

Московский государственный университет имени М.В. Ломоносова  
Факультет вычислительной математики и кибернетики

На правах рукописи

Игнатенко Алексей Викторович

**МЕТОДЫ ИНТЕРАКТИВНОЙ ВИЗУАЛИЗАЦИИ И ОБРАБОТКИ  
ТРЕХМЕРНЫХ ДАННЫХ НА ОСНОВЕ ИЗОБРАЖЕНИЙ**

Специальность 05.13.11 – «Математическое и программное  
обеспечение вычислительных машин, комплексов и компьютерных  
сетей»

Автореферат  
диссертации на соискание учёной степени  
кандидата физико-математических наук

Москва – 2005

Работа выполнена на кафедре Автоматизации Систем Вычислительных  
Комплексов факультета ВМиК МГУ

Научный руководитель: Кандидат физико-математических наук  
Баяковский Ю.М.

Официальные оппоненты: Доктор физико-математических наук  
Михайлюк М.В.

Кандидат физико-математических наук  
Березин С.Б.

Ведущая организация: Институт Системного Программирования  
РАН

Защита диссертации состоится « \_\_\_\_ » \_\_\_\_\_ 2005г. в \_\_\_\_ часов на  
заседании диссертационного совета Д 501.001.44 в Московском государственном  
университете им. М.В. Ломоносова по адресу: 119992, ГСП-2, Москва,  
Ленинские горы, МГУ, 2-й учебный корпус, факультет ВМиК, аудитория 685.

С диссертационной работой можно ознакомиться в библиотеке факультета  
ВМиК МГУ.

Автореферат разослан « \_\_\_\_ » \_\_\_\_\_ 2005г.

Ученый секретарь диссертационного  
совета, кандидат физико-  
математических наук, профессор



Н.П. Трифонов

## Общая характеристика работы

### Объект исследований и актуальность темы

Одна из основных задач в области разработки человеко-машинных интерфейсов – это задача синтеза изображений трехмерных геометрических моделей. В настоящее время особую актуальность приобретают задачи, связанные с алгоритмами интерактивного синтеза изображений объектов реального мира. Эти алгоритмы находят свое применение во многих прикладных областях, таких как автоматизированное проектирование, медицина, археология, мультимедийные, образовательные приложения, виртуальная реальность.

Процесс синтеза изображений трехмерной модели называется *экранизацией* (rendering). Традиционный процесс экранизации состоит из нескольких этапов: создание геометрической модели, моделирование материалов и условий освещения, синтез проекции модели для заданных параметров виртуальной камеры, растеризация геометрии, вычисление цвета пикселей результирующего изображения с учетом модели освещения.

Для реальных объектов входными данными для создания модели, как правило, являются цифровые фотографии, видеопоследовательности или карты дальности (глубины), получаемые с помощью специальных дальномерных устройств, позволяющих вводить в компьютер данные о форме интересующего объекта.

*Представлением* (representation) трехмерных данных будем называть класс (тип) цифровых трехмерных моделей, обладающих схожими признаками (структурой данных, свойствами и т.п.). Наиболее распространенным представлением трехмерных данных, используемым в интерактивной трехмерной графике, является граничное полигональное представление. Такое представление задает кусочно-линейную аппроксимацию поверхности объекта с помощью набора полигонов.

Современные системы экранизации реальных объектов реконструируют из входных данных поверхность объекта, создают модель в граничном представлении и экранизируют уже ее. Такая схема хорошо работает для синтетических, созданных вручную моделей и достаточно простых моделей освещения.

Однако для экранизации моделей объектов реального мира традиционный процесс недостаточно эффективен по следующим причинам:

- Для полигонального представления необходима полная реконструкция поверхности, провести которую автоматически для исходных данных в виде фотографий или видеопоследовательностей затруднительно или невозможно, а ручная реконструкция является чрезвычайно трудоемкой.

- Модели реального мира обычно имеют высокую детализацию и сложную форму, поэтому требуют огромного количества полигонов для качественной аппроксимации. Также проблемой является оптическая сложность – это сложность модели материала и освещения. На данный момент используемые в интерактивной экранизации модели материалов и освещения достаточно просты и не полностью передают свойства реальной поверхности. Использование более сложных моделей ограничено требованием интерактивности.
- В последнее время становится все более актуальной задача адаптации сложных данных для их экранизации на маломощных устройствах – например, наладонных компьютерах или мобильных телефонах. Для полигональных моделей решение этих задач связано с большими вычислительными затратами, что может быть неприемлемо для маломощных устройств.

Таким образом, чрезвычайно актуальной является задача исследования альтернативных представлений и методов экранизации, адаптированных под реальные данные.

В рамках диссертационной работы исследуется идея прямого использования входных данных о реальных объектах в качестве модели для экранизации. Поскольку на входе и выходе алгоритма экранизации находятся изображения объектов в том или ином виде, прямое использование изображений в качестве модели может преодолеть недостатки традиционного графического процесса.

В диссертации анализируются трехмерные представления на основе изображений. Рассматривается задача адаптации данных в этих представлениях, показывается их эффективность для решения этой задачи.

## **Цели работы**

- Исследование представлений трехмерных объектов на основе изображений, алгоритмов их обработки и экранизации.
- Разработка представления трехмерных данных на основе изображений, позволяющего эффективно хранить, передавать по сети и визуализировать трехмерные модели. Разработка алгоритмов обработки и синтеза изображений для этого представления.
- Исследование применимости представлений на основе изображений к задаче адаптации трехмерных данных для интерактивного показа на маломощных устройствах (карманных компьютерах, мобильных телефонах). Разработка эффективных методов экранизации трехмерных данных высокой сложности на маломощных устройствах.

## Научная новизна работы

- Предложено новое представление трехмерных данных на основе изображений, позволяющее эффективно хранить, передавать по сети и экранизировать трехмерные модели. Также предложены методы обработки моделей в этом представлении.
- Разработан алгоритм многомасштабной экранизации предложенного представления, обеспечивающий более высокую скорость экранизации по сравнению с известными алгоритмами.
- Предложен новый алгоритм клиент-серверной экранизации потока трехмерных данных, ориентированный на мобильные клиентские устройства (включая карманные ПК) в условиях низкой пропускной способности сети.

## Практическая значимость и реализация

Разработаны и доведены до реализации методы и алгоритмы решения актуальных задач компьютерной графики. Программные реализации описываемых в диссертации методов удовлетворяют требованиям и ограничениям, сформулированным в постановке задачи.

На основе предложенного представления трехмерных данных разработана спецификация формата данных, которая была включена в международный стандарт MPEG-4 AFX (Animated Framework Extension). Часть разработанных алгоритмов экранизации вошла в пакет стандартного программного обеспечения MPEG-4 AFX.

Разработана программная система Visual Framework, в рамках которой разрабатывались алгоритмы экранизации, а также клиент-серверная система экранизации потока трехмерных данных. Система предназначена для анализа различных методов представления и экранизации трехмерных данных. Компонентная архитектура системы позволяет добавлять новые форматы данных, представления и алгоритмы экранизации без изменения существующего кода. Система Visual Framework используется в учебном и научном процессе в лаборатории компьютерной графики и мультимедиа на факультете ВМиК МГУ для обучения методам экранизации по изображениям и разработки новых алгоритмов.

По тематике разработанных методов и алгоритмов, в рамках курса компьютерной графики, который читается на 2-м курсе факультета ВМиК МГУ, было подготовлено практическое задание (<http://graphics.cs.msu.ru/courses/cg02b/assigns/hw-6/hw6.html>).

Также было подготовлено учебное пособие по графической библиотеке OpenGL, которая использовалась для реализации разработанных алгоритмов (Ю. Баяковский, А. Игнатенко, А. Фролов, Графическая библиотека OpenGL).

2003, Москва: Издательский отдел факультета Вычислительной Математики и Кибернетики МГУ им. М.В. Ломоносова, 132с).

## Апробация работы и публикации

Результаты работы докладывались и обсуждались на:

- 13-й международной конференции по компьютерной графике и машинному зрению GraphiCon 2003, “A Framework for Depth Image-Based Modeling and Rendering” («Система для экранизации и моделирования на основе изображений с глубиной»), Россия, Москва, 2003.
- 14-й международной конференции по компьютерной графике и машинному зрению GraphiCon 2004, “A Real-Time 3D Rendering System with BRDF Materials and Natural Lighting” («Система интерактивной визуализации с поддержкой материалов с ДФОС и природным освещением»), Россия, Москва, 2004.
- Международной конференции по обработке изображений ICIP (International Conference on Image Processing), “Depth Image-based Representations for static and animated 3D objects” («Представление на основе изображений с глубиной для статических и анимированных трехмерных объектов»), США, Нью-Йорк, 2002.
- Семинаре по компьютерной графике и мультимедиа под руководством Ю.М. Баяковского (ф-т ВМиК МГУ).
- Научно-исследовательском семинаре по автоматизации программирования под руководством проф. М.Р. Шура-Бура (ф-т ВМиК МГУ).

По тематике диссертации получено четыре международных патента.

## Объем и структура работы

Диссертация состоит из введения, четырех глав, заключения, приложения и списка литературы. Содержание работы изложено на 148 страницах. Список литературы включает 96 наименований.

## Краткое содержание работы

Во **введении** формулируются цели и задачи диссертации, обосновывается ее актуальность, научная новизна и практическая значимость, кратко излагается содержание работы.

В **первой главе** приводится сравнительный анализ методов представления трехмерных данных и методов экранизации на основе изображений.

В **п.1.** даются общие понятия и определения, описываются сложности традиционного подхода, основанного на применении граничных или воксельных моделей.

Основными проблемами применения традиционного процесса моделирования к экранизации реальных объектов являются сложность моделирования, высокая геометрическая сложность моделей, затрудняющая интерактивную визуализацию, а также высокая оптическая сложность моделей реального мира, не позволяющая достичь фотореалистичного качества синтезируемых изображений.

*Экранизация на основе изображений* (Image-Based Rendering, IBR) – это подход к решению задачи экранизации, основанный на стремлении преодолеть трудности традиционного подхода путем прямого использования исходных данных (изображений, видеопоследовательностей) в процессе экранизации. С экранизацией на основе изображений тесно связано *моделирование на основе изображений* (Image-Based Modeling, IBM), задачей которого является создание моделей на основе изображений.

Анализировать представления и методы экранизации на основе изображений предлагается с помощью понятия пленоптик-функции.

Пленоптик-функция определяется как интенсивность светового излучения, проходящего через некоторую точку пространства  $(x, y, z)$ , для данного направления  $(\theta, \varphi)$ , данной длины световой волны  $\lambda$  и в заданное время  $t$ :

$$P_t = P(X, Y, Z, \theta, \varphi, \lambda, t)$$

Набор входных изображений может рассматриваться как конечный набор значений пленоптик-функции для некоторой области аргументов. Задача алгоритма экранизации – получение значений функции при заданных параметрах, в общем случае отличных от исходных. Это позволит построить изображение сцены для произвольного положения виртуальной камеры.

Реконструкция неизвестных значений пленоптик-функции может вестись как с восстановлением трехмерной геометрической информации (полным или частичным), так и без него (с помощью интерполяции между изображениями).

В п.2. приводятся критерии сравнения алгоритмов экранизации с использованием изображений:

- Удобство получения моделей
- Класс представляемых объектов (индивидуальные, сцены и т.п.)
- Объем данных и возможность сжатия
- Аппаратная поддержка экранизации
- Качество экранизации
- Поддержка адаптивности, эффективная передача по сети

В п.3. описаны основные классы методов экранизации. Методы делятся на четыре категории: экранизация без геометрии, экранизация с неявным использованием геометрии, экранизация с дискретной геометрией и экранизация с

явной реконструкцией геометрии. Это деление должно рассматриваться скорее как непрерывный спектр методов, чем дискретный набор категорий (см. Рисунок 1).



**Рисунок 1 Спектр методов экранизации по изображениям**

Методы, основанные изображениях без реконструкции геометрии, рассматриваются **п.4**. Эти методы синтезируют выходное изображение с помощью фильтрации и интерполяции изображений из большого набора исходных фотографий с известными положениями камеры. К достоинствам интерполирующих алгоритмов относится возможность воспроизведения практически любых материалов и форм поверхностей. К недостаткам можно отнести использование излишнего количества данных для увеличения качества выходных изображений. Это ведет к усложнению процесса получения фотографий, дублированию данных и очень большому объему используемой памяти.

В **п.5** анализируются методы следующей категории – использующие неявное представление геометрии. Термин «неявная геометрия» означает что трехмерные положения точек не восстанавливаются, а целевое изображение реконструируется с помощью манипуляций с соответствиями на изображениях. Методы этой категории не требуют явных геометрических моделей. Вместо них достаточно построить соответствие *особенностей* (например, точек или линий) на изображениях исходного набора. Для генераций таких соответствий используются технологии компьютерного зрения. Эти методы позволяют значительно уменьшить объем используемых данных (по сравнению с методами без реконструкции геометрии), однако на практике их применение ограничено из-за неустойчивости методов автоматического поиска соответствий, ведущей к некорректным результатам интерполяции пленоптик-функции.

В **п.6** описываются методы, использующие геометрию в дискретном представлении: карты глубины (карты дальности), многослойные карты глубины, деревья многослойных карт глубины. Методы этой категории используют алгоритмы *деформации изображений* (image warping) для синтеза новых изображений. Для заданной целевой камеры исходные изображения с глубиной деформируются таким образом, чтобы результирующее изображение совпадало с результатом восстановления трехмерных положений точек и их последующей проекции на экранную плоскость целевой камеры. Методы на основе

деформации используют алгоритмы обработки изображений и поэтому более эффективны, чем традиционный графический процесс.

Другой край спектра методов занимает традиционная экранизация с поддержкой текстурированной геометрии. Такие методы рассматриваются в **п.7**. Действительно, текстурирование относится к методам экранизации на основе изображений. Эти подходы основываются на точной геометрии и небольшом количестве изображений, используемых для создания текстур. Недостатки таких методов описаны выше.

Результаты проведенного анализа приводятся в **п.8**. Делается вывод о том, что по мере уменьшения количества и точности геометрии в модели, увеличиваются требования к количеству изображений, необходимых для качественного синтеза. С другой стороны, дискретные представления (в частности, на основе изображений) позволяют использовать более простые алгоритмы адаптации и многомасштабности, т.к. работа ведется в двумерном пространстве изображения. Делается заключение о том, что представления на основе изображений эффективно решают задачу экранизации моделей реального мира, в том числе многомасштабной экранизации.

Во **второй главе** приводится описание нового представления трехмерных данных на основе изображений, а также алгоритмов его обработки и экранизации. Представление основано на моделировании сцены с помощью набора проекций геометрии на заданные плоскости.

В **п.1** приводится постановка задачи разработки универсального представления и обосновывается ее актуальность.

Популярность полигонального представления обусловлена прежде всего наличием набора стандартных форматов для полигональных трехмерных данных. Это позволяет независимо развивать алгоритмы и системы создания данных и методы экранизации, обмениваться данными и т.д.

Несмотря на быстрое развитие и нарастающую популярность методов экранизации на основе изображений, до последнего времени не существовало международных стандартов представления трехмерных данных на основе изображений (за исключением Apple QuickTime VR – стандарта де-факто, охватывающего только самые простые представления на основе панорам и видео-последовательностей).

Была поставлена задача разработки универсального и гибкого представления, охватывающего широкий участок спектра методов представления и экранизации на основе изображений. Представление разрабатывалось с целью включения в развивающийся международный стандарт MPEG-4.

В **п.2** приводится описание структуры предлагаемого универсального представления, разработанного по результатам анализа существующих представлений и алгоритмов.

Модель представляет собой пару  $(P, K)$ , где  $P$  - набор проекций сцены на заданные плоскости, расположенные в трехмерном пространстве, а  $K$  - информация о связи проекций между собой. Проекция задается преобразованием  $S$  декартовых координат в пространстве в декартовы координаты на плоскости. В общем случае каждая проекция  $P_i$  может содержать один или несколько из следующих возможных атрибутов, записанных в виде изображений: карта цвета (фотография), карта глубины, карта нормалей, карта коэффициентов материалов.

Исходные изображения могут храниться в различных форматах, включая те, которые могут прогрессивно передаваться по сети (например, с вейвлетным сжатием). При этом изображения не обязательно должны иметь одинаковые размеры – понижение разрешения уменьшает размер модели при соответствующем ухудшении качества.

Каждая проекция может обладать различными атрибутами, что позволяет использовать такое представление для различных задач компьютерной графики. Например, для индивидуального объекта исходные изображения могут быть расположены как грани описывающего параллелепипеда. Для более детализированных частей модели могут создаваться карты глубины высокого разрешения.

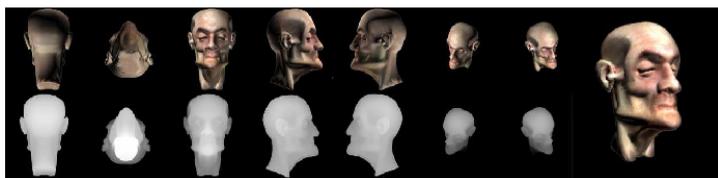


Рисунок 2 Пример модели, состоящей из карт цвета и карт глубины

Информацию  $K$  о связи проекций между собой допускается хранить в виде соответствий в форме точек либо примитивов (треугольников, многоугольников).

В п.3. описываются форматы данных, разработанных для универсального представления.

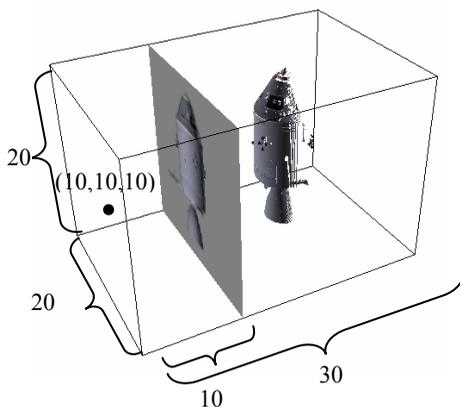
Были разработаны два формата:

- расширенный, на основе языка XML;
- ограниченный вариант, предложенный и включенный в спецификацию международного стандарта для мультимедийных данных MPEG-4 AFX (AFX – Animated Framework Extension) как часть набора представлений, основанных на изображениях (DIBR – Depth Image-Based Representation).

Расширенный формат позволяет с помощью языка XML описывать модель в разработанном представлении.

Пример записи файла в расширенном формате:

```
<xml version="1.0"
encoding="UTF-8"?>
<uibr version="1.0">
<map id=1>
  <camera id=1>
    <matrix rows=4 cols=4>
      <row>0.1 0 0 -1</row>
      <row>0 0.1 0 -1 </row>
      <row>0 0 -0.1 -1 </row>
      <row>0 0 0 1</row>
    </matrix>
  </camera>
  <colormap id=1>
    <image>color1.jpg</image>
  </colormap>
  <depthmap id=1>
    <nearclip-
plane>10</nearclipplane>
    <farclipplane>30</farclipplane>
    <image>depth1.jpg</image>
  </depthmap>
  <colormap id=1>
    <image>color2.jpg</image>
  </colormap>
</map>
</uibr>
```



**Рисунок 3** Пример модели на основе изображения с глубиной и ее запись в разработанном формате

Ограниченный вариант формата позволяет задавать представление на основе изображений с глубиной. В него вошли описания параметров проекций, причем каждая проекция может хранить только изображение с глубиной или видеопоследовательность с глубиной.

Для обоих форматов были разработаны программные модули для загрузки и сохранения.

В п.4. рассматриваются вопросы обработки и оптимизации моделей в представлении на основе изображений с глубиной (включенного в MPEG-4 AFX).

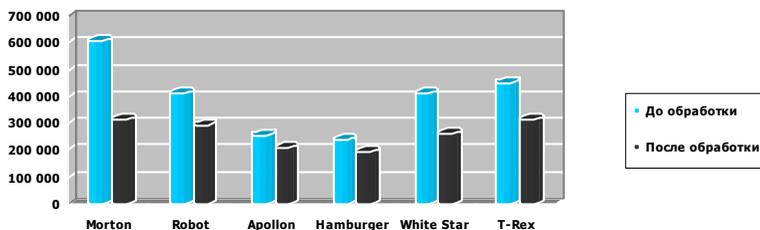
Задача удаления избыточности состоит в том, чтобы минимизировать число участков поверхности, встречающихся в различных изображениях объекта. *Дискрет* (область пространства, соответствующая пикселю на проекции) считается избыточным, если существует другой дискрет или набор дискретов, описывающий цвет той же части поверхности, но с лучшим качеством. Избыточные дискреты появляются в процессе работы трехмерных сканеров или алгоритмов машинного зрения.

В работе описывается разработанный алгоритм поиска и удаления таких дискретов из исходных изображений. Для каждого пикселя карты глубины

восстанавливается соответствующий трехмерный дискрет и вычисляется оценка, характеризующая качество описания дискрета на данной карте глубины (считается, что чем меньше угол между нормалью к поверхности дискрета и направлением на камеру, соответствующую данной карте, тем лучше пиксель карты глубины описывает дискрет).

Далее все дискреты, расположенные достаточно близко друг к другу в трехмерном пространстве (т.е. те, которые представляют одну и ту же часть трехмерной поверхности), сравниваются по их оценкам. Дискрет с наивысшей оценкой сохраняется, а остальные удаляются из соответствующих изображений. Также рассматриваются случаи, когда один дискрет полностью покрывает другой, в этом случае могут сохраняться несколько дискретов.

Результатом этого алгоритма является модифицированное представление, в котором каждая часть поверхности присутствует только на одном изображении. Тестирование показывает, что применение этого алгоритма ведет не только к улучшению качества и увеличению скорости экранизации, но и к уменьшению размера модели (см. Рисунок 4).



**Рисунок 4** Результаты удаления дубликатов для различных моделей. На диаграмме показано число дискретов до обработки (светлый столбец) и после обработки (темный столбец)

Достоинством представления на основе изображений является то, что для его сжатия, прогрессивной передачи по сети и т.п. могут быть использованы существующие алгоритмы работы с изображениями.

Методы сжатия изображений с потерями позволяют получить намного более высокую степень сжатия по сравнению с алгоритмами без потерь. Это достигается через использование свойства когерентности цветов соседних пикселей (соседние пиксели должны иметь близкие цвета). После сжатия с потерями цвета изображения меняются. Эти изменения малозаметны для человеческого глаза, но нарушают условие цветовой консистентности для моделей на основе изображений с глубиной, что приводит к проблемам во время экранизации и редактирования: цвета на границах объекта начинают смешиваться с фоном, «перетекать» со смежных внутренних границ объекта. Также проблемой является то, что внутренние пиксели могут изменить свой цвет по-разному на разных изображениях, что приведет к заметным артефактам во время экранизации.

Для минимизации артефактов сжатия был использован набор алгоритмов предобработки изображений. Основная идея подхода заключается в изменении цветов пикселей, не занятых объектом, таким образом, чтобы переход цвета на границе был незаметен. Новый цвет вычисляется в зависимости от используемого алгоритма сжатия изображения (основанный на блоках для JPEG, непрерывный для вейвлетного сжатия).

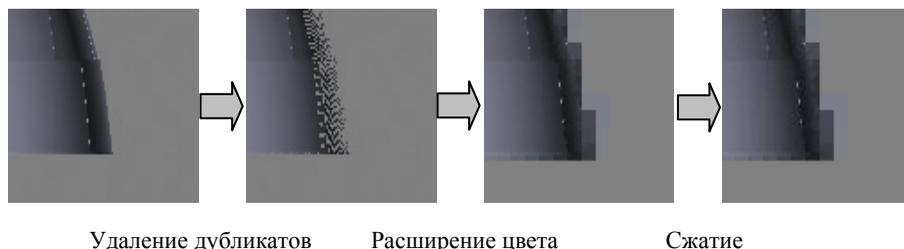


Рисунок 5 Этапы обработки модели

Также для предложенного представления были разработаны методы редактирования моделей. Разработанные алгоритмы работают по следующей схеме: пользователь модифицирует одно из изображений модели, остальные изображения изменяются автоматически, для поддержания консистентности модели.

Реализованы две операции редактирования: рисование цветом на поверхности объекта и так называемое «рисование светом». Последний метод редактирования предназначен для эмуляции освещения дополнительными источниками света. Пользователь задает положение, ориентацию и цвет источника света. Затем он выделяет часть пикселей исходного изображения, которые автоматически перекрашиваются в зависимости от выбранной модели освещения.

Разработанный многомасштабный метод экранизации моделей на основе изображений с глубиной описан в **третьей главе**.

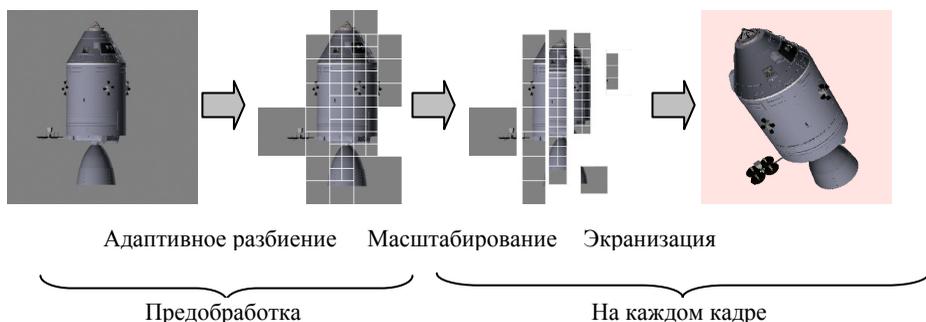
В **п.1.** обосновывается актуальность задачи разработки многомасштабного алгоритма экранизации изображений с глубиной и дается постановка задачи.

На входе алгоритма – модель в представлении на основе изображений с глубиной с полностью калиброванными исходными камерами (ортографическими или перспективными). Карты глубины соответствуют картам цвета, возможно, за исключением разрешения изображений. Разрешение изображений может быть высоким, число проекций – до нескольких десятков. Предполагается, что карты прошли этап оптимизации и не содержат дубликатов (или содержат пренебрежимо малое их количество). Задача заключается в интерактивной генерации изображения данной модели для произвольного положения виртуальной перспективной камеры.

В **п.2.** рассматриваются существующие методы экранизации изображений с глубиной.

Наиболее естественным алгоритмом экранизации изображений с глубиной является реконструкция полигональной поверхности для каждой карты глубины с последующей экранизацией обычными средствами для полигональной геометрии, например, с помощью графической библиотеки OpenGL. Методы, основанные на реконструкции полигональных поверхностей, могут давать хороший результат для небольшого набора карт глубины, при экранизации с помощью графических ускорителей. Однако алгоритмическая сложность таких методов не позволяет эффективно их использовать для экранизации больших наборов карт глубины высокого разрешения, особенно на платформах без аппаратной поддержки экранизации полигональной геометрии.

Другой класс методов экранизации изображений с глубиной использует регулярную структуру входных изображений. Для генерации нового вида исходных данных они используют алгоритмы деформации изображений.



**Рисунок 6** Схема разработанного алгоритма экранизации

Основной задачей для всех методов экранизации, основанных на изображениях, является необходимость реконструировать поверхность после растеризации, т.к. в силу дискретности исходных данных и их неполноты на финальном изображении появляются артефакты в виде пустот («дырок»). Реконструкция осуществляется либо с помощью интерполяции цветов известных пикселей, либо с помощью использования так называемых *сплатов* (splat): в этом случае в целевом изображении на месте пикселя рисуется небольшое изображение специальной формы и цвета. Перекрывающиеся сплаты смешиваются и заливают пустоты.

Существующие многомасштабные методы экранизации на основе изображений позволяют увеличивать скорость экранизации за счет использования иерархических структур, различные слои которых содержат модели разного уровня детализации. Однако такие методы характеризуются повышенным объемом хранимых данных и необходимостью разработки специфических алгоритмов сжатия.

Предлагаемый метод экранизации рассматривается в п.3. Разработанный алгоритм базируется на экранизации без реконструкции полигональной поверхности. В отличие от существующих аналогов, он не использует дополнительных иерархических структур и может применяться вместе с существующими методами экранизации изображений с глубиной.

Алгоритм состоит из трех этапов: этапа статической предобработки, этапа динамической предобработки и этапа репроекции.

На этапе статической предобработки из исходной модели создается эквивалентная новая, оптимизированная для экранизации. Для получения нового представления карты глубины разбиваются на участки разного размера, каждый из которых выбран с целью уменьшения вариации глубины в каждом (для этого используется метод главных компонент). Полученные области становятся новыми картами глубины.

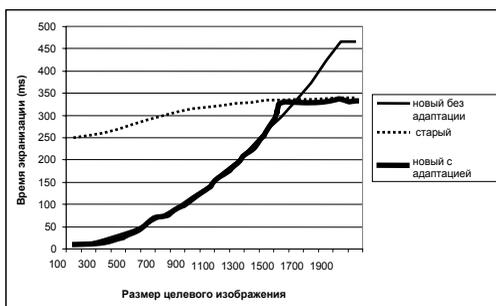


Рисунок 7 Сравнение времени синтеза одного кадра для многомасштабного алгоритма и алгоритма без масштабирования

Этап динамической предобработки выполняется непосредственно перед каждым кадром экранизации. Каждая карта глубины масштабируется таким образом, чтобы после проекции на экран максимальный размер дискрета карты глубины был сопоставим с пикселем экрана и чтобы скорость экранизации была оптимальной для данного аппаратного обеспечения. Особое внимание уделяется условиям, при которых исходный, не многомасштабный, метод работает быстрее. Это возможно, когда размеры исходного и уменьшенного изображения отличаются не сильно. Были разработаны специальные метрики и проведены тесты, позволяющие определить адаптивный критерий, при выполнении которого нет необходимости в масштабировании изображений. Таким образом, многомасштабный алгоритм с адаптацией гарантированно работает не медленнее, чем алгоритм без масштабирования (см. Рисунок 7).

На этапе репроекции выполняется экранизация обработанных карт глубины с помощью алгоритма репроекции с использованием OpenGL.

В п.4. рассматриваются вопросы реализации предложенного алгоритма и приводится архитектура разработанной системы Visual Framework, в рамках которой были реализованы алгоритмы.

Заключение дается в п.5. Достоинствами разработанного алгоритма является возможность интерактивной экранизации карт глубины большого размера с постоянной скоростью, регулируемое качество экранизации, автоматическая адаптация под возможности аппаратуры.



Рисунок 8 Примеры экранизации модели с разными уровнями детализации

**Четвертая глава** посвящена применению представления на основе изображений в клиент-серверной системе, предназначенной для адаптивной экранизации моделей на мобильных клиентах в условиях низкой пропускной способности сети.

В п.1. приводятся требования к системе, проводится обзор существующих решений и обосновывается актуальность работы.

В связи с бурным ростом сетевых приложений и возрастающей неоднородностью клиентских устройств (ПК, мобильные телефоны, карманные компьютеры и т.п.) особую актуальность приобретают технологии адаптивной и удаленной экранизации сложных трехмерных данных. Эти технологии могут использоваться в таких задачах как трехмерная навигация, научная визуализация, обучение, моделирование и т.п. Сложность заключается в том, что на данный момент мощностей и объемов памяти мобильных устройств не хватает для обработки больших объемов трехмерных данных, которые обычно используются на практике.

В диссертации рассматриваются следующие требования к системе:

- Данные хранятся на сервере в произвольном представлении, допускающем визуализацию с помощью аппаратных ускорителей. Модель представляет собой высокодетализированный трехмерный объект (статуэтка, модель телефона, ноутбука и т.п.).
- Клиент и сервер связываются по беспроводному соединению Wi-Fi (802.11b) или GPRS.
- В качестве клиента используется карманный компьютер на базе операционной системы Windows Mobile 2003.

- Пользователь должен иметь возможность вращать объект вокруг центральной точки, расположенной внутри описывающего объема объекта, иметь возможность изменить центр вращения (который должен оставаться внутри описывающего параллелепипеда модели), а также приближать/удалять объект.

В п.2. приводится обзор существующих решений. Их можно разделить на три класса: алгоритмы, основанные на передаче трехмерных данных клиентскому устройству с последующей локальной визуализацией, алгоритмы, основанные на потоковой передаче видео и алгоритмы, являющиеся комбинацией двух предыдущих подходов.

К недостаткам первого подхода относится то, что современные мобильные устройства не имеют возможности интерактивной экранизации сложных данных. Алгоритмы интерактивной генерации видеопоследовательности и ее точной передачи клиентскому устройству позволяют решить эту проблему, используя возможности сервера для синтеза изображения сложной модели. Однако такие методы очень чувствительны к качеству сетевого канала связи, что в условиях низкой скорости передачи данных приводит к заметным задержкам экранизации на клиентском устройстве. Комбинация первых двух подходов позволяет решать как проблему недостаточной мощности мобильного устройства (за счет генерации упрощенной модели на сервере и передачи ее клиенту), так и проблему задержек (т.к. во время ожидания данных устройство клиента может экранизировать имеющуюся частичную информацию). Разработанное решение использует именно этот подход.

В п.3. предлагается метод решения задачи, построенный на использовании изображений с глубиной (представления, описанного в главе 2) в качестве частичного представления. Показывается эффективность такого решения задачи.

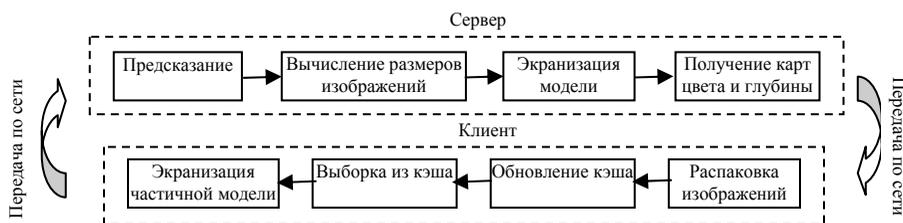


Рисунок 9 Общий вид разработанной системы

Предлагаемый метод заключается в интерактивном построении на сервере видеозависимой модели, основанной на изображениях с глубиной. Затем построенная модель сжимается и передается клиенту. Модель фактически представляет собой изображение синтезированной серверной модели вместе с картой глубины для текущего положения виртуальной камеры, определяемого пользователем клиентского устройства. Передача дополнительно карты глубины позволяет клиенту с тонким каналом самостоятельно строить промежуточ-

ные кадры, создавая плавную анимацию в период ожидания получения следующего кадра.

Использование представления на основе изображений ограничивает сложность алгоритма экранизации разрешением изображения. Поскольку обычно мобильные устройства обладают небольшим экраном, предлагаемый алгоритм позволяет интерактивно экранизировать на мобильном устройстве сколь угодно сложные сцены, интерактивный синтез которых возможен на стороне сервера.

Серверная часть системы описана в **п.4**. На сервере выполняется следующий алгоритм. Сначала обрабатывается информация о положении наблюдателя на клиенте, производится предсказание будущего положения с учетом средней задержки передачи. Затем производится обычная визуализация сцены с использованием OpenGL. Буфер цвета и буфер глубины сохраняются, формируя, таким образом, видеозависимое представление трехмерных данных. Изображения масштабируются с учетом средней задержки передачи и скорости экранизации на клиентском устройстве. На последнем этапе данные подвергаются обработке и сжатию, которые были описаны в главе 2, и пересылаются клиенту.

Реализация клиентской части системы, разработанной для карманного компьютера на базе операционной системы Windows Mobile 2003, приводится в **п.5**.

Пользователь может вращать, приближать или удалять модель, перемещаться и изменять направление взгляда. Изменения в положении наблюдателя передаются по обратному каналу на сервер. Во время синтеза изображения клиент использует методы деформации изображений с глубиной, описанные во второй главе, для построения последовательности промежуточных кадров. Также реализован алгоритм кэширования присылаемых сервером видов сцены (с обратной связью), что позволяет оптимизировать процесс пересылки данных от сервера за счет использования близких видов. Это особенно актуально в случае тонкого канала связи.

Пересылаемое представление актуально только для небольших отклонений от исходных параметров наблюдателя, поэтому в случае существенных задержек (очень узкого канала связи) могут появляться артефакты – пропадающие или искаженные части модели. Однако на практике при стабильном соединении артефакты малозаметны.

Программная реализация и результаты тестирования описаны в **п.6**. Тестирование проводилось на КПК HP iPaq 4150 (разрешение экрана 240x320). Разработанные алгоритмы позволяют создавать плавную анимацию с небольшим количеством артефактов даже при достаточно медленной связи (сравнимой с модемным соединением).



Рисунок 10: Примеры визуализации моделей на карманном компьютере (слева направо – высокое, среднее, низкое качество)

В заключении сформулированы основные результаты работы.

## Основные результаты работы

1. Предложено новое представление моделей трехмерных объектов на основе нескольких изображений с глубиной, позволяющее эффективно хранить, передавать по сети и визуализировать дискретные модели объектов.
2. Разработан алгоритм многомасштабной экранизации предложенного представления, обеспечивающий более высокую скорость экранизации по сравнению с известными алгоритмами.
3. Предложен алгоритм клиент-серверной экранизации потока трехмерных данных, ориентированный на мобильные клиентские устройства (включая карманные ПК) в условиях низкой пропускной способности сети.
4. На основе предложенных алгоритмов и разработанного представления моделей трехмерных объектов создана программная система их интерактивной обработки и визуализации.

## Работы автора по теме диссертации

1. А. Игнатенко, А. Конушин, «Система для экранизации и моделирования на основе изображений с глубиной», труды конференции Графикон-2003, с. 169-172., Москва, Россия, 2003.
2. Y. Bayakovski, L. Levkovich-Maslyuk, A. Ignatenko, A. Konushin, D. Timasov, A. Zhirkov, Mahnjin Han, and In Kyu Park, “Depth Image-based Representations for static and animated 3D objects”. In Proceedings of ICIP-2002, vol. 3, pp. 25-28, 2002.
3. Y. Bayakovski, L. Levkovich-Maslyuk, A. Ignatenko, A. Konushin, D. Timasov, A. Zhirkov, Mahnjin Han, In Kyu Park, “Depth Image-based Representations and Compression for Static and Animated 3D Objects”, IEEE Transactions on Circuits & Systems in Videotechnology, Special Issue on MPEG-4), pp. 1032-1045, 2004.
4. A. Ignatenko, B. Barladian, K. Dmitriev, S. Ershov, V. Galaktionov, I. Valiev, and A. Voloboy, “A Real-Time 3D Rendering System with BRDF Materials and Natural Lighting”. In Proceedings of Graphicon-2004, pp. 159-162, Moscow, Russia, 2004.