Visual Object Recognition

Neurobiology 230 – Harvard / GSAS 78454

Gabriel Kreiman

Email: gabriel.kreiman@tch.harvard.edu
Phone: 617-919-2530
Web site: http://tinyurl.com/vision-class
Dates: Mondays
Time: 3:30 – 5:30 PM
Location: Biolabs 1058
Visual system circuitry

Parietal Pathway

Temporal Pathway

Felleman and Van Essen. *Cerebral Cortex* 1991
From the retina to cortex

Primary visual cortex in Brodmann’s map

Brain shown from the side, facing left. Above: view from outside, below: cut through the middle. Orange = Brodmann area 17 (primary visual cortex)
Studies of gunshot lesions revealed topographic visual deficits.

The complex circuitry of cortex as drawn by Ramon y Cajal

GOLGI-STAINED NERVE TISSUE from the visual cortex of a rat was sketched by Cajal in 1888. The numbers along the right-hand margin identify cellular layers; the capital letters label individual neurons. One of Cajal's most important contributions to neuroscience was to establish the neuron as a discrete, well-defined cell rather than as part of a continuous network.
The gold standard to examine neuronal activity: microelectrode recordings

Edgar Adrian 1926

Neuronal resolution
Sub-millisecond temporal resolution
Direct examination of action potentials

Neurophysiological recordings from primary visual cortex

Hubel – Nobel Lecture

Hubel and Wiesel 1968
Receptive fields for simple cells in V1

Gabor function

\[ D(x,y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left[-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right]\cos(kx - \phi) \]

Spatial receptive field
Cat primary visual cortex (area 17)
Jones and Palmer 1987

Interlude: MATLAB
An easy way to write computer code

- “High-level” computer programming language
- Quite powerful!

\[
\theta_{\text{rad}} = (2\pi/360) \theta; \quad \theta \text{ angle in radians}
\]

\[
x = (-2*\sigma_x):\text{bin}:(2*\sigma_x); \text{nx} = \text{length}(x); \quad \% \text{define x axis}
\]

\[
y = (-2*\sigma_y):\text{bin}:(2*\sigma_y); \text{ny} = \text{length}(y); \quad \% \text{define y axis}
\]

\[
factor1 = 1/(2\pi*\sigma_x*\sigma_y);
\]

\[
\text{for i=1:nx}
\]

\[
\text{for j=1:ny}
\]

\[
curr_x = x(i)*\cos(\theta_{\text{rad}}) + y(j)*\sin(\theta_{\text{rad}});
\]

\[
curr_y = y(j)*\cos(\theta_{\text{rad}}) - x(i)*\sin(\theta_{\text{rad}});
\]

\[
factor2 = \exp(-\text{curr}_x^2/(2*\sigma_x^2) - \text{curr}_y^2/(2*\sigma_y^2));
\]

\[
factor3 = \cos(k*\text{curr}_x - \phi);
\]

\[
Ds(i,j) = factor1 * factor2 * factor3;
\]

\[
\text{end}
\]

\[
\text{end}
\]

\[
D(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right] \cos(kx - \phi)
\]
Interlude: MATLAB
An easy way to make plots

\[
D(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left[-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right] \cos(kx - \phi)
\]

\[
\begin{align*}
\text{sigma}_x &= 1; \\
\text{sigma}_y &= 1; \\
\text{bin} &= 0.05; \\
k &= 1/0.25; \\
\text{theta} &= 0; \\
i &= 0; \\
\phi &= 0; \\
\end{align*}
\]

\[
[\text{Ds}, \text{x}, \text{y}] = \text{mygabor1}(\text{sigma}_x, \text{sigma}_y, k, \phi, \text{theta}, \text{bin});
\]

\[
\text{subplot}(2, 2, 1); \\
\text{mesh}(\text{x}, \text{y}, \text{Ds}'); \\
\text{axis}([\min(\text{x}) \max(\text{x}) \min(\text{y}) \max(\text{y}) \min(\text{Ds}:) \max(\text{Ds}:)]);
\]

\[
\text{subplot}(2, 2, 2); \\
\text{contour}(\text{x}, \text{y}, \text{Ds}'); \\
\text{axis} \text{ square};
\]
A model for orientation tuning in simple cells

A feed-forward model for orientation selectivity in V1
(by no means the only model)

Text-fig. 4. Responses of a cell with a complex field to stimulation of the left (contralateral) eye with a slit $\frac{1}{4} \times 2^\circ$. Receptive field was in the area centralis and was about $2 \times 3^\circ$ in size. $A-D$, $\frac{1}{4}^\circ$ wide slit oriented parallel to receptive field axis. $E-G$, slit oriented at 45 and 90° to receptive-field axis. $H$, slit oriented as in $A-D$, is on throughout the record and is moved rapidly from side to side where indicated by upper beam. Responses from left eye slightly more marked than those from right (Group 3, see Part II). Time 1 sec.
Simple and complex cells in primary visual cortex

Fig. 1. Response patterns of a representative simple cell (A) and complex cell (B) to gratings drifted across their receptive fields. The response (peri-stimulus time histogram, PSTH) averaged over 20 repetitions of the sinusoidal stimulus is shown above a printout of the d.c. (mean rate of firing) and the first five harmonic components (amplitude/phase). The average maintained discharge in the absence of any visual stimulus is also displayed for each cell. Note that the simple cell’s response to the drifting grating shows a discharge pattern which modulates in synchrony with the fundamental temporal cycle of the stimulus, therefore most of the power appears in the 1st harmonic. The complex cell’s response, on the other hand, shows an overall increase in the mean rate of firing with little modulation, therefore the response appears in the d.c. component with little power in the harmonics.

Fig. 3. Distribution of the a.c./d.c. ratio for all cells measured (n = 343). Those cells whose d.c. is larger than the a.c. (that is, Y cells) fall between 0.0 and 1.0; those cells whose a.c. is larger than the d.c. fall above 1.0. It is clear that the distribution is best described as bimodal indicating the presence of two distinct populations of cells.

De Valois et al 1982
“Canonical” microcircuits in neocortex

Felleman and Van Essen 1991
Douglas and Martin 2004
Edges can take us a long way towards object recognition

MATLAB:
\[ I: \text{image} \]
\[ I_{\text{edges}} = \text{edge}(I); \]

Different methods:
Sobel, Prewitt, Roberts, Laplacian of Gaussian, Canny (determining how the gradients of I are computed)

1.9% of pixels > 0

2.9% of pixels > 0
Do we know what the early visual system does?

Up to 85% of “V1 function” has yet to be accounted for (Olshausen and Field 2005)

- Biased sampling of neurons
- Biased stimuli
- Biased theories
- Contextual effects
- Internal connections and feedback
- Joint activity

![Image](image.png)

Figure 8. Summary of the effect of finite sampling on predictions. Each bar indicates mean squared prediction correlation for the 49 neurons with greater than 2000 stimulus-response samples (error bars indicate standard error). Data for image domain STFs are shown at left, and for Fourier power STFs at right. $\rho^2$ indicates the mean squared prediction correlation actually measured for the STFs. $\rho^2_{\text{valmax}}$ indicates mean prediction after correcting for finite sampling of validation data. $\rho^2_{\text{ideal}}$ indicates the mean prediction after correcting for finite sampling of both estimation and validation data. Fourier power STFs perform consistently better than image domain STFs. After correcting for sampling limitations, Fourier power STFs can account for an average of 40% of the response variance in V1. The remaining portion of the response results from nonlinear response properties (‘unexplained variance’) not included in the Fourier power model.

David and Gallant, J.L. Network (2005)

Carandini et al J. Neurosci. 2005
Further reading


Original articles cited in class

Simoncelli and Olshausen. Annual Review of Neuroscience 2001
Felleman and Van Essen. Cerebral Cortex 1991.
De Valois et al. Vision Research 1982