CS4670: Computer Vision

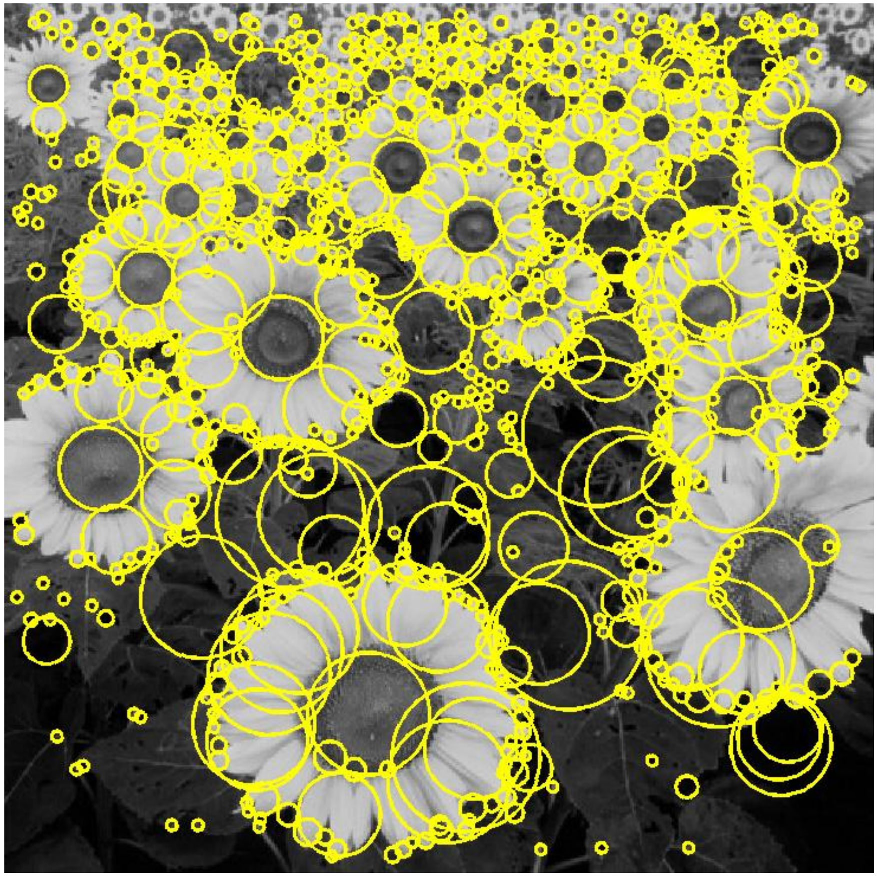
Noah Snavely

Lecture 6: Feature matching and alignment

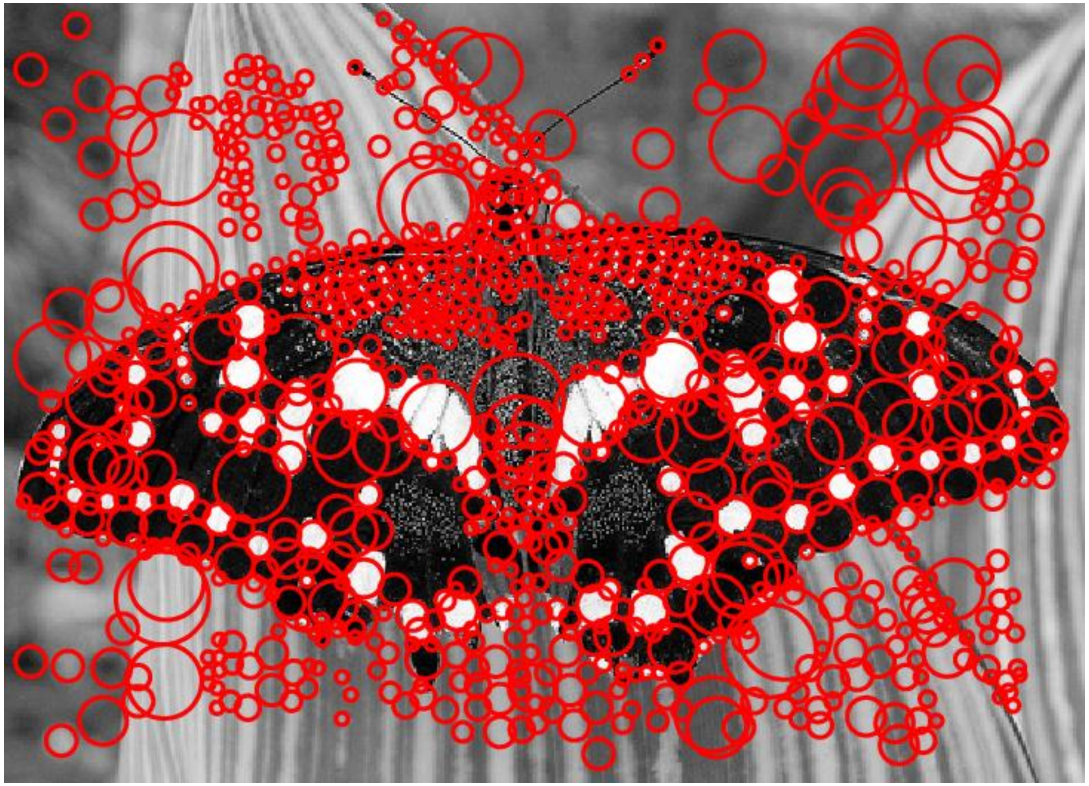
Reading

- Szeliski: Chapter 6.1

Last time: Corners and blobs



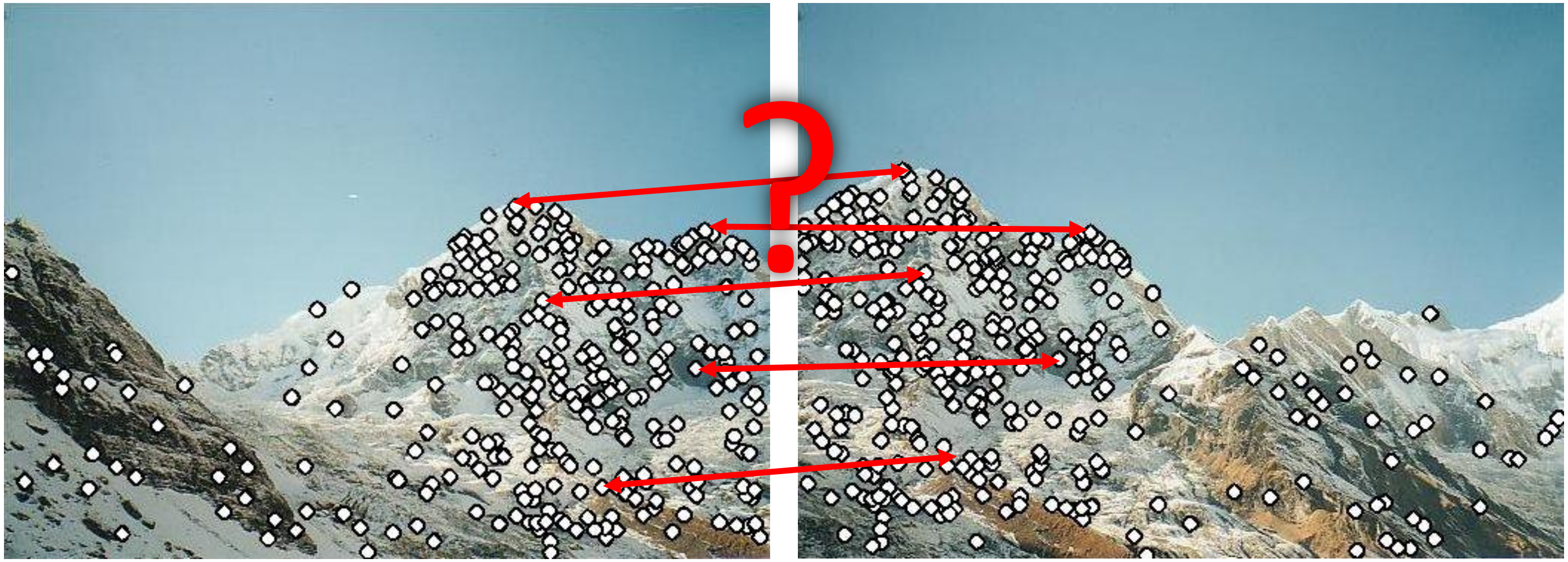
Scale-space blob detector: Example



Feature descriptors

We know how to detect good points

Next question: **How to match them?**



Answer: Come up with a *descriptor* for each point, find similar descriptors between the two images

How to achieve invariance

Need both of the following:

1. Make sure your detector is invariant
2. Design an invariant feature descriptor
   - Simplest descriptor: a single 0
     - What's this invariant to?
   - Next simplest descriptor: a square window of pixels
     - What's this invariant to?
   - Let's look at some better approaches…

Rotation invariance for feature descriptors

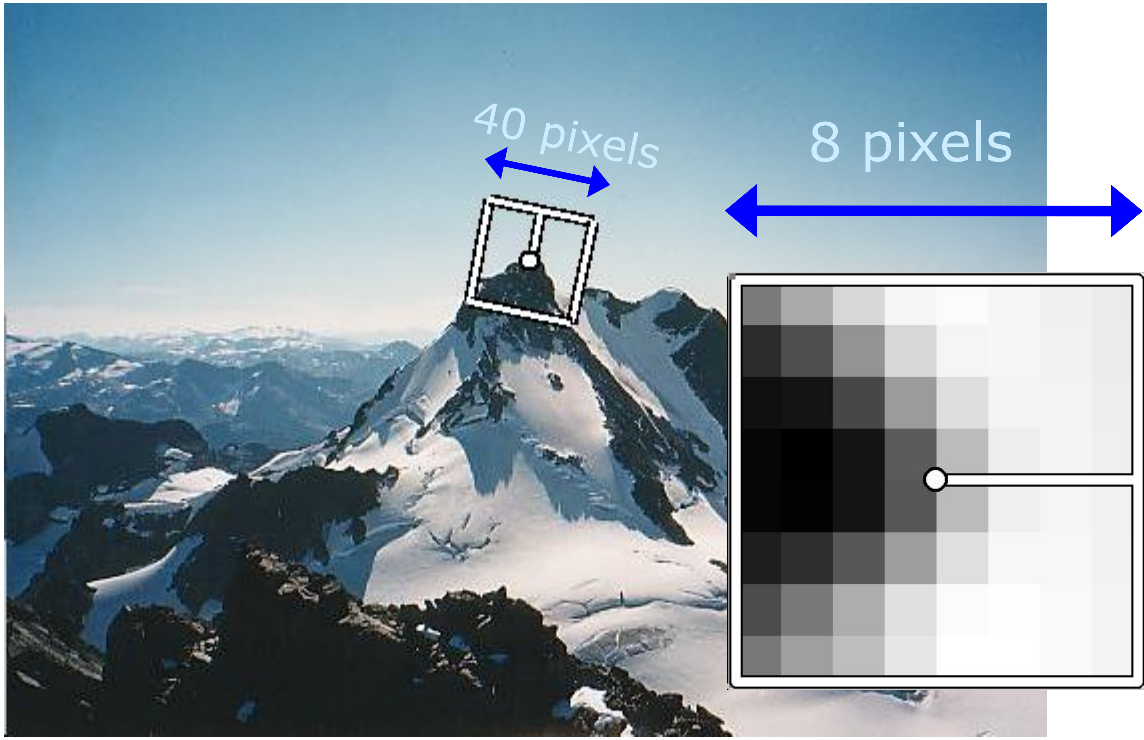
- Find dominant orientation of the image patch
  - This is given by $\mathbf{x}_{\mathbf{max}}$, the eigenvector of **H** corresponding to λ_{max} (the *larger* eigenvalue)
  - Rotate the patch according to this angle

Figure by Matthew Brown

Multiscale Oriented PatcheS descriptor

Take 40x40 square window around detected feature

- Scale to 1/5 size (using prefiltering)
- Rotate to horizontal
- Sample 8x8 square window centered at feature
- Intensity normalize the window by subtracting the mean, dividing by the standard deviation in the window



Adapted from slide by Matthew Brown

Detections at multiple scales

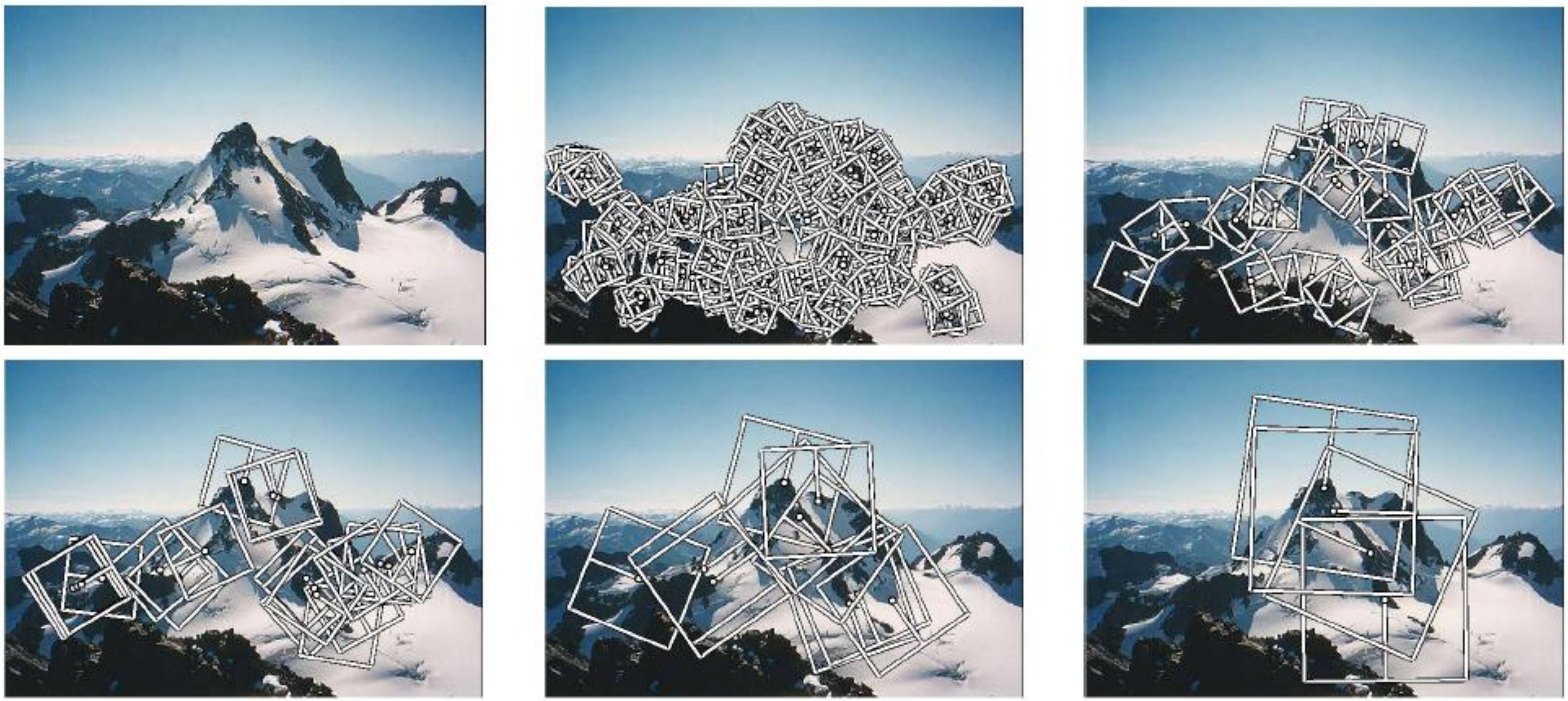


Figure 1. Multi-scale Oriented Patches (MOPS) extracted at five pyramid levels from one of the Matier images. The boxes show the feature orientation and the region from which the descriptor vector is sampled.

Scale Invariant Feature Transform

Basic idea:

- Take 16x16 square window around detected feature
- Compute edge orientation (angle of the gradient - 90°) for each pixel
- Throw out weak edges (threshold gradient magnitude)
- Create histogram of surviving edge orientations

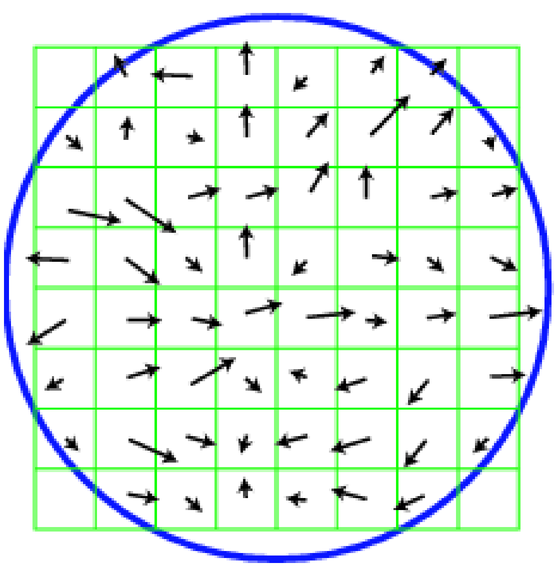
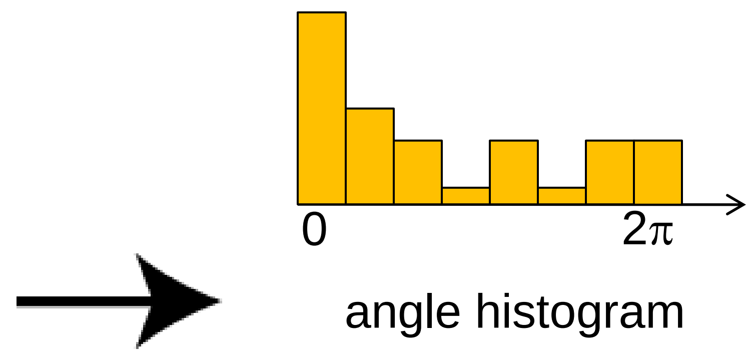
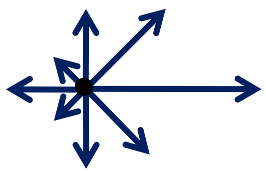


Image gradients



angle histogram

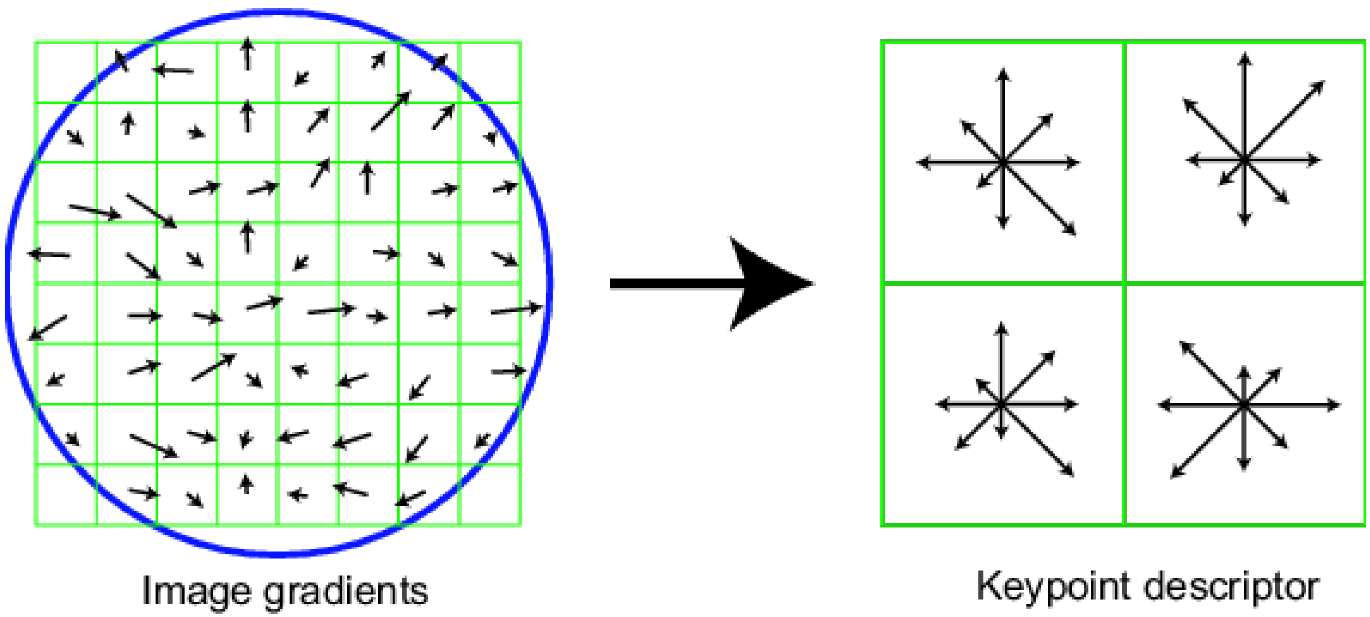


Adapted from slide by David Lowe

SIFT descriptor

Full version

- Divide the 16x16 window into a 4x4 grid of cells (2x2 case shown below)
- Compute an orientation histogram for each cell
- 16 cells * 8 orientations = 128 dimensional descriptor



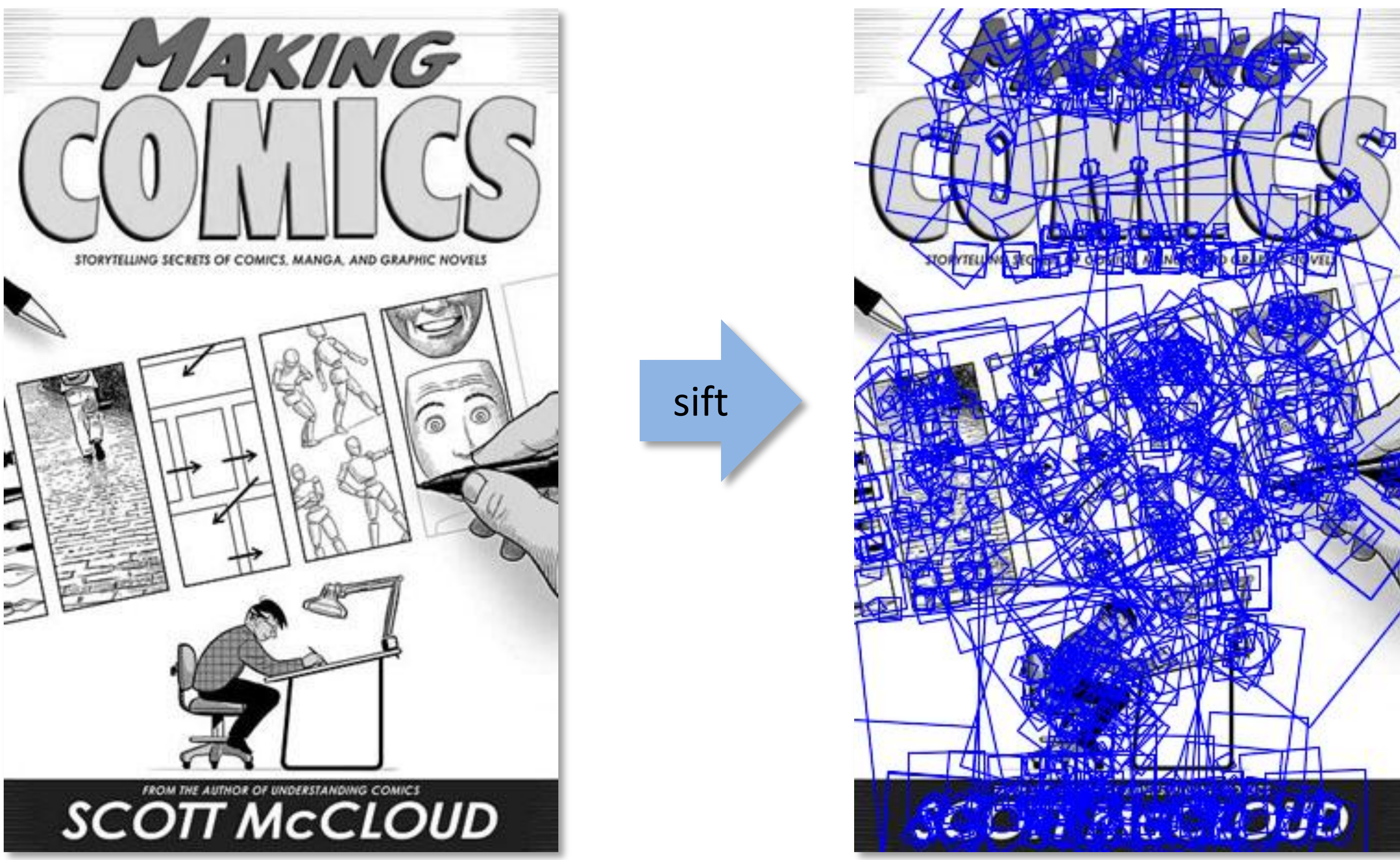
Adapted from slide by David Lowe

Properties of SIFT

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
  - http://people.csail.mit.edu/albert/ladypack/wiki/index.php/Known_implementations_of_SIFT

SIFT Example



868 SIFT features

Feature matching

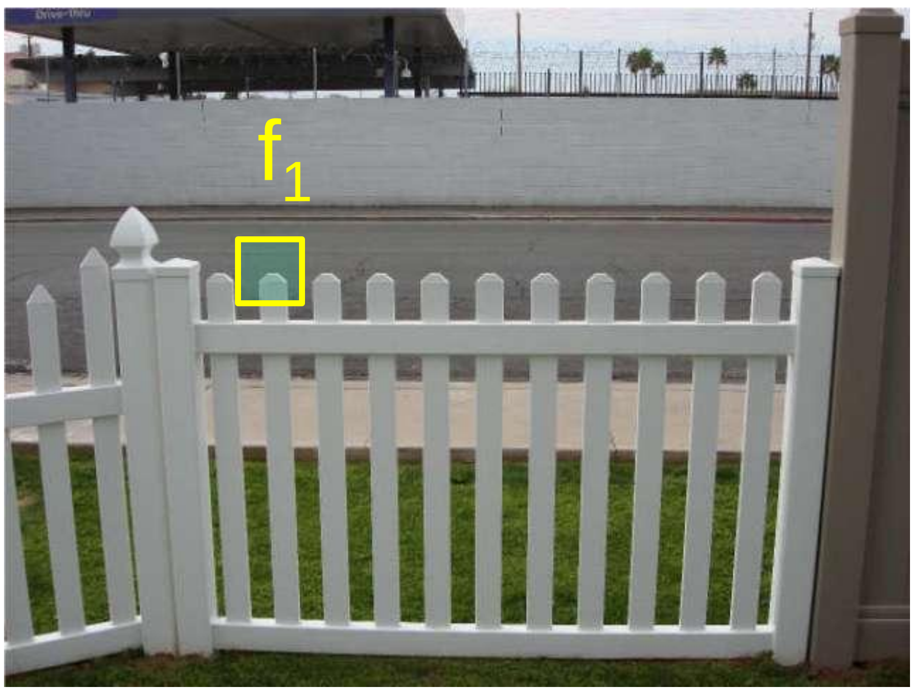
Given a feature in I_1, how to find the best match in I_2?

1. Define distance function that compares two descriptors
2. Test all the features in I_2, find the one with min distance

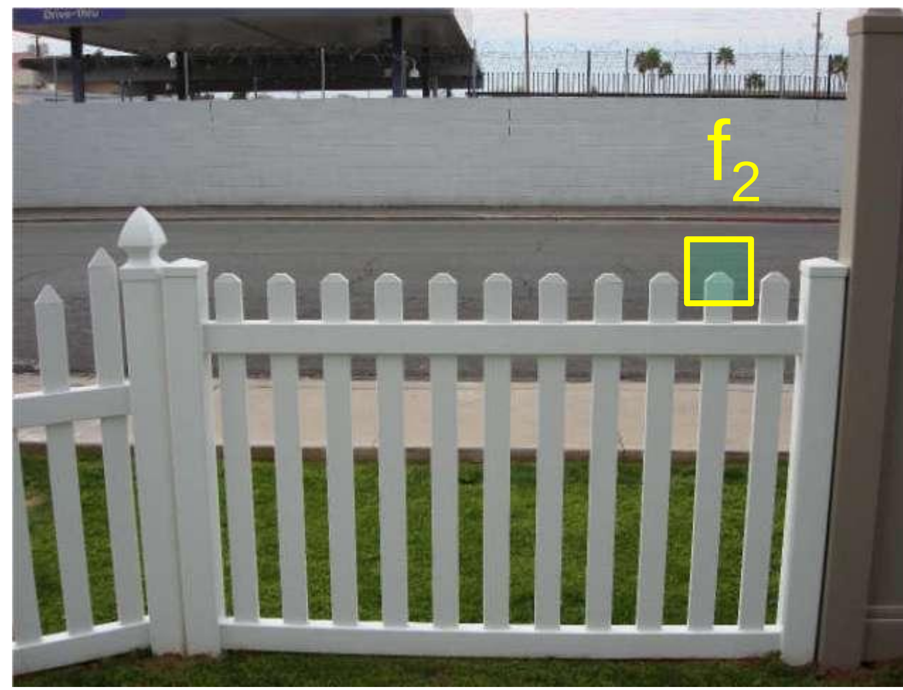
Feature distance

How to define the difference between two features f_1, f_2?

- Simple approach: L_2 distance, $||f_1 - f_2||$
- can give good scores to ambiguous (incorrect) matches



I_1

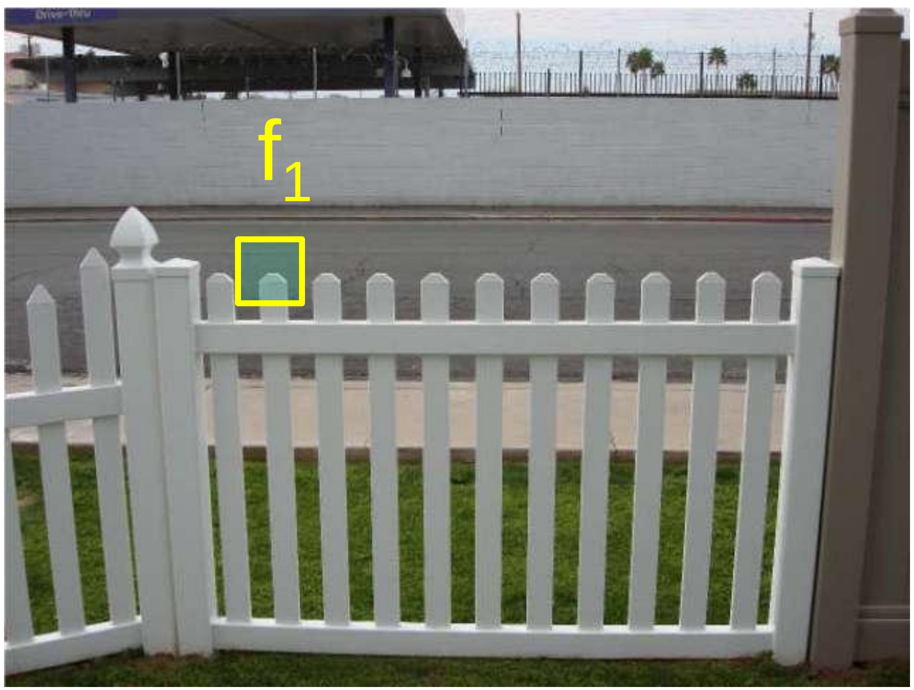


I_2

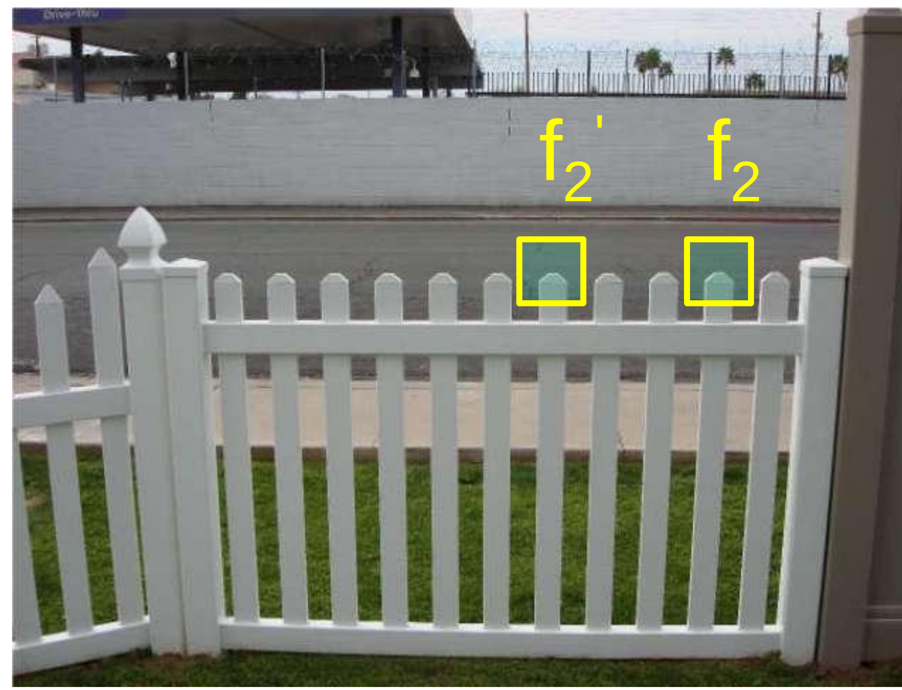
Feature distance

How to define the difference between two features f_1, f_2?

- Better approach: ratio distance = $||f_1 - f_2|| \, / \, || f_1 - f_2' ||$
  - f_2 is best SSD match to f_1 in I_2
  - f_2' is 2^{nd} best SSD match to f_1 in I_2
  - gives large values for ambiguous matches

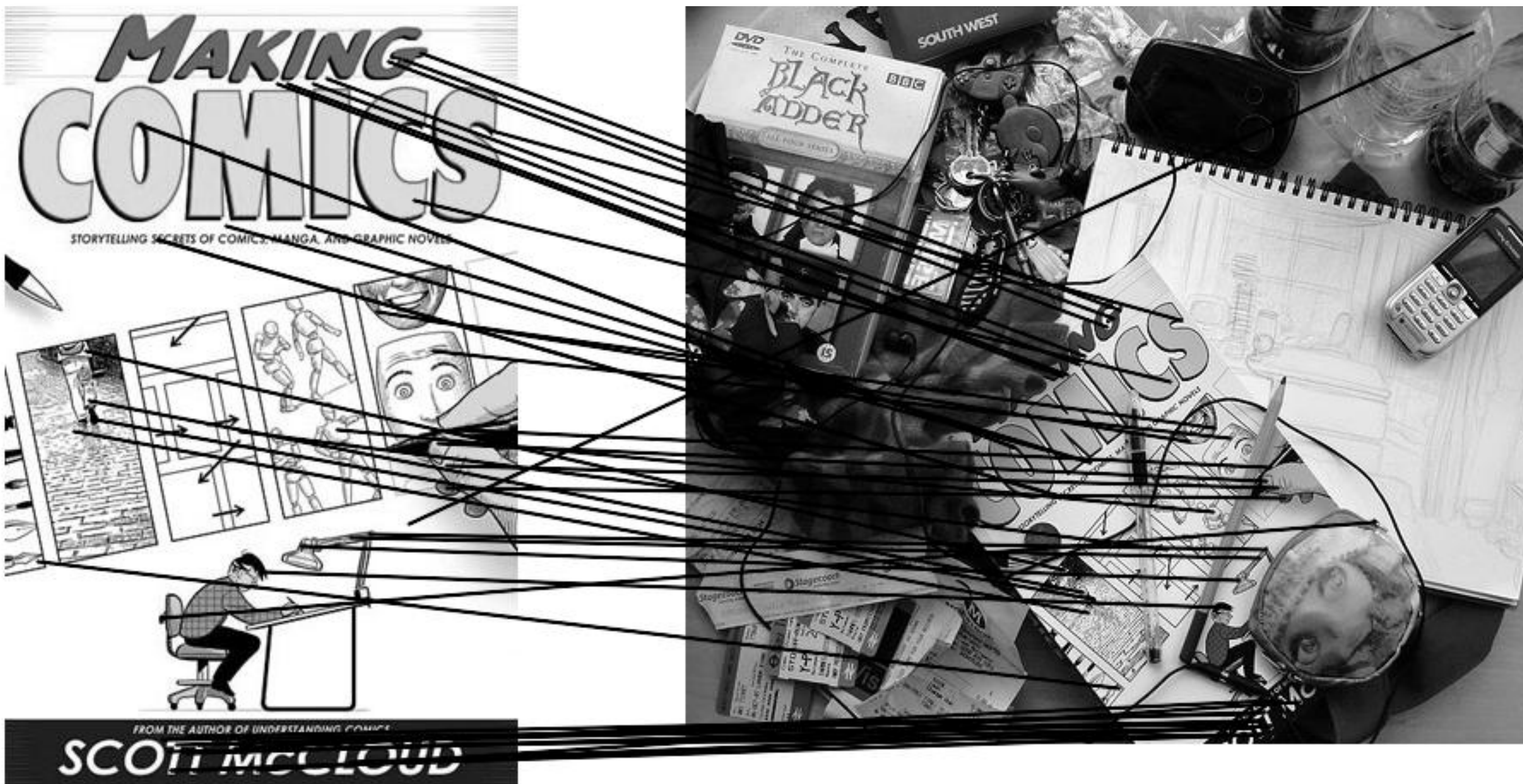


I_1



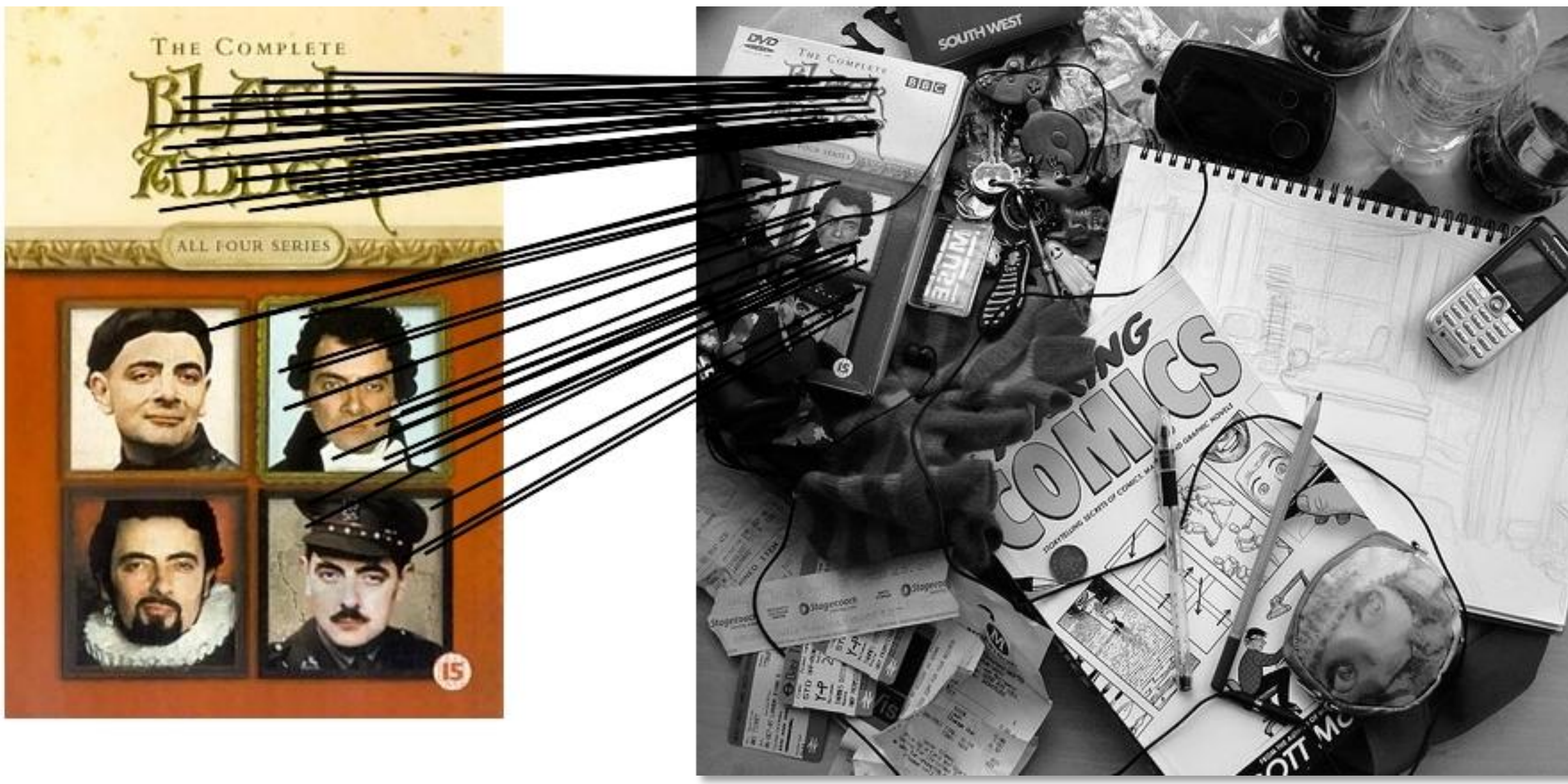
I_2

Feature matching example



51 matches

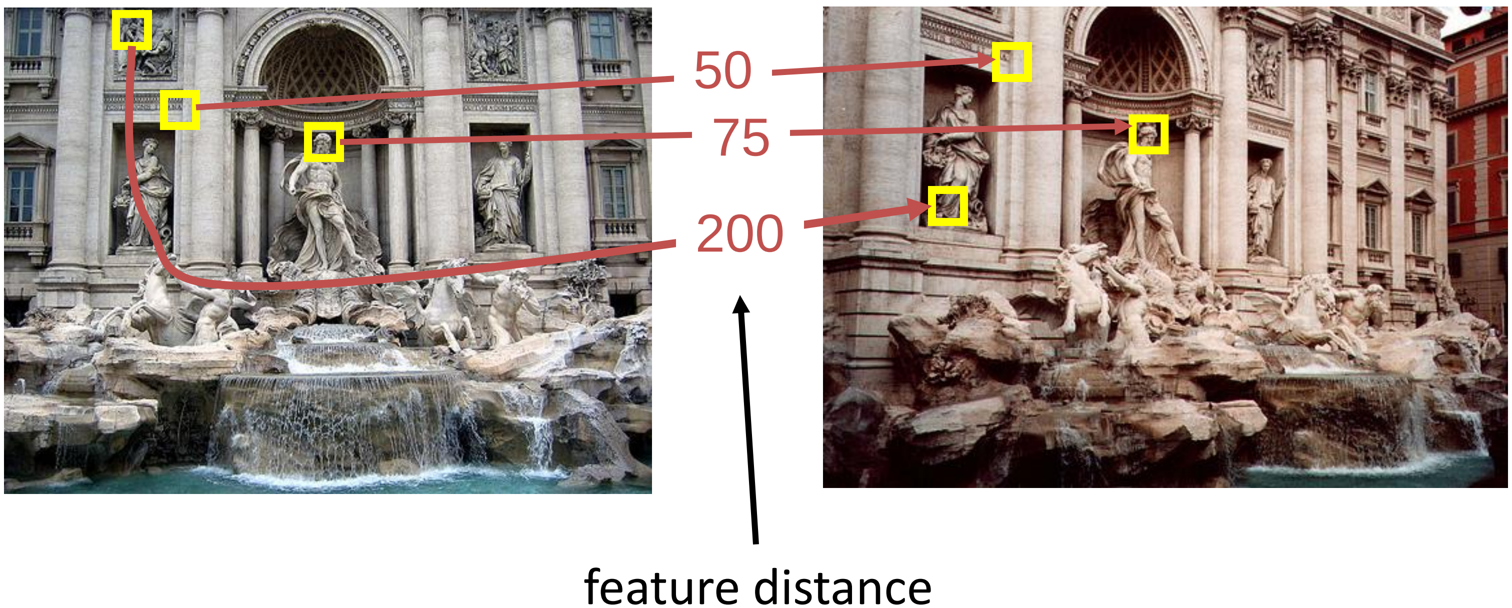
Feature matching example



58 matches

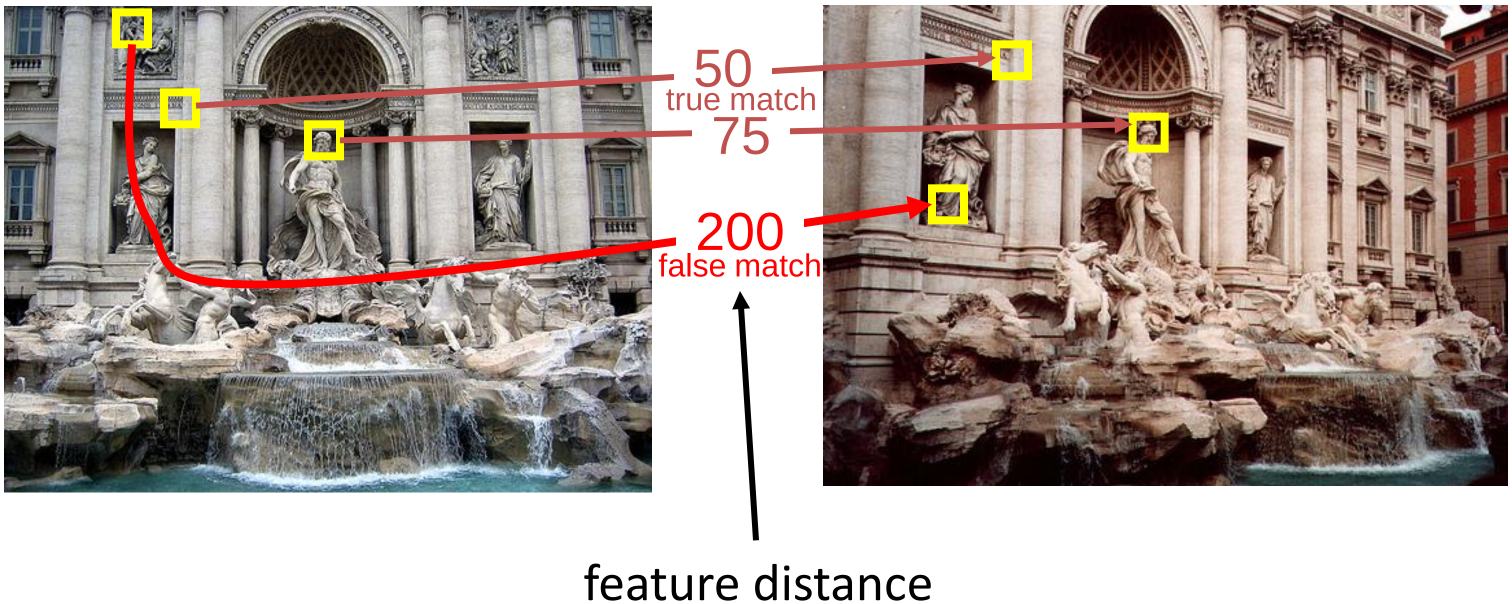
Evaluating the results

How can we measure the performance of a feature matcher?



True/false positives

How can we measure the performance of a feature matcher?

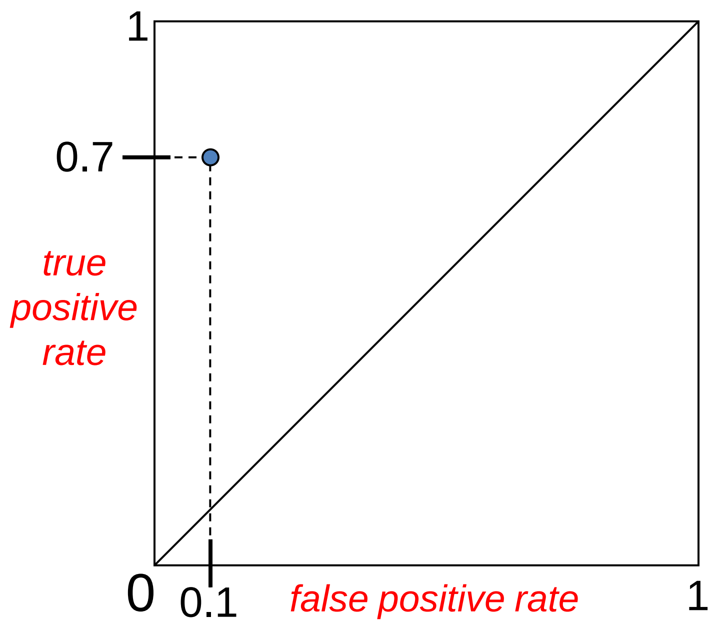


The distance threshold affects performance

- True positives = # of detected matches that are correct
  - Suppose we want to maximize these—how to choose threshold?
- False positives = # of detected matches that are incorrect
  - Suppose we want to minimize these—how to choose threshold?

Evaluating the results

How can we measure the performance of a feature matcher?



$$\frac{\text{\# true positives}}{\text{\# matching features (positives)}}$$

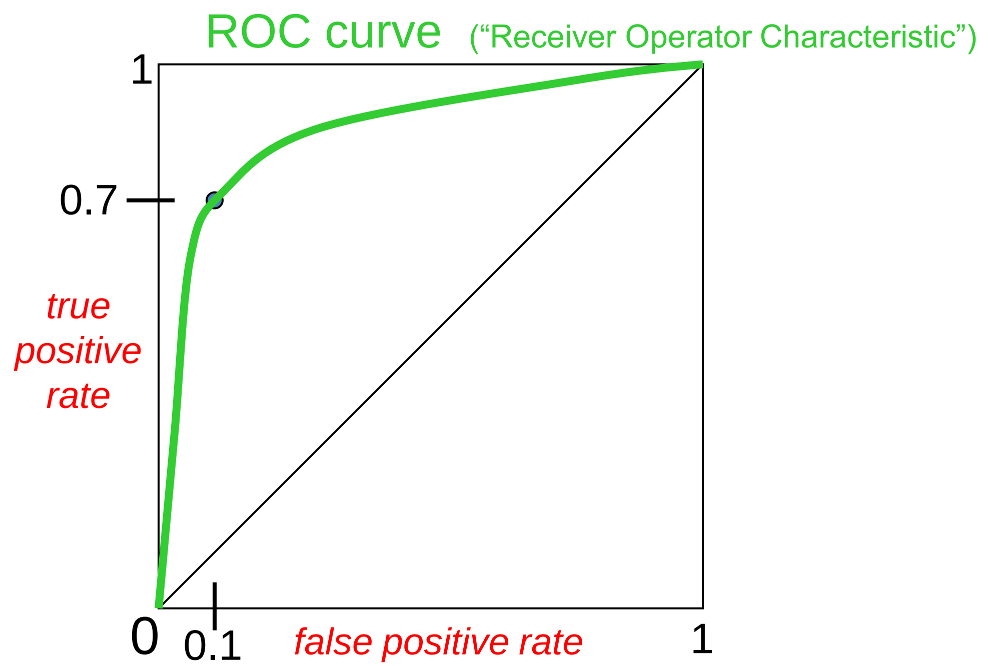
"recall"

$$\frac{\text{\# false positives}}{\text{\# unmatched features (negatives)}}$$

1 - "precision"

Evaluating the results

How can we measure the performance of a feature matcher?



true positives
matching features (positives)

"recall"

false positives
unmatched features (negatives)

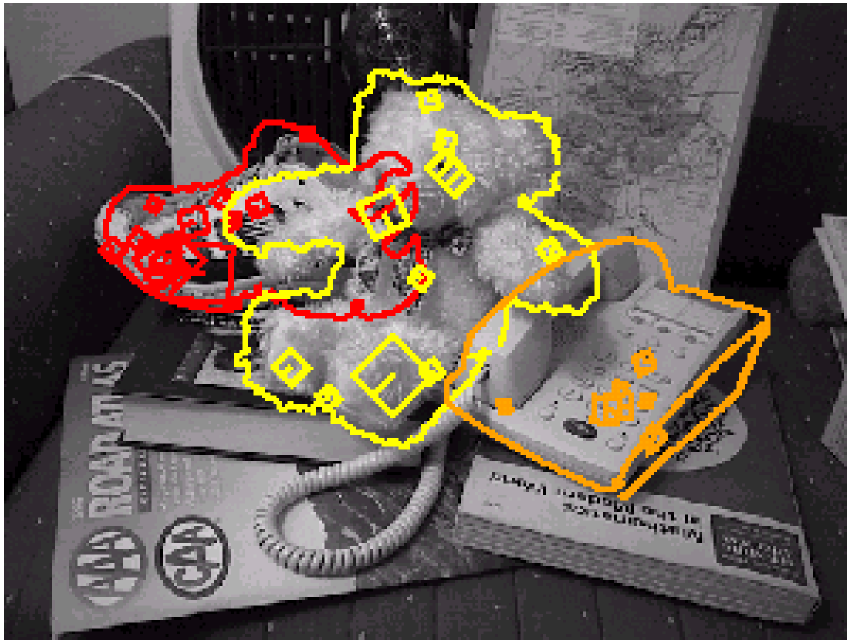
1 - "precision"

Lots of applications

Features are used for:

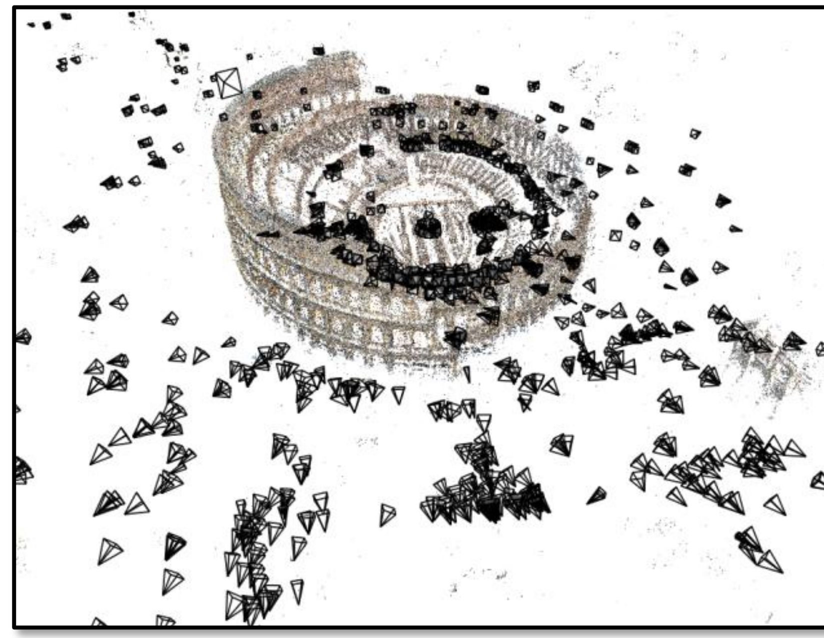
- Image alignment (e.g., mosaics)
- 3D reconstruction
- Motion tracking
- Object recognition (e.g., **Google Goggles**)
- Indexing and database retrieval
- Robot navigation
- … other

Object recognition (David Lowe)



3D Reconstruction

Internet Photos ("Colosseum")



Reconstructed 3D cameras and points

Sony Aibo

<u>SIFT usage:</u>

- Recognize charging station
- Communicate with visual cards
- Teach object recognition

Questions?

Image alignment

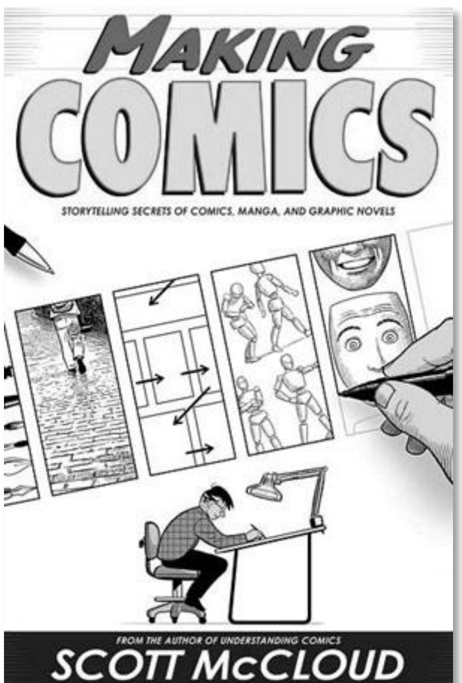
Image taken from same viewpoint, just rotated.

Can we line them up?

Image alignment

Why don’t these image line up exactly?

What is the geometric relationship between these two images?



?