

Лекция №17 Компьютерное зрение

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

Основные понятия

- **Компьютерное зрение** — теория и технология создания машин, которые могут видеть.
- Как научная дисциплина, компьютерное зрение относится к теории и технологии создания искусственных систем, которые получают информацию из изображений. Видеоданные могут быть представлены множеством форм, таких как видеопоследовательность, изображения с различных камер или трехмерными данными с медицинского сканера.

Примеры применения систем компьютерного зрения

- Системы управления процессами (промышленные роботы, автономные транспортные средства)
- Системы видеонаблюдения
- Системы организации информации (например, для индексации баз данных изображений)
- Системы моделирования объектов или окружающей среды (анализ медицинских изображений, топографическое моделирование)
- Системы взаимодействия (например, устройства ввода для системы человеко-машинного взаимодействия)

Задача компьютерного зрения

- Понять, что запечатлено на изображении



La Gare Montparnasse, 1895

Мы видим

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

Компьютер видит

История: первая презентация



Самая первая фотография
1825 год



Figure 5. J. N. Niepce.

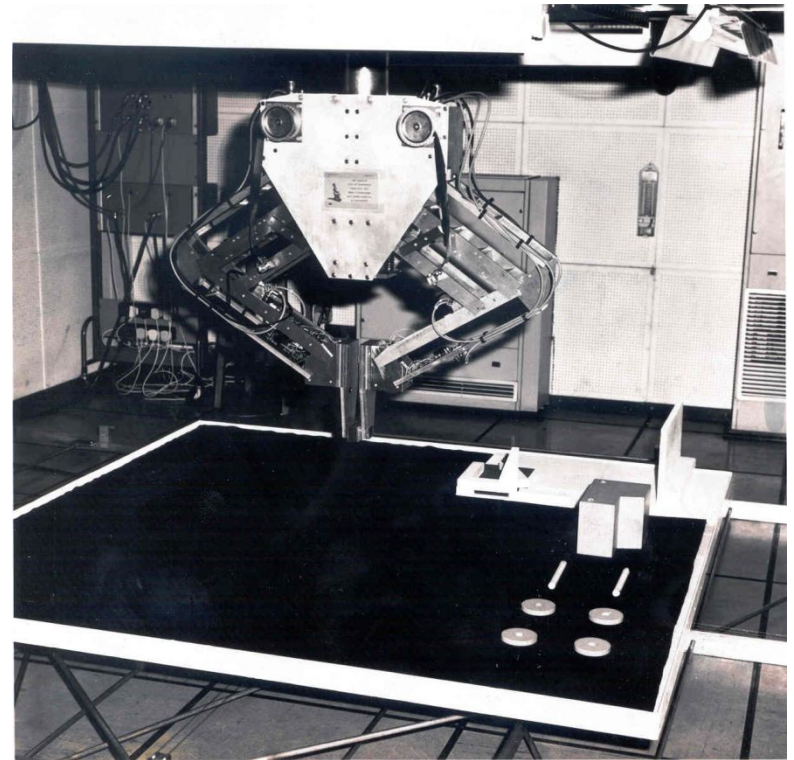
Whirlwind, MIT, 1951



- Первый компьютер, отображающий текст и графику в реальном времени на мониторе
- Точками карту, значком самолёт.
- «Световое перо» для взаимодействия с экраном (запрос информации об объекте)

Freddy II, 1973

- Университет Эдинбурга
- Один из первых роботов с системой машинного зрения
- 5 степеней свободы
- Умеет собирать машинки из кубиков, разбросанных по столу
- 384Кб RAM в управляющем компьютере



Решаемые задачи

- Изображения и видео повсюду
- Бурно растущая область
 - Обработка – улучшение качества, ретушь, изменение размера и формы, композиция
 - Интернет – поиск, аннотация, поиск дубликатов, распознавание объектов
 - Видеонаблюдение – отслеживание, распознавание объектов, распознавание жестов и событий
 - Промышленные системы – диагностика, контроль качества
 - Спецэффекты в кино – композиция, монтаж фонов, захват движения

Типичные функции систем компьютерного зрения

- Получение изображения
- Предварительная обработка
- Выделение деталей
- Детектирование/Сегментация
- Высокоуровневая обработка

Сглаживание изображений

- Всем устройствам получения изображений присуще наличие шумов в получаемых изображениях. Шумы могут порождаться как самим устройством (шум матрицы, шум обрабатывающих устройств и т.д.), так и носить объективный характер.
- Поэтому один из важных этапов в обработке изображения является устранение шумов или сглаживание изображения.

Сглаживание изображений

- Один из способов сглаживания изображения состоит в том, чтобы каждому пикселу присваивалось среднее значение характеристик его соседних пикселей. Такой способ обработки, как правило, исключает экстремальные значения.
- Основная проблема данного способа: *сколько нужно рассмотреть соседних пикселей?*

Фильтр с гауссовой характеристикой

$$G_{\sigma}(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{- (x^2 + y^2) / 2\sigma^2}$$

Под применением фильтра с гауссовой характеристикой подразумевается замена значения интенсивности $I(x_0/y_0)$ суммой по всем (x,y) пикселям значений $I(x,y) G_{\sigma}(d)$, где d — расстояние от (x_0,y_0) до (x,y) .

Такого рода взвешенная сумма применяется так часто, что для нее предусмотрено особое название – «свертка».

Значение σ , равное 1 пикселу, является достаточным для сглаживания шума с небольшой интенсивностью. Если же значение σ соответствует 2 пикселям, то происходит сглаживание шума с большей интенсивностью, но теряются некоторые мелкие детали.

Применение сглаживания для обнаружения роботов объектов сцены



Обнаружение регионов на изображении без и с использованием фильтра Гаусса

Обнаружение краев

Края(границы) — это такие кривые на изображении, вдоль которых происходит резкое изменение яркости или других видов неоднородностей.



Обнаружение краев

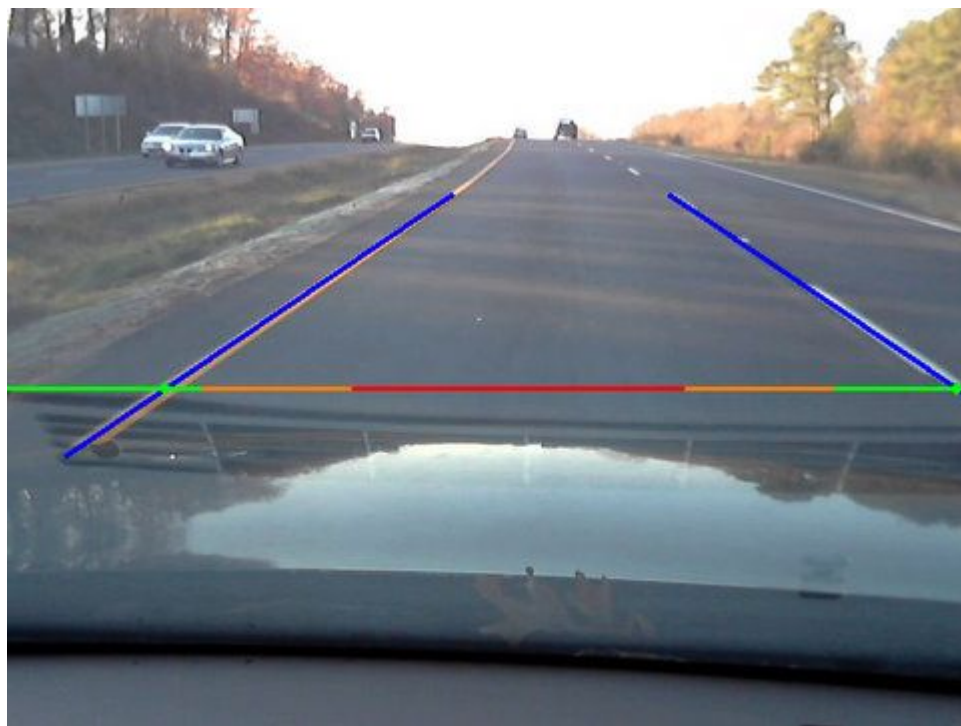
- Целью обнаружения краев является повышение уровня абстракции и переход от перегруженного подробностями мультимегабайтового изображения к более компактному, абстрактному представлению. Необходимость в выполнении такой операции обусловлена тем, что линии краев на изображении соответствуют важным контурам в сцене.

Детектор границ Кенни

- Хотя работа Кенни была проведена на заре компьютерного зрения (1986), детектор границ Кенни до сих пор является одним из лучших детекторов.
- Основные этапы детектора:
 - Убрать шум и лишние детали из изображения
 - Рассчитать градиент изображения
 - Сделать края тонкими (edge thinning)
 - Связать края в контура (edge linking)

Применение метода детектора границ Кенни

- реализация отслеживания разделительных полос на основе библиотеки OpenCV и использование детектора границ Кенни на примере робота-автомобиля от BMW



Вычисление градиента. Оператор Собеля.

- **Производная** (функции в точке) — это скорость изменения функции (в данной точке).
- Работая с изображением I , мы работаем с функцией двух переменных $I(x, y)$, т.е. со скалярным полем. Поэтому, более правильно говорить не о производной, а о градиенте изображения.
- **Градиент** (от лат. *gradiens* — шагающий, растущий) — вектор, показывающий направление наискорейшего возрастания некоторой величины, значение которой меняется от одной точки пространства к другой (скалярного поля).

Вычисление градиента. Оператор Собеля.

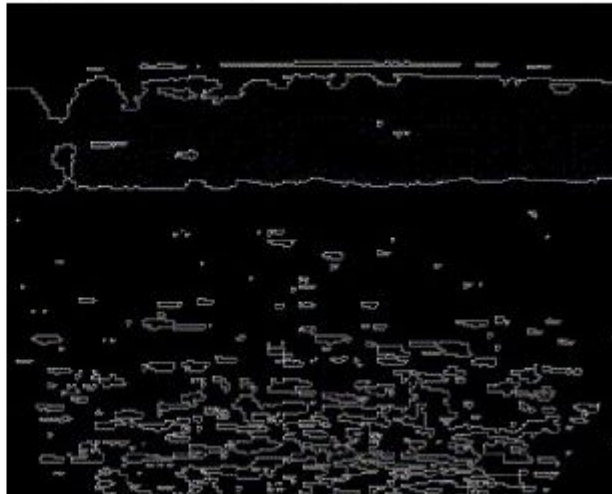
- Градиент для каждой точки изображения (функция яркости) — двумерный вектор, компонентами которого являются производные яркости изображения по горизонтали и вертикали.

$$\text{grad } I(x,y) = (dI/dx, dI/dy);$$

(a) Original image



(b) Sobel image



Вычисление градиента. Оператор Собеля.

- **Оператор Собеля** — это дискретный дифференциальный оператор, вычисляющий приближение градиента яркости изображения. Оператор вычисляет градиент яркости изображения в каждой точке. Так находится направление наибольшего увеличения яркости и величина её изменения в этом направлении. Результат показывает, насколько «резко» или «плавно» меняется яркость изображения в каждой точке, а значит, вероятность нахождения точки на грани, а также ориентацию границы.

Детектор границ Кенни

- Границы на изображении могут находиться в различных направлениях, поэтому алгоритм Кенни использует четыре фильтра для выявления горизонтальных, вертикальных и диагональных границ. Воспользовавшись оператором обнаружения границ (например, оператором Собеля) получается значение для первой производной в горизонтальном направлении (G_y) и вертикальном направлении (G_x). Из этого градиента можно получить угол направления границы:
$$\alpha = \arctan(G_x/G_y)$$

Детектор границ Кенни

- Угол направления границы округляется до одной из четырех углов, представляющих вертикаль, горизонталь и две диагонали (например, 0, 45, 90 и 135 градусов). Затем идет проверка того, достигает ли величина градиента локального максимума в соответствующем направлении.

Детектор границ Кенни

- Например, для сетки 3x3:
 - если угол направления градиента равен нулю, точка будет считаться границей, если её интенсивность больше чем у точки выше и ниже рассматриваемой точки,
 - если угол направления градиента равен 90 градусам, точка будет считаться границей, если её интенсивность больше чем у точки слева и справа рассматриваемой точки,
 - если угол направления градиента равен 135 градусам, точка будет считаться границей, если её интенсивность больше чем у точек находящихся в верхнем левом и нижнем правом углу от рассматриваемой точки
 - если угол направления градиента равен 45 градусам, точка будет считаться границей, если её интенсивность больше чем у точек находящихся в верхнем правом и нижнем левом углу от рассматриваемой точки.

Таким образом, получается двоичное изображение, содержащее границы (т.н. «тонкие края»).

Сегментация изображения

- Мозг человека не использует полученные им результаты восприятия в непосредственном виде, а организует эти результаты определенным образом, поэтому вместо коллекции значений яркости, связанных с отдельными фоторецепторами, мозг выделяет целый ряд визуальных групп, которые обычно ассоциируются с объектами или частями объектов.
- **Сегментация** — это процесс разбиения изображения на группы с учетом подобия характеристик пикселей.

Сегментация изображения

- Основная идея сегментации состоит в следующем: каждый пиксел изображения может быть связан с некоторыми визуальными свойствами, такими как яркость, цвет и текстура. В пределах одного объекта или одной части объекта эти атрибуты изменяются относительно мало, тогда как при переходе через границу от одного объекта к другому обычно происходит существенное изменение одного или другого из этих атрибутов. Необходимо найти вариант разбиения изображения на такие множества пикселов, что указанные ограничения удовлетворяются в максимально возможной степени.

Сегментация изображения

- Процесс сегментации, основанный исключительно на использовании низкоуровневых локальных атрибутов, таких как яркость и цвет, чреват существенными ошибками. Чтобы надежно обнаруживать границы, связанные с объектами, необходимо также использовать высокоуровневые знания о том, какого рода объекты могут повсей вероятности встретиться в данной сцене.

Сегментация изображения

- Методы сегментации можно разделить на два класса:
 - автоматические – не требующие взаимодействия с пользователем
 - интерактивные – использующие пользовательский ввод непосредственно в процессе работы.
- Задачи автоматической сегментации делятся на два класса:
 - выделение областей изображения с известными свойствами
 - разбиение изображения на однородные области

Пример сегментации



Оценка качества работы методов сегментации

- однородность регионов (однородность цвета или текстуры)
 - непохожесть соседних регионов
 - гладкость границы региона
 - маленькое количество мелких «дырок» внутри региона
 - и т.д.
- Более общий подход к оценке качества работы метода, не учитывающий конкретного приложения, состоит в тестировании методов на общей базе изображений, для которых известна «правильная» сегментация. Например, Berkeley Segmentation Dataset, насчитывает более 1000 изображений, отсегментированных вручную 30 разными людьми.

Кластеризация цветового пространства

- В постановке задачи сегментации прослеживается аналогия с задачей кластеризации. Для того чтобы свести задачу сегментации к задаче кластеризации, достаточно задать отображение точек изображения в некоторое пространство признаков и ввести метрику (меру близости) на этом пространстве признаков.

Кластеризация цветового пространства

- В качестве признаков точки изображения можно использовать представление ее цвета в некотором цветовом пространстве, примером метрики (меры близости) может быть евклидово расстояние между векторами в пространстве признаков. Тогда результатом кластеризации будет квантование цвета для изображения. Задав отображение в пространство признаков, можно воспользоваться любыми методами кластерного анализа.

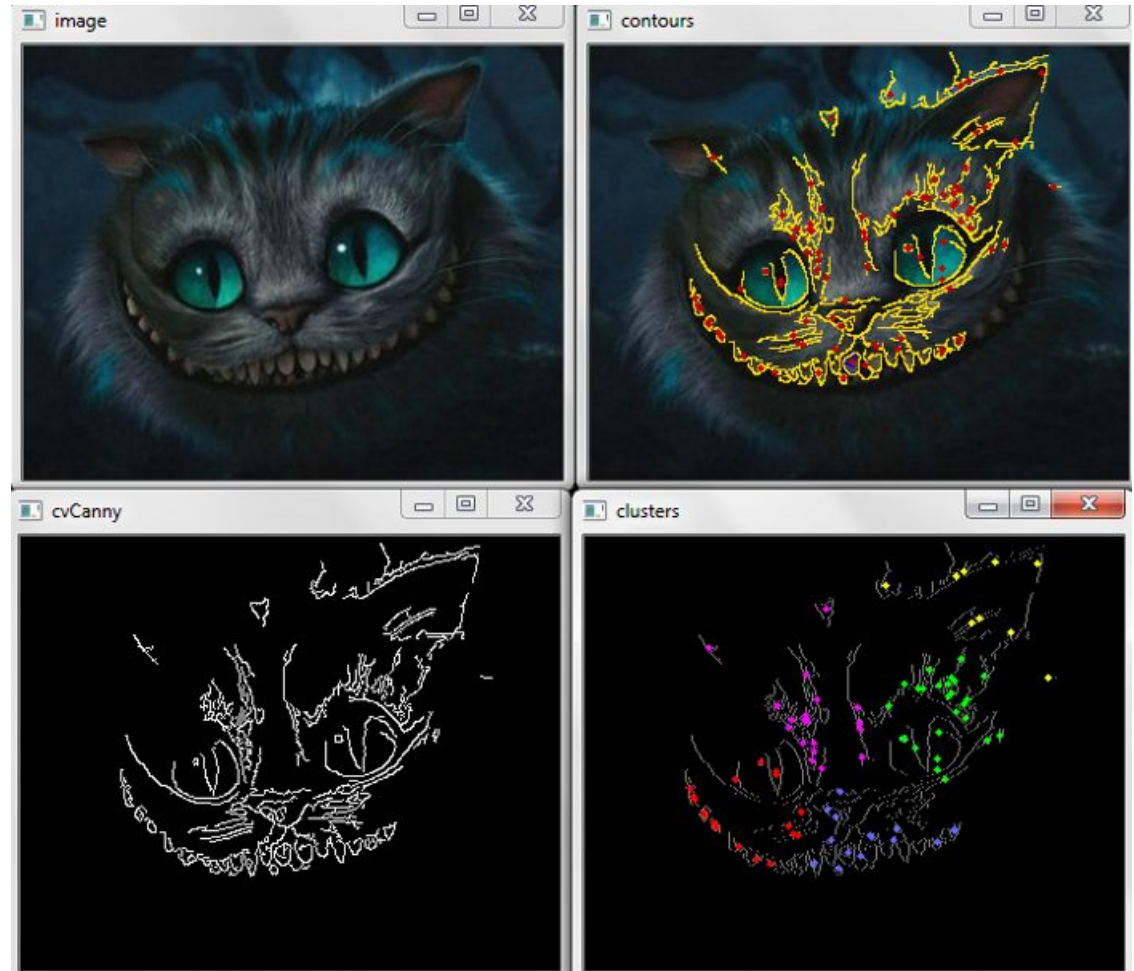
Кластеризация цветового пространства

- Основная проблема методов кластеризации, состоит в том, что пространственное расположение точек либо не учитывается совсем, либо учитывается косвенно (например, используя координаты точки как один из признаков).
- Методы кластеризации плохо работают на зашумленных изображениях: часто теряют отдельные точки регионов, образуется много мелких регионов, и. т. п.

Кластеризация цветового пространства

Кластеризация позволяет:

- лучше понять данные (выявив структурные группы)
- компактное хранение данных
- выявление новых объектов.



Метод выращивания регионов

- Методы этой группы учитывают пространственное расположение точек напрямую.
- Методы выращивания регионов основаны на следующей идее. Сначала по некоторому правилу выбираются центры регионов (seeds), к которым поэтапно присоединяются соседние точки, удовлетворяющих некоторому критерию. Процесс выращивания регионов (region growing) останавливается, когда ни одна точка изображения не может быть присоединена ни к одному региону.

Метод дробления-слияния

- Методы дробления-слияния состоят из двух основных этапов: дробления и слияния. Дробление начинается с некоторого разбиения изображения, не обязательно на однородные области. Процесс дробления областей происходит до тех пор, пока не будет получено разбиение изображения (пересегментация), удовлетворяющее свойству однородности сегментов. Затем происходит объединение схожих соседних сегментов до тех пор, пока не будет получено разбиение изображения на однородные области максимального размера.

Пирамиды изображений

- Пирамиды изображений - это коллекция изображений, получаемая из исходного изображения путём его последовательного сжимания пока не достигнута точка останова (естественно конечной точкой может быть один пиксель).

Пирамиды изображений

- Существует два вида пирамид - пирамиды Гаусса и Лапласа. Пирамиды Гаусса используются для сжатия изображения, а пирамиды Лапласа наоборот для восстановления изображения с повышенной дискретизацией из слоя в пирамиде.

Гауссовская пирамида

- Для получения слоя $(i+1)$ в Гауссовской пирамиде (обозначим этот слой G_{i+1}) из слоя G_i пирамиды, нам необходимо сначала свернуть этот слой с помощью Гауссовского ядра, а затем удалить все чётные строки и столбцы. Из этого следует что каждое последующее изображение будет занимать четверть площади его предшественника.

Пирамида Лапласа

- Изображение сначала преобразуется в два раза большее в обоих измерениях, с новыми строками заполненными нулями. А затем выполняется свёртка с заданным фильтром для аппроксимации значений отсутствующих пикселей. Операция получения пирамиды Лапласа **не является** обратной операцией получения гауссовской пирамиды. Т.к. в процессе получения каждого нового слоя в пирамиде Гаусса происходит потеря информации.

Пирамида́льная сегментации

- Операции построения пирамид Гаусса и Лапласа используются для выполнения пирамида́льной сегментации. Для этого на каждом новом слое похожие пиксели с предыдущего слоя объединяются в один, далее из нового слоя создаётся ещё один и т.д. до конца.

Сегментация преобразованием водораздела

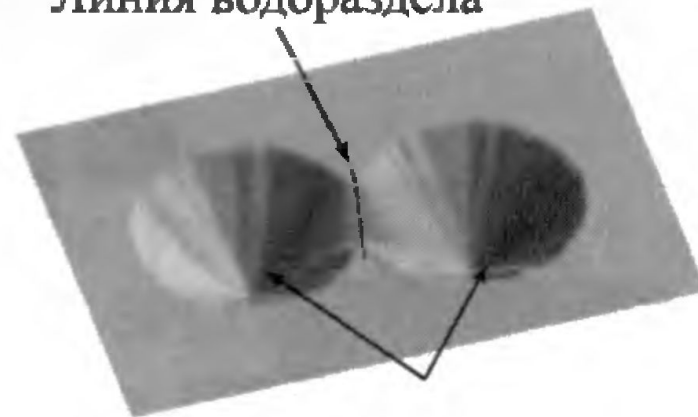
- В географии термин водораздел обозначает условную линию, которая разделяет области водосборов разных речных систем. Водосборный бассейн определяет географическую область, вода с которой собирается в одну реку или водохранилище.
- Сегментация преобразованием водораздела применяет эту идею к обработке монохромных изображений при решении задачи сегментации.

Сегментация преобразованием водораздела

a)



б)



Линия водораздела

Водораздельные бассейны

Виды сегментации по водоразделу

- Сегментация по водоразделам с помощью преобразования расстояния
- Сегментация по водоразделам с помощью градиентов
- С использованием маркеров при сегментации по водоразделам

Поиск объектов на изображении

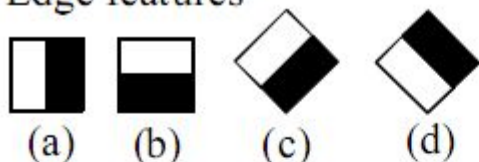
- Детектор Viola-Jones
 - основополагающий метод для поиска объектов на изображении в **реальном времени**
 - обучение очень медленное, но поиск очень быстрый

Детектор Viola-Jones

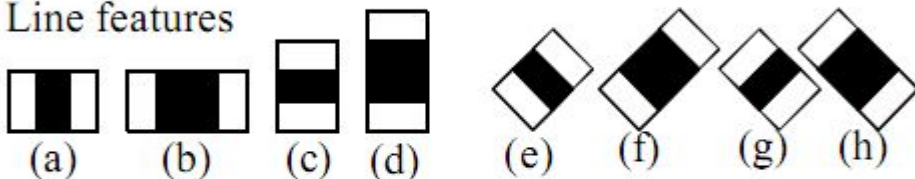
- Основные идеи:
 - Скользящее окно
 - Интегральные изображения для быстрого вычисления признаков
 - *Бустинг* для выбора признаков
 - *Каскад* для быстрой отбраковки окон без лица

Признаки Хаара

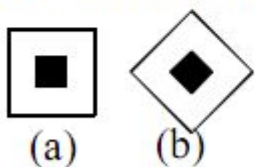
1. Edge features



2. Line features



3. Center-surround features



4. Special diagonal line feature used in [3,4,5]

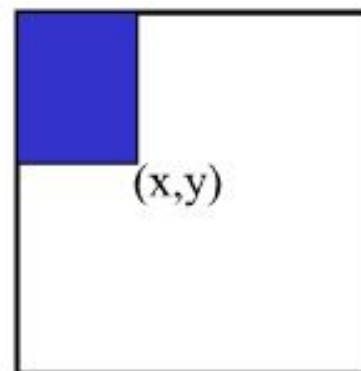


Value =

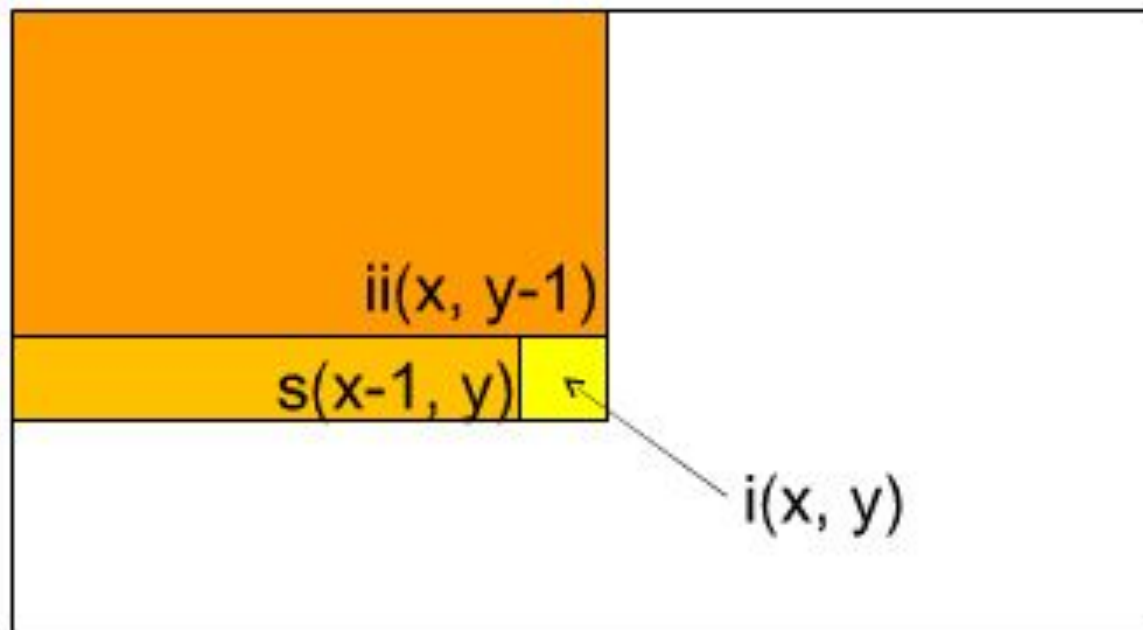
$$\sum (\text{pixels in white area}) - \sum (\text{pixels in black area})$$

Интегральные изображения

- Значение каждого пикселя (x, y) равно сумме значений всех пикселей левее и выше пикселя (x, y) включительно
- Интегральное изображение рассчитывается за один проход



Вычисление интегрального изображения



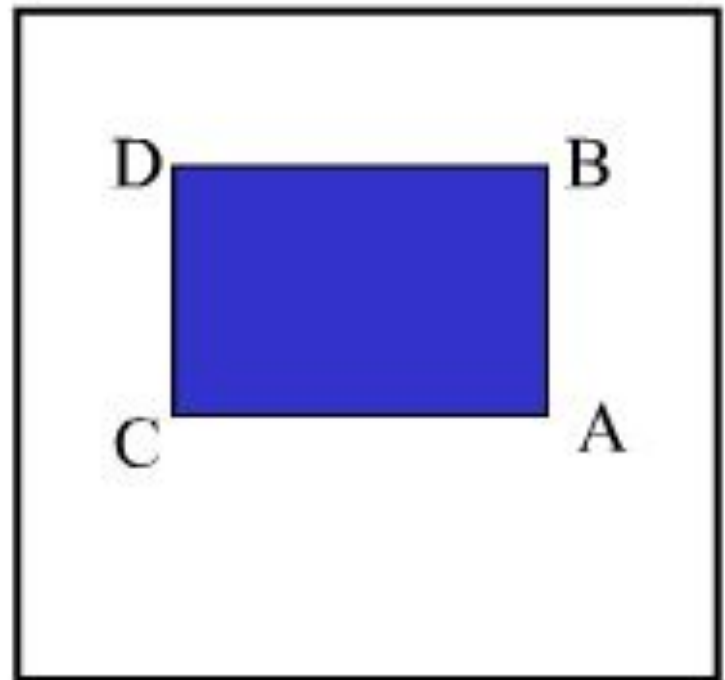
- Сумма по строке: $s(x, y) = s(x-1, y) + i(x, y)$
- Интегральное изображение: $ii(x, y) = ii(x, y-1) + s(x, y)$

Вычисление суммы в прямоугольнике

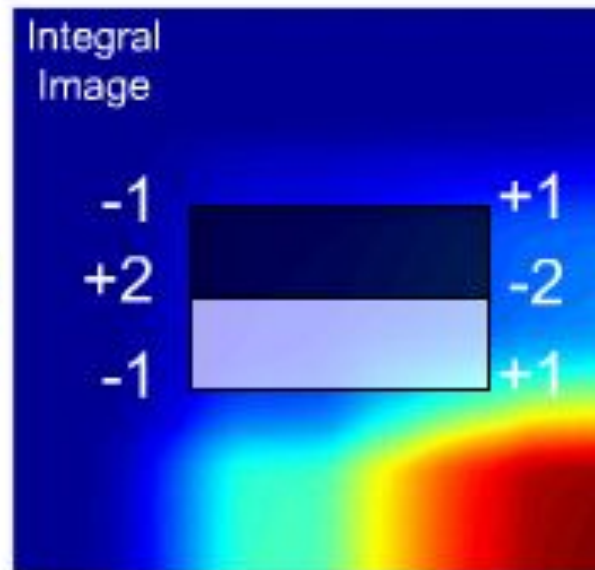
- Пусть A, B, C, D – значения интегрального изображения в углах прямоугольника
- Тогда сумма значений пикселей в исходном изображении вычисляется по формуле:

$$\text{sum} = A - B - C + D$$

- 3 операции сложения для любого прямоугольника

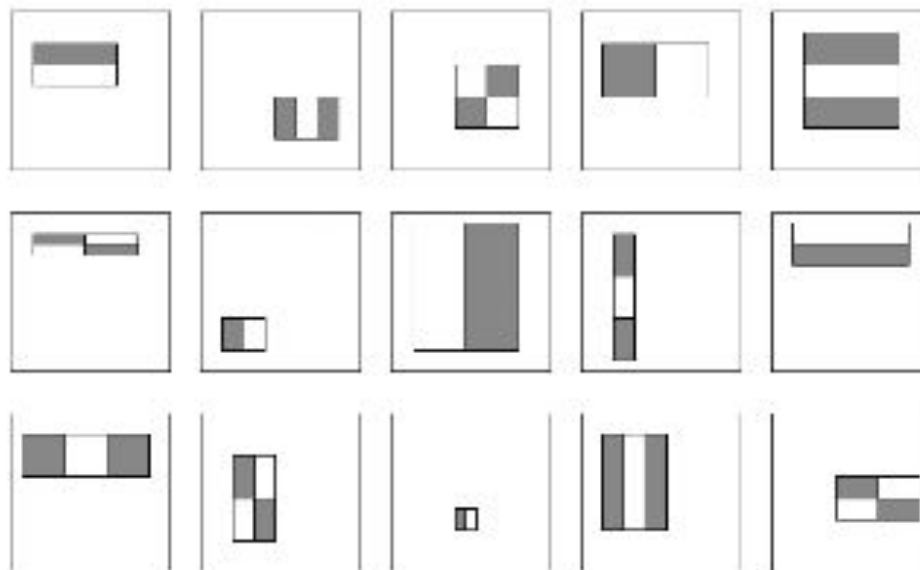


Пример интегрального изображения



Выбор признаков

- для окна поиска 24×24 число возможных прямоугольных признаков может достигать ~ 160000

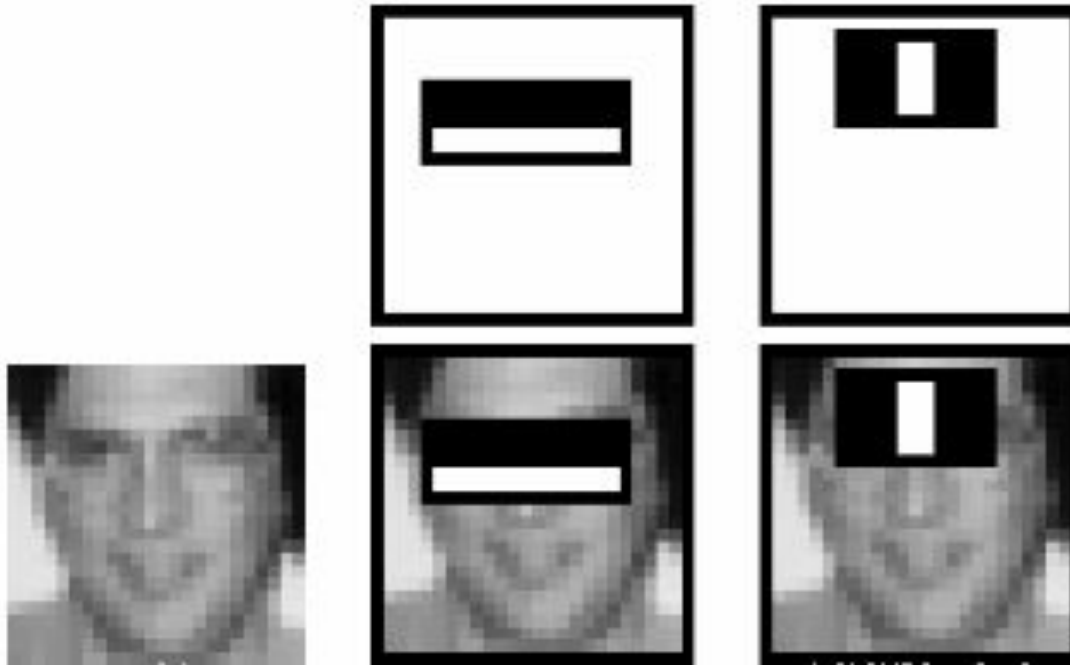


Бустинг

- Бустинг – схема классификации, основанная на комбинировании слабых классификаторов в более точный
- Обучение состоит из нескольких этапов усиления (boosting rounds)
 - на каждом этапе выбираем слабый классификатор, который лучше всех сработал на примерах, оказавшихся трудными для предыдущих классификаторов
 - «трудность» записывается с помощью весов, приписанных примерам из обучающей выборки
 - составляем общий классификатор как линейную комбинацию слабых классификаторов

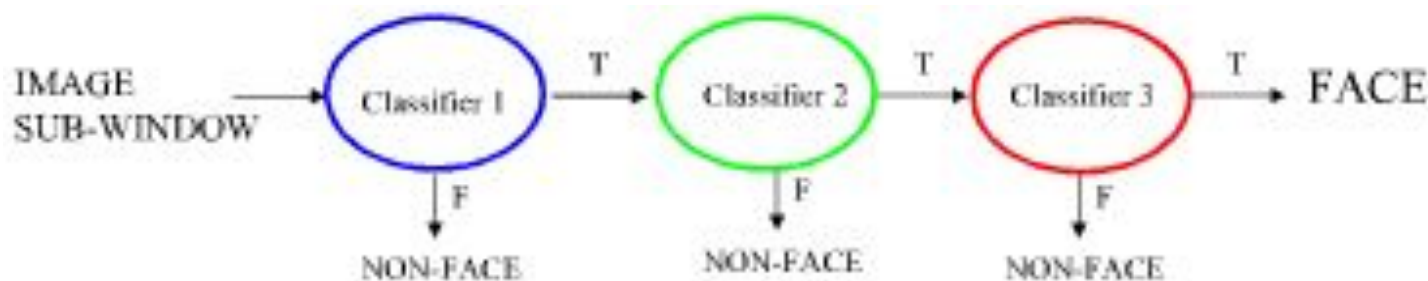
Бустинг для поиска лиц

- первый два признака, выбранные бустингом



Каскад

- Начинаем с простых классификаторов, которые отбрасывают часть отрицательных окон, при этом принимаю почти все положительные окна
- Положительный отклик первого классификатора запускает вычисление второго, более сложного, классификатора, и т.д.
- Отрицательный отклик на любом этапе приводит к немедленной отбраковке окна



Параметры каскада

- detection rate и false positive rate каскада вычисляются как произведение соответствующих ошибок каждого этапа
- detection rate 0.9 и false positive rate 10^{-6} достигаются с помощью каскада из 10 этапов, если на каждом этапе detection rate примерно равен 0.99 и false positive rate примерно 0.3

Обучение каскада

- Задаем требуемые значения detection rate и false positive rate для каждого этапа
- Добавляем признаки до тех пор, пока параметры текущего этапа не достигнут заданного уровня
- Если общий уровень false positive rate недостаточно низок, добавляем очередной этап
- Ложные обнаружения на текущем этапе используются как отрицательные примеры на следующем этапе