Lecture 1
CMSC426
Computer Vision
Course

• Mondays and Wednesdays
  – 3:30-4:45, CSIC 3120

• Grading
  – mid-term 25%
  – Final 35%
  – Homework, 30% (every other week)
  – Quizzes, 10% (surprise, 2)
Instructor

• Ramani Duraiswami
  – (ramani@umiacs.umd.edu) 301 405 6710

• Office hours
  – Thursday 11:00 – 12:30
  – By appointment

• No official TA

• If you have questions meet me
  – E-mail support is deprecated
Campus Policies

• Absences: Campus policy
  – Students claiming a excused absence must apply in writing and furnish documentary support (such as from a health care professional who treated the student) for any assertion that the absence qualifies as an excused absence. The support should explicitly indicate the dates or times the student was incapacitated due to illness. Self-documentation of illness is not itself sufficient support to excuse the absence. An instructor is not under obligation to offer a substitute assignment or to give a student a make-up assessment unless the failure to perform was due to an excused absence. An excused absence for an individual typically does not translate into an extension for team deliverables on a project.
Policies

• Collaborative work
  – Collaboration is encouraged on understanding assignments and formulating strategies for solution
  – Similarly web and other research is encouraged
  – However, the actual writing of solutions must be done independently, without access to any material from colleagues, or other sources.
  – Exams must be done independently
  – Any other collaboration is academic dishonesty, and subject to “academic integrity action.”
Textbook

- *Computer Vision: A Modern Approach*
- Readings from the text
- Recommended strongly that you get access to the book
- Homework problems may be assigned from it
Other Texts

UNDERGRADUATE

GRADUATE/ Specialized Reference
Online Resources

- Computer Vision home page
- CV Online texts
  - http://www.dai.ed.ac.uk/CVonline/
- Many others … any good ones please provide them to me
- Course home page
  http://www.umiacs.umd.edu/~ramani/cmsc426/index.html
  - Lecture Slides
  - Homework
  - Solutions (?)
Why study vision?

- Intellectually interesting
  - How do we figure out what objects are where?
- With cheap digital imaging many applications are now possible
  - Multi megapixel cameras are available everywhere
  - Chips are less than $2 --- every cellphone will have a camera
- Computers and software packages/libraries/architectures make it possible for people to build useful applications with relatively modest investment of time, effort and resources
- Large commercial interest
  - Movies, television, live sports broadcasts, advertisements
  - Post processing of video
  - Surveillance
Vision

- ``to know what is where, by looking.``
  (Marr).
- Where?
- What?
Why is Vision Interesting?

• Psychology
  – ~ 50% of cerebral cortex is for vision.
  – Vision is how we experience the world.

• Engineering
  – Want machines to interact with world.
  – Digital images are everywhere.
Vision is interdisciplinary

• Geometry
• Physics
• Prior Knowledge: The nature of objects in the world (This is the hardest part)
• Computer Science: to process, store and solve problems on images
• Numerical Methods/Scientific computing: to do these efficiently
• Probability/Statistics to deal with noise
What do we use vision for?

• Recognize objects
  – people we know
  – things we own

• Locate objects in space
  – to pick them up

• Track objects in motion
  – catching a baseball
  – avoiding collisions with cars on the road

• Recognize actions
  – walking, running, pushing
  – Emotions, gestures
Vision is

- Deceivingly easy
- Deceptive
- Computationally demanding
- Critical to many applications
Vision is deceivingly easy

• We see effortlessly
  – seeing seems simpler than “thinking”
  – we can all “see” but only select gifted people can solve “hard” problems like chess
  – we use half of our brains for visual perception!

• All “creatures” see
  – frogs “see”
  – birds “see”
  – snakes “see”

but they do not see alike
Vision is deceivingly easy

• The M.I.T. summer vision program
  – summer of 1965
  – point TV camera at stack of blocks
  – locate individual blocks
    • recognize them from small database of blocks
  – describe physical structure of the scene
    • support relationships
• Formally ended in 1985
Vision is deceptive

• Vision is an exceptionally strong sensation
  – vision is immediate
  – we perceive the visual world as external to ourselves,
  – but it is a reconstruction within our brains

• we regard how we see as reflecting the world “as it is;” but human vision is
  – subject to illusions
  – quantitatively imprecise
  – limited to a narrow range of frequencies of radiation
  – passive
Some Illusions
Higher illusions

A Bird
In The
The Bush
Vision is inferential: Light

Checker-shadow illusion:
The squares marked A and B are the same shade of gray.

(http://www-bcs.mit.edu/people/adelson/checkershadow_illusion.html)
Inference

- Humans have to make assumptions about illumination: bump (left) is perceived as hole (right) when upside down.

Illumination direction is unknown. It is assumed to come from above.
Spectral limitations of human vision

• We “see” only a small part of the energy spectrum of sunlight
  – we don’t see ultraviolet or lower frequencies of light
  – we don’t see infrared or higher frequencies of light
  – we see less than .1% of the energy that reaches our eyes

• But objects in the world reflect and emit energy in these and other parts of the spectrum
Non-human vision

- Infrared vision
- Polarization vision
  - navigation for birds
- Ultrasound vision
- X-ray vision!
- RADAR vision
Computationally Demanding

- Millions of pixels
- Hours of video
- Thousands of hours of film
- Millions of cameras
What is Computer Vision?

- Image Understanding (AI, behavior)
- Computer emulation of human vision
- A sensor modality for robotics
- Inverse of Computer Graphics

Computer Vision: [Trucco&Verri’98]

Figure 1.7  The book at a glance: method classes (white boxes), results (grey boxes), their interdependence, and where to find the various topics in this book.
Calibrate video sequences and insert new objects into them

Tom Hanks in Forrest Gump
Detect ground plane in video and introduce pictures on them

Images from: http://www.symah-vision.fr/
Application: Football first-down line

Requires (1) accurate camera registration; (2) a colour model for distinguishing foreground from background
Image Enhancement

• High dynamic range photography [Debevec et al.’97; Mitsunaga & Nayar’99]
  – combine several different exposures

![Figure 6: Sixteen photographs of a church taken at 1-stop increments from 30 sec to __ sec. The sun is directly behind the rightmost stained glass window, making it especially bright. The blue borders seen in some of the image margins are induced by the image registration process.](http://www.nips.cc/Conferences/2004/Tutorials/slides/szeliskiSlides.pdf)
Tracking
Panoramic Mosaics

Slide from: Richard Szeliski’s tutorial on Computer Vision at NIPS, 2004
Medicine/Genetics

• Diagnosis
  – radiology - read X rays, CAT scans
  – pathology - read biopsies

• Remote and tele-medicine

• Virtual reality surgical assistance
  – project images onto head during brain surgery

• Genetics: microarray analysis
Transportation

• Traffic safety and control
  – detection and ticketing of speeding vehicles
  – vehicle counting for flow control
• Robot drivers
  – convoys
• Advanced automobiles
  – autonomous parallel parking
  – road sign detectors and driver alerts
  – collision avoidance
  – smart cruise control
Agriculture

• Safety and quality inspection
  – sorting by size - peaches
  – sorting by shape - potatoes
  – identifying defects - blemishes on fruit, rot in potatoes
  – disease monitoring - chickens

• Robotic farming equipment
  – robotic harvesters - apple pickers, orange pickers
• So vision is interesting and useful
• We are going to study
Related Fields

• Graphics. “Vision is inverse graphics”.
• Visual perception.
• Neuroscience.
• AI
• Learning
• Math: eg., geometry, stochastic processes.
• Optimization.
Tools Needed for Course

• Math
  – Calculus
  – Linear Algebra.

• Computer Science
  – Algorithms
  – Programming, will use Matlab.

• Signal Processing (necessary parts will be covered).
Syllabus

Introduction to computer vision.

Cameras
  Projection and Photometry

Matlab introduction & Linear Algebra Refresher

Intensities – 2.1D vision.
  Boundary detection (edges)
  i. Linear filtering
  ii. Discontinuities (edge detection)
  iii. SNAKEs (finding the bounding contours of objects)

Perceptual grouping

Boundary detection (Regions).

Lightness constancy

Color

Texture

Features (corners, lines).

Perceptual Grouping

Mid term Exam
Syllabus

1. 3D Vision
   1. Stereo
   2. Structure-from-Motion
   3. i. Flow
   4. ii. 2 frame perspective structure-from-motion
   5. iii. Multi-frame affine structure-from-motion

2. Application: Image-based Rendering

3. Shading
   1. Photometric stereo for Lambertian
   2. Recognition as linear combinations of intensities.

4. Recognition
   1. Template matching
   2. Pose determination
   3. Invariance

5. Classification

6. Approaches to Vision

7. FINAL EXAM