Донейросетевые подходы к работе с изображениями

Пузанова Анастасия

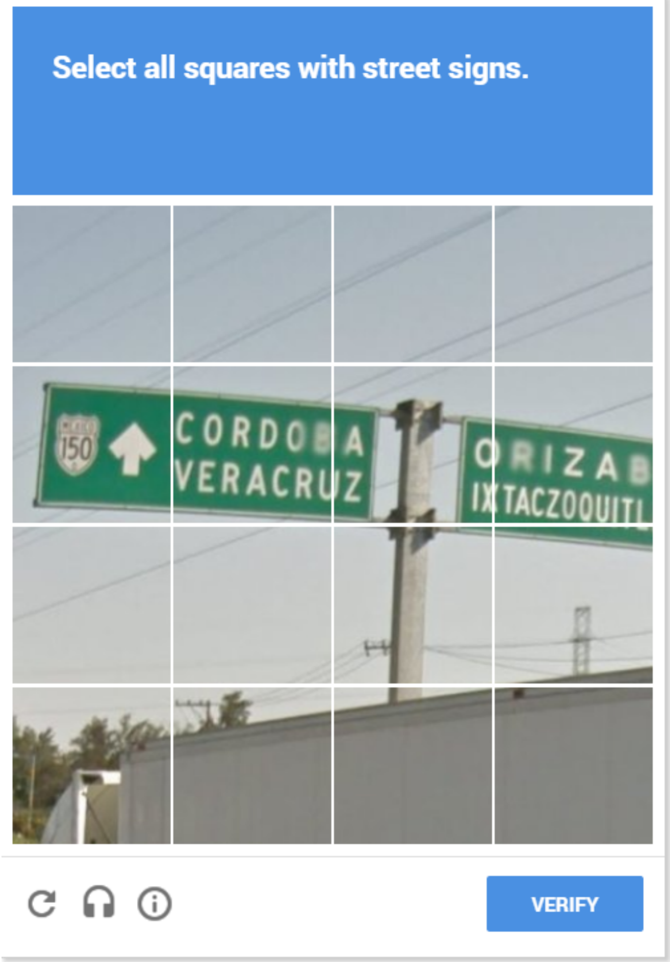
НИУ ВШЭ

12 октября 2018

Постановка проблемы

- Человек умеет работать с изображением на интуитивном уровне(сравнивать, выделять объекты, классифицировать)

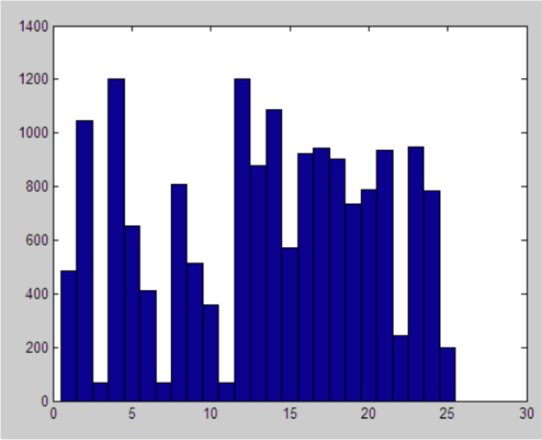
- Компьютер «видит» только матрицу пикселей

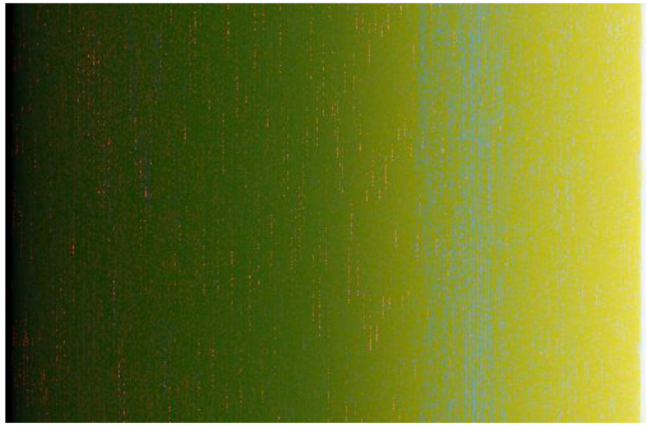


Идеи решения

- Для каждой точки изображения посчитаем значение какой-либо функции
- Получим характеристику изображения – вектор значений функции
- Будем сравнивать характеристики

Недостатки





Source: [4]

Недостатки

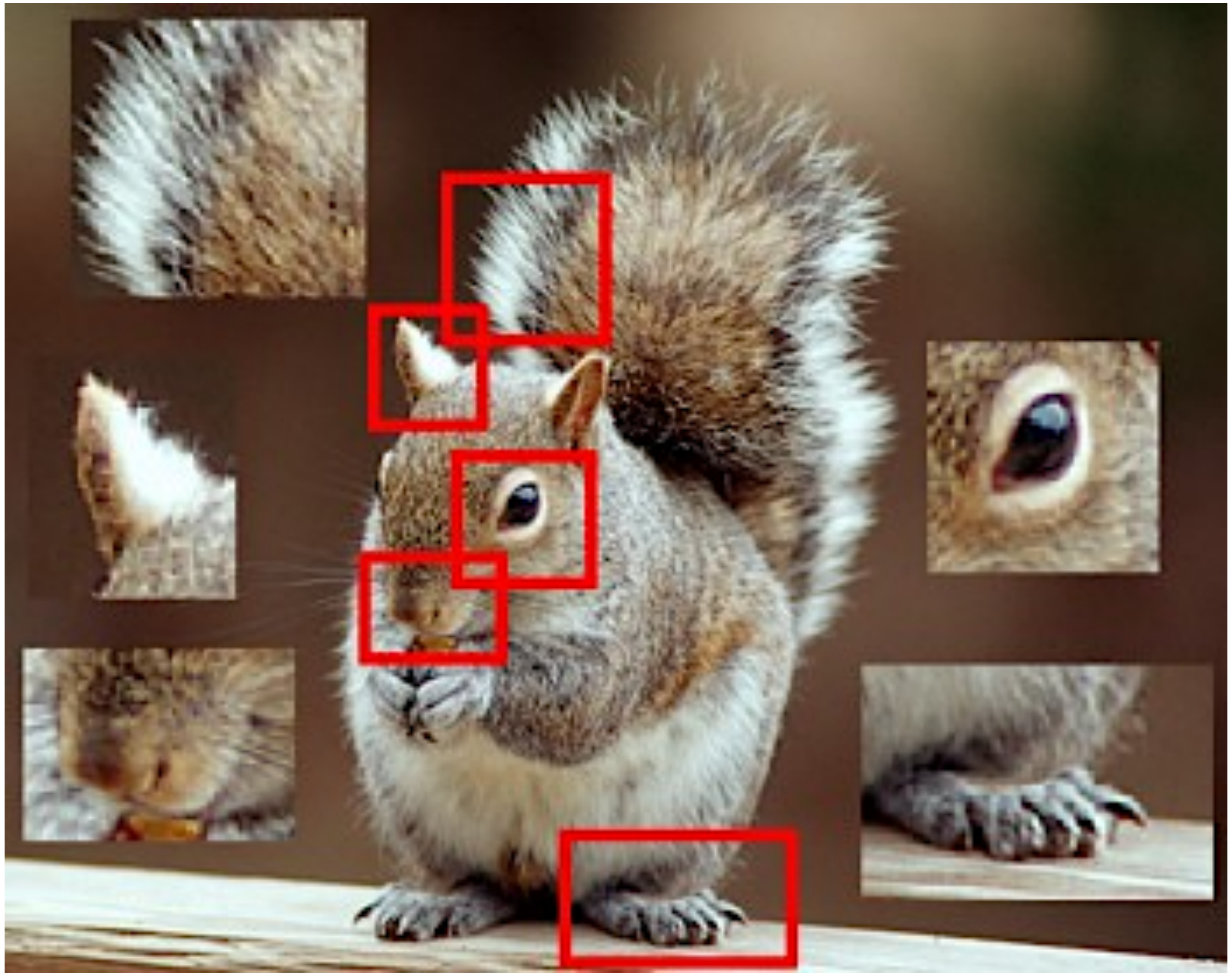
Алгоритм не инвариантен к следующим преобразованиям:

- Новые объекты на изображении, перекрытия объектов
- Шум
- Изменение масштаба
- Изменение композиции, ракурса
- Разница в освещении
- Аффинные преобразования изображения

Особенности(features)

Что изображено на слайде?

Особенности(features)



Особенности(features)

- особенности (features)
- характеристические точки (characteristic points)
- ключевые точки (keypoints)
- локальные особые точки (local feature points)

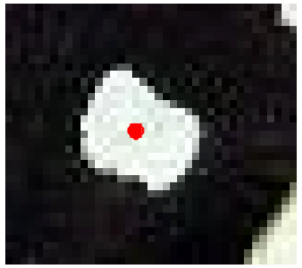
Свойства

- *Отличимость (distinctness)* – особая точка должна явно выделяться на фоне и быть отличимой (уникальной) в своей окрестности.
- *Инвариантность (invariance)* – определение особой точки должно быть независимо к аффинным преобразованиям.
- *Стабильность (stability)* – определение особой точки должно быть устойчиво к шумам и ошибкам.
- *Уникальность (uniqueness)* – кроме локальной отличимости, особая точка должна обладать глобальной уникальностью для улучшения различимости повторяющихся паттернов.
- *Интерпретируемость (interpretability)* – особые точки должны определяться так, чтобы их можно было использовать для анализа соответствий и выявления интерпретируемой информации из изображения.

Основные виды особенностей

уголки (corners)

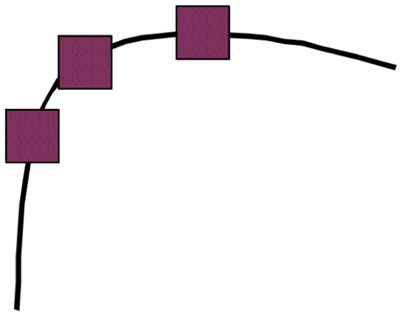
пятна (blobs)



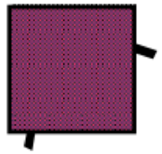
Source: [4]

Детектор блобов

- Масштаб изображения важен при поиске особенностей:



все эти точки являются углами



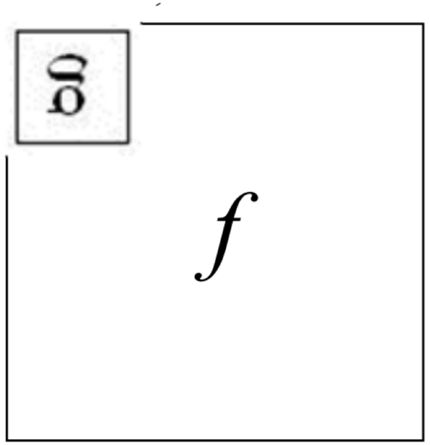
мы бы хотели видеть только один угол

- Цель: найти не только особую точку, но и ее окрестность (характеристический масштаб).
- Фактически, реализуем детектор блобов

Source: [4]

Фильтры

- Свертка:



$$(f * g)[x, y] = \sum_{i,j} f[x - i, y - j] g[i, j]$$

Где f – изображение, g – ядро свертки

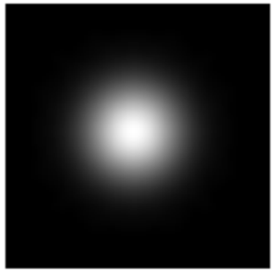
- Свойства:
  - Линейна
  - Инвариантна к сдвигу
  - Коммутативна, ассоциативна, дистрибутивна по сложению и т.д.
  - Операции свертки и дифференцирования ассоциативны

$$\frac{d}{dx}(f * g) = f * \frac{d}{dx} g$$

Source: [4]

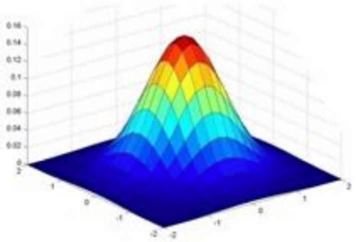
Какое ядро выбрать?

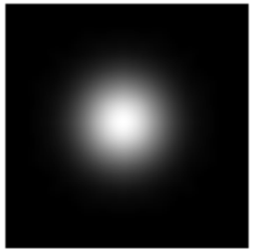
- Точка света в расфокусированном объективе выглядит как:



“fuzzy blob”

- Хотим использовать похожий фильтр:





Source: [4]

Фильтр Гаусса

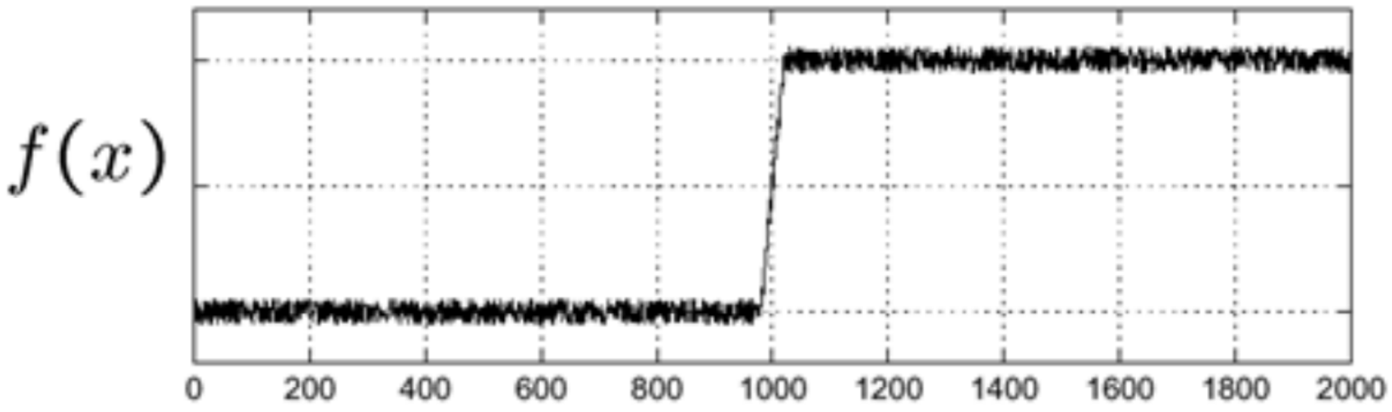
- Ядро:

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

- Используется для снижения уровня шума в изображении
- Размытие гауссовым фильтром похоже на изменение масштаба изображения

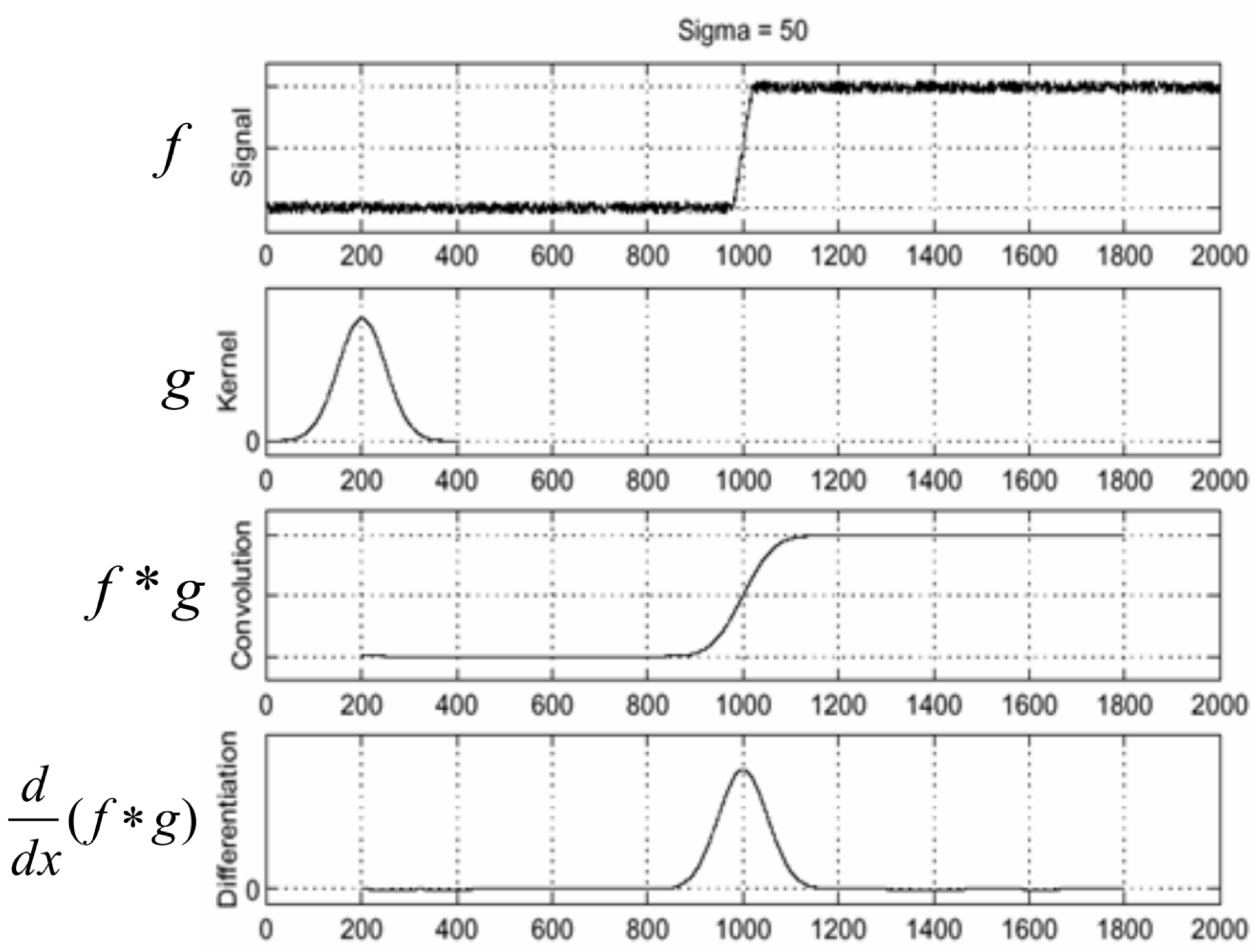
DoG

- Рассмотрим строку изображения(аналогично столбец)
- Нарисуем график зависимости яркости от x координаты
- Резкий перепад яркости соответствует краю на изображении



DoG

- Сглаживаем исходную кривую Фильтром Гаусса
- Дифференцируем по x
- Край – пик производной после применения фильтра



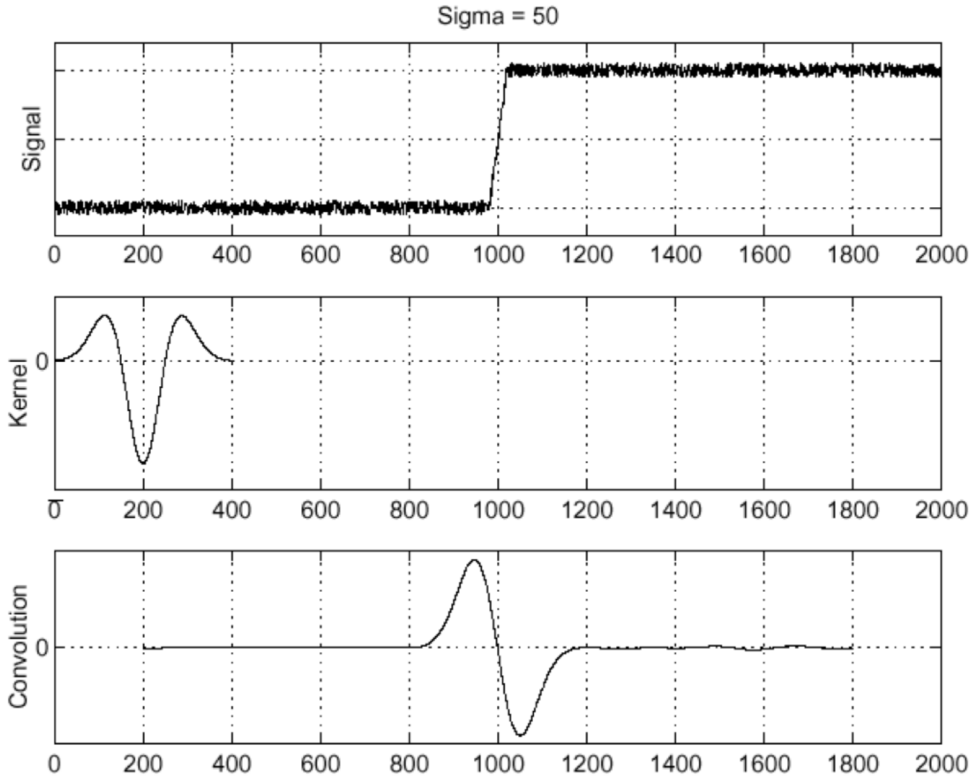
Source: [4]

DoG

$$f$$

$$\frac{d^2}{dx^2} g$$

$$f * \frac{d^2}{dx^2} g$$



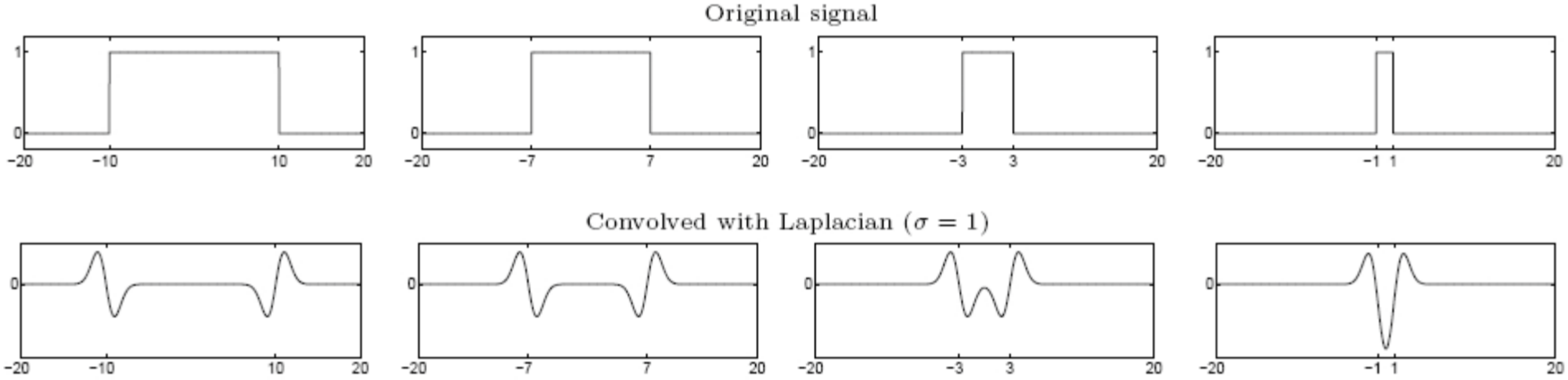
Край

Вторая производная Гауссианы (Лапласиан)

Край = переход через ноль второй производной

Source: [4]

DoG



- Размер блоба «соответствует» размеру Лапласиана Гауссиана с максимальной величиной отклика

Source: [4]

DoG

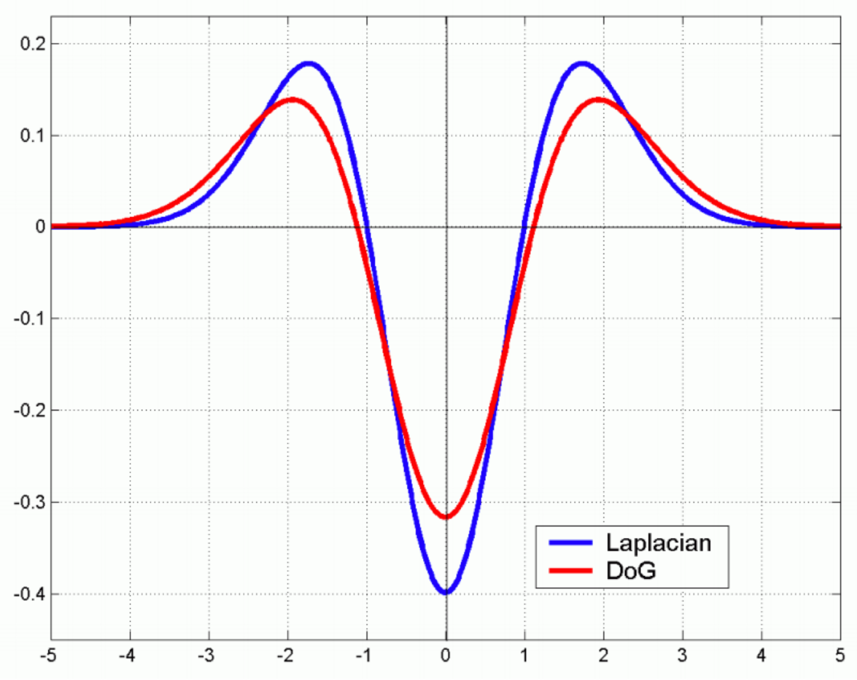
- Приближение Лапласиана разницей Гауссиан

$$L = \sigma^2 \left(G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma) \right)$$

(Лапласиан)

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

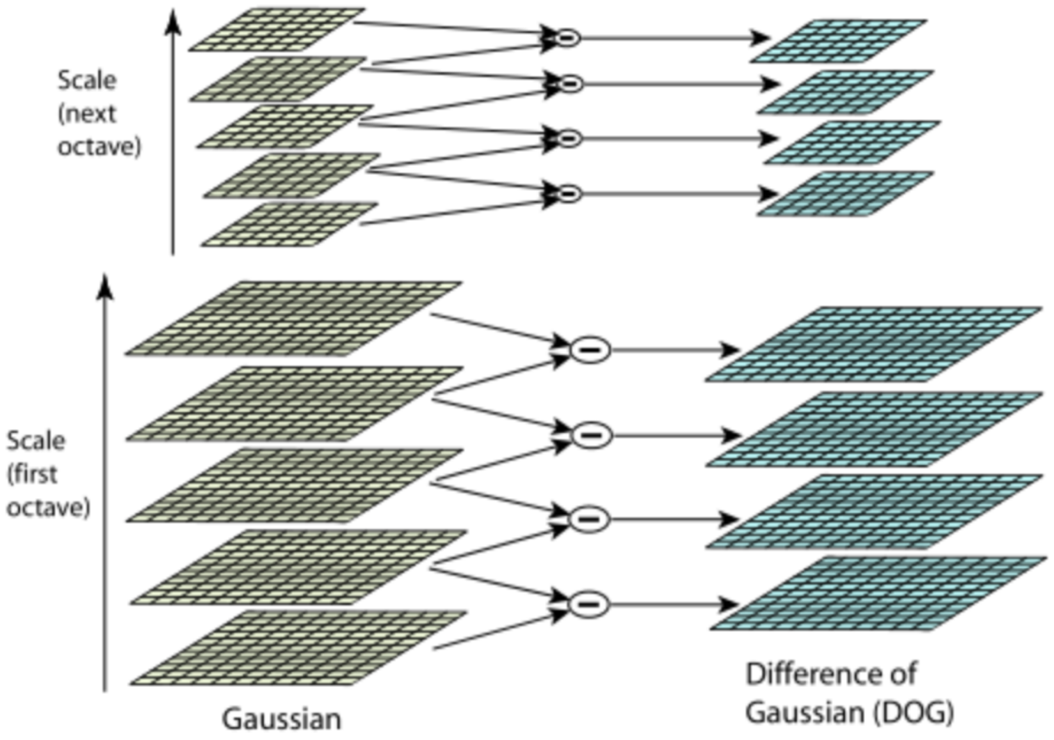
(Разница Гауссиан)



Source: [4]

DoG

- Строим пирамиду Гауссиан
- Параллельно строим пирамиду их разностей



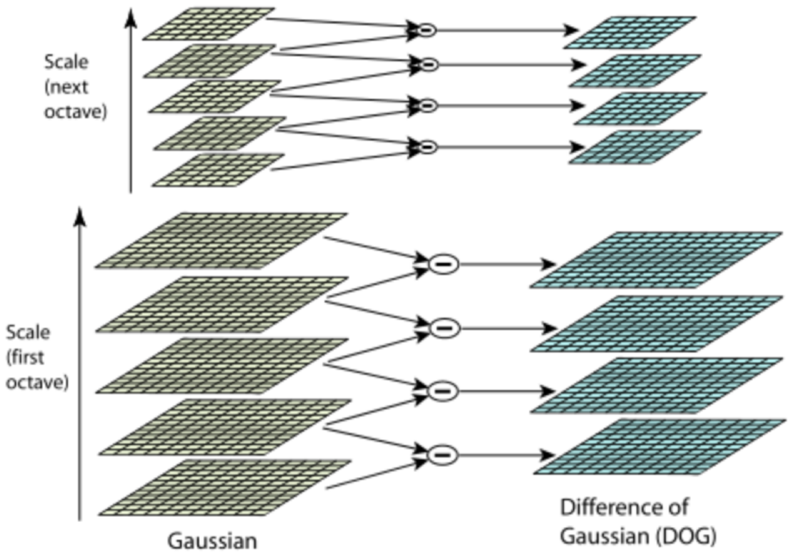
Source: [1]

DoG

- Каждая октава – набор изображений, размытых фильтром
- Если в октаве s интервалов, то k = $2^{1/s}$

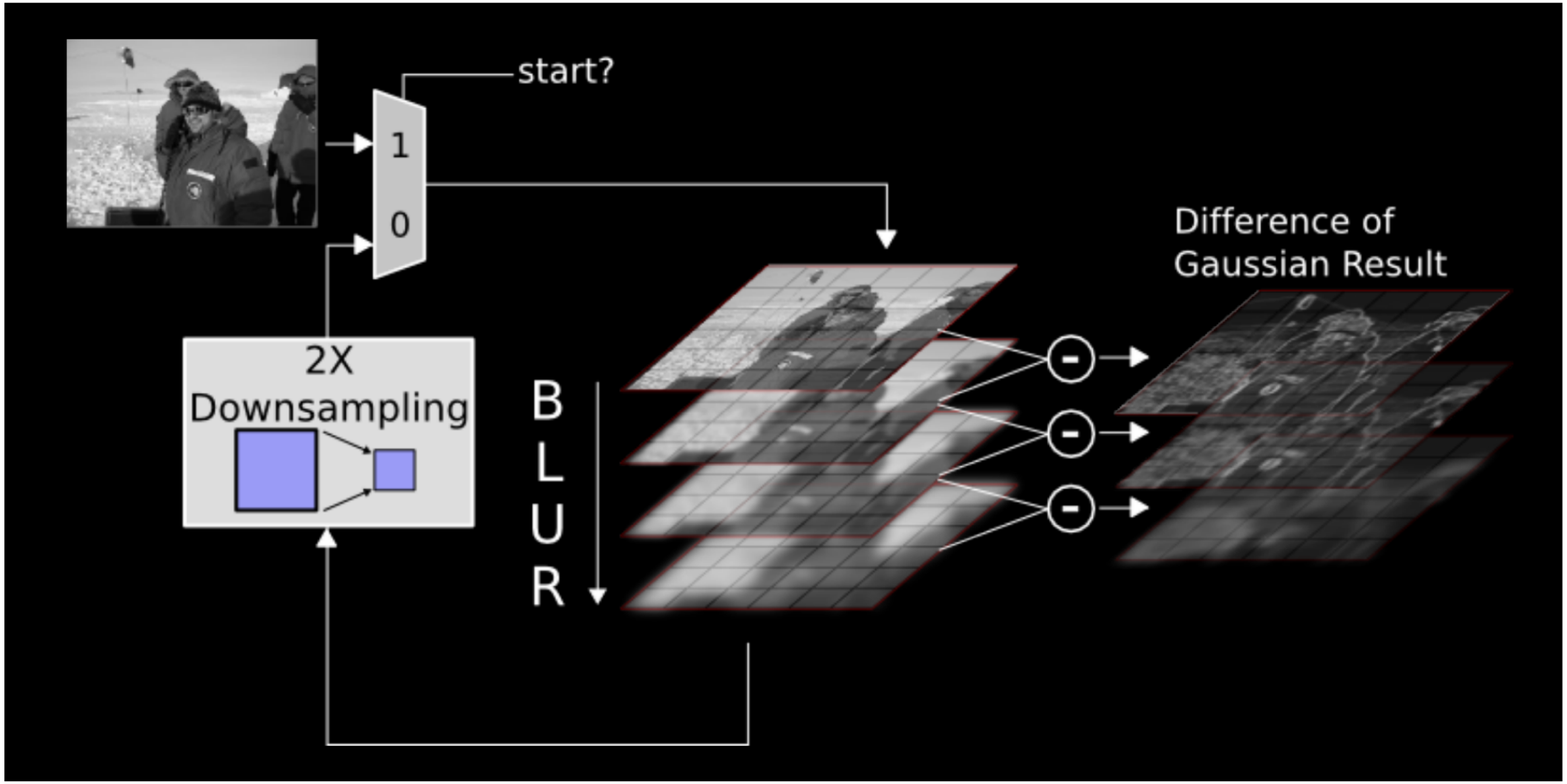
$$\boxed{DoG = G(x, y, k\sigma) - G(x, y, \sigma)}$$

- Для каждой октавы строим на 2 изображения больше (s+3)
- В следующей октаве изображения в 2 раза меньше
- Первое изображение текущей октавы и последнее изображения предыдущей имеют одинаковое размытие



Source: [1]

DoG

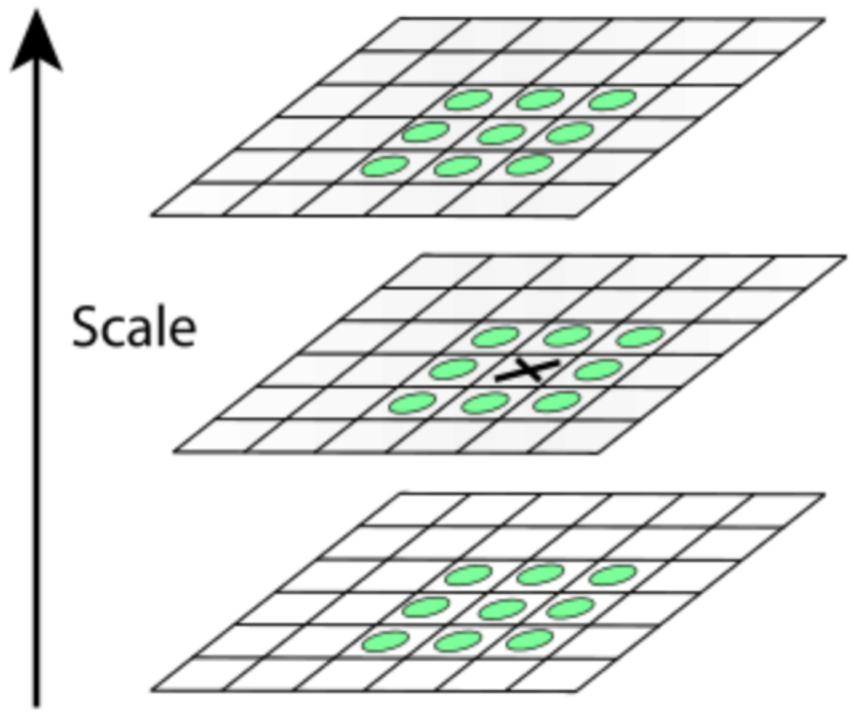


start?
1
0
2X
Downsampling
B
L
U
R
Difference of
Gaussian Result

Source: [3]

DoG

- Находим локальный экстремум в «3D»



Source: [1]

DoG

- Итог:
  - Аппроксимируем свертку Лапласианом
  - Строим пирамиду Гауссиан различных размеров и степеней размытия
  - Считаем разности между соседними изображениями
  - Ищем локальный экстремум
  - Характеристический масштаб определяется уровнем, на котором нашли экстремум

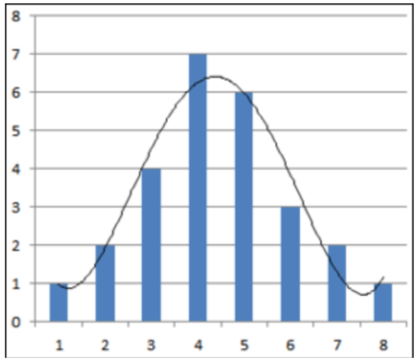
Уточнение особых точек

- Уточнение координат точки с субпиксельной точностью
- Удаление точек с малым контрастом
- Удаление точек на границе объектов

Уточнение координат точки

- DoG дискретна. Координаты строго по пикселям, размытие по уровням
- Аппроксимируем многочленом Тейлора 2 порядка в точке экстремума

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathbf{T}}\frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x} \quad , X = (x, y, \sigma)$$



- Находим экстремум, приравняв производную к 0

$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1}\frac{\partial D}{\partial \mathbf{x}}$$

- Считаем производные с помощью конечных разностей(по сетке значений)
- Если какая-то из компонент больше половины шага сетки, сдвигаемся в положительном направлении

Удаление точек с малым контрастом

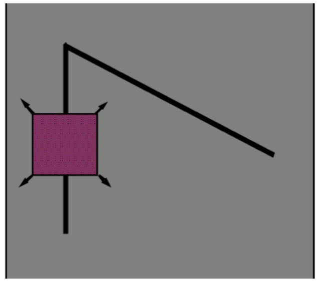
- Проверим абсолютное значение DoG в найденной уточненной точке

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2}\frac{\partial D}{\partial \mathbf{x}}^T \hat{\mathbf{x}}$$

- В оригинальной статье удаляли точки со значением меньше 0.03

Удаление точек на границе объектов

- Яркость таких точек не меняется вдоль одного направления, но сильно меняется вдоль других



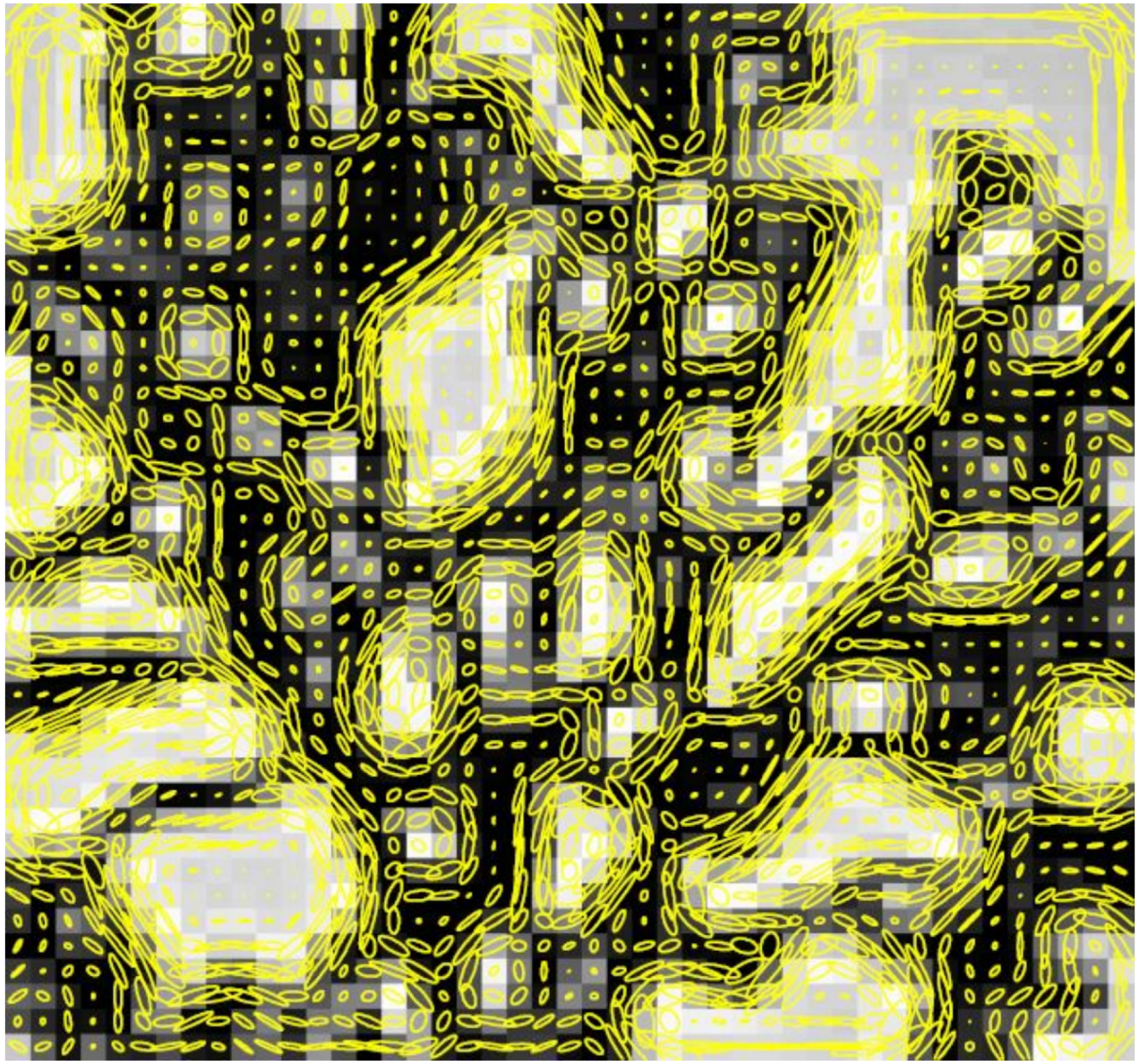
- Посчитаем матрицу Гессе $\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$
- Симметричная, значит можно привести к диагональному виду

$$R^{-1}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}R$$

- Собственные значения можно визуализировать в виде эллипса, направленного вдоль границы. Длины осей которого определены собственными значениями, а сам эллипс отражает изменение яркости.

Source: [4]

Удаление точек на границе объектов



Source: [4]

Удаление точек на границе объектов

- Если α и β – собственные числа и $\alpha = r\beta$, то чем меньше r, тем меньше «вытянут» эллипс
- $$\mathrm{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta,$$
  $$\mathrm{Det}(\mathbf{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta.$$
- $$\frac{\mathrm{Tr}(\mathbf{H})^2}{\mathrm{Det}(\mathbf{H})} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r}$$
- Зафиксируем r и будем выбирать только точки с $\frac{\mathrm{Tr}(\mathbf{H})^2}{\mathrm{Det}(\mathbf{H})} < \frac{(r+1)^2}{r}$

Заключение

- Для работы с изображениями необходимо выделить в них особые точки – характеристику изображения
- Найти особенности, а также соответствующий им масштаб окрестности можно с помощью DoG
- Уточним расположение особых точек
- Удалим из рассматриваемых точки, несущие мало информации

Источники

[1] - David G. Lowe. "Distinctive image features from scale-invariant keypoints.”, 2004

[2] - Chris Harris & Mike Stephens. “A combined corner and edge detector”, 1988

[3] – Chris Murphy, Ballard Blair. “Difference of Gaussian Scale-Space Pyramids for SIFT Feature Detection”, 2007

[4] – Антон Конушин. Курс “Компьютерное зрение”. Лекция 1 ”Локальные особенности”, 2017