

УДК 004.932.2+004.932.72'1

Д.Г. Мурадина, Н.С. Костюкова

Донецкий национальный технический университет, г. Донецк
кафедра прикладной математики и информатики

ИССЛЕДОВАНИЕ ОСНОВНЫХ МЕТОДОВ КЛАССИФИКАЦИИ КОЛЛЕКЦИЙ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Аннотация

Мурадина Д.Г., Костюкова Н.С. Исследование основных методов классификации коллекций цифровых изображений. Описаны базовые методы классификации изображений. Выполнен анализ этих методов, описаны варианты предварительной обработки изображений для достижения оптимального результата применения рассмотренных методов.

Ключевые слова: *методы классификации изображений, выделение границ объектов, оконтуривание изображений.*

Постановка проблемы. Человек около 90% информации об окружающем мире получает благодаря зрению. В сфере компьютерных технологий в качестве источника информации может выступать текст, аудио-/видеофайл или изображение. Очень часто, а в основном в сети Интернет, возникает необходимость найти то или иное изображение. В последние годы фотография стала популярным увлечением среди людей благодаря доступу к аппаратному обеспечению, как минимум, практически в каждый телефон встроена камера. Фотографии имеют свойство накапливаться и с течением времени может усложниться поиск по каталогу снимков. Результаты поиска изображений можно улучшить, подключив человека к процессу классификации. Именно такой подход применяется на многочисленных фотохостингах и фотосервисах, когда задачи описания содержимого изображений возлагаются непосредственно на пользователей, например, с помощью подробных «анкет» изображений. Но не у каждого хватает желания, терпения и свободного времени подробно описать («тэгами») весь перечень загружаемых снимков. Или, допустим, в интернет-магазинне одежды нужно построить поиск объекта по свойству, например, только одежды в клеточку. Для этого каждое изображение необходимо отнести к классу «клеточка».

Эти примеры доказывают необходимость изучения и развития методов классификации изображений.

Цель статьи – исследование основных методов классификации коллекций цифровых изображений.

Постановка задачи исследования. Анализ семантики образа на изображении является основной и последней задачей, которую необходимо

решать. Прежде нужно понять, с чем же конкретно придется работать, то есть декомпозировать изображение и соотнести признаки каждого объекта изображения с признаками объекта класса.

Существующие методы и подходы. Среди существующих методов решения задачи можно выделить несколько базовых подходов. Первый из них – распознавание на основе методов теории решений. Его подход основан на использовании решающих или дискриминантных функций. Пусть $x = (x_1, x_2, \dots, x_n)^T$ – n -мерный вектор признаков объекта. Предположим, что существует W классов образов w_1, w_2, \dots, w_W . Требуется найти W дискриминантных функций $d_1(x), d_2(x), \dots, d_W(x)$, таких, что если образ x принадлежит классу w_i , то дискриминантная функция с индексом i имеет большее значение, чем другие. В методах распознавания, основанных на сопоставлении, каждый класс представляется вектором признаков образа, являющегося прототипом этого класса. Незнакомый образ приписывается к тому классу, прототип которого оказывается ближайшим в смысле заранее заданной метрики. Простейший подход состоит в использовании классификатора, основанного на минимальном расстоянии, который, как ясно из названия, вычисляет евклидовы расстояния между вектором признаков неизвестного объекта и каждым вектором прототипа. Решение о принадлежности объекта к определенному классу принимается по наименьшему из таких расстояний. Метод корреляционного сопоставления состоит в том, что имеется эталон, который ищется на изображении методом «скользящего окна».

Ещё один подход – статистически оптимальные классификаторы (байесовский). Как и в большинстве областей, связанных с измерением и интерпретацией физических явлений, вероятностные подходы оказываются важными в задаче распознавания образов из-за случайностей, влияющих на порождение классов образов. Можно выработать такой метод классификации, который будет оптимальным в том смысле, что при его использовании будет достигаться наименьшая (в среднем) вероятность появления ошибок классификации. Байесовский подход является классическим в теории распознавания образов и лежит в основе многих методов. Он опирается на теорему о том, что если плотности распределения классов известны, то алгоритм классификации, имеющий минимальную вероятность ошибок, можно выписать в явном виде.

В рассмотренных подходах сущность обучения проста. Обучающие образы каждого класса используются для вычисления параметров дискриминантной функции, соответствующей этому классу. После того, как оценки необходимых параметров получены, структура классификатора становится фиксированной, и его окончательное качество зависит лишь от того, насколько хорошо реальные совокупности образов отвечают статистическим предположениям, изначально сделанным при выводе используемого метода классификации.

Существуют методы, в которых необходимые дискриминантные функции строятся непосредственно в ходе обучения. Архитектура многослойной нейронной сети (МНС) состоит из последовательно соединённых слоёв, где нейрон каждого слоя своими входами связан со всеми нейронами предыдущего слоя, а выходами - следующего. НС с двумя решающими слоями может с любой точностью аппроксимировать любую многомерную функцию. НС с одним решающим слоем способна формировать линейные разделяющие поверхности, что сильно сужает круг задач ими решаемы. НС с нелинейной функцией активации и двумя решающими слоями позволяет формировать любые выпуклые области в пространстве решений, в том числе и невыпуклой. При этом МНС не теряет своей обобщающей способности. Обучаются МНС при помощи алгоритма обратного распространения ошибки, являющегося методом градиентного спуска в пространстве весов с целью минимизации суммарной ошибки сети. При этом ошибки (точнее величины коррекции весов) распространяется в обратном направлении от входов к выходам, сквозь веса, соединяющие нейроны.

Анализ и продуктивное использование методов. Классификатор по минимуму расстояния хорошо работает в тех практических задачах, где расстояния между точками математических ожиданий классов велики по сравнению с диапазоном разброса объектов каждого класса, но на практике такие случаи встречаются редко, за исключением тех, когда само создание изображения и классов подразумевает в дальнейшем распознавание, например, набор символов (шрифт).

Корреляционная функция может быть нормирована относительно размеров изображения и размеров «скользящего окна» путем перехода к коэффициенту корреляции, но достичь нормировки относительно поворота или изменения размера достаточно сложно.

В реальных задачах статистические свойства классов образов зачастую неизвестны или не поддаются оценке. На практике для таких задач теории решений более эффективными оказываются методы, в которых необходимые дискриминантные функции строятся непосредственно в ходе обучения. Это устраняет необходимость использовать предположения о функциях плотности распределения вероятностей или о каких-то других вероятностных параметрах рассматриваемых классов. В качестве таких методов выступают нейросетевые методы распознавания изображений с многослойной архитектурой. Положительные свойства нейросетевых методов заключаются в том, что они обеспечивают быстрое и надёжное распознавание изображений. Однако при применении этих методов к изображениям трёхмерных объектов возникают трудности, связанные с пространственными поворотами и изменением условий освещённости.

Чтобы достичь положительных результатов любая система должна стремиться извлечь характеристики, инвариантные к внутриклассовым изменениям и максимально репрезентативные по отношению к межклассовым изменениям. А также к изображению могут быть применены различные методы улучшения «читаемости», то есть уменьшение количества шумов, выделение контуров, изменение контрастно-яркостных характеристик.

Выводы. Исходя из результатов анализа базовых методов классификации изображений можно сделать вывод, что наилучший результат может быть достигнут либо с помощью вмешательства человека, который сможет субъективно оценить степень сложности изображения и выбрать подходящий метод, либо описанные подходы могут использоваться в совокупности с итоговым результатом с заданной вероятностью.

Список литературы

1. Р.Гонсалес, Р.Вудс Цифровая обработка изображений: Пер. с англ. – М.: Издательский дом «Техносфера», 2005. – 1073 с.
2. Л.Шапиро, Дж.Стокман Компьютерное зрение: Пер. с англ. – М.: Издательский дом «Бином. Лаборатория знаний», 2006. – 762 с.
3. Классификация изображений / Интернет-ресурс. - Режим доступа: [www/URL: http://courses.graphicon.ru/files/courses/vision/2010/cv_2010_07.pdf](http://courses.graphicon.ru/files/courses/vision/2010/cv_2010_07.pdf)
4. Нейросетевые методы распознавания изображений / Интернет-ресурс. - Режим доступа: [www/URL: http://rusnauka.narod.ru/lib/author/briluk_d_b/1/](http://rusnauka.narod.ru/lib/author/briluk_d_b/1/)
5. Recognizing and Learning Object Categories / Интернет-ресурс. - Режим доступа: [www/URL: http://people.csail.mit.edu/torralba/shortCourseRLOC/index.html](http://people.csail.mit.edu/torralba/shortCourseRLOC/index.html)