Biomedical Imaging & Analysis

Lecture 6, Part 6. Fall 2014

Basic Image Processing / Filtering (VI)

*[**Text**: Ch: 10, Gonzalez and Woods, Digital Image Processing (3rd Edition) + Papers / Reading Assignments on Blackboard.]*

Prahlad G Menon, PhD

Assistant Professor

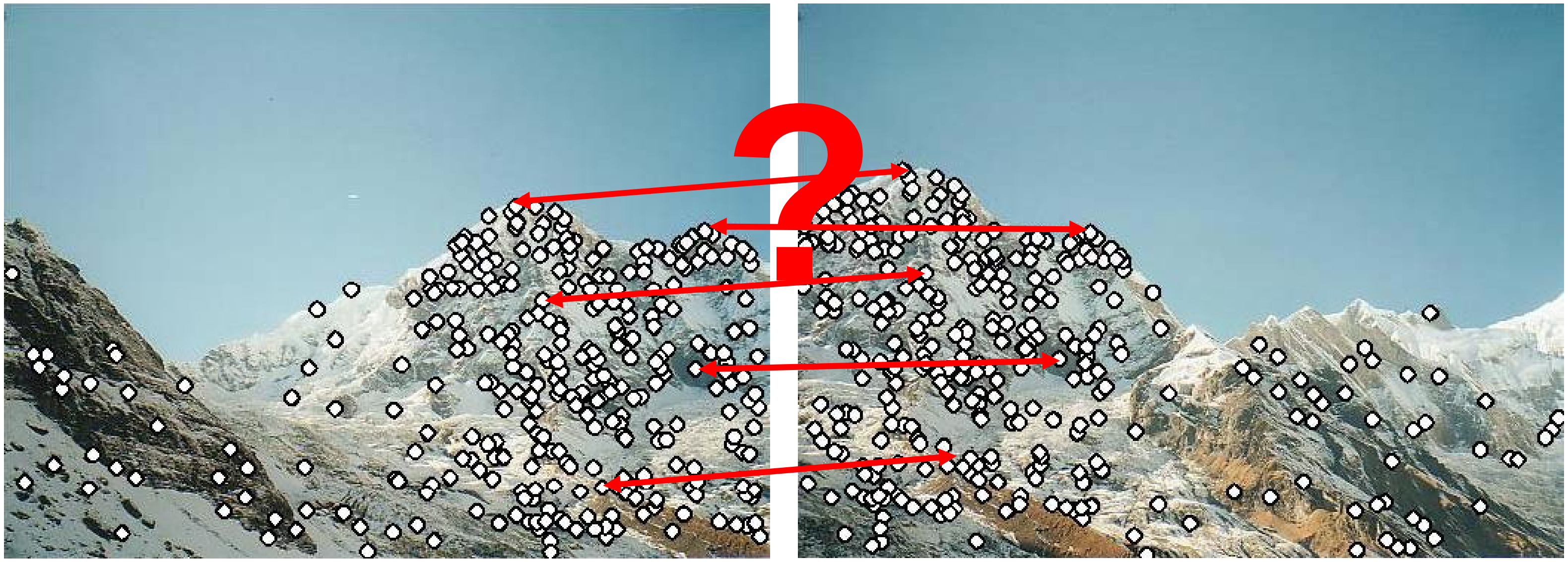
Sun Yat-sen University – Carnegie Mellon University (SYSU-CMU)

Joint Institute of Engineering

Point Descriptors

- We know how to detect points
- Next question:

How to match them?

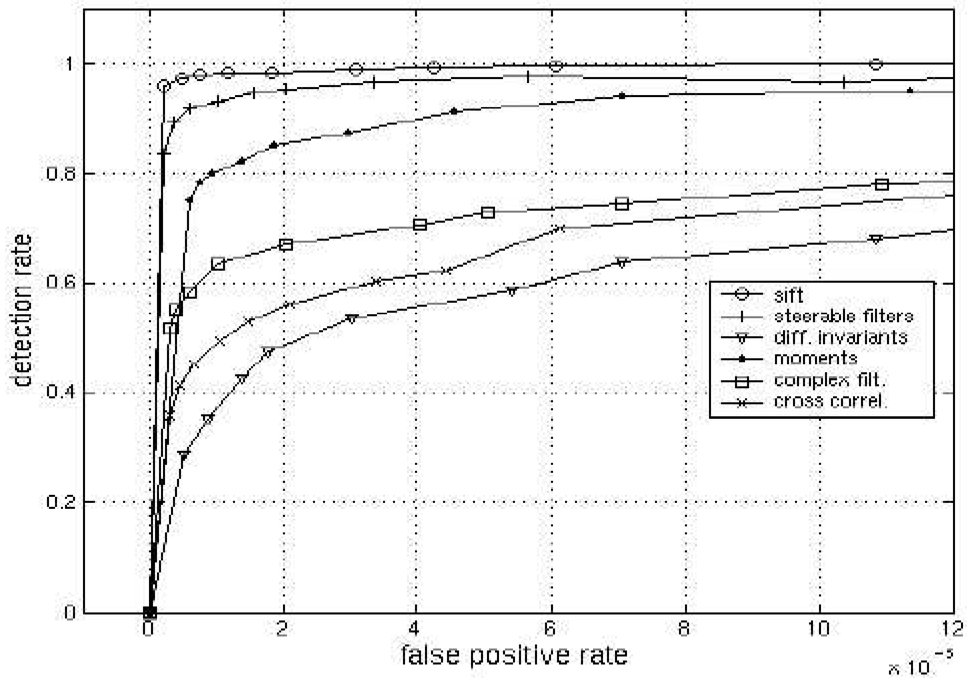


Point descriptor should be:
1. Invariant
2. Distinctive

SIFT – Scale Invariant Feature Transform[1]

- Empirically found[2] to show very good performance, invariant to *image rotation*, *scale*, *intensity change*, and to moderate *affine* transformations

Scale = 2.5
Rotation = 45^0

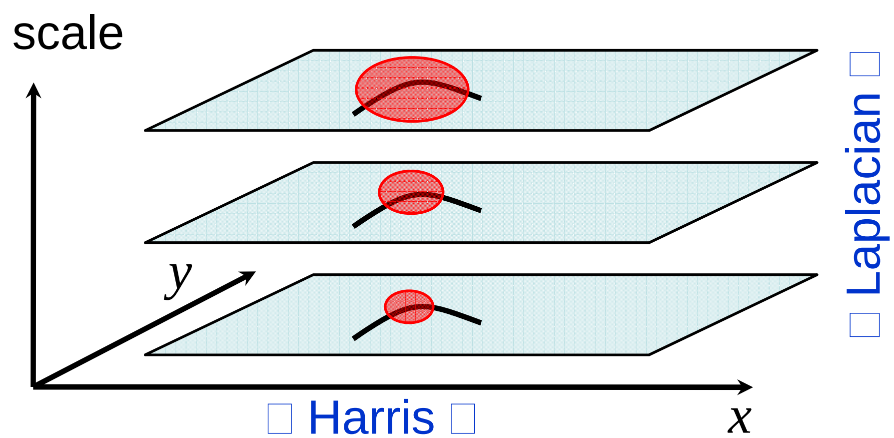


[1] D.Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. IJCV 2004
[2] K.Mikolajczyk, C.Schmid. “A Performance Evaluation of Local Descriptors”. CVPR 2003

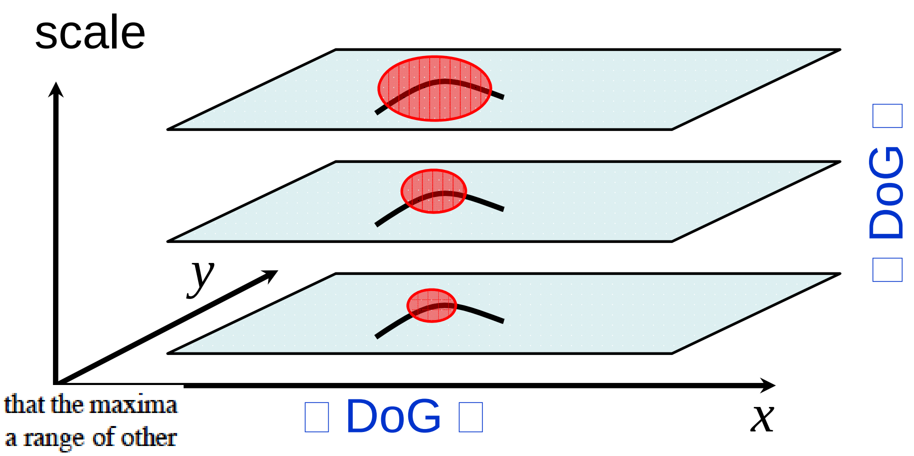
Scale Invariant Detectors

Darya Frolova, Denis Simakov The Weizmann Institute of Science

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



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

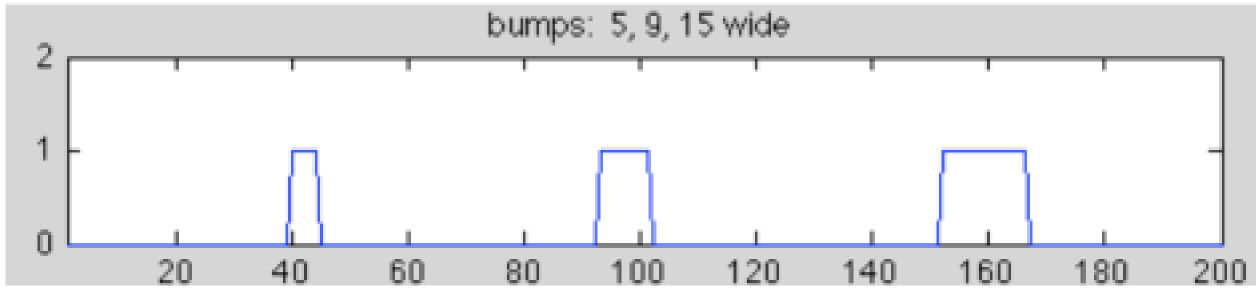


In detailed experimental comparisons, Mikolajczyk (2002) found that the maxima and minima of $\sigma^2\nabla^2 G$ produce the most stable image features compared to a range of other possible image functions, such as the gradient, Hessian, or Harris corner function.

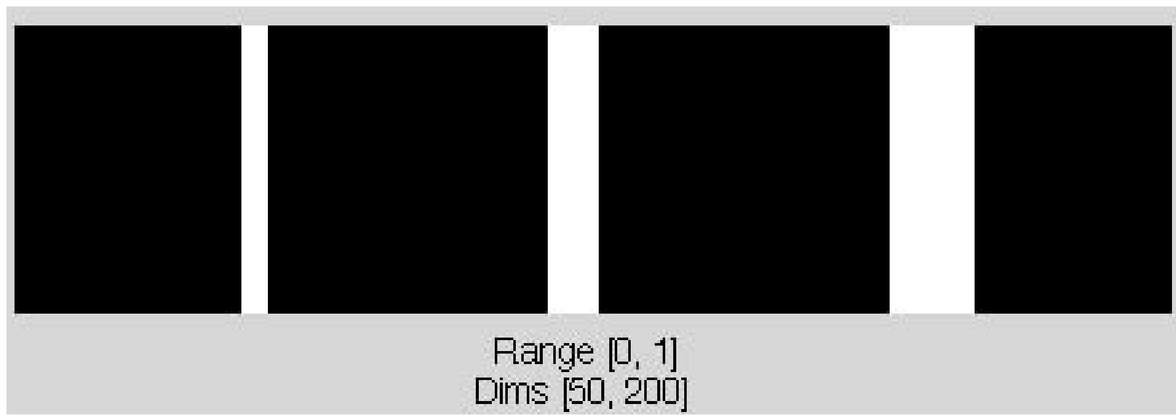
[1] K.Mikolajczyk, C.Schmid. "Indexing Based on Scale Invariant Interest Points". ICCV 2001
[2] D.Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". Accepted to IJCV 2004

Scale-space example: 3 bumps of different widths…

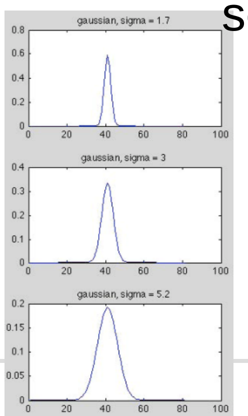
1-d bumps



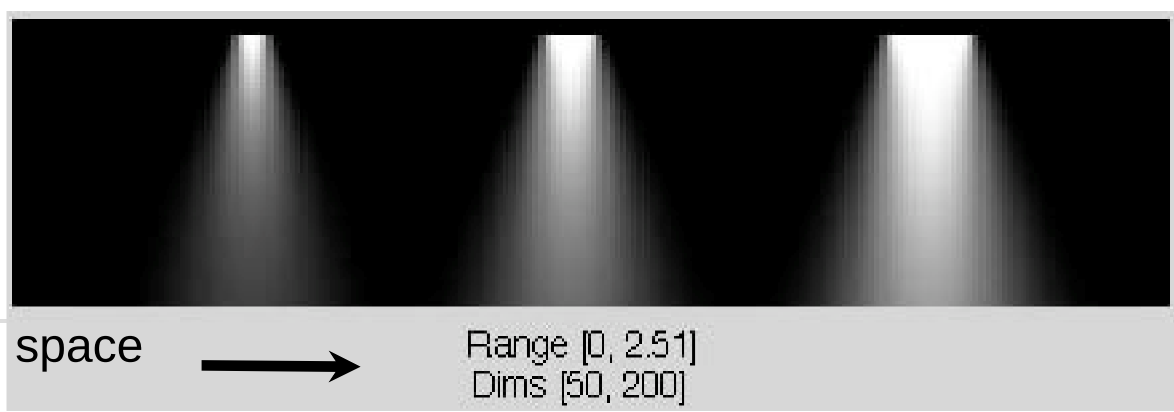
display as an image



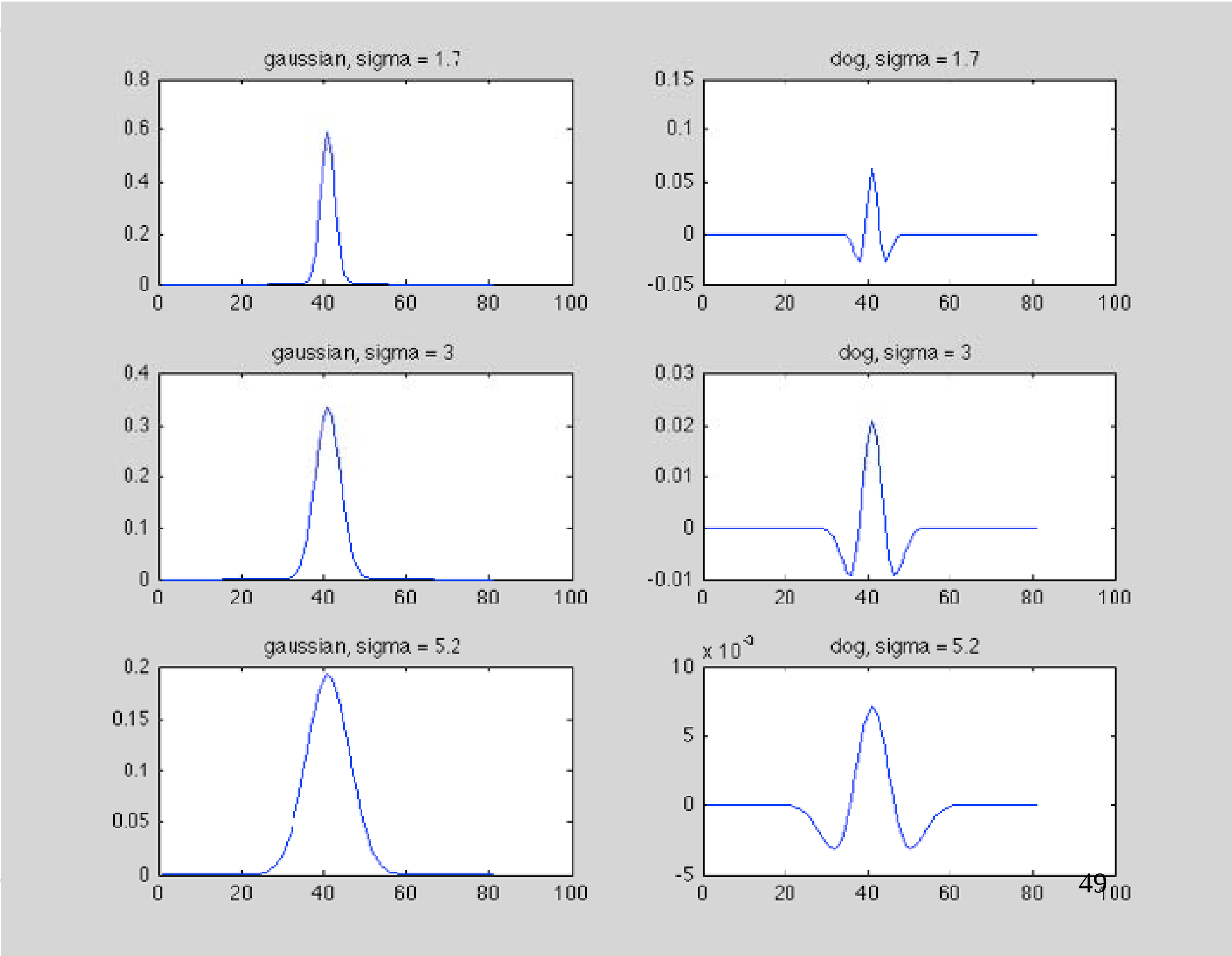
blur with Gaussians of increasing width



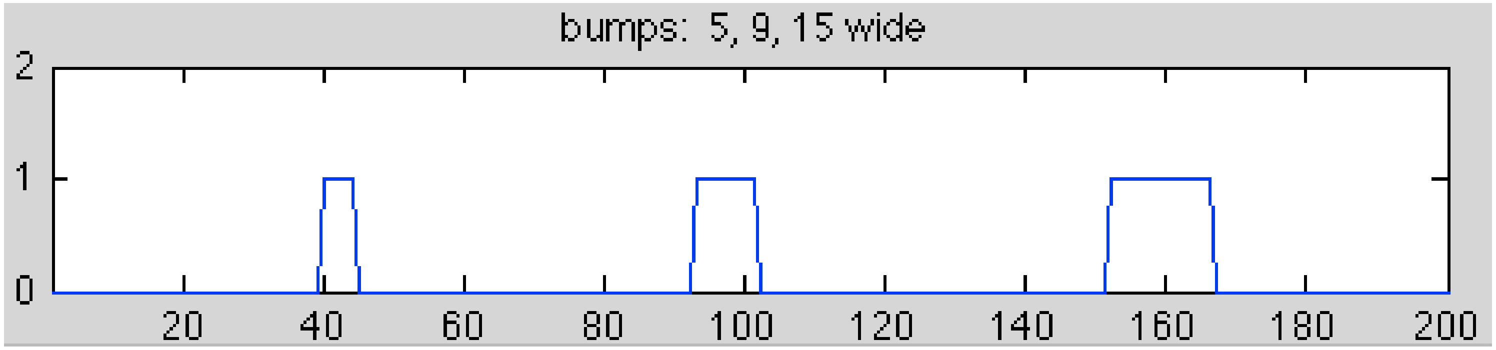
scale



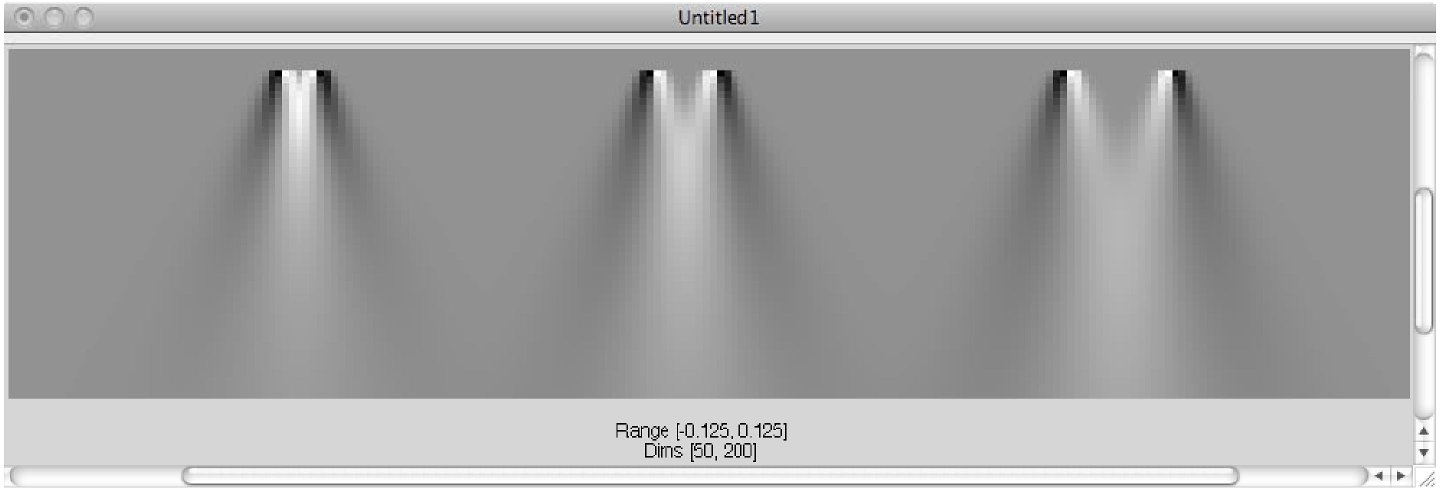
Gaussian v/s Difference-of-Gaussian filters



The bumps, filtered by Difference-of-Gaussian (DoG) filters

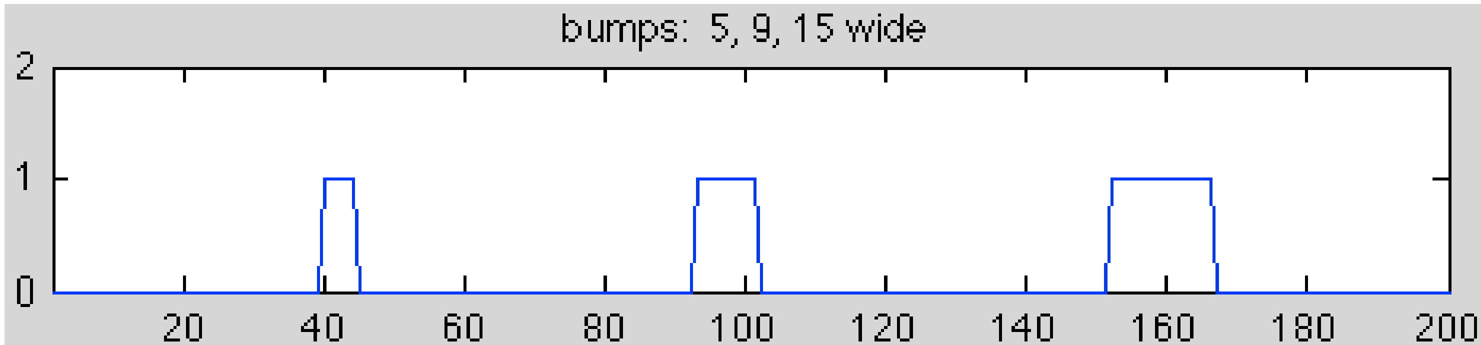


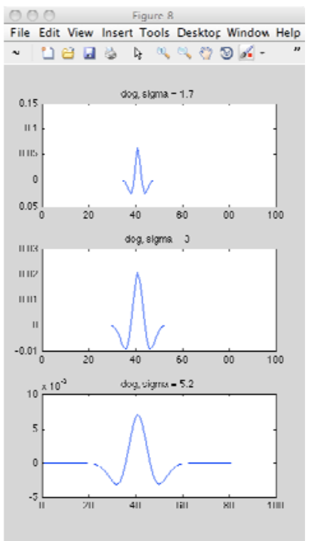
scale

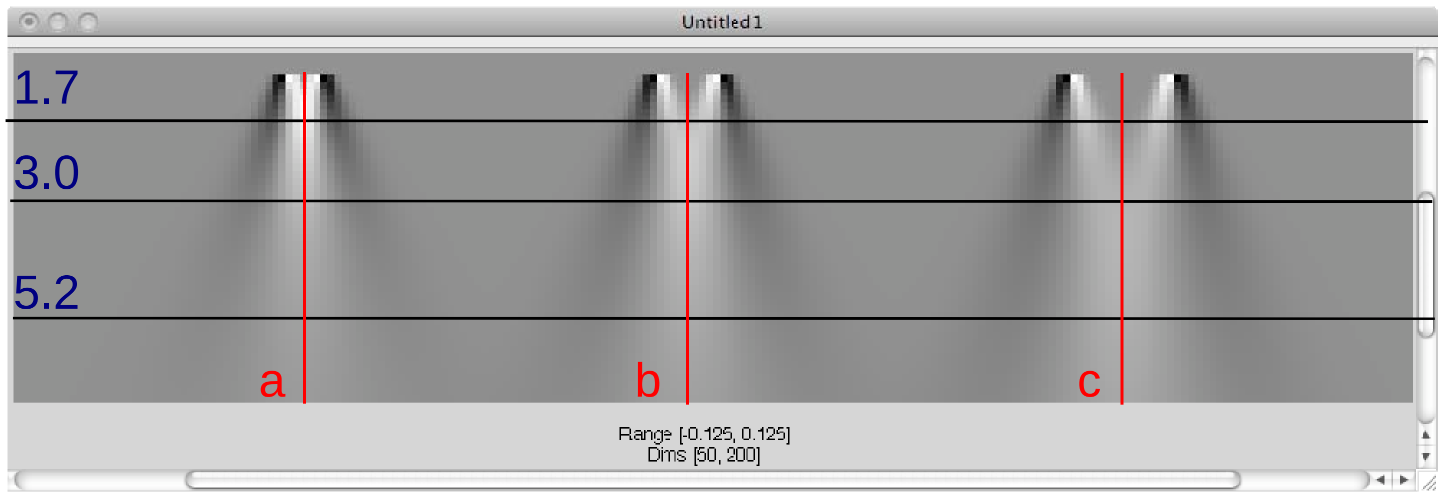


space

The bumps, filtered by difference-of-Gaussian (DoG) filters

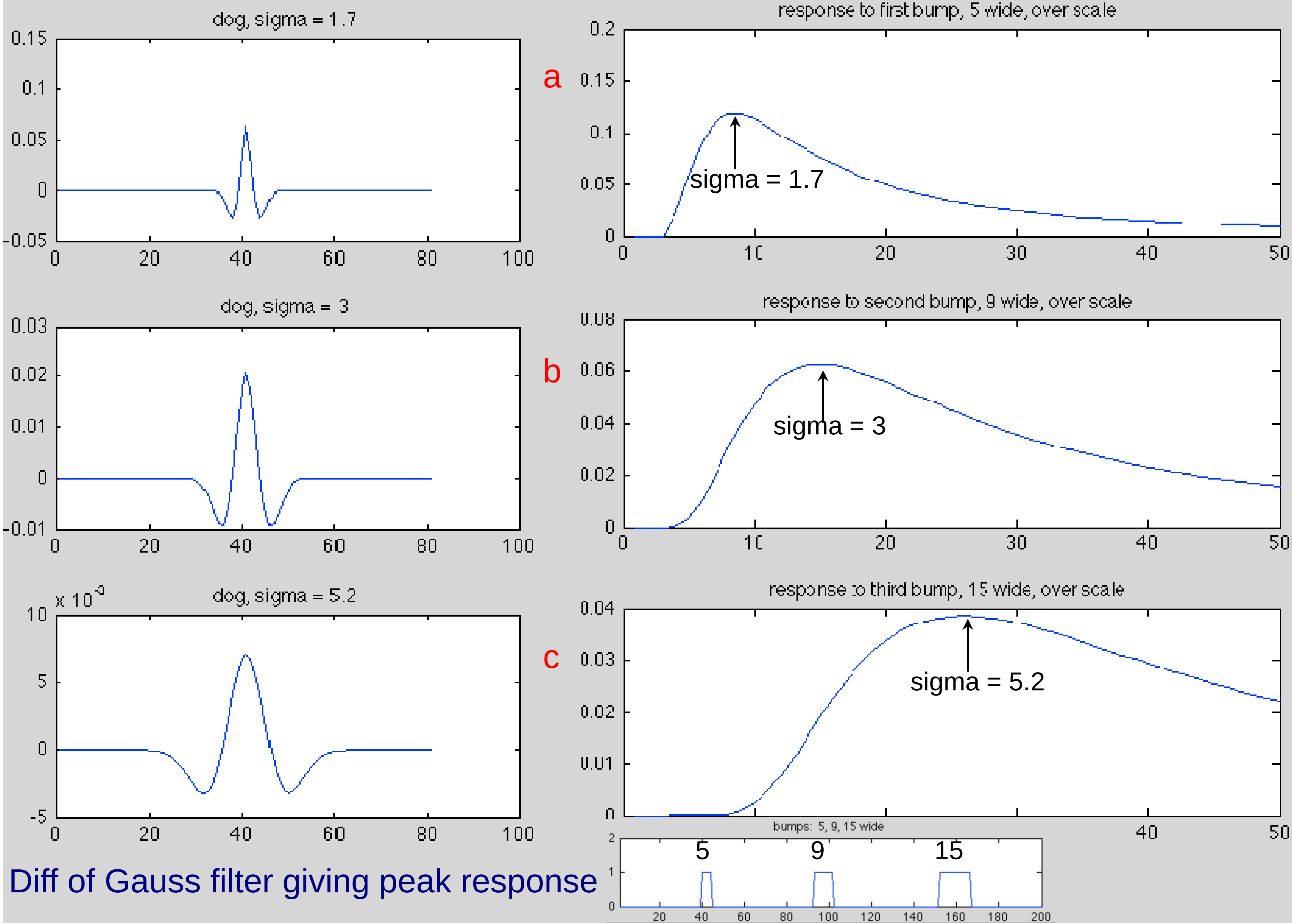






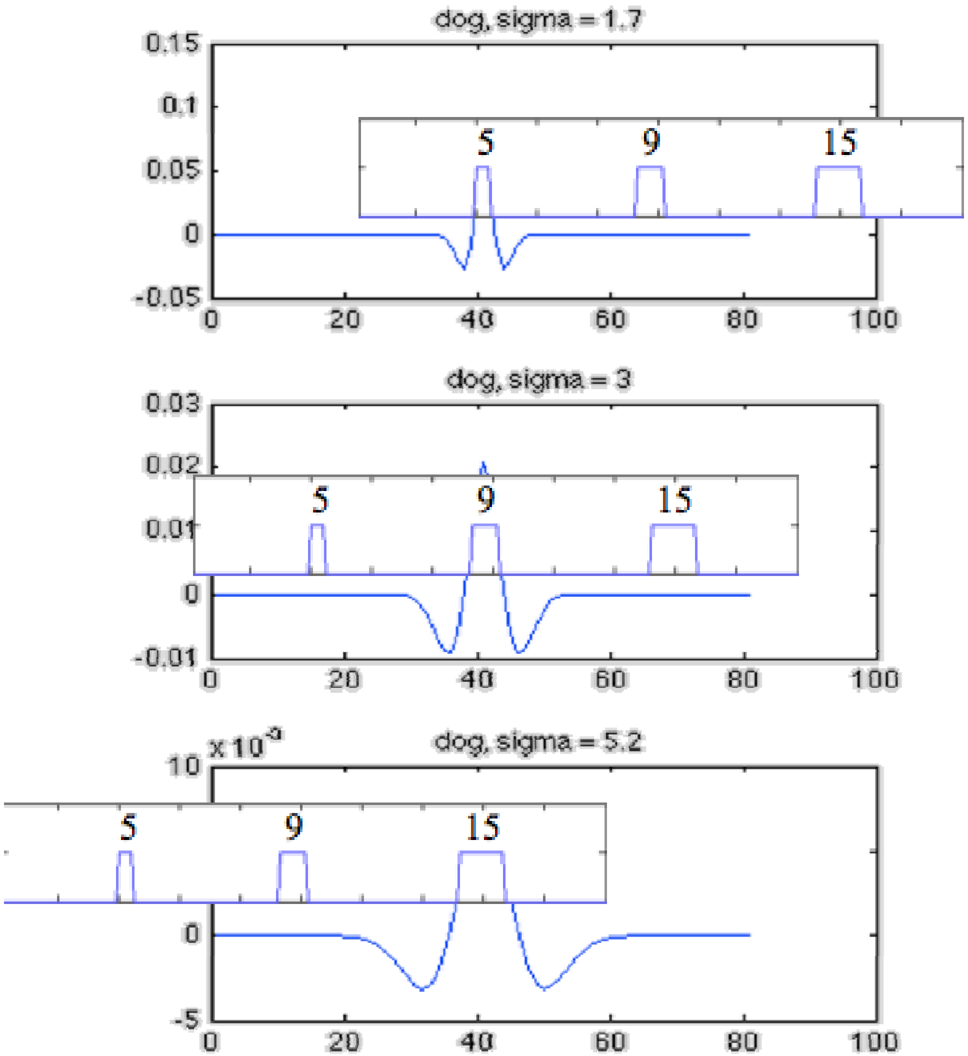
cross-sections along red lines plotted next slide

Scales of peak responses are proportional to bump width (the characteristic scale of each bump): [1.7, 3, 5.2] ./ [5, 9, 15] = 0.3400 0.3333 0.3467



Diff of Gauss filter giving peak response

The bumps, filtered by difference-of-Gaussian (DoG) filters



Scales of peak responses are proportional to bump width (the characteristic scale of each bump): [1.7, 3, 5.2] ./ [5, 9, 15] = 0.3400 0.3333 0.3467

Note that the max response filters each has the same relationship to the bump that it favors (the zero crossings of the filter are about at the bump edges). So the scale space analysis correctly picks out the "characteristic scale" for each of the bumps.

More generally, this happens for the features of the images we analyze.

Blob detection by Laplacian of Gaussian (LoG) filters

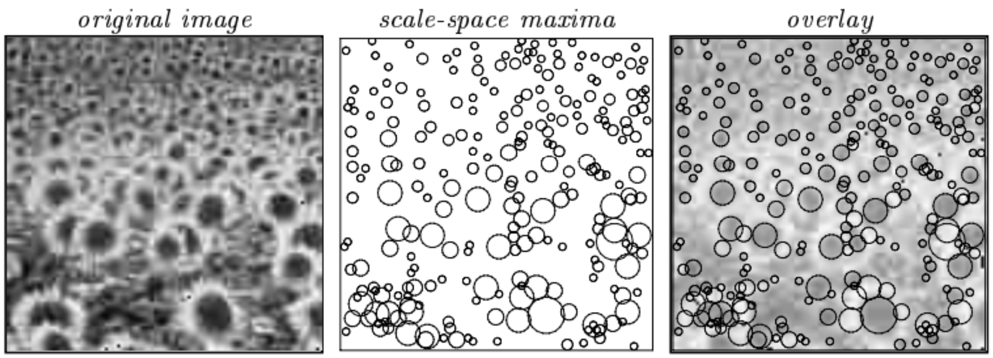


Figure 1.12: *Blob detection by detection of scale-space maxima of the normalized Laplacian operator: (a) Original image. (b) Circles representing the 250 scale-space maxima of $(\nabla_{\text{norm}} L)^2$ having the strongest normalized response. (c) Circles overlayed on image.*

Reading Assignment (PDF on Blackboard):

This will contribute to a Quiz in future. More reading assignments together will be mapped into the final quiz, so stay tuned!

$$\nabla^2_{\text{norm}} L = t(L_{xx} + L_{yy}) \tag{1.51}$$

assumes maxima with respect to space and scale. Such points are referred to as *scale-space extrema* of $(\nabla^2_{\text{norm}} L)^2$.

Scale space axis == *t*

Qualitative properties. For a Gaussian blob defined by

$$g(x, y) = h(x, y;\ t_0) = \frac{1}{2\pi t_0} e^{-(x^2+y^2)/2t_0} \tag{1.52}$$

it can be shown that the selected scale at the center of the blob is given by

$$\partial_t(\nabla^2_{\text{norm}} L)(0, 0;\ t) = 0 \quad \Longleftrightarrow \quad t_{\nabla^2 L} = t_0. \tag{1.53}$$

Hence, the selected scale directly reflects the width t_0 of the Gaussian blob.

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:

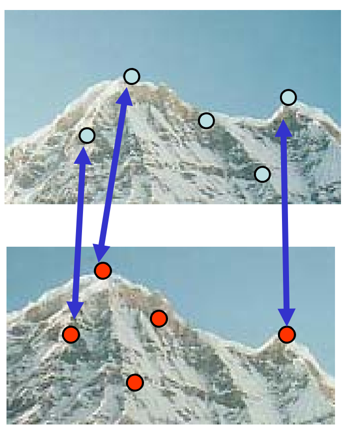
1. Harris-Laplacian [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. SIFT [Lowe]: maximize Difference of Gaussians over scale and space

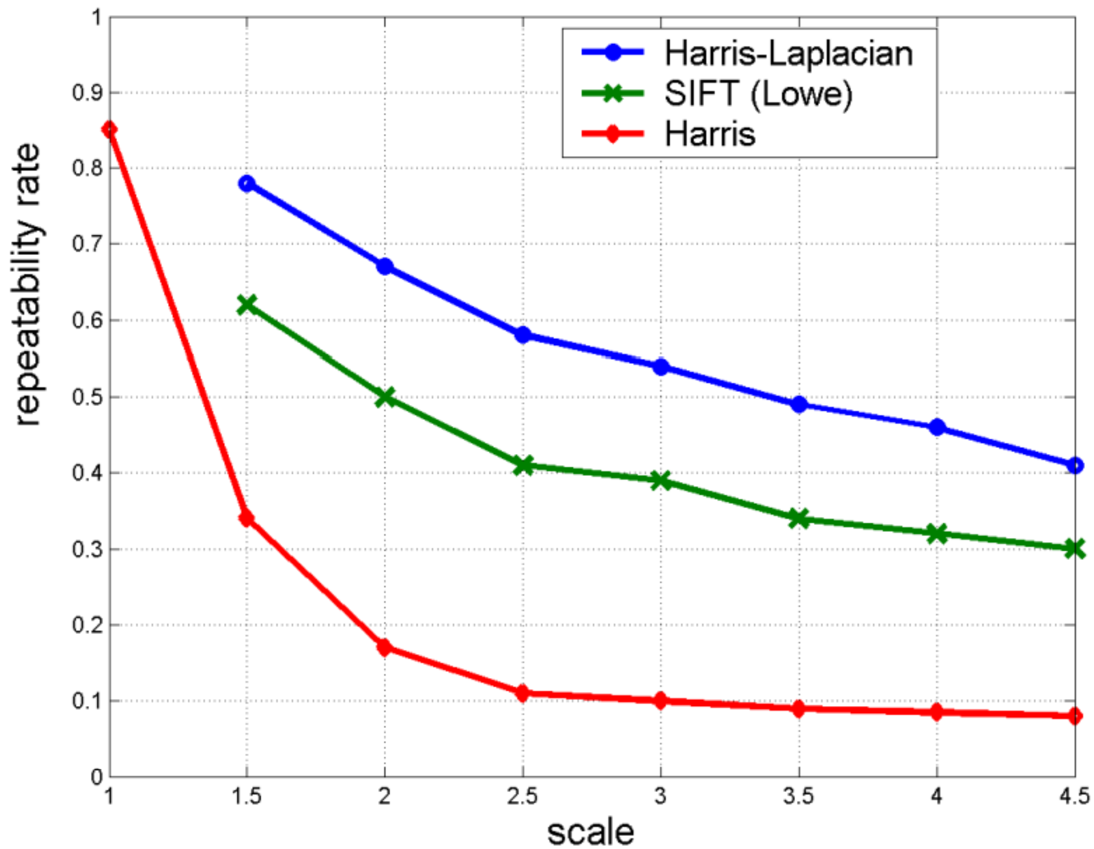
Scale Invariant Detectors

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

Repeatability rate:

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



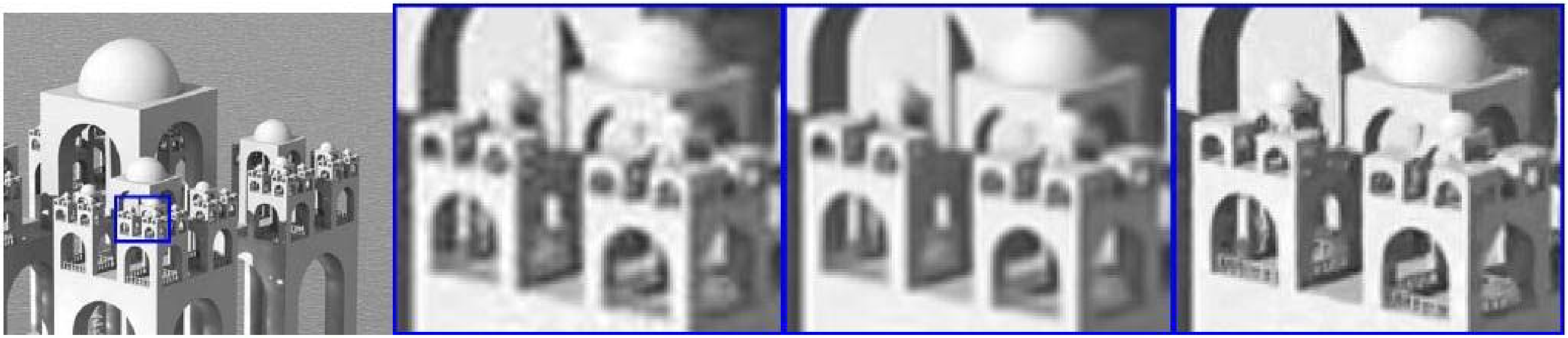


Darya Frolova, Denis Simakov The Weizmann Institute of Science

ppt

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

Super-Resolution (Learning from Repeated Patches in the same or multiple images)



(a) Input image (scaled for display). (b) Bicubic interpolation ($\times 2$). (c) Within image repetitions ($\times 2$). (d) Unified single-image SR ($\times 2$).

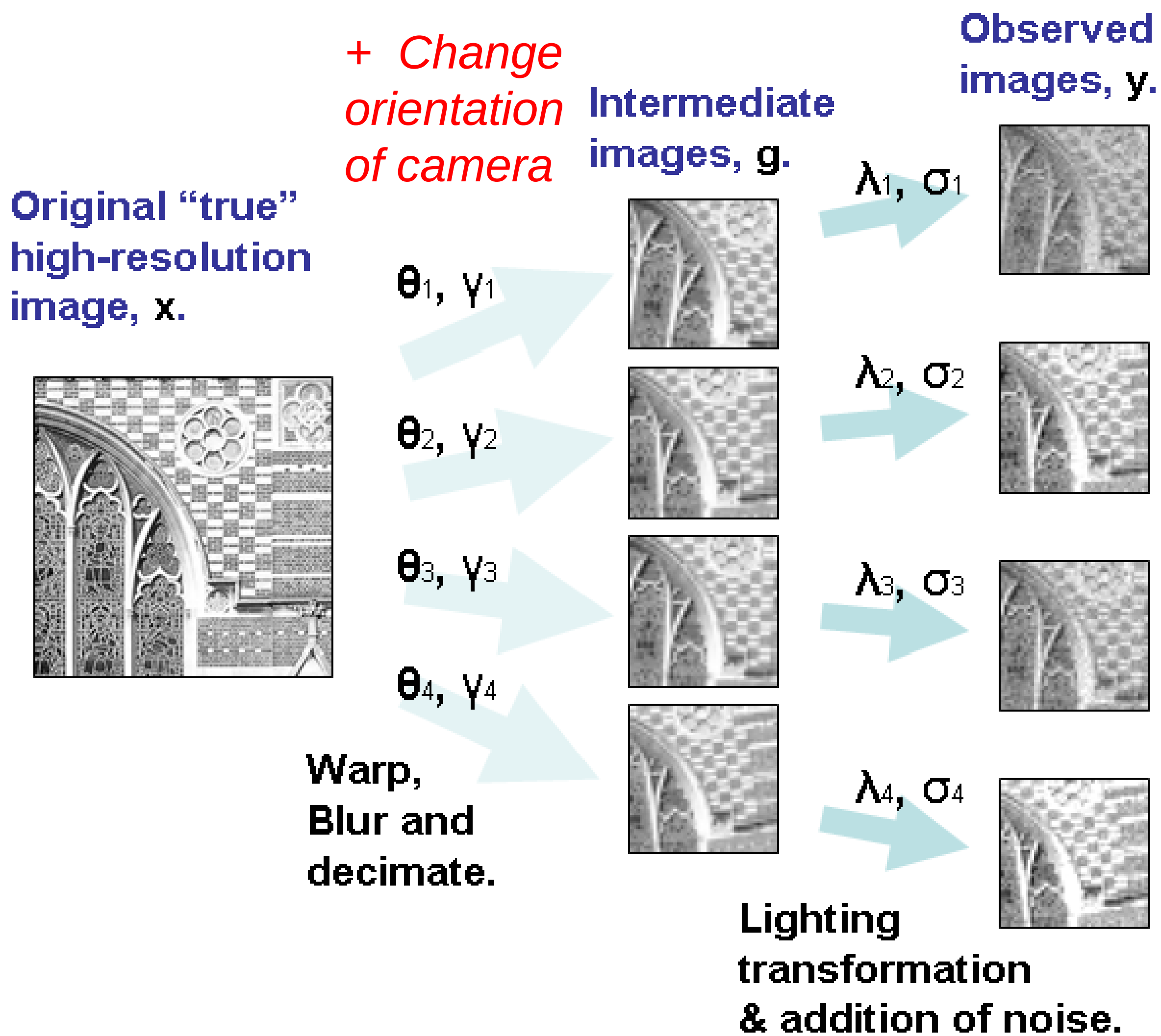
Glasner, D.; Bagon, S.; Irani, M.; , “Super-Resolution from a Single Image”, *12th International Conference on Computer Vision (ICCV), IEEE , pp. 349 – 356, 2009* URL: http://www.wisdom.weizmann.ac.il/~vision/SingleImageSR.html

- **Two types of Super-resolution:**

  (i) The classical multi-image super-resolution (combining images obtained at subpixel misalignments) + averaging ; and

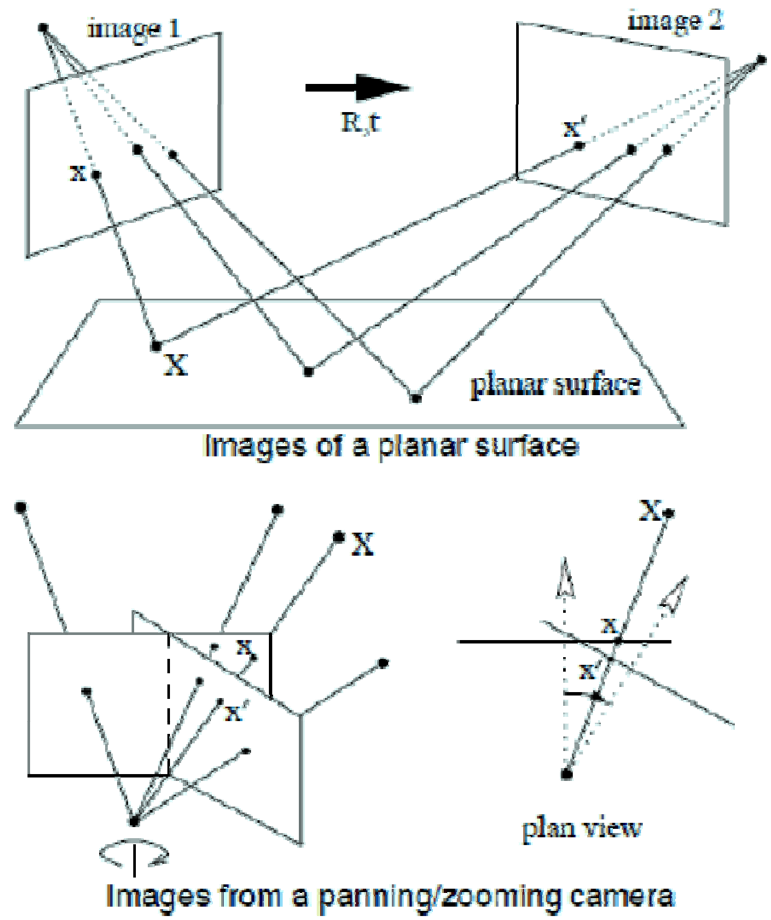
  (ii) Example-Based super-resolution (learning correspondence between low and high resolution image patches from a database).

Super-Resolution Image Enhancement

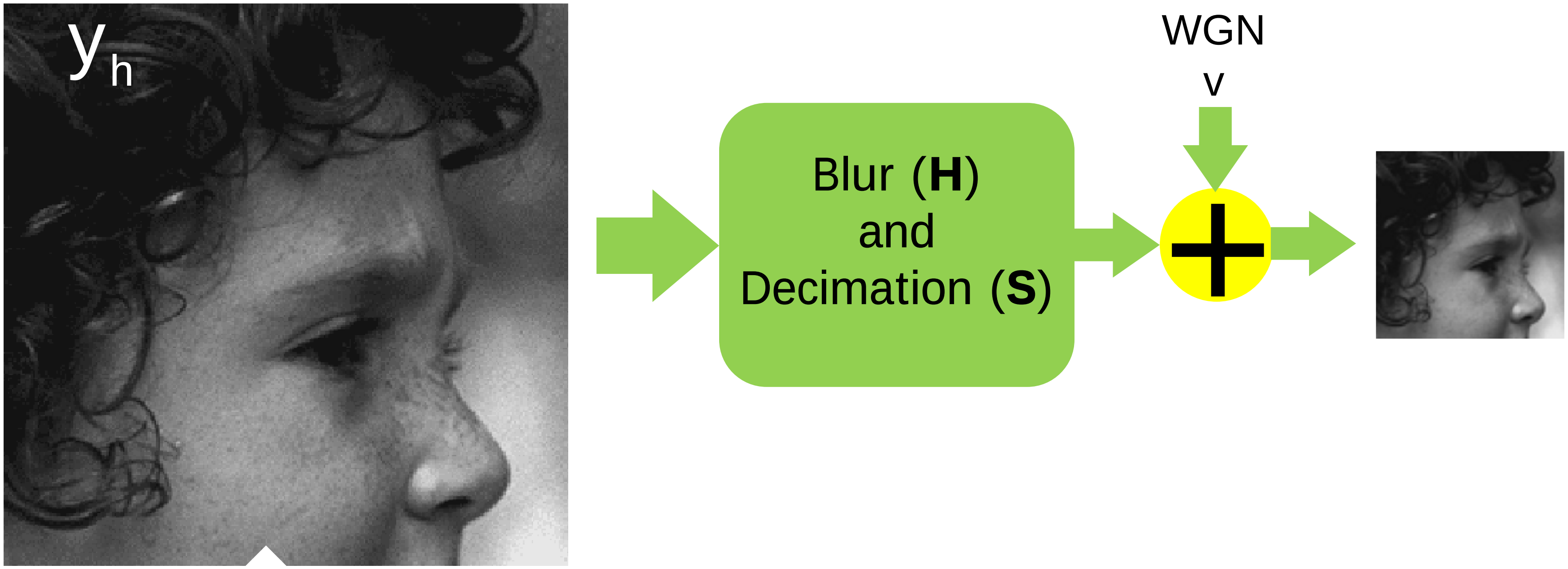


Super-Resolution Image Enhancement

- The multiple low resolution images can represent **different view-points of the same scene.**
- ***Image registration* deals with mapping corresponding points** in these images to the actual points in original scene and transforming data into one coordinate system. More on this later in BIA-2014.
- Several types of ***transformations*** could be required for registration of images like affine transformations, biquadratic transformations or planar homographic transformations.
- This alignment involves geometric component as well as photometric component – eg: brightness normalization / adjustment for constancy between images used for super-resolution, Blur adjustment, etc.



Single image Super-Resolution - this is a different problem!



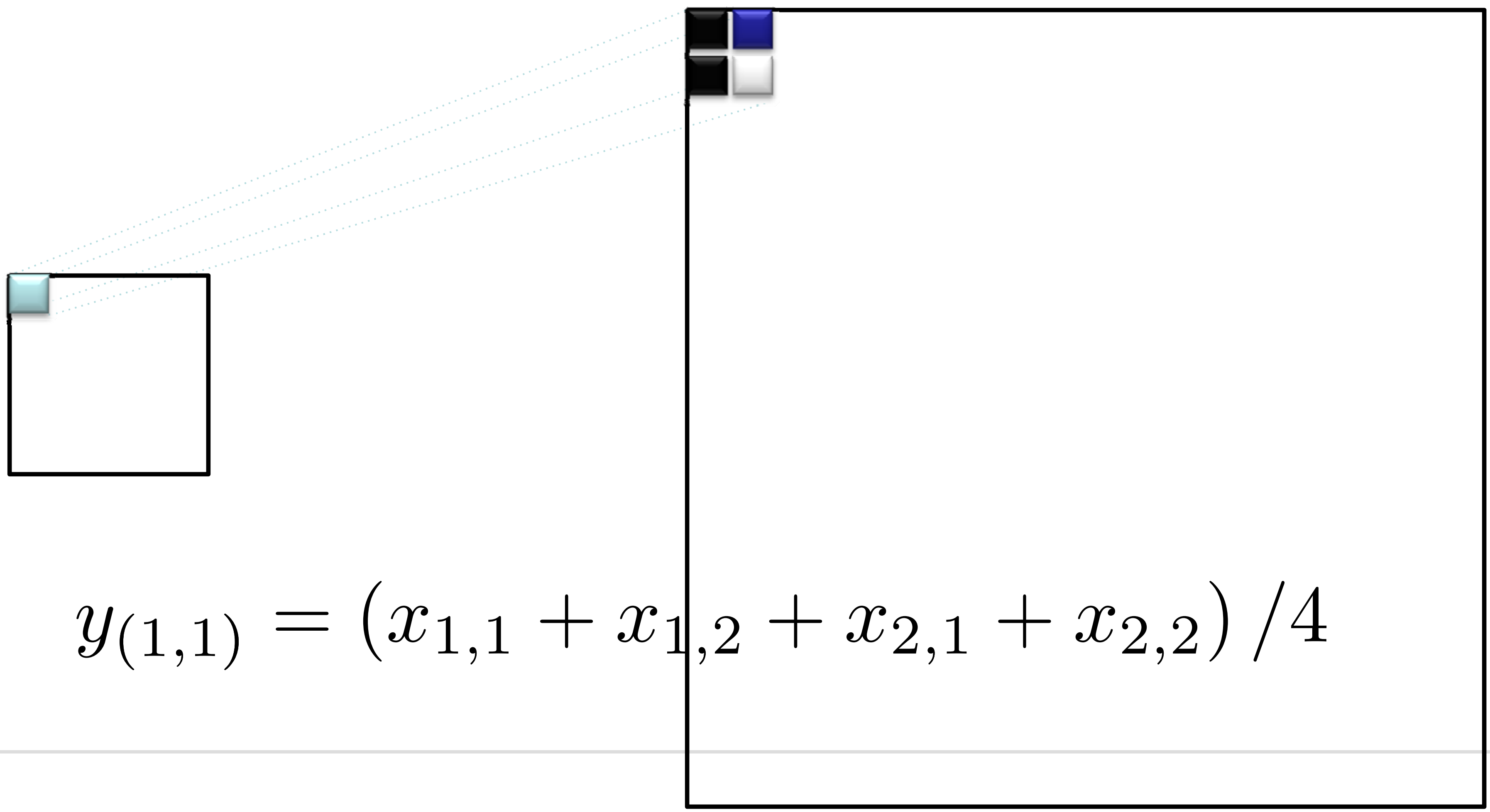
Our Task: Reverse the process – recover the high-resolution image from the low-resolution one .

Adapted from:
Image Super-Resolution Using Sparse Representation
By: Michael Elad, Technion University

Image super-resolution

$$y = \Phi x + e$$

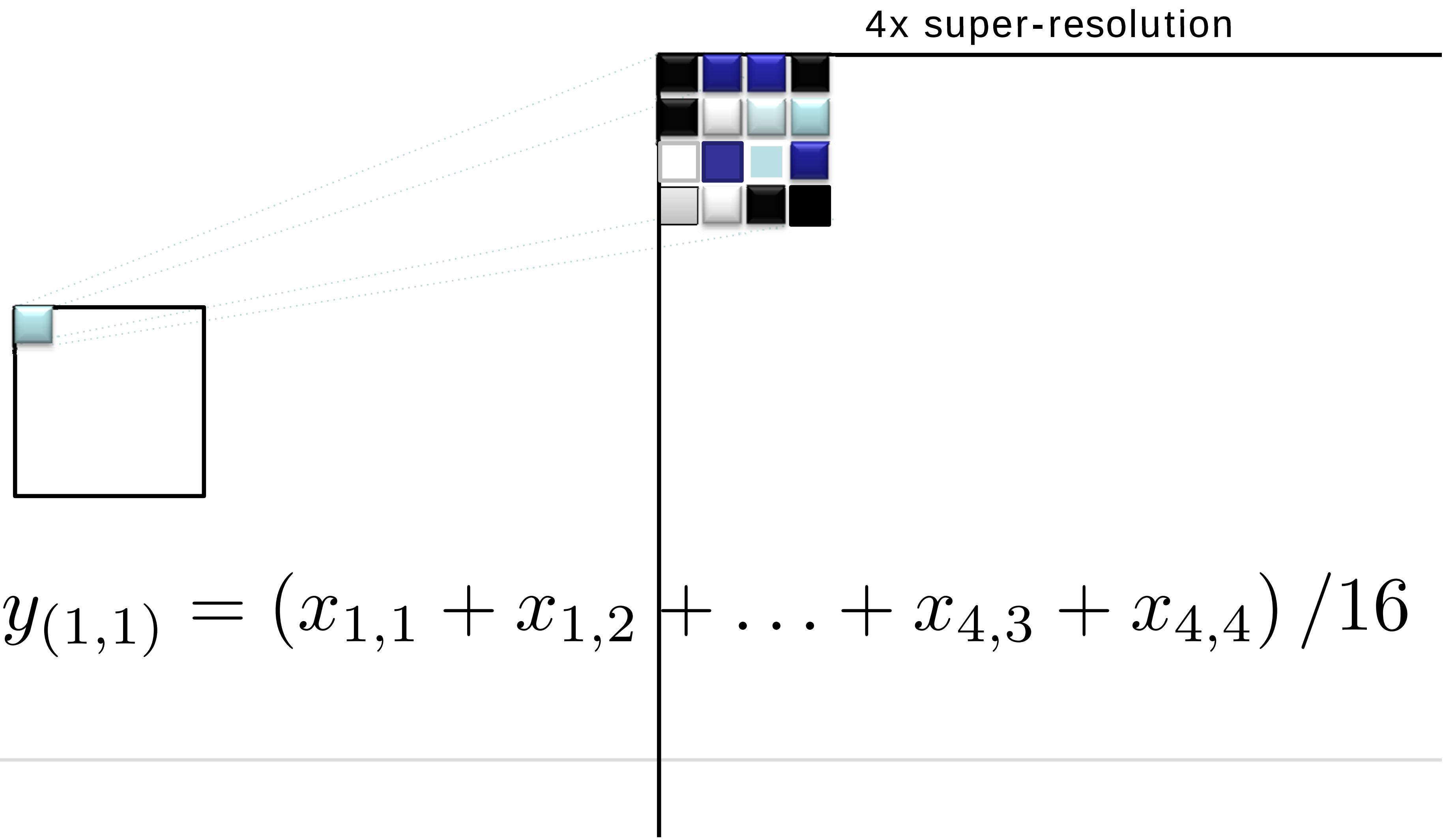
2x super-resolution



$$y_{(1,1)} = \left(x_{1,1} + x_{1,2} + x_{2,1} + x_{2,2}\right)/4$$

Image super-resolution

$$y = \Phi x + e$$



$$y_{(1,1)} = \left(x_{1,1} + x_{1,2} + \ldots + x_{4,3} + x_{4,4}\right)/16$$

Image super-resolution

$$y = \Phi x + e$$

Super-resolution factor *D*

Under-sampling factor $M/N = 1/D^2$

General rule:
The smaller the under-sampling, the more the unknowns and hence, the harder the super-resolution problem

Super-Resolution References

REFERENCES

[1] Capel, D.; Zisserman, A.; , "Computer vision applied to super resolution," *Signal Processing Magazine, IEEE , vol.20, no.3, pp. 75- 86, May 2003 doi:10.1109/MSP.2003.1203211 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1203211&isnumber=27099*

[2] Sung Cheol Park; Min Kyu Park; Moon Gi Kang; , "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE , vol.20, no.3, pp. 21- 36, May 2003 doi:10.1109/MSP.2003.1203207 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1203207&isnumber=2709*

[3] Glasner, D.; Bagon, S.; Irani, M.; , "Super-Resolution from a Single Image", *12th International Conference on Computer Vision (ICCV), IEEE , pp. 349 – 356, 2009* URL: http://www.wisdom.weizmann.ac.il/~vision/SingleImageSR.html

[4] Tsai, R.; Huang T.; , "Multi-frame image restoration and registration", *Advances in Computer Vision and Image Processing, vol. 1, no. 2, JAI Press Inc., Greenwich, CT, 1984, pp. 317–339*

OTHER RESOURCES

1. List of Super-Resolution publications
   http://decsai.ugr.es/~jmd/superresolution/publications.html [Date accessed: 31/03/2011]

2. Super-resolution SD-to-HD up-converter super-resolution (Avarex.ru)
http://www.youtube.com/watch?v=wxCleSGnji8 [Date accessed: 31/03/2011]

3. YUV Super Resolution vs plain static upscaling of video
http://www.youtube.com/watch?v=181c6DxDs6k [Date accessed: 31/03/2011]

4. SR Overview (Visual Geometry Group)
http://www.robots.ox.ac.uk/~vgg/research/SR/ [Date accessed: 31/03/2011]

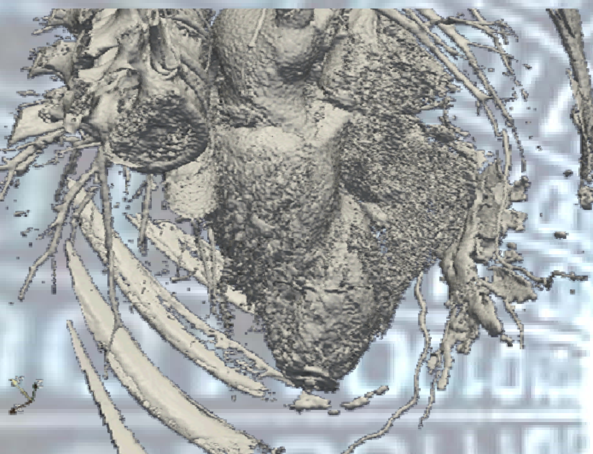
Scale invariance

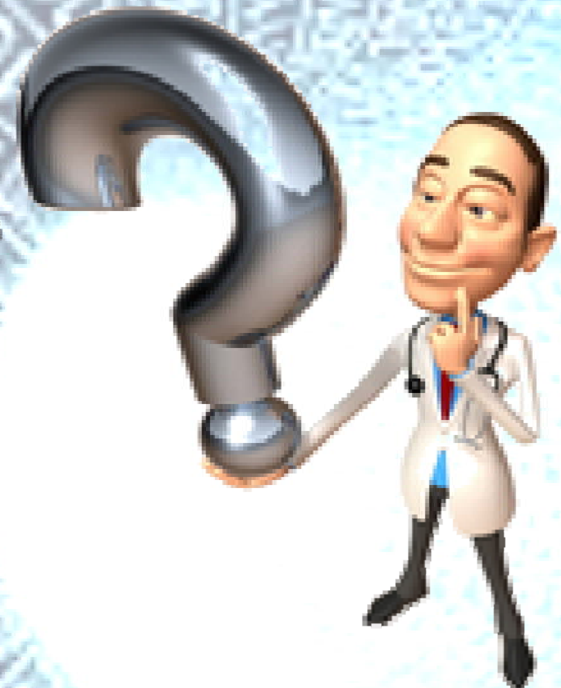
Requires a method to repeatably select points in location and scale:

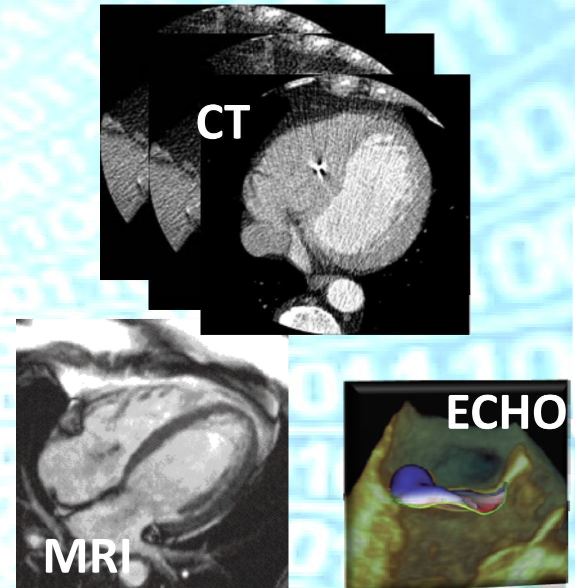
- The only reasonable scale-space kernel is a Gaussian (Koenderink, 1984; Lindeberg, 1994)
- An efficient choice is to detect peaks in the difference of Gaussian pyramid (Burt & Adelson, 1983; Crowley & Parker, 1984 – but examining more scales)
- Difference-of-Gaussian with constant ratio of scales is a close approximation to Lindeberg’s scale-normalized Laplacian (can be shown from the heat diffusion equation)

B I A 2 0 1 4

Carnegie Mellon University





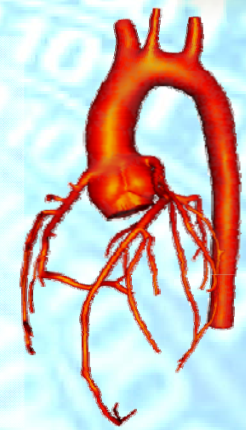


Prahlad G Menon, PhD

www.justcallharry.com

+1 412-259-3031

pgmenon@andrew.cmu.edu



Biomedical Imaging & Analysis

Lecture 6, Part 6. Fall 2014

Basic Image Processing / Filtering (VI)

*[**Text**: Ch: 10, Gonzalez and Woods, Digital Image Processing (3rd Edition) + Papers / Reading Assignments on Blackboard.]*

Prahlad G Menon, PhD

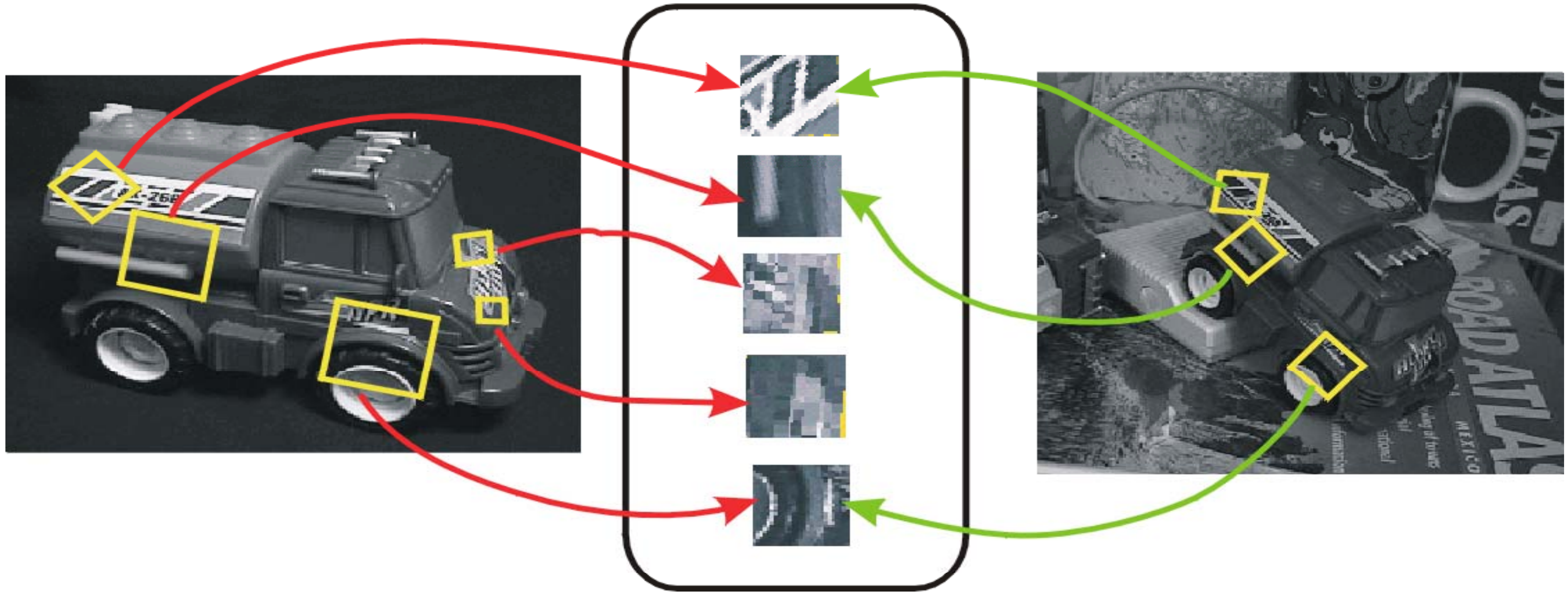
Assistant Professor

Sun Yat-sen University – Carnegie Mellon University (SYSU-CMU)

Joint Institute of Engineering

Recognition and Matching based on Local Invariant Features

- Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters



Advantages of invariant local features

- **Locality:** features are local, so robust to occlusion and clutter (no prior segmentation)
- **Distinctiveness:** individual features can be matched to a large database of objects
- **Quantity:** many features can be generated for even small objects
- **Efficiency:** close to real-time performance
- **Extensibility:** can easily be extended to wide range of differing feature types, with each adding robustness

Key point localization

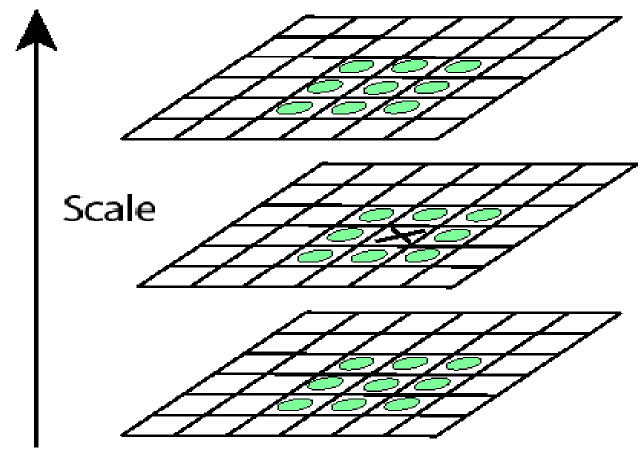
- Detect maxima and minima of difference-of-Gaussian in scale space.
- Fit a quadratic to surrounding values for sub-pixel and sub-scale interpolation (Brown & Lowe, 2002)
- Taylor expansion around point:

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^{\mathbf{T}}\frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x}$$

- Ratio of Principal Curvatures:

$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1}\frac{\partial D}{\partial \mathbf{x}}$$

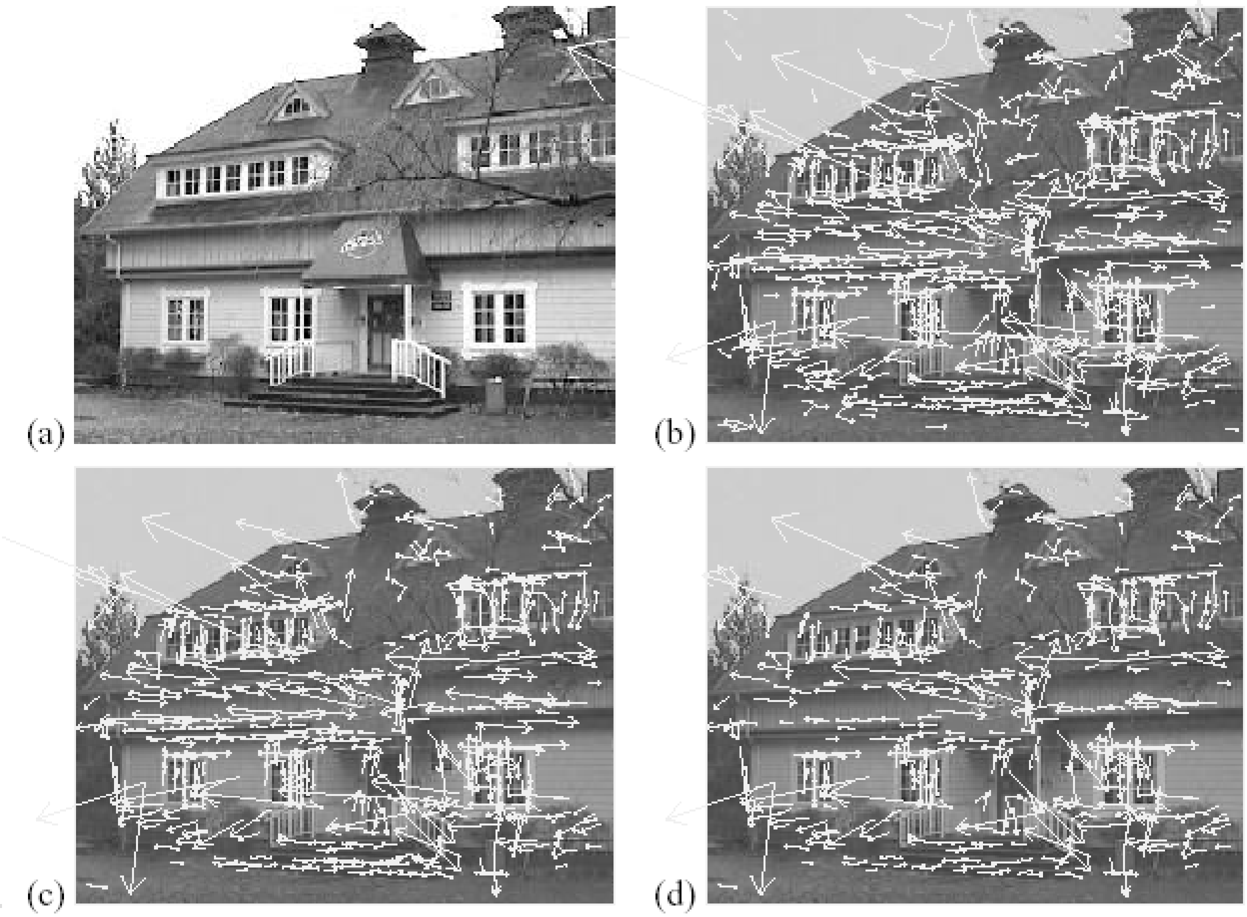
- The function value at the extremum, *D(^**x**)*, is useful for rejecting unstable extrema with low contrast (and therefore sensitive to noise). Eg: If *D(^x)* ***less than 0.03 , discard KeyPoint.***



Maxima and minima of the difference-of-Gaussian images are detected by comparing a pixel (marked with X) to its 26 neighbors in 3×3 regions at the current and adjacent scales (markedwith circles).

Example of Key-point detection

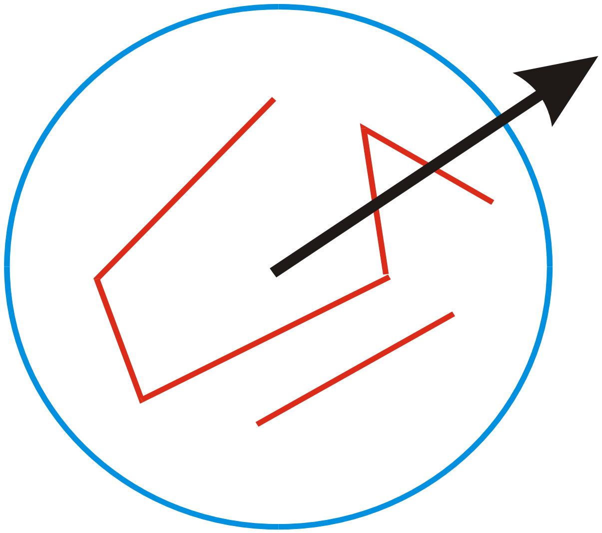
Threshold on value at DOG peak and on ratio of principle curvatures (Harris approach).

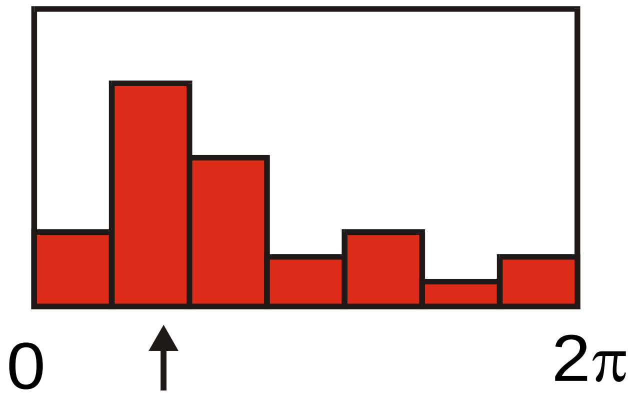


(a) 233x189 image
(b) 832 DOG extrema
(c) 729 left after peak value threshold
(d) 536 left after testing ratio of principle curvatures…

Select canonical **orientation**

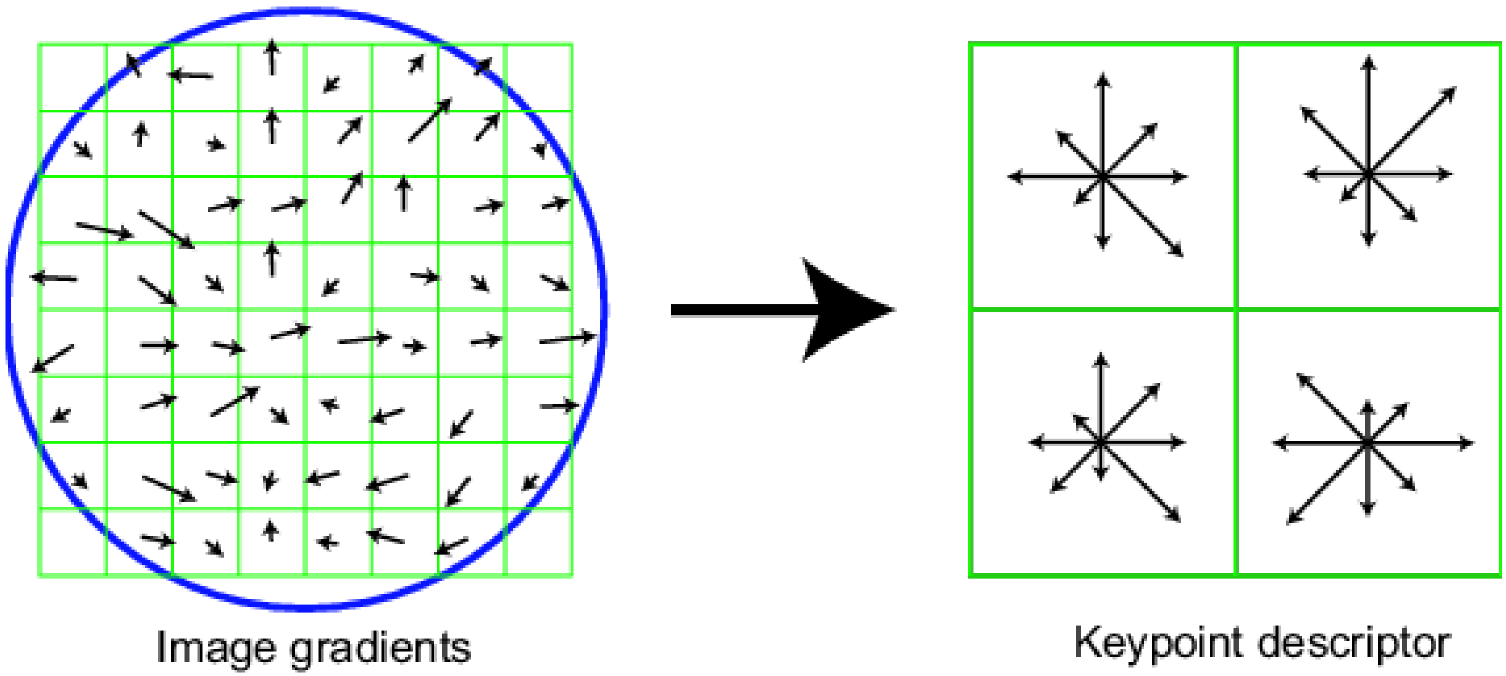
- By assigning a consistent orientation to each keypoint description based on local image properties, the keypoint descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation.
- Create histogram of local gradient directions computed at selected scale
- Assign canonical orientation at peak of smoothed histogram
- Each key specifies stable 2D coordinates (x, y, scale, orientation)





SIFT vector formation (example)

- Thresholded image gradients are sampled over 8x8 array of locations in scale space.
- Create array of orientation histograms.
- 8 orientations x 2x2 histogram array = 32 dimensions

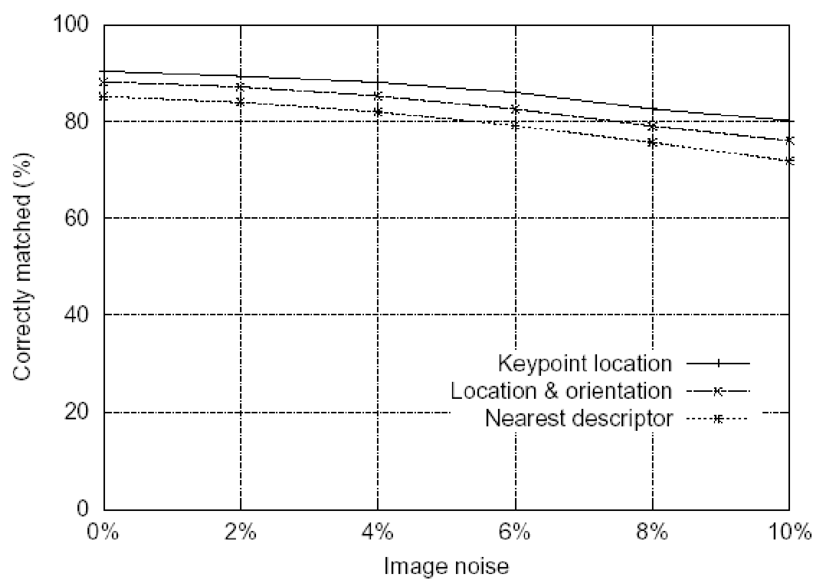


Feature Matching: **stability to noise**

- The best candidate match for each keypoint is found by identifying its nearest neighbor in the database of keypoints from training images. The nearest neighbor is defined as the keypoint with minimum Euclidean distance for the invariant descriptor vector **(READING ASSIGNMENT – Lowe 2004 – on Blackboard.**

Test case: Match features after random change in image scale & orientation, with differing levels of image noise.

***Example** application:* Find nearest neighbor in database of 30,000 features.



Example Application – finding hidden objects in images...

- To maximize the performance of object recognition for small or highly occluded objects – see Hough feature clustering in Lowe 2004. http://www.cs.toronto.edu/~fleet/courses/2503/fall11/Handouts/sift.pdf
- Check accuracy of detected KeyPoints using Affine Registration of one image onto the other (more in our Registration Lectures, later in the course)…

Figure 12: The training images for two objects are shown on the left. These can be recognized in a cluttered image with extensive occlusion, shown in the middle. The results of recognition are shown on the right. A parallelogram is drawn around each recognized object showing the boundaries of the original training image under the affi ne transformation solved for during recognition. Smaller squares indicate the keypoints that were used for recognition.

Good **SIFT features tutorials** / papers to learn from…

- http://www.cs.toronto.edu/~jepson/csc2503/tutSIFT04.pdf : By Estrada, Jepson, and Fleet.
- Reading assignment on **SIFT features for face recognition**, on Blackboard!
- ***Other references:***
  - Harris & Stephens (1988) http://www.bmva.org/bmvc/1988/avc-88-023.pdf
  - Lowe (2004) http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf
  - Mikolajczyk & Schmid (2005) http://lear.inrialpes.fr/pubs/2005/MS05/mikolajczyk_pami05.pdf
  - Brown & Lowe (2007) http://cvlab.epfl.ch/~brown/papers/ijcv2007.pdf