

Амеличев Глеб Эдуардович, студент-магистр, Калужский филиал ФГБОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»

Панина Виктория Сергеевна, студент-бакалавр, Калужский филиал ФГБОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»

Белов Юрий Сергеевич, к.ф. -м.н., доцент, Калужский филиал ФГБОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)»

РАСПОЗНАВАНИЕ ЛИЦ С ИСПОЛЬЗОВАНИЕМ КАСКАДОВ ХААРА

Аннотация: Обнаружение объектов – это компьютерная технология, связанная с компьютерным зрением, обработкой изображений и глубоким обучением, которая занимается обнаружением экземпляров объектов на изображениях и видео. В этой статье рассмотрены каскадные классификаторы Хаара, которые являются эффективным способом обнаружения объектов. Каскад Хаара – это подход, основанный на машинном обучении, при котором для обучения классификатора используется множество положительных и отрицательных образов.

Ключевые слова: распознавание лиц, компьютерное зрение, каскад Хаара.

Annotation: Object detection is a computer technology related to computer vision, image processing, and deep learning that deals with detecting instances of objects in images and videos. This article discusses cascading Haar classifiers, which are an effective way to detect objects. The Haar cascade is a machine

learning approach that uses a variety of positive and negative images to train the classifier.

Key words: face detection, computer vision, Haar cascade.

Компьютерное зрение является междисциплинарной областью, связанной, с искусственным интеллектом, машинным обучением, робототехникой, обработкой сигналов и геометрией. Цель компьютерного зрения – запрограммировать приложение так, чтобы оно «понимало» сцену или особенности изображения [4].

Существует различные практические применения компьютерного зрения:

- Автономные транспортные средства – это одно из наиболее важных приложений компьютерного зрения, где автомобили с автоматическим управлением должны собирать информацию об окружающей среде, чтобы решить, как себя вести.

- Распознавание лиц – это также очень важное приложение компьютерного зрения, где электронные средства используют технологию распознавания лиц, в основном чтобы подтвердить личность пользователя.

- Поиск изображений и распознавание объектов. Теперь существует возможность искать объекты на изображении, используя поиск по изображению. Хорошим примером является Google Lense (объектив). Google Lens (объектив) анализирует содержимое ваших фотографий и отображает соответствующую информацию, связанную с объектами [6].

- Робототехника. Большинству роботизированных машин, в производстве, часто необходимо «видеть» свое окружение, чтобы выполнить поставленную задачу. В производстве эти машины являются запрограммированными для осуществления узкоспециализированных заданий, таких как подсчеты объектов на конвейерах, чтения серийных номеров либо поиска поверхностных дефектов [7].

Обнаружение объектов с использованием каскадных классификаторов на основе признаков Хаара – это эффективный метод обнаружения объектов,

предложенный Полом Виолой и Майклом Джонсом. Это подход, основанный на машинном обучении, где каскадная функция обучается на основе множества положительных и отрицательных изображений. Затем эта функция используется для обнаружения объектов на других изображениях. Первоначально, алгоритм требует довольно много положительных изображений (изображений лиц) и отрицательных изображений (изображений без лиц) для обучения классификатора [1]. Далее необходимо извлечь из него особенности. Для этого используются признаки Хаара, показанные на рисунке ниже. Каждый объект представляет собой одно значение, полученное путем вычитания суммы пикселей под белым прямоугольником из суммы пикселей под черным прямоугольником [5; 8].

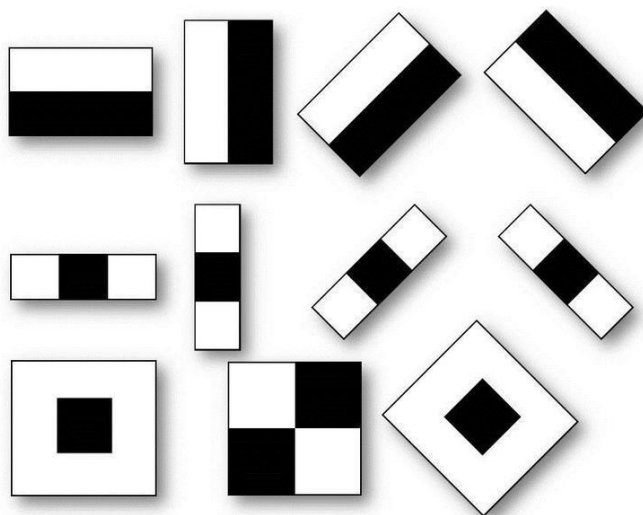


Рисунок 1. Признаки Хаара

Идея состоит в том, что каскад Хаара извлекает объекты из изображений с помощью какого-то фильтра. Эти фильтры называются функциями Хаара и на рисунке 2 можно увидеть теоретическую модель лица:

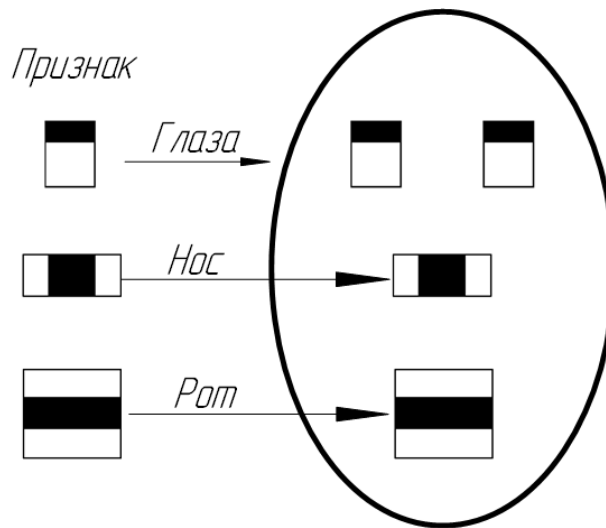


Рисунок 2. Теоретическая модель лица

Признаки Хаара состоят из смежных прямоугольных областей. Они позиционируются на изображении, далее суммируются интенсивности пикселей в областях, после чего вычисляется разность между суммами. Эта разность и будет значением определенного признака, определенного размера, определенным образом, помещенным на изображение [9; 10].

Идея этого метода состоит в том, чтобы проверять каждую область изображения одновременно. Затем для каждой области получается одно значение путем вычитания суммы пикселей под белым прямоугольником из суммы пикселей под черным прямоугольником. В идеале большое значение функции означает, что она актуальна. Ниже можно увидеть очень простой пример того, как применяются фильтры:

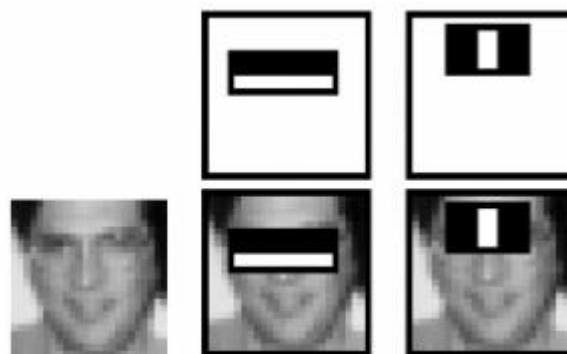


Рисунок 3. Применение фильтров Хаара

Две функции в верхнем ряду накладываются на тренировочную область в нижнем ряду. Как видно из рисунка, первая функция используется для измерения разницы в интенсивности между областью вокруг глаз и верхними щеками, поскольку область вокруг глаз зачастую темнее. Вторая сравнивает интенсивность в области глаза и переносицы.

Несмотря на то, что идея данного подхода хороша, ее практическая реализация нетривиальна. При столкновении с данным методом могут возникнуть проблемы с определением значимых функций.

Каскадирование – это метод последовательного объединения более сложных классификаторов в каскадной структуре, приводящий к значительному увеличению скорости детектора за счет фокусирования внимания на перспективных областях изображения. Известно, что если подэлементы плохо определяются в простом классификаторе, то они не будут хорошо определяться в более сложном классификаторе. В каскадном методе не будет применяться более сложный классификатор к подэлементам, пока простой классификатор не будет работать хорошо. Таким образом, каскадирование гарантирует, что ресурсы не тратятся на незначительные подэлементы поскольку они никогда не попадут в более сложные классификаторы.

Для расчета каждого объекта необходимо найти сумму пикселей под белыми и черными прямоугольниками. Чтобы ускорить этот процесс используется интегральное представление изображения. Это значительно упрощает вычисление суммы пикселей, поскольку для того, чтобы сложить большое число пикселей, достаточно использовать лишь четыре пикселя [11].

Основная идея распознавания лиц на основе каскадной стратегии заключается в быстром исключении большинства фоновых областей на ранней стадии, для того, чтобы уменьшить объем вычислений на более поздней стадии. На ранней стадии распознавания лица алгоритм должен соответствовать трем следующим требованиям:

1. Быстрое обнаружение: как известно, лицо имеет разные размеры и разное расположение, и это легко обнаруживается в результате

масштабирования и скользящего окна. В начале объем вычислений велик, так как изображения разных масштабов необходимо обнаружить в каждой позиции, поэтому скорость раннего обнаружения определяет скорость всей системы обнаружения [3].

2. Высокая частота обнаружения (DR). Раннее обнаружение представляет собой предварительная проверка. В ходе этого процесса область, рассматриваемая как не лицо, отбрасывается, а большая часть, где находится лицо сохраняется.

3. Высокая способность исключать не лицевую область. Чем больше не лицевой области можно исключить на ранней стадии, тем выше скорость всей системы распознавания лиц [2].

Требования 2 и 3 противоречивы, поэтому между ними должен быть установлен баланс. Для этого вводятся две концепции: скорость обнаружения лица (FDR) и скорость отклонения фона (BRR). FDR означает, какой процент обнаруженных образцов лица приходится на общее количество образцов; BRR означает, какой процент обнаруженных образцов без лица приходится на общее количество образцов без лица.

Для повышения эффективности работы существует способ, заключающийся в комбинировании слабых функций, которые строятся в ходе итеративного процесса, где на каждом шаге новая модель обучается с использованием данных об ошибках предыдущих. Таким образом, это метод получения одного очень точного классификатора путем объединения грубого и умеренно неточного классификатора [12].

Выводы. Каскадный подход Хаара имеет ряд преимуществ. Для обработки больших баз данных он является лучшим детектором с точки зрения скорости и надежности. Различное освещение является причиной плохого распознавания, если использовать каскадный классификатор Хаара, результаты распознавания лиц будут более точными.

Библиографический список:

1. Гришанов К.М., Белов Ю.С., Модель сверточной нейронной сети в задачах машинного зрения // Электронный журнал: наука, техника и образование, 2017, №СВ1(11), С. 100 – 106.
2. Кузнецов Г.С., Белов Ю.С. Обзор метода 3D распознавания лиц без преобразования лицевой поверхности // Электронный журнал: наука, техника и образование, 2016, №2(6)., С. 104 – 110.
3. Редько А.В., Молчанов А.Н., Белов Ю.С. Использование алгоритмов определения ключевых точек изображения в задаче реконструкции трехмерных сцен // Электронный журнал: наука, техника и образование, 2016, №1(5)., С. 94 – 101.
4. Biometrics Institute. "Biometrics Institute Privacy Code." (March 4, 2010) <http://www.biometricsinstitute.org/displaycommon.cfm?an=1&subarticlenbr=8>.
5. Mordini, Emilio and Massari, Sonia. "Body, Biometrics and Identity." Bioethics. Vol. 22, Issue 9, 2008. pp. 488-498.
6. Duarte T (2016) Biometric access control systems: A review on technologies to improve their efficiency. power Electronics and Motion Control Conference (PEMC).
7. Napoléon T., Alfalou A. Pose invariant face recognition: 3D model from single photo. Opt. Lasers Eng. 2017; 89:150–161. doi: 10.1016/j.optlaseng.2016.06.019.
8. 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; Kauai, HI, USA. 8–14 December 2001.
9. Ouerhani Y., Alfalou A., Brosseau C. Optics and Photonics for Information Processing XI. Volume 10395. SPIE; Bellingham, WA, USA: 2017. Road mark recognition using HOG-SVM and correlation; p. 103950Q. International Society for Optics and Photonics.
10. Shah J.H., Sharif M., Raza M., Azeem A. A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques. Int. Arab J. Inf. Technol. 2013; 10:536–545

11. Smach F., Miteran J., Atri M., Dubois J., Abid M., Gauthier J.P. An FPGA-based accelerator for Fourier Descriptors computing for color object recognition using SVM. *J. Real-Time Image Process.* 2007; 2:249–258. doi: 10.1007/s11554-007-0065-6.