WST665/CS770A: Web-Scale Image Retrieval
Scale Invariant Region Selection

Sung-Eui Yoon
(윤성의)

Course URL:
http://sglab.kaist.ac.kr/~sungeui/IR

KAIST

What we will learn today?

- Local invariant features
  - Motivation
  - Requirements, invariances
- Keypoint localization
  - Harris corner detector
  - Hessian detector
- **Scale invariant region selection**
  - **Automatic scale selection**
  - **Laplacian-of-Gaussian detector**
  - **Difference-of-Gaussian detector**
  - **Combinations**
- Local descriptors
  - An intro

Fei-Fei Li

From Points to Regions...

- The Harris and Hessian operators define interest points.
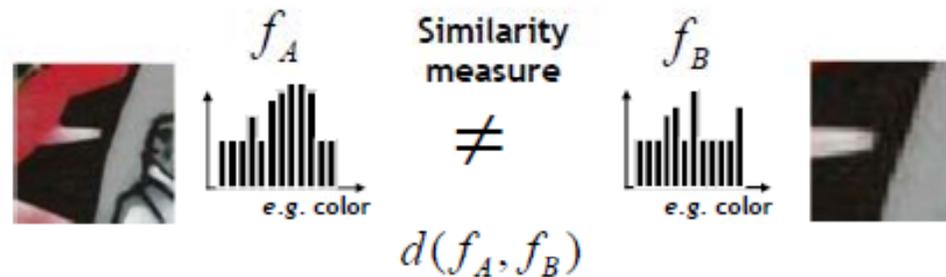 - Precise localization
  - High repeatability

- In order to compare those points, we need to compute a descriptor over a region.
  - How can we define such a region in a scale invariant manner?

- *I.e. how can we detect scale invariant interest regions?*
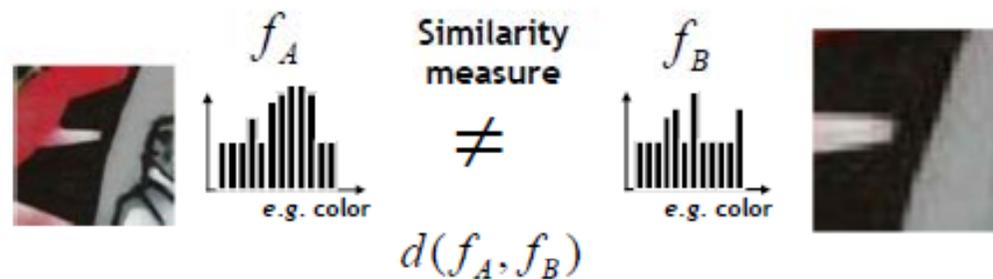
Fei-Fei Li

Naïve Approach: Exhaustive Search

- Multi-scale procedure
  - Compare descriptors while varying the patch size

$$f_A \qquad \text{Similarity measure} \qquad f_B$$

$$\neq$$

$$d(f_A, f_B)$$

Naïve Approach: Exhaustive Search

- Multi-scale procedure
  - Compare descriptors while varying the patch size

$$d(f_A, f_B)$$

Naïve Approach: Exhaustive Search

- Multi-scale procedure
  - Compare descriptors while varying the patch size

f_A Similarity measure f_B

e.g. color $\neq$ e.g. color

$$d(f_A, f_B)$$

Naïve Approach: Exhaustive Search

- Comparing descriptors while varying the patch size
    - Computationally inefficient
    - Inefficient but possible for matching
   - Prohibitive for retrieval in large databases
    - Prohibitive for recognition

f_A Similarity measure f_B

$$=$$

$$d(f_A, f_B)$$

Automatic Scale Selection

- Solution:
  - Design a function on the region, which is "scale invariant" (*the same for corresponding regions, even if they are at different scales*)

    > Example: average intensity. For corresponding regions (even of different sizes) it will be the same.
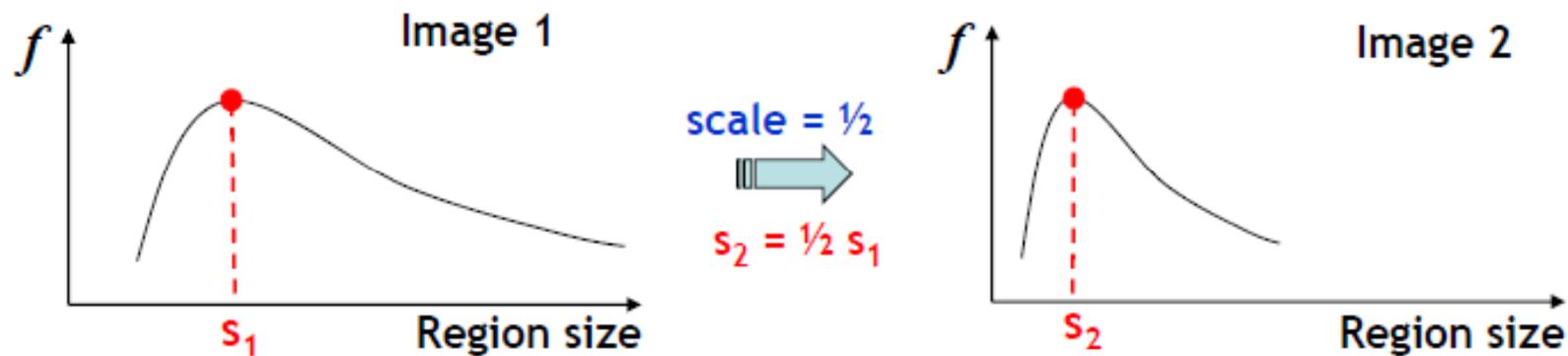
  - For a point in one image, we can consider it as a function of region size (patch width)
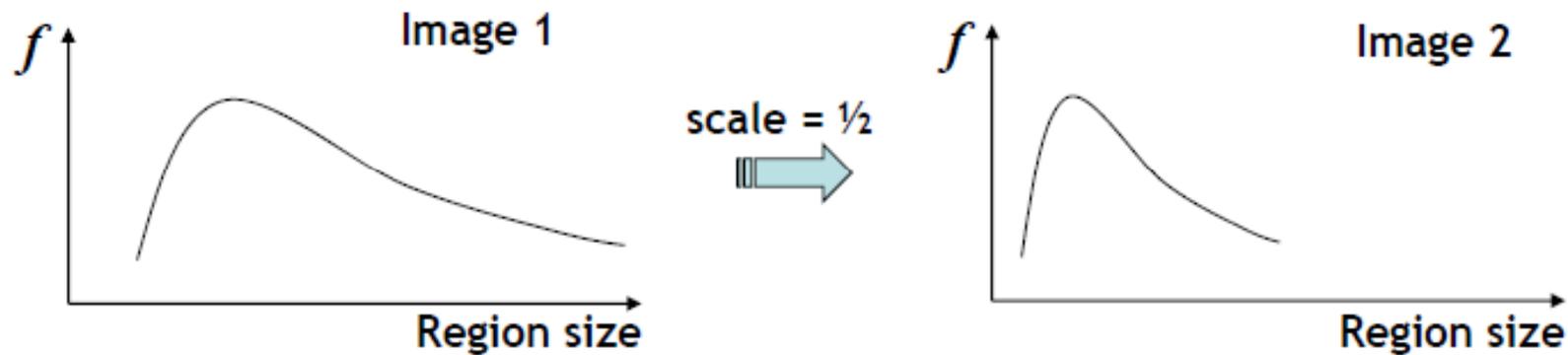
Slide credit: Kristen Grauman

Automatic Scale Selection

- Common approach:

  - Take a local maximum of this function.

  - Observation: region size for which the maximum is achieved should be *invariant* to image scale.
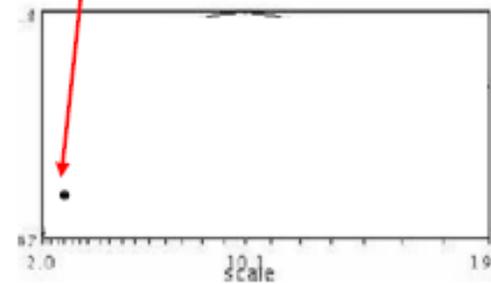
Automatic Scale Selection

- Function responses for increasing scale (scale signature)
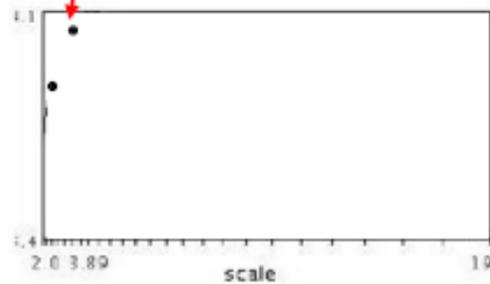

$$f(I_{i_1 \dots i_m}(x, \sigma))$$

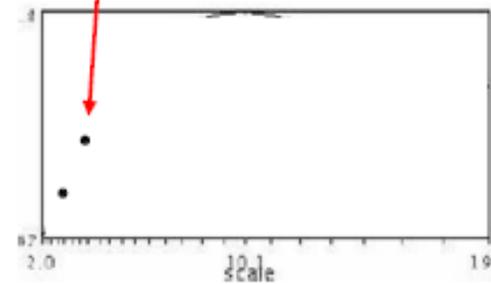$$f(I_{i_1 \dots i_m}(x', \sigma))$$

Automatic Scale Selection

- Function responses for increasing scale (scale signature)
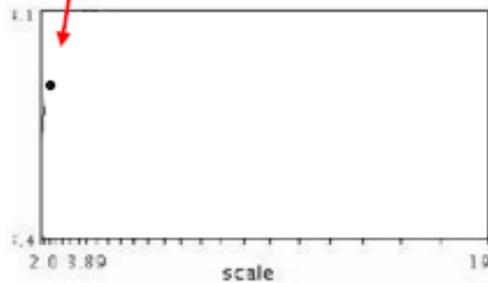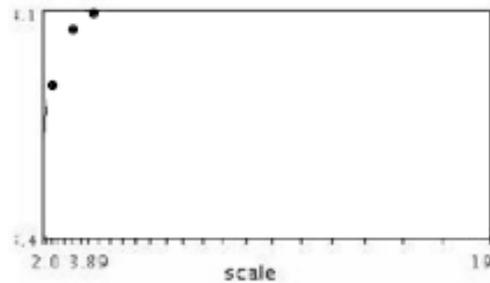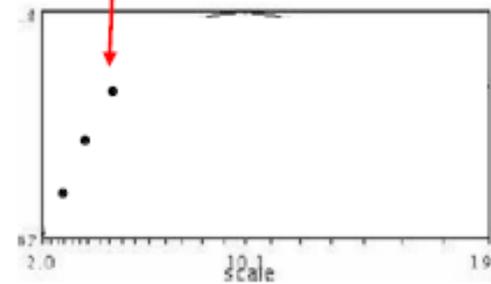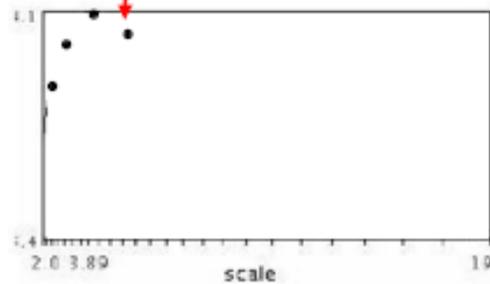

$$f(I_{i_1 \ldots i_m}(x, \sigma))$$

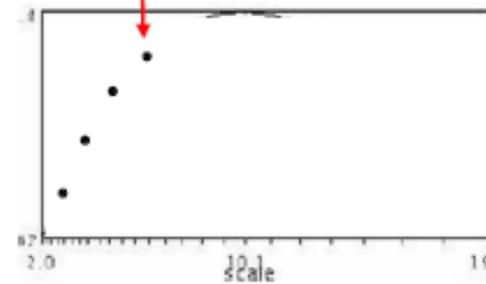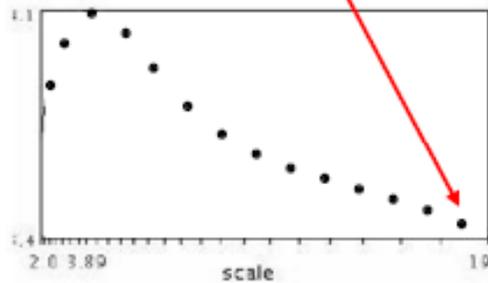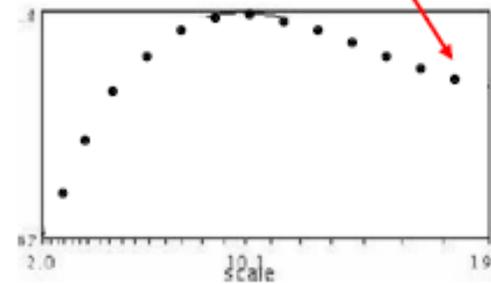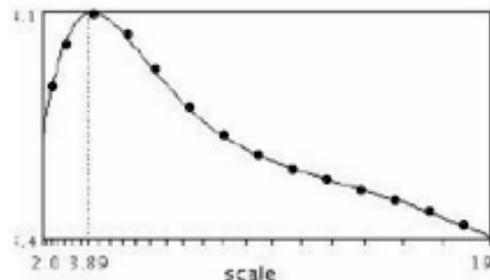$$f(I_{i_1 \ldots i_m}(x', \sigma))$$

Slide credit: Krystian Mikolajczyk

Automatic Scale Selection
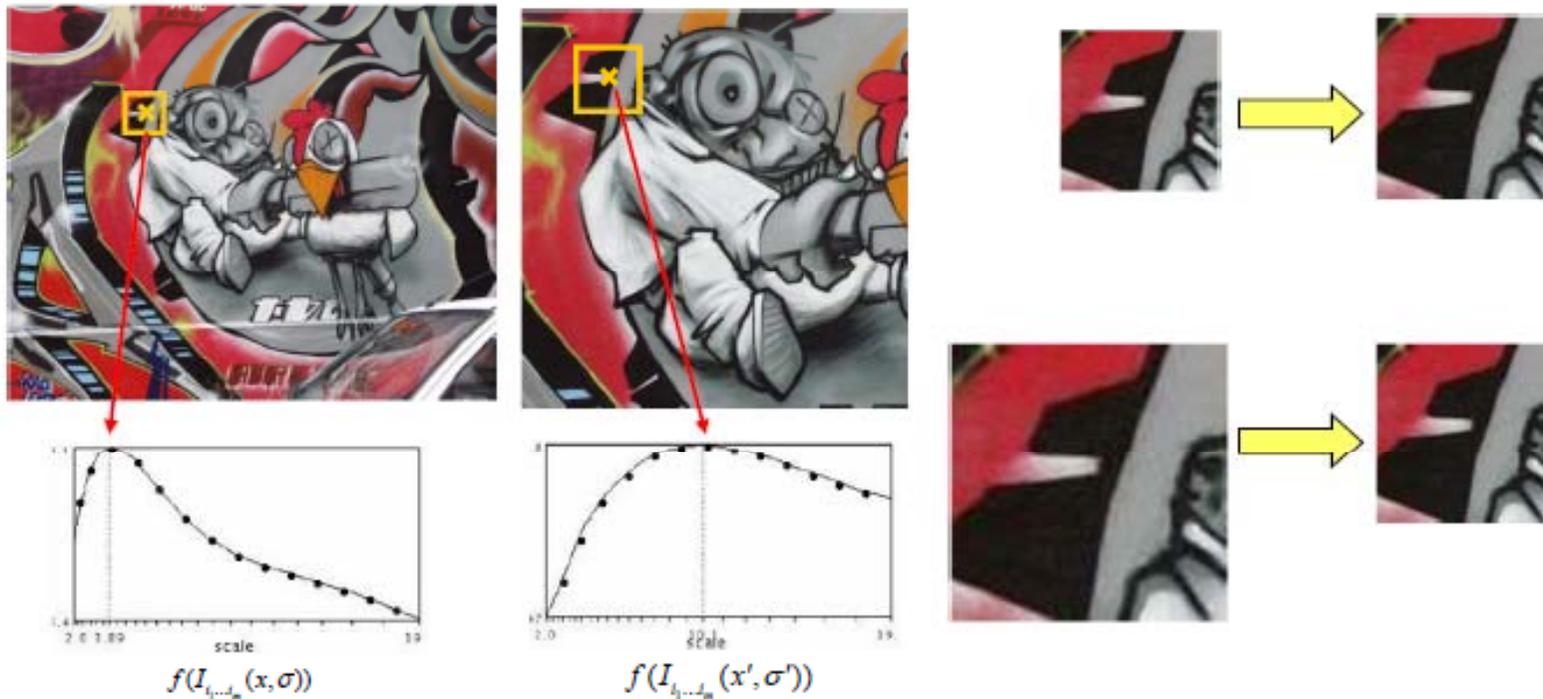
- Function responses for increasing scale (scale signature)

$$f(I_{i_1 \ldots i_m}(x, \sigma))$$

$$f(I_{i_1 \ldots i_m}(x', \sigma))$$

Automatic Scale Selection

- Function responses for increasing scale (scale signature)

$$f(I_{i_1 \ldots i_m}(x, \sigma))$$

$$f(I_{i_1 \ldots i_m}(x', \sigma))$$

Automatic Scale Selection

- Function responses for increasing scale (scale signature)

$$f(I_{i_1 \ldots i_m}(x, \sigma))$$

$$f(I_{i_1 \ldots i_m}(x', \sigma))$$

Automatic Scale Selection

- Function responses for increasing scale (scale signature)

$$f(I_{i_1 \dots i_m}(x, \sigma))$$
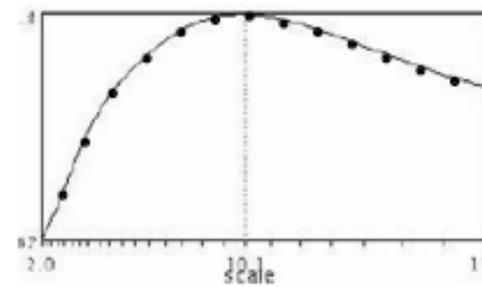
$$f(I_{i_1 \dots i_m}(x', \sigma'))$$

Automatic Scale Selection

- Normalize: Rescale to fixed size

$$f(I_{i_1\ldots i_m}(x,\sigma))$$

$$f(I_{i_1\ldots i_m}(x',\sigma'))$$

Automatic Scale Selection

- Normalize: Rescale to fixed size

$$f(I_{i_1 \ldots i_m}(x, \sigma))$$

$$f(I_{i_1 \ldots i_m}(x', \sigma'))$$
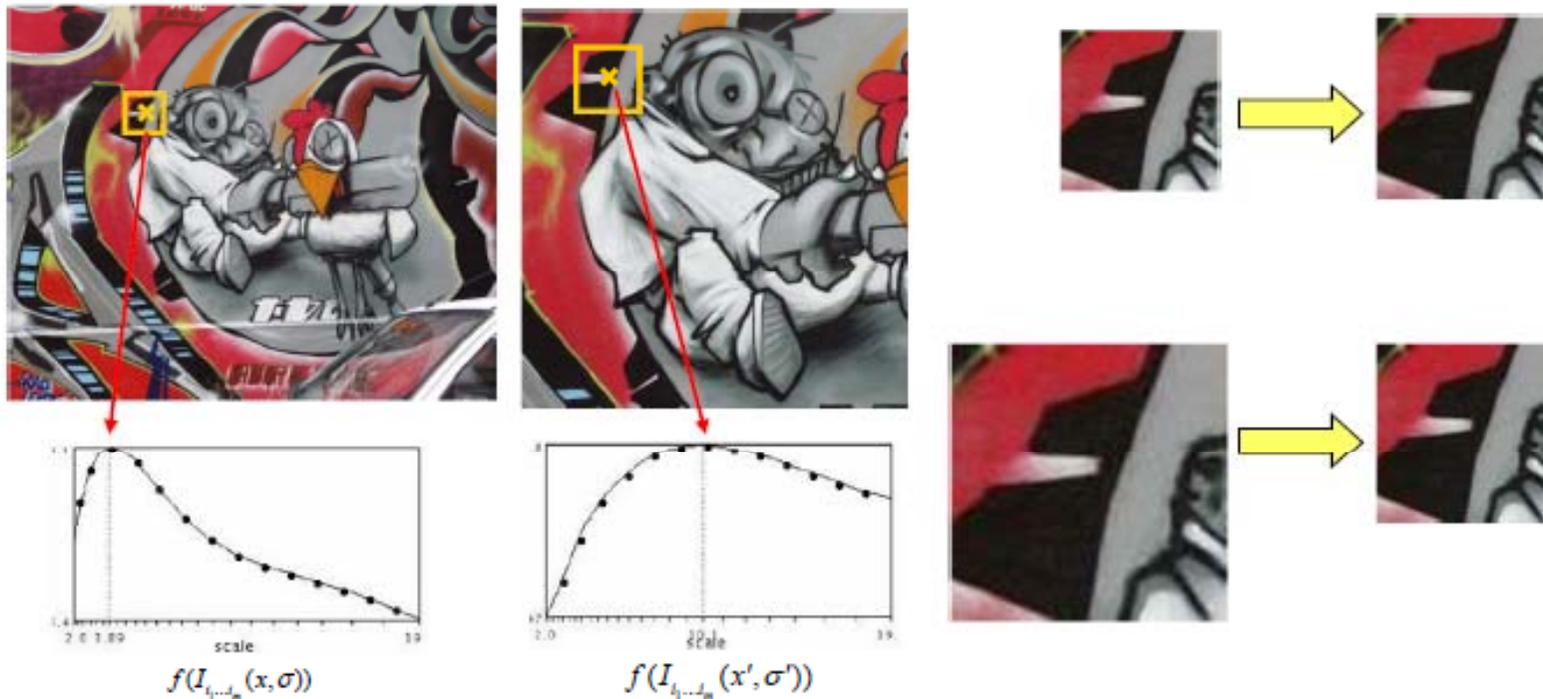
What Is A Useful Signature Function?

- Laplacian-of-Gaussian = "blob" detector
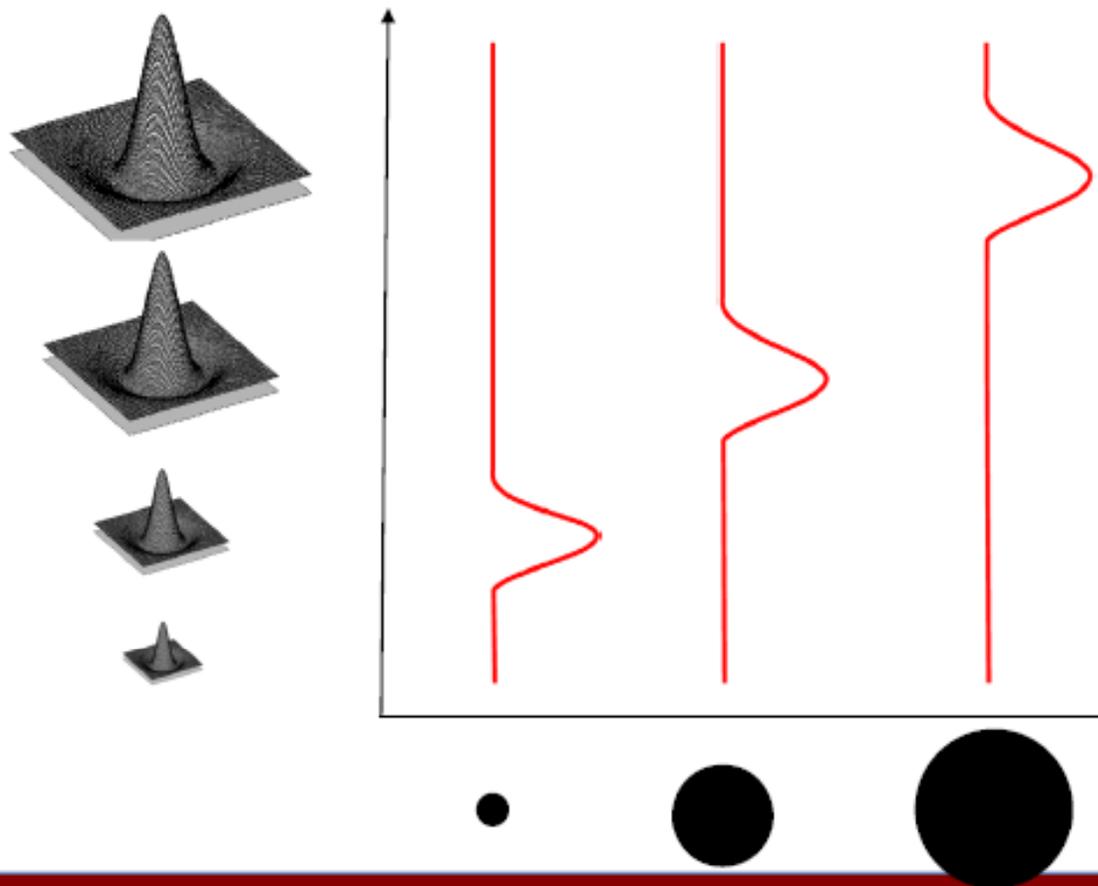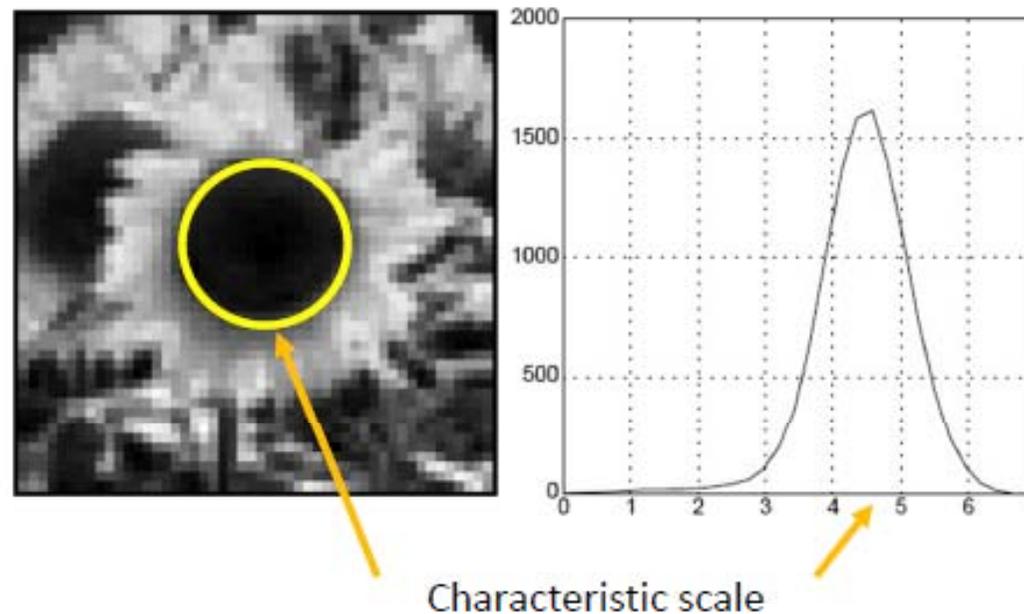
Characteristic Scale

- We define the *characteristic scale* as the scale that produces peak of Laplacian response


Characteristic scale

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

Laplacian-of-Gaussian (LoG)

- Interest points:
  - Local maxima in scale space of Laplacian-of-Gaussian

$$L_{xx}(\sigma) + L_{yy}(\sigma)$$

σ^5

σ^4

σ^3

σ^2

σ

Laplacian-of-Gaussian (LoG)

- Interest points:
  - Local maxima in scale space of Laplacian-of-Gaussian

$$L_{xx}(\sigma) + L_{yy}(\sigma)$$

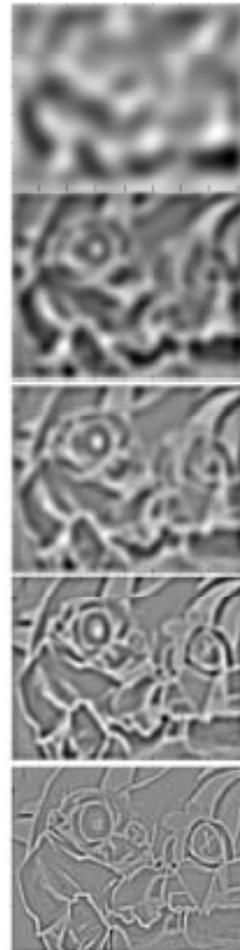σ^5

σ^4

σ^3

σ^2

σ

Scale

Laplacian-of-Gaussian (LoG)

- Interest points:
  - Local maxima in scale space of Laplacian-of-Gaussian

$$L_{xx}(\sigma) + L_{yy}(\sigma)$$

σ^5

σ^4

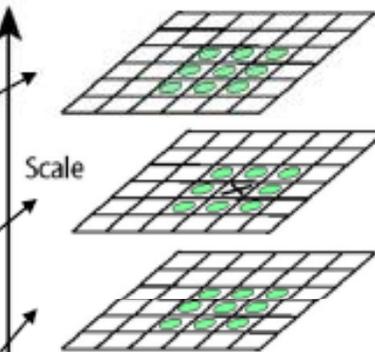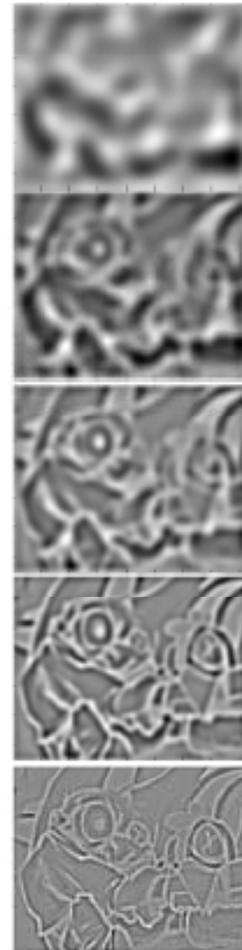σ^3

σ^2

σ
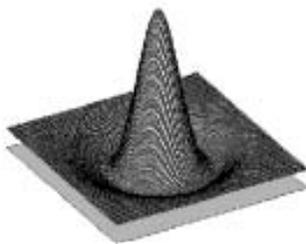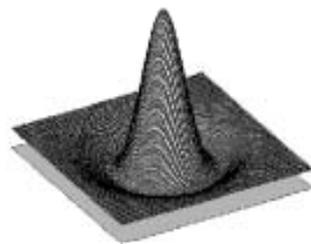
Scale

Slide adapted from

Laplacian-of-Gaussian (LoG)

- Interest points:
  - Local maxima in scale space of Laplacian-of-Gaussian
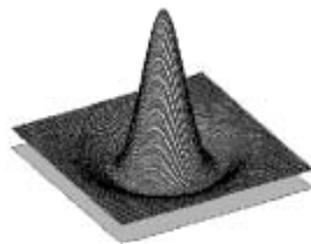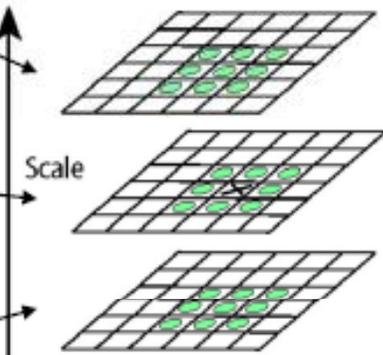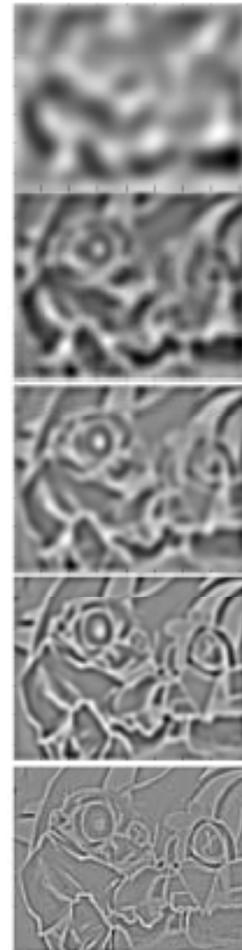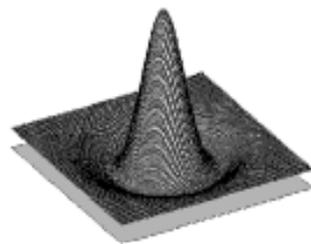

$$L_{xx}(\sigma) + L_{yy}(\sigma)$$

σ^5
σ^4
σ^3
σ^2
σ

Scale
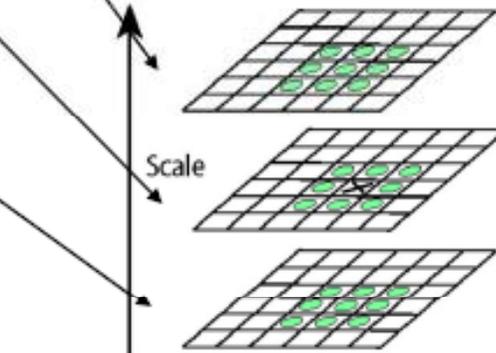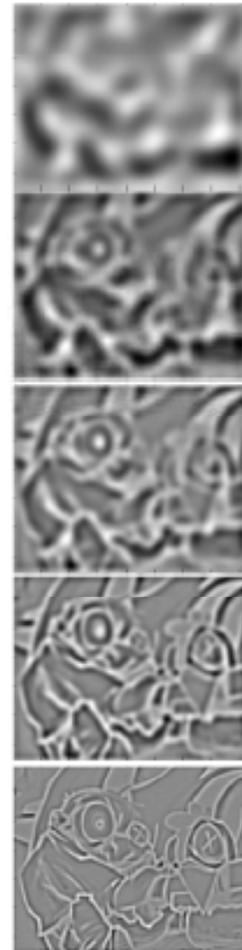
$\Rightarrow$ List of *(x, y, σ)*

LoG Detector: Workflow

LoG Detector: Workflow

sigma = 11.9912

LoG Detector: Workflow
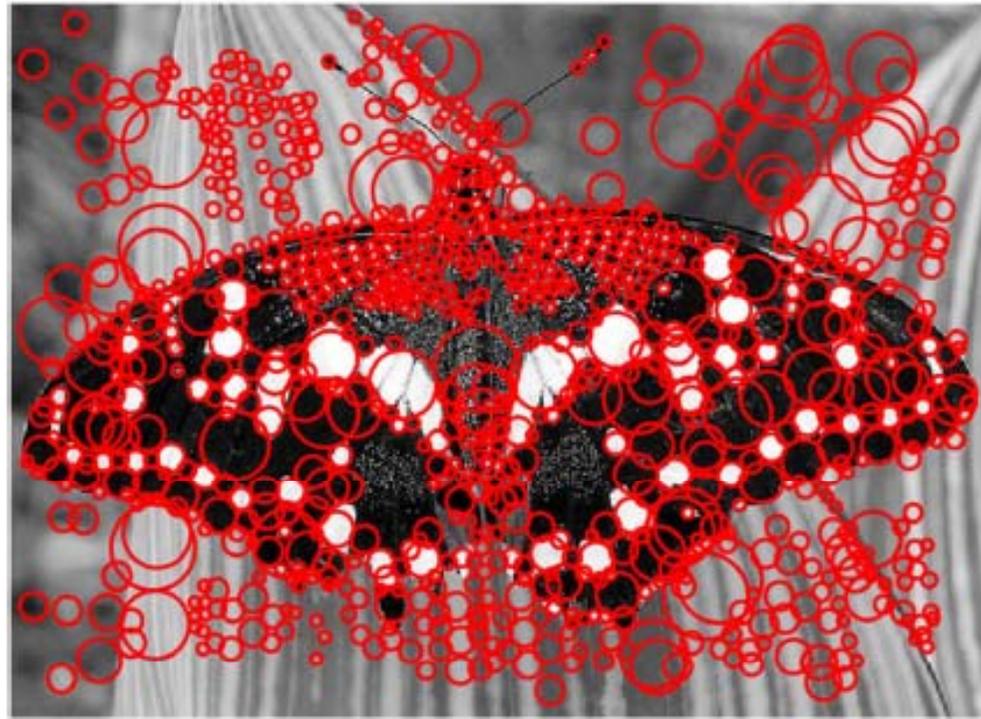
Technical Detail

- We can efficiently approximate the Laplacian with a difference of Gaussians:

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

(Laplacian)

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

(Difference of Gaussians)
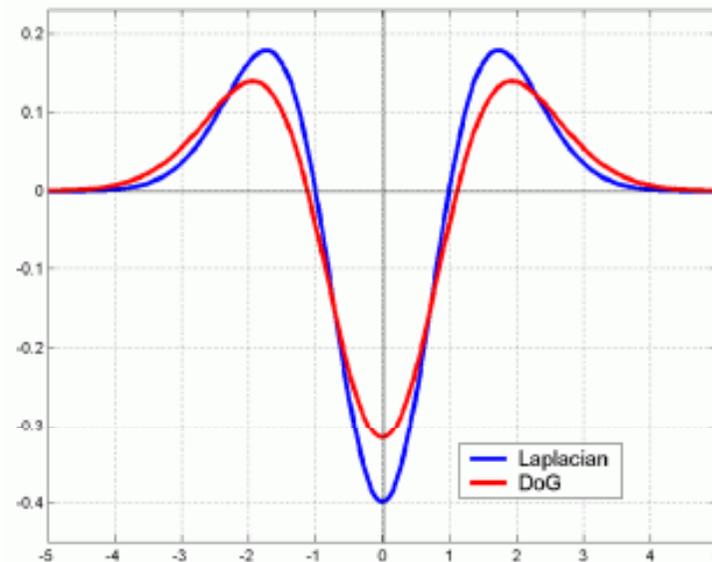
Difference-of-Gaussian (DoG)

- Difference of Gaussians as approximation of the LoG

  - This is used e.g. in Lowe's SIFT pipeline for feature detection.

- Advantages

  - No need to compute 2nd derivatives

  - Gaussians are computed anyway, e.g. in a Gaussian pyramid.
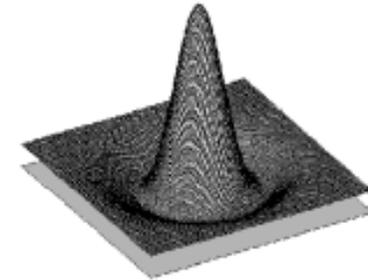
Key point localization with DoG

- Detect maxima of difference-of-Gaussian (DoG) in scale space

- Then reject points with low contrast (threshold)

- Eliminate edge responses

Scale

Candidate keypoints:
list of (x,y,σ)
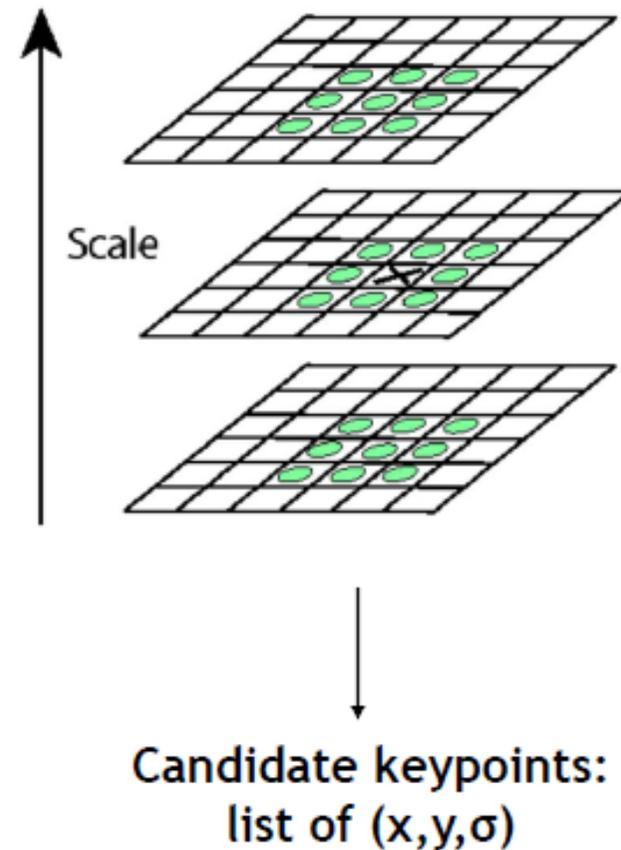
DoG – Efficient Computation

- Computation in Gaussian scale pyramid

Sampling with step $\sigma^4 = 2$

Original image

$\sigma = 2^{\frac{1}{4}}$

Scale (next octave)

Scale (first octave)

σ

Gaussian

Difference of Gaussian (DOG)

Results: Lowe's DoG

Example of Keypoint Detection

(a) 233x189 image

(b) 832 DoG extrema

(c) 729 left after peak value threshold

(d) 536 left after testing ratio of principle curvatures (removing edge responses)

Slide credit: David Lowe

Fei-Fei Li

Lecture 11 - 76

Harris-Laplace [Mikolajczyk '01]

1. Initialization: Multiscale Harris corner detection

Slide adapted from Krystian Mikolajczyk
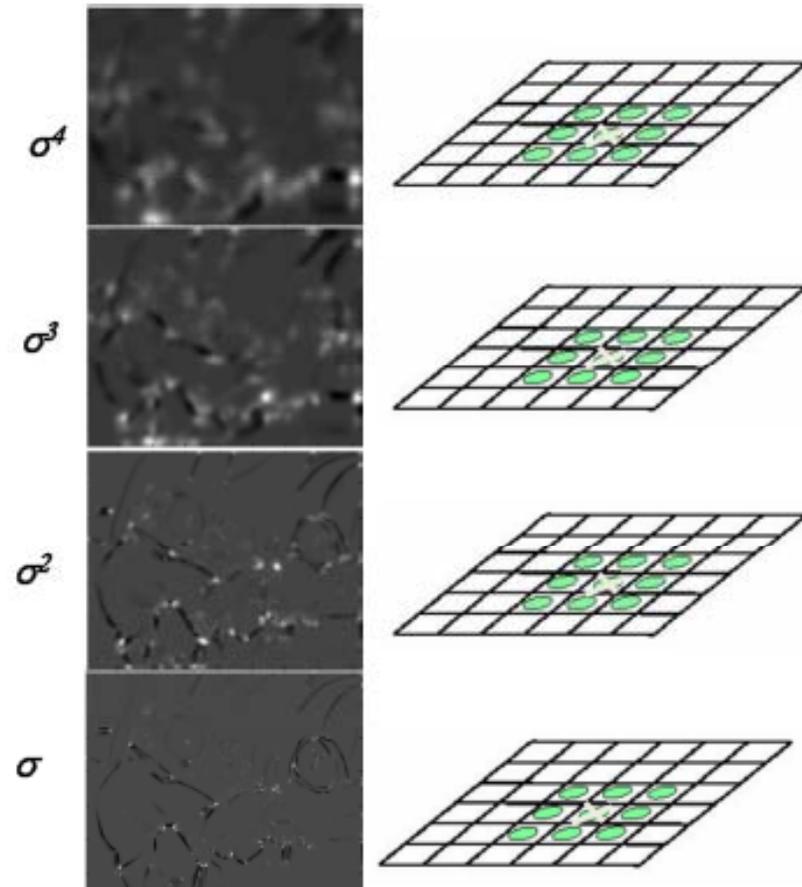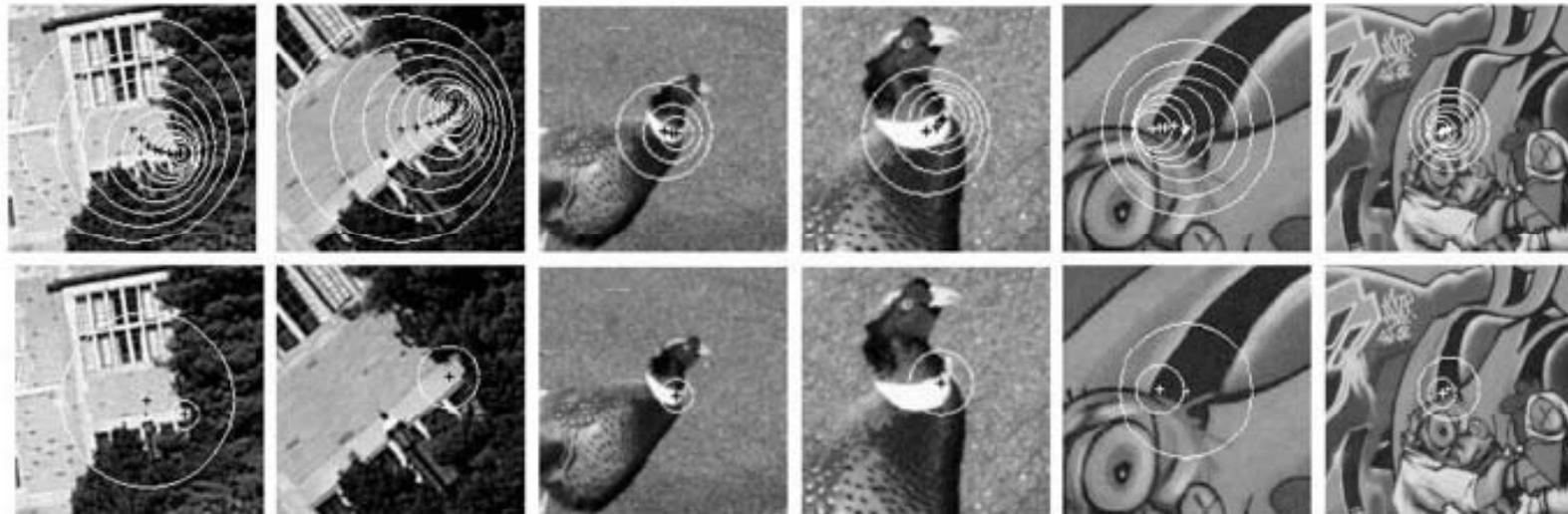
Computing Harris function   Detecting local maxima

Harris-Laplace [Mikolajczyk '01]

1. Initialization: Multiscale Harris corner detection

2. Scale selection based on Laplacian
   (same procedure with Hessian $\Rightarrow$ Hessian-Laplace)

Harris points

Harris-Laplace points

Summary: Scale Invariant Detection

- 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).

- Two strategies
  - Laplacian-of-Gaussian (LoG)
  - Difference-of-Gaussian (DoG) as a fast approximation

  - *These can be used either on their own, or in combinations with single-scale keypoint detectors (Harris, Hessian).*

What we will learn today?

- Local invariant features
  - Motivation
  - Requirements, invariances
- Keypoint localization
  - Harris corner detector
  - Hessian detector
- Scale invariant region selection
  - Automatic scale selection
  - Laplacian-of-Gaussian detector
  - Difference-of-Gaussian detector
  - Combinations
- **Local descriptors**
  - An intro

Local Descriptors

- We know how to detect points
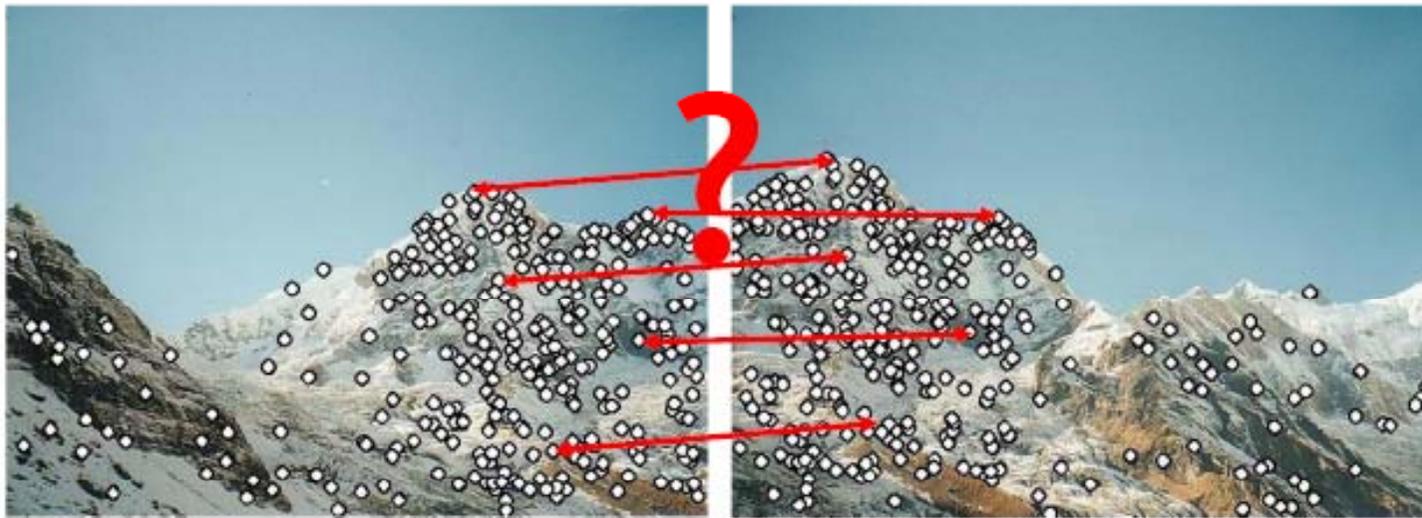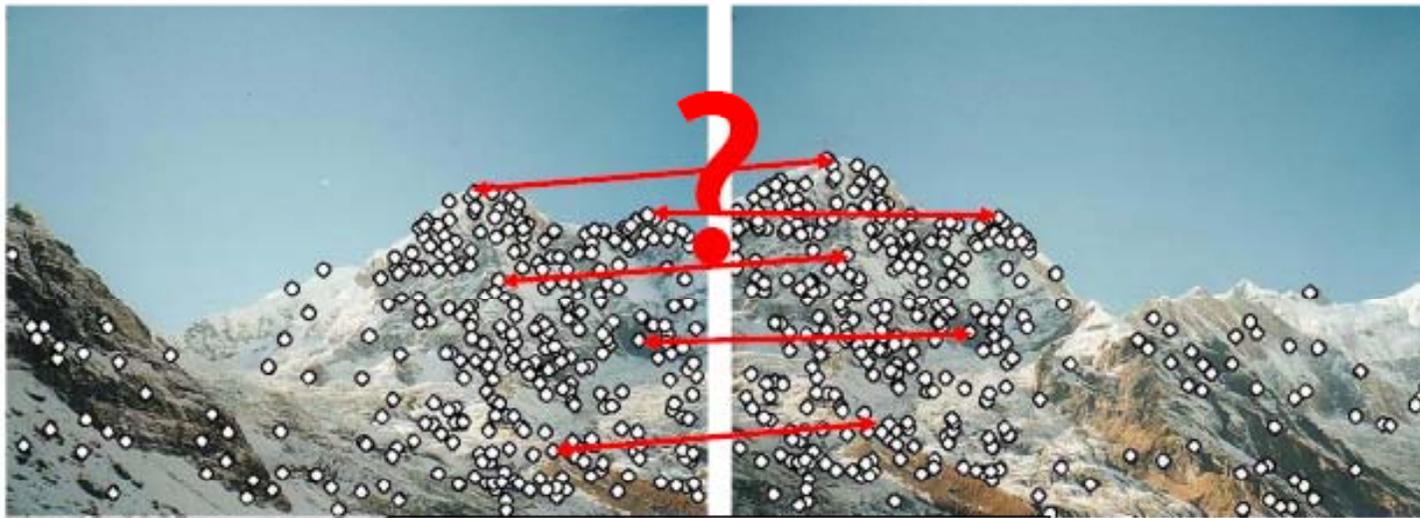
- Next question:

How to *describe* them for matching?

$\Rightarrow$ *Next lecture...*

Local Descriptors

- We know how to detect points

- Next question:

  How to *describe* them for matching?

Point descriptor should be:
1. Invariant
2. Distinctive

Slide credit: Kristen Grauman

Next Time

- Local descriptors (e.g., SIFT)

KAIST

Homework for Every Class

- Go over the next lecture slides
- Come up with one question on what we have discussed today and submit at the beginning of the next class
  - 0 for no questions
  - 2 for typical questions
  - 3 for questions with thoughts
  - 4 for questions that surprised me

KAIST