

# CS4495/6495

## Introduction to Computer Vision

---

8A-L1 *Introduction to recognition*

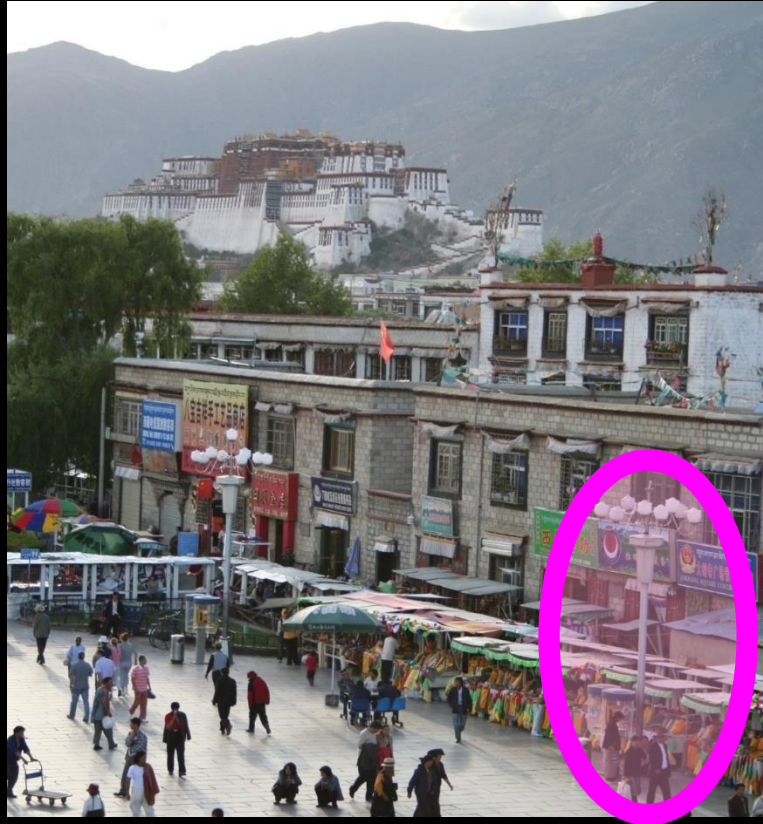


# What does recognition involve?

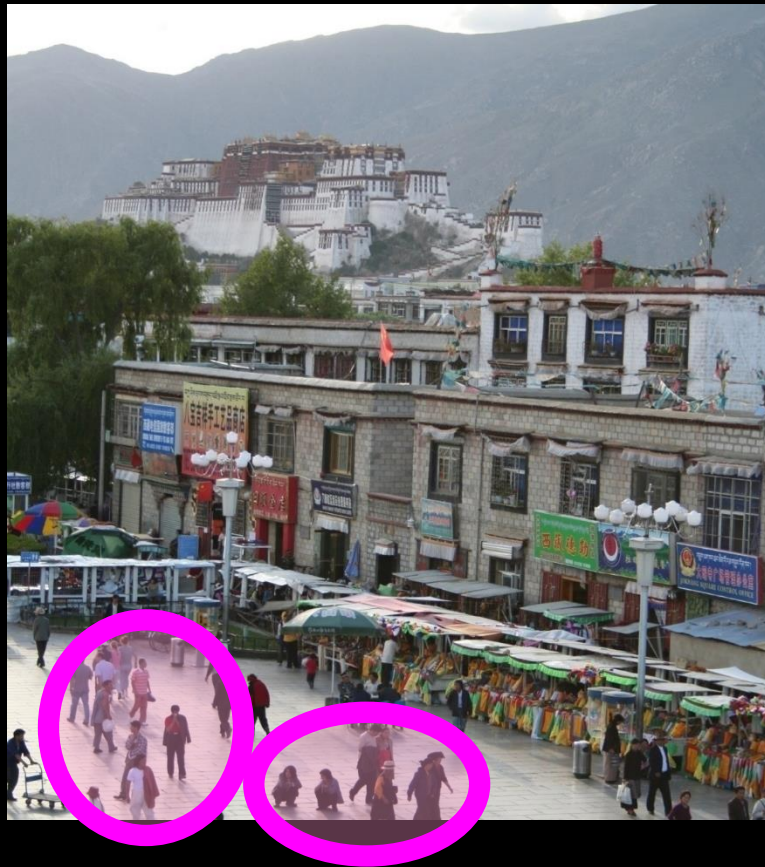


Source: Fei-Fei Li,  
Rob Fergus,  
Antonio Torralba.

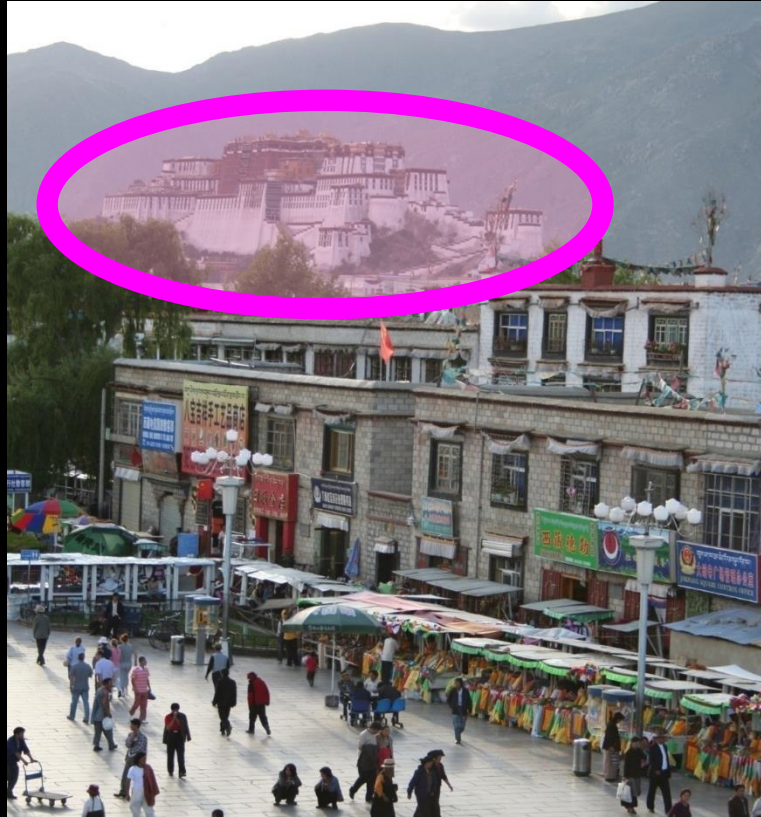
Verification: is that a lamp?



# Detection: are there people?



Identification: is that Potala Palace?

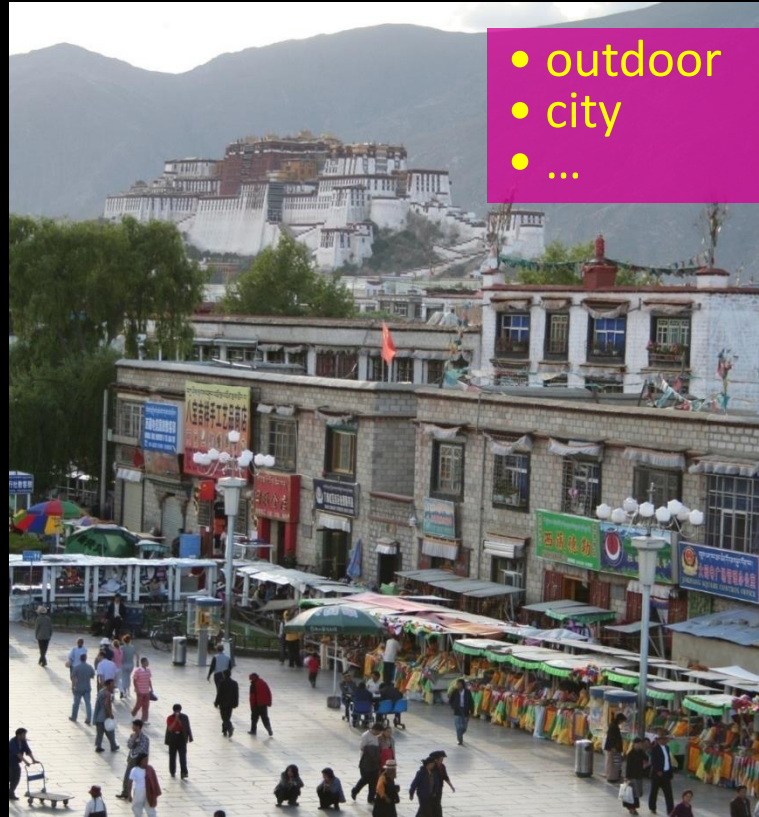




# *Object categorization*



# Scene and context categorization



# *Instance-level* recognition problem



John's car



# *Generic* categorization problem



# Object Categorization

*Task: Given a (small) number of training images of a category, recognize a-priori unknown instances of that category and assign the correct category label.*

# Object Categorization



Which categories are the best for visual identification?

# Visual Object Categories

## Basic Level Categories in human categorization

[Rosch 76, Lakoff 87]

- The highest level at which category members have similar perceived shape
- The highest level at which a single mental image reflects the entire category

# Visual Object Categories

## Basic Level Categories in human categorization

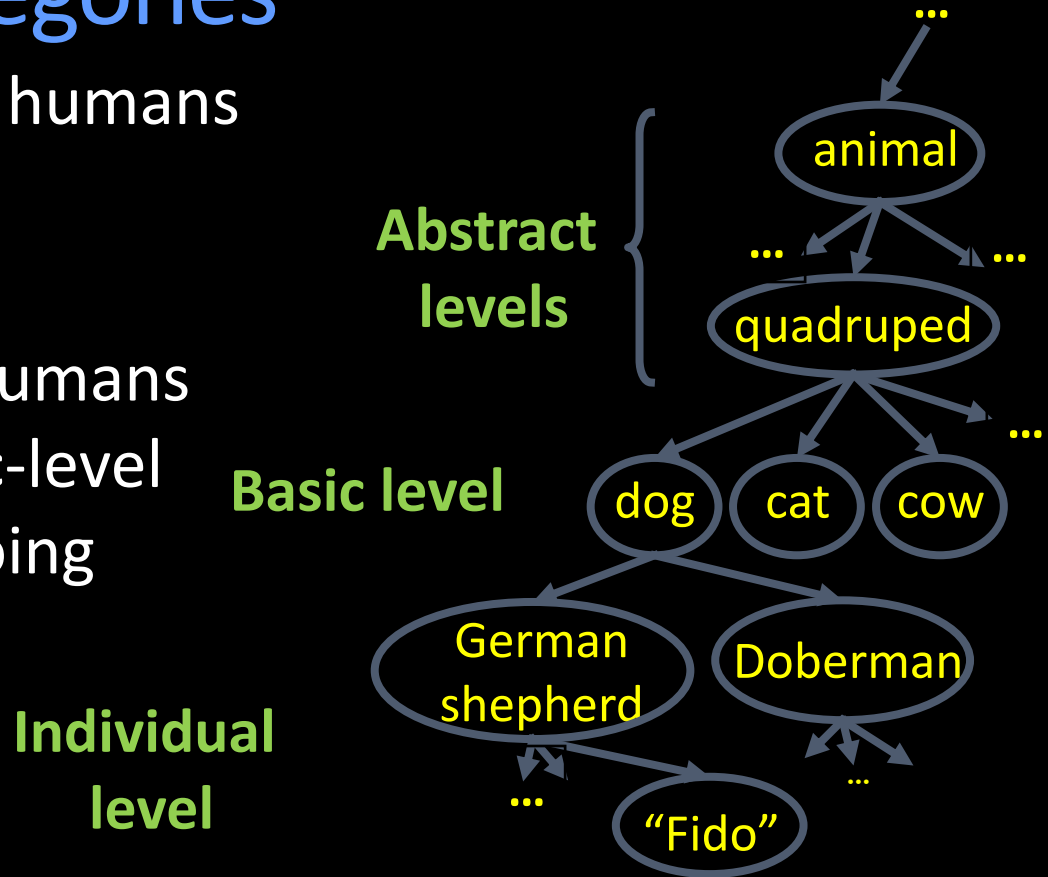
[Rosch 76, Lakoff 87]

- The level at which human subjects are usually fastest at identifying category members
- The first level named and understood by children
- The highest level at which a person uses similar motor actions for interaction with category members



# Visual Object Categories

- Basic-level categories in humans seem to be defined predominantly visually.
- There is evidence that humans (usually) start with basic-level categorization *before* doing identification.



DSpace@MIT : Natural Object Categorization - Mozilla Firefox

http://dspace.mit.edu/handle/1721.1/6964

Most Visited CoC Google Wells Fargo Home Page T-Square : Gateway : ... Readability Best Of The Big A Jazz24 Add GT Proxy U.S. Bank Healthcare P... xkcd.com computer vision g... Homepage of Alb... Generative vs Disc... Georgia Tech Libr... potala palace - Go... Bishop-Sicity-05.p... Georgia Tech Libr... DSpace@MIT : ... x +

DSpace@MIT MITLibraries Login

Home Computer Science and Artificial Intelligence Lab (CSAIL) Artificial Intelligence Lab Publications AI Technical Reports (1964 - 2004)

Search DSpace@MIT

☒ Search DSpace@MIT

☐ This Collection

[Advanced Search](#)

Browse

All of DSpace@MIT

[Communities & Collections](#)

[By Issue Date](#)

[Authors](#)

[Titles](#)

[Subjects](#)

This Collection

[By Issue Date](#)

[Authors](#)

[Titles](#)

[Subjects](#)

My Account

[Login](#)

[Register](#)

Links

[About DSpace@MIT](#)

Find: generative

☒ Next ☐ Previous ☐ Highlight all ☐ Match case

# Natural Object Categorization

[Show full item record](#)

<http://hdl.handle.net/1721.1/6964>

Bobick, Aaron F.

1987-11-01

Abstract:

claim about the structure of the world. We first design an evaluation function that measures how well a set of categories supports the inference goals of the observer. Entropy measures for property uncertainty and category uncertainty are combined through a free parameter that reflects the goals of the observer. Natural categorizations are shown to be those that are stable with respect to this free parameter. The evaluation function is tested in the domain of leaves and is found to be sensitive to the structure of the natural categories corresponding to the different species. We next develop a categorization paradigm that utilizes the categorization evaluation function in recovering natural categories. A statistical hypothesis generation algorithm is presented that is shown to be an effective categorization procedure. Examples drawn from several natural domains are presented, including data known to be a difficult test case for numerical categorization techniques. We next extend the categorization paradigm such that multiple levels of natural categories are recovered; by means of recursively invoking the categorization procedure both the genera and species are recovered in a population of anaerobic bacteria. Finally, a method is presented for evaluating the utility of features in recovering natural categories. This method also provides a mechanism for determining which features are constrained by the different processes present in a multiple modal world.

Technical Report 1001

# Natural Object Categorization

Aaron F. Bobick

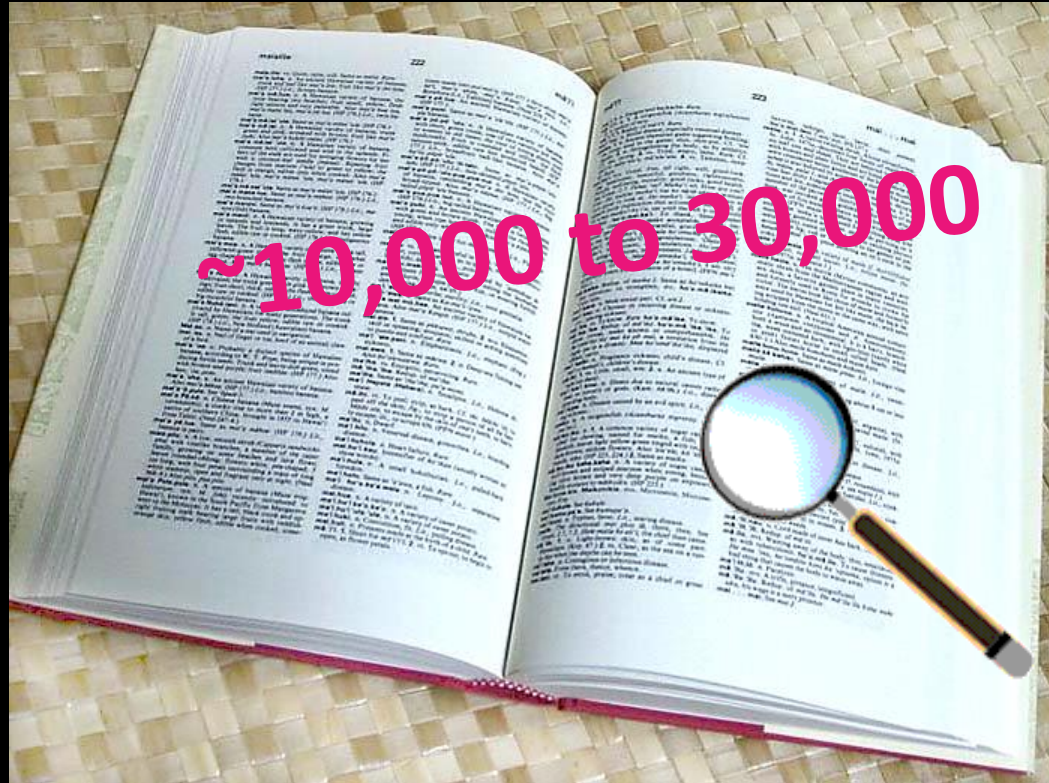
MIT Artificial Intelligence Laboratory

# Object Categorization



Which categories are the best for visual identification?

# How many object categories are there?



Biederman 1987



# Other Types of Categories

## Functional Categories

e.g. chairs = “*something you can sit on*”

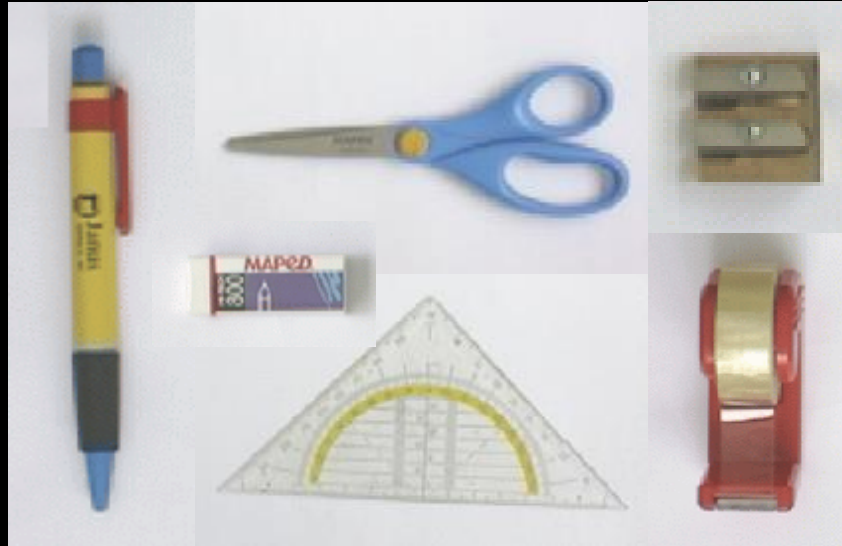


K. Grauman, B. Leibe

# Other Types of Categories

Ad-hoc categories

e.g. *“something you can find in an office environment”*



K. Grauman, B. Leibe

# Autonomous agents able to detect objects

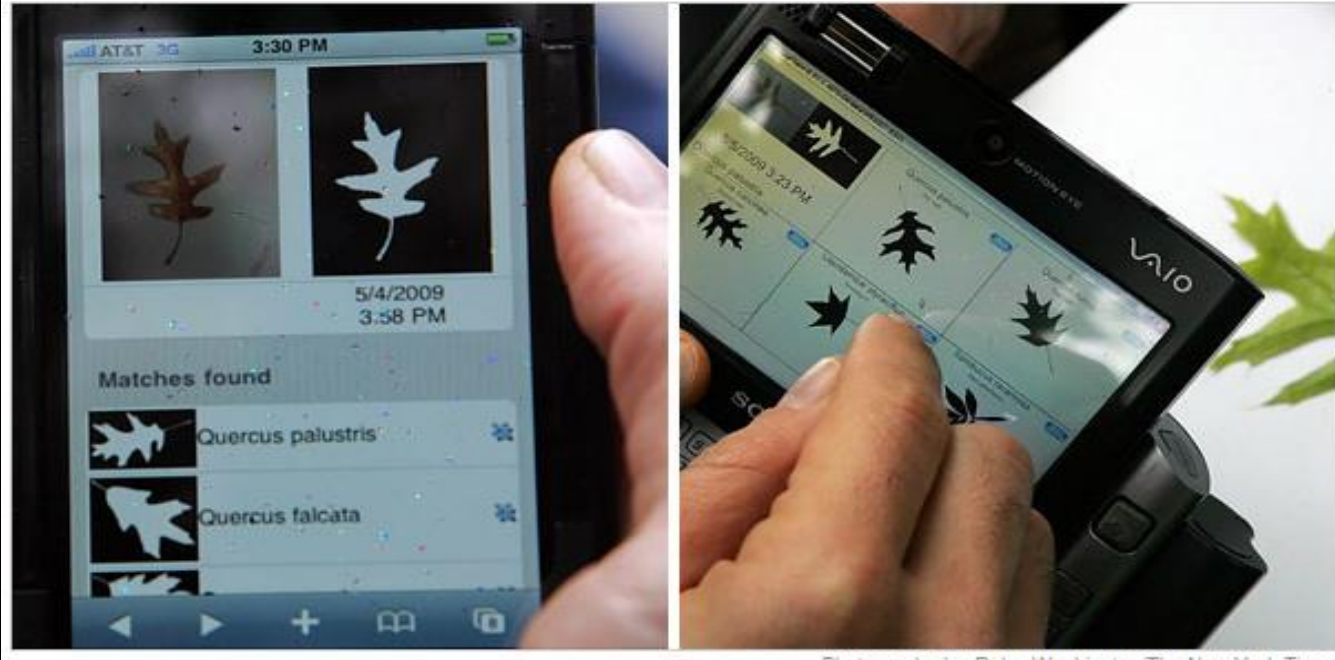


# Labeling people



# Posing visual queries


## Digital Field Guides Eliminate the Guesswork



Belhumeur et al.



# Finding visually similar objects


**like**  
visual shopping *alpha*


[My Like List](#) | [NewsLetter](#) | [Blog](#)


ALL SHOES BAGS WOMEN'S APPAREL MEN'S APPAREL KIDS ACCESSORIES JEWELRY & WATCHES HOLIDAY FOR THE HOME


IN Women's Shoes Search

Refine by Style


  
Pumps


  
Sandals


  
Flats


  
Patent

Refine by Color


  
crimson

  
taupe

  
scarlet

  
Clarks

Refine by Brand


  
Softt


Why is Like.com Different?

Like is a visual shopping engine that lets you find items by color, shape and pattern.

Click on [Likeness Search](#) to get started

Your Search Item

  
Which part of the image do you like?  
Draw a box on the item to focus your search on that area.



Cole Haan - Carma OT Air Pump



\$278.95

[More Details](#) [Save to LikeList](#)


[Shop at Zappos.com](#)

All Products > Shoes > Women's Shoes > Cole Haan > Cole Haan - Carma OT Air Pump

Search Results

Sort By [Likeness™](#) [Price](#) Change Your View:   Results 1 - 20 of 140,207


1 2 3 4 5 6 7 NEXT >>



Natural Comfort - LV58

a sexy classic pump with a pillow-like footbed to keep your feet happy. leather or patent leather upper. wrapped memory-foam footbed. covered heel. leather sole.


[Compare Prices](#) [More Details](#) [Save to LikeList](#)



Cole Haan 'Carma Air' Patent Leather Open Toe Pump

Open toe styles a sleek, cushioned pump with a wrapped heel and a mini platform. Color(s): black patent, dark chocolate suede, wine patent, black python, natural python, beige leather. Brand: Cole Haan.

[Compare Prices](#) [More Details](#) [Save to LikeList](#)



rsvp - Caitlyn

an easy on the eyes pump features craftsmanship to make it easy on your feet too. patent leather uppers. almond shaped toe. cushioned footbed. covered heel. leather outsole. made in brazil. 7 oz.

[More Details](#) [Save to LikeList](#)

\$99.95

[Shop at Zappos.com](#)

Free Shipping Available

Shop for more items like this:

[Likeness Search](#)

\$275.00

[Shop at MORDSTROM.com](#)

Shop for more items like this:

[Likeness Search](#)

\$89.95

[Shop at Zappos.com](#)

Free Shipping Available

Shop for more items like this:

[Likeness Search](#)

So why is this hard?

# Challenges: Robustness



**Illumination**



**Object pose**



**Clutter**



# Challenges: Robustness



**Occlusions**



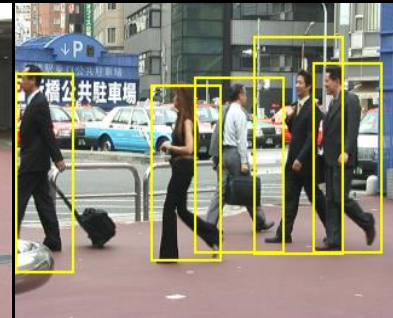
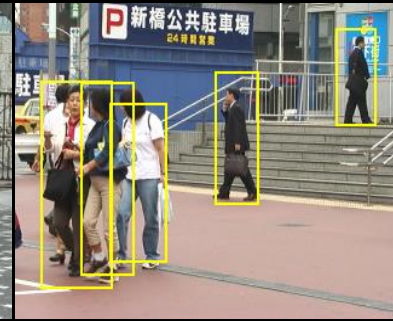
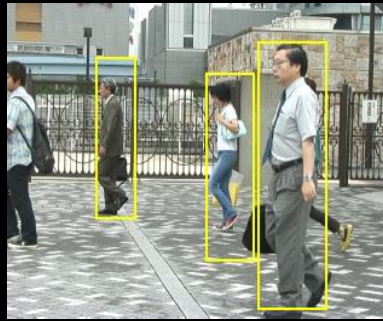
**Intra-class  
appearance**



**Viewpoint**

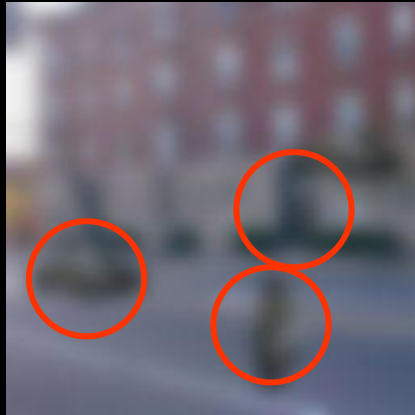
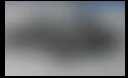
# Challenges: Robustness

Realistic scenes are crowded, cluttered, have overlapping objects.



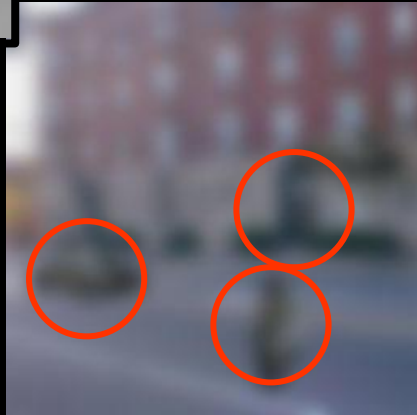
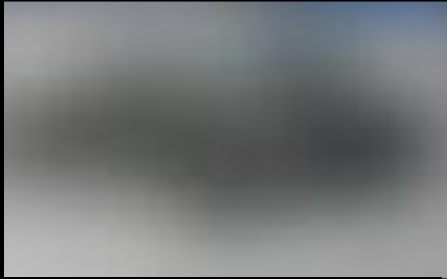


# Challenges: Importance of context



Fei-Fei,  
Fergus & Torralba

# Challenges: Importance of context



Fei-Fei,  
Fergus & Torralba

# Challenges: complexity

- Thousands to millions of pixels in an image
- 3,000-30,000 human recognizable object categories
- 30+ degrees of freedom in the pose of articulated objects (humans)

# Challenges: complexity

- Billions of images indexed by Google Image Search
- In 2011, 6 billion photos uploaded *per month*
- Approx one billion million camera phones sold in 2013
- About half of the cerebral cortex in primates is devoted to processing visual information  
[Felleman and van Essen 1991]

So what works?

# What worked most reliably “yesterday”

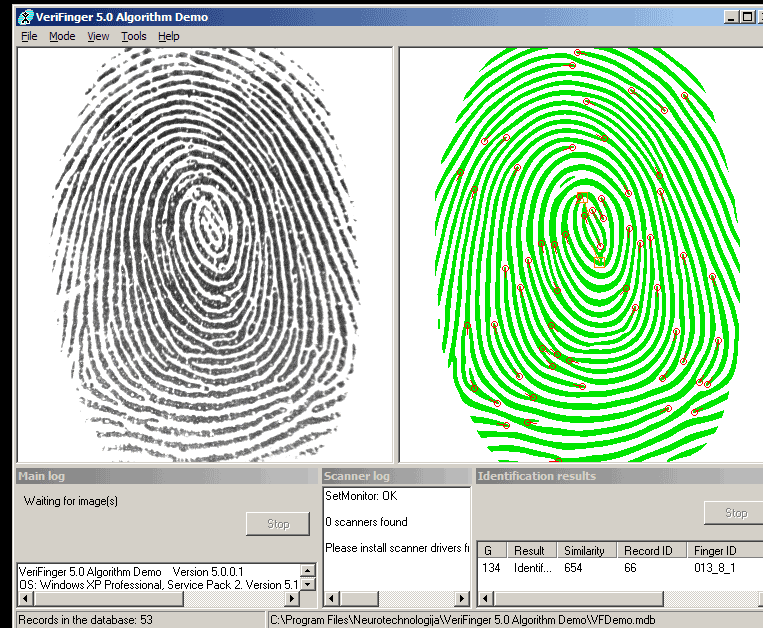
- Reading license plates (real easy), zip codes, checks





# What worked most reliably “yesterday”

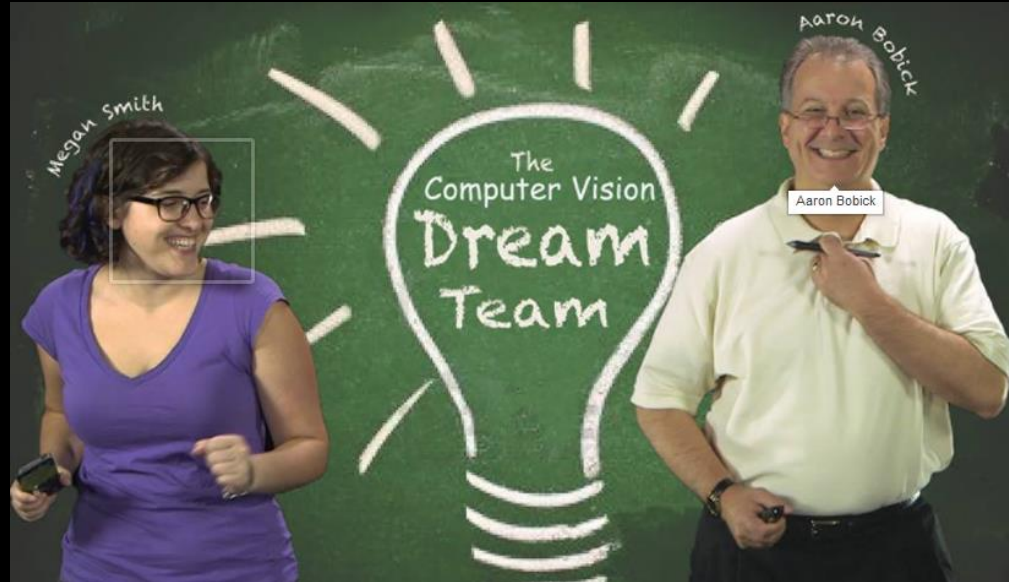
- Reading license plates, zip codes, checks
- Fingerprint recognition



Lana Lazebnik

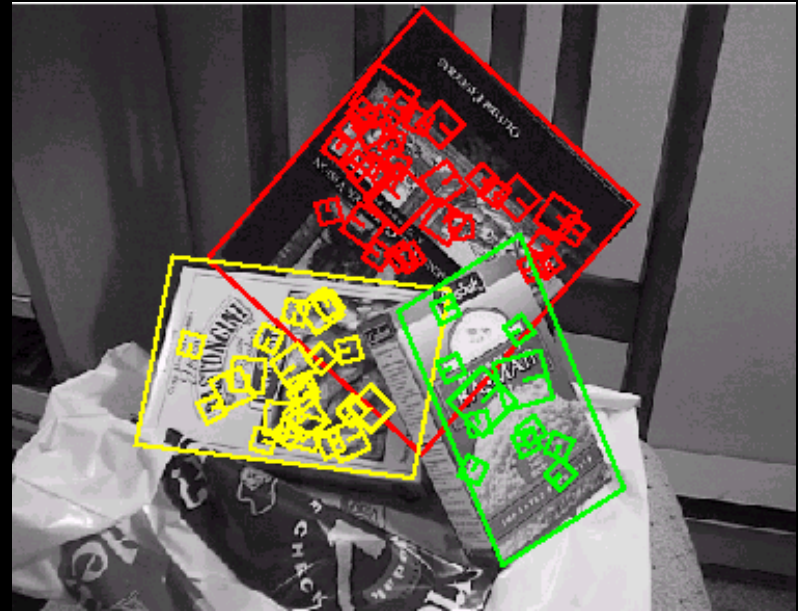
# What worked most reliably “yesterday”

- Reading license plates, zip codes, checks
- Fingerprint recognition
- Face detection  
(Today recognition)

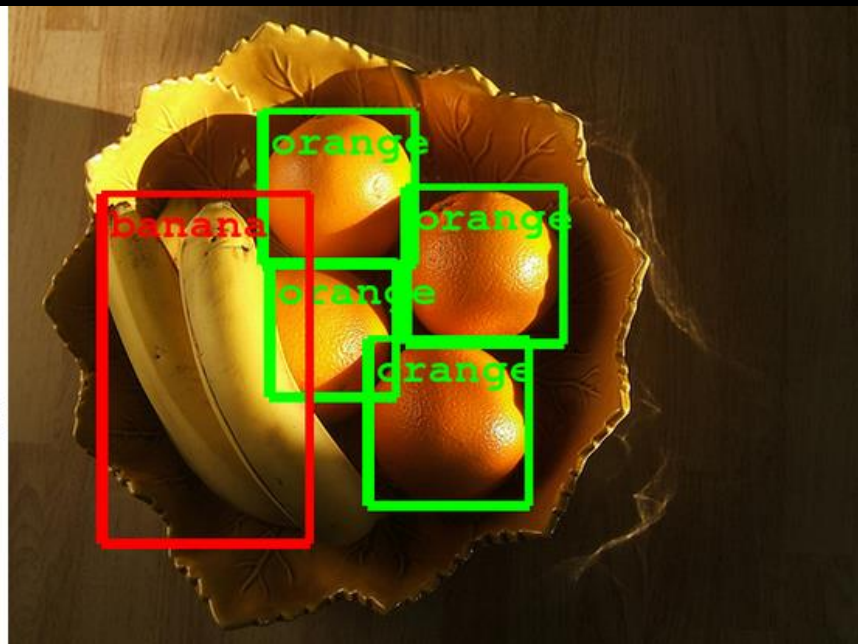
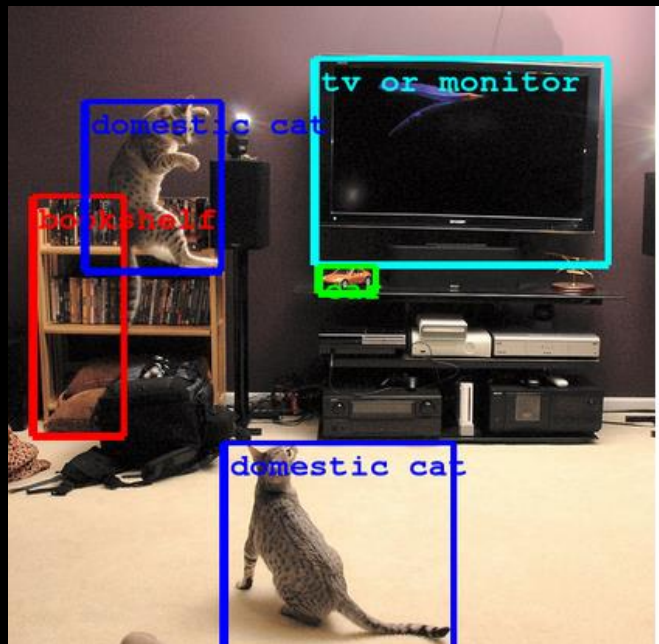


# What worked most reliably “yesterday”

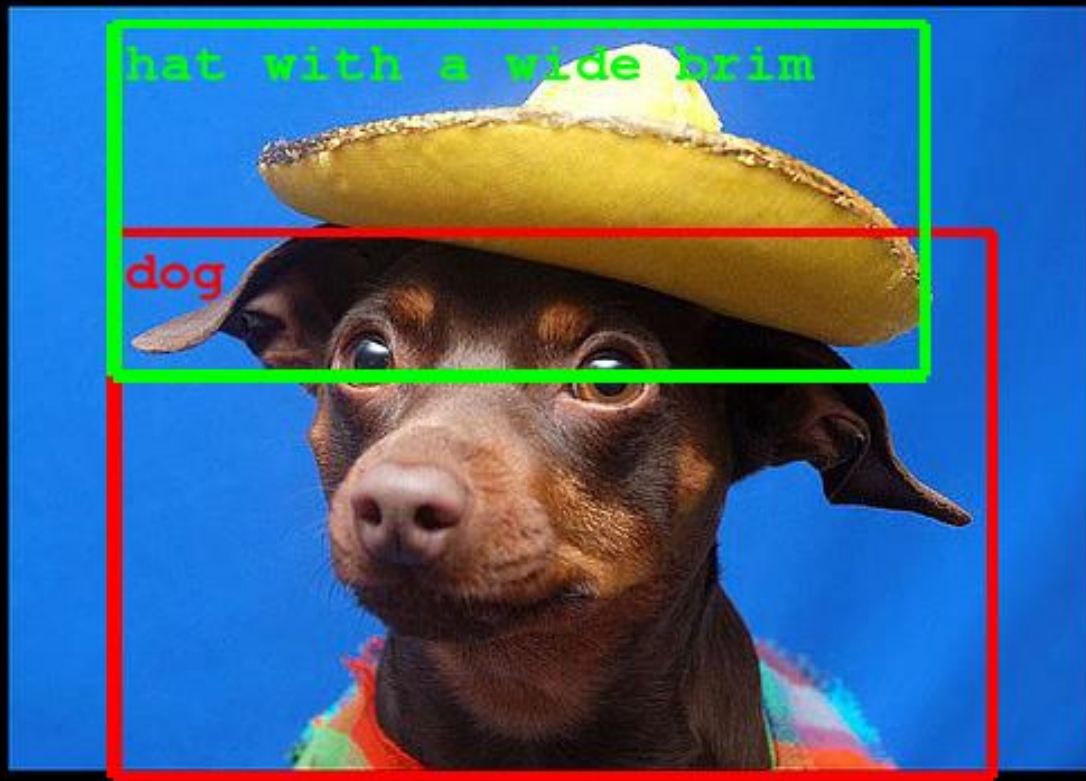
- Reading license plates, zip codes, checks
- Fingerprint recognition
- Face detection  
(Today recognition)
- Recognition of flat textured objects  
(CD covers, book covers, etc.)



# Just in: GoogleNet 2014



Just in: GoogleNet – no context needed?



# Going forward

- Much of strong label “recognition” is really machine learning applied to patterns of pixel intensities.
- We’ll focus on some principles of generative vs discriminative methods and image representations that they use.
- Then we’ll spend some time on *activity recognition* from video.