CMPSCI 370: Intro to Computer Vision

Image processing

Scale Invariant Feature Transform (SIFT)

University of Massachusetts, Amherst

March 03, 2015

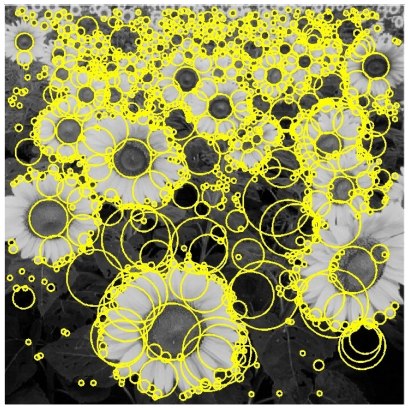
Instructor: Subhransu Maji

Administrivia

- Exam review session in next class
- Midterm in class (Thursday)
  - All topics covered till Feb 25 lecture (corner detection)
  - Closed book
- Grading issues
  - Include all the information needed to grade the homework
  - Keep the grader happy :-)
- Candy wrapper extra credit for participation (5%)

Scale invariant features

"blob detection"



Source: L. Lazebnik

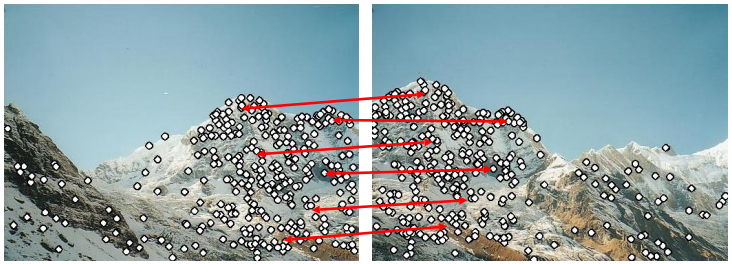
Why extract features?

- Motivation: panorama stitching
  - We have two images – how do we combine them?

Slide credit: L. Lazebnik

Why extract features?

- Motivation: panorama stitching
  - We have two images – how do we combine them?



Step 1: extract features

Step 2: match features

Why extract features?

- Motivation: panorama stitching
  - We have two images – how do we combine them?

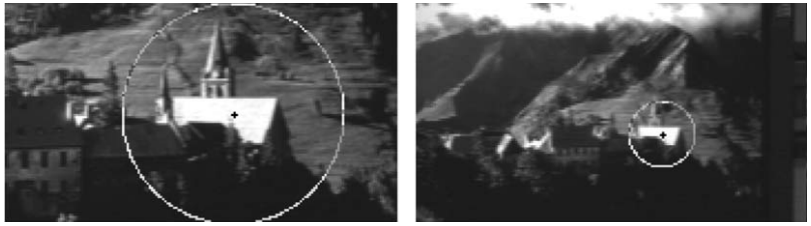
Step 1: extract features

Step 2: match features

Step 3: align images

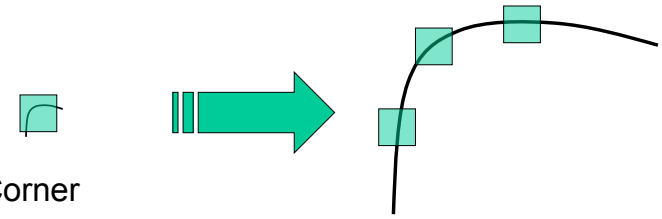
Feature detection with scale selection

- We want to extract features with characteristic scale that matches the image transformation such as scaling and translation (a.k.a. covariance)



Matching regions across scales

Scaling



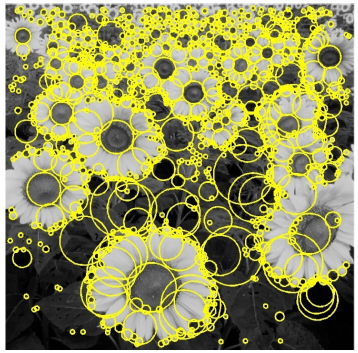
Corner

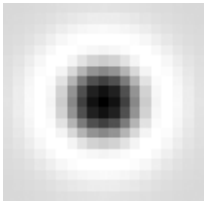
All points will be classified as edges

Corner detection is sensitive to the image scale!

Blob detection: basic idea

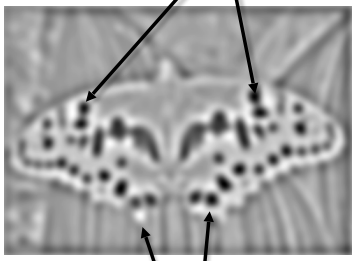
- Convolve the image with a “blob filter” at multiple scales
- Look for extrema (maxima or minima) of filter response in the resulting *scale space*
- This will give us a scale and space covariant detector





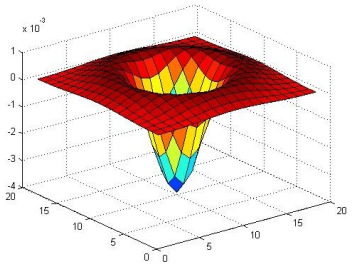
Blob detection: basic idea

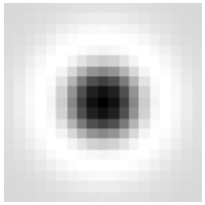
Find maxima ***and minima*** of blob filter response in space ***and scale***



Blob filter

Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D

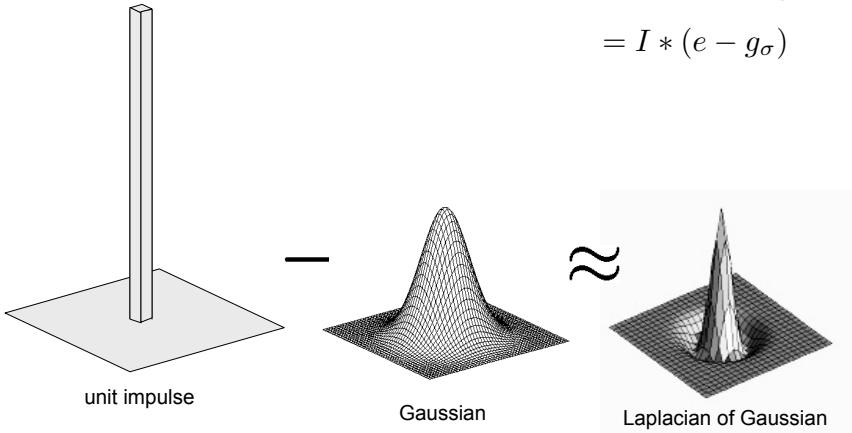




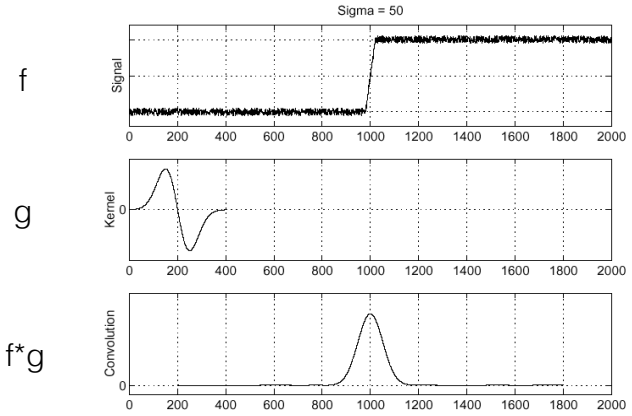
$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

Recall: sharpening filter

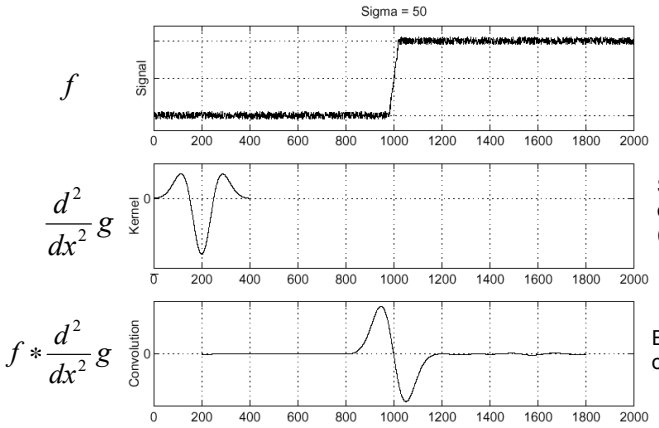
$$I = blurry(I) + sharp(I) \qquad sharp(I) = I - blurry(I)$$
$$= I * e - I * g_\sigma$$
$$= I * (e - g_\sigma)$$



Recall: edge detection

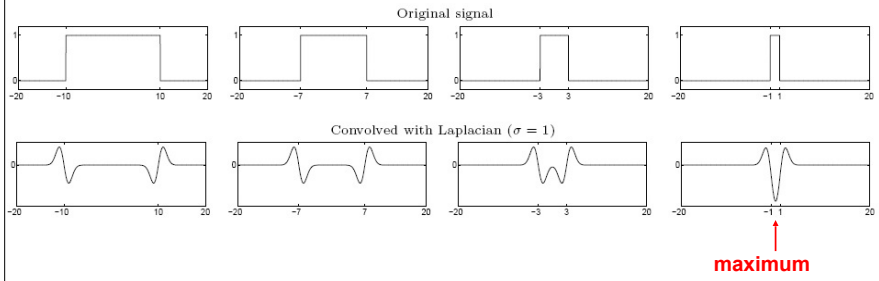


Edge detection using a Laplacian



From edges to blobs

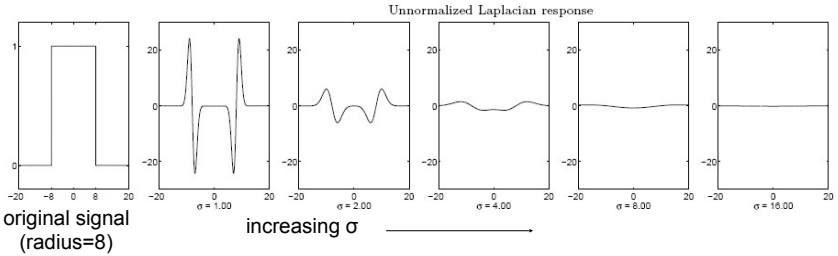
- edge = ripple
- blob = superposition of two ripples



Spatial selection: the magnitude of the Laplacian response will achieve a maximum at the center of the blob, provided the scale of the Laplacian is "matched" to the scale of the blob

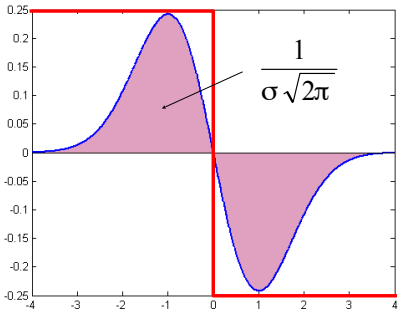
Scale selection

- We want to find the characteristic scale of the blob by convolving it with Laplacians at several scales and looking for the maximum response
- However, Laplacian response decays as scale increases:



Scale normalization

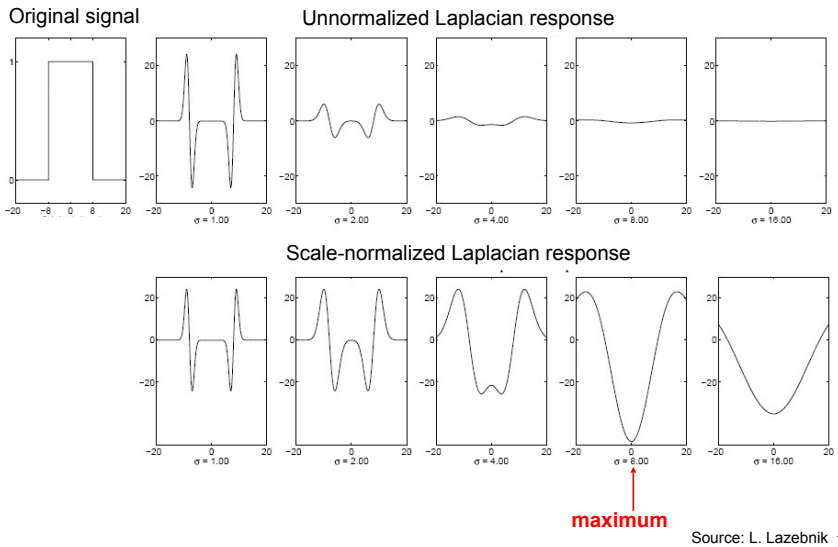
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases



Scale normalization

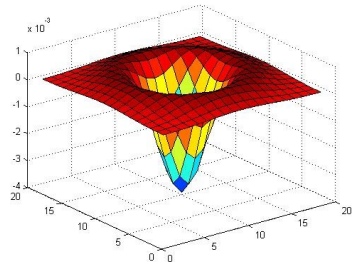
- The response of a derivative of Gaussian filter to a perfect step edge decreases as σ increases
- To keep response the same (scale-invariant), must multiply Gaussian derivative by σ
- Laplacian is the second Gaussian derivative, so it must be multiplied by σ^2

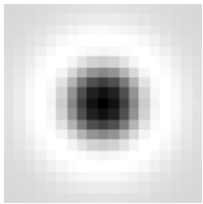
Effect of scale normalization



Blob detection in 2D

Laplacian of Gaussian: Circularly symmetric operator for blob detection in 2D

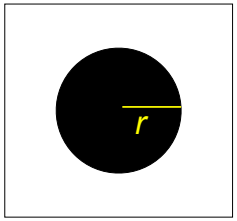




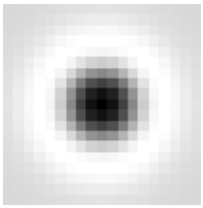
Scale-normalized: $\nabla^2_{\text{norm}} g = \sigma^2 \left(\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} \right)$

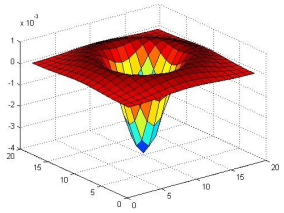
Scale selection

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



image





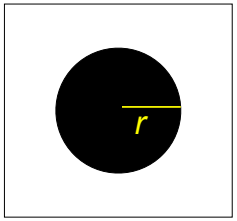
Laplacian

Scale selection

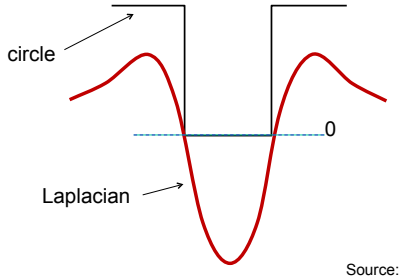
- At what scale does the Laplacian achieve a maximum response to a binary circle of radius r?
- To get maximum response, the zeros of the Laplacian have to be aligned with the circle
- The Laplacian is given by (up to scale):

$$(x^2 + y^2 - 2\sigma^2)\, e^{-(x^2+y^2)/2\sigma^2}$$

- Therefore, the maximum response occurs at $\sigma = r/\sqrt{2}.$

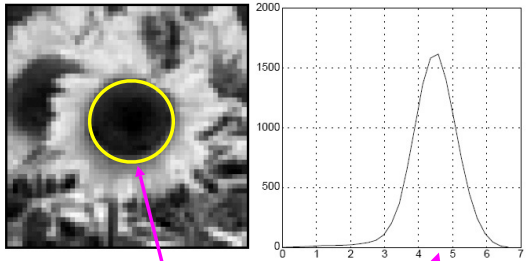


image



Characteristic scale

- We define the characteristic scale of a blob as the scale that produces peak of Laplacian response in the blob center



characteristic scale

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

Scale-space blob detector

1. Convolve image with scale-normalized Laplacian at several scales

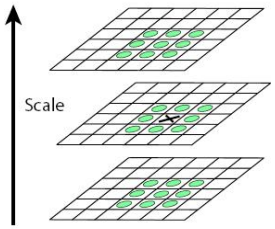
Scale-space blob detector: Example

Scale-space blob detector: Example

sigma = 11.9912

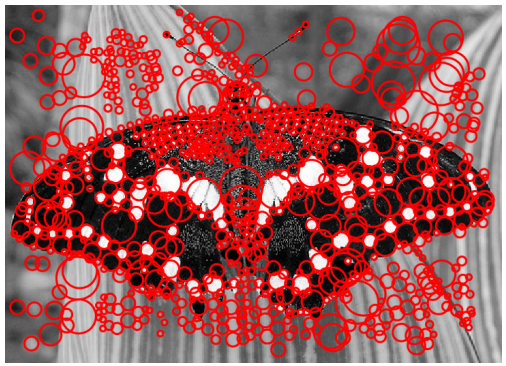
Scale-space blob detector

1. Convolve image with scale-normalized Laplacian at several scales
2. Find maxima of squared Laplacian response in scale-space



Source: L. Lazebnik

Scale-space blob detector: Example



Source: L. Lazebnik

Efficient implementation

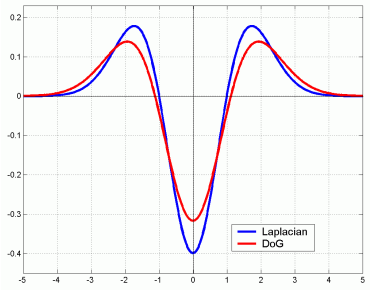
- Approximating the Laplacian with a difference of Gaussians:

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

(Laplacian)

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

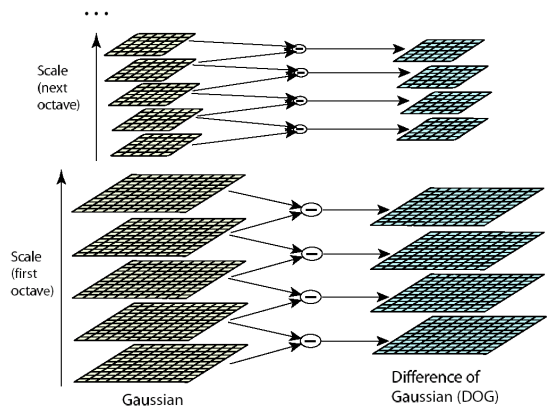
(Difference of Gaussians)



$$(x^2 + y^2 - 2\sigma^2)\, e^{-(x^2+y^2)/2\sigma^2}$$

Is the Laplacian separable?

Efficient implementation



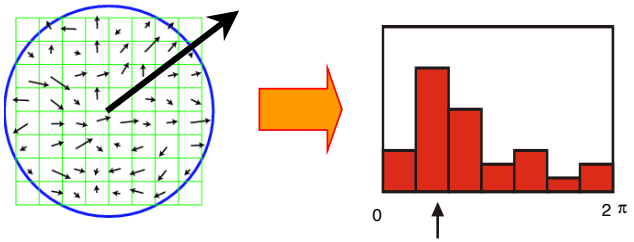
David G. Lowe. **"Distinctive image features from scale-invariant keypoints."** *IJCV* 60 (2), pp. 91-110, 2004.

From feature detection to description

- Scaled and rotated versions of the same neighborhood will give rise to blobs that are related by the same transformation
- What to do if we want to compare the appearance of these image regions?
  - **Normalization**: transform these regions into same-size circles
  - Problem: rotational ambiguity

Eliminating rotation ambiguity

- To assign a unique orientation to circular image windows:
  - Create histogram of local gradient directions in the patch
  - Assign canonical orientation at peak of smoothed histogram

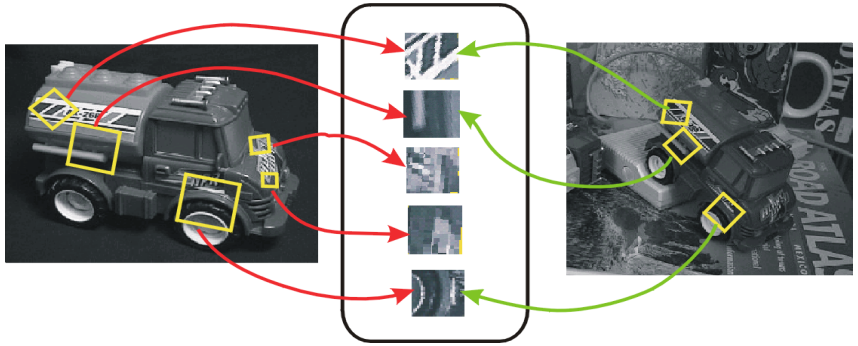


SIFT features

- Detected features with characteristic scales and orientations:

David G. Lowe. **"Distinctive image features from scale-invariant keypoints."** *IJCV* 60 (2), pp. 91-110, 2004.

From feature detection to description



how should we represent the patches?