

УДК 004.932.2

СОВРЕМЕННЫЕ МЕТОДЫ АВТОМАТИЧЕСКОГО ОБНАРУЖЕНИЯ ПРЯМОУГОЛЬНЫХ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Е.С. МАТУСЕВИЧ, И.Э. ХЕЙДОРОВ

*Белорусский государственный университет
Курчатова, 5, Минск, 220108, Беларусь**Поступила в редакцию 20 октября 2016*

Рассмотрены низкоуровневые и высокоуровневые методы выделения признаков на изображении и алгоритмы формирования прямоугольного объекта. Были исследованы и протестированы алгоритм обнаружения объекта на основе корреляционного анализа, а также алгоритм, содержащий в себе применение детектора границ Канни, обнаружение линий с помощью преобразования Хафа и преобразования Радона и далее, в зависимости от свойств объекта, объединением линий в прямоугольную область. Алгоритмы были протестированы на базе из 1000 паспортов для задачи точного обнаружения границ фотографии.

Ключевые слова: преобразование Хафа, преобразование Радона, коэффициент корреляции, прямоугольный объект, детектор границ Канни.

Введение

Обнаружение прямоугольных объектов является одной из базовых задач компьютерного зрения и используется для локализации зданий [1], номерных знаков [2], краев отсканированного документа [3].

Существует несколько категорий выделения признаков в области обработки изображений. Детекция границ и углов – примеры низкоуровневых техник выделения признаков, в то время как преобразование Хафа и Радона является примером высокоуровневых техник извлечения признаков на основе формы объекта. Как правило, детекторы границ дают удовлетворительный результат при наличии резких локальных изменений яркости. Однако широкий класс систем технического зрения эксплуатируется в условиях наличия шумов и помех. Шумы существенным образом искажают информацию о положении краевых точек, что приводит к появлению двух типов возможных ошибок. Первый тип связан с пропуском реально существующей граничной точки, а второй – с появлением ложной.

В большинстве работ для обнаружения прямоугольного объекта используется преобразование Хафа [4]. Этот метод быстрый и показывает хорошие результаты обнаружения линий, даже если границы объекта прерываются из-за шумов, однако результатом преобразования Хафа является набор бесконечных линий, граничные точки которых неизвестны и необходим дополнительный анализ, чтобы объединить линии в прямоугольный объект. В настоящее время не существует общепринятого алгоритма объединения линий. Полученные линии объединяются в объект, в зависимости от поставленной задачи. Например, на изображении необходимо обнаружить один самый представительный прямоугольный объект или все прямоугольные объекты с заданными характеристиками размером, цветом, ориентацией. Таким образом, можно констатировать, что проблема детекции прямоугольных объектов с высокой надежностью и скоростью остается актуальной.

Цель работы: исследование методов автоматического обнаружения прямоугольных объектов на изображении и выбор наиболее подходящего метода для обнаружения границ прямоугольного объекта фотографии. Авторам не удалось найти аналогов работ по

детектированию границ фотографии для сравнения результатов, так как это коммерчески закрытая информация.

Выделение границ изображения детектором Канни

Детекция границ является широко используемым методом для обнаружения признаков объекта. Она используется для идентификации точек на изображении с перепадами яркости. Основная цель выделения границ – уменьшить сложность изображения, сохраняя его основную структуру. Детектор границ Канни является одним из наиболее часто используемых детекторов границ [5]. Хотя работа Канни [6] была проведена на заре эры компьютерного зрения, детектор границ Канни до сих пор является одним из лучших детекторов. Алгоритм состоит из пяти отдельных шагов [7], описанных ниже.

На первом этапе производится сглаживание фильтром Гаусса, который обнаруживает и устраняет найденные разрывы, применяя перемещаемую по изображению маску. Далее рассчитываются градиенты с помощью дискретного дифференциального оператора Собеля. Затем применяется подавление не-максимумов, где пикселями границ объявляются пиксели, в которых достигается локальный максимум градиента в направлении вектора градиента. В результате подавления локальных неопределенностей толщина линии границы становится равномерной и тонкой, что увеличивает точность ее расположения. Следующий шаг – двойная пороговая фильтрация, то есть, если значение пикселя выше верхней границы порога, то он принимает максимальное значение (граница считается достоверной), если ниже – пиксель подавляется, точки со значением, попадающим в диапазон между порогами, принимают фиксированное среднее значение. Последний этап – трассировка области неоднозначности. Задача сводится к выделению групп пикселей, получивших на предыдущем этапе промежуточное значение, и отнесению их к границе (если они соединены с одной из установленных границ) или их подавлению (в противном случае). Пиксель добавляется к группе, если он соприкасается с ней по одному из восьми направлений.



Рис. 1. Пример входного изображения и результата после фильтра Канни

Таким образом, применение фильтра выделения границ (рис. 1) к изображению может существенно уменьшить количество обрабатываемых данных. Однако, применение алгоритма Канни в базовом варианте не всегда может давать те результаты, которые нужны для решения поставленной задачи. Одним из существенных недостатков алгоритма Канни является необходимость настройки пороговых значений детектора.

Обнаружение прямых линий на изображении с помощью преобразования Хафа

Общепринятым подходом для анализа и понимания структуры цифровых изображений является поиск простых объектов, таких как прямые. Предпосылкой для такого анализа является то, что точки, которые могут образовывать такие объекты, были извлечены, как правило, с использованием детектора границ. Задача сгруппировать полученные точки в соответствующие объекты, например, прямые, может быть решена с помощью преобразования Хафа. Его идея состоит в поиске кривых, которые проходят через достаточное количество точек интереса [8].

Поиск прямых линий. Прямую на плоскости можно задать следующим уравнением $x \cos \theta + y \sin \theta = R$, где R – длина перпендикуляра, опущенного на прямую из начала координат, θ – угол между перпендикуляром к прямой и осью абсцисс.

Введем разбиение в пространстве (R, θ) , где любая ячейка будет включать в себя набор прямых с близкими значениями R и θ . Во время выбора размера ячеек стоит принять во

внимание следующие факты. Если размер ячейки будет слишком большой, то за прямую линию может приниматься разрозненное множество точек. Если же наоборот, размер ячейки будет слишком мал, то существует вероятность, что ни одна прямая линия не будет найдена - все счетчики будут иметь небольшое значение.

К достоинствам метода можно отнести довольно высокую точность обнаружения объектов на изображении, а также высокую универсальность данного подхода – достаточно лишь выразить искомый геометрический объект в виде набора параметров. Однако из этого вытекает основной недостаток метода – чрезвычайная вычислительная сложность алгоритма, растущая в зависимости от количества параметров, поскольку алгоритм подразумевает перебор всех возможных вариантов комбинаций параметров искомой фигуры. Для увеличения быстродействия были разработаны некоторые модификации стандартного алгоритма [9].

Обнаружение линий на изображении с помощью преобразования Радона

Преобразование Радона [10] $R(k, b)$ непрерывной функции $f(x, y)$ вычисляется путем интегрирования значений f вдоль наклонной линии. Можно записать преобразование при помощи δ -функции Дирака:
$$R(k, b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(y - kx - b) dx dy.$$

Необходимо отметить, что преобразование Радона обладает некоторыми свойствами, очень важными для работы с изображениями, такими как свойство линейности, сдвига и масштабирования. Однако, интегральное преобразование имеет ограничение, связанное с поиском прямых на изображении, угол наклона которых равен или близок 90 градусов к оси абсцисс. Данное обстоятельство связано с бесконечно большим параметром k . Для того, чтобы избавиться от данной проблемы, можно перейти к заданию линии в нормальной форме:

$$R(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(\rho - x \cos \theta - y \sin \theta) dx dy.$$

Одномерное преобразование Фурье по переменной s от преобразования Радона функции f дает двумерное преобразование Фурье функции f .
$$F(w \cos \alpha, w \sin \alpha) = \int_{-\infty}^{\infty} ds e^{-iws} R(s, \alpha).$$

Таким образом, преобразование Радона можно вычислить через преобразование Фурье, что дает выигрыш производительности в ситуации, когда точек очень много. К тому же его можно применять к небинаризованному изображению в отличие от преобразования Хафа.

Обнаружение объекта на изображении с помощью корреляционного анализа

Одним из распространенных алгоритмов для поиска объектов является сравнение с эталоном [11] на основе корреляционного анализа. При полностью заданном эталоне многошаговая корреляция путем сканирования входного поля зрения является, по сути, полным перебором в пространстве сигналов. Поэтому эту процедуру можно считать базовой, потенциально наиболее помехоустойчивой, хотя и самой трудоемкой.

В данном подходе есть два изображения, первое изображение $I(x, y)$, а второе – $T(x', y')$ – назовем эталоном. Обычно эталон должен быть меньше размером, чем исследуемое изображение. Значение коэффициента корреляции изображения с эталоном в точке (x, y) исходного изображения определяется по формуле:
$$R_{coeff}(x, y) = \sum_{x', y'} T'(x', y') I'(x + x', y + y'),$$
 где

T' – эталон, полученный при вычитании из него математического ожидания, I' – изображение, полученное при вычитании из него математического ожидания $T'(x', y') = T(x', y') - \bar{T}$.

Таким образом, алгоритм поиска объекта по эталону состоит из следующих основных шагов: разбиение начального изображения на расчетные области (области интереса) равного размера, расчет карты коэффициентов корреляции и нахождение максимального отклика на карте. В качестве эталонов для поиска объекта можно выбрать 4 прямых угла (рис. 2), которые

являются более уникальными примитивами в отличие от линий. За эталон можно выбрать также и весь прямоугольный объект, но на реальных изображениях углы могут быть зашумлены другими объектами, поэтому поиск отдельно каждого угла – наиболее стабильный вариант поиска прямоугольного объекта.



Рис. 2. Эталоны для каждого из углов прямоугольного объекта

Недостатком данного метода является медленная скорость работы по сравнению с преобразованием Хафа и неинвариантность к углу разворота объекта, а главное достоинство в том, что результатом метода являются точные координаты углов объекта, в отличие от бесконечных линий преобразования Хафа, которые еще необходимо объединить в объект. Также достоинство данного метода в том, что для обнаружения прямоугольного объекта достаточно найти только два противоположных угла.

Обнаружение границ на изображении с помощью нейронных сетей

Искусственная нейронная сеть (ИНС) представляет собой математическую, программную или аппаратную модель, функционирующую по принципу биологических нервных клеток человеческого организма [12]. Искусственные нейронные сети имеют иерархическую структуру и состоят из простых взаимодействующих между собой процессоров, называемых нейронами. Текущее значение нейрона определяется как взвешенная сумма его входных значений: $S = \sum_{i=1}^N x_i w_{ij}$, где x_i – текущий выход i -го нейрона; w_{ij} – весовой коэффициент между i -м и j -м нейронами.

Каждый нейрон определяется функцией активации, которая предназначена для определения зависимости сигнала на выходе нейрона от взвешенной суммы сигналов на его входах. Пример стандартной логистической функции активации: $f(x) = \frac{1}{1 + e^{-s}}$, где s – взвешенная сумма входных значений нейрона.

Для всех данных из обучающей выборки, поступающих на вход нейронной сети, определяются их требуемые значения, которые затем сравниваются с текущими значениями и вычисляется ошибка. Основным принцип обучения ИНС заключается в сведении к минимуму эмпирической ошибки классификации [13]. После этого на основе метода градиентного спуска вычисляется величина коррекции синаптических коэффициентов между нейронами. Весовые коэффициенты изменяются до тех пор, пока ошибка не сведется к минимуму и станет меньше заданного порога. Данный тип обучения называется «обучение с учителем».

Нейронные сети часто используются для задач детекции краев объектов [14]. Для таких задач используются так называемые сверточные нейронные сети (СНС), представляющие собой особый класс многослойного персептрона, который обладает двумерной структурой и хорошо подходит для обработки изображений с высокой степенью инвариантности к смещению, поворотам, масштабированию и другим искажениям входных данных. Они показывают более эффективные результаты по сравнению с детектором краев Канни, однако для них необходимо обучение на базе позитивных и негативных образцов, а также скорость их работы в разы медленнее стандартного алгоритма Канни. К тому же СНС отличается сложностью настройки оптимальных параметров: количество слоев, плоскостей, нейронов и т.д.

Таким образом, для задачи обнаружения границ фотографии на различных паспортах в условиях зашумленности углов и(или) сторон защитными элементами, обучение нейросети и дальнейшее обнаружение объекта фотографии не является целесообразным.

Алгоритмы обнаружения прямоугольного объекта на изображении

Один из подходов для обнаружения прямоугольного объекта на примере отсканированного документа описан в статье [15]. Алгоритм состоит из следующих этапов: детектирование границ на изображении с помощью фильтра Канни, далее ведется поиск

прямых линий с помощью преобразования Хафа или Радона, затем производится поиск угловых точек и в итоге строится прямоугольная область, описывающая объект.

Первые два этапа были описаны в предыдущих частях. На следующем шаге из линий извлекаются угловые точки. Угловая точка – такая точка, где две обнаруженные линии в изображении пересекаются перпендикулярно (или близки к перпендикуляру). Еще одно требование к такой точке, чтобы линии были уверенно найденными, то есть их отклик был выше установленного порогового значения. Следующий этап – построение области, ограничивающей все угловые точки. На конечном шаге необходимо устранить точки пересечения, которые не соответствуют свойствам углов документа. Для этого анализируется средняя яркость окрестности каждой точки интереса.

С помощью алгоритма сравнения объекта с эталоном на основе корреляционного анализа [16] можно обнаружить прямоугольный объект, предварительно сформировав эталон каждого из углов объекта. Предложенный подход состоит из следующих этапов: формирование эталонов для каждого из 4 углов (рис. 2), далее ведется обнаружение хотя бы 2 противоположных углов, затем строится прямоугольный объект по полученным координатам. Такой алгоритм является достаточно затратным по времени, а также его недостатком является чувствительность к углу разворота документа. Для увеличения скорости обработки, изображение необходимо уменьшить до оптимального размера, в котором достигается компромисс скорость/качество.

Экспериментальные результаты

Исследованные алгоритмы были протестированы на базе из 1000 изображений паспортов, полученных с помощью аппаратного-программного комплекса «Регула» 70X4. На вход подавалась предполагаемая область нахождения объекта фотографии владельца паспорта размером 400×600 пикселей с целью точной локализации границ фотографии (см. таблицу). Углы фотографий были вручную промаркированы на базе изображений, что позволило получить статистику вероятностей правильной детекции границ фотографии. Эксперименты проводились на компьютере с ядром Intel Core 2 Duo и частотой процессора 2,93 ГГц.

Результаты алгоритмов детекции прямоугольного объекта фотографии

Название алгоритма	Вероятность правильной детекции границ фотографии	Время работы алгоритма, мс
Алгоритм на основе стандартного преобразования Хафа	92,1 %	62
Алгоритм на основе вероятностного преобразования Хафа	91,5 %	53
Алгоритм на основе преобразования Радона	92,7 %	67
Алгоритм на основе корреляционного анализа	96,4 %	94

Для первого алгоритма были протестированы стандартное и вероятностное преобразование Хафа, а также преобразование Радона. Подход, основанный на преобразовании Хафа, имеет высокую скорость работы. Однако он дает большее количество ложных срабатываний, так как линий на изображении может быть много и не всегда можно точно обнаружить край фотографии, также этот подход требует дополнительного поиска точек пересечения полученных линий, из которых уже формируется прямоугольный объект, а ложные прямоугольные объекты фильтруются по размеру. Также присутствие на краях фотографий защитных элементов у некоторых типов паспортов приводит к ошибкам обнаружения фотографии. Преобразование Радона работает немного точнее и медленнее, но в целом дает такой же результат, что и преобразование Хафа. Второй подход является более помехоустойчивым и точным. Одним из достоинств алгоритма является то, что если обнаружены 2 противоположных угла объекта, то можно не искать остальные, а на основе координат найденных углов построить границы прямоугольного объекта. Однако если фотография наклеена неаккуратно под углом, это приводит к ошибкам обнаружения объекта, а также данный подход более медленный по сравнению с преобразованием Хафа.

Заключение

Был исследован ряд низкоуровневых и высокоуровневых техник для обнаружения прямоугольных объектов. Общепринятым подходом является применение детектора Канни для

выделения границ объекта и уменьшения сложности изображения. Далее на бинарном изображении с помощью преобразования Хафа или Радона можно обнаружить линии на изображении и построить прямоугольный объект. Еще один подход – обнаружение объекта с помощью корреляционного анализа. Предложенные алгоритмы были протестированы на базе паспортов с целью обнаружения прямоугольного объекта фотографии. Вероятностное преобразование Хафа работает быстрее всего и дает результат правильной детекции выше, чем стандартное преобразование Хафа, точнее всего работает преобразование Радона, но оно медленнее. Также существует подход на основе нейростей, но для него требуется предварительное обучение, и также необходим этап дальнейшего объединения найденных линий в прямоугольный объект. Подход на основе корреляционного анализа дает самый высокий результат правильной детекции, однако он неустойчив к углу разворота документа и работает медленнее, чем преобразование Хафа.

MODERN METHODS OF AUTOMATIC RECTANGLE OBJECTS DETECTION

E.S. MATUSEVICH, I.E. KHEIDOROV

Abstract

Low-level and high-level feature extraction methods and algorithms for the image formation of a rectangular object were considered. The algorithm for object detection based on correlation analysis, as well as the algorithm containing the use of Canny edge detector, Hough and Radon transform for lines detection, and then, depending on the properties of the object lines combining in the rectangular area, were explored. The algorithms were tested on the base of 1000 passports for the problem of accurate photo edges detection.

Keywords: Hough transform, Radon transform, correlation coefficient, rectangular object, Canny edge detector.

Список литературы

1. *Noronha S., Nevatia R.* Detection and modeling of buildings from multiple aerial images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001.
2. *Юзов М.В., Пугачёв Е.К.* // Инж.вестн. № 12. 2014.
3. *Guerzhoy M., Zhou H.* Segmentation of Rectangular Objects Lying on an Unknown Background in a Small Preview Scan Image. IEEE Proceedings of Conference on Computer and Robot Vision, 2008.
4. *Jung C.R., Schramm R.* // SIBGRAPI '04 Proceedings of the Computer Graphics and Image Processing. Brazil, 2004.
5. *Nadernejad E., Sharifzadeh S., Hassanpour H.* // Edge detection techniques: Evaluations and comparisons. Applied Mathematical Sciences, 2008.
6. *Canny J.* Finding edges and lines in images. Cambridge, 1983.
7. *Шух О.Г., Рассомахин С.Г.* // Системы обработки информации. 2012. Вып. 3(1). С. 85–88.
8. *Дегтярева А., Вежневцев В.* // Компьютерная графика и мультимедиа. 2003. № 1 (2).
9. *Жданов И.Н.* // Обнаружение объектов со статистически зависимыми геометрическими параметрами на изображениях на основе теоретико-информационного обобщения преобразования Хафа: дисс. ... канд. техн. наук. СПб, 2015.
10. *Запьягаев С.А., Сорокин А.И.* // Прикладная информатика. 2009. № 4 (22). С. 76–86.
11. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. М., 2006.
12. *Farley B.G., Clark W.A.* // IRE Transactions on Information Theory. 1954. № 4. P. 76–84.
13. *Хайкин С.* Нейронные сети. М., 2006.
14. *El-Sayed M. A., Estaitia Y.A.* // International Journal of Advanced Computer Science and Applications (IJACSA). 2013. Vol. 4, Iss. 10.
15. *Toss T.* Automatic identification and cropping of rectangular objects in digital images. Master thesis. Uppsala University, 2012.
16. *Zhao F., Wei C., Wang J.* // J. of Software. 2011. Vol. 6, Iss. 5. P. 791.