

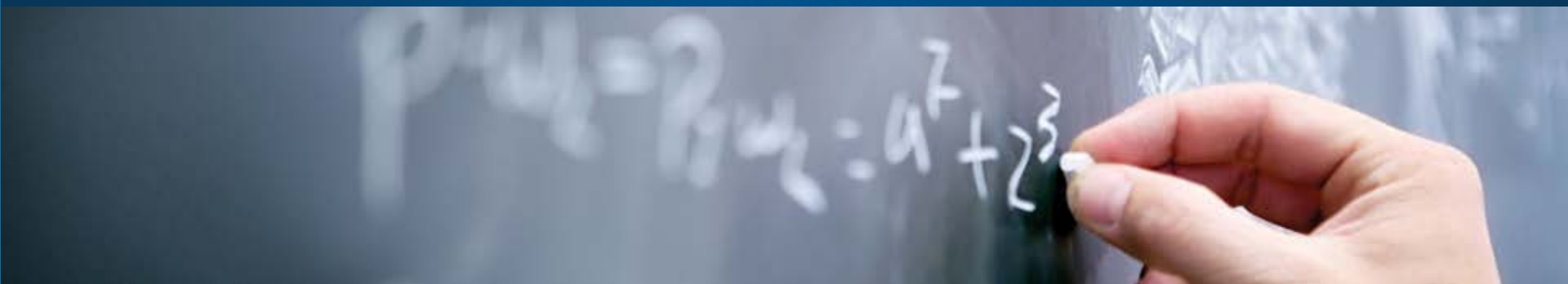
Анализ изображений и видео

Лекция 5: Построение признаков и сравнение изображений: локальные признаки

Наталья Васильева

nvassilieva@hp.com

HP Labs Russia



Признаки изображений



Описывают картинку целиком:

- средняя яркость;
- среднее значение по красному каналу;
- ...

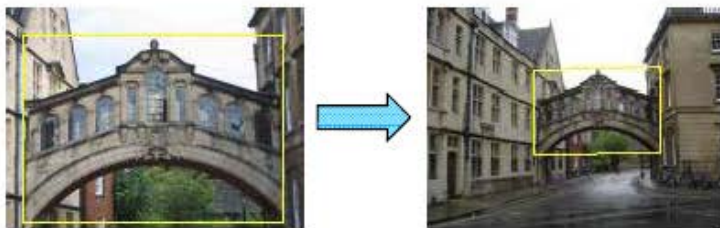
Обычно, вычисляются по всем пикселям изображения

Описывают часть картинки:

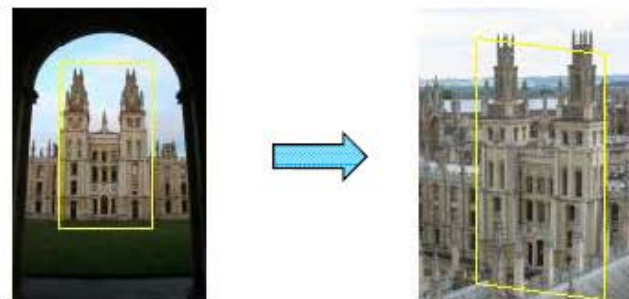
- средняя яркость верхней левой четверти;
- среднее значение по красному каналу в окрестности центра изображения;
- ...

Сегментация, поиск точек интереса, построение признаков по окрестностям точек интереса

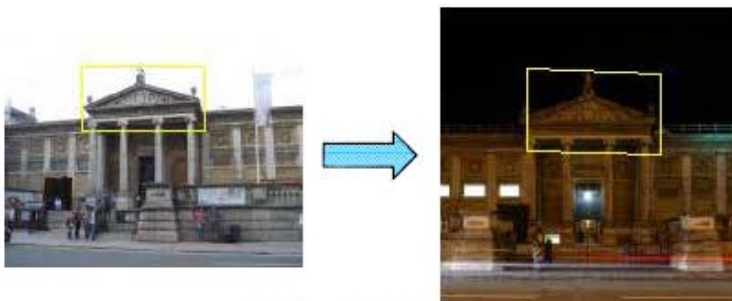
Когда глобальные признаки не работают?



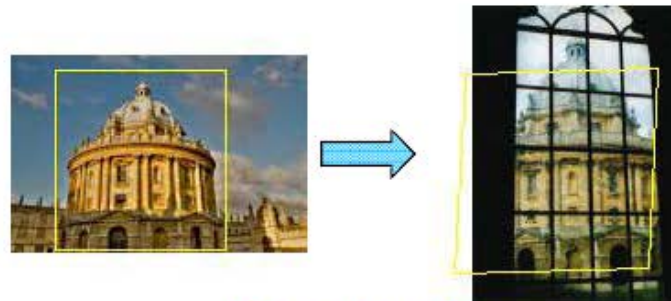
Scale



Viewpoint



Lighting

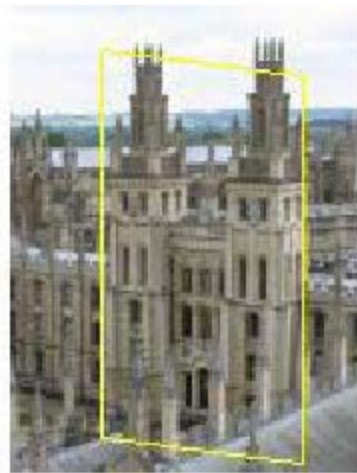
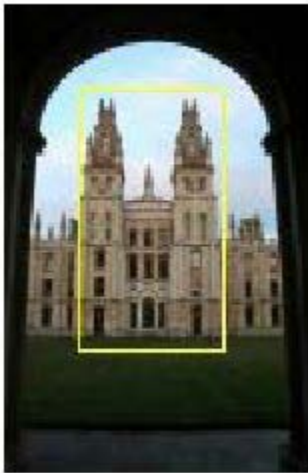


Occlusion

Slide credit: Josef Sivic

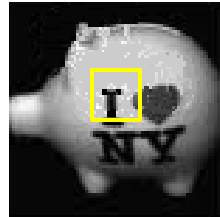
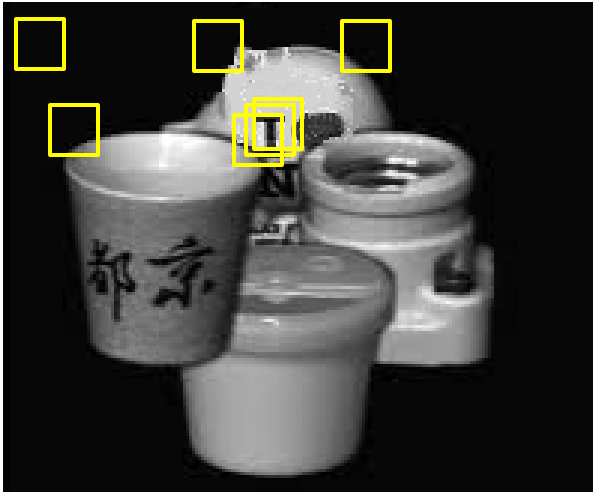


Решение: локальные признаки



- Хотя глобальные признаки отличаются, можно найти сопоставимые фрагменты изображений
- Как выбирать фрагменты?
- Как описывать фрагменты?

Как сопоставлять фрагменты?



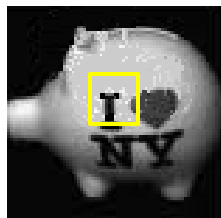
1. Сканирование (scanning)

Полный перебор (точно не пропустим пару сопоставимых фрагментов)

Полный перебор (медленно)

Слишком много пар

Как сопоставлять фрагменты?



- ★ Ключевые точки
- Точки интереса
- Точки внимания
- Особые точки
- Точечные особенности
- “salient”
- “keypoints”
- “representative”

1. Выборочное сравнение фрагментов (sparse matching)

Быстрее: меньше сравнений

Изолированные фрагменты (нет пересекающихся пар)

Нет гарантии, что всегда будут найдены все пары

Как искать «ключевые» точки???

Какими должны быть «особые» фрагменты?

1. Их должно быть немного
 - существенно меньше, чем пикселей на изображении
2. Информативные, репрезентативные, уникальные
 - Если окрестности двух точек не отличимы, будет сложно понять, какую из них сопоставить искомому фрагменту
3. Повторяемые
 - Одна и та же точка должна находиться на изображении вне зависимости от геометрических и фотометрических изменений объекта съемки
4. Локальные
 - Небольшого размера, устойчивы к частичному перекрыванию другим объектом



Сравнение изображений при помощи локальных признаков: основные шаги

1. Локализация особых точек
2. Выделение особых фрагментов – окрестности ключевых точек, инвариантные к изменению масштаба
3. Построение векторов признаков для найденных фрагментов
4. Сопоставление наборов локальных признаков для двух изображений

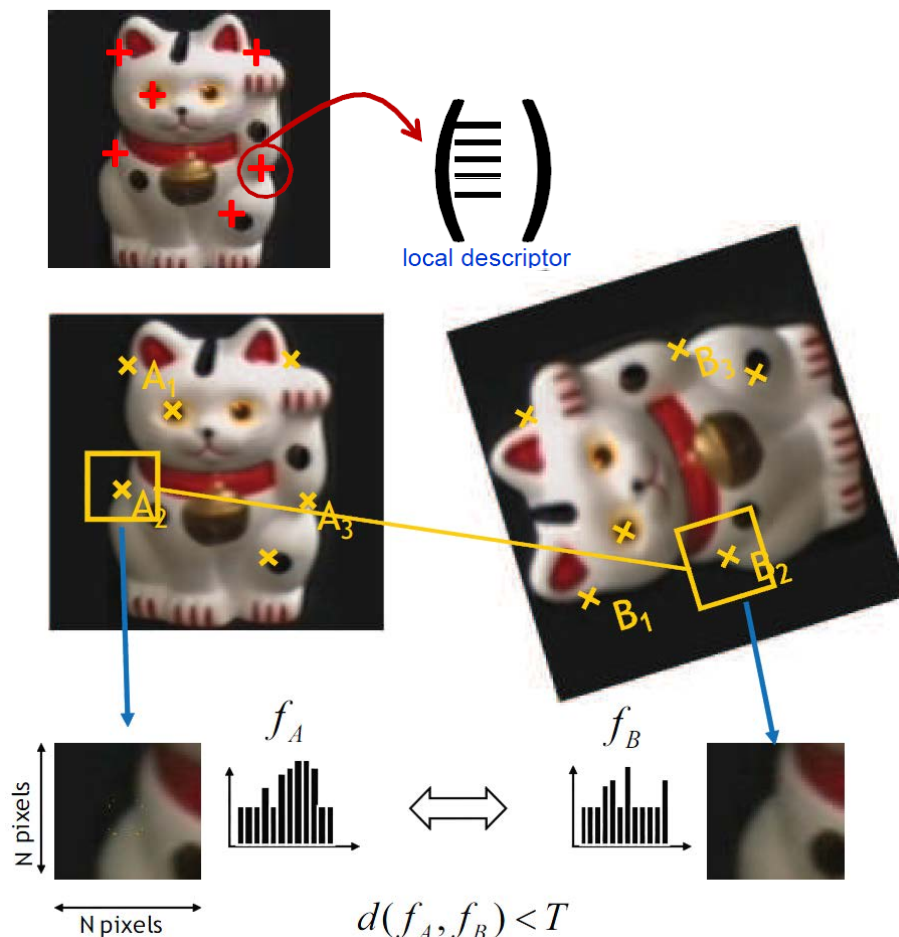
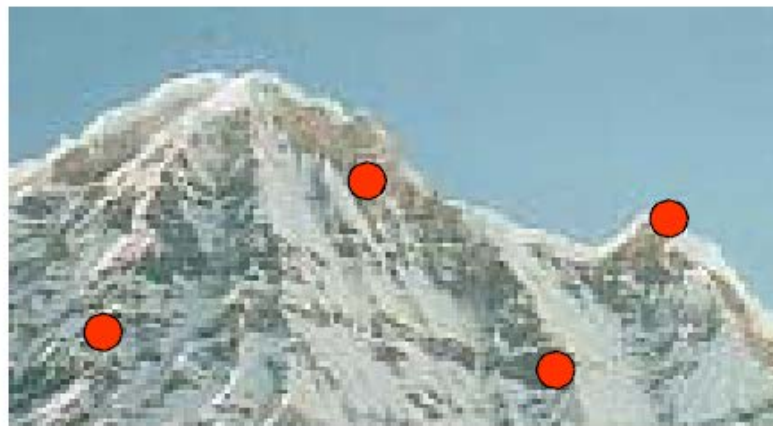


Fig. credit: K. Grauman, B. Leibe

Повторяемость особых точек



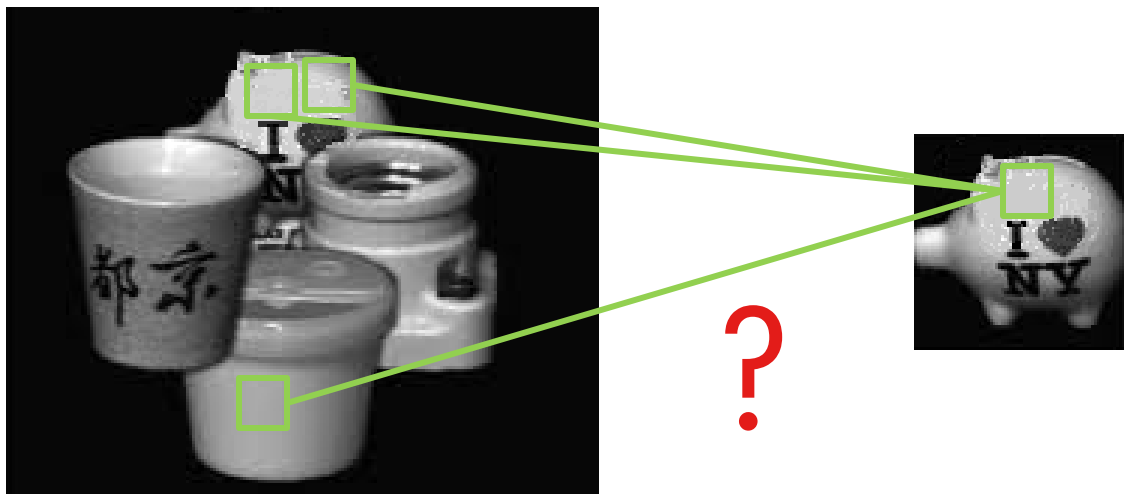
Нет совпадающих особых точек – нет возможности сопоставить изображения!

- Необходимо, чтобы хотя бы часть особых точек первого изображения была обнаружена на втором
- При этом обнаружение особых точек должно происходить независимо для каждого изображения

Slide credit: K. Grauman



Информативность, репрезентативность

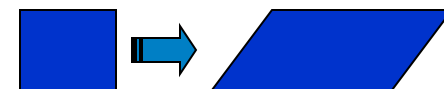
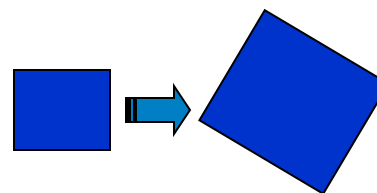
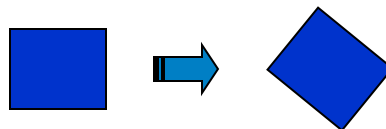


- Желательна однозначность в сопоставлении фрагментов
- Желательна инвариантность к геометрическим и фотометрическим трансформациям объекта на разных изображениях

Геометрические и фотометрические трансформации изображения

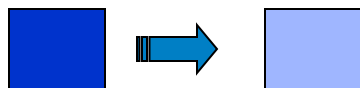
Геометрические:

- Поворот
- Поворот + изменение масштаба
- Аффинные преобразования



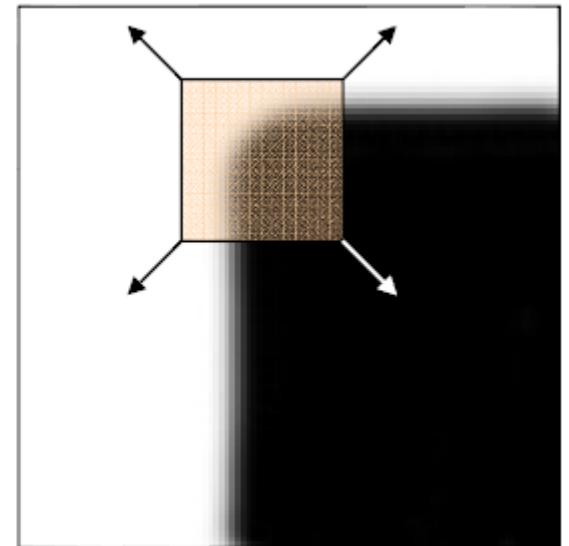
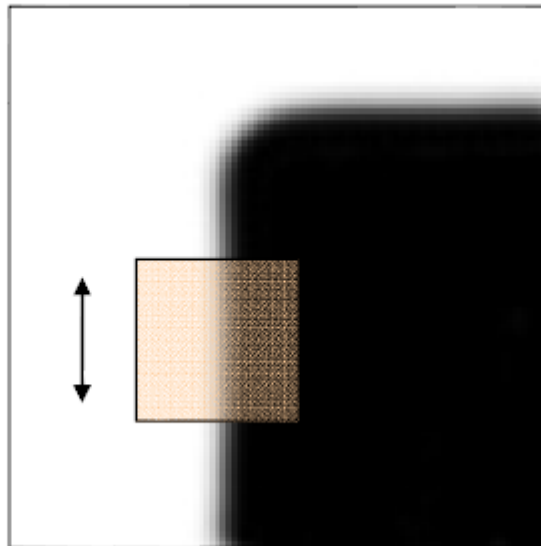
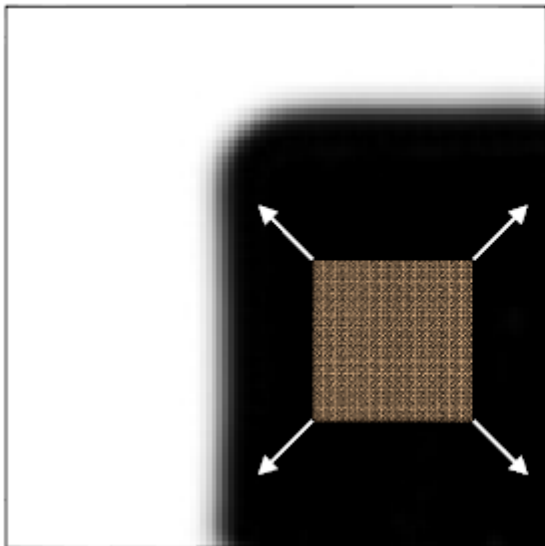
Фотометрические

- Аффинные преобразования интенсивности ($I \rightarrow aI + b$)

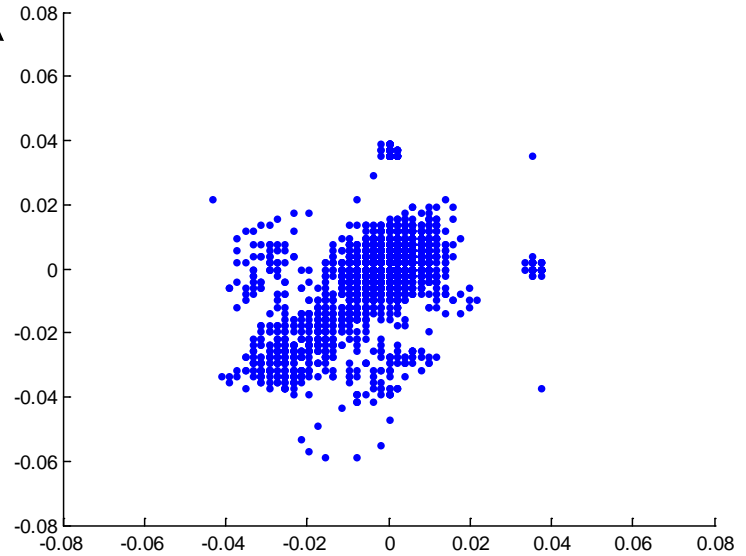
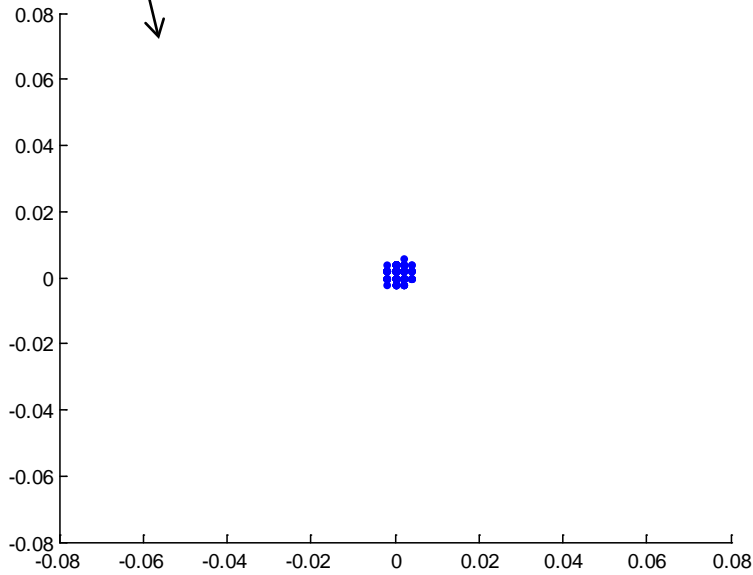
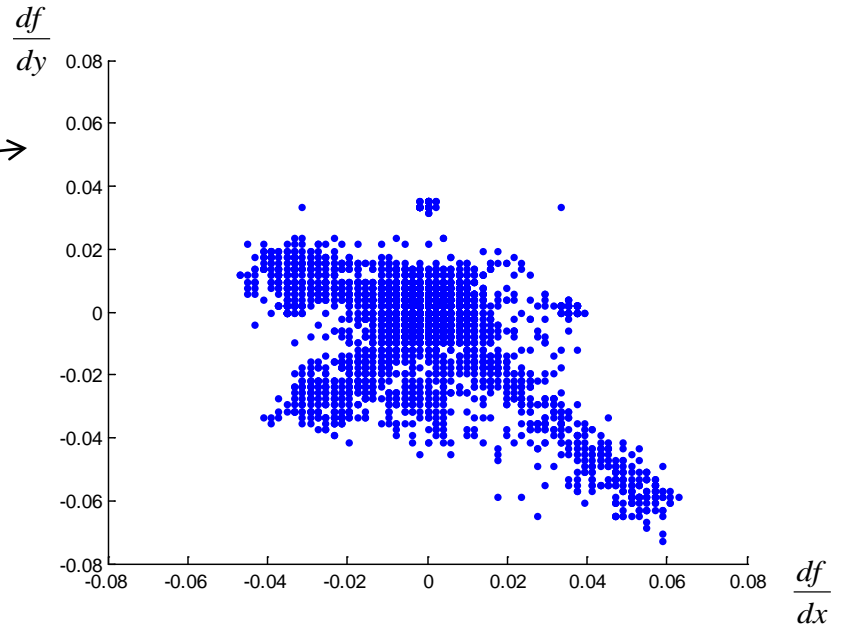
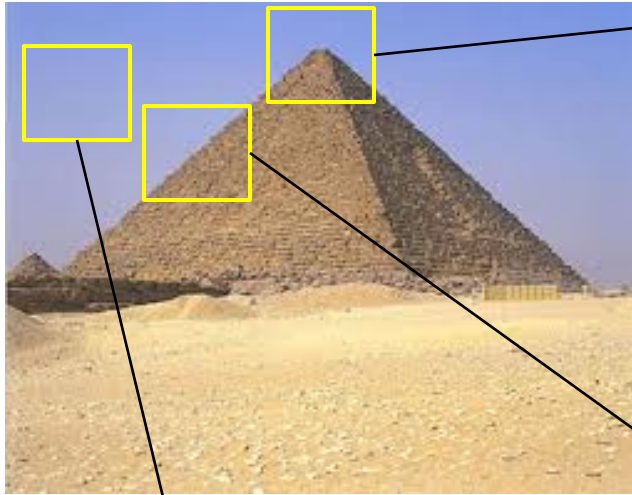


Угловые точки в качестве особых

- Просто обнаружить, используя маленькое окно
- Смещение окна в окрестности угловой точки в любом направлении приводит к изменению набора интенсивностей окна (алгоритм Моравика)
- В окрестности угловой точки – два направления градиентов (алгоритм Харриса)

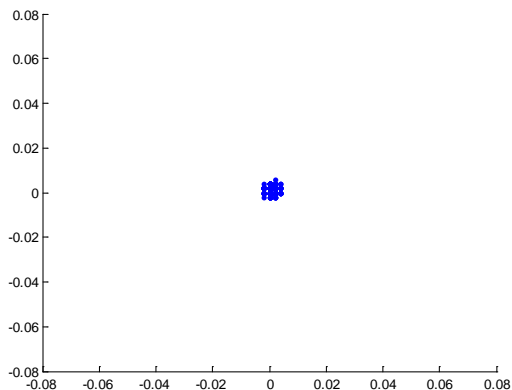


Детектор Харриса

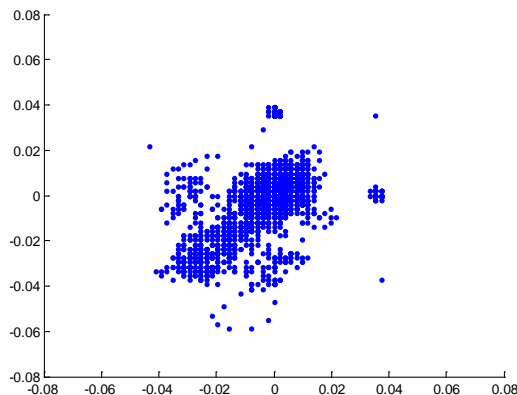


Детектор Харриса – основная идея

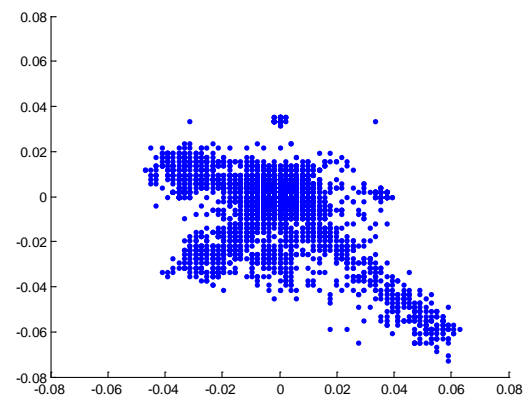
- Для каждого окна смотрим на собственные числа ковариационной матрицы градиентов пикселей
- Собственные вектора этой матрицы показывают «основные направления» данных
- Величина собственных чисел указывает на «степень выраженности» соответствующего направления



Нет «больших» собственных чисел



1 «большое» собственное число



2 «больших» собственных числа

Детектор Харриса: математика

Изменение интенсивности при сдвиге на $[u, v]$:

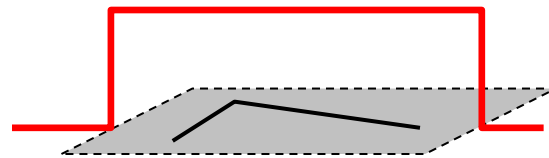
$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

Window function

Shifted intensity

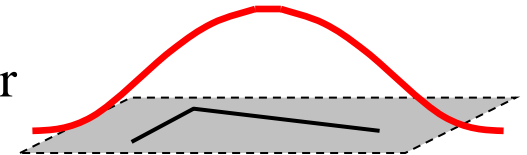
Intensity

Window function $w(x, y) =$



1 in window, 0 outside

or



Gaussian

Детектор Харриса: математика

Для небольших сдвигов $[u, v]$ можем аппроксимировать:

$$I(u + x, v + y) \approx I(x, y) + I_x(x, y)u + I_y(x, y)v$$

Это приводит к приближению:

$$E(u, v) \approx \sum_{x, y} w(x, y) (I_x(x, y)u + I_y(x, y)v)^2$$

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

где M - матрица 2×2 , состоящая из частных производных от интенсивности:

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



Детектор Харриса: математика

Угол характеризуется большим изменением E во всех направлениях вектора $[u, v]$: анализ собственных значений матрицы M

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad \lambda_1, \lambda_2 - \text{собственные значения } M$$

Measure of corner response:

$$R = \det M - k (\text{trace } M)^2$$

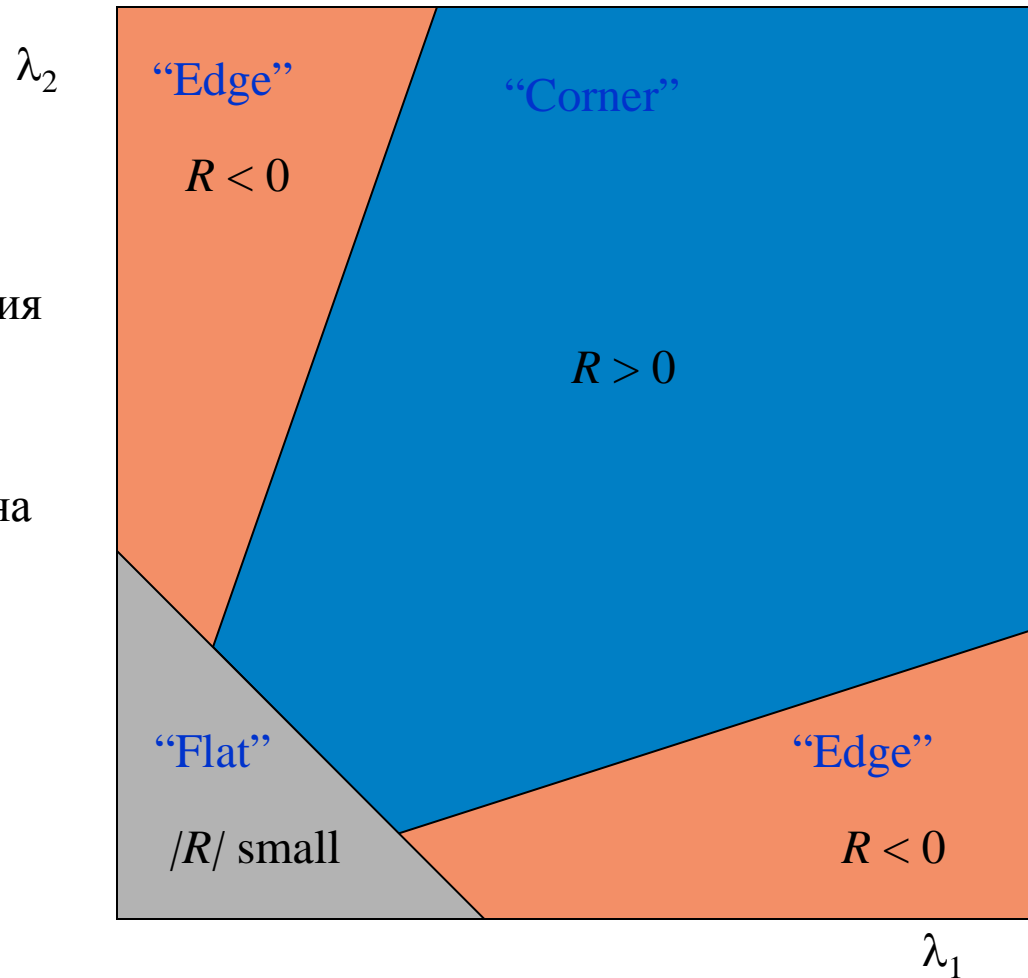
$$\det M = \lambda_1 \lambda_2$$

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

(k – эмпирическая константа, $k = 0.04-0.06$)

Детектор Харриса: математика

- R зависит только от значений собственных чисел λ_1, λ_2
- R принимает большие значения в угловых точках
- R принимает отрицательные значения, большие по модулю на контуре
- $|R|$ мало в плоском регионе



Алгоритм Харриса

Алгоритм:

- Найти точки с большим значением R ($R > \text{threshold}$)
- Выбрать среди них точки локального максимума R

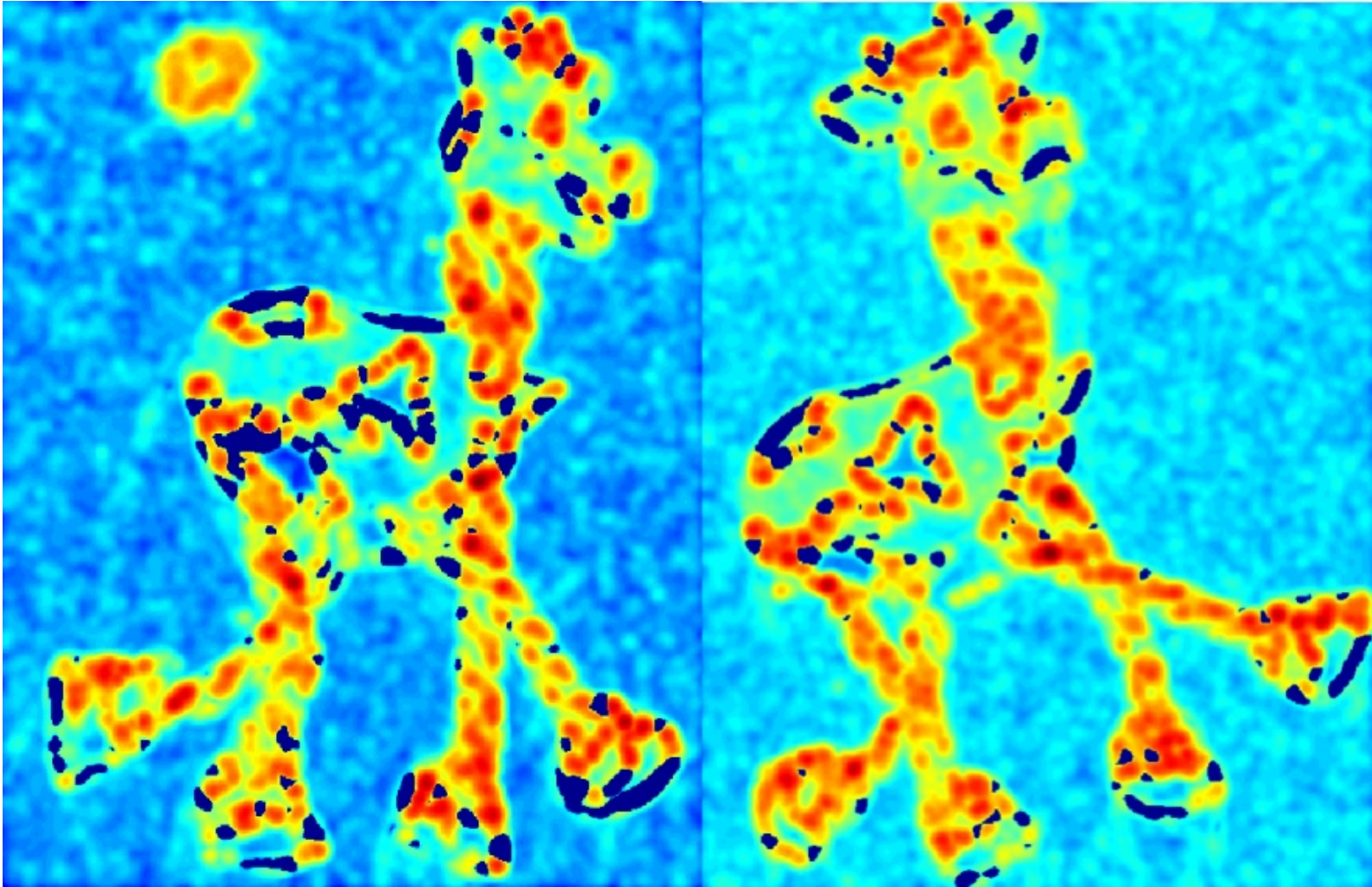


Пример



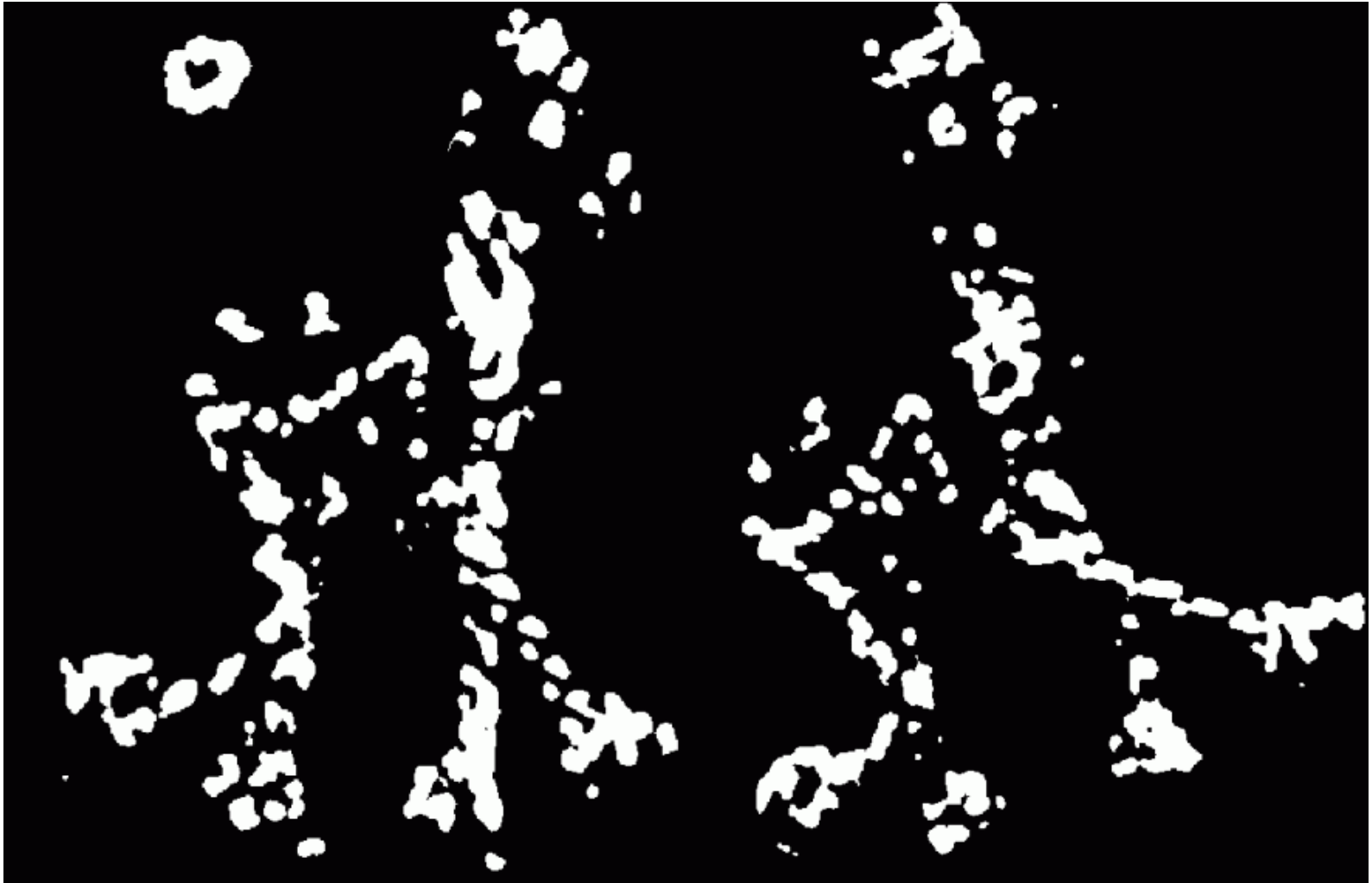
Пример

Вычисляем значение R



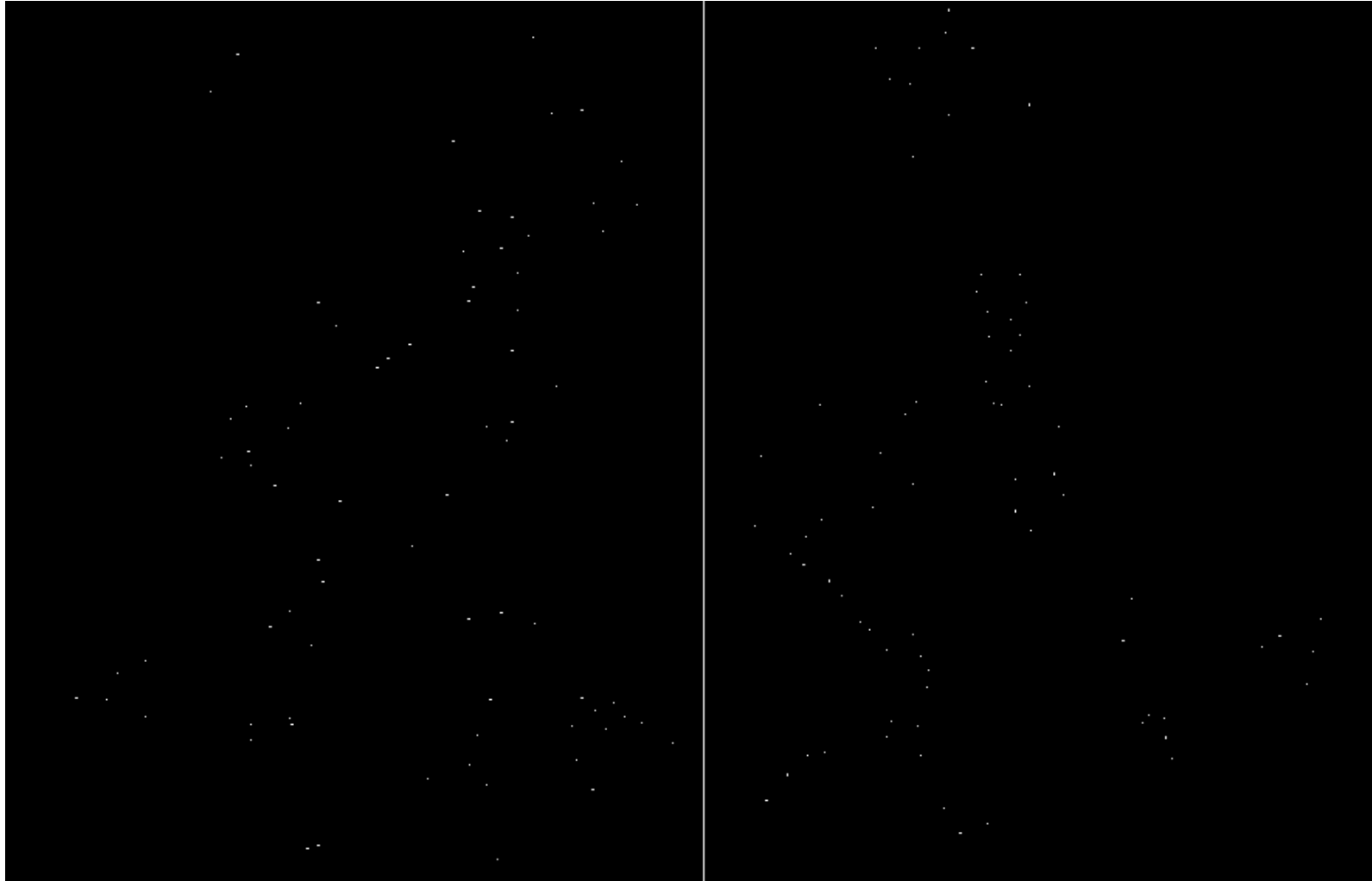
Пример

Выделяем точки с $R > \text{threshold}$



Пример

Оставим точки локального максимума R



Пример



Harris Detector: Summary

Average intensity change in direction $[u, v]$ can be expressed as a bilinear form:

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

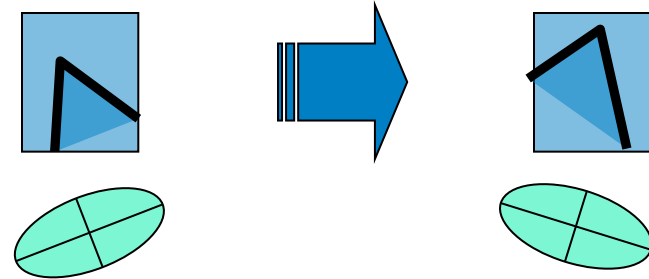
Describe a point in terms of eigenvalues of M :
measure of corner response

A good (corner) point should have a *large intensity change in all directions*,
i.e. R should be large positive

$$R = \lambda_1 \lambda_2 - k (\lambda_1 + \lambda_2)^2$$

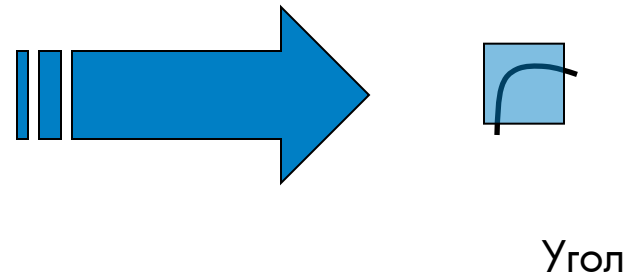
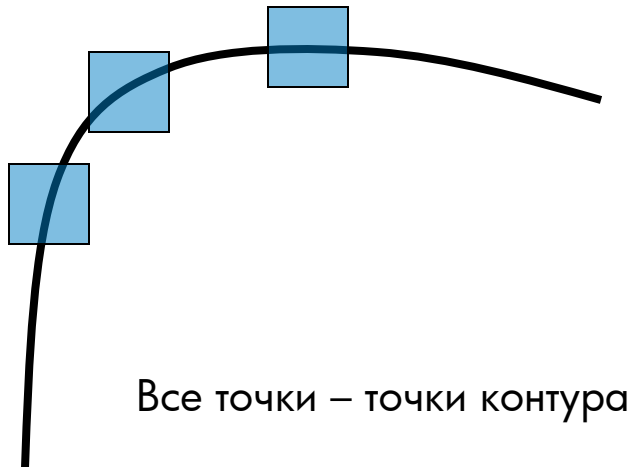
Свойства детектора Харриса

- Инвариантность к повороту



- Инвариантность к сдвигу интенсивности $I \rightarrow I + b$

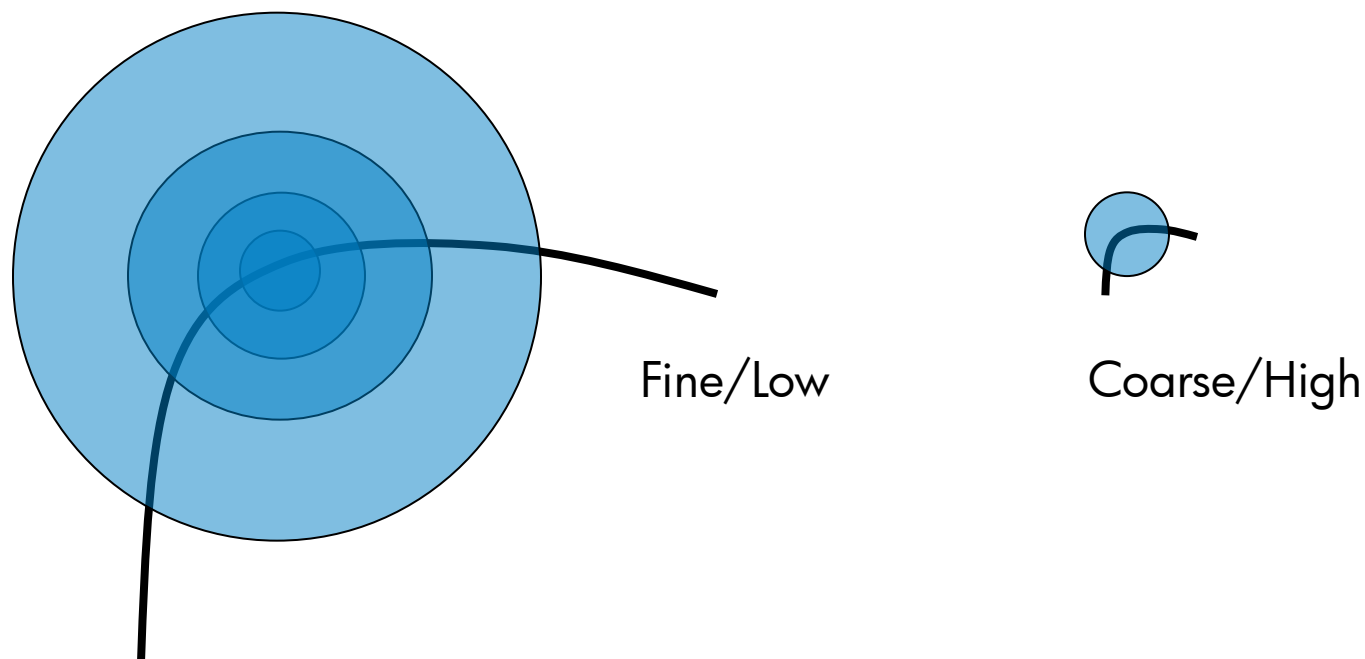
- Не инвариантен к изменению масштаба!



Инвариантность к изменению масштаба

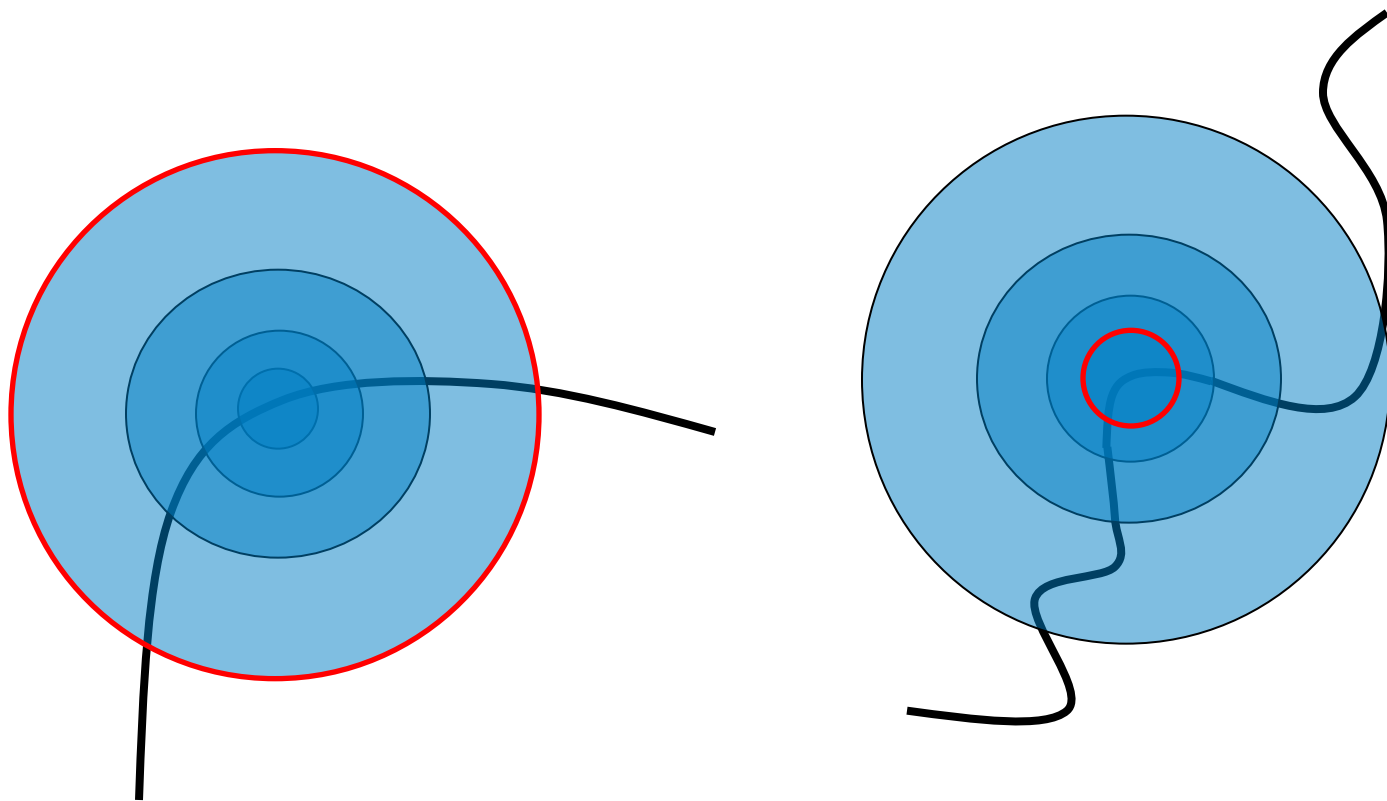
Рассмотрим фрагменты (например, круги) разного размера на разных изображениях

Для двух изображений найдутся соответствующие масштабы, при которых фрагменты будут выглядеть похоже



Инвариантность к изменению масштаба

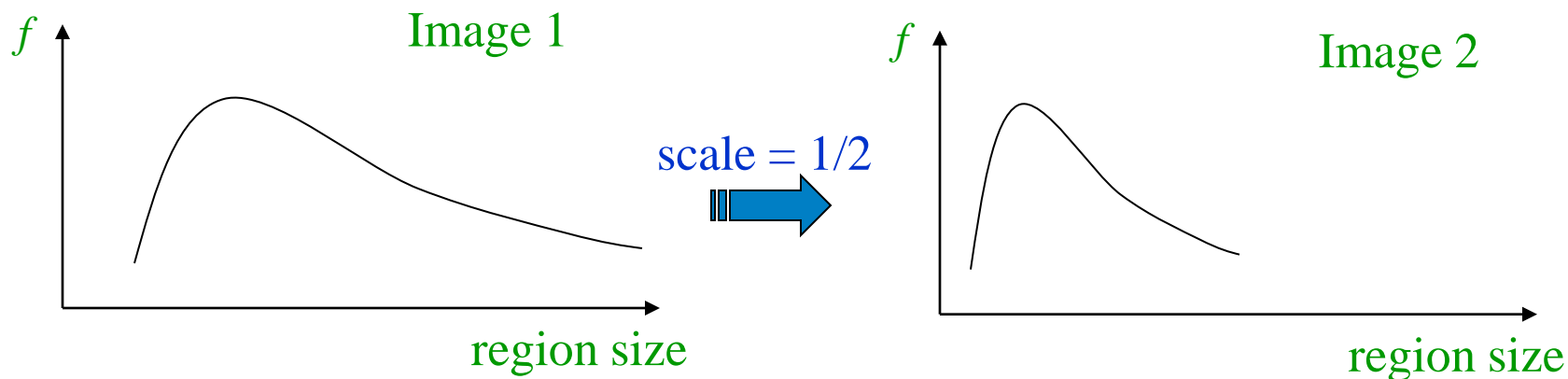
Задача: как выбрать размер фрагмента независимо для каждого изображения???



Инвариантность к изменению масштаба

Решение:

- Выберем функцию, заданную на фрагменте изображения, инвариантную к изменению масштаба – значение функции одинаково для сопоставимых фрагментов, даже если они разного масштаба (например, средняя интенсивность)
- В каждой точке изображения посмотрим на эту функцию, как на функцию от изменения размера фрагмента



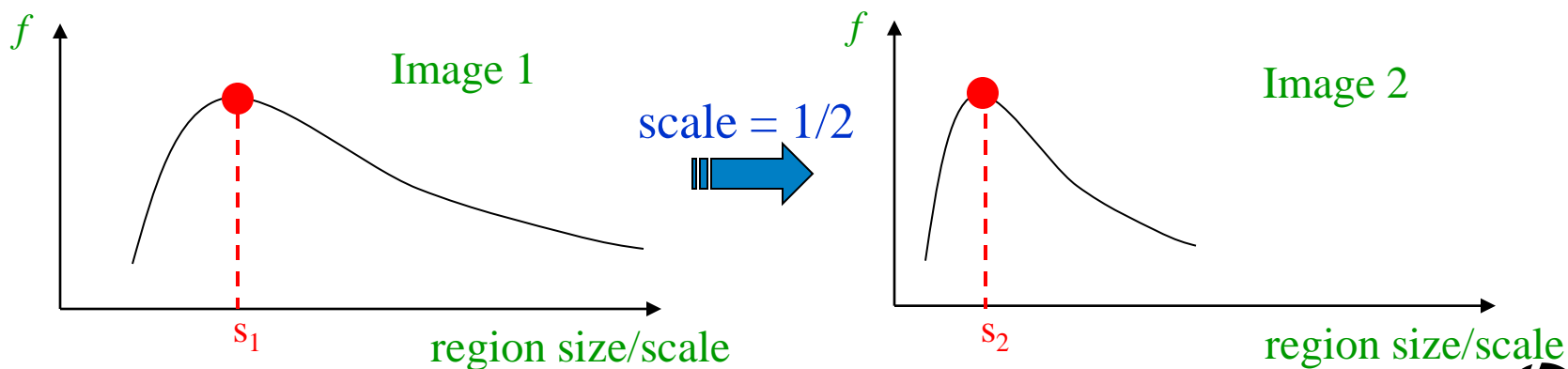
Инвариантность к изменению масштаба

Решение:

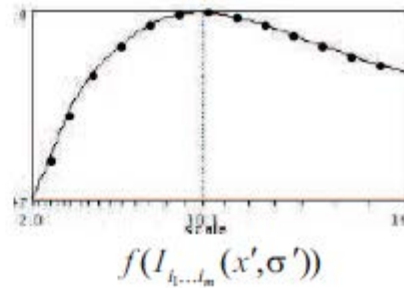
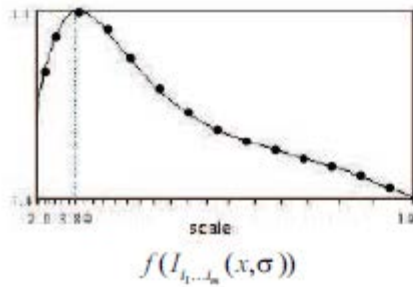
Найти локальный максимум такой функции

Наблюдение: точка локального максимума инвариантна к изменению масштаба

Важно: размер фрагмента, на котором достигается локальный максимум, находится для каждого изображения независимо!



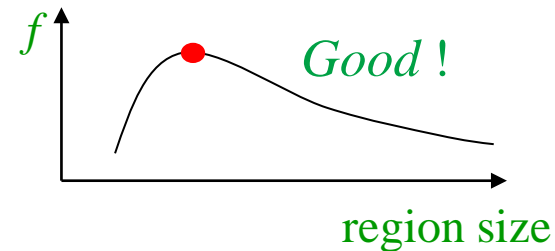
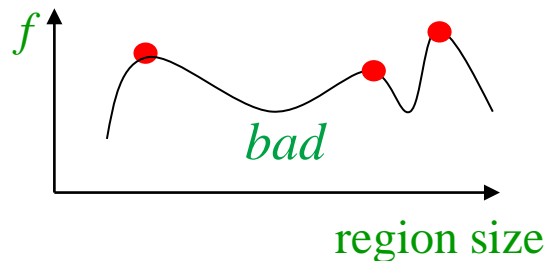
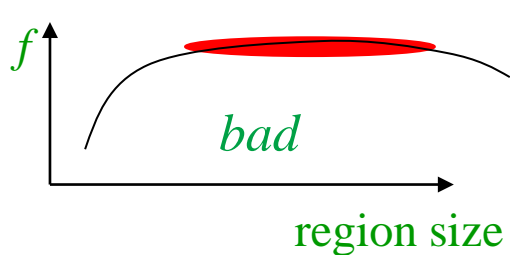
Characteristic Scale



- Точка локального максимума - characteristic scale
- $S1/S2$ равно соотношению масштабов между изображениями

Инвариантность к изменению масштаба

“Хорошая” функция для определения соотношения масштабов:



Scale Invariant Detection

Functions for determining scale

Kernels:

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

(Laplacian)

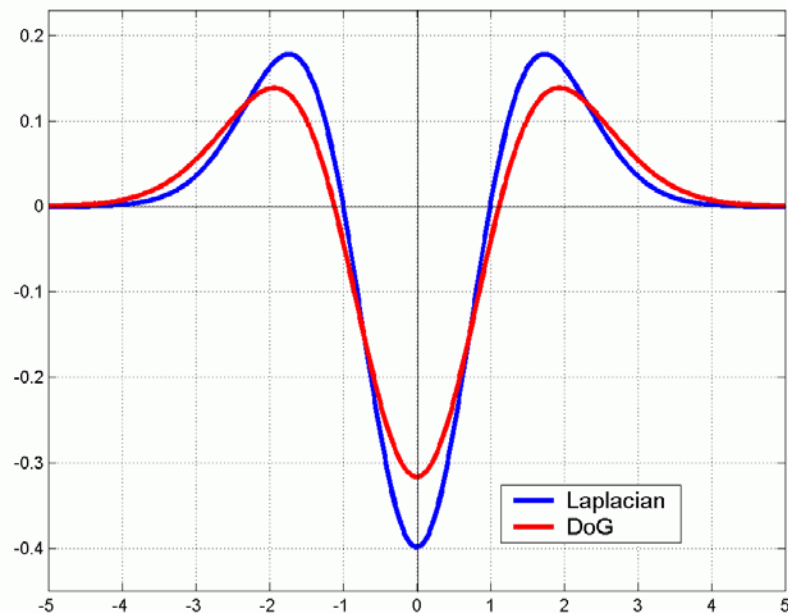
$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

(Difference of Gaussians)

where Gaussian

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

$$f = \text{Kernel} * \text{Image}$$



Note: both kernels are invariant to *scale* and *rotation*

Laplacian-of-Gaussian (LoG)

$$L(\mathbf{x}, \sigma) = \sigma^2 (I_{xx}(\mathbf{x}, \sigma) + I_{yy}(\mathbf{x}, \sigma))$$

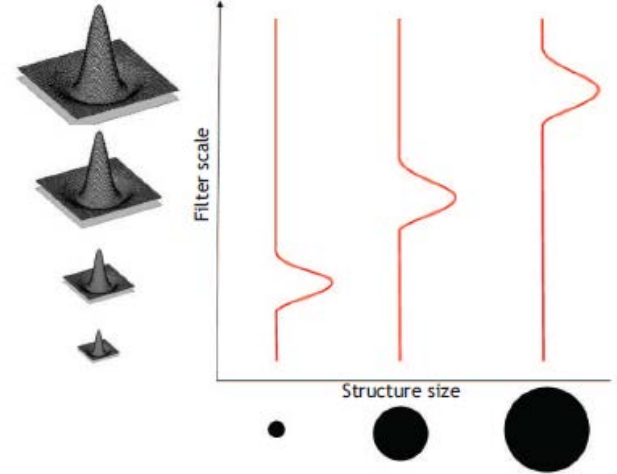
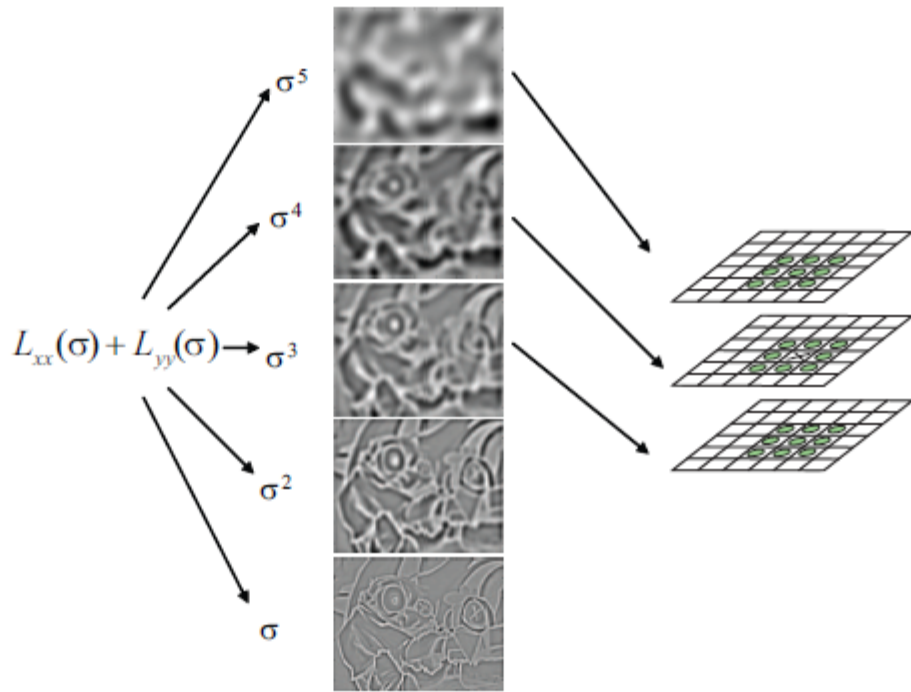
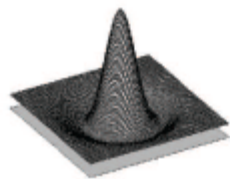


Figure 3.5: The Laplacian-of-Gaussian (LoG) detector searches for 3D scale space extrema of the LoG function.



Difference-of-Gaussian (DoG)

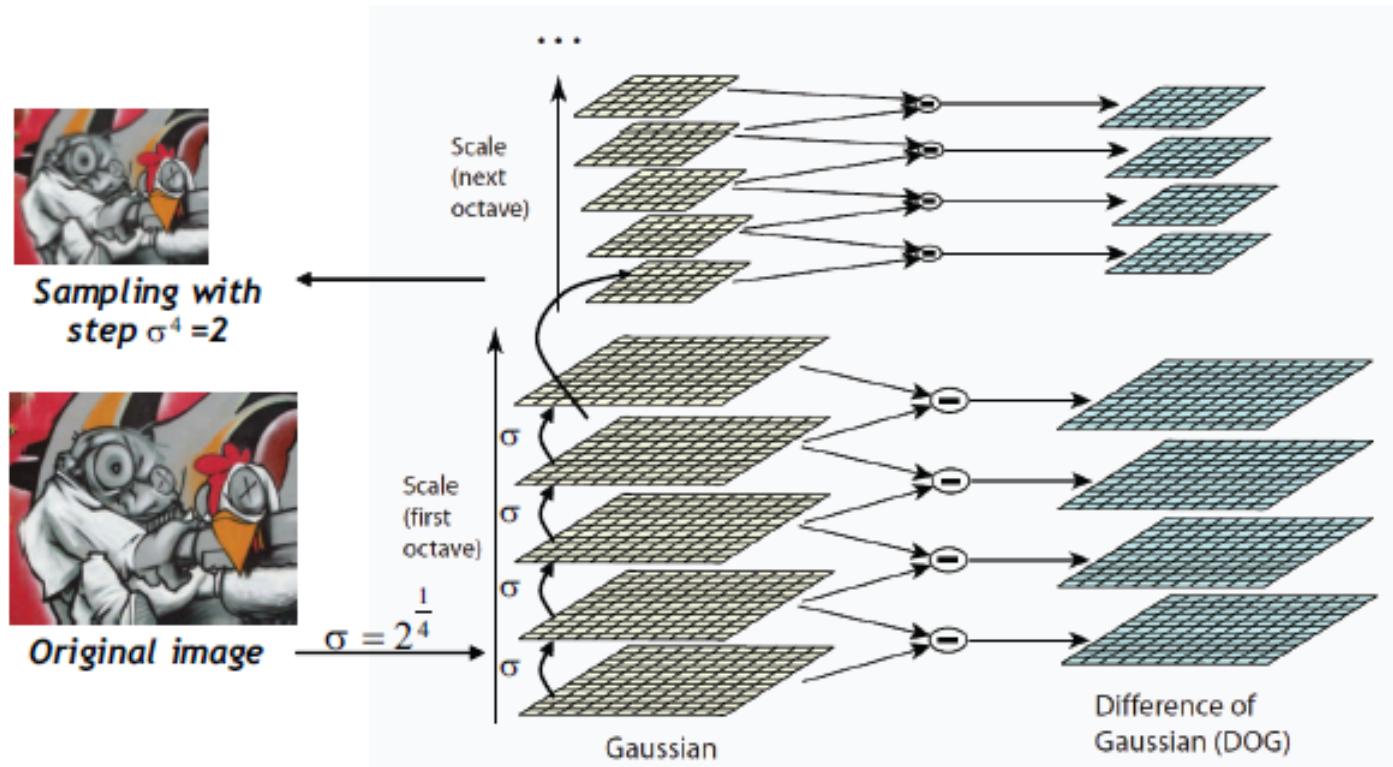
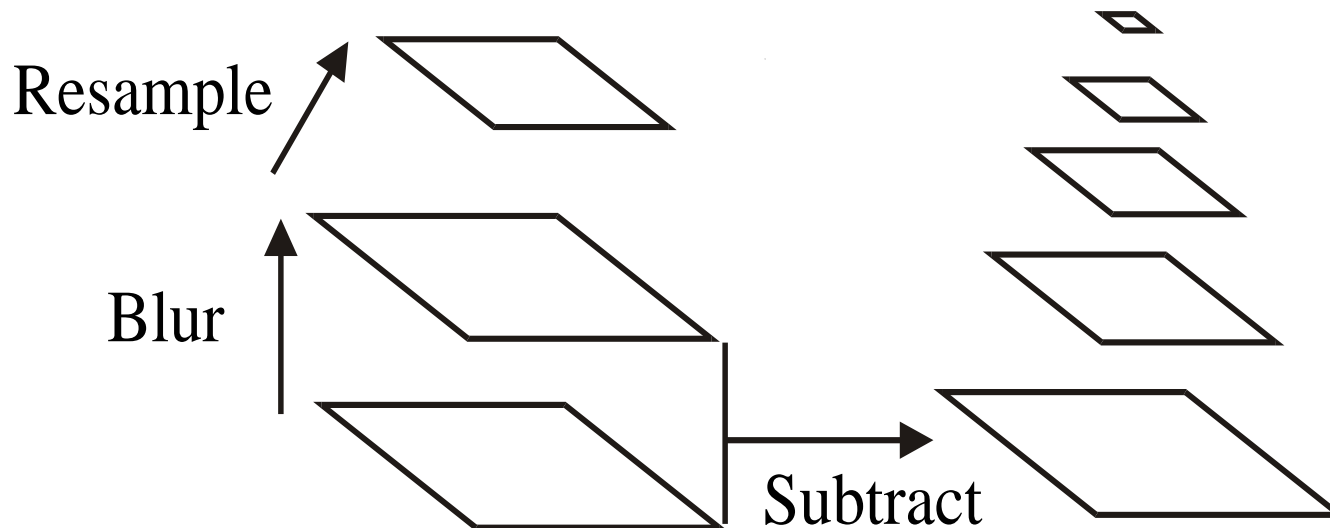


Figure 3.6: The Difference-of-Gaussian (DoG) provides a good approximation for the Laplacian-of-Gaussian. It can be efficiently computed by subtracting adjacent scale levels of a Gaussian pyramid. The DoG region detector then searches for 3D scale space extrema of the DoG function. *BL: Figure courtesy of [TM07]*

Build Scale-Space Pyramid

All scales must be examined to identify scale-invariant features
An efficient function is to compute the Difference of Gaussian (DOG) pyramid (Burt & Adelson, 1983) (or Laplacian)

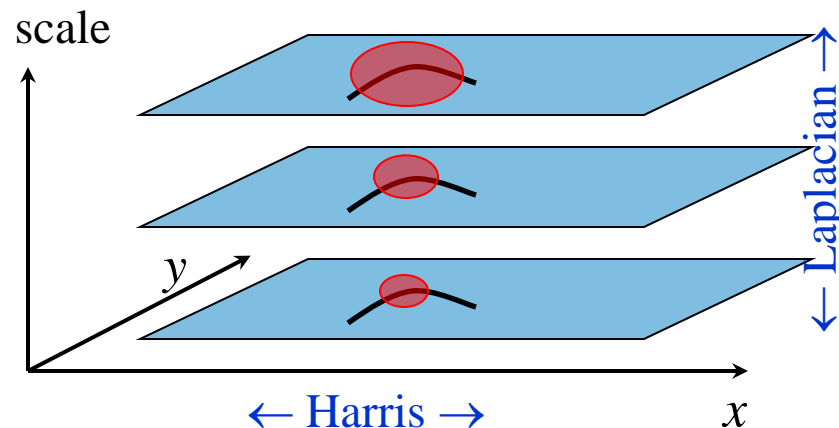


Key point localization

Harris-Laplacian¹

Find local maximum of:

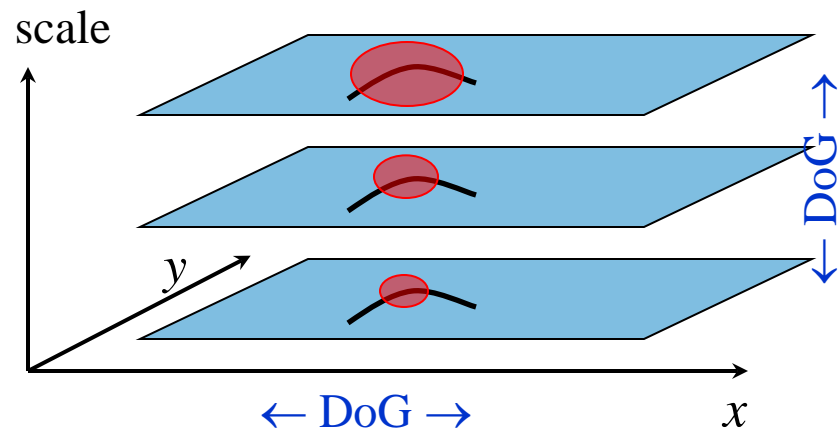
- Harris corner detector in space (image coordinates)
- Laplacian in scale



- **SIFT (Lowe)²**

Find local maximum of:

- Difference of Gaussians in space and scale



Scale Invariant Detection: Summary

Given: two images of the same scene with a large *scale difference* between them

Goal: find *the same* interest points *independently* in each image

Solution: search for *maxima* of suitable functions in *scale* and in *space* (over the image)

Methods:

1. **Harris-Laplacian** [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. **SIFT** [Lowe]: maximize Difference of Gaussians over scale and space



Локальные признаки

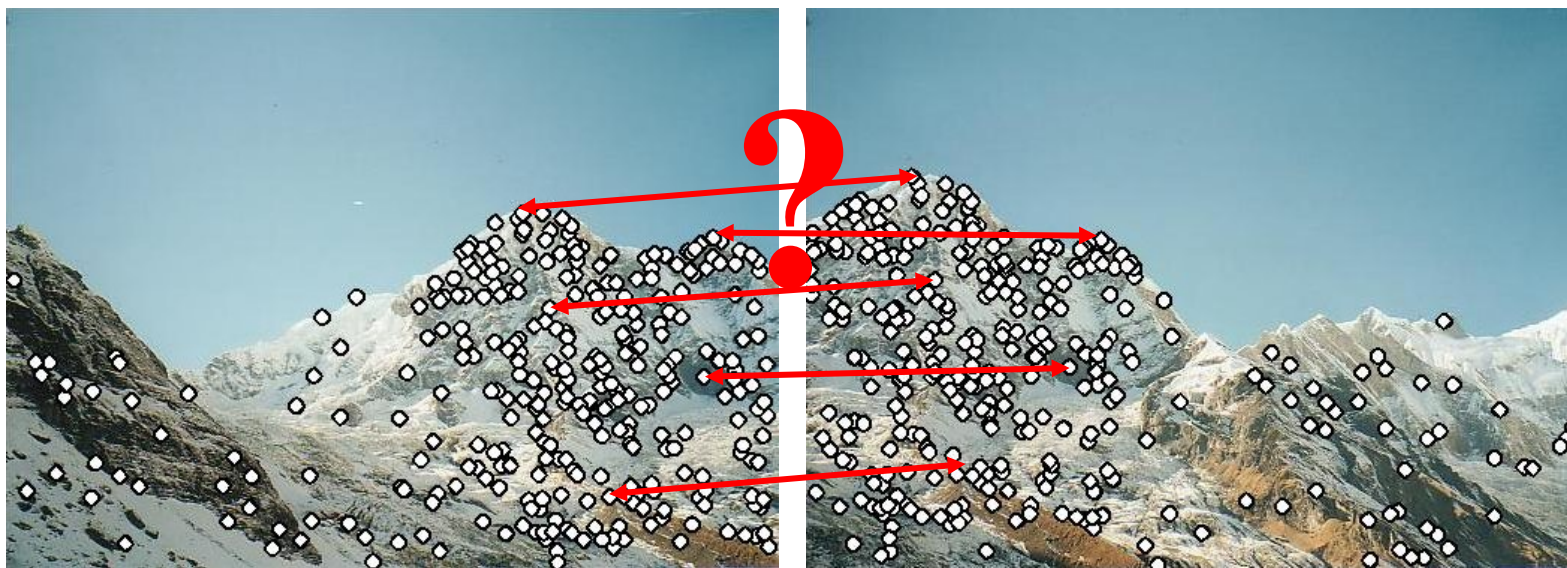
We know how to detect points

Next question:

How to match them?

Point descriptor should be:

1. Invariant
2. Distinctive

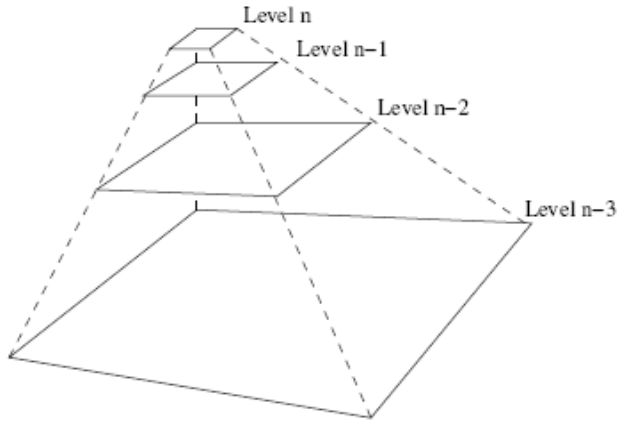


SIFT: main steps

1. Scale-space peak selection
 - Using Difference-of-Gaussians (DoG)
2. Keypoint localization
 - Elimination of unstable keypoints
3. Orientation assignment
 - Based on keypoint local image patch
4. Keypoint descriptor
 - Based upon the image gradients in keypoint local neighbourhood



Scale space



Build an image pyramid with resampling between each level



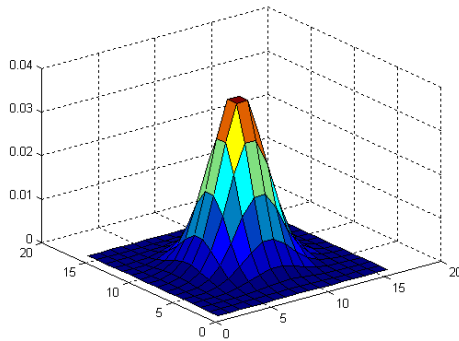
Difference-of-Gaussian



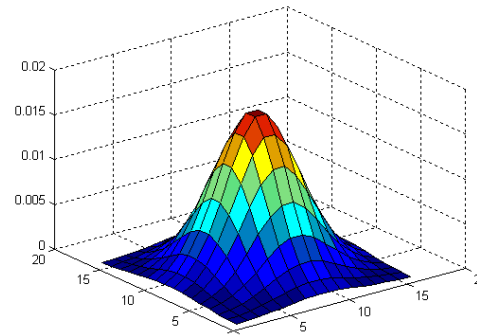
The input image is convolved with Gaussian function:

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$$

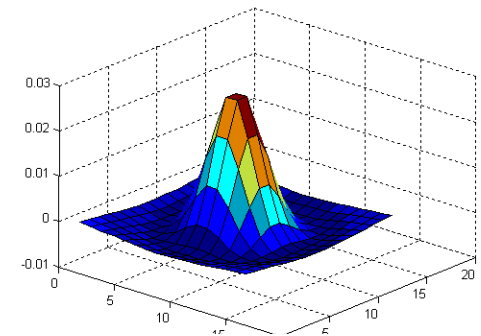
$\sigma_1 = 2$



$\sigma_2 = 4$



$\sigma_1 - \sigma_2$



Difference-of-Gaussian

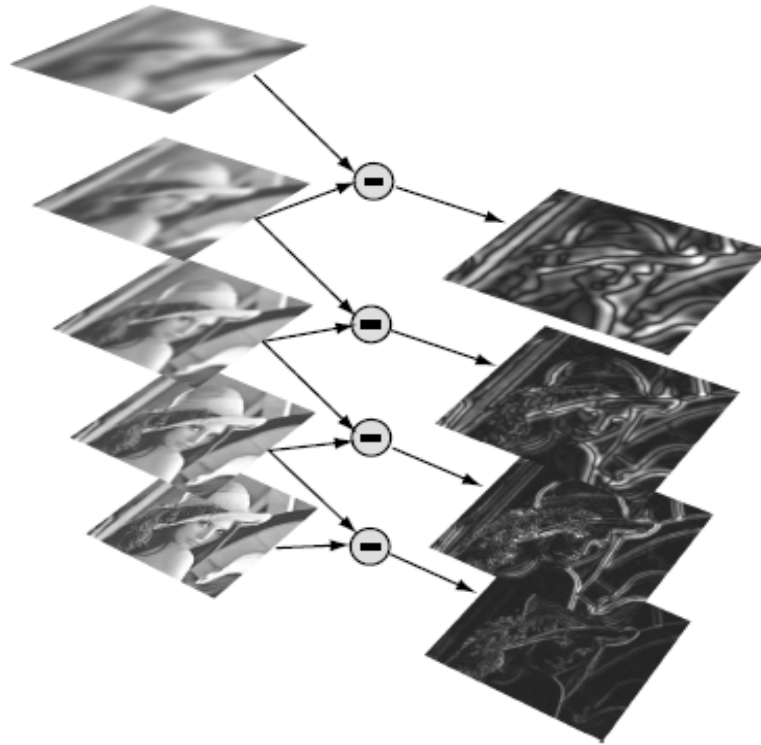
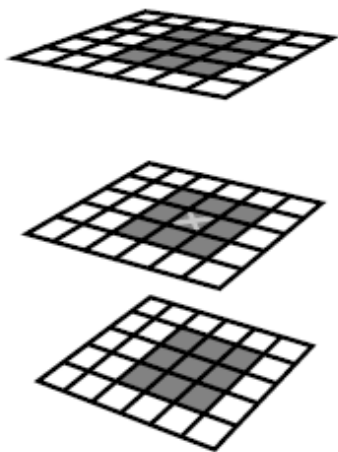


Figure 9.1: A Difference-of-Gaussian octave. The five images in the left stack are incrementally smoothed versions of the input image. The right stack shows the resulting DoG.

SIFT keypoints

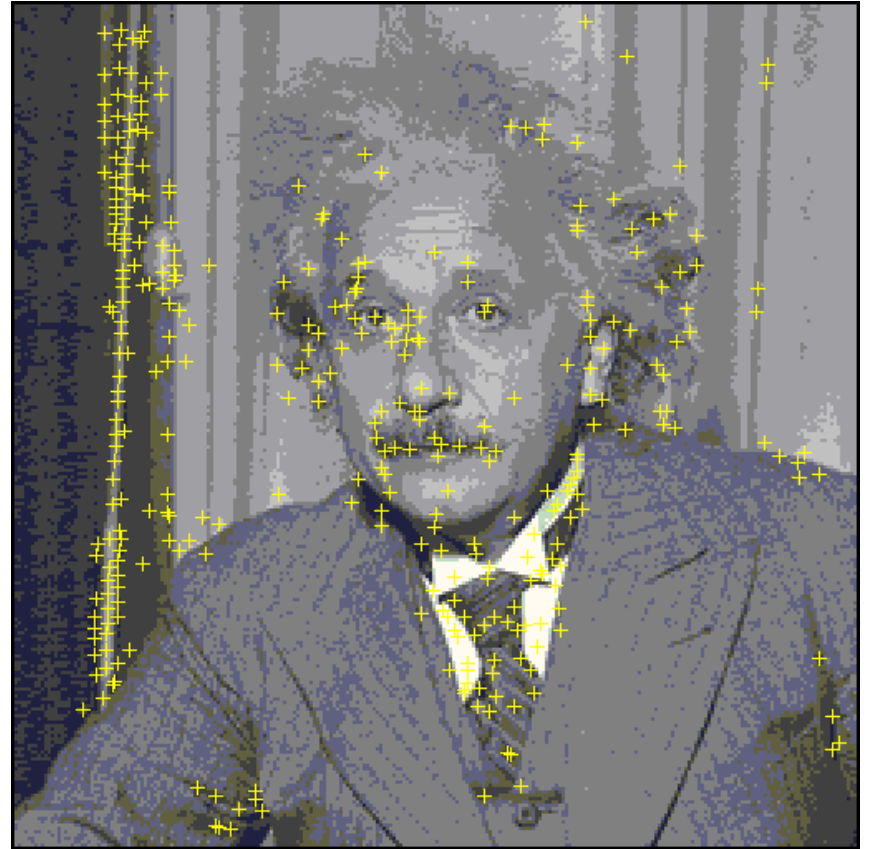
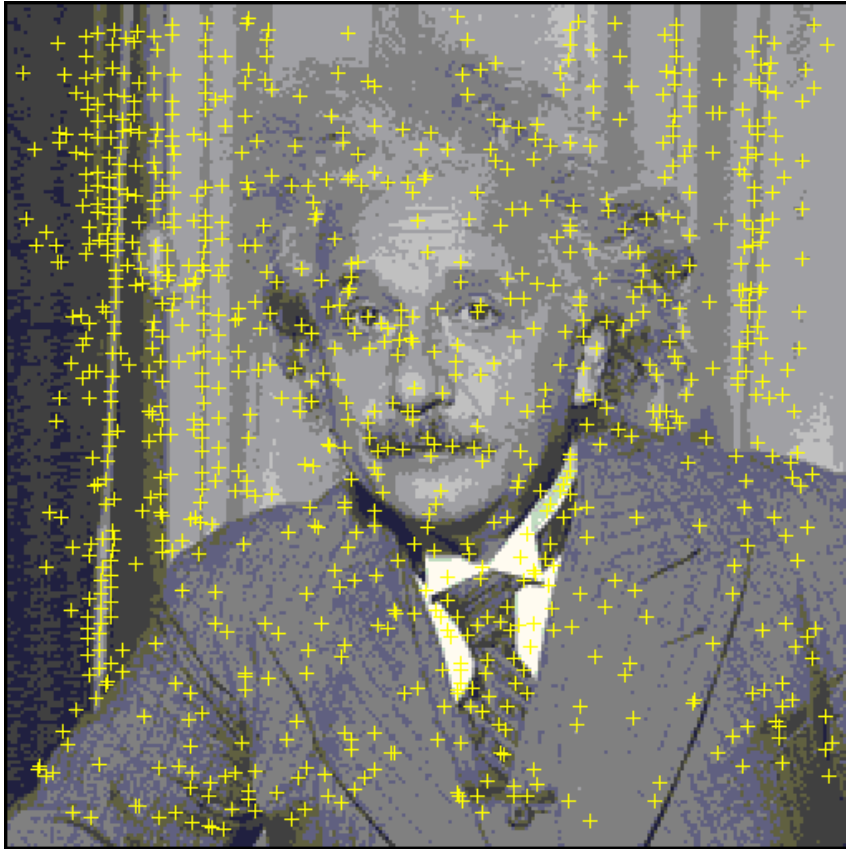


Maxima and minima of DoG applied in scale-space:

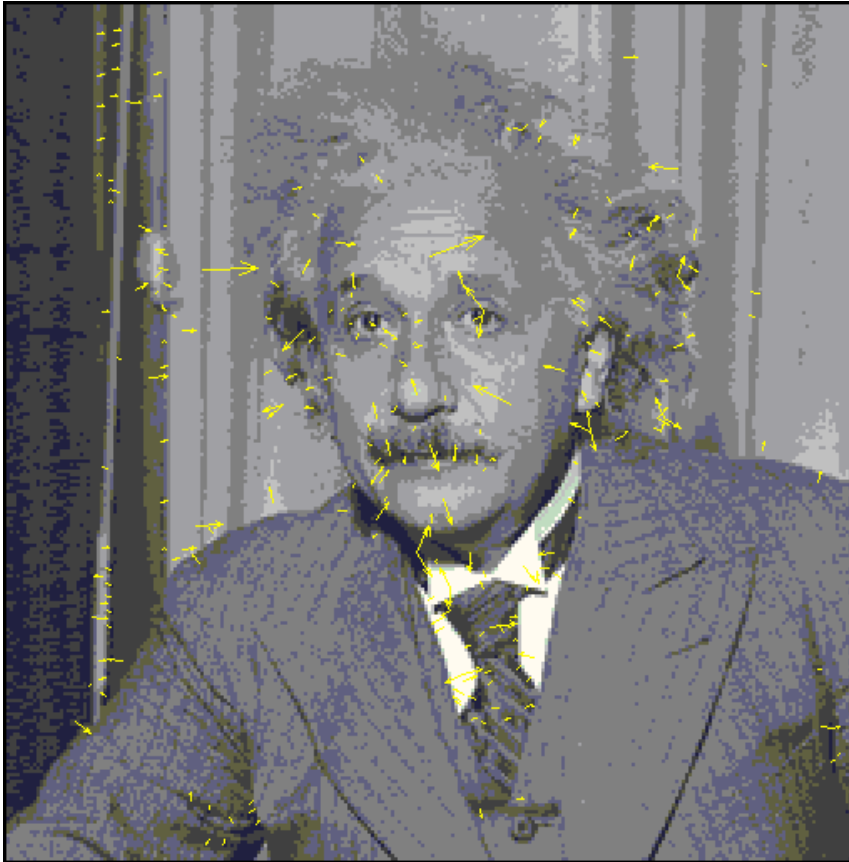
- 1) Extrema detection for the same scale
- 2) Check if it is stable for different scales



Scale-space extrema detection



Keypoints orientation and scale



- Extract image gradients and orientations at each pixel
- Each key location is assigned a canonical orientation
- The orientation is determined by the peak in a histogram of local image gradient orientations

$$M_{ij} = \sqrt{(A_{ij} - A_{i+1,j})^2 + (A_{ij} - A_{i,j+1})^2}$$

$$R_{ij} = \text{atan2}(A_{ij} - A_{i+1,j}, A_{i,j+1} - A_{ij})$$

SIFT descriptor

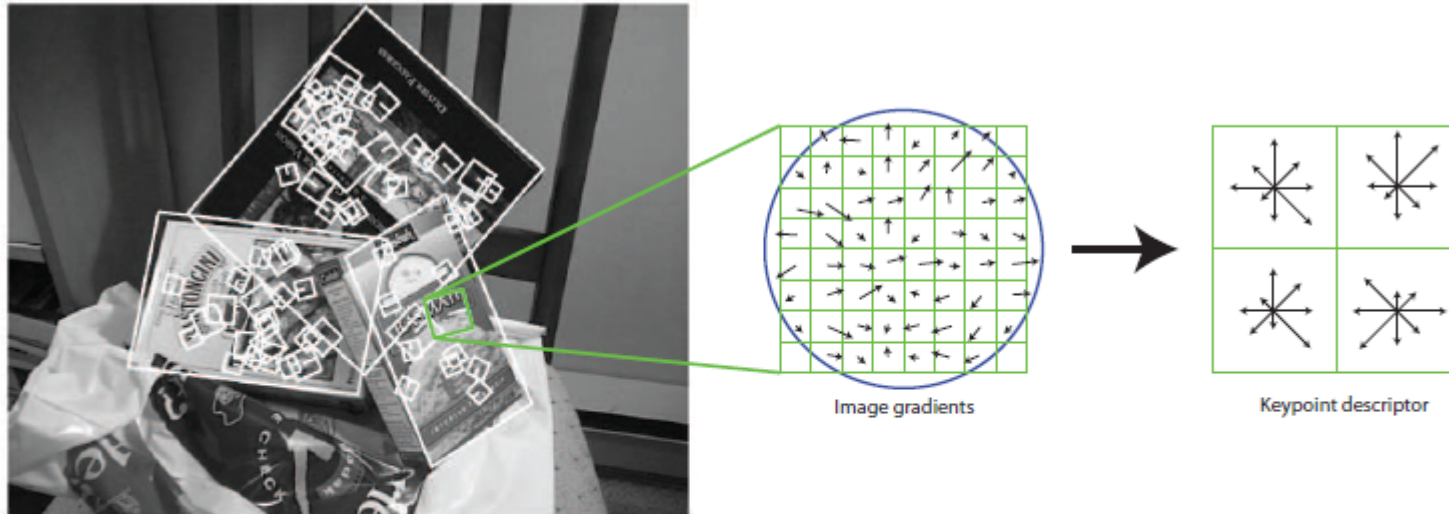
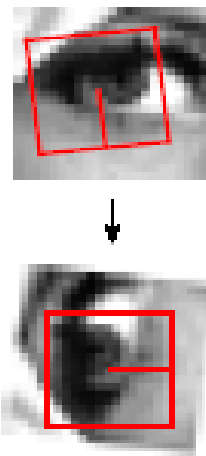


Figure 3.8: Visualization of the SIFT descriptor computation. For each (orientation-normalized) scale invariant region, image gradients are sampled in a regular grid and are then entered into a larger 4×4 grid of local gradient orientation histograms (for visibility reasons, only a 2×2 grid is shown here).

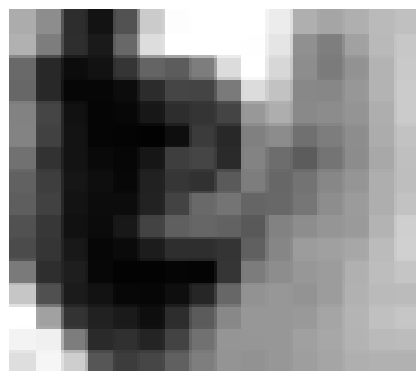
Example



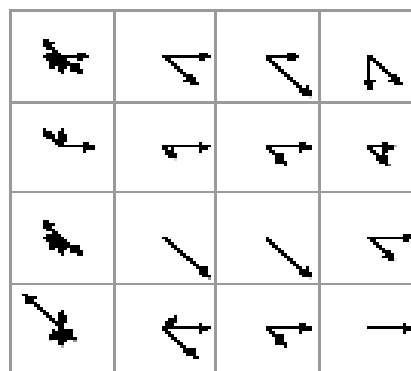
(a) A subset of the extracted interest points, and the associated regions used to create descriptors.



(b) An interest point region covering Lena's eye before and after rotation in respect to the reference orientation of the point of interest.



Interest point region



Interest point descriptor

(c) Computation of a descriptor by determining a 4×4 gradient orientation histogram array from a 16×16 pixels region around the interest point location.

Заключение

- Глобальные признаки не всегда работают
- Вычисление локальных признаков
 - Определение особых точек
 - Выделение особых фрагментов – инвариантность к масштабу!
 - Построение векторов признаков
 - Сопоставление локальных дескрипторов пары изображений

