

УДК 004.93'12:004.932.72'1

Распознавание лиц и глаз на фотографии

Воробьева К.О., студент

*Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
кафедра «Специальная робототехника и мехатроника»*

*Научный руководитель: Рубцов В.И., к.т.н., доцент
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
kafsm7@sm.bmstu.ru*

Введение

Задачи, связанные с распознаванием объектов, чрезвычайно сложны и актуальны в настоящее время. Распознавание образов для человека является задачей привычной и, казалось бы, простой, однако, при разработке алгоритмов, воспроизводящих функции человека нередко приходится столкнуться с большим количеством трудностей. Важно обеспечить точность, приемлемую скорость работы, а также независимость работы программы от внешних условий: от освещенности, от качества изображения, от поворота объекта и иных. В данной работе рассмотрен метод решения задачи распознавания лиц и глаз ввиду возрастающих потребностей широкого применения и использования лицевых детекторов. Так, в частности, система обнаружения подозрительных лиц представлена в качестве одного из методов решения проблем безопасности АГР «Умный город». Также функция распознавания лиц широко применяется в фотоаппаратах, охранных системах.

Постановка задачи

Целью работы является написание программы распознавания лиц и глаз на изображении. В качестве основных требований работы программы выступают точность и высокая скорость. Так, допустимая погрешность не может превышать 8 %, а время работы алгоритма – двух секунд. В результате необходимо получить программу обнаружения лиц и глаз на языке C++, которая бы удовлетворяла заданным требованиям точности и быстродействия, а также была бы устойчива к изменению прически, наличию усов, бороды, возрастным изменениям.

Анализ существующих методов решения

Задача распознавания лиц имеет значительную практическую перспективу и большой интерес среди исследователей, в связи с чем, на нынешний момент существует огромное количество алгоритмов распознавания лиц и глаз[2]. Среди них можно выделить:

1)Метод главных компонент. Опирается на векторы в линейном пространстве, сводится к вычислению главных компонент лица и последующим сравнением исследуемой фотографии и некоторых уже известных изображений. Главными недостатками данного метода являются необходимость иметь одинаковое освещение, а также требование к одинаковому положению головы на изображении, что существенно снижает его универсальность.

2)Сравнение эталонов. Необходимо выделить несколько областей на фотографии, а затем сравнивать их с имеющимися изображениями. В силу того, что данный метод – один из первых разработанных алгоритмов распознавания, он очень прост в реализации, однако, требует большого количества ресурсов, длительного времени работы, не допускает изменения условий съемки.

3)Методы, основанные на геометрических характеристиках. Изначально осуществляется поиск ключевых точек лица, а затем сравниваются с уже имеющимися признаками. Недостатками метода выступают строгие требования к условиям съемки, а также требования отсутствия помех на изображении.

4)Нейронные сети.Происходит сначала извлечение ключевых характеристик изображения, а затем их классификация. При высокой эффективности недостатками нейронных сетей выступают отсутствие формализованных алгоритмов, высокая сложность реализации, необходимость обучения.

5) Метод Виолы-Джонса. Данный метод будет рассмотрен далее, так как с его использованием и была решена поставленная задача.

Выбранный метод решения задачи

Алгоритм распознавания лиц и глаз человека на фотографии был реализован с помощью метода Виолы-Джонса на языке C++ с использованием стандартной библиотеки OpenCV, так как данный алгоритм обладает высокой эффективностью и значительной скоростью работы. Сама задача распознавания по указанному методу состоит в основном из двух частей, первая – это выделение лица на изображении, вторая – непосредственно идентификация выделенного как лица. Вторая часть наиболее трудоемкая в реализации, так как связана с полным сканированием изображения и с поиском признаков, отвечающих за степень схожести с лицом.

1. Алгоритм распознавания лиц и глаз

Для нахождения глаз необходимо определить местоположение лица. Это позволит снизить время работы алгоритма, а так же нивелировать количество ложных срабатываний. Алгоритм поиска лица Виолы-Джонса состоит из четырех этапов или последовательных концепций[1]:

- использование функций Хаара
- преобразование изображение в интегральное для быстрого поиска
- машинное обучение
- использование каскадного классификатора.

Рассмотрим каждую из этих концепций более подробно.

1.1.Использование функций Хаара

Функции Хаара базируются на вейвлетах Хаара, представляющие собой прямоугольные волны одинаковой длины, в двумерном пространстве – пара двух соседних прямоугольников: светлого и темного. Используемые алгоритмом функции не являются подлинными вейвлетами Хаара, однако, они представляют собой набор примитивов, схожих с ними, но наиболее удобных для работы с изображениями, почему и носят название функций Хаара[7].

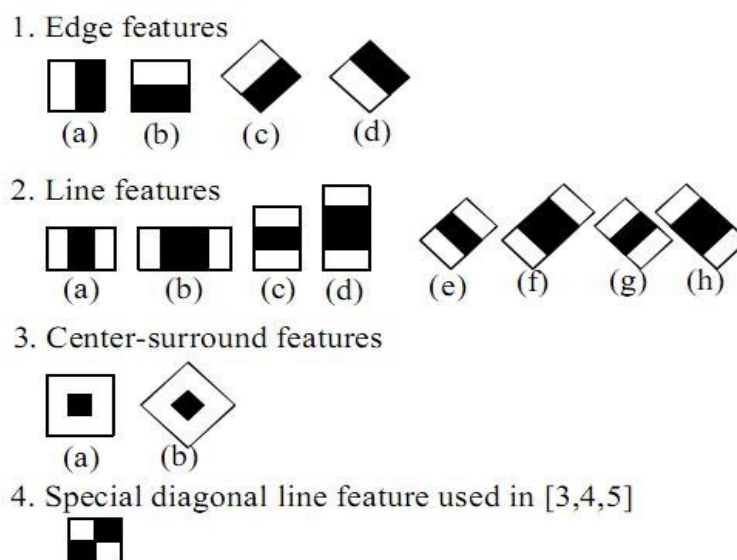


Рис.1. Функции Хаара

При работе с ними происходит следующее: в сканирующем окне алгоритм использует один из примитивов, накладывая его на изображение окна, затем считает сумму значений пикселей каждой области функции Хаара: $F = X - Y$, где X – сумма яркостей точек светлого прямоугольника, Y – сумма яркостей точек темного прямоугольника. Функции Хаара описывают значения перепада яркости на изображении.



Рис. 2. Накладывание функции Хаара на изображение

После чего получает обобщенную характеристику анизотропии части изображения. Исходя из того, что область носовой перегородки более светлая, чем область глаз, используется соответствующая функция Хаара. То есть задача алгоритма машинного обучения – выбрать те функции, которые наиболее эффективно детектируют рассматриваемый объект[5].

1.2. Преобразование исходного изображения в качестве интегрального

Интегральное преобразование представляется собой совпадающую по размеру с исходным изображением матрицу, состоящую из элементов – интенсивностей пикселей, расположенных выше и левее рассматриваемой области[3].

1	2
3	4

$$I(1234) = I(3) + I(2) - I(1) - I(4)$$

Рис.3. Интегральное преобразование

Расчет матрицы занимает линейное время, пропорциональное количеству пикселей в изображении. Используя подобную матрицу, можно вычислить сумму любого прямоугольника, произвольной площади.

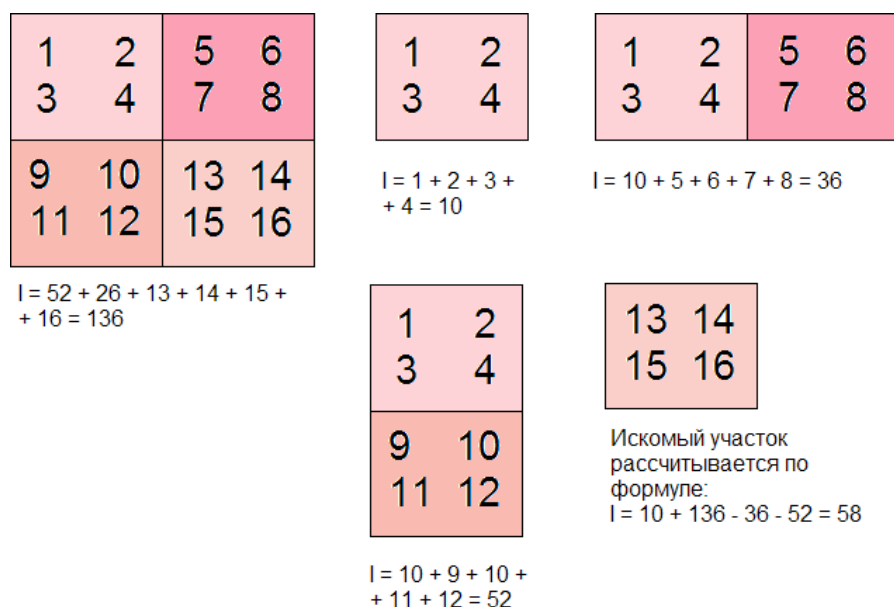


Рис. 4. Пример расчета интегральной матрицы

1.3. Метод машинного обучения AdaBoost и каскадный классификатор

Бустинг(англ. boosting- усиление) – совокупность методов, увеличивающих, усиливающих точность аналитических моделей. Алгоритмом Виолы-Джонса данный метод используется для выбора оптимального набора функций Хаара. AdaBoost представляет собой композицию алгоритмов машинного обучения для устранения ошибок предыдущего алгоритма. Данный метод комбинирует множество простых классификаторов, способный идентифицировать объект с вероятностью сходной со случайным принятием решений, с целью получить сложный и точный классификатор. AdaBoost присваивает каждому простому классификатору свой коэффициент. Таким образом, сложный классификатор представляет собой совокупность фильтров, каждый из которых обладает набором признаков лица. Изображение сканирующего окна последовательно проходит каждый из фильтров, если хотя бы один фильтр определит: «Рассматриваемое изображение - предположительно не лицо», то в этом случае сложный алгоритм определяет рассматриваемую область как не лицо и переходит к следующей итерации[4].

1.4. Распознавание глаз

Распознавание глаз происходит с использованием той же функции, что и в случае распознавания лиц. Используется каскад, содержащий функции Хаара, способствующий идентификации на изображении глаз. Как было отмечено ранее, детектирование глаз осуществляется на участке изображения, которое было идентифицировано как лицо. Это необходимо для устранения ложных срабатываний(ошибки первого рода), а также для

увеличения скорости алгоритма. Принцип работы воспроизводит выше рассмотренный метод:

- сканирующее окно размера 20x20 последовательно движется по изображению, проверяя на каждой итерации около 200 000 вариантов расположения функций Хаара,
- затем сканирующее окно масштабируется на 10 %, сканирование проводится с шагом в одну ячейку,
- классификатор рассматривает все признаки, для определения части изображения в качестве глаза необходимы положительные ответы от всех признаков.

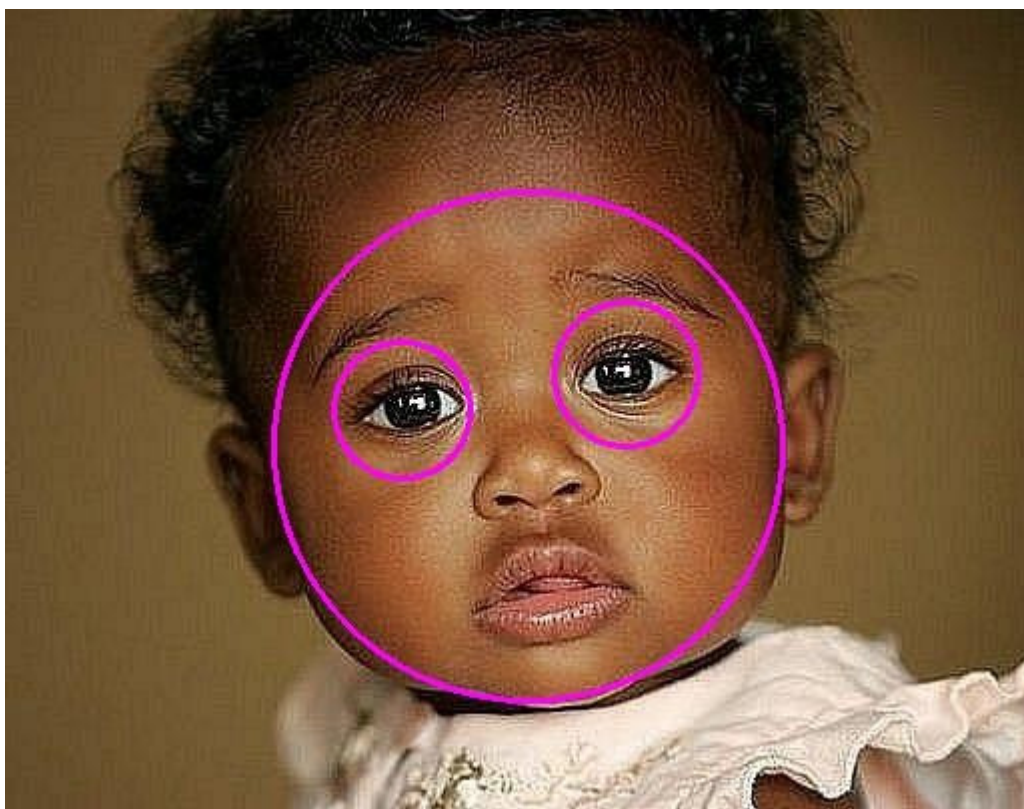


Рис. 5. Результат работы алгоритма

2. Функция cvHaarDetectObject

На основе рассмотренного алгоритма была написана программа с использованием библиотеки OpenCV. Автором используется лицевой детектор указанной библиотеки – классификатор Хаара. Данный классификатор использует данные, которые находятся в XML-файле, содержащим признаки распознаваемых объектов. Библиотека включает в себя три файла для обнаружения лиц анфас и один для лиц в профиль, также несколько файлов, позволяющих распознавать туловище человека. На вход данной функции подаются семь параметров: указатель на изображение, XML-файл, буфер памяти, минимальный масштаб поиска, минимальный соседний порог, коэффициент увеличения

масштаба, флаговую переменную и каскад Хаара. Минимальный масштаб - это минимальный размер лица, поиск которого будет осуществляться, стандартное значение размера лица: 24x24. Если нет необходимости поиска лиц подобного размера, имеет смысл увеличить размер окна до предполагаемого значения, ввиду того, что будет использовано значительное процессорное время. Также следует сохранить пропорцию данного параметра. Следующий параметр – минимальный соседний порог необходим для устранения, так называемых, ошибок второго рода, то есть идентификация части изображения как лица, которая в свою очередь лицом не является. Данная величина позволяет отбрасывать изолированные обнаружения. Стандартное значение в библиотеке – три, таким образом, пересечение трех и более прямоугольник образуют один, свидетельствующий о том, что рассматриваемый образ – лицо, меньшее количество найденных объектов будет игнорироваться алгоритмом. Коэффициент масштаба определяет, как сильно будет увеличиваться размер сканирующего окна. С возрастанием этого коэффициента увеличивается скорость работы метода, однако, снижается качество. Стандартное значение – 1.1, то есть с каждым проходом сканирующего окна, оно будет увеличиваться на 10 %. Шестой параметр используется для того, чтобы установить, что является наиболее приоритетным качество распознавания или скорость. Если флаг установлен, то классификатор пропускает те части изображения, где встретить лицо маловероятно. Установленный флаг увеличит время работы, однако, возможно, приведет к пропуску некоторых лиц. Каскад Хаара – каскад фронтальных лицевых детекторов, возможно использовать как собственный файл, так и стандартный[6].

```
cvHaarDetectObjects(small_img, cascade, storage, 1.1, 3, 0 |  
CV_HAAR_DO_CANNY_PRUNING);
```

3. Анализ полученных результатов

Как известно, в теориях распознавания объектов принято выделять ошибки двух видов: ошибки первого рода(ложное срабатывание) и второго рода(ложноотрицательное срабатывание). То есть при работе алгоритма объекты, найденные программой, но лицами не являющимися, будут называться ошибками первого рода, а в случае пропуска лиц – ошибками второго рода. Для тестирования написанной программы автором использовалась библиотека лиц Эссекского университета, содержащей 395 изображений лиц людей, по результатам чего была построена таблица «Точность работы программы».

Точность работы программы

	Правильно распознанные лица	Ошибки первого рода	Ошибки второго рода
Количество	393	14	2
Процент от общего числа	99.49	3.54	0.51

Главной негативной чертой рассматриваемого метода решения распознавания лиц является непригодность данного алгоритма в использовании при повороте лица. Так, при повороте лица на угол около 30° алгоритм перестает работать. Анализ данного качества работы метода представлен на рисунке 6.

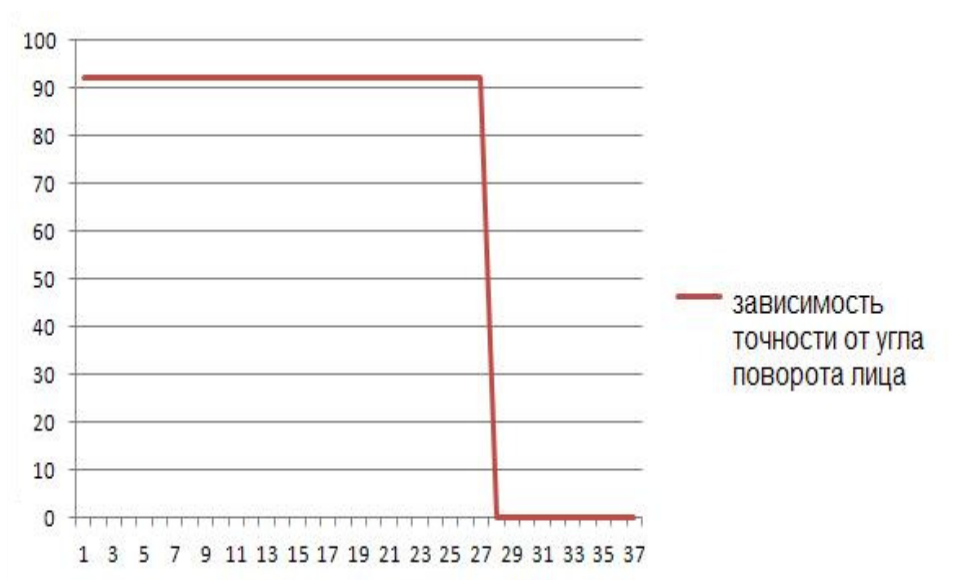


Рис. 6. График зависимости точности от угла поворота лица

Для распознавания лиц представляют существенное значение два качества работы алгоритма: точность и время работы. Как отмечалось выше, метод построен на принципе сканирующего окна, таким образом, время работы прямо пропорционально размеру изображения (рис. 7). Для обеспечения заданного требования по времени работы количество пикселей исследуемого изображения не должно превышать 20000 пикселей.

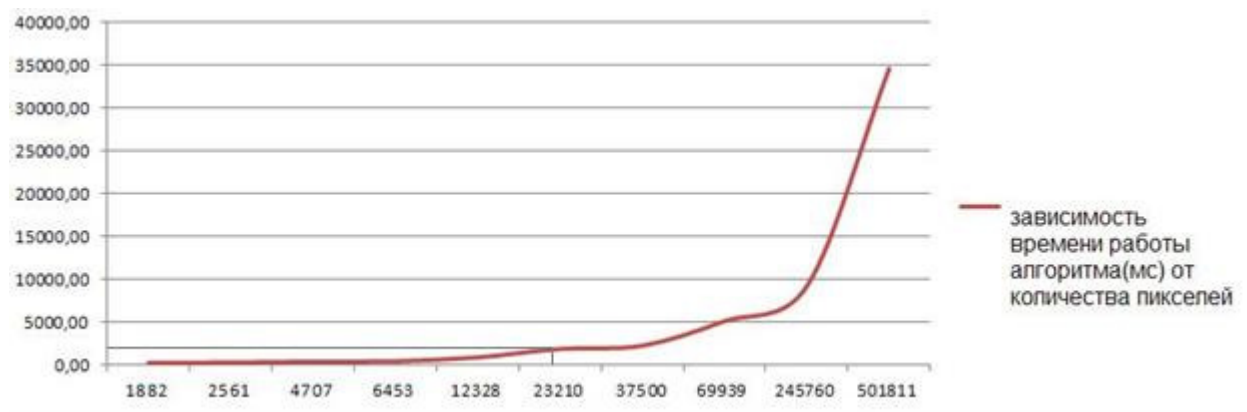


Рис. 7. График зависимости времени работы алгоритма от количества пикселей

Заключение

В данной статье был рассмотрен метод распознавания лиц и глаз на цветном изображении. Изложенный метод обладает простой реализацией, достаточно высокой точностью и скоростью. По результатам работы программы был проведен анализ точности и быстродействия, а также проанализированы условия работы алгоритма.

Список литературы

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. Р.Гонсалес, Р.Вудс, «Цифровая обработка изображений», ISBN 5-94836-028-8, изд-во: Техносфера, Москва, 2005. – 1072 с.
3. Метод Виолы-Джонса как основа для распознавания лиц / Хабрахабр. URL.<http://habrahabr.ru/post/133826/>(дата обращения: 14.06.13)
4. Распознавание лиц по методу Виолы-Джонса /SqFace. URL.<http://sqface.ru/>(дата обращения: 14.06.13).
5. Как работает детектирование лиц/ DigitalSky. URL.<http://www.digital-sky.ru/point-3/artcateg-17/article-10.html>(дата обращения: 14.06.13)
6. Зрение при помощи OpenCV. Часть 2 – нахождение лиц на изображениях/ DigitalSky. URL.<http://www.digital-sky.ru/point-3/artcateg-17/article-11.html> (дата обращения: 14.06.13).
7. В поисках НЛО. Детект объектов на изображении / Хабрахабр. URL.<http://habrahabr.ru/post/67937/>(дата обращения: 14.06.13).