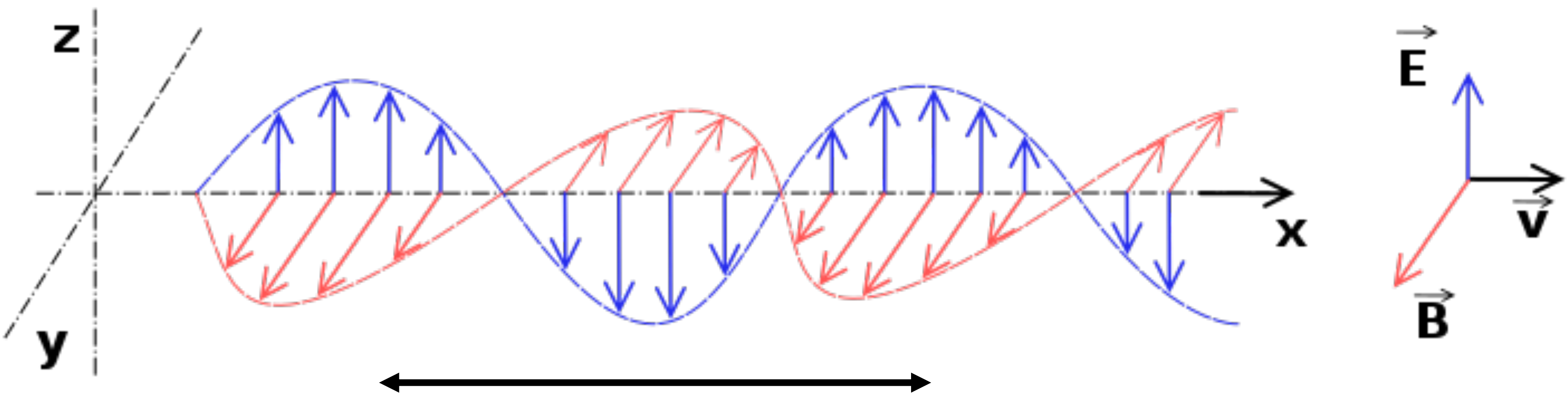
LECTURE 6: THE PHYSIOLOGY OF HUMAN VISION

Ehsan Aryafar
earyafar@pdx.edu
http://web.cecs.pdx.edu/~aryafare/VR.html

Recall: What is light?

- Light is a form of electromagnetic wave which can be detected by human eye
- Electromagnetic waves (EW) are created as a result of vibration of an electric filed and a magnetic filed
  - Travel at speed of light in vacuum (3×10^8 meters per second)
  - Does not need matter to travel
  - Example EWs: communication signals (LTE, WiFi), light, x-ray, gamma-ray



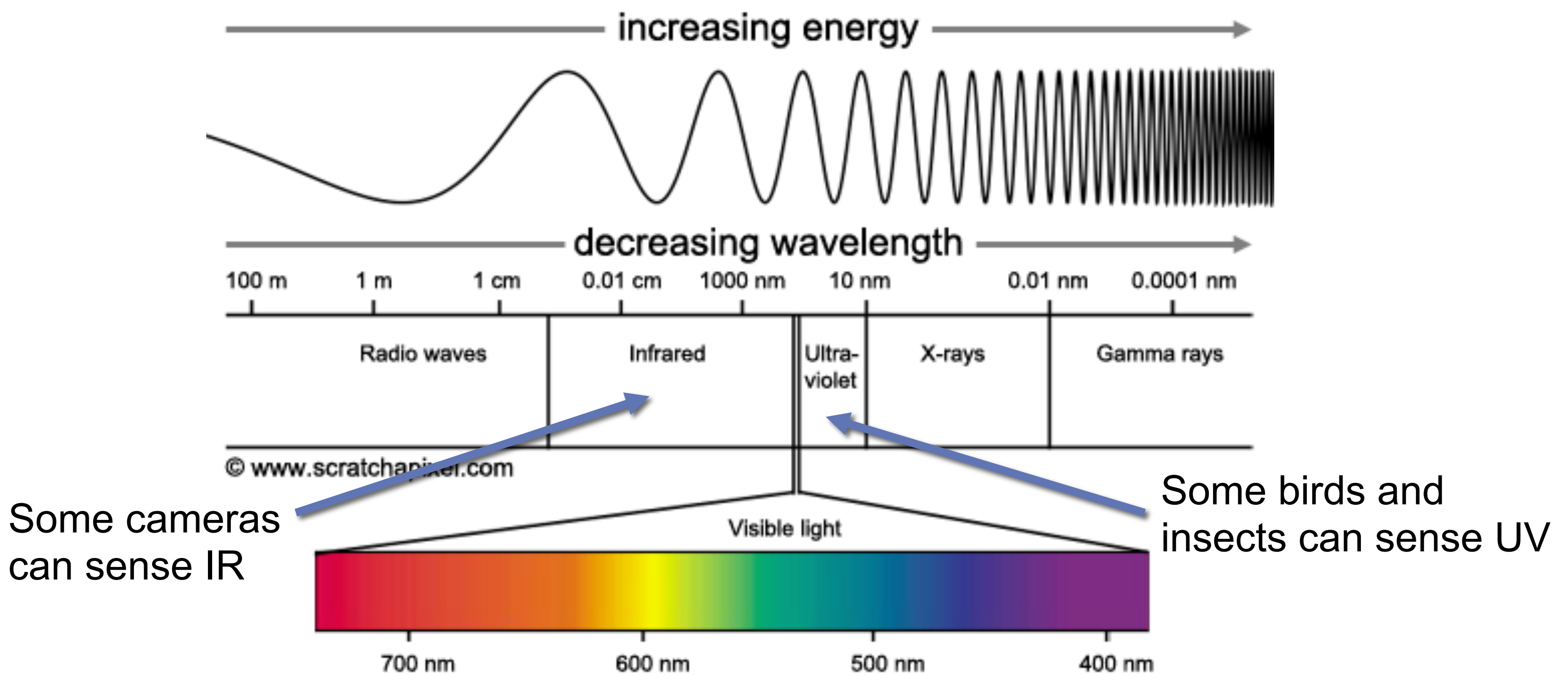
$$\lambda = c \,.\, T = c / f$$

λ: wavelength
c : speed of light
f: frequency (Hertz)
T: period/time (second)

What is the relationship between f and T?
Freq: # of times in 1 sec that a wave passes a fixed location

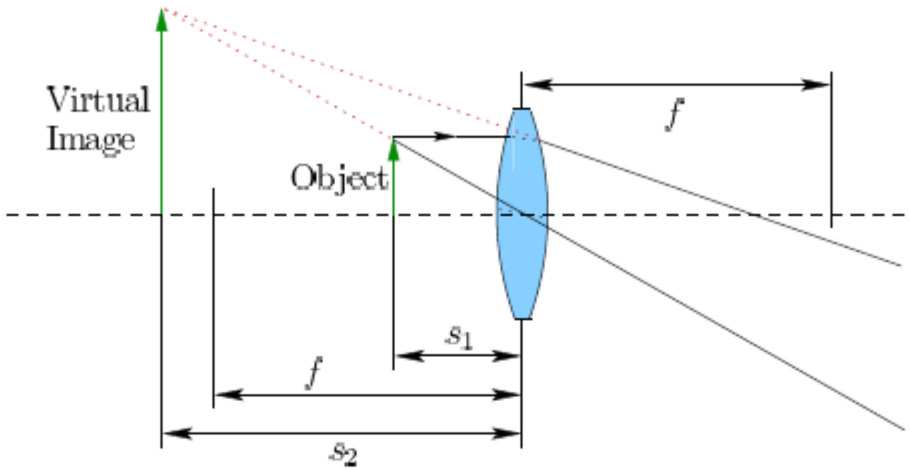
Recall: Spectrum

- The set of frequencies occupied by the EW



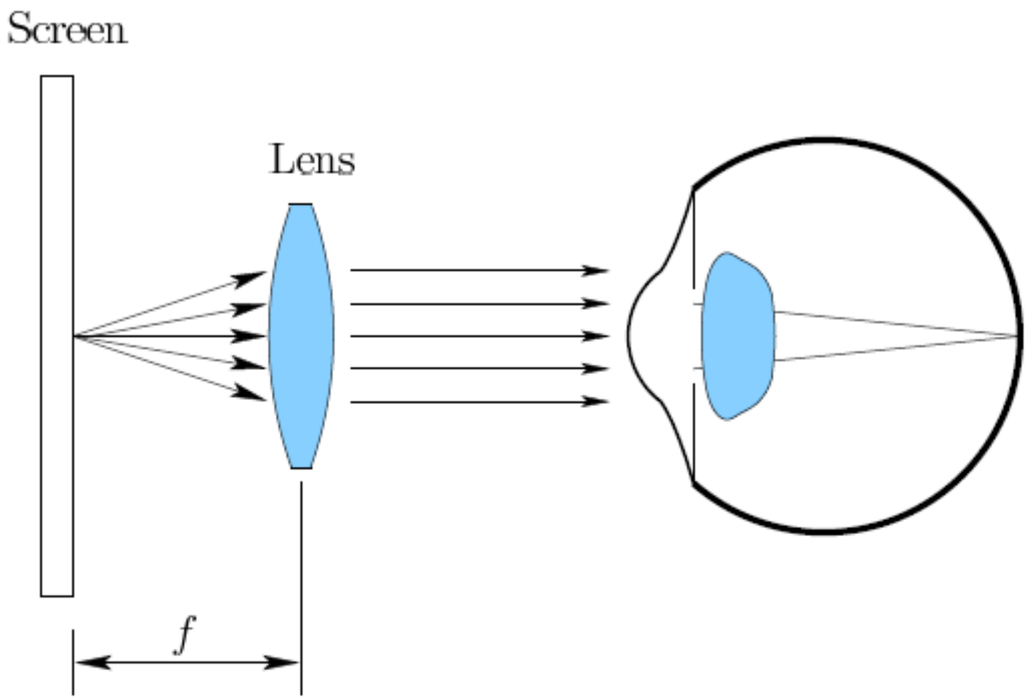
Different interpretations of EWs (e.g., light) help explain their different properties
Visible light is tied to humans!

Recall: Virtual Image



- If the object is very close to the lens, then the lens cannot force the outgoing light rays to converge to a focal point. In this case, a virtual image appears and the lens acts as a magnifying glass
  - This is the way lenses are commonly used in VR headsets
  - To the user looking through the lens, it appears as if the screen is infinitely far away and quite enormous!

Recall: A Simple VR Headset



- The lens is placed at its focal length. A virtual image (and enormous) is formed behind it. A real image is formed on your retina

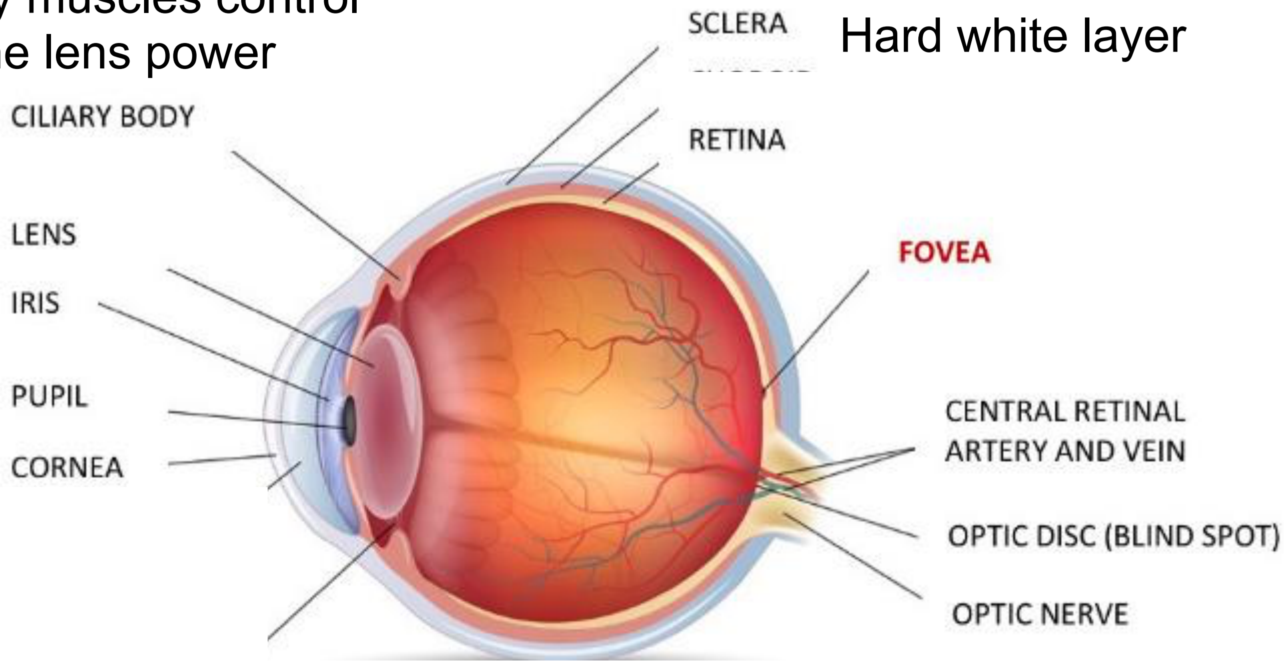
Outline

- Light around us form images on retinas that capture color, motion, and spatial relationships in the world
- Brain fills out void (which are many) to create an image of perfect clarity consistent with our life experiences
  - We are being fooled by our brains (mostly an illusion)!
- VR tries to fool our senses, so important to understand the physiology of our vision (and ears). We cover:
  - More anatomy of human eyes (photoreceptors as input pixels)
  - Neuroscience: from photoreceptors to visual cortex
  - Eye movements
  - Implication for VR display requirements!

Anatomy of Human Eye

Ciliary muscles control the lens power

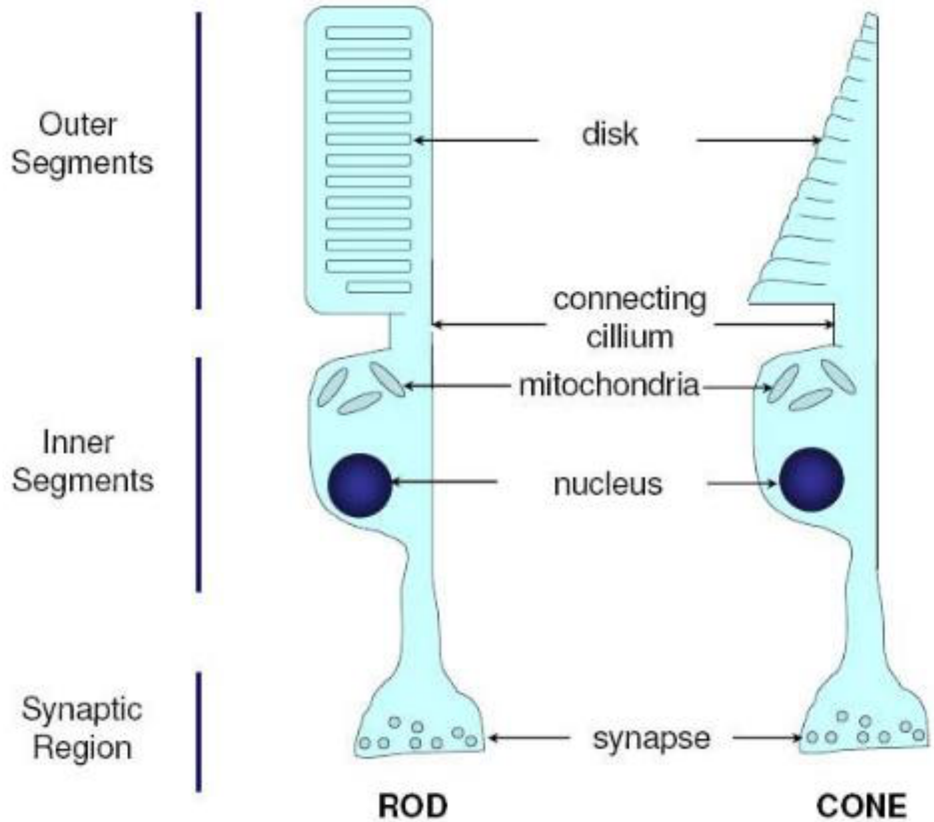
Hard white layer



Eye: spherical, 24 mm diameter; Interior is a gelatinous mass, allows light to penetrate
Cornea is a hard, transparent surface through which light enters (high optical power)
Light enters the lens by passing through pupil, the size of which controlled by Iris
Light eventually hits the retina, which is disc shaped (180 degrees)
Retina is covered by photoreceptors, which behave like input pixels
Fovea is the most important part of retina with highest visual activity
Optical disc is a hole in retina through which neural pulses are sent outside the eye

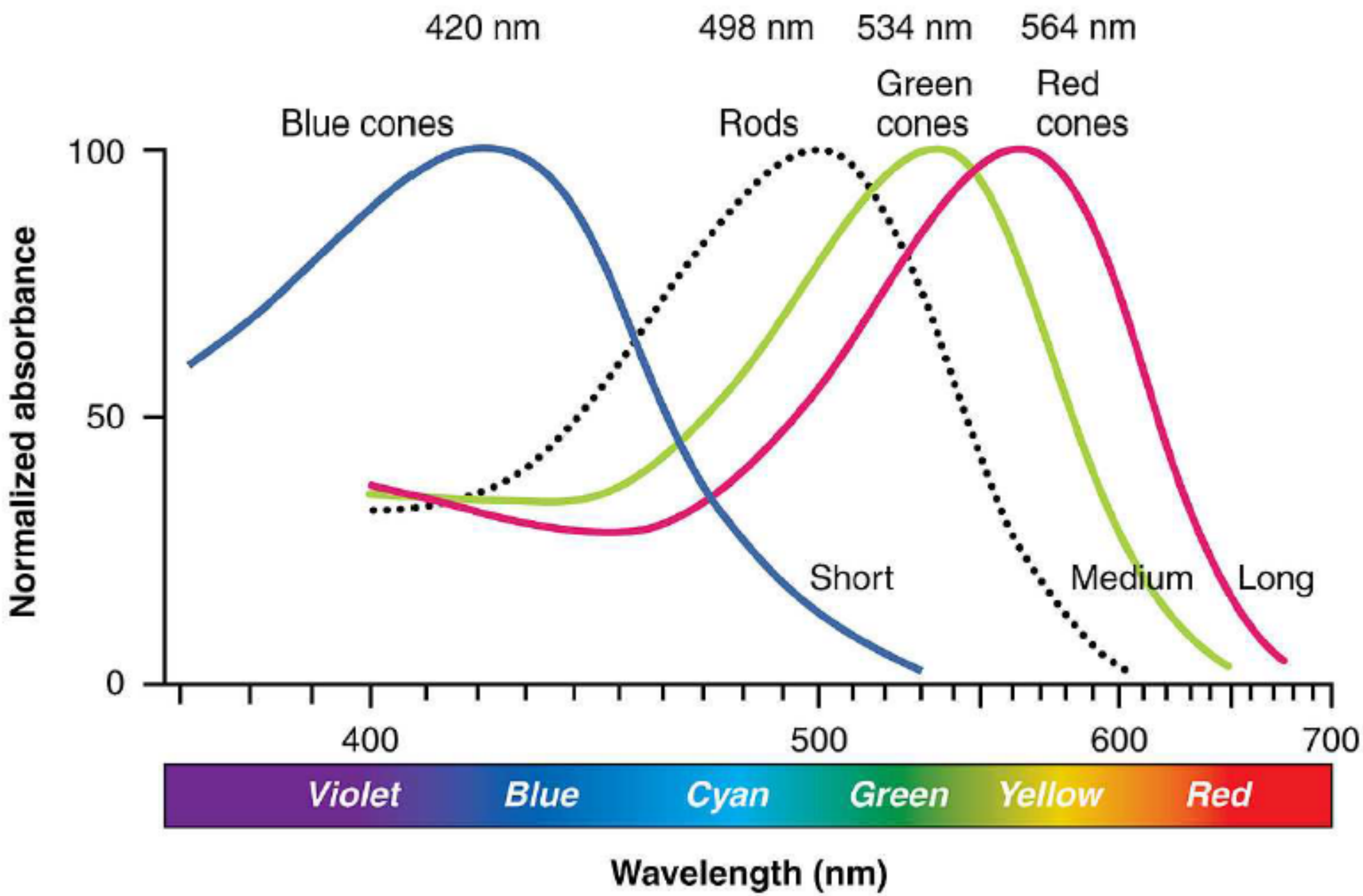
Photoreceptors

- The retina contains two types of photoreceptors
- Rods: triggered by very low levels of light
- Cones: Require more light for activation and distinguish color
  - 3 types of cones: red, blue, green
- Human eye has 120 million rods and 6 million cones
- The width of smallest cone is 1000 nm, close to light wavelength

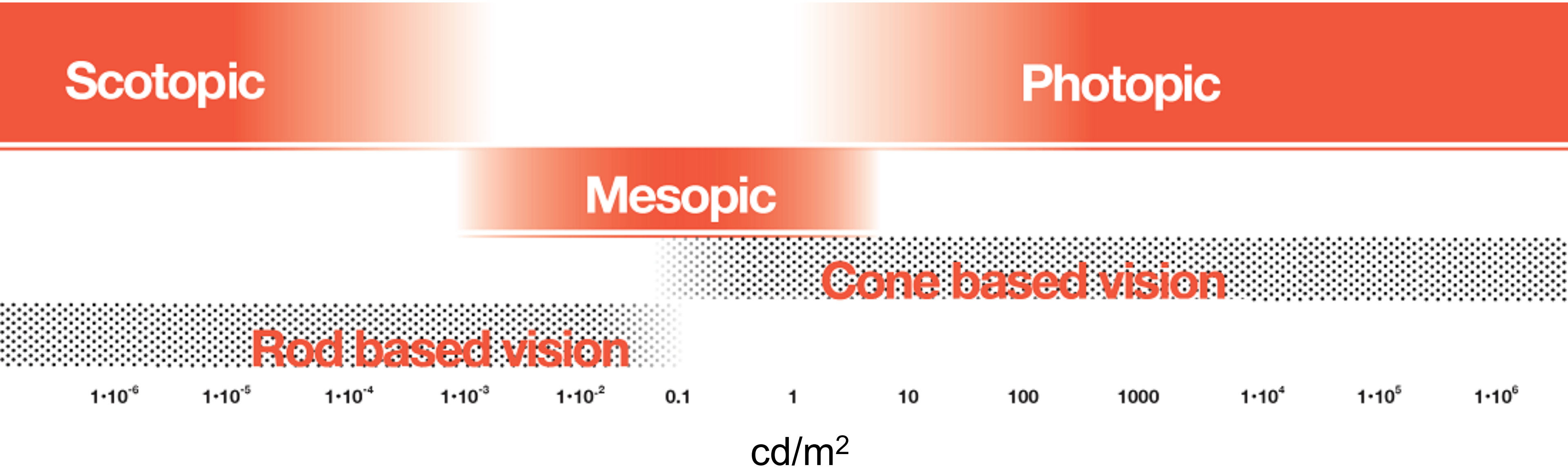


Outer segments contain photopigments that electrochemically respond when bombarded by photons

Sensitivity of Rods and Cones as a Function of Wavelength

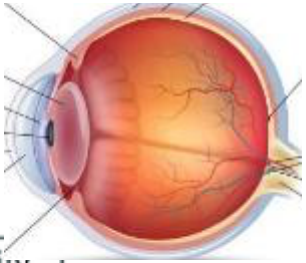


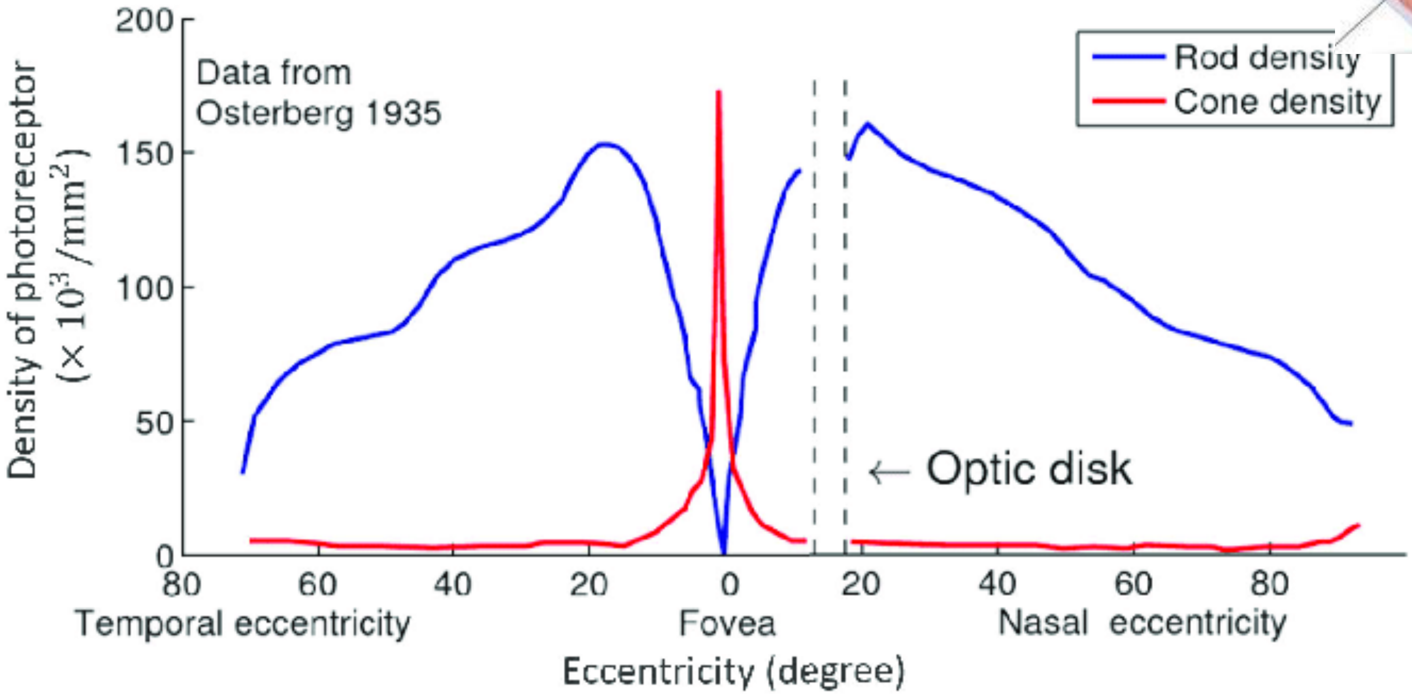
Rod vs Cone Based Vision



- Illuminance is measured in candles per square meter
- Example: paper in moonlight has 0.2 (cd/m^2), which is equivalent to 1 photon per receptor

Photoreceptor Density





- Cone density is highest at the fovea
  - Fovea has a diameter of only 0.5 mm (+/-0.85 degree angular range)
  - Eye must be pointed straight a target to see a sharp colored image
  - Corners have low cone density
    - We are good at detecting motion on the peripheries but not colors
- Blind spot is due to the canal through which eyes route neural signals to brain

Illusion of a Sharp Colorful Image

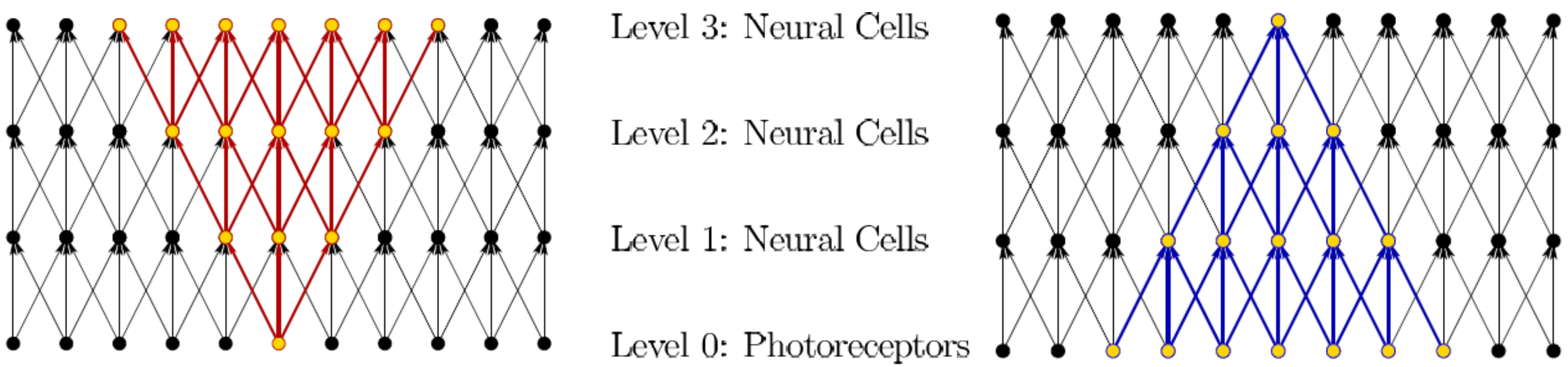
- If there is a blind spot and cones with narrow coverage, then why then do we perceive the world as a sharp colorful image with a wide angular range?

Illusion of a Sharp Colorful Image

- If there is a blind spot and cones with narrow coverage, then why then do we perceive the world as a sharp colorful image with a wide angular range?
  - Frequent eye movements
  - Having two eyes
  - Perceptual processes created by our brain fills the missing details using contextual information and memory/history

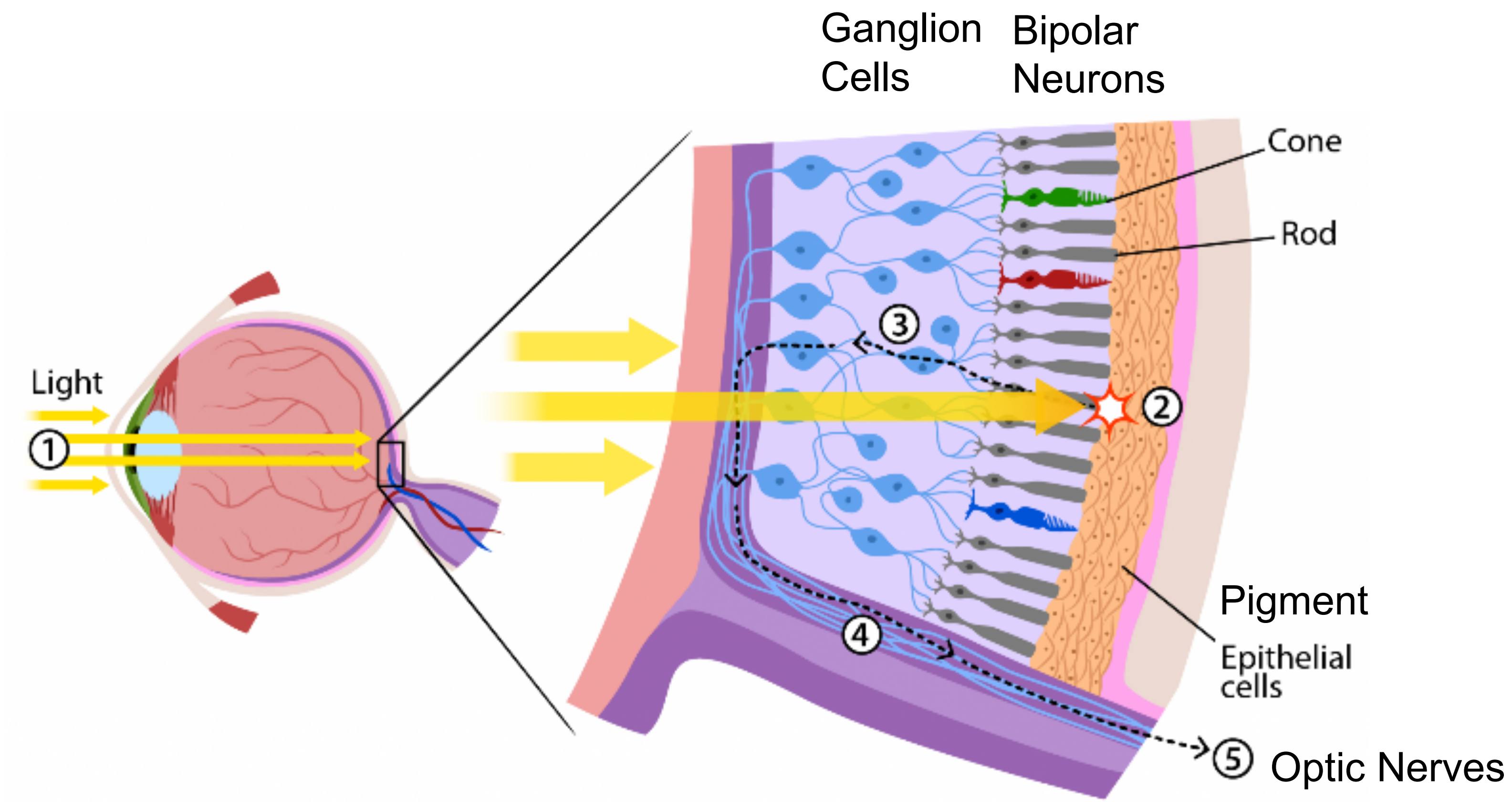
From Photoreceptors to Visual Cortex

- Photoreceptors convert the light energy into an electrical signal called a neural impulse
- Signals are propagated upward in a hierarchical manner
- Neural layers within eyes already perform some image processing tasks
- Brain neural layers fuse information from a lifetime of memories with information from all neurons to perceive something, e.g., recognize a face



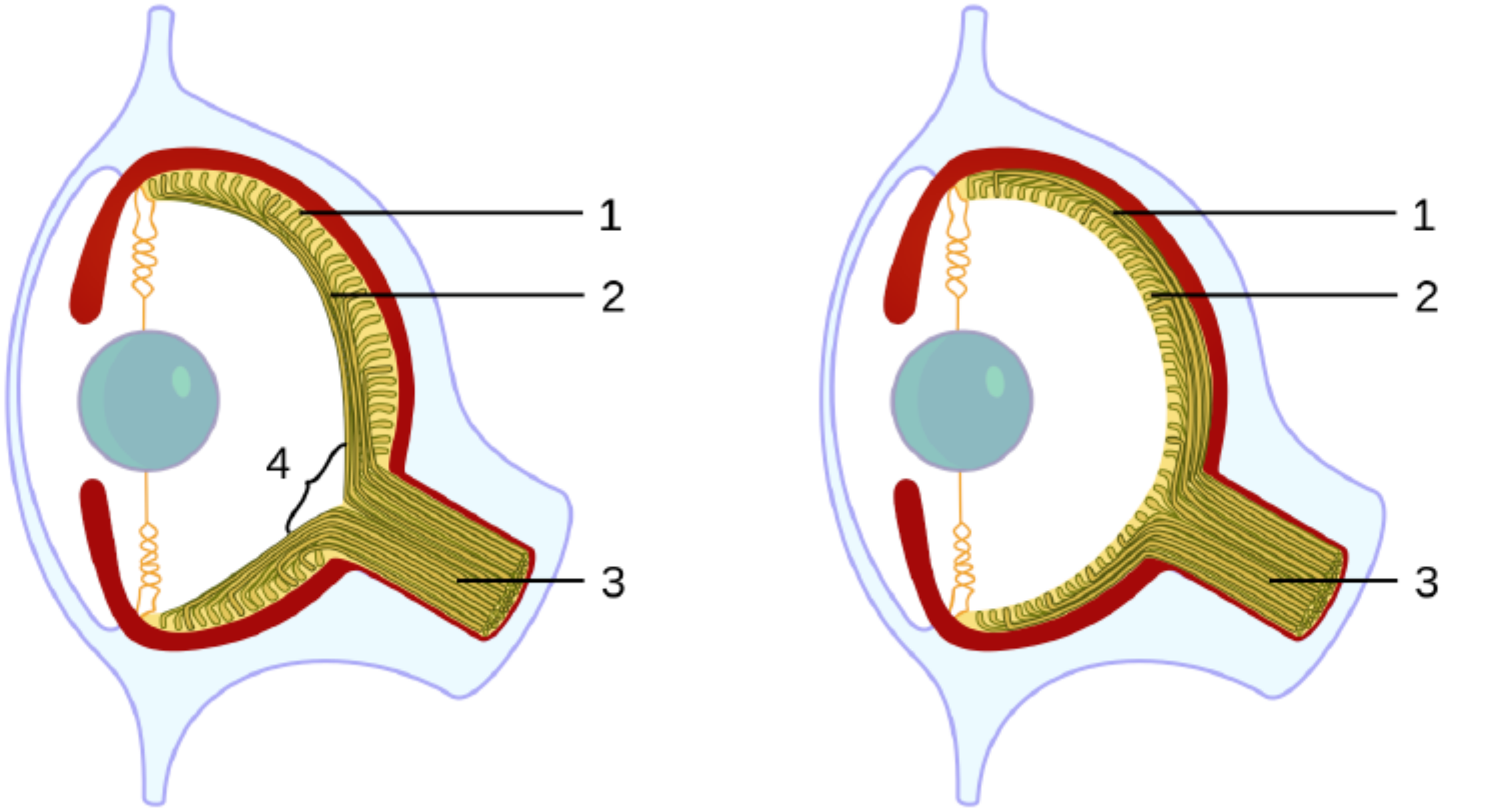
A simplified Model

Neuron Layers in the Eyes



- Light appears to be moving in the wrong directions (passing through neurons before hitting photoreceptors)
- A result of inside-out retina: a hole punctured into retinas to pass optical nerves to brain
- Some bipolar neurons connect only to cones, other to both rids and cones
- An on bipolar activates when the rate of photons hitting its photoreceptors increases
- An off bipolar neuron activates with a decreasing photon absorption
- Human vision even at the lowest layers is not totally understood
- The well understood parts greatly contribute to the design of VR systems

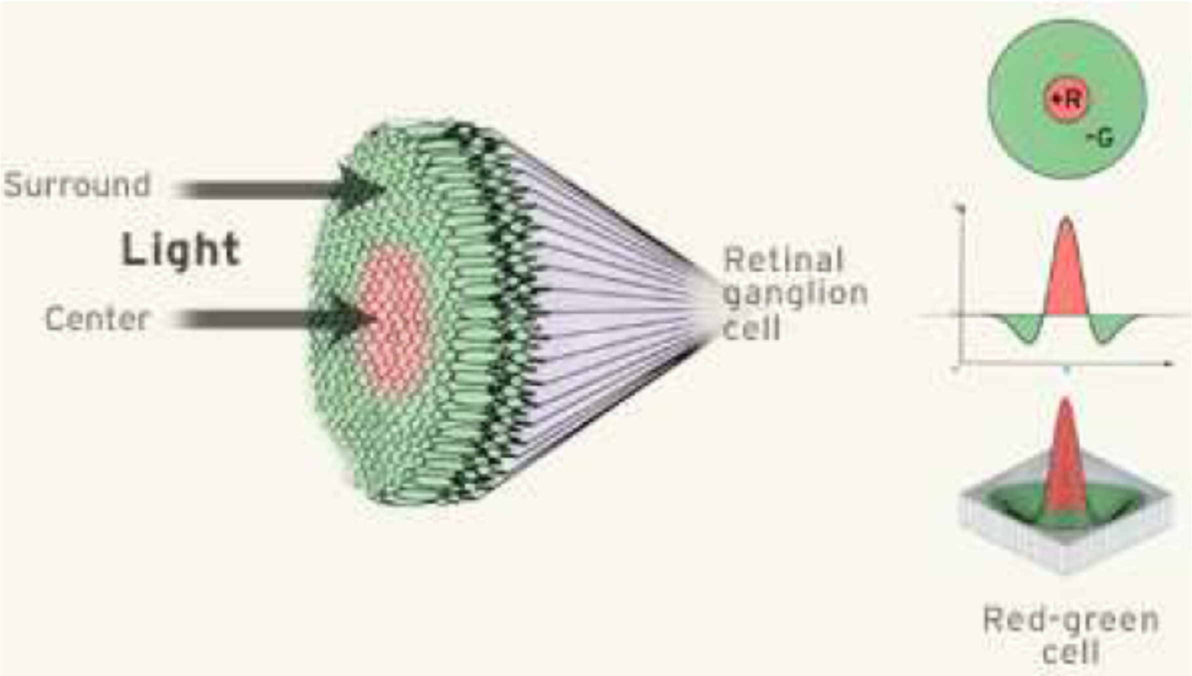
Retina in Different Species



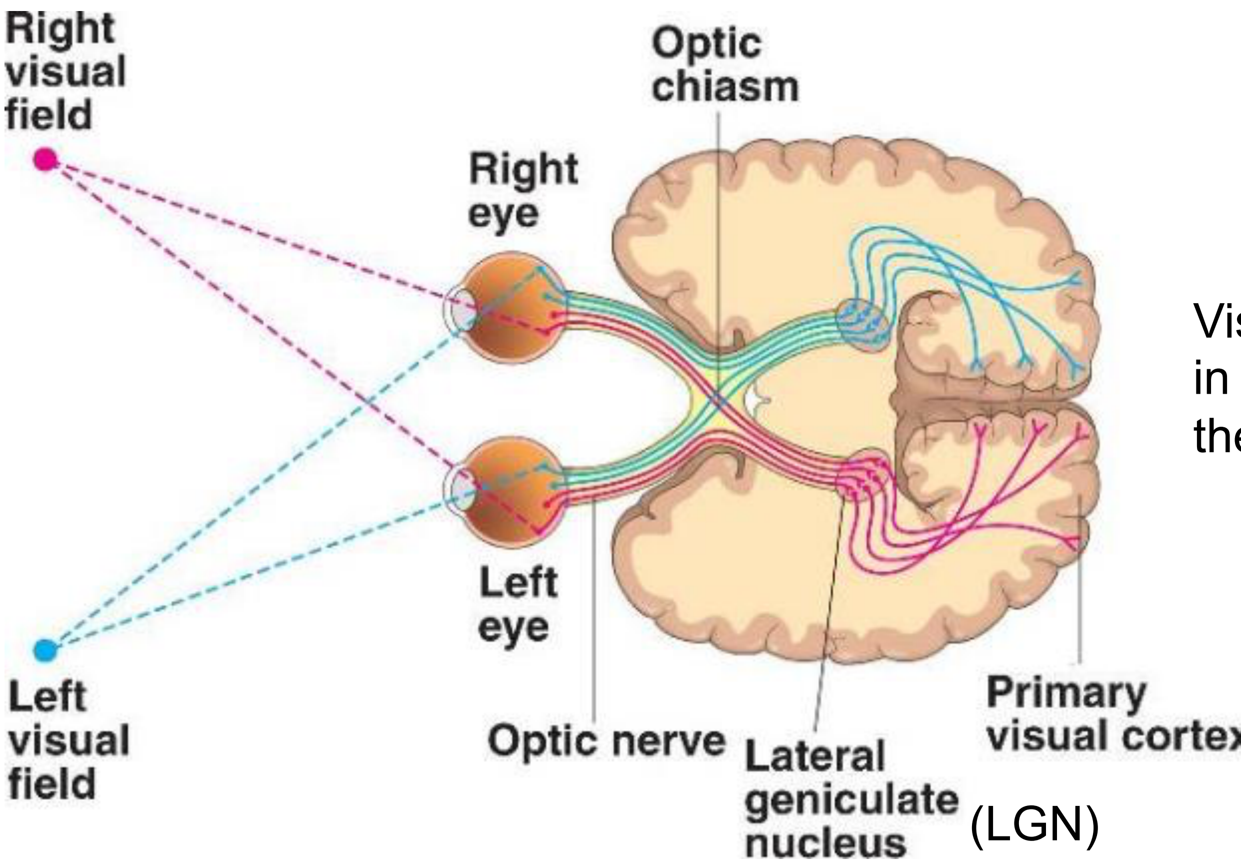
- Vertebrates (including humans) have inside-out retinas
- Nature got it right with octopuses and other cephalopods

Ganglion Cell Layer

- Several kinds of ganglion cells exist (some are well understood)
- Think of them as tiny image processing units that pick out local changes in time, space, and color
  - They respond to activations by photoreceptors
- Below is a ganglion cell that activates when center is red and surrounding is green
  - **Can pick up spatial changes in color**
- Note that as optical nerve leaves eyes, a significant amount of processing is done



From Photoreceptors to Visual Cortex



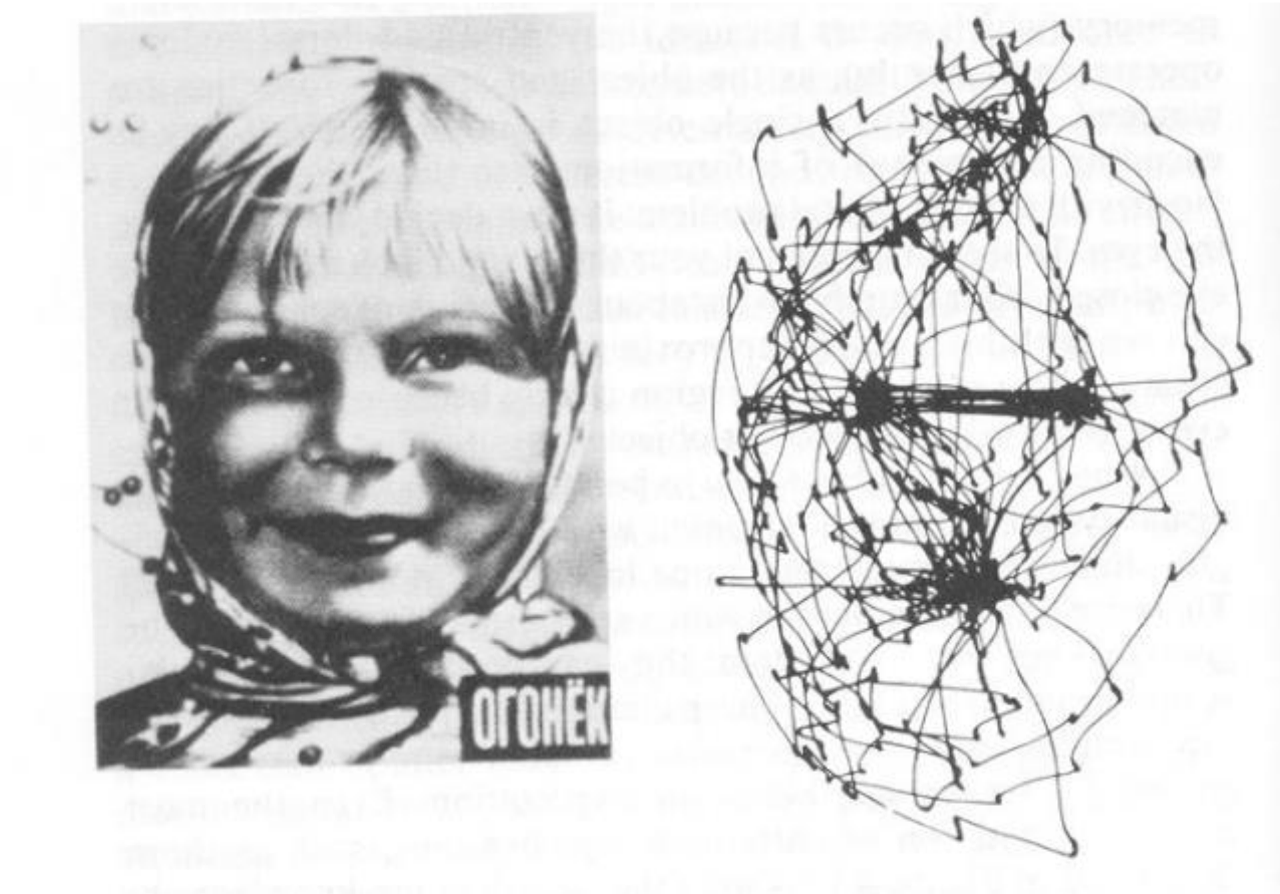
Visual cortex is in the back of the brain!

- LGN is part of thalamus, is a router that sends signals to brain (also some processing)
- Visual cortex processes data based on neural circuitry, stimulation of retinas, and information from other senses. This is called visual perception. Full characterization of visual perception is still an open field.

Eye Movements

- Eye Movements are both voluntarily and involuntarily
  - One reason for eye movement is to position feature of interest in fovea
    - Recall that fovea has a high concentration of cones with a narrow angle of view
  - It takes photoreceptors 10 ms to fully respond to stimuli (response lasts for up to 100 msec)
  - Eye movement also helps get a stereoscopic view and stop adaptation
    - It is shown that when eye motion is suppressed, visual perception disappear
- To get a coherent view of an object, eyes rapidly scan the object while fixating on point of interest
  - Fixating helps all photoreceptors to fully charge!

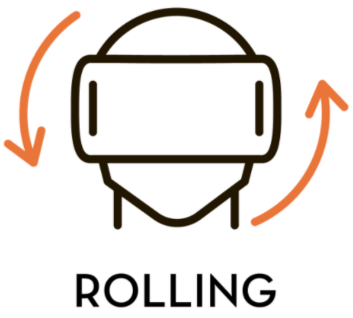
Eye Movement Studies

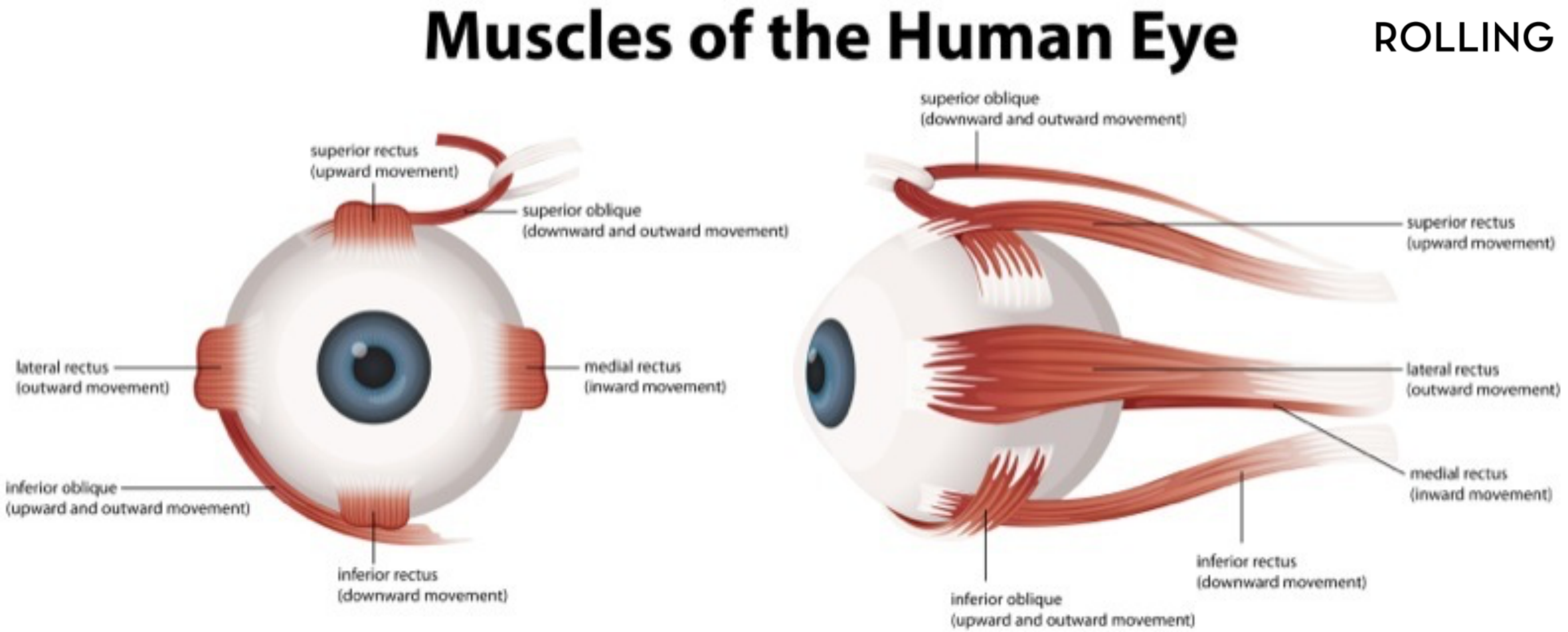


Eye movements while watching a girl's face
(early study by Yarbus, 1967)

Transition between fixations is achieved by saccades

Eye Muscles





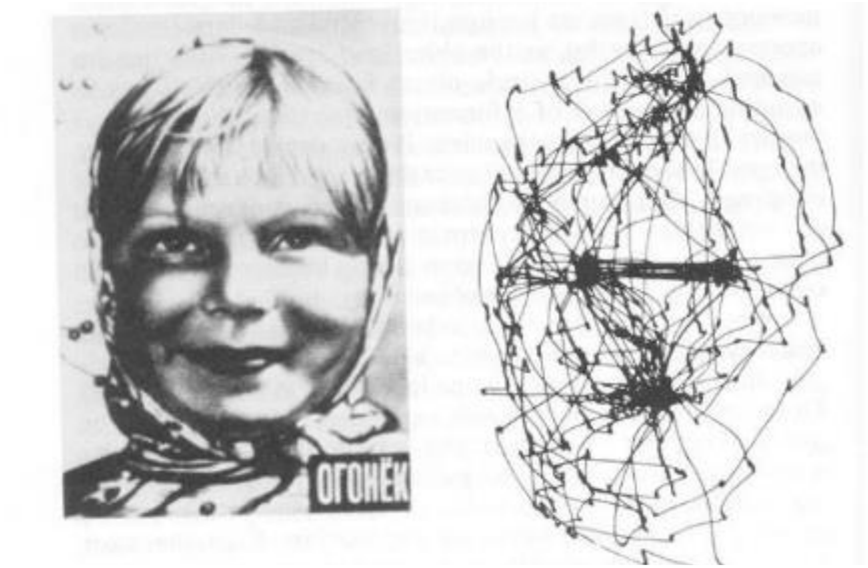
There are six muscle tendons attached to each eye so that yaw, pitch, and a small amount of roll becomes possible

To rotate, two muscles push and pull from opposite directions

Eye movements can mostly be approximated as a 2D set

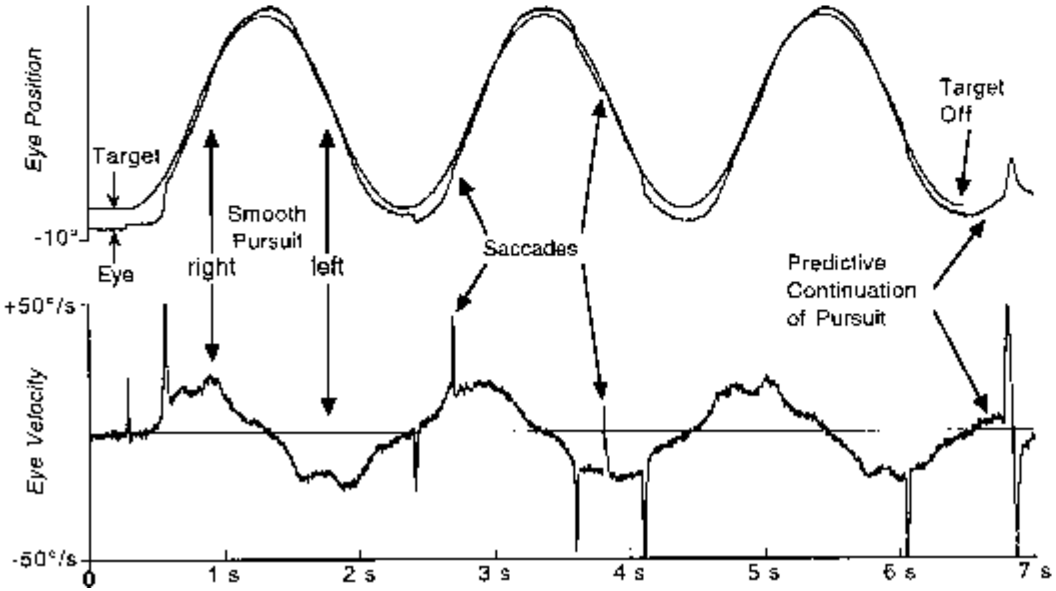
Types of Eye Movements

- (1) Saccades: rapid eye movements (e.g., between fixations)
  - Last less than 45 msec with 900 degrees/sec rotations
  - Brain masks the time between saccades (coherent vision)
  - We may have no awareness but can control them (what to focus)



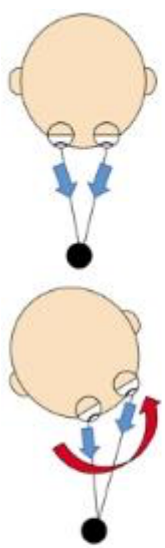
(2) Smooth Pursuit

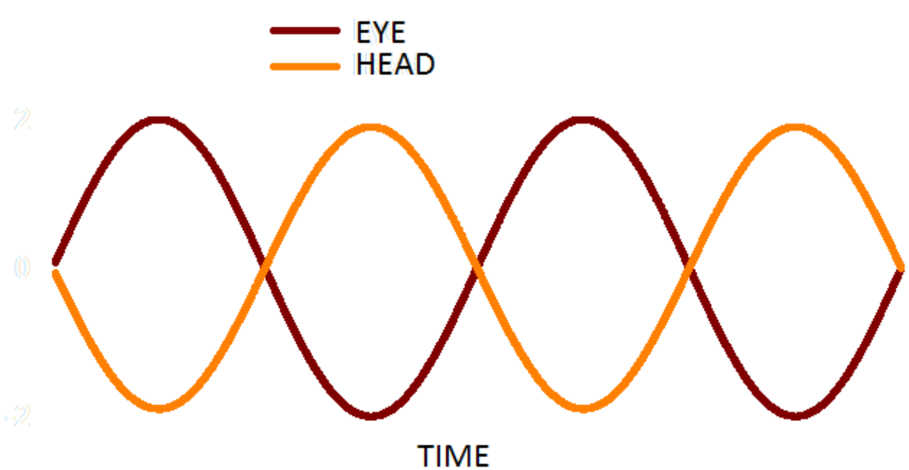
- The eye slowly rotates to track a moving target
  - Rotation rate is usually less than 30 degrees per second
  - Slow enough to reduce blur on retinas
    - Recall that photoreceptors have a 10 ms delay to activation
  - If the target is moving too fast, saccades may be inserted intermittently



(3) Vestibulo-Ocular Reflex (VOR)

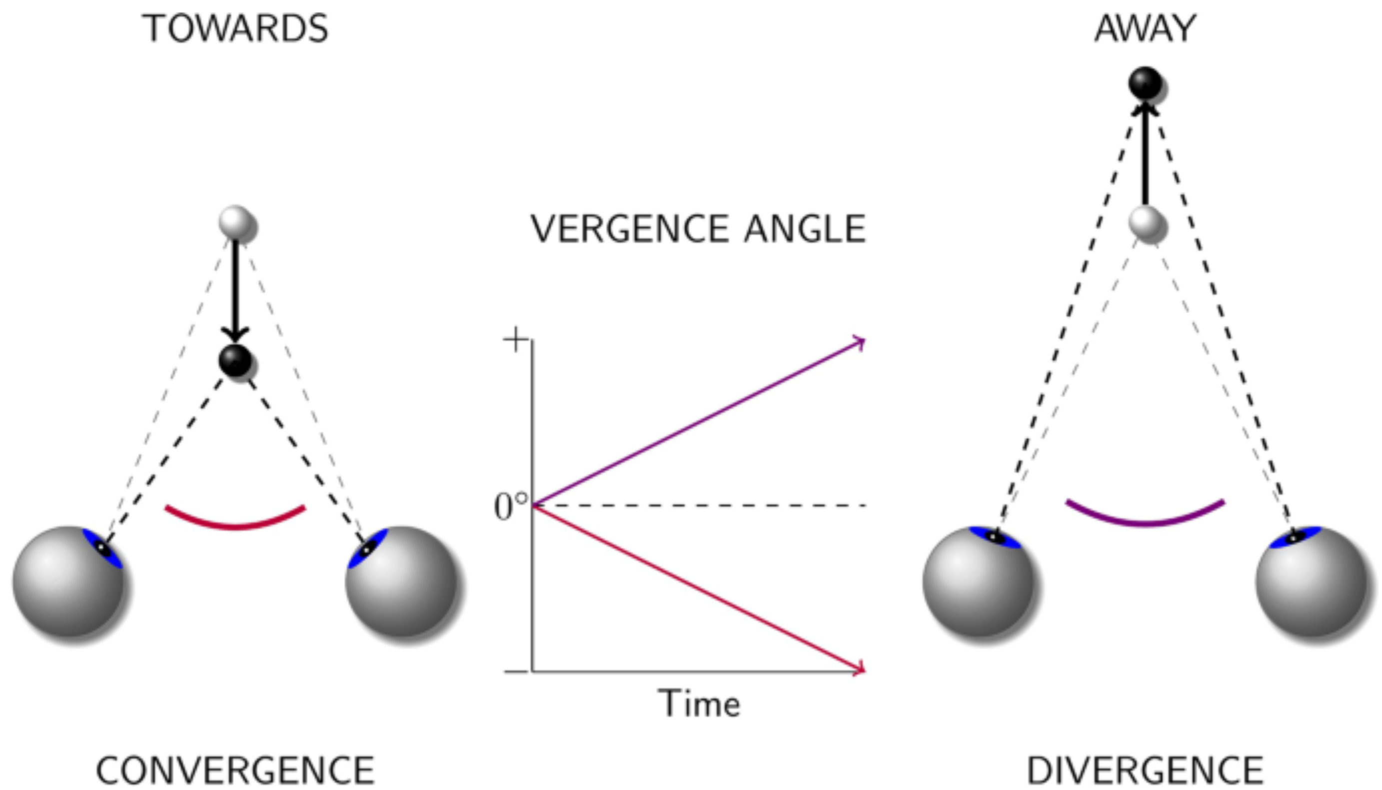
- Important to understand for VR
- Hold your finger in front of your eyes and fixate. Next yaw your head (like saying “no”)
  - Eyes effortlessly rotate to counteract the head movement
  - A reflex as eye movement bypasses higher brain functions
  - Main goal is to provide image stabilization





(4) Vergence

- Stereopsis is when both eyes fixate on an object
- Divergence/convergence happens when there is a transition in fixation from close to far and vice versa

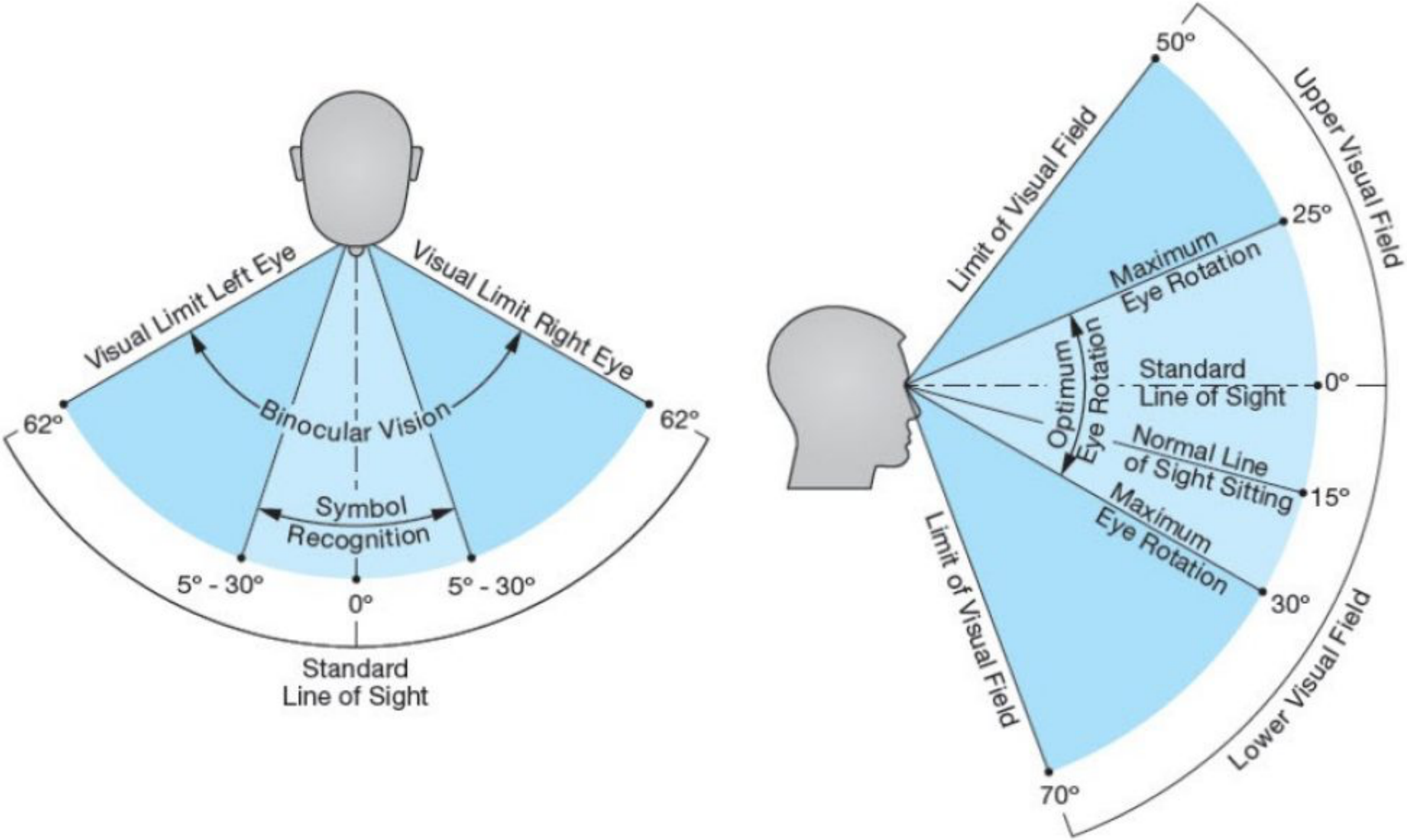


Other Motions

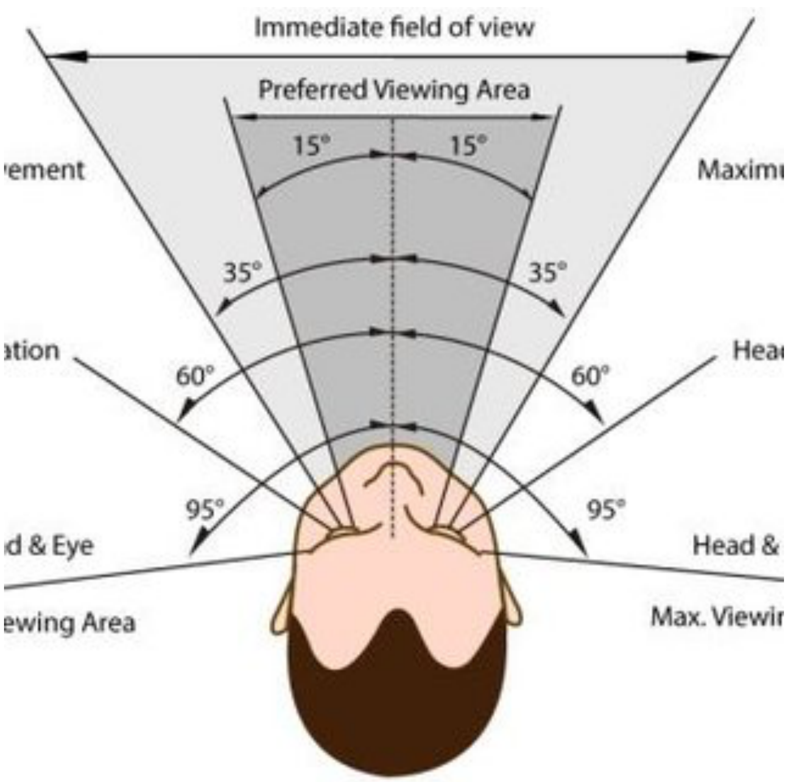
- (5) Optokinetic Reflex: happens when watching a fast moving train, while standing on a fixed ground
  - Eyes rapidly and involuntarily choose features for tracking the object

- (6) Microsaccades: small, involuntary jerks of less than one degree that trace out an erratic path
  - Believed to help with fixation control, improvement of visual activity, reduction of perceptual fading due to adaptation, and resolving perceptual ambiguities
    - Still an active area of research in neuroscience, biology, and psychology

Eye Field of View



Combined Head and Eye Field-of-View



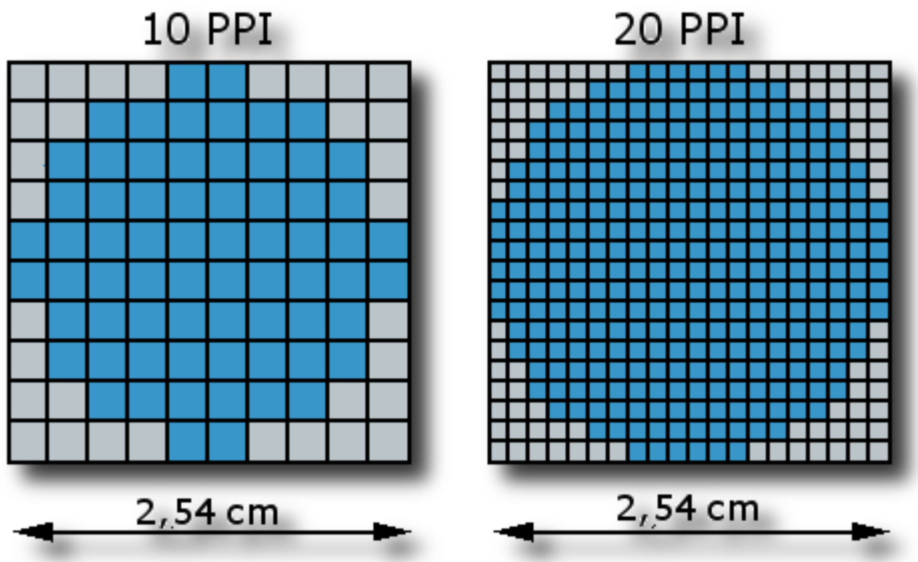
Implications for VR

- Basic physiological properties such as photoreceptor density or VOR circuitry directly impact the engineering requirement for visual display hardware
  - Engineering systems must be good enough to fool our senses
- Some question to answer
  - Intensity resolution and range: intensity resolution per pixel
    - Essentially the same as color (red, blue, and green) resolution
  - Spatial resolution: how many pixels per square meter?
  - Temporal resolution: how fast do displays need to change their pixels? (next lecture)

Intensity resolution and range

- Existing displays have only 256 intensity levels per color
- Photoreceptors, however, span seven orders of magnitude of light sensitivity
- Current displays do not even support scotopic vision (vision of the eye under low light conditions)
  - Very high intensity resolution would be needed at extremely low light conditions

What pixel density is Needed So that We Do Nor Perceive Individual Pixels?

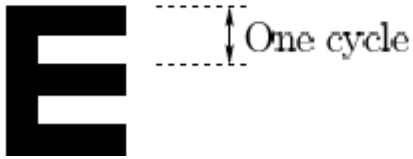


PPI: Pixel Per Inch

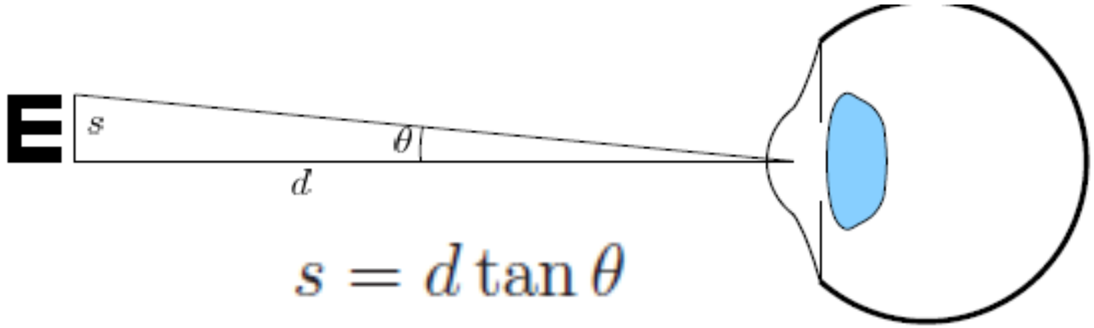
- In 2016, Steve Jobs claimed 326 PPI is enough, achieving what Apple called retina display!

Detecting Pixels

- Scientists usually want to know: “how small do pixels need to be so that a small white pixel is not detectable against a black background?”
- To derive this we need some calculations:
  - Cycles: number of stripes (or sinusoidal peaks) that can be seen as separate along a viewing arc
- A person with 20/20 vision can barely make out stipes at 30 cycles per degree



Determining Required PPI



$$s = d \tan \theta$$

- To generate 30 cycles per degree we need at least 60 pixels per degree (to create black and white)
- For a TV that is 20 feet away:

$$s = 20 * \tan 1^\circ = 0.349\text{ft, which is equivalent to } 4.189\text{in.}$$

- Thus, the sufficient PPI would be $60/4.189 = 14.32$

Determining PPI for Smartphone and VR Headset

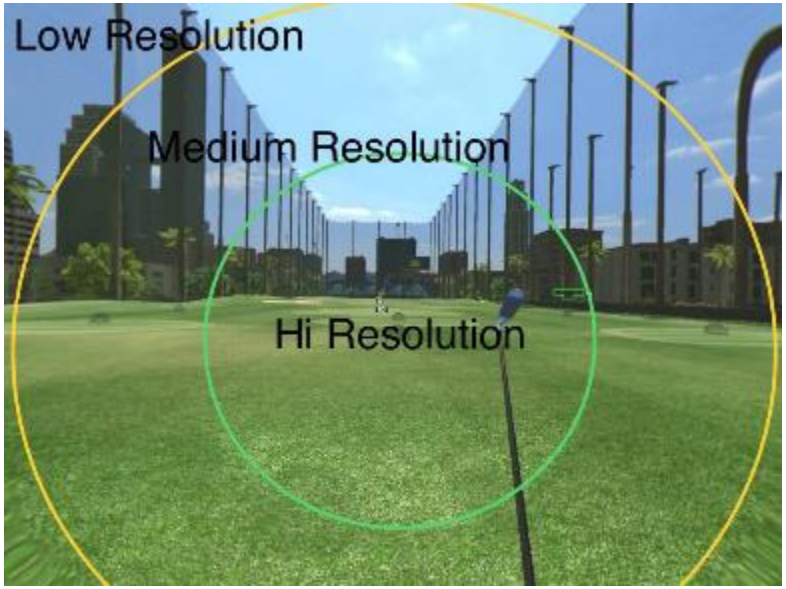
- A smartphone, is typically held at 12 inches from the user's eye
  - The screen would at least require 286 PPI
- VR lens magnifies the image!
- Suppose lens is positioned at its focal distance with respect to the display screen (1.5 inch)
  - The required PPI would be 2291
  - SONY Xperia XZ Premium has the highest display PPI (809)
- People with better eyesight may need even a much higher PPI!

Required Field-of-View

- Based on eye and head movement constraints, the maximum field-of-view is 270 degrees
- A head-mount VR display would need a less FoV (120 degrees)
  - A flat display may provide sufficient FoV
  - Lens aberrations at the peripheries limit the effective FoV

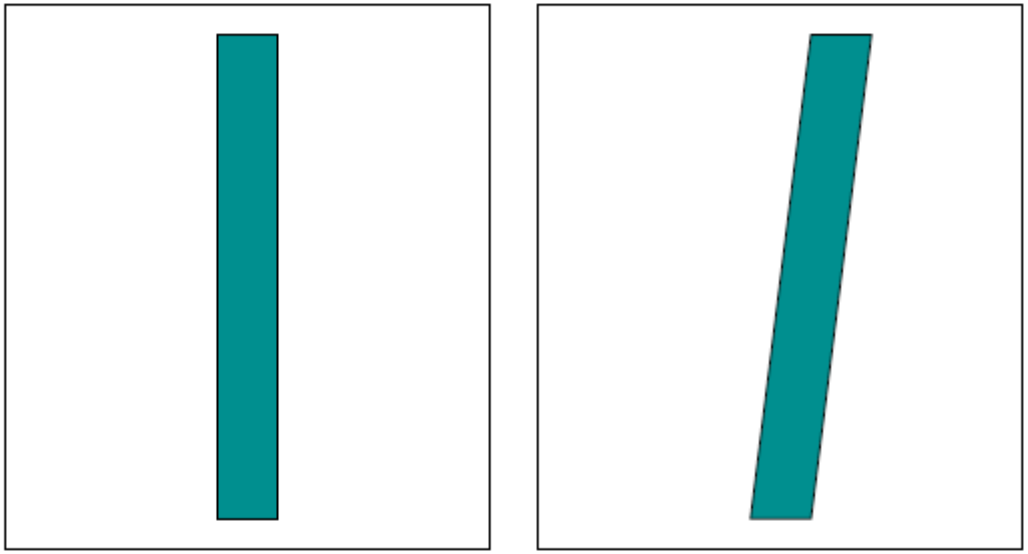
Research: Foveated Rendering

- So far we did not explore the fact that fovea has the highest density of cones
  - Photoreceptor density decreases away from the fovea
- **We could track where the eye is moving and have a tiny movable screen with a high density!**
- Or we could have a fixed display density **but focus the graphical rendering** to the place the eye is focusing



Display Scanout

- Some displays work like CRT monitors, updating pixels line by line
  - Due to the slow charge and response time of photoreceptors, we do not perceive the scan out pattern during normal use
  - When our eyes, features in the scene, or both are moving, then side effects of rolling scan outs become perceptible (e.g., a regular rectangle that moves in the VR world maybe distorted)
  - Another problem in displays is that pixels may take up to 20ms to switch. This may blur sharp edges!



Retina Image Slip

- During ordinary daily activities, the image of a feature may move due to motion and optical distortions
  - This is called retina image slip
- Once a VR headset is used, the motions of image on the retina may not match what would happen in the real world
  - Maybe due to optical distortions, tracking latency, and display scanout
  - Likely to contribute to fatigue and VR sickness