Introduction to Computer Vision

Today’s Class

- Introduction
- What is Computer Vision?
- Sample Problems

What is Computer Vision?

- Computer Vision is about understanding images. These images can be
  - Greyscale or Colour or Multi-spectral
  - Snapshots or video sequences
  - Taken with a static or moving camera
  - Taken of a stationary or dynamic scene
  - Taken with a calibrated or un-calibrated camera
- What Computer Vision aims to do is to extract some useful information from these images for…
  - Inspection purposes
  - Analysis purposes
  - Control purposes

Computer Vision

What kind of scene?

Where are the cars?

How far is the building?

…

Computer Vision and Nearby Fields

- Computer Graphics : Models to Images
- Comp. Photography : Images to Images
- Image Processing : Images to Images
- Computer Vision : **Images to Models**

Computer Vision Scope

- 3D Scene Reconstruction
  - Stereo, Multi View
  - Shape from X (motion, focus, shadow, texture)
  - Time-of-flight sensors
- 2D/3D Scene Analysis
  - Segmentation
  - Detection
  - Registration
  - Object Recognition
- Augmented Reality

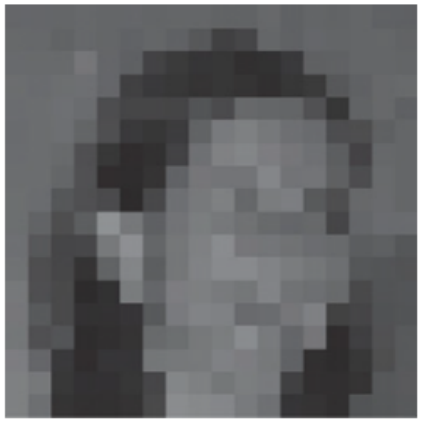
Vision is really hard

- 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 really hard

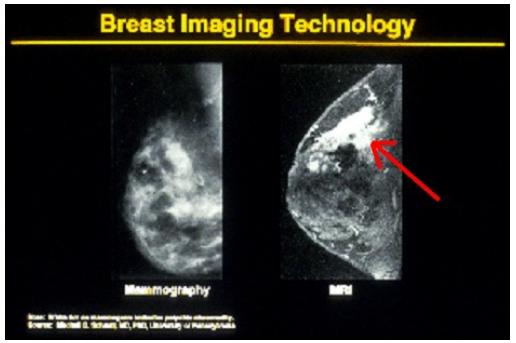
- "What we experience, apparently directly, is actually very different from what is recorded by our sense organs.“ [Perception: From Sense to Object by J. Wilding]

67	67	66	68	66	67	64	65	65	63	63	69	61	64	63	66	61	60
69	68	63	68	65	62	65	61	50	26	32	65	61	67	64	65	66	63
72	71	70	87	67	60	28	21	17	18	13	15	20	59	61	65	66	64
75	73	76	78	67	26	20	19	16	18	16	13	18	21	50	61	69	70
74	75	78	74	39	31	31	30	46	37	69	66	64	43	18	63	69	60
73	75	77	64	41	20	18	22	63	92	99	88	78	73	39	40	59	65
74	75	71	42	19	12	14	28	79	102	107	96	87	79	57	29	68	66
75	75	66	43	12	11	16	62	87	84	84	108	83	84	59	39	70	66
76	74	49	42	37	10	34	78	90	99	68	94	97	51	40	69	72	65
76	63	40	57	123	88	60	83	95	88	80	71	67	69	32	67	73	73
78	50	32	33	90	121	66	86	100	116	87	85	80	74	71	56	58	48
80	40	33	16	63	107	57	86	103	113	113	104	94	86	77	48	47	45
88	41	35	10	15	94	67	96	98	91	86	105	81	77	71	35	45	47
87	51	35	15	15	17	51	92	104	101	72	74	87	100	27	31	44	46
86	42	47	11	13	16	71	76	89	95	116	91	67	87	12	25	43	51
96	67	20	12	17	17	86	89	90	101	96	89	62	13	11	19	40	51
99	88	19	15	15	18	32	107	99	86	95	92	26	13	13	16	49	52
99	77	16	14	14	16	35	115	111	109	91	79	17	16	13	46	48	51



Why computer vision matters

Safety

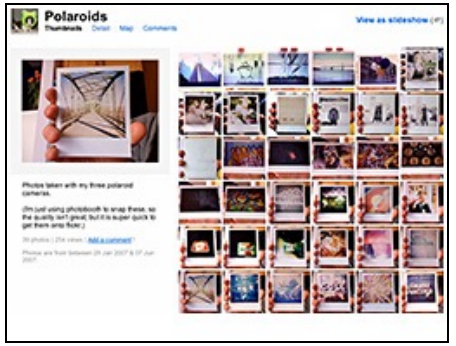


Health

Security

Comfort

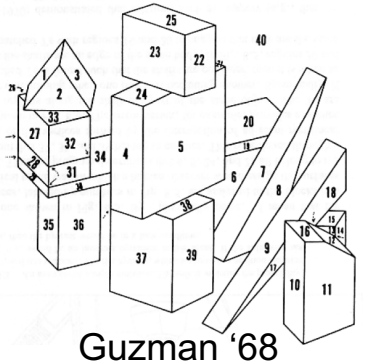
Fun



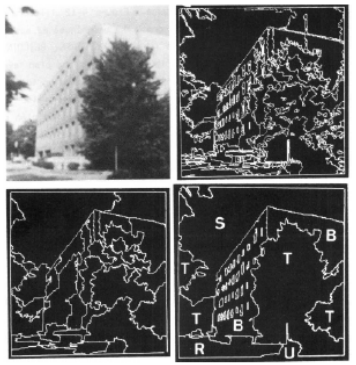
Access

Ridiculously brief history of computer vision

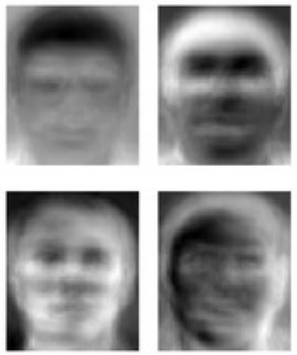
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- CNN, Deep Learning, Autonomous Vehicles



Guzman '68



Ohta Kanade '78



Turk and Pentland '91

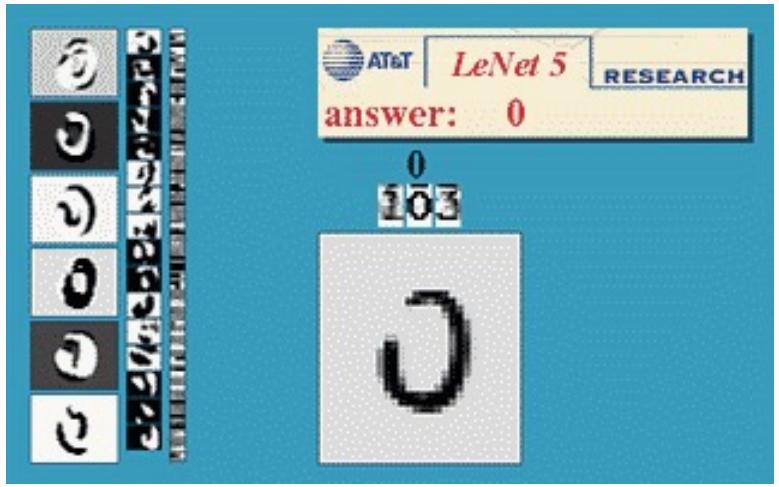
How vision is used now

- You just saw examples of current systems.
  - Many of these are less than 5-10 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

Optical character recognition (OCR)

Technology to convert scanned docs to text

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

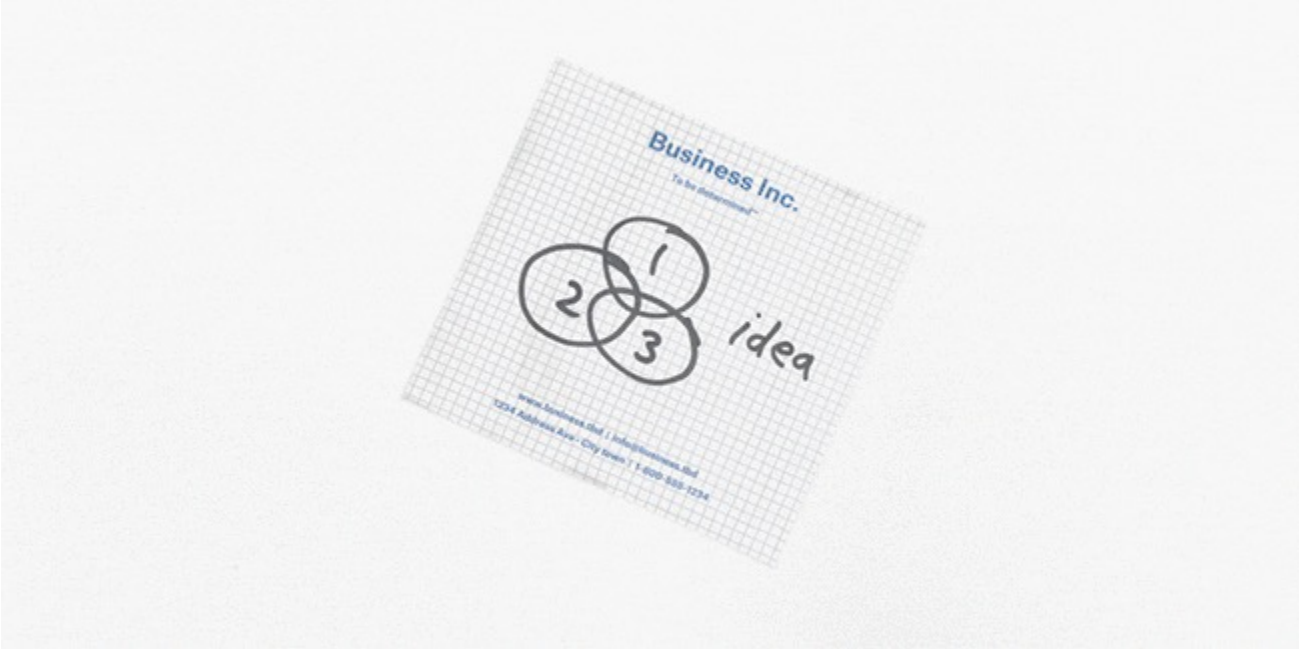


Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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

Dropbox OCR

- Convolutional Neural Nets (CNNs)



Optical character recognition (OCR)

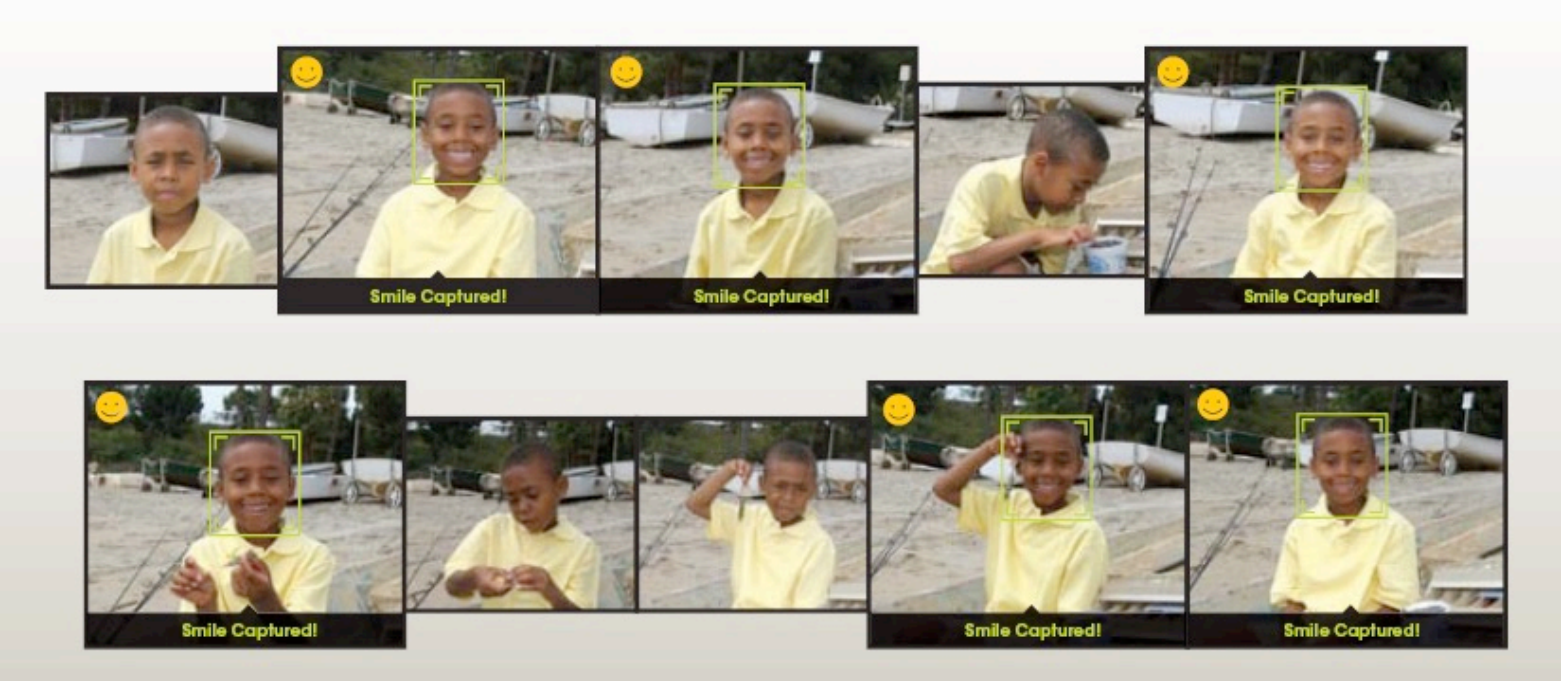
Google Street View (84.2% accuracy)

- street names, numbers, and businesses

AutoCars - Uber bought CMU’s lab

Smile detection

- The Smile Shutter flow



Sony Cyber-shot® T70 Digital Still Camera

3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009

3D Maps with LIDAR

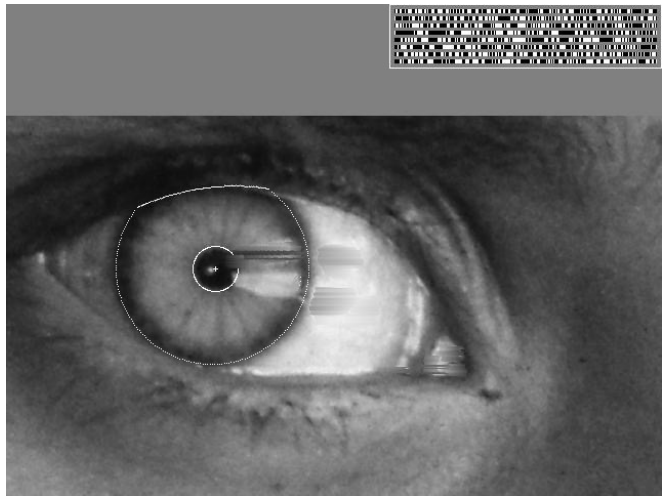
Object recognition (in supermarkets)

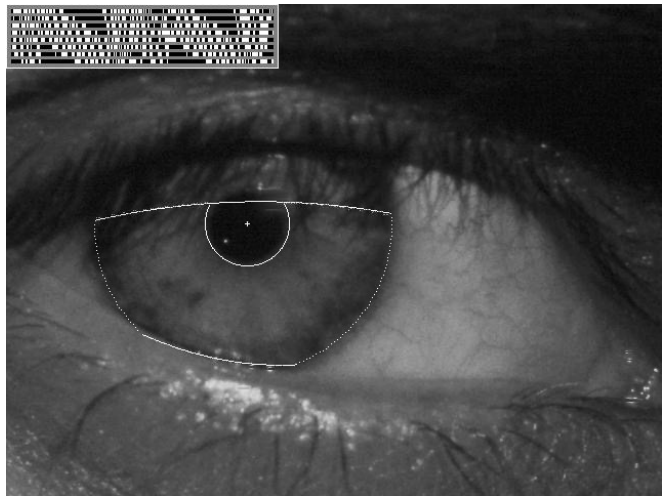
- LaneHawk by EvolutionRobotics

“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





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

Smart cars

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

Vision-based interaction (and games)

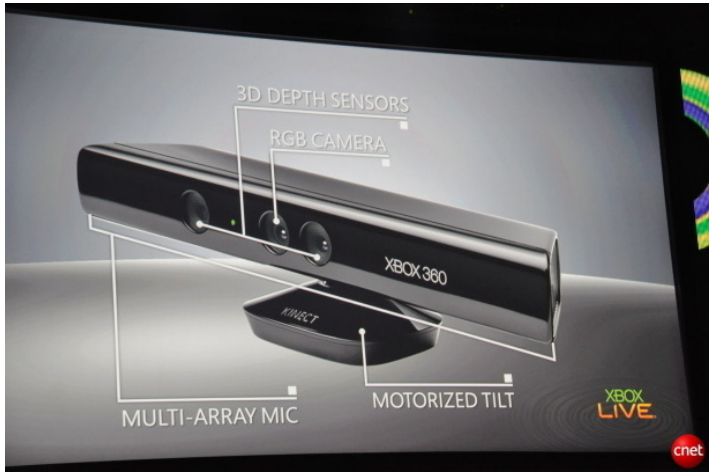
Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Digimask: put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNET
Camera tracking a crowd, based on this work.

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

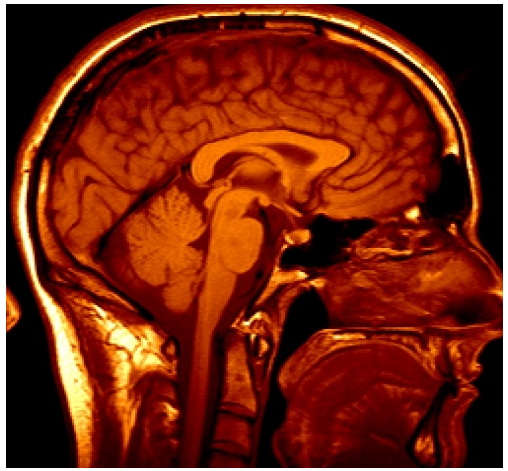
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Industrial robots

Vision-guided robots position nut runners on wheels

Augmented Reality

Medical imaging



Course contents

- Basics of image processing
  - Digitization
  - Color models
  - Noise and Smoothing
  - Histograms
  - Geometric Operations
- Edge Image Processing
  - Edge detection
  - Contour Following
  - Hough Transform
- Camera models
  - Calibration
- Features
  - Harris, FAST
  - SIFT
- Video
  - Motion detection
- Structure from X
  - Stereo
  - Motion
- Object Detection

General Comments

- Prerequisites
  - Data structures
  - A good working knowledge of a programming Language
    - (or willingness/time to pick it up quickly!)
    - Optional Matlab
  - Linear algebra
- Course does not assume prior imaging experience
  - image processing, graphics, etc.
- Book:
  - Computer Vision: Algorithms and Applications
  - A Practical Introduction to Computer Vision with OpenCV

Grading Policy

- 3 programming assignments
  - 10%, 2-3 weeks each
- 1 midterm exam
  - 35%
- 1 final exam/project
  - 35%, open ended of your choice, but needs
  - project proposal after 4 weeks
  - Final report and working demo