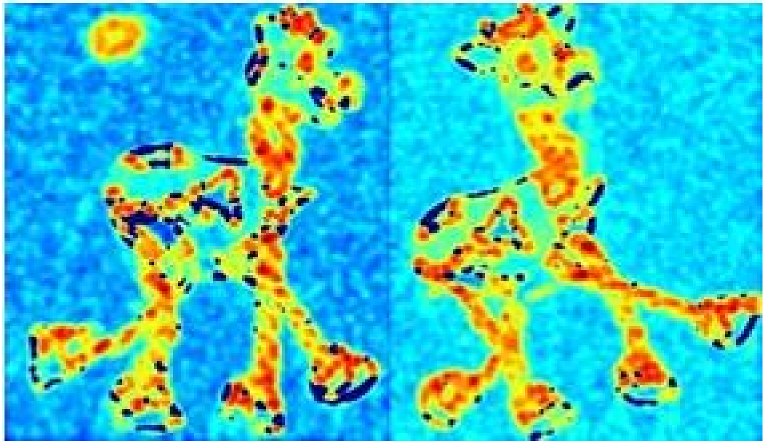
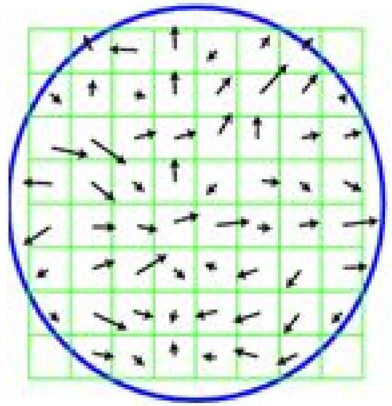
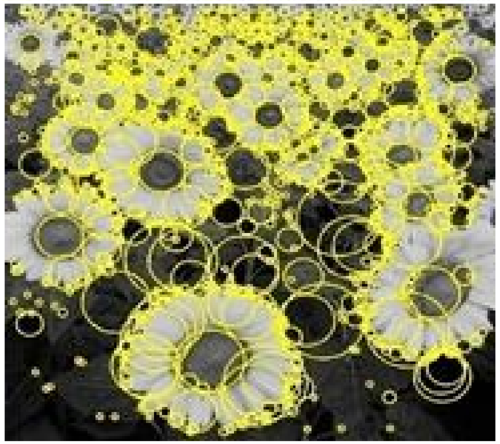
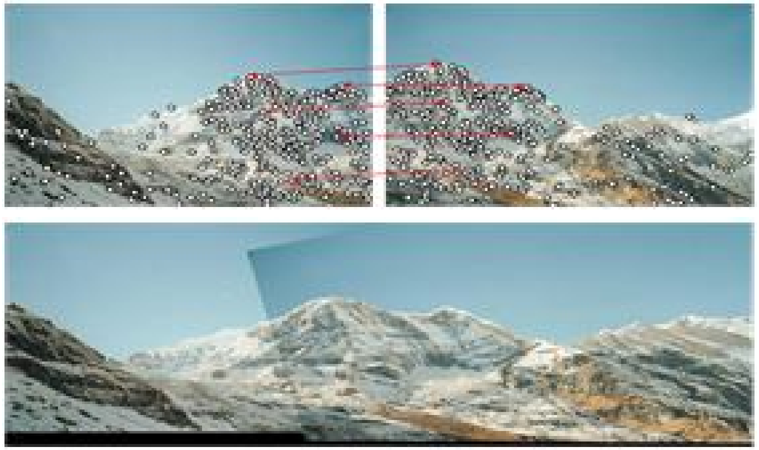
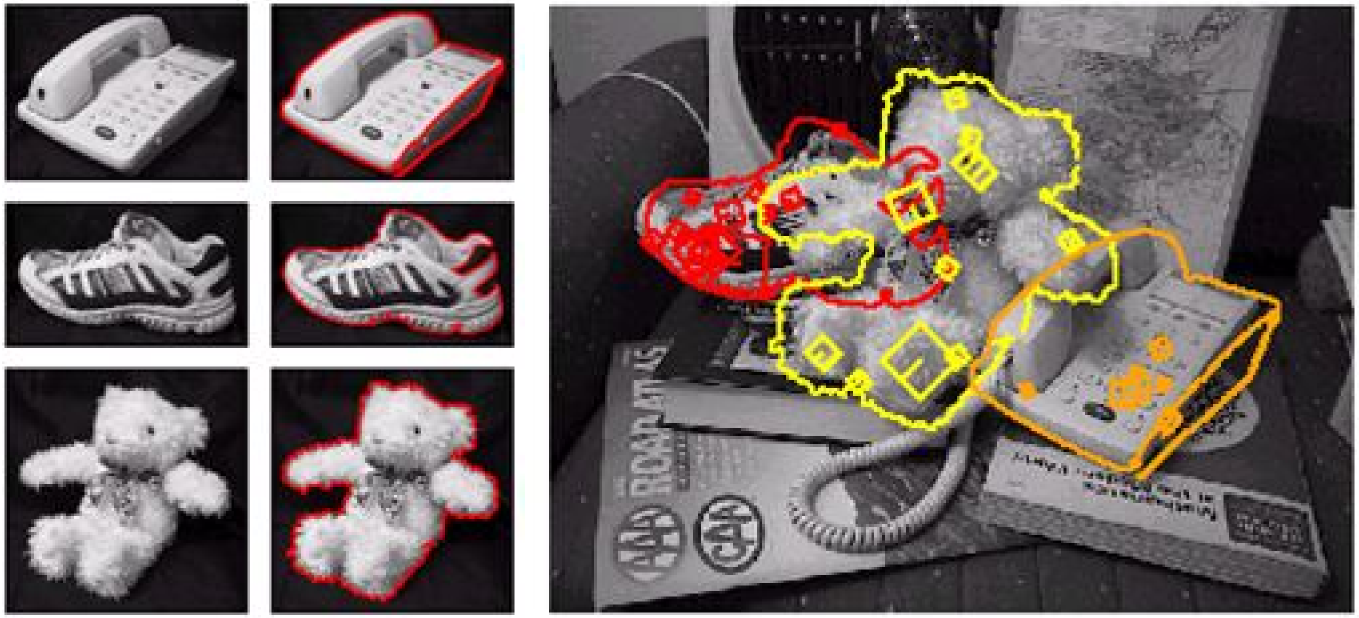
Локальные особенности и согласование изображений


Антон Конушин

Many slides adopted from Svetlana Lazebnik, Steve Seitz and Alexey Efros

Задача сопоставления изображений



- Есть несколько изображений конкретных объектов
- Хотим найти эти объекты на тестовом изображении
- Попробуем «сопоставить» изображения объектов с тестовым изображением
- Задача «image matching»

Вводный пример

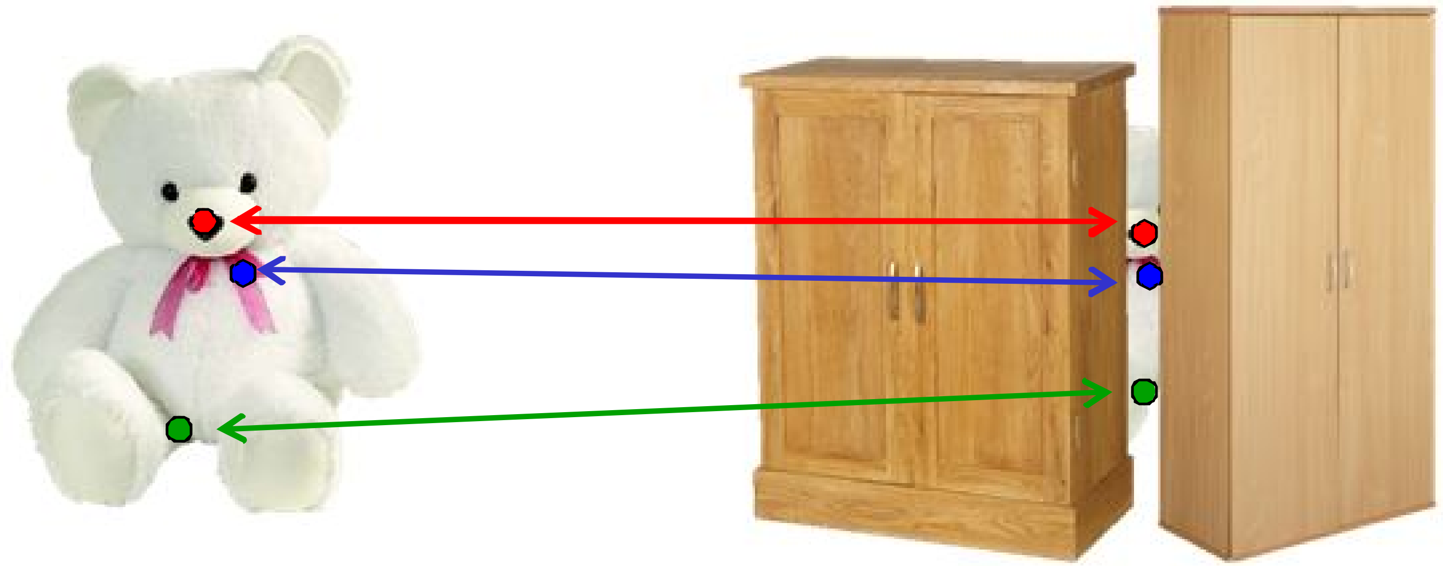
Где медведь? 

Задача №1

Задача №2

- Почему во втором случае было легче его найти?
- Были видны «характерные» фрагменты медведя

Особенности (features)

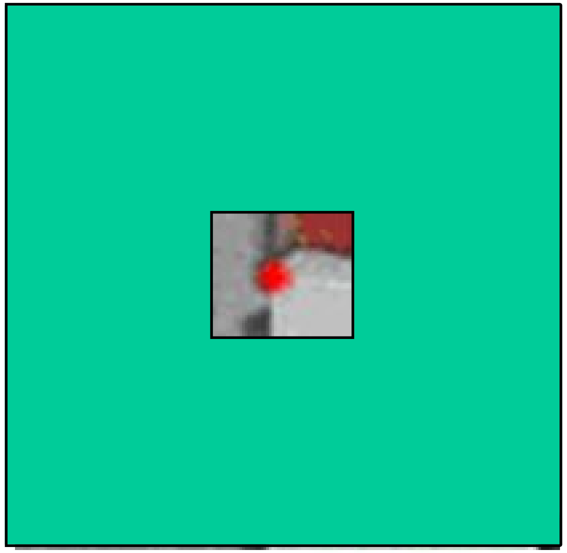


- «Хорошо различимые фрагменты» объекта
  - «особенности» (features)
  - «характеристические точки» (characteristic points)
  - «локальные особые точки» (local feature points)
- Характерные фрагменты позволяют справится с изменениями ракурса, масштаба и перекрытиями

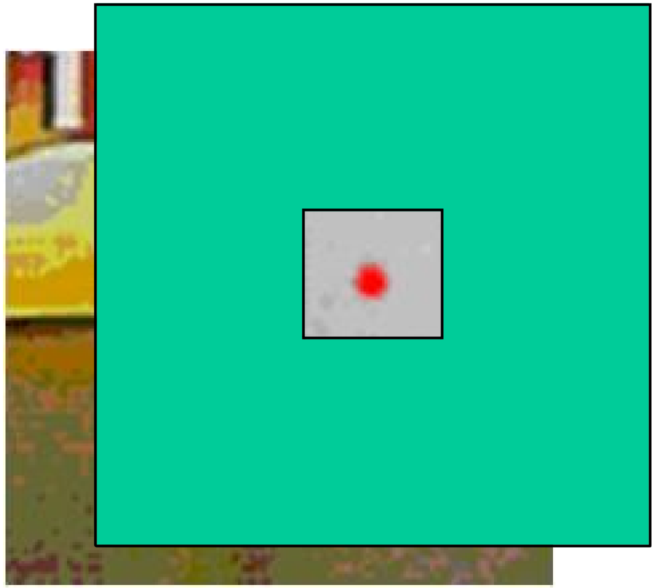
Требования

- Какие можно сформулировать требования к «хорошо различимым фрагментам» объекта?
- Отличаются от большинства других фрагментов объекта
- Инвариантны к изменению освещения
- Инвариантны к изменению ракурса
  - Можно находить одну и ту же точку на измененных изображениях
  - Можем «идентифицировать» эту точку

Локальные особенности



Пример особой точки



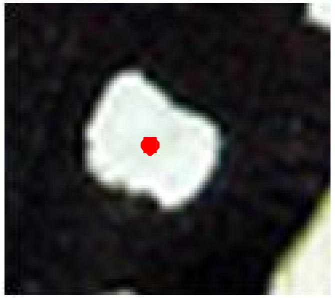
Пример точки, не являющейся особой

- Какая из двух точек является характерной («особой»)?
- Локальная (особая) точка *p* изображения *I* должна обладать «характерной окрестностью» D, т.е. отличаться от всех точек в некоторой окрестности *p*
- Для определения, является ли точка «характерной», нам достаточно только её окрестности

Два основных класса особенностей

уголки (corners)

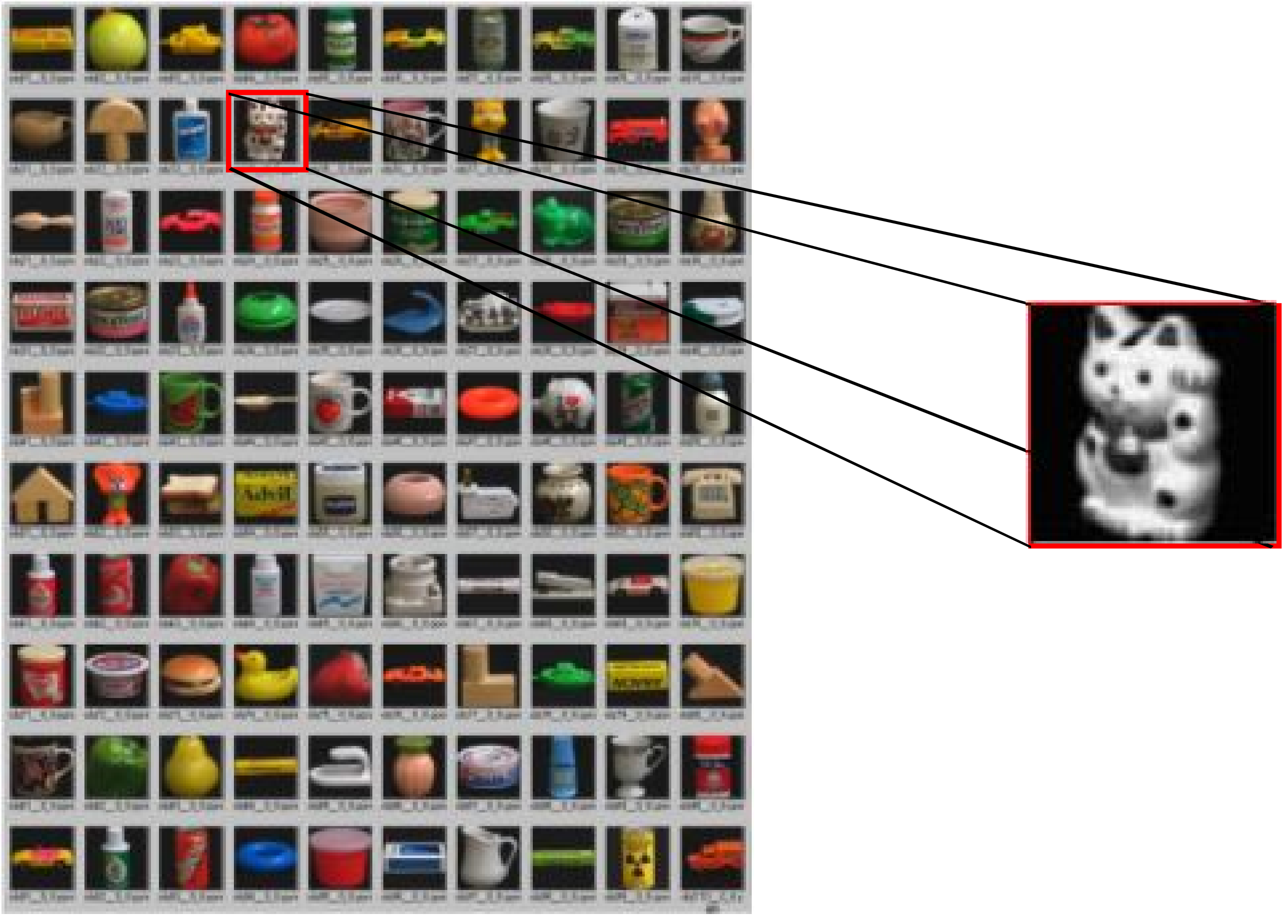
пятна (blobs)



Применение

Поиск и выделение объектов, распознавание изображений

Применение

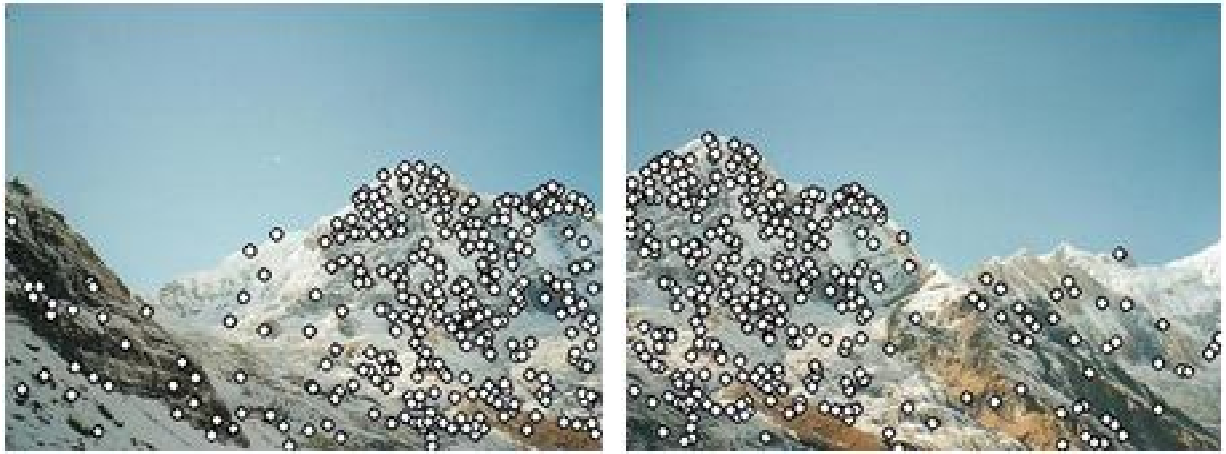


Поиск изображений по содержанию в базе изображений

Применение

Сопоставление изображений, построение панорам и трёхмерная реконструкция

Требования к особенностям



- Локальность (Locality)
  - Особенность занимает маленькую область изображения, поэтому работа с ней нечувствительна к перекрытиям
- Повторимость (Repeatability)
  - Особенность находится в том же месте объекта не смотря на изменения масштаба, положения, ракурса и освещения
- Значимость (Saliency)
  - Каждая особенность имеет уникальное (distinctive) описание
- Компактность и эффективность
  - Количество особенностей существенно меньше числа пикселей изображения

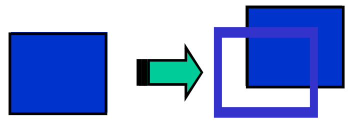
Повторимость

- Особенность должна находится в том же месте объекта не смотря на изменения масштаба, положения, ракурса и освещения изображения
- Как можно это проверить?
  - Выделим характерные особенности на изображении объекта
  - Применим к изображению геометрическое или цветовое преобразование
  - Выделим характерные особенности на изменённом изображении, они должны найтись в тех же местах объекта
- Метод выделения особенностей должен быть **«инвариантным»** к преобразованиям

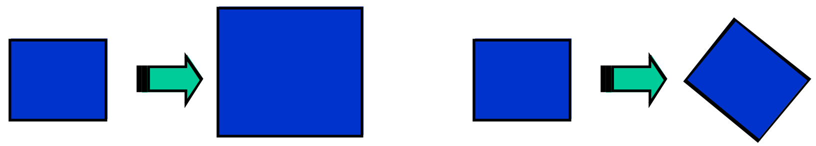
Slide by S. Lazebnik

Геометрические преобразования

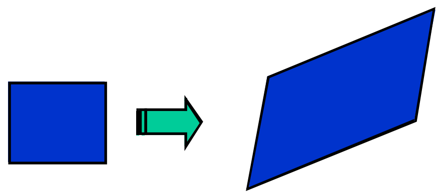
- Параллельный перенос



- Подобие (перенос, масштаб, поворот)



- Аффинное



Аффинное преобразование даёт хорошее приближение искажений, претерпеваемых небольшим плоским фрагментом объекта при малом изменении ракурса

Геометрические преобразования

- Параллельный перенос

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

- Евклидово преобразование (M – ортогональная матрица)

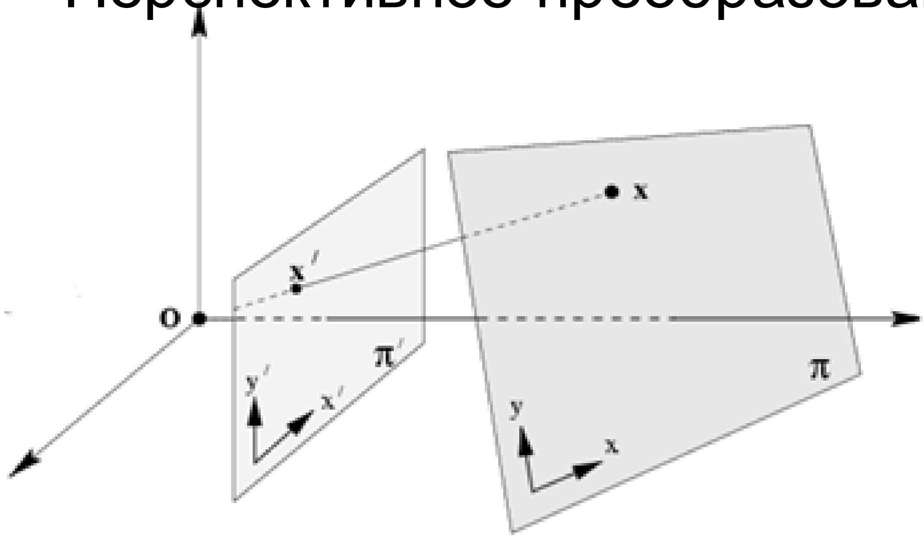
$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

- Аффинное преобразование

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

Гомография

Перспективное преобразование плоскости



$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \cong \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{12} & h_{33} \end{bmatrix} * \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Перевод в однородные координаты $(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$

Перевод из однородных координат $\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$

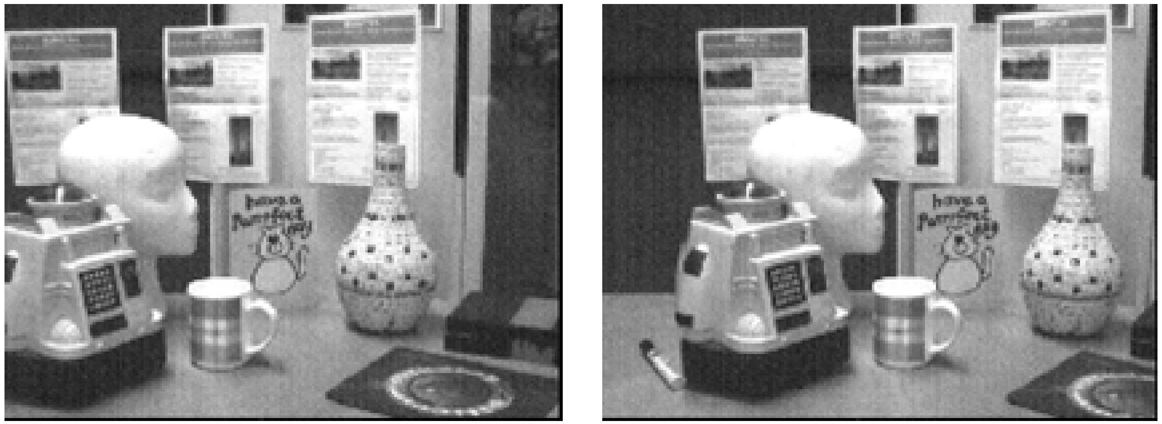
Удобнее представлять себе так:

$$\begin{bmatrix} wx & wy & w \end{bmatrix}^T \cong \begin{bmatrix} x & y & 1 \end{bmatrix}^T \cong \begin{bmatrix} x & y \end{bmatrix}^T$$

Гомография

- Преобразование между 2мя разными видами одной и той же плоскости

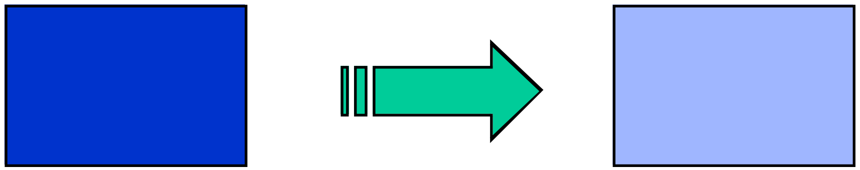
- Преобразование между видами с повернутой камеры (центр проекции общий)



Slide by S. Lazebnik

Фотометрическое преобразование

Аффинное изменение яркости ($I \to a\, I + b$)

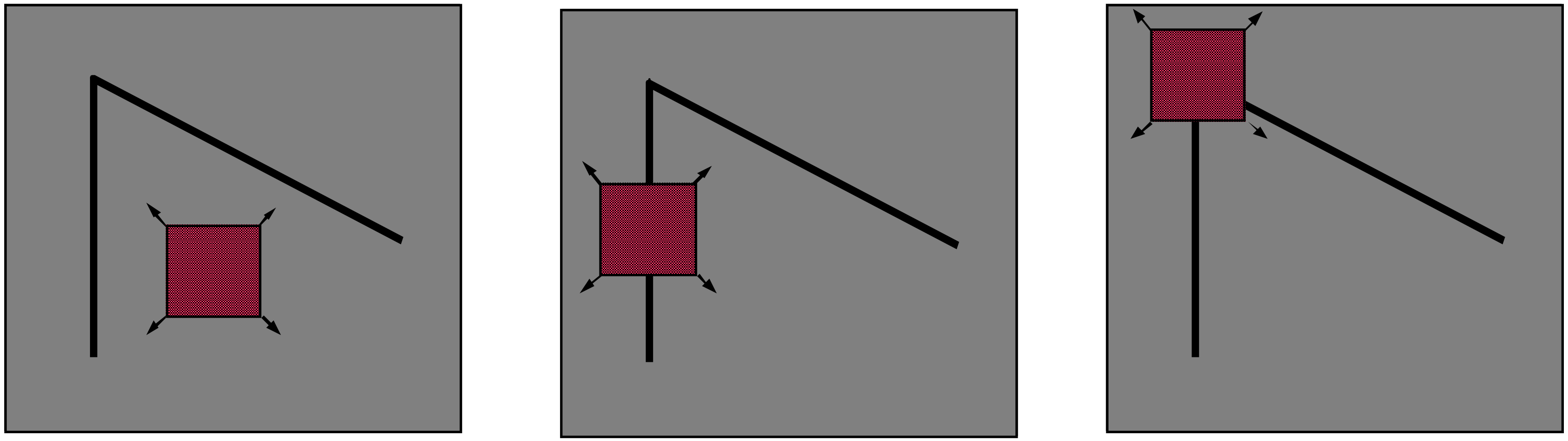


Его вполне достаточно для моделирования устойчивости методы выделения особенностей к изменению условий освещения

Особенно, если будем работать с серыми изображениями!

Локальные особенности

Проведём эксперимент, будем рассматривать разные точки на изображении и проверять, являются ли они локальными особенностями

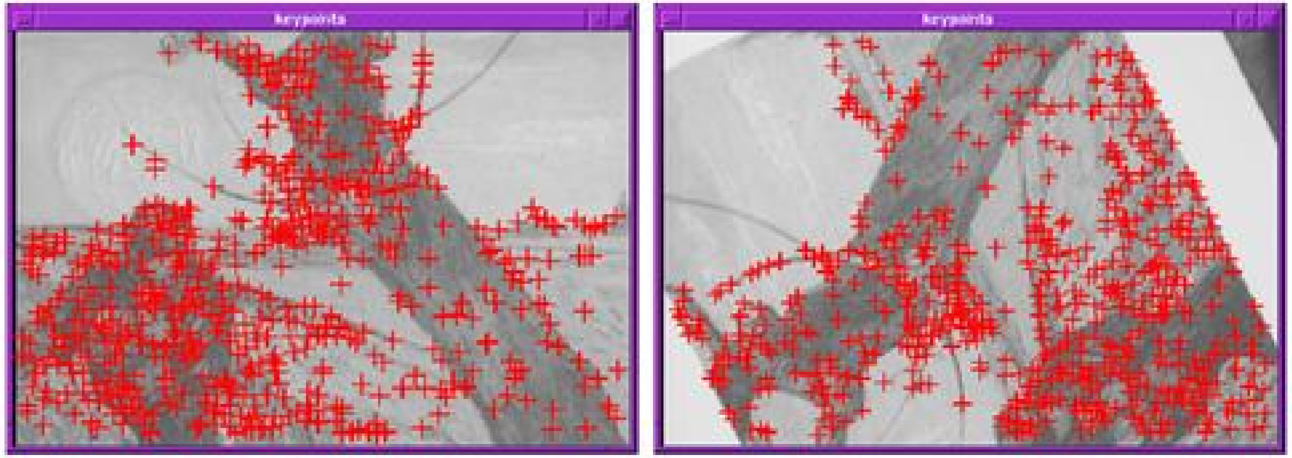


монотонный регион: в любом направлении изменений нет

«край»: вдоль края изменений нет

«уголок»: изменения при перемещении в любую сторону

Детектор Харриса



- Наиболее популярный детектор локальных особенность точек – детектор Харриса (Harris)
- Ищет такие точки (x,y), окрестность которых меняется при любом сдвиге $(x+u, y+v)$
- Такие точки обычно оказываются углами, поэтому метод ещё называют «детектор углов»

C.Harris and M.Stephens. "A Combined Corner and Edge Detector.“ *Proceedings of the 4th Alvey Vision Conference*: pages 147—151, 1988

Устройство метода

Изменение окрестности точки (x,y) при сдвиге [*u*,*v*]:

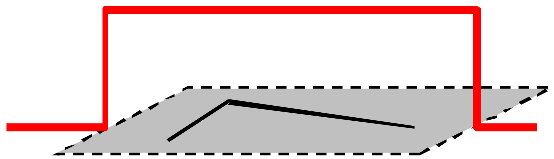
$$E(u,v)=\sum_{x,y} w(x,y)\left[I(x+u,y+v)-I(x,y)\right]^2$$

Функция окна

Сдвиг яркости

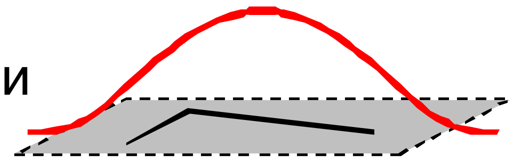
Яркость

Функция окна $w(x,y) =$



1 в окне, 0 вне

или



Гауссиана

Source: R. Szeliski

Устройство метода

Изменение окрестности точки при сдвиге [u,v]:

$$E(u,v) = \sum_{x,y} w(x,y)\left[I(x+u,y+v) - I(x,y)\right]^2$$

Разложение в ряд Тейлора 2го порядка $I(x,y)$ вокруг (x,y) (билинейная интерполяция при маленьких сдвигах)

$$[I(x+u,y+v) - I(x,y)]^2 \cong [I(x,y) + I_x u + I_y v - I(x,y)]^2$$

$$= [I_x u + I_y v]^2 = I_x^2 u^2 + 2 I_x I_y uv + I_y^2 v^2$$

$$= (u,v)\begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}(u,v)^T$$

Slide by S. Lazebnik

Устройство метода

Итого изменение окрестности можно свести к:

$$E(u,v) \approx [u \ \ v] \ M \begin{bmatrix} u \\ v \end{bmatrix}$$

где M – матрица 2×2 вычисленная по частным производным:

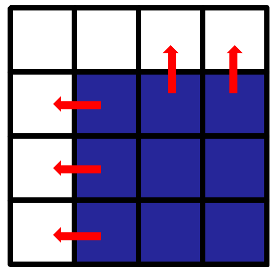
$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

$$M = \begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} = \sum \begin{bmatrix} I_x \\ I_y \end{bmatrix} [I_x \ I_y] = \sum \nabla I (\nabla I)^T$$

Slide by S. Lazebnik

Интерпретация матрицы моментов

Рассмотрим случай, когда градиенты выровнены по осям (вертикальные или горизонтальные)



$$M = \sum \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Если одно из λ близко к 0, тогда это не угол, и нужно искать другие точки

Общий случай

M – симметричная, поэтому её можно привести к диагональному виду:

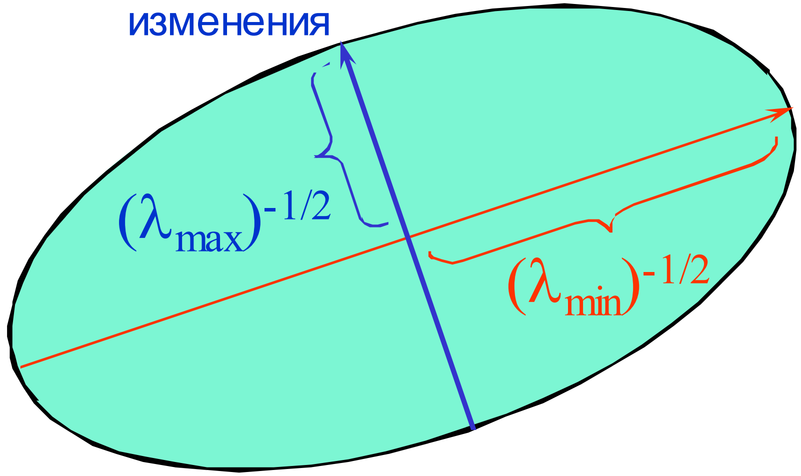
$$M = R^{-1}DR = R^{-1}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R$$

R – ортогональная матрица из собственных векторов M, D – диагональная из собственных значений M

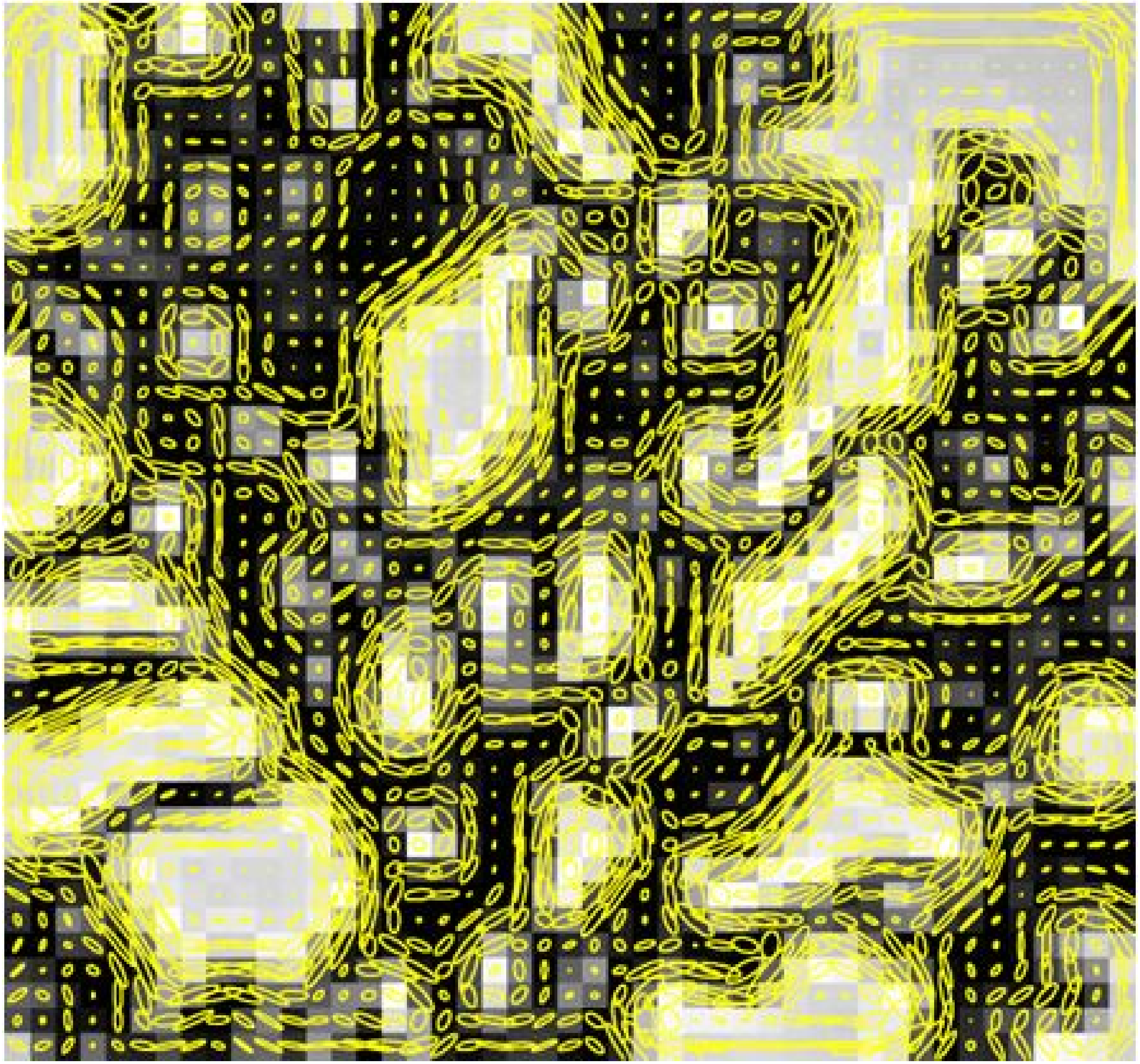
Матрицу *M* можно визуализировать в виде эллипса, у которого длины осей определены собственными значениями, а ориентация определена матрицей R

Уравнение эллипса:

$$[u \;\; v] \;\; M \begin{bmatrix} u \\ v \end{bmatrix} = \text{const}$$



Пример



Визуализация матриц вторых моментов (Гессианов)

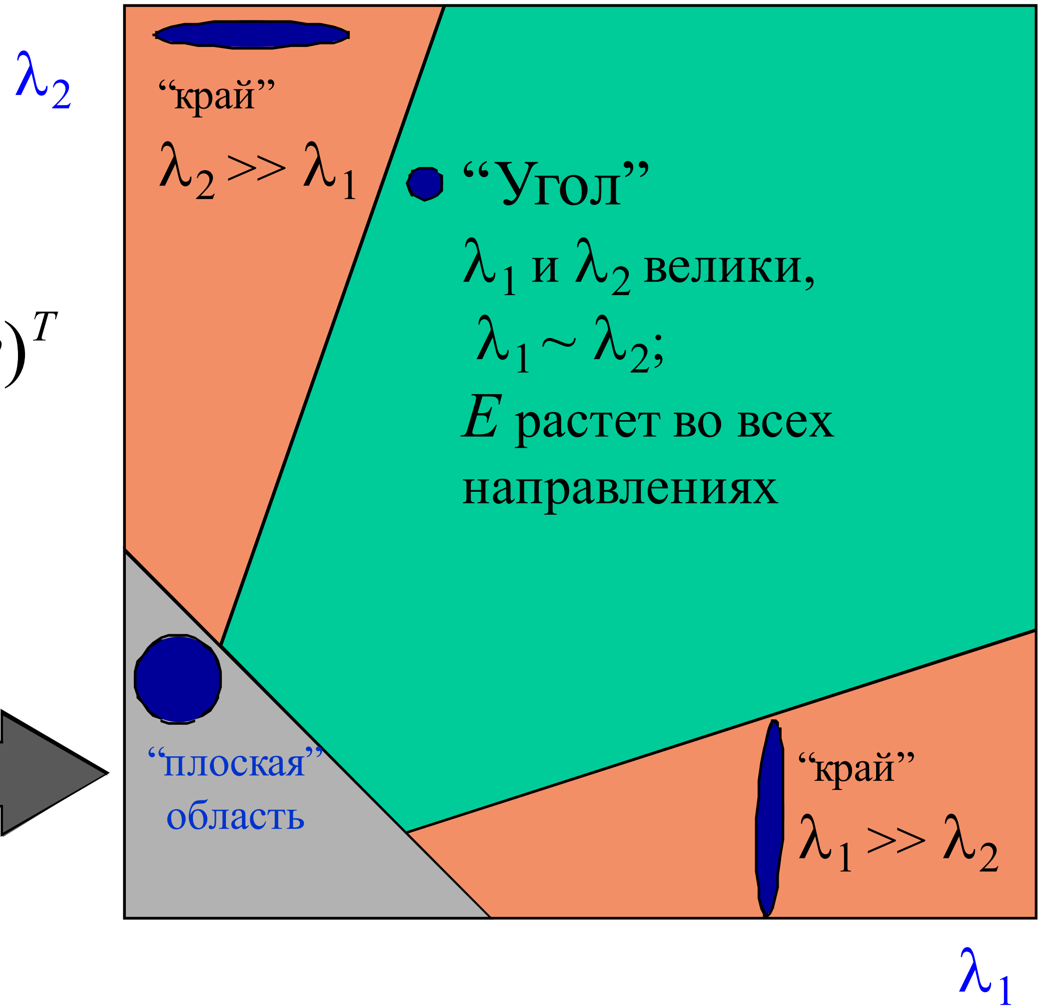
Зависимость E от λ

Классификация точек изображения по собственным значениям матрицы производных *M*

$$E(u,v) = (u,v)M(u,v)^T$$

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

λ_1 и λ_2 малы;
E не меняется по всем направлениям



Функции отклика углов

- Функция отклика угла по Харрису:

$$R = \det M - k\left(\operatorname{trace} M\right)^2$$

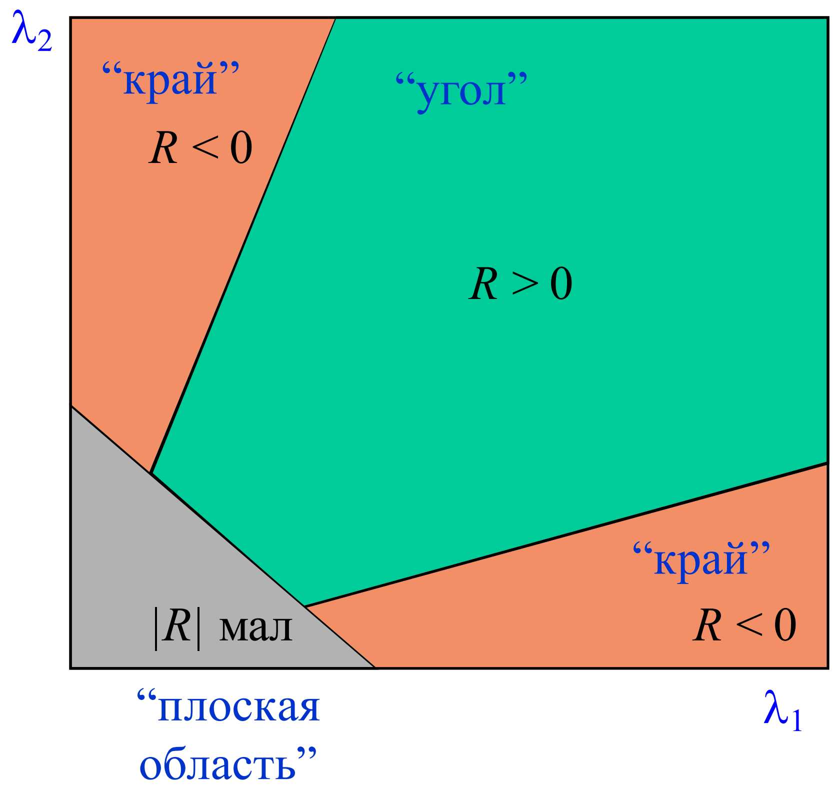
$$\det M = \lambda_1 \lambda_2$$

$$\operatorname{trace} M = \lambda_1 + \lambda_2$$

$(k = 0.04\text{-}0.06)$

- Функция по Фёрстнеру (Forstner):

$$R = \det M \,/\, \operatorname{trace} M$$

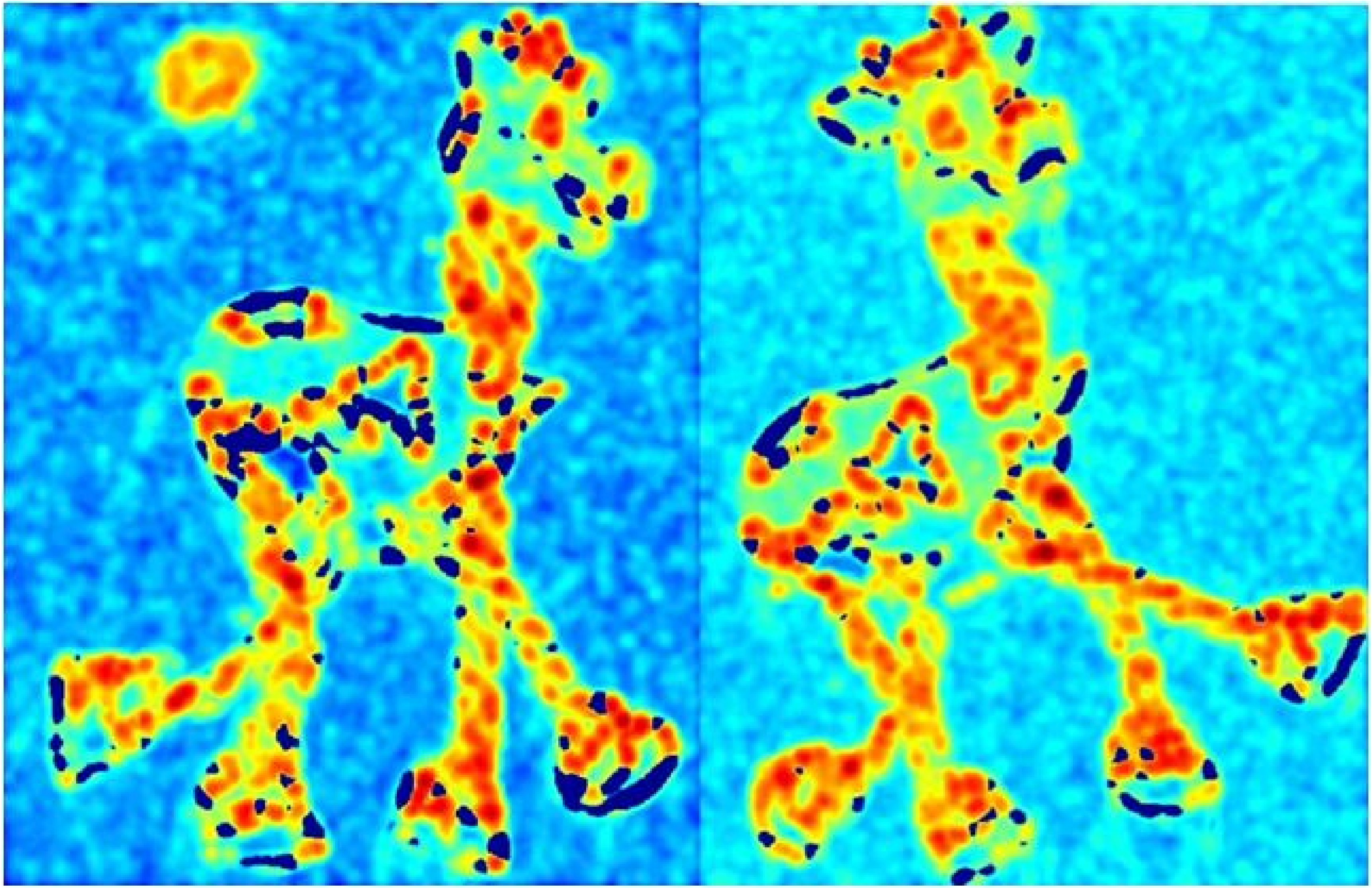


Foerstner, W; Gulch. "A Fast Operator for Detection and Precose Location of Distinct Points, Corners and Centres of Circular Features". *ISPRS 1997*

Демонстрация по шагам

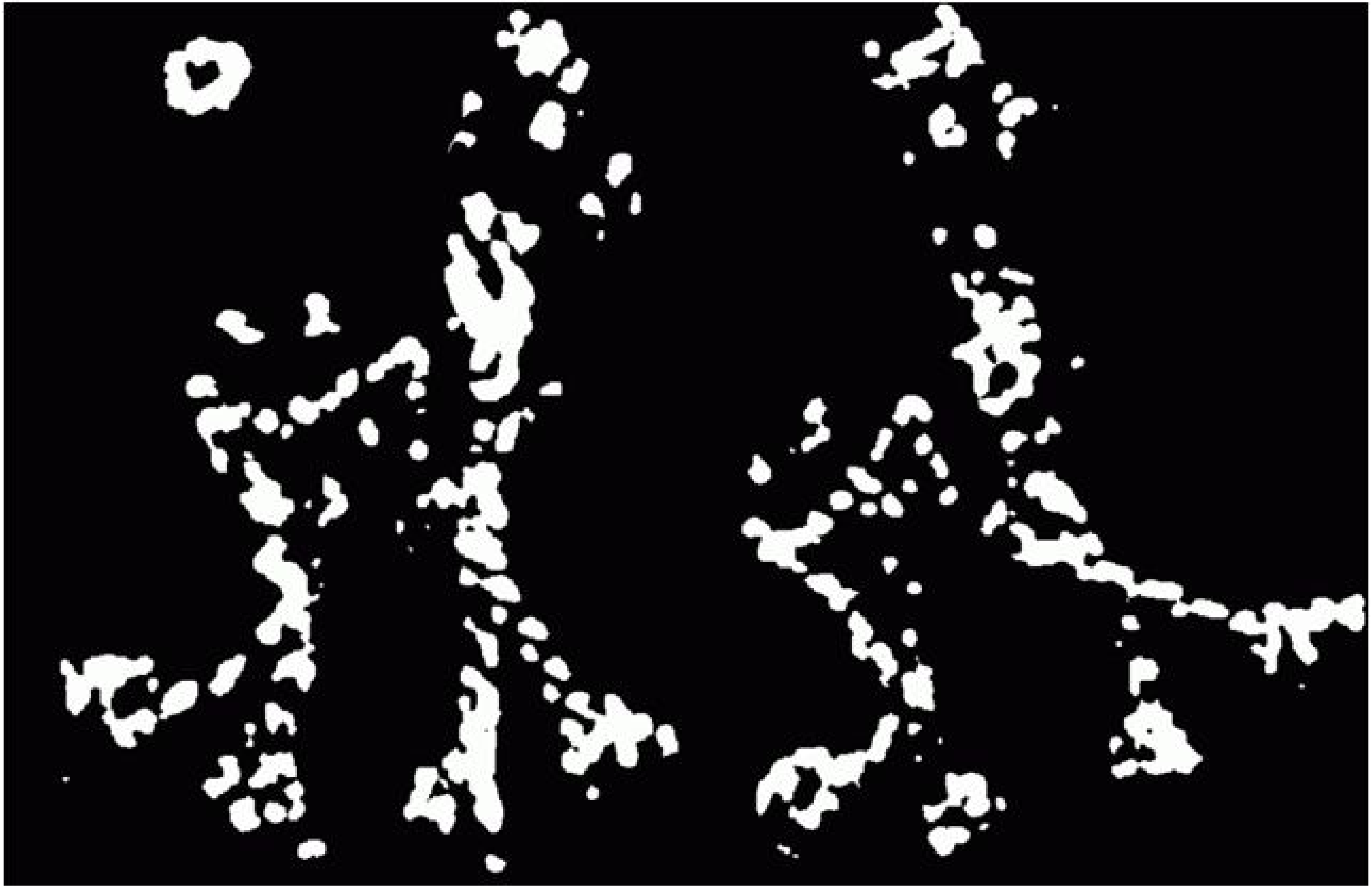
Демонстрация по шагам

Вычисление отклика угла *R*

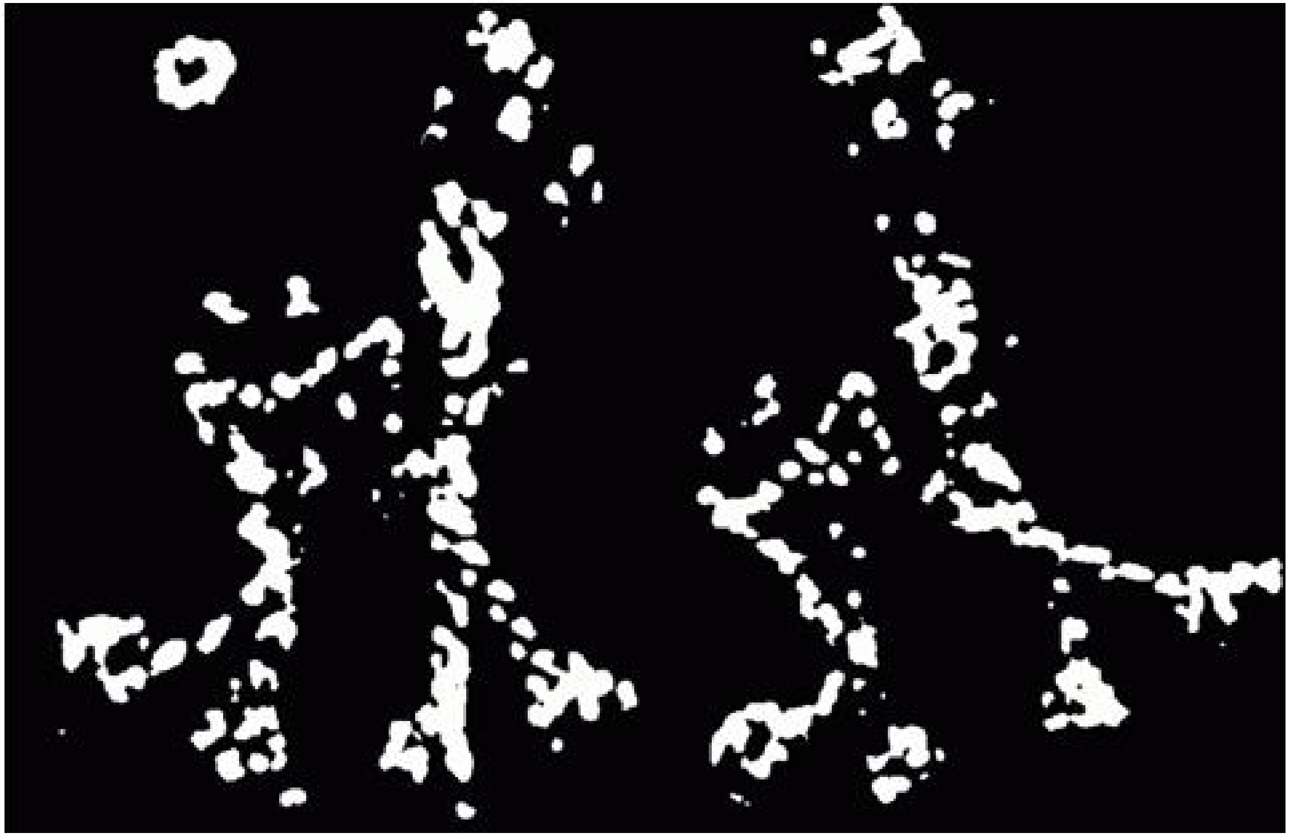


Демонстрация по шагам

Найдём точки с большим откликом R>порог



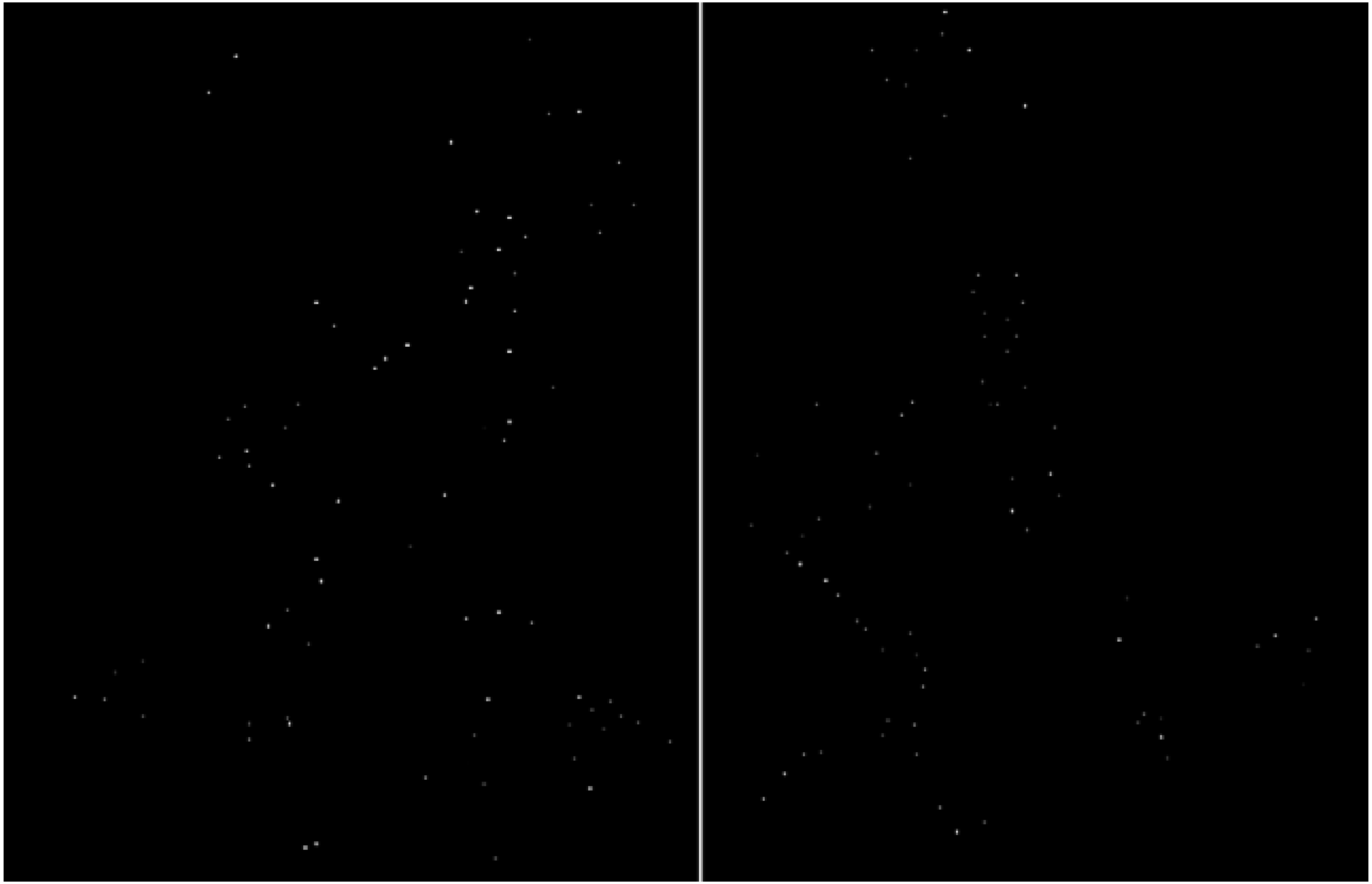
Демонстрация по шагам



- Как быть с тем, что функция отклика угла больше порога в некоторых областях?
- Как нам выбрать конкретные точки в областях?

Демонстрация по шагам

Оставим только точки локальных максимумов R



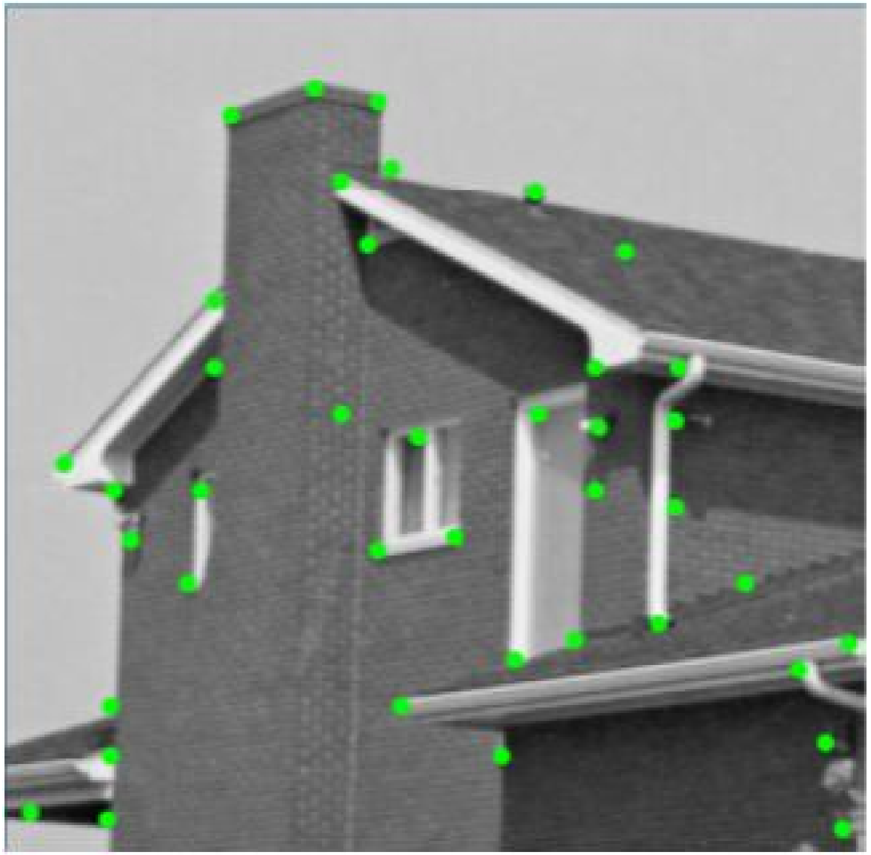
Демонстрация по шагам

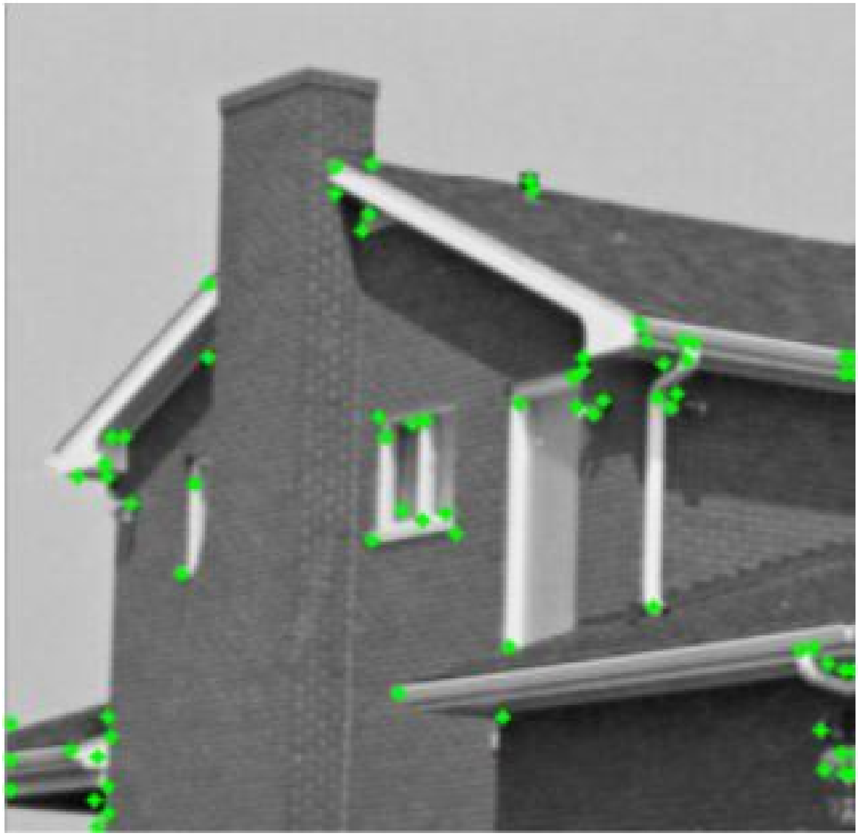
Алгоритм детектора Харриса

1. Вычислить градиент изображения в каждом пикселе
   - С использованием гауссова сглаживания
2. Вычислить матрицу вторых моментов M по окну вокруг каждого пикселя
3. Вычислить отклик угла *R*
4. Отсечение по порогу *R*
5. Найти локальные максимумы функции отклика (non-maximum suppression) по окрестности заданного радиуса
6. Выбор N самых сильных локальных максимумов

Slide by S. Lazebnik

Результат работы детектора

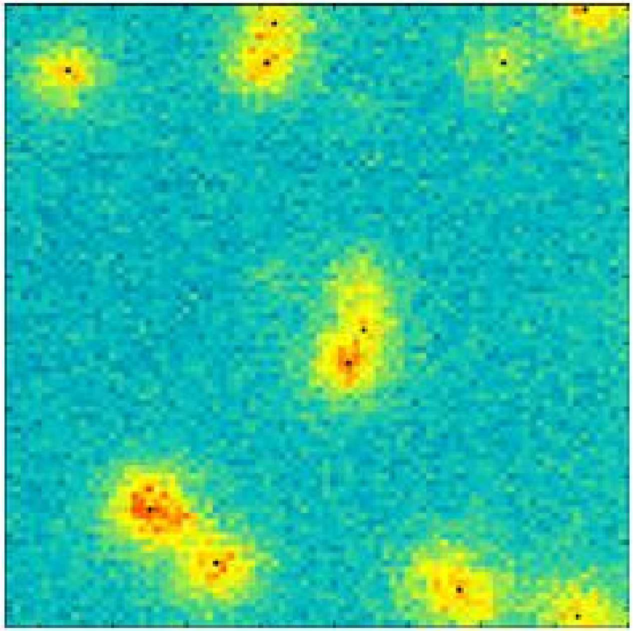




детектор Фёрстнера

детектор Харриса

Важный вывод

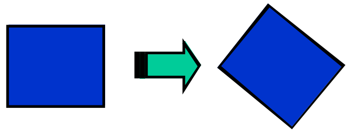


Детектор – суть некоторая функция, локальные максимумы которой мы используем в качестве особенностей изображений

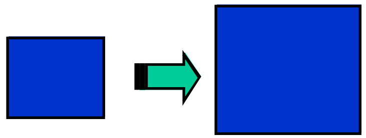
Инвариантность

Что у детектора Харриса с инвариантностью?

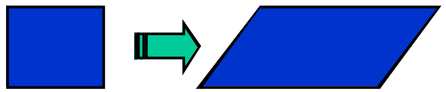
- Поворот



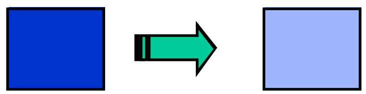
- Масштаб



- Аффинное



- Аффинное изменение яркости ($I \to a\,I + b$)

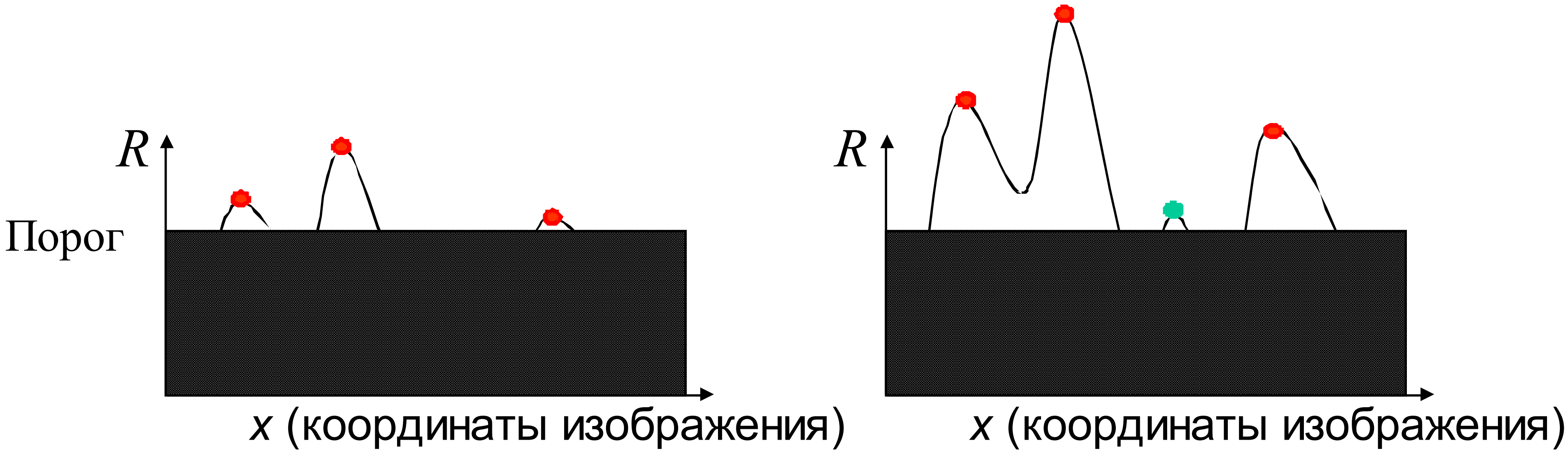


Детекторы Харриса

- Частичная инвариантность к изменению освещенности

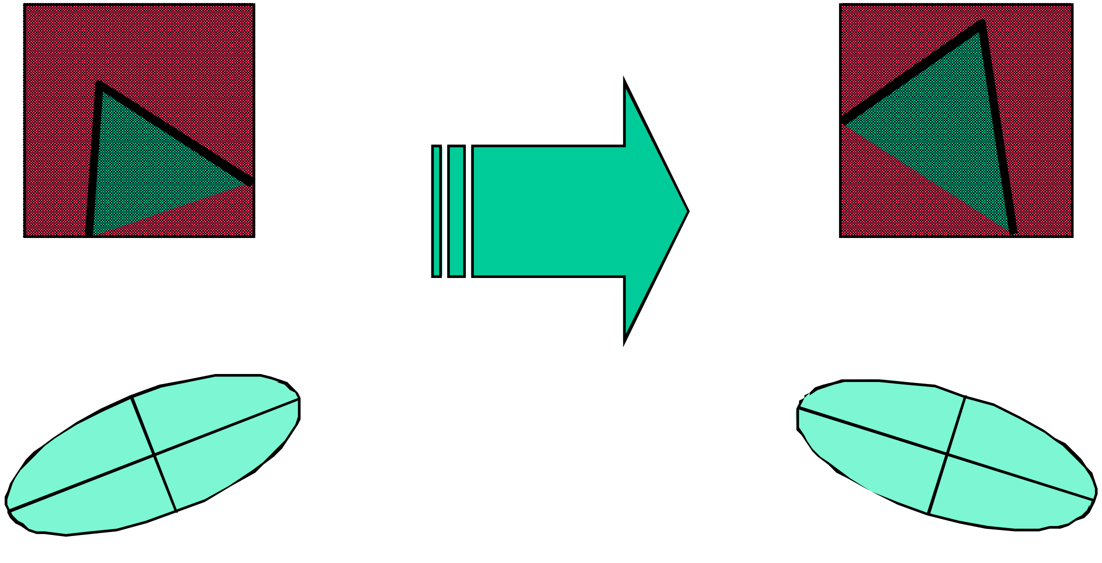
✓Используются только производные
=> инвариантность к сдвигу $I \to I + b$

✓Масштабирование: $I \to a\, I$



Детектор Харриса

Инвариантность к вращению изображения:



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

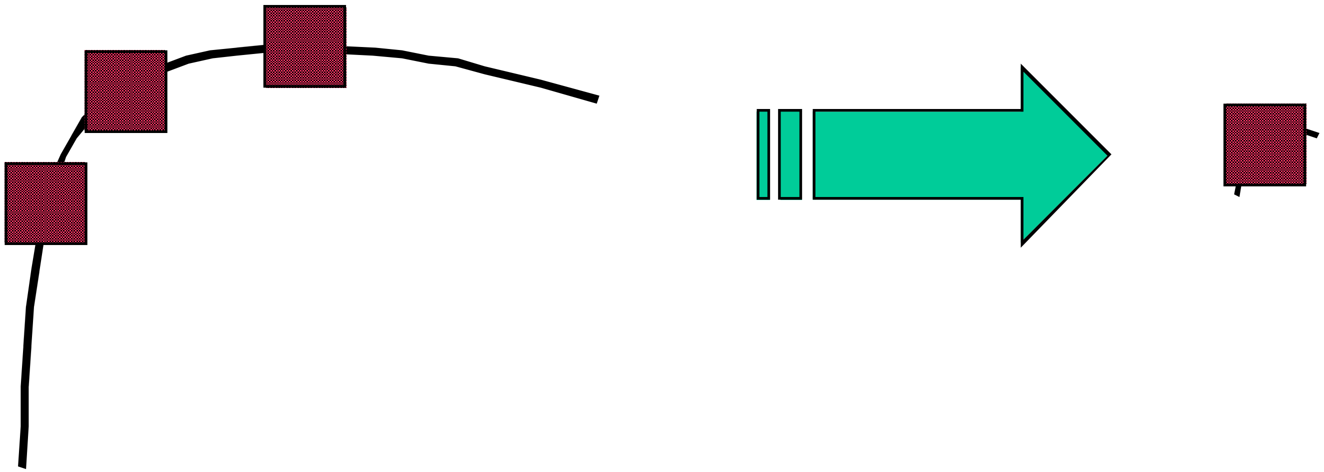
Эллипс вращается, но его форма (собственные значения) остаются неизменными

Отклик угла R инвариантен относительно вращению изображения

Slide by S. Lazebnik

Масштабирование?

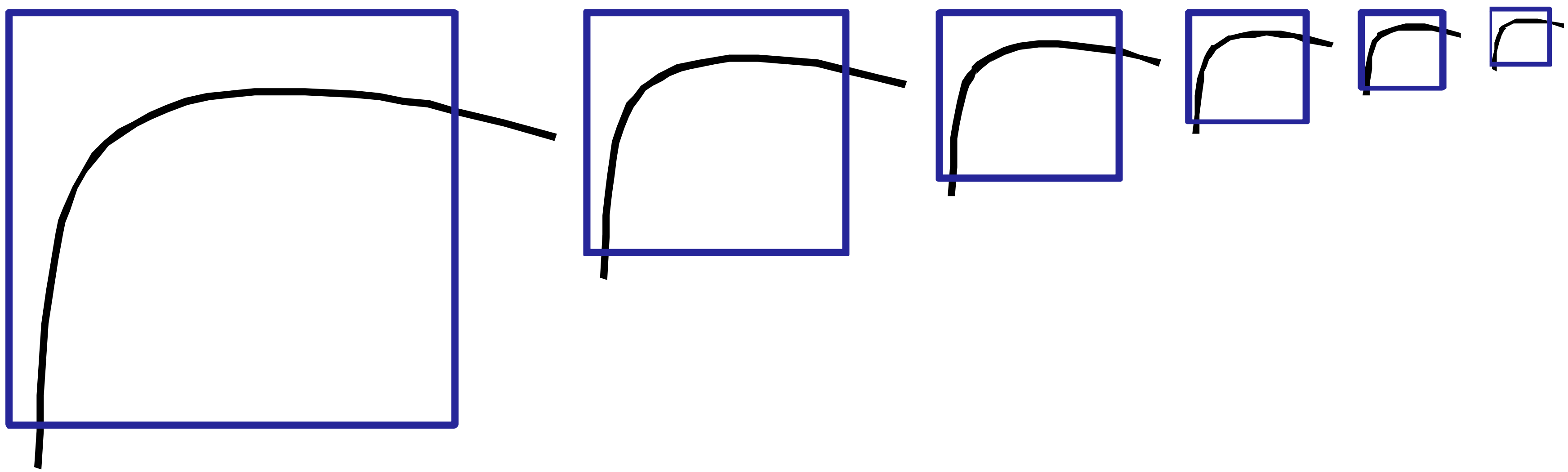
- Угол или нет? - Зависит от масштаба изображения!



Все эти точки будут помечены как ***края***

Угол !

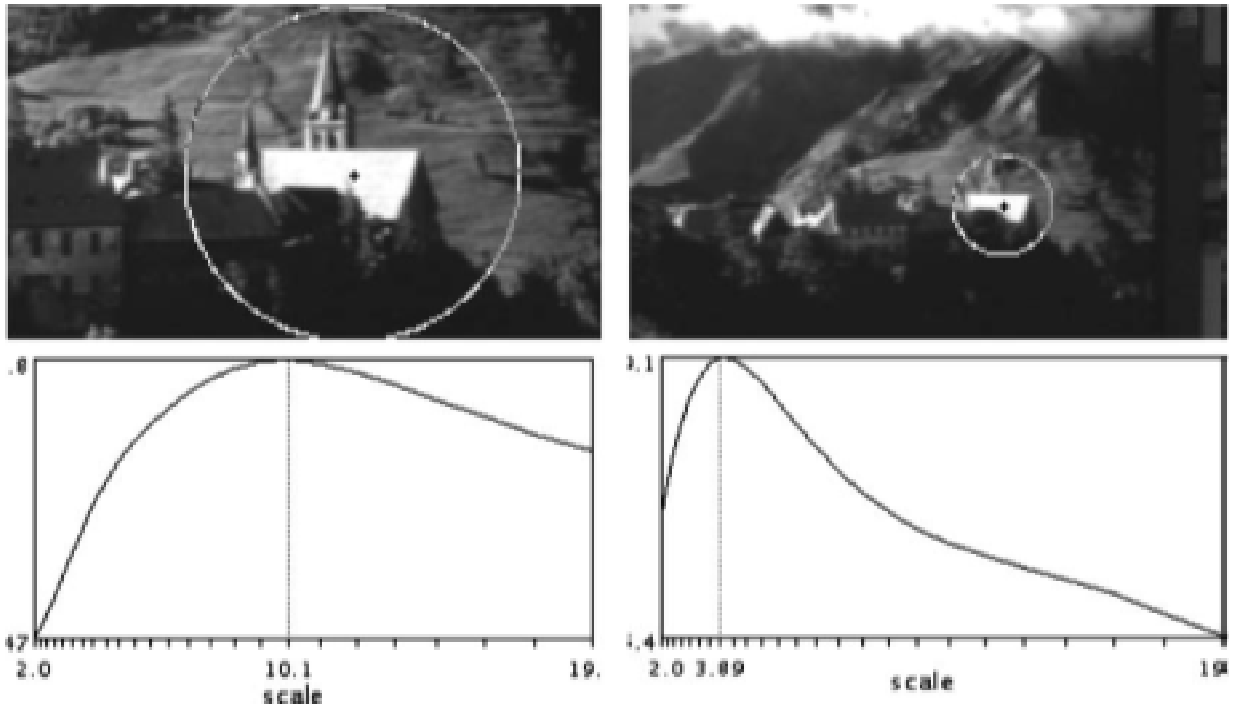
Характерный масштаб



- С какого момента фрагмент считается «углом»?
- Если наш детектор Харриса на нескольких соседних масштабах пометит точку как угол, то как нам быть?

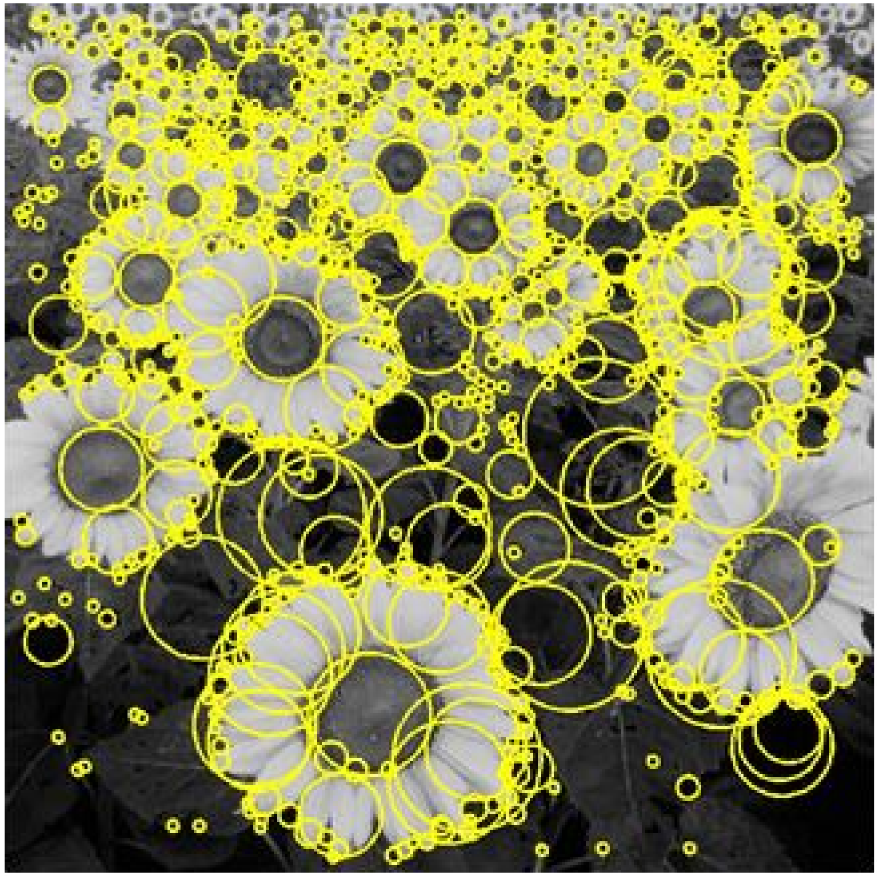
Инвариантность к масштабированию

- Цель: определять размер окрестности особой точки в масштабированных версиях одного и того же изображения
- Требуется метод выбора размера характеристической окрестности



Slide by S. Lazebnik

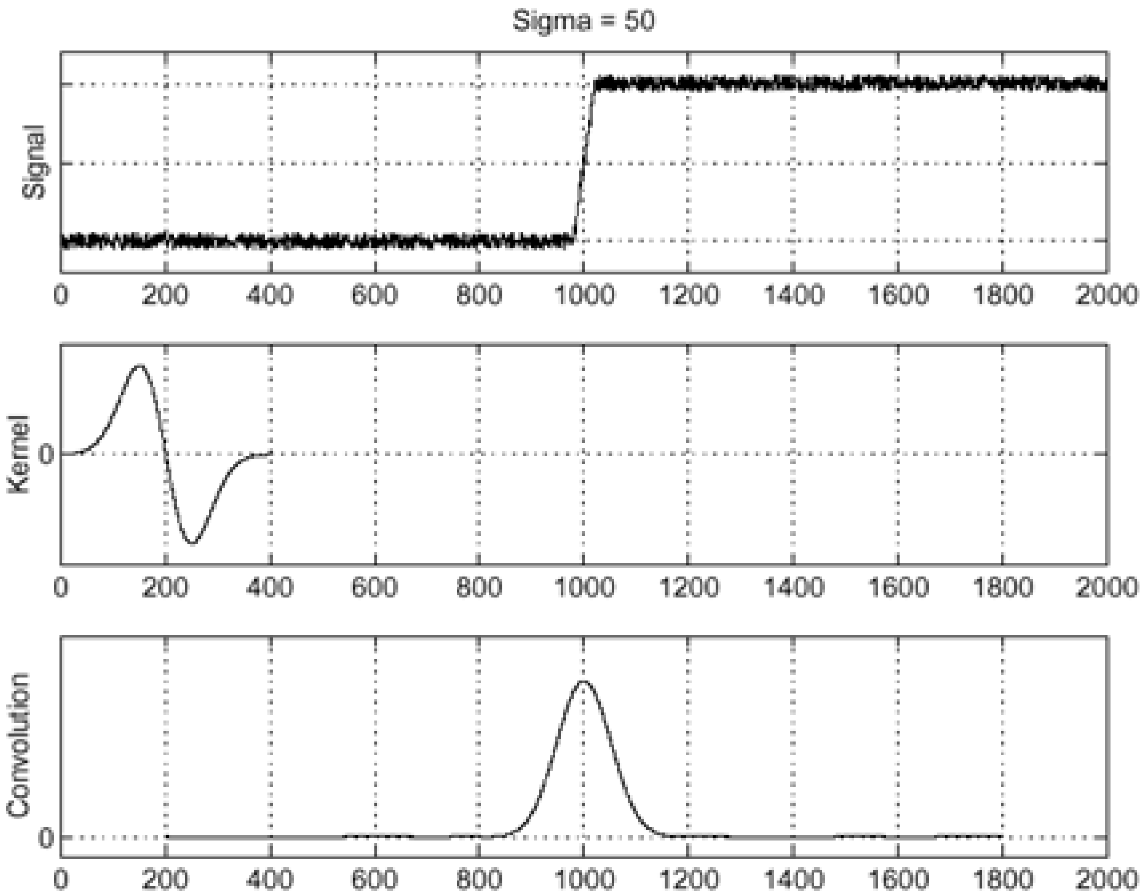
Блобы



«Пятна», «Капля», «Blob» - вначале для особенностей такого типа была разработана теория выбора характерного размера

Slide by S. Lazebnik

Поиск краев



$$f$$

Край

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

Производная гауссианы

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

Край = максимум производной

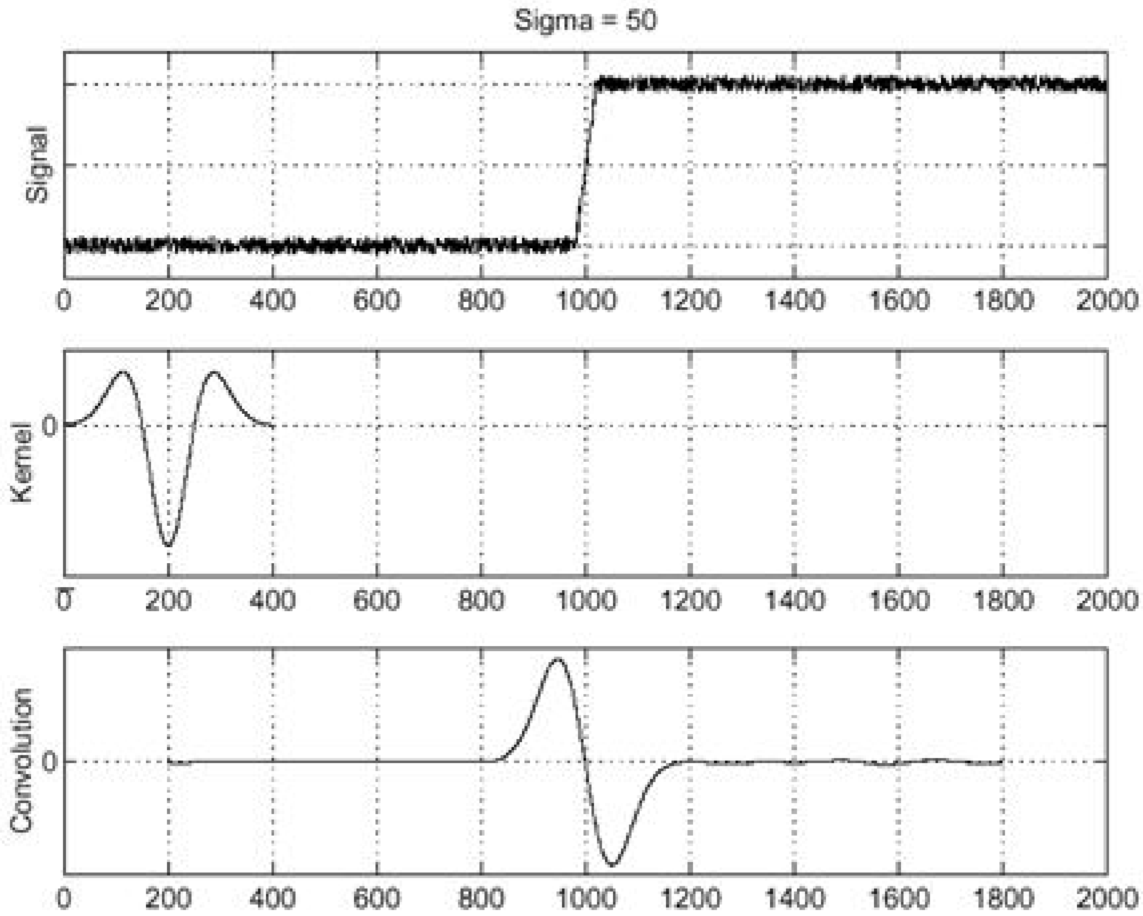
Source: S. Seitz

Второй проход

f

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

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



Край

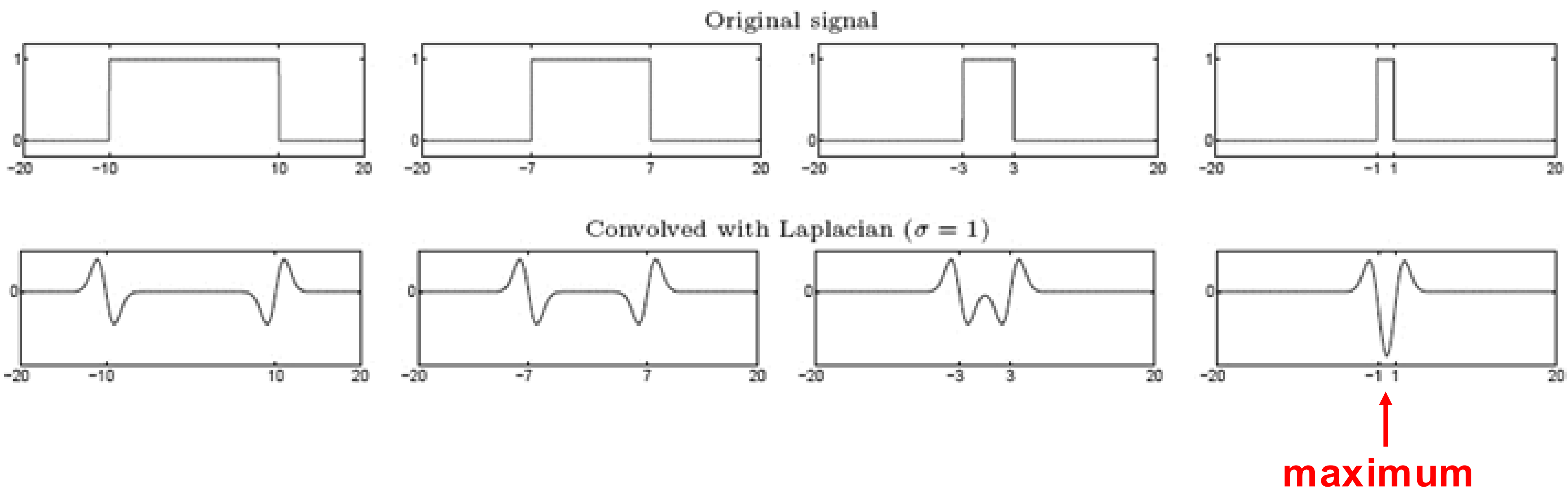
Вторая производная Гауссианы (Лапласиан)

Край = переход через ноль второй производной

Source: S. Seitz

От поиска краев к поиску блобов

- Край = «всплеск»
- Блоб = совмещение двух «всплесков»



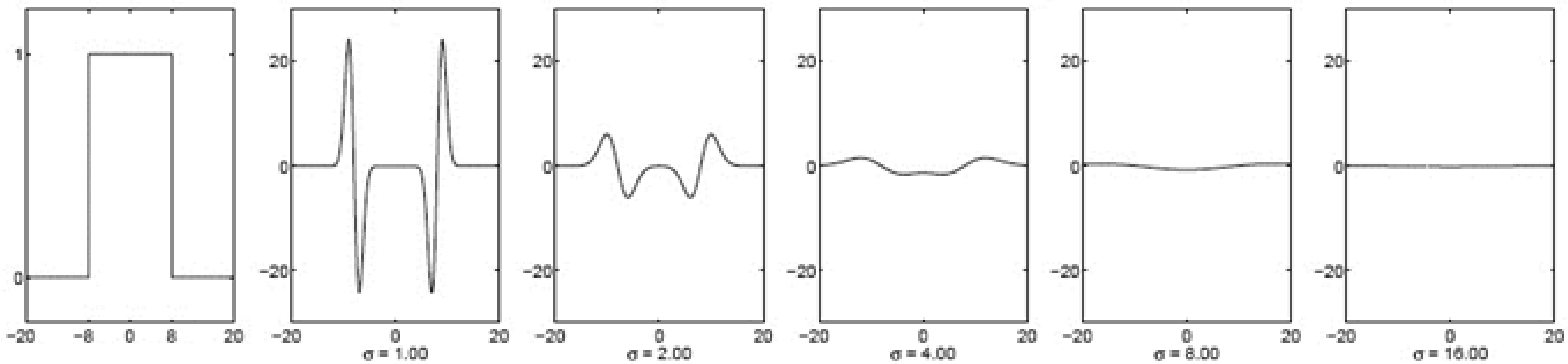
Выбор масштаба: величина отклика лапласиана Гауссиана достигает максимума в центре блоба в том случае, если размер лапласиана «соответствует» размеру блоба

Slide by S. Lazebnik

Выбор масштаба

- Нужно найти характеристический размер блоба путем свертки с Лапласианом в нескольких масштабах и найти максимальные отклики
- Однако, отклик Лапласиана затухает при увеличении масштаба:

Ненормализованный отклик Лапласиана



Исходный сигнал (радиус=8)

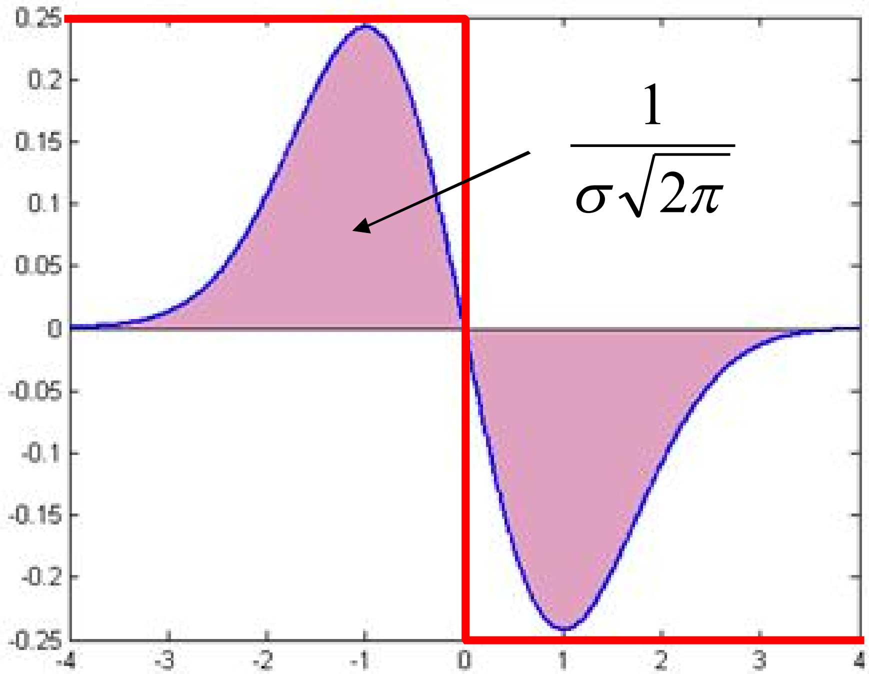
увеличение σ

Почему так происходит?

Slide by S. Lazebnik

Нормализация масштаба

- Отклик производной фильтра Гаусса на идеальный край затухает с увеличением масштаба σ



Slide by S. Lazebnik

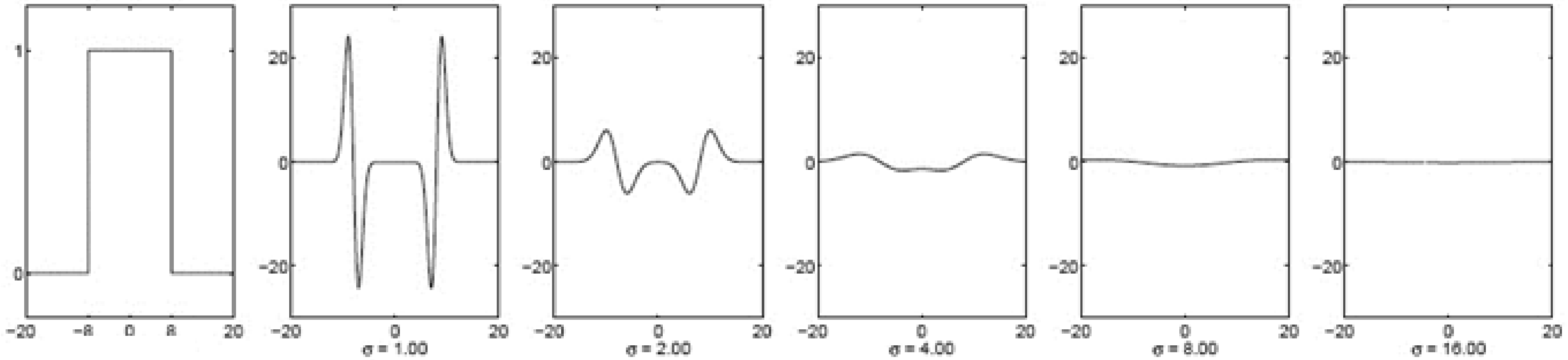
Нормализация масштаба

- Отклик производной фильтра Гаусса на идеальный край затухает при увеличении σ
- Нужно домножить производную на σ для достижения инвариантности к масштабу
- Лапласиан это вторая производная фильтра гаусса, поэтому домножаем на σ^2

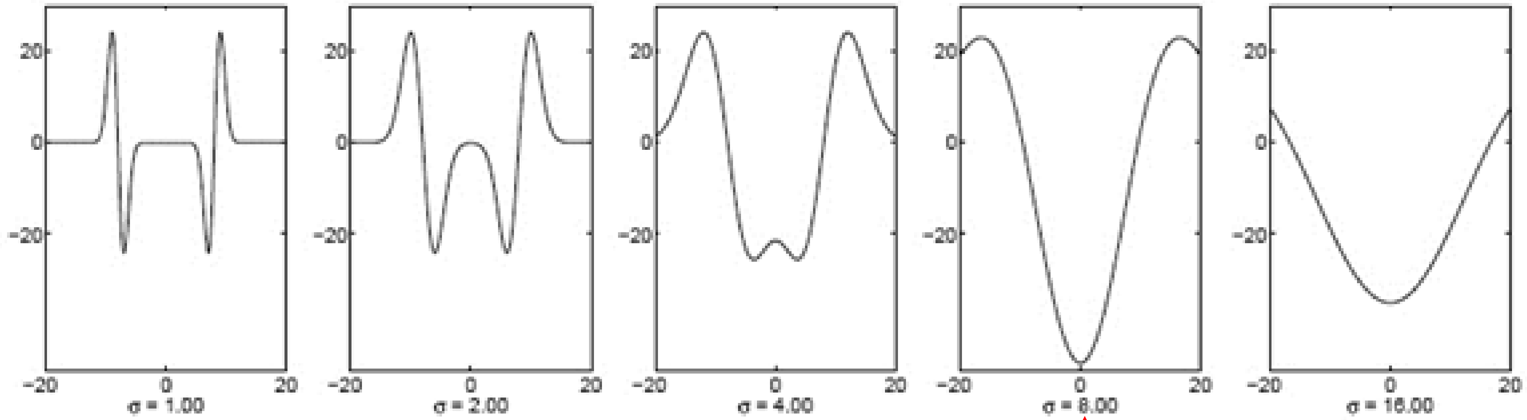
Эффект нормализации

Исходный сигнал

Ненормализованный отклик Лапласиана



Нормализованный по масштабу отклик Лапласиана

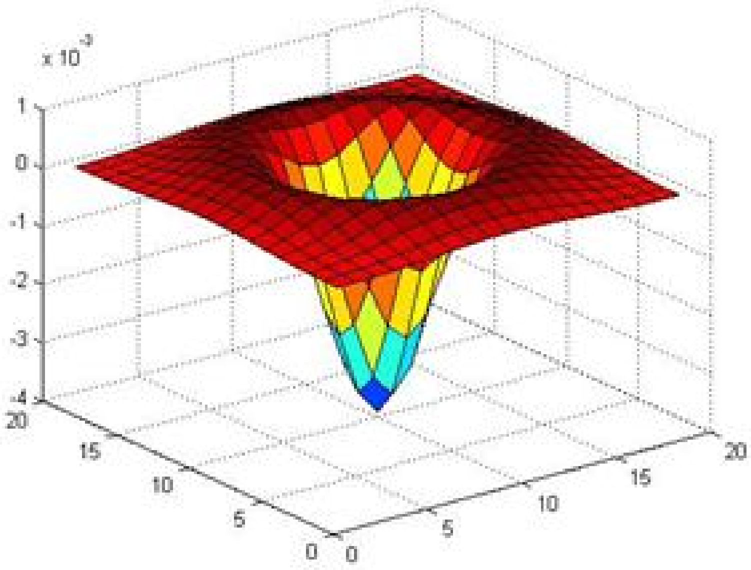


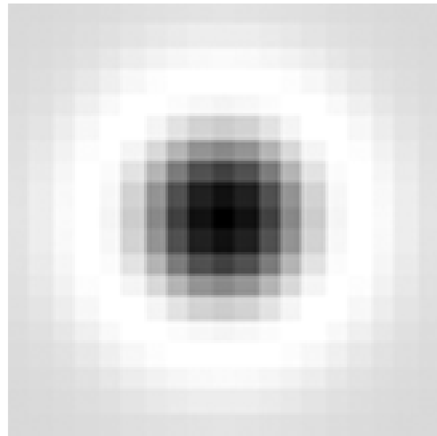
максимум

Slide by S. Lazebnik

Поиск блобов в 2D

Лапласиан Гауссиана: Центрально-симметричный оператор поиска блобов в 2D



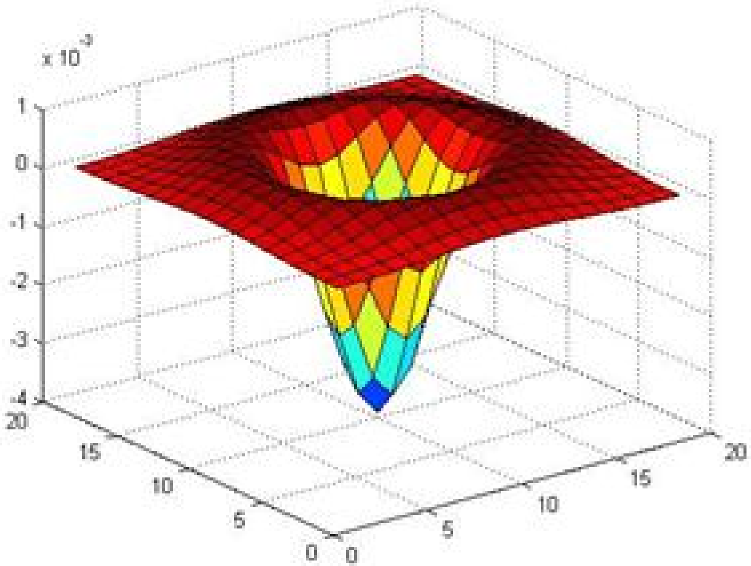


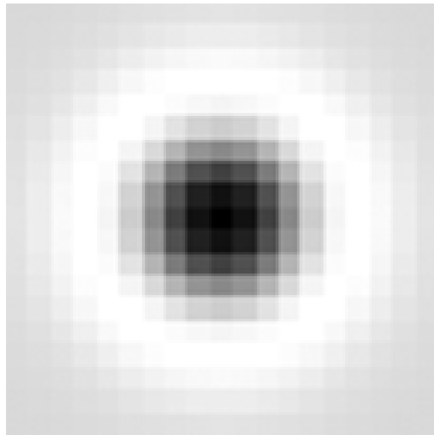
$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

Slide by S. Lazebnik

Поиск блобов в 2D

Лапласиан Гауссиана: Центрально-симметричный оператор поиска блобов в 2D



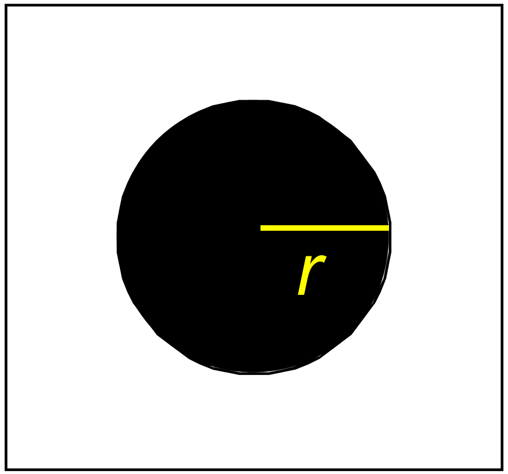


Нормализация: $$\nabla^2_{\text{norm}} g = \sigma^2 \left(\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} \right)$$

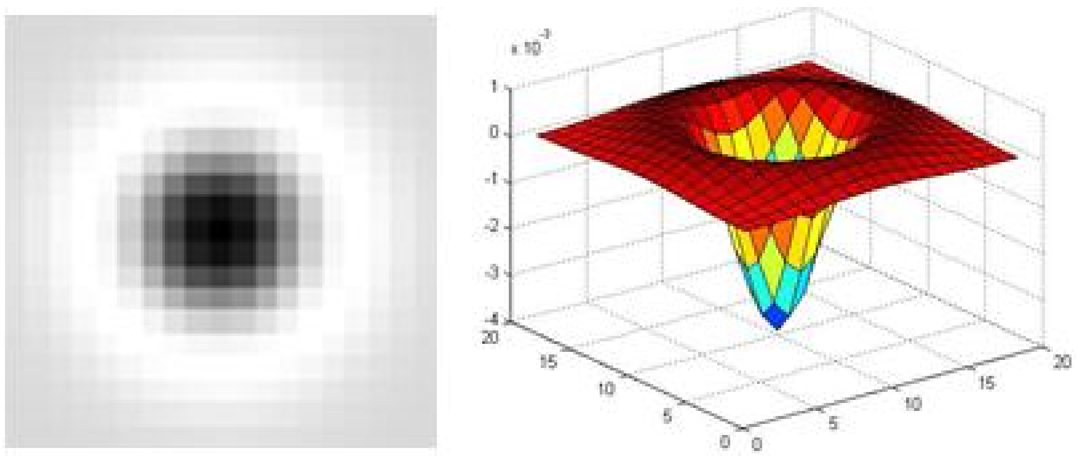
Slide by S. Lazebnik

Выбор масштаба

- На каком масштабе Лапласиан достигает максимума отклика на бинарный круг радиуса r?



изображение



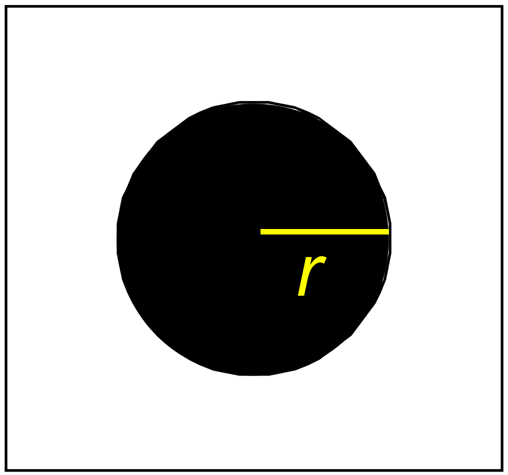
Лапласиан

Выбор масштаба

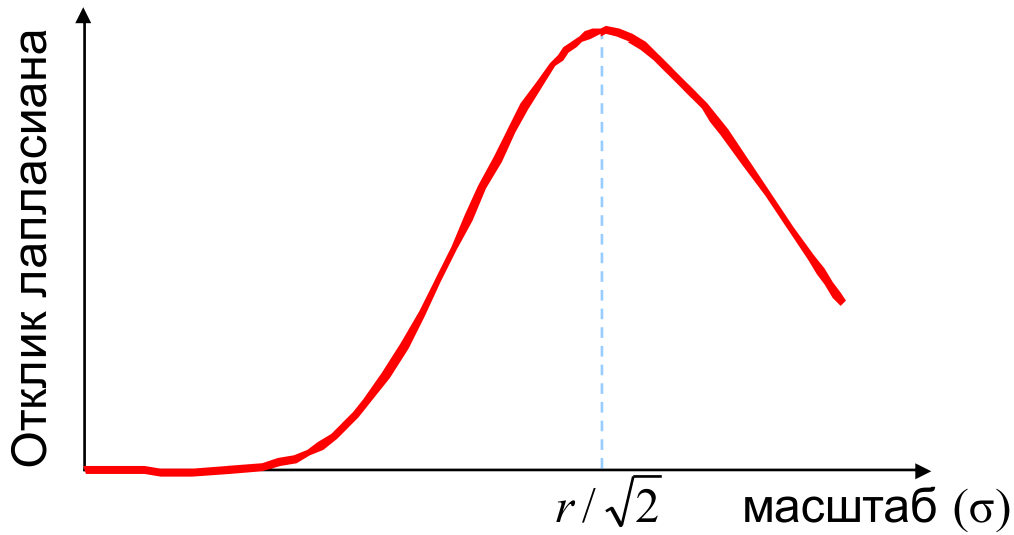
- 2D Лапласиан задается формулой:

$$(x^2 + y^2 - 2\sigma^2)\, e^{-(x^2+y^2)/2\sigma^2}$$ (с точностью до масштаба)

- Для бинарного круга радиуса r, Лапласиан достигает максимума в $\sigma = r/\sqrt{2}$



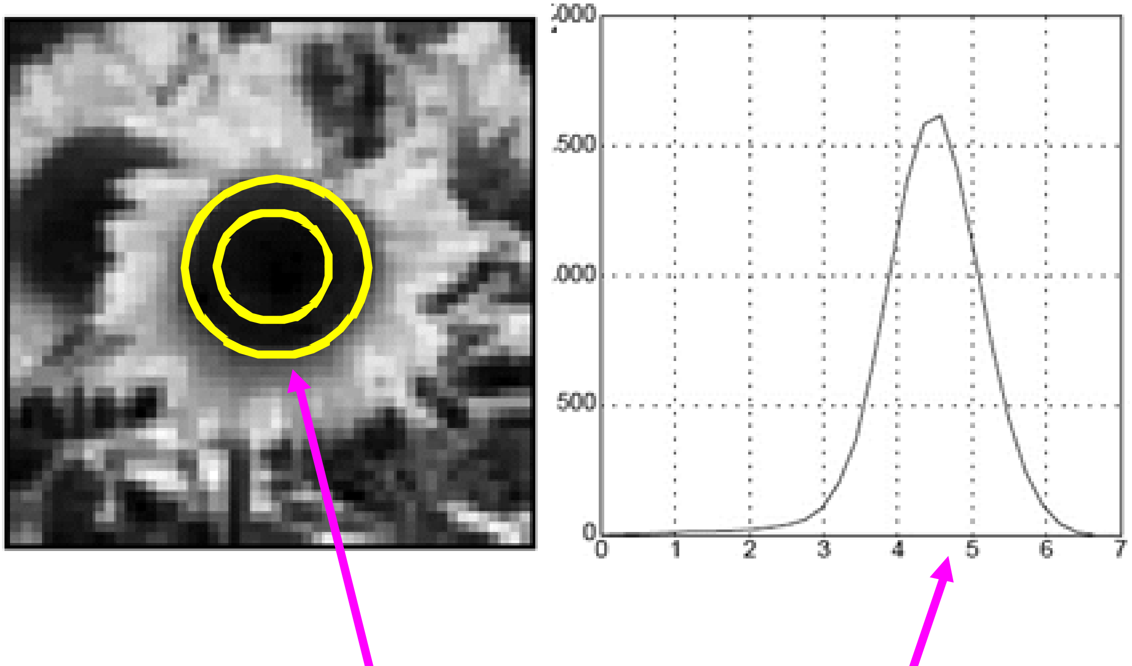
изображение



Slide by S. Lazebnik

Характеристический размер

- Характеристический размер определяется как масштаб, на котором достигается максимум отклика Лапласиана

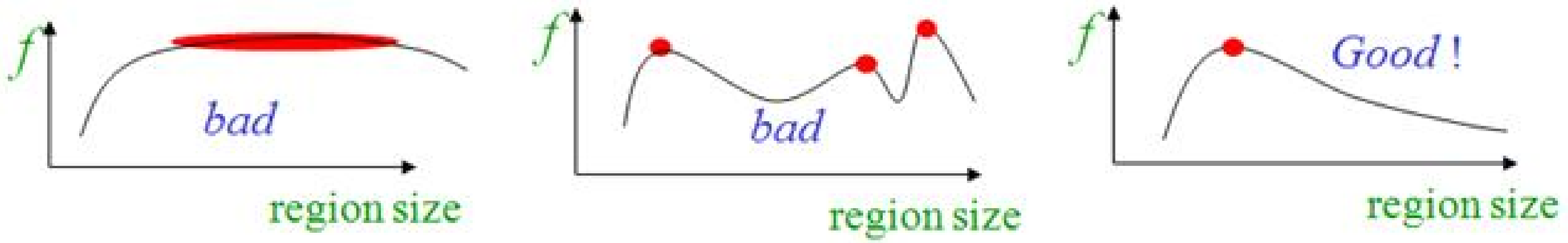


Характеристический масштаб

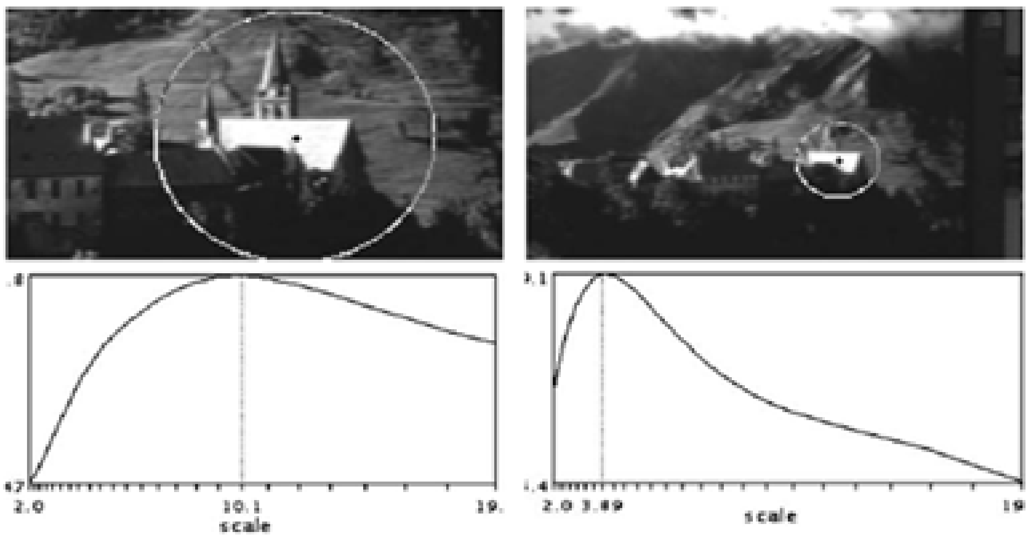
T. Lindeberg (1998). "Feature detection with automatic scale selection." *International Journal of Computer Vision* **30** (2): pp 77--116.

Slide by S. Lazebnik

Характеристический размер



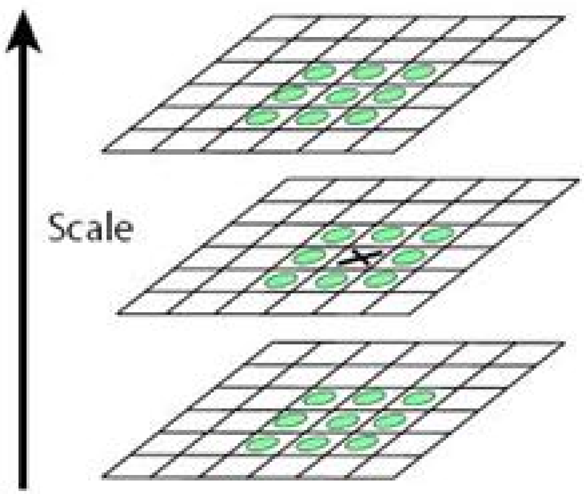
У «хорошего блоба»– один ярко выраженный пик функции



Slide by S. Lazebnik

Многомасштабный детектор блобов

1. Свертываем изображение нормализованным фильтром Лапласианом на разных масштабах
2. Ищем максимум отклика Лапласиана в 3D



Slide by S. Lazebnik

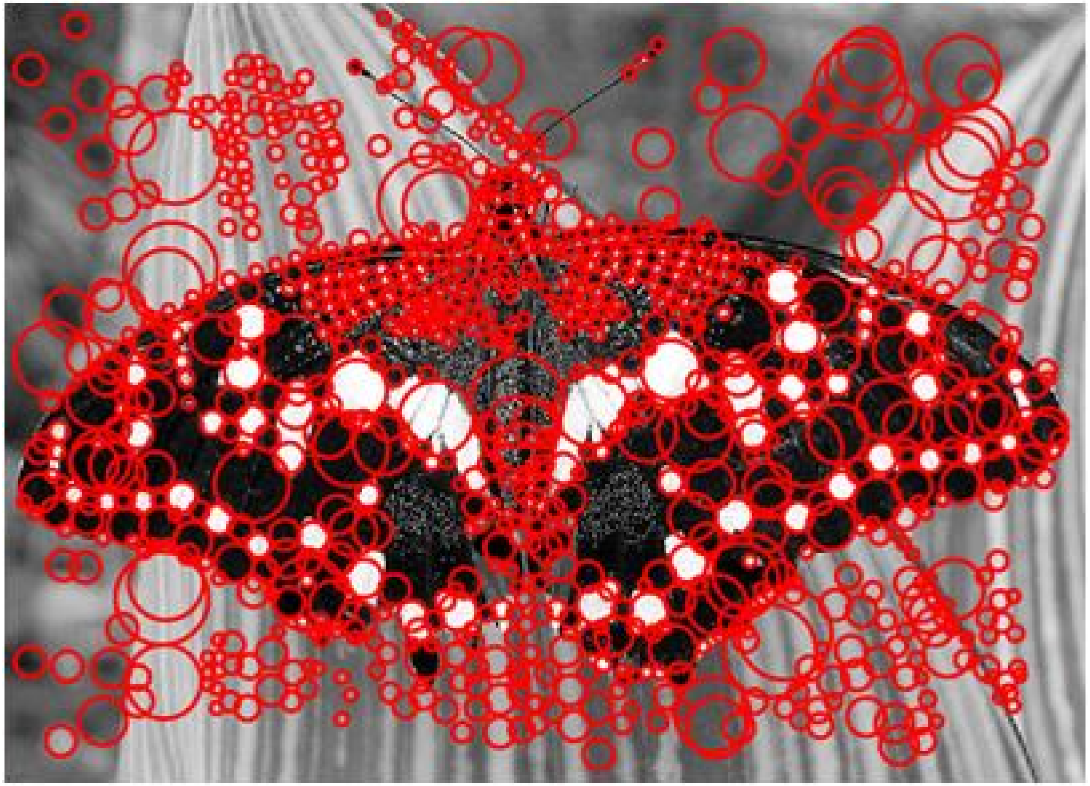
Пример

Slide by S. Lazebnik

Пример

sigma = 11.9912

Пример



Slide by S. Lazebnik

Эффективная реализация (DoG)

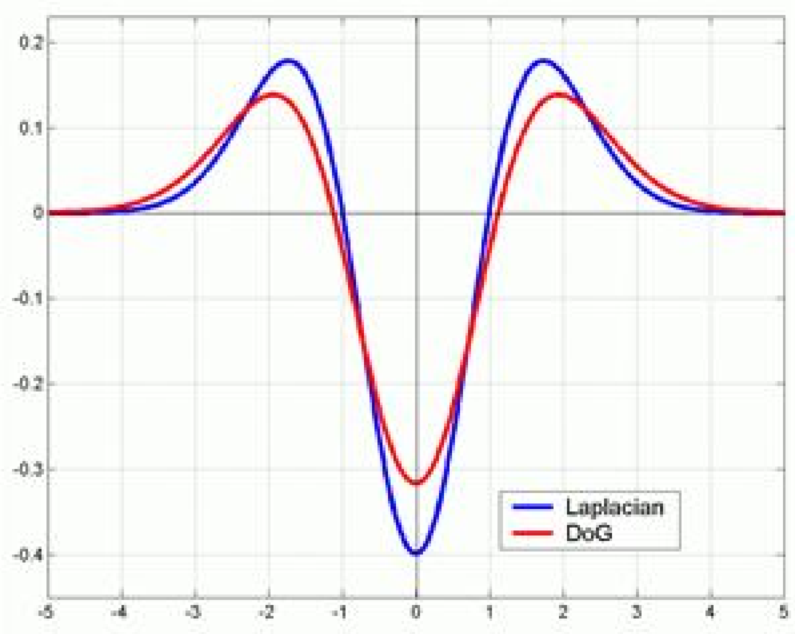
Приближение Лапласиана с помощью разницы гауссиан:

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

(Лапласиан)

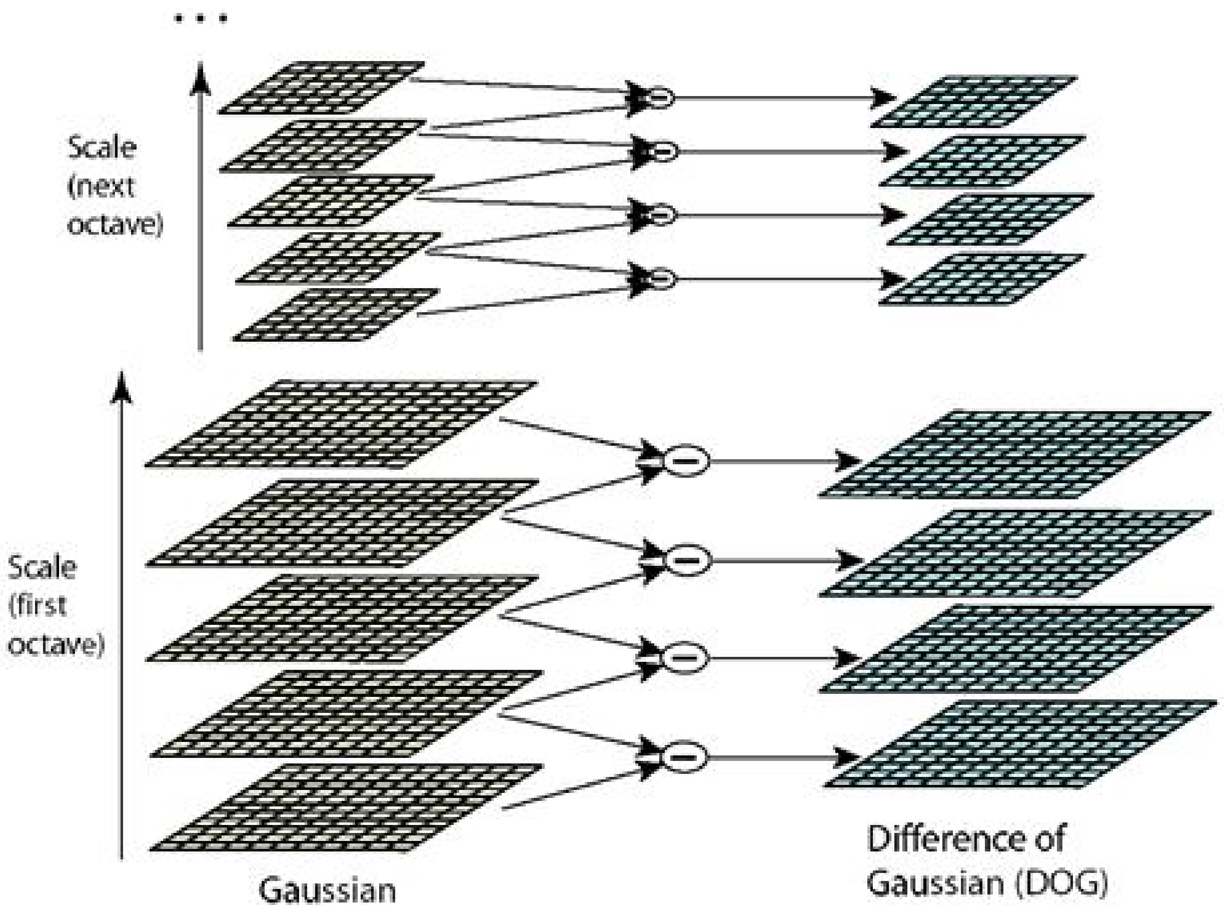
$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

(Разница Гауссиан)



Difference of Gaussian = DoG

Эффективная реализация (DoG)



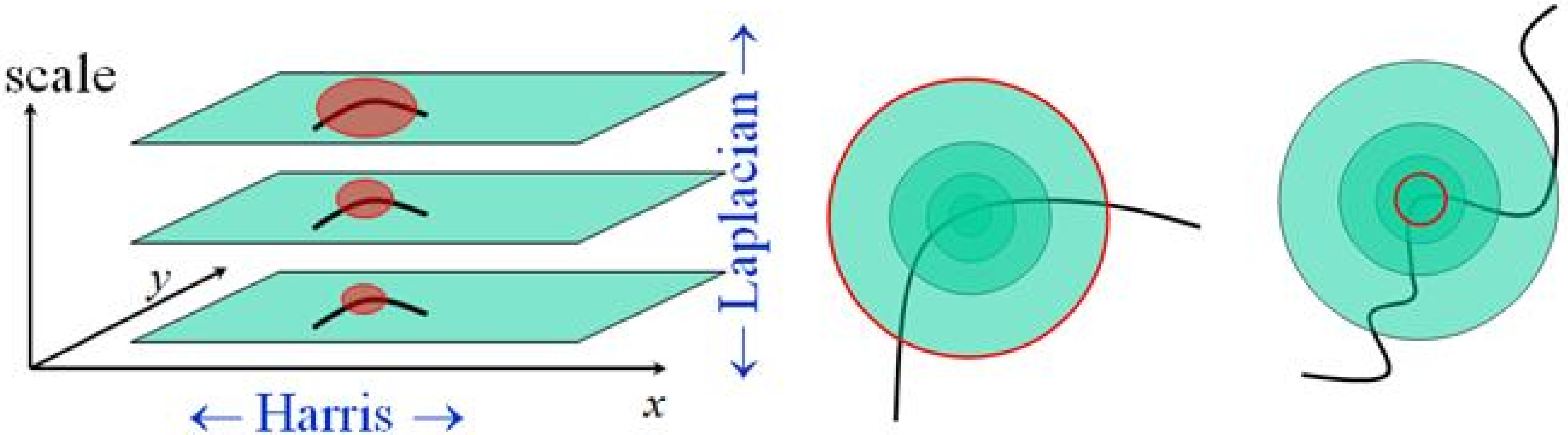
Детектор DoG также выделяет «блобы» на изображении

David G. Lowe. **"Distinctive image features from scale-invariant keypoints."**
IJCV 60 (2), pp. 91-110, 2004.

Slide by S. Lazebnik

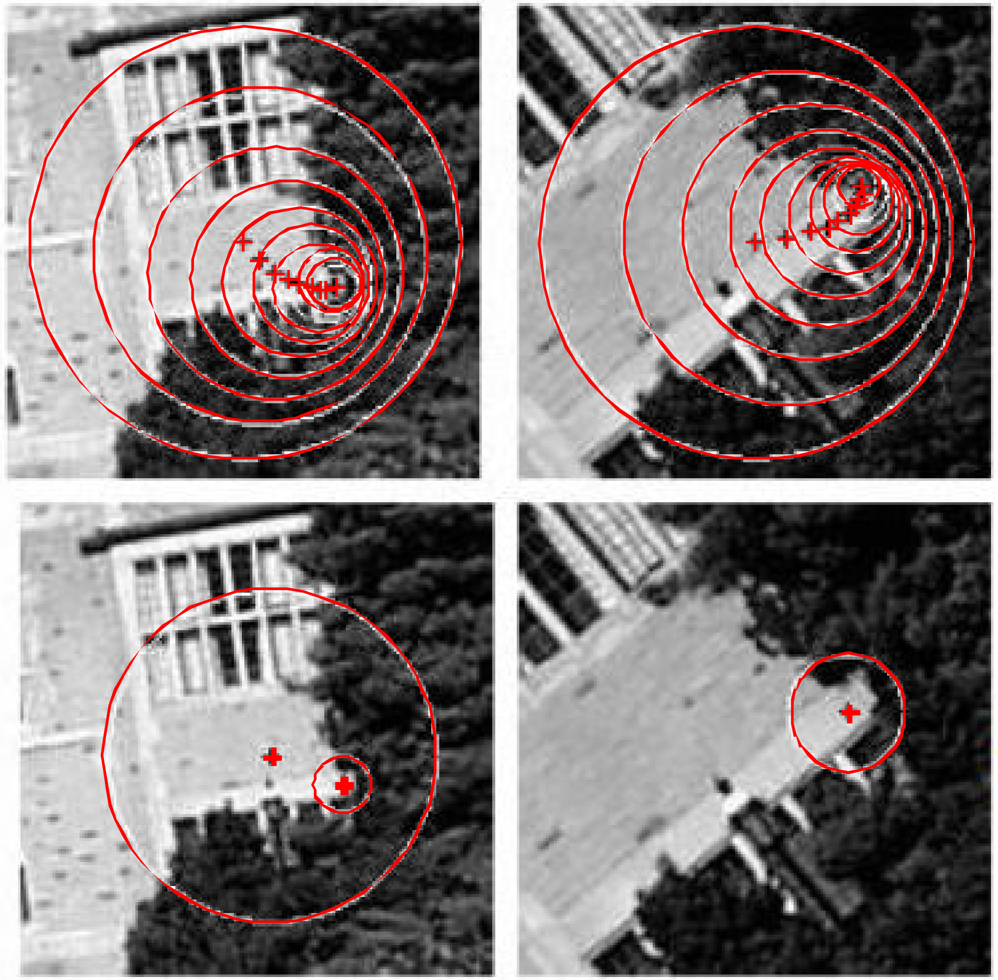
Детектор Harris-Laplacian

- Выделяем углы на изображении, но с характеристическим размером
- Максимизация:
  - По изображению – откликов углов Харриса
  - По масштабу – Лапласиана
- Разные варианты чередования вычисления функции Харриса и Лапласиана



Сравнение

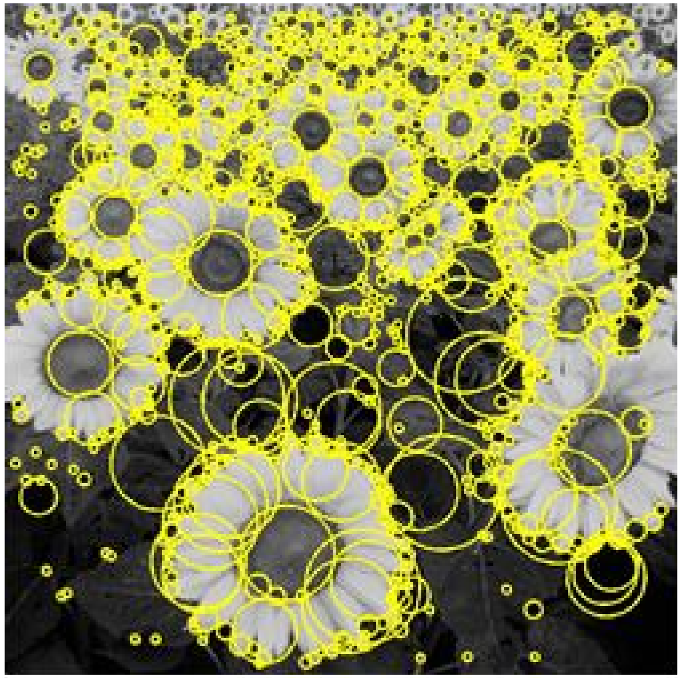
Сравнение простого детектора Харриса и Харрис-Лапласиана

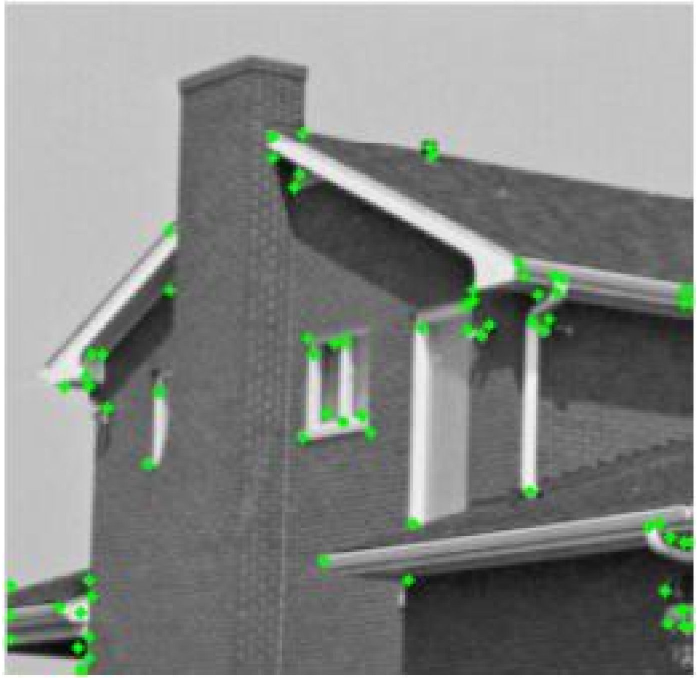


Харрис

Харрис-Лаплас

Углы и блобы

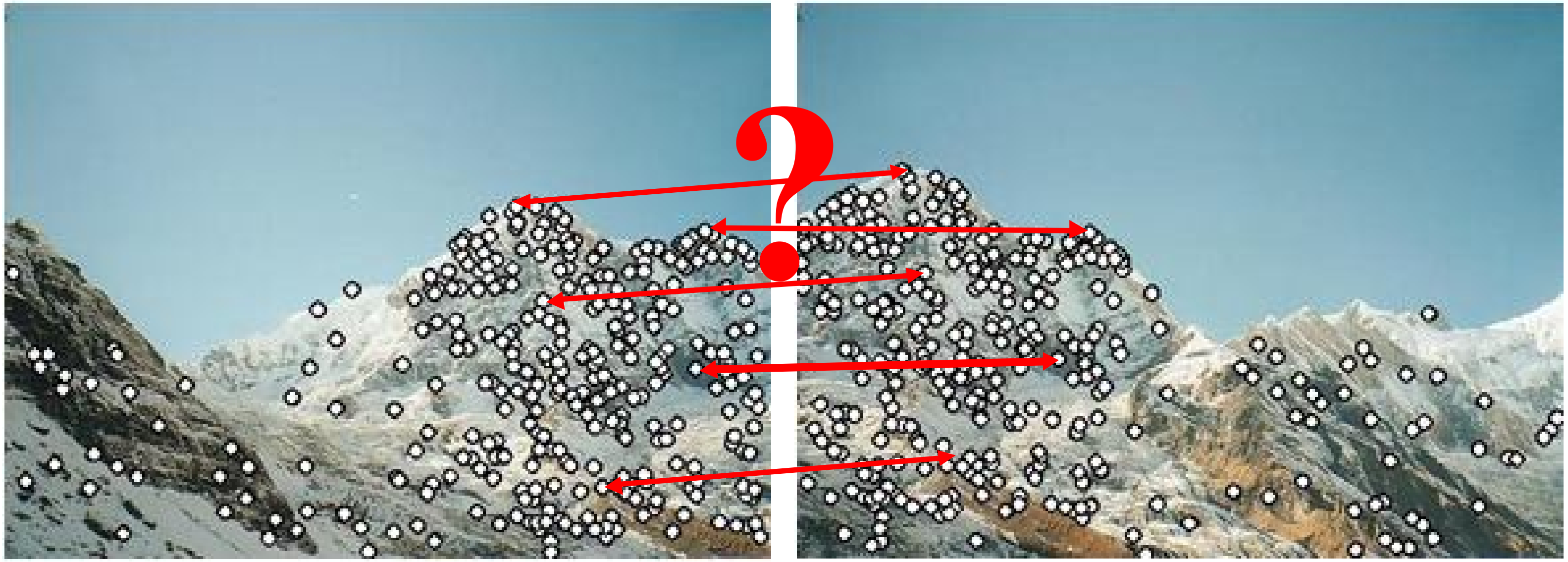




- Углы и блобы – разные виды локальных особенностей
- Детекторы Харрис-Лапласиан и LoG (DoG) выделяют разные множества особенностей
- Можно применять их одновременно

Дескрипторы

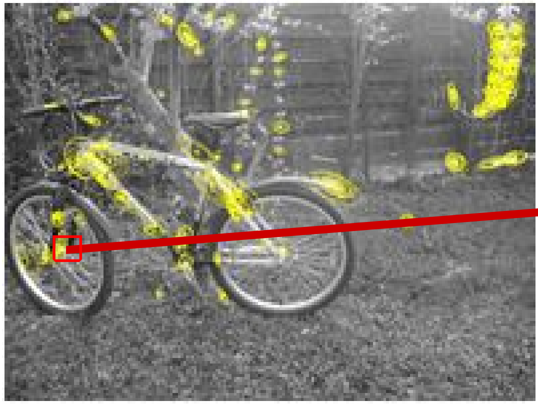
Точки найдены – как их сопоставить?

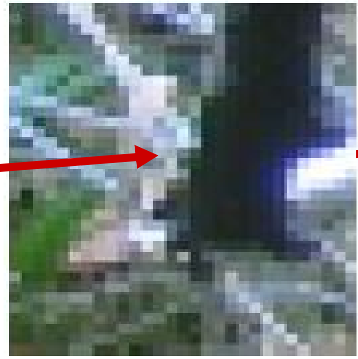


- Нужно как-то описать каждую точку, чтобы можно было отличать одну от другой!
- Дескриптор (Descriptor) - вектор признак окрестности точки

Дескрипторы

Необходимо каждую интересную точку описать набором параметров:



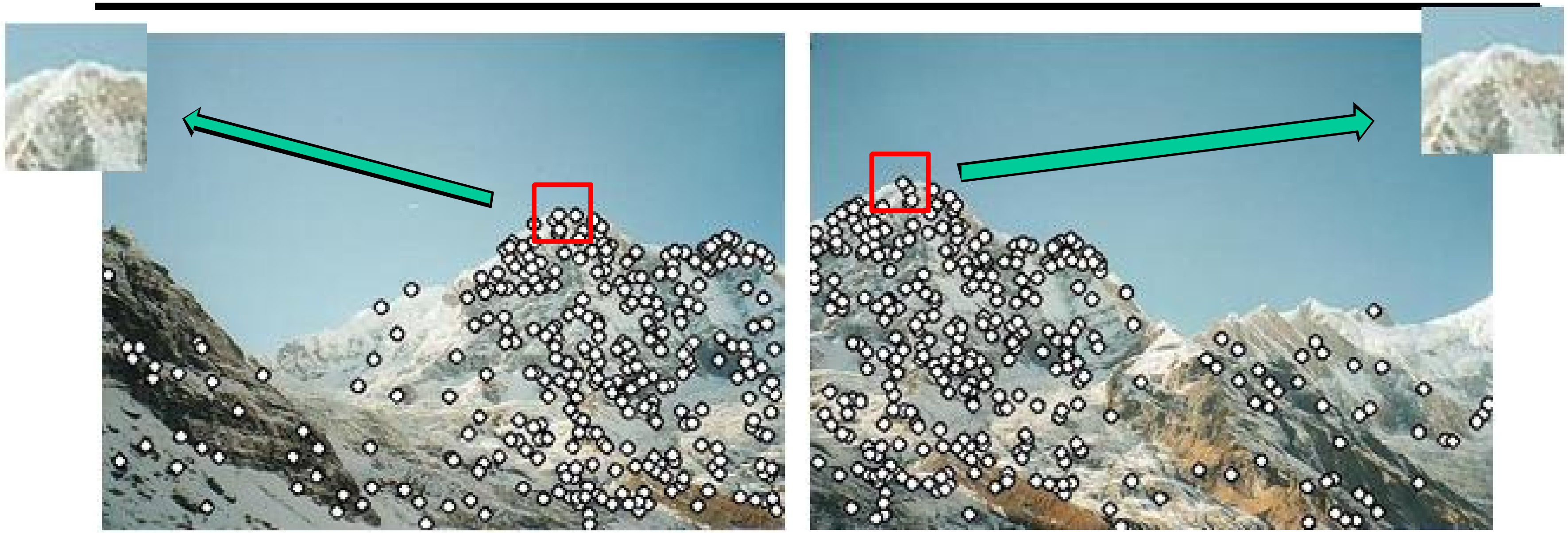


$$\mathbf{f}_n = (f_{n,1}, \cdots, f_{n,j})^T$$

Как будем поступать:

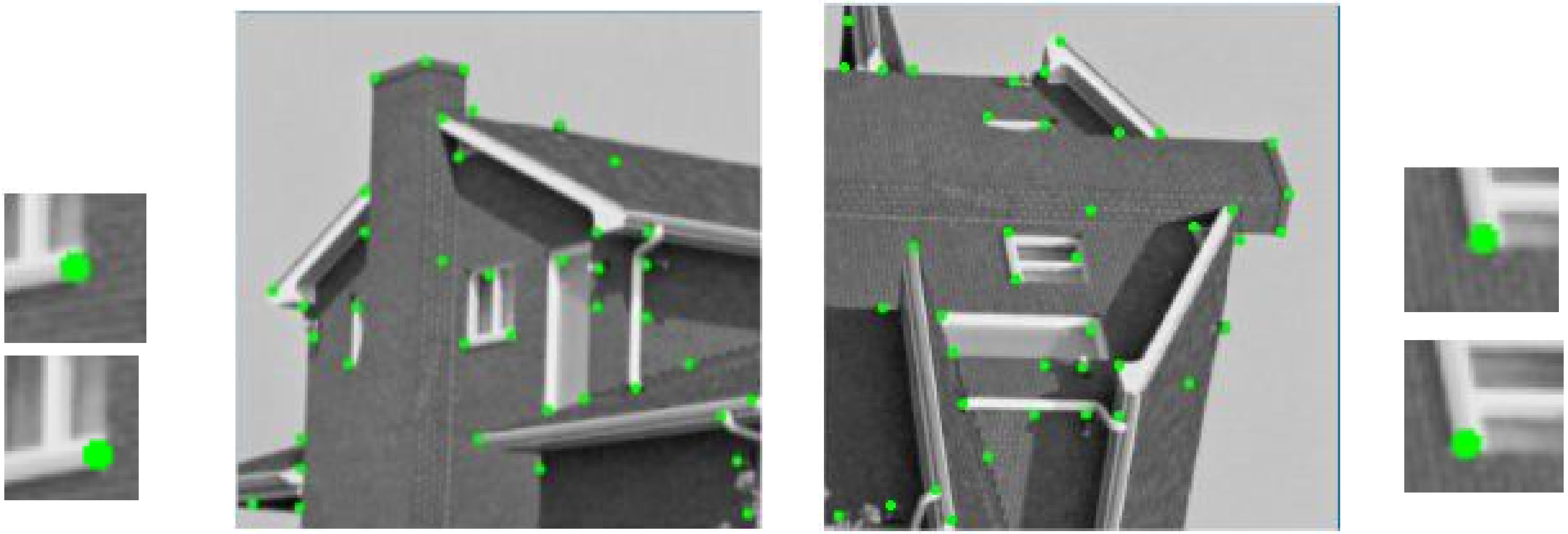
- Возьмём окрестность точки
  - Какой формы?
  - Какого размера?
- Вычислим по окрестности набор признаков
  - Какие?

Простейший подход



- Возьмём квадратные окрестности, со сторонами, параллельными строкам и столбцами изображения
- Яркости пикселов будут признаками
- Сравнивать будем как два изображения - попиксельно (SAD, SSD)
- Такая окрестность инвариантна только к сдвигу изображения

Недостаток простой окрестности



- Детектор точек инвариантен к повороту, а окрестность нет
- Небольшие сдвиги, т.е. ошибки в нахождении точки делают невозможным попиксельное сравнение

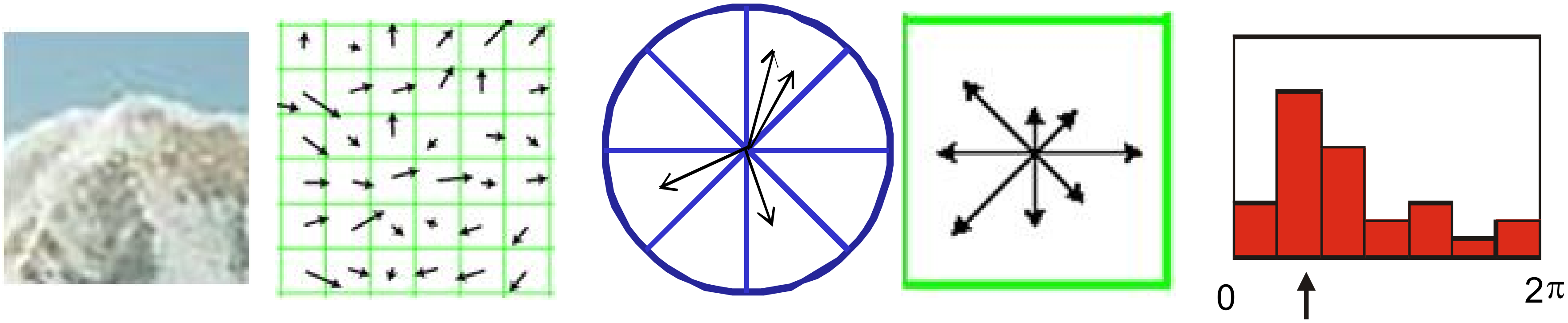
Метод SIFT

- Scale-Invariant Feature Transform:
  - Детектор DoG
    - Определение положения и характерного масштаба особенности
  - Ориентация
    - Определение доминантной ориентации особенности по градиентам
  - Дескриптор
    - Использование статистик по направлению градиентам

- Устойчив к изменениям освещенности и небольшим сдвигам

David G. Lowe. "Distinctive image features from scale-invariant keypoints." *IJCV* 60 (2), pp. 91-110, 2004.

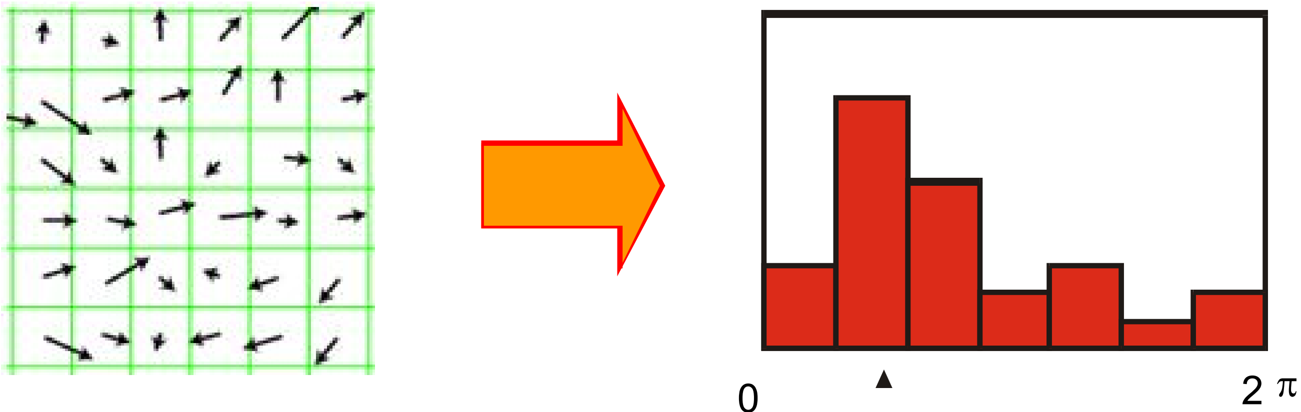
Гистограмма ориентаций градиентов



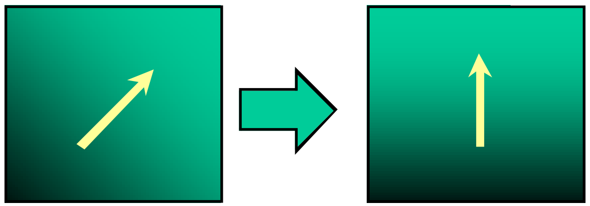
- Основа дескриптора SIFT – подсчёт гистограммы ориентаций градиентов
  - Вычислим направление градиента в каждом пикселе
  - Квантуем ориентации градиентов на 8 ячеек (направлений)
    - Пометим каждый пиксель номером ячейки
  - Посчитаем гистограмму направлений градиентов
    - Для каждой ячейки посчитаем количество пикселов с номером этой ячейки

Ориентация фрагмента

- Идея: найти основное (доминантное) направление градиентов пикселей в окрестности точки
- Выберем в гистограмме ячейку с максимальным значением, возьмём это направление как доминирующее



- Повернем фрагмент так, чтобы дом е направление градиента было направлен вверх

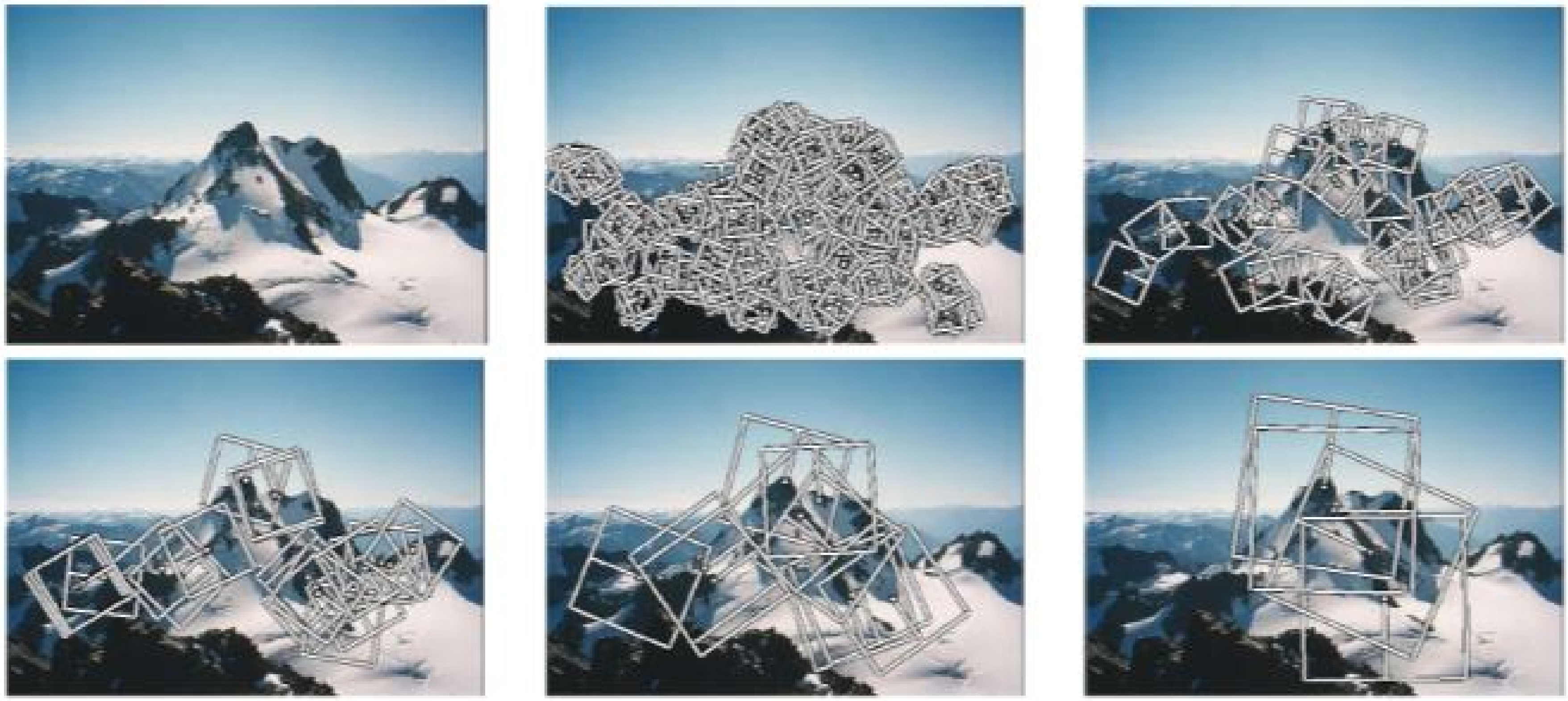


- Если локальных максимумов несколько – считаем, что несколько точек с разной ориентацией

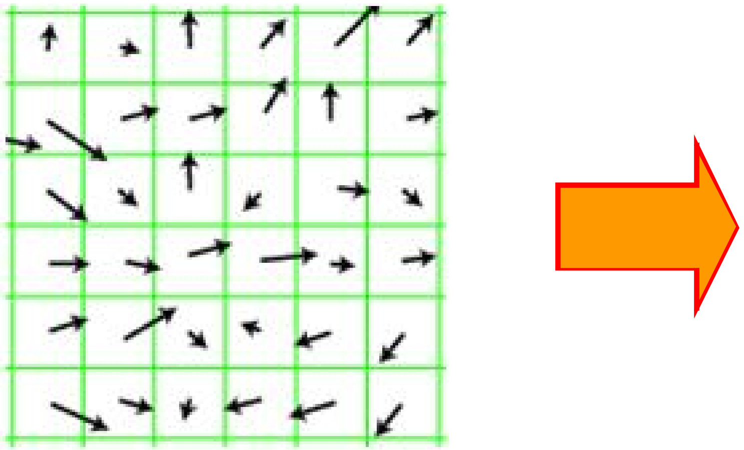
Окрестность особенности

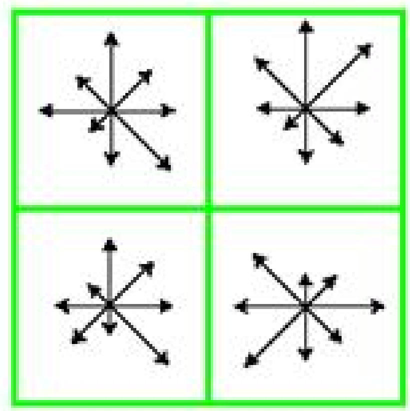
- Для каждой найденной особенности теперь знаем характеристические масштаб и ориентацию
- Выберем соответствующую прямоугольную окрестность
  - (Rotation Invariant Frame)
- Приведем окрестность к стандартному размеру (масштабируем)

Пример локальных особенностей



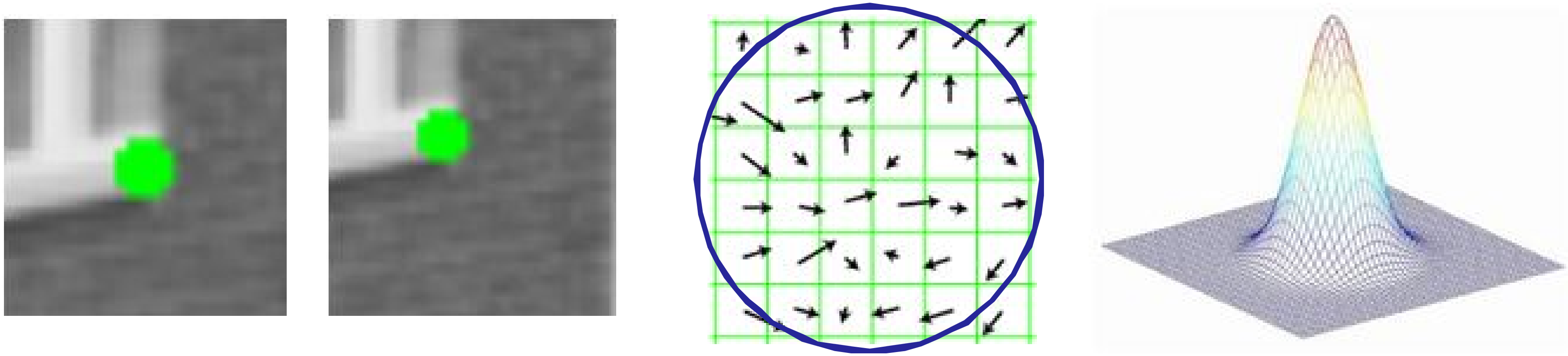
Построение дескриптора





- Для учета пространственного распределения свойств разделим окрестность на блоки сеткой, в каждом блоке посчитаем свою гистограмму градиентов
- Обычно – сетка 4x4, в каждой гистограмма с 8ю ячейками
- Стандартная длина вектора-дескриптора – 128 (4*4*8)
- Можем использовать обычную меру SSD для сравнения дескрипторов
- Можем использовать другие метрики, учитывающие, что дескриптор SIFT – это гистограмма

Устойчивость к сдвигам



- За счёт чего можно дескриптор сделать устойчивым к небольшим сдвигам?
- Использовать ядро при расчёте гистограммы
  - Взвешиваем вклад пикселей по ядрау
  - Веса рассчитываем в зависимости от близости к центру, по Гауссине
  - Небольшие изменения в локализации (положении, масштабе и ориентации) будут приводить к небольшим изменения дескриптора

Резюме SIFT

- Дескриптор SIFT весьма специфичен, устойчив к изменениям освещения, небольшим сдвигам
- Вся схема SIFT (детектор, выбор окрестностей, дескриптор) оказалась очень эффективным инструментов для анализа изображений
- Очень широко используется
- Запатентован
- Модель для многих других дескрипторов

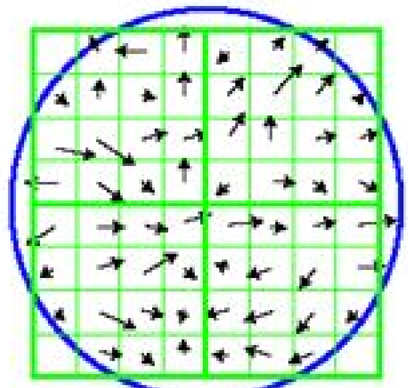


Image gradients

Пример выделения SIFT

Развитие

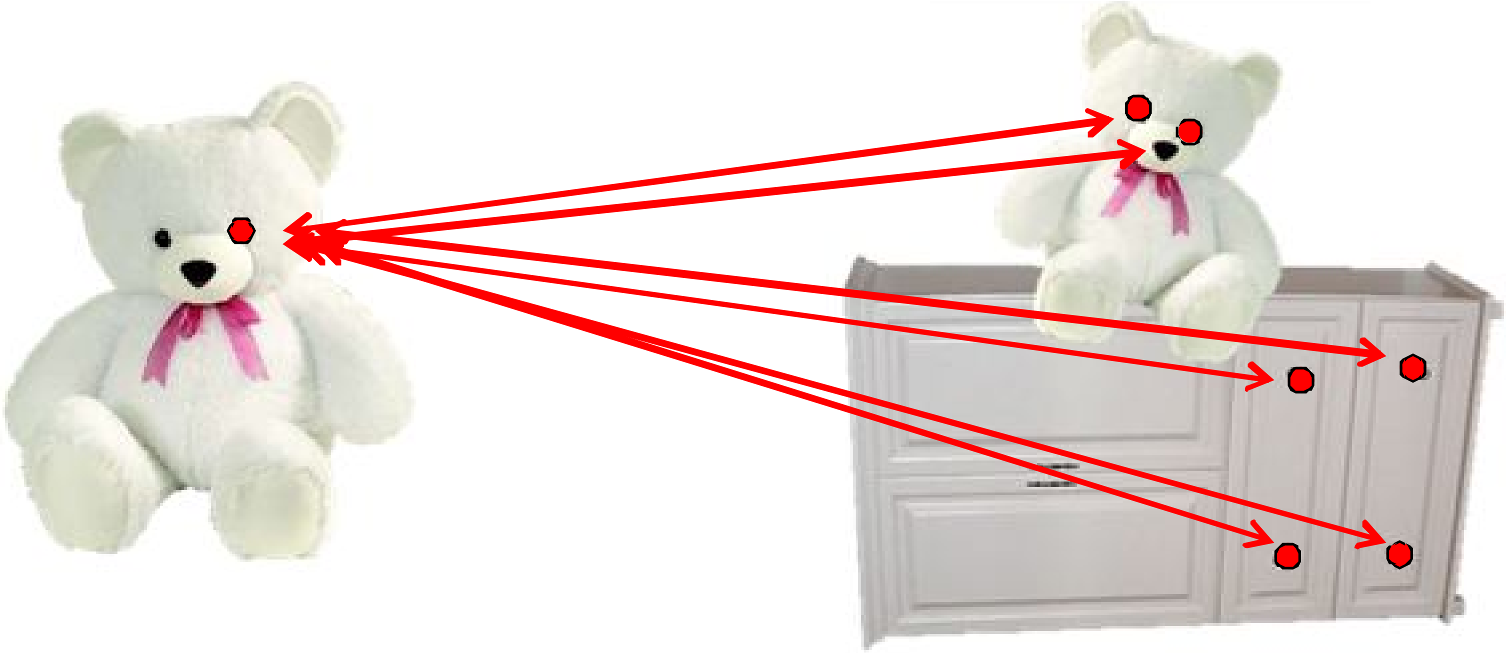
Дескрипторы продолжают активно исследоваться:

- Уменьшение размера
  - Разные формы сжатия
- Ускорение
  - Более простые признаки
- Обучение
  - Подбор оптимальных параметров

Резюме локальных особенностей

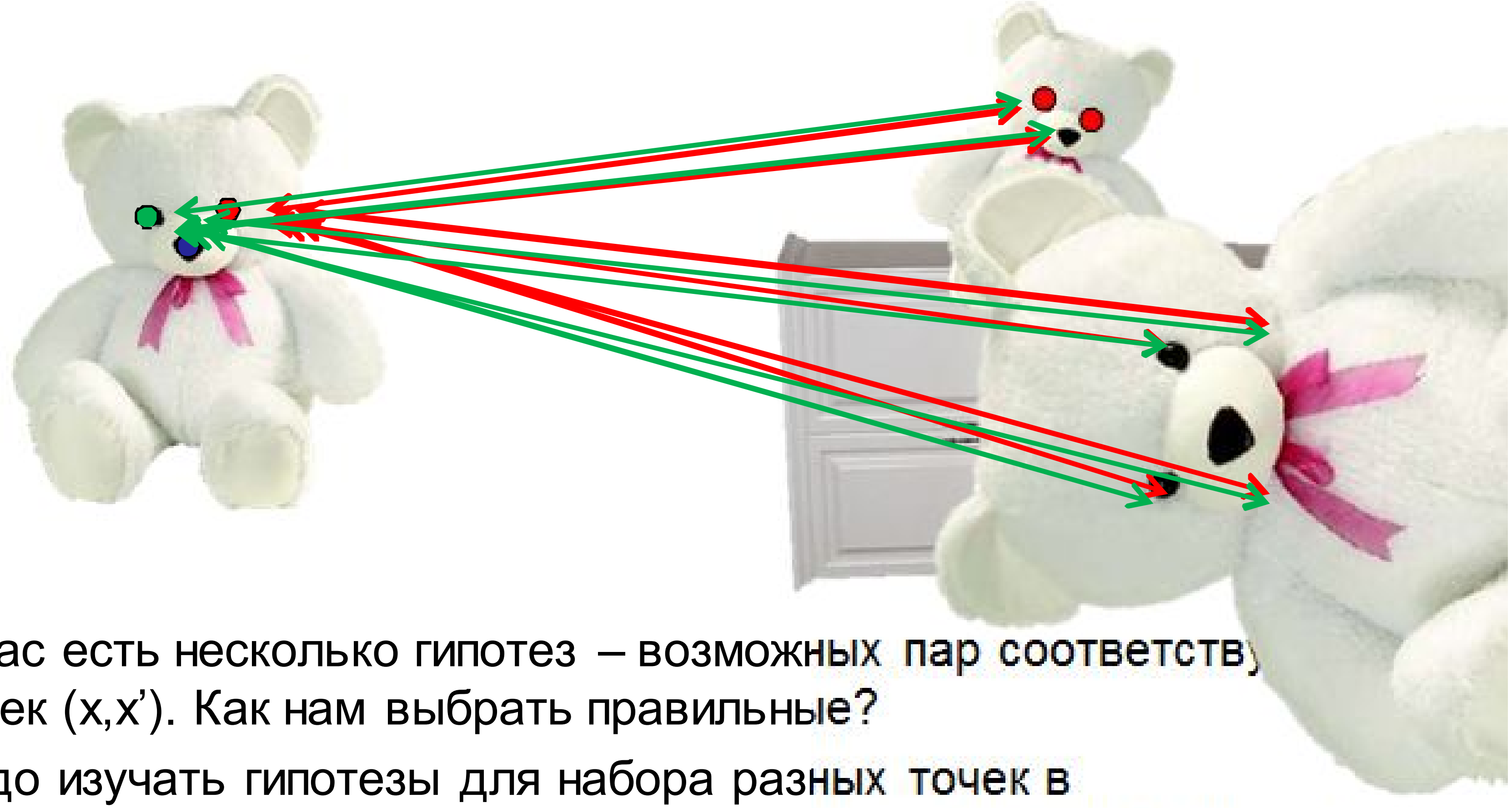
- Локальные характерные особенности - один из основных инструментов для анализа изображений
- Особенности должны быть устойчивы к изменению положения, масштаба, ракурса и освещения изображения
- Мы рассмотрели несколько методов:
  - Детекторы: Harris, LoG, DoG, Harris-Laplace
  - Дескрипторы: Окрестность, SIFT

Сопоставление изображений



- Ищем медведя на новой картинке с помощью особых точек
- Какие могут возникнуть проблемы?
- Несколько точек могут быть похожи
- Возникают ложные соответствия

Сопоставление изображений



- У нас есть несколько гипотез – возможных пар соответствующих точек (x,x’). Как нам выбрать правильные?
- Надо изучать гипотезы для набора разных точек в «совокупности», проверять, согласуются ли они друг с другом
- Что, в этом случае, значит «согласуются друг с другом»?
- Удовлетворяют одному геометрическому преобразованию!

Ошибки в данных

- Какая модель преобразования T «правильная» для медведя?
  - Определенная комбинация сдвига + поворота + масштабирования
- Обозначим одним цветом пару соответствующих точек
- Какие пары соответствующих точек («соответствия») на картинке согласуются друг с другом, а какие нет?
- Элементы данных, удовлетворяющие модели T называются «inlier» («вброс»)
- Не удовлетворяющие T – «outlier» («выброс»)

Общая схема работы

- Найти особые точки, вычислить дескрипторы
- Сопоставить точки по дескрипторам, получить соответствия («putative correspondence»)
- По набору соответствий, с учетом возможных ошибок измерения положения точек и выбросов, вычислить модель преобразования T
- Отфильтровать выбросы в соответствиях

RANSAC

- *Идея* – проведение оценки не по всем данным, а выборке, не содержащей выбросов
- Поскольку какие элементы данных – выбросы – заранее неизвестно, то…
  - …строим много выборок случайным образом!
- По каждой выборке строим гипотезу
- Среди всех гипотез выберем ту, которая лучше всего согласуется со всеми данными
- Random sample consensus (RANSAC)

M. A. Fischler, R. C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM, Vol 24, pp 381-395, 1981.

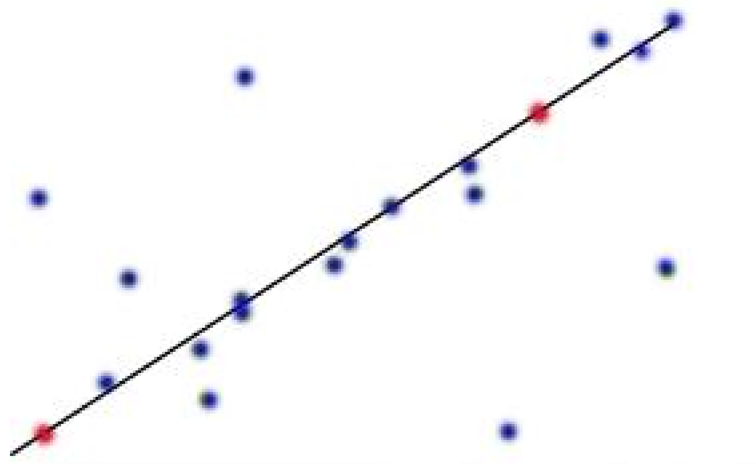
Выборка как основа схемы

• Основой RANSAC является оценка модели по небольшой выборке $S \subset x$

• Проблема - количество таких выборок огромно

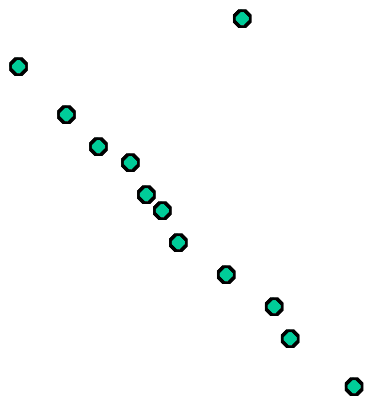
• Поэтому будем строить гипотезы по выборке минимального размера

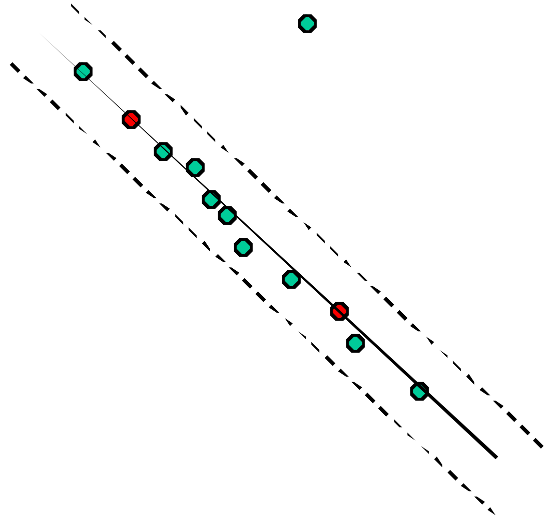
• Для прямой:



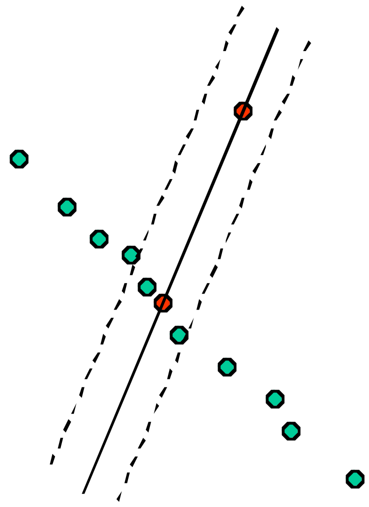
А для геометрических преобразований сколько?

Пример RANSAC

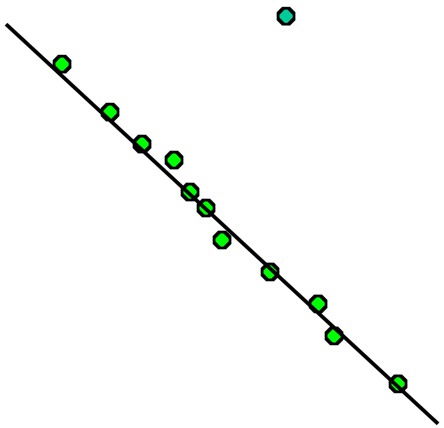




11 «голосов»



4 «голосов»



Базовая схема RANSAC

Повторяем N раз:

(i) Построение случайной выборки $S \subset X$

(ii) Построение гипотезы Θ по выборке S

(iii) Оценка качества гипотезы Θ по набору исходных данных X

(iv) Если гипотеза Θ лучше предыдущих, запоминаем её

После завершения итераций:

(i) Построение чистой выборки X' путём фильтрации выбросов, не удовлетворяющих Θ_{best}

(ii) Уточнение гипотезы Θ_{best} по X'

Сколько гипотез нужно проверить?

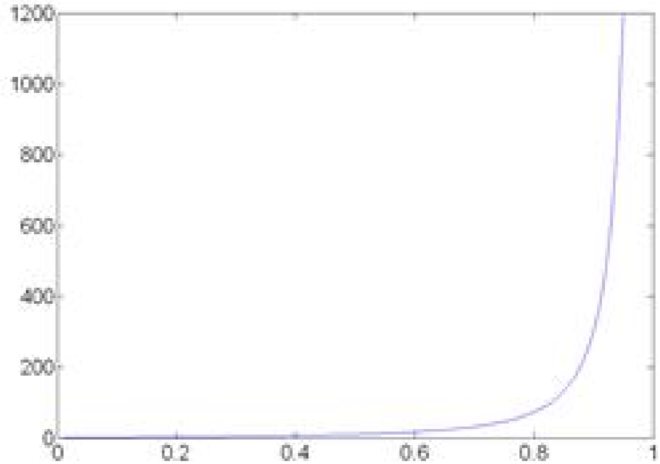
$$\boxed{\left(1-(1-e)^s\right)^N = 1-p}$$

$$N = \log(1-p) / \log\left(1-(1-e)^s\right)$$

- N – количество выборок
- p – вероятность получить хорошую выборку за N итераций;
- s – количество элементов в выборке
- ε – процент хороших точек в наборе

Количество итераций

			proportion of outliers e				
s	5%	10%	20%	25%	30%	40%	50%
2	2	3	5	6	7	11	17
3	3	4	7	9	11	19	35
4	3	5	9	13	17	34	72
5	4	6	12	17	26	57	146
6	4	7	16	24	37	97	293
7	4	8	20	33	54	163	588
8	5	9	26	44	78	272	1177



- Количество выборок быстро растет с ростом размера выборки и доли выбросов
- Как быть, если мы не знаем долю выбросов в данных?

Source: M. Pollefeys

Адаптивное завершение алгоритма

- Доля *e* обычно заранее неизвестна, поэтому начинаем с грубой оценки (пр.: 50%), затем вычисляем оценку хороших точек для каждой гипотезы

- Процедура:
  - $N=\infty$, *sample_count* =0
  - While *N* >*sample_count*
    - Строим выборку, гипотезу, оцениваем кол-во inliers
    - Установим e = 1 – (number of inliers)/(total number of points)
    - Перевычислим *N* по *e:*

$$N = \log(1-p) / \log\left(1-(1-e)^s\right)$$

    - Увеличим *sample_count* на 1

Source: M. Pollefeys

Функции качества

Первоначально было предложено 2 способа оценки качества гипотез:

RANSAC

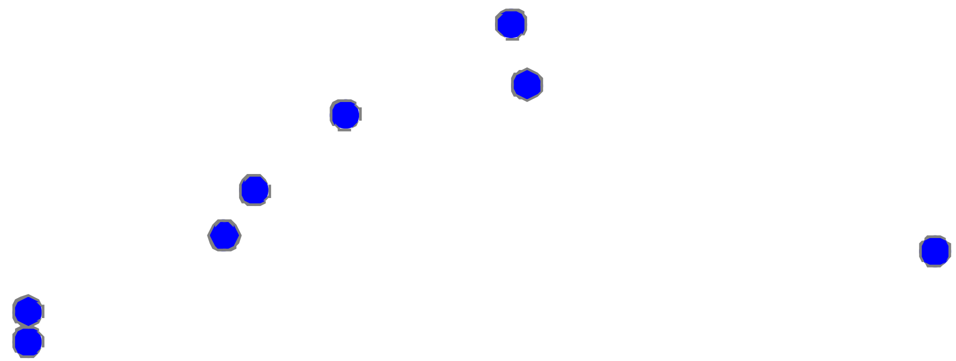
$$R(\theta) = \sum_i p(\varepsilon_i(\theta)^2), \;\; p(\varepsilon_i^2) = \begin{cases} 1 & \varepsilon_i^{\,2} \le T^2 \\ 0 & \varepsilon_i^{\,2} > T^2 \end{cases}, i = \overline{1,n}$$

$\varepsilon_i(\theta)$ - невязка *i*-ой точки и оцениваемой гипотезы

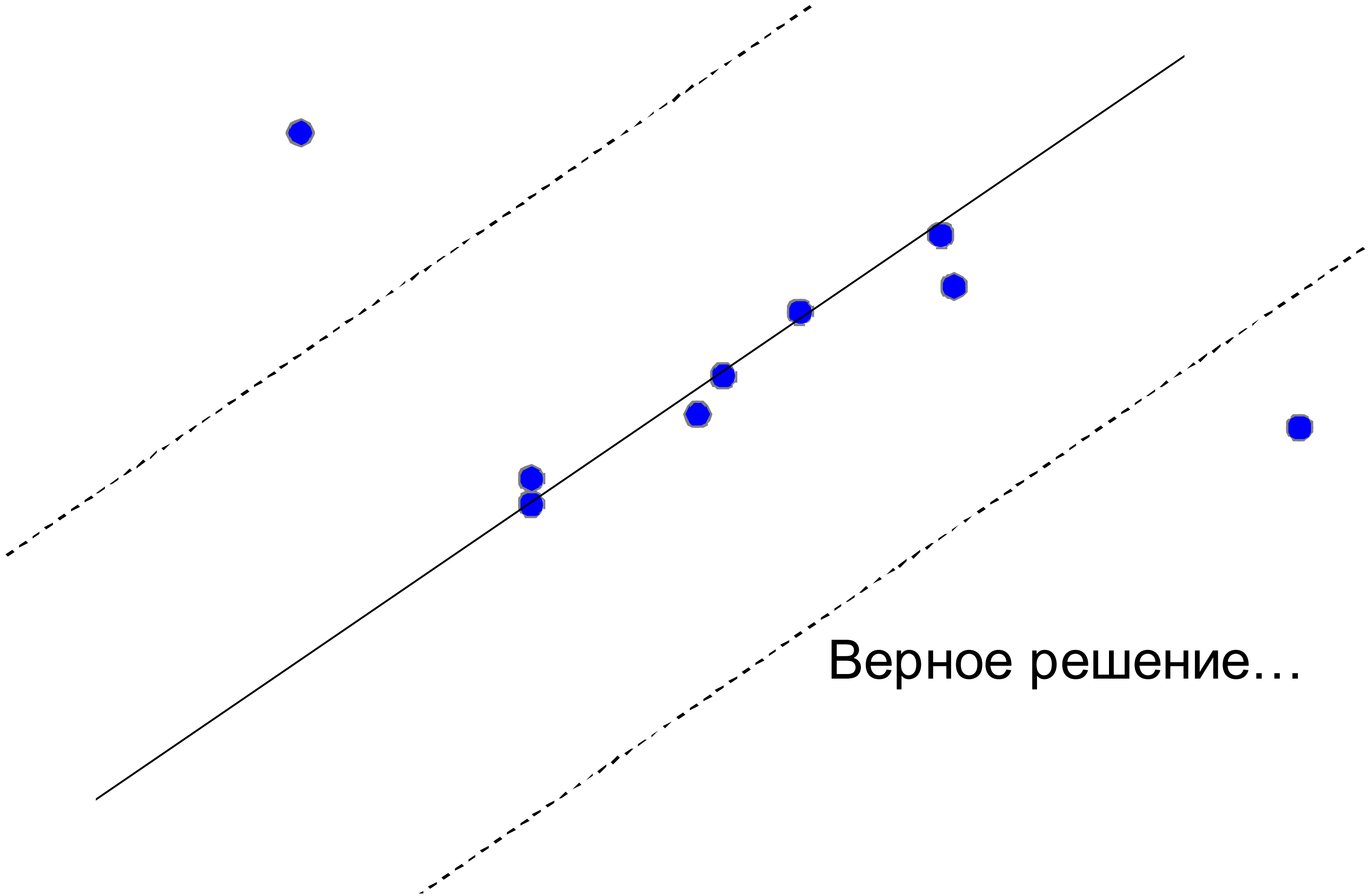
LMS (Least median squares)

$$R(\theta) = median(\varepsilon_i(\theta)^2), i = \overline{1,n}$$

Большой порог – проблема RANSAC!

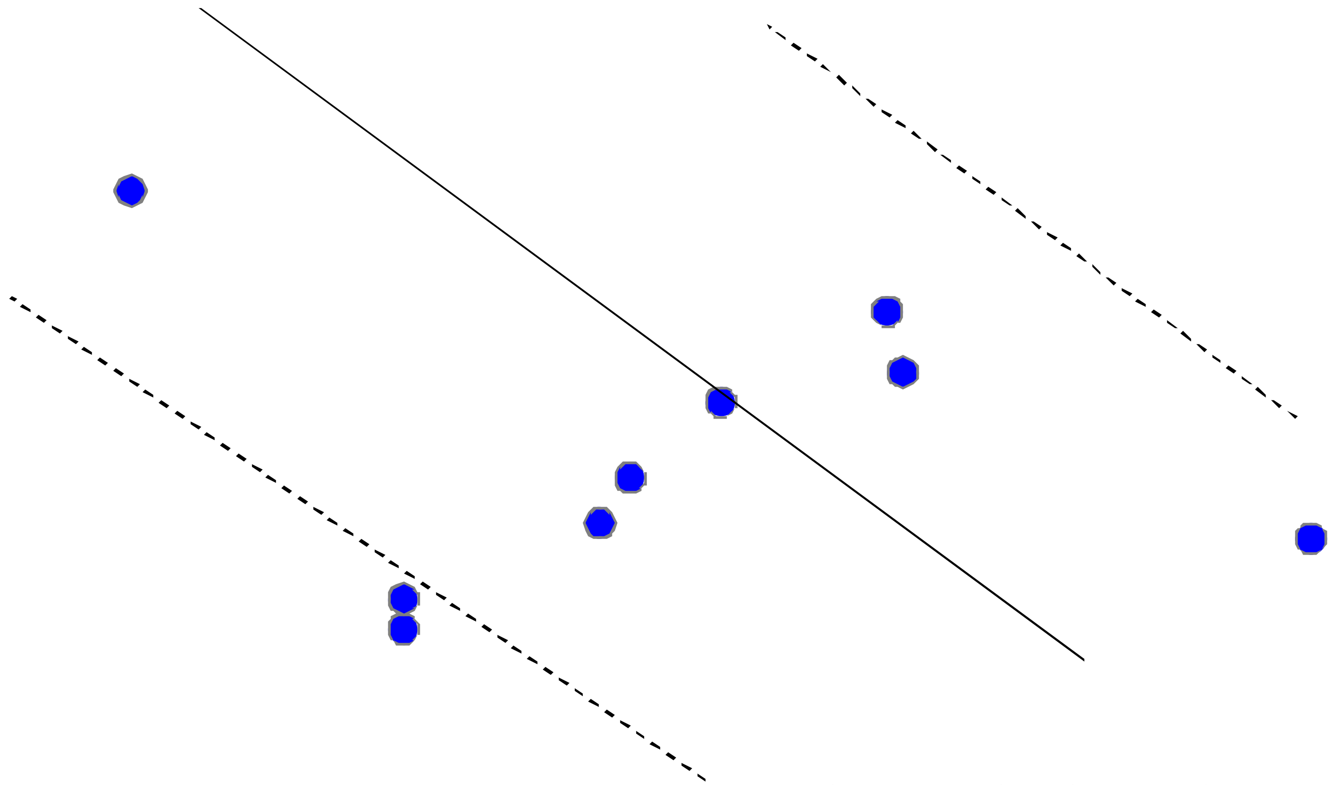


Большой порог – проблема!



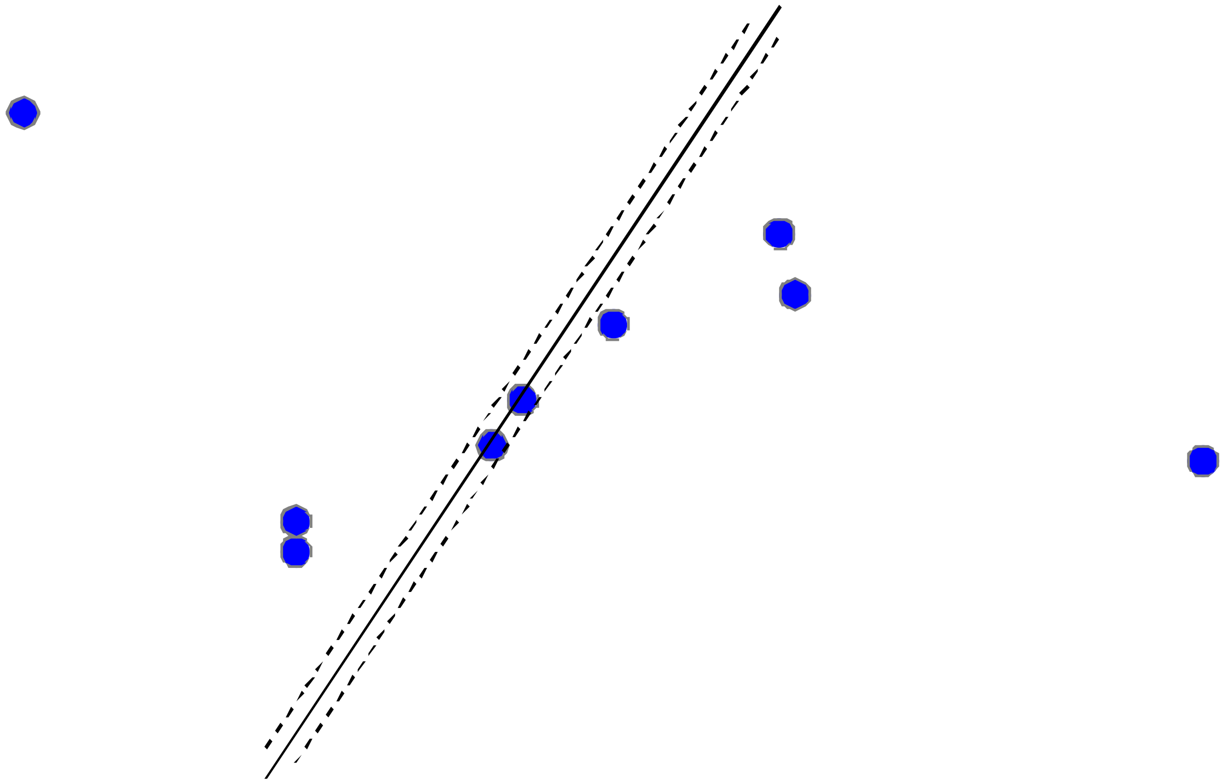
Верное решение…

Большой порог – проблема!

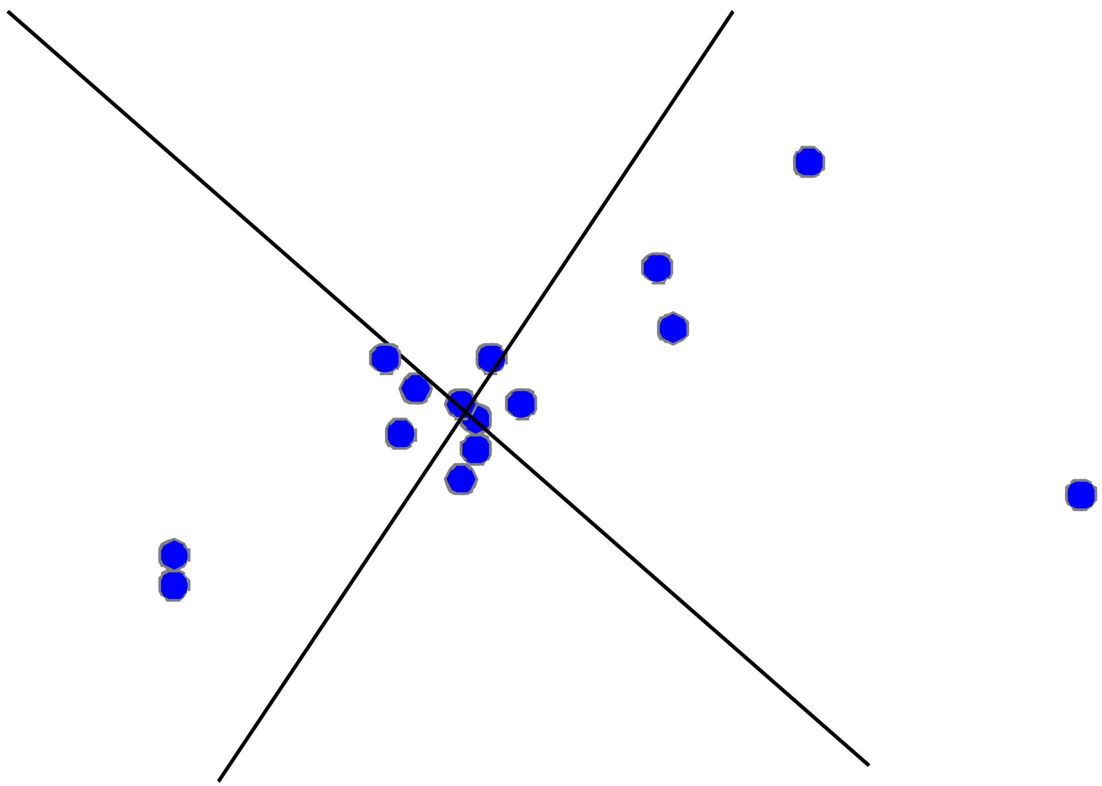


…эквивалентно
неверному

Маленький порог – проблема!



Проблема LMS



Нет хорошего решения, если выбросов >50%

Медиана ошибки одинакова для обоих решений

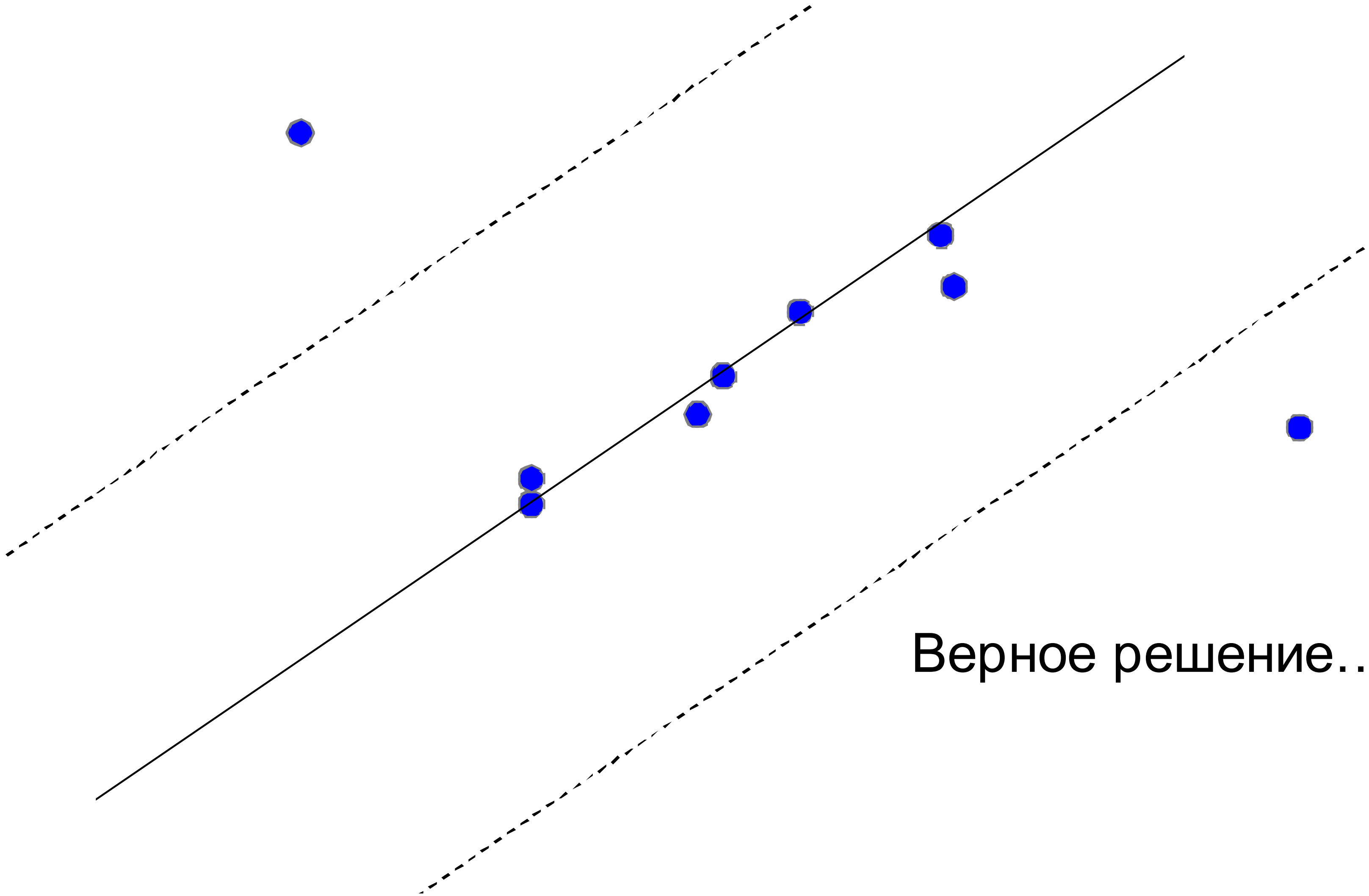
Функции качества

- M-SAC
  - Возьмём робастную функцию, называемую M-оценкой, в качестве целевой функции:

$$R(\theta) = \sum_i p(\varepsilon_i(\theta)^2), \;\; p(\varepsilon_i^2) = \begin{cases} \varepsilon_i^{\,2} & \varepsilon_i^{\,2} \le T^2 \\ T^2 & \varepsilon_i^{\,2} > T^2 \end{cases}, i = \overline{1,n}$$

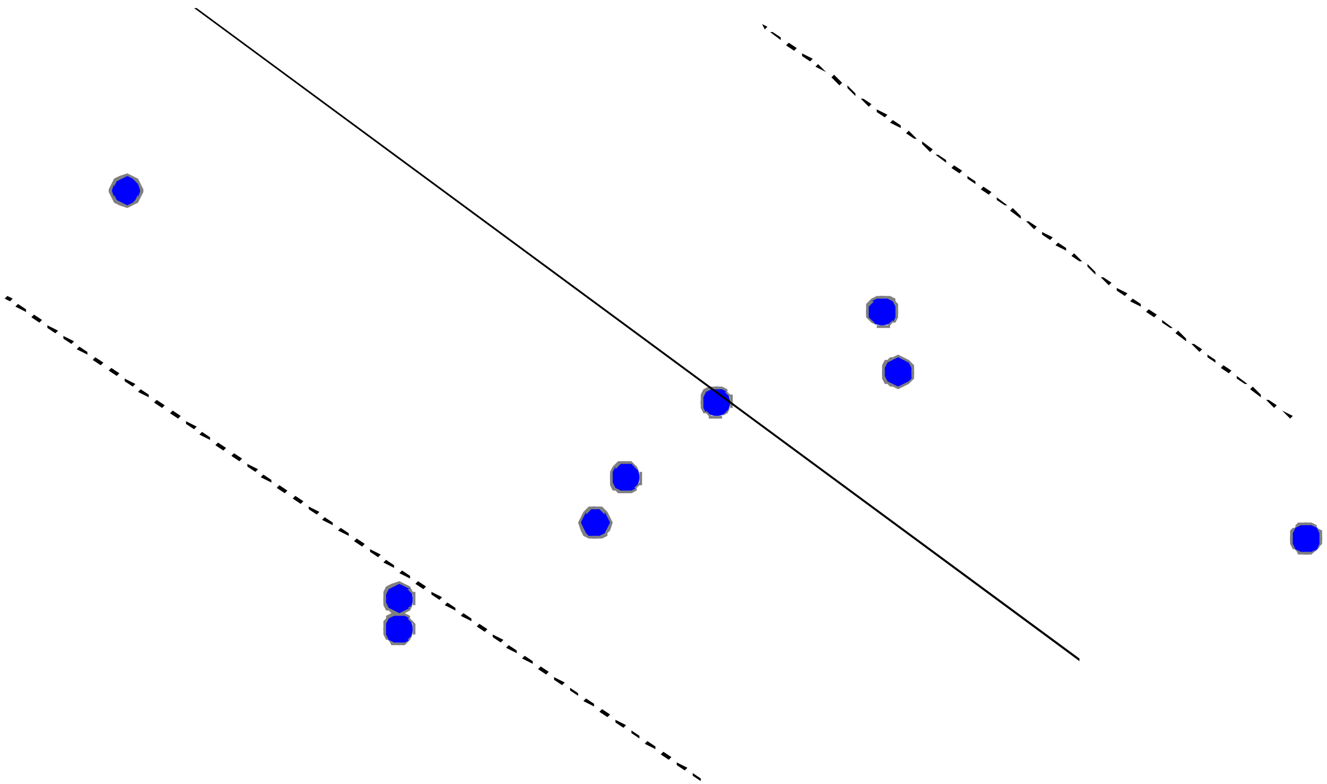
- M-SAC дает более точную оценку без увеличения вычислительной сложности

M-SAC



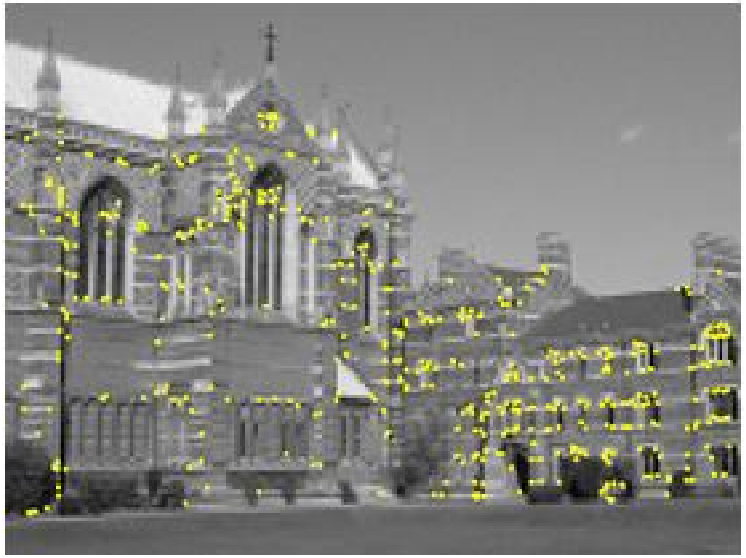
Верное решение…

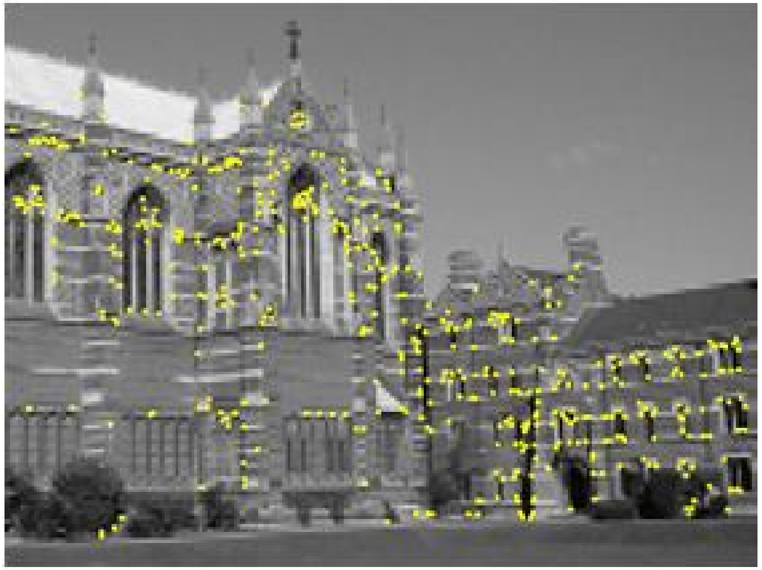
M-SAC



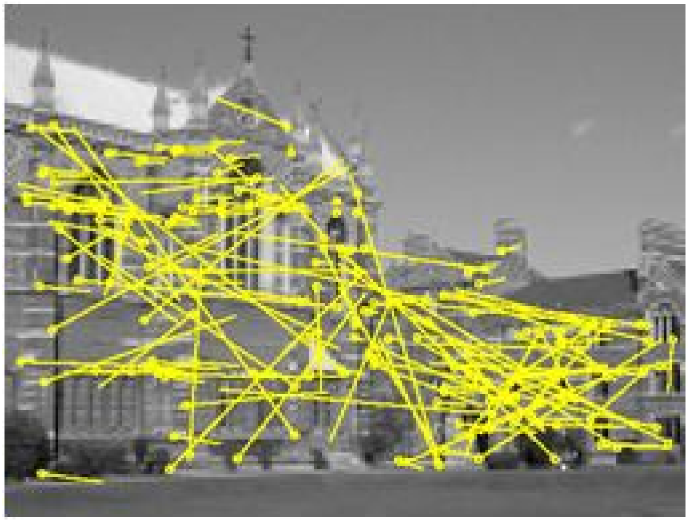
Лучше неверного

Пример использования





Сопоставление изображений

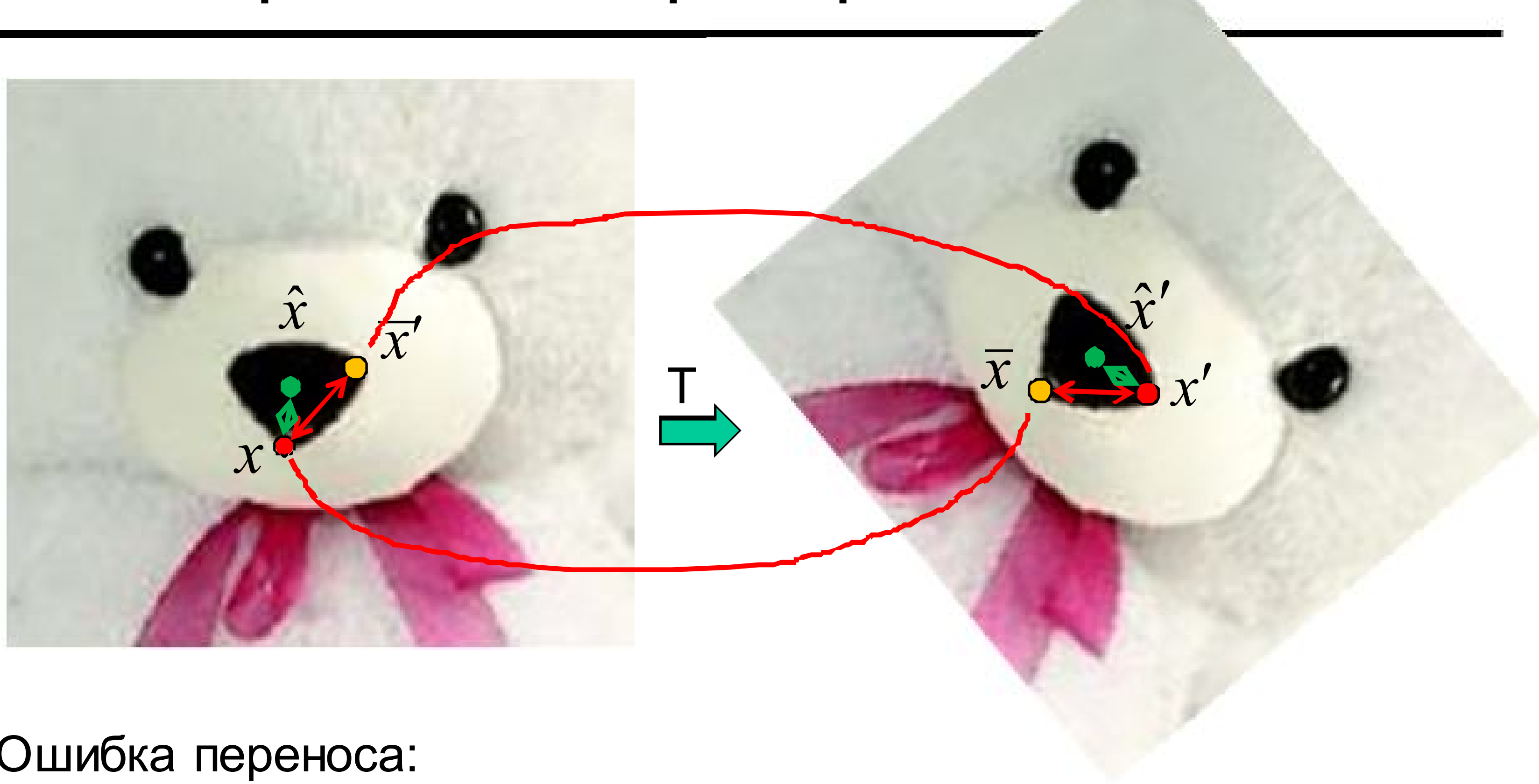


Сопоставление особенностей по дескрипторам – много ложных

Алгоритм сопоставления

- Дано {(x,x’)} – набор пар соответствующих точек на изображениях I и I’
- Вычислим модель преобразования Θ между изображениями I и I’ по ключевым точкам с помощью RANSAC
- Отфильтруем выбросы в {(x,x’)} по модели Θ
  - Если ошибка p(x,x’) > T – (x,x’) - выброс
- Уточним модель Θ
  - Метод наименьших квадратов по всем Inliers
  - Пересчёт inliers и outliers
  - Итеративный пересчёт
  - Нелинейная оптимизация по всем данным с М-функцией в качестве целевой

Геометрические преобразования



Ошибка переноса:

$$\sum_i d^2(x_i', \bar{x}_i) + d^2(x_i, \bar{x}_i') \qquad \bar{x} = Tx \qquad \bar{x}' = T^{-1}x'$$

Геометрическая ошибка:

$$\sum_i d^2(x_i', \hat{x}_i') + d^2(x_i, \hat{x}_i) \qquad \hat{x}_i' = T\hat{x}_i$$

Расчёт ошибок

- Ошибка переноса

$$\sum_i d^2(x_i', \bar{x}_i) + d^2(x_i, \bar{x}_i') \quad \bar{x} = Tx \quad \bar{x}' = T^{-1}x'$$

Рассчитывается достаточно просто

- Геометрическая ошибка

$$\sum_i d^2(x_i', \hat{x}_i') + d^2(x_i, \hat{x}_i) \quad \hat{x}_i' = T\hat{x}_i$$

Вычисление оптимальных $(\hat{x}_i, \hat{x}_i')$ зависит от конкретного преобразования.

Чаще всего геометрическая ошибка оптимизируется как нелинейная функция. Берутся начальные приближения для $(\hat{x}_i, \hat{x}_i')$ и итеративно оптимизируется.

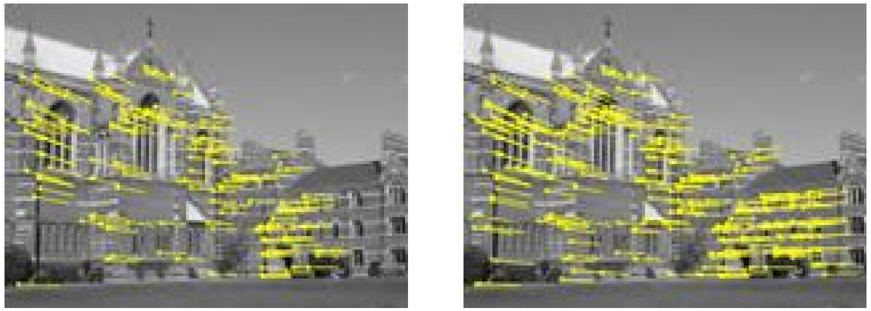
Медленный, но самый точный метод.

Ложные соответствия

500 особенностей

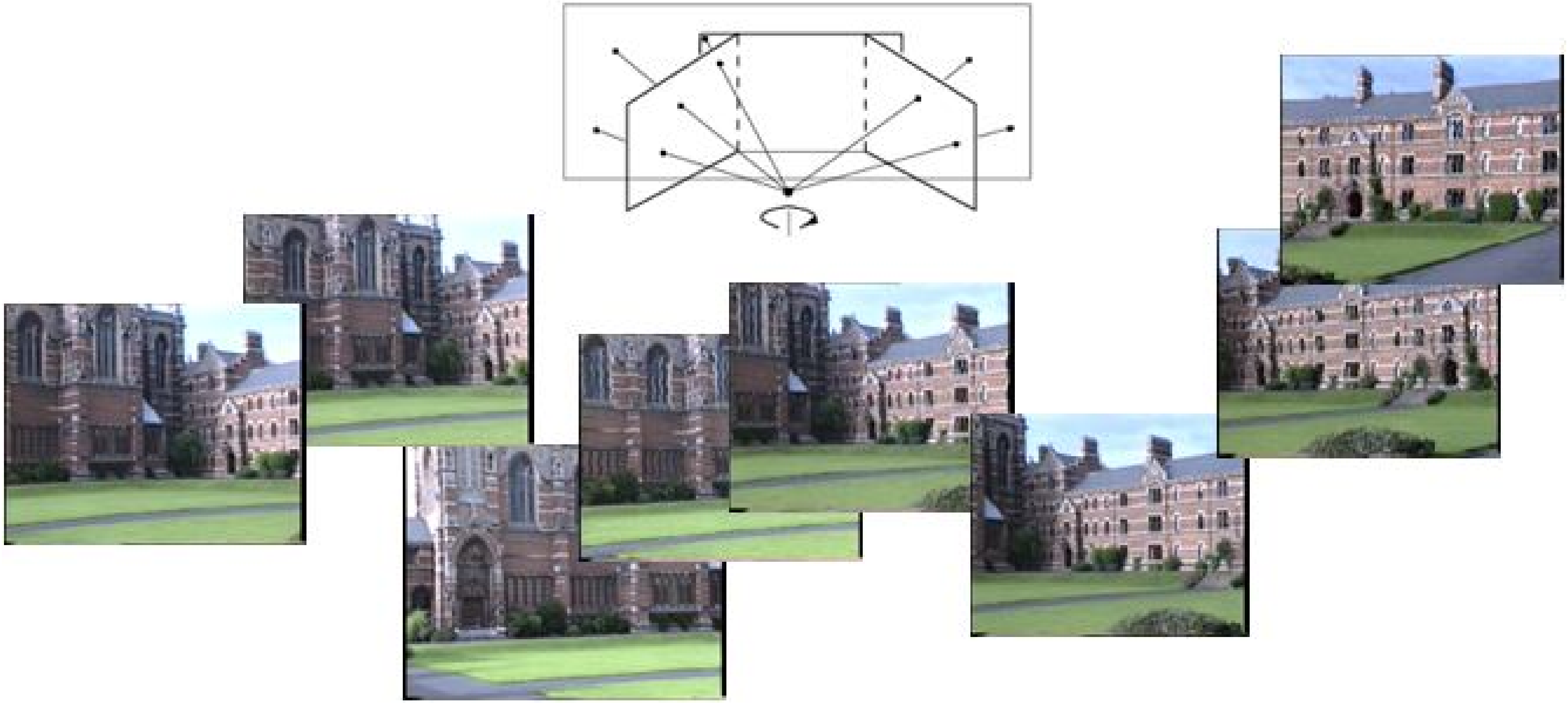
Соответствия (268)

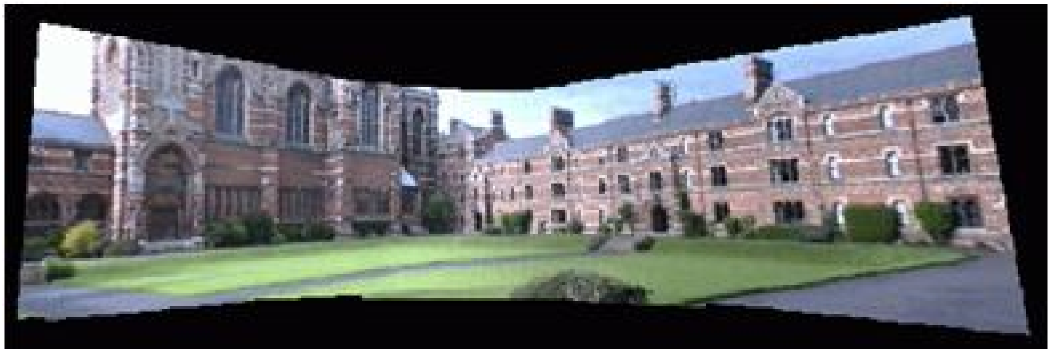
Выбросы (117)



Хороших соответствий (151)

Приложение: склейка панорам

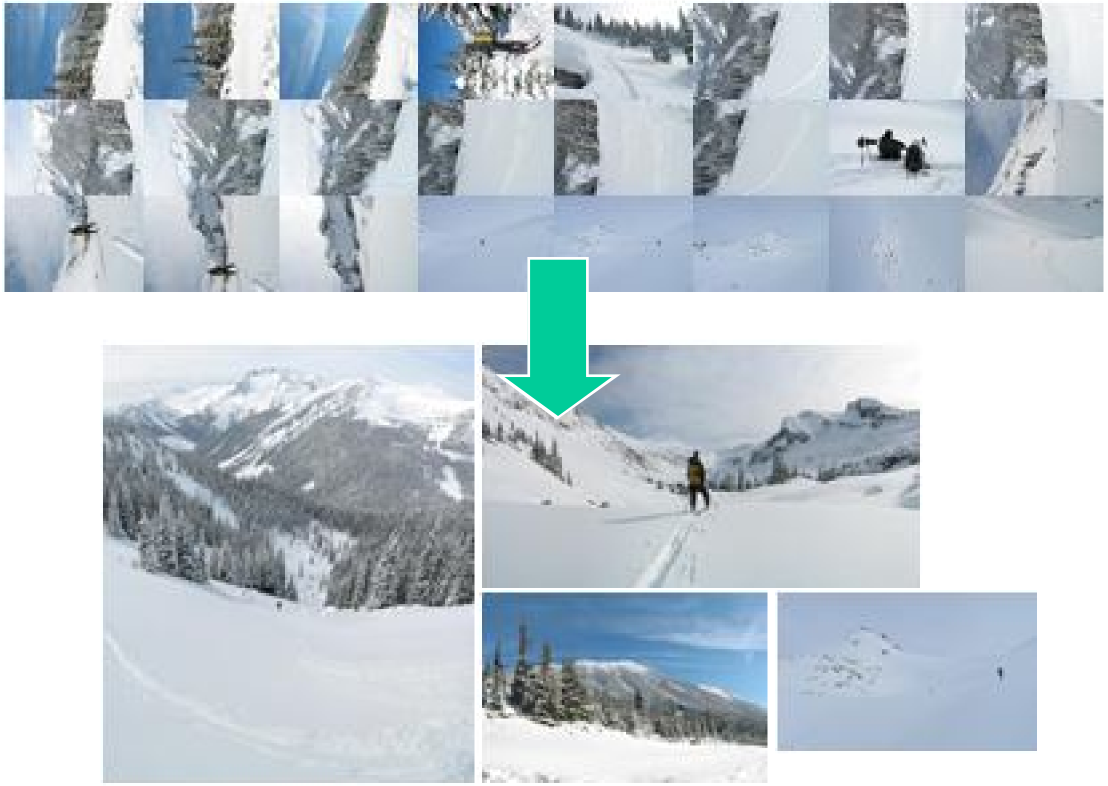




Slide by S. Lazebnik

Распознавание панорам

Благодаря мощи методов SIFT и RANSAC можно в неупорядоченном наборе фотографий определить, какие относятся к какой панораме, и сшить их



M. Brown and D. Lowe, *“Recognizing Panoramas,”* ICCV 2003.
http://www.cs.ubc.ca/~mbrown/panorama/panorama.html

Slide by S. Lazebnik

RANSAC pros and cons

- Плюсы
  - Простой и общий метод
  - Применим для множества задач
  - Хорошо работает на практике
- Минусы
  - Много настраиваемых параметров
  - Не всегда удается хорошо оценить параметры по минимальной выборке
  - Иногда требуется слишком много итераций
  - Не срабатывает при очень высокой доле выбросов
  - Часто есть лучший способ, нежели равновероятно выбирать точки