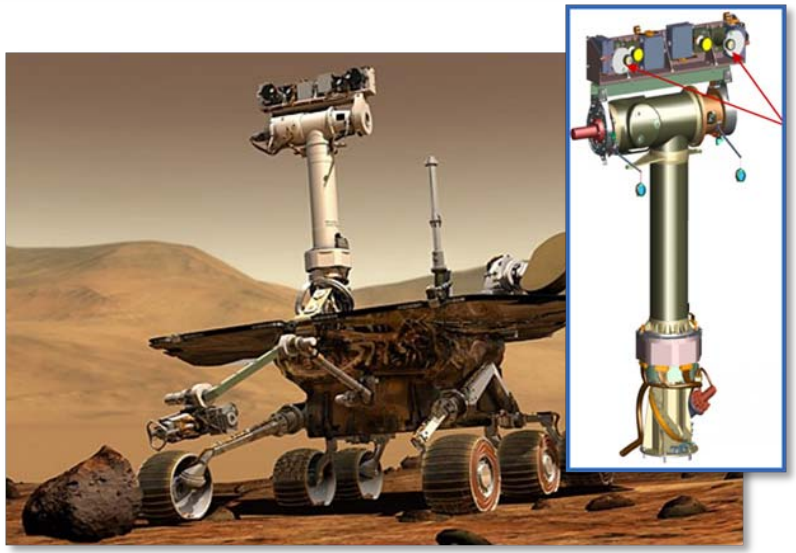
CS6670: Computer Vision

Noah Snavely



Instructor

- Noah Snavely (snavely@cs.cornell.edu)

- Office hours:

  Wednesdays 10:30 – noon, or by appointment

  (plus this Friday at 3-4pm)

- Research interests:
  - Computer vision and graphics
  - 3D reconstruction and visualization of Internet photo collections

Other details

- Textbook:

  R. Szeliski, *Computer Vision: Algorithms and Applications*

  online at:

  http://research.microsoft.com/en-us/um/people/szeliski/Book/

  (please check Web site periodically for updated drafts)

- Course webpage:
  http://www.cs.cornell.edu/courses/cs6670/2009fa/
- Mailing list: cs6670-l@lists.cs.cornell.edu

Today

1. Introduction to computer vision
2. Course overview
3. Basic image processing

- Note: Class will **not** meet next week
  - Next meeting: Tuesday, September 8

Today

- Readings
  - Szeliski, CV: A&A, Ch 1.0 (Introduction)

- Handouts
  - signup sheet
  - intro slides
  - image filtering slides

Announcement

- Today: Conway-Walker Distinguished Lecture

  Andy Wilson, Microsoft Research on
  Surface Computing

  Today, 4:15pm

  (right after class)

  B17 Upson Hall

Every image tells a story

- Goal of computer vision: perceive the story behind the picture
- Compute properties of the world
  - 3D shape
  - Names of people or objects
  - What happened?

The goal of computer vision

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

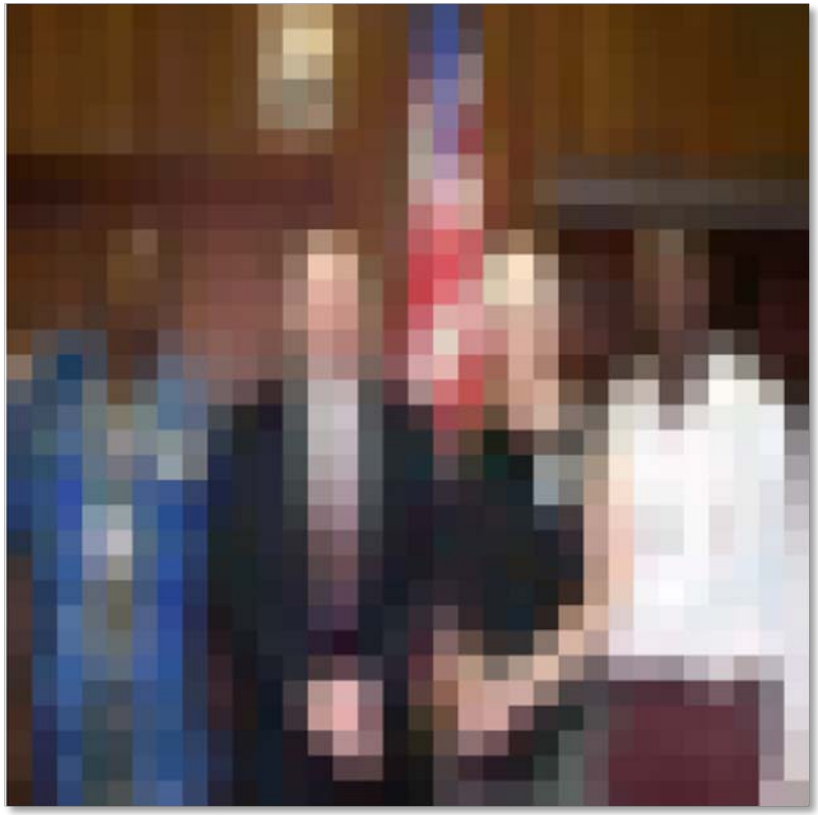
Can the computer match human perception?

- Yes and no (but mainly no, so far)
  - computers can be better at “easy” things
  - humans are much better at “hard” things

Human perception has its shortcomings

Sinha and Poggio, *Nature*, 1996

But humans can tell a lot about a scene from a little information…



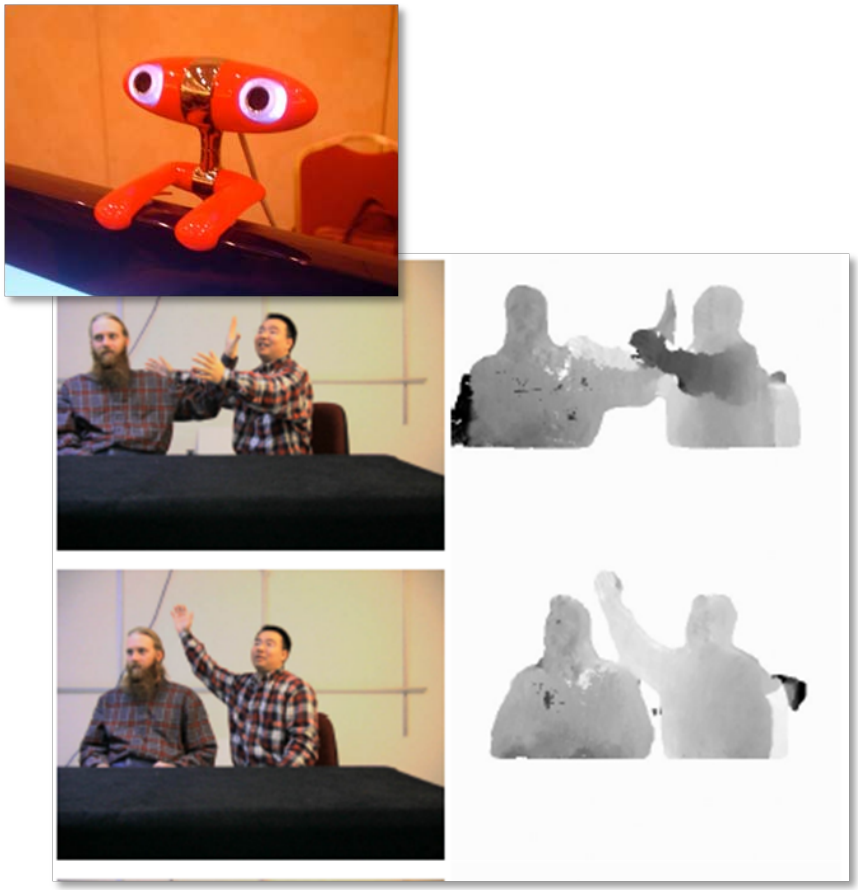
Source: “80 million tiny images” by Torralba, et al.

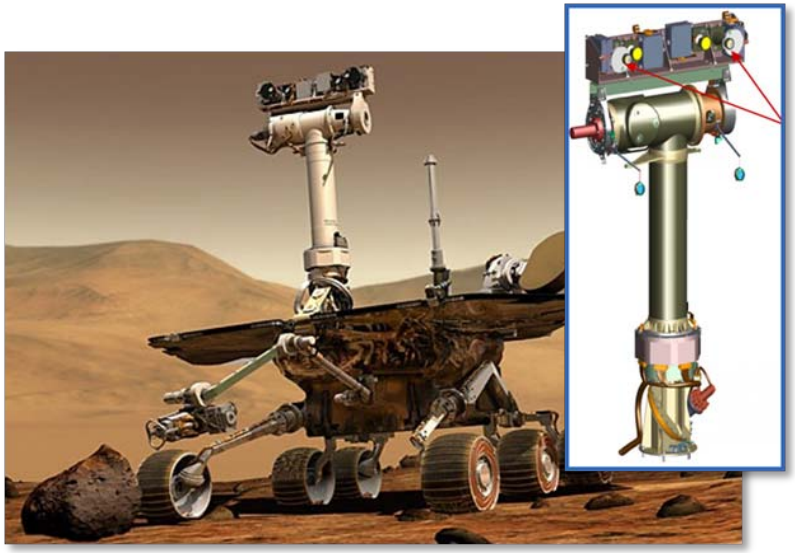
Source: Antonio Torralba

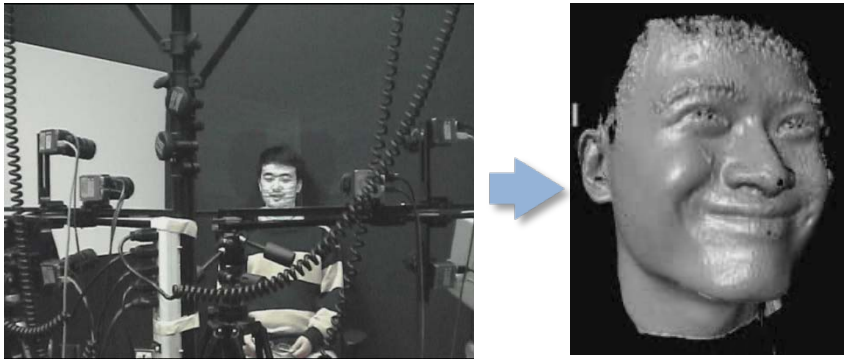
The goal of computer vision

The goal of computer vision

- Computing the 3D shape of the world

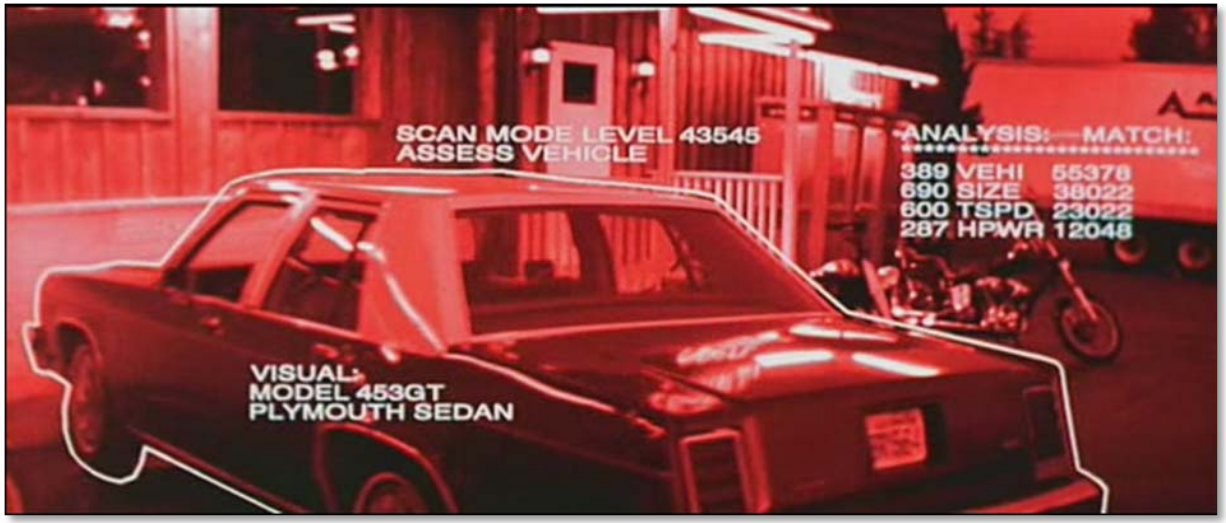






The goal of computer vision

- Recognizing objects and people



Vision as a source of semantic information

slide credit: Fei-Fei, Fergus & Torralba

Object categorization

The goal of computer vision

- “Enhancing” images

The goal of computer vision

- “Enhancing” images

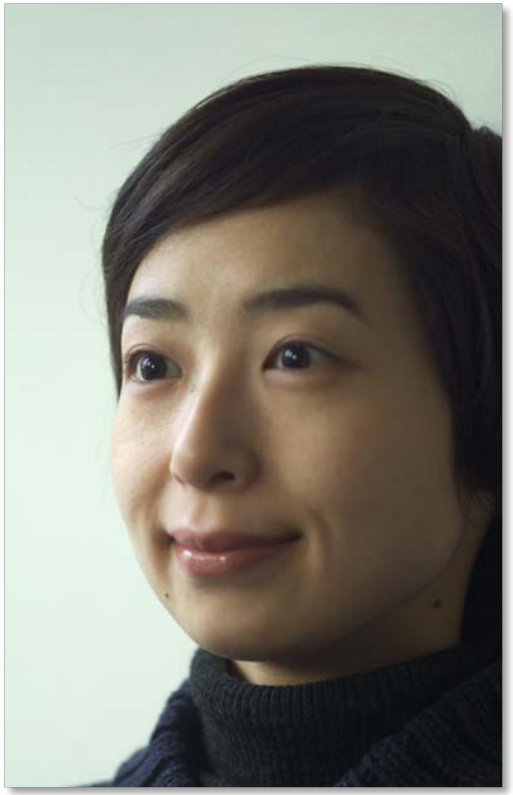
Super-resolution / denoising
(source: 2d3)

Texture synthesis / increased field of view (uncropping)
(image credit: Efros and Leung)

Inpainting / image completion
(image credit: Hays and Efros)

The goal of computer vision

- Forensics



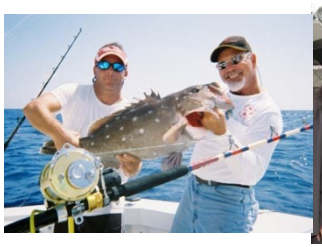
Source: Nayar and Nishino, “Eyes for Relighting”

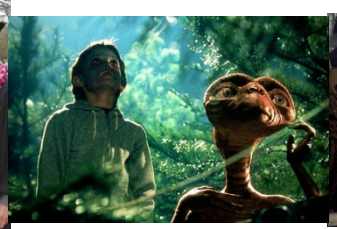
Source: Nayar and Nishino, “Eyes for Relighting”

Source: Nayar and Nishino, “Eyes for Relighting”

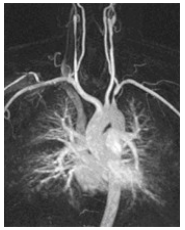
Why study computer vision?

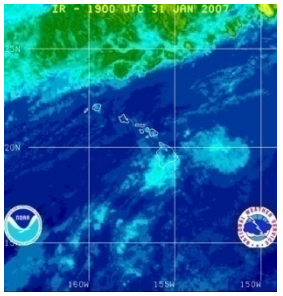
- Millions of images being captured all the time






Google Image Search™ Picasa™ flickr™ webshots™ picsearch™ YouTube Broadcast Yourself™


- Lots of useful applications
- The next slides show the current state of the art

Source: S. Lazebnik

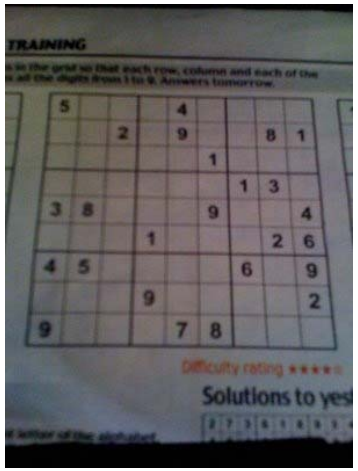
Optical character recognition (OCR)

- If you have a scanner, it probably came with OCR software

Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Automatic check processing



Sudoku grabber
http://sudokugrab.blogspot.com/

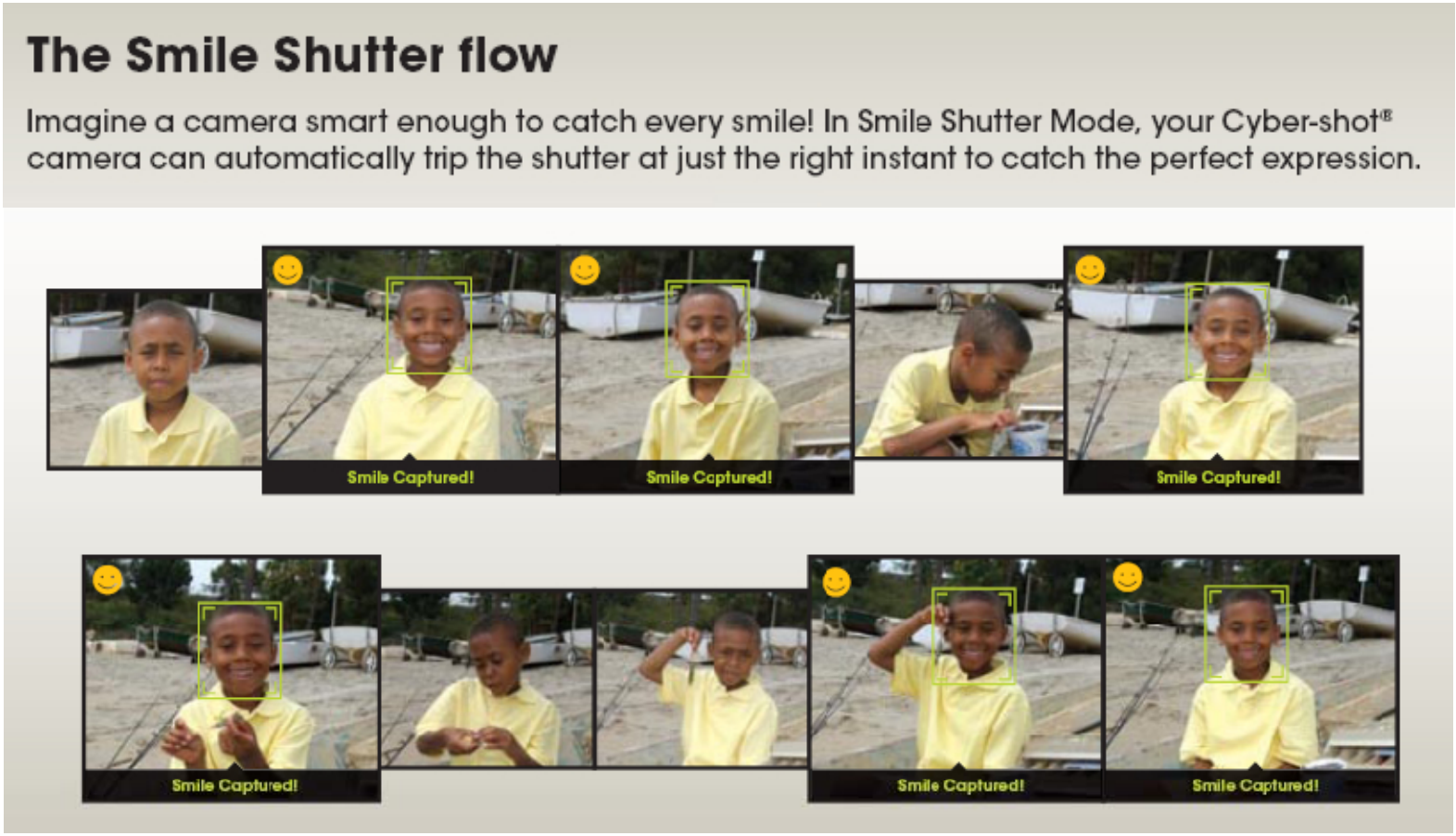
Source: S. Seitz

Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

Source: S. Seitz

Smile detection?



[Sony Cyber-shot® T70 Digital Still Camera]

Source: S. Seitz

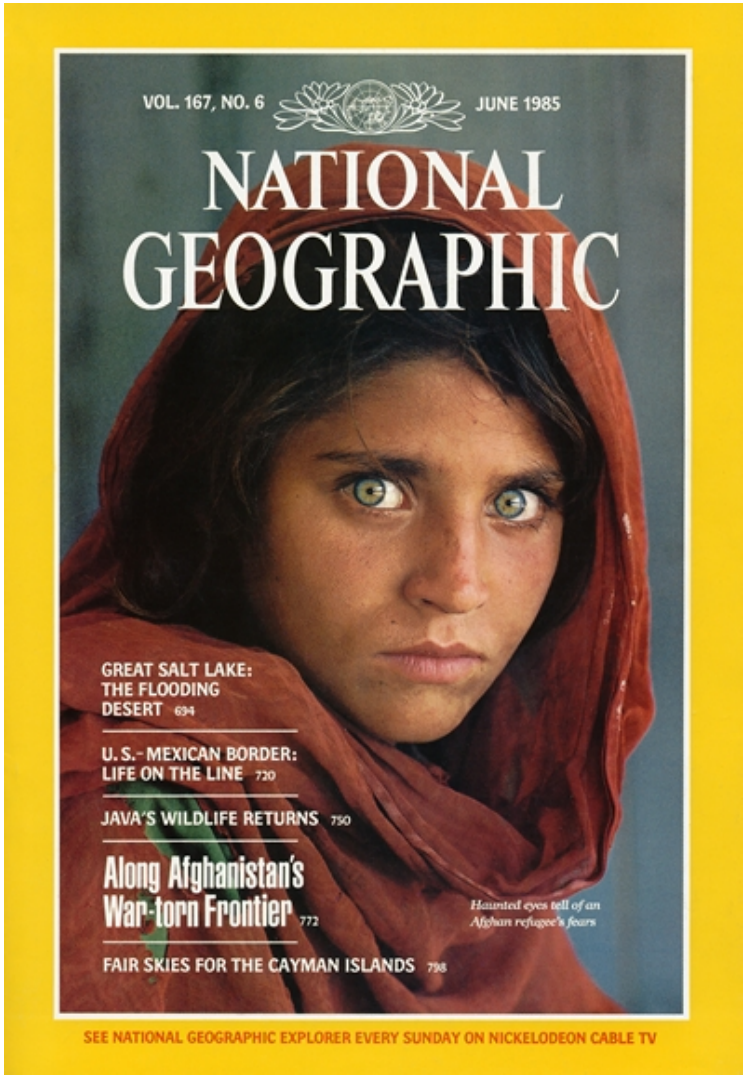
Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… “

Source: S. Seitz

Face recognition

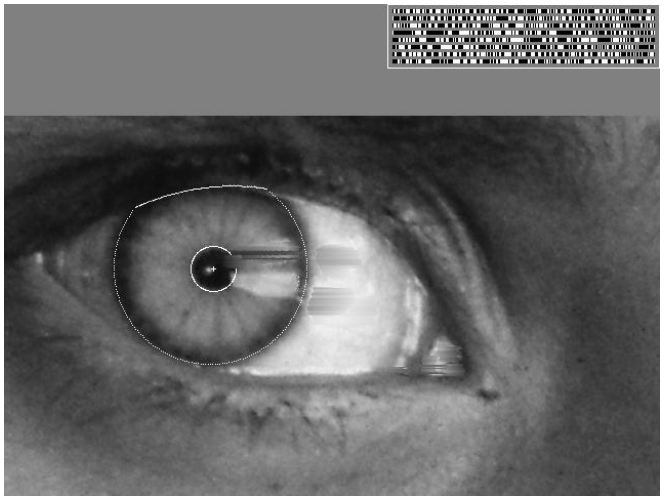


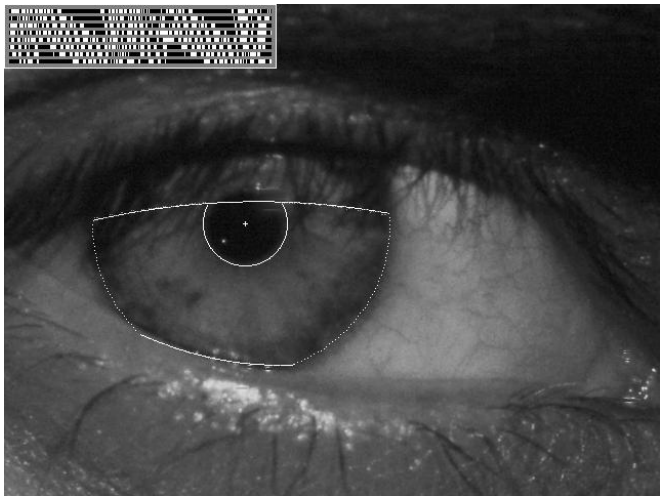
Who is she?

Source: S. Seitz

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story

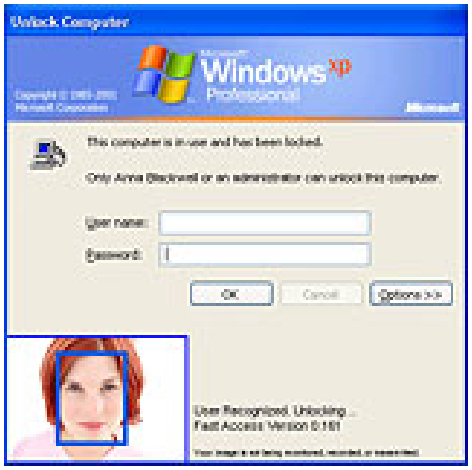




Source: S. Seitz

Login without a password…

Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Source: S. Seitz

Object recognition (in mobile phones)

- This is becoming real:
  - Lincoln Microsoft Research
  - Point & Find

Source: S. Seitz

iPhone Apps: kooaba (www.kooaba.com)

Source: S. Lazebnik

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Source: S. Seitz

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Source: S. Seitz

Special effects: camera tracking

Boujou, 2d3

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

Source: S. Seitz

Smart cars

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.

Sources: A. Shashua, S. Seitz

Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Sony EyeToy

Assistive technologies

Project Natal?

Source: S. Seitz

Vision in space

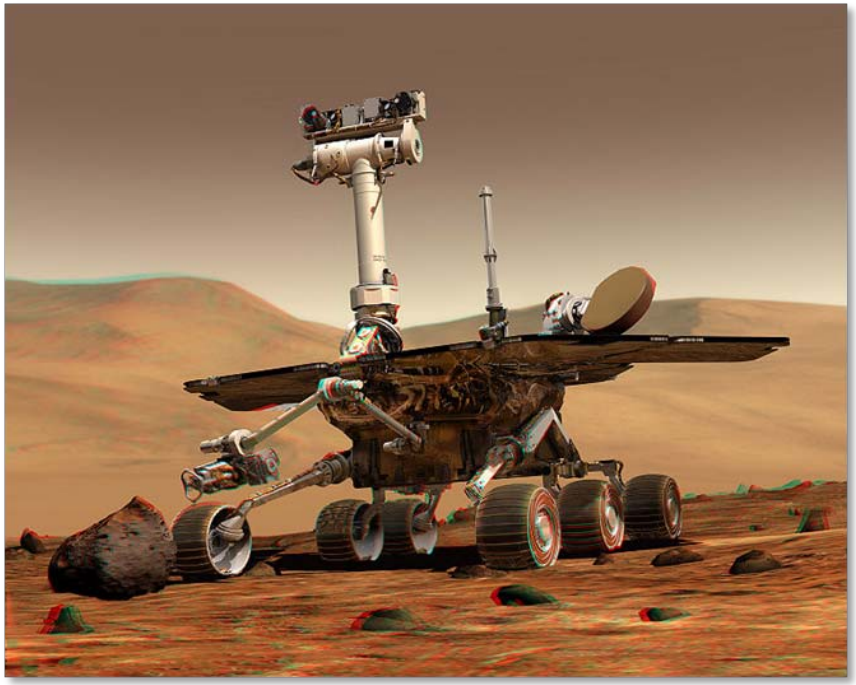
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Source: S. Seitz

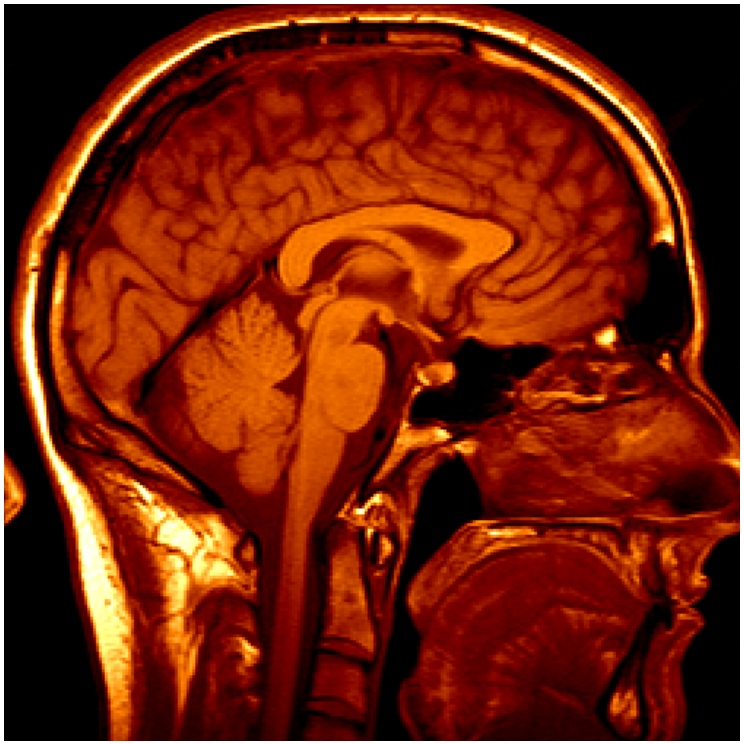
Robotics



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

Autonomous RC Car
http://www.cs.cornell.edu/~asaxena/rccar/

Medical imaging



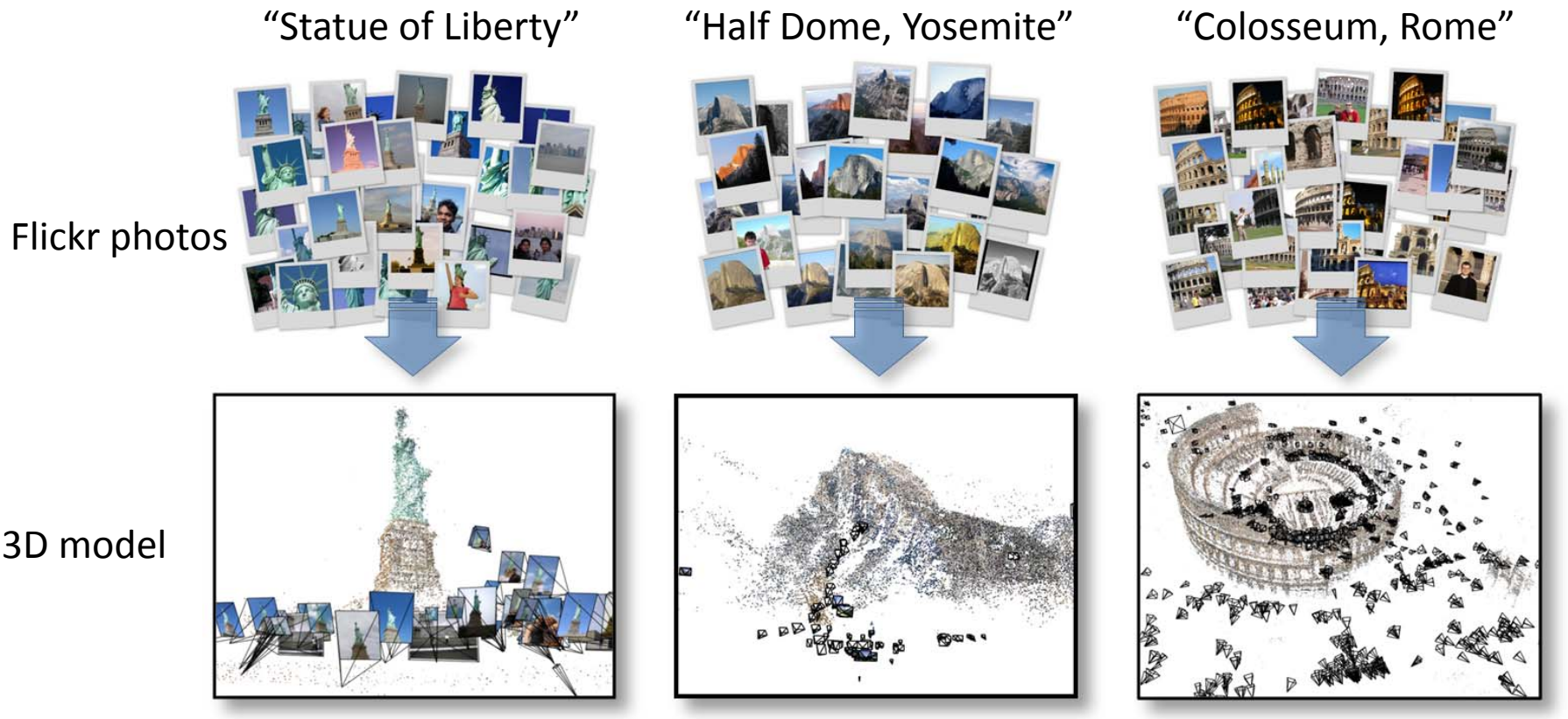
3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

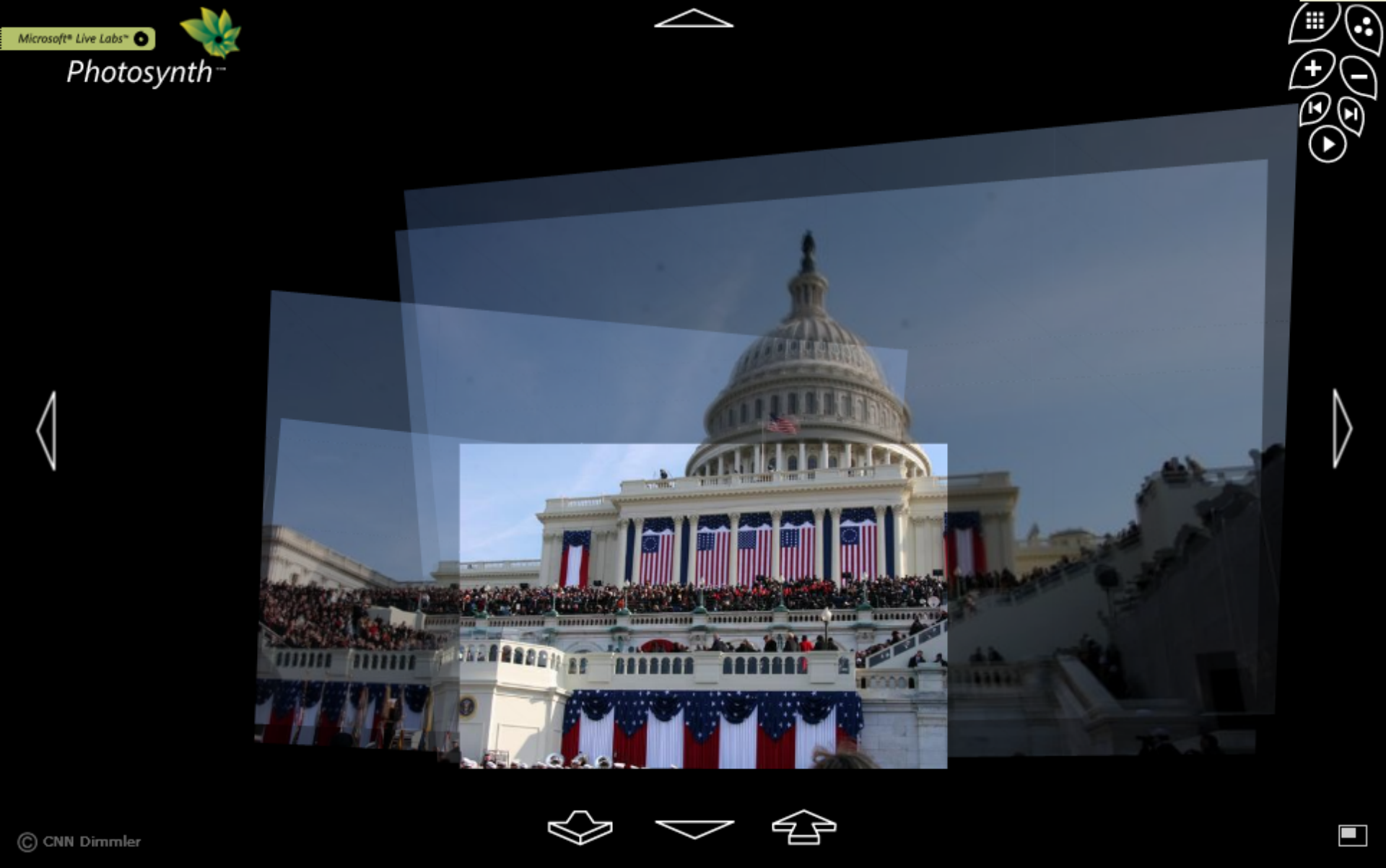
Source: S. Seitz

My own work

- Automatic 3D reconstruction from Internet photo collections



Photosynth



City-scale reconstruction

Reconstruction of Dubrovnik, Croatia, from ~40,000 images

Current state of the art

- You just saw examples of current systems.
  - Many of these are less than 5 years old

- This is a very active research area, and rapidly changing
  - Many new apps in the next 5 years

- To learn more about vision applications and companies
  - David Lowe maintains an excellent overview of vision companies
    - http://www.cs.ubc.ca/spider/lowe/vision.html

Why is computer vision difficult?

Viewpoint variation

Illumination

Scale

Why is computer vision difficult?

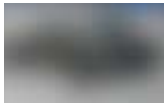
Intra-class variation

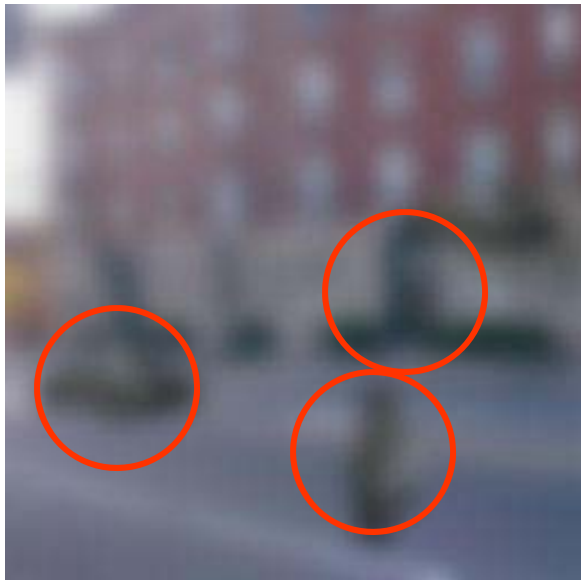
Motion (Source: S. Lazebnik)

Background clutter

Occlusion

Challenges: local ambiguity





slide credit: Fei-Fei, Fergus & Torralba

But there are lots of cues we can exploit…

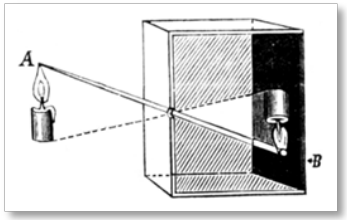
Source: S. Lazebnik

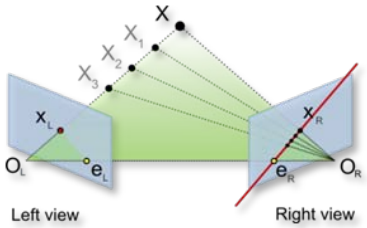
Bottom line

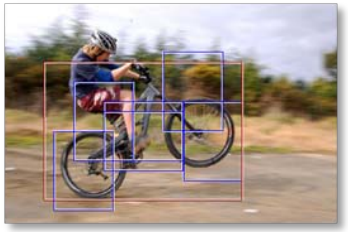
- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

  - We often need to use prior knowledge about the structure of the world

Course overview (tentative)



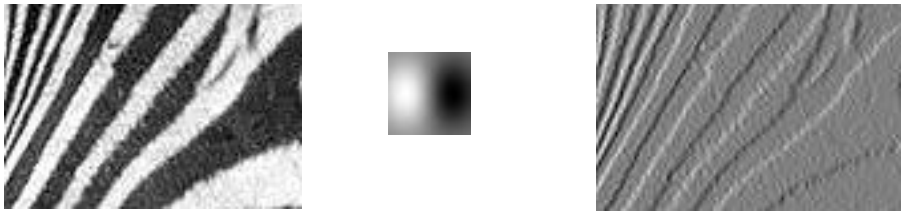




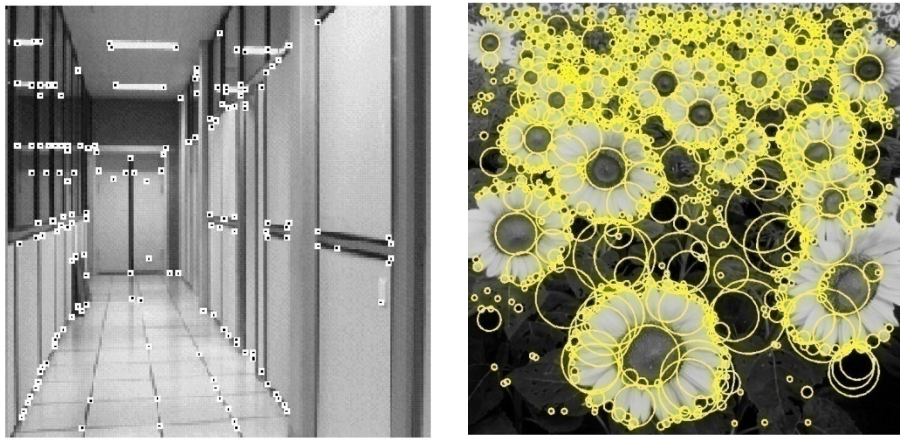
1. Low-level vision
   - image processing, edge detection, feature detection, cameras, image formation
2. Geometry and algorithms
   - projective geometry, stereo, structure from motion, Markov random fields
3. Recognition
   - face detection / recognition, category recognition, segmentation
4. Light, color, and reflectance
5. Advanced topics

1. Low-level vision

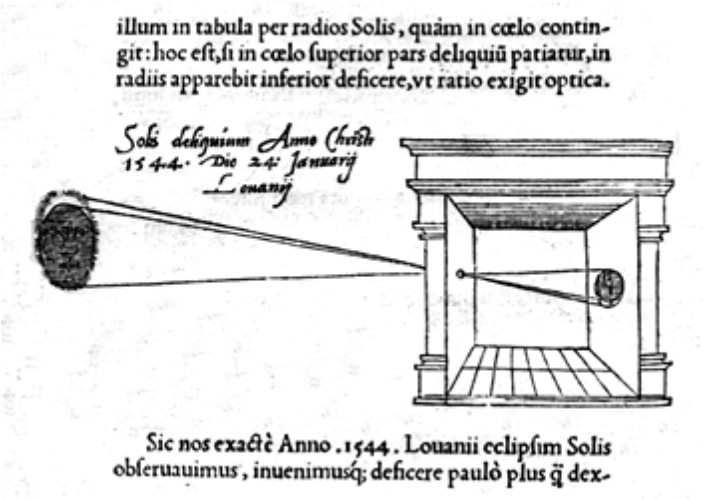
- Basic image processing and image formation

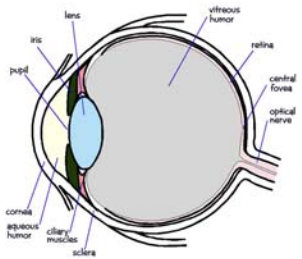


Filtering, edge detection



Feature extraction





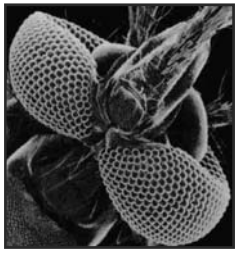
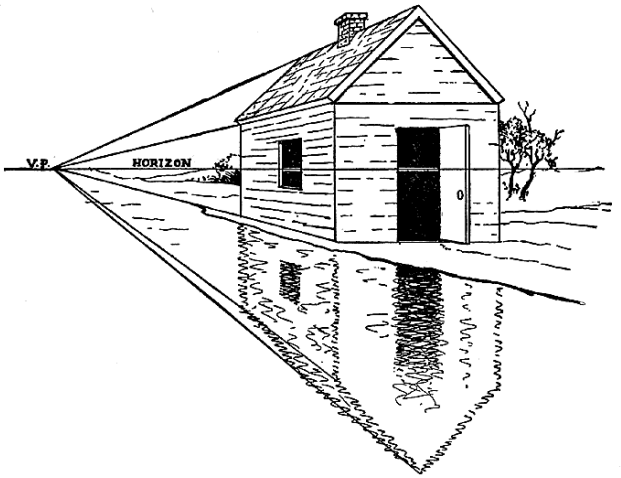


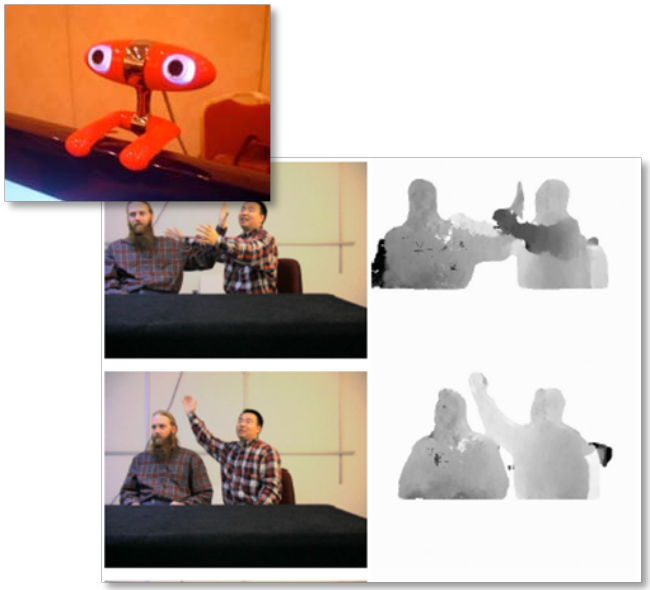
Image formation

Project 1: Feature detection and matching

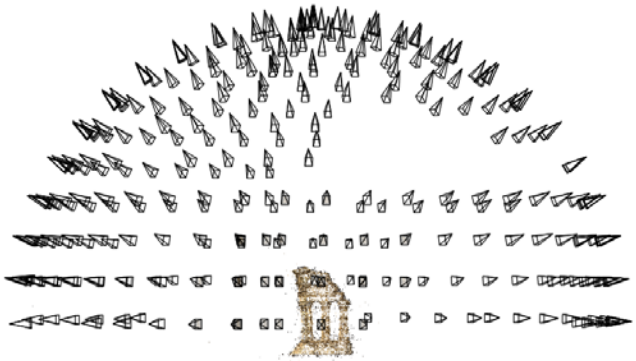
2. Geometry



Projective geometry



Stereo

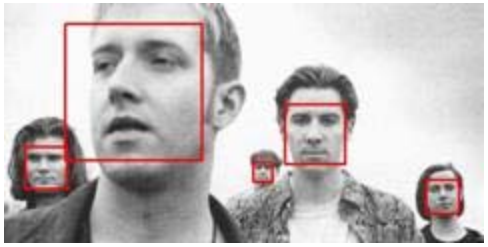


Multi-view stereo

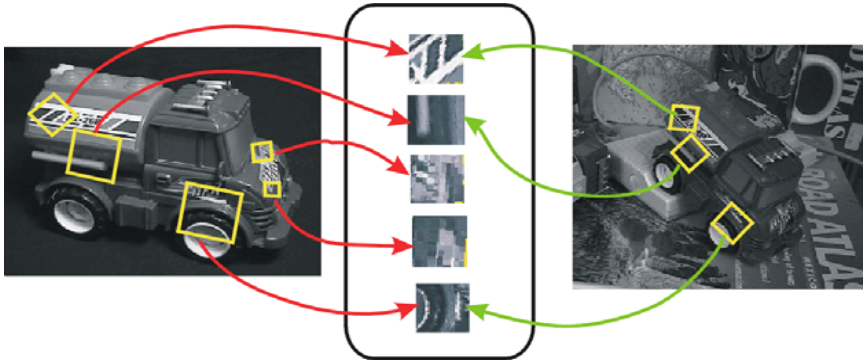
Structure from motion

Project 2: Creating panoramas

3. Recognition



Face detection and recognition

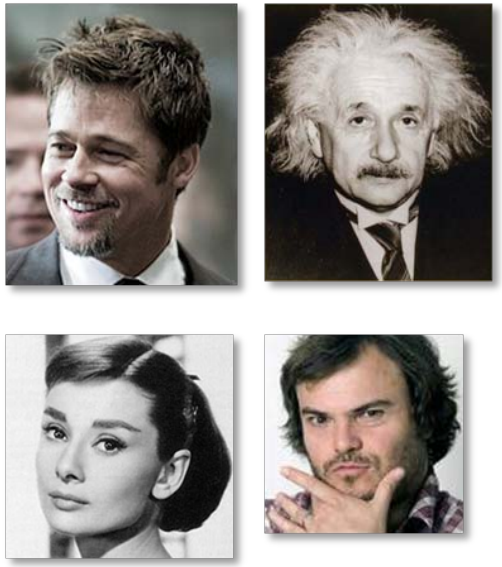


Single instance recognition

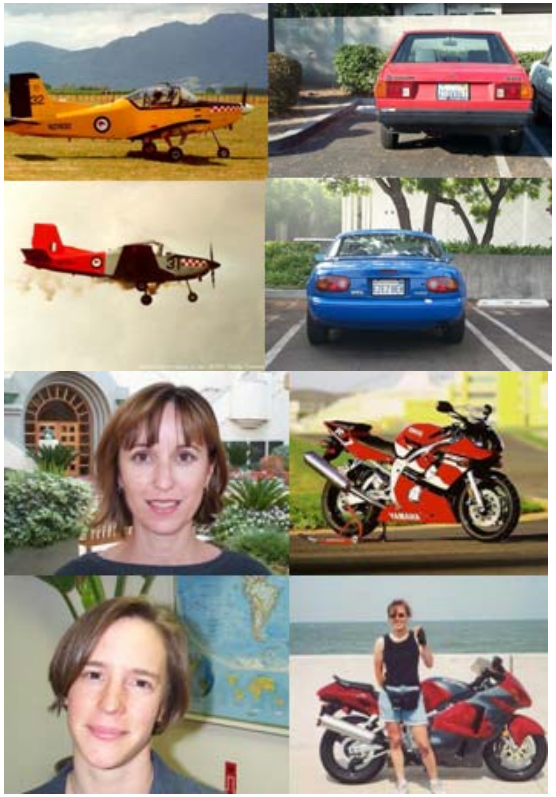
Category recognition

Project 3: Recognition challenge (TBA)

Location recognition

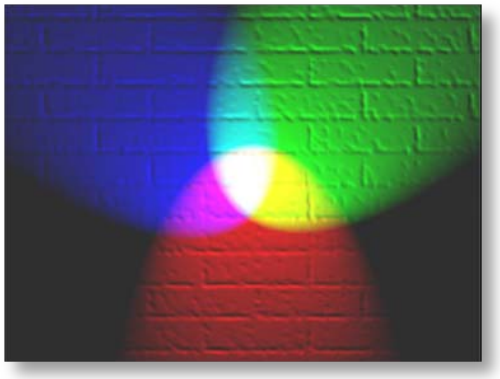


Face recognition

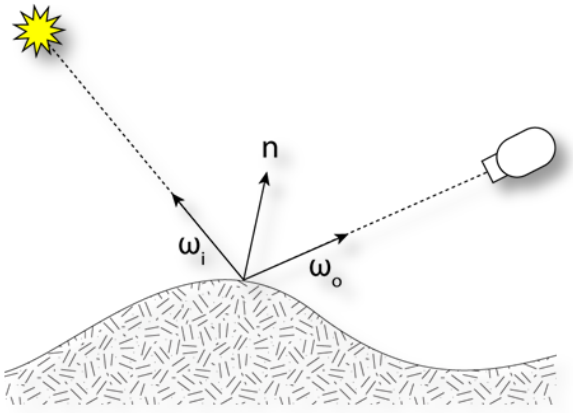


Object category recognition

4. Light, color, and reflectance

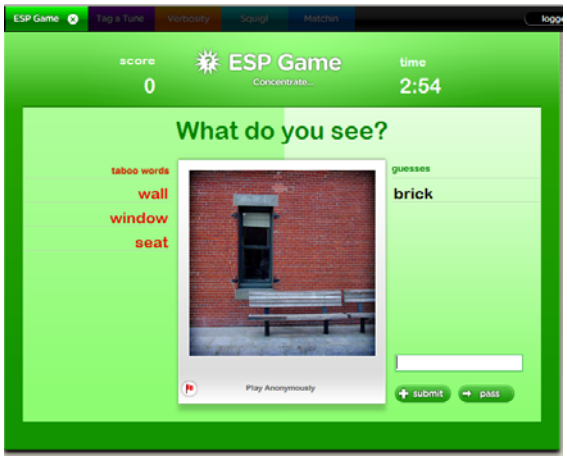


Light & Color

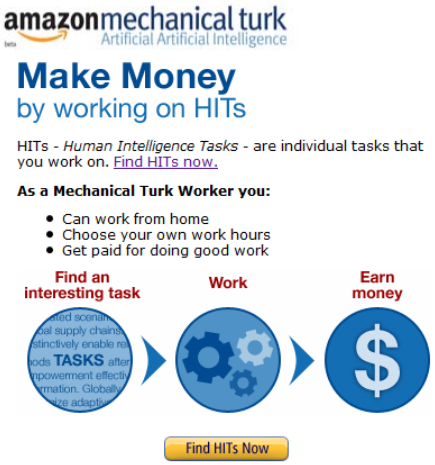


Reflectance

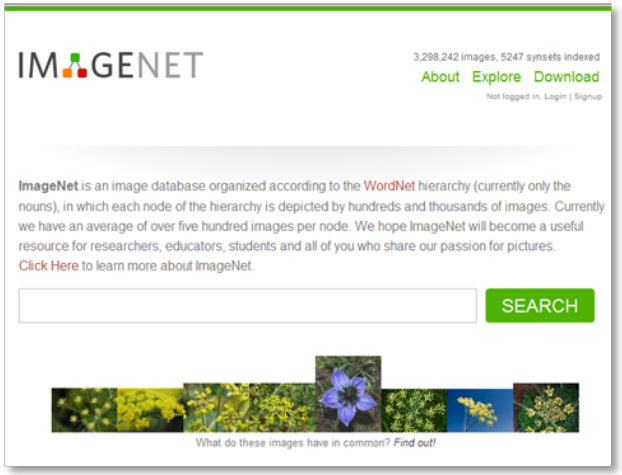
5. Advanced topics: Internet Vision



Human-aided computer vision



Turning the camera around



Internet datasets

Final project

- Either:
  1. Implement a recent computer vision paper (solo)

– or –

  2. Explore a new research problem (in groups of one or more)
     - Example research projects TBA

Course requirements

- Prerequisites—*these are essential*!
  - Data structures
  - A good working knowledge of C/C++ (or Matlab) programming
    - (or willingness/time to pick it up quickly!)
  - Linear algebra

- Course does ***not*** assume prior imaging experience
  - computer vision, image processing, graphics, etc.

Grading

- No exams; occasional quizzes (at the beginning of class)
- Quizzes: ~5-10%
- Programming projects: ~ 60%
- Final project: ~ 30%

3-minute break

- Next up: Images and image filtering