Computer Vision

CS 543 / ECE 549

University of Illinois

Instructor: Derek Hoiem

TAs: Jiabin Huang

Kevin Shih

photo source

Today’s class

- A little about me
- Intro to computer vision
- Course logistics
- Questions

About me

Raised in “upstate” NY

About me

1998-2002
Undergrad at SUNY Buffalo
B.S., EE and CSE

2002-2007
Grad at Carnegie Mellon
Ph.D. in Robotics

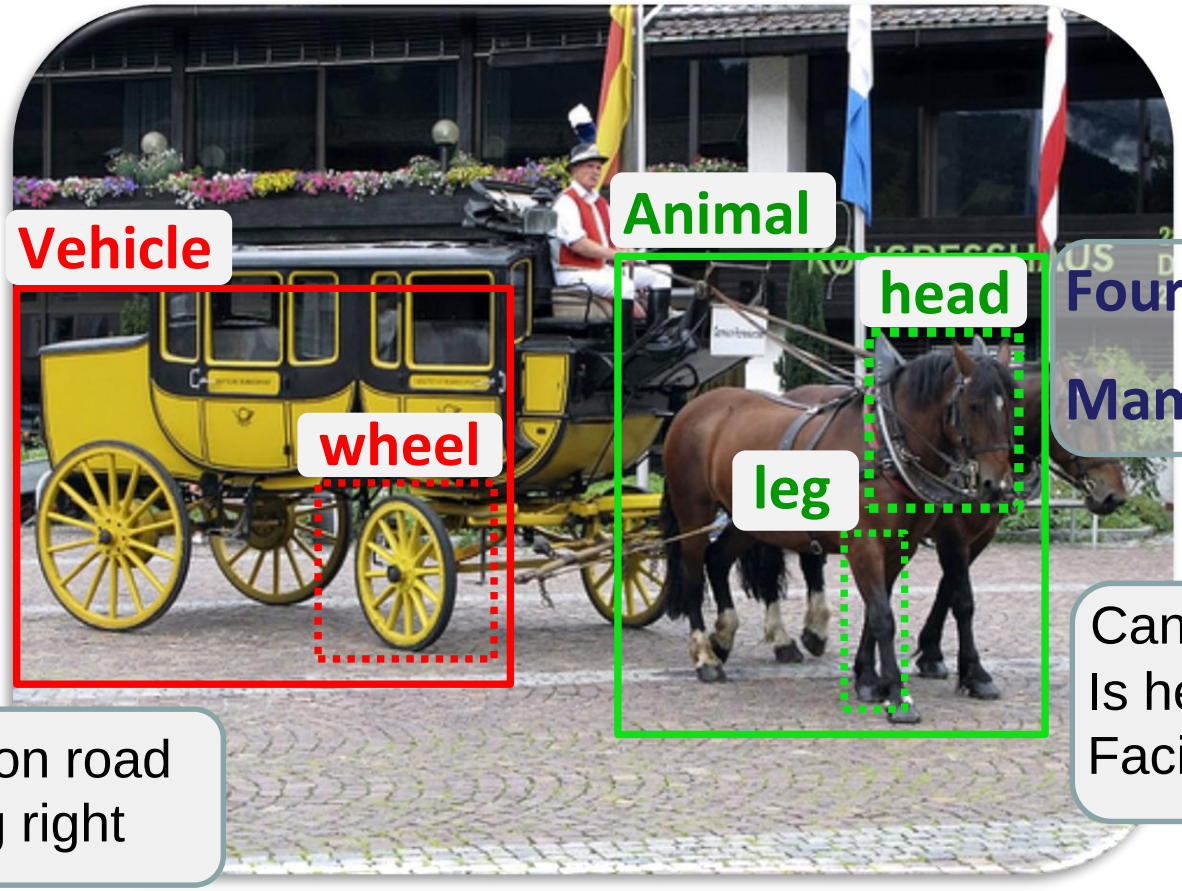
2007-2008
Postdoc at Beckman Institute

2009-
Assistant Prof in CS at UIUC

My research

with Efros, Hebert

My Research



with Farhadi, Endres

My Research

Input Depth Image

Match to exemplar,
Predict complete shape

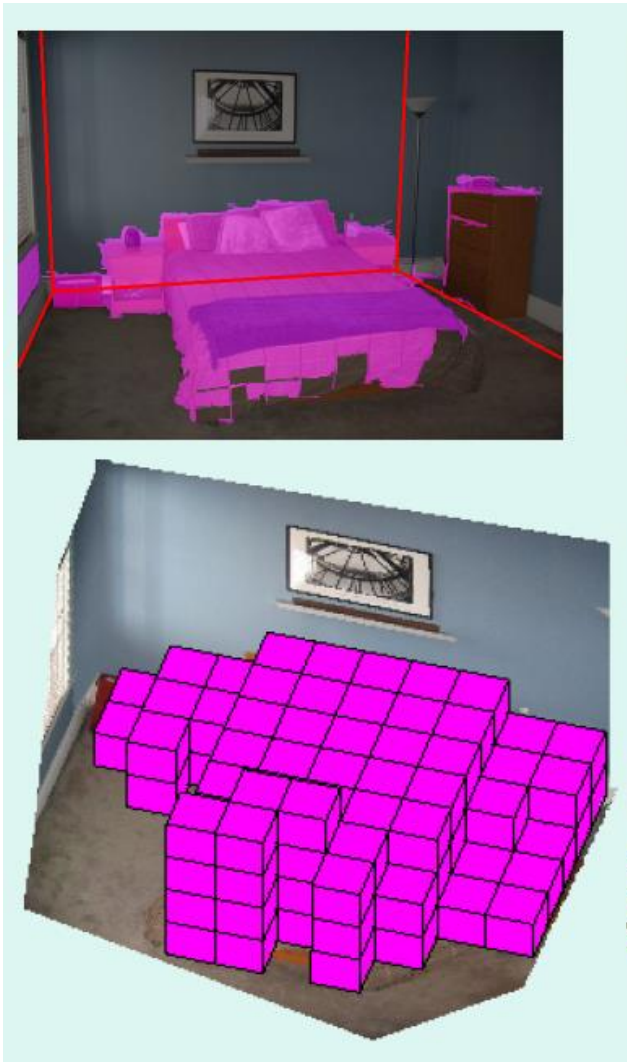
Estimated Shape

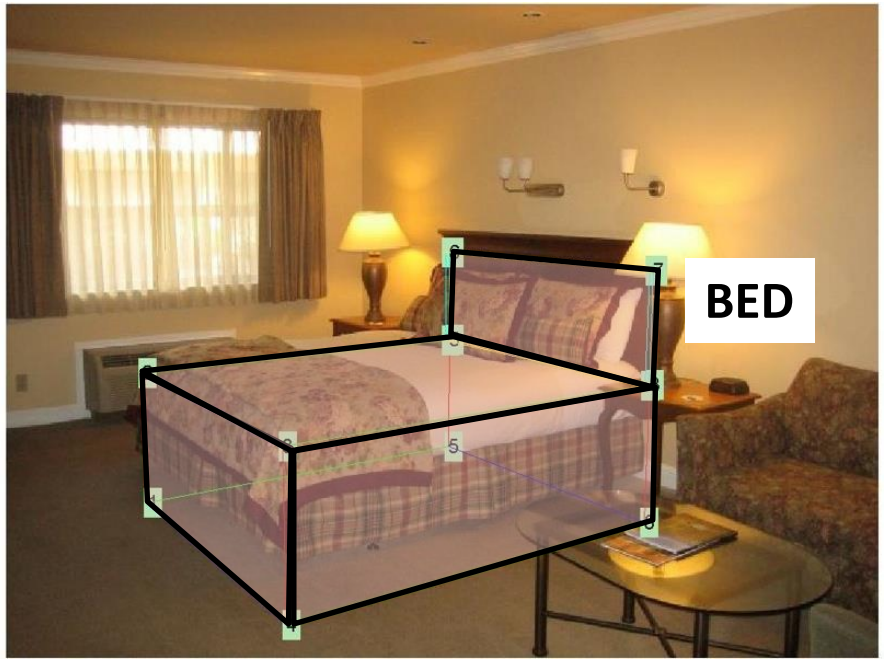
True
Shape

Rock et al., submitted CVPR 15

My Research

Recovering 3D layout and context



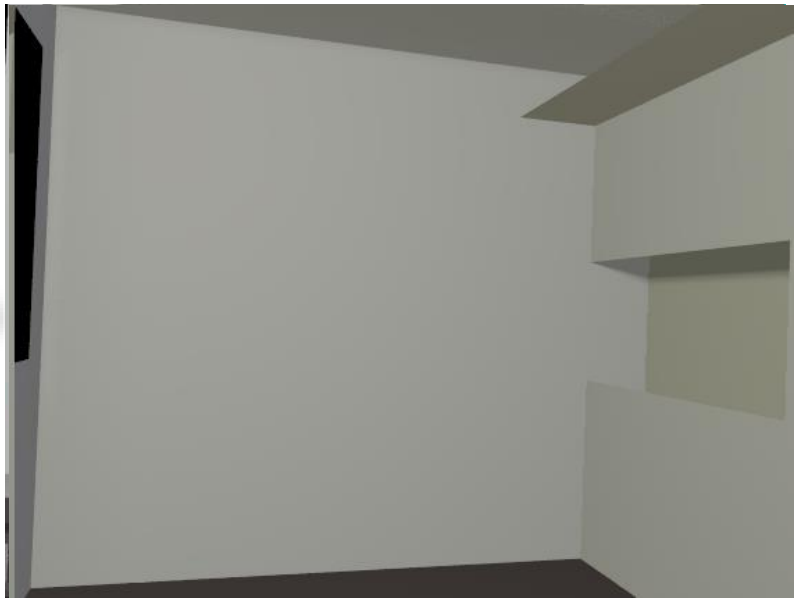


with Hedau, Forsyth

My Research

3D scene model from RGB+D image

RGBD Image



3D Model

Guo et al., submitted CVPR 15

My Research

My Research

Editing images as if they were 3D scenes

(video)

with Karsch, Hedau, Forsyth

Computer Vision

Make computers understand images and video.

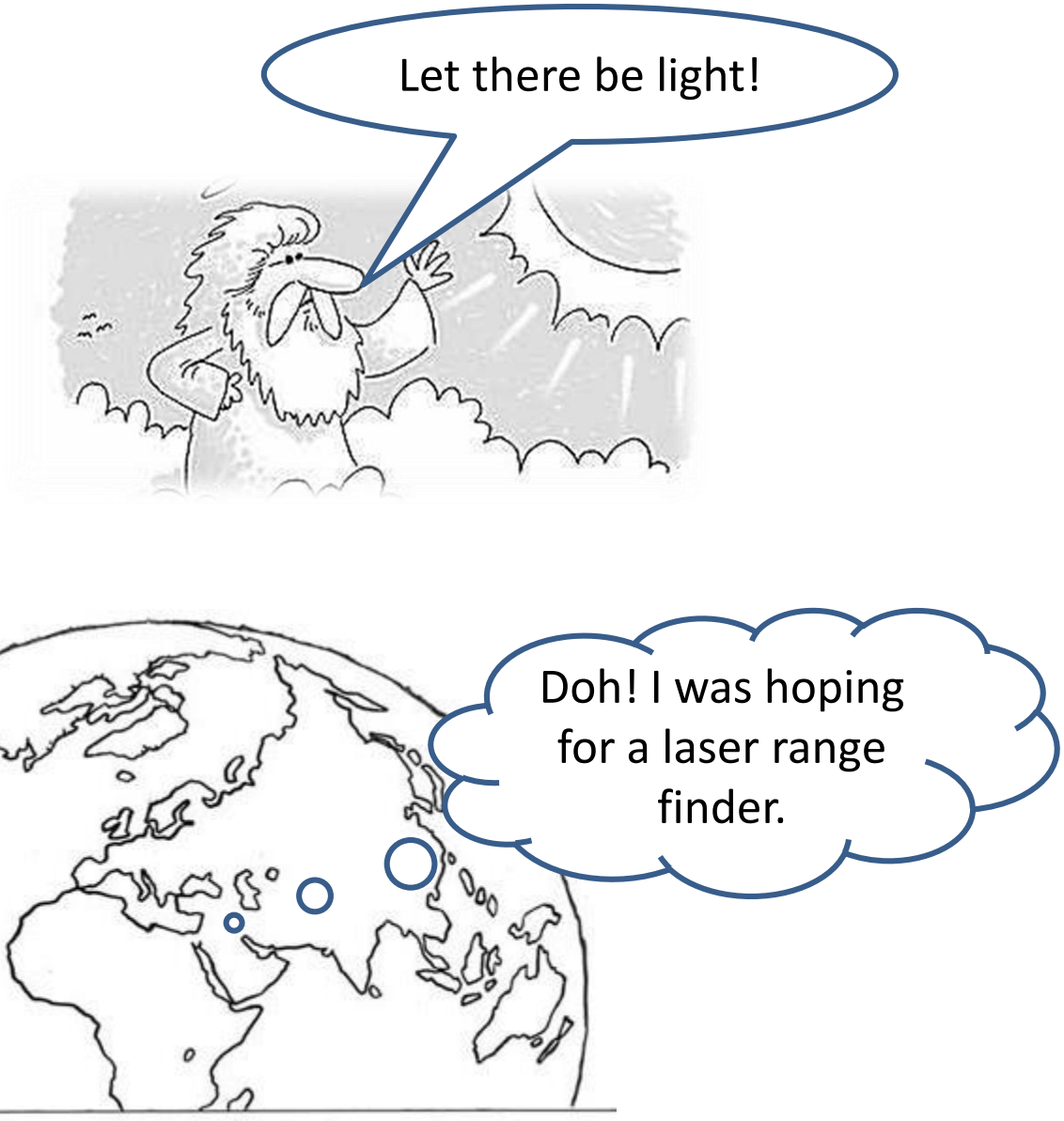
What kind of scene?

Where are the cars?

How far is the building?

…

The miracle of vision

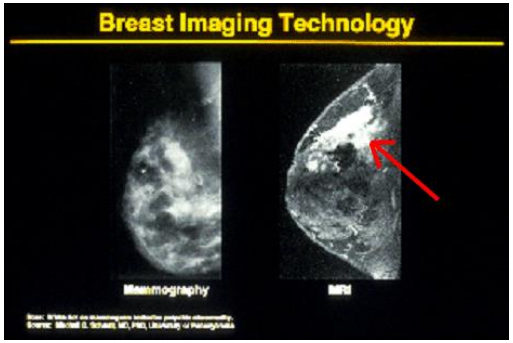


Vision is really hard

- Vision is an amazing feat of natural intelligence
  - More human brain devoted to vision than anything else

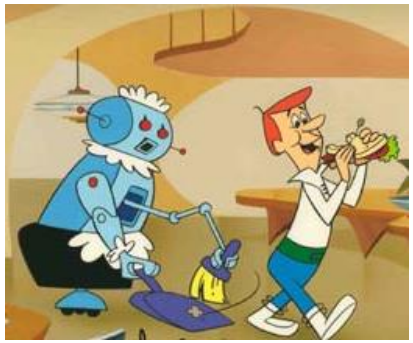
Computer vision matters

Safety



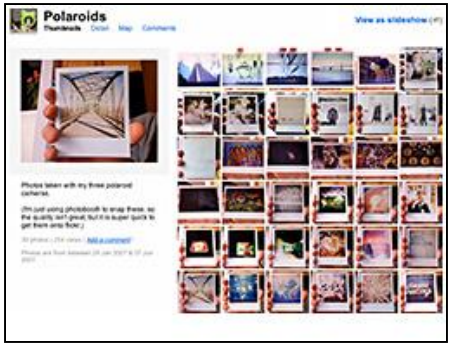
Health

Security



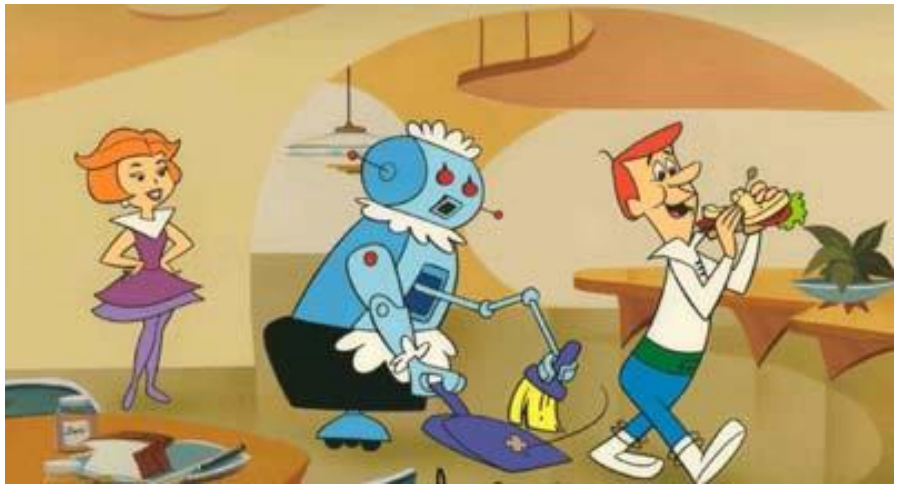
Comfort

Fun



Access

Two reasons for computer vision

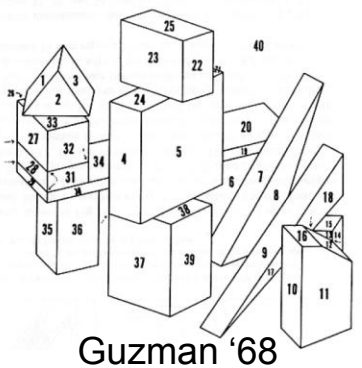


Household Robots

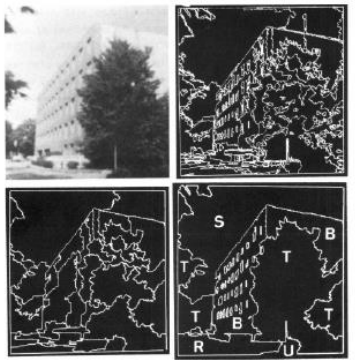
Assisted Driving

Ridiculously brief history of computer vision

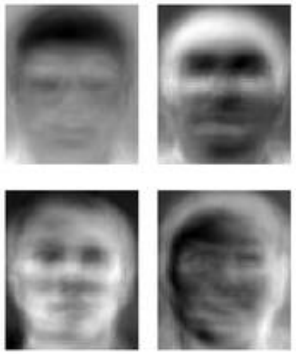
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: ANNs are back for big improvements in recognition; likely large increase in deployed vision systems
- 2020's: autonomous vehicles, the great robot rebellion?



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

How vision is used now

- Examples of state-of-the-art

Some of the following slides by Steve Seitz

Earth viewers (3D modeling)

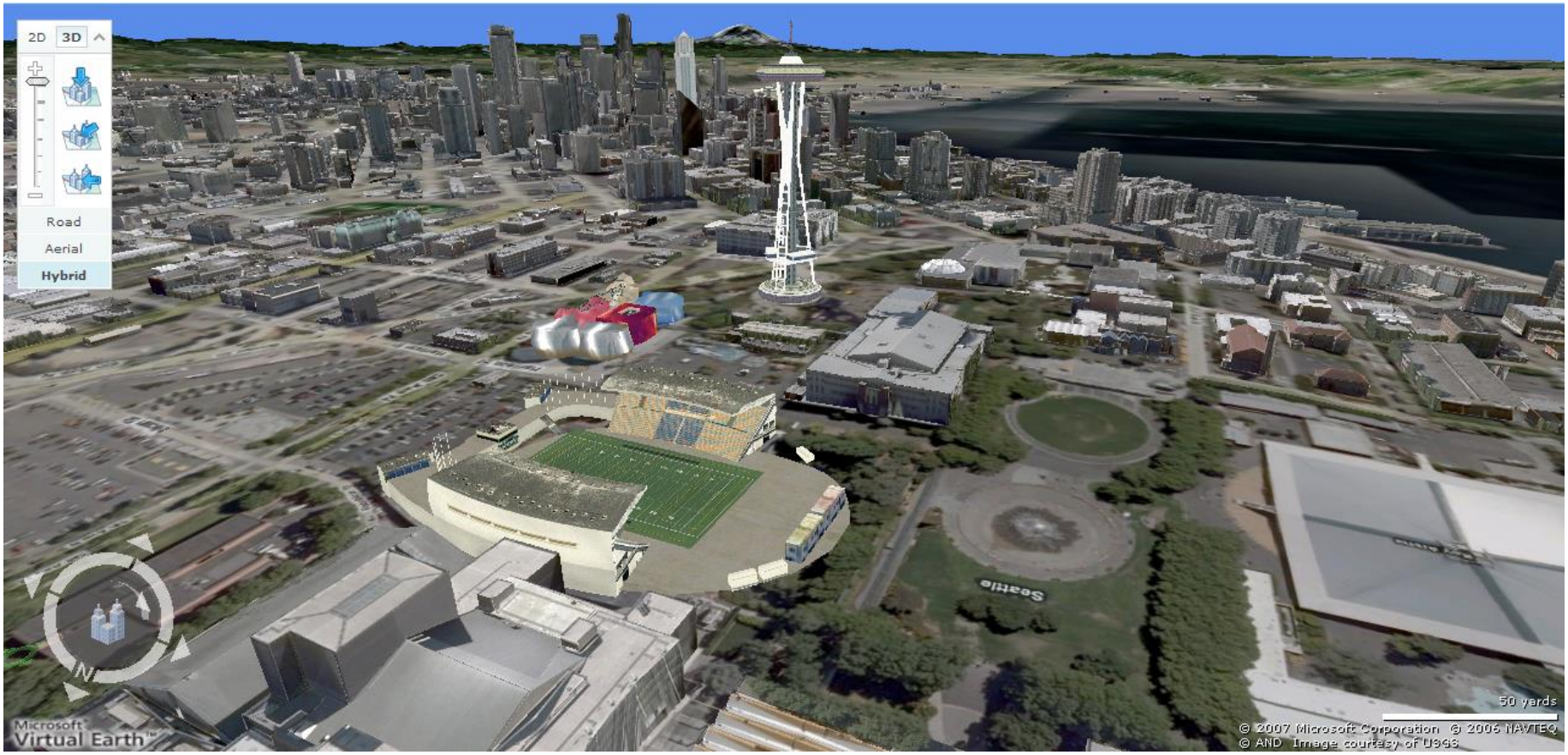


Image from Microsoft's Virtual Earth
(see also: Google Earth)

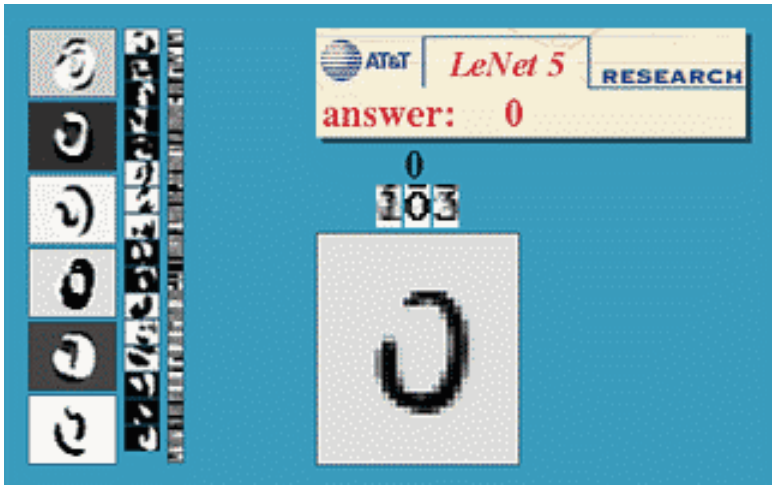
3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009

Optical character recognition (OCR)

Technology to convert scanned docs to text

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



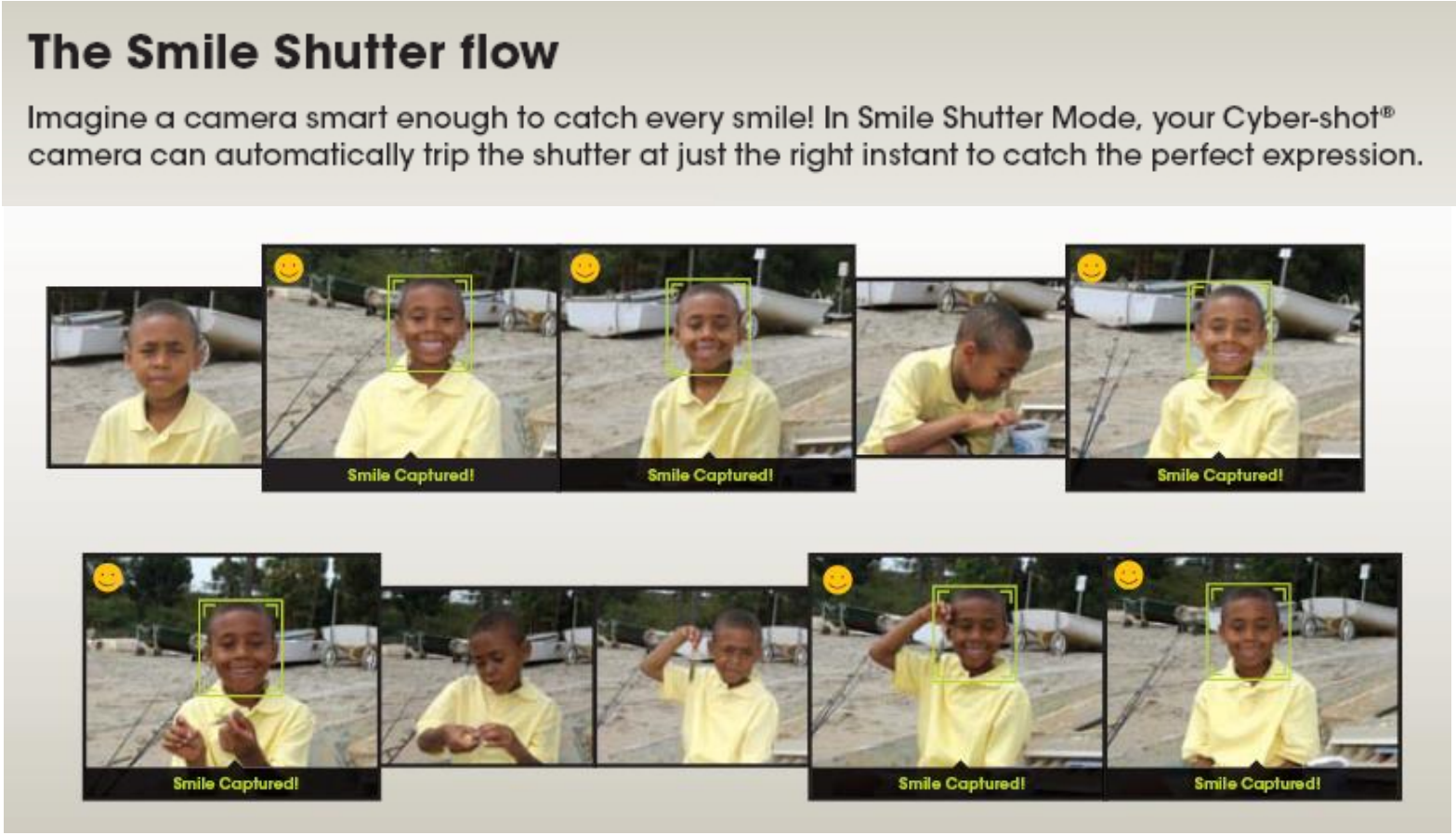
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Face detection

- Most digital cameras detect faces (and more)
  - Canon, Sony, Fuji, …

Smile detection



[Sony Cyber-shot® T70 Digital Still Camera]

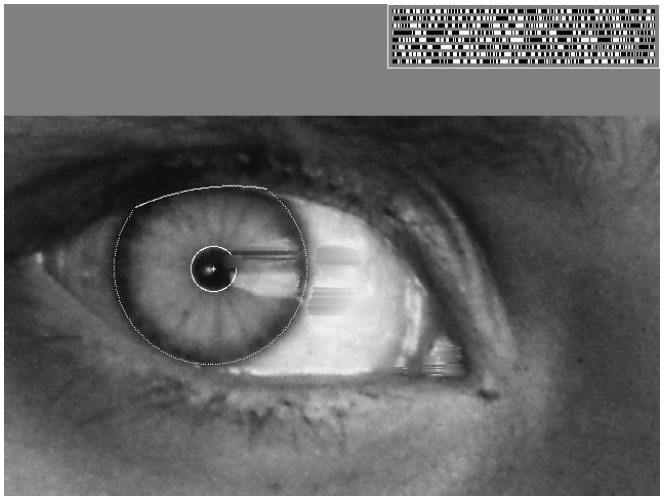
Object recognition (in supermarkets)

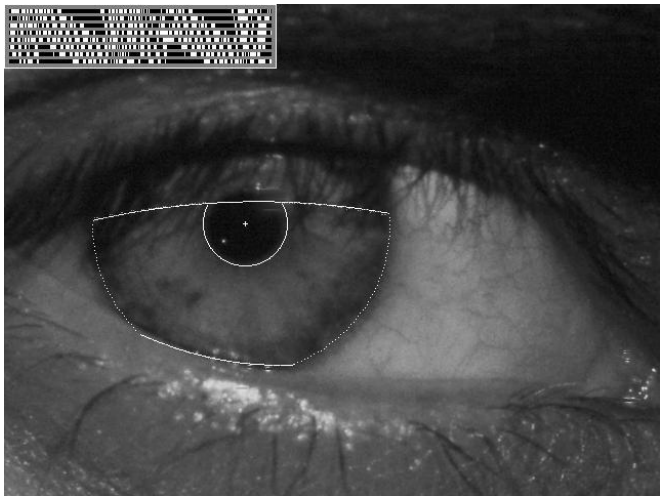
LaneHawk by EvolutionRobotics

“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… “

Vision-based biometrics

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





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/

Object recognition (in mobile phones)

[Point & Find](), [Nokia]()

[Google Goggles]()

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Sports

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

http://www.sportvision.com/video.html

Smart cars

- Mobileye: Vision systems currently in many cars

http://mobileye.com/technology/applications/vehicle-detection/forward-colision-warning/
http://mobileye.com/technology/applications/pedestrian-detection/pedestrian-collision-warning/

"Subaru thinks cameras are better than radar cruise"
http://www.roadandtrack.com/new-cars/news/a6852/subaru-camera-controlled-cruise/

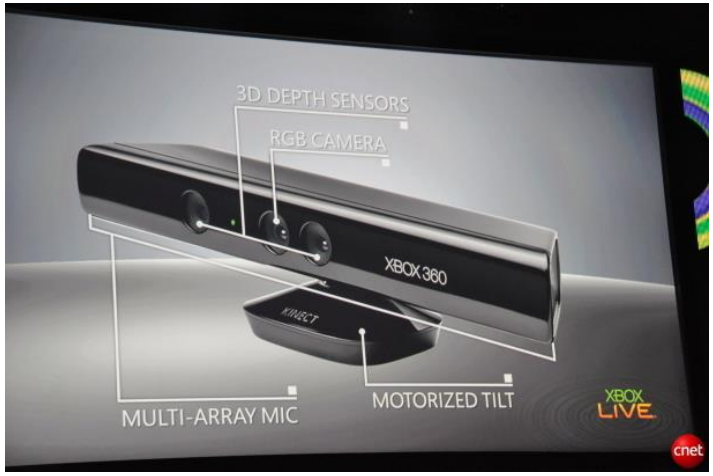
Google cars

Google in talks with Ford, Toyota and Volkswagen to realise driverless cars

http://www.theatlantic.com/technology/archive/2014/05/all-the-world-a-track-the-trick-that-makes-googles-self-driving-cars-work/370871/

Interactive Games: Kinect

- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



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

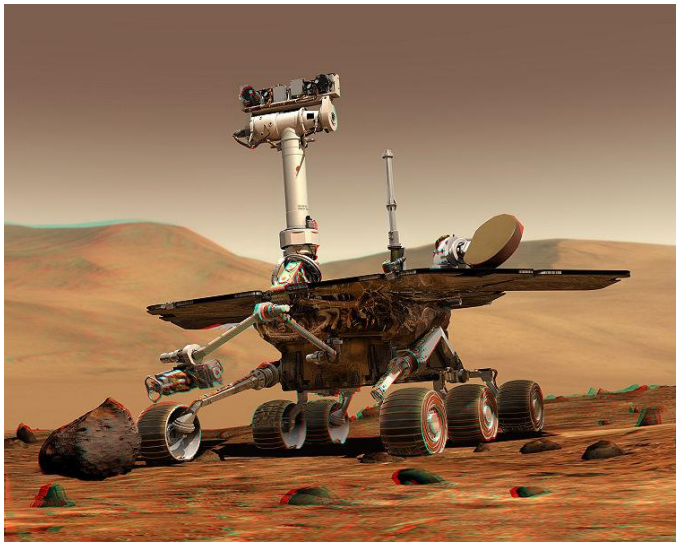
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Industrial robots

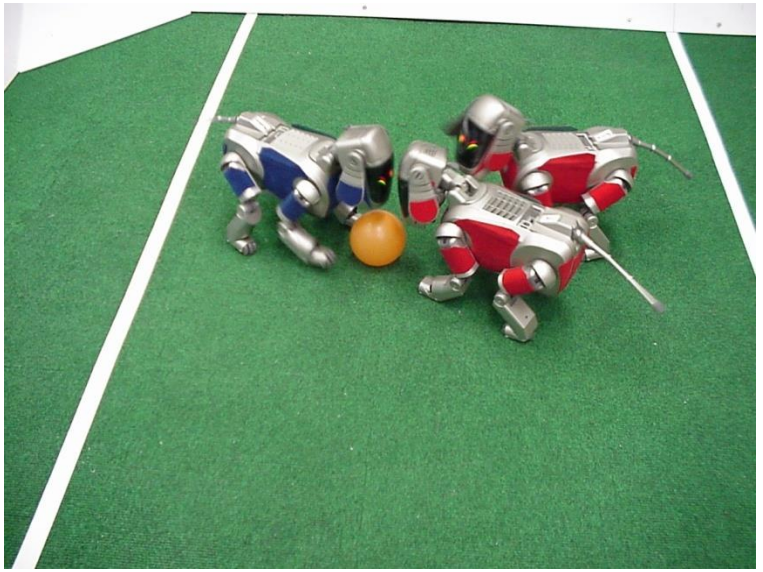
Vision-guided robots position nut runners on wheels

http://www.automationworld.com/computer-vision-opportunity-or-threat

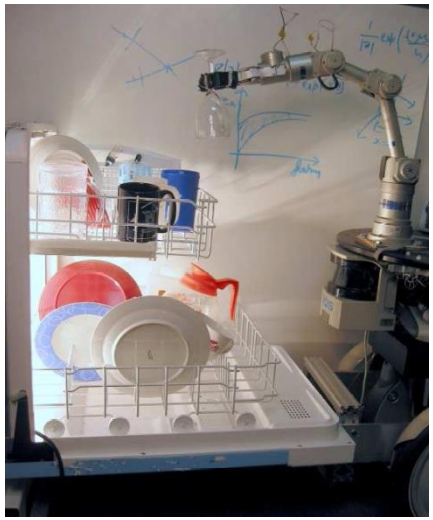
Mobile robots



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



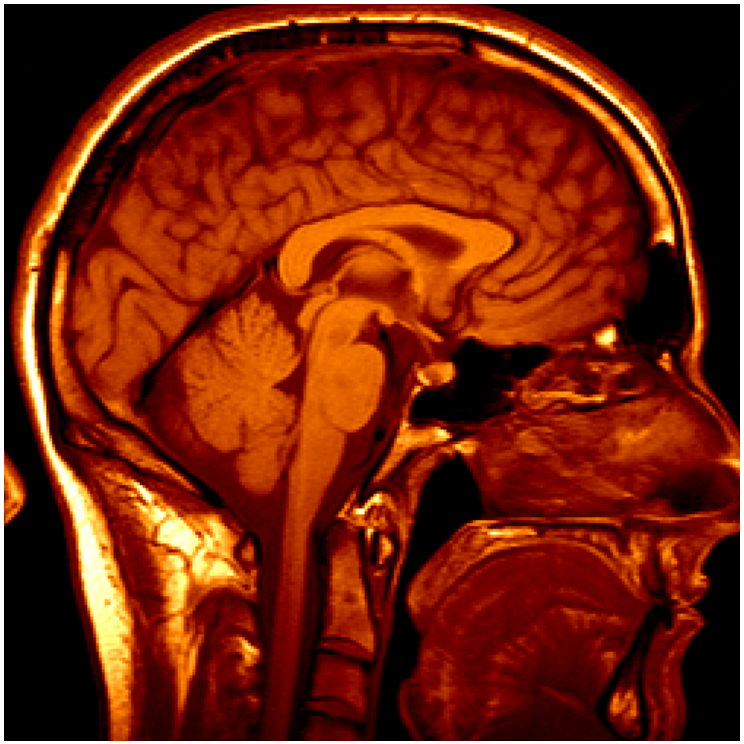
http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

http://www.youtube.com/watch?v=DF39Ygp53mQ

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

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

Course outline

Prof: Derek Hoiem (dhoiem@illinois.edu), SC3312

TA: Jiabin Huang (jbhuang1@Illinois.edu), ECE3034

Kevin Shih (kjshih2@Illinois.edu) SC3307

Web page:

https://courses.engr.illinois.edu/cs543/

Grades

- Homeworks (75%)
- Final project (25%)
- Attendance

Late policy

- 10% per day
- One late HW will be forgiven (up to one week)

Academic Integrity

- Can discuss hw with peers, but don’t copy
- Carefully document any sources within hw hand-in
- Don’t use code from Internet unless you have permission
  - If you’re not sure, ask

Getting help outside of class

Office hours

- Time, see website
- Otherwise, just stop by. If I'm not there, send me an e-mail.

Discussion board:
https://piazza.com/class/i4ohqjpy49s6ga

TA: Jiabin Huang (jbhuang1@Illinois.edu)
Kevin Shih (kjshih2@Illinois.edu)

Readings/Textbook

- Computer Vision: A Modern Approach (2nd edition) by David Forsyth and Jean Ponce (2011)
- See syllabus for other useful books

What to expect from this course

- Broad coverage (geometry, image processing, recognition, multiview, video)
- Background to delve deeper into any computer vision-related topic
- Practical experience
- Lots of work, tough material, fast pace, but hopefully lots of learning too!

Topics

- Interpreting Intensities
  - What determines the brightness and color of a pixel?
  - How can we use image filters to extract meaningful information from the image?
- Correspondence and Alignment
  - How can we find corresponding points in objects or scenes?
  - How can we estimate the transformation between them?
- Perspective and 3D Geometry
  - How can we map between the 3D world and the 2D image?
  - How can we recover 3D coordinates from images or video?
- Grouping and Segmentation
  - How can we group pixels into meaningful regions?
- Categorization and Object Recognition
  - How can we represent images and categorize them?
  - How can we recognize categories of objects?
- Advanced Topics
  - Action recognition, 3D scenes and context, CNNs, …

Prerequisites

- **Linear algebra**, basic calculus, and probability

- Experience with image processing or Matlab will help but is not necessary

Goals and Expectations

- My goal: maximize the learning effectiveness of your time

- What I expect from you
  - Attend and participate, when possible
  - Start assignments well before deadline
  - Tell me what's working and suggest improvements

Final comments

- Feedback

- To do
  - Sign up for newsgroup:
    https://piazza.com/class/i4ohqjpy49s6ga
  - Read syllabus, etc.

- Next class: light and color

- Questions?