COS 429: Computer Vision

Instructor: Szymon Rusinkiewicz
smr@cs.princeton.edu

TAs: Thiago Pereira, Yiming Liu
{tpereira, yimingl}@cs.princeton.edu

Web page:
http://www.cs.princeton.edu/~cos429/
What is Computer Vision?

• Input: images or video
• Output: description of the world
• But also: measuring, classifying, interpreting visual information
Low-Level or “Early” Vision

• Considers local properties of an image

“There’s an edge!”
Mid-Level Vision

• Grouping and segmentation

“There’s an object and a background!”
High-Level Vision

- Recognition
- Classification

“It’s a chair! It’s in a room!”
Big Question #1: Who Cares?

• Applications of computer vision
  – In AI: vision serves as the “input stage”
  – In robotics: localization, mapping, obstacle detection
  – In medicine: understanding human vision
  – In engineering: creating models of the world
Consumer Applications

Face-hunting cameras boost Nikon

Japanese camera maker Nikon has tripled its profits on the back of strong sales of digital cameras that automatically focus on human faces.
Vision and Other Fields

- Cognitive Psychology
- Artificial Intelligence
- Signal Processing
- Computer Vision
- Pattern Analysis
- Computer Graphics
- Metrology
Big Question #2: Does It Work?

• Situation much the same as AI:
  – Some fundamental algorithms
  – Large collection of hacks / heuristics
  – Continuous progress: more success than you might think!

• Vision is hard!
  – Especially at high level, physiology unknown
  – Requires integrating many different methods
  – Requires reasoning and understanding: “AI completeness”
Computer and Human Vision

- Emulating effects of human vision
- Understanding physiology of human vision
- Analogues of human vision at low, mid, and high levels
Image Formation

- Human: lens forms image on retina, sensors (rods and cones) respond to light

- Computer: lens system forms image, sensors (CCD, CMOS) respond to light
Low-Level Vision

Hubel
Low-Level Vision

- Retinal ganglion cells
- Lateral Geniculate Nucleus – visual adaptation?
- Primary Visual Cortex
  - Simple cells: orientational sensitivity
  - Complex cells: directional sensitivity
- Further processing
  - Temporal cortex: what is the object?
  - Parietal cortex: where is the object? How do I get it?
Low-Level Vision

- Net effect: low-level human vision can be (partially) modeled as a set of *multiresolution, oriented* filters
Low-Level Depth Cues

- Focus
- Vergence
- Stereo
- Not as important as popularly believed
Low-Level Computer Vision

• Filters and filter banks
  – Implemented via convolution
  – Detection of edges, corners, and other local features
  – Can include multiple orientations
  – Can include multiple scales: “filter pyramids”

• Applications
  – First stage of segmentation
  – Texture recognition / classification
  – Texture synthesis
Texture Analysis / Synthesis

Multiresolution Oriented Filter Bank

Original Image

Image Pyramid
Texture Analysis / Synthesis

Original Texture

Synthesized Texture

[Heeger and Bergen]
Low-Level Computer Vision

• Optical flow
  – Detecting frame-to-frame motion
  – Local operator: gradients over space and time

• Applications
  – First stage of tracking
Optical Flow

Image #1 → Optical Flow Field → Image #2
Low-Level Computer Vision

• **Shape from X**
  – Stereo
  – Motion
  – Shading
  – Texture foreshortening
3D Reconstruction

Tomasi + Kanade

Debevec, Taylor, Malik

Forsyth et al.

Phigin et al.
Mid-Level Vision

- Physiology unclear, but recent experiments with FMRI
- Observations by Gestalt psychologists
  - Proximity
  - Similarity
  - Common fate
  - Common region
  - Parallelism
  - Closure
  - Symmetry
  - Continuity
  - Familiar configuration

Wertheimer
Grouping Cues

- Not grouped
- Proximity
- Similarity
- Similarity
- Common Fate
- Common Region
- Parallelism
- Symmetry
- Continuity
- Closure
Grouping Cues
Grouping Cues
Grouping Cues
Mid-Level Computer Vision

• Techniques
  – Clustering based on similarity
  – Limited work on other principles

• Applications
  – Segmentation / grouping
  – Tracking
Snakes: Active Contours

Contour Evolution for Segmenting an Artery
Histograms
Expectation Maximization (EM)

Color Segmentation
Bayesian Methods

• Prior probability
  – Expected distribution of models

• Conditional probability $P(A|B)$
  – Probability of observation $A$ given model $B$
Bayesian Methods

- **Prior probability**
  - Expected distribution of models

- **Conditional probability** $P(A \mid B)$
  - Probability of observation $A$ given model $B$

- **Bayes’s Rule**
  $$P(B \mid A) = \frac{P(A \mid B) \cdot P(B)}{P(A)}$$
  - Probability of model $B$ given observation $A$

Thomas Bayes (c. 1702-1761)
Bayesian Methods

\[ P(X \mid a) \]

\[ P(X \mid b) \]

\# black pixels

\# black pixels
High-Level Vision

- Human mechanisms: ???
High-Level Vision

• Computational mechanisms
  – Bayesian networks
  – Templates
  – Linear subspace methods
  – Kinematic models
Template-Based Methods

Cootes et al.
Linear Subspaces
Principal Components Analysis (PCA)

Data

New Basis Vectors

Kirby et al.
Kinematic Models

• Optical Flow/Feature tracking: no constraints

• Layered Motion: rigid constraints

• Articulated: kinematic chain constraints

• Nonrigid: implicit / learned constraints
Real-world Applications

Osuna et al:

Figure 6. In this picture, circles represent faces, and squares represent non-faces.
Real-world Applications

Osuna et al:

Figure 5. Results from our Face Detection system.
Course Outline

- Image formation and capture
- Filtering and feature detection
- Motion estimation
- Segmentation and clustering
- Recognition and classification
- 3D shape acquisition
3D Scanning
Image-Based Modeling and Rendering

Debevec et al.

Manex
Reassembling the Thera Wall Paintings

- Shattered by earthquakes, volcanic eruption
Reassembling the Frescoes
Example Fragments
Even More Fragments...
And Still More Fragments
3-D Acquisition
Matching Results
Course Mechanics

• 70%: 4 written / programming assignments
  – Individual: all submitted work must be your own
  – 3 free late days

• 30%: Final project
  – Small groups – 2-3 people
  – Presentation / demo in January
  – Writeup due on Dean’s date
Course Mechanics

• Recommended book:
  *Computer Vision: Algorithms and Applications*
  Richard Szeliski
  – Also available online

• Assigned papers / other readings
MATLAB

• Some of the assignments use MATLAB

• School of Engineering is running a short course, Monday Sep 19 – Tuesday Sep 20

• Should we also do a precept next week?
Q&A

• We will use piazza for Q&A. Please direct all non-private questions there.

• Feel free to answer each others’ questions (we will monitor and endorse students’ answers) but keep in mind collaboration policy.
COS 429: Computer Vision

Instructor: Szymon Rusinkiewicz
smr@cs.princeton.edu

TAs: Thiago Pereira, Yiming Liu
{tpereira, yimingl}@cs.princeton.edu

Web page:
http://www.cs.princeton.edu/~cos429/