

## НОВЫЙ МЕТОД ДЕТЕКТИРОВАНИЯ ЧЕЛОВЕЧЕСКИХ ЛИЦ НА ЦИФРОВЫХ ИЗОБРАЖЕНИЯХ

ГРЕБНОВ И.В., асп.

**Предлагается новый метод детектирования человеческих лиц на цифровых изображениях, протестированный на тестовом наборе CMU/MIT [6], где превзошел все современные аналоги как по скорости, так и по качеству детектирования.**

*Ключевые слова:* алгоритм детектирования человеческих лиц, просеивающий фильтр, классификатор, каскадная модель, классифицирующие примитивы.

### NEW METHOD OF HUMAN FACES DETECTION ON DIGITAL PICTURES

I.V. GREBNOV, Post Graduate Student

**In this article the new method of human faces detection in digital pictures is represented. It has been tested by means of CMU/MIT test set. It overcomes the analogs of speed and also the analogs of detection quality.**

*Key words:* algorithm of human faces detection, sifting filter, classifier, cascade model, classifying primitives.

**Введение.** В настоящее время, в связи со стремительным развитием фото- и видеоэлектроники, очень перспективной является задача детектирования объектов на цифровых изображениях.

Несмотря на то, что в реальном мире существует огромное количество различных объектов, значительный интерес представляет разработка алгоритма обнаружения более узкого класса объектов – человеческих лиц. Это обусловлено растущей степенью автоматизации различных процессов и производственных систем. Так, конкретными практическими применениями для алгоритма детектирования лиц могут быть: системы автоматического учета числа посетителей в супермаркетах и развлекательных центрах; системы пропускного контроля в различных учреждениях, аэропортах и метро; автоматические системы предотвращения несчастных случаев, производящие мониторинг лица водителя транспортного средства; интеллектуальные интерфейсы человек-компьютер и др.

Существующие потребности в создании подобных систем накладывают жесткие ограничения на скорость работы алгоритмов, которые должны работать в реальном времени. Таким образом, перспективной является задача создания быстрого и надежного алгоритма детектирования человеческих лиц.

**Существующие подходы к задаче детектирования лиц.** Последние десять лет в области детектирования образов ведутся активные разработки и предложены различные методы детектирования: метод главных компонент, методы с использованием гистограмм, нейросетевые, байесовские сети, статистические методы и т.д. Часть этих алгоритмов детектирования инвариантны по отношению к объекту, другие используют такие априорные знания об объекте, как форма, цветовая гамма, взаимное расположение частей и др.

Впервые представленный в 2001 г. алгоритм Viola&Jones[1] является инвариантным и обладает высокой скоростью работы и хорошей точностью детектирования. Его работа состоит в следующем. Существует некоторый ограниченный набор из примитивов, каждый из которых представляет собой некий шаблон анализа изображения. На этапе обучения выделяются примитивы такого типа, размера и положения в окне поиска алгоритма, которые наилучшим образом разделяют изображения, содержащие объект обучения, и изображения, его не содержащие. По ито-

гам обучения получается несколько тысяч примитивов, объединенных в дерево классификации. Для детектирования образа исходное изображение просматривается скользящим окном, и для каждой позиции окна метод с помощью классификации определяет, является ли этот регион искомым образом или нет. Для детектирования образов разного размера и ориентации такой подход необходимо применять не только к исходному изображению, но и к его измененной копии.

Целью работы является создание нового метода детектирования человеческих лиц, изображенных на цифровых фотографиях, обладающего высокой скоростью работы и точностью детектирования. В основу нового метода положены идеи алгоритма Viola&Jones [1]. Найденные слабые места алгоритма, а также использование априорных знаний о форме и цвете человеческого лица позволили выявить возможные пути улучшения характеристик работы алгоритма и создать новый метод детектирования человеческих лиц.

**Описание нового метода.** При разработке нового метода исследования проводились по двум направлениям: повышение скорости и повышение точности детектирования. Сам метод предлагается формировать из серии фильтров. Каждый фильтр является самостоятельным методом детектирования и позволяет быстро отсекал регионы, не содержащие области лиц, для этого используются некоторые различные характеристики объекта. При этом каждая последующая часть обрабатывает только перспективные области изображения, полученные от предыдущих частей метода.

**Метод попиксельной фильтрации изображений.** Методы попиксельной фильтрации изображений представляют собой серию фильтрующих функций отображения. Каждому пикселю исходного изображения в зависимости от его интенсивности и интенсивности его соседей ставится в соответствие некоторое значение. Цель такой фильтрации заключается в том, чтобы отделить регионы, не содержащие области лиц, на основе информации о распределении значений искомого объекта.

**Использование информации о распределении цвета кожи лица.** Во многих случаях информация о распределении цвета искомого объекта может ускорить процесс детектирования объекта в несколько раз. В случае поиска лица на изображении подмножество цветового пространства RGB для регионов, содержащих области кожи (лица), составляет всего 6,1 % [2].

Проанализировав область распределения цвета кожи на цветных изображениях, была найдена элегантная формула, которая позволяет эффективно разделить регионы, которые не содержат изображения кожи (лица):

$$(R > 95) \text{ AND } (G > 40) \text{ AND } (B > 20) \text{ AND } (R - G > 15) \text{ AND } (R > B).$$

**Использование информации о форме объекта.** Во многих случаях информация о форме искомого объекта также может ускорить процесс детектирования. В качестве алгоритма анализа формы объекты был выбран алгоритм Canny Edge Detector [3]. Цель данного алгоритма заключается в поиске границ объектов на изображении. По интенсивности границ в некотором регионе можно оценить, насколько сложна или проста форма объекта. Проанализировав область распределения интенсивности границ лиц на изображениях, были найдены пороговые значения, которые позволяют эффективно отделить регионы, не содержащие изображения лица:

$$(\text{Intensity} > 10\%) \text{ AND } (\text{Intensity} < 60\%).$$

Примеры фильтрации на основе информации о цвете и форме лица представлены на рис. 1.

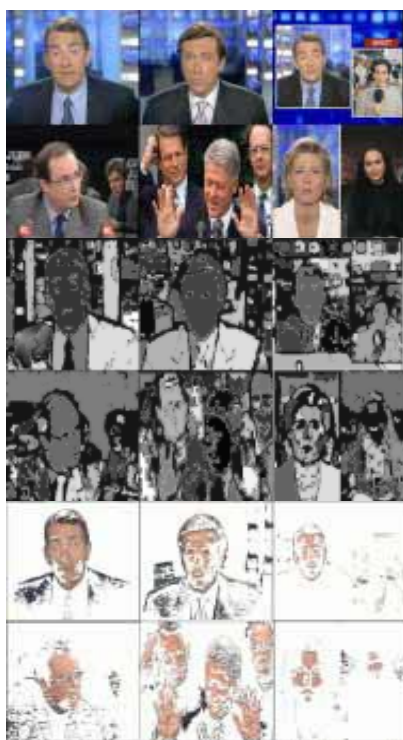


Рис. 1. Пример фильтрации на основе информации о цвете и форме объекта

**Метод усиления слабых классификаторов (AdaBoost).** Усиление слабых классификаторов – подход к решению задачи детектирования образов путем комбинирования слабых классификаторов в один сильный. Под силой классификатора в данном случае подразумевается эффективность (качество) решения задачи детектирования. В основе метода усиления простых классификаторов лежит простая предпосылка: скомбинировать некоторое количество элементарных (простых) признаков таким образом, чтобы получить один, но более мощный. Это можно сделать с помощью алгоритма AdaBoost [4]. В качестве слабого классификатора был выбран анизотропный гауссовый примитив, который превосходит по качеству детекти-

рования прямоугольный примитив Хора [1], предложенный авторами алгоритма Viola & Jones[1].

**Анизотропные гауссовы примитивы.** Анизотропные гауссовы примитивы впервые были описаны в работе Peotta [5], где предлагается применять их для сжатия изображений и аппроксимации сигналов. Базовая функция примитивов  $\varphi(x, y) : R^2 \rightarrow R^2$  имеет вид

$$\varphi(x, y) = xe^{-(|x|+y^2)}.$$

Дополнительно к базовой функции применяются функции преобразования, такие как перенос, поворот, изгиб, сжатие, позволяющие существенно расширить многообразие анизотропных гауссовых примитивов. На рис. 2 показаны примеры таких примитивов, полученных с помощью разных параметров переноса, поворота, изгиба и сжатия.

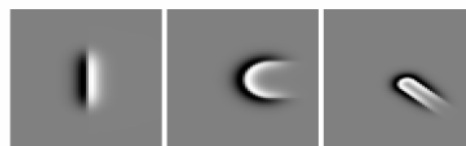


Рис. 2. Пример анизотропных гауссовых примитивов

Данный класс функций имеет ключевую способность хорошо анализировать не только контуры характерных участков образов, но и их градиенты. Причем, в отличие от прямоугольных примитивов Хора [1], способность к анализу контуров не ограничена только прямыми линиями на изображении.

Простой классификатор на основе анизотропного гауссова примитива строится следующим образом:

$$h_i(I) = \begin{cases} 1, f_i(I) < \theta_i, \\ -1, f_i(I) \geq \theta_i, \end{cases}$$

где  $I$  – сигнал анализируемого региона; параметр  $\theta_i$  выбирается исходя из правила Байеса; функция  $f_i(I)$  представляет собой скалярное произведение классифицируемого региона и функции анизотропного гауссова примитива:

$$f_i(I) = \iint_{X \times Y} \tilde{\varphi}_i(x, y) I(x, y) dx dy.$$

Конечно, сам по себе такой простой классификатор не может обеспечивать качественного детектирования образов. Назовем такой классификатор слабым. Для достижения более высокого качества детектирования необходимо объединить несколько таких слабых классификаторов в совокупность. Причем каждый слабый классификатор будет иметь свой вес. Назовем такую совокупность сильным классификатором:

$$H_M(I) = \frac{\sum_{m=1}^M w_m h_m(I)}{\sum_{m=1}^M w_m}$$

В качестве алгоритма построения такого сильного классификатора предлагается использовать алгоритм AdaBoost [4]. Это итерационный алгоритм, на каждом шаге которого слабый классификатор добавляется в сильный путем взвешенного линейного смешивания. Смешивание классификаторов производится в целях минимизации целевой функции. В нашем случае будем использовать функцию ошибки детектирования на тестовом наборе.

**Объединение сильных классификаторов в каскадную модель.** При объединении слабых анизотропных классификаторов в один большой сильный классификатор детектирующая способность повышается, но при этом повышается и время принятия решения. В некоторых случаях время, помимо качества распознавания, является ключевым фактором. Для этого случая предлагается построить каскадную модель из небольших сильных классификаторов, позволяющую повысить скорость детектирования, не ухудшая качество.

Каскадная модель сильных классификаторов – это дерево принятия решений, где каждый узел дерева построен таким образом, чтобы детектировать почти все интересующие образы и отклонять некоторую часть регионов, не являющихся образами. Помимо этого, узлы дерева размещены таким образом, что чем ближе узел находится к корню дерева, тем из меньшего количества анизотропных гауссовых примитивов он состоит и тем самым требует меньше времени на принятие решения. Данный вид каскадной модели хорошо подходит для обработки изображений, на которых общее количество детектируемых образов мало. В этом случае метод может быстрее принять решение о том, что данный регион не содержит образ, и перейти к следующему. Пример каскадной модели показан на рис. 3.

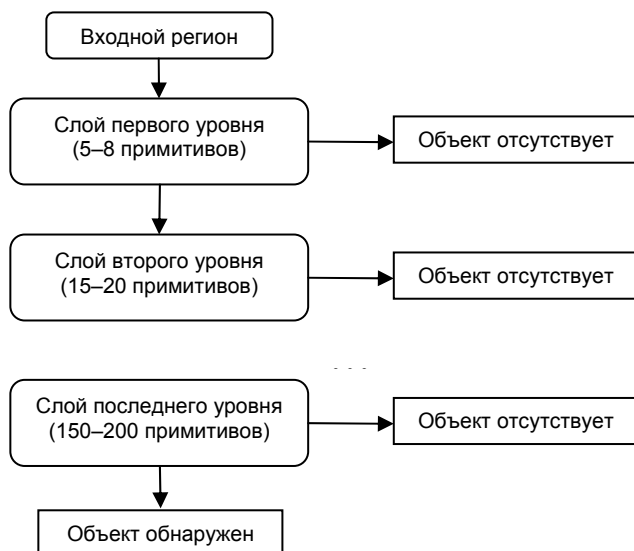


Рис. 3. Каскадная модель сильных классификаторов

**Гибридный метод на основе серии детекторов.** Качество и скорость работы метода усиления слабых классификаторов напрямую зависят от выбора классифицирующих примитивов. Анизотропные гауссовы примитивы обладают хорошим качеством детектирования, но при этом проигрывают прямоугольным примитивам Хора [1] по скорости.

Во многих случаях предварительный анализ региона можно выполнить с помощью более простых примитивов. Для достижения дополнительной производительности эффективно применять серию детекторов. В этом случае поиск лица на некотором шаге производится не во всем кадре, а в пределах полученной от предыдущего детектора области, взятой с некоторым запасом. Для поиска лиц на изображении эффективно применять серию из трех детекторов (рис. 4). Предварительный анализ осуществляется с помощью примитивов Хора в окне размера 36x36. Детектор следующего уровня осуществляет локализацию лица пу-

тем последовательного просмотра скользящим окном размера 24x24 также с помощью примитивов Хора. Далее непосредственно детектирование происходит с помощью анизотропных гауссовых примитивов.

Результат такой фильтрации представлен на рис. 5.

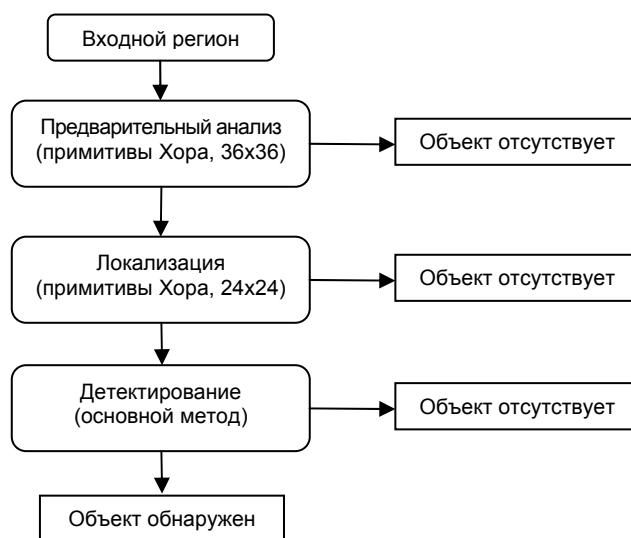


Рис. 4. Схема работы гибридного метода на основе серии детекторов

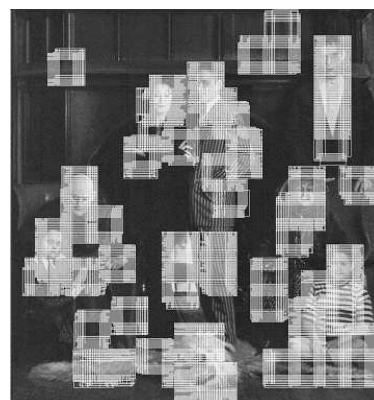


Рис. 5. Пример фильтрации на основе серии детекторов

**Результаты экспериментов.** Тестирование производилось на специализированном тестовом наборе CMU/MIT Test set [6]. Этот набор состоит из 130 изображений, содержащих в общей сложности 507 лиц. В качестве оппонента был выбран алгоритм Viola&Jones [1], построенный на прямоугольных примитивах Хора [1]. Данный алгоритм считается лучшим универсальным алгоритмом детектирования образов.

В первом эксперименте была проанализирована скорость детектирования лиц с помощью предложенных методов (табл. 1).

Таблица 1

Метод	Увеличение скорости, разы
Только метод усиления слабых классификаторов (AdaBoost)	3,8
AdaBoost + информации о форме и цвете объекта	21,7
AdaBoost + серия детекторов	6,5
Новый метод (гибридный)	28,90

Во втором эксперименте было проанализировано качество детектирования лиц с помощью предложенных методов (табл. 2).

Таблица 2

Метод	Ошибка, %
Только метод усиления слабых классификаторов (AdaBoost)	6,6
AdaBoost + информации о форме и цвете объекта	6,2
AdaBoost + серии детекторов	6,5
Новый метод (гибридный)	6,1

Анализ результатов эксперимента показал, что ошибка детектирования уменьшилась с применением дополнительных методов детектирования. Данное явление можно объяснить уменьшением количества ложных срабатываний при сохранении качества детектирования лиц в целом.

В третьем эксперименте (табл. 3) была проанализирована скорость детектирования нового метода в сравнении с алгоритмом Viola & Jones [1]. Предлагаемый метод превосходит по скорости детектирования метод, построенный на прямоугольных примитивах Хора.

Таблица 3

Метод	Скорость детектирования, к/с
Новый метод (гибридный)	28,90
Viola & Jones[1]	17,44

В заключительном эксперименте (табл. 4) было проанализировано качество детектирования нового метода в сравнении с алгоритмом Viola & Jones [1]. Результаты эксперимента свидетельствуют, что предлагаемый метод превосходит по качеству детектирования метод, построенный на прямоугольных примитивах Хора.

*Гребнов Сергей Викторович,*

ГОУВПО «Ивановский государственный энергетический университет имени В.И. Ленина»,  
аспирант кафедры программного обеспечения компьютерных систем,  
e-mail: Sergei.Grebnov@gmail.com

Таблица 4

Метод	Качество детектирования, %
Новый метод (гибридный)	93,90
Viola & Jones[1]	91,40

### Заключение

Представленный новый метод детектирования образов объединяет три идеи: использование совокупности анизотропных гауссовых примитивов, объединение полученных совокупностей в каскадную модель для увеличения скорости без ухудшения качества и применение серии дополнительных методов детектирования образов для предварительной фильтрации изображения. Предложенный метод, протестированный на стандартном тестовом наборе CMU/MIT Test set [6], превзошел по скорости и качеству детектирования все известные методы.

### Список литературы

1. **Paul Viola and Michael J. Jones.** Robust real-time face detection // *Int. J. Comput. Vision.* – 2004. – Vol. 57. – № 2. – P. 137–154.
2. **Kah Kay Sung and Tomaso Poggio.** Human Skin Color Detection in Images // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* – 1998. – Vol. 20. – № 1. – P. 39–51.
3. **Canny J.** A Computational Approach To Edge Detection // *Advances in Neural Information Processing Systems.*
4. **Peotta L., Granai L., and Vanderghyest P.** Very low bit rate image coding using redundant dictionaries // *Applications in Signal and Image Processing X.* – 2003. – Vol. 5207. – P. 228–239.
5. **Yoav Freund and Robert E. Schapire.** A decision-theoretic generalization of on-line learning and an application to boosting // *J. Comput. Syst. Sci.* – 1997. – Vol. 55. – № 1. – P. 119–139.
6. **Henry A. Rowley, Shumeet Baluja, and Takeo Kanade.** Neural network-based face detection // *IEEE Transactions on Pattern Analysis and Machine Intelligence.* – 1998. – Vol. 20. – № 1. – P. 23–38.