Vision and Color

Brian Curless
CSEP 557
Autumn 2017

Reading

Good resources:

Glassner, *Principles of Digital Image Synthesis*, pp. 5-32.

Palmer, *Vision Science: Photons to Phenomenology*.

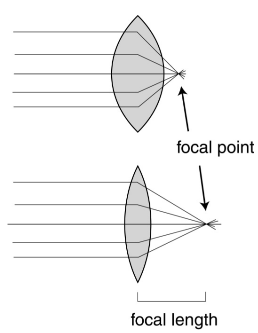
Wandell. *Foundations of Vision.*

Lenses

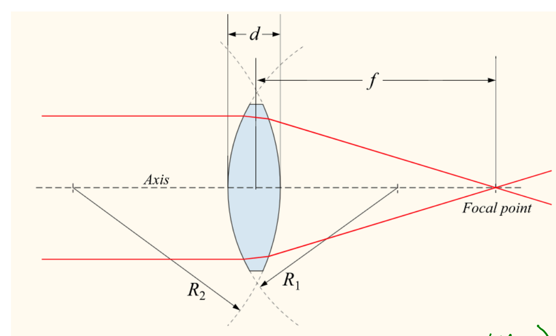
The human eye employs a lens to focus light.

To quantify lens properties, we'll need some terms from *optics* (the study of sight and the behavior of light):

- **Focal point** - the point where parallel rays converge when passing through a lens.
- **Focal length** - the distance from the lens to the focal point.



The lensmaker's formula



Positive (converging) lens

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

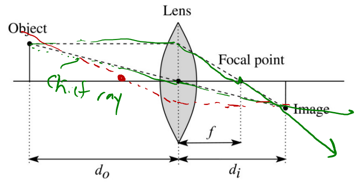
$d \approx 0$

(thickness is ≈ 0)

$$f = \frac{1}{n-1}\frac{R_1R_2}{R_2-R_1}$$

Optics, cont'd

By tracing rays through a lens, we can generally tell where an object point will be focused to an image point:



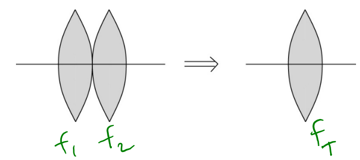
This construction leads to the Gaussian lens formula:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

Compound lenses

A compound lens is a sequence of simple lenses.

When simple, thin lenses are stacked right next to each other, they focus much like a single lens. We can compute the focal length of the resulting compound lens as follows:



$$\frac{1}{f_T} = \frac{1}{f_1} + \frac{1}{f_2}$$

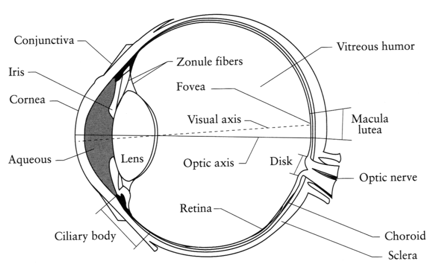
$$f_T = \frac{f_1 f_2}{f_1 + f_2}$$

It is convenient to define the **diopter** of a simple lens as the reciprocal of the focal length (in meters), 1/*f*.

Example : A lens with a "power" of 10D has a focal length of 0.1m.

Why is using diopters (1/*f*) convenient? add them up

Structure of the eye

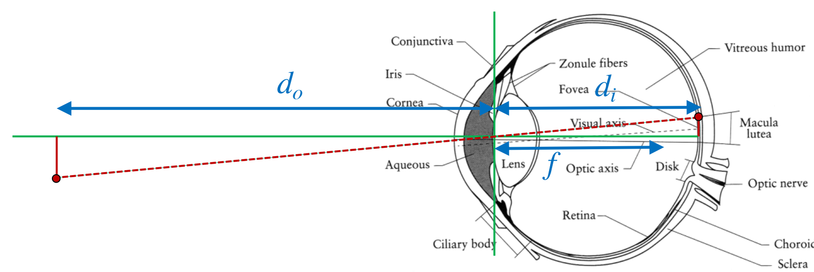


Physiology of the human eye (Glassner, 1.1)

The most important structural elements of the eye include:

- **Cornea** - a clear coating over the front of the eye:
  - Protects eye against physical damage.
  - Provides initial focusing (40D).
- **Crystalline lens** – provides additional focusing
- **Retina** – layer of photosensitive cells lining the back of the eye.

Structure of the eye



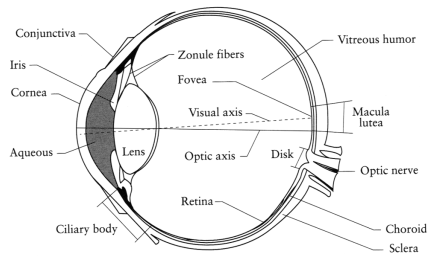
We can treat the cornea + crystalline lens as a compound lens, which roughly follows the Gaussian lens formula. Again, this is:

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}$$

Q: Given the three parameters (d_o, d_i, and f), how does the human eye keep the world in focus?

Change f

Structure of the eye, cont.

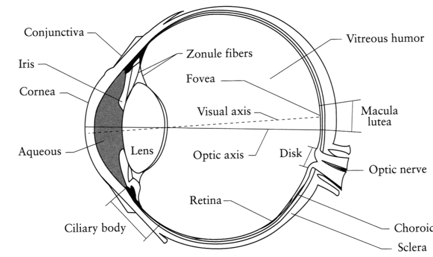


Physiology of the human eye (Glassner, 1.1)

- **Crystalline lens** - controls the focal distance:
  - Power ranges from 10 to 30D in a child.
  - Power and range reduces with age.
  - Focusing with the lens is called **accommodation**.
- **Ciliary body** - The muscles that compress the sides of the lens, controlling its power.

Q: As an object moves closer, do the ciliary muscles contract or relax to keep the object in focus?

Structure of the eye



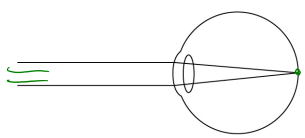
Physiology of the human eye (Glassner, 1.1)

The remaining important elements are:

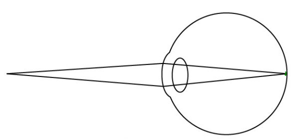
- **Iris** - Colored annulus with radial muscles.
- **Pupil** - The hole whose size is controlled by the iris.

The iris adjusts the size of the pupil according to the light levels in front of the subject.

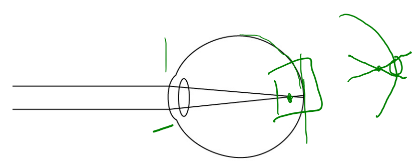
Eye geometry



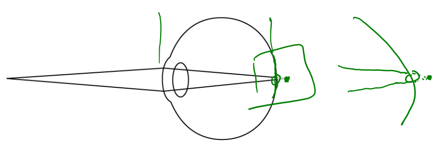
Emmotropic (focused at infinity)



Emmotropic (focused in close)



Myopic (unable to focus at infinity)



Hyperopic (unable to focus in close)

Eye geometry can account for near- and far- sightedness.

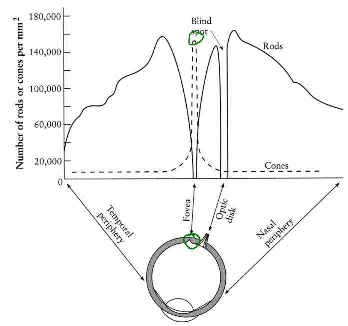
- **Emmetropic eye** - resting eye has focal point on retina.
- **Myopic eye** - eye too long (near-sighted).
- **Hyperopic eye** - eye too short (far-sighted).

Near- and far-sightedness can also result from deficiencies in focusing at the cornea or through the lens.

Presbyopia is loss of flexibility in the lens, reducing up-close focusing power. This happens naturally with age.

Q: Myopia and hyperopia are worse under low light. Why?

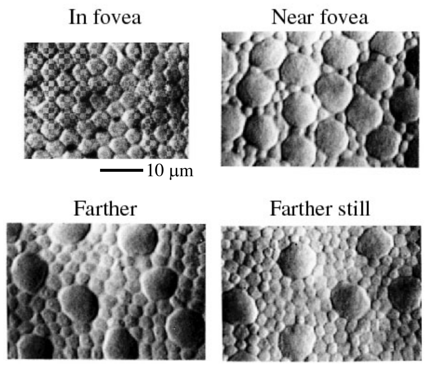
Retina



Density of photoreceptors on the retina (Glassner, 1.4)

- **Retina** - a layer of photosensitive cells covering 200° on the back of the eye.
  - **Cones** - responsible for color perception.
  - **Rods** - Limited to intensity (but 10x more sensitive).
- **Fovea** - Small region (1 or 2°) at the center of the visual axis containing the highest density of cones (and no rods).

The human retina

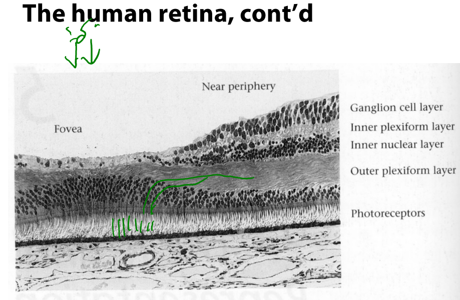


Photomicrographs at increasing distances from the fovea. The large cells are cones; the small ones are rods. (Glassner , 1.5 and Wandell, 3.4).

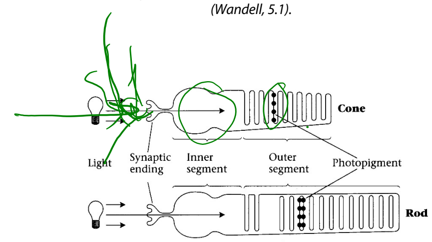
Photomicrographs at increasing distances from the fovea. In the fovea, all the cells are cones and are small and tightly packed.

Toward the periphery, there are fewer and fewer cones. The large cells are cones, and the small ones are rods, in the non-fovea figures above.

The human retina, cont'd



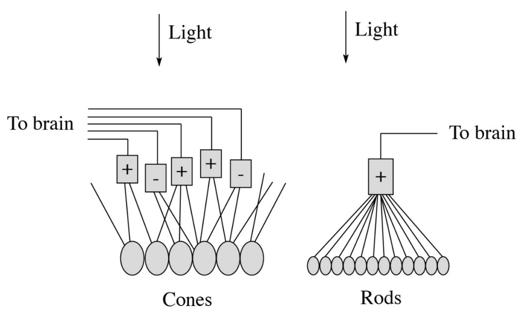
Photomicrograph of a cross-section of the retina near the fovea (Wandell, 5.1).



Light gathering by rods and cones (Wandell, 3.2)

Neuronal connections

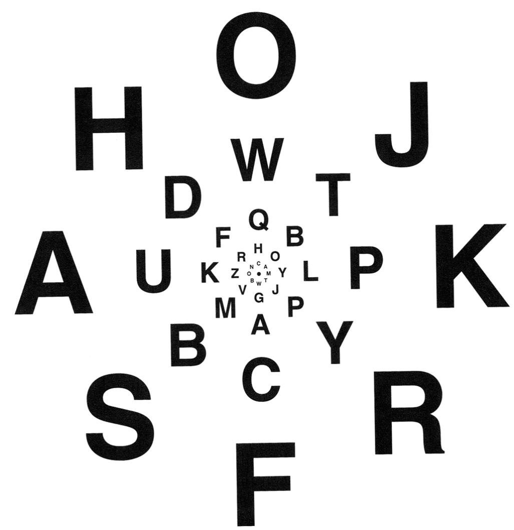
Even though the retina is very densely covered with photoreceptors, we have much more acuity in the fovea than in the periphery.



In the periphery, the outputs of the photoreceptors are averaged together before being sent to the brain, decreasing the spatial resolution. As many as 1000 rods may converge to a single neuron.

Accuity across visual field

With one eye shut, look at the center dot with the other eye. At the right distance, all of these letters should appear equally legible (Glassner, 1.7).



Blind spot

Close your left eye and focus on the "+" with your right eye. At the right distance with the right head rotation, the black dot disappears.

High resolution imaging?

Given that our vision is only high resolution over a very small range of our visual field…

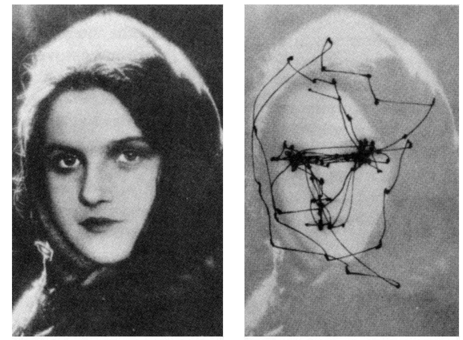
…how do we manage to see "everything" at high resolution?

Fixations and saccades

By scanning your eyes over a scene, you build a composite, high resolution image in our brain.

Fixations: our eyes pause at certain location to see the detail; these pauses are called **fixations**.
Saccades: between fixations, we scan rapidly with very jittery motion.

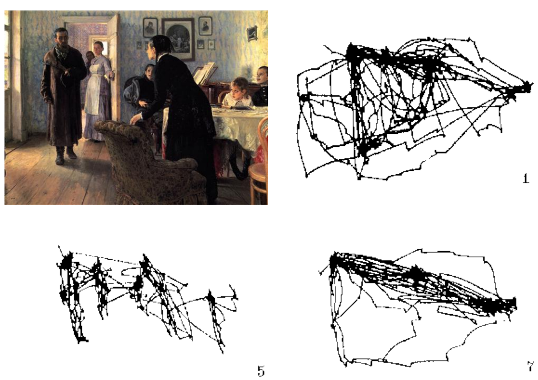
Through gaze tracking, scientists can study how we look at the world.



Yarbus, 1965

Saccades, cont'd

The saccadic behavior is task-specific:



Yarbus, 1965

1. Free examination.

5. Remember the clothes worn by the people

7. Estimate how long the "unexpected visitor" had been away from the family

Perceptual light intensity

The human eye is highly adaptive to allow us a wide range of flexibility.

One consequence is that we perceive light intensity as we do sound, I.e., on a *relative* or *logarithmic* scale.

Example: The perceived difference between 0.20 and 0.22 is the same as between 0.80 and 0.88.

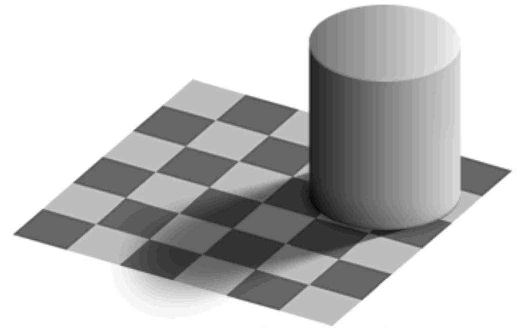
A related phenomenon is **lightness constancy**, which makes a surface look the same under widely varying lighting conditions.

Lightness contrast

The apparent brightness of a region depends largely on the surrounding region.

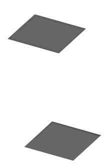
The **lightness contrast** phenomenon makes a constant luminance region seem lighter or darker depending on the surround:

Lightness contrast and constancy



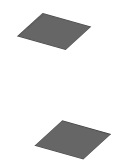
Checker Shadow Effect (Edward Adelson, 1995)

Lightness contrast and constancy



Checker Shadow Effect (Edward Adelson, 1995)

Lightness contrast and constancy



Checker Shadow Effect (Edward Adelson, 1995)

Adaptation

Adaptive processes can adjust the base activity ("bias") and scale the response ("gain").

Through **adaptation**, the eye can handle a large range of illumination:

Background	Luminance (cd/m²)
Moonless overcast night	0.00003
Moonlit covercast night	0.003
Twilight	3
Overcast day	300
Day with sunlit clouds	30,000

Some of our ability to handle this range comes from our ability to control the iris (aperture) of our eyes, and the fact that we have different types of photoreceptors.

However, much of the range comes from the adaptability of the photoreceptors themselves. This photoreceptor adaptation takes time, as you notice when going between very bright and very dark environments.

Mach bands

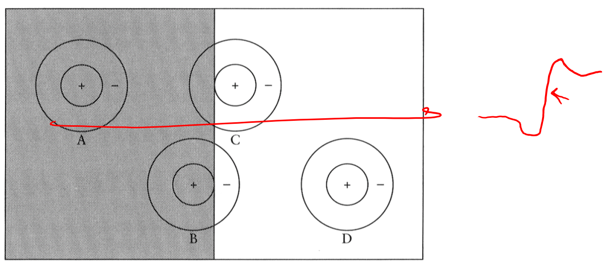
Mach bands were first dicussed by Ernst Mach, an Austrian physicist.

Appear when there are rapid variations in intensity, especially at C^0 intensity discontinuities:

And at C^1 intensity discontinuities:

Mach bands, cont.

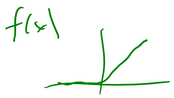
Possible cause: lateral inhibition of nearby cells.



Lateral inhibition effect (Glassner, 1.25)

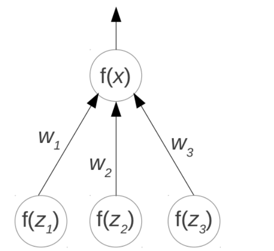
Q: What image processing filter does this remind you of?

Laplacian sharpening



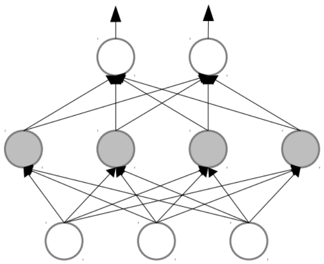
Neural Networks

- A neuron



$x = w_1 f(z_1) + w_2 f(z_2) + w_3 f(z_3)$

- A neural network



Convolution

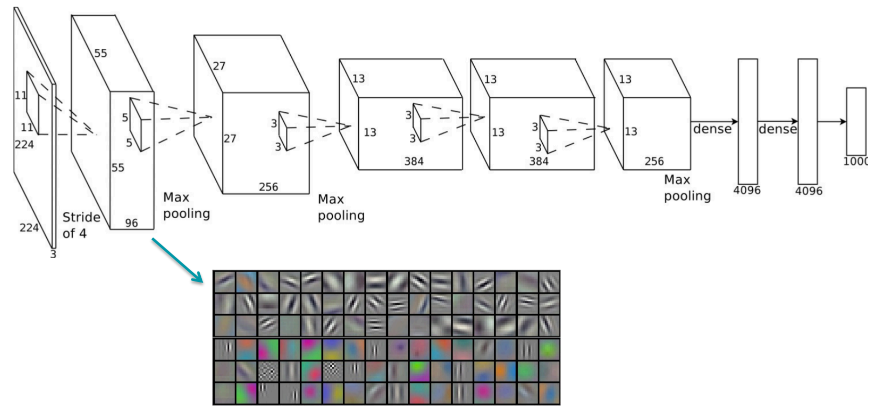
Image

128	54	9	78	100
145	98	240	233	86
89	177	246	228	127
67	90	255	148	95
106	111	128	84	172
221	154	97	69	94

Filter

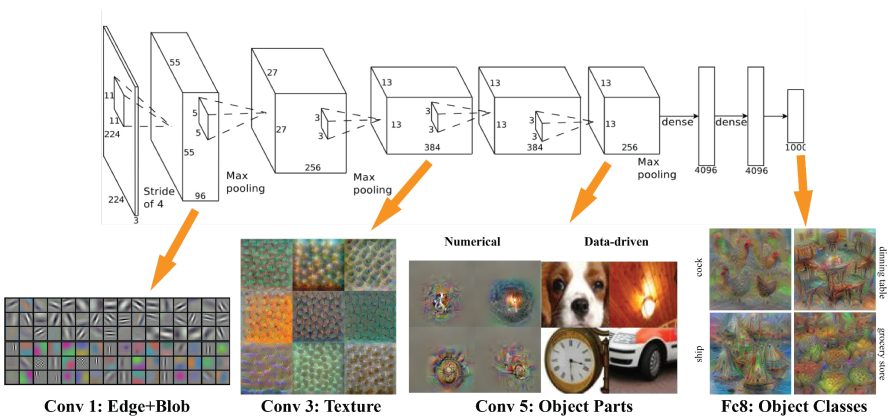
X 0.1	X 0.1	X 0.1
X 0.1	X 0.2	X 0.1
X 0.1	X 0.1	X 0.1

Convolutional Neural Network (CNN)



http://vision03.csail.mit.edu/cnn_art/index.html

Convolutional Neural Network (CNN)



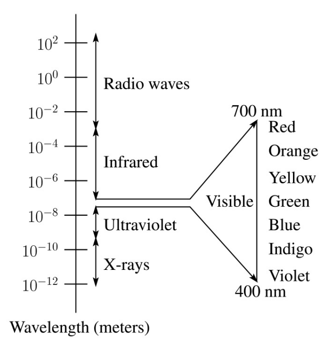
http://vision03.csail.mit.edu/cnn_art/index.html

Are we learning the filters and neural behavior of the human visual system? Maybe a little, a work in progress…

The radiant energy spectrum

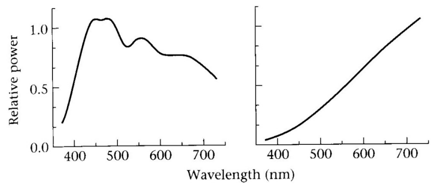
We can think of light as waves, instead of rays.

Wave theory allows a nice arrangement of electromagnetic radiation (EMR) according to wavelength:



Emission spectra

A light source can be characterized by an emission spectrum:



Emission spectra for daylight and a tungsten lightbulb (Wandell, 4.4)

The spectrum describes the energy at each wavelength.

What is color?

The eyes and brain turn an incoming emission spectrum into a discrete set of values.

The signal sent to our brain is somehow interpreted as *color*.

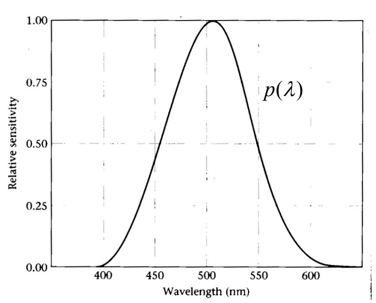
Color science asks some basic questions:

- When are two colors alike?
- How many pigments or primaries does it take to match another color?

Photopigments

Photopigments are the chemicals in the rods and cones that react to light. Can respond to a single photon!

Rods contain **rhodopsin**, which has peak sensitivity at about 500nm.

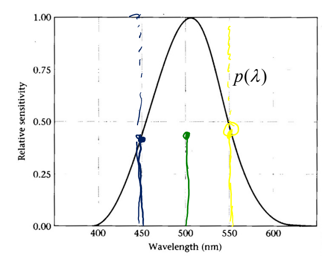


Rod sensitivity (Wandell ,4.6)

Rods are active under low light levels, i.e., they are responsible for **scotopic** vision.

What rods measure

A rod responds to a spectrum through its spectral sensitivity function, $p(\lambda)$.



The response to a test light, $t(\lambda)$, is simply:

$$P = \int t(\lambda)\, p(\lambda)\, d\lambda$$

Suppose we illuminate a rod with two different spotlights, one after the other:

455 nm blue laser of amplitude 1.0

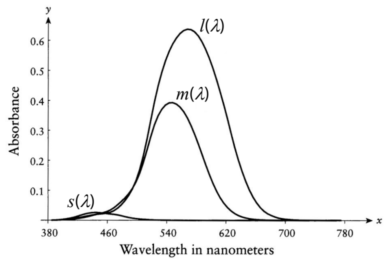
550 nm yellow laser of amplitude 1.0

Will these spots look different? No

What about a 500 nm green laser half as bright? Same

Cone photopigments

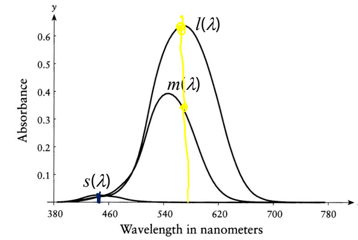
Cones come in three varieties: L, M, and S.



Cone photopigment absorption (Glassner, 1.1)

Cones are active under high light levels, i.e., they are responsible for **photopic** vision.

What cones measure



Color is perceived through the responses of the cones to light, written simply as:

$$L = \int t(\lambda) l(\lambda) d\lambda$$

$$M = \int t(\lambda) m(\lambda) d\lambda$$

$$S = \int t(\lambda) s(\lambda) d\lambda$$

Now suppose we illuminate a cone with two different spotlights, one after the other:

455 nm blue laser of amplitude 1.0

550 nm yellow laser of amplitude 1.0

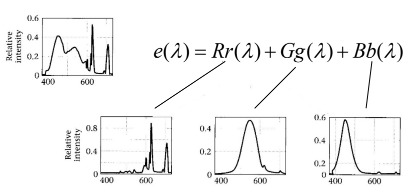
Will these spots look different? Yes

Primaries

Ultimately, the sensation of color happens by generating L, M, and S responses.

With three primaries (e.g., monochromatic red, green, blue laser light), we can adjust the power knobs on the lights and cause a wide range of L, M, and S responses.

In general, the primaries can be non-monochromatic, e.g., monitor phosphors from an old CRT:

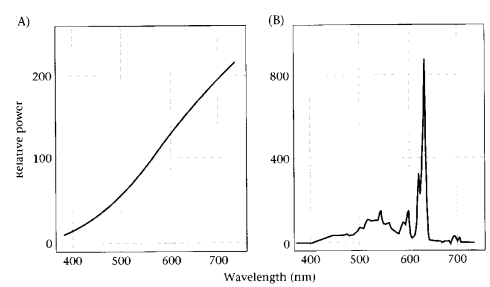


Emission spectra for RGB monitor phosphors (Wandell B.3)

Emission Spectrum is not color

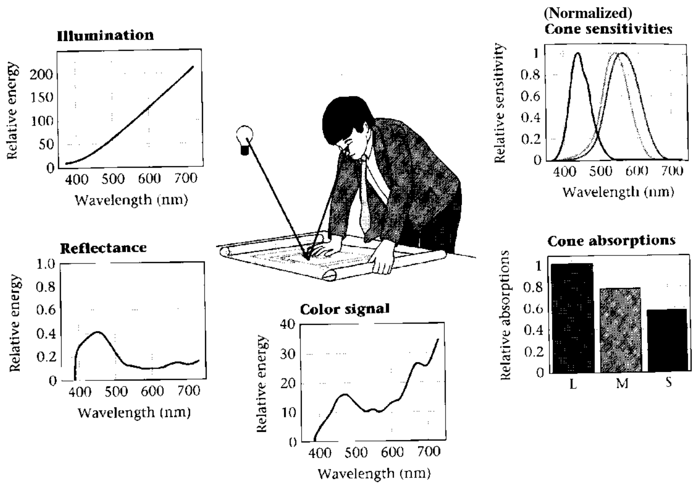
Although the cones give us some ability to distinguish some different spectra, they still convert every continuous spectrum into just three numbers – much information is lost!

Indeed, many different light sources can evoke exactly the same colors. Such lights are called **metamers**.



A dim tungsten bulb and an RGB CRT monitor set up to emit a metameric spectrum (Wandell 4.11)

Color Appearance of Light Reflection

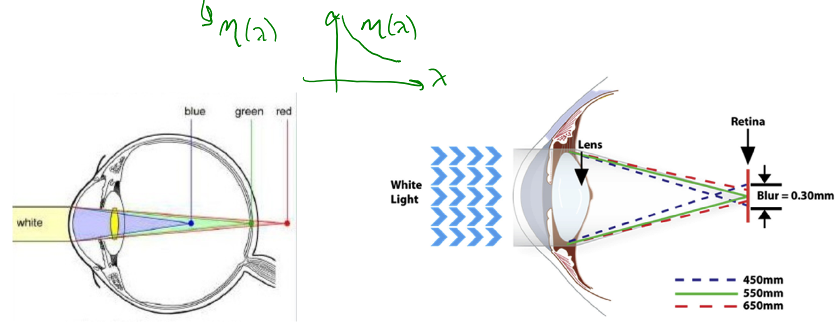


How light and reflectance become cone responses (Wandell, 9.2)

Chromatic aberration in the eye

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$n(\lambda)$

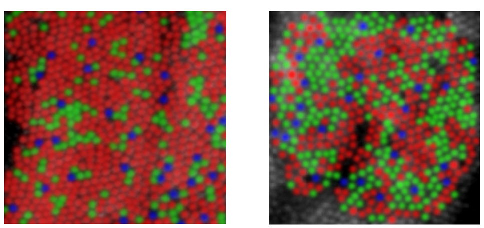


Cone distribution

How are cones distributed in the retina?

Is it about the same for everyone?

Here are images of near-fovea regions for two different human subjects, with colors to indicate the L (red), M (green) and S (blue) cones:



http://roorda.vision.berkeley.edu/ao_res.htm

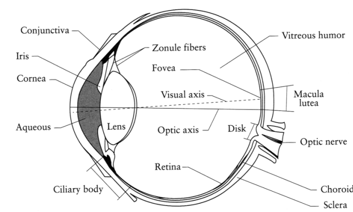
Remarkably, both subjects have normal color vision!

Note how there are very few S (blue) cones.

What does this mean for our ability to see blue things with high visual detail?

Human vision, perspective, and 3D

The human visual system uses a lens to collect light more efficiently, but records perspectively projected images much like a pinhole camera.



[Glassner, 1995]

Q: Why did nature give us eyes that perform perspective (and not orthographic) projections?

Q: Do our eyes “see in 3D”?