

Расширенный набор характеристик Хаара

Глеб Южаков

Факультет аэрокосмических исследований

Московский физико-технический институт (ГУ), Долгопрудный, Московская область, Россия
glebyu@gmail.com

Аннотация

В настоящее время большую известность приобрел алгоритм быстрого детектирования объектов на ускоренном каскаде элементарных характеристических классификаторов, известный как метод Виолы и Джонса. В данной работе будет рассмотрена техника выбора наиболее информативных характеристик из расширенного набора характеристик Хаара (haar-like features), используемых в данном методе.

Ключевые слова: Характеристики Хаара, Метод Виолы и Джонса.

1. ВВЕДЕНИЕ

В 2001 году П. Виола и М. Джонс [5] представили многоэтапную процедуру классификации, которая позволила существенно сократить вычислительное время. При этом качество осталось сопоставимо со многими более медленными и сложными одноэтапными классификаторами, такими как машина опорных векторов (SVM), случайный лес (Random Forest) и нейронные сети (ANN). Этапами алгоритма Виолы и Джонса являются классификаторы бустинга [2] над решающими деревьями, использующими в качестве признаков характеристики Хаара.

Определим характеристику Хаара как функцию f от суммарной интенсивности I_A и I_B двух прямоугольных участков изображения A и B , таких что участок B вложен в участок A . Прямоугольная форма участков выбрана затем, чтобы можно было применить технику интегральных изображений [4, 5] для расчета суммарных интенсивностей в них.

В современных работах используется характеристика вида $f(A, B) = \alpha I_A + \beta I_B$, где α и β – константы. Поэтому применение алгоритма Виолы и Джонса требует корректировки освещения:

$$I'(x, y) = \frac{I(x, y) - \mu}{c\sigma}, c > 0.$$

Здесь $I(x, y)$ – интенсивность в точке (x, y) , σ – оценочная дисперсия, μ – оценочное среднее значение интенсивности по некоторой окрестности, c – положительная константа, которую обычно полагают равной двум [4].

Вместо этого в данной работе для упрощения реализации алгоритма используется отношение суммарных интенсивностей:

$$f(A, B) = \frac{I_B}{I_A}.$$

Поскольку в данной статье производится сравнение нескольких подходов формирования характеристик, то эта модификация не может существенно повлиять на выводы.

Кроме того, в исходных работах [5] использовалось лишь несколько типов характеристик Хаара (рисунок 1).

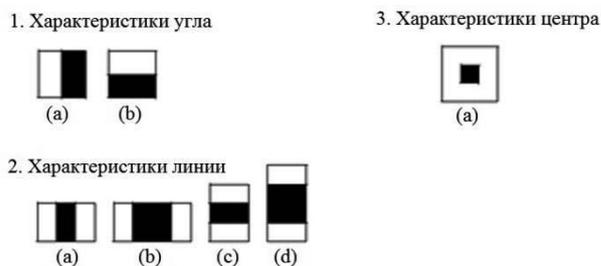


Рис. 1: Базовый набор характеристик Хаара.

В более поздних работах [4] набор характеристик Хаара был дополнен наклонными характеристиками (рисунок 2).

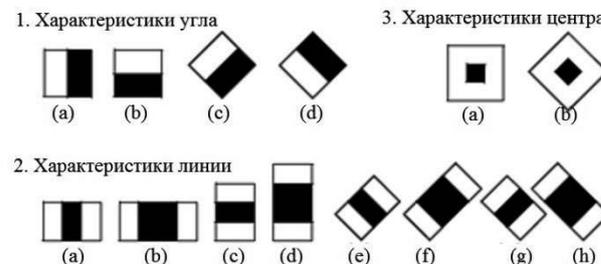


Рис. 2: Дополненный набор характеристик Хаара.

В данной работе предлагается отойти от идеи четкого задания нескольких типов характеристик Хаара и рассмотреть характеристики более общего вида, когда участок B имеет произвольное положение внутри участка A (рисунок 3).

2. ОПИСАНИЕ ХАРАКТЕРИСТИК

Рассмотрим скользящее окно, разбитое прямоугольной сеткой $W \times H$, внутри которого требуется проверить наличие изображения искомого объекта. Пусть имеется прямоугольник, заданный через координаты левого верхнего угла, ширину и высоту, $r = (x, y, w, h)$, где $0 \leq x \leq W - w, 0 \leq y \leq H - h, 0 < w \leq W, 0 < h \leq H$, $I(r)$ – сумма интенсивности пикселей в прямоугольнике r исследуемого изображения. Тогда множество характеристик Хаара можно представить в виде

$$F = \left\{ \frac{I(r_i)}{I(r_j)}, r_i \subset r_j \right\}.$$

Количество таких характеристик определяется по формуле

$$n = \sum_{w_A=1}^W \sum_{h_A=1}^H (W - w_A + 1)(H - h_A + 1)n(w_A, h_A),$$

$$\sum_{w_B=1}^{w_A} \sum_{h_B=1}^{h_A} (w_A - w_B + 1)(h_A - h_B + 1).$$

Отсюда

$$n = \frac{1}{576} W(W + 1)(W^2 + 5W + 6)H(H + 1)(H^2 + 5H + 6).$$

Но их слишком много, так как при размерах сетки 24×24 число характеристик достигает значения $n = 308002500$. Поэтому предлагается специальная техника выбора наиболее информативных характеристик Хаара за приемлемое вычислительное время.

3. ВЫЧИСЛЕНИЕ ХАРАКТЕРИСТИК

Под наиболее информативными характеристиками Хаара понимаются те, на которых достигается максимум критерия, характеризующего качество разделения объектов на положительные и отрицательные.

Так как характеристик слишком много, то техника выбора наиболее информативных характеристик состоит из трех ключевых моментов, позволяющих сократить временные и вычислительные затраты:

- 1) уточнение размера характеристик;
- 2) взятие случайного подмножества исследуемых характеристик;
- 3) аппроксимацией обучающей выборки обучающим подмножеством.

Рассмотрим их более подробно.

3.1 Уточнение размера характеристик

Пусть имеется две сетки, которыми разбиваются эталонные примеры: $W \times H$ и $2W \times 2H$. Тогда под уточнением понимается следующая процедура:

- 1) составление множества из k наиболее информативных (грубых) характеристик для сетки $W \times H$;
- 2) построение характеристик на сетке $2W \times 2H$, приближаемых полученными грубыми характеристиками;
- 3) выбор из этих характеристик наиболее информативных (точных) характеристик.

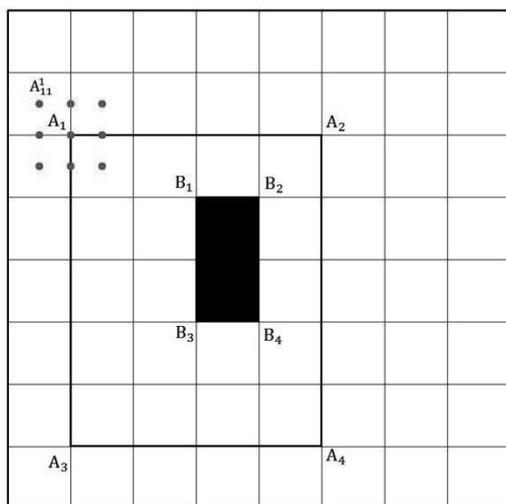


Рис. 3: Уточнение размера участков А и В.

На рисунке 3 прямоугольные участки A ($A_1A_2A_3A_4$) и B ($B_1B_2B_3B_4$) образуют грубую характеристику Хаара. Для уточнения рассматриваются все возможные смещения вершин составляющих грубую характеристику участков на половину размера ячейки сетки $W \times H$. В результате образуются участки вида $X_{ij}^1 X_{il}^2 X_{kj}^3 X_{kl}^4$, где $i, j, k, l = \{0, 1, 2\}$, $X = \{A, B\}$. Таким образом, каждая грубая характеристика образует до $9^4 = 6561$ уточняющих ее характеристик.

Поскольку требуется, чтобы участок B был вложен в A , уточняющих характеристик будет меньше. В случае грубого размера сетки 12×12 в среднем их число будет около 3000. При этом число различных (грубых) характеристик Хаара всего $n = 1863225$.

3.2 Взятие случайного подмножества исследуемых характеристик

Пусть каждая из возможных характеристик берется в рассмотрение с некоторой вероятностью p . В таком случае ожидаемое количество характеристик становится меньше: $n_p \approx pn$. Поскольку выбор характеристик производится при каждом обучении элементарного классификатора, данная процедура в среднем не ухудшает качество классификации, но уменьшает время обучения классификатора.

3.3 Аппроксимация обучающей выборки обучающим подмножеством

Еще одним методом ускорения выбора наиболее информативных характеристик является аппроксимация всей обучающей выборки некоторым его подмножеством. Так как у каждого обучающего примера имеется свой вес, то целесообразно в обучающее подмножество выбирать примеры с максимальным весом. Размер обучающего подмножества является фиксированным. В случае, когда имеется несколько примеров, претендующих попасть в обучающее подмножество и имеющих равные веса, решение об их принятии принимается случайно.

Пусть имеется положительная X_+ и отрицательная X_- обучающие выборки. Опишем процедуру выбора наиболее информативных характеристик с учетом аппроксимации:

- 1) составляются обучающие подмножества $T_+ \subset X_+$ и $T_- \subset X_-$ определенных размеров $n_+ = |T_+|$ и $n_- = |T_-|$;
- 2) определяется m наиболее информативных характеристик Хаара на примерах из обучающих подмножеств;
- 3) выбирается несколько наиболее информативных характеристик Хаара из этих m , но уже на примерах из всего обучающего множества X_+ и X_- .

Теперь остается на выбранных характеристиках Хаара построить очередной элементарный классификатор.

4. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Рассмотрим алгоритм быстрого детектирования объектов на ускоренном каскаде классификаторов бустинга над решающими деревьями, использующими характеристики Хаара. На рисунке 4 приведен сравнительный анализ качества (график точности от полноты precision – recall) работы данного алгоритма детектирования в зависимости от используемого метода выбора наиболее информативных характеристик Хаара на сетке размером 24×24 (в случае применения уточняющего метода бралась сетка размером 12×12).

В работе использовались открытые выборки LFW [3] и VOC [1]. В качестве источника положительных обучающих и тестовых примеров была взята открытая выборка LFW, содержащая лица. В качестве источника отрицательных обучающих и тестовых примеров была взята выборка VOC, не содержащая лица. Положительная обучающая выборка состояла из 1000 примеров, тестовая из 300. Обе отрицательные выборки состояли из 500 несовпадающих изображений, на которых производился поиск искомого объекта и выборов соответственно.

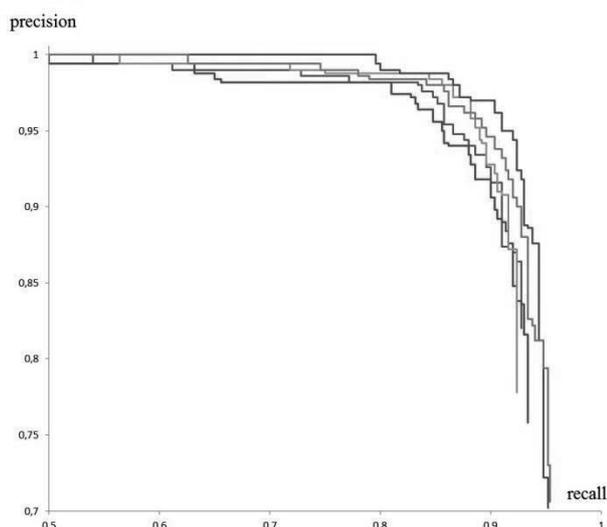


Рис. 4: Сравнительные результаты качества работы алгоритма детектирования в зависимости от метода выбора наиболее информативных характеристик Хаара. Информация о соответствии цвета кривой и метода содержится в таблице 1.

Табл. 1: Сравнительные результаты алгоритмов детектирования лиц в зависимости от метода выбора наиболее информативных характеристик Хаара (названия полей таблицы уточнены ниже).

Метод	Площадь под графиком	Время обработки, мс	Время обучения, часов
Синий	0.918	25.0	6.2
Голубой	0.942	22.8	4.5
Красный	0.917	27.6	4.0
Зеленый	0.921	28.4	3.3
Фиолетовый	0.945	23.7	6.3

В первом столбце таблицы содержится информация о соответствии метода и цвета, соответствующего ему графика точности от полноты на рисунке 4. А именно, синему цвету соответствует поиск наиболее информативных характеристик Хаара, использующий расширенный набор с уточнением и случайным взятием (10%), голубому – расширенный набор с уточнением, красному – базовый набор, зеленому – дополненный наклонными характеристикам набор, фиолетовому – дополненный и расширенный набор характеристик Хаара с уточнением и случайным взятием.

Во втором столбце дана площадь под графиком точности от полноты (precision – recall) соответствующего метода.

В третьем столбце приведено среднее время обработки одного изображения размером 500×500 , процессор Intel Core i3 – 2120, 3.3 GHz. При этом минимальный размер искомого объекта 96×96 .

В четвертом столбце приведено время в часах, которое требуется, чтобы алгоритм обучился.

Из таблицы 1 видно, что точность классификатора, использующего метод выбора наиболее информативных

характеристик из расширенного набора характеристик Хаара, выше, чем у классификатора, использующего при обучении дополненный и базовый набор характеристик Хаара, 0.942 в сравнении с 0.921. При этом он тратит меньше времени на обработку одного изображения, 22.8 мс в сравнении с 28.4 мс. Но зато ему требуется больше времени на обучение, 4.5 часа против 3.3.

Наилучшее качество (0.945) показал классификатор, использующий все три описанных метода выбора наиболее информативных характеристик, хотя он требует и наибольшее время на обучение (6.3 часа).

5. ЗАКЛЮЧЕНИЕ

Таким образом, описанная выше техника является эффективным методом выбора наиболее информативных характеристик Хаара. В сравнительном эксперименте при детектировании лиц у метода Виолы и Джонса, использующего данную технику, наблюдается более высокое качество детектирования и меньшее время, требующееся на обработку одного кадра. Единственным недостатком, но не очень существенным, является увеличение времени обучения классификатора.

6. ССЫЛКИ

- [1] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn and A. Zisserman, The PASCAL Visual Object Classes (VOC) Challenge, International Journal of Computer Vision, 88(2), 303-338, 2010.
- [2] Y. Freund, R. E. Schapire, A Short Introduction to Boosting, Shannon Laboratory, USA, 1999., pp. 771-780
- [3] G. B. Huang, M. Ramesh, T. Berg and E. Learned-Miller, Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. University of Massachusetts, Amherst, Technical Report 07-49, 2007.
- [4] 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.
- [5] P. Viola, M. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, – conference on computer vision and pattern recognition, 2001.

Об авторах

Глеб Южаков – аспирант ФАКИ МФТИ, научный сотрудник ФГУП ЦНИИХМ. Его адрес: glebyu@gmail.com.