

электронный журнал

МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК

Издатель: ФГБОУ ВПО «Московский государственный технический университет им. Н.Э. Баумана»

УДК 004.9

Обнаружение лиц на цветном растровом изображении с применением модифицированного метода Виолы-Джонса

МГТУ им. Н.Э. Баумана

Вязовых М.В. (maxvyaz@bmstu.ru),
Зайцев К.И. (kirzay@gmail.com),
Мухомтов М.В. (maxmix1205@ya.ru),
Перов А.Н. (temaperov@gmail.com)

Введение

В данной работе рассмотрен алгоритм обнаружения лиц на цветном растровом изображении, предложенный Полом Виолой и Майклом Джонсом [1]. Данный алгоритм был успешно реализован на языке C++ с помощью библиотеки машинного зрения OpenCV. Проведен анализ возможных ошибок алгоритма. Предложена его модификация, также реализованная на языке C++.

1 Постановка задачи

Задача выделения лица человека при естественном или искусственном освещении с последующей идентификацией всегда находилась в ряду самых приоритетных задач для исследователей, работающих в области систем машинного зрения и искусственного интеллекта. Однако, множество исследований, проводящихся в ведущих научных центрах всего мира в течение нескольких десятилетий, так и не привело к созданию реально работающих систем компьютерного зрения, способных обнаруживать и распознавать лицо человека в различных условиях.

2 Обзор существующих алгоритмов

В таблице 1 перечислены некоторые существующие алгоритмы, а так же их преимущества и недостатки.

Таблица 1. Сравнение существующих алгоритмов распознавания

Название	Преимущества	Недостатки
----------	--------------	------------

Поиск по цвету	Высокая скорость обнаружения, возможность селекции по цвету кожи, не требует обучения	Очень высокая вероятность ложного обнаружения
Корреляционные методы	Высокая точность обнаружения	Крайне низкая скорость работы, требуется большая обучающая выборка с жесткими условиями съемки
Поиск по характерным точкам лица	Высокая точность и скорость обнаружения для изображений лиц крупным планом	Низкая точность и скорость обнаружения для изображений с несколькими лицами

3 Обоснование выбора алгоритма Виолы-Джонса

Метод Виолы-Джонса является одним из лучших по соотношению показателей эффективности распознавания/скорость работы. Также этот детектор обладает низкой вероятностью ложного обнаружения лица. Алгоритм также хорошо работает и распознает черты лица под небольшим углом, примерно до 30 градусов, а так же при различных условиях освещенности. На момент написания алгоритм является основополагающим для поиска объектов на изображении в реальном времени в большинстве существующих методов распознавания и идентификации.

4 Основные понятия метода Виолы – Джонса

- 1) Представление изображения в интегральном виде, что позволяет вычислять быстро необходимые объекты [2];
- 2) Использование признаков Хаара [3], с помощью которых происходит поиск нужного объекта (в данном контексте, лица и его черт);
- 3) Бустинг (от англ. *boost*- улучшение, усиление) для выбора наиболее подходящих признаков для искомого объекта на данной части изображения [2];
- 4) Использование каскадов признаков для быстрого отбрасывания окон, где не найдено лицо.

4.1 Интегральное представление изображений

В интегральном представлении изображений формируется матрица, совпадающая по размерам с исходным изображением. В каждом ее элементе хранится сумма яркостей пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по следующей формуле:

$$L(x, y) = \sum_{i=0, j=0}^{i \leq x, j \leq y} I(i, j),$$

Где $I(i, j)$ – яркость пикселя исходного изображения.

Пусть в изображении есть прямоугольник АБВГ (рис. 1). Тогда сумму пикселей внутри прямоугольника можно выразить через разность смежных прямоугольников по следующей формуле:

$$S(ABVG) = L(B) + L(\Gamma) - L(A) - L(B)$$

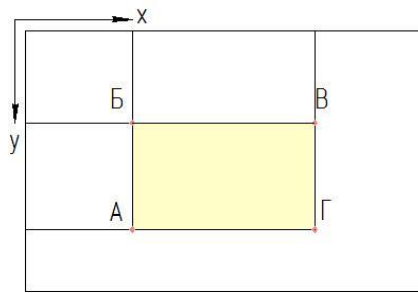


Рис 1. К расчету суммы яркостей пикселей в произвольном прямоугольнике

4.2 Признаки Хаара

Признак – отображение области изображения на множество допустимых значений признака т.е. $f: X \Rightarrow D_f$, где

X – область изображения для вычисления признака

D_f – множество допустимых значений признака

В стандартном методе Виолы-Джонса используются прямоугольные признаки, изображенные на рис. 2а, они называются примитивами Хаара. В расширенном методе Виолы-Джонса используются дополнительные признаки (рис. 2б).

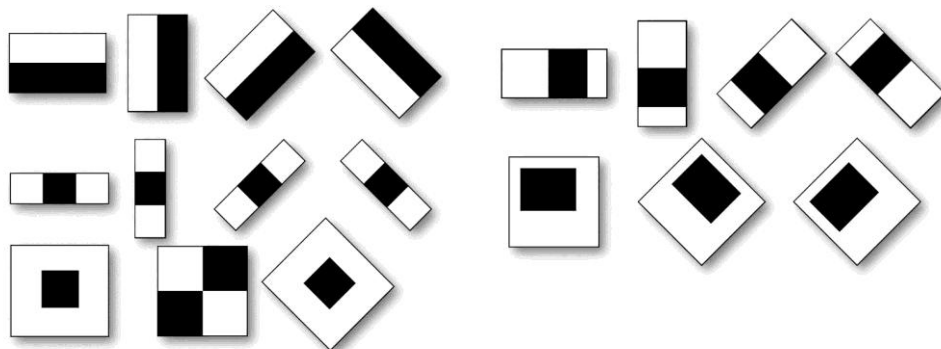


Рис.1. а) Стандартные примитивы Хаара б) Дополнительные примитивы Хаара

Вычисляемое значение признака будет значение $F=A-B$, где A – сумма яркостей точек закрываемых светлой частью признака, а B – сумма яркостей точек закрываемых темной частью. Для их вычисления используется понятие интегрального изображения, рассмотренное выше.

Если заданы признаки f_1, \dots, f_n , то вектор признаков $x=(f_1(x), \dots, f_n(x))$ называется признаковым описанием объекта $x \in X$. Признаковые описания допустимо отождествлять с самими объектами.

Метод Виолы-Джонса в общем виде ищет лица и черты лица по общему принципу сканирующего окна, схема работы которого представлена на рис. 3.

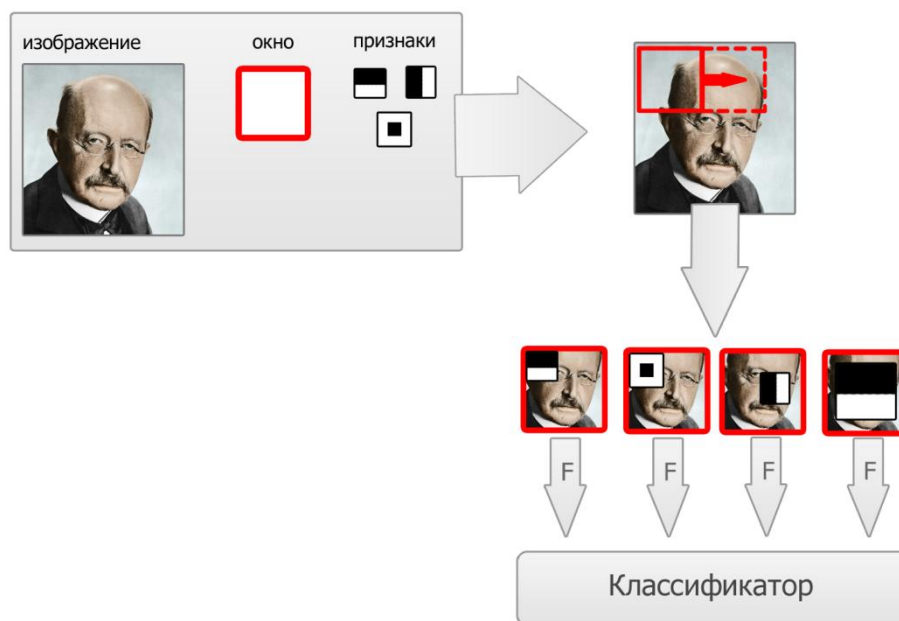


Рис. 3. Схема работы принципа сканирующего окна.

- 1) Сканирующее окно перемещается по изображению
- 2) В каждом положении вычисляются сотни тысяч значений признаков для их различных форм, размеров и расположения
- 3) Сканирование производится для различных масштабов окна

4.3 Бустинг

Бустинг — комплекс методов, способствующих повышению точности аналитических моделей. Бустинг (от англ. *boosting* – повышение, усиление, улучшение) означает дословно «усиление» «слабых» моделей – это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов.

Бустинг представляет собой жадный алгоритм построения композиции алгоритмов (*greedy algorithm*) — это алгоритм, который на каждом шагу делает локально наилучший выбор в надежде, что итоговое решение будет оптимальным. Бустинг над решающими деревьями считается одним из наиболее эффективных методов с точки зрения качества классификации. Во многих экспериментах наблюдалось практически неограниченное уменьшение частоты ошибок на независимой тестовой выборке по мере наращивания композиции.

Развитием данного подхода явилась разработка более совершенного семейства алгоритмов бустинга АдаБуст (*AdaBoost: adaptive boosting* – адаптированное улучшение), предложенная Йоавом Фройндом (*Freund*) и Робертом Шапиром (*Schapire*) в 1999 году [2], который может использовать произвольное число классификаторов и производить обучение на одном наборе примеров, поочередно применяя их на различных шагах.

4.4 Использование каскадов признаков

Дерево принятия решений — это дерево, в листьях которого стоят значения целевой функции, а в остальных узлах — условия перехода, определяющие по какому из ребер идти. Каскадная модель сильных классификаторов — это, по сути, дерево принятия решений, где каждый узел дерева построен таким образом, чтобы детектировать все интересующие образы и отклонять области, не являющиеся образами.

Каскад применяется к изображению по следующим правилам:

- работа с «простыми» классификаторами — при этом отбрасывается часть «отрицательных» окон;
- положительное значение первого классификатора запускает второй, более приспособленный и так далее;
- отрицательное значение классификатора на любом этапе приводит к немедленному переходу к следующему сканирующему окну, старое окно отбрасывается;
- цепочка классификаторов становится более сложной, поэтому ошибок становится намного меньше.

5 Модификация алгоритма

Эффективность метода Виолы-Джонса доказана многими исследованиями, однако он обладает рядом недостатков. Предлагаемый ниже метод позволяет внести следующие улучшения:

- уменьшение вероятности ложных обнаружений;
- снижение времени работы алгоритма

При анализе классического метода Виолы — Джонса было выявлено, что большинство времени работы алгоритма занимает операция вычисления расположения признаков Хаара, так что логично было предположить, что сужение области поиска лиц поможет алгоритму Виолы — Джонса работать эффективнее. У человеческого лица есть много особенностей, помогающих его детектировать. Одной из таких особенностей является цвет человеческой кожи. Определив цветовые координаты текущего пикселя, можно достаточно точно установить, относится ли данный участок к человеческой коже. Причем эта процедура не является проблемной с точки зрения вычислительных затрат.

Таким образом, суть реализованного улучшения заключается в предварительной обработке изображения, которая включает в себя следующие процедуры:

- 1) Поиск областей, содержащих изображение кожи человека;
- 2) Уточнение найденных областей;
- 3) Применение алгоритма Виолы — Джонса;



Рис. 4. Последовательность процедур работы модифицированного алгоритма Виолы – Джонса

6 Поиск областей содержащих изображение кожи человека

Данная процедура состоит из трёх этапов.

6.1. Децимация исходного изображения

Понижение частоты дискретизации в два раза, уменьшает размер исследуемого изображения в такое же количество раз. Цель нашего улучшения состоит в примерном выделении областей, содержащих кожу, так что такое «загрубление» исследуемого изображения не вносит каких – либо ошибок в алгоритм, напротив такая процедура позволяет увеличить скорость обработки на следующих этапах.

6.2 Обработка изображения низкочастотным фильтром

При разработке данного алгоритма было замечено, что в выбранный цветовой диапазон попадает не только кожа человека, но так же отдельные пиксели фона. И их достаточно много. Размытие изображения низкочастотным фильтром позволяет частично устранить фоновую составляющую. В данной реализации был выбран фильтр Гаусса с размеров ядра 11x11 пикселей.

6.3 Поиск областей содержащих изображение кожи человека

Проанализировав выборку из 100 фотографий, содержащих лица людей, было установлено, что цвет кожи европеоида можно описать в цветовом пространстве RGB следующей зависимостью:

$$\left\{ \begin{array}{l} R > 90 \\ G > 40 \\ B > 20 \\ R - G > 15 \\ R > B \end{array} \right. ,$$

где R, G, B - соответственно значения красной, зеленой и синей составляющих яркости пикселя. Используя данную зависимость, каждый пиксель изображения проверяется на принадлежность к заданному цветовому диапазону. Все пиксели попавшие в данный диапазон окрашиваются в белый цвет, не попавшие – в черный цвет, тем самым создается бинарная маска изображения.

7 Уточнение найденных областей

Результатом работы данной процедуры является устранение ложных областей и выделение участков на исходном изображении потенциально содержащих кожу. Процедура состоит из трёх этапов.

7.1 Морфологическая обработка изображения эрозия и дилатация

Применение эрозии к изображению (рис. 5а) позволяет исключить все, оставшиеся после фильтра Гаусса, шумовые области. Прimitив, используемый эрозией – квадрат с размерами 3x3 пикселя.

Однако эрозия помимо удаления шумовых областей, так же уменьшает интересные для нас участки. С этой целью применяют операцию дилатации (рис.5б), которая увеличивает найденные области. Прimitivesом является квадрат размером 11х11 пикселей. После выполнения данных операций бинарная маска интерполируется до размера исходного изображения.

7.2 Запись контуров

Запись контуров осуществляется при помощи цепного кода Фримена. Применение такого алгоритма позволяет представить все контуры в виде последовательности отрезков прямых линий определенного направления. В основе этого представления лежит 8-связная решетка. Из найденных контуров формируется список, включающий в себя следующие параметры: номер контура, цепной код, пиксели, принадлежащие контуру, с минимальными и максимальными координатами. Результат такой процедуры показан на рис. 5в.

7.3 Ограничение найденных контуров прямоугольниками

Все найденные контуры по очереди ограничиваются прямоугольниками, длина и ширина которых немного увеличивается (рис.5г, 5д). Это необходимо для повышения вероятности попадания лица в текущий прямоугольник. Затем координаты левой верхней вершины запоминаются, а область исходного изображения, ограниченная данным прямоугольником передается на следующий этап.

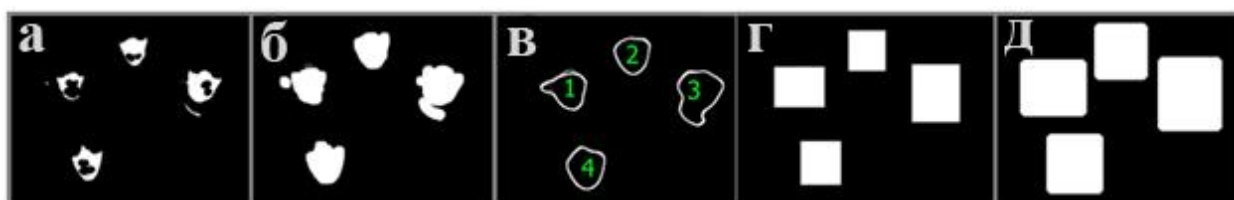


Рис. 5. а – эрозия бинарной маски; б- дилатация бинарной маски; в- оконтуривание областей; г - ограничение найденных контуров прямоугольниками; д – увеличение прямоугольников

8 Применение алгоритма Виолы – Джонса

Теперь алгоритм Виолы-Джонса применяется уже не ко всему изображению, а лишь к небольшим областям. Алгоритм Виолы – Джонса выносит вердикт, есть ли в этой области лицо или нет. Если оно все же присутствует, то на исходное изображение наносится метка, центр которой смещается с учетом координат запомненных на предыдущем этапе.

9 Экспериментальные данные

Для эксперимента была сделана выборка из 50 цветных фотографий с разрешением 640х480 пикселей, содержащих изображение лиц людей. Из рисунка 3 можно сделать вывод, что скорость работы алгоритма в большинстве случаев возросла более чем в 4 раза. Наиболее эффективно модифицированный алгоритм работает с групповыми фотографиями. В работе с портретными фотографиями метод проявил себя несколько хуже. Это связано с тем, что основная идея улучшения метода Виолы – Джонса заключается в уменьшении области поиска. При портретной же съемке лицо человека может занимать

значительную часть всего изображения. По всплескам на графике №2, приведенном на рис. 6 можно легко увидеть какие из фотографий в выборке являлись портретными.

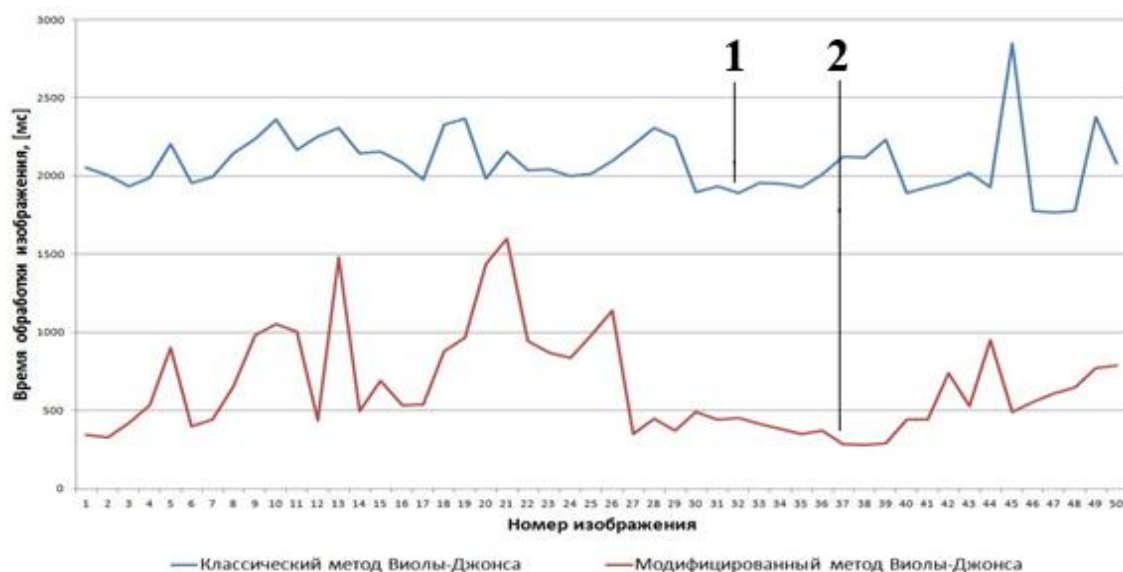


Рис. 6. Сравнение скорости работы алгоритмов

На рис. 7 приведена статистика ошибок двух типов: пропуск объекта и ложное обнаружение.

Можно заметить, что классический алгоритм в большей степени «страдает» от ошибки ложного обнаружения. Модифицированный же алгоритм практически не чувствителен к подобной ошибке.

Статистика пропуска объектов у обоих алгоритмов примерно одинакова.

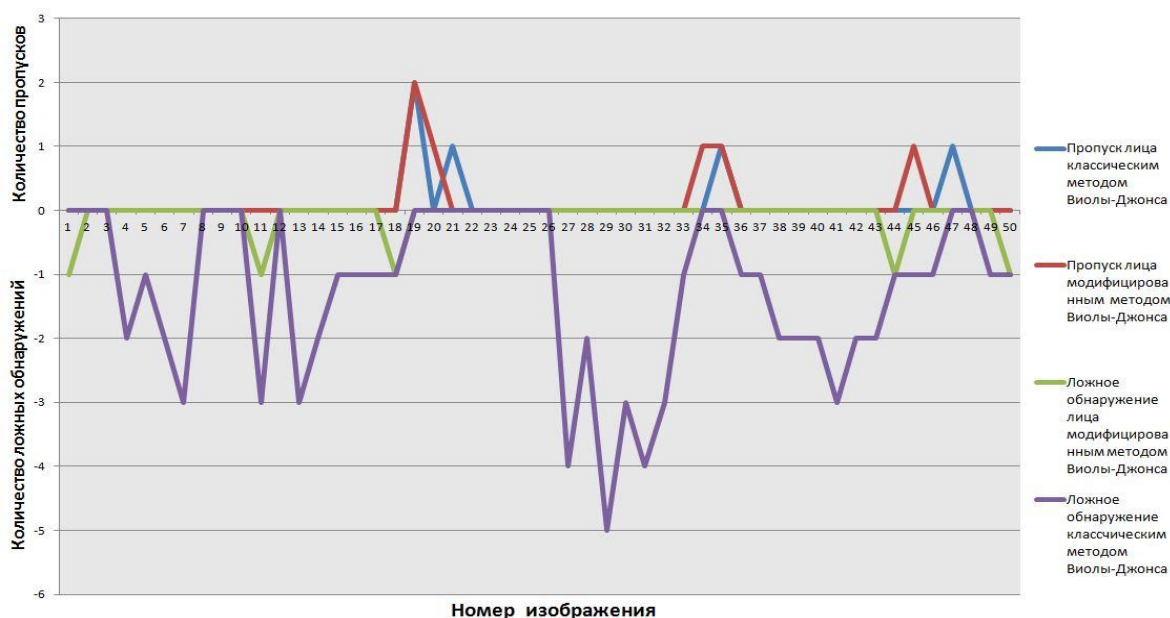


Рис. 7. Статистика ошибок

Заключение

В ходе работы были решены поставленные задачи: значительно снижена вероятность ложного обнаружения, а также повышена скорость работы (в среднем в три раза) по сравнению с классическим методом Виолы-Джонса. Для дальнейшего ускорения алгоритма можно применить распараллеливание

процессов вычисления признаков, что приведёт к возможности использования алгоритма в реальном масштабе времени.

Литература

1. P. Viola and M.J. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features», proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001
2. G. Bradski, A. Kaebler, «Learning OpenCV», Published by O'Reilly Media 2008, 495-512 p.
3. Р.Гонсалес, Р.Вудс, «Цифровая обработка изображений», ISBN 5-94836-028-8, изд-во: Техносфера, Москва, 2005. – 1072 с.