Video: Early Visual Processing from the Retina to LGN

[00:01] [slide 1] In the next two videos, you'll learn about the early stages of processing in the human visual system, drawing in part from studies of higher primates whose visual system is very similar to our own. We'll begin with the eye, focussing on the retina which is a thin sheet of cells around the back of the eye that sense the incoming light, process it, and send information up the optic nerve to an area called the LGN. Signals from the LGN travel to visual cortex at the back of the brain. This first video will take you from the incoming light to the doorstep of an area called primary visual cortex. Along the way, you'll learn about the structure of neurons and how information is encoded in neural signals. We'll also connect these first stages of processing to the computations you learned about earlier, smoothing and measuring derivatives of intensity in the image. In the second video, we'll examine the processing that takes place inside primary visual cortex, related to the detection of features like edges in the visual image.

[01:12] [slide 2] Light enters the eye through the pupil, which opens and closes to let in different amounts of light, depending on the overall brightness of the environment. It then passes through the lens, which focuses the image on the retina at the back of the eye. A diagram of a small snippet of the retina is shown enlarged on the right. Light actually passes through multiple layers of cells before it reaches the photoreceptors at the very back. There are two kinds of photoreceptors, rods and cones, that absorb the incoming photons of light through a chemical process, and convert them to electrical signals. Those signals are processed through the layers of cells that end with the ganglion cells whose output fibers form the optic nerve that passes out of the retina. There's a blind spot in your visual field that spans this small area of the retina where the optic nerve exits the eye and there are no photoreceptors.

[02:18] [slide 3] Before exploring the computations taking place in the retina, let's first examine the structure of individual neurons, and how they encode information in electrical signals. A typical neuron has dendrites where it receives inputs from other neurons. Those inputs are then integrated at the cell body where an output signal is generated, in the form of electrical impulses that propagate along an axon. The axon branches out at the end and transmits this output signal to other neurons at locations called synapses on their dendrites. Neurons are highly interconnected - they can have thousands of input or output connections to other neurons. The electrical signal that's generated by a neuron can be measured with a very fine microelectrode inserted into the neural tissue, like in this picture in the lower right corner. There are many neurons in the brain, on the order of 10^11 neurons, and it's estimated that there may be up to 1.5 million ganglion cells in the human retina transmitting information up the optic nerve.

[03:33] [slide 4] What do neural signals look like? They have the form of spikes of electrical activity that can occur in rapid succession over time, and information is encoded in the frequency of occurrences of the spikes. That's defined as firing rate, or number of spikes per some time interval. When neurons are at rest, they occasionally generate spikes down their axons, and the inputs to a neuron can have the effect of increasing its firing rate, a process

referred to as excitation, or it can decrease the firing rate, below its resting rate, a process called inhibition. The direct recording of this electrical activity of individual neurons is not done in humans, but in animals. We can especially learn a lot from single cell recordings in monkeys whose visual system is very similar to our own.

[04:38] [slide 5] So let's return to the retina and touch on a few properties of the rods and cones that initially sense the incoming light. First, these two types of photoreceptor are distributed across the retina in a very different way, as conveyed in the picture in the top middle. The cones are sensitive to color and they're portrayed as the red, green, and blue dots in this picture. They're packed in the center of the eye and their density falls off rapidly with distance from the center. This is also shown in the graph on the left - the horizontal axis is angular position in the visual field, where 0 deg corresponds to the center of your field of view and that's shown in the middle. The vertical axis is density, and the red curve shows that the density of cones peaks at the center of the visual field and drops off dramatically in the periphery. In contrast, the rods are depicted as the brown dots in the picture, and they're absent in the middle of the eye and mostly present in the periphery as the blue curve shows on the graph. This white bar labeled optic disk, that's the blind spot where there are no rods and cones. The central 2 deg of visual angle is referred to as the fovea, and it's where we have the highest acuity, or ability to distinguish fine detail. To get a feel for the size of the fovea, it roughly spans your thumbnail when you hold it out at arm's length.

[06:22] I mentioned that the cones are sensitive to color - there are three types of cones, originally called red, green, and blue cones, but now they're more commonly referred to as long, medium, and short wavelength cones, abbreviated L, M, and S. Each type of cone absorbs more photons of light over a particular range of wavelengths, as indicated by the solid white curves on this colorful figure on the right. The short wavelength cones respond more to light in the blue end of the spectrum, the long wavelength cones respond more to light at the red end, and the medium wavelength cones are in the middle, peaking around the greens. In addition to being sensitive to color, cones also operate at high light levels, such as in daylight, and they adapt quickly to changing light levels. If you step out of a dark room into bright light, your eyes adjust quickly to the light. Rods are quite different. They're not so selective for color, they operate at low light levels, at night or in dark spaces, and they adapt slowly. If you're in a bright area and step into a dark room, it takes a long time for your eyes to adjust, before you can clearly see what's in the room.

[07:44] [slide 6] Photoreceptors are the input to visual processing. Let's next turn our attention to the output of the retina, carried by the retinal ganglion cells. The activity of a single neuron is only affected by light in a very limited area of the visual field that's referred to as the receptive field of the neuron. The response of retinal ganglion cells to light stimulation within their receptive field has a center-surround structure - what does that mean? Imagine that this circular area here represents the receptive field of a particular cell. If we flash a tiny spot of light anywhere in the receptive field, the cell will respond in some way. For some cells, flashing a

light anywhere in the center will excite the cell, and increase its firing rate, and flashing light in the surround will inhibit the cell, and decrease its firing rate. That's illustrated in the diagram on the right which shows some hypothetical spikes generated by the cell over time. During the time window highlighted in yellow, if the light is turned on in the center, the frequency of spikes increases, but in the surround, the activity is suppressed, there's almost no spikes. Ganglion cells with this pattern of behavior are referred to as "on-center" cells. There are other cells that we refer to as "off-center" cells, where shining light in the center inhibits the cell's activity and light in the surround excites the cell. The amount of increase or decrease in response to light varies within the receptive field. The pattern is circularly symmetric, and if we plot a cross-section of the varying response in an on-center cell, for example, it looks like this graph here. This should be familiar from what you learned about detecting intensity changes in computer vision systems - it resembles the second derivative of a Gaussian. In studies of biological vision systems, it's often described as the difference of two Gaussians, and to see why, we'll look at how this behavior arises from the processing of signals from the photoreceptors within the retina. Through this exercise, we'll learn a valuable lesson, that biological vision systems and machines can perform very similar computations in different ways, and on very different hardware.

[10:34] [slide 7] In this diagram, at the top is a dense cluster of photoreceptors, and between them and the retinal ganglion cells is a layer of neurons called bipolar cells. These bipolar cells also have receptive fields with the center-surround structure that I just described - this is where that structure originates. They receive direct input from an area of receptors that's depicted in blue, and signals from receptors closest to the center of this blue area are weighed more heavily than those that are further away. There's another type of cell called a horizontal cell, that receives inputs from receptors over a wider area that spans both the red and blue receptors in this picture. Horizontal cells also weigh signals in the center of this area more heavily, but these cells inhibit the bipolar cells. The result of all this processing is that light intensities in the center of the receptive field are added together using a pattern of weights that resembles the narrow Gaussian curve here shown in blue. In the surrounding area, the light intensities are integrated with a pattern of weights inhibiting the bipolar cells, so I represented this by showing that wider Gaussian curve in red, and negative. The combination of these two effects is what produces the final response of bipolar cells to light stimulation in their receptive field. The key ideas here are first, the computation performed by these neurons is essentially the same convolution computation that you learned about earlier - each bipolar cell is weighing the receptor signals within a neighborhood by different amounts, and adding them together to produce a response. Imagine a whole grid of these cells performing the same computation at different locations, all in parallel. This mimics the convolution computation. The second key point is that the overall pattern of weights is very similar to the Laplacian of a Gaussian that we described in the context of computer vision systems, and we understand from that discussion, its important role in detecting intensity changes in the image.

[13:15] [slide 8] Here's the Wellesley lamppost and convolution of the image with the Laplacian-of-Gaussian operator. Retinal ganglion cells with their center-surround receptive fields are effectively performing the same computation, but you may be wondering, why are there on-center and off-center cells? Let's take a small snippet of the convolution and show it enlarged on the bottom. There are positive and negative values here, with the bright areas showing the positive values and dark areas show the negative values. The on-center and off-center cells are performing the same computation, but with the sign reversed, so the on-center cells are going to be active where we see the bright areas of this convolution, the positive parts, and the off-center neurons are active in the dark areas, the negative parts of the convolution. In a computer system, we can represent positive and negative values with equal precision, but that's not possible in neurons, particularly when you're sending information over long distances, like up the optic nerve. So in the biological system, there are two separate populations of neurons carrying the positive and negative parts of this convolution from the retina up to the brain.

[14:44] [slide 9] There's one last observation about this processing that I'd like to mention. We talked about using multiple operator sizes, with different amounts of smoothing to capture intensity changes at multiple scales. This also occurs in human vision. At each location of the visual field, there are neurons with different size receptive fields, and overall, receptive fields get larger as you move away from the center of the eye. When we look at a scene like this, we think we see everything in crystal clear detail, but really, the early stages of visual processing see something more like this. Imagine that you're looking directly at the siamese cat on the coffee cup tucked in the middle of the scene here. That area of the image is processed in detail, but the periphery is fuzzy. We can view other areas in detail by moving our eyes and bringing that high acuity fovea to different locations. When we do this, as we move our eyes around, we can remember all the fine details, but we can't actually see all this detail at once.

[slide 10] I mentioned at the beginning of the video that I would take us through the LGN. For the time being, we're going to think of the output of the LGN as being the same as the information conveyed by the retinal ganglion cells. Later in the course we might touch on the role of the LGN in helping to shift our attention around to different locations in a scene. In the next video, we'll peer into the early stages of processing by neurons in primary visual cortex and the detection of features like edges in the visual scene.