

ДЕСКРИПЦИОННАЯ ЛОГИКА ПОИСКА ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Кучуганов Александр Валерьевич

К.т.н., доцент, доцент кафедры АСОИУ, ФГБОУ ВО "Ижевский государственный технический университет имени М. Т. Калашникова" Министерства образования и науки Российской Федерации, 426069, г. Ижевск, ул. Студенческая, 7, e-mail: Aleks_KAV@udm.ru

Аннотация. Наиболее популярными подходами к проблеме поиска объектов на изображениях считаются: лингвистический подход, в рамках которого осуществляется синтаксическое распознавание объектов строго предопределенной структуры; искусственные нейронные сети; дескрипционные алгебры изображений; подходы, описывающие объекты с помощью логики предикатов; технологии CBIR (Content-Based Image Retrieval), опирающиеся на дескрипционные логики (ДЛ).

В статье представлен онтологический подход, основанный на дескрипционной логике с расширением на пространственную область данных, представленных в виде атрибутивных графов. Процесс анализа изображения управляется стратегией, содержащей: этап предварительного морфологического анализа; этап формирования гипотезы о категории найденного объекта; этап подтверждения гипотезы путем логического вывода о классе объекта. В ходе анализа строится дерево решений о категориях объектов. После выбора гипотезы формируется дерево вывода о соответствии объекта какому-либо ДЛ-определению из этой категории. При неудовлетворительной степени сходства и отличий происходит переход к другой ветви дерева решений о категориях. Показаны примеры работы системы анализа изображений.

Ключевые слова: атрибутивный граф, кластерный анализ, дескрипционная логика, гранулярная онтология, тезаурус, дерево решений, дерево вывода, степень сходства и отличий, граф объектов изображения.

Цитирование: Кучуганов А.В. Дескрипционная логика поиска объектов на изображениях // Информационные и математические технологии в науке и управлении. 2018. № 3 (11). С. 55–62. DOI:10.25729/2413-0133-2018-3-06

Введение. На сегодняшний день наиболее распространенными можно считать следующие подходы к анализу изображений:

В рамках лингвистического подхода [3] осуществляется синтаксическое распознавание объектов строго предопределенной структуры (буквы, символы, объекты простой формы), что существенно ограничивает область применения. При этом производительность сильно зависит от сложности структур.

Искусственные нейронные сети. Их недостатками считается необходимость длительного обучения каждому классу объектов, непредсказуемость обучения, непрозрачность принимаемого решения и ресурсоёмкость [Я. Бенджио, Л. Ботоу, Ю. ЛеЧин, С. Осовский, С. Хайкин, П. Хафнер].

Дескрипционные алгебры изображений описывают изображение в терминах дескрипторов – результатов вычисления различных интегрирующих и дифференцирующих функций, т.е. распознают объекты как целостные образы, но не осуществляют структурный анализ изображенных объектов [И.Б. Гуревич, Ю.И. Журавлев, И.В. Корябкина, Ю.О. Трусова, В.В. Яшина].

Технологии CBIR (Content-Based Image Retrieval) осуществляют контентный поиск изображений по образцу на основе интегральных характеристик, гистограмм ориентации, локальных дескрипторов и их комбинаций. Выявляются контуры и их расположение. Чтобы машина вывода знала, какие изображения считать похожими, применяют обучение с учителем. Недостатком существующих решений в рамках этого подхода следует считать ограниченность набора семантических признаков. Значения признаков субъективны, поскольку принадлежность совокупностей локальных особенностей тому или иному предмету - их семантику - назначает человек, т.е. возникает разрыв между синтаксисом и семантикой изображений [Heba Aboulmagd, G.A. Bilodeau, R. Bergevin, R. Balasubramaniam, Neamat El-Gayar, R. Krishnapuram, S. Medasani, Hoda Onsi, S.-H. Jung, Y.-S. Choi].

В настоящее время стали широко применяться и развиваться различные языки логического анализа и распознавания графической информации. В работе [1] предложен язык на базе языка ПРОЛОГ Ш для описания (деформируемых) логических моделей объектов. С помощью встроенных предикатов и пространственных функций библиотеки GEOS распознаются прямая, окружность, текстура и ее модель. Найденный объект, удовлетворяющий логической модели, оценивается функцией энергии, принимающей нечеткие значения. Логическая модель объекта представляет собой набор правил - высказываний. Работа интерпретатора основана на переборе с возвратами и методе ветвей и границ для отсечения неперспективных ветвей по заданным порогам. Показано дерево логического вывода. Осуществляется поиск в глубину.

Дескрипционные логики нашли широкое применение в различных аналитических и поисковых системах. Достоинство дескрипционных логик заключается в том, что с их помощью эксперт описывает термины и определения предметной области, а машина вывода (reasoner) системы осуществляет контроль корректности определений и поиск соответствующих экземпляров в разделе данных.

Испанские ученые [6] выделяют на изображении цветовые области, формируют их описание (форма, цвет) и пространственные характеристики (ориентация, топология/положение), а потом это описание переводят в дескрипционные логики с целью получения формального и явного смысла и классификации. Для описания концептов в дескрипционной логике в качестве атомарных они берут эти характеристики областей. В качестве примера анализируют ситуации, возникающие перед роботом, ориентирующимся в пространстве коридоров и комнат. Недостатком работы следует считать ограниченный набор характеристик изображения и сложность их формирования.

Французские ученые [5] описывают предложенную ими дескрипционную логику, названную $ALC(F)$, для которой ввели специальную пространственную область в виде набора функций, вычисляющих пространственные отношения. Доказывают свойства предложенного языка пространственных рассуждений: допустимость, разрешимость, выполнимость и т.п. С помощью этого языка они описывают структуру объектов на изображении. Авторы основываются на своей работе [4], где они описали онтологию пространственных

отношений, используя математическую морфологию. Основываясь на учебниках по анатомии, они создали набор функций, интерпретируемых на медицинских изображениях.

Таким образом, можно сделать вывод о том, что дальнейшее развитие моделей и методов поиска объектов на изображениях остается актуальным.

Ниже описывается онтологический подход, основанный на дескрипционной логике (ДЛ) с расширением на пространственную область данных, представленных в виде атрибутивных графов. Процесс анализа изображения управляется стратегией, содержащей: этап предварительного морфологического анализа; этап формирования гипотезы о категории найденного объекта; этап подтверждения гипотезы путем логического вывода о классе объекта. В ходе анализа изображения строится дерево решений о категориях объектов. После выбора гипотезы формируется дерево вывода о соответствии объекта какому-либо ДЛ-определению из этой категории. При неудовлетворительной степени сходства и отличий происходит переход к другой ветви дерева решений о категориях.

1. Логика $ALC(GI)$. Синтаксис концептов логики ALC (Attributive Language with Complement) [3–5]:

$$T \mid \perp \mid A \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \exists R.C \mid \forall R.C,$$

где: T, \perp – концепты *истина* и *ложь*; A – атомарный концепт; R – атомарная роль; C, D – производные концепты; $\exists R.C, \forall R.C$ – концепты, задающие общие ограничения на значения.

Расширение логики языка ALC на конкретную область данных в виде атрибутивных графов позволяет сократить область поиска пространственных данных. Например, вместо перебора множества пикселей изображения или множества отрезков будут проверяться только цепочки, соединяющие пары вершин. В логике $ALC(GI)$ атомарными концептами на графах служат понятия *Граф*, *Связный подграф*, *Вершина*, *Дуга*. Атомарные роли: *Начало дуги*, *Конец дуги*, *Исходящая дуга*, *Входящая дуга*. Из них составляются более сложные определения, известные в теории графов.

Домен «Изображение» является расширением домена «Граф» в части атрибутов вершин и ребер. В число непродеривированных (атомарных) понятий предлагаемой логики $ALC(GI)$ входят понятия, интерпретируемые на графах, и понятия (атрибуты вершин и ребер графа), вычисляемые на данных изображения, такие как: *Площадь сегмента*, *Периметр сегмента*, *Цвет Сегмента*, *Длина*, *Ширина*, *Угол*, *Удлиненность*, *Извилистость* и т.п. [2].

Примеры терминологических аксиом предметной области "Дешифрирование аэрокосмоснимков" в логике $ALC(GI)$:

Дорога – это цветовой сегмент, который имеет серый цвет, широкополосную форму, линейность границ – "смешанная", извилистость "несильная":

$$\text{Дорога} \equiv \text{Сегмент} \sqcap \neg \exists \text{ВнутриСегмента.} T \sqcap \exists \text{ЦветСегмента} = \text{Серый} \sqcap \exists$$

$$\text{ФормаСегмента} = \text{Широкополосный} \sqcap \exists \text{ЛинейностьГраницы} = \text{Смешанная} \sqcap \exists$$

$$\text{ИзвилистостьГраницы} = \text{Несильная}.$$

Дом, Здание – это цветовой сегмент малой площади, имеющий прямолинейные границы:

$$\text{Дом} \equiv \text{Здание} \equiv \text{Сегмент} \sqcap \exists \text{ФормаСегмента} = \text{Площадной} \sqcap \exists \text{ПлощадьСегмента} =$$

$$\text{Малая} \sqcap \exists \text{ЛинейностьГраницы} = \text{Прямолинейная}.$$

Озеро – это цветовой сегмент тёмносинего цвета, имеющий форму площадного, извилистость границы несильная:

$$\text{Озеро} \equiv \text{Сегмент} \sqcap (\exists \text{ЦветСегмента} = \text{Тёмносиний} \sqcap \exists \text{ФормаСегмента} = \text{Площадной} \sqcap \exists \text{ИзвилистостьГраницы} = \text{Несильная}).$$

Залив является частью озера, начинается и заканчивается отрицательным перегибом границы, отношение высоты образованного сегмента к длине хорды сегмента больше 1:

$$\text{Залив} \equiv \text{Озеро} \sqcap (\exists \text{ЧастьСегмента} \sqcap (\exists \text{НачальнаяВершина} \sqcap \exists \text{Угол} < 0) \sqcap (\exists \text{КонечнаяВершина} \sqcap \exists \text{Угол} < 0) \sqcap \exists \text{Отношение(ВысотаСегмента, ДлинаХорды, } K1).P1)$$

где атомарные конкретные атрибуты высота и хорда сегмента связаны предикатом $P1$ с константой $K1$.

Корабль – это область изображения, которая имеет форму, близкую к прямоугольнику, ее граница содержит один *не острый* угол и окружающий фон темный:

$$\text{Корабль} \equiv \text{Область} \sqcap (\exists \text{Форма} = \text{Прямоугольник}) \sqcap (\exists \text{Угол} = \neg \text{Острый}) \sqcap (\exists \text{Фон} = \text{Темный}).$$

Самолет – это область изображения, которая имеет форму, близкую к пятиугольнику, и ее граница содержит пять острых углов или имеет форму четырехугольника и ее граница содержит три острых угла:

$$\text{Самолет} \equiv \text{Область} \sqcap ((\exists \text{Форма} = \text{Пятиугольник}) \sqcap (\exists = 5(\text{Угол} = \text{Острый})) \sqcup (\exists \text{Форма} = \text{Четырехугольник}) \sqcap (\exists = 3(\text{Угол} = \text{Острый}))).$$

Автомобиль – это область изображения, которая имеет форму, близкую к прямоугольнику, площадь меньше площади самолета или корабля и окружающий фон темный:

$$\text{Авто} \equiv \text{Область} \sqcap (\exists \text{Форма} = \text{Прямоугольник}) \sqcap ((\exists \text{Площадь} < \text{Самолет.Площадь}) \sqcup (\exists \text{Площадь} < \text{Корабль.Площадь})) \sqcap (\exists \text{Фон} = \text{Темный}).$$

2. Онтология предметной области «Изображения объектов». Предлагаемый алгоритм осуществляет грануляцию информации по формам изображений объектов с помощью аппарата гранулярных онтологий и функций интерпретации семантики графической информации.

Гранулярную онтологию некоторого множества понятий формально определим как

$$CO = \langle C, R, \Omega, \Phi, \Gamma \rangle, \text{ где:}$$

C – базовое множество понятий – объектов и их атрибутов;

R – множество отношений на C ;

Ω – множество операций на C и R ;

Φ – множество функций фазсификации значений концептов-свойств;

Γ – множество функций интерпретации геометрических (пространственных) и иных понятий на множествах C и R .

Множество пространственных гранул формируется из базового множества C :

$$C = (C_1, C_2, \dots, C_n),$$

где $C = C_1 \cup C_2 \cup \dots \cup C_n$ и

$C_i \cap C_j = \emptyset, \forall i, j = 1, \dots, n, i \neq j$ – в случае четкой гранулярной онтологии.

Очевидно, при изменении функций фазификации изменяются и кластеры.

В качестве алфавита мы используем морфологические особенности изображений: отрезок прямой линии, отрезок дуги окружности, узел (разветвление или угол), а также цвет области, ограниченной прямыми и дугами. Каждый цветовой сегмент изображения считается словом. Предложение описывает некоторую более или менее компактную совокупность объектов изображения, например, дома и другие строения, автомобили, самолеты.

На одном изображении могут присутствовать объекты различных категорий, например: реки, дороги, водные объекты (озера, пруды, водохранилища), леса, поля, населенные пункты, автомобили и прочие. Каждая категория имеет, в общем случае, свой тезаурус, который, собственно, и характеризует соответствующую пространственную гранулу общей онтологии объектов изображений.

Множество объектов изображений, принадлежащих одной категории, представляется в виде таксономии понятий.

3. Общий алгоритм анализа изображений. Суть алгоритма заключается в построении дерева решений о категориях объектов. После выбора гипотезы формируется дерево вывода о соответствии объекта какому-либо ДЛ-определению из этой категории. При неудовлетворительной степени сходства и отличий происходит переход к другой ветви дерева решений о категориях. Результатом является атрибутивный граф объектов, найденных на изображении:

1. Цветовая сегментация (или тернарная сегментация черно-белого) изображения путем кластеризации палитры в пространстве HSV.

2. Выделение границ цветовых сегментов.

4. Аппроксимация границ круговыми дугами и прямыми.

5. Формирование синтаксических структур.

6. Формирование тезауруса путем нечеткой кластеризации слов.

7. Формирование списка категорий объектов на основе тезауруса изображения.

8. *Выбор* категории из списка.

8.1. Выбор области объекта и параметров локальной сегментации.

8.1.1. Локальная сегментация.

8.1.2. Уточнение синтаксических структур.

8.1.3. *Логический вывод* о классе объекта на основе ДЛ-определений.

8.1.4. Вычисление степени сходства и отличий.

Если «не удовлетворительно», то *Переход* к п. 8 для смены категории.

8.1.5. *Переход* к следующему объекту (п. 8.1).

8.2. *Переход* к следующей категории объектов (п. 8).

9. Формирование атрибутивного графа объектов изображения.

На рисунке 1 показаны примеры работы системы анализа изображений.

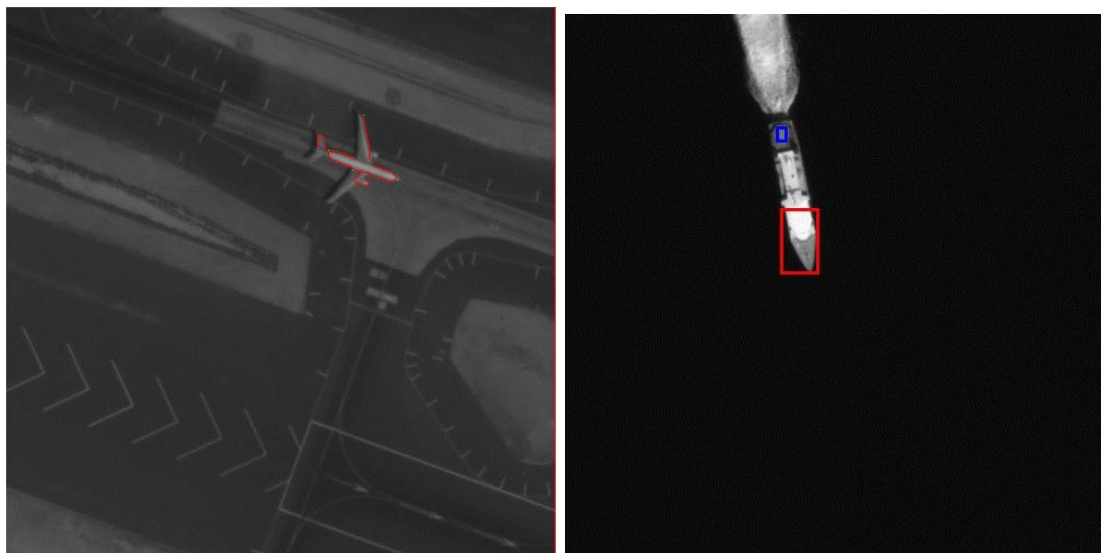


Рис. 1. Примеры работы системы анализа изображений: а) самолеты, автомобили; б) лодки, автомобили (Источник: <http://dataring.ru/competitions/>, 01.11.2016 – 04.04.2017).

Заключение. Таким образом, в условиях недостаточного качества изображений и четкости границ между объектами проблемой является не просто нахождение артефакта, но и отнесение его к известной категории и классу объектов. В описываемом подходе эта проблема решается путем кластерного анализа множества слов изображения и последующего определения принадлежности каждого слова тезаурусу какого-либо кластера. За счет этого сужается область поиска, осуществляется дополнительная подстройка параметров и фазификация значений.

СПИСОК ЛИТЕРАТУРЫ

1. Бычков И.В., Ружников Г.М., Фёдоров Р.К., Авраменко Ю.В. Интерпретатор языка SQL для обработки растровых изображений // Вычислительные технологии. Т. 21. №1. 2016. С. 49–59.
2. Кучуганов А.В. Поиск и описание искусственных объектов на аэрокосмоснимках // Седьмая международная конференция "Системный анализ и информационные технологии" САИТ - 2017 (13 - 18 июня 2017 г., г. Светлогорск, Россия): Труды конференции. ФИЦ ИУ РАН. 2017. С. 644–649. ISBN 978-5-904466-36-7.
3. Fu K.S. Syntactic Methods in pattern recognition. Academic Press. New York and London. 1974. Русский перевод: Фу К. Структурные методы в распознавании образов. М.: Мир. 1977. 319 с.
4. Hudelot C., Atif J., Bloch I. Integrating bipolar fuzzy mathematical morphology in description logics for spatial reasoning // European Conference on Artificial Intelligence ECAI 2010. Lisbon. Portugal (Aug 2010). Pp. 497–502.
5. Hudelot C., Atif J., Bloch I. Alc(f): a new description logics for spatial reasoning in images // 1st International workshop on computer vision + ontologyapplied cross-disciplinary technologies. 2014. Pp. 1–15.
6. Zoe Falomir, Ernesto Jimenez-Ruiz, M. Teresa Escrig, Lledó Museros. Describing Images using Qualitative Models and Description Logics / Preprint submitted to SPATIAL

UDK 004.93

DESCRIPTIVE LOGIC OF SEARCHING OBJECTS ON IMAGES

Aleksandr V. Kuchuganov

PhD, assistant professor, Department of Automated Data Processing and Control Systems, Kalashnikov Izhevsk State Technical University, 426069, city of Izhevsk, st. Studencheskaya, 7,
e-mail: Aleks_KAV@udm.ru

Abstract. The most popular approaches to the problem of searching objects on images are: a linguistic approach, within which syntactic recognition of objects of strictly predetermined structure is carried out; artificial neural networks; descriptive image algebras; approaches that describe objects using predicate logic; CBIR (Content-Based Image Retrieval) technologies, based on descriptive logics (DL).

The article presents an ontological approach based on descriptive logic with an extension to the spatial domain of data represented in the form of attributive graphs. The process of image analysis is controlled by a strategy containing: a preliminary morphological analysis stage; stage of the formation of a hypothesis about the category of the found object; stage of confirmation of the hypothesis by a logical conclusion about the class of the object. In the course of the analysis, a decision tree is constructed on the categories of objects. After selecting a hypothesis, a derivation tree is formed about the correspondence of the object to any DL-definition from this category. With an unsatisfactory degree of similarity and differences, a transition to another branch of the decision tree about categories occurs. Examples of the work of the image analysis system are shown.

Keywords: attributive graph, cluster analysis, descriptive logic, granular ontology, thesaurus, decision tree, output tree, degree of similarity and differences, graph of image objects.

References

1. Bychkov I.V., Ruzhnikov G.M., Fedorov R.K., Avramenko Y.V. Interpretator yazyka SOQL dlya obrabotki rastrovyykh izobrazheniy [The interpreter of the SOQL language for processing raster images] // Vychislitel'nyye tekhnologii = Computing technologies. T. 21. No. 1. 2016. Pp. 49–59. (in Russian)
2. Kuchuganov A.V. Poisk i opisaniye iskusstvennykh ob'yektov na aerokosmosnimkakh. [Search and description of artificial objects in aerospace vehicles] // Sed'maya mezhdunarodnaya konferentsiya "Sistemnyy analiz i informatsionnyye tekhnologii" SAIT - 2017 (13 - 18 iyunya 2017 g., g. Svetlogorsk, Rossiya): Trudy konferentsii. – FITS IU RAN = Seventh International Conference "System Analysis and Information Technologies" SAIT -

- 2017 (June 13 - 18, 2017, Svetlogorsk, Russia): Proceedings of the Conference. - FIC IU RAS. 2017. Pp. 644–649. ISBN 978-5-904466-36-7. (in Russian)
3. Fu K.S. Syntactic Methods in pattern recognition. Academic Press. New York and London. 1974.
 4. Hudelot C., Atif J., Bloch I. Integrating bipolar fuzzy mathematical morphology in description logics for spatial reasoning // European Conference on Artificial Intelligence ECAI 2010. Lisbon. Portugal (Aug 2010). Pp. 497–502.
 5. Hudelot C., Atif J., Bloch I. Alc(f): a new description logics for spatial reasoning in images // 1st International workshop on computer vision + ontologyapplied cross-disciplinary technologies. 2014. Pp. 1–15.
 6. Zoe Falomir, Ernesto Jimenez-Ruiz, M. Teresa Escrig, Lledó Museros. Describing Images using Qualitative Models and Description Logics / Preprint submitted to SPATIAL COGNITION AND COMPUTATION / Taylor & Francis. Volume 11. Issue 1. January 2011. Pp. 45–74 Available at: <http://dx.doi.org/10.1080/13875868.2010.545611>.