Sensory Coding

Maneesh Sahani

26 October 2016

1. **Essay: early visual and auditory processing.**

   This question will ask you to read about the mammalian visual and auditory pathways. As a first source, take a look at Kandel et al.: Principles of Neural Science; if you want further material, look at Zigmond et al.: Fundamentals of Neuroscience or ask for suggestions. Most of these questions don't actually have a right answer — just argue on the basis of what you read.

   - How many **synapses** are there between the receptors and the cortex in each system? Can the **subcortical** processing in the two pathways be compared?
   - Where does **bilateral convergence** happen? Why the difference?
   - Imagine you wanted to build **efficient** auditory and visual sensory systems. How do you think computational goals of processing in the two modalities might influence the anatomy of the system? Think in particular about the very early parts and bilateral convergence.

2. **Images seen through visual receptive fields**

   - Load the image provided into matlab using `imread('image.bmp')`.
   - Construct (in a MATLAB matrix) an on-centre difference-of-gaussians (DOG) centre-surround receptive field centred at 0:

     $$D(x,y) = \frac{1}{2\pi\sigma_c^2} e^{-(x^2+y^2)/2\sigma_c^2} - \frac{1}{2\pi\sigma_s^2} e^{-(x^2+y^2)/2\sigma_s^2}$$

     Make the receptive field on a 21-by-21 pixel grid, with a central Gaussian width of 1.5 pixels and a surround Gaussian width of 3 pixels. Note that the normalisation ensures that the RF has no DC component.
   - Suppose you had a cell with a receptive field like this centred at each pixel in the image. Show the image as represented by the activity of these cells (use `imagesc`), placing the cells in topographic order according to their centres. (Hint: this is effectively a 2D convolution. Why?).
   - Threshold the activity image (i.e. set all the values above some cutoff to 1, all below to 0). Does this look like the cells are detecting edges? Tune the parameters of the DOG RF and the threshold to improve the quality of the edge detection as much as you can.
   - Now construct a Gabor receptive field on the same 21-by-21 pixel grid:

     $$D(\vec{x}) = \exp\Big(- (\vec{k}(\theta)\cdot\vec{x})^2/2\sigma_l^2 - (\vec{k}_\perp(\theta)\cdot\vec{x})^2/2\sigma_w^2 \Big) \cos\Big(2\pi\frac{\vec{k}_\perp(\theta)\cdot\vec{x}}{\lambda} + \phi\Big)$$

     where $\vec{k}(\theta)$ is a unit vector with the orientation θ, $\vec{k}_\perp(\theta)$ is an orthogonal unit vector and θ, σ_l, σ_w, λ and ϕ parametrise the Gabor. Start with $\theta = \pi/2$, $\sigma_l = \sigma_w = 3$ pixels, $\lambda = 6$ pixels and $\phi = 0$.

- Show the image as seen by through receptive fields of this sort, again with one centred at each pixel in the image. Threshold the activities. Does this picture match what you expected?
- [Advanced] What parameter(s) determine(s) the orientation bandwidth of the Gabor (i.e. the range of orientations for which it will fire)? Derive a formal result using a cosine-grating stimulus with spatial frequency $\Omega_s = 2\pi/\lambda$ matching that of the Gabor, and orientation ϕ:

$$s(x, y) = \cos\left(\Omega_s \cos(\phi)x + \Omega_s \sin(\phi)y - \psi_s\right)$$

  [Hints: This is much easier in the 2D Fourier transform space. By symmetry you can consider the case of $\vec{k}$ parallel to the y-axis. It may be helpful to draw pictures (as in lecture) to keep track of the shape of the functions. Remember both the convolution theorem and Parseval's theorem (the latter tranforms dot products in visual space to dot products in Fourier space). You can use symmetry to restrict your consideration to a single lobe of the Fourier transform.]
- Adjust the parameter(s) you identified in the part above [or ask someone who did] to narrow the bandwidth, and look at the resulting image. Did it work?
- Construct 3 Gabors with $\theta = 0, \pi/4, \pi/2$ and the other parameters as above. Sum the thresholded outputs of all three types of cell. Does this image look any better than that obtained by the RGC? Why do you think this is?
- How might you improve the quality of the edge detection? (Hint: one thing to think about is boosting responses of cells that line up consistently along an edge).

3. **Contrast saturation and nonspecific suppression.**

- Assume a V1 cell has a response kernel,

$$f_{\alpha,a,\psi}(x, y) = r_{max} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \cos\left(a\cos(\psi)x + a\sin(\psi)y - \alpha\right)$$

  where α is the preferred phase, ψ the preferred orientation and a the preferred frequency. We stimulate it with the following (static) stimulus:

$$s(x, y) = B\cos\left(b\cos(\phi)x + b\sin(\phi)y - \beta\right)$$

  Plot the response

$$L_{\alpha,a,\psi}(\phi, \beta, b) = \int dx\, dy\, f_{\alpha,a,\psi}(x, y)s(x, y)$$

  for $\alpha = \beta$ as a function of orientation ϕ for $\psi = 0$ and as a function of frequency b for $a = 1$, showing that this cell is tuned to both spatial frequency and orientation. Hint: $\cos(x) = (e^{ix} + e^{-ix})/2$
- Unlike the prediction from this model, responses of cells in visual cortex saturate at high contrasts and they also adapt. Furthermore, presentation of a grating at an orientation to which the cell shows no response prevents the cell from responding to a grating presented at its preferred stimulus orientation (nonspecific suppression). Heeger (1992) proposed a very simple modification of this model that accounts for these two effects. Let simple $SC_{\alpha,a,\psi}$ and complex cells $CC_{a,\psi}$ respond as:

$$SC = \frac{[L]_+^2}{F_{1/2}^2 + [L]_+^2}$$

$$CC = \frac{\sum_{\alpha=0,90,180,270}[L_\alpha]_+^2}{G_{1/2}^2 + \sum_{\alpha=0,90,180,270}[L_\alpha]_+^2}$$

where $[f]_+ = f$ if $f > 0$ and $= 0$ otherwise; $F_{1/2}$ and $G_{1/2}$ are constants, and we have suppressed the subscripts where possible. Show analytically that this formulation leads to saturation at high stimulus contrasts B. How might you modify the two models to produce nonspecific suppression? What might these expressions imply about cortical architecture?

- Finally, consider spatiotemporally inseparable receptive fields: $f_{\alpha,a,\psi}(x, y, t)$, i.e. receptive fields that can't be written as a product of a purely temporal and purely spatial receptive field. How would each of the following cells respond to a drifting sinusoidal grating $s(x, y, t) = s(x, y)\cos(ct)$: a simple cell with spatiotemporally separable receptive field, a simple cell with inseparable receptive field, and a complex cell. What is the difference between complex cells and simple cells with an inseparable spatiotemporal receptive field?
- Suppose that natural scenes are composed of sums of Gabor functions. Assuming that the properties of V1 cells are matched to the statistics of natural scenese, what might the normalization structure implied by Heeger's model imply about those statistics?