CSCE 643: Introduction to Computer Vision

Jinxiang Chai

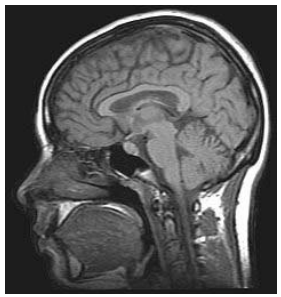
What is Computer Vision?

What is Computer Vision?

- **Computer vision** is the science and technology of machines that see.

- Concerned with the theory for building artificial systems that obtain information from images.

- The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medi r



Computer Vision

Make computers understand images and videos.

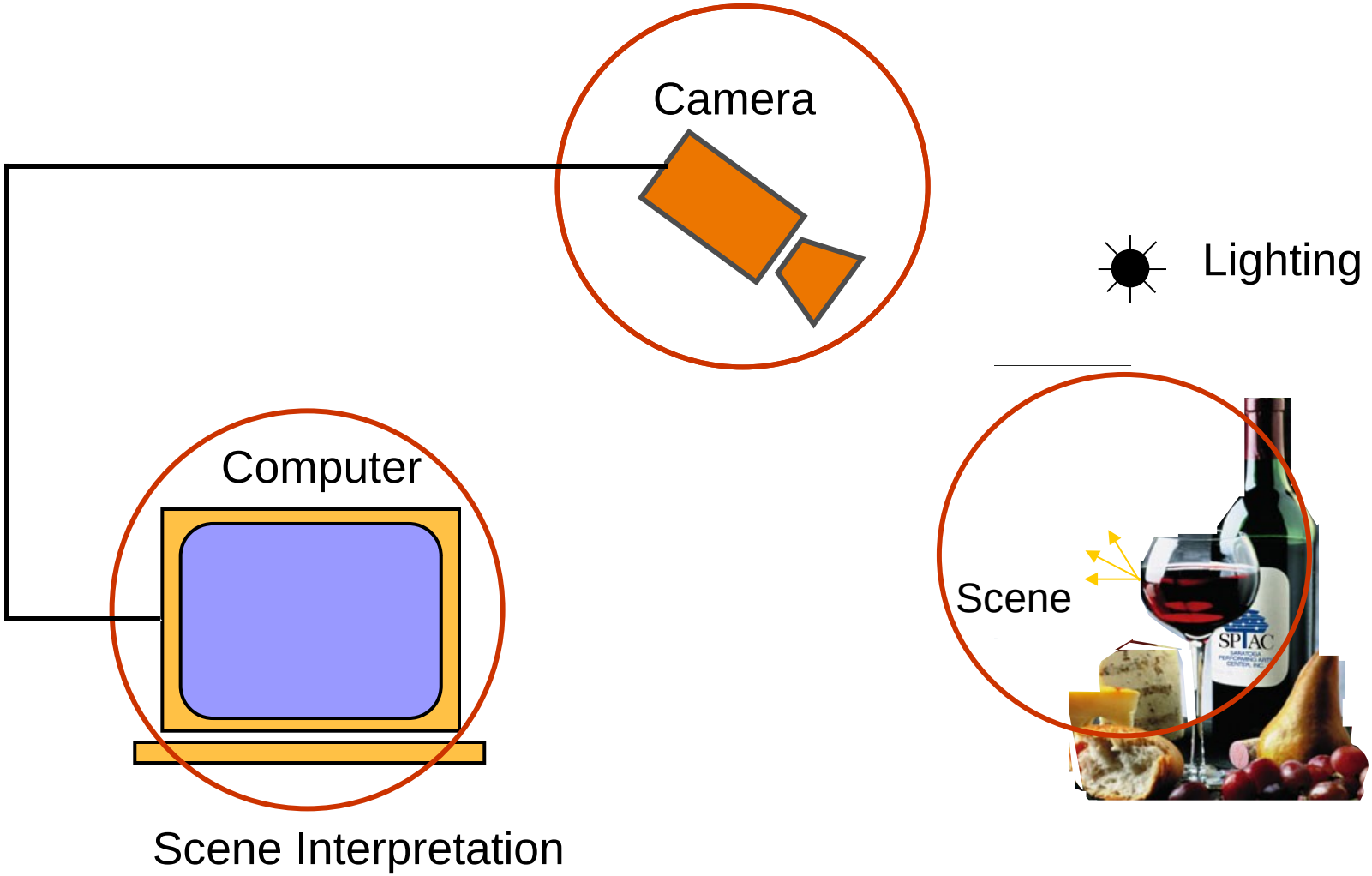
What kind of scene?

Where are the cars?

How far is the building?

…

Components of a computer vision system



Srinivasa Narasimhan's slide

Computer vision vs human vision

What we see

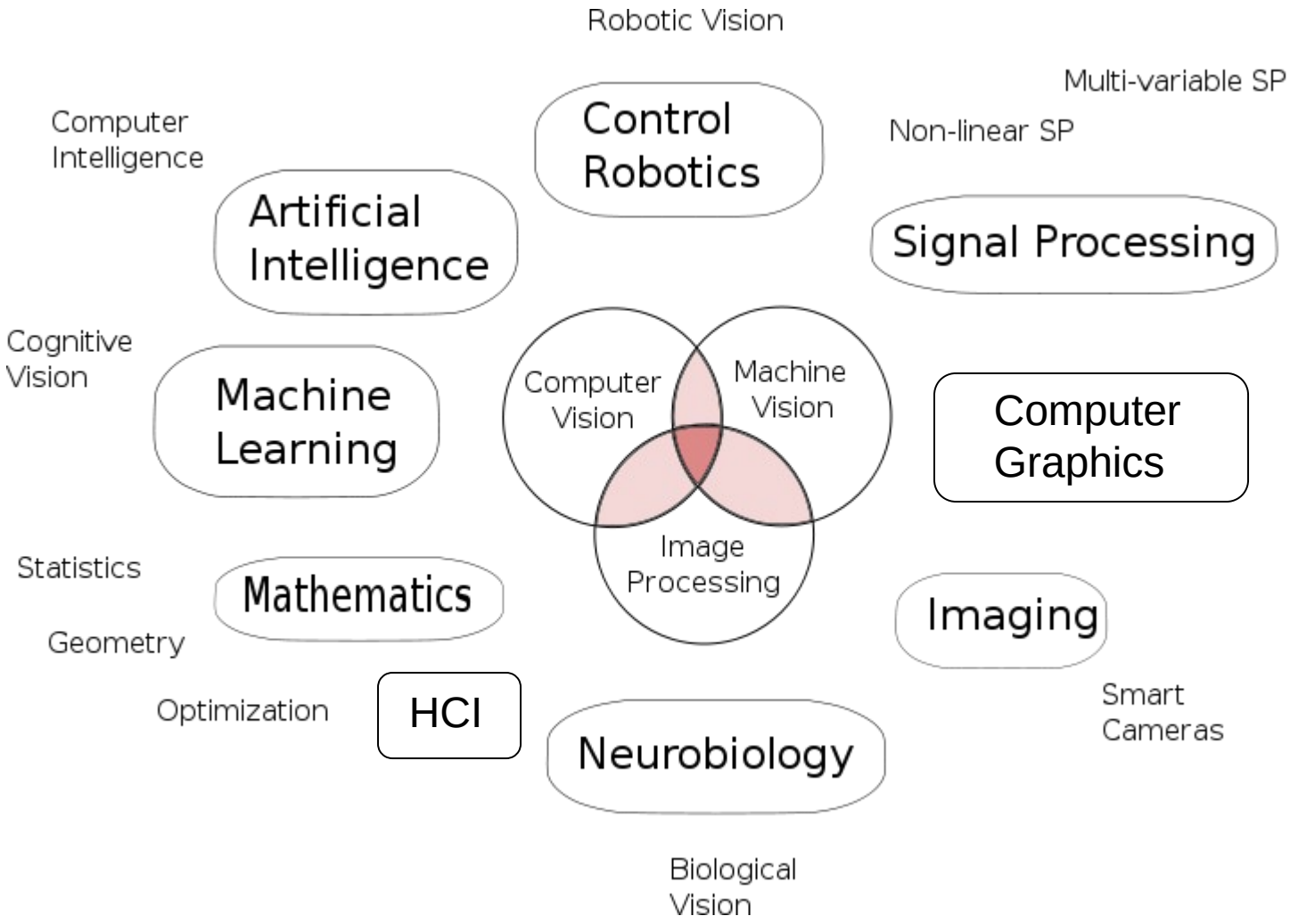
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of Macaque brain
  - More human brain devoted to vision than anything else

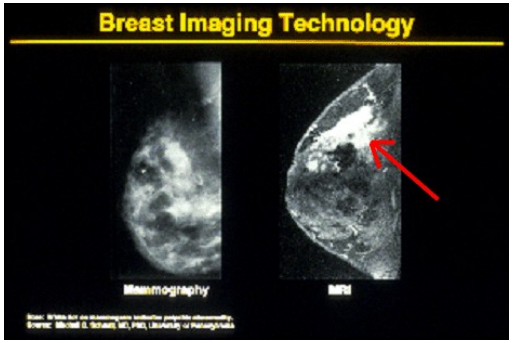
Vision is multidisciplinary



From wiki

Why computer vision matters

Safety

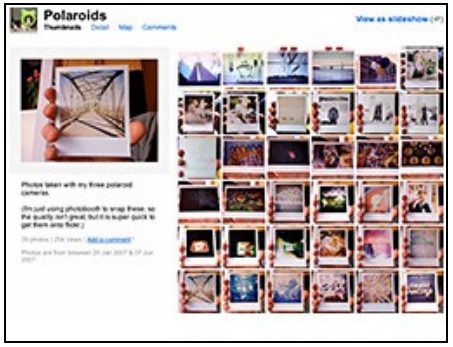


Health

Security

Comfort

Fun



Access

A little story about Computer Vision

In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”. We now know that the problem is slightly more difficult than that. (Szeliski 2009, Computer Vision)

A little story about Computer Vision

Founder, MIT AI project

In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”. We now know that the problem is slightly more difficult than that.

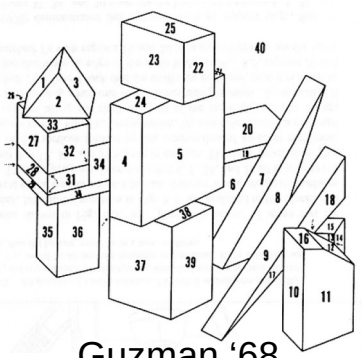
A little story about Computer Vision

In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”. We now know that the problem is slightly more difficult than that.

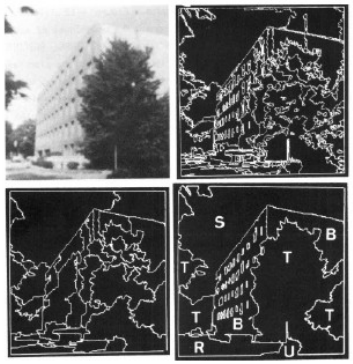
Image Understanding

Ridiculously brief history of computer vision

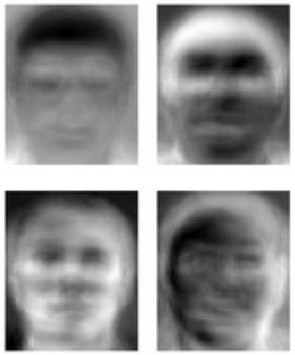
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960’s: interpretation of synthetic worlds
- 1970’s: some progress on interpreting selected images
- 1980’s: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990’s: face recognition; statistical analysis in vogue
- 2000’s: broader recognition; large annotated datasets available; video processing starts; vision & graphis; vision for HCI; internet vision, etc.



Guzman ‘68



Ohta Kanade ‘78



Turk and Pentland ‘91

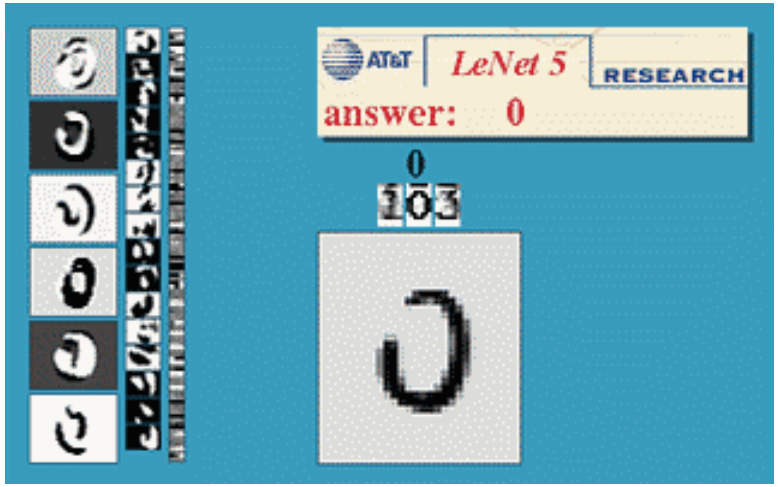
How vision is used now

- Examples of state-of-the-art

Optical character recognition (OCR)

Technology to convert scanned docs to text

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



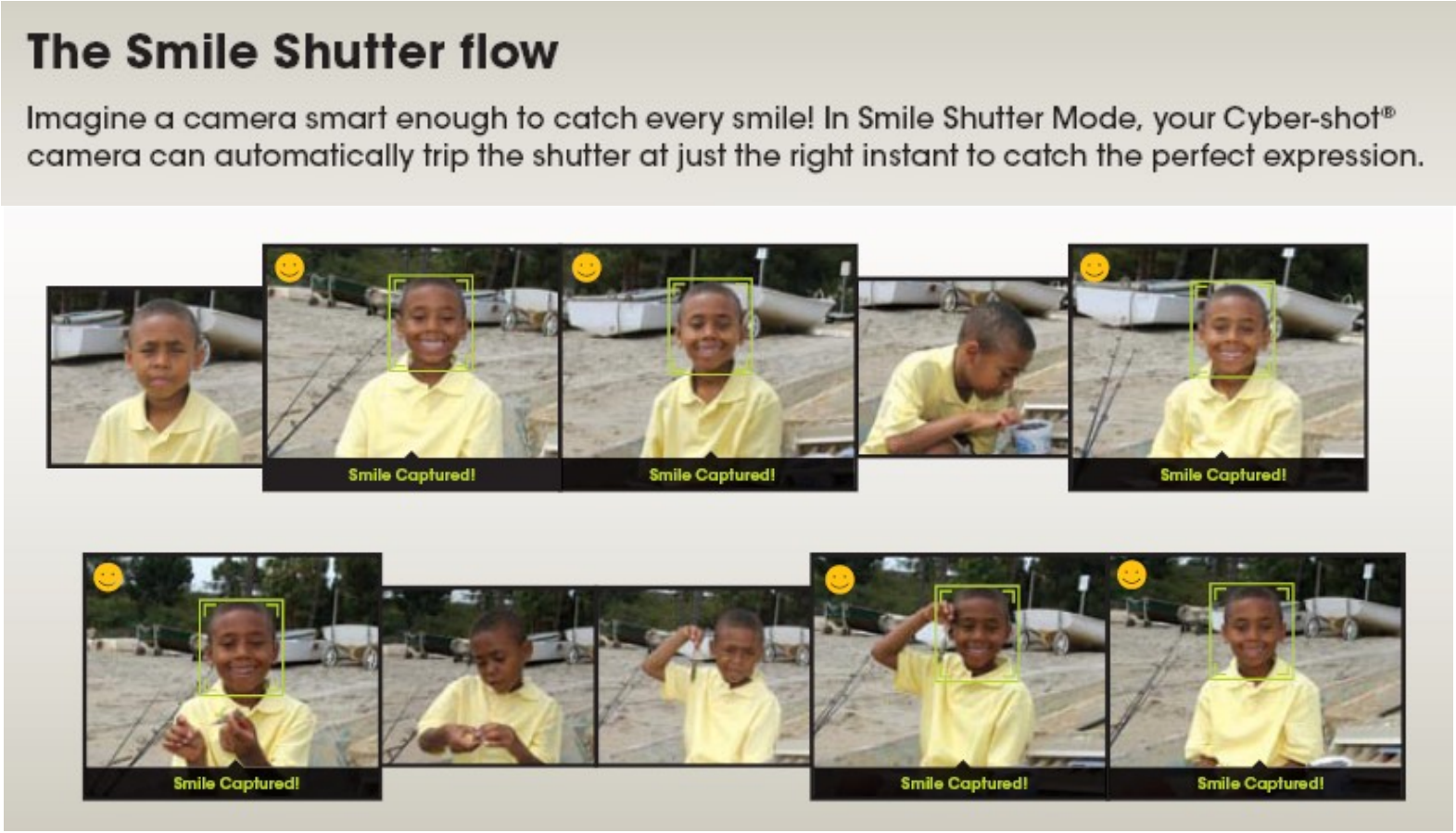
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Face detection

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

Smile detection



Sony Cyber-shot® T70 Digital Still Camera

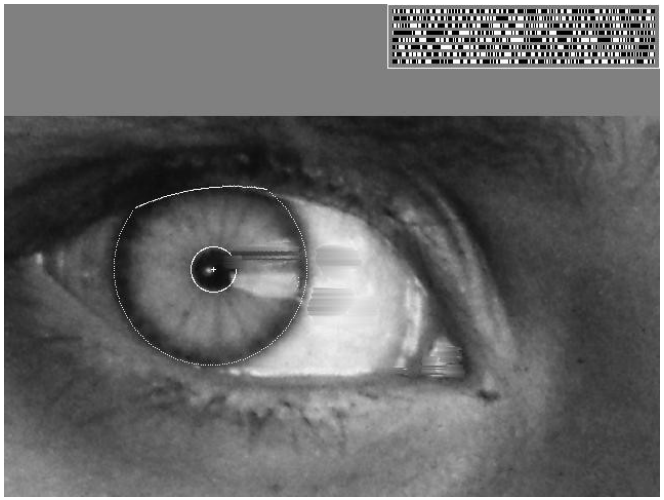
Object recognition (in supermarkets)

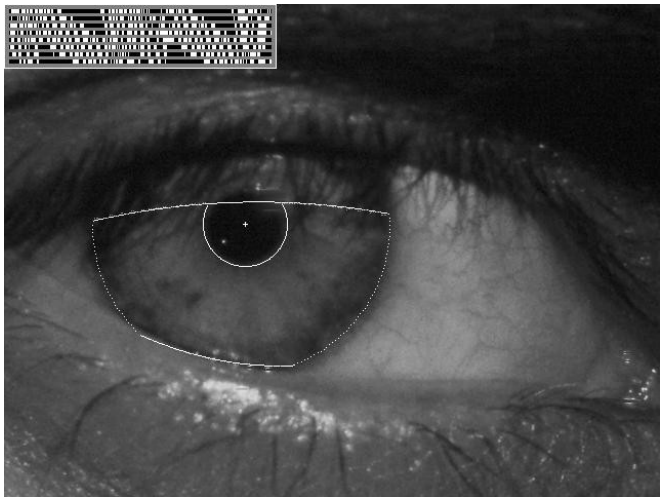
LaneHawk by EvolutionRobotics

“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

Slide content courtesy of Amnon Shashua

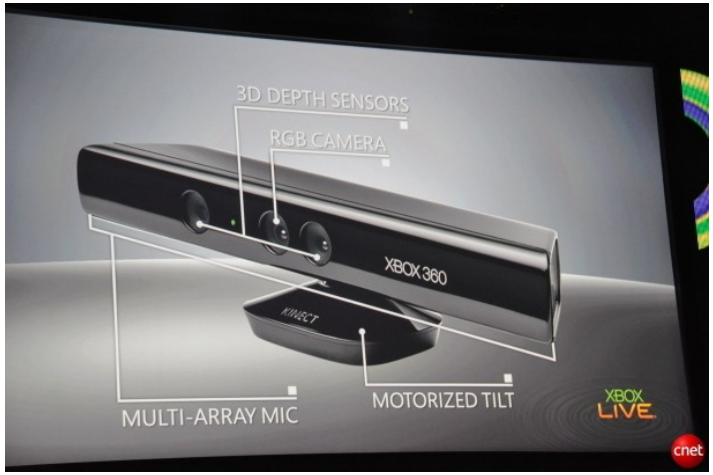
- Mobileye [wiki article]
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.

Google cars

http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence

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
- 3D tracking, reconstruction, and interaction: http://research.microsoft.com/en-us/projects/surfacerecon/default.aspx



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.

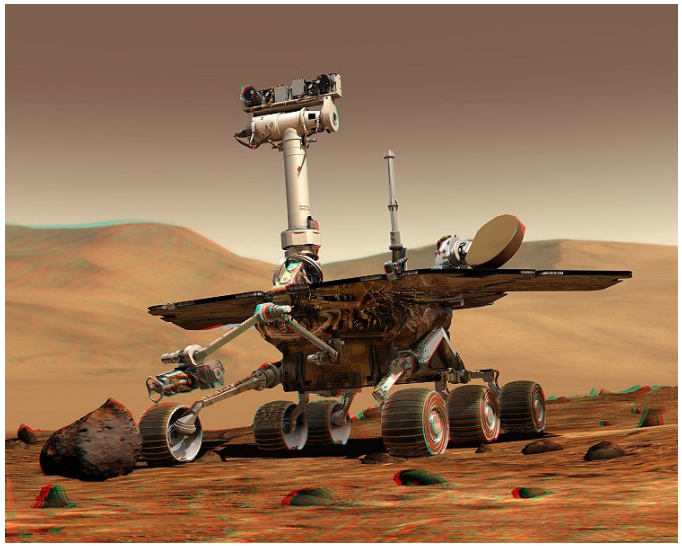
Vision systems (JPL) used for several tasks

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

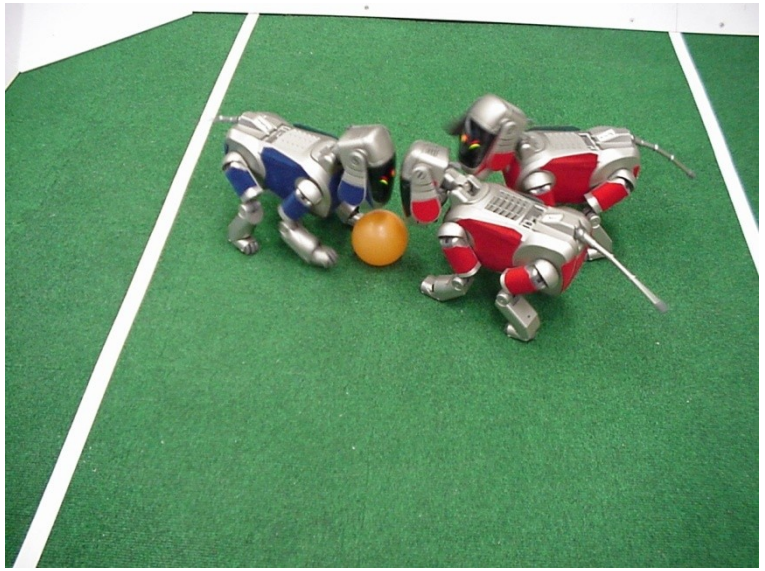
Industrial robots

Vision-guided robots position nut runners on wheels

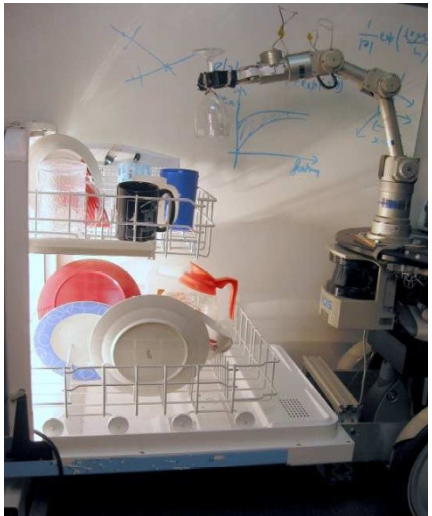
Mobile robots



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

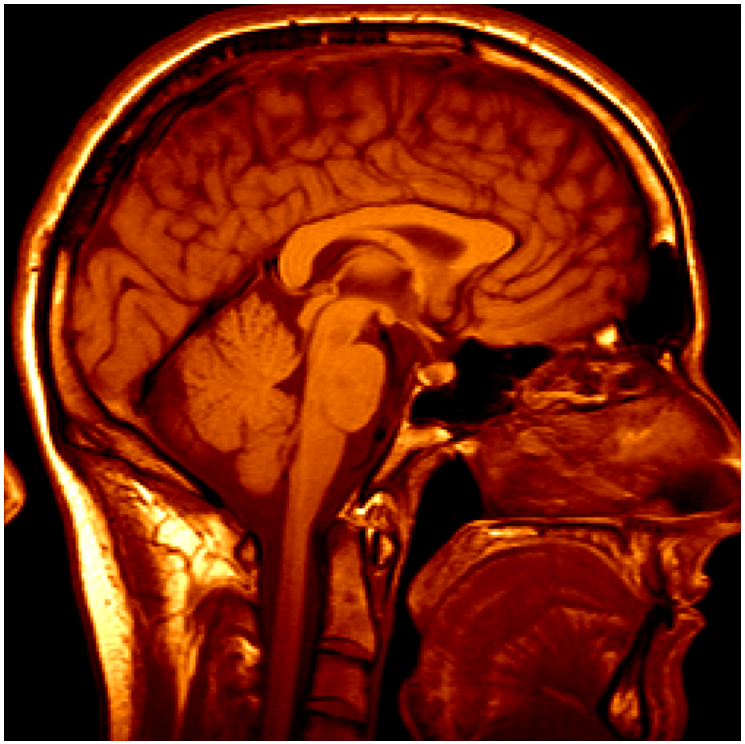


http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Prerequisites

- A good working knowledge of C/C++, Java or Matlab
- A good understand of math (linear algebra, basic calculus, basic probability)
- Willing to learn new stuffs (optimization, statistical learning etc.)

Grading Schemes

- Assignments (25%)
- Final project (40%+)
- Class participation/discussion (5%)
- Paper readings and presentation (25%)

Final Project

Approved by the professor

Student can work in a group of two

Submit your code and final project report

Final presentation & in class demos

Late policy: 20% reduction per day if you do not have good reasons

Other Information

My email: jchai@cs.tamu.edu

My homepage: http://faculty.cs.tamu.edu/jchai

My office: Rm 527D Bright

Office hours: Wed 4:00-5:00 Pm or by appointment

Course webpage: http://www.cs.tamu.edu/jchai/csce643/

Textbook: **Computer vision: Algorithms and Applications,** Richard Szelisk

Email Me Today

Your background

- Vision, Graphics, machine learning, image processing
- Math (linear algebra, statistics, calculus, optimization, etc.)
- Coding (C++, java, matlab, etc.)

Your research Interest?

Master/Ph.D. (year)?

Why do you take this class?