CS 395T
Visual Recognition

Thursday, January 18

Today

- Course logistics
- Overview
- Volunteers, prep for next week

Administration

- **Class**: Tues / Thurs 12:30-2 PM
- **Instructor**: Kristen Grauman
  grauman at cs.utexas.edu
  TAY 4.118
- **Office hours**: Tues / Thurs 3-4 PM
- **Class page**: link from
  http://www.cs.utexas.edu/~grauman
  ** Check for updates to schedule.

Course content

- Focus on current research in visual category and object recognition
- High-level vision and learning problems

We will *not* spend much time on low-level image processing, video-based techniques, particular vision systems, human vision system.

Expectations

- **Discussions** will center on recent papers in the field
  - Paper reviews, prepared discussion points
- **Student presentations**
  - Paper content, demos (extra credit)
- **Projects**
  - Research-oriented

Paper reviews

- For each class, choose **one** of the 2-3 papers we are covering to review
- Reviews due via email to me **before class**
- Posted for our class (anonymously)

Also, prepare (write down) a few discussion points to have on hand in class about all of the papers you read.

Paper review guidelines

- Brief (2-3 sentences) summary
- Main contribution
- Strengths? Weaknesses?
- How convincing are the experiments? Suggestions to improve them?
- Extensions?
- Additional comments.

More is not necessarily more.

Presentation guidelines

- Approx. 25 minutes
- Clear overview of the paper
- Consider:
  - Main problem, motivation
  - Assumptions
  - Technical approach: high level and intuition
  - Important technical details
  - Experiments
  - Connections to other papers

Demo guidelines

Implement/download code for a main idea in the paper and show us toy example(s):

- Experiment with different types of (mini) training/testing data sets
- Evaluate sensitivity to parameter settings
- Show (on a small scale) an example in practice that highlights a strength/weakness of the approach
- …

Projects

Possibilities:

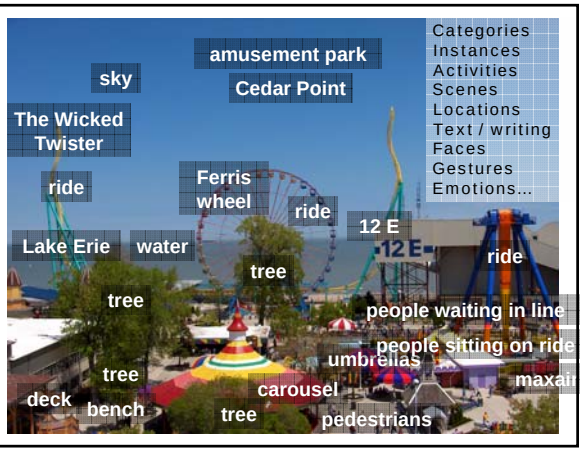
- Extend a technique studied in class
- Empirical evaluation and analysis of a few related techniques
- Design and evaluate a novel approach

- May be possible to tie it into your research
- Work in pairs
- Proposal due at midterm (March 8)
- Short presentation at end of term, paper

What is visual recognition?

- Perception of familiar objects
- Given image data, determine what's in it, and where

- Detection, categorization, identification



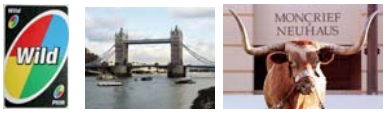
Why recognition?

- Fundamental problem in computer vision
- Area is rich with very challenging questions
- Applications …

Our focus in this class (primarily)

butterfly butterfly building building

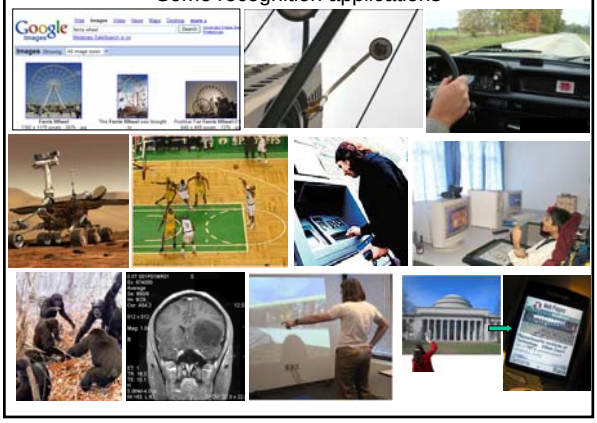
Categories



Wild card Tower Bridge Bevo

Specific

Some recognition applications



Key challenges: robustness


Illumination



Object pose


Clutter


Occlusions


Intra-class appearance


Viewpoint

Key challenges: efficiency

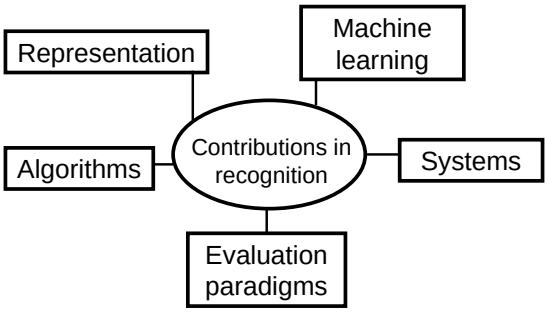
- Thousands to millions of pixels in an image
- 3,000-30,000 human recognizable object categories
- 30+ degrees of freedom in the pose of articulated objects (humans)
- Billions of images indexed by Google Image Search
- 18 billion+ prints produced from digital camera images in 2004
- 295.5 million camera phones sold in 2005

Disciplines

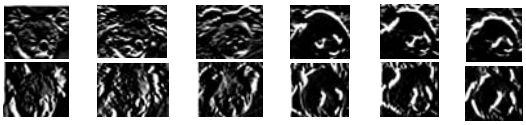
Recognition (and vision in general) draws from

- Machine learning
- Probability
- Geometry, physics
- AI
- Algorithms
- Image processing
- Cognitive science

Recognition research



Global image representations



Map image to a single vector based on overall characteristics

- vector of pixel intensities
- grayscale / color histogram
- bank of filter responses …

Local image representations

Describe component regions or patches separately


SIFT [Lowe]


Shape context [Belongie et al.]

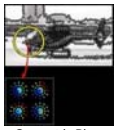

Superpixels [Ren et al.]


Maximally Stable Extremal Regions [Matas et al.]

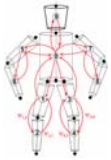

Salient regions [Kadir et al.]


Harris-Affine [Schmid et al.]

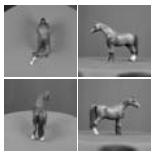

Spin images [Johnson and Hebert]


Geometric Blur [Berg et al.]

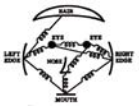
Representing object shape, geometry



3-D models



View-based models



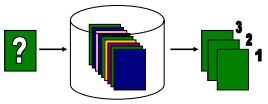
Parts + structure

Bag of features

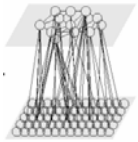
Image from Sigal et al. | Images from ETH-80 | Image from Fischler and Elschlager | Image from ICCV short course

Algorithms

optimization

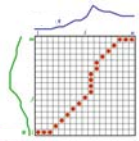


indexing



graphs

matching



distances

ASM image from T. Cootes, Graph image from S. Yu

Learning

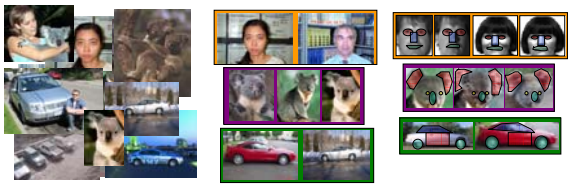
- What defines a category/class?
- What distinguishes classes from one another?
- How to understand the connection between the real world and what we observe?
- What features are most informative?
- What can we do without human intervention?
- Does previous learning experience help learn the next category?

Inputs/outputs/assumptions

- What **input** is available?
  - Static grayscale image
  - 3D range data
  - Video sequence
  - Multiple calibrated cameras
  - Segmented data, unsegmented data
  - Labeled data, unlabeled data, partially labeled data
- What is the **goal**?
  - Say yes/no as to whether an object present in image
  - Categorize all objects
  - Forced choice from pool of categories
  - Bounding box on object
  - Full segmentation
  - Build a model of an object category

Spectrum of supervision

Less — **More**



Category recognition: state-of-the-art

What's possible now?

What's difficult now?

One way to measure: benchmark data sets.

Category recognition: state-of-the-art

PASCAL Visual Object Classes Challenge 2006

- 10 categories
- Unsegmented, realistic images
- Supervised setup
- Classification task: For each class, predict presence/absence of an example of that class in the test image.
- 26 teams/methods competed

Bicycles

- Highest ranked class images

Images thanks to Mark Everingham

Bicycles

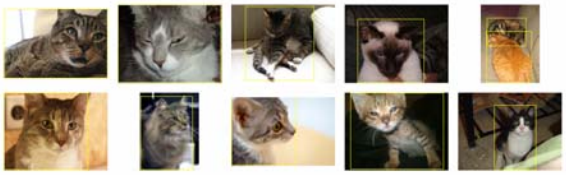
- Lowest ranked class images

- Highest ranked non-class images

Images thanks to Mark Everingham

Cats

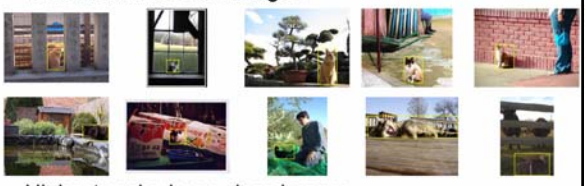
- Highest ranked class images



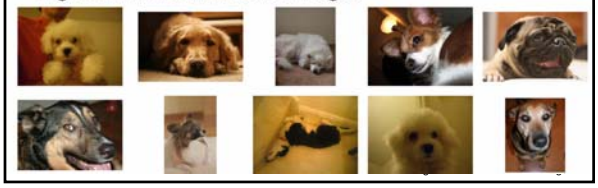
Images thanks to Mark Everingham

Cats

- Lowest ranked class images

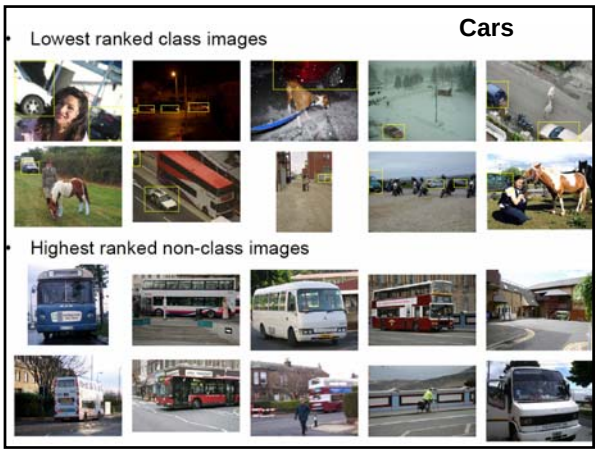


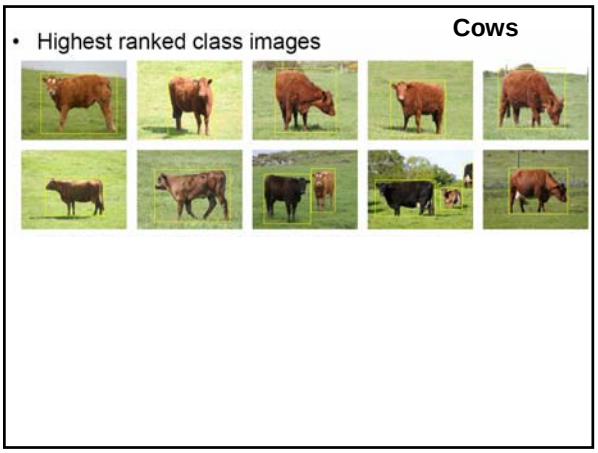
- Highest ranked non-class images

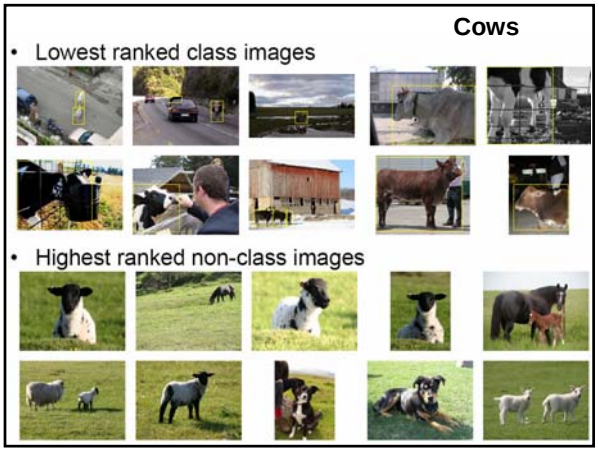


Cars

- Highest ranked class images





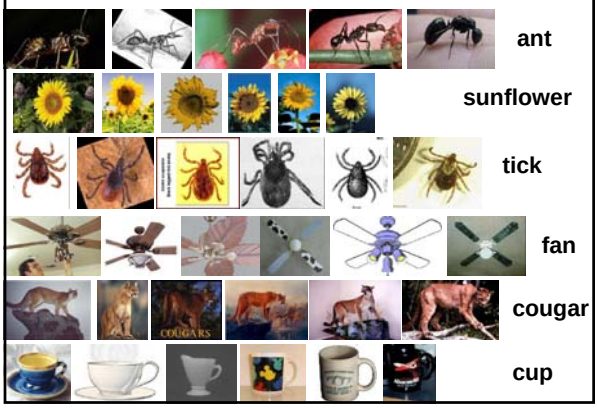


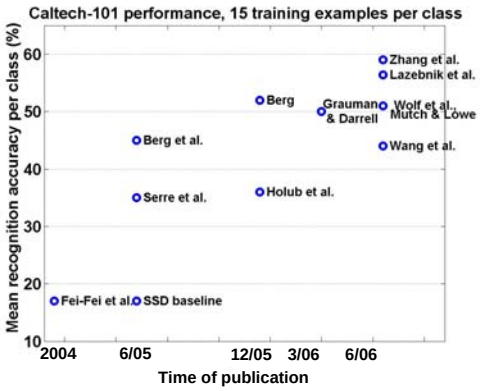
Category recognition: state-of-the-art

Caltech-101 Database

- 101 categories
- Wide appearance variation
- Images fairly centered and scaled similarly
- Supervised setup
- Classification task: predict class for test images
- Around 12 methods tested in literature

Example images from the Caltech-101 database





Topics

Through readings in recent vision literature…

- part-based models for recognition
- invariant local features
- bags of features and feature vocabularies
- spatial constraints and geometry
- shape descriptors and matching
- learning similarity measures
- fast indexing methods
- recognition with text and images
- the role of context in recognition
- unsupervised category discovery

Goals of this course

- Understand current approaches
- Analyze
- Identify interesting research questions

Coming up

- For tomorrow:
  - Send me 4-5 paper preferences for presentations up to spring break
- For Tuesday Jan 23
  - *Face Recognition Using Eigenfaces* by Turk and Pentland
  - *Face Recognition Using Active Appearance Models* by Edwards et al.
  - Bring discussion points
  - Review one of the papers (email to grauman@cs by Tuesday 12:30)
  - Demos preferences
- For Thursday Jan 25
  - *Rapid Object Detection Using a Boosted Cascade of Simple Features* by Viola and Jones
  - *Face Recognition by Humans* by Sinha et al.
  - Bring discussion points
  - Review on Viola and Jones

Coming up

- Presentation volunteers
- Quick survey