

УДК 004

АНАЛИЗ МЕТОДОВ ПОИСКА ИЗОБРАЖЕНИЙ В ЦИФРОВЫХ ХРАНИЛИЩАХ ДАННЫХ

Е.П. Доморацкий, Т.Н. Байбикова

Аннотация. В статье приведены обзор и анализ подходов, применяемых в системах поиска оцифрованной информации.

Ключевые слова: цифровое изображение, семантическая классификация изображений, анализ изображения.

Abstract. There is a review and analysis of methods for digital image retrieval in this paper.

Keywords: digital image, image semantic classification, image analysis.

Поиск мультимедийных документов в больших хранилищах данных вызывает с каждым годом все больший интерес. Таким образом, разработка новых парадигм поиска, удовлетворяющих запросам пользователей, является актуальной задачей.

Анализ литературы показывает, что для поиска оцифрованной информации в больших хранилищах данных, таких как, например, Интернет, применяются следующие подходы:

1. Индексирование текстовых описаний.

Довольно значительный период времени преимущественно использовался поиск визуальной информации, базирующийся на индексировании текстовых описаний, ассоциированных с изображением (поиск графических изображений в поисковых системах Яндекс и Google).

Данный способ поиска информации имеет ряд недостатков:

- неоднозначность между визуальным содержанием изображения и его текстовым описанием, снижающая показатели точности и полноты поиска;
- невозможность словесного описания ряда изображений, таких, как абстрактная живопись и т.п.;
- лексическая многозначность слов (нос, как часть лица и нос, как часть корабля);

■ при перемещении изображения необходимо перемещать одновременно с ним текстовое описание и т.п.

2. Поиск изображений по содержанию.

С начала 1980-х гг. активно разрабатываются методы поиска изображений по содержанию (CBIR – Content-based image retrieval). В CBIR-системах изображения индексируются по их визуальному содержанию (цвет, текстура, форма и т.д.).

Примерами таких программ являются:

■ **PIRIA.** Программа для индексирования и поиска изображения по сходству. По запросу-изображению находит подобные изображения, используя цвет, текстуру, форму.

■ **ScribR.** Система автоматической аннотации изображений, предоставляющая общую информацию о природе изображения и о его содержании.

В данной системе определяемыми атрибутами являются:

- «клипарт», схемы, рисунки, фотографии;
- цветные/черно-белые изображения;
- определение сцены внутри/снаружи помещений;
- нахождение лиц (фронтальных) и указание их местоположения.

Указанная система распознавания способна идентифицировать в изображениях следующие объекты: здания, траву, землю, песок, небо, снег, деревья, воду. При этом указывается процентное содержание указанных объектов по отношению ко всему изображению.

В настоящее время в свободном доступе находятся также следующие системы: Anaktisi, MFIRS, CIRES, Tiltomo, INRIA, Retrievr, Alipr, SIMPLIcity, Viper, FS. Однако в целом данные системы уступают системам, производящим поиск по текстовым аннотациям примерно в два-три раза.

3. Комбинированный подход, использующий и текстовые описания, и поиск изображений по содержанию. Например, поиск похожих изображений в поисковой системе Google Similar Images: на первом этапе применяется традиционный способ поиска изображения по текстовому запросу, а затем в качестве запроса применяют одно из найденных изображений и производят поиск по визуальному подобию.

В настоящее время все указанные выше подходы находят широкое применение. Однако в области поиска оцифрованной графической информации по-прежнему ведутся многочисленные исследования, что вызвано широким применением данного вида информации, как в научных целях, так и в повседневной жизни.

Направления исследований в этой области можно разбить на две группы:

- исследования, направленные на ускорение поиска информации;
- исследования, направленные на повышение качества найденной информации.

Здесь под качеством найденной информации подразумевается способность найденной информации удовлетворять поисковому запросу.

Поисковые запросы различаются по требованиям, предъявляемым к искомой информации. Запросы могут быть:

1. *Точные запросы*, требующие нахождения вполне определенных графических данных. Пример: поиск дубликатов изображений. Так как графическая информация занимает значительный объем в системах хранения данных, то хранение дубликатов изображений требует слишком много дополнительных ресурсов. Размещение информации в больших хранилищах данных может производиться многими пользователями. Таким образом, нахождение повторяющихся графических данных и их удаление из системы сокращает объем потребляемых ресурсов.

2. *Неточные запросы*, требующие нахождения подобной информации. Пример: нахождение нечетких дубликатов. В качестве нечетких дубликатов могут пониматься:

- изображения, отличающиеся разрешением;
- изображения, отличающиеся наличием шума;
- изображения, подвергшиеся незначительным фотометрическим преобразованиям, снимки одной и той же сцены, выполненные с несущественными изменениями ракурса камеры.

Также к нахождению нечетких дубликатов относят поиск:

- изображений, подвергшихся сильным искажениям;
- изображений, содержащих отдельные одинаковые сегменты;
- снимки одного объекта со значительно различающихся позиций.

С точки зрения человека, данные графические изображения идентичны, а с точки зрения поисковой системы – это два различных графических объекта.

Приведенные выше примеры показывают, что при решении задач, даже имеющих значительное сходство с точки зрения пользователя, с точки зрения поисковой системы должны использоваться совершенно разные механизмы поиска. Таким образом, алгоритмы, применяемые в том и другом случае, имеют различную природу. Алгоритмы для

нахождения дубликатов могут производить поиск путем попиксельного сравнения различных графических данных, размещенных в хранилищах. Алгоритмы для нахождения нечетких дубликатов должны оперировать с набором признаков и сравнивать эти признаки по заложенному в них правилу. Здесь часто применяются знания из теории нечетких множеств. А сами алгоритмы называют робастными.

В последнее время возрастает потребность в выполнении запросов, требующих наличия робастных алгоритмов.

Следует отметить, что к задачам поиска изображений по содержанию близко стоят задачи распознавания образов. Однако между ними имеется существенная разница. В задачах распознавания образов должна быть априорная информация о распознаваемых образах. Все образы делятся на классы с известным набором признаков. Распознаваемое изображение должно быть отнесено к какому-либо предопределенному классу, или оно остается неопознанным. Например, в системе распознавания лиц в базе данных находится набор фотографий определенного числа людей, сфотографированных в разных ракурсах. В том случае, если образ лица, находящегося перед камерой, находится в базе данных системы, лицо считается опознанным.

В задаче поиска изображений по содержанию требуется найти изображения, обладающие визуальным сходством с запросом. Изначально требования к идентификации класса входного изображения не ставится. Таким образом, задачи поиска изображений по содержанию изначально оперируют с такими понятиями, как подобие, сходство.

Несмотря на указанные различия между задачами поиска по содержанию и задачами распознавания, и те и другие в своей работе используют схожие алгоритмы, что приводит к тому, что эти задачи часто, не совсем обоснованно, относят в одну группу.

Также в обоих случаях широко используется применение робастных алгоритмов. Однако в случае распознавания изображений потребность в таких алгоритмах возникает из-за природы распознаваемых объектов, наличия помех, возникающих при приеме сигнала, проходящего через какую-либо физическую среду, например, воздушную. Задача робастных алгоритмов в этом случае состоит в том, чтобы при наличии шума иметь возможность отнести рассматриваемый объект к какому-либо классу. В случае поиска изображений по содержанию робастные алгоритмы работают с характеристиками объектов и проверяют близость данных характеристик, или, точнее, вектора характеристик на попадание в определенный диапазон.

В настоящее время наметились различные подходы, позволяющие повысить качество поисковых запросов. В целом проявляется тенденция использовать несколько различных парадигм поиска в одном способе, так что дать строгую классификацию способов не представляется возможным. Рассмотрим использование некоторых наиболее характерных методов.

СОЗДАНИЕ ОНТОЛОГИЙ

Онтологии необходимы для улучшения обработки пользовательских текстовых запросов.

В настоящее время широкую известность получил язык *Ontology Web Language (OWL)* – язык описания онтологий для семантической паутины. Язык позволяет описывать классы и отношения между ними, присущие веб-документам и приложениям. В настоящее время язык *OWL* рекомендован консорциумом Всемирной паутины.

В работе [1, р. 435–439] отмечается, что ручная разработка много-масштабных онтологий для таких динамичных ресурсов, как Интернет, не позволяет эффективно решить проблему организации знаний, в связи с чем предлагаются подходы, основанные на автоматических или полуматематических методах создания онтологий.

Разрабатываются подходы, в основе которых, помимо использования онтологий, также применяется обработка изображений [2]. Примером может послужить алгоритм, вначале индексирующий изображения, а затем объединяющий их в кластеры. Основные шаги работы данного алгоритма заключаются в следующем:

а) индексация изображений. Используется подход, основанный на классификации пикселей, находящихся на границе изображенного объекта, и пикселей, содержащихся внутри данного объекта [3, р. 1695–1709]. В результате вычисляется вектор, содержащий 128 элементов для каждого изображения;

б) кластеризация изображений. Индексированные изображения объединяются в кластер. Для этого применяется алгоритм *k-SNN (Shared Nearest Neighbors, k ближайших соседей)*, описанный в работе [4].

Таким образом, к достоинствам подходов, основанных на онтологиях, добавляются достоинства подходов, основанных на цифровой обработке изображений.

К недостаткам подходов, аналогичных указанному выше, можно отнести:

- трудность применения алгоритма для составного изображения, на котором изображено сразу несколько объектов;
- неточное описание объекта, представленного в неудачном ракурсе, или изображенного частично;
- возможное совпадение вектора признаков у несхожих объектов и т.п.

К тому же при работе с любыми онтологиями необходимо привлечение экспертов, что вносит элементы субъективизма в разрабатываемые методики.

ПОИСК, ОСНОВАННЫЙ НА АНАЛИЗЕ ИЗОБРАЖЕНИЙ

Существует обширная группа методов, которая для распознавания изображений производит анализ данных, составляющих изображение. В процессе анализа одной из первых его стадий является сегментация изображения – процесс разделения изображения на однородные области. Результаты сегментации влияют на все последующие операции, производимые при дальнейшем анализе. Для разных видов графических данных приходится подбирать алгоритмы, дающие приемлемые результаты именно для данного вида изображений. Таким образом, алгоритмы сегментации изображений (АСИ) не являются универсальными.

В работе [6, с. 163–172] приводится обобщенная классификация алгоритмов сегментации изображений на основе подробного анализа существующих решений.

Приведем наиболее известные классификации.

Классификация АСИ Скарбека и Кошана. В соответствии с данной классификацией все алгоритмы разделены на следующие 4 группы:

1. алгоритмы, основанные на анализе свойств пикселей;
2. алгоритмы, основанные на анализе свойств области;
3. алгоритмы, основанные на выделения краев областей;
4. алгоритмы, учитывающие физические свойства объектов, представленных на изображении.

Классификация АСИ Лючиса и Митра. Здесь выделено три класса алгоритмов:

1. алгоритмы, основанные на анализе свойств на пространстве признаков;
2. алгоритмы, основанные на анализе свойств областей;

3. алгоритмы, учитывающие физические свойства объектов, представленных на изображении.

Приведенные выше классификации дают представление о работе алгоритмов данного вида в целом. В основе лежит обработка пикселей, анализ его свойств, а также свойств пикселей, соседних с ним, или находящихся в непосредственной близости, также в отдельных алгоритмах учитывается частота отдельно взятого цвета и т.п.

Недостатками данных алгоритмов являются:

- отсутствие универсальности;
- оценка правильности работы многих алгоритмов требует привлечения экспертов, которые сегментируют исходное изображение вручную, а затем сравнивают результаты работы алгоритмов с результатами ручной сегментации.

ПОИСК ПО ЦИФРОВОМУ ОБРАЗЦУ

В начале 1990-х гг. началась активная разработка методов поиска на основе визуального образца. При этом критериями для признания изображений визуально похожими считалось:

- наличие общей цветовой гаммы;
- наличие похожих по форме объектов.

В качестве признаков изображений, используемых в данных методах, выступают:

- цветовые характеристики;
- текстурные характеристики;
- дескрипторы формы;
- коэффициенты вейвлет-преобразования.

Существуют методы, основанные на других принципах, например, байесовских вероятностных моделях и т.п.

Данный вид поиска имеет много схожего с предыдущим, здесь также производится анализ изображений, однако выраженной особенностью является наличие цифрового образца. Таким образом, оптимизация запроса производится посредством обработки графической информации.

При поиске по цифровому образцу поиск может производиться:

- по уменьшенной копии, или пиктограмме;
- по сжатому экземпляру изображения;
- по отдельному фрагменту;
- поиск изображений с различными пропорциями и т.п.

Недостатком подходов, в которых в качестве поискового запроса применяется графический образец, является большой объем обрабатываемых данных для поиска информации. К тому же в случае обработки уменьшенных копий и сжатых изображений неизбежно присутствует погрешность вычислений, т. к. характеристики пиктограммы и сжатого изображения всегда отличны от оригинала.

Для улучшения данного вида методов предлагается ряд подходов. В качестве примера можно привести психофизиологический подход, при котором улучшение достигается тем, что в качестве образца выбирается не миниатюрная копия изображения, в которой возможна потеря информации, значимой для пользователя, а часть изображения, представляющая область интереса для наблюдателя.

В работе [5] подход к поиску изображения основан на том, что внимание человека в процессе наблюдения изображения концентрируется не на всей его области, а на некоторых вполне определенных участках изображения. При этом выбор области (или нескольких областей), привлекающей внимание человека, в большей степени обусловлен свойствами данной области. В указанной работе выполнено исследование возможности применения моделей восходящего внимания для нахождения области изображения, приблизительно соответствующей объекту. Таким образом, для визуального запроса возможно использование изображения, значительно меньшего размера, которое, к тому же, представляет собой область интереса.

СЕМАНТИЧЕСКАЯ КЛАССИФИКАЦИЯ ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЙ

В последнее время появилось значительное количество работ, в которых предпринимается попытка установить связь между низкоуровневыми характеристиками изображения и его семантикой. Под низкоуровневыми характеристиками понимаются цвет, граница, текстура изображения. Пользователь взаимодействует с более высокоуровневыми характеристиками, с абстрактными и субъективными понятиями. Таким образом, то, что для человека выглядит подобным, например, один и тот же пейзаж в разное время суток, для распознающей системы представляется различными изображениями.

Методы, производящие семантическую классификацию, часто основаны на обучении системы распределять изображения с различны-

ми характеристиками (цвет и т.п.) по различным группам. В некоторых случаях первоначально строят соответствия между низкоуровневыми характеристиками и лексическими характеристиками, затем связывают лексические характеристики с семантикой изображения. Например, выделяют низкоуровневые характеристики, обозначающие на лексическом уровне реку, лес и т.п., в результате семантикой изображения будет река в лесу и т.п.

Недостатки данных методов:

- требуется привлечение экспертов;
- требуется обучение системы;
- существует проблема «*семантического разрыва*» – отсутствия однозначной связи между низкоуровневыми характеристиками и семантикой изображения и т.п.;
- различные семантические пространства, в которых оперируют пользователь и система поиска изображений. Система оперирует низкоуровневыми характеристиками, а пользователь оперирует более высокоуровневыми, абстрактными и субъективными понятиями.

МЕТОДЫ АВТОМАТИЧЕСКОЙ ИДЕНТИФИКАЦИИ ЛИЦ НА ЦИФРОВЫХ ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЯХ И В ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ

Данная группа методов является наиболее изученной и описана в ГОСТ [7, с. 379–382], что позволяет выделить ее в отдельную группу.

В указанных методах изображение (image) рассматривается, как двумерное представление яркости и текстуры объекта в определенных условиях освещения. Для поиска лиц (идентификации биометрических объектов) используются так называемые контрольные точки (feature points), под которыми подразумеваются «заранее определенные точки на изображении лица, используемые алгоритмами распознавания лица». Примером таких контрольных точек являются точки центров глаз.

Применение контрольных точек широко используется в алгоритмах, применяемых в медицинских системах, и вообще в системах, в которых имеется предположение о форме распознаваемого объекта.

Несмотря на все многообразие методов обработки цифровых графических изображений, в основе их работы лежат общие подходы, которые можно разделить на группы:

1. Методы, основанные на глобальных характеристиках изображения. Данные методы группируют пиксели на основе анализа данных в пространстве признаков. Примерами методов данной группы являются метод кластеризации и метод, использующий пороговую гистограмму.

2. Методы, основанные на локальных характеристиках изображения. Данные методы анализируют спектральные характеристики соседних пикселей. Примерами методов данной группы являются методы обнаружения краев и методы выделения областей.

3. Адаптивные методы (как развитие локальных) В этом случае применяются алгоритмы, обрабатывающие окрестности пикселей не для всего изображения, а для отдельных областей, так называемых зон интереса.

Подводя итог, можно отметить, что на сегодняшний день не существует универсальных методов, способных производить поиск изображений.

Развитие существующих методов не способно дать значимых результатов и может дать улучшения лишь для какого-либо отдельно взятого вида изображений.

По мнению автора статьи, для устойчивого поиска изображений в цифровых хранилищах данных необходимо вносить словесные описания изображений, и при этом также использовать наборы характеристик, рассчитанных по анализируемому изображению.

В работе [8] предлагаются некоторые способы внедрения дополнительной информации в цифровые графические изображения, которые возможно использовать для реализации идеи. Данные способы разработаны в МИЭМ НИУ ВШЭ (Московском институте электроники и математики национального исследовательского университета Высшей школы экономики). Имеется положительное решение о выдаче патента.

Основное отличие подхода, предлагаемого автором статьи, заключается в том, что данные описания предполагается хранить в самом изображении, таким образом, необходимо реализовать подход, при котором изображение «знает» о самом себе как можно больше информации. При этом смысл, заключенный в изображении, заложен на этапе его создания, без необходимости сопоставления низкоуровневых характеристик с лексикой, а лексики со смыслом.

Данный подход позволяет упростить поиск изображений, имеющих разный масштаб, отличающиеся разрешением, подвергшиеся незначи-

тельным фотометрическим преобразованиям, снимки одной и той же сцены, выполненные с несущественными изменениями ракурса камеры и т.п.

Данный метод необходимо использовать в совокупности с рядом других.

Дальнейшее направление работы – проверка работы этого метода в комбинации с другими группами методов и выделение наиболее успешных комбинаций.

ЗАКЛЮЧЕНИЕ

В статье произведен анализ методов поиска изображений в цифровых хранилищах данных. Описаны основные направления развития данных методов, указаны их слабые места.

Также предложено направление дальнейшей работы для повышения эффективности обработки поисковых запросов.

Подводя итог, можно сказать, что поиск цифровых изображений в хранилищах данных является актуальной задачей, в результате решения которой увеличивается количество успешных поисковых запросов и качество найденной информации.

ЛИТЕРАТУРА

1. *Cimiano P., Hotho A., Staab S.* Comparing Conceptual, Divisive and Agglomerative Clustering for Learning Taxonomies from Text, in Proc. of ECAI 2004, Valencia, Spain, 2004.

2. *Ertoz L., Steibach M., Kumar V.* Finding Topics in Collections of Documents: A Shared Nearest Neighbor Approach, In: Wu, W., Xiong, H., Shekhar, S.(eds.): Clustering and Information Retrieval, Kluwer, 2003.

3. *Kartikeyan, B., Sarkar, A. and Majumder, K.L.*, 1998. A Segmentation approach to classification of remote sensing imagery. International Journal of Remote Sensing, 19(9).

4. *Popescu A., Millet C., Moëllic P-A.* Ontology Driven Content Based Image Retrieval, CIVR 2007 – posters session, July 9 – 11, 2007, Amsterdam, The Netherlands.

5. ГОСТ Р ИСО/МЭК 19794-5-2006. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными. Часть 5. Данные изображения лица.

6. *Левашкина А.О., Поршнев С.В.* Универсальная классификация алгоритмов сегментации изображений // Журнал научных публикаций аспирантов и докторантов, 2008. № 3.

7. Левашкина А.О., Поршнев С.В. Вычислительный алгоритм поиска на изображении протообъекта // Математические методы распознавания образов: 14-я Всероссийская конференция. Владимирская обл., г. Суздаль, 21–26 сентября 2009 г.: Сборник докладов. – М., 2009.

8. Патент – 2495494 РФ, МПК6 G06T 3/00. Способ внесения дополнительной информации в цифровые графические изображения / Т.Н. Байбикова, А.К. Зыков, П.В. Орлов, А.А. Соснин; Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Национальный исследовательский университет «Высшая школа экономики». № 2011141314; Заяв. 12.10.2011; Оpubл. 10.10.2013. Бюл. № 28.

Е.П. Доморацкий

д-р техн. наук, профессор,

Национальный исследовательский университет «Высшая школа экономики»

E-mail: edomoratsky@hse.ru

Т.Н. Байбикова

ст. преподаватель,

Национальный исследовательский университет «Высшая школа экономики»

E-mail: tbaibicova@hse.ru