

УДК 004.93.1

Моругов А.М., Волков С.В.

ФГБОУ ВО «Пензенский государственный университет», Пенза, Россия

МЕТОДЫ РАСПОЗНАВАНИЯ СИМВОЛОВ

В статье проанализированы распространенные методы распознавания символов на примере поэтапного процесса распознавания автомобильных номеров. Приведены достоинства и недостатки каждого из методов, предложены варианты применения методов в зависимости от поставленной задачи

Ключевые слова:

распознавание, алгоритмы, анализ, автомобильный номер

Введение

В настоящее время актуальна задача идентификации автомобиля по регистрационному номерному знаку. Использовать систему идентификации можно в различных сферах, касающихся автотранспорта. Примером применения могут послужить контроль въезда на территорию предприятия, автозаправочные станции, автомобильные стоянки, контроль нарушений ПДД и т.п. [10,13].

Процесс идентификации автомобильного номера можно разделить на два этапа: предварительный поиск номера (локализация номерной пластины) и распознавание символов.

Этап 1. Локализация номерной пластины

Существует несколько способов предварительного поиска номера.

Анализ границ и фигур, контурный анализ

Самый очевидный способ выделения номера – поиск прямоугольного контура. Данный способ работает только при наличии ясно читаемого контура – с достаточно высоким разрешением и с ровной границей, при отсутствии преград между камерой и номерной пластиной.

Суть контурного анализа: производится фильтрация изображения для нахождения границ, затем производится выделение всех найденных контуров и их анализ.

Данный способ непрактичен, поскольку контур номерного знака может быть нечетким (грязным, сливаясь с цветом машины, наличие помех и т.д.).

Анализ части границ

Куда интереснее, стабильнее и практичнее представляется подход, где анализируется только часть рамки номерного знака. Выделяются контуры, после чего ищутся все вертикальные прямые. Для любых двух прямых, расположенных недалеко друг от друга, с небольшим сдвигом по оси Y, с правым отношением расстояния между ними к их длине, рассматривается гипотеза того, что номер располагается между ними.

Гистограммный анализ регионов

Одним из самых популярных методов подхода является анализ гистограмм изображения. Подход основывается на предположении, что частотная характеристика региона с номером отлична от частотной характеристики окрестности.

На изображении выделяются высокочастотные пространственные компоненты изображения. Строится проекция изображения на ось y (иногда и на ось x). Максимум полученной проекции может совпасть с расположением номера. У такого подхода есть существенный минус – машина по размеру должна быть сопоставима с размером кадра, т. к. фон может содержать надписи или другие детализированные объекты [1].

Статистический анализ, классификаторы

Недостатки предыдущих методов заключаются в том, что на реальных номерах, запачканных грязью, нет ни выраженных границ, ни выраженной статистики.

Лучшие, но редко используемые методы – это методы, которые опираются на различные классификаторы. Эти методы дают возможность анализировать заданную область на предмет наличия в ней характерных для номера отношений, точек или градиентов и позволяют находить номерной знак в сложных и нетипичных условиях. Например, обученный каскад Хаара. Его работа производится по методу Виолы-Джонса, в основе которого лежат при-

митивы Хаара, представляющие собой разбивку заданной прямоугольной области на наборы разнотипных прямоугольных подобластей:

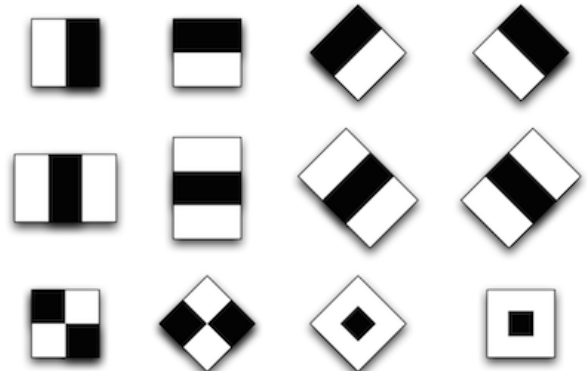


Рисунок 1 – Примеры разбивки прямоугольной области на наборы разнотипных подобластей

В реальных алгоритмах многие методы косвенно или напрямую базируются на наличии границ номера. Даже если при детектировании номера границы не используются, они могут использоваться в дальнейшем анализе.

Случаи с довольно чистым номером в светлой или хромированной рамке на белой машине встречаются очень редко, в сравнении с авто с грязными номерами и их число может оказаться недостаточным для качественного обучения. В таких случаях данный метод неэффективен [2].

Применение Вейвлет-преобразований

В общем случае данный метод состоит из трех частей:

- вейвлет-преобразование изображения;
- фильтрация в вейвлет-домене;
- нахождение и выделение области номера.

Вейвлет-преобразование изображения является крайне необходимым, поскольку оно позволяет увеличить отношение полезного сигнала (номера) к шуму (остальным деталям изображения). Так, например, применения вейвлет-преобразования и последующей фильтрации позволило распознать по рутинным алгоритмам мало контрастное изображение номера автомобиля с включенными фарами.

Для выполнения задачи фиксируется масштаб вейвлет-преобразования в значении, определяемом размерами области номера, частотой чередования черного и белого в этой области и порядком выбранного вейвлета. Например, для гауссова вейвлета 6-го порядка для описания номера достаточно значение масштаба, равное 2.

При исследовании нескольких типов как дискретных, так и непрерывных вейвлет-преобразований, различных семейств и порядков, выяснилось, что более результативными, как и предполагалось, оказались симметричные вейвлеты. Наилучший результат среди дискретных дали вейвлеты семейства койфлетов, а среди непрерывных – гауссовы вейвлеты. Фиксируя порядок каждой строки исходного изображения при выполнении вейвлет-преобразования, мы получаем, таким образом, не двумерный, а одномерный спектр.

Чтобы описать чередование черных и белых пикселей в области номера, соответствующие вейвлет-коэффициенты должны быть достаточно большими. Это отражается в виде характерного плато с крутыми склонами. Данное свойство вейвлет-образа и используется для поиска области номера. Был найден критерий, отражающий наличие плато определенного размера, симметричного и с крутыми

склонами с обеих сторон, так как простое пороговое обрезание по величине вейвлет-коэффициентов привело к тому, что преобразование от изображения решетки радиатора дало большие значения коэффициентов.

Крутизна границ области вейвлет-коэффициентов, упомянутая ранее, соответствующая распознаваемому номеру может быть детектирована по большим значениям второй производной. Вычисление второй производной реализуется с помощью метода бинаризации вейвлет-пикселей с последующим применением специального клеточного автомата (КА) для кластеризации. При оцифровке изображения все пиксели вейвлет-изображения упорядочиваются по уровням серого. Расположенные в упорядоченной последовательности снизу 70% пикселей становятся черными (=0), а оставшиеся 30% – белыми (=1). Выживание каждой белой клетки КА в зависимости от восьми ее соседей организована так, чтобы «выжила» слитная группа единичных пикселей, соответствующая области номера, а окружающие её отдельные единицы вымирали. После применения КА на периферии итогового вейвлет-изображения могут присутствовать малые участки неправильной формы. Эти элементы легко можно вычислить по площади (числу пикселей в них) и отношению длины к ширине [8].

Этап 2. Методы распознавания символов Tesseract OCR

Это метод, использующий открытое программное обеспечение, выполняет автоматическое распознавание символов (букв), а также распознавание текста. Tesseract существует для любых ОС, стабильно работает и легко обучаем – в этом его достоинство. Но он очень плохо работает с битым, грязным и деформированным текстом. При распознавании номеров с помощью данного ПО правильно распознаются порядка 20-30% из существующей базы: самые чистые и прямые.

K-nearest

Очень простой для понимания метод распознавания символов, который, несмотря на свою примитивность, часто может побеждать не самые удачные реализации SVM или нейросетевых методов.

Принцип работы:

- 1) предварительно записывается определенное количество изображений реальных символов, корректно разбитых на классы
- 2) вводится мера расстояния между символами (если изображение бинаризовано, то операция XOR будет оптимальна)
- 3) затем, при попытке распознавания символа, поочередно рассчитывается дистанция между распознаваемым символом и всеми символами в базе. Среди k ближайших соседей, возможно, будут представители различных классов. Выбор класса для символа производится по классности соседних символов. Символ получает такой же класс, как и у большинства соседних символов

В теории, при наличии большой базы с примерами символов, зафиксированных в различных условиях, K-nearest – это все, что нужно. Недостатком может послужить лишь то, что нужно очень быстро рассчитывать дистанцию между изображениями, а, значит, бинаризовать его и использовать XOR, а бинаризация совершенно непредсказуемо изменяет символ. Поэтому в случае с загрязненными или потерянными номерами будут проблемы.

Во многих случаях очень важно понимать, как работает ваш алгоритм. У данного метода есть одно очень важное в этом плане преимущество: он простой и прозрачный, а значит, легко отлаживается и настраивается на оптимальный результат.

Корреляционный метод распознавания

– метод распознавания образов, при котором для каждого класса распознаваемых объектов в декартовом пространстве признаков задается эталонная область, и любой распознаваемый объект относится к классу, соответствующему ближайшей эталонной области. Последняя формируется путем допустимых преобразований одного или нескольких эталонных векторов класса [6].

Шрифт на автомобильном номере известен, шум фотокамеры и/или пыль на номере приближены к гауссовскому. Если изображение не бинаризовано, и амплитуда сигнала неизвестна, соответственно, неизвестна и яркость символа.

Распознавание сводится к операции расчета ковариации входного сигнала с гипотетическим (с учетом заданных смещений и поворотов):

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)],$$

где X – входной сигнал, Y – гипотеза. Обозначение E – математическое ожидание.

При необходимости выбора из различных символов, гипотезы по повороту и смещению строятся отдельно для каждого символа. Если точно известно, что входное изображение содержит символ, то максимум ковариации по всем гипотезам определяет символ, его смещение и наклон. Здесь можно столкнуться с проблемой близости изображений различных символов («р» и «в», «о» и «с» и др.). Самый простой способ решения этой проблемы – введение для каждого символа весовой матрицы коэффициентов.

Иногда такие методы называются «template-matching» (сопоставление шаблонов), что полностью отражает их суть: выбор образцов – затем сравнение входного изображения с образцами. В случае возникновения неопределенности по параметрам, производится перебор всех возможных вариантов, либо применяются адаптивные подходы.

Достоинства метода:

- предсказуемый и хорошо изученный результат, если шум, даже в малой степени, соответствует выбранной модели;
- при строго заданном шрифте, данный метод позволяет распознать сильно пыльный/грязный/потертый символ.

Недостатки:

- немалые затраты на процедуры вычисления.
 - Распознавание скелетных образов
- В первую очередь распознаваемый символ подвергается процедуре скелетизации (утонения). Существуют разнообразные методы получения скелета символа, отличные друг от друга.

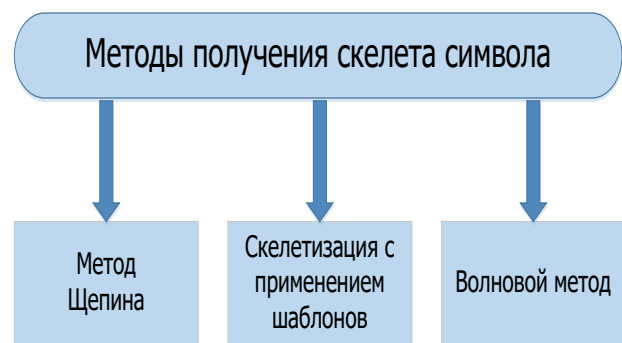


Рисунок 2 – Методы получения скелета символа

Метод Щепина

Состоит в определении исходных верхних левых точек для каждого внешнего и внутреннего контура изображения. Далее рассматривается конфигурация восьми соседей для каждой очередной точки контура. Если точка не оказывается концевой, и после ее удаления соседние точки все еще остаются связным множеством, то точка удаляется. Таким образом анализируется каждая точка и ее соседи, следующая точка для анализа выбирается таким образом, чтобы оставаться на границе изображения. После анализа точки и ее соседей, и возможного удаления точки осуществляется переход к следующей точке контура таким образом, чтобы остаться на границе изображения. Последовательно анализируя каждую точку на границе изображения по слоям, удаляем слои до тех пор, пока не останется множество неудаляемых точек (рисунок 3). Они и составят скелет изображения.

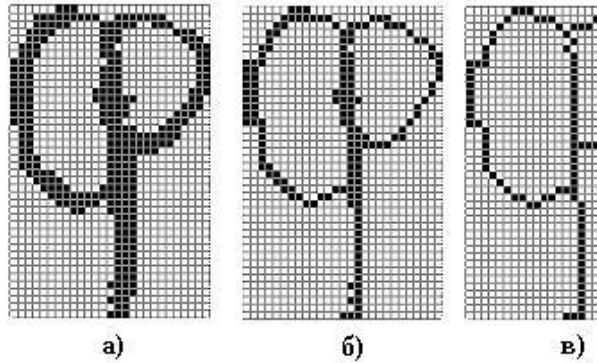


Рисунок 3 – Скелетизация буквы «Ф», состоящей из одного внешнего и двух внутренних контуров. а) исходное изображение; б) удаление одного слоя; в) удаление второго слоя

Скелетизация с применением шаблонов

Существуют шаблоны, которые можно использовать для удаления лишних пикселей и получения скелетного изображения. Знаком «X» в них обозначаются любые цветные пиксели (рисунок 4). Черный пиксель находящийся в центре, удаляется в любой области, соответствующей хотя бы одному из шаблонов. Таким образом делается несколько проходов по изображению – до тех пор, пока не останется пикселей, которые необходимо удалить.

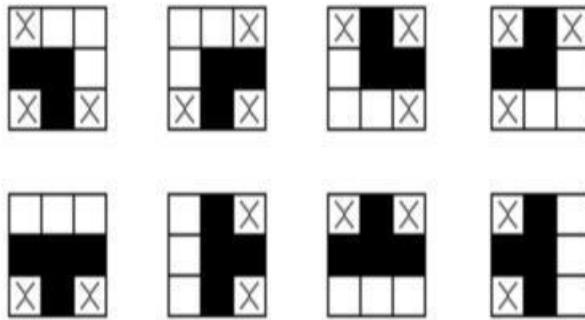


Рисунок 4. Шаблоны скелетизации

Подготовительный этап перед распознаванием – выделение в полученном скелете изображения ключевых точек (точек, в которых соединяются три или четыре ребра скелета, а также концевых точек).

Алгоритм работы.

Создается пустой стек, в котором будут сохранены координаты ребер (начала и конца) и точек ветвления скелета.

2. В созданный стек заносятся указанные выше координаты любой точки скелета.

3. Следующие действия (шаги 4-7) повторяются до тех пор, пока стек не пуст.

4. Выбирается одна из точек, находящихся в стеке.

5. Из выбранной точки строится последовательность ребер, до тех пор, пока не произойдет ветвление скелета, или не будет достигнута концевая точка.

6. В случае достижения концевой точки, либо достижения отмеченного ранее ребра, в массив данных заносится пройденный путь.

7. В случае достижения точки ветвления скелета, найдено место соединения ребер. В массив данных заносится последовательность ребер. В стек заносятся координаты точки ветвления скелета.

8. Переходим к пункту 3.

Полученное описание скелета проходит предварительную обработку, при которой удаляются короткая линия скелета и объединения близких троедов.

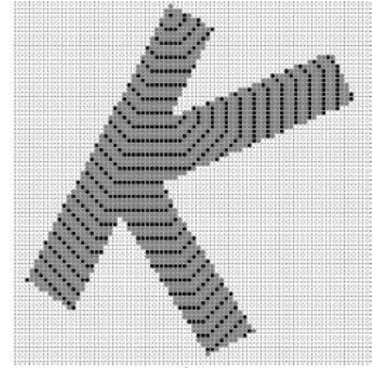
Волновой МЕТОД

В данном методе рассматривается путь прохождения по изображению сферической волны (рисунок 5, а), т.е. пошагово анализируется смещение центра масс точек, которое образует новую генерацию волны, относительно его предыдущих положений.

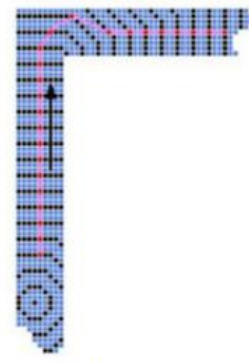
Этапы, из которых состоит метод:

- с помощью сферической волны производится построение скелета изображения;
- производится оптимизация полученного скелета.

Линии изображения отслеживаются путем отслеживания перемещения центра отрезка, образуемого крайними точками генерации волны (рисунок 5, б). После того, как произведено отслеживание линий, возможно сглаживание полученных отрезков.



а)



б)

Рисунок 5. а) Прохождение сферической волны по изображению

б) отслеживание линий изображения

В результате работы метода получается не оптимальный скелет. Основная причина этого в том, что работа ведется с растровым изображением, т.е. с имеющим искажения, которые возрастают по мере уменьшения размера изображения в символах. Поэтому для оптимизации скелета изображения, полученного таким методом, необходима оптимизация. Отрезок изображения в скелете может быть представлен некоторой последовательностью ребер. Оценивая отклонение получающейся линии от прямой в последовательности ребер, можно от этого избавиться. Если отклонение находится в допустимых пределах, последовательность ребер может быть заменена на одно ребро.

В процессе оптимизации скелета изображения просматриваются точки, в которых происходит разделение волны на полуволны (т.е. окрестности выделенных точек соединения отрезков) [4].

Адаптивное распознавание

Существует два подхода:

Шрифтовый (шрифтозависимый)

Безшрифтовый (шрифтонезависимый)

Первый подход.

Программа измеряет и анализирует различные характеристики шрифта и заносит их в свою базу эталонных характеристик. По окончании этого процесса шрифтовая программа оптического распознавания символов (ОРС) готова к распознаванию данного конкретного шрифта. Этот процесс условно можно назвать обучением программы. Далее обуче-

ние повторяется для некоторого множества шрифтов, которое зависит от области применения программы.

К недостаткам данного подхода можно отнести следующие факторы:

- алгоритм должен заранее знать шрифт;
- для работы программы распознавания необходим блок настройки на конкретный шрифт.

С другой стороны, у шрифтового подхода есть существенное преимущество: имея детальную априорную информацию о символах, можно построить достаточно точные и надежные алгоритмы распознавания.

Второй подход.

Эти алгоритмы измеряют и анализируют различные характеристики (признаки), присущие буквам как таковым безотносительно шрифта и абсолютного размера (кегля), которым они напечатаны.

К недостаткам данного подхода можно отнести следующие факторы:

- качество распознавания существенно ниже, чем у шрифтовых алгоритмов;
- низкие значения коэффициента надежности распознавания.

Основные достоинства:

- универсальность: алгоритм одинаково работает с разными шрифтами
- технологичность: простота обучения; возможность реализовать автоматизированное обучение [5]

Нейронные сети

Искусственная нейронная сеть (ИНС, нейронная сеть) – это набор нейронов, соединенных между собой. Работа нейронной сети состоит в преобразовании входного вектора в выходной вектор, причем это преобразование задается весами нейронной сети.

Практически любую задачу можно свести к задаче, решаемой нейронной сетью [9].

Для того, чтобы нейронная сеть начала работать, ее необходимо обучить. Для этого в нейроны записывается определенный эталон, затем подается входной сигнал, который сравнивается с эталоном и различие между эталоном и входным воздействием фиксируется (добавляется к эталону). Чем больше будет входных воздействий, тем лучше обучится нейросеть.

Достоинства:

- при правильной настройке и обучении может работать лучше других методов;
- при большом обучающем массиве данных появляется устойчивость к искажениям символов.

Недостатки:

- сложность;

– в многослойных сетях невозможна диагностика аномального поведения.

Необходимо отметить, что для решения задачи распознавания существуют уже готовые библиотеки распознавания автомобильных номеров. Рассмотрим некоторые из них.

Проект Opos

Реализован на C# [3]. Для распознавания номера используется порт библиотеки компьютерного зрения OpenCV на C#. Для распознавания символов номера используется обертка на C# над библиотекой OCR Cuneiform – Puma.NET. Недостаток реализации заключается в затруднении использования библиотеки в системах реального времени, т.к. изображение сначала сохраняется на диск, а потом уже передается в Puma.NET.

Проект JavaANPR

Релизован на Java. Преимуществом данной библиотеки является то, что она подходит для любых ОС. Все алгоритмы написаны на Java без использования нативных библиотек, что сильно упрощает использование. Данную библиотеку можно использовать в системах реального времени, так как она имеет достаточно высокую скорость распознавания: 0,2–0,8 сек. на один номерной знак. Недостатком можно назвать то, что в случае зашумленных изображений или изображений с плохой освещенностью библиотека не всегда дает хороший результат [3].

Также возможно реализовать систему распознавания в среде графического программирования National Instruments (NI) LabView с помощью платформы машинного зрения NI Vision, которая состоит из технологии захвата изображения IMAQ (Image Acquisition) и программной технологии его обработки и анализа [7]. С помощью IMAQ можно работать с изображениями любого спектрального диапазона, с разной скоростью ввода, с различной глубиной оцифровки, т.к. большей частью функций модулей захвата изображения можно управлять программно, а это упрощает ввод изображения с различных видеопреобразователей [11,12].

Заключение

Сравнивая представленные методы и алгоритмы распознавания автомобильных номеров можно сказать, что самым действенным методом локализации рамки автомобильного номера является каскад Хаара. Если же рассматривать методы и способы распознавания символов, то здесь преимущество над всеми будет иметь нейронная сеть, так как нейронная сеть может обучаться в процессе работы. Недостатками данных методов является сложность реализации.

ЛИТЕРАТУРА

1. <https://habrahabr.ru/company/recognitor/blog/221891/>
2. <https://habrahabr.ru/company/recognitor/blog/228195/>
3. <https://www.ibm.com/developerworks/ru/library/os-avto/>
4. Лукошенко Г. Н. Распознавание скелетных образов. [Электронный ресурс]. – <http://www.ocrai.narod.ru/skeletonrecognize.html>
5. Арлазаров В. Л., Троянker В.В., Котович Н.В. Адаптивное распознавание символов. [Электронный ресурс]. – <http://www.ocrai.narod.ru/adaptive.html>.
6. Глушков В.М., Амосов Н.М., Артеменко И.А. Энциклопедия кибернетики. Том 1. Киев, 1974 г.
7. 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.
8. Стадник А.В. Диссертация. «Использование искусственных нейронных сетей и вейвлет-анализа для повышения эффективности в задачах распознавания и классификации». Ивановский Государственный университет. Иваново, 2004г.
9. Саймон Хайкин. Нейронные сети. Полный курс. – М.: Вильямс, 2005. – 1104 стр.
10. Волков С.В., Колдов А.С., Захарова О.В., Чапаев В.С. Система автоматического контроля и управления параметрами объекта. // Надежность и качество: Труды международного симпозиума: в 2-х т. / Под ред. Н.К. Юркова. – Пенза: Изд-во ПГУ, 2014 – 2-т., стр. 91–94.
11. Бободжонов Д.А., Волков С.В., Чапаев В.С. Виртуальный прибор для систем видеомониторинга. // Надежность и качество: Труды международного симпозиума: в 2-х т. / Под ред. Н.К. Юркова. – Пенза: Изд-во ПГУ, 2012 – 2-т., стр. 95–99.
12. Волков С.В., Рафиков Р.Р. Система видеомониторинга. // Надежность и качество: Труды международного симпозиума: в 2-х т. / Под ред. Н.К. Юркова. – Пенза: Изд-во ПГУ, 2013 – 2-т., стр. 247–249.
13. Современные технические решения и проблемы в обеспечении комплексной безопасности Волков С.В., Кулапин В.И., Светлов А.В. Надежность и качество сложных систем. 2014. № 4 (8). С. 61–68.