CS 534: Computer Vision

Spring 2004
Ahmed Elgammal
Dept of Computer Science
Rutgers University

Outlines

- Vision What and Why ?
- Human vision
- Computer vision
- General computer vision applications
- Course Outlines
- Administrative

What is vision?

- What does it mean to see ?

" The plain man's answer (and Aristotle's too) would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is *" David Marr, Vision 1982*

What is vision?

- 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

Vision is

- Deceivingly easy
- Deceptive
- Computationally demanding
- Critical to many applications

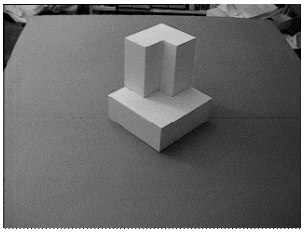
Vision is deceivingly easy

- 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 nearly 70% 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

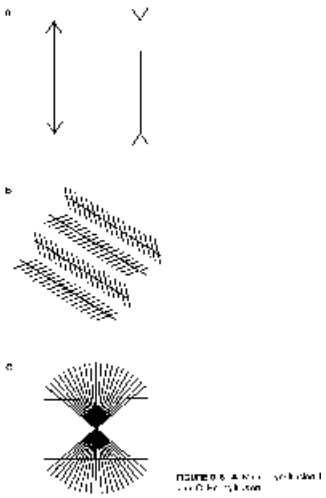


Vision is deceptive

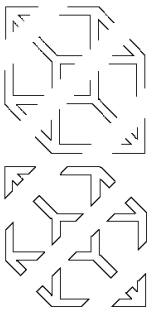
- Vision is an exceptionally strong sensation
  - vision is immediate
  - we perceive the visual world as external to ourselves, but it is a reconstruction within our brains
  - we regard how we see as reflecting the world “as it is;” but human vision is
    - subject to illusions
    - quantitatively imprecise
    - limited to a narrow range of frequencies of radiation
    - passive

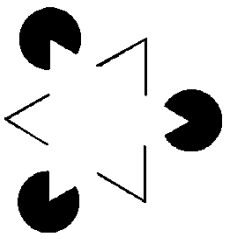
Vision is deceptive

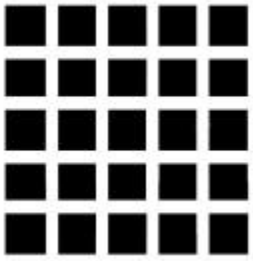
- Human vision is
  - subject to illusions
  - quantitatively imprecise
  - limited to a narrow range of frequencies of radiation
  - passive



More illusion

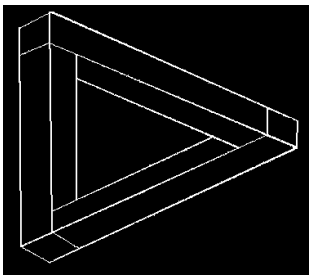


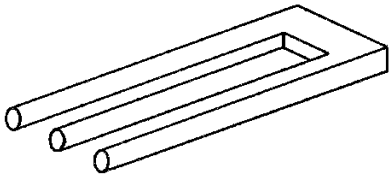




- Subjective contours
- Depth, reversibility, Figure completion

More illusion





- We can see impossible figures

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

Infrared vision

- Vision systems exist that can see reflected and emitted infrared light
  - visual system of the pit viper
  - infrared cameras used for night vision
- Why don't we see the infrared?
  - we would see the blood flow through the capillaries in the eye

Human vision is passive

- It relies on external energy sources (sunlight, light bulbs, fires) providing light that reflects off of objects to our eyes
- Vision systems can be "active" - carry their own energy sources
  - Radars
  - Bat acoustic imaging systems

According to Marr:

- Vision is an information-processing task
- But not just a process
- Our brain must somehow be capable of representing this information.

" vision study … not only the study of how to extract from images the various aspects of the world that are useful to us, but also an inquiry into the nature of the internal representations by which we capture this information and thus make it available as a basis for decisions about our thoughts and actions"

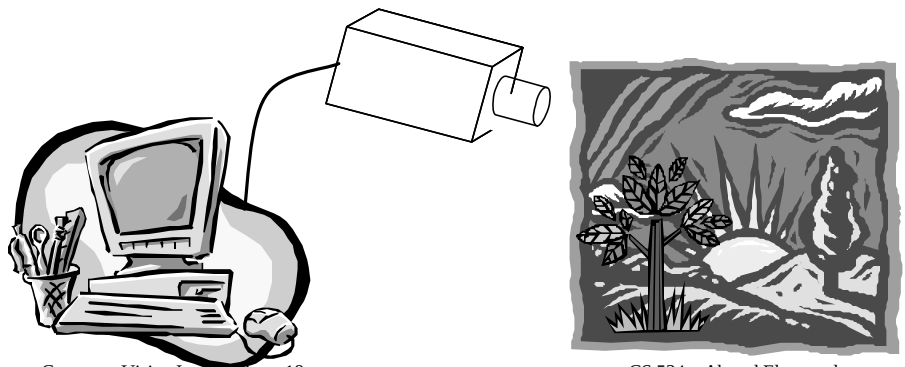
Representation + Processing

"if vision is an information-processing task, then I should be able to make my computer do it, provided that it has sufficient power, memory, and some way of being connected to a home television camera."

We wants to know how to program vision.

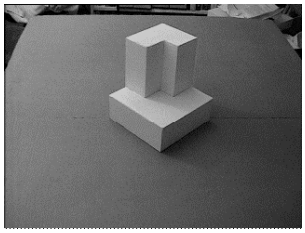
Computer Vision

- Understanding the content of images and videos



Vision is deceivingly easy = Computer Vision is hard

- 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



" The first great revelation was that the problems are difficult. Of course, these days this fact is a commonplace. But in the 1960s almost no one realized that machine vision was difficult. The field had to go through the same experience as the machine translation field did in its fiascoes of the 1950's before it was at least realized that here were some problems that had to be taken seriously." D. Marr, Vision, 1982.

Understanding and Recognition

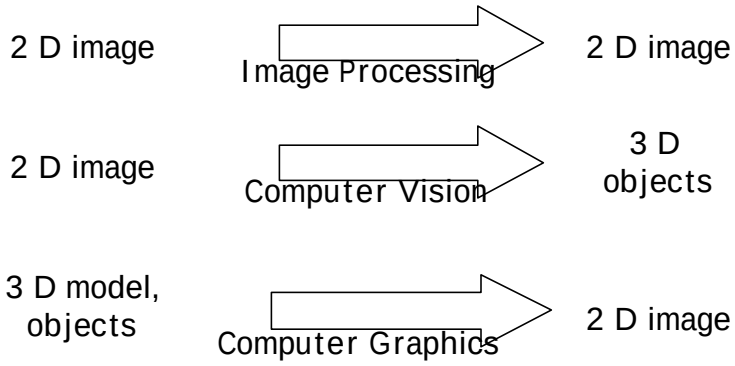
- People draw distinctions between what is seen
  - "Object recognition"
  - This could mean "is this a fish or a bicycle?"
  - It could mean "is this George Washington?"
  - It could mean "is this poisonous or not?"
  - It could mean "is this slippery or not?"
  - It could mean "will this support my weight?"
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties

What are the problems in recognition?

- Which bits of image should be recognized together?
  - Segmentation.
- How can objects be recognized without focusing on detail?
  - Abstraction.
- How can objects with many free parameters be recognized?
  - No popular name, but it's a crucial problem anyhow.
- How do we structure very large model-bases?
  - again, no popular name; abstraction and learning come into this

Why study Computer Vision?

- Images and movies are everywhere
- Fast-growing collection of useful applications
  - building representations of the 3D world from pictures
  - automated surveillance (who's doing what)
  - movie post-processing
  - face finding
- Various deep and attractive scientific mysteries
  - how does object recognition work?
- Greater understanding of human vision



Related Fields: AI, pattern recognition, machine learning, signal processing, neural networks, cognitive vision.

Critical to many applications in

- Manufacturing
- Communications
- Medicine
- Transportation
- Entertainment
- Agriculture
- Defense

Manufacturing

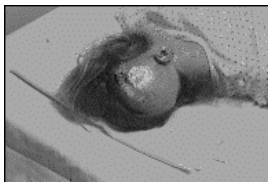
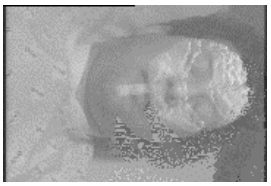
- Visual inspection for quality control
  - during the manufacture of parts in the automotive industry
  - inspection of semiconductors
- Visual control of robots
  - during assembly of parts from pieces
  - during calibration of robot control systems

Communications

- Smart document readers
  - character recognition
  - discrimination of text from graphics and images
  - reading cursive script
  - "language" recognition
- Virtual teleconferencing
- Virtual reality

Medicine

- 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

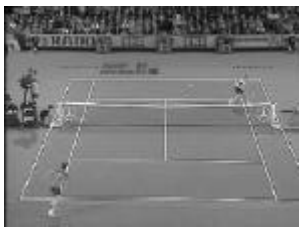
Transportation

- 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

Pittsburgh to San Diego!

Entertainment

- Acquisition of 3D computer models for graphical manipulation
- Control of animation through vision
- Indexing tools for video databases

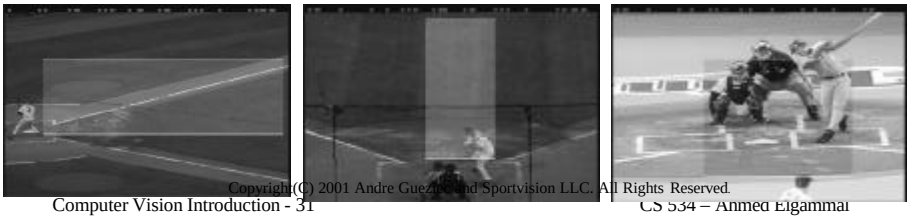



- Detect ground plan in video and introduce pictures on them

Images and videos from: SYMAH VISION, Easily Virtual www.symah-vision.fr

Tracking Baseball Pitches for Broadcast Television

- K Zone: system developed by Sportvision for ESPN.
- The system is used by ESPN for its Major League Baseball broadcast.

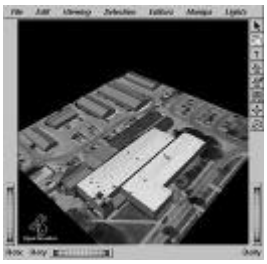


Agriculture

- 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

Defense

- Automatic target recognition systems
  - cruise missiles
  - air to surface "smart" missiles
- Reconnaissance
  - monitoring strategic sites
- Simulation
  - acquisition of terrain models from imagery
  - model acquisition of buildings, roads, etc.



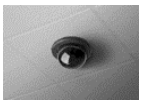
Looking at People

- Human detection
- Human tracking
- Human recognition and biometrics
  - Face recognition
  - Gait recognition
  - Iris, etc.
- Gesture recognition
- Facial expression recognition
- Activity recognition

Applications

- Human Computer Interaction
  - Keyboard and mouse are restrictive
- Driver Assistance, Autonomous driving
  - Pedestrian detection
  - Traffic signs detection/recognition
  - Lane detection
  - Occupant detection
- Motion Capture
- Video editing, archival and retrieval.
- Surveillance: security, safety, resource mangement

Visual Surveillance




Consider a visual surveillance system

State of the art: archive huge volumes of video for eventual off-line human inspection

Goal : Automatic understanding of events happening in the site.

- Efficient archiving
- Automatic Annotation
- Direct human attention
- Reduce bandwidth required for video transmission and storage.

Course Outline

- Part I: The Physics of Imaging
  Image formation and image models: Cameras, light, color
- Part II: Early Vision in One Image
  Edges and texture
- Part III: Early Vision in Multiple Images
  Stereopsis, structure from motion
- Part IV: Mid-Level Vision
  Finding coherent structure in images and movies: Segmentation, Tracking
- Part V: High Level Vision (Geometry)
  The relations between object geometry and image geometry: Model-based vision
- Part VI: High Level Vision (Probabilistic)
  Using classifiers and probability to recognize objects

Textbook

"Computer Vision: A Modern Approach" By David Forsyth and Jean Ponce, Prentice Hall 2002
ISBN 0-13-085198-1

- Other reading materials will be provided.

- Other useful references:
  - L. Shapiro, G. C. Stockman “Computer Vision”, Prentice Hall.
  - O. Faugeras “Three-Dimensional Computer Vision: A Geometric View Point”, MIT press.
  - Horn “Robot Vision”, MIT press.
  - D. Marr “Vision”, Freeman.

- **Course Load**
  - Homework assignments: (20%) 3-4 assignments, which might contain some Matlab programming.
  - Term Project: (40%) Details below.
  - Exams: Midterm (15%) and Final (25%).

- **Class Project:**

  Students are expected to work on a class research project throughout the semester to explore a recent computer vision research topic. Students are to choose their own projects and are encouraged to find a project related to their own research. The project ideas are expected to be innovative, experimental and feasible to be done within the semester time frames.

Project time line

- *Week 4 ~ Feb 9th - Project Pre-proposal (abstract):* One page abstract about the project you intend to do
- *Week 4 ~ Feb 11th:* Abstract Feedback.
- *Week 6 ~ Feb 25th - Project Proposal:* detailed project proposal including problem definition, brief description of related work, approach to be taken, expectation, specification of evaluation data and evaluation methods. (10 pages maximum).
- *Week 7 ~ March 1st* - Proposal feedback.
- *Week 11 ~ March 31st - Midterm progress report / proposal revision:* Short report describing the progress you achieved and the problems you are facing. It's also a chance to revise your proposal if you need to do so.
- *Week 15 ~ April 26th - Project Final Report: a research paper (format to be provided) describing your work, results, and conclusions.*
- *Final Week: Project Presentations.*

Useful Computer Vision Resources:

- Computer Vision Home Page (CMU): http://www-cgi.cs.cmu.edu/afs/cs/project/cil/ftp/html/vision.html
- CV online: http://www.dai.ed.ac.uk/CVonline/
- Keith Price Annotated Computer Vision Bibliography at USC: http://iris.usc.edu/Vision-Notes/bibliography/contents.html
- Open CV: Intel Open Source CV library. http://www.intel.com/research/mrl/research/opencv/
- Matlab Image Processing toolbox: http://www.mathworks.com/access/helpdesk/help/toolbox/images/images.shtml

Sources

- D. Marr "Vision", Freeman.
- Slides by Prof. Larry Davis at UMD