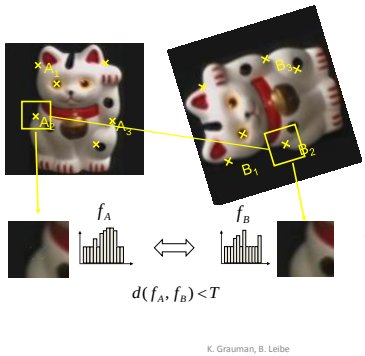
EEE422/6082 Computational Vision

Feature Detectors and Descriptors

Ling Shao

Many keypoint slides from Grauman&Leibe

Overview of Keypoint Matching



1. **Find a set of distinctive key-points**
2. **Define a region around each keypoint**
3. **Extract and normalize the region content**
4. **Compute a local descriptor from the normalized region**
5. **Match local descriptors**

K. Grauman, B. Leibe

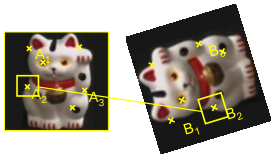
Main challenges

- Change in position and scale
- Change in viewpoint
- Occlusion
- Articulation

Goals for Keypoints

Detect points that are *repeatable* and *distinctive*

Key trade-offs



Localization

More Points
Robust to occlusion
Works with less texture

More Repeatable
Robust detection
Precise localization

Description

More Robust
Deal with expected variations
Maximize correct matches

More Selective
Minimize wrong matches

Choosing interest points

- If you wanted to meet a friend would you say
  a) "Let's meet on campus."
  b) "Let's meet on Green street."
  c) "Let's meet at Green and Wright."
  – Corner detection
- Or if you were in a secluded area:
  a) "Let's meet in the Plains of Akbar."
  b) "Let's meet on the side of Mt. Doom."
  c) "Let's meet on top of Mt. Doom."
  – Blob (valley/peak) detection

Choosing interest points

- Corners
  – "Let's meet at Green and Wright."
- Peaks/Valleys
  – "Let's meet on top of Mt. Doom."

Many Existing Detectors Available

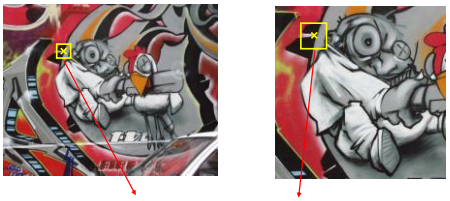
Hessian & Harris	[Beaudet '78], [Harris '88]
Laplacian, DoG	[Lindeberg '98], [Lowe 1999]
Harris-/Hessian-Laplace	[Mikolajczyk & Schmid '01]
Harris-/Hessian-Affine	[Mikolajczyk & Schmid '04]
EBR and IBR	[Tuytelaars & Van Gool '04]
MSER	[Matas '02]
Salient Regions	[Kadir & Brady '01]
Others…	

K. Grauman, B. Leibe

Harris Detector – Responses [Harris88]

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

Automatic Scale Selection

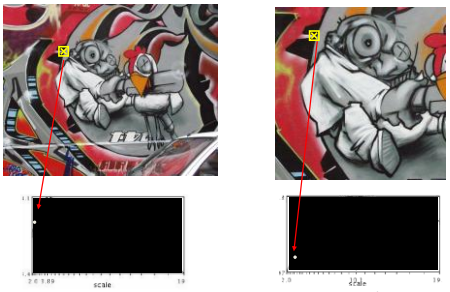


$$f(I_{i_1\ldots i_m}(x,\sigma)) = f(I_{i_1\ldots i_m}(x',\sigma'))$$

How to find corresponding patch sizes?

Automatic Scale Selection

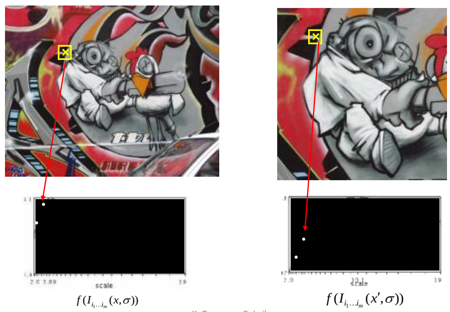
- Function responses for increasing scale (scale signature)



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

Automatic Scale Selection

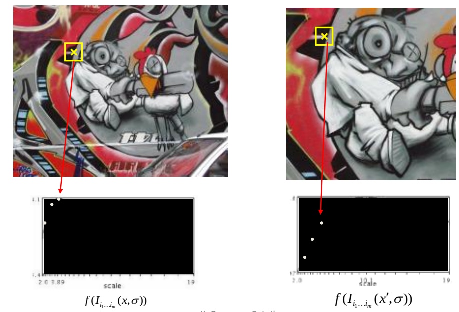
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

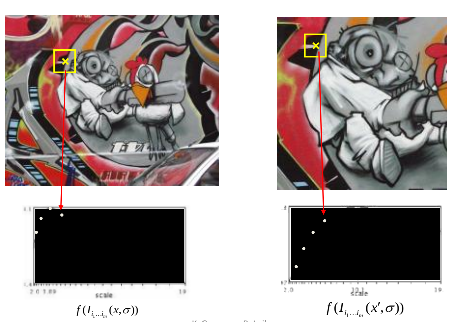
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

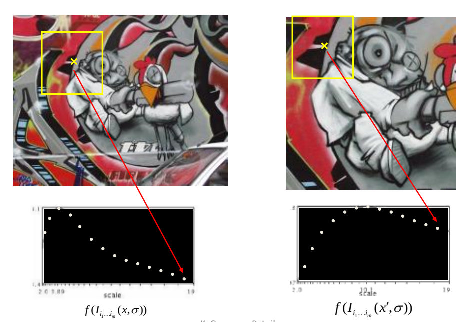
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

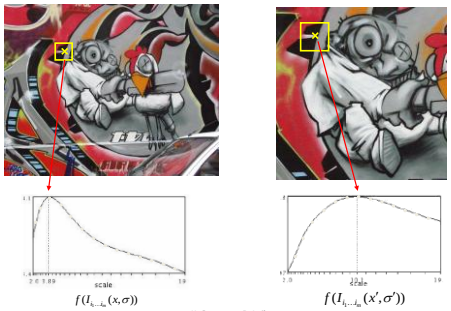
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

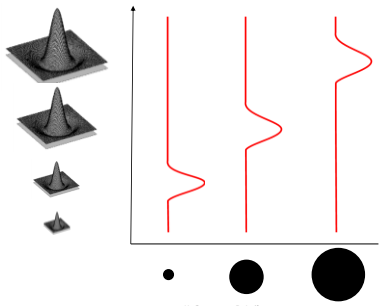
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

What Is A Useful Signature Function?

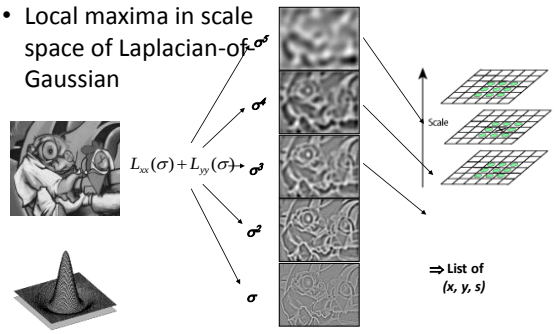
- Laplacian-of-Gaussian = "blob" detector



K. Grauman, B. Leibe

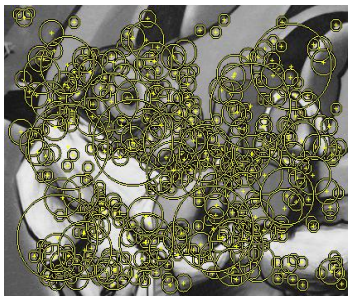
Laplacian-of-Gaussian (LoG)

- Local maxima in scale space of Laplacian-of-Gaussian



K. Grauman, B. Leibe

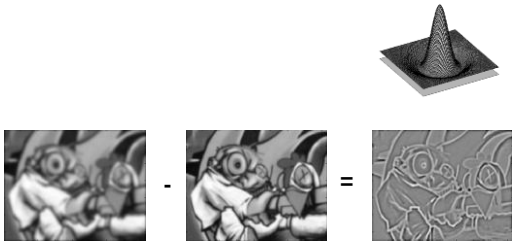
Results: Laplacian-of-Gaussian



K. Grauman, B. Leibe

Difference-of-Gaussian (DoG)

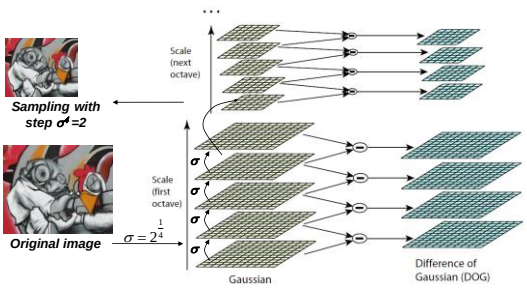
- Difference of Gaussians as approximation of the Laplacian-of-Gaussian



K. Grauman, B. Leibe

DoG – Efficient Computation

- Computation in Gaussian scale pyramid



K. Grauman, B. Leibe

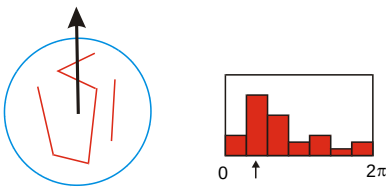
Results: Lowe's DoG

K. Grauman, B. Leibe

Orientation Normalization

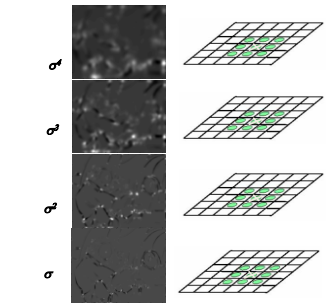
- Compute orientation histogram
- Select dominant orientation
- Normalize: rotate to fixed orientation

[Lowe, SIFT, 1999]



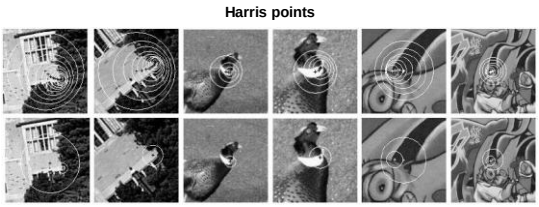
Harris-Laplace [Mikolajczyk '01]

1. Initialization: Multiscale Harris corner detection



Harris-Laplace [Mikolajczyk '01]

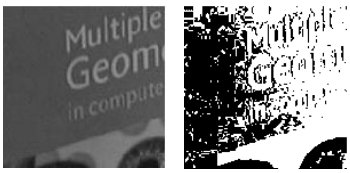
1. Initialization: Multiscale Harris corner detection
2. Scale selection based on Laplacian
   (same procedure with Hessian ⇒ Hessian-Laplace)



K. Grauman, B. Leibe

Maximally Stable Extremal Regions [Matas '02]

- Based on Watershed segmentation algorithm
- Select regions that stay stable over a large parameter range



K. Grauman, B. Leibe

Example Results: MSER

30

Available at a web site near you…

- For most local feature detectors, executables are available online:
  - http://robots.ox.ac.uk/~vgg/research/affine
  - http://www.cs.ubc.ca/~lowe/keypoints/
  - http://www.vision.ee.ethz.ch/~surf

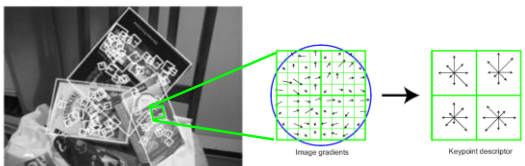
K. Grauman, B. Leibe

Local Descriptors

- The ideal descriptor should be
  - Robust
  - Distinctive
  - Compact
  - Efficient
- Most available descriptors focus on edge/gradient information
  - Capture texture information
  - Color rarely used

K. Grauman, B. Leibe

Local Descriptors: SIFT Descriptor



Histogram of oriented gradients

- **Captures important texture information**
- **Robust to small translations / affine deformations**

[Lowe, ICCV 1999]

K. Grauman, B. Leibe

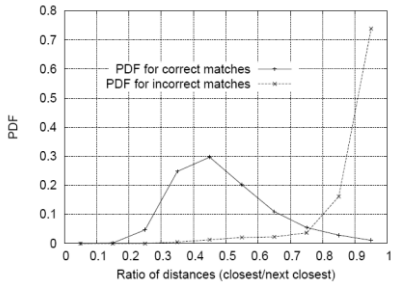
Details of Lowe's SIFT algorithm

- Run DoG detector
  - Find maxima in location/scale space
  - Remove edge points
- Find all major orientations
  - Bin orientations into 36 bin histogram
    - Weight by gradient magnitude
    - Weight by distance to center (Gaussian-weighted mean)
- For each (x,y,scale,orientation), create descriptor:
  - Sample 16x16 gradient mag. and rel. orientation
  - Bin 4x4 samples into 4x4 histograms
  - Threshold values to max of 0.2, divide by L2 norm
  - Final descriptor: 4x4x8 normalized histograms

Lowe IJCV 2004

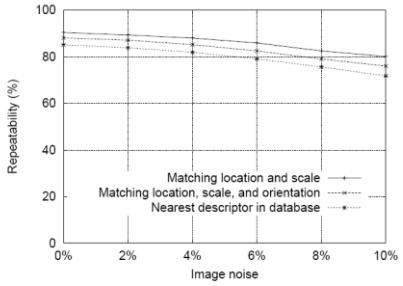
Matching SIFT Descriptors

- Nearest neighbor (Euclidean distance)
- Threshold ratio of nearest to 2nd nearest descriptor



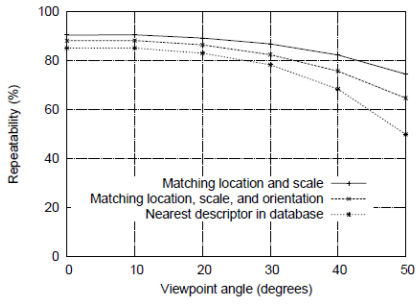
Lowe IJCV 2004

SIFT Repeatability

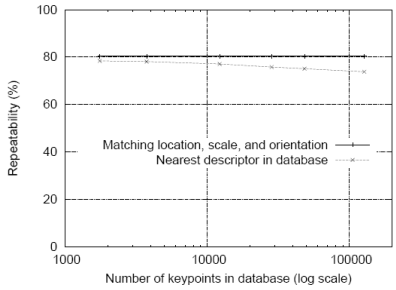


Lowe IJCV 2004

SIFT Repeatability

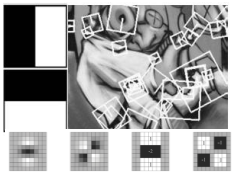


SIFT Repeatability



Lowe IJCV 2004

Local Descriptors: SURF



Fast approximation of SIFT idea

Efficient computation by 2D box filters & integral images
⇒ 6 times faster than SIFT
Equivalent quality for object identification

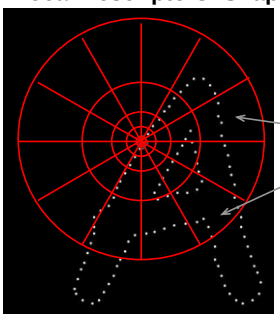
GPU implementation available

Feature extraction @ 200Hz (detector + descriptor, 640×480 img)
http://www.vision.ee.ethz.ch/~surf

[Bay, ECCV'06], [Cornelis, CVGPU'08]

K. Grauman, B. Leibe

Local Descriptors: Shape Context



Count the number of points inside each bin, e.g.:

Count = 4
⋮
Count = 10

Log-polar binning: more precision for nearby points, more flexibility for farther points.

Belongie & Malik, ICCV 2001

K. Grauman, B. Leibe

Choosing a detector

- What do you want it for?
  - Precise localization in x-y: Harris
  - Good localization in scale: Difference of Gaussian
  - Flexible region shape: MSER
- Best choice often application dependent
  - Harris-/Hessian-Laplace/DoG work well for many natural categories
  - MSER works well for buildings and printed things
- Why choose?
  - Get more points with more detectors
- There have been extensive evaluations/comparisons
  - [Mikolajczyk et al., IJCV'05, PAMI'05]
  - All detectors/descriptors shown here work well

Comparison of Keypoint Detectors

Table 7.1 Overview of feature detectors.

Feature Detector	Corner	Blob	Region	Rotation invariant	Scale invariant	Affine invariant	Localization Repeatability	Localization accuracy	Robustness	Efficiency
Harris	√			√			+++	+++	+++	++
Hessian		√		√			++	++	++	+
SUSAN	√			√			++	++	++	+++
Harris-Laplace	√	(√)		√	√		+++	+++	++	+
Hessian-Laplace	(√)	√		√	√		+++	+++	+++	+
DoG	(√)	√		√	√		++	++	++	++
SURF	(√)	√		√	√		++	++	++	+++
Harris-Affine	√	(√)		√	√	√	+++	+++	++	++
Hessian-Affine	(√)	√		√	√	√	+++	+++	+++	++
Salient Regions	(√)	√		√	√	(√)	+	+	++	+
Edge-based	√			√	√	√	+++	+++	+	+
MSER			√	√	√	√	+++	+++	++	+++
Intensity-based			√	√	√	√	++	++	++	++
Superpixels			√	√	(√)	(√)	+	+	+	+

Tuytelaars Mikolajczyk 2008

Choosing a descriptor

- Again, need not stick to one
- For object instance recognition or stitching, SIFT or variant is a good choice

Things to remember

- Keypoint detection: repeatable and distinctive
  - Corners, blobs, stable regions
  - Harris, DoG

- Descriptors: robust and selective
  - spatial histograms of orientation
  - SIFT