- Mapping of the Visual Pathway
- Receptive Fields
- Modeling Receptive Fields
- Linear, Shift-Invariant Systems
- Impulse Response of LSI Systems
- The Power of Convolutions
- Convolutions : The Basics

1 Mapping of the Visual Pathway

This lecture discussed the neurophysiology of receptive fields, and their mathematical representations. As previously discussed, the first gathering of visual data from the scene occurs in the retina of the eye, where photoreceptors gather luminance values. The ganglion cells of the retina project back through the optic nerve to the LGN and then to V1, the first visual processing area of the brain. The left side of the visual field projects back to the right half of V1 and the right side of the visual field projects back to the left half of V1. LGN will loosely be considered as a "relay station".

A camera has pixels and then a map of processors that corresponds to these pixels. In humans, the photoreceptors of the retina can be thought of as the pixels and V1 contains the map of processors. This retinotopic mapping was demonstrated by Tootell, Silverman, Switkes and DeValois in 1982. They conducted a study using macaque monkeys. The monkeys were injected with radioactively labeled glucose and then shown a stimulus. When cells in V1 respond to a stimulus, they use more energy and take up more glucose. As a result, all the cells that responded to the display used radioactively labeled glucose, allowing the experimenters to do a mapping of V1 (figure 1). It is clear that the processors in V1 are a retinotopic map, but the foveal area (central visual angle) is overrepresented. The rings in the display are increasing in exponentially in diameter, but the processors in V1 are equally spaced. With modern technology, we can use alternate imaging techniques to study! brain responses to visual stimuili (e.g. MRI).

Main Conclusions from DeValois study:

1) More processors(25%) in V1 are dedicated to the fovea (central 1-2 degrees of visual field). The fovea is where we have the best resolution and where we most easily recognize objects.

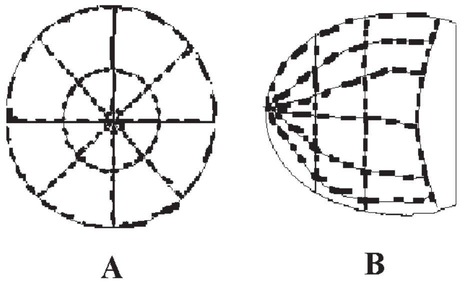

Figure 1: **A.** Stimulus used for mapping the retinotopic projection onto striate cortex (V1). **B.** Pattern of 2-DG uptake in macaque striate cortex produced by the stimulus pattern.

2) There is a logarithmic fall-off of processing from the fovea towards the periphery.

3) DeValois et.al. successfully marked of sectors of the visual field.

This processor mapping suggests possible processing strategies for computer vision.

2 Receptive Fields

The next questions are, "What are the processors doing? What do they respond to?". At the first stage of processing is the retina, which consists of photoreceptors (rods and cones), bipolar cells and ganglion cells (and others such as horizontal and amacrine cells, figure 2). The wiring of these cells create receptive fields that respond differently to stimuli. One example is the "center-surround" receptive field (figure 3). A cell with this type of receptive field will respond optimally to a bright light in the center of its receptive field and a dark light in the surround of its receptive field. These neurons are referred to as "spot detectors". If a larger bright light is introduced to the receptive field, negative input from the horizontal cells will cancel some of the excitation from the surround going to the bipolar cell. So, to a first approximation, the bipolar cell can be thought of as doing a linear summation of input across its receptive field:

$$S = w_{ij} I_j$$

where I is the intensity input and w_{ij} is the synaptic strength from i to j. Reversed connections would respond optimally to a dark spot. Such receptive fields would be quite useful to say, a FROG, who wants to track the dark spot, the FLY, as a matter of survival. Bipolar cells are the first cells to have receptive fields in monkeys (and humans). The receptive fields of the processors get more complex as we progress through the visual system.

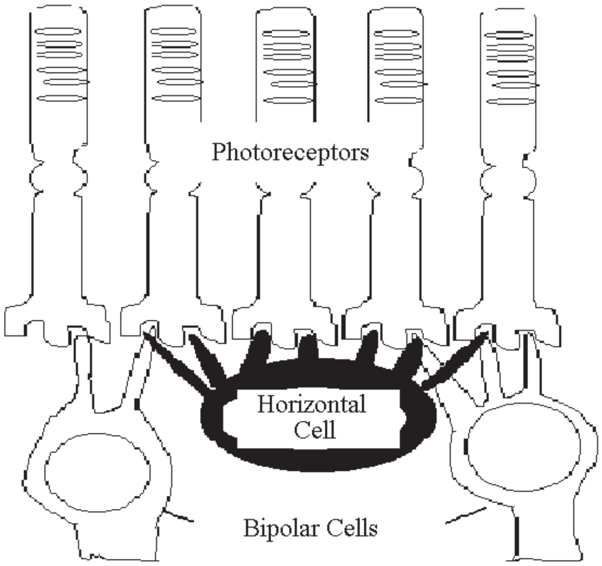

Figure 2: Retinal physiology. The horizontal cells provide summing over many photoreceptors, while the bipolars get information from very few receptors.

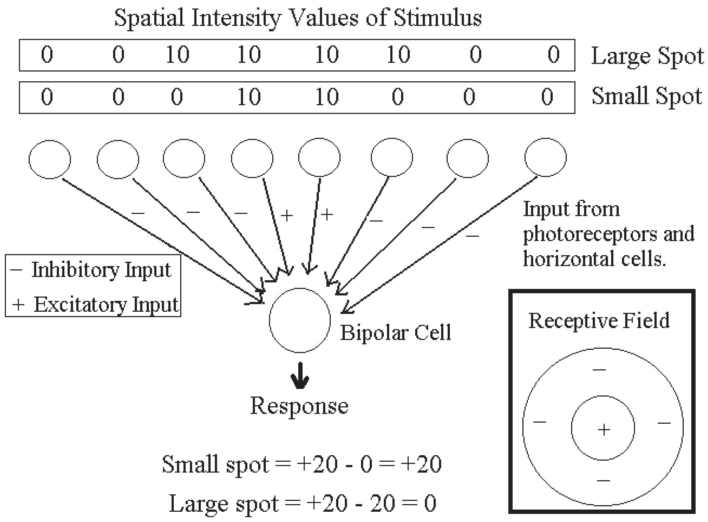

Figure 3: Schematic of possible recpetive field wiring.

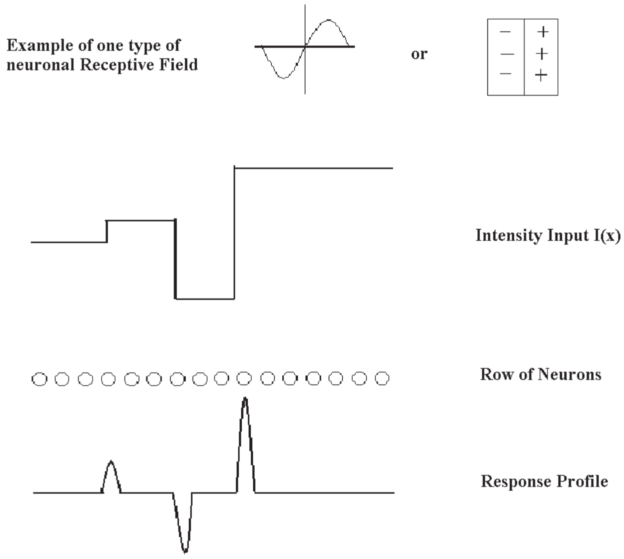

Figure 4: Response of edge detecting receptive fields to a stimulus.

3 Modeling Receptive Fields

Receptive fields can be modeled mathematically and treated as filters of visual stimuli. Consider the brightness input $I(x)$ in figure 4. Neurons with this type of receptive field and associated response could be considered "edge detectors", where the response of the neurons is simply a convolution between the stimulus (I) and the filter (receptive field, RF):

$$R = I * RF, \text{ or equivalently, } R(x) = \int I(u)RF(x-u)du$$

Refer to the end of these notes for a discussion on convolutions.

The important assumption we make by using convolutions is that these neurons are behaving linearly. This isn't exactly true, but it is a reasonable approximation. Now consider cells with different receptive fields and what they respond to (figure 5). The weights of the different inputs are determined by synaptic strength. From this point on, we can consider neurons as a sheet of processors that are capable of performing convolutions. This puts us back in the realm of pure mathematics!

4 Linear, Shift-Invariant Systems

An example of an LSI system is shown in figure 6a. Such systems must obey the following properties:

1) If a system is linear

$$\alpha v_1 + \beta v_2 \;\; -> \;\; \alpha w_1 + \beta w_2$$

The response to a linear combination of known inputs is a linear combination of known outputs for their inputs.

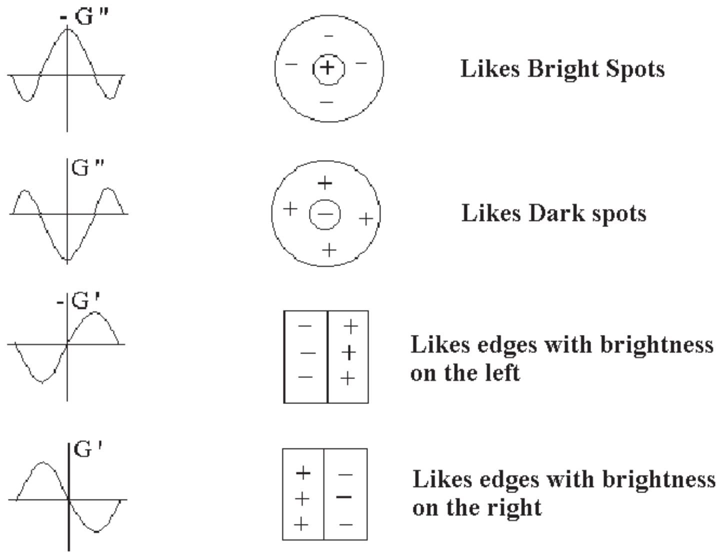

Figure 5: Example of receptive fields derived from a normalized Gaussian (G).

2) Shift Invariance. Instead of $v(t)$, give the input $v(t-\tau)$ and the output will be shifted by the same amount, $w(t-\tau)$.

These properties can be extended to include two-dimensional systems.

5 Impulse Response of LSI Systems

Theorem: Any LSI system can be completely characterized by its impulse response function, h. For an input v, the output is given by $v * h$ (figure 6b). The impulse response function h is the response of the system to an impulse, the dirac delta function δ (figure 7a). The dirac delta function can be considered as the limit of a normalized Gaussian (area = 1) as sigma tends toward zero.

$$G_\sigma(x) = 1/\sqrt{(2\pi)}\sigma \exp^{-x^2/2\sigma^2}, \text{ where } Area = \int G_\sigma(x)dx = 1$$

The result is an infinite pulse at $x = 0$ with $area = 1$ (figure 7b). A 2D dirac delta function $\delta(x,y)$ looks like a very tiny, very bright spot. The impulse response function can be measured experimentally, thereby complete characterizing the LSI system. An intuition that this is true can be built by understanding that every input can be divided into an infinite number of dirac delta functions that are scaled and shifted (figure 8). A convlution is a weighted sum, just like the function shown in figure 8. We make the assumption that neurons are LSIs so that we can characterize them with just an impulse response. This is not entirely correct for neurons, because they are non-linear in that they have saturation points (e.g. max firing rate around 300 Hz).

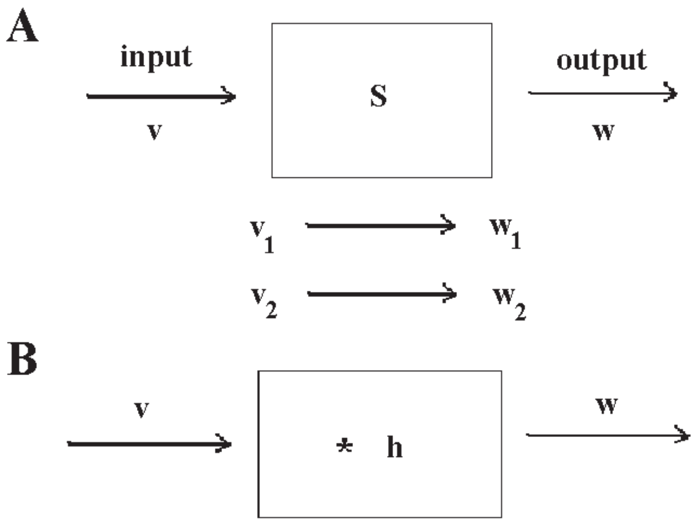

Figure 6: **A.** LSI System. **B.** Role of Impulse Response Function.

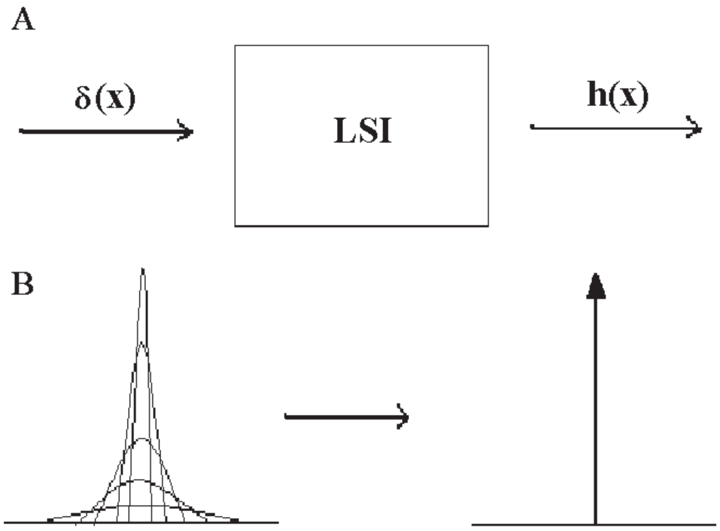

Figure 7: **A.** Impulse Response Function. **B.** The dirac delta function is the limit of a normalized Gaussian (area = 1) as the sigma of the Gaussian tends toward zero.

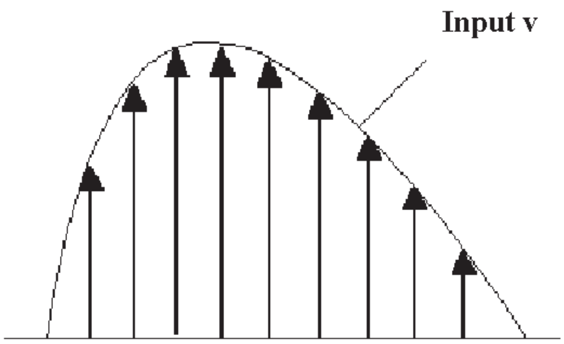

Figure 8: Functions can be considered as a summation of scaled and shifted dirac delta functions.

6 The Power of Convolutions

Convolving any function with a Gaussian results in a smoothing of that function. As sigma of the Gaussian increases, more smoothing occurs. Convolving with the derivative of a Gaussian allows for edge detection, and convolving with the second derivative of a Gaussian allows for spot detection (figure 12). This type of linear filtering is a reasonable approximation of what's going on in the human visual system. Some important properties of convolutions are:

1) Commutativity: $f * g = g * f$

2) Associativity: $(f * g) * h = f * (g * h)$

3) Identity element for convolution is the dirac delta function : $f * \delta = f$

4) Commutivity with differentiation: $(f * g)' = f' * g = f * g'$

7 Convolutions : The Basics

A convolution of two functions, $f(x) * g(x)$ is defined as a new function $h(x)$, such that

$$h(x) = \int f(u)g(x-u)du.$$

The output of a given location x is a weighted sum of the two functions. This is a standard operation in signal processing. In the case of receptive fields, we have the functions $R(x)$, the response at position x; $I(x)$, the input or intensity at position x; and $RF(x)$, the receptive field at position x.

$$R(x) = \int I(u) RF(x-u) du$$

$RF(x)$ in this formula is centered at x, flipped around and shifted by u to do the weighting. These fomulae can also be expanded to two dimensions, i.e. $R(x, y)$. Convolutions can only be used with linear, shift invariant systems.