



OSTIS-2014

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514; 004.932

РАЗБОР ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ ОНТОЛОГИЧЕСКИМ КОМПОНЕНТОМ ГИБРИДНОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ

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

Научно-Исследовательский Центр Электронной Вычислительной Техники, Москва, Россия

gkhakhalin@yandex.ru

curbatow.serg@yandex.ru

lobzin@rambler.ru

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

Ключевые слова: гибридная система; концептуальный разбор изображения; предметная онтология; семантический гиперграф.

Введение

Проблематика машинного зрения столь привлекательна для современных исследователей по той причине, что аппаратные возможности, предоставляемые в данной области последними достижениями электроники и вычислительной техники, достигли такого уровня, что они уже во многом приближаются к «техническим характеристикам» глаза человека [Визильтер и др., 2007]. Но главная задача в машинном зрении – задача «понимания» изображений – сегодня еще далека от своего окончательного решения.

При решении задачи понимания изображений необходимо сместить цели распознавания от только наименования объектов к их структурно-логическим описаниям. Это позволит не просто именовать знакомые объекты, но также сообщать неизвестные аспекты известных объектов; говорить что-либо о неизвестных объектах; и обучать распознавать новые объекты на небольшом количестве зрительных примеров или вообще при их отсутствии [Farhadi et al., 2009].

Для процессов такой обработки изображений даже введен термин «разбор изображений» (image parsing). В настоящее время большинство систем разбора изображений [Endres et al., 2013; Shrivastava et al., 2011] реализуется с помощью одной системы распознавания статистического типа, в которой используются кроме дискриминантных признаков семантические свойства объектов. Последние

характеризуют не только различия одного класса объектов от другого, но и их сходство.

В области машинного зрения принято выделять следующие основные этапы обработки данных: предобработка изображений; сегментация; выделение геометрической структуры; определение относительной структуры и семантики. Связанные с этими этапами уровни обработки обычно называются соответственно: обработка нижнего уровня, среднего уровня, высокого уровня. В то время как алгоритмы обработки нижнего уровня (кодирование, сглаживание, фильтрация) могут рассматриваться как хорошо проработанные и детально изученные, алгоритмы среднего уровня (сегментация) продолжают сегодня оставаться центральным полем приложения инженерных и исследовательских усилий. Методы обработки высокого уровня, относящиеся собственно к «пониманию (разбору) изображений», находятся еще в начальной фазе развития и по-прежнему представляют собой «вызов» для сообщества исследователей в области компьютерного зрения и искусственного интеллекта [Визильтер и др., 2007].

Исходя из различной проработанности этих этапов, оптимальной нагруженности и невозможности на пиксельном уровне в общем случае решить задачу априорной сегментации, в данной работе предлагается гибридная система понимания (разбора) изображения, состоящая из двух компонентов: PR-система (Pattern Recognition) для распознавания «непроизводных» элементов при использовании эффективных дискриминантных (числовых) методов, и KB-система (Knowledge

Based) для распознавания композиционных объектов в целом, обладающих существенной структурной информацией, где ведущая роль принадлежит структурным (символьным) методам. Типология методов распознавания и обоснованность гибридного подхода представлены в [Курбатов и др., 2012b; Хахалин и др., 2013].

Под задачей понимания (разбора) изображения подразумевается описание распознаваемых объектов и их расположение на изображении в терминах и в структурах языка прикладной онтологии. Основным результатом разбора является описание зрительной ситуации, состоящей из экземпляров распознанных классов объектов (с означенными характеристиками) и отношений между ними.

1. О гибридной системе анализа изображений

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

PR-система непосредственно работает с реальным изображением. KB-система работает с моделью изображения (по результатам работы PR-системы), но с мощным языком символьного (концептуального) описания среды. Они работают последовательно с возможностью управления процессом распознавания на нижнем уровне с концептуального уровня, реализуя т.н. «эксклюзивный» анализ (см. ниже). Функциональная схема двухуровневой системы анализа изображений представлена на рис. 1.

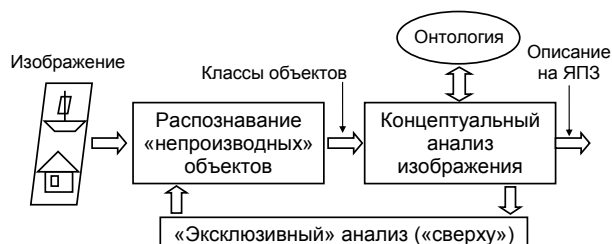


Рисунок 1 – Схема двухуровневой системы разбора изображений

Данная гибридная система разбора 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 [Курбатов и др., 2012b]. Взаимодействие этих систем осуществляется на основе общей прикладной онтологии [Khakhalin et al., 2012].

Для иллюстраций в выбранной предметной области мы используем реальную PR-систему MyScript Notes [MyScript Notes, 2007]. Эта система решает стандартные задачи предварительной обработки изображений, включая кодирование, сглаживание, фильтрацию, сегментацию и т.д., избавляя от этих процедур KB-систему. На вход программы MyScript Notes подаются нарисованные от руки геометрические рисунки. На выходе получаем для каждого класса распознаваемых объектов: имя класса, вероятность распознавания и значения геометрических параметров, привязанных к полю изображения (координатной сетке). Кроме того дается альтернативная информация распознавания, когда приводятся возможные другие кандидаты на распознавание данного объекта с соответствующей вероятностью.

Результатом распознавания будет список (его фрагмент дан на рис. 2), где представлены имена предполагаемых классов объектов с указанием вероятности распознавания и конкретные характеристики экземпляров объектов, привязанные к системе координат изображения (подчеркнутые 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

```

Рисунок 2 – Фрагмент результатов распознавания PR-системой

Результаты распознавания композиции с привязкой к координатной сетке приведены на рисунке 3. На нем представлены распознанные объекты (некоторые *прямоугольники, треугольники* и др.), неправильно распознанные объекты (другие *прямоугольники, эллипсы* и др.), и не распознаваемые на этом уровне классы объектов (*домик, кораблик, паровозик, зайчик* и др.).

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

2. KB-система

На вход KB-системы подается результат распознавания «непроизводных» объектов PR-системой (имена объектов с конкретными характеристиками на координатной сетке).

Основным результатом работы KB-системы является описание на языке представления знаний, состоящее из экземпляров распознанных классов объектов и отношений между этими экземплярами.

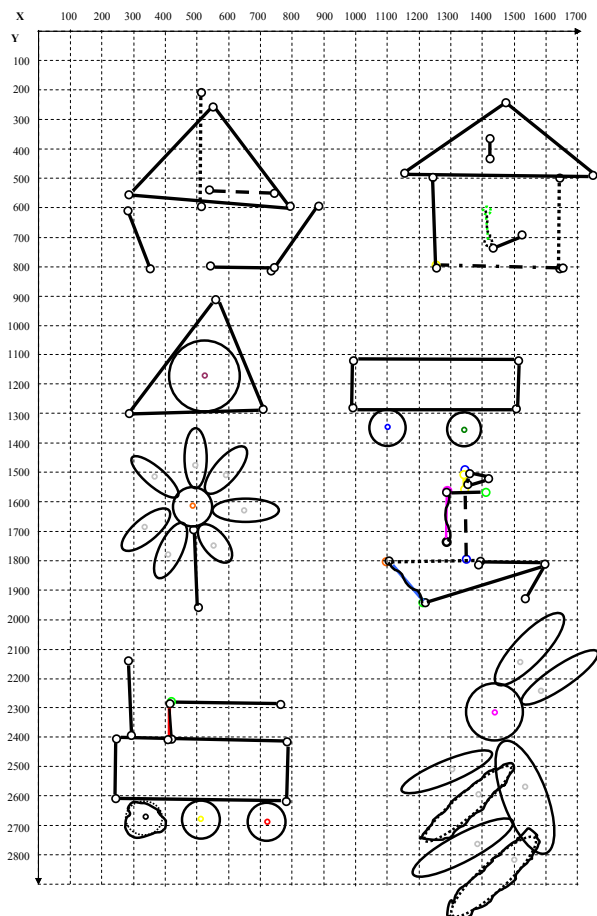


Рисунок 2 – Результаты распознавания композиции фигур в координатной сетке

Поскольку система работает в комплексе, результат может быть также представлен в виде:

- перечня имен всех «понятых» (распознанных) объектов композиции;
- текста на естественном языке (какие объекты и в каких отношениях друг с другом присутствуют на изображении);
- изображения, полученного с помощью системы концептуального синтеза изображений с теми же самыми характеристиками объектов и/или «аналогичного» с характеристиками тех же классов, но со значениями, не выходящими за пределы этих классов.

2.1. Структура прикладной онтологии

Прикладная онтология представлена в виде семантического гиперграфа со структурами, каждая из которых соответствует обобщенному описанию некоторого класса объектов, и взаимосвязями с другими классами объектов. Фрагмент прикладной онтологии в виде семантического гиперграфа представлен на рис. 4., где большие кружки – концепты онтологии (объекты среды), тонкие стрелки – родовидовые связи концептов, маленькие кружки и тонкие стрелки – свойства объектов, замкнутые области – описание структур объектов, толстые стрелки – связь типа «часть-целое» («входит_в_структуру»). Каждая структура объекта представляет собой подгиперграф, отражающий

взаимоотношения его частей. Вершины и дуги (ребра) онтологии на языке семантического гиперграфа помечаются цепочками символов некоторого алфавита (в данном случае, словами русского языка) для различения вершин (концептов) и связей между ними (отношений). На рисунке дана только часть разметок.

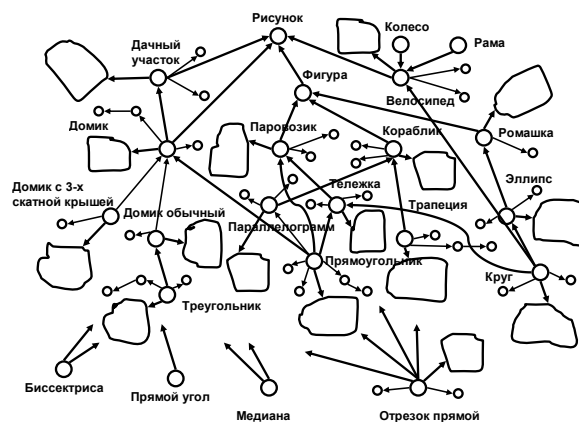


Рисунок 4 – Фрагмент прикладной онтологии

В прикладной онтологии представлены 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.).

2.2. Функции концептуального компонента

КВ-система выполняет следующие функции по концептуальному разбору изображения:

- представление результатов работы PR-системы в виде концептов «непроизводных» объектов с означенными характеристиками (*точки*, *отрезки прямых*, *прямоугольники*, *окружности*, *дуги* и т.д.) – получение «сырого» графа изображения с привязкой к координатной сетке;
- «вычисление» путей в онтологии для формирования списка гипотез на сопоставление (по связям «вх_в_структуру», «род-вид»), т.е. формирование ранжированного списка гипотез;
- сопоставление «сырого» графа изображения с графами структур комплексных объектов прикладной онтологии (поиск полного либо частичного изоморфизма);
- вычисление параметров «ареала» (области) по структурному описанию объекта из онтологии для проведения поиска на изображении гипотетического непроизводного объекта с помощью PR-системы;
- подтверждение результатов сопоставления, осуществляемое на реальном изображении при посредстве «эксклюзивного» анализа.

2.3. Онтологический разбор композиции

КВ-система первоначально преобразует

результаты, полученные PR-системой, в граф, вершины которого помечены именами классов производных объектов. То есть задается модель изображения в виде графа производных понятий со своими атрибутами, которые может распознать PR-система (например, *отрезок прямой* с координатами концевых точек и длиной отрезка; *окружность* с координатами центра окружности и радиусом).

По этому графу составляется ранжированный список производных понятий $L_{v_1}(v_1, v_2, \dots, v_k)$, где на первых местах стоят более «крупные» понятия (*треугольники, эллипсисы, квадраты* и т.д.), а на последних – *точки, отрезки, дуги* и *произвольные кривые (drawing)*. Ранг определяется иерархией прикладной онтологии.

Для выбранного понятия v_i с наибольшим рангом по онтологии формируется список сложных, комплексных понятий $L_{s1}(h_1, h_2, \dots, h_n)$, в структуру которых входит выбранное понятие (используются онтологическое отношение «вх_в_структуру»). Выбираем из этого списка гипотезу h_1 . В онтологии с этим именем выделяем структуру онтологического понятия и частично означиваем ее по атрибутам производного понятия v_1 . По значениям этих атрибутов с учетом числовых отношений в структуре вычисляем ареал (область) на координатной сетке, в котором могут присутствовать другие производные объекты данной онтологической структуры.

По списку L_{v_1} составляем список производных понятий, которые попадают в этот ареал. Тем самым из модельного графа изображения выделяем некоторый подграф. И этот подграф используется для сопоставления со структурой гипотезы h_1 (граф структуры онтологического понятия).

Если граф структуры онтологического понятия полностью сопоставлен хотя бы с частью подграфа модели изображения, то комплексный объект считается идентифицированным. И тогда: 1) формируется означенный экземпляр онтологического понятия; 2) из L_{v_1} исключаются все понятия, которые вошли в структуру h_1 ; 3) осуществляется переход к следующему производному понятию в модифицированном списке L_{v_1} ; и 4) формируется список $L_{v_2}(h_1)$ – список подтвержденных гипотез для следующей итерации (для более сложных объектов и ситуаций).

Если граф структуры онтологического понятия не полностью сопоставлен с подграфом модели изображения, то переходим на эксклюзивный анализ реального изображения, при этом: 1) уточняется ареал по распознанным составляющим (т.е. определяется «где надо искать»); 2) по структуре гипотезы h_1 задается список производных объектов, которые необходимо искать в этом ареале (т.е. определяется «что искать»).

Далее работает PR-система в пределах только вычисленной области на реальном изображении.

После работы PR-системы получаем модифицированные результаты распознавания выделенного ареала изображения. По этим результатам модифицируется подграф модельного графа и осуществляется сопоставление структуры гипотезы h_1 с этим подграфом.

Если после эксклюзивного анализа сопоставление найдено, то считаем, что комплексный объект распознан. В противном случае, переходим к следующей гипотезе h_2 .

Далее в цикле обрабатываются оставшиеся понятия из множества $\{v_i\}$.

Если в процессе всего концептуального анализа ничего не смогли распознать, то получаем пустой список $L_{v_2}()$.

Если частично идентифицированы комплексные рисунки (это общий случай), то получаем список $L_{v_2}(h_1, h_2, \dots, h_m, v_1, v_2, \dots, v_p)$, где подсписок (h_1, h_2, \dots, h_m) – комплексные рисунки, а подсписок (v_1, v_2, \dots, v_p) – «остатки» в виде производных объектов. Для более сложных композиций (если есть описания в прикладной онтологии) этот список берется в качестве начального для следующей итерации.

Если на изображении только комплексные рисунки и KB-система их идентифицировала, получаем список $L_{v_2}(h_1, h_2, \dots, h_m)$. Это говорит о том, что все сложные объекты распознаны и кроме них на изображении ничего нет (это крайне редкий случай).

Вообще целью концептуального разбора изображения является получение списка $L_{v_2}(h_1, h_2, \dots, h_m, v_1, v_2, \dots, v_p)$ с минимизацией мощности множества $\{v_1, v_2, \dots, v_p\}$. По данному списку формируется основной результат разбора изображения, а уже на его основе – другие формы представления результата: текст на естественном языке и/или синтезированное изображение композиции.

Проиллюстрируем процесс эксклюзивного анализа на следующем примере. Предположим, что на композиционном изображении кроме всего прочего присутствует объект *Домик* (рис 5а).

PR-система распознала *треугольную крышу* – *треугольник*, а все остальные элементы (*избу, окошко* и *слуховое окно*) не распознала (рис 5б). Тогда KB-система, выйдя на структуру понятия *домик* и найдя частичное сопоставление (часть вершин этой структуры будет означено), выдвинула гипотезу о том, что в определенном месте изображения (вычисляется ареал по местоположению и размерам *треугольника*) могут присутствовать недостающие для *домика* элементы (формируется список требуемых элементов) с определенными ограничениями по размерам (вычисляются на основе размеров *треугольника*). Подав эту информацию на вход PR-системы, происходит повторное распознавание в только

определенной части реального изображения, что повышает робастность всей системы.

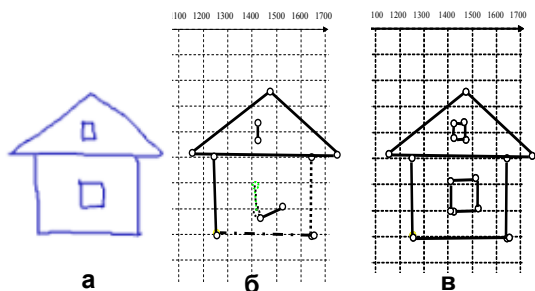


Рисунок 5 – Иллюстрация эксклюзивного анализа

Процесс сегментации при таком методе обработки изображения носит более адекватный характер, т.е. сегментация осуществляется на основе частичных результатов распознавания, а не априори – только с учетом пиксельных характеристик изображения.

Реальные модельные эксперименты показывают более надежные результаты даже на «нижнем» уровне распознавания (на уровне PR-системы), когда сегменты изображения «под конструкции» выделяются более осмысленно. Результат, полученный PR-системой по схеме эксклюзивного распознавания, показан на рис. 5в.

3. Использование концептуального синтеза

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

Первый случай подробно рассмотрен в [Хахалин и др., 2013]. Здесь рассмотрим кратко процесс модификации и пополнения онтологии.

Естественно, что КВ-система «видит только то, что знает», т.е. распознает комплексные объекты, информация о которых присутствует в онтологии. Поэтому важными вопросами являются задачи пополнения и коррекции прикладной онтологии.

В интегральной системе модификация и расширение онтологии предполагает использование двух каналов: для инженера по знаниям (для разработчика) – непосредственный ввод структур новых концептов и/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 это дедуктивное обучение не приводило к «хаосу», необходимо применить методику наполнения онтологии с верификацией вводимых описаний с помощью системы концептуального синтеза изображений [Курбатов и др., 2012а; Литвинович, 2013].

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 синтезированных изображений не будет корректным.

Проиллюстрируем это на примере концепта *Тачка*. До ввода структуры объекта *Тачка* в онтологию гибридная система распознавания представит его описание в виде взаимосвязанного набора *равнобоковой трапеции*, *окружности* и *отрезка прямой*. Введя в онтологию описание структуры понятия *Тачка*, как это представлено на рис. 7, и получив множество синтезированных изображений по этой структуре, представленных на рис. 8, можно предположить, что теперь гибридная система будет распознавать изображение как целостный объект планиметрической среды.

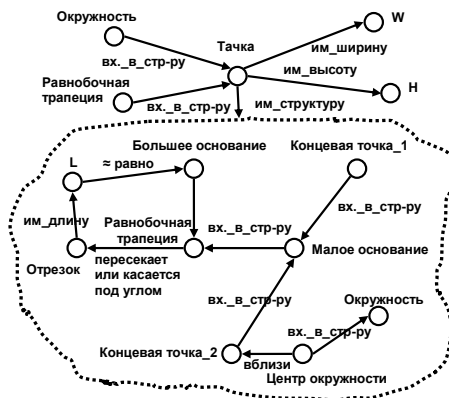


Рисунок 7 – Онтологическая структура концепта *Тачка*

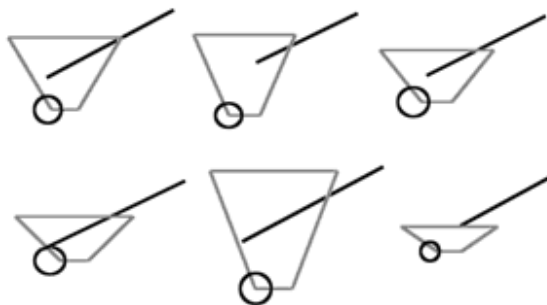


Рисунок 8 – Множество синтезированных изображений концепта *Тачка*

Заключение

Модельные эксперименты показывают, что наиболее оптимальными для решения задачи понимания изображений будут системы, использующие преимущества статистических и структурных методов анализа.

Если учитывать погружение гибридной системы в интегральную систему, то обучение онтологического компонента может происходить без визуальных примеров по описаниям классов на

естественном языке. Для верификации такого процесса обучения привлекается концептуальный синтезатор изображений.

Будущие исследования гибридной системы предполагают решение задач анализа перекрывающихся (маскирующихся) рисунков и поиска объекта на изображении по запросу пользователя.

По мере развития гибридной системы предполагается переход от контурных рисунков к изображениям реальных фрагментов среды (например, для решения задач навигации роботов в реальной среде; для распознавания лиц и т.п.).

Интересным вопросом является «передача знаний» с онтологического уровня к статистическому, позволяющая подключать параллельно несколько PR-систем, обученных на разные ограниченные множества классов.

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

Библиографический список

[Endres et al., 2013] Ian Endres, Kevin Shih, Johnston Jiaa, and Derek Hoiem. Learning Collections of Part Models for Object Recognition. CVPR, 2013.

[Farhadi et al., 2009] A. Farhadi, I. Endres, D. Hoiem, and D.A. Forsyth. Describing Objects by their Attributes. CVPR 2009.

[Khakhalin et al., 2012] 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.

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

[Shrivastava et al., 2011] Abhinav Shrivastava, Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros. Data-driven visual similarity for cross-domain image matching. ACM Trans. Graph. 30(6): 154 (2011).

[Визильтер и др., 2007] Визильтер Ю. В., Желтов С. Ю., Князь В. А., Ходарев А. Н., Моржин А. В. Обработка и анализ цифровых изображений с примерами на LabVIEW IMAQ Vision. – М.: ДМК Пресс, 2007. – 464 с.

[Курбатов и др., 2012а] 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.

[Курбатов и др., 2012б] Курбатов С.С., Лобзин А.П., Найденова К.А., Хахалин Г.К. Гибридная схема анализа изображений // Труды Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» OSTIS-2012. 16-18 февраля 2012, г. Минск, Беларусь. - 2012. – С. 327-334.

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

[Хахалин и др., 2013] Хахалин Г.К., Курбатов С.С., Лобзин А.П. Концептуальный компонент в гибридной системе анализа изображений // Труды 23-й Международной конференции по компьютерной графике и зрению, Владивосток, Россия. – Владивосток: Институт автоматизации и процессов управления ДВО РАН, 2013. – С. 222-225.

PARSING OF GRAPHIC IMAGES ON THE BASIS OF APPLIED ONTOLOGY IN HYBRID SYSTEM OF RECOGNITION

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

*Research Centre of Electronic Computing
Engineering, Moscow, Russia*

gkhakhalin@yandex.ru

curbatow.serg@yandex.ru

lobzin@rambler.ru

It is considered conceptual component of the hybrid system of an image analysis. The knowledge base for this component is the applied ontology which is presented in the semantic hypergraph language

Introduction

Now systems of image understanding it is realized by means of one system of recognition only statistical type in which semantic properties of objects are used except for discriminant attributes.

The description of distinguished objects and their arrangement is meant a problem of understanding (parsing) of the images in terms and in structures of applied ontology language. The basic result of analysis is the description of the visual situation consisting of occurrences of distinguished classes of objects and relations between them.

Main Part

For the decision of a problem of images understanding the hybrid system of the analysis of the images, consisting of a statistical and ontology components is proposed.

The statistical component directly works with the real image. The ontological component works with model of the image (by results of work of statistical system).

The algorithm of work of the analysis of a composition on the image is described and the examples illustrating process of understanding are given.

Functions of an ontological component are listed. The algorithm of work of the analysis of a composition on the image is described and the examples illustrating process of understanding are given.

The scheme of automation of process of ontology augmentation with attraction of naturally language interface and conceptual system of image synthesis is considered.

Conclusion

The future researches of hybrid system assume the decision of problems of the analysis of overlapped (masking) figures and search of object on the image by inquiry of the user.