Vision and Color

Brian Curless
CSE 557
Autumn 2014

Reading

Good resources:

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

Palmer, *Vision Science: Photons to Phenomenology*.

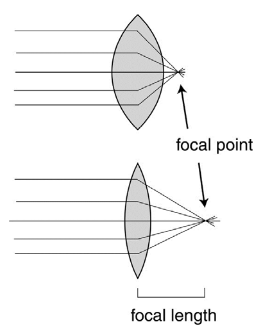
Wandell. *Foundations of Vision.*

Lenses

The human eye employs a lens to focus light.

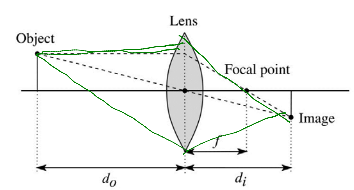
To quantify lens properties, we'll need some terms from *optics* (the study of sight and the behavior of light):

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



Optics, cont'd

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



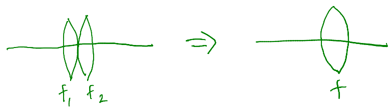
This construction leads to the Gaussian lens formula:

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

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 single lens. We can compute the focal length of the resulting compound lens as follows:



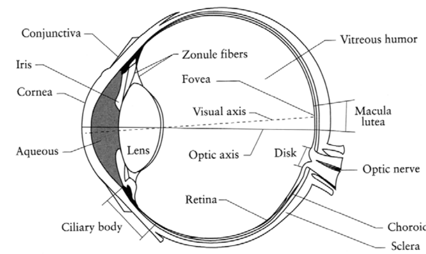
$$\frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f} \qquad f = \frac{f_1 f_2}{f_1 + f_2}$$

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 this convenient? Add diopters for compound lenses (~ in contact)

Structure of the eye

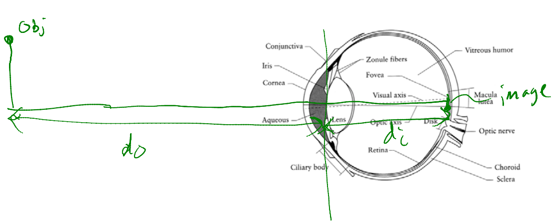


Physiology of the human eye (Glassner, 1.1)

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



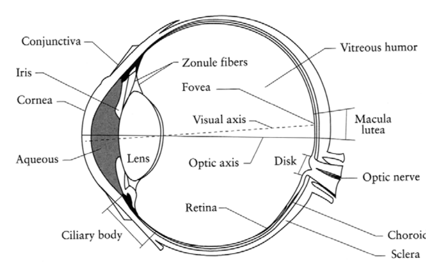
We can treat the cornea + crystalline lens as a compound lens, which roughly follows the Gaussian lens formula. Again, this is:

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

Q: Given these three parameters, how does the human eye keep the world in focus?

Change f

Structure of the eye, cont.

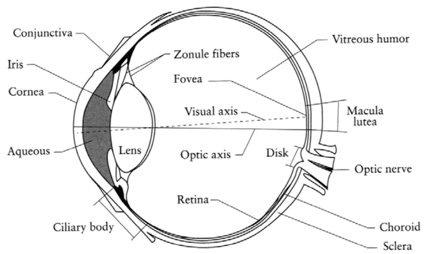


Physiology of the human eye (Glassner, 1.1)

- **Crystalline lens** - controls the focal distance:
  - Power ranges from 10 to 30D in a child.
  - Power and range reduces with age.
- **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? Contract

Structure of the eye



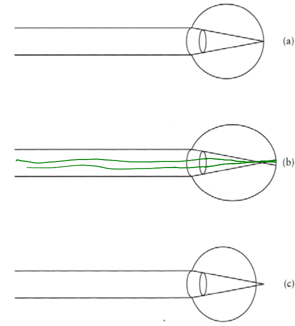
Physiology of the human eye (Glassner, 1.1)

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



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

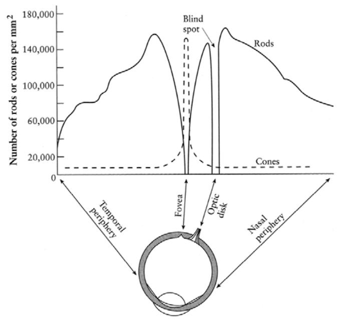
- **Emmetropic eye** - resting eye has focal point on retina.
- **Myopic eye** - eye too long (near-sighted).
- **Hyperopic eye** - eye too short (far-sighted).

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? Blur ~ aperture diameter

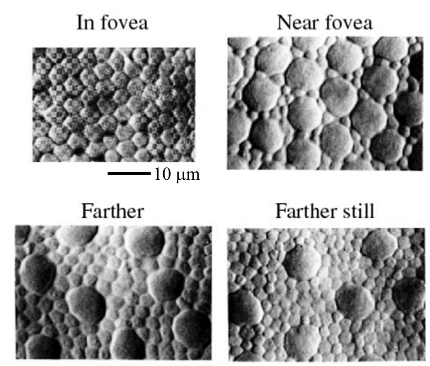
Retina



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

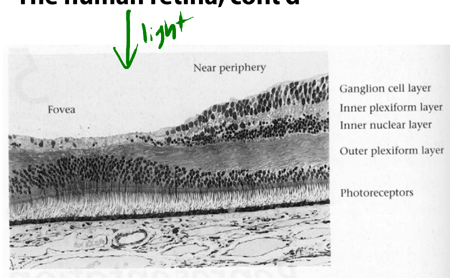


Photomicrographs at incresasing distances from the fovea. The large cells are cones; the small ones are rods. (Glassner , 1.5 and Wandell, 3.4).

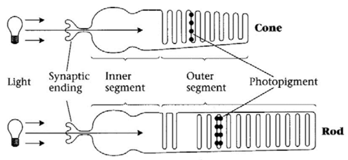
Photomicrographs at increasing distances from the fovea. In the fovea, all the cells are cones and are small and tightly packed.

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



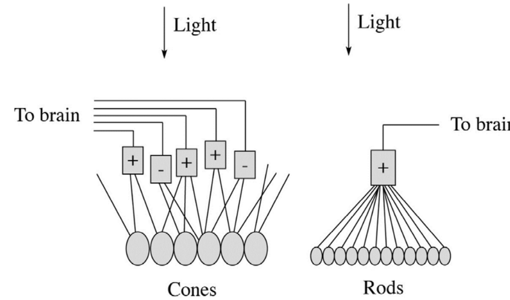
Photomicrograph of a cross-section of the retina near the fovea (Wandell, 5.1).



Light gathering by rods and cones (Wandell, 3.2)

Neuronal connections

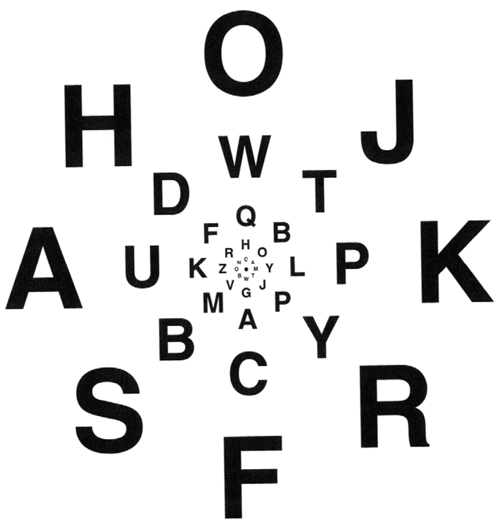
Even though the retina is very densely covered with photoreceptors, we have much more acuity in the fovea than in the periphery.



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…

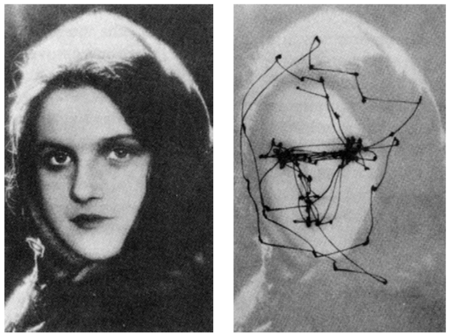
…how do we manage to see "everything" at high resolution?

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.

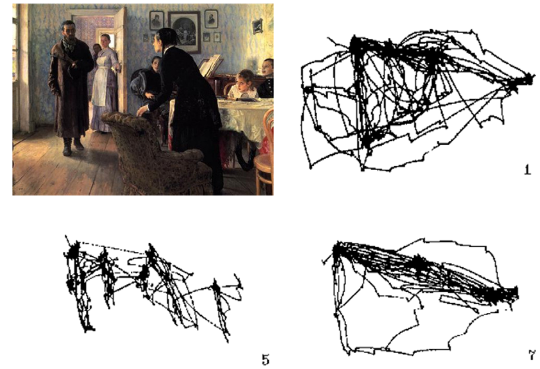
Through gaze tracking, scientists can study how we look at the world.



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.

$\frac{I_2}{I_1}$ $\frac{10 I_2}{10 I_1}$

$\log I_2 - \log I_1$

$\log(10 I_2) - \log(10 I_1)$

$\log I_2 + \log 10 - \log I_1 - \log 10$

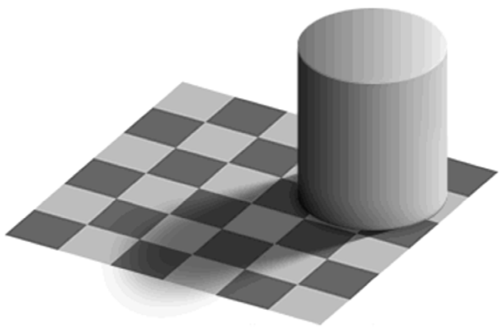
$\log I_2 - \log I_1$

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)

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/m^2)**
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.

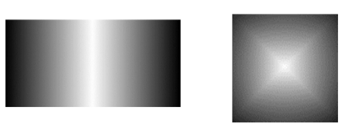
However, much of the range comes from the adaptability of the photoreceptors themselves. This photoreceptor adaptation takes time, as you notice when going between very bright and very dark environments.

Mach bands

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

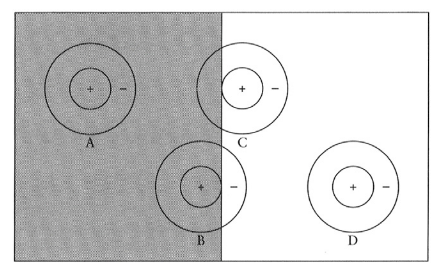
Appear when there are rapid variations in intensity, especially at C^0 intensity discontinuities:

And at C^1 intensity discontinuities:



Mach bands, cont.

Possible cause: lateral inhibition of nearby cells.



Lateral inhibition effect (Glassner, 1.25)

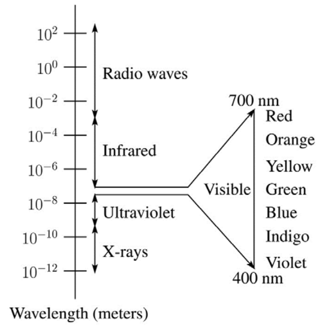
Q: What image processing filter does this remind you of?

~ Laplacian sharpen

The radiant energy spectrum

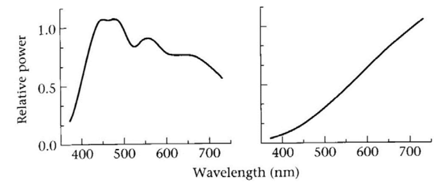
We can think of light as waves, instead of rays.

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:

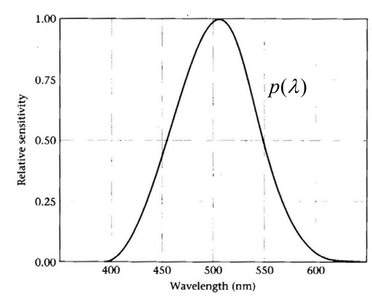
- When are two colors alike?
- How many pigments or primaries does it take to match another color?

One more question: why should we care?

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.

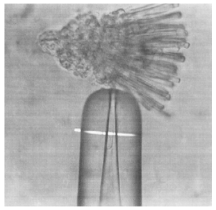


Rod sensitivity (Wandell ,4.6)

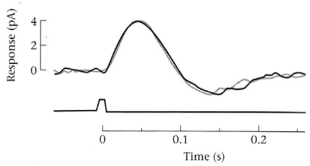
Rods are active under low light levels, i.e., they are responsible for **scotopic** vision.

Univariance

Principle of univariance: For any single photoreceptor, no information is transmitted describing the wavelength of the photon.



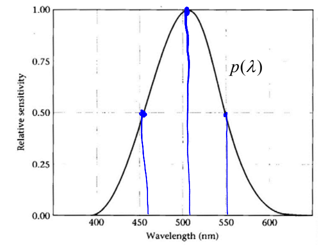
Measuring photoreceptor photocurrent (Wandell, 4.15)



Photocurrents measured for two light stimuli: 550nm (solid) and 659 nm (gray). The brightnesses of the stimuli are different, but the shape of the response is the same. (Wandell 4.17)

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:

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

Suppose a rod sees three light spots:

455 nm blue laser of amplitude 1.0 → 0.5

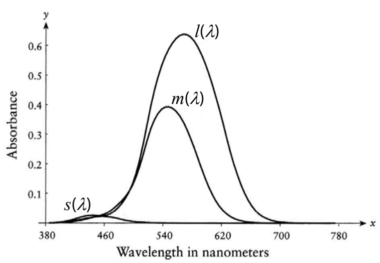
505 nm green laser of amplitude 0.5 → 0.5

550 nm yellow laser of amplitude 1.0 → 0.5

Will these spots look different? No

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.

The response of each cone can be written simply as:

$$L_t = \int t(\lambda)\,l(\lambda)\,d\lambda$$

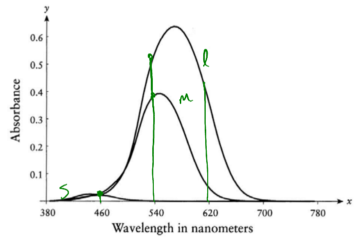
$$M_t = \int t(\lambda)\,m(\lambda)\,d\lambda$$

$$S_t = \int t(\lambda)\,s(\lambda)\,d\lambda$$

These are the only three numbers used to determine color.

What cones measure

Consider the sensitivity spectra again:



Suppose we show three light spots with unit intensity lasers at 460nm, 540nm, and 620nm. What will the cones measure?

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.025 \\ 0.025 \\ 0.025 \end{bmatrix} \quad \begin{bmatrix} 0.55 \\ 0.4 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0.4 \\ 0.05 \\ 0 \end{bmatrix}$$

460nm 540nm 620nm

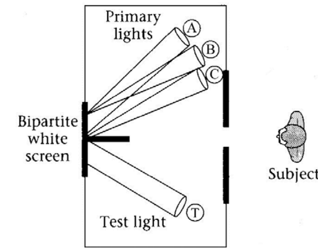
Can I turn up the intensity of one of the lights to mimic another? No

The color matching experiment

We can actually distinguish all of the individual wavelengths as different colors. Does this mean our eyes are full spectral sensors?

Unfortunately, no. To show this, we can perform a color matching experiment.

The idea is to see if we can match a given test light using a finite number of lights called **primaries** with power control knobs.



The color matching experiment (Wandell, 4.10)

The primary spectra are $a(\lambda)$, $b(\lambda)$, $c(\lambda)$.

The power knob settings are A, B, C.

Matching the test light

With the knob settings, we can produce spectra of the form:

$$Aa(\lambda) + Bb(\lambda) + Cc(\lambda)$$

Can we match the *L, M, S* responses of the test light?

First some notation:

$$\langle f, v \rangle \equiv \int f(\lambda) v(\lambda) d\lambda$$

$$\langle \alpha f, v \rangle = \int (\alpha f(\lambda)) v(\lambda) d\lambda = \alpha \int f(\lambda) v(\lambda) d\lambda = \alpha \langle f, v \rangle$$

$$\langle f + g, v \rangle = \int (f(\lambda) + g(\lambda)) v(\lambda) d\lambda = \int f(\lambda) v(\lambda) d\lambda + \int g(\lambda) v(\lambda) d\lambda = \langle f, v \rangle + \langle g, v \rangle$$

$$\langle \alpha f + \beta g, v \rangle = \langle \alpha f, v \rangle + \langle \beta g, v \rangle = \alpha \langle f, v \rangle + \beta \langle g, v \rangle$$

$$\langle \alpha f + \beta g + \gamma h, v \rangle = \alpha \langle f, v \rangle + \beta \langle g, v \rangle + \gamma \langle h, v \rangle$$

Matching the test light

Now I can write the cone responses to the test stimulus as:

$$\begin{bmatrix} L_t \\ M_t \\ S_t \end{bmatrix} = \begin{bmatrix} \int t(\lambda) l(\lambda) d\lambda \\ \int t(\lambda) m(\lambda) d\lambda \\ \int t(\lambda) s(\lambda) d\lambda \end{bmatrix} = \begin{bmatrix} \langle t, l \rangle \\ \langle t, m \rangle \\ \langle t, s \rangle \end{bmatrix}$$

The response to the combination of primaries $Aa(\lambda)+Bb(\lambda)+Cc(\lambda)$ is then:

$$\begin{bmatrix} L_{ABC} \\ M_{ABC} \\ S_{ABC} \end{bmatrix} = \begin{bmatrix} \langle Aa + Bb + Cc, l \rangle \\ \langle Aa + Bb + Cc, m \rangle \\ \langle Aa + Bb + Cc, s \rangle \end{bmatrix} = \begin{bmatrix} A\langle a, l \rangle + B\langle b, l \rangle + C\langle c, l \rangle \\ A\langle a, m \rangle + B\langle b, m \rangle + C\langle c, m \rangle \\ A\langle a, s \rangle + B\langle b, s \rangle + C\langle c, s \rangle \end{bmatrix}$$

$$= \begin{bmatrix} \langle a, l \rangle & \langle b, l \rangle & \cdots \\ \langle a, m \rangle & \ddots & \\ \langle a, s \rangle & & \langle c, s \rangle \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$

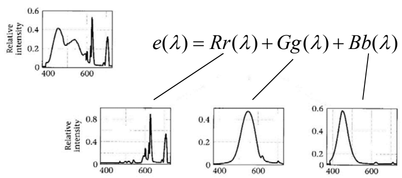
Thus, choosing the primary knobs to match a test light amounts to multiplying a matrix!

Choosing Primaries

The primaries could be three color (monochromatic) lasers.

But, they can also be non-monochromatic, e.g., monitor phosphors from an old CRT:

$$e(\lambda) = Rr(\lambda) + Gg(\lambda) + Bb(\lambda)$$

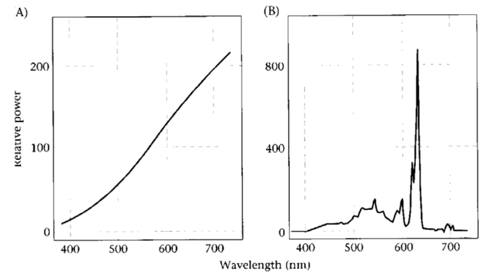


Emission spectra for RGB monitor phosphors (Wandell B.3)

Emission Spectrum is not color

Clearly, information is lost in this projection step…

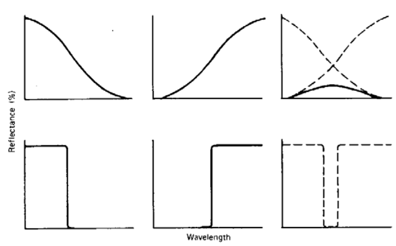
Different light sources can evoke exactly the same colors. Such lights are called **metamers**.



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

Colored Surfaces

So far, we've discussed the colors of lights. How do *surfaces* acquire color?



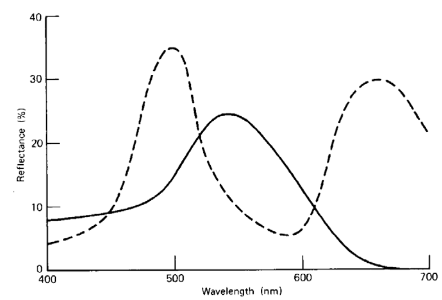
Subtractive colour mixing (Wasserman 2.2)

A surface's **reflectance**, $\rho(\lambda)$, is its tendency to reflect incoming light across the spectrum.

Reflectance is combined "**subtractively"** with incoming light. Actually, the process is *multiplicative*:

$$I(\lambda) = \rho(\lambda)t(\lambda)$$

Subtractive Metamers

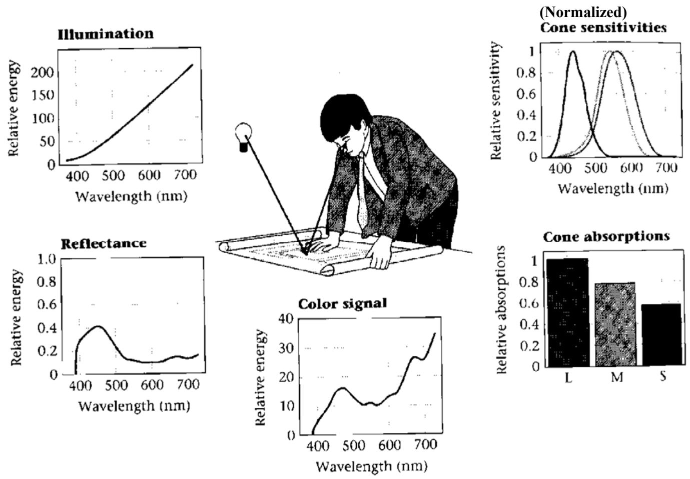


Surfaces that are metamers under only some lighting conditions (Wasserman 3.9)

Reflectance adds a whole new dimension of complexity to color perception.

The solid curve appears green indoors and out. The dashed curve looks green outdoors, but brown under incandescent light.

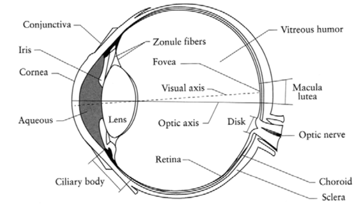
Illustration of Color Appearance



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

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 projections?

Q: Do our eyes "see in 3D"?

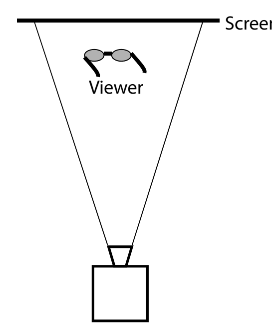
3D Displays

So-called 3D displays are all the rage now for movies and soon for televisions.

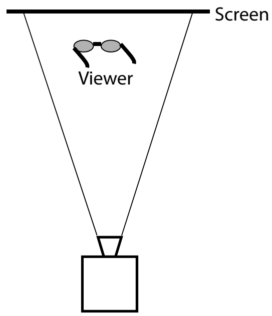
Much of our perception of 3D comes from stereo vision: each eye sees a different view of the world.

So, to create the illusion of 3D, we only need to show each eye an image of a scene created from that eye's point of view!

3D Displays, cont'd



3D Displays, cont'd



3D Displays, cont'd