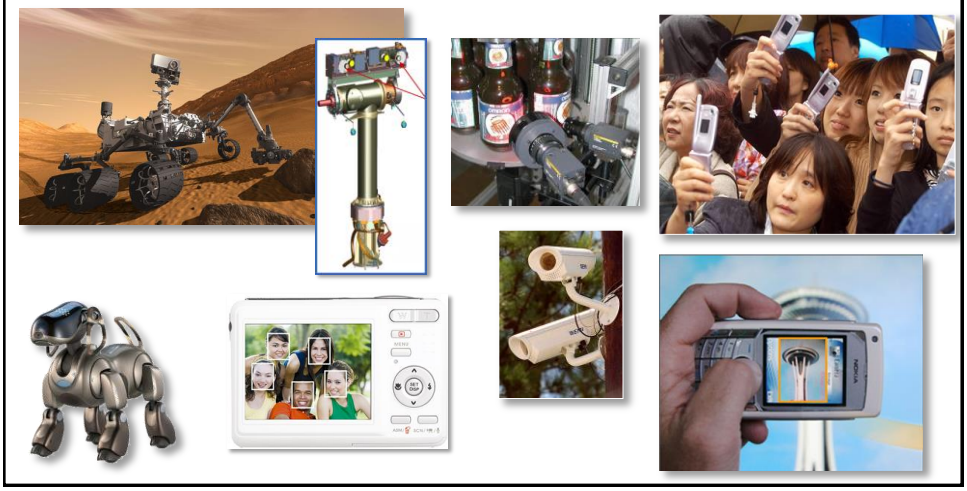
CS4670/5670: Intro to Computer Vision

Instructor: Noah Snavely



Instructor

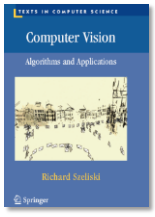
- Noah Snavely (snavely@cs.cornell.edu)
- Office hours:

  Wednesdays 1:30 – 3pm, or by appointment
- Research interests:
  - Computer vision and graphics
  - 3D reconstruction and visualization of Internet photo collections

Important personnel

- TAs:
  - Daniel Cabrini Hauagge
  - Scott Wehrwein
- Office hours TBA

Other administrative details



- Textbook:

  Rick Szeliski, *Computer Vision: Algorithms and Applications*

  online at: http://szeliski.org/Book/

- Course webpage (content coming soon):

  http://www.cs.cornell.edu/courses/cs4670/2013fa/

- Announcements/grades via Piazza/CMS

  https://piazza.com/class#fall2013/cs46705670

  https://cms.csuglab.cornell.edu/

Course requirements

- Prerequisites—*these are essential*!
  - Data structures
  - A good working knowledge of C/C++ programming
  - Linear algebra
  - Vector calculus
- Course does ***not*** assume prior imaging experience
  - computer vision, image processing, graphics, etc.

Today

1. What is computer vision?
2. Course overview

Today

- Readings
  - Szeliski, Chapter 1 (Introduction)

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

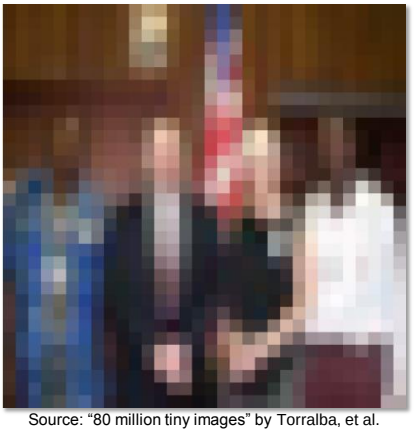
Can the computer match human perception?

- Yes and no (mainly no)
  - computers can be better at "easy" things
  - humans are much better at "hard" things
- But huge progress has been made
  - Especially in the last 10 years
  - What is considered "hard" keeps changing

Human perception has its shortcomings

Sinha and Poggio, *Nature*, 1996

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

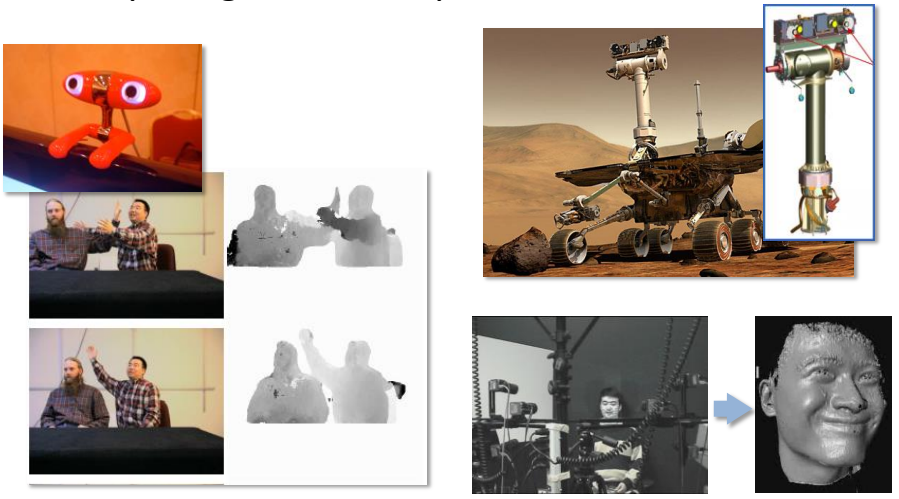


Source: "80 million tiny images" by Torralba, et al.

The goal of computer vision

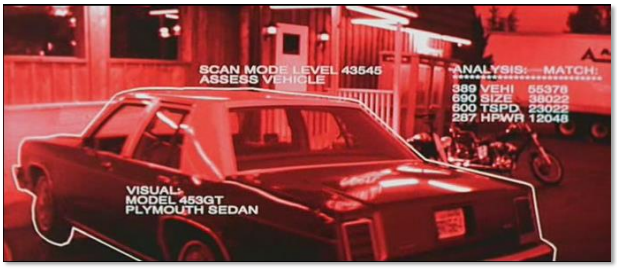
The goal of computer vision

- Computing the 3D shape of the world



The goal of computer vision

- Recognizing objects and people




中华人民
和国万岁
世界
民大团结万岁

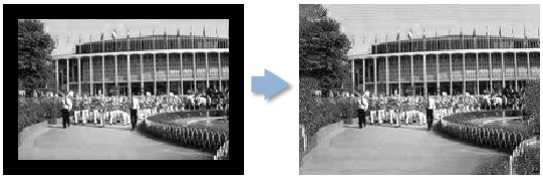
The goal of computer vision

- “Enhancing” images

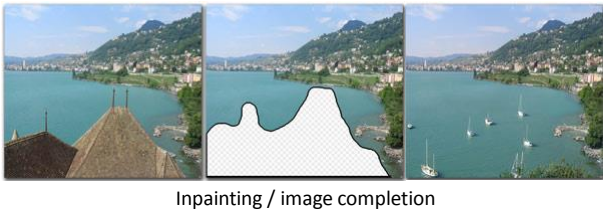
The goal of computer vision

- "Enhancing" images (c.f. Computational Photography)

Super-resolution / denoising
(source: 2d3)



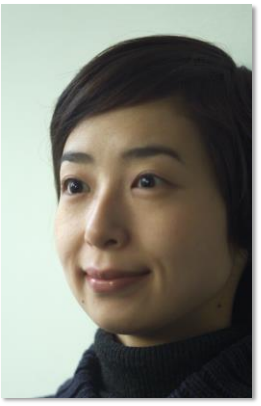
Texture synthesis / increased field of view (uncropping)
(image credit: Efros and Leung)



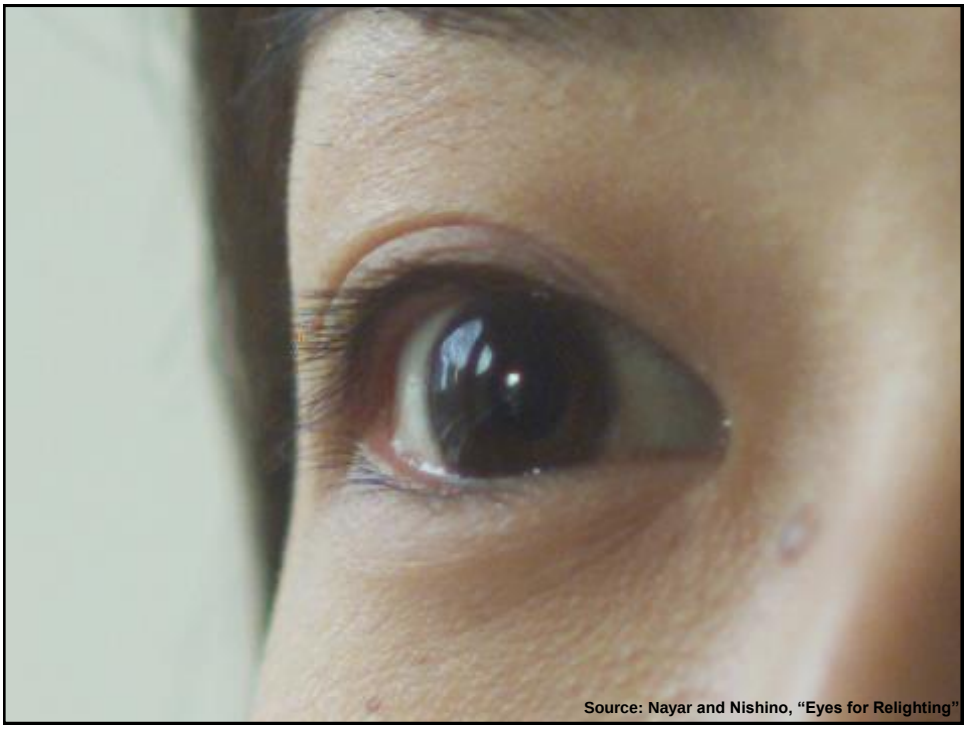
Inpainting / image completion
(image credit: Hays and Efros)

The goal of computer vision

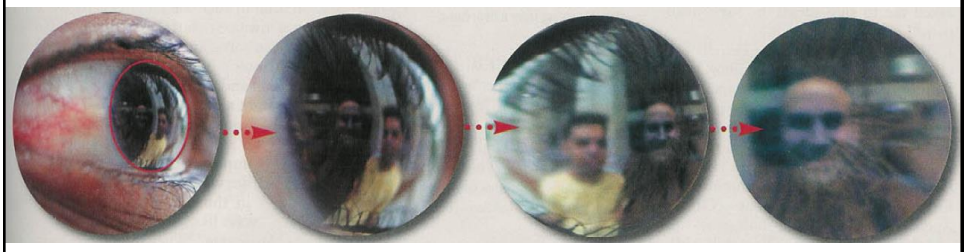
- Forensics



Source: Nayar and Nishino, “Eyes for Relighting”



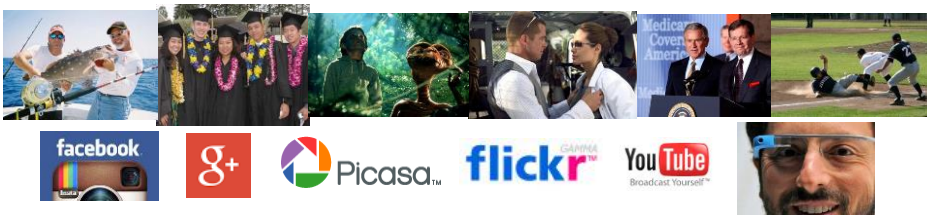
Source: Nayar and Nishino, “Eyes for Relighting”



Source: Nayar and Nishino, "Eyes for Relighting"

Why study computer vision?

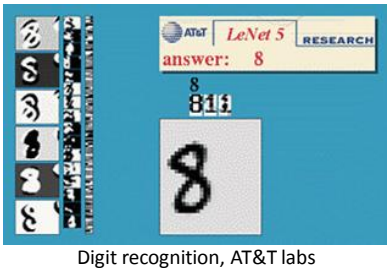
- Millions of images being captured all the time



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

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

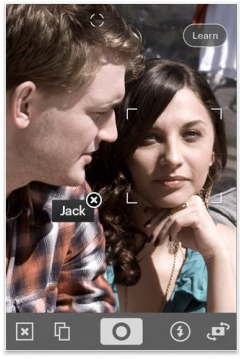
- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

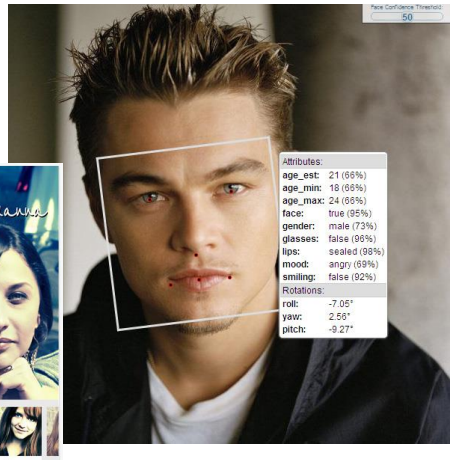
Source: S. Seitz

Face Recognition

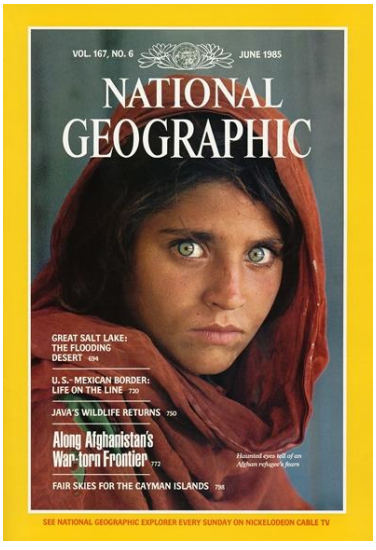
face.com

http://developers.face.com/tools/





Face recognition

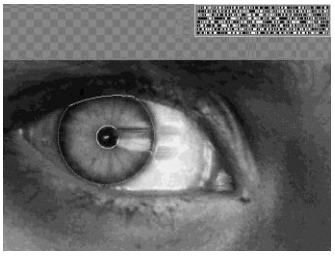


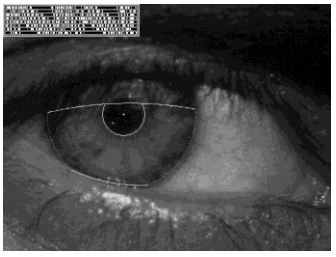
Who is she?

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 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

Object recognition (in mobile phones)

Source: S. Seitz

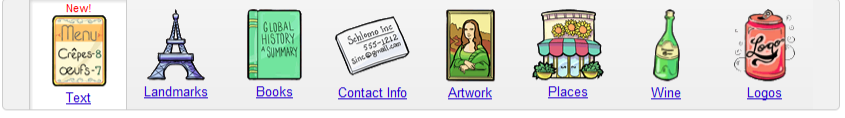
iPhone Apps: kooaba (www.kooaba.com)

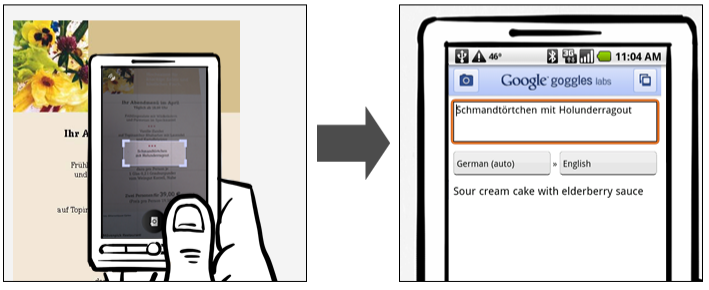
Source: S. Lazebnik

Google Goggles

Google Goggles in action

Click the icons below to see the different kinds of objects and places you can search for using Google Goggles.





Google Search by Image

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

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!

Assistive technologies

Kinect

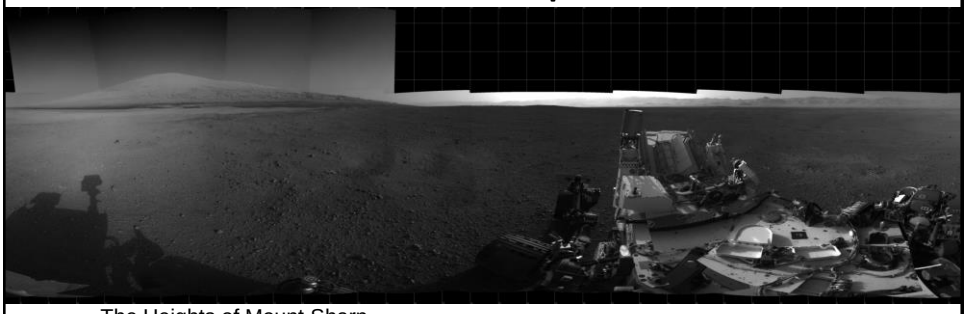
Smart cars

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models

Sources: A. Shashua, S. Seitz

Smart cars

Vision in space



The Heights of Mount Sharp
http://www.nasa.gov/mission_pages/msl/multimedia/pia16077.html
Panorama captured by Curiosity Rover, August 18, 2012 (Sol 12)

Vision systems (JPL) uses for several tasks

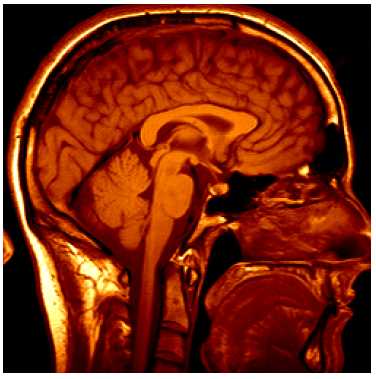
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Robotics

NASA's Mars Curiosity Rover (Mars Science Laboratory)
http://en.wikipedia.org/wiki/Spirit_rover

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

Medical imaging



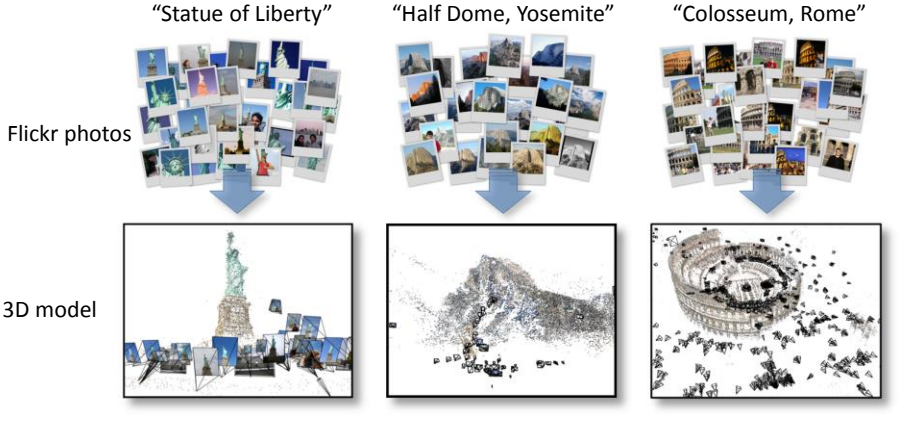
3D imaging
MRI, CT

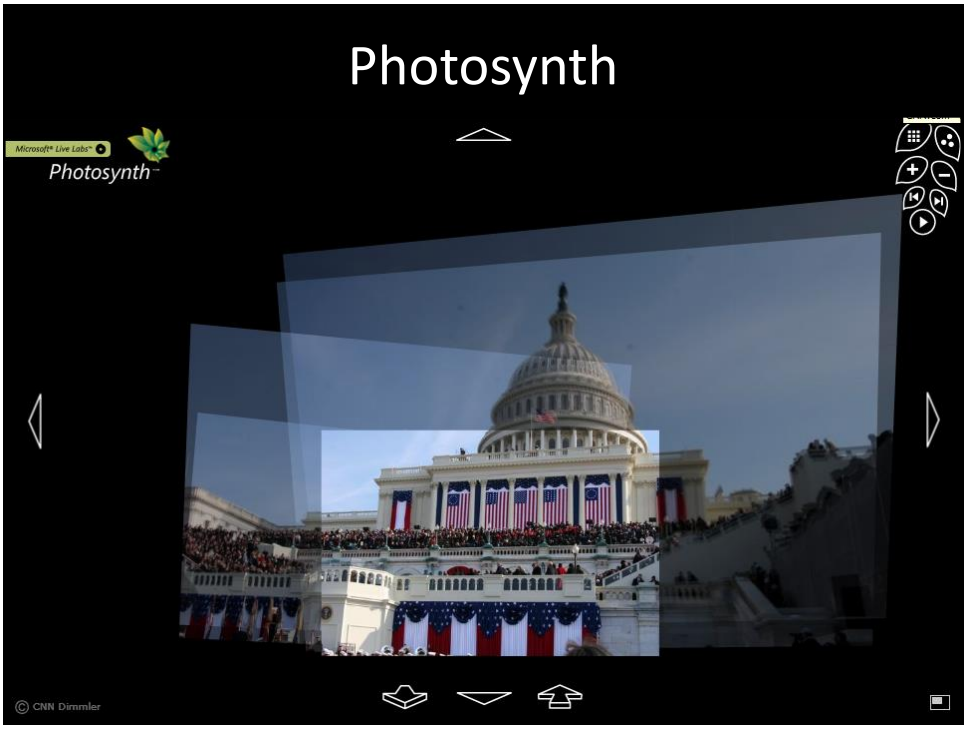
Image guided surgery
Grimson et al., MIT

Source: S. Seitz

My own work

- Automatic 3D reconstruction from Internet photo collections





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…

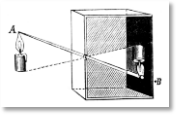
Source: S. Lazebnik

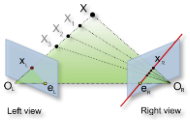
Bottom line

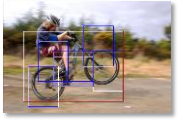
- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

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

Course overview (tentative)





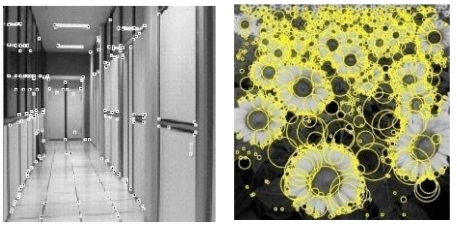


1. Low-level vision
   - image processing, edge detection, feature detection, cameras, image formation
2. Geometry and algorithms
   - projective geometry, stereo, structure from motion, Markov random fields
3. Recognition
   - face detection / recognition, category recognition, segmentation
4. Light,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

1. Low-level vision

- Basic image processing and image formation

Filtering, edge detection



Feature extraction

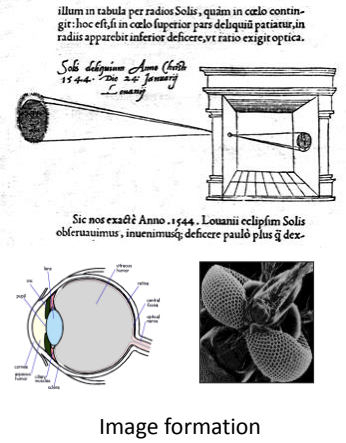
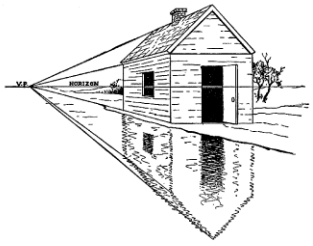


Image formation

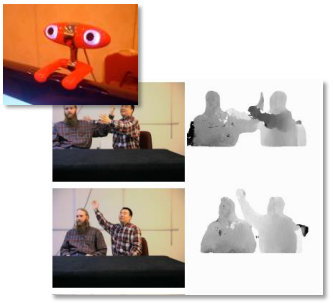
Project: Image Scissors

Project: Feature detection and matching

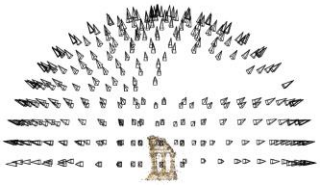
2. Geometry



Projective geometry



Stereo

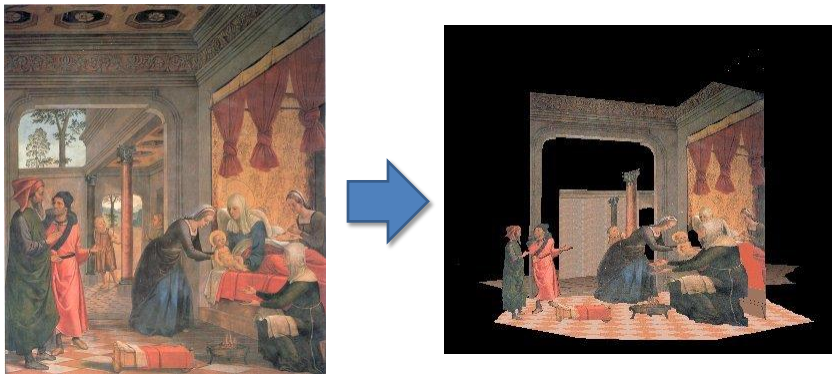


Multi-view stereo

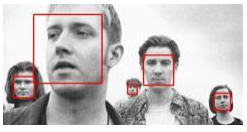
Structure from motion

Project: Creating panoramas

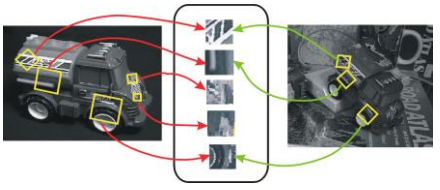
Project: Single-View Modeling



3. Recognition



Face detection and recognition

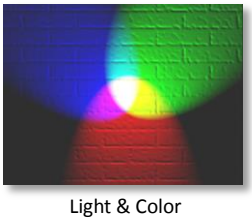


Single instance recognition

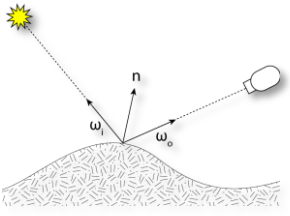
Category recognition

Project: Pedestrian Detection

4. Light, color, and reflectance



Light & Color



Reflectance

Grading

- Occasional quizzes (at the beginning of class)
- One prelim, one final exam

- Rough grade breakdown:
  - Quizzes: 5%
  - Midterm: 20%
  - Programming projects: 50%
  - Final exam: 20%

Late policy

- Three “late days” will be available for the semester
- Late projects will be penalized by 25% for each day it is late, and no extra credit will be awarded.

Academic Integrity

Questions?