EEE422/6082 Computational Vision

Introduction

Ling Shao

Many slides from Derek Hoiem

Computer Vision

Make computers understand images and video.

What kind of scene?

Where are the cars?

How far is the building?

…

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

Why computer vision matters

Safety

Health

Security

Comfort

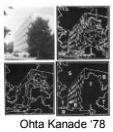
Fun

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

Guzman '68



Ohta Kanade '78

Turk and Pentland '91

Current state of the art

- Some examples of what current vision systems can do

Many of the following slides by Steve Seitz

Earth viewers (3D modeling)

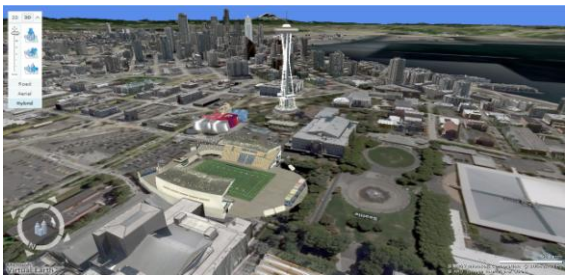


Image from Microsoft's Virtual Earth
(see also: Google Earth)

Photosynth.net

Based on Photo Tourism
by Noah Snavely, Steve Seitz, and Rick Szeliski

3D from multiple images

Building Rome in a Day: Agarwal et al. 2009

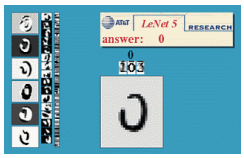
3D from one image

Hoiem Efros Hebert SIGGRAPH 2005

Optical character recognition (OCR)

Technology to convert scanned docs to text

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



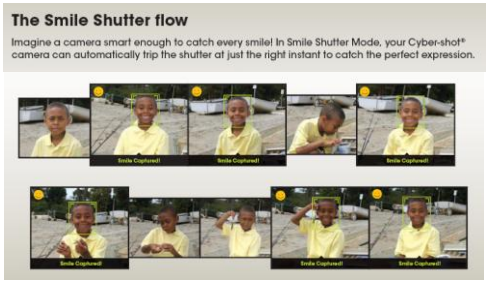
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

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)

- This is becoming real:
  - Point & Find, Nokia

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic
Click here for interactive demo

Sports

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

Smart cars

Slide content courtesy of Amnon Shashua

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

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.

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

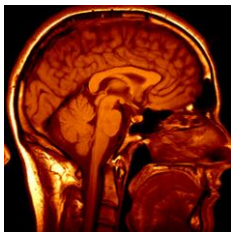
Mobile robots

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

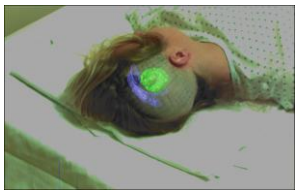


Image guided surgery
Grimson et al., MIT

Recent news

The New York Times

Television Begins a Push Into the 3rd Dimension

BY BRIAN STELTER AND BRAD STONE

Ralph Kramden can finally buy a television.

It was more than half a century ago, in a 1955 episode of "The Honeymooners," that Kramden, the parsimonious bus driver played by Jackie Gleason, told his wife, Alice, that he had not yet bought a new television because "I'm waiting for 3-D."

The wait will soon be over. A full-fledged 3-D television turf war is brewing in the United States, as manufacturers unveil sets capable of 3-D and cable programmers rush to create new channels for them.

Many people are skeptical that consumers will sud-

1 of 2

The 3-D TV sets to be unveiled this week may start at $2,000.

Recent news

The New York Times

Military Is Awash in Data From Drones

BY CHRISTOPHER DREW

HAMPTON, Va. — As the military rushes to place more spy drones over Afghanistan, the remote-controlled planes are producing so much video intelligence that analysts are finding it more and more difficult to keep up.

Air Force drones collected nearly three times as much video over Afghanistan and Iraq last year as in 2007 — about 24 years' worth if watched continuously. That volume is expected to multiply in the coming years as drones are added to the fleet and as some start using multiple cameras to shoot in many directions.

A group of young analysts already watches every second of the footage live as it is streamed to Langley Air

1 of 4

An MQ-1 Predator drone returned from a mission to Bagram Air Base in Afghanistan in 2008. (Master Sgt. Demetrius Lester/U.S. Air Force, via EPA)

Recent news

The New York Times

Giving Electronic Commands With Body Language

BY ASHLEE VANCE
JANUARY 12, 2010

LAS VEGAS — The technology industry is going retro — moving away from remote controls, mice and joysticks to something that arrives without batteries, wires or a user manual.

It's called a hand.

In the coming months, the likes of Microsoft, Hitachi and major PC makers will begin selling devices that will allow people to flip channels on the TV or move documents on a computer monitor with simple hand gestures. The technology, one of the most significant

Microsoft demonstrated its Project Natal, a new way to play Xbox without touching any device, at an electronics show in June. (Casey Rodgers/AP Images for XBox)

Current state of the art

- You just saw examples of current systems.
  - Most 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

Focus of this module

Aspects of a modern vision (recognition) algorithm and applications:

- Feature detection
- Feature description and representation
- Classification
- Face recognition
- Image categorization
- Object detection
- Action recognition

Course logistics

- Webpage: http://hercules.shef.ac.uk/eee/teach/resources/eee422/eee422.html
- Coursework: 40%
- Exam: 60%

E-mail: ling.shao@sheffield.ac.uk
Tel: 0114 222 5841
Office: F160