

УДК 004.4

ПРИМЕНЕНИЕ ПРЕОБРАЗОВАНИЯ ХАФА ДЛЯ РАСПОЗНАВАНИЯ ТЕКСТА*Гайдуков Н.П., Савкова Е.О.**Донецкий национальный технический университет**кафедра автоматизированных систем управления**E-mail: gaydukov.nikita@gmail.com*

Был рассмотрен алгоритм преобразования Хафа, как один из основных инструментов для распознавания рукописного текста. Также были приведены краткие характеристики методов распознавания, был выбран оптимальный для использования в магистерской работе.

Введение

Преобразование Хафа позволяет находить на монохромном изображении плоские кривые, заданные параметрически, например: прямые, окружности, эллипсы, и т.д. Монохромным изображением считается изображение, состоящее из точек двух типов: фоновых точек и точек интереса. Задача преобразования Хафа состоит в выделении кривых, образованных точками интереса.

Преобразование Хафа

Это метод по извлечению элементов из изображения, используемый в анализе, обработке изображения и компьютерном зрении. Данный метод предназначен для поиска объектов, принадлежащих определённому классу фигур с использованием процедуры голосования. Процедура голосования применяется к пространству параметров, из которого и получают объекты определённого класса фигур по локальному максимуму в, так называемом, накопительном пространстве (accumulator space), которое строится при вычислении трансформации Хафа.

При автоматизированном анализе цифровых изображений очень часто возникает проблема определения простых фигур, таких как прямые, круги или эллипсы. Во многих случаях используется алгоритм детектирования границ в качестве предобработки для получения точек, находящихся на кривой в изображении. Однако, либо из-за зашумлённости изображения, либо из-за несовершенства алгоритма детектирования границ могут появиться «потерянные» точки на кривой, так же как и небольшие отклонения от идеальной формы прямой, круга или эллипса. По этим причинам часто довольно сложно сгруппировать выделенные границы в соответствующий набор прямых, кругов и эллипсов. Назначение преобразования Хафа в том, чтобы разрешить проблему группировки граничных точек путём применения определённой процедуры голосования к набору параметризованных объектов изображения.

В простейшем случае преобразование Хафа является линейным преобразованием для обнаружения прямых. Прямая может быть задана уравнением $y = mx + b$ и может быть вычислена для любой пары точек на изображении (x, y) . При преобразовании Хафа главная идея – учесть характеристики прямой не как точек изображения, а в терминах её параметров, то есть m – коэффициента наклона и b – точки пересечения. Основываясь на этом факте, прямая $y = mx + b$ может быть представлена в виде точки с координатами

(b, m) в пространстве параметров. Однако, есть одна проблема в том, что вертикальные прямые имеют бесконечные значения для параметров m и b [1][2]. Соответственно для удобства вычислений лучше всего представить прямую с помощью других параметров, более известных как r и θ (тета). Параметр r представляет собой длину радиус-вектора ближайшей к началу координат точки на прямой, а θ – это угол между этим вектором и осью координат. Таким образом, уравнение прямой можно записать так:

$$y = \left(-\frac{\cos \theta}{\sin \theta}\right)x + \left(\frac{r}{\sin \theta}\right), \quad (1)$$

что может быть преобразовано в $r = x \cos \theta + y \sin \theta$.

Поэтому возможно связать с каждой прямой на изображении пару (r, θ) которая является уникальной при условии если $\theta \in [0, \pi]$ и $r \in \mathbb{R}$, или если $\theta \in [0, 2\pi]$ и $r \leq 0$. Плоскость (r, θ) иногда называется Пространство Хафа для набора прямых 2-мерном случае. Это представление делает трансформацию Хафа концептуально очень близкой к 2-мерному преобразованию Радона.

Бесконечное число прямых может проходить через одну точку плоскости. Если эта точка имеет координаты (x_0, y_0) в изображении, то все прямые, проходящие через неё соответствуют следующему уравнению:

$$r(\theta) = x_0 \cdot \cos \theta + y_0 \cdot \sin \theta. \quad (2)$$

Это соответствует синусоидальной кривой в (r, θ) пространстве, которая, в свою очередь, уникальна для данной точки. Если кривые соответствующие двум точкам накладываются друг на друга, то точка (в пространстве Хафа) где они пересекаются, соответствует прямой (в оригинальном месте изображения), которые проходят через обе точки. В общем случае, ряд точек, которые формируют прямую линию, определяют синусоиды, которые пересекаются в точке параметров для той линии. Таким образом, проблема обнаружения коллинеарных точек может быть сведена к проблеме обнаружения пересекающихся кривых.

Словесное описание алгоритма

Алгоритм преобразования Хафа использует массив, называемый аккумулятором, для определения присутствия прямой $y = mx + b$. Размерность аккумулятора равна количеству неизвестных параметров пространства Хафа. Например, для линейной трансформации нужно использовать двумерный массив, так как имеются два неизвестных параметра: m и b . Два измерения аккумулятора соответствуют квантизированным значениям параметров m и b . Для каждой точки и её соседей алгоритм определяет достаточен ли вес границы в этой точке. Если да, то алгоритм вычисляет параметры прямой и увеличивает значение в ячейке аккумулятора, соответствующей данным параметрам.

Потом, найдя ячейки аккумулятора с максимальными значениями, обычно поиском локального максимума в пространстве аккумулятора, наиболее подходящие прямые могут быть определены. Самый простой способ – это пороговая фильтрация. Однако в разных ситуациях разные методы могут давать разные результаты. Так как полученные прямые не содержат информацию о длине, следующим шагом является нахождение частей изображения соответствующих найденным прямым. Более того, из-

за ошибок на этапе детектирования границ в пространстве аккумулятора также будут содержаться ошибки. Это делает поиск подходящих линий нетривиальным.

Методы распознавания

Далее применяется один из методов распознавания. Система распознавания реализуется как классификатор. Существуют несколько методов реализации классификаторов:

- шаблонные (растровые);
- признаковые;
- структурные.

В классификаторе первого типа с помощью критерия сравнения определяется, какой из шаблонов выбрать из базы. Самый простой критерий - минимум точек, отличающих шаблон от исследуемого изображения.

Достоинства: К достоинствам шаблонного классификатора относятся хорошее распознавание дефектных символов («разорванных» или «склеенных»), простота и высокая скорость распознавания.

Недостатки: Недостатком является необходимость настройки системы на типы и размеры шрифтов.

Наиболее распространены признаковые классификаторы. Анализ в них проводится только по набору чисел или признаков, вычисляемых по изображению. Таким образом, происходит распознавание не самого символа, а набора его признаков, т. е. производных данных от исследуемого символа.

Достоинства: Этот метод позволяет распознавать различные начертания символов, т.е. различные подчерки шрифты и т.д.

Недостатки: Этот метод неизбежно вызывает некоторую потерю информации.

Структурные классификаторы переводят шейп символа в его топологическое представление, отражающее информацию о взаимном расположении структурных элементов символа. Эти данные могут быть представлены в графовой форме.

Достоинства: Такой способ обеспечивает инвариантность относительно типов и размеров шрифтов.

Недостатки: Недостатками являются трудность распознавания дефектных символов и медленная работа.

Структурно-пятенный метод

Однако наиболее оптимальным среди всех методов является так называемый структурно-пятенный эталон и его фонтанное (от англ. font - шрифт) представление. Оно имеет вид набора пятен с попарными отношениями между ними. Подобную структуру можно сравнить со множеством шаров, нанизанных на резиновые шнуры, которые можно растягивать. Данное представление нечувствительно к различным начертаниям и дефектам символов.

Алгоритм основан на сочетании шаблонного и структурного методов распознавания образов. При анализе образца выделяются ключевые точки объекта – так называемые «пятна».

В качестве пятен, например, могут выступать:

- концы линий;
- узлы, где сходятся несколько линий;
- места изломов линий;
- места пересечения линий;
- крайние точки.

После выделения «пятен» определяются связи между ними – отрезок, дуга. Таким образом, итоговое описание представляет собой граф, который и служит объектом поиска в библиотеке «структурно-пятенных эталонов».

При поиске устанавливается соответствие между ключевыми точками образца и эталона, после чего определяется степень деформации связей, необходимая чтобы привести искомый объект к сравниваемому эталонному образцу. Меньшая степень необходимой деформации предполагает большую вероятность правильного распознавания символа.

Заключение

Эффективность преобразования Хафа в большой степени обусловлена качеством входных данных: границы должны быть четко определены. Использование преобразования Хафа на зашумленных изображениях затруднено. Для зашумленных изображений необходим этап предобработки с целью подавления шума.

Перечень источников

- [1] Абраменко А. Принципы распознавания // Компьютер-пресс. 1997. – № 12.
- [2] Загоруйко Н.Г. Методы распознавания и их применение. – М.: Сов. радио, 1972. – 208 с.
- [3] Линдсней П., Норман Д. Переработка информации у человека: Пер. с англ. – М.: Мир, 1974. – 550 с.
- [4] Шамис А.Л. Принципы интеллектуализации автоматического распознавания изображений и их реализация в системах оптического распознавания символов // Новости искусственного интеллекта. 2000.
- [5] Cognitive Forms – система массового ввода структурированных документов [Электронный ресурс] / В.В. Арлазаров, В.В. Постников, Д.Л. Шоломов.
- [6] Уинстон П. Искусственный интеллект / П. Уинстон. – М.: Мир, 1980. – 519 с.
- [7] Хант Э. Искусственный интеллект / Э. Хант. – М.: Мир, 1978. – 558 с.
- [8] Логический подход к искусственному интеллекту / под ред. Г.П. Гаврилова – М.: Мир, 1998. – 493 с.
- [9] Эндрю А. Искусственный интеллект / Под ред. Поспелова Д.А. – М.: Мир, 1985. – 265 с.
- [10] Дарья Калинкина, Дмитрий Ватолин. Проблема подавления шума на изображениях и видео и различные подходы к ее решению // Компьютерная графика и мультимедиа : журнал. — 2005. — № 3 (2).

-
- [11] Распознавание рукописных текстов, Ижевск 2006, А.В. Кучуганов, Г.В. Лапинская
- [12] Экспертные системы. Учебное пособие для вузов, Воронеж 2007, А.Ю. Телков
- [13] G.A.Carpenter and S. Grossberg, Pattern Recognition by SelfOrganizing Neural Networks, MIT Press, Cambridge, Mass., 1991.
- [14] “The First Census Optical Character Recognition System Conference”, R.A.Wilkinson et al., eds., . Tech. Report, NISTIR 4912, US Deop. Commerce, NIST, Gaithersburg, Md., 1992.
- [15] J. Hertz, A. Krogh, and R.G. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, Reading, Mass., 1991.
- [16] S. Haykin, Neural Networks: A Comprehensive Foundation, MacMillan College Publishing Co., New York, 1994.
- [17] G.A.Carpenter and S. Grossberg, Pattern Recognition by SelfOrganizing Neural Networks, MIT Press, Cambridge, Mass., 1991.
- [18] “The First Census Optical Character Recognition System Conference”, R.A.Wilkinson et al., eds., . Tech. Report, NISTIR 4912, US Deop. Commerce, NIST, Gaithersburg, Md., 1992.