

НЕКОТОРЫЕ СПОСОБЫ РЕАЛИЗАЦИИ КОМПЬЮТЕРНОГО ЗРЕНИЯ, ИСПОЛЬЗУЮЩИЕ МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Д. Р. Алиев,^{*} Д. А. Возжаева,[†] А. С. Рябцева,[‡] В. С. Васильев[§]

Компьютерное зрение - это одна из самых захватывающих областей искусственного интеллекта, которая помогает создавать системы, способные воспринимать и обрабатывать изображения и видео. Она существует уже на протяжении многих лет, и в последнее время наблюдается резкий рост развития технологий и алгоритмов в этой области. Сегодня компьютерное зрение используется в различных сферах жизни, начиная от медицины и производства, заканчивая транспортом и безопасностью. Однако, несмотря на то, что эти технологии уже достаточно распространены, компьютерное зрение по-прежнему является актуальной темой, которая представляет большой интерес для исследователей и разработчиков в этой области.

Ключевые слова: компьютерное зрение, машинное зрение, машинное обучение

ВВЕДЕНИЕ

Для начала, разберемся с определениями. Итак, что же такое компьютерное зрение? По определению, компьютерное зрение это теория и технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов на основе фотографий или видео. Также вы наверняка слышали термин “машинное зрение” - научное направление в области искусственного интеллекта и связанные с ним технологии получения изображений объектов реального мира, их обработки и использования полученных данных для решения разного рода прикладных задач без участия человека. На первый взгляд может показаться, что эти определения эквивалентны. Дело в том, что компьютерное зрение это набор методов анализа визуальных данных, а машинное зрение в свою очередь их прикладное применение. Сегодня мы поговорим о нескольких способах реализации компьютерного зрения, а именно метод Виолы-джонса, бинаризация, оператор Собеля, теория графов и метод 68 точек.

ГЛАВА 1 | МЕТОД ВИОЛЫ-ДЖОНСА

Интергральное представление изображения

Интергральное представление изображения - это матрица, совпадающая по размеру с исходным изображением, при этом каждый её элемент равен сумме яркостей пикселей, расположенных выше и левее данного. В данном случае под яркостью пикселя подразумевается присвоенное ему число в диапазоне от 0 до 255 включительно, характеризующее яркость цвета (для черно-белого изображения - черного, для RGB изображения - каждого из составляющих цветов).

$$A_{m \times n} = \left\| \sum_{i=0}^m \left(\sum_{j=0}^n a_j^i \right) \right\| \quad (1)$$

Рассмотрим на простом примере (Рис.1): пусть белому цвету пикселя соответствует значение "0 серому - "1" и наконец черному - "2". Тогда матрица интегрального представления будет выглядеть следующим образом:

Нетрудно заметить, что данное преобразование является линейным, а значит его сложность - $O(m \cdot n)$, где $m \times n$ - размер исходного изображения. Такое представление изображения позволяет легко вычислить за константное время яркость определенного участка изображения. Чтобы вычислить яркость прямоугольника с вершинами $(i_1; j_1)$ и $(i_2; j_2)$

^{*} aliev08@mail.ru

[†] vozzhayeva.darya@bk.ru

[‡] alinarabceva416@gmail.com

[§] 6aculuu@mail.ru

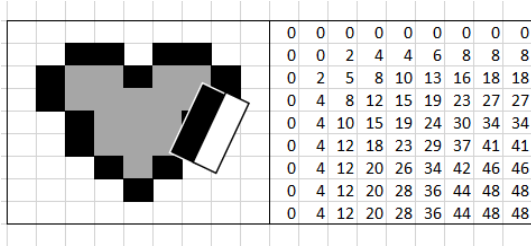


Рис. 1. Интегральное представление изображения

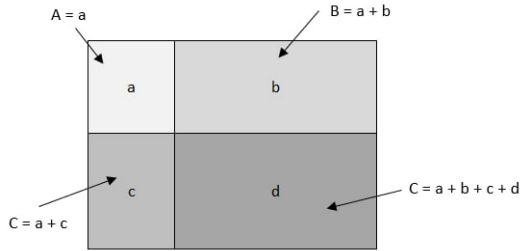


Рис. 2. Яркость прямоугольника в центре

(где $i_1 \leq i_2$ и $j_1 \leq j_2$), необходимо воспользоваться следующей формулой :

$$B_{(i_1:j_1);(i_2:j_2)} = a_{j_2}^{i_2} + a_{j_1-1}^{i_1-1} - a_{j_2}^{i_1-1} - a_{j_1-1}^{i_2} \quad (2)$$

Эта формула берет начало из вычисления площади прямоугольника (рис.2):

$$d = S_D + S_A - S_B - S_C$$

$$d = (a + b + c + d) + a - (a + b) - (a + c) \quad (3)$$

$$d = d - \text{истина}$$

Признаки Хаара

Признаки Хаара - это математические функции, используемые в компьютерном зрении для распознавания объектов на изображении. Признаки Хаара являются парными прямоугольниками с различными весами (коэффициентами), которые перемещаются по всему изображению. Эти прямоугольники могут быть различной формы и размера, а их положение может быть настроено под конкретные цели алгоритма обработки изображений (рис.3).

Используя признаки Хаара, система может извлечь информацию о текстуре, цвете, контрасте и других свойствах обрабатываемого

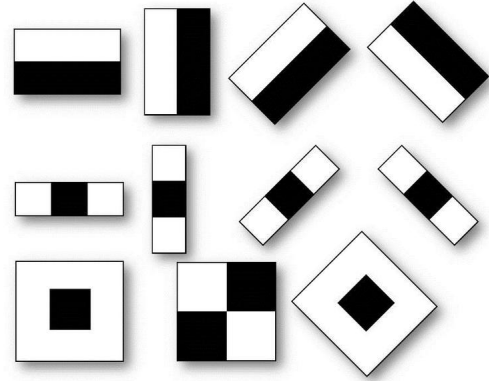


Рис. 3. Примеры признаков Хаара



Рис. 4. Пример применения признаков Хаара

изображения (рис.4). Одной из главных особенностей признаков Хаара является их простота и быстрота вычисления. Это обусловлено тем, что при вычислении признака Хаара для определенной области изображения, достаточно суммировать значения пикселей в ней, используя заранее рассчитанные интегральные изображения.

Это делает их подходящими для решения задач реального времени, например, для распознавания лиц в видеопотоке. Однако, признаки Хаара имеют свои недостатки. Например, они не могут обрабатывать изображения с высоким уровнем шума или с измененной освещенностью, что может привести к неправильному распознаванию объекта. Тем не менее, признаки Хаара являются мощным инструментом для обработки изображений, ко-

торый продолжает широко использоваться в различных областях компьютерного зрения.

Применение

Таким образом, на этих двух методах основана работа `opencv`. Это библиотека с открытым исходным кодом, которая предназначена для обработки и анализа изображений, компьютерного зрения, машинного обучения и более широкой области робототехники. `OpenCV` была создана в 1999 году и в настоящее время является одной из наиболее популярных библиотек для анализа и обработки изображений. Она может использоваться на различных платформах, таких как Windows, Linux, Android и iOS, а также поддерживает различные языки программирования, в том числе C++, Python и Java. Библиотека `OpenCV` имеет более 2500 оптимизированных алгоритмов, которые могут использоваться для распознавания объектов, определения и анализа форм, обнаружения движения, работы с видеопотоками, а также более широкой обработки изображений. Одним из применений этой библиотеки является программа, анализирующая язык жестов. Именно она эффективно и наиболее точно может отслеживать и распознавать людей, их лица и движения тела. Но у нее есть и недочет, при повороте на 30 градусов точность резко падает.

ГЛАВА 2 | МЕТОД 68 ТОЧЕК

Проблему ограниченного угла поворота лица решает наш следующий метод, а именно метод 68 точек. С помощью алгоритмов компьютерного зрения, происходит распознавание 68 ключевых точек на лице. Они показывают расположение различных отдельных элементов лица, таких как глаза, нос, рот и т.д. (рис.5). Каждая точка на лице имеет свой уникальный номер, и каждый номер связан с определенным анатомическим элементом лица. Например, точки №37 и №38 отвечают за внешние края глазных впадин, а точка №33 - за раздвоение носа. Этот метод имеет несколько преимуществ. Во-первых, при таком методе

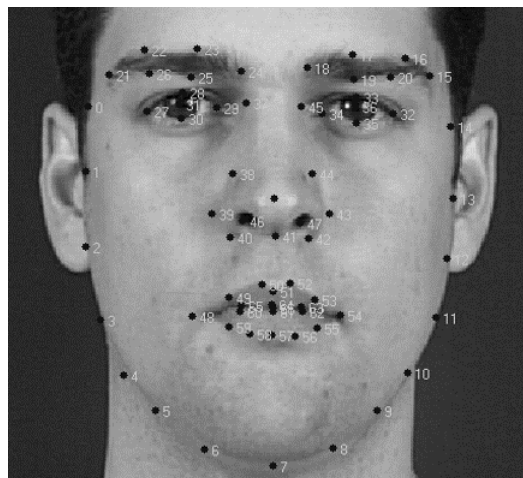


Рис. 5. Пример расположения 68 точек на лице человека

есть возможность распознавать лицо, повернутое от плоскости камеры более чем на 30 градусов. Главным же плюсом является практически отсутствие возможности обмануть подобный алгоритм. На основе микро смещений этих точек можно отделить например живого человека от его фотографии.

ГЛАВА 3 | ТЕОРИЯ ГРАФОВ

Теория графов широко используется в компьютерном зрении для решения различных задач. Одно из применений теории графов в компьютерном зрении - это задача сегментации изображений. Для этого изображение представляется в виде графа, где каждый узел соответствует пикселю, а связи между узлами - это связи между соседними пикселями. Затем на графе выполняется алгоритм кластеризации, который разбивает граф на несколько кластеров, каждый из которых соответствует одному сегменту на изображении. Рассмотрим этот метод подробнее на примере предложения "The new play is a great draw". Для начала выделяем грамматическую основу: "play подлежащее, "is draw". Далее мы имеем определения, "new относящееся к подлежащему и "great" к сказуемому. Остались артикли и предлоги, их распределим аналогично. Таким образом мы получили бинарное дерево связей слов в предложении, полностью

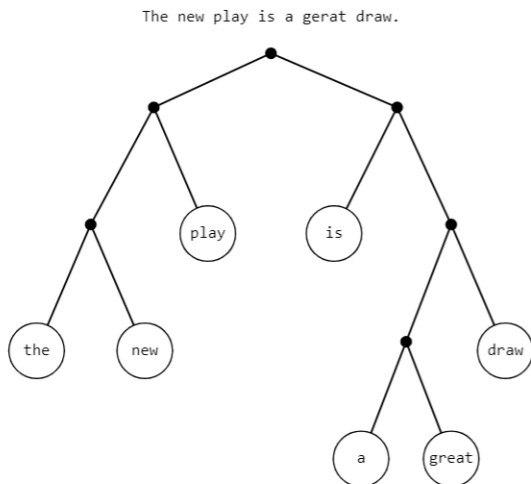


Рис. 6. Бинарное дерево предложения

характеризующее ее основу.

Мы рассмотрели построение двоичного дерева предложения, при этом вершинам являются его наименьшие составные единицы - слова. В случае составления аналогичного дерева изображения, вершинами будут являться пиксели, из которых состоит наша картинка. Теория графов может быть использована для решения задачи распознавания объектов на изображении. Для этого объекты на изображении выделяются в отдельные сегменты, которые затем представляются в виде графа. Затем на графе выполняется алгоритм поиска наибольшей клики, который позволяет определить, какие узлы графа соответствуют объекту, а какие - фону. Таким образом, применение теории графов в компьютерном зрении позволяет решать различные задачи более эффективно и точно, что делает этот подход все более популярным в современной науке и технологиях.

ГЛАВА 4 | БИНАРИЗАЦИЯ И ОПЕРАТОР СОБЕЛЯ

Бинаризация

Заключительными методами, которые рассмотрим, будут бинаризация и оператор Собеля. Бинаризация изображения - это про-



Рис. 7. Бинаризованное изображение

цесс преобразования изображения из многоцветного или градиентного изображения в изображение, содержащее только два возможных цвета - черный и белый (рис.6). Обычно этот процесс используется с целью упрощения изображения и выделения объектов на фоне. В процессе бинаризации изображения каждый пиксель присваивается значению белого или черного цвета в зависимости от того, соответствует ли его цветовое значение определенным пороговым значениям. Если значение пикселя превышает порог, то он присваивается белому цвету, а если значение пикселя меньше или равно пороговому значению, то он присваивается черному цвету.

Оператор Собеля

Далее применяется оператор Собеля — матричный оператор, который применяется для обнаружения границ на изображениях. Он использует две матрицы, которые называются ядрами, и проходится по всем пикселям изображения. Эти ядра присутствуют в вертикальном и горизонтальном направлениях. Для каждого пикселя входного изображения оператор Собеля вычисляет величину градиента, используя свертки с вертикальным и горизонтальным ядрами. Затем, результаты этих сверток объединяются для определения направления границы пикселя. Если граница находится в вертикальном направлении, то больший вклад вносит вертикальное ядро, если в горизонтальном, то больший вклад вносит горизонтальное ядро. Для получения финального изображения границ необходимо определить пороговое значение, которое будет

использоваться для фильтрации градиентов. Пиксели с градиентом большим, чем пороговое значение, будут отображаться как границы на выходном изображении.

Применение

Эта пара методов наиболее эффективна для распознавания контуров на изображении. Некоторые области, где применяют оператор Собеля, включают в себя:

Распознавание лиц: Оператор Собеля может быть использован для извлечения контуров лица, что помогает в процессе распознавания лиц.

Медицинское изображение: В области медицинского изображения оператор Собеля используется для обработки различных типов данных, таких как рентгеновские снимки, магнитно-резонансные изображения и томограммы.

Обнаружение границ: В компьютерном зрении оператор Собеля используется для выделения границ на изображениях.

Детектирование объектов: Оператор Собеля помогает выделить контуры объектов на изображении и обнаружить их на этапе детектирования объектов.

Автоматическое вождение: Оператор Собеля применяется в системах автоматического вождения для обнаружения машин и других объектов на дороге.

СПИСОК ЛИТЕРАТУРЫ

1. Н.В. Маниюкова Компьютерное зрение как способ извлечения информации из видеоряда // Математические структуры и моделирование. - 2015. - №4
2. A.V. Alyoshintsev, A.N. Sak Modeling of information systems of computer vision for machine translation by methods of the theory of graphs // Computer science. - 2018. - №12
3. Е.Е. Истратова, Д.Н. Достовалов, Е.А. Бухамер Разработка интеллектуальной системы для распознавания лиц на основе нейронных сетей // 2021. - №5
4. Д.С. Засышкин, Ю.С. Белов Алгоритмы обнаружения человека в opencv// Технические науки. - 2019
5. М.Е. Пажетнов, А.А. Каштанов, М.Г. Гриф Разработка методов распознавания жестов глухих на основе открытых пр // Современные научные исследования: теория, методология, практика . - 2022