Local invariant features part 2

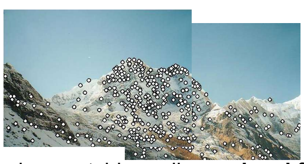
Thurs Mar 1

Kristen Grauman
UT Austin

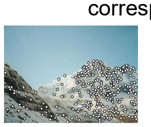
Announcements

- Reminder - Midterm next Thursday Mar 8
  - Closed book
  - One 8.5x11" sheet of notes allowed (both sides)
  - Content up to and including Thurs 3/1 lecture
- Practice midterm handout today

Important tool for multiple views: Local features



Multi-view matching relies on **local feature** correspondences.



How to *detect* which local features to match?
How to *describe* the features we detect?

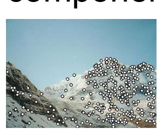
Slide credit: Kristen Grauman

Outline

- **Last time**: Interest point detection
  - Harris corner detector
- **Today**:
  - Laplacian of Gaussian, automatic scale selection
  - Local descriptors for image patches
  - Matching sets of features

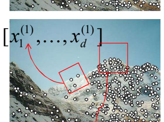
Local features: main components

1) Detection: Identify the interest points



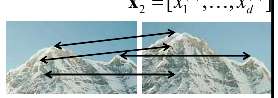
2) Description: Extract vector feature descriptor surrounding each interest point.

$\mathbf{x}_1 = [x_1^{(1)}, \ldots, x_d^{(1)}]$



$\mathbf{x}_2 = [x_1^{(2)}, \ldots, x_d^{(2)}]$

3) Matching: Determine correspondence between descriptors in two views



Slide credit: Kristen Grauman

Recall: Harris corner detector

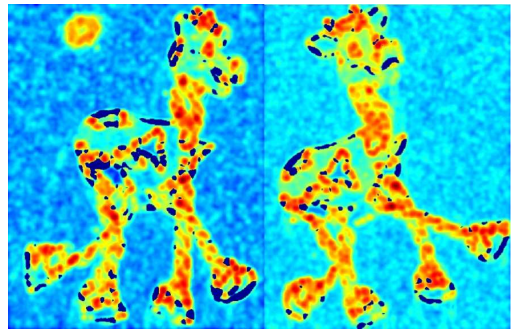
$$M = \sum w(x,y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

1) Compute *M* matrix for each image window to get their *cornerness* scores.
2) Find points whose surrounding window gave large corner response (*f*> threshold)
3) Take the points of local maxima, i.e., perform non-maximum suppression

Recall: Harris Detector: Steps

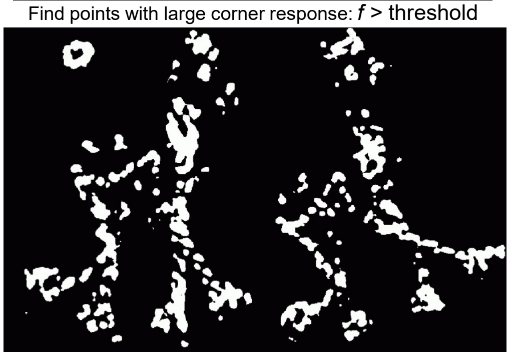
Recall: Harris Detector: Steps

Compute corner response *f*



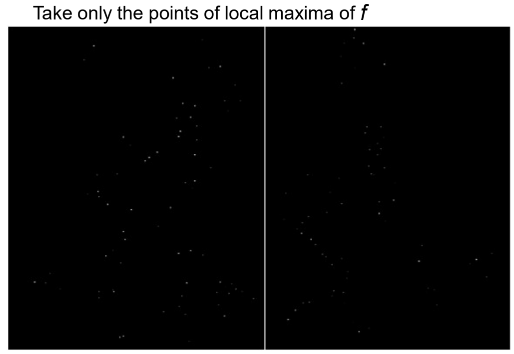
Recall: Harris Detector: Steps

Find points with large corner response: *f* > threshold

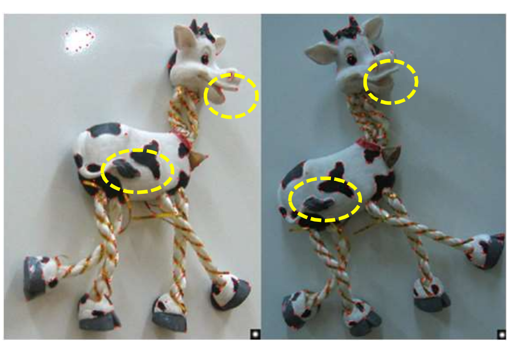


Recall: Harris Detector: Steps

Take only the points of local maxima of *f*



Recall: Harris Detector: Steps



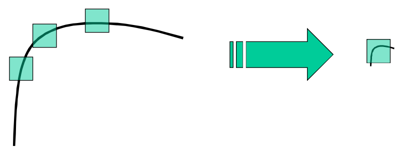
Review questions

- What properties should an interest operator have?
- What will determine how many interest points a given image has?
- What changes for Harris corner detections for in-plane or out-of-plane image rotations?

Properties of the Harris corner detector

Rotation invariant? Yes

Scale invariant? No



All points will be classified as edges

Corner !

Scale invariant interest points

How can we independently select interest points in each image, such that the detections are repeatable across different scales?

Slide credit: Kristen Grauman

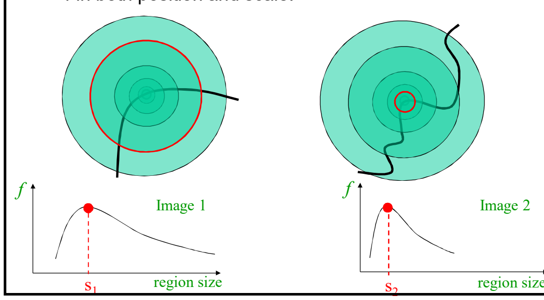
Automatic Scale Selection

How to find corresponding patch sizes, with only one image in hand?

K. Grauman, B. Leibe

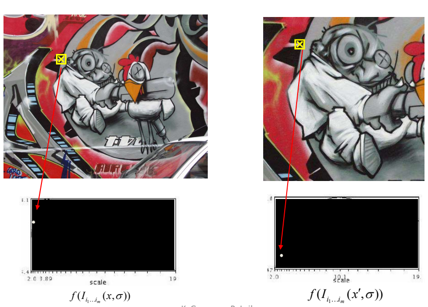
Automatic scale selection

Intuition:
- Find scale that gives local maxima of some function *f* in both position and scale.



Automatic Scale Selection

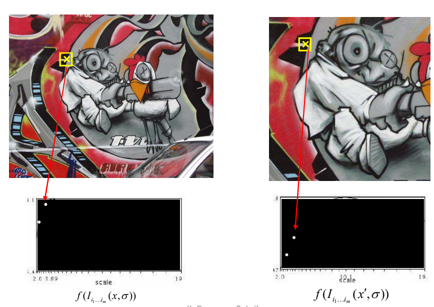
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

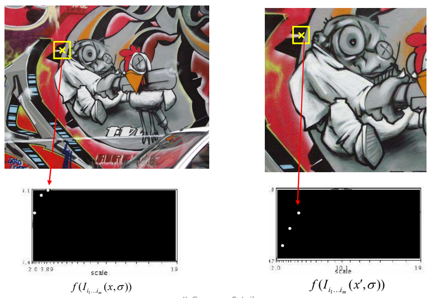
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

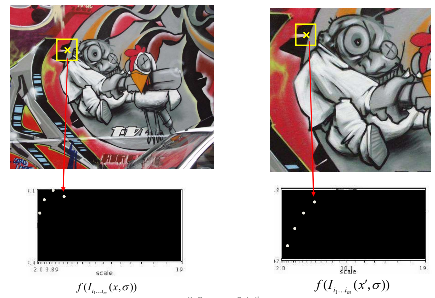
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

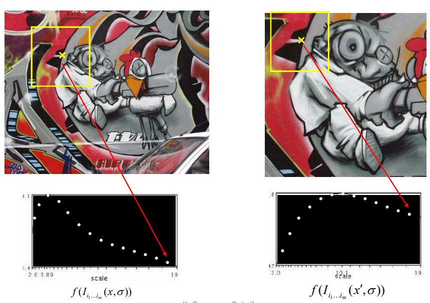
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

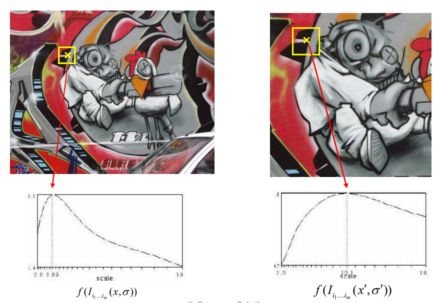
- Function responses for increasing scale (scale signature)



K. Grauman, B. Leibe

Automatic Scale Selection

- Function responses for increasing scale (scale signature)

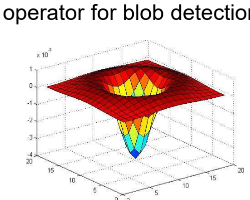


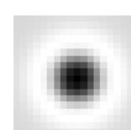
K. Grauman, B. Leibe

What can be the “signature” function?

Blob detection in 2D

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

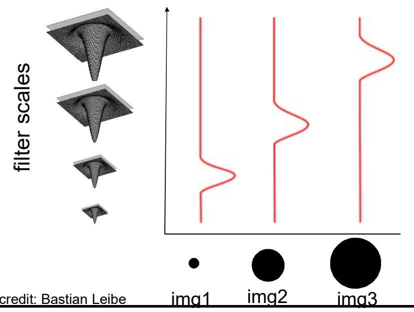




$$\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$$

Blob detection in 2D: scale selection

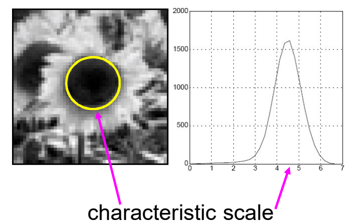
Laplacian-of-Gaussian = "blob" detector $\nabla^2 g = \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2}$



Slide credit: Bastian Leibe

Blob detection in 2D

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



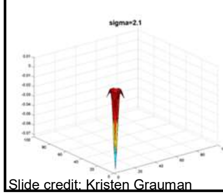
Slide credit: Lana Lazebnik

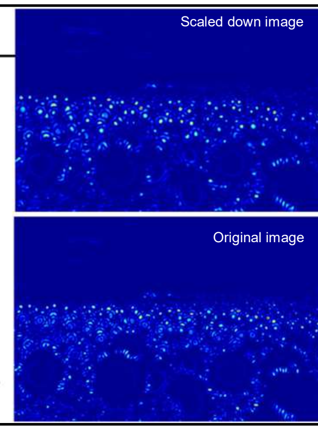
Example

Original image at ¾ the size

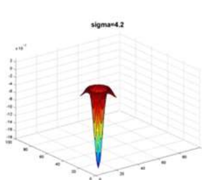
Slide credit: Kristen Grauman

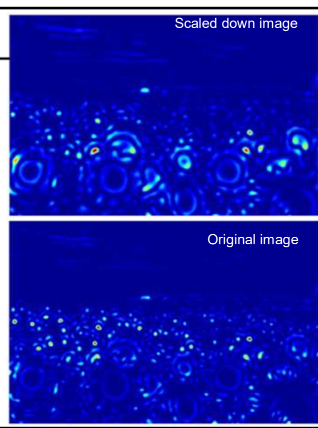
Original image at ¾ the size



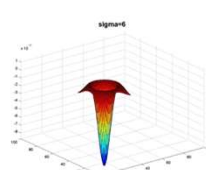


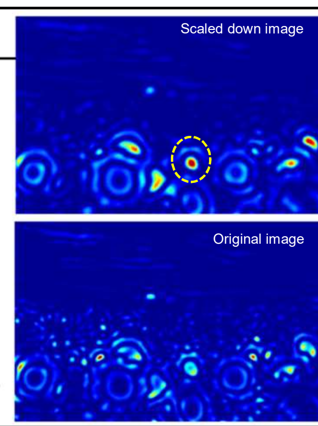
Slide credit: Kristen Grauman



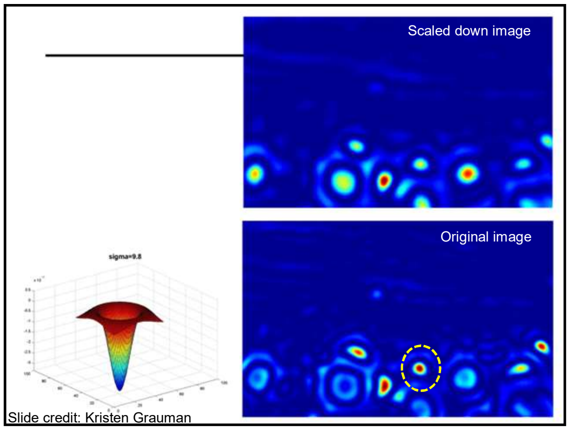


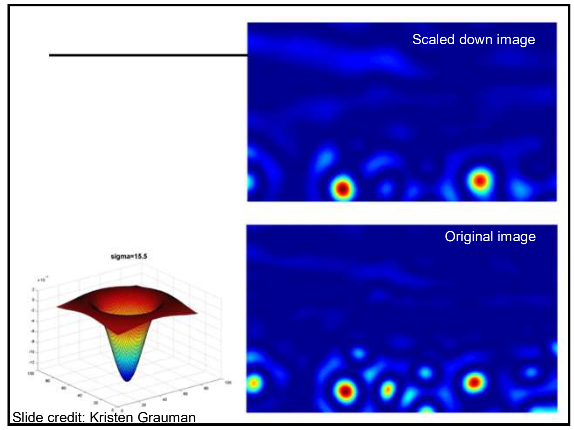
Slide credit: Kristen Grauman

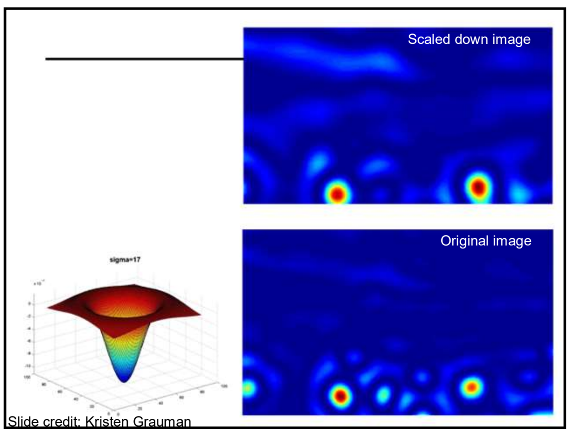


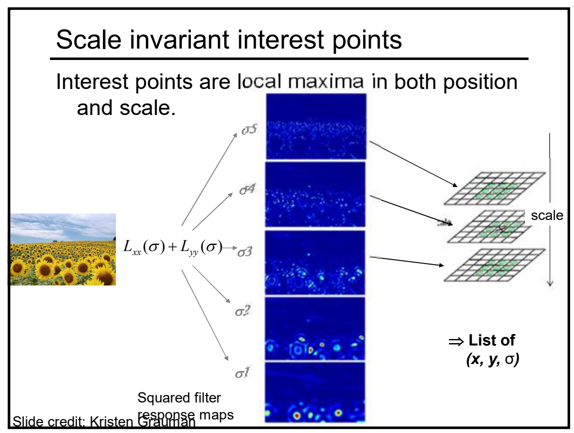


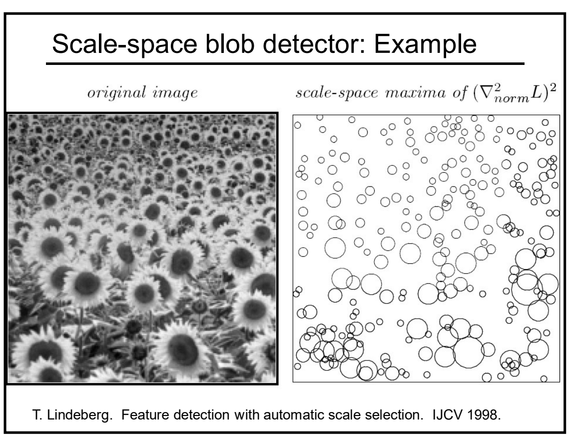
Slide credit: Kristen Grauman

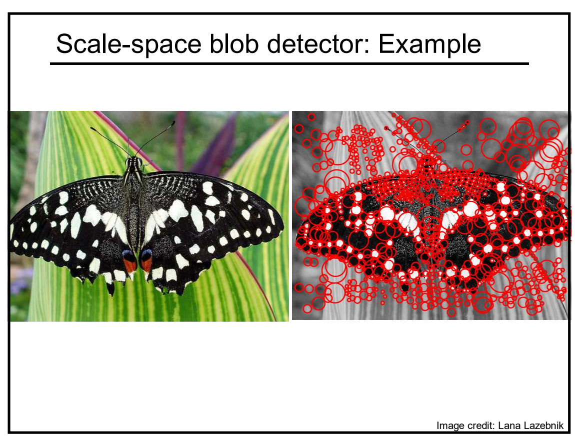

Scaled down image
Original image
sigma=9.8
Slide credit: Kristen Grauman


Scaled down image
Original image
sigma=15.5
Slide credit: Kristen Grauman


Scaled down image
Original image
sigma=17
Slide credit: Kristen Grauman


Scale invariant interest points
Interest points are local maxima in both position and scale.
$L_{xx}(\sigma) + L_{yy}(\sigma)$
σ5
σ4
σ3
σ2
σ1
scale
⇒ List of (x, y, σ)
Squared filter response maps
Slide credit: Kristen Grauman


Scale-space blob detector: Example
original image
scale-space maxima of $(\nabla^2_{norm} L)^2$
T. Lindeberg. Feature detection with automatic scale selection. IJCV 1998.

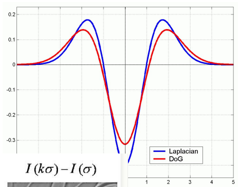

Scale-space blob detector: Example
Image credit: Lana Lazebnik

Technical detail

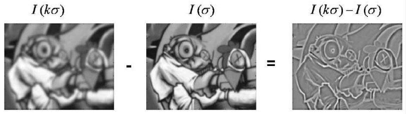
We can approximate the Laplacian with a difference of Gaussians; more efficient to implement.

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

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



$I(k\sigma)$ - $I(\sigma)$ = $I(k\sigma) - I(\sigma)$



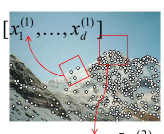
Summary so far

- Desirable properties for local features for correspondence
- Basic matching pipeline
- Interest point detection
  - Harris corner detector
  - Laplacian of Gaussian, automatic scale selection

Local features: main components

1) Detection: Identify the interest points

2) Description: Extract vector feature descriptor surrounding each interest point.



3) Matching: Determine correspondence between descriptors in two views

Slide credit: Kristen Grauman

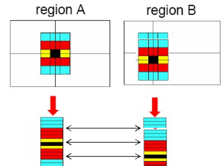
Geometric transformations

e.g. scale, translation, rotation

Photometric transformations

Figure from T. Tuytelaars ECCV 2006 tutorial

Raw patches as local descriptors



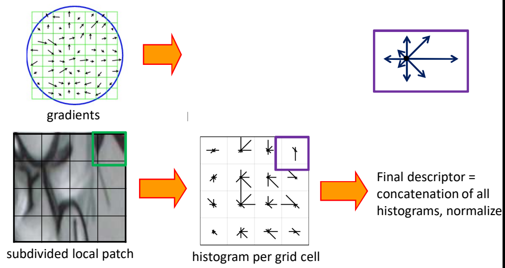
The simplest way to describe the neighborhood around an interest point is to write down the list of intensities to form a feature vector.

But this is very sensitive to even small shifts, rotations.

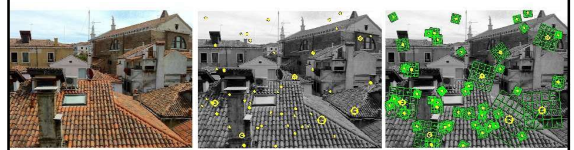
Figure: Andrew Zisserman

Scale Invariant Feature Transform (SIFT) descriptor [Lowe 2004]

- Use histograms to bin pixels within sub-patches according to their orientation.



Scale Invariant Feature Transform (SIFT) descriptor [Lowe 2004]

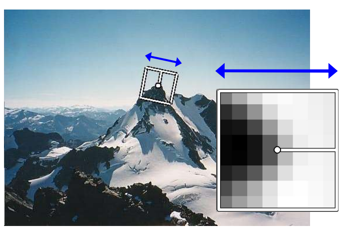


Interest points and their scales and orientations (random subset of 50)

SIFT descriptors

http://www.vlfeat.org/overview/sift.html

Making descriptor rotation invariant

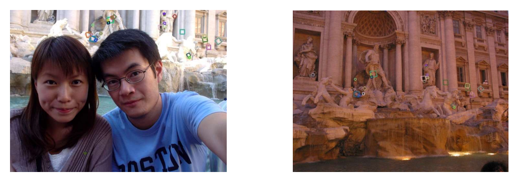


- Rotate patch according to its dominant gradient orientation
- This puts the patches into a canonical orientation.

Image from Matthew Brown

SIFT descriptor [Lowe 2004]

- Extraordinarily robust matching technique
  - Can handle changes in viewpoint
    - Up to about 60 degree out of plane rotation
  - Can handle significant changes in illumination
    - Sometimes even day vs. night (below)
  - Fast and efficient—can run in real time
  - Lots of code available, e.g. `http://www.vlfeat.org/overview/sift.html`



Steve Seitz

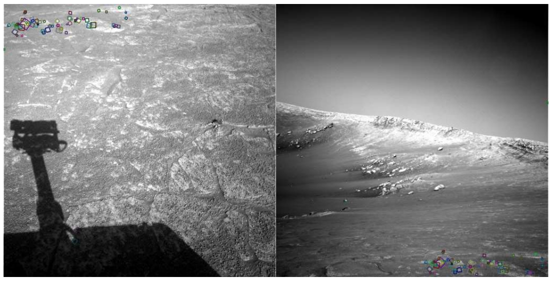
SIFT properties

- Invariant to
  - Scale
  - Rotation
- Partially invariant to
  - Illumination changes
  - Camera viewpoint
  - Occlusion, clutter

Example

NASA Mars Rover images

Example



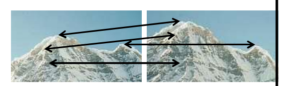
NASA Mars Rover images
with SIFT feature matches
Figure by Noah Snavely

Local features: main components

1) Detection: Identify the interest points

2) Description: Extract vector feature descriptor surrounding each interest point.

3) Matching: Determine correspondence between descriptors in two views



Slide credit: Kristen Grauman

Matching local features

Slide credit: Kristen Grauman

Matching local features

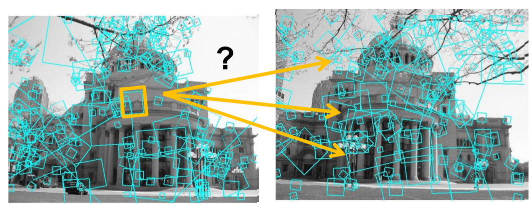


Image 1 Image 2

To generate **candidate matches**, find patches that have the most similar appearance (e.g., lowest SSD)

Simplest approach: compare them all, take the closest (or closest k, or within a thresholded distance)

Slide credit: Kristen Grauman

Ambiguous matches

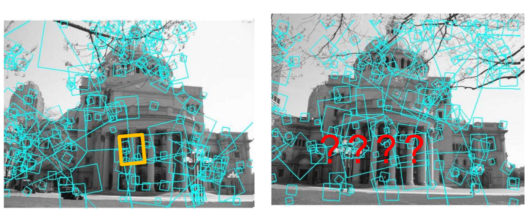


Image 1 Image 2

At what SSD value do we have a good match?

To add robustness to matching, consider **ratio** :

dist to best match / dist to second best match

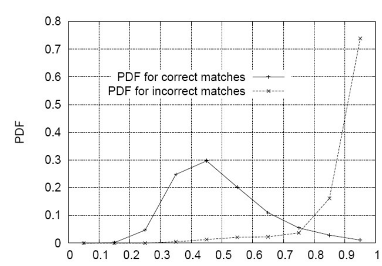
If low, first match looks good.

If high, could be ambiguous match.

Slide credit: Kristen Grauman

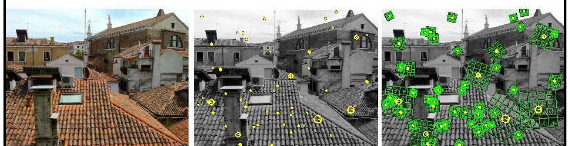
Matching SIFT Descriptors

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



Lowe IJCV 2004

Scale Invariant Feature Transform (SIFT) descriptor [Lowe 2004]



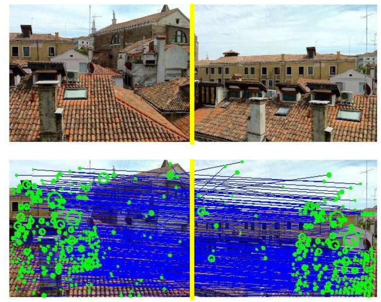
Interest points and their scales and orientations (random subset of 50)

SIFT descriptors

http://www.vlfeat.org/overview/sift.html

SIFT (preliminary) matches

img1 img2



img1 img2

http://www.vlfeat.org/overview/sift.html

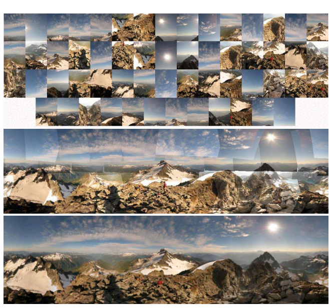
Value of local (invariant) features

- Complexity reduction via selection of distinctive points
- Describe images, objects, parts without requiring segmentation
  - Local character means robustness to clutter, occlusion
- Robustness: similar descriptors in spite of noise, blur, etc.

Applications of local invariant features

- Wide baseline stereo
- Motion tracking
- Panoramas
- Mobile robot navigation
- 3D reconstruction
- Recognition
- …

Automatic mosaicing

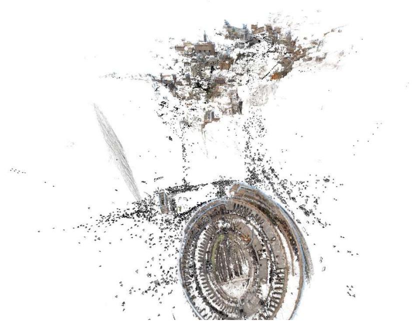


Matthew Brown
http://matthewalunbrown.com/autostitch/autostitch.html

Wide baseline stereo

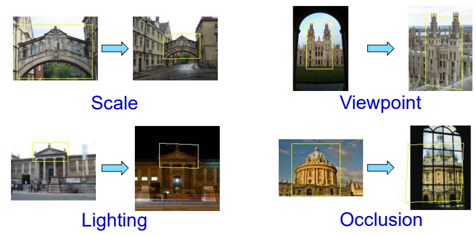
[Image from T. Tuytelaars ECCV 2006 tutorial]

Photo tourism [Snavely et al.]



Slide credit: Noah Snavely

Recognition of specific objects, scenes



Slide credit: J. Sivic

Google Goggles

Summary

- Interest point detection
  - Harris corner detector
  - Laplacian of Gaussian, automatic scale selection
- Invariant descriptors
  - Rotation according to dominant gradient direction
  - Histograms for robustness to small shifts and translations (SIFT descriptor)

Coming up

Additional questions we need to address to achieve these applications:

- Fitting a parametric transformation given putative matches
- Dealing with outlier correspondences
- Exploiting geometry to restrict locations of possible matches
- Triangulation, reconstruction
- Efficiency when indexing so many keypoints