

Клюев В.В.

студент магистратуры

1 курс, институт информатики и телекоммуникаций

Сибирский государственный университет науки и технологий имени

академика М.Ф. Решетнева

Россия, г. Красноярск

ОБНАРУЖЕНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ С ПОМОЩЬЮ ГИСТОГРАММЫ НАПРАВЛЕННЫХ ГРАДИЕНТОВ

Аннотация: В статье рассматривается метод обнаружения объектов на изображениях, основанный на гистограмме направленных градиентов.

Ключевые слова: компьютерное зрение, обнаружение объектов, HOG, SVM.

Annotation: The article discusses a method for detecting objects in images based on a histogram of oriented gradients (HOG).

Keywords: computer vision, object detection, HOG, SVM.

Идея алгоритма заключается в том, что внешний вид и форма объекта на участке изображения могут быть описаны распределением градиентов интенсивности или направлением краев [2]. Изображение делится на небольшие связанные области, называемые ячейками, и для пикселей в каждой ячейке составляется гистограмма направлений градиента. Объединение этих гистограмм является дескриптором.

Основные этапы расчета дескриптора HOG:

1. Расчет значений градиентов.

Наиболее часто применяемый метод – использование одномерной дифференцирующей маски с использованием фильтрующих ядер $[-1, 0, 1]$ и

$[-1, 0, 1]^T$. Используя полученные матрицы можно с легкостью вычислить величину (1) и направление (2) градиента.

$$g_r = \sqrt{g_x^2 + g_y^2} \quad (1)$$

$$g_\varphi = \tan^{-1} \frac{g_y}{g_x} \quad (2)$$

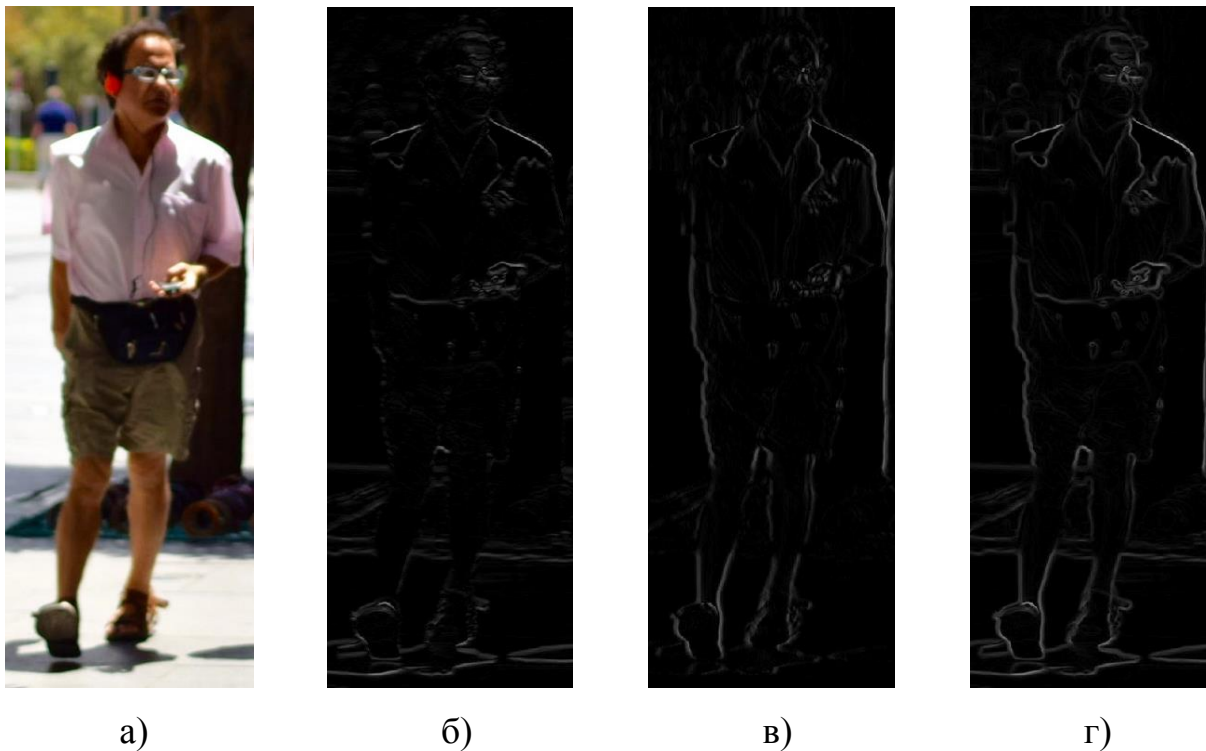


Рисунок 1. Пример вычисления градиентов: а – исходное изображение; б – абсолютное значение у-градиента; в – абсолютное значение х-градиента; г – абсолютная величина градиента.

2. Группировка направлений.

На данном этапе вычисляются гистограммы ячеек. В статье Далала и Триггса [1] оптимальным размером ячейки для нахождения пешеходов является 8×8 . Используется каждое значение градиента, находящееся в вычисляемой ячейке. Каналы гистограммы равномерно распределены по диапазону от 0 до 180 градусов или от 0 до 360 градусов, в зависимости от того, является ли градиент знаковым.

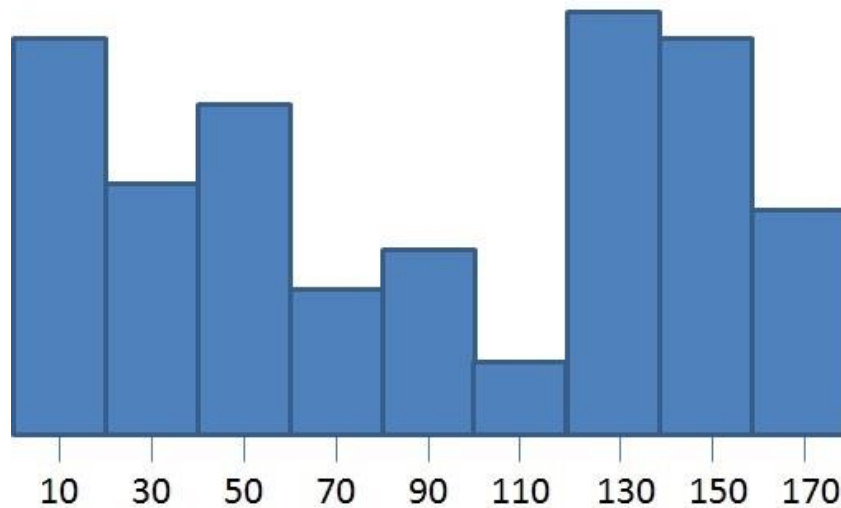


Рисунок 2. Пример гистограммы ячейки [3]

3. Группировка ячеек в блоки для дальнейшей нормализации.

Чтобы учесть изменения освещенности и контраста, силы градиента должны быть локально нормализованы, что требует группировки ячеек в более крупные, пространственно связанные блоки. Эти блоки обычно наслаиваются один на другой, что означает, что каждая ячейка вносит вклад в окончательный дескриптор более одного раза.

Далал и Триггс определили [1] 4 способа поиска нормировочного множителя: L2-норма (3), L2-норма, ограниченная сверху, L1-норма (4), корень из L1-нормы.

$$f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \quad (3)$$

$$f = \frac{v}{(\|v\|_1 + e)} \quad (4)$$

Все предложенные методы нормализации улучшают результаты, однако установлено [1], что L1-норма работает менее надежно, чем остальные.

В результате работы данного алгоритма, изображение преобразуется в характеристический вектор. Длина вектора зависит не только от размера входного изображения, а также от принятых размеров блока и ячейки.



(а)



(б)

**Рисунок 3. Результат построения гистограмм направленных градиентов:
а – исходное изображение; б – визуализация работы алгоритма.**

4. Классификация дескрипторов при помощи системы обучения с учителем.

Для тренировки алгоритма требуется большая выборка изображений как содержащих искомый объект, так и на которых его нет. Существует множество алгоритмов обучения, однако основная идея заключается в бинарной классификации характеристических векторов. Дихотомическое разделение дескрипторов на имеющие обнаруживаемый объект и на которых он отсутствует.

Наиболее популярным является метод опорных векторов (SVM). Характеристические векторы представлены точками в p -мерном пространстве. Каждая точка может принадлежать только одному из двух классов. Задача алгоритма состоит в том, чтобы найти гиперплоскость с размерностью $(p-1)$, которая разделит все точки на два класса.

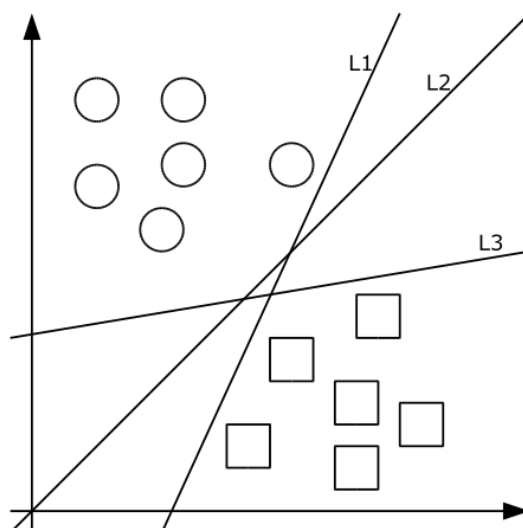


Рисунок 4. Пример классифицирующих прямых в двумерном пространстве [4]

Гистограмма направленных градиентов является мощным способом преобразования изображения в характеристический вектор (дескриптор). Основанный на ней метод распознавания является крайне эффективным для объектов, которые не демонстрируют существенное изменение в форме. Если ориентация рассматриваемого объекта постоянно различается или он вращается, то данный метод не рекомендуется к использованию.

Использованные источники:

1. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 1, pages 886–893, June 2005.
2. Histogram of oriented gradients // Wikipedia. URL: https://en.wikipedia.org/wiki/Histogram_of_oriented_gradients (дата обращения: 3.02.2019).
3. HOG Person Detector Tutorial // Chris McCormick. URL: <http://mccormickml.com/2013/05/09/hog-person-detector-tutorial/> (дата обращения: 3.02.2019).

4. Support-vector machine // Wikipedia. URL:
https://en.wikipedia.org/wiki/Support-vector_machine (дата обращения:
3.02.2019).