

# Информационные технологии

## Квантильная калибровка мягкого каскада

Д.П. МАТАЛОВ<sup>I,II</sup>

<sup>I</sup> ООО «Смарт Энджинс Сервис», г. Москва, Россия

<sup>II</sup> Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва, Россия

**Аннотация.** В статье представлен оригинальный алгоритм калибровки мягкого каскада Виолы и Джонса, построенный на базе метода Multiple Instance Pruning. Эффективность предложенного алгоритма продемонстрирована на примере практической задачи построения локализатора третьей страницы паспорта гражданина РФ на цифровых изображениях, полученных в неконтролируемых условиях. Предложенный в работе алгоритм продемонстрировал значительно меньшую ошибку ложного пропуска и идентичный уровень ложных срабатываний по сравнению с оригинальным алгоритмом за счет анализа статистических особенностей распределения откликов классификатора на всем множестве допустимых окон обнаружения целевого объекта. Представленный алгоритм является обобщением Multiple Instance Pruning, позволяющим строить более робастные к параметрам сканирования изображения детекторы.

**Ключевые слова:** машинное обучение, детектирование объектов, локализации объектов, метод Виолы и Джонса, мягкий каскад, soft cascade, распознавание документов.

**DOI:** 10.14357/20790279240103 **EDN:** OTBOEY

### Введение

На сегодняшний день технологии компьютерного зрения уже по-настоящему глубоко интегрированы в процессы жизнедеятельности цивилизованного человека, причем скорость экспансии таких технологий только увеличивается [1]. Практически каждый житель современного мегаполиса сталкивается с системами компьютерного зрения, реализующие распознавание лиц [2] для обеспечения безопасности и контроля доступа, распознавание документов в финансовой сфере и экономике совместного потребления [3] для ввода данных в информационные системы, автоматизированный анализ изображений в здравоохранении [4] для первичной диагностики. В процессе функционирования таких систем регу-

лярно возникает задача локализации и детектирования объектов.

На сегодняшний день метод Виолы и Джонса [5] является одним из наиболее ценных инструментов для решения промышленных задач распознавания. За свою 20-летнюю историю метод существенно эволюционировал, за счет чего получил свое широчайшее практическое распространение. Метод изначально был предложен на примере задачи локализации лиц и предназначался для решения задач локализации объектов на цифровых изображениях. Простота его дизайна, объединяющего множество простых инженерных идей, подкрепленных мощным теоретически обоснованным алгоритмом AdaBoost [6], а также революционная для того времени производительность решения за-

дачи, обеспечили высочайший интерес среди научного и индустриального сообщества.

Метод Виолы и Джонса решает задачу построения детекторов объектов на цифровых изображениях с помощью машинного обучения. В методе можно выделить 4 основных идеи: 1) вычислительно эффективное признаковое пространство; 2) обучение сильного классификатора путем построения алгоритмической композиции слабых классификаторов; 3) каскадная структура сильных классификаторов; 4) сканирующее окно. Локализация объекта производится путем сканирования областей изображения на различных масштабах бинарным каскадным классификатором на предмет наличия целевого объекта в исследуемой области, и последующей агрегацией положительных ответов классификатора для вычисления итоговых параметров прямоугольника, задающих координаты найденных объектов. Интуитивно понятно, что общее количество исследуемых областей на одном изображении может достигать нескольких сотен тысяч, причем областей изображения, содержащих целевой объект, на порядки меньше. Заприметив такую особенность, Виола и Джонс предлагают использовать каскадную структуру сильных классификаторов, которая призвана решить задачу быстрого отказа на «пустых» областях изображения, не содержащих целевой объект. В качестве признакового пространства используются признаки Хаара, отображающие структуру локального контраста областей изображения. Вычисление таких признаков может быть эффективно реализовано с помощью преподсчета интегральных изображений и целочисленной арифметики. Выбор наиболее значимых признаков для целевого класса объектов и обучение сильных классификаторов производится хорошо известным алгоритмом машинного обучения AdaBoost [6].

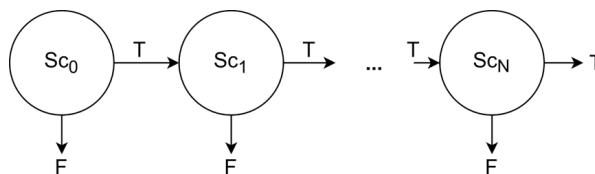
Независимыми исследователями было предложено множество модификаций и расширений оригинального метода для эффективного решения различных индустриальных задач распознавания. Такие модификации можно условно разделить на относящиеся к расширению признакового пространства классификатора [7–11], обучающему алгоритму [12–15] и высокоуровневой архитектуре классификатора [16,17]. Отдельно стоит отметить мягкий каскад [16], который впоследствии получил наиболее пристальное внимание со стороны научного и технического сообщества [8,18].

Данная работа посвящена модификации алгоритма калибровки мягкого каскада [18]. Предложенная модификация позволяет увеличить полноту детектирования объектов за счет обучения более

робастных классификаторов. Работа организована следующим образом. Для полноты повествования и обеспечения целостной картины в первом разделе представлено краткое описание и содержательный анализ классического каскада сильных классификаторов. Во втором разделе приводится описание высокоуровневой архитектуры мягкого каскада, которая является целевым объектом предлагаемого в работе алгоритма. Третий раздел содержит описание предлагаемого алгоритма калибровки и идеи, лежащие в его основе. Обширное описание экспериментальных результатов, показывающие эффективность разработанной модификации по сравнению с оригинальным алгоритмом калибровки на примере задачи локализации паспорта гражданина РФ приводится в четвертом разделе.

## 1. Классический каскад сильных классификаторов

Классическая каскадная архитектура сильных классификаторов позволяет существенно увеличить производительность детектирования объектов и обеспечивает жадную фильтрацию обучающей выборки для алгоритма AdaBoost. На рис. 1 представлена визуализация процесса классификации одного региона изображения на предмет нахождения целевого объекта. Каждый сильный классификатор представляет вычисление линейной комбинации определяющих его слабых классификаторов с последующим сравнением с пороговым значением, а по результату сравнения производится «отказ» или вычисление следующего сильного классификатора. Итоговой положительный вердикт достигается только при положительном вердикте (если значение линейной комбинации больше порога) последнего сильного классификатора.



**Рис. 1.** Процедура классификации одного региона изображения классическим каскадом сильных классификаторов

Такая высокоуровневая организация классификатора обладает рядом существенных недостатков. Во-первых, информация, полученная при вычислении текущего сильного классификатора, никаким образом не используется при вычислении следующего сильного классификатора. Со-

ответственно, решение о выходе или дальнейшем исследовании рассматриваемой области изображения на предмет наличия целевого объекта не зависит от того, насколько далеко признаковое описание рассматриваемого региона изображения находилось от границы принятия решений предыдущего сильного классификатора. Такой подход может привести к переобучению и неробастным детекторам, где небольшое отклонение от «выученного» распределения признаков целевых объектов приводит к ложным пропускам на тестовых наборах данных.

Другим существенным недостатком классической каскадной архитектуры является высокая сложность подбора обучающих параметров и вычислительная дороговизна эксперимента по обучению [19]. Существует 2 способа задать останов обучения сильного классификатора: по качественным характеристикам, формулируемым в терминах FPR и FNR — доля ложных срабатываний и пропусков соответственно, и по максимально допустимому количеству слабых классификаторов, по достижении которого сильный классификатор признается обученным. Терминальность обучения по качественным характеристикам нередко приводит к достаточно длинным композициям, а останов по длине композиции не позволяет предсказать уровень ложных пропусков итогового каскада (если исходить из того, что каскад обучается до тех пор, пока есть ложные срабатывания на всей обучающей выборке). Более того, сильный классификатор является пороговым, что означает – этот порог можно и нужно обучать для получения оптимального классификатора, что дополнительно усложняет процесс обучения.

Наконец, жесткий каскад сильных классификаторов оказывается слабо применимым в современных индустриальных распознающих системах, сопряженных с необходимостью постоянной настройки в связи с расширением обучающей выборки. В процессе развития распознающей системы, на отдельных этапах регулярно возникает множество ошибочно классифицированных примеров. Часть таких примеров, или их синтезированные изображения, используются для переобучения и донастройки системы, что является крайне ресурсозатратной задачей.

## 2. Мягкий каскад

Для того чтобы частично нивелировать недостатки каскадной архитектуры, Боурдевым и Брандтом был предложен мягкий каскад [16]. По своей сути, мягкий каскад представляет собой

один сильный классификатор, где принятие решения о дальнейшем исполнении классификатора или отказе производится совместно с вычислением очередного слабого классификатора. Правило классификации может быть представлено в виде:

$$SoftCascade(x) = \prod_{n=1}^N [\sum_{n=1}^N \alpha_n h_n(x) > t_n],$$

где  $\alpha_n$  – вес классификатора,  $h_n$  – ответ классификатора,  $t_n$  – порог на частичную сумму линейной комбинации  $n$  классификаторов. На рис. 2 представлена визуализация процесса вычисления и классификации.

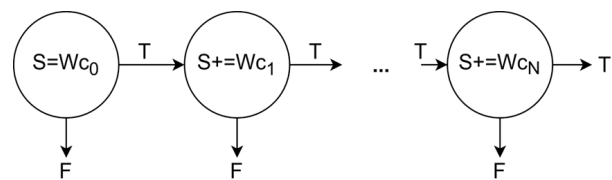


Рис. 2. Устройство архитектуры мягкого каскада

Алгоритм обучения состоит из двух этапов. Первый этап организован как классическое обучение одного сильного классификатора с помощью алгоритма бустинга, сопровождающееся итеративным обновлением не только весов обучающих примеров, но и самой обучающей выборки. Второй этап принято называть калибровкой классификатора. Задача калибровки заключается в отборе оптимального подмножества слабых классификаторов и вычислении наиболее эффективных порогов на частичные суммы.

Подобно классическому каскаду Виолы и Джонса, калибровка итеративно «пропускает» положительные примеры, чтобы жадно минимизировать ложные срабатывания. В работе [16] авторы замечают, что подобные детекторы стоит оценивать в терминах ROC-поверхности, где в качестве дополнительного измерения (помимо TPR – доля верных срабатываний и FPR – доля ложных срабатываний) выступает AWC (average weak classifiers) – математическое ожидание количества слабых классификаторов для классификации одного окна. В своем алгоритме калибровки авторы используют так называемый Rejection Distribution Vector – распределение, описывающее минимально допустимую долю пропусков для определенного уровня классификатора. Для построения данного распределения предлагается использовать некоторое параметризуемое семейство экспоненциальных функций. Главной проблемой такого алгоритма калибровки является выбор параметров. В данном случае выбор параметров экспоненциальной кривой представляется еще менее интуитивно понятным, чем выбор параметров доли ложных сра-

батований и, пропусков и количества слабых классификаторов при обучении классического каскада. Другим существенно более простым и эффективным алгоритмом калибровки является алгоритм Multiple-Instance Pruning [18].

### 3. Предлагаемый метод калибровки

Предлагаемый метод калибровки является модификацией алгоритма Multiple-Instance Pruning (MIP) [18]. Этот алгоритм имеет только один параметр – порог на частичную сумму для последнего классификатора. Алгоритм исходит из простой идеи – классификатор делает самые неэффективные ложные пропуски при вычислении последнего слабого классификатора. Под неэффективным подразумевается то, что для вычисления ошибочного вердикта вычисляются все слабые классификаторы. Поэтому авторы предлагают учитывать на этапе калибровки только те положительные примеры, которые прошли финальный уровень обученного на первом этапе сильного классификатора.

Метод Виолы и Джонса представляет собой разновидность метода сканирующего окна, где регионы изображения в различных масштабах сканируются на предмет наличия целевого объекта. Для каждого объекта целевого класса на изображении можно задать множество допустимых окон обнаружения (МДОО), где при положительном вердикте классификатора объект считается корректно обнаруженным. На рис. 3 представлена визуализация множества допустимых окон обнаружения на примере третьей страницы паспорта гражданина РФ.

Способ вычисления МДОО достаточно нетривиальная задача ввиду того, что далеко не каждый объект на изображении может быть хорошо аппроксимирован прямоугольником фиксированного соотношения сторон, что дополнительно усугубляется моделью формирования изображений и проективными искажениями. Как правило, для формирования МДОО в практических задачах задается максимально допустимое отклонение для линейных размеров и сдвигов сканирующего окна от истинного прямоугольника. Так, в работе [18] окно обнаружения является допустимым, если разница линейных размеров окна-кандидата и координатное расстояние от истинного не превышает 50%. Для каждого уровня алгоритм жадно выбирает наибольшее значение отклика среди всех откликов на допустимых окнах обнаружения соответствующего уровня, а порог для уровня выбирается в соответствии с минимумом по всем объектам. Такой алгоритм гарантирует, что полностью калиброванного классификатора будет такой же,

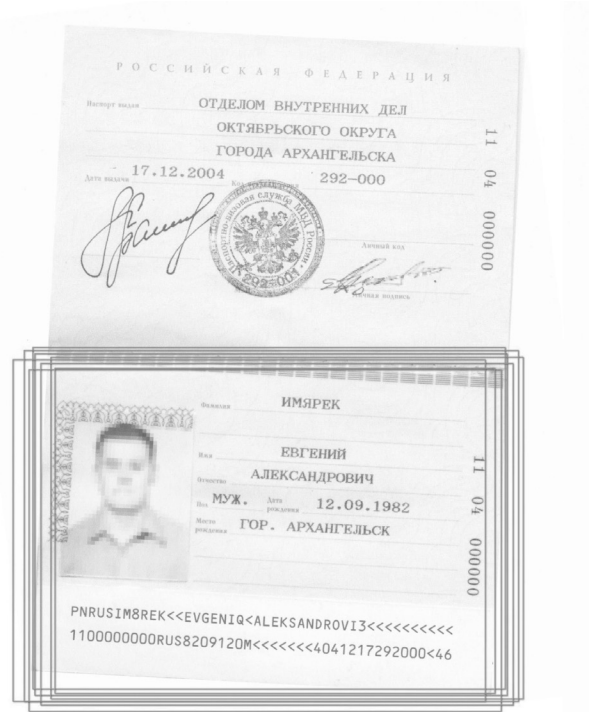


Рис. 3. Множество допустимых окон обнаружения для классификатора третьей страницы паспорта гражданина РФ

как и некалиброванного, а такой классификатор является оптимальным по скорости. Основным недостатком алгоритма является чрезмерная жадность стратегии выбора порогов на частичные суммы. Стратегия выбора максимального отклика никак не учитывает особенности распределения откликов классификатора на допустимых окнах обнаружения, которое существенно зависит от настроек сканирования изображения. На рис. 4 представлены графики функции распределения отклика

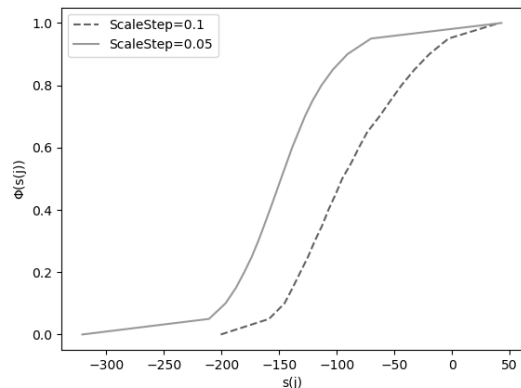


Рис. 4. Функция распределения откликов некалиброванного классификатора на МДОО для разных параметров масштабирования изображения

некалиброванного классификатора на множестве допустимых окон обнаружения одного и того же объекта, полученные с различными параметрами масштабирования изображения.

Из графика видно, что параметр масштабирования существенно влияет на форму и область допустимых значений функции распределения откликов классификатора. Другой, вычислительно простой, но в то же время существенно более информативной характеристикой распределений в статистике является значение квантиля. По определению, квантилем является значение, которое случайная величина не превышает с фиксированной вероятностью. Рассматривая отклик классификатора в качестве случайной величины, можно построить менее жадную стратегию калибровки классификатора, основанную на квантиле, выбирая значения порогов в соответствии с заданным уровнем полноты детектирования МДОО. Сохраняя нотацию алгоритма MIP, модифицированный алгоритм можно представить в виде, проиллюстрированном на рис. 5.

#### 4. Экспериментальные результаты

Оценка применимости метода калибровки, основанного на квантилях, проводилась на примере задачи построения детектора третьей страницы паспорта гражданина РФ. В качестве экспериментальных данных было отобрано 5000 изображений паспорта, полученных с помощью различных устройств – планшетных сканеров, копировальных аппаратов, производящие черно-белые изображения с дизерингом, малоформатных веб-камер и камер мобильных устройств. Для того чтобы компенсировать проективные искажения, сопро-

вождающие съемку документов с рук мобильным устройством, все изображения были предварительно нормализованы [20] таким образом, чтобы плоскость и границы страницы документа были параллельны плоскости и сторонам изображения, соответственно. Данный экспериментальный набор изображений был разделен на обучающую и тестовую выборки размером 1000 и 4000 изображений, соответственно. В качестве признакового пространства использовались граничные признаки Хаара, состоящие из двух прямоугольников [21]. В отличие от классических, такие признаки вычисляются поперек нормы градиента изображения, что обеспечивает эффективное обобщение целевых объектов, содержащих большое количество границ. В табл. 1 представлены качественные характеристики полученных детекторов на тестовом наборе данных.

Табл. 1

Качественные характеристики детектирования третьей страницы паспорта гражданина РФ

	0.0 (MIP)	0.25	0.5	0.75	1.0
FNR (%)	8.1	8.1	7.7	6.9	5.6
FP	2	2	2	2	6
AWC	1.473	1.473	1.477	1.519	1.569

Первый эксперимент (1-й столбец табл. 1), где значение квантиля равно 0, соответствует оригинальному алгоритму Multiple-Instance Pruning [18], где порог на частичную сумму классификаторов проставлялся исходя из максимального отклика всех окон объекта. Такой классификатор продемонстрировал наименьшую полноту детектирования (2-я строка табл. 1) и в то же время

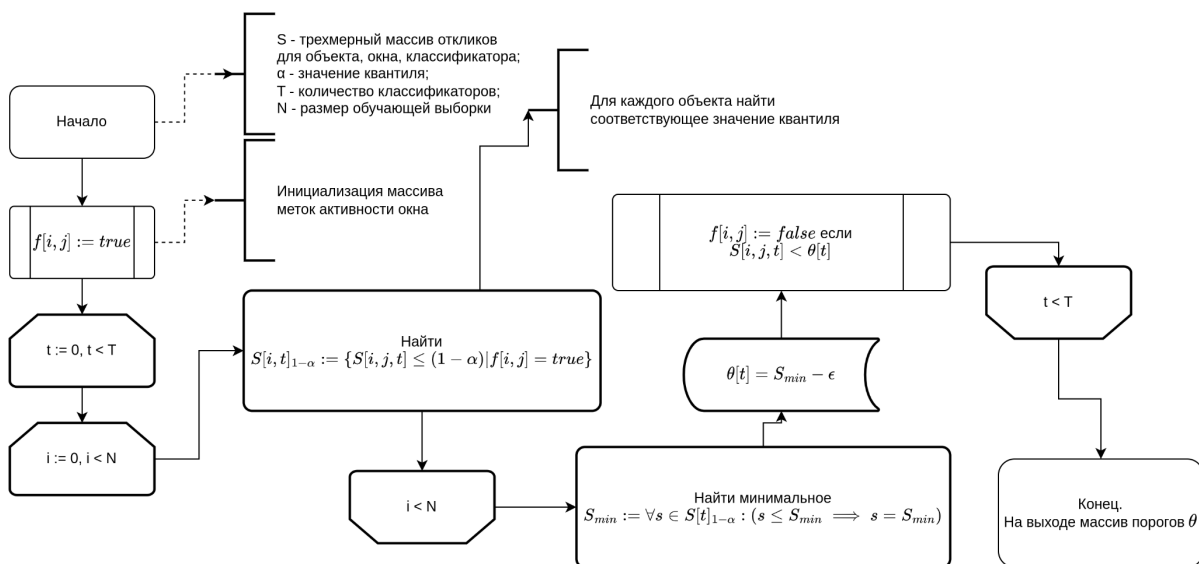


Рис. 5. Схема квантильной калибровки мягкого каскада



наилучшую скорость классификации одного региона изображения среди всех полученных детекторов. При другом граничном значении квантиля равном 1 (последний столбец табл. 1), детектор продемонстрировал существенно большую полноту, однако количество ложных срабатываний увеличилось в 3 раза.

Наилучшие качественные характеристики продемонстрировал детектор, полученный калибровкой с параметром квантиля равным 0.75 (5-й столбец). Такой детектор показал минимальное (такое же что и МПР) количество ложных срабатываний, при существенно большей полноте (частота ложных пропусков сократилась на более чем 13%). Ожидание вычисляемых слабых классификаторов увеличилось незначительно и является закономерным ввиду увеличения полноты классификации, так как при положительном ответе вычисляются все слабые классификаторы. Стоит отметить, что при значении квантиля детектор показал идентичную производительность МПР, что может свидетельствовать о низкой чувствительности признакового пространства к небольшим сдвигам и высокой робастности классификатора Виолы и Джонса.

### Заключение

В данной работе предложена модификация алгоритма калибровки мягкого каскада на примере задачи построения локализатора третьей страницы паспорта гражданина РФ. Метод рассматривает отклики классификатора на множестве допустимых окон обнаружения в качестве случайной величины, а пороговые значения на частичные суммы выбираются в соответствии со значением квантилей распределений откликов, что позволяет увеличить устойчивость итогового детектора к параметрам обхода и масштабирования изображения и обеспечить большую полноту детектирования целевых объектов.

Дальнейшее развитие предложенного в работе метода можно вести в двух направлениях. Первое состоит в более глубокой проработке вероятностного подхода к анализу распределения откликов классификатора на множестве допустимых окон обнаружения. Например, учитывать дисперсию распределений и их моменты. Другое направление заключается в разделении откликов по масштабам, а значимость каждого определять исходя из «отклонения» параметров окна от идеала. Так можно построить схему, гарантирующую полноту детектирования на нескольких переборных масштабах.

### Литература

1. *Goode A.* Digital identity: solving the problem of trust // *Biometric Technology Today*. MA Business London. 2019. Vol. 2019, № 10. P. 5–8.
2. *Lin W.-H., Wang P., Tsai C.-F.* Face recognition using support vector model classifier for user authentication // *Electronic Commerce Research and Applications*. 2016. Vol. 18. P. 71–82.
3. *Gai K., Qiu M., Sun X.* A survey on FinTech // *Journal of Network and Computer Applications*. 2018. Vol. 103. P. 262–273.
4. *Bulatov K.B. et al.* Reducing radiation dose for NN-based COVID-19 detection in helical chest CT using real-time monitored reconstruction // *Expert Systems with Applications*. 2023. Vol. 229. P. 120425.
5. *Viola P., Jones M.* Rapid object detection using a boosted cascade of simple features // *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2001. IEEE Comput. Soc.
6. *Freund Y., Schapire R.E.* A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting // *Journal of Computer and System Sciences*. 1997. Vol. 55. № 1. P. 119–139.
7. *Гладилин С.А. et al.* Построение устойчивых признаков детекции и классификации объектов, не обладающих характерными яркостными контрастами // *ИТиВС / ed. Соломонович П.Ю. ФИЦ ИУ РАН*. 2014. № 1. P. 53–60.
8. *Dollar P., Belongie S., Perona P.* The Fastest Pedestrian Detector in the West // *Proceedings of the British Machine Vision Conference 2010*. 2010.
9. *Li S.Z. et al.* Statistical Learning of Multi-view Face Detection // *Computer Vision — ECCV 2002 / ed. Heyden A. et al. Berlin, Heidelberg: Springer Berlin Heidelberg*. 2002. Vol. 2353. P. 67–81.
10. *Lienhart R., Maydt J.* An extended set of Haar-like features for rapid object detection // *Proceedings. International Conference on Image Processing*. Rochester. NY. USA: IEEE. 2002. Vol. 1. P. I-900-I-903.
11. *Ramirez G.A., Fuentes O.* Multi-Pose Face Detection with Asymmetric Haar Features // *2008 IEEE Workshop on Applications of Computer Vision*. Copper Mountain. CO. USA: IEEE. 2008. P. 1–6.
12. *Поляков И.В. et al.* Построение оптимальных каскадов Виолы-Джонса при помощи “жадных” алгоритмов перебора управляющих параметров с промежуточным контролем по валидационной выборке // *Сенсорные системы / ed. Островский М.А. ООО «ИКЦ Академкнига»*. 2016. Vol. 30. № 3. P. 241–248.

13. *Chang Huang et al.* Vector boosting for rotation invariant multi-view face detection // Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. Beijing, China: IEEE, 2005. Vol. 1. P. 446-453.
14. *Hutchison D. et al.* Multi-stage Sampling with Boosting Cascades for Pedestrian Detection in Images and Videos // Computer Vision – ECCV 2010 / ed. Daniilidis K., Maragos P., Paragios N. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. Vol. 6316. P. 196–209.
15. *Jones M., Viola P.* Fast multi-view face detection. 2003.
16. *Bourdev L., Brandt J.* Robust object detection via soft cascade // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005.
17. *Minkina A. et al.* Generalization of the Viola-Jones method as a decision tree of strong classifiers for real-time object recognition in video stream // Seventh International Conference on Machine Vision (ICMV 2014). SPIE. 2015. Vol. 9445. P. 944517.
18. *Zhang C., Viola P.* Multiple-Instance Pruning For Learning Efficient Cascade Detectors.
19. *Wang Y.-Q.* An Analysis of the Viola-Jones Face Detection Algorithm // Image Processing On Line. Image Processing On Line. 2014. Vol. 4. P. 128–148.
20. *Shemiakina J. et al.* Automatic cropping of images under projective transformation // Eleventh International Conference on Machine Vision (ICMV 2018) / ed. Nikolaev D.P. et al. Munich, Germany: SPIE. 2019. P. 117.
21. *Usilin S.A., Slavin O.A., Arlazarov V.V.* Memory Consumption and Computation Efficiency Improvements of Viola–Jones Object Detection Method for Remote Sensing Applications // Pattern Recognition and Image Analysis. Pleiades journals. 2021. Vol. 31, № 3. P. 571–579.

**Маталов Даниил Павлович.** Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва, Россия. Младший научный сотрудник. ООО «Смарт Энджинс Сервис» Научный сотрудник-программист. Область научных интересов: обработка изображений, распознавание образов, машинное обучение. E-mail: d.matalov@smartengines.com.

## Quantile based soft cascade calibration

D.P. Matalov<sup>I,II</sup>

<sup>I</sup> Smart Engines Service LLC, Moscow, Russia

<sup>II</sup> Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia

**Abstract.** This paper presents an algorithm based on the Multiple Instance Pruning for calibrating the Viola and Jones soft cascades. The effectiveness of the proposed algorithm is showed in the practical problem of training detector of the 3rd page of a passport of a citizen of the Russian Federation in digital images obtained in uncontrolled conditions. The resultant detector showed a significantly lower false negative rate and the same false positive rates compared to the original Multiple Instance Pruning algorithm. To achieve that, the algorithm follows statistical approach by considering the distributions of classifier responses on the entire set of acceptable detection windows. The proposed algorithm is a generalization of Multiple Instance Pruning, which allows to train detectors robust to different image scanning parameters.

**Keywords:** *machine learning, object detection, object localization, Viola-Jones method, soft cascade, identity document recognition*

**DOI:** 10.14357/20790279240103 **EDN:** OTBOEY

### References

1. *Goode A.* Digital identity: solving the problem of trust. *Biometric Technology Today*. 2019; 2019(10):5–8.
2. *Lin WH, Wang P, Tsai CF.* Face recognition using support vector model classifier for user authentication. *Electronic Commerce Research and Applications*. 2016 Jul;18:71–82.
3. *Gai K, Qiu M, Sun X.* A survey on FinTech. *Journal of Network and Computer Applications*. 2018 Feb;103:262–73.
4. *Bulatov KB, Ingacheva AS, Gilmanov MI, Chukalina MV, Nikolaev DP, Arlazarov VV.* Reducing radiation dose for NN-based COVID-19 detection in helical chest CT using real-time monitored reconstruction. *Expert Systems with Applications*. 2023 Nov;229:120425.
5. *Viola P, Jones M.* Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001*. IEEE Comput. Soc;
6. *Freund Y, Schapire RE.* A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*. 1997 Aug;55(1):119–39.
7. *Gladilin S., Kotov A., Nikolaev D. and Usilin S.* “Postroenie ustoychivyykh priznakov detektsii i klassifikatsii obektov, ne obladayuschikh kharakternymi yarkostnymi kontrastami,” *ITiVS*, no 1, pp. 53-60, 2014.
8. *Dollar P, Belongie S, Perona P.* The Fastest Pedestrian Detector in the West. In: *Proceedings of the British Machine Vision Conference 2010*. 2010.
9. *Li SZ, Zhu L, Zhang Z, Blake A, Zhang H, Shum H.* Statistical Learning of Multi-view Face Detection. In: *Heyden A, Sparr G, Nielsen M, Johansen P, editors. Computer Vision – ECCV 2002* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2002 [cited 2023 Sep 13]. p. 67–81. (Goos G, Hartmanis J, Van Leeuwen J, editors. *Lecture Notes in Computer Science*; vol. 2353). Available from: [http://link.springer.com/10.1007/3-540-47979-1\\_5](http://link.springer.com/10.1007/3-540-47979-1_5)
10. *Lienhart R, Maydt J.* An extended set of Haar-like features for rapid object detection. In: *Proceedings International Conference on Image Processing* [Internet]. Rochester, NY, USA: IEEE; 2002 [cited 2023 Sep 13]. p. I-900–I-903. Available from: <http://ieeexplore.ieee.org/document/1038171/>
11. *Ramirez GA, Fuentes O.* Multi-Pose Face Detection with Asymmetric Haar Features. In: *2008 IEEE Workshop on Applications of Computer Vision* [Internet]. Copper Mountain, CO, USA: IEEE; 2008 [cited 2023 Sep 13]. p. 1–6. Available from: <http://ieeexplore.ieee.org/document/4544013/>
12. *Polyakov I., Kuznetsova E., Usilin S. and Nikolaev D.* “Training optimal Viola-Jones detectors using greedy algorithms for selecting control parameters with intermediate validation on each level,” *Sensory systems*, vol. 30, no 3, pp. 241-248, 2016.
13. *Chang Huang, Haizhou Ai, Yuan Li, Shihong Lao.* Vector boosting for rotation invariant multi-view face detection. In: *Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1* [Internet]. Beijing, China: IEEE; 2005 [cited 2023 Sep 13]. p. 446-453 Vol. 1. Available from: <http://ieeexplore.ieee.org/document/1541289/>
14. *Hutchison D, Kanade T, Kittler J, Kleinberg JM, Mattern F, Mitchell JC, et al.* Multi-stage Sam-



- pling with Boosting Cascades for Pedestrian Detection in Images and Videos. In: Daniilidis K, Maragos P, Paragios N, editors. Computer Vision – ECCV 2010 [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010 [cited 2023 Sep 13]. p. 196–209. (Lecture Notes in Computer Science; vol. 6316). Available from: [http://link.springer.com/10.1007/978-3-642-15567-3\\_15](http://link.springer.com/10.1007/978-3-642-15567-3_15)
15. Jones M, Viola P. Fast multi-view face detection. 2003 Dec;
  16. Bourdev L, Brandt J. Robust object detection via soft cascade. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005.
  17. Minkina A, Nikolaev D, Usilin S, Kozyrev V. Generalization of the Viola-Jones method as a decision tree of strong classifiers for real-time object recognition in video stream. In: Seventh International Conference on Machine Vision (ICMV 2014). SPIE; 2015. p. 944517.
  18. Zhang C, Viola P. Multiple-Instance Pruning For Learning Efficient Cascade Detectors.
  19. Wang YQ. An Analysis of the Viola-Jones Face Detection Algorithm. Image Processing On Line. 2014 Jun 26;4:128–48.
  20. Shemiakina J, Zhukovsky A, Konovalenko I, Nikolaev D. Automatic cropping of images under projective transformation. In: Nikolaev DP, Radeva P, Verikas A, Zhou J, editors. Eleventh International Conference on Machine Vision (ICMV 2018) [Internet]. Munich, Germany: SPIE; 2019 [cited 2023 Sep 13]. p. 117. Available from: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11041/2523483/Automatic-cropping-of-images-under-projective-transformation/10.1117/12.2523483.full>
  21. Usilin SA, Slavin OA, Arlazarov VV. Memory Consumption and Computation Efficiency Improvements of Viola–Jones Object Detection Method for Remote Sensing Applications. Pattern Recognition and Image Analysis. 2021 Jul 1;31(3):571–9.

**Matalov D.P.** Researcher-programmer at Smart Engines Service LLC, Moscow, Russia. Junior researcher at the Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia. Number of publications: 18. Research interests: image processing, pattern recognition, machine learning. E-mail: [d.matalov@smartengines.com](mailto:d.matalov@smartengines.com).