

Анализ изображений и видео.

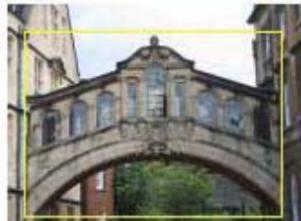
Часть 2

Локальные признаки

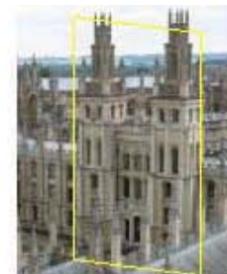
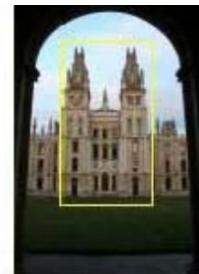
Дескрипторы изображений

- Глобальные
 - Вычисляются по всему изображению
- Локальные
 - Вычисляются по окрестностям ключевых точек на изображении.

Глобальные признаки иногда работают плохо.



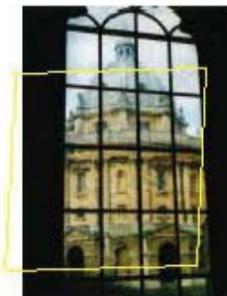
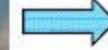
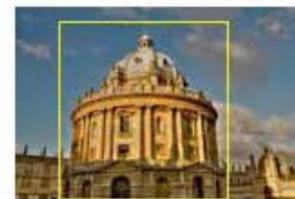
Изменение масштаба



Изменение точки съемки



Изменение освещения



Перекрытие

Свойства особых точек

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

Повторяемость особых точек

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

Информативность, репрезентативность

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

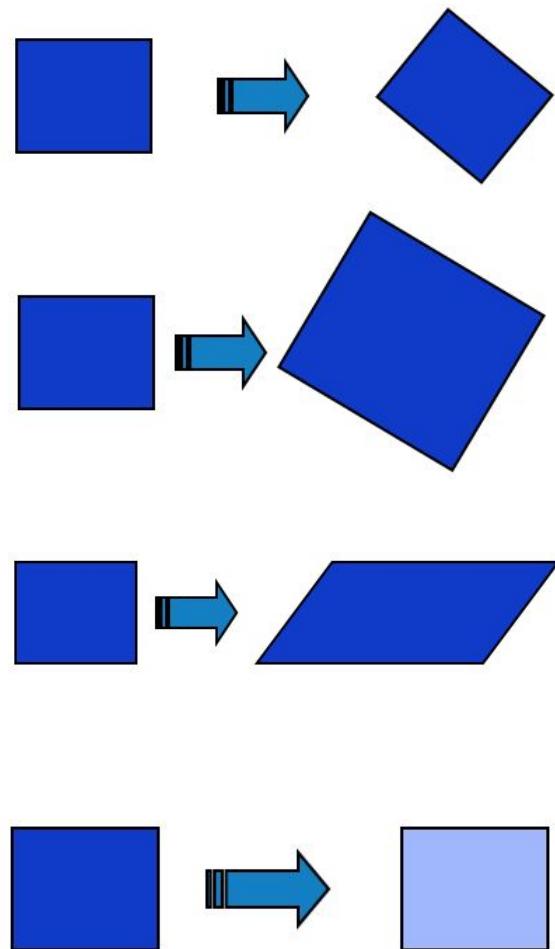
Геометрические и фотометрические трансформации изображения.

- Геометрические:

- Поворот
 - Поворот + изменение масштаба
 - Афинные преобразования

- Фотометрические

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



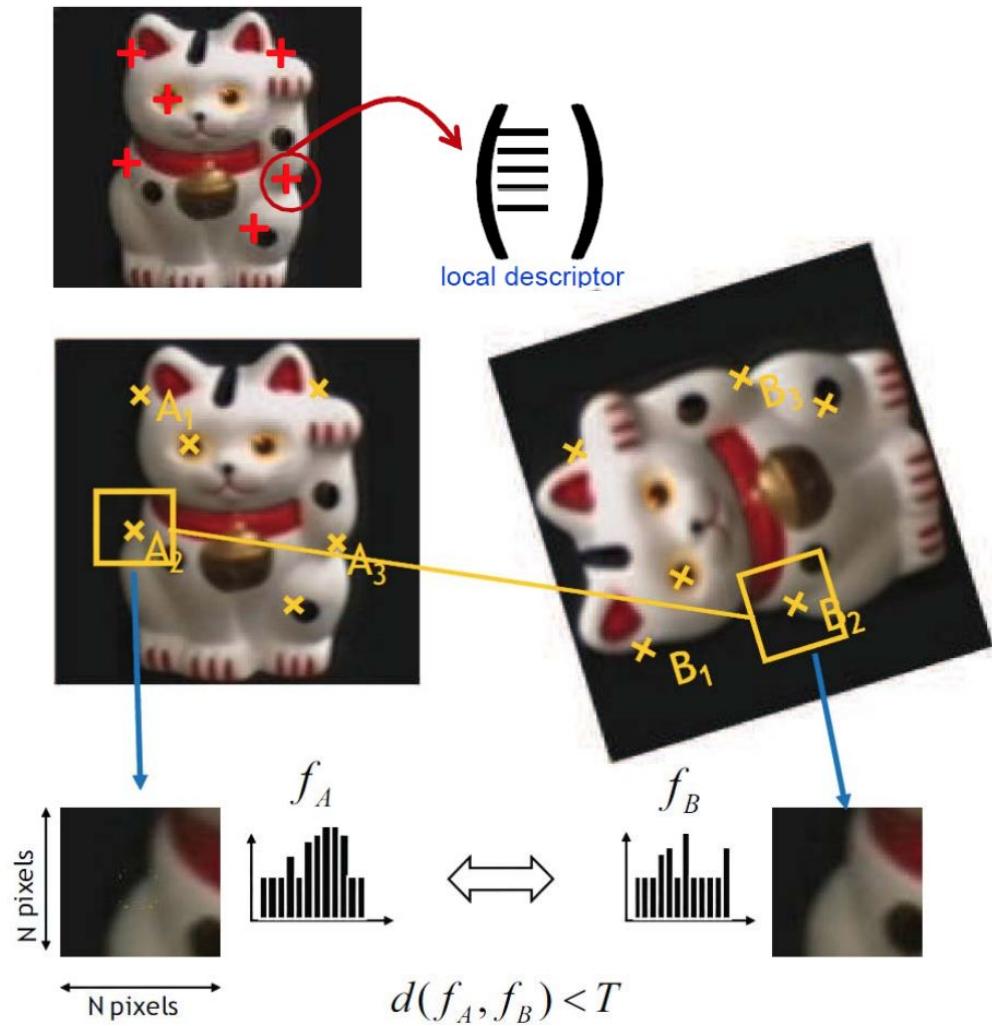
Сравнение изображений при помощи локальных признаков: основные шаги

1. Локализация особых точек

2. Выделение особых фрагментов – окрестности ключевых точек, инвариантные к различного рода преобразованиям

3. Построение векторов признаков для найденных фрагментов

4. Сопоставление наборов локальных признаков для двух изображений

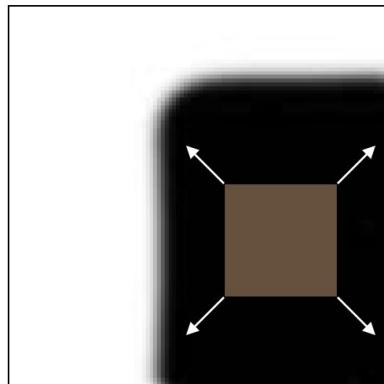


Угловые точки, в качестве особых.

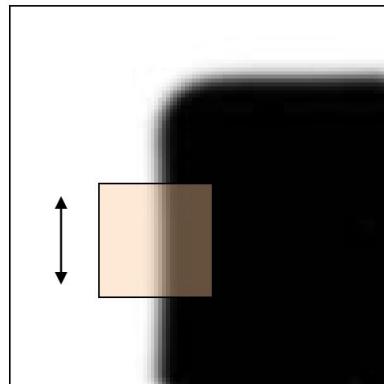
- Угловые точки хорошо локализованы на объекте.
- Угловые точки легко найти
- Семантически значимые участки изображения (края краев =))

Детектор угловых точек.

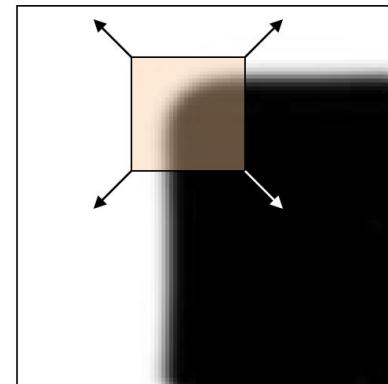
- Просто обнаружить, используя маленькое окно
- Смещение окна в окрестности угловой точки в любом направлении приводит к существенному изменению набора интенсивностей окна (алгоритм Моравица)
- В окрестности угловой точки – два доминирующих направления градиента (алгоритм Харриса)



“монотонная” область:
нет изменений ни в
каком направлении

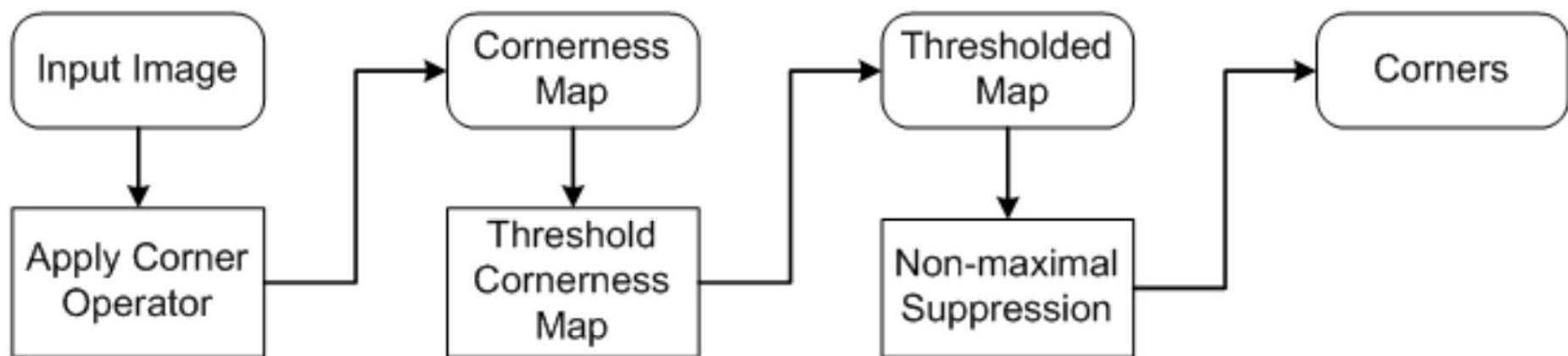


“край”:
нет изменений по одному
из направлений – вдоль
края

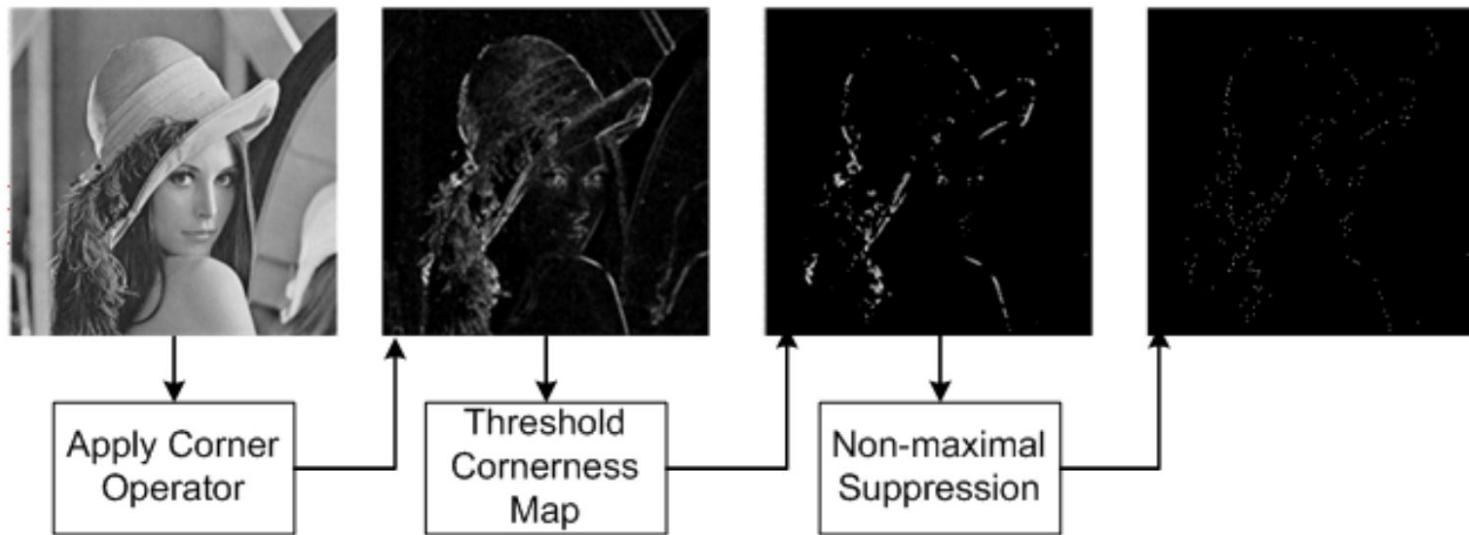


“угол”:
существенные
изменения по всем
направлениям

Схема работы детекторов угловых точек

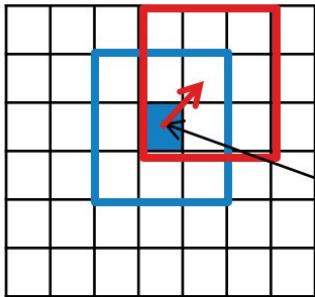


Пример схемы работы детектора угловых точек



Corners superimposed on
Input Image

Детектор Моравица (Moravec)



(u, v) – направление смещения (всего 8 направлений)

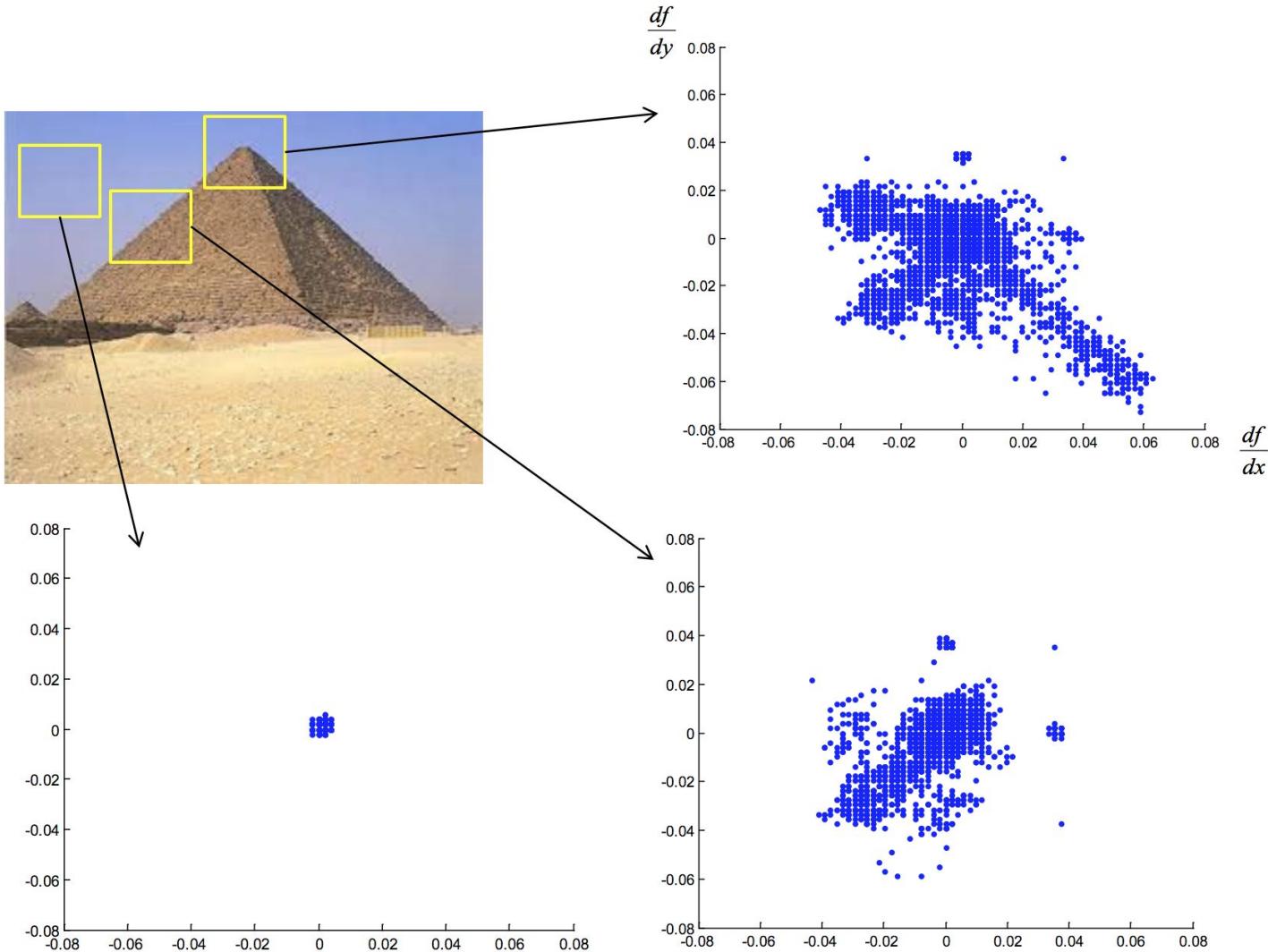
(x, y) – центр окна, размер окна $(2n+1) \times (2n+1)$, $n=1, 2, 4$

- Для каждого направления смещения вычисляется изменение интенсивности:
$$S(x, y, u, v) = \sum_{a=x-n, b=y-n}^{a=x+n, b=y+n} (I(a, b) - I(a + u, b + v))^2$$
- Строится карта силы угла в каждой точке:
$$C(x, y) = \min_{u, v} \{S(x, y, u, v)\}$$
- Отсекаются неугловые точки по порогу (значение силы угла меньше порога)
- Отсекаются точки, не являющиеся локальным максимумом (non-maximal suppression)

Недостатки детектора Моравица

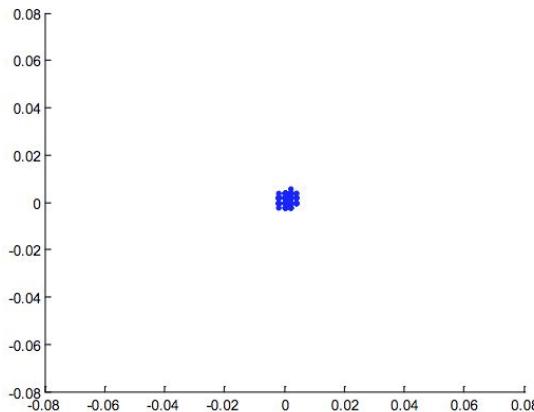
1. Не является инвариантным к повороту из-за дискретности рассматриваемого множества направлений смещений
2. Выдает высокий отклик вдоль при наличии даже небольшого шума
3. Неточен в оценке локального изменения интенсивности из-за использования квадратной бинарной маски окна

Детектор Харриса

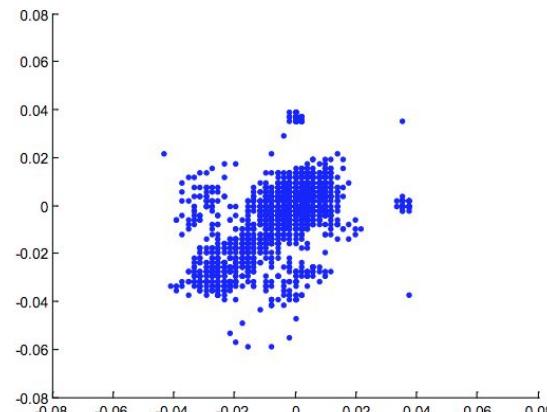


Детектор Харриса – основная идея

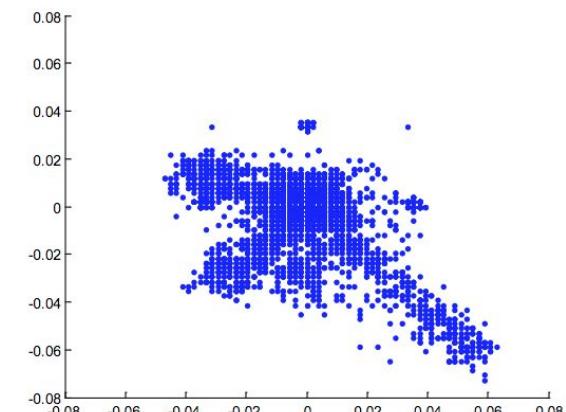
- Для каждого окна смотрим на собственные числа ковариационной матрицы градиентов пикселей
- Собственные вектора этой матрицы показывают «основные направления» данных
- Величина собственных чисел указывает на «степень выраженности» соответствующего направления



Нет «больших» собственных чисел



1 «большое» собственное число



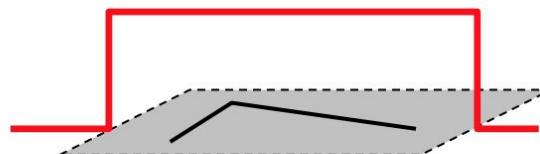
2 «большых» собственных числа

Детектор Харриса

Изменение интенсивности в окрестности точки (x,y) при сдвиге $[u,v]$:

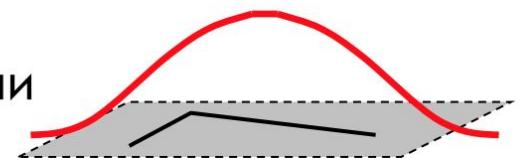
$$E(u, v) = \sum_{x, y} w(x, y) [I(x + u, y + v) - I(x, y)]^2$$


Функция окна $w(x,y) =$



1 в окне, 0 снаружи

или



Гауссиан

Детектор Харриса

Для небольших сдвигов $[U, V]$ можем аппроксимировать (ряд Тейлора):

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

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:

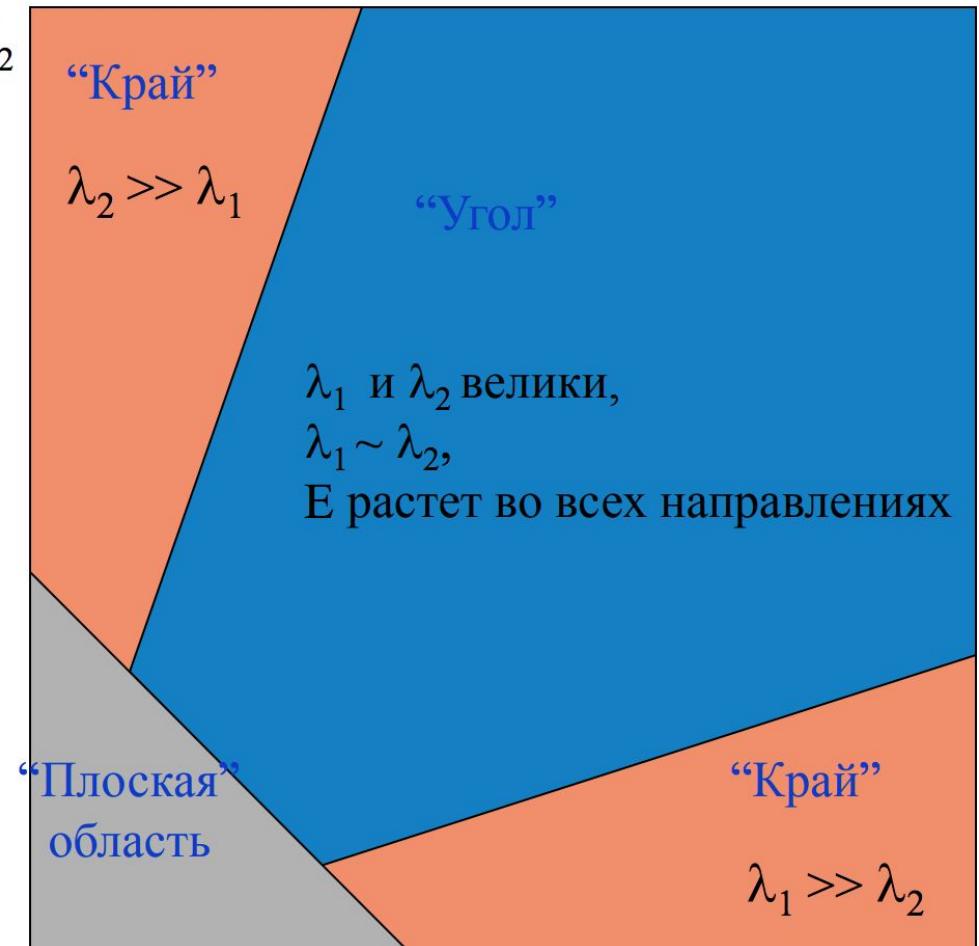
$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Детектор Харриса

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

λ_1, λ_2 – собственные значения M

λ_1 и λ_2 малы;
 E почти постоянна по
всем направлениям



Детектор Харриса: мера отклика

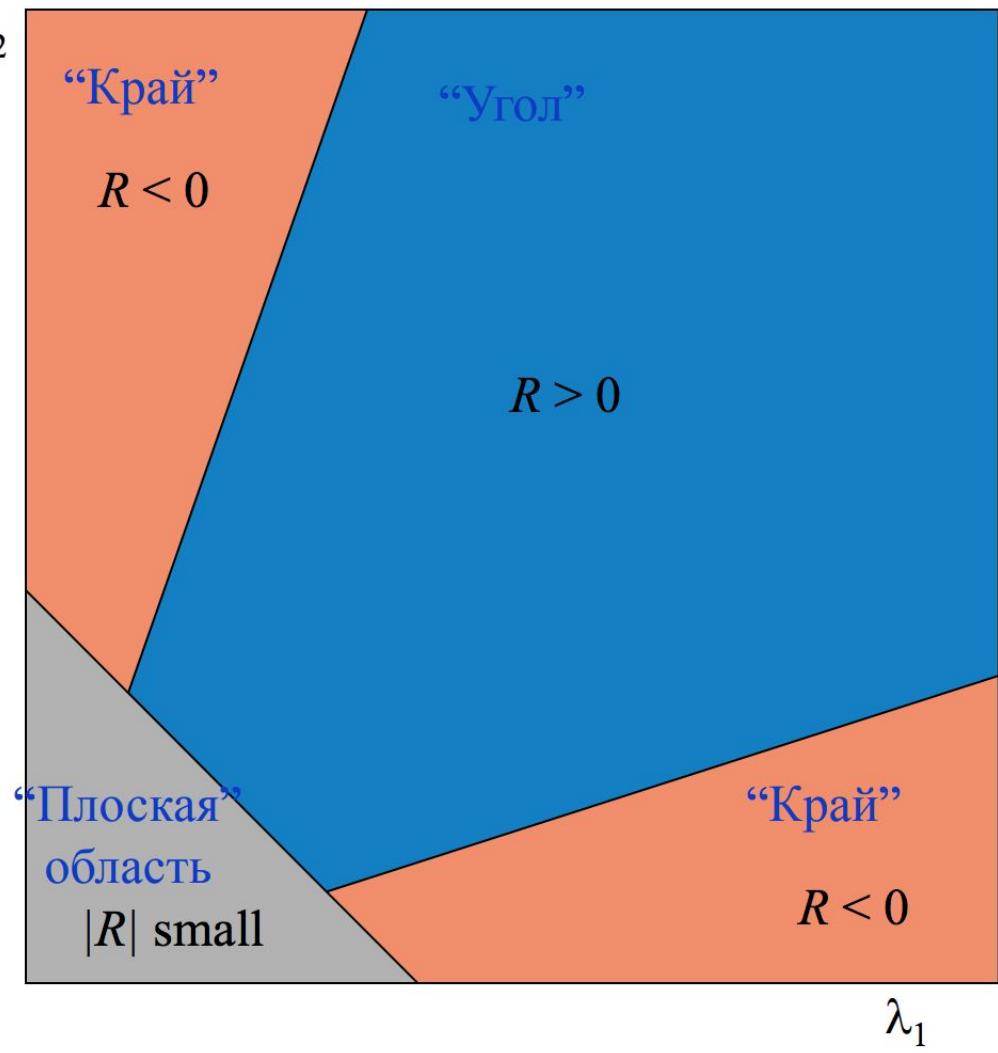
$$R = \det M - k (\operatorname{trace} M)^2$$

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

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

(k – эмпирическая константа,
 $k = 0.04\text{-}0.06$)

- 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



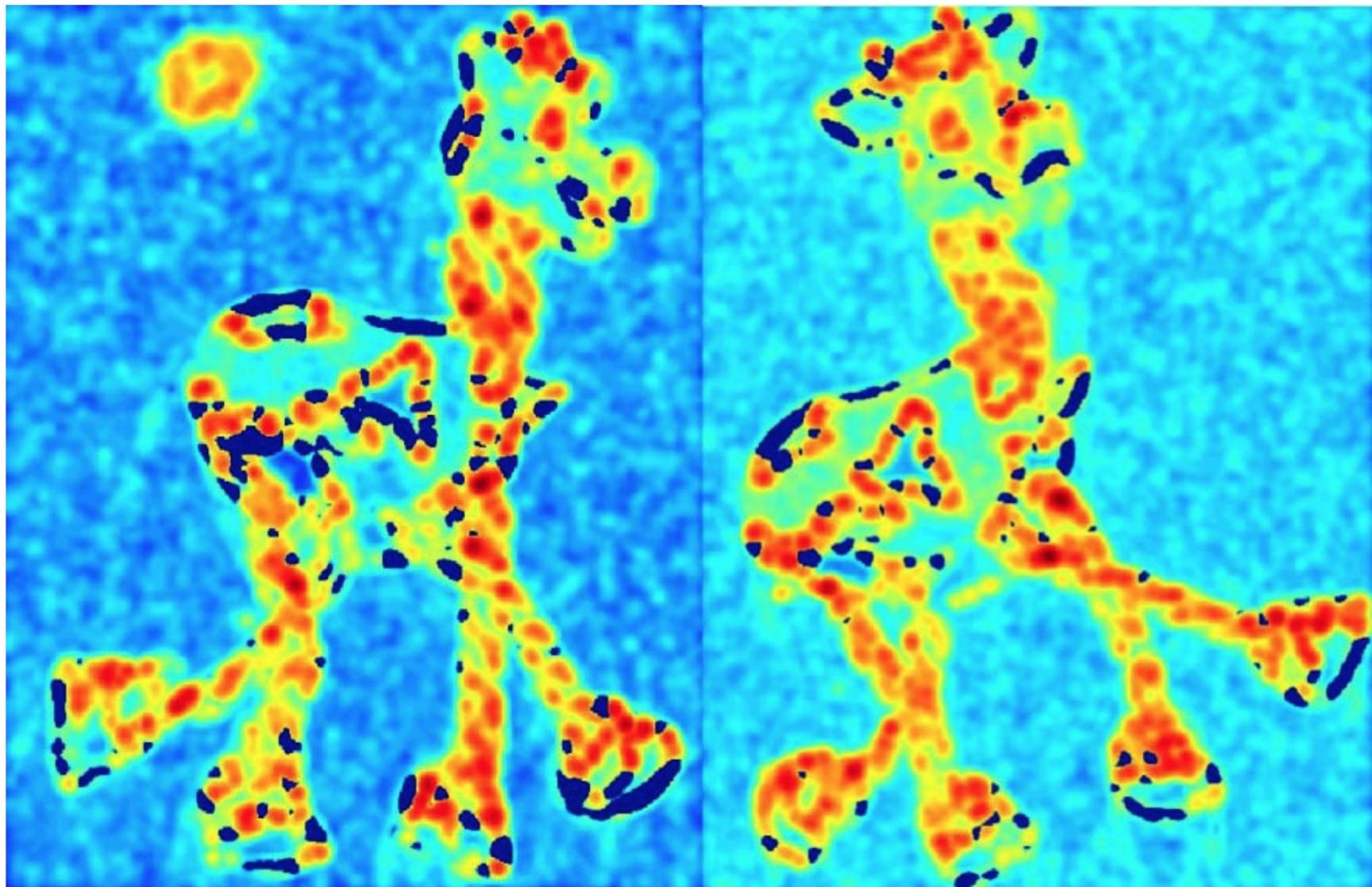
Детектор Харриса: алгоритм

1. Вычислить градиент изображения в каждой точке с использованием Гауссова сглаживания
2. Вычислить матрицу M с весовыми коэффициентами Гауссова окна для окрестности каждой точки
3. Вычислить меру отклика угла R
4. Найти точки с большим значением R (отсечение по порогу)
5. Найти локальные максимумы меры отклика угла (nonmaximum suppression)

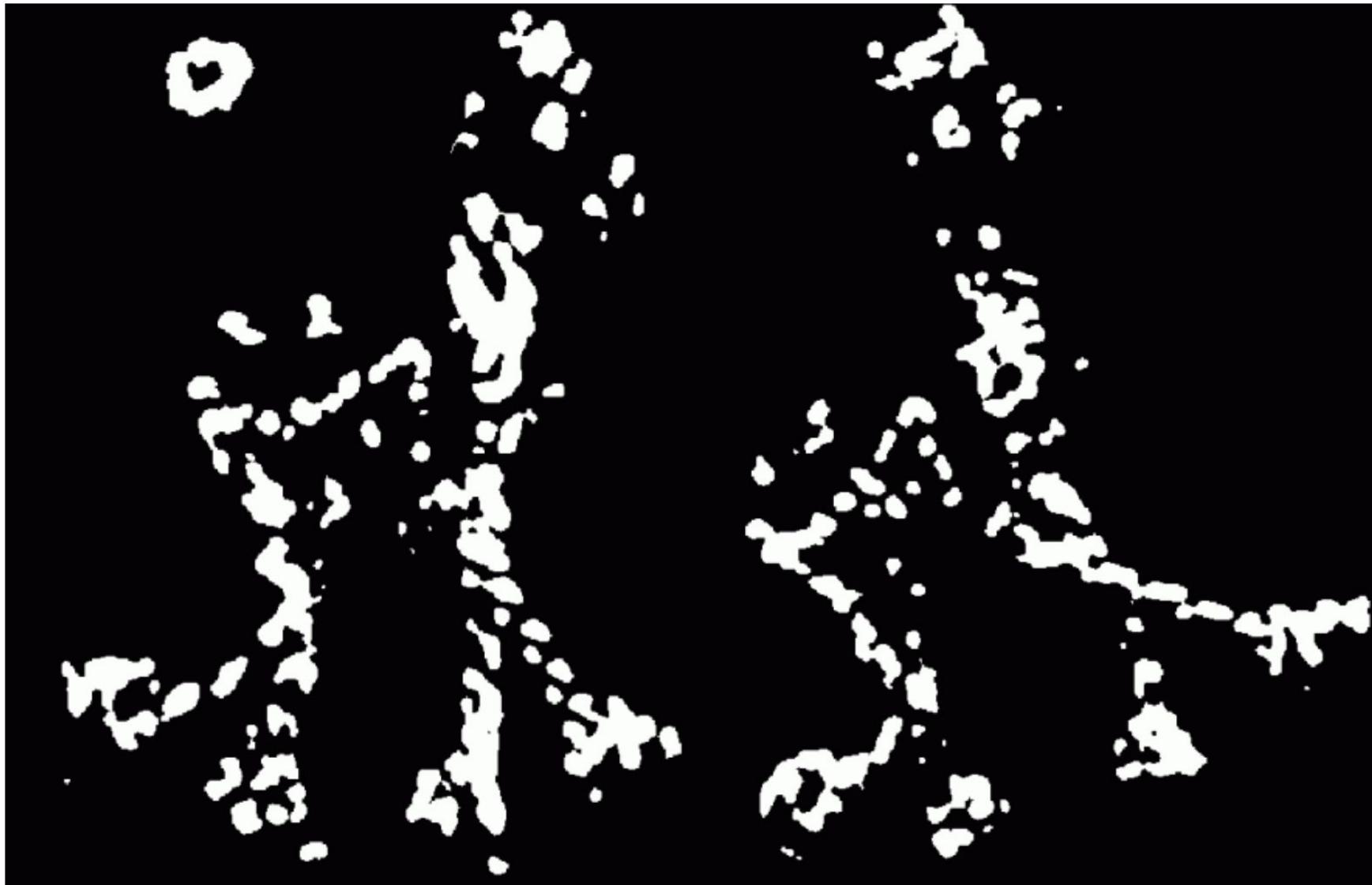
Пример



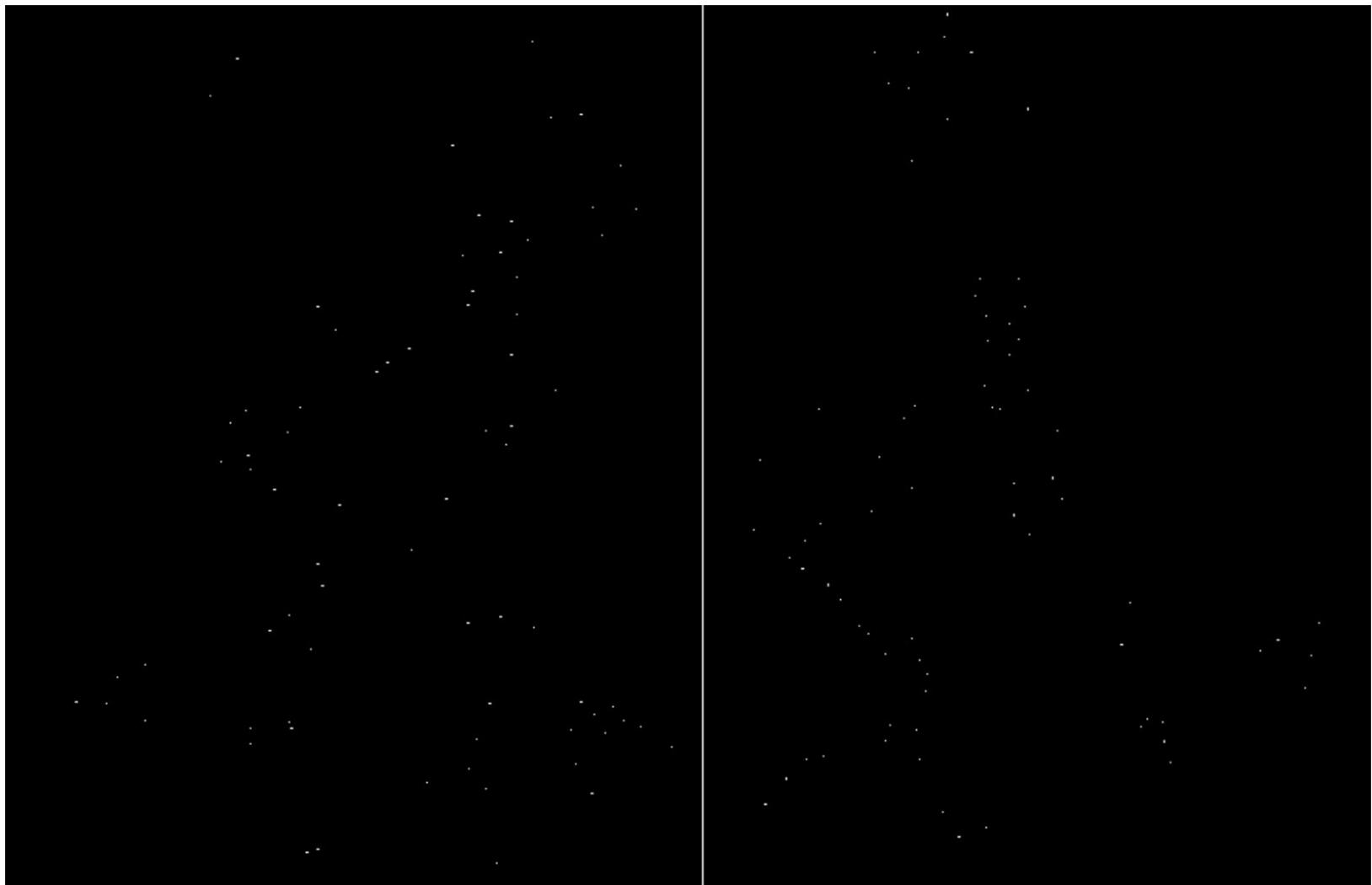
Пример: Вычисляем R



Пример: $R > T$



Пример: non-max suppression



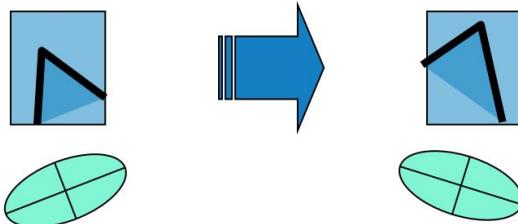
Пример: результат.



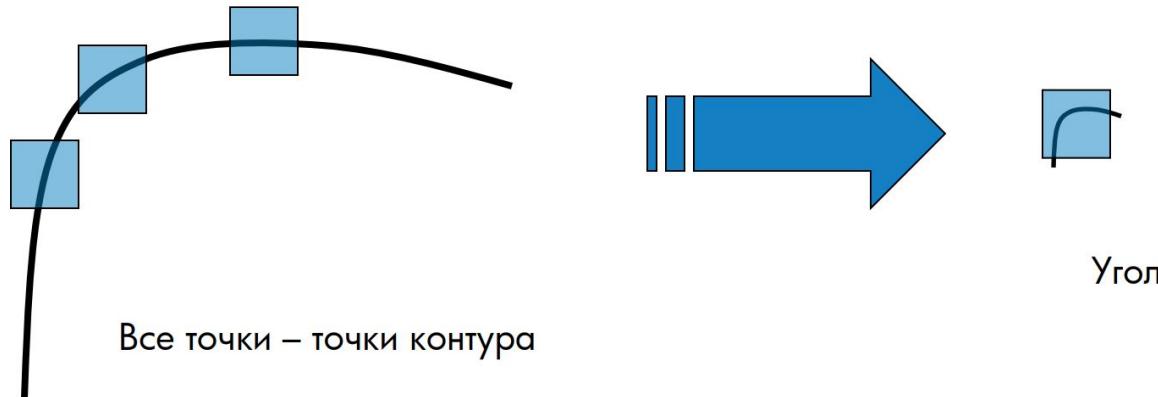
Свойства детектора Харриса

Свойства детектора Харриса

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

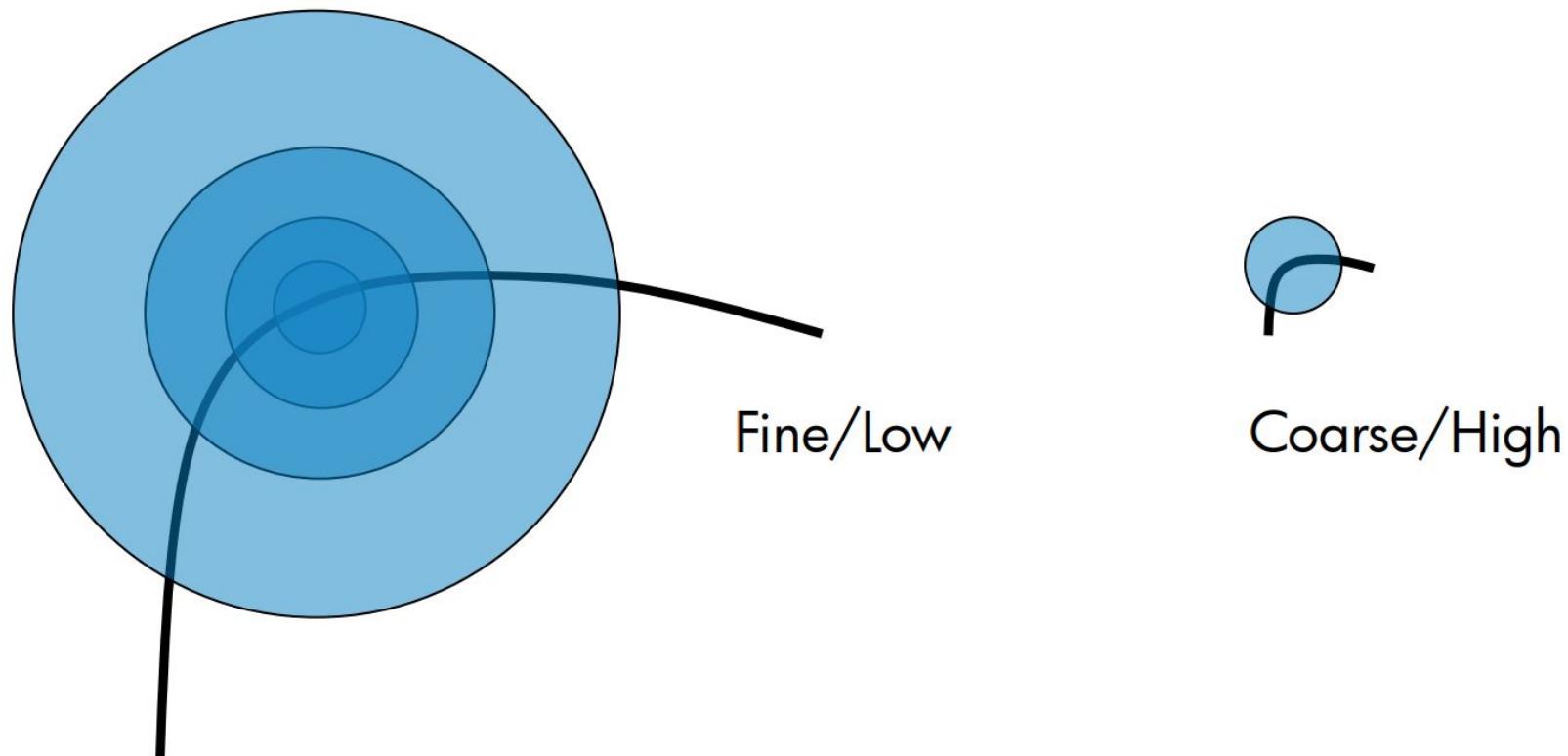


- Не инвариантен к изменению масштаба!

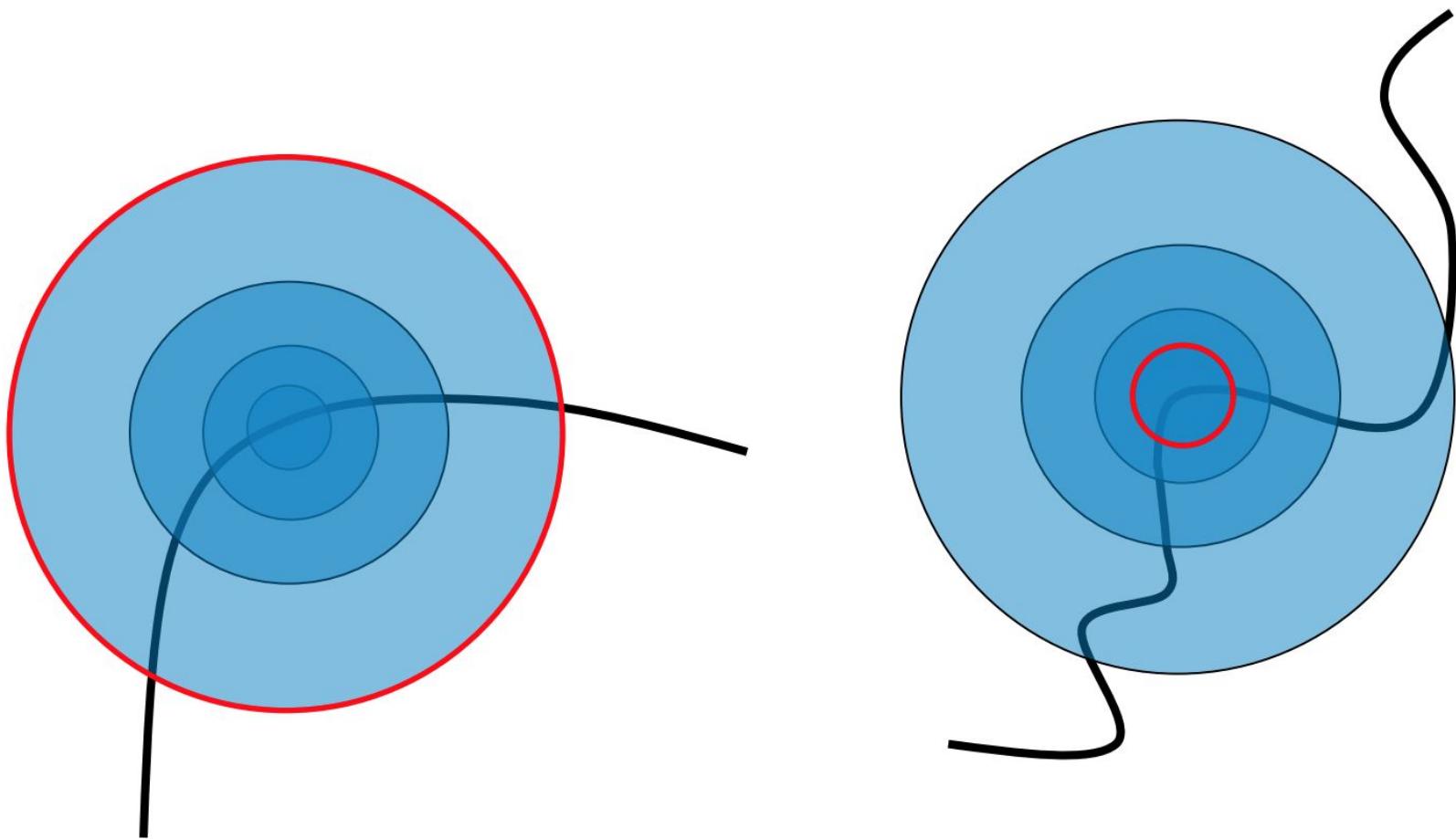


Инвариантность к изменению масштаба

Рассмотрим фрагменты (например, круги) разного размера на разных изображениях. Для двух изображений найдутся соответствующие масштабы, при которых фрагменты будут выглядеть похоже



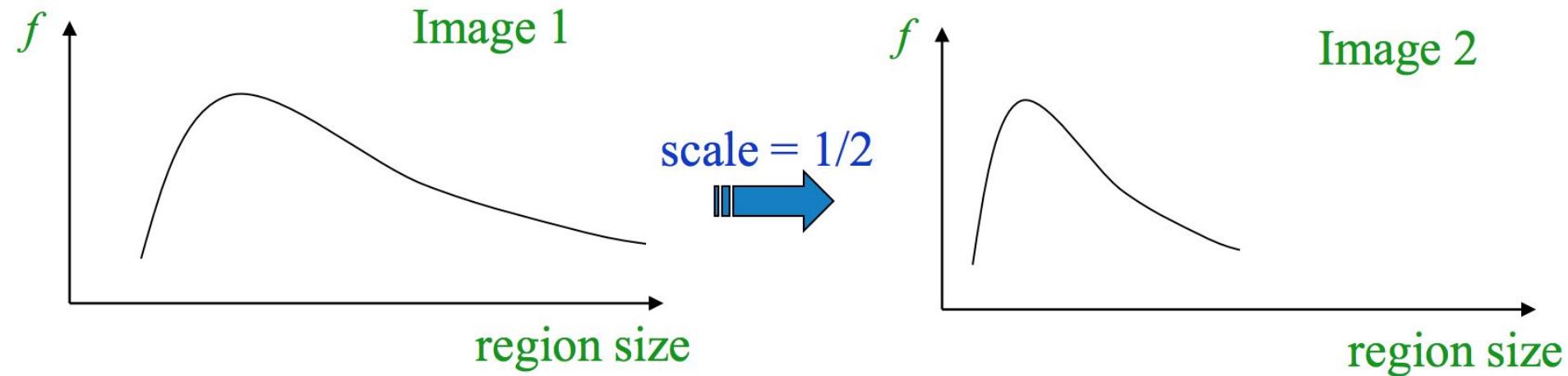
Задача: как выбрать размер фрагмента независимо для каждого изображения?



Инвариантность к изменению масштаба

Решение:

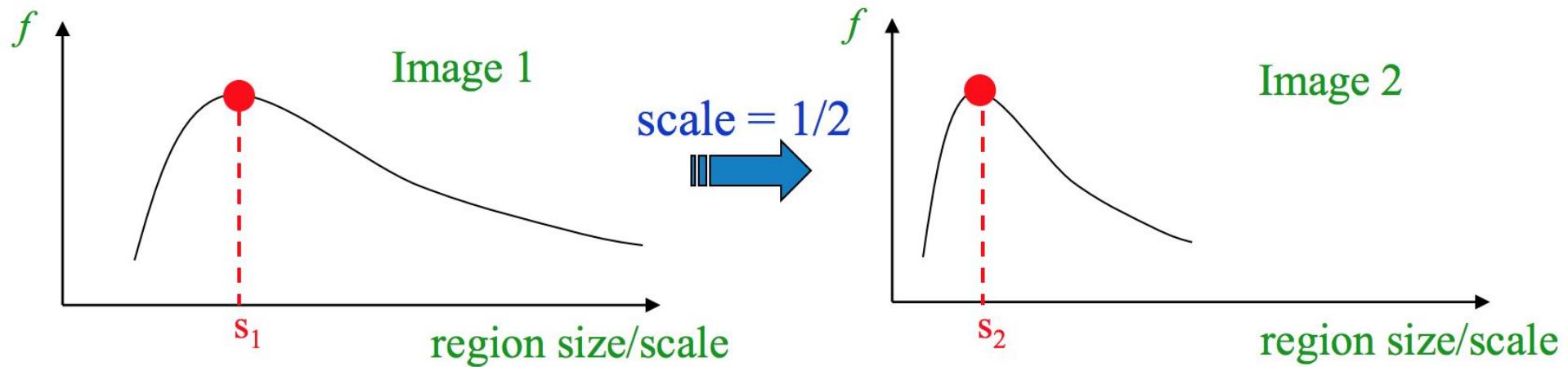
- Выберем функцию, заданную на фрагменте изображения, инвариантную к изменению масштаба – значение функции одинаково для сопоставимых фрагментов, даже если они разного масштаба (например, средняя интенсивность)
- В каждой точке изображения посмотрим на эту функцию, как на функцию от изменения размера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



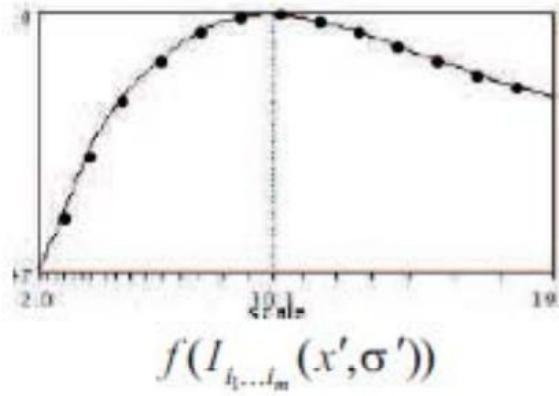
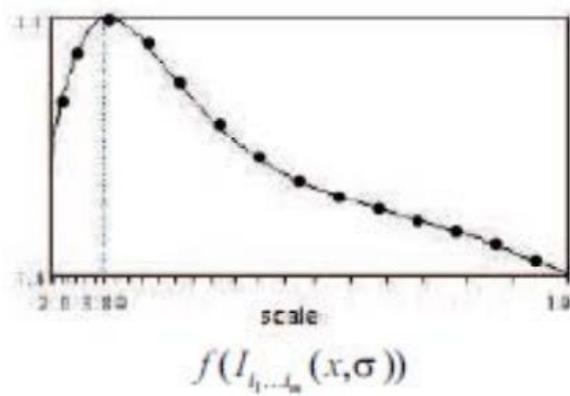
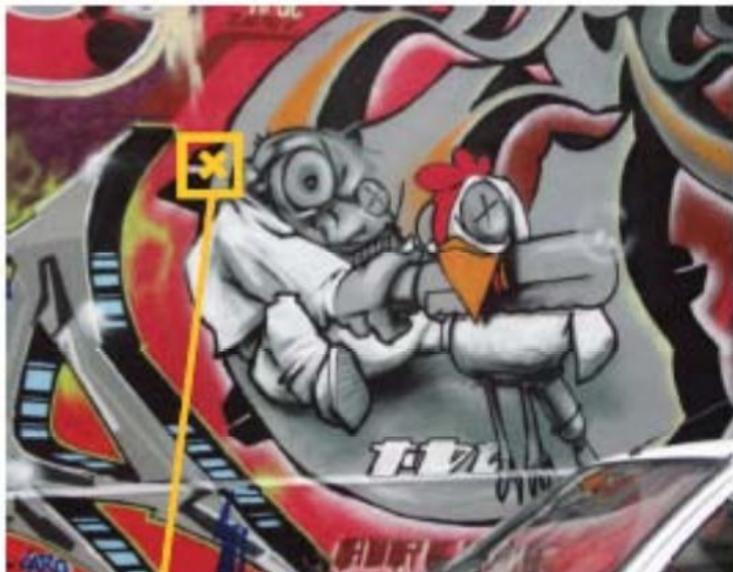
Инвариантность к изменению масштаба

Решение:

1. Найти локальный максимум такой функции
2. Наблюдение: точка локального максимума инвариантна к изменению масштаба
3. **Важно: размер фрагмента, на котором достигается локальный максимум, находится для каждого изображения независимо!**

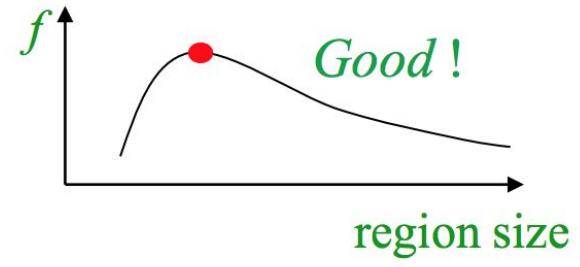
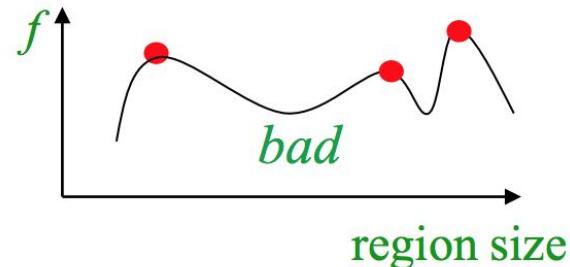
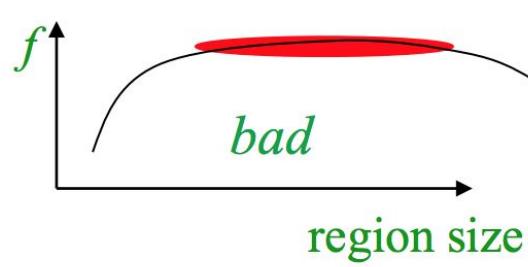


Characteristic Scale (Характеристический размер)



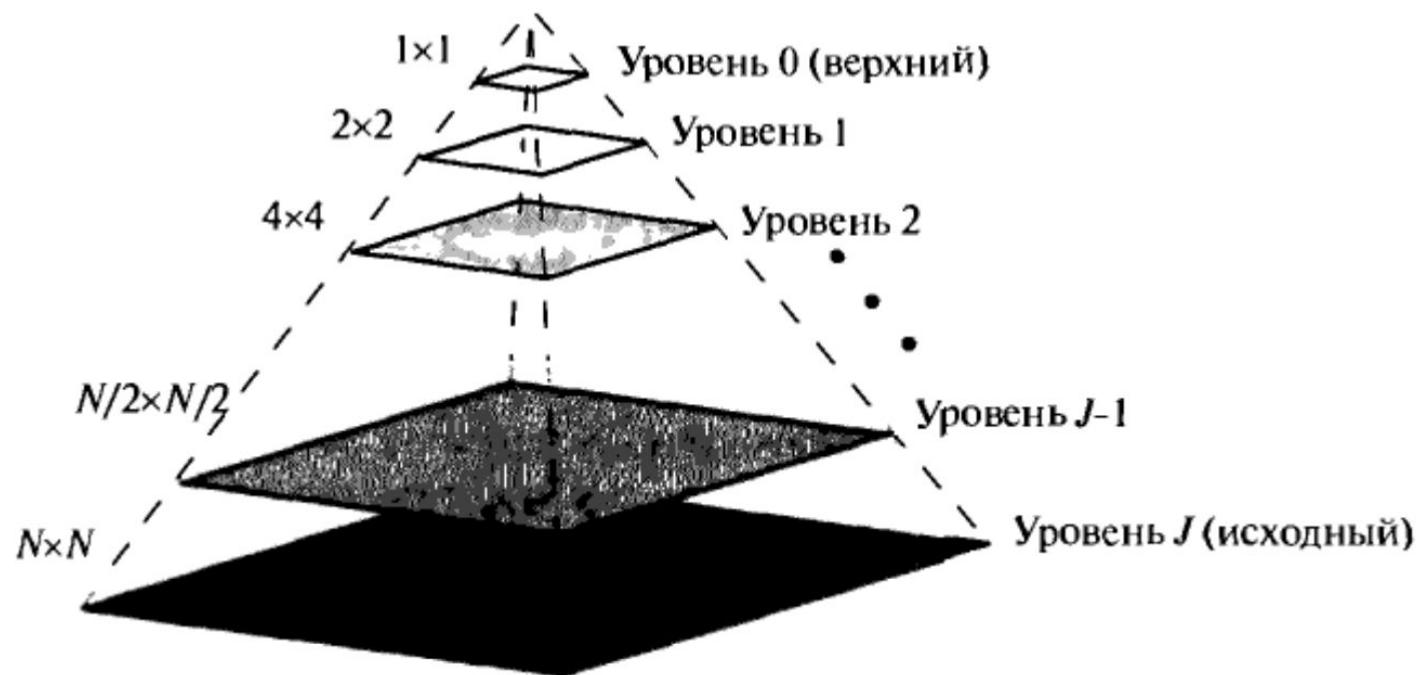
Инвариантность к изменению масштаба

“Хорошая” функция для определения соотношения масштабов фрагментов:

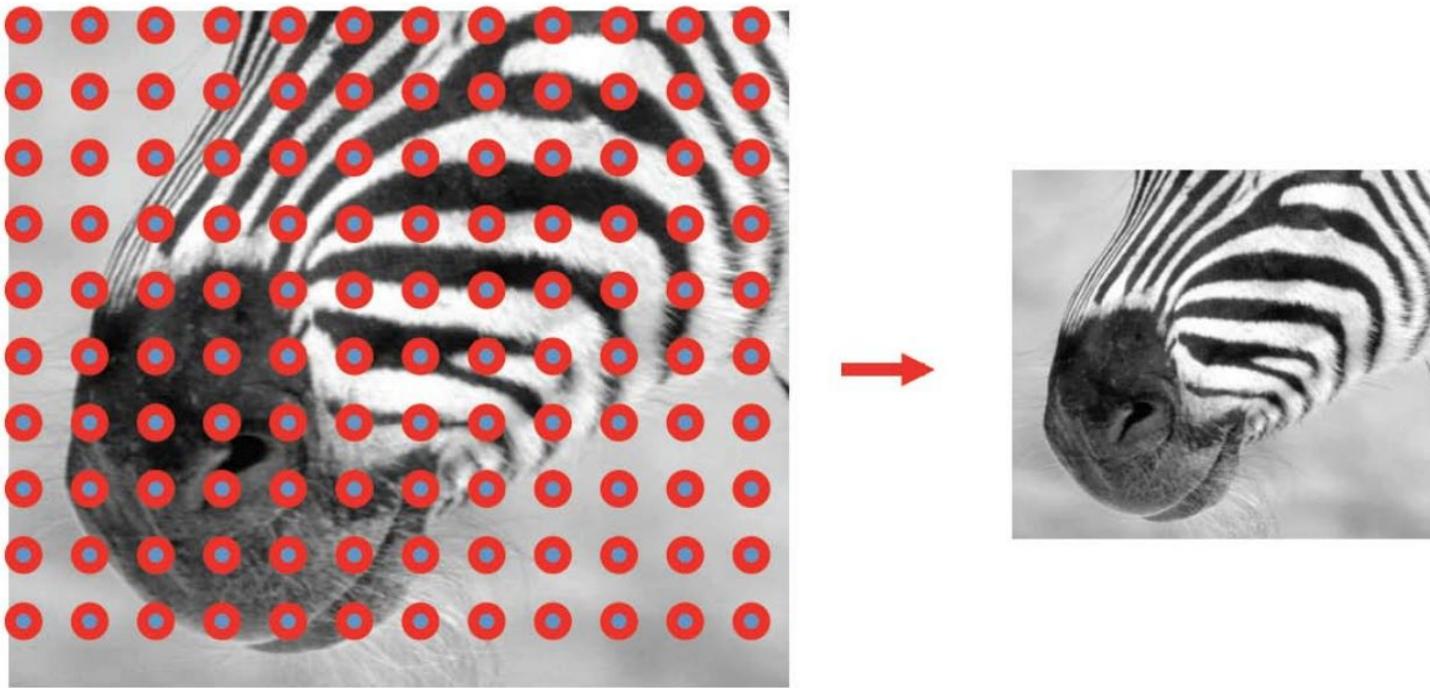


Нужен один хорошо выраженный пик функции.

Многомасштабное представление изображений



Прореживающая выборка (sampling)



$$g[n, m] = f[2n, 2m] = \begin{bmatrix} \ddots & & & & \\ & f[-2, 2] & f[0, 2] & f[2, 2] & \vdots \\ \dots & f[-2, 0] & \underline{f[0, 0]} & f[2, 0] & \dots \\ f[-2, -2] & f[0, -2] & f[2, -2] & & \ddots \\ & & \vdots & & \end{bmatrix}$$

Алиасинг (aliasing)

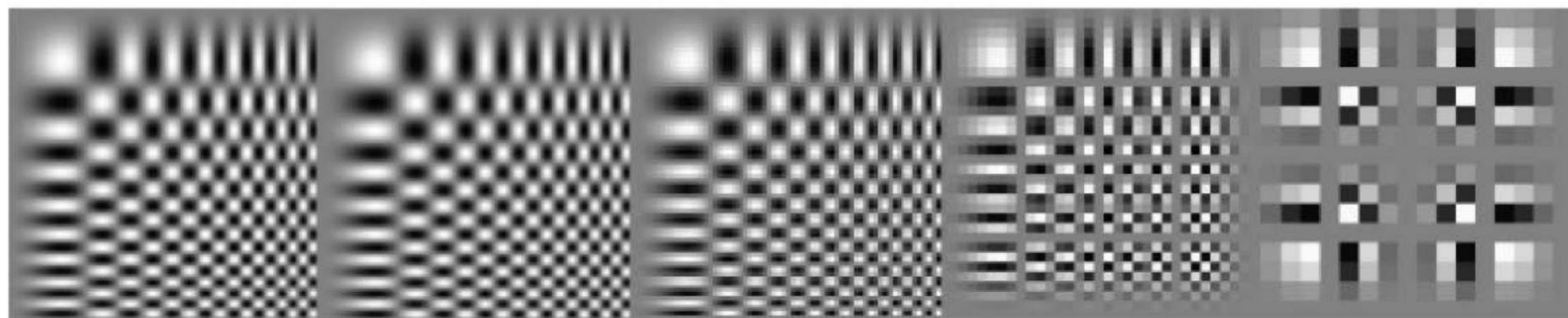
256x256

128x128

64x64

32x32

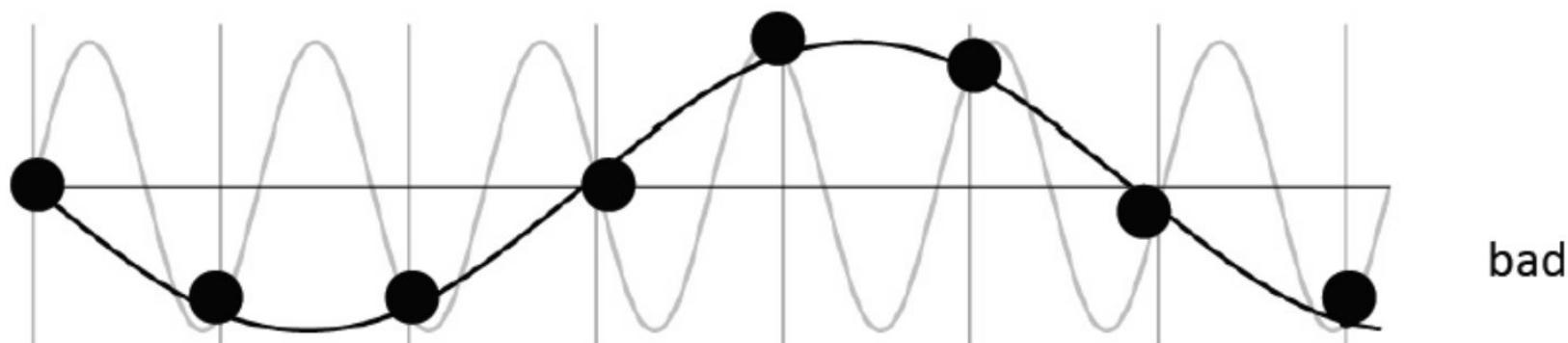
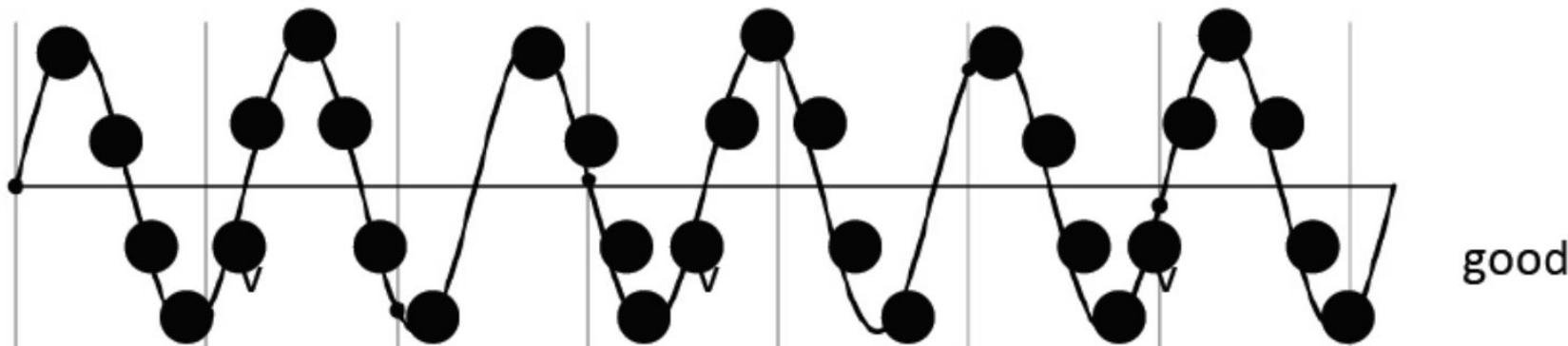
16x16





Sampling Theorem (Nyquist)

- When sampling a signal at discrete intervals, the sampling frequency must be $\geq 2 \times f_{\max}$
- f_{\max} = max frequency of the input signal.



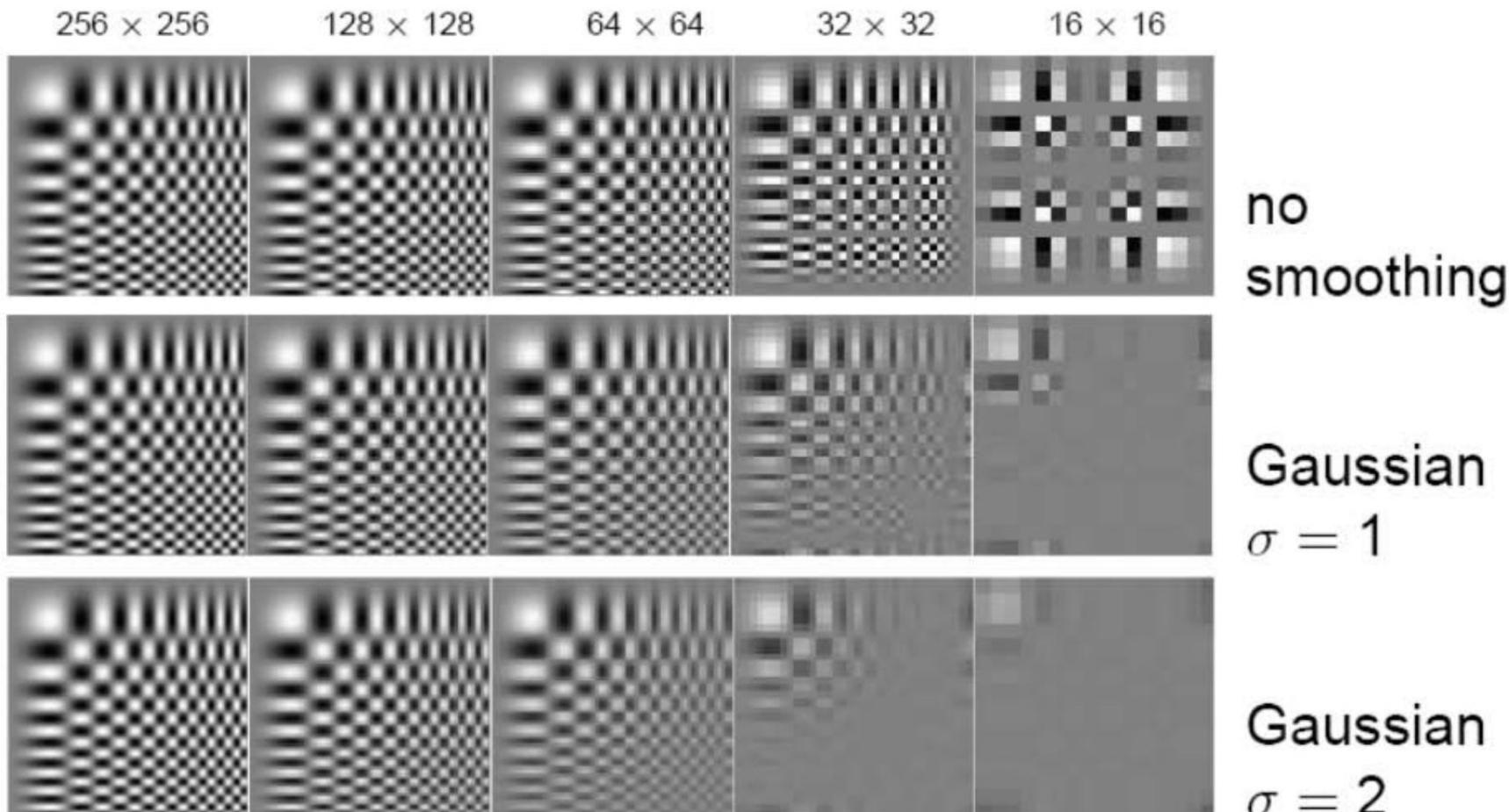
Антиэлайзинг

Решение: избавится от высоких частот (Потеряем часть информации, но это лучше, чем алиасинг)

Для удаления высоких частот используется сглаживание

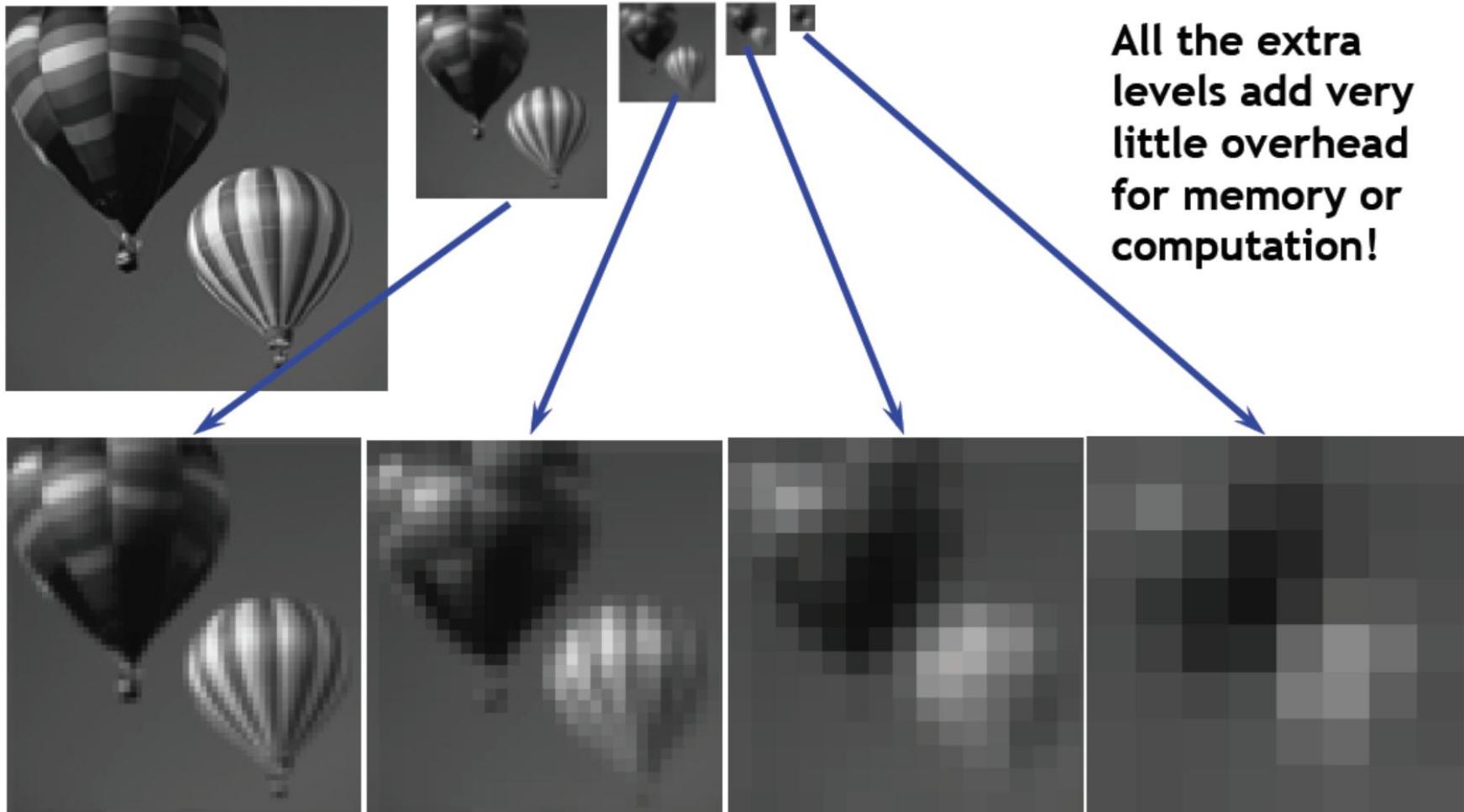
Если сглаживаем при помощи фильтра Гаусса, получим Гауссову пирамиду

Resampling with Prior Smoothing



- Note: We cannot recover the high frequencies, but we can avoid artifacts by smoothing before resampling.

Гауссова Пирамида



Масштабно-независимые детекторы

Функции для определения характеристического размера:

Лапласиан:

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

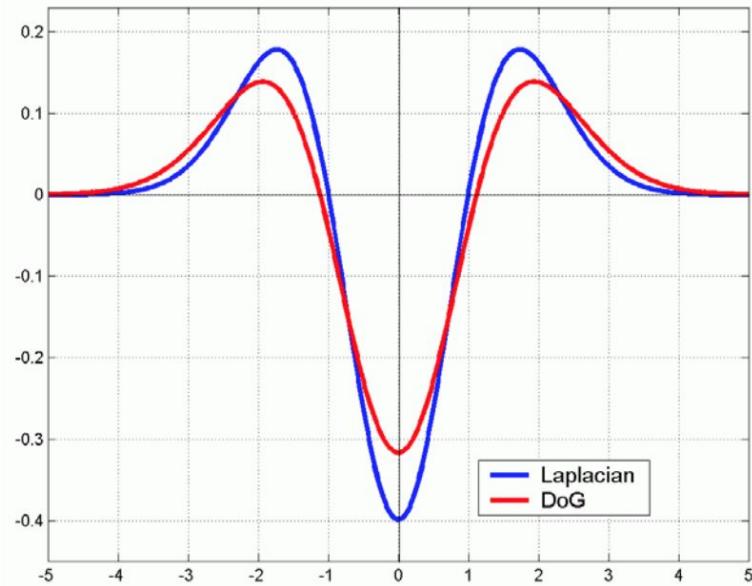
Разность Гауссиан
(приближение Лапласиана):

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

где

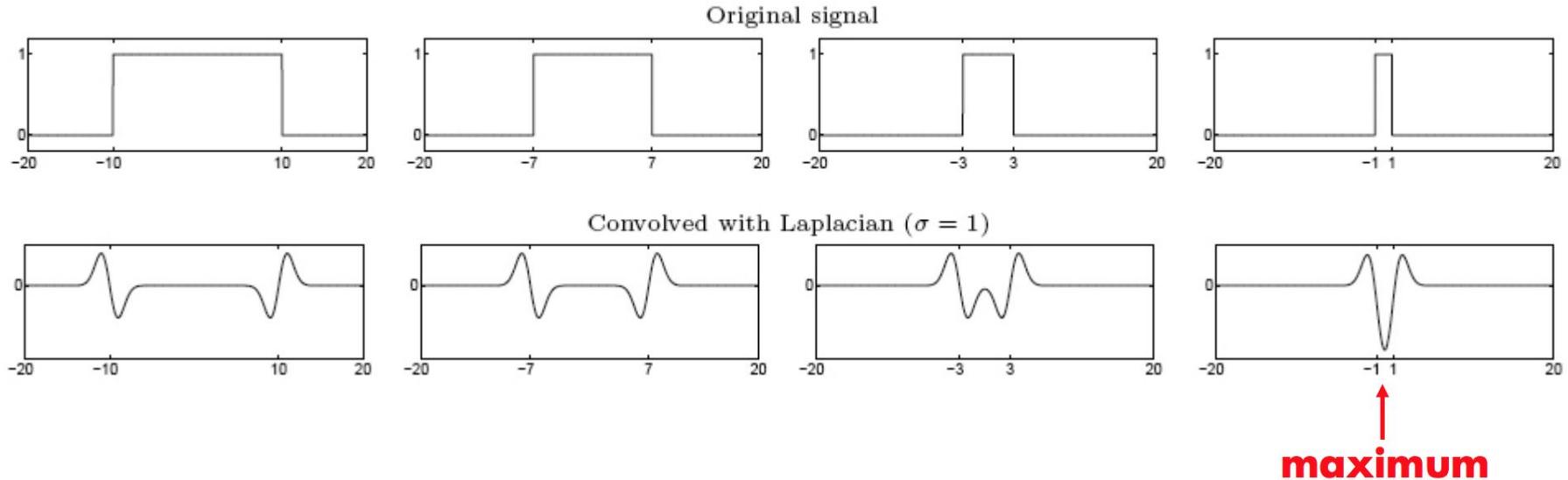
$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

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



Оба ядра инварианты к изменению
масштаба и повороту

Свертка с Лапласианом



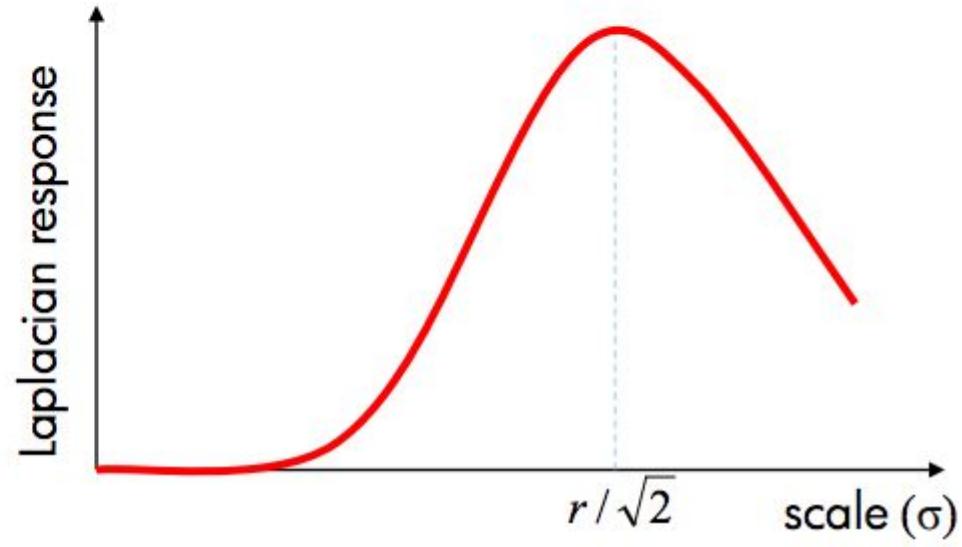
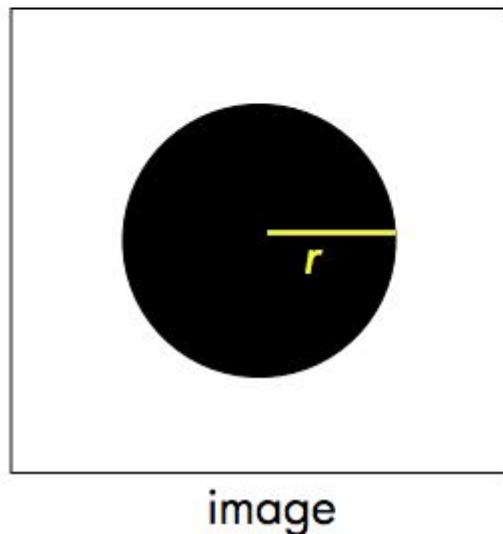
Spatial selection: the magnitude of the Laplacian response will achieve a maximum at the center of the blob, provided the scale of the Laplacian is “matched” to the scale of the blob

Масштаб фильтра.

- 2D Laplacian задается формулой:

$$(x^2 + y^2 - 2\sigma^2) e^{-(x^2+y^2)/2\sigma^2}$$

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



Laplacian-of-Gaussian (LoG)

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

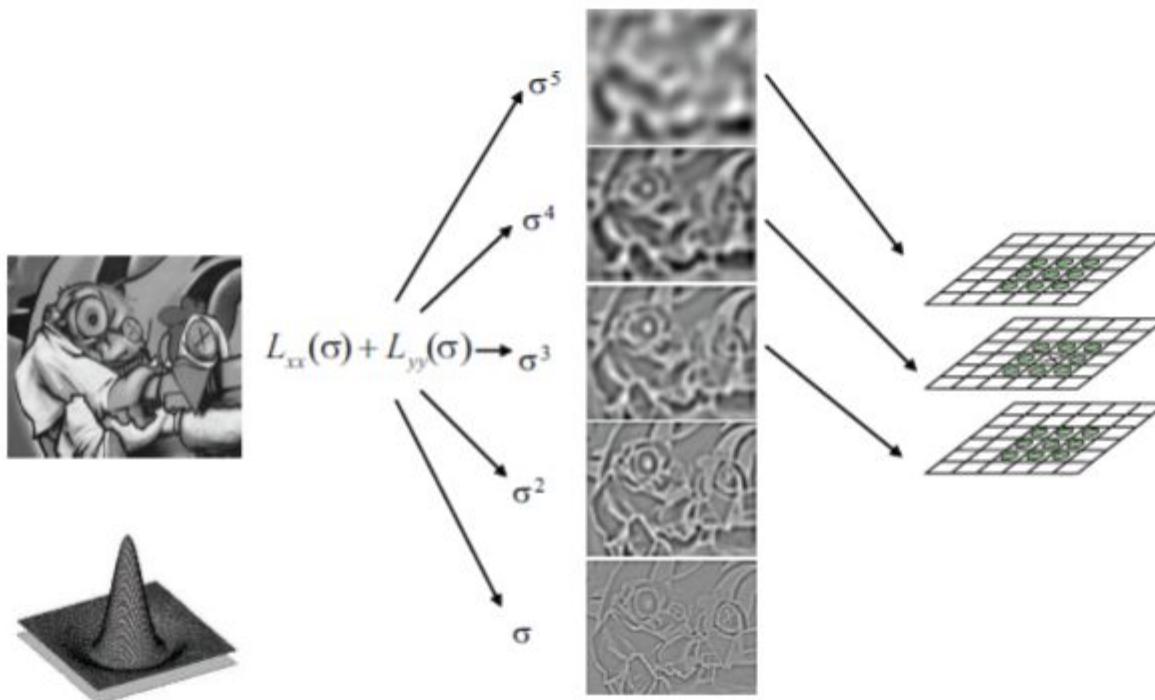
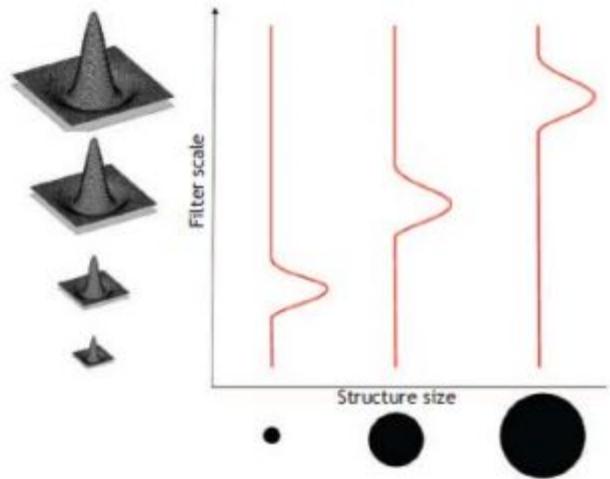


Figure 3.5: The Laplacian-of-Gaussian (LoG) detector searches for 3D scale space extrema of the LoG function.

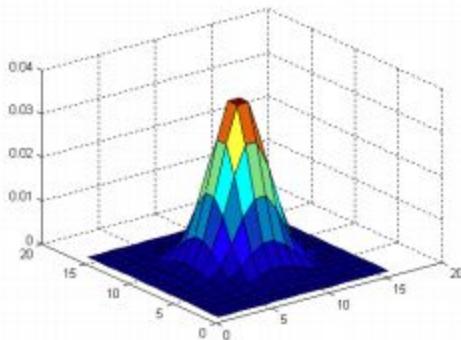
Difference-of-Gaussian

The input image is convolved with Gaussian function:

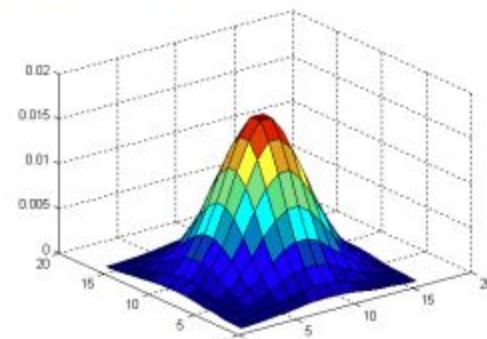
$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$$



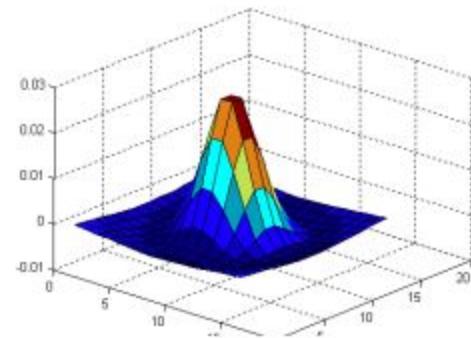
$$\sigma_1 = 2$$



$$\sigma_2 = 4$$



$$\sigma_1 - \sigma_2$$



DoG

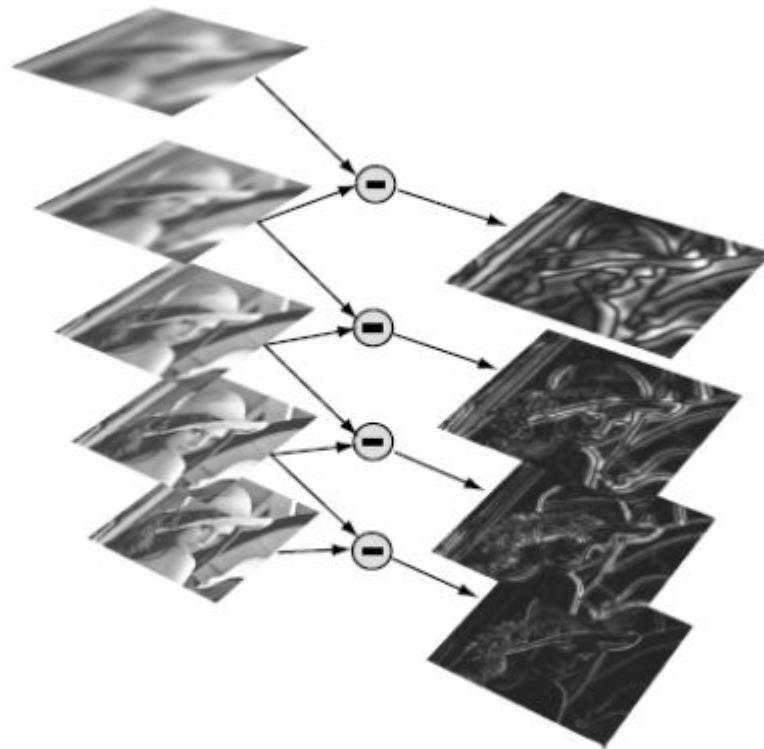


Figure 9.1: A Difference-of-Gaussian octave. The five images in the left stack are incrementally smoothed versions of the input image. The right stack shows the resulting DoG.

DoG

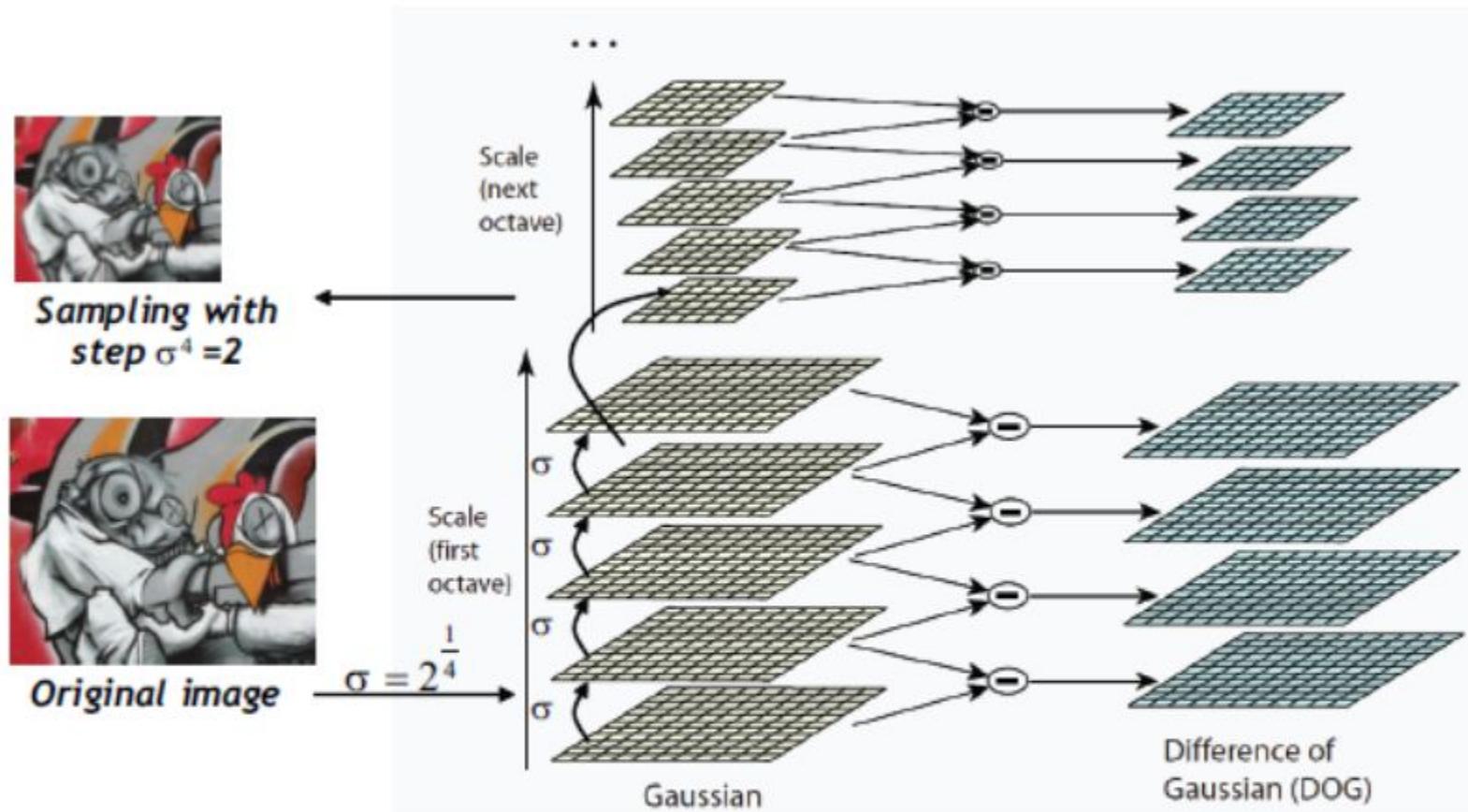


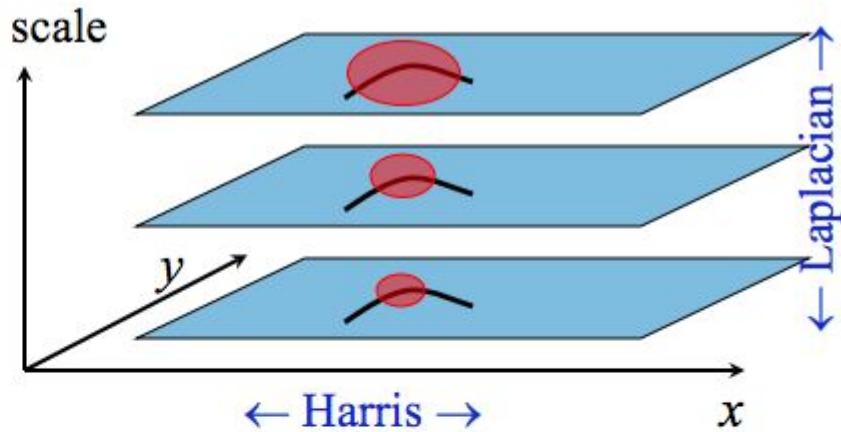
Figure 3.6: The Difference-of-Gaussian (DoG) provides a good approximation for the Laplacian-of-Gaussian. It can be efficiently computed by subtracting adjacent scale levels of a Gaussian pyramid. The DoG region detector then searches for 3D scale space extrema of the DoG function. *BL: Figure courtesy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

Поиск особых точек. 2 подхода.

Harris-Laplacian¹

Find local maximum of:

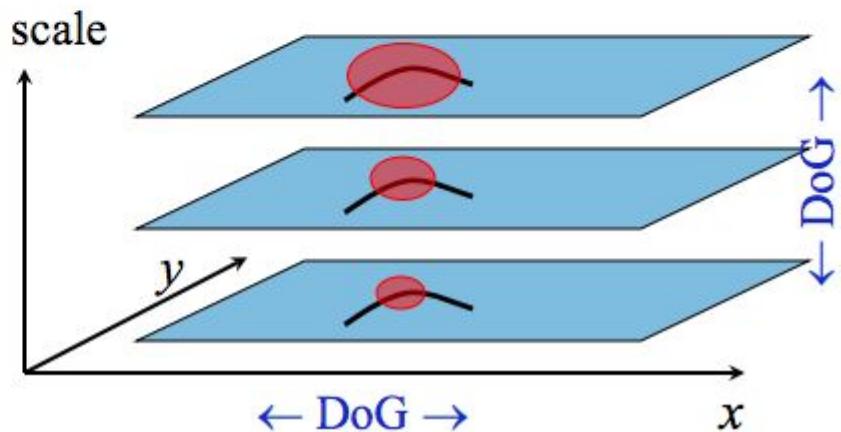
- Harris corner detector in space (image coordinates)
- Laplacian in scale



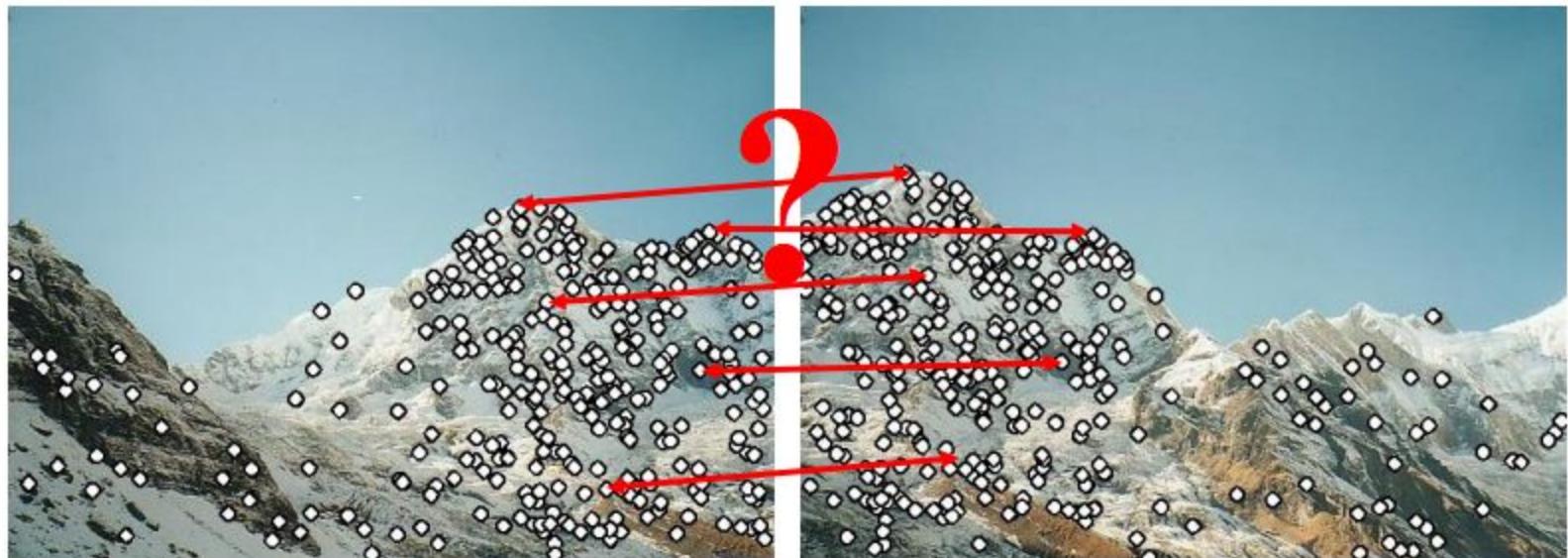
SIFT (Lowe)²

Find local maximum of:

- Difference of Gaussians in space and scale



Сопоставление особых точек.



SIFT дескриптор

1. Scale-space peak selection

- Using Difference-of-Gaussians (DoG)

2. Keypoint localization

- Elimination of unstable keypoints

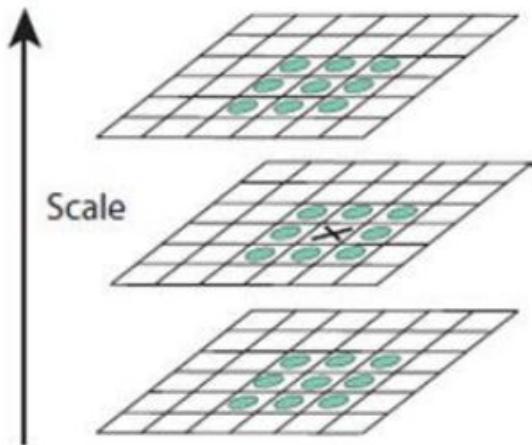
3. Orientation assignment

- Based on keypoint local image patch

4. Keypoint descriptor

- Based upon the image gradients in keypoint local neighbourhood

SIFT: stable maximum in different scales.

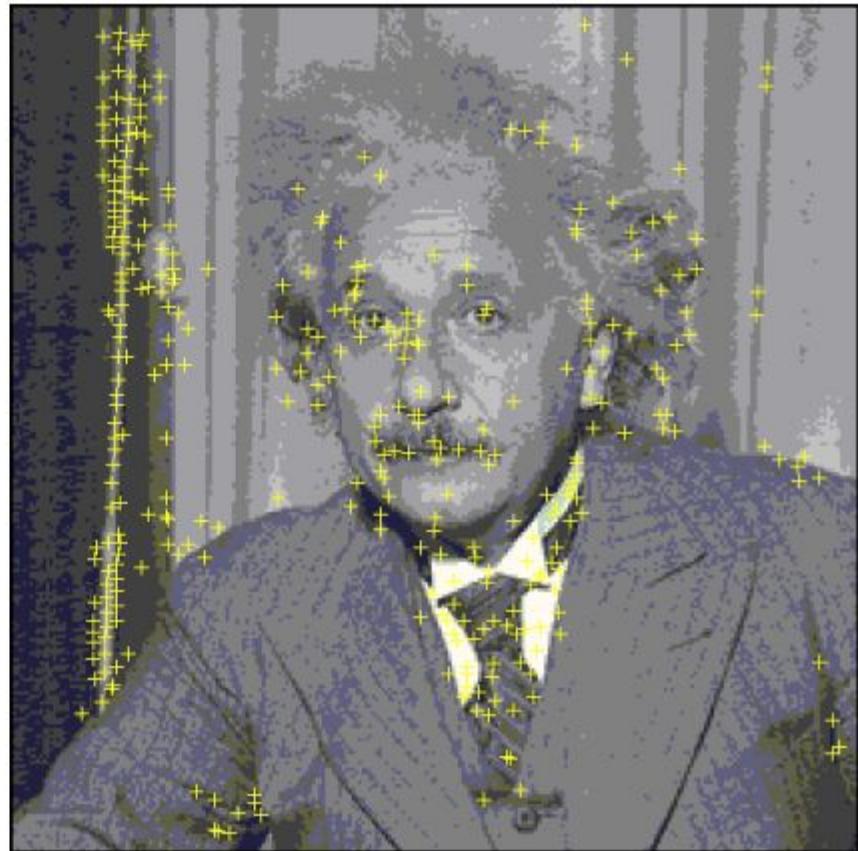
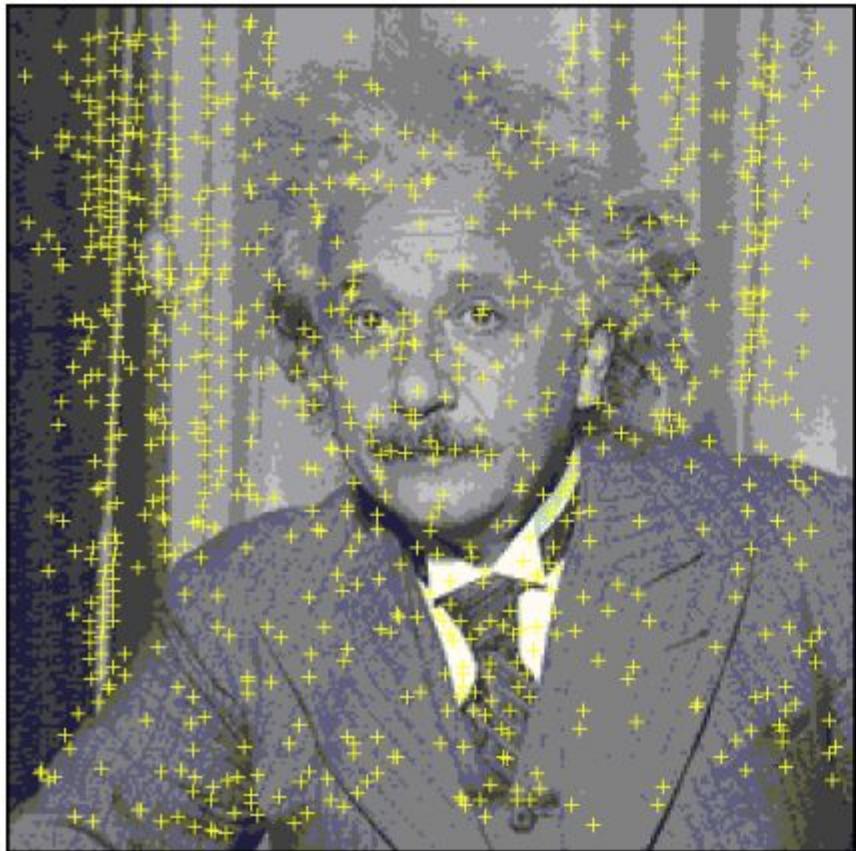


Maxima and minima of DoG applied in scale-space:

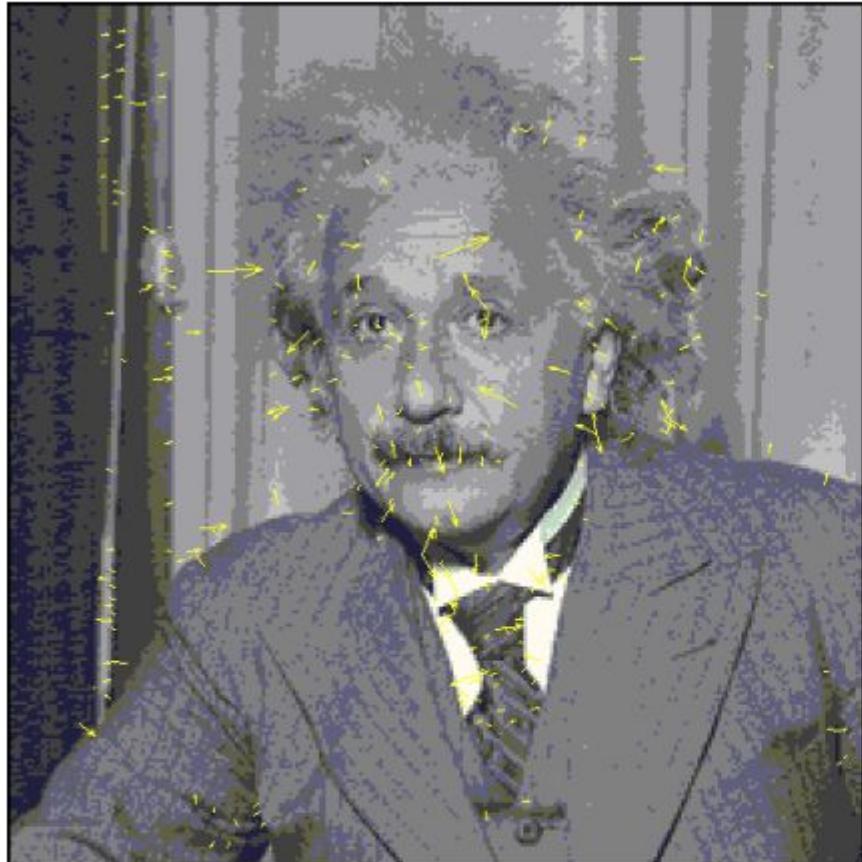
- 1) Extrema detection for the same scale
- 2) Check if it is stable for different scales



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

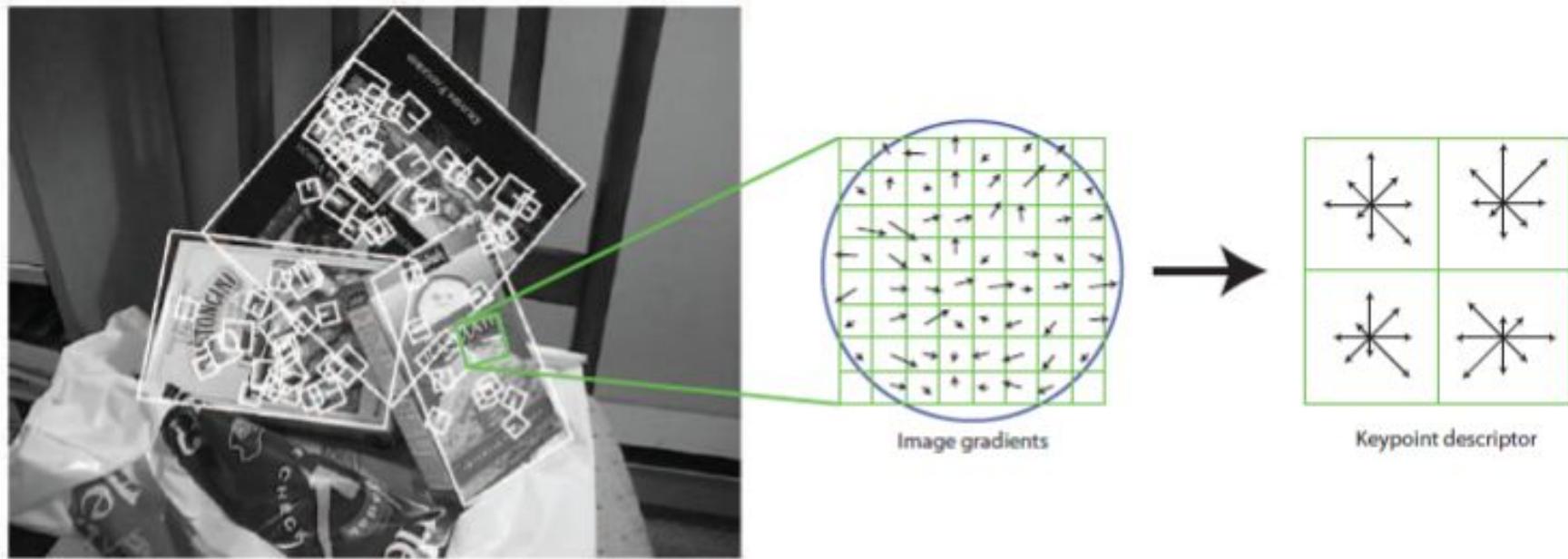
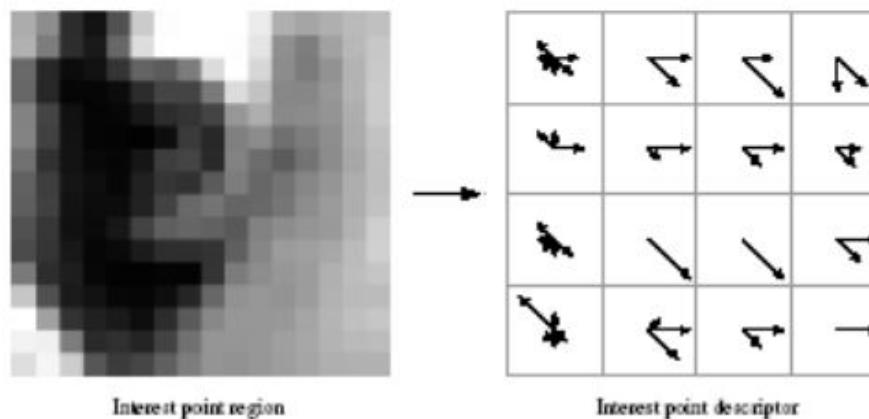


Figure 3.8: Visualization of the SIFT descriptor computation. For each (orientation-normalized) scale invariant region, image gradients are sampled in a regular grid and are then entered into a larger 4×4 grid of local gradient orientation histograms (for visibility reasons, only a 2×2 grid is shown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



(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.



(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.