Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neurobiology 130/230. Harvard College/GSAS 78454

Web site:  http://tinyurl.com/visionclass  (Class notes, readings, etc)
Location:  Biolabs 1075
Time:  Mondays 03:30 – 05:30
Dates:  Friday 09/04*, Mondays 09/14, 09/21, 09/28, 10/05, 10/19, 10/26, 11/02, 11/09, 11/16, 11/23, 11/30, 12/07*

Lectures:
Faculty:  Gabriel Kreiman and invited guests

Contact information:
Gabriel Kreiman
gabriel.kreiman@tch.harvard.edu
617-919-2530
Office Hours: After Class. Mon 05:30-06:30
<table>
<thead>
<tr>
<th>Class</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>Sep-04</td>
<td>Introduction to pattern recognition. Why is vision difficult?</td>
</tr>
<tr>
<td>Class 2</td>
<td>Sep-14</td>
<td>Visual input. Natural image statistics. The retina.</td>
</tr>
<tr>
<td>Class 3</td>
<td>Sep-21</td>
<td>Psychophysics of visual object recognition [Ken Nakayama]</td>
</tr>
<tr>
<td>Class 4</td>
<td>Sep-28</td>
<td>Lesion studies in animal models. Neurological studies of visual deficits in humans.</td>
</tr>
<tr>
<td>Class 5</td>
<td>Oct-05</td>
<td>Introduction to the thalamus and primary visual cortex [Camille Gomez-Laberge]</td>
</tr>
<tr>
<td>Oct-12</td>
<td></td>
<td>Columbus Day. No class.</td>
</tr>
<tr>
<td>Class 6</td>
<td>Oct-19</td>
<td><strong>Adventures into terra incognita.</strong> Neurophysiology beyond V1 [Hanlin Tang]</td>
</tr>
<tr>
<td>Class 7</td>
<td>Oct-26</td>
<td>First steps into inferior temporal cortex [Carlos Ponce]</td>
</tr>
<tr>
<td>Class 8</td>
<td>Nov-02</td>
<td>From the highest echelons of visual processing to cognition [Leyla Isik]</td>
</tr>
<tr>
<td>Class 9</td>
<td>Nov-09</td>
<td>Correlation and causality. Electrical stimulation in visual cortex.</td>
</tr>
<tr>
<td>Class 11</td>
<td>Nov-23</td>
<td>Computer vision. Towards artificial intelligence systems for cognition [Bill Lotter]</td>
</tr>
<tr>
<td>Class 12</td>
<td>Nov-30</td>
<td>Computational models of visual object recognition.</td>
</tr>
</tbody>
</table>
How do we go from oriented lines to complex shapes?

Hubel and Wiesel (1959) *J. Physiol.* **148**: 574-591


Divide and conquer strategy: multiple small steps are required to solve a complex task.
Adventures into *terra incognita*

Felleman and Van Essen. *Cerebral Cortex* 1991
The curse of dimensionality

Exhaustive exploration of the high dimensional image space is not possible with current techniques
Response latency increases along the visual hierarchy

Each additional processing step takes ~15 ms

Schmolesky et al 1998
Receptive field size increases along the ventral visual stream

Fig. 2. Schematic diagram showing convergence achieved by the forward projections in the visual system, and the types of representation that may be built by competitive networks operating at each stage of the system from the primary visual cortex (V1) to the inferior temporal visual cortex (area TE) (see text). Area TEO forms the posterior inferior temporal cortex. The receptive fields in the inferior temporal visual cortex (e.g. in the TE areas) cross the vertical midline (not shown). Abbreviation: LGN, lateral geniculate nucleus.

Wallis and Rolls 1997
Receptive field size increases along the ventral visual stream

Freeman and Simoncelli 2013

Fig. 2. Responses of neurons in area 18 of the monkey visual cortex to edges, bars, and stimuli producing illusory contours. The stimuli (insets) (10) were moved back and forth across the receptive fields (neuron 1, 1° at 1 Hz; neurons 2 and 3, 2° at 1 Hz). Each was presented 8 (I), 16 (J), or 24 (A to H) times; blocks of eight repetitions were alternated in pseudo-random order. For neurons 1 and 2, the response fields (the regions in the visual field where the neurons could be activated by a bar or edge) are represented by ellipses, and the fixation point is marked by crosses in A and F; the responses are represented by rows of dots; mean numbers of spikes per stimulus cycle are indicated on the right. Neuron 1, which responded to the lower right edge of the light bar (A), was activated also when only the illusory contour passed over its response field (B). Either half of the stimulus failed to evoke a response (C and D); (E) spontaneous activity. Neuron 2 responded to a narrow bar (F) and, less strongly, to the illusory bar stimulus (G). When the ends of the “bar” were intersected by thin lines, however, the response was nearly abolished (H). In neuron 3, the border between two abutting gratings elicited a strong response. The orientation tuning curves show corresponding peaks for bar and illusory contour (I). When the lines inducing the contour were reduced in number to less than three, the response disappeared (J); compare the lines...
Visual area V4

V4 lesions:
- moderate impairment in simple 2D shape discrimination
- Large deficit in 3D object recognition
- Loss of color constancy
- Deficits in the ability to detect less salient objects

V4 implicated in many visual functions
Neurons in V4 show color selectivity


FIG. 2.  A: example of the responses of a neuron in area V4 to patches of different isoluminant colors presented against a gray background. Spike rasters and response histograms are plotted with reference to the CIE chromaticity diagram. The bar under each histogram shows the duration of stimulus presentation (1 s). The neuron responded best to the blue stimulus. B: histogram showing the distribution of spectral preferences of all the V4 neurons recorded in these experiments. C: schematic representation of the part of the brain from which recordings were made. The dashed line on the lateral view of the monkey brain shows the approximate position of the coronal section (right). □, spectrally tuned neurons.
Probing the responses of V2 and V4 neurons

Varied responses along the ventral visual stream

Increase in “complexity” of feature preferences along the ventral visual stream

Smax = maximum response to “simple stimulus”
MAX = max response to all stimuli
Smax/MAX = 1 → “simple responses”
Smax/MAX = 0 → “complex responses”


FIG. 10. Distribution histograms of the ratio of Smax/MAX and the size of the receptive field in the 4 regions. The size of the receptive field is given by the square root of the area of the receptive field. See METHODS for the method of determining the border of the receptive field and the method of calculation for the area. Filled areas in right histograms: cells having
Increase in “complexity” of feature preferences along the ventral visual stream

<table>
<thead>
<tr>
<th></th>
<th>V2</th>
<th>V4</th>
<th>posterior IT</th>
<th>anterior IT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>![Shape 1]</td>
<td>![Shape 2]</td>
<td>![Shape 3]</td>
<td>![Shape 4]</td>
</tr>
<tr>
<td></td>
<td>![Shape 5]</td>
<td>![Shape 6]</td>
<td>![Shape 7]</td>
<td>![Shape 8]</td>
</tr>
<tr>
<td></td>
<td>![Shape 9]</td>
<td>![Shape 10]</td>
<td>![Shape 11]</td>
<td>![Shape 12]</td>
</tr>
<tr>
<td></td>
<td>![Shape 13]</td>
<td>![Shape 14]</td>
<td>![Shape 15]</td>
<td>![Shape 16]</td>
</tr>
<tr>
<td></td>
<td>![Shape 17]</td>
<td>![Shape 18]</td>
<td>![Shape 19]</td>
<td>![Shape 20]</td>
</tr>
</tbody>
</table>

Neurophysiological recordings in the human brain

- Patients with pharmacologically intractable epilepsy
- Multiple electrodes implanted to localize seizure focus
- Targets typically include the temporal lobe (inferior temporal cortex, fusiform gyrus), medial temporal lobe (hippocampus, entorhinal cortex, amygdala and parahippocampal gyrus)
- Patients stay in the hospital for about 7-10 days
Shape selectivity in human extrastriate visual cortex

Liu et al. Neuron 2009
Visual shape selectivity is largely focused along the ventral visual stream

Responsive

2205 electrodes
27 subjects

Selective
Further reading


Original articles cited in class (see lecture notes for complete list)