Vision and Color

Brian Curless
CSEP 557
Autumn 2017

Reading

Good resources:
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.

\[
\frac{1}{f} = \left(\frac{1}{R_1} - \frac{1}{R_L}\right)
\]

Positive (converging) lens

\[
\frac{1}{f} = \frac{1}{R_1} + \frac{1}{R_2}
\]

\[
f > \frac{1}{m-1} \left(\frac{R_1 R_2}{R_L}\right)
\]

(micros is ~ 0)
**Optics, cont’d**

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

![Diagram](image)

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:

\[
\frac{1}{f_{new}} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{1}{f_3}
\]

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?

**Structure of the eye**

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.

![Eye Diagram](image)
Structure of the eye, cont.

- **Crystalline lens** - controls the focal distance:
  - Power ranges from 10 to 30 diopters (D) 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

- **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?

Retina

- **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.

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.

Saccades, cont’d

The saccadic behavior is task-specific:

1. Free examination.
2. Free examination.
3. Free examination.
5. Remember the clothes worn by the people
6. Remember the clothes worn by the people
7. Estimate how long the "unexpected visitor" had been away from the family
8. 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 0.83.

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)
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:

<table>
<thead>
<tr>
<th>Background</th>
<th>Luminance (cd/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moonless overcast night</td>
<td>0.00003</td>
</tr>
<tr>
<td>Moonlit overcast night</td>
<td>0.003</td>
</tr>
<tr>
<td>Twilight</td>
<td>3</td>
</tr>
<tr>
<td>Overcast day</td>
<td>300</td>
</tr>
<tr>
<td>Day with sunlit clouds</td>
<td>30,000</td>
</tr>
</tbody>
</table>

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 discussed by Ernst Mach, an Austrian physicist.

Appear when there are rapid variations in intensity, especially at C₀ intensity discontinuities:

And at C₁ intensity discontinuities:

Mach bands, cont.

Possible cause: lateral inhibition of nearby cells.

Q: What image processing filter does this remind you of? Lateral inhibition (Glassner, 1.25)

Neural Networks

- A neuron

Q: What image processing filter does this remind you of? Lateral inhibition (Glassner, 1.25)
Convolution

<table>
<thead>
<tr>
<th>Image</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>54</td>
</tr>
<tr>
<td>145</td>
<td>98</td>
</tr>
<tr>
<td>89</td>
<td>177</td>
</tr>
<tr>
<td>67</td>
<td>90</td>
</tr>
<tr>
<td>106</td>
<td>111</td>
</tr>
<tr>
<td>221</td>
<td>154</td>
</tr>
</tbody>
</table>

Convolutional Neural Network (CNN)

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:

The spectrum describes the energy at each wavelength.

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

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.

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? \( \sqrt{2} \)

What about a 500 nm green laser half as bright? \( \frac{1}{2} \)
Cone pigments

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

Cone pigment 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:

\[
L(\lambda) = \int 2g(\lambda) \, d\lambda \\
M(\lambda) = \int 2g(\lambda) \, d\lambda \\
S(\lambda) = \int 2g(\lambda) \, d\lambda
\]

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?

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 4.13)
**Color Appearance of Light Reflection**

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

**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:

![Cone distribution images](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.

![Human vision diagram](Glassner, 1995)

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

Q: Do our eyes “see in 3D”?