

Концептуальный компонент в гибридной системе анализа изображений

Геннадий Хахалин, фрилансер

gkhakhalin@yandex.ru

Сергей Курбатов, Алексей Лобзин

Научно-Исследовательский Центр Электронной Вычислительной Техники, Москва, Россия

curbatow.serg@yandex.ru

lobzin@rambler.ru

Аннотация

В работе рассматривается концептуальный компонент гибридной системы анализа изображений. В качестве базы знаний для этого компонента служит прикладная онтология, которая представлена на языке семантического гиперграфа.

Предлагается достаточно гибкое взаимодействие двух компонент гибридной системы. Показано, каким образом привлекаются другие части интегральной системы искусственного интеллекта (синтез изображений и анализ, и синтез естественного языка) для «инспектирования» процесса распознавания и концептуального «дообучения». Многие моменты этих процессов иллюстрированы примерами из прикладной области.

Ключевые слова: гибридная система, концептуальный анализ изображения, предметная онтология, семантический гиперграф.

1. ВВЕДЕНИЕ

При решении многих задач распознавания изображений однокомпонентные системы оказываются неэффективными. И если при распознавании «непроизводных» элементов эффективными оказываются дискриминантные (числовые) методы, то при распознавании композиционных объектов в целом, обладающих существенной структурной информацией, ведущая роль принадлежит структурным (символьным) методам [3].

2. ГИБРИДНАЯ СИСТЕМА АНАЛИЗА

Под задачей полного анализа изображения подразумевается описание распознаваемых объектов и их расположение на изображении в терминах и в структурах языка прикладной онтологии. Результатом анализа должно быть описание ситуации на изображении, состоящей из экземпляров распознанных классов объектов с их означенными характеристиками и отношениями между ними.

На входном изображении в общем случае задается не один образ, а целая ситуация (композиция) и ее надо описать в терминах взаимосвязей означенных структур классов объектов.

В качестве предметной среды рассматривается мир плоских двумерных изображений объектов: планиметрические фигуры и «детские» рисунки (прямоугольники, трапеции, окружности и т.п., и кораблики, домики, ромашки и т.п.).

Выбрана двухуровневая система: нижний уровень, который непосредственно работает с изображением, и верхний уровень, который работает с моделью изображения (по результатам работы системы нижнего уровня), но с мощным языком символьного (концептуального) описания среды. Систему нижнего уровня назовем PR-система (Pattern Recognition), а систему верхнего уровня – KB-система (Knowledge Based). Они работают последовательно с возможностью управления процессом распознавания на нижнем уровне с концептуального уровня, реализуя т.н. «экслюзивный» анализ (см. ниже).

Функциональная схема двухуровневой системы анализа изображений представлена на рис. 1.

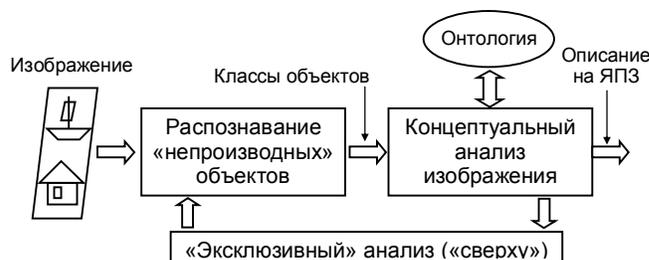


Рис 1: Схема двухуровневой системы анализа изображений.

Данная гибридная система анализа изображений не является изолированной, а представляет собой элемент интегральной системы, в которую входят дополнительно системы концептуального синтеза изображений и анализа и синтеза текстов естественного языка [6]. Взаимодействие этих систем осуществляется на основе общей прикладной онтологии [1].

Для иллюстраций в выбранной предметной области мы используем реальную PR-систему MyScript Notes [2]. Эта система решает стандартные задачи предварительной обработки изображений, включая кодирование, сглаживание, фильтрацию, сегментацию и т.д., избавляя от этих процедур верхний уровень гибридной системы. На вход программы MyScript Notes подаются нарисованные геометрические рисунки. На выходе получаем для каждого класса распознаваемых объектов: имя класса, вероятность распознавания и значения геометрических параметров, привязанных к полю изображения (координатной системе). Кроме того дается альтернативная информация распознавания, когда приводятся возможные другие кандидаты на распознавание данного объекта с соответствующей вероятностью.

Множество распознаваемых объектов ограничивается перечнем планиметрических фигур (отрезок прямой,

треугольник, прямоугольник, стрелки, ромб, дуга, окружность и др.) и классом неизвестных объектов (рисунки произвольной, нераспознаваемой формы - *drawing*).

Пример изображения композиции приведен на рисунке 2. На нем представлены как объекты, распознаваемые программой MyScript Notes (*прямоугольник, окружность* и др.), так и нераспознаваемые на этом уровне классы объектов (*домик, кораблик, паровозик* и др.).

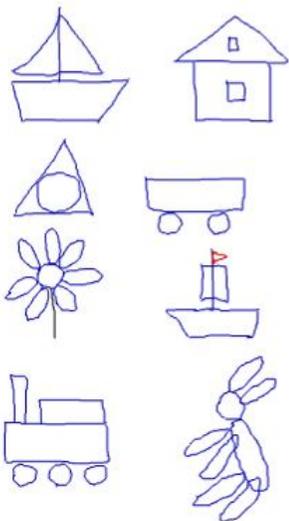


Рис 2: Композиция фигур.

Результатом распознавания, например *Паровозика*, будет список, представленный фрагментом на рис. 3, где указаны имена предполагаемых классов с указанием вероятности распознавания и конкретных характеристик объектов, привязанных к системе координат изображения (подчернутые имена классов имеют максимальную вероятность).

rectangle 0.753047 625.545 395.106 764.597 390.574 777.322 780.953 638.269 785.485
 parallelogram 0.752861 624.92 390.086 764.34 396.322 777.918 785.459 638.498 779.224
rectangle 0.894784 877.211 785.921 876.636 587.349 1391.45 585.859 1392.02 784.432
 parallelogram 0.89469 880.424 785.418 872.933 586.889 1388.14 586.386 1395.63 784.915
rectangle 0.956212 1440.37 1090.52 584.024 1099.19 580.978 798.668 1437.32 789.989
 parallelogram 0.95617 1445.39 1091.08 589.294 1098.52 575.977 798.106 1432.07 790.666
circle 0.67984 1281.93 1187.42 75.2383
 ellipse 0.679622 1281.93 1187.42 78.6176 71.8591 -1.4689
 arc 0.679404 1281.93 1187.42 78.6176 71.8591 -1.4689 0 6.28319
circle 0.609414 746.45 1192.97 82.0498
 ellipse 0.609176 746.45 1192.97 83.6068 80.4929 -0.325574
 arc 0.608938 746.45 1192.97 83.6068 80.4929 -0.325574 0 6.28319
circle 0.609228 1024.86 1177.54 75.0216
 ellipse 0.60899 1024.86 1177.54 80.0398 70.0035 -0.590248
 arc 0.608752 1024.86 1177.54 80.0398 70.0035 -0.590248 0 6.28319

Рис 3: Фрагмент результата работы PR-системы.

Для нераспознанных и нераспознаваемых PR-системой объектов как раз и служит верхний уровень гибридной системы – KB-система.

3. КОНЦЕПТУАЛЬНЫЙ АНАЛИЗ ИЗОБРАЖЕНИЙ

Основным результатом работы KB-системы является описание, состоящее из экземпляров распознанных классов объектов на языке представления знаний. Поскольку система работает в комплексе, результат может быть также

представлен в виде синтезированного изображения композиции или в виде текста на естественном языке.

В качестве основы для системы анализа верхнего уровня выбран метод контекстного целенаправленного анализа изображений, который подробно описан в [1].

Построение описаний в онтологии предполагает наличие достаточно выразительных средств представления знаний о изображениях. Прикладная онтология организуется в виде множества структур, каждая из которых соответствует обобщенному описанию некоторого класса объектов на изображении с его внутренней организацией и взаимосвязями с другими классами.

На наш взгляд наиболее адекватным языком представления концептуальных знаний является язык гиперграфов в качестве расширения семантических сетей, где естественным образом представляются *n*-арные отношения, которые позволяют задавать не только атрибуты объектов, но и представлять структурные, «целостные» описания объектов с взаимосвязями своих компонентов. Здесь вершины соответствуют элементам описываемого изображения, а дуги – отношениям между этими элементами.

Если элементам гиперграфа приписаны цепочки символов (имена понятий и отношений онтологии), то такой гиперграф называется *семантическим гиперграфом* [4, 8].

Фрагмент прикладной онтологии в виде семантического гиперграфа представлен на рис. 3., где красные кружки – концепты онтологии (объекты среды), красные стрелки – родовидовые связи концептов, синие кружки и стрелки – свойства объектов, зеленые замкнутые области – описание структур объектов, бирюзовые стрелки – связь типа «часть-целое» («входит_в_структуру»).

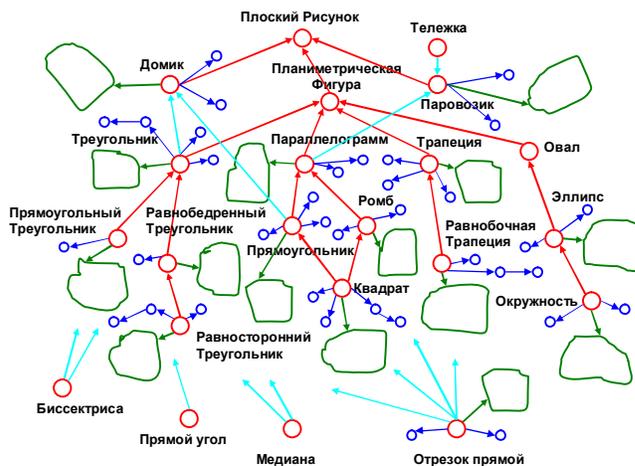


Рис 4: Фрагмент прикладной онтологии.

В прикладной онтологии представлены и понятия нижнего уровня (*точка, треугольник* и т.д.), дублирующие классы PR-системы, для исправления ошибок распознавания и для анализа тех классов, которые PR-системой не распознаются в принципе (например, *трапеция*), и структурные образы рисунков, составленных из классов PR-системы (*кораблики, домики, ромашки* и т.д.).

Алгоритм работы KB-системы заключается в следующем.

Представление результатов работы PR-системы в виде концептов «непроизводных» объектов – получение гиперграфа модели изображения.

Выдвижение кандидатов-гипотез из онтологии на сопоставление (поиск пути в семантическом гиперграфе).

Сопоставление графа модели изображения с графами структур объектов прикладной онтологии (поиск изоморфизма в семантическом гиперграфе). Требуемые по онтологической структуре отношения проверяются на координатной сетке с помощью вызова присоединенных процедур, связанных с этими отношениями.

В случае удачи полного сопоставления – получение означенного экземпляра онтологической структуры понятия.

В случае частичного сопоставления – выделение «ареала» на изображении и поиск гипотетического непроизводного элемента с помощью PR-системы («что хотим найти» - предполагаемый непроизводный объект, «где хотим найти» - ареал, т.е. «место обитания») этого гипотетического объекта, и «с какими ограничениями» - необходимые диапазоны изменения параметров: вся эта информация берется из онтологической структуры).

Эффективный поиск путей в онтологии и выбор наиболее оптимальных гипотез осуществляется «хождением» по семантическому гиперграфу: по связям «вх_в_структуру», «род-вид» и др.

3.1 Эксклюзивный анализ

Эксклюзивный анализ работает в двух режимах:

- Последовательный анализ: PR-система полностью обрабатывает изображение композиции, а потом включается KB-система и при «коллизиях» она обращается к PR-системе.
- Итеративный анализ: PR-система частично обрабатывает изображение, после этого подключается KB-система, выдвигает гипотезы «чего и где искать» и передает управление на PR-систему.

В первом случае KB-система выдвигает гипотезу о «дораспознавании» на реальном изображении. Для вызова соответствующих процедур PR-системы она формирует управляющую информацию, которая включает: что надо искать (это известно из онтологической структуры-гипотезы); где надо искать (вычисленный сегмент-«ареал» по модели изображения); с какими ограничениями, т.е. с необходимыми параметрами (из онтологической структуры).

После получения управляющей информации PR-система реально анализирует выделенную область и реализует процедуру распознавания только в этом сегменте, как будто на целом изображении присутствуют только компоненты этого объекта. Реальные эксперименты показывают, что присутствие композиции усложняет процесс распознавания отдельных объектов, поскольку глобального априорного решения задачи сегментации не существует.

Если в результате повторного распознавания будет выделен предполагаемый объект, то KB-система корректно выполнит свое предназначение. В противном случае выдвигаются

другие гипотезы или система распознавания обращается к пользователю за советом.

При итеративном анализе реализуется стратегия «сверху-вниз» (от онтологии к изображению) и при «стандартном» распознавании. По частичным результатам PR-системы выдвигаются гипотезы, которые последовательно проверяются на реальном изображении. При этом процесс сегментации носит «осмысленный» целенаправленный характер, т.к. она проводится на основе частичных результатов распознавания, выдаваемых PR-системой, в рамках проверяемых гипотез, которые формируются KB-системой.

3.2 Использование концептуального синтеза

Концептуальный синтезатор изображений по означенному онтологическому описанию графической ситуации строит и визуализирует на плоскости отображения графическое изображение. Он состоит из планировщика и визуализатора. Планировщик реализует «обход» онтологической структуры и выделение описаний подобъектов для визуализатора; согласование параметров подобъектов и означивание их свойств; генерация текста для визуализации. В качестве визуализатора выбран пакет OpenGL, на вход которого поступает текст интерпретации планировщика, а выходом является 3D или 2D изображение (более подробно см. в [5, 7]).

Этот синтезатор используется в двух ипостасях: как средство отображения результатов работы гибридной системы распознавания и как средство отладки при модификации и пополнении онтологии.

В первом случае визуализируются результаты распознавания. Небольшой фрагмент визуализации распознанной композиции (рис. 2) дан на рис. 5 (представлен конкретный экземпляр образа *Паровозик*).



Рис 5: Фрагмент визуализации результата распознавания.

Наименование концепта необходимо, поскольку без него не совсем понятно на каком уровне распознан объект: на уровне составляющих (три прямоугольника и три окружности распознает PR-система) или на уровне целостного объекта (распознает KB-система). Синтезированное изображение при этом будет одно и то же.

Другое использование концептуального синтезатора описано в следующем пункте.

3.3 Модификация и пополнение онтологии

Естественно KB-система «видит только то, что знает», т.е. распознает комплексные объекты, информация о которых присутствует в онтологии. В интегральной системе модификация и расширение онтологии предполагает использование двух каналов: для инженера по знаниям (для

разработчика) – непосредственный ввод структур новых концептов или модификацию старых на языке семантического гиперграфа; для пользователя – то же самое на естественном языке. В отличие от стандартных методов обучения для статистических систем такое обучение носит «дедуктивный» характер. Чтобы такое дедуктивное обучение не приводило к «хаосу», необходимо применить методику наполнения онтологии с верификацией вводимых описаний с помощью системы концептуального синтеза изображений. По новой введенной структуре система синтезирует множество изображений. Если с точки зрения учителя (разработчика или пользователя) множество корректно, то структура фиксируется в онтологии. Если среди синтезированных объектов присутствуют изображения, не относящиеся к данному классу, то структура модифицируется. И так до тех пор, пока множество не будет корректным.

Проиллюстрируем это на примере концепта *Паровозик*. До ввода структуры объекта *Паровозик* в онтологию гибридная система распознавания представит его описание в виде взаимосвязанного набора прямоугольников и окружностей. Введя в онтологию описание структуры понятия *Паровозик*, как это представлено на рис. 6., и получив множество синтезированных изображений, данных на рис. 7. по этой структуре, можно предположить, что теперь гибридная система распознавания будет распознавать изображение как целостный объект планиметрической среды.

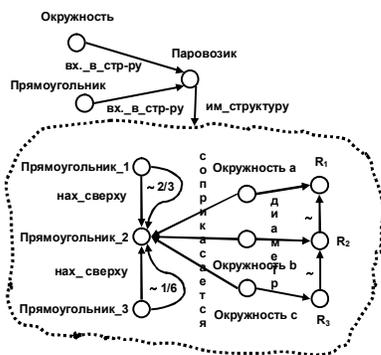


Рис 6: Онтологическая структура концепта *Паровозик*.

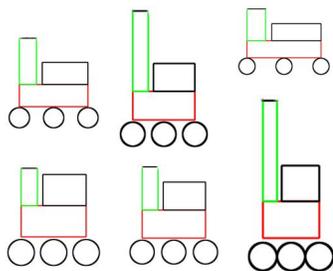


Рис 7: Множество синтезированных изображений концепта *Паровозик*.

4. ЗАКЛЮЧЕНИЕ

Подобная гибридная система дает возможность говорить о «понимании» изображения, если рассматривать этот процесс в рамках интегральной системы.

Дальнейшая работа в рамках данной системы направлена на расширение онтологии для других областей и на более скрупулезную проработку вопросов верификации онтологии.

5. БЛАГОДАРНОСТИ

Работа была выполнена при поддержке гранта РФФИ, проект № 12-07-00531а.

6. ССЫЛКИ

[1] Khakhalin G., Kurbatov S., Naidenova K., Lobzin A. *Integration of the Image and NL-text Analysis/Synthesis Systems. In book: "Intelligent Data Analysis for Real-Life Applications: Theory and Practice" (Ed. Rafael Magdalena et al.). – N.Y., USA: IGI Global, 2012, p. 160-185.*

[2] MyScript Notes, 2010. *Режим доступа: http://www.visionobjects.com/en/webstore/myscript-studio/description/*. – [Электронный ресурс].

[3] Roli Fabio, Serpico Sebastiano B., and Vernazza Gianni. *A Hybrid System for Two-Dimensional Image Recognition // Proceedings of the IEEE, VOL. 84, NO. 11, November 1996, pp. 1659-1680.*

[4] Баранович А.Е. *Семиотико-хроматические гиперсети: унифицированная модель представления знаний // Труды Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» OSTIS-2011, Минск, Беларусь. – 2011. – С. 71-86.*

[5] Курбатов С.С., Литвинович А.В., Лобзин А.П., Хахалин Г.К. *Концептуальный синтез графических образов по структурам прикладной онтологии // Труды XIII национальной конференции по Искусственному Интеллекту, Т. 2.. – Белгород: БГТУ, 2012. – С. 120-127.*

[6] Курбатов С.С., Лобзин А.П., Найденкова К.А., Хахалин Г.К. *Гибридная схема анализа изображений // Труды Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» OSTIS-2012. Минск, Беларусь. – 2012. – С. 327-334.*

[7] Литвинович А.В. *Система синтеза изображений по тексту на естественном языке // Динамика сложных систем — XXI век, № 2, 2013, С. 65-68.*

[8] Хахалин Г.К. *Прикладная онтология на языке гиперграфов // Труды второй Всероссийской Конференции "Знания-Онтология-Теория" (ЗОНТ-09). Новосибирск, 2009, с. 223-231.*

Об авторах

Геннадий Хахалин – фрилансер. Его адрес: gkhakhalin@yandex.ru.

Сергей Курбатов – к.т.н., НИЦЭВТ. Его адрес: curbatow.serg@yandex.ru.

Алексей Лобзин – с.н.с., НИЦЭВТ. Его адрес: lobzin@rambler.ru.