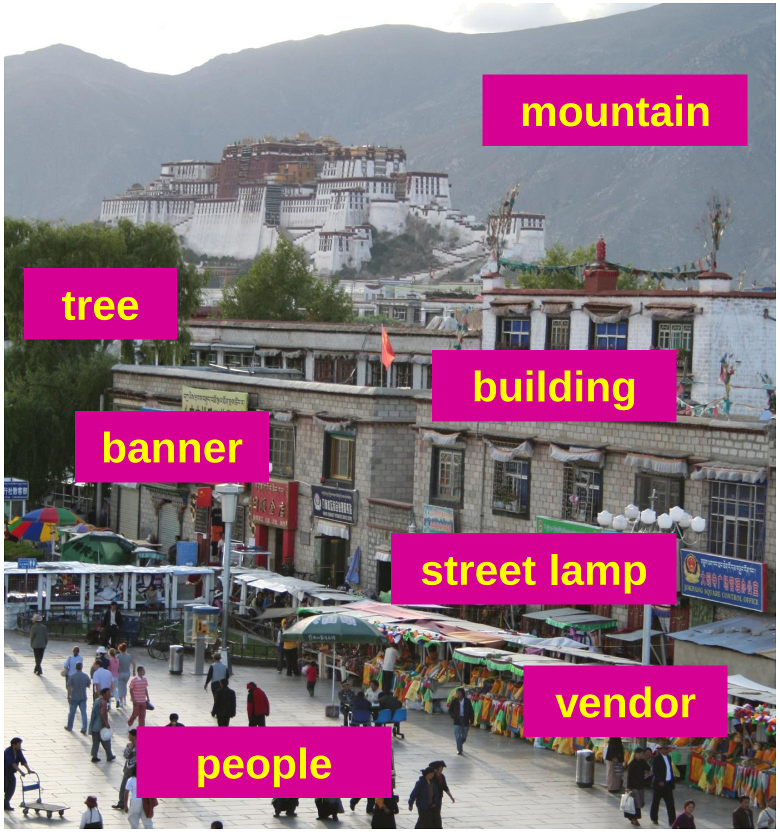
CS5670: Intro to Computer Vision

Noah Snavely

Introduction to Recognition



Announcements

- Project 4 (Stereo) has been released (demo)
  - Due next Friday, April 28, by 11:59pm
  - To be done in pairs

Where we go from here

- What we know: Geometry
  - What is the shape of the world? How does that shape appear in images?
- What's next: Recognition
  - What are we looking at?

What do we mean by “object recognition”?

Next 15 slides adapted from Li, Fergus, & Torralba’s excellent short course on category and object recognition

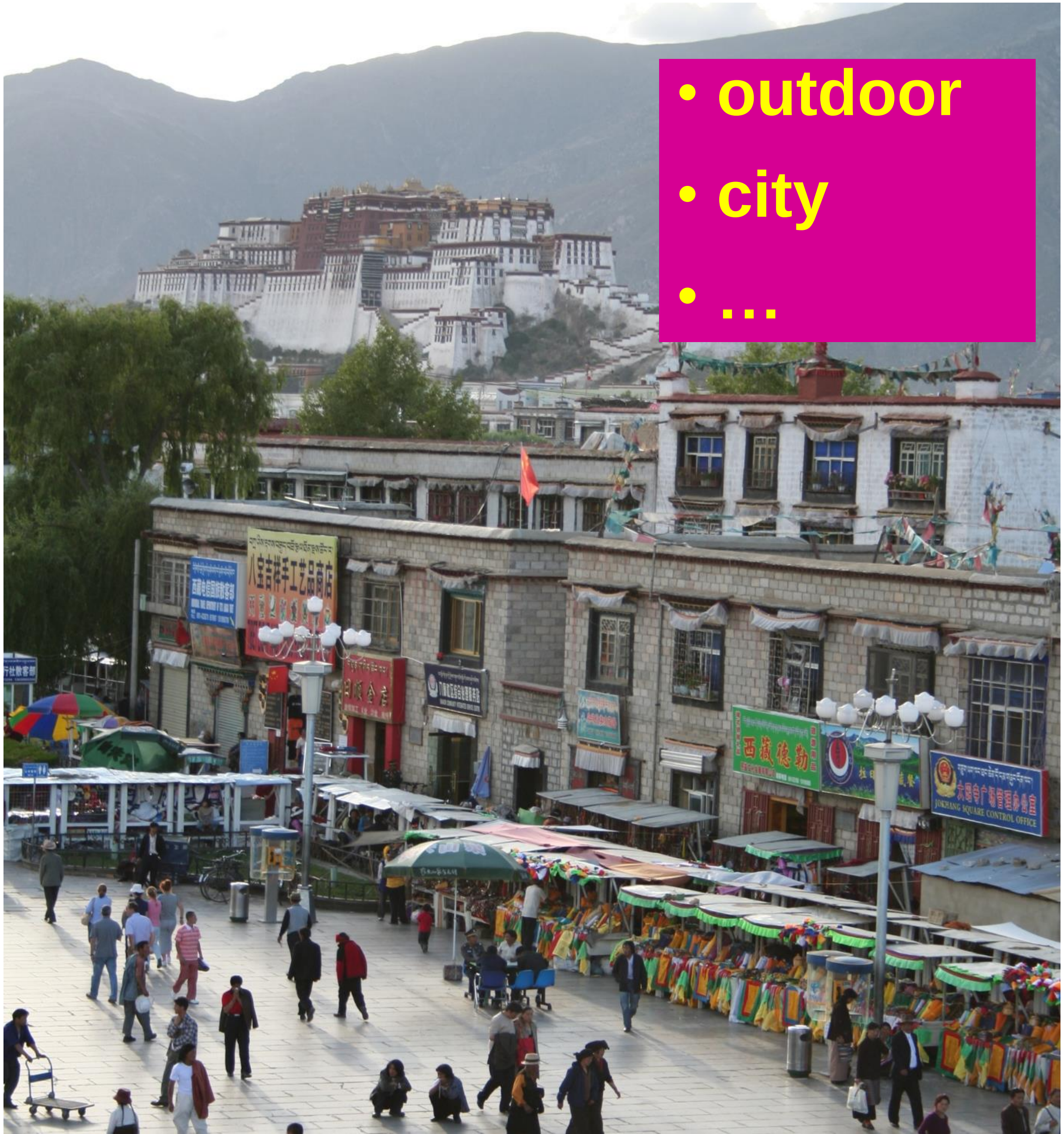
Verification: is that a lamp?

Detection: are there people?

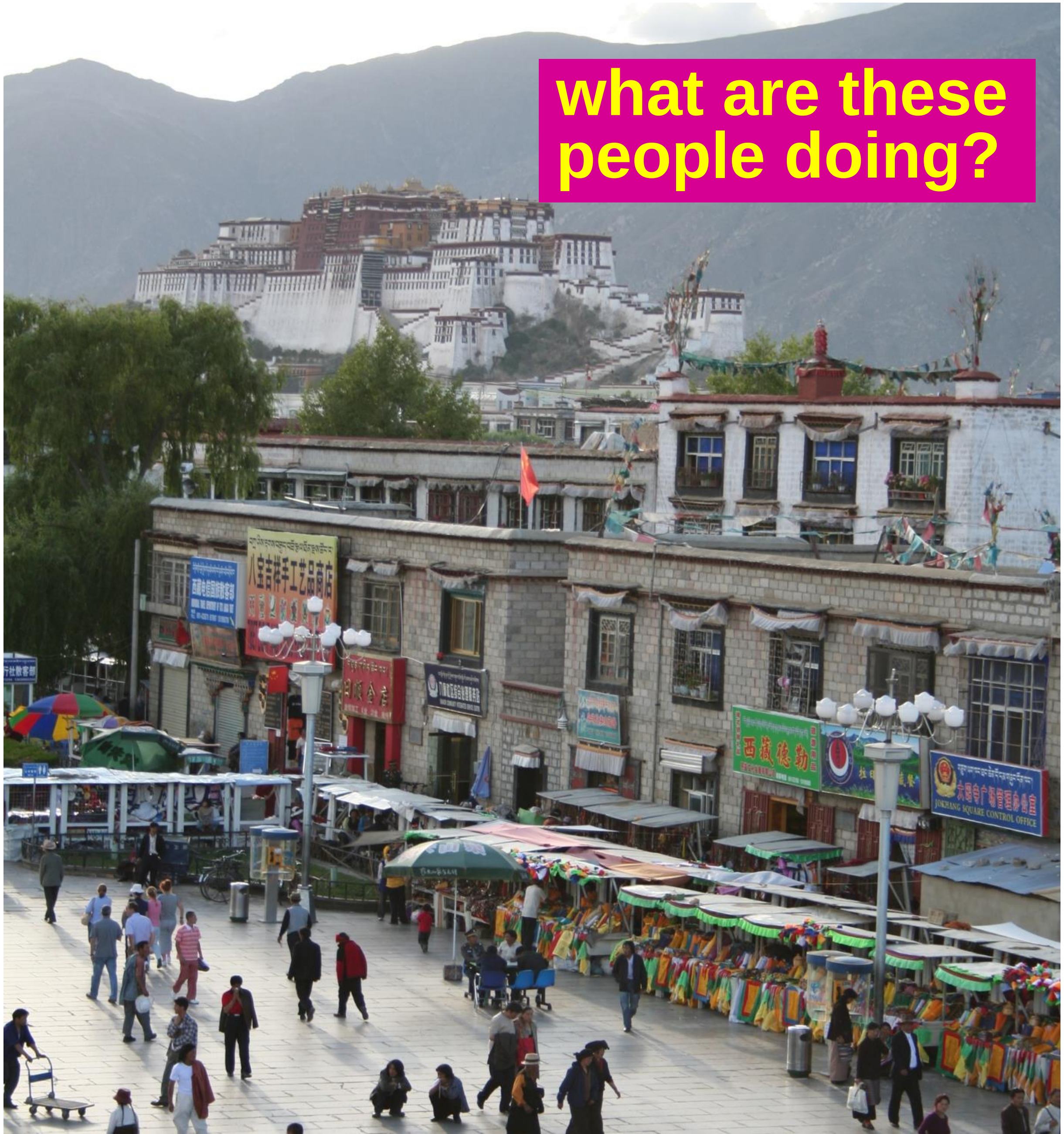
Identification: is that Potala Palace?

Object categorization

Scene and context categorization

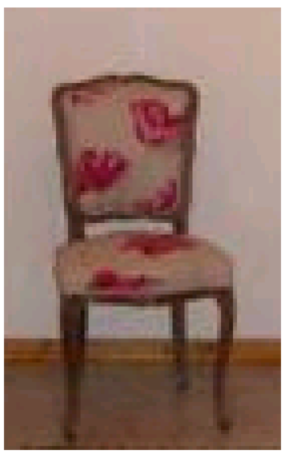


Activity / Event Recognition



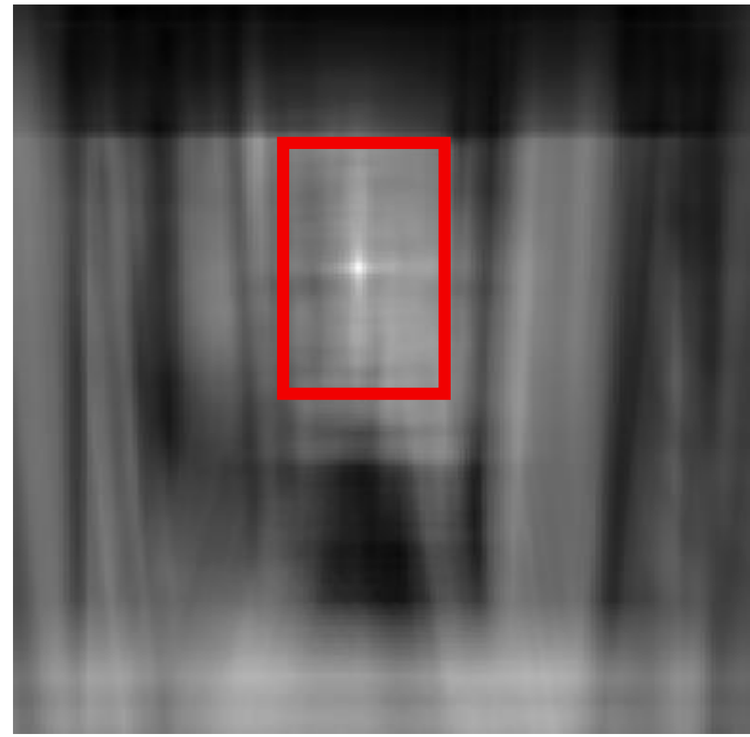
Object recognition
Is it really so hard?

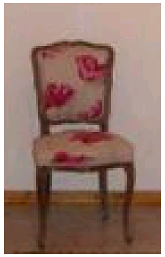
This is a chair



Find the chair in this image

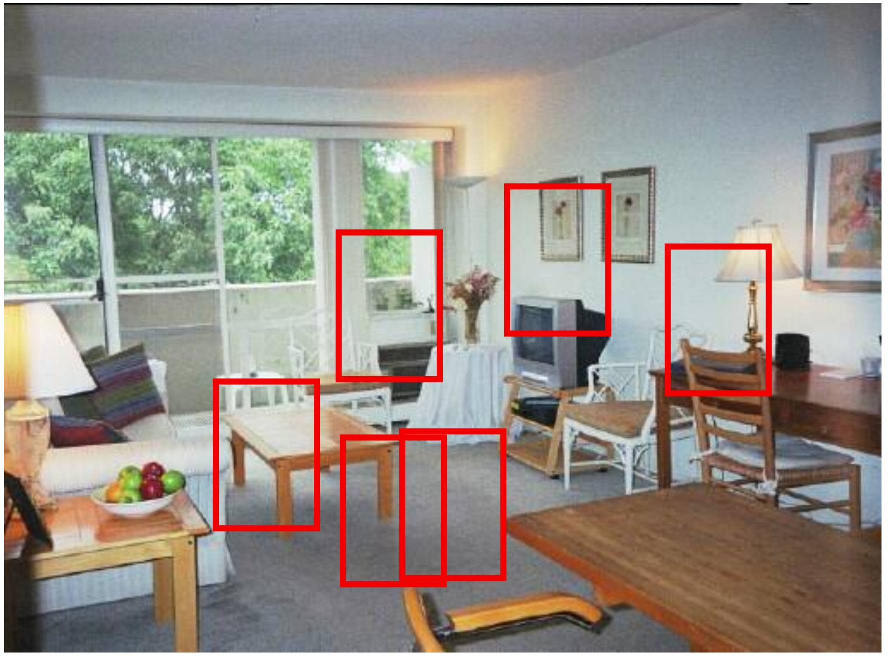
Output of normalized correlation

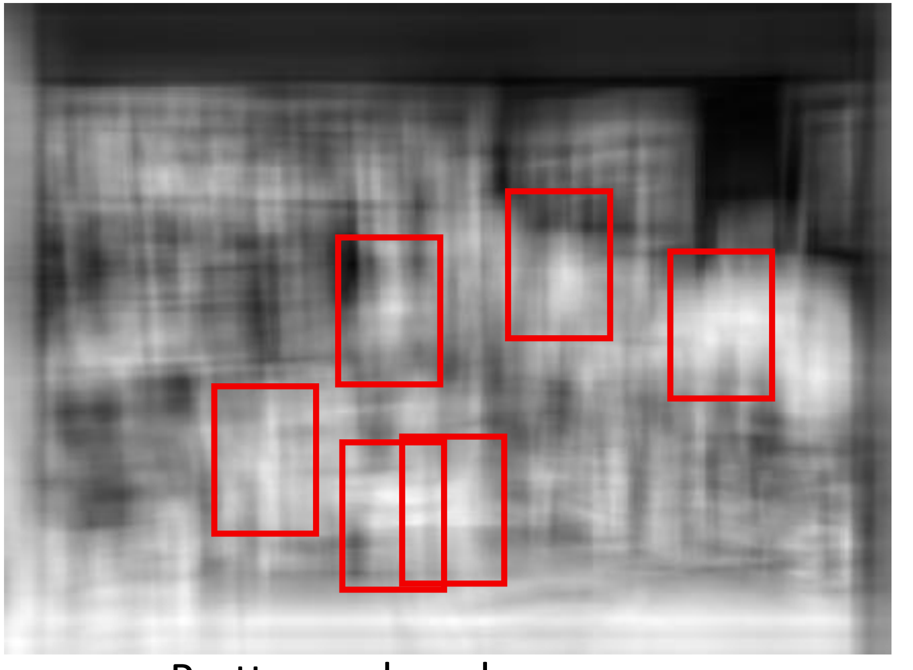




Object recognition
Is it really so hard?

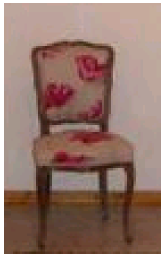
Find the chair in this image





Pretty much garbage

Simple template matching is not going to make it



Object recognition
Is it really so hard?

Find the chair in this image

A “popular method is that of template matching, by point to point correlation of a model pattern with the image pattern. These techniques are inadequate for three-dimensional scene analysis for many reasons, such as occlusion, changes in viewing angle, and articulation of parts.” Nivatia & Binford, 1977.

Why not use SIFT matching for everything?

- Works well for object *instances*

- Not great for generic object *categories*



And it can get a lot harder

Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. J Vis, 3(6), 413-422

How do human do recognition?

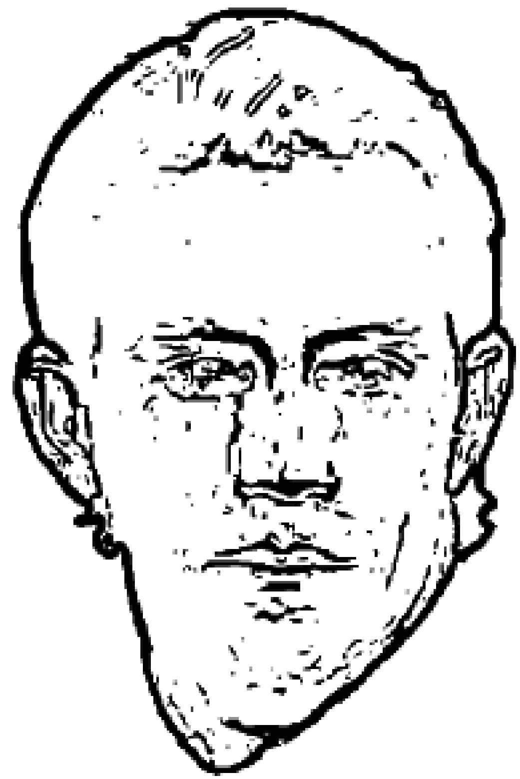
- We don’t completely know yet
- But we have some experimental observations.

Observation 1

- We can recognize familiar faces even in low-resolution images

Observation 2:

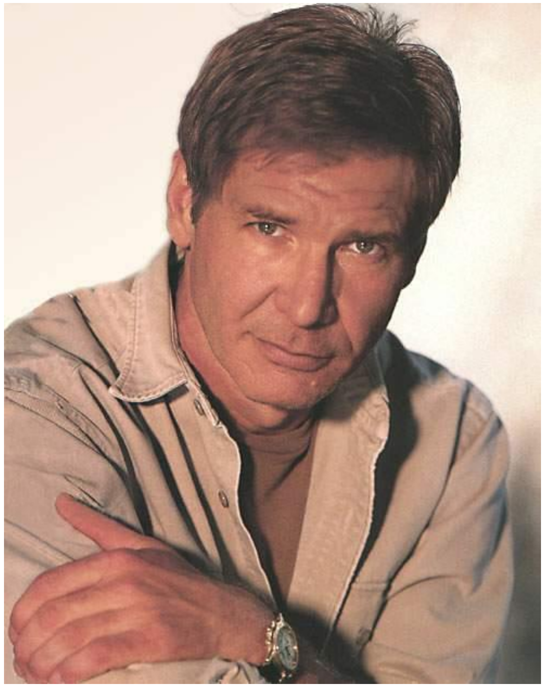
Jim Carrey



Kevin Costner

- High frequency information is not enough

What is the single most important facial features for recognition?



Observation 4:

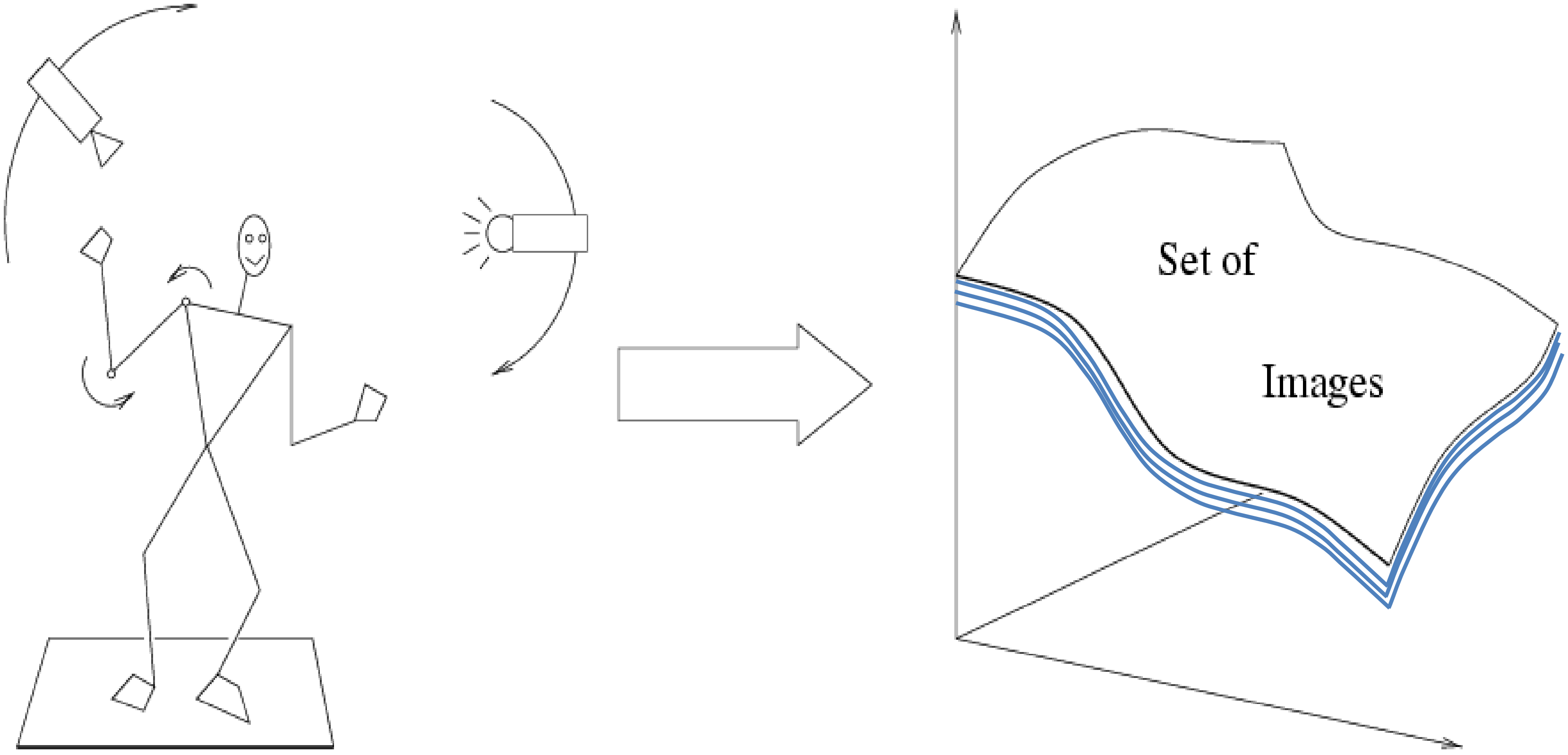
- Image Warping is OK

The list goes on

Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About

- http://web.mit.edu/bcs/sinha/papers/19results_sinha_etal.pdf

Why is this hard?



Variability: Camera position
Illumination
Shape parameters

How many object categories are there?

~10,000 to 30,000

Challenge: variable viewpoint

Michelangelo 1475-1564

Challenge: variable illumination

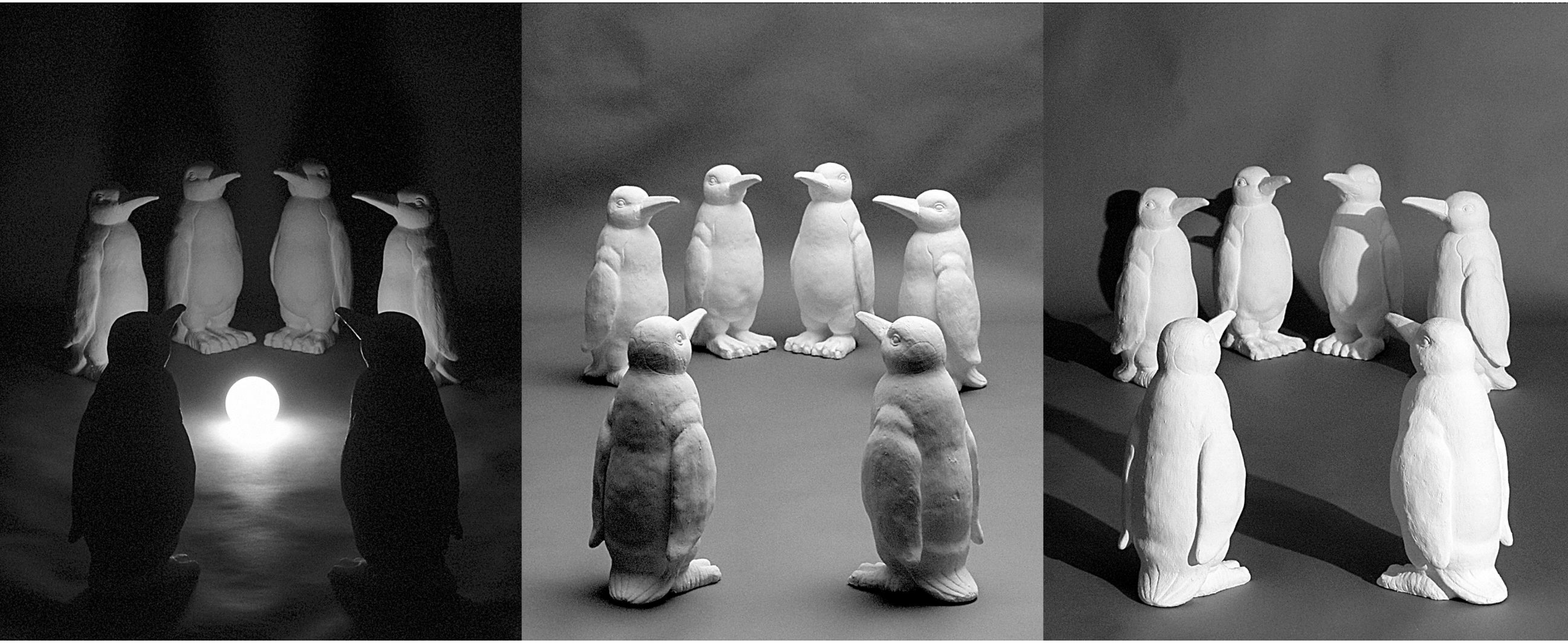
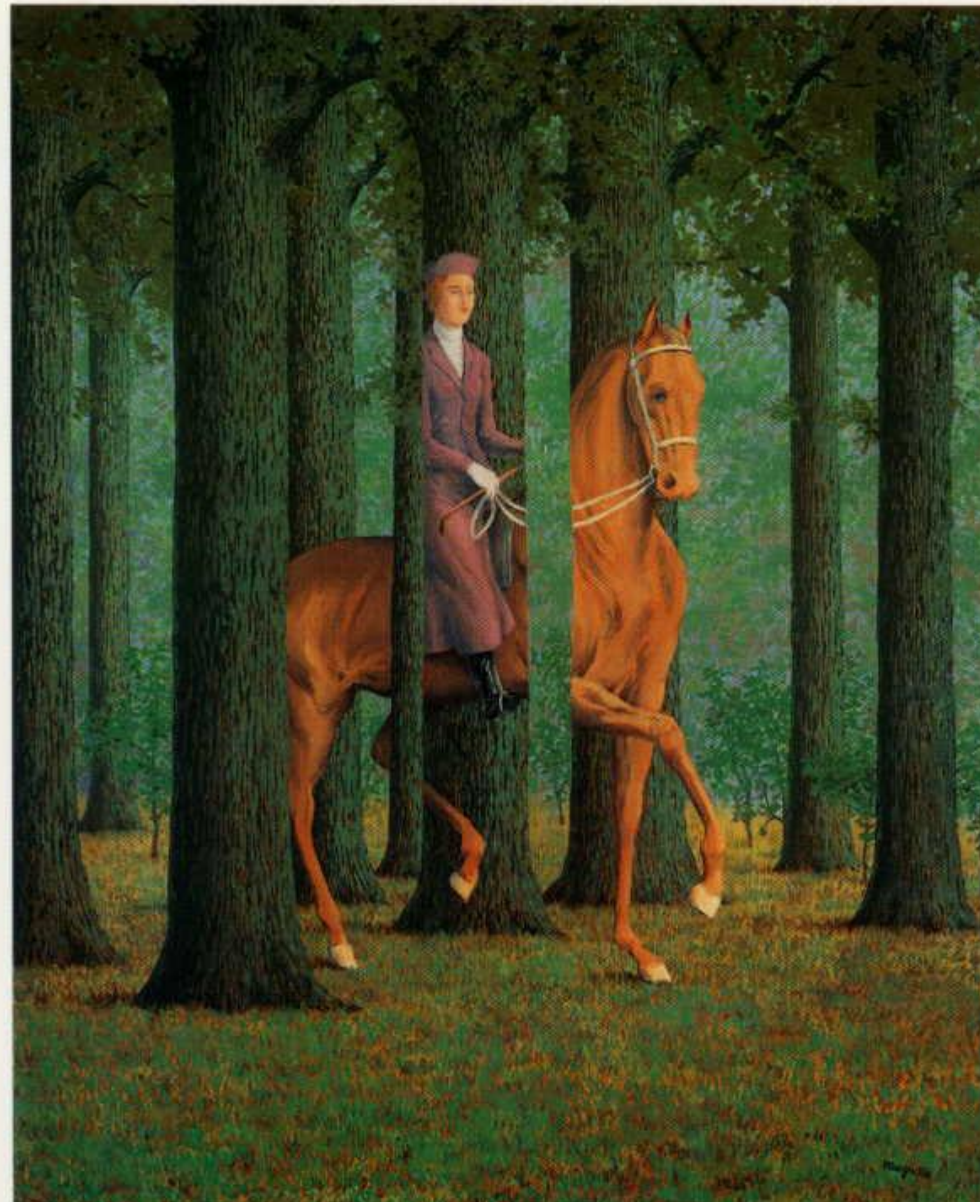


image credit: J. Koenderink

Challenge: scale

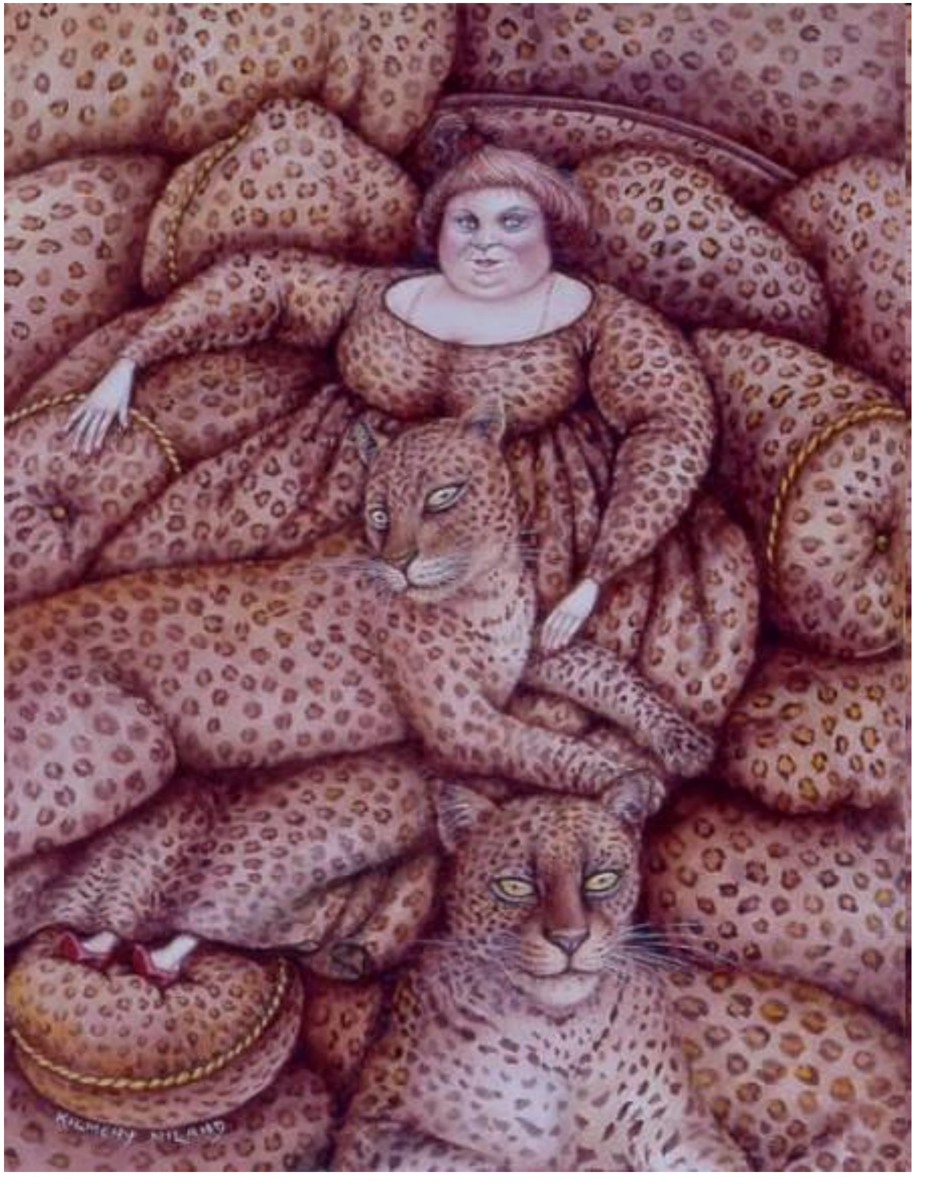
Challenge: deformation

Challenge: Occlusion



Magritte, 1957

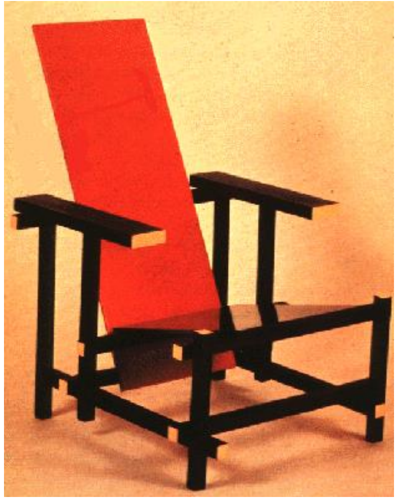
Challenge: background clutter



Kilmeny Niland. 1995

Challenge: intra-class variations

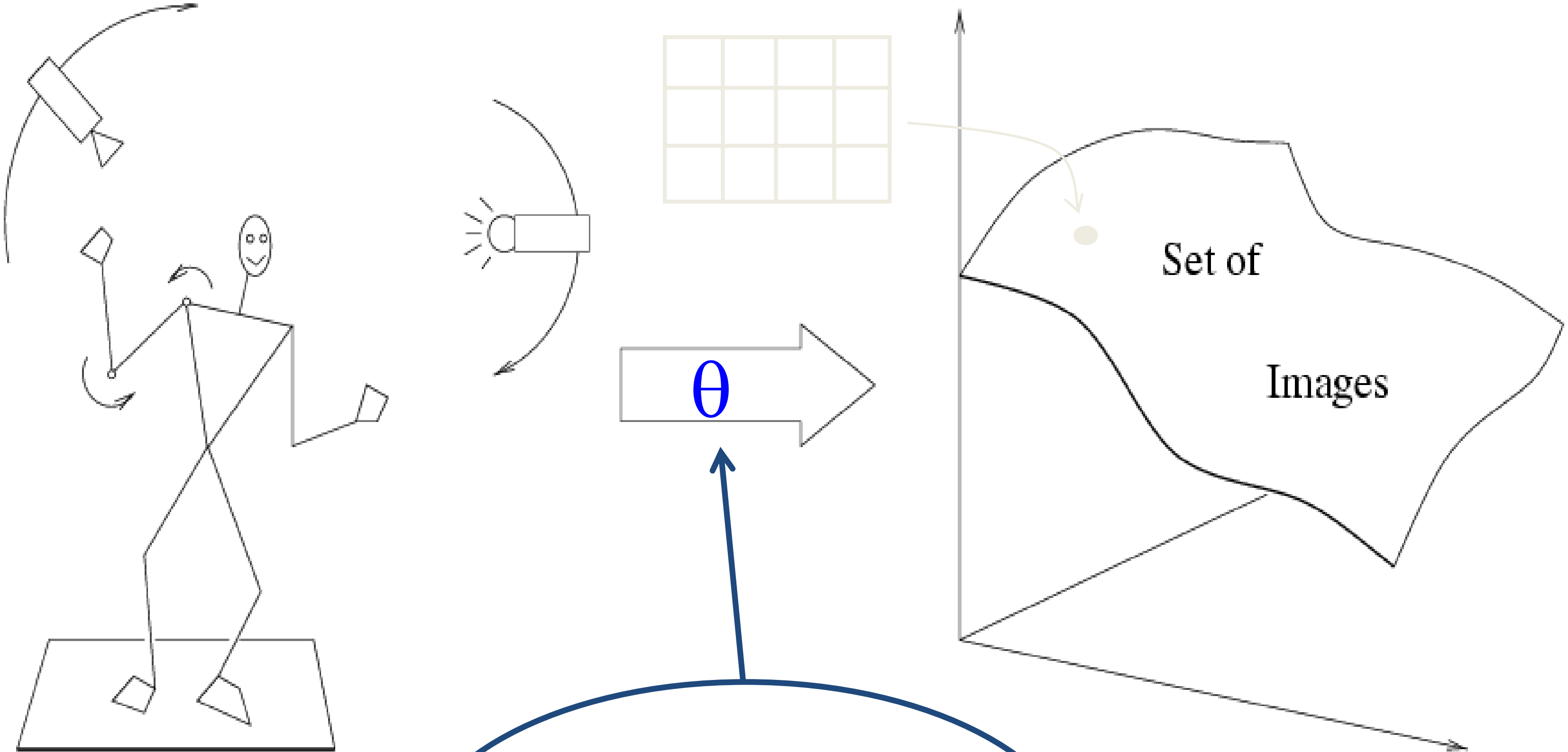


Svetlana Lazebnik

History of ideas in recognition

- 1960s – early 1990s: the geometric era



Variability: Camera position
Illumination

Alignment

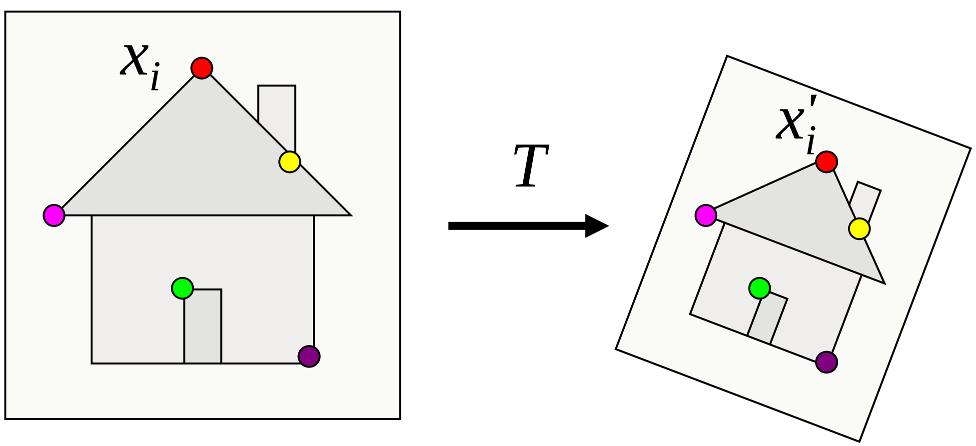
Shape: assumed known

Roberts (1965); Lowe (1987); Faugeras & Hebert (1986); Grimson & Lozano-Perez (1986); Huttenlocher & Ullman (1987)

Svetlana Lazebnik

Instance Recognition

- Alignment: fitting a model to a transformation between pairs of features (*matches*) in two images



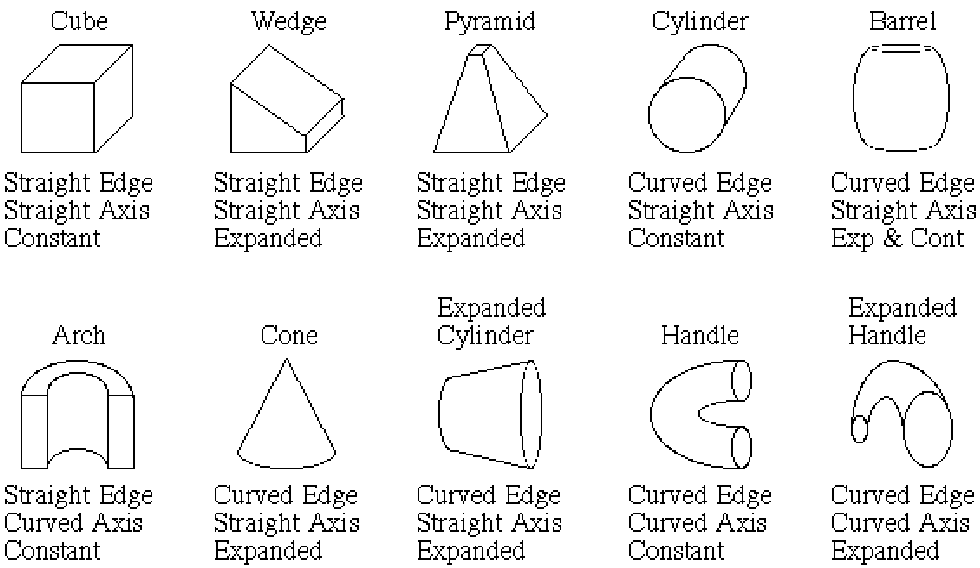
Find transformation *T* that minimizes

$$\sum_i \text{residual}\,(T(x_i), x_i')$$

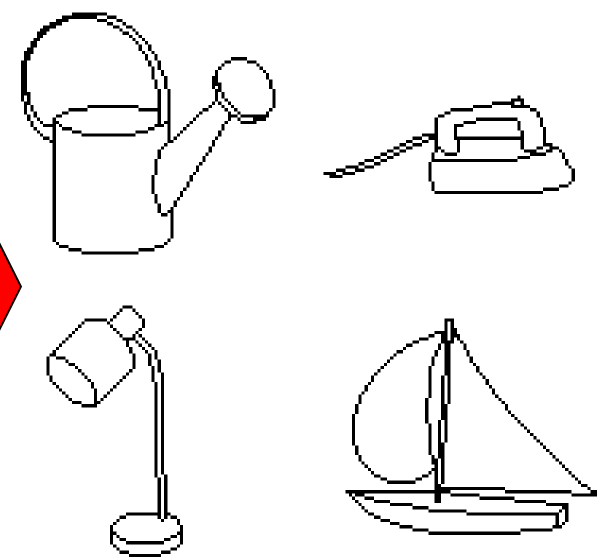
Recognition by components

Biederman (1987)

Primitives (geons)



Objects



http://en.wikipedia.org/wiki/Recognition_by_Components_Theory

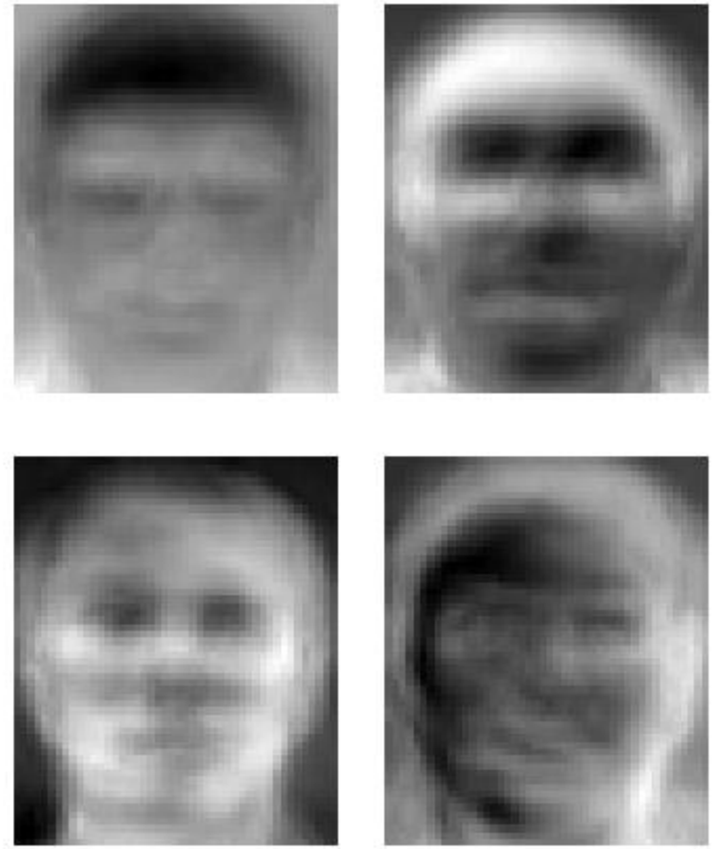
Svetlana Lazebnik

History of ideas in recognition

- 1960s – early 1990s: the geometric era
- 1990s: appearance-based models

Svetlana Lazebnik

Eigenfaces (Turk & Pentland, 1991)



Experimental	Correct/Unknown Recognition Percentage		
Condition	Lighting	Orientation	Scale
Forced classification	96/0	85/0	64/0
Forced 100% accuracy	100/19	100/39	100/60
Forced 20% unknown rate	100/20	94/20	74/20

Svetlana Lazebnik

Limitations of global appearance models

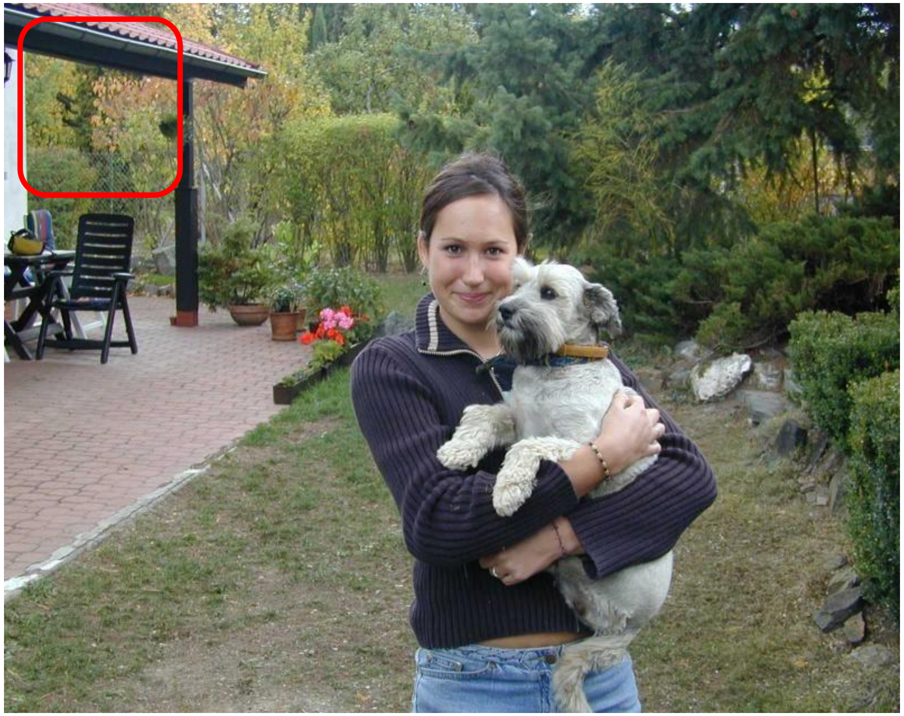
- Requires global registration of patterns
- Not robust to clutter, occlusion, geometric transformations

Svetlana Lazebnik

History of ideas in recognition

- 1960s – early 1990s: the geometric era
- 1990s: appearance-based models
- 1990s – present: sliding window approaches

Sliding window approaches



Sliding window approaches

- Turk and Pentland, 1991
- Belhumeur, Hespanha, & Kriegman, 1997
- Schneiderman & Kanade 2004
- Viola and Jones, 2000

- Schneiderman & Kanade, 2004
- Agrawal and Roth, 2002
- Poggio et al. 1993

History of ideas in recognition

- 1960s – early 1990s: the geometric era
- 1990s: appearance-based models
- Mid-1990s: sliding window approaches
- Late 1990s: local features

Svetlana Lazebnik

Local features for object instance recognition

D. Lowe (1999, 2004)

History of ideas in recognition

- 1960s – early 1990s: the geometric era
- 1990s: appearance-based models
- Mid-1990s: sliding window approaches
- Late 1990s: local features
- Early 2000s: parts-and-shape models

Parts-and-shape models

- Model:
  - Object as a set of parts
  - Relative locations between parts
  - Appearance of part

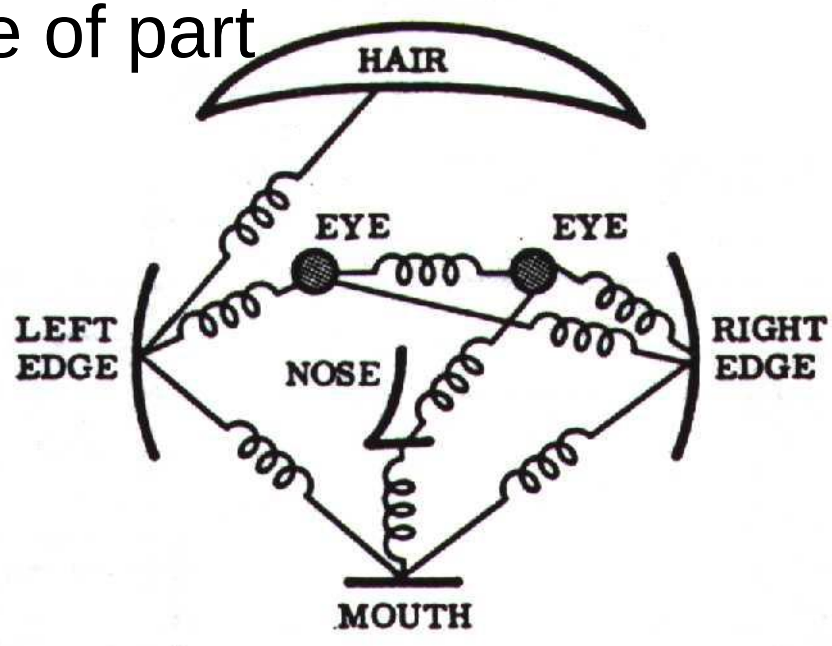
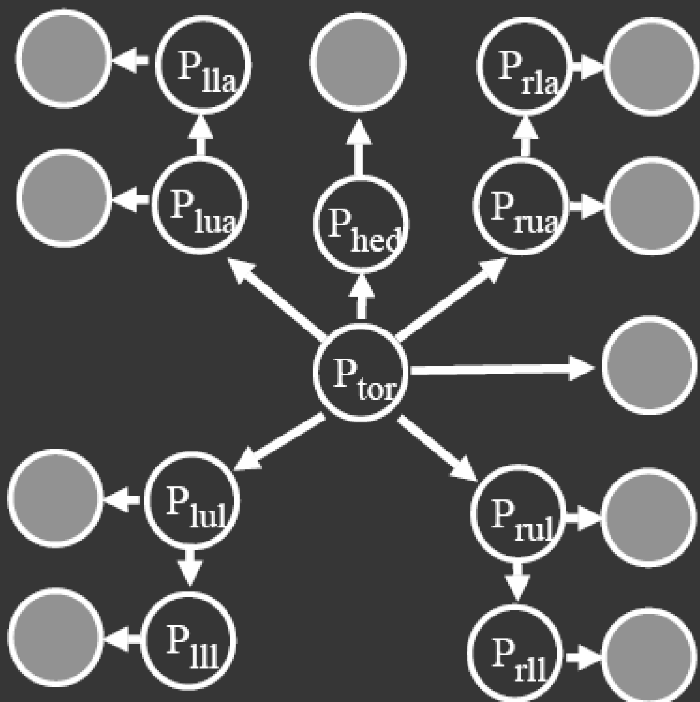
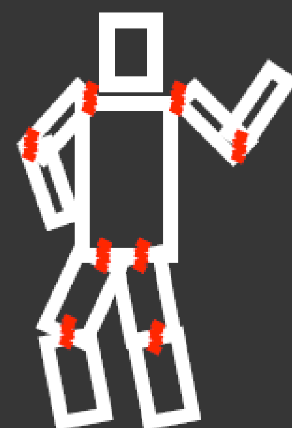


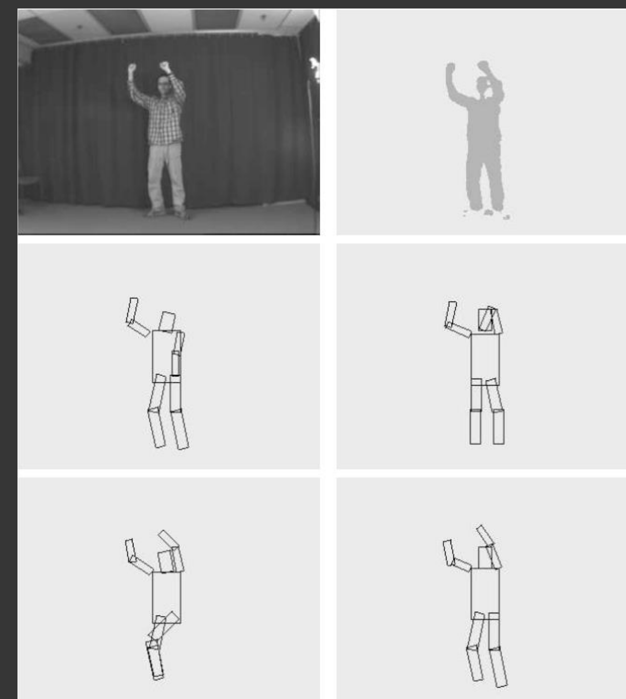
Figure from [Fischler & Elschlager 73]

Pictorial structure model

Fischler and Elschlager(73), Felzenszwalb and Huttenlocher(00)





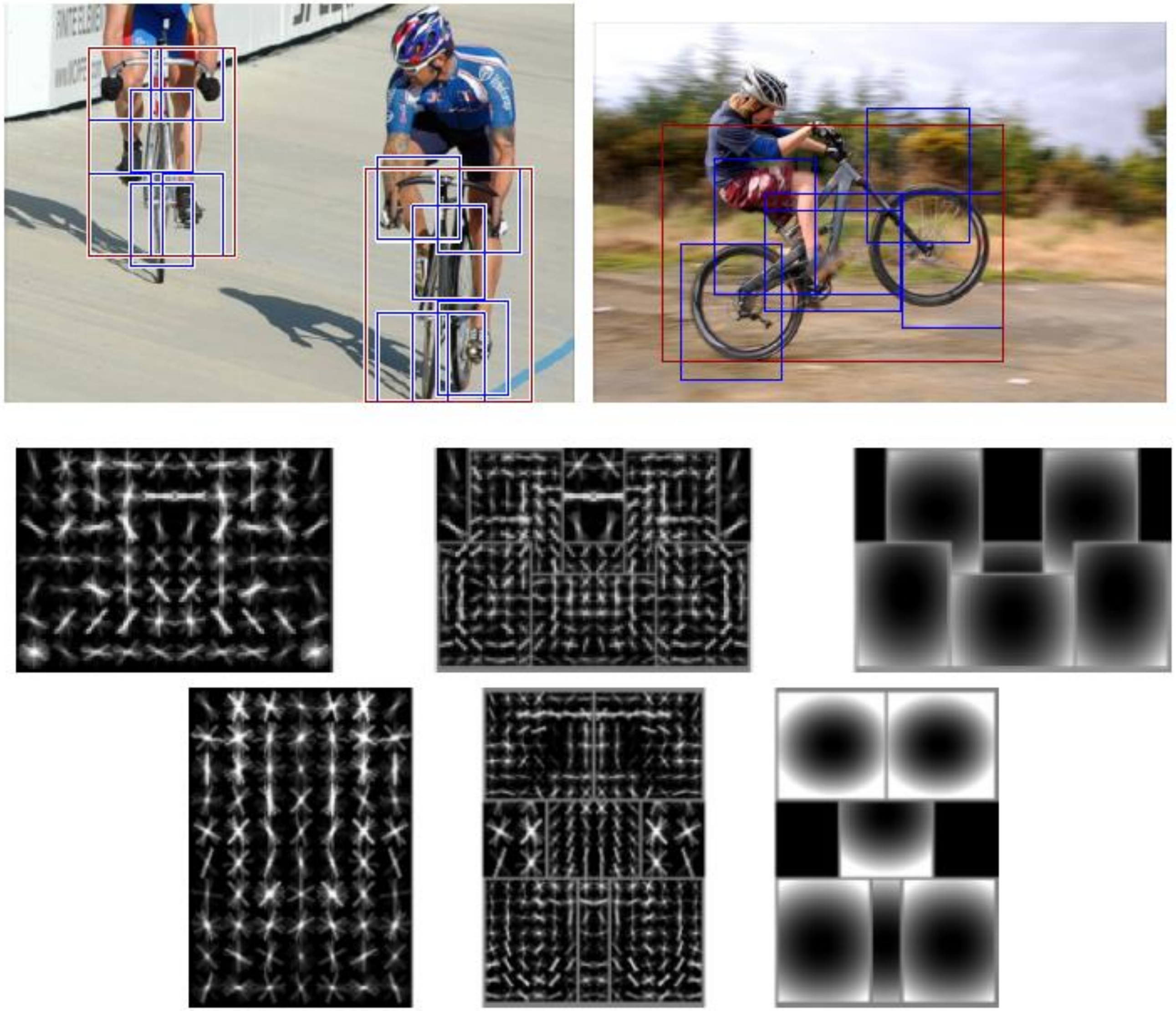


$$\mathrm{Pr}(\mathrm{P}_{\mathrm{tor}},\mathrm{P}_{\mathrm{arm}},\ldots|\mathrm{Im}) \quad \propto \prod_{i,j} \mathrm{Pr}(P_i \mid P_j) \prod_{i} \mathrm{Pr}(\mathrm{Im}(P_i))$$

part geometry

part appearance

Discriminatively trained part-based models



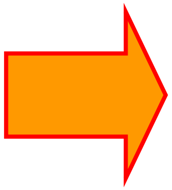
P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, **["Object Detection with Discriminatively Trained Part-Based Models,"]** PAMI 2009

History of ideas in recognition

- 1960s – early 1990s: the geometric era
- 1990s: appearance-based models
- Mid-1990s: sliding window approaches
- Late 1990s: local features
- Early 2000s: parts-and-shape models
- Mid-2000s: bags of features

Svetlana Lazebnik

Bag-of-features models



Svetlana Lazebnik

Bag-of-features models

Object → **Bag of ‘words’**

Svetlana Lazebnik

History of ideas in recognition

- 1960s – early 1990s: the geometric era
- 1990s: appearance-based models
- Mid-1990s: sliding window approaches
- Late 1990s: local features
- Early 2000s: parts-and-shape models
- Mid-2000s: bags of features
- Present trends: data-driven methods, **deep learning**