

Разработка алгоритма детектирования лиц с использованием комбинаций каскадов Хаара

Е.А. Рудинская¹, Р.А. Парингер¹

¹Самарский национальный исследовательский университет имени академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. Данная статья посвящена проблеме детектирования лиц на изображениях. Основное внимание было уделено поиску наиболее точного и оптимального метода детектирования, параметрам, влияющим на точность детектирования, и подбору пары каскадов Хаара, ввиду нецелесообразности их самостоятельного использования.

1. Введение

В настоящее время в связи со стремительным развитием социальных сетей и разных систем наблюдений стала актуальной задача детектирования объектов, в частности - лиц, на изображениях или видео (например, в охранных системах слежения). Все чаще можно встретить тот или иной метод детектирования лица на изображении. Рассмотрим некоторые из уже существующих методов:

Метод гибкого сравнения эластичных графов (Elastic graph matching)

Суть метода заключается в эластичном сопоставлении графов, описывающих изображения лиц. [14][19] Эти лица представляются в виде графов, с взвешенными вершинами и ребрами. Один граф (эталонный) всегда остается неизменным, когда другой изменяется с целью наиболее точной подгонки к первому. В таких системах детектирования графы могут представляться как прямоугольной решеткой, так и структурой, образованной антропометрическими точками, т.е. характерными точками лица.

Нейронные сети

В настоящее время существует около десятка разновидностей нейронных сетей. Самым востребованным вариантом является сеть, построенная на многослойном перцептроне. Это позволяет классифицировать поданное на вход изображение (или сигнал) в соответствии с предварительной настройкой (обучением) сети. [21-22]

Обучение нейронных сетей происходит на наборе обучающих примеров. Суть заключается в настройке весов межнейронных связей в процессе решения оптимизационной задачи методом градиентного спуска. В процессе обучения нейронных сетей происходит автоматическое извлечение ключевых признаков, определение их важности и построение взаимосвязей между ними. Ожидается, что обученная нейронная сеть при помощи уже накопленного опыта будет применять опыт на неизвестные образы за счет обобщающих способностей.

Скрытые Марковские модели (СММ, НММ)

Одним из статистических методов детектирования лиц являются скрытые Марковские модели с дискретным временем. [15-17] Эти модели используют статистические свойства сигналов и учитывают непосредственно их пространственные характеристики. Элементами модели являются: множество скрытых состояний, множество наблюдаемых состояний, начальная вероятность состояний, матрица переходных вероятностей, где каждому соответствует своя Марковская модель. При детектировании объекта проверяются Марковские модели,

сгенерированные для заданной базы объектов, а также ищется максимальная вероятность (из наблюдаемых) того, что последовательность наблюдений для данного объекта и модель совпадают. На сегодняшний день примера коммерческого применения скрытых Марковских моделей для детектирования лиц нет.

Active Appearance Models (AAM) и Active Shape Models (ASM)

Active Appearance Models (AAM) [23]

Активные модели внешнего вида (Active Appearance Models, AAM) — это статистические модели изображений, которые разного рода деформаций могут быть подогнаны под реальное изображение путем разного рода. Она содержит два типа параметров: параметры формы, и параметры внешнего вида, связанные со статистической моделью пикселей изображения или текстурой. Перед использованием модель обучается на множестве заранее размеченных изображений, где разметка производится вручную. Каждая метка имеет свой номер и определяет характерную точку, которую модель должна будет найти во время адаптации к новому изображению.

Active Shape Models (ASM) [18]

Суть метода ASM заключается в учете статистических связей между расположением антропометрических точек на имеющейся выборке изображений лиц, снятых в «анфас». Эксперт размечает расположение антропометрических точек на каждом изображении, причем точки пронумерованы в одинаковом порядке. Для того чтобы привести координаты на всех изображениях к единой системе обычно выполняется обобщенный прокрустов анализ, в результате которого все точки приводятся к одному масштабу и центрируются, далее для всего набора образов вычисляется средняя форма и матрица ковариации. На основе этой матрицы вычисляются собственные вектора, которые затем сортируются в порядке убывания соответствующих им собственных значений. Модель ASM определяется матрицей Φ и вектором средней формы \bar{s} .

Целью создания данного алгоритма детектирования являлась попытка оптимизировать работу уже известного метода, посмотреть, как будет вести себя алгоритм при использовании пар каскадов и изучить влияние параметров на верное детектирование.

2. Каскады Хаара

Каскад Хаара — способ обнаружения объектов на изображении, основанный на машинном обучении, идея которого была предложена в статье за авторством Пола Виолы (Paul Viola) и Майкла Джонса (Michael Jones) [1-2]. Принимая на вход изображение, обученный каскад Хаара определяет, есть ли на нем искомый объект, т.е. выполняет задачу классификации, разделяя входные данные на два класса (есть искомый объект, нет искомого объекта). Правильно обученный каскад Хаара имеет хорошую скорость выполнения классификации, а также неплохую устойчивость к отклонениям разного рода. [4-6]

Признаки Хаара

Признак – отображение $f: X \Rightarrow D_f$, где D_f — множество допустимых значений признака. Вектор признаков $x = (f_1(x), \dots, f_n(x))$ называется признаковым описанием объекта $x \in X$, если заданы признаки f_1, \dots, f_n . Признаковые описания допустимо отождествлять с самими объектами. При этом множество $X = D_{f_1} * \dots * D_{f_n}$ называют признаковым пространством. Признаки делятся на следующие типы в зависимости от множества D_f :

- a) бинарный признак, $D_f = \{0, 1\}$;
- b) номинальный признак: D_f — конечное множество;
- c) порядковый признак: D_f — конечное упорядоченное множество;
- d) количественный признак: D_f — множество действительных чисел.

Для решения прикладных задач с разнотипными признаками подходят далеко не все методы.

В стандартном методе Виолы – Джонса используются прямоугольные признаки, изображенные на рисунке ниже, они называются примитивами Хаара 'Рисунок 1':

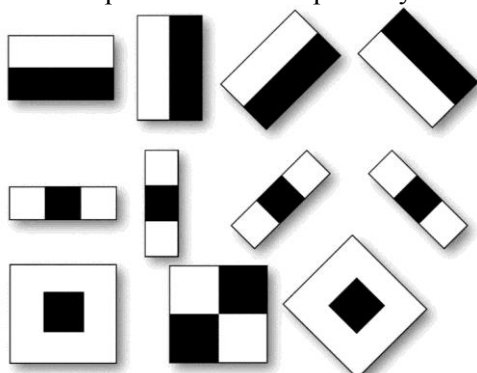


Рисунок 1. «Примитивы Хаара».

В расширенном методе Виолы – Джонса, используемом в библиотеке OpenCV используются дополнительные признаки 'Рисунок 2':

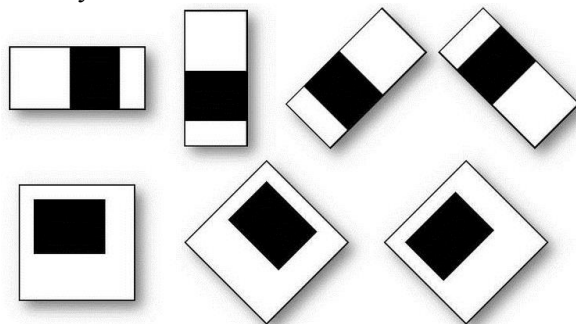


Рисунок 2. «Дополнительные примитивы Хаара».

Вычисляемым значением такого признака будет

$$F = X - Y,$$

где X – сумма значений яркостей точек закрываемых светлой частью признака, а Y – сумма значений яркостей точек закрываемых темной частью признака.

Масштабный фактор (scalefactor) – параметр, который определяет размер изображения при каждом его шаге. Он используется для создания масштабной пирамиды (представление изображений таким образом, чтобы мы могли обнаружить как маленькие, так и большие лица, используя одно и то же окно обнаружения). От его значения зависит тщательность исследования каждой области.

Размер скользящего окна (minsize) – параметр, определяющий минимальный возможный размер объекта, т.е. области меньше его игнорируются. Этот размер устанавливается самостоятельно.

Параметр, определяющий количество возможных соседей (minNeighbors) – отвечает за количество соседей, которое может иметь каждый прямоугольник. Этот параметр влияет на качество обнаруженных лиц: чем больше значение, тем меньше обнаружений, но выше качество. Оптимальные значения для него 3-6.

3. Результаты исследования

Для проведения исследований был составлен набор из 100 различных фотографий одного разрешения (1600 x 1200 пикселей), который содержит в себе различные изображения с количеством лиц от 1 до 10 и несколько фотографий толпы. [3] Число лиц на каждом

изображении было оценено с использованием экспертной оценки. Общее число лиц на всём исследуемом наборе данных варьировало между одним, двумя, группой от 3 до 6 человек и неопределенным количеством лиц - толпой. Эксперименты проводились с использованием библиотеки OpenCV [7-10] и входящих в её состав подготовленных каскадов Хаара[20]. При проведении экспериментов значение параметров масштабирования равнялось 1.2, скользящее окно квадратной формы с длиной стороны 30 пикселей и параметром, определяющим количество возможных соседних окон равного 3.

Первый эксперимент заключался в оценке точности детектирования объектов на изображениях с использованием каждого из готовых каскадов Хаара. Полученные результаты представлены в таблице 1.

Таблица 1. Результаты детектирования объектов с использованием отдельных каскадов.

| Название Каскада | eye | eye_tree_eyeglasses | frontalcatface_extended | Frontalface | frontalface_alt | frontalface_alt2 | frontalface_alt_tree | frontalface_default | fullbody | lefteye_2splits | lowerbody | profileface | smile | upperbody |
|--------------------------------------|-----|---------------------|-------------------------|-------------|-----------------|------------------|----------------------|---------------------|----------|-----------------|-----------|-------------|-------|-----------|
| Полученные результаты детектирования | 20 | 23 | 15 | 15 | 36 | 36 | 19 | 43 | 12 | 26 | 13 | 20 | 0 | 13 |

Анализируя результаты первого эксперимента, можно сделать выводы:

- Одиночные каскады при использовании указанных параметров не позволяют детектировать лица с достаточной точностью.
- Каскад “smile” самостоятельно использовать не стоит, т.к. он не дает ни одного точного детектирования. Его можно использовать для уточнения результатов уже найденных областей другими каскадами.
- Если использовать каскад для детектирования глаз (каскад “eye”), то все результаты нужно проверять лично или делить пополам, т.к. когда находятся оба глаза на одном лице, он считает его за два разных. Его также можно использовать для уточнения детектирования найденной области другими каскадами.

Второй эксперимент заключался в оценке точности детектировании лиц при использовании различных пар каскадов Хаара. В рамках эксперимента использовались только каскады Хаара, используемые для детектирования лиц. Поскольку каскады применяются независимо, для проведения второго эксперимента необходимо было реализовать алгоритм объединения результатов работы, полученных с использованием различных каскадов. Каждый из каскадов возвращает множество областей изображения, соответствующих детектируемым объектам. Поэтому чтобы исключить двойной учёт одного и того же объекта, был разработан и реализован соответствующий алгоритм:

Брались верхний левый и правый нижний углы двух найденных областей и сравнивались по положению относительно друг друга по двум пунктам:

1. Правый нижний угол первого выше или левее верхнего левого второго
2. Левый верхний первого ниже или правее нижнего правого второго

Если выполнялось хоть одно из этих условий, то прямоугольники не пересекаются.

Ниже представлена схема алгоритма (рисунок 3), которая описывает действия, производимые над изображением, перед тем, как в результате оно будет сохранено в новой папке с новым названием. Также приведена таблица с указанием используемых каскадов и результатов обработки.

В подсчете найденных лиц применялось условие: если выделенные области пересекаются между собой, то в конечный список добавляется одна область большая по размеру. Если же области не имеют пересечений, то добавляем те, которые не имели пересечений ни с одним прямоугольником. Это позволяет избежать принятия одного и того же лица за несколько разных.

Таблица 2. Группировка каскадов.

| Использованные каскады | frontalcatface_exten ded | frontalface | frontalface_alt | frontalface_alt2 | frontalface_alt_tree | frontalface_default | profileface |
|--------------------------------|-----------------------------|-------------|-----------------|------------------|----------------------|---------------------|-------------|
| frontalcatface_extended | - | 16 | 41 | 36 | 21 | 37 | 23 |
| frontalface | 16 | - | 37 | 35 | 18 | 36 | 22 |
| frontalface_alt | 41 | 37 | - | 39 | 34 | 38 | 41 |
| frontalface_alt2 | 36 | 35 | 39 | - | 35 | 36 | 27 |
| frontalface_alt_tree | 21 | 18 | 34 | 35 | - | 36 | 23 |
| frontalface_default | 37 | 36 | 38 | 36 | 36 | - | 35 |
| profileface | 23 | 22 | 41 | 27 | 23 | 35 | - |

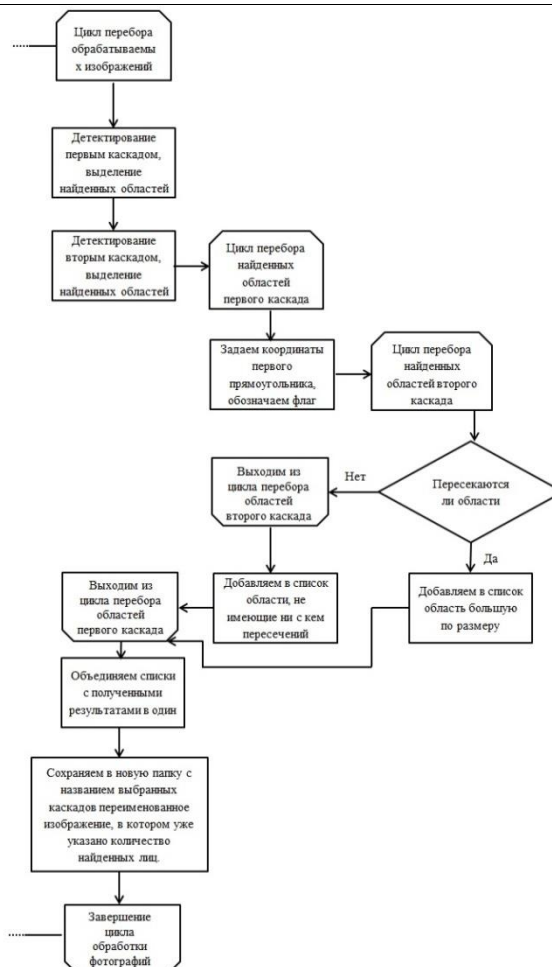


Рисунок 3. Схема алгоритма детектирования.

4. Выводы

Были рассмотрены существующие методы детектирования лиц и исследованы разные комбинации и параметры для более точного распознавания. Также было выяснено, что использование каскадов самостоятельно нецелесообразно, т.к. они дают малую точность (например, “smile” не дает ни одного верного определения). При комбинировании каскадов нужно учитывать особенности каждого (например, какие признаки используются для поиска, значения параметров и т.д.), чтобы детектирование было более качественным. Обязательно надо учесть возможность пересечения найденных областей, чтобы не посчитать одно и то же лицо дважды. Возможно, что использование третьего каскада (например, “eye” или “smile”) для определения лиц на уже найденных областях даст лучшие результаты. Также нужно исследовать влияние значений параметров на качество детектирования изображений.

5. Литература

- [1] Viola, P. Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M.J. Jones // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2001.
- [2] Viola, P. Robust real-time face detection / P. Viola, M.J. Jones // Journal of Computer Vision. – 2004. – Vol. 57(2). – P. 137-154.
- [3] Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.
- [4] Местецкий, Л.М. Математические методы распознавания образов. – М.: МГУ, ВМиК, 2002-2004. – С. 42-44.
- [5] Sochman, J. AdaBoost / J. Sochman, J. Matas // Center for Machine Perception, Czech Technical University, Prague, 2010.
- [6] Freund, Y. A Short Introduction to Boosting / Y. Freund, R.E. Schapire. – Shannon Laboratory, USA, 1999. – P. 771-780.
- [7] Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. – 2004. – 928 с.
- [8] Sirovich, L. Low dimensional procedure for characterization of human faces / L. Sirovich, M. Kirby // Journal of the Optical Society of America A. – 1987. – Vol. 4. – P. 519.
- [9] Turk, M. Face recognition using eigenfaces / M. Turk, A. Pentland // Proc. IEEE Conference on Computer Vision and Pattern Recognition. – 1991. – P. 586-591.
- [10] Beymer, D. Face Recognition From One Model View / D. Beymer, T. Poggio // Proc. Fifth Int'l Conf. Computer Vision, 1995.
- [11] Viola, P. Rapid object detection using a boosted cascade of simple features / P. Viola, M.J. Jones // Accepted Conference on Computer Vision and Pattern Recognition, 2001.
- [12] Vetter, T. Linear Object Classes and Image Synthesis From a single Example Image / T. Vetter, T. Poggio // IEEE Trans. on PAMI. – 1997. – Vol. 19(7). – P. 733-742.
- [13] Zhao, W. Face Recognition: A Literature Survey / W. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips // Technical Report CS-TR-4167, Center For Automation Research, University of Maryland, 2000.
- [14] Wiskott, L. Face Recognition by Elastic Bunch Graph Matching / L. Wiskott, J.-M. Fellous, C. von der Malsburg // IEEE Trans. on PAMI. – 1997. – Vol. 19. – P. 775-799.
- [15] Двойной, И.Р. Методы распознавания изображения лица человека по цветовым признакам и идентификации личности на основе скрытых Марковских моделей в системах видеонаблюдения, 2013. – 150 с.
- [16] Гульятеева, Т.А. Применение скрытых Марковских моделей с одномерной топологией к задаче распознавания лиц // Российская научно-техническая конференция “Информатика и проблема телекоммуникации”. – Новосибирск: СибГУТИ. – 2006. – Т. I. – С. 150-154.
- [17] Ara, V. Face Detection and Recognition Using Hidden Markov Models / V. Ara, H. Monson, 1999. – P. 5.
- [18] Zhao, W.Y. Imagebased Face Recognition Issues and Method / W.Y. Zhao, C. Rama, 2002. – P. 25.

- [19] Wiskott, L. Face Recognition by Elastic Bunch Graph Matching / L. Wiskott, J.-M. Fellous, N. Kruger, C. von der Malsburg // Intelligent Biometric Techniques in Fingerprint and Face Recognition. – Publ. CRC Press, 1999. – Vol. 11. – P. 355-396.
- [20] Каскады Хаара [Электронный ресурс]. – Режим доступа: <https://github.com/opencv/opencv/tree/master/data/haarcascades>.
- [21] Duffner, S. Face Image Analysis With Convolutional Neural Networks, 2008. – P.176.
- [22] Lawrence, S. Face Recognition: A Convolutional Neural-Network Approach / C.L. Giles, A.Ch. Tsoi, A.D. Back // IEEE Transactions on Neural Networks. – 1997. – Vol. 8(1).
- [23] Edwards, G.J. Wolfson Image Analysis Unit / G.J. Edwards, T.F. Cootes, C.J. Taylor // Computer Vision – ECCV. – 1998. – P. 581-595.

Development of face detection algorithm using combinations of Haar cascades

E.A. Rudinskaya¹, R.A. Paringer¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. This paper is devoted to the problem of detecting faces in images. The main attention was paid to the search for the most accurate and optimal recognition method, parameters affecting the accuracy of detection on images; selection of a pair of Haar cascades, due to the inexpediency of their independent use.