

Построение устойчивых признаков детекции и классификации объектов, не обладающих характерными яркостными контрастами

С.А. Гладилин, А.А. Котов, Д.П. Николаев, С.А. Усилин

Аннотация. В работе исследован новый вид признаков для задач детекции и классификации объектов на изображении. Показана устойчивость данного класса признаков в ситуациях со значительными вариациями освещенности. Исследована точность модифицированного алгоритма Виолы-Джонса в задачах идентификации колесных осей транспортных средств и распознавания цифр, нанесенных выдавливанием или чеканкой.

Ключевые слова: компьютерное зрение, детекция и классификация объектов, устойчивые признаки, алгоритм Виолы-Джонса.

Введение

Одна из наиболее распространенных задач компьютерного зрения – задача классификации. В рамках этой задачи требуется определить, принадлежит ли некоторое изображение или его часть к определенному классу. Примером такой задачи является задача автоматического поиска и локализации объектов на изображении или в видеопотоке. Метод, разработанный и представленный в 2001 году Полом Виолой и Майклом Джонсом [1], является одним из самых известных методов поиска объектов на изображении в реальном времени. Он стоит на стыке таких областей, как машинное обучение (Machine Learning) и компьютерное зрение (Computer Vision) и объединяет в себе множество алгоритмов.

Изначально методика разрабатывалась и применялась в области построения алгоритмов для детектирования лиц, и в настоящее время

алгоритм Виолы-Джонса является, пожалуй, основным методом поиска лиц на изображении. Различные модификации алгоритма получили широкое применение в задачах детектирования пешеходов [2], автомобилей в потоке [3], документов с фиксированной геометрией [4], запрещенных предметов на рентгене в аэропорту, опухолей на медицинских снимках.

Одной из важнейших частей алгоритмов типа Виолы-Джонса является набор признаков.

Наиболее часто используются признаки типа Хаара [6]. В работах по детектированию пешеходов и автомобилей в потоке использовались признаки Хаара, построенные над яркостным изображением. Такой подход позволяет не только решать задачу обнаружения (бинарной классификации), но и задачу многоклассовой классификации объектов на изображении: определение марки автомобиля [7], идентификация изображений документов с фиксированной геометрией.

¹ Работа выполнена при финансовой поддержке РФФИ в рамках научных проектов №12-07-31256, №13-07-00870 и №13-01-12106.



Рис. 1. Пример извлечения признаков лица для классификатора Виолы-Джонса: исходное изображение лица и результат применения вейвлет-преобразования Хаара

В работе [8] авторы предлагают использовать локальные бинарные паттерны (Local Binary Pattern) для конструирования признаков алгоритма Виолы-Джонса в задаче распознавания лиц.

Авторы работы [9] на примере прикладной задачи поиска «ромашкоподобных» объектов в траве использовали более простой набор признаков, который состоит из разностей яркостей изображения в двух пикселах (x_1, y_1) и (x_2, y_2) , для всех возможных пар пикселей:

$$feature(x_1, y_1, x_2, y_2) = image(x_1, y_1) - image(x_2, y_2) \quad (1)$$

В отличие от признаков типа Хаара и локальных бинарных паттернов, этот набор признаков существенно не локализован. Подобные признаки ранее использовались рядом авторов для исследования задач распознавания пола человека по фотографии лица [10].

В работе [11] используются вейвлет-преобразования Хаара и Добеши [12] для извлечения признаков изображения лиц и построения классификатора Виолы-Джонса (Рис. 1).

Авторы статьи [13] предложили модификацию метода Виолы-Джонса с использованием детекторов особенностей на основе конечных автоматов Мили [14] с 8 состояниями вместо детекторов прямоугольных особенностей. Применение автоматов позволило сократить число уровней в каскаде и число детекторов на каждом уровне за счет того, что детектор, управляемый автоматом, может обнаружить более сложные особенности, нежели простые прямоугольники.

В данной работе предлагается использовать признаки типа Хаара, построенные над картиной направленных краев. Такой подход позволяет обучить классификатор, который будет определяться формой распознаваемых объектов, а не их яркостной характеристикой.

При этом сохраняются преимущества, характерные для признаков Хаара. Во-первых, существует метод быстрого вычисления значения признака вне зависимости от размера с помощью интегрального изображения [1]. Во-вторых, такого рода признаки слабо чувствительны к незначительным искажениям обучающей выборки, что помогает акцентировать внимание на характерных особенностях классифицируемых объектов, снижая при этом влияние шума. В-третьих, прямоугольные признаки линейно масштабируемы, что облегчает задачу поиска объектов различных размеров (масштабируя обученные классификаторы, а не исходное изображение).

1. Предлагаемый подход к построению признаков

Как показал анализ, изображения колес автомобилей не содержат достаточного количества устойчивых контрастных областей предопределенного размера. Казалось бы, что центр образа колеса должен быть светлее шин, но это далеко не всегда так (Рис 2). То же самое и с фоном – асфальт может быть как темнее шин, так и светлее их. Такая ситуация определяется не только различными видами колес, кузовов автомобилей, но и условиями освещенности. Изображение одного и того же колеса может кардинально поменяться при смене условий освещенности (яркое солнце, тень, дым).

Классификатор Виолы-Джонса с использованием признаков Хаара, построенных по яркостному изображению, чувствителен к характерным перепадам яркости – когда одна часть изображения всегда темнее другой (например, область глаз на изображении всегда темнее

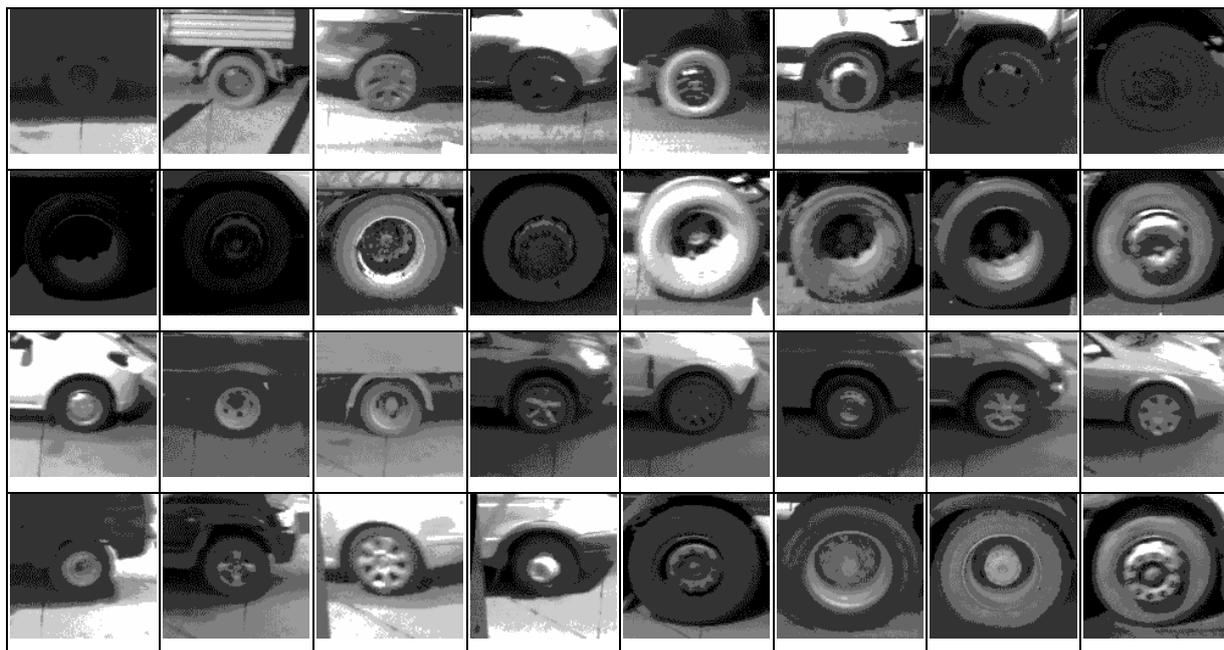


Рис. 2. Примеры изображений колес

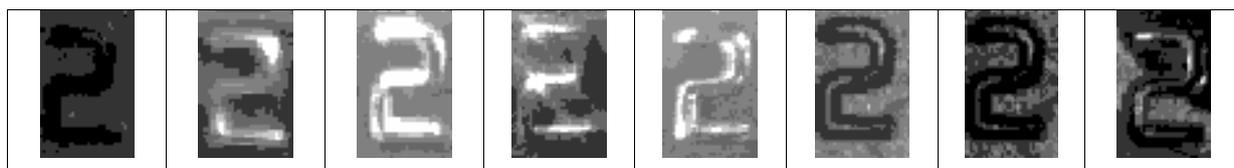


Рис. 3. Примеры изображений цифры 2

области лба). Для задач детектирования колес, таким образом, он не подходит, как и для задачи распознавания цифирных символов, нанесенных выдавливанием или чеканкой (Рис. 3).

Решением является такое преобразование изображения (что равносильно переходу в такое пространство признаков), при котором значимыми становятся направления границ, а не яркостный перепад на них. Таким преобразованием является, например, детектор Канни [20], который изображениям колес, существенно отличающимся друг от друга по цветовым характеристикам (Рис. 4 слева), ставит в соответствие очень похожие контурные изображения (Рис. 4 в центре).

Однако для построения признаков, аналогичных признакам Хаара на яркостных изображениях, на контурном изображении необходимо не просто выделить границы, но и присвоить им «знаки» (положительный и отрицательный). В качестве таких «знаков» было

решено использовать направление границ – тогда признаки смогут реагировать, например, на условия вида «граница, ориентированная вдоль основной диагонали расположена левее другой границы, ориентированной вдоль побочной диагонали изображения».

Анализ задачи детектирования колес показал, что на изображениях автомобилей имеется большое количество вертикальных и горизонтальных границ (днище автомобиля, бензобак грузовика и т.д.), в то время как границ, направленных по диагоналям мало – и значительное количество из них как раз относится к колесам.

Итак, для задачи детектирования колес построим картину направленных краев при помощи детектора границ Канни и будем рассматривать только границы вдоль главной и побочной диагоналей. При этом на контурном изображении пиксели граничных точек вдоль главной диагонали «покрасим» черным, а вдоль побочной – белым, все остальные точки – серым. Та-

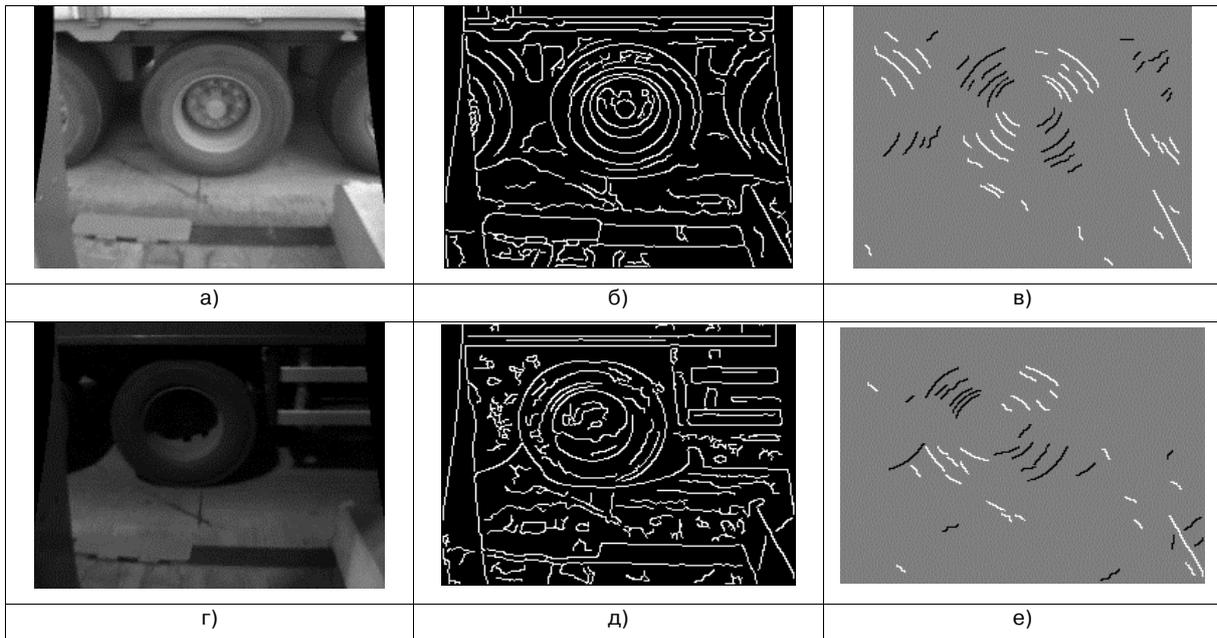


Рис. 4. Исходные изображения колес (а, г), результат детектора границ Канны (б, д) и картина направленных краев (в, е)

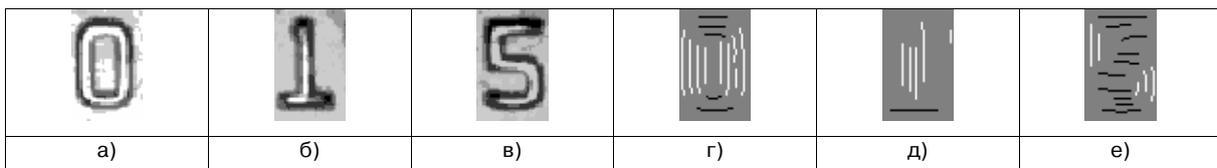


Рис. 5. Цифирные символы (а-в) и соответствующие им контурные изображения (г-е)

ким образом, если на исходном изображении был образ колеса, то на контурном изображении на его месте мы получим определенную устойчивую картину (Рис. 4 справа), по которой можно обучить классификатор Виолы-Джонса.

Анализ задачи детектирования и распознавания цифирных символов, нанесенных выдавливанием и чеканкой, показал, что для нанесения данных символов, как правило, используется небольшой набор шрифтов, в нашей конкретной задаче – шрифты, большинство направлений границ в которых ориентированы вдоль вертикали и горизонтали (Рис. 3). В то же время направление контрастов фонового изображения примерно равномерно распределен по всем направлениям. Поэтому для задачи детектирования цифр будем «красить» вертикальные границы белым, горизонтальные – черным, остальные оставлять серыми (Рис. 5).

2. Алгоритм построения контурного изображения

Подробно опишем алгоритм построения контурного изображения с диагональными границами (Рис. 6). Алгоритм является модификацией детектора Канны [20]. Построение изображения с прямыми границами полностью аналогично.

- Сглаживаем изображение фильтром Гаусса [21].
- Строим градиент изображения.
- Отмечаем потенциальные границы, то есть те пиксели, в которых достигается локальный максимум градиента в направлении вектора градиента, при условии, что данный максимум выше некоторого (нижнего) порога и что вектор градиента направлен вдоль выделенных границ (в задаче детектирования колес – вдоль главной или побочной диагонали, в задаче распознавания цифр – вдоль вертикали и горизонтали).

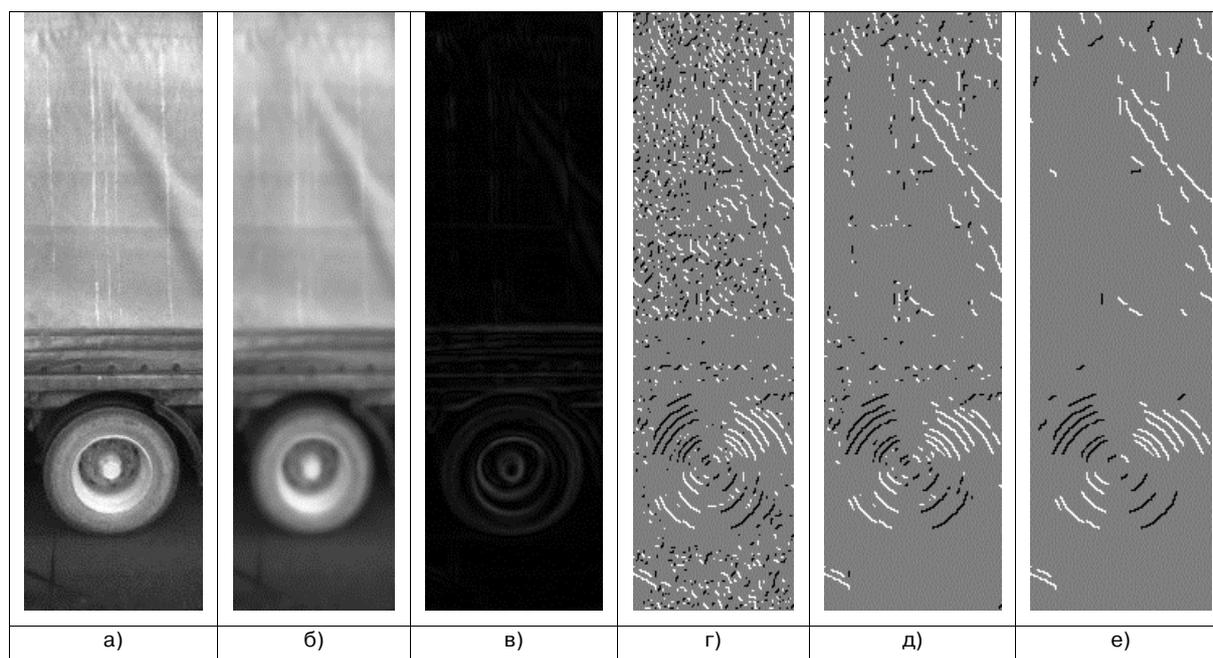


Рис. 6. Исходное изображение (а); сглаженное по Гауссу изображение (б); модуль градиента (в); потенциальные границы (г), подавление слабых компонент связности (д); подавление коротких компонент (е)

- Собираем полученные пиксели в компоненты связности (независимо вдоль выделенных направлений).
- Подавляем компоненты, максимальное значение градиента которых меньше некоторого (верхнего) порога.
- Подавляем компоненты, длина которых меньше определенной величины.

3. Численный эксперимент и результаты

В рамках задачи распознавания образов колес обучены и протестированы классификаторы двух типов. Для их обучения были подготовлены наборы положительных и отрицательных образцов. Первый классификатор («яркостной») обучался по яркостным изображениям. При этом набор положительных образцов составляли, в основном, изображения «светлых» (нормальные условия освещенности) колес (всего 460 изображений). Набор же негативных примеров для первого классификатора состоял из фрагментов изображений, не содержащих изображения колес (7420 изображений). Соответствующие наборы для второго («контурно-» классификатора были получены из наборов

первого применением детектора диагональных границ.

При обучении использовались следующие параметры:

- false positive rate на слой $f = 0,2$
- detection rate на слой $d = 0,98$
- количество уровней (слоев) $S = 15$

Обучение производилось с учетом поиска объектов различных линейных размеров. Задавались размер распознавателя 60×60 пикселей и набор масштабов (от 1,0 до 2,5 с шагом 0,1), соответственно при тренировке классификатора все положительные объекты масштабировались к размеру 60×60 пикселей. Помимо этого, в отрицательную выборку на каждом слое входили фрагменты изображений с размером распознавателя, вырезанные из исходных отрицательных изображений, взятых во всех масштабах (сжатых в 1.0, 1.1, ..., 2.5 раз). Количество отрицательных изображений на один масштаб в одном слое выбрано равным 500.

Чтобы оценить качество детектирования, обученные классификаторы подверглись тестированию. Первый — на выборке, состоящей из 10477 изображений, при этом в тестовой выборке были изображения колес при всевозможных условиях освещенности (при обычных

условиях, яркое солнце, колеса частично или полностью покрываемые тенью). Второй — на выборке, состоящей из соответствующих контурных изображений. При тестировании использовались те данные, которые не использовались при обучении. Результаты тестирования для различных порогов классификации представлены в Табл. 1 и Табл. 2. Величины «полнота» и «точность» определяются следующим образом:

$$\text{полнота} = \frac{\text{кол. - во прав. срабатываний}}{\text{кол. - во истин. объектов}} \quad (2)$$

$$\text{точность} = \frac{\text{кол. - во прав. срабатываний}}{\text{полное кол. - во срабатываний}} \quad (3)$$

По тестовой выборке оценим ROC-кривые (Receiver Operating Characteristic curve), отражающие зависимость полноты классификаторов (detection rate) от ошибки второго рода. Полученные ROC-кривые представлены на Рис. 7.

Из графика видно, что при одном и том же уровне ложных срабатываний полнота «контурного» классификатора существенно больше полноты «яркостного». Таким образом, классификатор на основе предложенного подхода более устойчив к различным условиям освещенности.

В рамках задачи классификации цифрных символов для каждой цифры также тренировались бинарные классификаторы двух типов. Происходило обучение принципу «один против всех», то есть при обучении классификатора на поиск какой-то цифры изображения остальных цифр помещались в отрицательную выборку. Потом на основе нескольких таких бинарных классификаторов можно сконструировать общий распознаватель, который будет соотносить показанное ему изображение с одним из классов, соответствующих цифрам.

Табл. 1. Результаты работы «яркостного» классификатора

ложные срабатывания	полнота	точность
32	0,49	0,99
78	0,53	0,99
99	0,56	0,98
134	0,57	0,98
197	0,59	0,97
298	0,65	0,96
352	0,66	0,95
426	0,68	0,94
686	0,71	0,91
1344	0,75	0,84
1599	0,76	0,82
2662	0,78	0,74

Табл. 2. Результаты работы «контурного» классификатора

ложные срабатывания	полнота	точность
32	0,68	0,99
37	0,7	0,99
50	0,72	0,99
60	0,74	0,99
79	0,74	0,99
175	0,78	0,98
327	0,82	0,96
443	0,85	0,95
605	0,87	0,94
961	0,89	0,9
1235	0,9	0,88
1339	0,9	0,87
1564	0,91	0,85
2485	0,92	0,79

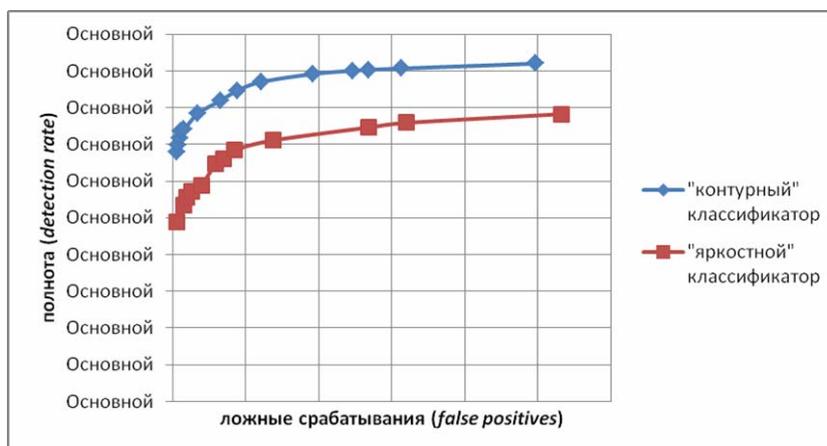


Рис. 7. ROC-кривые классификаторов

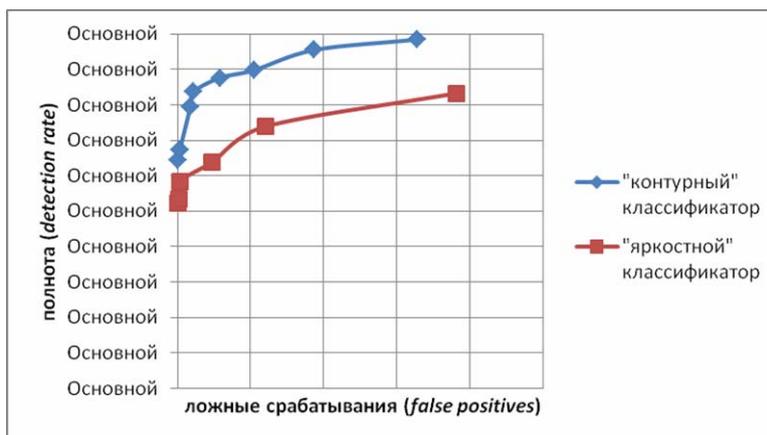


Рис. 8. ROC-кривые классификаторов цифры 3

При обучении использовались следующие параметры:

- false positive rate на слой $f = 0,3$
- detection rate на слой $d = 0,97$
- количество уровней (слоев) $S = 15$
- размер распознавателя 30×46
- масштаб от 1,0 до 1,2 с шагом 0,05
- количество отрицательных изображений на один масштаб $n=500$.

В положительные примеры для каждой цифры отбирались те изображения цифр, где сама цифра светлая, а фон темный. В тестовой выборке для каждой цифры были представлены изображения с совершенно разными условиями освещенности.

Как и в первом эксперименте, «контурные» классификаторы для каждой цифры продемонстрировали более высокое качество распознавания, чем соответствующие «яркостные» классификаторы. Худший результат среди «контурных» классификаторов у детектора цифры 3. На Рис. 8 приведены ROC-кривые для «яркостного» и «контурного» классификаторов тройки.

Для остальных цифр приведем уровни распознавания двух типов классификаторов при практически нулевой ошибке II рода (Табл. 3).

Полученные результаты объясняются тем, что в тестовой выборке примерно половина изображений была похожа на положительные образцы, а оставшиеся изображения существенно отличались от них по своим цветовым характеристикам. Таким образом, «яркостные» классификаторы распознали только первую

Табл. 3. Полнота распознавания классификаторов при практически нулевой ошибке II рода

цифра	полнота «яркостного» классификатора	полнота «контурного» классификатора
0	0,45	0,91
1	0,57	0,89
2	0,49	0,72
3	0,52	0,64
4	0,54	0,96
5	0,52	0,83
6	0,54	0,86
7	0,68	0,78
8	0,48	0,86
9	0,41	0,75

часть тестовой выборки, в то время как «контурные» классификаторы показали хороший результат в целом на всей выборке.

Проведенные эксперименты показывают, что предложенный подход построения базовых признаков позволяет обучить классификатор с высоким качеством (уровнем распознавания), который более устойчив к шуму и различным условиям освещенности, чем оригинальный классификатор Виолы-Джонса.

Заключение

Новый класс признаков для алгоритма Виолы-Джонса показал себя эффективным и устойчивым в случаях, когда детектируемые объекты не обладают характерными яркостными контра-

стами, но имеют границы характерной формы. При этом сохраняются преимущества классических признаков Хаара – устойчивость к шуму, полностью автоматическое обучение классификатора, вычислительная эффективность.

Литература

1. Viola P. and Jones M. Robust Real-time Object Detection, *International Journal of Computer Vision*. – 2001.
2. P. Viola, M. Jones, and D. Snow, Detecting Pedestrians Using Patterns of Motion and Appearance, *International Journal of Computer Vision*, vol. 63, no. 2, 2005, pp. 153–161.
3. D.C. Lee and T. Kanade, Boosted Classifier for Car Detection, 2007.
4. С. Усилин, Д. Николаев, В. Постников, Локализация, ориентация и идентификация документов с фиксированной геометрией на изображении, Труды Института системного анализа Российской академии наук. Обработка информационных и графических ресурсов, 2010, стр. 248-261.
5. A. Grigoryev, T. Khanipov, D. Nikolaev, Determination of axle count for vehicle recognition and classification, 8th Open German-Russian
6. С.Р. Papageorgiou, M. Oren, and T. Poggio, A general framework for object detection, *International Conference on Computer Vision*, 1998, pp. 555–562.F.
7. Копылов И. Е., Тассов К. Л., Определение марки автомобиля по видеокадру с использованием модифицированного алгоритма Виолы-Джонса, Наука и Образование – июнь, 2012
8. Ahonen T, Hadid A, Pietikainen M, Face Description with Local Binary Patterns: Application to Face Recognition, *Pattern Analysis and Machine Intelligence*, 2006, p. 2037 – 2041
9. Ф. А. Корнилов, К. В. Костоусов, Д. С. Первалов, Применение алгоритма Виолы-Джонса с двухточечными признаками для поиска антропогенных объектов в траве, Техническое зрение в системах управления, 2012, с. 214216
10. Baluja S., Rowley H. A. Boosting sex identification performance // *Intern. J. Computer Vision*. 2007. V. 71(1). P. 111–119.
11. Буй Тхи Тху Чанг, Фан Нгок Хоанг, В.Г. Спицын, Распознавание лиц на основе применения метода Виолы-Джонса, вейвлет-преобразования и метода главных компонент // *Известия Томского политехнического университета*. – 2012. – Т. 320. – № 5. – с. 5459
12. Добеши И. Десять лекций по вейвлетам. — Ижевск: РХД, 2001. — 464 с.
13. П.А. Скорынин, Детекторы особенностей в методе Виолы-Джонса, построенные на конечных автоматах // *Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики*, 2011, № 2 (72), с. 4044
14. Mealy George H. A Method to Synthesizing Sequential Circuits. — *Bell Systems Technical Journal*. — P. 1045–1079
15. Papageorgiou C.P., Oren M., and Poggio T. A general framework for object detection // *International Conference on Computer Vision*. – 1998. – pp. 555–562.
16. Crow, Summed-area tables for texture mapping, *Proceedings of SIGGRAPH*, vol. 18, no. 3, 1984, pp. 207–212.Y.
17. Freund and R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *Proceedings of the Second European Conference on Computational Learning Theory*, Barcelona, March, 1995, pp. 23–37.
18. Schapire R. E. and Singer Y. Improved Boosting Algorithms Using Confidence-rated Predictions // *Machine Learning*. – Vol. 37. – No. 3. – 1999. – pp. 297–336.
19. Nock R. and Nielsen F. A Real generalization of discrete AdaBoost // *Artificial Intelligence*. – Vol. 171. – No. 1. – 2007. – pp. 25–41.
20. J. Canny, A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. pami-8, no. 6, 1986, pp. 679–698.
21. Б. Яне, Цифровая обработка изображений. – Москва: Техносфера, 2007 – 584 с.

Гладилин Сергей Александрович. Научный сотрудник ИППИ РАН. Окончил МГУ в 2002 году. Кандидат физико-математических наук. Автор 9 печатных работ. Область научных интересов: зрительный интеллект, алгоритмы обработки изображений. E-mail: gladilin@iitp.ru

Котов Александр Александрович. Младший научный сотрудник ИППИ РАН. Окончил МФТИ в 2013 году. Автор 12 печатных работ. E-mail: kotov.alex.22@gmail.com

Николаев Дмитрий Петрович. Зав. сектором ИППИ РАН. Окончил МГУ в 2000 году. Кандидат физико-математических наук. Автор 113 печатных работ. Область научных интересов: быстрые алгоритмы обработки изображений. E-mail: dimonstr@iitp.ru

Усилин Сергей Александрович. Научный сотрудник ИСА РАН. Окончил МФТИ в 2009 году. Автор 16 печатных работ. Область научных интересов: цифровая обработка изображений и распознавание образов. E-mail: usilin.sergey@gmail.com