Human Visual System

CSE 152
Lecture 19

Final Exam

- Closed book
- No Calculator
- One cheat sheet
  - Single piece of paper, handwritten, no photocopying, no physical cut & paste.
  - You can start with your sheet from the midterm.
- What to study
  - Basically material presented in class, and supporting material from text
  - If it was in the reading, but NEVER mentioned in class, it is very unlikely to be on the exam
- Whole course, weighted toward second half
- Question style:
  - Short answer
  - Some longer problems to be worked out.
Training and Loss Function

- Training data: \( \{(x_i, y_i): 1 \leq i \leq n\} \)

\[
x_i \rightarrow f(x, w) \rightarrow \hat{y}_i \rightarrow L(y_i, \hat{y}_i) \rightarrow y_i
\]

- Total Loss: \( \sum_{i=1}^{n} L(f(x_i; w), \hat{y}) \)

- Training: Find \( w \) that minimizes the total loss.

The loss function

- The loss function is really important. It’s how we compare the network output to the training labels.

- Common loss functions:
  - Regression problems:
    - Distance: \( L(y, \hat{y}) = ||y - \hat{y}||^p \), usually \( p = 1 \) or \( 2 \)
  - Classification: Softmax + cross entropy
    - Softmax: \( \hat{y}_i(z) = \frac{e^{z_i}}{\sum_{j=1}^{c} e^{z_j}} \)
    - Cross entropy between \( y \) and \( \hat{y} \) is \( H(y, \hat{y}) = \sum_{i=1}^{n} y_i \log \frac{1}{\hat{y}_i} = -\sum_{i=1}^{n} y_i \log \hat{y}_i \)
    - where: \( y \) is a vector with one 1 and the rest 0’s.
    - \( \hat{y} \) is a vector with positive floats that sum to 1
Many kernels yielding many features!

Color images
3 channels
- Convolution kernel is 4-D: For each output channel, kernel goes across input dimensions and channels
- Size = width x height x # input channels x # output channels

Conv layer features
9 channels

[Diagram of convolution layer structure]

VGG16 [Simonyan and Zisserman, 2014]
Why should we study Biological Vision in a Computer Vision Course?

- Computer vision systems mostly operate under the same environmental constraints as biological systems
- Both systems exploit the same cues for interpreting scenes
- There is immense diversity in biological vision systems which can serve as models for machine systems.
- It’s easier to copy (or be inspired by) a working system than invent it from scratch.
Kepler

Kepler, 1604
Eye as an optical instrument
Image is inverted on retina
First such experiment by Scheiner, 1625

Ways to study human vision

1. Physiological
2. Phenomenological/Psychophysical
3. Cellular recordings
4. Functional MRI
5. Computational modeling
Physiological level

What does this do?

Can we readily understand whole from understanding pieces?
Ways to study human vision

1. Physiologically
2. Phenomenological/Psychophysical
3. Cellular recordings
4. Functional MRI
5. Computational modelling

Psychophysical Testing of Subjects
Example:
Show gratings with different spatial frequencies
Move them in different directions and measure human perception

2) Spatial Frequency (cycles/deg)

Gradients/Motion
Occlusion provides a different organization
Ways to study human vision

1. Physiologically
2. Phenomenological/Psychophysical
3. Cellular recordings
4. Functional MRI
5. Computational modeling

Single Cell Recordings
Activation in the right fusiform gyrus.
[ Tarr, Cheng 2003]

Ways to study human vision

1. Physiologically
2. Phenomenological/Psychophysical
3. Cellular recordings
4. Functional MRI
5. Computational modeling
Computational Modeling

What is being computed and why?

Structure of the eye
Focus

The range of lighting

<table>
<thead>
<tr>
<th>Lighting Condition</th>
<th>Lux</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct sun</td>
<td>100'000</td>
</tr>
<tr>
<td>Sunny day</td>
<td>50'000</td>
</tr>
<tr>
<td>Cloudy day</td>
<td>5'000</td>
</tr>
<tr>
<td>Office</td>
<td>400</td>
</tr>
<tr>
<td>Home lighting</td>
<td>10</td>
</tr>
<tr>
<td>Street lamps</td>
<td>1</td>
</tr>
<tr>
<td>Full moon</td>
<td>0.1</td>
</tr>
<tr>
<td>Quarter moon</td>
<td>0.01</td>
</tr>
<tr>
<td>Clear moonless night</td>
<td>0.001</td>
</tr>
<tr>
<td>Cloudy moonless night</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

1 lux = 1 lum/m²
Rods and cones

Rods
- Cylindrical
- $10^8$
- Periphery
- Broad Spectral Response

Cones
- Conical
- $5 \times 10^6$
- Fovea
- Three types (color)

3.4 THE SPATIAL MOSAIC OF THE HUMAN CONES. Cross sections of the human retina at the level of the inner segments showing (A) cones in the fovea, and (B) cones in the periphery. Note the size difference (scale bar = 10 μm), and that, as the separation between cones grows, the rod receptors fill in the spaces. (C) Cone density plotted as a function of distance from the center of the fovea for seven human retinas; cone density decreases with distance from the fovea. Source: Curcio et al., 1990.
Three types of cones: R, G, B

**Response of k’th cone** = \( \int \rho_k(\lambda)E(\lambda)d\lambda \)

There are three types of cones:
- **S**: Short wave lengths (Blue)
- **M**: Mid wave lengths (Green)
- **L**: Long wave lengths (Red)

- Three attributes to a color
- Three numbers to describe a color

---

Distribution of Rods & Cones

3.1 THE DISTRIBUTION OF ROD AND CONE PHOTORECEPTORS across the human retina. (A) Degrees of visual angle relative to the position of the fovea for the left eye; the position of the blind spot is also shown. (B) The cone receptors are concentrated in the fovea. The rod photoreceptors are absent from the fovea and reach their highest density between 10 and 20 degrees peripheral to the fovea. No photoreceptors are present in the blind spot.
Retinal Neuron

5.2 RETINAL NEURONS have many different shapes and sizes. (A) The cell body of a bipolar cell resides in the outer nuclear layer. Its dendrites make contact with the photoreceptors and horizontal cells and its axon carries the output of the bipolar cell to the inner plexiform layer (see Figure 5.1), where it contacts the dendritic field of a ganglion cell. (B) The retinal ganglion cell bodies reside in the ganglion cell layer of the retina (see Figure 5.1). The axons of the retinal ganglion cells comprise the optic nerve.

Several types of retinal ganglion cells can be distinguished based on the properties of their dendritic fields, their interconnections, and their cell bodies. The cell shown here was called a parasol cell by Stephen Policky (1941, 1957). Sources: A from Yamashita and Wilkie, 1991; B from Dacey and Peterson, 1992.
Other Eyes
Trilobite Visual System

- Most ancient known visual system.
- Compound eye with single crystal for each lens.

Electron Micrograph of Holochroal eye

Good trilobite eye info at: http://www.aloha.net/~smgon/eyes.htm

Scallop eyes

- Hundreds of primitives eyes, mirror in back
- Changes in light and motion and very rough images are registered on the retinas of the mollusk.
- Nice material at: http://soma.npa.uiuc.edu/courses/bio303/Ch11b.html
Stomatopod eyes

- Dumb bell shaped, compound eyes (next slide)
- Stereo vision with just one eye;
- Each eye is up on a stalk, with a wide range of motion;
- Stomatopods have up to 16 visual pigments
- can see ultra-violet and infra-red light
- some can see polarized light.
- See http://www.ucmp.berkeley.edu/aquarius/

Mantis Shrimp

Trinocular vision?
Visual resolution of human eye is 60 cycles per deg whereas Wedge-Tailed Eagle is 140 cycles per deg

Ocular Dominance Columns

Single Cell Recordings
6.23 AN ANATOMICAL/PERCEPTUAL MODEL OF THE VISUAL CORTEX. In this speculative model, visual streams within the cortex are identified with specific perceptual features. The anatomical streams are identified using anatomical markers; the perceptual properties are associated with the streams by applying the neuron doctrine. Source: Livingstone and Hubel (1988).
Subjective Contours
Kanizsa’s Triangle

Shading Cues
Shading Cues

(A)

(B)

Shading Cues

(A)

(B)
Which square is darker?
Global vs. Local information: Fraser’s Spiral
Context

Who is taller?  Who is taller?
Context: Whose faces do you see?

It works for Republicans too….
A picture of a man

In this shot, what is his facial expression?
In this shot, what is his facial expression?

Thatcher illusion

Hidden Human Face