Today

• Intros
• Computer vision overview
• Course overview
• Image processing

Readings

• Book: Richard Szeliski, Computer Vision: Algorithms and Applications
  – Intro: Ch 1.0

Slides based on UW vision course
What is computer vision?
What is computer vision?

Terminator 2
Every picture tells a story

Goal of computer vision is to write computer programs that can interpret images
Can computers match (or beat) human vision?

Yes and no (but mostly no!)
  • humans are much better at “hard” things
  • computers can be better at “easy” things
The goal of computer vision
To extract “meaning” from pixels

Source: S. Narasimhan
Human perception has its shortcomings…

Challenges or opportunities?

• Images are confusing, but they also reveal the structure of the world through numerous cues
• Our job is to interpret the cues!

Image source: J. Koenderink
Slides: Lana Lazebnik
Depth cues: Linear perspective
Depth cues: Aerial perspective
Depth ordering cues: Occlusion

Source: J. Koenderink
Shape cues: Texture gradient
Position and lighting cues: Cast shadows

Source: J. Koenderink
Grouping cues: Similarity (color, texture, proximity)
Grouping cues: “Common fate”
Inherent ambiguity of the problem

• Many different 3D scenes could have given rise to a particular 2D picture (how large are the objects?)

• Possible solutions
  – Bring in more constraints (more images)
  – Use prior knowledge about the structure of the world

• Need a combination of geometric and statistical methods
Connections to other disciplines

- Artificial Intelligence
- Robotics
- Machine Learning
- Computer Vision
- Cognitive science
- Neuroscience
- Computer Graphics
- Image Processing
Origins of computer vision

Current state of the art

The next slides show some examples of what current vision systems can do
Earth viewers (3D modeling)

Image from Microsoft’s Virtual Earth
(see also: Google Earth)
Application: Augmented Reality

Application areas:
- Film production (the “match move” problem)
- Heads-up display for cars
- Tourism
- Architecture
- Training

Technical challenges:
- Recognition of scene
- Accurate sub-pixel 3-D pose
- Real-time, low latency
What if your photo collection was an entry point into the world, like a wormhole that you could jump through and explore…

The Photosynth Technology Preview is a taste of the newest - and, we hope, most exciting - way to view photos on a computer. Our software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed three-dimensional space, showing you how each one relates to the next.

http://labs.live.com/photosynth/

Based on Photo Tourism technology developed in CSE (UW) by Noah Snavely, Steve Seitz, and Rick Szeliski
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?

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

Sony Cyber-shot® T70 Digital Still Camera
Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics  (TECHNOLOGY from UBC – David Lowe)
“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… “
Face recognition

Who is she?
Vision-based biometrics

“How the Afghan Girl was Identified by Her Iris Patterns” Read the story
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:

- **Lincoln** Microsoft Research
- **Point & Find, Nokia**
google street view
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.howstuffworks.com)
Smart cars

Mobileye

- Vision systems currently in high-end BMW, GM, Volvo models
- By 2010: 70% of car manufacturers.
- Video demo
Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See [Lee's work at CMU](http://www.cs.cmu.edu/~lee/) on clever tricks on using it to create a [multi-touch display](http)!

[Digimask](http): put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNET
Camera tracking a crowd, based on [this work](http)
Vision-based interaction: Xbox Kinect
Mobile visual search: Google Goggles

Google Goggles in Action
Click the icons below to see the different ways Google Goggles can be used.

- Landmark
- Book
- Contact Info
- Artwork
- Places
- Wine
- Logo

![Image of Google Goggles in Action](image-url)
Vision in space

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.

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.
Robotics

NASA’s Mars Spirit Rover

http://www.robocup.org/
Medical imaging

3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT
But then …. Deep Learning!

Convolution nets (cnets) have lead to enormous progress in

- Image classification
- Object detection
- Object localization
- Image captioning
- Video understanding
- …
Amazon Go

“weave the most advanced machine learning, computer vision, and AI”
Autonomous Vehicles: Mobileye

http://www.mobileye.com/our-technology/
But not all is about deep-learning…

Dynamic Fusion…. (CVPR 2015)

Dynamic Fusion
LCI computer vision/robotics

Areas of interest/applications

Mobile robotics
  Curious George
  CanWheel – the intelligent wheelchair
  People and Planet Friendly Home w/ ICICS members

Actions and Activity Understanding

Sports
  Robotics

Scene Parsing

Nearest Neighbors
Curious George Generations
Today - Robot Recognition
Results 2010

Scoring:
8 of 12 specific instances
4 of 8 generic categories

Observations:
Simple environment makes 3D segmentation very effective
Most instances missed due to bad pictures
Categories missed since learning generic mode from web data is hard!
Intelligent Wheelchair
ICICS/TELUS
People & Planet Friendly Home Initiative

October 19, 2012
3.2 ACCESSIBILITY & INTELLIGENT WAY-FINDING IN AN INDOOR-OUTDOOR ENVIRONMENT

Professors: Alan Mackworth and Jim Little
Computer Science

• Objective: Improve way-finding for users of powered wheelchairs, both indoors and outdoors, who would otherwise be confined to manual wheelchairs.
• Develop a smart robotic wheelchair that uses computer vision and other sensors to
  – allow users with physical or cognitive impairments to safely navigate in their homes and other indoor environments.
  – recognize typical user intentions and tasks.
  – allow safe navigation in the home as well as retrieval of a typical object.
Challenges of Clutter
Willow Garage PR2
Charlie goes for a walk
Sports

Understanding Sports Video
Reading Groups

Robuddies - robotics and computer vision


CVRG - Computer Vision Reading Group


Machine Learning Reading Group

This course

http://www.cs.ubc.ca/~little/505.html
General Comments

Prerequisites—*these are essential!*

- Data structures
- A good working knowledge of programming
  
  We’ll use Matlab (or whatever you like! Python…)
- Linear algebra
- Vector calculus

Course does *not* assume prior imaging experience

- computer vision, image processing, graphics, etc.