CS688/WST665: Web-Scale Image Retrieval
Descriptors

Sung-Eui Yoon
(윤성의)

Course URL:
http://sglab.kaist.ac.kr/~sungeui/IR

KAIST

Announcements

- **17 students take the course**
- **Presentations**
  - One presentation for each person: 25min talk and 12min for its Q&A
  - Deeper understanding on a paper is required; go over its related papers carefully
  - Paper/its presentation date selection: Oct-16
- **Projects**
  - Only 2 or more are allowed; clear role for each student!
  - Final presentation: Dec. 11 & 13
  - Mid-term review: Nov. 13 & 15
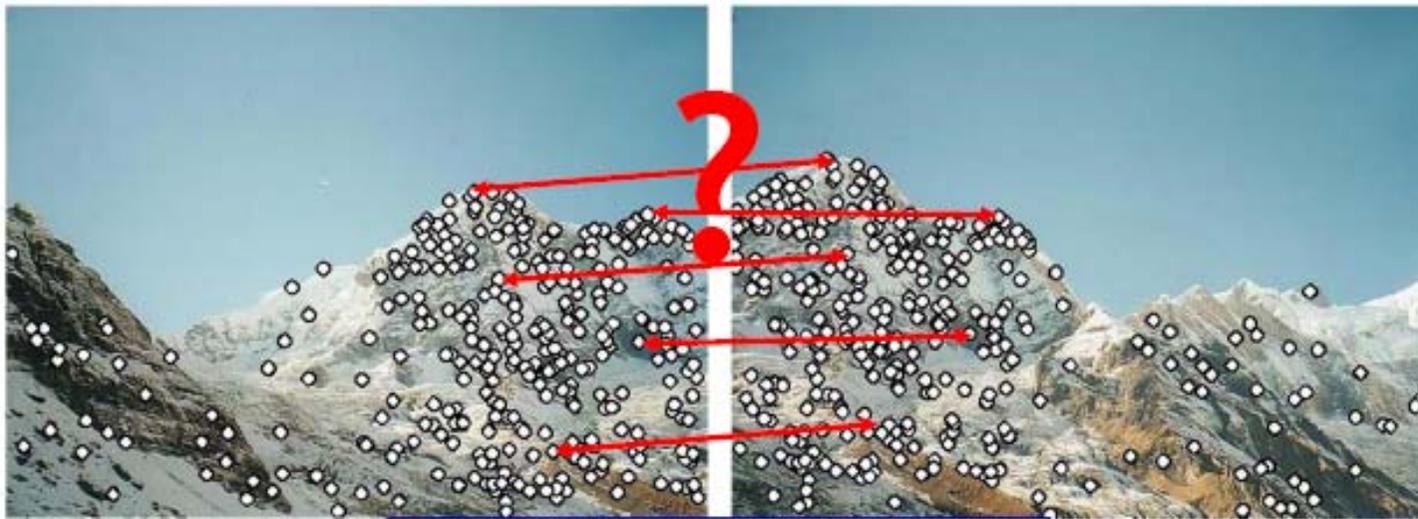 - Team formation: Oct - 16

KAIST

What we will learn today

- Local descriptors
  - SIFT
  - An assortment of other descriptors
  - Applications

Local Descriptors

- We know how to detect points

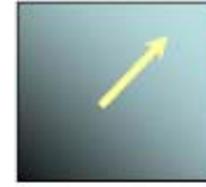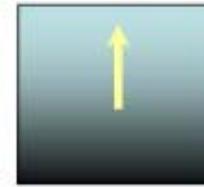- Next question:

How to *describe* them for matching?

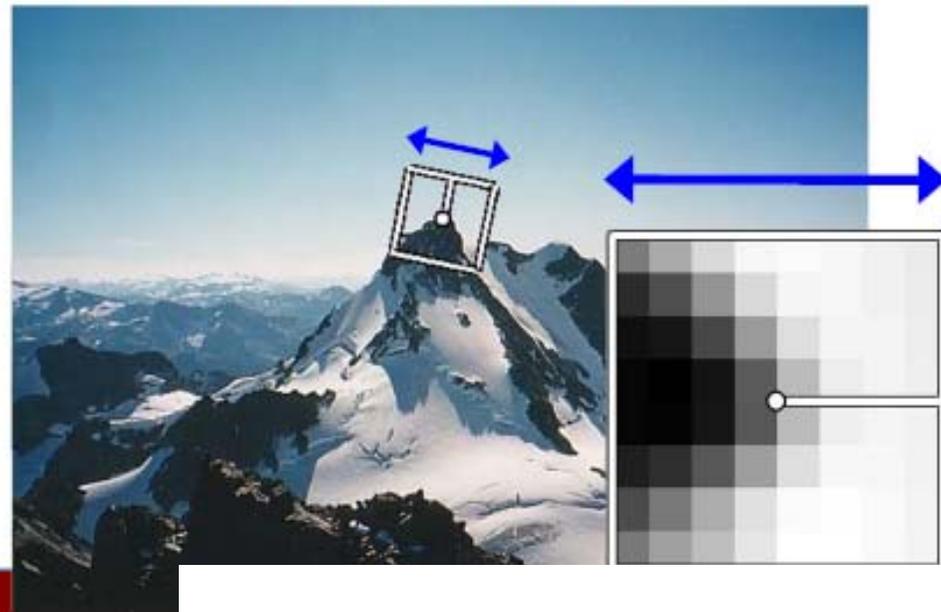**Point descriptor should be:**
1. **Invariant**
2. **Distinctive**

Fei-Fei Li

Rotation Invariant Descriptors

- Find local orientation
  - Dominant direction of gradient for the image patch

- Rotate patch according to this angle
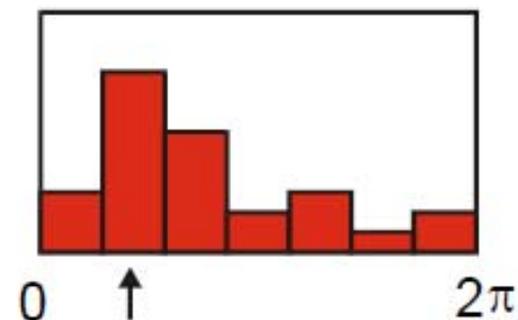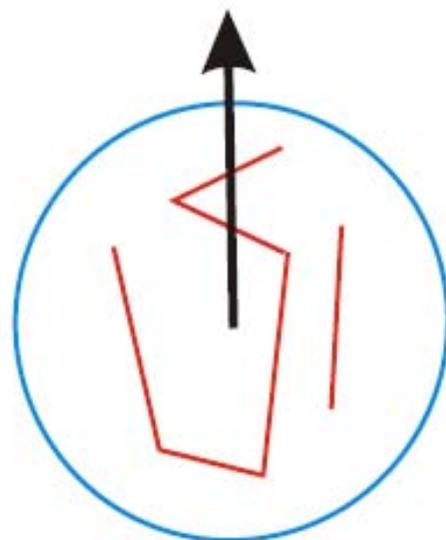 - This puts the patches into a canonical orientation.

Orientation Normalization: Computation

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

0   2π

Fei-Fei Li

The Need for Invariance
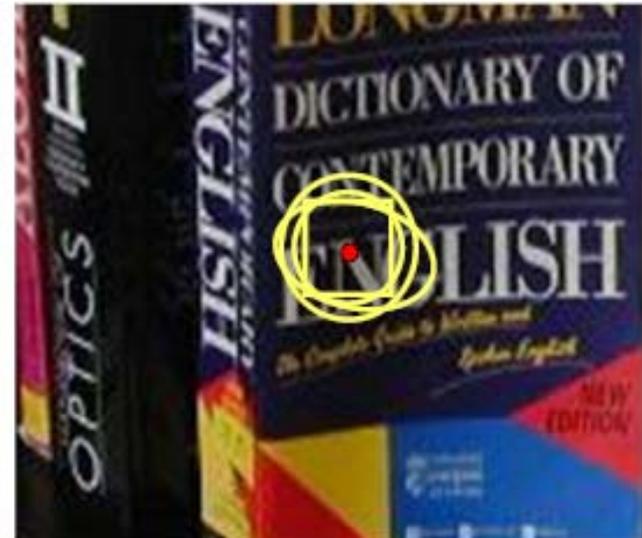
- Up to now, we had invariance to
  - Translation
  - Scale
  - Rotation
- Not sufficient to match regions under viewpoint changes
  - For this, we need also affine adaptation

Fei-Fei Li

Affine Transformation

- ## Matrix representation
  - ### Less general types than perspective transformation

$$\begin{bmatrix} \vec{y} \\ 1 \end{bmatrix} = \begin{bmatrix} & A & & \vec{b} \\ 0,\ldots,0 & & 1 \end{bmatrix} \begin{bmatrix} \vec{x} \\ 1 \end{bmatrix}$$

- ## Geometric interpretation
  - ### Rotation + scaling
  - ### Shearing

KAIST

Affine Adaptation

- Problem:
  - Determine the characteristic shape of the region.
  - Assumption: shape can be described by "local affine frame".

- Solution: iterative approach
  - Use a circular window to compute second moment matrix.
  - Compute eigenvectors to adapt the circle to an ellipse.
  - Recompute second moment matrix using new window and iterate...

Fei-Fei Li
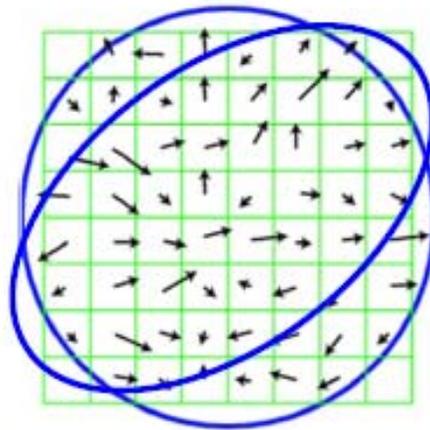
Iterative Affine Adaptation
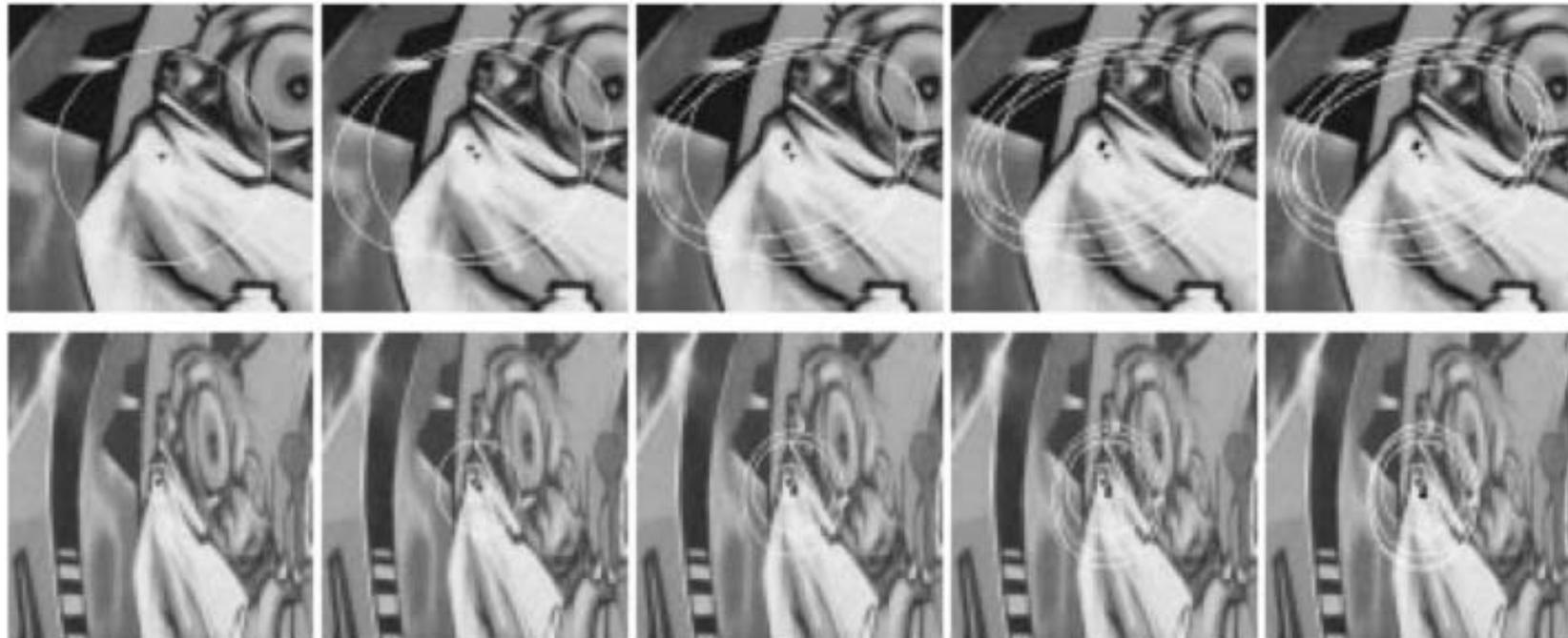
1. Detect keypoints, e.g. multi-scale Harris
2. Automatically select the scales
3. Adapt affine shape based on second order moment matrix
4. Refine point location

K. Mikolajczyk and C. Schmid, Scale and affine invariant interest point detectors, IJCV 60(1):63-86, 2004.

Fei-Fei Li

Affine Normalization/Deskewing
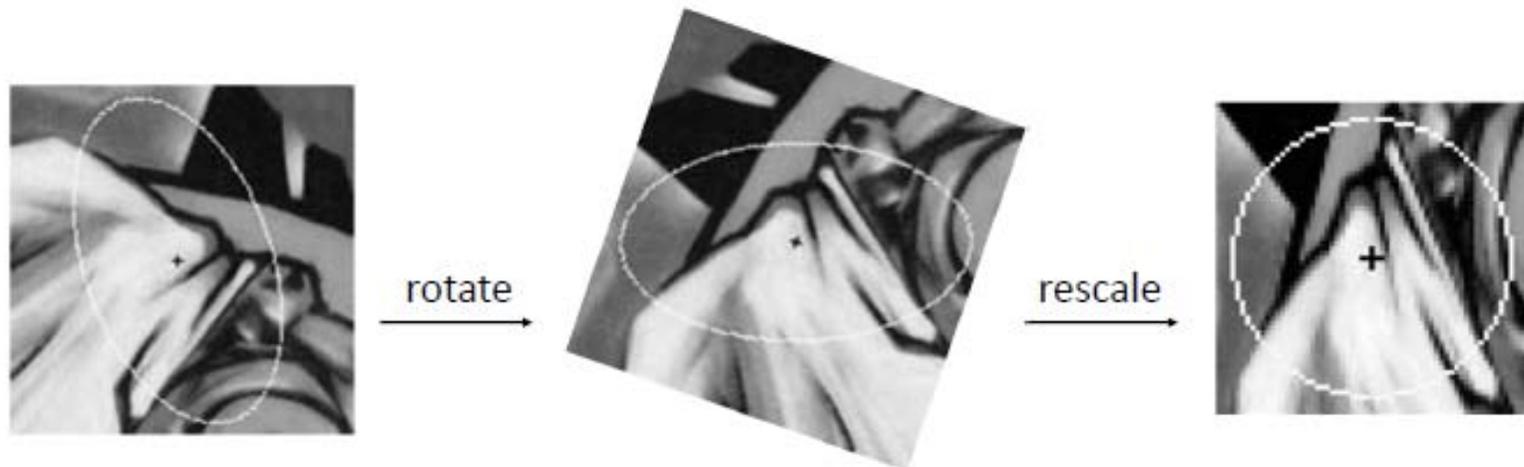
rotate

rescale

- Steps
  - Rotate the ellipse's main axis to horizontal
  - Scale the x axis, such that it forms a circle

Fei-Fei Li

Affine Adaptation Example
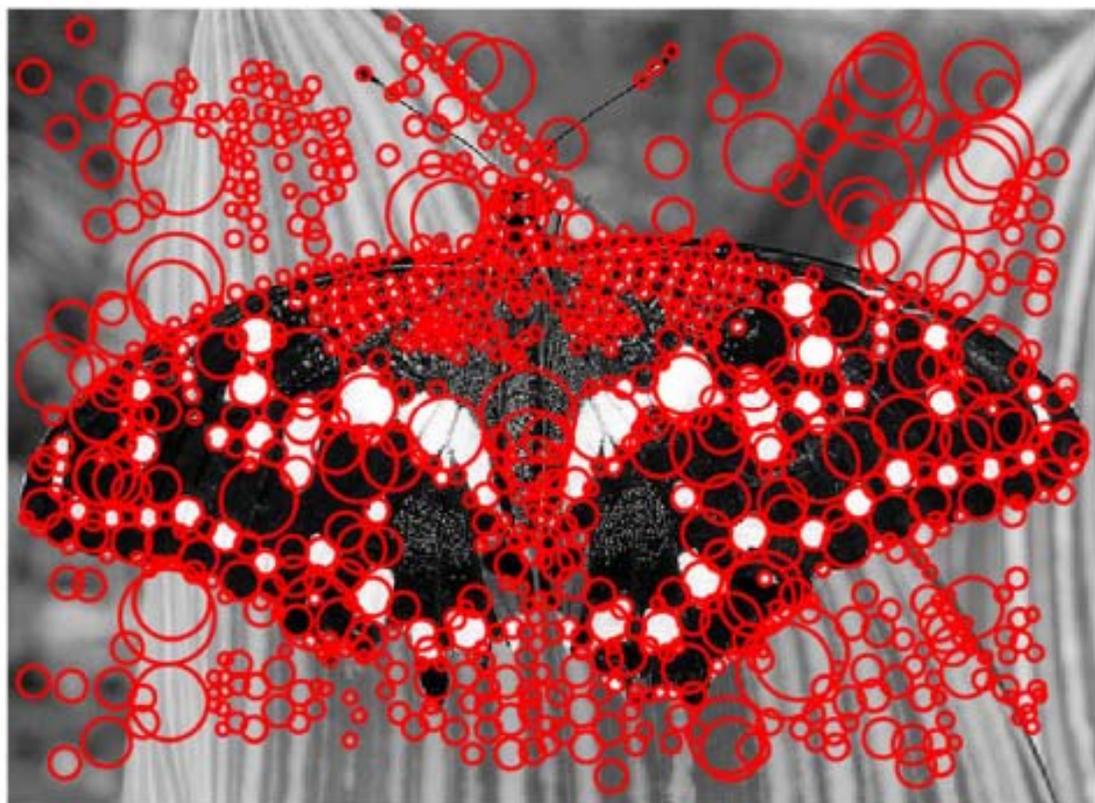
Scale-invariant regions (blobs)

Fei-Fei Li

Affine Adaptation Example

Affine-adapted blobs

Fei-Fei Li

Summary: Affine-Inv. Feature Extraction

Extract affine regions → **Normalize regions** → **Eliminate rotational ambiguity** → Compare descriptors
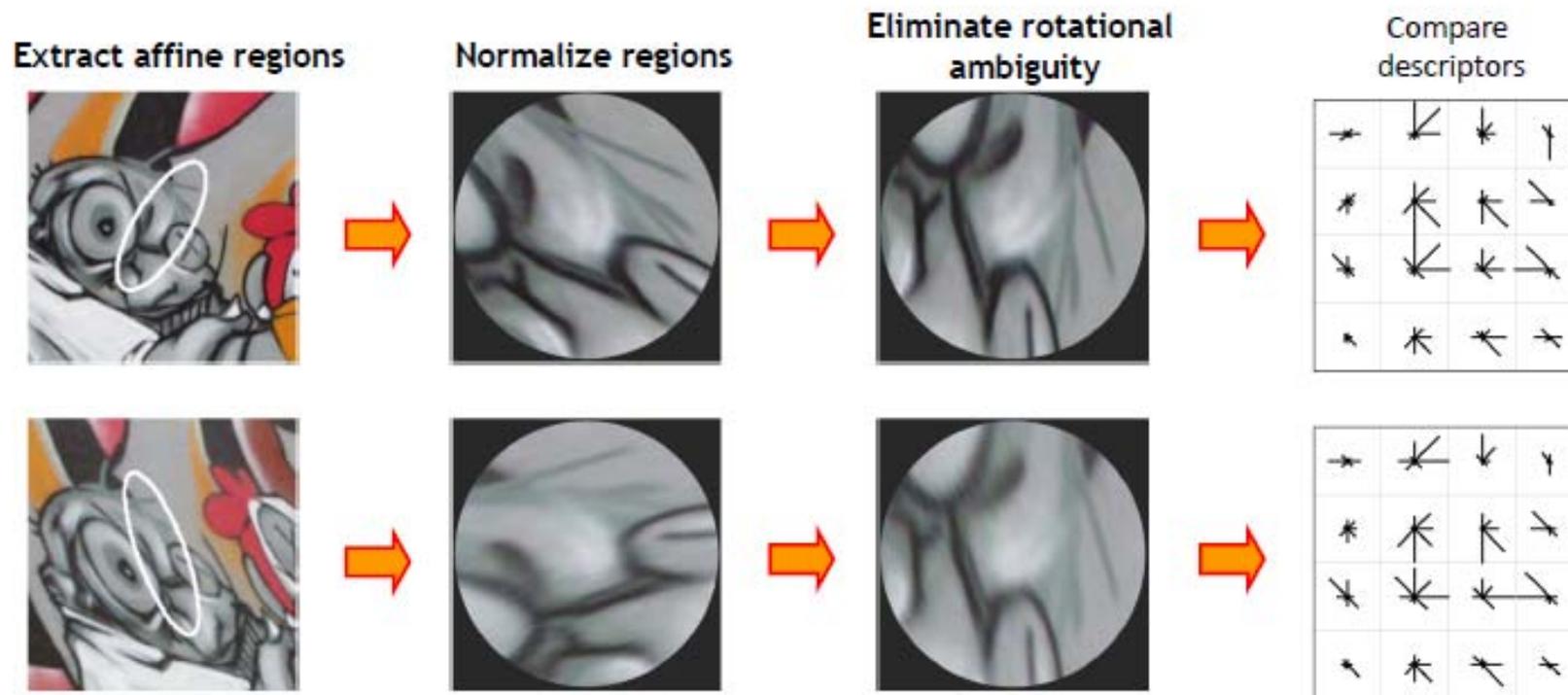
Slide credit: Svetlana Lazebnik

Fei-Fei Li

Local Descriptors

- We know how to detect points

- Next question:

How to *describe* them for matching?

Point descriptor should be:
1. Invariant
2. Distinctive

Fei-Fei Li
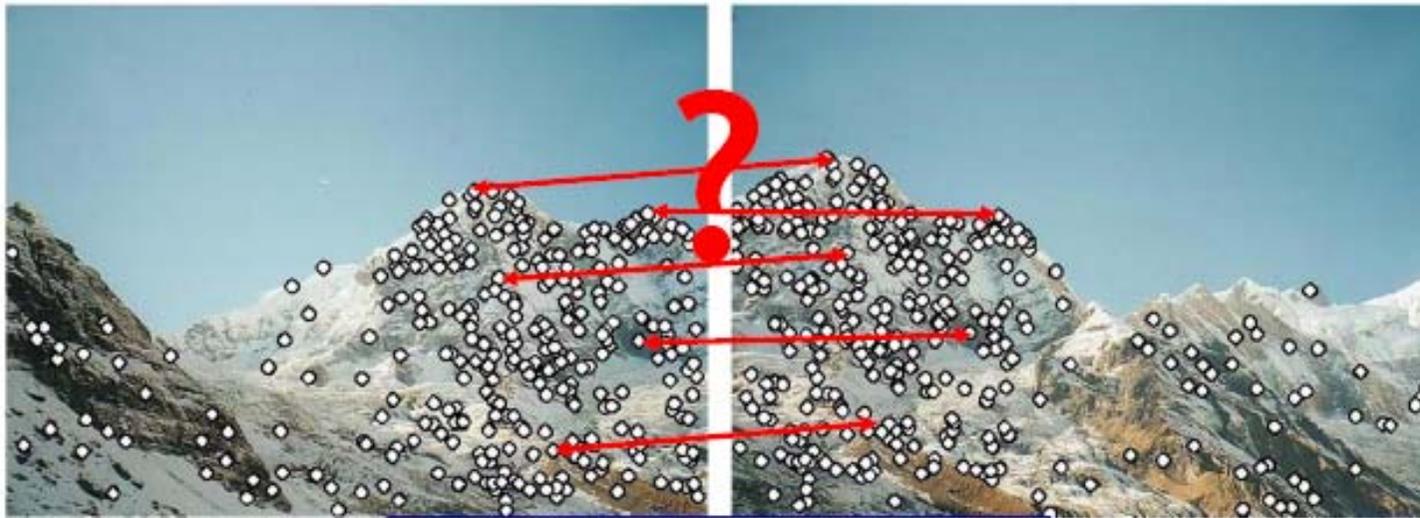
Local Descriptors

- Simplest descriptor: list of intensities within a patch.
- What is this going to be invariant to?

Write regions as vectors

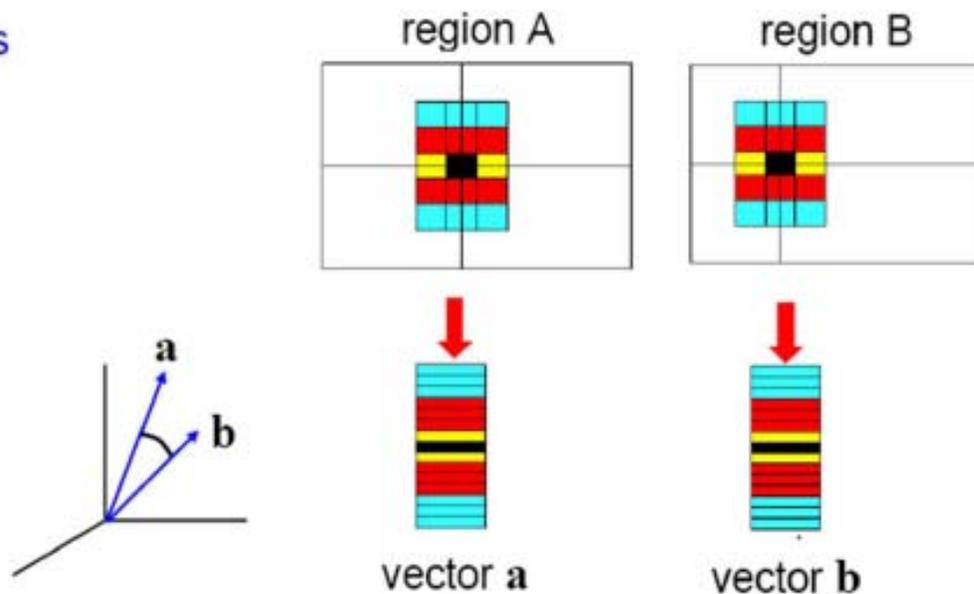$$A \rightarrow a, \quad B \rightarrow b$$

region A   region B

vector a   vector b

Fei-Fei Li

Feature Descriptors

- Disadvantage of patches as descriptors:
  - Small shifts can affect matching score a lot

- Solution: histograms

Fei-Fei Li

Feature Descriptors: SIFT

- **S**cale **I**nvariant **F**eature **T**ransform

- Descriptor computation:
  - Divide patch into 4x4 sub-patches: 16 cells
  - Compute histogram of gradient orientations (8 reference angles) for all pixels inside each sub-patch
  - Resulting descriptor: 4x4x8 = 128 dimensions
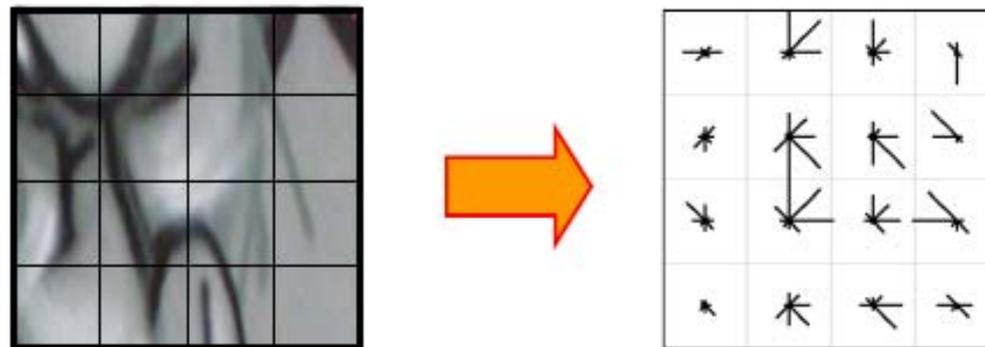
David G. Lowe. "Distinctive image features from scale-invariant keypoints." *IJCV* 60 (2), pp. 91-110, 2004.

Fei-Fei Li

Overview: SIFT

- Extraordinarily robust matching technique
  - Can handle changes in viewpoint up to ~60 deg. 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

Fei-Fei Li

Working with SIFT Descriptors
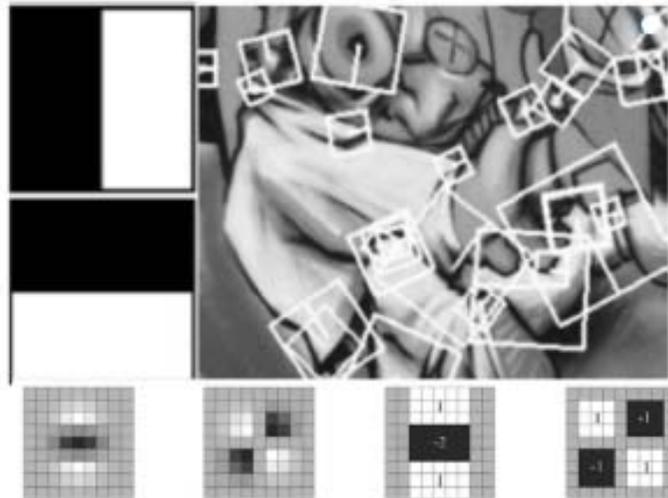
- One image yields:
  - n 128-dimensional descriptors: each one is a histogram of the gradient orientations within a patch
    - [n x 128 matrix]
  - n scale parameters specifying the size of each patch
    - [n x 1 vector]
  - n orientation parameters specifying the angle of the patch
    - [n x 1 vector]
  - n 2D points giving positions of the patches
    - [n x 2 matrix]

Fei-Fei Li

Local Descriptors: SURF

Fast approximation of SIFT idea

Efficient computation by 2D box filters & integral images
$\Rightarrow$ 6 times faster than SIFT

Equivalent quality for object identification

http://www.vision.ee.ethz.ch/~surf

GPU implementation available

Feature extraction @ 100Hz
(detector + descriptor, 640×480 img)

http://homes.esat.kuleuven.be/~ncorneli/gpusurf/

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

http://www.vision.ee.ethz.ch/~surf

Fei-Fei Li

Other Descriptors

- **Gray-scale intensity**

- **GIST**
- **Many others**

KAIST
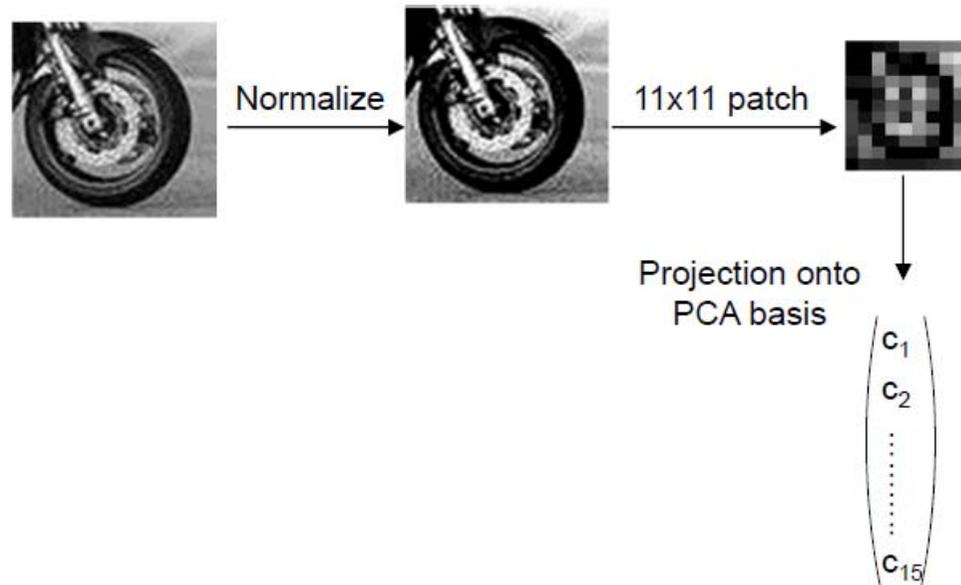
Applications of Local Invariant Features

- Wide baseline stereo
- Motion tracking
- Panoramas
- Mobile robot navigation
- 3D reconstruction
- Recognition
  - Specific objects
  - Textures
  - Categories
- ...

Fei-Fei Li

Wide-Baseline Stereo

Image from T. Tuytelaars ECCV 2006 tutorial

Fei-Fei Li

Automatic Mosaicing

[Brown & Lowe, ICCV'03]

Fei-Fei Li
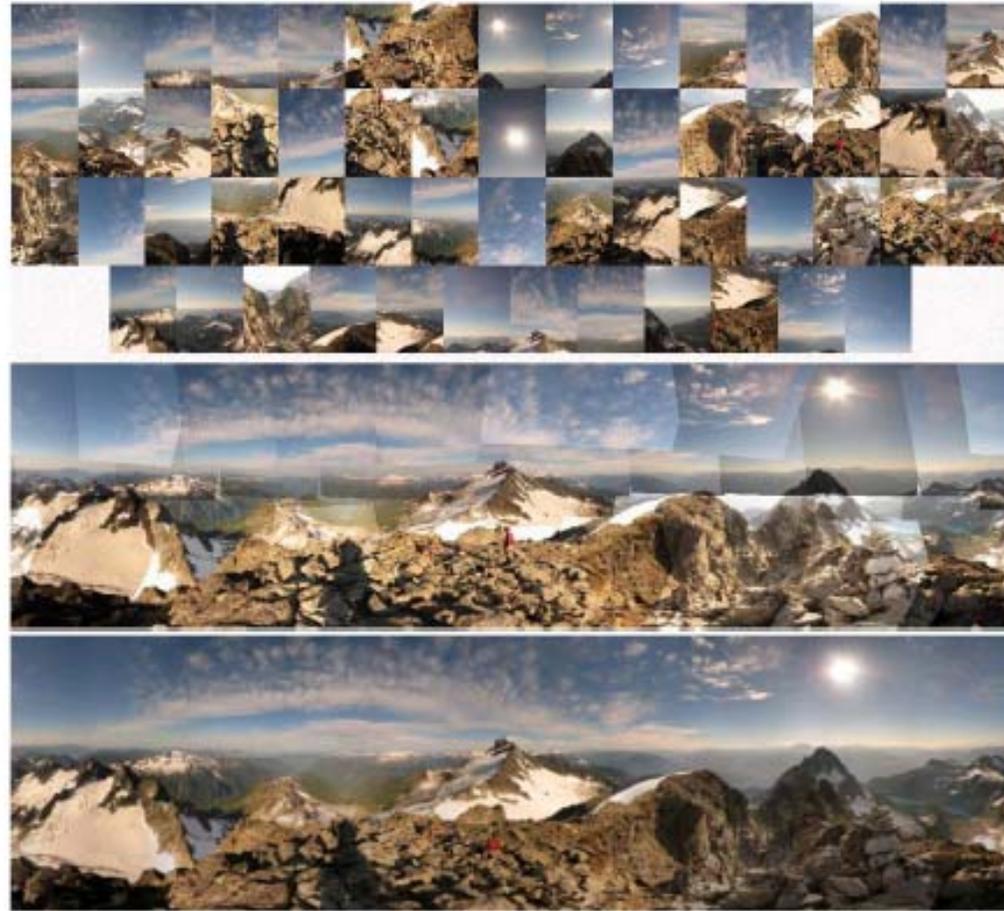
Panorama Stitching

(a) Matier data set (7 images)

iPhone version available

(b) Matier final stitch

[Brown, Szeliski, and Winder, 2005]

http://www.cs.ubc.ca/~mbrown/autostitch/autostitch.html

Fei-Fei Li

Recognition of Specific Objects, Scenes

Schmid and Mohr 1997

Sivic and Zisserman, 2003

Rothganger et al. 2003

Lowe 2002

Fei-Fei Li

Value of Local Features

- Advantages
  - Critical to find distinctive and repeatable local regions for multi-view matching.
  - Complexity reduction via selection of distinctive points.
  - Describe images, objects, parts without requiring segmentation; robustness to clutter & occlusion.
  - Robustness: similar descriptors in spite of moderate view changes, noise, blur, etc.

- How can we use local features for such applications?
  matching and recognition

Fei-Fei Li

Alignment Problem

- Fit different images into one canonical image

Alignment Problem

- **Many different approaches exist**

- **Simple fitting procedure in the linear least square sense**
  - Approximates viewpoint changes for roughly planar objects and roughly orthographic cameras
  - Can be used to initialize fitting for more complex models

- **We do not discuss this issue here**
 - Will be discussed in a computer vision course

KAIST

Time for a Demo...

Automatic panorama stitching

Matthew Brown: http://cvlab.epfl.ch/~brown/autostitch/autostitch.html

Fei-Fei Li

References and Further Reading

•**More details on the alignment problem can be found in:**

- R. Hartley, A. Zisserman
  Multiple View Geometry in Computer Vision
  2nd Ed., Cambridge Univ. Press, 2004

- Details about the DoG detector and the SIFT descriptor can be found in
  - D. Lowe, Distinctive image features from scale-invariant keypoints, *IJCV* 60(2), pp. 91-110, 2004

- Try the available local feature detectors and descriptors
  - http://www.robots.ox.ac.uk/~vgg/research/affine/detectors.html#binaries
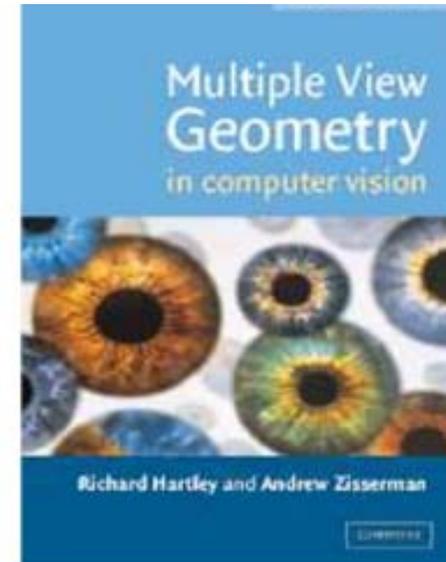
What we have learned today

- Local descriptor
  - SIFT
  - An assortment of other descriptors
  - Applications

Next Time…

- Object recognition
- Bag-of-Words (BoW) models

KAIST

PA1

- **Objective**
  - **Understand how to extract SIFT features and to use related libraries**

- **Deadline**
  - **Oct-2(Tue.) (before 11:59pm)**

KAIST

Homework for Every Class

- **Go over the next lecture slides**
- **Come up with one question on what we have discussed today**
  - 0 for no questions
  - 2 for typical questions
  - 3 for questions with thoughts
  - 4 for questions that surprised me

KAIST