

Анализ принципа распознавания объектов на изображении методом Виолы–Джонса

Национальный аэрокосмический университет им. Н. Е. Жуковского “ХАИ”

Описан метод детектирования объектов П. Виолы и М. Джонса. Рассмотрены концепции: интегральное изображение, признаки Хаара, машинное обучение AdaBoost, каскадный классификатор. Проанализированы сильные и слабые стороны метода Виолы–Джонса. Создано ПО FindFases на основе проанализированного алгоритма. Приведена структура диаграммы классов ПО. Приведены результаты работы программы, реализующей метод Виолы–Джонса на языке C++.

Ключевые слова: метод Виолы–Джонса, П. Виола и М. Джонс, распознавание объектов на изображении, компьютерное зрение, каскады Хаара, вейвлеты Хаара.

Введение

OpenCV – это библиотека компьютерного зрения, которая имеет множество возможностей, таких, как распознавание объектов на изображении и видео, анализ движения и др. К тому же, готовые варианты реализации OpenCV есть под большинство существующих систем. Одной из главных задач, решаемых ею, является распознавание объекта на изображении. Самым распространенным инструментом распознавания объектов на изображении является метод Виолы–Джонса.

Детектор объектов, используемый в OpenCV, опирается на метод П. Виолы и М. Джонса, который был опубликован в 2001 [1]. Этот подход к детектированию объектов на изображении комбинирует четыре ключевые концепции:

- представление изображения в интегральном виде для быстрого вычисления функций;
- использование простых прямоугольных функций, называемых признаками Хаара;
- обучение системы распознавания объектов на основе метода «машинного обучения AdaBoost»;
- организация алгоритма распознаваемых объектов в виде «каскадного классификатора»;

Преимуществами данного метода является высокая скорость работы и процент вероятности верного детектирования объекта на изображении. Таким образом, данный метод можно применять во многих сферах деятельности, например, для обнаружения распознавания номерных знаков на автомобилях (для дальнейшего их анализа) или обнаружения лиц на изображении. В данной статье будет рассмотрено применение этого метода на примере распознавания лиц.

1. Интегральное представление изображения

Чтобы эффективно определить наличие или отсутствие множества вейвлетов Хаара на огромном количестве областей изображения разного размера, П. Виола и М. Джонс применили метод интегрального представления изображения. Вес каждого пикселя представляет собой его яркость. Интегральное значение для каждого пикселя есть сумма всех значений над ним и слева от него плюс его собственный вес. Совершая обход начиная с левого верхнего угла вправо и вниз,

все изображение может быть интегрировано с несколькими целочисленными операциями.

Как показано на рис. 1, а, после интеграции значение каждого пикселя с координатами (x, y) содержит сумму всех пиксельных значений внутри прямоугольной области, которая имеет один угол в левой верхней части изображения и другой в положении (x, y) . Чтобы найти среднее пиксельное значение в этом прямоугольнике, необходимо только разделить значение в прямоугольнике (x, y) на площадь прямоугольника. Далее найдем суммарные значения для некоторых других прямоугольников, которые не имеют ни одного угла в левой верхней части изображения. Рис. 1. б иллюстрирует данное решение. Предположим, что нужно узнать суммированные значения в области D. И что это сумма значений пикселей в комбинированном прямоугольнике, $A+B+C+D$, минус сумма в прямоугольниках $A+B$ и $A+C$, плюс сумма пиксельных значений в прямоугольнике A. Так, с интегральным изображением можно найти сумму пиксельных значений для любого прямоугольника в первоначальном изображении всего с помощью трех целочисленных операций.

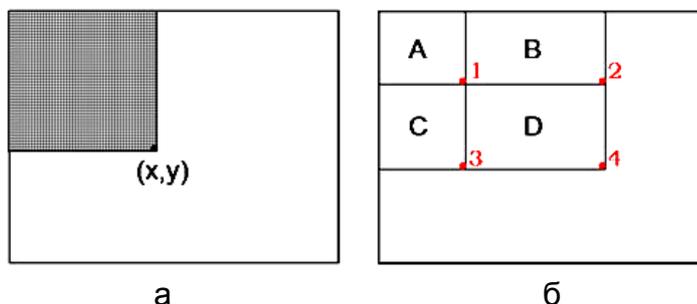


Рис. 1: Интегральное изображение

2. Использование вейвлетов Хаара

Особенности, которые использовали П. Виола и М. Джонс, базируются на каскадах признаков Хаара. Основной причиной, почему в основу метода Виолы–Джонса легли примитивы Хаара, стала попытка уйти от пиксельного представления с сохранением скорости вычисления признака. Такое название они получили из-за того, что их основой являются вейвлеты Хаара. Каскады Хаара представляют собой прямоугольные области, которые составлены из нескольких соседних прямоугольных областей, отмеченных как светлая или темная. На рис. 2 показаны функции, используемые в OpenCV (в оригинальном методе Виолы–Джонса наклонных областей не было).

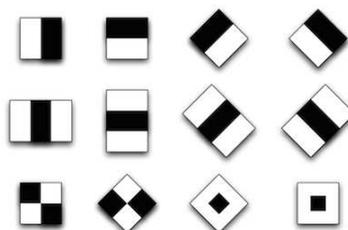


Рис. 2. Примеры функций Хаара, используемых в OpenCV

Из значений пары пикселей сложно вынести какую-либо осмысленную информацию для классификации, в то время как из двух признаков Хаара строится, например, первый каскад системы по распознаванию лиц, который имеет вполне осмысленную интерпретацию (рис. 3) [2]. Наличие функции Хаара определяется

посредством вычитания среднего значения области темных пикселей из среднего значения области светлых пикселей. Если разница превышает порог (определяемый в процессе обучения), то говорят, что функция является существующей.

Еще одним важным преимуществом каскадов перед пиксельным представлением является то, что сложность вычисления признака, как и получения значения пикселя, остается $O(1)$.

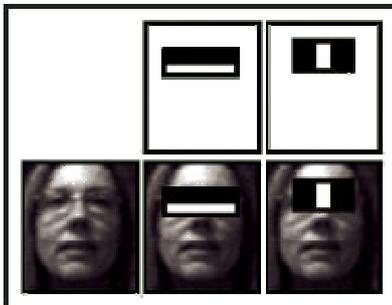


Рис. 3. Первые две Хаар-функции в оригинальном каскаде Виолы–Джонса

3. Метод машинного обучения AdaBoost

Для выбора конкретных используемых функций Хаара и установления пороговых уровней, П. Виола и М. Джонс используют метод машинного обучения AdaBoost. Он комбинирует много «слабых» классификаторов с целью создания одного «сильного». «Слабый» здесь означает такой классификатор, который получает правильный ответ ненамного чаще, чем случайное угадывание, что является недостатком. Однако имея множество таких слабых классификаторов, каждый из которых «выдвинул» окончательный ответ немного в верном направлении, можно получить серьёзную комбинированную силу для достижения корректного решения. AdaBoost выбирает набор слабых классификаторов для объединения и присваивает каждому из них свой вес. Эта взвешенная комбинация и является сильным классификатором. П. Виола и М. Джонс объединили серии классификаторов AdaBoost как последовательность фильтров, что особенно эффективно для классификации областей изображения. Каждый фильтр является отдельным классификатором AdaBoost с достаточно малым числом слабых классификаторов.

4. Каскадный классификатор

Порядок фильтров в каскаде основывается на весовых значениях, которые присваивает AdaBoost. Более тяжелые взвешенные фильтры располагаются в начале, для более быстрого устранения (отбрасывания) областей изображения, не содержащих данный объект изображения («не лицевых» областей). На рис. 3 показано наложение первых двух функции из каскада Виолы–Джонса на лицо. Первая функция основана на факте того, что область глаз более светлая, чем область верхней части щек, а вторая проверяет наличие более светлой переносицы между двумя темными областями глаз.

Принятый порог на каждом уровне устанавливается достаточно низким, чтобы пройти все (или почти все) лицевые образцы в тренировочном наборе. Фильтры на каждом уровне могут классифицировать тренировочные изображения, которые прошли все предыдущие этапы. Если во время работы какой-то из этих фильтров не пропускает область изображения, то тогда область сразу же классифицируется как «не лицо». Когда фильтр пропускает область изображения,

она переходит к следующему фильтру в последовательности. Области изображения, прошедшие через все фильтры, классифицируются как «лицо». П. Виола и М. Джонс назвали это фильтрацией цепи каскада. Каждый последующий уровень цепи каскада содержит все больше классификаторов, например каскад для распознавания лиц в анфас на первых уровнях содержит только один или два самых «сильных» классификатора, а последние – до 200, но «слабых».

5. Практическое применение метода Виолы–Джонса

Алгоритм распознавания лиц по методу Виолы–Джонса может применяться в различных сферах для решения самых разнообразных задач. Его самая распространенная реализация принадлежит библиотеке компьютерного зрения OpenCV.

Классическим примером для демонстрации работы метода Виолы–Джонса является программа распознавания лиц в фас на фотографии[3]. Диаграмма классов подобной программы изображена на рис. 4. Данная программа позволяет определить координаты центров лиц, расположенных на обрабатываемом изображении. Ее преимуществом является возможность быстро распознавать объекты типа «лицо» на изображении, что позволяет использовать подобный алгоритм при анализе данных, поступающих в реальном времени, как при обработке видео.

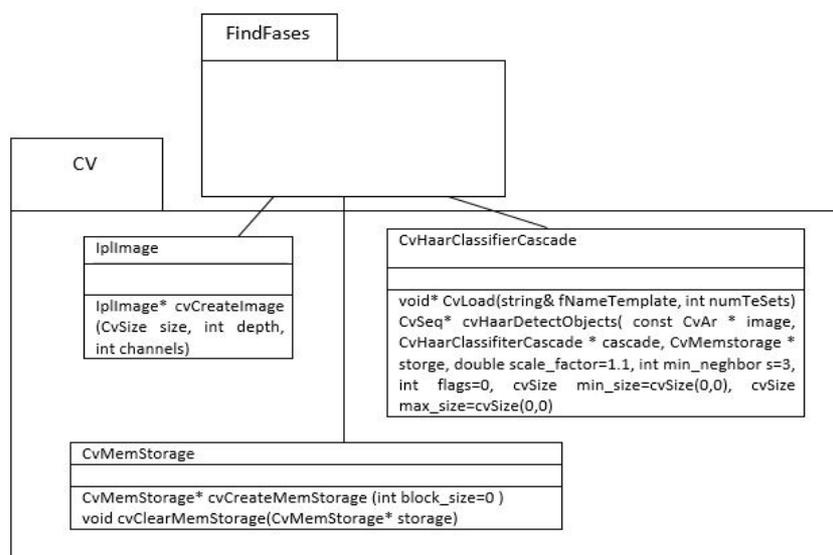


Рис. 4. Диаграмма классов программы, реализующей метод Виолы–Джонса с использованием библиотеки OpenCV для языка C++

Преимуществами такого метода являются высокая вероятность верного распознавания объектов и возможность сравнительно простого изменения типа распознаваемого объекта (при наличии классификатора), например, на номерные знаки. Недостатками же являются резкое ухудшение вероятности обнаружения объекта при усилении шумов на изображении, а также высокая сложность создания новых классификаторов распознавания других объектов, которые будут использованы в процессе машинного обучения.

Выводы

В ходе исследования метода Виолы–Джонса с использованием библиотеки OpenCV был реализован программный код [3]. Проведен эксперимент по распо-

знаванию лиц на изображениях с разрешениями от 700 x 513 до 2078 x 1294. Полученные результаты: из 116 лиц распознано 108 лиц, при этом было произведено 7 ложных распознаний. Средняя скорость обработки изображения была значительно меньше одной секунды, что дает возможность использовать данный метод для обработки изображений в реальном времени, например, чтобы находить лица на видео для определения количества людей в помещении с помощью камеры.

Список литературы

1. Viola. P. Robust Real-Time Object Detection: tech. report / P. Viola, M. Jones. – (city: Cambridge) Cambridge, 2001. – 320 p.
2. Как работает детектирование лиц [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://www.digital-sky.ru/point-3/artcateg-17/article-10.html>, свободный – Загл. с экрана.
3. opencv-extension-library [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://opencv-extension-library.googlecode.com/svn/trunk/QtOpenCV/example/facedetect/facedetect.c>, свободный – Загл. с экрана.
4. Работа каскада Хаара в OpenCV [Электронный ресурс]. – Электрон. дан. – Режим доступа: <http://habrahabr.ru/company/recognitor/blog/228195/>, свободный – Загл. с экрана.

Поступила в редакцию 10.02.2015

Аналіз принципів розпізнання об'єктів на зображенні за методом Віоли–Джонса

Описано метод детектування об'єктів П. Віоли та М. Джонса. Розглянуто концепції: інтегральне зображення, ознаки Хаара, машинне навчання AdaBoost, каскадний класифікатор. Проаналізовано сильні й слабкі сторони методу Віоли–Джонса. Створено ПЗ FindFases на основі проаналізованого алгоритму. Наведено структуру діаграми класів ПЗ. Наведено результати роботи програми, що реалізує метод Віоли–Джонса на мові C ++.

Ключові слова: метод Віоли–Джонса, П. Віола та М. Джонс, розпізнавання об'єктів на зображенні, комп'ютерний зір, каскади Хаара, вейвлети Хаара.

Analysis the principle of objects recognition on the basis of Viola – Jones

Describes a method for detecting objects P.Viola and M.Dzhons. We consider the concept: in-integrand image Haar-like features, machine learnin AdaBoost, cascade classifier Analyzed the strengths and weaknesses of the method of Viola–Jones. Create by FindFases analyzed on the basis of the algorithm. The structure of the class diagram software. The results of the work program implementing the method of Viola–Jones C ++.

Keywords: Viola-Jones method, P.Viola and M.Dzhons, recognition of objects in the image, computer vision, cascades Haar, Haar wavelets.