CS510: Image Computation

Bruce A. Draper
Jan 18, 2017

Class Goals

- Prepare you to do research in computer vision
  - Provide "big picture" (comparison to humans)
  - Give you experience reading papers
  - Familiarize you with SOA methods/algorithms
  - Familiarize you with open problems
  - Familiarize you with experimental methodology

Class Project

- Identify objects in videos
  - Moving objects: potential actors
  - Still objects: potential related objects
- Identify actions in videos
  - E.g. walk, run, stand, pick up, dig, etc.
- Describe short videos in simple English sentences
  - $S \rightarrow N \ V \ [DO] \ [IO] \ [PP]$
- Note: this was SOA in 2012

Class Project (II)

- I/O (PA1)
- Attention
  - Still object attention (PA2) [DoG]
  - Moving object attention (PA3) [ViBE / MoG]
  - Tracking (PA3) [Kalman / Particle / MOSSE]
- Classification
  - Object identification (PA4) [CNN]
  - Action identification (PA5) [ANN / BoW]
- Sentence Generation (PA6)
  - Information fusion [Markov]
  - Relations [Grammars / Schemas]
- Expertise
  - PCA, Manifolds (PA7)

Administrivia

- Workload:
  - Programming assignments (40%)
    - At least 6 (hopefully 7)
    - Roughly every two weeks
    - Large assignments – teams of 2 or 3
    - Assignments build on each other
      - A single project
  - Tests (40%)
    - 3 tests
    - All in-class
    - No "exam"
Administrivia (II)

- Readings
  - Ungraded
  - Emphasized on tests
  - Make connections to project presentations
- Project presentations (20%)
  - Details may vary by team (size, etc.)
  - Relate projects to readings, class lectures
- Class web site: http://www.cs.colostate.edu/~cs510
  - Progress page has lectures & reading assignments
  - Assignments page has assignments
  - No Canvas page for this class

Misconceptions about Vision

- Vision is passive
  - Bad metaphor: video recorder
  - Better metaphor: blind-folded touch
- Vision is 3D
  - Bad metaphor: 3D sensor
    - E.g. Kinect, ladar range finder, etc.
  - Better metaphor: theater with backdrop

Illusions

Kanizsa Triangle

Why Illusions?

- Collectively, what do these illusions tell us?
  - Vision (sensing) is active
  - We select what we attend to
  - We construct percepts from what we sense
- What is the lesson for artificial vision?
  - Must actively seek information
  - Must be “like people”

Input: The Eye(s)

Start at the beginning:
- Lens focuses light
- Iris serves as aperture
- Retina contains receptors
- Optic nerve transmits to brain

Gaze, lens, iris are controlled by muscles under the control of the brain
Retina as Processor

- Five cell types:
  - receptor (rod/cones)
  - horizontal
  - bipolar cells
  - amacrine cells
  - ganglion cells

- Its inside out!
- Blind spot where optic nerve passes through retina

Fields of View & Stereo

- Right hemisphere receives the left visual field from both eyes
  - And vice-versa
  - Splitting the field of view supports disparity computations

- High resolution in fovea, lower elsewhere
  - Fovea is ±2° (thumbnail at arms length)

Projections (LGN & S.C.)


Primary Visual Cortex (V1)

- First cortical visual area
  - Columnar (like all cortex)
  - Retinotopically mapped
  - Ocular dominance columns
  - Edges (Gabor filters), color, disparity & motion maps
  - Connects to other retinotopic areas (V2, V3, MT)

Proof of Retinotopic Mapping

Pattern flashed (like a strobe) in front of monkey injected with sugar dye
Left primary visual cortex of the same monkey

V1 Connections

V1 is the starting point of cortical visual processing.
Dorsal projections lead to somatosensory and motor control areas
Ventral projections lead toward associative memories
Anatomical Maps of Visual Cortex

1983 Version

1990 Version

Colorado State University

Visualizing Two Subsystems

D. Milner & M. Goodale, The Visual Brain in Action, p. 22

A Model of Human Vision

Associative Memories

Information Shunting

Attention Shifting

Spatial Properties Processing

Object Properties Processing

Visual Buffer

Attention Window

S. Kosslyn, 2006

(The same model of) Human Vision

Attention Shifting

Spatial Properties Processing

LT Associative Memory

Information Shunting

Object Properties Processing

Visual Buffer

Attention Window

ST Associative Memory

Colorado State University