

УДК 004.93

<https://doi.org/10.36906/AP-2020/36>

ИСПОЛЬЗОВАНИЕ КАСКАДОВ ХААРА ДЛЯ РАСПОЗНАВАНИЯ ОБРАЗОВ

Туманов В. И.

*Нижневартровский государственный университет**г. Нижневартовск, Россия*

Катермина Т. С.

*канд. техн. наук**Нижневартровский государственный университет**г. Нижневартовск, Россия*

Аннотация. Данная статья посвящена каскадам Хаара и их использованию для определения объектов на изображении, изучению принципов для уменьшения времени работы алгоритмов распознавания, проведению тестирования работы алгоритма, использующего каскады Хаара и анализу работоспособности метода.

Ключевые слова: распознавание образов; признаки Хаара; классификация.

В мире около 4,6 миллиарда человек являются пользователями интернета — это составляет 59% населения планеты. Количество зарегистрированных пользователей в Instagram на конец 2019 г — 2,8 млрд. человек, что является 36% населения планеты и 61% пользователей Интернета. Ежеминутно публикуется около 80 тыс. постов (<https://www.internetlivestats.com/>). Основным контентом является фото- и видеоматериалы, которые затрагивают зрительное восприятие информации человеком.

С введением дистанционного обучения проявились проблемы ведения занятий для людей с ограниченными возможностями здоровья из-за отсутствия физического взаимодействия между участниками процесса. Существующие сервисы Text-To-Speech (Текст-В-Голос) не способны осуществлять распознавание изображений так, как это может человек, поэтому текстовая и графическая информация будет восприниматься хуже.

Согласно данным медицинского научного журнала The Lancet количество слепых к 2050 году может достигнуть 115 миллионов человек [4]. Поэтому важно начинать разработку и изучение технологий, способных упростить жизнь людям с ограничениями по зрению [1].

Существуют различные подходы к распознаванию образов [2]:

1. По форме

Шаблон, имеющий форму объекта поиска, перемещается по изображению и в каждой позиции рассчитывается характеристика и, при прохождении определенного порога, строится предположение, что в данной области может находиться искомый объект. Для расчета характеристики чаще всего используется среднеквадратичная разность значений яркости шаблона и исходного изображения:

$$y = \frac{\sqrt{\sum_{i=0}^{\infty} (pixel1_i - pixel2_i)^2}}{n - 1}$$

Недостатком является ресурсоемкость, так как шаблон проходится по большому числу позиций, а также проблемы определения при различных размерах шаблона и искомого объекта.

Поэтому существует менее ресурсоемкий подход, использующий аналитическое описание формы вместо шаблона. Набору точек на изображении ставится в соответствие геометрическая форма и сравниваются их границы.

2. По положению

Подход требует существование какой-то «точки отсчета», которая является легко искомой меткой, и относительно этой точки строится предположение о положении остальных объектов на изображении. Например, если посчитать рамку автомобильных госномеров как метку, то можно сделать вывод, что само значение номера находится внутри нее. Или использовать нос на лице в качестве метку, чтобы определить положение глаз и рта.

3. По цвету

Некоторые объекты можно определить при помощи их цвета. Их окраска может быть определена достаточно четко. Используя различные цветовые базисы, данный подход дает высокую результативность и позволяет практически безошибочно распознать объект. Недостатком является необходимость выбора подходящего цветового базиса для конкретной ситуации.

Метод Виолы-Джонса строится на двух из вышеперечисленных подходов и основывается на примитивах Хаара, которые представляют собой наборы прямоугольных областей с различными характеристиками яркостей пикселя [3]. Метод использует монохромные изображения, так как это дает прирост в эффективности по времени распознавания.

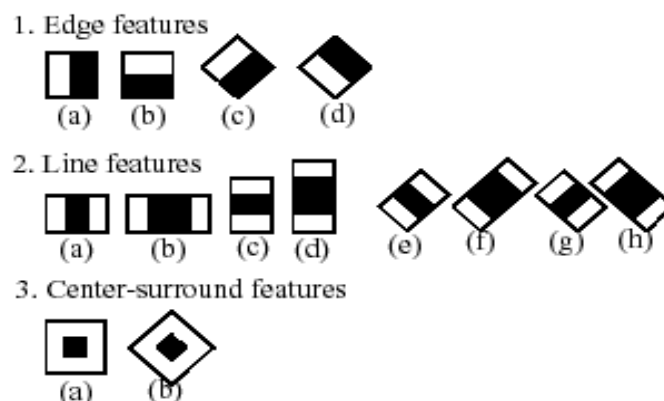


Рис. 1. Представление каскадов Хаара

Так как изображение имеет большое число таких примитивов, используется модифицированный алгоритм AdaBoost, позволяющий сфокусироваться на небольшом числе важных примитивов.

Затем отобранные примитивы формируются в каскадную структуру для их использования в более сложных алгоритмах и снижения нагрузки на вычислительную машину.

На рисунке 2 первый и второй признак были выбраны алгоритмом AdaBoost и наложены на изображение на нижней строчке. Первый признак замеряет разницу яркости региона пикселей глаз и регионом у верхней части щек. Признак строится на том, что региона глаз обычно более темный, чем регион щек. Второй признак действует подобным способом, но уже строится на том, что регионы глаз более темные, чем регион переносицы.

Для проверки эффективности был проведен тест по определению лиц и автомобильных госномеров.

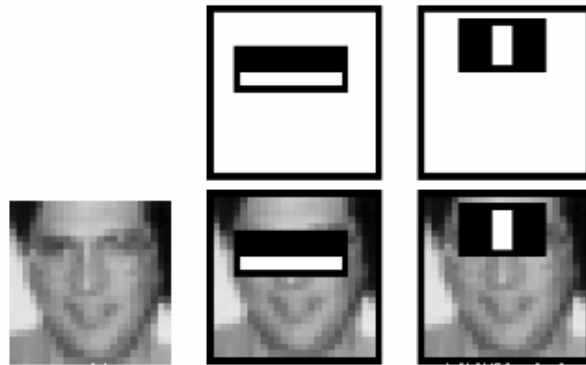


Рис. 2. Каскады Хаара на примере распознавания лица

Таблица 1

Тесты изображений с лицами

№ изображения	1	2	3	4	5	6	7	8	9	10
Распознано лицо	+	+	+	+	-	+	+	+	+	+
Распознаны глаза	++	++	++	++	-	+-	--	-+	--	++
Ложные срабатывания	+	+	-	-	-	-	-	-	-	-

Тесты изображений с лицами показали высокий результат, не удалось определить лицо лишь на одной фотографии. Причинами может являться то, что область рта прикрыта шарфом и лицо частично повернуто в сторону.



Рис. 3. Изображение, на котором не удалось распознать лицо

Таблица 2

Тесты изображений с автомобильными госномерами

№ изображения	1	2	3	4	5	6	7	8	9	10
Распознан госномер	+	+	+	+	+	+	+	+	+	+
Ложные срабатывания	-	+	-	-	-	-	-	-	-	-

Тесты изображений с госномерами показали еще лучший результат, среди выбранных изображений на всех удалось определить область госномера. И лишь на одном изображении было два ложных срабатывания. Причиной может являться то, что ложные области имели более темную границу по краям, светлую часть внутри границы и темные части в центре, что достаточно похоже на представление госномера.

Литература

1. Разинкин В. Б., Катермина Т. С. Распознавание лица по фотографии // International journal of advanced studies. 2018. №1-2. С. 171-180.
2. Ту Дж., Гонсалес Р. Принципы распознавания образов Принципы распознавания образов. М.: Мир, 1978.
3. 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, 2001. Vol. 1. P. I-I. <https://doi.org/10.1109/CVPR.2001.990517>
4. Flaxman S. R. et al. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis // The Lancet Global Health. 2017. Vol. 5. №12. P. e1221-e1234. [https://doi.org/10.1016/S2214-109X\(17\)30393-5](https://doi.org/10.1016/S2214-109X(17)30393-5)

©Туманов В. И., Катермина Т. С., 2020