Computer Vision

Jan 19, 2011

Introductions

- **Instructor**: Prof. Kristen Grauman
  grauman@cs.utexas.edu
- **TA**: Shalini Sahoo
  shalini@cs.utexas.edu

Today

- Course overview
- Requirements, logistics

What is computer vision?

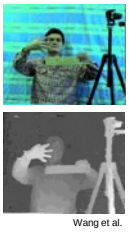
Done?

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.

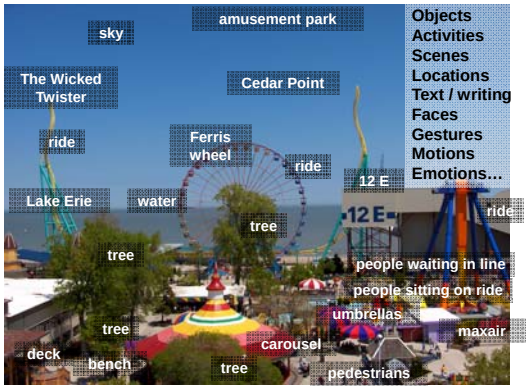
Tracking

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)*

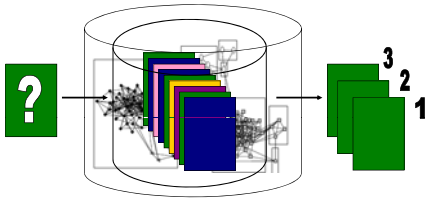
2. Vision for perception, interpretation



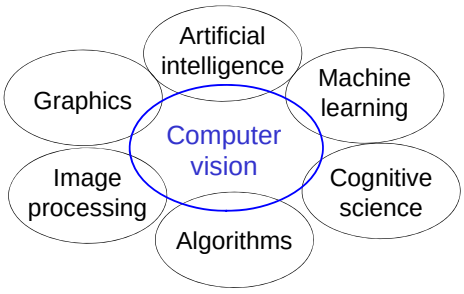
Computer Vision

- 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*)

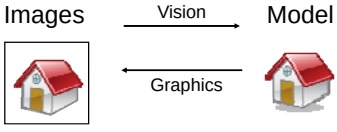
3. Visual search, organization



Related disciplines

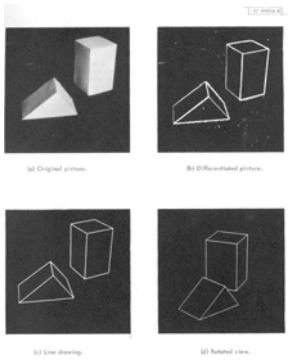


Vision and graphics



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 2011

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

MSR Lincoln

kooaba

Video-based interfaces

Human joystick, NewsBreaker Live

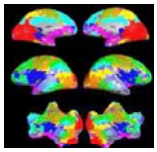
Assistive technology systems
Camera Mouse, Boston College

Microsoft Kinect

What else?

Vision for medical & neuroimages

Image guided surgery
MIT AI Vision Group



fMRI data
Golland et al.

Special visual effects

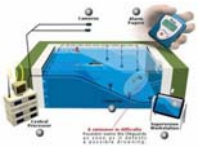
The Matrix

What Dreams May Come

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

Safety & security

Navigation,
driver safety

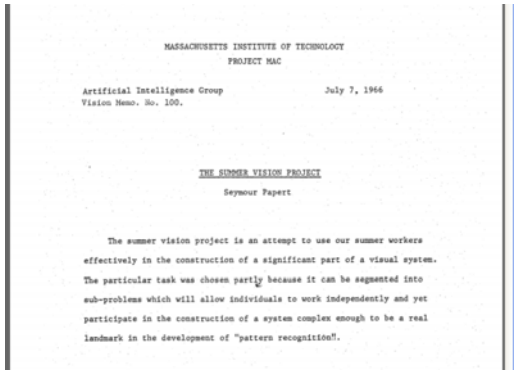


Monitoring pool
(Poseidon)

Pedestrian detection
MERL, Viola et al.

Surveillance

Obstacles?



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

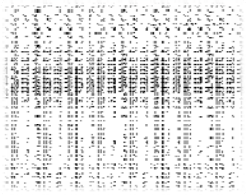
July 7, 1966

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

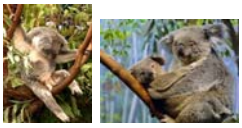


Why is vision difficult?

- 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

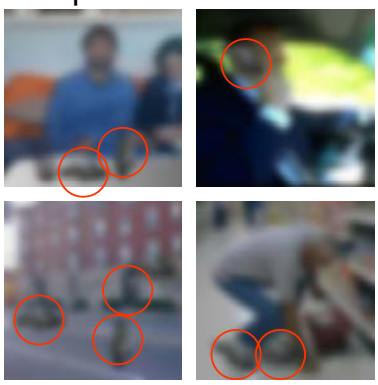
Intra-class appearance

Viewpoint

Challenges: intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: importance of context

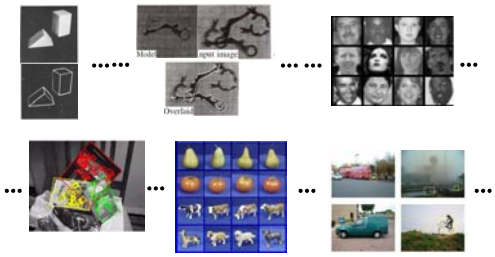


slide credit: Fei-Fei, Fergus & Torralba

Challenges: complexity

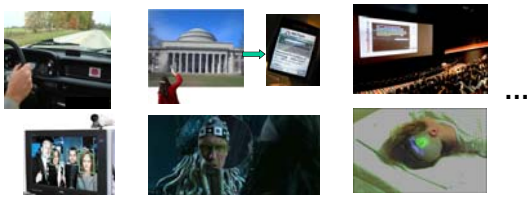
- Thousands to millions of pixels in an image
- 3,000-30,000 human recognizable object categories
- 30+ degrees of freedom in the pose of articulated objects (humans)
- Billions of images indexed by Google Image Search
- 18 billion+ prints produced from digital camera images in 2004
- 295.5 million camera phones sold in 2005
- About half of the cerebral cortex in primates is devoted to processing visual information [Felleman and van Essen 1991]

- Ok, vision is very challenging…
- Yet also active research area with exciting progress!



Brainstorm

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



...

Goals of this course

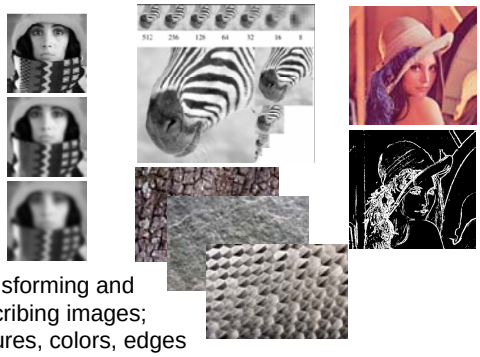
- Upper division undergrad course
- Introduction to primary topics
- Hands-on experience with algorithms
- Views of vision as a research area

Topics overview

- Features & filters
- Grouping & fitting
- Multiple views and motion
- Recognition
- Video processing

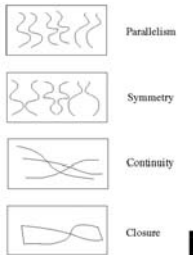
Focus is on algorithms, rather than specific systems.

Features and filters



Transforming and describing images; textures, colors, edges

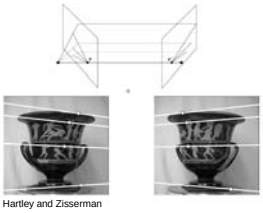
Grouping & fitting



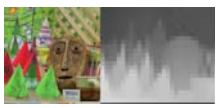
[fig from Shi et al]

Clustering, segmentation, fitting; what parts belong together?

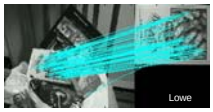
Multiple views and motion



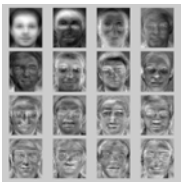
Hartley and Zisserman

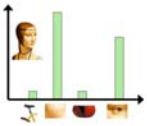


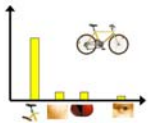
Multi-view geometry, matching, invariant features, stereo vision



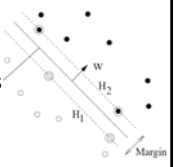
Recognition and learning







Recognizing objects and categories, learning techniques



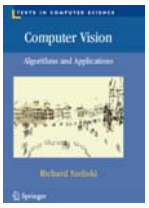
Video processing

Tracking objects, video analysis, low level motion, optical flow

Tomas Izo

Textbooks

- Recommended book:
  - Computer Vision: Algorithms and Applications
  - By Rick Szeliski
  - http://szeliski.org/Book
- And others on reserve at PCL



Requirements / Grading

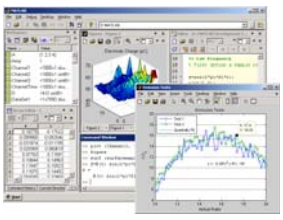
- Problem sets (50%)
- Midterm exam (20%)
- Final exam (20%)
- Class participation, including attendance (10%)
  - A quote from a student evaluation:

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

Problem sets

- Some short answer concept questions
- Programming problem
  - Implementation
  - Explanation, results
- Code in Matlab – available on CS Unix machines (see course page)
- These assignments are substantial.
- They will take significant time to do.
- Start early.

Matlab



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

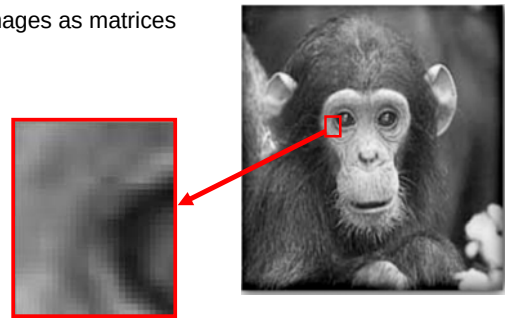
Pset 0

- Pset 0: Matlab warmup + basic image manipulation
- Out Fri Jan 21, Due Fri Jan 28

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

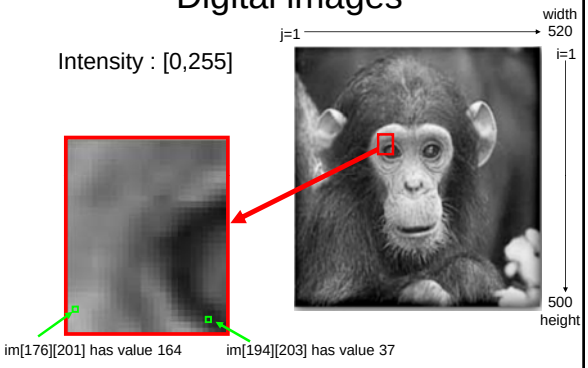
Digital images

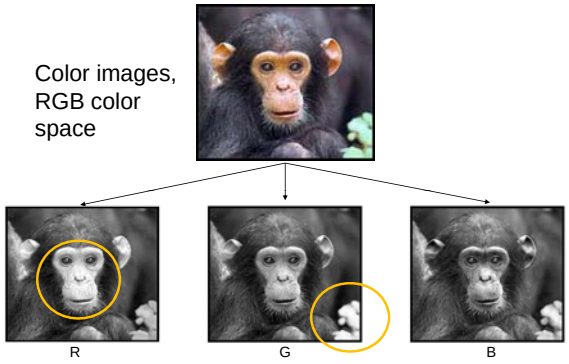
Images as matrices



Digital images

Intensity : [0,255]





Preview of some problem sets

Grouping

Preview of some problem sets

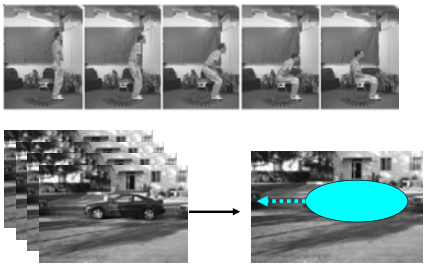
Image mosaics / stitching

Image from Fei-Fei Li

Preview of some problem sets

Object search and recognition

Preview of some problem sets



Tracking, activity recognition

Collaboration policy

All responses and code must be written individually.

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

Assignment deadlines

- Assignments in by11:59 PM on due date
  - Follow submission instructions given in assignment regarding hardcopy/electronic.
  - Deadlines are firm. We'll use turnin timestamp.
- 3 free late days, total for the term.
- Use as you want, but note that first two assignments lighter than rest.
- If your program doesn't work, clean up the code, comment it well, explain what you have, and still submit.

Miscellaneous

- Check class website regularly
- We'll use Blackboard to send email
- No laptops, phones, etc. open in class please.
- Use our office hours!

Coming up

- Now: check out Matlab tutorial online
- Friday 21st: Pset 0 out
- Monday 24th : first lecture on linear filters
- Friday 28th : Pset 0 due