Human Vision

Lecture #4: Tuesday, 28 January 2003
Lecturer: Greg Humphreys
Scribe: Brenden Schubert
Reviewer: Nolan Goodnight

1 Why Study Vision

Some computer-generated images are meant for computer processing, but most are intended for human viewing. To successfully use an image to convey information, we must understand how it will be percieved by the viewer. We can improve rendering performance by not computing things that will be imperceptible in the resultant image or can improve rendering efficacy by eliminating artifacts that would impair perception of the resultant image.

2 Ocular Anatomy and Physiology

2.1 Structure

The human visual system is comprised of the eyes and the brain. We know much more about the workings of the eyes than the brain.

The Nodal Point is the point in the eye where light rays converge. The Visual Angle is the angle subtended by an object being viewed, as seen from the Nodal Point. A diopter is a unit of focusing strength equal to the inverse of the focal length in meters.

- Cornea: The clear covering over the front of the eye; has a focusing strength of 40 diopters (40D).

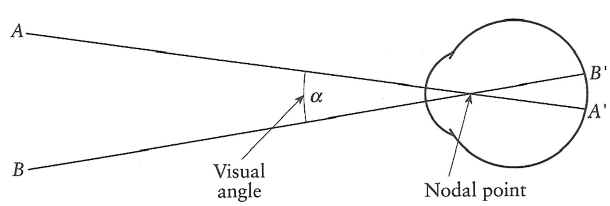


Figure 1: Figure 1: Nodal Point and Visual Angle

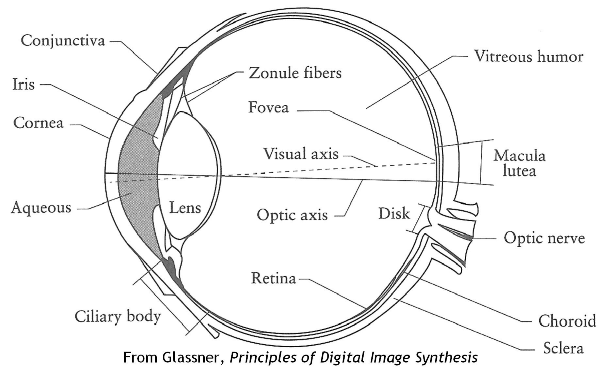


Figure 2: Figure 2: Anatomical Structures of the Human Eye

- Iris: Colored ring around pupil; can change the size of the pupil, regulating the amount of light entering the eye
- Pupil: Hole in the middle of the iris; only light going through the pupil contributes to seeing
- Lens: Transparent focusing mechanism; surrounded by muscles which can change its focusing power, an ability called accommodation. The range of accommodation is 10D-30D in a child, but becomes much smaller with age.
- Retina: layer of cells covering about 200 degrees on the back of the eyeball; contains two types of photosensitive cells, rods and cones.

Cones are sensitive to color, and rods only to intensity. Rods have roughly ten times the sensitivity of cones. The Fovea is a 1-2 degree region in the center of the retina which contains only a very dense covering of cones - 147,000 cones per millimeter. Hawks have 1M cones per millimeter, resulting in vision about eight times as sharp as ours. The retina is radially symmetric around the fovea, with a total of 120M rods and 6M cones. There are no rods or cones where the optic nerve attaches to the retina, resulting in a "blind spot". The optic nerve has only 1M fibers in it, so some processing of information must occur in the eye itself.

2.2 Spectra

The eye is sensitive to a narrow band of the electromagnetic spectrum which typically ranges from 380-780nm. All rods react equally to a range of wavelengths centered around 500nm. There are three types of cones, S, M, and L, which react to small, medium, and large wavelengths of light. The only information generated by a cone is that a photon has arrived, not its wavelength, so the visual system must count up the indications of all three types of cones to determine the wavelength of light that would be most likely to cause such a response.

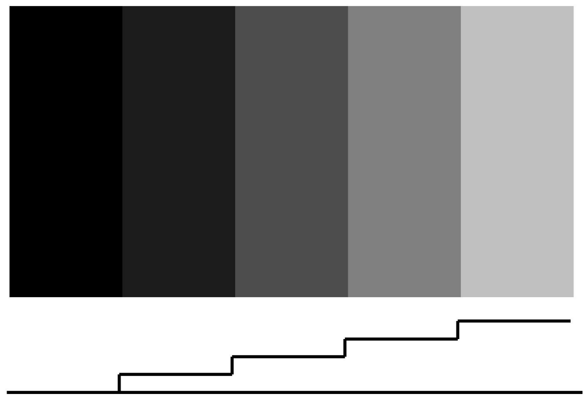


Figure 3: Figure 3: Mach Bands

2.3 Visual Phenomena

Photoreceptors average the stimuli they recieve over a few milliseconds, which acts as a low pass filter over the time response signal. This results in the Critical Flicker Frequency - the highest frequency of flickering that is perceptible - which is around 60Hz for humans. The sensitivity of the eye to light can change drastically based on the overall luminance of incoming light. The Just Noticable Difference is the lowest noticable difference in neighboring regions' intensities. It is constantly about .02 over a wide range of intensities. The eye is relatively insensitive to high-frequency noise.

When viewing bordering regions of different intensities, Mach bands are what the eye perceives as a gradient across the band of uniform intensity, due to the contrast with the bordering regions. Surrounding regions also affect our perception of absolute and relative intensity.

3 Depth Perception

3.1 Monocular Depth

The human visual system is well-suited to extracting distance information to build a mental 3-D model to represent what it sees.

- Oculomotor Depth Perception: the use of how much muscle contraction of the lens is required to focus on an object; only useful at distances of ¡6m.

- Visibility/Interposition: if something is not visible because something else is concealing it, then the concealing object must be closer; this is a very strong cue for depth and even overtrumps stereo vision when they conflict.

$$C(\lambda) = \bar{X}x(\lambda) + \bar{Y}y(\lambda) + \bar{Z}z(\lambda)$$

$$\bar{X} = \int_{380}^{780} C(\lambda)x(\lambda)d\lambda \quad \bar{Z} = \int_{380}^{780} C(\lambda)z(\lambda)d\lambda$$

$$\bar{Y} = \int_{380}^{780} C(\lambda)y(\lambda)d\lambda$$

- Perspective: the size of an object decreases as it gets farther away; we can use the expected size of something to estimate how far away it is.
- Motion Parallax: the apparent rate of relative movement of objects to a moving observer decreases as objects are farther away. Also, as we turn our head, the apparent rotation of objects decreases the more distant they are.

4 Color

4.1 Channels

Color is transmitted from the eyes to the brain in three channels which are composed of combinations of the outputs of S, M, and L receptors:

- Achromatic = M + L
- Red/Green = M - L
- Blue/Yellow = S - M - L = S - A

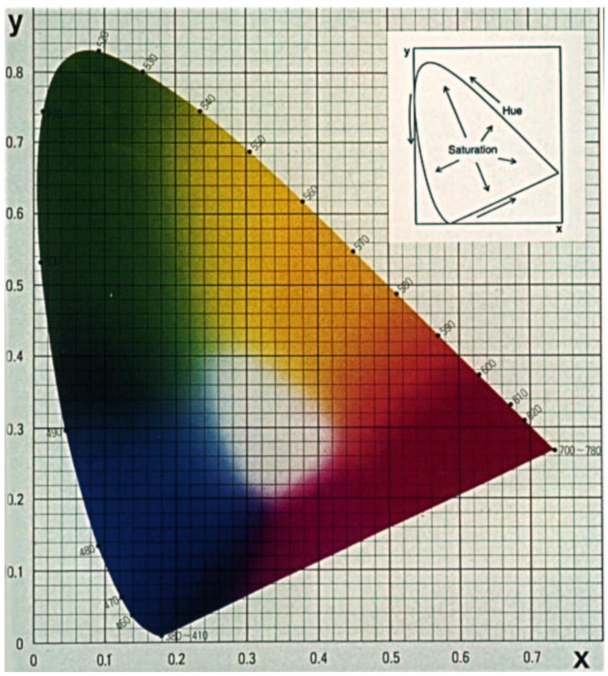
This suggests a three-dimensional color space, with compoments of intensity, red/green, and blue/yellow. Many attributes of the observer and the object being viewed can cause varying perceptions of color. The CIE (Commission Internationale de l'Eclairage) performed a series of trials where different users were shown a target color and allowed to mix arbitrary amounts of three primary colors (red, green, and blue) until they felt the mixture matched the target. A positive contribution mixed with the other primary colors, and a negative contribution added onto the target color. The results were averaged for many users and became the perceptual response of the "average" observer. Because there are a finite number of primary colors, many different spectra can have the same matching amounts of the primaries; these are called metamers. Because the red primary spectrum actually goes negative, the CIE generated three new primary colors, X, Y, and Z, with always-positive matching functions x(λ), y(λ), and z(λ). Given a spectrum C, we can decompose it into X, Y, and Z using these functions:

We can project the X-Y-Z color space onto the plane X+Y+Z=1 to give x, y, and z as shown:

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$

$$z = \frac{Z}{X+Y+Z} = 1 - x - y$$



This gives the Chromaticity Diagram: