

УДК 004.93

АВТОМАТИЗАЦИЯ РАСПОЗНАВАНИЯ ЛИЦ ЛЮДЕЙ ПО ФОТОГРАФИИ

И.В. Тимохин, Н.Б. Осипенко

Гомельский государственный университет им. Ф. Скорины

AUTOMATION OF PEOPLE'S FACES RECOGNITION BY PHOTOGRAPH

I.V. Tsimokhin, N.B. Osipenko

F. Scorina Gomel State University

Описывается алгоритм распознавания людей по фотографиям их лиц, состоящий из двух шагов. На первом шаге на исходных изображениях определяются области, содержащие лица; на втором шаге из полученных областей формируются признаковые описания людей, используемые затем для сравнения их между собой и принятия решения о наличии одинаковых или разных людей на фотографиях.

Ключевые слова: распознавание лиц, метод Виолы-Джонса, метод главных компонент.

An algorithm for faces recognition is described. The algorithm has two stages. The aim of the first stage is to detect faces on source images; the aim of the second stage is to create attributes describing detected faces. Those attributes are used to compare faces and decide whether there is the same or different person on different images.

Keywords: face recognition, Viola-Jones method, principal component analysis.

Введение

Создание фотографий в последнее время не требует больших затрат, объемы обработки изображений растут. Одно из действий, которое может применяться при обработке множества изображений – выделение информации о людях, находящихся на изображении. Такое действие требуется выполнять в различных сферах человеческой деятельности, а из-за доступности фотографий, эти сферы не ограничиваются только такими профессиональными сферами, как криминалистика, журналистика и т. п., но также затрагивают и повседневный быт.

Задача, описываемая в статье, может быть сформулирована следующим образом: задан набор изображений, которые могут содержать лица различных людей; требуется получить список людей на изображениях и для каждого человека список изображений, на которых они находятся. Рассматриваемая задача состоит из нескольких подзадач. Для каждого изображения требуется определить, есть ли на нём изображение лица, и если есть, то получить содержащую его область. Эта подзадача является частным случаем задачи нахождения объектов на изображениях (object detection), для которой существуют разные решения, включая метод Виолы-Джонса [1]. Существуют различные модификации метода Виолы-Джонса, включая возможность поворота используемых признаков на 45 градусов [2]. Второй подзадачей является определение: одинаковые или разные люди приведены на различных изображениях. Эта подзадача может быть решена различными способами, одним из которых является применение метода главных компонент [3], [4].

1 Описание алгоритма

Алгоритм для распознавания людей на фотографиях состоит из двух шагов. На первом шаге выделяются области, содержащие лица. На втором шаге из выделенных областей формируются признаки, характеризующие лицо, и полученные признаки используются для сравнения лиц между собой и принятия решений о принадлежности двух различных лиц на фотографиях одному и тому же человеку.

Для первого шага алгоритма применен метод Виолы-Джонса. Была использована реализация метода из библиотеки OpenCV [5], которая предоставляет возможность настраивать параметры, используемые при обучении модели. Параметры, при которых модель дает наилучший результат, определяются эмпирически. Метод Виолы-Джонса, по сравнению с некоторыми другими методами, является более быстрым за счет использования так называемых интегральных изображений [1]. Для обучения модели и оценки ее точности в качестве выборки положительных образцов использовалась выборка лиц Georgia Tech face database [6], содержащая по 15 изображений 50 людей, сделанных в течение двух-трех фотосессий. Кроме самих изображений выборка содержит координаты областей, содержащих лица. Общий размер данных в выборке положительных образцов составляет 128 МБ. В качестве выборки отрицательных образцов была использована выборка, составленная из изображений, взятых с web-сайта reXels.com, не содержащих лиц, и доступная для применения в коммерческих и некоммерческих целях [7]. В выборке отрицательных образцов содержалось 134

изображения в высоком разрешении (до 12626 на 8417 пикселей, медиана размера файлов – 3,6 МБ). Общий размер данных в выборке составляет 661 МБ.

Для второго шага алгоритма использовался метод главных компонент [8], который позволяет упрощать описание выборочных данных за счет уменьшения числа описывающих объекты признаков. Для сравнения объектов с меньшим числом признаков было выбрано евклидово расстояние. Различные лица соотносятся с одним и тем же человеком, если расстояние между ними меньше заданной константы, которая определяется эмпирически в процессе построения модели. Для обучения модели и оценки ее точности использовалась выборка лиц, полученных из выборки Georgia Tech face database после первого этапа алгоритма.

Результатом выполнения обоих этапов алгоритма в рабочем режиме будет множество лиц, для каждого из которых известно на каких изображениях оно присутствует. Так как в рамках алгоритма невозможно определить имена или какие-то другие понятные идентификаторы для людей на изображениях, то для конечного приложения был создан графический интерфейс с использованием стандартной библиотеки Tkinter языка программирования Python, позволяющий вводить текстовые метки для распознанных лиц. На рисунке 1.1 приводится блок-схема работы конечного приложения.

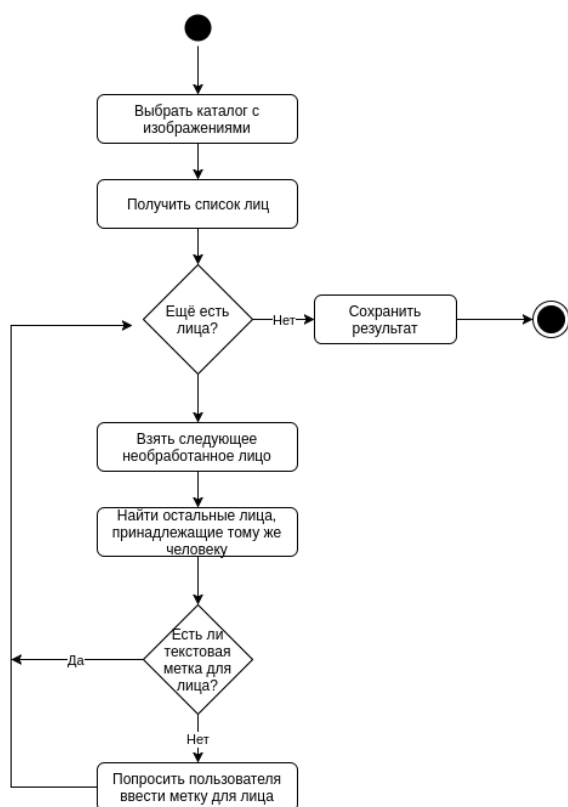


Рисунок 1.1 – Блок-схема конечного приложения

2 Модель для определения областей, содержащих лица

При построении, обучении и верификации модели для определения областей, содержащих лица, использовалась реализация метода Виолы-Джонса из библиотеки OpenCV [5].

Метод Виолы-Джонса позволяет находить на изображениях объекты, которые задаются набором признаков Хаара. Набор признаков, описывающих требуемый объект, получается в процессе обучения модели. Численным значением признака Хаара является разность между суммой пикселей в светлой и тёмной областях признака Хаара, накладываемого на используемое изображение. Изображения признаков Хаара приведены на рисунке 2.1. Для поиска объектов на большом изображении перебираются все различные области с последующей передачей в классификатор Виолы-Джонса, который дает заключение, является ли сканируемая область требуемым объектом или нет. При обучении классификатора из множества всех возможных признаков Хаара выбираются такие, которые разделяют выборку на положительные (содержащие объект) и отрицательные (не содержащие объект) образцы.

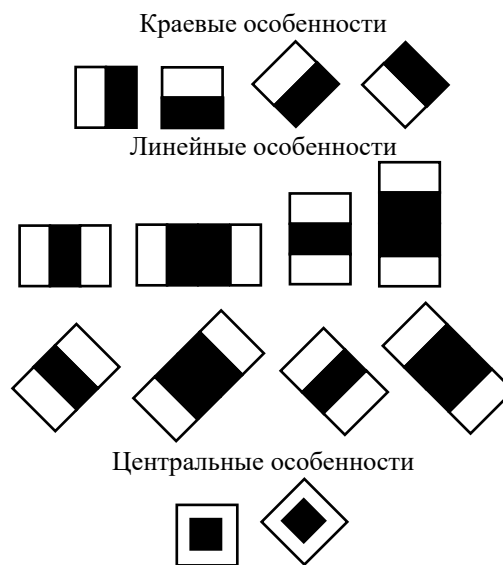


Рисунок 2.1 – Признаки Хаара [2]

Библиотека OpenCV предоставляет возможность задавать значения параметров, используемых при обучении модели. Параметры w , h отвечают за ширину и высоту окна, в котором находятся признаки Хаара. При выполнении работы окна выбирались квадратного размера, поэтому параметры w и h равны. Параметр $mode$ задает набор используемых признаков. Если значение $mode$ равняется BASIC, то используются только основные признаки (без наклона). Для использования всех признаков значение параметра должно равняться ALL. Для задания числа итераций обучения применяется параметр numStages. Число

итераций равняется числу каскадов в созданном классификаторе. Параметры numPos и numNeg используются для задания числа используемых положительных и отрицательных образцов, которые выбираются на каждом этапе соответственно; это число может отличаться от числа образцов в выборке.

Для определения зависимости между параметрами, используемыми при обучении, и точностью получаемой модели было обучено несколько моделей с различными параметрами. Результаты их обучения приведены в таблице 2.1. Все

модели обучались с параметром numPos равным 500, при размере выборки положительных образцов равным 750. Выборка отрицательных образцов содержала 134 образца.

Можно заметить, что при одинаковых остальных параметрах точность определения лица зависит от параметра numStage следующим образом: с увеличением параметра numStage уменьшается точность верного определения лица на изображении и увеличивается точность верного определения отсутствия лица на изображении. Это можно видеть на рисунке 2.2.

Таблица 2.1 – Оценка точности классификаторов в зависимости от параметров $w, h, mode, numNeg, numStages$

№	Значения параметров				Оценка точности	
	w, h	mode	numNeg	numStages	Верно найдено, %	Верно не найдено, %
1	10	BASIC	100	10	91,06	5,36
2	10	BASIC	100	15	11,73	69,64
3	10	BASIC	100	20	4,13	85,71
4	10	BASIC	134	10	95,07	1,79
5	10	BASIC	134	15	33,07	50,00
6	10	BASIC	134	20	2,80	96,43
7	10	ALL	100	10	96,80	1,79
8	10	ALL	100	15	33,33	23,21
9	10	ALL	100	20	8,80	82,14
10	10	ALL	134	10	92,13	3,57
11	10	ALL	134	15	21,47	48,21
12	10	ALL	134	20	2,00	94,64
13	20	BASIC	100	10	100,00	0,00
14	20	BASIC	100	15	90,80	14,28
15	20	BASIC	100	20	56,26	71,42
16	20	BASIC	134	10	100,00	0,00
17	20	BASIC	134	15	95,47	7,14
18	20	BASIC	134	20	61,20	62,50
19	20	ALL	100	10	100,00	0,00
20	20	ALL	100	15	96,93	7,14
21	20	ALL	100	20	75,33	57,14
22	20	ALL	134	10	100,00	0,00
23	20	ALL	134	15	82,67	30,36
24	20	ALL	134	20	37,73	87,50

Таблица 2.2 – Оценка точности классификаторов в зависимости от параметров $w, h, mode, numNeg$ при выборе оптимального numStages

№	Параметры			Оптимальное numStage	Точность	
	w, h	mode	numNeg		Верно найдено, %	Верно не найдено, %
1	10	BASIC	100	14	34,26	41,07
2	10	BASIC	134	14	39,6	32,14
3	10	ALL	100	15	33,33	23,21
4	10	ALL	134	14	26,8	28,57
5	20	BASIC	100	19	54,00	66,07
6	20	BASIC	134	20	61,20	62,50
7	20	ALL	100	21	70,27	57,14
8	20	ALL	134	17	59,60	60,71
9	25	BASIC	100	21	65,07	62,5
10	25	BASIC	134	21	63,33	71,43
11	25	ALL	100	23	60,80	71,43
12	25	ALL	134	21	63,07	67,86

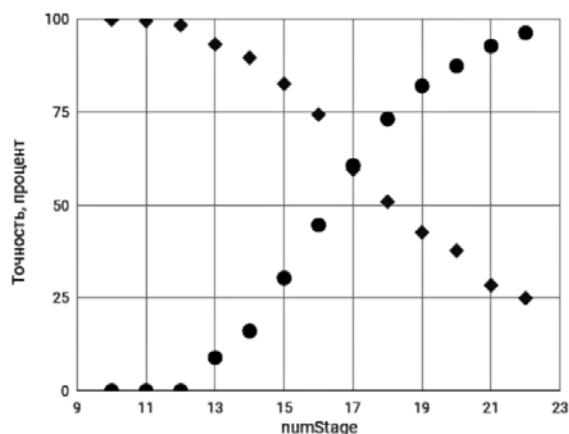


Рисунок 2.2 – Точность классификатора в зависимости от параметра numStage (◆ – точность верного определения лица; ● – точность верного определения отсутствия лица)

Таким образом, для каждого набора параметров можно подобрать такое оптимальное numStage, чтобы точность верного определения присутствия лица и точность верного определения отсутствия лица отличались минимально.

С одной стороны, можно заметить, что с увеличением значений параметров w и h , вообще говоря, увеличивается точность распознавания. С другой стороны, понятно, что с увеличением размеров окна (а именно за его размеры отвечают эти параметры) уменьшается число объектов, которые могут быть найдены; т.к. объекты меньшего размера чем окно не будут определяться. В качестве конечной модели была выбрана та, которая обучалась с параметрами w, h равными 20, mode равным BASIC, numNeg равным 134.

3 Модель для идентификации лиц

Изображения лиц, переданные в модель, уменьшались до размера 20 на 20 пикселей, после чего представлялись в виде вектора. Таким образом, каждый элемент вектора соответствует одному пикселю изображения. Понятно, что не каждый пиксель изображения содержит информацию, необходимую для идентификации лица. Кроме того, некоторые пиксели могут дублировать информацию. Для того, чтобы уменьшить размерность данных и оставить только значимые признаки, использовался метод главных компонент. В методе главных компонент для построения признакового описания объекта в пространстве меньшей размерности имеющийся вектор описания исходного объекта умножается на матрицу, составленную из нескольких первых собственных векторов матрицы ковариации признаков. Для выполнения операций над матрицами использовались функции библиотеки Numpy [9]. Для идентификации объекты с меньшим

числом признаков сравнивались между собой с использованием евклидова расстояния. Близкие по расстоянию вектора малой размерности считались соответствующими изображениями лиц одного и того же человека.

При создании итоговой модели осуществлялось тестирование моделей при разных числах компонент. Также каждая модель проверялась в зависимости от того, применялся ли поворот по линии между центрами глаз или нет. В таблице 3.1 приводятся полученные результаты.

Таблица 3.1 – Оценка точности распознавания при разных параметрах

Число компонент	Верное распознавание, %	
	Без поворота	С поворотом по линии между центрами глаз
1	2	2
1	9,16	8,37
2	15,54	15,94
3	30,68	29,08
4	44,23	44,62
5	44,62	45,82
6	46,61	47,01
7	47,81	50,60
8	51,00	54,18
9	52,99	53,39
10	53,78	53,78
11	54,58	53,78
12	53,78	55,79
13	56,57	54,58
14	57,37	56,97
15	56,18	58,96
16	59,36	60,56
17	59,36	60,96
18	60,56	60,56
19	62,15	59,76
20	62,55	59,76
21	61,75	60,56

Можно видеть, что до определенного момента с увеличением числа компонент увеличивается точность работы алгоритма. Кроме того, было замечено, что, несмотря на то, что при малых значениях числа компонент поворот изображений по линии между центрами глаз улучшает точность, с увеличением числа компонент такое предварительное преобразование изображений перестает давать лучший результат.

В качестве конечной модели была выбрана та, которая при использовании метода главных компонент уменьшает размерность до 20 признаков и не применяет преобразования к изображениям. При программной реализации использовалась библиотека Numpy [9] для языка Python, предоставляющая функции, связанные с работой с векторами и матрицами.

Заключение

В статье описан один из вариантов решения задачи распознавания лиц на изображениях, состоящий из двух этапов. Каждый из этапов рассмотренного алгоритма реализовывался независимо от другого. Для различных вариантов моделей эмпирически были подобраны параметры, при которых достигается большая точность. Описаны выявленные закономерности между значениями параметров и точностью получаемой модели.

ЛИТЕРАТУРА

1. *Viola, P. Robust Real-time Object Detection / Paul Viola and Michael Jones // International Journal of Computer Vision. – 2001. – № 4. – P. 34–47.*
2. *Leinhart, R. An Extended Set of Haar-like Features for Rapid Object Detection / R. Lienhart, J. Maydt // ICIP02. – 2002. – P. 900–903.*
3. *Turk, M. Face recognition using eigenfaces / M.A. Turk, A.P. Pentland // Computer Vision and Pattern Recognition. – 1991. – P. 586–591.*
4. *Delac, K. Appearance-based Statistical Methods for Face Recognition / K. Delac, M. Grgic,*

P. Liatsis // Proceedings of the 47th International Symposium ELMAR-2005 focused on Multimedia Systems and Applications – Zagreb. – 2005 – P. 151–158.

5. *OpenCV Library* [Electronic resource] – URL: <http://opencv.org/>. Date of access: 01.05.2017.

6. *Georgia Tech face database* [Electronic resource] – URL: http://www.anefian.com/research/face_reco.htm. – Date of access: 10.03.2017.

7. *Pexels. Free stock photos* [Electronic resource] – URL: <https://www.pexels.com>. – Date of access: 01.05.2017.

8. *Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности: справочное издание / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков. – М.: Финансы и статистика, 1989. – 605 с.*

9. *Numpy and Scipy Documentation* [Electronic resource] – URL: <https://docs.scipy.org/doc/>. – Date of access: 01.05.2017.

Поступила в редакцию 28.06.17.