CS688: Web-Scale Image Search
Keypoint Localization

Sung-Eui Yoon
(윤성의)

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

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

Homework for Every Class

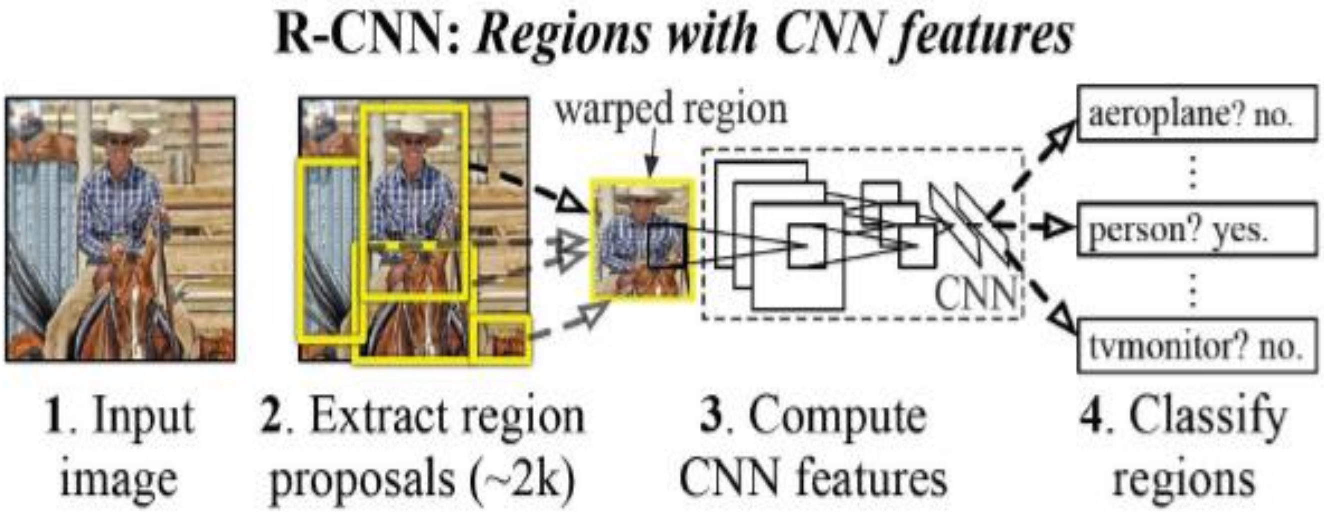
- **Go over recent papers on image search**
  - **High quality papers: Papers published at the top-tier conf. or close it can be presented; e.g., CVPR, ICCV, ECCV, ACM ICMR, ACM MM, ACM SIGGRAPH**
  - **Recent publication: papers published since 2013**
  - **Find and browse two papers, and submit your 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..**

Computer Vision Field: CVPR, ICCV, ECCV

- Handle various computer vision problems

- Get various machine learning techniques from ICML

Example: R-CNN [CVPR 14, oral]



- 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

Example: Localization Networks

- **DenseCap, CVPR 16 (oral)**

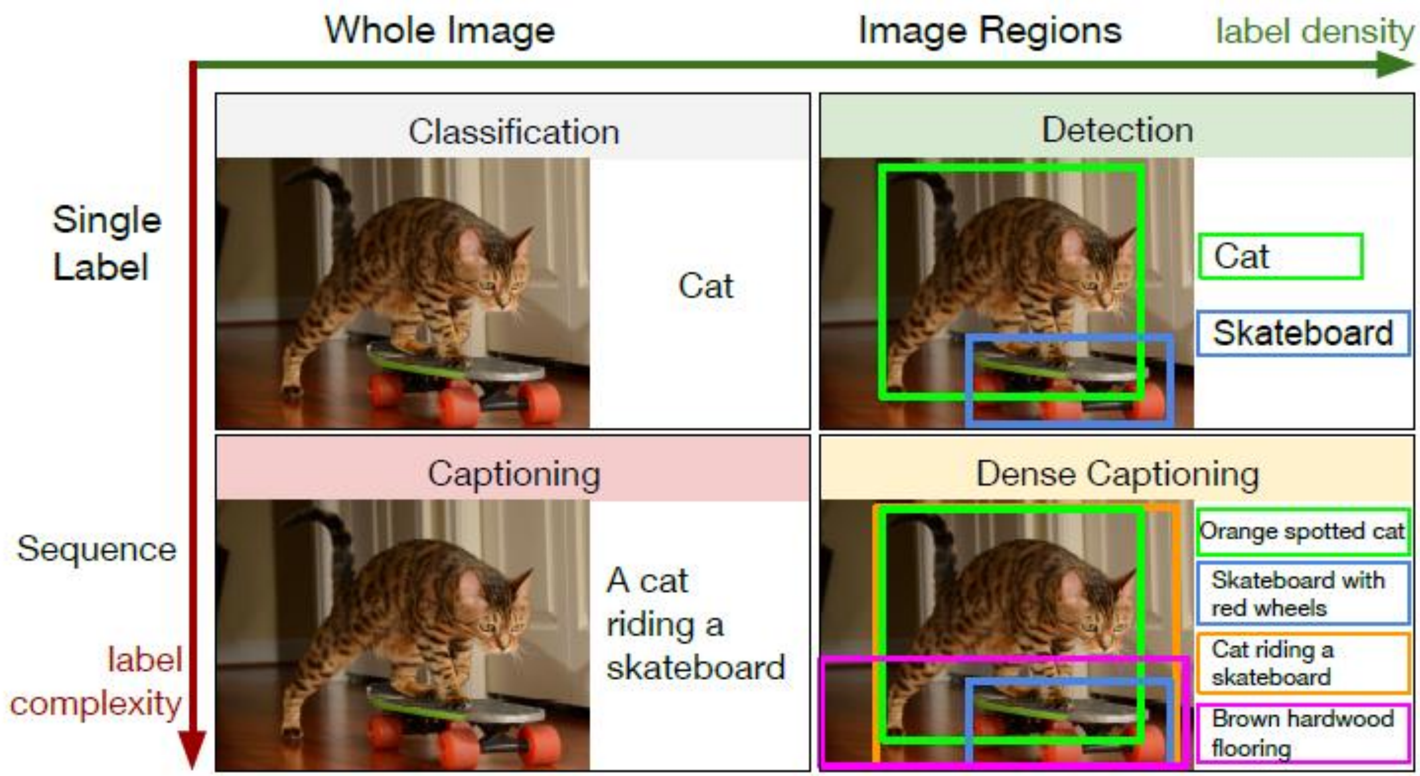
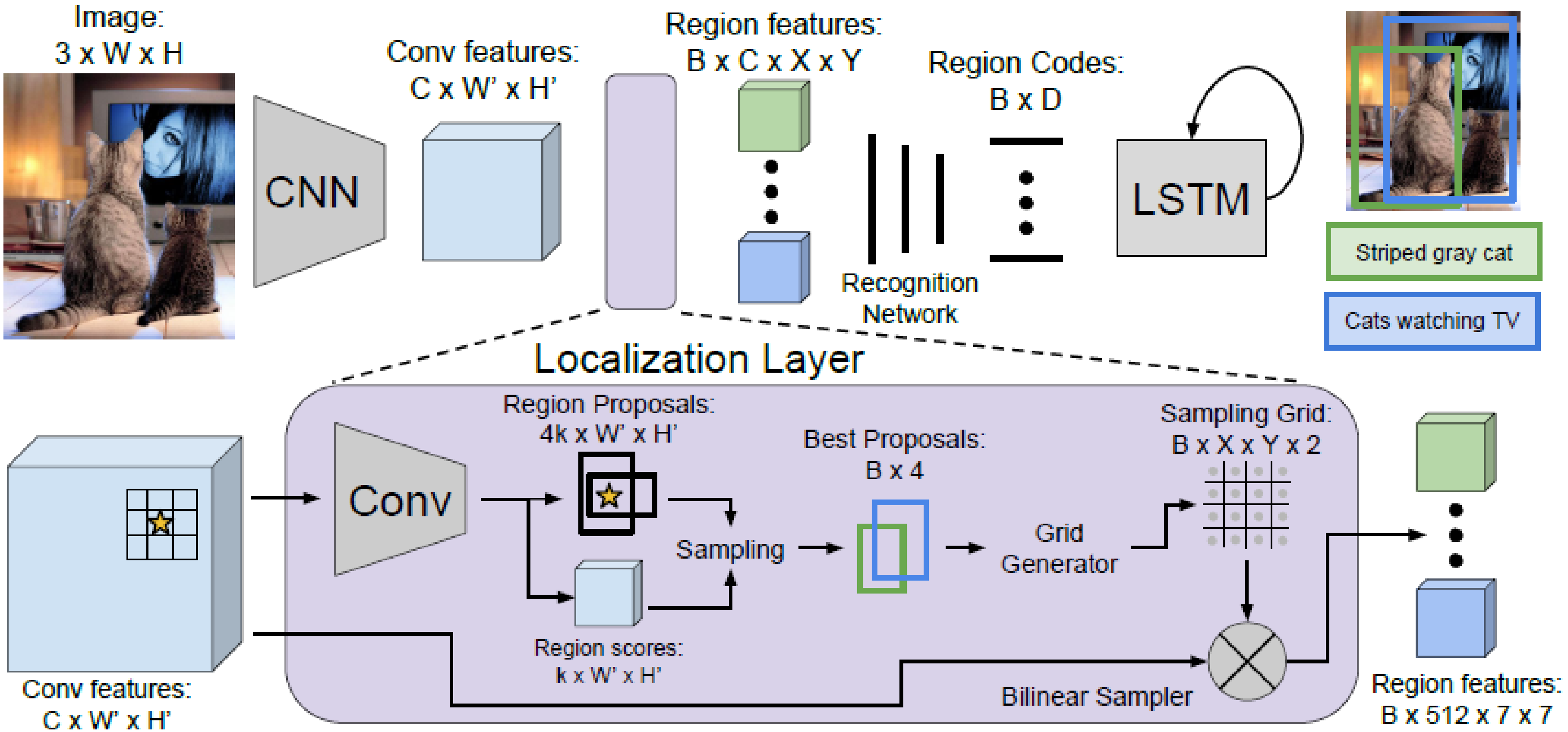


Figure 1. We address the Dense Captioning task (bottom right)

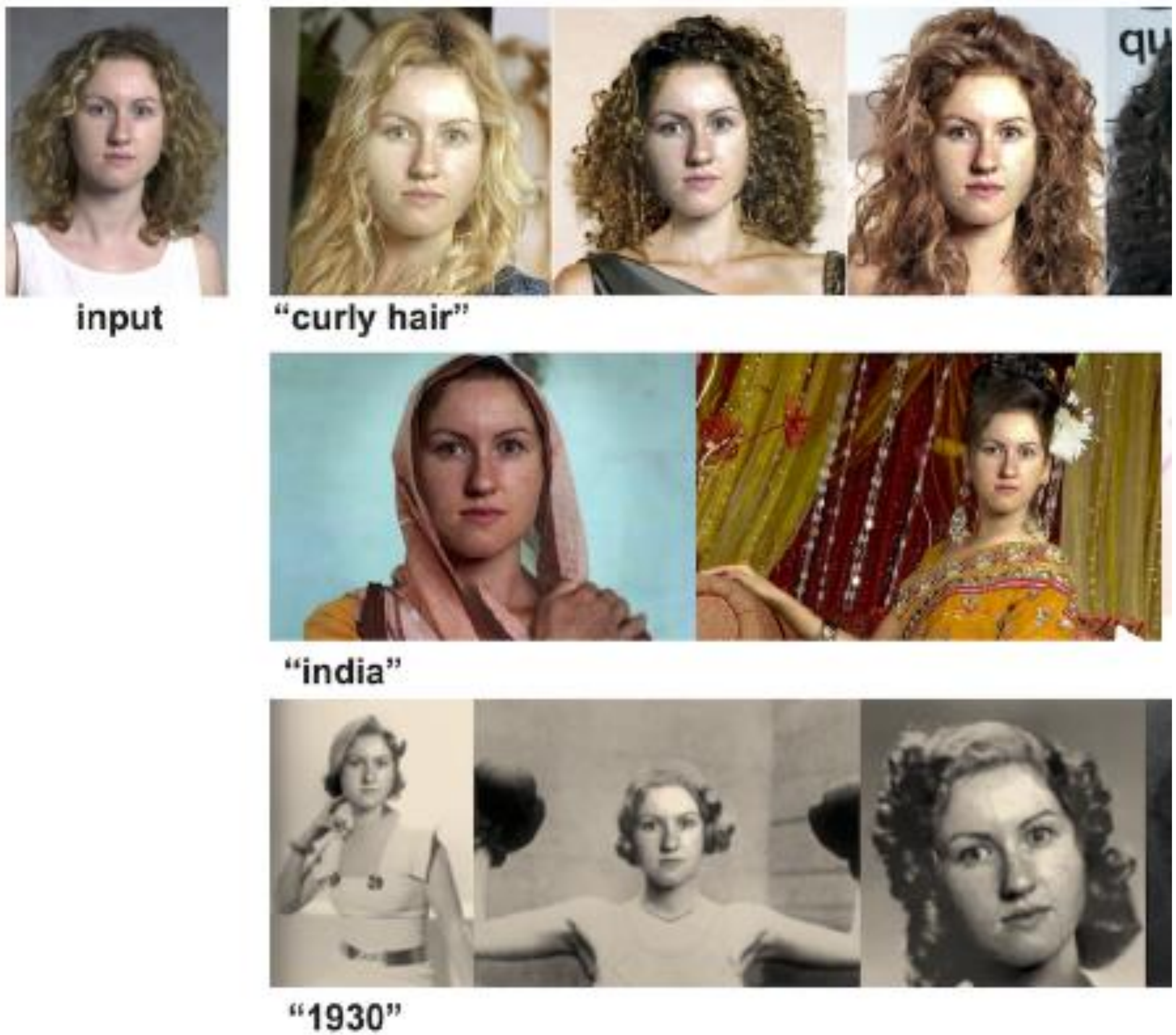


- **Use bi-linear interpolation that is differentiable and can be used for back-propagation**

SIGGRPH

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

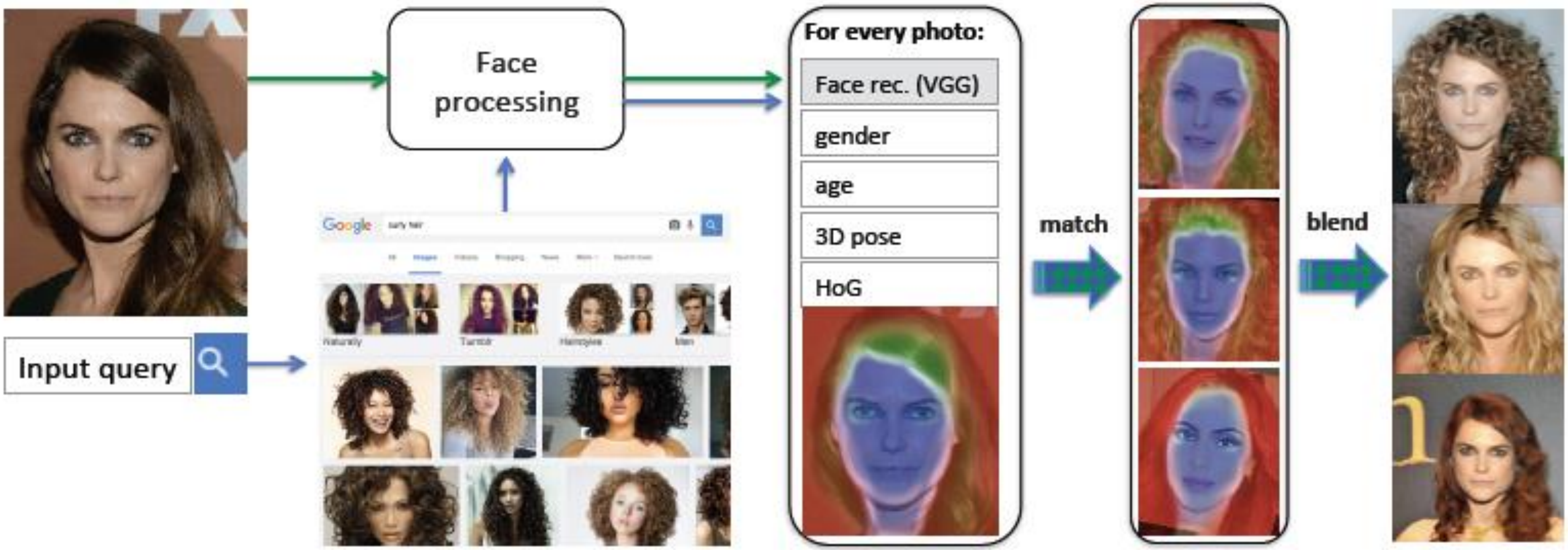
Example: Transfiguring Portraits [SIG. 16]



Overall System

Various feature extractions (vision tech.)

Image process tech.



Input image & text

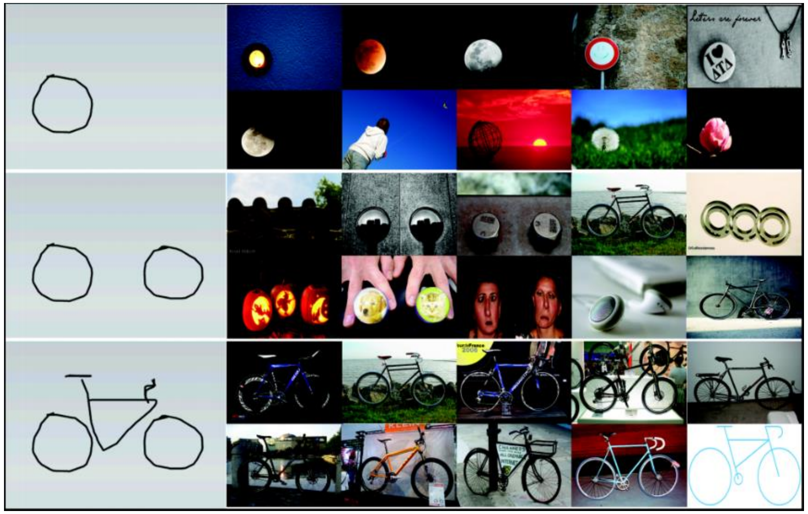
Search tech.

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

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



Class Objective

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

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**

Image Representations

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

Image Retrieval

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

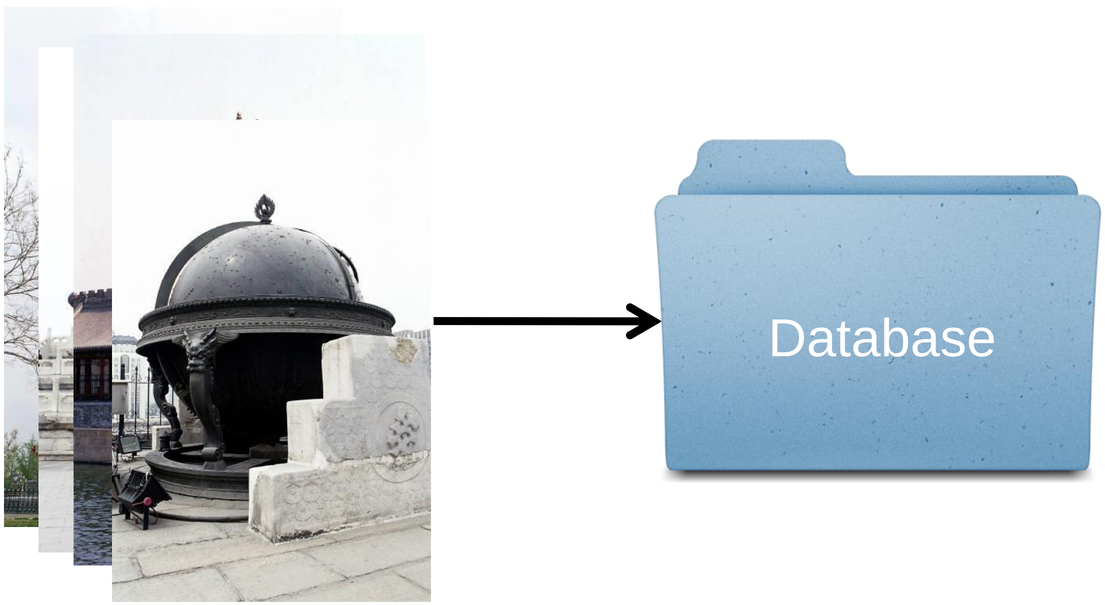
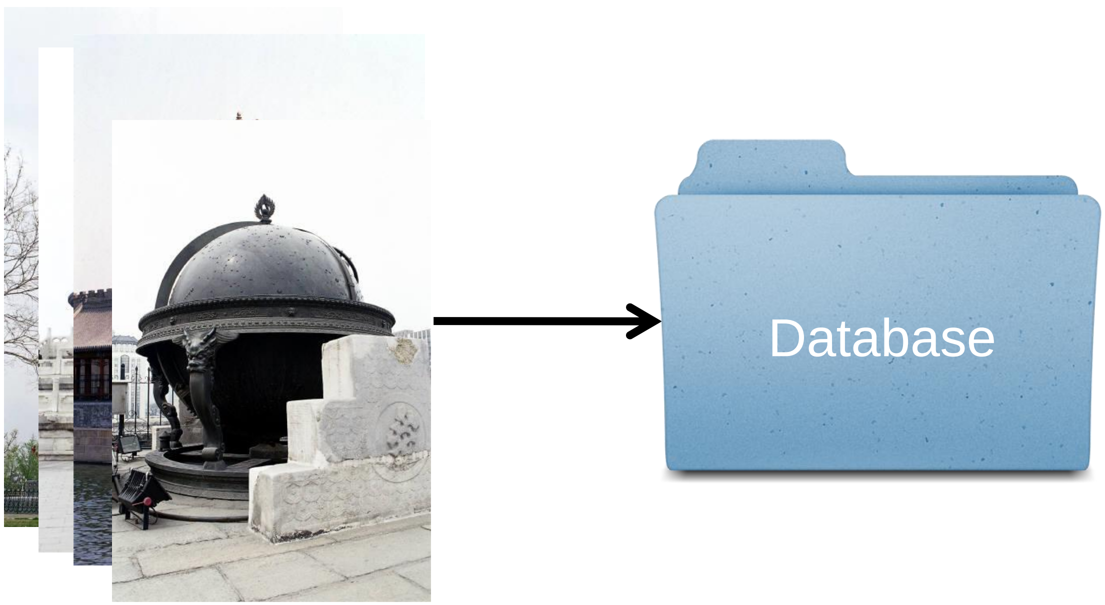


Image Retrieval

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



Index schemes: vocabulary trees, hashing, and inverted files

Image Retrieval: Runtime Procedure

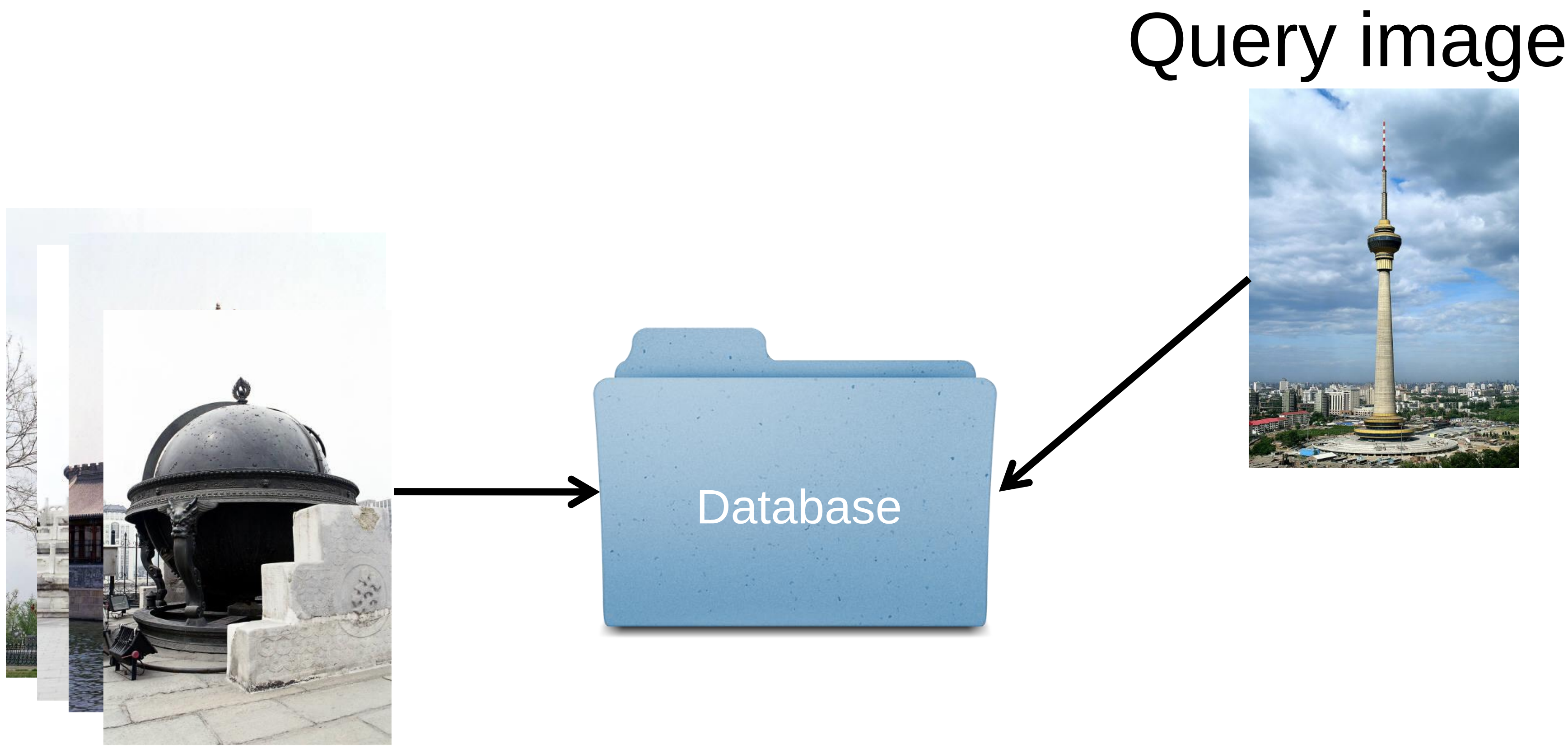
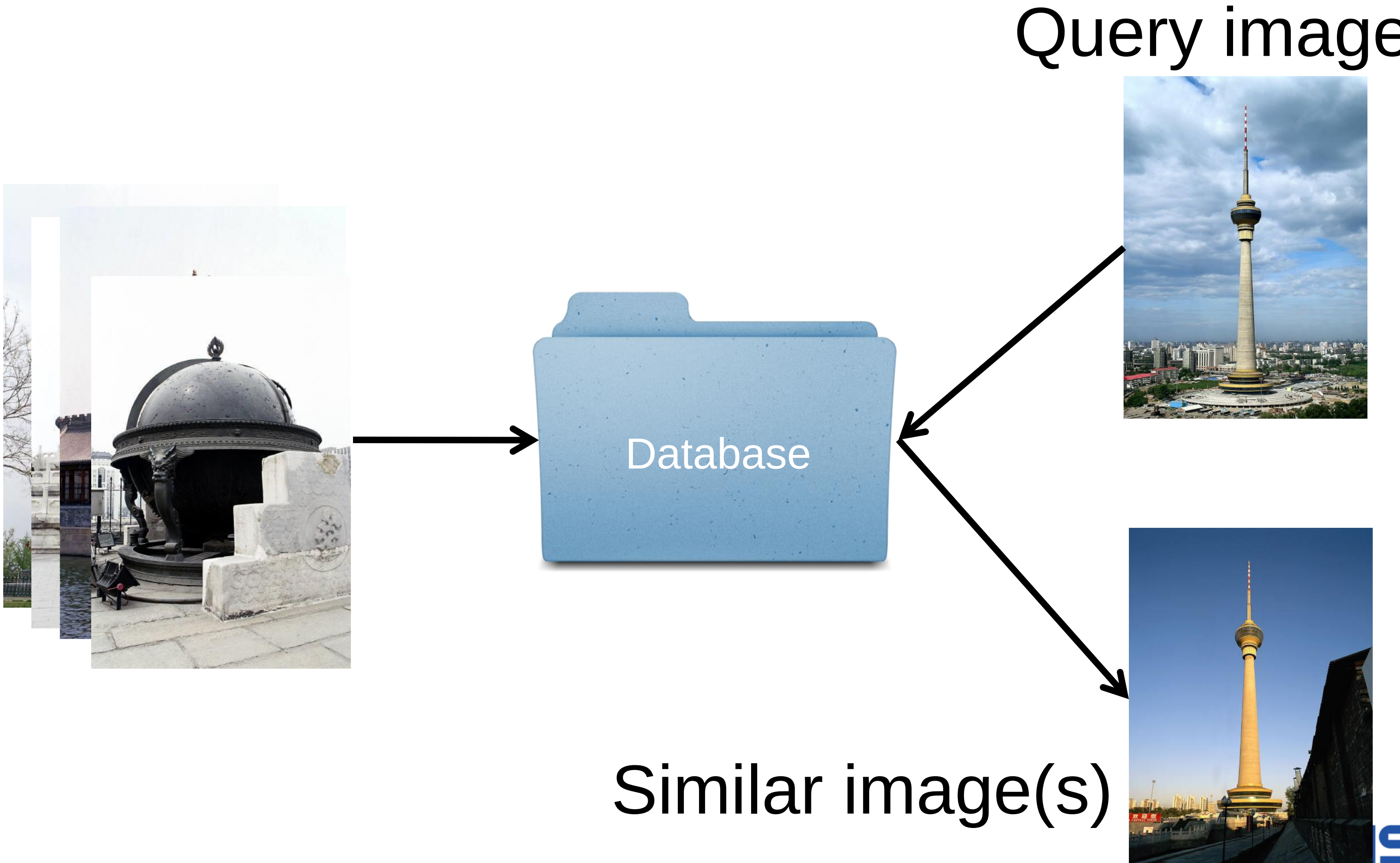


Image Retrieval: Runtime Procedure



Post-Processing

Image Retrieval with Spatially Constrained Similarity Measure

[Xiaohui Shen, Zhe Lin, Jon Brandt, Shai Avidan and Ying Wu, CVPR 2012]

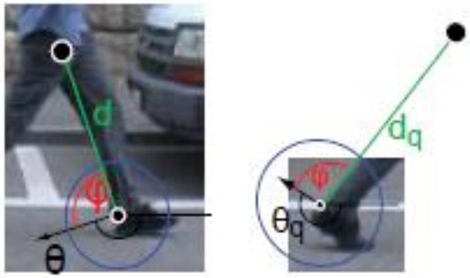
Classification through Image Search

- Image search
  - Find images that have smaller distances to the query
- Classification
  - Fine classes that have smaller distances to the query
  - Utilize labels
- Classification using image search
  - Naïve Bayes Nearest Neighbor (NBNN)
  - Image classification and Retrieval are ONE

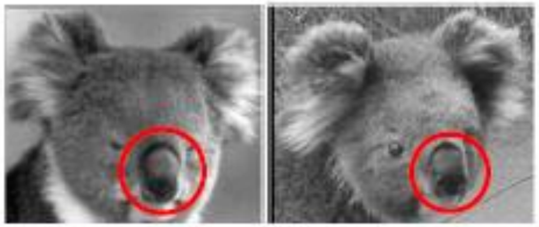
Motivation

- 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

Challenges: illumination

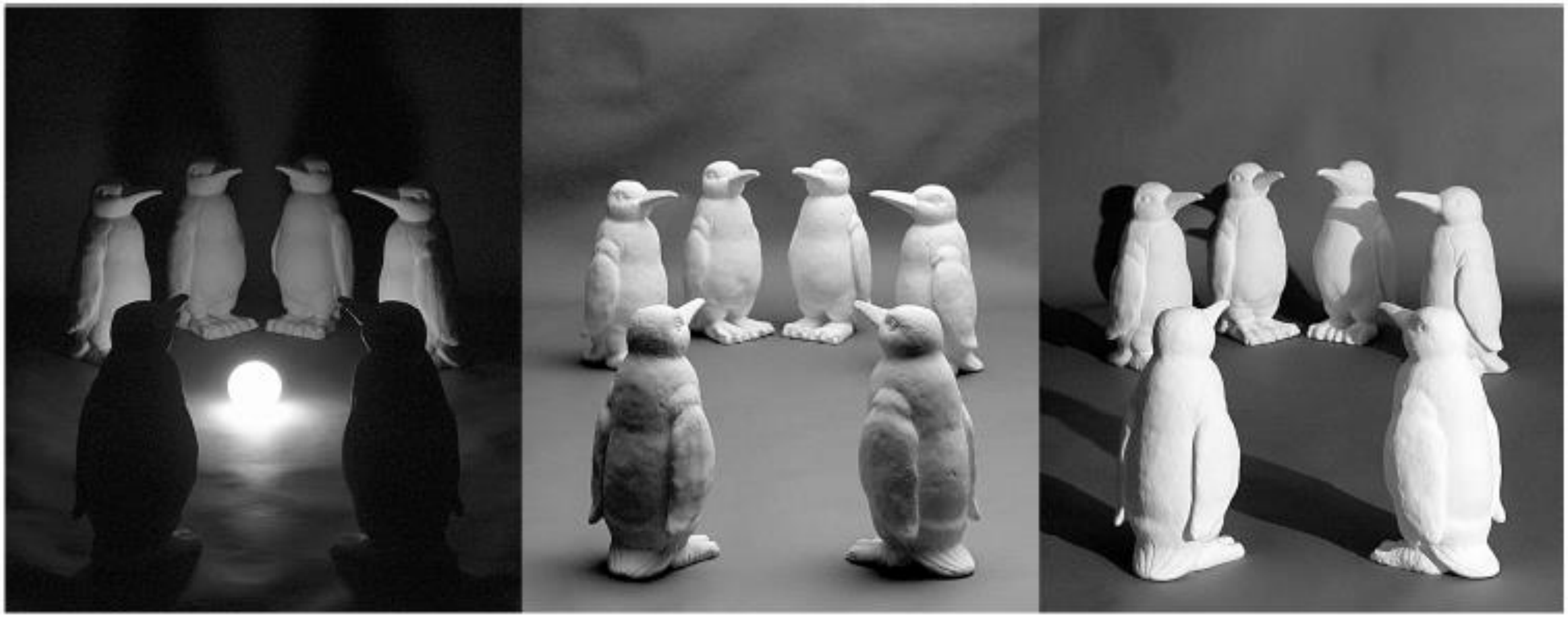
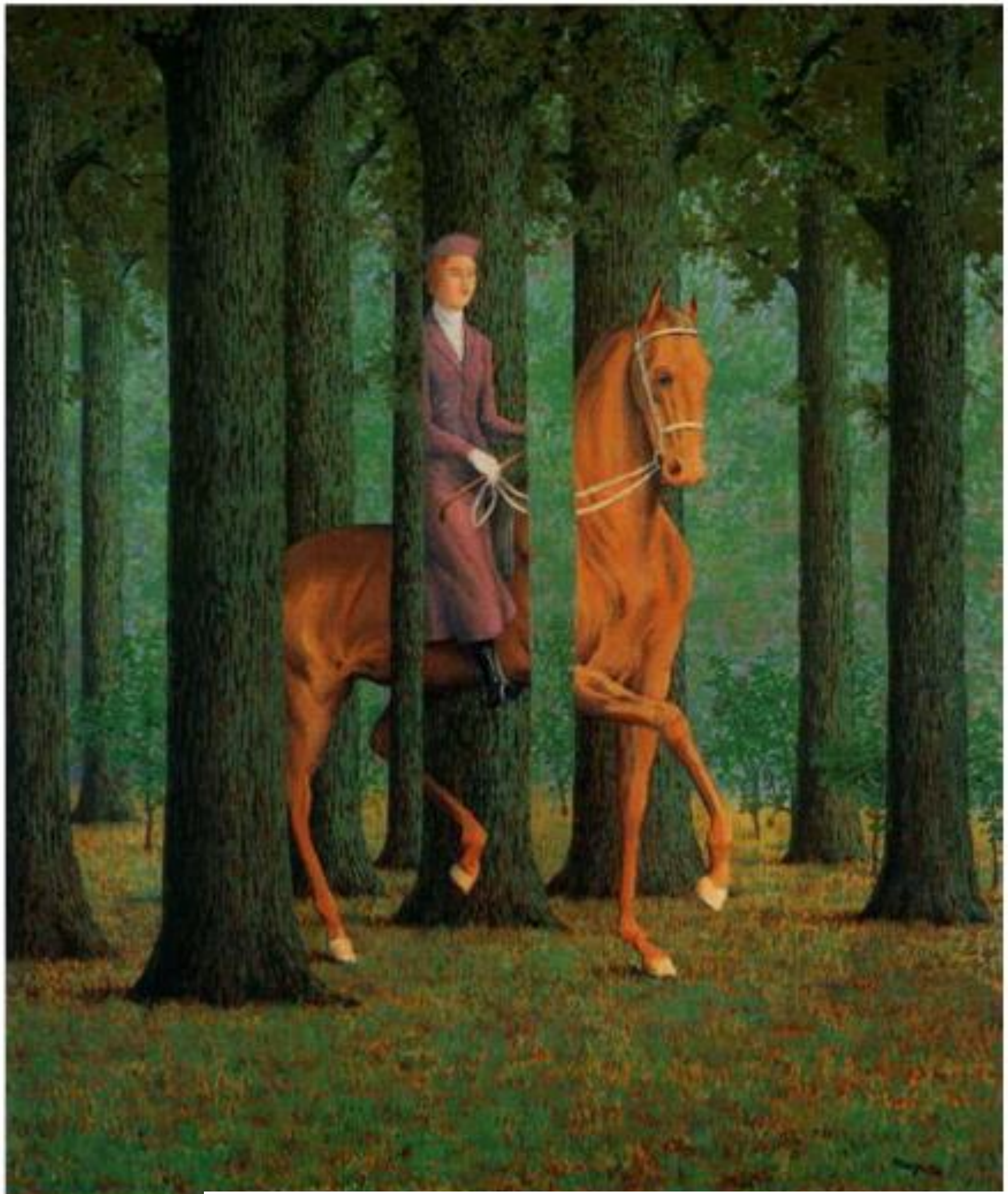


image credit: J. Koenderink

Challenges: scale

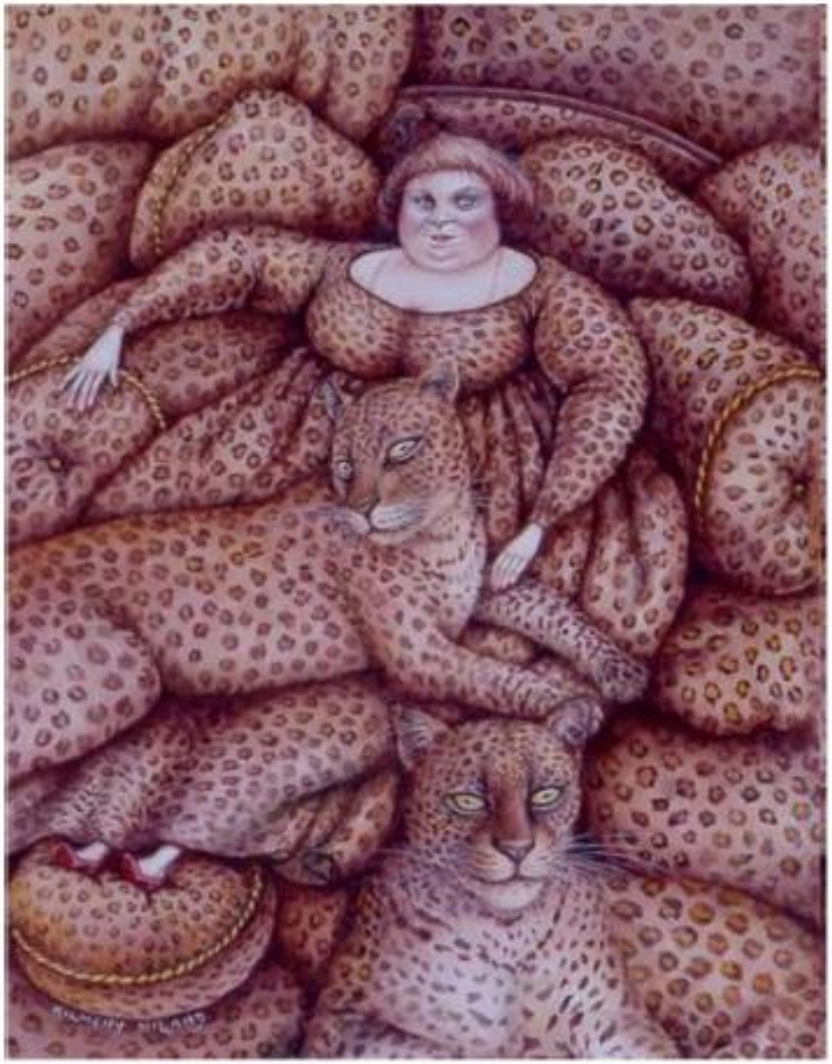
Challenges: deformation

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

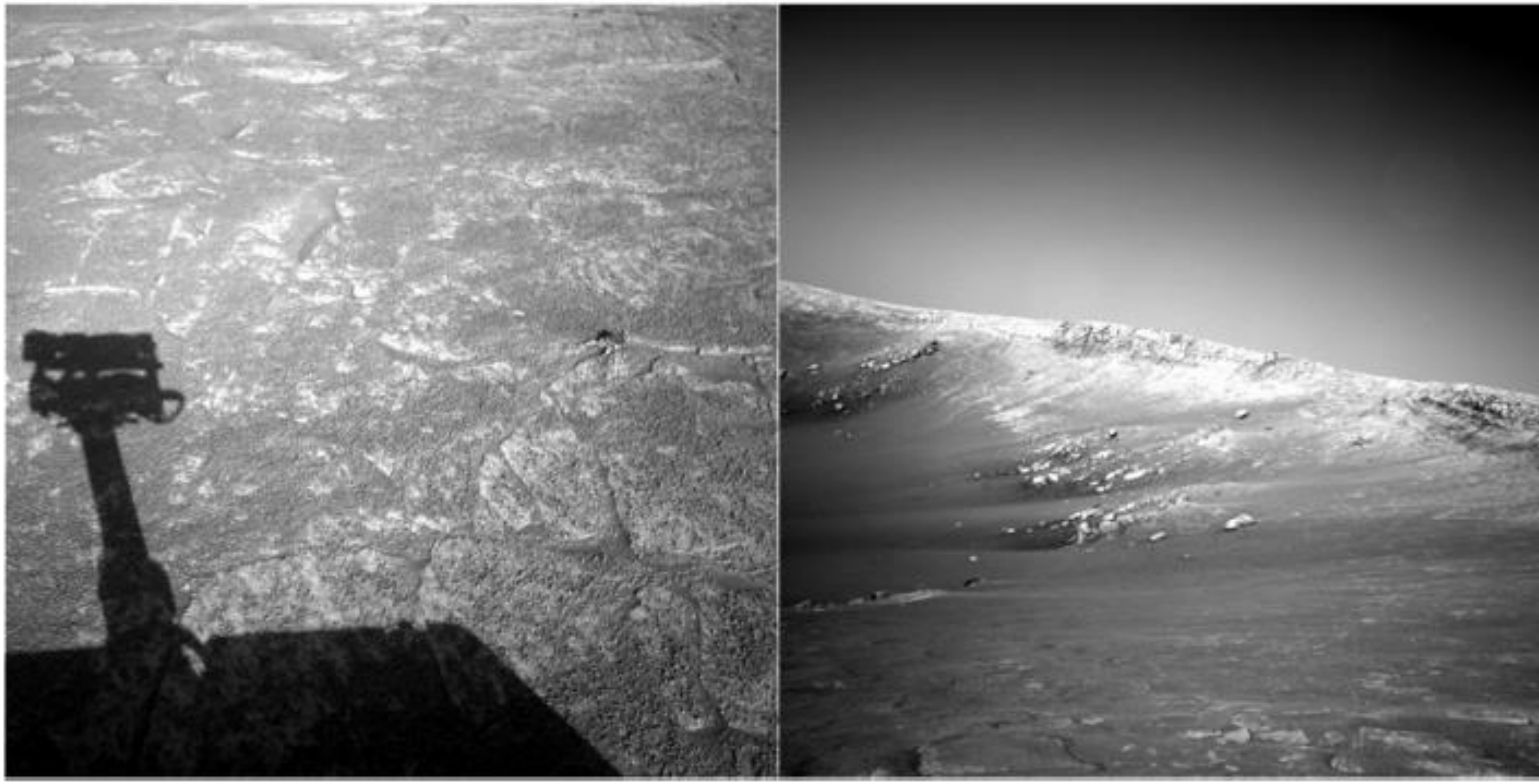
Harder Case

by Diva Sian

by scgbt

Slide credit: Steve Seitz

Harder Still?



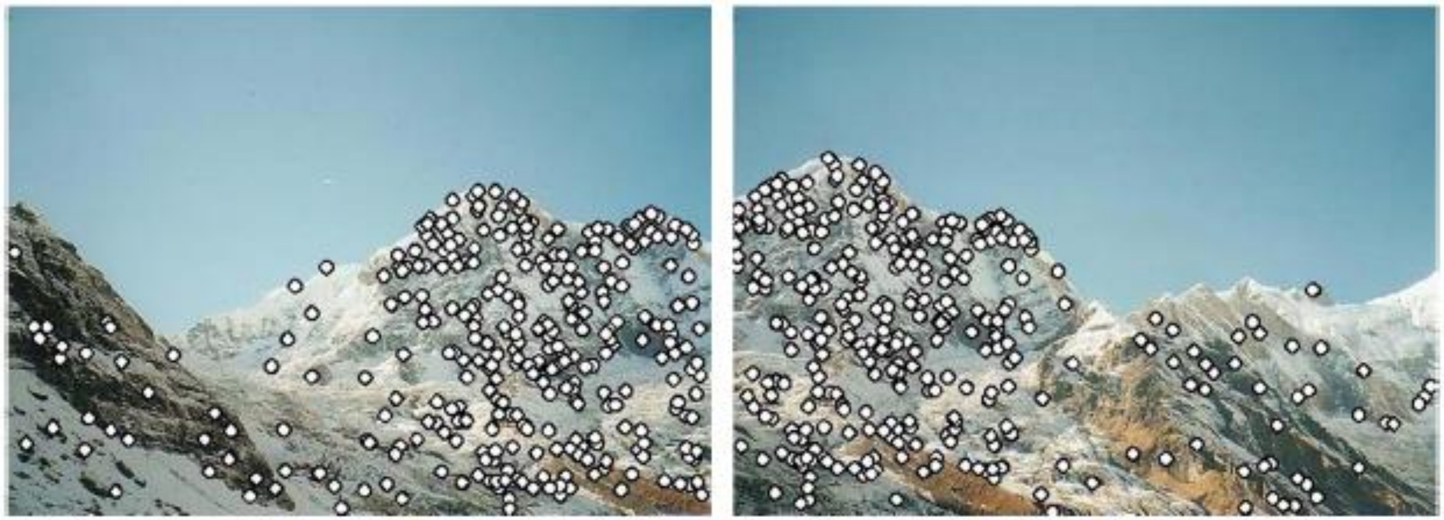
NASA Mars Rover images

Slide credit: Steve Seitz

Application: Image Stitching

Slide credit: Darya Frolova, Denis Simakov

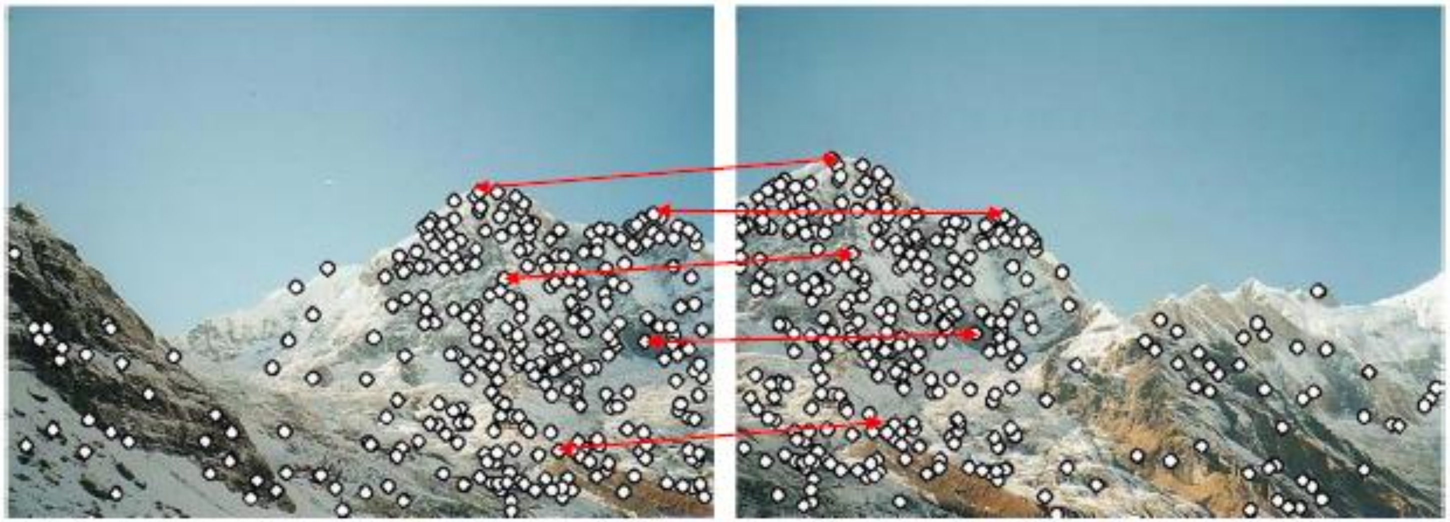
Application: Image Stitching



- Procedure:
  - Detect feature points in both images

Slide credit: Darya Frolova, Denis Simakov

Application: Image Stitching



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

Slide credit: Darya Frolova, Denis Simakov

Application: Image Stitching

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

Slide credit: Darya Frolova, Denis Simakov

Common Requirements

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

No chance to match!

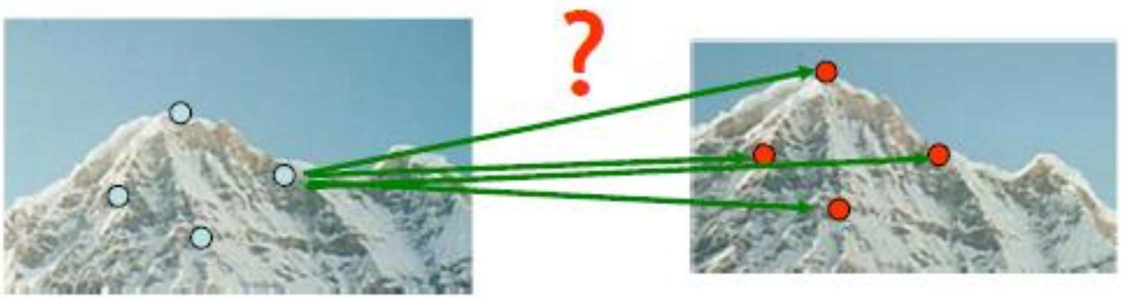
This lecture

We need a repeatable detector!

Slide credit: Darya Frolova, Denis Simakov

Common Requirements

- Problem 1:
  - Detect the same point *independently* in both images
- Problem 2:
  - For each point correctly recognize the corresponding one



Next lecture

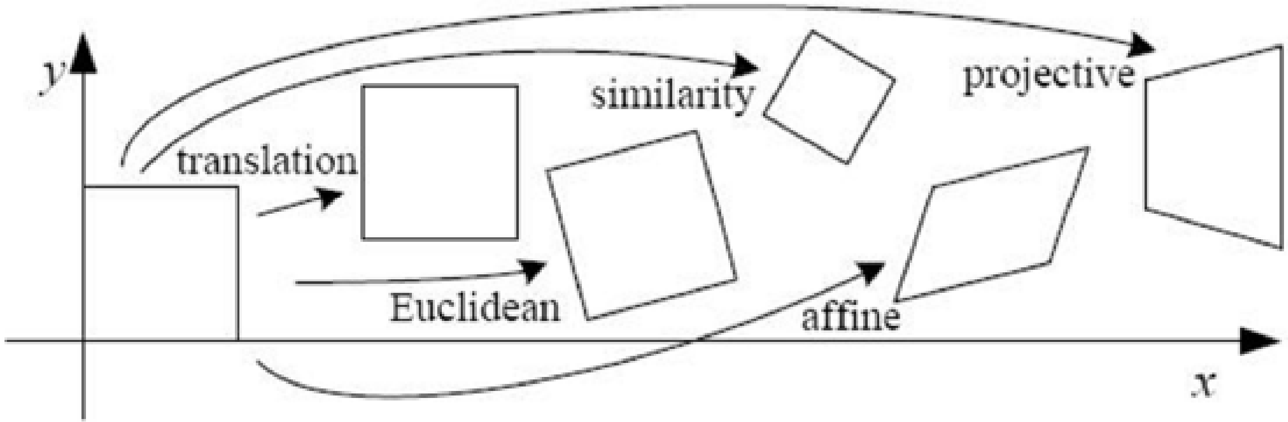
We need a reliable and distinctive descriptor!

Slide credit: Darya Frolova, Denis Simakov

Invariance: Geometric Transformations

Slide credit: Steve Seitz

Levels of Geometric Invariance



Slide credit: Bastian Leibe

Invariance: Photometric Transformations

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

Slide credit: Tinne Tuytelaars

Requirements

- Region extraction needs to be repeatable and accurate
  - Invariant to translation, rotation, scale changes
  - Robust or covariant to out-of-plane (≈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.

Slide credit: Bastian Leibe

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

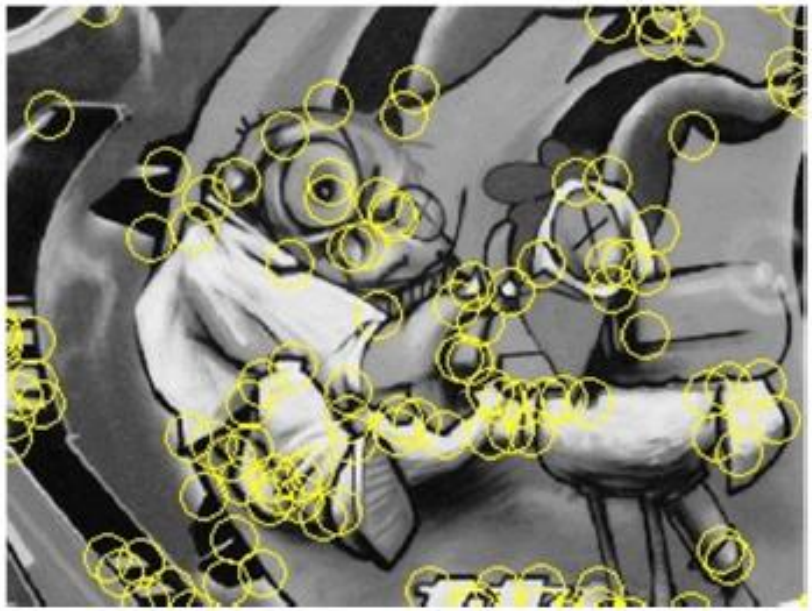
Many Existing Detectors Available

- 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.*

Slide credit: Bastian Leibe

Keypoint Localization

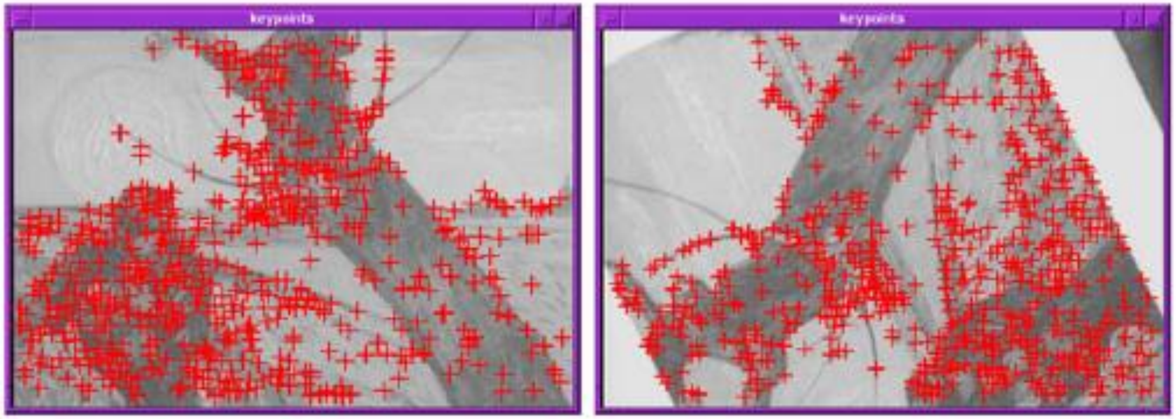


- Goals:
  - Repeatable detection
  - Precise localization
  - Interesting content

$\Rightarrow$ *Look for two-dimensional signal changes*

Slide credit: Bastian Leibe

Finding Corners



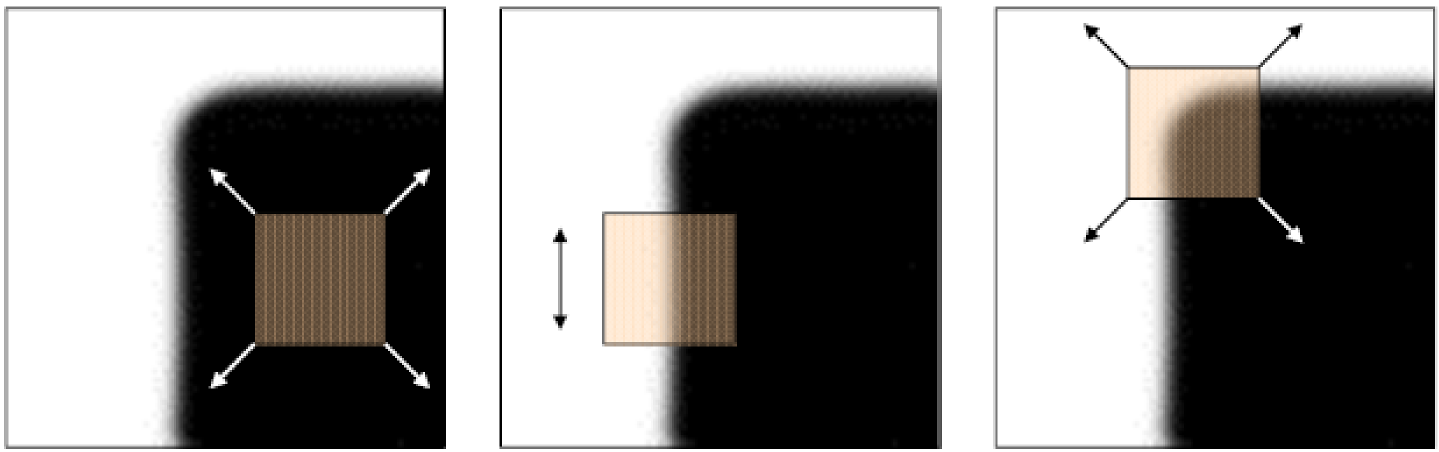
- Key property:
  - In the region around a corner, image gradient has two or more dominant directions
- Corners are *repeatable* and *distinctive*

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

Slide credit: Svetlana Lazebnik

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

Slide credit: Alyosha Efros

Harris Detector Formulation

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

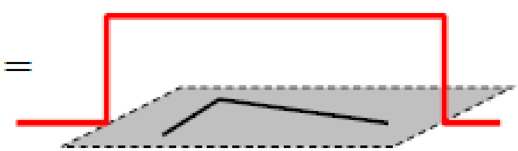
$$E(u,v) = \sum_{x,y} w(x,y)\left[I(x+u,y+v) - I(x,y)\right]^2$$

Window function

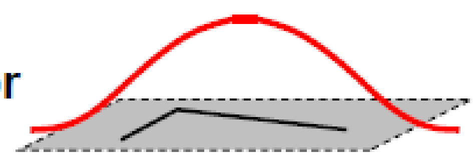
Shifted intensity

Intensity

Window function $w(x,y) =$



or



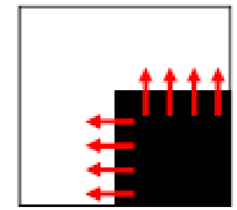
1 in window, 0 outside

Gaussian

Slide credit: Rick Szeliski

What Does This Matrix Reveal?

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



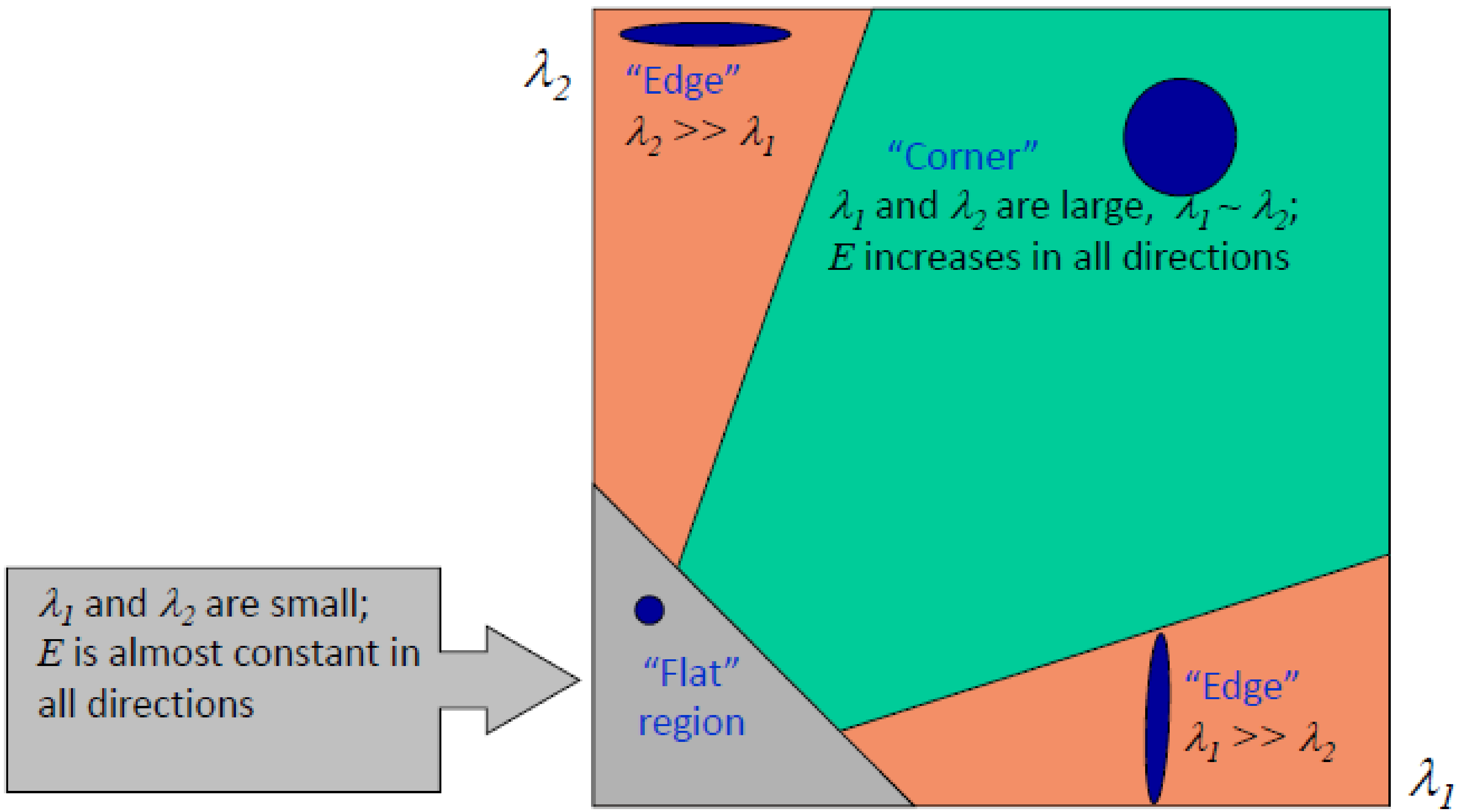
$$M = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

- 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?

Slide credit: David Jacobs

Interpreting the Eigenvalues

- Classification of image points using eigenvalues of M:

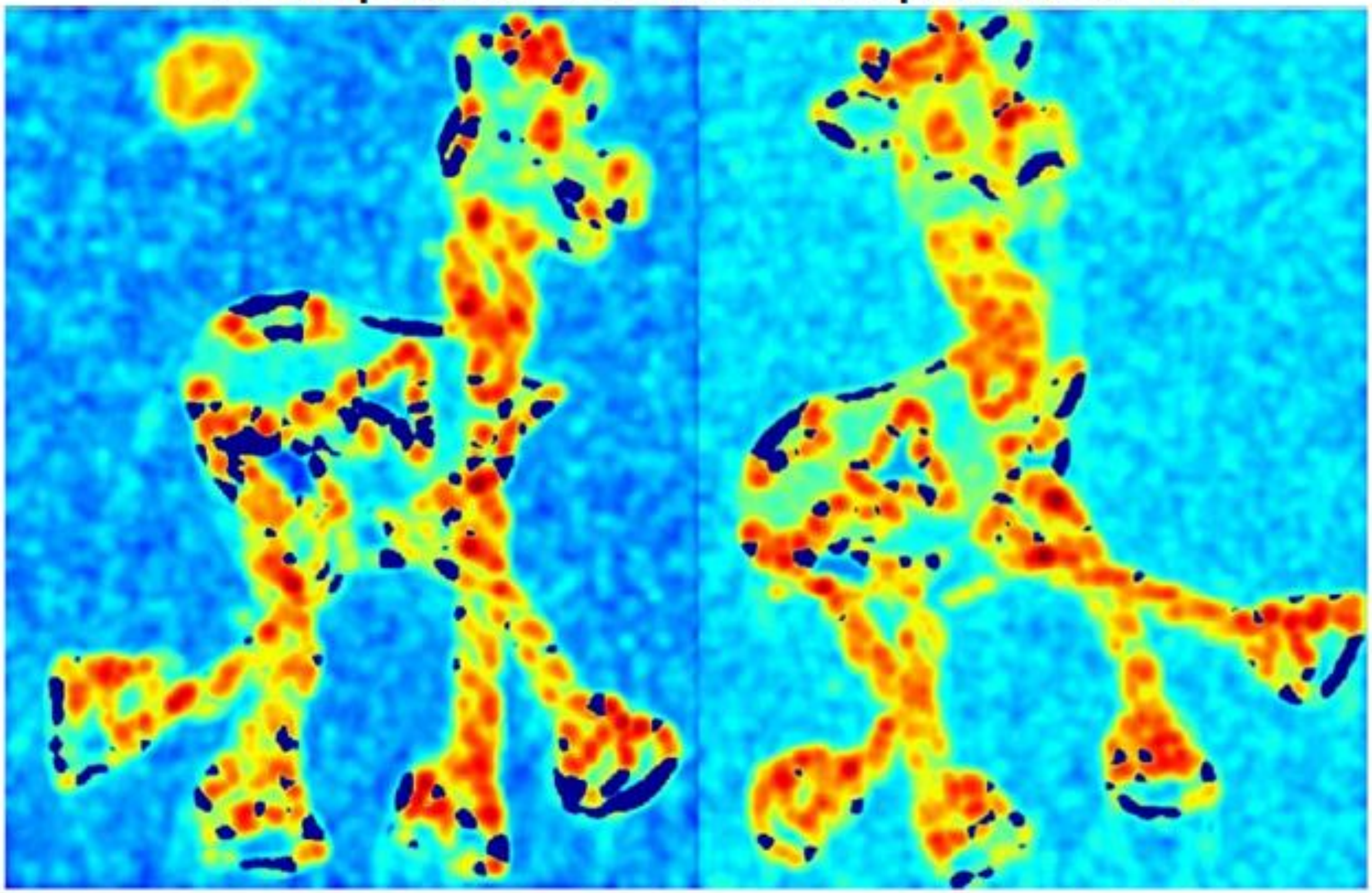


Slide credit: Kristen Grauman

Harris Detector: Workflow

Slide adapted from Darya Frolova, Denis Simakov

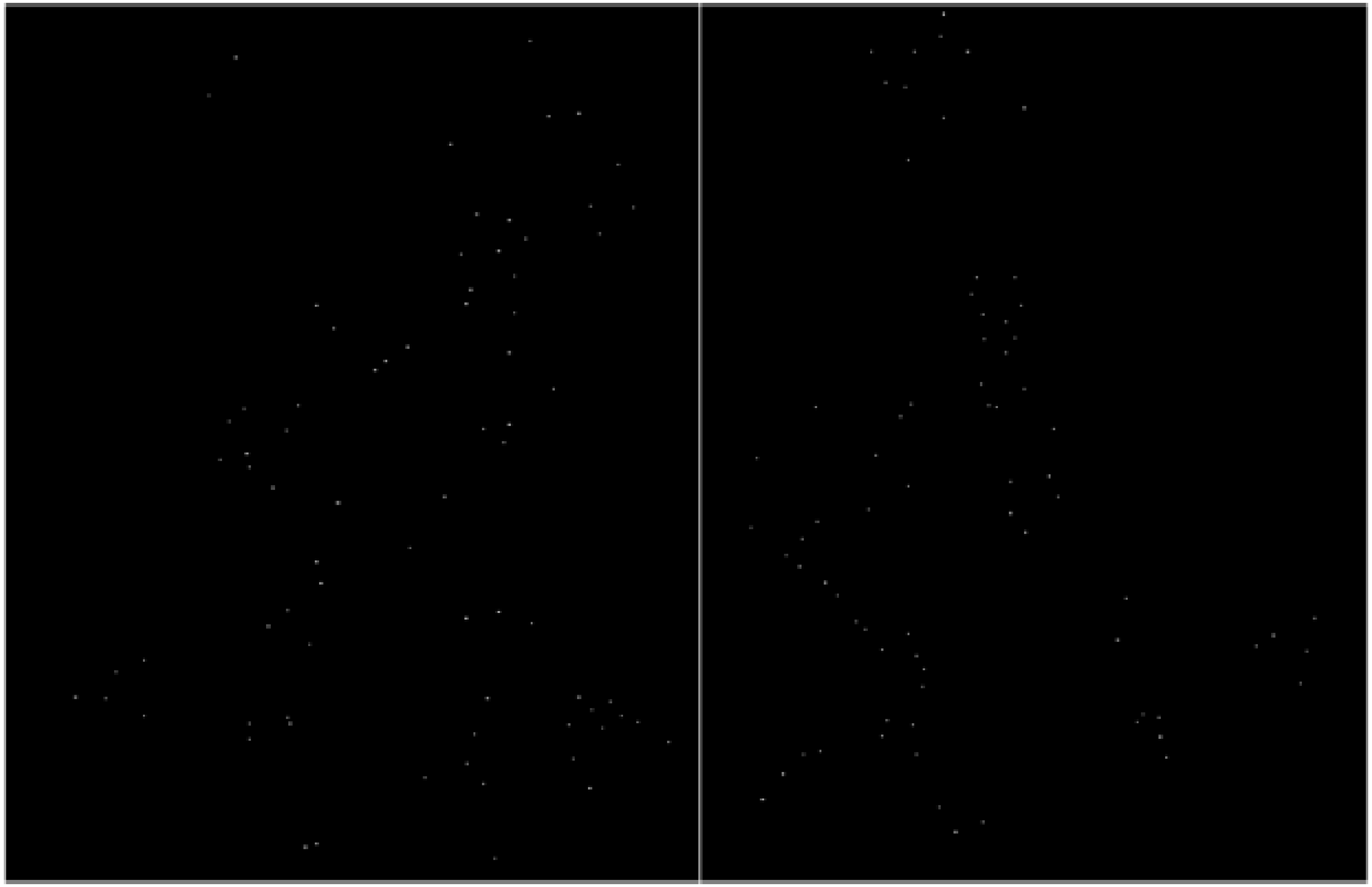
Harris Detector: Workflow
- computer corner responses R



Slide adapted from Darya Frolova, Denis Simakov

Harris Detector: Workflow

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

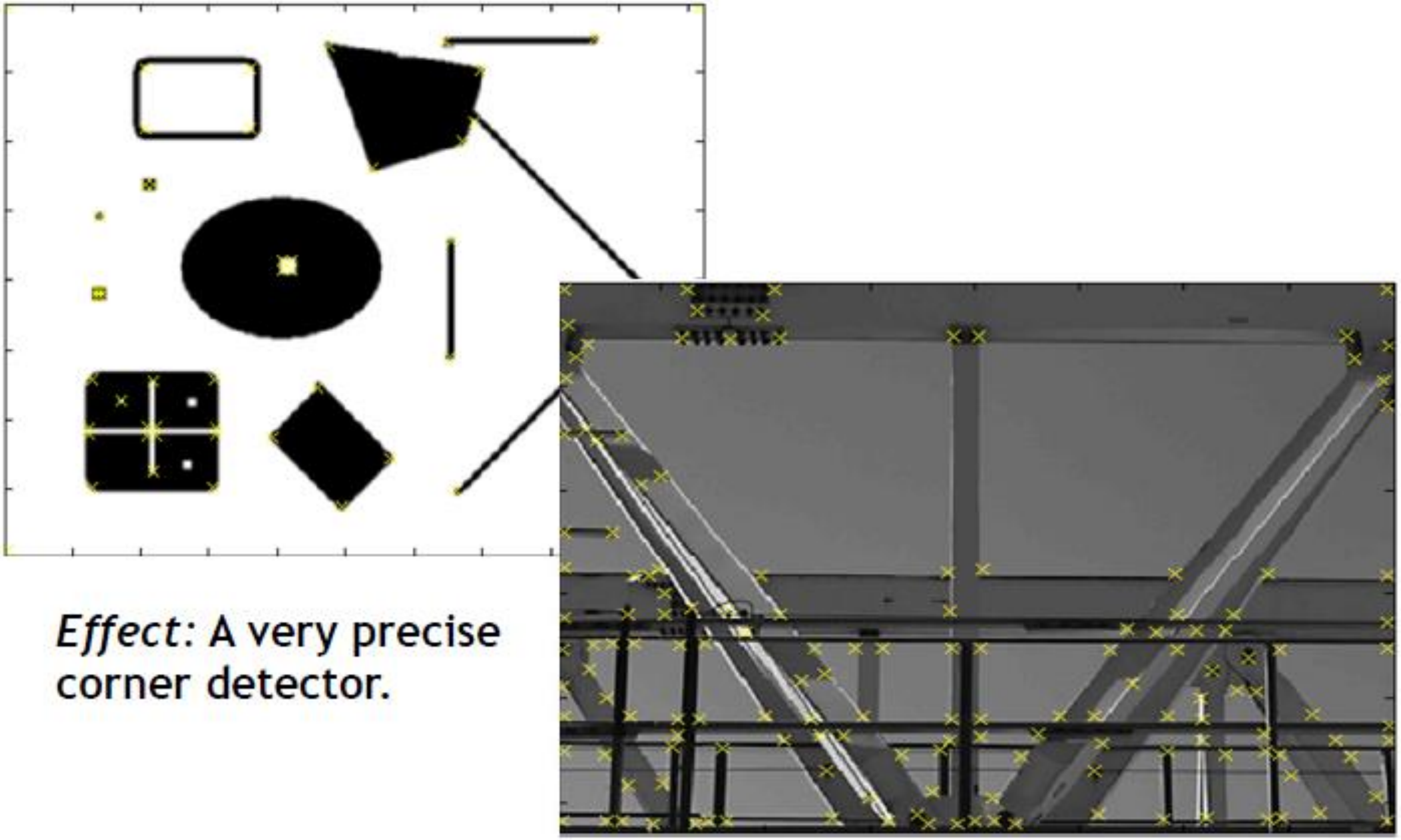


Slide adapted from Darya Frolova, Denis Simakov

Harris Detector: Workflow
- Resulting Harris points

Slide adapted from Darya Frolova, Denis Simakov

Harris Detector – Responses [Harris88]



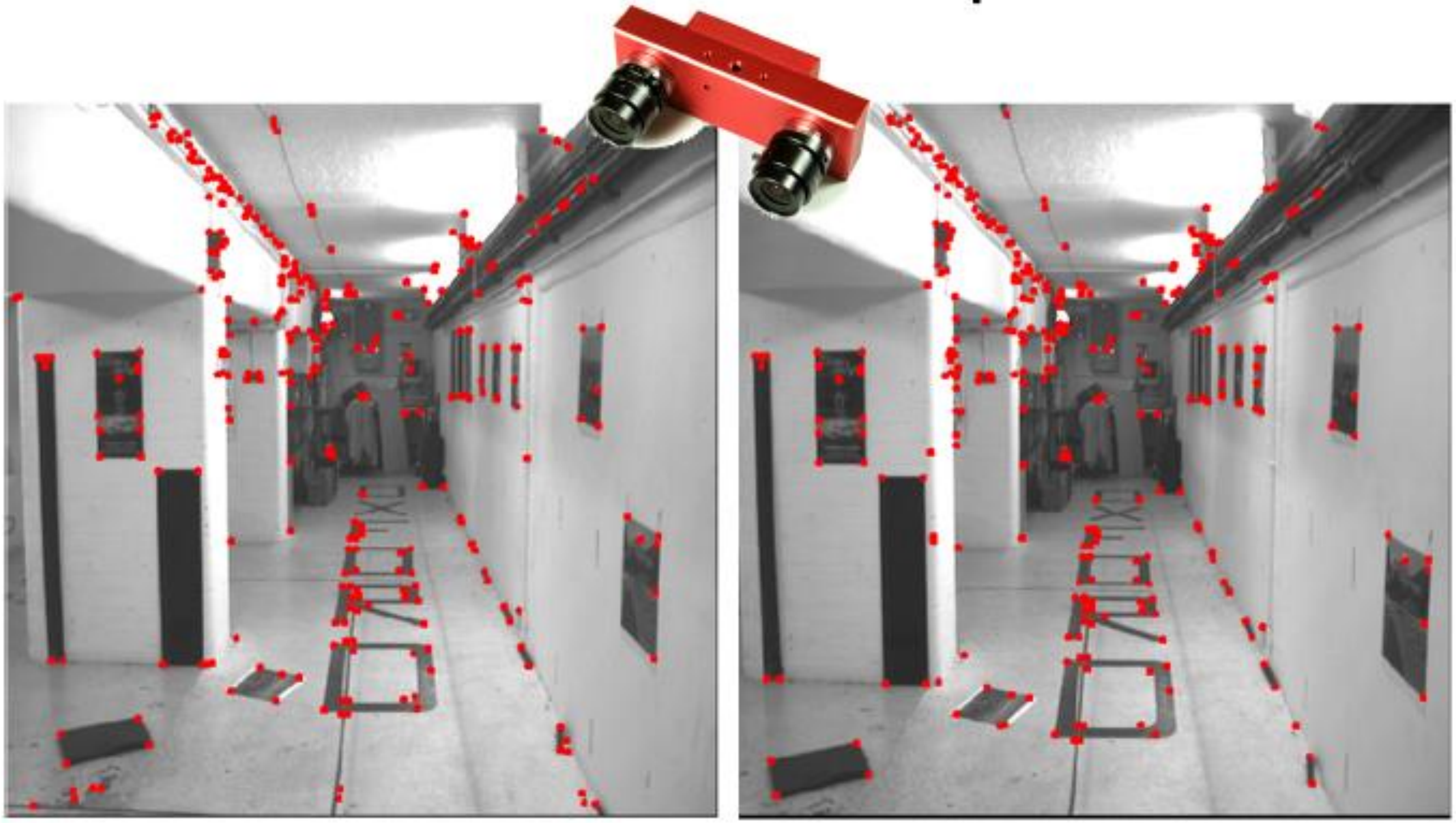
Effect: A very precise corner detector.

Slide credit: Krystian Mikolajczyk

Harris Detector – Responses [Harris88]

Slide credit: Krystian Mikolajczyk

Harris Detector – Responses [Harris88]

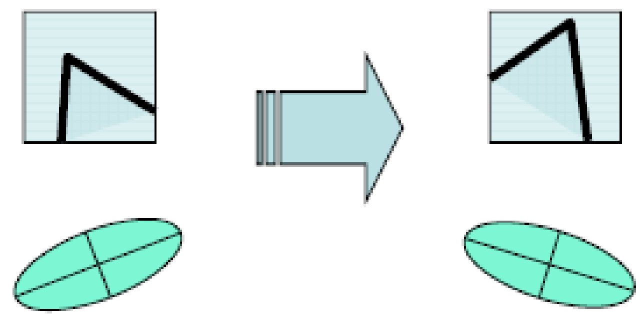


- Results are well suited for finding stereo correspondences

Slide credit: Kristen Grauman

Harris Detector: Properties

- Rotation invariance?



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

Corner response R **is invariant to image rotation**

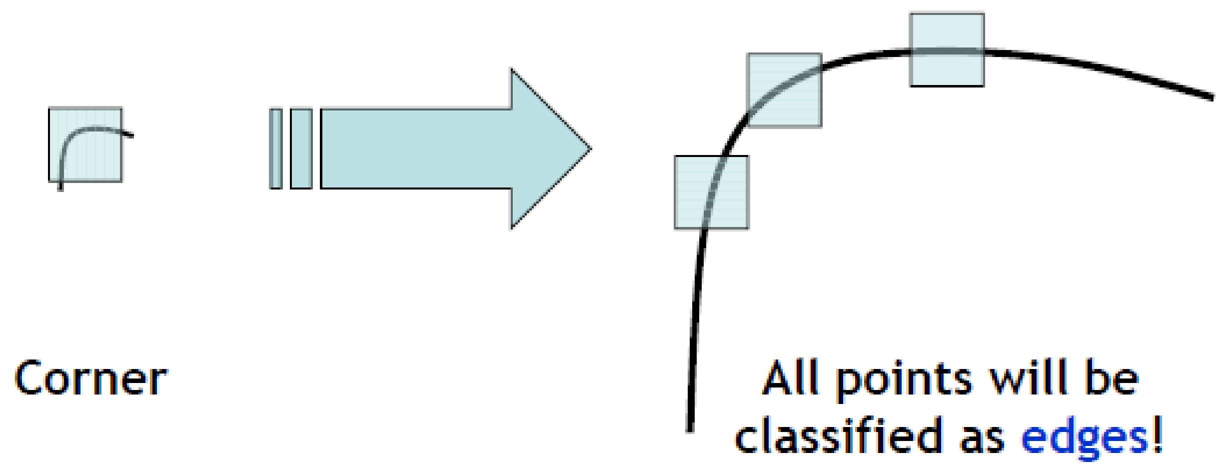
Slide credit: Kristen Grauman

Harris Detector: Properties

- Rotation invariance
- Scale invariance?

Harris Detector: Properties

- Rotation invariance
- Scale invariance?



Not invariant to image scale!

Slide credit: Kristen Grauman

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**

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

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, MM, SIGGRAPH
  - Recent publication : papers published since 2013
  - 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..