

Визуальные признаки (image features)

Занятие 6

Computer Vision

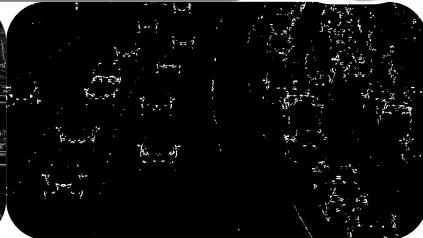
Лектор
Ян Колода

r_d

ВИЗУАЛЬНЫЕ ПРИЗНАКИ

Содержат важную информацию для распознавания и интерпретации сцен

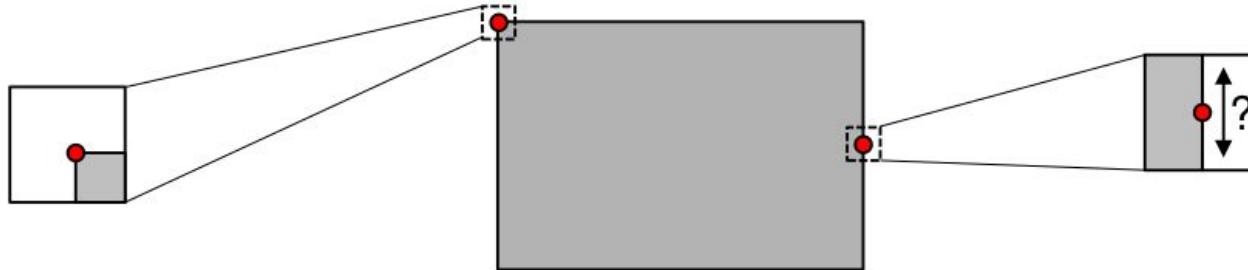
- Цвет
- Освещение
- Границы
- Углы
- ...



KEYPOINTS

Визуальные признаки с конкретными координатами (локализация)

- Границы локализованы только в одном направлении → углы?



KEYPOINT DETECTOR

Желательные свойства детектора keypoint

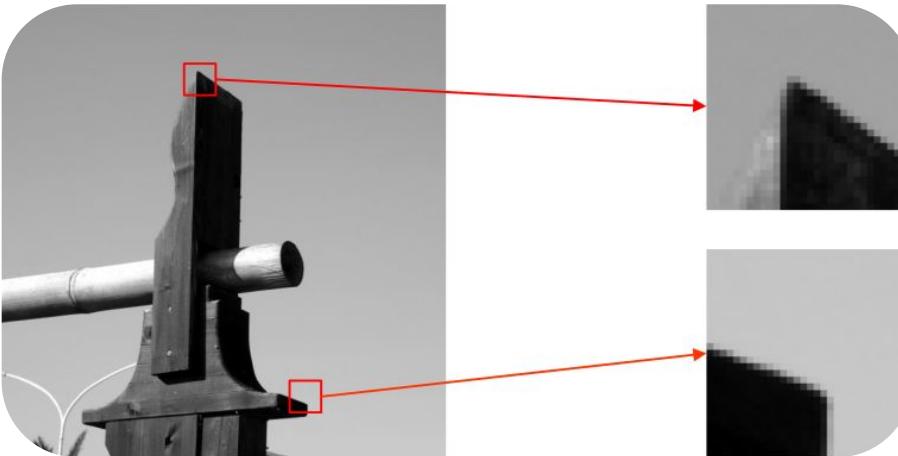
- Аккуратная локализация
- Invariance against shift, rotation, scale, brightness changes
- Robust against noise



ЧТО ТАКОЕ УГЛЫ?

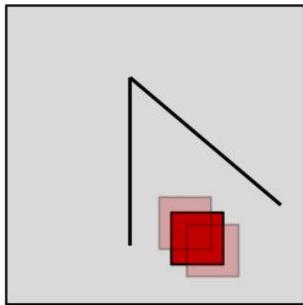
Хорошие и стабильные keypoints

- Узловой пункт (стык границ)
- Высокая визуальная вариация во всех направлениях

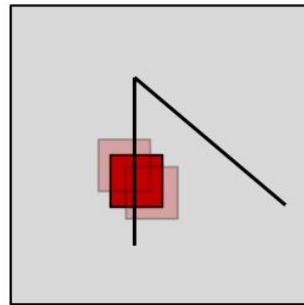


ДЕТЕКТОР УГЛОВ ХАРРИСА

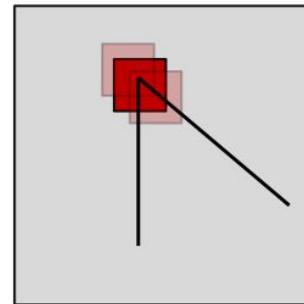
Harris corner detector: интуитивный подход



Flat region: no change
in all directions



Edge: no change along
edge direction



Corner: significant
change in all directions

Детектор Харриса

- Математический инструмент для определения вышеописанных ситуаций

ДЕТЕКТОР УГЛОВ ХАРРИСА

Sum of squared differences for shift (u, v) :

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

First order Taylor approximation:

$$G(x + u, y + v) \approx G(x, y) + uG_x(x, y) + vG_y(x, y)$$

Therefore:

$$S(u, v) \approx \sum_{x,y} w(x, y)[uG_x(x, y) + vG_y(x, y)]^2$$

ДЕТЕКТОР УГЛОВ ХАРРИСА

In matricial notation: $S(u, v) \approx [u \quad v] \mathbf{M} \begin{bmatrix} u \\ v \end{bmatrix}$

Harris matrix M is a 2×2 structure matrix (second-moment matrix)

$$\mathbf{M} = \begin{bmatrix} \sum_{x,y} w(x, y) G_x^2(x, y) & \sum_{x,y} w(x, y) G_x(x, y) G_y(x, y) \\ \sum_{x,y} w(x, y) G_x(x, y) G_y(x, y) & \sum_{x,y} w(x, y) G_y^2(x, y) \end{bmatrix}$$

Weighted window w , usually Gaussian

- Isotropic response

ДЕТЕКТОР УГЛОВ ХАРРИСА

Detection based on eigenvalues of \mathbf{M}

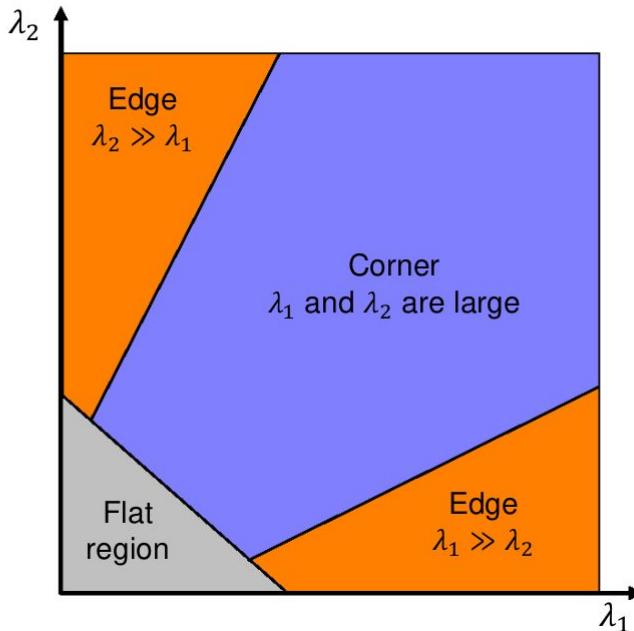
- λ_1, λ_2

Computation of eigenvalues

- Expensive

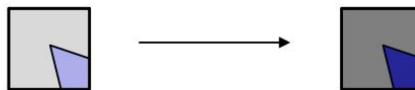
Measure of “cornerness” instead

$$\begin{aligned} C(x, y) &= \det(\mathbf{M}) - \kappa (\text{Tr}(\mathbf{M}))^2 \\ &= \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2) \end{aligned}$$

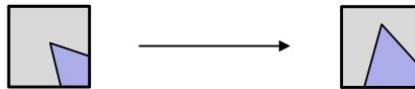


ROBUST HARRIS-LAPLACE DETECTOR

Invariant to brightness offset



Invariant to shift and rotation



Not invariant to scaling



SCALE SPACE FEATURES

Некоторые признаки видны только в определенном масштабе

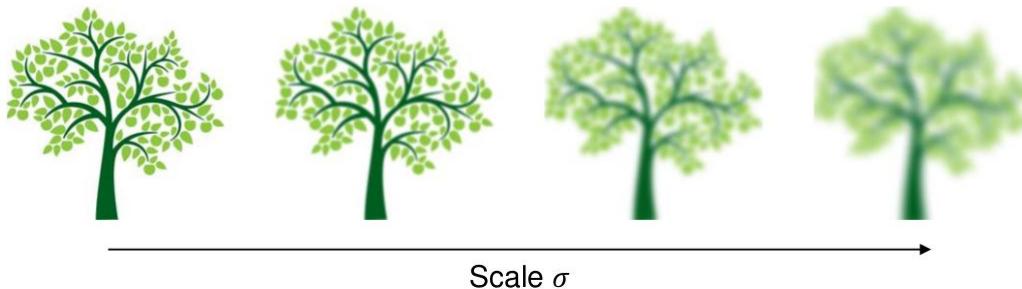
- Нужна масштабная инвариантность



SCALE SPACE REPRESENTATION

Параметрическая группа изображений, размытых фильтром Гаусса

$$f(x, y, \sigma) = G(x, y, \sigma) * f(x, y) \quad \text{with} \quad G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)}$$

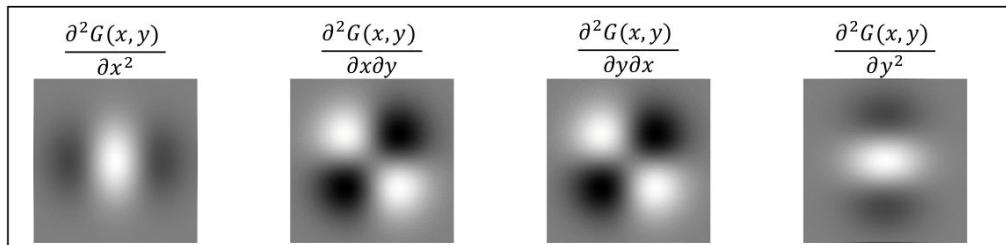
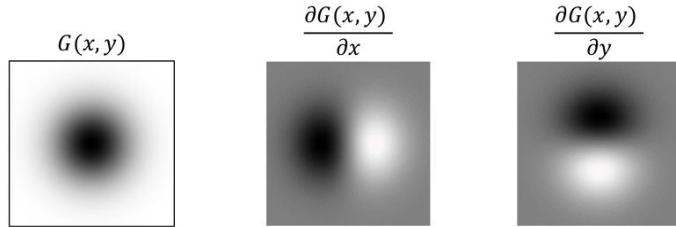


Хотите смотреть на листья или на дерево?

GAUSSIAN DERIVATIVES

Фильтры Гаусса всегда положительные

- Не могут детектировать keypoints



HESSIAN MATRIX

Square matrix of second order derivatives

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f(x, y)}{\partial x^2} & \frac{\partial^2 f(x, y)}{\partial x \partial y} \\ \frac{\partial^2 f(x, y)}{\partial x \partial y} & \frac{\partial^2 f(x, y)}{\partial y^2} \end{bmatrix}$$

Special case: **Hessian of Gaussian**

$$\mathbf{H} = \left(\begin{array}{cc} \text{Blurry image} & \text{Blurry image} \\ \text{Blurry image} & \text{Blurry image} \end{array} \right)$$

SIFT (SCALE INVARIANT FEATURE TRANSFORM)

Масштабно-инвариантная трансформация признаков

- Высокоотличительные признаки



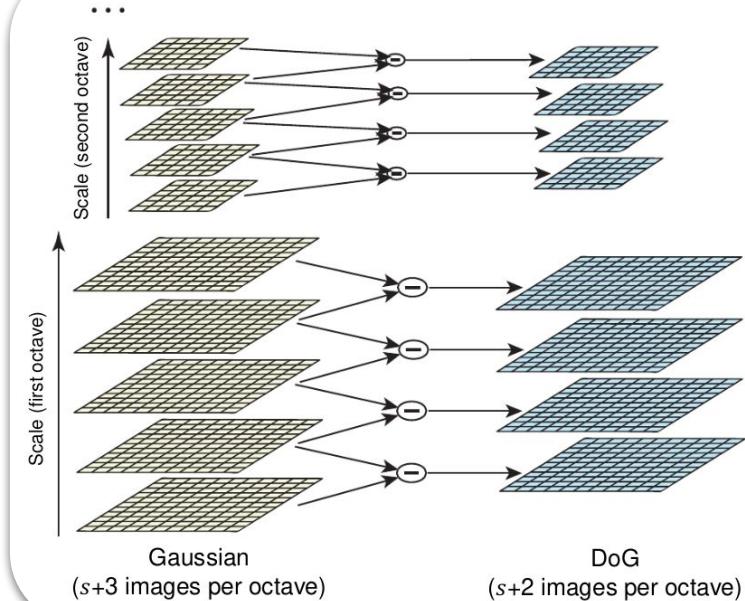
SIFT (SCALE INVARIANT FEATURE TRANSFORM)

Вычисляем максимальные пункты

- Позиция, масштаб, величина

Величина вычисляется так, как у Харриса

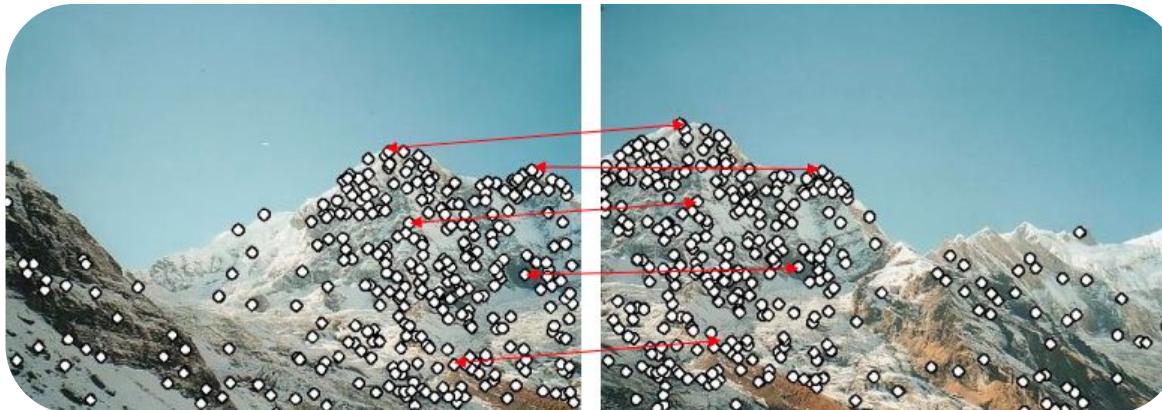
- Но на основе матрицы Гессса



SIFT (SCALE INVARIANT FEATURE TRANSFORM)

Fundamental for feature/image matching

- Кадры, снятые разными камерами и/или в разных интервалах



SIFT

SIFT — не только детектор, но также дескриптор

- Reference orientation → invariance to rotation

Для каждого keypoint'a

- Вычисли градиент вокруг

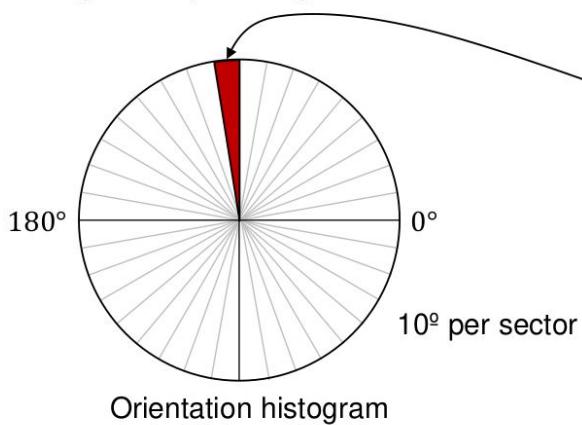
$$m(x, y) = \sqrt{(I(x + 1, y, \sigma) - I(x - 1, y, \sigma))^2 + (I(x, y + 1, \sigma) - I(x, y - 1, \sigma))^2}$$

$$\varphi(x, y) = \text{atan} \frac{I(x, y + 1, \sigma) - I(x, y - 1, \sigma)}{I(x + 1, y, \sigma) - I(x - 1, y, \sigma)}$$

-2,-2	-1,-2	0,-2	1,-2	2,-2
-2,-1	-1,-1	0,-1	1,-1	2,-1
-2,0	-1,0	x, y	1,0	2,0
-2,1	-1,1	0,1	1,1	2,1
-2,2	-1,2	0,2	1,2	2,2

SIFT (ORIENTATION ASSIGNMENT)

Orientation assigned by voting



121°	122°	128°	131°	121°
120°	120°	124°	120°	123°
98°	90°	95°	102°	110°
97°	95°	101°	101°	70°
84°	79°	80°	80°	76°

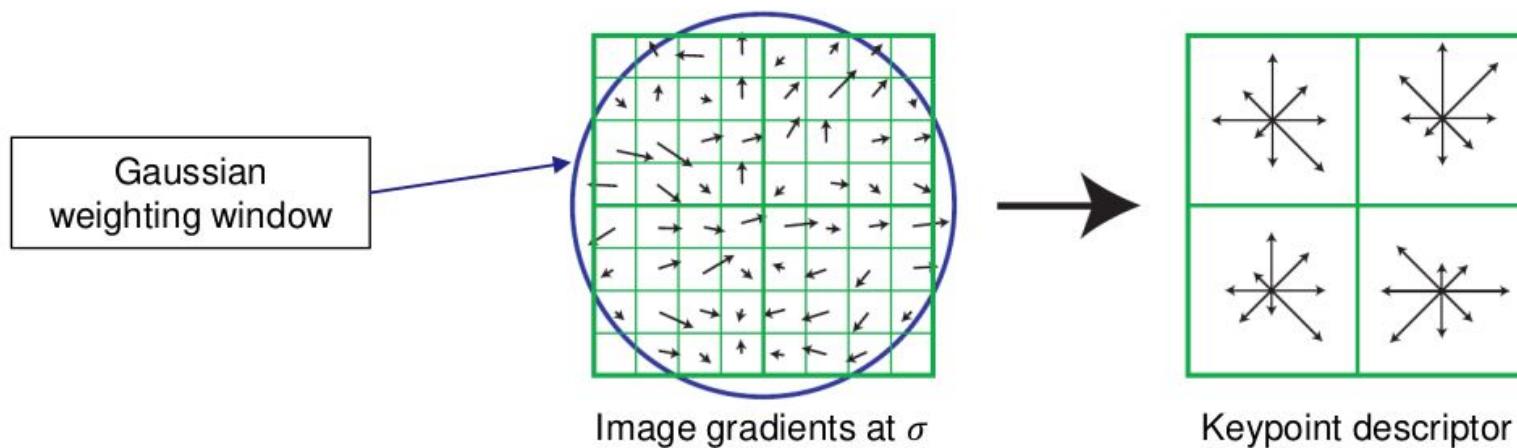
Votes proportional to:

- gradient magnitude
- distance from keypoint (Gaussian window with 1.5σ)

SIFT (DESCRIPTOR)

Состоит из 16 гистограмм направления

- Высокоотличительный и инвариантный



SIFT (DESCRIPTOR)

Есть разные детекторы и дескрипторы признаков

- SIFT, SURF, BRISK и т. д.

Детекторы и дескрипторы можно комбинировать

- Например, использовать Харриса для детекции и SIFT для описания (дескрипции)

Feature matching

- Два признака (из разных кадров) соответствуют друг другу, если их дескрипторы похожи

Q&A



r_d

ЗАВЖДИ Є КУДИ
ЗРОСТАТИ