Lecture 2: CS574

Aug 28, 2014
Review: Admin

• Enrollment
  – Larger room is not available so capacity cannot be expanded
• Grading
  – Assignments 30%, 2 exams 30% ea, participation 10%
  – Tentative Dates: Exam1:~Oct 14, note change from before; Exam2: Dec 4, both during class periods
• Pre-requisites
  – Math: Calculus, analytical geometry, linear algebra, basic probability theory
  – Programming in C/C++, data structures, basic algorithms
• HW0 posted, no due date, no credit, will not be graded
• Course material being posted on http://courses.uscden.net
Class Schedule

• Professor needs to be on travel Sept 9, 11, 16
• No classes on Sept 9 and 11
• Additional classes on Fridays Aug 29 and Sep 5 to make up
  – Time: A.M.;
  – place Studio 100B (in a DEN studio)
  – Will be recorded for those unable to attend makeup in person
• Sept 16, regular class but taught by TA
  – Will cover use of OpenCV library, a key software component
• Regular schedule resumes Sept 18
• Remember next class is tomorrow!
Class Schedule

• Professor needs to be on travel Sept 9, 11, 16
• No classes on Sept 9 and 11
• Additional classes on Fridays Aug 29 and Sep 5 to make up
  – Friday Aug 29, 9:30-10:50 A.M., OHE 100B
  – Attendance not required
  – Will be recorded for those unable to attend makeup in person
• Sept 16, regular class but taught by TA
  – Will cover use of OpenCV library, a key software component
• Regular schedule resumes Sept 18
Today’s Plan

- Challenges of Computer Vision
- Survey of some recent successes
- Start on understanding of Image Formation process
Why is Vision Hard?

- Seems easy to us, no conscious effort is needed by human viewers
- Small variations in human population in ability to see/perceive
- Can’t we recognize objects based on “how they look”?
  - Isn’t a pen (a chair) a pen (chair) because it looks like a pen (chair)?
  - What does a pen (chair) look like?
Same Object Class?
Some Issues: Representation

• What is representation of an object
• Objects of same class can have large variations in shape, size, color, material and other properties
  – Think about every day objects, such as chairs, coffee mugs, telephones…
• What is representation of an action (say throw an object)?
• Same action can be performed in different ways by different actors or even the same actor at different times or in different contexts
Viewpoint Change Examples
Illumination Change Examples
Depth Ambiguity and Occlusions

- World is 3-D, images are 2-D
  - There is an inherent loss of information; process is not truly invertible
  - Our perception of 3-D from single 2-D images must take advantage of some regularities of the natural world

- Occlusion is (almost) ever-present
  - Objects occlude one another
  - Self-occlusion
Multiple Objects in a Scene
How many objects in this image?

What can we say about each?

What can we say about this scene?
Every picture tells a story

We would like to infer such stories, not just object relations
Video Analysis

• Adds difficulties of detecting and tracking moving objects
  – If camera also moves, we need to distinguish between object and background motion
• We want to detect not only objects but also events in the video
• Make inferences about the intentions/plans of actors
• This course will cover motion analysis only briefly or not at all (though motion analysis is the major topic of research in the USC vision group at this time).
Designing Perceptual Systems

• No coherent, complete theory exists

• Some considerations
  – Directly from sensors to semantics?
    • Many successful applications do this, using Machine learning techniques
  – Intermediate layers may help reduce complexity
    • How to specify these layers?
    • Should they be “transparent” (to humans) or learned from data?
  – How to make representations invariant to viewpoint, occlusion, lighting changes etc.
  – How to account for noise and missing data?
  – What can we learn/borrow from biological vision?
  – How much will depth/range sensors help?
Topics to be studied in this class

- **Introduction** (1 week)
  Background, requirements and issues, human vision.

- **Image formation: geometry and photometry** (2 weeks)
  Geometry, brightness, quantization, camera calibration, photometry (brightness and color)

- **Image segmentation** (2 weeks)
  Region segmentation, Edge and line finding

- **Multi-view Geometry** (3 weeks)
  Shape from stereo and motion, feature matching, surface fitting, Active ranging

- **Image classification** (2 weeks)
  Pixel classification, region classification, face detection and identification

- **Object Recognition** (2 weeks)
  Alignment methods, Shape descriptions

- **Motion analysis** (1 week)
  Motion detection and tracking, Inference of human activity from image sequences

- **Applications survey, Review** (1 week)
  Industrial, navigation, mapping, multimedia
Current state of the art

• The next slides show some examples of what current vision systems can do
  – Most taken from class page of Prof. Seitz/Szeliski
Earth viewers (3D modeling)

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

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

Based on Photo Tourism technology
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 new digital cameras now detect faces
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.
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
  - Point & Find, Nokia
Special effects: shape capture

*The Matrix* movies, ESC Entertainment, XYZRGB, NRC
Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com
Smart cars
Slide content courtesy of Amnon Shashua

- http://www.mobileye.com/
- Vision systems currently in many medium to high-end models
Pedestrian Detection
Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in.

Microsoft Kinect (no images here)

Digimask: put your face on a 3D avatar.

“Game turns moviegoers into Human Joysticks”, CNET
Camera tracking a crowd, based on this work.
Microsoft Kinect
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
Robotics

NASA’s Mars Spirit Rover

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

3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT
Next Class

• Read Sections 1.1 and 1.2 of the Forsyth-Ponce Book