Lecture 6. Feature Descriptors

Scale invariant keypoint detection

Juan Carlos Niebles and Jiajun Wu

CS131 Computer Vision: Foundations and Applications

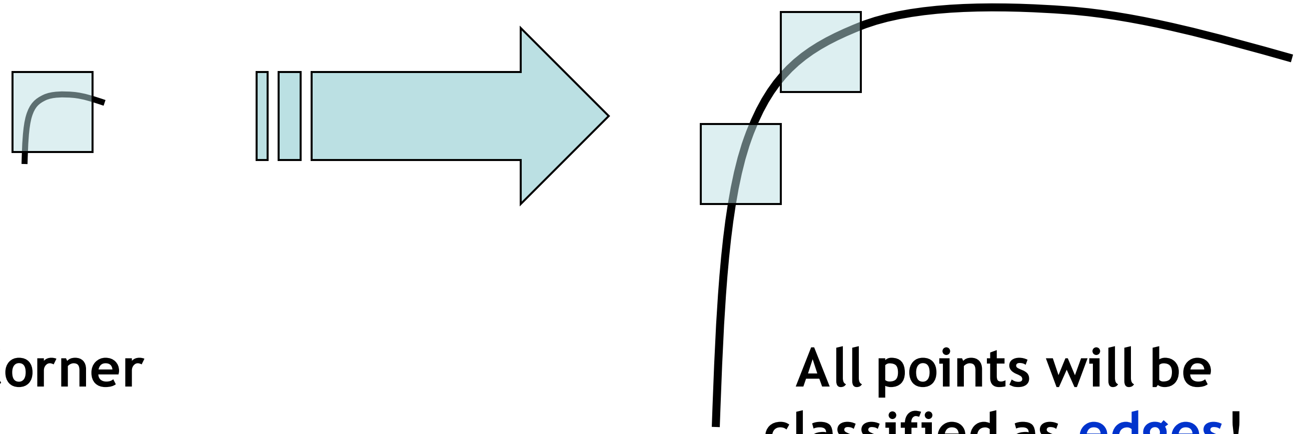
CS 131 Roadmap

Pixels	Segments	Images	Videos	Web
Convolutions Edges Features	Resizing Segmentation Clustering	Recognition Detection Machine learning	Motion Tracking	Neural networks Convolutional neural networks

What will we learn today?

- Scale invariant keypoint detection
  - Automatic scale selection
  - Harris-Laplace detector
  - Difference-of-Gaussian (DoG) detector

The Harris corner detector is not scale invariant



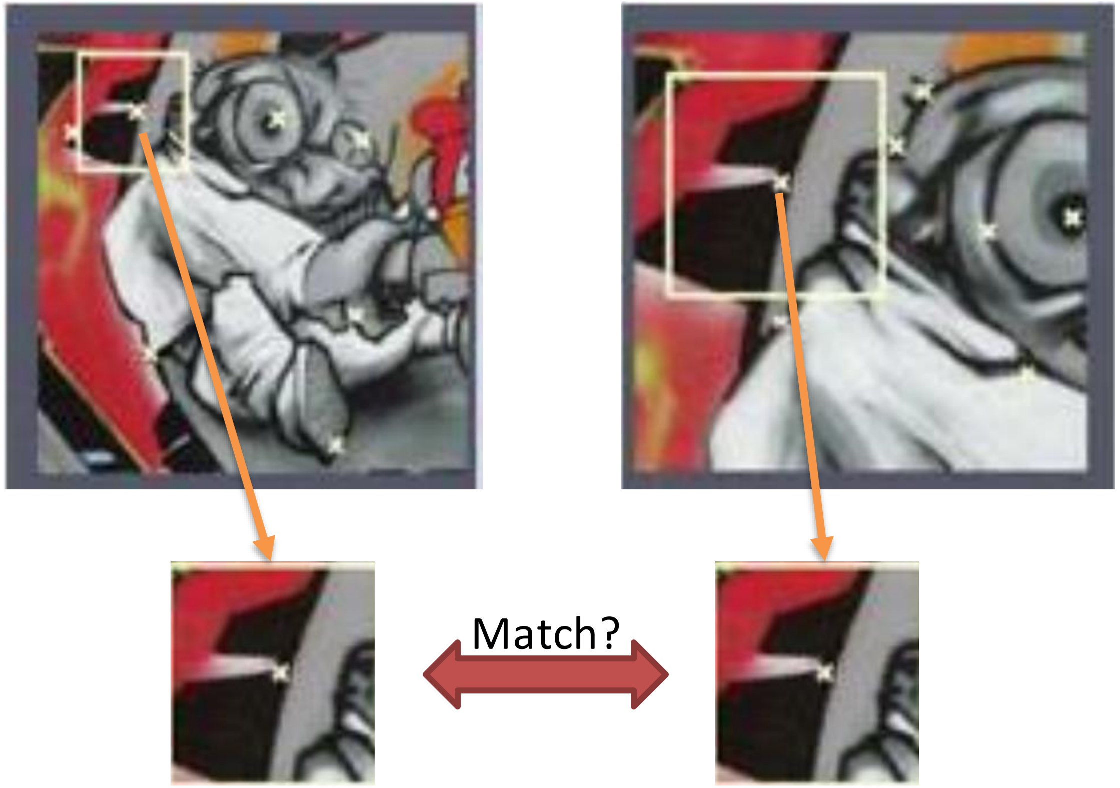
Corner

All points will be classified as edges!

Not invariant to image scale!

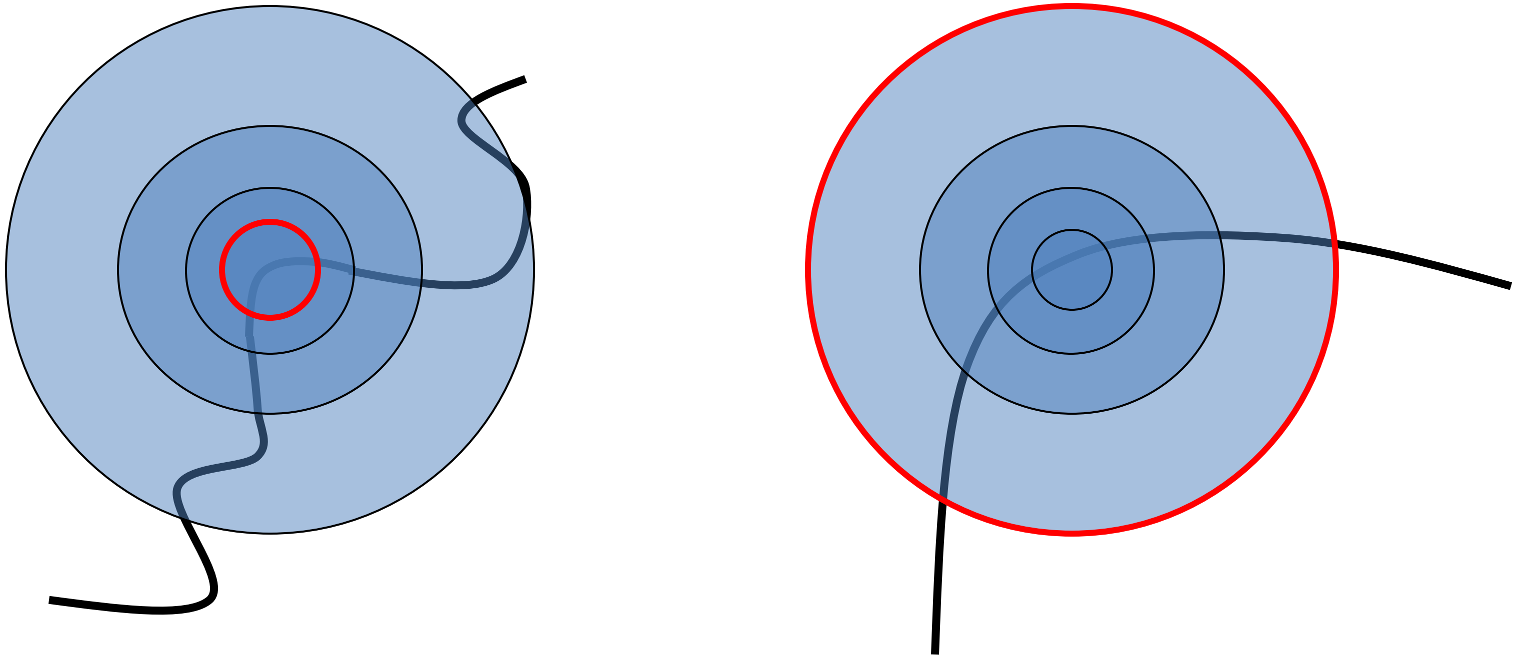
Does scale matter?

- For feature matching, it is important to estimate the size of the neighborhood that can lead to best matching between images of different scale.



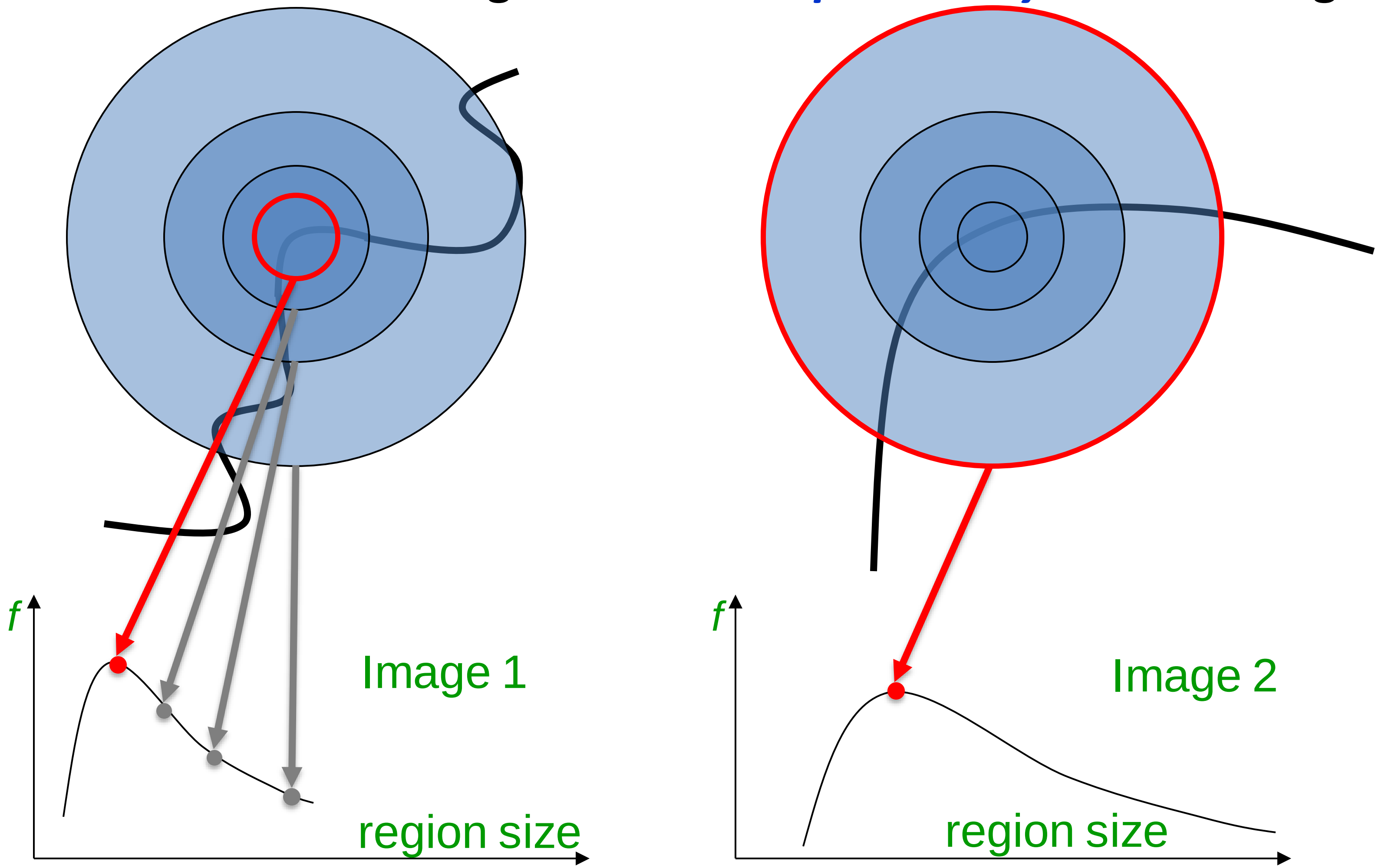
Scale Invariant Detection

- Consider regions (e.g. circles) of different sizes around a point
- What region size do we choose, so that the regions look the same in both images?



Scale Invariant Detection

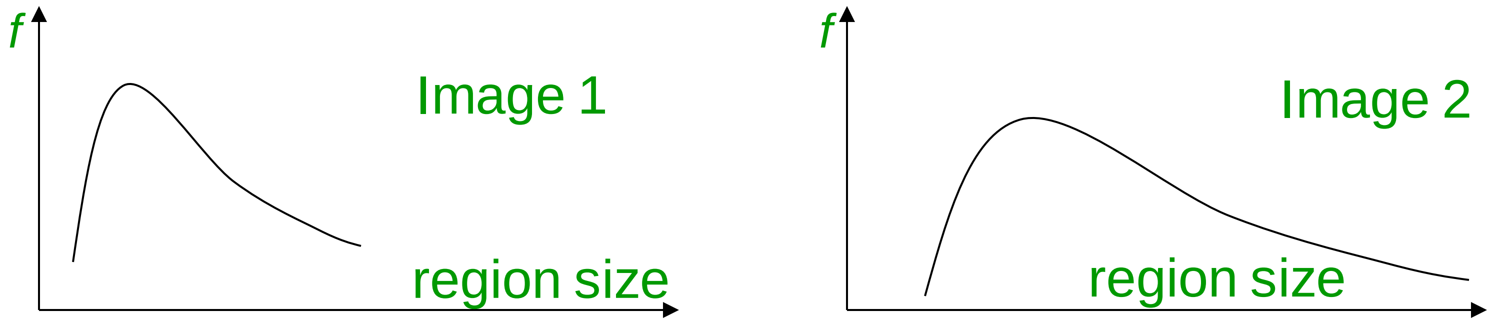
- Problem: How do we choose region sizes ***independently*** in each image?



Scale Invariant Detection

- Solution:
  - Design a function on the region (circle), which is “scale invariant”: has the same value 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.

  - Given a point in one image, we can think of it as a function of region size (circle radius)

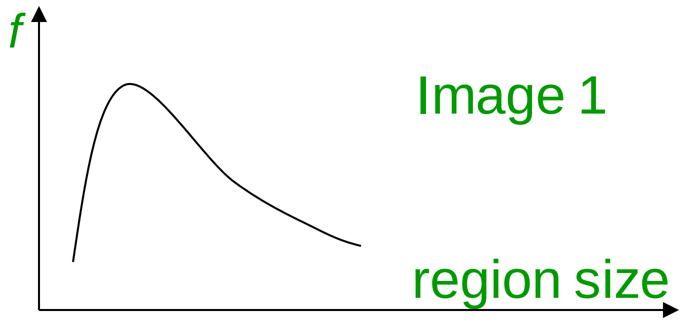


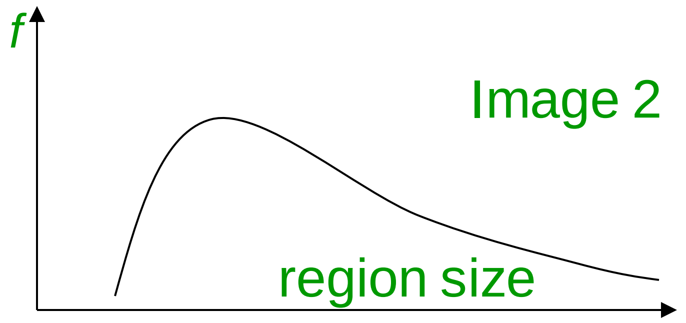
Scale Invariant Detection

- Common approach:
  - Take a local maximum of this function

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

- Observation: region size, for which the maximum is achieved, should be *co-variant* with image scale.



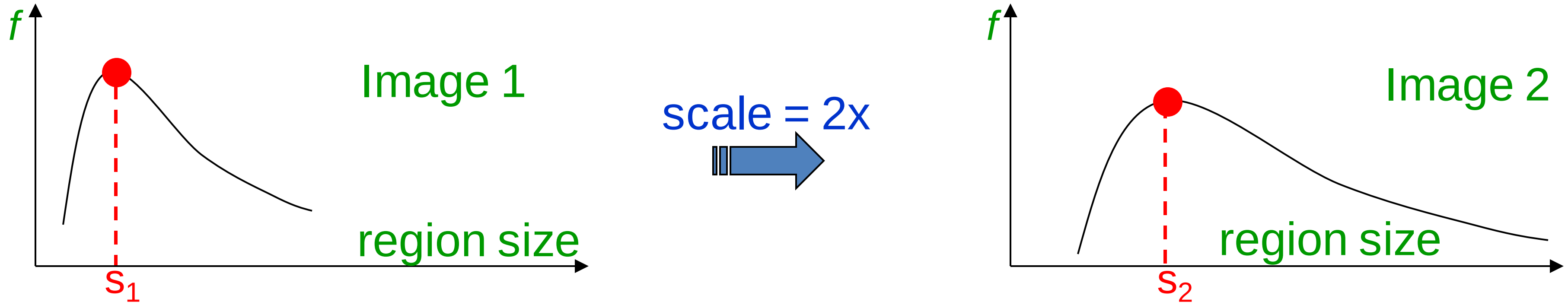


Scale Invariant Detection

- Common approach:
  - Take a local maximum of this function

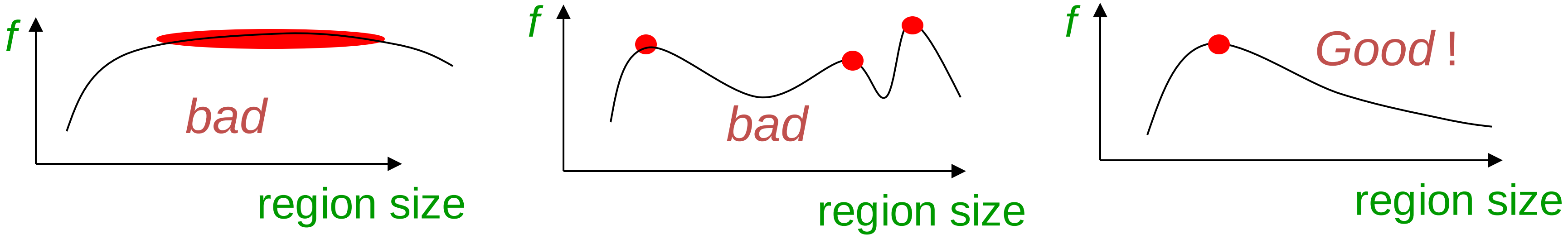
- Important: this scale invariant region size is found in each image independently!

- Observation: region size, for which the maximum is achieved, should be *co-variant* with image scale.



Scale Invariant Detection

- A "good" function for scale selection has one stable sharp peak



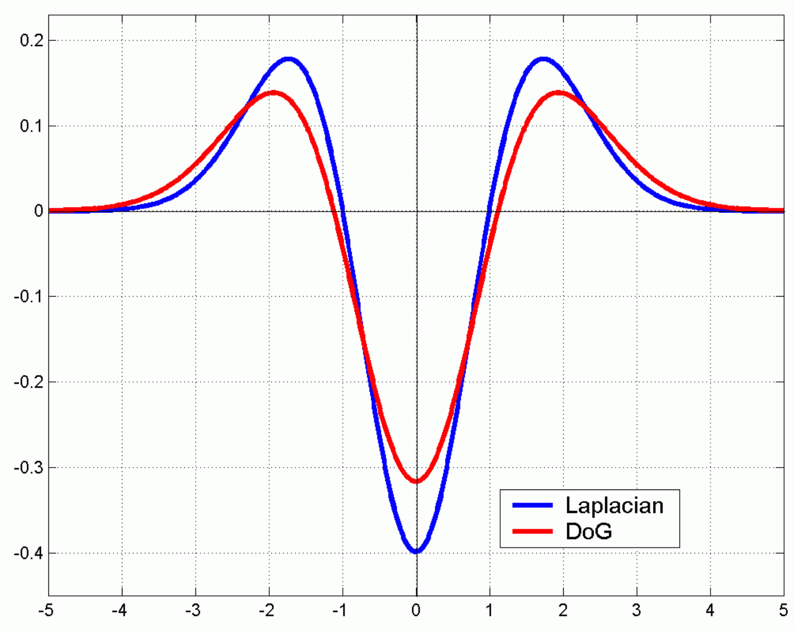
- For usual images: a good function would be one which responds to contrast (sharp local intensity change)

Scale Invariant Detection

- We will use functions for determining scale that have the form of $f = \text{Kernel} * \text{Image}$
- Choices of Kernel:
  - Laplacian:
    $$L = \sigma^2 \left(G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma) \right)$$
  - Difference of Gaussians (DoG):
    $$\text{DoG} = G(x, y, k\sigma) - G(x, y, \sigma)$$

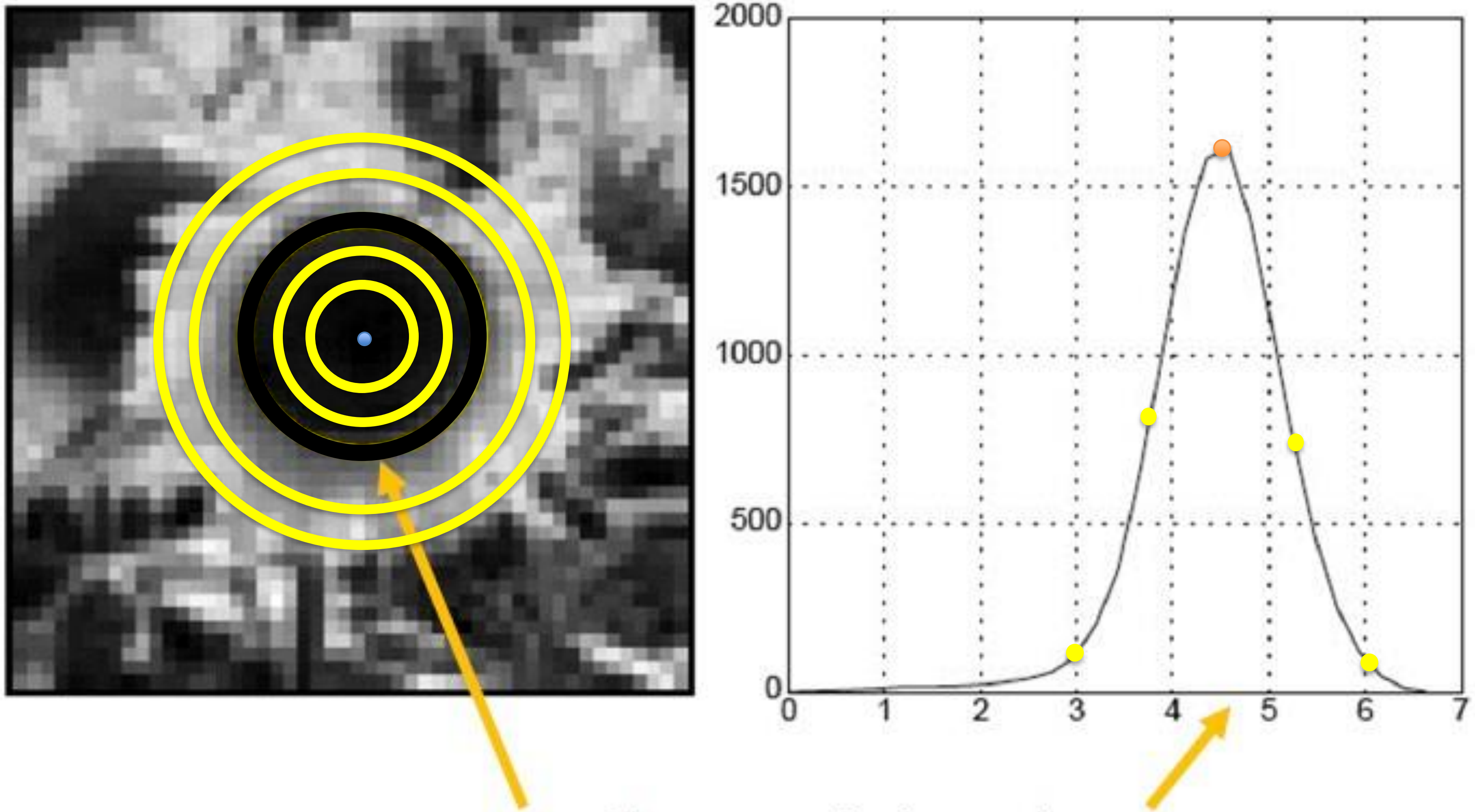
  where G is the Gaussian function:

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



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

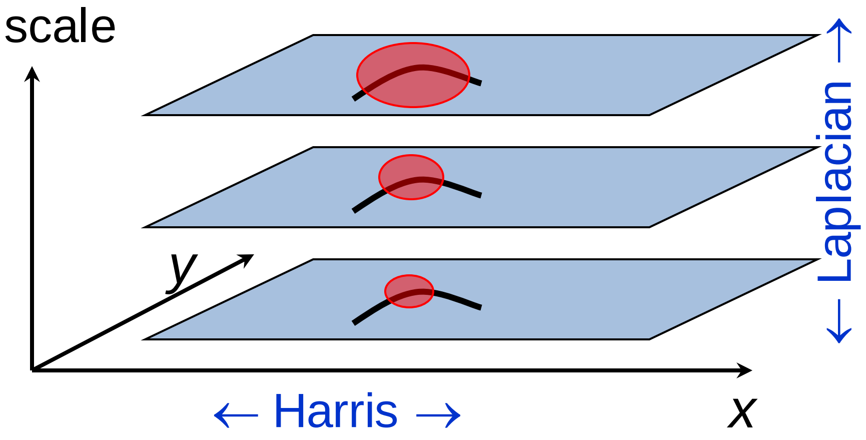
Laplacian



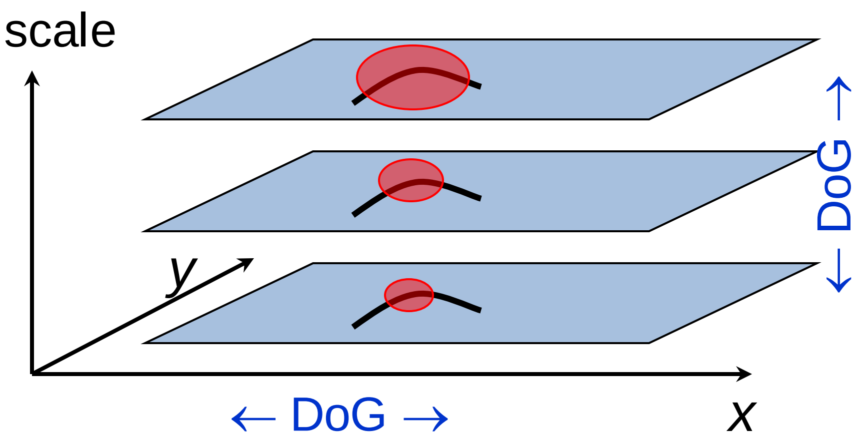
Characteristic scale

Scale Invariant Detectors

- Harris-Laplacian[1]
  Find local maximum of:
  - Harris corner detector in space (image coordinates)
  - Laplacian in scale



- DoG (from SIFT by Lowe)[2]
  Find local maximum of:
  - Difference of Gaussians in space and scale



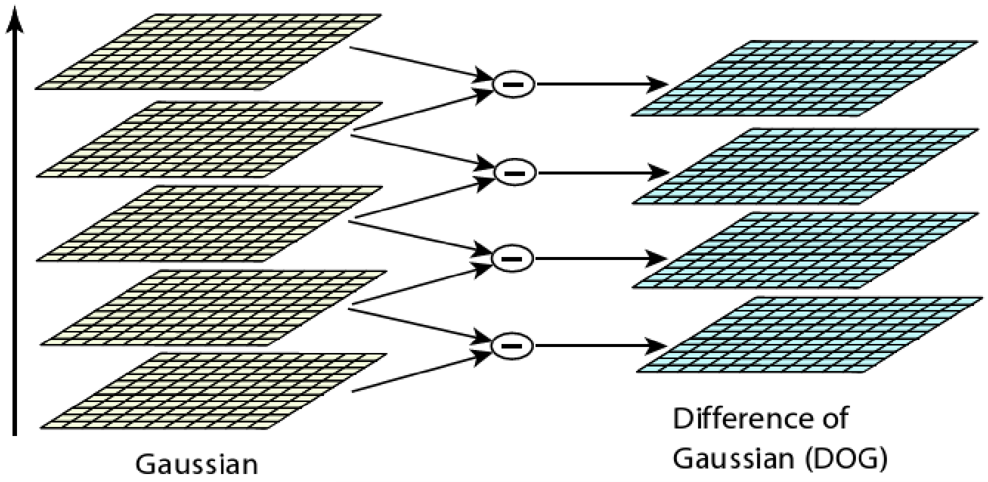
[1] K.Mikolajczyk, C.Schmid. “Indexing Based on Scale Invariant Interest Points”. ICCV 2001
[2] D.Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. IJCV 2004

Difference-of-Gaussians

$G(k^2\sigma) * I$

$G(k\sigma) * I$

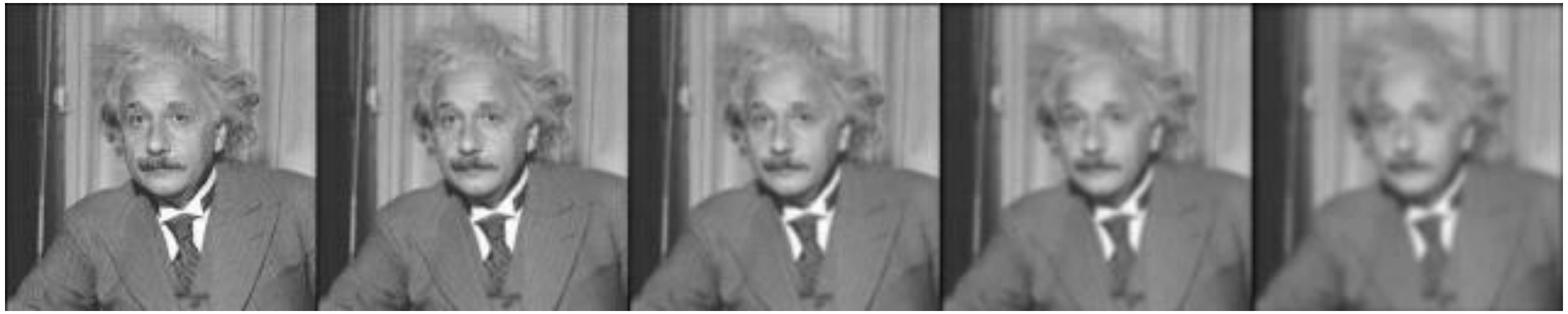
$G(\sigma) * I$



$$D(\sigma) \equiv (G(k\sigma) - G(\sigma)) * I$$

Scale, σ

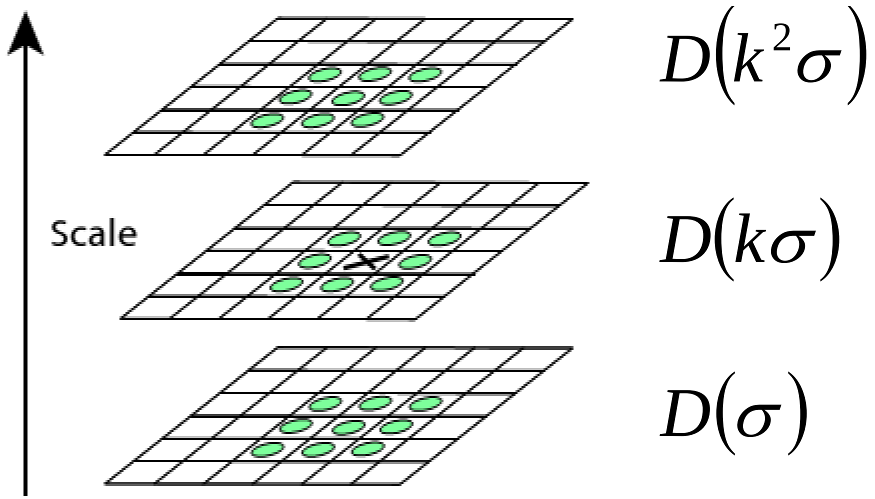
Gaussian:



DoG:

Scale-Space Extrema

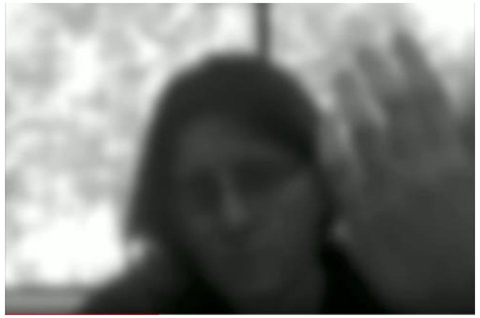
- Choose all extrema within 3x3x3 neighborhood.



X is selected if it is larger or smaller than all 26 neighbors

Difference of Gaussians (DoG) example

Original video



Blurred with a Gaussian kernel

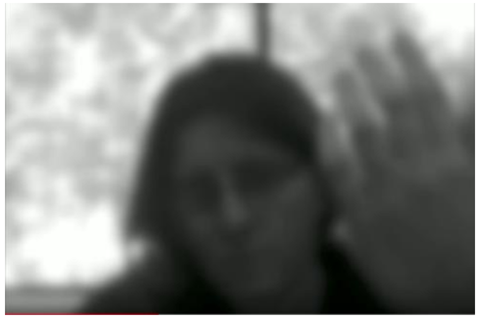
Blurred with a different Gaussian kernel

What happens if you subtract one blurred image from another?

Source: https://www.youtube.com/watch?v=oTud1De_W4s

Difference of Gaussians (DoG) example

Original video



Blurred with a Gaussian kernel: k_1

Blurred with a different Gaussian kernel: k_2

DoG: $k_1 - k_2$

DoG: $k_1 - k_3$

DoG: $k_1 - k_4$

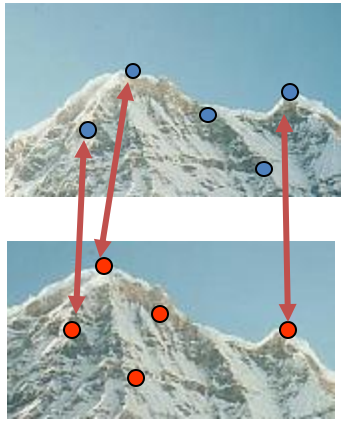
Source: https://www.youtube.com/watch?v=oTud1De_W4s

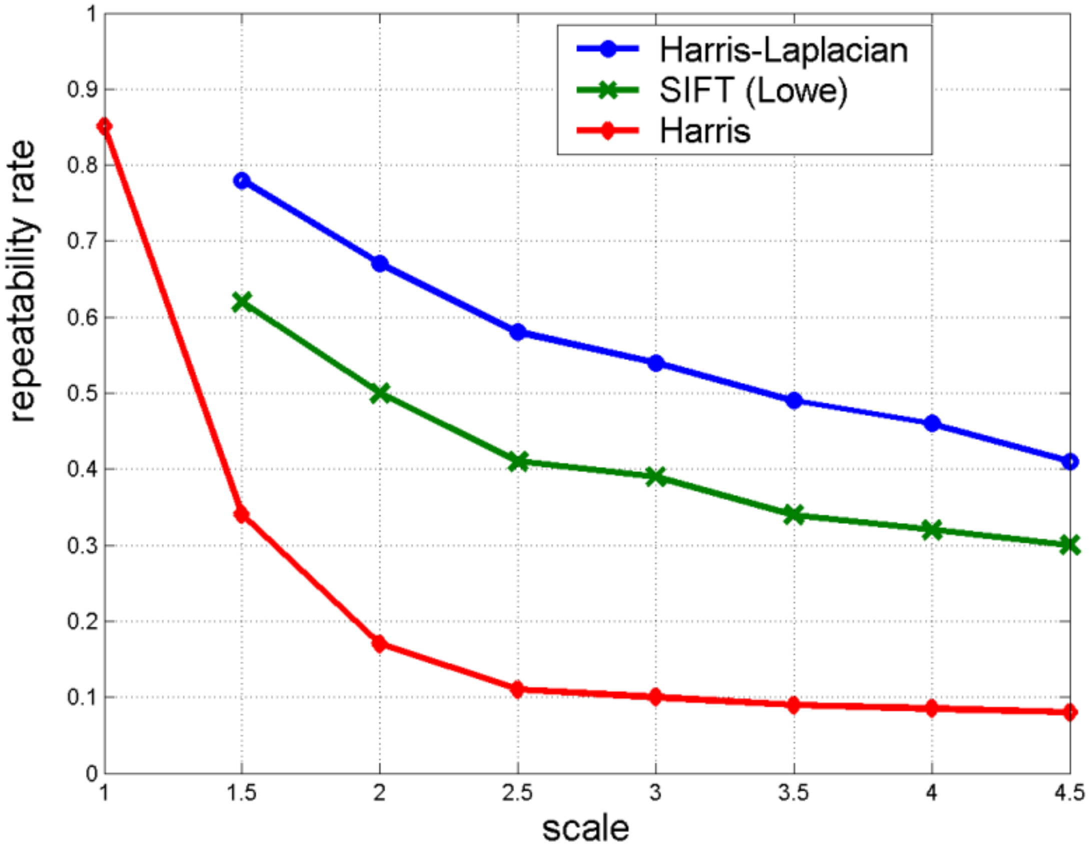
Scale Invariant Detectors

- Experimental evaluation of detectors w.r.t. scale change

Repeatability rate:

$$\frac{\#\ \text{correspondences}}{\#\ \text{possible correspondences}}$$





K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001

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

What’s next?

We now can detect keypoints at varying scales. But what can we do with those keypoints?

Things we would like to do:

- Search:
  - We would need to find similar key points in other images
- Panorama stitching
  - Match keypoints from one image to another.
- Etc…

For all such applications, we need a way of `describing` the keypoints.

Summary

- Scale invariant keypoint detection
  - Automatic scale selection
  - Harris-Laplace detector
  - Difference-of-Gaussian (DoG) detector