Introduction

Aykut Erdem
Dept. of Computer Engineering
Hacettepe University
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. [Marr, 1982]
Why is vision hard?

• A typical image includes many objects organized in many different configurations.
• Vision requires solving ill-posed problems.
• Images are both complicated and highly ambiguous.
• Same object can generate very different images.
• Different objects can generate similar images.

Figures: Steven Pinker, How the Mind Works, 1997
Challenges: Illumination

Figure: J. Koenderink

Slide credit: L. Fei-Fei
Challenges: viewpoint variation

Michelangelo 1475-1564

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba
Challenges: Scale

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba
Challenges: Deformations

Xu, Beihong 1943

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba
Challenges: Occlusion

The Blank Check, by René Magritte

Slide credit: L. Fei-Fei, R. Fergus and A. Torralba
Challenges: background clutter
Challenges: Motion
Challenges: Some things have strong variations in appearance
Challenges: local ambiguity
Why does vision appear easy to humans?

- Our brains are specialized to do vision.
- ~50% of the cortex in a human brain is devoted for visual processing
  (cf. motor control ~20-30%, language ~10-20%)

[David Heeger, 2006] [Felleman and van Essen, 1991]
Visual Pathways

One pathway goes from the eyes to the striate cortex and one from the eyes to each superior colliculus. The distortion in the brain mapping in the striate cortex reflects the emphasis given to analysing the central region of the retinal image, so much so that the tiny representation of the child's hand can hardly be seen in this figure. See 1.7 for details.

Surrounds the question of what cells in the lateral geniculate nuclei do. They receive inputs not only from the eyes but also from other sense organs, so some think that they might be involved in filtering messages from the eyes according to what is happening in other senses. The lateral geniculate nuclei also receive a lot of fibers sending messages from the cortex. Hence there is an intriguing two-way up-down traffic going on in this visual pathway and we discuss its possible functions in later chapters.

Before we go on to discuss the way fiber terminations are laid out in the striate cortex, note that the optic nerves provide visual information to two other structures shown in 1.6—the left and right halves of the superior colliculus. This is a
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!
Measuring light vs. measuring scene properties

We perceive two squares, one on top of each other.

Slide credit: B. Freeman and A. Torralba
Measuring light vs. measuring scene properties

by Roger Shepard ("Turning the Tables")

Depth processing is automatic, and we can not shut it down...

Slide credit: B. Freeman and A. Torralba
Measuring light vs. measuring scene properties
Measuring light vs. measuring scene properties

Edward H. Adelson

Slide credit: B. Freeman and A. Torralba
Measuring light vs. measuring scene properties

(c) 2006 Walt Anthony

Slide credit: B. Freeman and A. Torralba
Measuring light vs. measuring scene properties

Figure: Richard L. Gregory, Phil. Trans. R. Soc. B, 2005
Illusory contours

- Continuity of discontinuities

Figures: Steven Pinker, How the Mind Works, 1999
Assumptions can be wrong!

Ames room

Slide credit: B. Freeman and A. Torralba
Some things know that you have eyes


Slide credit: B. Freeman and A. Torralba
# Computer Vision

What we see

What a computer sees

<table>
<thead>
<tr>
<th>0</th>
<th>3</th>
<th>2</th>
<th>5</th>
<th>4</th>
<th>7</th>
<th>6</th>
<th>9</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Slide credit: S. Narasimhan
Related Disciplines

- Optics/imaging
- Image processing
- Computer graphics
- Visual perception
- Visual neuroscience
- Machine learning
- Robotics
THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".
Origins of computer vision

Vision as an information processing task, David Marr

“Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information”

- 3 levels of understanding:
  1. Computational theory
     - What is computed? Why it is computed?
  2. Representation and Algorithm
     - How it is computed?
     - Input, Output, Transformation
  3. Physical Realization
     - Hardware
Marr’s observation: Studying vision at 3 levels
Marr’s observation: Studying vision at 3 levels
Reading Assignment

• David Marr. Vision. Chapter 1
  – Prepare a short (about ~2, 3 pages) summary.
  – Be careful to use your own words!
  – Due next week (Feb 26). Print out and bring to the class.

• Some other additional readings are also available on the course webpage!
Course Information

• Instructor: Aykut ERDEM
  aykut@cs.hacettepe.edu.tr
  Office: 111
  Tel: 297 7500, 146

• Lectures: Wednesdays 13:00-15:45_D7

• Office Hour: By appointment

• Prerequisites: Good math background (Calculus, Linear Algebra, Statistical Methods) and programming skills (MATLAB, C/C++)

• Course Website: http://web.cs.hacettepe.edu.tr/~aykut/classes/spring2014/bil719/
Communication

• We will be using Piazza for course related discussions and announcements.

• Please enroll it by following the link https://piazza.com/hacettepe.edu.tr/spring2014/bil719
Reference Books

Supplemental Reading Material

• Lecture notes and handouts
• Papers and journal articles
Related Conferences

• IEEE International Conference on Computer Vision (ICCV)
• European Conference on Computer Vision (ECCV)
• IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
• British Machine Vision Conference (BMVC)
• Asian Conference of Computer Vision (ACCV)
• IEEE International Conference on Pattern Recognition (ICPR)
• Advances in Neural Information Processing Systems (NIPS)
Related Journals

• International Journal of Computer Vision (IJCV)
• IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI)
• Computer Vision and Image Understanding (CVIU)
• Pattern Recognition (PR) Journal of
• Mathematical Imaging and Vision (JMIV)
• Image and Vision Computing (IMAVIS)
Grading Policy

• 20% Programming Assignments
• 20% Quizzes
• 20% Paper Presentations/Class participation
• 40% Project and final term paper
Programming Assignments

- A total of 4 programming assignments
- Related to the topics covered in the class.
- Each one will involve
  - implementing an algorithm (usually in Matlab),
  - carrying out a set of experiments, and
  - writing up a report

- All assignments have to be done individually, unless stated otherwise.
Paper presentations and Quizzes

• The students will be required to present at least one research paper either of their choice or from the suggested reading list.

• These papers should be read by every student as the quizzes about the presented papers will be given on the weeks of the presentations.
Project

• The students taking the course will be required to do a project in computer vision, which should be done individually.

• This project may be
  – An original implementation of a new or published study
  – A detailed empirical evaluation of two or more related methods not covered in the class

• March 12: Project proposals
• April 23: Project progress reports
• June 04: Project final reports
Course Overview

• (1 week) Introduction, image formation
• (2 weeks) Filtering and texture
• (3 weeks) Grouping and fitting
• (3 weeks) Multiple geometry and stereo
• (3 weeks) Recognition
• (1 week) Motion
Image formation

• Color
• Pinhole cameras
• Lenses

Filtering and texture

Slide credit: K. Grauman
Grouping and fitting

Alignment

Segmentation

Fitting: Least squares
Hough transform
RANSAC

Slide credit: S. Narasimhan
Multiple geometry and stereo

Stereo

Epipolar geometry

Tomasi & Kanade (1993)

Affine structure from motion

Projective structure from motion

Slide credit: S. Narasimhan
Recognition

Patch description and matching

Clustering and visual vocabularies

Bag-of-features models

Classification

Slide credit: S. Narasimhan, D. Lowe, L. Fei-Fei
Motion