Lecture 22 Object and Feature detection (chapter 12)

Yuyao Zhang, Xiran Cai PhD

zhangyy8@shanghaitech.edu.cn caixr@shanghaitech.edu.cn

SIST Building 2 302-F/302-C

Course piazza link：
piazza.com/shanghaitech.edu.cn/spring2021/cs270spring2021

上海科技大学 ShanghaiTech University

Outline

- Template matching.
- Image feature.
- Harris corner and Shi-tomasi corner detection
- SIFT – Scale Invariant feature transform

Template matching

➢ Correlation between template and image

➢ Template: $\omega(x, y)$; Image: $I(x, y)$.

➢ Correlation coefficient:

➢ $\gamma(x, y)=\frac{cov(\omega,I)}{\sigma_\omega \sigma_I}=\frac{E(\omega-\overline{\omega})(I-\bar{I}))}{\sigma_\omega \sigma_I}$

➢ $\bar{\omega}$: average value of template; $\bar{I}_{xy}$: average value of image inside window; $\gamma \in [-1, 1]$; $\gamma = 1$: template perfectly match the window;$\gamma = 0$: no correlation/no match.

Image

Template

Image

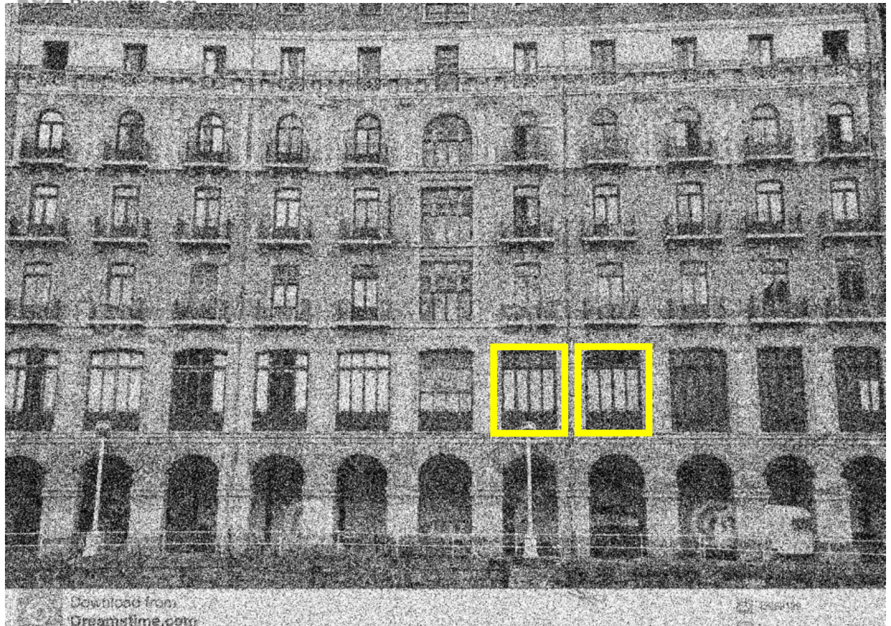
Template

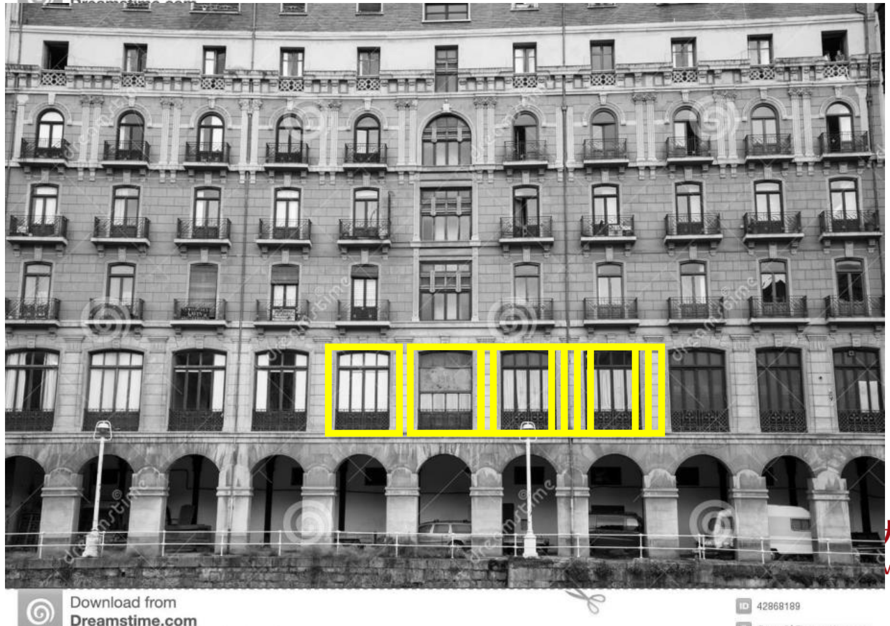
Thresh=0.9

Thresh=0.6

Template matching

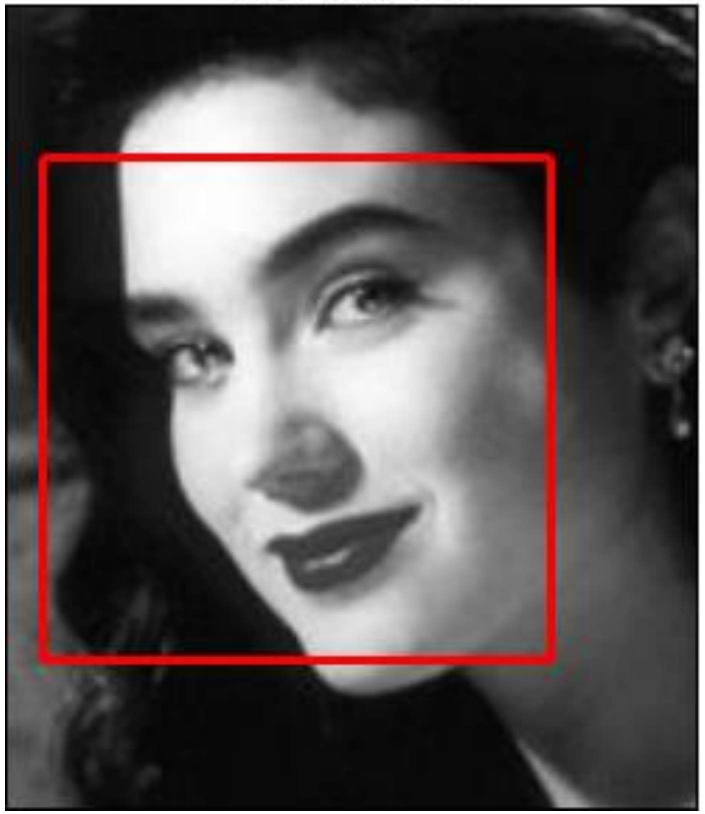
Template selected



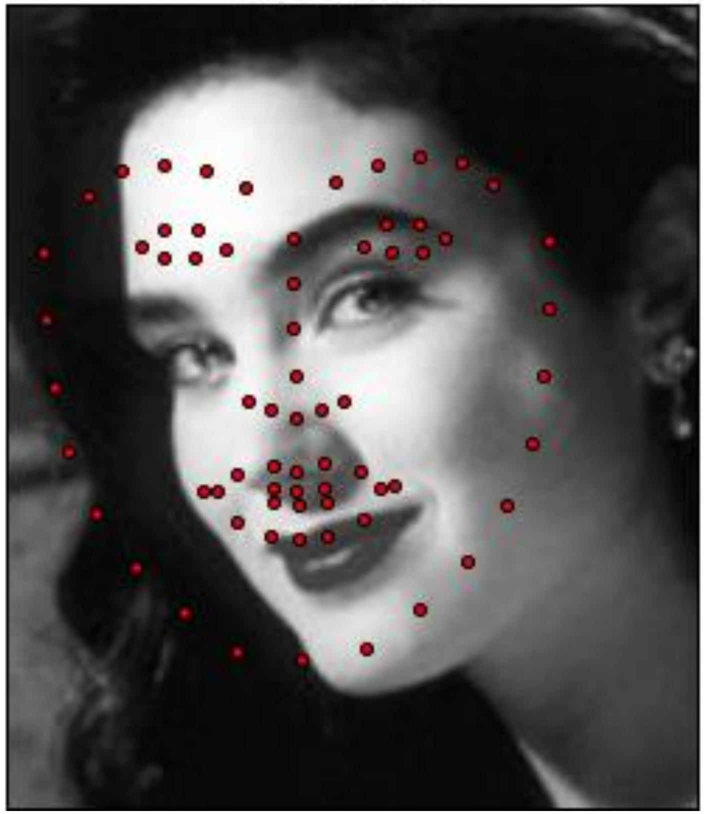


Application

Bounding box



Initial shape



Final shape

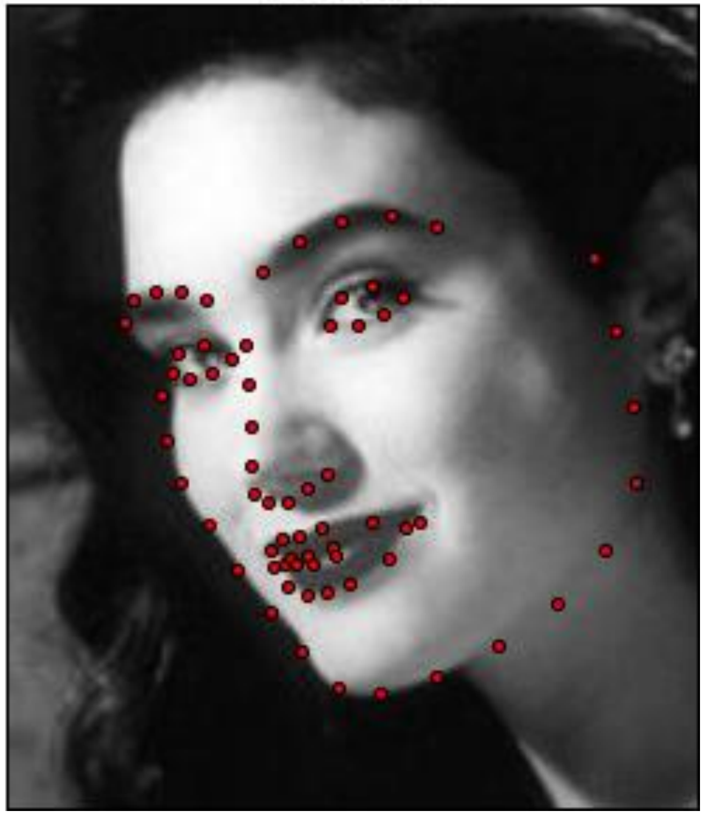
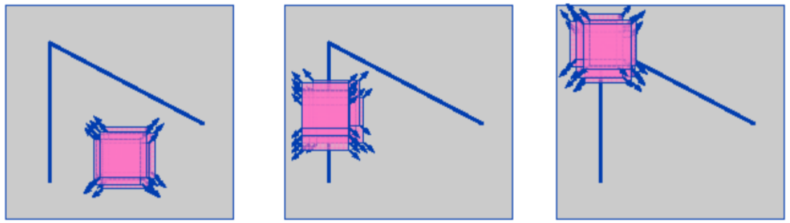


Image features

- Idea: describe object as a collection of smaller features.
- What makes a good feature?
- Distinctive.
- (a) flat; (b) single edge; (c) dual-direction edge. A good feature should have lots of edge strength in 2 directions.

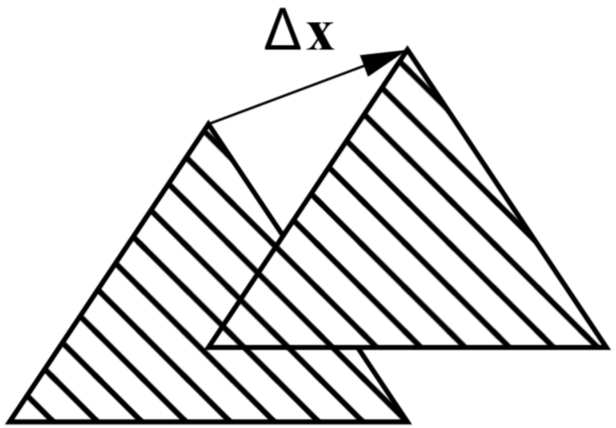
Harris Corner



"flat" region:
no change in
all directions

"edge":
no change along
the edge direction

"corner":
significant change
in all directions



- A shifted corner produces some difference in the image.
- A shifted uniform region produces no difference.
- So, look for large difference in shifted image.

Harris Corner

Suppose an image patch $\boldsymbol{W}$ at $\boldsymbol{x}$ is shifted by a small amount $\Delta\boldsymbol{x}$. Then, the sum-squared difference at $\boldsymbol{x}$ is

$$E(\mathbf{x}) = \sum_{\mathbf{x}_i \in W} \left[I(\mathbf{x}_i) - I(\mathbf{x}_i + \Delta\mathbf{x}) \right]^2 .$$

$$E(x, y) = \sum_{(x_i, y_i) \in W} \left[I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y) \right]^2 .$$

This is also called the auto-correlation function. Apply Taylor's series expansion to $I(x_i + \Delta x)$:

$$\begin{aligned} I(x_i + \Delta x, y_i + \Delta y) &= I(x_i, y_i) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y \\ &= I(x_i, y_i) + I_x \Delta x + I_y \Delta y \\ &= I(\mathbf{x}_i) + (\nabla I)^\top \Delta\mathbf{x} \end{aligned}$$

$$\nabla I = (I_x, I_y)^\top .$$

Harris Corner

Then we have:

$$\begin{aligned} E(\mathbf{x}) &= \sum_W [I_x \Delta x + I_y \Delta y]^2 \\ &= \sum_W [I_x^2 \Delta^2 x + 2 I_x I_y \Delta x \Delta y + I_y^2 \Delta^2 y] \\ &= (\Delta \mathbf{x})^\top \mathbf{A}(\mathbf{x}) \Delta \mathbf{x} \end{aligned}$$

where the auto-correlation matrix A is given by (Exercise)

$$\mathbf{A} = \begin{bmatrix} \sum_W I_x^2 & \sum_W I_x I_y \\ \sum_W I_x I_y & \sum_W I_y^2 \end{bmatrix}.$$

Harris Corner

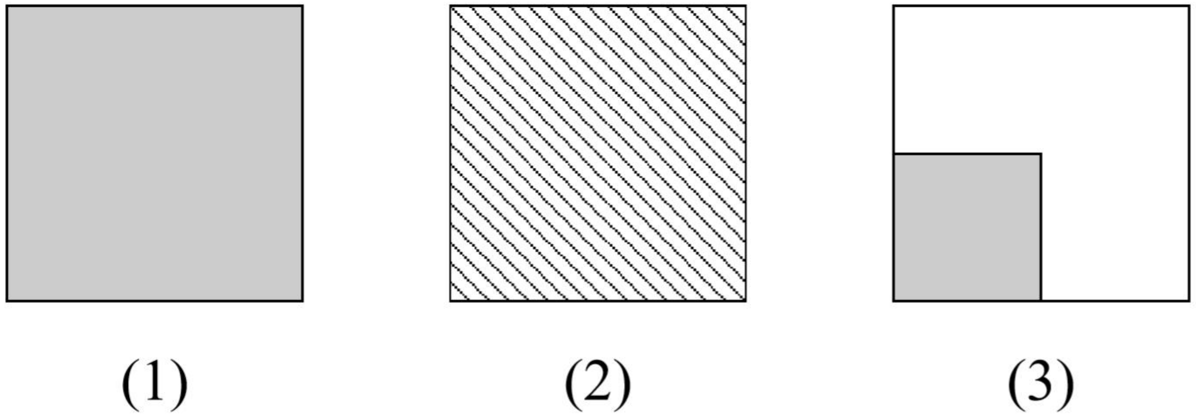
- $\boldsymbol{A}$ is a 2×2 matrix. This means there exist scalar values $\boldsymbol{\lambda_1}, \boldsymbol{\lambda_2}$ and vectors $\boldsymbol{v_1}, \boldsymbol{v_2}$ such that

$$\mathbf{A}\,\mathbf{v}_i = \lambda_i \mathbf{v}_i \ , \quad i = 1, 2$$

- $\boldsymbol{vi}$ are the orthonormal eigenvectors, i.e.,

$$\mathbf{v}_i^\top \mathbf{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

- $\boldsymbol{\lambda_i}$ are the eigenvalues; expect $\boldsymbol{\lambda_i} \geq \mathbf{0}$.



(1) If both λ_i are small, then feature does not vary much in any direction. $\Rightarrow$ uniform region (bad feature)

(2) If the larger eigenvalue $\lambda_1 \gg \lambda_2$, then the feature varies mainly in the direction of $\mathbf{v}_1$. $\Rightarrow$ edge (bad feature)

(3) If both eigenvalues are large, then the feature varies significantly in both directions. $\Rightarrow$ corner or corner-like (good feature)

(4) In practice, I has a maximum value (e.g., 255).
So, λ_1, λ_2 also have an upper bound.
So, only have to check that $\min(\lambda_1, \lambda_2)$ is large enough.

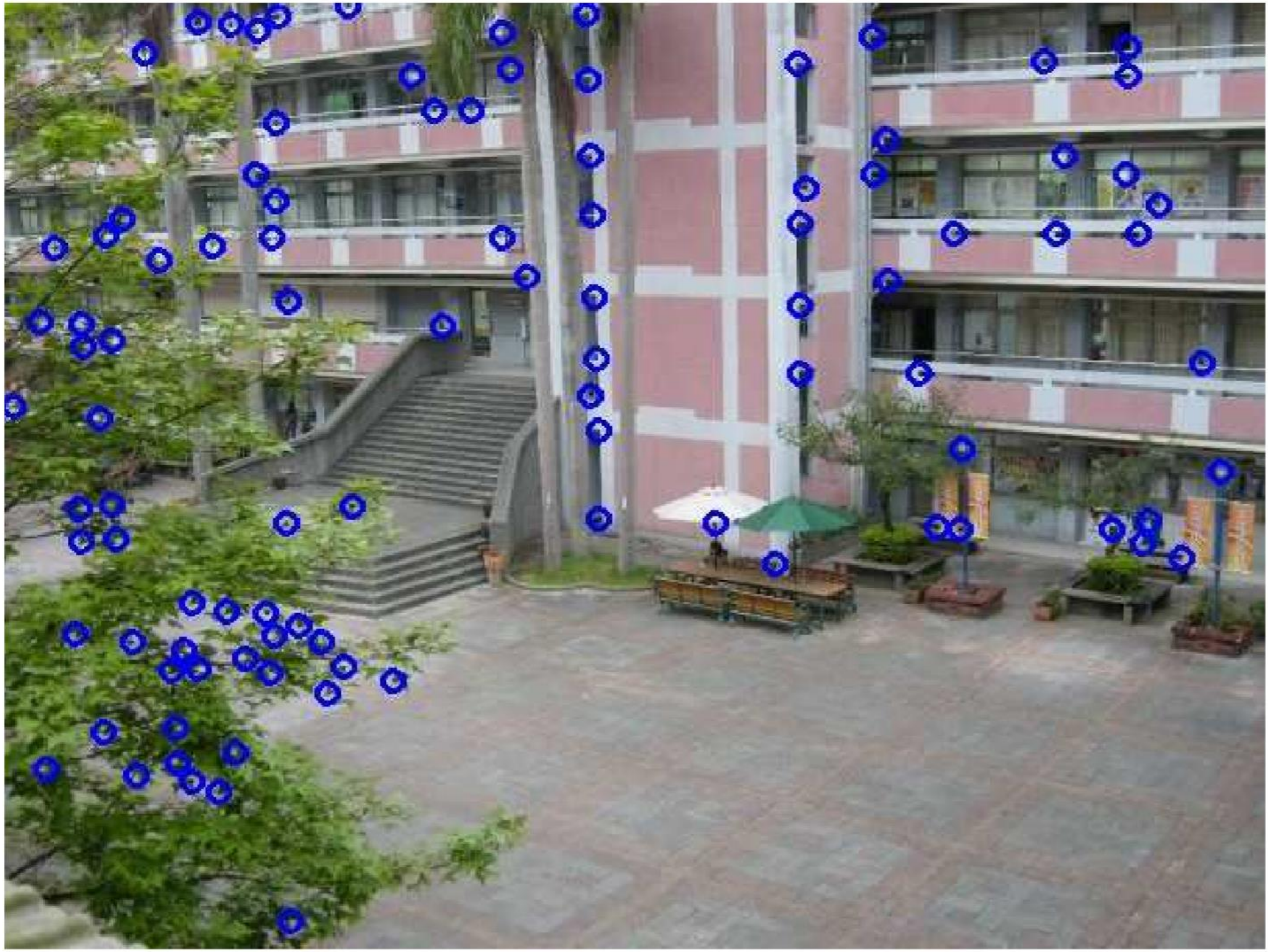
Harris Corner

- 1) Compute $\boldsymbol{g_x}$, $\boldsymbol{g_y}$ gradients at each point in image.
- 2) For every *NxN* block *A* of pixels.
  - (a) create 2x2 matrix $\boldsymbol{A} = \begin{bmatrix} \sum_{(x,y)\in B} g_x^2 & \sum_{(x,y)\in B} g_x g_y \\ \sum_{(x,y)\in B} g_x g_y & \sum_{(x,y)\in B} g_y^2 \end{bmatrix}$
  - (b) Compute eigenvalues $\boldsymbol{\lambda_1, \lambda_2}$ of this matrix.
  - (c) if $\boldsymbol{\lambda_1, \lambda_2}$ are both >$\boldsymbol{\tau}$, accept $\boldsymbol{A}$ as good feature.
  - Matlab commond: Corner.

上海科技大学 ShanghaiTech University

Sample result (large response):

Many corners are detected near each other. So, better to find local maximum.



With non-maximum suppression, detected corners are more spread out.

Shi-tomasi

- Also referred as harris corner detector, is good for finding at a certain scale.
  - There may be many of these corners.
  - Only at (small) certain scale. (11x11-15x15)
- Better features than simple corners.
  - Multi-scale feature (window of different sizes).
  - “Best” scale for a feature.
  - Viewpoint/rotation invariant neighborhood to describe feature.
  - SIFT features.

SIFT: Key point extraction

- **Stands for scale invariant feature transform**
- **Patented by University of British Columbia**
- **Similar to the one used in primate visual system (human, ape, monkey, etc.)**

- D. Lowe. Distinctive image features from scale invariant key points., International Journal of Computer Vision 2014

Scale Invariance

- In many applications, the scale of the object of interest may vary in different images.



- Simple but inefficient solution:

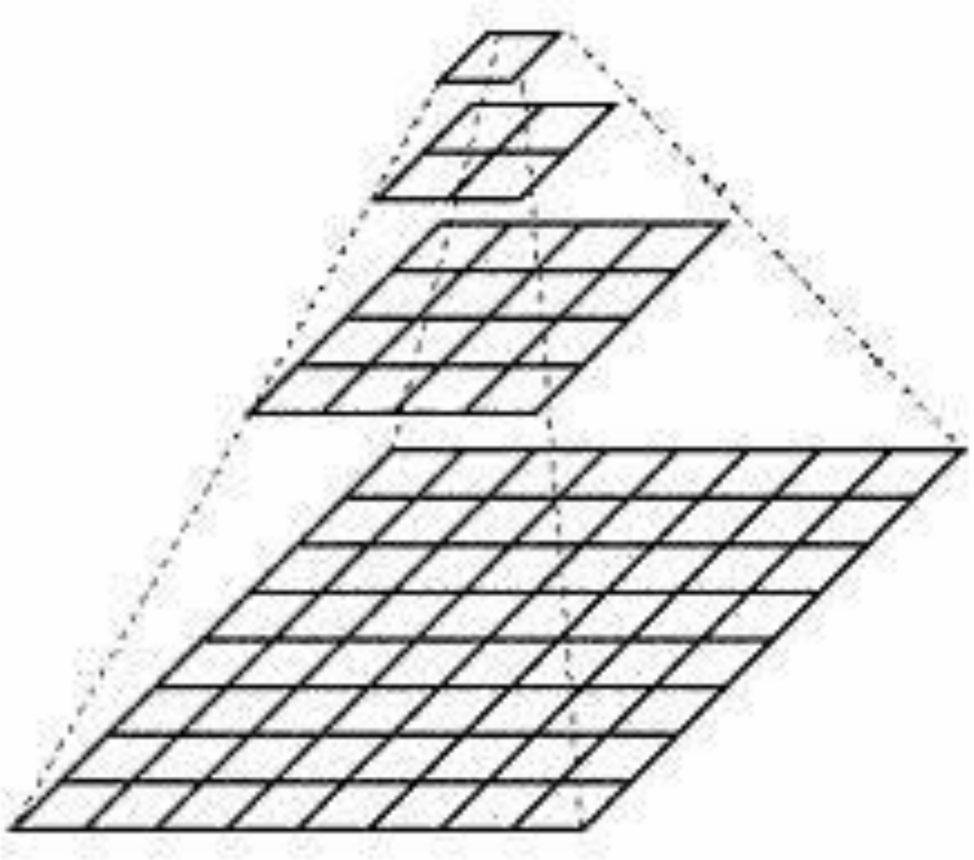
  Extract features at many different scales.

  Match them to the object's known features at a particular scale.

- More efficient solution:

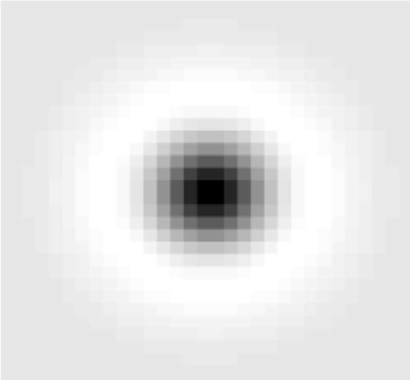
  Extract features that are invariant to scale

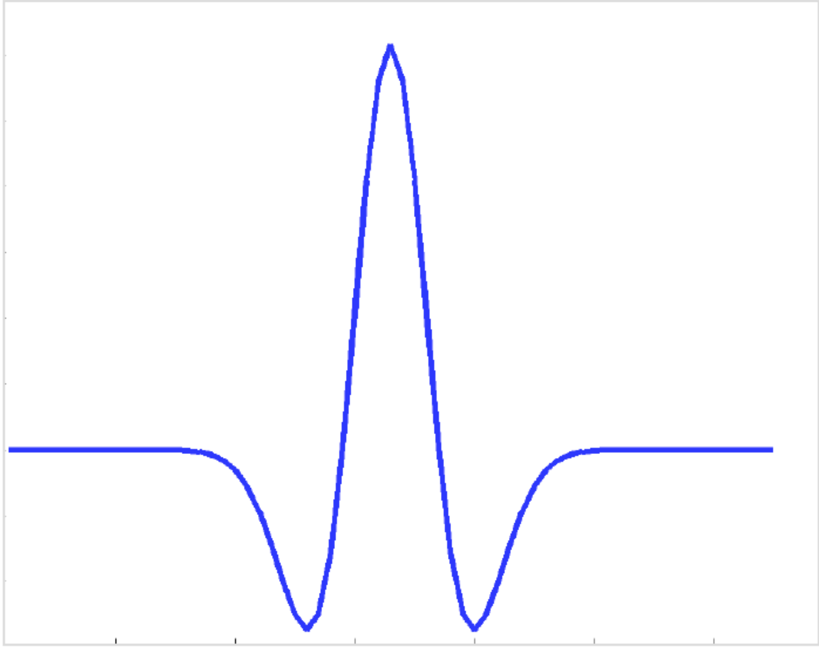
Image Pyramid



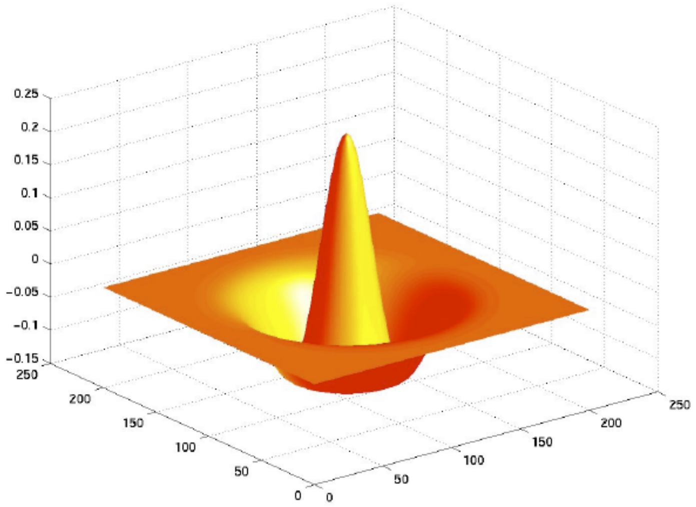
Log filter： Second derivative of a Gaussian

$$G''(x,y) = -\frac{1}{\pi\sigma^4}\left(1-\frac{x^2+y^2}{2\sigma^2}\right)e^{-\frac{x^2+y^2}{2\sigma^2}}$$





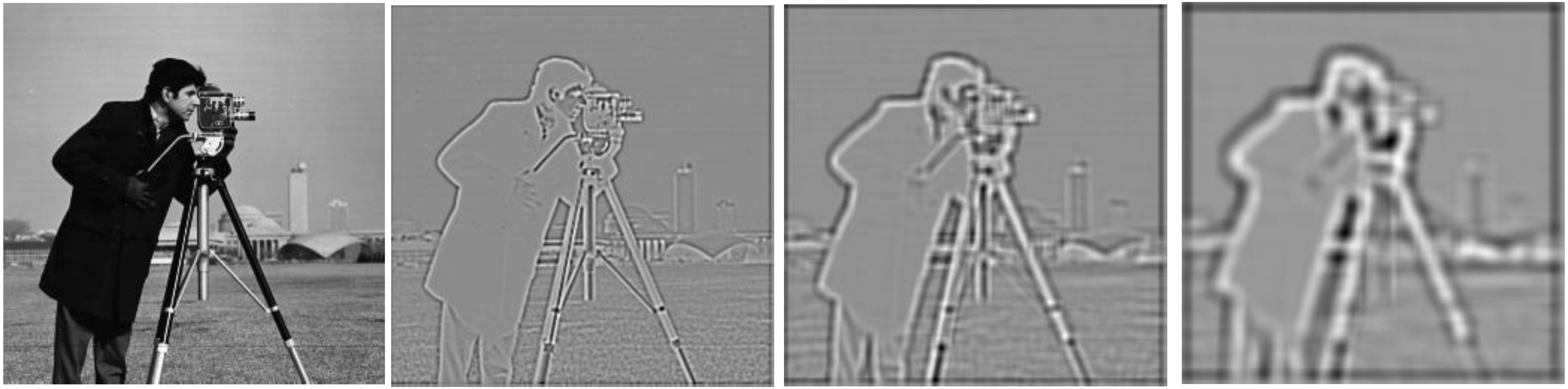
2D analog



LoG "Mexican Hat"

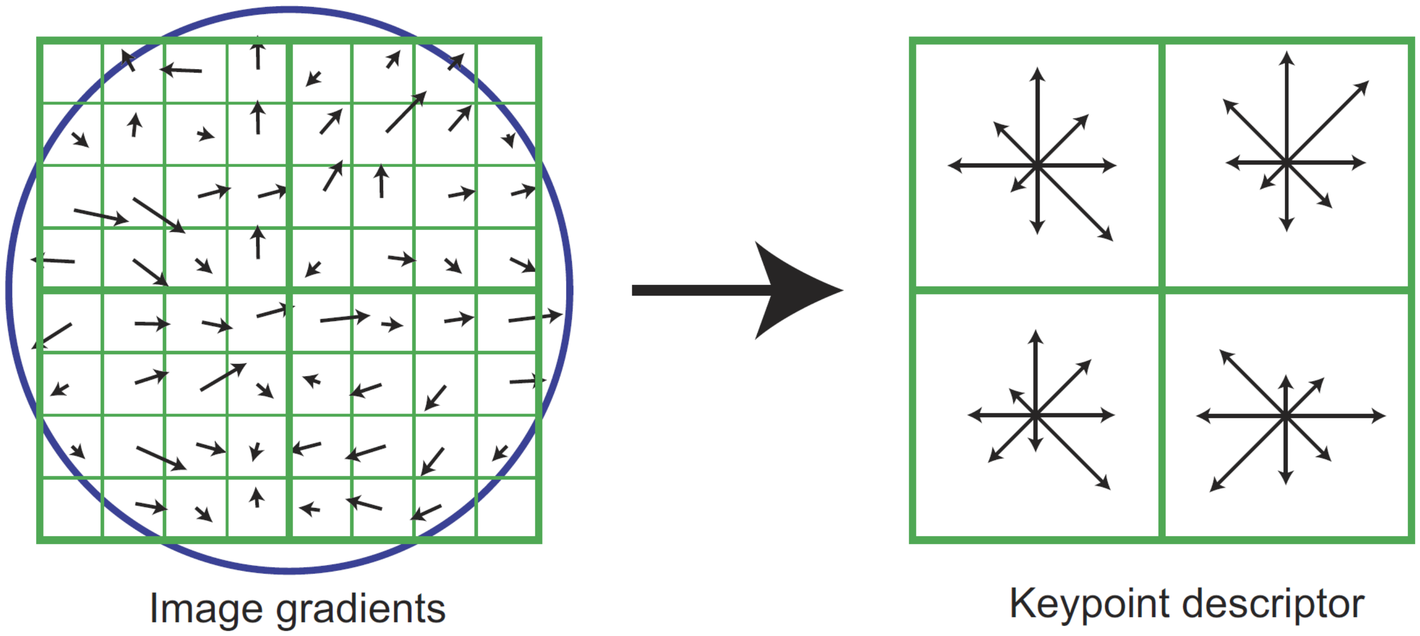
Effect of LoG Filter

Sigma = 1 Sigma = 4 Sigma = 10



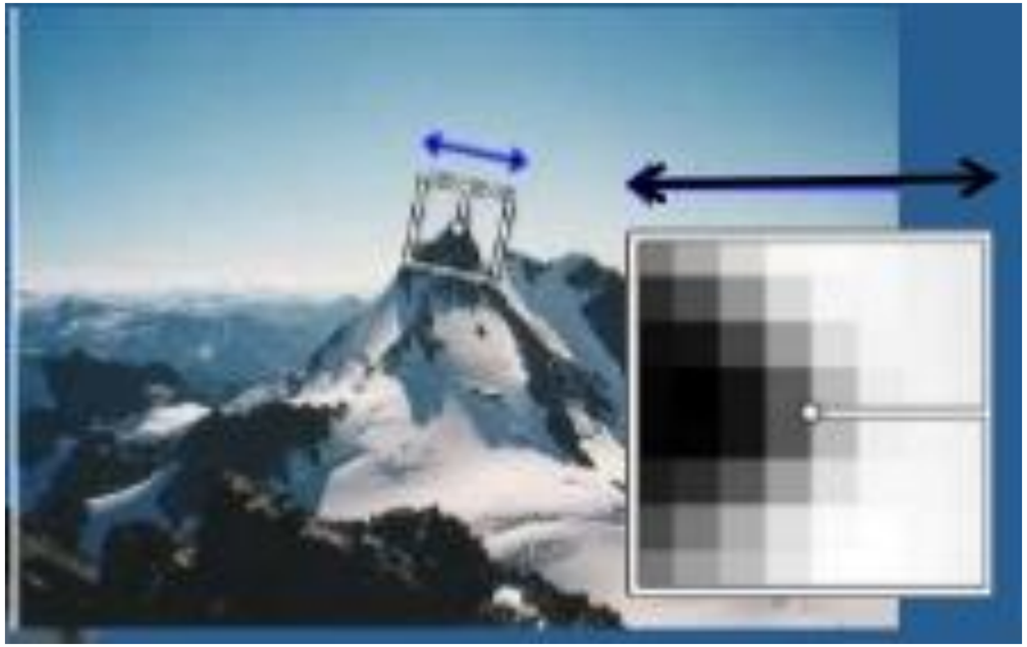
Band-Pass Filter (suppresses both high and low frequencies)

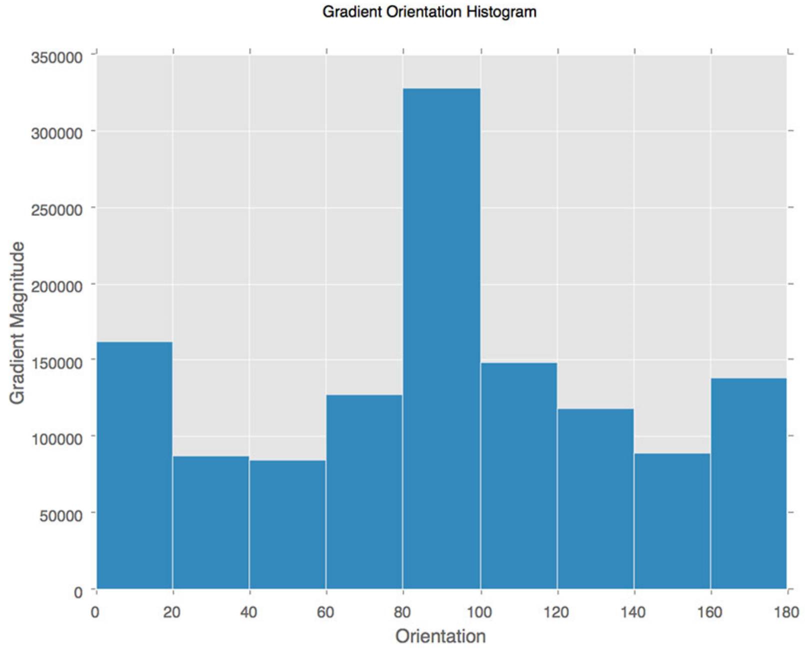
Keypoint descriptors



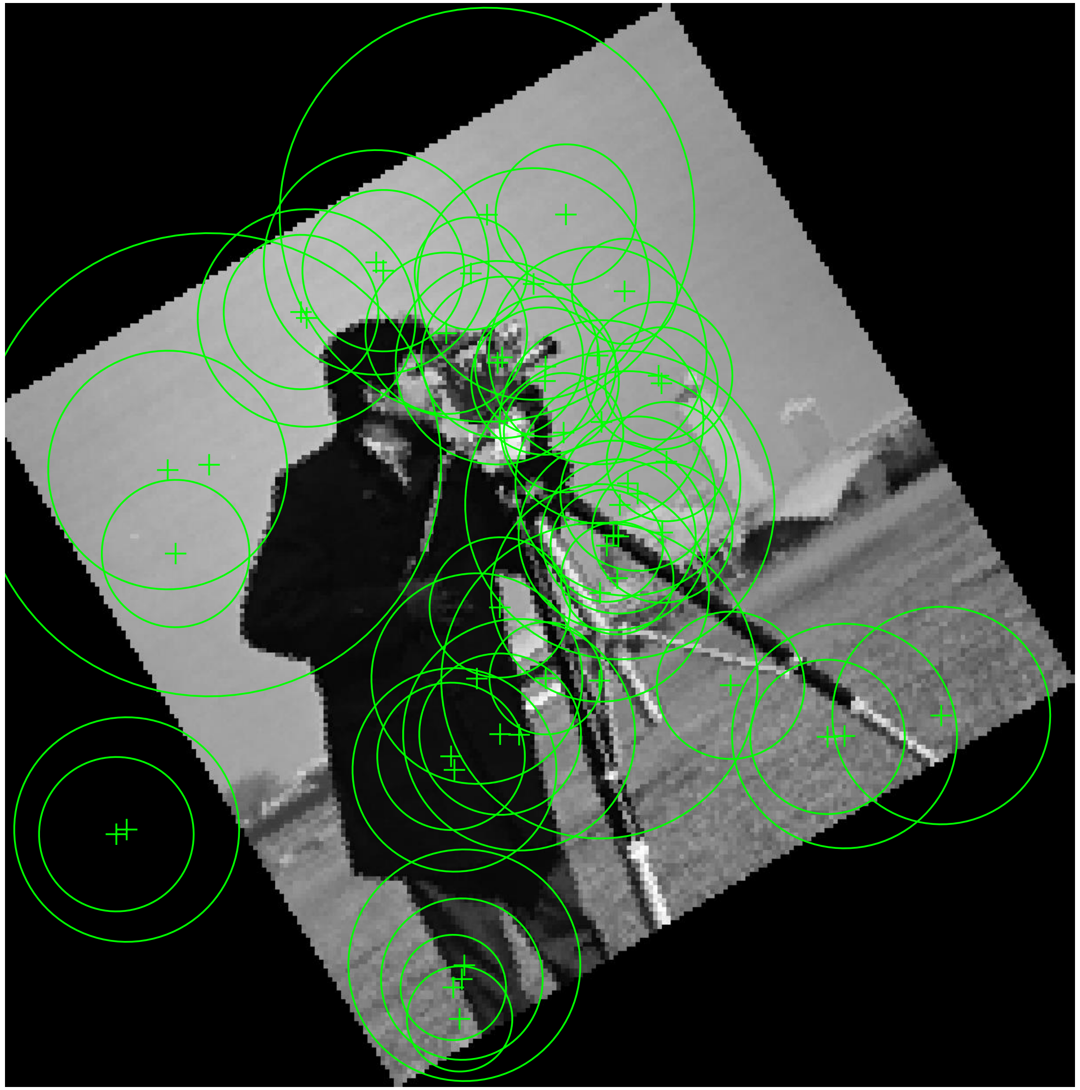
- Crop key point local feature to 4x4 sub-regions.
- Compute the quantized orientation histogram in each subregion.
- This operation allows for significant shift and rotation for the key point.

Orientation assignment





- Compute orientation histogram, based on orientation quantization.
- Find the dominant orientation for key point.
- Extract local region around key point and orient the region to its right direction.



Take home message

- Template matching is good for similar structures.
- Harris corner detector and Tomasi's algorithm find corner points.
- SIFT keypoint: invariant to scale.
- SIFT descriptors: invariant to scale, orientation, illumination change.
- Variants of SIFT: PCA-SIFT, SURF, GLOH.