CS688: Web-Scale Image Search
Keypoint Localization

Sung-Eui Yoon
(윤성의)

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

KAIST

Homework for Every Class

- **Go over the next lecture slides**

- **Come up with one question on what we have discussed today**
  - 1 for typical questions (that were answered in the class)
  - 2 for questions with thoughts or that surprised me

- **Write questions at least 4 times before the mid-term**
  - Multiple questions in one time will be counted as one time

- Common questions are addressed at my draft
  - Some of questions will be discussed in the class

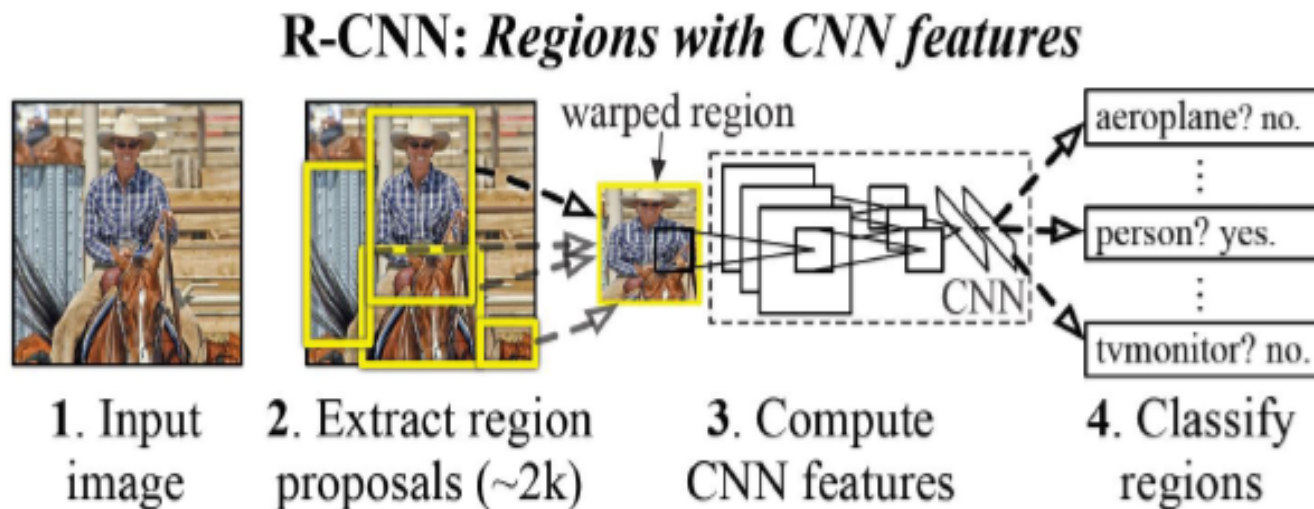- If you want to know the answer of your question, ask me or TA **on person**

KAIST

Homework for Every Class

- **Go over recent papers on image search**
  - **High quality papers: Papers published at the top-tier conf. or close it can be presented; e.g., CVPR, ICCV, ECCV, ACM ICMR, ACM MM, ACM SIGGRAPH**
  - **Recent publication: papers published since 2015**
  - **Find and browse two papers, and submit your summary before every beginning of the Thur. class; submit two summaries**
  - **Online submission is possible**
- **Think about possible team members**
- **Too late if you think them later..**

KAIST

Computer Vision Field: CVPR, ICCV, ECCV

- **Handle various computer vision problems**

- **Get various machine learning techniques from ICML, NIPS**

KAIST

Example: R-CNN [CVPR 14, oral]

R-CNN: *Regions with CNN features*

warped region

aeroplane? no.

person? yes.

tvmonitor? no.

CNN

1. Input image

2. Extract region proposals (~2k)

3. Compute CNN features

4. Classify regions

- Three Modules
  - ① Category-independent region proposals
  - ② CNN that extracts a fixed-length feature vector from each region (Blackbox feature extractor)
  - ③ Class-specific linear SVMs

Rich feature hierarchies for accurate object detection and semantic segmentation, Slide is from Mr. Lee

KAIST

Example: Localization Networks
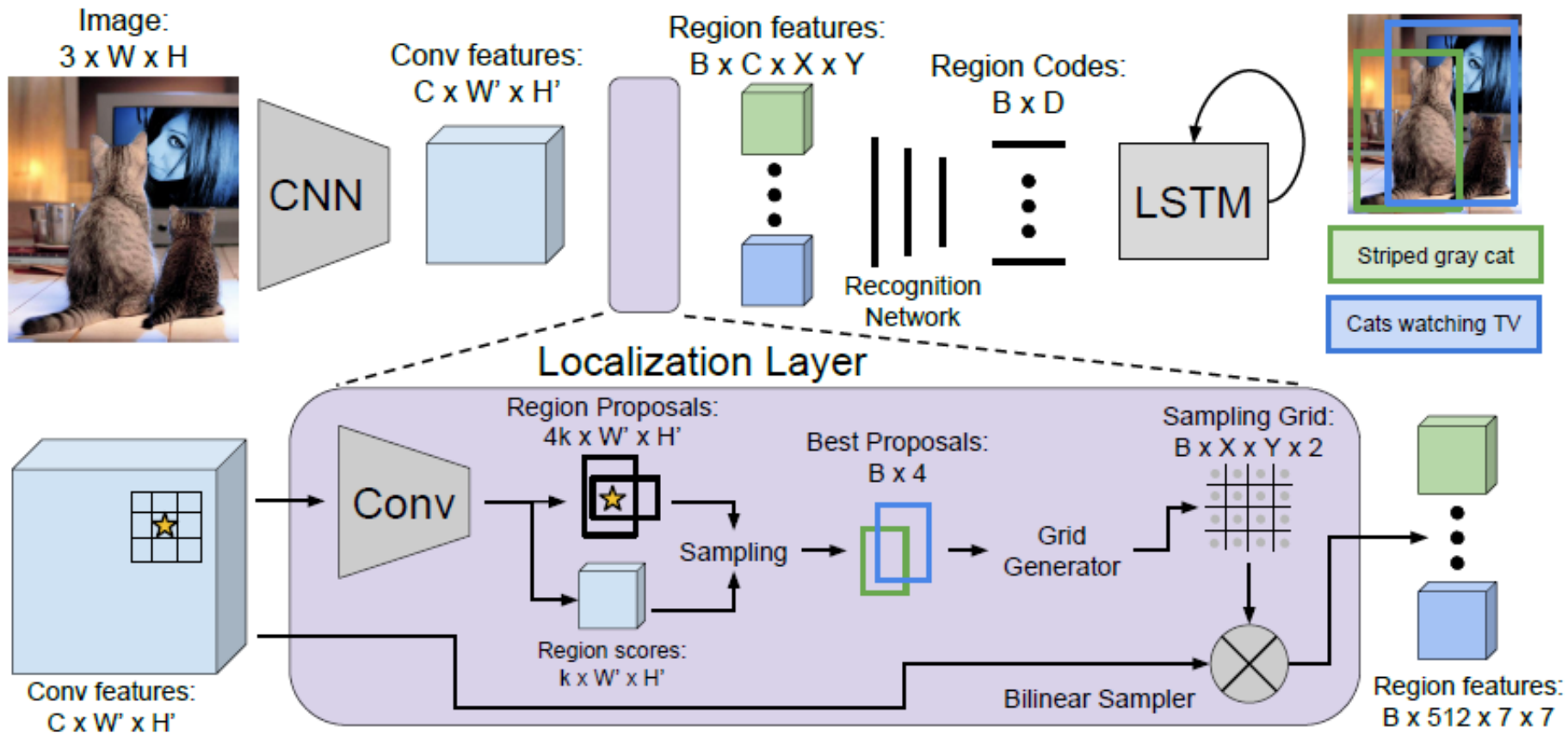
- **DenseCap, CVPR 16 (oral)**
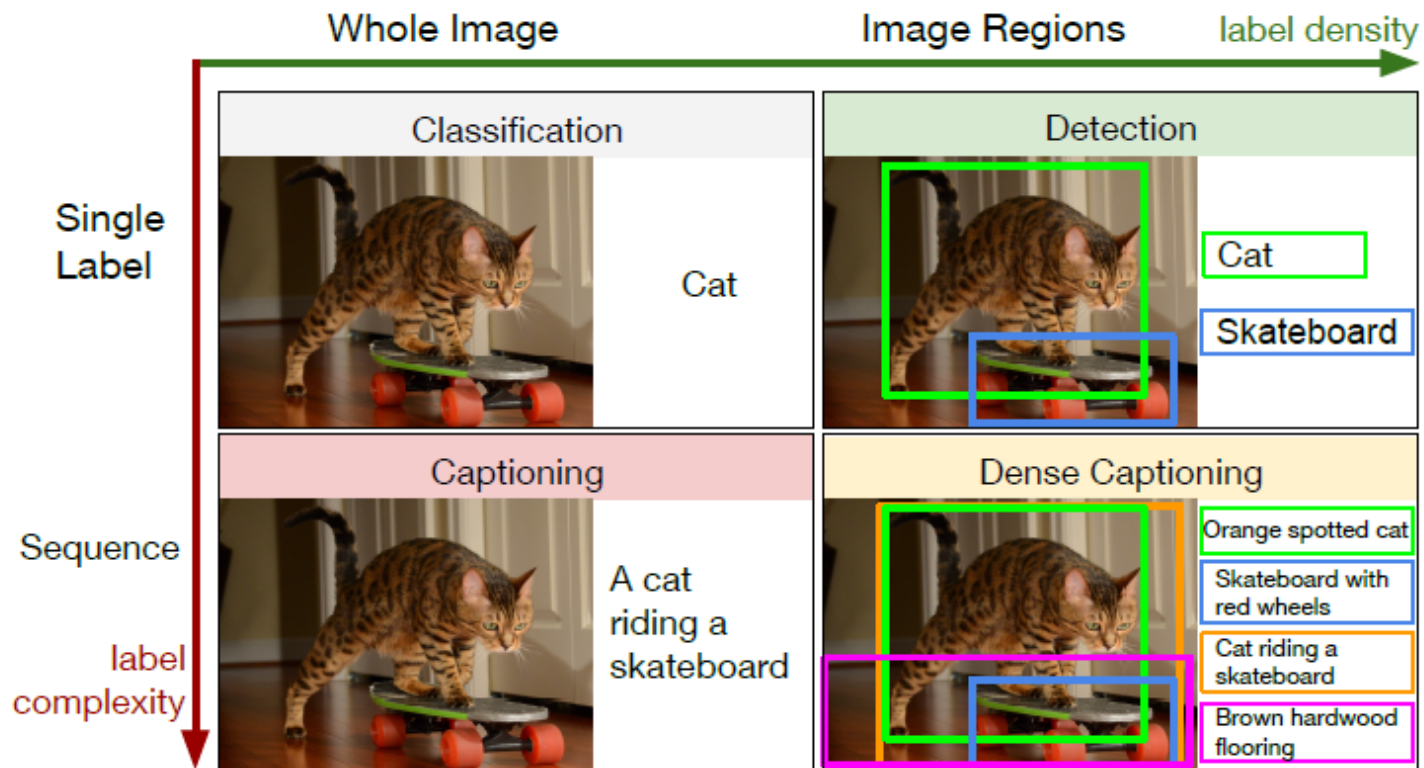
Figure 1. We address the Dense Captioning task (bottom right)

- **Use bi-linear interpolation that is differentiable and can be used for back-propagation**

SIGGRPH

- **Focus more on useful applications**
  - **Wow factor is important**

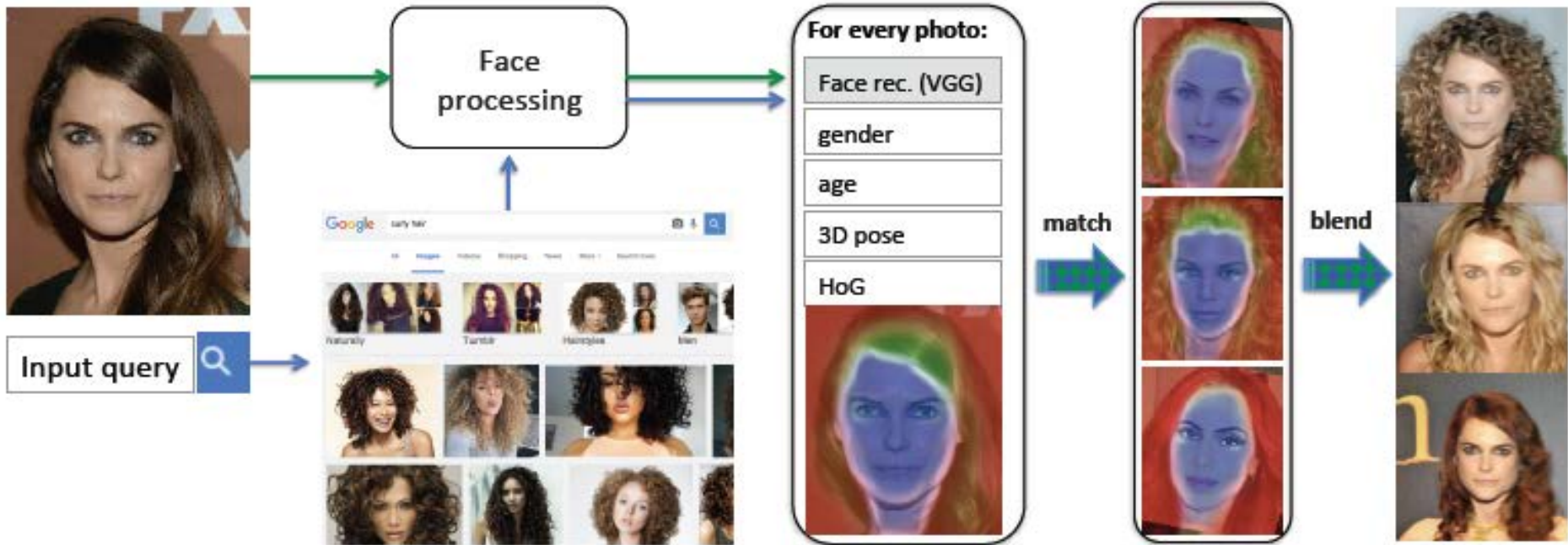# Example: Transfiguring Portraits [SIG. 16]

input

"curly hair"

"india"

"1930"

KAIST

Overall System

**Various feature extractions (vision tech.)**

Image process tech.

Input image & text

Search tech.

ACM Multimedia and ACM ICMR

- **ICMR (Multimedia retrieval)**
  - A recently created conf. since 2011
  - Many papers on image/video search and analysis
- **IEEE multimedia**
  - The top-tier conf. in multimedia
  - Many different topics related to image/video

KAIST

Example: MindFinder, Finding Images by Sketching

- **Sketch-based Image Retrieval via Shape Words. ICMR 2015**

- **Representation for Sketch-Based 3D Model Retrieval. IEEE Signal Processing Letters, 2014**

- **Indexing Billions of Images for Sketch-based Retrieval. ACM Multimedia 2013**

- **Efficient Image Contour Detection using Edge Prior. ICME 2013**

- **The Scale of Edges, in CVPR 2012**

KAIST
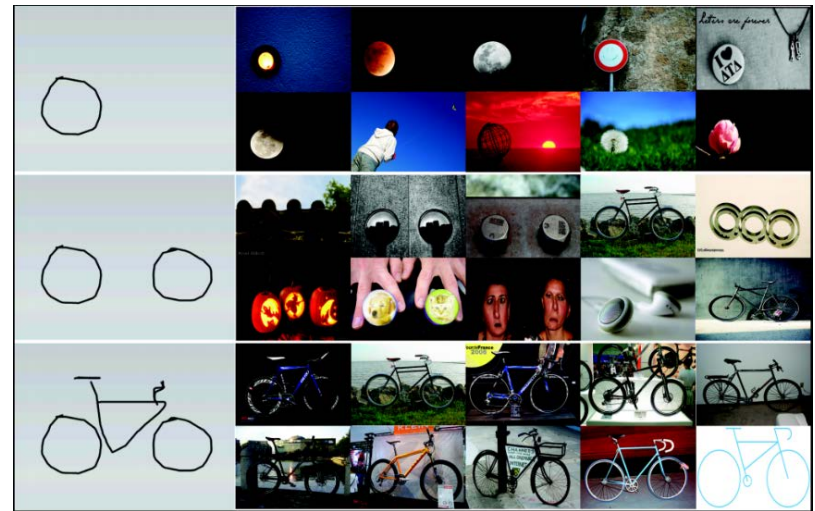
Class Objective (Ch. 2.2)

- **Understand locally invariant features**
  - **Key point localization**
  - **Harris detector**

KAIST

Content-Based Image Retrieval (CBIR)

- **Identify similar images given a user-specified image or other types of inputs**

Extract image descriptors (e.g., SIFT)

Web-scale image database

Input
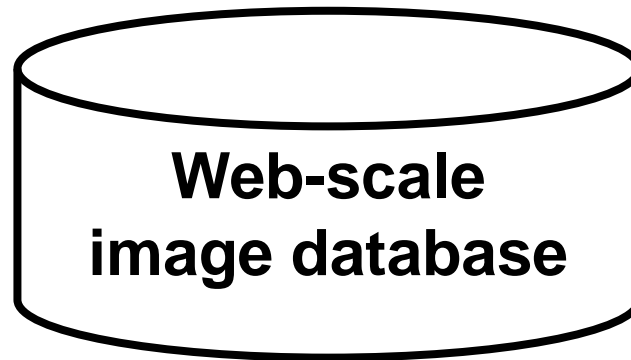
Output

KAIST

Key Components of Image Search

- **Image representations**
- **Indexing algorithms**
- **Matching methods**
- **Classification, Localization, etc.**
  - **Apply image search (or nearest neighbor search)**
  - **Data-driven approach**

Offline Stage

Image Crawling → Image Database → Image Representation → Database Indexing

Zhou et al., arxiv

Online Stage

User Intention → Query Formation → Image Representation → Image Scoring → Search Reranking → Retrieval Browsing

ST
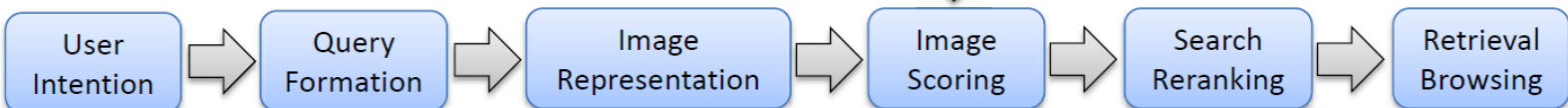
Image Representations

- **SIFT, GIST, CNN, etc.**
  - **Invariant to different transformations**

Quality of Images

- **Search results depend highly on quality of query images**
  - **Suggest re-takes in mobile search**
  - **Select potential candidates from initial results with high precisions in the server side**
  - **Learning to judge image search results, ACM Multimedia**

(b) The top-10 images returned on query "White House".

Quality varies depending on algorithms

Image Retrieval

- **At pre-processing, build a database for efficient retrieval at runtime**
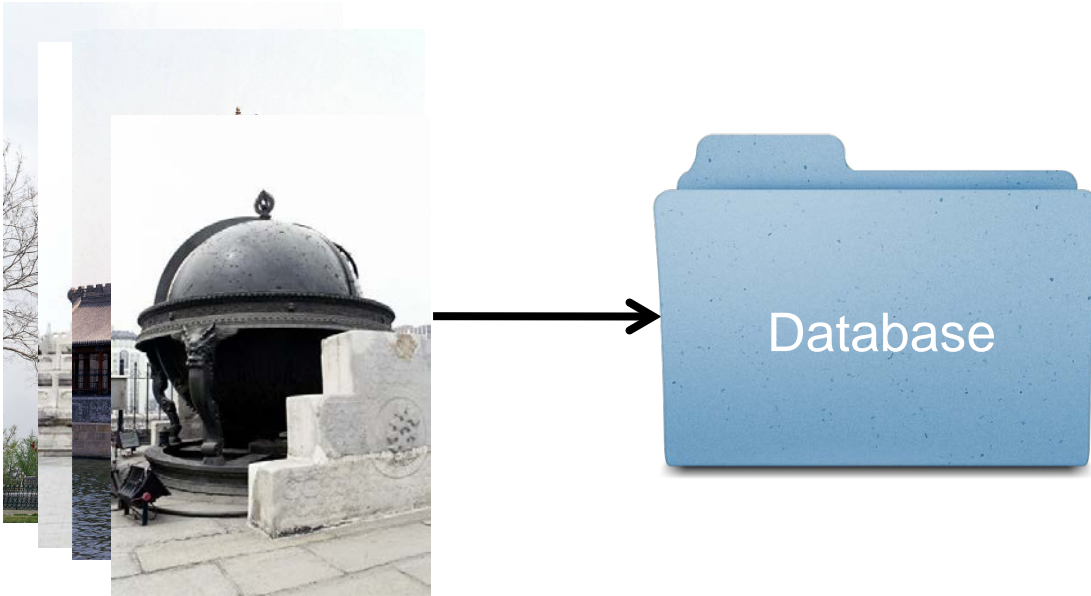
Database

Image Retrieval

- **At pre-processing, build a database for efficient retrieval at runtime**

Database
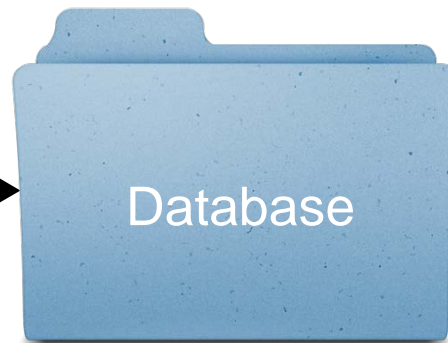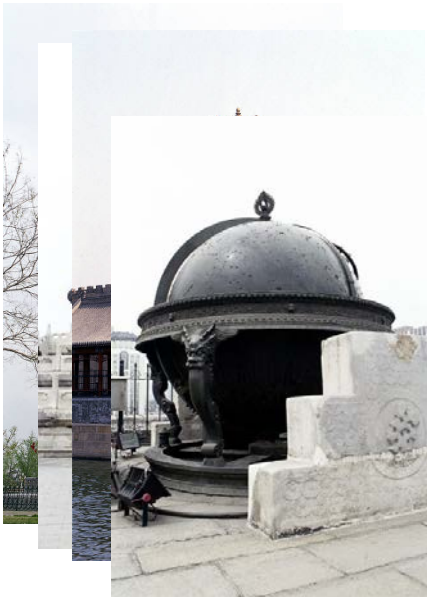
Index schemes: vocabulary trees, hashing, and inverted files
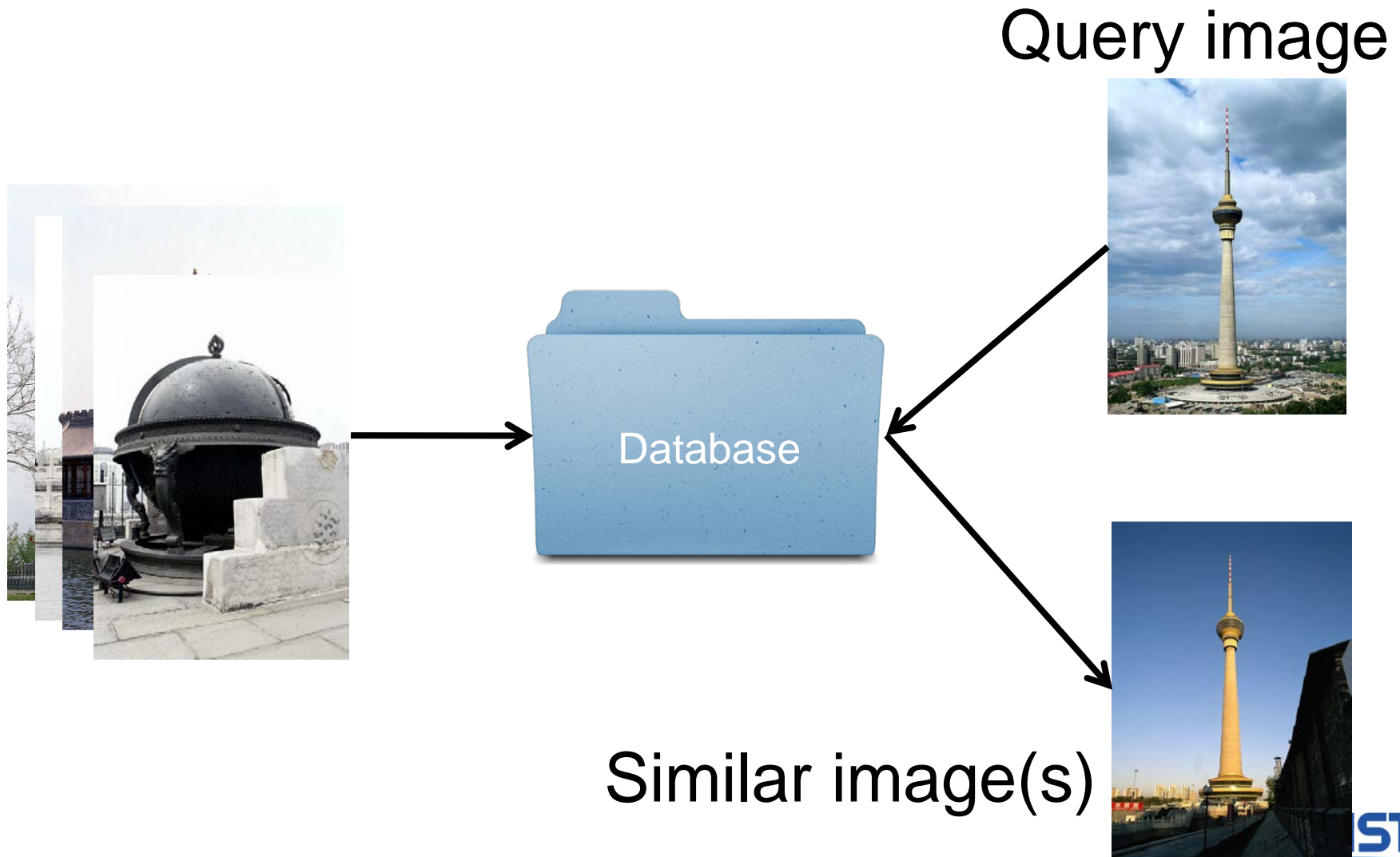
KAIST

Image Retrieval: Runtime Procedure

Query image

Database

KAIST

Image Retrieval: Runtime Procedure

Query image

Database

Similar image(s)
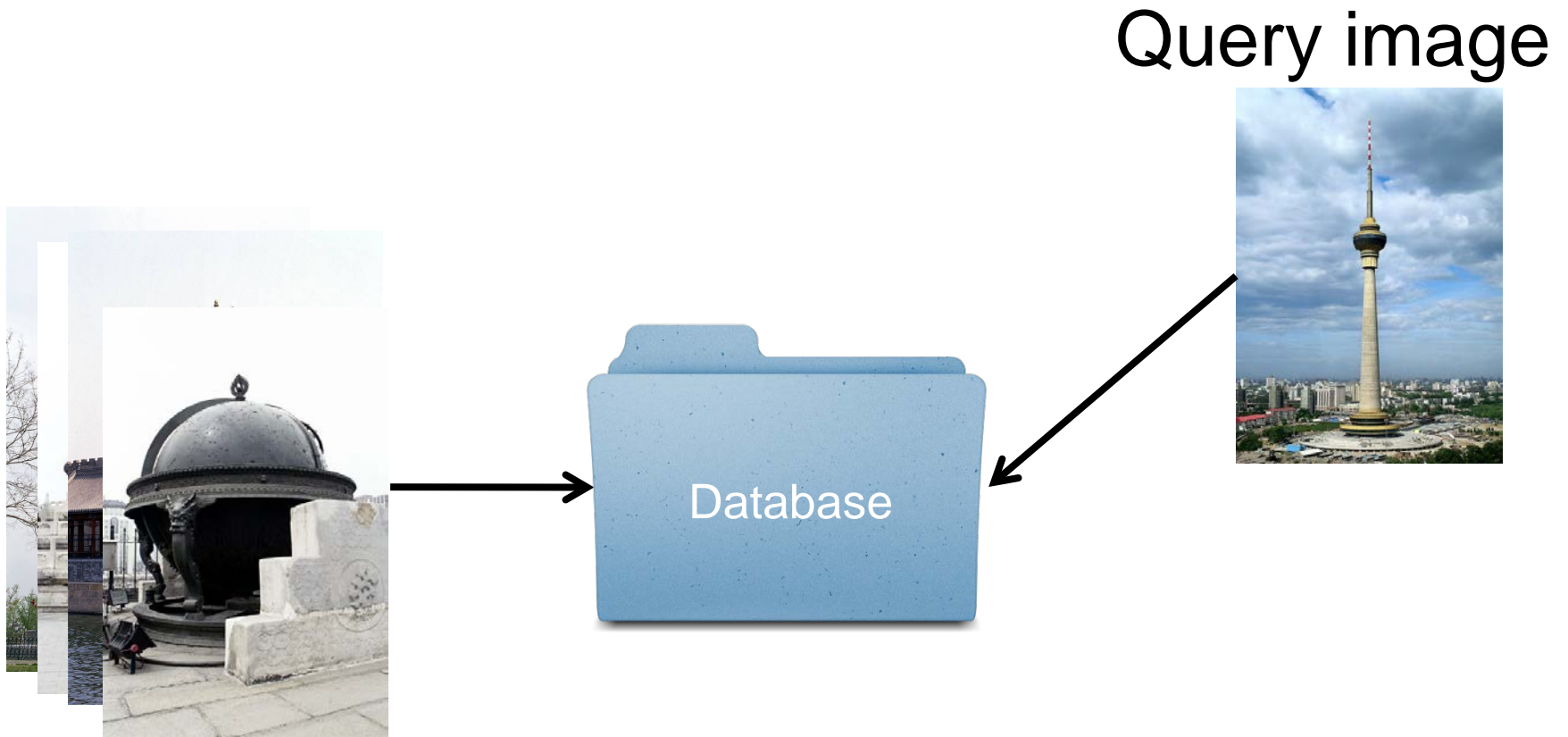
Post-Processing

Query image

Shortlist (e.g., 100 images)

Re-ranking
(spatial verification)

Motivation

- Global representations have major limitations
- Instead, describe and match only local regions
- Increased robustness to
  - Occlusions

  - Articulation
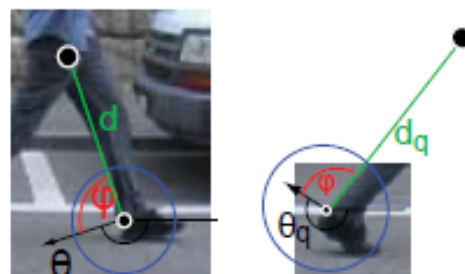
  - Intra-category variations
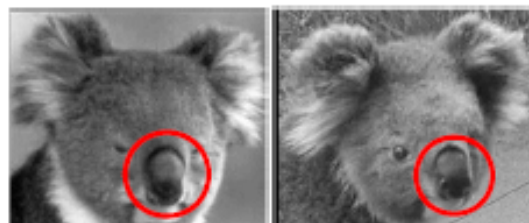
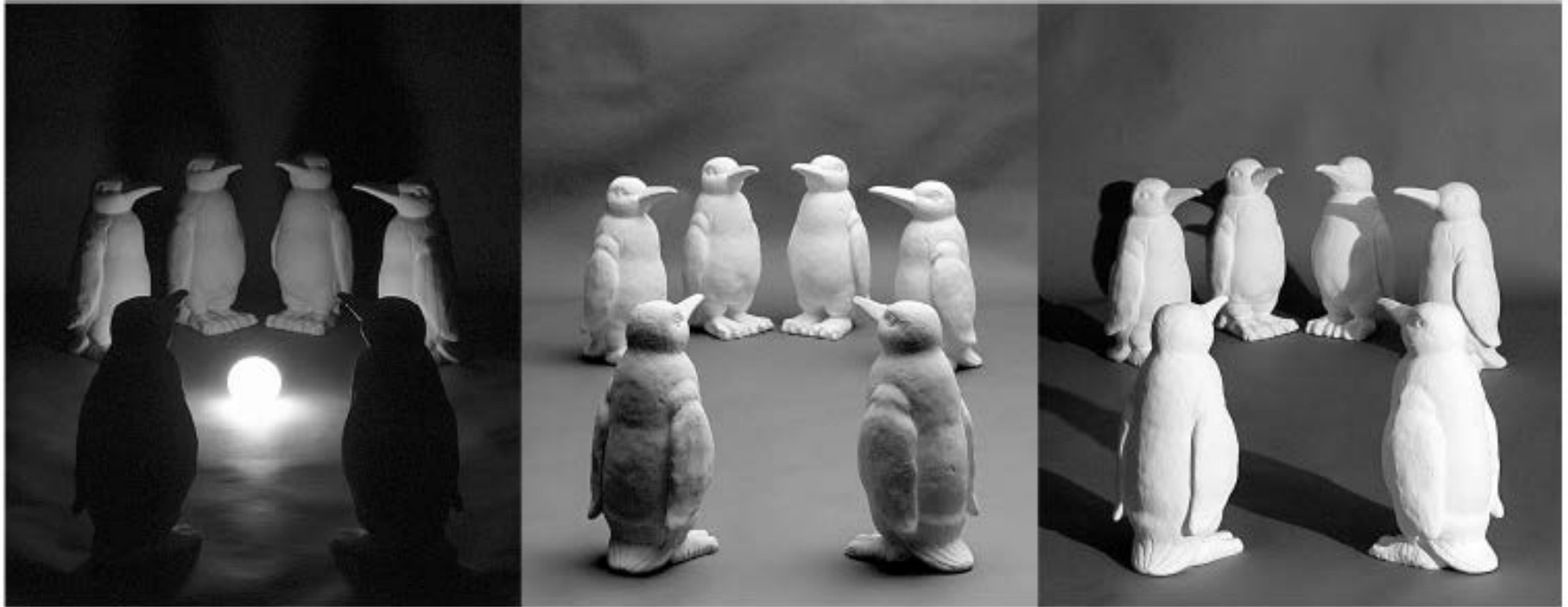# Challenges: viewpoint variation

Michelangelo 1475-1564

Fei-Fei Li

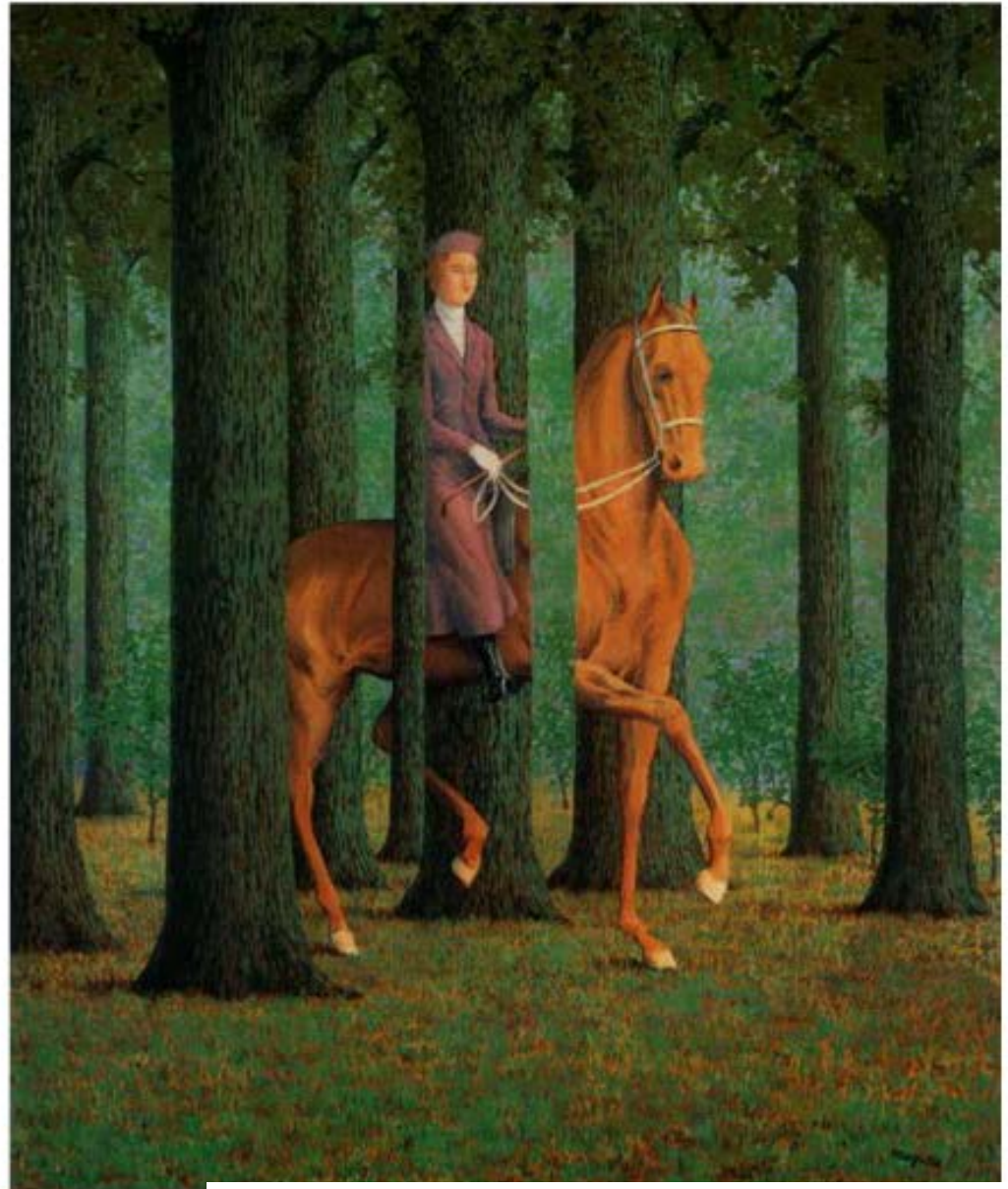# Challenges: illumination

image credit: J. Koenderink

Fei-Fei Li

Challenges: scale

Challenges: deformation

Challenges:
occlusion

Magritte, 1957

Challenges: background clutter

Kilmeny Niland. 1995

Challenges: intra-class variation

Application: Image Matching

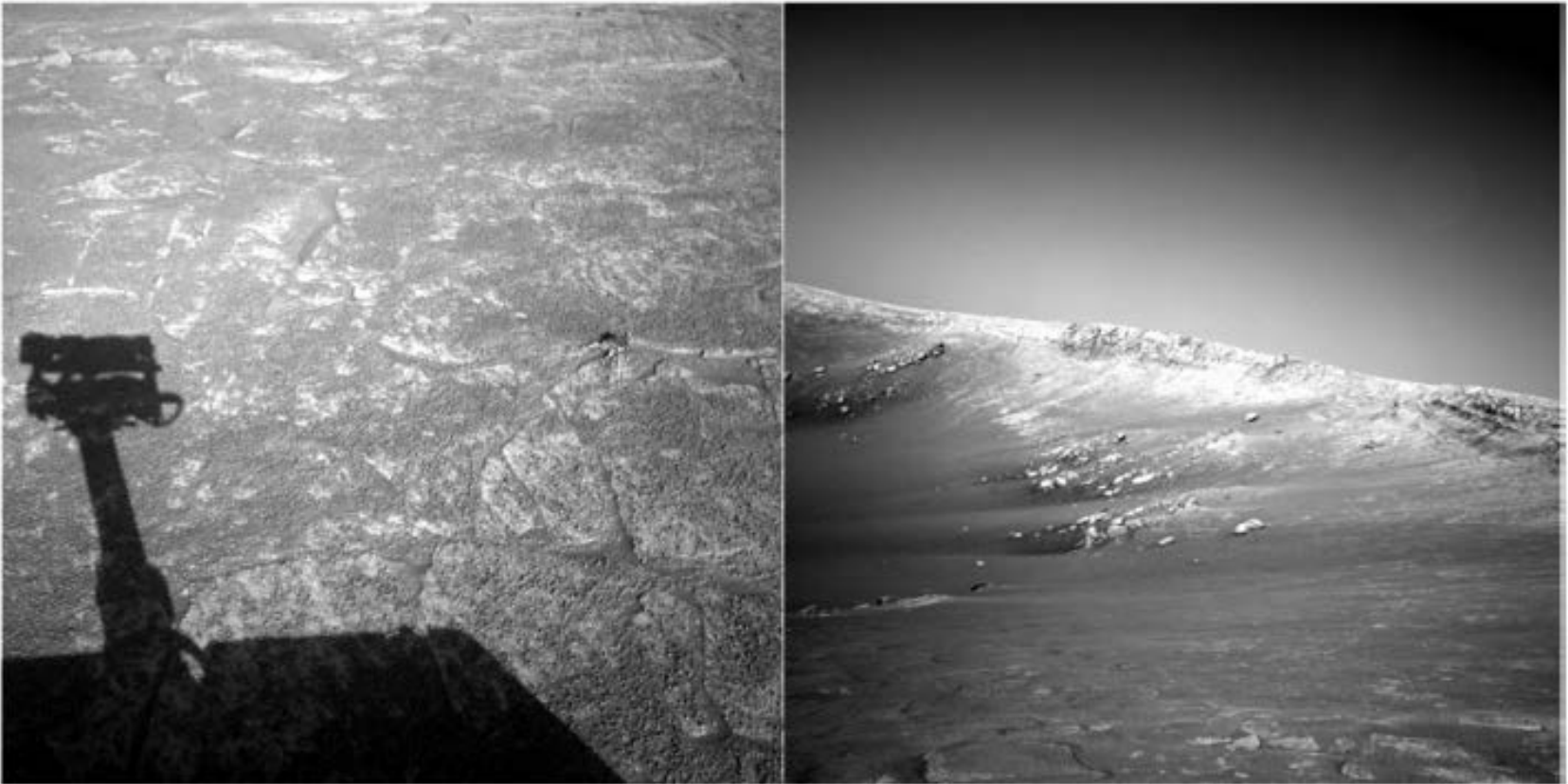by **Diva Sian**

by **swashford**

Slide credit: Steve Seitz

Fei-Fei Li

Harder Case

by **Diva Sian**

by **scgbt**

Fei-Fei Li

Harder Still?

NASA Mars Rover images

Fei-Fei Li

Application: Image Stitching

Fei-Fei Li

Application: Image Stitching

- Procedure:
  - Detect feature points in both images

Fei-Fei Li
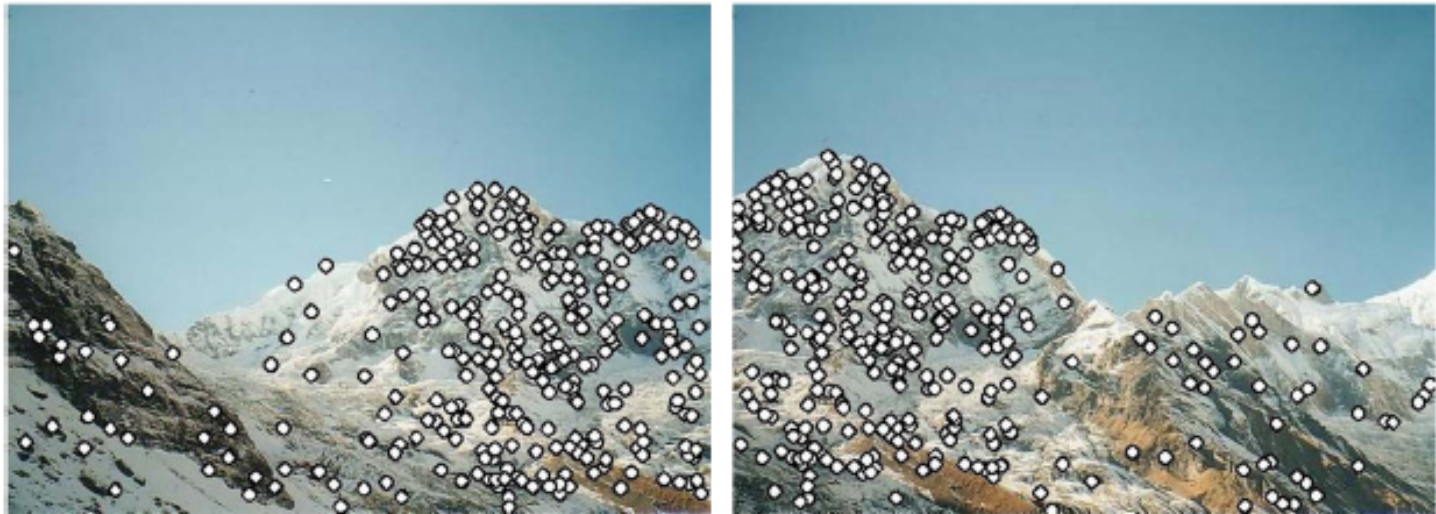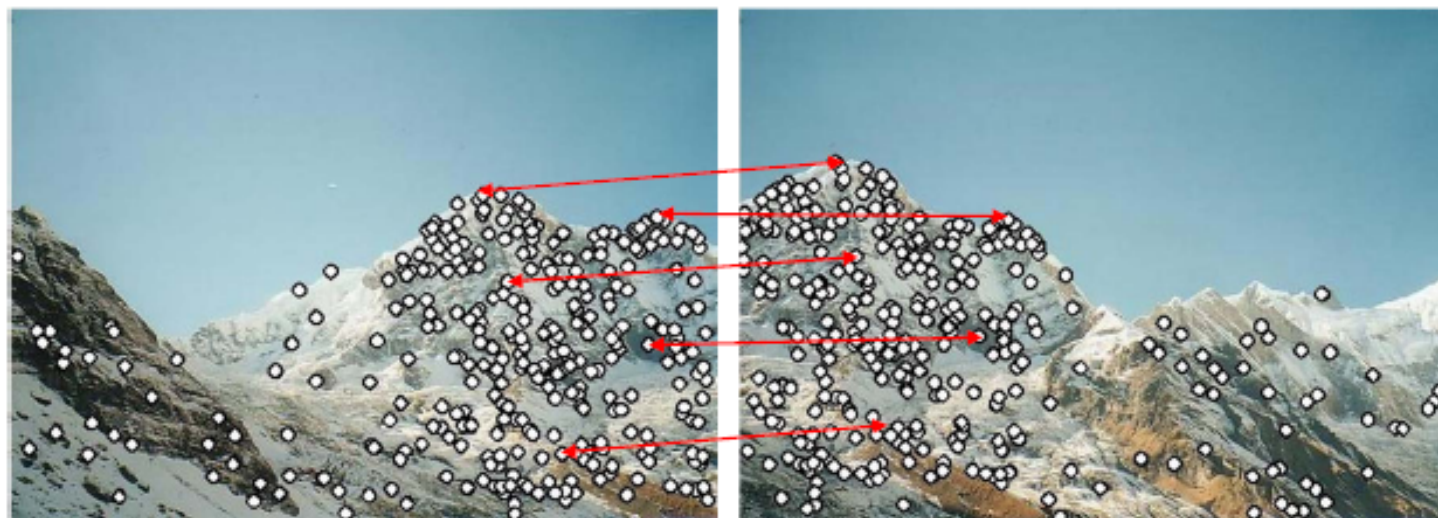
Application: Image Stitching

- Procedure:
  - Detect feature points in both images
  - Find corresponding pairs

Application: Image Stitching

- Procedure:
  - Detect feature points in both images
  - Find corresponding pairs
  - Use these pairs to align the images

Common Requirements

- Problem 1:
  - Detect the same point *independently* in both images

No chance to match!

This lecture

We need a repeatable detector!
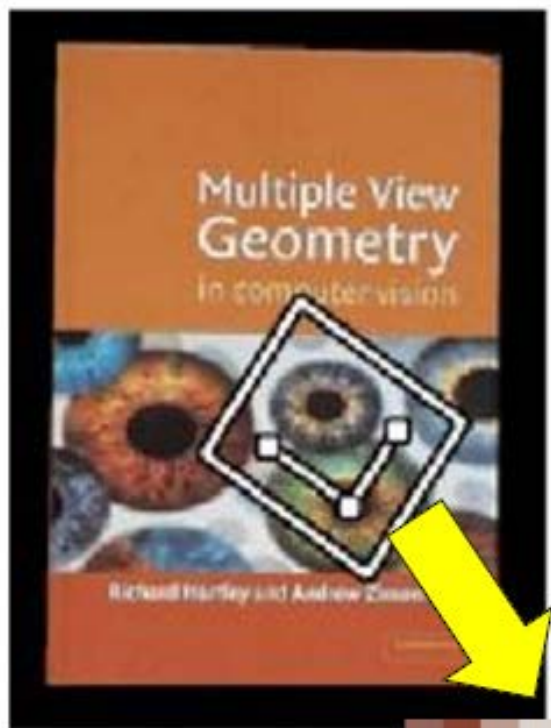
Common Requirements

- Problem 1:
  - Detect the same point *independently* in both images

- Problem 2:
  - For each point correctly recognize the corresponding one
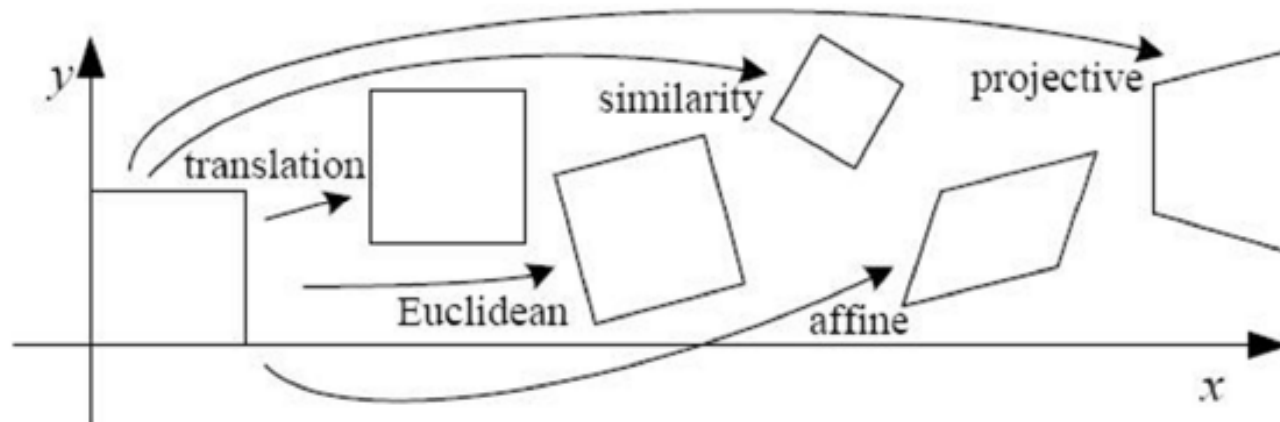
?

Next lecture

We need a reliable and distinctive descriptor!

Fei-Fei Li

Invariance: Geometric Transformations

Fei-Fei Li

Levels of Geometric Invariance

Fei-Fei Li

Invariance: Photometric Transformations

- Often modeled as a linear transformation:
  - Scaling + Offset

Fei-Fei Li

Requirements

- Region extraction needs to be repeatable and accurate
  - Invariant to translation, rotation, scale changes
  - Robust or covariant to out-of-plane ($\approx$affine) transformations
  - Robust to lighting variations, noise, blur, quantization

- Locality: Features are local, therefore robust to occlusion and clutter.

- Quantity: We need a sufficient number of regions to cover the object.

- Distinctivenes : The regions should contain "interesting" structure.

- Efficiency: Close to real-time performance.

Fei-Fei Li

Two Different Directions

- **Classical approaches**
  - **Manually designed in image processing and computer vision fields**

- **Deep learning approaches**
  - **Learned approaches, but are inspired by many prior (manually crafted) approaches**

- **In this class**
  - **We first talk about the classical approaches, followed by deep learning approaches**

KAIST

# Many Existing Detectors Available

- Hessian & Harris   [Beaudet '78], [Harris '88]
- Laplacian, DoG   [Lindeberg '98], [Lowe '99]
- 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...

- *Those detectors have become a basic building block for many recent applications in Computer Vision.*

Fei-Fei Li

Keypoint Localization
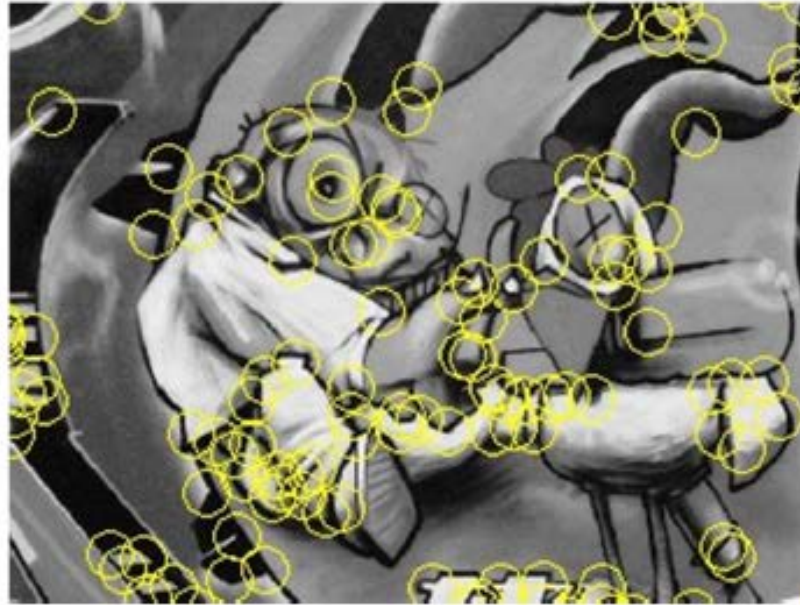
- Goals:
  - Repeatable detection
  - Precise localization
  - Interesting content
 $\Rightarrow$ *Look for two-dimensional signal changes*

Finding Corners
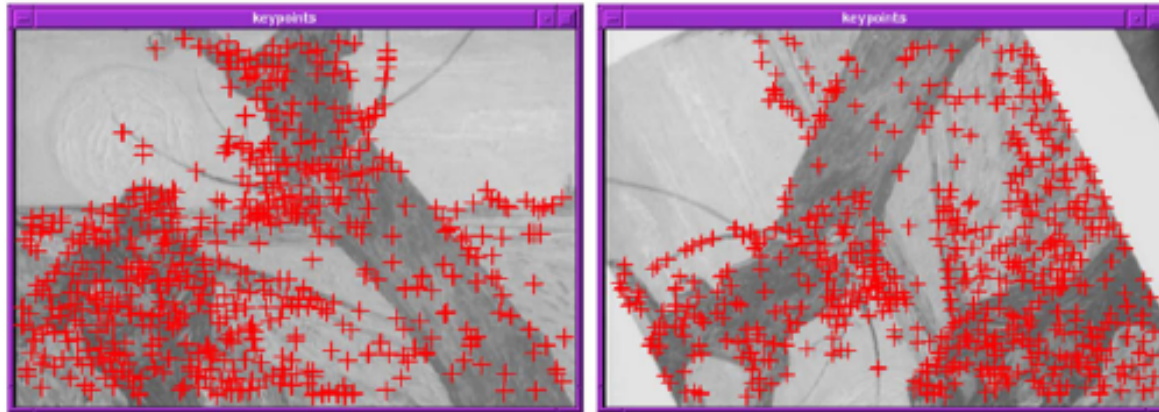
- Key property:
  - In the region around a corner, image gradient has two or more dominant directions
- Corners are *repeatable* and *distinctive*

C.Harris and M.Stephens. "A Combined Corner and Edge Detector."
Proceedings of the 4th Alvey Vision Conference, 1988.

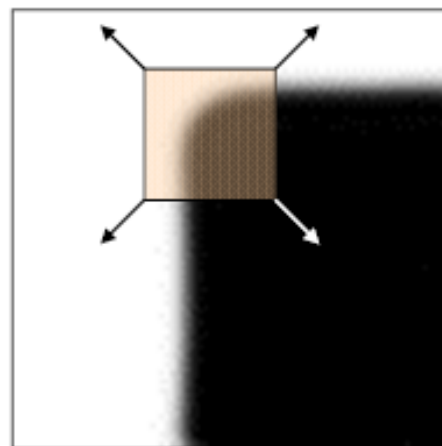Fei-Fei Li

Corners as Distinctive Interest Points

- Design criteria
  - We should easily recognize the point by looking through a small window (*locality*)
  - Shifting the window in *any direction* should give *a large change* in intensity (*good localization*)
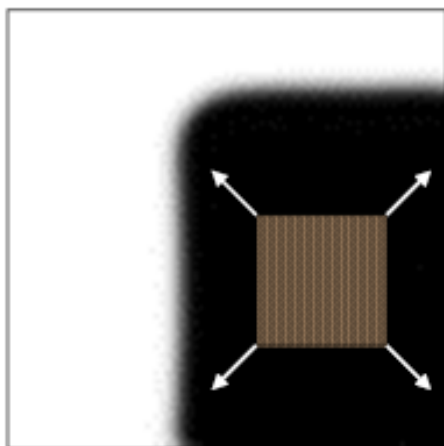
"flat" region:
no change in all
directions
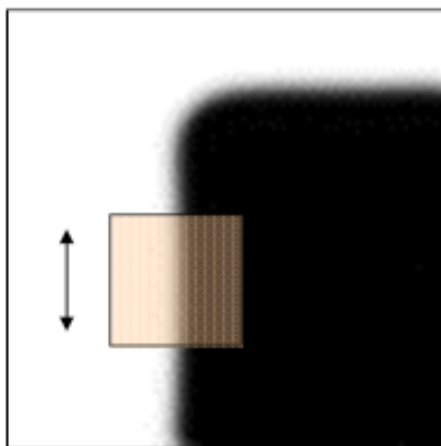
"edge":
no change along
the edge direction

"corner":
significant change
in all directions

Harris Detector Formulation
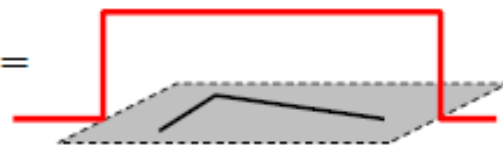
- Change of intensity for the shift [u,v]:

$$E(u,v) = \sum_{x,y} w(x,y)\left[I(x+u, y+v) - I(x,y)\right]^2$$
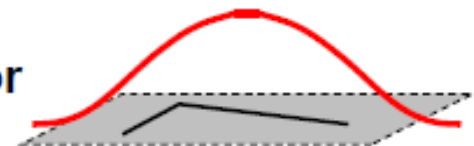
Window function

Shifted intensity

Intensity

Window function $w(x,y) =$
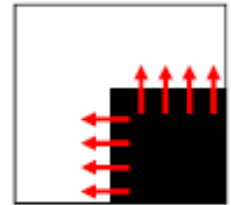
1 in window, 0 outside   or   Gaussian

Fei-Fei Li

What Does This Matrix Reveal?

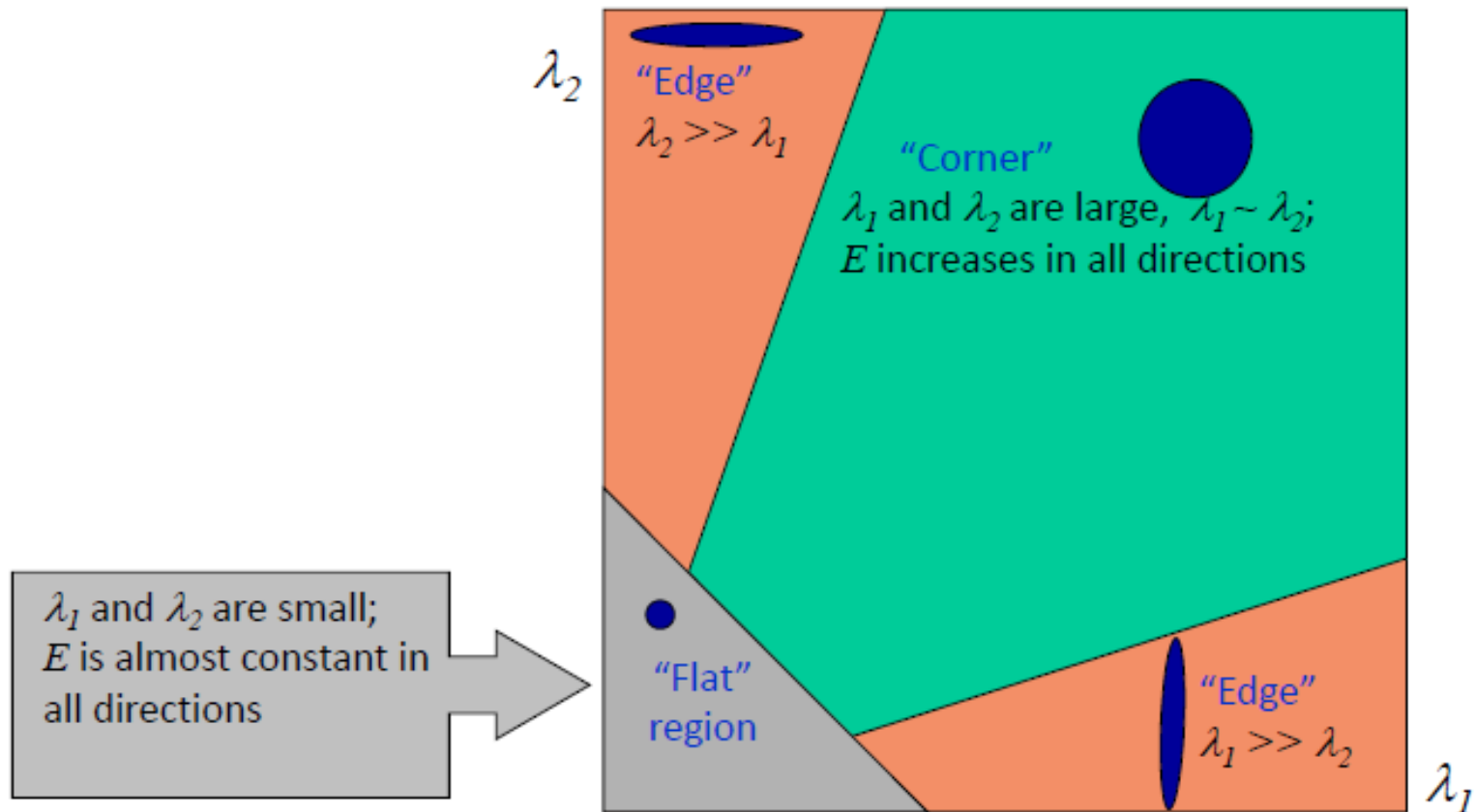- First, let's consider an axis-aligned corner:

$$M = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

- This means:
  - Dominant gradient directions align with x or y axis
  - If either λ is close to 0, then this is not a corner, so look for locations where both are large.

- What if we have a corner that is not aligned with the image axes?

Interpreting the Eigenvalues

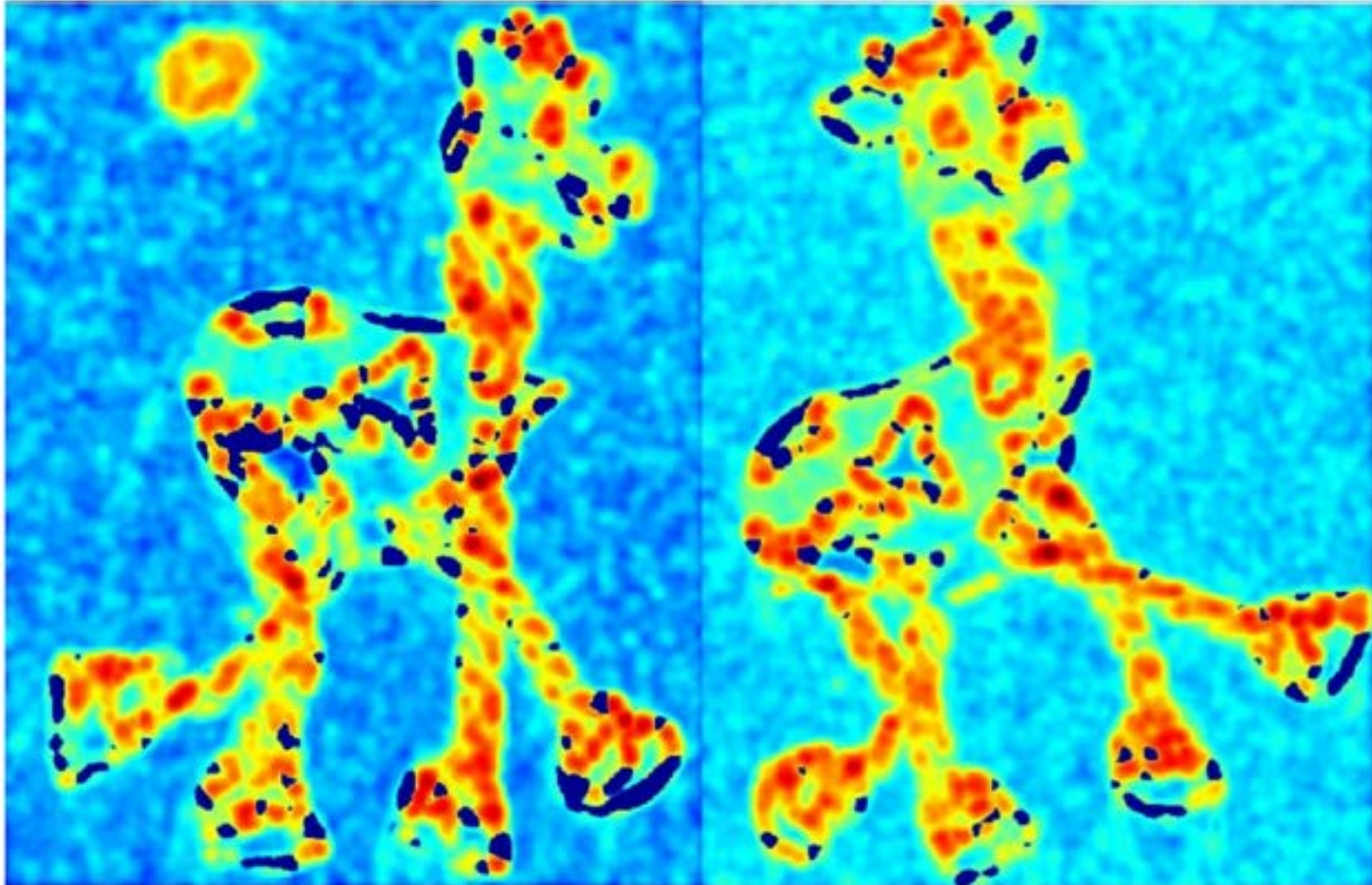- Classification of image points using eigenvalues of M:

λ_2

"Edge"
$\lambda_2 >> \lambda_1$

"Corner"
λ_1 and λ_2 are large, $\lambda_1 \sim \lambda_2$;
E increases in all directions

λ_1 and λ_2 are small;
E is almost constant in all directions

"Flat"
region

"Edge"
$\lambda_1 >> \lambda_2$

λ_1

Fei-Fei Li

Harris Detector: Workflow

Fei-Fei Li

Harris Detector: Workflow
- computer corner responses R

Harris Detector: Workflow

- Take only the local maxima of R, where R>threshold

Fei-Fei Li
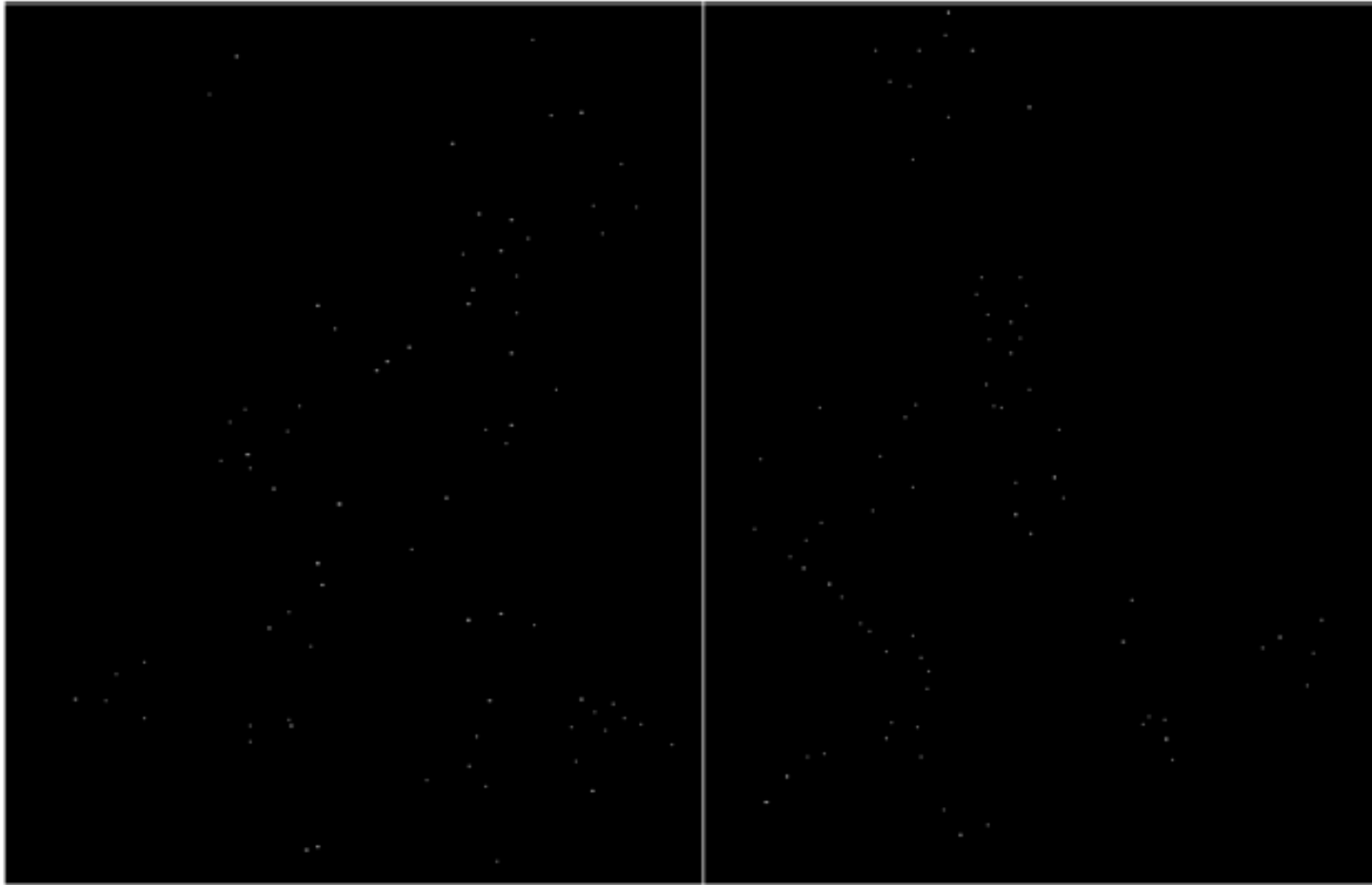
Harris Detector: Workflow
- Resulting Harris points

Fei-Fei Li

Harris Detector – Responses [Harris88]
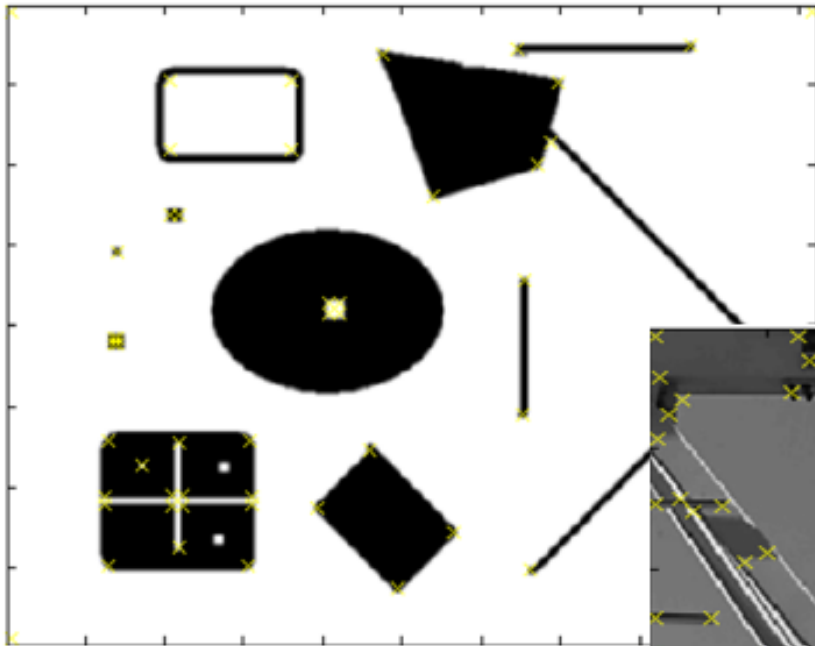
Effect: A very precise corner detector.
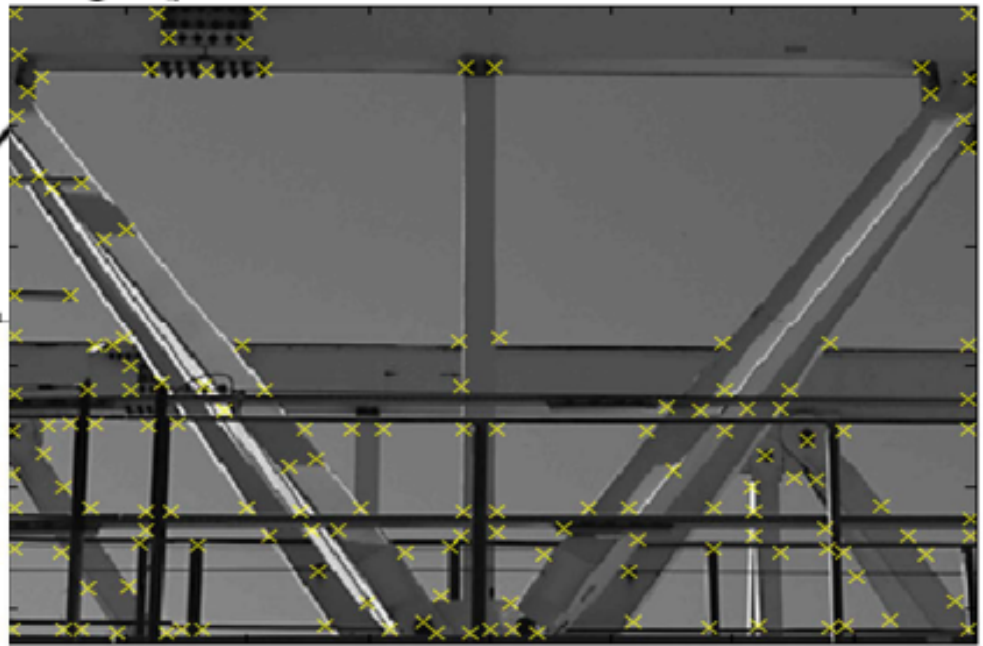
Slide credit: Krystian Mikolajczyk

Fei-Fei Li

Harris Detector – Responses [Harris88]

Fei-Fei Li

Harris Detector – Responses [Harris88]
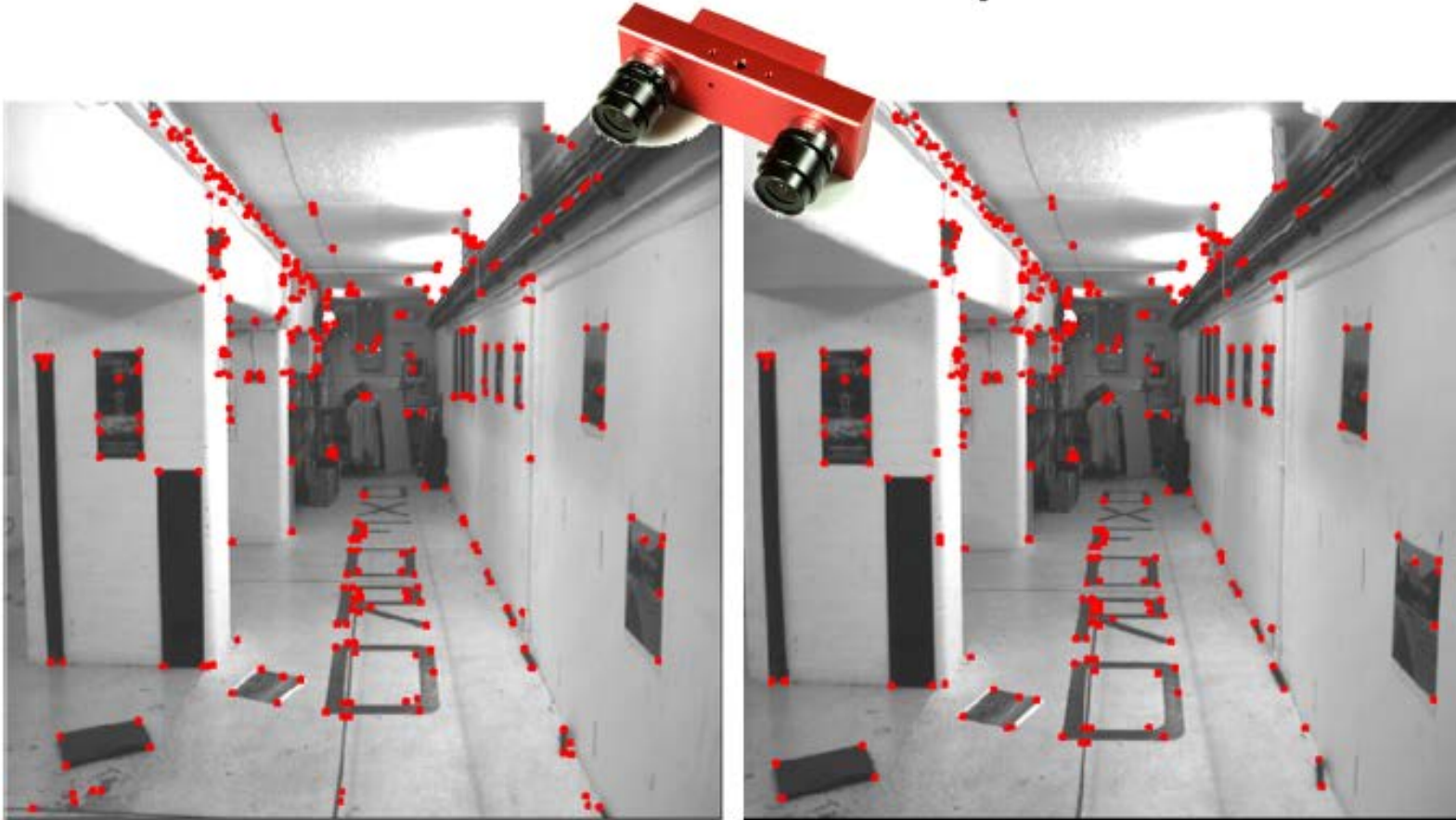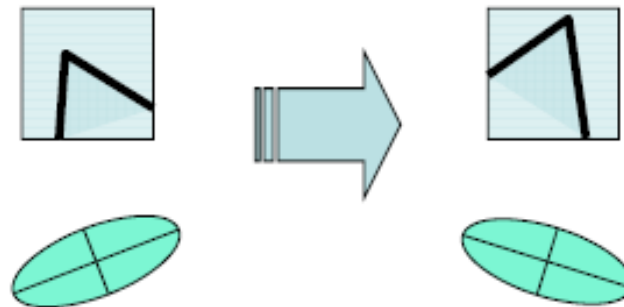

- Results are well suited for finding stereo correspondences

Fei-Fei Li

Harris Detector: Properties

- Rotation invariance?

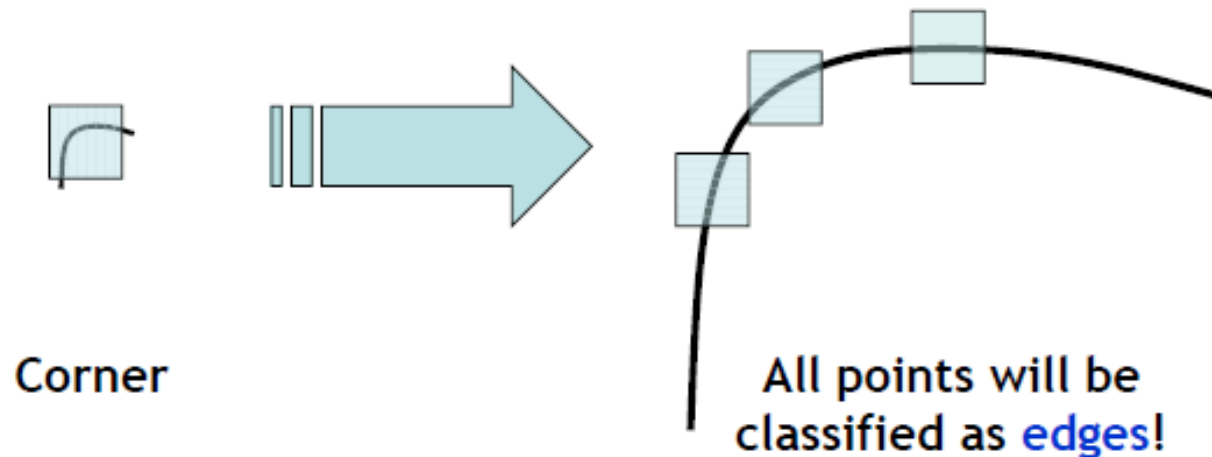Ellipse rotates but its shape (i.e. eigenvalues) remains the same

Corner response R is invariant to image rotation

Harris Detector: Properties

- Rotation invariance

- Scale invariance?

Harris Detector: Properties

- Rotation invariance

- Scale invariance?

Corner

All points will be
classified as **edges**!

Not invariant to image scale!

Fei-Fei Li

Class Objective were:

- **Understand locally invariant features**
  - **Key point localization**
  - **Harris detector: manually designed detector → automatically learned detector using deep learning**

Next Time..

- **Scale invariant region selection**

KAIST

Homework for Every Class

- **Go over the next lecture slides**

- **Come up with one question on what we have discussed today**

- **Go over recent papers on image search, and submit their summary**

KAIST

Homework for Every Class

- Go over recent papers on image search
  - High quality papers: Papers published at the top-tier conf. or close it can be presented; e.g., CVPR, ICCV, ICMR, ECCV, MM, SIGGRAPH
  - Recent publication : papers published since 2015
  - Find and browse two papers, and submit your summary before every beginning of the Thur. class

- Think about possible team members
- Too late if you think them later..

KAIST