CS651: Computer Vision

Thanks to Chris Bregler & Szymon Rusinkiewicz
what is human vision?

- input: visual stimuli (spatial/temporal)
- output: model of world
- goal: measure, classifying, interpret visual information
what is computer vision?

- computational methods and algorithms for simulating human vision
- input: images and video
- output: model of world
- goal: measure, classifying, interpret visual information
who cares?

- many diverse applications
  - AI: vision serves as the “input stage”
  - medicine: understanding human vision/augmented surgery
  - engineering: model extraction
  - graphics: content generation/authoring
applications
low-level or “early” vision

- local properties of an image

What do you see?
mid-level vision

- grouping and segmentation

What do you see?
What do you see?

high-level vision

- recognition
vision and other fields

- Cognitive Psychology
- Signal Processing
- Artificial Intelligence
- Computer Vision
- Computer Graphics
- Pattern Analysis
- Metrology
does it work?

- similar situation as AI:
  - some fundamental algorithms (we will learn them)
  - many hacks/heuristics (we will implement them)
- computer vision is hard!
  - especially at higher level, physiology unknown
  - requires integrating many different methods
  - requires reasoning and understanding: “AI-completeness”
does it work?

Completely Automated Public Turing test to tell Computers and Humans Apart

[von Ahn]
detour: RECAPTCHA

melsenly seems
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 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

Original Image

Multiresolution Oriented Filter Bank

Image Pyramid
texture analysis / synthesis
low-level computer vision

- optical flow
  - detecting frame-to-frame motion
  - local operator: look for gradients
- 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.

Debevec, Taylor, Malik

Phigín et al.
mid-level vision

- physiology unclear
- observations by Gestalt psychologists
  - proximity
  - similarity
  - common fate
  - common region
  - parallelism
  - closure
  - symmetry
  - continuity
  - familiar configuration

Wertheimer
emergence

holistic aspect of perception
multistability

multiple valid interpretations
invariance

recognition independent of rotation, scale, deformations, shading, depiction
grouping cues

- Not grouped
- Proximity
- Similarity
- Similarity
- Common Fate
- Common Region
- Parallelism
- Symmetry
- Continuity
- Closure
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
- sampling probability $P(A|B)$
  - probability of observation $A$ given model $B$
Bayesian methods

- prior probability
- expected distribution of models
- sampling probability $P(A|B)$
- probability of observation $A$ given model $B$
- Bayes’s rule:
  
  $P(B|A) = \frac{P(A|B) \ P(B)}{P(A)}$

- probability of model $B$ given observation $A$
Bayesian methods

\[ P(X|a) \]

\[ P(X|b) \]

# black pixels

# black pixels
high-level vision

- human mechanisms ??
- computational mechanisms:
  - Bayesian networks
  - templates
  - linear subspace models
  - kinematic models
  - hierarchical probabilistic models
template-based methods

Cootes et al.
linear subspaces
Principal Components Analysis

input images

PCA

new basis images
kinematic models

- Optical Flow/Feature tracking: no constraints

- Layered Motion: rigid constraints

- Articulated: kinematic chain constraints

- Nonrigid: implicit / learned constraints
course outline

- image formation and capture
- filtering and feature detection
- segmentation and clustering
- PCA
- motion estimation
- shape acquisition
course mechanics

✦ 70%: 4 written / programming assignments
✦ 30%: final group project
course mechanics

- assignments:
  - edge and corner detection (due 9/18)
  - face detector (10/04)
  - optical flow (10/23)
  - stereo (11/06)
  - group project proposals due (11/08)
  - group project presentations & reports (TBD)
course mechanics

🔹 book:
  - *Introductory Techniques for 3-D Computer Vision*
    Emanuele Trucco and Alessandro Verri

🔹 research papers
course mechanics

✦ office hours:
- Olsson 212 WF 3-5 and by appointment

✦ course mailing list:
- cs651-jdl@cs.virginia.edu

✦ course web page:

✦ lectures will be combination of slides and chalk
MATLAB

✦ some of the assignments use MATLAB
✦ available from ITC:
  - http://www.itc.virginia.edu/research/matlab/
picture time