terminal vertices of the *B*-polygon which coincide with the terminal vertices of the *B*-spline curve segment. The tangent vectors at the endpoints are orthogonal to the radius vectors from the center (0, 0). The intermediate vertices of the *B*-polygon are defined in such a way as to obtain a configuration with a constant leg length and a fixed angle between the legs.

Below is a table with the distance values of the evolute points farthest from the zero center (0, 0), of the Bézier curve and the *B*-spline curve.

Table 1. Comparison of approximation accuracy

Degree (*n*)	**Bézier curve**	***B*-spline curve**
3	8.17283×10^{-1}	8.33333×10^{-1}
5	8.06272×10^{-2}	1.11371×10^{-3}
7	4.65855×10^{-2}	7.85284×10^{-6}
9	3.19524×10^{-2}	5.92221×10^{-8}

As can be seen from Table 1, an increase in the degree results in an extremely slow increase in the approximation accuracy of a circular arc by Bézier curve. Even for a degree of $n = 9$, the Bézier curve does not provide acceptable accuracy (in modern CAD systems, the NURBS degree does not exceed a tenth degree). Concurrently, for $n = 9$, the *B*-spline curve achieves accuracy acceptable for engineering calculations in CAD systems. This test result shows the advantage of using the Mineur-Farin method for constructing a curve of constant curvature by means of an *S*-polygon in a float format. This statement is also true in the case of modeling curves with a monotone change in curvature. Consider the configuration of a polygon of five points with a monotonous change in the length of the leg with a scaling factor of two units and a fixed angle between the legs equal to 90°. Let us define the Bézier curve of the fifth degree on this configuration and on the same configuration, as on the *S*-polygon in the float format, we define the segment of the *B*-spline curve. Analysis of the evolute graphs shows that the curvature of the Bézier curve is not monotonic, and the *B*-spline curve is defined with a monotone change in the curvature (Fig. 2).

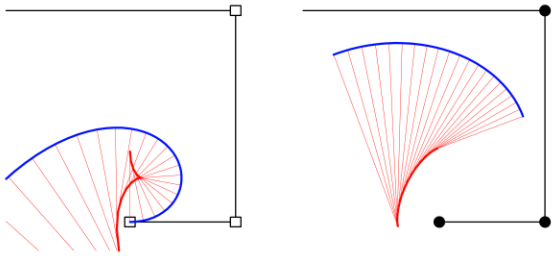


Fig. 2. Bézier curve (left) and *B*-spline curve (right) segments, defined on the same control polygon.

5. Construction of a high-quality approximating B-spline curve

Welters and Farin [3] proposed an approximation method taking advantage of the geometric proximity of two curves (the original curve and the approximating curve). In Welters and Farin's method, geometric proximity is estimated by examining the relative equivalence of two curves with respect to the parameters of curvature and torsion. Moreover, the authors introduced the term *shape approximation*, and noted that successful approximation guarantees the accuracy of the shape in the sense that the geometry of the original curve is approximated neatly [3].

To approximate the form, we propose to use additional, highly detailed characteristics of the internal geometry of the curve, proposed in [12]:

- Curvature of the first level is the curvature function itself $\kappa(s)$.
- The torsion of the first level is the torsion function itself $\tau(s)$.
- The second level curvature is a function of curvature of the curvature function $\kappa^{(2)}(s) = (\kappa \circ \kappa)(s)$
- The second level curvature of the torsion is a function of curvature of the torsion: $\kappa_\tau{}^{(2)}(s) = (\kappa \circ \tau)(s)$.

We define of the curvature of the *n*th level from the curvature of a curve as follows

$$\kappa^{(n)}(s) = \left(\kappa^{(n-1)} \circ \kappa\right)(s),$$

and the curvature of the *n*th level from the torsion of the curve is

$$\kappa_\tau^{(n)}(s) = \left(\kappa_\tau^{(n-1)}(s) \circ \tau\right)(s).$$

Under shape preserving approximation of *n*-level of equivalence we mean the equivalence of curvatures to the *n*-th level and of the curvatures of torsion to the *n*-th level for the original curve and the approximating curve. Equivalence of the shape of the curves means that the number and order of inflection and extrema points in the curves are the same. Thus, for high-quality shape approximation, the curvature of the curvature and the curvature of the torsion of the original and approximating curve should be equivalent in shape, at least until the second level. These arguments lead us to the following technique of isogeometric shape approximation of an analytic curve by *B*-spline curve of high-degree *n*, where $n>5$. The vertices of an *S*-polygon in a float format of a *B*-spline curve are generated directly on the analytic curve. With a proper distribution of points and the distribution density on the analytic curve, it is possible to ensure the equivalence of the shape of the analytic curve segment and the *B*-spline curve not lower than the second level.

As an example, a shape approximation of a conical spiral by a *B*-spline curve of the eighth degree is given (Fig. 3). The conical spiral is determined by the parametric equation

$$x(s) = 2 + s\ \sin s$$
$$y(s) = 2 + s\ \cos s$$
$$z(s) = s$$

The vertices of the *S*-polygon are computed on an analytic curve with the initial value of the parameter $s = 0$ and with a step $h = 1$ in the number of twenty points.

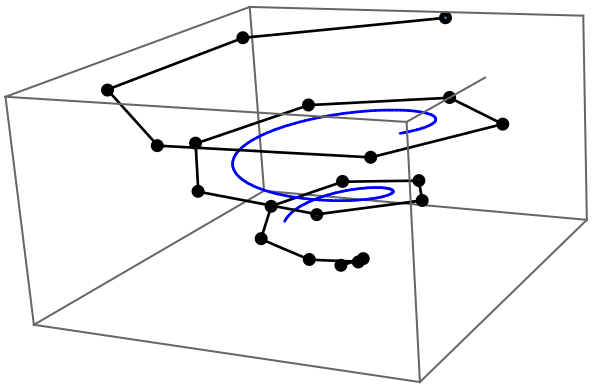


Fig. 3. Approximation of the shape of a conical spiral.

The curvature graph as well as the graph of curvature of torsion of the *B*-spline curve up to the second level are equivalent in shape to the second level curvatures of the conical spiral.

This technique will also be useful for approximating the log-aesthetic curves [13, 14] and Ziatdinov's superspirals [15]. In accordance with the multi-criteria approach to the evaluation of

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

log-aesthetic curves [16], when importing log-aesthetic curves to CAD systems in the form of NURBS templates, it is important to preserve the characteristics of the internal geometry using smoothness criteria.

6. Composite curves

Existing methods of modern CAD systems provide an order of smoothness of no more than G^4. With the float format of the *S*-polygon, it is easy to construct composite log-aesthetic curves by segments of *B*-spline curves of the same degree of *n* approximating log-aesthetic curves not lower than the second level of shape isogeometry. The *S*-polygons of the *B*-spline curves are positioned and matched by additional vertices so that the composite *S*-polygon has a harmonious, regular shape. Hence, a composite *B*-spline curve will have an order of smoothness of *n*-1.

7. Conclusion

This article demonstrates the techniques of applying *S*-polygons in a float format for modeling a high-quality *B*-spline curve using smoothness criteria. In particular, the technique of modeling the non-closed *B*-spline curve is proposed with the provision of accurate positioning of the end points of the curve and the directions of the tangent vectors at the end points. In addition, a technique for modeling a *B*-spline curve with a monotonic curvature through an *S*-polygon with a Mineur-Farin configuration is proposed, as well as a technique for approximating analytic curves with the equivalence of curvature and curvature of torsion to a level *m*. A technique has been developed for constructing smooth connections of *B*-spline curves with a high order of smoothness of the composite curve.

The proposed techniques are implemented in the program FairCurveModeler [17].

Acknowledgements

This manuscript was copy-edited by volunteers Mr. Duncan Griffiths (Canada), Mr. Anthony Dowds (UK) and Mr. Marty Smith (USA). Additionally, we thank Ms. Kimberley Hoffman (USA), Ms. Linda Sherwood (USA) and Mr. Seth I. Rich (USA) for useful remarks and suggestions. Without their kind assistance, this publication would not be possible.

References

[1] Mineur, Y., Lichah, T., Castelain, J.M., Giaume, H. A shape controlled fitting method for Bézier curves. *Computer Aided Geometric Design* 15 (9), 1998, pp. 879-891.

[2] Farin, G. Class A Bézier curves. *Computer Aided Geometric Design*, 23, 2006, pp. 573-581.

[3] Welters, H.J, Farin, G. Geometric curve approximation. *Computer Aided Geometric Design*, 14, 1997, pp. 499-513.

[4] Cox M. G. The numerical evaluation of *B*-splines, *Journal of the Institute of Mathematics and its Applications*, 10, 1972, pp. 134-149.

[5] De Boor, C., et al. A practical guide to splines. Vol. 27. New York: Springer-Verlag, 1978.

[6] Boehm, W. Cubic *B*-spline curves and surfaces in computer-aided geometric design. *Computing*, 19, 1977, pp. 29-34.

[7] Sabloniere, P. Spline and Bézier polygons associated with a polynomial spline curve. *Computer-Aided Design*, 10 (4), 1978, pp. 257-261.

[8] Prautzsch, H., Boehm, W., Paluszny, M. Bézier and *B*-spline techniques. Springer, 2002, 190 p.

[9] Muftejev, V.G., Romanyuk, A.N., Mardanov, A.R., Farkhutdinov, I.M. Geometrically stable modeling of NURBS curves and surfaces of arbitrary degrees. *Applied Geometry*, 11, pp. 19-77.

[10] Muftejev, V.G., Mardanov, A.R., Romanyuk, A.N., Turta, V.G., Farkhutdinov, I.M. Program of isogeometric modeling of high-quality curves. Web-application for CAD systems. Proceedings of the international scientific and technical internet-conference on computer graphics and image recognition. Vinnytsa, Ukraine. 2012. pp. 127-139.

[11] Golovanov, N.N. Geometric modeling. Moscow, Fizmatlit, 2002 (in Russian).

[12] Ziatdinov, R., Muftejev, V.G. On a class of plane curves with equal curvatures of the *n*-th level. In Systems of design, technological preparation of production and management of the life cycle of an industrial product (CAD/CAM/PDM - 2017). Institute of Control Problems, Russian Academy of Sciences. p.133.

[13] Yoshida, N., Saito, T. Interactive aesthetic curve segments. *The Visual Computer*, 22 (9-11), 2006, pp. 896-905.

[14] Miura, K.T. A general equation of log-aesthetic curves and its self-affinity. *Journal of the Japan Society for Precision Engineering*, 72 (7), 2006, pp. 857-861.

[15] Ziatdinov, R. Family of superspirals with completely monotonic curvature given in terms of Gauss hypergeometric function. *Computer Aided Geometric Design*, 29 (7), 2012, pp. 510-518.

[16] Muftejev, V.G., Ziatdinov, R. Functionality and aesthetics of curves in industrial design: a multi-criteria approach to assessing the quality of shapes in CAD systems of the future. *Vestnik Mashinostroyeniya*, Issue 7, 2018, pp. 23-27 (in Russian).

[17] FairCurveModeler app for AutoCAD v. 2019 https://apps.autodesk.com/ACD/en/Detail/Index?id=4526969846340104233&appLang=en&os=Win32_64&mode=preview&autostart=True&loginRequired=True

[18] Wang, A., Zhao, G. Counter examples of "Class A Bézier curves". *Computer Aided Geometric Design*, 61, 2018, pp. 6-8.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Использование чертежей с разметкой для проверки электронных чертежей и трехмерных моделей

А. А. Бойков, А. М. Федотов
albophx@mail.ru | fam1950@bk.ru
Ивановский государственный энергетический университет имени В. И. Ленина, Иваново, Россия

В статье предлагается использовать чертежи с разметкой для задания параметров поиска геометрических элементов в проверяемых чертежах и трехмерных моделях.

Ключевые слова: *автоматизированная проверка чертежей, автоматизированная проверка электронных моделей, автоматизированное проектирование.*

Using of drawings with markup for verifying of drawings and 3D-models

A. A. Boykov, A. M. Fedotov
albophx@mail.ru | fam1950@bk.ru
Ivanovo State Power University, Ivanovo, Russia

It is proposed to use drawings with markings to define parameters for searching geometric elements in drawings and 3D models while verifying them.

Keywords: *Computer aided verifying of drawings, computer aided verifying of 3D models, CAD.*

1. ВВЕДЕНИЕ

В работах [1, 4] отмечается актуальность контроля качества электронных моделей и разработки средств автоматизации такого контроля, но предлагаемые в них способы ориентированы на использование данных трехмерных геометрических моделей и сборок. В условиях, когда большая часть документации представлена еще электронными чертежами, контроль качества моделей требует значительного ручного труда (еще большего труда требует перевод архивов чертежей в трехмерные модели и сборки и способы контроля качества из [1, 4] оказываются неприменимы).

Предлагается способ частичной автоматизации работ по контролю качества геометрических моделей, представленных в виде электронных чертежей или трехмерных моделей, на основе данных электронных чертежей.

2. ПОСТАНОВКА ЗАДАЧИ ПРОВЕРКИ

Пусть требуется проверить возможность соединения (совместного использования) детали, информация о которой представлена в виде электронного чертежа или трехмерной модели, с узлом или деталью, представленной электронным чертежом (рабочим или сборочным). Возможность такого соединения обеспечивается, в общем случае, возможностью совмещения частей проверяемой детали с частями детали или узла, выполняющего роль образца.

Автоматизировать проверку можно если рассматривать ее как проверку решения конструктивной задачи инженерной геометрии (КЗ) [2, 3], как задачи создания конструктивной геометрической модели (КГМ, электронный чертеж и электронная трехмерная модель – частные случаи) некоторой обобщенной фигуры по заданным условиям. Условия могут касаться элементов в составе проверяемой фигуры (решения КЗ), но, главным образом, определяют связи ее с другими, внешними, фигурами.

В соответствии с алгоритмом проверки (способ наложения) автоматизация проверки требует [2]:

1. Формализовать требования к решению в виде эталона.
2. Установить соответствие между элементами решения и эталона и определить «правильность» как меру различия.

3. СОЗДАНИЕ ЭТАЛОНА ДЛЯ ПРОВЕРКИ

Эталон представляет собой совокупность геометрических и структурно-логических данных. Геометрическая часть эталона – это параметры фигур, по отношению к которым проверяется выполнение заданных условий. В рассматриваемой задаче геометрическую часть эталона составляют объекты исходного чертежа. Структурно-логическая часть определяет проверку заданных условий и составляется с использованием специальных элементов [2]:

- *and* – объединяет группу целевых фигур .
- *or* – задает альтернативные варианты фигуры или группы.
- *transform* – задает возможность геометрического преобразования (сдвиг, поворот и др.) для элемента или группы.
- *instance* – задает образец, сравнение с которым осуществляется прямой проверкой геометрических условий.

Типовая структура эталона для проверки трехмерной модели, с учетом того, что проверяемая фигура может произвольным образом располагаться в моделируемом пространстве, выглядит так – *transform (and (элементы модели))*. КЗ с такими эталонами были выделены в класс задач *со свободным расположением* (ЗСР). Типовая структура эталона для проверки многовидового чертежа, с учетом того, что изображения обладают свободой перемещения относительно друг друга в соответствии с ГОСТ 2.305, выглядит так – *and (transform (and (элементы изображения$_1$)), transform (and (элементы изображения$_2$)), ...)*. КЗ, решения которых состоят из относительно независимых групп элементов, выделены в класс *комбинированных задач* (ЗК).

Структура эталонов ЗСР и ЗК может быть сформирована автоматически, поэтому создание эталона состоит в разметке исходного чертежа, в ходе которой указываются целевые фигуры и связанные с ними правила. Для этого используются выделение, маркеры и фигуры-заместители:

- *выделение* позволяет задать группе фигур общее при проверке свойство, для выделения удобно использовать раскрашивание или перемещение в особый слой.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

- *маркер* позволяет задать элементу особые правила проверки, например, набор допустимых значений диаметра для отверстия или направление присоединения (рис. 1, *а*). В качестве маркеров удобно использовать элементы-обозначения (размерные объекты, линии-выноски) или блоки.
- *фигура-заместитель* позволяет ввести в эталон объект, отсутствующий среди фигур геометрического редактора, и особым образом ведущий себя в ходе проверки. В качестве фигур-заместителей могут выступать блоки или специально помеченные обычные фигуры. Например, прямоугольный блок *zoneObject* позволяет явным образом указать одно изображение многовидового чертежа, а блок *formatObject* ввести в эталон рамку чертежа, как целое (рис. 1, *б*).

Чертеж с разметкой преобразуется в эталонную КГМ, в ходе сравнения с которой осуществляется проверка модели.

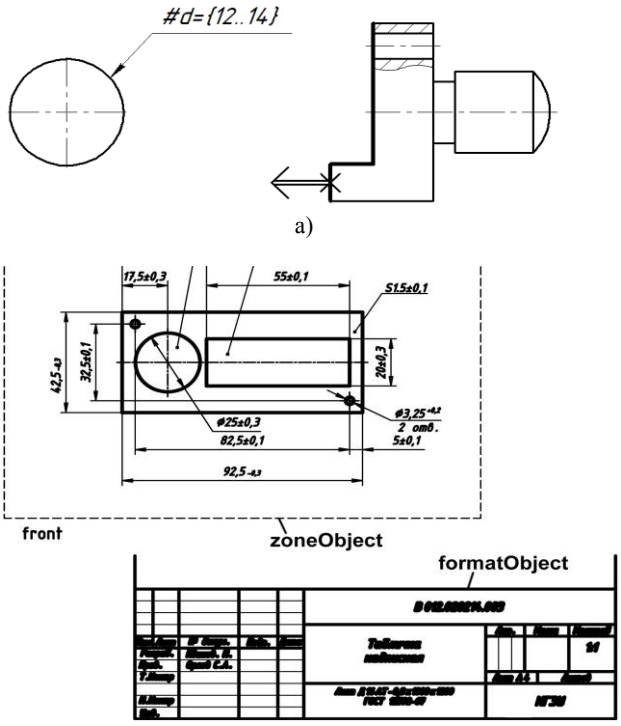


Рис. 1. Элементы разметки чертежа

4. ПРОВЕРКА ЭЛЕКТРОННОГО ЧЕРТЕЖА

В ходе проверки система последовательно обходит элементы структуры эталона. В элементах, соответствующих фигурам, производится поиск наиболее подходящих фигур в проверяемом чертеже и определяется мера различия. В специальных элементах *and, or, transform*, в соответствии с соотношениями, приведенными в [2], упорядочиваются комбинации соответствий, чтобы по окончании обхода всего эталона мера различия была минимальной.

Рассмотрим работу системы проверки на примере простой одновидовой задачи.

На рис. 2, а приведен чертеж детали «подушка». Эталонный чертеж представлен в формате *DXF*. Необходимо установить, могут или нет детали, представленные на проверяемых чертежах, использоваться совместно с нею. Такую возможность обеспечивает, в первую очередь, соответствующее расположение и размеры отверстий, служащих для соединения деталей. Поскольку все требуемые элементы расположены на одном виде, имеем задачу класса ЗСР.

Некоторые элементы чертежа выделены – перенесены в специальный слой – как «отображаемые» (показаны утолщенной линией), при проверке они отрисовываются поверх проверяемого чертежа. «Проверяющие» элементы обозначены. Обозначение помещено на полке линии-выноски и начинается с символа «#». Текст после «#» используется для их настройки: так, последовательность символов «#d>14» показывает, что в проверяемом чертеже соответствующий элемент (окружность) должен сопоставляться с окружностями диаметром больше 14 мм, при этом центры эталонной окружности и проверяемой должны совпадать в пределах общего преобразования эталона; последовательность символов «#d=$» требует равенства диаметра проверяемой окружности диаметру эталонной (значение $), элементы с символами «#lim:x» и «#lim:y» задают габаритные размеры и т. п. Размеченный один раз эталон может использоваться многократно.

Проверка выполняется, как показано в [2]. Центры окружностей служат в качестве опорных точек преобразования. Для эталона на рис. 2, а таких точек 6. Для определения параметров разрешенного преобразования (поворот и масштаб) требуется две точки. Таким образом, в эталоне имеется 30 пар опорных точек.

Рассмотрим проверку чертежа детали «плита направляющая» (рис. 2, б). На чертеже – 10 окружностей с шестью различными центрами, что также дает 30 пар опорных точек и 900 вариантов преобразования (матрицы преобразования ввиду громоздкости не приводятся), из которых 394 различных (два – с приоритетом 30, по четыре – с приоритетом 12 и 6 и т. д.). При проверке используются параметры преобразования, имеющего наибольший приоритет. Два преобразования с приоритетом 30 соответствуют двум способам наложения эталонных окружностей на окружности проверяемого чертежа. Один из вариантов показан на рис. 2, б. После того, как эталонные фигуры преобразованы, для каждой из них выполняется поиск элемента, соответствующего по значениям параметров и дополнительным правилам (см. выше). В конце проверяемый чертеж отображается, и схема дополняется изображениями элементов, помеченных в эталоне как «отображаемые».

Достоинством предлагаемого способа проверки является то, что визуальная схема проверки, дополненная элементами эталона (после совмещения) и маркерами ошибок, может быть сохранена в виде файла в формате *DXF* для последующего анализа.

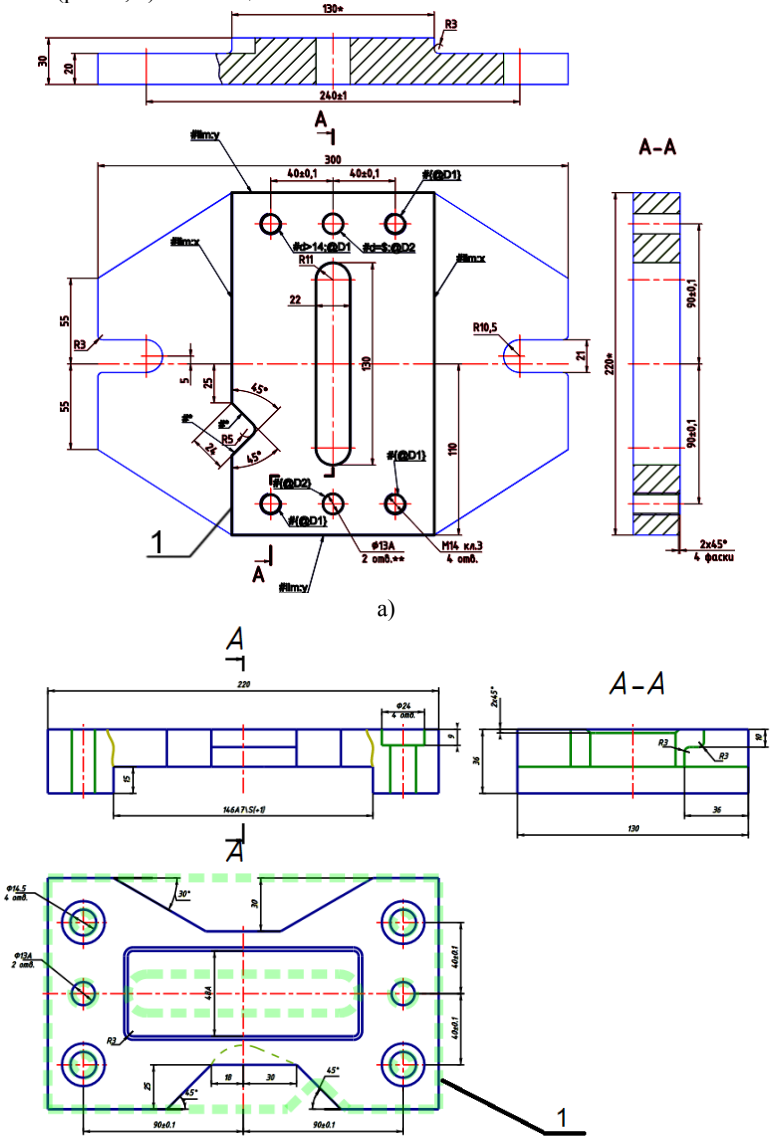
На рис. 3 приведен пример проверки чертежа детали «матрица», выполненного в масштабе 2:1, при помощи того же эталона. В чертеже удалено одно отверстие и диаметры двух других заданы равными 12 (в эталоне – 13). Чертеж содержит 14 окружностей (выполненное «вручную» обозначение выносного элемента содержит окружность, которая рассматривается вместе с остальными) с десятью различными центрами, что дает 90 пар опорных точек, 2700 вариантов преобразования, из которых 1314 – различные: два имеют приоритет 20, два – 12 и т. д. Поиск на чертеже элементов, соответствующих эталонным, показывает, что окружности, диаметры которых равны 12 вместо 13, помечены ошибкой «Нарушение правила (d=$)», отсутствующий на чертеже эталонный элемент (вставлен в проверяемую модель с ключом «+192») помечен ошибкой «Элемент отсутствует», что подтверждает правильность результатов проверки.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Приведенный способ проверки может быть также полезен для организации поиска в электронном архиве чертежей, если выполнить проверку нескольких чертежей подряд с одним и тем же эталоном.

5. ПРОВЕРКА ТРЕХМЕРНОЙ МОДЕЛИ

Рассмотрим проверку трехмерной электронной модели детали «матрица» в формате *IGES* (рис. 4, а) с помощью шаблона, показанного выше. Чтобы иметь возможность сравнения элементов эталона с элементами проверяемой модели, спроецируем ее. Центры окружностей эталона, входящие в определитель преобразования (сдвиг, поворот и масштаб), дополним *z*-координатой, равной 0.



а)

б)

Рис. 2. Эталонный чертеж детали «подушка» с разметкой в окне редактора *AutoCAD* (а) и визуальная схема результатов проверки (штриховки скрыты) детали «плита направляющая» (б):
1 – контур, составленный из отображаемых элементов эталона

Пространственное преобразование движения и масштаба требует указания соответствующих пар трех точек, не лежащих на одной прямой. В эталоне таких троек, не считая перестановок, – 18.

В качестве точек, определяющих преобразование пространственной модели используем центры оснований полных цилиндров. В модели «матрицы» их 18, что дает слишком большое число вариантов определителя (верхняя оценка – 4896). Для сокращения перебора можно наложить дополнительные условия. Например, задав отношение длин сторон опорного треугольника равным одному из 18 треугольников эталона, получаем 212 вариантов определителя. Это дает 8 различных вариантов преобразования, 4 – имеют приоритет 52 – это четыре способа совместить цилиндрические отверстия с их проекциями в картинной плоскости. Один из вариантов приведен на рис. 4, б. Проверка подтвердила «правильность» модели.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

6. ЗАКЛЮЧЕНИЕ

В работе показан способ решения частной задачи проверки электронных чертежей и трехмерных моделей на основе геометрических данных электронного чертежа – проверки возможности соединения (совместного использования) детали с данной деталью или узлом. В качестве эталона могут быть использованы чертежи деталей, сборочные чертежи или отдельные изображения. Усложнение правил проверки достигается при помощи специальной разметки. Планируется совершенствование этого способа для работы с многовидовыми эталонами.

Разметка геометрических элементов (например, раскрашивание граней) может быть использована и для задания параметров сравнения трехмерных моделей.

Предлагаемый в работе способ проверки геометрических моделей может быть полезен при выполнении некоторых проектных работ, для организации поиска чертежей и моделей в электронных архивах, а также для автоматической проверки решений некоторых учебных задач в автоматизированных обучающих системах.

7. ЛИТЕРАТУРА

[1] Блинова А.А., Гаврилова Н.Ю., Пащенко О.Б. Методы контроля твердотельных электронных моделей машиностроительного изделия на всех этапах его жизненного цикла // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. – 2012. – № 2. – С. 80–84.

[2] Бойков, А., Федотов, А. Автоматическая проверка решений задач инженерной геометрии // GraphiCon 2016. Труды Международной научной конференции. – Москва-Протвино, 2016. – С. 352–355.

[3] Бойков, А., Федотов, А. Использование полифигур для валидации и верификации вычислительных моделей в системе компьютерной проверки решений задач инженерной геометрии // GraphiCon 2017. Труды Международной научной конференции. – Пермь, 2017. – С. 335–338.

[4] Пащенко О. Б., Серебрякова И. Л. Метод контроля качества электронной модели машиностроительного изделия с использованием информационных технологий // Инженерный вестник. – 2014. – №12. – С. 125–130.

ОБ АВТОРАХ

Федотов Александр Михайлович, к.т.н., доцент кафедры конструирования и графики Ивановского государственного энергетического университета. E-mail fam1950@bk.ru.

Бойков Алексей Александрович, доцент кафедры конструирования и графики Ивановского государственного энергетического университета. E-mail albophx@mail.ru.

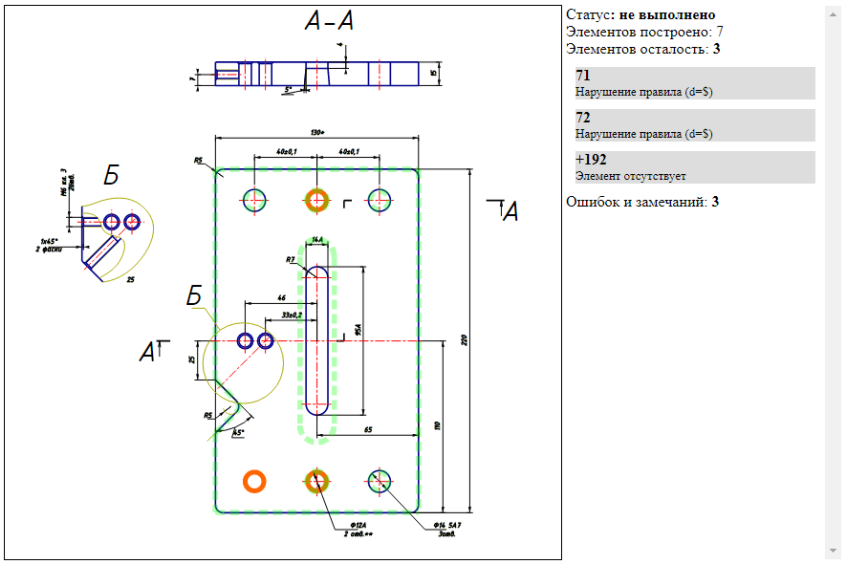
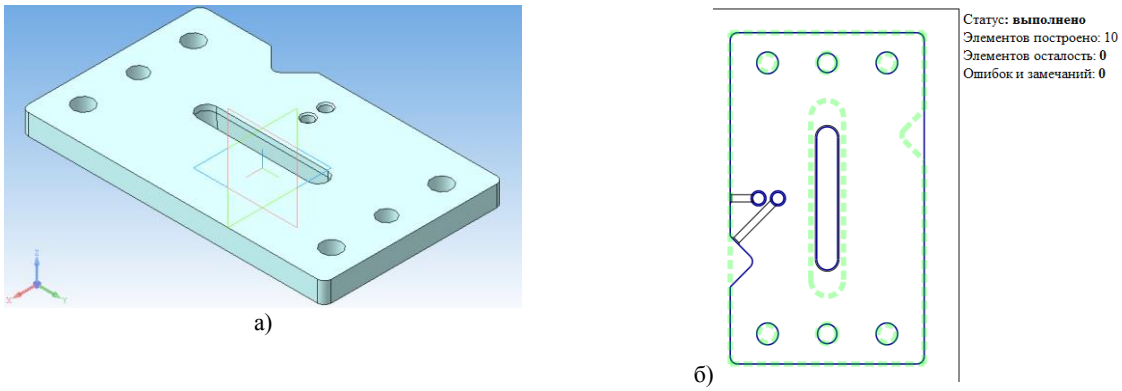


Рис. 3. Визуально-текстовая схема результатов проверки чертежа детали «матрица»



а) б)

Рис. 4. Трехмерная модель детали в окне САПР «Компас-3D» (а) и визуально-текстовая схема результатов проверки (б)

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

АВТОМАТИЗИРОВАННАЯ ТЕХНОЛОГИЯ ВВОДА ГРАФИЧЕСКИХ ДОКУМЕНТОВ СО СЛАБО ФОРМАЛИЗОВАННЫМ ОПИСАНИЕМ ОБЪЕКТОВ

Васин Д.Ю.[1], Громов В.П.[1], Кустов Е.А.[1], Ротков С.И.[2] , Ясаков Ю.В.[1]
dm04@list.ru, gvp1941@bk.ru, kustov@pmk.unn.ru, rotkov@nngasu.ru, yuri961@yandex.ru
Нижний Новгород, Россия,
[1]ИТММ, Нижегородский Государственный Университет им. Н.И. Лобачевского
[2]Нижегородский Государственный Архитектурно-Строительный Университет

В докладе в рамках развития комбинаторно–геометрического подхода к обработке пространственно–распределенных данных изложены базовые технологические этапы автоматизированной технологии ввода графических документов со слабо формализованным описанием объектов. В качестве примера рассмотрено автоматическое распознавание объектов на планах поэтажной застройки зданий.

Ключевые слова: *растровые изображения, планы поэтажной застройки зданий, комбинаторно–геометрический подход к обработке пространственно–распределенных данных, большеформатные, сложноструктурированные графические документы со слабо формализованным описанием объектов, технология автоматического распознавания*

THE AUTOMATED TECHNOLOGY OF INPUT OF GRAPHIC DOCUMENTS WITH POORLY FORMALIZED DESCRIPTION OF OBJECTS

Vasin D.Yu.[1], Gromov V. P.[1], Kustov E.A. [1], Rotkov S.I. [2], Yasakov Yu.V.[1]
dm04@list.ru, gvp1941@bk.ru, kustov@pmk.unn.ru, rotkov@nngasu.ru, yuri961@yandex.ru
Nizhny Novgorod, Russia,
[1] ITMM Nizhny Novgorod State University. N.I. Lobachevsky
[2]Nizhny Novgorod State Architectural and Construction University

In the report within development of combinatory and geometrical approach to processing of the spatial distributed data basic technological stages of the automated technology of input of graphic documents with poorly formalized description of objects are stated. As an example automatic recognition of objects on plans of floor-by-floor building of buildings is considered.

***Keywords**: raster images, plans of floor-by-floor building of buildings, .kombinatorno-geometrical approach to processing of the spatial distributed data, bolsheformatny, slozhnostrukturirovanny graphic documents with poorly formalized description of objects, technology of automatic recognition*

1. Введение

В данной работе в качестве исходной информации рассматриваются только изображения большеформатных, семантически–насыщенных, сложноструктурированных графических документов (БФГД). К ним отнесем топографические карты и планы, планы поэтажной застройки зданий, машиностроительные чертежи, технологические схемы и пр. На такие документы наносятся условные обозначения объектов четырех классов: точечных (дискретных), линейных, двумерных (площадных) и изображения символов, определяя вкупе пространственно–распределенные данные (ПРД) [4, 5]. Объекты ПРД распределены в пространстве, имеют координатную привязку (X,Y,Z), метрическое описание объектов сложноструктурировано, может включать в себя, полностью или частично, другие объекты, менять свойства (атрибуты) в зависимости от их местонахождения. Объекты могут взаимодействовать между собой, имея либо пространственно−логические (ПЛС), либо тополого−метрические связи (ТМС), число которых может значительно изменяться.

Характеристики (атрибуты) объектов также имеют пространственную привязку, могут меняться в зависимости от местонахождения объекта, иметь сложноструктурированный характер, то есть состоять (включать в себя) большой набор уточняющих свойств.

Значительный подкласс БФГД, включающий в себя документы массового типа (машиностроительные чертежи, схемы, планы поэтажной застройки зданий и пр.), которые; как правило, исполнены ручным способом на бумажных носителях среднего и плохого качества, имеют существенное разнообразие знакового наполнения, большую плотность нанесения графических знаков на поверхность ГД, значимые отклонения от нормативного изображения графических объектов, их произвольную ориентацию, содержат эффекты межобъектных наложений, примыканий и пересечений, прочие искажения при нанесении знаков, приводящие к касанию и слипанию знаков и пр. На БФГД типа поэтажных планов зданий базовыми графическими элементами являются в основном ортогональные линейные объекты и окружности (дуги окружностей), которые находятся в сложных топологических отношениях друг с другом

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

(примыкания, наложения, пересечения), ярко выражена топологическая нагрузка на базовые графические элементы – линейные отрезки, выражающаяся в возможном многократном использовании одного и того же отрезка в описании нескольких графических знаков. В зависимости от качества носителя БФГД и правильности выбора параметров оцифровки, степень зашумленности растрового изображения графического документа (РИГД), а также метрическая точность отдельных графических знаков варьируется широком диапазоне значений.

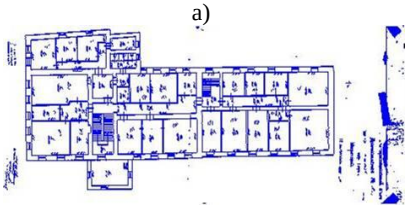
Указанные особенности позволяют выделить эти БФГД в отдельный подкласс со слабо формализованным начертанием объектов (СФГД) и обуславливают повышенные требования к их геометрическому моделированию.

Ниже представлена автоматизированная технология создания цифровых баз данных СФГД на примере поэтажных планов зданий, которая, однако, может быть легко адаптирована к созданию аналогичных баз данных для СФГД других типов.

2. Технологические этапы обработки СФГД

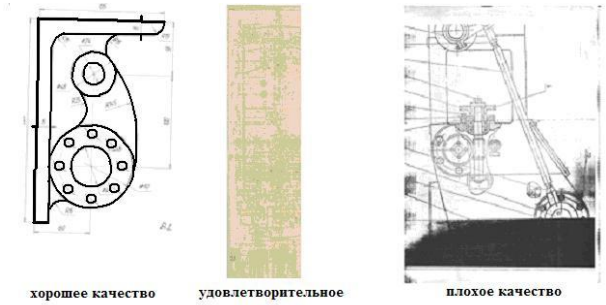
Первичный анализ ГД. На данном этапе выполняется визуальная оценка качества исходного ГД и принимается решение о возможности его автоматической или интерактивной векторизации (рис. 1 – 3). Если качество исходного ГД такое, что применение автоматизированных технологий ввода признается нецелесообразным, то оцифровка и ввод такого документа исключительно интерактивные. Круг возникающих при этом проблем и способов их разрешения лежит за пределами обсуждения данной статьи.

a)



b)

Рисунок 1 план поэтажной застройки здания: a) плохого качества b) удовлетворительного качества

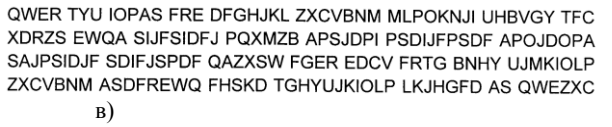


a) b) c)

Рисунок 2 технические чертежи: a) хорошего, b) удовлетворительного и c) плохого качества

а)

б)



в)

Рисунок 3 а) планшет гидрографической съемки хорошего качества; б) планшет гидрографической съемки среднего качества; в) отсканированный текст

Сканирование. Перевод бумажного чертежа в электронный вид осуществляется посредством планшетных сканеров с разрешением 100–300 dpi в полноцветном режиме с последующей цветовой фильтрацией и бинаризацией. Несмотря на то, что большинство ГД указанного класса выполнятся черно–белыми, изначальное сканирование их в режиме монохромного изображения может привести к существенным искажениям, т.к. довольно малое количество информации представлено чистыми черным и белым цветами, более того, некоторые ГД представлены так называемыми «синьками», основные цвета которых лежат в бордово–фиолетовой части спектра. Для их минимизации требуется постепенное уменьшение глубины цвета путём преобразования цвета всех пикселей изображения к требуемому диапазону.

Представляется предпочтительным использование растрового формата tif в качестве базового для хранения исходного РИГД. Данный формат имеет теговую структуру и допускает создание собственных информационных тегов, что исключает необходимость создания дополнительных информационно–описательных файлов в процессе автоматизированной обработки. Кроме того, данный формат имеет хорошую программную поддержку, существует достаточное количество свободно–распространяемых библиотек с открытым кодом для поддержки данного формата. Открытость кода позволяет вносить в него необходимые изменения для поддержки новых тегов.

Преобразование исходного РИГД в оттенки серого. Изначально СФГД сканируется с глубиной цвета 24 бит/пиксель. Данный режим сканирования является избыточным, поскольку источниками исходных данных являются различные СФГД с основными цветами их представления белый (цвет фона) и чёрный (цвет объектов), но он позволяет выполнить последующую программную бинаризацию РИГД с минимизацией возможных искажений. Далее осуществляем программную конвертацию полноцветного РИГД в оттенки серого путем применения к каждому пикселю исходного РИГД преобразования вида:

Y = 0.299 R + 0.587 G + 0.144 B,

в котором R, G, B – красная, зелёная и синяя компонента цвета исходного РИГД соответственно, а Y – полученная яркость в оттенках серого.

Предварительная обработка, контрастирование и бинаризация растра. С целью устранения различных искажений входного растрового изображения, обусловленных невысоким качеством исходных СФГД, особенно исполненных на бумажных носителях, в технологии предусмотрены этапы линейного контрастирования исходного РИГД.

Линейное контрастирование. Использовалось линейное поэлементное преобразование вида:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}}(y_{\max} - y_{\min}) + y_{\min}$$, где y – выходная яркость контрастированного пикселя, y_{min} и y_{max} желаемые значения минимальной и максимальной выходной яркости при экстремальных значениях исходной яркости x_{min}. и x_{max}

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Бинаризация. Очевидно, что качество формируемой векторной модели СФГД напрямую зависит от качества выполнения операции бинаризации исходного полутонового РИГД, поскольку неудачный выбор порога бинаризации приводит к появлению различных шумов, среди которых можно выделить неправомочное увеличение количества выходных линейных объектов, появляющиеся вследствие наличия на исходном РИГД ложных разрывов линейных растровых объектов, а также различные нарушения топологической целостности из–за ложных слипаний объектов, или, наоборот, при отсутствии необходимой связности растровых объектов. Данные помехи вносят значительный вклад в снижение качества итогового результата векторизации.

В зависимости от вида СФГД предлагается использовать три широко известных метода бинаризации: Оцу[2], двойного ограничения порога[16] и Брэдли–Рота[15]. Метод Оцу показал удовлетворительные результаты на полутоновых изображениях чертежей и планов поэтажной застройки зданий хорошего качества. Однако, результат обработки имеет и явный недостаток: при обработке объектов, имеющих различные размеры, но близкие цвета, после бинаризации Оцу получаются темные пятна. Достоинствами метода Брэдли-Рота являются простота реализации и высокая скорость выполнения; для большинства случаев нет нужды подбирать параметры. Метод хорошо работает с неоднородным фоном. Из последнего достоинства логически вытекает недостаток метода Брэдли, а именно, плохая чувствительность к низкоконтрастным деталям изображения.

Интерактивное редактирование – осуществляется любым доступным растровым графическим редактором. Основная задача редактирования – интерактивное разделение неправомочно слипшихся объектов на РИГД, либо наоборот восстановление их связности, если это необходимо. С целью повышения эффективности выполнения интерактивных редакторских функций был разработан ряд специализированных интеллектуальных редакторов как растровых, так и векторных графических данных [9–11].

Штриховая форма представления РИГД. Поскольку растровые данные существенно информационно избыточны, то для их обработки желательно использовать те или иные представления, уменьшающие степень этой избыточности**.** Одной из форм экономичного представления бинарного растрового изображения (БРИ) является штриховая (рис.4). Штрих – есть одномерный кластер заданного цвета вдоль строки сканирования растра. Для БРИ штрих описывается тройкой

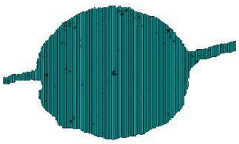


Рисунок 4 Штриховая форма представления РИГД

$\{N, K, Y\}$, где $N = j \,\&\, \overline{P}_{ij} \,\&\, P_{ij}$, $K = j \,\&\, P_{ij} \,\&\, \overline{P}_{ij}$, $Y = i$,

N, K – координаты начала и конца штриха черного цвета, расположенного в строке с номером Y, P_{ij} – код цвета пикселя с координатами (i, j). Использование штрихового формата позволяет сократить объем исходного растрового описания в 10–20 раз. Степень сжатия зависит от насыщенности РИГД разноцветными связными компонентами. Одновременно, большинство алгоритмов локальной фильтрации, сглаживания, форматирования и векторизации также в среднем в 10–20 раз эффективнее по времени для штрихового описания РИГД по сравнению с пиксельным[3 – 5].

Геометрическая коррекция растра. На данном этапе устраняются геометрические искажения (бумагу вставили в сканер с небольшим угловым отклонением), неправильная ориентация («боком» или «вверх ногами»), нестандартный формат (отсканированный с перекосом документ будет иметь нестандартные размеры) и т.п. Допускается автоматическая коррекция по 4 узловым точкам или по сетке узлов. Кроме того, выполнение операции нормализации позволяет в случае необходимости создавать различные мозаики (сшивать, накладывать друг на друга) из обрабатываемых СФГД [6].

Объектовая фильтрация растра. Для минимизации побочного эффекта масочных фильтров, выражающегося в уменьшении резкости границ объектов, был разработан алгоритм объектовой фильтрации БРИ в штриховой форме [3, 9-11, 14]. Алгоритм позволяет отфильтровывать СРК, которые по своим геометрическим характеристикам удовлетворяют параметрам помехи. Важным достоинством данного алгоритма является тот факт, что указанная фильтрация не вносит никаких искажений в те связные области, которые не распознаются алгоритмом как шумовые. На рис 5a представлены исходный растр поэтажного плана здания в штриховом формате, а результат его объектовой фильтрации на рис. 5b.

a b

Рисунок 5 a) исходное растровое изображение поэтажного плана здания в штриховом формате, b) результат его фильтрации объектовым фильтром

Векторизация – является базовой операцией в большинстве систем обработки и анализа графической информации. В частности, решение задачи синтеза 3D-модели объектов по наборам их 2D-проекций осуществляется исключительно по векторным описаниям. Если исходное РИГД не содержит искажений и шумов, то существующие локальные алгоритмы вполне удовлетворительно справляются с задачей векторизации, хотя, получаемые при этом векторизованные объекты, требуют дополнительного сглаживания, либо аппроксимации, а это не всегда удается сделать успешно, т.е. не всегда удается одновременно удовлетворить требованиям метрической точности аппроксимации и геометрической достоверности векторизуемых объектов. В результате выполнения операции векторизации получается контурная модель изображения (рис. 6)[3-5].

Рисунок 6 Векторная (контурная) модель поэтажного плана здания

Поскольку, для рассматриваемого класса ГД характерно изображение сюжетов рисунка с помощью отрезков линий с четко выраженным направлением их ориентации, то если при их сканировании направление линий

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

сканирования выбрать в соответствии с преимущественной ориентацией отрезков на изображении, то штриховая модель отсканированных данных позволяет построить простые и надежные алгоритмы векторизации и осуществить эффективное геометрическое моделирование РИГД [7, 8, 12].

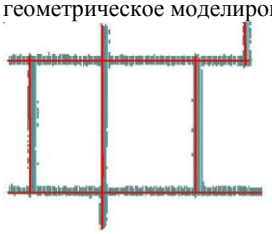


Рисунок 7 Линейная модель изображения

Для данного класса ГД с учетом указанных особенностей изображений на этапе векторизации эффективнее построение линейной модели изображения (рис.7).

Автоматическое распознавание условных знаков конструктивных элементов поэтажного плана здания. Ключевым моментом повышения эффективности преобразования различной бумажной конструкторской документации, в частности, планов поэтажной застройки зданий в цифровой вид, является автоматическое распознавание условных знаков конструктивных элементов поэтажного плана здания. Для поэтажных планов зданий в автоматическом режиме распознаются следующие конструктивные элементы: *Капитальная стена*, *Перегородочная стена*, *Дверь*, *Дверь внутри капитальной перегородки, Окно, Стенной шкаф, Ниша в стене, Помещения*, *Комната, Лестница*, После этого автоматически формируются объекты *Лестничный марш* и *Лестничная площадка*. Все оставшиеся нераспознанные контура получают выходной классификационный тип *Контур* для их дальнейшей интерактивной идентификации.

Поскольку один и тот же сегмент может одновременно входить в метрику сразу нескольких условных знаков, то решение задачи установления тополого–метрических связей между различными объектами поэтажных планов существенно упрощается и в некоторых случаях решается уже на этапе распознавания.

3. Благодарности

Работа выполнена при финансовой поддержке Российского Фонда Фундаментальных Исследований, гранты РФФИ № 17–07–00543 и № 18–07–00715.

4. Литература

1. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. – М.: Физматлит, 2001. – 784 с.
2. N. Otsu. A thresholding selection method from gray-scale histogram // IEEE Transactions on System, Man, and Cybernetics, 1979. No. 9. P.62-66
3. Васин Ю.Г., Васин Д.Ю., Громов В.П. Структурное описание растровых данных. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 23–29.
4. Васин Ю.Г., Башкиров О.А., Чудинович Б.М. Комбинаторно–геометрический подход в задачах анализа сложной графической информации. //Автоматизация обработки сложной графической информации: Межвуз. сб. науч. тр. / Под ред. Ю.Г. Васина: Горьк. гос. ун–т. Горький, 1987
5. Васин Ю.Г., Башкиров О.А., Рудометова С.Б. Математические модели структурированного описания графических изображений. //Автоматизация обработки сложной графической информации: Межвуз. сб. науч. тр. / Под ред. Ю.Г. Васина: Горьк. гос. ун–т. Горький, 1984
6. Громов В.П. Формирование нормализованных растров графических документов. //Методы и средства обработки сложной графической информации: 8 Всероссийская научная конференция, 12–16 сентября 2005г., Н.Новгород, стр. 51–54.
7. Горбунов А.А. Громов В.П. Робастная векторизация растровых линейных объектов. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 69–70.
8. Громов В.П. Построение геометрической модели графических документов с выраженной ориентацией линейных объектов. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 70 – 71.
9. Васин Д.Ю., Громов В.П. Разработка и создание базовых интеллектуальных функций редактора растровых графических данных. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 19 – 20.
10. Васин Д.Ю., Громов В.П. Интеллектуальный редактор графических данных. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 21 – 23.
11. Васин Ю.Г., Васин Д.Ю. Развитие возможностей интеллектуального редактора в задаче оперативного обновления растровых изображений. //Методы и средства обработки сложной графической информации: 8 Всероссийская научная конференция, 12–16 сентября 2005г., Н.Новгород, стр. 8 – 10.
12. Громов В.П., Смирнов А.Ф. Векторизация линейных объектов по штриховой модели бинарного растра. //Методы и средства обработки сложной графической информации: 8 Всероссийская научная конференция, 12–16 сентября 2005г., Н.Новгород, стр. 55 – 56.
13. Беляева Ю.В., Васин Ю.Г., Васин Д.Ю., Громов В.П. Подсистема создания цифровых баз данных поэтажных планов. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 14 – 16.
14. Васин Ю.Г., Васин Д.Ю., Громов В.П. Подсистема редактирования и обработки растровых графических изображений. // Труды III конференции «Распознаваниеобразов и анализ изображений: новые информационные технологии». РОАИ–97. Н.Новгород, 1–7 декабря 1997г., стр. 19–22.
15. Derek Bradley. Adaptive Thresholding Using the Integral Image / Derek Bradley, Gerhard Roth // Carleton University, Canada 2007.
16. M.I. Sezan, A peak detection algorithm and its application to histogram-based image data reduction //Computer Vision, Graphics, and Image Processing, 1990. Vol. 49 (1) P.36-51.

Об авторах

Васин Дмитрий Юрьевич – к.т.н., ведущий научный сотрудник Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского)
E–mail: dm04@list.ru.

Громов Владимир Петрович – ведущий программист Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского)

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

E–mail: gvp1941@bk.ru.

Кустов Евгений Александрович – ведущий программист Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского)

E–mail: kustov@pmk.unn.ru.

Ротков С.И. – д.т.н., профессор, зав. кафедрой инженерной геометрии, компьютерной графики и автоматизированного проектирования ННГАСУ им. В.П. Чкалова

E–mail: rotkov@nngasu.ru.

Ясаков Юрий Васильевич – старший научный сотрудник Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского)

E–mail: , yuri961@yandex.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

МОДЕЛИ ПРЕДСТАВЛЕНИЯ РАСТРОВЫХ ГРАФИЧЕСКИХ ДОКУМЕНТОВ СО СЛАБО ФОРМАЛИЗОВАННЫМ ОПИСАНИЕМ ОБЪЕКТОВ

Васин Д.Ю.[1], Громов В.П.[1], Ротков С.И.[2]
dm04@list.ru|gvp1941@bk.ru|rotkov@nngasu.ru
Нижний Новгород, Россия,
[1]ИТММ, Нижегородский Государственный Университет им. Н.И. Лобачевского
[2]Нижегородский Государственный Архитектурно-Строительный Университет

В докладе в рамках развития комбинаторно–геометрического подхода к обработке пространственно–распределенных данных изложены оригинальные модели представления растровых и векторных графических документов со слабо формализованным описанием объектов, предложены алгоритмы их формирования и поддержки, а так же разработанные на основе указанных моделей алгоритмы фильтрации растровых данных и распознавания объектов по растровым данным планов поэтажной застройки, символов на растровых изображениях текстовых документов.

Ключевые слова: *растровые изображения, векторные изображения, планы поэтажной застройки зданий, гиперспектральные растровые данные ДЗЗ, комбинаторно–геометрический подход к обработке пространственно–распределенных данных, большеформатные, сложноструктурированные графические документы со слабо формализованным описанием объектов, фильтрация растровых данных, автоматическое распознавание символов.*

MODELS OF SUBMISSION OF RASTER GRAPHIC DOCUMENTS WITH POORLY FORMALIZED DESCRIPTION OF OBJECTS

Vasin D.Yu.[1], Gromov V. P.[1], Rotkov S.I. [2]
dm04@list.ru|gvp1941@bk.ru|rotkov@nngasu.ru
Nizhny Novgorod, Russia,
[1]ITMM Nizhny Novgorod State University. N.I. Lobachevsky
[2]Nizhny Novgorod State Architectural and Construction University

In the report within development of combinatory and geometrical approach to processing of the spatial distributed data original models of submission of raster and vector graphic documents with poorly formalized description of objects, the developed algorithms of their formation and support, and algorithms of filtration of raster data and recognition of objects which are also developed on the basis of the specified models according to raster data of plans of floor-by-floor building and symbols narastrovy images of text documents are stated.

Keywords: *raster images, vector images, plans of floor-by-floor building of buildings, hyper spectral raster data of remote sensing of the earth, combinatory and geometrical approach to processing of the spatial distributed data, bolsheformatny, slozhnostrukturirovanny graphic documents with poorly formalized description of objects, filtration of raster data, automatic recognition of symbols.*

Введение

Цифровая обработка графической информации (ГИ) относится к числу наиболее трудоемких задач современной кибернетики, информатики и вычислительной техники. При этом ГИ является наиболее естественным носителем исходной информации практически во всех областях науки и техники. Цифровая обработка графической информации широко используется при решении многих важных отраслевых задач, автоматизации проектирования (САПР), автоматизации научных исследований (АСНИ), в робототехнике, медицинской и технической диагностике, мониторинге природных ресурсов, геоинформационных технологиях (ГИС) и т.д.

Расширение сфер, требующих автоматизации обработки графической информации, привело к качественному усложнению графических документов (ГД)– большие размеры, отсутствие строгих ограничений на шрифты и типоразмеры, наличие фона, большое количество пространственно–логических и топологических связей и метрических отношений между линейными, площадными и дискретными объектами. Практикой выдвинута проблема обработки большеформатных, сложноструктурированных графических данных (БФГД) БФГД – это материальный объект (конструкторские проекты, географические и топографические карты и планы, графические материалы медицинской диагностики и поэтажные планы БТИ, растровые данные дистанционного зондирования Земли – ДЗЗ и др.), содержащий образно–знаковые модели действительности в форме графических изображений и терминов естественного языка. Размеры документов не менее 1000 мм по обеим осям, объекты могут изменяться от 0.1 мм до нескольких метров, точность обработки составляет порядка 0.1 мм. Это предопределяет огромные размеры исходных данных, которые превышают десятки и сотни

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

мегабайт. При этом источником, как правило, являются графические изображения на бумажных носителях, а создание цифровых документов необходимо обеспечить в терминах соответствующей проблемной области. Однако, значительный класс БФГД выполняется ручным способом со слабо формализованным начертанием объектов, произвольной ориентацией, произвольными размерами символов, наличием большого количества межобъектных наложений, примыканий и пересечений, образуя подкласс БФГД со слабо формализованным описанием объектов (СФГД). На такие документы наносятся условные обозначения объектов четырех классов: точечных (дискретных), линейных, двумерных (площадных) и изображения символов – пространственно–распределенные данные (ПРД).

Объекты ПРД распределены в пространстве, имеют координатную привязку (X,Y,Z), метрическое описание объектов сложноструктурировано, может включать в себя, полностью или частично, другие объекты, менять свойства (атрибуты) в зависимости от их местонахождения. Объекты могут взаимодействовать между собой, имея либо пространственно–логические (ПЛС), либо тополого–метрические связи (ТМС), число которых может значительно изменяться [2, 3].

Характеристики (атрибуты) объектов также имеют пространственную привязку, могут меняться в зависимости от местонахождения объекта, иметь сложноструктурированный характер, то есть состоять (включать в себя) большой набор уточняющих свойств.

В процессе первичного анализа документов, относящихся к СФГД, выявляются характерные особенности, присущие данному классу документов, а именно: СФГД типа машиностроительных чертежей, схем, планов поэтажной застройки зданий являются документами массового типа; как правило, данный класс документов исполнен на бумажных носителях среднего и плохого качества; на планах и схемах отсутствует жесткий стандарт и часто не соблюдены правила отрисовки отдельных графических знаков; в зависимости от качества носителя СФГД и правильности выбора параметров оцифровки, степень зашумленности растрового изображения графического документа (РИГД) может варьироваться в широких пределах; метрическая точность графических знаков также варьируется широком диапазоне значений.

Таким образом, основные особенности ПРД, влияющие на выбор моделей и способов описания объектов, можно сформулировать следующим образом [2, 3]:

- переменное сложно–структурированное метрическое описание информационных объектов;
- пространственная привязка атрибутивных данных, число которых может значительно изменяться;
- переменное число различных внутри и межобъектных связей с их пространственной локализацией;
- переменное число внутриобъектных и межобъектных топологических связей.

Определим основные графические объекты, встречающиеся на БФГД [2, 3]:

- *линейный* – это объект, характеризующийся некоторой протяженностью (линией) на местности и БФГД;
- *площадный* – это объект, площадь которого выражается в масштабе БФГД и ограничивается контуром (внешней границей). Для их обозначения соответствующие области заполняются определенным цветом, штриховкой или структурными элементами (точками, значками и т.п.). Очертания областей достаточно точно отражают очертания реальных объектов в масштабе БФГД;
- *дискретный* – это объект, имеющий "точечный" характер или площадь которого не выражается в масштабе БФГД. К этому классу относятся различные условные знаки и специальные символы. Совокупность дискретных объектов может определять характеристики объектов остальных классов.
- *без метрики* – это объект, не имеющий пространственной локализации, или имеющий неопределенную локализацию на местности и БФГД.

Каждый объект содержит определенный набор характеризующих его признаков. В общем случае все признаки могут быть разделены на две основные группы [2, 3]:

1. *метрическая информация* – содержит сведения о местоположении объекта на БФГД;
2. *семантическая информация* – содержит сведения о свойствах данного объекта.

Метрическая информация включает координаты точек объекта. Метрика дискретных объектов содержит от 1 до 3 точек, метрика площадных объектов – упорядоченную последовательность точек, описывающих внешний и внутренний контуры соответствующей области. Основное требование к метрической информации – точность. Это означает, что пространственные характеристики с требуемой точностью определяют положение объекта в системе координат и относительно других объектов.

К семантической информации относят непосредственно качественные и количественные характеристики объектов:

- временные характеристики, определяющие время исследования объекта и иногда показывающие зависимость изменения свойств объекта с течением времени;
- тематические характеристики, описывающие свойства объекта, не включенные в пространственные и временные. Это могут быть экономические, статистические, технические, организационные, управленческие и прочие виды данных. Основное требование к тематическим данным – полнота. Полнота тематических данных означает, что этих данных достаточно для решения практических задач и нет необходимости проводить дополнительный сбор данных.

Временные и особенно тематические данные фиксируют с помощью разных описательных характеристик, частично или полностью формализованных. Поэтому эти данных определяют в виде кодов, дескрипторов, словарей и т.п.

Все это усложняется еще и тем обстоятельством, что необходимо производить обработку ГИ в реальном масштабе времени и при ограниченных ресурсах памяти.

Указанные особенности рассматриваемого класса БФГД обуславливают повышенные требования к геометрическому моделированию отсканированных изображений такого рода графических документов.

На протяжении более 40 лет в НИИ ПМК ННГУ (ныне Центр информатики и интеллектуальных информационных технологий ИТММ ННГУ– ЦИИИТ ИТММ ННГУ) разрабатываются программно–аппаратные комплексы обработки БФГД. Необходимо отметить, что технология автоматического анализа графической информации (ГИ) – сложный многоэтапный процесс, включающий в себя большое количество методов и алгоритмов обработки – фильтрации, сжатия, хранения и поиска, анализа и принятия решений. Для эффективной работы этого «конвейера» необходимо, чтобы все математические модели, методы и алгоритмы, а также структуры представления данных, что очень важно, были взаимоувязаны и взаимоэффективны, так как очевидно, что сколь угодно высокая эффективность на каком-то отдельном участке обработки может быть сведена на нет на других этапах. Таким образом, модели и методы

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

обработки должны быть технологичными и удовлетворять некоторым общим требованиям, предопределенным эффективностью решения задач анализа ГИ в целом. Однако, в настоящее время большинство методов и математических моделей описания растровых изображений, позволяющих решать частную задачу, малоэффективны для работы в «конвейере» обработки ГИ. Это особенно важно на нижних уровнях иерархии. Проблемы еще более усугубляются в ГИС в связи с необходимостью осуществлять обработку очень большого объема растровой информации в реальном масштабе времени, при ограниченных ресурсах оперативной памяти и необходимости их естественной интеграции в геоинформационные технологии и системы. Таким образом, основные требования, предъявляемые к моделям описания, методам и алгоритмам обработки растровой информации ГД в ГИС следующие [2, 3]:

- технологичность;
- высокая емкостная и временная эффективность;
- естественная интегрируемость в общую схему ГИС.

Все это определяет актуальность проблемы развития новых математических моделей описания и методов обработки растровых изображений БФГД.

В работах нижегородских ученых Васина Ю.Г., Башкирова О.А., Чудиновича Б.М. и др. (НИИ ПМК ННГУ) [2, 3] был предложен комбинаторно-геометрический подход к обработке растровых изображений графической информации пространственно-распределенных данных (ПРД) в основе которого весь сложный комплекс задач, связанный с анализом видеоданных, рассматривается с единой точки зрения построения иерархии взаимосвязанных математических моделей описания, структур представления и принятия решений, нижний уровень которых обрабатывает растровую информацию со сканирующего устройства, а верхний уровень соответствует описанию ГД в терминах пользователя. Были предложены математические модели нижних уровней иерархии для описания изображений, в которых графические изображения представлялись в виде растра (набора пикселей), совокупности линий, контуров и точек (контурная, линейно-контурная модели) [3]. Линейно-контурная модель изображения ГД рассматривалась как базовая для всех последующих уровней моделей. На базе контурной и линейно-контурной моделей строились модели высших уровней. Для моделей высших уровней были развиты эффективные методы вычислительной геометрии (метод от общего к частному на базе иерархических структур представления векторных данных), методы распознавания графических объектов (корреляционно-экстремальный метод) и пр.

Однако, практика обработки БФГД, а особенно СФГД, особенно ГД, взятых из архивов, вызвало новые серьезные проблемы, связанные с переходом от нижнего (пиксельного) представления к векторному, что, в свою очередь, привело к необходимости дальнейшего развития иерархии моделей описания растровых изображений, особенно на нижних уровнях, где наиболее велики объемы обрабатываемых данных и временные затраты.

Низкоуровневая структурированная модель растровых данных

В рамках комбинаторно–геометрического подхода [2] под математической моделью изображения будем понимать тройку вида $M_V^\alpha = \{E_V^\alpha, C_V^\alpha, R_V^\alpha\}$,

где:

$E_V^\alpha = \{e_1^\alpha, e_2^\alpha, ..., e_S^\alpha\}$ – множество непроизводных элементов модели ранга α;

$R_V^\alpha = \{r_1^\alpha, r_2^\alpha, ..., r_t^\alpha\}$ – множество допустимых отношений между непроизводными элементами модели ранга α;

$C_V^\alpha = \{c_1^\alpha, c_2^\alpha, ..., c_n^\alpha\}$ – множество характеристик непроизводных элементов модели ранга $\alpha=1,2,3,...N$.

Исходя из этого, предлагаются *класс растровых* и *класс векторных моделей* БФГД [1].

Класс растровых моделей (v=1):

- $M_1^1 = \{E_1^1, C_1^1, R_1^1\}$ – пиксельная модель (ПМ);
- $M_1^2 = \{E_1^2, C_1^2, R_1^2\}$ – штриховая модель на базе «линейных» штрихов (ШМ);
- $M_1^3 = \{E_1^3, C_1^3, R_1^3\}$ – модель простых растровых объектов (ПРО);
- $M_1^4 = \{E_1^4, C_1^4, R_1^4\}$ – модель растровых составных объектов (РСО);
- $M_1^5 = \{E_1^5, C_1^5, R_1^5\}$ – линейно-площадная растровая модель (ЛПРМ).

ПМ – это общепринятая в цифровой обработке изображений пиксельная прямоугольная матрица $[P_{ij}]$ $(i= 1, 2,..., N; j= 1, 2,..., M)$, элемент p - пиксель, который является кодом цвета точки изображения с координатой $X_i=i* hX$ и $Y_i= j*hY$, где hX и hY - шаг пространственной дискретизации изображения соответственно по координатам X и Y.

В общем случае пиксель, как элементарный непроизводный элемент РИГД, можно описать четверкой типа $p=\{x,y,c,sw\}$, где:

- x, y -растровые координаты пикселя(номер строки и номер столбца пиксельной матрицы);
- c – структурная переменная, описывающая цветовое значение пикселя. В зависимости от цветовой гаммы исходного изображения эта переменная принимает следующие значения:
  - для бинарных растровых изображений (БРИ) это бинарная переменная, принимающая два фиксированных значения $c =\{0,1\}$, при этом $c=0$ (черное) соответствует не фоновому пикселю изображения, а $c=1$ (белое) – фоновому пикселю;
  - для серых полутоновых изображений это дискретная переменная, заданная на некотором интервале значений $c =\{0,1,... R\}$;
  - для цветных изображений используется *RGB* формат описания цвета пикселя и для него структурная переменная с имеет вид трехкомпонентного вектора, то есть $c=\{r, g, b\}$, где каждая компонента принимает значение из заданного интервала.
- sw – коэффициент связи, измеренный в четырех– или восьмиэлементной окрестности пикселя P и описывающий его топологические свойства (рис. 1).

2	3	4
1	P	5
8	7	6

а

	2	
1	P	3
	4	

б

Рисунок 1 Восьми– (а) и четырехсвязная (б) области пикселей и условная их нумерация в зоне единичной окрестности пикселя P

Например, при выборе четырехсвязной области $sw=\{0-4\}$. Причем, если

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

$$sw = \begin{cases} 0, & P - изолированный\ пиксель, не\ имеющий\ соседей\ своего\ цвета; \\ 1, & P - граничный\ пиксель\ с\ одним\ соседом\ своего\ цвета; \\ 2, & P - граничный\ пиксель\ с\ двумя\ соседями\ своего\ цвета; \\ 3, & P - граничный\ пиксель\ с\ тремя\ соседями\ своего\ цвета; \\ 4, & P - внутренний\ пиксель, неимеющий\ соседей\ не\ своего\ цвета. \end{cases}$$

На базе понятия пикселя и его топологических свойств можно построить структурные компоненты РИГД: штрих, ПРО, РСО, являющиеся результатами одномерной, либо двухмерной кластеризации связных пикселей [1].

Штрих – одномерный кластер связных пикселей заданного цвета, представляет набор чисел *H=(S, N, K ,C)*, где: *S* – текущий номер линии сканирования РИГД; *N, K* – координаты начала и конца совокупности подряд идущих пикселей одного цвета вдоль линии растра; *C* – код цвета связанных пикселей, образующих штрих. Для бинарных изображений кодируются только «объектовые» связные области черного цвета, поэтому штрих описывается тройкой типа *H=(t, N, K)*, где: *t* - номер строки растра; *N, K* – координаты начала и конца штриха. Совокупность штрихов для заданной строки растра *t* определяет строку штрихов *SH={Hi}(i=1,2,...m)*, где H_i – текущий штрих в строке. Описание растра с помощью набора строк штрихов является однозначным и задает векторно-растровый, или штриховой, формат представления пиксельных матриц. Необходимо отметить, что штриховой формат представления и хранения РИГД является в *L* раз более компактным, чем пиксельный, где *L* – средняя длина штриха для выбранного направления линии сканирования данного растра. Опыт применения штрихового формата для описания БФГД типа топографических карт показал, что *L=10-20* и зависит от насыщенности растра разноцветными компонентами. Кроме того, многие алгоритмы локальной фильтрации, сглаживания, форматирования и векторизации также в среднем в *L* раз эффективнее по времени для штрихового формата по сравнению с пиксельным.

Топологические свойства штриха *(t, N, K)* определяются через его связность со штрихами смежных строк. Формально можно дать следующее определение меры связности штриха *(t, N, K).* Пусть sw_p, sw_s определяют топологические свойства штриха, то есть его связность со штрихами смежных строк, а *t* является номером текущей строки штрихов растра. Тогда, если в предыдущей строке *p=t-1* нет ни одного штриха, для которого выполняется условие

$$N_p \geq N_t \wedge N_p \leq K_t \vee K_p \geq N_t \wedge K_p \leq K_t \quad (1)$$

то $sw_p=0$; иначе, sw_p равно числу раз выполнения условия (1), где sw_p – количество штрихов строки *p*, связных с текущим штрихом строки *t*. Если в следующей строке *s=t+1* нет ни одного штриха для которого выполняется условие

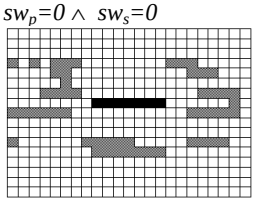
$$N_s \geq N_t \wedge Ns \leq Kt \vee K_s \geq N_t \wedge K_s \leq K_t \quad (2)$$

то $sw_s=0$, иначе sw_s равно числу раз выполнения условия (2), где sw_s - количество штрихов строки *s*, связных с текущим штрихом строки *t*.

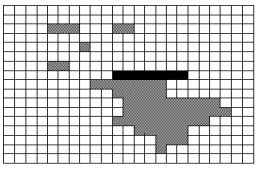
Здесь N_t, K_t - координаты начала и конца исследуемого на связность штриха текущей строки *t*, N_p, K_p – координаты начала и конца очередного штриха предыдущей смежной строки *p*, N_s, K_s - координаты начала и конца текущего штриха последующей смежной строки *s*.

По суперпозиции значений sw_p и sw_s можно ввести следующую классификацию графических ситуаций (Рис.2).

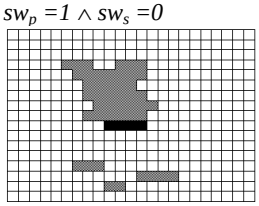
Штрих изолированный (Ши): $sw_p=0 \wedge sw_s=0$



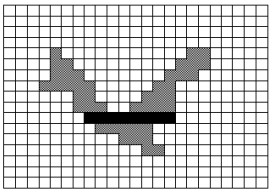
Штрих начала растрового объекта (Шн): $sw_p=0 \wedge sw_s=1$



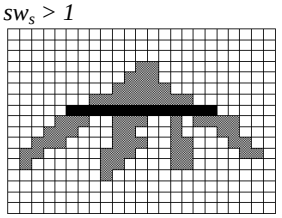
Штрих конца растрового объекта (Шк): $sw_p=1 \wedge sw_s=0$



Штрих слияния растровых объектов (Шс): $sw_p > 1$



Штрих расщепления растровых объектов (Шр): $sw_s > 1$



Штрих слияния и расщепления растровых объектов (Шср): $sw_p > 1 \wedge sw_s > 1$.

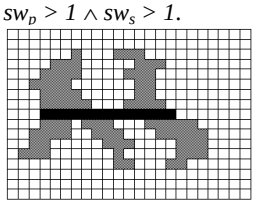


Рисунок 2 Классификация графических ситуаций положения штриха в текущей строке

На базе классификации положения штриха в строке определим структурные компоненты РИГД [1]:

- *растровый простой объект РПО* – кластер связанных штрихов, не содержащий графических ситуаций типа слияния и расщепления, то есть *РПО* – это такой двумерный (плоский) кластер связанных штрихов, для любого штриха которого выполняется условие

$$sw_p < 2 \wedge sw_s < 2 \quad (3)$$

- *растровый составной объект РСО* – кластер связанных штрихов, для каждого штриха которого выполняется условие

$$sw_p > 0 \vee sw_s > 0 \quad (4)$$

и для любых двух элементов (штрихов или пикселей) найдется хотя бы одна соединяющая их пиксельная траектория, состоящая из связных смежных элементов. Отметим, что построение структурных растровых компонент типа РСО можно производить на основе и пиксельного и штрихового описания, а выделение РПО можно осуществлять только по совокупности штрихов растра.

Линейно-площадная модель РИГД – это представление исходного РИГД как логической суммы двух РИГД, один из которых состоит из пикселей *растровых линейных* объектов (РЛО), а второй - из пикселей *растровых площадных* объектов (РПО). Такое расщепление исходного РИГД на две компоненты возможно при задании каждому пикселю исходного РИГД свойства протяженной связности L_{sw}. Определим L_{sw} для пикселя как длину минимального штриха, полученного по всем возможным направлениям линии сканирования, проходящей через этот пиксель, то есть $L_{sw} = \min_u(L_u)$, $u=1,2,...k$, где u – направление линии сканирования, L_u – длина штриха, которому принадлежит исследуемый пиксель. Сравнивая полученную величину с пороговой пиксельной толщиной линейных растровых объектов,

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

можно осуществить бинарную классификацию на два класса: пиксель, принадлежащий РЛО, или пиксель, принадлежащий РПО, простым пороговым решающим правилом: если $L_{sw} < Por$, то пиксель принадлежит РЛО, иначе, пиксель принадлежит РПО.

На рис. 3 представлен кластер связных пикселей, состоящий из одного РПО и двух РЛО. Пиксельная толщина растровой линии выбрана равной 2 пикселям.

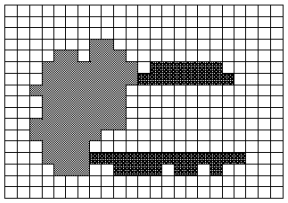


Рисунок 3 кластер связных пикселей, состоящий из одного РПО и двух РЛО

Предложенная иерархия моделей описания РИГД на нижних уровнях позволила перейти к структурированному представлению (описанию) растровых изображений БФГД (Рис.4). При этом растровое изображение представляется совокупностью *связных растровых компонент* (СРК): РПО, РСО, а также узловых штрихов и штрихов начала и конца объектов.

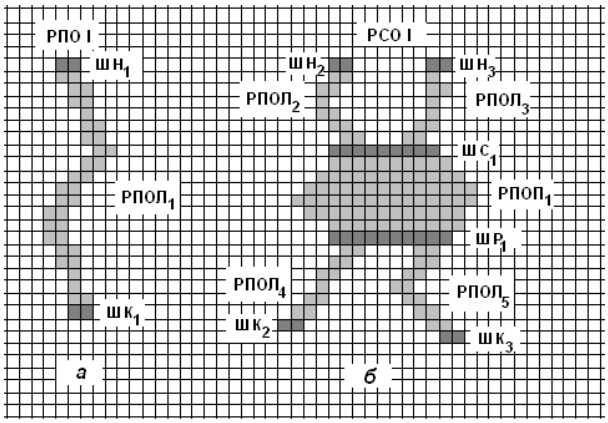


Рисунок 4 структурированное описание растровых изображений БФГД

РПО разделяются на *линейные* (РПОЛ) и *площадные* (РПОП). Множество непроизводных элементов (НЭ) состоит из РПОЛ, РПОП, и штрихов (Шн, Шк, Шс, Шр, Шср). РПО примыкают к штрихам. На рис. 4 изображены две СРК: на рис. 4а – $РПОЛ_1$, а на рис. 4б – $РСО_1$, состоящий из $РПОЛ_2$, $РПОЛ_3$, $РПОЛ_4$, $РПОЛ_5$ и $РПОП_1$, двух штрихов начала - $Шн_2$, $Шн_3$, двух штрихов конца - $Шк_2$, $Шк_3$, одного штриха слияния $Шс_1$, одного штриха расщепления $Шр_1$. $Шс_1$ – имеет три связи с $РПОЛ_2$, $РПОЛ_3$, $РПОП_1$, $Шр_1$ – имеет три связи с $РПОП_1$, $РПОЛ_4$, $РПОЛ_5$

Для РСО и РПО, а также штрихов начала и конца объектов и узловых штрихов, вычисляются характерные признаки следующих типов:

- номер РСО и РПО;
- количество РПО и штрихов (Шн, Шк, Шс, Шр, Шср) в РСО;
- координаты описывающего прямоугольника;
- площадь и периметр этого прямоугольника;
- площадь РПО и РСО;
- отношение площади описывающего прямоугольника к площади РПО и РСО;
- номер штриха;
- длина штриха;
- количество связей и др.

Построенная иерархия растровых моделей структурированного описания изображений БФГД позволяет осуществлять:

- структурирование РИГД;
- распараллеливание обработки растровых объектов;
- расширение класса используемых методов обработки РИГД на нижних уровнях описания;
- использование не только локальных, но и интегральных критериев обработки;
- распознавание линейных, площадных и дискретных растровых объектов;
- сокращение емкостной и вычислительной сложности алгоритмов обработки РИГД.

Структурная модель векторных данных

В качестве дальнейшего развития представленной выше иерархии растровых моделей, перейдем к усовершенствованному классу моделей векторного уровня. При этом, в добавление к имеющимся контурной и линейно-контурной моделям, предлагается линейная и, получаемая на ее основе, сегментно-узловая модель. В результате *класс векторных моделей (v=2)* включает [1,3]:

- $M_2^1 = \left\{E_2^1, C_2^1, R_2^1\right\}$ – контурная модель изображения (КМ);
- $M_2^2 = \left\{E_2^2, C_2^2, R_2^2\right\}$ – линейно-контурная модель изображения (ЛКМ);
- $M_2^3 = \left\{E_2^3, C_2^3, R_2^3\right\}$ – линейная модель изображения (ЛМ);
- $M_2^4 = \left\{E_2^4, C_2^4, R_2^4\right\}$ – сегментно – узловая модель изображения (СУМ).

Контурная модель изображения. Пусть $K = \{x_1, y_1, x_2, y_2, \ldots, x_n, y_n\}$ – последовательность точек контура связной фигуры. Тогда непроизводный элемент контурной модели – i-ый вектор метрики контура: $\left\{x_i^н, y_i^н, x_i^к, y_i^к\right\}$, отношение примыкания векторов контура:

$$\forall_t \left\{ \left(x_t^к = x_{t+1}^н \right) \wedge \left(y_t^к = y_{t+1}^н \right) \wedge \left(x_t^н = x_{t-1}^к \right) \wedge \left(y_t^н = y_{t-1}^к \right) \right\}$$

вектор всегда примыкает к единственному последующему вектору и к единственному предыдущему, так что «черное» всегда остается справа; характеристика вектора: вектор может быть внутренним и внешним.

Последовательность векторов контура образует контур (внешний или внутренний), т.е. объект изображения – это внешний контур и множество внутренних контуров связной фигуры. Совокупность контуров для всего растра определяет контурную модель растровых данных.

С точки зрения растрового изображения, контур – это линия границы растровой области заданного цвета, то есть слева и справа от линии контура расположены пиксели разных цветов. Направление обхода контура показывает цвет раскраски его внутренней области. Например, для бинарных изображений, если направление обхода контура по часовой стрелке, то внутренняя область контура (справа по ходу движения) – черная, иначе – белая. По этому признаку все контура разделяются на внешние (обход по часовой стрелке) и внутренние (обход против часовой стрелки).

Контурная модель находится во взаимно – однозначном соответствии с РИГД. Имеет самостоятельное значение, но может рассматриваться и как подготовительный этап формирования линейно–контурной модели.

Линейно-контурная модель изображения. Множество непроизводных элементов: вектор, образованный точками метрики контура (вектор контура), вектор, образованный точками метрики линии (вектор

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

линии), черная точка на белом фоне, белая точка на черном фоне. Набор отношений: вектор контура всегда примыкает к предыдущему и последующему векторам контура и может примыкать к вектору линии, вектор линии может примыкать к последующему и предыдущему векторам линии и вектору контура или иметь один, или два свободных конца, точки не примыкают ни к каким объектам – изолированы. Набор характеристик: вектор линии имеет толщину, вектор контура может быть внешним или внутренним, точка имеет диаметр. Объекты изображения: линия – последовательность векторов линии; контур – последовательность векторов контура, черная и белая точки. Линии и контуры не пересекаются, контуры замкнуты. Линейно-контурная модель описывает границы площадных и линейных графических объектов, и поэтому имеет, в отличие от контурной модели, более сложную топологическую структуру.

Линейная модель изображения – может рассматриваться как вырожденный случай линейно–контурной модели. Множество непроизводных элементов: вектор линии, Набор отношений: вектор линии может примыкать к последующему и предыдущему векторам линии или иметь один или два свободных конца. Набор характеристик: вектор линии имеет толщину. Объекты изображения: линия – последовательность векторов линии. Линии могут пересекаться. Наиболее эффективно ее использование при обработке БФГД с явно выраженной линейной структурой обрабатываемых объектов. С другой стороны, линейные объекты появляются в задачах геометрического моделирования РИГД при аппроксимационной замене протяженных СРК на линейные объекты. Построение этой модели возможно как непосредственно по РИГД, так и по контурной модели.

Сегментно-узловая модель изображения – является векторной моделью, целиком описывающей все изображение в виде набора связных множеств. Непроизводные элементы: вектора – сегменты и точки–узлы, описываемые координатами $\{x,y\}$ и параметром Sw_L, указывающим на L-сегмент, связанный с этим узлом. Наборы отношений и характеристик совпадают с аналогичными наборами линейной модели.

Контурная, линейно- контурная и линейная модели изображений связаны с теми или иными алгоритмами векторизации СРК и обеспечивают геометрическую интерпретацию изображений в задачах анализа сцен и распознавания, а также метрическое описание информационных компонент растра.

Для представления созданных моделей разработаны оригинальные структуры и форматы хранения.

Построенная иерархия растровых моделей структурированного описания изображений ГД на растровом уровне, позволяет осуществлять:

- структурирование растровых изображений;
- распараллеливание обработки растровых объектов;
- расширение класса используемых методов обработки растровых изображений на нижних уровнях описания;
- использование не только локальных, но и интегральных критериев обработки;
- распознавание линейных, площадных и дискретных растровых объектов;
- сокращение емкостной и вычислительной сложности алгоритмов обработки растровых изображений ГД.

Алгоритмы построения иерархии моделей штрихового и векторного уровней описания РИГД

Алгоритмы построения иерархии моделей описания РИГД основаны на вычислении признаков связности пикселей в строке для формирования штриховой модели, либо признака связности штрихов в двух смежных с ним строках (предыдущей и последующей) – для формирования РПО. В результате формируется массив штрихов $H_i(x_н, x_к, y)$, или связных штрихов $H_i(x^i_н, x^i_к, sw_p, sw_s)$. По массиву связных штрихов строится таблица параметров РПО (ТПРПО), в которой хранятся адреса начала размещения РПО, количество штрихов в РПО, площадь РПО, координаты и площадь описывающего прямоугольника и др. Используя данные ТПРПО, строится модель растрового составного объекта (РСО), который представляется совокупностью связанных РПО. Важной отличительной особенностью этих алгоритмов является их локальность. На каждом шаге алгоритмов производится обработка полосы исходных данных, состоящей не более чем из трех строк. При этом выявляются все возникающие значимые ситуации и осуществляется их разрешение. Другая важная особенность состоит в параллельности. Построение моделей осуществляется за один просмотр входной информации и состоит из двух основных операций: обработки информации и отслеживании размеченных связных штрихов [4 – 7].

Для СФГД типа чертежей, схем и планов характерно изображение сюжетов рисунка с помощью отрезков линий с четко выраженным направлением их ориентации. Если при сканировании этих документов направление сканирования выбрать в соответствии с преимущественной ориентацией этих отрезков, то штриховая модель отсканированных данных позволяет построить простые и надежные алгоритмы векторизации и осуществить эффективное геометрическое моделирование растрового описания этих документов.

Векторизация является базовой операцией в большинстве систем обработки и анализа графической информации. Если растровое описание векторизуемых объектов не содержит искажений и шумов то существующие локальные алгоритмы векторизации вполне удовлетворительно справляются с этой задачей, хотя надо иметь ввиду, что получаемые при этом векторизованные объекты требуют дополнительного сглаживания либо аппроксимации, а это не всегда удается сделать успешно, то есть одновременно vдовлетворить требованиям метрической точности аппроксимации и геометрической достоверности векторизованного объекта.

Линейная модель растрового изображения строится только по штрихам, принадлежащим линейным РПО. Изначально выполняется этап предварительной обработки РИГД, включающий следующие операции над исходным растром [5, 6]:

1) определение преимущественно выраженного направления ориентации РЛО, как направление наиболее протяженных РЛО на данном РИГД;
2) поворот исходного РИГД, таким образом, чтобы новое направление сканирования графического документа совпало или было близко к найденной на шаге 1 ориентации РЛО;
3) конвертация исходного РИГД в штриховую модель.
4) разбиение полученного множества штрихов на набор РСО.
5) формирование линейно-площадной модели РИГД в виде двух РИГД, содержащих наборы РЛО и РПО исходного РИГД. Для этого исходный набор РСО сегментируем на кластеры связных коротких и длинных штрихов по пороговому значению $L_p = k * dLine$, где

d_Line – средняя толщина растровой линии, k = {1 – горизонтальные, 2 – наклонные, 3 – вертикальные}– коэффициент наклона растровых линий на документе. Далее выполняем слияние кластеров заданного типа методом ε-кластеризации, то есть формируем РСО с заданными свойствами и полученные РСО считаем исходными растровыми объектами для тех или иных алгоритмов векторизации.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Как было отмечено, для РИГД типа поэтажных планов, схем, технических чертежей характерно наличие большого количества РЛО. Поэтому далее остановимся подробно на алгоритмах векторизации именно этого вида растровых объектов, рассмотрев два алгоритма векторизации РЛО – глобальный и локальный [5 – 7].

В глобальном алгоритме очередной исходный РЛО рассматривается как некоторое множество точек плоскости $P = \{x_i, y_i\}$, $(i=1, 2, ..., N)$, где каждая точка $\{x_i, y_i\}$ является векторизованным описанием текущего пикселя обрабатываемого РЛО. Для множества P методом наименьших квадратов (МНК) строится кусочно – линейная аппроксимирующая с заданным ε функция, которая и является выходным результатом векторизации текущего РЛО.

В локальном алгоритме производится последовательное отслеживание узлов линии методом самоцентрируемого скользящего окна с использованием следующих алгоритмических операций:

1) центр растрового окна заданного размера $H_w = d_Line$ помещается в текущий свободный пиксель РЛО;

2) для координат n пикселей, попавших в окно, вычисляется среднее значение и их линейная ориентация

$\tilde{x} = \frac{1}{n}\sum_{i=1}^{n} x_i; \tilde{y} = \frac{1}{n}\sum_{i=1}^{n} y_i; U_x, U_y$, где U_x, U_y – компоненты первого собственного вектора, полученного по МНК, и линейно аппроксимирующего множество пикселей окна;

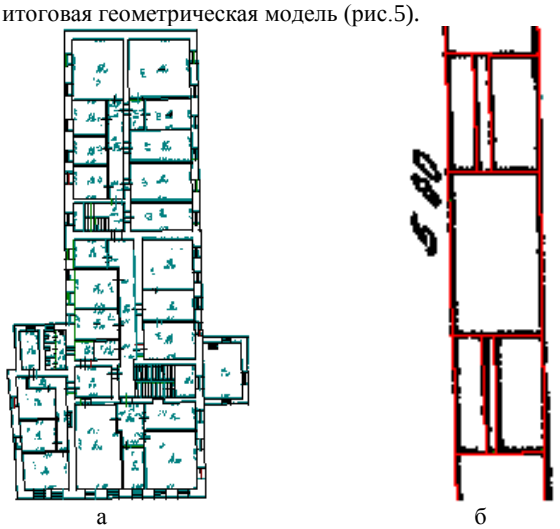
3) среднее значение принимаем за текущий узел векторизуемого РЛО;

4) перемещаем окно по вектору линейной ориентации до появления первого свободного пикселя;

5) шаги 2 – 4 выполняем до тех пор, пока в перемещаемом окне будет хотя бы один свободный пиксель, в противном случае считаем, векторизацию текущего РЛО законченной

Для множества отрезков, полученных из текущего РСО, выполняется операция метрического и геометрического согласования метрики отрезков.

По согласованной метрике отрезков формируется итоговая геометрическая модель (рис.5).



а б

Рисунок 5 а) итоговая геометрическая модель; б) ее увеличенный фрагмент

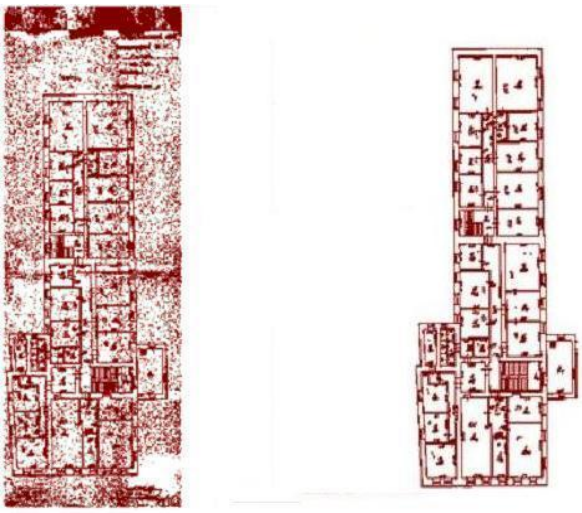
При построении контурной и линейно−контурной моделей за счет использования информации о РПО и РСО удалось сократить емкостную сложность алгоритмов в 4–5 раз (отпадает необходимость хранить всю таблицу связей координат точек контура или линии), а также временную сложность в 2–3 раза.

Построение сегментно−узловой модели описания осуществляется в два этапа: сначала, по сформированным признакам близости отрезков и углов направления близко расположенных отрезков линейной модели, осуществляется сборка длинных линий, затем, по заданным критериям проводится поиск узлов связи построенных линий (точки пересечения и примыкания линий). Алгоритмы построения контурной и линейно−контурной моделей описания осуществляют отслеживание внешних и внутренних контуров площадных РПО и РСО, а также линий для линейных РПО, и точек связи линий и контуров [6].

На базе изложенных алгоритмов создано оригинальное программное обеспечение, используемое при автоматическом цифровании схем поэтажной планировки в задачах создания электронных кадастров.

Объектовая фильтрация растра

Для минимизации побочного эффекта масочных фильтров, выражающегося в уменьшении резкости границ объектов, был разработан алгоритм объектовой фильтрации БРИ в штриховой форме суть которого в следующем. Входными параметрами являются код цвета удаляемой помехи и размеры описанного вокруг помехи прямоугольника со сторонами, параллельными осям координат, а также процент его заполненности пикселями цвета помехи. Работа алгоритма основана на выделении по исходному БРИ набора РСО, состоящих из пикселей с цветом, равным заданному, с последующим удалением (перекрашиванием в альтернативный цвет) тех сформированных РСО, которые по своим геометрическим характеристикам удовлетворяют параметрам помехи. Важным достоинством данного алгоритма является тот факт, что указанная фильтрация не вносит никаких искажений в те связные области, которые не распознаются алгоритмом как шумовые. На рис 6а представлены исходный растр поэтажного плана здания в штриховом формате, а результат его объектовой фильтрации на рис. 6b.



а б

Рисунок 6 а) исходное растровое изображение поэтажного плана здания в штриховом формате, б) результат его фильтрации объектовым фильтром

Автоматическое распознавание условных знаков конструктивных элементов поэтажного плана здания

Ключевым моментом повышения эффективности преобразования различной бумажной конструкторской документации, в частности, планов поэтажной застройки зданий в цифровой вид, является автоматическое распознавание условных знаков конструктивных элементов поэтажного плана здания. Для решения данной

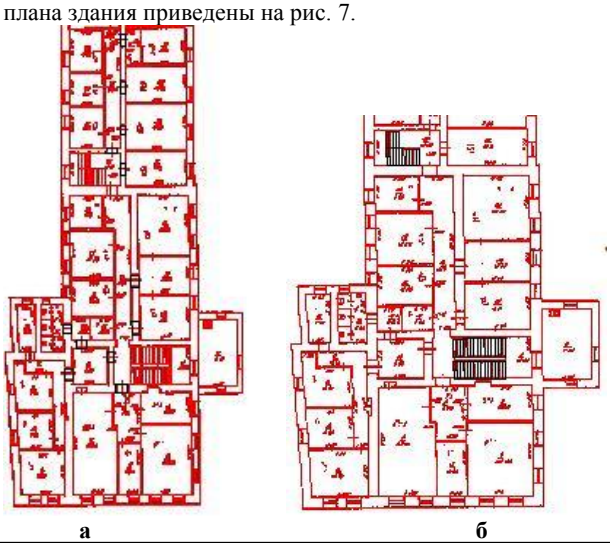
Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

задачи был разработан алгоритм, состоящий из следующих шагов [8]:

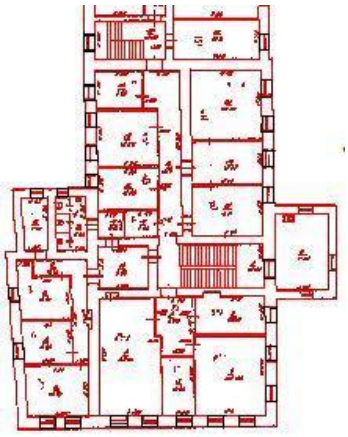
- используя линейную модель исходного документа, формируется массив точек пересения линейных объектов (узлов) и набор соединяющих их отрезков–сегментов;
- производится автоматическая разметка сегментов по связности их концов между собой;
- осуществляется кластеризация и разметка сегментов, являющихся элементами условного знака *Лестница*;
- выполняется разметка линий со свободными концами, являющихся элементами условного знака *Дверь*;
- по контурной модели документа производится предварительная разметка контуров, принадлежащих *Помещениям*, по наличию внутри них различных текстовых символов;
- осуществляется кластеризация сегментов для формирования фрагментов, относящихся к условному знаку *Капитальная стена*;
- выполняется поиск внутренних элементов объекта *Капитальная стена* (дверей, окон, стенных шкафов, ниш);
- по набору сформированных, подряд стоящих сонаправленных фрагментов условного знака *Капитальная стена*, методами кластерного анализа собираются полноценные элементы, относящиеся к данному условному знаку;
- подвергаются кластеризации оставшиеся неразмеченные сегменты и выполняется построение элементов, относящихся условному знаку *Перегородочная стена*;
- выполняется поиск элементов, относящихся к условному знаку *Дверь внутри капитальной прегородки*;
- выполняется согласование границ условных знаков распознанных стен между собой и другими объектами;
- производится автоматическое распознавание условных знаков *Комната* с учетом уже имеющихся границ условных знаков стен;
- осуществляется автоматическое распознавание условных знаков *Лестница* и построение объектов *Лестничный марш* и *Лестничная площадка*;
- все оставшиеся нераспознанные контура получают выходной классификационный тип *Контур* для их дальнейшей интерактивной идентификации.

Поскольку один и тот же сегмент может одновременно входить в метрику сразу нескольких условных знаков, то решение задачи установления тополого–метрических связей между различными объектами поэтажных планов существенно упрощается и в некоторых случаях решается уже на этапе распознавания.

Результаты автоматического распознавания некоторых условных знаков конструктивных элементов поэтажного плана здания приведены на рис. 7.



а **б**

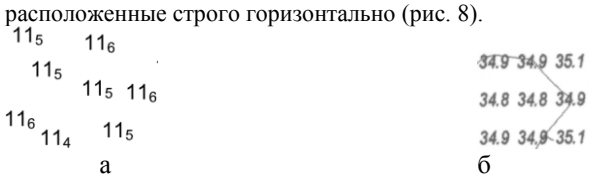


в

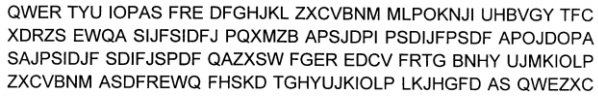
Рисунок 7 Результаты автоматического распознавания отдельных условных знаков конструктивных элементов поэтажного плана здания: а) дверных проемов; б) лестниц, лестничных маршей и площадок; в) окон

Автоматическое распознавание символов на СФГД

В качестве исходных документов были взяты 24–битные растровые изображения текстов на английском языке, а также планшеты гидрографической съемки, расположенные строго горизонтально (рис. 8).



а б



в

Рисунок 8 Примеры исходных изображений: а) планшет гидрографической съемки хорошего качества; б) планшет гидрографической съемки среднего качества; в) отсканированный текст

Особенностью СФГД является то, что, несмотря на отклонение от нормативных требований к изображению объектов, они обладают некоторой стилизованной формой представления. Это означает, что возможно создание набора эталонов объектов, основанных на работе с низкоуровневыми моделями графических изображений, максимально адаптированных для данного класса документов, с учетом сильной зависимости эффективности признаков классификации от искажений входных данных. В дальнейшем, по мере поступления новых документов, этот набор может соответствующим образом дополняться.

Ранее, в практику обработки ГИ были введены моментные характеристики, нашедшие широкое применение во многих задачах обработки цифровых изображений [9].

Момент – характеристика контура изображения, объединённая (суммированная) со всеми пикселями контура.

Момент $m_{p,q}$ для непрерывного изображения *f(x,y)* определяется как: $m_{p,q} = \iint x^p y^q f(x,y)dxdy$, в случае дискретного растрового изображения: $m_{p,q} = \sum\sum x^p y^q f(x,y)$, где p,q – мощность, на которой соответствующий компонент взят в сумме с

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

другими отображенными. Если функция *f(x,y)* рассматривается как бинарное изображение, тогда момент m_{00} вычисляет площадь чёрных пикселей этого изображения.

Для вычисления характеристик, инвариантных относительно сдвига, необходимо центрировать относительно средних значений $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$ – центр тяжести фигуры.

Центральный момент определяется как: $\mu_{p,q} = \sum\sum (x-\bar{x})^p (y-\bar{y})^q f(x,y)$.

Центральные моменты выражаются через обычные, по формуле:

$$\mu_{p,q} = \sum_{i=0}^{p} \sum_{j=0}^{q} C_p^i \, C_s^j \, (-\bar{x})^{(p-i)} (-\bar{y})^{(q-j)} m_{ij},$$

где $C_n^k = \frac{n!}{k!(n-k)!}$.

С помощью центральных моментов определяются характеристики, инвариантные к повороту изображения (моментные инварианты). *Моментные инварианты* – это признаки изображения, построенные на основе степенных моментов и описывающие силуэт некоторого объекта, являются инвариантными к аффинным преобразованиям изображения, не зависят от ориентации, угла обзора и других изменений, также они нечувствительны к различным геометрическим преобразованиям (сдвиг, поворот, масштабирование). Это определяет их эффективность при использовании в качестве признаков в задаче обнаружения и распознавания на изображении объектов неизвестной ориентации [9, 10].

Имея набор характеристик, можно определить следующие аффинные преобразования моментных инвариантов, которые инвариантны относительно сдвигов, поворотов, а также растяжений и сжатий: $I^1 = \frac{\mu_{20} + \mu_{02}}{m_{00}^2}$,

$$I^2 = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{m_{00}^4},$$

$$I^3 = \frac{(\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{30})^2}{m_{00}^5}$$

$$I^4 = \frac{(\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2}{m_{00}^5},$$ которые используются в качестве признаков классификации.

*Признаки классификации***.** Для каждого объекта будем фиксировать следующие признаки [11 – 13]:

- количество внутренних областей (ВО);
- отношение сторон описанного вокруг символа прямоугольника со сторонами, параллельными осям координат (отношение аспекта);
- отношение площади символа к площади описанного вокруг символа прямоугольника со сторонами, параллельными осям координат (процент заполненности);
- количество РПО. Также для каждого РПО вычислялись:
  - количество штрихов, входящих в РПО;
  - средняя длина штриха;
- значения моментных инвариантов $I^1 – I^4$.
- количество особых штрихов: Шн, Шк, Шс, Шр;
- координаты центра тяжести РСО, формирующего символ.

Рассмотренный набор признаков не является окончательным и может быть расширен [11 – 13].

Число внутренних областей достаточно эффективный признак, позволяющий сразу разделить все объекты на 3 группы: у которых: ВО = 0 (К, Х, S, 1, 2 и др.), ВО = 1 (R,O,P, 4,6,9 и др.) и ВО = 2 (В,8). Учтем, что «В» – единственный среди символов английского алфавита, а цифра 8 – единственная цифра среди арабских, имеющие ВО=2. Дальнейшее распознавание объекта можно выполнять в пределах этих групп. Отношение аспекта и процент заполненности более устойчивы к различным искажениям распознаваемых данных. Количество РПО, а также их параметры, достаточно устойчиво характеризуют каждый символ. Например, символ “А” состоит из 6 РПО, символ “К” – из 5 РПО.

Алгоритм автоматического распознавания символов

На вход алгоритма поступает вектор признаков связного объекта. На выходе распознанный символ алфавита базы эталонов или пустой символ. В алгоритме введено понятие штрафа – положительная величина, пропорциональная отношению расстояния до ближайшей границы интервала к длине этого интервала. Для признаков, сформированных по штриховому описанию: если значение признака меньше значения нижней границы интервала, то сразу начисляется штраф, больший порогового значения и класс убирается из дальнейшего рассмотрения [11 – 13].

Изначально формируем массив размером по количеству классов. В цикле по номеру признака:

- для каждого класса в соответствующую ему ячейку массива добавляется штраф, если значение признака не входит в допустимый интервал для данного класса;
- если штраф объекту для класса превосходит порог, то класс исключается из дальнейшего рассмотрения;
- если в рассмотрении остался один единственный класс, то соответствующий ему символ возвращается как результат распознавания;
- если были исключены все классы, то в качестве результата работы алгоритма возвращается пустой символ.

Если после завершения цикла осталось несколько классов, штраф которых не превзошел порог, то в качестве результата возвращается символ, соответствующий классу с минимальным штрафом.

Практические эксперименты по распознаванию текста на английском и русском языках обнаружили качество распознавания не ниже 98%, для планшетов гидрографической съемки качество оказалось чуть ниже, порядка 93–95%.

Снижение качества распознавания на данном типе документов объясняется наличием «рассыпанных» объектов–цифр и наложением на объекты–цифры тонких линейных объектов. Интерактивной или автоматической коррекции таких искажений в рамках данного эксперимента не проводилось.

В целом, можно утверждать, что значения моментных инвариантов достаточно устойчивы, при этом, значений I^1 и I^2 достаточно, чтобы выделить среди всех классов группы схожих, а значения I^3 и I^4 необходимы для более тонкой классификации внутри группы. На практике, после использования моментных признаков остаётся 2 – 5 возможных классов, к которым может принадлежать распознаваемый объект. Для окончательного решения об отнесении объекта к определённому классу используются признаки, полученные из штрихового описания. При этом учитывается экспериментально установленный факт, что отдельные признаки, полученные по штриховому описанию, обладают достаточной устойчивостью. Например, у символа “М” никогда не может быть менее трех Шк, двух Шн и Шр, одного Шс и четырех РПО.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Данное утверждение справедливо для любого из распознаваемых нами объектов, при условии, что они представляются связными растровыми объектами. Указанная особенность позволяет еще более повышать временную эффективность процедуры распознавания.

Представленные модели, алгоритмы и созданное на их основе СПО, успешно внедрены в технологии и программно–аппаратные комплексы, разрабатываемые в ЦИИИТ ИТММ ННГУ для картографических производств ОАО «Роскартография» и ВТУ МО РФ.

Обработка растровых данных ДЗЗ

В последние годы сотрудники ЦИИИТ ИТММ ННГУ принимают активное участие в разработке автоматизированной системы для дешифрирования (распознавания) объектов по растровым данных ДЗЗ, включая новейший и один из перспективных типов таких данных – гиперспектральные изображения (ГСИ ДЗЗ). Одной из основных причин использования ГСИ является максимально полное описание пространственного распределения спектрального состава регистрируемого излучения. В отличие от полутоновых, цветных и мультиспектральных изображений, ГСИ характеризуются более широким спектральным диапазоном, и высоким спектральным разрешением по шкале длин волн 5–10нм. Гиперспектральный сенсор одновременно формирует изображения на всех участках спектрального диапазона и получает спектры всех пикселей сцены. Для такого сенсора важно не количество измеряемых спектров, а узость и последовательность измерений [14]. Таким образом, ГСИ ДЗЗ обладают достаточно большой информационной емкостью, поэтому актуальна и задача их компрессии без потерь, желательно в режиме реального времени. Представляется очевидным, что использование предложенной иерархической модели РИГД применительно к ГСИ ДЗЗ позволит создавать эффективные в емкостном и временном смыслах алгоритмы их предобработки и распознавания.

Благодарности

Работа выполнена при финансовой поддержке Российского Фонда Фундаментальных Исследований, гранты РФФИ № 17–07–00543 и № 18–07–00715.

Литература

1. Васин Ю.Г., Васин Д.Ю., Громов В.П. Структурное описание растровых данных. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 23–29.
2. Васин Ю.Г., Башкиров О.А., Чудинович Б.М. Комбинаторно–геометрический подход в задачах анализа сложной графической информации. //Автоматизация обработки сложной графической информации: Межвуз. сб. науч. тр. / Под ред. Ю.Г. Васина: Горьк. гос. ун–т. Горький, 1987
3. Васин Ю.Г., Башкиров О.А., Рудометова С.Б. Математические модели структурированного описания графических изображений. //Автоматизация обработки сложной графической информации: Межвуз. сб. науч. тр. / Под ред. Ю.Г. Васина: Горьк. гос. ун–т. Горький, 1984
4. Громов В.П. Формирование нормализованных растров графических документов. //Методы и средства обработки сложной графической информации: 8 Всероссийская научная конференция, 12–16 сентября 2005г., Н.Новгород, стр. 51–54.
5. Горбунов А.А. Громов В.П. Робастная векторизация растровых линейных объектов. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 69–70.
6. Громов В.П. Построение геометрической модели графических документов с выраженной ориентацией линейных объектов. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 70 – 71.
7. Громов В.П., Смирнов А.Ф. Векторизация линейных объектов по штриховой модели бинарного растра. // Методы и средства обработки сложной графической информации: 8 Всероссийская научная конференция, 12–16 сентября 2005г., Н.Новгород, стр. 55 – 56.
8. Беляева Ю.В., Васин Ю.Г., Васин Д.Ю., Громов В.П. Подсистема создания цифровых баз данных поэтажных планов. //Методы и средства обработки сложной графической информации: 6 Всероссийская конференция с участием стран СНГ, 25–27 сентября 2001г., Н.Новгород, стр. 14 – 16.
9. J. Hull, “Document Image Skew Detection: Survey and Annotated Bibliography,” Document Analysis Systems II, World Scientific, pp. 40–64, 1998.
10. Dan S. Bloomberg, Gary E. Kopec and Lakshmi Dasari. Measuring document image skew and orientation. // Xerox Palo Alto Research Center. – Access mode: http://www.leptonica.com/papers/skew-measurement.pdf, free.
11. Васин Д.Ю., Редькин М.А. Автоматическое распознавание символов по структурной модели растровых бинарных изображений. // Труды Международной научной конференции «Ситуационные центры и ИАС4i для мониторинга и безопасности» 21-24 ноября 2016 г, Протвино, ЦарьГрад, Московская область, Россия, стр. 44–50.
12. Д.Ю. Васин, Редькин М.А. Распознавание символов на большеформатных растровых изображениях документов со слабо формализованным описанием объектов. // Труды 27–ой Международной конференции GraphiCon 2017, 24-28 сентября 2017 г, ПГНИУ, Пермь, Россия, стр.303 – 308.
13. Yu. G. Vasin, D.Yu. Vasin, V.P. Gromov. Intellectual Information Technology for Symbol Extraction from Ill-sructured Graphical Documents // Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence, Montreal, May 13-17, 2018. Published by CENPARMI Centre for Pattern Recognition and Machine Intelligence Concordia University, Montreal, Quebec H3G 1M8, Canada, pp. 724-732.
14. Перспективные информационные технологии дистанционного зондирования Земли: моногр. / под ред. В.А. Сойфера. – Самара: Новая техника, 2015. – 256 с

Об авторах

Васин Дмитрий Юрьевич – к.т.н., ведущий научный сотрудник Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского)
E–mail: dm04@list.ru.

Громов Владимир Петрович – ведущий программист Центра информатики и интеллектуальных информационных технологий Института Информационных технологий, математики и механики Нижегородского Госуниверситета им. Н.И. Лобачевского (ИТММ ННГУ им. Н.И. Лобачевского)
E–mail: gvp1941@bk.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Ротков Сергей Игоревич – д.т.н., профессор, зав. кафедрой инженерной геометрии, компьютерной графики и автоматизированного проектирования ННГАСУ
E–mail: rotkov@nngasu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Концепция мультимедийного дидактического комплекта для морфологического анализа в обучении студентов-дизайнеров

Е.В. Вехтер, В.Ю. Радченко, Т.Д. Казакова
vehter@tpu.ru|rad@tpu.ru|k.t.d2@mail.ru
Национальный исследовательский Томский политехнический университет, Томск, Россия

В статье предлагается концепция дидактического комплекта для морфологического анализа, предназначенного для развития изобразительных навыков бакалавров-дизайнеров, оформленного в мультимедиа-формате. Обосновывается роль морфологического анализа в развитии пространственного мышления будущих дизайнеров. Описывается содержание мультимедийного комплекта.

Ключевые слова: *компьютерная графика, мультимедийный дидактический комплект, морфологический анализ, развитие изобразительных навыков, дизайн-образование.*

1. Введение

Одной из тенденций развития современного высшего образования является активное использование мультимедийных технологий, компьютерной графики. Компьютерная графика в обучении выполняет две основные функции: иллюстративную и когнитивную [3]. Иллюстративная функция заключается, в первую очередь, в визуализации объектов, явлений и процессов действительности; когнитивная функция обусловлена тем, что компьютерная графика способствует интеллектуальному процессу получения учащимися новых знаний за счет выполнения ими различных действий (манипуляций) в виртуальном пространстве. Мультимедийные средства обеспечивают в процессе обучения возможность изменения объектов, управления их формой, цветом, размерами и т. п.

Наиболее активно элементы компьютерного обучения применяются при реализации образовательных программ подготовки дизайнеров. В связи с тем, что современный дизайнер должен уметь пользоваться компьютерными средствами визуализации, учебные планы по разным профилям направления подготовки «Дизайн» в различных вузах включают такие дисциплины, как «Компьютерная графика», «Компьютерные технологии в дизайне» и т. п., где студенты учатся использовать программы растровой и векторной графики, 3D-визуализации, анимации, которые в дальнейшем пригодятся им непосредственно в профессиональной деятельности. Однако компьютерные технологии при работе со студентами-дизайнерами могут быть использованы и на более ранних этапах обучения (на первом-втором курсах) в качестве вспомогательного инструмента при развитии базовых изобразительных навыков. Цель данной статьи – представить концепцию мультимедийного дидактического комплекта, предназначенного для развития изобразительных навыков бакалавров-дизайнеров, в частности для развития пространственного мышления.

2. Роль морфологического анализа в развитии пространственного мышления

Пространственное мышление представляет собой специфический вид мыслительной деятельности, которая направлена на решение задач, связанных с ориентацией в пространстве (теоретическом и практическом, видимом и воображаемом). Это мышление образами, основанное на фиксировании пространственных свойств и отношений [5]. Оно предполагает умение видеть особенности пространственных характеристик объекта. Пространственное мышление позволяет трансформировать, видоизменять наглядные образы и на их основе создавать новые, оно базируется на умении видеть мир объемным, трехмерным. В основе процесса пространственного мышления лежат операции сравнения, анализа, синтеза, абстракции, обобщения, конкретизации.

Для дизайнера пространственное мышление имеет особое значение, т. к. дизайнер должен уметь видеть объект проектирования в различных проекциях и ракурсах, с разных точек зрения, изнутри и снаружи; уметь преобразовывать трехмерный объект в двумерное изображение и наоборот [4]. Хорошо развитое пространственное мышление позволяет дизайнеру создавать уравновешенные объемно-пространственные композиции. Для качественного выполнения своих задач дизайнер должен уметь свободно оперировать пространственными образами.

Одним из основных методов развития пространственного мышления в художественном образовании и в дизайн-образовании является использование наглядного дидактического оборудования: геометрических гипсовых фигур, учебных гипсовых моделей сложной формы, наборов каркасных геометрических фигур. Работа с данным оборудованием (наблюдение, анализ, совмещение, трансформация, срисовывание и т. п.) позволяет учащимся развивать навыки построения сложных форм.

Однако недостатком данного классического оборудования является малая информативность: цельные фигуры не дают учащимся возможности увидеть внутренние грани объекта, что необходимо для точного построения формы. Работая с цельными фигурами, обучающиеся не могут изменять их конфигурацию, рассматривать их с определенного ракурса. Кроме того, классическое наглядное дидактическое оборудование в основном выполнено из гипса, такие модели громоздки, тяжелы и не безопасны в хранении и применении.

Указанные недочеты обусловили два главных принципа разрабатываемого в данном исследовании дидактического оборудования: 1) данный комплект должен обеспечивать формирование у обучающихся навыков морфологического анализа (анализа формы, внутреннего строения сложных объектов); 2) комплект выполнен в мультимедиа-формате, что расширяет возможности его применения при развитии изобразительных навыков – комплект может быть использован как для очной, так и для заочной и дистанционной форм обучения.

Идея морфологического анализа при развитии изобразительных навыков основывается на том, что любую форму сложного объема можно структурировать на составляющие ее простые геометрические тела: куб, пирамида, сфера, цилиндр, конус, призма. Геометрические тела представляют собой наиболее подходящие объекты, позволяющие понять законы формообразования предметов

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

и усвоить принципы их построения в рисунке [2]. Для успешной проектной деятельности дизайнеру необходимо уметь увидеть в сложном объекте простые геометрические тела, их взаимосвязь. Морфологический анализ рассматривается в данном исследовании как метод изучения формы, конструкции и взаимного расположения частей изделия [1]. Развитию умений морфологического анализа служит специальное дидактическое оборудование, позволяющее обучающимся увидеть, понять принципы построения сложных форм и научиться их использовать. В данном исследовании для развития умений морфологического анализа у студентов-дизайнеров разработана наглядная дидактическая модель (рис. 1), выполненная в двух материалах (фанера и оргстекло).

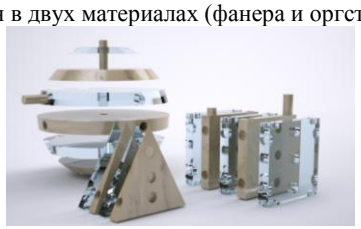


Рис. 1. 3D-модель дидактического комплекта морфологического анализа.

Основой данной модели являются простые геометрические фигуры (куб, сфера, конус, цилиндр), разбитые на плоскости, которые могут скрепляться между собой с помощью деревянных или металлических стержней, образуя, таким образом, различные сложные многоуровневые фигуры. На основе данного оборудования разработан мультимедийный дидактический комплект.

3. Содержание мультимедийного комплекта для развития пространственного мышления

Мультимедийный дидактический комплект для морфологического анализа при обучении студентов-дизайнеров включает два основных элемента: 1) видеоролики, выполняющие функцию демонстрации и объяснения материала; 2) пакет интерактивных заданий, предназначенных для тренировки развиваемых умений и контроля их сформированности.

1. Видеоролики. Предполагается десять 60-тисекундных видеороликов, в которых представлены варианты сборки дидактического оборудования – от простых геометрических фигур (четыре видеоролика) до сложных композиций (шесть видеороликов). Видеоматериалы представляются в двух форматах: 1) записанный на цифровую видеокамеру процесс ручной сборки комплекта с закадровым пояснением действий (для первого этапа работы с мультимедийным комплектом); 2) созданное в программе 3D-моделирования изображение процесса соединения элементов комплекта в различных вариантах (второй этап работы с мультимедийным комплектом). На рис. 2 представлен скриншот фрагмента видеоролика, демонстрирующего соединение деталей куба в геометрическую композицию. Видеоролик представляет собой 3D-сцену со сборкой основной конструкции.

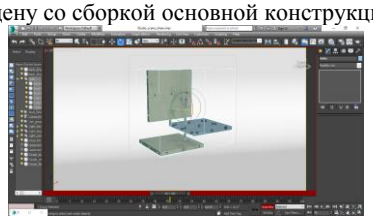


Рис. 2. Скриншот фрагмента видеоролика.

2. Пакет интерактивных заданий. Пакет интерактивных заданий предполагается создать на базе интернет-сервисов (LMS Moodle, LearningApps, Google Класс и т. п.). Каждое из последовательно усложняющихся заданий содержит заданную преподавателем конфигурацию элементов комплекта для морфологического анализа, из которых студентам необходимо составить различные варианты композиции в соответствии с заданием. Онлайн-сервисы позволят студентам самостоятельно компоновать необходимую для работы фигуру. Данный тип заданий (третий этап работы с мультимедийным комплектом) развивает пространственное мышление и позволяет изучать необходимую геометрическую фигуру, а также помогает студентам отрабатывать технику построения композиции. На завершающем этапе работы с мультимедийным дидактическим комплектом студентам предлагается перенести получившуюся композицию на бумагу, что позволяет формировать такие элементы изобразительных навыков, как техника рисунка, художественно-образное мышление, а также готовит студентов к работе по созданию авторских пространственных композиций с заданными параметрами.

4. Заключение

Таким образом, в данном исследовании разработана концепция мультимедийного дидактического комплекта для морфологического анализа для обучения студентов-дизайнеров. Работа с комплектом основана на принципах комплексности, последовательности и системности. Разработанный комплект учитывает выявленные недостатки классического дидактического оборудования, соответствует современным тенденциям в образовании, доступен и удобен для преподавателей и обучающихся. Мультимедийный комплект может быть как отдельным самостоятельным средством обучения, используемым в аудиторной и внеаудиторной работе студентов-дизайнеров, так и частью целого онлайн-курса по дисциплине «Пропедевтика» для студентов, обучающихся по образовательной программе бакалавриата «Дизайн».

5. Литература

[1] Горбунова Т.Г. Морфология как метод анализа вещественных археологических источников // Актуальные вопросы истории Сибири. – Барнаул, 2005. – С. 249–250.

[2] Мамугина В.П. Рисование геометрических форм и композиций: метод. разработки. – Тамбов: Изд-во Тамб. гос. техн. ун-та, 2009. – 32 с.

[3] Романюк А.Н., Зайдуллина С.Г., Поддубецкая М.П. Роль компьютерной графики в образовании // Новые образовательные технологии в вузе: материалы X международной научно-методической конференции. – Екатеринбург, 2013. – URL: http://elar.urfu.ru/bitstream/10995/26617/1/notv_2013_161.pdf.

[4] Щеглов А.В. Роль и место объемно-пространственной композиции в процессе обучения студентов-дизайнеров // Известия Тульского государственного университета. – 2012. – № 1–2. – С. 351–358.

[5] Якиманская И.С. Развитие пространственного мышления школьников: Автореф. дис. … д-ра психол. наук. – М., 1980. – 41 с.

Об авторах

Вехтер Евгения Викторовна – канд. пед. наук, доцент, Национальный исследовательский Томский политехнический университет. E-mail: vehter@tpu.ru.

Радченко Валерия Юрьевна – старший преподаватель, Национальный исследовательский Томский политехнический университет. E-mail: rad@tpu.ru.

Казакова Тамара Дмитриевна – магистрант, Национальный исследовательский Томский политехнический университет. E-mail: k.t.d2@mail.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Об уточнении некоторых понятий конструктивной геометрии

Д.В. Волошинов
denis.voloshinov@yandex.ru
Санкт-Петербургский государственный университет телекоммуникаций, Санкт-Петербург, Россия

Статья посвящена разработке рассмотрению нескольких геометрических конструкций, связанных с построением радикальной оси двух окружностей. Геометрические свойства, обнаруживающиеся в этих конструкциях, позволяют сделать ряд обобщений и уточнений в определении такого понятия, как радикальная ось, окружность бесконечно большого радиуса, преобразование n-мерного пространства в инверсии относительно сфер размерности n.

Ключевые слова: *радикальная ось, инволюция в точечных рядах, инверсия, конструктивная геометрия, геометрическое моделирование, Симплекс.*

The refinement of some terms in constructive geometry method

D.V. Voloshinov
denis.voloshinov@yandex.ru
Saint-Petersburg state University of Telecommunications, Saint-Petersburg, Russia

The article is devoted to the development of several geometric structures associated with the construction of the radical axis of two circles. The geometric properties found in these constructions allow us to make a number of generalizations and refinements in the definition of such concepts as radical axis, circle of infinitely radius, transformation of n-dimensional space in inversion in relation to n-dimensional spheres.

Keywords: *radical axis, involution in linear point series, inversion, constructive geometry, geometric modeling, Simplex.*

1. Введение

Геометрические объекты, которые получили название радикальных пространств, играют важную роль в решении многих геометрических задач, в частности, задач сопряжения n-мерных сфер в пространствах n измерений. Понятие «радикальный» тесно связан с понятием степени точки относительно некоторой выбранной n-сферы [1]. В плоскости под радикальной осью двух окружностей обычно понимают линию, несущую на себе точки, проявляющие одинаковую степень по отношению к этим окружностям. В литературе указывается, что данная линия является обычной прямой линией. Несложно показать, что получаемая линия проходит через точки пересечения окружностей, причем, если это пересечение явное, то прямая проходит через действительные точки пересечения; если же явного пересечения нет, то радикальная ось проходит через две мнимые комплексно-сопряженные точки определяющих ее окружностей. Существует несложное геометрическое построение, которое позволяет путем введения дополнительной третьей окружности, явно пересекающей исходные, построить радикальную ось, не используя при этом точки пересечения исходных окружностей. Поэтому, для окружностей, имеющих вещественные центр и радиус, построение радикальной оси даже вручную не составляет особого труда.

Общепринятое определение радикальной оси основывается на метрических понятиях геометрии. Однако при разработке средств автоматизации проектирования конструктивных геометрических моделей требуется учитывать не только метрические, но и глубинные проекционные свойства плоскости, поскольку проекционные преобразования геометрических образов не должны быть подвержены каким-либо исключениям, противоречащим общим положениям теории. Ситуация усугубляется, если задача построения радикальной оси ставится в отношении окружностей, хотя бы одна из которых определена на мнимом центре или радиусе.

Ответ на вопрос, какой линией будет представлена радикальная ось, существует ли она вообще и как ее построить, если она действительно существует, требует дополнительного изучения, поскольку ответы на подобные вопросы без дополнительного экспериментирования с геометрическими моделями в вопросах автоматизации не всегда очевидны. Более того, нельзя исключить вероятность того, что, несмотря на кажущуюся простоту и очевидность алгоритмов построения радикальной оси, за этим понятием скрываются и иные, нежели прямая линия, образы. Такие ситуации могут иметь место в тех случаях, когда из виду упускаются свойства бесконечно удаленных образов, неопределенности, вырождения или иные геометрические явления, подобные перечисленным. Анализ ряда алгоритмов и преобразований, проведенных автором статьи, позволил сделать заключение о том, что проявления свойств радикальной оси в конструктивной геометрической связи с определяющими ее объектами изучены недостаточно полно. Полученные результаты основываются не на аналитическом описании задачи, а на исследовании поведения обобщенных геометрических схем и взаимодействия их объектов. Из результатов следует неожиданный вывод: радикальная ось не является прямой линией в обычном понимании. Геометрические эксперименты [2–4] с конструктивными моделями показывают, что радикальная ось – это некая «предельная» окружность особого рода, геометрическая интерпретация которой возможна за счет единения двух прямых линий: собственной и бесконечно удаленной, причем несобственную точку «пересечения» этих прямых следует трактовать, как двойную.

Естественно, полученный результат становится возможным транзитивно распространить и на пространства многих измерений.

2. Особенности построения радикальной оси

Рассмотрим способ построения радикальной оси как прямой линии, проходящей через точки пересечения двух исходных окружностей. Пусть окружности a и b пересекаются в двух точках P и Q, а радикальная ось x проходит через эти точки. Тогда, выбрав на плоскости

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

точку M такую, что $M \not\subset a; M \not\subset b; M \not\subset x$, мы вправе провести через нее окружность $m: P \subset m; Q \subset m; M \subset m$. При этом радикальная ось x является таковой для любой пары окружностей, построенной на множестве $\{a,b,m\}$. Таким образом, произвольной точке M ставится одна единственная окружность из пучка окружностей с центрами P и Q (рис. 1).

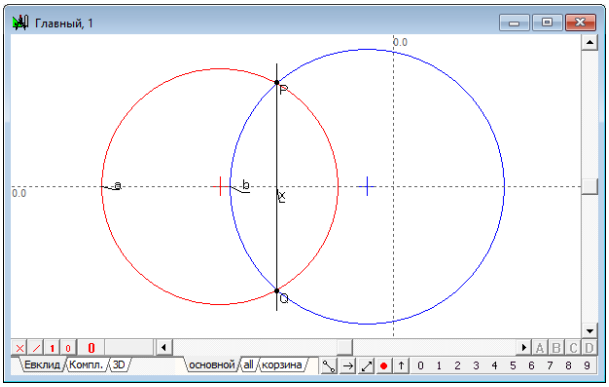


Рис.1. Радикальная ось как прямая, проходящая через точки пересечения окружностей.

Обращает на себя внимание тот факт, что ось x является одним из компонентов этого пучка. Причем ее образ может быть получен в тех случаях, если точка M располагается на x, в результате чего говорят, что окружность вырождается в прямую линию. Но этого же результата можно добиться, если точку M расположить на бесконечно удаленной прямой, замыкающей плоскость на себя. Причем сделать это можно, перемещая точку M в бесконечность по любому направлению. С одной стороны, мы приходим к выводу о том, что окружность бесконечно большого радиуса выродилась в прямую, с другой стороны, мы предстаем перед фактом, что эта прямая будет проведена через три точки, не лежащие на одной прямой. Получаемое логическое противоречие является весомым аргументом к тому чтобы, задуматься о геометрической природе объекта, ассоциируемого с понятием радикальной оси.

Второе соображение основывается на том, что любой пучок окружностей в пересечении с произвольной прямой линией l, образует на каждом элементе этого пучка пару точек (вещественных или мнимых), которые, являются двойными точками инволюций, индуцированных окружностями пучка в точечных рядах на линии l. Более того, рассматривая прямую l, как одномерное пространство, мы можем сделать еще одно важное обобщение: возникающая точечная пара в своем единстве – суть окружность одномерного пространства. Однако, понимая под элементом x прямую линию, мы получаем в пересечении с прямой l лишь единственную точку (рис. 2), что является причиной нарушения общей закономерности. В этом случае не возникает повода говорить и об «одномерной» окружности.

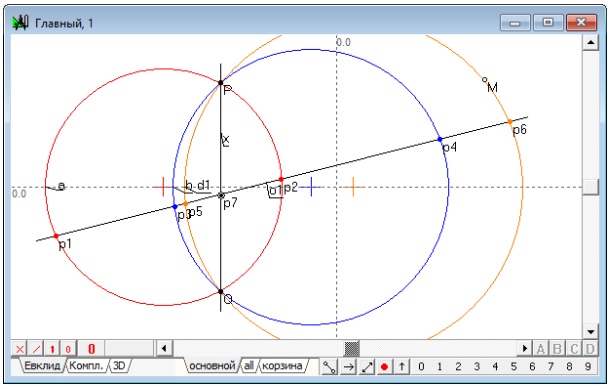


Рис.2. Проявление конструктивной «неоднородности» пучка окружностей.

Однако, можно заметить, что, если к существующей в общепринятом смысле радикальной оси причислить несобственную прямую плоскости, возникающее логическое несоответствие устраняется. Действительно, точек пересечения прямой l с реконструированным объектом становится две и, что более важно, на получаемых точках в парах окружностей можно построить инволюционное соответствие в рядах точек прямой l. Несложно показать, что двойные точки любых таких инволюционных соответствий одинаковы, в том числе и на точках, полученных от пересечения прямой l с прямой x и с бесконечно удаленной прямой (рис. 3).

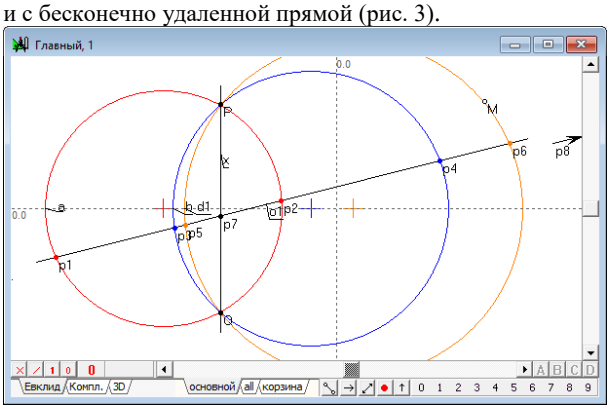


Рис.3. Иллюстрация дополнения радикальной оси бесконечно удаленной прямой.

Наблюдаемый результат можно трактовать и следующим образом. Как известно, все окружности плоскости, являясь частным (аффинным) проявлением кривых второго порядка, проходят через две фиксированные мнимые циклические точки, находящиеся на бесконечно удаленной прямой. Тогда радикальная ось, дополненная бесконечно удаленной прямой, с одной стороны, является окружностью, поскольку проходит через две циклические точки. В то же время радикальная ось проявляет свойства предельной вырожденной гиперболы, обычно интерпретируемой двумя «пересекающимися» прямыми линиями, одна из которых совпадает с бесконечно удаленной прямой.

3. Конструктивная равноценность «ветвей» радикальной оси

Покажем теперь, что с точки зрения конструктивного подхода к решению задач обе «ветви» расширенной радикальной оси равнозначны. Для этого решим задачу преобразования двух конических сечений в окружности и

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

найдем две коллинеации, осуществляющие такие преобразования.

Пусть на плоскости заданы две коники f и g. Эти коники индуцируют в плоскости два поляритета. Переведем коники в поляритетах относительно друг друга в коники f' и g' (рис. 4).

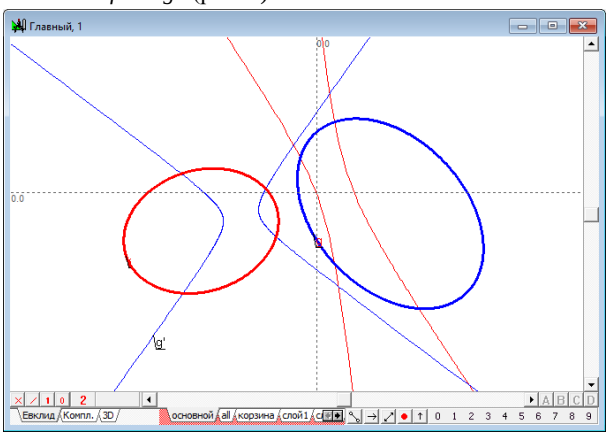


Рис.4. Исходные дачи к решению задачи о преобразовании двух коник в окружности.

Найдем точки пересечения коник $A_1, A_2, A_3, A_4 = f \times g'$. Строя на точках A_1, A_2, A_3, A_4 четырехвершинник, определим положение точек A, B и C – вершин автополярного треугольника, равнозначно индуцируемого обеими кониками f и g.

Проведем через вершины A и B прямую $ab = A \circ B$ и построим точки пересечения этой прямой с исходными кониками: $P_1, P_2 = f \times ab$ и $Q_1, Q_2 = g \times ab$. Найдем также центры коник f и g: соответственно P и Q. Через бесконечно удаленную точку T^{∞} прямой ab и точку C проведем прямую ab' (рис. 5).

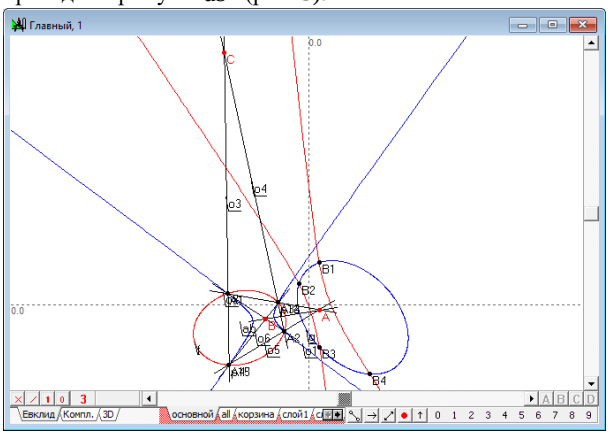


Рис.5. Построение вершин автополярного треугольника.

На двух наборах объектов (P_1, P, P_2, ab', C) и (Q_1, Q, Q_2, ab', C) построим коники m и n, полагая что прямая ab' является общей касательной к этим коникам в точке C (рис. 6).

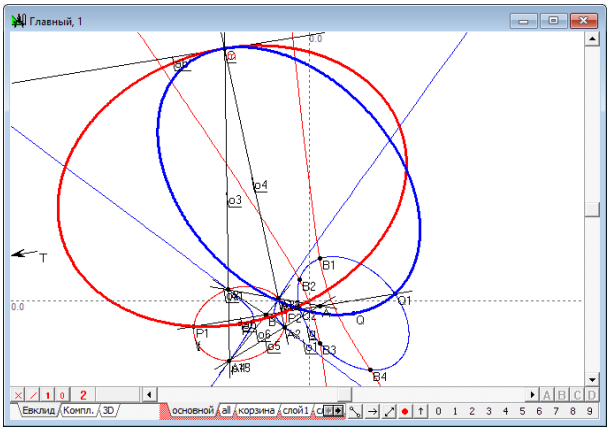


Рис.6. Построение вспомогательных коник m и n.

Имея общую двойную точку, коники m и n пересекаются в двух различных точках $U, V = m \times n$. Соединим точку C с точками U и V прямыми линиями $u = C \circ U$ и $v = C \circ V$ (рис. 7). Прямые u и v проходят через точки пересечения коник f и g (в примере мнимые), а это значит, что обе коники индуцируют на этих прямых пару одинаковых инволюционных соответствий η_u и η_v.

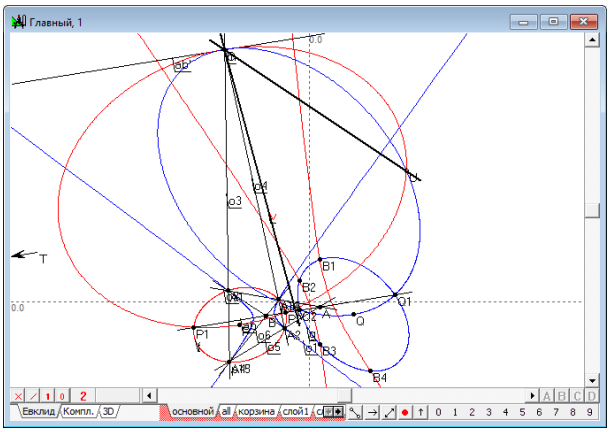


Рис.7. Прямые равных инволюций, индуцируемых исходными кониками.

Для определения коллинеаций, переводящих пару коник в пару окружностей, необходимо построить две пары точек Лагерра для прямых u и v соответственно. Это несложно сделать, переведя пару произвольных точек на каждой из прямых u и v в их образы в соответственных инволюциях η_u и η_v и построив на образах и прообразах этих точек, взятых как диаметральные, по паре окружностей. Точки пересечения соответственных окружностей и будут являться точками Лагерра, на которых следует построить две необходимые гомологии. Выполним эти действия раздельно для каждой их прямых u и v.

Возьмем на прямой u две произвольные точки U_1 и U_2 построим их образы: $U_1' = \eta_u(U_1)$ и $U_2' = \eta_u(U_2)$. Точки Лагерра LU_1 и LU_2 получим, как результат пересечения двух окружностей, построенных на парах точек $(U_1 - U_1')$ и $(U_2 - U_2')$, взятых в качестве диаметральных.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Выберем на прямой u две произвольные точки R_1 и R_3 и, спроецировав их, например, из центра LU_1 на бесконечно удаленную прямую, получим, соответственно, точки R_2 и R_4. Проведем также прямую линию $u' \| u$, на которой выберем две произвольные точки S_1 и S_2. Выбранный набор геометрических объектов позволяет задать гомологию $\chi_u \left| \begin{matrix} R_1, S_1, R_2, S_2 \\ R_3^\infty, S_1, R_4^\infty, S_2 \end{matrix} \right.$, переводящую пару коник f и g в две соответствующие им окружности f_u и g_u. При этом коллинеация χ_u переводит прямую v в линию x_v, а линию u в бесконечно удаленную прямую, то есть «расширенную» радикальную ось окружностей f_u и g_u (рис. 8).

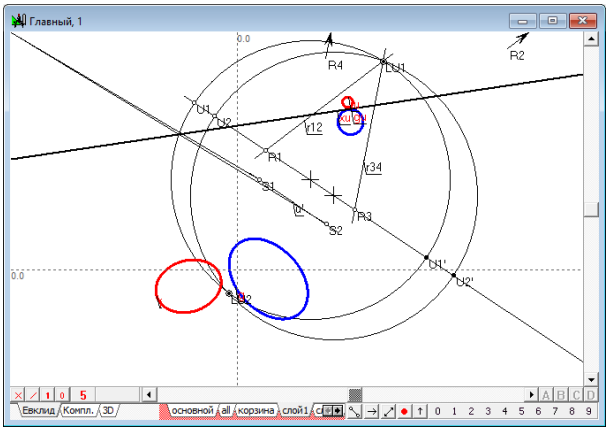


Рис.8. Построение двух окружностей и их радикальной оси с помощью коллинеации, определенной прямой u.

Второе построение аналогично первому лишь с той разницей, что оно выполняется относительно прямой v. Уже не рассматривая его подробно, отметим лишь, что коллинеация $\chi_v \left| \begin{matrix} K_1, S_3, K_2, S_4 \\ K_3^\infty, S_3, K_4^\infty, S_4 \end{matrix} \right.$ переводит прямую u в линию x_u, а линию v в бесконечно удаленную прямую, то есть «расширенную» радикальную ось окружностей f_v и g_v (рис. 9).

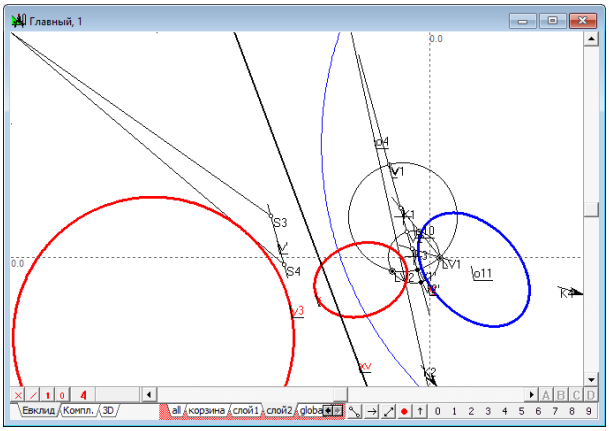


Рис.9. Построение двух окружностей и их радикальной оси с помощью коллинеации, определенной прямой v.

Из приведенных построений видно, что они конструктивно равнозначны, следовательно, предположение о том, что понятие радикальной оси допускает расширенное толкование, подтверждается единством подхода к реализации общей геометрической модели трансформации коник в окружности, можно считать правомерным.

4. Заключение

В результате проведенного исследования и проведенных экспериментов с конструктивными моделями выработаны рекомендации по изменению и расширению функционального состава системы геометрического моделирования Симплекс [5], позволившие устранить возникавшие из-за неполноты теории исключения при решении задач, основанных на понятии радикальных пространств.

5. Литература

[1] Бакельман И.Я. Инверсия. М.: Наука. 1966, 79 с.

[2] Волошинов Д.В. Геометрическая лаборатория. Закладываем основы // Качество графической подготовки: проблемы, традиции и инновации: Материалы VII международной Интернет-конференции. Февраль - март 2017 г. Пермь, 2017.

[3] Волошинов Д.В. Геометрическая лаборатория. Инструменты ортогональности // Качество графической подготовки: проблемы, традиции и инновации: Материалы VII международной Интернет-конференции. Февраль - март 2017 г. Пермь, 2017.

[4] Волошинов Д.В. Геометрическая лаборатория. Новый геометрический инструмент // Качество графической подготовки: проблемы, традиции и инновации: Материалы VII международной Интернет-конференции. Февраль - март 2017 г. Пермь, 2017.

[5] Волошинов Д.В. Конструктивное геометрическое моделирование. Теория, практика, автоматизация: монография / Д.В.Волошинов. – Saarbrücken: Lambert Academic Publishing, 2010. – 355 с.

Об авторе

Волошинов Денис Вячеславович, д.т.н., заведующий кафедрой информатики и компьютерного дизайна Санкт-Петербургского государственного университета телекоммуникаций им. М.А.Бонч-Бруевича. E-mail: denis.voloshinov@yandex.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Геометрическое моделирование и оптимизация конструктивных характеристик котлоагрегата

О.С. Воронова[1]
kornilova.oly@mail.ru
[1]ГОУ ВПО «Донбасская национальная академия строительства и архитектуры»,
г. Макеевка, Донецкая Народная Республика

В данной статье предложена геометрическая модель и вычислительный алгоритм на основе точечных уравнений зависимости конструктивных характеристик конвективной части котлоагрегата с учётом невязки теплового баланса. Полученная геометрическая модель даёт возможность определить оптимальное соотношение внутреннего диаметра конвективных труб и их количества с минимальным значением невязки теплового баланса.

Ключевые слова: *конструктивные характеристики, котлоагрегат, конвективная часть, невязка теплового баланса, геометрическая модель, факторы влияния, поверхность отклика, точечное уравнение.*

Geometrical modeling and optimization the constructive characteristics of the boiler

O.S. Voronova[1]
kornilova.oly@mail.ru
[1]Donbas National Academy of Civil Engineering and Architecture, Makeyevka, Donetsk People's Republic

In this article have been proposed the geometric model and the computational algorithm based on the point equations for the dependence of the constructive characteristics convective part of the boiler, taking into account the heat balance residual. The resulting geometric model makes it possible to determine the optimum ratio of the internal diameter of convective pipes and their number with the minimum value of the heat balance residual.

Keywords: *constructive characteristics, boiler, convective part, heat balance residual, geometric model, influence factors, response surface, point equation.*

1. Введение

В процессе проектирования котлоагрегатов (КА) большое значение имеет расчёт и оптимизация конструктивных характеристик КА. На сегодняшний день в инженерной практике существуют различные подходы к моделированию и расчёту конструктивных характеристик КА, основанных на тепловом балансе с учётом тепломассообменных процессов в топочной камере и конвективной части [5]. Кроме классических методов моделирования и оптимизации, также распространение получили инновационные методы с использованием современной вычислительной техники. Так в работе [6] для математического моделирования и оптимизации КА используются теория графов и численные методы моделирования, среди которых в первую очередь автор предлагает использовать метод Ньютона. В работах [7, 8] используются системы конечно-элементного анализа для проектирования и оптимизации жаротрубных КА. Основным недостатком предложенных подходов оптимизации является сложный и трудоёмкий вычислительный алгоритм моделирования и расчёта тепломассообменных процессов, требующий от проектировщика специальной подготовки. В работе [4] была предложена упрощенная методика численного расчета конструктивных параметров КА на примере двухконтурного пароконденсационного котла. Но даже упрощённая методика предполагает достаточно сложный расчёт конвективного теплообмена, что значительно затрудняет оптимизацию конструктивных элементов КА.

В данной работе приводится пример использования методов геометрического моделирования и оптимизации конструктивных размеров КА с использованием поверхности отклика и методов математического анализа. Для аналитического описания геометрической модели и её оптимизации используется математический аппарат БН-исчисление (точечное исчисление Балюбы-Найдыша [1-3]).

2. Исходные данные для моделирования

Жаротрубный КА представляет собой топочную камеру и конвективный газоход, и является сложным теплообменным аппаратом, где происходит процесс горения топлива, а также теплопередача от дымовых газов к теплоносителю.

Поскольку габариты котла менять нежелательно для оптимизации работы КА обычно используют конструктивные элементы конвективных газоходов. Основными конструктивными характеристиками конвективного газохода жаротрубного КА являются следующие параметры:

d_e – внутренний диаметр конвективных труб, м;

z – количество конвективных труб, шт.

Профессором Лукьяновым А.В. была предложена методика [5] поверочного расчёта теплообмена в топке и конвективной части жаротрубных КА, которая используется для инженерно-технических работников предприятий и проектных организаций котлостроения, а также для студентов, обучающихся по специальности «Телогазоснабжение и вентиляция» ГОУВПО «Донбасская национальная академия строительства и архитектуры». Эта методика основана на тепловом балансе КА, для соблюдения которого невязка теплового баланса ΔQ должна быть в пределах 2%.

Воспользуемся методикой, предложенной Лукьяновым А.В., для сбора исходной экспериментально-статистической информации, необходимой для геометрического моделирования и оптимизации конструктивных размеров КА. В качестве примера

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

воспользуемся паспортными данными теплогенератора КВ-0,3 с номинальной теплопроизводительностью 0,3 МВт.

Другие исходные данные:

1. Температура уходящих газов – 150 °С.
2. Параметры теплоносителя – 95/65.
3. Топливо – природный газ.
4. Газопровод – Средняя Азия – Центр.

В данном случае внутренний диаметр конвективных труб и их количество будут выступать факторами, влияющими на функцию отклика – невязку теплового баланса. Ограничим факторы влияния в следующих пределах: $d_в$=0,03–0,07 м; z=10–50 шт. Далее при помощи офисного пакета *MS Excel* был выполнен расчёт 36 последовательных комбинаций факторов влияния с вычислением значения функции отклика. Результаты расчёта представлены в таблице 1.

Таблица 1. Исходные экспериментально-статистические данные для геометрического моделирования и оптимизации конструктивных размеров КА.

№ п/п	$d_в$, м	z, шт	ΔQ, %	№ п/п	$d_в$, м	z, шт	ΔQ, %
1.	0,03	10	17,26	19.	0,054	10	26,28
2.	0,03	18	32,01	20.	0,054	18	16,7
3.	0,03	26	42,22	21.	0,054	26	9,96
4.	0,03	34	50,18	22.	0,054	34	4,62
5.	0,03	42	56,79	23.	0,054	42	0,13
6.	0,03	50	62,48	24.	0,054	50	3,78
7.	0,038	10	2,8	25.	0,062	10	33,77
8.	0,038	18	9,52	26.	0,062	18	25,01
9.	0,038	26	18,08	27.	0,062	26	18,8
10.	0,038	34	24,78	28.	0,062	34	13,85
11.	0,038	42	30,36	29.	0,062	42	9,65
12.	0,038	50	35,18	30.	0,062	50	5,97
13.	0,046	10	16,41	31.	0,07	10	39,65
14.	0,046	18	5,7	32.	0,07	18	31,51
15.	0,046	26	1,78	33.	0,07	26	25,68
16.	0,046	34	7,67	34.	0,07	34	20,98
17.	0,046	42	12,6	35.	0,07	42	16,98
18.	0,046	50	16,87	36.	0,07	50	13,44

3. Геометрическая модель конструктивных характеристик жаротрубного котлоагрегата

Составим геометрическую схему моделирования и конструктивных характеристик жаротрубного КА. Поскольку факторов влияния два, то геометрическим объектом, соответствующим моделируемому процессу будет двухпараметрическое множество точек, т.е. поверхность. Поэтому выберем для моделирования трехмерную декартову систему координат (рис. 1), установив соответствие между осями и значениями факторов влияния, а также функцией отклика. Так оси *Ox* декартовой системы координат соответствует количество конвективных труб; оси *Oy* – внутренний диаметр конвективных труб; оси *Oz* – невязка теплового баланса.

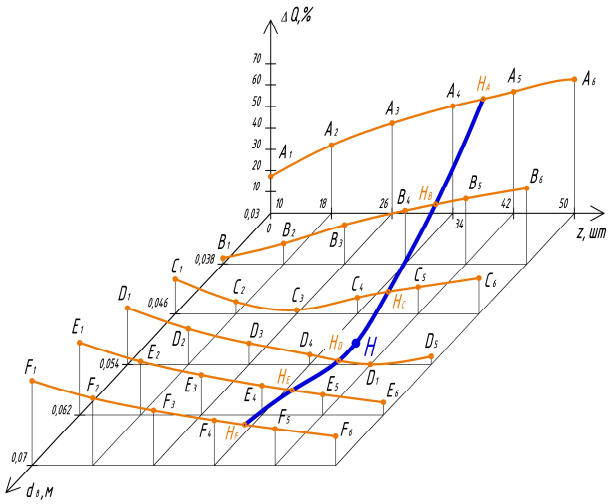


Рис. 1. Геометрическая схема моделирования поверхности отклика.

Для наглядного представления геометрической схемы (рис. 1) осям декартовой системы координат был присвоен произвольный масштаб.

В соответствии с геометрической схемой (рис. 1) сформируем 6 направляющих линий H_A, H_B, H_C, H_D, H_E, H_F по 6 точек в каждой, которым соответствуют значения исходных экспериментально-статистических данных (табл. 1). Определим направляющие дуги кривых с помощью дуги кривой 5-го порядка, проходящей через 6 наперёд заданных точек [3].

Для направляющей дуги кривой линии H_A точечное уравнение имеет следующий вид:

$$\begin{aligned} H_A = {} & A_1\left[\bar{u}^5 - 6,4\bar{u}^4u + 11,2\bar{u}^3u^2 - 6,4\bar{u}^2u^3 + \bar{u}u^4\right] + \\ & + A_2\left[25\bar{u}^4u - 60,4\bar{u}^3u^2 + 38,5\bar{u}^2u^3 - 6,3\bar{u}u^4\right] + \\ & + A_3\left[-25\bar{u}^4u + 122,9\bar{u}^3u^2 - 95,8\bar{u}^2u^3 + 16,7\bar{u}u^4\right] + \\ & + A_4\left[16,7\bar{u}^4u - 95,8\bar{u}^3u^2 + 122,9\bar{u}^2u^3 - 25\bar{u}u^4\right] + \\ & + A_5\left[-6,3\bar{u}^4u + 38,5\bar{u}^3u^2 - 60,4\bar{u}^2u^3 + 25\bar{u}u^4\right] + \\ & + A_6\left[\bar{u}^4u - 6,4\bar{u}^3u^2 + 11,2\bar{u}^2u^3 - 6,4\bar{u}u^4 + u^5\right], \end{aligned} \quad (1)$$

где u – параметр, который определяет положение текущей точки на соответствующей направляющей,

$\bar{u} = 1 - u$ – дополнение параметра до единицы.

Аналогичным образом определим точечные уравнения ещё 5-ти направляющих дуг кривых, которые также будут иметь общий согласовывающий текущий параметр u.

Дуга кривой H, которая является образующей отсека поверхности отклика, также определяется при помощи точечного уравнения дуги кривой 5-го порядка, проходящей через 6 наперёд заданных точек:

$$\begin{aligned} H = {} & H_A\left[\bar{v}^5 - 6,4\bar{v}^4v + 11,2\bar{v}^3v^2 - 6,4\bar{v}^2v^3 + \bar{v}v^4\right] + \\ & + H_B\left[25\bar{v}^4v - 60,4\bar{v}^3v^2 + 38,5\bar{v}^2v^3 - 6,3\bar{v}v^4\right] + \\ & + H_C\left[-25\bar{v}^4v + 122,9\bar{v}^3v^2 - 95,8\bar{v}^2v^3 + 16,7\bar{v}v^4\right] + \\ & + H_D\left[16,7\bar{v}^4v - 95,8\bar{v}^3v^2 + 122,9\bar{v}^2v^3 - 25\bar{v}v^4\right] + \\ & + H_E\left[-6,3\bar{v}^4v + 38,5\bar{v}^3v^2 - 60,4\bar{v}^2v^3 + 25\bar{v}v^4\right] + \\ & + H_F\left[\bar{v}^4v - 6,4\bar{v}^3v^2 + 11,2\bar{v}^2v^3 - 6,4\bar{v}v^4 + v^5\right], \end{aligned} \quad (2)$$

где v – параметр, который определяет положение текущей точки на образующей поверхности отклика,

$\bar{v} = 1 - v$ – дополнение параметра до единицы.

Таким образом, получим вычислительный алгоритм формирования поверхности отклика, проходящей через 36

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

наперёд заданных точек, соответствующих исходным экспериментально-статистическим данным из табл. 1. Следует отметить, что полученный вычислительный алгоритм представлен в точечной форме. Для перехода от точечных уравнений к параметрическим, необходимо выполнить покоординатный расчёт, в результате которого получим систему параметрических уравнений проекций на оси декартовой системы координат (рис. 1).

Для визуального анализа результатов моделирования на рис. 2 представим визуализацию полученной геометрической модели зависимости невязки теплового баланса от внутреннего диаметра конвективных трубок и их количества при помощи программного пакета *Maple*.

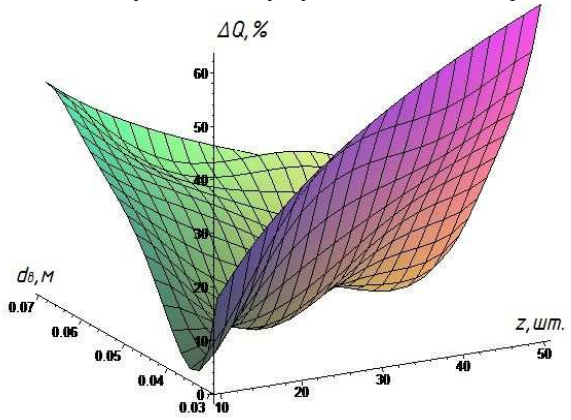


Рис. 2. Визуализация геометрической модели конструктивных характеристик КА.

4. Анализ результатов моделирования и оптимизация конструктивных характеристик котлоагрегата

Для оптимизации работы КА необходимо минимизировать невязку теплового баланса. Поскольку функция отклика соответствующая невязке теплового баланса представлена в виде двухпараметрического множества точек, то её минимизация относится к задачам теории математического анализа функции двух переменных, в соответствии с которой определим в программном пакете *Maple* точки подозрительные на экстремум. В результате решения системы уравнений:

$$\begin{cases} \dfrac{\partial H}{\partial u} = 0 \\ \dfrac{\partial H}{\partial v} = 0 \end{cases}$$

были получены комбинации значений параметров u и v, из которых были выбраны комбинации удовлетворяющие условиям: $0 \le u \le 1$ и $0 \le v \le 1$. Далее подстановкой в уравнение (2) были отобраны 4 минимальных значения невязки теплового баланса. Полученные значения четырех возможных вариантов соотношения количества конвективных труб и их диаметра проверим с помощью методики поверочного расчёта КА, которая была реализована при помощи офисного пакета *MS Excel*. В результате получим следующие значения.

Таблица 2. Анализ подбора оптимального соотношения внутреннего диаметра конвективных труб и их количества.

№ п/п	Параметры	z, шт	$d_в$, м	ΔQ_1, %	ΔQ_2, %
1.	u=0,13 v=0,26	15	0,04	5,83	1,35
2.	u=0,36 v=0,4	24	0,046	1,5	0,1
3.	u=0,57 v=0,51	33	0,05	4,06	0,41
4.	u=0,87 v=0,63	45	0,055	0,65	0,04

где ΔQ_1 – значения невязки теплового баланса полученные по геометрической модели,

ΔQ_2 – значения невязки теплового баланса, которые получены по расчету в *MS Excel*.

Как видно из результатов, представленных в таблице 2, целых 4 значения невязки теплового баланса, удовлетворяют условию $\Delta Q \le 2\%$, что даёт возможность проектировщику выбрать наиболее оптимальный вариант, поскольку не всегда вариант, дающий минимальную невязку теплового баланса, является оптимальным для проектирования.

5. Заключение

Предложенная геометрическая модель конструктивных характеристик КА позволяет оптимизировать конструктивные размеры конвективной части жаротрубного КА.

В данном примере полного совпадения геометрической модели конструктивных характеристик котлоагрегата с поверочным расчетом КА, основанном на тепловом балансе, достигнуто не было. Однако, предложенная геометрическая модель конструктивных характеристик котлоагрегата, позволяет найти оптимальные значения соотношения количества конвективных труб и их диаметра. Оказалось, что для данного КА таких значений, в пределах изменения параметров, может быть 4. Также просматривается некоторая зависимость между значениями внутреннего диаметра конвективных трубок и их количеством, удовлетворяющими требуемой невязке теплового баланса, что является перспективой дальнейших исследований.

6. Литература

[1] Балюба, И. Г. Конструктивная геометрия многообразий в точечном исчислении [Текст]: дис. … доктора техн. наук: 05.01.01 / Балюба И. Г. – Макеевка, 1995. – 227 с.

[2] Балюба, И.Г. Точечное исчисление [Текст]: [учебное пособие] / И.Г. Балюба, В.М. Найдыш; под ред. В.М. Верещаги. – Мелитополь: МГПУ им. Б. Хмельницкого, 2015. – 236 с.

[3] Бумага, А. И. Геометрическое моделирование физико-механических свойств композиционных строительных материалов в БН-исчислении [Текст]: дис. … канд. техн. наук: 05.23.05, 05.01.01 / Бумага А. И. – Макеевка, 2016. – 164 с.

[4] Жидилов К.А. Методика математического моделирования газового водогрейного котла [Текст] / Жидилов К.А., Киселев В.Ф., Проворов В.В., Ротков С.И. – Вестник Ижевского государственного технического университета. – Ижевск: изд-во Ижевского государственного технического университета им. М.Т. Калашникова. – № 3 – С.33-36.

[5] Лук'янов О.В. Теплогенератори для локальних систем теплопостачання [Текст] / Лук'янов О.В. – Макіївка.: ДонДАБА, 2003. – 149 с.

[6] Попырин, Л.С. Математическое моделирование и оптимизация теплоэнергетических установок [Текст] / Попырин Л.С. – М. Энергия, 1978 – 416с.

[7] Хаустов С.А., Компьютерное моделирование гидродинамики жаротрубного котла с использованием конечно-элементного анализа [Текст] / Хаустов С.А., Хаустов П.А., Максимова Е.И. – Современные проблемы

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

науки и образования. 2014 -№. 6. С. 1 – 6. Режим доступа: https://elibrary.ru/item.asp?id=22877366
[8] Хаустов С.А. Современные тенденции проектирования жаротрубных котлов [Текст] / Хаустов С.А., Заворин А.С. – Вестник науки Сибири. – 2014. – №. 2 (12). – С. 21 – 28.

Об авторах

Воронова Ольга Сергеевна, ассистент кафедры специализированных информационных технологий и систем, строительного факультета Донбасской национальной академии строительства и архитектуры.
Её e-mail kornilova.oly@mail.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Решение дифференциальных уравнений методами геометрического моделирования

Е.В. Конопацкий[1]
e.v.konopatskiy@mail.ru
[1]ГОУ ВПО «Донбасская национальная академия строительства и архитектуры»,
г. Макеевка, Донецкая Народная Республика

В статье приводятся теоретические основы и принципиальный вычислительный алгоритм, необходимые для решения дифференциальных уравнений с помощью геометрического моделирования объектов многомерного аффинного пространства, проходящих через наперёд заданные точки. Также приводится пример решения простейших дифференциальных уравнений модифицированным методом конечных элементов.

Ключевые слова: *геометрическое моделирование, многомерное пространство, дифференциальные уравнения, геометрический объект, функция отклика, метод конечных элементов.*

Solving Differential Equations by Geometric Modelling Methods

E.V. Konopatskiy[1],
e.v.konopatskiy@mail.ru
[1] Donbas National Academy of Civil Engineering and Architecture, Makeyevka, Donetsk People's Republic

The theoretical principles and the basic computational algorithm necessary for solving differential equations with the help of geometric modelling of objects a multidimensional affine space passing through given points are given in the article. An example of solving the simplest differential equations by a modified finite element method is also given.

Keywords: *geometric modeling, multidimensional space, differential equations, geometric object, response function, finite element method.*

1. Введение

Математическое и, как следствие, компьютерное моделирование, давно заняло одно из главенствующих положений в системе формирования знаний, накопленных человечеством, как источник новых знаний о протекании тех или иных процессов и явлений живой и не живой природы, техники, технологии и т.п. Однако отдельным направлением можно выделить создания моделей для решения непосредственно математических задач. В этом случае моделирование выступает как отдельный инструмент решения сугубо математических задач, которые впоследствии могут быть использованы уже для решения каких-либо прикладных и инженерных задач. В качестве подобного примера рассмотрим в данной статье использование методов геометрического моделирования с применением многомерной геометрии для решения дифференциальных уравнений в частных производных.

Уравнения в частных производных нашли широкое применение в разделе математической физики [4] для аналитического описания моделей различных физических явлений. Вместе с тем возникает задача решения полученных уравнений в частных производных, большинство из которых не выражаются через элементарные функции. В таких случаях используют численные методы решения дифференциальных уравнений [7], к которым, например, относятся метод конечных элементов [5] и метод конечных разностей [7]. Эти методы имеют огромное прикладное значение, поскольку являются основой для всех существующих современных САПР. Основным же их недостатком является огромный объём вычислений. Здесь следует отметить, что задачи, решаемые методом конечных элементов, могут иметь линейную и нелинейную постановку. При этом в линейной постановке решение задачи методом конечных элементов обладает большим быстродействием, однако полученный результат не всегда соответствует истине, поскольку и в природе и в технике линейные процессы и явления встречаются крайне редко. Чаще всего они как раз и носят нелинейный характер. Причём сама нелинейность может иметь различную природу. Например, в строительной механике разделяют нелинейность 4-х видов:

1) геометрическая;
2) физическая;
3) конструктивная;
4) генетическая.

В результате конечно-элементный анализ конструкции с учётом хотя бы одной из выше перечисленных нелинейностей – это решение отдельной инженерной задачи, связанной с большими затратами временных ресурсов как на составление корректной конечно-элементной схемы, так и на сами вычисления. Поэтому усовершенствование существующих вычислительных методов решения дифференциальных уравнений в частных производных можно считать актуальной научно-практической задачей.

2. Исходная идея

Рассмотрим обыкновенное дифференциальное уравнение:

$$F\left(x, y, y', y'', ..., y^{(n)}\right) = 0.$$

Что значит решить это уравнение? Это значит необходимо установить такую функциональную зависимость между переменной x и функцией отклика y, которая удовлетворяла бы условию уравнения. Геометрически это означает, что нужно определить однопараметрическое множество точек, т.е. построить такую дугу кривой, которая на определенном участке могла бы удовлетворить условию исходного уравнения. Здесь возникает несколько вариантов решения. Можно использовать единую дугу кривой, проходящую через наперед заданные точки, а можно составную, состоящую из нескольких, каким-либо образом, состыкованных между собой дуг. Если искомая кривая носит монотонный и гладкий характер, то конечно удобно искать уравнение

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

единой дуги кривой. С другой стороны, если искомая дуга кривой носит сложный незакономерный характер, то единственно возможным способом её описания будет составная кривая. Таким образом, переходим к идее метода конечных элементов, которая заключается в том, чтобы разбить искомую область на конечное число элементов и аппроксимировать искомую кривую дугами более простых кривых. Говоря другими словами, нужно подобрать такую дугу кривой на каждом участке, чтобы выполнялось условие исходного уравнения на всем интервале решения задачи. При этом стыковка дуг осуществляется за счёт составления глобальных матриц численного решения дифференциальных уравнений.

Здесь следует остановиться на уравнениях дуг аппроксимационных кривых. Одним из наиболее простых и вместе с тем эффективных способов – это использование дуг алгебраических кривых проходящих через наперед заданные точки [3], которые были получены в БН-исчислении (точечное исчисление Балюбы-Найдыша [1, 2, 6]) на основе полиномов Бернштейна. Использование дуг кривых, проходящих через наперёд заданные точки, даёт возможность создавать геометрические объекты, носителем которых являются узловые точки. При этом порядок кривой напрямую зависит от степени производной исходного дифференциального уравнения. Дабы не усложнять решение и уменьшить количество вычислений, рекомендуется выбирать кривую аппроксимации таком образом, чтобы порядок кривой был на единицу больше порядка производной функции. Хотя в некоторых случаях есть необходимость в использовании кривых более высокого порядка. Всё зависит от исходного дифференциального уравнения.

Если рассматривать другие дифференциальные уравнения в частных производных, то, с точки зрения геометрического моделирования, необходимо выделить лишь количество переменных, влияющих на функцию отклика, а значит и на выбор аппроксимирующего геометрического объекта. Например, двухпараметрическое множество точек – поверхность отклика, позволяет решать дифференциальные уравнения следующего вида:

$$F\left(z,x,x',x'',...,x^{(n)},y,y',y'',...,y^{(n)}\right)=0.$$

Трёхпараметрическое множество точек – гиперповерхность отклика позволяет решать дифференциальные уравнения с большим количеством переменных:

$$F\left(t,x,x',x'',...,x^{(n)},y,y',y'',...,y^{(n)},z,z',z'',...,z^{(n)},...\right)=0.$$

Таким образом, модификация метода конечных элементов заключается в том, что решение дифференциального уравнения представляется в виде совокупности геометрических объектов функции отклика, проходящих через наперед заданные точки. Это даёт возможность обобщения для вычислительного решения различных дифференциальных уравнений в виде совокупности геометрических объектов функции отклика, проходящих через наперед заданные точки и принадлежащих многомерному аффинному пространству. Одним из принципиальных отличий от классического метода конечных элементов является то, что вместо увеличения степени полинома на каждом участке для достижения необходимого количества параметров увеличивается размерность пространства, в котором расположен геометрический объект, выбранный для аппроксимации решения исходного дифференциального уравнения. Основное же отличие от классического метода конечных элементов заключается в том, что на конечные элементы разбивается не исследуемый объект, а геометрический объект функции отклика, проходящий через наперёд заданные точки, конечная форма которого как раз и определяется в результате вычислений.

3. Покоординатный расчёт для решения дифференциальных уравнений

В классической литературе дифференциальные уравнения в частных производных представляются в явном виде, реже в неявном виде, как было показано выше. Для построения функций отклика в БН-исчислении, используются геометрические объекты, проходящие через наперед заданные точки в виде точечных уравнений и вычислительных алгоритмов, которые по своей сути являются символьной записью. Для перехода от точечных уравнений к параметрическим необходимо выполнить покоординатный расчёт. Здесь возникает отдельная задача, как продифференцировать функцию отклика по переменным, если она задана параметрически? Для дуги кривой, как однопараметрического множества, эта задача не представляет сложности и хорошо известна из курса дифференциального исчисления:

$$y'_x=\frac{y'_t}{x'_t}\ \text{ или }\ \frac{\partial y}{\partial x}=\frac{\partial y}{\partial t}\Big/\frac{\partial x}{\partial t}.$$

С поверхностью уже немного сложнее. Предположим, задан отсек поверхности отклика в трёхмерном аффинном пространстве системой параметрических уравнений:

$$\begin{cases}x=f_1(u,v),\\ y=f_2(u,v),\\ z=f_3(u,v).\end{cases}$$

Необходимо определить частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$.

Принципиальный подход к решению поставленной задачи представлен в работе [8] и сводится к решению однородной системы двух дифференциальных уравнений первой степени:

$$\begin{cases}\frac{\partial F}{\partial x}\cdot\frac{\partial f_1}{\partial u}+\frac{\partial F}{\partial y}\cdot\frac{\partial f_2}{\partial u}+\frac{\partial F}{\partial z}\cdot\frac{\partial f_3}{\partial u}=0,\\ \frac{\partial F}{\partial x}\cdot\frac{\partial f_1}{\partial v}+\frac{\partial F}{\partial y}\cdot\frac{\partial f_2}{\partial v}+\frac{\partial F}{\partial z}\cdot\frac{\partial f_3}{\partial v}=0,\end{cases}$$

где $F(x,y,z)=0-$ уравнение поверхности отклика в неявном виде.

Решая полученную систему дифференциальных уравнений методом Крамера относительно неизвестных $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ и $\frac{\partial F}{\partial z}$, получим:

$$\frac{\partial F}{\partial x}=k\left(\frac{\partial f_2}{\partial u}\cdot\frac{\partial f_3}{\partial v}-\frac{\partial f_3}{\partial u}\cdot\frac{\partial f_2}{\partial v}\right),$$
$$\frac{\partial F}{\partial y}=k\left(\frac{\partial f_3}{\partial u}\cdot\frac{\partial f_1}{\partial v}-\frac{\partial f_1}{\partial u}\cdot\frac{\partial f_3}{\partial v}\right),$$
$$\frac{\partial F}{\partial z}=k\left(\frac{\partial f_1}{\partial u}\cdot\frac{\partial f_2}{\partial v}-\frac{\partial f_2}{\partial u}\cdot\frac{\partial f_1}{\partial v}\right),$$

где $k-$ некоторый коэффициент пропорциональности.

Отсюда определяем искомые частные производные:

$$\frac{\partial z}{\partial x}=\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}}=\frac{\frac{\partial f_2}{\partial u}\cdot\frac{\partial f_3}{\partial v}-\frac{\partial f_3}{\partial u}\cdot\frac{\partial f_2}{\partial v}}{\frac{\partial f_1}{\partial u}\cdot\frac{\partial f_2}{\partial v}-\frac{\partial f_2}{\partial u}\cdot\frac{\partial f_1}{\partial v}},$$
$$\frac{\partial z}{\partial y}=\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}=\frac{\frac{\partial f_3}{\partial u}\cdot\frac{\partial f_1}{\partial v}-\frac{\partial f_1}{\partial u}\cdot\frac{\partial f_3}{\partial v}}{\frac{\partial f_1}{\partial u}\cdot\frac{\partial f_2}{\partial v}-\frac{\partial f_2}{\partial u}\cdot\frac{\partial f_1}{\partial v}}.$$

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Обобщая рассмотренный выше способ дифференцирования, для трёхпараметрической гиперповерхности отклика, принадлежащей 4-мерному пространству и заданной в неявном виде: $F(x,y,z,t)=0$, где $x=f_1(u,v,w)$, $y=f_2(u,v,w)$, $z=f_3(u,v,w)$, $t=f_4(u,v,w)$, получим:

$$\frac{\partial t}{\partial x}=\frac{\begin{vmatrix}\frac{\partial f_4}{\partial u}&\frac{\partial f_2}{\partial u}&\frac{\partial f_3}{\partial u}\\ \frac{\partial f_4}{\partial v}&\frac{\partial f_2}{\partial v}&\frac{\partial f_3}{\partial v}\\ \frac{\partial f_4}{\partial w}&\frac{\partial f_2}{\partial w}&\frac{\partial f_3}{\partial w}\end{vmatrix}}{\begin{vmatrix}\frac{\partial f_1}{\partial u}&\frac{\partial f_2}{\partial u}&\frac{\partial f_3}{\partial u}\\ \frac{\partial f_1}{\partial v}&\frac{\partial f_2}{\partial v}&\frac{\partial f_3}{\partial v}\\ \frac{\partial f_1}{\partial w}&\frac{\partial f_2}{\partial w}&\frac{\partial f_3}{\partial w}\end{vmatrix}},\quad \frac{\partial t}{\partial y}=\frac{\begin{vmatrix}\frac{\partial f_1}{\partial u}&\frac{\partial f_4}{\partial u}&\frac{\partial f_3}{\partial u}\\ \frac{\partial f_1}{\partial v}&\frac{\partial f_4}{\partial v}&\frac{\partial f_3}{\partial v}\\ \frac{\partial f_1}{\partial w}&\frac{\partial f_4}{\partial w}&\frac{\partial f_3}{\partial w}\end{vmatrix}}{\begin{vmatrix}\frac{\partial f_1}{\partial u}&\frac{\partial f_2}{\partial u}&\frac{\partial f_3}{\partial u}\\ \frac{\partial f_1}{\partial v}&\frac{\partial f_2}{\partial v}&\frac{\partial f_3}{\partial v}\\ \frac{\partial f_1}{\partial w}&\frac{\partial f_2}{\partial w}&\frac{\partial f_3}{\partial w}\end{vmatrix}},$$

$$\frac{\partial t}{\partial z}=\frac{\begin{vmatrix}\frac{\partial f_1}{\partial u}&\frac{\partial f_2}{\partial u}&\frac{\partial f_4}{\partial u}\\ \frac{\partial f_1}{\partial v}&\frac{\partial f_2}{\partial v}&\frac{\partial f_4}{\partial v}\\ \frac{\partial f_1}{\partial w}&\frac{\partial f_2}{\partial w}&\frac{\partial f_4}{\partial w}\end{vmatrix}}{\begin{vmatrix}\frac{\partial f_1}{\partial u}&\frac{\partial f_2}{\partial u}&\frac{\partial f_3}{\partial u}\\ \frac{\partial f_1}{\partial v}&\frac{\partial f_2}{\partial v}&\frac{\partial f_3}{\partial v}\\ \frac{\partial f_1}{\partial w}&\frac{\partial f_2}{\partial w}&\frac{\partial f_3}{\partial w}\end{vmatrix}}.$$

Аналогичным образом можно продифференцировать гиперповерхность отклика, принадлежащую $n-$мерному пространству. При этом будет увеличиваться только порядок определителя матрицы, составленной из частных производных. С другой стороны, для определения производной $k-$го порядка достаточно продифференцировать функцию k раз.

4. Принципиальный вычислительный алгоритм модифицированного метода конечных элементов

Собственно говоря, как и в классическом методе конечных элементов, каждая инженерная задача носит отдельный характер и обладает своими особенностями. Тем не менее, можно выделить принципиальный алгоритм решения дифференциальных уравнений модифицированным методом конечных элементов, который состоит из 6 этапов.

1) Выбрать дугу аппроксимирующей кривой, которая будет основой для создания геометрического объекта, проходящего через наперёд заданные точки.
2) Выполнить покоординатный расчёт и продифференцировать его результат в соответствии с исходным дифференциальным уравнением.
3) Поочередно подставлять значения параметров в узловых точках, формируя тем самым локальную систему линейных алгебраических уравнений (СЛАУ).
4) При необходимости повторяем первые 3 этапа несколько раз, накапливая, таким образом, локальные СЛАУ для формирования глобальной СЛАУ.
5) Решаем полученную СЛАУ и определяем необходимые значения функции отклика в узловых точка. После чего подставляем результат вычислений в параметрические уравнения аппроксимации.
6) Анализируем полученный результат и проверяем его достоверность. В случае получения недостаточно точных результатов увеличиваем количество геометрических объектов функции отклика, которые в нашем случае играют роль конечного элемента.

Таковы основные этапы решения дифференциальных уравнений модифицированным методом конечных элементов. Интересно, что для решения некоторых дифференциальных уравнений достаточно использовать всего один геометрический объект, т.е. один конечный элемент.

5. Маленький пример

В качестве примера предложенной модификации метода конечных элементов рассмотрим решение простейшего дифференциального уравнений первого порядка: $y'=2x^2+1$ на интервале $x\in[-2,4]$.

В качестве аппроксимирующей дуги кривой выберем модифицированную дугу кривой Безье 3-го порядка, которая описывается следующим точечным уравнением:

$$M=M_1\left[\bar{t}^3-2{,}5\bar{t}^2t+\bar{t}t^2\right]+M_2\left[9\bar{t}^2t-4{,}5\bar{t}t^2\right]+$$
$$+M_3\left[-4{,}5\bar{t}^2t+9\bar{t}t^2\right]+M_4\left[\bar{t}^2t-2{,}5\bar{t}t^2+t^3\right].$$

Выполнив покоординатный расчёт, получим следующую систему параметрических уравнений:

$$\begin{cases}x=x_1\left[\bar{t}^3-2{,}5\bar{t}^2t+\bar{t}t^2\right]+x_2\left[9\bar{t}^2t-4{,}5\bar{t}t^2\right]+\\ +x_3\left[-4{,}5\bar{t}^2t+9\bar{t}t^2\right]+x_4\left[\bar{t}^2t-2{,}5\bar{t}t^2+t^3\right].\\ y=y_1\left[\bar{t}^3-2{,}5\bar{t}^2t+\bar{t}t^2\right]+y_2\left[9\bar{t}^2t-4{,}5\bar{t}t^2\right]+\\ +y_3\left[-4{,}5\bar{t}^2t+9\bar{t}t^2\right]+y_4\left[\bar{t}^2t-2{,}5\bar{t}t^2+t^3\right].\end{cases}$$

Продифференцировав полученную систему параметрических уравнений и подставив результат дифференцирования в исходное дифференциальное уравнение, получим:

$$\frac{\left[\begin{array}{l}y_1(-5.5\bar{t}^2+7\bar{t}t-t^2)+y_2(9\bar{t}^2-27\bar{t}t+4.5t^2)+\\ +y_3(-4.5\bar{t}^2+27\bar{t}t-9t^2)+y_4(\bar{t}^2-7\bar{t}t+5.5t^2)\end{array}\right]}{\left[\begin{array}{l}x_1(-5.5\bar{t}^2+7\bar{t}t-t^2)+x_2(9\bar{t}^2-27\bar{t}t+4.5t^2)+\\ +x_3(-4.5\bar{t}^2+27\bar{t}t-9t^2)+x_4(\bar{t}^2-7\bar{t}t+5.5t^2)\end{array}\right]}=$$
$$=2\left(\begin{array}{l}x_1\left[\bar{t}^3-2{,}5\bar{t}^2t+\bar{t}t^2\right]+x_2\left[9\bar{t}^2t-4{,}5\bar{t}t^2\right]+\\ +x_3\left[-4{,}5\bar{t}^2t+9\bar{t}t^2\right]+x_4\left[\bar{t}^2t-2{,}5\bar{t}t^2+t^3\right]\end{array}\right)^2+1.$$

Для определения граничных условий разобьем интервал $x\in[-2,4]$ на 3 равные части. Таким образом, получим следующие значения координаты x узловых точек: $x_1=-2$, $x_2=0$, $x_3=2$ и $x_4=4$. В данном случае было использовано равномерное распределение точек, но на самом деле это совсем не обязательно. Просто в основу точечного уравнения дуги кривой Безье 3-го порядка заложено простое отношение трёх точек, которое является инвариантом покоординатного расчёта (отличительная особенность БН-исчисления). Поэтому в случае равномерного распределения точек кубическое уравнение преобразуется в уравнение прямой: $x=6t-2$, из которого легко выделить значение параметра t:

$$t=\frac{x+2}{6}.$$

Далее воспользуемся значениями параметров, заложенными в уравнение аппроксимации.

При $t=0$, получим: $-5.5y_1+9y_2-4.5y_3+y_4=54$.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

При $t=\frac{1}{3}$, получим: $-y_1-1.5y_2+3y_3-0.5y_4=6.$

При $t=\frac{2}{3}$, получим: $0.5y_1-3y_2+1.5y_3+y_4=54.$

При $t=1$, получим: $-y_1+4.5y_2-9y_3+5.5y_4=198.$

В результате получаем следующую СЛАУ:

$$\begin{cases}-5.5y_1+9y_2-4.5y_3+y_4=54.\\-y_1-1.5y_2+3y_3-0.5y_4=6.\\0.5y_1-3y_2+1.5y_3+y_4=54.\\-y_1+4.5y_2-9y_3+5.5y_4=198.\end{cases}$$

Проанализируем полученную СЛАУ. Главный определитель системы $\Delta=0$, т.е. система имеет множество решений. Однако не всё так однозначно. Дело в том, что все вычисления были выполнены в программном пакете Maple с высокой степенью точности, в СЛАУ же попали округлённые значения коэффициентов уравнений, поэтому они и дают значение определителя $\Delta=0$. На самом деле, без учёта округлений, $\Delta\neq 0$ и система имеет единственное решение:

$$y_1=-28{,}8;\ y_2=-21{,}47;\ y_3=-14{,}13;\ y_4=25.2.$$

Используя вычисленные значения координаты y узловых точек, получим следующее параметрическое уравнение:

$$y=144t^3-144t^2+54t-28{,}8.$$

С учётом полученной ранее зависимости параметра t от переменной x, получим:

$$y=0{,}(6)x^3+x-21{,}467\approx\frac{2}{3}x^3+x-21{,}467.$$

Решим теперь исходное дифференциальное уравнение традиционным способом:

$$dy=\left(2x^2+1\right)dx\Rightarrow y=\int\left(2x^2+1\right)dx\Rightarrow y=\frac{2}{3}x^3+x+C.$$

Таким образом, путём вычислений было получено частное решение дифференциального уравнения, для которого константа $C=-21{,}467$.

Возникает вопрос: зачем решать таким громоздким способом, если можно легко воспользоваться таблицей первообразных? Дело в том, что в данном случае был приведен простой для понимания пример и выполнена проверка предложенного метода на работоспособность, которую всегда легче проверять на простых задачах. Используя предложенный в работе алгоритм, обобщённый на многомерное пространство, можно вычислительно решать сложные дифференциальные уравнения, для которых решение не выражается в элементарных функциях.

6. Заключение

На данном этапе разработаны теоретические основы и принципиальный вычислительный алгоритм модификации классического метода конечных элементов и его использования для решения дифференциальных уравнений на основе методов геометрического моделирования. При этом можно выделить следующие преимущества предложенной модификации:

1) Криволинейные геометрические объекты многомерного аффинного пространства являются одновременным носителем нескольких точек, что значительно уменьшает «кусочность» итоговой функции и количество необходимых вычислений.
2) Модифицированный метод конечных элементов может быть использован на любой сети точек, сохраняя при этом криволинейную составляющую, что обеспечивает учёт геометрической нелинейности при решении задач на прочность и устойчивость. При этом сохраняется возможность геометрического моделирования физической, конструктивной, генетической и другой нелинейности.
3) Благодаря инвариантным свойствам параметра точечных уравнений отсутствует необходимость в использовании направляющих косинусов.
4) Результат моделирования представлен не дискретным набором точек, для которых вычислены значения функции отклика удовлетворяющие исходному дифференциальному уравнению, а непрерывным геометрическим объектом многомерного аффинного пространства. Такой подход исключает необходимость проведения интерполяции для вычисления значений функции отклика в промежуточных точках.
5) На основе метода многомерной интерполяции пространства выполнено обобщение вычислительного способа решения дифференциальных уравнений с частными производными как в сторону увеличения количества переменных, так и в сторону увеличения порядка производных. Причём количество переменных ограничивается только размерностью пространства, а порядок производных – степенью алгебраической кривой, которая является исходной составляющей аппроксимируемого геометрического объекта.

В перспективе планируется выпуск цикла статей с подробным решением дифференциальных уравнений математической физики на основе модифицированного метода конечных элементов.

7. Литература

[1] Балюба, И. Г. Конструктивная геометрия многообразий в точечном исчислении [Текст]: дис. … доктора техн. наук: 05.01.01 / Балюба И. Г. – Макеевка, 1995. – 227 с.

[2] Балюба, И. Г. Точечное исчисление [Текст]: [учебное пособие] / И.Г. Балюба, В.М. Найдыш; под ред. В.М. Верещаги. – Мелитополь: МГПУ им. Б. Хмельницкого, 2015. – 236 с.

[3] Бумага, А. И. Геометрическое моделирование физико-механических свойств композиционных строительных материалов в БН-исчислении [Текст]: дис. … канд. техн. наук: 05.23.05, 05.01.01 / Бумага А. И. – Макеевка, 2016. – 164 с.

[4] Владимиров, В.С. Что такое математическая физика? [Текст] / В.С. Владимиров // Препринт, Математический институт им. В. А. Стеклова РАН. – М.: МИАН, 2006. – 20 с.

[5] Галлагер, Р. Метод конечных элементов. Основы [Текст] / Р. Галлагер // Пер. с англ. – М.: Мир, 1984. – 428 с.

[6] Найдыш, В. М. Алгебра БН-исчисления [Текст] / В.М. Найдыш, И.Г. Балюба, В.М. Верещага // Прикладна геометрія та інженерна графіка: Міжвідомчий науково-технічний збірник. – К.: КНУБА, 2012. – Вип. 90. – С. 210-215.

[7] Ортега, Дж. Введение в численные методы решения дифференциальных уравнений [Текст] / Дж. Ортега, У. Пул // Пер. с англ.; под ред. А.А. Абрамова. – М.: Наука, 1986. – 288 с.

[8] Смирнов, В.И. Курс высшей математики [Текст] / В.И. Смирнов – М.: Наука, 1974. – Т. 2. – 656 с.

Об авторах

Конопацкий Евгений Викторович, к.т.н., доцент кафедры специализированных информационных технологий и систем, строительного факультета Донбасской национальной академии строительства и архитектуры. Его e-mail e.v.konopatskiy@mail.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Классификация букв в изображениях на основе медиального представления

А.Л. Липкина[1], Л.М. Местецкий[1]
lipkina96@mail.ru|mestlm@mail.ru
[1]Московский государственный университет имени М.В. Ломоносова, Москва, Россия

В статье предлагается математическая модель понятия графемы для символов кириллических шрифтов и метод построения графем на основе непрерывного медиального представления букв в цифровых изображениях. Также предлагается метод распознавания изображения печатного текста на базе математической модели графемы, используемой при генерации признаков и для построения классификатора. Приводятся результаты экспериментов, подтверждающих работоспособность графемного подхода, высокое качество распознавания текста в различных вариантах шрифта и в различных качествах изображения текста.

Ключевые слова: *цифровое изображение текста, цифровой шрифт, графема, медиальное представление, агрегированный скелетный граф*

Letters Classification in Images Based on Medial Representation

A.L. Lipkina[1], L.M. Mestetskiy[1]
lipkina96@mail.ru|mestlm@mail.ru
[1]Lomonosov Moscow State University, Moscow, Russia

In the article we propose a mathematical model of grapheme for Cyrillic characters and a method of constructing graphemes based on the continuous medial representation of letters in digital images. We also suggest the recognition method of the printed text image on the basis of mathematical model of the grapheme used at generation of features and for classifier construction. The results of experiments confirming the efficiency of the grapheme approach, high quality of text recognition in different font variants and in different qualities of the text image are presented.

Keywords: *digital text image, digital font, grapheme, medial representation, aggregated skeleton graph*

1. Введение

Когда грамотный человек читает текст, он по виду символа может сразу определить, какую букву этот символ изображает. Причём может это сделать вне зависимости от различных вариантов художественного начертания символа (с засечками, наклонного, прямого, декоративного и т. д. [7]). То есть существует некий «образ» буквы, который легко узнаваем человеком и легко отличим от «образов» других букв. Такой «образ» называется *графемой* [11].

Определение 1.1. *Буква* — отдельный символ алфавита.

В процессе развития письменности и скорописи [8, 12] у букв появилось несколько начертаний: строчное и прописное, а позднее — различные варианты написания одной и той же буквы. Зачастую эти написания могут довольно сильно различаться, хотя и обозначают они произношение одного и того же звука, например: **А** и **а**. Для описания таких различий вводится понятие графемы:

Определение 1.2. *Графема* — единица письменности, некоторый графический примитив, имеющий вид геометрического графа и изображающий канонические начертания буквы.

Графему можно представлять себе как изображения буквы в тонком шрифте, например, как на рис. 1.

Рис. 1. Изображения букв в шрифте Lato и выделенные вершины геометрических графов

Графемы должны обладать следующими свойствами:

1. Любые две графемы хорошо отличимы между собой.
2. Пусть изображения I_1 и I_2 представляют одну и ту же графему. Тогда различие между I_1 и I_2 *несущественное*. Таким образом, похожесть определяется некоторой *мерой сходства* между I_1 и I_2.

Рис. 2. Различные варианты графемы прописной буквы «А»

Понятие графемы вводится дизайнерами и шрифтовиками в словесной форме, через «общую» конструкцию начертаний графем. В настоящей статье предлагается математическое описание этой «общей» конструкции и проверяется гипотеза о том, что такого опи-

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

сания достаточно для распознавания букв в большинстве шрифтов.

2. Структура алгоритма классификации букв

Алгоритм фактически состоит из двух частей:

1. Построение математической модели графемы.
2. Разработка алгоритма на основе построенной модели.

Определим несколько базовых понятий.

Определение 2.1. *Фигура* — множество точек на плоскости.

Определение 2.2. *Пустой круг фигуры* — круг, целиком лежащий в фигуре.

Определение 2.3. *Вписанный пустой круг фигуры* — пустой круг фигуры, не содержащийся ни в каком другом пустом круге фигуры.

Определение 2.4. *Скелетное представление фигуры* — множество центров всех вписанных пустых кругов фигуры (см. рис. 3).

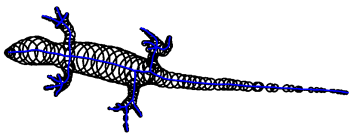


Рис. 3. Скелетное представление фигуры

В действительности скелетное представление фигуры представляет из себя граф $\mathcal{S}$, называемый *скелетом (скелетным графом)* фигуры. Вершинами графа являются центры вписанных пустых кругов, имеющих либо одну, либо три общих точки с границей фигуры, а рёбрами — линии из центров вписанных пустых кругов, касающихся границы ровно в 2 точках. Более подробно скелетное представление фигуры рассматривается в [10].

Определение 2.5. *Силуэт скелетного графа* — фигура, состоящая из объединения всех вписанных пустых кругов, центры которых лежат в скелетном графе $\mathcal{S}$. Обозначение: $\mathcal{V}_{\mathcal{S}}$.

Определение 2.6. *Стрижка скелетного графа (с параметром α)* — процесс регуляризации скелетного графа $\mathcal{S}$, основанный на удалении несущественных рёбер из скелетного графа (см. рис. 4). В процессе такого удаления возникает минимальный подграф $\mathcal{S}'$ исходного скелетного графа, для которого выполняется $H(\mathcal{V}_{\mathcal{S}}, \mathcal{V}_{\mathcal{S}'}) \leqslant \alpha$, где $H(\mathcal{V}_{\mathcal{S}}, \mathcal{V}_{\mathcal{S}'})$ — хаусдорфово расстояние [2] между силуэтом скелетного графа $\mathcal{S}$ и силуэтом скелетного графа $\mathcal{S}'$.

Рис. 4. Пример скелета без стрижки (слева) и со стрижкой (справа)

3. Построение математической модели графемы

Для построения математической модели графемы предлагается сделать два шага:

1. Сегментация изображения текста на изображения отдельных символов (графем).
2. Выделение структурного описания (математической модели) изображения каждой графемы.

Второй шаг разделяется на следующие действия:

1. Получение скелетного графа из изображения буквы.
2. Агрегирование скелетного графа.
3. Обработка скелетного графа, а именно удаление шумовых рёбер.

Получение скелетного графа

Построение скелетного графа фигуры подробно описывается в [10]. После его построения производится последующая стрижка с параметром α.

Агрегирование скелетного графа

Полученный скелетный граф содержит только следующие типы вершин: вершины степени 1 (листья), вершины степени 2, вершины степени 3 (развилки).

Основную информацию о скелетном графе несут листья и развилки, а также виды связей между ними. Чтобы выделить эти связи, производится операция *агрегирования скелетного графа:* «склеивание» в одну цепь всех таких последовательных рёбер, инцидентные вершины которых имеют степень либо 1, либо 2. После такой «склейки» в качестве вершин остаются только листья и развилки (см. рис. 5).

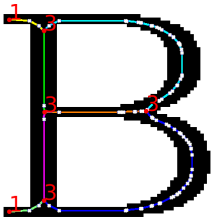


Рис. 5. Пример агрегированного скелетного графа. Белым цветом отмечены вершины степени 2 исходного графа.

После агрегирования скелетный граф представляет собой гиперграф $\mathcal{S}_{agg,1}$, вершинами которого являются листы и развилки, а рёбрами — выделенные в процессе цепочки.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Обозначения и понятия

1. Рассматривается входное бинарное изображение символа. $\mathcal{B}$ — минимальная по площади прямоугольная рамка с горизонтальными и вертикальными сторонами, ограничивающая данный символ. $\mathcal{B}_H$ и $\mathcal{B}_W$ — высота и ширина рамки $\mathcal{B}$ соответственно.
2. Пусть e — ребро неагрегированного скелетного графа $\mathcal{S}$. $v_1(e), v_2(e)$ — концевые вершины этого ребра без учёта какого-либо порядка.
3. $l(e)$ — длина ребра e. Она вычисляется через евклидово расстояние между двумя точками $v_1(e)$ и $v_2(e)$:
$$l(e) = \sqrt{\big(v_1(e)_x - v_2(e)_x\big)^2 + \big(v_1(e)_y - v_2(e)_y\big)^2}.$$
4. Для ребра (цепочки) e_{agg} гиперграфа $\mathcal{S}_{agg}$ через $v_1(e), v_2(e)$ обозначаются концевые вершины этой цепочки.
5. Ребро e_{agg} обозначенного гиперграфа $\mathcal{S}_{agg}$ состоит из n последовательных рёбер исходного графа $\mathcal{S}$, попавших в данную цепочку e_{agg}: $\{e^1_{agg}, e^2_{agg}, \dots, e^n_{agg}\}$.
6. $l(e_{agg})$ — длина цепочки e_{agg}. Она считается как сумма длин всех входящих в данную цепочку рёбер:
$$l(e_{agg}) = \sum_{i=1}^{n} l(e^i_{agg}).$$
7. $\deg v$ — степень вершины v.

Определение 3.1. Пусть дан отрезок $d = [v_1(e_{agg}), v_2(e_{agg})]$. Среди всех вершин цепочки e_{agg} находится вершина v_h, наиболее удалённая от отрезка d. По трём точкам $v_1(e_{agg}), v_2(e_{agg}), v_h$ строится окружность. Тогда *аппроксимирующей дугой* называется дуга наименьшей длины, ограниченная точками $v_1(e_{agg}), v_2(e_{agg})$ (см. рис. 6).

Определение 3.2. *Угол кривизны цепочки* — центральный угол её аппроксимирующей дуги.

Замечание. В случае, когда три точки $v_1(e_{agg}), v_2(e_{agg}), v_h$ лежат на одной прямой или когда в цепочке e_{agg} нет вершин, угол кривизны цепочки полагается равным 0.

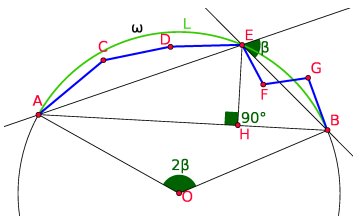
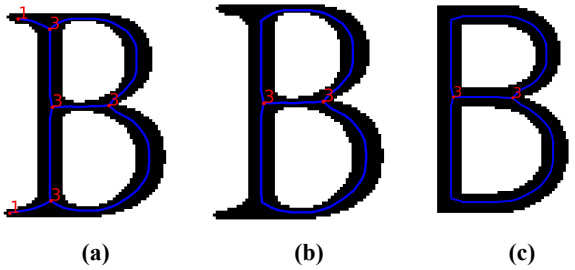


Рис. 6. Пример аппроксимирующей цепочку $[A, C, D, E, F, G, B]$ дуги L и центрального угла BOA

Удаление шумовых рёбер

После регуляризации и агрегации скелетного графа и в $\mathcal{S}_{agg,1}$ все ещё могут содержаться шумовые рёбра. Это проявляется на буквах, изображённых в шрифтах с засечками [7].

Засечки это некое украшение для буквы, и их наличие или отсутствие не мешает человеку распознать, какая именно буква изображена. Таким образом в модель графемы буквы засечки не должны входить. Поэтому следующий этап построения математической модели графемы — удаление из $\mathcal{S}_{agg,1}$ рёбер, являющихся засечками (см. рис. 7). Множество рёбер гиперграфа $\mathcal{S}_{agg,1}$, являющихся засечками, обозначим через $\mathcal{E}_{\mathcal{S}}$.



(a) **(b)** **(c)**

Рис. 7. 7a: скелет буквы в шрифте с засечками; 7b: тот же скелет с удалёнными рёбрами из $\mathcal{E}_{\mathcal{S}}$; 7c: скелет буквы в шрифте без засечек.

Для множества $\mathcal{E}_{\mathcal{S}}$ можно выделить следующие признаки:

1. $|\mathcal{E}_{\mathcal{S}}| \geqslant 2$, то есть если засечки в скелетном графе присутствуют, то их не меньше двух.
2. $\forall\, e_{agg} \in \mathcal{E}_{\mathcal{S}}$ характерны следующие признаки:
   - ровно одна из вершин $\{v_1(e_{agg}), v_2(e_{agg})\}$ является листовой, и ровно одна из них является развилкой;
   - длина ребра $l(e_{agg})$ не превышает некоторого порога $\mathcal{L}(\mathcal{B})$;
   - центральный угол 2β аппроксимирующей e_{agg} дуги не меньше некоторого порога $\mathcal{A}$.

Алгоритм удаления шумовых рёбер из $\mathcal{S}_{agg,1}$:

1. Определение множества $\mathcal{E}_{\mathcal{S}}$ на основе его признаков.
2. Удаление из агрегированного скелетного графа $\mathcal{S}_{agg,1}$ всех рёбер из $\mathcal{E}_{\mathcal{S}}$.

Гиперграф, полученный после удаления рёбер и повторной агрегации, обозначим через $\mathcal{S}_{agg,2}$. Он и является предлагаемой математической моделью графемы.

4. Распознавание графем

На этом этапе из $S_{agg,2}$ будут выделяться признаки для последующего построения классификатора графем.

Генерация признаков

В данном методе предлагается выделять 2 типа признаковых описаний: *признаки верхнего уровня* $\mathcal{F}_a$ и *признаки нижнего уровня* $\mathcal{F}_d$. Они обладают следующими свойствами:

- Если из гиперграфов $\mathcal{S}'_{agg,2}, \mathcal{S}''_{agg,2}$ выделяются одинаковые признаки верхнего уровня $\mathcal{F}'_a = \mathcal{F}''_a$,

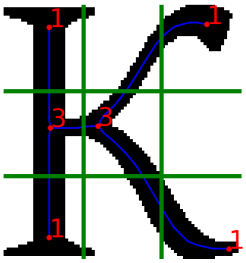
The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

то признаки нижнего уровня $\mathcal{F}'_d$ и $\mathcal{F}''_d$ лежат в одном признаковом пространстве.

— Если из гиперграфов $\mathcal{S}'_{agg,2}, \mathcal{S}''_{agg,2}$ выделяются различные признаки верхнего уровня $\mathcal{F}'_a \neq \mathcal{F}''_a$, то признаки нижнего уровня $\mathcal{F}'_d$ и $\mathcal{F}''_d$ лежат в разных признаковых пространствах.

Признаки верхнего уровня

Рамка $\mathcal{B}$, в которую заключена графема, разбивается на n равных частей горизонтальными прямыми и на m равных частей — вертикальными. В каждом из получившихся $n \cdot m$ прямоугольников считается количество листьев и количество развилок, и эти количества добавляются в признаковое описание верхнего уровня. Помимо этого в качестве признака верхнего уровня рассматривается количество компонент связности графемы внутри рамки $\mathcal{B}$ (см. рис. 8).



(a)

1	0	1
0	0	0
1	0	1

(b)

0	0	0
1	1	0
0	0	0

(c)

Рис. 8. 8a: скелет $\mathcal{S}_{agg,2}$ буквы «K» и разбиение рамки на 9 прямоугольников ($n = m = 3$); 8b: количества листьев в каждом из прямоугольников; 8c: количества развилок в каждом из прямоугольников.

Признаки нижнего уровня

Считаем, что признаковое описание верхнего уровня $\mathcal{F}_a$ зафиксировано. Это означает, что фактически зафиксировано строение гиперграфа $\mathcal{S}_{agg,2}$: для каждого из $n \cdot m$ прямоугольников разбиения известно количество листьев и развилок, попавших в него, и известно количество рёбер гиперграфа, связанных с каждым прямоугольником разбиения. Таким образом, теперь появляется возможность генерировать фиксированное число признаков для каждого из $n \cdot m$ прямоугольников.

Признаковое описание нижнего уровня $\mathcal{F}_d$ предлагается генерировать из структуры рёбер $\mathcal{S}_{agg,2}$.

Генерация признаков из ребра

Пусть $[A, B]$ — ребро гиперграфа $\mathcal{S}_{agg,2}$. Фиксируется маска разбиения этого ребра на k частей:

$$\mathcal{Z}_k = [z_1, z_2, \ldots, z_k], \quad z_j \in (0, 1) \ \ \forall j = \overline{1, k}.$$

Фиксируется стартовая вершина (не ограничивая общности считаем, что это A). Применяем разбиение $\mathcal{Z}_k$ к ребру $[A, B]$, начиная от вершины A, следующим образом: ребро $[A, B]$ разбивается k точками, считая от точки A, на $k + 1$ отрезков s_i так, что:

$$\sum_{i=1}^{j} l(s_i) = z_j l([A, B]) \ \ \forall j = \overline{1, k}.$$

Пусть концы отрезков s_i имеют координаты C_{i-1}, C_i:

$$s_i = [C_{i-1}, C_i] \ \ \forall i = \overline{1, k+1}.$$

Заметим, что $C_0 = A$ и $C_{k+1} = B$. Также отдельно обозначим $\overrightarrow{b} = \overrightarrow{AC_1}$.

Пример разбиения можно увидеть на рис. 9.

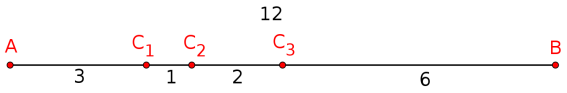


Рис. 9. Пример разбитого маской $\mathcal{Z}_3 = \left[\frac{1}{4}, \frac{1}{3}, \frac{1}{2}\right]$ ребра $[A, B]$ длины 12 со стартовой вершиной A

Предлагается выделять следующие признаки нижнего уровня:

1. Рассматриваются вектора $\overrightarrow{AC_i}, \ i = \overline{1, k+1}$. Пусть $m_i = ||\overrightarrow{AC_i}||_2, \ i = \overline{1, k+1}$. Эти вектора нормируются на свои длины соответственно:
$$\overrightarrow{c_i} = \frac{\overrightarrow{AC_i}}{m_i}, \ i = \overline{1, k+1}.$$
В качестве признаков берутся последовательно (по i) координаты получившихся векторов $\overrightarrow{c_i}, \ i = \overline{1, k+1}$.
2. Пусть $\overrightarrow{g} = (1, 0)$ (см. рис. 10). В качестве признаков добавляются последовательно (по i) следующие ориентированные углы:
$$\angle(\overrightarrow{g}, \overrightarrow{AC_i}), \ i = \overline{1, k+1}.$$
3. Следующие ориентированные углы последовательно (по i):
$$\angle(\overrightarrow{C_iC_{i-1}}, \overrightarrow{C_iC_{i+1}}), \ i = \overline{1, k}.$$
4. Отношения длин соседних векторов:
$$\frac{m_i}{m_{i-1}}, \ i = \overline{2, k+1}.$$

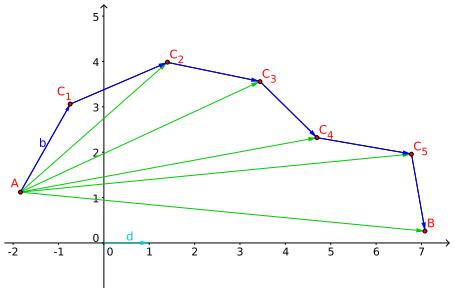


Рис. 10. Пример разбитого ребра $[A, B]$ со стартовой вершиной A

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Генерация признаков для одного прямоугольника
Пусть $\mathcal{R}$ — текущая рассматриваемая прямоугольная область из разбиения рамки $\mathcal{B}$.

Сначала рассмотрим все листья, затем все развилки. Во всех случаях стартовой вершиной будет являться текущая рассматриваемая вершина v.

1. v — лист. Тогда генерируются признаки f_e для соответствующего ей ребра e, и они добавляются к итоговому признаковому описанию $f_{\mathcal{R}}$.
2. v — развилка. Рассмотрим соответствующие трём выходящим рёбрам вектора $\overrightarrow{b_1}, \overrightarrow{b_2}, \overrightarrow{b_3}$. Выходящие рёбра сортируются в порядке возрастания ориентированных углов $\angle(\overrightarrow{b_i}, \overrightarrow{g})$, $i = 1, 2, 3$, после чего для них генерируются признаки f_e. Полученные признаки добавляются в порядке сортировки рёбер к итоговому признаковому описанию $f_{\mathcal{R}}$.

Генерация признаков для графемы

Признаки нижнего уровня F_d для графемы получаются путём объединения признаков $f_{\mathcal{R}}$ в порядке упорядочивания прямоугольных областей $\mathcal{R}$.

Обучение классификатора

Теперь в рамках каждого признака верхнего уровня F_a имеется возможность обучить свой классификатор — каждый на своём признаковом пространстве, соответствующем своему признаковому пространству нижнего уровня F_d.

На этапе обучения берётся размеченная тренировочная выборка $(\mathbb{X}_{tr}, \mathbb{Y}_{tr})$, где $x \in \mathbb{X}_{tr}$ — бинаризованное изображение символа, $y \in \mathbb{Y}_{tr}$ — соответствующий класс изображения.

Алгоритм обучения состоит из следующих шагов:

1. По всей тренировочной выборке $(\mathbb{X}_{tr}, \mathbb{Y}_{tr})$ выделить признаки верхнего уровня и по ним построить словарь классификации $\mathcal{D}$.
2. Для каждого уникального признака верхнего уровня $\mathcal{F}_a$ выделить те объекты, которые имеют данный признак $\mathcal{F}_a$. Для каждого из таких объектов построить признак нижнего уровня $\mathcal{F}_d$. В итоге получается новая подвыборка объектов из признакового пространства $\mathcal{F}_d$, при помощи которой обучается классификатор (см. рис 11).

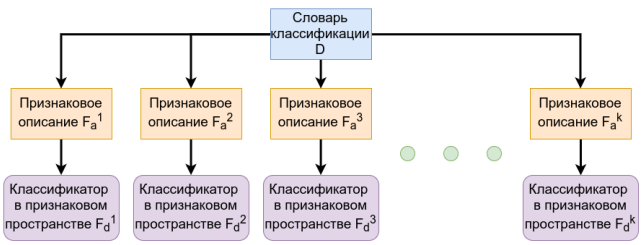


Рис. 11. Устройство словаря классификации $\mathcal{D}$

Алгоритм классификации

Пусть на вход приходит новый объект x (бинаризованное изображение одного символа), и его необходимо классифицировать. Для его классификации необходимо проделать следующие шаги:

1. Выделение математической модели графемы $\mathcal{S}_{agg,2}$ из x.
2. Построение признаков верхнего уровня $\mathcal{F}_a$ из $\mathcal{S}_{agg,2}$.
3. Проверка наличия признака $\mathcal{F}_a$ в словаре классификации $\mathcal{D}$, полученном на этапе обучения. Если признак не присутствует, то производится операция пост-обработки скелетного графа. Если присутствует — переход к следующему шагу.
4. Построение признаков нижнего уровня $\mathcal{F}_d$, применение к нему соответствующего обученного классификатора и получение ответа.

Идея пост-обработки заключается в следующем: продолжение поиска подграфа, который возможно будет классифицировать по обученному словарю классификации $\mathcal{D}$. Если же такой граф не был найден, то будет возвращен отказ от классификации.

Итоговый алгоритм можно увидеть на рис. 12.

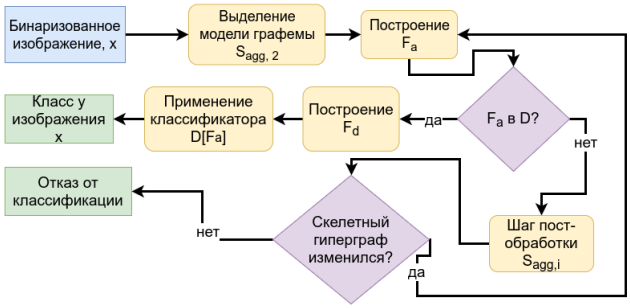


Рис. 12. Алгоритм классификации бинаризованного изображения

Метрика качества

В качестве метрики качества используется *точность классификации (accuracy)*.

5. Вычислительные эксперименты

Обучающая выборка

В качестве обучающей выборки было выбрано 88 различных шрифтов, из каждого сгенерировано по 33 буквы русского алфавита в строчном и прописном вариантах (то есть всего 66 графем) в трёх размерах шрифта: 30, 50, 100 пикселей. Генерация изображений производилась без сглаживания, то есть сразу в бинарном формате. Размер тренировочной выборки $(\mathbb{X}_{tr}, \mathbb{Y}_{tr})$ составляет $n_{tr} = 17424$ бинаризованных изображений букв. В качестве истинного класса $\mathbb{Y}^i_{tr}$ для объекта $\mathbb{X}^i_{tr}$ обучающей выборки бралась изображенная буква *в нижнем регистре*.

Параметры предложенного алгоритма

1. Параметр стрижки $\alpha = 0.06 \cdot \mathcal{B}_H$.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2. Порог для отсечения по длине:
$$\mathcal{L}(\mathcal{B}) = \frac{2}{7}\max(\mathcal{B}_H, \mathcal{B}_W).$$
3. Порог для отсечения по длине на стадии постобработки увеличивается в 1.8 раза:
$$\mathcal{L}'(\mathcal{B}) = 1.8 \cdot \mathcal{L}(\mathcal{B}) \quad (\kappa = 1.8).$$
4. Порог для отсечения по центральному углу:
$$\mathcal{A} = \frac{\pi}{5}.$$
5. На стадии извлечения признаков верхнего уровня полагается $n = m = 3$.
6. Фиксируемая сетка полагается равной:
$$\mathcal{Z}_8 = \left[\frac{1}{50}, \frac{1}{5}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{4}{5}\right].$$
7. В качестве классификаторов нижних уровней рассматривается *случайный лес (Random forest)* [3].

Базовый алгоритм

В качестве базового алгоритма была выбрана свёрточная нейронная сеть [6][1], архитектура которой представлена на рис. 13:

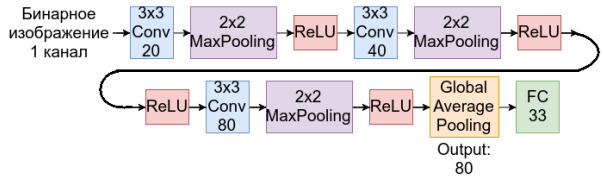


Рис. 13. Архитектура нейронной сети

Из соображений решения задачи классификации над выходным слоем $(x^1, x^2, \ldots, x^{33})$ из 33 нейронов производится softmax-активация.

В качестве оптимизируемой функции потерь берётся кросс-энтропия.

Эксперимент 1

В качестве тестовой выборки берутся те же самые 88 шрифтов, которые использовались при обучении, но другого размера шрифта, который равен 80 пикселям. Таким образом, $n_{te} = 5800$. Изображения букв генерируются при помощи программы, то есть изображения качественные, без шумов и бинаризованные. Результаты работы двух методов (структурный анализ — описываемый в статье метод распознавания) представлены в таблице 1:

	Структурный анализ	**Нейронная сеть**
Качество, Q	0.99689	**0.99862**
Доля отказов от классификации	0.00086	0

Таблица 1. Результаты работы двух методов

Эксперимент 2

В качестве тестовой выборки берутся 50 шрифтов, которые *не использовались* при обучении [5]. Размер шрифта равен 80 пикселям, $n_{te} = 3300$. Изображения букв генерируются при помощи программы. Результаты работы двух методов представлены в таблице 2:

	Структурный анализ	**Нейронная сеть**
Качество, Q	**0.97**	0.96515
Доля отказов от классификации	0.01364	0

Таблица 2. Результаты работы двух методов

Эксперимент 3

В качестве тестовой выборки берутся те же самые 50 шрифтов, что и в предыдущем эксперименте 5, и того же размера. Сначала генерируется документ (.doc) со всеми буквами из тестовой выборки, затем этот документ конвертируется в изображение в формате png с разрешением 300 dpi (см.рис. 14). Сначала изображения из цветного представления RGB переводились в серые тона Y по формуле:
$$Y = 0.299R + 0.587G + 0.114B.$$
Затем изображения были бинаризованы при помощи метода Оцу [4].

Рис. 14. Пример буквы из входного изображения

Результаты работы двух методов представлены в таблице 3:

	Структурный анализ	**Нейронная сеть**
Качество, Q	**0.94818**	0.94454
Доля отказов от классификации	0.01485	0

Таблица 3. Результаты работы двух методов

Эксперимент 4

В этом эксперименте в качестве тестовой выборки берутся 18 выборочных шрифтов из 50 шрифтов эксперимента 5. Размер шрифтов полагается равным 80 пикселям, $n_{te} = 1188$. Генерируется документ (.doc) со всеми буквами из тестовой выборки, затем этот документ распечатывается на принтере. Далее полученные образцы сканируются с разрешением 300 dpi. То есть изображения получаются менее качественные, нежели чем в предыдущем случае (см.рис. 15).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Рис. 15. Пример буквы из входного изображения

Результаты работы двух методов представлены в таблице 4:

	Структурный анализ	**Нейронная сеть**
Качество, Q	**0.95538**	0.94696
Доля отказов от классификации	0.01263	0

Таблица 4. Результаты работы двух методов

Анализ экспериментов

Проведённые эксперименты показывают, что: по качеству предлагаемый метод не уступает выбранному базовому алгоритму и он имеет небольшую долю отказов от классификации, которая растёт с ухудшением качества изображения.

6. Выводы

В данной статье предложена формализация понятия «графема», а именно предложена математическая модель графемы. На основе этой модели предложен метод генерации признаков, использующихся для последующего построения алгоритма классификации изображений букв (то есть определена мера сходства между математическими моделями графем). По этому подходу к извлечению признаков был сделан доклад в [9]. Также в данной статье предложен алгоритм распознавания текста на изображении.

Достоинства предлагаемого метода распознавания букв: независимость от размера, типа шрифта и типа начертания буквы; выделение общей структуры (математической модели графемы) букв, которой достаточно для распознавания букв в новых шрифтах; интерпретируемость признаков.

Недостатки метода: наличие отказов от классификации и зависимость качества распознавания от качества бинаризации изображения.

Проведённые эксперименты подтверждают, что предлагаемая математическая модель графемы показала свою работоспособность.

7. Благодарности

Работа выполнена при поддержке РФФИ, грант 17-01-00917.

8. Литература

[1] Bishop C. M. Pattern recognition and machine learning. — Springer. — 2006.

[2] Felix Hausdorff. Grundzüge der mengenlehre. — 1914.

[3] Kam, H. T. Random decision forests / Ho Tin Kam // Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC. — 1995.

[4] Otsu, N. A threshold selection method from gray-level histograms / N. Otsu // IEEE Trans. Sys., Man., Cyber. — 1979.

[5] Fonts database. — https://www.fontsquirrel.com/ .

[6] Gradient-based learning applied to document recognition / LeCun Yann, Léon Bottou, Yoshua Bengio, Patrick Haffner // Proceedings of the IEEE. — 1998.

[7] ParaType. Цифровые шрифты. — ParaType. — 2008.

[8] А. А. Зализняк. Русское именное словоизменение приложением избранных работ по современному русскому языку и общему языкознанию. — М.:Языки славянской культуры. — 2002.

[9] А. Л. Липкина. Генерация структурных признаков для классификации букв в цифровых изображениях / А. Л. Липкина // Сборник тезисов XXV Международной научной конференции студентов, аспирантов и молодых ученых «ЛОМОНОСОВ-2018». — Макс-Пресс, 2018.

[10] Л. М. Местецкий. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. — М.:ФИЗМАТЛИТ. — 2009.

[11] О. В. Осетрова. СЕМИОТИКА ШРИФТА. — Вестник Воронежского государственного университета. Серия:Филология. Журналистика. — 2006.

[12] А. Соломоник. О языке и языках. — М.: Издательство «Спутник+». — 2017.

Об авторах

Липкина Анна Львовна, бакалавр кафедры математических методов прогнозирования факультета вычислительной математики и кибернетики Московского государственного университета. lipkina96@mail.ru.

Местецкий Леонид Моисеевич, д. т. н., профессор кафедры математических методов прогнозирования факультета вычислительной математики и кибернетики Московского государственного университета. mestlm@mail.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Исследование временных характеристик алгоритма восстановления каркасной модели по проекционным изображениям

Т.В. Мошкова[1], С.А. Роменский[1], С.И. Ротков[1], В.А. Тюрина[1]
ng.forever.mtv@gmail.com|romensky.serge@gmail.com|rotkov@nngasu.ru|55555_73@mail.ru
[1]ННГАСУ, Нижний Новгород, Россия

В данной статье рассматриваются экспериментальные данные вычислительной сложности алгоритма формирования трехмерной каркасной модели.

Ключевые слова: *синтез модели, пространственный объект, проекции, чертеж, каркасная модель*.

Research for time characteristics of algorithm of reconstructing wireframe model from projection images

T.V. Moshkova[1], S.A. Romensky[1], S.I. Rotkov[1], V.A. Tyurina[1]
ng.forever.mtv@gmail.com|romensky.serge@gmail.com|rotkov@nngasu.ru|55555_73@mail.ru
[1]NNGASU, Nizhny Novgorod, Russia

In this article considered experimental data for computational complexityof algorithm of reconstructing 3D wireframe model.

Keywords: *model synthesis, spatial object, projections, drawing, wireframe model*.

1. Введение

Проблема преобразования бумажных архивов чертежно-конструкторской документации является одной из ключевых задач внедрения и реализации проекта «Цифровое производство» [8]. Среди множества задач, которые направлены на решение поставленной проблемы, является задача синтеза геометрической модели пространственного объекта по его проекциям. В разное время были осуществлены попытки программной реализации указанной задачи, но только одна была доведена до уровня действующей программы [7]. Эта разработка имела ограничения на количество вершин и ребер синтезируемой модели, что обусловлено имевшимися на тот период времени возможностями как ЭВМ, так и средств разработки программ. Для решения проблемы преобразования бумажных архивов возникла необходимость полной программной переработки как алгоритма, так и его программной реализации. В работе [2] проведен детальный анализ имеющихся в мире алгоритмических и программных разработок по проблеме синтеза модели пространственного объекта по изображениям на поле многовидового технического чертежа, однако ни одна из имеющихся публикаций по этой проблеме не затрагивает вопрос о вычислительных характеристиках.

Данная статья посвящена исследованию временных характеристик алгоритма восстановления каркасной модели по проекционным изображениям, который является частью большого программного комплекса, предназначенного для формирования электронной модели изделия на основе бумажного архива чертежно-конструкторской документации [8]. Формально алгоритм восстановления каркасной модели на основе проекционных изображений описан в [1,6], в этом докладе лишь конкретизируются моменты, связанные с его программной реализацией.

Необходимые двумерные геометрические данные разработанная программная реализация получает при помощи COM-объектов из системы автоматизированного проектирования «Компас-3D». Данный способ и его реализация описаны в [4,3,5].

2. Описание алгоритма формирования трехмерной каркасной модели и доказательство его вычислительной сложности.

Первый этап формирования трехмерной каркасной модели – получение двумерной геометро-графической информации прикладной программой. Учитывая, что способ получения двумерных данных в каждой конкретной программной реализации может отличаться, при определении временных характеристик не будем рассматривать данный этап. В основном получаемая двумерная информация имеет свою локальную систему координат (ЛСК), не связанную с будущей трехмерной моделью. Чтобы иметь возможность построения каркасной модели, необходимо преобразовать имеющуюся систему координат проекционных изображений в эпюр Монжа, содержащий информацию уже о трех (x,y,z) координатах. Укрупненный алгоритм данного преобразования координат представлен на рис. 1

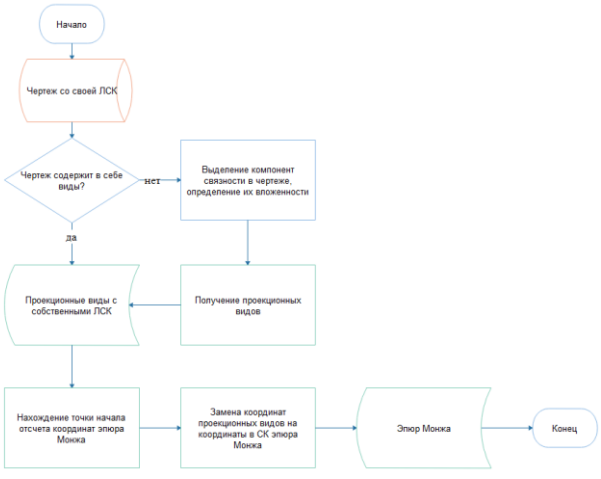


Рис. 1. Блок-схема алгоритма получения геометро-графической информации на основе эпюра Монжа.

Вычислительная сложность данного этапа напрямую зависит от способа хранения геометро-графической информации внутри прикладной программы. В экспериментальной программной реализации все вершины на проекционных изображениях хранятся в виде односвязного списка. В среднем, на каждом проекционном изображении находится равное количество точек, то есть

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

при обработке только одно проекционного изображения используется треть от общего числа вершин.

В реализованном алгоритме этап получения геометро-графической информации из внешнего источника представляет собой чтение координат всех вершин на проекционных изображениях и их преобразование в систему координат эпюра Монжа, что соответствует вычислительной сложности O(m), где m – количество вершин на проекционных изображениях.

Вторым этапом получения трехмерной каркасной модели является построение точечной трехмерной модели, которая представляет собой облако точек, никак не связанных между собой. Каждая вершина данной модели получается следующим образом:

1. на одном из проекционных изображений рассматривается какая-либо точка, например, на виде спереди, получая таким образом сведения о координатах (x_1, z_1) будущей трехмерной вершины;
2. следует найти все пары подходящих по одной из координат (x_1 или z_1 в нашем примере) вершин на другом проекционном изображении. Например, на виде сверху есть две подходящие вершины – (x_1, y_1) и (x_1, y_2);
3. на третьем проекционном изображении необходимо найти точки с координатами из двух предыдущих шагов. В нашем примере это вершины с координатами (y_1, z_1) и (y_2, z_1);
4. если данные вершины существуют на проекционном изображении, то данные трехмерные точки есть и в трехмерной точечной модели. В примере это точки с координатами (x_1, y_1, z_1) и (x_1, y_2, z_1).

Учитывая, что поиск по списку имеет сложность O(l), где l – длина списка, вычислительная сложность этапа формирования трехмерной точечной модели определяется следующим выражением: $O(m/3)*(O(m/3)+O(m/3))=2*O(m/3)^2=O(m^2/9)=O(m^2)$, где m - количество вершин на проекционных изображениях.

На третьем этапе путем добавления ребер в точечную модель получается искомая каркасная трехмерная модель. Ребра добавляются, исходя из следующего утверждения: если для точек (x_1, y_1, z_1), (x_2, y_2, z_2) верно что на соответствующих проекционных изображениях существуют ребра (x_1, z_1)-(x_2, z_2); (x_1, y_1)-(x_2, y_2); (y_1, z_1)-(y_2, z_2) или соответствующие точки совпадают, то, скорее всего, существует ребро (x_1, y_1, z_1)-(x_2, y_2, z_2).

Решение этой задачи возможно двумя способами.

Первый способ формирования списка ребер трехмерной каркасной модели состоит в поиске ребер между всеми вершинами:

1. для каждой вершины взять все последующие за ней в списке трехмерной точечной модели;
2. для каждой пары вершин проверить связь на всех трех проекциях, то есть выполнить поиск на всех трех проекционных изображениях проекций обеих точек и сверить списки их ребер.

Вычислительная сложность первого этапа – $O(n^2)$, сложность второго этапа – 3*O(m), таким образом сложность данного способа формирования списка ребер каркасной модели – $O(n^2)*3*O(m)=O(n^2*m)$, где n – количество вершин в трехмерной точечной модели

Второй способ формирования списка ребер трехмерной каркасной модели, использованный в реализованном алгоритме, состоит в реализации следующих шагов:

1. найти проекцию на две координатные плоскости для каждой вершины из списка точечной модели – сложность этапа 2*O(m);
2. для каждой проекции, используя список связанных вершин, найти восстановленные в точечной модели вершины – сложность этапа O(k*n) (k – среднее количество ребер, исходящих из вершины);
3. проверить наличие связи на трёх проекционных изображениях начальной вершины и каждой из вершин, полученных на втором этапе при помощи сверки списка проекций ребер – сложность этапа 3*O(m).

Таким образом получаем, что вычислительная сложность данного способа формирования списка ребер трехмерной каркасной модели равна $2*O(m)*O(k*n)*3*O(m) = 6*O(m^2*k*n) = O(m^2*k*n)$.

Исходя из вышеописанных этапов формирования каркасной трёхмерной модели на основе проекционных изображения, получаем следующую вычислительную сложность реализованного алгоритма:

$O(m)+O(m^2)+O(m^2*k*n) = O(m^2*k*n)$,

где m – количество точек на проекционных изображениях; k – среднее количество ребер, исходящих из одной вершины; n – количество вершин в трехмерной точечной модели.

3. Подтверждение сложности алгоритма экспериментальными данными

Первый вид эксперимента – сравнение скорости построения трехмерной модели путем ручного ввода геометро-графической информации в САПР и автоматизированного построения трехмерной каркасной модели. В среднем на создание простой трехмерной модели тратится от 2 до 5 минут. Экспериментальная программная реализация показывает результат уже через 0.5-3 секунды после запуска процесса чтения двумерной информации на простых трехмерных моделях.

Вышеизложенные результаты эксперимента позволяют говорить об ускорении процесса получения трехмерной модели по проекционным изображениям в 60-200 раз по сравнению с предлагаемой в системах геометрического моделирования и компьютерной графики технологией формирования пространственной модели из 3D примитивов с применением аппарата булевых операций и операций формообразования (выдавливания, вращения и кинематических) в зависимости от сложности восстанавливаемой модели и мощности центрального процессора. Такое ускорение процесса обработки позволит обработать большое количество архивных бумажных чертежей за достаточно приемлемое время.

Второй вид эксперимента – сравнение времени, которое тратится алгоритмом на выполнение своих различных составных частей: реализацию операций чтения чертежа, получения точечной трехмерной модели и её преобразования в каркасную.

Первая часть исследования проводилась на несложных трехмерных моделях – до 100 вершин. Эксперимент показал, что самый большой процент времени тратится на получение геометро-графической информации из внешнего источника – до 99%. После выполнения операции чтения вся двумерная геометрическая информация помещается в виртуальное адресное пространство разработанного программного комплекса и соответственно на доступ к ней тратится существенно меньше времени. На экспериментальных чертежах время восстановления точечной и каркасной моделей не превышало 1 мс.

Вторая часть исследования включала в себя более сложные модели – более 100 вершин. Данные, полученные в ходе исследования, приведены в таблице 1.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Таблица 1. Временные характеристики разработанного алгоритма

Количество вершин	Общее время восстановления, сек.	Время чтения двумерных данных, сек.	Время восстановления точечной модели, сек.	Время получения каркасной модели, сек.
844	29.859	28.125	0.5	1.234
2440	206.547	185.140	1.282	20.125
15616	424.969	95.156	14.78	315.735

Из проведенного исследования можно сделать следующие выводы:

1. на моделях до 1000 вершин около 90% времени занимает чтение геометро-графической информации из внешнего источника и лишь 10% - собственно формирование точечной и каркасной моделей;
2. экспериментальные данные показывают сложность формирования точечной трехмерной модели O(n), где n – количество вершин в обрабатываемой модели;
3. экспериментальные данные показывают сложность алгоритма формирования каркасной трехмерной модели O(k*n), где n – количество вершин в обрабатываемой модели, k-среднее количество ребер, исходящих из одной вершины.

4. Заключение

Экспериментальная программная реализация позволяет ускорить процесс создания трехмерной каркасной модели в 60-200 раз (в зависимости от сложности модели) по сравнению с созданием этой модели ручным способом в одной из САПР.

5. Благодарности

Работа выполнена в рамках гранта РФФИ № 17-07-00543.

6. Литература

[1] Котов И.И., Полозов В.С., Широкова Л.В. Алгоритмы машинной графики. М., Машиностроение, 1977 г., 270 стр.

[2] Т. В. Мошкова, С. И. Ротков, В. А. Тюрина Проблема синтеза модели 3D объекта по его проекционным изображениям. Аналитический обзор. // Журнал «Научная визуализация», 2018, т.10, №1, стр.135-156

[3] Роменский, С.А. Обеспечение информационной совместимости между различными системами автоматизированного проектирования // Сборник трудов конференции регионального значения XX Нижегородская сессия молодых ученых. Естественные, математические науки.: сб.тр. рег.конф., 19-22 мая 2015г., г. Арзамас, 2015г.

[4] Роменский, С.А. О проблеме передачи данных между пользовательскими программами в информационных технологиях //Сборник трудов аспирантов, магистрантов и соискателей. Технические науки. Науки о Земле. Экология. - Н.Новгород: ННГАСУ, 2015. С. 119-126

[5] Роменский, С.А. Передача геометро-графической информации из системы автоматизированного проектирования в прикладную программу на примере САПР “Компас-График” // Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта (CAD/CAM/PDM-2015): сб. тр.

[6] Роменский С.А., Ротков С.И., Смычёк М.М., Тюрина В.А.. Синтез 3D модели объекта по изображениям на поле чертежа // 26-я Международная конференция (GraphiCon 2016), Нижний Новгород, 19-23 сентября 2016 г.

[7] Ротков, С.И. Средства геометрического моделирования и компьютерной графики пространственных объектов для CALS-технологий: Дис. докт. техн. наук: 05.01.01 / С.И. Ротков. – Н. Новгород, 1999. – 287 с.

[8] Ротков С.И., Попов Е.В., Мошкова Т.В., Тюрина В.А., Васин Д.Ю., Роменский С.А. Макаров Н.Л., Чепкасов В.Л. Проблема преобразования бумажных архивов чертежно-конструкторской документации в электронную модель изделия и связанные с ней геометро-графические задачи // Труды 27 Международной научно-практической конференции GraphiCon 2017 (Пермь, 24-28 сентября 2017 г.) – Пермь: ПГНИУ, 2017, С. 20.

Об авторах

Мошкова Татьяна Владимировна, к.т.н., доцент, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета. E-mail: ng.forever.mtv@gmail.com.

Роменский Сергей Александрович, аспирант кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета. E-mail: romensky.serge@gmail.com.

Ротков Сергей Игоревич, д.т.н., проф., зав. кафедрой Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета. E-mail:rotkov@nngasu.ru.

Тюрина Валерия Александровна, к.т.н., доцент, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета. E-mail: 55555_73@mail.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Циклография. Аспекты теории и практических применений

К.Л. Панчук, Е.В. Любчинов, Т.М. Мясоедова
panchuk_kl@mail.ru|lubchinov.e.v@yandex.ru|mtm44mtm44@mail.ru
Омский государственный технический университет, Омск, Российская Федерация

В настоящей работе рассматриваются вопросы развития теории и расширения областей практического применения одного из классических методов геометрического моделирования - циклографии. Кратко приведены известные основные понятия, геометрические конструкции и их математическое описание, используемые в циклографическом моделировании. Рассмотрено получение новых теоретических результатов, развивающих циклографию и расширяющих ее возможности в решении разнообразных задач науки и практики.

Ключевые слова: *циклография, отображение, образ и прообраз, геометрическая оптика, гиперэпюр Наумович .*

Cyclography. Aspects of theory and practical applications

K.L. Panchuk, E.V. Lyubchinov, T.M. Myasoedova
panchuk_kl@mail.ru|lubchinov.e.v@yandex.ru|mtm44mtm44@mail.ru
Omsk State Technical University, Omsk, Russian Federation

In the present paper, questions of development of the theory and expansion of areas of practical application of one of the classical methods of geometric modelling – cyclography - are considered. Briefly given known basic concepts, geometric constructions and their mathematical description, used in cyclic modeling. Consideration is given to obtaining new theoretical results that develop the cyclography and expand its possibilities in solving various problems of science and practice.

Keywords: *cyclography, mapping, image and protoimage, geometric optics, Naumovich hyperdrawing*

1. Введение

Теоретические основы циклографии заложены в трудах западноевропейских геометров: W. Fiedler [5], L. Eckhart [4], E. Müller и J. L. Krames [7]. Современный уровень теоретического развития и области практических применений циклографического моделирования отражены в работах [1–3, 11–15].

Изначально циклография рассматривалась как изоморфное отображение *C* пространства R^3 на плоскость (*xy*): $C: (x, y, z) \leftrightarrow (x, y, r = z)$. Этим отображением приводятся во взаимно-однозначное соответствие параметрически равные множества точек пространства R^3 и циклов (ориентированных) окружностей на плоскости (*xy*). При этом само отображение может быть выполнено математически (отображение C_A) и конструктивно-геометрически (отображение C_G). В классических работах по циклографии [4, 5, 7] были получены и исследованы циклографические образ линейных объектов (прямых, плоскостей и их множеств) евклидова пространства, а также положено начало циклографического моделирования нелинейных объектов (кривых и поверхностей).

Вопросам теории и практических приложений циклографического моделирования объектов евклидова пространства посвящены, кроме вышеуказанных, работы [8–10, 18].

Учитывая известные результаты исследований в циклографическом моделировании, рассмотрим более общие подходы к решению его теоретических и практических задач.

2. Развитие теории и практические приложения циклографического моделирования

2.1. Обобщенная циклографическая модель кривой линии

Существующий циклографический образ a_α: $x_\alpha = x_\alpha(t), y_\alpha = y_\alpha(t)$ кривой линии пространства a: $x = x(t)$, $y = y(t)$, $z = z(t)$, $t \in R$ получен на основе отображения C_G, в котором полуугол при вершине конуса отображения с вершиной в текущей точке кривой постоянен и равен $\alpha = 45°$. Вся предшествующая настоящему времени циклография использовала именно такой конструктивно-метрический вариант отображения. Авторами работы предложено использовать переменный полуугол, например, в виде некоторой функции $\beta = f(t)$, $t \in R$, где *t* – текущий параметр, определяемый геометрией исходной кривой – прообразом. Обобщённый циклографический образ a_β описывается параметрическими уравнениями более общего вида [10]: $a_\beta: x_\beta = x_\beta(t), y_\beta = y_\beta(t), t \in R$ в сравнении с известными уравнениями образа a_α [2, 3]. Оба циклографических образа: существующий a_α и предложенный a_β, представляют собой огибающие соответствующих однопараметрических множеств циклов – ориентированных круговых оснований конусов отображения C_G.

Обобщённый циклографический образ a_β линии $a \subset R^3$ положен в основу математического моделирования автомобильной дороги. На рис. 1 представлена визуализация результатов вычислений в примере циклографического моделирования этого объекта [10].

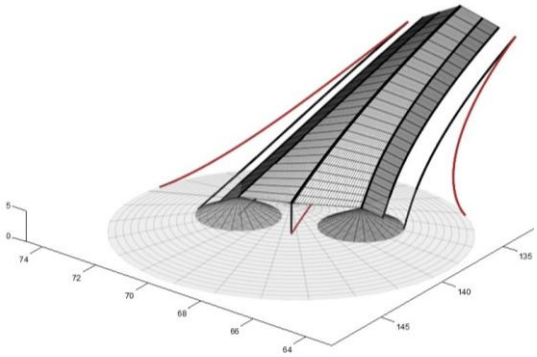


Рис.1. Циклографическое моделирование проезжей части и обочины автомобильной дороги

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

2.2. Циклографическое моделирование в проектировании обработки карманных поверхностей изделий машиностроения

Оптимизация траектории движения режущего инструмента для обработки карманных поверхностей на станках с ЧПУ традиционно выполняется средствами геометрического моделирования на плоскости [6]. При этом используется медиальная ось (биссекторная линия) контура карманной поверхности как множество точек самопересечений линий *ОС* (“*Offset Curves*” [2]) равных расстояний δ_i по отношению к исходному контуру. Линии *ОС* рассматриваются как траектории движения инструмента, обрабатывающего карманную поверхность. Определение линий *ОС* выполняется аналитически по параметрическим уравнениям $x_\delta = x_\delta(t), y_\delta = y_\delta(t), t \in R$, полученным из параметрических уравнений исходного контура a: $x_a = x_a(t), y_a = y_a(t)$ [2, 3] с последующим анализом и корректировкой самопересечений линий *ОС* для определения медиальной оси и обеспечения необходимой траектории движения инструмента.

Авторами работы рассмотрено пространственное решение задачи построения множества линий *ОС* на основе циклографического отображения в пространстве R^3 контура карманной поверхности, определения в пространстве образа медиальной оси и образов множества линий *ОС* с последующим построением соответствующих им прообразов в плоскости исходных данных. Визуализация решения задачи определения множества *ОС* позволяет более качественно выполнить анализ полученного множества линий при построении траектории режущего инструмента. Рисунки 2 - 6 демонстрируют на примере эллиптического контура последовательность выполнения предлагаемого решения задачи построения множества линий *ОС*.

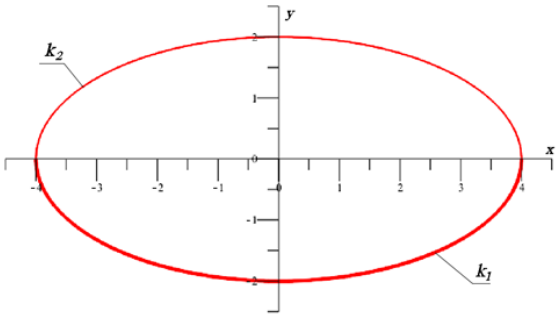


Рис.2. Исходный эллиптический контур

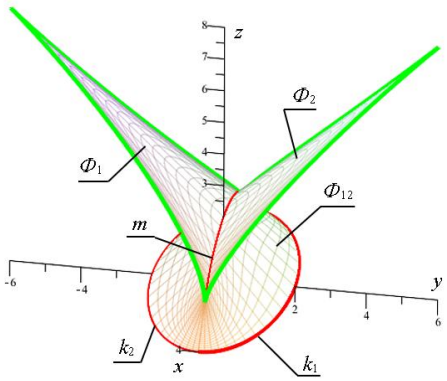


Рис.3. Составной линейчатый образ эллиптического контура

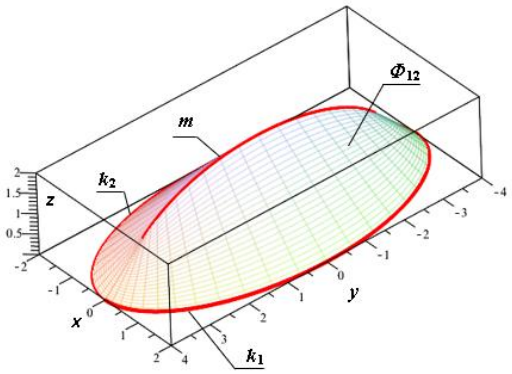


Рис.4. Линейчатый образ эллиптического контура без “рожковых” частей

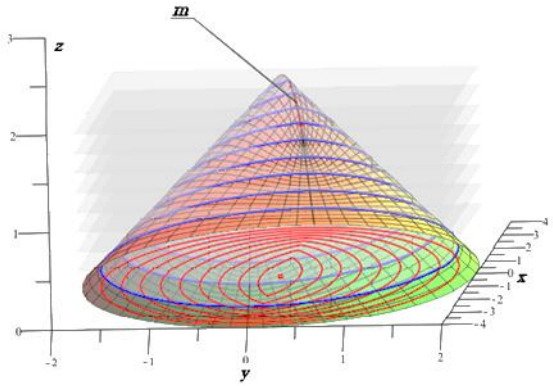


Рис.5. Образ множества линий *ОС*

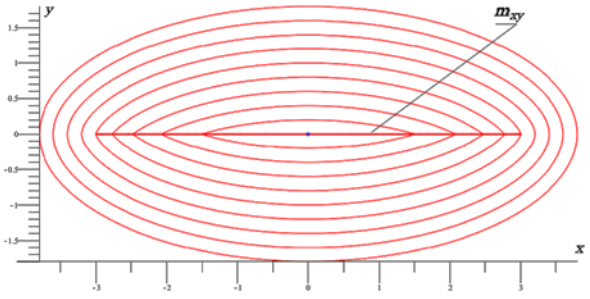


Рис.6. Прообраз - множество линий *ОС*

2.3. Циклографическое моделирование решений задач геометрической оптики

Задачи геометрической оптики можно условно разделить на два типа:

1. Прямая задача, в которой требуется определить отражатель (зеркало), преобразующий один пучок лучей с носителем пучка в другой, также определенный носителем.
2. Обратная задача, когда требуется по заданному пучку лучей и отражателю определить носитель отраженного пучка.

Обе задачи имеют известные в геометрической оптике решения лишь для частных случаев, когда носителями пучков являются точка или прямая линия. Для общего случая, когда носителем одного из пучков является кривая линия, существующие решения отличаются громоздкостью и приближёнными результатами вычислений. Авторами работы предложены алгоритмы решений прямой и обратной задачи, основанные на циклографическом отображении с выходом из плоскости задания условий задач в пространство. Алгоритмы охватывают все возможные варианты геометрических форм носителей в парах пучков и характеризуются аналитическим решением обеих задач. На рис. 7 приведены результаты

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

циклографического решения прямой задачи для двух нелинейных пучков с носителями b_0 и d_0 [9, 18].

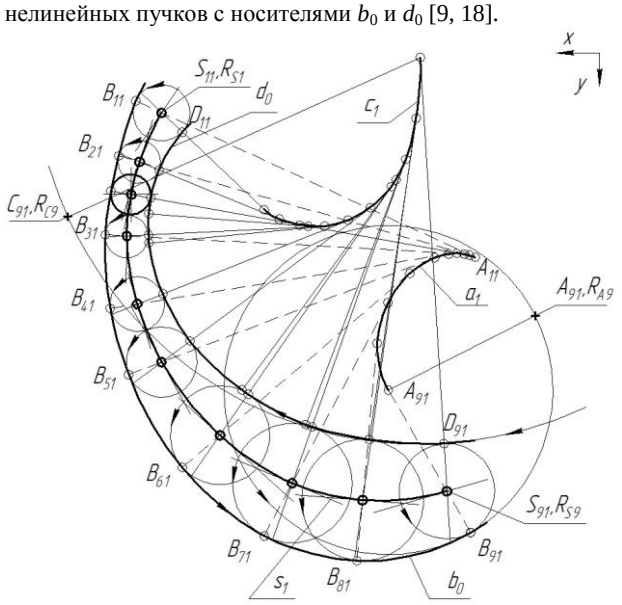


Рис. 7. Циклографическое решение прямой задачи для случая двух нелинейных пучков

2.4. Пространственная циклография и ее реализация на гиперэпюре Наумович

Возможность построения пространственной циклографии рассмотрена в работах [7, 15]. Суть ее заключается в построении в пространстве R^{n-1} циклографических образов объектов пространства $R^n, n > 3$. Так, например, используя отображение C_A можно поставить во взаимно-однозначное соответствие точке (*x, y, z, v*) пространства R^4 сферу $S(x, y, z, r = v)$ пространства R^3 (рис. 8). Очевидно, кривой линии $a: x(t), y(t), z(t), v(t), t \in R$ пространства R^4 будет соответствовать в пространстве R^3 каналовая поверхность $\Phi_\alpha : x_\alpha = x_\alpha(t,u), y_\alpha = y_\alpha(t,u), z_\alpha = z_\alpha(t,u)$, $u \in R$; 2-поверхности $\Phi^2 \subset R^4$ будет соответствовать огибающая $\Phi_\alpha^2 \subset R^3$ двухпараметрического множества сфер [11].

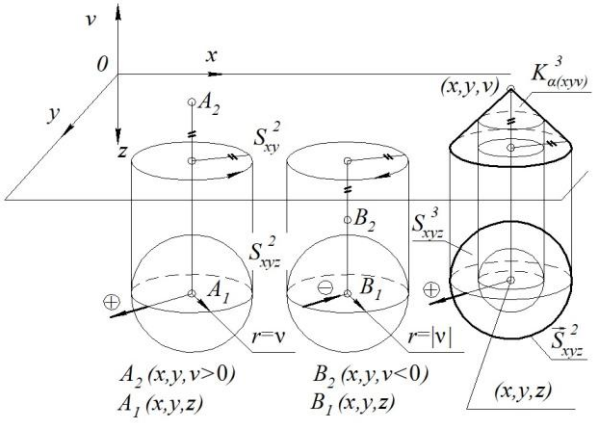


Рис.8. Циклографическое моделирование точки пространства R^4 на гиперэпюре Наумович

Авторами работы предложено выполнять циклографическое отображение C_G объектов пространства R^4 на основе гиперэпюра Наумович [17]. В этом отображении точка $A(x, y, z, v) \in R^4$ отображается α-гиперконусом K^3_α с вершиной в этой точке на гиперплоскость проекций (*xyz*) в двумерный цикл $\vec{S}^2_{xyz}$ (направленную сферу). Очевидно, кривая линия $a \subset R^4$ отобразится огибающей всех *α*-гиперконусов K^3_α ее точек – линейчатой *α*-гиперповерхностью Φ^3_α на гиперплоскость проекций (xyz) в каналовую поверхность Φ_α, которая представляет собой огибающую всех двумерных циклов $\vec{S}_{xyz}$ (рис.9).

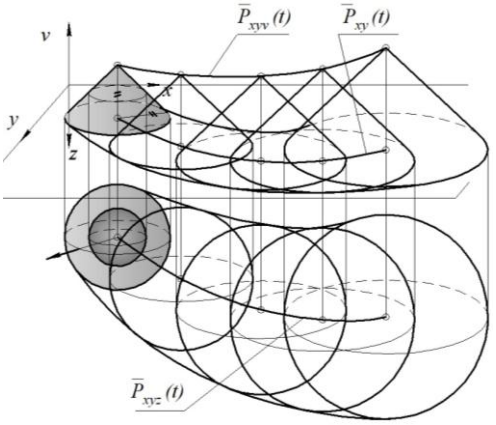


Рис.9. Циклографическое моделирование линии пространства R^4 на гиперэпюре Наумович

2.5. Циклографическое моделирование решения задачи измерения псевдодальностей в спутниковых системах определения местоположения

В теории измерения псевдодальностей [16] используется квадратичное уравнение для выполнения псевдодальномерных измерений, в котором неизвестными являются координаты (x_{np}, y_{np}, z_{np}) точки стояния приёмника и поправка на уход часов приёмника δt_{np}. Для определения этих неизвестных составляется система из четырех квадратных уравнений, соответствующих расстояниям до четырех спутников. С учетом замены переменных: $x_{np} \to x_1, y_{np} \to x_2, z_{np} \to x_3, \delta t_{np} \to x_4$ эту систему можно представить в виде

$$F_i(x_1, x_2, x_3, x_4) = (x_1 - a_{i1})^2 + (x_2 - a_{i2})^2 + (x_3 - a_{i3})^2 - (x_4 - a_{i4})^2 = 0,$$
$$i = 1, 2, 3, 4.$$

Систему 4-х квадратных уравнений можно алгебраическими преобразованиями привести к эквивалентной системе

$$F_1 = 0, F_1 - F_2 = 0, F_1 - F_3 = 0, F_1 - F_4 = 0,$$

в которой первое уравнение квадратное, а три последующие – линейные. Первое уравнение описывает гиперконус вращения K^3_α с телесным полууглом при вершине, равным 45^0, и направленным основанием – двумерным циклом $\vec{S}^2_{xyz}$, а остальные уравнения описывают гиперплоскости P^3_i, $i = 1, 2, 3$ пространства R^4. Геометрический анализ системы уравнений показывает, что её полным решением являются 16 точек (x_1, x_2, x_3, x_4) пространства R^4, которым соответствуют 16 сфер пространства R^3. На рисунках 10 и 11 приведена визуализация результатов двух решений в примере рассматриваемой задачи для определенной ориентации гиперконуса K^3_α и трех гиперплоскостей P^3_i [8].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Использование циклографических образов в пространстве R^3 объектов пространства R^4 позволяет получить аналитическое решение известной задачи об измерении псевдодальностей в спутниковых системах определения местоположения [8] в отличие от существующего приближенного решения [16]. Вариант постановки этой практической задачи и исследование её решений изложены в работе H. Stachel [15].

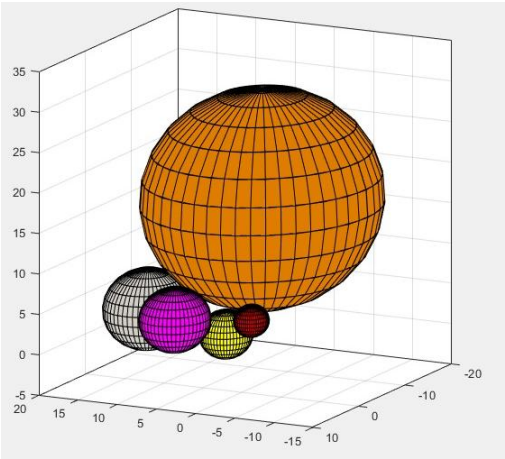


Рис.10. Геометрическая интерпретация первого решения

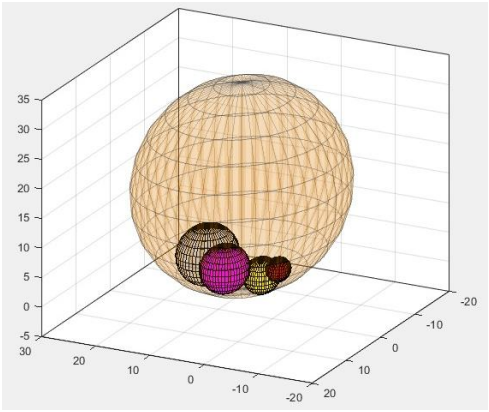


Рис.11. Геометрическая интерпретация второго решения

3. Заключение

Рассмотрены вопросы теории циклографического метода геометрического моделирования применительно к решениям разноплановых практических задач. Решения достаточно убедительно подтверждают работоспособность и демонстрируют потенциальные возможности развивающегося классического метода.

4. Литература

[1] Cho H.Ch., Choi H.I., Kwon S.-H., Lee D.S. and Wee N.S. Clifford algebra, Lorentzian geometry and rational parametrization of canal surfaces. Computer Aided Geometric Design, 21:327–339, 2004.

[2] Choi H.I., Han C.Y., Moon H.P., Roh K.H., Wee N.S. Medial axis transform and offset curves by Minkowski Pythagorean hodograph curves, Computer-Aided Design 31 (1999), 5 p.p. 9–72.

[3] Choi H. I., Choi S. W. and Moon H. P. Mathematical theory of medial axis transform. Pacific J. Math., 181(1):56–88, 1997.

[4] Eckhart L. Konstructive abbildungsverfahren. Wien. Verlag von Julius Springer, 1926, 120 p.

[5] Fiedler W. Cyklographie oder Construction der Aufgabenüber Kreise und Kugeln und elementare Geometrie der Kreis- und Kugelsysteme. – Leipzig, Druckund Verlag von B.G. Teubner, 1882. – 284 p.

[6] Held M. On the Computational Geometry of Pocket Machining. Lecture Notes in Computer Science, Vol 500, Springer Verlag. Berlin 1991, p.184.

[7] Dr. Emil Muller. Vorlesungen über Darstellende Geometrie. II. Band: Die Zyklographie. Edited from the manuscript by Dr. Josef Leopold Krames. Leipzig and Vienna, Franz Deuticke, 1929. - 476 pp.

[8] Panchuk K. L., Lyashkov A. A., Lyubchinov E. V. Geometric model of pseudo-distance measurement in satellite location systems. IOP Conf. Series: Journal of Physics: Conf. Series 998 (2018) 012021. doi:10.1088/1742-6596/998/1/012021.

[9] Panchuk K. L., Lyubchinov E. V., Krysova I. V. Surface triads with optical properties. IOP Conf. Series: Journal of Physics: Conf. Series 944 (2018) 012086. doi :10.1088/1742-6596/944/1/012086.

[10] Panchuk K.L., A. S. Niteyskiy, E. V. Lyubchinov. Cyclographic Modeling of Surface Forms of Highways. IOP Conf. Series: Materials Science and Engineering 262 (2017) 012108. doi:10.1088/1757-899X/262/1/012108.

[11] Peternell M. Rational two-parameter families of spheres and rational offset surfaces. J. Symbolic Computation 45 (2010), 1-18.

[12] Peternell M., Pottmann H. Computing rational parametrizations of Canal Surfaces. J. Symbolic Computation 23 (1997), 255–266.

[13] Pottmann H., Peternell M. Applications of Laguerre Geometry in CAGD, Comp. Aided Geometric Design 15 (1998), 165–186.

[14] Pottmann H., Wallner J. Computational Line Geometry. Berlin. Heidelberg: Springer Verlag, 2001. 565 p.

[15] Stachel H. (1996) “Why Shall We also Teach the Theory behind Engineering Graphics”, Technical Report No. 35, TU-Wien, Institute for Geometry, 5p.

[16] Генике А. А. Глобальные спутниковые системы определения местоположения и их применение в геодезии / А. А. Генике, Г. Г. Побединский. Изд. 2-е, перераб. и доп. – М.: Картгеоцентр, 2004. – 355 с.

[17] Короткий В. А. Компьютерное моделирование фигур четырехмерного пространства. Вестник компьютерных и информационных технологий. 2014. №7(21). С.14-20.

[18] Панчук К. Л. Циклографическая начертательная геометрия: монография/ К. Л. Панчук, Н. В. Кайгородцева. – Омск: Изд-во ОмГТУ, 2017. – 232 с.

Об авторах

Панчук Константин Леонидович, д.т.н., профессор кафедры «Инженерная геометрия и САПР» Омского технического университета. Его e-mail: panchuk_kl@mail.ru.

Любчинов Евгений Владимирович, аспирант, преподаватель кафедры «Инженерная геометрия и САПР» Омского технического университета. Его e-mail: lubchinov.e.v@yandex.ru.

Мясоедова Татьяна Михайловна, старший преподаватель кафедры «Инженерная геометрия и САПР» Омского технического университета. Ее e-mail: mtm44mtm44@mail.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Аналитическое задание четырехмерной области разрешенных конфигураций руки андроидного робота при наличии запретной зоны в рабочем пространстве

Ф.Н. Притыкин, Д.И. Нефедов
pritykin@mail.ru|3demon@bk.ru
Омский государственный технический университет, Омск, Россия

Приведены результаты исследований, связанных с разработкой модуля для автоматического обнаружения столкновений механизма руки андроидного робота, с запретной зоной при виртуальном моделировании движения. С этой целью для определения условия столкновений предложено использовать область разрешенных конфигураций. В работе предложен аналитический способ задания области разрешенных конфигураций в пространстве обобщенных координат при наличии заранее известной запретной зоны в рабочем пространстве. Пространство обобщенных координат при этом задано осями, по направлению которых откладывают углы, задающие угловые перемещения во вращательных кинематических парах. Приведены результаты расчетов совокупностей двумерных сечений задающих область разрешенных конфигураций. Данные сечения предложено задавать с помощью использования совокупностей треугольников. В работе представлены результаты численных расчетов, связанных с виртуальным моделированием движения механизма руки андроидного робота при наличии известной запретной зоны и использовании разработанного модуля. Полученные аналитические зависимости позволяют задавать область разрешенных конфигураций, являющейся базой знаний при интеллектуальном управлении движением механизма манипулятора в заранее известной внешней среде. Использование области разрешенных конфигураций при синтезе траектории в пространстве обобщенных координат позволяет проводить коррекцию движения руки андроидного робота с целью предвидения и исключения тупиковых ситуаций при синтезе движений по вектору скоростей.

Ключевые слова: *пространство обобщенных координат, область разрешенных конфигураций, виртуальное моделирование движений, запретные зоны, база знаний, управление движением робота.*

Analytical defining of the four-dimensional region of allowed configurations of the android robot arm in the presence of a restricted area in the working space

F.N. Pritykin, D.I. Nefedov
pritykin@mail.ru | 3demon@bk.ru
Omsk State Technical University, Omsk, Russia

The results of research related to the development of the module for automatic collision detection mechanism of android robot hand with a restricted area in the virtual simulation of the motion are presented. For this purpose, collision detection conditions proposed to use the allowed configurations. An analytical method for defining the region of allowed configurations in the space of generalized coordinates has been developed in the presence of a previously known forbidden zone in the working space. Generalized coordinates space with axes defined rectangular coordinate system, the direction of which the lay angles of the angular displacement in the rotational kinematic pairs. The results of calculations of the sets of sections defining the region of allowed configurations are presented. These sections are proposed to be specified using the sets of triangles. The paper presents the results of numerical calculations associated with a virtual simulation of the motion of android robot arm mechanism in the presence of a known restricted area and using the developed module. The obtained analytical dependencies allow us to define the region of allowed configurations, which is the knowledge base for the intelligent control of the movement of the manipulator mechanism in a previously known external environment. The use of the region of allowed configurations in the synthesis of the trajectory in the space of generalized coordinates makes it possible to correct the manipulator's movement with the aim of foreseeing and eliminating deadlocks in the synthesis of motions along the velocity vector.

Keywords: *space of generalized coordinates, the region of allowed configurations, virtual motion modeling, forbidden zones, knowledge base, robot motion control.*

1. Введение

В настоящее время ведутся работы, связанные с созданием андроидных роботов. В большинстве случаев данные роботы управляются человеком оператором. Вместе с тем так же ведутся работы связанные с созданием андроидных роботов оснащенных интеллектуальными системами управления. Данные системы позволяют обеспечить целенаправленное поведение и планирование действий приводящих в конечном итоге к выполнению основной двигательной задачи без участия человека оператора. Автономный робот – это техническая система, способная автономно перемещать объекты манипулирования при наличии препятствий в рабочем пространстве [1, 4].

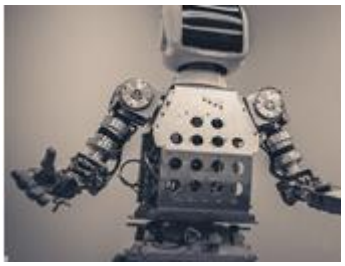
Одной из наиболее распространенных аварийных ситуаций, возникающих при функционировании автономного робота, является его столкновение со средой [2, 5]. Для обеспечения успешного функционирования автономных андроидных роботов в сложноорганизованных средах необходимо совершенствовать способы анализа положения механизма руки в известной и неизвестной окружающей среде. Поэтому существует необходимость в разработке модуля автоматического обнаружения столкновений с запретными зонами. С геометрической точки зрения столкновению двух тел на виртуальном уровне соответствует наличие пересечения двух граней пространственных примитивов задающих в заданные моменты времени положения звеньев механизма и окружающей среды. Таким образом, для обнаружения пересечения манипулятора со средой необходимо перебрать все возможные комбинации граней примитивов

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

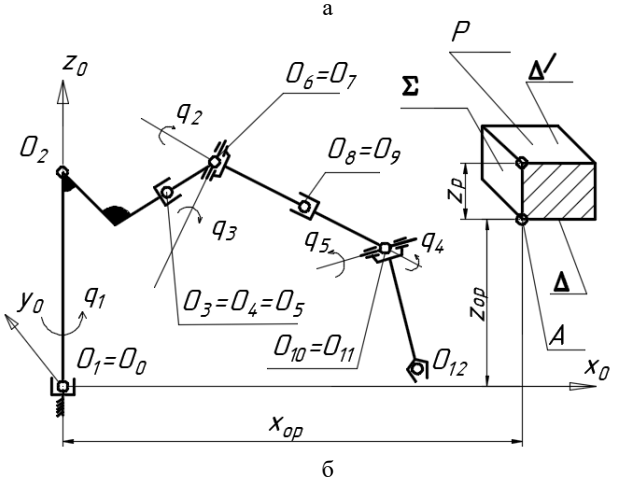
до тех пор, пока не будет обнаружено пересечение между ними. Если это условие при окончании процедуры обнаружения пересечений не будет выполнено, то столкновение отсутствует. Однако на практике использование указанного алгоритма вызывает некоторые сложности в связи со слишком большим количеством комбинаций граней. Для сокращения времени расчетов в работах [2, 3, 6] первоначально предлагается определять условие не пересечения траектории движения в пространстве обобщенных координат с границами области разрешенных конфигураций. В указанных работах предложен алгоритм построения движения механизма манипулятора, основанный на анализе положения точек пространства обобщенных координат, задающих разрешенные и запрещенные конфигурации. Все разрешенные конфигурации в пространстве обобщенных координат, могут быть заданы n - мерным геометрическим объектом. Где n – определяет число обобщенных координат. Предложенный алгоритм позволяет существенно повысить быстродействие процедуры обнаружения столкновений.

2. Постановка задач исследования

На рисунке 1аб изображены общий вид и кинематическая схема механизма руки андроидного робота AR-600 E и положение запретной зоны *P*. Высота базовой точки *A* запретной зоны для рассматриваемого примера в системе координат O_o принята равной z_{op} = 400 мм, минимальное безопасное удаление основания манипулятора от препятствия принято равным x_{op} = 450 мм (рис. 1б).



а



б

Рис. 1. Андроидный робот AR-600 E: а – общий вид робота; б – кинематическая схема механизма руки

Высоту запретной зоны (стеллажа) определяет параметр z_p = 100 мм. Геометрические параметры механизм манипулятора, задают длины отрезков O_1O_2 = 300 мм, O_2O_3 = 120 мм, O_3O_6 = 100 мм, O_6O_8 = 80 мм, O_8O_{10} = 150 мм и $O_{10}O_{12}$=250 мм. На рисунке обобщенные координаты обозначены q_i. Точки O_1 - O_{12} определяют начала систем координат, используемых при задании модели кинематической цепи механизма руки андроидного робота. Для того чтобы определить область разрешенных конфигураций (которую обозначим Λ), с учетом положения запретной зоны *P*, построим совокупность двумерных сечений указанной области в четырехмерном пространстве обобщенных координат. Отдельные положения звеньев механизма руки будем исследовать при q_1 = 0. Запретная зона *P*, заданная двумя горизонтальными и одной профильной плоскостями уровня Δ, Δ/ и Σ (рис. 2), задает в пространстве обобщенных координат L_q некоторую область разрешенных конфигураций Λ. Исследуем сечения этой области, а так же определим аналитический метод нахождения принадлежности точек пространства обобщенных координат указанной области Λ. Способ вычисления точек сечений области разрешенных конфигураций Λ для создания базы знаний о прошлом опыте синтеза движений механизма манипулятора мобильного робота ранее был предложен в работах [2,3].

3. Аналитический способ задания сечений области разрешенных конфигураций

Функционирование механизма манипулятора происходит в пределах ограниченной области Ω (четырехмерного пространства обобщенных координат L_q при q_1 = 0). Область Ω, в общем случае заданная гиперпараллелепипедом, имеет такой вид, что для любого значения вектора $\boldsymbol{q}$ (q_2, …, q_5) выполняются неравенства:

$$q_i^{min} \le q_i \le q_i^{max}, \qquad (1)$$

где q_i^{min}, q_i^{max} – нижние и верхние предельные значения обобщенных координат. Для рассматриваемого примера при q_1 = 0 значения параметра *i* удовлетворяют условиям $2 \ge i \ge 5$. Значения q_i^{min} и q_i^{max} соответственно равны q_i^{min} (0^o, -25^o, -120^o, -120^o,) и q_i^{max} (120^o, 120^o, 120^o, 120^o).

На рисунке 2 представлено множество разрешенных конфигураций, построенных с шагом $\Delta q_i = 20^o$ при наличии запретной зон *P*. На рисунке запретная зона изображена на горизонтальной и фронтальной проекциях в виде четырехугольников P_1 и P_2. Условимся четырехмерную область разрешенных конфигураций Λ задавать совокупностью двумерных сечений. В связи со сложной и различной формой двумерных сечений области Λ андроидного робота предложенная в работах [2,3] методика аналитического задания сечений не может быть применена. При шаге сетки $\Delta q_i = 20^o$ в проведенных исследованиях изучена форма восьмидесяти четырех сечений. В таблице представлены некоторые из указанных сечений области разрешенных конфигураций Λ. Для задания точек сечений разработан обобщенный метод задания их совокупностями треугольников l_1 и l_2, и т.п., которые будут задавать двумерные области Ω_1 и Ω_2 и т.д. (рис. 3). При этом области Ω_1 и Ω_2 и т.д. получают на основе использования операций объединения теории множеств. На рисунке 3 представлено сечение области разрешенных конфигураций Λ соответственно при значении обобщенных координат q_2 = 40^o, q_3 = -100^o, при этом x_{op} = 450 мм., z_{op} = 400 мм и z_p = 100 мм. На рисунке 3 изображены треугольники, используемые для задания сечения области Λ.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

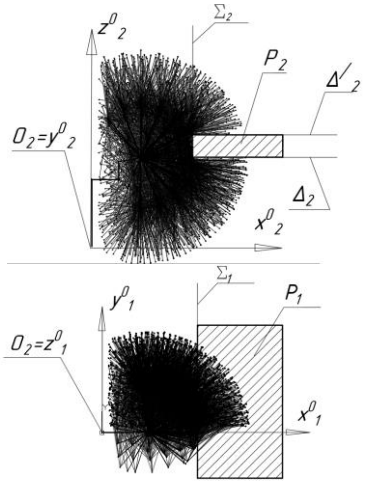
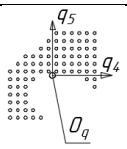
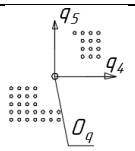
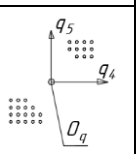
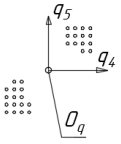
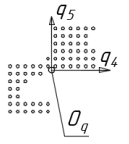
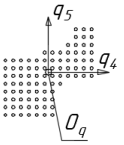
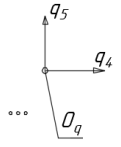
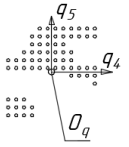
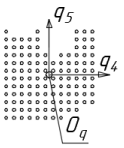


Рис. 2. Множество конфигураций, построенных при наличии запретной зоны *P*, заданной в виде прямоугольного параллелепипеда z_{op} = 400 мм, x_{op} = 450 мм, z_{op} = 400 мм;

Таблица

Некоторые сечения области Λ при различных значениях параметров q_2 и q_3

	q_3 = -100°	q_3 = 20°	q_3 = 100°
q_2 = -20°			
q_2 = 20°			
q_2 = 60°			

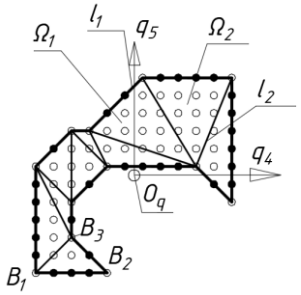


Рис. 3. Разделение сечения области разрешенных конфигураций Λ на треугольники при q_2 = -20° и q_3 = -100°

Для разделения сечения на треугольники (триангуляция [7]) необходимо выполнить три основных этапа:

1) Определить контур сечения (на рис. 3 данный контур изображен утолщенной линией):
- точкой, определяющей контур сечения, является любая точка, у которой отсутствует любая из 4 соседних точек (непосредственно сверху, снизу, слева или справа);
- точка, у которой отсутствуют все 4 соседние точки, считается не принадлежащей сечению;

2) Определить вершины сечения (отбросить промежуточные точки контура, лежащие на сторонах сечения. На рис. 3 данные вершины обозначены черной заливкой):
- определить последовательность точек контура (у каждой точки контура в радиусе 30 градусов будут находится еще две, и эти три точки считаются лежащими последовательно на контуре);
- для каждой точки контура определить, лежит ли она на прямой, определяемой двумя соседними точками;
- на данном этапе если непосредственно соседняя точка уже удалена, соседней считается следующая по контуру;

3) Провести собственно триангуляцию «ушным методом» [7]:
- взять три соседние вершины на контуре;
- определить, является ли треугольник, определяемый этими вершинами, «ухом» (треугольник является «ухом» многоугольника, если две его стороны являются сторонами многоугольника, а третья полностью лежит внутри него и не пересекает другие стороны многоугольника);
- если треугольник является «ухом», отрезать его от многоугольника. Три вершины добавить в базу знаний как часть исходного многоугольника. Свободную вершину (лежащую напротив ребра многоугольника) удалить из контура, определив новый многоугольник;
- повторять предыдущие шаги, пока в новом многоугольнике не останется три вершины, добавить их в базу знаний как последнюю часть многоугольника (линии сечения на рис. 3 обозначены тонкими линиями).

Области треугольников Ω_i задают полуплоскостями, определяемыми неравенствами:

$$\sum_{l=3}^{4} d_{l1} q_l \geq b_1, \ldots, \sum_{l=2}^{4} d_{lm} q_l \geq b_m \quad , \qquad (2)$$

где m – число прямых ограничивающих треугольник области Λ (для рассматриваемого примера $3 \geq m \geq 1$), $l = 2$ – параметр определяет размерность пространства сечений области разрешенных конфигураций при $q_1 = 0$, q_2 = const и q_3 = const; d_{11}, d_{12}, …, d_{lm}, b_1, b_2, …, b_m – коэффициенты уравнений прямых, проходящих через вершины B_i треугольников l_i.

4. Результаты виртуального моделирования движения руки андроидного робота

Совокупность сечений области разрешенных конфигураций и их аналитическое задание в виде множества треугольников определяется заранее на основе известного положения запретных зон (например, стеллажей) и механизма руки. Область Λ при этом задает базу знаний и модель окружающей рабочей обстановки. На основе использования области Λ появляется возможность создания алгоритма с целью прогноза развития событий при выполнении рукой робота двигательных задач. Алгоритм расчета траектории движения механизма в пространстве обобщенных координат с использованием построения движений по вектору скоростей и области Λ представлен в работах [2, 5]. Прогноз возникновения различных ситуаций при движении манипулятора в данных работах решается благодаря использованию масштаба отображения вектора приращений обобщенных координат, а так же анализу положения траектории по отношению к области Λ. На рисунке 4а представлены результаты моделирования движения руки андроидного робота с возникновением тупиковой ситуации. При этом центр

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

захвата смещается по отрезку прямой, которая соединяет начальную и конечную точки. Указанные точки на рис. 4аб обозначены $B_н$ и $B_ц$.

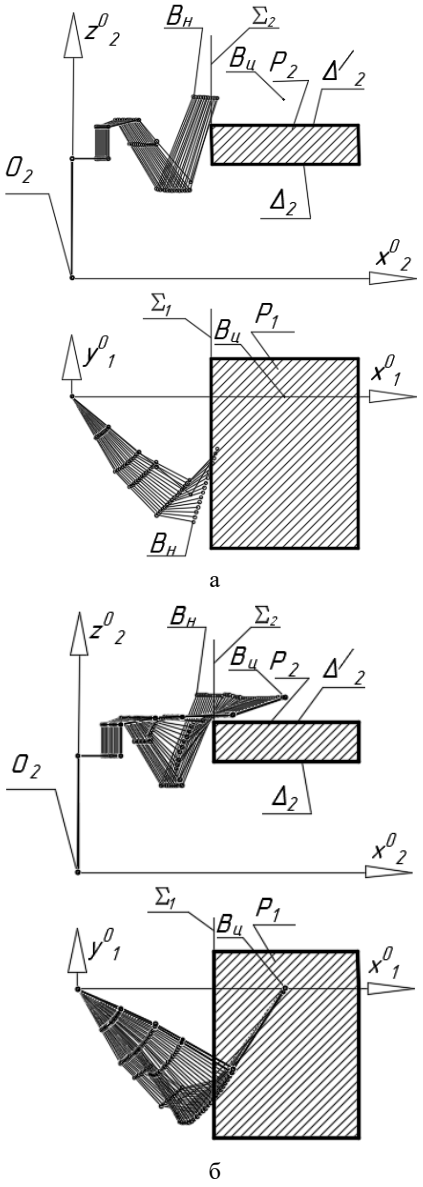


Рис. 4. Моделирование движения руки андроидного робота: а – моделирование с возникновением тупиковой ситуации, б – моделирование с использованием вектора приращений обобщенных координат и области Λ

На рис. 4б на основе использования области Λ и значения масштаба вектора приращений обобщенных координат смоделировано движение на всем участке заданной траектории.

5. Заключение

Результаты построения сечений области разрешенных конфигураций андроидного робота показали, что форма наружных контуров сечений может значительно отличаться друг от друга. Так же сечения могут состоять из двух и более отдельных контуров. Для определения точек, принадлежащих сечениям области разрешенных конфигураций, наиболее рационально задавать указанные сечения множествами треугольников.

Аналитическое задание области разрешенных конфигураций руки андроидного робота, позволяет создать модуль автоматического обнаружения столкновений со средой на виртуальном уровне. Для приобретения и правильного представления знаний о возможных положениях руки с учетом заранее заданной запретной зоны использована область разрешенных конфигураций. Данная область позволяет представлять знания о множестве положений руки в такой форме, которая понятна интеллектуальной системе. При этом формализация знаний на основе использования области Λ позволяет создание модели внешней среды с описанием множества ситуаций.

Результаты исследований могут быть использованы при разработке интеллектуальных систем управления автономно функционирующих роботов, в частности при создании баз знаний о прошлом опыте синтеза движений.

6. Литература

[1] Лохин, В. М. Повышение адаптивных свойств автономных роботов на базе интеллектуальных технологий / В. М. Лохин, С. В. Манько, М. П Романов. // Экстремальная робототехника. – 2015. – № 1 (1). – С. 59–67.

[2] Притыкин, Ф. Н. Геометрический анализ текущих ситуаций, характеризующих положение манипулятора и окружающей среды на основе использования области разрешенных конфигураций [Электронный ресурс] / Ф. Н. Притыкин , В. Г. Хомченко, Д. И. Нефедов // Программные системы и вычислительные методы. – 2017. – № 2. – Режим доступа: http://e-notabene.ru/ppsvm/article_22685.html

[3] Притыкин, Ф. Н. Исследование поверхностей, задающих границы области разрешенных конфигураций механизма мобильного манипулятора при наличии запретных зон / Ф. Н. Притыкин, Д. И. Нефедов // Мехатроника, автоматизация, управление. – 2016. – № 6, – т. 17. – С. 404–413.

[4] Юсупов, Р. М. Интеллектуализация процессов управления и навигации робототехнических систем / Р. М. Юсупов, А. В. Тимофеев // Экстремальная робототехника. – 2014. – Т. 1, № 1. – С. 16–21.

[5] Whitney, D. E. The Mathematics of Coordinated Control of Prosthetic Arms and Manipulators / D. E. Whitney // J. Dyn. Sys., Meas., Control. – 2010. – № 94 (4). – P. 303–309.

[6] Pritykin F.N. Computation of manipulator mechanism path in joint coordinate space with working range forbidden regions / F.N.Pritykin, D.I. Nefedov // Dynamics of Systems, Mechanisms and Machines (Dynamics), 2016, 15-17 Nov. 2016, DOI: 10.1109/Dynamics.2016.7819065

[7] Eberly, D. Triangulation by Ear Clipping. / David Eberly // Geometric Tools, Redmond WA 98052

Об авторах

Притыкин Федор Николаевич, д.т.н., профессор кафедры «Инженерная геометрия и САПР» Омского Государственного Технического Университета. Его e-mail pritykin@mail.ru.

Нефедов Дмитрий Игоревич, аспирант кафедры «Инженерная геометрия и САПР» Омского Государственного Технического Университета. Его e-mail 3demon@bk.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

«Доступная» графика в довузовском образовании

И.Д. Столбова, М.А. Ширинкина
stolbova.irina@gmail.com | shirinkina-masha@mail.ru
ФБГОУ ВО «Пермский национальный исследовательский политехнический университет», Пермь, Россия

На сегодняшний день в России ярко выражена тенденция технологизации, развития промышленности и, как следствие, инженерного дела. В рамках расширения и совершенствования данных направлений деятельности крайне актуальным становится развитие знаний и навыков по дисциплинам «Черчение» и «Начертательная, инженерная и компьютерная графика». Несколько лет назад эти дисциплины были выведены из программ общеобразовательных школ или сокращены до нескольких тематических часов в рамках сопутствующих дисциплин. Это способствовало появлению большой пропасти знаний, а также отсутствию у учащихся (абитуриентов) умений и навыков в сфере инженерного дела.

Кроме того, развитие интернета и компьютерных технологий в нашем XX1 веке направляет процесс образования в сторону развития умений находить нужную информацию и применять её, а применение электронных ресурсов становится неотъемлемой частью обучения. Поэтому широкое применение современной компьютерной графики для освоения вышеупомянутых дисциплин является объективной необходимостью. Всё это провоцирует возникновение непонимания между учащимися и «старшим» поколением, которое считает эталоном образования выученные наизусть стандарты.

Целью данной статьи является пример внедрения смешанного курса довузовский графической подготовки с использованием электронных ресурсов, а также результаты прохождения учащимися ознакомительного курса «Основы инженерной графики».

Ключевые слова: *довузовское образование, черчение, инженерная графика, компьютерная графика, информация, электронные ресурсы, процесс обучения.*

«Available» graphics in pre-university education

I.D. Stolbova[1], M.A. Shirinkina[2]
stolbova.irina@gmail.com | shirinkina-masha@mail.ru
[1]Perm national research polytechnic university, Perm, Russia;
[2]Perm national research polytechnic university, Perm, Russia

Today in Russia there is a bright tendency of technologization, industrial development and, as a result, engineering. As part of the expansion and improvement of these activity areas it is becoming extremely actual to develop knowledge and skills in "Drawing" and "Descriptive, engineering and computer graphics". A few years ago these disciplines were excluded from regular school curriculum or reduced to several thematic hours within related subjects. This has contributed to the emergence of great knowledge gap, as well as students (entrants) have a lack of skills and abilities in engineering sphere.

In addition, development of the Internet and computer technologies in our XXI century is directing the education process in the ability development to find the necessary information and use it, and application of electronic sources is becoming an integral part of the education. Therefore, the widespread use of modern computer graphics to master the above disciplines is an objective necessity. All this leads to misunderstanding between students and the "older" generation, which considers that learnt by heart standards are the sample of education.

The purpose of this article is an example of the implementation of electronic sources in the pre-University education process, and also, the results of students, who copes with the "Fundamental engineering graphics" course.

Keywords: *pre-university education, drawing, engineering graphics, computer graphics, information, electronic sources, education process.*

1. Введение

Передовая экономическая модель характеризуется высоким уровнем автоматизации деятельности инженера, смещением грани между проектированием и производством. На первый план выступает компьютерное моделирование, которое сопровождает весь производственный процесс, а в основе проектирования находится виртуальная модель изделия [16]. Компьютеризация общества, ярко выраженная технологизация, развитие промышленности и инновации в области инженерного дела привели к изменениям и в образовательной сфере, направленных на подготовку специалистов технического профиля на всех ступенях образования. В рамках расширения и совершенствования вышеперечисленных аспектов крайне актуальным является наличие достаточных знаний и навыков по дисциплинам «Черчение» и «Начертательная геометрия, инженерная и компьютерная графика», в том числе и в рамках довузовского обучения.

К сожалению, несколько лет назад эти дисциплины были выведены из программ общеобразовательных школ или значительно сокращены до нескольких тематических часов в рамках сопутствующих дисциплин. В отведённое время имеется возможность рассмотреть лишь незначительную часть учебного материала, дающего представление об основах современной инженерной деятельности. К тому же, актуальные знания не только не доносятся до учащихся, но и не являются багажом знаний самого учителя, что в свою очередь становится причиной деформации знаний по предмету [1]. Конечно, есть категория учителей, обладающих техническим образованием и довольно хорошо владеющих правилами выполнения чертежей. Но для организации процесса обучения, а также формирования целостной геометро-графической части образовательной программы

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

учителю необходимо знать современные технологии проектирования, а также владеть продвинутой методической системой передачи своих знаний по преподаваемому предмету. Это способствует появлению все большей пропасти между актуальной довузовской геометро-графической подготовкой и действующими учебными программами, а также отсутствию у учащихся (абитуриентов) первоначальных знаний и умений в сфере инженерного дела.

В условиях глобальной информатизации производства, повышения наукоемкости современных технологий расширяются возможности использования компьютерных технологий в области инженерно-конструкторской деятельности. Это предполагает усиление роли компьютерной графики в рамках геометро-графической подготовки и не только в качестве развития навыков автоматизации чертежно-графических работ, но и ознакомления с передовыми технологиями 3d-моделирования.

С развитием интернета и компьютерных технологий отличительной чертой современной жизни стала доступность информации. Хорошо известно, что неотфильтрованная информация не всегда является достоверной. В связи с этим, в сегодняшних реалиях процесс образования необходимо направлять в сторону развития умений учащихся находить нужную и достоверную информацию, обрабатывать ее и анализировать, а также эффективно применять полученные знания в процессе обучения [5, 6].

Кроме того, развитие интернета способствует появлению огромного количества электронных информационных и образовательных ресурсов. Это позволяет осуществлять процесс обучения более мобильно, интересно, насыщенно и информативно, и, конечно же, делает его более доступным для широкого круга обучающихся. Сегодня применение электронных ресурсов становится неотъемлемой частью процесса обучения [2, 9]. Молодое поколение подобные инновации воспринимает с увлечением, но не всегда адекватно оценивает их полезность.

Недостаточна и подготовленность взаимодействия с информационными ресурсами педагогических кадров. Вносимые изменения в процесс образования и внедрение электронных ресурсов провоцируют возникновение непонимания между учащимися и более «старшим» поколением, которое зачастую считает эталоном образования выученные наизусть определения, стандарты, нормативы. Для таких педагогов не очевидны имеющиеся явные преимущества применения компьютерных технологий в образовании.

2. Актуальные проблемы и пути совершенствования довузовского графического образования

Качество довузовской подготовки абитуриентов напрямую влияет на сферу высшего профессионального образования, а уровень графической подготовки на начальном этапе обучения определяет успешность адаптации студентов в высшей технической школе. Анализ литературы, обобщение данных многолетних наблюдений и собственный практический опыт авторов в организации довузовской графической подготовки учащихся позволяет выделить ряд насущных проблем [3, 4, 15].

1. Современный старшеклассник отличается выраженным отсутствием осведомленности о своей будущей профессиональной деятельности, отсутствием стремления к профессиональному росту и самосовершенствованию. Кроме того, ему присуща низкая самооценка, пониженный уровень зрелости и социальной ответственности, равнодушие по отношению к происходящим в обществе процессам [13, 14].
2. В рамках школьной программы крайне мало предметов, направленных на развитие творческого потенциала учащихся, так как основное направление обучения сводится к «натаскиванию» мышления школьников на решение тестов и заданий ЕГЭ [10]. Это, в свою очередь, заведомо исключает развитие творческого мышления и способности к решению нестандартных задач, в том числе требующих пространственного воображения.
3. Доступность в получении высшего образования, его коммерциализация, а также массовость и полифункциональность образовательных услуг привели к увеличению спроса на профессии гуманитарных областей – экономической и юридической сфер. Одновременно с этим значительно сократилось число желающих поступить на инженерные специальности, а уровень обучаемости студентов-технарей стал значительно ниже [11].
4. У поступающих в высшие технические заведения отмечается крайне невысокий уровень геометро-графической подготовки, а это вызывает неспособность к визуализации и пониманию проблем графической сферы. У поступающих крайне слабо развито образное мышление и абсолютно не сформировано пространственное воображение.

Уровень графической подготовки старшеклассников во многом влияет на их дальнейшее профессиональное развитие. Знание материалов черчения и графики являются необходимыми для получения практически всех технических профессий: инженера, электрика, технолога, строителя, механика, промышленного дизайнера, архитектора и многих других. Высокий уровень геометро-графической подготовки в рамках школьной программы позволяет сформировать у старшеклассников устойчивый интерес к инженерным и техническим профессиям, а также даёт возможность учащимся быть более успешным в профильных средне-профессиональных и высших учебных заведениях [1, 18]. Кроме того высокая начальная подготовка заведомо способствует повышению качества получаемого инженерного образования, а также позволяет внедрять инновации в учебные программы и благоприятствует решению проблем в подготовке инженерных кадров.

При оценке уровня графической подготовки необходимо руководствоваться более глубокими смысловыми понятиями [7, 12]. Достигнутый уровень должен определяться не техническим качеством выполнения графических работ, а подвижностью и развитостью пространственного и образного мышления, способностью определять и формулировать проблему, готовностью к мысленным трансформациям. Исследователями обоснованы положения, подтверждающие, что владение пространственным мышлением и умение оперировать пространственными образами лежат в основе достижения значительных результатов в профессиональной художественно-графической и конструкторско-технологической деятельности.

В прошлом образованию в области графической подготовки всегда придавалось существенное значение. Преподавать черчение на постоянной основе в нашей стране начали в начале XVIII века, в связи с промышленными и образовательными реформами Петра I. Как результат возникла целая плеяда высококвалифицированных специалистов, создавших условия для последующего промышленного развития России [15]. В советской школе техническому развитию

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

учащихся также уделялось должное внимание: изучение черчения было распределено на 2-3 учебных года, что способствовало формированию готовности школьников к изучению устройства машин и механизмов, устраняло вопросы профессиональной ориентации учащихся на инженерные специальности, а также способствовало пониманию и освоению смежных предметов – физики, геометрии и технологии [8].

Сегодня основным звеном преемственности геометро-графической образованности между средней и высшей ступенью являются подготовительные курсы или так называемы «инженерные» школы, сформированные на базе факультетов ВУЗов [17]. У обучающихся появляется возможность иметь в качестве учителей преподавателей университета, а соответственно наставников с более глубоким опытом преподавательской деятельности и техническим кругозором. Таким образом, устанавливаются более рациональные связи, учитывающие специфику дальнейшего обучения абитуриентов. Неслучайно, после прохождения подготовительных курсов, абитуриенты, как правило, лучше адаптированы к процессу обучения в ВУЗе.

Использование информационных и компьютерных технологий в рамках профильного довузовского образования не только способствуют формированию представления о будущей профессии инженера-конструктора, но и развивают первичные навыки и умения современной проектно-конструкторской деятельности, содействуют в профессиональном самоопределении личности учащегося, а также демонстрируют преимущества в доступности необходимых ресурсов для собственного профессионального развития [6, 16].

В рамках довузовской подготовки внимание уделяется и вопросам цифрового образования, новым моделям обучения на базе online-платформ, использованию адаптивных и мобильных приложений, и как следствие, внедрению гибридного и смешанного процессов обучения [5]. На сегодняшний день осуществляется активное реформирование образовательных программ, трансформирование классических образовательных моделей преподавания в e-Learning-модели и все больше используется электронный способ представления информации. Как показывает практика, в рамках такой цифровой модернизации учебного процесса визуальный способ усвоения информации становится более эффективным для процесса подготовки компетентных специалистов, особенно в сфере инженерного дела. Заметим, что для нынешних школьников информационно-коммуникационные технологии уже являются рабочим инструментом. Поэтому основополагающим ресурсом при формировании программ подготовки абитуриентов является соответствующая гибкая система образования, адекватная современному инновационному высокотехнологичному обществу.

3. Пример внедрения «доступной» графики в довузовскую подготовку

Идея повышения «доступности» графической подготовки абитуриентов включает в себя два понятия.

В первом случае *доступная графика* понимается как легкая для понимания визуальная составляющая информации, которая позволяет в довольно короткое время максимально точно понять суть преподаваемого материала. Кроме того новые наукоемкие технологии позволяют решать комплексные задачи в традиционных и смежных областях и формируют необходимость обновления парадигм в области инженерной деятельности и инженерного образования.

Для осуществления данной функции чаще всего используются мультимедийные возможности, видеоматериалы и оборудованные помещения. Наглядность и возможность личного участия позволяет максимально вовлечь обучающихся в процесс обучения, используемые интерактивные технологии 3D-моделирования и визуализации виртуальной реальности позволяют организовать образовательную программу учащихся с акцентом на гибкость и индивидуальные потребности обучаемых.

Во втором понимании *доступность графической подготовки* – это возможность в любом месте при необходимости получить доступ к нужной информации при минимальных трудовых и временных затратах.

Современные реалии стремительного внедрения цифрового обучения способствуют созданию образовательных программ, основанных на виртуальной и мобильной средах обучения. Внедрение дистанционного обучения (использование электронных ресурсов и облачных площадок) дает возможность совместного обсуждения возникающих вопросов и учебной работы в удобных пространствах независимо от местонахождения каждого участника. Это существенная поддержка образования, большая возможность для его совершенствования.

Предлагаемый ознакомительный курс «Основы инженерной графики». был организован по схеме смешанного обучения – совмещение очных занятий и дистанционной online-площадки. При таком подходе учебный процесс предусматривает наличие ряда образовательных модулей с интегрированной балльно-рейтинговой системой оценивания текущих результатов.

Организация сетевой площадки была осуществлена на базе электронного ресурса «Canvas», который представляет собой удобную облачную платформу для формирования развитой архитектуры курса с возможностями обеспечения учащихся необходимой теоретической информацией, выдачей практических заданий, предоставления иллюстративного материала и контроля полученных знаний.

При реализации курса обучения 2017/2018 проведение входного контроля выявило, что около 80% участников обладали нулевыми знаниями по изучаемому курсу и не имели представления об основных понятиях предметной области.

Проводимый курс был рассчитан на ознакомление учащихся с основными понятиями по дисциплине «Инженерная графика» и, безусловно, предусматривал начальную инструментальную подготовку учащихся в рамках раздела «Компьютерная графика» (работа с CAD-пакетом «Компас 17.1»). Стартовая страница ознакомительного курса представлена на рис. 1. Со стартовой страницы осуществляется полная навигация по курсу, отслеживание календаря занятий, доступ к теоретическому материалу, сопровождению практических работ, содержанию индивидуальных заданий и тестового контроля текущих знаний.

Рис. 1. Стартовая страница ознакомительного курса для школьников «Основы инженерной графики».

Выполняемые задания были направлены на становление у учащихся пространственно-образного мышления, формирование у учащихся навыков чтения

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

чертежей, анализ геометрических форм и технических конструкций, создание эскизов проектируемых деталей и геометрическое моделирование в технологии 3d.

Аудиторные занятия проводились в мультимедийной аудитории и компьютерном классе Аэрокосмического факультета ПНИПУ. Самостоятельная работа поддерживалась возможностями доступа учащихся в облачной сети. Проверка выполнения индивидуальных заданий осуществлялась на аудиторных и компьютерных занятиях.

Материалы по вопросам теории были представлены в формате документов, содержащих иллюстрации и необходимые текстовых описаний. Помимо основных источников в каждом модуле учащимся предоставлялись ссылки на дополнительные (достоверные!) источники информации.

С активным использованием компьютерных технологий при обучении необходимо совершенствовать мониторинговые процедуры контроля, а именно, организацию контрольных мероприятий и анализ получаемых данных. В рамках курса были разработаны тесты по всем тематическим модулям. Различные типы вопросов используемой площадки позволили в максимальной степени оценить понимание пройденного материала и выявить «слабые места» учащихся. Использовались вопросы типа «Верно/Неверно», «один правильный ответ», «соответствие частей» и др.

4. Заключение

Проблема совершенствования подготовки учащихся общеобразовательных школ для их дальнейшего профессионального становления в области инженерной деятельности является актуальной и своевременной и на сегодняшний день напрямую связана с использованием современных информационных и компьютерных технологий.

Описанный в статье пример внедрения «доступной» графики при подготовке старшеклассников к поступлению в высшие технические заведения в рамках разработанного курса «Основы инженерной графики» показал высокие результаты. Формирование у обучающихся геометро-графической компетентности реализуется за счёт комплексных обучающих заданий, сочетающих профессиональную направленность и инструментальную подготовку, что в свою очередь влияет на повышение уровня самостоятельности и самоорганизованности. Были созданы условия для проявления интеллектуальных способностей каждого обучаемого и его творческой активности. Широкое применение компьютерных технологий усилило позиции интегративности отдельных разделов графической подготовки и повысило эффективность образовательного процесса в целом.

Основной задачей проводимого эксперимента стало создание гибкого и индивидуально настраиваемого курса. Поэтому примечателен акцент на удобство использования электронной площадки для получения домашних заданий, их отправки и возможности изучения дополнительных материалов.

На рис. 2 представлены результаты входного и итогового контроля при оценке успеваемости контингента учащихся. Диаграммы показывают успешность обучения при практически полном отсутствии начальной графической подготовки обучаемых.

Кроме того, необходимо отметить положительную эмоциональную реакцию обучаемых на построение учебного процесса, их вовлеченность при приобретении знаний и активность при выполнении включенных в программу курса индивидуальных заданий, в том числе дополнительных повышенной сложности.

Рис. 2. Успеваемость учащихся про результатам прохождения ознакомительного курса

Главный результат – повышение мотивации учащихся к получению инженерного и технического образования, формирование перспективы будущего профессионального развития. Удовлетворение от посещения курса, удобства использования электронной площадки и 3D-моделирования объектов по достоинству отметили все участники. Используемые нововведения в методику обучения, безусловно, повлияли на формирование высокого уровня довузовской геометро-графической компетентности будущих абитуриентов

5. Литература

[1] Александрова Е.П. Практическая реализация проектно-ориентированной деятельности студентов в ходе графической подготовки / Е.П. Александрова, К.Г. Носов, И.Д. Столбова // Открытое образование. 2015. №5.

[2] Анисимова Г.А. К вопросу повышения уровня геометро-графической подготовки абитуриентов высших технических учебных заведений [Электронный ресурс] / Г.А. Анисимова, О.Г. Мелкумян, В.О. Москаленко // Наука и образование: электронный научно-технический журнал. 2013. №7.С.1-6.

[3] Выборова Н.Н. Формирование функционально-графической грамотности как необходимый компонент профессиональной подготовки студентов технических специальностей / Н.Н. Выборова, М.Ю. Пермякова // Педагогическое образование в России. 2016. №6.

[4] Губарева О.Н. Организационно-педагогическое проектирование деятельности довузовской подготовки абитуриентов в современных условиях: автореф. дис. … кан. пед. наук. Киров, 2012.

[5] Иванов В.Г. Инженерное образование в цифровом мире / В.Г. Иванов, А.А. Кайбияйнен, Л.Т. Мифтахутдинова // Высшее образование в России. 2017. №12 (218). С. 136-143.

[6] Иванов В.Г. Инженерное образование для гибкого, жизнеспособного и стабильного общества / В.Г. Иванов, А.А. Кайбияйнен, Л.Т. Мифтахутдинова // Высшее образование в России. 2015. №12 (218). С. 60-69.

[7] Князевский Д.А. Специфика довузовской подготовки школьников в учреждении дополнительного образования аэрокосмического профиля: дис. … канд. пед. наук. Ульяновск, 2005.

[8] Кроули Э.Ф. Переосмысление инженерного образования. Подход CDIO / Э.Ф. Кроули, Й. Малмквист, С. Остлунд, Д.Р. Бродер, К. Эдстрем; перевод с англ. С. Рыбушкиной; под науч. ред. А. Чучалина / М.: Изд. дом ВШЭ, 2015. 504 с.

[9] Лях В.И. Проектирование довузовской подготовки на основе идеологии CDIO // Высшее образование России. 2016. №2 (198). С. 68-74.

[10] Мишина Н.В. Методы развития графических умений подростков в процессе художественно-проектной деятельности [Электронный ресурс] / Н.В. Мишина //

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Современные проблемы науки и образования: электронный научный журнал. 2014. №4.
[11] Осипов П.Н. Инженерная педагогика: от сотрудничества к синергии // Высшее образование в России. 2017. №11 (217). С. 54-60.
[12] Охтя Н.В. Повышение объективности оценки знаний слушателей подготовительных куров САМ-ГТУ // Вестник Самарского гос. технического ун-та. Серия: Психолого-педагогические науки. 2010. №3. С. 139-144.
[13] Павлова А.А. Проблемы формирования и развития начальной графической грамотности школьника и студента-первокурсника / А.А. Павлова, Е.И. Корзинова // Преподаватель XXI век. 2012. №2.
[14] Пантелеева О.О. Результативные способы профессионально ориентированных форм и методов подготовки будущего специалиста на основе анализа довузовского образования строительного университета // Педагогическое образование в России. 2015. №8.
[15] Ройтман И.А. Методика преподавания черчения / И.А. Ройтман. М.:ВЛАДОС, 2002. 237 с.
[16] Столбова И.Д. Графическое образование как составляющая проектно-конструкторской подготовки специалиста / И.Д. Столбова, Е.П. Александрова, К.Г. Носов // Справочник. Инженерный журнал с приложением. 2017. № 3 (240). С. 40-46.
[17] Черных А.И. Современный рынок труда и довузовская подготовка школьников // Сибирский педагогический журнал. 2008. №4. С. 194-197.
[18] Шевченко О.Н. Довузовское геометро-графическое образование абитуриента в стратегии подготовки бакалавров технических направлений // ВЕСТНИК Оренбургского государственного университета. 2017. №5. С. 33-38 .

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Инструмент автоматизации создания и визуализации геометрии корпуса аудиосистемы рупорного типа

М. С. Тарских, А.В. Шкляр
tarskydesign@gmail.com | shklyarav@mail.ru
г. Томск, Томский Политехнический Университет

В данной работе предложен способ автоматизации процесса создания и изменения геометрии коридора тела рупора с возможностью визуализации и внесения корректировок геометрии в режиме реального времени.

Ключевые слова: *рупорная аудиосистема, компьютерная графика, геометрическая модель, автоматизация.*

Instrument for automating the creation and visualization of the geometry of the housing of a horn-type audio system

M. S. Tarskikh, A.V. Shklyar
tarskydesign@gmail.com | shklyarav@mail.ru
Tomsk, Tomsk Polytechnic University;

In this paper, we propose a method for automating the process of creating and changing the geometry of the corridor of the horn body with the possibility of visualization and making corrections to geometry in real time.

Keywords: *horn audio system, computer graphics, geometric model, automation.*

1. Введение

Tapered Quarter-Wave Pipe или Расширяющаяся Четвертьволновая Труба впервые описывается Paul Voight в 1930 г. Фактически, такого рода оформление является рупорным, но оно настроено на частоты резонансов динамической части головки. Серийное производство акустик типа TQWP принадлежало известной в кругах ценителей хорошего звучания франшизе Lauther, которая устанавливала в нее широкополосный динамик с высокой характеристикой чувствительности. Подобного рода решение обладает рядом преимуществ с электроакустической точки зрения. Аудиосистема (в дальнейшем АС) проектируется без разделительного фильтра и поэтому входные характеристики её импеданса совпадают с характеристиками импеданса головок. Это качество облегчает сопряжения систем с любым усилителем. Повышенная чувствительность обладает уменьшенной нагрузкой для маломощных усилителей, что крайне важно для усилителей лампового и транзисторного типа, особенно, если их работа происходит в А-классе. Также, наличие всего одной динамической головки улучшает ФЧХ системы, а это положительно отражается на локализациях источника звука и его возможности преобразовывать быстрый сигнал.

Следует отметить, что существует большое количество вариантов «сворачивания» коридора TQWP рупора.

Рисунок 1. Пример оформления TQWP

Расчет для дальнейшего изготовления конструкции подобного тела рупора – нетривиальная и крайне трудоемкая конструкторская задача, радикально увеличивающая окончательную стоимость изделия. В связи с чем появляется необходимость в создании инструмента, способного автоматизировать и упросить процесс расчета и создания корректной векторной геометрии.

2. Создание алгоритма реализации формообразования тела рупора в интерактивной трехмерной среде

Для создания трехмерной модели тела рупора с учетом факторов, отвечающих за формообразование, была выбрана программная среда Autodesk 3Ds Max.

Выбор данной программной среды был обусловлен возможностями и инструментами данного программного обеспечения, дающих возможность:

- Автоматизировать расчеты, необходимые для создания тела рупора в соответствии с ограничивающими факторами
- Просмотреть в интерактивной трехмерной среде выходные результаты расчетов с изменением входных данных в режиме реального времени
- Создать интуитивный и дружелюбный пользовательский интерфейс для управления входных и выходных данных
- Воплотить рассчитанную трехмерную модель тела рупора в виде высококачественного фотореалистичного изображения.

Автоматизация расчета тела рупора произведена с помощью встроенного в программную среду Autodesk 3Ds Max компонента MAXScript (Рисунок 2), позволяющего создавать скрипты на собственном языке программирования либо на языке программирования Python.

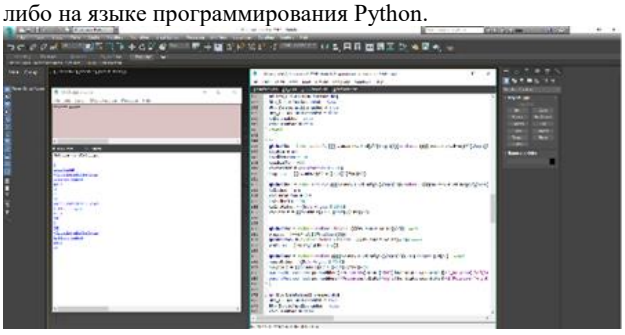


Рисунок 2. Интерфейс программного обеспечения Autodesk 3Ds Max с частью скрипта, выполненного в программной среде MAXScript

Входными данными для расчетов являются размеры наиболее важных элементов самого тела рупора и цифро-аналоговых преобразователей, расположенных в корпусе рупора, такие как:

- Габаритные размеры тела рупора

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

- Размер динамика, расположенного в начале коридора тела рупора
- Пропорциональное соотношение площадей начала коридора тела рупора и его выхода.

Далее, в скрипте происходят автоматические вычисления соотношений входных данных, приводящие к образованию «правильной» геометрии коридора рупора.

Также, опираясь выходные данные вычислений, скрипт автоматически подбирает необходимую толщину стенок тела рупора в соответствии с наиболее распространёнными на рынке толщинами листовых материалов, равных 6,10,12,16 и 22 мм.

3. Создание интерфейса управления интерактивной трехмерной моделью

Для получившегося скрипта необходимо запрограммировать визуальную оболочку с управляющими элементами, так же называемую пользовательским интерфейсом (Рисунок 3), для удобности обращения к скрипту с целью введения входной информации и произведения манипуляций с выходной расчетной информацией. Эта операция производится так же посредством написания программного кода в среде MAXScript.

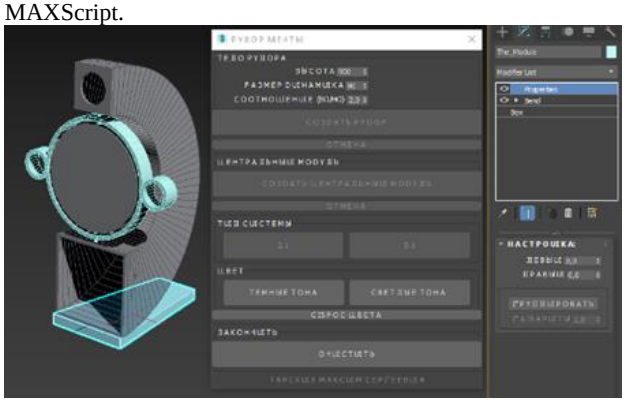


Рисунок 3. Пользовательский интерфейс

В ходе создания инструмента проектирования, были проанализированы основные факторы, отвечающие за формообразование тела рупора для рупорной акустической системы. В ходе аналитики принято решение использовать уменьшение габаритных размеров коридора движения звукового давления посредством сворачивания рупора в спираль (применённый в данном примере тип геометрии спирали рупора именуется «подкова»). Была рассчитана геометрия образования «правильной» формы коридора с учетом характерных особенностей конструкции подобного корпуса для аудиосистем с последующей автоматизацией обработки входных данных.

Дополнительно были внесены алгоритмы, отвечающие за толщину стенок корпуса, т.е. за ступенчатое изменение размеров, равных 6,10,12,16 и 22 мм и соответствующих наиболее распространённым предложениям на рынке листовый пиломатериалов, в зависимости от габаритов, соотношения протяженности участков коридора рупора к площади боковой стенки.

Написанный скрипт был привязан к интерфейсу управления, интуитивно понятному и позволяющему пошагово создавать тело рупора с вводом входных данных и изменением выходных данных в реальном времени с последующим отображением изменений на трехмерной цифровой модели.

Данное решение позволяет значительно сократить время на создание виртуальной модели корпуса рупорного типа для аудиосистемы с сохранением «правильной геометрии», последующего редактирования получившейся формы.

Также, реализация такого решения позволяет наглядно рассмотреть со всех сторон любую форму корпуса, являющуюся продуктом действия предложенного алгоритма с возможностью максимально представить будущее изделие, в будущем изготавливаемое на производстве благодаря фотореалистичному рендеру.

Рисунок 4. Результаты использования созданного инструмента формообразования в виде разногабаритных корпусов с различающимися входными значениями.

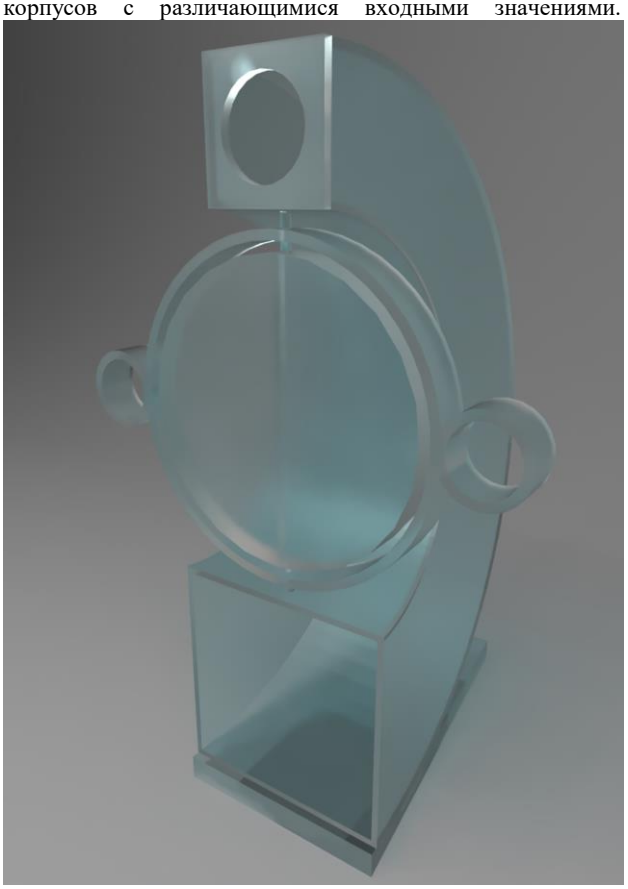


Рисунок 5. Материальная визуализация модели с малым количеством входных данных, описывающих геометрию коридора рупора.

Введение дополнительных информационных блоков дает неограниченное уточнение сложной геометрии рупора проектируемой аудиосистемы, наравне с усложнением геометрии, описывающей опорно-габаритные функции, позволяя в дальнейшем внедрить в программу заготовленные дизайнерами и инженерами варианты оформления конструкционных решений, несвязанных с расчетом внутренней геометрии корпуса АС.

Также, использование программной среды Autodesk 3Ds Max дает возможность конвертации синтезированных объектов в другие программные среды для дальнейшей конструкторской работы.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

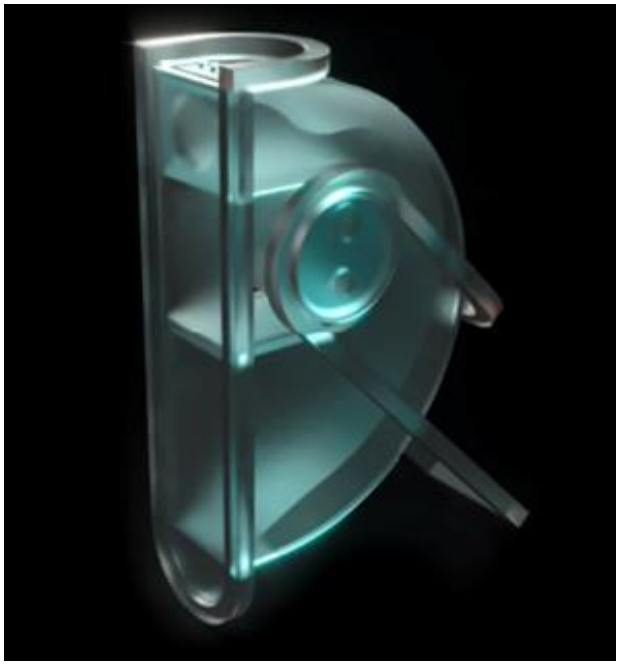


Рисунок 6. Результат введения дополнительных входных данных и углубленной дизайнерско-конструкторской проработки конструкции, созданной посредством инструмента автоматизации формообразования, представленный в виде материальной визуализации.

4. Практическое применение результатов работы созданного инструмента

Данные, поученные в ходе использования инструмента для автоматизации создания геометрии коридора TQWP рупора для АС, можно применять не только с целью наглядной визуализации проекта.

Посредством конвертации выходных данных, описывающих конструкционные элементы, такие, как векторная геометрия корпуса, либо полная трехмерная модель, в распространённые цифровые форматы (DFX, DWG, OBG, STL и т.п.) можно продолжить работу с проектом, например, создав необходимую конструкторскую документацию, либо проведя физические исследования в программных средах, позволяющих проводить соответствующие симуляции.

Также, наиболее полезным аспектом является возможность прямой конвертации для создания управляющей программы для станков с ЧПУ с целью дальнейшего материального воплощения проекта.

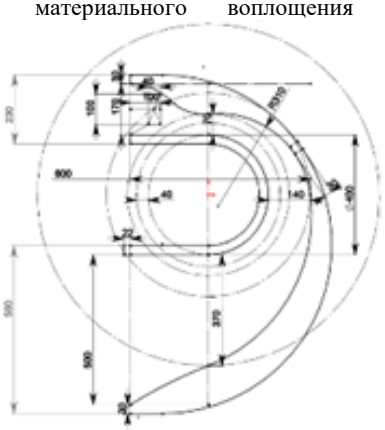


Рисунок 7. Анализ векторной геометрии в программной среде SolidWorks в комплектном дополнении физических симуляций с целью проверки корректности построения геометрии коридора рупора.

5. Реализация проекта

Созданный инструмент автоматизации проектирования позволил сократить большое количество трудоемких процессов, давая возможность сразу приступить к изготовлению опытного образца на ЧПУ платформе.

Рисунок 8. Процесс создания проекта, как результата использования разработанного инструмента автоматизации.

Также, выходные данные в виде трёхмерных моделей позволяют создать полноценную высококачественную фотореалистичную визуализацию проекта, интегрированного в интерьерную среду для дальнейшего их использования в рекламной и полиграфической продукции.

Рисунок 9. Использование выходных данных в полиграфии

7. Заключение

Использование описанной в данной статье разработки в виде инструмента, позволяющего автоматизировать и упростить многие трудоемкие конструкторские задачи, позволит оптимизировать дальнейшее производство изделий подобного рода, сократив время, необходимое на расчет конструкции, создание конструкторской документации и общие финансовые затраты фирмы-изготовителя, связанные с данным родом интеллектуальной деятельности.

Вышеперечисленные положительные аспекты также могут сказаться на уменьшении стоимости изделия для конечного потребителя, а, следовательно, на глобальной популяризации изделия типа рупорная TQWP аудиосистема.

8. Литература

1. Акустическая_система [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Акустическая_система –. Заглавие с экрана. –(Дата обращения: 14.01.2018).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

2. Эффект_Доплера [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Эффект_Доплера –. Заглавие с экрана. –(Дата обращения: 14.01.2018).

3. Диффузоры_Шрёдера [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Диффузоры_Шрёдера –. Заглавие с экрана. –(Дата обращения: 14.01.2018).

4. Акустика. Типы акустических систем - AB Клуб [Электронный ресурс]. – Режим доступа: http://www.avclub.pro/articles/audio-video-ot-a-do-ya/akustika-tipy-akusticheskikh-sistem –. Заглавие с экрана. –(Дата обращения: 20.04.2018).

5. Информационный ресурс основ акустики [Электронный ресурс]. – Режим доступа: http://baseacoustica.ru/akustika/4-materialy-korpusov/4-akusticheskoe-oformlenie.html –. Заглавие с экрана. –(Дата обращения: 20.04.2018).

6. Плиты древесно-стружечные. Технические условия [Электронный ресурс]: ГОСТ 10632-2014. – Введ. 01.07.2015. — Режим доступа: Система Кодекс-клиент.

9. Об авторах

Тарских Максим Сергеевич, студент инженерной школы информационных технологий и робототехники (бывш. Институт кибернетики) Томского политехнического университета. E-mail tarskydesign@gmail.com

Шкляр Алексей Викторович, старший преподаватель, отделение автоматизации и робототехники, Томского политехнического университета. E-mail shklyarav@tpu.ru

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Базовые принципы преподавания инженерной и компьютерной графики будущим специалистам.

Д.Е. Тихонов – Бугров, С.Н. Абросимов, В.А. Дюмин,
tbde@mail.ru|abr-sn@yandex.ru|diumin.v@yandex.ru
Балтийский государственный технический университет “Военмех», Санкт-Петербург, РФ

Показано, что обучение инженерной и компьютерной графике должно быть построено как проектное с максимально возможным многообразием форм. Отмечается, что кафедры, преподающие графические дисциплины, не должны терять теоретическую базу, сбиваться на обучение работе с графическими пакетами как инструментом. Декларируется необходимость включения в учебный процесс основ инженерного творчества и эвристических приёмов. Отмечается особая важность соревновательного характера учебного процесса, в частности, олимпиад. Показаны недостатки региональной олимпиады в Санкт-Петербурге.

Ключевые слова: *инженерная и компьютерная графика, проектное обучение, конструирование, САПР, олимпиада*

Basic principles of engineering and computer graphics teaching to future professionals

D.E. Tikhonov – Bugrov, S.N. Abrosimov V.A. Diumin.
tbde@mail.ru|abr-sn@yandex.ru|diumin.v@yandex.ru
Baltic State Technical University “Voenmeh”, Saint-Petersburg, Russian Federation

It is shown that engineering and computer graphics training should be built as a project with the greatest possible variety of forms. It is noted that the departments teaching graphic disciplines should not lose theoretical basis, be bogged down at learning to work with graphic packages as a tool. The need to include in the educational process the foundations of engineering creativity and heuristic methods is declared. The special importance of the competitive nature of the educational process, in particular, the Olympiads, is noted. The shortcomings of the regional Olympiad in St. Petersburg are shown.

Key words*: engineering and computer graphics, project-based teaching, design engineering, CAD, olympiads.*

Аксиомой является тот факт, что для обеспечения качественной графической подготовки таковая должна быть сквозной.

В этом случае на кафедры, обеспечивающие графическую подготовку, ложится обязанность обеспечения не только базовой подготовки, но и сопровождения конструкторских разработок студентов вплоть до окончания обучения. Очевидно, что такие кафедры не должны быть гибридом общетехнических дисциплин (теоретическая механика, сопротивление материалов, материаловедение, графика, детали машин…) и должны иметь хороший контакт с профилирующими кафедрами. Очень важна и политика руководства вуза, связанная с обеспечением качества дипломных работ и проектов в графическом отношении, в частности.

Однако, изменение ряда образовательных стандартов, исключивших начертательную геометрию из инженерного образования, ставят кафедры, ведущие графическую подготовку в сложную ситуацию, о которой говорится в статье [7]. Авторы отмечают, что описанная выше ситуация лишает научности дисциплины кафедры и ставят вопрос о принципиальном изменении содержания и статуса дисциплины инженерная и компьютерная графика, и статуса самой кафедры в том случае, если научность будет потеряна окончательно.

Действительно, назрела необходимость рассматривать проектирование и документирование как единый процесс. Если в настоящее время обучение сводится к обучению работе с графическим пакетом, как инструментом, то логично поставить вопрос о существовании кафедры, занимающейся, по сути, инструментальной подготовкой среднего технического персонала.

О том, что именно так примитивно часто строится учебный процесс обучения компьютерной графике, свидетельствуют публикации [9]. В ней и в подобных публикациях, рассматривается практически техника работы в пакетах КОМПАС-3D [2], AutoCAD, Solid Works и др.

Попытки введения в дисциплину «Инженерная и компьютерная графика» вопросов, связанных с теорией проектирования часто вызывают возражения со стороны профильных кафедр, которые, тем не менее, обычно обучают проектированию только конкретных изделий. Как справедливо замечено в [7], должна существовать дисциплина, формирующая общий подход к проектированию и конструированию, задающая общие рамки, общие требования, общие постановки, развивающие философию проектирования. В связи с этим, авторы напоминают о технетике, как науке о документированной реальности, о технической форме существования материи, о её идеальном отображении.

Неким приближением к заявленной дисциплине должно являться постепенное включение в неё, связанную с компьютерной графикой, проектного обучения с подключением методологии инженерного творчества, эвристических приёмов [8], необходимости стремления к «идеальному конечному результату»[1].

Среди большого количества проблем, стоящих перед современным учебным процессом, связанным с компьютерной графикой, выделим проблему обеспечения его проектного характера и такую важную составляющую технологии обучения, как соревновательность. Убеждены, что отмеченные вопросы решающим образом влияют на качество обучения.

Элементы проектного обучения мы включаем уже на начальном этапе обучения. В разделе «Проекционная графика» студент решает задачу модификации стандартной детали общемашиностроительного применения [6]. Например, ставится задача: используя стандартную деталь подкласса 4200, модифицировать её за счёт введения шпоночного паза или лыски, что требует анализа

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

соответствующих стандартных элементов, конструктивных нюансов.

В рабочей тетради, в свою очередь, содержатся задачи на формирование чертежа простой детали с заданными параметрами (толщина стенки, размер под ключ, габариты и.т.п.).

На следующей стадии используется широкий спектр конструкторских заданий:

- Создание модели сравнительно простого устройства по словесному описанию и теоретической схеме (рис. 1).
- Доконструирование – создание детали, отсутствующей в реальной сборочной единице, формирование модели сборки в полном комплекте (рис. 2а), распечатка детали на 3D-принтере, проверка собираемости (рис.2б).

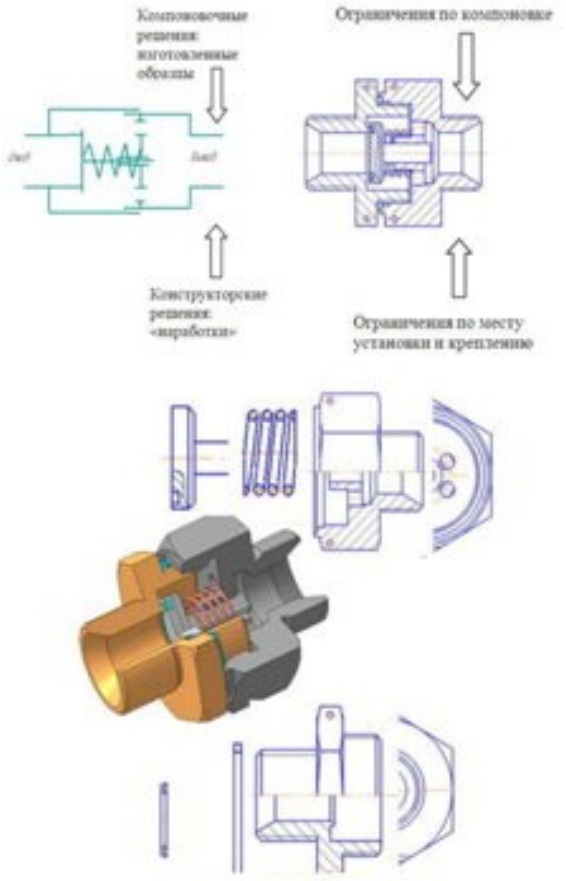


Рис.1.Создание модели по словесному описанию и теоретической схеме.

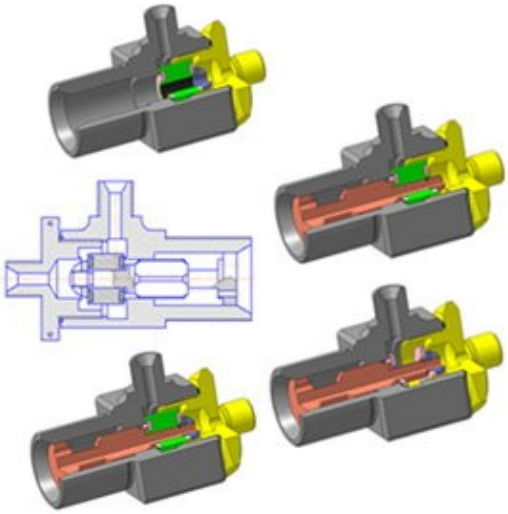
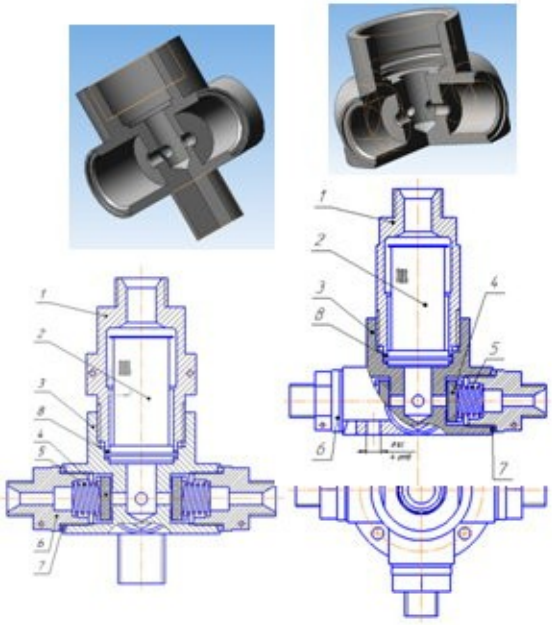


Рис.2а. Доконструирование.

Рис. 2б. Распечатка деталей на 3-D принтере и проверка собираемости.

- Переконструирование - процесс выявления недостатков конструкции и выбор необходимых решений для их ликвидации. Задания данного характера являются базой для построения учебного процесса, о котором говорилось выше – приближением к дисциплине с условным названием «Основы инженерного проектирования». На данной стадии студент знакомится с основами инженерного творчества, ТРИЗ. Задания первого уровня содержат «подсказки» где искать недостатки конструкции. Задания второго уровня предполагают самостоятельный поиск изъянов конструкции, работу с прототипами, информационный



поиск (рис.3а и рис.3б).

Рис.3а. Изменения узла крепления корпуса.

Рис.3б. Изменение габаритов и узла крепления изделия.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

- Особое место занимает конструирование надувных и воздухоопорных сооружений т.к. наиболее убедительно демонстрирует студенту значение начертательной геометрии в процессе проектирования. Выполняются модели сооружений, выкройки, распечатки на 3D принтере в уменьшенном масштабе.

Важным этапом развития описываемого учебного процесса и его логического продолжения является дисциплина Основы САПР, читаемая, в зависимости от специальности, на 3, 4 или 5 семестрах. Ставится задача сформировать у студента понимание основ формирования системы автоматизированного проектирования, закрепить владение построением трёхмерных моделей объектов, научить ориентироваться в программно-аппаратных средствах автоматизированного проектирования.

Студент должен усвоить, что эффективность использования аппаратно-программной среды определяется и сервисом, и хорошим интерфейсом, приближающим пользователя непосредственно к предметной области разработки.

На наш взгляд, технологии обучения, всегда игравшие значительную роль, в настоящее время приобретают роль ключевую. Моделирование отношений в творческих коллективах, в том числе конкурентных, задача весьма актуальная.

В перечне компетенций в последнее время всё чаще выделяют умение работать в коллективе, как важный аспект профессиональной деятельности инженера. В этом отношении мы уделяем внимание бригадной форме работы над заданием, в случае его значительного объёма. При этом имеет место соревнование между «творческими коллективами», итоги которого определяются при защите проектов. Соревновательность в учебном процессе – важный аспект, своего рода модель реальных отношений в проектных организациях. Стремление к повышению квалификации, к победе в профессиональных конкурсах – естественная деятельность перспективного специалиста.

По этой причине активно развиваем рейтинги, внутренние олимпиады, участвуем во всероссийских мероприятиях среди которых необходимо выделить детище В.И.Вышнепольского – Всероссийскую открытую олимпиаду по начертательной геометрии, инженерной и компьютерной графике[5]. В этом комплексном, системном соревновании превалирует творчество и что очень важно и для воспитательного процесса, объективность, даже щепетильность в оценке результатов.

Отбор на всероссийские олимпиады по положению должен осуществляться на олимпиадах региональных. К сожалению, в Санкт-Петербурге региональная олимпиада по компьютерной графике начисто не выполняет своих функций, что, по нашему мнению, наносит большой ущерб развитию учебного процесса в данной области. Ведь очевидно, что олимпиада – школа творчества, существенно влияющая на развитие учебного процесса в вузах, принимающих активное в ней участие.

Этот факт подтверждается многолетним опытом проведения городской олимпиады по начертательной геометрии, опытом участия в упомянутой выше Всероссийской открытой олимпиаде. Эти мероприятия повлияли и постоянно влияют на содержание учебного процесса, технологию обучения, в том числе, одарённых студентов, что представляется особенно важным.

Однако региональная олимпиада Санкт-Петербурга по компьютерной графике является типичным соревнованием «инструментального» характера, в котором в готовый сборочный чертёж необходимо ввести крепёж и выполнить модели и чертежи двух деталей и аксонометрию. Творческий элемент отсутствует начисто, а содержание ориентировано на программу обучения в вузе-организаторе, где конструирование заключается в конструировании корпуса прибора или оси ротора. Кредо организаторов этой олимпиады изложено в[3], где откровенно заявлено, что «оптимизация» обучения заключается в исключении из обучения элементов начертательной геометрии, лекций по компьютерной графике в угоду более обстоятельному формированию умений по созданию твёрдотельных моделей.

Анализ результатов региональной олимпиады Санкт-Петербурга с 2000 по 2017 годы показывает следующее: за этот период команда, представителем которой является председатель жюри, побеждала 9 раз и не была призёром только один раз, не смотря на то обстоятельство, что в данном вузе нет кафедры графики как таковой и на изучение данной дисциплины отведен малый ресурс. Данный факт объясняет претензии ряда вузов к содержанию и объективности при проведении олимпиады, и критериям оценки результатов. Эти критерии изложены в [4]. Значительная их часть сформулирована таким образом, что их можно трактовать, так как это выгодно заинтересованной стороне.

В течение ряда лет в олимпиаде в знак протеста не участвовали крупнейшие инженерные вузы, где компьютерной графике придаётся большое значение -университеты: «Горный», Политехнический, Военмех.

Количество участников за последние пять лет не превышает семи вузов среди которых отсутствуют «Горный» и Военмех, представители которых доказали свою компетентность на всероссийских олимпиадах. А вот победители Санкт-Петербургской олимпиады (за исключением бойкотирующих её Горного и Военмеха) ни разу не участвовали во всероссийских олимпиадах, вероятно понимая, что там другой уровень.

В данной ситуации важную миссию по восстановлению статуса олимпиады по компьютерной графике, развитию её содержания, коллективного обсуждения проблем среди профессионального сообщества на базе Дома Учёных им. Горького, взял на себя ГУТ. Идейным вдохновителем интересной олимпиады «Тотальный чертёж», которая становится всё более популярной, стал профессор Волошинов В.А.

По нашему убеждению, для повышения уровня подготовки инженеров в области компьютерной графики необходимо развитие кафедр, осуществляющих данную подготовку в области общей методологии проектирования технических объектов, построение учебного процесса на базе тотального проектного обучения, активизации элементов соревновательности, развития олимпийского движения, построенного, на всех уровнях, таким образом, чтобы оно стало катализатором эвристического мышления студентов.

Необходимо, как минимум, сформировать следующую устойчивую цепочку учебного процесса: геометрические основы формирования моделей и ортогонального чертежа, стандарты ЕСКД; основы инженерного проектирования, основы технологии изготовления изделий (1-3 семестры обучения); изучение основ САПР (3-4 семестры); тотальный нормоконтроль конструкторской документации при выполнении курсовых (желательно межпредметных) и дипломных разработок.

Заключение

Отмеченные принципы и подходы уже апробировались в учебной практике и показали свою целесообразность. А соревновательность придает учебному процессу то, чего сейчас не хватает технически ориентированной молодежи и

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

необходимо в становлении современного технического специалиста. Роль грамотно построенных предметных олимпиад в этом смысле трудно переоценить.

Литература

[1] Г.Альтшуллер. Найти идею. М.Альпинапаблишер. 2015.

[2] В.П.Большаков, А.В.Чагина. Тестовые задания по основам трёхмерного моделирования// Геометрия и графика.-2016.-Т.4.-№4.-С.60-67.

[3] В.П.Большаков. Приближение содержания курса «Инженерная и компьютерная графика» к уровню развития современныхCAD-систем. Материалы 24 международной научно-методической конференции «Современное образование: содержание, технологии, качество»,СПб.-2018. Т.1.-С.119-122.

[4] В.П.Большаков, И.В.Климов, А.В.Чагина. Содержание сайта сопровождения региональной студенческой олимпиады по инженерной и компьютерной графике// Геометрия и графика.- 2017.-Т.5.-№4.-С. 83-93.

[5] В.И.Вышнепольский. Открытая Всероссийская олимпиада по начертательной геометрии, инженерной и компьютерной графике 2015г. //Геометрия и графика.-2016.-Т.4.-№1.-С.-73-89.

[6] Иллюстрированный определитель деталей общемашиностроительного применения. Руководящий технический материал. Издательство стандартов. М. 1977.

[7] М.Н.Лепаров, М.Х.Попов. Инженерная графика – ТО BE, OR NOT TO BE.Материалы 3 международной научно-практической конференции «Проблемы качества графической подготовки».- Пермь, 2012.

[8] А.И.Половинкин. Основы инженерного творчества. М. Машиностроение. 1988.

[9] М.М.Харах, И.А.Козлова, Б.М.Славин. Конструирование сборочного чертежа изделия методом моделирования как завершающий этап изучения инженерной и компьютерной графики. //Геометрия и графика.-2014.-Т.2.-№3.-С.36-40.

Об авторах

Тихонов-Бугров Дмитрий Евгеньевич, к.т.н., доцент, зав. кафедрой «Инженерная и машинная геометрия и графика». Балтийского государственного технического университета «Военмех», г. Санкт-Петербург. Его e-mail tbde@mail.ru.

Абросимов Сергей Николаевич, к.т.н., профессор кафедры «Инженерная и машинная геометрия и графика». Балтийского государственного технического университета «Военмех» , г. Санкт-Петербург.
Его e-mail abr-sn@yandex.ru.

Дюмин Владимир Андреевич, к.т.н., доцент кафедры «Инженерная и машинная геометрия и графика». Балтийского государственного технического университета «Военмех», г. Санкт-Петербург.
Его e-mail diumin.v@yandex.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Синтез предварительной каркасной модели 3D-объекта по техническому чертежу, содержащему разрезы и сечения

В.А. Тюрина[1], М.М. Смычек [2], С.И. Ротков[3], Т.В. Мошкова[4]
55555_73@mail.ru | mariasmychek@gmail.com | rotkovs@mail.ru | ng.forever.mtv@gmail.com
[1]ННГАСУ, Нижний Новгород, Россия;
[2]ННГАСУ, Нижний Новгород, Россия;
[3]ННГАСУ, Нижний Новгород, Россия;
[4]ННГАСУ, Нижний Новгород, Россия.

Проблема синтеза 3D-модели по чертежу относится к разряду сложных для формализации задач и, как показывает анализ литературы, алгоритмы, реализующие учёт всей информации, содержащейся на техническом чертеже, до сих пор не разработаны. Каркасные модели, полученные известными способами автоматического синтеза, часто несут в себе ложную геометрическую информацию, что существенно затрудняет дальнейшее преобразование модели и её последующую визуализацию. Этап формирования каркасной модели является одним из наиболее значимых в задаче восстановления объекта. Использование всей многоплановой информации, содержащейся на техническом чертеже, является сложным процессом и с точки зрения алгоритмической реализации. Существующие алгоритмы решения задачи имеют ограничения на входную информацию и ни один из них не предусматривает работу с чертежом, выполненным в соответствии с требованиями ЕСКД. Ввиду необходимости преобразования архивов чертежей на электронных и бумажных носителях в электронные 3D-модели исследования по расширению видов входной информации в задаче синтеза весьма актуальны. В данной статье представлены результаты исследования, направленного на разработку алгоритмов, реализующих автоматизированное чтение чертежей, имеющих в своём составе разрезы и сечения.

Ключевые слова: *электронная геометрическая модель изделия, каркасная модель, предварительная каркасная модель, технический чертеж, разрезы и сечения, автоматизированное проектирование.*

Synthesis of the preliminary wireframe model of the 3D object according to the technical drawing, containing incisions and sections

V.A. Tyurina [1], M.M. Smychek [2], S.I. Rotkov[3], T.V. Moshkova[4]
55555_73@mail.ru | mariasmychek@gmail.com | rotkovs@mail.ru | ng.forever.mtv@gmail.com
[1]NNGASU, Nizhny Novgorod, Russia;
[2]NNGASU, Nizhny Novgorod, Russia;
[3]NNGASU, Nizhny Novgorod, Russia;
[4]NNGASU, Nizhny Novgorod, Russia.

The problem of synthesis of the 3D model according to the drawing belongs to the category of difficult tasks for formalization and, as analysis of the literature shows, algorithms that implement the accounting of all information contained in the technical drawing have not yet been developed. Wireframe models obtained by known automatic synthesis techniques often carry false geometric information, which makes it difficult to further transform the model and its subsequent visualization. The stage of the formation of the wireframe model is one of the most significant in the object restoration task. The use of all the multifaceted information contained in the technical drawing is a complex process from the point of view of algorithmic implementation. The existing algorithms for solving the problem have limitations on the input information and none of them provides for working with a drawing executed in accordance with the requirements of the ESKD. In view of the need to convert the archives of drawings on electronic and paper media into electronic 3D models, studies on expanding the types of input information in the synthesis problem are very relevant. This article presents the results of a study aimed at developing algorithms that implement automated reading of drawings that have incisions and sections in their composition.

Keywords: *electronic geometric model of the product, wireframe model, preliminary wireframe model, technical drawing, incisions and sections, computer-aided design.*

1. Введение

Практически все системы геометрического моделирования позволяют осуществлять генерацию чертежей по трёхмерной модели. Однако получение трёхмерной модели по чертежу вызывает у проектировщиков затруднения, связанные с отсутствием математического и программного обеспечения.

Создание системы, выполняющей автоматическую реконструкцию трёхмерных моделей по техническому чертежу, позволило бы во много раз сократить время проектирования объектов различного назначения. Поэтому разработка алгоритмов для функционирования системы автоматической реконструкции трёхмерных моделей объектов по техническому чертежу является весьма актуальной задачей для разработчиков САПР.

Проблема синтеза 3D модели объекта по ортогональным проекционным изображениям решалась в работах В.С. Полозова [3], С.И. Роткова [4], G. Markowsky, M.A. Wesley [5], а также других авторов.

Однако существующие алгоритмы решения задачи имеют ограничения на входную информацию и ни один из них не предусматривает работу с чертежом, выполненным в соответствии с требованиями ЕСКД (Единой Системы Конструкторской Документации).

Каркасные модели, полученные известными способами автоматического синтеза [2, 4, 5], часто несут в себе ложную геометрическую информацию, что существенно затрудняет дальнейшее преобразование модели и её

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

последующую визуализацию. Этап формирования каркасной модели является одним из наиболее значимых в задаче восстановления объекта. Использование всей многоплановой информации, содержащейся на техническом чертеже, является сложным процессом с точки зрения алгоритмической реализации. Ввиду необходимости преобразования архивов чертежей на электронных и бумажных носителях в электронные 3D-модели востребованы исследования по расширению классов обрабатываемой входной информации, в том числе необходима разработка алгоритмов, реализующих автоматизированное чтение чертежей, имеющих в своём составе разрезы и сечения.

Классический подход к программной реализации алгоритма восстановления оригинала по его проекционным изображениям состоит в его разбиении на шесть основных шагов [3]:

1) анализ изображений;
2) создание массивов 3-х мерных координат;
3) создание каркасной модели;
4) анализ ложных геометрических элементов;
5) создание граничной модели;
6) создание конструктивной модели.

В данной статье представлены результаты исследования и разработки алгоритма, для случая, когда технический чертеж содержит в своем составе простые разрезы.

2. Общая схема анализа изображений технического чертежа, включающих разрезы и сечения

В основе разработки общей схемы автоматического анализа изображений технического чертежа, включающих разрезы и сечения, лежит ГОСТ 2.305-2008 «Изображения – виды, разрезы, сечения». В соответствии с методом эвристического моделирования (процесс описания поведения человека в ходе решения неформальной задачи) были проведены эксперименты с «чтением» технического чертежа. Такой подход делает возможным выявить исходные данные задачи и наложенные на них ограничения, определить признаки правильности решения и т.д.

С точки зрения возможности восстановления трёхмерных образов геометрических объектов наиболее значимыми являются следующие изображения:

1) основные виды;
2) сечения;
3) разрезы.

Стандартами предусмотрено формирование чертежей при условии минимального количества изображений, входящих в чертёж. Напомним, что в процессе создания чертежа формируется параметрический граф [3], включающий в себя размеры и геометрические условия, позволяющие заменить размеры. Выбор изображений оригинала осуществляется так, что на них в совокупности отображаются без искажений все элементы параметрического графа. Для минимизации количества изображений на каждом из них должно присутствовать максимально возможное количество элементов графа, показанных без искажения. Очевидно, что чем сложнее геометрическая форма объекта, тем большее количество изображений наносится на чертёж при его формировании. Для несложных объектов бывает достаточно двух или трёх основных видов, а для некоторых тел требуется добавление и разрезов, и выносных и местных видов, и сечений.

Основные изображения являются наиболее важной и обязательной информацией для получения 3D образов проектируемых изделий и деталей, а вспомогательные изображения – это дополнительная информация, которая требует уточнений представления о проектируемом объекте.

Многообразие геометрических форм предметов окружающего нас мира определяет высокую сложность формализации задачи автоматического чтения полного многовидового технического чертежа. Алгоритмы, позволяющие считывать и анализировать всю возможную графическую информацию, встречающуюся на чертежах, пока не созданы. Поэтому основная работа ведётся в области расширения классов графической информации, доступной для обработки. Этот процесс делает возможным автоматический синтез всё более сложных геометрических объектов.

В данной работе предлагается расширить спектр обрабатываемой входной информации, содержащейся на техническом чертеже, а именно, за счёт возможности анализировать информацию о разрезах и сечениях объекта. В результате проведенного исследования было выявлено, что формирование предварительной каркасной модели по чертежу, на котором присутствуют разрезы, дает возможность получать предварительную каркасную модель с большим количеством истинных элементов (вершин и рёбер), и позволяет сократить количество процедур по распознаванию и удалению ложных геометрических элементов, присутствующих в предварительной каркасной модели.

Такой подход был применен для разработки общей схемы алгоритма чтения чертежа, содержащего разрезы и сечения [1].

Проведённые наблюдения за поведением человека, перед которым стоит задача «прочитать» чертёж, позволили выделить следующие признаки, присутствующие на чертеже и позволяющие выявить изображения, на которых присутствуют разрезы и сечения:

1) наличие на чертеже объекта «Штриховка»;
2) наличие на чертеже буквенно-цифровых обозначений типа «А-А» или «1-1»;
3) наличие на изображениях чертежа линий обрыва;
4) наличие на чертеже условного графического обозначения «повернуто»;
5) наличие на чертеже изображения в разрыве между частями одного и того же вида;
6) наличие на чертеже изображения, связанного с видом нормальной штрихпунктирной тонкой линией.

Для определения типов разрезов и сечений человек, осуществляющий «чтение» чертежа, в первую очередь обращает внимание на следующие признаки:

1) способ оформления области штриховки;
2) обозначения секущей плоскости;
3) как расположен след секущей плоскости относительно системы координат;
4) какова структура следа секущей плоскости;
5) количество секущих плоскостей.

Анализ данных признаков позволяет определить виды разрезов и сечений, присутствующих на чертеже, понять их назначение и, в результате, получить полное представление об особенностях формы изображенного на чертеже объекта.

Электронный чертёж объекта, выполненный в той или иной системе геометрического моделирования, позволяет получить информацию о приведённых выше признаках в формальном (цифровом) представлении и даёт возможность после процедур преобразования полученных данных перейти к построению предварительной каркасной модели объекта в автоматизированном режиме.

Работа над программной реализацией описанного подхода к решению поставленной задачи требует подробного эвристического анализа каждого шага алгоритма и приведения его к наиболее полному

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

формальному виду. Процедуры преобразования данных электронного чертежа объекта могут отличаться конкретной формой реализации, зависящей от вида системы, используемой для построения чертежа, но общая схема синтеза электронной 3D модели от вида системы не зависит и в перспективе может быть использована и для работы с растровыми архивными изображениями.

3. Алгоритм синтеза предварительной каркасной модели по чертежу, имеющему в своём составе простые разрезы

Изображение является сложной графической структурой, в состав которой входят более простые графические элементы, а именно – отрезки, дуги, окружности, области штриховки, буквенно-цифровые обозначения типа «А-А» и «1-1» и т.д. Эти простые элементы выполняются согласно «Общим правилам выполнения чертежей и содержат в себе данные об изображённых деталях и изделиях, в том числе и информацию, позволяющую распознавать и восстанавливать 3D-объект.

Разработанный алгоритм синтеза ПКМ по техническому чертежу продемонстрирован на двух примерах. Для визуализации алгоритма разработана блок-схема, представленная на рис. 1.

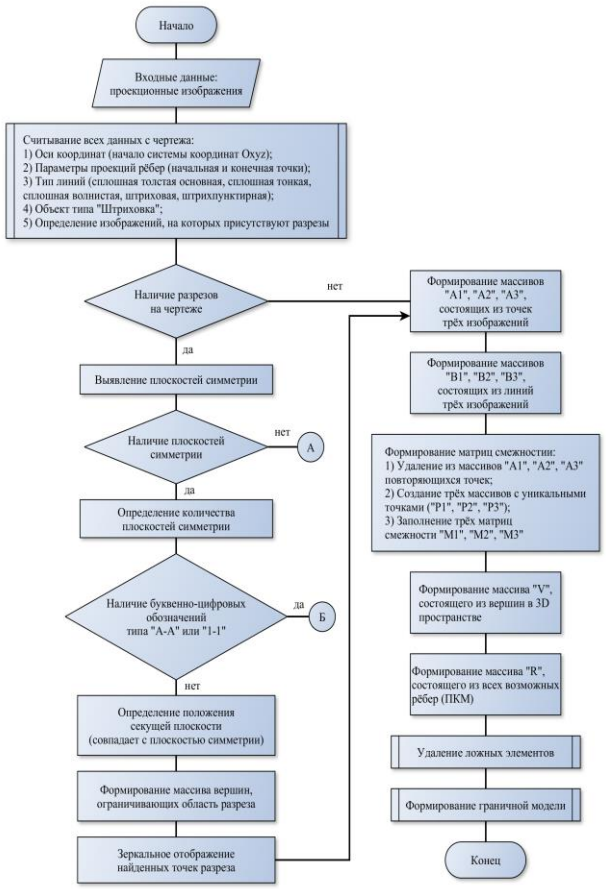


Рис. 1. Блок-схема процесса «чтения чертежа» с простыми разрезами.

Объект, на примере которого демонстрируется работа алгоритма, специально выбран простым для простоты восприятия и представляет собой четырёхугольную призму с цилиндрическим отверстием. На рис. 2 изображены три проекционных изображения: «вид спереди», совмещённый с простым разрезом, «вид сверху» и «вид слева», составляющие в совокупности входную графическую информацию.

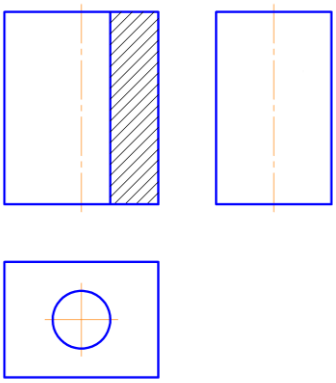


Рис. 2. Пример проекционных изображений с разрезом, совмещённым с главным видом.

После получения входной электронной графической информации происходит считывание всех данных с технического чертежа, а именно:

1) оси координат (начало системы координат Oxyz);
2) параметры проекций рёбер (начальная и конечная точки);
3) тип линий (сплошная толстая основная, сплошная тонкая, сплошная волнистая, штриховая, штрихпунктирная);
4) объект типа «Штриховка»;
5) определяется изображение, на котором присутствует разрез.

Необходимо отметить, что область штриховки содержит в себе информацию о «заполненности» материалом конкретной части объекта, чей пространственный образ необходимо восстановить, то есть позволяет определить, где в объекте «материал», а где «отверстие». Эта информация очень важна для отсечения ложной геометрической информации при формировании каркасной модели 3D объекта, так как её анализ позволяет обнаружить рёбра и вершины, не лежащие на поверхности тела.

Описанный алгоритм реализуется на языке Java SE 1.8 в среде разработки IntelliJ IDEA Community Edition 2016.2.5.

Одним из значимых этапов алгоритма является этап обработки массива с исходными данными, в результате которого происходит поиск «уникальных» точек с помощью разработанного метода unique_points() и добавление их в 3 отдельных массива, соответствующие трём разным видам на чертеже. Демонстрационный пример работы алгоритма показан на рис. 3.

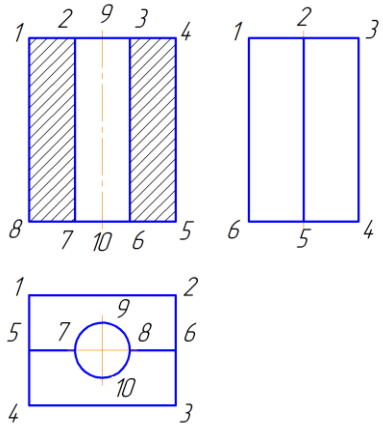


Рис. 3. Поиск «уникальных» точек.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Формируются три матрицы смежности соответствующие каждому виду. Далее, с помощью разработанной процедуры, формируется массив точек разреза, состоящий из 14 вершин на 3-х изображениях, см. рис.4.

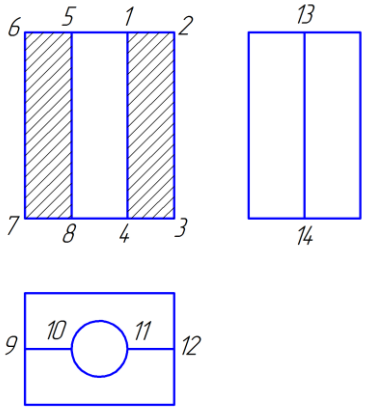


Рис. 4. Формирование вершин, принадлежащих разрезу.

Затем формируется массив с точками в 3D пространстве, учитывающий разрезы (см. рис. 5), и массив R, состоящий из всех возможных рёбер. В рассматриваемом примере насчитывается 32 ребра. На следующем этапе алгоритма происходит удаление из массива R ложных элементов.

В предварительной каркасной модели рёбра или вершины могут быть истинными или ложными, то есть отсутствующими у реального объекта.

За счёт дополнения графа каркасной 3D модели графом разреза появляются рёбра, которые со 100% вероятностью будут являться истинными, так как разрез даёт точную информацию о структуре объекта в секущей плоскости.

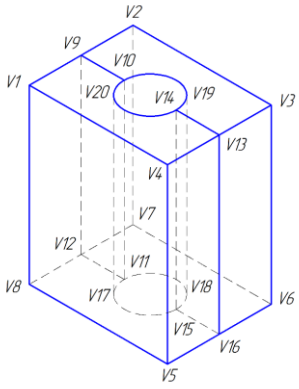


Рис. 5. Полный набор 3D-вершин объекта.

На основе информации о разрезах и сечениях ускоряется поиск и удаление ложных элементов (ложные вершины, ложные рёбра, ложные грани) каркасной модели объекта. Таким образом, расширение входной информации алгоритма задачи синтеза 3D-модель объекта по техническому чертежу за счёт включения в анализ информации о разрезах, позволяет эффективнее выявлять ложные геометрические элементы предварительной каркасной модели в сравнении с алгоритмом, работающим без учёта данной информации.

4. Заключение

Проведённое исследование позволило сделать вывод, что использование в качестве входной информации чертежа, на котором присутствуют простые разрезы, даёт возможность получить каркасную модель с меньшим количеством ложных элементов, в сравнении с каркасной моделью, восстановленной по техническому чертежу, включающему только основные виды. Этот результат позволяет облегчить процесс дальнейших преобразований каркасной модели, необходимых для решения различных задач, в том числе визуализации генерируемого объекта.

Перспектива дальнейших исследований в данной области заключается в разработке алгоритмов распознавания сложных разрезов, сечений, дополнительных видов и т.д. для получения возможности полного решения задачи синтеза 3D-модели по чертежу, выполненному в соответствии с требованиями ЕСКД.

5. Благодарности

Работа выполнена по гранту РФФИ №17-07-00543.

Исследование проводилось в рамках магистерской научно-исследовательской работы Смычёк М.М., обучающейся по направлению 09.04.01 Информационные системы и технологии, по профилю Инженерная геометрия и компьютерная графика в 2016-2018 гг. в ННГАСУ.

6. Литература

[1] Дергунов В.И., Формирование предварительной каркасной модели 3D-объекта по техническому чертежу, содержащему разрезы и сечения / В.И. Дергунов, М.В. Лагунова, М.М. Смычёк, В.А. Тюрина // Сборник трудов Междунар. Научно-практ. конф., Протвино, 2018.

[2] Зудин, А.А. Новая технология геометрического моделирования твердых тел / А.А. Зудин, С.И. Ротков // Сб. тезисов Междунар. конф. «VAI-91». – Новосибирск, 1991.

[3] Котов, И. И. Алгоритмы машинной графики / И.И. Котов, В.С. Полозов, Л.В. Широкова. – М.: Машиностроение, 1977. – 232 с.

[4] Ротков, С. И. Средства геометрического моделирования и компьютерной графики пространственных объектов для CALS-технологий: Дис. докт. техн. наук: 05.01.01 / С.И. Ротков. – Н. Новгород, 1999. – 287 с

[5] Markowsky, G. Generation of solid models from two-dimensional and three- dimensional data / G. Markowsky, M.A. Wesley // in Pickett, MS and Boyse, J M (eds). Solid modelling by computer: from theory to application: Plenum, 1986. – P. 23–51.

Об авторах

Тюрина Валерия Александровна, к.т.н., доцент, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.

Его e-mail: 55555_73@mail.ru.

Смычёк Мария Михайловна, ассистент кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета. Его e-mail: mariasmychek@gmail.com.

Ротков Сергей Игоревич, д.т.н., профессор, заведующий кафедрой Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.

Его e-mail: rotkovs@mail.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Мошкова Татьяна Владимировна, к.т.н., доцент, профессор кафедры Инженерной геометрии, компьютерной графики и автоматизированного проектирования Нижегородского государственного архитектурно-строительного университета.

Его e-mail: ng.forever.mtv@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Компьютерная графика, светотехника и дизайн

Computer Graphics, Lightning Engineering and Design

GraphiCon 2018

24 – 27 сентября 2018
Томск, Россия

Модель рассеивающего слоя с диффузной подложкой и френелевской границей

А.Ю. Басов, В.П. Будак
callia12@rambler.ru|budakvp@gmail.com
ФГБОУ ВО «НИУ «МЭИ», Москва, Россия

Настоящая работа посвящена созданию математической модели плоскопараллельного рассеивающего слоя с диффузной подложкой и верхней френелевской границей. Основой модели является уравнение переноса излучения. Для решения уравнения предлагается метод, основанный на представлении искомой функции в виде суммы анизотропной части, содержащей дельта-особенность, и гладкой части. Анизотропная часть решения определяется с помощью малоугловой модификации метода сферических гармоник. Гладкая часть находится с помощью метода дискретных ординат.

Ключевые слова: *уравнение переноса излучения, метод сферических гармоник, метод дискретных ординат*

Modeling of a scattering slab with a diffuse bottom and top reflecting by the Snell law

A.Y. Basov, V.P. Budak
callia12@rambler.ru|budakvp@gmail.com
National Research University "Moscow Power Engineering Institute", Moscow, Russia

The present paper is devoted to the creation of a mathematical model of plane parallel scattering slab with a diffuse bottom and top reflecting by the Snell law. The model is based on the radiative transfer equation solution. The developed solution method is based on the representation of the function as the sum of an anisotropic part that contains all singularities and a smooth regular part. The anisotropic part could be found by the small-angle modification of the spherical harmonics method. The smooth part is determined by the discrete ordinate method.

Keywords: *radiative transfer equation, spherical harmonics method discrete ordinate method.*

1. Введение

Рассеивающие среды часто представляют собой сложные нерегулярные структуры. Расчёт световых характеристик в таких средах является непростой задачей, связанной с учётом влияния многих факторов: отражения на границах различной формы, рассеяния на частицах, поляризации. Математическая модель рассеивающего слоя с диффузной подложкой и френелевской верхней границей является первым шагом в создании более сложной модели, учитывающей все перечисленные факторы.

Пространственные свойства отражения асфальтобетонных покрытий играют ключевую роль при расчётах яркостных характеристик дорожных покрытий. Нерегулярная структура, недостаточная точность измерительной техники и требования норм в обеспечении яркости при угле наблюдения в 1° [1] создают сложности для исследования отражательных свойств асфальтобетонных покрытий. Существующие нормы основаны на результатах измерений, проведённых в 60-ых годах прошлого века [2]. Есть сомнения, что характеристики современных покрытий аналогичны, что также подтверждается расхождениями в результатах расчётов яркости и результатах измерений яркости на дорогах. Есть надежда, что создаваемая модель будет корректно описывать процессы переноса излучения, проходящие в асфальтобетонных покрытиях. Возможно, модель будет применима и к лакокрасочным и отделочным материалам. За основу работы взята программа, описанная в [3]. Сделано допущение, что дно отражает диффузно, а верхняя поверхность является случайно неровной и отражает по закону Френеля. На данном этапе реализовано отражение от плоской верхней границы.

2. Особенности решения уравнения переноса

Для начала рассмотрим краевую задачу переноса излучения в рассеивающем слое, облучаемом плоским мононаправленным источником под углом Θ_0:

$$\begin{cases} \mu\dfrac{\partial L(\tau,\hat{\mathbf{l}})}{\partial\tau}+L(\tau,\hat{\mathbf{l}})=\dfrac{\Lambda}{4\pi}\oint x(\hat{\mathbf{l}}',\hat{\mathbf{l}})L(\tau,\mu',\varphi')d\hat{\mathbf{l}}', \\ L(\tau,\hat{\mathbf{l}})\big|_{\tau=0,\,\mu>0}=\delta(\hat{\mathbf{l}}-\hat{\mathbf{l}}_0),\ L(\tau,\hat{\mathbf{l}})\big|_{\tau=\tau_0,\,\mu<0}=0, \end{cases} \quad (1)$$

где $\tau=\int_0^z\varepsilon(z')dz'$ – оптическая глубина,

$\tau_0=\int_0^d\varepsilon(z')dz'$,

$\hat{\mathbf{l}}_0=\{\sqrt{1-\mu_0^2}\cos\varphi_0,\sqrt{1-\mu_0^2}\sin\varphi_0,\mu_0\}$,

$\mu_0=\cos\theta_0$.

Оптимальным методом устранения проблемы, связанной с анизотропией, является представление решения в виде суммы анизотропной и гладкой части:

$$L(\tau,\hat{\mathbf{l}})=L_a(\tau,\hat{\mathbf{l}})+\tilde{L}(\tau,\hat{\mathbf{l}}). \quad (2)$$

Малоугловая модификация метода сферических гармоник позволяет представить анизотропную часть решения в трёхмерном пространстве в виде [3]:

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

$$L_a(\tau,\hat{\mathbf{l}}) = L_a(\tau,\mu,\varphi) = \sum_{l=0}^{\infty}\sum_{n=-l}^{l}\frac{2l+1}{4\pi}Z_l^n(\tau)\mathrm{Q}_l^n(\mu)\mathrm{e}^{in\varphi}. \quad (3)$$

Для нахождения гладкой части решения необходимо решить краевую задачу с источником $S(\tau,\mu,\varphi)$:

$$\begin{cases}\mu\dfrac{\partial\tilde{L}}{\partial\tau}+\tilde{L}(\tau,\mu,\varphi)=\dfrac{\Lambda}{4\pi}\oint x(\hat{\mathbf{l}}',\hat{\mathbf{l}})\tilde{L}(\tau,\mu',\varphi')d\hat{\mathbf{l}}'+S(\tau,\mu,\varphi),\\ \tilde{L}(\tau,\mu,\varphi)\big|_{\tau=0,\,\mu>0}=0,\ \tilde{L}(\tau,\mu,\varphi)\big|_{\tau=\tau_0,\,\mu<0}=-L_a(\tau_0,\mu,\varphi).\end{cases} \quad (4)$$

Краевая задача (4) для регулярной части разрешима произвольным численным методом. Представим искомую функцию в виде ряда Фурье:

$$\tilde{L}(\tau,\mu,\varphi)=\sum_{m=-M}^{m=M}C^m(\tau,\mu)\mathrm{e}^{im\varphi}, \quad (5)$$

что приведёт к:

$$\mu\frac{dC^m(\tau,\mu)}{d\tau}=-C^m(\tau,\mu)+\frac{\Lambda}{2}\sum_{k=m}^{N_x}(2k+1)x_k\cdot$$
$$\cdot\mathrm{Q}_k^m(\mu)\int_{-1}^{1}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'+\sum_{k=0}^{\infty}\frac{2k+1}{4\pi}f_k^m(\tau)\mathrm{Q}_k^m(\mu). \quad (6)$$

Главная задача на этом этапе – замена интеграла суммой. Наиболее эффективным методом для этой задачи является метод дискретных ординат. Интеграл можно заменить на [3]:

$$\int_{-1}^{1}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'\approx\frac{1}{2}\sum_{j=1}^{N/2}w_jC_j^-(\tau)\mathrm{Q}_k^m(\mu_j^-)+$$
$$+\frac{1}{2}\sum_{j=1}^{N/2}w_jC_j^+(\tau)\mathrm{Q}_k^m(\mu_j^+), \quad (7)$$

где $\mu_j^-=(\mu_j-1)/2,\ \mu_j^+=(\mu_j+1)/2$,

μ_j – нули полиномов Лежандра $P_\kappa(\mu)$,

w_j – весовые коэффициенты гауссовой квадратуры.

Систему (6) можно заменить системой обыкновенных дифференциальных уравнений. Для решения системы применяется матрично-операторный метод, что удобно для реализации в программе Matlab. После преобразований системы получаем [3]:

$$\begin{bmatrix}\vec{C}_-(0)\\ \vec{C}_+(\tau_0)\end{bmatrix}=\begin{bmatrix}\vec{F}_-\\ \vec{F}_+\end{bmatrix}+\begin{bmatrix}\vec{\vec{R}}_- & \vec{\vec{T}}_-\\ \vec{\vec{T}}_+ & \vec{\vec{R}}_+\end{bmatrix}\begin{bmatrix}\vec{C}_+(0)\\ \vec{C}_-(\tau_0)\end{bmatrix}. \quad (8)$$

Выражение (8) определяет перенос излучения через слой: отраженное $\vec{C}_-(0)$ и прошедшее $\vec{C}_+(\tau_0)$ излучение через излучение падающее на границы слоя сверху $\vec{C}_+(0)$ и снизу $\vec{C}_-(\tau_0)$.

3. Диффузно отражающая подложка

Для решения задачи переноса излучения в слое с диффузно отражающей подложкой перепишем граничные условия на дне в виде:

$$L(\tau_0,\hat{\mathbf{l}})\big|_{\mu<0}=\frac{\rho}{\pi}\int_{(\mu>0)}L(\tau_0,\hat{\mathbf{l}}')\mu d\hat{\mathbf{l}}' \quad (9)$$

Если выразить интеграл (9) через гауссовы квадратуры, то получим в матричной форме:

$$m=0:\ \vec{C}_-(\tau_0)=2\rho\vec{\vec{R}}_L\vec{C}_+(\tau_0);$$
$$\forall m>0:\ \vec{C}_-(\tau_0)=\vec{0}. \quad (10)$$

Таким образом, для реализации программы с диффузной подложкой в программе, реализованной в [3], необходимо поменять только блок, относящийся к нулевой ординате (m=0). Результат работы программы представлен на рис. 1.

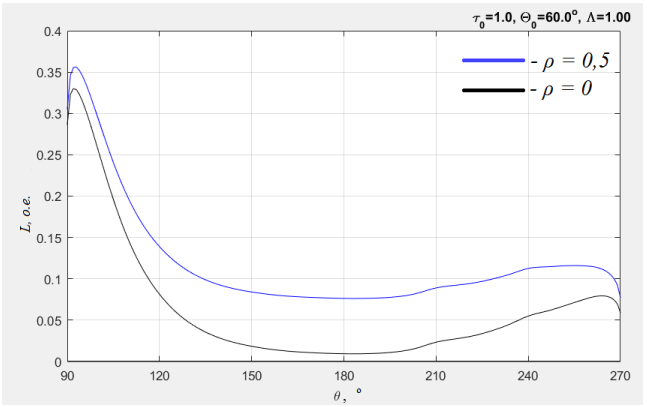


Рис. 1. Реализация учёта диффузного отражения от подложки.

4. Френелевская верхняя граница

Для учёта отражения от верхней границы необходимо учесть преломление:

$$n_1\sin\theta_1=n_2\sin\theta_2. \quad (11)$$

Необходимо учесть, что при $\mu_2<\mu_t\equiv\sqrt{1-1/n_2^2}$ происходит полное внутренне отражение. Тогда интеграл (7) можно рассмотреть по зонам:

$$\int_{-1}^{1}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'=\int_{-1}^{-\mu_t}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'+$$
$$+\int_{-\mu_t}^{\mu_t}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'+\int_{\mu_t}^{1}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'. \quad (12)$$

Первый и последний интегралы связаны с областью преломления, а второй относится к области total reflection. Для второго интеграла выполним преобразование

$$\int_{-\mu_t}^{\mu_t}\mathrm{Q}_k^m(\mu')C^m(\tau,\mu')d\mu'=\int_{-1}^{1}\mathrm{Q}_k^m(\nu)C^m(\tau,\nu)d\nu, \quad (13)$$

где $\mu'=\mu_t\nu$,

что позволяет применить двойную гауссову квадратуру с Nt узлами и впоследствии перейти в этой зоне к двум потоком ординат $\vec{C}_+^t$, $\vec{C}_-^t$, которые на границе связаны идеальным зеркальным отражением.

Для первого и последнего интегралов в (13) сделаем преобразование переменной интегрирования.

При переходе к дискретным ординатам устанавливается полное соответствие между атмосферными ординатами $\vec{C}_+^1$, $\vec{C}_-^1$ и соответствующими $\vec{C}_+^2$, $\vec{C}_-^2$. При этом изменяются веса квадратурной формулы:

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

$$w_i^2 = w_i^1 \frac{\mu_i}{\sqrt{n_2^2 - (1-\mu_i^2)}}, \qquad (14)$$

где μ_i – корни полинома порядка N.

Если ввести вектора $\vec{C}_+^2 = \left[\vec{C}_+^t;\ \vec{C}_+^2\right]$, $\vec{C}_-^2 = \left[\vec{C}_-^2;\ \vec{C}_-^t\right]$, где квадратные скобки обозначают объединение векторов в один суммарной размерности в соответствии с правилами системы Matlab, то для них будет справедливо соотношение (8). Для введенных величин можно записать и условие на границе раздела атмосфера – океан:

$$\begin{gathered}\begin{bmatrix}\vec{C}_-^1\\ \vec{C}_+^2\end{bmatrix} = \begin{bmatrix}\overleftrightarrow{R} & \overleftrightarrow{T}_{12}\\ \overleftrightarrow{T}_{21} & \overleftrightarrow{R}_{22}\end{bmatrix}\begin{bmatrix}\vec{C}_+^1\\ \vec{C}_-^2\end{bmatrix},\\ \overleftrightarrow{T}_{12} \equiv [\overleftrightarrow{T}\quad \overleftrightarrow{0}],\ \overleftrightarrow{T}_{21} \equiv \begin{bmatrix}\overleftrightarrow{0}\\ \overleftrightarrow{T}\end{bmatrix},\ \overleftrightarrow{R}_{22} \equiv \begin{bmatrix}\overleftrightarrow{0} & \overleftrightarrow{1}\\ \overleftrightarrow{R} & \overleftrightarrow{0}\end{bmatrix},\end{gathered} \qquad (15)$$

где $\overleftrightarrow{R}$, $\overleftrightarrow{T}$ – френелевские матрицы отражения.

5. Заключение

В работе реализована модель рассеивающего слоя с диффузной подложкой и френелевской границей. В дальнейшем планируется учесть отражение от случайно неровной границы и поляризацию. Если также учесть многообразие частиц как по форме, так и по размеру и по составу, то модель сможет описывать процессы переноса излучения в асфальтобетонных покрытиях. Для оценки точности моделирования результаты моделирования будут сравниваться с результатами измерений.

6. Литература

[1] ГОСТ Р 55706–2013. Освещение наружное утилитарное. Классификация и нормы. – Введ. 01.07.14. – М.: Стандартинформ, 2014. 12 с.

[2] Островский М.А. Разработка новых норм уличного освещения: отчёт о НИР (промежуточ.). ВНИСИ, – 1971. 35 с.

[3] Budak VP, Korkin SV. On the solution of a vectorial radiative transfer equation in an arbitrary three-dimensional turbid medium with anisotropic scattering. J Quant Spectrosc Radiat Trans 2008; 109:220–34.

Об авторах

Басов Александр Юрьевич, аспирант кафедры светотехники ФГБОУ ВО «НИУ «МЭИ». Его e-mail callia12@rambler.ru.

Будак Владимир Павлович, д.т.н., профессор кафедры светотехники ФГБОУ ВО «НИУ «МЭИ». Его e-mail budakvp@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Освещение лесопарковой зоны. Чувственные аспекты темы

А. А. Бородина[1], Т. С. Хруль[2]
nastia23b@mail.ru|hts@tpu.ru

[1]Томский Политехнический Университет, Томск, Россия;
[2]Томский Политехнический Университет, Томск, Россия

В статье изложена концепция световой инсталляции, объединяющей в себе функцию непосредственного освещения лесопарковой зоны, а также функцию стимуляции человеческих чувств, реализующейся за счет включения человека в светопроцесс. Активация этого процесса предполагает тактильный контакт наблюдателя с деревьями, которые, в свою очередь, являются основой инсталляции.

Ключевые слова: *чувства, природный ландшафт, сенсор, стимуляция, баланс, естество.*

Illumination of urban forest. Sensual aspects of the theme

A. A. Borodina[1], T. S. Khrul[2]
nastia23b@mail.ru|hts@tpu.ru

[1]Tomsk State University, Tomsk, Russia;
[2]Tomsk State University, Tomsk, Russia

The aim of the article is to provide the reader with some material on the conception of light installation, which combines the forest park illumination and the stimulation of the human senses, which is achieved by incorporating people into the process. The process is meant to start when the tactile contact with the trees is made, the trees being the main part of installation.

Keywords: *senses, natural landscape, sensor, stimulation, balance, nature.*

1. Введение

Городская среда современных, особенно крупных, городов действует разрушительно на баланс всех систем организма человека. Давно отмечено, что жители городов живут меньше, чем жители загородной зоны. Это связано с совокупностью факторов, формирующих городскую среду: резкие шумовые эффекты, высокая загазованность, геометричные контуры строений, скученность, захламленность, незначительное озеленение, внутреннее напряжение, связанное с множественными опасностями города. Природная среда, напротив, оказывает на человека благоприятное воздействие, помогая расслабиться и получить эстетическое наслаждение. Сочетание природных цветов, звуков, запахов является успокаивающим фактором, закрепленным в генетической памяти.

Свет оказывает большое влияние на наше благополучие, наше настроение и эмоции, на нашу индивидуальность, он определяет, как мы видим мир вокруг нас. Свет также обладает способностью создавать и усиливать связь с определенным местом – качество, которое стало почти столько же важным, сколько просто обеспечение видимости объектов [5-6]. Безусловен тот факт, что наличие света является фактором, обеспечивающим нас чувством безопасности, освещенность близлежащей территории если не гарантирует чувство всецелой защищенности, то хотя бы предоставляет нам возможность визуального контроля над тем, что нас окружает.

В России большое количество населенных пунктов, в том числе и город Томск, находятся в зоне дефицита солнечного света, приводящего к депрессиям и ухудшению психического состояния людей, поэтому создание технических продуктов, увеличивающих время уличной инсоляции и одновременно положительно воздействующих на эмоциональный и сенсорный фон, активно разрабатываются дизайнерами [2].

2. Концепция

Человек, являясь частью природы, всегда испытывал потребность контакта с ней и по мере развития и укрупнения городов стал вовлекать ее в черту города, оставляя островки леса или создавая парки и сады, сознавая, что что-то в нем самом напрямую зависит от возможности контактировать с природой. Эстетика природной среды вызывает чувство успокоения, запахи растений благоприятно воздействуют на дыхательную систему и мозг, приятное глазу сочетание природных цветов снимает напряжение и раздражение. Пребывание в этой среде делает чувства более тонкими, способствует полезному отрешению и расконцентрации внимания, что просто позволяет привести в баланс свои ощущения, «настроить» тот механизм, с помощью которого человек чувствует, воспринимает мир.

Нередко потенциал таких мест в черте города не раскрыт - природные ландшафты служат лишь для созерцания, как декорации, тогда как дизайнерские идеи по обустройству парковых и лесопарковых зон не стоят на месте, современные решения все чаще и чаще вовлекают человека во взаимодействие с природой. Помимо этого, все возможности таких городских лесов не раскрываются, поскольку они часто остаются небезопасными для пребывания горожан - крайне скудная оснащенность этих зон уличными осветительными приборами сильно ограничивает время посещения. Эти проблемы обходят стороной такие рекреационные зоны, как парки и городские сады в силу того, что обычно они не так густо засажены растительностью, а значит, требуют меньших усилий на обустройство. Лесопарковые зоны гораздо чаще остаются в первозданном виде, не облагораживаются в достаточной степени, хотя и имеют не меньшую привлекательность для горожан по сравнению с более мелкими формами природно-рекреационных зон.

В качестве такой локации был выбран городской лес города Томска – Лагерный сад, площадь которого составляет почти 40 га. Выбор в пользу данного объекта был сделан по ряду причин: он находится в центре города

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

на берегу реки Томи и обладает большой привлекательностью для горожан, рядом с ним расположено место проведения городских праздников, он обладает развитой системой тропинок и дорожек (с учетом оптимальных нагрузок на природную среду и рационального распределения потоков посетителей), однако основной лесной массив (который является проходной зоной) совсем не оснащен уличными осветительными приборами, что не раз становилось причиной различных происшествий.

Предлагаемая концепция освещения лесопарковой зоны заключается в том, что человек может стать частью светопроцесса, который, с одной стороны, направлен на стимуляцию человеческих чувств, а с другой стороны, делает участие человека в нем безопасным, благодаря непосредственному освещению локации, выбранной для расположения инсталляции (то есть сама инсталляция будет ее обеспечивать). Еще с юных лет мы знаем, что «используем наши чувства», чтобы исследовать мир вокруг нас. Но наши чувства больше, чем просто определяют мир вокруг нас: помимо простого восприятия, наши чувства играют интегральную роль в эмоциональной обработке информации, обучении и интерпретации физической среды. С другой стороны, известно, что среда генерирует и стимулирует наши эмоции. Таким образом, существует сильная социальная и эмоциональная связь между нашими чувствами и тем, как мы реагируем на окружающую среду. То есть, способность видеть свет является «доминирующей» в процессе идентификации окружающей среды [3-4].

Концепция предполагает размещение на стволе дерева некоторых технических элементов, которые во взаимодействии друг с другом могли бы воспроизводить спроектированный сценарий освещения (рис.1).

Рис. 1. Схема размещения технических элементов.

Так, на верхушке дерева закреплен датчик освещенности, который включает слабое рассеянное освещение с наступлением сумерек. На стволе дерева от земли до высоты среднего человеческого роста закрепляются датчики приближения, при соприкосновении с которыми передается сигнал прожектору (на солнечной батарее) и парогенератору (на солнечной батарее, с системой сбора и очистки дождевой воды), закрепленным на верхушке дерева, которые в свою очередь плавно подают яркий поток света и небольшое количество пара, имитируя пробуждение и дыхание дерева - реакцию на взаимодействие.

После взаимодействия датчик присутствия, считывает движение в радиусе 50 см вокруг дерева, если никого нет, то через 5 минут система возвращается в режим слабого рассеянного освещения до следующего контакта. С рассветом датчик освещенности выключает систему (рис.2).

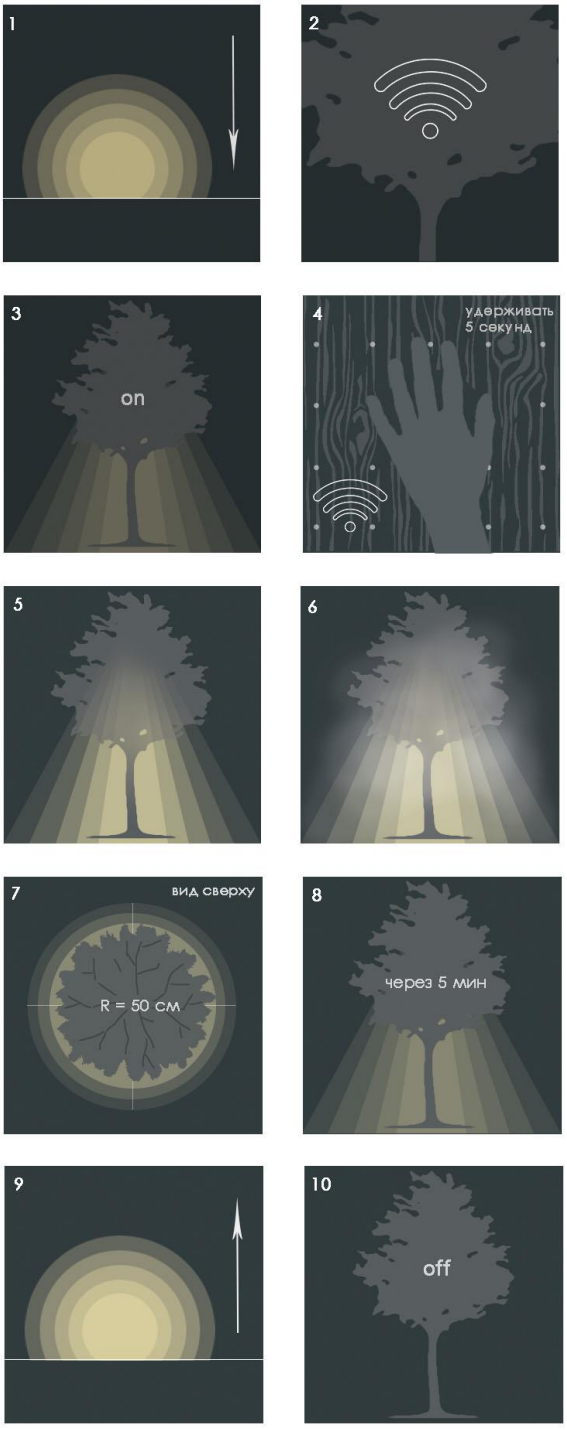


Рис. 2. Сценарий освещения.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Посредством соприкосновения с деревом и датчиками приближения, закрепленными на стволе дерева, человек активирует прожектор и парогенератор, закрепленные на верхушке дерева. Плавная подача небольшого количества пара имитирует дыхание дерева, а мягкое включение света выражает пробуждение дерева, подчеркивает его «реакцию» на взаимодействие с человеком.

В качестве основного активного сенсора выступает осязание – касаясь ствола дерева, человек чувствует текстуру коры. Задействуются и другие чувственные аспекты: слух – человек слышит «дыхание» дерева за счет парогенератора, который подает пар с характерным звуком, имитирующим выдох; обоняние – чувствует запах влажной растительности, благодаря парогенератору, который подает пар непосредственно на ветки деревьев; зрение – видит поток света, спускающийся с верхушек деревьев сквозь пар, подчеркивающий грандиозность и величественность деревьев (рис.3).

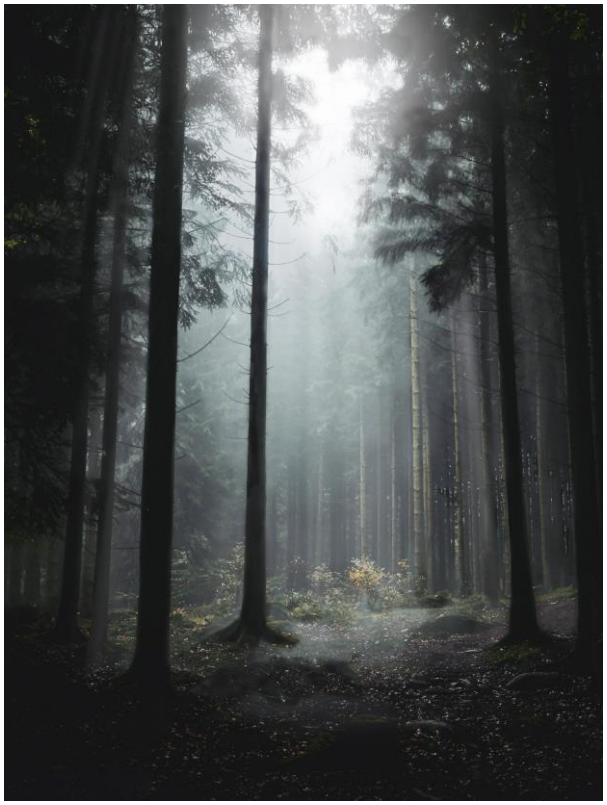


Рис. 3. Предполагаемый визуальный эффект.

Цветовой спектр применяемого искусственного света имеет определяющее значение - он должен быть максимально приближен к естественному свету за счет применения самых современных светодиодных систем, которые позволяют регулировать световые спектры в течение суток. Светодиодные осветительные системы последнего поколения способны хорошо копировать индекс естественной цветопередачи CRI (англ. colour rendering index) практически в любой среде. Искусственное освещение, по возможности, должно быть подобно естественному, так как это не только благоприятно сказывается на организме человека, но и позволяет приблизить саму инсталляцию к естественной среде [1].

Предлагаемое взаиморасположение «дышащих» деревьев учитывает примерный радиус распространения световых лучей от одного дерева. Это позволяет избежать затемненных участков лесопарковой зоны в пределах размещения всей инсталляции, что является важным фактором в обеспечении чувства безопасности в темное время суток. Так, специальные установки предполагается располагать на деревьях, которые находятся друг относительно друга не дальше радиуса в 15 метров (рис.4).

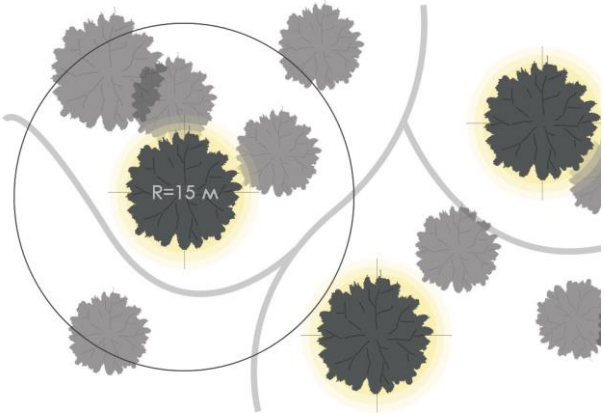


Рис. 4. Предполагаемое взаимное расположение «дышащих» деревьев.

Совокупность сенсоров, задействованных в инсталляции, позволяет наблюдателю почувствовать спектр, казалось бы, знакомых ощущений, но представленных в ином ракурсе. Инсталляция незаметна, она есть лишь техническое оснащение. Она не вынуждает прохожих совершать какие-то действия, но находится в среде, там же, где пребывает человек, который может увидеть огромные деревья в струящемся свете, падающем откуда-то сверху, просто прикоснувшись к нему, может получить необычные впечатления. Суть концепции заключается в том, что через прямое взаимодействие с человеком эта инсталляция особым образом «настраивает» чувства, делает их более точными и тонкими, потому что лес – это естественная среда человека.

3. Заключение

Сформулированная концепция предлагает дополнение существующей лесопарковой зоны светоинсталляцией, способствующей стимуляции человеческих чувств, а также обеспечивающей эту лесопарковую зону освещением.

Предложенное решение хорошо вписывается в современную экологическую повестку – его функциональность предполагает активные меры для сохранения ресурсов окружающей среды, а именно пресной воды и энергетических ресурсов. Также оно хорошо вписывается в Федеральный приоритетный проект «Формирование комфортной городской среды» и может быть реализовано в любых лесопарковых зонах, как Томска, так и любого другого населенного пункта или в лесных природных зонах отдыха.

Размещение точечных зон регулируемой освещенности, инициируемой посетителями, будет расширять временные рамки безопасного пребывания в природной среде, так как прогулку здесь можно будет осуществить и в вечернее время.

Реализация концепта позволит обеспечить не только полноценный отдых в природной среде, но и психоэмоциональную разгрузку, а также компенсацию дефицита солнечного света за счет сочетания природных факторов и прицельного воздействия на органы чувств.

В результате работы была переосмыслена роль общественных природных пространств и роль освещения для них.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4. Литература

[1] Андропова Т.В., Волкотруб Л.П. Гигиеническая оценка современных источников света // Бюллетень сибирской медицины. 2010. Т. 9. № 5. С.155-159.

[2] Афанасьева Р.Ф., Гаврилкина И. Н. Методика и облучательная техника компенсации «солнечного голодания» населения страны // Светотехника. 2004. №4. С. 20–24.

[3] Варфоломеев Л. П. Элементарная светотехника М.: Световые технологии, 2008. – 220 с.

[4] Джост Дж. Маркеза Техника профессионального освещения М.: VERLAG PHOTOGRAPHIE, 1996. - 210 с.

[5] Карлсон В. Настольная книга осветителя / В. Карлсон, С. Карлсон: Пер. с анг. Д. М. Демуровой под ред. С. В. Шульца – М.: ГИТР: Флинта, 2004. – 320 с.

[6] Килпатрик Д. Свет и освещение: Перевод с английского к. т. н. С. Ф. Костромина под ред. к. т. н. А. В. Шеклиана. М.: Мир, 1988 - 223 с.

Об авторах

Бородина Анастасия Александровна – студентка группы 8Д62 Отделения автоматизации и робототехники Инженерной школы информационных технологий и робототехники Томского Политехнического Университета. Ее e-mail nastia23b@mail.ru.

Хруль Татьяна Сергеевна – старший преподаватель Отделения автоматизации и робототехники Инженерной школы информационных технологий и робототехники Томского Политехнического Университета. Ее e-mail hts@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

ОСОБЕННОСТИ ПРОЕКТИРОВАНИЯ РАЗВИВАЮЩЕЙ МОДУЛЬНОЙ МЕБЕЛИ ДЛЯ ДЕТЕЙ

Бринюк И.А., Радченко В.Ю.
ibrinyuka@gmail.com|rad@tpu.ru
Национальный исследовательский Томский Политехнический университет, г.Томск, Россия

Недостаточный уровень активности детей обусловлен несоответствием окружающей среды особенностям развития ребенка. Существующая среда не подталкивает ребенка к самостоятельности, проявлению инициативности. Это в свою очередь приводит к затруднению эмоционально-оценочного восприятия и осмысления информации, овладение разными способами ее получения. Главным, при проектировании детской модульной мебели, является сохранение для ребенка ощущение доступности всех игровых элементов в комнате, обеспечения удобства подхода к игровой мебели. В данной статье предложен подход к организации и моделированию предметно-пространственной среды с применением модульных объектов, а также представлены требования к их созданию.

Ключевые слова: Развивающая среда, восприятие, разработка, предметно-пространственной среды.

DESIGN FEATURES OF DEVELOPING MODULAR FURNITURE FOR CHILDREN

Brinyuk I.A, Radchenko V.Yu.
ibrinyuka@gmail.com|rad@tpu.ru
National Research Tomsk Polytechnic University, Tomsk, Russia

Insufficient level of activity of children is caused by discrepancy of environment to features of development of the child. The existing environment does not encourage the child to self-sufficiency, manifestation of initiative. This, in turn, leads to difficulties in the emotional and evaluative perception and understanding of information, mastering different ways to obtain it. The main thing in the design of modular children's furniture, is to keep the child a sense of accessibility of all the game elements in the room, ensuring the convenience of approach to gaming furniture. This article proposes an approach to the organization and modeling of the object-spatial environment using modular objects, as well as the requirements for their creation.

Keywords: Development Environment, perception, working out, the object-spatial environment.

1. Введение

В современном мире остро стоит проблема нехватки удобной детской мебели. Детская мебель играет важную роль в развитии личности ребенка, влияя на развитие его умственных и физических возможностей. Окружающая среда воздействует на ребенка, является источником познания действительности. Именно поэтому важно использовать подходящую детскую мебель в совокупности ее свойств и связей для развития наблюдательности, а также исследовательского подхода к доступным объектам развивающей среды. Это открывает перед ребенком широкие возможности действовать активно, разумно и самостоятельно.

2. Принципы создания детской мебели

Целью проекта, являлось создание детской мебели посредством модульной конструкции, которая удовлетворяла потребности каждого ребенка, помогая ему реализовать свои способности. Предлагаемая детская мебель, может служить и мебелью, и игрушкой и использоваться детьми и родителями в разных целях: хранение вещей, организация игрушек, а также объектом игры. Данную детскую мебель можно собрать перед игровой деятельностью, а после игры разобрать и сложить все в один модуль [1]. Это и является главным преимуществом нашей детской мебели. Использование модульной мебели позволяет обеспечить детям интересные впечатления, позитивные и эстетические эмоции. При проектировании детской модульной мебели, следует учитывать принципы создания модульной мебели и их использование в работе с детьми.

- Принцип информативности. Данный принцип проявляется в удовлетворении интереса ребенка к новизне. Ребенок получает и использует различную информацию, в процессе деятельности.

- Содержательно-организационные принципы. Данные принципы отражают особенности комплектации модульной мебели, способы создания композиций и организации работы с ними.

- Принцип эмоциональной насыщенности. Данный принцип способствует в получении детьми различных, меняющихся впечатлений.

- Принцип поли-функциональности. Суть данного принципа проявляется в мобильности конструкции. Детали могут быть использованы в разных видах деятельности ребенка. Изменяя дополнительные детали, дети могут изменить назначение игрового модуля.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Данные принципы обеспечивают возможность моделировать мебель, используя модульные объекты [2].

При проектировании детской модульной мебели, необходимо вызвать максимальный интерес у ребенка. Поэтому было решено, что создаваемый комплект будет являться модульной конструкцией. В игре ребенок легче поддается обучению, поскольку он сам заинтересован в результате и сам направляет усилия на достижение цели. Разрабатываемая модульная мебель привлечет к себе внимание ребенка с различными физическими возможностями, а также будет многопользовательской, командной.

Для более точного представления о функциях, необходимых детской мебели, также был проведен анализ методов воспитания и обучения детей, в результате которого было выявлено, что игра оказывает большее положительное влияние на развитие, и является самым эффективным средством развития ребенка дошкольного возраста. Роль игры в развитии ребенка очевидна, так как с ее помощью можно успешно корректировать, улучшать, умственные, а также физические способности ребенка.

Для полного погружения в игровой процесс, детская мебель должна быть многовариантной, она должна развивать разные интеллектуальные качества: память, пространственное представление, умение находить зависимости, т. е. создавать новые комбинации из имеющихся элементов, деталей, предметов, но также не должна ограничивать играющего. Таким образом, для решения проблемы с организованностью, детская мебель должна быть модульной.

3. Реализация

На этапе проработки концептуального решения, с учетом изученной литературы и анализа существующих решений детской мебели, была создана эскизная модель (Рис.1), которая состоит из модулей (сот), соединяемых в единую конструкцию.

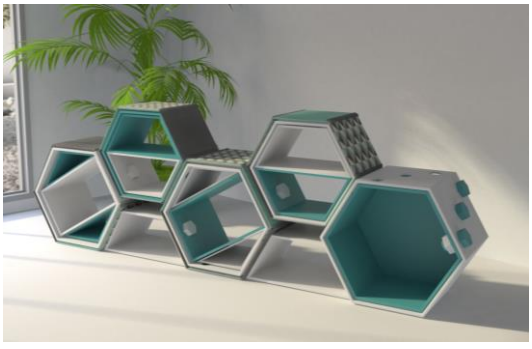


Рис.1. Эскизная модель

Важным, в данной конструкции, является соединительный механизм. было принято решение использовать болтовое соединение. Данный вид соединения, помогает развивать физические и умственные способности ребенка.

Реализация данной детской мебели началась с создания модулей, которые вставляются в форму (сота). В форме модули закрепляются с помощью болтов. Главным преимуществом является то, что если вынуть любой модуль (стол, стулья), целостность конструкции не нарушится (Рис.2).

Рис. 2. Модули

С помощью данной детской мебели удается получать различные игровые зоны и занимать минимальное пространство в помещении. Таким образом можно развивать способность к комбинированию элементов (Рис.3).

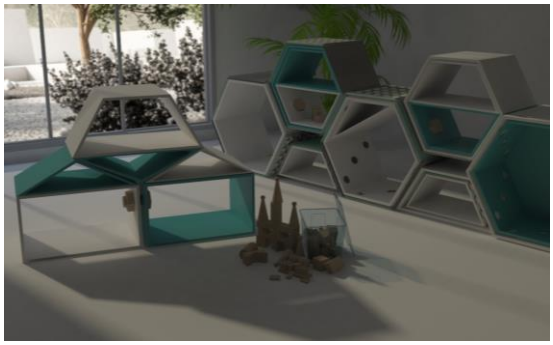


Рис. 3. Комбинирование модулей

Основной идеей, является разнообразное соединение модулей в различные геометрические конструкции. Это помогает развивать последовательное мышление и зрительную память, а также, данная мебель позволяет повысить интерес к игрушке и не ограничивать ребенка в его фантазии.

Немало важным факторам при проектировании детской мебели, является - внешний вид и материал, из которого будут изготовлены модули. Более безопасным материалом для детской мебели - является тамбурат. Данный материал, имеет композитную структуру. Все свои достоинства этот материал получил благодаря специфической конструкции. Основными преимуществами данного материала являются:

Легкость обработки – данный материал можно легко резать любым инструментом, благодаря чему можно получить детали сложной формы.

Прочность конструкции. Соты из плотного картона помогают тамбурату выдерживать сильные нагрузки.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Именно поэтому тамбурат часто используют при изготовлении выпуклых элементов конструкций.

Малый вес. Благодаря особому составу, внешние габариты тамбурата совершенно не совпадают с массой материала. Также следует отметить, что данный материал является одним из доступных материалов в своей категории. Изделие из данного материала в несколько раз дешевле по сравнению с аналогами.

4. Заключение

Таким образом, разработанная детская мебель, существенно отличается от аналогов, а также имеет преимущества. В рамках данной работы, удалось добиться максимального функционала, с минимальным зонированием пространства. Данная мебель является модульной конструкцией, которая имеет различное количество вариантов сборки, что значительно повышает заинтересованность ребенка. Разработанная детская мебель является инструментом для обучения за счет модульности и развития разных интеллектуальных качеств, а также интересна не только ребенку, но и его родителям. Модульная детская мебель развивает у ребенка логическое мышление и мелкую моторику пальцев и рук. Созданный комплект детской мебели не зациклен на одном физическом действии, а позволяет развивать ребенка, как физически, так и умственно.

5. Список литературы

[1]. Моделирование предметно-пространственной среды [Электронный ресурс] // 2017. URL: https://cyberleninka.ru/article/v/modelirovaniepredmetnoprostranstvennoy-sredy-posredstvom-modulny konstruktorov (Дата обращения: 03.05.2018).

[2]. Разработка детской модульной мебели [Электронный ресурс] // 2018. URL: http://docplayer.ru/65915502-Razrabotka-modulnoy-detskoy-mebeli-dlya-doshkolnogo-obrazovatelnogo-uchrezhdeniya.html (Дата обращения: 04.05.2018).

[3]. Тамбурат: материал для производства мебели [Электронный ресурс] // 2017. URL: http://derevo-s.ru/material/mebelnye/tamburat (Дата обращения: 04.05.2018).

Об авторах

Автор: Бринюк Илья Андреевич, студент 3 курса ИШИТР ТПУ, e-mail: ibrinyuka@gmail.com.

Руководитель: Радченко Валерия Юрьевна, старший преподаватель ИШИТР ТПУ, e-mail: rad@tpu.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Влияние учета поляризации света на результат при расчете многократных отражений

В.П. Будак[1], А.В. Гримайло[2]
BudakVP@gmail.com|GrimailoAV@gmail.com
[1]ФГБОУ ВО "НИУ "МЭИ", Москва, Россия;
[2]ФГБОУ ВО "НИУ "МЭИ", Москва, Россия

Предложена математическая модель отражения света от поверхности на основе использования вектора Стокса и матриц Мюллера. Приведены уравнение глобального освещения и локальная оценка метода Монте-Карло для вектора Стокса. В первом приближении получены результаты на основе предложенной модели.

Ключевые слова: *математическая модель, вектор Стокса, матрица Мюллера, поляризация, многократные отражения.*

The influence of the light polarization account on the result of multiple reflections calculation

V.P. Budak[1], A.V. Grimailo[2]
BudakVP@gmail.com|GrimailoAV@gmail.com
[1]National Research University "Moscow Power Engineering Institute", Moscow, Russia;
[2]National Research University "Moscow Power Engineering Institute", Moscow, Russia

The surface light reflection mathematical model based on Stokes vector and Mueller matrices is proposed. The global illumination equation and the local estimation of the Monte-Carlo method are given for Stokes vector. The first approximation results based on the proposed mathematical model are obtained.

Keywords: *mathematical model, Stokes vector, Mueller matrix, polarization, multiple reflections.*

1. Введение

За последние годы визуализация светораспределения стала неотъемлемой частью проектирования осветительных установок. В этом исключительно важную роль играет расчет многократных отражений света от поверхностей объектов, составляющих сцену.

Традиционно в светотехнике при расчетах не учитывается поляризация света [3]. В частности, программы DIALux и Relux, ставшие стандартом в отрасли при расчетах осветительных установок, полностью пренебрегают поляризацией. Если при расчете малого числа отражений отсутствие учета поляризации практически не вносит ошибку в конечный ответ, то при моделировании многократных отражений света, особенно от поверхностей со значительной долей френелевского отражения, это может привести к большому отличию полученных результатов от истинных значений интересующих величин. Данный вопрос на сегодняшний день остается малоизученным.

Исходя из этого, авторы сочли необходимым провести исследование влияния учета поляризации света на результат при расчете многократных отражений.

2. Математическая модель отражения света с учетом поляризации

Для создания математической модели отражения от поверхности, которая бы содержала характеризующие поляризацию света параметры, удобным представляется использование вектора Стокса (описывающего параметры луча) и 4×4 матриц Мюллера (описывающих свойства отражающих поверхностей).

Фотометрический смысл вектора Стокса состоит в том, что первый его элемент представляет собой полную яркость луча, а три других — разность яркостей этого луча, прошедшего через соответствующие поляризаторы [4]. Необходимо отметить, что компоненты вектора Стокса можно получить экспериментально. Это, несомненно, является его большим достоинством.

Результат взаимодействия луча, описываемого вектором Стокса с поверхностью может быть представлен в виде:

$$\boldsymbol{L}(\boldsymbol{r}', \hat{\boldsymbol{l}}') = \vec{\boldsymbol{\rho}}(\boldsymbol{r}, \hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}')\boldsymbol{L}(\boldsymbol{r}, \hat{\boldsymbol{l}}),$$

где $\boldsymbol{L}(\boldsymbol{r}, \hat{\boldsymbol{l}})$ и $\boldsymbol{L}(\boldsymbol{r}', \hat{\boldsymbol{l}}')$ — соответственно векторы Стокса до и после взаимодействия с поверхностью, $\vec{\boldsymbol{\rho}}$ — матрица отражения (Мюллера).

Результат ряда последовательных преобразований будет получен путем применения соответствующей матрицы $\vec{\boldsymbol{\rho}}$, которая является произведением матриц частичных преобразований. В свою очередь матрица параллельных преобразований представляет собой сумму матриц частичных преобразований.

Таким образом, руководствуясь теми же соображениями, что и при выводе уравнения глобального освещения [1], можно получить аналогичное уравнение для вектора Стокса:

$$\boldsymbol{L}(\boldsymbol{r}, \hat{\boldsymbol{l}}) = \boldsymbol{L}_0(\boldsymbol{r}, \hat{\boldsymbol{l}}) + \frac{1}{\pi}\int \vec{\boldsymbol{R}}(\hat{\boldsymbol{l}}' \times \hat{\boldsymbol{l}}, \hat{\boldsymbol{N}} \times \hat{\boldsymbol{l}})\vec{\boldsymbol{\rho}}(\boldsymbol{r}, \hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}') \times \vec{\boldsymbol{R}}(\hat{\boldsymbol{l}}' \times \hat{\boldsymbol{N}}, \hat{\boldsymbol{l}}' \times \hat{\boldsymbol{l}})\boldsymbol{L}(\boldsymbol{r}, \hat{\boldsymbol{l}})\left|(\hat{\boldsymbol{N}}, \hat{\boldsymbol{l}})\right| d\hat{\boldsymbol{l}}', \quad (1)$$

где $\boldsymbol{L}(\boldsymbol{r}, \hat{\boldsymbol{l}})$ – вектор Стокса, $\vec{\boldsymbol{R}}(\hat{\boldsymbol{l}}' \times \hat{\boldsymbol{N}}, \hat{\boldsymbol{l}}' \times \hat{\boldsymbol{l}})$ – матрица поворота плоскости референции.

Для решения уравнения (1) используется численный метод Монте-Карло как зарекомендовавший себя при решении подобного рода задач.

Во время реализации данного метода был сделан выбор в пользу локальных оценок вместо прямого моделирования ввиду его неудобства и малой эффективности при расчете многократных отражений. Для этого способом, аналогичным тому, что приведен в [2], было получено выражение для локальной оценки метода Монте-Карло:

$$\boldsymbol{I}_{\varphi} = M\sum_{n=0}^{\infty}\boldsymbol{Q}_n k(\boldsymbol{r}, \boldsymbol{r}'), \quad (2)$$

где $k(\boldsymbol{r}, \boldsymbol{r}') = \frac{\left|(N(r),(r-r'))(N(r'),(r-r'))\right|}{(r-r')^4}\Theta(\boldsymbol{r}, \boldsymbol{r}')$ – ядро уравнения излучательности [2], $\boldsymbol{Q}_n$ – векторный вес луча, $\Theta(\boldsymbol{r}, \boldsymbol{r}')$ – функция видимости.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

3. Реализация предложенной модели и результаты

Приведенная выше математическая модель была реализована с помощью пакета MATLAB. Для упрощения алгоритма и уменьшения времени расчета было принято допущение, согласно которому направления отраженных лучей от всех поверхностей распределяются по диффузному закону. Это дает возможность оценить в первом приближении влияние учета поляризации на результат.

В качестве сцены для моделирования было использовано «помещение» с размерами $1 \times 1 \times 1$. В центре «потолка» находится ламбертовский светящий диск диаметром 0,1. На «полу» расположены точки, в которых оценивается освещенность (рис. 1).

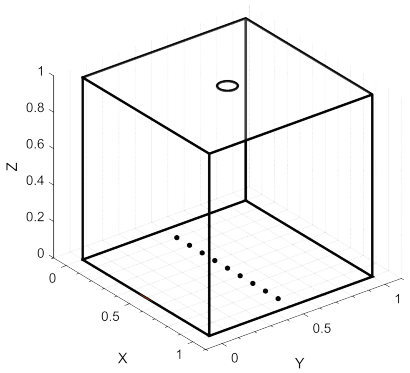


Рис. 1. Сцена для моделирования.

При расчете в качестве матрицы отражения использовалась сумма двух матриц:

$$\vec{\boldsymbol{\rho}} = a\vec{\boldsymbol{\rho}}_F + (1-a)\vec{\boldsymbol{\rho}}_L,$$

где $\vec{\boldsymbol{\rho}}_F$ – матрица френелевского отражения (Мюллера), $\vec{\boldsymbol{\rho}}_L$ – матрица ламбертовского отражения, a – доля френелевского отражения ($0 < a < 1$).

Матрица $\vec{\boldsymbol{\rho}}_L$ представляет собой нулевую матрицу с единственным ненулевым элементом $\vec{\boldsymbol{\rho}}_{L\,11}$, равным коэффициенту отражения.

Были заданы следующие параметры поверхностей: коэффициент отражения – 0,5, коэффициент преломления – 1,5, параметр a изменялся от 0 до 1. На рис. 2 приведены полученные результаты, нормированные относительно количества выпущенных из диска лучей.

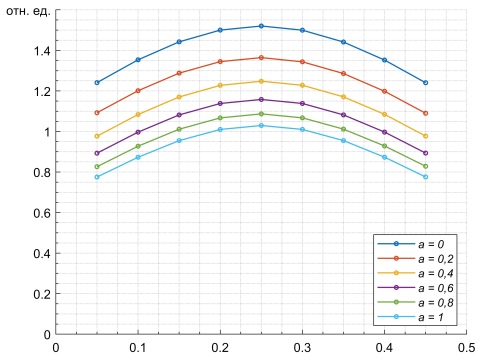


Рис. 2. Результаты расчета.

4. Заключение

Как видно из приведенных выше результатов расчета, уже при $a = 0{,}2$ различие между полученными значениями составляет порядка 10%, а при $a = 0{,}6$ — более 20%. Из этого можно сделать вывод, что учет поляризации вносит существенные изменения в результат при расчете многократных отражений. Следовательно, в ряде случаев (например, при расчете осветительных установок и моделировании отражающих свойств поверхностей) необходимо учитывать состояние поляризации света.

Исследования других авторов [5], посвященные вопросам визуализации с учетом поляризации (но не принимающие во внимание светотехнические аспекты) показывают, что изменяется также положение и направление бликов. Это очень важно при оценке дискомфорта и ослепленности, которые определяют качество освещения сцены.

Стоит отметить, что предложенная математическая модель остается в рамках стандартных фотометрических представлений. Отличием является то, что яркость теперь становится векторной величиной, а коэффициент отражения превращается в матрицу. Также появляется необходимость в матрицах поворота плоскости референции при каждом взаимодействии луча с поверхностью.

Дальнейшими направлениями работы в этой области могут быть разработка более полной модели отражения поляризованного света от поверхностей исследуемой сцены, включающей отражение и преломление на случайно-неровной поверхности, рассеяние света в подверхностном слое среды, и ее экспериментальная проверка.

5. Литература

[1] Будак В.П. Визуализация распределения яркости в трехмерных сценах наблюдения. – М.: Издательство МЭИ, 2000. – 136 с.

[2] Желтов В.С. Моделирование осветительных установок на основе решения уравнения глобального освещения локальными оценками метода Монте-Карло: диссертация … кандидата технических наук: 05.09.07 – Москва, 2008. – 97 с.

[3] Коркин С.В. Математическая модель отражения поляризованного излучения при дистанционном зондировании мутных сред: диссертация … кандидата технических наук: 05.11.07 — Москва, 2009. — 150 с.

[4] Шерклифф У. Поляризованный свет. — М.: Издательство «Мир», 1965. — 264 с.

[5] Wolff L. B., Kurlander D. J. Ray tracing with polarization parameter // IEEE Computer Graphics and Applications. 1990. V. 10, No. 6, P. 44-55.

Об авторах

Будак Владимир Павлович, д.т.н., профессор кафедры светотехники факультета электронной техники института радиотехники и электроники НИУ «МЭИ». Его e-mail BudakVP@gmail.com.

Гримайло Антон Валентинович, студент кафедры светотехники факультета электронной техники института радиотехники и электроники НИУ «МЭИ». Его e-mail GrimailoAV@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Оценка качества внутреннего освещения в сценах с неравномерными блескими источниками

Владимир Будак, Виктор Желтов, Виктор Чембаев, Татьяна Мешкова
budakvp@mpei.ru | zheltov@list.ru | chembervint@gmail.com | tvmesh@mail.ru
кафедра светотехники
Национальный исследовательский университет "МЭИ", Москва, Россия

Современные методы моделирования уравнения глобального освещения позволяют рассчитывать простанственно-угловое распределение яркости, на основе которого появляется возможность получения качественных характеристик освещения сцены при светотехническом проектировании. Существующие нормативные базы предлагают лишь объединенный показатель дискомфорта, который имеет ряд ограничений в своем применении. В данной работе описывается критерий, учитывающий вклад от неравномерных протяженных источников бликов на общее качество сцены внутреннего освещения.

Ключевые слова: *светотехника, компьютерная графика, критерий качества освещения, уравнение глобального освещения.*

Interior lighting quality estimation for scenes with non-uniform light sources

Vladimir Budak, Victor Zheltov, Victor Chembaev, Tatina Meshkova
budakvp@mpei.ru | zheltov@list.ru | chembervint@gmail.com | tvmesh@mail.ru
department of Lighting technology
Moscow Power Engineering Institute, Moscow, Russia

Modern methods of global illumination equation simulation enables us to compute spatial angular radiance distribution, which could be used in determining qualitative characteristics of illumination scene for lighting design. Currently existing regulatory environments includes only Unified Glare Rating (UGR), which has a number of limitations in its application. This article considers new criteria, which takes into account the contribution from non-uniform extended glare sources to the overall quality of the indoor lighting scene.

Keywords: *lighting technology science, computer graphics, illumination quality criteria, global illumination equation.*

1. Введение

Вопрос качества внутреннего освещения занимает человечество с момента возникновения первых осветительных приборов (ОП). И, как и все процессы, неизбежно проходит сквозь этапы эволюции: выживание, социализация, удовольствие. И действительно, если первые осветительные установки в виде костров, помогали первобытным людям выживать и освещать свои жилища, далее, в античные времена появляются факелы, которые делают ОП мобильными и дают возможность создавать первые осветительные установки (ОУ). Активное развитие искусственного освещения приходится на 15-19 века, когда основным источником света являются всевозможные свечи и масляные лампы, которые позволяют людям работать в темное время суток, а также украшать свои дома и улицы. Ну и, разумеется, самый активный рост данной отрасли начинается с появлением электричества, когда уже по-настоящему можно сказать, что сфера искусственного освещения перешла на третий этап своей эволюции.

Теория расчета и нормирования ОУ практически полностью уже была изложена в учебных материалах к концу 70-х годов. В начале появляются методы расчета ОУ по определенным количественным характеристикам, таким как освещенность, яркость, световой поток, сила света. Постепенно формируется нормативная база. Математика, лежащая в основе этих методов не особо сложна с алгоритмической точки зрения, и появившиеся в конце 70х компьютеры общего назначения, уже могут справляться с задачами расчета ОУ. Особый прогресс начинается в 90гг с появлением персональных компьютеров, ростом их мощности и появлением систем автоматического проектирования (САПР). Однако все эти методы расчета и инженерные практики пока еще не учитывают важнейший вопрос качества освещения.

2. Компьютерная графика и светотехника

В основе как компьютерной графики, так и расчета ОУ в светотехнике лежит уравнение глобального освещения (УГО), полученное еще в 1986 году J. Kajia [4]

$$L(\mathbf{r},\hat{\mathbf{l}}) = L_0(\mathbf{r},\hat{\mathbf{l}}) + \frac{1}{\pi}\int L(\mathbf{r},\hat{\mathbf{l}}')\sigma(\mathbf{r};\hat{\mathbf{l}},\hat{\mathbf{l}}')\left|(\hat{\mathbf{N}},\hat{\mathbf{l}}')\right| d\hat{\mathbf{l}}', \qquad (1)$$

где $L(\mathbf{r},\hat{\mathbf{l}})$ – яркость светового поля в точке $\mathbf{r}$ по направлению $\hat{\mathbf{l}}$, $\sigma(\mathbf{r};\hat{\mathbf{l}},\hat{\mathbf{l}}')$ – коэффициент яркости элемента поверхности в точке $\mathbf{r}$ в направлении $\hat{\mathbf{l}}$ при его направленном освещении параллельным пучком света в направлении $\hat{\mathbf{l}}'$, часто называемая двунаправленной функцией рассеяния (отражения или пропускания), L_0 – прямая компонента яркости, непосредственно от источников, $\hat{\mathbf{N}}$ – нормаль в точке $\mathbf{r}$ к элементу поверхности сцены. Крышечкой над вектором мы обозначаем единичные вектора. Контурный интеграл в уравнении означает интегрирование по полному телесному углу, элемент которого обозначен через $d\hat{\mathbf{l}}'$.

Моделируя УГО, мы получаем пространственно-угловое распределение яркости, т.е. визуализацию 3М сцен. Уравнение представляет собой интегральное уравнение второго рода, и, соответственно, в общем виде не имеет аналитического решения (за исключением некоторых частных случаев, которые могут быть применены для

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

оценки того или иного численного метода. (Фотометрическая сфера и “задача Соболева”). На сегодняшний день существует ряд численных методов решения УГО, однако, все они имеют ряд допущений, отражающихся в конечном счете на достоверности расчета.

На решении УГО построено и современное моделирование ОУ, что так тесно связывает компьютерную графику и светотехническое проектирование ОУ. Однако наиболее распространенные инженерные методы моделирования ОУ на сегодняшний день основываются не на решении УГО, а на решении уравнения излучательности [3] методом конечных элементов [7]:

$$M(\mathbf{r}) = M_0(\mathbf{r}) + \frac{\sigma(\mathbf{r})}{\pi}\int_{\Sigma} M(\mathbf{r}')F(\mathbf{r},\mathbf{r}')\Theta(\mathbf{r},\mathbf{r}')d^2\mathbf{r}', \quad (2)$$

где $M(\mathbf{r})$ – светимость в точке $\mathbf{r}$, $M_0(\mathbf{r})$ – начальная светимость в точке $\mathbf{r}$ непосредственно от источников света в приближении одного отражения, $\Theta(\mathbf{r},\mathbf{r}')$ – функция видимости точки $\mathbf{r}$ из точки $\mathbf{r}'$ (1 если видна и 0 если не видна), $F(\mathbf{r},\mathbf{r}')$ – элементарный форм-фактор [7], $\sigma(\mathbf{r})$ – коэффициент отражения.

При светотехническом проектировании ОУ одним из важнейших аспектов является определение нормируемых характеристик. Эти характеристики можно классифицировать на 2 типа: количественные и качественные. К количественным относятся такие характеристики, как освещенность, яркость, а также производные от них величины – средняя освещенность, неравномерность и т.п. Данные характеристики зарегулированы нормативными документами, на основании которых на сегодняшний день и производится проектирование и расчет ОУ.

Однако, человеческий глаз реагирует на яркость, падающего на него излучения, но не на освещенность, в то время как большинство существующих норм нормируют освещенность. Эта ситуация объясняется тем, что еще до недавнего времени численные методы и имеющие вычислительные средства не позволяли моделировать пространственно-угловое распределение яркости с достаточным уровнем достоверности. Поэтому самые распространенные компьютерные программы для инженерного светотехнического моделирования осветительных установок, такие как DIALux [9] и Relux [10], основываются как раз на методе излучательности.

За последние несколько лет численные методы решения уравнения глобального освещения, а кроме того, и вычислительные мощности компьютеров получили значительное развитие. Огромный вклад в развитие алгоритмов решения УГО внес взрывной рост трехмерной компьютерной графики, как отрасли, во многом благодаря задачам САПР и компьютерных игр. Так или иначе наличие этих самых методов позволяет уже сегодня перейти к моделированию непосредственно яркости. Вопрос в том, что с ней дальше делать? Еще в 1915г американская ученая Клэренс Фери в своей работе [1] показала, что именно пространственно-угловое распределение яркости определяет дискомфорт, а значит и качество освещения. Поэтому логичным кажется переход от проектирования ОУ по нормируемым количественным характеристикам к проектированию на заданное качество освещения. На сегодняшний день практически еще не существует нормативной базы для качества освещения. Среди нормируемых качественных характеристик в инженерной практике используется лишь объединенный показатель дискомфорта UGR:

$$\mathrm{UGR} = 8\lg\left[\frac{0{,}25}{L_a}\sum_{i=1}^{N}\frac{L_i^2\omega_i}{p_i^2}\right], \quad (3)$$

где L_i - яркость блеского источника, кд/м^2; ω_i - угловой размер блеского источника, стер; p_i - индекс позиции блеского источника относительно линии зрения; L_a - яркость адаптации, кд/м^2.

UGR учитывает яркость точечных источников света, их угловой размер и положение в поле зрения. Таким образом, данный критерий позволяет учитывать так называемую блескость от малоугловых источников, однако в реальной жизни, источники часто бывают протяженными и не равнояркими. Кроме того, UGR не учитывает и блики на диффузно-направленных отражающих поверхностях. В настоящее время ведутся активные работы по исследованию возможностей нивелирования текущих ограничений UGR [2].

3. Качество внутреннего освещения

Качество внутреннего освещения непрерывно связано с такими понятиями как дискомфорт и блескость. Для глаза человека, как для приемника, воспринимающего сцену освещения, яркость, создаваемая осветительными установками, может, как помогать выполнению поставленной задачи, т.е. осветить тот или иной предмет, текст, рабочую поверхность и т.д., так и мешать – например, в случае если отраженная яркость предмета будет значительно превосходить яркость адаптации (фона), или будут появляться блики. Наглядным примером может служить Панно из майолики “Метростроевцы”, рис. 1, которое, как предмет искусства, было призвано украшать переход станции московского метро Комсомольская, рассказывая исторический сюжет проектирования станции, однако, за счет невозможности рассчитать качество освещения данной сцены на стадии проектирования ОУ, мы видим, что светильники на потолке создают яркие блики на хорошо отражающем кафеле, что мешает прочитать саму картину.

Рис. 1. Панно из майолики «Метростроевцы».

Стоит отметить, что и на сегодняшний день этот эффект не может быть учтен средствами САПР, которые из качественных критериев учитывают только UGR.

Другим примером может быть освещение в учебном классе. На рис. 2 мы видим насколько блики на доске мешают учащимся прочитать информацию с доски.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Рис. 2. Изображение классной доски.

Что же такое качество освещения? В работе [6] была произведена попытка сформулировать определение критерия качества освещения: "*освещение можно считать качественным, если оно способствует повышению зрительной работоспособности человека и не препятствует в решении задач, поставленных в рамках той или иной сцены освещения*". Кроме того, авторы предложили новый подход к определению качества освещения. Подход основан на представлении качества освещения в виде одного числа.

$$Q = \frac{1}{AK_{пор}} \int K(x,y)\,p(x,y)\,h(x,y)\,dxdy \ , \qquad (4)$$

где $K_{пор}$ - пороговое значение контраста, $p(x)$ - некоторая весовая функция, учитывающая различный вклад в реакцию глаз точек, расположенных в центре поля зрения и на периферии, так как плотность колбочек наибольшая у зрительной оси, $h(x)$ - функция учитывающая значимость той или иной точки для качества освещения.

Как уже упоминалось выше на сегодняшний день существует только один нормируемый критерий качества освещений - UGR, однако, он имеет ряд ограничений в применении. Новый единый критерий качества внутреннего освещения позволяет сравнивать разные сцены освещения на предмет удобства выполнения зрительной работы в них. Критерий учитывает полное пространственно-угловое распределение яркости для заданной точки наблюдения, поэтому позволяет учитывать вклад не только точечных блеских источников, но и протяженные неравнояркие источники, в т.ч. и вклад от многократных переотражений.

Для нормирования критерия и проверки его работоспособности было проведено ряд экспериментов на лабораторной установке, рис. 3, собранной на кафедре светотехники НИУ МЭИ. Установка представляет матово-белую панель с установленными на ней светодиодными модулями разной формы и конфигурации. Подробное описание экспериментальной установки, а также методик проведения экспериментов изложено в статье [8].

Рис. 3. Лабораторная установка оценки границы качества освещения.

Первым экспериментом на установке был поставлен опыт, повторяющий опыт Лекиша и Гутта [5], по установке границы комфорт-дискомфорт. Это было необходимо сделать, чтобы верифицировать установку. После этого был поставлен эксперимент по оценки комфортности освещения в зависимости от контраста между яркостью освещения рабочего пространства к яркости адаптации, рис. 4. Кроме того, было исследовано влияние формы источника света на уровень дискомфорта выполнения зрительной работы.

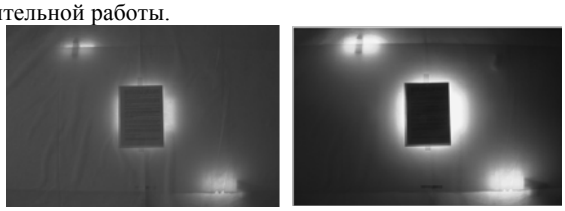


Рис. 4. Эксперимент по установлению границы комфорт-дискомфорт.

При исследовании критерия качества освещения испытуемые решали стандартную зрительную задачу – чтение текста. Вначале в течение 2-х минут проходила адаптация при средней яркости в фовиальной части 150 кд/м2. В поле зрения при этом вокруг листа формировался протяженный яркий блик. После чего яркость блика последовательно увеличивалась и задачей испытуемого было определение порогов смены ощущений при решении зрительной задачи по шкале: безразлично, дискомфортно, неудобно и невыносимо.

В следующим эксперименте, рис. 5, мы добавили блики в периферийной части зрения наблюдателя.

Рис. 5. Эксперимент с бликами в периферийной части зрения.

В момент регистрации одновременно с наблюдателями производилась съемка в формате RAW, что позволило получить пространственно-угловое распределение яркости в сцене наблюдения. В результате были получены распределения критерия качества от яркости блика.

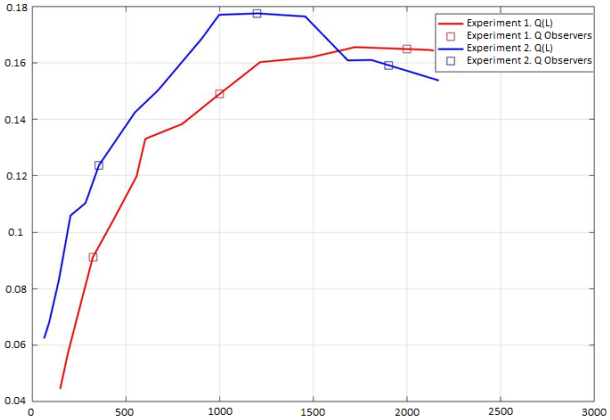


Рис. 6. Результаты экспериментов.

Как видно на рис. 6, в результате обоих экспериментов с различной яркостью адаптации мы получили близкие распределения, что подтверждает тот факт, что наш критерий качества может применяться для оценки сцен освещения с неравномерными и протяженными источниками света.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

4. Заключение

В данной работе мы провели верификация критерия оценки качества внутреннего освещения на примерах реальной зрительной работы. Результаты экспериментов говорят о том, что критерий, скорее всего, может быть использован для оценки качества освещения при светотехническом проектировании. Однако, на текущий момент, статистика еще довольно мала, и очевидно, что перед нами стоит еще большая работа по экспериментальной апробации критерия и его нормировке. Кроме того, ранее нами был разработан алгоритм расчета видо-независимого пространственно-углового распределения яркости. Применяя данный алгоритм с полученным критерием оценки качества, видится логичным попытаться автоматизировать оценку качества сцены освещения для реальных задач. Например, понятно, что наблюдатель может быть не один, и в одном и том же помещении у него может быть несколько различных зрительных задач. Таким образом в нашей будущей работе нам бы хотелось попытаться разработать методику проектирования ОУ на заданное качество освещения, основанную на полученном критерии.

5. Литература

[1] Ferree C и Rand G, The efficiency of the eye under different conditions of lighting // Tr. IES, 10, 407-415.
[2] Funke C., Schierz Ch., Extension of the unified glare rating formula for non-uniform LED luminaires // Lux junior, 2015
[3] Goral C., Torrance K., Greenberg D., Battaile B. Modeling the interaction of light between diffuse surfaces // Computer Graphics, 1984. Vol. 18, No. 3. P.213-222.
[4] Kajiya J.T. The rendering equation // Computer Graphics (Proc. SIGGRAPH'86), 1986. V.20, N4. P.143-150.
[5] Luckiesh M., Guth S.K. Brightness in the visual field at borderline between comfort and discomfort // Illuminating Engineering, 1949. V.44, No. 11. P.650
[6] Будак В.П., Желтов В.С., Мешкова Т.В., Нотфуллин Р.Ш. Оценка качества освещения на основе пространственно-углового распределения яркости // Светотехника. 2017. № 3. С. 17-22.
[7] Галлагер Р. Метод конечных элементов. Основы. —М.: Мир, 1984
[8] Vladimir P. Budak, Victor S. Zheltov, Tatiyana V. Meshkova, and Renat Sh. Notfullin, Evaluation of illumination quality based on spatial-angular luminance distibution // Svetotekhnika No. 3, 2017, pp. 17–22
[9] www.dialux.de
[10] www.relux.biz

Об авторах

Будак Владимир Павлович, д.т.н., профессор кафедры светотехники национального исследовательского университета "МЭИ".

Желтов Виктор Сергеевич, к.т.н., ассистент кафедры светотехники национального исследовательского университета "МЭИ".

Чембаев Виктор Дмитриевич, аспирант кафедры светотехники национального исследовательского университета "МЭИ".

Мешкова Татьяна Валерьевна, аспирант кафедры светотехники национального исследовательского университета "МЭИ".

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Разработка концепции источника освещения на базе Arduino Uno, автоматически поддерживающего заданный уровень освещенности

Бычкова И.В,[1], Колина В.[2]
irinka100@gmail.com|vkolina@gmail.com
[1]НИТУ «МИСиС», Москва, Россия

В данной статье рассматривается алгоритм создания автоматизированного светодиодного источника освещения, позволяющего поддерживать заданный уровень освещенности рабочего пространства. Описывается принципиальный механизм работы и точность измерений, определенная эмпирическим путем. Это не законченный рабочий проект, а его концептуальная разработка.

Ключевые слова: *автоматизация освещения, микроконтроллер, освещенность, светодиоды*

Development of concept of light source based on Arduino Uno, which automatically keeps indicated level of illumination

In the article there is considered an algorithm of development of automatic photodiode light source, which would keep indicated level of work space illumination. A principal working mechanism is described and accuracy in measurement is defined the empirical way. This is not finished working project, but it's conceptual lay-out.

Key words: lightning automation, microcontroller, illumination, photodiodes

1. Введение

Автоматизация освещения, занимающего значительную долю в современном энергопотреблении, позволяет более эффективно использовать энергетические ресурсы, что приводит к их рациональной экономии [1]. Одним из способов сокращения избыточных электрозатрат является контроль режима работы источников освещения посредством использования различных датчиков, что часто применяется в проектах «умного дома».

В данной статье рассматривается концепция автоматизации работы источника света, применимая для стационарных светильников. Использование данной схемы освещения позволяет контролировать уровень освещенности в рамках заданного значения путем изменения входного напряжения.

2. Оценка уровня освещенности

Необходимый уровень освещенности рабочей зоны зависит от параметров, определяющих уровень зрительной работы, таких как точность и разряд зрительной работы, характер фона и размер объекта различения [2]. Данный параметр является стандартизированным для каждого рода деятельности, и может считаться эталонным значением, константой, к которой необходимо стремиться. Эта величина должна быть определена предварительно при проектировании режимов работы светоисточника.

Для оценки изменения уровня освещённости используют один из трёх типов устройств: фоторезистор, фототранзистор или фотодиод [3]. В данном исследовании средством оценки освещённости выступал цифровой 16-битный датчик освещённости bh1750, в корпус которого встроен фотодиод. Преимуществами данного датчика является его высокая чувствительность и слабая подверженность влиянию инфракрасного излучения, что говорит о близости воспринимаемого им спектрального диапазона к тому, который воспринимается человеческим глазом. Значения на выходе передаются уже в люксах, в отличие от большинства альтернативных устройств, результаты работы которых поступают в неопределенных единицах измерения. Исходя из технической документации, bh1750 способен измерять освещенность в диапазоне от 1 до 65 535 лк (на практике, при использовании датчика «как есть», без внесения модификаций в стандартную библиотеку, максимальное значение освещенности, которое удалось зафиксировать, составило 54 612 лк).

Принципиальная схема подключения датчика к плате указана на рис. 1.

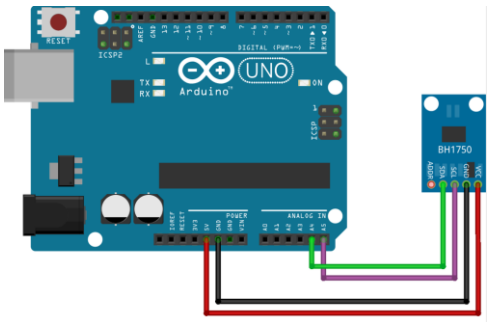


Рис. 1. Схема подключения датчика освещенности к Arduino Uno

3. Регулирование яркости светодиода

Как известно, освещённость – это величина, зависящая от силы света источника освещения (I), угла падения света (γ) и расстояния от освещаемой им поверхности (r) [4]:

$$E = \frac{I}{r^2}\cos\gamma. \quad (1)$$

Для упрощения расчета и условий проведения эксперимента, в данном исследовании предполагается расположение источника освещения перпендикулярно освещаемой им поверхности, что приводит формулу (1) к следующему виду:

$$E = \frac{I}{r^2}. \quad (2).$$

Сила света, в свою очередь, находится в зависимости от телесного угла (Ф) и светового потока (w):

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

$$I = \frac{d\Phi}{dw}. (3).$$

Где телесный угол можно определить, зная двойной угол половинной яркости (θ):

$$w = 2\pi(1 - \cos(\theta/2). (4).$$

А световой поток, в свою очередь, зависит от напряжения. Следовательно, если принять положение источника освещения статичным, уровень освещенности изменится вследствие изменения входного напряжения источника питания.

В рамках исследования был проведен ряд замеров освещенности с использованием четырех различных светодиодов при подаче напряжения от 0 до 5 В с шагом 0,5 В (рис. 2 и 3). Для проведения измерений светодиоды помещались в закрытую камеру, в которую не попадал свет извне, в которой на фиксированной высоте был установлен датчик освещённости bh175.

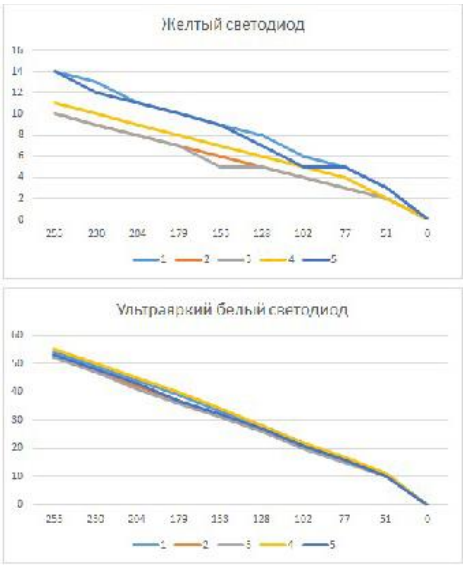


Рис. 2. Замеры освещенности желтого и белого светодиодов

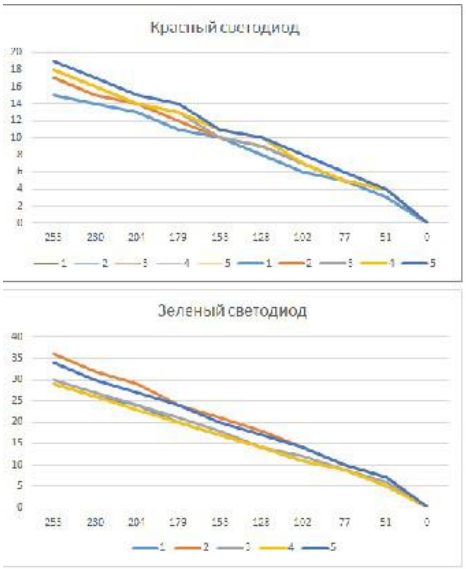


Рис. 3. Замеры освещенности красного и зеленого светодиодов

Результаты измерений показали, что при одинаковых условиях разница показаний люксметра для одного и того же светодиода может колебаться в пределе 4 лк. Данная погрешность обусловлена выбранным режимом работы датчика, позволяющим производить замер освещенности с максимальной скоростью: 16 мс/замер.

Для проверки точности показаний датчика, с использованием заданных производителем светодиодов значений светового потока – количества света излучаемого источником света (5-7 лм) и двойного угла половинной яркости (110°-120°), а также заданным в рамках опыта расстоянием от источника света до датчика (0,2 м), был произведен расчет освещенности, который показал, что в среднем освещенность от такого источника освещения должна составлять 46,7…65,3 лк, что соответствует экспериментальным значениям.

Помимо уровня освещенности была произведена оценка коэффициента пульсаций, по формуле:

$$К = \frac{Emax - Emin}{Ecp} 100\%. (5),$$

где *Emax* максимальное значение освещенности, лк,
Emin - минимальное значение освещенности, лк,
Ecp - среднее значение освещенности, лк.

Был произведен расчет значений освещенности для пяти одновременно работающих сверхъярких светодиодов для всех входных значений напряженности (с шагом 0,5 В), всего 185 замеров для каждой позиции. С использованием формулы (5) был произведен расчет коэффициента пульсаций, показавший, что данный показатель находится в пределах нормы (не превышает 10%), результаты представлены в Табл 1.

U, В	1	1,5	2	2,5	3	3,5	4	4,5	5
Emax	48	70	93	114	137	160	179	200	220
Emin	46	70	91	114	135	157	177	199	219
Ecp	46,93478	70	92,02717	114	136,0163	158,375	178	199,1413	219,0707
К	4,261232	0	2,173271	0	1,470412	1,894238	1,123596	0,502156	0,456474

Табл. 1. Расчет коэффициента пульсации

4. Разработка опытного образца

Для имитации принципа работы функции автоматизированного поддержания освещенности был собран опытный образец на базе Arduino Uno. К плате было подключено пять сверхъярких светодиодов белого цвета и датчик освещенности.

Для работы одного такого светодиода необходим ток 20 мА, а его рабочее напряжение составляет 3..3,2 В. Поскольку с одного порта платы Arduino Uno можно получить не более 40 мА, каждый светодиод был подключен к отдельному порту, дающему на выходе ШИМ-сигнал (широтно-импульсная модуляция).

Был произведен расчет сопротивления резистора с учетом рабочего напряжения платы 5 В по закону Ома для последовательного соединения цепи:

$$I = \frac{U}{R} (5),$$

где U – напряжение на резисторе, В
I – сила тока в цепи, А.

В соответствии с подученным значением 100 Ом, оказалось, что рекомендуемый для применения к другим светодиодам резистор номиналом 220 Ом является избыточным для данного светодиода, что приводит к его работе вполсилы.

Для работы микроконтроллера был запрограммирован следующий алгоритм:

1) датчик считывает уровень освещенности с интервалом в 1 с.;
2) полученное значение сравнивается с заданной константой (в данном случае значение константы было равно 300 лк – требуемый уровень освещенности для офиса);

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

3) если полученное значение освещенности больше или равно требуемому, микроконтроллер не принимает никаких действий и возвращается к шагу 1; если полученное значение меньше требуемого, на светодиоды начинает подавать напряжение (с шагом 0,5 В) до тех пор, пока не будет достигнут требуемый уровень освещенности или подаваемое напряжение не составит максимальо возможные 5 В.

5. Заключение

По результатам проведенного исследования стало понятно, что желаемая концепция устройства, регулирующего уровень освещенности может быть разработана на базе Arduino Uno. Описанный алгоритм работы является пробным вариантом и требует определенного ряда доработок, таких как автоматизация измерения расстояния от источника света до поверхности, увеличение точности измерений, создание возможности переключения между режимами работы, для того, чтобы являться полноценным рабочим образцом, однако уже на этом этапе можно сказать, что подобный подход позволяет осуществить основную задачу – поддержание заданного уровня освещенности. Следовательно, исследования в этом направлении будут продолжаться.

6. Литература

[1] Кунгс Я. А. Автоматизация управления электрическим освещением. М: Энергоатомиздат, 1989, 5 с.
[2] СНиП 23-05-95 «Естественное и искусственное освещение»
[3] Киреев П. С. Физика полупроводников. М.: Высшая школа, 1975, 584 с.
[4] Кнорринг Г.М., Фадин И.М., Сидоров В.Н. Справочная книга для проектирования электрического освещения. СПб. Энергоатомиздат, 1992, 23 с.

Об авторах

Бычкова Ирина Владимировна, научный руководитель, старший преподаватель кафедры автоматизации проектирования и дизайна факультета информационных технологий и автоматизированных систем управления национального исследовательского технологического университета, irinka100@gmail.com.

Колина Виктория, магистрант кафедры автоматизации проектирования и дизайна факультета информационных технологий и автоматизированных систем управления национального исследовательского технологического университета «МИСиС», vkolina@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Разработка математического обеспечения по расчету и оценке качества изображения оптической системы

Валиев Д.Т., Россомахина Н.Е., Тюлькин Е.В., Агапов Н.А.
rubinfc@tpu.ru | rossomahina96@mail.ru | anikolja@tpu.ru
Национальный исследовательский Томский политехнический университет, г. Томск, Россия

Работа посвящена демонстрации возможностей прикладного программного обеспечения ПО «Оптика» для синтеза, расчета и анализа оптических систем, разработанного в Томском политехническом университете. Продемонстрированы основные принципы работы с программой, ее возможности. Приведен пример расчета оптической системы линзового объектива

Ключевые слова: *программное обеспечение (ПО), оптическая система (ОС), расчет ОС.*

1. Введение

Оптические приборы и устройства применяются во многих отраслях науки и техники, где требуются высокоточные измерения и контроль, высококачественная и информативная связь и наблюдение [1]. Повышение требований, предъявляемых к современным оптическим системам (ОС), как в количественном (увеличение светосилы, полей зрения), так и качественном (повышение разрешающей способности, улучшение измерительных свойств) отношении требует повышения качества расчетов на основе создания математических моделей, позволяющих описывать реальные ОС и волновые фронты, а также учитывать векторный характер электромагнитного поля, преобразуемого ОС.

Расчет ОС – сложный процесс, особенностью которого являются большое количество трудоемких вычислений. [2, 11]. Проектирование ОС немыслимо без прикладного программного обеспечения (ПО). Сегодняшнее состояние автоматизации проектирования характеризуется переходом от отдельных программ и пакетов программ, обеспечивающих трудоемкие расчеты к системам инженерного анализа (CAE), дающим возможность осуществлять проектирование ОС на более высоком, качественном уровне. Широкий набор ПО на современном этапе (как правило, зарубежных CAE – систем, таких как Optics Studio (Zemax), CodeV, Oslo и т.д.) представляют собой конкурирующие программные продукты, позволяющие решать широкий круг задач вычислительной оптики. Несмотря на то, что функциональные возможности ПО для расчета ОС во многом совпадают, каждая из них имеет свои особенности, обусловленные основными направлениями развития и методиками решения задач, наличием собственной школы оптиков; имеет свои достоинства и недостатки, т.к. успешное решение многих задач вычислительной оптики зависит от того, насколько удачно выбран и разработан соответствующий алгоритм.

Практическое использование зарубежного ПО показывает, что главным его недостатком является реализация отдельных стадий проектирования и ориентация на зарубежную элементную базу, нормы и стандарты, что не позволяет полноценно использовать их в отечественных научно-исследовательских организациях и на предприятиях оптической отрасли. Кроме того, подходы, терминология, используемые при проектировании оптических систем и анализе качества изображения, могут отличаться от принятых отечественных. Немаловажным фактором, ограничивающим использование зарубежных САПР, является высокая стоимость программных продуктов, а также англоязычная терминология.

Среди отечественных систем автоматизированного проектирования ОС можно выделить программы DEMOS (Design, Evaluation and Modeling of Optical Systems), CAPO (Система Автоматизированного Расчета Оптики), разработанные в разные годы ВНЦ “ГОИ им. С. И. Вавилова”, а также пакет ОПАЛ (ОПтических АЛгоритмов), разработанный в ЛИТМО [12].

Однако необходимо отметить, что такие системы являются аппаратно-зависимыми и доступ к ним ограничен. Бурное развитие и быстрая смена компьютерной техники привели к тому, что использование этой системы на сегодняшний день неэффективно и невозможно. Современные условия требуют более широкой интеграции ПО и поддержки жизненного цикла изделия на всех этапах. Наличие теоретической базы и создание математического обеспечения являются хорошей основой для решения задач автоматизации проектирования ОС на совершенно другом уровне.

Целью работы является демонстрация возможностей разработанного в Томском политехническом университете современного конкурентоспособного отечественного ПО «Оптика», позволяющего решать комплексную поддержку работы оптика-разработчика, начиная с этапа синтеза исходной оптической схемы или поиска аналога, включая автоматизированную оптимизацию и анализ оптических систем.

2. Возможности программы

Алгоритмы, на которых базируется ПО «Оптика», опубликованы в работах [5-10]. ПО “Оптика” позволяет рассчитать:

- гауссовы характеристики ОС с осевой симметрией и с двумя плоскостями симметрии;
- световые диаметры оптических элементов;
- точечные диаграммы (ТД) на поверхностях ОС и в любой заданной плоскости;
- ход меридионального, сагиттального и полного пучка лучей;
- ход нескольких пучков;
- ход гауссового пучка через ОС;
- функцию концентрации энергии (ФКЭ);
- волновые аберрации и аберрации третьего порядка ОС;
- положение плоскости наилучшей установки (ПНУ);
- кардинальные элементы ОС в области реальных лучей в меридиональном и сагиттальном сечениях;
- зрачки ОС по заданному положению и величине апертурной диафрагмы;
- люки ОС по заданному положению и величине полевой диафрагмы;
- технологические параметры асферических поверхностей;
- конструктивные параметры линзы в воздухе;
- конструктивные параметры объектива Кассегрена;
- конструктивные параметры двухлинзового склеенного ахроматического объектива (предмет в бесконечности и на конечном расстоянии);
- конструктивные параметры двухлинзового несклеенного ахроматического объектива (предмет в бесконечности и на конечном расстоянии)

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

- сканирующую систему из двух плоских зеркал;
- сканирующую систему из двух оптических клиньев.

Кроме того ПО «Оптика» позволяет произвести:

- асферизацию последней поверхности ОС;
- оптимизацию параметров ОС.

Имеет базу данных на объективы.

Возможные типы систем:

- Пространственная – Р;
- Центрированная – С;

Типы поверхностей:

- Поверхности второго порядка с осевой симметрией;
- Поверхности второго порядка с двумя плоскостями симметрии;
- Поверхности второго порядка с зональной ошибкой;
- Поверхности высшего порядка;
- Конические поверхности;
- Цилиндрические поверхности;
- Тороидальные поверхности;
- Плоские дифракционные решетки;
- Диафрагмы и экраны круглые, прямоугольные, полуплоскости.

Большинство оптических систем являются «центрированными». С точки зрения разработчика на схемотехническом уровне такая оптическая система использует в качестве функциональных элементов оптические среды и оптические поверхности. Она полностью описывается такими параметрами, как радиусы кривизны поверхностей (при необходимости — коэффициенты асферики), осевые расстояния между поверхностями, показатели преломления или марки оптических материалов, из которых состоят среды, разделяющие эти поверхности, световые диаметры. Результатом проектирования является оптическая схема («оптический выпуск»), содержащая численные значения и допускаемые отклонения указанных параметров.

Форма поверхности второго порядка в программе «Оптика» задается уравнением в системе координат, расположенной в вершине поверхности:

$$x^2 + y^2 = 2 \cdot R \cdot z + (e^2 - 1) \cdot z^2$$

где R – радиус кривизны при вершине поверхности, e – эксцентриситет. Вершиной поверхности считается точка пересечения оси симметрии поверхности с самой поверхностью. В программе Оптика правило знаков принимают согласно ГОСТ 7427-76 «Геометрическая оптика» [10]. Ход луча слева направо считается положительным, справа налево – отрицательным.

3. Пользовательский интерфейс ПО «Оптика»

Внешний вид программы Оптика представляет собой две таблицы (Рис. 1):

1. Верхняя таблица определяет общие характеристики оптической системы (название системы, количество поверхностей, количество лучей, единицы измерения углов, диаметр и положение входного зрачка, положение и величину предмета, показатель преломления и направление хода луча в пространстве предметов).
2. В нижнюю таблицу заносятся конструктивные параметры ОС (радиусы кривизны поверхности, осевые расстояния, эксцентриситеты, показатели преломления, направление хода луча и другие параметры, используемые для описания формы поверхности и ее положения в пространстве).

Основное меню программы располагается в ее верхней части. Здесь можно рассчитать гауссовы характеристики, гауссовы пучки, световые высоты, оптимизировать оптическую систему, рассчитать положение зрачков и люков системы и т.д. Также программа может произвести графическое построение точечной диаграммы в заданной плоскости (ЗП), плоскости Гаусса (ПГ) и в плоскости наилучшей установки (ПНУ).

Одной из особенностей программы является автоматическая запись отчета о проведенных расчетах, которая формирует все результаты в единый документ.

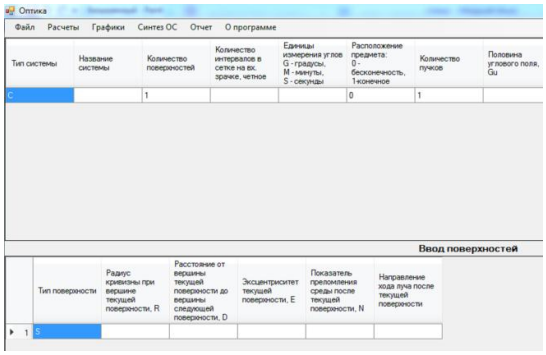


Рис. 1. Фрагмент внешнего вида ПО «Оптика».

4. Расчет линзового объектива

В качестве примера работы с программой произведем расчет линзового объектива, состоящего из четырех поверхностей [9]. Схема линзового объектива приведена на Рис. 2. Параметры объектива: $2\omega = 5^0$, $\forall = 1:2$, $\lambda = 10$ мкм, материал – Ge.

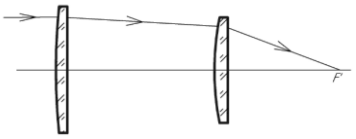


Рис. 2. Оптическая схема линзового объектива.

Отчет о проведенных расчетах

Данные по системе:

Тип системы: С

Название системы: Линзовый объектив

Количество поверхностей: 4

Количество интервалов в сетке на входном зрачке: 50

Единицы измерения углов: G (градусы)

Расположение предмета: 0 (бесконечность)

Количество пучков: 1

Половина углового поля, G_u:0

Диаметр входного зрачка: 50

Координата центра входного зрачка по оси X, Xc: 0

Координата центра входного зрачка по оси Y, Yc: 0

Показатель преломления среды в пространстве предметов: 1

Направление хода луча в пространстве предметов: 1

Данные по поверхностям приведены в таблице 1.

Таблица 1- Данные по поверхностям линзового объектива

Тип поверхности	R	d	e	n	δ
S	1025,64	3,37	0	4,0052	1
S	0	104,8	0	1	1
S	293,69	4,67	0	4,0052	1
S	0	0	0	1	1

где S – поверхности второго порядка,

R – радиус кривизны при вершине текущей поверхности, мм,

d – расстояние от вершины текущей поверхности до верши-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

ны следующей поверхности, мм,
e - эксцентриситет поверхности,
n – показатель преломления среды после текущей поверхности,
δ – направление хода луча после текущей поверхности.

Результаты расчетов

1) Гауссовы характеристики

Таблица 2 - Гауссовы характеристики линзового объектива

Переднее фокусное расстояние	-100,047268
Заднее фокусное расстояние	100,047268
Передний фокальный отрезок	8,102011
Задний фокальный отрезок	67,912950
Расстояние от вершины передней поверхности до передней главной точки	108,149278
Расстояние от вершины задней поверхности до задней главной точки	-32,134318
Расстояние от вершины передней поверхности до передней узловой точки	108,149278
Расстояние от вершины задней поверхности до задней узловой точки	-32,134318
Величина гауссового изображения	0

2) Полный пучок. Точечная диаграмма в ЗП
Количество лучей на входе системы: 1961
Количество лучей на выходе системы: 1961
Координаты X, Y центра тяжести пятна:
X=0, Y=0
Размеры кружка рассеяния (координаты границ точечной диаграммы):
X_{min}= -16,948985, Y_{min}=16,948985
X_{max}= -16,948985, Y_{max}=16,948985

3) Полный пучок. Точечная диаграмма и функция концентрации энергии (ФКЭ) в ПГ (Рис. 3-4)
Количество лучей на входе системы: 1961
Количество лучей на выходе системы: 1961
Координаты X, Y центра тяжести пятна:X=0, Y=0
Размеры кружка рассеяния (координаты границ точечной диаграммы):
X_{min}= -0,537609, Y_{min}=0,537609
X_{max}= -0,537609, Y_{max}=0,537609

4) Полный пучок. Точечная диаграмма и ФКЭ в ПНУ (Рис. 3-4)
Расстояние до ПНУ: 66,5201369967688
Координаты X, Y центра тяжести пятна:
X=0, Y=0
Размеры кружка рассеяния (координаты границ точечной диаграммы):
X_{min}= -0,178980, Y_{min}=0,178980
X_{max}= -0,178980, Y_{max}=0,178980

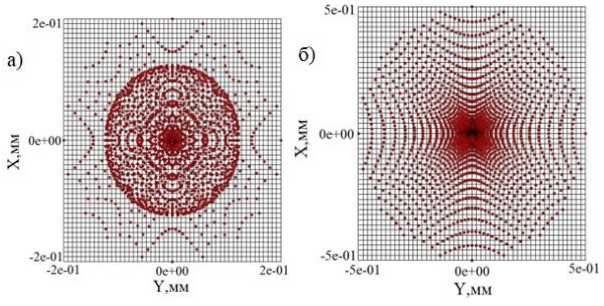


Рис. 3. Точечная диаграмма в ПНУ (а) и ПГ (б).

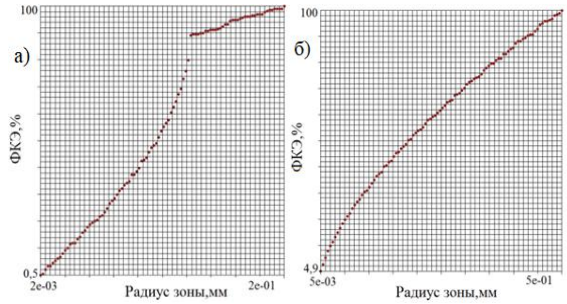
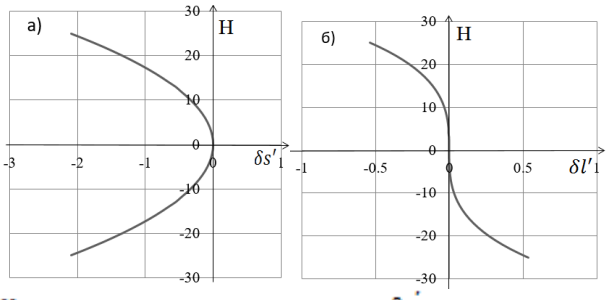


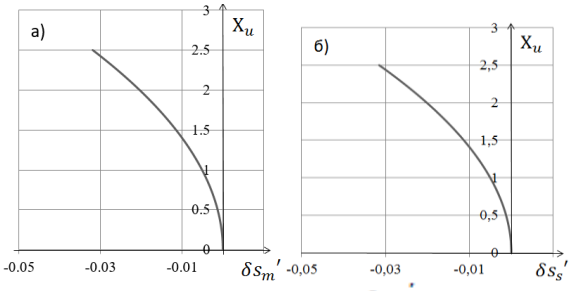
Рис. 4. ФКЭ в ПНУ (а) и в ПГ (б).

5) Расчет аберраций Зейделя третьего порядка. (Рис. 5-8). Плоскость предмета расположена в бесконечности.



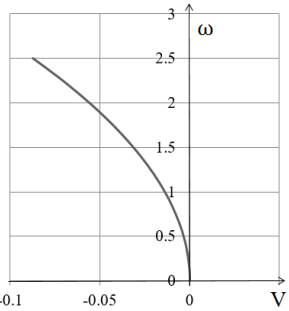
H — высота луча на входном зрачке, $\delta s'$ – продольная сферическая аберрация, $\delta l'$ – поперечная сферическая аберрация

Рис. 5. Продольная (а) и поперечная (б) сферическая аберрация.



X_u– координата точки предмета, $\delta s_m'$ –меридиональный астигматизм, $\delta s_s'$ –сагиттальный астигматизм

Рис. 6. Меридиональный (а) и сагиттальный (б) астигматизм.



ω – угловая величина предмета (градусы), V-дисторсия

Рис. 7. Дисторсия.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

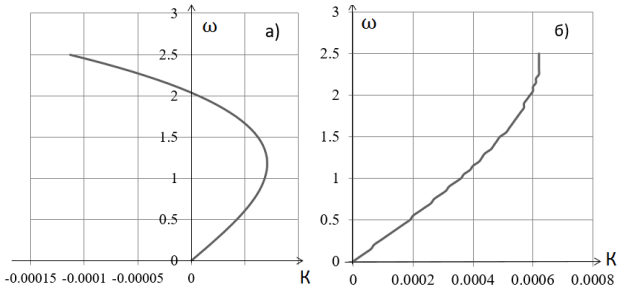


Рис. 8. Меридиональная кома в ПГ (а) и в ПНУ (б).

6) Расчет волновых аберраций.

Для примера рассчитаем волновые аберрации линзового объектива в меридиональном сечении в ПНУ и в ПГ. Точка предмета находится в бесконечности и на оси (Рис. 9). Волновые аберрации рассчитываются относительно двух сфер сравнения: центр первой сферы сравнения расположен в центре тяжести ТД в ПГ, а вершина – в точке пересечения главного луча с выходным зрачком; центр второй сферы сравнения расположен в центре тяжести ТД в ПНУ, а вершина также – в точке пересечения главного луча с выходным зрачком.

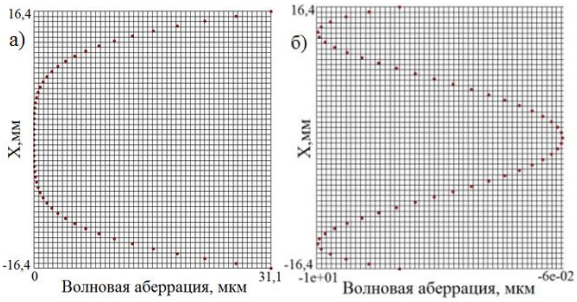


Рис. 9. Меридиональный профиль волновых аберраций в ПГ (а) и в ПНУ (б).

7) Расчёт производных.

Осуществляется расчёт производных от заднего фокального отрезка по расстояниям между поверхностями и по радиусам кривизны поверхностей в параксиальном приближении. Результаты расчета приведены в Таблице 3.

Таблица 3 – Производные от заднего фокального отрезка

№	Производные $\frac{\partial S'_{F'}}{\partial d_i}$	Производные $\frac{\partial S'_{F'}}{\partial R_i}$
1	-0,0214557212018727	0,0285952530805739
2	-0,0859344545577407	–
3	-0,249675421951463	0,166259441408024
4	–	–

Зная производные, можно рассчитать допуски на расстояния между поверхностями и радиусы кривизны поверхностей при заданном допуске на задний фокальный отрезок.

5. Заключение

Значимость работы обусловлена ее направленностью на решение как фундаментальных так и практических вопросов расчета ОС и обработки результатов измерений в процессе контроля оптических деталей, узлов и материалов фотоники. Полученные в работе результаты позволяют:

-на этапе проектирования получать полную информацию о качестве ОС, оценить допуски и технологичность ОС;

-применять отечественное программное обеспечение, обеспечивающее решение основных задач по расчету, синтезу и анализу качества изображения оптических систем;

-обеспечить возможность подготовки высококвалифицированных кадров в области фотоники (расчет и конструирование оптических систем): включение в образовательный процесс.

Литература

[1] Philip C. D. Hobbs Building Electro-Optical Systems: Making it All Work. 2000 John Wiley & Sons, Inc. p. 710.

[2] Warren J. Smith Modern Optical Engineering 3rd ed. McGraw-Hill, 2008, p. 641

[3] Агапов Н.А. Двухзеркальная сканирующая система. // ОМП, 1986, №6, 11-13 с.

[4] Агапов Н.А. К вопросу о главных поверхностях оптических систем. Ред. журн. «Изв. вузов. Физика». - Томск, 2010. - Деп. в ВИНИТИ 20.12.2010, Рег. № 707-В2010, 24 с.

[5] Агапов Н.А., Кулешов В.К., Корнев В.А. Расчет хода лучей через конические поверхности второго порядка, не обладающие осевой симметрией. // Образование. Наука. Научные кадры, 2013, №1, с. 123-133.

[6] Агапов Н.А., Лялин Ю.В., Калошина Х.С. Свидетельство об официальной регистрации программы для ЭВМ ППП “ОПТИКА” № 960502, заявка № 960412. Зарегистрировано в Реестре программ для ЭВМ 22.11.96.

[7] Агапов Н.А. Оптика реальных лучей: монография / Н.А. Агапов, В.К. Кулешов; под ред. д-ра техн. наук Н.А. Агапова. – Томск: Изд-во Томск. гос. ун-та систем упр. и радиоэлектроники, 2017. 100 с.

[8] Герцбергер М. Современная геометрическая оптика. – М.: ИЛ, 1962, 487 с.

[9] Ефремов В.С. Оптические устройства оптико-электронных приборов: методические указания по курсу «Конструирование оптико-электронных приборов», Изд-во НИИГАиК, 1991. – 37 с.

[10] ГОСТ 7427-76 Геометрическая оптика. Термины, определения и буквенные обозначения; М.: Издательство стандартов, 1988, – 19 с.

[11] Шеннон Р., Вайант Дж. Проектирование оптических систем, Мир, 1983. 422 с.

[12] Шехонин А.А. Методология проектирования оптических приборов: учеб. пособие / А.А. Шехонин, В.М. Домненко, О.А. Гаврилина.– СПб: СПбГУ ИТМО, 2006. – 91 с.

Об авторах

1) Валиев Дамир Талгатович, к. ф.-м. н., доцент отделения Материаловедения Томского политехнического университета. E-mail: rubinfc@tpu.ru

2) Россомахина Наталья Евгеньевна, студент отделения Материаловедения Томского политехнического университета. E-mail: rossomahina96@mail.ru.

3) Тюлькин Евгений Викторович, магистрант Томского политехнического университета

4) Агапов Николай Афанасьевич, д. т. н., профессор отделения Материаловедения Томского политехнического университета. E-mail: anikolja@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Моделирование световых приборов автомобилей с использованием трассировки лучей

Грачёв Д.Г.[1], Будак В.П.[2]

genflare@gmail.com|budakvp@gmail.com

[1]Московский Энергетический Институт, кафедра светотехники, Москва, Россия;

[2]Московский Энергетический Институт, кафедра светотехники, Москва, Россия

Выявлены основные преимущества использования адаптивного светораспределения перед с традиционным на примере автомобильной фары. Создана оптическая система, демонстрирующая возможность управления геометрией светового пучка за счёт использования в качестве источника излучения матрицы СИД. Обозначены функции, реализуемые только при данной конструкции фары. Фотографии, представляющие собой поле зрение водителя в ночное время суток, были совмещены с результатами анализа распределения освещённости, выполненного в программном комплексе TracePro, при включении светодиодов в определённой комбинации.

Ключевые слова: *матрица СИД, автомобильная фара, адаптивное светораспределение, TracePro.*

1. Постановка задачи и определение способа её решения

Совершенствование систем автомобильного освещения было и остаётся важным вопросом при обеспечении безопасности движения и эффективности перевозок ночью. При нынешней динамике вождения и плотности движения очевидна невозможность постоянного использования режима дальнего света. Безгранично увеличивать яркость автомобильной фары недопустимо, так как после достижения некоторой величины работоспособность органа зрения водителя снижается вследствие возникающего зрительного утомления. Одновременно это приносит дискомфорт другим участникам движения. В связи с этим технология автомобильного освещения сегодня развивается в направлении повышения количества освещённого пространства, не достигая при этом уровня избыточной яркости. Современные разработки направлены, главным образом, на системы, включающие в себя функцию обнаружения других участников движения и всевозможных ориентиров (сигнальные столбики, разметка, дорожные знаки) с помощью камеры, радара, тепловизора и прочими способами, для последующего изменения светотеневой границы [2].

Матричный световой прибор содержит в себе определённое количество светодиодов, каждый из которых управляется электроникой. Вся эта система лишена поворотных механизмов, а изменение геометрии светового пучка осуществляется микропроцессором, которой по отдельности меняет яркость светодиодов либо отключает их. Это позволяет плавно изменять параметры освещения, благодаря чему постоянно сохраняется освещенность, максимально достижимая при данной конструкции фары. Использование в качестве источника излучения кластера светодиодов позволяет реализовать уникальные в своем роде функции и иначе подойти к вопросам автомобильного освещения. Наличие подобной светотехники позволяет автономно выполнять следующие функции:

- предотвращать ослепление водителей впередиидущих и встречных транспортных средств;

- плавно изменять параметры освещения, благодаря чему освещенным остается всё пространство перед автомобилем за исключением зон, занимаемых другими транспортными средствами;

- адаптироваться к дорожной топологии (регулировка угла наклона фар относительно продольной плоскости автомобиля);

- освещать дорожные знаки, пешеходов;

- сигнализировать о наличии животных на проезжей части;

- освещать повороты на основании навигационных данных.

2. Моделирование оптической части светового прибора

С помощью программного комплекса TracePro [1] была создана простейшая оптическая система (далее ОС), демонстрирующая возможность управления светораспределением за счёт использования матрицы СИД в качестве источника излучения. Первоначально было выполнено моделирование светодиода, имеющего светоизлучающую поверхность квадратной формы площадью 1 мм^2 и отражательный коллиматор, по форме напоминающий простейший диффузор. КСС таким образом представляет собой гауссоиду, а излучаемый поток составляет 200 лм. Затем была установлена первичная оптика в виде плосковыпуклой линзы, радиус кривизны которой подобран таким образом, чтобы при её установке сверху коллиматора световые пучки от соседних элементов перекрывались в плоскости вторичной оптики примерно наполовину.

После этого был сформирован кластер из данных элементов размерностью 3x11, а перед ним установлена вторичная оптика в виде двояковыпуклой линзы (рис.1). Она фокусирует излучение светодиодов на плоскости анализа, установленной на расстоянии 150 м, в пятно размерами 25 м по горизонтали и 7 м по вертикали при полной комбинации включения, чего достаточно для заполнения светом всего дорожного пространства вместе с обочиной.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

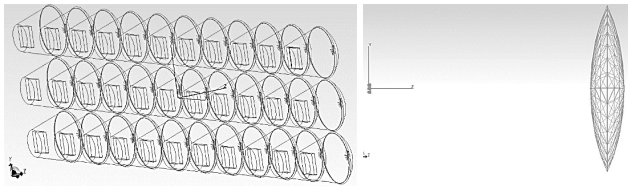


Рисунок 1. Слева - кластер СИД, справа – к ОС добавлена вторичная оптика.

3. Результаты

Изменение комбинации включения светодиодов позволяет получать светораспределения различной формы. Кроме этого, можно изменять и яркость каждого элемента (аналогично диммированию светодиода за счёт изменения величины подводимого тока).

Для большей наглядности, фотографии, представляющие собой поле зрения водителя в ночное время суток, были совмещены с результатами распределения освещённости, полученными в TracePro при включении светодиодов в определённой комбинации.

Рисунок 2. Совмещение изображения поля зрения водителя с результатом анализа распределения освещённости.

4. Заключение

Приведённые изображения показывают, что световой прибор, формирующий адаптивную светотеневую границу, можно использовать в качестве замены штатной системы освещения транспортного средства. Но прежде его необходимо перестроить под данную задачу по множеству параметров, определяемых ГОСТ Р 41.48-2004. Также можно задуматься о его оптимизации:

1. В первую очередь, на основе пороговых характеристик органа зрения в условиях темновой адаптации (определив, по крайней мере, максимальное разрешение, после которого более мелкое разбиение пространства не сегменты будет уже нецелесообразным).

2. Во-вторых, следует обеспечить неравномерность разбиения пространства перед автомобилем на сегменты, по крайней мере, в левой половине поля зрения водителя, где светотеневая граница изменяется наиболее активно вследствие частого появления объектов, в направлении которых распространение света необходимо исключать. Плотность разбиения целесообразно снижать от центра к периферии, так как в этом направлении скорость перемещения объектов в поле зрения возрастает при движении автомобиля по проезжей части.

После комплексного анализа требований, предъявляемых к системам освещения транспортного средства, можно приступать к разработке настоящей автомобильной фары.

5. Литература

[1] Aber O., Geimer M., Muller S. *Direct and Fast Ray Tracing of NURBS Surfaces* // IEEE Symposium on Interactive Ray Tracing, 2006. P.161-168. – ISBN 1-4244-0693-5

[2] Kauschke R., Eichhorn K., Wallaschek J. *Innovative optical concepts for future car headlights* // PAL 2003, Volume 10. – TU Darmstadt, 2003.

Об авторах

Грачёв Дмитрий Геннадьевич – студент Московского Энергетического института, кафедры светотехники. Контактный адрес genflare@gmail.com.

Будак Владимир Павлович – д.т.н. Московского Энергетического института, кафедры светотехники. Контактный адрес budakvp@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Снижение ресурсоемкости дизайн - проектирования

Ю.С. Дё, А.В. Шкляр
dyojulias@gmail.com|shklyarav@mail.ru
г. Томск, Томский Политехнический Университет

В данной работе рассматривается проблема снижения ресурсоемкости дизайн – проектирования за счет своевременной апробации дизайн решения. Своевременная апробация дизайн решения приводит к снижению ресурсоемкости процессов производства, что в свою очередь влияет на себестоимость продукта. В распространенной схеме дизайн – проектирования решение апробируется на завершающей стадии – прототипирование. Прототипирование это изготовление изделия с использованием технологий и материалов близких к производственным. При обнаружении ошибки в следствии тестировании прототипа необходимо возвращаться к предыдущим ступеням производственной цепочки. Применение метода проектирования, за счет использования виртуального тестирования позволяет апробировать решение на любом из необходимых этапов в режиме реального времени, что приводит к снижению ресурсоемкости проектирования.

Ключевые слова: *виртуальное тестирование, компьютерная графика, ресурсоемкость, снижение себестоимости.*

Reduction of resource intensity - design

J.S. Dyo, A.V. Shklyar
dyojulias@gmail.com|shklyarav@mail.ru
Tomsk, Tomsk Polytechnic University

In this scientific article, we consider the problem of reducing the resource intensity of design-design due to the timely approbation of the design of the solution. Timely approbation of the design solution leads to a decrease in the resource intensity of production processes, which in turn affects the cost of the product. In a common design design scheme, the solution is tested at the final stage - prototyping. Prototyping is the manufacture of a product using technologies and materials close to production. If an error is detected during the testing of the prototype, it is necessary to return to the previous stages of the production chain. Application of the design method, through the use of virtual testing, allows you to test the solution at any of the necessary stages in real time, which leads to solving the resource-intensive design.

Keywords: virtual testing, computer graphics, resource intensity, cost reduction.

1. Введение

Общероссийский классификатор продукции по видам экономической деятельности является нормативным документом, описывающим какие работы, товары или услуги необходимо получить от поставщика.

Услуги, предоставляемые в области промышленного дизайна, по ОКПД имеют код 74.10.12.000 и описаны как:

« …создание и разработка дизайна и спецификаций для оптимизации использования, стоимости и внешнего вида продуктов, включая определение материалов, конструкции, механизма, формы, цвета и поверхностной обработки продукта, принимая во внимание человеческие характеристики и потребности, требования безопасности, привлекательности для потребителя и эффективности распространения, использования и технического обслуживания» [1].

Если ознакомиться с официальными документами, которые определяют не только саму трактовку понятия "промышленный дизайн", но и требования к знаниям и умениям у специалистов в этой области, можно сделать вывод и о задачах, которые стоят при разработке устройства:

- повышение эстетических качеств, для ориентирования на выбранную или заданную сегментацию рынка;
- снижение ресурсоемкости процессов производства, для уменьшения себестоимости устройства;
- повышение функциональности устройства для улучшения потребительских свойств;

Рассмотрим вопрос о снижении ресурсоемкости процессов производства.

2. Этапы разработки промышленных изделий

При разработке промышленных объектов используют распространенный алгоритм производственной цепочки, в котором учувствует ряд специалистов.

Стоимость каждого из этапов и специалистов оплачивается отдельно. Следовательно, чем более ресурсоемкая разработка, тем дороже себестоимость разработки.

Для того что бы снизить ресурсоемкость дизайн - проектирования рассмотрим алгоритм производственной цепочки, представленный студией промышленного дизайна «Форма». [2] (Рисунок 1)

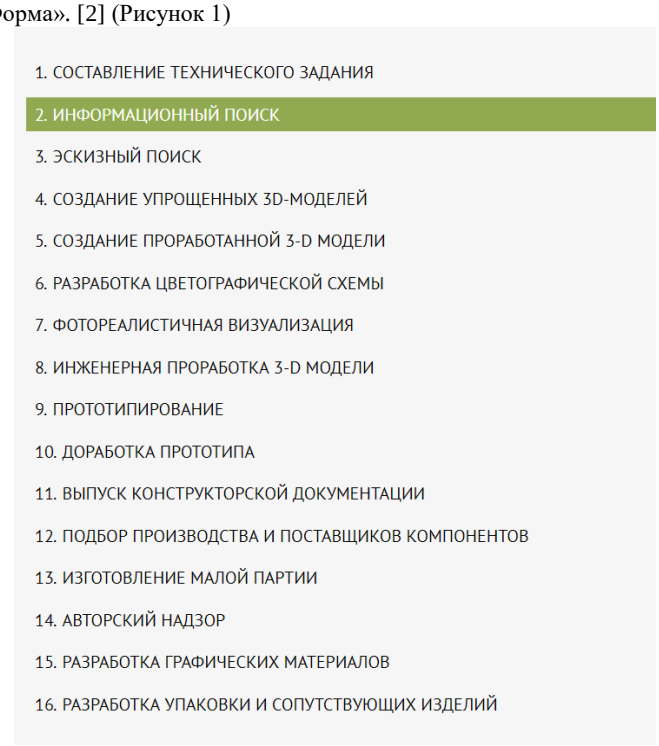


Рисунок 1. Технологическая цепочка, представленная на сайте студии «Forma»

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Подтверждение было найдено в методическом указании Московского государственного технического университета имени Н.Э. Баумана, «Введение в дизайн-проектирование» [3].

Для выявления недостатков технологической цепочки необходимо проанализировать каждый из этапов.

Пункт 2-7 относиться к задачам дизайна, 8-10 к инженерным задачам.

1) Эскизный проект.

Эскиз выполняется от руки в глазомерном масштабе. [4] Задача этой стадии – поиск различных концепций для последующей проработки.

На этапе эскизирования невозможно апробировать решение, потому что информация о формообразовании и свойствах изделия являться неполной.

Таким же недостатком обладает следующий этап - этап трехмерного моделирования.

2) Трехмерное моделирование

Трехмерные модели создаются на основе наиболее удачных эскизов.

На этапе трехмерного моделирование отсутствует полноценное представление о свойствах изделия.

Для апробации физических свойств на этапе трехмерного моделирования можно применить виртуальное тестирование или макетирование. В дизайне термин «виртуальное тестирование» не определён до конца, и чаще всего встречается в области инжиниринга, где применяются программные средства автоматизированного проектирования, включающие системы для моделирования твердотельных объектов, расчёта динамики твёрдых и упругих тел, расчёт прочности и деформаций конструкции и т.п. [4]

Макет выполняется из материалов, позволяющих постоянно вносить изменения в форму (пеноплэкс, пластилин, картон, бумага и т.д.) [5].

Так как материал и способ изготовления влияет на физические свойства изделия, жесткость, мягкость, радиопрозрачность, электропроводимость, изломостойкость и т.д.

Макетирование не дает полного представления о реальных физических свойствах объекта и позволяет оценить лишь отдельные характеристики изделия.

Возникают трудности в апробации решения.

3) Разработка цветографической схемы

Для объекта подбираются самые выигрышные цвета, наносятся графика, логотипы и все необходимые надписи. Данный этап можно отнести к этапу 3д моделирования, или эскизирования.

4) Фотореалистичную визуализацию нужно вынести после инженерных работ в случае необходимости наличия презентационной картинки продукта. При инженерных работах форма может быть скорректирована и смысла делать фотореалистичную визуализацию – нет.

5) Прототипирование.

Итоговый шаг в создании дизайна – прототипирование. Прототипирование – это физическая апробации решения.

Прототип изготавливается из материалов и технологий близким по свойствам к реальным. Потому что материал и способ изготовления влияет на физические свойства изделия, жесткость, мягкость, радиопрозрачность, электропроводимость, изломостойкость и т.д. [6]

Если прототип не отвечает заданным характеристикам, то нужно возвращаться к предыдущим этапам. А изготовление прототипа дорогостоящий и время затратный процесс, который может потребовать многократной апробации.

3. Поиск решения задач снижения ресурсоемкости дизайн – проектирования на примере медицинского инъектора

Недостаток рассмотренной технологической цепочки - отсутствие проверки решения на каждом из этапов. Поэтапная проверка позволит избежать не нужных действий и корректировать ошибки сразу, следовательно, снизить ресурсоемкость при проектировании изделия.

Выдвинутое утверждение подтверждается решением задач дизайн - проектирования на примере разработки индивидуального, медицинского инъектора.

Медицинский инъектор - прибор для инвазивного введения лекарственных препаратов.

Дизайнеру было выдано техническое задание от студии промышленного дизайна «Карфидов Лаб» [9] - готовое конструктивное решение (Рисунок 2), к которому нужно было разработать корпус.

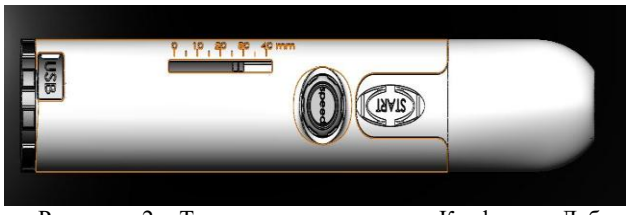


Рисунок 2. Техническое задание «Карфидов Лаб». Конструктивное решение.

Для разработки персонального медицинского инъектора необходимо было создать ряд эскизов, удовлетворяющих требования технического задания. В результате проведенной работы, были предложены эскизы оболочки персонального медицинского инъектора, которые соответствовали выдвинутым требованиям.

Рисунок 3. Эскизный проект

Эскизный проект был согласован и принят заказчиком на инженерную доработку.

Однако в ходе проведенной работы, был выявлен ряд недочетов, касающихся требований к дизайну устройства, указанных в техническом задании. Возникла необходимость проведения дизайн – исследования, результаты которого необходимы для преодоления выявленных недочетов

Так как конструкция была предоставлена, дизайнерское решение ограничивалось компоновкой и расположением органов управления, габаритными размерами.

При этом задачи дизайна и конструктивные решения вступают в конфронтацию. Например, параметр мобильности не сопоставим с габаритными размерами конструкции, фиксированные органы управления и

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

заданные габаритные размеры не позволяют работать с эргономикой устройства.

Следовательно, необходимо изменить параметры конструкции, а для этого дизайнер должен обладать специальными навыками и знаниями и иметь возможность апробировать решение, при этом управлять процессом тестирования, для внесения своевременных правок.

Актуальным решением выступает виртуальное тестирование.

4. Виртуальное тестирование как инструмент снижения ресурсоемкости дизайн – проектирования

В дизайне термин «виртуальное тестирование» не определён до конца, и чаще всего встречается в области инжиниринга, где применяются программные средства автоматизированного проектирования, включающие системы для моделирования твердотельных объектов, расчёта динамики твёрдых и упругих тел, расчёт прочности и деформаций конструкции и т.п. [7]

Определим виртуальное тестирование в дизайне как процедуру исследования свойств виртуальной модели дизайн - объекта, на базе компьютерных технологий.

Создание симуляции реальных условий физического мира и интерактивное управление процессом тестирования позволил осуществить программный пакет 3D-моделирования – autodesk 3ds Max, на основе встроенного плагина Mass FX**.** [8]

Так как этап инжиниринга был определен как первоочередный, дизайнер должен апробировать предполагаемые конструктивные решения.

5. Постановка эксперимента

Инъектор совершает три последовательных действия

- Ввод иглы
- Ведение лекарственного препарата за счет движения поршня
- Вывод иглы

Данный алгоритм действий выполняется за счет манипулирования тремя элементами управления.

Для предотвращения ошибки при эксплуатации пользователем, улучшения эргономических характеристик, объединим заданные функции в один элемент управления, за счет разности сопротивления.

Создана сцена, имитирующая выполнение процедуры инъекциирования. (Рисунок 4)

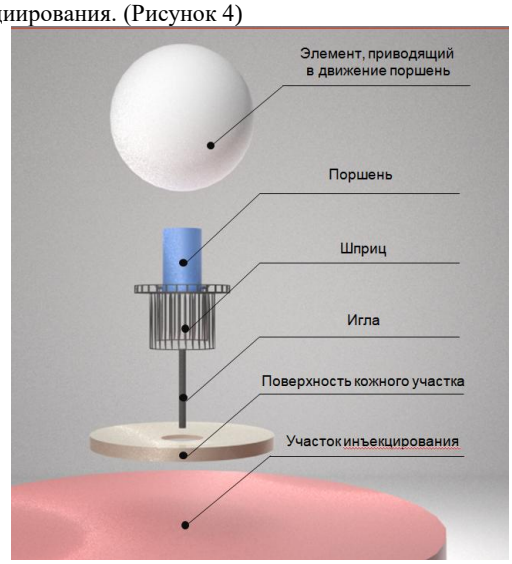


Рисунок 4. Смоделированная сцена для проведения виртуального тестирования.

Зададим условия тестирования:

- Необходимо соблюсти габаритные размеры шприца, иглы и поршня
- Шприц должен совершить обратно поступательное движение, без каких либо резких скачков, и отклонений от вертикальной оси хода шприца.
- Необходимо создать три элемента сопротивления. (Рисунок 4)
- При совершении инъекциирования игла должна полностью войти на заданную глубину, при этом поршень не должен начать ход до тех пор, пока игла не будет полностью введена.
- После введения иглы поршень начинает движение - ввод лекарственного препарата.
- После полного введения лекарственного препарата игла должна полностью выйти. Выход иглы не должен происходить за счет вытягивания шприца за поршень, потому что существует вероятность всасывания лекарственного препарата обратно в шприц. Следовательно, сопротивление возвратному ходу необходимо устанавливать под шприцом.

Было установлено три сопротивления (Рисунок 5):

1) сопротивление поршня (введения лекарственного препарата);
2) сопротивление кожи при введении иглы;
3) обратное движение (вывод иглы);

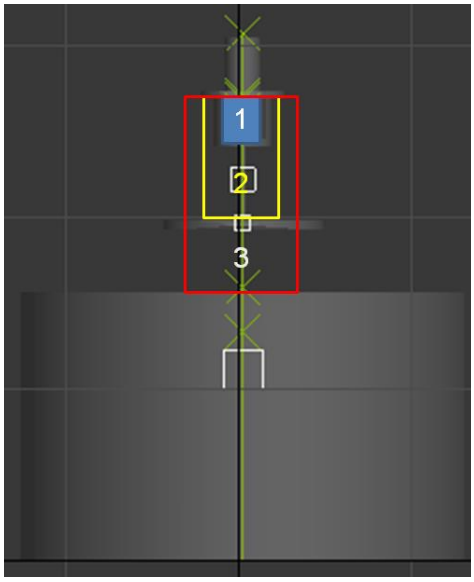


Рисунок 5. Обозначение зон сопротивления

В ходе эксперимента были соблюдены все условия тестирования, и сделан вывод о работе конструкции: разница в сопротивлении между 1 и 2 пружиной должна быть меньше сопротивление третьей.

С учетом того что принцип работы конструкции с соблюдением габаритных размеров был успешно апробирован в программе 3д моделивания, трехмерное моделирование возможно выполнять в той же сцене. Это позволит при необходимости совместить проведение испытаний корпуса с симуляцией работы конструкции.

7. Заключение

В распространенной схеме дизайн - проектирования апробация решения выполняется за счет прототипирования. При этом в распространенной технологической цепочке этап прототипирования является завершающим. И для

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

корректировки необходимо обращаться к предыдущим этапам проектирования.

Этап прототипирования может проводиться неоднократно. А это вложение как временного, так и финансового ресурса. Что в свою очередь влияет на себестоимость разрабатываемого продукта.

В связи с этим виртуальное тестирование имеет ряд преимуществ. Так как может проводиться на разных этапах дизайн - проектирования, имеет возможность интерактивного управления процессом для внесения своевременных правок.

Снижение ресурсоемкости дизайн – проектирования при использовании метода виртуального тестирования очевидно.

8. Литература

[1] Общероссийский классификатор экономической деятельности. URL: http://www.okvad.ru/razdel-m.html (дата обращения 10.04.2018)

[2] Студия «Forma» промышленный, предметный дизайн и инжиниринг. URL: http:// formaforma.ru/ (дата обращения 17.04.2018)

[3] МГТУ имени Н.Э. Баумана, «Введение в дизайн - проектирование». URL:http://design.bmstu.ru/ru/metodichki/Bakalavriat/Vvedenie%20v%20professiiu.pdf

[4] Михаил Куприков, Леонид Маркин. Инжинерная графика.изд. Москва Дрофа 2011

[5] В.А Смирнов Профессиональное макетирование и техническое моделирование. Краткий курс. Учебник пособие : Изд. "Проспект", 2016

[6]Формлаб.Как и где изготовить прототип корпуса. URL: https://formlab.ru/tehnologii-dlya-prototipa

[7] А. И. Боровков. Компьютерный инжиниринг : учеб. пособие / [и др.]. - СПб. : Изд-во Политехн. ун-та, 2012. - 93

[8] 3ds Max Learning Center. URL: http: // help.autodesk.com/view/3DSMAX/2018/ENU/?caas=caas/discussion/t5/3ds-Max-Forum/Mass-FX-is-gone/td-p/7586985.html (дата обращения 17.04.2018)

[9]Студия промышленного дизайна «Карфидов Лаб». URL: https://karfidovlab.com/ (17.04.2018)

9. Об авторах

Шкляр Алексей Викторович, старший преподаватель, отделение автоматизации и робототехники, Томского политехнического университета. E-mail shklyarav@tpu.ru.

Дё Юлия Станиславовна, студентка инженерной школы информационных технологий и робототехники (бывш. Институт кибернетики) Томского политехнического университета. E-mail dyojulias@gmail.com.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Расчет и конструирование оптико-механической системы зрительной трубы Ньютона с ортоскопическим окуляром

Р.Г. Калинин, Д.Т. Валиев, Агапов Н.А.
kalininrostislav@gmail.com|rubinfc@tpu.ru|anikolja@tpu.ru
Национальный исследовательский Томский политехнический университет, г. Томск, Россия

Работа посвящена расчету и проектированию оптико-механической системы зрительной трубы Ньютона с ортоскопическим окуляром. Для расчета и анализа оптической системы использовалось отечественное прикладное программное обеспечение (ПО) «ОПТИКА». 3D – моделирование осуществлялось с использованием CAD системы «Solidoworks».

Ключевые слова: зрительная труба Ньютона, ПО «ОПТИКА», расчет оптических систем

1. Введение

Оптические приборы и системы представлены практически во всех областях научной и технической деятельности. Так, например, для наблюдения за удалёнными объектами используются телескопические системы различной конструкции. Телескопические системы позволяют не только наблюдать за удалёнными объектами, но и измерять различные параметры наблюдаемого объекта. Чаще всего телескопические системы применяются в астрономии, как в профессиональной, так и в любительской.

Современные телескопы являются сложными оптическими системами. Проектирование таких систем представляет собой достаточно трудоемкий процесс, который включает в себя как этапы теоретического исследования, так и практические этапы [1, 5]. Как правило, зеркальная оптика рассчитывается совместно с линзовыми или зеркальными устройствами меньшего размера корректорами поля, редукторами фокусного расстояния, адаптивными системами и т.д. [7]. При этом проектировщик должен учитывать множество факторов, в том числе технологические возможности производства разрабатываемых оптических систем, стоимость используемых материалов и возможность их поставки и др. [3, 6]. Кроме того, специалисты, способные квалифицированно разработать оптические системы для решения задач, поставленных перед оптико-электронным прибором, должны обладать фундаментальными знаниями по общеобразовательным и техническим дисциплинам, а также в совершенстве знать основы оптики и оптические системы приборов.

Работа посвящена расчету и проектированию несущей оптико-механической системы зрительной трубы Ньютона с ортоскопическим окуляром. Авторы ставили целью продемонстрировать возможности разработанного в Томском политехническом университете прикладного программного обеспечения ПО «Оптика», позволяющего решать комплексную поддержку работы оптика-разработчика, начиная с этапа синтеза исходной оптической схемы или поиска аналога, включая автоматизированную оптимизацию и анализ оптических систем.

2. Расчёт оптической системы зрительной трубы Ньютона

В общем виде зрительная труба Ньютона представляет собой трубу, установленную на монтировке, снабжённой осями для наведения на объект наблюдения и слежения за ним. Основными элементами данного вида зрительных труб являются объектив и окуляр (рис. 1). В зависимости от типа оптической схемы большинство телескопов делятся на линзовые (рефракторы), зеркальные (рефлекторы) и зеркально-линзовые.

На сегодняшний день существует несколько типов оптических систем зеркальных телескопов, среди которых рефлекторы с системой Ньютона получили наибольшее распространение среди любительских телескопов, что обусловлено относительной простотой конструкции и высоким соотношением цена/качество изображения в сравнении с другими системами телескопов.

В конструкцию зрительной трубы входят следующие основные элементы:

- Труба (тубус);
- Главное зеркало – объектив;
- Диагональное зеркало;
- Окуляр;
- Узел диоптрийной подстройки окуляра;
- Монтировка + штатив (тренога).

Одной из главных задач при проектировании оптических приборов и систем является расчёт и анализ оптической системы прибора. Для данного прибора была разработана оптическая схема зрительной трубы Ньютона с ортоскопическим окуляром, с учётом следующих начальных условий (Таблица 1-3):

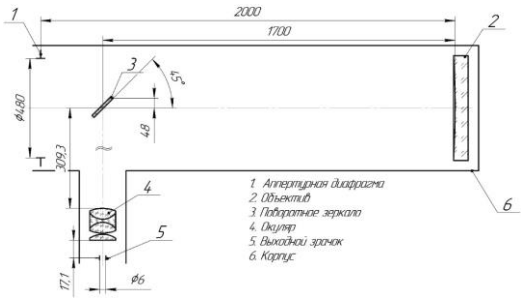


Рис. 1. Принципиальная оптическая схема зрительной трубы Ньютона

Таблица 1 – Исходные данные

Объектив	Зеркало - параболоид
Положение входного зрачка (апертурной диафрагмы)	В фокальной плоскости параболоида
Окуляр	Ортоскопический с задним фокусным расстоянием f'=25 мм
Поле зрения окуляра, 2·ω'	40°
Линейное перемещение окуляра для диоптрийной настройки	от -5 мм до +5 мм
Расположение объекта наблюдения	В бесконечности
Нормальное увеличение телескопа, $Г_н$	80

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Таблица 2 – Объектив, главное зеркало

Уравнение параболы	$y^2 = -8000z$
Фокусное расстояние f', мм	–2000
Диаметр зеркала D, мм	506
Световой диаметр , мм	498
Тип зеркального покрытия и способ нанесения	1И. 41И
Толщина по оси d, мм	63,3

Таблица 3 – Ортоскопический окуляр

№	r	d	n_e	Стекло
1	32,21			
		10,54	1,518291	К8
2	-15,996			
		3,0	1,629835	БФ12
3,	15,996			
		10,54	1,518291	К8
4	-32,21			
		1,0	1	
5	22,28			
		5,84	1,606258	ТК13
6	∞			

Расчёт оптической системы производился в ПО «ОПТИКА». Прикладное ПО позволяет моделировать, анализировать, производить расчет оптических систем. Интуитивно понятный интерфейс ПО «Оптика» разработан таким образом, чтобы максимально облегчить работу пользователя с данной программой.

Пакет «ОПТИКА» позволяет рассчитать:

- Гауссовы характеристики;
- Ход пучка лучей через систему;
- Кардинальные элементы оптической системы;
- Волновые аберрации и аберрации третьего порядка;
- Конструктивные параметр.

А также произвести:

- Оптимизацию системы;
- Асферизацию последней поверхности ОС.

Типы поверхностей:

- Поверхности второго порядка с осевой симметрией;
- Поверхности второго порядка с двумя плоскостями симметрии;
- Поверхности второго порядка с зональной ошибкой;
- Поверхности высшего порядка;
- Конические поверхности;
- Цилиндрические поверхности;
- Тороидальные поверхности;
- Плоские дифракционные решетки;
- Диафрагмы и экраны круглые, прямоугольные, полуплоскости.

Результаты расчёта оптической системы зрительной трубы Ньютона приведены в таблицах 4-5.

Расчет аберраций показал, что полевые аберрации объектива для поля зрения 2ω = 0,5º очень малы, так например, астигматическая разность составляет порядка $2 \cdot 10^{-8}$. Результаты расчета полевых аберраций объектива приведены в таблице 6. Аберрации окуляра представлены на рисунках 2-4. Для окуляра наибольшее влияние на качество изображения оказывают сферические аберрации. Максимальное значение продольной и поперечиной сферической аберрации равны соответственно: $\delta_{s'} = -0{,}4$, $\delta_{h'} = 0{,}04$.

Таблица 4 – Рассчитанные гауссовы характеристики

S_F, мм	$S'_{F'}$, мм	$\emptyset_{св}$, мм	D, мм	$\delta s'_{хр.}$, мм
9,3	17,1	22,8	25	0,06

Таблица 5 – Оптические характеристики системы

Характеристика	значение
Нормальное увеличение	80
Угловое поле зрения 2 ω, угл. град.	0,5
Диаметр выходного зрачка, мм	6
Угловое разрешение, угл. сек.	0,9
Удаление выходного зрачка от последней поверхности, мм	17,1

Аберрации объектива для поля зрения 2ω = 0,5º

Таблица 6 – полевые аберрации объектива

Параметр	значение
Меридиональный астигматизм, $\delta_{s'm}$	$-0{,}02$
Сагиттальный астигматизм,	$-0{,}021$
Средний астигматизм, $\delta_{s'}$	0,021
Астигматическая разность, Δ	$2 \cdot 10^{-8}$
Дисторсия V, %	$-5 \cdot 10^{-4}$

Аберрации окуляра для поля зрения 2ω = 40º

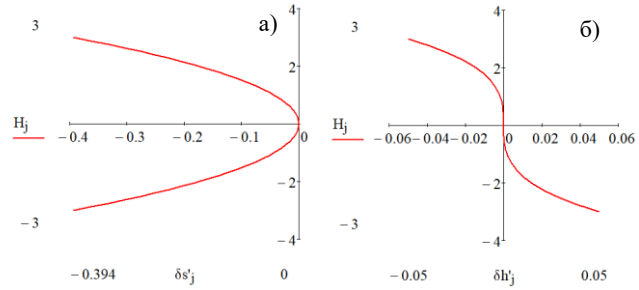


Рис. 2. Сферическая аберрация: а) продольная, б) поперечная

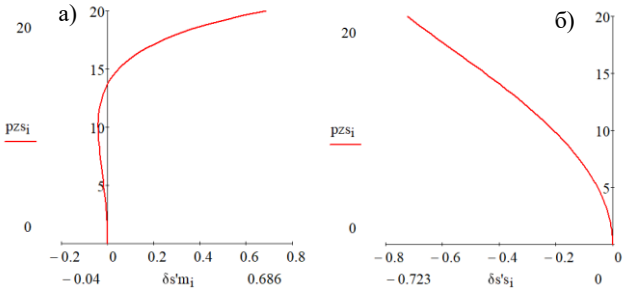


Рис. 3. Астигматизм: а) меридиональный, б) сагиттальный

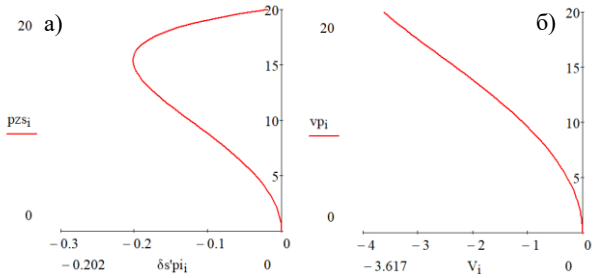


Рис. 4. Полевые аберрации: а) средний астигматизм, б) дисторсия

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

3. Разработка несущей механической системы зрительной трубы Ньютона

Фокусное расстояние главного зеркала $f' = -2000$ мм, диаметр главного зеркала $D = 506$ мм. Таким образом, телескоп имеет сравнительно ощутимые параметры массы и габаритов. Следовательно, наиболее целесообразно для данного телескопа использовать монтировку Добсона, что обеспечит простоту и низкую стоимость несущей конструкции (рис. 5).

Вращение вокруг вертикальной оси осуществляется поворотом всего корпуса монтировки за рукоятку, расположенную на трубе телескопа. Скорость вращения вокруг вертикальной оси регулируется путем изменения величины момента затяжки болта (гайки).

Для регулировки угла наклона используются 4 подшипниковые опоры, на которые опирается тубус телескопа. Регулировка угла наклона по горизонту осуществляется за счет вращения тубуса телескопа вокруг оси подшипников качения. Фиксация телескопа в заданном положении осуществляется с помощью рукояток.

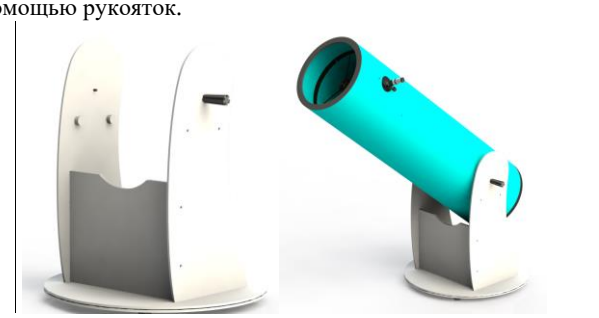


Рис. 5. Монтировка зрительной трубы Ньютона

Объектив зрительной трубы Ньютона представлен главным зеркалом параболической формы. Для крепления зеркал, участвующих в построении изображения и находящихся в основном пучке системы обязательными, являются следующие требования:

- статической определенности элементов, участвующих в построении изображения;
- компенсации воздействия температуры;
- возможностью регулировки замыкающих усилий.

В работе для обеспечения статической определённости конструкции (опора «на три точки») использовано крепление с помощью Z-образных лапок (Рис. 6). Так как планки жесткие то эластичность крепления обеспечивается специальными прокладками, устанавливаемыми между планками и зеркалом. Усилие прижима необходимой величины обеспечивается за счет подбора толщины прокладок.

Рис. 6. Узел объектива

Ортоскопический окуляр состоит из трех склеенных линз и одной удаленной линзы [8] (Рис. 7). Для данного окуляра применено крепление резьбовым кольцом. Данное крепление является разъёмным. Таким образом, данный тип крепления позволяет заменять части окуляра в случае их поломки или растрескивания. Однако, в отличие от крепления завальцовкой оно является жестким, что несколько ограничивает допустимый температурный режим работы окуляра [2, 4]. Посадочное отверстие 1,25″ (31,75 мм).

Постоянство взаимного положения всех оптических элементов обеспечивается тубусом. В нижнем конце трубы крепится главное зеркало (объектив). Верхний конец трубы несет диагональное зеркало и узел окуляра.

Рис. 7. Узел окуляра.

Тубус зрительной трубы Ньютона выполнен в виде сплошной трубы (Рис. 8) благодаря чему обеспечивается защита оптических элементов от пыли, мелких частиц и т.д. Однако зачастую для объективов большого диаметра в качестве меры снижения общей массы трубы возможно применение конструкции каркасного типа.

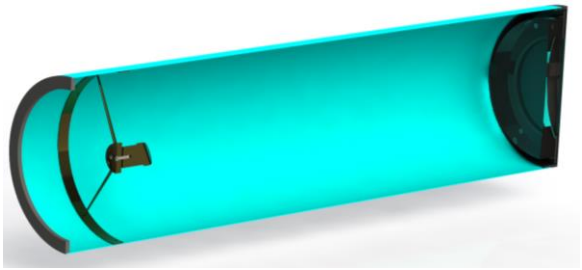


Рис. 8. Труба зрительной трубы Ньютона

Анализ массо-габаритных характеристик показал, что общий вес трубы составляет ≈ 80 кг. Большая масса телескопа в первую очередь обусловлена массой узла объектива (≈ 56 кг).

Таким образом, данный тип зрительной трубы целесообразно использовать в стационарных условиях. При этом наиболее удобно осуществлять транспортировку телескопа в разобранном виде, что возможно накладывает ряд дополнительных операций, связанных с последующей настройкой и сборкой зрительной трубы.

4. Заключение

Данная работа посвящена расчету и проектированию зрительной трубы Ньютона с ортоскопическим окуляром. В ходе работы продемонстрированы некоторые возможности прикладного ПО «ОПТИКА», а именно расчет и проектирование последовательных компонентов с возможностью изменения направления хода луча. При разработке несущей механической системы использовались только типовые соединения деталей, что обеспечивает простоту конструкции, а также возможности по автоматизации процесса производства. В качестве CAD системы при разработке несущей механической системы использовалось ПО Solidworks.

Достоинства зеркальных телескопических систем:

- в силу сравнительно большой апертуры превосходно работают для наблюдений тусклых объектов далекого космоса – галактик, туманностей, звездных скоплений;
- дают яркие изображения с малыми искажениями, отсутствует хроматическая аберрация.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Однако такие системы имеют и недостатки:

- центральное экранирование и растяжки вторичного зеркала снижают контраст деталей изображения;
- требуется периодическая подстройка положений зеркал (юстировка или коллимация), склонная утрачиваться при транспортировке и эксплуатации.

5. Литература

[1] Запрягаева Л.А., Свешникова И.С. Расчет и проектирование оптических систем. М. Логос, 2000, с. 581.
[2] Латыев, С. М. Конструирование точных (оптических) приборов: Москва: «Лань», 2015.
[3] Мальцев М.Д. Расчет допусков на оптические детали. М., Машиностроение, 1974, с. 168.
[4] Панов В. А. Справочник конструктора оптико-механических приборов, Под ред. В. А. Панова. — 3-е изд., перераб. и доп. — Ленинград: «Машиностроение», 1980. — 742 с.: ил. — Библиогр.: с. 728-732.
[5] Русинов М. М. Вычислительная оптика. Справочник под общ. ред. М. М. Русинова. Изд. 3-е. М., Книжный дом «ЛИБРОКОМ», 2009.
[6] Сокольский М.Н. Допуски и качество оптического изображения. Л. «Машиностроение», 1989, с. 221.
[7] Теребиж В.Ю. Современные оптические телескопы. М. Физматлит, 2005, с. 80.
[8] Турыгин, И. А. Прикладная оптика. Геометрическая оптика и методы расчета оптических схем: учебное пособие, Москва: «Машиностроение», 1965, с. 362.

Об авторах

Калинин Ростислав Георгиевич, студент, отделения материаловедения Томского политехнического университета. E-mail kalininrostislav@gmail.com.

Валиев Дамир Талгатович, к.ф.-м.н., доцент отделения материаловедения Томского политехнического университета. E-mail rubinfc@tpu.ru.

Агапов Николай Афанасьевич, д. т. н, профессор отделения материаловедения Томского политехнического университета. E-mail anikolja@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Создание панорамного видео

И.А. Канаева[1], Ю.А. Болотова[1]
iap15@tpu.ru|jbolotova@tpu.ru
[1]Томский политехнический университет, г. Томск, Россия;

Автоматическое создание бесшовного панорамного видео из набора перекрывающихся видов сцены, полученных с различных камер, является комплексной задачей. В данной работе представлен общий алгоритм создания панорамного видео, рассмотрены основные методы и подходы этапов вычисления параметров коррекции искажений, определения матриц гомографии, выбора цветокоррекции, смешения изображений в панораме. В работе предлагается применять проецирование изображений в цилиндрическую или сферическую систему координат, вычислять матрицу гомографии по соответствующим особым точкам смежных изображений, использовать метод цветопереноса для коррекции цветности и яркости изображений и вычислять градиентное смешение вдоль найденной оптимальной границы сшивки изображений.

Ключевые слова: *сшивка панорамных изображений, обработка многокадровых изображений, поиск особых точек, гомография, смешение цветов.*

Panoramic video creation

I.A. Kanaeva[1], Yu.A. Bolotova[1]
iap15@tpu.ru|jbolotova@tpu.ru
[1]Tomsk Polytechnic University, Tomsk, Russia;

The article is devoted to analysis of methods for stitching a seamless panoramic video from an overlapping image sequence. In following text general algorithm for creating a panoramic video is presented. More over main methods and approaches for calculating distortion correction parameters, determining homographic matrices, selecting color correction and blending images in panorama will be discussed. It is proposed to apply the projection of images into a cylindrical or spherical coordinate system, calculate the homography over overlapping areas, use the perceptual color transfer method to correct tones and brightness of the images and calculate the gradient blending along the optimum path.

Keywords: *panoramic image stitching, multiframe image processing, feature point detection, homography estimation, color blending.*

1. Введение

На сегодняшний момент панорамные изображения и видеоданные применяются в различных сферах деятельности благодаря наличию большого угла обзора и хорошего разрешения. Существуют два подхода к созданию панорамного видео: техническое решение, основанное на использовании панорамной видеокамеры и программное решение, основывающееся на обработке исходных видеоданных с перекрывающимися видами со стандартных видеокамер, совмещенных в единую систему. Панорамные видеокамеры с большим разрешением выходного кадра являются весьма дорогостоящими, нуждаются в обслуживании и калибровке.

Целью данной работы является обзор существующих и разработка новых программных методов создания панорамных изображений и видеозаписей из последовательности перекрывающихся снимков сцены. В начале статьи будет приведен общий алгоритм создания панорамного видео, а затем будут рассмотрены основные методы и походы каждого этапа отдельно.

2. Алгоритм создания панорамного видео

При съемке панорамного видео с системы камер необходимо, чтобы их положение и фокусное расстояние были фиксированы и не менялись в процессе записи, а также присутствовало перекрытие видов между соседними камерами. Итеративный алгоритм создания панорамного видео с нескольких видеокамер включает в себя три стадии: регистрация изображений, инициализация параметров и создание панорамы кадра. На рис. 1 приведен алгоритм создания панорамного видео с пяти горизонтально последовательно расположенных камер с широкоугольными объективами.

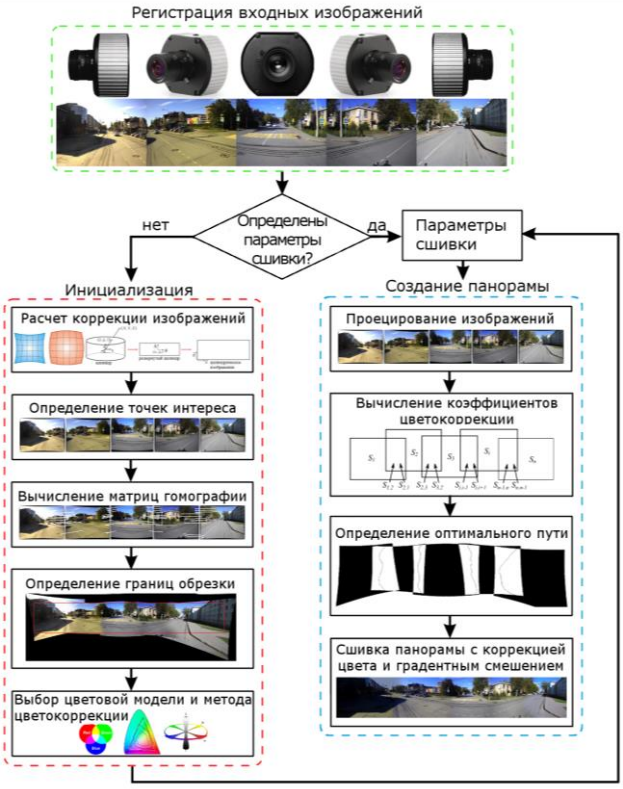


Рис. 1. Алгоритм создания панорамного видео.

В блоке инициализации по начальным кадрам видеопоследовательностей последовательно определяются коэффициенты исправления дисторсии, параметры проецирования изображения на цилиндр или сферу. Затем на всех спроецированных изображениях отмечаются точки

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

интереса. Между соседними изображениями вычисляются наборы коррелированных точек интереса, по которым с помощью метода RANSAC определяются матрицы гомографии изображений. При этом положение центрального изображения фиксируется, соседние левое и правое изображения преобразуются относительного центрального, а затем крайние левое и правое изображения преобразуются относительно полученных. Внутри полученной панорамы находится максимальный вписываемый прямоугольник, по которому будет производится обрезка панорамы. Затем выбирается цветовое пространство и задается метод цветокоррекции. После нахождения всех необходимых параметров в блоке создания исходные кадры вначале проецируются, затем внутри областей перекрытия соседних изображений вычисляются коэффициенты цветокоррекции. По скорректированным по цвету областям перекрытия с помощью динамического программирования определяется оптимальный путь сшивки и адаптивно выбирается ширина смешения. Работа по созданию панорамного кадра завершается проходом внутри границ обрезанного панорамного изображения и установки пикселей с учетом цветокоррекции и градиентного смешения.

Затем полученный панорамный кадр записывается в выходную видеопоследовательность, загружаются новые входные изображения и запускается блок создания с определенными на первом кадре параметрами.

3. Коррекция изображений

Существует несколько типов искажений, вносимых объективом камеры. К ним могут относиться затемнение углов кадра, искривление исходно прямых линий или цветная кайма вокруг контрастных границ. Несмотря на то, что они могут быть не очень заметны в исходном снимке, польза от их компенсации есть всегда.

Искривление исходно прямых линий на изображении, которое может повлиять на отображение объёма, описывается термином дисторсия. Хотя искажения могут быть беспорядочными и иметь различные формы, наиболее часто встречаются радиально симметричные искажения или очень близкие к таким, являющиеся следствием симметрии фотографической линзы. Эти искажения, как правило, классифицируются на бочкообразные и подушкообразные. Часто применяется следующая упрощенная модель исправления дисторсии [4]:

$$x' = x(1 + k_1 r^2 + k_2 r^4),\ y' = y(1 + k_1 r^2 + k_2 r^4),$$

где (x, y) – исходные координаты точки изображения, (x', y') – координаты точки с коррекцией дисторсии, $r^2 = x^2 + y^2$ – функция расстояния, а k_1 и k_2 называются параметрами радиального искажения. Более сложные модели искажений также включают тангенциальные (декантирующие) искажения.

Для коррекции геометрических искажений объектива типа «рыбий глаз» применяется следующий алгоритм [11]:

1. Читаем пиксель изображения $p = (x, y)$, нормализуем координаты $x = x - x_c, y = y - y_c$, где $c = (x_c, y_c)$ – координаты оптического центра.
2. Вычисляем $\rho = \sqrt{x^2 + y^2}$ – расстояние от центра до пикселя p и $z = a_0 + a_2\rho^2 + a_3\rho^3 + a_4\rho^4$ [6].
3. Вычисляем новые координаты пикселя $x' = \lambda \cdot f \cdot \frac{x}{z}, y' = \lambda \cdot f \cdot \frac{y}{z}$, где f – фокусное расстояние в пикселях, λ – масштабный коэффициент.
4. Преобразуем полученные координаты в координаты изображения $x' = x' - x_c, y' = y' - y_c$ и записываем пиксель $p' = (x', y')$ в выходное скорректированное изображение.

Параметры радиального искажения, коэффициенты полинома a_0, a_2, a_3, a_4, фокусное расстояние и оптический центр определяются в процессе калибровки камер, а также при коррекции дисторсии производится билинейная интерполяция.

Ввиду того, что оптические центры камер различны, исходные изображения содержат в себе ошибку параллакса, т. е. идеального совмещения как близких, так и дальних объектов не произойдет. Для сборки изображений в панораму прибегают к цилиндрической или сферической системе координат, т. к. прямоугольная проекция существенно преувеличивает перспективу по мере увеличения угла обзора и приводит к растягиванию стыкуемых изображений [10].

Цилиндрические проекции сохраняют более точные относительные размеры объектов, чем прямоугольные, однако достигается это за счёт искривления линий, параллельных линии зрения. Для проецирования изображения на цилиндр принимают, что оптическая ось соответствует оси z, а ось y вертикальна. Пиксель *(x, y)* соответствует точке *(X, Y, Z)* в трехмерном пространстве. При проецировании на цилиндрическую поверхность единичного радиуса точка на этой поверхности переводится в координаты угла и высоты цилиндра *(θ, h)*, т.е. *(sinθ, h, cosθ) = (x, y, f)*. Центр координат развернутого цилиндра лежит оптическом центре изображения (x_c, y_c), как показано на рис. 2.

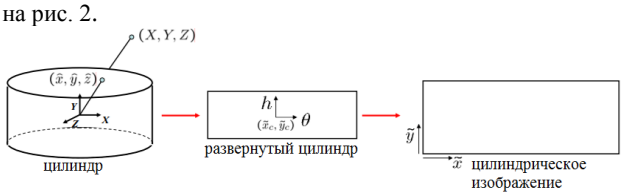


Рис. 2. Проецирование изображения на цилиндр.

Формула проецирования изображения на цилиндр, согласно [7], имеет вид:

$$x' = s\theta = s\tan^{-1}\frac{x}{f},\ y' = sh = \frac{sy}{\sqrt{x^2+f^2}},$$

где f – фокус, s – произвольный масштабный коэффициент (иногда называемый радиусом цилиндра), который может быть установлен равным f, чтобы минимизировать искажение (масштабирование) вблизи центра изображения.

Изображения можно также проецировать на сферическую поверхность, что полезно, если панорама включает полную сферу или полусферу изображений. В этом случае сфера задается двумя углами *(θ, φ)* с трехмерными сферическими координатами, заданными уравнением *(sinθcosφ, sinφ, cosθcosφ) = (x, y, f)*. Соответствие между координатами задается формулой:

$$x' = s\theta = s\tan^{-1}\frac{x}{f},\ y' = s\varphi = s\tan^{-1}\frac{y}{\sqrt{x^2+f^2}}.$$

4. Вычисление матрицы гомографии

Гомография – это перспективное преобразование плоскости, которое описывает, что случиться с воспринимаемыми положениями наблюдаемых объектов, когда изменяется точка зрения наблюдателя. Гомография представляется в виде матрицы 3×3 с восемью степенями свободы.

$$H = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix},\ x' = H\,x,\ x' = \frac{ax+by+c}{gx+hy+1}, y' = \frac{dx+ey+f}{gx+hy+1},$$

где H – матрица гомографии.

При отсутствии информации о внутренних и внешних параметрах камер используется вычисление матриц гомографии по коррелированным парам характерных точек изображений. Для этого на перекрывающихся частях соседних изображений выделяются характерные точки с помощью алгоритма Харриса или метода SURF.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Альтернативой является применение метода фазовой корреляции, но это требует большого процента перекрытия изображений порядка 90% [12].

Для создания матрицы гомографии из набора коррелированных в работе используется метод RANSAC (RANdom SAmple Consensus) – стабильный метод оценки параметров модели на основе случайных выборок [1].

Для создания горизонтальной панорамы стыковка происходит параллельно вправо и влево от центрального изображения согласно проекционным матрицам.

5. Цветокоррекция изображений

Последовательность исходных изображений для создания панорамы может сильно отличаться по яркости, тону и экспозиции, а также иметь такие искажающие эффекты, как виньетирование. Для уменьшения яркостных и цветовых отличий между смежными кадрами панорамы целесообразно применить методы предварительной цветокоррекции исходных изображений.

На рис. 3 представлена последовательность изображений S_1, S_2, …, S_n, где перекрывающиеся области соседних изображений, принадлежащие изображению S_i, обозначаются как $S^o_{i,i-1}$ – левая область, и $S^o_{i,i+1}$ – правая.

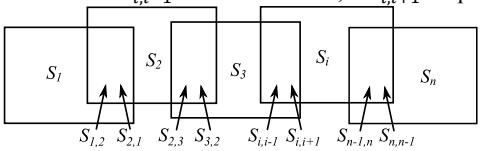


Рис. 3. Перекрывающиеся области в последовательности изображений.

Новым подходом к цветокоррекции панорам является применение перцептивного переноса цвета одного изображения S на другое I [5]. Идея метода заключается в вычислении статистик двух исходных изображений, таких как среднее значение μ_c и стандартное отклонение σ_c в каждом канале c цветовой модели, а затем изменение значения каждого пикселя исходного изображения I по следующей формуле:

$$I_c(x,y) = \frac{\sigma_c^S}{\sigma_c^I}(I_c(x,y) - \mu_c^I) + \mu_c^S.$$

В работе [2] показано, что для восстановления темных изображений путем цветопереноса лучшие результаты достигаются в цветовом пространстве RLAB [3].

Для уменьшения вычислений параметров цветокррекции панорам путем переноса цвета предлагается параллельно рассчитывать гистограммы каждой области перекрытия $S_{i,i-1}$ и $S_{i,i+1}$ (рис. 3), а затем по полученным гистограммам определять статистические параметры областей.

Если принять центральное изображение за идеальное, то изображения справа от центрального будут рассчитываться относительно статистик правой перекрывающейся области центрального кадра, а изображения слева от него – относительно статистик левой перекрывающейся области. Результат такого цвеопереноса в пространстве RLAB приведен на рис. 4.

Рис. 4. Результат переноса цвета от центрального кадра в RLAB.

Другим вариантом реализации является выбор области с наиболее однородной гистограммой и изменение остальных изображений относительно её. Для исходных кадров на рис. 1 наиболее однородной является левая область перекрытия пятого изображения (рис. 5).

Рис. 5. Результат переноса цвета от наиболее однородной области в RLAB.

При сравнении результатов можно отметить, что визуально плавность переходов цветов и яркости панорамы на рис. 4 лучше, однако крайне левая область слишком затемнена, чего не происходит на рис. 5.

Применение метода переноса цветовых статистик одного изображения на другие для улучшения создания бесшовной панорамы является новым подходом и имеет наилучшее соотношение качество/сложность.

6. Смешение изображений

При создании панорамного изображения необходимо устранять видимые границы стыковки соседних изображений, которые появляются не только из-за цветовых и яркостных различий изображений, но и из-за параллакса при съемке. На рис. 6*а* показана полученная из пяти исходных изображений панорама с явно видимыми границами стыковки, а на рис. 6*б* черным цветом внутри панорамы показаны области, в которых соседние изображения перекрываются.

а)

б)

Рис. 6. Панорама без смешения цветов и перекрывающиеся области в последовательности изображений.

Самым простым и распространенным вариантом смешения смежных изображений является градиентное смешение от центра области перекрытия изображений. Для этого от каждого перекрывающегося пикселя определяются расстояния d_1 и d_2 от его положения до границ области смешения δ (рис. 7).

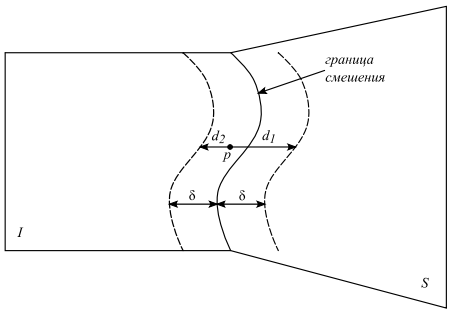


Рис. 7. Схематичное представление градиентного смешения.

На линии смешения значение цветового канала определяется как среднее между двумя соответствующими значениями двух пикселей перекрывающихся изображений, а при удалении от линии весовой коэффициент одного изображения увеличивается до 1, а другой уменьшается до 0 согласно следующей формуле [9]:

$$C_I'(p) = \frac{d_1^n C_I(p) + d_2^n C_s(p)}{d_1^n + d_2^n},$$

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

где $C_{I'}(p)$ – новое значение цвета пикселя p в панорамном изображении I', $C_I(p)$ – цвет пикселя p в исходном изображении I, $C_S(p)$ – цвет пикселя p в перекрывающем изображении S. В зависимости от степени n формы переходов цветов отличаются.

Для создания горизонтальной бесшовного панорамного видео лучше всего подходит метод градиентного линейного смешения 2-ой степени, благодаря его простоте и скорости работы, а также неплохим результатам для панорам из малых последовательностей изображений. Чтобы устранить эффект «призрака» при смешении по линейной границе, находят оптимальную границу сшивки двух перекрывающихся изображений, такую, для которой сумма квадратов ошибок яркостей вдоль границы будет минимальна.

Для этого в [8] с помощью динамического программирования находят путь минимальной стоимости (оптимальную границу смешения) на поверхности ошибок перекрытия двух смежных изображений – сканируют построчно область перекрытия снизу вверх и вычисляют поверхность кумулятивных минимальных ошибок. После оптимальный путь может быть получен путем прохождения сверху вниз по полученной поверхности и выборе минимального значения. На рис. 8*а* приведена панорама, сшитая с оптимальными границами, которые показаны черным цветом, а на рис. 8*б* показан результат градиентного смешения 2-ой степени в пределах данных зон.

а)

б)

Рис. 8. Панорама с оптимальными границами сшивки и результат градиентного смешения 2-ой степени

По рис. 8*б* видно, что несмотря на применение смешения цветов по оптимальным границам между изображениями, видны цветовые различия между частями панорамы. На рис. 4 и 5 показан результат бесшовной сшивки после применения цветокоррекции изображений.

7. Заключение

В рамках задачи создания панорамного видео был представлен алгоритм покадровой автоматической сшивки перекрывающихся изображений сцены в панорамное видео. На первой стадии инициализации параметров обработки изображений вычисляются коэффициенты дисторсии для стандартных объективов и типа «рыбий глаз», находятся параметры проецирования изображений в цилиндрическую или сферическую систему координат. Вычисление матриц гомографии основано на поиске соответствующих особых точек смежных изображений. Наилучшие результаты цветокоррекции достигаются путем оценки статистик областей перекрытия и цветопереноса в пространстве RLAB. Смешение изображений производится вдоль найденной с помощью динамического программирования оптимальной границы градиентным смешением яркостей 2-ой степени.

Благодаря оптимизации обработки изображений и параллельному исполнению основных алгоритмов среднее время создания одного панорамного кадра из пяти входных кадров размером 2048x1536 на двухъядерном процессоре Intel Core i5 с частотой ядер 3.2 ГГц составляет 3.4 с. Ведется работа по распараллеливанию задач обработки изображений на графическом процессоре для повышения быстродействия обработки. Предложенный алгоритм создания панорамного видео позволяет получать качественный выходной результат за приемлемое время.

Работа выполнена в рамках программы повышения конкурентоспособности ТПУ и при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А.

8. Литература

[1] Brown, M. Automatic panoramic image stitching using invariant features / M. Brown, D. G. Lowe / Int. J. Comput. Vision. – 2007. – Vol. 74(1). P. 59-73. – ISSN 0920-5691.

[2] Cepeda-Negrete, J. Dark image enhancement using perceptual color transfer / J. Cepeda-Negrete, R. Sanchez-Yanez, F. Correa-Tome, R. Lizarraga-Morales / IEEE Access. – 2018. – Vol.6. P. 14935-14945.

[3] Fairchild, Mark D. Color Appearance Models / Mark D. Fairchild // The Wiley-IS&T Series in Imaging Science and Technology - John Wiley & Sons, 2005. – 408 p.

[4] Gao D. Computing a complete camera lens distortion model by planar homography / D. Gao, F. Yin / Optics & Laser Technology, 2013, 49: 95–107

[5] Reinhard, E. Color transfer between images / E. Reinhard, B. Gooch, M. Adhikhmin, P. Shirley / IEEE Comput. Graph. Appl. – 2001. – Vol. 21(5). P. 34-41.

[6] Scaramuzza, D. A Toolbox for Easy Calibrating Omnidirectional Cameras / D. Scaramuzza, A. Martinelli, R. Siegwart / Proceedings to IEEE International Conference on Intelligent Robots and Systems, (IROS). Beijing, China, October 7–15, 2006.

[7] Szeliski, R. Computer Vision Algorithms and Applications. Texts in Computer Science / R. Szeliski. – London: Springer-Verlag London Limited, 2011. - 812 p.

[8] Xiong Y, Pulli K. Fast and high-quality image blending on mobile phones. 7th IEEE Consumer Communications and Networking Conference, Las Vegas, NV 2010: 1-5. DOI: 10.1109/CCNC.2010.5421736

[9] Xiong, Y. Fast panorama stitching for high-quality panoramic images on mobile phones / Y. Xiong, K. Pulli // IEEE Transactions on Consumer Electronics. – 2010. Vol. 56(2). P. 298-306. ISSN: 0098-3063.

[10] Коррекция искажений, вносимых объективом / [Электронный ресурс] / Cambridge in Colour, 2018. URL: http://www.cambridgeincolour.com/ru/tutorials-ru/lens-corrections.html, дата доступа: 05.05.18 г.

[11] Лазаренко В.П. Алгоритм преобразования изображений, полученных всенаправленными оптико-электронными системами / В.П. Лазаренко, Т.С. Джамийков., В.В. Коротаев, С.Н. Ярышев / Научно-технический вестник информационных технологий, механики и оптики. 2015. Том 15. № 1. С. 30–39.

[12] Лосев С. В. Устойчивость метода фазовой корреляции к возмущающим воздействиям при создании панорамных цифровых изображений / Изв. вузов. Приборостроение. 2015. Т. 58, № 11. С. 908–914.

Об авторах

Канаева Ирина Александровна, 1994 года рождения, обучается в магистратуре ТПУ по направлению «Информатика и вычислительная техника». Область научных интересов: обработка и анализ изображений и видео, рекуррентные нейронные сети. Email: iap15@tpu.ru.

Болотова Юлия Александровна, 1986 года рождения, в 2009 году окончила ТПУ по специальности «Программное обеспечение вычислительной техники и автоматизированных систем», к.т.н., доцент отделения информационных технологий. Область научных интересов: обработка изображений, распознавание образов, биологически-подобные алгоритмы. E-mail: jbolotova@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Применение метода художественного формообразования в проектировании интерьерных часов «К вершине»

Канкеева Н.С., Давыдова Е.М.
Томский политехнический университет, Томск, Россия
Natasha.kank@gmail.com | davydova@tpu.ru

В данной статье рассмотрено применение метода художественного-формообразования на основе проектирования интерьерных часов «К вершине». А также рассмотрен процесс создания концепции и создание итогового прототипа.

Ключевые слова: *конференция, компьютерная графика, метод художественного-формообразования, предметное проектирование.*

Application of the method of artistic form building in the design of the interior clock "To the top"

Kankeeva N.S., Davydova E.M.
Tomsk Polytechnic University, Tomsk, Russia
Natasha.kank@gmail.com | davydova@tpu.ru

This article is considered application of artistic form building on the example on the design of the "To the top" interior clock. In addition, the process of creating concepts and creating a final prototype is analyzed.

Key words: *conference, computer graphics, art forming method, objective design*

1. Введение

Процесс проектирования предмета является важным этапом в формировании внешнего облика и направлен на поиск новых решений. Он позволяет всецело раскрыть идею, которую дизайнер закладывает в концепцию.

Создание предметов дизайна возможно при условии понимания особенностей данной деятельности и владения определенной базой знаний [1]. Поэтому знание методов, которые позволяют осуществлять процесс проектирования, имеет большое значение.

Любой метод, который использует дизайнер, является сложной системой, совокупностью определенных действий, которые помогают достигнуть цели [2]. В современной дизайнерской деятельности можно встретить большое разнообразие методов. Поэтому какими пользоваться решает сам дизайнер, исходя из существующей ситуации.

Одним из основных и наиболее популярных методов проектирования – метод художественного формообразования. Так, на примере отдельно взятого предмета дизайна – часов, можно раскрыть процесс проектирования и показать применение данного метода.

Часы – это элемент декора с помощью, которого человек определяет время суток или измеряет временные интервалы для разных явлений.

Данный предмет является неотъемлемой частью жизни человека и повсеместно применяется во многих интерьерах. Часы позволяют быстро определить время, а также могут включать в себя дополнительный функционал (часы с полками, календарем, часы с подсветкой и т.п.).

Сегодня можно найти большое разнообразие данного продукта на рынке. Часы могут различать по материалам и габаритам, но главной проблемой остается – однообразие формы. Поэтому с помощью метода художественного-формообразования можно создать более привлекательный образ.

2. Метод художественного формообразования

Многие методы разрабатываются исходя из прагматичных и ориентированных на пользователя подходов. Развитие происходит в процессе проектирования, при их применении и анализе [5].

Существует множество методов, которые позволяют реализовать процесс проектирования, но наиболее полно его раскрывает метод художественного формообразования, который основан на индивидуальном творческом процессе и применении художественных принципов.

Метод художественного формообразования позволяет создать красивые и эстетически приятные образы для проектируемого изделия. А главной задачей промышленного дизайнера является умение сочетать полет творческой мысли и возможность воплощения в производстве.

Метод построен на основе таких понятий, как целостность образа, его гармонии и выразительности. Каждый критерий дополняет другой, тем самым создавая сложную, но уравновешенную систему.

Данный метод является решающим этап в процессе проектирования, он позволяет создать полный образ и воплотить его в качестве концепции с эскизами или другими графическими средствами [3].

Он включает в себя такие этапы как:

1. Постановка проблемы (сбор и изучение необходимой информации, выявление предполагаемых трудностей при производстве, анализ аналогов и т.д.)
2. Разработка концепции
3. Создание технического задания, включающее в себя перечень заданных требований
4. Разработка художественного образа изделия
5. Создание действующего прототипа.

3. Создание концепции

Создание проекта начинается с разработки основной идеи и определения проблемы, от которой идет дальнейшая работа.

Алгоритм создания концепции состоит из нескольких этап: описания идеи, определения ряда проблем, цели и задач проекта, критерии и описание их достижения, обозначение ключевых свойств, которыми будет обладать объект [4].

За основу образа, с которым велась последующая работа, были взяты горы. Горы – это не просто форма рельефа с ярко

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

выраженными склонами, подножиями и вершинами, горы – это прежде всего наслоение времени, которое спустя миллионы лет рождает великие и массивные скалы.

Как известно, в географии на различных чертежах, картах они обозначаются тонкими линиями, которые называются изолинии. Они используются для обозначения постепенно изменяющихся в пространстве величин. Поэтому данное географическое обозначение послужило хорошим образом для стилизации самих гор, а также они отображают то течение времени, благодаря которому они достигают своих высот.

Название «К вершине», выбранное для предмета, обусловлено связью времени и ландшафта: вершины достигаются временем. Каждый слой горы создавался тысячелетиями и каждая вершина – это великий труд стихии и времени.

4. Этапы моделирования и цветового обоснования

Первый этап – это визуализация объекта. Он заключается в создании 3D модели. Она создавалась в несколько шагов. Первый заключался в отрисовке формы в графическом редакторе, который позволяет создавать векторные изображения. Затем уже подготовленный объект импортировался в программу (импорт производился с помощь формата DWG) и далее уже создавалась сама модель с помощью различных функций. Созданный объект состоял из 7 пластин (Рис. 1) с градацией цвета от более светлого до более темного. Данный переход подчеркивает заданную идеи и позволяет разграничить контур. При визуализации накладывался материал имитирующий дерево.

Рис. 1. Визуализация часов

Габаритные размеры данного объекта: высота – 245 мм, ширина – 216 мм, глубина – 21 мм. Так же в ходе моделирования было выявлено обязательное условие – наличие отверстии для помещения самого часового механизма (квадрат высотой 65 мм и шириной 65 мм) и для оси крепления стрелок (окружность диаметром 8 мм).

Для любого объекта необходимо цветовое решение, которое позволяет значительно расширить область применения. В случае такого предмета как настенные часы областью применения будет являться интерьер.

На географических картах можно встретить различные варианты гор – обычные вершины, снежные, горы с лесополосой и множество других разновидностей.

В прототипе используется градация цвета – от наиболее светлого к более темному. Это нужно для того, чтобы поддержать образ, а также показать разность высот выделить каждую пластину.

Для изготовления оболочки часов была выбрана фанера. Она имеет ряд свойств, которые являлись определяющими при выборе материала:

1. Влагостойкость, за счет которой не подвержена деформации
2. Устойчива к воздействию высоких температур
3. Поверхность практически не восприимчива к загрязнениям

Верхний слой фанеры – шпон дерева, расцветку которого можно менять различными тонирующими средствами (морилки, пропитка). В дальнейшей работе данное свойство поможет сделать градацию цвета от верхней самой светлой пластины к нижней самой темной. Такой эффект получается в результате наложения большего количества слоев на последующие пластины.

Было выбрано два варианта для цветового решения – обычные горные вершины и снежные, что позволяет расширить линейку продукта.

Как было упомянуто ранее, в качестве материала для изготовления была выбрана фанера. Поэтому чтобы поддержать материал, для первого цветового решения была выбрана морилка с оттенком дуба.

Для поддержания образа снежных вершин была выбрана пропитка с белым, серыми и коричневыми оттенками. Так же данное решение позволит расширить список применимых интерьеров, потому что серо-белая гамма является нейтральной и часто используется.

Перечисленные цветовые решения представлены ниже на рис. 2 в качестве графических эскизов. Представлены шкалы градаций цвета, что достигается тонировкой каждой последующей пластины большим количеством слоев пропитки.

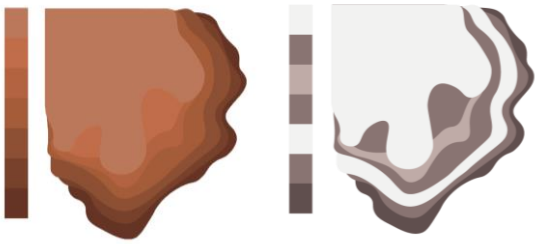


Рис. 2. Цветовое решение

Таким образом, на заключительном этапе проектирования мы получили действующий прототип интерьерных часов «К вершине».

5. Заключение

В процессе реализации проекта был изучен и применен метод художественного формообразования, выполнена визуализация проекта и изготовлен проект. Так же полученные знания были применены в процессе проектирования, который состоял из нескольких этапов – разработки концепции, поиска цветового решения, создания 3D модели, изготовления прототипа. На каждом этапе был изучен большой объем информации и применен на практике. В итоге проделанной работе был получена новая оболочка для часового механизма, выполненная из экологически чистого материала и вписывающийся во многие стили интерьера. Это позволяет расширить выбор продукта в данной сфере рынка для потребителя.

6. Литература:

[1] Jan Fredrik Schønheyder , Kjetil Nordby, 'The use and evolution of design methods in professional design practice', Institute of Design, The Oslo School of Architecture and Design, Oslo, Norway, 2018.

[2] Мартынов Ф. Т., Основные законы и принципы эстетического формообразования и их проявление в архитектуре и дизайне. Учебное пособие / Ф. Т. Мартынов – Екб.: Уральский архитектурно-художественный институт, 1992, 107 с.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

[3] Овчинникова, Р.Ю. Дизайн-проектирование: теоретические основания и специфика // Омский научный вестник. — 2012. —№1.— С. 267—270.
[4] Отличия концепции от устава проекта [электронный ресурс]: http://projectimo.ru/iniciaciya-proekta/koncepciya-proekta.html (дата обращения 03.05.2018).
[5] Прокопьева, И.А. Проблема выбора методов формообразования в дизайне // Архитектон: известия вузов. —2012.— № 38.

Об авторах

Давыдова Евгения Михайловна старший преподаватель кафедры инженерной графики и промышленного дизайна инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail: davydova@tpu.ru.

Канкеева Наталья Сергеевна студент направления промышленного дизайна инженерной школы информационных технологий и робототехники Томского политехнического университета.

E-mail: Natasha.kank@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Формализация описания класса изображений, полученных в различных условиях освещения

Касаткин Семён[1], Калайда Владимир[1]
antimaterialny@rambler.ru|kvt@ioa.ru
[1]Национальный исследовательский
Томский государственный университет, Томск, Россия

В работе рассматривается задача формализации описания класса изображений лица человека в различных условиях освещения. Исследования направлены на поиск методов, повышающих точность описания.

Ключевые слова: *Формализованное описание, класс изображений, непараметрическая статистика, коэффициент размытия, освещение.*

Formalization of the description of the class of images obtained under different lighting conditions

Kasatkin Semyon[1], Kalayda Vladimir[1]
antimaterialny@rambler.ru|kvt@ioa.ru
[1] National Research
Tomsk State University, Tomsk, Russia

The paper deals with the task of formalizing the description of a class of images of a person's face in various lighting conditions. Research aims to find methods that improve the accuracy of the description.

Keywords: *Formalized description, image class, nonparametric statistic, blur coefficient, lighting.*

1. Введение

Так как среда и используемые средства регистрации светового потока вносят искажения в сигнал, то изображение можно рассматривать как случайное поле. Под классом будем понимать множество реализаций одного и того же образа. Примером является изображения одного человека полученных в различных условиях: ракурс наблюдения, мимика, освещение и прочее.

Аппаратные искажения сигнала могут быть компенсированы за счёт использования в качестве формализованного описания вектором параметров (например, моментов) инвариантных к аффинным и проективным искажениям [3,6,10]. Однако наиболее существенные ошибки, обусловленные средой передачи сигнала и ее освещенностью, компенсировать сложнее. Для устранения яркостных искажений, при описании класса, удобно использовать плотность вероятности распределения параметров вектора, описывающего элемент класса [2,4,8].

Так как в прикладных системах видеонаблюдения и контроля проблематично получить «достаточное» количество наблюдений и сложно априори сделать предположение о виде функции плотности, то для её восстановления целесообразно использовать непараметрическую форму представления [1,2,4,5,6,8,9]. Данный подход не использует явно гипотезу о виде функции распределения. Ещё одним преимуществом непараметрического подхода является наличие в таких функциях плотности распределения свободных параметров (коэффициентов размытия), изменением которых можно регулировать меру приближения к сформулированным требованиям (повышение точности описания) [1,2,4,8].

Исходя из вышесказанного, для описания класса изображений целесообразно использовать методы непараметрической статистики. В качестве меры точности описания будем использовать суммарные ошибки первого и второго рода.

2. Непараметрический метод

Для формализованного описания класса плотностью вероятности, в байесовские функции подставим непараметрические оценки Розенблатта–Парзена:

$$\overline{p}(x)=\frac{1}{nc}\sum_{i=1}^{n}\Phi\left(\frac{x-x^{i}}{c}\right)$$

где Φ – ядро (чётная, нормированная функция), c – коэффициент размытости, n – количество наблюдений (количество изображений).

Из байесовской решающей функции следует, что исследуемый образ принадлежит тому классу, плотность вероятности которого больше.

Параметрами данной модели является вектор коэффициентов размытия для функций плотностей вероятностей. Тогда задача построения решающего правила сводится к оцениванию $n\times d$ параметров, где n – количество классов, d - размерность пространства признаков.

В современных работах, для упрощения задачи оптимизации принимается, что коэффициент размытия представляется в виде произведения некоторого общего коэффициента и оценок среднеквадратических отклонений. Таким образом задача оптимизации сводится к поиску оптимального общего коэффициента. Подробно изложение метода можно найти в книге [6].

Однако, использование общего коэффициента размытия может привести к пересечениям функций плотности вероятностей. Этого можно избежать (или по крайней мере значительно снизить значения суммарных ошибок первого и второго рода), если использовать альтернативные методы поиска параметров оптимизируемой функции [2].

Поскольку восстановленные плотности вероятности, в зависимости от выбранных ядерных функций (например, треугольной), могут иметь разрывы первого рода,

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

необходимо использовать безградиентные методы, не вычисляющие производные (методы нулевого порядка), например, методы: Хука-Дживса, Розенброка, Нелдера-Мида, Пауэлла.

Ни один из алгоритмов не оказался значительно лучше с точки зрения поиска минимума, выбор метода определяется конкретной задачей и опытом исследователя [9]. Метод Хука-Дживса в случае сильно вытянутых, изогнутых или обладающих острыми углами линий уровня целевой функции может оказаться неспособным обеспечить продвижение к точке минимума. Эффективность метода Росенброка зависит от выбора начальной точки поиска экстремума. Для метода Пауэлла так же, как и для предыдущих, количество проб целевой функции линейно растёт при увеличении размерности пространства, что негативно сказывается на эффективности при минимизации сложной целевой функции [9]. Метод Нелдера-Мида на каждой итерации алгоритма производит, как правило, одно-два вычисления значений функции, что чрезвычайно эффективно. В связи с этим, для оценки коэффициентов размытия целесообразно использовать алгоритм Нелдера-Мида.

3. Оценка коэффициента размытия

В качестве целевой функции используем суммарную ошибку первого и второго рода, для плотностей вероятностей, восстановленных непараметрическим методом.

Для обеспечения непрерывности функций в области определения, начальное значение коэффициента размытия установим равным половине максимального значения стандартной девиации для интервала:

$$c_d^0 = \max_i \left(x_{i,d}^j - x_{i-1,d}^j \right)$$

$$j = \overline{1,N},\ i = \overline{2,M},\ d = \overline{1,D}$$

где d – индекс коэффициента размытия, i – индекс изображения внутри класса, j – индекс класса, N – это количество классов, M – это количество изображений для класса, D – длина вектора коэффициентов размытия (размерность пространства признаков), σ_d^j – среднеквадратичное отклонение значений характеристик для класса с индексом j и коэффициентом размытия с индексом d.

Для обеспечения этого условия в целевую функцию вводиться штраф $e(\mathbf{c})$:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(p_1^{ij}(\mathbf{c}) + p_2^{ij}(\mathbf{c}) \right) + e(\mathbf{c}) \to \min$$

$$e(c) = \begin{cases} 1, & \exists \left(c_i - c_i^0 < 0 \right),\ где\ i = \overline{1,D} \\ 0, & иначе \end{cases}$$

где c^0 – стартовое, минимальное значение коэффициента размытия.

4. Результаты расчетов

В тестовой задаче при описании изображений в качестве вектора характеристик используются 9 инвариантных моментов:

$$\left\{ m_{1,1}, m_{2,0}, m_{2,2}, m_{3,3}, m_{4,0}, m_{4,2}, m_{4,4}, m_{6,0}, m_{6,4} \right\}$$

По данным векторам восстанавливалась функция плотности вероятности с использованием подхода [6] и методом Нелдера–Мида. Оценивались суммарные ошибки первого и второго рода для плотностей вероятностей (Таблица 1).

В качестве классов изображений были использованы:

1. CMU AMP face expression (13 классов изображений лиц людей по 75 полутоновых изображений разрешением 64x64 каждая. Общее количество изображений равно 975).
2. Georgia Tech face. (50 классов изображений лиц людей по 15 цветных изображений различного размера. Общее количество изображений равно 750).

Таблица 1. Сравнение суммарных ошибок

Метод оценки	Суммарная ошибка	
	CMU AMP	Georgia Tech
Классический	$1{,}04 \cdot 10^{-9}$	0,023
Нелдера–Мида	$4{,}49 \cdot 10^{-10}$	0,009

5. Заключение

На основании результатов расчетов с использованием баз изображений CMU AMP face expression и Georgia Tech face можно сделать предположение о том, что оптимизация индивидуальных коэффициентов размытия методами нулевого порядка вместо оптимизации общего коэффициента размытия позволяет уменьшить суммарные ошибки первого и второго рода на порядок.

6. Литература

[1] Агеева Н.М., Антонов В.Н., Калайда В.Т. и др. Наука и инновации в современном мире: техника и технологии. В 3 книгах. К 3.: монография // Одесса: Куприенко СВ, 2017 г., 157с.

[2] Калайда В.Т., Тарбаев Б.Б., Касаткин С.С. Распознавание полутонового изображения лица на основе формализации описания класса непараметрическими статистиками // Мир науки и инноваций., 2016. Т. 10. № 1(3), С. 44 – 56.

[3] Касаткин С.С., Калайда В.Т. Программная система для вычисления формализованных характеристик полутонового изображения, не зависимых от условий съемки. // Научная сессия ТУСУР-2014 Часть 4., 14–16 мая 2014 г., г. Томск, С. 47 – 49.

[4] Касаткин С.С., Калайда В.Т. Идентификация полутонового изображения лица человека. // Научная сессия ТУСУР-2015 Часть 2., 13–15 мая 2015 г., г. Томск, С. 205 – 208.

[5] Касаткин С.С., Калайда В.Т. Анализ методов распознавания изображений // Наука и инновации в современном мире: техника и технологии, 2017. Т. 3, С. 110 – 116.

[6] Лапко А.И., Лапко В.А. Непараметрические модели и алгоритмы обработки информации: учеб. пособие // Сиб. Гос. Аэрокосмич. Ун-т. – Красноярск, 2010 г., 220 с.

[7] Тарбаев Б.Б., Калайда В.Т. Разработка алгоритма программной системы формализованного описания полутоновых изображений, инвариантных к аффинным преобразованиям // Материалы Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР–2014» Часть 4., 2014 г. г. Томск, С. 52 – 45.

[8] Тарбаев Б.Б., Калайда В.Т. Вычисление формальных инвариантных признаков, описывающих изображение Б.Б. Тарбаев // Материалы Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР–2015. Часть 2. 2015 г., г. Томск, С. 244 – 247.

[9] Химельблаю Д. Прикладное нелинейное программирование // МИР – Москва, 1975 г., 536 с.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

[10] Abdul Kadir, Lukito Edi Nugroho, Adhi Susanto, P. Insap Santosa Experiments of zernike moments for leaf identification // Journal of Theoretical and Applied Information Technology., 15 July 2012., Vol. 41, No.1, P. 82-93.

[11] Kasatkin S.S. The methods of pattern recognition: A Review // АПР: Труды Международной молодежной научной школы. 25–26 сентября 2017 г., г. Томск, С. 128-131.

Об авторах

Касаткин Семён Сергеевич, аспирант кафедры оптико-электронных систем и дистанционного зондирования радиофизического факультета Томского государственного университета. Его e-mail antimaterialny@rambler.ru.

Калайда Владимир Тимофеевич, д.т.н., профессор кафедры оптико-электронных систем и дистанционного зондирования радиофизического факультета Томского государственного университета. Его e-mail kvt@ioa.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Основы использования цвета в рамках геймдизайна

А.Ю. Ларионова, Ю.С. Ризен
akalarik@gmail.com|yulja-vit@tpu.ru
Национальный исследовательский Томский политехнический университет (НИ ТПУ), Томск, Россия

Статья посвящена исследованию цвета, как инструмента, который использует дизайнер в проектировании видеоигр, а также исследованию особенностей психологии восприятия цвета для игровых элементов в контексте различных стилистик и жанров. Практическая значимость данной статьи заключатся в выявлении оптимальной последовательности действий для создания визуального стиля игры и формирования цветовых палитр.

Ключевые слова: *геймдизайн, цвет, цветовая палитра, цветовое решение, визуальная иерархия, компьютерная игра.*

Basics of using colour within the game design

A.Y. Larionova, Y.S. Rizen,
akalarik@gmail.com|yulja-vit@tpu.ru
Tomsk Polytechnic University (TPU), Tomsk, Russia;

The article investigates the colour, as a tool used by a designer in the design of video games, as well as to study the features of the psychology of colour perception for game elements in the context of various styles and genres. The practical significance of this article is to identify the optimal sequence of actions to create a visual style of the game and the formation of color palettes.

Keywords: *game design, colour, colour palette, colour scheme, visual hierarchy, video game.*

1. Введение

В современном игровом дизайне встречается огромное множество визуальных стилей, и с развитием технологий и появлением новых модных тенденций их становится всё больше. В этих условиях перед геймдизайнерами ставится задача определения целевой аудитории, способов её привлечения и удержания. Поэтому существенной частью подготовки игры становится создание её «правильной атмосферы», вызывающей эмоциональный отклик, который так необходим для популярности всего проекта. В данной статье рассматриваются основы использования одного из самых полезных инструментов игрового дизайна – цвета.

Актуальность данной темы обусловлена интересом к проектированию визуального стиля как начинающих геймдизайнеров, так и давно работающих в этой сфере людей. Зачастую перед художником стоит задача разработки авторского стиля для всей игры и выбор соответствующей цветовой палитры, усиливающей эмоциональную составляющую игры и упрощающей распознавание объектов. Таким образом художник сталкивается с проблемой выбора оптимального цветового решения. Цель статьи – выявить основные функции цвета в видеоиграх, определить какие задачи ставили перед собой дизайнеры, используя определённые цветовые палитры в разных игровых проектах, а также рассмотреть аспекты вопроса о восприятии цвета игроком с точки зрения психологии. Задачами исследования являются изучение и анализ опыта применения цвета в геймдизайне, определение особенностей психологии восприятия цветовых палитр существующих игр. Практическая значимость данной статьи заключатся в представлении художнику сведений, необходимых для принятия оптимального решения с выбором цвета в процессе проектирования, представление информации необходимой для создания наиболее выразительного художественного стиля игры, способного вызвать эмоциональную реакцию у широкой аудитории игроков.

2. Функции цвета в геймдизайне

Цвет в играх выполняет несколько ключевых функций визуальной составляющей проекта. Большинство таких функций перешли в гейм-индустрию из изобразительного искусства, дизайна, кино. Главной из них является то, что цвет служит для распознавания элементов игры. Благодаря цвету игрок может отличить искомый объект в сцене или же обратить внимание на жизненно важную шкалу в игровом интерфейсе, как показано на примере (рис. 1) [4].

Рис. 1. Экран игры Castle Crashers.

Другая немаловажная функция цвета – способность вызывать у зрителя эмоции. К примеру по-разному окрашенная игровая сцена может вызвать абсолютно противоположные эмоции у игрока. Посредством цвета происходят смена дня и ночи, смена температуры воздуха, различные физические явления или даже то, как воспринимается мир героем игры в текущий момент времени. Самым популярным приёмом для настройки цвета, активно использующимся художниками в наше время, является цветокоррекция (рис. 2). Такой приём позволяет художнику быстро манипулировать цветом [1, 4].

Рис. 2. Использование цветокоррекции.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Выбор общей цветовой палитры для всего проекта играет важную роль в узнаваемости бренда. Примерами могут послужить использование пастельных родственных оттенков в игре Journey, кислотного сочетания голубого и розового в Hotline Miami, фиолетовый цвет, как символика отряда Святых в Saints Row, красный голубой и белый, ставшие визитной карточкой Mirror's Edge, приведёны в пример на рис. 3, родственные оттенки красного, оранжевого и жёлтого, против бирюзового в Transistor.

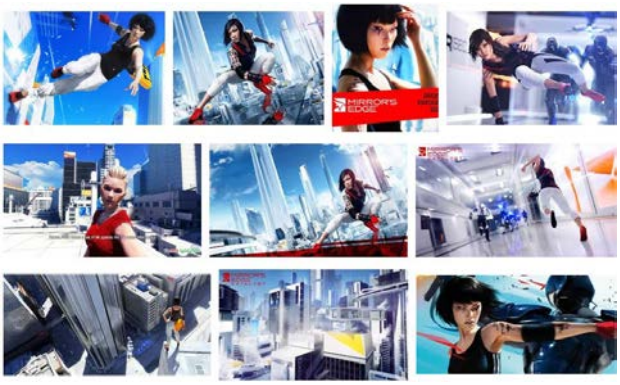


Рис. 3. Цвета, помогающие в узнаваемости бренда Mirror's Edge.

Определённые цветовые гаммы нацелены на разную аудиторию игроков, так более сочные, тёплые и насыщенные оттенки, характерны для казуальных игр, а утончённые цветовые сочетания нацелены на привлечения аудитории для более сложных игр [5]. На выбор цвета также существенно влияют модные тенденции, так для большинства поздних проектов присущи приглушенные тона, добавляющие играм большей реалистичности [1, 3].

Во всех играх присутствует особая визуальная иерархия: в первую очередь игрок должен видеть своего персонажа, во вторую своих врагов, далее объекты, с которыми можно взаимодействовать и фон. Цвет может помочь зрителю в определении этой иерархии, может сфокусировать внимание на главном и отвлечь от второстепенного, что очень важно для интерактивной деятельности. Элементы можно выделять и скрывать при помощи изменения цветового тона, светлоты и насыщенности [2]. В таких играх как Mirror's Edge и Portal, частью игрового процесса которых является решение логических задач, игрок полностью сосредоточен на выделенных цветом игровых элементах, а нейтральный белый фон почти не занимает его внимание. Примером неудачного использования цвета для визуальной иерархии может послужить, такой игровой проект как Cuphead: фон и игровые объекты почти одинаковы по яркости, многочисленные цвета отвлекают внимание игрока от приближающихся снарядов, что перегружает зрительное восприятие и значительно увеличивает сложность прохождения. Сравнение удачного и неудачного цветового решения для визуальной иерархии представлено на рис. 4.

Рис. 4. Сравнение цветового решения для визуальной иерархии игр Mirror's Edge и Cuphead.

В игровых интерфейсах цвет используется как идентификатор. Цвет подсказывает игроку о возможности взаимодействия с объектами, например, подсветка объектов для сбора в играх серии Ведьмак. Цветовое различие сообщает свойства объекта, помогает с их категоризацией, как показано на рис. 5, такой приём часто используется в панели инвентаря и книгах заклинаний многих РПГ-игр. Также цвет помогает отличить своих от чужих в командных играх и стратегиях.

Рис. 5. Отличающиеся по свойствам ингредиенты зелий выделены разным цветом.

Цвет также может быть и частью игровой механики. Существует множество игр построенных на решении логических задач, связанных с цветом объектов. Например, цвет может наделять объекты, или даже игровых персонажей, определёнными физическими силами и свойствами, помогающими игроку достичь цели, как в Thomas Was Alone, или такой вариант – игра-головоломка Bejeweled (рис. 6), где задача игрока заключается в составлении объектов в наборы определённого цвета. Во многом экспериментальный проект, квест с элементами выживания Тургор, использует цвет, как способ воздействия на окружающий мир, игрок «раскрашивает» игровое окружение, чтобы взаимодействовать с объектами игры и персонажами [1, 3].

Рис. 6. Цвет, как часть игровой механики в Bejeweled.

Вариативность цвета может создавать у игрока ощущение разнообразности контента. Особенно это важно для игр с абстрактной графикой. Чаще всего такие изменения затрагивают неактивные объекты окружения и фон. Их цвет меняется от уровня к уровню, чтобы игрок не

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

успевал заскучать. Такой приём был использован при разработке минималистичной игры Thomas Was Alone (рис. 6). Существуют специальные программные алгоритмы для генерации случайных цветов, не нарушающих визуальную иерархию игры [1].

Рис. 7. Цветовое различие уровней в Thomas Was Alone.

3. Цветовые палитры

Что касается выбора цветовой палитры для игры, то в этом аспекте геймдизайн во многом заимствует опыт изобразительного искусства, кино, фотографии и многих других сфер визуальной культуры. Цветовая гамма тёплых тонов будет использоваться для создания атмосферы уюта или же агрессии, это зависит от контекста самой игры (рис. 8, а). Холодные яркие тона используются в sci-fi играх для описания бездушного стерильного будущего (рис. 8, б). Для игр, основными темами для которых являются вестерн и антиутопия, основной гаммой служат пыльные и землистые оттенки, подчеркивающие атмосферу пустоты, запылённости и заброшенности (рис. 8, в).

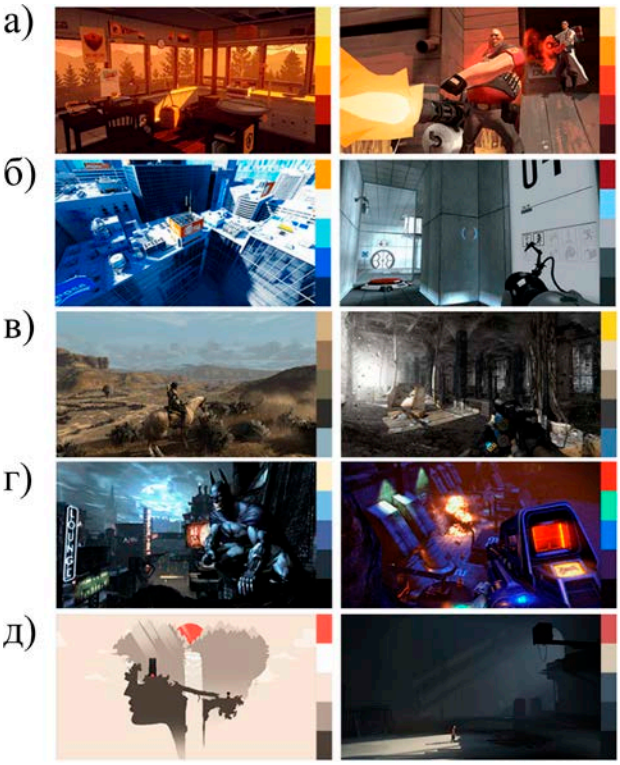


Рис. 8. Наборы цветовых палитр игр разных тематик.

Картинка в нуаре и киберпанке, как правило, темная, и в ней присутствуют яркие светлые участки, можно использовать спокойные оттенки цветов, монохром, или кислотные тона, в зависимости от времени и места событий игры (рис. 8, г). Монохромные и ограниченные палитры, упрощающие восприятие, используются для создания минималистических проектов, ключевую роль в которых играет композиция (рис. 8, д) [2].

4. Процесс работы с цветом

Вначале работы художник должен определиться с общей цветовой гаммой игры: определить характерные цвета для игровой тематики, её стилистические особенности согласно целевой аудитории, время и место происходящих событий, в зависимости от этого будут использоваться теплые или холодные оттенки, ахроматическая цветовая гамма.

Согласно выбранной цветовой схеме, заготавливаются необходимые палитры цветов. Для упрощения этой задачи многие художники пользуются референсами, либо используют такой приём, как наложение фильтра на цветовой круг.

Когда основные палитры были определены, можно переходить к следующему этапу работы, в рамках которого будет полезно создать фейк-скриншот (рис. 9) – это экран игры, с наиболее часто встречающейся ситуацией, создаваемый для того, чтобы проверить, может ли игра быть такой, какой художник её себе представляет. На такой собранной картинке разбрасываются примерные цвета фонового окружения, персонажей, интерфейса, текстур, намечаются предполагаемые освещённые и теневые участки.

Рис. 9. Фейк-скриншот.

Исходя из доминирующих оттенков выбранной палитры, выбирается основной цвет фонового окружения. Необходимо учесть, что такой цвет не должен концентрировать на себе внимание игрока, отвлекать или раздражать его. Количество цветов на фоне будет ограниченным, будет использована родственная гармония. Чем дальше будет располагаться объект, тем более светлым и менее насыщенным он будет, так создастся имитация воздушной перспективы, что даст игроку ощущение глубины картинки.

Важно учесть, что главный герой и другие персонажи должны выделяться на фоне окружения. Наилучшим вариантом будет, если цвета героев игры и фона будут в контрастной паре, то есть будут находиться в противоположных частях цветового круга. Для выделения персонажей обычно используются наиболее яркие и насыщенные оттенки. Как правило, художники стараются подчеркнуть верхнюю часть каждого персонажа, так как в этой области чаще всего располагаются самые важные элементы, позволяющие идентифицировать героя, для них стоит использовать самые насыщенные оттенки. При сильной детализации персонажа различие его форм подчёркивается тоновым контрастом.

Когда основные цвета для персонажей и фонового окружения были распределены, следует определить цвет для интерфейса. Интерфейс, как и персонажи, должен

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

выделяться на фоне, при этом он не должен излишне концентрировать на себе внимание игрока, так как интерфейс играет вспомогательную роль. В связи с психологией восприятия, обычно для создания интерфейсов используют тёплые или нейтральные цвета. Необходимо соблюдать контраст межу подложками и расположенными слоем выше текстом и иконками.

Далее художник переходит к более детальной составляющей сцены – к текстурам. Для них используются оттенки одного цвета схожие по тону, это поможет избежать чрезмерной яркости при освещении и резкого выделения деталей текстуры, во избежание цикличности.

Последним этапом для художника в определении цветов игровой сцены будет введение в неё света и теней. Для того, чтобы среда игры выглядела естественной, при создании света и теней используются хроматические оттенки. Яркость цветов не должная быть более 80%, такие области будут плохо различимы при освещении, повысится риск засвечивания. Негативное влияние на вид сцены при её освещении также могут оказать чисто-белый и чисто-чёрный цвета [6]. При расстановке света и теней важно помнить об их главных функциях в процессе игры: с их помощью можно направлять игрока и скрывать маловажные участки уровня игры.

5. Заключение

Поскольку цвет является неотъемлемой частью художественной стороны игрового дизайна, то практика его применения должна отталкиваться от накопленного опыта и от понимания функций, которые несёт цвет в этой сфере. Цветовое решение влияет на психологическое восприятие игры, и при его грамотном использовании контент игры будет казаться многообразней, игровые сцены будут эмоционально выразительными, а внимание игрока будет сосредоточено на ключевых моментах геймплея. В свою очередь, своеобразие цветовой палитры способно создать особую атмосферу видеоигры, сделать её отличной от других игровых проектов.

При создании визуального стиля опытные художники придерживаются определённого ряда действий при работе с цветом. Каждый из них имеет свой индивидуальный подход. В результате исследования были проведены анализ и обобщение их алгоритмов работы, а также сформирована оптимальная последовательность действий при работе с цветом:

1. Определение общей цветовой гаммы игры.
2. Заготовка необходимых для дальнейшей работы цветовых палитр.
3. Создание фейк-скриншота.
4. Определение цвета:
   а. фона;
   б. персонажей;
   в. интерфейса;
   г. текстур;
   д. света и теней.

Опираясь на выявленную последовательность и существующий опыт работы с цветом в других областях (кино, изобразительное искусство и т.д.) были адаптированы основные законы колористики для игровой индустрии. Важно заметить, что работа с цветом в перспективе развития геймдизайна и технологий становится мощным инструментом, влияющим как на восприятие целевой аудитории, так и на погружение в игровой процесс.

Работа с цветом позволяет художнику, сохраняя концепцию игры, подчеркнуть её жанровые и стилистические особенности, а также грамотно выстроить визуальную иерархию.

6. Литература

[1] Color in games: An in-depth look at one of game design's most useful tools: [Электронный ресурс] // Gamasutra - The Art & Business of Making Games
URL: https://www.gamasutra.com/blogs/HermanTulleken/20150729/249761/Color_in_games_An_indepth_look_at_one_of_game_designs_most_useful_tools.php. (Дата обращения: 21.02.2018).

[2] Color in Video Games: How to Choose a Palette: [Электронный ресурс] // Gamasutra - The Art & Business of Making Games
URL: https://www.gamasutra.com/blogs/DougStewart/20170411/295806/Color_in_Video_Games_How_to_Choose_a_Palette.php. (Дата обращения: 21.02.2018).

[3] Parkin S. An illustrated history of 151 video games. – London: Anness Publishing Limited, UK, 2012. – 255 p.

[4] Schell J. The art of game design. A book of lenses. – Amsterdam; Boston: Morgan Kaufmann Publishers, 2008. – 489 p.

[5] Казакова Н. Ю., Назаров Ю. В. Целевая аудитория гейм-дизайна и игровой процесс // Вестн. МГХПА. – 2015. – № 1. – С. 393–414.

[6] Фиалка К. Подходы и методы в разработке визуального стиля игрового проекта: [Электронный ресурс] // App2Top.ru
URL: http://app2top.ru/game_development/podhody-i-metody-v-razrabotke-vizual-nogo-stilya-igrovogo-proekta-63222.html. (Дата обращения: 12.05.2018).

Об авторах

Ларионова Анна Юрьевна, студент инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail: akalarik@gmail.com.

Ризен Юлия Сергеевна, старший преподаватель инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail: yulja-vit@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Методы дизайн-проектирования портативного медицинского устройства

А.М. Минчук[1], Е.М. Давыдова[2]
Doksa_2008@mail.ru|davydova@tpu.ru
[1]ТПУ, Томск, Россия;
[2]ТПУ, Томск, Россия

В данной статье рассмотрен процесс проектирования портативного медицинского устройства, а точнее поиск наилучшего формообразования. Кроме того, рассмотрены основные методы дизайн-проектирования. Также продемонстрированно применение некоторых из методов на практике с описанием полученных результатов. Сделан вывод по определению оптимального варианта, максимально соответствующего требуемым критериям.

Ключевые слова: *конференция, компьютерная графика, методы дизайн-проектирования, проектирование.*

Methods for designing a portable medical device

A.M. Minchuk[1], E.M. Davydova[2]
Doksa_2008@mail.ru|davydova@tpu.ru
[1]TPU, Tomsk, Russia;
[2]TPU, Tomsk, Russia

This article discusses the process of designing a portable medical device, more precisely, the search for the best form-building. In addition, the main methods of design are considered. It has also been demonstrated the application of some of the methods in practice with a description of the results obtained. A conclusion is drawn by the definition of the optimal variant, which maximally meets the required criteria.

Key words: *conference, computer graphics, design-design methods, design.*

1. Введение

Беременность и рождение малыша – это особенное состояние женщины. В этот период весь ее организм перестраивается, готовясь к самому главному событию - рождению нового человека. Именно поэтому к планированию беременности стоит подходить ответственно. На сегодняшний день существует множество различных медицинских устройств, помогающих женщине при планировании беременности, а также в процессе самой беременности и после рождения малыша. Каждое из рассматриваемых устройств имеет свой корпус, это условие создает необходимость приобретать каждое устройство отдельно. Гипотеза разработчиков заключается в более удобном использовании всех рассматриваемых приборов, при совмещении всех технических характеристик в одном.

Проектируемое портативное медицинское устройство предполагает использование его женщинами в разный период времени: до беременности (монитор овуляций), во время беременности (фетальный доплер), после беременности (монитор для новорожденных).

Исходные данные предмета проектирования. Монитор овуляций позволяет отслеживать фертильные дни цикла, благоприятные для зачатия. Фетальный монитор - устройство, позволяющее одновременно и надежно обеспечивать регистрацию частоты сердцебиения плода (ЧССП), частоту сердцебиения матери (ЧССМ), дыхание матери, сократительную деятельность матки и двигательной активности плода с 20 недели беременности. Монитор для новорожденных – устройство, отслеживающее сердечный ритм младенца, дыхание младенца, температуру его тела [1].

Особенность проектируемого устройства заключается в его многофункциональности, длительном использовании на разных участках тела женщины и ребенка. В формообразовании учитываются не только основная задача объекта, но и связь с человеком. Этот фактор также зависит от функции устройства, но главное в нем - обеспечение через соответствующую форму удобства и безопасности пользования, т.е. учет эргономических требований к предмету.

Большую часть времени устройство носится под одеждой, поэтому форма должна быть такой, чтобы не доставлять дискомфорт при взаимодействии с телом человека. Также при работе над формой необходимо учитывать ее взаимодействие с одеждой пользователя, одежда не должна цепляться за устройство, и одежда не должна срывать устройство с места измерения данных.

Еще одним важным критерием является экономическая целесообразность, одним из необходимых условий, выхода проектируемого устройства на рынок является его конкурентоспособная стоимость.

Для формирования выводов, способствующих определению критериев разработки формы были изучены работы А. И. Захарова доктора психологических наук, профессора кафедры психологической, и Н.П. Коваленко доктора психологических наук, занимающейся исследованиями перинатальной психологии на кафедре медицинской психологии и психофизиологии в Санкт- Петербургском государственном университете.

Основываясь на работах данных авторов можно сделать вывод, что эмоции, которые испытывает беременная женщина, напрямую влияют на протекание беременности и родов, на отношение к ребенку, на отношение к себе самой. Благоприятные качества, формирующие положительные эмоции во время беременности - это отсутствие страха, адекватная самооценка, уравновешенность. С древних времен беременным женщина рекомендовалось созерцание красивых вещей, природы, картин [2]. Поэтому проектируемое устройство должно обеспечивать

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

правильное психоэмоциональное состояние женщины, вызывая у нее положительные эмоции и укрепляя ее самооценку в положительную сторону.

2. Методы проектирования в дизайне

При проектировании объекта необходимо последовательно продумывать и планировать действия для достижения поставленных целей. При этом проводить поиск новых идей, это обуславливает использование методов дизайн-проектирования. При работе над данным устройством было принято решение использовать несколько методов проектирования для нахождения оптимального варианта формообразования.

Эвристика (от греч. Heurisko – отыскиваю, открываю) – это наука, изучающая продуктивное творческое мышление [3]. Применение на практике эвристических методов проектирования помогает раскрыть творческий потенциал, развить логическое мышление в профессиональном направлении, а также дает возможность контролировать и интенсифицировать процесс творческого поиска.

Основные эвристические методы проектирования:

Метод ассоциаций – метод формирования идеи, построенный на образно-ассоциативном мышлении. Данный метод проектирования эффективен, если есть достаточный опыт взаимодействия с окружающей действительностью. Развитие данного образа мышления позволяет быстро реагировать на окружающий мир, черпать оттуда вдохновение и ассоциации. Данный метод заключается в преобразовании ассоциаций в реальную форму, при этом ассоциации могут быть как предметные, абстрактные, так и психологические, ирреальные [4].

Метод аналогий – данный метод базируется на заимствовании решений из народного промысла, национальной культуры, инженерных и конструкторских решений и т.д. Изначально рассматриваемый метод относился к функциональному проектированию, например, проектировать не чайник, а способ кипячения воды [5].

Бионический метод- суть данного метода заключается в анализе объектов живой природы. Метод позволяет находить неординарные решения конструкторских узлов, поверхностей, фактур и форм.

Метод декомпозиции и принцип последовательного приближения – смысл данного метода разложение одной сложной задачи на составляющие, которые необходимо решить последовательно для получения результата.

Метод наводящей задачи - решение исходной задачи и определение тех показателей, которые мешают решению данной задачи. Данный метод позволяет решать поставленные задачи, используя опыт проектирования других дизайнеров. Суть метода заключается в поиске чужих идей, имеющих схожую задачу проектирования, и анализе их достоинств и недостатков [6].

Метод эмпатии – суть данного метода заключается во «вхождении в роль» проектируемого изделия. На первый взгляд метод достаточно абсурдный, но с его помощью можно получить необычное правильное решение.

Метод агрегативности – данный метод используется для создания изделий, базирующийся на геометрической и функциональной взаимозаменяемости отдельных элементов и узлов.

Метод «мозговой атаки» - метод, при котором определенный коллектив людей генерирует огромное количество идей в сжатые временные сроки. Такой подход к проектированию позволяет предположить, что среди множества различных идей может оказаться несколько отличных решений.

Метод художественного формообразования – данный метод отличается тем, что основывается на индивидуальном творческом процессе, а также художественных принципах проектирования.

3. Этап эскизирования и выбор варианта дизайн-решения

На данном этапе осуществляется разработка начальных представлений о проектируемом объекте: формируется представление о пропорциях и размерах проектируемых элементов, формируется внешний образ объекта.

В данном дизайн-проекте исходя из назначения проектируемого объекта, из обзора методов проектирования были разработаны варианты формообразования.

Первый вариант разработан с использованием метода агрегатирования (Рисунок 1). Предлагаемая форма лишена художественно-образного решения и основывается только на геометрической и функциональной взаимосвязи технических элементов конструкции. Полученную форму можно назвать рациональной, логически обоснованной. В ней прослеживается прямая, самая тесная связь формы с ее функциональным содержанием. Такое формообразование демонстрирует глубокое решение сугубо утилитарных задач. Данное решение формообразования делает устройство не только удобным, но и комфортным в своих функциональных качествах. Кроме того, можно подчеркнуть важный экономический аспект данного метода формообразования. Здесь следует говорить о целесообразной экономии материальных и финансовых средств, используемых в создании дизайнерской формы. Однако в данном формообразовании рациональная составляющая превалирует, поэтому форма выглядит «сухой», неинтересной и маловыразительной. Что касается эргономики, то данная форма имеет небольшие размеры, что позволяет получать измерения с поверхности любой кривизны, но она имеет приподнятые края, что увеличивает возможность срывания устройства с тела одеждой. Особенно, при ношении устройства на животе в последнем триместре беременности.

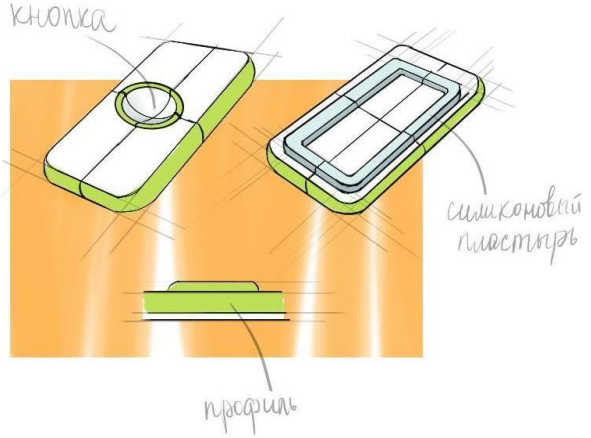


Рис. 1. Метод агрегатирования

Второй вариант разработан при помощи метода ассоциаций (Рисунок 2). Ассоциативный ряд выстраивался следующим образом – беременность, единство, часть, жизнь, вода, капля. При создании данной формы особое внимание уделялось объективным и субъективным составляющим характеру художественной формы. Объективное содержание образа есть некое идеальное представление о форме, как объекте действительности. Оно носит абсолютный характер, сложившийся на основе отношения к этому объекту многих людей. Такое содержание не зависит от мнения отдельного человека.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Субъективное содержание выражает мысль автора по поводу разрабатываемой формы. Оно носит относительных характер, который выражает мнение каждого воспринимающего форму зрителя. В этом смысле образ субъективен. Устранение противоречия между объективным и субъективным содержанием образа и есть искомый момент гармонизации [7,8]. При таком устроении образ приобретает правдиво-выразительный характер.

Эргономика данной формы позволяет производить замеры с поверхностей различной кривизны. С одной стороны, форма имеет наклонную поверхность, что существенно снижает возможность срывания устройства с тела одеждой. Но из-за сложной конфигурации формы и вынесенной на другой уровень кнопки, значительно увеличивается экономический аспект.

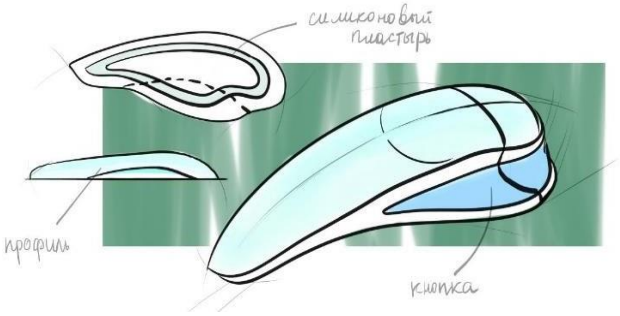


Рис. 2. Метод ассоциаций

Третий вариант разработан с использованием бионического метода (Рисунок 3). Основой для создания данной формы послужил анализ формы листа дерева. Рассматриваемый принцип формообразования позволяет четко и глубоко продемонстрировать раскрытие в композиции поставленной художественной идеи. Выраженный в форме образ наполняет ее глубоким духовным содержанием, делает впечатляющей. Образная форма оказывает на пользователя более сильное и глубокое эмоционально-эстетическое воздействие, чем простая утилитарная форма[9].

Эргономика данного формообразования позволяет производить замеры с поверхностей различной кривизны. Обтекаемая с обеих сторон форма, с опущенными вниз краями обеспечивает плотное прилегание к телу, сводя к минимуму возможность срывания устройства с тела одеждой. К тому же, силиконовая поверхность, имеющая прорези, обеспечивает надежный контакт устройства с местом снятия необходимых показателей. Экономический аспект данного формообразования незначительно увеличивается из-за криволинейных поверхностей.

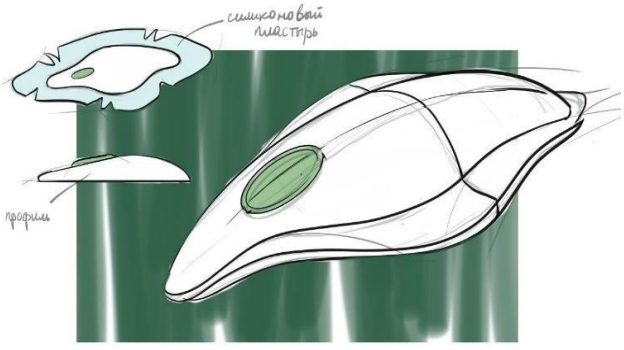


Рис. 3. Бионический метод

4. Сравнительный анализ проектируемых форм

Для наглядности была составлена таблица, где каждому из вариантов присвоен балл от 0 до 1 по каждому из критериев (Рисунок 4).

По результату анализа можно сделать вывод, что наиболее подходящей формой, более соответствующей заявленным критериям для проектируемого медицинского устройства, является третий вариант.

Критерии	1 вариант	2 вариант	3 вариант
Эстетика	0	1	1
Эргономика	0,5	0,7	1
Экономика	1	0,5	0,7

Рис. 4. Таблица сравнительного анализа

5. Проектирование дополнительной функции

После того, как была определена наилучшая форма устройства, соответствующая эстетическим, эргономическим и экономическим показателям, было принято решение усилить критерий эстетики, возложив на форму дополнительную функцию декоративного украшения.

Каждая женщина желает быть особенной, иметь уникальные аксессуары, которые могли бы выделить ее из толпы. Таким ярким акцентом вполне может стать брошь. К основному устройству, носимому на теле, при помощи магнитов можно будет прикрепить различные броши поверх одежды. Такого рода украшения маскировали бы устройство, находящееся под одежной, а также доставляли приятные эмоции женщине, благоприятно влияя на ее психоэмоциональное состояние (Рисунок 4).

Рис. 4. Дополнительная функция декоративного украшения

При разработке украшений был использован метод художественного формообразования. Варианты решений при использовании данного подхода проектирования могут быть не обусловлены структурно-техническими параметрами.

Продукт дизайнерского формообразования представляется как продукт художественного формообразования, имеющий самоценность вне зависимости от целей проектирования. Характер данного метода, часто не имеет жесткой фиксации, сложно поддается анализу и структурированию [10,11]. Метод художественного формообразования абстрагирован от производства, однако максимально полно позволяет отразить и реализовать эстетическую выразительность.

С учетом дополнительной функции декоративного украшения была доработана форма считывающего

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

устройства. Формообразование дорабатывалось с применением композиционного ключа.

Композиционный ключ - это композиция, которая характеризует главные признаки объекта при помощи различных средств в композиции: пропорции, масштаба, симметрии, асимметрии и т.д. [12,13].

При проектировании основного корпуса измерительного устройства был проведен поиск композиционных решений.

Композиционный ключ был разработан с применением окружностей и настройкой симметрии. В нем отражалось больше логически обоснованных решений, нежели в первом варианте. Форма получилась более обтекаемой, острый уголок был сглажен за счет появления большой окружности, однако все также оставалось много недочетов. Маленькие окружности, образующие «крылья» формы были на неопределенном расстоянии от круга в центре

В дальнейшем были внесены уточнения. Видно, что центры окружностей располагаются на одной оси, все окружности, стоящие с одной стороны от оси симметрии, взаимодействуют между собой (Рисунок 5).

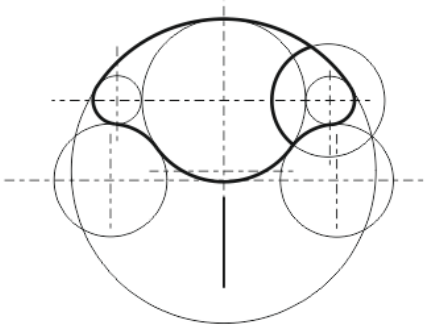


Рис. 5. композиционного ключа измерительного устройства

Кроме того, что данная форма обоснована композиционным ключом, она содержит в себе семантическое значение. Образ беременной женщины, символизирующий начало жизни, единство с внутреннем содержанием, предвкушение, гармонию. Данный образ отлично подходит под функциональные особенности, разрабатываемого многофункционального портативного устройства (Рисунок 7)

Рис. 7. Семиотический образ измерительного устройства

6. Заключение

При работе над формообразованием могут быть использованы различные методы дизайн-проектирования. При этом могут быть получены абсолютно разные результаты. Но при наличии определенных критериев можно сравнить варианты и определить оптимальный для конкретного медицинского устройства. Но наилучшего результата можно добиться, комбинируя и сочетая грамотно эти методы.

7. Литература

1. Абрамченко В.В., Коваленко Н.П. Перинатальная психология: Теория, методология, опыт. – Петрозаводск: ИнтелТек, 2004. – 350 с.
2. Влияние эмоций на беременность [электронный ресурс] https://psy.wikireading.ru/20783, дата обращения [10.04.18]
3. Методы проектирования в дизайне [электронный ресурс]: http://mydocx.ru/10-71206.html дата обращения [10.04.18]
4. Джонс Дж.К. Методы проектирования: Пер. с англ. – 2-е изд., доп.- М.: Мир, 1986.
5. Концепция и методы проектирования в дизайне [электронный ресурс] :https://studfiles.net/preview/2470636/ дата обращения [10.04.18]
6. Голубева, О. Л. Основы композиции / О. Л. Голубева. - М. : Искусство, 2004.
7. H. Alpay Er & John Langrish, 'Industrial design in developing countries: a review of the literature', IAS Research Papers, RP-66, Institute of Advanced Studies, Manchester Metropolitan University, 1993.
8. H. Alpay Er, 'Industrial design in newly industrialised countries: an explorotary study of the factors influencing the development of local design capabilities', MS Research Papers, RP-72, Institute of Advanced Studies, Manchester Metropolitan University,1993.
9. Волкотруб, И. Т. Основы художественного проектирования / И.Т.Волкотруб. - Киев Высшая школа, 1988.
10. Фрилинг, Г. Человек, цвет, пространство / Г. Фрилинг, К. Ауэр. - М. :Стройиздат, 1973.
11. TDesign and the state and the state of the design', Design, no. 495, 1990.
12. Устин, В. Б. Композиция в дизайне / В. Б. Устин. - М.: АСТ: Астрель, 2007.
13. Методика художественного конструирования. — М.: ВНИИТЭ, 1983. — 166 с.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Методы оценки и подготовки изображений в видеопотоке к распознаванию объектов

С.Г. Небаба[1]
stepanlfx@tpu.ru
[1]Томский политехнический университет, Томск, Россия

В работе рассматривается проблема выбора методов оценки и подготовки изображений в видеопотоке для систем распознавания объектов реального времени. Рассмотрены основные способы компенсации условий съемки и оптимизации подготовки изображений к распознаванию объектов, такие как оценка контрастности изображений, резкости, нормализация освещения, поиск и отбраковка похожих изображений. Проведен анализ существующих алгоритмов сравнения изображений, оценки контрастности и резкости изображения.

Ключевые слова: *компьютерное зрение, распознавание объектов, яркость, контрастность, резкость, видеопоток, режим реального времени.*

Methods of evaluation and preparation of images in the video stream for object recognition

S.G. Nebaba[1]
stepanlfx@tpu.ru
[1]Tomsk Polytechnic Univercity, Tomsk, Russia

The problem of choosing methods for estimating and preparing images in the video stream for real-time object recognition systems is considered in this paper. The main methods for compensating shooting conditions and optimizing the preparation of images for object recognition, such as estimating the contrast of images, sharpening, lighting normalization, searching and rejecting similar images, are considered. The current estimators of similar images search, the contrasting effect and image sharpness analyzed.

Keywords: *computer vision, object recognition, brightness, contrast, sharpness, video stream, real-time mode.*

1. Введение

Дисциплины компьютерного зрения и робототехники являются одними из наиболее активно развивающихся областей науки. Одной из фундаментальных проблем в этих сферах является оценка и нормирования уровней яркости, контрастности и цветности изображений. Эффективное решение этих проблем упрощает задачи слежения за объектами, их идентификации, определения и моделирования их свойств и внешнего вида [1].

С развитием информационных технологий, средств коммуникации, мобильных устройств и робототехнических комплексов наметился бурный прогресс в решении такого рода задач. Однако на сегодняшний день все еще не существует универсального способа быстрой и надежной идентификации объектов, даже с привлечением 3D моделей высокого разрешения на основе систем, требующих обработки данных в режиме реального времени.

Системы технического зрения (СТЗ) являются основной частью многих интеллектуальных систем, а также управляемых и автономных роботов. Сложность задач, решаемых такими системами, постоянно увеличивается, а требования к их техническим характеристикам повышаются. При этом все острее встают вопросы надежности и точности подобных автоматизированных систем, особенно для таких критичных направлений как безопасность предприятий, автономные транспортные средства, мобильные роботы и медицинские технологии. Продолжает расти потребность в быстрой некооперативной идентификации личности по изображению лица, в особенности в сфере безопасности, где оперативность реакции на событие имеет решающее значение [2].

В связи с этим возникает необходимость систематизации существующих и разработке новых методов и алгоритмов высокоскоростной оценки и обработки изображений. В перспективе это способствует повышению точности распознавания объектов в видеопотоке в режиме реального времени по его изображению, в том числе для автономных мобильных систем с возможностью массового применения и реализации.

В работах [3,4] была предложена комплексная технология обработки и подготовки изображений лиц в видеопотоке к распознаванию личности, рассчитанная на работу в режиме реального времени (25 кадров в секунду). Предложенная технология позволила повысить точность распознавания личности по изображению лица по сравнению с математическими классификаторами и добиться высокой скорости обработки видеопотока, превосходя по производительности нейросетевые методы распознавания. Однако, несмотря на эффективность предложенной технологии и повышение точности распознавания личности в режиме реального времени, проблема выбора методов оценки и нормализации освещения сцены, яркости, контрастности и цветности изображений в видеопотоке была рассмотрена недостаточно подробно и существует потенциал для улучшения системы в этом направлении.

Таким образом, возникает задача исследования методов, а также создания системы, которая будет способна осуществлять подбор наиболее эффективных методов оценки и нормирования изображений для конкретных задач распознавания объектов с сохранением высокой скорости обработки видеопотока.

2. Информативность изображения в видеопотоке

Для систем распознавания объектов чаще всего существуют неформализованные требования, что условия освещения должны позволять регистрировать изображения объектов, на которых могут быть различим основной рельеф (достаточный контраст). Важным условием

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

является наличие прямого или близкого к нему угла падения лучей света на фронтальную плоскость объектов, а также наличие рассеянного источника света. Заметное отрицательное влияние на качество распознавания оказывает боковая засветка объектов.

В большей части систем видеонаблюдения используется частота поступления кадров не менее 25 кадров в секунду. Очевидно, что изображения анализируемого объекта на двух соседних кадрах будут почти неотличимы друг от друга. Анализ таких практически идентичных кадров не оправдан с точки зрения использования вычислительных ресурсов. В связи с этим имеет смысл проводить отбраковку мало отличимых кадров.

Предварительный обзор позволил выделить 3 основных подхода сравнения изображений, получивших широкое распространение.

1. Применение хэш-функций к уменьшенной копии изображения и их сравнение. Данный метод обладает хорошим быстродействием, но дает очень грубую оценку, которая напрямую зависит от уменьшения изображения, поэтому может применяться только для поиска отдаленного визуального сходства объектов, например, поиску похожих картинок в коллекции. Метод относится к поиску изображений по содержанию (англ. Content-based image retrieval (CBIR)) [5] – разделу компьютерного зрения, решающему задачу поиска изображений, которые имеют требуемое содержание, в большом наборе цифровых изображений.
2. Вычисление коэффициента корреляции по одной из формул. Этот метод достаточно чувствителен к небольшим изменениям изображения, однако требователен к вычислительным ресурсам. Его применение ограничено специфическими областями работы с изображениями [6].
3. Сравнение изображений через построение и анализ SURF-дескрипторов, которые описывают некоторые небольшие области изображения. Исследования показывают, что данный метод показывает наилучшие результаты при сравнении изображений объектов, снятых под разными углами и в разных условиях освещения, однако он требует в несколько раз больше вычислительных ресурсов, чем сравнительно трудоемкий метод вычисления коэффициента корреляции [7]. Существует модификация данного метода (BRISK), которая обладает высоким быстродействием [8], однако данный метод сложно применить к поиску малозаметных отличий в кадрах.

Таким образом, задача сравнения кадров не имеет простого и эффективного решения, применимого к видеопотоку высокого разрешения, однако остается возможность разработать некий гибридный метод на основе рассмотренных выше.

3. Оценка контрастности

Перед тем, как улучшать изображение, необходимо дать оценку его качеству. Человек, бросив один взгляд на изображение, может сказать яркое оно или тёмное, чёткое или размытое, контрастное или нет и т.д. Алгоритмы же работают детально, анализируя изображение попиксельно или небольшими группами пикселей. Поэтому, на основании работы алгоритма, тяжело дать общую оценку качеству изображения.

Контраст (фр. contraste) – в оптике (сенситометрии и фотометрии) разница в характеристиках различных участков изображения, способность фотографического материала или оптической системы воспроизводить эту разницу, а также характеристика чувствительности глаза (зрительной системы) относительно яркости и цвета.

Контрастность (также, в различных контекстах употребляется и само слово контраст и коэффициент контраста) – степень контраста, чаще всего выражается безразмерной величиной, отношением или логарифмом отношений.

Мерой контрастности является величина $k = \frac{B_b - B_s}{B_s}$, где B_b – яркость фона, B_s – яркость предмета или изображения.

Контрастность изображения является показателем его информативности, так как этот показатель учитывает освещённость и различимость объектов на изображении. Этот параметр недостаточен, чтобы можно было признать изображение качественным, так как боковая засветка объектов дает высокую оценку контраста, но затрудняет выделение деталей изображения.

В случае, если изображение удается нормализовать, контрастность становится более полезной оценкой.

Различают 2 типа контрастности: яркостная и тоновая.

Яркостная контрастность представляет собой разницу между физической или видимой яркостью отдельных участков изображения. Вообще говоря, вычисление физической или видимой яркости можно рассматривать как конвертацию цветного изображения в ахроматические цвета. Поэтому яркостная контрастность – это сравнение двух участков изображения, приведенных к ахроматическим цветам.

Если проанализировать RGB-гистограммы, то можно сделать вывод, что у контрастного изображения количество тёмных и светлых пикселей должно быть приблизительно равным, разница в их яркости – значительна, а основное место сосредоточения пикселей – возле границ диапазона.

Тоновая контрастность учитывает не только яркостные переходы, но и переходы цвета. Ее имеет смысл применять при анализе цветных изображений, так как, несмотря на возрастание сложности вычислений, итоговая оценка становится более точной.

Подходящим критерием оценки яркостной контрастности может быть дисперсия яркости пикселей изображения [9]:

$$\sigma^2 = \frac{1}{N}\sum_{p=1}^{N}(Y_p - Y)^2 , \qquad (1)$$

где Y – среднее значение яркостной контрастности всего изображения, Y_p – значение яркостной контрастности в точке p, N – общее число точек изображения.

Оценка контрастности (1) может быть нормирована путём вычисления отношения среднеквадратического отклонения к максимально возможному значению яркости:

$$C = \frac{2\sigma}{Y_{\max}} . \qquad (2)$$

Диапазон изменения значения C – [0;1]. Значение 0 соответствует однотонному изображению, значение 1 — максимально контрастному. Оптимальное значение контрастности (2) зависит от типа объекта, представленного на изображении.

Контрастность изображения не может являться точным показателем качества изображения: размытые изображения менее информативные, но при этом имеют более высокую оценку контрастности, чем четкие изображения с меньшим числом точек максимальной и минимальной яркости.

4. Оценка резкости

Резкость чаще всего трактуется как качество воспроизведения на фото достаточно крупных фрагментов. В этом, условно, ее отличие от четкости изображения, характеризующей различимость мелких деталей.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Повышение резкости – один из самых распространенных способов улучшения фотографий. Обычно всегда требуется в той или иной степени усиливать резкость первоначального изображения.

Объекты на изображении отличаются друг от друга по яркости или оптической плотности, что является причиной их различимости. Граница же между объектами всегда имеет ту или иную степень размытости.

Ширина размытости теоретически резкой границы между светлыми (яркими) и темными участками (объектами) изображения характеризует резкость изображения. Причем это характерно для любых изображений, как цветных, так и черно-белых.

Четкость изображения – это параметр, который характеризует качество полученного изображения и показывает, насколько отчетливо различимы мелкие объекты на изображении и определяется минимальным расстоянием между ними.

Другими словами, четкость показывает полноту воспроизведения мелких объектов на изображении и насколько у них резкие контуры.

Изображение высокой четкости имеет резкие, четко видимые границы.

Резкость может выступать еще одним критерием информативности изображения, она в общем случае может быть определена наличием контуров контрастных переходов. Однозначного метода для определения резкости не существует, но есть несколько алгоритмов, которые позволяют вычислить оценку резкости из различных характеристик изображения.

Главное различие всех предлагаемых методов состоит формулах, которые определяют количественную оценку контрастности изображения. В целом все методы можно разделить на 3 группы:

1. методы, вычисляющие среднее значение дифференциальных операторов изображения;
2. методы, вычисляющие эксцесс амплитудного спектра изображения;
3. методы, вычисляющие отношение значимых и незначимых точек амплитудного спектра изображения.

Первая группа методов вычисляет разностную характеристику напрямую из исходного изображения, например с помощью оператора Лапласа или оператора Робертса, вторая и третья группы вычисляют частотные характеристики изображения (преобразование Фурье, либо Wavelet-преобразование), и вычисляет количество контрастных переходов из спектра изображения.

Идея метода, использующего дифференциальные операторы, заключается в вычислении среднего значения яркости от преобразованного оператором изображения. Такой метод чувствителен к качеству изображения, а также к характеру изображенных объектов, то есть он может дать высокую оценку резкости для плохого изображения со сложным объектом, и меньшую – для хорошего изображения с малым числом контрастных переходов.

Оператор Лапласа – дифференциальный оператор, сумма вторых частных производных по координатам изображения, эквивалентен операциям градиента и дивергенции:

$$Laplace(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Оператор Робертса – свертка изображения с ядрами

$$\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \text{ и } \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}.$$

Оператор Робертса по сравнению с оператором Лапласа требует меньше вычислений, более точно выделяет тонкие контуры, но очень чувствителен к шуму.

В работе [10] рассмотрен метод оценки резкости на основе преобразования Фурье, суть которого заключается в вычислении эксцесса изображения в частотной области:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}, \quad (3)$$

где μ_2 – второй центральный момент изображения, μ_4 – четвертый центральный момент, σ – стандартное отклонение.

В работе [11] также используется преобразование изображения в частотную область, но в качестве критерия оценки резкости предлагается отношение

$$FM = \frac{T_H}{M \cdot N}, \quad (4)$$

где T_H – число значимых точек, $M*N$ – размер изображения амплитудного спектра. Значимые точки определяются как точки, значение в которых превышает некоторую долю максимума амплитуды спектра.

Исследование методов показало, что оценка резкости с использованием оператора Лапласа даёт заметный разброс результатов при схожих условиях освещения и ракурса, что связано отсутствием нормирования результата и способом выделения области объекта. Оценка (3) показывает более стабильные результаты, но практически не реагирует на фактическое снижение резкости, полученное размытием по Гауссу. Оценка (4) наиболее точно соответствует изначальной задаче выделения поиска резких изображений, она корректно реагирует на изменения резкости при движении и отсутствие фокуса, и помимо этого имеет строгую нормировку, позволяющую по заданному уровню отсеивать изображения неудовлетворительного качества.

Несмотря на достоинства оценки (4), она отчасти подвержена тем же проблемам, что и оценка с использованием Лапласиана. Основная причина этого в том, что резкость изображения оценивается количественно, и потому связана с общим числом контрастных переходов, а это число зависит от объекта на изображении. Исходя из этого, можно говорить только об относительной резкости изображения, которая оценивается для одного объекта в треке (серии изображений).

5. Компенсация освещения

Повысить применимость методов (1)-(4) может компенсация изображения объекта по условиям освещения.

Для нормирования исходного изображения по условиям освещения существует несколько методов, например DoG-фильтрация [12] и частотная фильтрация полосовым фильтром [13], убирающим низкие и высокие частоты. Оба метода позволяют эффективно выравнивать изображение при дисбалансе яркости по освещению.

DoG обнаруживает края путем применения размытия изображения по Гауссу (Gaussian blur) с указанным значением стандартного отклонения. Результирующее изображение является размытой версией исходного. Далее применяется другое размытие с меньшим значением стандартного отклонения, что дает более резкое изображение, чем в предыдущем случае. Финальное изображение вычисляется заменой каждого пикселя разностью от пикселей двух размытых изображений.

Для нормализации яркости изображения с этим методом часто используется гамма-коррекция.

Частотный полосовой фильтр при нормализации освещения подавляет высокочастотные шумы и сглаживает неравномерное освещение изображений, кодируемое в изображении низкими частотами.

Оба метода повышают четкость изображения, но не устраняют полностью влияние неравномерного освещения.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Также для повышения контраста изображения применяется адаптивное выравнивание гистограммы с ограничением (CLAHE) [14]. Первоначально изображение разбивается на прямоугольные области (блоки). В каждом блоке вычисляется гистограмма распределения яркости. Чтобы избежать перенасыщения монотонных участков изображения, ограничивается наклон функции распределения яркости путем введения ограничения на максимальное значение частоты яркости. Пусть M – число строк в блоке, N – число столбцов в блоке, α – параметр, задающий ограничение. Предельное значение гистограммы вычисляется в соответствии с формулой:

$$hist_{\lim} = hist_{\min} + \alpha(MN - hist_{\min}),$$

где $hist_{\min}=MN/(L+1)$, L – максимальное значение яркости.

После ограничения для каждой области формируется своя функция преобразования яркости на основании выравнивания гистограммы распределения яркости. Для входного изображения вычисляется значение яркостей выходного изображения в соответствии с формулой:

$$LUT(k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n},$$

где $LUT(k)$ - значение яркости выходного изображения, соответствующее яркости r_k входного изображения, $k \in [0,L]$; r_j – яркость элемента входного изображения, n_j – число элементов входного изображения со значением r_j, n – общее число элементов изображения; $p_r(r_j)$ – вероятность появления элемента входного изображения с яркостью r_j.

Такое преобразование позволяет равномернее распределить уровни яркости по частоте, в основе преобразования лежит предположение, что в контрастных черно–белых изображениях значения яркости распределены равномерно.

Чтобы разрушить блочную структуру полученного изображения, выходное изображение формируется путем билинейной интерполяции по четырем ближайшим LUT.

6. Заключение

В работе проведен обзор существующих методов подготовки изображений к процессу распознавания объектов и критериев оценки качества изображения.

Проведен анализ существующих алгоритмов сравнения похожих изображений, а также алгоритмов оценки контраста, которые подходят для проверки условий освещения, позволяя предотвратить использование системы распознавания объектов на малоконтрастных изображениях и, как следствие, снизить вероятность ложноположительных срабатываний.

Проведены обзор и сравнение алгоритмов оценки резкости изображения. Оценка резкости позволяет ранжировать изображения в видеопотоке по количеству контрастных переходов, в общем случае давая возможность выбрать из всего количества кадров более качественные.

Рассмотрен метод компенсации условий освещения для изображений с неравномерной яркостью.

Применение всех рассмотренных методов и алгоритмов в подходящих условиях позволит снизить ресурсоемкость систем распознавания объектов по изображению, а также устранить влияние неблагоприятных для процесса распознавания факторов.

7. Благодарности

Работа выполнена в рамках госзадания № 2.1642.2017/4.6 на выполнение проекта по теме «Когнитивные методы визуализации и анализа многомерных данных при моделировании нелинейных динамических систем».

8. Литература

[1] Hartley R., Zisserman A. Multiple view geometry in computer vision. – Cambridge university press, 2003.

[2] Sharif M. et al. Face Recognition: A Survey // Journal of Engineering Science & Technology Review. – 2017. – Т. 10. – №. 2.

[3] Небаба С.Г. Тестирование технологии подготовки изображений лиц к распознаванию личности в видеопотоке в режиме реального времени // Современная наука: актуальные проблемы теории и практики. Серия «Естественные и технические науки», 2017. №3-4. С. 73-77.

[4] Небаба С.Г. Захарова А.А. Алгоритм построения деформируемых 3D моделей человеческого лица и обоснование его применимости в системах распознавания личности // Труды СПИИРАН. 2017. Вып. 52. С. 157-179.

[5] Charles E. Jacobs A. Finkelstein D.H. Sales in Fast Multiresolution Image Querying // Department of Computer Science and Engineering, University of Washington, Seattle, Washington 98195.

[6] Альтман Е.А., Захаренко Е.И. Быстрый алгоритм вычисления двумерной корреляции для видеообработки // Доклады ТУСУР. 2015. №2 (36) С.119-124.

[7] Bay H., Ess A., Tuytelaars T., Van Gool L. SURF: Speeded up robust features // Computer Vision and Image Understanding. – 2008. – V. 110. – P. 346–359.

[8] Leutenegger S. BRISK: Binary Robust invariant scalable keypoints / S. Leutenegger, M.Chli, R.Y. Siegwart // Proceedings of the 2011 International Conference on Computer Vision (ICCV '11), 6 November 2011, P. 2548-2555.

[9] Фисенко В.Т. Компьютерная обработка и распознавание изображений: учебное пособие / В.Т. Фисенко, Т.Ю. Фисенко - СПб: СПбГУ ИТМО, 2008. – 192 с.

[10] Монич Ю.И. Оценки качества для анализа цифровых изображений / Ю. И. Монич, В. В. Старовойтов. - Минск: Государственное научное учреждение ОИПИ НАН Беларуси, 2008.

[11] Kanjar D. Image Sharpness Measure for Blurred Images in Frequency Domain / D. Kanjar, V. Masilamani // International Conference on Design and Manufacturing. - Procedia Engineering, 2013. - P. 149 – 158.

[12] Tan X. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions / X. Tan, B. Triggs // IEEE Transactions on image processing, Vol. 19, № 6, June 2010. – P. 1635-1650.

[13] Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс // Москва: Техносфера, 2012. – 1104 с.

[14] Фисенко Т.Ю. Исследование и разработка методов улучшения подводных изображений / Т. Ю. Фисенко, В. Т. Фисенко – Материалы международной конференции «Прикладная оптика», 15-19 октября 2012. - С. 294-298.

Об авторах

Небаба Степан Геннадьевич, к.т.н., инженер научно-образовательной лаборатории 3D моделирования и промышленного дизайна Инженерной школы информационных технологий и робототехники Томского политехнического университетета. Его e-mail stepanlfx@tpu.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Реверсивный инжиниринг палеоантропологических объектов для визуализации и исследований

М.М. Новиков[1], В.А. Князь[2], Р.М. Галеев[3]
novikov@rambler.ru | knyaz@gosniias.ru | ravil.galeev@gmail.com
[1]Институт проблем лазерных и информационных технологий РАН – филиал Федерального государственного учреждения «Федеральный научно-исследовательский центр «Кристаллография и фотоника» Российской академии наук», г.Шатура Московской обл., РФ;
[2]Федеральное государственное унитарное предприятие «Государственный научно-исследовательский институт авиационных систем», г. Москва, РФ;
[3]Федеральное государственное бюджетное учреждение науки Институт этнологии и антропологии им. Н.Н. Миклухо-Маклая Российской академии наук , Москва, РФ

Возможности компьютерной графики для визуализации сложных структур и процессов позволяют с успехом применять их для решения сложных задач в широком спектре приложений. Использование их потенциала для решения задач палеоантропологических исследований позволяет разрабатывать и применять новые методы для обработки, анализа, визуализации и обмена уникальными данными. Представлены результаты разработки и применения методов точной фотореалистичной 3D реконструкции и обратного инжиниринга для создания цифровых трёхмерных моделей палеоантропологических объектов и их твёрдых копий. Приведено сравнение точности измерения антропологических параметров разработанными методами с результатами стандартных ручных измерений.

Ключевые слова: *палеоантропологические материалы, бесконтактные методы измерений, цифровая модель, аддитивные технологии, компьютерная томография, фотограмметрия.*

Reverse engineering of paleoanthropological objects for visualization and research

M.M. Novikov[1], V.A. Knyaz[2], R.M. Galeev[3]
novikov@rambler.ru | knyaz@gosniias.ru | ravil.galeev@gmail.com
[1]Research Center Crystallography and Photonics RAS, Shatura, Russia;
[2]State Research Institute of Aviation System (GosNIIAS), 125319 Moscow, Russia;
[3]Institute of Ethnology and Anthropology RAS, Moscow, Russia

The possibilities of computer graphics for visualization of complex structures and processes make it possible to successfully apply them to solve complex problems in a wide range of applications. Using their potential in paleoanthropological research allows to develop and apply new methods for processing, analyzing, visualizing and exchanging unique paleoanthropological data. The results of development and application of methods of accurate photorealistic 3D reconstruction and reverse engineering for creation of digital 3D models of paleoanthropological objects and their hard 3D copies are presented. Comparison of the accuracy of measurement of anthropological parameters by the developed methods with the results of standard manual measurements is given.

Keywords: *palaeoanthropology, non-contact measurements, digital model, additive technologies, computed tomography, photogrammetry.*

1. Введение

Инженерная графика и её современное цифровое воплощение – компьютерная графика – служат универсальным языком общения инженеров и учёных. Возможности компьютерной графики для визуализации сложных структур и процессов позволяют ставить и решать новые задачи в различных областях знаний, в том числе, гуманитарных, таких как палеоантропология.

Особенность исследований в палеоантропологии связана с уникальностью и сложностью объектов исследований, существующих в единичных экземплярах, имеющих сложную структуру и обычно требующих очень бережного обращения при выполнении исследований. Данные факторы существенно затрудняют и замедляют процесс исследований, ограничивают возможности доступа к объектам и обмена информацией между исследователями, сдерживают применение новых методов обработки данных. Создание геометрически точных и визуально адекватных цифровых компьютерных моделей методами обратного инжиниринга позволяют в значительной степени решить указанные проблемы.

Настоящая работа представляет результаты исследований по разработке методов и программно-аппаратных средств обратного инжиниринга для создания точных и визуально корректных цифровых моделей палеоантропологических объектов и их твёрдых копий для задач исследований и обучения.

Одной из основных задач палеоантропологии является морфологическое описание костных останков с

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

использованием ряда измерительных методов (краниометрия, остеометрия и т.д.), позволяющих проводить широкие сравнительные исследования, цель которых описать разнообразие антропологических типов, как в древности, так и в современности.

В современной физической антропологии, несмотря на значительное развитие современной измерительной техники, используется ручной инструментарий рис. 1. Отчасти, это связано с тем, что разработанные измерительные методики палеоантропологических объектов могут быть использованы с конкретным набором измерительного инструмента – толстотного, координатного, скользящего циркуля и других инструментов.

Рис. 1. Механические инструменты для измерения антропологических объектов.

Кроме того, достоверность ручных замеров в полной мере соответствует необходимой точности (точность измерений проводится до 0,1 см). При этом у ручного измерения есть ряд очевидных недостатков – это, прежде всего, трудоемкость, времязатратность, нередко повреждения объекта исследования, и, как ни удивительно, крайне высокая стоимость антропологических измерительных инструментов.

2. Требования к цифровым моделям палеоантропологических объектов

Палеоантропология предъявляет особые требования к создаваемым цифровым моделям [4]. К основным требованиям следует отнести:

1. высокая геометрическая точность (выше точности стандартных ручных измерений) на уровне 0.01 ... 0.05 мм;
2. высокий уровень детализации на уровне 0.05 ... 0.1 мм, обеспечивающий возможность локализации заданных антропологических ориентиров [3];
3. высокая степень визуального соответствия, позволяющая исследователю получать достоверную информацию о визуальных параметрах объекта (цвет объекта, заданные контуры и т.п.);
4. высокая точность нанесения фотореалистичной текстуры, обеспечивающая возможность проведения измерений с использованием координат измеряемых точек на текстуре [2];
5. представление 3D модели в форме, пригодной для изготовления на установке быстрого прототипирования.

Сформулированные требования не могут быть комплексно реализованы с помощью коммерческих продуктов. Поэтому для исследований разработана оригинальная система бесконтактных оптических измерений и построения 3D моделей, отвечающая сформулированным требованиям. Она создана на основе имеющихся технических решений и оригинальных программных библиотек и обеспечивает возможность развития новых методов обработки данных на основе расширения функций системы.

3. Система построения фотореалистичных 3D моделей

Для решения специфических задач палеоантропологических исследований создана специализированная система 3D сканирования, основанная на фотограмметрических принципах трёхмерных измерений. Она является развитием существующей системы [1] и дополнена аппаратными и программными средствами для работы с палеоантропологическими объектами.

Система включает 2 цифровых камеры высокого разрешения, программно-управляемый проектор структурированного света для подсветки образца и цифровую камеру для получения текстуры высокого разрешения. Оригинальное программное обеспечение позволяет выполнять 3D сканирование объекта и получение геометрических данных в автоматизированном режиме (рис. 2). Также система обеспечивает оперативное изменение конфигурации для сканирования в требуемом рабочем пространстве с целью обеспечения заданных требований по точности и подробности измерений.

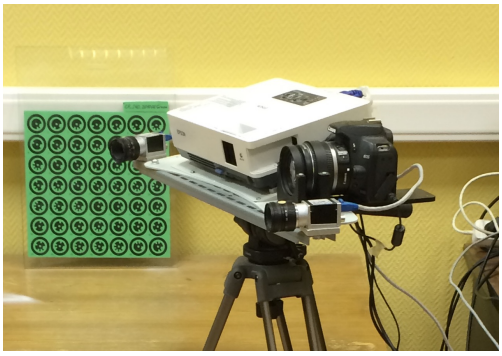


Рис. 2. Фотограмметрическая система.

Оригинальная процедура калибровки обеспечивает точность сгенерированных 3D моделей на уровне 0.01 ... 0.05 мм в зависимости от размеров рабочей области сканирования, что соответствует требованиям по точности для задач антропологического анализа.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Текстурированная цифровая модель предоставляет эксперту больше информации (рис. 3), так как некоторые особенности могут быть найдены только на цветном изображении объекта. Так как калибровка фотограмметрической системы выполняется одновременно для всех трех камер с использованием одного калибровочного поля, то это обеспечивает точное нанесение текстуры (соответствие координат изображения трёхмерным точкам поверхности объекта).

Рис. 3. Изображение объекта (слева) и его текстурированная цифровая модель (справа).

Качество получаемых моделей позволяет их использовать для палеоантропологических исследований и решения таких задач, как:

1. измерение морфометрических антропологических параметров (расстояния, углы, путь вдоль поверхности для заданных антропометрических точек объекта);
2. виртуальная реставрация (восстановление утерянных фрагментов) палеоантропологических объектов;
3. виртуальная реконструкция внешнего облика;
4. создание специализированных баз данных цифровых компьютерных моделей палеоантропологических объектов, предоставляющих возможность работы с ценными находками широкому кругу исследователей;
5. создание точных твёрдых копий палеоантропологических объектов для целей обучения, для физической реставрации и реконструкции внешнего облика.

В качестве альтернативы традиционной методике предложен комплекс алгоритмов автоматического обнаружения и распознавания ориентиров. Они используют как цифровую модель объекта, так и текстуры высокого разрешения для распознавания набора необходимых ориентиров.

4. Автоматизированные измерения антропологических параметров

Оценка точности и удобства применения разработанных методов автоматических морфометрических измерений проводилась на 3D-моделях палеоантропологических объектов, полученных с помощью разработанной системы построения моделей. Ряд стандартных морфометрических измерений был выполнен с помощью традиционных (ручных) и автоматизированных методов (рис. 4).

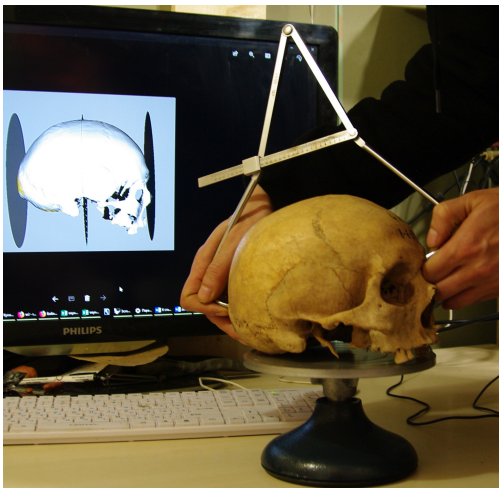


Рис. 4. Измерение антропологических параметров

Результаты измерений заданных антропологических параметров приведены в таблице ниже.

#	Параметр	Ручн.	3D ручн.	3D авт.
1	d_{long}, mm	197	197.3	197.2
8	d_{tr}, mm	154	154.1	154.5
9	$w_{min}^{forehead}, mm$	131	131.3	131.0
10	$w_{max}^{forehead}, mm$	136	136.2	135.9
17	d_{height}, mm	119	119.2	118.9
32	$\alpha_{forehead}, ^\circ$	82	82.4	81.6
72	$\alpha_{profile}, ^\circ$	95	94.7	95.3
75	$\alpha_{nose}, ^\circ$	53	52.5	52.9

Сравнение результатов измерений подтверждает высокую точность 3D реконструкции палеоантропологических объектов и возможность применения интеллектуальных методов обработки данных в палеоантропологических исследованиях.

5. Реконструкция с применением аддитивных технологий

Важным преимуществом применения цифровых моделей в практике палеоантропологических исследований является возможность получения их точных твёрдых копий с помощью аддитивных технологий [5]. Такие копии палеоантропологических объектов могут быть использованы для художественной реконструкции внешнего облика, для создания справочных коллекций в целях обучения.

Аддитивные технологии создания твёрдых копий цифровых моделей задают к ним специальное требование – требование «водонепроницаемости» или «твердотельности» трёхмерной модели. Другими словами, для изготовления твёрдой копии цифровой модели требуется, чтобы её любое сечение было замкнуто. Для

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

выполнения данного требования разработаны методы преобразования 3D моделей, полученных фотограмметрической системой, в «твердотельную» форму.

Пластиковые копии палеоантропологических (рис. 5, слева) объектов по преобразованным в «твердотельную» форму цифровым моделям были изготовлены на лазерном стереолитографе ЛС250, разработанным в ИПЛИТ РАН.

Также были исследованы качество и точность 3D-печати с точки зрения возможности применения таких копий в палеоантропологических исследованиях. Для оценки точности стереолитографическая модель сканировалась фотограмметрической системой, и полученная цифровая модель сравнивалась с исходной моделью. Результаты сравнения этих двух моделей показаны на (рис. 5, справа). Стандартное отклонение для разности поверхностей составило 1.76 мм.

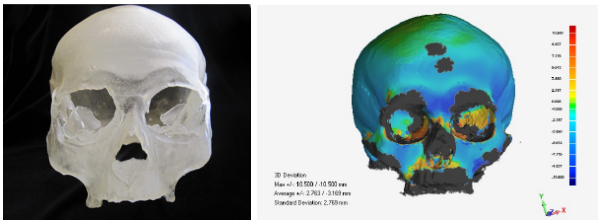


Рис. 5. Твёрдая копия палеоантропологического объекта (слева) и сравнение с исходной моделью (справа).

Результаты оценок точности реконструкции стереолитографии показывают, что модели, созданные методом стереолитографии, могут быть использованы для решения таких антропологических задач, как реконструкция внешнего вида человека и восстановление объекта.

6. Заключение

Разработана замкнутая технология построения точных фотореалистичных цифровых моделей палеоантропологических объектов, обеспечивающая геометрическую точность и визуальную адекватность моделей, необходимые для проведения на них палеоантропологических исследований. Технология основана на фотограмметрических методах измерений, обеспечивающих требуемую точность 3D реконструкции и фотореалистичное текстурирование модели, и базируется на оригинальных аппаратно-программных средствах бесконтактных измерений и обработки данных.

Разработаны методы автоматических измерений ряда антропологических параметров по цифровым компьютерным моделям и проведено сравнение результатов автоматических и ручных измерений заданных параметров. Сравнение показало высокую точность разработанных методов и их применимость для палеоантропологических исследований.

Изучена применимость использования стереолитографических копий палеоантропологических объектов для исследований и обмена редкими данными. Оценка точности полученных стереолитографических копий показывает, что их можно успешно использовать в образовательных целях и для задач реконструкции внешнего вида.

7. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 17-29-04509 офи-м.

8. Литература

[1] V. A. Knyaz. Image-based 3d reconstruction and analysis for orthodontia. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B3:585–589, 2012.

[2] V. A. Knyaz, S. Y. Zheltov, and A. Chibunichev. Photogrammetric techniques for analysis and visualization of changes in 2d and 3d data: Plastic surgery application. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W4:109–114, 2017.

[3] Gérard Subsol, Bertrand Mafart, Alain Silvestre, and M.A. De Lumley. 3D Image Processing for the Study of the Evolution of the Shape of the Human Skull: Presentation of the Tools and Preliminary Results. In Bertrand Mafart, Hervé Delingette, and Gérard Subsol, editors, *Three-Dimensional Imaging in Paleoanthropology and Prehistoric Archaeology*, pages 37–45. BAR International Series 1049, 2002.

[4] Gerhard W. Weber. *Virtual Anthropology and Biomechanics*, pages 937–968. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.

[5] С.С. Абрамов, Н.И. Болдырев, М.М. Новиков, А.В. Евсеев, Е.В. Коцюба, В.Я Панченко, Н.М. Семешин, and В.П. Якунин. О возможности применения метода лазерной стереолитографии в судебной медицине. *Судебно-медицинская экспертиза*, 41(3):13–17, 1998.

Об авторах

Новиков Михаил Михайлович, заведующий лаборатории ЛСОИ, Институт проблем лазерных и информационных технологий РАН – филиал Федерального государственного учреждения «Федеральный научно-исследовательский центр «Кристаллография и фотоника» РАН», ул. Святоозерская, д. 1, г. Шатура, Московская область, РФ, novikov@rambler.ru тел.: (49645) 22200 доб.459.

Князь Владимир Александрович, кандидат технических наук, зам. руководителя отдела, ФГУП «Государственный научно-исследовательский институт авиационных систем», г. Москва, РФ.

Галеев Равиль Марветович, младший научный сотрудник, ФГБУН Институт этнологии и антропологии им. Н.Н. Миклухо-Маклая РАН, г. Москва, РФ.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Подходы к созданию культиватора микроводорослей для домашнего использования

Романенко С.А.[1], Шакиров Е.М.[1], Туранов С.Б.[1]
sofia.romanenko98@gmail.com|shakirov@tpu.ru|tyrsb@yandex.ru
[1]Национальный исследовательский Томский политехнический университет, г. Томск, Россия;

Авторами выявлена проблема доступности повсеместного культивирования микроводорослей при наличии высокого спроса со стороны населения. В целях решения указанной проблемы предложен концепт компактного культиватора микроводорослей для домашнего использования. Приведены основные требования к форме установки. Выявлена значимость учета параметров освещения при формообразовании устройства

Ключевые слова: *микроводоросли, хлорелла, спирулина, культиватор, фотобиореактор.*

The concepts of creating a microalgae cultivator for domestic use

S. Romanenko[1], E. Shakirov[1], S. Turanov[1]
sofia.romanenko98@gmail.com|shakirov@tpu.ru|tyrsb@yandex.ru
[1]National Research Tomsk Polytechnic University, Tomsk, Russia;

Authors have determined the issue in the widespread cultivation of microalgae due to increased demand. To solve this problem, the concept of compact microalgae cultivator for domestic use has been proposed. Main claims to the construction have been introduced. It has been pointed out, that criteria of lightning should be taken into account when the generation of geometry is occuring.

Keywords: *microalgae, chlorella, spirulina, cultivator, photobioreactor.*

1. Введение

За последние десятилетия идея здорового образа жизни среди интересов современного населения приобрела первостепенное значение, в связи с чем становятся крайне востребованными всевозможные биологически активные пищевые добавки. Данный факт обуславливает и возросший спрос среди потребителей. Так, к примеру, в период с января по март 2018 года был выявлен прирост спроса на биологически активные добавки среди граждан РФ в размере 337% [8]. Схожие показатели наблюдаются и в других странах. Одним из наиболее перспективных представителей данной категории добавок являются микроводоросли, которые обладают рядом полезнейших свойств.

Так, например, хлорелла - один из представителей рода одноклеточных зеленых водорослей - является активным продуцентом биомассы и содержит высокий уровень полноценных белков, жиров, углеводов и витаминов, что, безусловно, вызывает интерес у населения. Хлорелла входит в категорию так называемых «суперпродуктов» (продуктов с высоким содержанием витаминов, микроэлементов и биологически активных веществ). Среди многочисленных употребляемых в пищу растений хлорелла стоит на первом месте по очень многим показателям. При анализе биомассы хлореллы установлено следующее процентное соотношение: 40-60 % белков, 30-35 % углеводов, 5-10 % липидов и до 10 % минеральных веществ[5].

В настоящее время микроводоросли активно культивируются более чем в 70 странах мира и применяются в пищевой промышленности в виде порошков и таблеток. Известно, что наилучшей усвояемостью в организме человека по сравнению с порошковой формой продукта обладают суспензии [2]. Таким образом, для получения максимальной пользы необходим прием микроводорослей в виде суспензии. В связи с тем, что данный продукт производится преимущественно в промышленных масштабах, покупатели среднего и низкого уровня достатка далеко не всегда могут позволить себе приобретение микроводорослей, поскольку производитель в связи с затратностью производства реализует свою продукцию по высокой цене и объемом упаковки, не рассчитанной на единичного потребителя [7]. Как следствие, появляется заинтересованность в создании домашней установки, которую смогут использовать потребители разного уровня достатка.

2. Проектирование культиватора

На данный момент, для культивирования микроводорослей применяются специальные устройства, обычно называемые фотобиореакторами, которые создают оптимальные условия для их жизни и размножения. Они, в свою очередь, разделяются на открытые и закрытые. Открытые фотобиореакторы используются преимущественно в промышленности, тогда как закрытые можно использовать как на производстве, так и в бытовых условиях. Производительность процесса выращивания в последних в значительной степени зависит как от параметров самого устройства (культиватора), так и от внешних условий.

Одним из важнейших параметров, оказывающих значительное влияние на процесс роста микроводорослей, является свет. Так, в качестве источника света в закрытых установках традиционно применяют лампы накаливания, в том числе кварцевые галогенные с отражателями, зеркальные лампы, люминесцентные. Кроме того, наблюдается также использование дуговых, ртутных, люминесцентных, ксеноновых и натриевых ламп. По сравнению с естественными источниками света искусственные источники смогут обеспечивать необходимую освещенность в течение долгого периода

времени вне зависимости от погодных условий, что выгодно отличает последние от естественных источников света [1]. Помимо этого, при проектировании фотобиореакторов также представляется возможным применение светодиодных источников света, которые отличаются большей долговечностью, компактностью и энергоэффективностью.

Наиболее важной задачей при разработке оптимального в сложившейся ситуации культиватора стало формообразование. Данный процесс не стоит воспринимать только лишь как проектирование художественной формы, необходимо учесть и значимость структурирования реальной среды с учетом жизненных процессов и рациональных принципов организации бытового пространства [3]. Устройство должно органично вписываться в повседневную обстановку и иметь эстетически приятный для потребителя вид.

Более того, с точки зрения освещения, форма должна обеспечивать наибольшую освещенность при наименьших потерях. Проект ожидаемого распределения светового излучения внутри формы в зависимости от расположения светильника представлен на рис. 1.

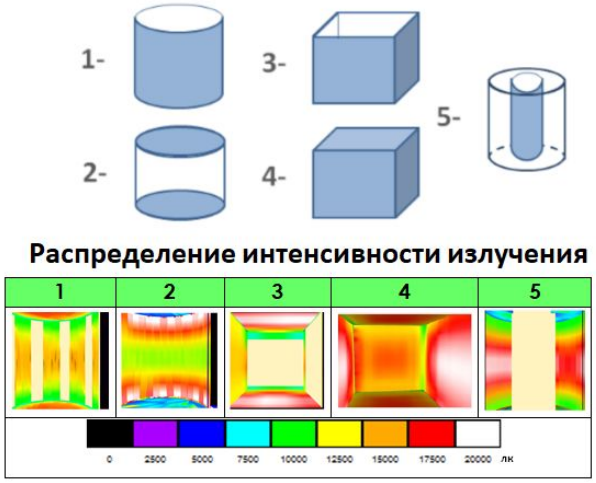


Рис. 1. Распределение интенсивности излучения в зависимости от формы.

Исходя из вышеизложенного, в основу конструкции лег кувшин цилиндрической формы с погружным светодиодным светильником. Одной из главных особенностей данного культиватора является возможность регуляции уровня досветки микроводорослей, что позволяет уменьшить количество потребляемой электроэнергии.

Кроме того, на продуктивность влияют и размеры фотобиореактора. Исходя из рекомендаций, профилактическая доза микроводорослей в сухом виде – 2 г/л [6]. Ожидаемая концентрация сухого вещества в суспензии – 3 г/л [4]. Таким образом, из расчета на среднестатистическую семью из 4 человек выявлено, что наиболее оптимальным объемом суспензии является 3 л. С учетом объема, занимаемого погружным светильником и другими элементами конструкции, конструкция должна иметь габариты 27,5×19×18,5 см и общий объем 4,2 л.

Предлагаемый культиватор микроводорослей для домашнего использования представлен на рис. 2.

Рис. 2. Концепт домашнего культиватора для микроводорослей.

3. Заключение

Предложенный концепт культиватора имеет значительное преимущество перед аналогами, поскольку он отличается компактностью, простотой конструкции, энерго- и ресурсоэффективностью, а также приятен с эстетической точки зрения. На пике популярности различных пищевых добавок такое устройство легко найдет свое место в быту.

Дальнейшая работа в данной области будет направлена на исследование влияния параметров разрабатываемого фотобиореактора и на оптимизацию его параметров для достижения максимальной эффективности.

4. Литература

[1] Алексеев М.А., Арьянова Э.Д., Иванова С.С. Установка для культивирования микроводоросли хлореллы// Архитекторы будущего: сб. статей. – Томск, 2014. – С. 377 – 381.

[2] Белоусов Ю.Б. Гуревич К.Г. Клиническая фармакокинетика. Практика дозирования лекарств. – М.: Литтерра, 2005. – 288 с.

[3] Бондарев Ю.И., Степанова-Третьякова Н.С. Формообразование как основа дисциплин «Дизайн-проектирование» и «Рисунок»// Наука. Искусство. Культура. – 2016. – №4(12). – С. 111 – 120.

[4] Мещерякова Ю.В., Нагорнов С.А., Ерохин И.В. Накопительное культивирование микроводоросли хлорелла в закрытом фотобиореакторе// Наука в Центральной России. – 2015. – № 2(14). – С. 92 – 100.

[5] Музафаров А.М. Таубаев Т.Т. Культивирование и применение микроводорослей. – Т.: ФАН Узбекской ССР, 1984. – 122 с.

[6] Смирнова Г.А. Рекомендации по применению спирулины [Электронный ресурс] / Г.А. Смирнова. – Электрон. текстовые дан. – Режим доступа: http://www.spirulina.com.ua/stati/1recomendacii.html, свободный.

[7] Живая спирулина [Электронный ресурс]. – Режим доступа: http://spirulina-s.ru/spirulina_kupit, свободный.

[8] Названы самые популярные товары у россиян весной [Электронный ресурс]. – Электрон. период. изд. – Москва: [б.и.], 2018. – Режим доступа: https://lenta.ru/pressrelease/2018/04/12/priceru/, свободный.

Особенности дизайн-проектирования установки для переработки полимеров

Сафьянникова В.И., Вехтер Е.В.
victoriasafren@gmail.com|vehter@tpu.ru
Томский политехнический университет, Томск, Россия

В статье рассматриваются особенности создания дизайн-проекта технического объекта, в данном случае это установка по переработке полимеров. До стадии проектирования необходимо провести аналитический обзор и учесть все характеристики объекта для дальнейшего определения критерий проектирования. По итогам работы была создана модель установки по переработке полимеров с учетом технических особенностей, методов формообразования и модульности объекта.

Ключевые слова: *дизайн, проектирование, установка, полимеры, дизайн-проект.*

Features of designing an installation of polymers recycling

Safyannikova V.I., Vekhter E.V.
victoriasafren@gmail.com|vehter@tpu.ru
Tomsk Polytechnic University, Tomsk, Russia

In the article features of designing process of a technical device have been evaluated, in that case it is an installation of polymers recycling. Before design process there is a need in an analytical review of information and object's characteristics, which leads to determination of criteria. As a result there was created a 3D model of an installation with considering technical features, methods of creating a shape and modularity of an object.

Keywords: *design, project, installation, polymers, design project.*

1. Введение

При проектировании любого изделия необходимо придерживаться определенной стратегии по реализации этапов и объекта в целом. Это поможет избежать ошибок в проектировании и сэкономить временные и трудовые ресурсы, а также получить всесторонне проработанный результат.

В статье рассматриваются особенности дизайн-проектирования технического комплекса с учетом подходов к дизайн проектированию и методов формообразования. На основе этого обзора выявились критерии и ограничения, на котором основывалось проектирование.

2. Системный подход к проектированию

Установка по переработке полимеров является сложным техническим комплексом, в котором несколько разноплановых объектов связаны друг с другом электрическими сетями. Дизайн каждого из таких модулей должен быть разработан в отдельности и в последствии представлять единую модульную структуру. Изначально выявляются цели, задачи и основная проблема дизайн-проектирования. После определяются средства достижения цели, далее следуют методы, выбранные для реализации объекта проектирования. Существует несколько подходов к проектированию, которые выбираются для разработки в зависимости от назначения будущего объекта.

Для такого типа устройств самым рациональным для дизайн-проектирования является системный подход. На основе него выработалась структура разработки установки по переработке полимеров.

Системный подход – это подход, при котором любая система (объект) рассматривается как совокупность взаимосвязанных элементов (компонентов), имеющая связь с внешней средой и обратную связь. Согласно теории систем, каждый объект в процессе его исследования должен рассматриваться как большая и сложная система и, одновременно, как элемент более общей системы [1] .

Согласуясь с системным подходом, необходимо:

- определить основные составляющие, влияющие на проблемы;
- установить (найти) связи между ними;
- определить степень их взаимовлияния;
- очертить возможное количество направлений, в которых эти проблемы могут решаться;
- выработать критерии для выбора направлений;
- определить наиболее перспективные направления;
- выстроить приоритетную стратегию.

3. Методы формообразования

Методы формообразования также должны учитываться в процессе проектирования. Основными методами являются: инженерные, художественные, научные. Все они имеют разные подходы, свои плюсы и минусы. Для разработки собственного метода необходимо выбрать определенные особенности каждого из них и включить в процесс системного дизайн проектирования [2].

- инженерные методы формообразования:
  «плюсы» – наличие жесткого алгоритма проектной деятельности, безусловная реализация утилитарной функции, учет возможностей производства;
  «минусы» – эстетическая выразительность объекта лишь попутный продукт формообразования, узкая специализация.
- художественные методы формообразования:
  «плюсы» – безусловная эстетическая выразительность и содержательность формы;
  «минусы» – абстрагированность от системы производства, нет алгоритма проектной деятельности.
- научные методы формообразования:
  «плюсы» – логичность и комплексность решения задач формообразования, возможность автоматизации и оптимизации;

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

«минусы» – высокая трудоемкость, отсутствие проектной концепции [3].

Из технического формообразования это учет технической специфики конструкции объекта, его функциональность, конструктивность и технологичность. Аспекты художественного формообразования учитывают эстетическую выразительность и содержательность формы. Научный метод формообразования является наиболее подходящим, так как включает в себя комплексность решения задач, возможность оптимизации, что является ключевыми показателями в разработке дизайна установки по переработке полимеров. Однако в отличие от научного метода формообразования, в данном исследовании должна получиться итоговая проектная концепция.

Таким образом можно выделить следующие ступени развития проекта:

- Изучение технологии и принципа работы устройства.
- Анализ имеющихся аналогов.
- Определение основных критериев и ограничений для разработки.
- Определение важных особенностей конструкции.
- Разработка технической структуры и способ взаимодействия всех технических устройств между собой.
- Исследование возможностей формообразования с точки зрения художественной композиции.
- Адаптация формы под конструктивные решения.
- Анализ модульности и вариативность формы в связи с добавлением дополнительных модулей [4].

Эти шаги проектирования легли в основу разработки дизайн-проекта.

4. Критерии и задачи проектирования

Объект имеет множество технических параметров, которые описывают характер работы установки, процессы, протекающие внутри, и критерии для разработки дизайн-проекта.

Такие критерии можно определить по нескольким составляющим:

Экономическая эффективность:

- Рациональный подбор материалов и комплектующих.
- Оптимизация пространства и материалов.
- Возможность модернизации установки без изменения корпуса.

Технологический уровень:

- Внутреннее устройство и организация компонентов.
- Наглядное представление результатов исследования.
- Автоматический процесс управления возникновения плазмы.
- Доступ к компонентам установки для ремонта и замены
- Возможность автоматической откачки газа из реактора

Удобство эксплуатации:

- Эффективная панель управления.
- Эргономичное место оператора.
- Компактность.
- Удобство транспортировки.
- Легкость монтажа/демонтажа.
- Маркировка потенциально опасных объектов.

Внешний вид установки:

- Соответствие оптимальным размерам исходя из экономических характеристик.
- Единство стиля.
- Эргономичная форма цельного объекта и его компонентов.
- Компоновка частей установки по удобству и внешнему восприятию.
- Варианты доступа к внутренним компонентам.

В связи с вышеуказанными критериями можно выделить задачи для проектирования.

Задачи проектирования:

- Разработать логистику в компоновке установки по переработке полимеров для улучшения показателей эффективности.
- Провести эргономический анализ.
- Рассмотреть процесс формообразования каждой из частей установки в отдельности.
- Выявить оптимальные материалы для разных частей комплекса.
- Спроектировать внешний вид с учетом поставленных требований.
- Создать единый комплекс для проведения экспериментов с плазмой и показов результатов на выставках, отличающегося надежностью, модульностью и доступностью.

Для создания формы объекта, которая будет учитывать все вышеуказанные характеристики, необходимо в первую очередь разместить все комплектующие по удобству эксплуатации и возможностям расположения. Все объекты должны быть разбиты на подгруппы для создания для них отдельных модулей, которые в комплексе будут образовывать единую установку.

5. Дизайн-проектирование

С учетом рассмотренных ранее критериев принято решение создать основной блок и дополнительные модули. Основной конструкцией является центральный блок, в который вмонтирован компьютер, выдвижной рабочий стол, реактор, а внутри располагаются силовые приборы и электроника (Рисунок 2). Доступ к внутренним частям осуществляется с задней стороны. На лицевой стороне размещается экран, панель управления и стеклянная часть реактора. Также необходимо предусмотреть комфортное рабочее место: выдвижной стол будет фиксироваться в определенном положении, а в нижней части блока для размещения ног есть вырез.

Возможность присоединения модулей осуществляется с боковых сторон центрального блока, таким образом модули влияют на размер установки только по одной оси. Присоединение модулей производится через специальные крепления.

В связи с вышеуказанными критериями, требованиями и пожеланиями по разработке установки по переработке полимеров разработан эскиз, на основании которого создана 3D модель (Рисунок 1).

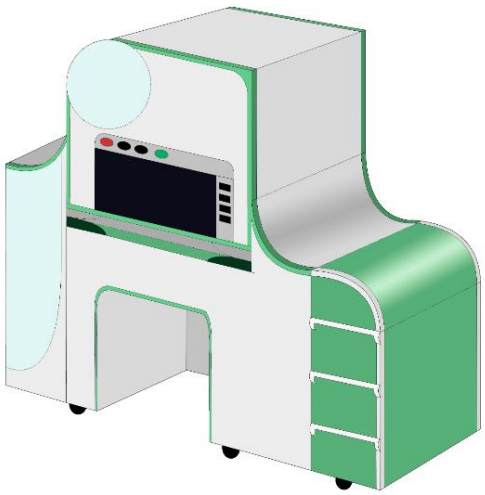


Рис. 1. Эскиз установки

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

В представленном эскизе основой является асимметрия и динамика, которые придают конструкции легкость и движение. В такой конструкции можно добавлять модули разных размеров, которые будут выглядеть общим целым со всей конструкцией.

В разработке единого комплекса нужно детально рассмотреть каждую составляющую в отдельности. В составляющие входят не только основные блоки, а также крепления конструкций, способы открывания дверок и дизайн ручек, соединительные механизмы и варианты внутреннего доступа. Отдельной разработкой является внутренняя организация реактора, а также комплектация силовых и регистрационных блоков внутри установки. За счет съемных панелей соединения кабелями можно производить по любой удобной конфигурации.

Дизайн основного блока не отличается от предварительно разработанного в эскизе, так как комплектация является удобной и функциональной. В процессе работы добавлена открывающаяся часть для доступа к экструдеру, куда будет засыпаться пластик. Модули под компрессор и шкафы также остались неизменными относительно разработанного ранее эскиза (Рисунок 2).

В модулях под газовые баллоны верхняя часть имеет скругление, которое изменено по отношению к финальному эскизу. Такое решение принято из-за увеличения объема модуля в следствие изменения типа скругления. В этом случае можно использовать баллоны большего объема.

Рис. 2. Взрыв схема с внутренней комплектацией, 1-Электроды, 2-Экструдер, 3-Газовые баллоны, 4-Силовой блок, осциллограф, фильтры, экран моноблока, 5-Компрессор

Для быстрой и удобной транспортировки объекта выбраны уже разработанные колеса из литого полиуретана с кронштейном из стального листа или нержавеющей стали. Данный вид является надежным при высоких механических нагрузках, которые в данной разработке обуславливаются весом всей установки.

Задние панели всех модулей и основного корпуса крепятся винтами к каркасу и снимаются в случае необходимости ремонта и замены оборудования [5]. У основного корпуса также снимаются боковые панели для более удобного доступа к компьютеру изнутри.

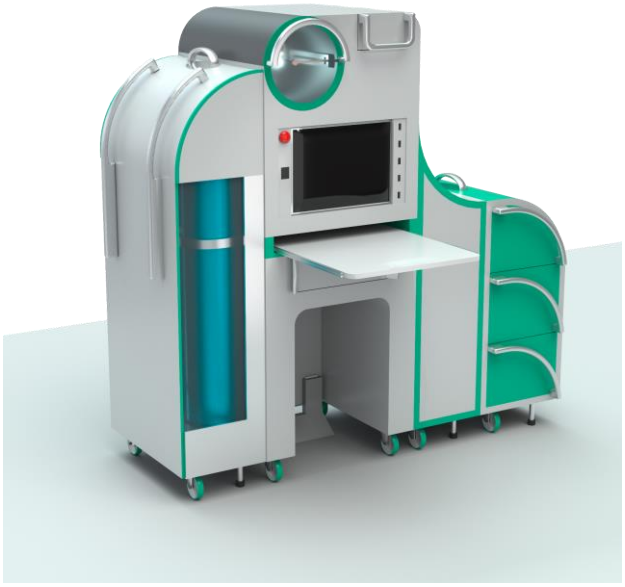


Рис. 3. Финальный рендер

Цветовое решение для установки выбрано в цветах: белый-зеленый-серый. Такие цвета лаконично сочетаются друг с другом и являются отсылкой к экологической направленности проекта.

6. Заключение

По итогам проектирования был разработан дизайн-проект установки по переработке полимеров и создана 3D модель (Рисунок 3). Конструкция является модульной, легко транспортируемой, эргономичной и учитывает все технические особенности установки.

7. Литература

[1] Михеева М.М., Основы системного дизайна, методическое указание, Москва:МГТУ им. Н.Э. Баумана, 2010 г.-59с.

[2] Минервин Г.Б. Дизайн. Иллюстрированный словарь-справочник / Г.Б. Минервин, В.Т. Шимко, А.В. Ефимов и др. – М.: Архитектура – С, 2004, С. 190

[3] Прокопьева И.А. Проблема выбора методов формообразования в дизайне, Архитектон: известия вузов №38, 2012

[4] Методика художественного конструирования. – М.: ВНИИТЭ, 1978. – 336 с.

[5] Сафьянникова, В. И., Вехтер Е. В.; Проблемы выбора материала при проектировании дизайн оболочки плазменной установки [Электронный ресурс]; под ред. С. С. Михалевич [и др.]. — Томск: Изд-во ТПУ, 2017. — [С. 343-344].

Об авторах

Вехтер Евгения Викторовна, к.п.н., доцент отделения автоматизации и робототехники Инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail vehter@tpu.ru.

Сафьянникова Виктория Игоревна, магистрант отделения автоматизации и робототехники Инженерной школы информационных технологий и робототехники Томского политехнического университета. E-mail victoriasafren@gmail.com.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

СОЗДАНИЕ РАЗВИВАЮЩЕГО ИНСТРУМЕНТА ДЛЯ ОБЕСПЕЧЕНИЯ БЕЗБАРЬЕРНОГО ИНКЛЮЗИВНОГО ОБУЧЕНИЯ

Сотников М.Е., Ризен Ю.С.

Национальный исследовательский Томский Политехнический университет, г.Томск, Россия

Инклюзивное образование существует уже очень давно. Инструменты для обучения детей с ограниченными возможностями здоровья и обычных детей по-прежнему остаются разными. Существующие аналоги не являются универсальными инструментами обучения для всех. Решить эту проблему можно введя игровой элемент, поскольку в игре человек добровольно соглашается с ее правилами и заинтересован в их соблюдении. Созданный инклюзивный инструмент (игрушка) развивает мелкую моторику и логическое, последовательное мышление, а также является образовательным инструментом для каждого участника игры. Игрушка не ориентирована на одну физическую особенность, что позволяет объединить людей с ОВЗ и людей без них, а также имеет перспективы для дальнейшей трансформации.

Ключевые слова: Инклюзивное образование, игра, ОВЗ, доступная среда, промышленный дизайн.

DESIGN A DEVELOPING TOOL TO PROVIDE A PERSISTENT INCLUSIVE TRAINING

Sotnikov M.E., Rizen Yu.S.

National Research Tomsk Polytechnic University, Tomsk, Russia

Inclusive education has existed for a very long time. Tools for teaching children with disabilities and ordinary children are still different. Existing analogues are not universal learning tools for all. To solve this problem, you can enter the game element, because in the game a person voluntarily agrees with its rules and is interested in their observance. The created inclusive tool (toy) develops fine motor skills and logical, consistent thinking, and is also an educational tool for each participant of the game. The toy is not oriented to one physical feature, which allows to unite people with HIA and people without them, and also has prospects for further transformation.

Keywords: Inclusive education, game, HIA, accessible environment, industrial design.

1. Введение

Инклюзивное образование — это совместное обучение и воспитание детей с ограниченными возможностями здоровья (ОВЗ) и детей, не имеющих таких ограничений. На данный момент, оно недостаточно эффективно, поскольку, преодолев барьер, разделяющий обычные и специальные школы, оно создало другой, не менее серьезный – разные инструменты обучения. Это создает серьезный дискомфорт, как для детей с ОВЗ, так и для обычных детей, а также объясняет актуальность создания инклюзивного инструмента для обучения. Предлагаемый для внедрения в учебную программу инструмент (игрушка) будет подходить людям с ограничениями и без, ликвидируя данный барьер.

2. Обзор существующих аналогов

Рассмотрев достоинства и недостатки существующих аналогов, было выявлено, что любой, названный инклюзивным, инструмент предназначен для конкретной особенности, игнорируя другие [1,3]. Это оказывает отталкивающий эффект для обычных детей, или детей с другими особенностями, полностью отбивает интерес к инструментам обучения и взаимодействия, делая человека не сосредоточенным, рассеянным. Такие игрушки вовсе нельзя назвать инклюзивными. Необходимо создать инклюзивный инструмент с учетом существующих методов воспитания и обучения, но при этом способный заинтересовать все группы людей.

3. Методы воспитания и обучения детей

Кроме привлекательности в процессе создания общего инструмента требуется еще и учитывать особенности взаимодействия людей с ОВЗ и людей без них. Поэтому было решено, что создаваемый инструмент будет являться игрушкой. В игре человек легче поддается обучению, поскольку он сам заинтересован в результате и сам направляет усилия на достижение цели. Разрабатываемая игрушка сосредоточит на себе внимание людей с различными физическими возможностями, а также будет многопользовательской, командной.

Для более точного представления о функциях, необходимых инклюзивной игрушке, был проведен анализ методов воспитания и обучения детей, в результате которого было установлено, что игра оказывает большее положительное влияние на развитие, и является самым эффективным средством усвоения знаний. Роль игры в воспитании и обучении очевидна, так как с ее помощью можно успешно корректировать, улучшать, развивать важнейшие *психические свойства*, а также *личностные качества* ребенка (ответственность, активность, творчество, самостоятельность и др.), *физические и творческие способности*. Особое значение в

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

коррекционно-воспитательной работе с детьми всех категорий имеют дидактические и развивающие игры, направленные на коррекцию и развитие основных *высших психических функций ребенка*, а также формирование таких качеств как: речь, внимание, память, воображение, мыслительная деятельность. К числу таких относятся головоломки и логические игры с кубиками. Поэтому, было решено, что разрабатываемая игрушка будет являться головоломкой.

Чаще всего, инклюзивные игры имеют только один вариант взаимодействия человека с игрушкой. Для повышения интереса к игровому процессу, она должна быть многовариантной, символичной, развивать разные интеллектуальные качества: память, особенно зрительную, пространственное представление, умение находить зависимости и закономерности, классифицировать и систематизировать материал, способность к комбинированию, т. е. умению создавать новые комбинации из имеющихся элементов, деталей, предметов, но также не должна ограничивать играющего [4]. Таким образом, для решения проблемы с заинтересованностью, игрушка должна иметь широкий функционал.

4. Проектирование

На этапе разработки решения, на основе изученной литературы и анализа существующих инклюзивных игрушек, было создано несколько эскизных идей (Рис.1), которые основываются на модулях (кубиках и карточках), соединяемых в единую конструкцию.

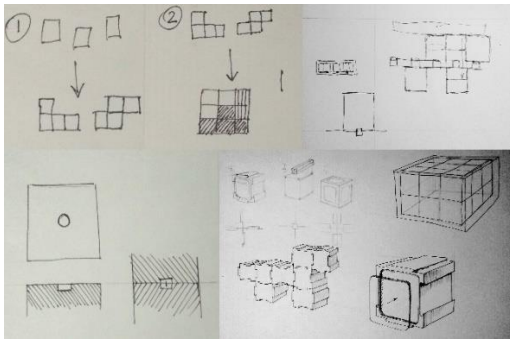


Рис.1. Эскизы

Рассмотрев соединительные механизмы, было решено осуществлять стыковку модулей магнитами, помещенными в них на каждую грань (рис.2). На фоне других, данный соединительный элемент выделяется легкостью внедрения в модули. Также, данный вид соединения поможет легче комбинировать модули людям, которым тяжело двигаться и/или у которых слабо развита моторика пальцев.

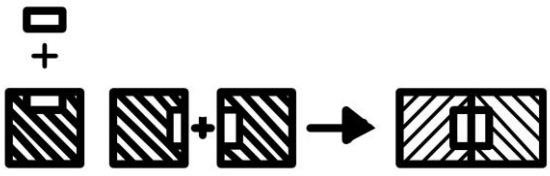


Рис. 2. Соединение магнитами.

На карточки также помещены магниты (рис.3), что позволяет быстро и легко менять рисунок на кубиках. Сами карточки были изготовлены из пластика.

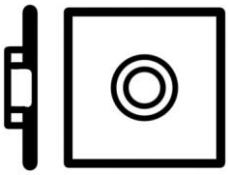


Рис.3. Карточка с магнитом

5. Реализация

Реализация игрушки началась с создания модулей (кубиков), которые вставляются в прямоугольную форму (рис.4). В форме кубики закрепляются магнитами с двух сторон, что позволяет повернуть кубики в форме. На остальные 4 закрепляются карточки с буквами и картинками. Эта идея направлена на изучение детьми алфавита, а также соединение рисунков, крепящихся на кубики, в единый сюжет (мама идет в сад – мама забирает ребенка из сада).

Рис. 4. Модули, вставленные в рамку

Далее была рассмотрена идея с теми же модулями, но соединяющимися в единую картину (рис.5). Таким образом можно развивать способность к комбинированию элементов.

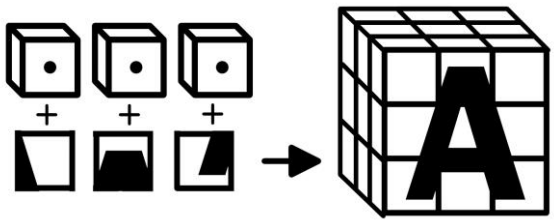


Рис. 5. Модули, соединяющиеся в картинку

Основной идеей стало последовательное соединение модулей, сначала в некие геометрические формы, а затем в единый куб (рис.6). Это помогает развивать последовательное мышление и зрительную память. Также можно создавать другие формы, что позволяет повысить интерес к игрушке и не ограничивать человека в их выборе.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

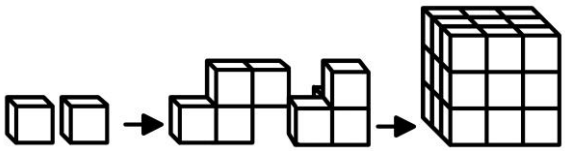


Рис.6. Модули с последовательным соединением в куб

Для слабо- или незрячих людей на модули были нанесены выпуклые элементы для создания рисунка (в виде геометрических фигур) на грани собранного куба (подобно рис.5) Для ясности зрячим выпуклые области были выделены цветом. Таким образом будет выстроено взаимодействие игроков с различными физическими возможностями.

Не менее важны внешний вид и материал, из которого будут изготовлены модули. Контакт с природными материалами, такими как: песок, вода, дерево - оказывает успокаивающее влияние и помогает сконцентрироваться, привлекает, развивает тактильные ощущения. В игрушках из дерева выделяют кубики их хвойных пород древесины. Поэтому в ходе работы модули были выполнены из сосны, так как этот материал не является токсичным, не вызывает дискомфорт и является безопасным.

С учетом эргономических требований были определены оптимальные размеры модулей для разных рук. Одним модулем будет являться куб, размером 40х40х40мм. С каждого ребра будет снята фаска в виде полувалика, радиусом в 5мм. В центре каждой грани будет выполнено отверстие, глубиной в 2мм, в которое помещается магнит радиусом в 2,5мм.

5. Эксперимент

После изготовления концептуального макета игрушка (рис.7) была передана для испытания группам разных возрастов, с ограниченными возможностями, без ограничений и смешанным.

Рис.7. Разработанная игрушка

В результате, более 75% испытуемых остались довольны игрушкой: контакт с игрушкой был приятен и детям, и взрослым; в смешанной группе дети согласовывали действия по сборке игрушки, была заметна легкость сбора модулей. Четверть испытуемых выразили беспокойство по поводу того, что магниты могут выпасть из пазов, что было в дальнейшем исправлено при помощи клея. Участники всех экспериментальных групп отметили, что игрушка оказалась интересной для каждого из них, что доказывает решение проблемы малой заинтересованности в инструментах для обучения у людей разных возрастов и особенностей развития. Также, треть испытуемых без особенностей развития выразили желание создать уменьшенную версию игрушки. По их мнению, это позволило бы сильнее развивать мелкую моторику, но в то же время они согласны с разработчиками в том, что такое решение сделает игрушку не инклюзивной.

6. Заключение

Таким образом, разработанный инструмент имеет перспективы для трансформации в приложение на электронные устройства с головоломками для развития логики, а также в уменьшенный не инклюзивный вариант для кинестетиков (люди, которые воспринимают большую часть информации через обоняние, осязание, с помощью движений и др.) как анти-стресс. Но в рамках данной работы рассматривается именно инклюзивность. Игрушка является модульной головоломкой, которая имеет как определенный алгоритм сбора куба, так и широкое количество других вариантов форм, что повышает заинтересованность всех участников игры. Она является инклюзивным инструментом для обучения и за счет многовариантности и развития разных интеллектуальных качеств интересна широкой возрастной аудитории. Головоломка позволяет развивать у ребенка логическое мышление и мелкую моторику пальцев кистей рук. Созданный инструмент не зациклен на одной физической особенности, чем позволяет объединить людей с различными видами ОВЗ и людей без них, тем самым устранив существующий барьер.

7. Литература

[1]. Основные виды развивающих игрушек для детей [Электронный ресурс] // Fainaidea. 2017. URL: http://www.fainaidea.com/jeto-interesno-znat/osnovnye-vidy-razvivayushhih-igrushek-dl-32453.html (Дата обращения: 29.11.2017).

[2]. Агостон Ж. Теория цвета и ее применение в искусстве и дизайне: Пер. С англ. - М.: Мир, 1982. - 184 с., ил.1.Ханников А.А Автомеханик/ А.А. Ханников.- 2010.- 384 с. (Дата обращения: 02.12.2017).

[3]. Виды развивающих игрушек [Электронный ресурс] // Minimum price. 2017. URL: http://www.minimum-price.ru/articles/vidu_razvivayushih_igrushek/ (Дата обращения: 02.12.2017).

[4]. Алмазова О.В, Брызгалова С.О, Тенкачева Т. Р. Проблемы интеграции и инклюзивного образования, Научная электронная библиотека «КИБЕРЛЕНИНКА» [Электронный ресурс] URL: https://cyberleninka.ru/ (Дата обращения: 17.02.2018).

Об авторах

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Автор: Сотников Михаил Евгеньевич, студент 3 курса ОАР ИШИТР ТПУ, e-mail: Soticov.mishutka@gmail.com

Руководитель: Ризен Юлия Сергеевна, Старший преподаватель ОАР ИШИТР ТПУ, e-mail: yulja_vit@tpu.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Выделение объекта в видеопотоке в различных условиях освещения

Тарбаев Б.Б.[1], Калайда В.Т.[1]
tarbaevbb@gmail.com|kvt@iao.ru
[1] Национальный исследовательский Томский государственный университет, Томск, Россия

Целью данной работы является выделение объекта в видеопотоке в режиме реального времени. Для достижения данной цели предлагается модификация SIFT, в которой уменьшается вычислительная сложность гауссового размытия. Предлагается аппроксимировать этот алгоритм с помощью Box linear filter.

Ключевые слова: *видеопоток, выделение объекта, фильтр Гаусса.*

Selecting an object in the video stream under different lighting conditions

Tarbaev B.B.[1], Kalayda V.T.[1]
tarbaevbb@gmail.com|kvt@iao.ru
[1]the National Research Tomsk State University, Tomsk, Russia

The purpose of this work is to select an object in the video stream in real time. To achieve this goal, a modification of SIFT is proposed, in which the computational complexity of the Gaussian blur is reduced. It is suggested to approximate this algorithm using the Box linear filter.

Keywords: *video stream, object tracking, Gaussian filter.*

1. Введение

В связи с широким применением систем видеонаблюдения возникает проблема повышения эффективности обработки и использования «огромных» объёмов информации предоставляемыми ими. Полученные данные также включают различные шумы (освещённость, резкость, контрастность и т.д.). Существующие в настоящее время подходы к решению задачи выделения полезной информации в видеопотоке обладают двумя недостатками: невысокое быстродействие и невысокая для практического применения точность [14].

Среди всех задач выделения информации из видеопотока наибольшее практическое значение имеет задача поиска заданного объекта в кадре в условиях отмеченных помех, которая естественным образом возникает в самых разных областях: от компьютерного зрения до задач обеспечения безопасности.

При поиске объекта в видеопотоке необходимо решать две задачи:

1. быстрый поиск областей, претендующих на дальнейшее сравнение, схожих с образцом;
2. оценка мера близости между выделенной областью и образцом.

Исходя из вышесказанного, целью данной работы является разработка алгоритма и программная реализация системы поиска объекта в видеопотоке.

2. Метод определения ключевых точек

На первом этапе определяются инвариантные характеристики двух соседних изображений путём выделения особых точек и их дескрипторов. Ключевая точка (особая точка) – это точка изображения, окрестность которой можно отличить от окрестности любой другой точки изображения в некоторой другой окрестности ключевой точки.

Дескриптор – это идентификатор ключевой точки, выделяющий её из остального множества ключевых точек.

Для обнаружения особых точек используется модифицированный алгоритм преобразования масштабно-инвариантных характеристик (англ. Scale Invariant Feature Transform, SIFT). Алгоритм SIFT показывает лучшие результаты обнаружения особых точек по качеству чёткости и контрастности цвета и освещения изображения, чем другие алгоритмы обнаружения точек [6,16]. Недостатком SIFT является его вычислительная сложность, что ограничивает его применение в режиме постобработки [3].

Одним из способов преодоления указанных недостатков является разработка эффективных численных методов поиска точечных особенностей, расчета дескрипторов, вычисления расстояний между дескрипторами, поиска ближайших дескрипторов. К таким методам относятся методы приближенного анализа дифференциальной структуры изображений на основе вейвлетов и интегральных изображений [1], методы индексирования дескрипторов [2,10,19], методы понижения размерности дескрипторов [7].

В работах, посвященных сравнительному анализу методов-детекторов [5,11,12], отмечается, что используемые детекторы имеют относительно невысокую степень воспроизводимости, т.е. точечные особенности, найденные в эталонном изображении, не всегда могут быть найдены в преобразованном изображении, хотя качество сопоставления изображений остается высоким. При этом число невоспроизводимых точечных особенностей может достигать 40% от общего числа найденных точечных особенностей. Это наблюдение говорит о том, что, с точки зрения задачи сопоставления изображений, невоспроизводимые точечные особенности являются избыточными данными. В то же время в известных алгоритмах сопоставления изображений все точечные особенности обрабатываются единообразно. Таким образом, обработка избыточных данных представляет собой источник дополнительной вычислительной сложности, исключение которого не повлияет на результат сопоставления изображений.

В данной работе алгоритм SIFT будет рассмотрен с точки зрения теории алгоритмов.

Алгоритм SIFT [4,6,8,15,18] состоит из четырех главных этапов:

1) Определение максимума масштабируемого пространства: первый этап вычислений. Он реализуется посредством вычисления функции разности Гауссиана для определения потенциальных ключевых точек

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

инвариантных к масштабированию [13] и изменению ориентации.

Инвариантность относительно масштаба достигается за счет нахождения ключевых точек для исходного изображения, взятого в разных масштабах. Для этого строится пирамида гауссианов: все масштабируемое пространство разбивается на некоторые участки — октавы, причем часть масштабируемого пространства, занимаемого следующей октавой, в два раза больше части, занимаемой предыдущей. К тому же, при переходе от одной октавы к другой размеры изображения уменьшаются вдвое. Естественно, что каждая октава охватывает бесконечное множество гауссианов изображения, поэтому строится только некоторое их количество N, с определенным шагом по радиусу размытия. С тем же шагом достраиваются два дополнительных гауссиана (всего получается $N+2$), выходящие за пределы октавы.

Гауссианом является изображение (рис. 1, слева):

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y),$$
$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2},$$

где L – значение гауссиана в точке с координатами (x,y),

σ – радиус размытия,

G – гауссово ядро,

I – значение яркости исходного изображения,

$*$ – операция свертки.

Разностью гауссианов называют изображение, полученное путем попиксельного вычитания одного гауссиана исходного изображения из гауссиана с другим радиусом размытия (рис. 1, справа):

$$D(x,y,\sigma) = \big(G(x,y,k\sigma) - G(x,y,\sigma)\big) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma),$$

где $k = 2^{\frac{1}{s}}$ обычно имеет значение $\sqrt{2}$,

$s \geq 1$ и является целым числом.

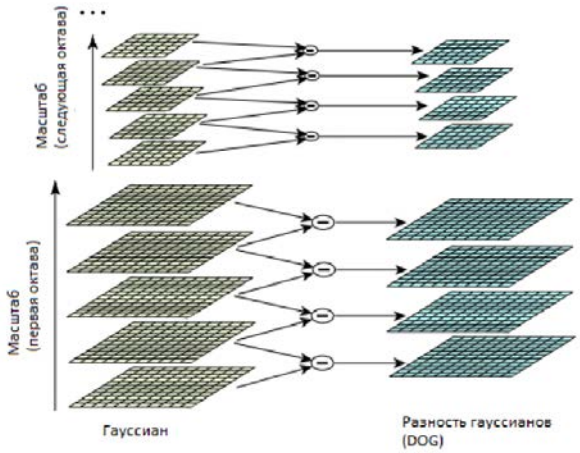


Рис. 1. Пирамида Гауссианов

2) Определение ключевых точек. Ключевые точки определяются на основе меры их стабильности.

Будем считать точку ключевой, если она является локальным экстремумом разности гауссианов.

Для поиска экстремумов используется метод, схематично изображенный на рис. 2. Если значение разности гауссианов в точке, помеченной X, больше (меньше) всех значений в точках, помеченных окружностями, то эта точка считается точкой экстремума. Для того, чтобы обнаружить локальные максимумы и минимумы $D(x,y,\sigma)$ каждая точка сравнивается с ее ближайшими соседями в текущем изображении и с девятью соседями в масштабах выше и ниже.

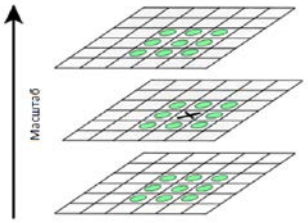


Рис. 2. Поиск экстремумов разности Гауссианов

Следующий шаг по поиску ключевых точек – проверка пригодности точки экстремума на роль ключевой. Сначала определяются координаты ключевой точки с субпискельной точностью. Потом происходит удаление точек с малым контрастом и точек, находящихся на границе объектов.

3) Определение ориентации. Для каждой ключевой точки может быть определено одно или несколько направлений, основанных на локальных направлениях градиентов. Все последующие операции выполняются на изображении, которое было преобразовано в соответствии с ориентацией и масштабом каждой ключевой точки, обеспечивая тем самым инвариантность к этим преобразованиям.

Направление, соответствующее характерной точке, вычисляется исходя из направлений градиентов точек, соседних с характерной. Все вычисления градиентов приводятся на изображении в пирамиде гауссианов, с масштабом наиболее близким к масштабу характерной точки. Величина и направление градиента в точке (x,y) вычисляются по формулам.

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2}$$
$$\theta(x,y) = \tan^{-1}\big((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))\big),$$

где m – величина градиента,

θ – его направление.

4) Построение дескрипторов.

Как и направление ключевой точки, дескриптор вычисляется на гауссиане, ближайшей по масштабу к характерной точке, и исходя из градиентов некотором окне характерной точки. Перед вычислением дескриптора это окно поворачивают на угол направления ключевой точки, чем достигается инвариантность относительно поворота.

3. Оценка сложности алгоритма

Пусть размер изображения $N \times M$. Обозначим за α долю пикселей, которая окажется экстремумами, выразим ее как дробь $0 < \alpha < 1$. Таким образом αNM – количество экстремумов. Только небольшая часть из этих экстремумов будет квалифицирована как ключевые точки. Обозначим их как $0 < \beta < 1$. Итого имеем $\alpha\beta NM$ ключевых точек. Далее проверяем не были ли пропущены ключевые точки, обозначим эту долю пикселей γ. Если такие точки нашлись, то они становятся γNM новых ключевых точек, таким образом всего у нас $(\alpha\beta + \gamma)NM$ ключевых точек. Далее эти точки будут преобразованы в дескрипторы. Их число будет $(\alpha\beta + \gamma)NM$. α, β, γ зависят от рассматриваемого изображения [17].

В таблице 1 представлена вычислительная сложность и число операций каждого из этапов алгоритма.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Таблица 1. Вычислительная сложность операций

Этап	Сложность	Число операций
Гауссовое размытие	$0(NMw^2s)$	$4NMw^2s$
Разница гауссиан	$0(sNM)$	$4NMs$
Обнаружение экстремумов	$0(sNM)$	$104sNM$
Обнаружение ключевых точек	$0(\alpha sNM)$	$100s\alpha NM$
Назначение ориентации	$0(sNM(1-\alpha\beta))$	$48sNM$
Генерация дескриптора	$0(x^2NM(\alpha\beta+\gamma))$	$1520x^2(\alpha\beta+\gamma)NM$

Как видно из таблицы, наибольшей вычислительной сложностью обладает операция получения гауссового размытия. Таким образом следующим шагом является уменьшение вычислительной сложности алгоритма.

4. Уменьшение сложности гауссового размытия

Пусть f – изображение, g – ядро гаусса, тогда фильтр Гаусса можно представить, как:

$$Gb(i,j) = \sum_{y=j-r}^{j+r}\sum_{x=i-r}^{i+r} f(x,y) * g(x,y).$$

Алгоритм Box linear filter (BLF) позволяет аппроксимировать размытие по гауссу [9]. Обозначим половину размера поля как r – радиус окна, тогда постоянное значение веса равно $\frac{1}{(2r)^2}$. Здесь вес (ядро) постоянный и лежит в квадрате (поле). Особенностью алгоритма является то, что повторное применение BLF будет аппроксимировать гауссовое размытие. Определим BLF как:

$$Bb(i,j) = \sum_{y=j-r}^{j+r}\sum_{x=i-r}^{i+r} f(x,y)/(2r)^2.$$

Сложность этого алгоритма по-прежнему является квадратичной $0(NMr^2)$. Но в данном случае нам не нужно вычислять новые веса. Для аппроксимации гауссового размытия, данный алгоритм обычно применяют 3 раза.

Определим горизонтальное размытие и глобальное размытие:

$$b_h(i,j) = \sum_{x=i-br}^{i+br} f(i,y)/(2r),$$
$$b_g(i,j) = \sum_{y=j-br}^{j+br} b_h(x,j)/(2r),$$

$$b_g(i,j) = \sum_{y=j-r}^{j+r}\left(\sum_{x=i-r}^{i+r} f(x,y)/(2r)\right)/(2r) = \sum_{y=j-r}^{j+r}\sum_{x=i-r}^{i+r} f(x,y)/(2r)^2.$$

Отсюда следует, что глобальный фильтр является фильтром BLF. Полное размытие как и горизонтальное имеют сложность $0(NMr)$, поэтому фильтр BLF тоже имеет сложность $0(NMr)$.

Попытаемся вычислить одномерное размытие еще быстрее. Если мы хотим вычислить горизонтальное размытие, то мы вычисляем: $b_h(i,j)$, $b_h(i+1,j)$, $b_h(i+2,j)$ и т.д. Но эти значения почти одинаковы. Единственное различие в том, что они равны в самом левом и правом значении:

$$b_h(i+1,j) = b_h(i,j) + f(i+r+1,j) - f(i-r,j).$$

Таким образом, мы можем создать буфер, в который мы поместим значение самой левой ячейки. Затем мы вычислим следующие значения, просто изменив предыдущее. Сложность этого действия будет $0(NM)$. Но его надо выполнить дважды, чтобы получить размытие ящика, которое в свою очередь надо выполнить 3 раза, чтобы получить гауссовое размытие. Таким образом, сложность алгоритма равна $6 \cdot 0(NM)$.

5. Результаты

В таблице 2 приведены результаты оценки времени выполнения и вычислительная сложность алгоритмов для изображения 800x200 пикселей.

Таблица 2. Время выполнения и вычислительная сложность алгоритмов

Алгоритм	Время: мс, r=5	Время: мс, r=10	Вычисл-ая сложность
Фильтр Гаусса	7077	27021	$0(NMr^2)$
Box filter	586	2437	$0(NMr^2)$
Box filter оптимизация 1	230	435	$0(NMr)$
Box filter оптимизация 2	32	34	$0(NM)$

6. Заключение

В статье представлен метод для оптимального вычисления размытия по Гауссу для использования в алгоритме SIFT. Проведённые вычислительные эксперименты подтверждают эффективность и устойчивость работы алгоритма. Различия изображений, полученных стандартным фильтром Гаусса и его аппроксимацией, не превышает 0,2%.

В дальнейшем планируется: оптимизация работы программы для ускорения обработки данных, проведения тестирования модифицированного SIFT.

7. Литература

[1] Кудряшов А. П. Извлечение и сопоставление точечных особенностей. Электронный научный журнал Исследовано в России, т.10, 2007. http://masters.donntu.org/2012/iii/chigarev/library/article1.htm

[2] Пименов В. Ю. Метод поиска нечетких дубликатов изображений на основе выявления точечных особенностей. Труды РОМИП2007-2008, с. 145-158. СПб.:НУ ЦСИ, 2008.

[3] Со С.А. Реализация цифровых технологий обработки изображений в адаптивных системах управления сварочными процессами: диссертация кандидата техн. наук. МГТУ им. Н.Э. Баумана, 2015, 145 с.

[4] Chu E., Hsu E., Yu S. Image-Guided Tours: Fast-Approximated SIFT with U-SURF Features, Stanford University. 2007. pp. 6-12.

[5] Frintrop S. The High Repeatability of Salient Regions. Proc. Of ECCV Workshop «Vision in Action: Efficient Strategies for Cognitive Agents in Complex Environment». 2008.

[6] Juan L., Gwun O. A Comparison of SIFT, PCA-SIFT and SURF // International Journal of Image Processing (IJIP). 2009. Vol. 3. No. 4. PP. 143-15.

[7] Ke Y., Suthanakar R. PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. In Computer Vision and Pattern Recognition, volume 2, pages 506-513, 2004.

[8] Khan N., McCane B., Wyvill G. SIFT and SURF performance evaluation against various image deformations on benchmark dataset // International conference on digital image computing: techniques and applications. 2011. pp. 109-114.

[9] Kovesi P. Fast Almost-Gaussian Filtering // Centre for Exploration Targeting School of Earth and Environment.

[10] Lowe D. Distinctive Image Features from Scale-Invariant Keypoints. In International Journal of Computer Vision, volume 60, pages 91-110, 2004.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

[11] Mikolajczyk K., Schmidt C. A Performance Evaluation of LocalDescriptors. IEEE Trans. Pattern Anal. Mach. Intell., volume 27 (10), pages 1615-1630. 2005.
[12] Mikolajczyk K. et al. A Comparison of Affine Region Detectors. Int. J. of Comput. Vis.volume 65 (1-2), pages 43-72, 2005.
[13] Mikolajczyk K., Tuytelaars T., Schmid C., Zisserman A., Matasm J., Schaffalitzkym F., Kadir T., Van Gool L. A comparison of affine region detectors // Intelligent journal of computer vision, Vol. 65 (1-2), 2005. pp. 43-72.
[14] Prajapati D., Galiyawala H. J. A Review on Moving Object Detection and Tracking // International Journal of Computer Application (2250-1797) Volume 5– No. 3, April 2015 PP. 168-175.
[15] Tang C.Y., Wu Y.L., Hor M.K. Modified sift descriptor for image matching under interference // Machine learning and cybernetics, international conference on, Vol. 6, July 2008. pp. 3294 – 3300.
[16] Valgren C., Lilienthal A.J. SIFT, SURF & Seasons: Appearance-based Long-term Localization in Outdoor Environments // Robotics and Autonomous Systems. 2010. Vol. 58. No. 2. PP. 149–156. DOI: http://dx.doi.org/10.1016/j.robot.2009.09.010.
[17] Vinukonda P. A study of the scale-invariant feature transform on a parallel pipeline // B.Tech., JNTU University, 2007 May 2011.
[18] Zhang Z. A Flexible new technique for camera calibration. technical report MSRTR // Microsoft Research. December 1998. pp. 98-71.
[19] Zhao W.-L., Ngo C.-W., Tan H.-K., Wu X. Near-Duplicate Keypoint Identification with Interest Point Matching and Pattern Learning. In IEEE Transactions on Multimedia, volume 9(5), pages 1037-1048, 2007.

Об авторах

Тарбаев Баир Баясхаланович, аспирант кафедры оптико-электронных систем и дистанционного зондирования радиофизического факультета Томского государственного университета. Его e-mail tarbaevbb@gmail.com.

Калайда Владимир Тимофеевич, д.т.н., профессор кафедры оптико-электронных систем и дистанционного зондирования радиофизического факультета Томского государственного университета. Его e-mail kvt@iao.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Применение световых панно для преодоления депрессии

Хмелевский Ю.П., Серяков В.А., Мамонтов Г.Я., Хруль Т.С.
hmelevskiy@tpu.ru, seryakov@tpu.ru, mgy@tpu.ru, hts@tpu.ru
ТПУ, Томск, Россия

Актуальность исследования обусловлена потребностью поиска новых методов преодоления депрессии с помощью применения световых панно, изготавливаемых на основе LED-технологии. Целью работы является обоснование метода положительного воздействия яркого света и цветовых изображений, соответствующий темпераменту человека, проходящего сеанс. Выявление влияния света и цветовых сочетаний фотоизображения на улучшение эмоционального состояния человека.

Ключевые слова: *световое панно, яркий свет, цветовые изображения, депрессия, комфорт, темперамент человека.*

.

Application of light panels for overcoming a depression

Khmelevsky Y.P., Seryakov V.A., Mamontov G.Ya., Khrul T.S
hmelevskiy@tpu.ru, seryakov@tpu.ru, mgy@tpu.ru, hts@tpu.ru
TPU, Tomsk, Russia

The relevance of a research is caused by requirement of search of new methods of overcoming a depression by means of the application of light panels made to a basis on LED technology. The purpose of work is justification of a method of positive impact of bright light and the color images suitable to temperament of the person passing a session. Identification of influence of light and color combinations of the facsimile on improvement of an emotional condition of the person.

Key words*: light panel, bright light, color images, depression, comfort, temperament of people.*

1. Введение

По данным департамента психического здоровья и зависимости от веществ, г. Женева в мире 450 миллионов человек страдают психическим или поведенческим расстройством. Ежегодно около 1 миллиона человек совершают самоубийство. Четыре из шести ведущих причин к этому шагу связаны с инвалидностью и с нейропсихиатрическими расстройствами, такими как депрессия или расстройства основной причиной которых является злоупотребление алкоголем [1].

Психотерапевты и психологи всего мира пытаются разобраться и предложить эффективные методы решения данной проблемы. Решение этой очень сложной задачи требует широких исследований и комплексного подхода.

Одним из методов способствующим решению представленной проблемы может служить светотерапия и цветотерапия. Известно, [2] что уровень серотонина так называемого «гормона счастья» значительно повышается в моменты повышения настроения и падает во время депрессии. Около 5-10% серотонина синтезируется шишковидной железой из жизненно необходимой аминокислоты триптофана. Для его производства абсолютно необходим солнечный свет, именно поэтому в солнечные дни наше настроение на высоте. Этим же процессом можно объяснить и общеизвестную зимнюю депрессию. В 1980 году американский психиатр Норман Розенталь и его команда применяли методы [3] лечения с помощью яркого света которые помогли пациентам успешно справиться с депрессией. В 1984 году результаты были опубликованы официально, в них описывались сезонное аффективное расстройство (САР) и впервые светотерапия как эффективный метод лечения рис. 1.

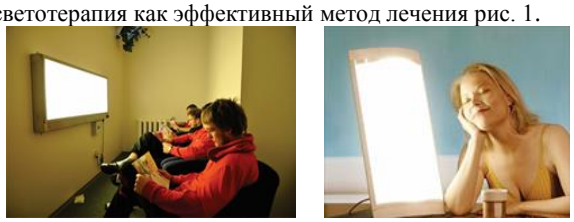


Рис. 1. Сеансы светотерапии.

2. Теоретические основы

Для решения проблемы нехватки солнечной энергии используются специальные лампы, излучение которых очень близко к солнечному свету как по яркости, так и спектральному составу. Индекс цветопередачи показывает насколько в свете лампы видны другие цвета, измерение его производят в процентном соотношении от 0% до 100%. Показатель светопередачи, комфортный для человека должен быть более 80%. Применение светодиодов позволяет решить, как проблему яркости, так и спектрального состава света, на рис. 2 приведены сравнительные характеристики спектров солнечного света и светодиодных ламп.

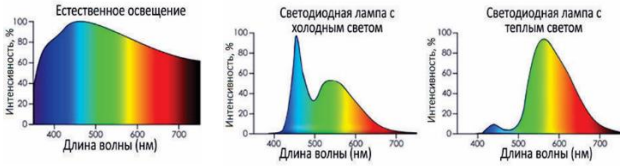


Рис. 2. Сравнительные характеристики различных источников света.

Светодиодные лампы с теплым светом и требуемым уровнем яркости позволяют имитировать естественный свет солнца и наиболее подходят для использования в световых панно.

3. Исследовательская часть

Для создания освещенной плоскости предполагается использование светодиодной матрицы. Кроме функции положительного светового воздействия на человека предлагается добавить к проектируемому световому панно функцию цветотерапии, то есть возможности установки в качестве полупрозрачного цветового фильтра панелей с различными изображениями по цветовой гамме.

В ходе исследования и анализа, данных опроса были получены следящие результаты. В исследовании приняли участие 100 человек проживающих в городе Томске, из которых 65 % женщины и 35 % мужчины. Средний возраст

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

мужчин от 18 до 35 лет, средний возраст женщин от 18 до 45 лет.

Процентное соотношение мужчин по типу темперамента:

- меланхолик, сангвиник, холерик – 26%,
- флегматик – 22%.

Процентное соотношение женщин по типу темперамента:

- меланхолик - 20%,
- флегматик – 16%.
- сангвиник – 24%;
- холерик – 40%.

Как представлено в таблице №1 мужчинам всех типов темперамента психологически комфортно наблюдать фотографии, на которых изображено летнее время года. Иная же тенденция прослеживается у опрашиваемых женщин. Женщины с типом темперамента меланхолик и холерик отдают предпочтение фотографиям с изображениями весеннего времени года в отличии от женщин с типом темперамента флегматик и сангвиник, выбравших чаще виды летнего времени.

Таблица №1.Зависимость выбора времени года от типа темперамента

Тип темперамента человека	Изображение времени года, %			
	Весна	Лето	Осень	Зима
мужчина				
меланхолик	22	40	22	16
холерик	23	32	21	24
флегматик	26	32	24	18
сангвиник	24	35	18	23
женщина				
меланхолик	43	33	10	14
холерик	35	30	20	15
флегматик	26	32	24	18
сангвиник	27	33	27	13

В таблице №2 в ходе исследования были определены основные цвета, соответствующие типу темперамента. Как видно из таблицы исследование подтвердило уже существующие теории соответствия цвета типу темперамента:

- меланхолики выбирают фиолетовый цвет и некоторые оттенки сине-зеленого, голубого;
- холерики, люди, которые быстро возбуждаются и теряют над собой контроль - зеленый или сине-зеленый, особенно красный, но никогда не выбирают белый.
- флегматики выбирают между белым, цветом слоновой кости или зелёным.
- сангвиники уверены в себе, живут ярко и выбирают красно-оранжевый, терракотовый, желтый [4].

Красный цвет повышает давление и усиливает напряжение мышц, ускоряет пульс, возбуждает. Желтый придаёт равновесие эмоциям, а синий, особенно фиолетовый, заметно уменьшают давление крови, частоту пульса и дыхание, тормозят и даже подавляют эмоции. Зелёный слегка снижает давление и уравновешивает эмоции. Аналогичные результаты были получены в работе [5].

Таблица №2. Основные цвета каждого типа темперамента

Меланхолик		Холерик		Флегматик		Сангвиник	
муж	жен	муж	жен	муж	жен	муж	жен

Наиболее повторяющееся изображение для всех типов темперамента при проведении свето- и цвето- терапии в качестве изображения на панели светового панно представлено на рис.3.

Рис. 3. Изображение лета, выбор большинства мужчин.

Представленному изображению лета на рисунке 3 отдали предпочтение наибольшее число мужчин: меланхоликов – 57 %, флегматиков – 67%, сангвиников – 63 %, холерик – 50 %.

На рисунке 4 представлено изображение лета, ему отдали предпочтение наибольшее число женщин:

- меланхоликов – 70,00 %, флегматиков – 57,14 %, сангвиников – 58,33 %, холерик – 60,00 %.

Рис. 4. Изображение лета, выбор большинства женщин.

Примеры набора изображений для светотерапии приведены на рис. 5.

меланхолик – синий цвет

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

холерик – красный цвет

флегматик – зеленый цвет

сангвиник – желтый

Рис. 5. Изображения для светотерапии.

Заключение

Опираясь на результаты исследования можно сделать вывод, что изображения с видами зелено-голубой глади океана, зеленой тропической растительности и голубого небосклона наиболее предпочтительны для людей всех типов темперамента в городе Томске из которых 65% женщин и 35% мужчин. Средний возраст мужчин от 18 до 35 лет, средний возраст женщин от 18 до 45 лет. На основе этих данных можно сказать что изображения с видами океана и тропической растительности могут быть наиболее востребованы для использования в качестве цветотерапии на панели светового панно. Кроме того, опираясь на данные исследования изображения видов весны у женщин меланхоликов и холериков являлись наиболее популярными в следствии этого для проведения сеансов цветотерапии у данных групп рекомендуется использование на панелях светового панно использовать фотографии весны.

Представленное световое панно предполагается использовать в жилых помещениях и специализированных кабинетах светотерапии. Дизайн съемных панелей может быть разнообразен для возможности размещения их в помещения отличающихся по стилю рис. 6 [6].

Рис. 6. Интерьерное решение с использованием светильника для арт-терапии.

Кроме перечисленных помещений сеансы светотерапии и цветотерапии предлагается использовать комплексно с физическими занятия на реабилитационных тренажёрах. Люди, перенесшие инсульт, инфаркт, тяжелые травмы и операции часто находятся в состоянии депрессии, это состояние не способствует эффективной реабилитации. Для улучшения эмоционального состояния данную задачу предлагается решить с помощью светового панно [7].

Список использованной литературы:

[1] Department of Mental Health and Substance Dependence. Инвестирование в душевное здоровье [Электронный ресурс] - Режим доступа: http: //www.who.int/mental_health/media/investing_mnh.pdf

[2] Гормон счастья: 95% серотонина находится в кишечнике [Электронный ресурс] - Режим доступа: https://econet.ru/articles/125549-gormon-schastya-95-serotonina-nahoditsya-v-kishechnike

[3] Rosenthal NE, Sack DA, Gillin JC, Lewy AJ, Goodwin FK, Davenport Y, Mueller PS, Newsome DA, Wehr TA (1984) Seasonal affective disorder: A description of the syndrome and preliminary find-ings with light therapy Archives of General Psychiatry, 41: 72-80

[4] Ватерман Г. Дизайн вашей квартиры / Г. Ватерман перевод с нем. Т.А. Набатникова [и др.]. – из-во «Фалькен» (Германия), 1992. – с. 125.

[5] V. A. Seryakov, Yu. P. Khmelevsky, G. Ya. Mamontov The design features of the body of the portable electrocardiograph "ECG-EXPRESS" [Electronic resource] /IOP Conference Series: Materials Science and Engineering. - 2015. - Vol. 93: Modern Technique and Technologies (MTT'2015). - [012076, 7 p.].

[6] Khmelevsky Yu. P., Seryakov V.A., Mamontov G.Ya., Environment as a basis for the design of advertising structures by forming [Electronic resource] / [et al.] // Journal of Physics: Conference Series. - 2017. - Vol. 803: Information Technologies in Business and Industry (ITBI2016). — [012068, 5 p.].

[7] Кухта М.С. Куманин В.И., Соколова М.Л., Гольдшмидт М.Г. Промышленный дизайн: учебник / – Томск: Изд-во ТПУ, 2013.

Об авторах

Хмелевский Юрий Петрович старший преподаватель отделения автоматизации и робототехники Инженерной школы информационных технологий и робототехники Томского Политехнического университета

Серяков Вадим Александрович, кандидат технических наук, доцент отделения автоматизации и робототехники Инженерной школы информационных технологий и робототехники Томского Политехнического университета

Мамонтов Геннадий Яковлевич, доктор физико-математических наук, профессор отделения автоматизации

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

и робототехники Инженерной школы информационных технологий и робототехники Томского Политехнического университета

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Решение задач эргономики с помощью трехмерной модели рабочей зоны

Л.Н. Черепова, А.А. Черепов, А.В. Шкляр
lni1@tpu.ru|andrcherepov@gmail.com|shklyarav@tpu.ru
Томский Политехнический Университет, Томск, Россия

В работе рассматриваются особенности проведения соматографического анализа и составления соматографических моделей в двумерном и трехмерном пространстве. Выявляется потребность использования средств трехмерного моделирования для построения антропометрически точных моделей с целью проведения соматографического анализа на примере рабочего места хирурга. Описывается способ построения трехмерной соматографической модели и её применение для построения трехмерной модели рабочей зоны.

Ключевые слова: *эргономика, соматография, соматографическая схема, 3d моделирование, антропометрический анализ, антропометрические характеристики*

Solving of ergonomic problems using the three-dimensional model of the working area

L.N. Cherepova, A.A. Cherepov, A.V. Shklyar
lni1@tpu.ru|andrcherepov@gmail.com|shklyarav@tpu.ru
Tomsk Polytechnic University, Tomsk, Russia

The paper deals with the features of performing somatographic analysis and composing of somatographic models in two-dimensional and three-dimensional space. The need to use 3D modeling tools to construct anthropometrically accurate models is identified with the purpose of performing somatographic analysis using the example of a surgeon's workplace. A method for constructing a three-dimensional somatographic model and its application for constructing a three-dimensional model of a working zone are described.

Keywords: *ergonomics, somatography, somatographic scheme, 3d modeling, anthropometric analysis, anthropometric characteristics*

1. Введение

При решении эргономических задач по определению доступной моторной зоны используются соматографические модели [9,10]. Соматографическая модель среды – это модель, использующая схематические изображения человеческого тела с соблюдением анатомических пропорций между его частями, для определения соотношения размеров человеческого тела и среды. Соматографические модели находят применение в процессе проектирования дизайна рабочего места, что обуславливается следующими проблемами, которые решаются с их помощью:

1. Необходимостью проверки соотношения человеческой фигуры, размеров и формы проектируемого рабочего места.
2. Необходимостью определения границ моторного пространства человека, трудящегося на рабочем месте.
3. Необходимостью определения обзора на рабочем месте.

В данной статье рабочим местом называется область пространства, в которой располагается работник в процессе осуществления производственных операций, а также те области пространства, которые могут быть задействованы или задействуются работником в этом процессе.

В процессе проектирования рабочего места также применяются трехмерные графические пакеты для создания трехмерных эскизов и дальнейшей визуализации предполагаемого дизайна. Под моторным пространством в контексте проектирования дизайна рабочего места понимается то пространство рабочего места, в котором будут осуществляться двигательные действия человека с актуальными для деятельности средствами труда. Существующие двухмерные соматографические модели не позволяют в полной мере оценить моторное пространство при создании трехмерных эскизов, поскольку показывают лишь его отображение на плоскость, в которой проводится эскизирование рабочего места.

Для облегчения проектирования дизайна рабочего места в пакетах трехмерной графики предлагается создать трехмерную модель, позволяющую отобразить доступное работнику моторное пространство на рабочем месте.

Для построения такой модели необходимо знать основные положения человеческого тела, возникающие при работе за рабочим местом, возможные движения в суставах, участвующих в построении моторной зоны, и иметь данные о распределении антропометрических величин. Это позволит определить общие границы моторной зоны при различных положениях человеческого тела.

В работе приводятся способы определения основных рабочих позиций, описывается способ построения трехмерной модели досягаемости рабочей зоны человека с использованием антропометрически точной модели его тела и найденных позиций.

2. Определение рабочих поз

Здесь и далее в статье под рабочей позой подразумевается положение тела и конечностей работника относительно рабочего места при выполнении производственных операций.

Определение основной или основных рабочих поз, а также возможности их смены при выполнении рабочих операций, является отдельной задачей, которая решается за рамками создаваемой модели моторного пространства, но решение которой необходимо для корректного и объективного её построения. Термин «основное» в данном контексте означает, что во время работы человек проводит большую часть времени в рассматриваемом положении, но

при этом не исключает возможность изменения рабочей позы путем наклона или поворота тела, или смену позы на другую.

Возможными решениями задачи определения рабочих поз, возможности и частоты их изменения могут быть:

1. Частота и возможность смены рабочей позы может диктоваться существующими нормами, ГОСТами [3-5] и т.д.
2. Личные наблюдения за рабочим процессом, беседа с пользователями помогут собрать данные для эмпирического анализа.
3. Проведение опросов с целью выявить желания и предпочтениях пользователей рабочего места.

Определение часто используемых рабочих поз и степени их изменения является основой для построения трехмерной модели моторного пространства, поскольку позволяет выявить и наложить требуемые ограничения на возможные положения тела человека в пространстве при работе, выделив только необходимые.

Примером для построения модели рабочей зоны будет служить рабочее место хирурга при проведении открытых операций. Анализ видеоматериалов [11], показывающих работу хирурга при проведении открытых операций показал, что в работе хирурга, при проведении этого типа операций, преобладает положение стоя. При этом хирург не должен упускать из поля зрения место проведения операции, из-за чего его тело все время направлено в сторону операционного стола. Моторная зона хирурга при проведении открытых операций обладает определенными ограничениями, позволяющими продемонстрировать применение предложенной модели, и при этом практически исключает смену рабочей позы, что позволяет сосредоточиться на описании модели, а не особенностей профессии.

3. Построение антропометрически точной модели человеческого тела в 3d редакторе

Антропометрически точная модель человеческого тела используется для определения и наглядного отображения положения рук и тела человека в определенных ранее рабочих позах. Для построения такой модели предлагается использовать инструментарий современных 3d редакторов для построения скелетных моделей. Для построения модели используется редактор Autodesk 3ds Max[1], построение возможно и в других редакторах, обладающих инструментарием, позволяющим создание аналогичной скелетной модели, например, в продукте Blender Foundation, Blender[2].

Размер доступной человеку моторной зоны находится в прямой зависти от длины отдельных частей тела и их подвижности в суставных соединениях. Используя информацию из актуальных антропометрических таблиц [8] можно выбрать размеры тех или иных частей тела, а также определить максимальные углы вращения суставов. При построении антропометрической модели следует располагать все кости в одной плоскости, поскольку это упрощает её построение благодаря отсутствию необходимости перемещать камеру.

Отдельного внимания заслуживает построение скелетной модели позвоночника, поскольку он является сложной скелетной структурой, состоящей из 24 позвонков, каждый из которых имеет свои максимальные углы поворота относительно соседних позвонков. Однако, для целей определения объема движений позвоночника, его разбивают на три отдела: шейный, грудной и поясничный. Шейный отдел позвоночника включает в себя область от верхушки позвоночника до уровня плеч, грудной отдел от уровня плеч до уровня локтя опущенной руки в положении стоя, поясничный отдел всю область ниже этой линии. Поэтому для упрощения моделирования можно построить в костной модели только эти три области.

Поскольку все суставы человека имеют ограниченный угол поворота, то после построения всех требуемых костей необходимо смоделировать ограничения подвижности суставов, что, применительно к скелетной модели в 3d редакторе, означает ограничение диапазона допустимых углов поворота каждой кости.

В медицине выделяют три главных взаимно перпендикулярных оси вращения, относительно которых производятся сгибательные и вращательные движения. Это поперечная ось, вокруг которой происходит сгибание и разгибание, вертикальная, вокруг которой происходит супинация-вращение наружу и пронация — вращение внутрь. Третья ось — передне-задняя. Вокруг нее происходит отведение и приведение. Чем больше осей вращения имеет сустав, тем он подвижней. Различают суставы трехосные, двухосные и одноосные [8]. Принятые в медицине плоскости относительно человеческого тела и движения в суставах показаны на рисунке 1.

В качестве основы для антропометрических размеров берутся размеры 5-ой перцентили, если необходимо оценить зону досягаемости элементов рабочего окружения, или 95-ой перцентили, если требуется понять минимальные габаритные размеры. При использовании обеих моделей появляется возможность оценить диапазон возможных рабочих областей в выбранных положениях тела для 90% пользователей [10].

Также возможно дополнительное ограничение подвижности суставов, или редактирование скелетной модели в случае необходимости построения рабочей области для людей, у которых в силу обстоятельств такие ограничения присутствуют. Например, необходимо рассмотреть работу людей с загипсованной или ампутированной конечностью.

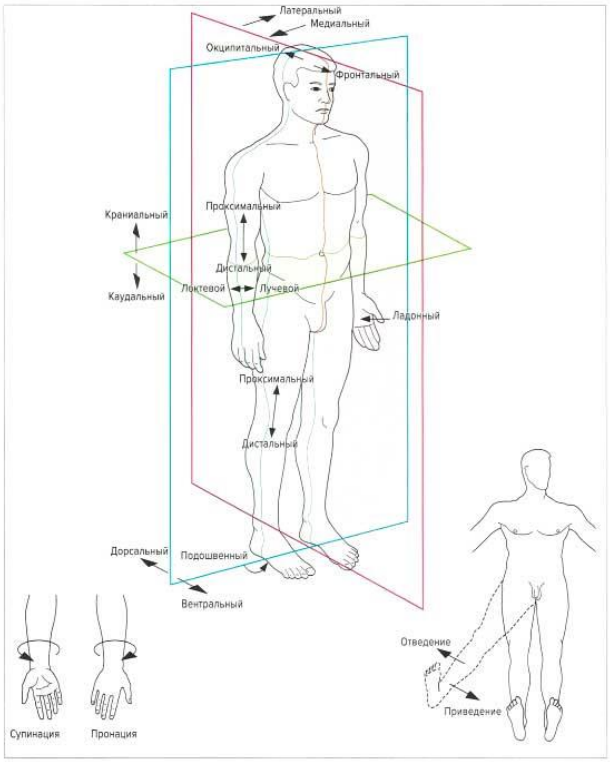


Рис 1. Плоскости человеческого тела и движения конечностей.

4. Создание трехмерной модели, отображающей досягаемость рабочей зоны

Досягаемость рабочей зоны в каждой из основных рабочих поз определяется набором крайних положений рук работника в пространстве с учетом необходимости, возможности, или невозможности изменения положения тела в момент выполнения производственной операции. Поэтому, для получения представления о доступном моторном пространстве, можно отобразить крайние положения рук работника в трехмерной сцене, а также промежуточные положения рук при перемещении. Поскольку на этапе анализа могло быть выявлено несколько основных поз, то отображать крайние положения необходимо для каждой из этих поз.

Поскольку ручное создание копий антропометрической модели в каждой из крайних точек и в промежуточных положениях может быть утомительным из-за повторяющихся операций, следует воспользоваться инструментарием, позволяющим автоматизировать этот процесс. Промежуточные положения при перемещении модели между крайними точками могут быть просчитаны инструментами для создания анимации в 3d редакторах. Построение анимационной модели смены положения рук и тела работника может также являться предпочтительным решением, поскольку позволяет не только автоматизировать построение промежуточных положений, но, с небольшими правками, использовать результат в презентационных целях.

Построение модели доступной рабочей области в данной статье выполняется при помощи инструмента Snapshot, имеющегося в Autodesk 3ds Max. Этот инструмент позволяет создать копию объекта, находящуюся в той же точке пространства, что и исходный объект в определенный момент времени. Snapshot можно использовать либо для создания одиночной копии объекта, либо для создания набора копий объекта на временном интервале с указанным промежутком. Полученный таким образом набор копий, может отображать перемещение объекта во время анимации.

При построении модели рабочей зоны хирурга берется во внимание тот факт, что ему нельзя упускать из внимания область проведения операции. Таким образом, допускается только небольшой наклон тела вперед с сохранением зрительного контакта с областью операции. Результат моделирования моторной области на примере рабочего места хирурга можно увидеть на рисунке 2.

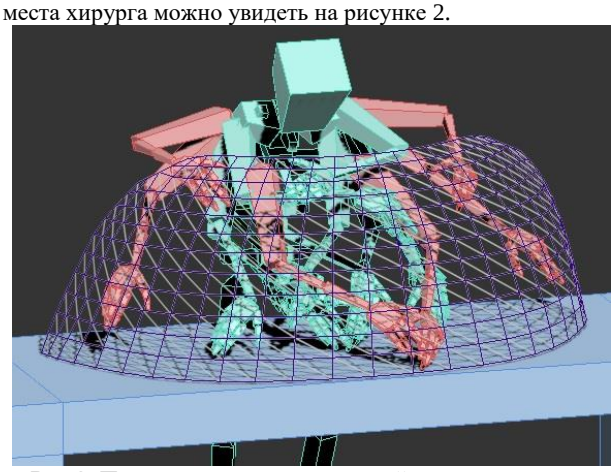


Рис 2. Полученная модель моторной зоны рук хирурга при проведении операции.

5. Заключение

Рассмотрены проблемы проектирования, решаемые при помощи соматографических моделей, а также предложено построение трехмерной соматографической модели, позволяющей облегчить проектирование в пакетах трехмерной графики.

Построенная модель тела человека позволяет при помощи настройки передвижения отдельных её частей определять доступную рабочую зону и строить её трехмерную модель. Возможность изменения подвижности соединений костей в модели позволяет проводить анализ нестандартных антропометрических ситуаций [7], когда подвижность отдельной части тела сильно ограничена, например, в случае с людьми, использующих протез верхней конечности. Данная модель может найти применение в профессиональной и исследовательской деятельности промышленных дизайнеров, а также в процессе обучения студентов.

Также данную модель можно использовать для определения моторной зоны нижних конечностей, в случае если это влияет на комфорт использования рабочего места.

Полученная модель человеческого тела была успешно применена для построения трехмерной модели доступного моторного пространства на примере рабочего места хирурга при проведении операции в положении стоя.

6. Литература

[1] Autodesk 3ds Max. https://www.autodesk.ru/products/3ds-max

[2] Blender Foundation Blender. https://www.blender.org/

[3] ГОСТ Р 56274-2014 Общие показатели и требования в эргономике

[4] ГОСТ Р ИСО 6385-2007 Эргономика. Применение эргономических принципов при проектировании производственных систем

[5] ГОСТ Р ИСО 9241-5-2009 Эргономические требования к проведению офисных работ с использованием видеодисплейных терминалов (VDT). Часть 5. Требования к расположению рабочей станции и осанке оператора

[6] ГОСТ Р ИСО 15534-3-2007 Эргономическое проектирование машин для обеспечения безопасности. Часть 3. Антропометрические данные

[7] ГОСТ Р ИСО 28803-2013 Эргономика физической среды. Применение требований стандартов к людям с особыми потребностями

[8] М. Ф.Иваницкий. Анатомия человека.

[9] М.С. Кухта, В.И. Куманин, М.Л. Соколова, М.Г. Гольдшмидт. Промышленный дизайн: учебник

[10] А.И. Фех. Эргономика: учебное пособие

[11] https://www.youtube.com/channel/UC_hZJxe37v_SzxhDliESxjg

7. Об авторах

Шкляр Алексей Викторович, старший преподаватель, отделение автоматизации и робототехники, Томского политехнического университета. E-mail shklyarav@tpu.ru

Черепова Лилия Николаевна, студентка инженерной школы информационных технологий и робототехники (бывш. Институт кибернетики) Томского политехнического университета. E-mail lni1@tpu.ru

Черепов Андрей Андреевич, студент инженерной школы информационных технологий и робототехники (бывш. Институт кибернетики) Томского политехнического университета. E-mail andrcherepov@gmail.com

Автоматизация проектирования, тренажеры и симуляторы

CAD/CAM/CAE, trainers and simulators

GraphiCon 2018

---

24 – 27 сентября 2018
Томск, Россия

Adaptive mesh generation using shrink wrapping approach

A. Malyshev[1], A. Zhidkov[2], V. Turlapov[1]
al.s.malyshev@gmail.com|artem.zhidkov@opencascade.com|vadim.turlapov@itmm.unn.ru
[1] Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia
[2] Open Cascade, Guyancourt, France

We consider the problem of mesh preparation for subsequent simulations. The Cartesian shrink-wrapping technique is used to construct a triangular mesh on dirty input geometry. Supported imperfections are gaps, overlaps, inconsistent normal orientations and self-intersections. Wrapping is stored using octree. An octree is refined adaptively to follow a form of the original mesh. Initial watertight mesh which is extracted from octree is improved using nodes projection to original triangulation and by local operations such as swap edge, collapse edge, and Laplacian smoothing.

Keywords: *shrink wrapping, octree, bounding volumes hierarchy.*

1. Introduction

Significant performance boost of the computers attracted close attention to automatic mesh generation. Mesh processing covers such engineering practices as mesh healing, mesh construction from triangles soup, adapting design models for simulation etc. The necessity of healing often emerges after data exchange different CAD application. Diverse tolerance handling approaches may ruin model for further use obtained via neutral formats like STEP or IGES.

Usually, healing is performed manually, and this process is tedious and time-consuming. The proposed approach allows reducing the number of manual operations related to detecting and fixing of self-intersections, gaps, overlaps and inconsistent normal orientations.

Generally, there are two main ways to deal with imperfections. Methods of the first group modify existing mesh or CAD. Another group is intended to build a new object, which is close to the original one. The second approach gives more freedom, so it is used in most cases.

On the CAD level, some clean-up simplification techniques were introduced using a virtual topology concept [14]. For meshes, various re-meshing articles were published ([3] and [12]). The common point between them is that they cannot deal with all kinds of topological and geometric problems simultaneously. It is necessary to produce a pipeline of methods where each algorithm fixes only some defects.

The robust industrial healing algorithms are often based on the volumetric representation. Initially, Kobbelt suggested shrink-wrapping concept in [8]. His idea was to wrap a thin deformable surface around the initial mesh. Wang [18] carried out research based Cartesian grid approach to deal with some imperfections. In that study, all the Cartesian cells intersecting the entity are refined recursively to satisfy cell size criterion. Bischoff [1] presented a method to extract manifold meshes from architectural models using dual contouring [5].

Lee [10] overcame the most of these issues using a Cartesian grid approach. Grid size functions ([19] and [20]) were utilized to refine cell adaptively. They took outer (or inner) wrapping constructed on the Cartesian grid of the input mesh as the original watertight deformable surface. After that, the watertight mesh is projected to the surface. At the end, the mesh is refined to obtain better skewness and aspect ratio. The method is based on the Wang's [18] interior-to-boundary approach.

Recent research of Juretić [6] was connected to holes detection based on the intensity of heat fluxes in various places of the input surface mesh. Another way of holes handling was proposed by Martineau [11] where the wrapping and volumetric mesh generator are combined.

Octree usage ([13], [15]) is the well-established technique to represent arbitrary free-form objects in the three-dimensional space. Octree fills the space using hierarchically structured axis-aligned boxes. Lee [10] states that octree construction requires more CPU time compared to the Cartesian grid but the lower amount of memory. In our study, we use octrees to build the adaptive representation of input mesh due to the following reasons:

- Construction simplicity.
- Effective parallelization. Nodes of octree do not depend on each other so that they can be processed in a parallel mode.
- Low memory consumption.

Many computational problems such as initial watertight nodes projection, self-intersection state checking or neighbors detection for a particular node can be effectively solved using accelerating data structures like bounding volumes hierarchy ([2], [9] and [16]). The idea of BVH is to subdivide a set of primitives built on input object recursively until some target criterion is met. In case of mesh input, as we have, no mesh conversion is needed since it is possible to reflect the mesh triangle as a primitive of BVH.

Usually, the result of wrapping is not directly suitable for analysis because of insufficient quality of mesh triangles. Local operators like edge swap, edge collapse [17] are used to improve the quality of the resulting mesh. Also, it is possible to use some decimation techniques [7], but they are not investigated in this study. If the quality of the resulting mesh is not an aim, this algorithm can be applied as automatic feature suppressor [11].

In our study, we present a "boundary-to-interior" method according to Wang terminology [18]. The idea is to convert the initial model to the volumetric form presented as octree and extract initial outer watertight mesh. After that, initial mesh is projected to the original mesh, and possible self-intersections are eliminated. At the end of the algorithm, mesh quality is improved.

2. Algorithm

The input for our algorithm is a mesh presented in stereolithography (STL) format. The majority of CAD systems provide the possibility to save faceted representation in that format due to its simplicity and portability.
The algorithm has one parameter to tune – minimal allowed voxel size. This parameter determines the maximal gap which algorithm can cover.
The algorithm has three main stages:

- Construction of original watertight mesh.
- Projection + Self-intersections elimination.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

- Mesh quality improvements.

2.1 Watertight mesh construction

The octree is a tree-like data structure where each non-leaf node has 8 children. Figure 1 illustrates the octree structure.

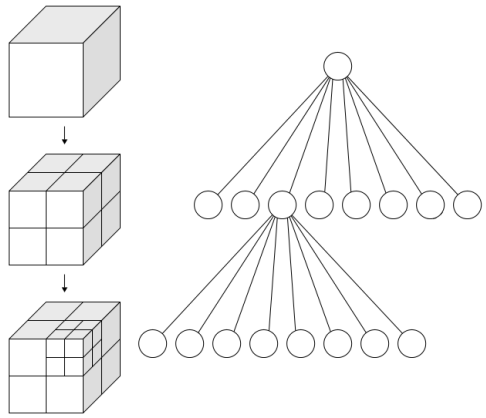


Figure 1: Octree structure.

The root node of voxel octree corresponds to the axis-aligned bounding box of dirty input mesh. The tree is constructed layer by layer and nodes of fixed depth are checked for subdividing in a parallel way. The number of nodes grows exponentially (up to 8^k nodes on the k layer). Adaptivity may significantly decrease this value. Our experiments show linear speed up on trees with more than one hundred thousand nodes. Subdividing is made taking into account the following principles:

- It is not necessary to subdivide nodes which do not intersect input triangles. Subdivision requires effective box-mesh intersection checks. In our research, it is done using BVH where triangles of the input dirty mesh are used as primitives. Separation axis theorem is used to check intersection state between triangles and boxes. Let N be the number triangles in input mesh and M – the number of nodes in the octree. BVH allows checking intersection state in $O(log(N))$ time for each box in case of proper balancing. The tree itself can be built on linear-logarithmic time. So, the full complexity of this operation in case of BVH usage is $O(N * log(N) + M * \log(N))$ while direct naive implementation requires $O(N * M)$ operations.
- Subdivision of intersecting nodes is not obligatory in case of quasi-planar intersecting triangles. MCAD models deal with canonical geometry, and it is a usual case to deal with (Figure 2 and Figure 3). Subdivision check is implemented by calculating the maximal angle between normals of triangles intersecting the current node. Besides, it is possible to construct gauss maps to speed up this check.

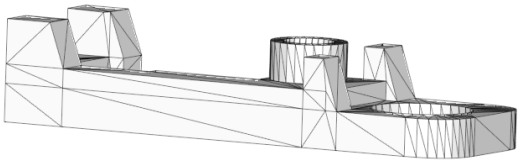


Figure 2: Faceted representation of MBB Gehause rohteil model with planar faces.

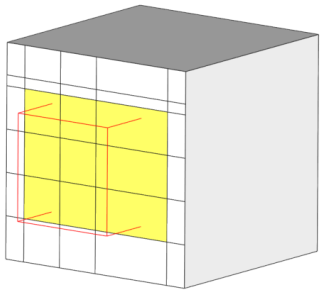


Figure 3: Demonstration of octree node which should not be subdivided (red lines) and intersecting facets of input model (yellow color).

Outer mesh extraction is hardly possible for arbitrary octree. The nature of this problem is connected to possible size jumps of adjacent nodes. In general, their sizes can have 2^n: 1 gradation (Figure 4).

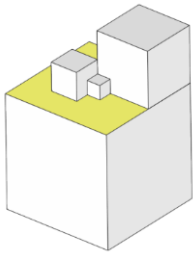


Figure 4: Octree with 8:1 relation. Outer contour extraction will add many points to the yellow face.

Again, effective neighbor detection is based on BVH data structure. For that, leaf nodes are converted to BVH primitives, and simple box intersections are used to check intersection. All neighbors for a particular node can be detected as all nodes intersecting enlarged bounding box of the current node. Since we know the minimal allowed cell size, it is easy to avoid numerical stability problems increasing bounding box by some value smaller than minimal size ("0.1 * minimal size" for example). The result of that process is demonstrated by Figure 5.

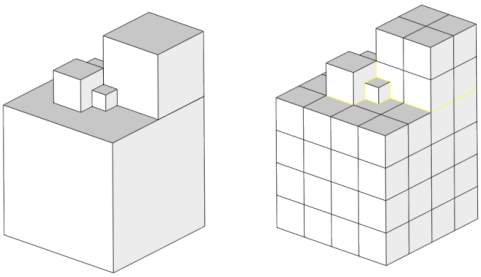


Figure 5: Model conversion to 2:1 form. Original node connections are marked by yellow color, from left to right.

One additional problem should be solved before outer contour extraction. Usually, adjacent nodes have non-manifold connections. In our paper conformity restoration is performed by adding new cells. The resulting tree might break 2:1 gradation rule and its restoration is needed. In general, initial polygons can have up to 8 points. Polygons are not tessellated during initial wrapping construction; it is performed after projection. It allows getting better quality mesh since in projection optimal tessellation may change. To conclude, watertight mesh construction involves the following steps:

1. Construct initial octree.
2. Convert octree to 2:1 form.
3. Resolve non-manifold joints.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

4. Convert octree to 2:1 form again.
5. Extract outer or inner wrapping.

2.2 Projection and self-intersection elimination

The first and the most critical step in watertight mesh improvement is projection. Nowadays, typical watertight wrapped mesh consists of million triangles. The efficiency of the projection procedure is crucial for the construction of a real industrial application. Direct exhaustive implementation of the projection algorithm requires $O(M * N)$ operations. As an alternative to it, BVH-driven projection gives about $M * \log(N)$ time complexity.

As mentioned above, the projection procedure comprises several healing steps like sticking polygons treatment. Usually, sticking polygons appear when input geometry has a substantial amount of planar places. The list of problems and their origins is presented below:

- Two polygons with equal nodes – move polygon nodes back in projection directions. It happens when the whole polygon is positioned at shell part of input mesh.
- A polygon with wrong normal – remove it. Non-manifold joints resolving adds new nodes some of them are placed inside the input geometry. Another source of that problem is sharp and complicated input geometry.
- A polygon with sticking nodes – collapse degeneracy. It arises when an edge of a polygon is orthogonal to input mesh.
- A polygon without area – remove it. It comes when the original non-projected polygon is entirely orthogonal to input mesh.

It turns out that nearest projection leads to a situation when neighbor nodes are projected in opposite directions. Sometimes, it indicates insufficient amount of octree refinement steps but, typically, the best cell size is unknown. That is why this issue is addressed. First, the accumulated length of node's outgoing edges is calculated, and it is compared with a perimeter of faces containing these nodes. If these two values are too far from each other new position will be assigned to the node using smoothing. Node is projected to original mesh once again after smoothing to sure that node is lies on the input triangulation. Wrong projection incident is shown in Figure 6.

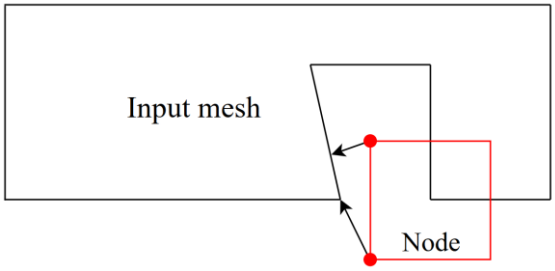


Figure 6: Wrong projection.

The last step of projection routine is self-intersection correction. Unlike the previous problems before, this issue is more geometric than topological. It means that implementation should be robust enough to fight with numerical issues raised from floating point operations performed using IEEE 754 standard. This idea should be kept in mind while developing algorithms for detection and fixing of self-intersections. Self-intersections are obtained using BVH tree to reduce complexity from quadratic to linear-logarithmic. The intersecting pairs are grouped into clusters according to their adjacency. Each cluster is processed independently in a parallel mode that decreases execution time. Intersections are amended using smoothing. The idea is to keep cluster boundary untouched and modify only internal nodes.

2.3 Mesh quality improvements

Usually, outer wrapping is not applicable for analysis after projection due to near degenerated polygons. The ordinary way to improve mesh quality is to apply local topological operators like edge swap, face collapse etc. The following metrics are utilized to check mesh quality:

- Skewness. $skewness = 1 - \frac{2r}{R}$, where r is the radius of the inscribed circle and R is the radius of the circumscribed circle.
- Aspect ratio. $aspect\ ratio = \frac{L}{l}$, where L is the length of the longest side, l is the length of the shortest side.

Thanks to polygons usage on the previous steps of the algorithm, it is possible to build different triangulations and get better mesh quality. The triangulation algorithm is based on the ear-slicing approach [4]. It tessellates arbitrary convex or concave polygon to triangles but in a non-optimal manner. Greedy-algorithm is used to determine the best ear to slice since brute-force check of all possible tessellations is computationally ineffective. Despite the execution time, some workflows like volumetric mesh generation may benefit from optimal tessellation because the mesh is constructed by increasing dimensions from lower to higher and excellent quality in the previous dimension is mandatory to build fine mesh in upper dimensions.

The local mesh operators like edge collapse are used to improve the quality of tessellated triangular mesh. In our algorithm the following operators are implemented:

- Collapse edge. The lengths of all neighbors are calculated for the target edge. If an edge is several times shorter than all its neighbors, the target edge will be collapsed.
- Split edge. Neighbor lengths are calculated like in case before. If the target edge's length is significantly bigger than neighbor lengths a new node will be introduced in edge, while two adjacent faces containing the target edge will be split into four triangles.
- Swap edge. Sometimes two faces containing target edge may be improved by connecting non-shared nodes of triangles and removing existing target edge.

All mentioned operations are safe from self-intersection perspective. It is checked that the mesh is not corrupted during modification. In general, this leads to a situation when arbitrary mesh cannot be converted to an optimal state.

3. Mesh generation examples

Our algorithm is capable of generating watertight meshes for many types of imperfections like non-manifold topology or overlaps. The examples presented in that section are intended to demonstrate obtained performance and applicability of our method for industrial use.

3.1 Geometric primitives: sphere, cylinder, and cube

These cases are aimed to demonstrate basic wrapping abilities and robustness against common defects. The initial mesh, shown in Figure 7, has a missing triangle, that situation emulates erroneous data translation from one format to another. The sphere has a radius of 0.1 m, and the longest side of the missing triangle has a 0.0018 m length. This model is processed with 0.005 m cell size, and the result is shown in Figure 8.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

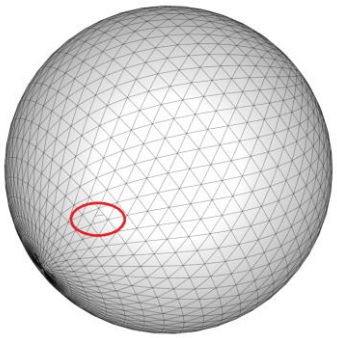


Figure 7: Sphere with missing triangle (marked by the red ellipse).

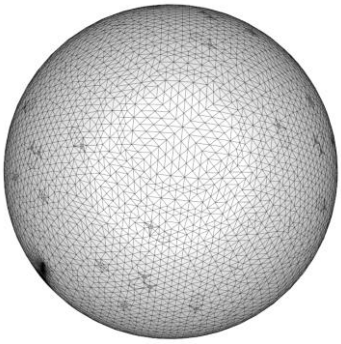


Figure 8: Wrapped mesh.

The mesh presented in the Figure 9 demonstrates the result of inaccurate modeling when two objects are mutually intersecting. This artificial model is constructed using two orthogonal cylinders with a radius of 0.01 m and 0.1 m in length. Wrapped mesh generated using 0.001 m cell size fixes the problem. It should be noted, that this case also shows the conversion possibility from CAD-based mesh to its computer graphics equivalent. The result of wrapping is shown in Figure 10.

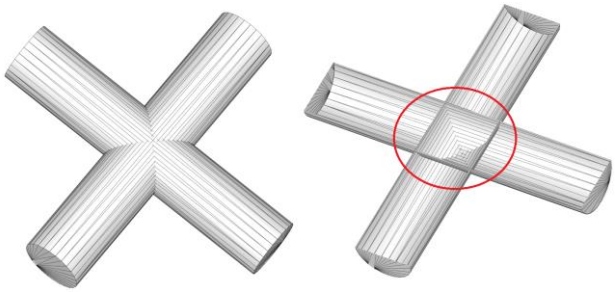


Figure 9: Cylinder models and cylinders with clipping planes (intersection is marked by red color), from left to right.

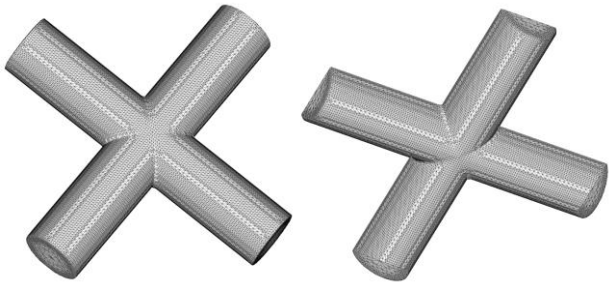


Figure 10: Wrapped cylinders and clipped cylinders, from left to right.

The cube model displayed on the left side of Figure 11 has 0.1 m size length. This model is used to compare execution time, the number of triangles and peak memory consumption of uniform and adaptive versions of the wrapping. Table 1 shows almost linear increasing of all tracked parameters; however, Table 2 demonstrates square complexity for uniform octree construction.

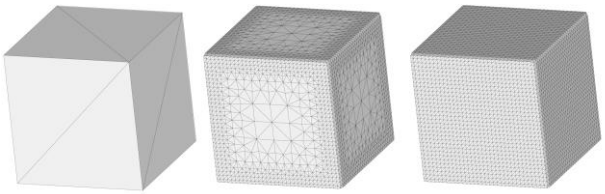


Figure 11: Original box, adaptive wrapping and uniform wrapping from left to right.

Cell size, mm:	Execution time, sec:	Max used memory, MiB:	Number of triangles:
0.5	0.91	44	6912
0.25	2.08	54	16224
0.125	4.50	75	35520
0.0625	9.76	125	74784

Table 1: Adaptive wrapping of the cube model.

Cell size, mm:	Execution time, sec:	Max used memory, MiB:	Number of triangles:
0.5	1.35	51	12288
0.25	5.86	94	49152
0.125	23.60	249	196608
0.0625	96.37	959	786432

Table 2: Uniform wrapping of the cube model.

3.2 Industrial model: engine

The following example is designed to reveal envelope construction capabilities on the real industrial model. The model shown in Figure 12 is the wright whirlwind radial engine. The model sizes are equal to 250, 300 and 550 mm for X, Y and Z axis correspondingly. The CAD model containing 1002 parts is meshed, and the input mesh has 1635339 triangles. This mesh is wrapped with cell size equal to 1.0 mm. The resulting mesh presented in Figure 13 has 1120166 triangles and 559839 nodes. Multiple problems with non-valid topology, self-intersecting mesh are detected and healed during envelope construction.

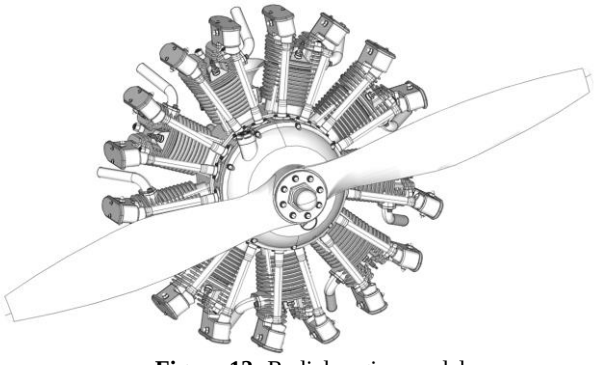


Figure 12: Radial engine model.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

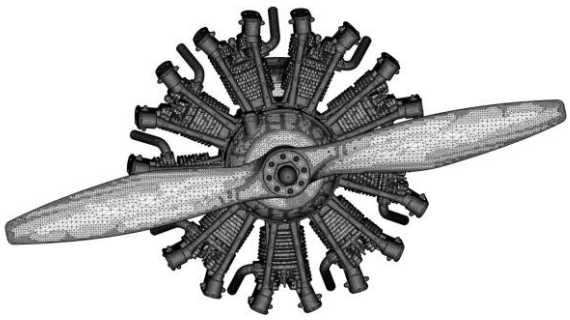


Figure 13: Engine model wrapping, constructed in 48 minutes and 44 seconds on core i5-3450 CPU.

4. Conclusion and further work

This article presents a mesh generation method for surface wrapping on dirty models. No assumptions about quality of input mesh are made. The proposed approach fixes both geometrical and topological problems such as overlaps, open boundary edges, and non-manifold configurations. The octree is used to construct adaptive outer wrapping. The quality of the resulting mesh is controlled by the minimal allowed cell size, and this value defines the gap filling capability of the algorithm. Initial outer wrapping is refined to reach computational mesh quality using local operators, such as edge collapse, edge split, and edge swap. Laplacian smoothing is used for eliminating self-intersections and deeper refinement of the mesh quality.

The method can be further improved in several ways. Some sub-routines should be investigated to be used in parallel mode. Memory usage is another room for improvement since existing implementation is not optimized from a memory point of view.

In many practical cases, model features like sharp edges preserving should be kept, but no delicate algorithms are investigated in that study. The present approach smooths input geometry which is considered as an automatic feature suppression mechanism.

The topological state of the result of the wrapping and meshing can be enhanced by the usage of proximity size functions proposed by Zhu [20]. Specialized algorithms for efficient proximity computation should be developed.

Finally, the proposed algorithm can be adapted for various case studies like protection of intellectual properties, model simplification, and preparation for thermal, fluid or aero analysis.

5. References

[1] Aichholzer, O. and Aurenhammer, F. 1996. Straight skeletons for general polygonal figures in the plane. Computing and Combinatorics, Lecture Notes in Computer Science Volume 1090. (1996), pp 117-126.

[2] Chen, X.D., Yong, J.H., Wang, G., Paul, J.C. and Xu, G. 2008. Computing the minimum distance between a point and a NURBS curve. CAD Computer Aided Design. 40, 10–11 (2008), 1051–1054.

[3] Coupez, T., Digonnet, H. and Ducloux, R. 2000. Parallel meshing and remeshing. Applied Mathematical Modelling. 25, 2 (2000), 153–175.

[4] ElGindy, H., Everett, H. and Toussaint, G. 1993. Slicing an ear using prune-and-search. Pattern Recognition Letters. 14, 9 (1993), 719–722.

[5] Ju, T., Losasso, F., Schaefer, S. and Warren, J. 2002. Dual contouring of hermite data. ACM Transactions on Graphics. 21, 3 (2002).

[6] Juretić, F. and Putz, N. 2014. Applications of heat-diffusion equation for surface wrapping: Hole detection and normal orientation. Engineering with Computers. 30, 3 (2014), 363–374.

[7] Kobbelt, L., Campagna, S. and Seidel, H.-P. 1998. A General Framework for Mesh Decimation. Graphics Interface. (1998), 43–50.

[8] Kobbelt, L.P., Vorsatz, J., Labsik, U. and Seidel, H.-P. 1999. A Shrink Wrapping Approach to Remeshing Polygonal Surfaces. Computer Graphics Forum. 18, 3 (1999), 119–130.

[9] Lauterbach, C., Garland, M., Sengupta, S., Luebke, D. and Manocha, D. 2009. Fast BVH construction on GPUs. Computer Graphics Forum. 28, 2 (2009), 375–384.

[10] Lee, Y.K., Lim, C.K., Ghazialam, H., Vardhan, H. and Eklund, E. 2010. Surface mesh generation for dirty geometries by the Cartesian shrink-wrapping technique. Engineering with Computers. 26, 4 (2010), 377–390.

[11] Malyshev, A., Slyadnev, S. and Turlapov, V. 2017. Graph-based feature recognition and suppression on the solid models. GraphiCon 2017. (2017), 319–322.

[12] Peyré, G. and Cohen, L.D. 2006. Geodesic remeshing using front propagation. International Journal of Computer Vision. 69, 1 (2006), 145–156.

[13] Schneiders, R., Schindler, R. and Weiler, F. 1996. Octree-based Generation of Hexahedral Element Meshes. Meshing Roundtable. December 1999 (1996), 205–216.

[14] Sheffer, A., Bercovier, M., Blacker, T. and Clements, J. 2000. Virtual topology operators for meshing. International Journal of Computational Geometry & Applications. 10, 3 (Jun. 2000), 309–331.

[15] Shephard, M.S. and Georges, M.K. 1991. Automatic three-dimensional mesh generation by the finite octree technique. International Journal for Numerical Methods in Engineering. 32, 4 (1991), 709–749.

[16] Teschner, M., Kimmerle, S., Heidelberger, B., Zachmann, G., Raghupathi, L., Fuhrmann, A., Cani, M.P., Faure, F., Magnenat-Thalmann, N., Strasser, W. and Volino, P. 2005. Collision detection for deformable objects. Computer Graphics Forum. 24, 1 (2005), 61–81.

[17] Wang, D., Hassan, O., Morgan, K. and Weatherill, N. 2006. Enhanced remeshing from STL files with applications to surface grid generation. Communications in Numerical Methods in Engineering. 23, 3 (Sep. 2006), 227–239.

[18] Wang, Z.J. and Srinivasan, K. 2002. An adaptive Cartesian grid generation method for “dirty” geometry. International Journal for Numerical Methods in Fluids. 39, 8 (2002), 703–717.

[19] Zhu, J. 2003. A new type of size function respecting premeshed entities. Proceedings of the 11th international meshing roundtable. (2003), 403–413.

[20] Zhu, J., Blacker, T. and Smith, R. 2002. Background overlay grid size functions. Proceedings of the 11th International Meshing Roundtable. (2002), 65–73.

About authors

Vadim Turlapov (vadim.turlapov@itmm.unn.ru): Professor of Computer Science, Lobachevsky Nizhny Novgorod University.

Alexander Malyshev (al.s.malyshev@gmail.com): Ph.D. student, Lobachevsky Nizhny Novgorod University.

Artem Zhidkov (artem.zhidkov@opencascade.com): Ph.D., OpenCascade, France.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Особенности использования сценариев в автоматическом тестировании систем оптического моделирования

Копылов М.С. [1], Бирюков Е.Д. [1], Барладян Б.Х. [1]
kopylov@gin.keldysh.ru|birukov@gin.keldysh.ru|obb@gin.keldysh.ru
[1]ИПМ им. М.В. Келдыша РАН, Москва, Россия

В работе рассмотрен опыт применения сценариев, написанных на языках Python и Visual Basic, для автоматического тестирования отдельных программных компонентов комплекса оптического моделирования. Также рассмотрены аспекты отдельного тестирования пользовательского интерфейса такого комплекса. Приведены примеры подобного тестирования.

***Ключевые слова**: автоматическое тестирование, сценарии, графический интерфейс.*

Usage of scenarios for automatic testing of optics CAD systems

Kopylov M.S. [1], Birukov E.D. [1], Barladian B.Kh. [1]
kopylov@gin.keldysh.ru|birukov@gin.keldysh.ru|obb@gin.keldysh.ru
[1] Keldysh Institute of Applied Mathematics, Moscow, Russia

This work describes specific features of Python and Visual Basic scenarios usage for automatic testing of individual program components of the 3D optical simulation software package. Also there are described aspects of a special testing of graphical user interface for such package. Examples of such testing are provided.

***Keywords**: automatic testing, scripts, graphical interface.*

1. Введение

Увеличение сложности программного обеспечения требует максимального использования всех доступных средств автоматизации разработки. Тестирование программного обеспечения, как одна из дорогостоящих процедур, нуждается в автоматизации на всех этапах жизненного цикла. Это особенно критично в условиях ограниченных ресурсов и времени на выпуск новых версий продукта.

Для программных комплексов, выполняющих расчеты различных физических процессов, к которым относятся системы оптического моделирования, наиболее важным видом тестирования является тестирование корректности результатов, которое может выполняться путем сравнения с аналитическими результатами или результатами полученными другими независимыми программами.

Регрессионное тестирование основано на сравнении полученных результатов в текущей версии программы с аналогичными результатами, полученными в предыдущей версии программы, которые считаются заведомо правильными.

Кроме того, в системах автоматизированного проектирования требуется тестирование корректности работы пользовательского интерфейса, зачастую имеющего сложную разветвленную структуру.

2. Использование сценариев при тестировании компонентов ядра комплекса

Комплекс оптического моделирования Lumicept, разрабатываемый нашим коллективом в ИПМ им. М. В. Келдыша РАН представляет собой модульную систему, состоящую из различных компонентов, которые можно разделить на две подгруппы – модули ядра и модули графического интерфейса. При этом множество компонентов ядра содержит в себе почти весь функционал комплекса, поэтому задача тестирования этой функциональности является первоочередной.

Комплекс Lumicept также интегрирован с существующей системой автоматизированного проектирования CATIA, которая является одной из популярных САПР в области машиностроения, в частности, проектирования автомобилей и самолетов. Интеграция выполнена в виде специальных дополнений (add-ons) к системе CATIA, которые вызывают общие модули ядра (в некоторых случаях – и общие модули графического интерфейса) [1].

Основные модули комплекса Lumicept имеют встроенную систему обработки сценариев на базе языка Python. Подмножество внутренних объектов и классов данного комплекса доступно из этих сценариев через Python API [2]. Данная особенность даёт возможность автоматизировать действия пользователя по работе с различными объектами данного комплекса, такими как объекты сцены, модули рендеринга, различные симуляторы и.т.д. К настоящему времени общее число подобных объектов весьма велико, например одних только объектов сцены существует более сотни. Так же необходимо отметить, что многие из этих объектов имеют достаточно большое количество различных настроек, влияющих на режимы их работы и вычислений.

Сценарии можно запускать как в режиме командной строки, так и в режиме графического интерфейса. Как правило, при автоматическом тестировании используется режим командной строки. Запуск сценариев осуществляется с помощью консольной программы ipython.exe, которая представляет собой специальный модуль комплекса Lumicept, содержащий в себе интерпретатор языка Python с возможностью доступа ко всем остальным объектам комплекса. Пример простейшего скрипта на языке Python приведен ниже.

```
def Render(file):
    kernel = Kernel()
    kernel.LoadScene(file)
    rp = RenderParams()
    rp.res = (800, 600)
    rp.depth = 5
    res = kernel.Render(rp)
    res.SaveImageToFile(…)
    res.CompareImageWithEthalon(…)

files = [“file1”, “file2”, …, “file99”]
```


The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

```
for f in files:
    Render(f)
```

В данном примере в ядро комплекса Lumicept последовательно загружаются файлы, содержащие в себе различные модели (сцены). Затем производится рендеринг каждой модели и результат (как правило, это графическое изображение) сохраняется в новый файл. На заключительном шаге производится сравнение получившегося изображения с эталонным изображением, с помощью специальной подпрограммы.

Работа со сценариями состоит из нескольких обязательных этапов - подготовка исходных данных, выполнение расчётов, сохранение и анализ результатов. В качестве необязательного этапа можно добавить установку или настройку различных параметров модулей, которые задействованы в вычислениях. Сценарии не имеют каких-либо ограничений по длине (количеству операторов), а также могут вызвать другие сценарии, быть вложенными, использовать рекурсию и.т.д. При этом из любого сценария можно получить доступ как к одному, так и к нескольким модулям комплекса оптического моделирования одновременно.

Подобная широкая функциональность, несомненно, является весьма удобной для автоматизации вычислений. В рамках тестирования же, как правило, ограничиваются работой только с одним конкретным модулем в каждый момент времени, то есть для каждого модуля или компонента модуля пишется отдельный тестовый сценарий, который используется для тестирования работоспособности этого компонента или модуля. Укрупнение подобных сценариев тоже возможно, например, объединив их в пакет подпрограмм в рамках одного сценария.

Особенностью реализации Python API, разрабатываемой нами для написания тестовых сценариев, является строгая проверка на ошибки, возникающие по ходу выполнения сценариев. Допустим, если по каким либо причинам исходный файл не сможет загрузиться, например, из-за ошибки в тестируемом модуле, то Python API немедленно вернёт код ошибки и выполнение сценария будет остановлено, что в свою очередь будет явно сигнализировать о возможных проблемах в данном модуле. Важно отметить, что даже успешное завершение тестового сценария не гарантирует правильности работы тестируемого модуля. Поэтому любой тестовый сценарий должен дополнительно выполнять анализ получившихся результатов. Обычно такой анализ представляет собой сравнение различных численных и графических данных, полученных при оптических симуляциях в загруженной сцене, например, результатов рендеринга.

Дополнительно требуется анализ скорости работы тестового сценария и количества используемой памяти в процессе работы. Выбор того или иного типа анализа зависит непосредственно от формата исходных данных, а также от теста (функциональный или регрессионный). Учитывая богатые возможности языка сценариев Python в плане обработки данных и в наличии множества дополнительных пакетов расширения самого разного назначения, таких как numpy, scipy, matplotlib, imageio, задачи анализа получившихся результатов могут быть реализованы непосредственно в сценарии тестирования, обычно в виде отдельной подпрограммы. Например, если результатами некоторых симуляций являются графические изображения, то их можно сравнивать с помощью библиотеки scikit-image (часть пакета scipy) для поиска различающихся регионов в изображениях [7]. В некоторых случаях (например, при проведении трассировки лучей методом Path Tracing) в вычислениях используются случайные величины, поэтому результаты всегда будут немного отличаться от запуска к запуску. В этих случаях для успешного прохождения теста среднеквадратичное отклонение между сравниваемыми изображениями не должно превышать некоторого определённого значения.

3. Многопоточность в сценариях

При автоматическом тестировании компонентов комплекса оптического моделирования весьма важным критерием является эффективность данной процедуры с точки зрения затраченного времени. К примеру, в настоящий момент, для выполнения всех тестов использующих сценарии может понадобиться от 12 до 15 часов (на компьютере со следующими характеристиками: Процессор: Intel Core i7-3770 3,4 ГГц; RAM: 32 Гб; ОС: Windows 8.1 x64). Поэтому в рамках работы с комплексом были выявлены задачи тестирования, выполняющиеся не оптимально с точки зрения использования доступных вычислительных ресурсов, характеризующиеся неполной загрузкой всех доступных процессоров или процессорных ядер. В основном это вызвано такими причинами, как собственные ограничения алгоритмов обработки графических данных в многопроцессорных/многоядерных конфигурациях и конфигурациях с неоднородным доступом к памяти (NUMA), или же из-за использования тестовыми сценариями некоторых специфических методов ядра комплекса, которые являются однопоточными по своей природе.

К счастью, язык сценариев Python обладает достаточно продвинутой поддержкой многопоточного программирования, позволяющей весьма сильно увеличить эффективность выполнения вышеперечисленных групп сценариев. Данным языком, например, поддерживаются оба базовых вида распределения вычислительной нагрузки - многопоточность и многопроцессность, которые реализованы в модулях threading и multiprocessing соответственно. При этом стоит отметить, что унификация интерфейсов данных модулей делает их весьма удобными в использовании. В некоторых случаях в исходном сценарии достаточно изменить лишь несколько строк кода, чтобы перейти от использования одного модуля к другому и наоборот. Более того, оба модуля поддерживают один и тот же набор примитивов синхронизации, что упрощает разработку программ. Выбор того или иного способа реализации многопоточности в тестовом сценарии, как правило, зависит от конкретных задач, решаемых им, и выбирается разработчиком данного сценария.

Как показывает наш опыт, модуль threading имеет некоторые ограничения при работе с потоками в плане эффективности, если эти потоки в основном работают с общими данными, такими как массивы, словари, списки. Другое ограничение связано с возможностью ситуации, при которой из двух потоков одновременно будет вызвана одна и та же функция ядра комплекса. Учитывая, что многие функции ядра комплекса не являются потокобезопасными, существует реальная угроза аварийного завершения при выполнении многопоточных тестовых сценариев, или возникновения ошибок в результатах вычислений или тестирования. В связи с этим, использование модуля multiprocessing является более предпочтительным при использовании в тестовых сценариях, так как при использовании процессов можно добиться большего параллелизма, особенно при задачах пакетной обработки данных, при этом данный модуль не имеет каких либо ограничений [6].

Пример многопоточного тестового сценария, накладывающий фильтр на множество отдельных изображений, приведен ниже. В данном примере каждый

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

экземпляр подпрограммы наложения фильтра на исходные изображения, выполняется в отдельном процессе, благодаря использованию библиотеки multiprocessing позволяющей вынести код любой функции сценария в отдельный процесс.

```
def ApplyFilter(name):
    pp = PostProcessor(name)
    res = pp.ApplyAverageFilter()
    res.SaveImageToFile(…)
    res.CompareImageWithEthalon(…)

if __name__ == '__main__':
    from multiprocessing import Process
    files = ["file1", "file2", … ]
    for f in files:
        p = Process(target = ApplyFilter, args = f)
        p.Start()
```

4. Особенности работы сценариев в САПР CATIA

Существует реализация программы Lumicept, которая представляет собой дополнения к САПР CATIA. Эта САПР использует технологию COM для работы встроенной системы сценариев. Данная технология предполагает, что исходный код всех объектов, которые существуют в самой системе CATIA, а также те, которые были добавлены к CATIA в различных дополнениях, должен быть написан по определенному стандарту. При этом та часть исходного кода для каждого объекта, которая содержит объявление этого объекта со списком находящихся в нем свойств и методов, является открытой. Эта часть написана на специальном языке IDL. Все объекты, написанные по стандартам COM, доступны с помощью пользовательских сценариев, при этом теоретически возможно использование любого языка сценариев, реализация которого поддерживает стандарт COM. Имена для объектов, их свойств и методов при написании сценариев берутся из исходного кода на языке IDL, приложенному к системе CATIA и к дополнениям к ней.

Существуют три способа вызвать запуск выполнения сценария в системе CATIA. Первые два способа предполагают использование команд самой системы CATIA для запуска сценария – непосредственно с помощью команды «Запуск сценария» из пользовательского интерфейса, либо при запуске CATIA из командной строки в качестве дополнительного параметра. Третий способ предусматривает вызов внешнего интерпретатора сценариев, который уже самостоятельно запускает систему CATIA и обеспечивает передачу обращений к объектам системы CATIA с помощью технологии COM. Этот последний способ позволяет использовать любой язык сценариев, поддерживающий стандарт COM, но, к сожалению, позволяет запустить систему CATIA только в консольном режиме, без отображения пользовательского интерфейса, поэтому не может использоваться для тестирования самого пользовательского интерфейса. Первые два способа могут использоваться как при консольном, так и при графическом режиме отображения системы CATIA, однако, в этом случае доступны только те языки сценариев, которые встроены в саму систему CATIA. В настоящее время такие языки представлены только несколькими разновидностями языка Basic.

Поскольку для программы Lumicept стандартом сценариев является Python, то этот язык применяется и при тестировании дополнения Lumicept для системы CATIA в тех случаях, когда это возможно. Однако в тех случаях, когда использование языка Python не представляется возможным (в первую очередь – при тестировании графического интерфейса модулей Lumicept в системе CATIA) – тогда используется язык CATVBs.

5. Методы тестирования пользовательского интерфейса САПР

Автоматическое тестирование пользовательского интерфейса, как правило, проводится с помощью отдельной программы. Для тестирования комплекса оптического моделирования Lumicept используется инструмент AutoIt, который имеет собственный встроенный язык сценариев, похожий на Basic [3]. Для имитации действий пользователя эти сценарии используют систему идентификаторов элементов управления в интерфейсе, и команд, которые можно посылать данным элементам управления. В операционных системах семейства Windows используются дескрипторы окон (при этом под окнами понимаются любые элементы управления), и сообщения, которые можно послать окнам с соответствующими дескрипторами с помощью функции SendMessage [4].

Если каждый элемент управления однозначно реагирует на действия мыши и клавиатуры, при этом достаточно лишь чтобы указатель мыши находился над любой точкой элемента управления, то такая система автоматического тестирования обеспечивает полную эмуляцию действий пользователя. Однако отличительной особенностью систем автоматизированного проектирования является необходимость выделять с помощью мыши не только стандартные оконные элементы управления, но также различные объекты, отображенные в основном окне просмотра. С точки зрения операционной системы окно отображения сцены является одним элементом управления, и объекты, нарисованные в этом окне, не имеют никаких внешних идентификаторов, по которым их можно найти сторонней программой. В большинстве систем автоматизированного проектирования существует то или иное отображение иерархической структуры объектов в сцене в виде дерева, обычно отображаемое в отдельном окне. Зачастую такое окно является стандартным элементом управления (например, TreeViewControl в операционной системе Windows), что позволяет выделять объекты в сцене с помощью такого дерева. Однако система автоматизированного проектирования CATIA отличается от других подобных систем тем, что дерево объектов сцены отображается в том же окне, что и сама сцена [5]. Стандартное окно системы CATIA с отображенным в нем деревом сцены показано на рис. 1.

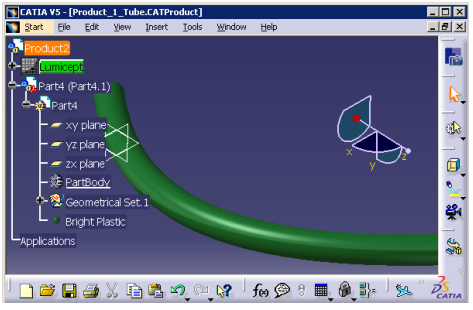


Рис. 1. Окно системы CATIA с загруженной сценой и отображенным деревом иерархии сцены

Ни само дерево, ни его отдельные элементы не имеют внешних идентификаторов. Поэтому в любом случае необходимы некоторые дополнительные действия для

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

выделения в окне просмотра некоторой точки, соответствующей заданному объекту. Автоматизация тестирования пользовательского интерфейса в этом случае может быть произведена несколькими способами.

Первый способ выделения мышью определенной точки на элементе интерфейса предусматривает жесткое прописывание в сценарии теста конкретных координат этой точки. Второй способ заключается в использовании систем распознавания изображений. Оба этих способа имеют свои недостатки: в первом случае – необходимость использования строго определенного разрешения экрана и масштаба элементов управления, во втором случае – общую неточность работы систем распознавания изображений, особенно, если на экране присутствуют несколько похожих друг на друга объектов.

Третий способ, выбранный авторами в качестве основного, представляет собой комбинацию использования сценариев, встроенных в тестируемую программу, и внешней программы автоматических тестов пользовательского интерфейса. Основу тестов составляет сценарий, выполняемый в самой тестируемой программе. Этот сценарий вызывает внешнюю программу тестирования пользовательского интерфейса (в нашем случае – AutoIt) непосредственно в тех местах, где требуется тестирование отдельных элементов интерфейса.

Рассмотрим для примера тестирование диалогового окна редактирования параметров материала в системе CATIA. Данное диалоговое окно является частью самой системы CATIA, однако в рамках комплекса оптического моделирования, разработанного нашим коллективом, к нему были добавлены дополнительные параметры. При тестировании используется сценарий на языке CATVBs (разновидность языка Visual Basic), который вызывается из самой системы CATIA, а также внешний инструмент AutoIt и его встроенный язык сценариев.

Тестирование состоит из следующих этапов:

1. Запуск системы CATIA в стандартном режиме с графической оболочкой, но с автоматическим запуском сценария. Для запуска используется пакетный командный файл (BAT), в котором заданы параметры – необходимая оболочка (Workbench) системы CATIA (в данном случае – оболочка Lumicept), а также путь к сценарию, который будет запущен сразу после запуска. Следует обратить внимание на то, что в данном случае система CATIA будет открыта в стандартном графическом режиме, хотя, как правило, при автоматическом запуске сценариев используется консольный режим.

2. Сценарий (назовем его сценарий 1) на языке CATVBs, который выполняет действия над объектами сцены. Данный сценарий производит поиск материала с заданным именем, добавляет его в список выделенных объектов, вызывает стандартную команду Properties(), которая открывает диалог параметров выделенного объекта, после чего вызывает программу AutoIt для автоматического тестирования диалога параметров. Для вызова AutoIt используется команда SystemService.ExecuteProcessus() из набора системных процедур CATIA. Эта команда ждет завершения работы программы AutoIt; на протяжении этого времени диалог параметров остается открытым. Его закрытие произойдет уже далее, с помощью сценария AutoIt, который будет имитировать нажатие пользователем кнопки OK. После закрытия диалога параметров управление снова переходит к сценарию 1, который считывает значения требуемых параметров из объекта материала и проверяет, что они были правильно заданы с помощью диалога параметров.

3. Сценарий (назовем его сценарий 2) для работы инструмента AutoIt. Этот сценарий находит диалоговое окно с заголовком "Properties", проверяет, что это окно принадлежит системе CATIA с помощью сравнения идентификаторов главного окна системы CATIA и данного диалогового окна. Затем AutoIt производит активацию этого окна, после чего передает фокус ввода в элемент управления (поле ввода) с именем "SpnDifReflPos", записывает в это поле текст "0.5" и инициирует нажатие кнопки OK. Управление снова передается самой системе CATIA, в которой продолжает работу сценарий 1. Он проверяет правильность установленных данных.

6. Заключение

Система автоматических тестов на базе сценариев, описанная в данной статье, в настоящее время используется при разработке системы оптического моделирования и реалистичной компьютерной графики, которая ведется в ИПМ им. М.В. Келдыша РАН. Также начала внедряться система тестирования пользовательского интерфейса того же программного комплекса. При этом тестированию, в том числе, стало возможным подвергать ту часть комплекса, которая представляет собой дополнения к системе автоматизированного проектирования CATIA.

7. Литература

[1] Барладян Б.Х., Дерябин Н.Б., Шапиро Л.З. Интеграция модуля компьютерной графики в систему CATIA // Новые информационные технологии в автоматизированных системах: материалы шестнадцатого научно-практического семинара - Моск. Ин-т электроники и математики национального исследовательского университета «Высшая школа экономики». М., 2013, с. 11-19.

[2] Н.Б. Дерябин, Д.Д. Жданов, В.Г. Соколов. Внедрение языка сценариев в программные комплексы оптического моделирования // Программирование, 2017, № 1, с. 40-53.

[3] Денисов Е.Ю., Волобой А.Г., Калугина И.А. Автоматическое тестирование графического интерфейса интерактивных программных комплексов // Новые информационные технологии в автоматизированных системах: материалы двадцатого первого научно-практического семинара - М.: ИПМ им. М.В.Келдыша, 20 апреля 2018, с. 83-88.

[4] Atif Memon, Adithya Nagarajan, 2003. GUI Ripping: Reverse Engineering of Graphical User Interfaces for Testing. In Proceedings of the 10th Working Conference on Reverse Engineering (WCRE '03)

[5] Dassault Systemes, Inc. CATIA Vertsion 5-6 R2015 Documentation, 2014

[6] Singh, Navtej, Lisa-Marie Browne, and Ray Butler. 2013. "Parallel astronomical data processing with Python: Recipes for multicore machines." Astronomy and Computing 2: 1-10.

[7] Van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T. scikit-image: image processing in Python. PeerJ. 2014 Jun 19;2:e453.

Об авторах

Михаил Сергеевич Копылов, ст. инженер, ИПМ им. М.В. Келдыша РАН. E-mail: pmk@gin.keldysh.ru.

Елисей Дмитриевич Бирюков, м.н.с., ИПМ им. М.В. Келдыша РАН. E-mail: peb@gin.keldysh.ru.

Борис Хаимович Барладян, к.т.н., с.н.с., ИПМ им. М.В. Келдыша РАН. E-mail: obb@gin.keldysh.ru.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Автоматизированное упрощение машиностроительных CAD-моделей и сборок без использования истории построения

С.Е. Сляднев, А.С. Малышев, В.Е. Турлапов
sergey.slyadnev@gmail.com | al.s.malyshev@gmail.com | vadim.turlapov@itmm.unn.ru
Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия

Дается обзор методов автоматизированного упрощения CAD-моделей, включая одиночные детали и сборки. Модели предполагаются заданными в открытом формате без истории построения, что делает рассматриваемые методы независимыми от конкретных систем проектирования. Упрощение геометрии сопровождается приведением в непротиворечивое состояние ассоциированных метаданных, таких как имена, структура сборки, допуски, графические атрибуты и проч. Предложенные алгоритмы могут использоваться для эффективного обмена данными между инженерными коллективами, преобразования конструкторских моделей в расчетные, подготовки сцен виртуальной и дополненной реальности, а также в других приложениях, где требуется упрощение существующих геометрических объектов.

Ключевые слова: *упрощение CAD-моделей, оптимизация цифровых двойников, геометрическое моделирование, прямое редактирование, Analysis Situs.*

Automated history-free simplification of mechanical CAD models and assemblies

S.E. Slyadnev, A.S. Malyshev, V.E. Turlapov
sergey.slyadnev@gmail.com | al.s.malyshev@gmail.com | vadim.turlapov@itmm.unn.ru
Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation

We present a review of methods for automatic and semiautomatic simplification of CAD models and assemblies. The digital representation of a product comes from an open format such as STEP. Therefore, the proposed approaches are independent of specific MCAD systems. CAD geometry is usually accompanied with engineering metadata, such as product names, assembly structure, visualization attributes, dimensions and tolerances, etc. The considered methods are implemented in a way to keep data consistent during the simplification process. The proposed algorithms are used in a commercial application for AV/VR scenes preparation, optimizing models for mechanical simulations, IP protection during collaboration, and in other engineering workflows which require simplification of existing digital mock-ups.

Keywords: *simplification of CAD models and assemblies, digital mock-up, geometric modeling, direct editing, Analysis Situs.*

1. Введение

Конструкторские модели машиностроительных изделий, как правило, максимально детализированы. Само их назначение состоит в том, чтобы дать исчерпывающее цифровое описание всех узлов и агрегатов проектируемого объекта. В то же время, многие инженерные работы не требуют высокой детализации, а, напротив, оказываются затруднены, либо попросту невозможны на оригинальных конструкторских моделях. Возникает потребность в упрощении геометрии (Рис. 1) при сохранении необходимой для дальнейших операций инженерной семантики: имен деталей, значений допусков, графических атрибутов и проч.

Практика показывает, что инженерные подразделения редко используют единственную систему автоматизированного проектирования, предпочитая инструментальное разнообразие для решения проектных задач. При этом известно, что обмен данными между разнородным программным обеспечением является существенной проблемой, которая в настоящее время далека от решения. Перенос модели из одной системы в другую нередко сопровождается потерей или искажением данных. Более того, широко используемые открытые форматы (STEP, IGES, JT и проч.), предпочтительные с точки зрения долговременного хранения информации, не содержат истории построения модели и ее конструктивных элементов в явном виде. Тем не менее, использование именно открытых форматов позволяет уйти из-под «опеки» тех или иных компаний-поставщиков коммерческого ПО. Следовательно, программные решения на базе открытых форматов можно считать наиболее общими, а в ряде случаев – единственно приемлемыми.

Мы предполагаем, что исходная модель задана в нейтральном формате STEP (ISO 10303). Отсутствие истории построения и конструктивных элементов (в явном виде) диктует руководящую идею для реализации операторов упрощения, выносимых в настоящий обзор. Этой идеей является так называемое «прямое редактирование», т.е. редактирование без использования истории.

1.1 Цель исследования

Упрощение модели задействует два уровня абстракции: уровень *сборки* и одиночной *детали.* Кроме того, в зависимости от цели упрощения, модификация геометрии может осуществляться либо на точном (граничном) представлении детали, либо на ее полигональном приближении. Достаточно богатая палитра инструментов делает возможным их комбинирование в «рецепты» упрощения произвольной сложности.

Мы даем обзор наиболее эффективных и востребованных операторов упрощения CAD-моделей. В качестве геометрического ядра, определяющего формат представления инженерных данных, нами выбрана библиотека OpenCascade [1]. Операторы упрощения, описываемые в настоящей статье, дополняют набор доступных операторов этой открытой библиотеки (Рис. 2).

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

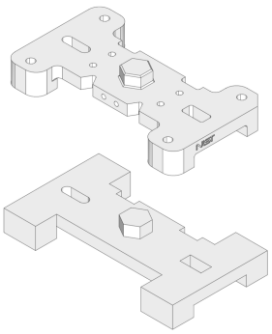


Рис. 1. Упрощение тестовой детали.

1.2 Другие обзоры

Авторы [2] дают подробную компиляцию техник упрощения CAD-моделей для инженерного анализа, причем обсуждаются различные варианты представления геометрической формы. В отличие от [2], мы рассматриваем исключительно граничное представление (B-Rep, boundary representation).

Работа [3] посвящена упрощению сборок, представляющих оборудование для комплектации сложных промышленных объектов, таких как корабли или нефтеплатформы. Ценность этой работы, на наш взгляд, состоит в том, что авторы предъявляют *критерии упрощения* моделей, полученные «из первых рук» по результатам интервьюирования глав инженерных подразделений. Развитие заявленных идей упрощения содержит статья тех же авторов [4]. Укажем, однако, что работы [3] и [4] основаны на использовании геометрического ядра ACIS [5], которое, хотя и содержит богатую алгоритмическую палитру, является коммерческим продуктом с ограниченным доступом для инженерных и академических сообществ.

Актуальность задачи упрощения сборок возрастает с проникновением в индустрию концепции «цифрового двойника» (Digital Mock-Up, DMU). Мы говорим о механических сборках, получаемых как «временной срез» из PLM-системы предприятия. DMU заменяет физический макет изделия, выполняя роль «справочной» модели, пригодной для выполнения некоторых видов инженерного анализа. В частности, речь идет о проверке сборочных допусков или анализе самой возможности собрать изделие из помещенных в DMU деталей. Подробный обзор, касающийся адаптации цифровых макетов для численного моделирования, содержит диссертация Ф. Буссюж [6]. Поскольку наша работа имеет непосредственное отношение к упрощению больших машиностроительных сборок, мы считаем работу [6] наиболее релевантной нашему исследованию. Дополнительная литература будет разобрана далее в соответствующих специальных разделах.

2. Основные определения

Конструктивный элемент (КЭ, «фичер») – набор граней, имеющих некоторый инженерный смысл. Часто под КЭ понимают грани, образованные одной операцией механообработки на станке с ЧПУ. Наиболее полное описание принципов проектирования в терминах конструктивных элементов можно найти в классической монографии [7].

«Немая» геометрия – граничное представление детали с утерянной историей построения. В немой геометрической модели отсутствует информация о конструктивных элементах («конструкторском замысле»). Как правило, немая модель есть результат преобразования данных из параметрической САПР в нейтральный формат.

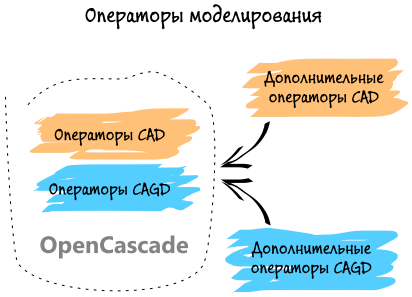


Рис. 2. Операторы упрощения дополняют функциональность выбранного геометрического ядра.

Метаданные -- инженерная семантика, связанная с геометрической моделью. Под метаданными понимают структуру сборки, имена деталей, графические атрибуты, технологические допуски и т.п.

Топологический граф -- ациклический ориентированный граф, представляющий вложенные друг в друга абстрактные топологические элементы, такие как вершины, ребра, грани, оболочки и т.п. Объектная модель геометрического ядра является топологическим графом сама по себе, будь то OpenCascade, ACIS, Parasolid или другая библиотека моделирования, реализующая традиционное разделение структур данных на чисто *геометрические* и *топологические*.

Атрибутированный граф смежности – вспомогательный неориентированный граф, вершины которого отвечают граням модели, а дуги представляют отношение смежности. С вершинами и дугами связываются произвольные пользовательские атрибуты. В нашей работе атрибутированный граф смежности является «носителем» конструктивных элементов геометрической модели.

Иерархический граф сборки – ациклический ориентированный граф, вершинами которого являются подсборки и одиночные детали, а дуги представляют отношение «вхождения» с некоторой локальной трансформацией.

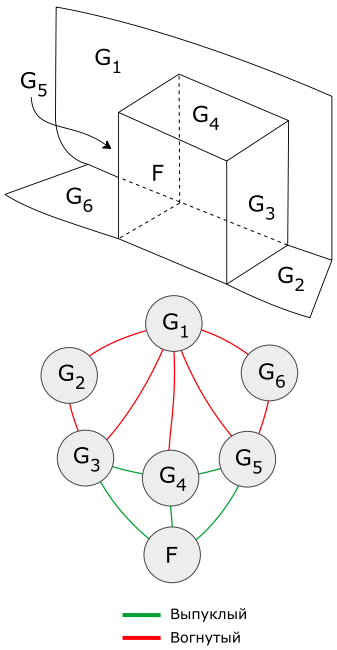


Рис. 3. Грани твердотельной модели и соответствующий атрибутированный граф смежности.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

3. Упрощение в контексте детали

Упрощение одиночной детали может заключаться в подавлении ее конструктивных элементов, либо в *идеализации* путем снижения размерности соответствующего топологического многообразия. Операторы идеализации в настоящем обзоре не фигурируют.

Подавление конструктивных элементов мы рассматриваем преимущественно как геометрическую задачу. Однако в общем случае нельзя упускать из внимания влияние операторов упрощения на результаты инженерного анализа. Во избежание потери точности моделирования, некоторые конструктивные элементы подавляться не должны. Работы [8] и [9], на наш взгляд, дают хорошее представление о том, как учесть вклад КЭ в результирующую точность.

Операторы упрощения детали реализуются в парадигме *распознавания конструктивных элементов,* то есть сопровождаются извлечением того или иного типа КЭ из немой геометрии. Исчерпывающий обзор основных подходов к распознаванию КЭ содержится в работах [10] и [11]. Наш подход состоит в абстрагировании вспомогательных структур данных из граничного представления модели и применении набора *правил* для поиска типовых конструктивных элементов по этим структурам. В классификации [10] такой подход относится к «графовым» с той оговоркой, что в нашей работе правила зафиксированы и не могут задаваться извне. Помимо подходов, основанных на графах, интерес представляют также методы объемной декомпозиции [12], поскольку они позволяют распознавать и подавлять сложные неизолированные КЭ.

Операторы упрощения имеют различную сложность в зависимости от типа конструктивного элемента и топологической реализации его вхождения в деталь (изолированное или с затрагиванием границ). Мы рассматриваем здесь два оператора упрощения: подавление изолированных конструктивных элементов и подавление цепочек скруглений. Математическим языком, адекватным в задачах распознавания КЭ, является язык теории графов. Все методы, представленные в данном разделе, оперируют с *атрибутированным графом смежности* (AAG, Attributed Adjacency Graph) граней [14], представляющим исходную деталь (Рис. 3). Ключевым атрибутом, связанным с дугами графа, является тип соответствующего двугранного угла: «выпуклый», «вогнутый» или «гладкое сопряжение». Мы рассматриваем также дополнительные атрибуты, связанные с вершинами графа. Эти атрибуты содержат признаки принадлежности той или иной грани некоторому КЭ. Таким образом, в результате процедуры распознавания исходный граф оказывается «обогащен» информацией о найденных в модели конструктивных элементах.

3.1 Изолированные КЭ

«Изолированность» означает, что конструктивный элемент реализован только через внутренние циклы «базовых граней» [15]. Удаление базовых граней изолированных КЭ приводит к увеличению числа компонент связности атрибутированного графа смежности.

3.1.1 Распознавание

Рассмотрим атрибутированный граф смежности граней G исходной детали. Требуется получить граф G^* как результат «обогащения» графа G атрибутами, выражающими принадлежность граней некоторым конструктивным элементам. На вход алгоритма распознавания поступает граф G и оператор P. Оператор P реализует проверку некоторого правила для окрестности вершины f в атрибутированном графе смежности. Некоторые эвристики для распознавания наиболее распространенных типов КЭ могут быть найдены в нашей предыдущей работе [16]. Правило P применяется последовательно ко всем вершинам f, выбранным в качестве кандидатов по некоторому признаку (например, могут быть выбраны все грани, имеющие внутренние циклы). В результате применения правила P происходит «обогащение» графа G новыми атрибутами. Кроме того, те грани, что были «посещены» оператором, исключаются из дальнейшего рассмотрения.

3.1.2 Подавление

Мы рассматриваем только *изолированные* конструктивные элементы поскольку их подавление может быть осуществлено на уровне топологии модели, то есть без применения вычислительно дорогостоящих и ненадежных операторов редактирования геометрии. Подавление изолированных конструктивных элементов осуществляется путем «иссечения» найденных граничных элементов из топологического графа модели (Рис. 4).

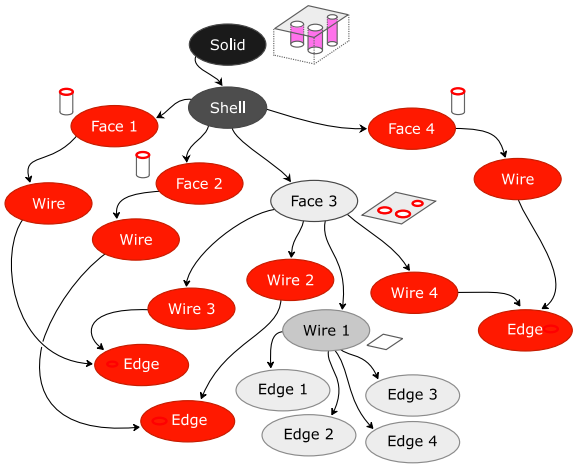


Рис. 4. Подавление изолированных конструктивных элементов путем «иссечения» вершин топологического графа. Красным отмечены вершины графа, удаление которых приводит к подавлению трех сквозных отверстий предполагаемой детали.

3.2 Цепочки скруглений

Подавление скруглений может существенно сказаться на сложности итоговой модели, особенно, если в качестве метрики сложности используется количество фасетов (см. Раздел 5). В нашей работе распознавание и подавление скруглений реализуется согласно принципам, изложенным в докладах [17] и [18]. Важной частью алгоритма подавления являются эйлеровы операторы [19], гарантирующие топологическую целостность итоговой геометрической модели. Применение эйлеровых операторов происходит сообразно результатам локального топологического анализа вблизи граничных элементов скруглений. Привлечение топологического анализа и аппарата эйлеровых операторов позволяет реализовать подавление скруглений как *локальную операцию* упрощения (см. главу 4 монографии [20] о сути «локальных операторов»). Дополнительные ценные сведения об упрощении граничных моделей путем подавления скруглений могут быть найдены в работах [21].

Пример подавления двух цепочек разного радиуса с вершинным скруглением показан на Рис. 5. Ниже приведен листинг соответствующей процедуры для выполнения в среде «Анализ Положения» [24].

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

```
init-naming

kfmv -name face_8
kev -edge edge_2 edge_18 edge_12
kef -face face_3 -kedge edge_8 -sedge edge_9
kef -face face_10 -kedge edge_22 -sedge edge_16
kef -face face_7 -kedge edge_19 -sedge edge_7

rebuild-edge -name edge_1 edge_4 edge_8
                   edge_14 edge_16 edge_7
                   edge_6 edge_11 edge_17
```

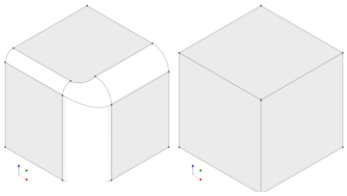


Рис. 5. Подавление цепочки скруглений разных радиусов.

4. Упрощение в контексте сборки

Для представления сборок мы используем реализацию структуры данных типа *иерархический граф сборки* (HAG, Hierarchical Assembly Graph). Описание этой структуры данных было предложено А. Раппопортом в 1993 году [25]. HAG – направленный ациклический граф, вершинами которого являются детали или подсборки, а дуги выражают отношение вхождения (instancing). Другие отношения между компонентами сборки, такие как сопряжения, мы не рассматриваем в силу известной ограниченности нейтральных форматов, используемых для передачи инженерных данных. Для выражения прочих отношений предлагается использовать расширенные графы сборок [26], либо дополнительные структуры данных.

А. Раппопорт отмечает важную специфику HAG-представления сборок, состоящую в неоднозначности привязки метаданных к ее элементам. Действительно, различные варианты вхождения одной детали в сборку представлены не вершинами и дугами, а путями в графе. На Рис. 6 показана тестовая сборка **l-bracket-assembly**, содержащая три вхождения подсборки **nut-bolt-assembly**, состоящей из болта **bolt** и гайки **nut.** Соответствующий иерархический граф сборки показан на Рис. 7.

В отличие от работы [25], где вхождения соответствуют дугам, для представления вхождений мы используем вершины графа (узлы нашей вариации HAG отвечают объектам **next_assembly_usage_occurence** формата STEP). Из Рис. 7 видно, что привязка метаданных к конкретной вершине или дуге графа не позволяет дифференцировать отдельные вхождения болтов и гаек. Для однозначной адресации вхождений детали или подсборки следует использовать полные пути графа, набранные из уникальных идентификаторов вершин.

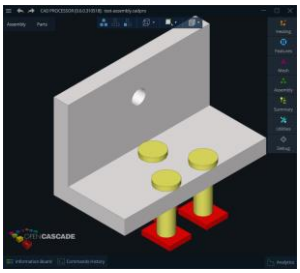


Рис. 6. Тестовая сборка **l-bracket-assembly,** содержащая три вхождения подсборки **nut-bolt-assembly,** представленной болтом **bolt** и гайкой **nut.**

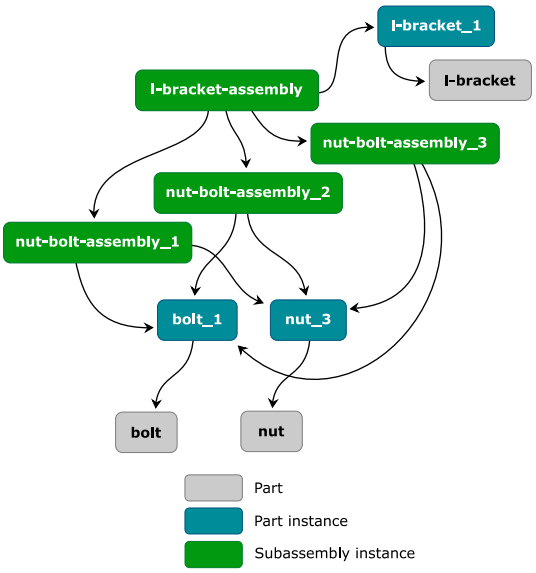


Рис. 7. Граф тестовой сборки **l-bracket-assembly.**

4.1 Удаление невидимых частей

Эффективным средством упрощения сборок является удаление ее невидимых частей. Эта операция востребована в ситуациях, когда интерес представляет только внешняя форма продукта, например, при обмене данными между инженерными подразделениями, при проверке коллизий, а также для упрощения трехмерной модели, погружаемой в контекст более сложного продукта для проектирования сопряженных узлов и агрегатов. Реализация данного оператора основана на испускании лучей с поверхности каждой детали и подсчете количества пересечений с другими компонентами сборки.

4.2 Замена детали примитивом

Проектирование сборки «сверху-вниз» предполагает, что точная геометрическая форма изделия определяется не сразу, но на более поздних этапах работы конструктора [27]. Действительно, первичными являются функциональные требования к продукту, позволяющие начать проектирование еще без детальной геометрической проработки. Оригинальные мысли относительно роли геометрии в машиностроительном проектировании изложены в замечательном эссе [29]. Упрощение сборки является процессом, в некотором смысле обратным проектированию «сверху-вниз». От точного описания формы совершается переход к грубому геометрическому представлению «абстрактного узла» в контексте сборки, которое часто оказывается достаточным для решения прикладных задач.

Простейшим способом перейти от точной формы детали к абстрактной является замена ее исходной геометрической модели примитивом, таким как параллелепипед, цилиндр, шар и т.п. При этом можно сохранить ориентацию детали в пространстве, используя, например, ее главные оси инерции.

4.3 Слияние вхождений

Граф сборки – эффективная структура данных, поскольку идентичная геометрия деталей в ней не дублируется, а реализуется посредством *вхождений* (instances). На практике, однако, встречаются сборки, представленные неоптимальным образом. В худшем случае, иерархический граф может оказаться вырожденным

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

в дерево, где каждое вхождение детали реплицирует уже существующую геометрию.

Возникает задача, обратная задаче *выделения вхождения* (instance singling), рассмотренной А. Раппопортом [25]. Имея две независимые детали, требуется проверить их идентичность и, в случае таковой идентичности, осуществить *слияние вхождений* с заново рассчитанными матрицами трансформаций. Сравнение моделей осуществляется при помощи простых геометрических эвристик, таких как кардинальное число топологического графа, типы граничных элементов и т.п. Для расчета матрицы трансформации используется сингулярное разложение [30].

5. Переход к сеточным моделям

Мы выделяем два класса поверхностной триангуляции: *фасеты* для визуализации и *сетки* расчетного качества. Фасеты распределены по граням геометрической модели детали, тогда как расчетная сетка теряет эту ассоциативность и помещается в сборке на том же уровне, что и точное представление (Рис. 8).

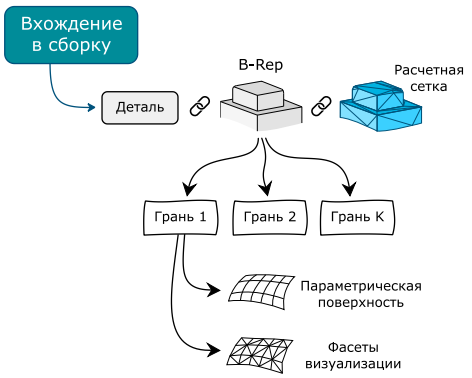


Рис. 8. Принцип организации сеточных данных.

Переход от граничных моделей к сеточным открывает дополнительные возможности для упрощения. Например, на граничных моделях можно осуществить подавление конструктивных элементов, после чего сконструировать расчетную сетку и продолжить упрощение с использованием полигональных операторов, таких как децимация [31] или сглаживание по Лапласу [34].

5.1 Создание расчетной сетки

Конструирование сетки по граничному представлению CAD-модели – стандартная функция геометрического ядра. Теория триангуляции описана в таких монографиях как [35] и [36]. Также существуют открытые программные компоненты, позволяющие создавать расчетные сетки из граничных моделей ядра OpenCascade [37].

5.2 Оборачивание

Оборачивание – специфический способ триангуляции граничного или фасетного представления CAD-модели (сборки). Оборачивание сглаживает дефекты и особенности геометрической модели, требуя лишь, чтобы она описывала некоторый замкнутый объем в пространстве. Этот оператор часто используется для подготовки упрощенных моделей к аэродинамическому анализу. На Рис. 9 представлен результат оборачивания сборки, состоящей из 140 уникальных деталей. Подробно оператор оборачивания описан в работах [39] и [40].

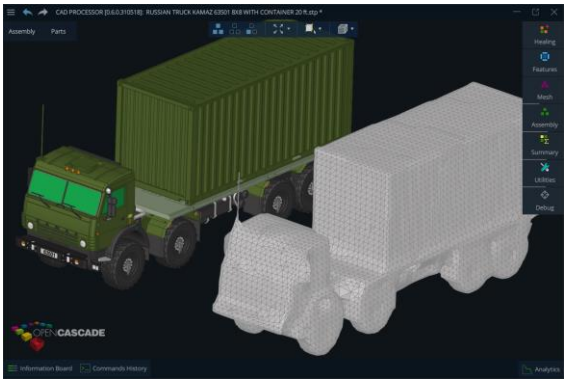


Рис. 9. Результат оборачивания модели «КАМАЗа». Сборка предоставлена Антонино Раписадра (Rapisarda Antonino) в библиотеке GrabCAD (www.grabcad.com).

5.3 Сеточное представление проводов

Некоторые типы деталей не допускают упрощения путем подавления конструктивных элементов или заменой примитивом. К такому классу можно отнести кинематические поверхности, моделирующие, например, жгуты с поперечным сечением в виде окружности. Упрощение подобных моделей, как правило, реализуется понижением количества фасетов [41]. С другой стороны, «наивные» методы, состоящие в перестроении сетки с заведомо грубыми параметрами точности, как правило, дают неадекватные результаты (Рис. 10).

В работе [42], авторы решают задачу оптимизации сетки с привлечением таких топологических операторов, как **edge-collapse, edge-split** и **edge-swap.** Применение оператора полагается разрешенным, если результат его действия не изменяет топологический тип данной сетки. В частности, не допускается возникновение самопересечений. В отличие от традиционных техник децимации [31], алгоритм Хоппа [42] конструирует сетку, максимально близкую к исходной, имеющую притом существенно меньшую сложность и тот же топологический тип. Метод Хоппа применим для любых геометрических моделей. Однако для кинематических форм, моделирующих электрические провода или гибкие трубки, можно рассмотреть специальные операторы упрощения, дающие больше контроля над результатом.

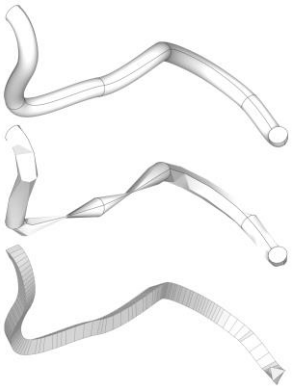


Рис. 10. Варианты геометрического представления провода. Сверху-вниз: оригинальное граничное представление; «грубая» фасетизация с самопересечениями; полиэдральная модель, построенная нашим алгоритмом.

Суть метода состоит в переходе от исходных поверхностей свободной формы к полиэдральному представлению (Рис. 10), фасетная модель которого создается путем разбиения каждой «косой плоскости» по диагонали.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

6. Заключение и план исследований

Описанные подходы для упрощения деталей и сборок были реализованы в коммерческом программном обеспечении [43]. В качестве среды для прототипирования большинства алгоритмов применена свободно распространяемая программа «Анализ Положения» [24], показавшая свою эффективность.

Укажем основные направления наших дальнейших исследований в области упрощения геометрических моделей.

6.1 Расширенное распознавание

Планируется расширить модуль распознавания КЭ, перейдя от ограниченного набора правил к формальным грамматикам, как описано в докладе Веткатарамана [15]. Расширенный модуль принимает в качестве исходных данных символьное описание конструктивного элемента, используемое универсальным алгоритмом поиска подграфа в атрибутированном графе смежности. Следующий этап состоит в проверке изолированности КЭ с целью применения оператора топологической редукции, как описано в Разделе 3.1.2.

Во многих сценариях автоматического упрощения требуется распознавание не конструктивных элементов, но специфических типов самих деталей. Так, для применения оператора упрощения проводов (см. Раздел 5.3) требуется предварительно их отыскать среди всех элементов сборки.

6.2 Прямое редактирование

Особый интерес представляет расширение функциональности упрощения механизмами *прямого редактирования* [45] для обеспечения гибкой интерактивной модификации моделей.

6.3 Упрощение сеток

Некоторые геометрические модели изначально лишены точного граничного представления. Для их упрощения требуется разработка дополнительных операторов, таких, например, как оператор оптимизации сеток [42].

6.4 Подготовка данных

Использование операторов упрощения предполагает, что исходная модель лишена геометрических и структурных патологий. Однако опыт показывает, что передача данных между различными инженерными системами нередко сопровождается нарушением целостности модели. Возникает задача разработки специальных операторов для «исправления» модели с целью ее дальнейшего редактирования. Данной проблеме посвящены такие работы, как [46], а также теоретические исследования [50]. Отправной точной для этих работ является фундаментальный труд А. Рекуича [53].

6.5 Метрики процесса упрощения

Подход, описанный в Разделе 3, может применяться для подготовки CAD-моделей к инженерному анализу. В этом случае подавление КЭ должно сопровождаться оценкой ошибки, которую вносит оператор упрощения в результат расчета. Те КЭ, удаление которых приводит к существенному понижению точности моделирования, должны специальным образом помечаться. Трудность состоит в том, чтобы дифференцировать все КЭ, имея в виду их возможные взаимопроникновения. В этом отношении перспективной представляется работа [54], расширяющая «графовый» метод для нахождения неизолированных фичеров.

7. Благодарности

Авторы признательны руководству компании OPEN CASCADE за предоставленную возможность публикации результатов внутренних исследований.

8. Литература

[1] S. Slyadnev, Обзор Open CASCADE Technology (2014). URL http://isicad.ru/ru/articles.php?article_num=17367

[2] A. Thakur, A. G. Banerjee, S. K. Gupta, A survey of CAD model simplification techniques for physics-based simulation applications, Computer-Aided Design 41 (2) (2009) 65-80.

[3] S. Kwon, B. C. Kim, D. Mun, S. Han, Simplification of feature-based 3D CAD assembly data of ship and offshore equipment using quantitative evaluation metrics, Computer-Aided Design 59 (2015) 140-154.

[4] S. Kwon, B. C. Kim, D. Mun, S. Han, User-assisted integrated method for controlling level of detail of large-scale B-Rep assembly models, International Journal of Computer Integrated Manufacturing (2018) 1-12.

[5] J. Corney, T. Lim, 3D modeling with the ACIS Toolkit, Saxe-Coburg Publications, 2001.

[6] F. Boussuge, Idealization of CAD assemblies for FE structural analyses (2006).

[7] J. J. Shah, M. Mantyla, Parametric and feature-based CAD/CAM: concepts, techniques, and applications, 1st Edition, John Wiley & Sons, Inc., New York, NY, USA, 1995.

[8] M. Li, S. Gao, Estimating defeaturing-induced engineering analysis errors for arbitrary 3D features, Computer-Aided Design 43 (12) (2011) 1587-1597.

[9] R. Ferrandes, P. Marin, J.-C. Leon, F. Giannini, A posteriori evaluation of simplification details for finite element model preparation, Computers & Structures 87 (1-2) (2009) 73-80.

[10] J. H. J. Han, M. Pratt, W. Regli, Manufacturing feature recognition from solid models: a status report, IEEE Transactions on Robotics and Automation 16 (6) (2000) 1-31.

[11] J. J. Shah, D. Anderson, Y. S. Kim, S. Joshi, A discourse on geometric feature recognition from CAD models, Journal of Computing and Information Science in Engineering 1 (1) (2001) 41.

[12] С. Сляднев, В. Турлапов, Метод декомпозиции машиностроительных твердотельных моделей на элементы объема изъятия: GraphiCon 2016, Нижний Новгород, 2016, pp. 58-63.

[13] B. C. Kim, D. Mun, Feature-based simplification of boundary representation models using sequential iterative volume decomposition, Computers and Graphics (Pergamon) 38 (1) (2014) 97-107.

[14] S. Joshi, T. C. Chang, Graph-based heuristics for recognition of machined features from a 3D solid model, Computer-Aided Design 20 (2) (1988) 58-66.

[15] S. Venkataraman, M. Sohoni, V. Kulkarni, A graph-based framework for feature recognition, in: Proceedings of the sixth ACM symposium on Solid modeling and applications - SMA '01, ACM Press, New York, New York, SA, 2001, pp. 194-205.

[16] A. Malyshev, S. Slyadnev, V. Turlapov, Graph-based feature recognition and suppression on solid models, in: GraphiCon'17 proceedings, 2017, pp. 319-322.

[17] S. Venkataraman, M. Sohoni, G. Elber, Blend recognition algorithm and applications, in: Proceedings of the sixth ACM symposium on Solid modeling and applications - SMA '01, ACM Press, New York, New York, USA, 2001, pp. 99-108.

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

[18] S. Venkataraman, M. Sohoni, R. Rajadhyaksha, Removal of blends from boundary representation models, in: Proceedings of the seventh ACM symposium on Solid modeling and applications - SMA ’02, ACM Press, New York, New York, USA, 2002.
[19] Mantyla, Sulonen, GWB: A solid modeler with Euler operators, IEEE Computer Graphics and Applications 2 (7) (1982) 17-31.
[20] I. Stroud, H. Nagy, Solid modelling and CAD systems, Springer London, London, 2011.
[21] H. Zhu, C. Menq, B-Rep model simplification by automatic fillet/round suppressing for efficient automatic feature recognition, Computer-Aided Design 34 (2) (2002) 109-123.
[22] X. Cui, S. Gao, G. Zhou, An efficient algorithm for recognizing and suppressing blend features, Computer-Aided Design and Applications 1 (1-4) (2004) 421-428.
[23] J.-Y. Lai, C. Wong, T. T. Huynh, M.-H. Wang, C.-H. Hsu, Y.-C. Tsai, C.-Y. Huang, Small blend suppression from B-Rep models in computer-aided engineering analysis, Journal of the Chinese Institute of Engineers 39 (6) (2016) 735-745.
[24] S. Slyadnev, A. Malyshev, V. Turlapov, CAD model inspection utility and prototyping framework based on OpenCascade, in: GraphiCon’17 proceedings, 2017, pp. 323-327.
[25] A. Rappoport, A scheme for single instance representation in hierarchical assembly graphs, in: B. Falcidieno, T. L. Kunii (Eds.), Modeling in Computer Graphics, Springer Berlin Heidelberg, Berlin, Heidelberg, 1993, pp. 213-223.
[26] K. Lee, D. C. Gossard, A hierarchical data structure for representing assemblies: part 1, Computer-Aided Design 17 (1) (1985) 15-19.
[27] M. Mantyla, A modeling system for top-down design of assembled products, IBM Journal of Research and Development 34 (5) (1990) 636-659.
[28] E. C. Libardi, J. R. Dixon, M. K. Simmons, Computer environments for the design of mechanical assemblies: A research review, Engineering with Computers 3 (3) (1988) 121-136.
[29] V. Shapiro, H. Voelcker, On the role of geometry in mechanical design, Research in Engineering Design 1 (1) (1989) 69-73.
[30] O. Sorkine-Hornung, M. Rabinovich, Least-squares rigid motion using SVD.
[31] W. J. Schroeder, J. A. Zarge, W. E. Lorensen, Decimation of triangle meshes, SIGGRAPH Comput. Graph. 26 (2) (1992) 65-70.
[32] M. Knapp, Mesh decimation using VTK (2002) 1-8.
[33] L. Kobbelt, S. Campagna, H.-P. Seidel, A general framework for mesh decimation, Graphics Interface (1998) 43-50.
[34] A. Nealen, T. Igarashi, O. Sorkine, M. Alexa, Laplacian mesh optimization, SIGGRAPH (2006) 381.
[35] P. J. Frey, P.-L. George, Mesh generation: application to finite elements, ISTE, 2007.
[36] S. Cheng, T. Dey, J. Shewchuk, Delaunay mesh generation, Chapman & Hall/CRC Computer and Information Science Series, CRC Press, 2016.
[37] J. Schoberl, NETGEN: an advancing front 2D/3D-mesh generator based on abstract rules, Computing and Visualization in Science 1 (1) (1997) 41-52.
[38] C. Geuzaine, J.-F. Remacle, Gmsh: a three-dimensional finite element mesh generator with built-in pre-and post-processing facilities, International Journal for Numerical Methods in Engineering 79(11) 0 (2009) 1309-1331. arXiv:1010.1724.
[39] Y. K. Lee, C. K. Lim, H. Ghazialam, H. Vardhan, E. Eklund, Surface mesh generation for dirty geometries by the Cartesian shrink-wrapping technique, Engineering with Computers 26 (4) (2010) 377-390.
[40] A. Malyshev, A. Zhidkov, V. Turlapov, Adaptive mesh generation using shrink-wrapping approach, in: GraphiCon’18 proceedings, 2018, preprint.
[41] Z. Qiu, Y. Wong, J. Fuh, Y. Chen, Z. Zhou, W. Li, Y. Lu, Geometric model simplification for distributed CAD, Computer-Aided Design 36 (9) (2004) 809 - 819.
[42] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, W. Stuetzle, Mesh optimization, in: Proceedings of the 20th annual conference on Computer graphics and interactive techniques - SIGGRAPH ’93, Vol. d, ACM Press, New York, New York, USA, 1993, pp. 19-26.
[43] OPEN CASCADE, Digital mock-up simplification. URL https://www.opencascade.com/content/digital-mockup-simplification
[44] OPEN CASCADE, CAD Processor. URL https://www.opencascade.com/content/cad-processor
[45] K.-P. Fahlbusch, T. D. Roser, HP PE/SolidDesigner: dynamic modeling for three-dimensional Computer-Aided Design 46 (1995) 6-13.
[46] F. Frischmann, Topological and geometric healing on solid models, Master’s thesis (2011).
[47] J. Qi, V. Shapiro, epsilon-Solidity in Geometric Data Translation, Tech. rep., Spatial Automation Laboratory, University of Wisconsin-Madison (2004).
[48] C. Hoffmann, V. Shapiro, V. Srinivasan, Geometric interoperability via queries, Computer-Aided Design 46 (1) (2014) 148-159.
[49] J. Qi, V. Shapiro, Geometric interoperability with epsilon solidity, Journal of Computing and Information Science in Engineering 6 (3) (2006) 213.
[50] J. Qi, V. Shapiro, Epsilon-regular sets and intervals, in: International Conference on Shape Modeling and Applications 2005 (SMI’ 05), Spatial Automation Laboratory, University of Wisconsin-Madison, IEEE Comput. Soc, 2004, pp. 308-317.
[51] J. Qi, V. Shapiro, epsilon-Topological formulation of tolerant solid modeling, Computer-Aided Design 38 (4) (2006) 367-377.
[52] V. Shapiro, Reliable implementation of real number algorithms: theory and practice, Springer-Verlag, Berlin, Heidelberg, 2008, Ch. Homotopy conditions for tolerant geometric queries, pp. 162-180.
[53] A. G. Requicha, Representations for rigid solids: theory, methods, and systems, ACM Computing Surveys 12 (4) (1980) 437-464.
[54] S. Gao, J. Shah, Automatic recognition of interacting machining features based on minimal condition subgraph, Computer-Aided Design 30 (9) (1998) 727-739.

Об авторах

Турлапов Вадим Евгеньевич, д.т.н., профессор кафедры математического обеспечения и суперкомпьютерных технологий ННГУ им.Н.И.Лобачевского,
vadim.turlapov@itmm.unn.ru

Малышев Александр Сергеевич, аспирант 2 года,
ИИТММ ННГУ им.Н.И.Лобачевского,
email: al.s.malyshev@gmail.com

Сляднев Сергей Евгеньевич, аспирант 3 года,
ИИТММ ННГУ им.Н.И.Лобачевского,
email: sergey.slyadnev@gmail.com

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Реконструкция параметрической модели лопатки турбины из набора аэродинамических сечений с использованием техники сглаживания

С.Е. Сляднев, В.Е. Турлапов
sergey.slyadnev@gmail.com | vadim.turlapov@itmm.unn.ru
Национальный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Нижний Новгород, Россия

Описан подход к восстановлению параметрической твердотельной модели лопатки турбины из заданного набора аэродинамических сечений. Новизна подхода состоит в применении техники сглаживания при построении поверхности пера лопатки. Промежуточный этап сглаживания позволяет избавиться от геометрических осцилляций, возникающих из-за неточности во входных данных. Предложенный метод находит применение в обратном инжиниринге трехмерных моделей турбомашин.

Ключевые слова: *обратный инжиниринг, геометрическое моделирование, турбостроение, сглаживание, параметрическая модель, Analysis Situs, OpenCascade.*

Reverse engineering of turbine blades from airfoil sections using curve fairing operator

S.E. Slyadnev, V.E. Turlapov
sergey.slyadnev@gmail.com | vadim.turlapov@itmm.unn.ru
Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation

We present an enhanced method for reconstructing parametric solid models of turbine blades from the given set of airfoil sections. The novelty of our approach is using a fairing operator as an intermediate stage of reconstruction. The curve fairing operator eliminates high-frequency geometric oscillations induced by noisy input data. The proposed method can be used for reverse engineering of CAD models in turbomachinery field.

Keywords: *reverse engineering, geometric modeling, turbomachinery, fairing, parametric model, Analysis Situs, OpenCascade.*

1. Введение

Типовой цикл проектирования лопатки турбомашины состоит в последовательном подборе параметров изделия до тех пор, пока не будут достигнуты целевые эксплуатационные качества при допустимой стоимости изготовления. Как правило, изменение параметров модели происходит сообразно результатам промежуточных аэрогидродинамических расчетов. Поведение потока, предсказанное расчетом, используется для оптимизации формы через посредство имеющихся переменных проектирования. На Рис. 1 показан классический цикл оптимизации [1].

Традиционно, моделирование лопатки начинается с параметризации ее аэродинамического профиля (сечения). «Стопка» сечений формирует перо лопатки, чья эталонная геометрия может быть получена методом скиннинга [2]. Альтернативой такому моделированию «снизу-вверх» является восстановление формы лопатки из имеющихся данных о форме реального изделия или потоке вблизи него. Так, информация о потоке используется в задачах обратного (инверсного) проектирования. Цель инверсного проектирования состоит в получении геометрии лопатки исходя из желаемого поведения потока. Как в прямой, так и в обратной задаче, параметрическая модель изделия есть необходимый ингредиент общего процесса проектирования. Параметрическая модель выступает в роли «поставщика» эталонной формы изделия для данного набора переменных проектирования $\mathbf{a} \in \mathbb{R}^N$ (N – количество переменных). Из эталонной формы, в свою очередь, будет построена расчетная сетка для численного моделирования физики процесса. Настоящая работа направлена на создание геометрического инструментария для эффективной реконструкции трехмерных параметрических моделей турбомашин из облака точек.

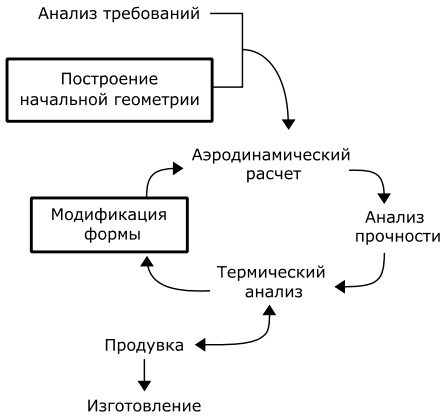


Рис. 1. Традиционный цикл оптимизации модели турбомашины. Жирным обведены подзадачи, имеющие чисто геометрическую подоплеку и, потому, находящиеся в фокусе нашего внимания.

Несмотря на то, что задача геометрического моделирования компонент турбомашин считается многократно решенной, мы видим существенные пробелы в ранее опубликованных работах и доступном программном обеспечении. Во-первых, обратный инжиниринг, хотя и предполагает использование «джентльменского набора» геометрических операторов, но требует специальных подходов для борьбы с неточностью входных данных. Поэтому классические «прямые» операторы моделирования применимы к задачам реконструкции лишь с оговорками. Во-вторых, восстановлению подлежит не конкретная геометрическая модель, но параметрическое семейство таких моделей $\mathrm{S}(\mathbf{a})$, сохраняющее в своей структуре *замысел проектировщика*. Большинство же методов, публикуемых по поводу реконструкции, ограничивается воссозданием «немой» формы изделия. В-

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

третьих, для эффективной параметрической оптимизации недостаточно иметь только лишь параметрическую модель. При высокой размерности пространства поиска $N \gg 1$ эффективными становятся градиентные методы оптимизации [3], требующие вычисления вектора $\nabla S(\mathbf{a})$. Компоненты вектора ∇S есть частные производные формы по параметрам проектирования. Такие производные называют «чувствительностью» модели, поскольку они отражают вклад каждой переменной проектирования в итоговую форму. В данной работе делаются первые шаги по направлению к созданию геометрического ядра системы, предоставляющей интерфейс, схематично изображенный на Рис. 2. Создание подобных систем геометрического моделирования является актуальной задачей, широко востребованной промышленностью.

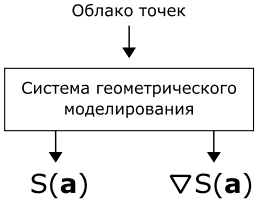


Рис. 2. Интерфейсы целевой системы геометрического моделирования. Для эффективной параметрической оптимизации целевая система должна возвращать не только параметрическую модель, но и ее градиент (в общем случае – производные высших порядков).

Из практики замечено, что в «грамотно» спроектированной модели $S(\mathbf{a})$ варьирование вектора параметров $\mathbf{a}$ сохраняет не только геометрическую, но и концептуальную целостность формы. Говорят, что такая модель отражает *замысел проектировщика*. Автоматическая реконструкция – сложная задача, во многом из-за того, что требуется *алгоритмически* распознать и воссоздать этот замысел.

Как справедливо указано в [4], чтобы обратный инжиниринг вообще имел смысл, необходимо привлечь в процессе реконструкции как можно больше инженерного знания не о форме, но о самой структуре модели. В настоящей работе мы накладываем дополнительные ограничения на входные данные, требуя организации облака точек по продольным сечениям лопатки. Это допущение позволяет зафиксировать структуру параметрической модели, обеспечивая, тем самым, соблюдение наперед заданной конструктивной концепции. Заметим, что указанное ограничение не отменяет общности самих геометрических операторов, реализованных в рамках настоящего исследования.

Мы описываем процесс реконструкции элементов рабочего колеса турбины, состоящего из основания, набора лопаток и обода (Рис. 3).

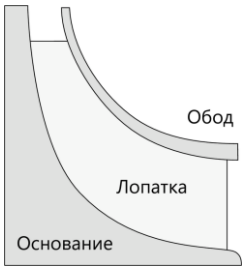


Рис. 3. Меридиональная проекция рабочего колеса турбины с лопаткой и ободом.

Статья организована следующим образом. Раздел 2 содержит обзор текущего положения дел в области реинжиниринга применительно к турбостроению. Формат исходных данных кратко описан в Разделе 3, где вводятся основные ограничения метода. В Разделе 4 изложены этапы реконструкции модели. Первостепенное внимание уделено оператору скиннинга со сглаживанием. Раздел 5 содержит промежуточные выводы и указывает направление дальнейших исследований.

2. Обзор подходов

Цель реинжиниринга состоит в преобразовании геометрического «артефакта» в инженерную концепцию [4]. Реинжиниринг лопаток турбомашин затруднен тем, что восстановлению подлежат поверхности свободной формы. Как правило, такие поверхности получают оператором скиннинга, то есть путем интерполяции аэродинамических сечений лопатки. В процессе реконструкции требуется принять решение о параметризации сечения и зафиксировать, тем самым, конструктивную концепцию восстанавливаемой модели.

Способов параметризации аэродинамического профиля существует множество. Так, авторы [4] используют последовательность дуг окружностей для представления формы сечений. Другие варианты включают явное использование B-кривых, распределение толщины пера вдоль срединной линии, а также комбинированные приемы [5]. Параметризация с использованием срединной линии предпочтительна в циклах оптимизации, поскольку она позволяет избежать самопересечений сторон лопатки в процессе поиска оптимальных значений параметров. В то же время существуют подходы для реконструкции пера лопатки непосредственно из облака точек, минуя этап восстановления сечений [6]. Мы не рассматриваем такие подходы, поскольку они не позволяют «ухватить» конструктивную концепцию детали, а значит параметрическая оптимизация будет существенно затруднена. По аналогии с тезисом из [4] заметим, что чисто геометрический подход к промышленному реинжинирингу есть всего лишь преобразование одного геометрического «артефакта» в другой.

В работе [7] дается модифицированный вариант оператора скиннинга, в котором вместо интерполяции авторы задействуют аппроксимацию методом наименьших квадратов. Анализ распределения кривизн показывает, что результирующая форма лопатки оказывается более гладкой. Кроме того, авторы предлагают использование скалярных энергетических метрик для оценки качества итоговой геометрии. Наш подход близок к работе [7] с той разницей, что вместо аппроксимации мы используем оператор сглаживания.

Оператор сглаживания как составная часть скиннинга был реализован также в работе [8]. Метод начинается с выбора общего узлового вектора кривых-сечений так, чтобы задача интерполяции оставалась недоопределенной. Затем контрольные точки рассчитываются путем минимизации квадратичного функционала энергии кривых, что позволяет понизить сложность итоговой поверхности в сравнении с традиционными методами скиннинга.

3. Исходные данные

Алгоритм реконструкции принимает упорядоченный набор аэродинамических сечений, распределенных по предполагаемым поверхностям вращения (Рис. 4). Согласно [1], поверхность вращения дает хорошее приближение к реальному поведению потока. В этом смысле, точки, образующие сечения, можно понимать как дискретно заданные линии тока, огибающие перо реконструируемой лопатки. Каждый профиль лопатки определен в меридиональной плоскости $r - z$. Две дополнительные угловые координаты φ_1 и φ_2

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

«размыкают» меридиональную кривую, формируя обвод лопатки. Мы предполагаем, что точки каждого сечения изначально отсортированы вдоль линии тока. В противном случае, требуется использовать вспомогательные техники сортировки, например, изложенную в [9].

4. Алгоритм реконструкции

4.1 Интерполяция сечений

На первом этапе дискретно заданные сечения лопатки интерполируются для получения трехмерных кривых. Для выбора параметров интерполируемых точек используется широко известная *центростремительная параметризация* [10], позволяющая сгладить незначительные осцилляции формы при «извилистом» распределении исходных точек. Спина и корытце каждого сечения интерполируются по отдельности для повышения точности в зонах со значительным изменением кривизны. После этого кривые стыкуются с порядком гладкости C^2 путем удаления узлов соответствующей объединенной B-кривой.

Как показано в [2], оператор скиннинга на независимо подготовленных сечениях приводит к «взрывному» росту количества контрольных точек в результирующей B-поверхности. Такие поверхности избыточны для моделирования искомой формы, и, кроме того, чрезмерная сложность их представления негативно сказывается на производительности и надежности последующих операторов моделирования (например, построения скруглений). Причиной квадратичного роста сложности поверхности с увеличением количества сечений является алгоритм унификации узловых векторов A5.4, описанный в Книге NURBS [11]. Однако, если сечения интерполировать на одном узловом векторе, выбранном изначально, то унификация не потребуется [7].

В результате интерполяции получаем кривые $\mathbf{c}_k(u)$, $k = \overline{0,K}$, где $(K+1)$ – количество сечений лопатки.

4.2 Скиннинг со сглаживанием

Оператор скиннинга широко используется для протягивания поверхности через упорядоченный набор B-кривых $\mathbf{c}_k(u)$ степени p, заданных на едином узловом векторе $\mathbf{U}$. Каждая кривая имеет следующий вид:

$$\mathbf{c}_k(u) = \sum_{i=0}^{n} N_{i,p}(u)\mathbf{P}_{i,k}$$

Здесь $k = \overline{0,K}$; $N_{i,p}(u)$ – базисные сплайны степени p; $(m+1)$ – количество узлов с повторениями; $(n = m + p + 1)$ – индекс последней контрольной точки $\mathbf{P}_{i,k}$.

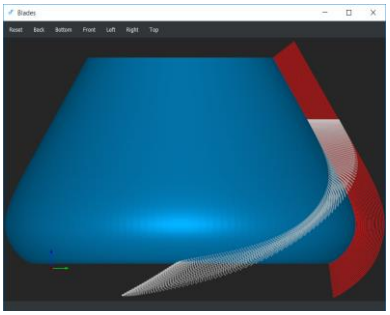


Рис. 4. Исходные данные, содержащие «стопку» сечений, представленных упорядоченными наборами точек. Каждое сечение определено на собственной несущей поверхности вращения (на рисунке показана только нижняя поверхность). Образующие кривые показаны красным цветом.

Из теории [11] известно, что контрольные точки результирующей поверхности скиннинга образуются путем интерполяции контрольных точек $\mathbf{P}_{i,k}$ кривых-сечений по направлению их следования друг за другом. Этот факт вытекает из требования совпадения изолиний поверхности с опорными кривыми, т.е. $\mathbf{s}(u, v_k) = \mathbf{c}_k(u)$, где $\mathbf{s}(u,v)$ – результирующая поверхность.

Исходные данные для реконструкции снимаются с некоторой погрешностью, присущей измерительным приборам. Поскольку оператор скиннинга в своей сути является оператором интерполяции, то результирующая поверхность лопатки неизбежно подвергается воздействию шума. В то же время, в области турбостроения даже незначительная вариация формы изделия приводит к существенному изменению его эксплуатационных качеств. Возникает потребность в сглаживании. Известно, однако, что сглаживание, как таковое, может нанести ущерб замыслу проектировщика, поскольку оно в принципе уничтожает не только шум, но и конструктивные элементы модели. Следовательно, этот инструмент должен применяться локально. Для лопатки турбины эффективное сглаживание без потери конструктивной концепции можно осуществлять на промежуточных кривых-интерполянтах вдоль оси лопатки.

Пусть $\mathbf{g}_i(v)$ – набор промежуточных интерполяционных кривых, используемых для получения итоговых контрольных точек $\mathbf{Q}_{i,j}$ финальной поверхности $\mathbf{s}(u,v)$. По определению, для всех $i = \overline{0,n}$ и $k = \overline{0,K}$ имеем:

$$\mathbf{g}_i(v) = \sum_{j=0}^{K} N_{j,q}(v)\mathbf{Q}_{i,j}\,;\ \mathbf{g}_i(v_k) = \mathbf{P}_{i,k}$$

Здесь q – степень поверхности вдоль оси лопатки, задаваемая пользователем (для получения гладкости C^2 требуется $q \geq 3$). Для устранения высокочастотных осцилляций поверхности мы используем оператор сглаживания F (fairing) на промежуточных кривых $\mathbf{g}_i$:

$$\mathbf{h}_i = \mathrm{F}[\mathbf{g}_i]$$

Здесь $\mathbf{h}_i \equiv \mathbf{h}_i(v)$ – сглаженные кривые, имеющие следующий вид ($i = \overline{0,n}$):

$$\mathbf{h}_i(v) = \textstyle\sum_{j=0}^{K} N_{j,q}(v)\mathbf{R}_{i,j};\ \mathbf{R}_{i,0} = \mathbf{Q}_{i,0};\ \mathbf{R}_{i,K} = \mathbf{Q}_{i,K}$$

Оператор сглаживания F устраняет высокочастотные осцилляции кривой (Рис. 5).

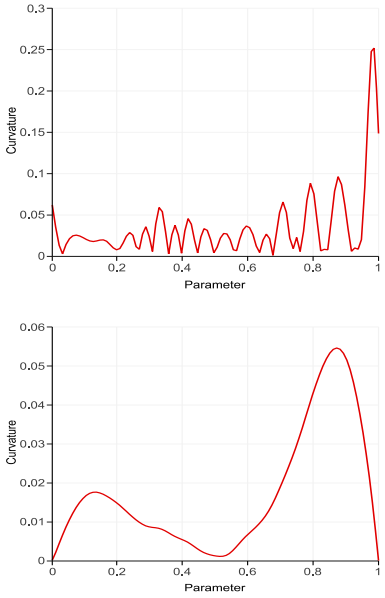


Рис. 5. Графики кривизны $\mathbf{g}_i$ (сверху) и $\mathbf{h}_i$ (снизу).

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Оператор F реализован в соответствии с работой [12] с той разницей, что вместо поверхности сглаживаются кривые. Рассмотрим функционал $\mathrm{J} \equiv \mathrm{J}[\mathbf{c}]$ следующего вида:

$$\mathrm{J} = \int \{\lambda[\mathbf{c}'']^2 + [\mathbf{c} - \mathbf{c_0}]^2\}\, du$$

Здесь $\lambda > 0$ – коэффициент сглаживания; $\mathbf{c_0} \equiv \mathbf{c_0}(u)$ – исходная кривая; $\mathbf{c} \equiv \mathbf{c}(u)$ – результирующая кривая. Интегрирование осуществляется по всей области определения кривых $[u_0, u_1]$. Функционал J агрегирует геометрическую невязку с членом, аппроксимирующим энергию деформации кривой. В общем виде этот способ восстановления кривых и поверхностей из неупорядоченного набора точек хорошо известен. В частности, работы [13] содержат ценные указания относительно стратегии реализации методов аппроксимации со сглаживанием.

Кривую $\mathbf{c}$ представим как результат возмущения контрольных точек кривой $\mathbf{c_0}$:

$$\mathbf{c} = \mathbf{c_0} + \sum_i \mathbf{D}_i N_i$$

Здесь $N_i \equiv N_{i,q}(u)$ – базисные сплайны степени q. Тогда выражение для J можно переписать в следующем виде:

$$\int \left\{ \lambda[\mathbf{c}'']^2 + \left(\sum_i \mathbf{D}_i N_i\right)\left(\sum_j \mathbf{D}_j N_j\right) \right\} du = \\ = \lambda \int [\mathbf{c}_0'']^2\, du + 2 \sum_i \lambda \left(\int N_i'' \, \mathbf{c}_0'' du \right) \mathbf{D}_i \\ + \sum_{i,j} \left(\lambda \int N_i'' \, N_j'' du + \int N_i N_j \, du \right) \mathbf{D}_i \mathbf{D}_j$$

Введем обозначения:

$$a_{i,j} = \lambda \int N_i'' \, N_j'' du + \int N_i N_j \, du$$

$$b_j = \lambda \int N_i'' \, \mathbf{c}_0'' du$$

$$c = \lambda \int [\mathbf{c}_0'']^2 \, du$$

Тогда формула для энергии деформации принимает вид:

$$\mathrm{J} = \sum_{i,j} a_{i,j} \mathbf{D}_i \mathbf{D}_j + 2 \sum_i b_i \mathbf{D}_i + c \to min$$

Из необходимого условия экстремума имеем систему линейных уравнений:

$$\sum_i a_{i,j} \mathbf{D}_i + b_j = 0;\ i, j \in \mathbf{K}$$

Здесь $\mathbf{K}$ – множество индексов контрольных точек $\mathbf{D}_i$, определяющих возмущение исходной кривой. Если $\mathbf{K} = \{0, K\}$, то сглаживание происходит без ограничений. В таком случае кривая $\mathbf{c}(u)$ при достаточно больших значениях коэффициента λ приближается к прямой линии. Для скиннинга лопатки со сглаживанием, но без потери конструктивной концепции, мы используем $\mathbf{K} = \{1, K-1\}$, то есть точки первого и последнего сечений остаются зафиксированными.

Коэффициенты $a_{i,j}$ и b_j находятся методами численного интегрирования. Работа [15] содержит краткие практические сведения о применении некоторых методов численного интегрирования в задаче построения гладких поверхностей (поверхностей *минимальной вариации кривизны*). Мы использовали квадратурную формулу Гаусса [16], реализация которой доступна в открытой библиотеке геометрического моделирования OpenCascade. Отметим, что интегрирование следует осуществлять независимо на каждом узловом интервале.

4.3 Проверка кривизны

Существует множество способов оценки качества результирующей геометрии. Так, широкое распространение получила техника отраженного света («зебра», Рис. 6) [17].

Рис. 6. Поверхность лопатки, построенная оператором скиннинга без сглаживания (сверху) и со сглаживанием при $\boldsymbol{\lambda = 10^{-5}}$ (снизу). Для анализа дефектов формы использовалась «зебра».

Мы оценивали качество поверхности пера лопатки при помощи карты гауссовых и средних кривизн (Рис. 7). Исследование дифференциальных свойств кривых осуществлялось при помощи вспомогательных инструментов, а именно графика и «гребешка» кривизны в соответствии с меморандумом Д. Фарина [18].

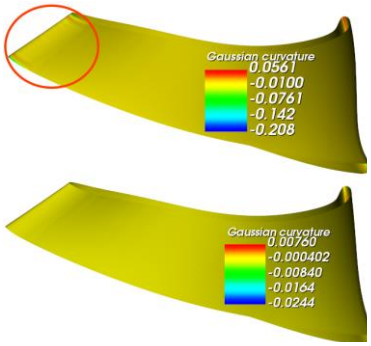


Рис. 7. Поверхность лопатки, построенная оператором скиннинга без сглаживания (сверху) и со сглаживанием при $\boldsymbol{\lambda = 0.001}$ (снизу). Разброс гауссовой кривизны для поверхности со сглаживанием меньше примерно в 10 раз. На поверхности без сглаживания видны дефекты формы, которых нет на сглаженной поверхности.

Используемые средства инспекции доступны в разрабатываемой нами исследовательской платформе с открытыми исходными кодами «Анализ Положения» [19].

5. Выводы и планируемые исследования

Наш метод позволяет создавать параметрические модели лопаток турбин из упорядоченных облаков точек (некоторые результаты показаны на Рис. 8). Переменными проектирования здесь являются контрольные точки B-кривых, представляющих сечения пера лопатки. Для задач параметрической оптимизации предпочтительно, однако, использовать срединные линии сечений с заданными на них функциями толщин. Введение такой параметризации мы оставляем на будущее.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Рис. 8. Результаты реконструкции рабочего колеса турбины с одной лопаткой.

Описанный алгоритм строится на двух основных операторах: *интерполяции* и *сглаживания*. Было показано, что оператор сглаживания, примененный на промежуточных кривых оператора скиннинга (интерполяции), позволяет компенсировать осцилляции формы лопатки без потери конструктивного замысла в целом.

Методы обратного инжиниринга, даже для узкого класса задач, предполагают использование богатого набора геометрических операторов. В дальнейших исследованиях мы рассмотрим методы аппроксимации для восстановления кривых-сечений [20], а также введем дополнительные операторы формообразования, такие как построение скруглений с контролем гладкости стыка. Конечной целью исследования является создание геометрической системы, позволяющей эффективно решать задачи параметрической оптимизации с использованием данных о «чувствительности» модели (Рис. 2).

6. Литература

[1] P. L. Miller, Blade geometry description using B-splines and general surfaces of revolution, Ph.D. thesis, Iowa State University (2000).

[2] L. Piegl, W. Tiller, Surface skinning revisited, The Visual Computer 18 (4) (2002) 273-283.

[3] M. Banovic, O. Mykhaskiv, S. Auriemma, A. Walther, H. Legrand, J.-D. Muller, Algorithmic differentiation of the Open CASCADE Technology CAD kernel and its coupling with an adjoint CFD solver, Optimization Methods and Software (2018) 1-16.

[4] K. Mohaghegh, M. Sadeghi, A. Abdullah, Reverse engineering of turbine blades based on design intent, International Journal of Advanced Manufacturing Technology 32 (9-10) (2007) 1009-1020.

[5] I. Sanchez Torreguitart, T. Verstraete, L. Mueller, Optimization of the ls89 axial turbine profile using a CAD and adjoint based approach, in: Proceedings of 12th European Conference on Turbomachinery Fluid dynamics and Thermodynamics, 2017, p. 1-13.

[6] A. Seiler, D. Grossmann, B. Juttler, Spline surface fitting using normal data and norm-like functions, Computer Aided Geometric Design (June).

[7] S. F. Hosseini, B. Moetakef-Imani, Improved B-spline skinning approach for design of hawt blade mold surfaces, Journal of Mechanics 33 (04) (2017) 427-433.

[8] H. Park, Lofted B-spline surface interpolation by linearly constrained energy minimization, Computer-Aided Design 35 (14) (2003) 1261-1268.

[9] X. Lyu, H. Yu, J. Wu, Surface reconstruction for thin aero engine blade from disorganized contact measured points, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 232 (5) (2018) 746-758.

[10] E. Lee, Choosing nodes in parametric curve interpolation, Computer-Aided Design 21 (6) (1989) 363-370.

[11] L. Piegl, W. Tiller, The NURBS Book, Springer, New York, 1997.

[12] M. Kallay, Constrained optimization in surface design, in: Modeling in Computer Graphics, Springer Berlin Heidelberg, 1993, pp. 85-93.

[13] T. Varady, R. Martin, Chapter 26 - reverse engineering, in: G. Farin, J. Hoschek, M.-S. Kim (Eds.), Handbook of Computer Aided Geometric Design, North-Holland, Amsterdam, 2002, pp. 651-681.

[14] V. Weiss, L. Andor, G. Renner, T. V´arady, Advanced surface fitting techniques, Computer Aided Geometric Design 19 (2002) 19-42.

[15] H. P. Moreton, C. H. Sequin, Functional optimization for fair surface design, ACM SIGGRAPH Computer Graphics 26 (2) (1992) 167-176.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3rd Edition, Cambridge University Press, New York, NY, USA, 2007.

[17] E. Bagci, Reverse engineering applications for recovery of broken or worn parts and re-manufacturing: Three case studies, Advances in Engineering Software 40 (6) (2009) 407-418.

[18] G. Farin, Curvature combs and curvature plots, Computer-Aided Design 80 (2016) 6-8.

[19] S. Slyadnev, A. Malyshev, V. Turlapov, CAD model inspection utility and prototyping framework based on OpenCascade, in: GraphiCon'17 proceedings, 2017, pp. 323-327.

[20] L. A. Piegl, W. Tiller, Least-squares B-spline curve approximation with arbitrary end derivatives, Engineering with Computers 16 (2) (2000) 109-116.

Об авторах

Турлапов Вадим Евгеньевич, д.т.н., профессор кафедры математического обеспечения и суперкомпьютерных технологий ННГУ им.Н.И.Лобачевского,
vadim.turlapov@itmm.unn.ru

Сляднев Сергей Евгеньевич, аспирант 3 года,
ИИТММ ННГУ им.Н.И.Лобачевского,
email: sergey.slyadnev@gmail.com

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Разработка имитационной модели сократительной функции сердца в среде Autodesk Maya

М.В. Титова, Т.Н. Томчинская
m-titova@list.ru | tomchinskaya@ mail.ru
Нижегородский государственный технический университет им. Р.Е. Алексеева, Нижний Новгород, Россия

Представлена динамическая имитационная модель сократительной функции сердца. Сократительная функция реализована на основе моделирования управляющих волокон трех типов по Атласу анатомии человека. Определен алгоритм локальной параметризации сократительной функции сердца, открытый для реализации управления параметрами через данные о локальном кровоснабжении или воздействии лекарственных средств или через их имитацию. Алгоритм действия имитационной модели показан на примере только левого желудочка сердца, но встроен в полную трёхмерную модель желудочкового комплекса сердца. Имитационная модель реализована как твердотельная параметризованная модель в инструментальной среде Autodesk Maya. Планируется продолжение работы и реализация полной модели четырёхкамерного сердца человека.

Ключевые слова: *имитационная модель, сердце человека, желудочковый комплекс, сократительная функция.*

Development of a simulation model of the contractile function of the heart in an Autodesk Maya environment

M.V. Titova, T.N. Tomchinskaya
m-titova@list.ru | tomchinskaya@ mail.ru

Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod, Russia

A dynamic simulation model of the contractile function of the heart is presented. The contract function is realized on the basis of modeling of control muscle fibers of three types according to the Atlas of human anatomy. An algorithm for the local parametrization of the contractile function of the heart is open to realize the management of parameters through data on local blood supply or the effect of drugs or through their imitation. The algorithm of the simulation model is shown for the example of the left ventricle of the heart only, but it is built into the full three-dimensional model of the ventricular complex of the heart. The simulation model is implemented as a solid-state parametrized model in the Autodesk Maya tool environment. It is planned to continue the work and realize the full model of the four-chambered human heart.

Keywords: *imitation model, human heart, ventricular complex, contractile function.*

1. Введение

Еще в 2009 году лидеры фирмы Microsoft видели участие в создании вычислительной супермодели человека с полным циклом его физиологии [1] как одну из важнейших задач фирмы. Экономические причины этого сегодня понятны, но решение этой задачи является очень наукоемким и требующим времени. Так, представляется, что комплексные модели органов, способные не только моделировать физиологические свойства органа, но и изменение геометрии органа и его функций при изменении положения тела человека, появятся еще не скоро.

Модель должна допускать ее вариацию для исследования многих типовых групп пациентов (по возрасту, по полу, по стадиям заболевания, по сопутствующим заболеваниям). Во всех случаях необходима персонализация модели каждого органа и каждой системы по данным объективных медицинских исследований и, прежде всего, 3D данным томографии.

Персонализация модели попутно может решать также задачу детектирования и количественной оценки аномалий органа и, на этой основе, автоматизации диагностики и отслеживания состояния пациента, в том числе на ранних стадиях заболевания.

2. Обзор моделей сердца

В зависимости от целей применения модели сердца может меняться состав ее элементов, состав и сложность математических моделей, набор функций. Модели для цифровой медицины требуют биологически и физиологически корректной реализации, и визуализации. Велик сегодня интерес к динамическим моделям электрической и мышечной активности сердца.

Моделирование электрической активности требует адекватного, допуская конечно и упрощения, геометрического моделирования сердца [2]. А анизотропное моделирование сократительной функции сердца требует знания мышечной структуры и ориентации мышечных волокон во всех тканях сердца [3,4]. Задача является сложной и в функциональном моделировании чаще всего исследователи ограничиваются моделированием только левого желудочка.

В последние годы стала активно решаться задача персонализации модели по данным объективных медицинских исследований, прежде всего по данным томографии [5]. Начали применяться и методы автоматизации цифровой реконструкции сердца. Так в [6] на основе данных КТ реконструируется воксельный мышечный каркас всех 4-х камер сердца, затем этот каркас сглаживается и используется для построения конечноэлементной сетки. Такая сетка требует далее еще и сегментации, отражающей ее тканевое наполнение.

В работе [7] исследуется автоматизация построения геометрической модели левого желудочка сердца собаки по данным МРТ (магниторезонансной томографии), которая бы позволила моделировать кровоток (гемодинамику) левого желудочка в здоровом сердце и пораженном болезнью. Публикация [8] исследует модели сокращения сердца, построенные на основе знания типа ориентации мышечных волокон в конкретном участке мышечной ткани.

Работа по созданию геометрических моделей сердца ведется достаточно давно. Ее первой целью было создание

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

трехмерных анатомических атласов. К настоящему времени создание геометрических моделей приобрело новой дыхание в связи с востребованностью динамических симуляторов конкретных функций органов и мастер-моделей для их персонализации по данным пациента. Наработан достаточно серьезный электронный ресурс моделей сердца [9]. Среди моделей этого ресурса есть как статические, так и динамические модели. Но нет таких, которые меняли бы свое поведение при дифференцированном внешнем воздействии. В частности, нет динамических моделей мышечной активности сердца, которые бы учитывали направления волокон сердечной мышцы и допускали при этом управление мышечной активностью.

Задача создания такой модели сердца решается в настоящей работе. Результатом должна явиться параметризация сокращений сердца параметрами ведущих мышечных волокон сложной структуры.

3. Требования, предъявляемые к модели

Несмотря на наличие в Интернет готовых моделей органов человека, разного уровня детализации, многие компании предпочитают создавать собственные модели. Существует ряд типовых причин этого:

1) ориентированность многих моделей только на визуализацию, а отсюда: представлена только поверхность (не представлен объем), низкая степень детализации, отсутствие сегментации модели на анатомические компоненты (можно решить вопрос, доработав модель);
2) проблема с получением патента;
3) ограниченность функционирования вне среды разработки модели;
4) статичность или отсутствие параметризации модели, и, следовательно, отсутствие возможности построить семейство моделей.

Сейчас существует большое количество программных сред для создания 3D моделей, и не все они подходят для вышеперечисленных задач. Подход сегодняшнего дня предъявляет к модели в их исполнении следующий ряд типовых требований:

1) модель должна быть твердотельной;
2) высокая детализация;
3) она должна быть параметризованной и анимированной;
4) текстурирование близкое к фотореализму;
5) доступный формат представления модели.

Однако, если для виртуальной реальности этого может хватить, то для цифровой медицины этих требований далеко недостаточно. Рассмотрим (табл.1) лучшие модели сердца из набора, представленного на сайте [9].

Таблица 1. Характеристики существующих моделей

Характеристики	Модель 1	Модель 2	Модель 3	Модель 4	Модель 5
Геометрия	Polygonal Quads	Polygonal	Polygonal Quads/Tris	Polygonal	Polygonal Quads/Tris
Число полигонов	57 244	17 910	70 510	26 958	255 877
Число вершин	57 338	18 421	90 095	26 824	288 575
Материалы текстур	Да	Да	Да	Нет	Да
Анимация биения сердца	Да (22 кадра)	Нет	Нет	Нет	Нет
Фото-реализм	Да	Нет	Да	Да	Да
Внутрен. элементы (желудочки, предсердия, клапаны)	Нет	Да	Нет	Да	Да
Анимация внутрен. элементов	Нет	Нет	Нет	Нет	Нет
Стоимость	299$	299$	119$	299$	139$
Soft	Modo 9	Maya 2008	Maya 2013	Cinema 4D 11.5	Maya 2014

Как видно из характеристик моделей, ни одна из них не соответствует полностью требованиям с точки зрения виртуальной реальности.

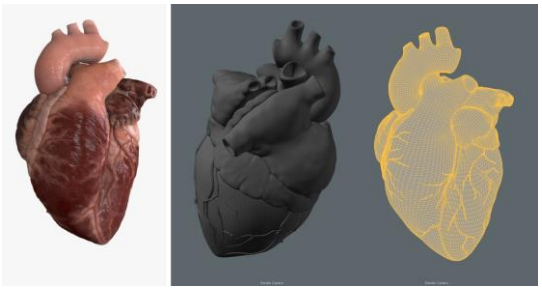


Рис. 1. Рендер, сетка и модель сердца (вариант 1)

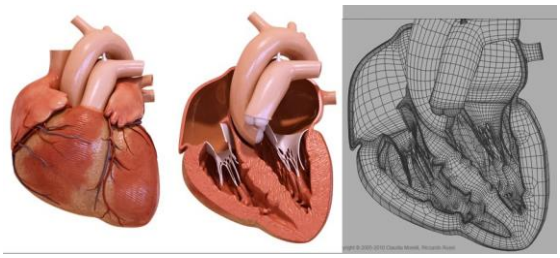


Рис. 2. Рендер и сетка модели сердца (вариант 2)

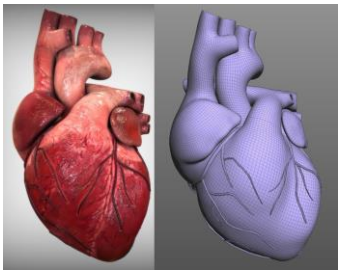


Рис. 3. Рендер, сетка и модель сердца (вариант 3)

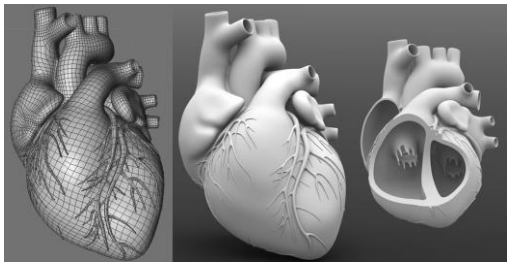


Рис. 4. Сетка и модель сердца (вариант 4)

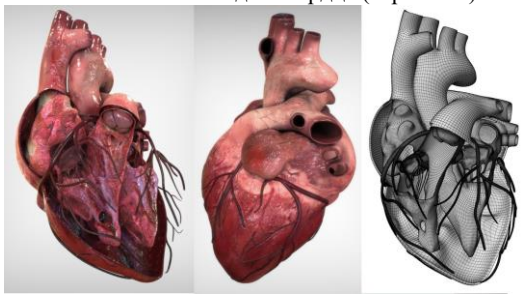


Рис. 5. Рендер, сетка и модель сердца (вариант 5)

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

Анализируя выбранные модели, можно выделить следующие недостатки. Все модели платные от 119$ до 299$. Две модели не имеют внутренних элементов (рис.1, рис.3). Высоко детализированная модель не является твердотельной (рис.5), фотореалистичные текстуры и материалы отсутствуют у модели на рис.4, отсутствует анимация у моделей на рис.2-рис.5.

Цифровая медицина выдвигает следующие дополнительные требования:

1) присутствие в модели тех систем, которые исследуются (проводящей, сосудистой, мышечной)
2) топографическая и функциональная взаимосвязанность систем
3) возможность параметризации и персонализации модели в исследуемых параметрах
4) возможность экстракции параметров формы и плотности тканей модели из данных трехмерных медицинских исследований (томографии);
5) анатомическая полнота модели: присутствие всех необходимых для исследования систем сердца;
6) функциональная адекватность систем и управляемость систем исследуемыми факторами [3].

4. Геометрическая модель сердца

Как видно из обзора работ, модели мышечных тканей сердца, используемые сегодня в практике отечественных и зарубежных исследований, практически не реализуют реальные структуры мышечных волокон, заменяя их приближенными полями направлений.

Для контроля реалистичности нашей модели с точки зрения ее анатомии использован «Атлас анатомии человека» (авторы: Р.Д. Синельников, Я.Р. Синельников). Построение модели сердца выполнено в профессиональном программном продукте Autodesk Maya. Модель сердца состоит из множества элементов, таких как мышцы, хорды, желудочки, предсердия, клапаны, легочная артерия и корень аорты. Желудочковый комплекс сердца является одним из основных компонентов сердца. Каждый элемент моделировался здесь отдельно. Объединение элементов выполнено средствами функций Attach и Merge.

Известно, что мышечные волокна сердца имеют весьма сложную макроструктуру (рис. 6).

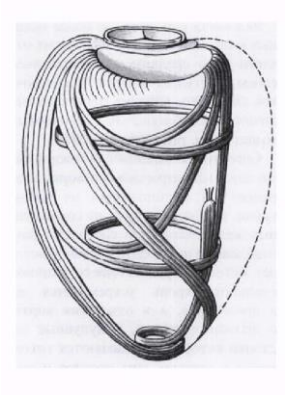


Рис. 6. Направление мышечных пучков стенки левого желудочка (схема, по Атласу анатомии человека)

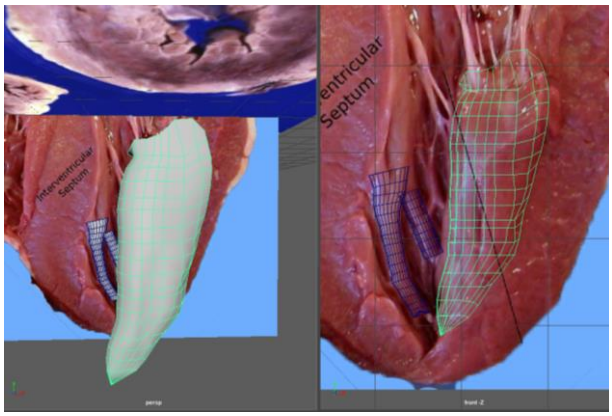
Основой создаваемой модели является специальные референсные поверхности левого желудочка (ЛЖ). Референсная поверхность топологически эквивалентна сфере и может быть создана на основе стандартного элемента «Сфера», путем перемещения точек, а также добавления новых ребер при помощи команды Multi-cut. При корректировке созданной формы использовался инструмент «Skulpt a geometry object». Таким образом можно контролировать любой изгиб и шероховатость поверхности. Результат практически эквивалентен модели, использованной в диссертации С.Ф.Правдина «Математическое моделирование структуры и функции левого желудочка сердца» [10], построенной аналогично экспериментальным данным и модели [11] внешней (epi) и внутренней (endo) поверхностей ЛЖ, как поверхности вращения в цилиндрических координатах (ρ,ψ,z):

$$\rho_{\text{endo; epi}}(\psi) = r_{a;b}(\varepsilon \cdot \cos\psi + (1-\varepsilon)(1-\sin\psi));$$

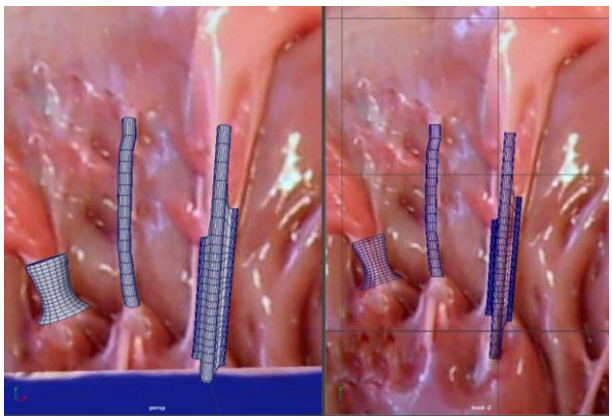
$$z_{\text{endo; epi}}(\psi) = z_{a;b}(1-\sin\psi)$$

где $\varepsilon \in [0,1]$ - параметр, задающий форму стенки ЛЖ от конуса при ε=0, до эллипсоида вращения, при ε=1; r_a, r_b – радиусы внутренней и внешней поверхностей на экваторе; ψ – широта (от 0 до 90). Созданная на данной поверхности параметрическая сетка затем будет играть важную роль управляющую сокращением желудочка.

Созданные далее методом полигонального моделирования необходимые элементы модели левого желудочка представлены на рисунке 7.



а)



б)

Рис.7. а) крупные и мелкие сосочковые мышцы; б) хорды.

Согласно рис.6 смоделировано фиброзное кольцо, к которому происходит крепление всех мышечных волокон. На данном этапе принято решение реализовать управление сокращением левого желудочка через изменение длины трех типовых мышечных волокон, задающих движение соответствующих зон сердечной мышцы (рис.6). Для моделирования каждого из волокон создается NURBS-кривая осевой линии волокна $P(t)$, все контрольные точки которой расположены на соответствующей референсной поверхности желудочка. В среде Maya для этого используем привязку Make the selection object live. В текущем варианте

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

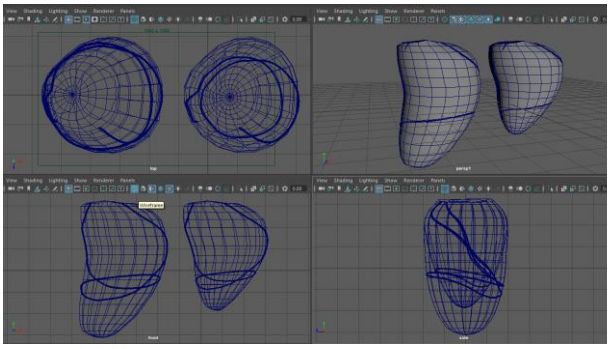


Рис. 9. Визуализация ЛЖ сердца в состояниях диастолы и систолы, созданных изменением длины управляющих волокон

осевая линия имеет 15 (m=15) узлов P_i в однородных координатах и 14 сегментов:

$$P(t) = \sum_{i=1}^{m} B_{i,3}(t) P_i w_i$$

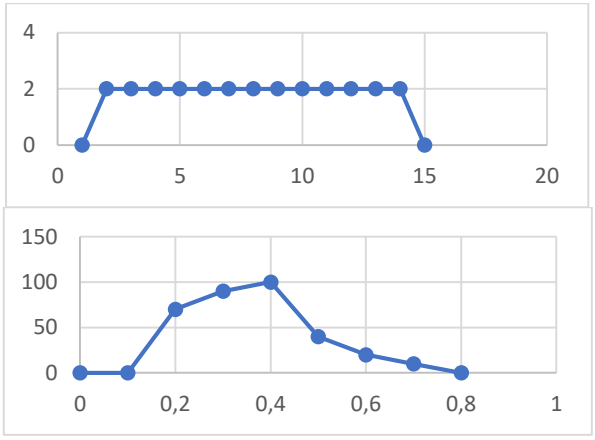


Рис.8 Пример графиков управления: а)максимальная величина коэффициента сжатия сегмента в управляющих точках сплайна; б)изменение коэффициента сжатия (%) во времени (время в сек.): желудочковая систола и диастола

где $B_{i,3}$- коэффициенты Кокса-де-Бура для NURBS 3 степени. Распределение абсолютной величины сжатия по направлениям сетки поверхности в каждой точке осевой линии зависит от наклона сплайна к вертикальной оси поверхности. Например, если волокно образует горизонтальное кольцо, то сжатие происходит только в горизонтальной плоскости; если направлено параллельно оси Z, то сжатие только вертикальное; если направлено под углом, то сжатие вдоль оси волокна раскладывается на вертикальную и горизонтальную проекции. Управление величиной сжатия может осуществляться как посегментно, так и непрерывно в каждой точке осевой линии мышечного волокна. В текущем варианте коэффициент сжатия участков сплайна задан графиком и в норме считается постоянным по всей длине за исключением начального и конечного сегментов (Рис.8а). Изменением коэффициента сжатия во времени управляет еще один график (Рис.8б).

Если управляющие волокна каждого из трех типов представлены одним экземпляром, то можно принять его воздействие однородным по всему горизонтальному слою сетки. Вся вертикальная ось желудочка должна быть поделена на три зоны влияния (для каждого типа – своя зона) с условиями непрерывности стыковки. На рисунке 9 показан ЛЖ сердца в диастоле и систоле, в качестве примера управления, реализованного таким образом.

Еще более интересна возможность локального управления сократимостью волокон в каждой точке миокарда, связанная с кровоснабжением участка или воздействием лекарств. В этом случае необходимо замощение поверхности желудочка такими волокнами с некоторым шагом. Это может быть сделано подобно аппарату, представленному в [10], для образов хорд сплайна на полукруге радиуса $R = \varphi_{max} \cdot (r + l)/\pi$:

$$\rho(\Phi) = \frac{Y}{\sin\Phi} \cdot \left(r + l\frac{\Phi}{\pi}\right),$$
$$\varphi(\Phi) = \varphi_{max} \cdot \Phi/\pi,$$
$$z(\Phi) = z_{sp}(\rho(\Phi), \varphi(\Phi)),$$

где разные значения параметра $Y \in (0, R)$ отвечают разным волокнам, $\Phi \in \left[\arcsin\frac{Y}{R}, \pi - \arcsin\frac{Y}{R}\right]$, $z_{sp}(\rho, \varphi)$ — явное выражение спиральной поверхности в цилиндрической системе координат, r - внутренний радиус ЛЖ на экваторе; l — толщина стенки ЛЖ на экваторе; z — аппликата внутренней поверхности ЛЖ, $\varphi_{max} > \pi$ – угол закрутки миокарда.

Твердотельный образ мышечных волокон создаётся стандартным выдавливанием (extrud) полигонального круга вдоль осевой линии волокна.

После построения имитационной модели левого желудочка аналогично создается модель правого желудочка (Рис.10,11).

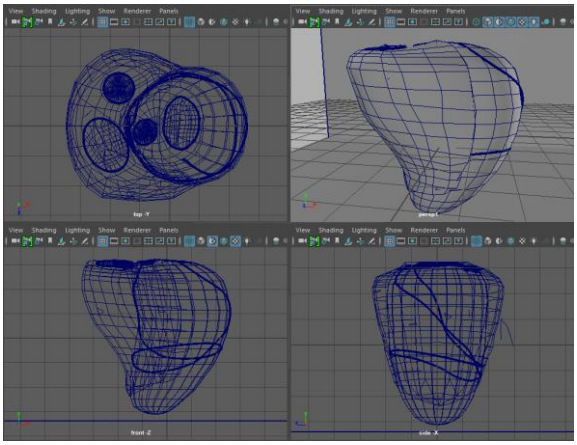


Рис. 10. Трёхмерная модель желудочкового комплекса

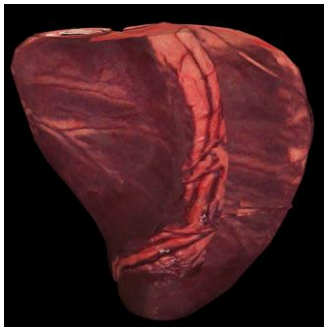


Рис. 11. Визуализация модели желудочкового комплекса

Работа опубликована при финансовой поддержке РФФИ, грант 18-07-20045\18

5. Заключение

В работе представлена динамическая имитационная модель сократительной функции сердца. Сократительная функция реализована на основе моделирования управляющих волокон трех типов по Атласу анатомии человека. Определен алгоритм локальной параметризации сократительной функции сердца, открытый для реализации управления параметрами через данные о локальном кровоснабжении или воздействию лекарственных средств или их имитацию. Алгоритм действия имитационной модели показан на примере только левого желудочка сердца, но встроен в полную трёхмерную модель желудочкового комплекса сердца. Имитационная модель реализована как твердотельная параметризованная модель в инструментальной среде Autodesk Maya. Планируется продолжение работы и реализация полной модели четырёхкамерного сердца человека.

6. Литература

[1] Nuala Moran. The virtual physiological human. The search for computing's supermodel //Futures. 2009 (4), p.28-30 (http://www.wisdom.weizmann.ac.il/~yaki/FUTURES_04.pdf) – (Дата обращения: 28.02.2018).

[2] С.Ф. Правдин. Неосесимметричная математическая модель анатомии левого желудочка сердца // Российский журнал биомеханики. Вып. 17, № 4 (62). 2013. С. 84–105.

[3] S.F. Pravdin, V.I. Berdyshev, A.V. Panfilov, L.B. Katsnelson, O. Solovyova, V.S. Markhasin. Mathematical model of the anatomy and fibre orientation field of the left ventricle of the heart // Biomedical Engineering Online, 12:54, 2013. 21 p.

[4] S.F. Pravdin, H. Dierckx, L.B. Katsnelson, O. Solovyova, V.S. Markhasin, A.V. Panfilov. Electrical wave propagation in an anisotropic model of the left ventricle based on analytical description of cardiac architecture // PLOS One. 2014. PLoS ONE 9(5): e93617.

[5] Матвеенко В.П., Шардаков И.Н., Шестаков А.П. Алгоритм создания трехмерных образов органов человека по томографическим данным // Российский журнал биомеханики. – 2011. – Т. 15, № 4. –С. 20–32.

[6] И.Н. Шардаков, А.П. Шестаков. Построение четырехкамерного геометрического образа сердца человека на основе рентгеновской томографии// Российский журнал биомеханики. 2015. Т. 19, № 4: 372–384

[7] Choi Y.J. A new MRI-based model of heart function with coupled hemodynamics and application to normal and diseased canine left ventricles / Y.J.Choi, J.Constantino, V.Vedula, N.Trayanova and R.Mittal // Frontiers in Bioengineering and Biotechnology, 2015, V.3, Art.140, p.1-13. DOI: 10.3389/fbioe.2015.00140 https://www.frontiersin.org/articles/10.3389/fbioe.2015.00140/full

[8] Molitoris,J.M. Precisely parameterized experimental and computational models of tissue organization/ J.M. Molitoris, S.Paliwal, R.B.Sekar, R.Blake, J.Park, N.A.Trayanova, L.Tung and Andre Levchenko// Integrative Biology, 2016, 8, 230-242. DOI: 10.1039/c5ib00270b

[9] Human Heart 3D Models [Электронный ресурс] - https://www.turbosquid.com/3d-model/human-heart . - (Дата обращения: 08.08.2018).

[10] Gavrilov N.I., Vasiliev E.P., Khramov I.V., Getmanskaya A.A., Turlapov V.E. Visual analytics and segmentation of color biomedical high resolution cryo-imaging scans // Scientific Visualization, (2017) 9 (5), pp.105-116 (http://sv-journal.org/2017-5/09.php?lang=en)

[11] Правдин С.Ф. Математическое моделирование структуры и функции левого желудочка сердца. Автореф. дисс. на соиск. ст. канд.ф.-м.н., 2015. 20с.

[12] Streeter Jr DD. Gross morphology and fiber geometry of the heart. In: Berne RM, Sperelakis N, editors. Handbook of physiology: Sec.2. The heart (American Physiology Society), vol. 1. 1979. p. 61—112.

[13] The helical ventricular myocardial band band: global, three-dimensional, functional architecture of the ventricular myocardium (Review)/M.J.Kocica,…, F.Torrent-Guasp/ European Journal of Cardio-thoracic Surgery, 29S (2006) S21—S40.

Об авторах

Титова Мария Владимировна, магистрант кафедры графических информационных систем института радиоэлектроники и информационных технологий при Нижегородском государственном техническом университете им. Р.Е. Алексеева; e-mail: m-titova@list.ru.

Томчинская Татьяна Николаевна, к.т.н., доцент кафедры графических информационных систем института радиоэлектроники и информационных технологий при Нижегородском государственном техническом университете им. Р.Е. Алексеева; e-mail: tomchinskaya@mail.ru.

The publication of the conference proceedings was supported by the RFBR, grant No. 18-07-20045\18

Авторский указатель

A
Akhmetshin R. 324

D
Danilov V.V. 252
Demeshko M. 66
Demin A. 66

E
Ershov S.V. 20

F
Frolov V.A. 26, 31

G
Galaktionov V.A. 26
Gerget O.M. 252

K
Kharlova O. 258
Khvostikov A. 258
Kolesnik M. 70
Krylov A.S. 258, 262

M
Malkov P. 258
Malyshev A. 479
Manakov R.A. 252
Mardanov A. 324
Melgani F. 252
Melnikova A.A. 262
Mikhailov I. 258
Muftejev V. 324

N
Nabiyev R. 324

O
Oleynikova N. 258
Oliva D. 66

R
Ryabinin K. 70

S
Sanzharov V.V. 26, 31
Skirnevskiy I.P. 252

T
Turlapov V. 479

V
Voloboy A.G. 20

Z
Zhdanov D.D. 20
Zhidkov A. 479
Ziatdinov R. 324

А
Абросимов С.Н. 389
Авербух В.Л. 77, 103
Авербух Н.В. 74
Агапов Н.А. 418, 428
Акопян В.С. 285
Андреев С.В. 82, 86
Андросов К.Ю. 90
Арбузов В.А. 119, 123
Арбузов Э.В. 119
Арсеев С.П. 218

Б
Багутдинов Р.А. 94
Барладян Б.Х. 36, 56, 484
Басов А.Ю. 399
Бахтерев М.О. 103
Белобородов Д.С. 191
Белоусов К.И. 167
Бердников В.С. 119
Береснев А.П. 196, 200
Бирюков Е.Д. 36, 484
Богданов Н.Н. 43, 56
Бойков А.А. 328
Болотова Ю.А. 230, 238, 432
Бондарев А.Е. 82, 98
Бондарева Н.А. 86
Бордодымов А.Н. 142
Борисова М.А. 155
Бородина А.А. 402
Бринюк И.А. 406
Будак В.П. 399, 409, 411, 422
Бычкова И.В. 415

В
Валиев Д.Т. 418, 428
Васин Д.Ю. 15, 332, 337
Васёв П.А. 103
Вехтер Е.В. 127, 348, 460
Вин А.М. 40, 107
Волобой А.Г. 36, 226
Волошинов Д.В. 350
Воронова О.С. 354
Выголов О.В. 142

Г
Гаврилов Д.А. 266
Галактионов В.А. 43, 98
Галеев Р.М. 454
Галкин Т.П. 111
Гвоздарев И.Л. 77
Глазырин Ф.Н. 132
Горбунов А.Н. 90
Гордеев Д.С. 115
Грачёв Д.Г. 422
Григорьева М.А. 111
Гримайло А.В. 409
Громов В.П. 332, 337
Гудков В.Ю. 214

Д
Давыдов С.В. 90
Давыдова Е.М. 436, 446
Данилов С.Ю. 142
Довганич А.А. 270
Дорощенко И. 135
Дубнищев Ю.Н. 119, 123
Дубровская М.В. 275, 289
Дюмин В.А. 389
Дё Ю.С. 424

Е
Евсютин О.О. 185, 297
Ерёмченко Е.Н. 15, 301, 304

Ж
Жданов А.Д. 43, 56
Жданов Д.Д. 43, 56, 61
Желтов В. 411

З
Завьялов Д.А. 307
Закиров Э.И. 266
Захарова А.А. 94, 127, 304, 307
Знаменская И.А. 132, 135
Зоев И.В. 196, 200
Золотухин Т.А. 138

И
Исаев Р.А. 158
Исхаков А. 311
Исхаков С. 311

К
Казакова Т.Д. 348
Казиев А.Б. 238
Калайда В.Т. 439, 467
Калинин Р.Г. 428
Калиновский И.А. 204
Камаев А.Н. 319
Канаева И.А. 432
Канкеева Н.С. 436
Капустин В.В. 222
Касаткин С. 439
Касьянов В.Н. 138
Катаев М.Ю. 15
Кислицын С.А. 119
Климентов А.А. 111
Клышинский Э.С. 98
Князь В.А. 454

Князь В.В. 142
Кокурина А.С. 297
Колина В. 415
Конопацкий Е.В. 358
Конушин А.С. 208, 234, 242
Конушин В.С. 208
Копылов М.С. 484
Коротеева Е.Ю. 132
Корчуганова Т.А. 111
Крылов А.С. 15, 270, 280, 285
Крысько А.В. 172
Кузнецов А. 135
Купляков Д.А. 208
Курячий М.И. 222
Кустов Е.А. 332

Л
Лаврентьева Г.М. 204
Ларионова А.Ю. 442
Лебедев Л.И. 315
Левчук Г.И. 77
Леманов В.В. 123
Лепихова Д.Н. 214
Ливший И.Л. 43
Липкина А.Л. 362
Ломов Н.А. 218
Лузин И.С. 285
Лукашов В.В. 123
Луцкий А. 135
Любчинов Е.В. 372

М
Малышев А.С. 488
Мамаев Н.В. 280
Мамонтов Г.Я. 471
Манаков Д.В. 103, 146
Марков Н.Г. 196, 200
Мелерзанов А.В. 266
Мелёхина О.С. 119
Местецкий Л.М. 247, 362
Мешкова Т. 411
Мещеряков Р.В. 297, 311
Мещеряков С.В. 151
Мильман И.Е. 111
Минчук А.М. 446
Мовчан А.К. 222
Мошкова Т.В. 369, 393
Муравьев С.В. 155
Мясоедова Т.М. 372

Н
Насонов А.В. 285
Насонова А.А. 285
Наумов Д. 135
Небаба С.Г. 94, 450
Нефедов Д.И. 376
Новиков М.М. 454
Новожилов М.М. 275, 289
Новоторцев Л.В. 226
Носова С.А. 292

П
Палташев Т.Т. 15
Панчук К.Л. 372
Папкова И.В. 172
Пермяков С.С. 167
Пилюгин В.В. 111
Подвесовский А.Г. 158
Попова Е.С. 230
Поройков А.Ю. 107
Потемин И.С. 43, 56
Праздникова Т.Д. 176
Притыкин Ф.Н. 376

Р
Радченко В.Ю. 348, 406
Расковская И.Л. 163
Ризен Ю.С. 442, 463
Ринкевичюс Б.С. 107, 163
Романенко С.А. 458
Роменский С.А. 369
Россомахина Н.Е. 418
Ротков С.И. 332, 337, 369, 393
Рудникович А.С. 48
Рябинин К.В. 167

С
Савицкий Ю.В. 238
Салтыкова О.А. 172
Сафьянникова В.И. 460
Свириденко М.И. 52
Семенова Н.С. 285
Серяков В.А. 471
Сляднев С.Е. 15, 488, 495
Смычек М.М. 393
Соколов В.Г. 56
Соколова А.И. 234
Сорокин А.А. 319
Сорокин М.И. 61
Сотников М.Е. 463
Спицын В.Г. 230, 238
Столбова И.Д. 380

Т
Тарбаев Б.Б. 467
Тарских М.С. 385
Тикунов В.С. 304
Титов М.А. 111
Титова М.В. 500
Тихомирова А.В. 163
Тихонов–Бугров Д.Е. 389
Толкачев А.В. 163
Томчинская Т.Н. 500
Трубаков А.О. 176
Туранов С.Б. 458
Турлапов В.Е. 292, 488, 495
Тюлькин Е.В. 418
Тюрина В.А. 369, 393

У
Урманов И.П. 319

Ф
Федотов А.М. 328
Филоненко Д.Ю. 103
Форгани М.А. 103

Х
Харинов М.В. 180
Хмелевский Ю.П. 471
Хруль Т.С. 402, 471
Хрушков П.В. 242

Ч
Чембаев В. 411
Черепов А.А. 475
Черепова Л.Н. 475
Чуприна С.И. 167

Ш
Шакиров Е.М. 458
Шальнов Е.В. 208
Шапиро Л.З. 36, 98
Шахлан А.О. 315
Шахуро В.И. 242
Шипунова К.В. 48
Ширинкина М.А. 380
Шкляр А.В. 127, 385, 424, 475
Шолохова Т.Н. 247
Шумская О.О. 185

Щ
Щелкунов Н.Н. 266
Щемелинин Д.А. 151

Ю
Юрин Д.В. 270, 280

Я
Ягунов Т.А. 200
Ясаков Ю.В. 332

Научное издание

GraphiCon 2018

28-я Международная конференция
по компьютерной графике
и машинному зрению

Труды конференции

Издается в авторской редакции
Компьютерная вёрстка: *Д.А. Завьялов, С.Г. Небаба*

Национальный исследовательский Томский политехнический университет
634050, г. Томск, ул. Ленина, 30