Методы распознавания образов

#### **R-CNN**

Первые модели интуитивно начинают с поиска области, а затем выполняют классификацию. В R-CNN метод выборочного поиска, разработанный Дж.Р.Р. Уйлингс и др. (2012) является альтернативой полному поиску на изображении для фиксации местоположения объекта. Он инициализирует небольшие области изображения и объединяет их в иерархическую группу.

#### **Fast R-CNN**

Цель быстрой сверточной сети на основе регионов (Fast R-CNN), разработанной Р. Гиршиком (2015), состоит в том, чтобы сократить затраты времени, связанные с большим количеством моделей, необходимых для анализа всех предложений регионов.

#### **Faster R-CNN**

Предложения регионов, обнаруженные с помощью метода выборочного поиска, по-прежнему были необходимы в предыдущей модели, которая требовала значительных вычислительных ресурсов. С. Рен и др. (2016) представили Сеть предложений по регионам (RPN) для прямого создания предложений по регионам, прогнозирования ограничивающих рамок и обнаружения объектов. Более быстрая сверточная сеть на основе регионов (Faster R-CNN) представляет собой комбинацию между RPN и моделью Fast R-CNN.

#### **Region-based Fully Convolutional Network (R-FCN)**

Методологии Fast и Faster R-CNN заключаются в обнаружении предложений регионов и распознавании объекта в каждом регионе. Региональная полностью сверточная сеть (R-FCN)) представляет собой модель только со свёрточными слоями, обеспечивающую полное обратное распространение для обучения и логического вывода. Авторы объединили два основных шага в одну модель, чтобы одновременно учитывать обнаружение объекта (инвариант местоположения) и его положение (вариант местоположения).

#### **You Only Look Once (YOLO)**

Модель YOLO (J. Redmon et al., 2016)) напрямую предсказывает ограничивающие рамки и вероятности классов с помощью одной сети в одной оценке. Простота модели YOLO позволяет делать прогнозы в реальном времени.

#### **Single-Shot Detector (SSD)**

Подобно модели YOLO, W. Liu et al. (2016) разработали однократный детектор (SSD) для одновременного прогнозирования всех ограничительных рамок и вероятностей классов с помощью сквозной архитектуры CNN.

#### **Mask Region-based Convolutional Network (Mask R-CNN)**

Еще одно расширение модели Faster R-CNN добавленной параллельной ветви к обнаружению ограничивающей рамки, чтобы предсказать маску объекта. Маска объекта — это его сегментация по пикселям на изображении. Эта модель превосходит современную в четырех задачах COCO: сегментация экземпляра, обнаружение ограничивающей рамки, обнаружение объекта и обнаружение ключевой точки.

**Алгоритмы OpenCV**

Статья аспиранта на эту тему  
[https://asp.niisi.ru/%D0%91%D0%B5%D1%81%D1%88%D0%B0%D0%BF%D0%BE%D1%88%D0%BD%D0%8%D0%BA%D0%BE%D0%B2%20%D1%80%D0%B5%D1%84%D0%B5%D1%80%D0%B0%D1%82%20OpenCV.pdf](https://asp.niisi.ru/%D0%91%D0%B5%D1%81%D1%88%D0%B0%D0%BF%D0%BE%D1%88%D0%BD%D0%258%D0%BA%D0%BE%D0%B2%20%D1%80%D0%B5%D1%84%D0%B5%D1%80%D0%B0%D1%82%20OpenCV.pdf)

Краткое описание основных алгоритмов:  
  
**Детектор границ Кэнни**

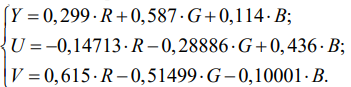
<https://cyberleninka.ru/article/n/matematicheskie-i-kompyuternye-metody-analiza-fotoizobrazheniy/viewer>

<https://cyberleninka.ru/article/n/matematicheskie-i-kompyuternye-metody-analiza-fotoizobrazheniy/viewer>

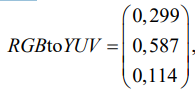
https://ru.wikipedia.org/wiki/%D0%9E%D0%BF%D0%B5%D1%80%D0%B0%D1%82%D0%BE%D1%80\_%D0%9A%D1%8D%D0%BD%D0%BD%D0%B8#:~:text=%D0%9E%D0%BF%D0%B5%D1%80%D0%B0%D1%82%D0%BE%D1%80%20%D0%9A%D1%8D%D0%BD%D0%BD%D0%B8%20(%D0%B4%D0%B5%D1%82%D0%B5%D0%BA%D1%82%D0%BE%D1%80%20%D0%B3%D1%80%D0%B0%D0%BD%D0%B8%D1%86%20%D0%9A%D1%8D%D0%BD%D0%BD%D0%B8,%D1%88%D0%B8%D1%80%D0%BE%D0%BA%D0%BE%D0%B3%D0%BE%20%D1%81%D0%BF%D0%B5%D0%BA%D1%82%D1%80%D0%B0%20%D0%B3%D1%80%D0%B0%D0%BD%D0%B8%D1%86%20%D0%B2%20%D0%B8%D0%B7%D0%BE%D0%B1%D1%80%D0%B0%D0%B6%D0%B5%D0%BD%D0%B8%D1%8F%D1%85.

Оператор Кэнни (алгоритм Кэнни) – оператор обнаружения границ изображения. Данный алгоритм удовлетворяет трем критериям: 1) улучшенное обнаружение; 2) точная локализация; 3) лишь один отклик на одну границу (рис. 1)[1].

Перед применением метода для уменьшения вычислительных затрат исходное изображение подвергается предварительной обработке. Изначально рисунок, представленный в цветовой модели RGB, преобразуют в модель YUV (или HSL, HSV и т. д.), в которой цвет представляется как три компоненты – яркость (Y) и две цветоразностных (U и V). Для этого каждый из составляющих цвета умножается на коэффициенты перевода, являющиеся постоянными в связи с особенностями человеческого восприятия.



В связи с тем, что цветовая составляющая не используется в обработке, можно свернуть данный перевод до векторной функции:



где R, G и B – это компоненты (каналы) цветовой модели RGB. Для дальнейшей обработки важна только Y-составляющая, отвечающая за яркость [2].

Для того чтобы обрабатывать различные изображения методом Канни с ожидаемым одинаковым результатом без ручного задания параметров алгоритма отдельно для каждого изображения, необходимо иметь одинаковую яркость и насыщенность изображений. В этих целях применяется алгоритм нормализации гистограммы изображения.

Гистограмма яркости – это количественное распределение пикселей по значениям яркости. Другими словами, это представление распределения полутоновой яркости изображения в виде графика, где по горизонтальной оси представлена яркость в градациях серого (0…255), а по вертикальной оси – относительное число пикселей для конкретной величины яркости [3].

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

Изображение, переведенное в оттенки серого, сглаживается за счет использования специального фильтра. Этот фильтр обнаруживает и устраняет найденные разрывы, применяя перемещаемую по изображению маску. Обычно ее называют ядром, которое представляет собой квадратную матрицу. Элементами этой матрицы являются пиксели изображения, над которыми в данный момент находится маска. Согласно значениям яркости этих пикселей изменяется яркость пикселя под центром маски.

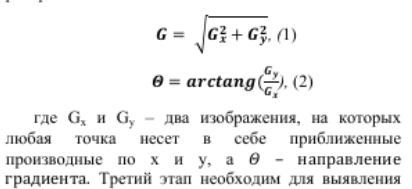
Оператор Канни использует фильтр Гаусса. Размытие по Гауссу – это характерный фильтр размытия изображения, который использует нормальное распределение (также называемое гауссовым распределением) для вычисления преобразования, применяемого к каждому пикселю изображения.

Уравнение распределения Гаусса в N измерениях имеет вид:

где 2 2 ( ) x y + – это радиус размытия; σ – стандартное отклонение распределения Гаусса. В случае двух измерений эта формула задает поверхность, имеющую вид концентрических окружностей, с распределением Гаусса от центральной точки.

Пиксели, где распределение отлично от нуля, используются для построения матрицы свертки, которая применяется к исходному изображению. Значение каждого пикселя становится средневзвешенным для окрестности. Исходное значение пикселя принимает наибольший вес (имеет наивысшее Гауссово значение), и соседние пиксели принимают меньшие веса, в зависимости от расстояния до них.

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



В применяемом методе используются специализированные ядра фильтра, являющиеся разновидностью оператора Собеля. Именно такой оператор использует маски, с помощью которых сворачивает исходное изображение и вычисляет приближенное значение производной по вертикали, горизонтали или диагоналям [4].

Градиент определяет, как быстро изменяется яркость изображения в каждой точке, что дает возможность определить границу и ее ориентацию. При применении оператора Собеля в области постоянной яркости, т. е. области, где нет резких перепадов или границ, будет получен вектор очень маленький или близкий к нулю. В точке, относящейся к границе областей, имеющих различную яркость, величина вектора будет значительно больше, с направлением в сторону увеличения яркости.

После вычисления оператора Собеля определяется угол направления вектора границы. Направления вектора округляется до ближайшего угла, кратного 45° (0, 45, 90 и 135°) ° (рис. 3). Затем проверяется достижение локального максимума величиной градиента в найденном направлении вектора [2].

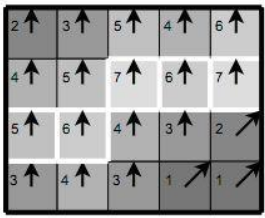
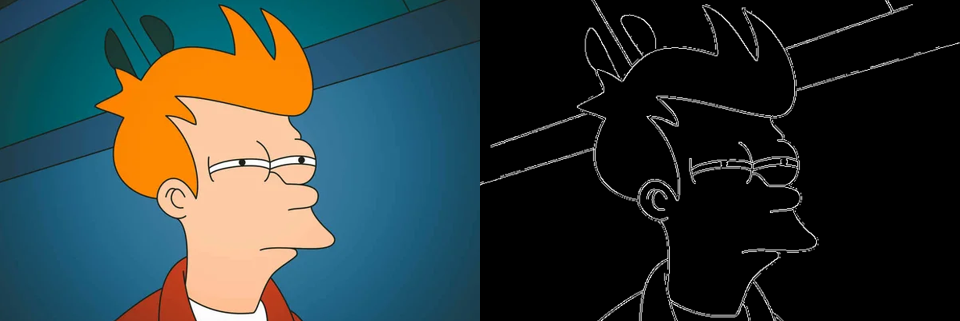


Рис. 3. Принцип подавления «ложных» максимумов

Метод использования порогов заключается в определении областей на изображении, где могут располагаться границы, за счет пороговых значений. Чем сильнее пороги, тем меньше границ будет обнаружено, но и шум на изображении будет восприниматься значительно слабее.

Алгоритм Канни 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 [2].

Все пиксели, прошедшие предыдущие этапы, проверяются на близкое расположение друг к другу. Пиксели, лежащие обособленно или отдаляющиеся от границы, подавляются





1. Вежневец А., Баринова О. Методы сегментации изображений: автоматическая сегментация / Графика и мультимедия: офиц. сайт. – URL: http://cgm.computergraphics.ru/content/view/147 (дата обращения: 25.04.2013); 2. John Canny. A computational approach to edge detection // IEEE Transactions on pattern analysis and machine intelligence. – 1986. – V. 8, № 6. – P. 679–698. 3. Otsu, N. A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on Systems, Man, and Cybernetics. – 1979. – V. 9, № 1. – P. 62–66. 4. Гонсалес Р.С., Вудс Р.Э. Цифровая обработка изображений: пер. с англ. – М.: Техносфера, 2005. – 1070 с.

Хотя его работа была проведена на заре компьютерного зрения (1986 г.), детектор границ Канни до сих пор является одним из лучших детекторов. Кроме особенных частных случаев трудно найти детектор, который бы работал существенно лучше, чем детектор Канни.

**Оператор Собеля**

Оператор Собеля – дискретный оператор, вычисляющий приближенное значение градиента яркости изображения. Результатом применения оператора Собеля в каждой точке изображения является либо вектор градиента яркости в этой точке, либо его норма. (Много любимых слов Аниси)

https://www.wikiwand.com/ru/%D0%9E%D0%BF%D0%B5%D1%80%D0%B0%D1%82%D0%BE%D1%80\_%D0%A1%D0%BE%D0%B1%D0%B5%D0%BB%D1%8F  
  
**Преобразование Хафа**

Доклад из бгуира  
<https://doklady.bsuir.by/jour/article/download/816/816>

Преобразование Хафа – численный метод, применяемый для поиска объектов, принадлежащих определенному классу фигур, с использованием процедуры голосования.

Преобразование Хафа позволяет находить на монохромном изображении плоские кривые, заданные параметрически, например: прямые, окружности, эллипсы, и т.д. Монохромным изображением называется изображение, состоящее из точек двух типов: фоновых и точек интереса. Задача преобразования состоит в выделении кривых, образованных точкками интереса.

<https://cyberleninka.ru/article/n/matematicheskie-i-kompyuternye-metody-analiza-fotoizobrazheniy/viewer>

Назначение преобразования Хафа состоит в том, чтобы выделить на изображении группы точек, образующие те или иные геометрические фигуры (прямые, окружности, эллипсы и др.), и найти параметры данных объектов.

Классический алгоритм преобразования Хафа имеет дело с идентификацией прямых в изображении, но позже алгоритм был расширен возможностью идентификации позиций кругов, эллипсов, а также произвольных фигур [2, 3].

Алгоритм Хафа распознавания прямых

Прямая на плоскости описывается уравнением y=kx+b и может быть задана парой несовпадающих точек. Однако во втором случае существует проблема описания вертикальных прямых, для которых значения параметров k и b бесконечны. Поэтому в данной ситуации удобнее представить прямую с помощью двух других параметров ρ и θ. Параметр ρ − это длина перпендикуляра, опущенного на прямую из начала координат, а θ − это угол между данным перпендикуляром и осью x (см. рис. 1).

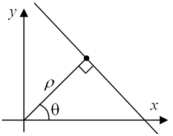


Рис. 1. Задание прямой на плоскости параметрами ρ и θ

Через одну точку декартовой плоскости можно провести бесконечное число прямых (см. рис. 2,а). Если эта точка имеет координаты (x0, y0) на изображении, то все прямые, проходящие через неё, соответствуют следующему уравнению:

ρ(θ) = x0⋅cosθ+y0⋅sinθ.

Это соответствует синусоидальной кривой в пространстве (ρ, θ) (см. рис. 2,б)

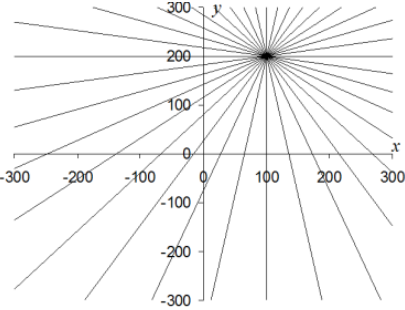


Рис. 2,а. Соответствующая прямым синусоида

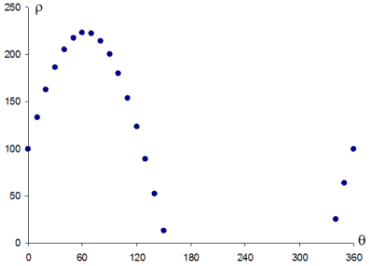


Рис. 2,б. Соответствующая прямым синусоида

Если синусоиды, соответствующие двум точкам декартовой плоскости, наложить друг на друга, то точка (в пространстве Хафа), где они пересекутся, будет соответствовать параметрам прямой, проходящей через обе эти точки. Таким образом, ряд точек, которые формируют прямую линию (см. рис. 2,в), определяют синусоиды, которые пересекаются в точке параметров (ρ0, θ0) для этой линии (см. рис. 2,г). Поэтому, проблема обнаружения коллинеарных точек может быть сведена к проблеме обнаружения пересекающихся кривых.

Каждой точке (ρ0, θ0) пространства (ρ, θ) можно поставить в соответствие счетчик, соответствующий количеству точек (x, y), лежащих на прямой x⋅cosθ0+y⋅sinθ0=ρ0.

Вследствие дискретности растрового представления исходного изображения, каждой прямой в координатах (x, y) будет соответствовать не одна точка, а сгущение точек в координатах (ρ, θ). Таким образом, достаточно выбрать на изображении, построенном в полярных координатах, самые «жирные пятна» и для их центров произвести преобразование в полярные координаты, получив тем самым параметры соответствующей прямой. Схема алгоритма Хафа для нахождения прямых линий на изображении приведена на рис. 3.

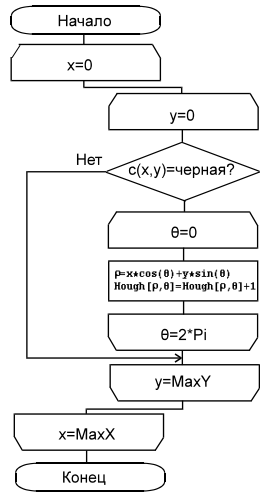
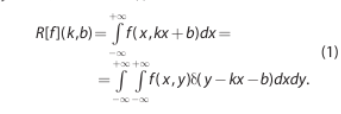


Рис. 3. Схема алгоритма Хафа распознавания прямых

**Преобразование Радона**

Преобразование Радона – это хорошо известное интегральное преобразование, сопоставляющее фукнции F в исходном пространстве функцию на множестве плоскостей, задаваемую интегралам от f вдоль этих плоскостей.

Определение преобразования Радона: если в пространстве двух переменных задана функция f(x, y), то интегральное преобразование Радона вдоль прямой линии y=k\*x+b имеет вид:



Здесь b(x) – дельта функция Дирака, а параметры линии b и k образуют пространство данного интегрального преобразования.

Необходимо отметить, что преобразование Радона обладает некоторыми свойствами, очень важными для работы с изображениями, такими как свойство линейности, сдвига и масштабирования. Однако, интегральное преобразование имеет ограничение, связанное с поиском прямых на изображении, угол наклона которых равен или близок 90 градусов к оси абсцисс. Данное обстоятельство связано с бесконечно большим параметром k . Для того, чтобы избавиться от данной проблемы, можно перейти к заданию линии в нормальной форме:



**Метод Виолы-Джонса**

Метод Виолы-Джонса – алгоритм, позволяющий обнаруживать объекты в реальном времени.

