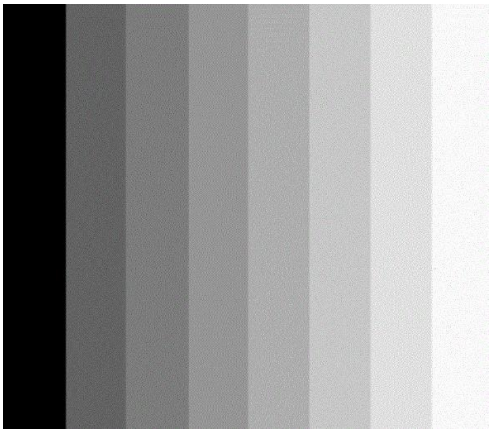
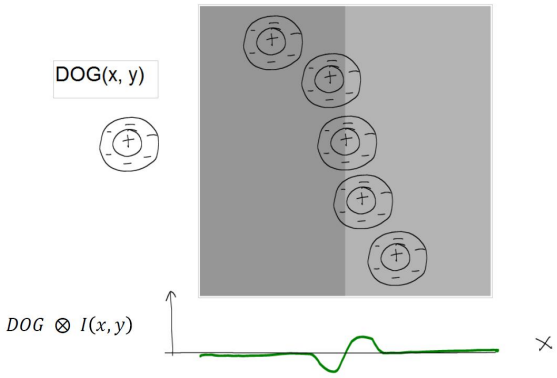
Today we'll examine how orientation information such as edges and lines is encoded in early visual processing. There is much to say about this topic before one gets to the stage in visual processing where individual cells are sensitive to oriented structures though. I'll keep this preliminary discussion short and just discuss a specific phenomenon called Mach Bands[1].

Mach Bands

If you look at the image on the left which consists of a set of stripes, each of constant shade of gray, you will notice that the boundaries between the stripes appear to have slight rise (when the stripe goes from light to dark) or fall (when it goes from dark to light). This *edge enhancement* effect is believed to be an artifact of how our eye and brain codes the image. It causes us to fail to perceive the intensities as they really are.





Many have argued that Mach bands are the result of the center-surround coding mechanism, in particular, the DOG "filtering" that happens in the retina. This idea is illustrated in the sketch above right. As we move the DOG template across the edge, it begins in a uniform region and gives 0 response since the ON and OFF regions cancel. Then it encounters a rise in intensity in an OFF region, which leads to an overall negative linear response. When the DOG straddles the edge, the left and right halves of the DOG each have uniform intensities and because of symmetry the ON and OFF regions in each half are balanced just as the cell's overall ON and OFF regions are balanced, so again the cell give no response. As the DOG template continues beyond the edge, the intensity rises because the tailing edge of the OFF region falls on the lower intensity region – less OFF contribution leads to an increase in response.

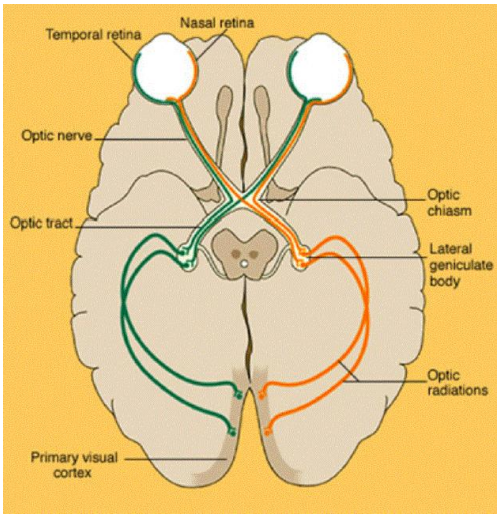
Mach bands are not just a curiosity. They have practical applications, for example, when people examine images and need to make subtle distinctions between grey levels. The best example of this is dentists or radiologists who examine radiographs. Such professionals are well acquainted with the effects of Mach bands. They cannot change their visual systems, but they can learn when and when not to believe what their eyes are telling them.

[1] named after Ernst Mach who was a 19th century scientist

Early visual pathway: retina to cortex

Let's move further into the brain. The axons of the retinal ganglion cells of each eye are bundled together into the optic nerve which sends the signals to the lateral geniculate nucleus (LGN) on each side of the brain. The two LGNs which are in the thalamus are located near the center of the brain. Note that the optic nerve from each eye needs to split into two in order to send signals to both halves of the brain. See figure below.

Cells in the LGN relay the signals to the surface of the back end of the brain. The surface of the brain in general is called the *cortex*, and the surface of the brain at the back of the head is called the *primary visual cortex* (V1) because this is the first area of the cortex to receive visual inputs.



[ASIDE: I say the LGN cells "relay" the retinal signals to the cortex, but there is more going on than that. The LGN receives axons from the retinal ganglion cells (about 10^6 of them), but it receives far more axonal inputs (about 10^7 of them) from the visual cortex – that is, there is a feedback loop between the visual cortex and the LGN.]

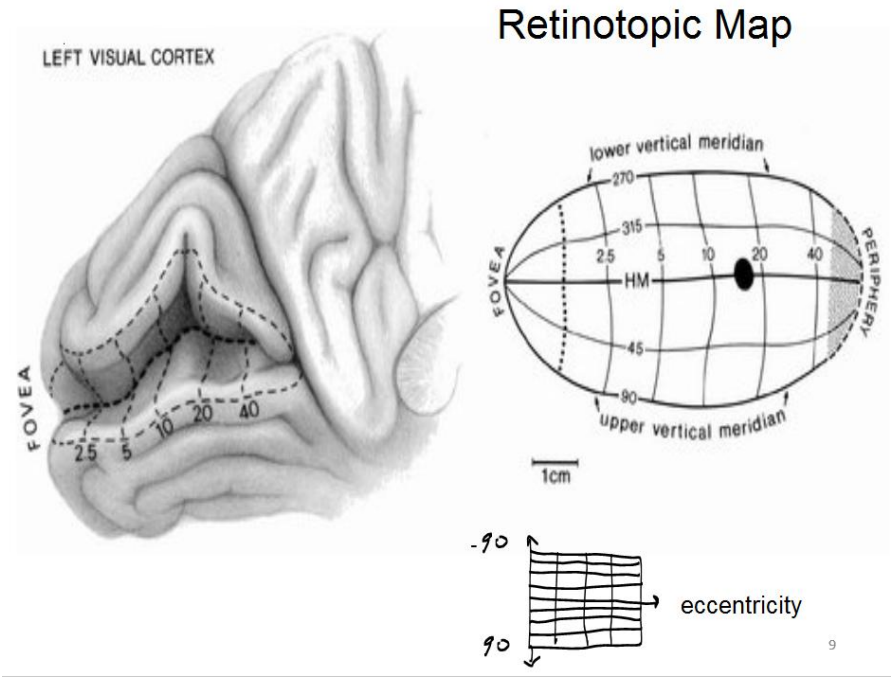
The figure above illustrates how the two halves of the visual field are coded by the two halves of the brain. The left half of each retina codes the right visual field and the axons from these retinal cells terminate in the LGN on the left side of the brain. Relay cells in the LGN then send their axons to the left half of the primary visual cortex (V1). Thus, the left half of V1 receives the image code from both eyes for the right visual field only. Similarly, the right half of V1 receives the image code of left visual field. The cells that encode the "seam" along a vertical meridian between the left and right halves of the visual field can be found in both halves of the brain. So there is an overlap in the representation of the vertical meridian.

Retinotopic maps

When the axons from the retina are bundled into the optic nerve, their spatial arrangement is preserved (to some extent). These cells terminate in the LGN. When you measure the receptive fields from neighboring cells in the LGN, you typically find that they encode the intensities of nearby visual directions, or equivalently, nearby retinal positions. In this sense, the LGN is said have a

retinotopic map: nearby points on the retina map to nearby points in the LGN. Similarly, cells in the LGN project to V1 and if you measure the receptive fields of nearby cells in V1, you generally find that they encode intensities of nearby visual directions (i.e. nearby positions on the retina). So V1 also has a retinotopic map.

Here I give just a few details about the retinotopy in LGN and V1. There are six layers in each LGN, and each relays information from just one eye. There are also differences in the receptive field properties of different LGN layers. In some layers, the cells have relatively large receptive fields but are not sensitive to color differences, and these cells respond to the time variations in the stimuli. These cells are involved in motion processing which I'll get to in a few lectures. In other LGN layers, the cells have smaller receptive fields which encode color and intensity differences. These cells do not seem to be involved in motion processing. The details of the different LGN layers are not crucial for our understanding. My main point here is that within each layer of the LGN, the cells are arranged in a retinotopic map. They then relay signals to V1.



Because the receptive fields[2] of retinal ganglion cells in the fovea are so much smaller than in the periphery, it is possible to pack many more retinal ganglion cells per mm^2 in the center of the retina. The signals get relayed from LGN to V1, and so the inputs to V1 are dominated by the cells near the center of the visual field. This requires a deformation of the retinotopic map.

One simple way to think about this deformation of the retinotopic map in V1 is to use polar coordinates (r, θ) for visual direction instead of (x, y): one coordinate r is eccentricity and the other coordinate θ is an angle away from say the x axis. This polar coordinate system (r, θ) captures

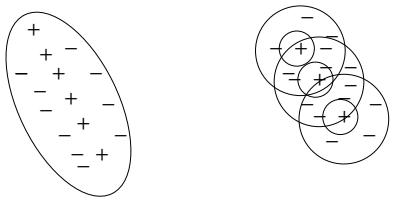
[2]The receptive field size doesn't just depend on the cell body. It also depends on the width span of the cell dendrites (branches) that the cell uses to "read" signals from its neighbors.

the distortion of the retinotopic map, namely that the number of cells that represent a given visual angle increases toward the center of the image (toward $r = 0$.) The figure above illustrates the distortion. V1 in the left cortex is flattened out into an elongated ellipsoid (close to a rectangle). The directions θ from -90 to 90 degrees (or 270 deg to 90 deg) are represented in the map. These cover half of the visual field.

The slides show another example which is based on fMRI images. The point there is that the central part of the visual field is coded using a relatively large part of V1. We still have a retinotopic map, but it is distorted.

Orientation selectivity in primary visual cortex

What are the receptive field properties of cells in the primary visual cortex? The first experiments to successfully address this question were carried out in the late 1950's by David Hubel and Torsten Wiesel. (For this and subsequent work, these two researchers were awarded the Nobel Prize.) Hubel and Wiesel examined the responses of single cells in primary visual cortex of anaesthetized cats. They found that each cell responded to an small area of the visual field but, unlike in the retina and LGN, the receptive fields in V1 were not radially symmetric. Instead the cells were tuned to a particular orientation, such as in the sketch below. The response of this cell can be thought of as a weighted average of the image intensity over the ellipsoidal region shown. The weights are positive along a center stripe parallel to the elongation and negative along the two flanking stripes. Such cells might be thought of as *line detectors*. Cells are also found of the opposite sign, namely negative along the center stripe and positive along the flanking regions.



Hubel and Wiesel discovered the orientation properties quite accidently. Hubel describes the discovery here: `https://www.youtube.com/watch?v=IOHayh06LJ4`
For a longer video showing the mapping of the receptive field, see
`https://www.youtube.com/watch?v=Cw5PKV9Rj3o`.

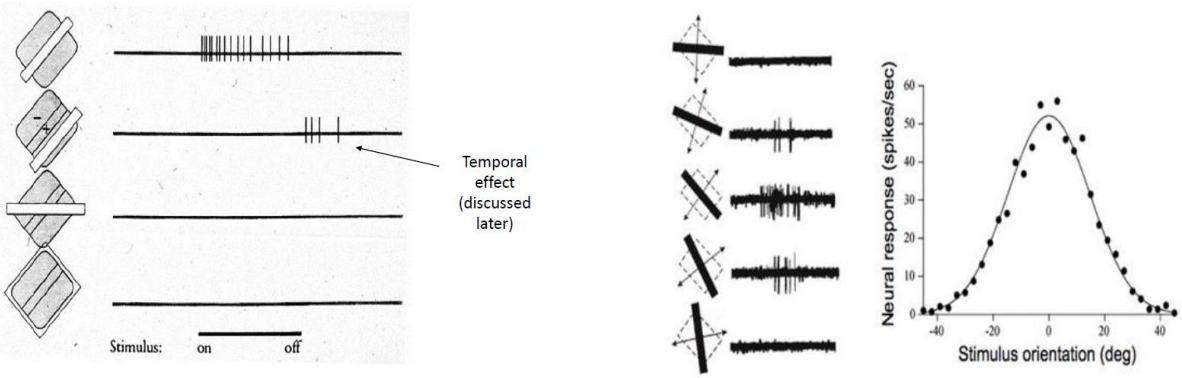
Simple cells

Hubel and Wiesel discovered a number of different types of cells in the primary visual cortex. The type I described above are called *simple cells*. These cells have well defined ON and OFF regions that are elongated such that an oriented bright line can either excite the cell (in the ON region) or inhibit the cell (in the OFF region). I will often call the ON and OFF regions "excitatory" and "inhibitory" respectively. See figure below left. The first shows a white line on the elongated excitory region of the cell. When this white line stimulus is turned on, the cell spiking rate goes up and when the line turns off the cell stops spiking. The second example is more subtle. The line is placed over an inhibitory region. There is no response shown until the white line stimulus turns off,

as if a removal of inhibition acts as an excitation. The models that we will discuss later today do *not* handle these temporal effects. Next week when we discuss motion processing, we will consider temporal effects.

The third example on the left shows a white line of the wrong orientation and the fourth example shows a very thick white bar, the same width as the receptive field. In both of these cases, there is no response from the cell.

The figure on the right is called an *orientation tuning curve*. It shows how *one cell*'s response varies as the orientation of a line varies. (Note that this has a different meaning than saying that a fixed line stimulus produces responses to *different cells* that have the same receptive field position and size and are tuned to different orientations.)



Hubel and Wiesel proposed that simple cells are formed by summing the inputs from a set of center-surround LGN cells whose receptive field centers fall along a line (see slides). It has also been found that simple cells have a large variety of profiles. Some are ON center OFF surround (with orientation preference, as always); others are OFF in the central elongated region and ON in the flanking regions. Still others have an edge like receptive field structure so they are ON on the left side and OFF on the right side, or vice-versa. Finally, simple cells are also sensitive to color. For example, there are double opponent simple cells that might be R+G- on one half of their oriented edge profile and R-G+ on the other half.

As with the DOG functions from last lecture, simple cell receptive field profiles define either positive or negative linear responses, depending on whether the white line stimulus is on the ON or OFF region, respectively. But neurons cannot have negative responses and so a non-linearity must be used to model the negative response e.g. half wave rectification as we discussed last lecture. As long as one has both ON center cells and OFF center cells, one will not lose information because of half wave rectifiation since one of the two will carry any non-zero response.

Gabor model

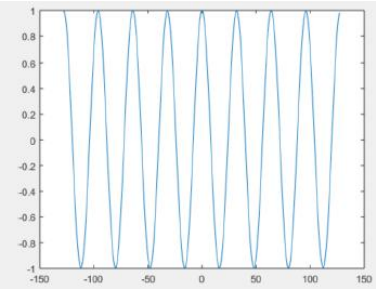
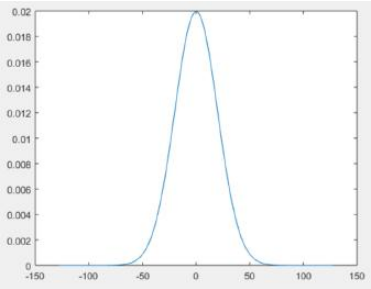
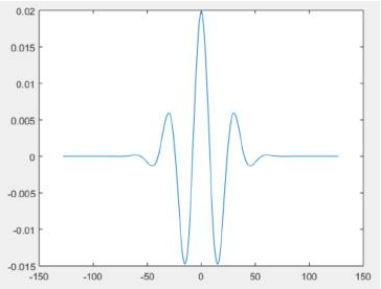
The standard mathematical model of simple cell receptive fields is the Gabor function. Let's define this function first in the 1D case, and then in 2D. Consider a cosine function which is sampled on

a sequence of N uniformly space points

$$\cos(\frac{2\pi}{N}(k_0 x))$$

where k_0 is the spatial frequency which has units of number of cycles per N samples. Typically it is an integer between 0 and $N-1$. Notice that as x goes from 0 to N, the cosine argument goes from 0 to $2\pi k_0$ radians which is indeed k_0 cycles or "times around the circle".

A 1D cosine Gabor is defined by multiplying the cosine function by a Gaussian function of some standard deviation σ. There is no fixed relationship between k_0 and σ. One is free to vary them at will. Increasing σ for a fixed k_0 will increase the number of side lobes. One can define a sine Gabor similarly.

To model the shapes of (2D) simple cell receptive fields, one uses a 2D cosine or sine function and a 2D Gaussian. Consider a 2D cosine function of size $N \times N$,

$$\cos(\frac{2\pi}{N}(k_0 x + k_1 y))$$

where k_0 and k_1 are fixed integers between 0 and $N-1$. This family of 2D cosine functions can define a range of frequencies and orientations. To understand how, note the expression $\frac{2\pi}{N}(k_1 x + k_2 y)$ has a constant value c along a line,

$$\frac{2\pi}{N}(k_0 x + k_1 y) = c.$$

For example, if $c = 0$, the line passes through $(x, y) = (0, 0)$. For different c, one gets different lines and the cosine takes different values. The cosine variation occurs in a direction perpendicular to these lines, namely, in direction (k_0, k_1). One can define a 2D sine function similarly.

Another way to understand 2D sinusoid functions is to note that if you fix x to have a particular value so that you are looking along only a vertical line (column) in the (x, y) domain, then the argument $\frac{2\pi}{N}(k_0 x + k_1 y)$ has k_1 cycles as y goes from 0 to N. Similarly, if you fix x then you are looking along a horizontal line (row) and the argument has k_0 cycles as x goes from 0 to N.

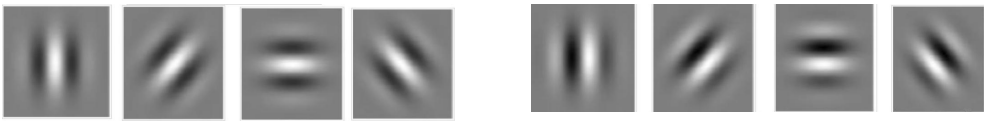
To define a 2D Gabor function, we multiply a 2D cosine function by a 2D Gaussian:

$$cosGabor(x, y, k_0, k_1, \sigma) \equiv G(x, y) \cos(\frac{2\pi}{N}(k_0 x + k_1 y).$$

We define a *sine Gabor* similarly:

$$sinGabor(x, y, k_0, k_1, \sigma) \equiv G(x, y) \sin(\frac{2\pi}{N}(k_0 x + k_1 y).$$

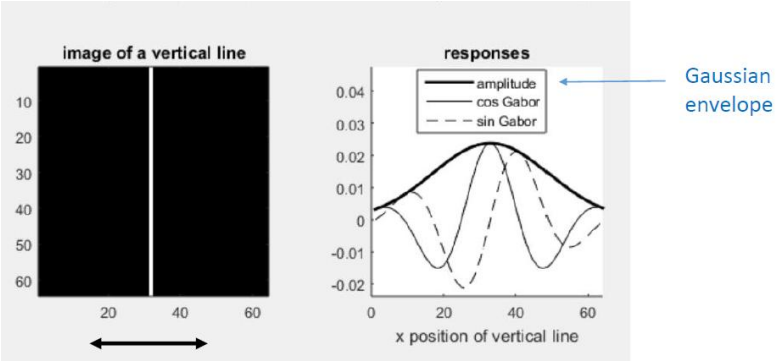
Four examples of cosine (left) and sine (right) Gabors are shown below.



Let's examine how a sine or cosine Gabor cell responds to the position of a line across the receptive field. This is similar to the orientation tuning curve shown above but now we vary the position rather than orientation of the line stimulus. The figure below shows the linear response of a cell as a function of the x position x_{line} of the line, namely the inner product of the Gabor template with the image:

$$< cosGabor(x, y, ...), I(x, y; x_{line}) > \quad = \sum_{(x',y')} cosGabor(x', y', ...) I(x', y'; x_{line})$$

where $I(x', y'; x_{line})$ has value 0 everywhere except on $x = x_{line}$, and the $<\quad>$ notation here is for inner product. Note the response follows the shapes of a 1D sine and cosine Gabor in x. (See Exercises.) The figure below also indicates a "Gaussian envelope". I will discuss this next lecture.



We next examine the response of a family of sine or cosine Gabors to a single image. Here I show just the response to a family of cosine Gabors. (See the slides for responses to a family of sine Gabors.) By "response", I mean the cross correlation of the cosine Gabor and the image. Here are the results for four different cosine Gabors. The filtered image on the upper left shows the results for vertical Gabor.

Notice how the vertical Gabor gives a good response along the pole on the left side of the image, but the details of where the response is a large and positive number (white) versus large and negative (black) vary along the pole. The right diagonal Gabor (top right) picks out the diagonal shadows in the image. Do you understand why the diagonal shadow of the pole is black and flanked by two bright white diagonal regions? (If not, then try to think it through and ask me if you don't get it.) Examine the other two filtered images and identify which parts give a large response.