

МАШИННОЕ ОБУЧЕНИЕ И ЕГО ПРИМЕНЕНИЕ

Ярмак В.С., Юхимук А.М.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Анисимов В.Я. – канд. физ.-мат. наук

В данной научной работе рассматривается применение машинного обучения и анализ четырех алгоритмов, входящих в библиотеку OpenCV для работы в этой области. Работа включает экспериментальное исследование, в ходе которого описанные методы и алгоритмы применяются на практике. Алгоритмы, рассмотренные в работе, включают детектор границ Кэнни, преобразование Хафа и метод Виолы-Джонса.

Машинное обучение - это раздел искусственного интеллекта, который разрабатывает и применяет алгоритмы и модели, способные учиться на данных и использовать их для принятия решений и выполнения задач без явного программирования. Рассмотрим некоторые алгоритмы, которые применяются для обработки изображений в этой сфере.

Выделение границ изображения детектором Кэнни

Оператор Кэнни (алгоритм Кэнни) – оператор обнаружения границ изображения. Перед применением метода исходное изображение подвергается обработке. Изначально рисунок, представленный в цветовой модели RGB, преобразуют в модель YUV [1]. Затем применяется метод нормализации гистограммы изображения для согласования яркости [2].

После подготовки изображения, алгоритм нахождения границ Кэнни использует следующие шаги: 1) сглаживание; 2) поиск градиентов; 3) подавление «ложных» максимумов; 4) двойная пороговая фильтрация; 5) трассировка области неоднозначности.

Для сглаживания изображения перед применением оператора Кэнни используется фильтр Гаусса. Размытие по Гауссу – это фильтр размытия изображения, который использует нормальное распределение для вычисления преобразования, применяемого к каждому пикселю изображения (1).

$$Gauss(x, y, \sigma) = \frac{e^{-\frac{x^2+y^2}{2\sigma^2}}}{2\pi\sigma^2} \quad (1)$$

где x^2+y^2 – это радиус размытия; σ – стандартное отклонение распределения Гаусса.

Выбор границ происходит в тех местах, где наблюдается максимальное значение градиента, которое зависит от направления поиска. Поэтому в алгоритме Кэнни выделяются четыре ядра фильтра, отвечающие за горизонтальное, вертикальное и два диагональных направления.

$$G = \sqrt{G_x^2 + G_y^2} \quad (2)$$

$$\theta = \arctg\left(\frac{G_y}{G_x}\right), \quad (3)$$

где G_y и G_x - два изображения, на которых любая точка имеет в себе приближенные производные по x и y , а θ - направление градиента. Градиент определяет, как быстро изменяется яркость изображения в каждой точке, что дает возможность определить границу и ее ориентацию.

Затем определяется угол направления вектора границы. Направления вектора округляется до ближайшего угла, кратного 45° . Затем проверяется достижение локального максимума величиной градиента в найденном направлении вектора [1].

Метод использования порогов заключается в определении областей на изображении, где могут располагаться границы, за счет пороговых значений.

Алгоритм Кэнни использует двойную пороговую фильтрацию: только пиксели, попавшие в средний диапазон, т. е. выше нижнего порога и ниже верхнего, относятся к границе. Все другие пиксели не исключаются [1]. Все пиксели, прошедшие предыдущие этапы, проверяются на близкое расположение друг к другу. Пиксели, лежащие обособленно подавляются.



Рисунок 1 – Пример входного изображения и результата после фильтра Кэнни

Поиск прямых линий с помощью преобразования Хафа

Классическое преобразование Хафа используется для того, чтобы выделить на изображении группы точек, образующие те или иные геометрические фигуры, и найти их параметры [3].

Прямая на плоскости описывается уравнением $y=kx+b$ и может быть задана парой несовпадающих точек. Во втором случае существует проблема описания вертикальных прямых, т.к.

для них параметры k и b бесконечны. Поэтому удобнее представить прямую с помощью других параметров r и q , где r - это длина перпендикуляра, опущенного на прямую из начала координат, а q - это угол между данным перпендикуляром и осью x .

Все прямые, проходящие через точку с координатой (x_0, y_0) , соответствуют следующему уравнению: $r(\theta) = x_0 \cos \theta + y_0 \sin \theta$. Это соответствует и синусоидальной кривой в пространстве (r, θ) .

Проблемы обнаружения коллинеарных точек можно упростить, преобразуя их в задачу обнаружения пересекающихся синусоид. Каждая точка в пространстве (r, θ) соответствует параметрам прямой, проходящей через нее. Путем анализа плотности точек в пространстве можно определить прямые на изображении. Дискретное растровое представление изображения приводит к сгущению точек в пространстве (r, θ) , где находятся параметры прямой. Выбор наиболее "плотных" областей на изображении позволяет определить параметры соответствующих прямых.

Алгоритм детектирования лица Виолы-Джонса

Метод Виолы-Джонса — алгоритм, позволяющий обнаруживать объекты на изображениях в реальном времени. Основной задачей алгоритма при его создании было обнаружение лиц.

Детектор Виолы-Джонс обладает высокой скоростью распознавания и низкой вероятностью ложных обнаружений. Однако он имеет ограничение в распознавании лиц под углом более 30° , что затрудняет его использование в системах с повышенными потребностями.

Рассмотрим принцип сканирующего окна:

Имеется изображение, на котором есть искомые объекты. Оно представлено двумерной матрицей пикселей размером $w \times h$, в которой каждый пиксель имеет значение: 1) от 0 до 255, если это черно-белое изображение; 2) от 0 до 2553, если это цветное изображение (компоненты R, G, B).

в результате своей работы, алгоритм должен определить лица и их черты и пометить их – поиск осуществляется в активной области изображения прямоугольными признаками, с помощью которых и описывается найденное лицо и его черты: $rectangle = \{x, y, w, h, a\}$.

где x, y – координаты центра i -го прямоугольника, w – ширина, h – высота, a – угол наклона прямоугольника к вертикальной оси изображения.

Интегральное изображение $ii(N \times N)$, построенное по исходному изображению $I(N \times N)$ – матрица, каждый элемент которой равен сумме значений яркости соответствующих пикселей на исходном изображении, расположенных слева и сверху от него:

$$ii = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (5)$$

Мы протестировали описанные алгоритмы, которые используются в библиотеке OpenCV, на примере программы, распознающей номера машин, результат ее работы приведен на рисунке 2.

В данном случае используется алгоритм Кэнни для обнаружения границ изображения. Затем с помощью алгоритма преобразования Хафа были найдены все прямоугольники на изображении, а затем выбран необходимый. Далее, используя фреймворк easyOCR текст был проанализирован.



Рисунок 3 – Поэтапные результаты выполнения программы

Машинное обучение активно развивается и прогрессирует в современном мире. С появлением новых методов и алгоритмов, а также соответствующих вычислительных ресурсов, возможности машинного обучения значительно расширяются.

Список использованных источников:

1. John Canny. A computational approach to edge detection // *IEEE Transactions on pattern analysis and machine intelligence*. – 1986. – V. 8, № 6. – P. 679–698. .
2. Otsu, N. A Threshold Selection Method from Gray-Level Histograms // *IEEE Transactions on Systems, Man, and Cybernetics*. – 1979. – V. 9, № 1. – P. 62–66.
3. Анна Дегтярева, Владимир Вежневцев. Преобразование Хафа (Hough transform). Компьютерная графика и мультимедиа. Выпуск №1(1)/2003. [Электронный ресурс]. – Режим доступа : <http://cgm.computergraphics.ru/content/view/36>.