Today we will examine the measurment of light by photoreceptor cells in the eye. We will look at how sensitive photoreceptors are to different wavelengths of light and also how sensitive they are too different levels of light e.g. night versus day. We will examine how different 'colored' lights can be distinguished, and touch on phenomena such as color blindness.

Light spectra

A good place to start is with Isaac Newton and his prism experiment (late 17th century). Newton observed that when a beam of sunlight is passed through a prism, the beam is spread out into a fan of different color lights – like a rainbow. He argued based on this experiments that light from common radiating source such as the sun or a candle flame is composed of a mixture of colors. The theory that explains Newton's experiments and many other optics experiments has come a long way since then. In a nutshell, we now know that light just is electromagnetic waves, with wavelengths ranging from 400-700 nanometers. (A nanometer is 10^{-9} meters. Thus, you need about 2000 wavelengths of light to extend a distance of one millimeter.)

For any beam of light, we can write the distribution of power in that beam as a function of wavelength. More generally, any function of wavelength can be referred to as a *spectrum*. The light emitted from a source (sun, light bulb, candle) has an *emission* spectrum. A surface that is illuminated has a *reflectance* spectrum, which specifies for each wavelength what is the fraction of light that arrives at the surface that is reflected. Note that light that arrives at a surface doesn't change its wavelength upon reflection. Rather, for each wavelength, some is reflected and the rest is absorbed or transmitted through the medium.

Transmission and absorption spectra are both important in vision. Transmission spectra arise in the context of filters, for example, red and cyan filters that are a cheap way to view 3D images. (More on this later.) Such filters also reflect light and absorb light. Typically we are concerned with how well they transmit light, rather than how much they absorb versus reflect. *Absorption* spectra are especially relevant for understanding photoreceptors, which we discuss next.

Photoreceptors: Rods and cone

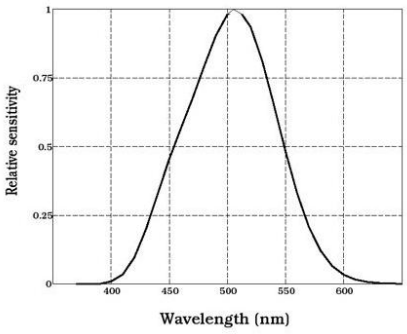
There are two general classes of photoreceptors in the human retina. One class is specialized for discriminating between very low levels of light (night vision). The receptors are long and thin, and are called *rods*. The second class is specialized for discriminating between high light levels (day vision), and also between different spectra. These are called *cones* since their shape is conical.

At very low light levels, namely at night when there is only moonlight, only the rod system is functioning. All rods have the same spectral sensitivity, and so there is no way to compare the spectral distributions of light at two different parts of the retina. Thus at night, we only see shades of grey from black to white. During the day when light levels are high, our rod system shuts down and only the cones are operating, and we can see color. Of course, since level of light is a continuum, there must be some in between levels in which both rods and cones are operating. In these levels (twilight, or night with some artificial light), one can still see in color but not as well as at levels in which the cones are fully operating.

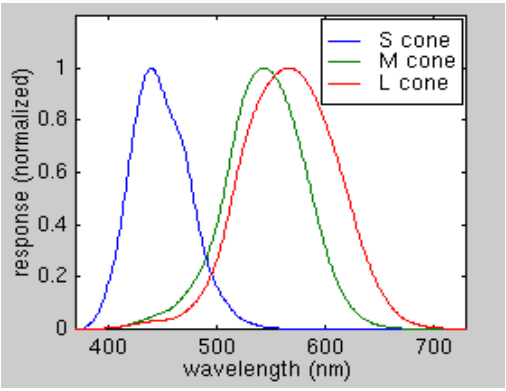
There are three subclasses of *cone* cells, called L, M, and S where L is for long wavelength (red), M is for medium wavelength (green) and S is for short wavelength (blue). Each cone type is defined by a "pigment" (a protein) that absorbs certain wavelengths of light better than others.

The curves below shows how the response (or sensitivity) of the rods and cones depends on wavelength. Each of the curves has been normalized so that its maximum is 1. That is, for each photoreceptor type, there is a certain wavelength for which that photoreceptor type responds best. These curves are called *spectral sensitivity functions*.

Let's just consider the cones for now. The L and M cones have quite similar spectral sensitivities and that the range of wavelengths for L and M is almost non-overlapping with the range for S. This has consequences for how these three "channels" are encoded in subsequent processing stages, which I'll discuss next lecture.



rods



cones

Given some spectrum $E(\lambda)$ of the light arriving at an L, M, or S cone, we can write out a measure of the light absorbed by the photoreceptor as:

$$I_{LMS} = \int C_{LMS}(\lambda)\ E(\lambda) d\lambda$$

which is effectively over 400 to 700 nm since the C functions are non-zero only there.

We can loosely think of this quantity as the responses of L,M, or S cells to the spectrum $E(\lambda)$. But we should keep in mind that the response of a photoreceptor is a complicated thing: it involves changes in membrane potential as well as release of neurotransmittors. One can measure the former but not the latter, and even measuring the former is quite difficult to do.[1] The main idea I want you to get across here is that, for a given tiny local neighborhood on the retina where all three cones are present, there will be triplet of values I_{LMS} that we can associate with the response of the cell. It is the *values* that matter, since our goal here is to build computational models (starting next week).

Let's returning to the above equation. Although λ is a continuous variable, in practice one represents such a spectrum by breaking the interval 400 to 700 nanometers into N_λ bins, for example, 30 bins each having a 10 nanometer range. These bins are small enough and the functions are smooth enough that the functions are approximately constant within each bin. Let E_λ be this spectral intensity function, where now λ is discrete.

[1] Such measurements have been made both in live animals as well as in photoreceptor cells that have been isolated and kept alive.

Similarly, we can characterize the response of the three cone types to this beam of light using a $3 \times N_\lambda$ matrix $\mathbf{C}$ whose rows are C_L, C_M, and C_S, respectively. Each row specifies the relative sensitivity of the cone at each wavelength. i.e. each curve has been normalized to have a maximum value of 1. Therefore, we model the responses of the three cones with a discrete approximation of the integral:

$$\begin{bmatrix} I_L \\ I_M \\ I_S \end{bmatrix} = \begin{bmatrix} C_L \\ C_M \\ C_S \end{bmatrix} E(\lambda) \tag{1}$$

A key implication of this model is that information is lost when a cone measure the light. The discretized spectrum $E(\lambda)$ has dimension N_λ whereas a cone only has one response. The fact that each cone cell's response is just one variable (one dimensional), and all information about the spectrum collapses to that one variable is so important that it is given a name: the *principle of univariance*. The same principle applies to rod cells too, of course.

Metamers and color blindness

One implication of the linear model is that if two spectra $E_1(\lambda)$ and $E_2(\lambda)$ produce the same integrating intensity triplets at a point

$$\mathbf{C}\ E_1(\lambda) = \mathbf{C}\ E_2(\lambda)$$

then these two spectra will be visually indistinguishable. In this case, these two spectra are called *metamers.* Metamers occur often, especially in scenes with many surfaces and different reflectances, but we are (by definition) unaware when they occur.

One important example of metamerism, which we do notice, is color blindness. Many people (2 % of males) are missing a gene for one of the three cone pigments. This leads to three types of "color blindness", depending on which type is missing. "Color blind" doesn't mean the person can't see any colors. Rather, it means that they cannot distinguish some spectra that color normal people can distinguish. Such spectra are metamers for the color blind person.

The model of color blindness follows immediately from the above matrix model. With color blindness, one has only two classes of cones and so the matrix is $2 \times N_\lambda$ rather than $3 \times N_\lambda$. One only has two variables by which spectra can be distinguished rather than three. Many professions do not allow color blindess (police officer, baggage handler, electrician, pilot or driver)[2]

[BEGIN ADDED: Jan. 19]

Another type of color blindness – and indeed a very common one – is that one of the pigments for the three cones has a different spectral absorption than normal. This typically occurs with either L or M cone. For example, the abnormal (anomolous) cone, say L, has an absorption spectrum that is closer to the M's absorption spectrum than a normal person's L cone is. Such a person still has a three dimension color vision (*trichromacy*) but has trouble distinguishing red from green. Such a person is said to be an *anomolous trichromat*. Notice that if the absorption spectrum of the L cone happens to be very similar to that of the normal M cone, then such a person is essential a *dichromat* i.e. having just two cone types (M and S).

[2] `http://wereadbetter.com/7-jobs-that-you-are-prohibited-from-with-colorblindness/`

In case you wish to read more on this, here is some of the basic terminology. A person who is missing one type of cone is said to have _____*anopia* where the prefix to fill the underline specifies which of the three cones is *missing*: *prot* for L, *deuter* for M, or *trit* for S which are Greek roots for first, second, third. So, for example, a person missing the S cone is said to have tritanopia. A person who has *abnormal* cone absorptivity is said to have a _____*anomaly*. So, for example, someone with an abnormal L is said to have a protanomaly.

One student asked in class what the term *red green color blindness* referred to. It refers to any type of problem with the L or M cones, namely one could be either missing or just anomolous. Problems with the L or M cones are much more common than problems with the S cones. Another student asked if some people are missing two of the three cones. I looked it up and the answer is yes, but it is very rare.

[END ADDED: Jan. 19]

Rod vision is an extreme case of metamerism. In sufficiently dim conditions in which the cones are not operating, one no longer perceives color. One does still perceive shades of gray though. We would say that two surfaces that are placed side by side and that produce the same rod response levels would be metameric. Note that rods are most sensitive to wavelengths in the middle of the spectrum, which we roughly associate with say green. This does not mean that the world at night looks green. Rather, it means that if you have red, blue and green objects that appear roughly equally bright during the day, then the green object will appear brighter at night.[3]

Color displays

One type of spectrum where this theory finds an application is electronic color displays (projectors, computer monitors, TVs, cell phones). Is it easy to characterize the spectra of light coming from each pixel of a display by adding together the three spectra that are determined by the RGB values of that pixel. More precisely, let the spectra of light emitted by each of the RGB color elements of a display be represented by the columns of an $N_\lambda \times 3$ matrix $\mathbf{P}$. (For old TVs, $\mathbf{P}$ stood for "phosphor".) Let $\mathtt{e}$ be a 3×1 vector that specifies a scalar weight for each spectra. So the spectra $E(\lambda)$ that results from a pixel can be written as the following weighted sum:

$$E(\lambda) = \mathbf{P}\ \mathbf{e}$$

For simplicity,[4] let's take $\mathbf{e}$ to be the RGB values in [0,1] at a pixel. In the slides I just wrote RGB instead of $\mathbf{e}$.

What is the set of LMS values that can be produced by such a color display ? If we let the three components of $\mathbf{e}$ be in [0,1], then we can look at how those $\mathbf{e}$ vectors map to triplets of intensities absorbed by the LMS cones:

$$I_{LMS} = \mathbf{C}\ E(\lambda) = \mathbf{C}\ \mathbf{P}\ \mathbf{e}$$

The matrix product $\mathbf{C}\ \mathbf{P}$ is a 3×3 matrix which maps from the unit cube of $\mathbf{e}$ values to LMS space.

[ASIDE: I did not mention this in the lecture but ... from basic linear algebra, we can see that the three columns represent respectively the LMS coordinates of the R,G, and B emitters

[3] There is a technical sense in which we can compare the brightness of a red versus green or blue object, but we don't have the tools for explaining that yet in the course.

[4] I say "for simplicity" because usually there is also a non-linear transformation called *gamma* from the RGB value to the $\mathbf{e}$ value.

(on maximum intensity, i.e. value 1). According to this simple model (which is basically correct, ignoring the issue of monitor gamma), the LMS triplets that can be reached are a distorted cube in LMS space, namely the linear transformation of the points in the unit cube in RGB space.]

Transmission spectra, and anaglyphs

An interesting example in which transmission spectra matter is the case of colored glass or plastic. One vision application is *anaglyph images* which can be used to produce perception of 3D. Anaglyphs are composed of a pair of grey level images that are presented in the different color channels. For example, one image might be presented in the R channel only with R having some value ρ that varies across the image, and the other image might be presented in the G and B channels with value ψ that varies across the image.

The key idea of 3D stereo using anaglyphs to film (or photograph, in the case of a still image) a scene from two neighboring camera positions. Then, when presenting the scene as an image as described briefly above, place color filters in front of each eye that will only let the light from one of the two images through. Typically anaglyph glasses have a red filter over the left eye and a cyan filter over the right eye, so the left eye will see the $(\rho, 0, 0)$ red image and the right eye will see the $(0, \psi, \psi)$ cyan image, where ρ and ψ will vary with position in the image. This gives a 3D effect since the images that reach the left and right eye correspond to the images that were capture by a left and right camera in a 3D scene – namely the binocular disparities are consistent with that 3D scene. We will discuss this again later. In the meantime, see the example in the slides and see the exercises.

Temporal effects on image measurement

As mentioned earlier, we are being somewhat loose in refering to the response of a photoreceptor as the meaurable potential difference across the cell membrane, as there is more going on than that. But we'll work with that loose definition anyhow.

Let's briefly discuss the *temporal* properties of the cell response to light. If you flash a pulse of light briefly on a photoreceptor, it doesn't respond instantly but rather it takes several milliseconds to respond and then the response continues. The membrane potential decreases (becomes more negative – see slide) for a short time and then climbs back to its resting state. The magnitude of the response (size of the potential drop) and the duration of response will depend on the length and magnitude of the pulse of light that was used. Not surprisingly, a longer duration and higher magnitude pulse will produce a greater response.

There is another factor that determines the response of a photoreceptor and that is the intensity of the light over the recent past, which affects the current state of the cell. If the cell was continuously exposed to a bright light for a few seconds or even minutes and then it was stimulated with the pulse mentioned above, it will have less of a response to that pulse than if the cell was exposed to darkness in the previous several minutes before the pulse. The main concept here is that, at any time, the cell will have some operating range over which its response depends on the brightness of a short pulse of light. If the brightness of the pulse is too low, there will be no measurable response. If the brightness is too high, then the response will max out. (One refers to the response as being saturated.) Often the response obeys a sigmoid (S) shaped curve. This response curve itself will shift as the background level of light changes.

Camera's also have this sigmoid shaped response curve. For any camera setting, if the image captured is too dark (e.g. because the scene is dark or because the exposure time is too short) then the image will have RGB values of 0. If the scene is too bright, then it will have values of 255 (maximum 8 bit value). There is some operating range in the middle in which the camera will measure distinct image intensities in each RGB channel. Indeed part of the technical challenge of photography is choosing the camera settings so that you don't have too many 0 or 255 values.

Getting back to photoreceptors and vision, we refer to the shifting of the response as *adaptation*. Adaptation occurs not just in the photoreceptors; it occurs throughout the vision system and indeed throughout all sensory systems. I presented an example in the lecture slides of how we adapt to a white or black square on a grey background. If you look at the dot between the two squares for say 30 seconds and then you look to the right, you will see a blurry black and white square in the visual direction where the white and black square were, respecively. Roughly what is happening here is that the cells that encode for those parts of the image are adjusting their "code" for what is dark versus light. The part of the image that adapted to the black square is now procssing grey, and so that part of the visual field looks brighter, since grey is brighter than black. Similarly, the part of the image that adapted to the white square is now procssing grey, and so that part of the visual field looks darker, since grey is darker than white.

What's happening here is that your visual system does not just provide information about what is out there. It also (simultaneously as part of the its "code") provides information about *changes* in what was out there. While the visual system obviously makes mistakes in judging brightness, the visual system seems to have an overall benefit in adapting because it allows us to move our eyes from dark parts of scenes to bright parts of scenes and adjust our operating range. The adjustments can occur not just at times scales of seconds, but rather the adjustments can occur over minutes. For example, when we walk from the bright sunlight into a cave, the drop in intensity can easily be a factor of one million or more. This extreme adaptation is handled by more than just shifting the operating range of photoreceptions; it is handled by switching from the cone system to the rod system.

Yet another form of adaptation is the pupillary response. If the overall light level in a scene goes up suddenly or if you look at a brighter region of the scene – then your pupil may shrink to reduce the amount of light to come in. This provides a *global* adaptation, which is different from the *local* adaptation for the squares above. Note that the diameter of the pupil can range from say 2 mm up to say 8 mm. Considering that the area grows like diameter squared, the changing pupil size can lead to a roughly factor of 16 range of intensity of light reaching the retina.

Next lecture I will finish up the discussion of photoreceptors and then we will move on to other parts of the retina.

