



ТЕХНИЧЕСКИЕ НАУКИ

ИССЛЕДОВАНИЕ И КЛАССИФИКАЦИЯ МЕТОДОВ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ В СИСТЕМАХ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Др. В. Гриценко, Н. С. Дорошенко

THE INVESTIGATION AND CLASSIFICATION OF IMAGE RECOGNITION METHODS IN COMPUTER VISION SYSTEMS

Gritsenko A. V., Doroshenko N. S.

The systematization of existing approaches to the solution of image recognition problems is given in the article from the positions of two main aspects – the type of used image presentation and recognition quality criterion; the advantages and disadvantages and the applicability domain of each method are also determined. The investigations have been fulfilled in the frames of the “Scientific and Scientific-Pedagogical Personnel of the Innovation Russia” Federal Designated Project.

Key words: computer vision, image recognition, recognition quality criteria.

В статье проведена систематизация существующих подходов к решению задач распознавания изображений с позиций двух основных аспектов – вида используемого представления изображения и критерия качества распознавания, а также определены преимущества, недостатки и область применимости каждого метода. Исследования выполнены в рамках ФЦП «Научные и научно-педагогические кадры инновационной России».

Ключевые слова: компьютерное зрение, распознавание изображений, критерии качества распознавания.

УДК 004.93

В связи со сложностью решения общей задачи распознавания изображения и отсутствия четкого представления об универсальных подходах, возникает большое количество частных методов. При решении той или иной задачи возникает вопрос об эффективности существующих методов, их сходствах и различиях.

Для применения того или иного метода к изображению, которое изначально представляется как совокупность пикселей, а точнее массив пикселей, вводят так называемые уровни представлений изображений [1]. При этом под уровнем представления понимается система, в которой содержатся алгоритмы получения описаний объектов изображения заданных в определенной форме. Исходя из этого, процесс распознавания объектов можно интерпретировать как процесс последовательного преобразования информации, от исходного изображения, представленного в виде массива пикселей, до семантического описания изображения через некоторую совокупность промежуточных представлений.

На основании существующих типов представлений, выделяют следующие группы методов распознавания изображения [2]:

- низкоуровневые;
- признаковые;
- контурные;
- структурные.

Также необходимо рассмотреть критерии качества описания изображений, так как они в большей степени характеризуют мето-



ды распознавания, ведь именно они определяют, какое именно описание изображения будет сформировано в рамках используемого представления.

Выделяют следующие основные подходы к заданию критерия качества [3]:

- эвристический;
- байесовский;
- энтропийный;
- теоретико-информационный.

Рассмотрим указанные типы представлений, а также подходы к заданию критерия качества и на основе полученных данных проведем анализ возможности применения используемых методов в задачах распознавания изображений.

1. Низкоуровневые методы

Если представить изображения (массивы пикселей) в виде элементов математического пространства, то становится возможным использование строгих и внутренне непротиворечивых средств анализа и преобразования изображений. Таким образом, появляется возможность «измерять» графические элементы, что в дальнейшем служит основой для их распознавания.

Выделяют два следующих основных класса математических моделей изображений [4]: функциональные представления и представления в виде случайных полей.

1.1. Функциональные модели

В функциональных моделях изображение интерпретируется как функция из некоторого пространства, в котором становится возможным проведение таких операций над изображениями, как пространственное преобразование (масштабирование, вращение, искривление и др.), преобразование интенсивности (яркость, контраст), фильтрация (сглаживание, дифференцирование).

При распознавании широко распространено использование корреляционных методов, с помощью которых осуществляется поиск смещения между эталонным и распознаваемым изображением, при наличии условия линейности данных преобразований.

К минусам функциональных моделей можно отнести то, что представления изображений, как правило, инвариантны (т. е. описание изображения неизменно при некоторых его преобразованиях) к сравнительно простым преобразованиям изображений, в большинстве случаев пространственным преобразованиям ограниченных классов, следовательно область их применения очень ограничена.

1.2. Модели на основе случайных полей

Концепцией данных моделей является представление изображений как выборочных функций (реализаций) случайного процесса. Эти модели более общие, в сравнении с функциональными, так как представляют изображения в виде некоторых случайных функций, по отношению к которым детерминированные функции являются частным подклассом.

Стохастические модели позволяют использовать взаимную информацию, находимую через энтропию, вместо корреляции как в функциональных моделях. Данная мера сходства по величине взаимной информации более универсальна, так как является так же инвариантной по отношению к произвольным глобальным преобразованиям яркости изображений, а не только к линейным преобразованиям. Но это приводит к увеличению размерности пространства параметров и может стать проблематичным при преобразовании подобия.

В силу того, что стохастические модели изображений характеризуются определенными параметрами, становится возможным отделять случайную составляющую изображений от регулярной составляющей (параметров модели). Появляется возможность сопоставления изображения с эталоном по этим параметрам. Но все же описание изображения через параметры модели следует рассматривать в качестве признакового представления изображений.

Исходя из этого, можно сделать выводы о том, что функциональные и стохастические модели в основном используются на предварительном этапе обработки изображений и для построения описаний после-



дующих уровней, либо для задач распознавания изображений, обладающих малой изменчивостью.

2. Признаковые методы

Признак выражает некоторую значимую характеристику объекта в численной форме и определяется как функция от значений пикселя или группы пикселей

Признаки можно разделить на следующие [5]:

- Общие признаки: признаки, независимые от приложения (цвет, текстура, форма).
- Предметно-зависимые признаки: признаки, зависящие от приложения (например, описывающие лица людей, отпечатки пальцев и т. д.). Эти признаки формируются на основе общих признаков для конкретной предметной области.

Признаковый подход очень часто используется в распознавании объектов или структур, так как описание изображения или его частей совокупностью признаков позволяет применять дискриминантные методы.

Сами признаковые методы могут применяться только в том случае, когда удастся составить инвариантные признаки. Разработчикам приходится составлять набор признаков для каждого конкретного случая. Впоследствии по выбранным признакам возможно применение математического аппарата распознавания образов.

Использование признаковых методов позволяет осуществлять распознавание объектов довольно в широком спектре задач, однако описание сложных сцен или объектов, имеющих сложный тип изменчивости, в рамках данного подхода оказывается затруднительным, как и распознавание по глобальным признакам объектов по их фрагментам.

3. Контурные методы

В условиях высокой изменчивости низкоуровневые и признаковые методы оказываются трудно формализующимися. В данном случае, как показывает мировой опыт исследователей, оказываются крайне полезными в задачах сопоставления и распозна-

вания изображений контурные и структурные методы.

Под контуром, в частности, понимается местоположение локального изменения или резкого перепада яркости на изображении. При этом подразумевается, что такие перепады возникают на границах объектов.

При построении контурных описаний изображений выделяют глобальные и локальные процедуры построения [6]. Для глобальных процедур характерно разбиение изображения на однородные области, на основе которых и строятся контуры, как границы этих областей либо как их срединные оси.

Локальные же процедуры основываются либо на определении цепочек максимумов на градиентном поле изображения (либо пересечений нуля второй производной), либо на непосредственной аппроксимации яркостных переходов в рамках функциональных или стохастических моделей изображений.

Однако существует проблема формализации понятия контура, о чем свидетельствует разнообразие подходов к обнаружению контуров.

Обычно в пользу использования контурных методов приводят следующие аргументы:

- контуры объекта практически неизменны на изображениях, полученных в разное время, при разных ракурсах или при смене датчика, в отличие от его остальных точек;
- контур является концентратором информации в изображении;
- контур полностью характеризует форму объектов;
- уменьшение обрабатываемых точек изображения (обрабатываются лишь точки контура), значительно сокращает объем вычислений.

В итоге, хоть контурные методы распознавания и обладают высокой степенью робастности, в них оказывается затруднительным поиск пространственных преобразований сложного типа, а также распознавание объектов при наличии большого числа ложных контуров.

4. Структурные методы

В качестве основы структурного представления изображения могут быть использованы [5]:

- точки интереса (точки, в которых значение некоторого локального признака достигает экстремального значения), строящиеся в рамках признакового подхода;

- выделяемые геометрические элементы, на основе крайних точек, не объединенных в связанные контуры;

- выделения элементов, на основе контуров, путем сегментации, их аппроксимации, либо обнаружения точек максимальной кривизны.

В последствии на основе непроектируемых элементов формируются составные структурные элементы.

Поиск соответствий между структурными элементами производится с учетом взаимного расположения данных элементов, их размеров, взаимной ориентации и др. Таким образом, вместо поиска параметров пространственного преобразования производится поиск вариантов соответствий структурных элементов.

К сожалению, в большинстве случаев привлекаемые структурные представления не полностью инвариантны по отношению условиям формирования изображений. В связи с этим, на изображениях, полученных в разных условиях, структурные элементы выделяются по-разному, что уменьшает вероятность корректного сопоставления и снижает его точность.

5. Критерии качества

Во многих методах распознавания изображений используются критерии, применимость которых для решения конкретной задачи строго не обосновывается, либо на нее накладываются серьезные ограничения (мера Хаусдорфа для контуров, коэффициент корреляции для значений яркости).

Хотя данные критерии и имеют математическое обоснование, их применение несет творческий характер [1]. Такие методы очень разнообразны и обладают низкой достоверностью, поэтому их применение оправдано

только для простых задач распознавания или на начальном этапе проектирования.

Байесовские методы определения критерия качества распознавания изображений являются наиболее разработанными и широко применяемыми на практике. Решение в них строится на основе апостериорной вероятности, вычисляемой по правилу Байеса. Сложность в данном случае заключается в неизвестности априорной вероятности решения и трудности явного задания распределения вероятности. Эти недостатки в совокупности с проблемой задания плотности распределения условных вероятностей накладывают некоторые ограничения на использование стохастических моделей с некоторыми типами представлений.

Энтропийный подход основывается на понятии энтропии, введенной в шенноновской теории информации. В качестве критерия выбора решения выступает энтропия значений, хранящихся в пикселях изображения. Согласно принципу минимум энтропии, лучшим значением считается то, для которого энтропия минимальна. Этот принцип можно дополнить принципом максимума энтропии. Согласно ему, если два распределения хорошо описывают частоты значений случайной величины, то выбрать следует распределение, обладающее максимальной энтропией. Сочетание этих принципов приводит к минимуму энтропии, когда решение заключается в поиске компромисса между точностью и простотой модели.

Теоретико-информационные методы базируются на принципе минимальной длины описания. Подход основывается на алгоритмической теории информации, в котором количество информации, содержащейся в некотором наборе данных, равно длине минимальной подпрограммы, способной воспроизвести эти данные. Распознавание изображений происходит без прямого обращения к алгоритмической теории информации. В рассматриваемом подходе задается описание (длина которого заменяет значение правдоподобия), по которому можно восстановить исходное изображение, используя явное представление изображения (содержащее априорное распределение вероятностей).



Большое разнообразие методов является следствием широкого диапазона предметных областей, в которых они применяются. Низкоуровневые и признаковые методы наиболее эффективны в узких предметных областях, т. к. остро реагируют на изменчивость изображений, причинами которой могут служить изменение ракурса съемки, ос-

вещения, типа камеры). Контурные и структурные методы являются более универсальными, т. е. применимыми к широким предметным областям с высокой изменчивостью изображений. Но вместе с тем, они обладают рядом недостатков, наиболее значительным из которых является высокая вычислительная сложность.

Методы	Преимущества	Недостатки
Низкоуровневые	Простота реализации, высокое качество распознавания даже в условиях нескольких объектов на изображении	Минимальная степень инвариантности к факторам изменчивости изображений
Признаковые	Быстрое распознавание изолированных объектов и категоризация изображений	Низкая степень инвариантности к факторам изменчивости изображений
Контурные	Достаточно высокая эффективность и простота реализации при незначительной зависимости от изменчивости изображений	Низкое качество распознавания при сложных пространственных изменениях
Структурные	Возможность распознавания при наличии перспективных искажений и сложных типах изменчивости объектов	Плохая применимость к объектам малого размера, высокая ресурсоемкость

Критерии выбора качества распознавания изображений также различаются по применимости к различным предметным областям. Эвристические методы нашли свое место в работе с простыми представлениями благодаря наименее трудоемкой реализации. Теоретико-информационные методы представляются наиболее перспективными для наиболее сложных задач распознавания. Байесовский и энтропийный подходы могут применяться для большинства задач с достаточно приемлемым результатом, но для изображений с высокой степенью изменчивости и сложным представлением нахождение решения может иметь очень высокую вычислительную сложность.

Таким образом, для решения узконаправленных задач, т. е. при распознавании объектов четко определенных типов (лица, люди, здания, символы) рассмотренные подходы дают вполне приемлемый результат. Сложность возникает при распознавании произвольных объектов, например, когда необходимо распознать объекты нескольких заранее неопределенных типов. Существующие методы позволяют лишь в незначительной степени решить эту проблему. Поэтому создание эффективных универсальных решений требует дальнейшей работы в этом направлении.

ЛИТЕРАТУРА

1. Потапов А. С., Гуров И. П., Васильев В. Н. *Математические методы и алгоритмическое обеспечение анализа и распознавания изображений в информационно-телеком-*

муникационных системах // Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы». – 2008. – 46 с.

2. Rares A., Reinders M. J. T., Hendriks E. A. *Imageinterpretationsystems // TechnicalReport (MCCWS 2.1.1.3.C), MCCWSproject, InformationandCommunicationTheoryGroup. TUDelft. – 1999. – 32 p.*
3. *Потапов А. С. Исследование представлений изображений на основе принципа репрезентационной длины описания // Изв. вузов. Приборостроение. – 2008. – Т. 51. – № 7. – С. 3–7.*
4. Chan T. F., Shen J., and Vese L. *Variational PDE models in image processing // Notice Amer. Math. Soc. – 2003. – V. 50. – P. 14–26.*
5. Lei B. J., Hendriks E. A., Reinders M. J. T. *On feature extraction from images // Technical Report, Deliverable 2.1.1.2.A+B, MCCWS project, Information and Communication Theory Group. TU Delft. – 1999. – 57 p.*
6. Brown L. G. *A survey of image registration techniques // ACM Computing surveys. – 1992. – V. 24. – P. 325–376.*
7. Linying S., Sharp B., Chibelushi C. *Knowledge-based image understanding: a rule-based production system for X-ray segmentation // Int. Conf. on Enterprise Information Systems (ICEIS). – 2002. – P. 530–533.*

Об авторах

Гриценко Артем Владимирович, ГОУ ВПО «Ставропольский государственный университет», аспирант 1 года обучения, специальность «Методы и системы защиты информации, информационная безопасность». Сфера научных интересов – компьютерное зрение (интеллектуальный анализ изображений), распределенные высокопроизводительные вычисления, технологии обнаружения аномалий сетевого трафика.

gricenko.ar@stavsu.ru

Дорошенко Николай Сергеевич, ГОУ ВПО «Ставропольский государственный университет», аспирант 1 года обучения, специальность «Математическое моделирование, численные методы и комплексы программ». Сфера научных интересов – технологии компьютерного зрения, методы обработки изображений, интеллектуальные системы безопасности.

Doroshenko@stavsu.ru

