Honors Machine Vision

Jan 17, 2017

Kristen Grauman, University of Texas at Austin

Introductions

- **Instructor**: Prof. Kristen Grauman
- **TA**: Dongguang You

Today

- Course overview
- Requirements, logistics

What is computer vision?

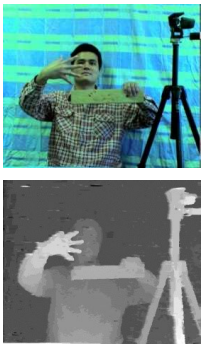
Done?

Computer Vision

- Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data *(measurement)*

1. Vision for measurement

Real-time stereo



Wang et al.

Structure from motion

Snavely et al.

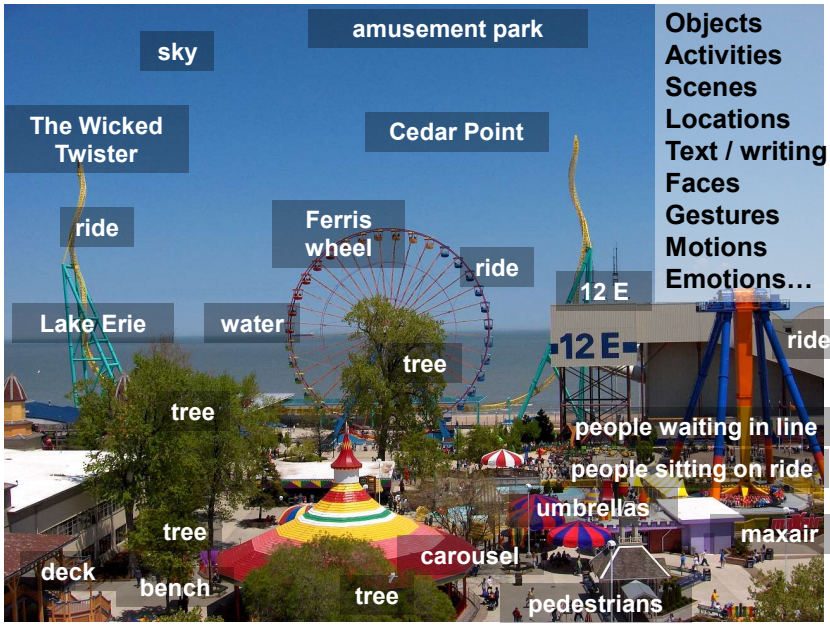
Tracking

Demirdjian et al.

Computer Vision

- Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data *(measurement)*
  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*

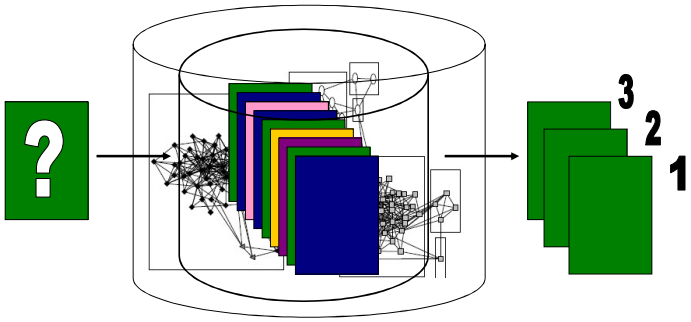
2. Vision for perception, interpretation



Computer Vision

- Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data *(measurement)*
  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*
  3. Algorithms to mine, search, and interact with visual data (*search and organization*)

3. Visual search, organization

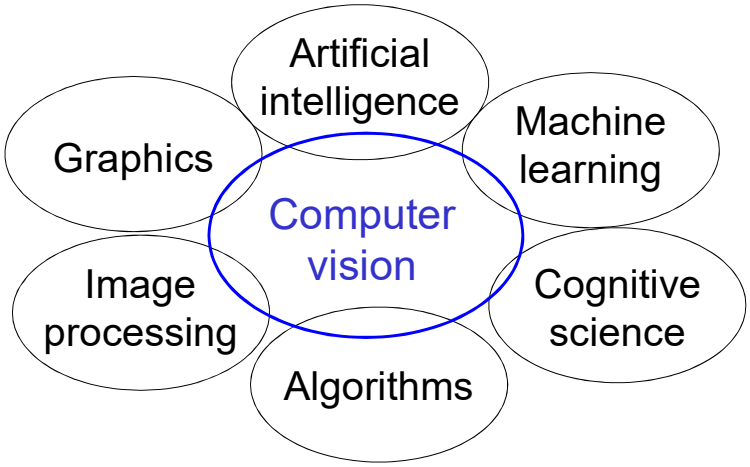


Computer Vision

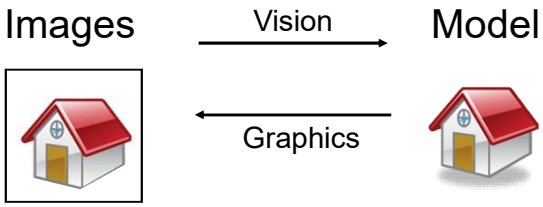
- Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data *(measurement)*
  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*
  3. Algorithms to mine, search, and interact with visual data (*search and organization*)

Course focus

Related disciplines

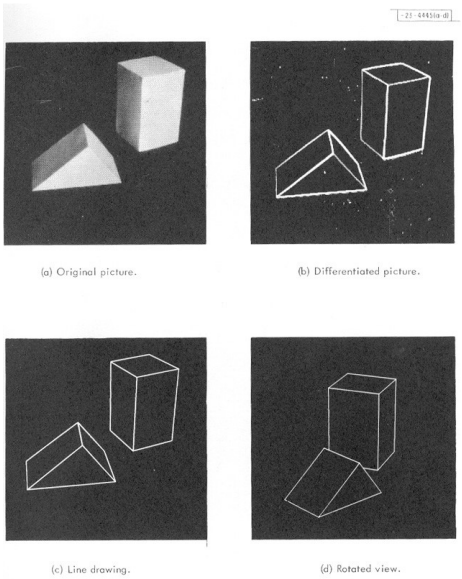


Vision and graphics



Inverse problems: analysis and synthesis.

Visual data in 1963



L. G. Roberts, *Machine Perception of Three Dimensional Solids*, Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

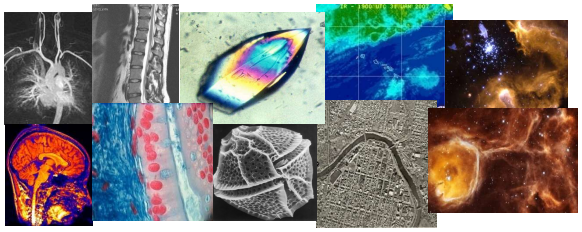
Visual data in 2017

Personal photo albums

Movies, news, sports

Google Image Search Picasa flickr webshots picsearch YouTube Broadcast Yourself

Surveillance and security



Medical and scientific images

Slide credit; L. Lazebnik

Why vision?

- As image sources multiply, so do applications
  - Relieve humans of boring, easy tasks
  - Enhance human abilities
  - Advance human-computer interaction, visualization
  - Perception for robotics / autonomous agents
  - Organize and give access to visual content

Faces and digital cameras

Camera waits for everyone to smile to take a photo [Canon]

Setting camera focus via face detection

Linking to info with a mobile device

Situated search
Yeh et al., MIT

Google Goggles

MSR Lincoln

kooaba

Video-based interfaces

Human joystick, NewsBreaker Live

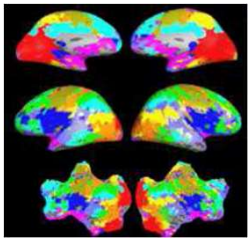
Assistive technology systems
Camera Mouse, Boston College

Microsoft Kinect

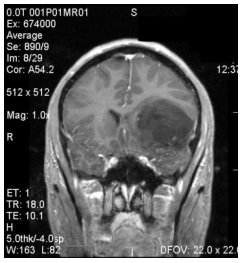
What else?

Vision for medical & neuroimages

Image guided surgery
MIT AI Vision Group



fMRI data
Golland et al.



Special visual effects

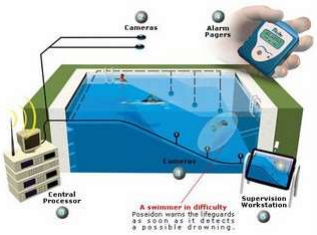
The Matrix

What Dreams May Come

Mocap for *Pirates of the Carribean*,
Industrial Light and Magic
Source: S. Seitz

Safety & security

Navigation,
driver safety



Monitoring pool
(Poseidon)

Pedestrian detection
MERL, Viola et al.

Surveillance

Obstacles?

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

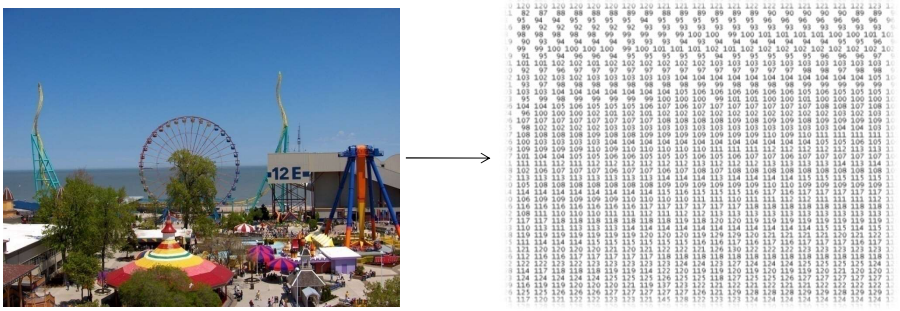
July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

What the computer gets



Why is vision difficult?

- Ill-posed problem: real world much more complex than what we can measure in images
  - 3D → 2D
- Impossible to literally “invert” image formation process

Challenges: many nuisance parameters

Illumination

Object pose

Clutter

Occlusions

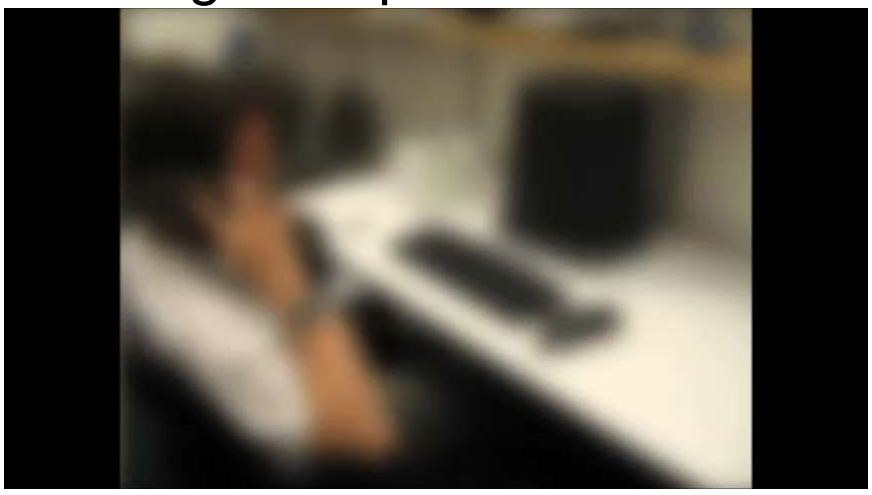
Intra-class appearance

Viewpoint

Challenges: intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: importance of context



Challenges: importance of context

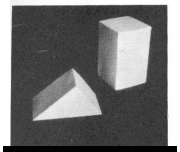
Challenges: importance of context

slide credit: Fei-Fei, Fergus & Torralba

Challenges: complexity

- Millions of pixels in an image
- 30,000 human recognizable object categories
- 30+ degrees of freedom in the pose of articulated objects (humans)
- Billions of images online
- 144K hours of new video on YouTube daily
- …
- About half of the cerebral cortex in primates is devoted to processing visual information [Felleman and van Essen 1991]

Progress charted by datasets



Roberts 1963

COIL

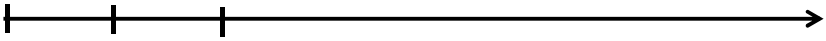
1963 ⋯ 1996

Progress charted by datasets

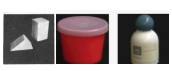
MIT-CMU Faces

INRIA Pedestrians

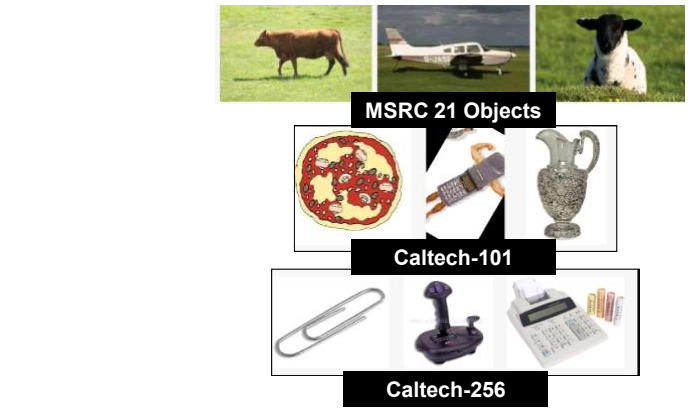
UIUC Cars



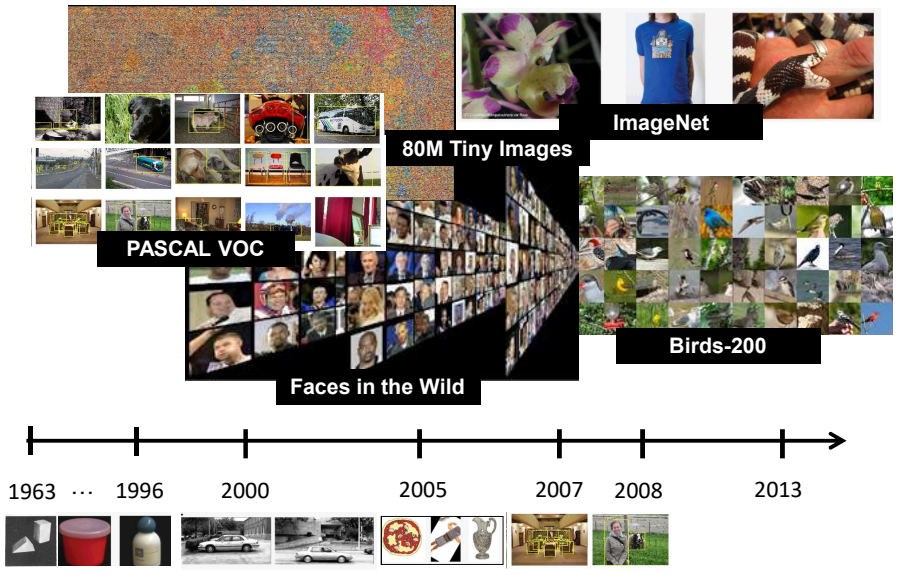
1963 ⋯ 1996 2000



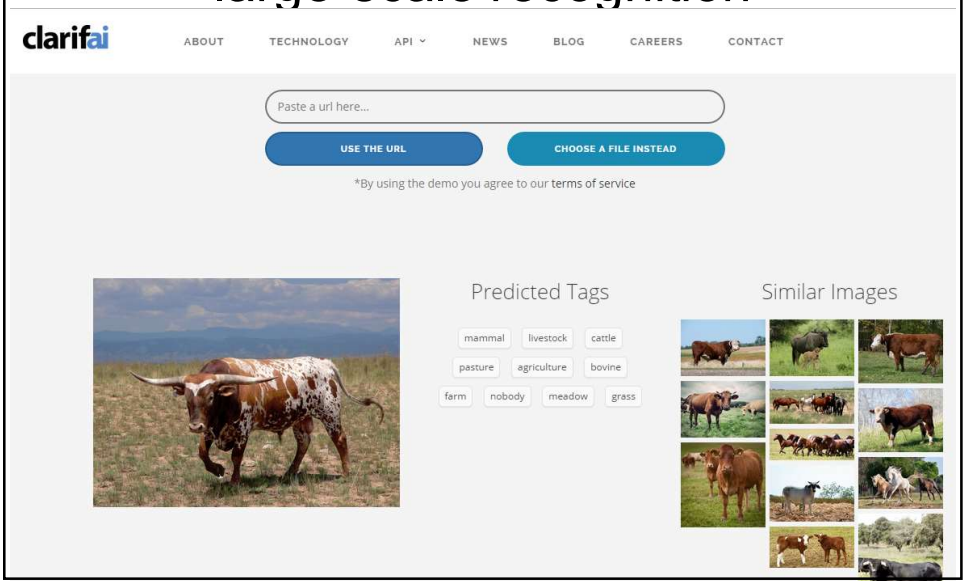
Progress charted by datasets



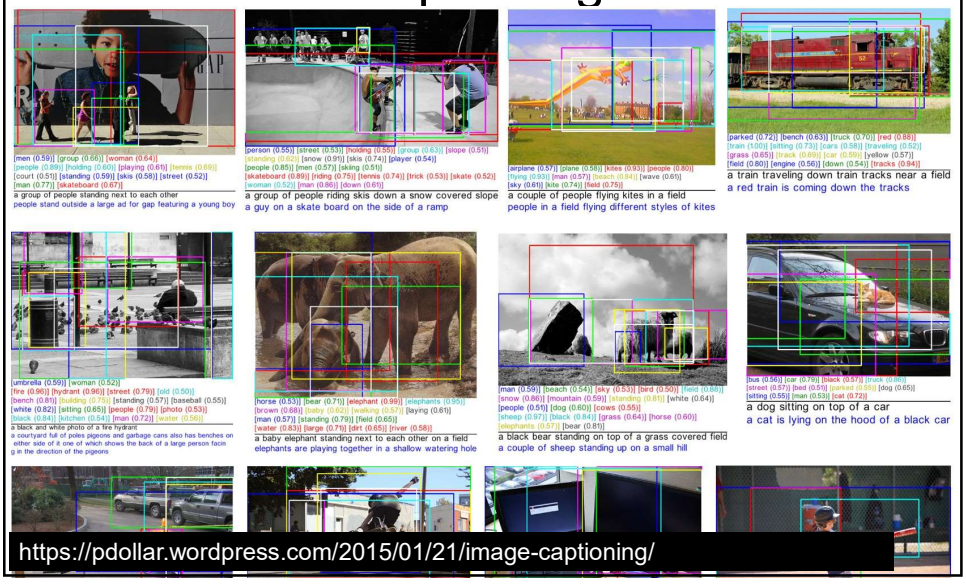
Progress charted by datasets



Expanding horizons: large-scale recognition

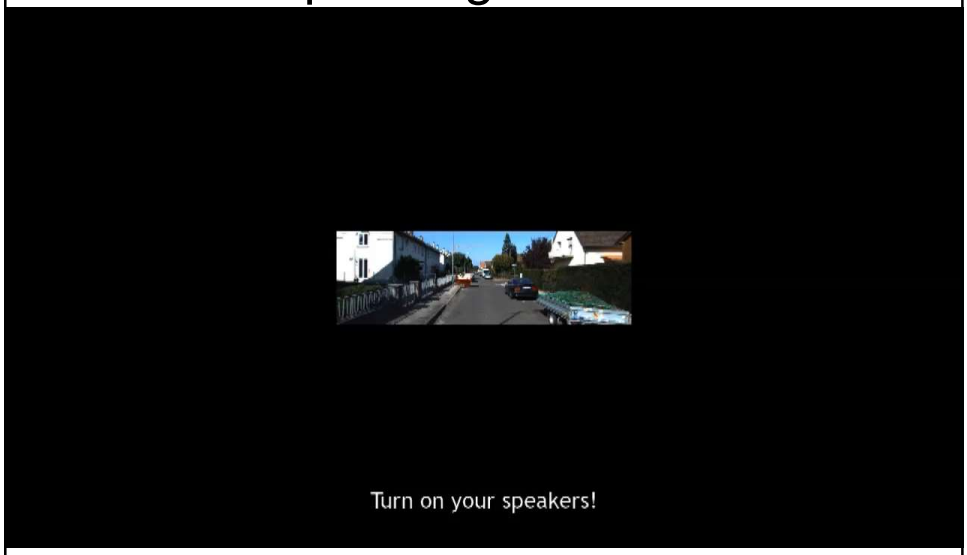


Expanding horizons: captioning



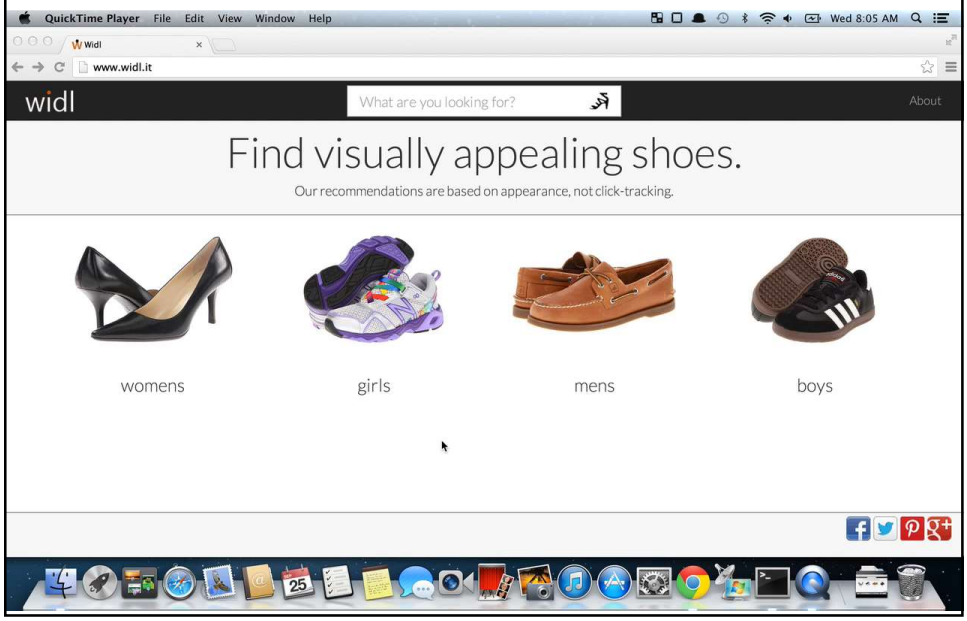
https://pdollar.wordpress.com/2015/01/21/image-captioning/

Expanding horizons:

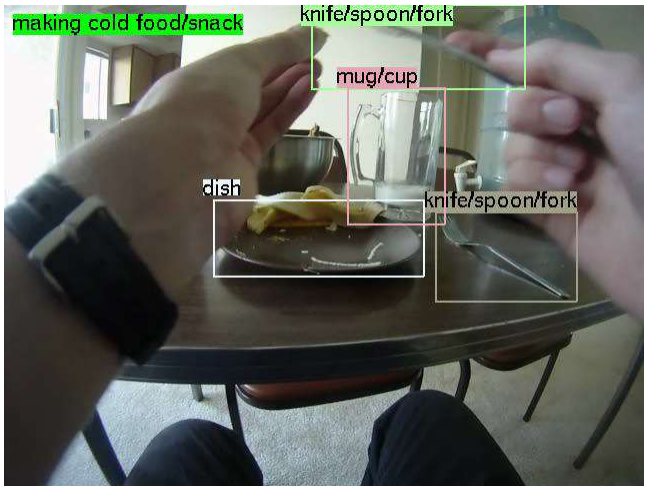


KITTI dataset – Andreas Geiger et al.

Expanding horizons: interactive visual search



Expanding horizons: first-person vision



Activities of Daily Living – Hamed Pirsiavash et al.

Brainstorm

Pick an application or task among any of those we've described so far.

1. What functionality should the system have?
2. Intuitively, what are the technical sub-problems that must be solved?

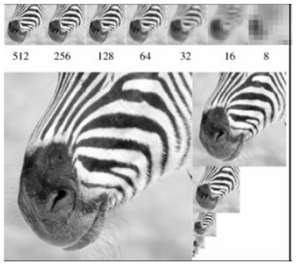
Goals of this course

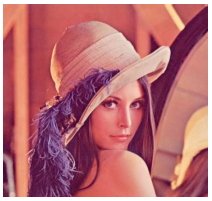
- Upper division honors undergrad course
- Introduction to primary topics
  - Fundamentals of computer vision – image processing, grouping, multiple views
  - Recognition - emphasis on *learning* (~last third of the course)
- Hands-on experience with algorithms
- Views of vision as a research area

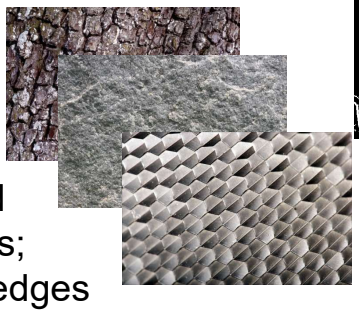
Topics overview

- Features & filters
- Grouping & fitting
- Multiple views
- Recognition

Features and filters

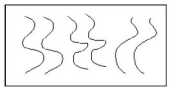




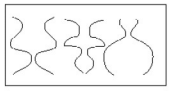


Transforming and describing images; textures, colors, edges

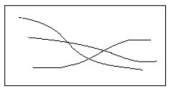
Grouping & fitting



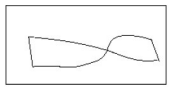
Parallelism



Symmetry



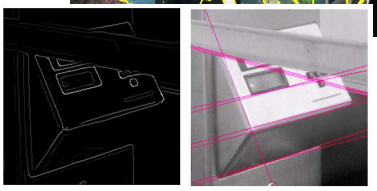
Continuity

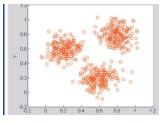


Closure

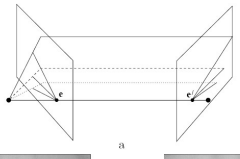
Clustering, segmentation, fitting; what parts belong together?

[fig from Shi et al]



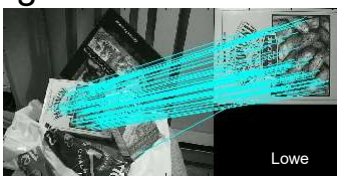


Multiple views

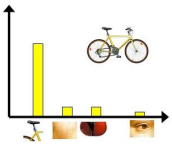


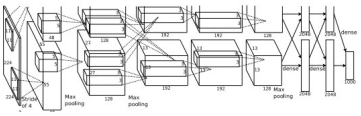
Hartley and Zisserman

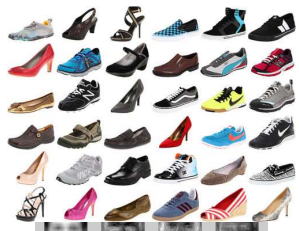
Matching, invariant features, stereo vision, instance recognition

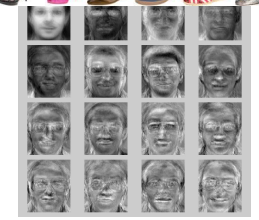


Recognition and learning

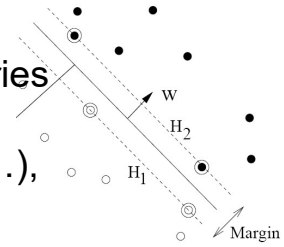






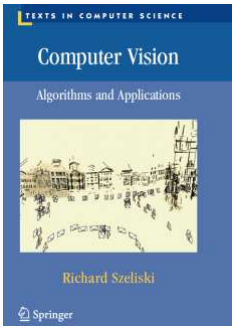


Recognizing categories (objects, scenes, activities, attributes…), learning techniques



Textbooks

- Recommended book:
  - Computer Vision: Algorithms and Applications
  - By Rick Szeliski
  - http://szeliski.org/Book/



Requirements / Grading

- Problem sets (50%)
- Midterm exam (15%)
- Final exam (25%)
- Class participation, including attendance (10%)
- Check grades on Canvas

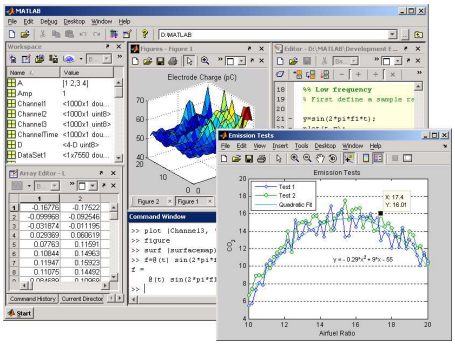
  – A quote from a prior student evaluation:

  "To be honest, I think without going to class, the course would be very hard. "

Assignments

- Majority - Programming problem
  - Implementation
  - Explanation, results
- Code in Matlab – available on CS Unix machines (see course page)
- Optional Latex templates
- Most of these assignments take significant time to do. We recommend starting early.

Matlab



- Built-in toolboxes for low-level image processing, visualization
- Compact programs
- Intuitive interactive debugging
- Widely used in engineering

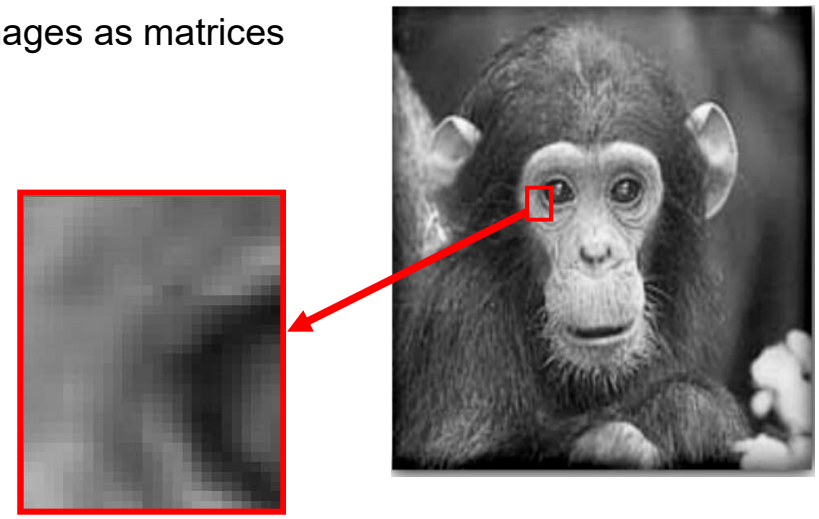
Assignment 0

- A0: Matlab warmup + basic image manipulation
- Out today, due Fri Jan 27

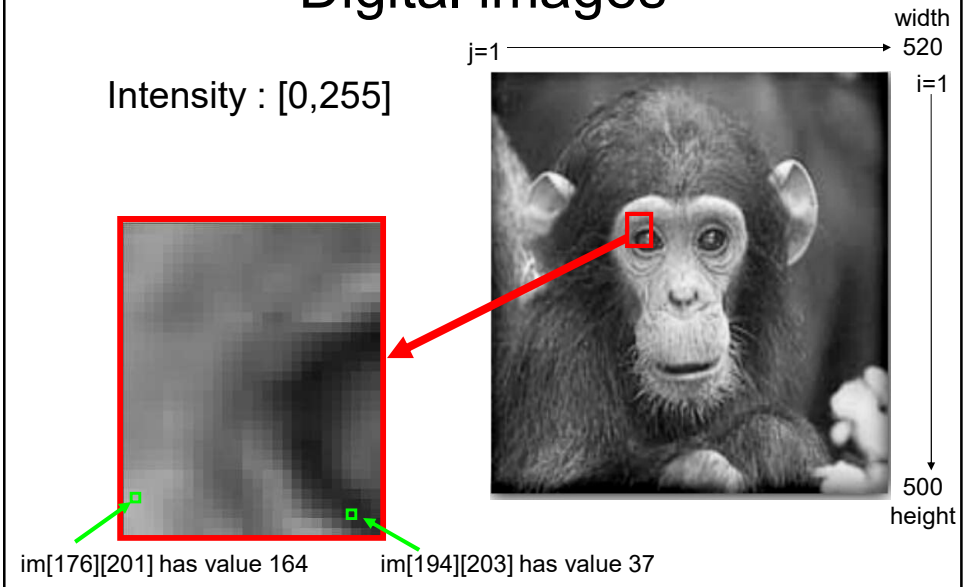
- Verify CS account and Matlab access
- Look at the tutorial online

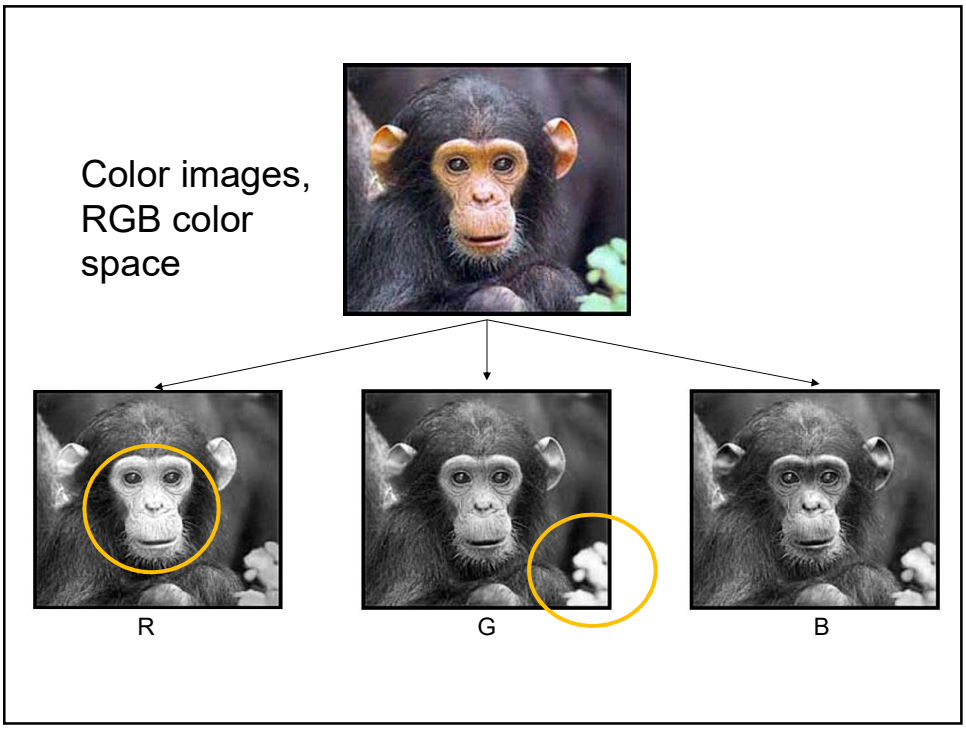
Digital images

Images as matrices

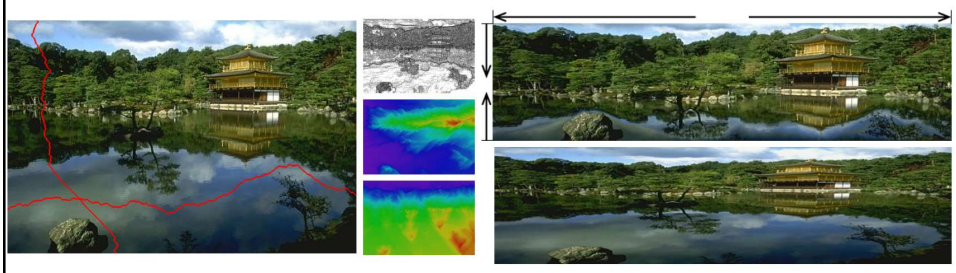


Digital images





Preview of assignments



Seam carving

Preview of assignments

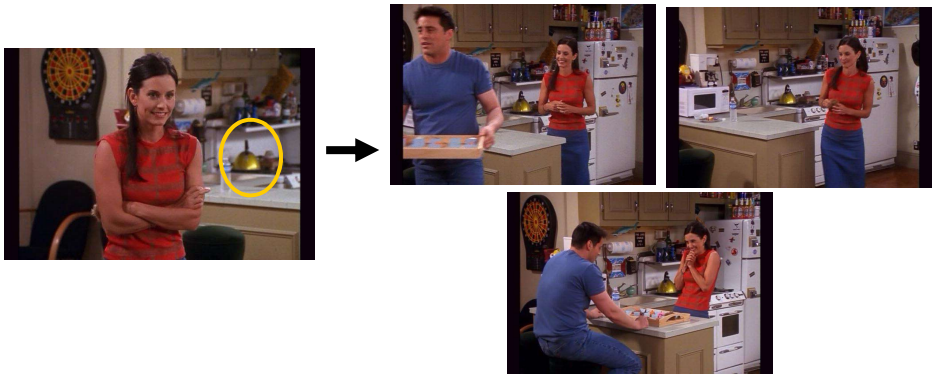
Grouping for segmentation

Preview of assignments

Image mosaics / stitching

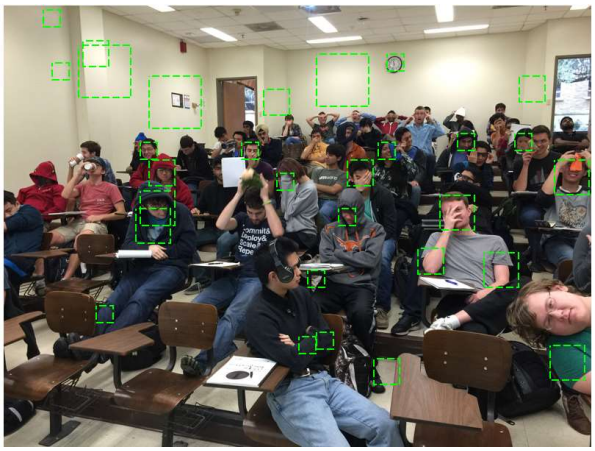
Image from Fei-Fei Li

Preview of assignments



Matching and recognition

Preview of assignments



Object detection

Collaboration policy

All responses and code must be written individually unless otherwise specified.

Students submitting answers or code found to be identical or substantially similar (due to inappropriate collaboration) risk failing the course.

Assignment deadlines

- Due about every two weeks
  - tentative deadlines posted online but could slightly shift depending on lecture pace
- Assignments in by 11:59 PM on due date
  - Submit on Canvas, following submission instructions given in assignment.
  - Deadlines are firm. We'll use timestamp.
- Use Piazza, office hours for questions

Miscellaneous

- Slides, announcements via class website
- No laptops, phones, tablets, etc. open in class please.

Coming up

- Now: check out Matlab tutorial online
- A0 due Fri Jan 27

- Textbook reading posted for next week