



OSTIS-2012

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

ГИБРИДНАЯ СХЕМА АНАЛИЗА ИЗОБРАЖЕНИЙ

С.С. Курбатов *, А.П. Лобзин *, К.А. Найденова **, Г.К. Хахалин *

** Научно Исследовательский Центр Электронной Вычислительной Техники, г. Москва, Россия*

curbatow.serg@yandex.ru

lobzin@rambler.ru

gkhakhalin@yandex.ru

*** Военно-медицинская академия, г. Санкт-Петербург, Россия*

naidenovakseniya@mail.ru

В работе рассматривается схема анализа изображений, соединяющая преимущества числового и символического методов распознавания. Гибридизация двух подходов при использовании прикладной онтологии и при наличии в интегральной системе лингвистического транслятора и синтезатора изображений позволяет выйти на качественно более высокий уровень понимания изображений. В качестве языка представления знаний для символического метода используется семантический гиперграф.

Ключевые слова: гибридная система, прикладная онтология, семантический гиперграф, числовые и символические методы распознавания.

свойств подобной системы.

ВВЕДЕНИЕ

Разработчики систем распознавания вообще и, в частности, систем распознавания изображений чаще всего придерживаются тенденции разрабатывать системы, основанные на каком то одном из подходов, и пытаются доказать преимущества выбранного подхода над другими. Такой взгляд не всегда приводит к успеху, поскольку для решения проблемы анализа изображений часто требуются различные виды представления и обработки данных. В этой ситуации наиболее адекватным представляется попытка для решения задач восприятия и анализа изображений интегрировать разные подходы, разработанные в этой области, в «гибридную систему», чтобы полностью использовать преимущества каждой составляющей, а не пытаться решить их с помощью единственного подхода.

В данной работе кратко анализируется спектр существующих методов распознавания, обосновывается выбор составляющих гибридной системы и рассматриваются функции системы анализа изображений.

Изложение иллюстрируется примерами из проблемной области «Планиметрия», включающей как стандартные плоские фигуры, так и «детские» рисунки, составленные из этих фигур. В заключении дается характеристика синергетических

1. Типология методов распознавания образов и анализа изображений

Наши представления о компьютере как о «числовой мельнице» породили классические числовые методы распознавания (первое направление в этой области), основанные на цифровом представлении изображений. Успехи в исследованиях по искусственному интеллекту рассматривают компьютер как «интеллектуальный символический процессор», способный запоминать и обрабатывать информацию, не выраженную числами. Такой взгляд породил символический подход (второе направление) к обработке изображений, который позволяет решать не только задачи распознавания, но и задачи понимания изображений, включающие вывод по символической информации. Третьим направлением в обработке изображений является нейросетевой подход, который, как полагают, подражает человеческой системе восприятия и представляется более эффективным, чем современные компьютеризированные системы.

Коротко опишем «наполнение» каждого из этих направлений, отмечая преимущества и недостатки. Распределение методов распознавания по направлениям представлено на рисунке 1.

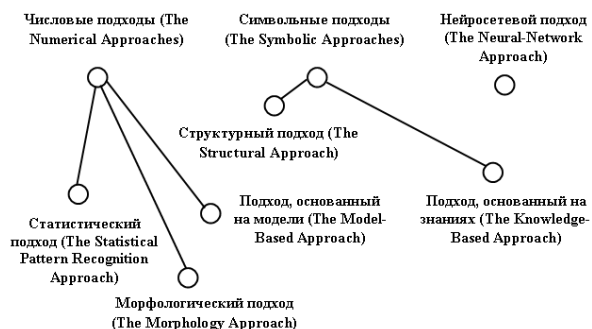


Рисунок 1 – Типология методов распознавания

Статистическое распознавание образов (The Statistical Pattern Recognition) определяет проблему распознавания изображений как задачу классификации на признаках, выделенных на изображении. Объекты представляют собой n -мерные векторы в пространстве признаков, а процесс принятия решения основывается на вычислении меры близости наблюдаемого вектора к области определения искомого объекта. В качестве возможных мер близости используются: расстояние, дискриминантная функция, вероятностные функции и т.п. Классификация проводится в соответствии со статистическими методами по многомерным функциям плотности условной вероятности [Горелик и др., 2004].

Фактически при таком подходе не требуется каких-либо конкретных знаний о прикладной области. Эти преимущества оплачиваются типичной трудностью в использовании априорных знаний о прикладной области, включая контекстуальную и структурную информацию.

Преимуществами подобных систем распознавания являются относительная эффективность, «прозрачная» постановка задачи, наработанный потенциал и имеющие признание методы обучения.

Подход, основанный на эталонах (моделях) (The Model-Based Approach). Обычно цели этого подхода – распознать объекты и оценить их положение в зашумленных средах (например, содержащих частично перекрывающиеся объекты при анализе сцен). Распознавание выполняется сопоставлением данных изображения с моделями (эталонами) объектов. Модели описывают формы объектов, которые обычно устойчивы. Эталоны чаще всего основываются на пространственно локализованных признаках, например, ограничивающие грани, типы углов и т.п. Процесс поиска выполняется в пространстве всех возможных соответствий между данными и моделями. Процесс принятия решения представляет собой сравнение со всеми эталонами и вычисление критерия согласованности (например, корреляционной функции). Этот подход хотя и прост по логике (в основном над исходными изображениями производится лишь предварительная обработка с целью очистки изображения от шумов), но требует больших

вычислительных затрат, поскольку сравнение приходится проводить для всех возможных положений объекта в поле изображения [Катыс, 1990].

Морфологический подход (The Morphology-Based Approach). Под морфологическим анализом понимаются методы решения задач узнавания, классификации объектов, выделения отличий в сценах по их изображениям, оценивания параметров объекта по его изображению, основанные на исследовании математических моделей изображений [Животников и др., 2005]. Так как на практике некоторые детали модели, связывающей объект с его изображением, как правило, неизвестны, (например, условия освещения, оптические свойства поверхностей объекта, характеристики системы формирования изображения), то можно указать лишь множество изображений, порожденное данной сценой при всех возможных параметрах модели формирования изображений. Это множество изображений и называется его формой. Если это множество выпукло и замкнуто в нормированном пространстве всех изображений, а расстояние между изображениями – сильно выпуклая функция, то с формой изображения сцены можно однозначно связать проектор на это множество. В терминах этих проекторов и решаются названные задачи анализа изображений. Подобные методы достаточно ресурсоемкие, но их развитие, основанное на нелинейных проекторах и теоретико-возможностных моделях, оказываются достаточно эффективны при неформализованном характере данных и ограниченном размере обучающей выборки [Пытьев, 2007].

Структурное распознавание образов (The Structural Pattern Recognition) имеет целью распознавать объекты на изображении на основе синтаксического и контекстуального описания [Горелик и др., 2004]. Эти методы предусматривают включение в процесс распознавания информации о структуре объекта и требуют установить, какие элементы-подобразы можно выделять в объекте, насколько они могут или должны быть элементарны и в каких отношениях находятся. Идея структурных методов состоит в изучении и использовании структур взаимоотношений характерных элементов объекта; она реализуется построением описания сложного объекта в виде иерархической структуры более простых подобразов.

При реализации структурного метода используется аналогия между построением объекта с помощью соединения различными способами подобразов и построением фраз и предложений языка с помощью соединения слов, состоящих, в свою очередь, из букв.

В рамках этого подхода для решения задач распознавания используются различные лингвистические конструкции, образованные словарем признаков и грамматикой — правилами конструирования фраз, в совокупности

описывающих классы и подлежащий распознаванию объект. В результате каждый объект получает свое представление с помощью некоторого набора непроемких элементов и ряда фиксированных синтаксических операций.

Система должна обладать способностью обнаруживать синтаксические связи, существующие в объекте. Решение о синтаксической правильности представления объекта, т. е. о принадлежности его к определенному классу, задаваемому определенной синтаксической системой или грамматикой, вырабатывается синтаксическим анализатором (блоком грамматического разбора). При выполнении грамматического разбора анализатор обычно воспроизводит полное синтаксическое описание объекта в виде дерева грамматического разбора, если соответствующий объект является синтаксически правильным. Структурные методы распознавания базируются на порождающей грамматике [Новикова, 2008].

Подход, основанный на знаниях (The Knowledge-Based Approach), интенсифицирует использование различных видов явных знаний, выраженных в символьной форме, и управление знаниями с помощью методов символьной обработки [Roli et al., 1996]. Такие знания касаются описаний объектов и их возможных конфигураций, стратегий распознавания изображений («управляющие знания») и т.д. В области искусственного интеллекта существуют методы для эффективного представления вышеперечисленных знаний (семантические сети, фреймы, продукционные системы и онтологии).

Модели и данные об изображениях обычно представляются на нескольких иерархических уровнях абстракции. Простыми примерами уровней являются уровень частей объекта и уровень целостных объектов. В этом случае модели частей объекта сопоставляются с признаками, тогда как группы смежных признаков («сложные признаки») сопоставляются с моделями целостного объекта. Обычно применяется парадигма «гипотеза-проверка»: гипотезы распознавания генерируются на основе доступных знаний и проверяются с помощью процесса сопоставления. Генерирование гипотез обычно выполняется по схемам вывода, которые реализуются по правилам типа «ЕСЛИ-ТО».

Подход, основанный на знаниях, использует более обширные и конкретные знания о приложении. Такие знания позволяют решать сложные задачи распознавания изображений. С другой стороны, использование знаний, специфических для конкретного приложения, часто делает эти системы распознавания подходящими только для ограниченных областей. Тем не менее, для задач распознавания, включающих много атрибутов и классов, этот подход позволяет получать оптимальные решения за разумную цену при использовании соответствующих экспертных знаний о проблемной области, декомпозиции задачи

распознавания на подзадачи и представление данных и моделей на различных уровнях абстракции.

Этот подход вбирает многое из структурного распознавания, только языки, используемые в данном подходе, являются более мощными и гибкими.

Нейросетевой подход (The Neural-Network Approach) предназначен для построения параллельных компьютерных архитектур, называемых искусственными нейронными сетями, моделирующими биологические нейронные сети. Обычно каждый «нейрон» сети, реализует нелинейную вход/выходную характеристику и связывается с другими элементами посредством связей с весами. Алгоритмы обучения позволяют значения весов настраивать на характеристики входных данных. Были разработаны различные нейросетевые модели, каждая отличающаяся топологией сети, вход/выходными характеристиками нейронов и используемым алгоритмом обучения [Николаев и др., 2003].

При данном подходе, также как и при статистическом, не закладывается каких-либо конкретных знаний о прикладной области. В этом случае знания включаются в значения весов и в топологию сети. А, следовательно, нет возможности в явном виде использовать априорные знания о прикладной области.

2. Задача анализа изображений

В системах анализа изображений используются широкий спектр описаний: от набора признаков до структурных (символьных) представлений. В нашем случае 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, в разной степени «полноты» в зависимости от внешних критериев.

Построение описаний в онтологии (модели) предполагает наличие достаточно выразительных средств описания (модель, язык и форма представления описания изображений). Модель организуется в виде прикладной онтологии со структурами, каждая из которых соответствует обобщенному описанию некоторого класса объектов на изображении с его структурной организацией и взаимосвязями с другими классами. В качестве языка и формы представления описаний изображений в онтологии выбран семантический гиперграф ([Хахалин, 2009], [Баранович, 2011]), вершины которого соответствуют элементам описываемого изображения, а дуги – отношениям между этими элементами.

На входном изображении в общем случае задается не один образ, а целая ситуация 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, паровозики и т.п.).

3. Гибридная схема анализа изображений

Цель разработки системы для решения задачи полного анализа изображений – совместить достаточно эффективный процесс распознавания с помощью программ, работающих непосредственно с изображениями, но дающих результат на «примитивном» концептуальном уровне (имя класса + параметры экземпляра) при использовании дискриминантных, статистических методов, и достаточно эффективный с языковой точки зрения процесс концептуального анализа изображений, используя методы, основанные на знаниях. Для этого мы выбрали двухуровневую систему: нижний уровень, который непосредственно работает с изображением, и верхний уровень, который работает с моделью изображения, но с мощным языком символического описания среды. Они работают последовательно с возможностью управления процессом распознавания на нижнем уровне с концептуального уровня (см. ниже).

3.1. Система статистического распознавания (PR-система)

Примеров систем статистического распознавания много (CSoft [Raster Arts, 2008], IcadGT [IcadGT, 2007], MyScript Notes [MyScript Notes, 2007]). Для достаточно разнообразных и сложных сред (таблицы, контурные карты, схемы, архитектурные планы, печатные платы, чертежи по механике, аэрофотоснимки и т.д.) можно использовать первые две системы. Но специфические средства представления информации в этих системах, рассчитанные на конкретные области приложения, не позволяют достаточно простое использование в комплексе с концептуальной системой распознавания.

В качестве опоры для иллюстраций в выбранной предметной области мы используем реальную систему MyScript Notes. 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 (заранее определенных), 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 классов объектов ограничивается перечнем планиметрических фигур: *отрезок прямой, треугольник, прямоугольник, ромб, дуга, окружность, эллипс* и др. + класс неизвестных объектов (рисунки произвольной, нераспознаваемой формы - *drawing*).

Для входного изображения прямоугольника и окружности (см. на рисунке 2, слева) в выходной информации представлены имена классов *rectangle* и *circle*, вероятность распознавания, значения соответствующих координат и возможные дополнительные кандидаты на распознавание (*parallelogram, drawing, circle* и др.). Справа на рисунке 2 дано синтезированное программой MyScript Notes изображение по результатам распознавания.

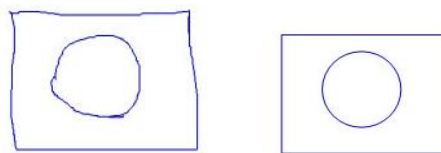


Рисунок 2 – Пример входного и синтезированного изображения

Пример более сложного изображения для гибридной системы распознавания приведен на рисунке 3. На нем представлены как объекты, распознаваемые программой MyScript Notes (*прямоугольник, ромб* и др.), так и не распознаваемые на этом уровне классы объектов (*домик, кораблик, машинка*).

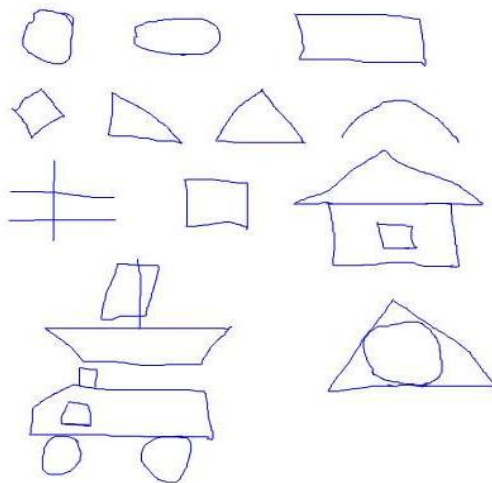


Рисунок 3 – Представители объектов для распознавания

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

3.2. Система символьного анализа изображений (КВ-система)

В качестве основы для системы анализа верхнего уровня выбран метод контекстуального целенаправленного анализа изображений, который подробно описан в [Курбатов и др., 2010] и [Khakhalin et al., 2012]. Подчеркнем только главные дополнения для этого уровня.

1. Анализ осуществляется не в поле реального изображения, а по модели изображения, полученной по результатам числового (статистического) распознавания. Первоначально анализ осуществляется по распознанным PR-системой классам объектов. Если в модели изображения присутствуют только «простые» объекты (из множества распознаваемых PR-системой), то функция КВ-системы заключается в означивании структур этих объектов из онтологии и в «наведении» геометрических отношений между ними, которые в неявном виде присутствуют в выходных данных PR-системы. Например, для рисунка 2 выходом КВ-системы будет описание «простой» ситуации в виде означенных структур «Прямоугольник» и «Окружность» с отношением «находится внутри».

Если же в модели изображения присутствуют не только те объекты, которые PR-система может распознавать, но и те, которые распознать не может, то КВ-система осуществляет поиск в онтологии структур «сложных» объектов и сопоставление с фрагментами модели изображения. Например, для объекта «кораблик», результат распознавания которого дан на рисунке 4, КВ-система, используя информацию о дополнительных кандидатах распознавания вместо нераспознаваемых (кандидаты подчеркнуты), выделит образ «трапеция». И далее, поскольку параллелограмм-«парус» и перпендикулярный отрезок-«мачта» распознаны, после процедуры «наведения» отношений сопоставит этому изображению структуру в онтологии с именем *кораблик*.

```
parallelogram 0.841063 453.52 495.446 672.868 491.344 765.673
137.967 546.324 142.07
drawing 0.5
triangle 0.31167 422.79 725.928 831.228 140.91 533.597 139.68
ellipse 0.290373 610.072 309.389 211.211 122.372 -1.14976
arc 0.290167 610.072 309.389 211.211 122.372 -1.14976 0 6.28319
line 0.731522 278.853 890.619 881.88 876.762
drawing 0.5
arc 0.100776 561.153 888.015 253.049 6.93823 -0.0208331
2.56482 6.28319
drawing 0.5
line 0.436261 192.309 655.127 287.242 894.925
arc 0.201896 232.866 775.533 118.11 9.32409 1.19305 3.09251
6.28319
line 0.834599 186.09 667.422 1004.87 650.702
drawing 0.5
arc 0.125702 585.368 655.953 392.703 6.51846 -0.00942964 -
1.11674 3.12932
drawing 0.5
line 0.423253 870.282 895.919 1011.53 654.31
arc 0.162266 936.106 773.071 125.247 7.98859 -1.02929 -0.589049
3.15386
line 0.957515 652.818 71.9868 653.825 669
drawing 0.5
```

Рисунок 4 – Результат распознавания «кораблика»

2. В случае частичного сопоставления фрагмента модели изображения со структурой онтологии выдвигается гипотеза о дополнительном анализе на уровне PR-системы и на этот уровень КВ-система передает управляющую информацию для числового распознавания. Проиллюстрируем это на примере распознавания *домика*, представленного на рисунке 3. Если по фрагменту результатов числового распознавания синтезировать с помощью программы MyScript Notes изображение, то получим фигуру, представленную на рисунке 5.

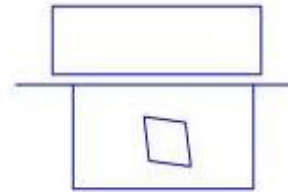


Рисунок 5 – Синтезированное изображение по результатам распознавания PR-системы

Совершенно очевидно, что PR-системой была допущена ошибка: она распознала *крышу домика* не как *треугольник*, а как *прямоугольник*. Ошибка распознавания определяется, по-видимому, «наводками» при распознавании сложного изображения, которые затрудняют процесс правильной априорной сегментации. Доказательством тому служит эксперимент: если на вход программы подать изображение только *домика*, то MyScript Notes выдаст правильные результаты (*прямоугольник*, *параллелограмм*, *треугольник* с соответствующими параметрами).

В этом случае КВ-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.

Следует отметить, что здесь проявляется более оптимальное решение задачи сегментации, когда сегментация осуществляется на основе частичных результатов распознавания нижнего уровня. В дальнейшем расширение этой стратегии можно использовать более «плотно», т.е. реализовать стратегию «сверху-вниз» (от онтологии к изображению) и при «стандартном» распознавании.

Если в результате повторного распознавания будет выделен объект класса *треугольник*, то в выходных данных КВ-системы для этого объекта

будет представлена онтологическая структура, как показано на рисунке 6 (означенные параметры выделены сплошными кружками).

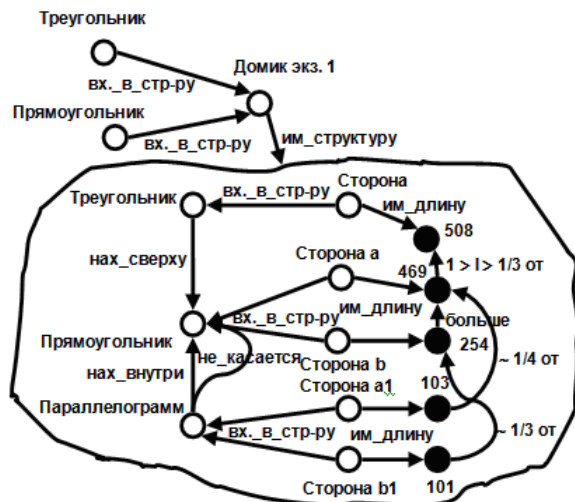


Рисунок 6 – Означенная онтологическая структура объекта «домик»

3. Трансформация описания структуры образа для передачи описания в терминах «непроизводных объектов» на уровень числового распознавания.

В такой гибридной системе необходимо иметь средства трансформации из одного (символьного) представления в другое (числовое) представление. При передаче информации от KB-системы к PR-системе необходимо из онтологических описаний структур объектов уметь вычленять существенные признаки, образовывать из них кортежи и «придумывать» обучающее множество примеров и контр примеров. Контролировать этот процесс может пробное распознавание при наблюдении и сопровождении учителя. Для этого можно привлекать синтезатор изображений интегральной системы.

4. Обучение подсистем анализа изображений

Гибридная система работает в двух режимах: обучение (символьное и статистическое – разные методы) и собственно распознавание.

Для обучения PR-системы существует много методов обучения. Приведем краткое описание процедур обучения с учителем [Найденова и др., 2008].

Методы машинного обучения (Machine Learning) являются методами обучения классификации объектов, представленных описаниями в признаковых пространствах. Цель обучения есть получение необходимых и достаточных правил, с помощью которых можно произвести классификацию новых объектов, сходных с теми, которые составляли обучающую выборку (обучение с учителем – supervised learning). При этом каждый обучающий пример (описание объекта) имеет метку, показывающую, к какому классу он принадлежит.

Ключевыми моментами машинного обучения являются:

1. Выбор и формирование признакового пространства;
2. Проверка гипотез о различимости/схождении объектов и классов объектов; задание бинарных операций сходства-различия объектов; задание мер сходства-различия для классов объектов;
3. Формирование обучающей выборки;
4. Формирование контрольной выборки;
5. Адекватный выбор алгоритма обучения.

Если выбор признакового пространства определяет задачу обучения, главным образом, содержательно, то формирование обучающей и контрольной выборок отвечает за точность, быстроту и эффективность обучения. С помощью правильно выбранных примеров можно направлять процесс обучения. Пошаговые процедуры обучения и выбор последовательности примеров (от простого к сложному) позволяет также минимизировать число примеров, необходимых для обучения. Контрольная выборка необходима не только для проверки правильности работы классификатора, но и для целенаправленного «доучивания» классификатора, его исправления, модификации, придания ему требуемых свойств.

К широко известным методам относятся вероятностные методы машинного обучения.

Метод Байеса определяет вероятность наступления события в условиях, когда наблюдается лишь некоторая частичная информация о событиях, например, по наблюдаемым признакам определяется принадлежность некоторого объекта к одному из заданных классов.

Метод опорных векторов (Support Vector Machine) относится к линейным разделяющим методам. Каждый вектор признаков (объект) представляется точкой в многомерном пространстве признаков. При заданной классификации объектов строятся две параллельные гиперплоскости (границы), разделяющие объекты разных классов, таким образом, что расстояние между этими гиперплоскостями максимизируется. Примеры, расположенные вдоль гиперплоскостей, называют опорными векторами. Значения признаков – вещественные числа.

К вероятностным методам относится скрытая Марковская модель (Hidden Markov chains). Это статическая модель, имитирующая некоторый последовательный процесс, в котором на наблюдаемые переменные оказывают влияния скрытые состояния. Переход из одного состояния в другое происходит с некоторой вероятностью. По последовательности наблюдений можно получить информацию о последовательности состояний.

Для обучения KB-системы можно применить логический метод и метод концептуального обучения [Naidenova, 2009].

К логическим методам относятся методы, при

которых в процессе обучения строятся логические правила в форме продукций или в форме решающих деревьев, в узлах которых проверяются значения отобранных при обучении признаков и принимается решение о разбиении объектов на подклассы. Логические методы работают на описаниях, которые имеют как символьную, так и целочисленную природу, или в пространстве булевых признаков. С этой точки зрения эти методы можно отнести к методам концептуальных классификаций или методам построения правил рассуждений на уровне концептов. Но формирование правил в практических задачах часто влечет за собой громадную ручную работу и включает в себе творческую и неавтоматизированную компоненту методов машинного обучения.

К методам концептуального обучения (извлечение концептов и их иерархических отношений) относится направление машинного обучения, называемое формальным концептуальным анализом. Концептуальное обучение представляет собой особый класс методов, основанных на порождении и использовании концептуальных знаний, элементами которых являются объекты, атрибуты (значения атрибутов), классификации (разбиения объектов на классы) и связи между ними. Эти связи выражаются через импликативные отношения вида «объект \leftrightarrow класс», «объект \leftrightarrow значение атрибута (свойство)», «значение атрибута (ов) \leftrightarrow класс», «подкласс \leftrightarrow класс».

Индуктивный вывод концептуальных знаний сводится к задаче вывода хороших диагностических тестов для заданных классификаций объектов. Хороший диагностический тест определяется как совокупность атрибутов (значений атрибутов), которая порождает наилучшую аппроксимацию заданной классификации (разбиения) некоторого заданного множества объектов. К поиску хороших тестов сводятся почти все хорошо известные логические методы машинного обучения. Эти методы охватывают вывод импликативных и функциональных зависимостей из данных, вывод логических правил (правила «если-то», «грубые» множества), вывод ассоциативных правил, конструирование решающих деревьев, извлечение иерархических концептуальных классификаций и ряд других. Все перечисленные задачи отличаются по форме представления выделяемых отношений включения на множестве всех подмножеств объектов и решаются с использованием одной и той же структуры данных и одних и тех же алгоритмов.

Подход *lazy learning* основывается на гипотезе, что решение когнитивных задач (обучение языку в частности) базируется на построении выводов на основе аналогий, а не на основе абстрактных правил, полученных из экспериментов. Этот подход используется в различных дисциплинах искусственного интеллекта и лежит в основе таких методов, как выводы на основе сходства, выводы на

основе примеров, выводы на основе аналогии, вывод на основе прецедентов (*case-based reasoning*) и пр.

При "ленивом" обучении обучающие примеры добавляются в память без обобщений и реструктурирования. Сходство нового примера с остальными вычисляется по метрике сходства, и правило большинства сходных примеров используется как базовое для предсказания категории нового примера.

Основными методами подхода *greedy learning* являются обучение на основе деревьев решений и индуктивного вывода. Обучение на основе деревьев решений основывается на предположении, что сходство примеров может быть использовано для автоматического построения деревьев решений, на основе которых порождаются обобщения и объяснения. Целью индуктивного вывода является построение ограниченного множества интерпретируемых правил на основе обучающих примеров или деревьев решений.

Обучение KB-системы осуществляется с помощью: 1) разработчика онтологии; 2) пользователя посредством ЕЯ-интерфейса и при посредстве системы синтеза изображений. Разработчик вводит новые описания в онтологию непосредственно на языке представления знаний (язык семантического гиперграфа). Пользователь может вводить новые описания посредством текстов на ЕЯ. Но для проверки корректности введенных описаний и в том и другом случае используется система онтологического синтеза изображений, по результатам работы которой разработчик и пользователь корректируют свои описания.

Для иллюстрации этого процесса на рисунке 7 приведено множество синтезируемых изображений, по которым видно, что в онтологической структуре объекта «домик» (в отличие от рисунка 6) не заданы некоторые ограничения по параметрам подобъектов (например, соотношение длин сторон «сруба домика» и «крыши»).

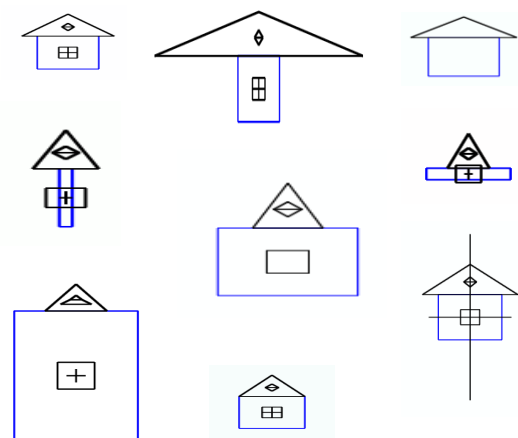


Рисунок 7 - Множество синтезируемых изображений

ЗАКЛЮЧЕНИЕ

Подчеркнем особенности гибридной системы анализа (двумерных) изображений:

- Интеграция числового (статистического) PR-подхода с KB-подходом, и использование преимуществ и того и другого.
- Новые + старые методы обучения (индуктивные и дедуктивные методы обучения). Машинные методы обучения концептуальным знаниям представляют собой модель правдоподобных индуктивных и дедуктивных рассуждений, в которых вывод знаний и их использование неотделимы друг от друга.
- Процесс анализа изображений на PR-уровне может управляться KB-системой.
- Трансформация структурного описания класса образов (объектов) в атрибутивное описание и передача этой информации с верхнего на нижний уровень распознавания.
- Для верификации процессов обучения и самого распознавания привлечение системы концептуального синтеза изображений.
- Автоматизация процесса пополнения онтологии с использованием процедуры верификации (синтеза изображений): а) за счет введения структурных описаний разработчиком непосредственно на языке семантических гиперграфов; б) за счет введения описаний пользователем на проблемно ориентированном ЕЯ.
- Интеграция анализа/синтеза изображений с ЕЯ-анализатором (лингвистическим транслятором).
- Онтологические описания ситуаций, получаемые от изображений или с текстом на ЕЯ, представляются на общем языке представления знаний – семантическом гиперграфе.
- Достаточно мощный и гибкий язык представления знаний – семантический гиперграф, как развитие семантических сетей.
- Процедура сегментации изображения на фрагменты, соответствующие распознаваемым объектам, в отличие от априорной 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.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [Баранович, 2011] Баранович, А.Е. Семиотико-хроматические гипертопосети: унифицированная модель представления знаний // OSTIS-2011. Труды конференции. С. 71-86
- [Горелик и др., 2004] Горелик, А.Л., Скрипкин, В. А. Методы распознавания. Учебное пособие. – М.: Высшая Школа, 2004. - 261 с.
- [Животников и др., 2005] Животников, Г.С., Пытьев, Ю.П., Фаломкин, И.И. Итеративный морфологический алгоритм локализации неизвестного объекта на изображении в реальном времени // ММРО-2005. Труды конференции. С. 322-325
- [Журавлёв, 2000] Журавлёв, Ю.И. Распознавание образов и анализ изображений / Ю.И. Журавлёв, И.Б. Гуревич // Искусственный интеллект: Модели и методы. Т. 1. – М.: Радио и связь, 2000. – 310 с.

[Катыс, 1990] Катыс, Г.П. Обработка визуальной информации. – М.: Машиностроение, 1990. – 320 с.

[Курбатов и др., 2010] Курбатов, С.С., Найденова, К.А., Хахалин, Г.К. О схеме взаимодействия в комплексе «анализ и синтез естественного языка и изображений» // КИИ-2010. Труды конференции. С. 234-242

[Курбатов и др., 2011] Курбатов, С.С., Найденова, К.А., Хахалин, Г.К. Интегрирование интеллектуальных систем анализа/синтеза изображений и текста: контуры проекта INTEGRO // OSTIS-2011. Труды конференции. С. 213-232

[Найденова и др., 2008] Найденова, К.А., Невзорова, О.А. Машинное обучение в задачах обработки естественного языка: обзор современного состояния исследований // Новости Казанского университета, 2008, № 1, С. 3-24

[Николаев и др., 2003] Николаев, А.Б., Фоминых, И.Б. Нейросетевые методы анализа и обработки данных. Учебное пособие. – М.: МАДИ (ТГУ), 2003, 95 с.

[Новикова, 2008] Новикова, Н.М. Структурное распознавание образов. Учебно-методическое пособие для вузов. – Воронеж: Издательско-полиграфический центр Воронежского ГУ, 2008. – 29 с.

[Павлидис, 1986] Павлидис, Т. Алгоритмы машинной графики и обработки изображений / Т. Павлидис; пер. с англ. Н.В. Петрова; под ред. М.С. Гуревича. – М.: Радио и связь, 1986. – 400 с.

[Пытьев, 2007] Пытьев, Ю.П. Математические методы и адаптивные алгоритмы эмпирического построения теоретико-возможностной модели стохастического объекта // ММРО-13. Труды конференции. С.54–56.

[Хахалин, 2009] Хахалин, Г.К. Прикладная онтология на языке гиперграфов // ЗОНТ-09. Труды конференции. С. 223-231

[Bhanu et al., 2005] Bhanu Bir, Peng Jing, Huang Thomas and Draper Bruce A.. Introduction to the Special Issue on Learning in Computer Vision and Pattern Recognition // IEEE Transactions on Systems, Man and Cybernetics—Partb: Cybernetics, vol. 35, no. 3, June 2005, pp. 391-395

[IcadGT, 2007] Режим доступа: <http://www.recsoft.ru>. – [Электронный ресурс].

[Khakhalin et al., 2012] Khakhalin, G., Kurbatov, S., Naidenova, K., and 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 (в печати).

[MyScript Notes, 2007] Режим доступа: <http://www.visionobjects.com/en/webstore/myscript-studio/description/>. – [Электронный ресурс].

[Naidenova, 2009] Naidenova, X. A. Machine Learning Methods for Commonsense Reasoning Processes: Interactive Models. N.Y., USA: IGI Global. 2009.

[Pellegratti et al., 1994] Pellegratti Paolo, Roli Fabio, Serpico Sebastian B., and Vemazza Gianni. Supervised Learning of Descriptions for Image Recognition Purposes // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 1. January 1994, pp. 92- 98

[Raster Arts, 2008] Режим доступа: <http://www.csoft.ru>. – [Электронный ресурс].

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

THE HYBRID CIRCUIT OF THE IMAGE ANALYSIS

S.S. Kurbatov, A.P. Lobzin, K.A. Naidenova, G.K. Khakhalin

An image analysis scheme connecting the advantages of numerical and symbolic methods of pattern recognition is examined. The crossbreeding of two approaches with the use of applied ontology in an integrated system of linguistic translator and synthesizer of images makes it possible to achieve the higher level of understanding images. Semantic hypergraph is used as the language for knowledge representation in the framework of symbolic methods of image processing.