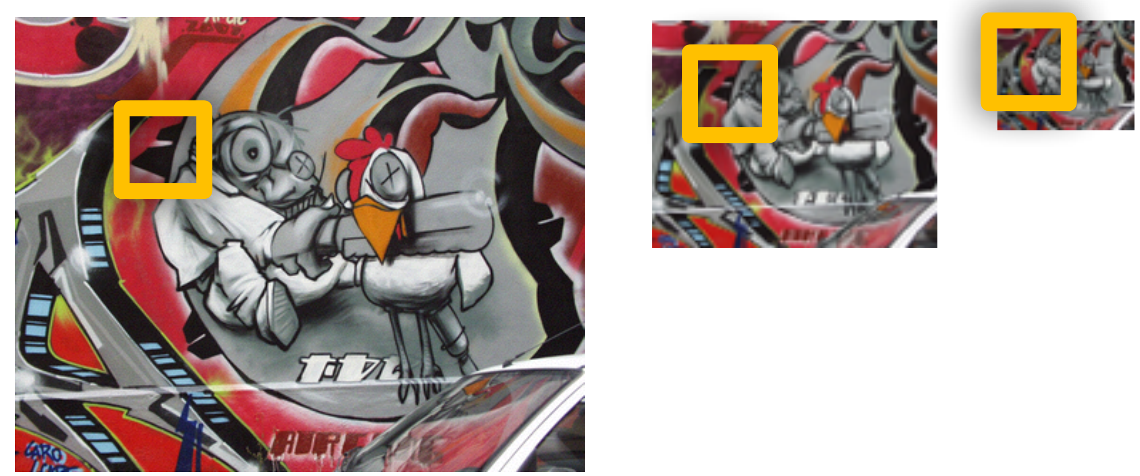
CSC 589 Introduction to Computer Vision

Lecture 17
Scale Invariance and Feature description

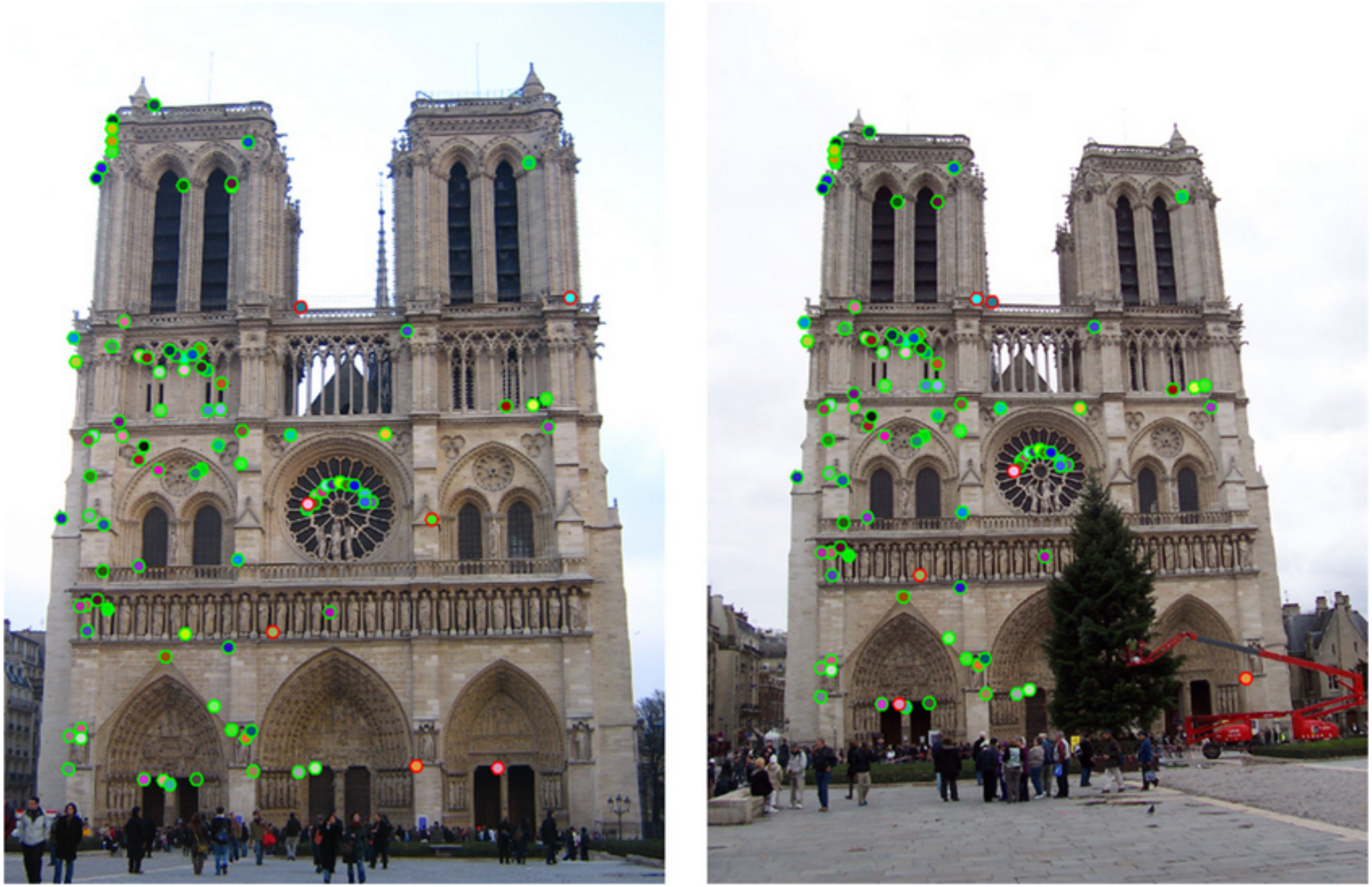


Bei Xiao

Spring, 2014

American University

Project 3: Feature Matching



The top 100 most confident local feature matches from a baseline implementation of project 2. In this case, 93 were correct (highlighted in green) and 7 were incorrect (highlighted in red).

Project 4: Automatic Paranoma

http://www.panoramas.dk/

Project 3: Feature Description and Matching (next three lectures)

- Local feature detection: Harris Corner, Chapter 4.1
- Local feature description: MOPS, SIFT, Chapter 4.12
- Local feature matching: Sum of squared distance or ratio test, 4.1.3.

Detections at multiple scales

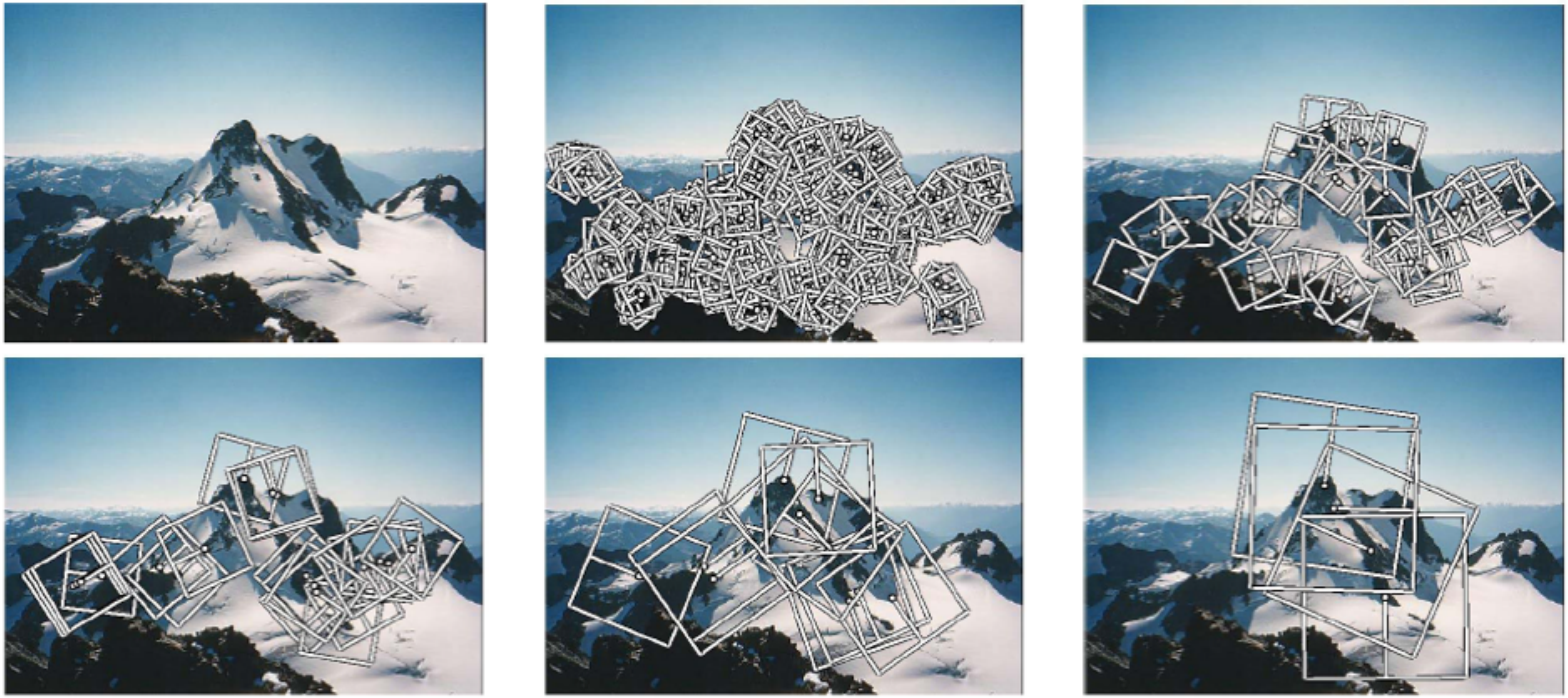
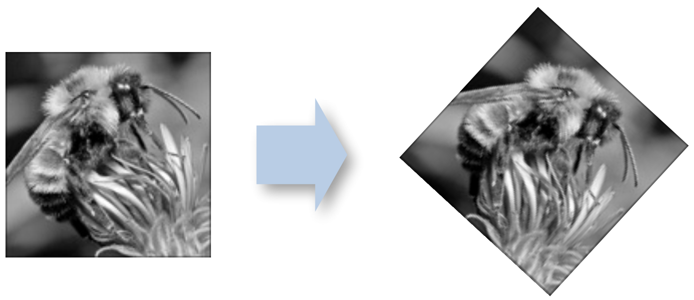


Figure 1. Multi-scale Oriented Patches (MOPS) extracted at five pyramid levels from one of the Matier images. The boxes show the feature orientation and the region from which the descriptor vector is sampled.

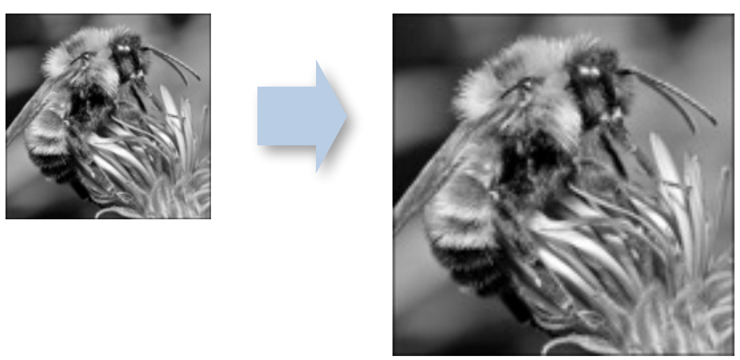
Image transformations

- Geometric

Rotation

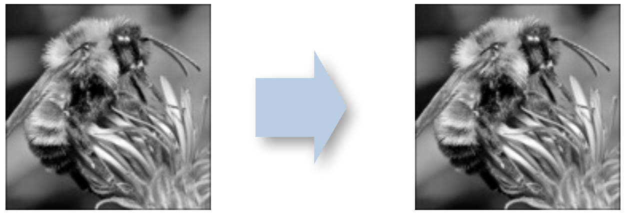


Scale



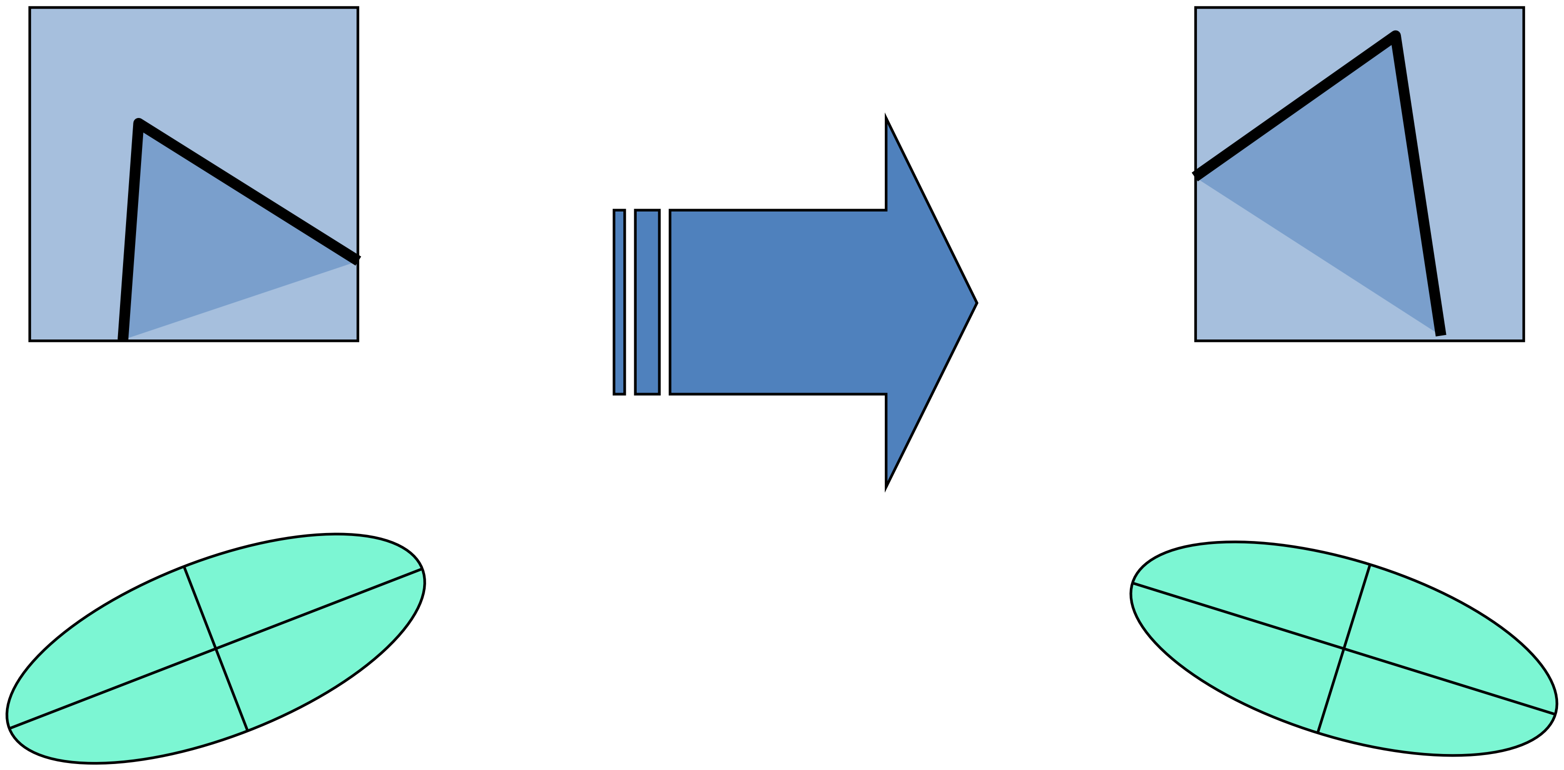
- Photometric

Intensity change



Harris Detector: Invariance Properties

- Rotation



Ellipse rotates but its shape (i.e. eigenvalues) remains the same

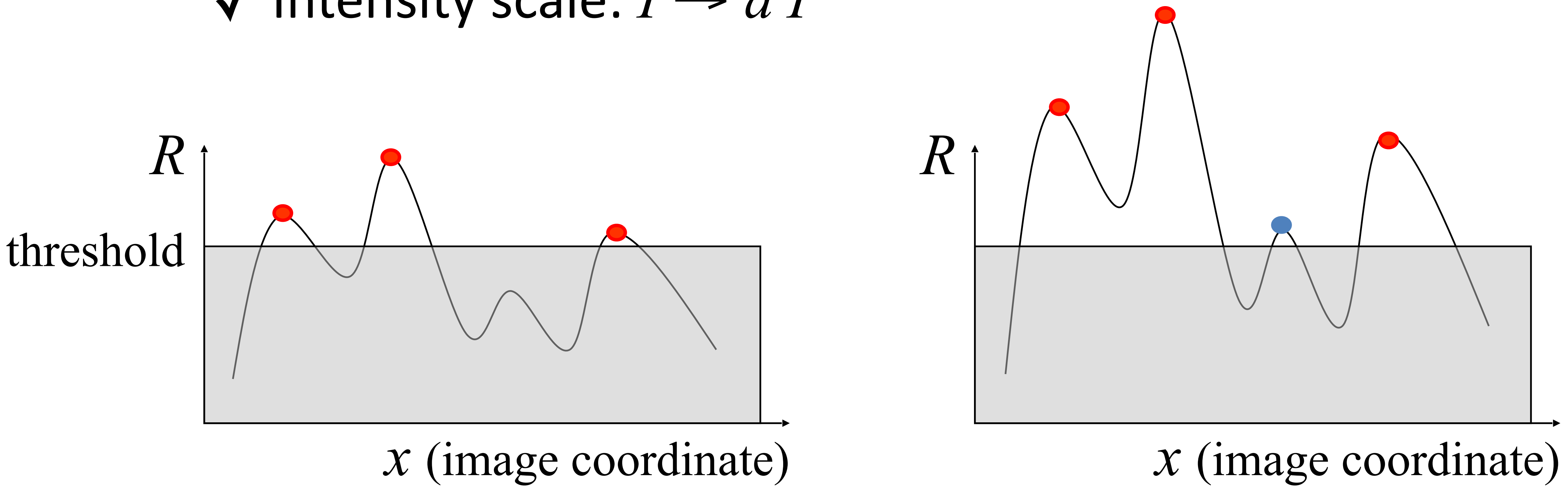
Corner response is invariant to image rotation

Harris Detector: Invariance Properties

- Affine intensity change: $I \rightarrow aI + b$

✓ Only derivatives are used => invariance to intensity shift $I \rightarrow I + b$

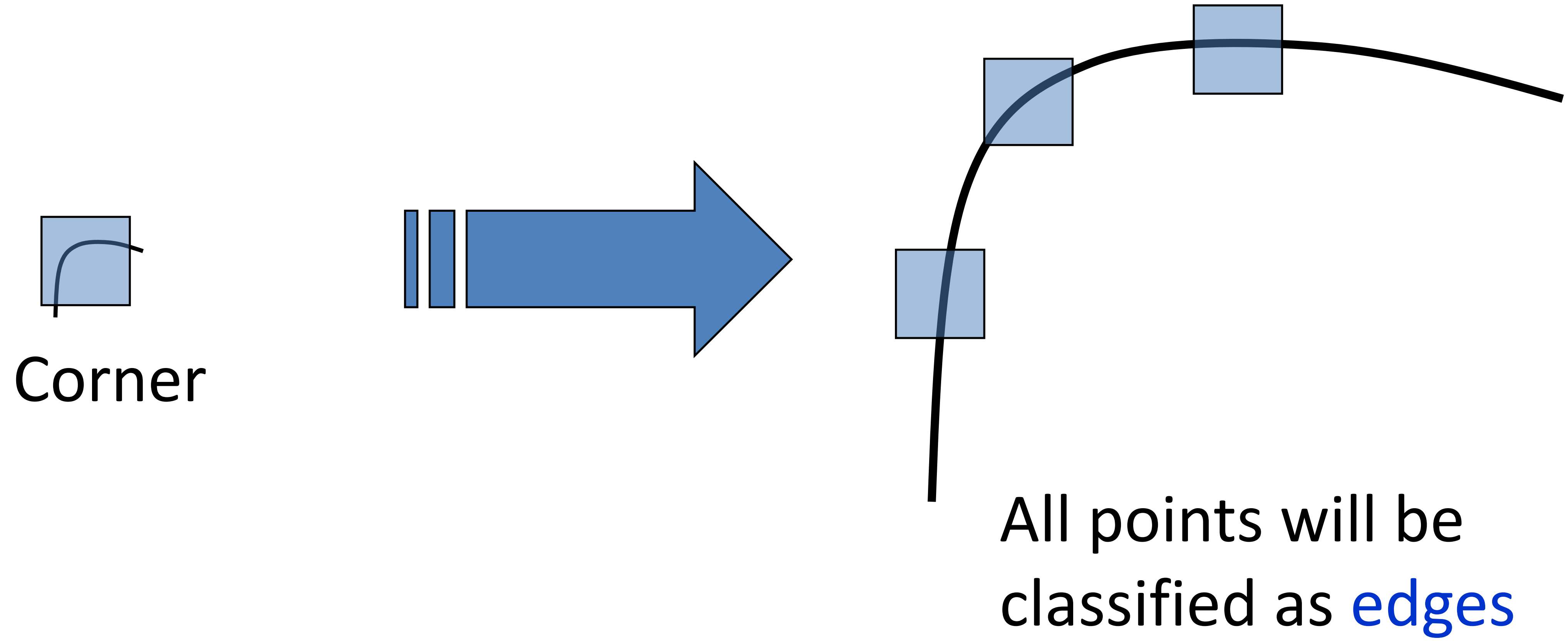
✓ Intensity scale: $I \rightarrow a\,I$



Partially invariant to affine intensity change

Harris Detector: Invariance Properties

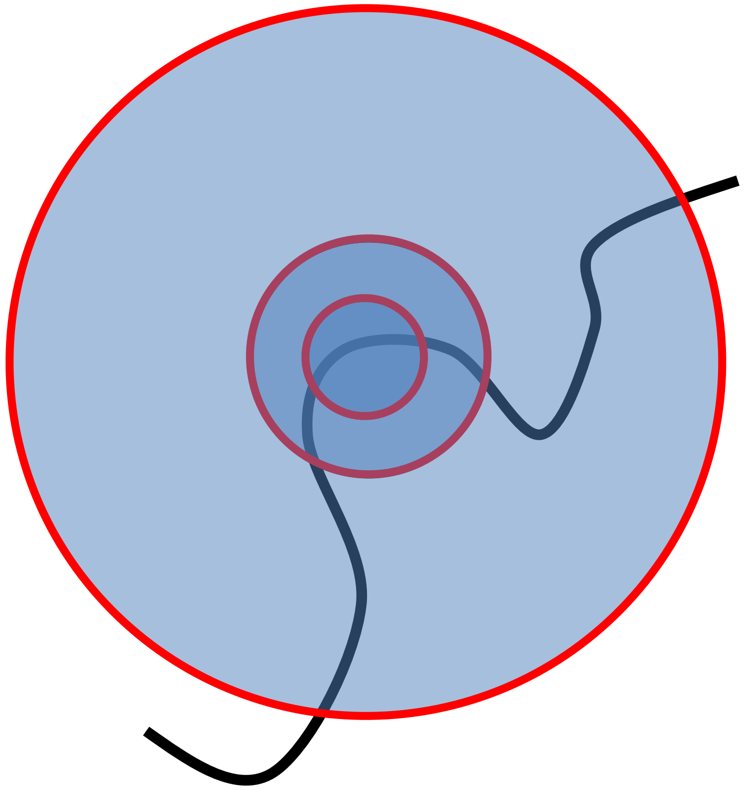
- Scaling (not invariant!) why?



Not invariant to scaling

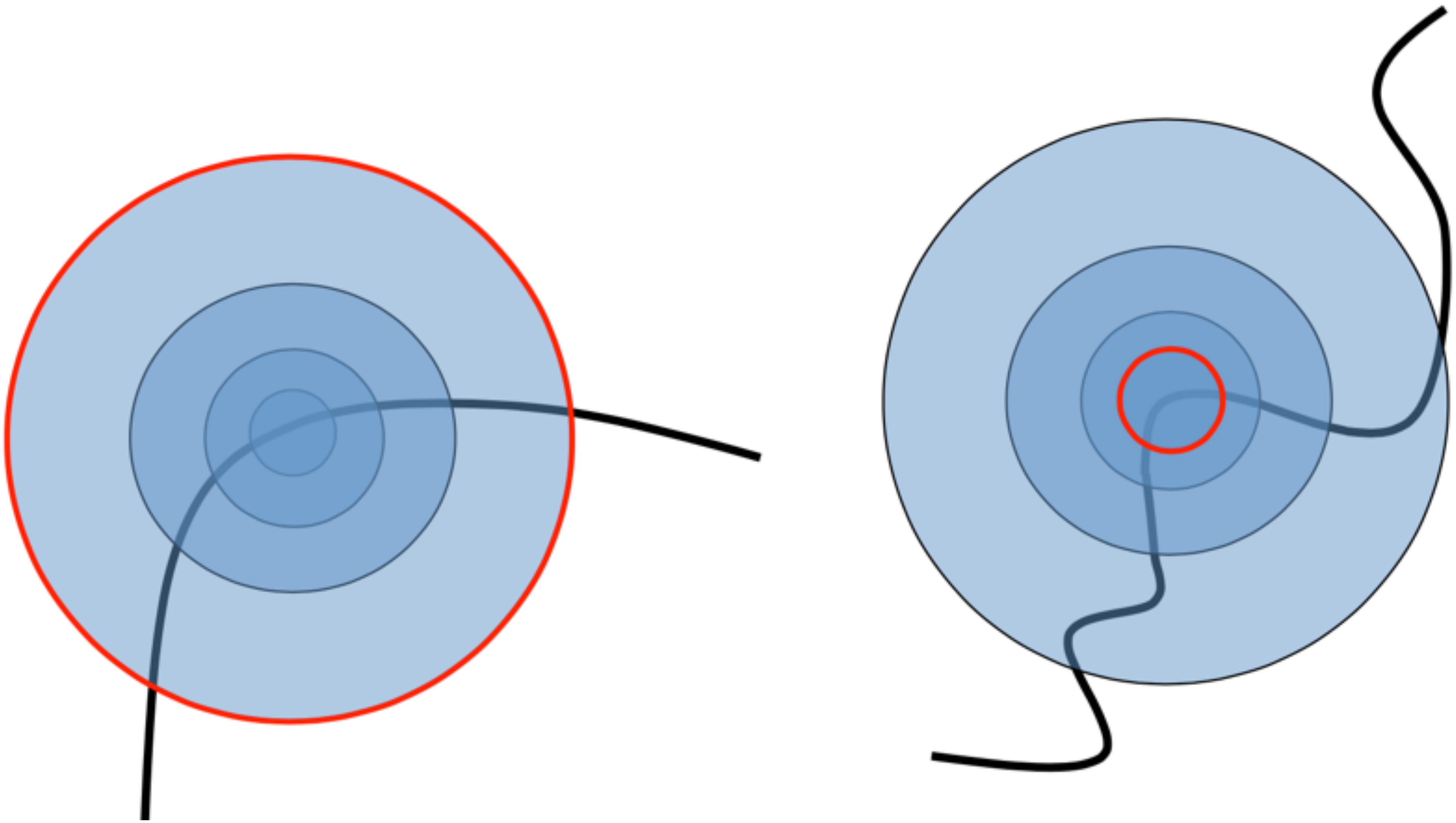
Scale invariant detection

Suppose you're looking for corners



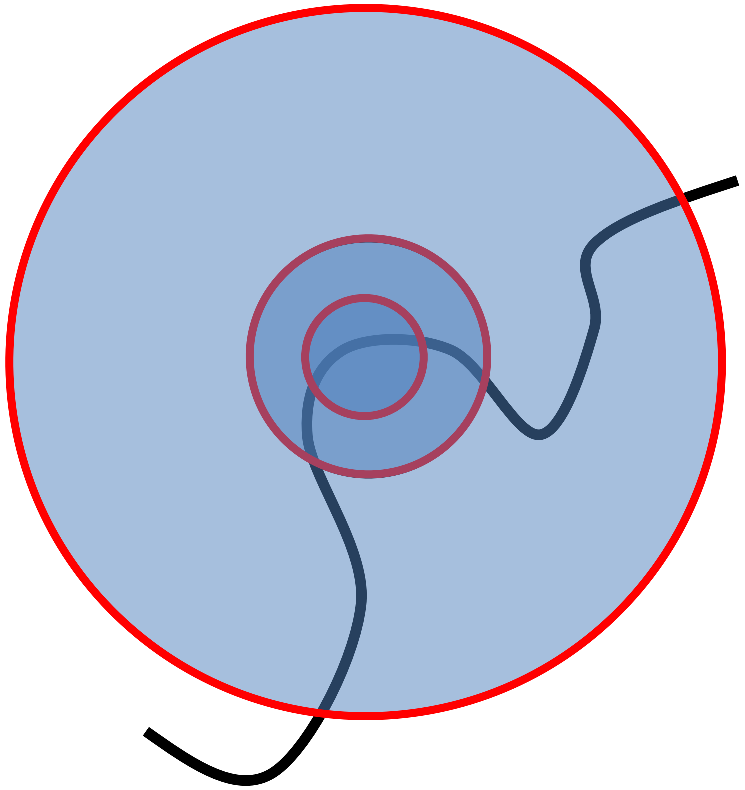
Q: How to find circle of right size?

- The problem: how do we choose corresponding circles ***independently*** in each image?



Scale invariant detection

Suppose you're looking for corners



Q: How to find circle of right size?

Key idea: find scale that gives local maximum of *f*

- in both position and scale
- One definition of *f*: the Harris operator

Solution

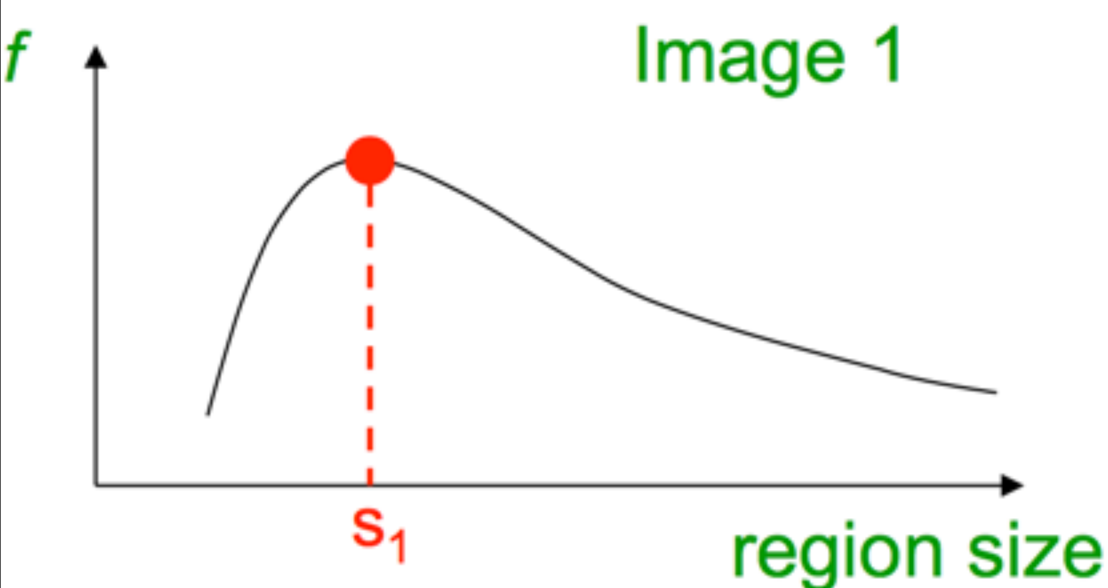
- Design a function on the region (circle) which is “scale invariant”
  - i.e., the same for corresponding regions, even if at different scales
  - E.g., average intensity. Same even for different sizes

- For a point in one image, consider it as a function of region size (circle radius)

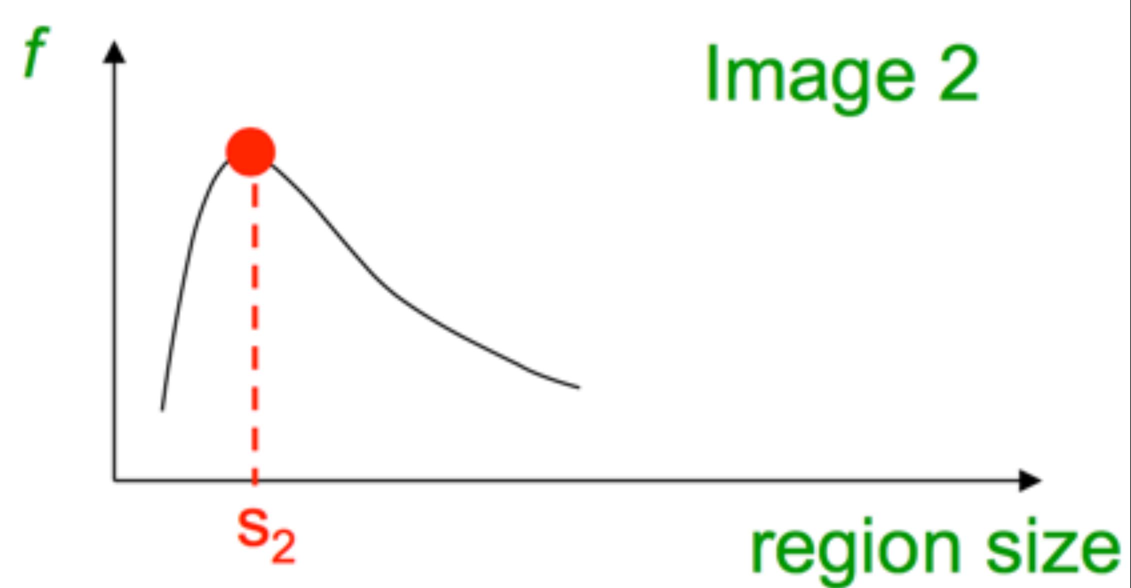
- Common approach:
  Take a local maximum of this function

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

> Important: this scale invariant region size is found in each image independently!

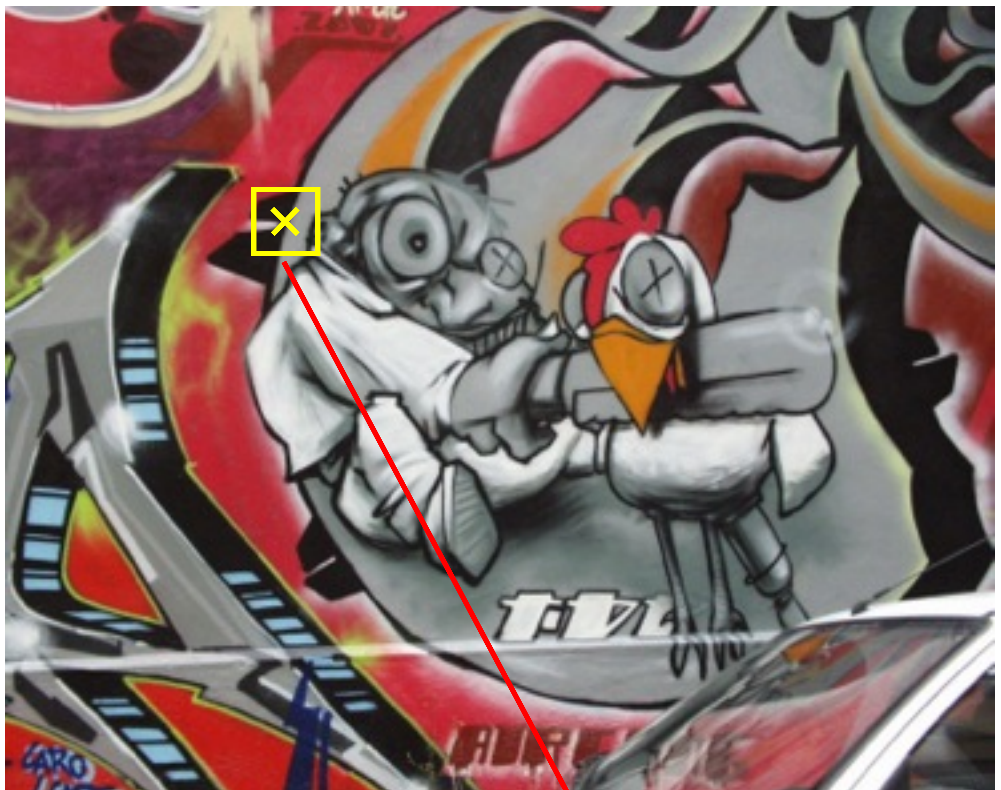


scale = 1/2



So far: can localize in x-y, but not scale

Automatic Scale Selection



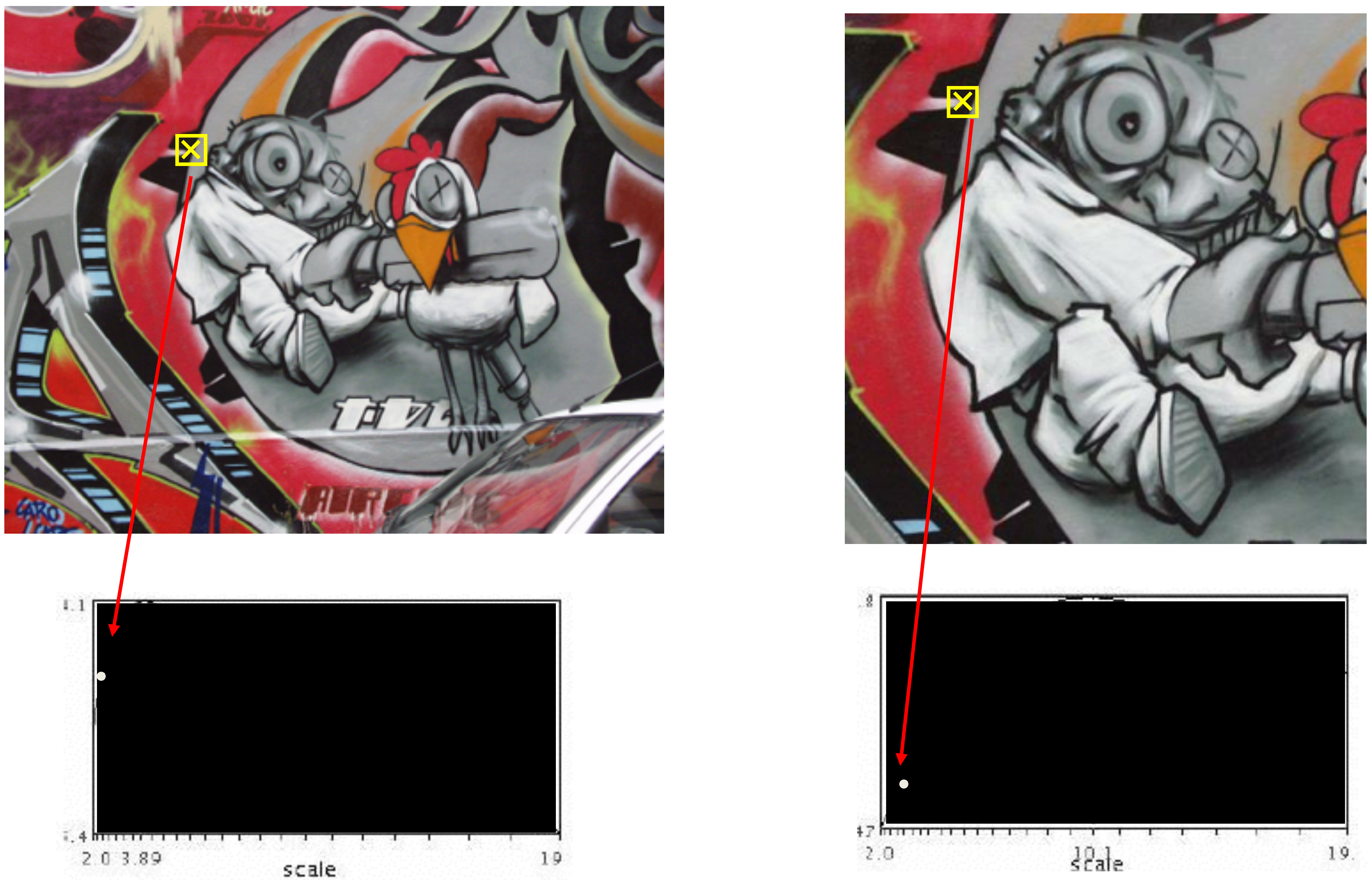
$$f(I_{i_1 \ldots i_m}(x,\sigma)) = f(I_{i_1 \ldots i_m}(x',\sigma'))$$

How to find corresponding patch sizes?

K. Grauman, B. Leibe

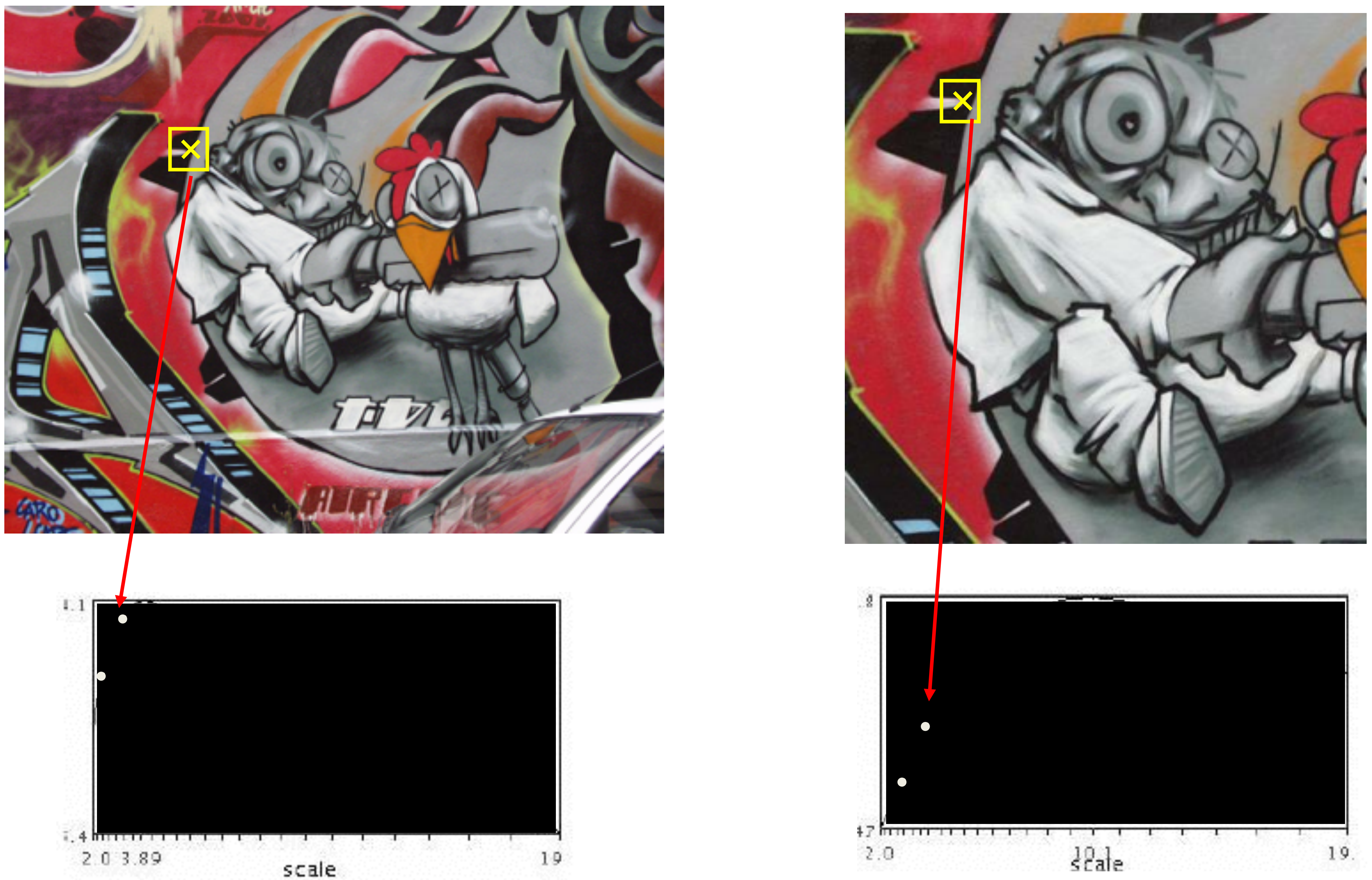
Automatic Scale Selection

- Function responses for increasing scale (scale signature)



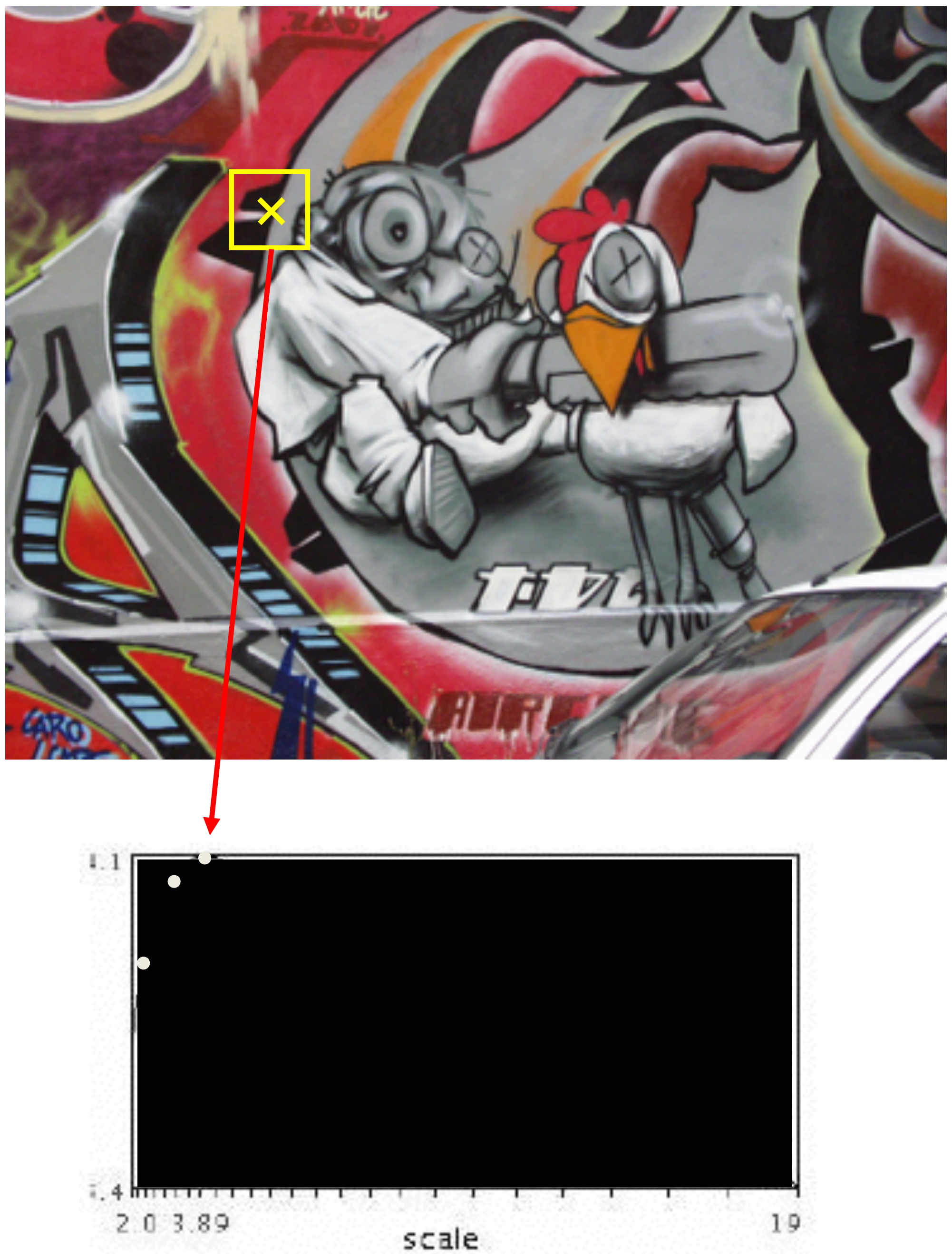
$f(I_{i_1 \ldots i_m}(x, \sigma))$ $\quad$ $f(I_{i_1 \ldots i_m}(x', \sigma))$

Automatic Scale Selection



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

Automatic Scale Selection

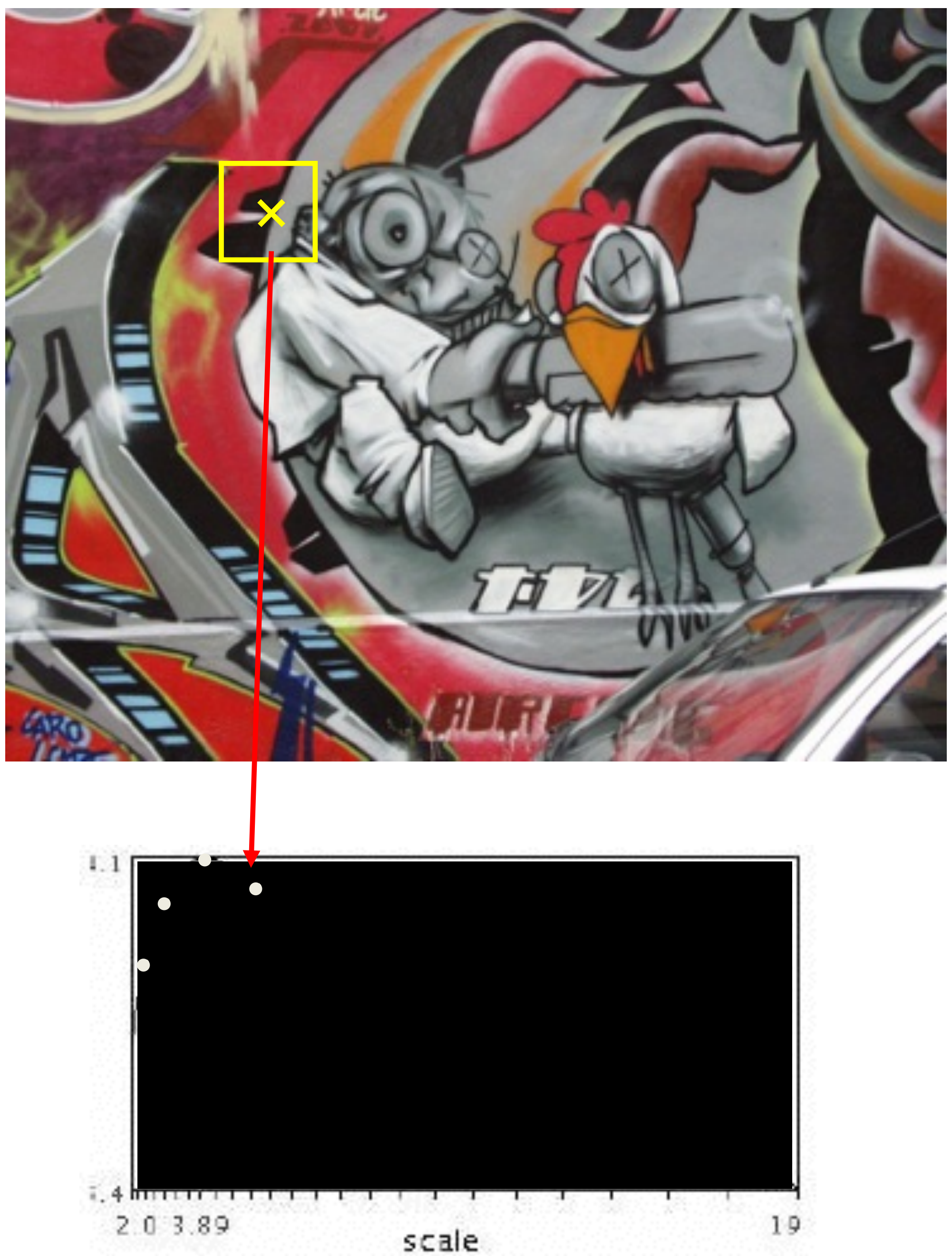


$f(I_{i_1 \ldots i_m}(x,\sigma))$

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

K. Grauman, B. Leibe

Automatic Scale Selection

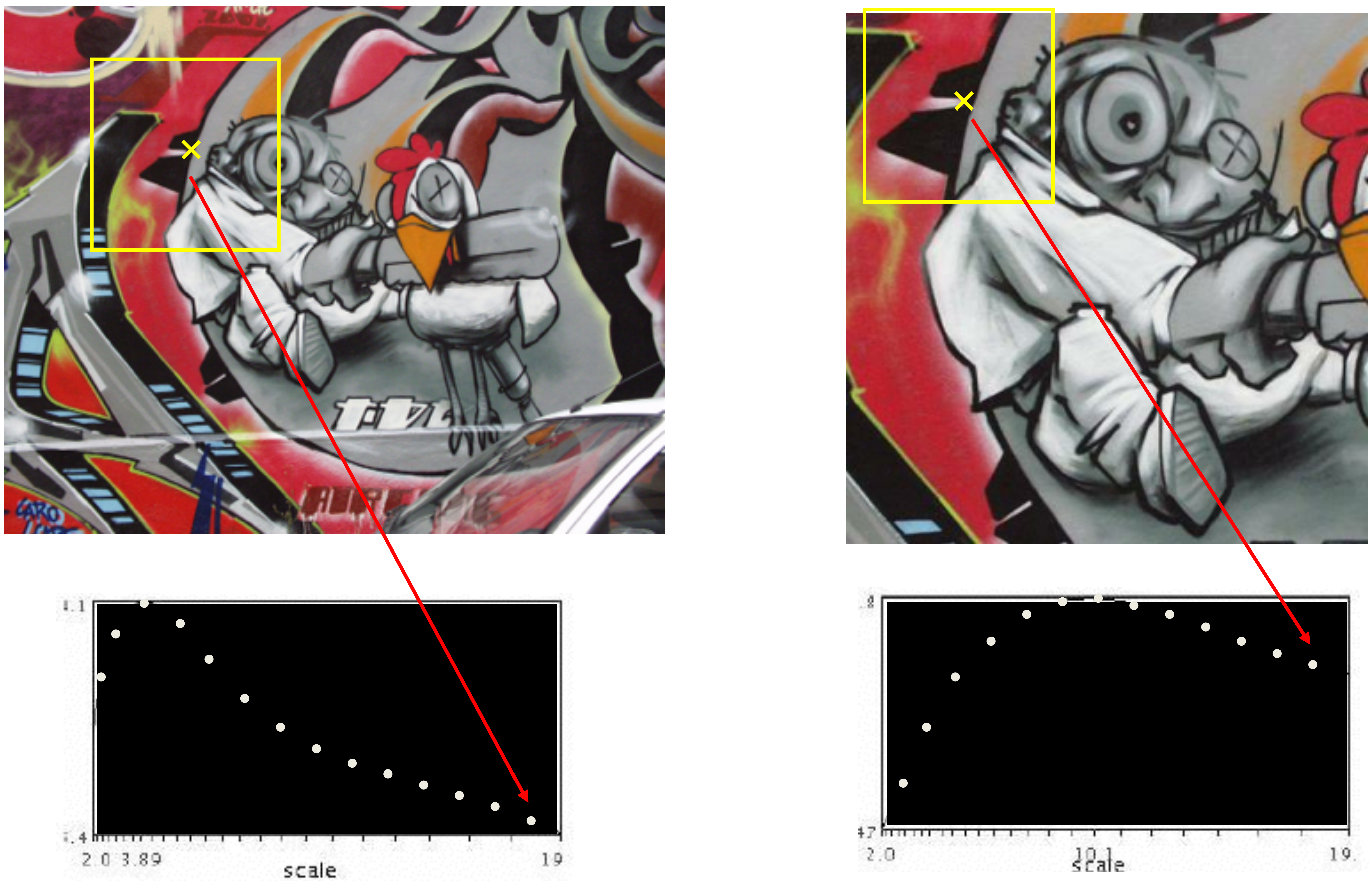


$f(I_{i_1 \ldots i_m}(x,\sigma))$

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

K. Grauman, B. Leibe

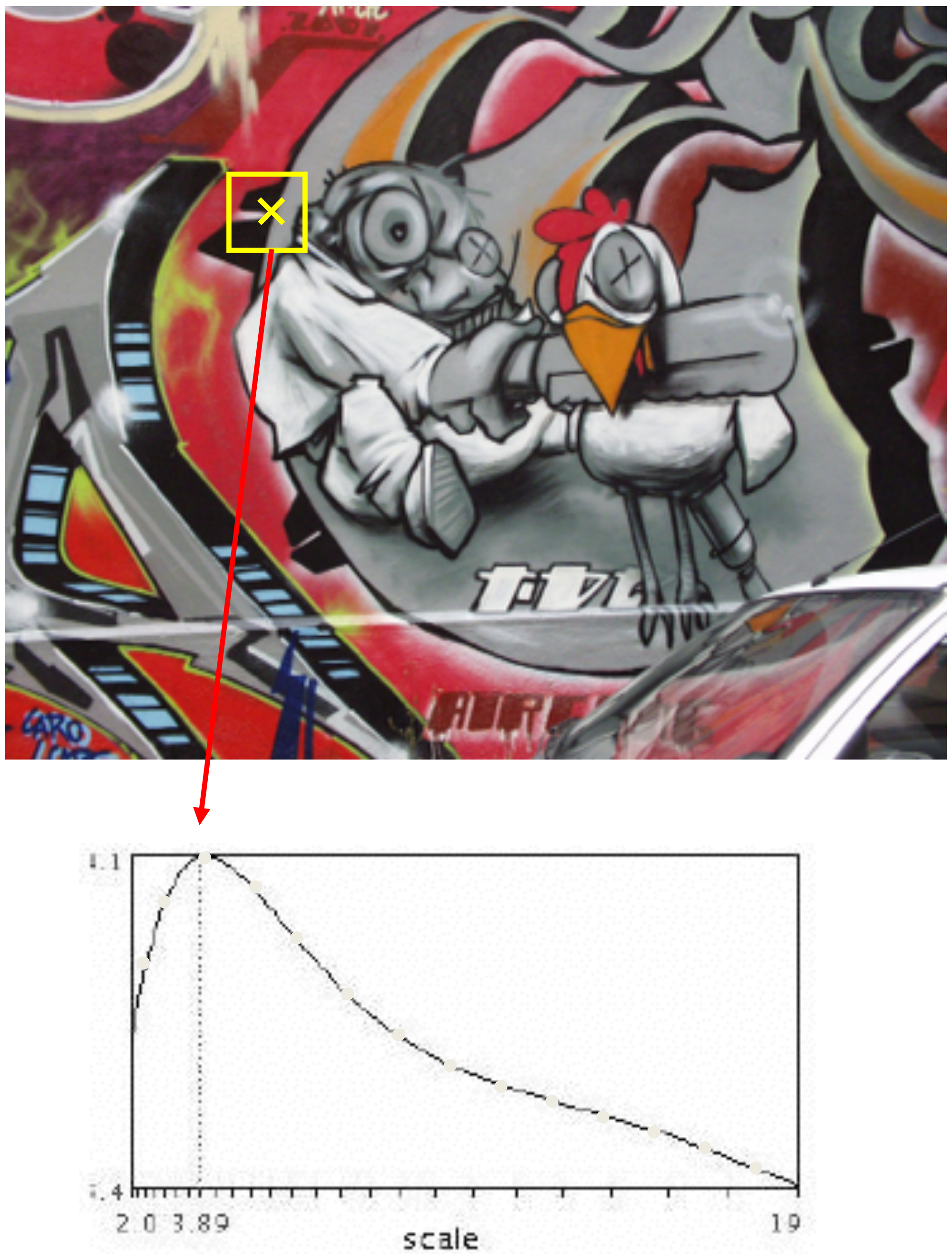
Automatic Scale Selection



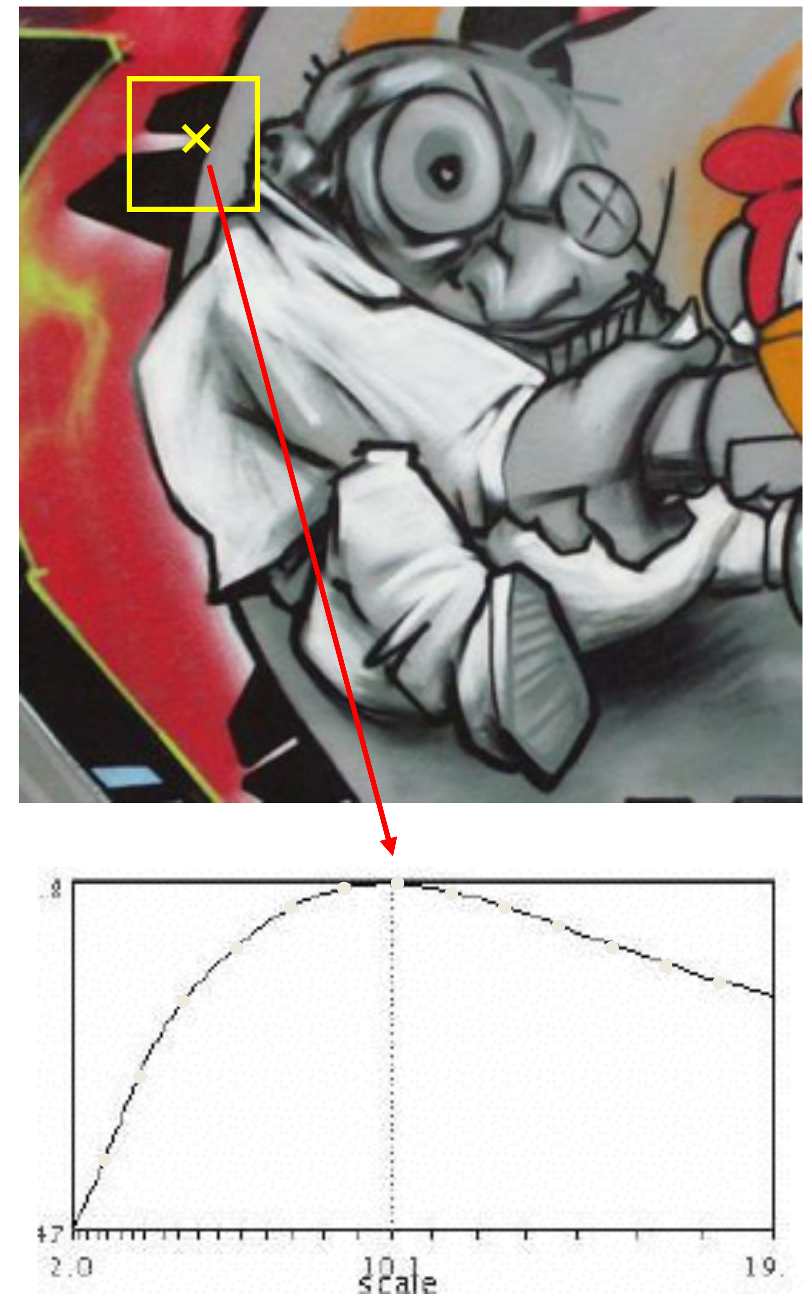
$f(I_{i_1 \dots i_m}(x, \sigma))$ $f(I_{i_1 \dots i_m}(x', \sigma))$

K. Grauman, B. Leibe

Automatic Scale Selection



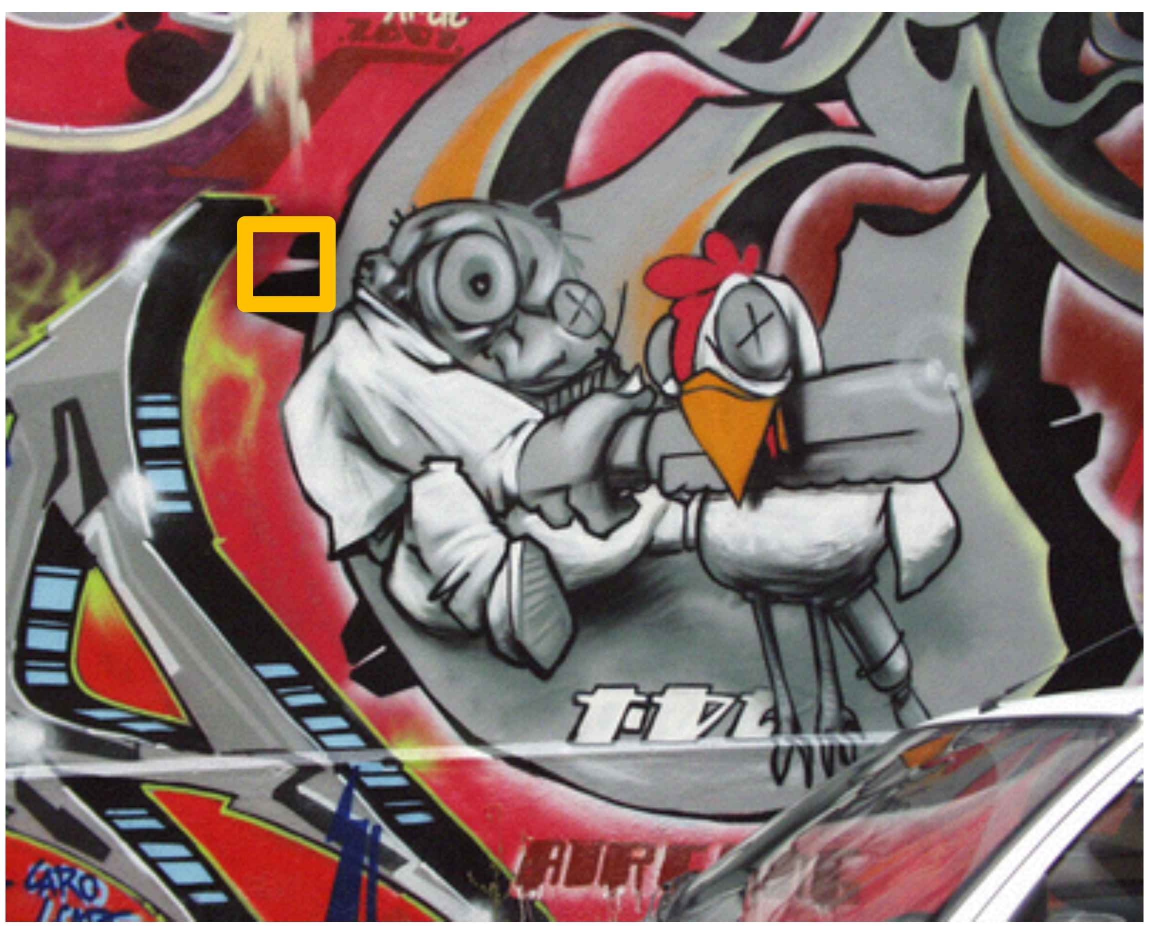
$f(I_{i_1 \ldots i_m}(x, \sigma))$

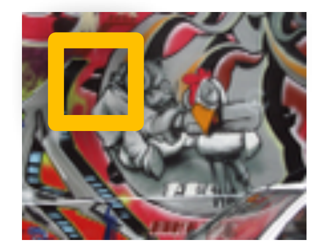


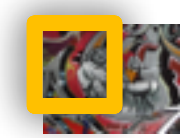
$f(I_{i_1 \ldots i_m}(x', \sigma'))$

Implementation

- Instead of computing *f* for larger and larger windows, we can implement using a fixed window size with a Gaussian pyramid





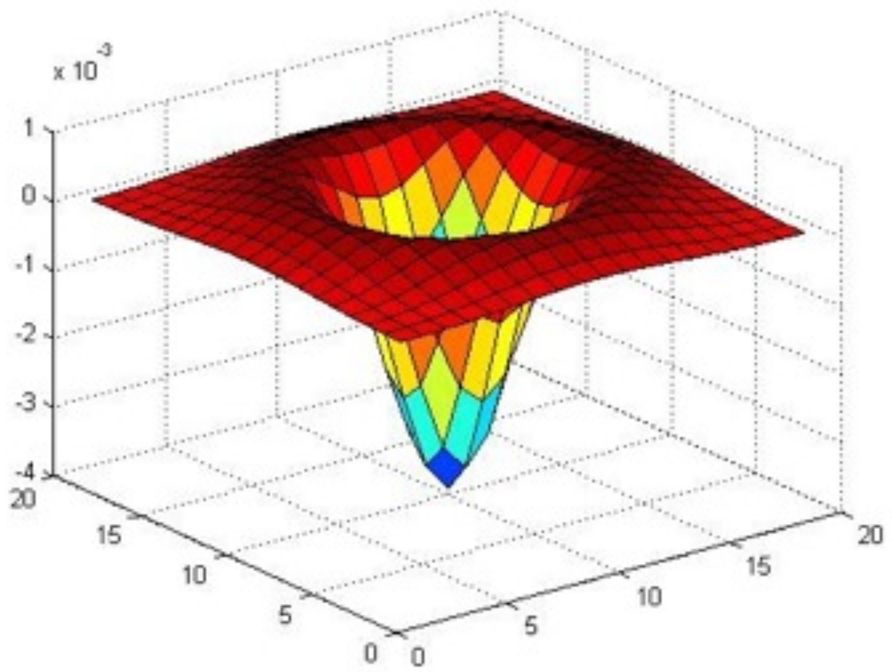


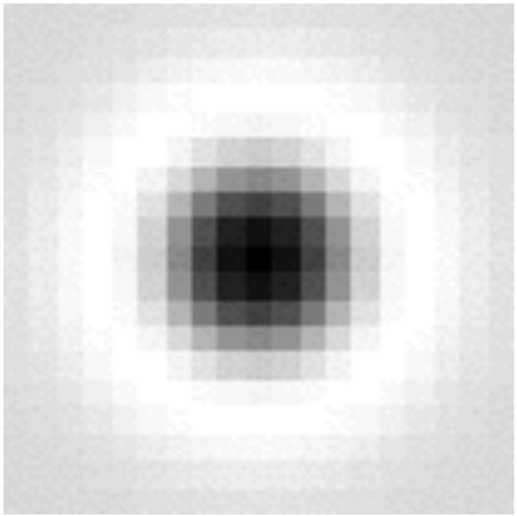
(sometimes need to create in-between levels, e.g. a ¾-size image)

Questions?

Another type of feature

- The *Laplacian of Gaussian* (*LoG*)





$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

(very similar to a Difference of Gaussians (DoG) – i.e. a Gaussian minus a slightly smaller Gaussian)

Slide source: Kavita Bala

Scale Invariant Detection

- Functions for determining scale $f = \text{Kernel} * \text{Image}$

Kernels:

$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

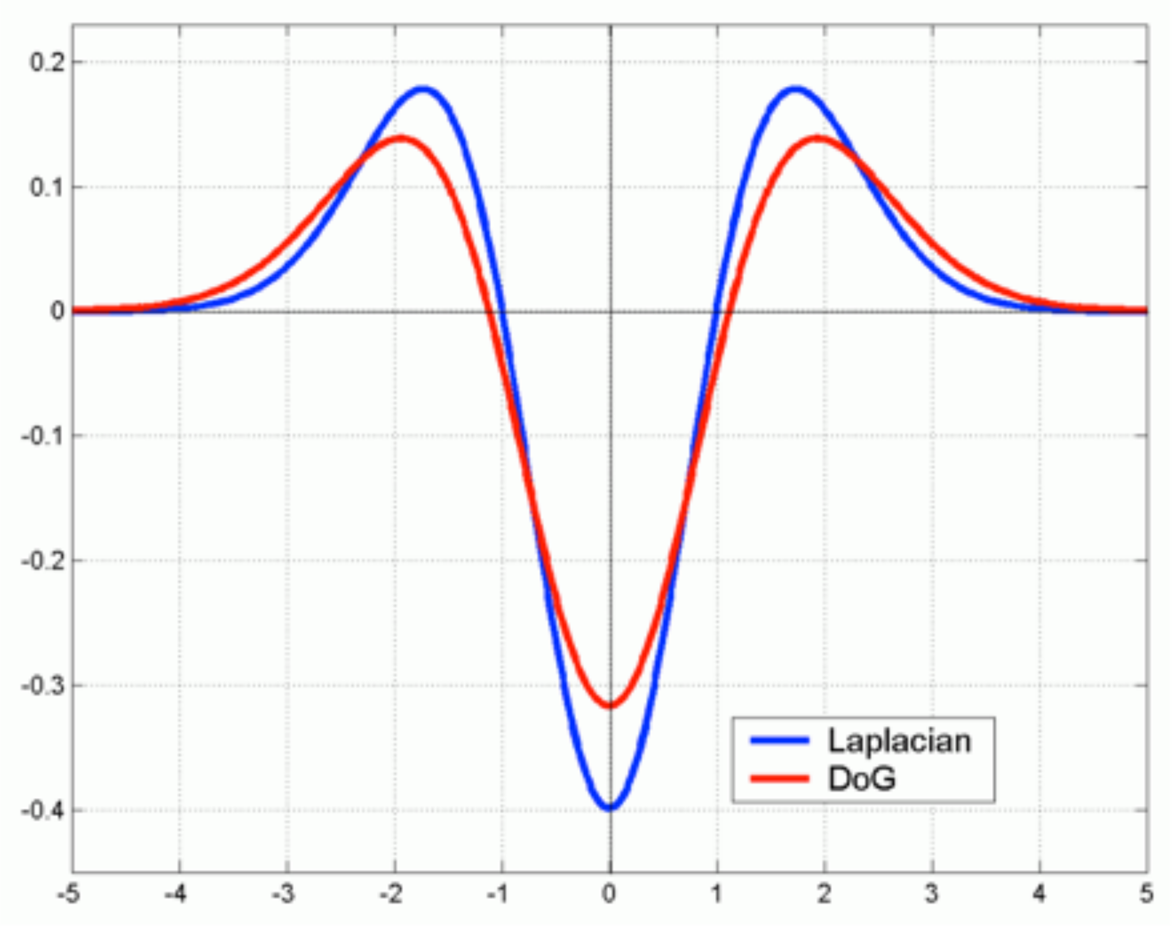
(Laplacian)

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

(Difference of Gaussians)

where Gaussian

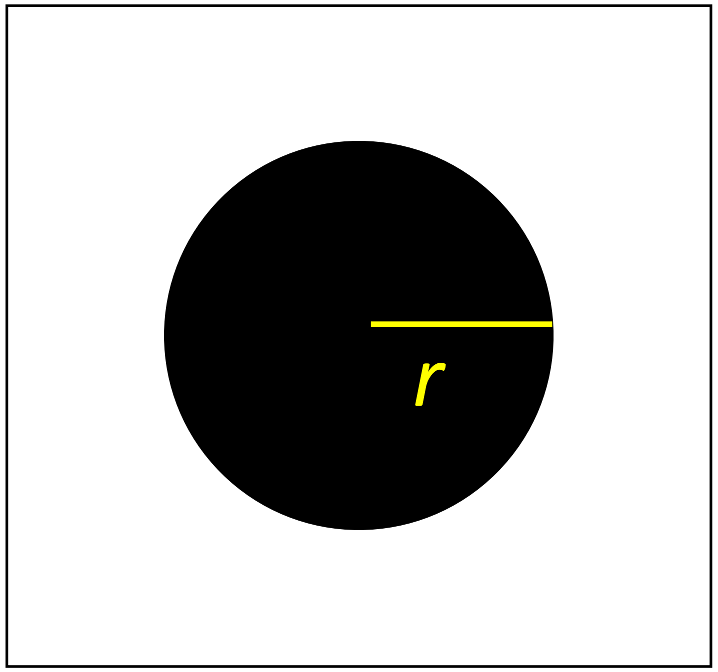
$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$



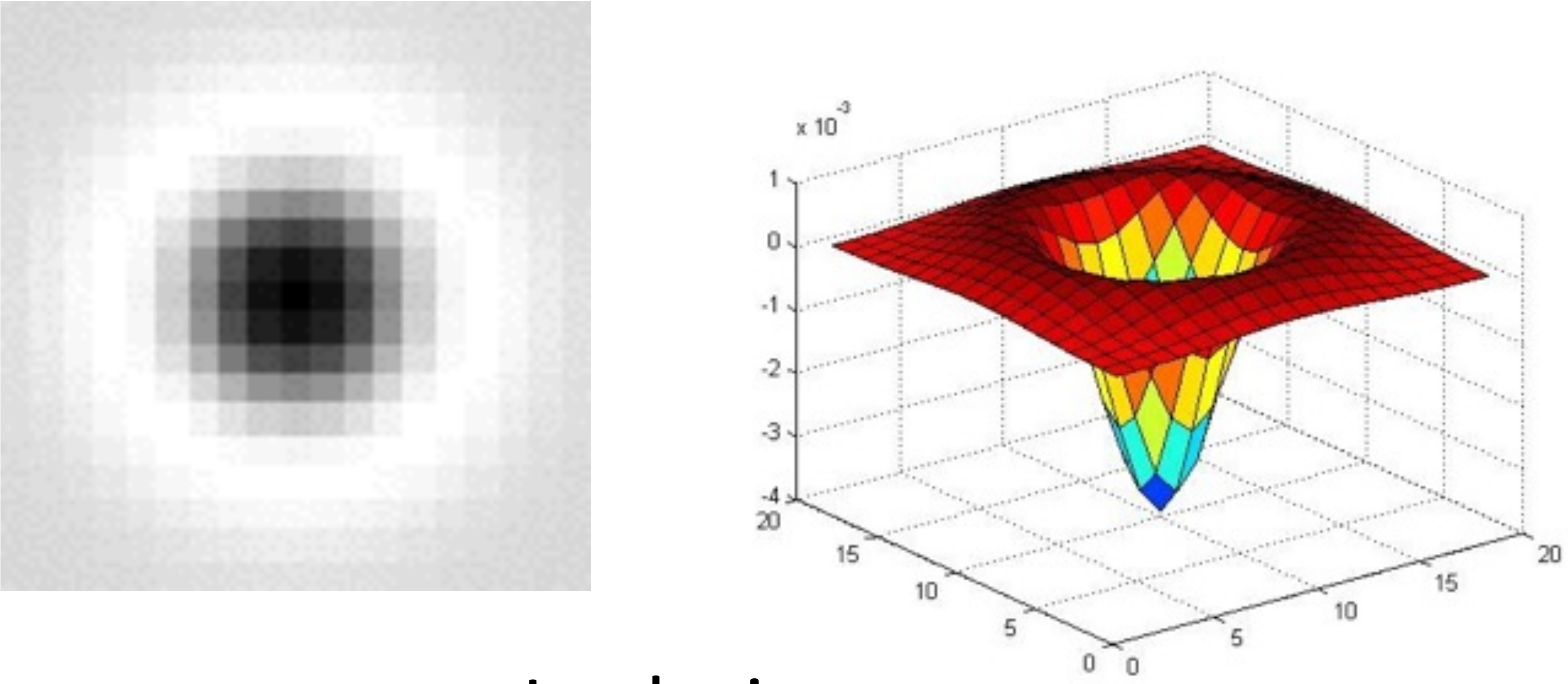
Note: both kernels are invariant to *scale* and *rotation*

Scale selection

- At what scale does the Laplacian achieve a maximum response for a binary circle of radius r?



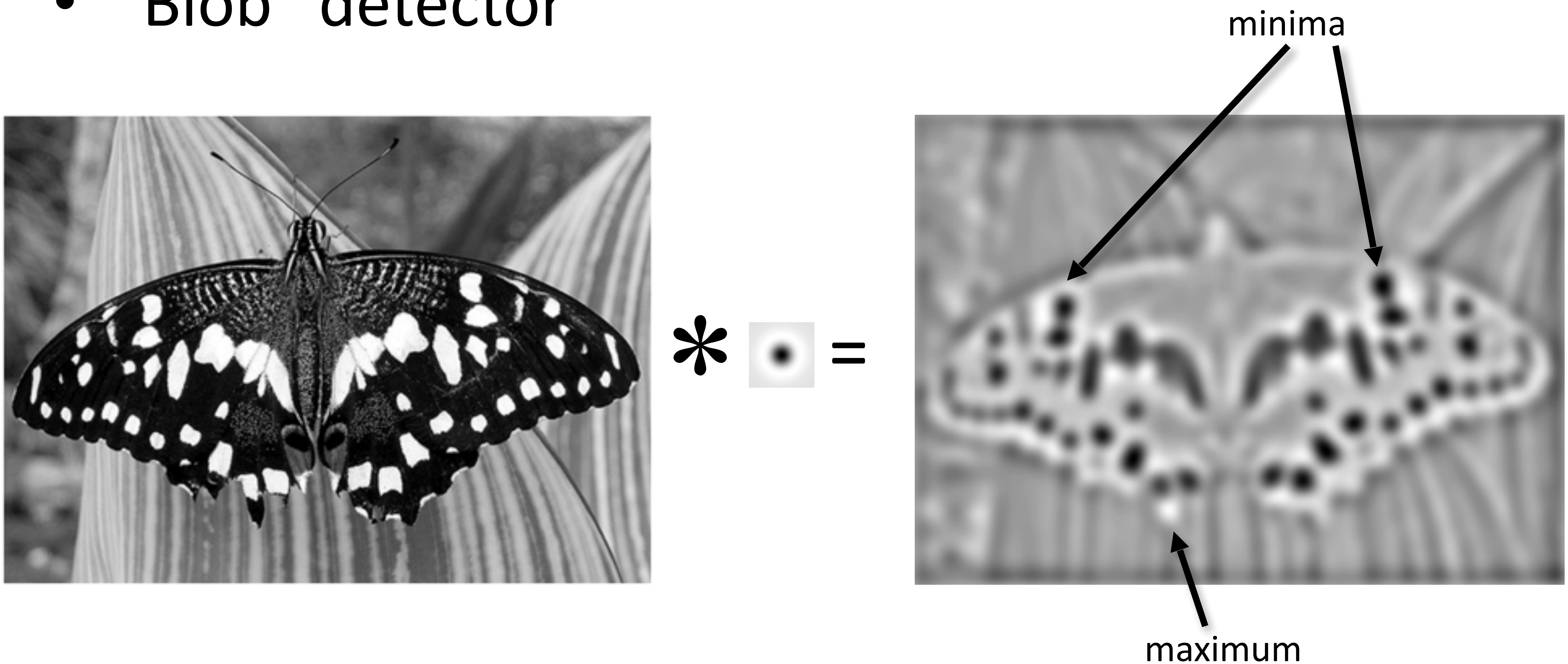
image



Laplacian

Laplacian of Gaussian

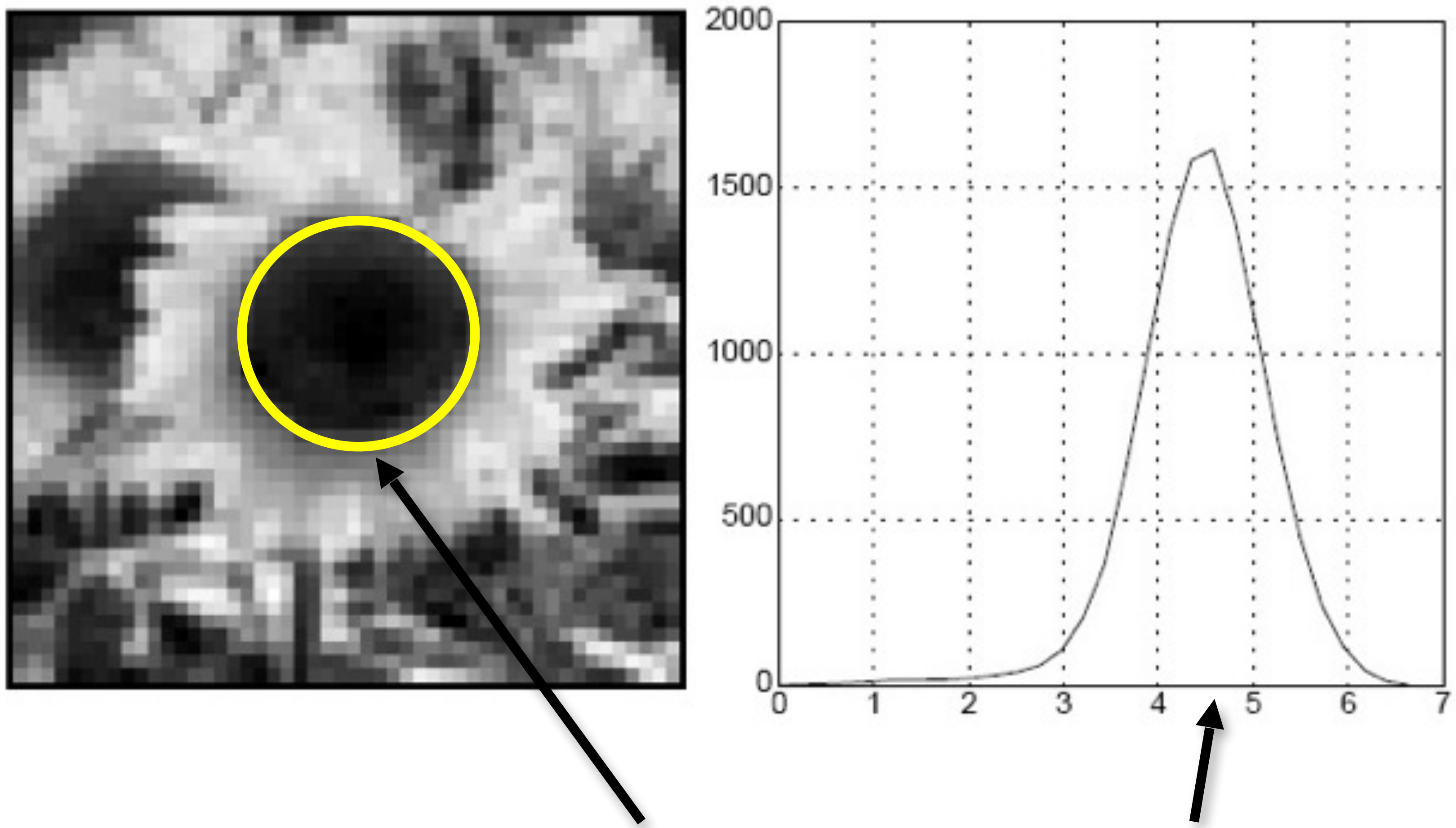
- "Blob" detector



- Find maxima *and minima* of LoG operator in space and scale

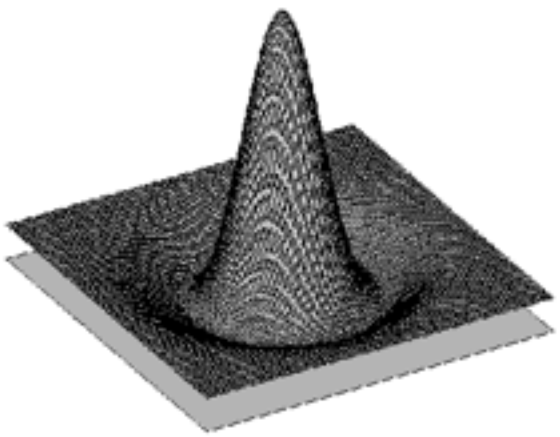
Characteristic scale

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



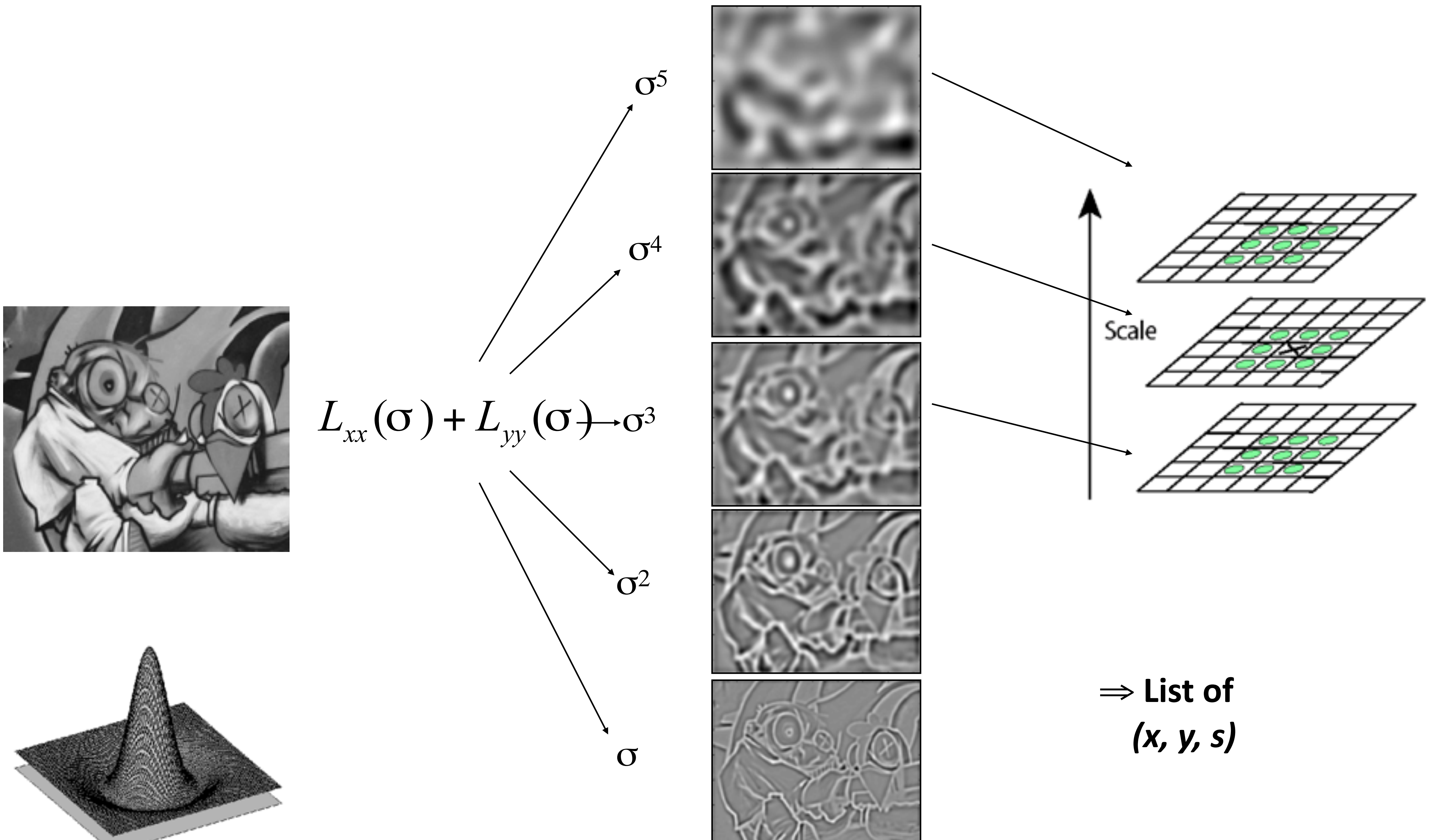
T. Lindeberg (1998). ["Feature detection with automatic scale selection."]() *International Journal of Computer Vision* **30** (2): pp 77--116.

Difference-of-Gaussian (DoG)



K. Grauman, B. Leibe

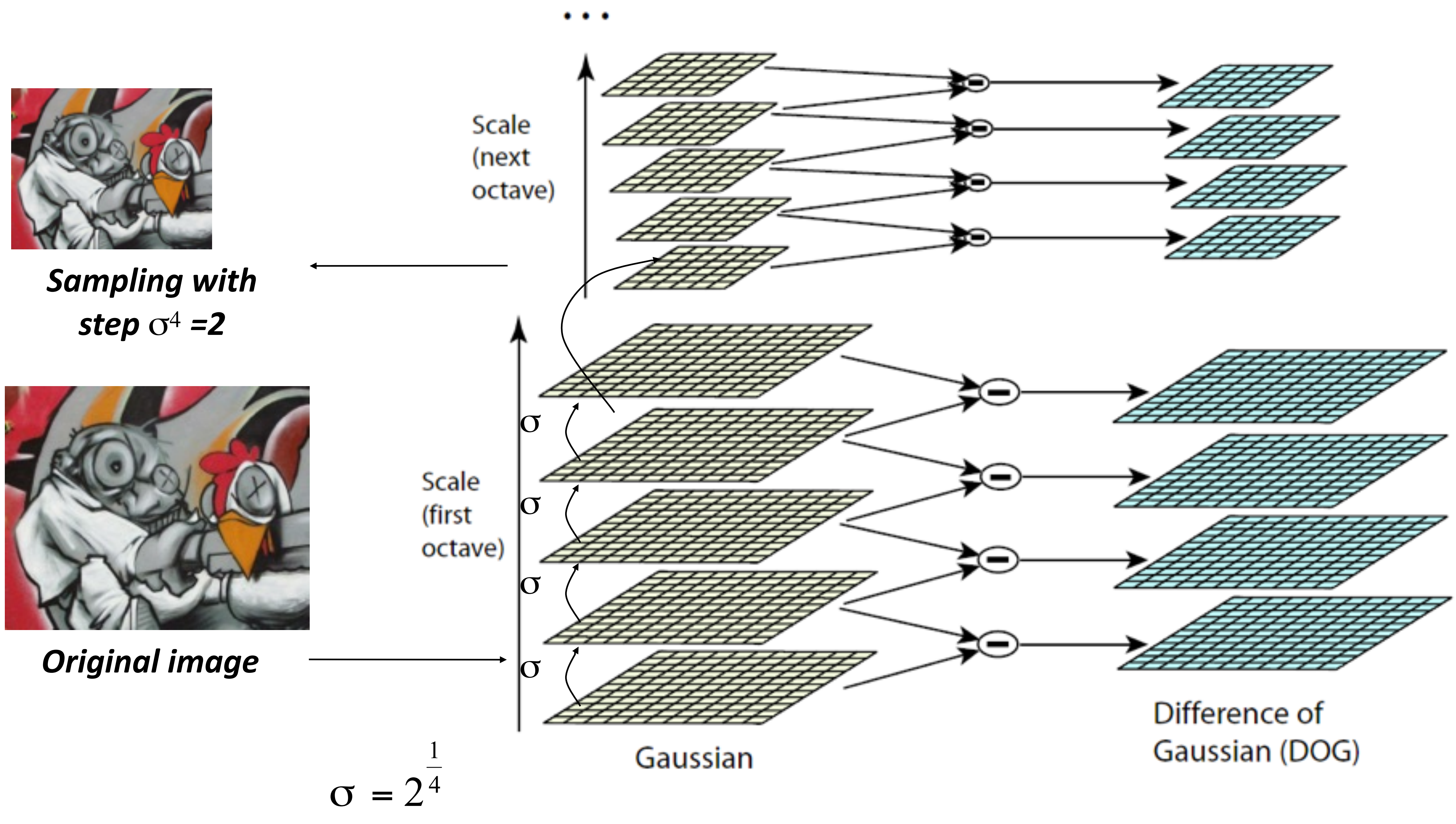
Find local maxima in position-scale space of Difference-of-Gaussian



K. Grauman, B. Leibe

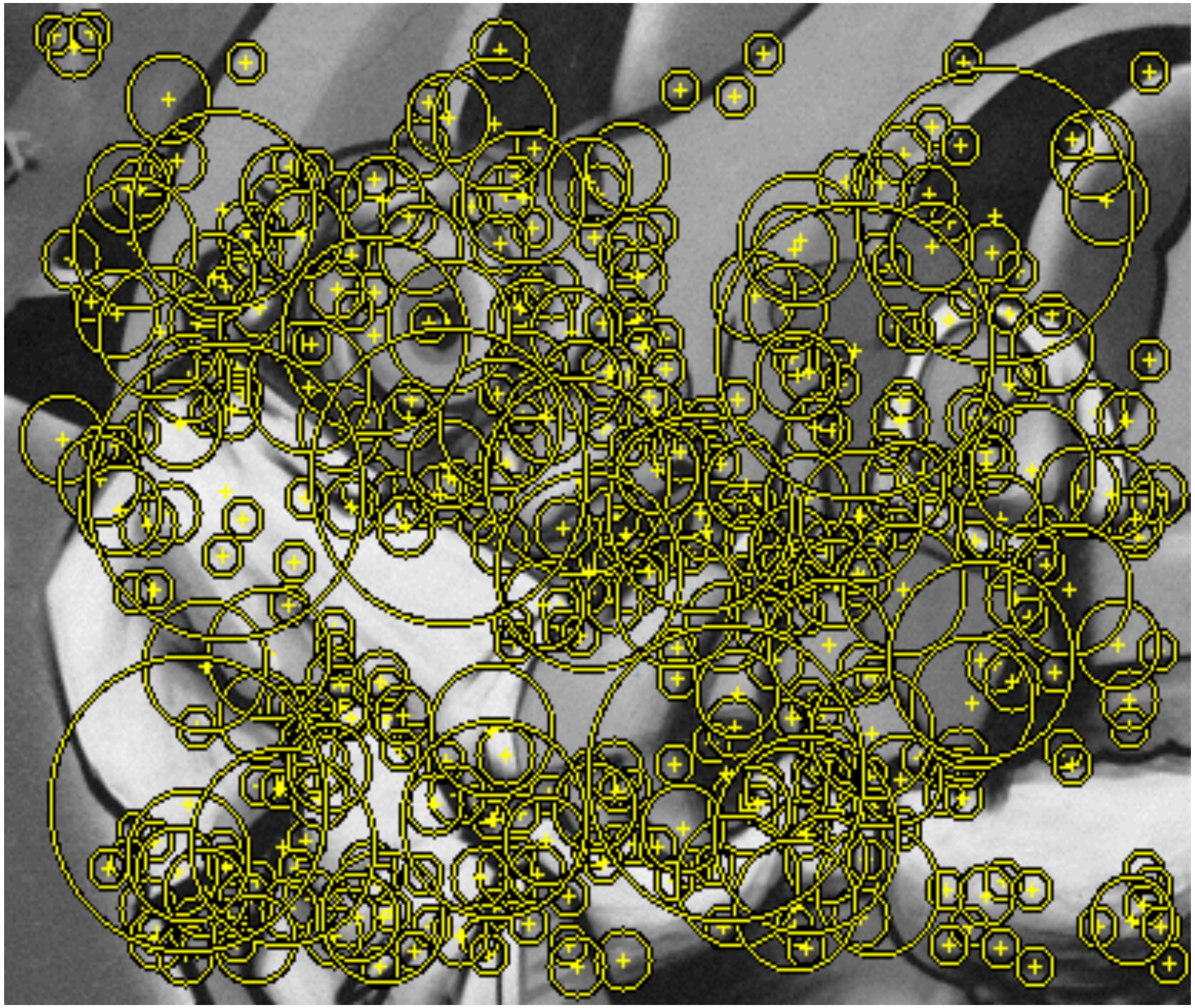
DoG – Efficient Computation

- Computation in Gaussian scale pyramid

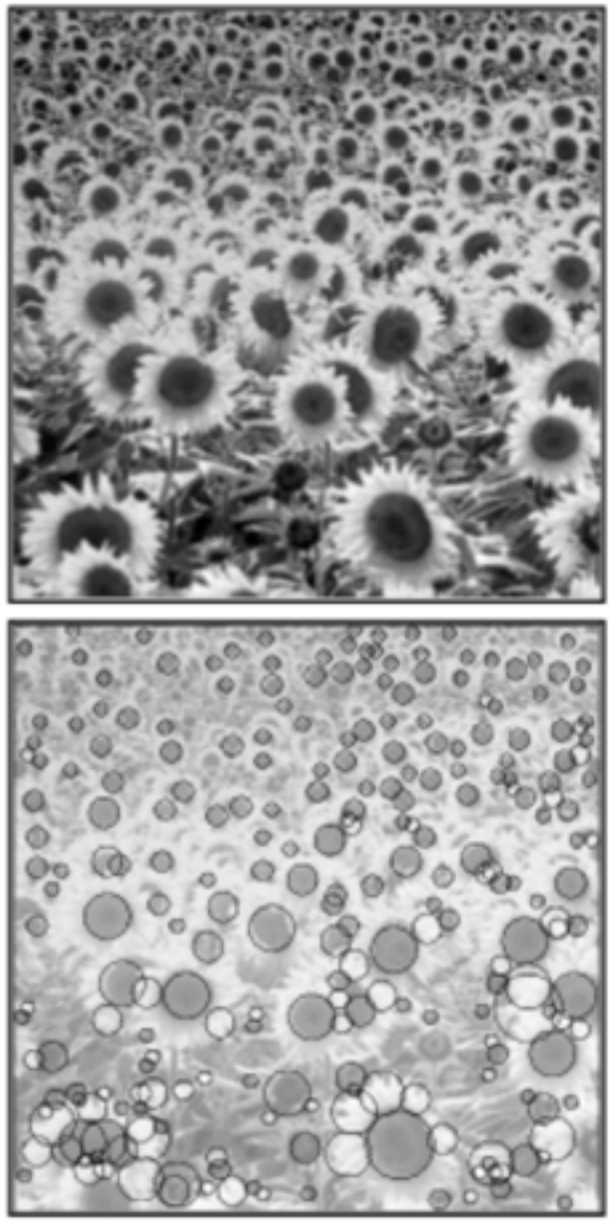


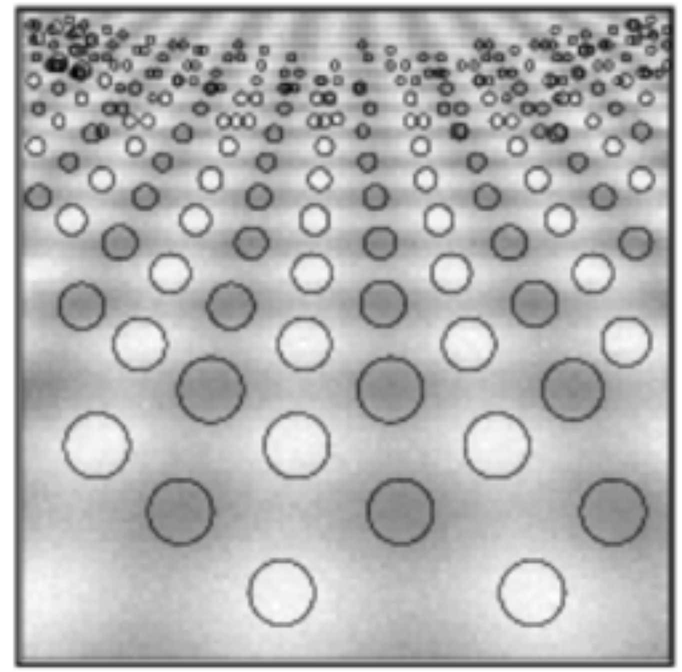
K. Grauman, B. Leibe

Results: Difference-of-Gaussian



K. Grauman, B. Leibe





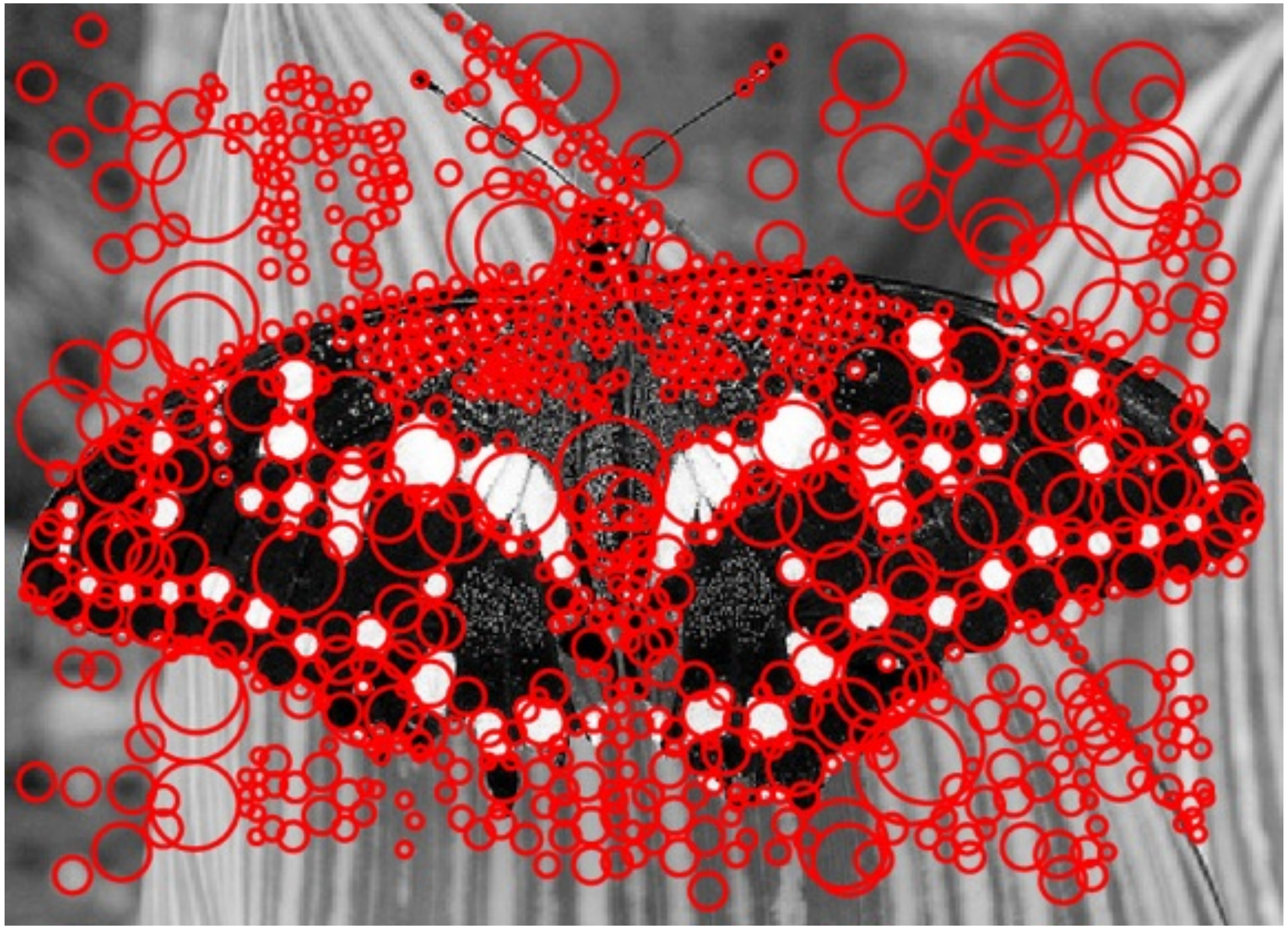
But first, we have to talk about detecting blobs at one scale...

Scale-space blob detector: Example

Scale-space blob detector: Example

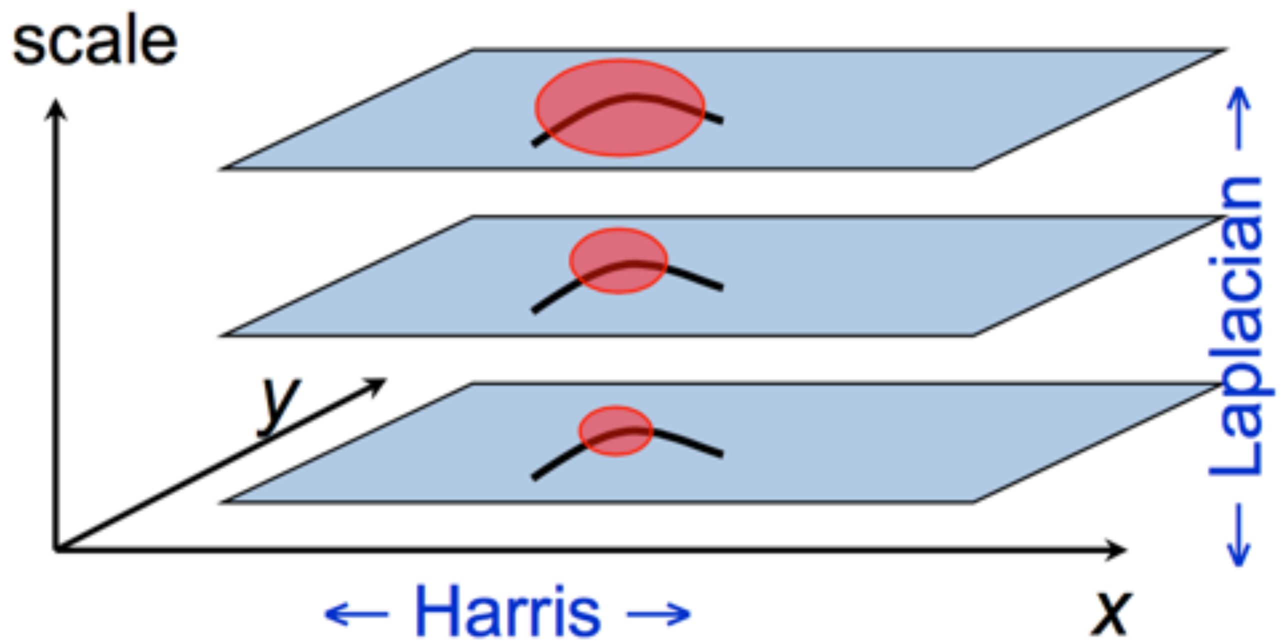
sigma = 11.9912

Scale-space blob detector: Example

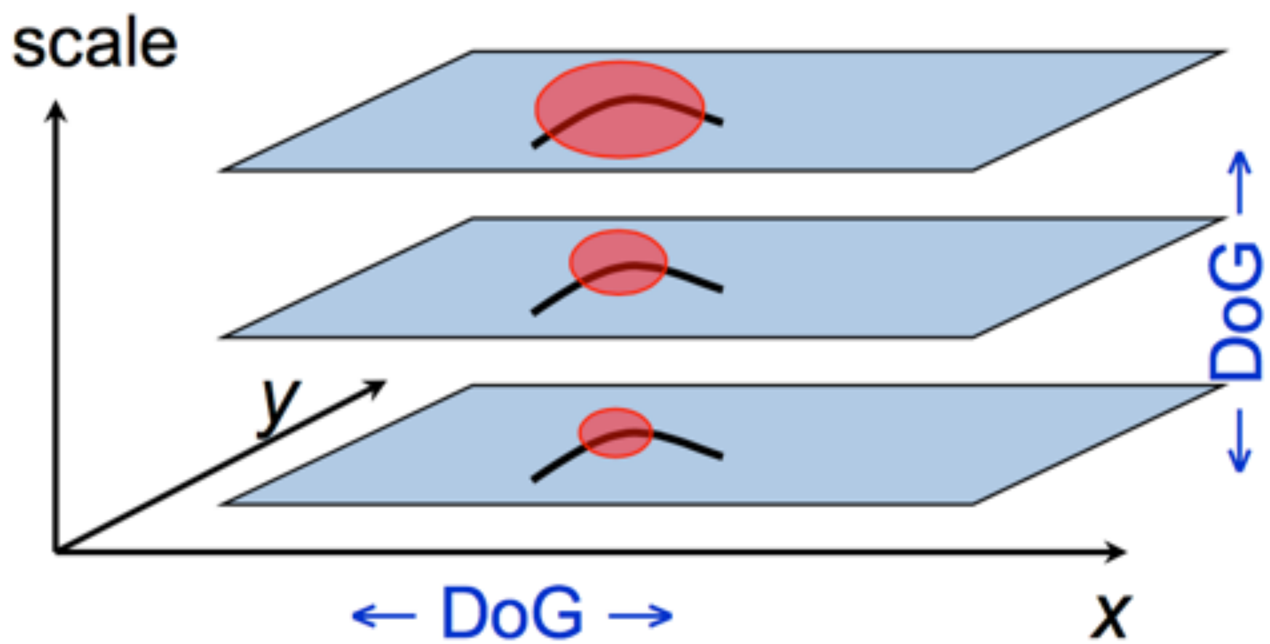


Scale Invariant Detectors

- Harris-Laplacian[1]
  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



[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001
[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". IJCV 2004

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:

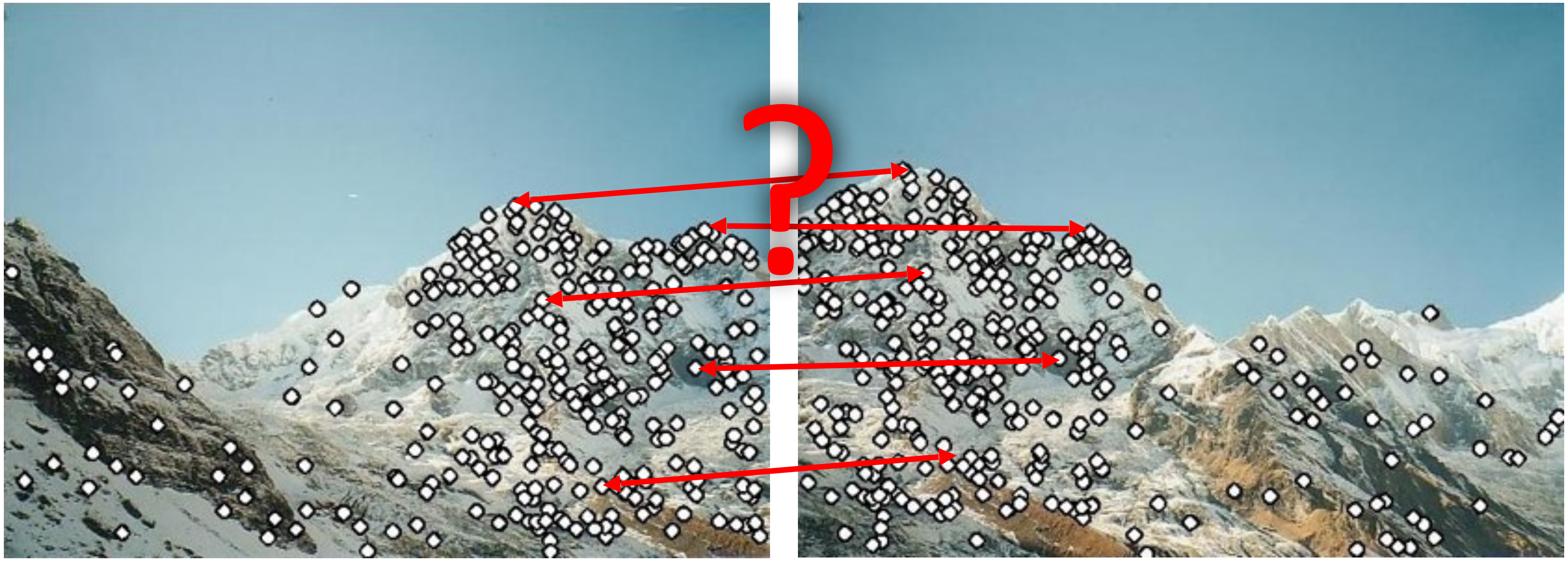
1. Harris-Laplacian [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. SIFT [Lowe]: maximize Difference of Gaussians over scale and space

Questions?

Feature descriptors

We know how to detect good points

Next question: **How to match them?**



Answer: Come up with a *descriptor* for each point, find similar descriptors between the two images

Invariance vs. discriminability

- Invariance:
  - Descriptor shouldn't change even if image is transformed

- Discriminability:
  - Descriptor should be highly unique for each point

Invariance

- Most feature descriptors are designed to be invariant to
  - Translation, 2D rotation, scale

- They can usually also handle
  - Limited 3D rotations (SIFT works up to about 60 degrees)
  - Limited affine transformations (some are fully affine invariant)
  - Limited illumination/contrast changes

How to achieve invariance

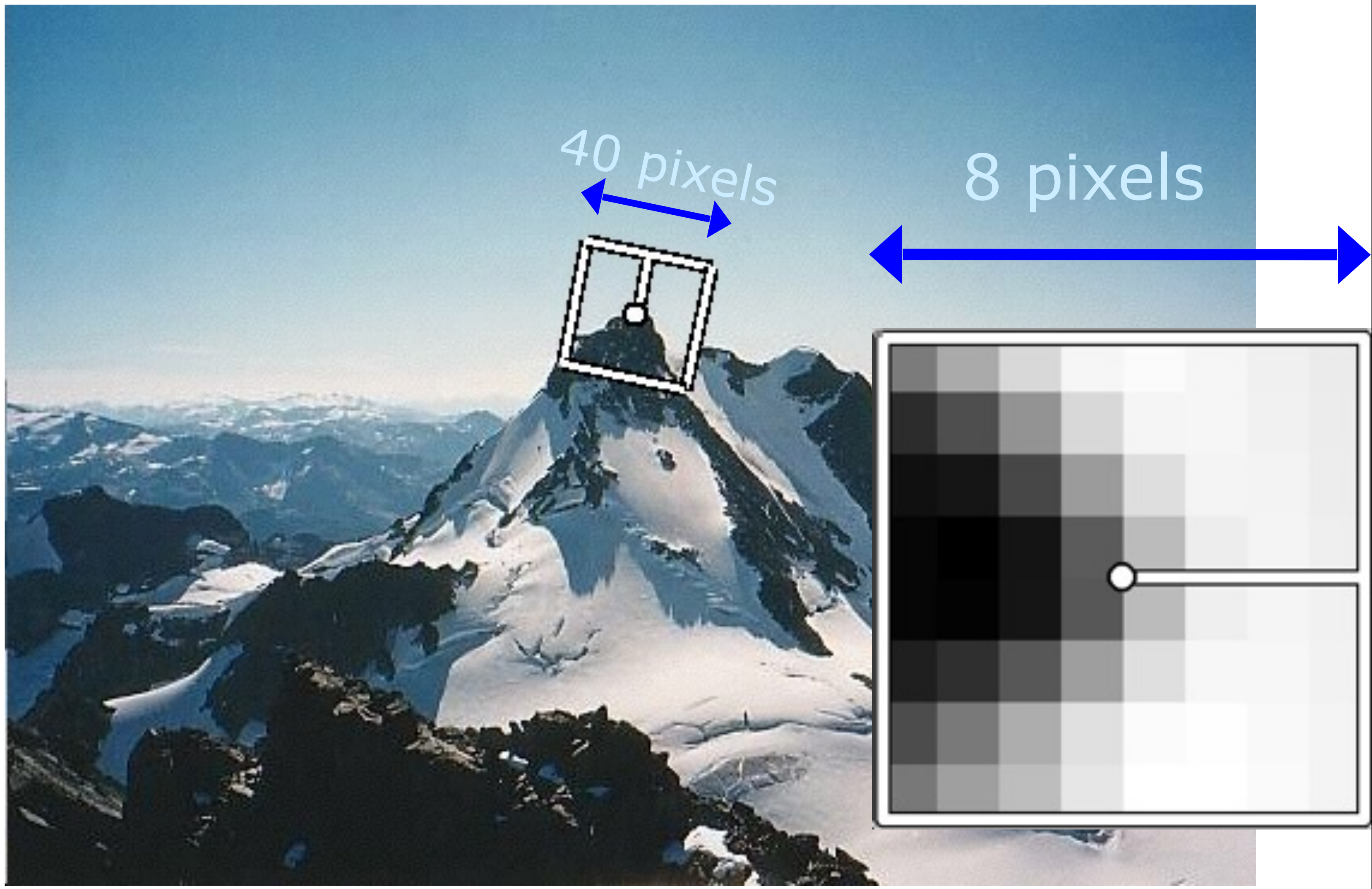
Need both of the following:

1. Make sure your detector is invariant
2. Design an invariant feature descriptor
   - Simplest descriptor: a single 0
     - What's this invariant to?
   - Next simplest descriptor: a square window of pixels
     - What's this invariant to?
   - Let's look at some better approaches…

Multiscale Oriented PatcheS descriptor

Take 40x40 square window around detected feature

- Scale to 1/5 size (using prefiltering)
- Rotate to horizontal
- Sample 8x8 square window centered at feature
- Intensity normalize the window by subtracting the mean, dividing by the standard deviation in the window



Adapted from slide by Matthew Brown

Detections at multiple scales

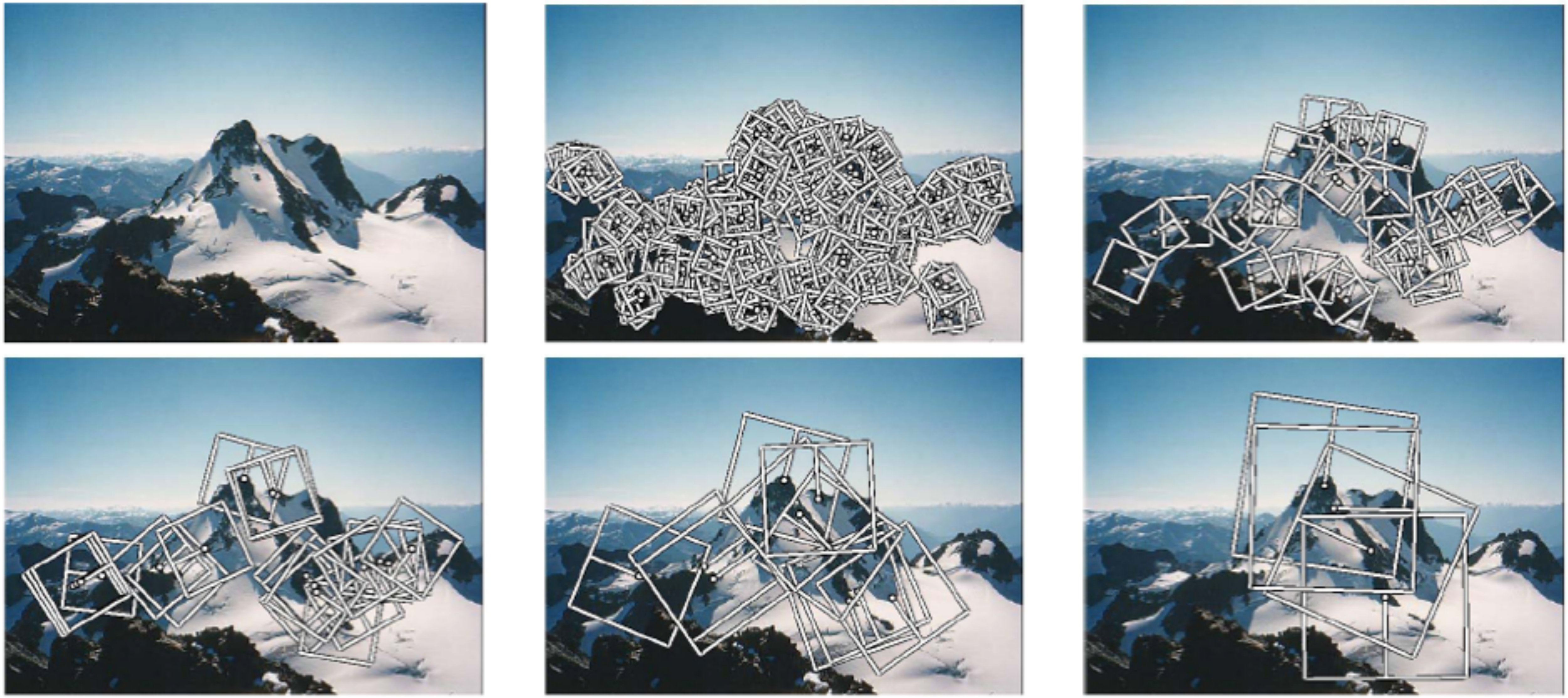


Figure 1. Multi-scale Oriented Patches (MOPS) extracted at five pyramid levels from one of the Matier images. The boxes show the feature orientation and the region from which the descriptor vector is sampled.