

ПОДБОР ПАРАМЕТРОВ АЛГОРИТМА ВИОЛЫ–ДЖОНСА ПОД КОНКРЕТНЫЕ УСЛОВИЯ

А. Д. Егоров, А. Н. Штанько, П. Е. Минин

В статье рассматривается поиск лиц на изображении методом Виолы–Джонса. Работа метода определяется рядом параметров, влияющих на баланс между количеством ложных срабатываний и пропущенных лиц, а также на быстрдействие метода. Подбор параметров под практические задачи является достаточно трудоёмким, так как для оптимального выбора требуется множество экспериментов на большом объёме данных. Данное исследование призвано дать априорную информацию о влиянии параметров на работу метода, используя которую можно достаточно эффективно подбирать параметры без необходимости в предварительном сборе тестового множества данных.

Ключевые слова: компьютерное зрение, Виола–Джонс, поиск объектов, поиск лиц.

Введение. Анализ изображения и видеопотока на наличие тех или иных объектов является одной из важнейших задач в обеспечении безопасности периметра охраняемого объекта, общественного правопорядка, дорожного движения [1].

Типичной задачей как в проектах в области обеспечения безопасности, так и в бизнес проектах является задача поиска объектов на изображении, в частности, задача нахождения людей (лиц людей) на изображении. Однако поиск сложных объектов, таких как лица, не является простой задачей из-за большого разнообразия вариантов освещения и положения объекта в пространстве.

За прошедшие несколько десятков лет с момента появления науки компьютерное зрение было создано множество различных алгоритмов поиска объектов. Метод Виолы–Джонса [2, 3], предложенный П. Виолой и М. Джонсом в 2001 году, является универсальным методом для поиска объектов на изображении. Этот метод основан на прин-

ципе скользящего окна и классификации слабыми классификаторами, объединенными в каскад, что позволяет быстро отсеивать большинство кандидатов.

Результат применения суммарного классификатора на основе алгоритма AdaBoost и является результатом применения алгоритма Виолы–Джонса к рассматриваемому изображению.

Несмотря на широкое описание различных практических реализаций алгоритма Виолы–Джонса, не существует общепринятых методов для подбора параметров алгоритма под конкретные условия (освещенность, максимально расстояние поиска и т.д.). Цель данной работы – подбор наиболее оптимальных для конкретной ситуации параметров алгоритма Виолы–Джонса.

Метод Виолы–Джонса. Первым шагом для работы метода, в случае если используются признаки Хаара – исследуемое изображение (матрица) I преобразуется в интегральное изображение, которое представляет собой некоторую матрицу L , размеры которой совпадают с размерами исходного изображения. При этом каждый элемент матрицы L является суммой элементов I , расположенных выше и левее данного элемента матрицы L :

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

где $I(i, j)$ – яркость (интенсивность) соответствующих пикселей исходного изображения. Кроме того, каждый элемент матрицы можно рассчитать, используя следующее рекуррентное соотношение:

$$L(x, y) = I(x, y) - L(x - 1, y - 1) + L(x, y - 1) + L(x - 1, y).$$

Главным преимуществом использования интегрального изображения является очевидная возможность быстрого подсчёта суммы внутри любого прямоугольного (с вершинами A, B, C, D , которые можно видеть на рис. 1(а)) элемента исходного изображения I . Для этого можно воспользоваться следующей формулой [4]:

$$S(ABCD) = L(A) + L(C) - L(B) - L(D).$$

Характеристики Хаара. Характеристика Хаара – функция f от суммарной интенсивности I_A и I_B двух прямоугольных участков изображения A и B , причём участок B вложен в участок A . Таким образом, для расчёта суммарной интенсивности можно

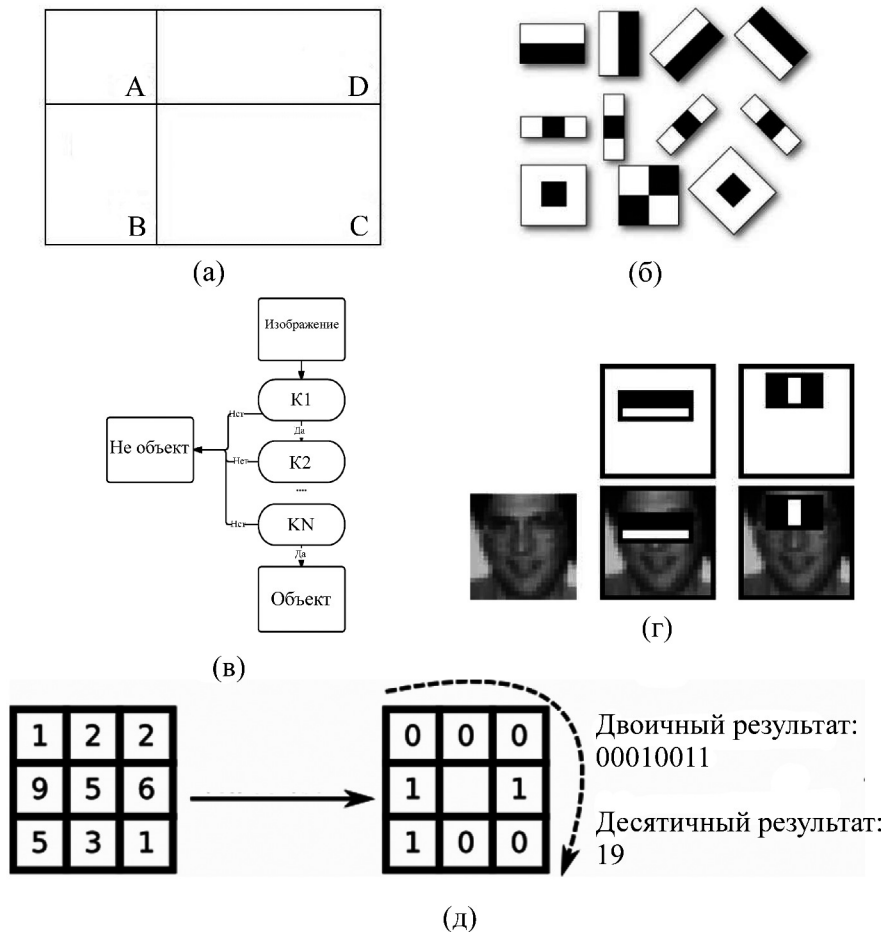


Рис. 1: Части алгоритма Виолы-Джонса: (а) Комбинация областей. (б) Признаки Хаара. (в) Каскадный классификатор. (г) Применение признаков Хаара к изображению. (д) Признак LBP.

использовать интегральное изображение. В современных работах используется характеристика вида $f(A, B) = aI_A + bI_B$, где a и b константы [5] и комбинации как из двух, так и из трех и более прямоугольников.

Характеристики или примитивы Хаара являются прямоугольными признаками, которые используются в методе Виолы-Джонса. Некоторые стандартные признаки Хаара представлены на рис. 1(б).

Характеристики LBP. В ряде работ[6–8] вместо признаков Хаара было предложено использовать признаки, построенные на основе локальных бинарных шаблонов (LBP). Локальный бинарный шаблон рассчитывается для каждого конкретного пикселя и представляет собой один байт информации. Характеристика для каждого пикселя

рассчитывается на основании данных о соседях пикселя в окрестности 3×3 с рассматриваемым пикселем в центре. Пусть мы рассматриваем пиксель с координатами x и y на исходном изображении I , тогда значение шаблона LBP_{xy} рассчитывается по следующей формуле:

$$\begin{aligned} LBP_{xy} = & 1 \cdot (I(x-1, y-1) > I(x, y)) + 2 \cdot (I(x, y-1) > I(x, y)) + 4 \cdot (I(x+1, y-1) > I(x, y)) + \\ & + 8 \cdot (I(x+1, y) > I(x+1, y+1)) + 16 \cdot (I(x+1, y+1) > I(x, y)) + 32 \cdot (I(x, y+1) > I(x, y)) + \\ & + 64 \cdot (I(x-1, y+1) > I(x, y)) + 128 \cdot (I(x-1, y) > I(x, y)). \end{aligned}$$

Пример расчёта можно видеть на рис. 1(д).

Алгоритм агрегации слабых классификаторов. Как было сказано ранее, в случае, если ведётся работа с характеристиками или признаками Хаара (со «слабыми» классификаторами), необходимо использовать очень большое число классификаторов для того, чтобы сделать вывод – является ли рассматриваемая область искомым объектом или нет.

Для выбора конкретных используемых признаков Хаара и установления пороговых уровней используется метод машинного обучения под названием AdaBoost [9]. AdaBoost комбинирует много “слабых” классификаторов с целью создания одного “сильного” классификатора. Если имеется множество слабых классификаторов, и каждый из них говорит, что рассматриваемая область – искомый объект, то существует достаточно большая уверенность в том, что результат является верным.

AdaBoost выбирает набор слабых классификаторов для объединения и присваивает каждому из них свой вес. Эта взвешенная комбинация и является сильным классификатором. Виола и Джонс объединили серии классификаторов AdaBoost как последовательность фильтров, показанных на рис. 1(в), что особенно эффективно для классификации областей изображения. Каждый фильтр является отдельным классификатором AdaBoost с достаточно небольшим числом слабых классификаторов. Пример работы алгоритма можно видеть на рис. 1(г).

Параметры реализации алгоритма Виолы-Джонса. У алгоритма Виолы-Джонса существуют следующие параметры:

- 1). Используемый каскад для поиска объектов.
- 2). Коэффициент изменения размера сканирующего окна (объекта) при переходе от одной итерации алгоритма к другой.
- 3). Количество распознанных рядом друг с другом объектов, на основании которого может быть сделан вывод о наличии в данной области искомого объекта.

4). Начальный размер окна поиска (сканирующего окна).

5). Максимальный размер окна поиска (сканирующего окна).

Для перебора всех указанных выше параметров была разработана система, в которой можно было оценить параметры алгоритма.

Эксперименты с разработанной системой и различными базами объектов. Для вычисления различных параметров алгоритма были взяты три базы лиц: FERET, MIT+CMU, Faces in the Wild. Базы были по очереди поданы на вход системы оценки параметров алгоритма. Для оценки качества работы алгоритма с текущими параметрами было использовано три показателя: количество верно найденных лиц, количество ошибочных срабатываний и среднее время обработки на одно лицо. После чего ко всей совокупности полученных данных была применена операция получения нормированного отклонения по каждому из параметров.

$$\vartheta_n = \frac{\vartheta - \bar{\vartheta}}{\sigma},$$

где ϑ – значение показателя для текущего набора параметров, $\bar{\vartheta}$ – среднее значение показателя для всех возможных комбинаций параметров, σ – стандартное отклонение.

Для выявления наилучшего набора параметров были проведены тесты для всех возможных комбинаций среди: 1. Каскад: lbr, default, alt, alt2, alt_tree (свободно распространяются в интернете).

2. Коэффициент масштабирования = 1.1, 1.2, 1.3, 1.4 (MIT+CMU); 1.1, 1.25, 1.4 (для FERET).

3. Число соседей = 1, 3, 5.

4. Начальный размер = 20, 30, 40 (MIT+CMU); 50, 80, 110 (FERET).

В результате перебора всех возможных значений были получены результаты, которые представлены на рис. 2. Каждой точке соответствует определенный набор параметров из списка выше.

На основании полученных результатов можно сделать следующие выводы:

- Каскад default даёт наибольшее количество как ложных, так и верных срабатываний, независимо от параметров. При этом каскад работает достаточно медленно.

- Каскад lbr работает максимально быстро для всех возможных параметров, однако даёт достаточно малое количество верных срабатываний, при этом не давая и большого числа ложных срабатываний.

- Наиболее точными оказываются каскады alt и alt2, которые дают достаточно много верных срабатываний, однако, при этом, работают достаточно долго, в сравнении с lbr.

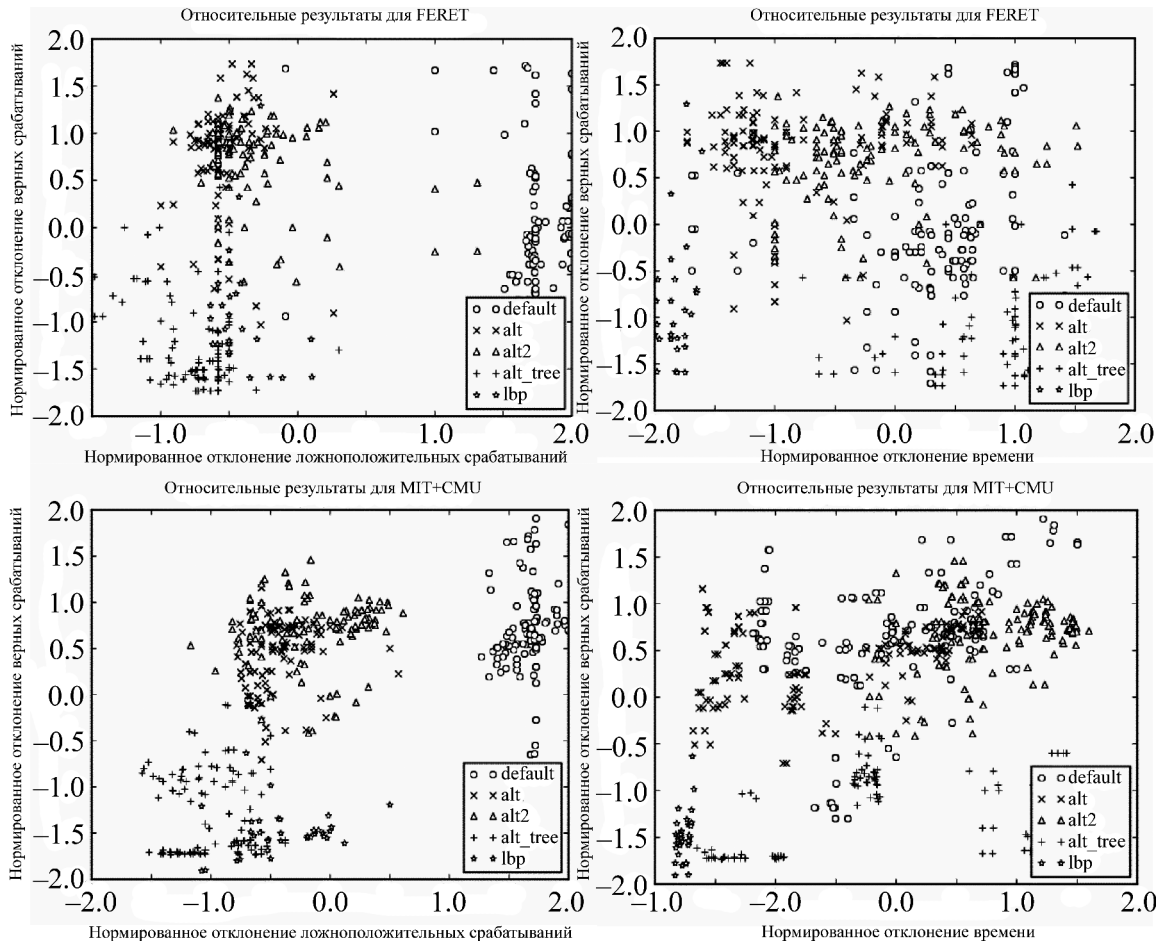


Рис. 2: Нормированные отклонения показателей.

• Каскад alt_tree дает слишком малое количество верных срабатываний, большое количество ложных, а также работает дольше остальных каскадов.

Кроме того, для базы Faces in the Wild были выявлены лучшие показатели работы каскадов с использованием F -меры в качестве главного критерия оценки качества работы.

$$F = \frac{2PR}{P + R},$$

где P – отношение количества верно найденных объектов (лиц) к общему числу найденных лиц, R – отношение количества верно найденных объектов (лиц) к общему количеству объектов в тестовом множестве (лиц в базе лиц).

Результаты сравнения можно видеть в табл. 1.

Т а б л и ц а 1

Наилучшие параметры работы алгоритма Виолы–Джонса для *Faces in the Wild*

Каскад	alt	alt	alt2	alt	alt	alt	alt	alt2
Коэффициент изменения размера	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
Количество соседей	3	3	5	3	5	3	3	5
Начальный размер	40	40	40	60	40	50	30	50
Верно найденные	1165	1165	1142	1135	1128	1148	1166	1124
Неверно найденные	57	58	35	29	22	44	73	28
Ненайденные	37	37	60	67	74	54	36	78
Двойные срабатывания	0	0	0	0	0	0	0	0
Среднее время в мс	97	91	104	40	91	61	156	70
Точность	0.953	0.953	0.970	0.975	0.981	0.963	0.941	0.976
Полнота	0.969	0.969	0.950	0.944	0.938	0.955	0.970	0.935
<i>F</i> -мера	0.961	0.961	0.960	0.959	0.959	0.959	0.955	0.955

Исходя из анализа таблицы, можно сделать вывод о том, что наибольшее качество работы достигается в случае работы в случае работы с каскадами alt и alt2, однако при этом время работы алгоритма очень сильно различается от времени.

Заключение. В данной работе рассматривается алгоритм Виолы–Джонса и приводятся сравнение параметров его реализации при использовании различных каскадов. Были рассмотрены самые часто используемые каскады для различных значений параметров алгоритма. По итогам анализа можно сделать вывод, о том, что наибольшую точность распознавания дают каскады alt и alt2. При этом важно учитывать, что время работы алгоритма сильно варьируется в зависимости от параметров при использовании этих каскадов. Наибольшую точность при наименьшем времени работы дают каскады LBP. Использование каскадов default и alt_tree не рекомендуется ни в какой ситуации.

Л И Т Е Р А Т У Р А

- [1] Charles Poynton, Digital Video and HDTV Algorithms and Interfaces (Elsevier, San Francisco, 2002).

- [2] P. Viola and M. J. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features”, proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001, <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf>
- [3] P. Viola and M.J. Jones, Int. J. Comp. Vision, **57**(2), 137 (2004).
- [4] А.О. Макаров, В.В. Старовойтов. Быстрая обработка изображений на основе интегральных матриц изображений. “Искусственный интеллект” **3**, 597 (2006).
- [5] R. Lienhart, A. Kuranov, V. Pisarevsky, Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection With an Extended Set of Haar-like Features, – Intel Technical Report, 2002. http://link.springer.com/chapter/10.1007%2F978-3-540-45243-0_39.
- [6] Jin H. [and others], Face detection using improved LBP under Bayesian framework. // In Proc. Third International Conference on Image and Graphics (ICIG), 2004, p. 306-309. http://ieeexplore.ieee.org/xpl/login.jsp?tp=arnumber=1410446url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D1410446
- [7] Rodriguez Y. Face Detection and Verification using Local Binary Patterns: PhD Thesis / École Polytechnique Fédérale de Lausanne, 2006.
- [8] Chang-yeon J. Face detection using lbp features / Stanford University, Final Project Report, 2008. <http://cs229.stanford.edu/proj2008/Jo-FaceDetectionUsingLBPfeatures.pdf>
- [9] Ji Zhu , Hui Zou , Saharon Rosset, Trevor Hastie. Statistics and Its Interface **2**, 349 (2009).
- [10] О. С. Амосов, Ю. С. Иванов, *Модифицированный алгоритм детекции лиц в видеопотоке и его программная реализация*. <http://cyberleninka.ru/article/n/modifitsirovannyu-algoritm-detektsii-lits-v-videopotoke-i-ego-programmnaya-realizatsiya>

По материалам IV Международной молодежной научной школы-конференции “Современные проблемы физики и технологий”.

Поступила в редакцию 12 мая 2015 г.