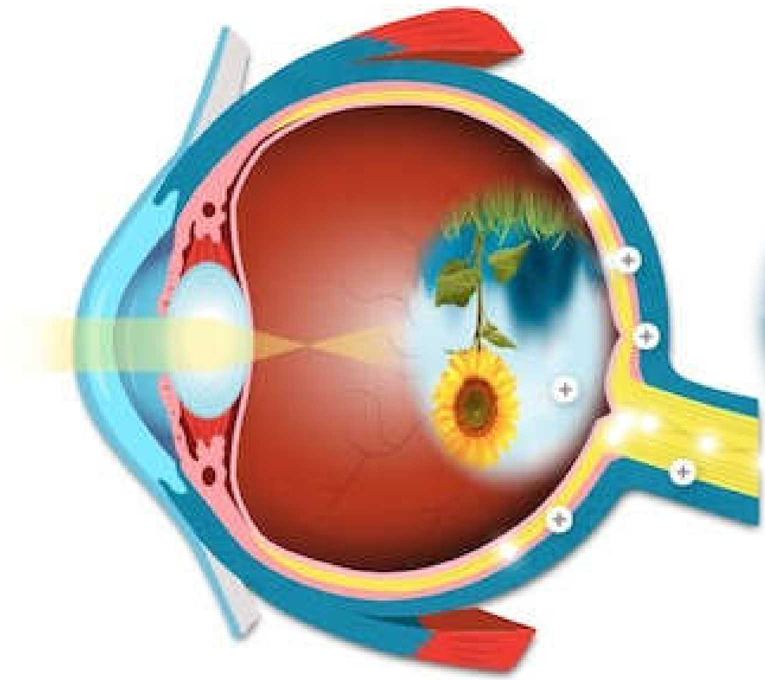
Physiology of Human Vision and Visual Perception - Implications for Virtual Reality (VR)

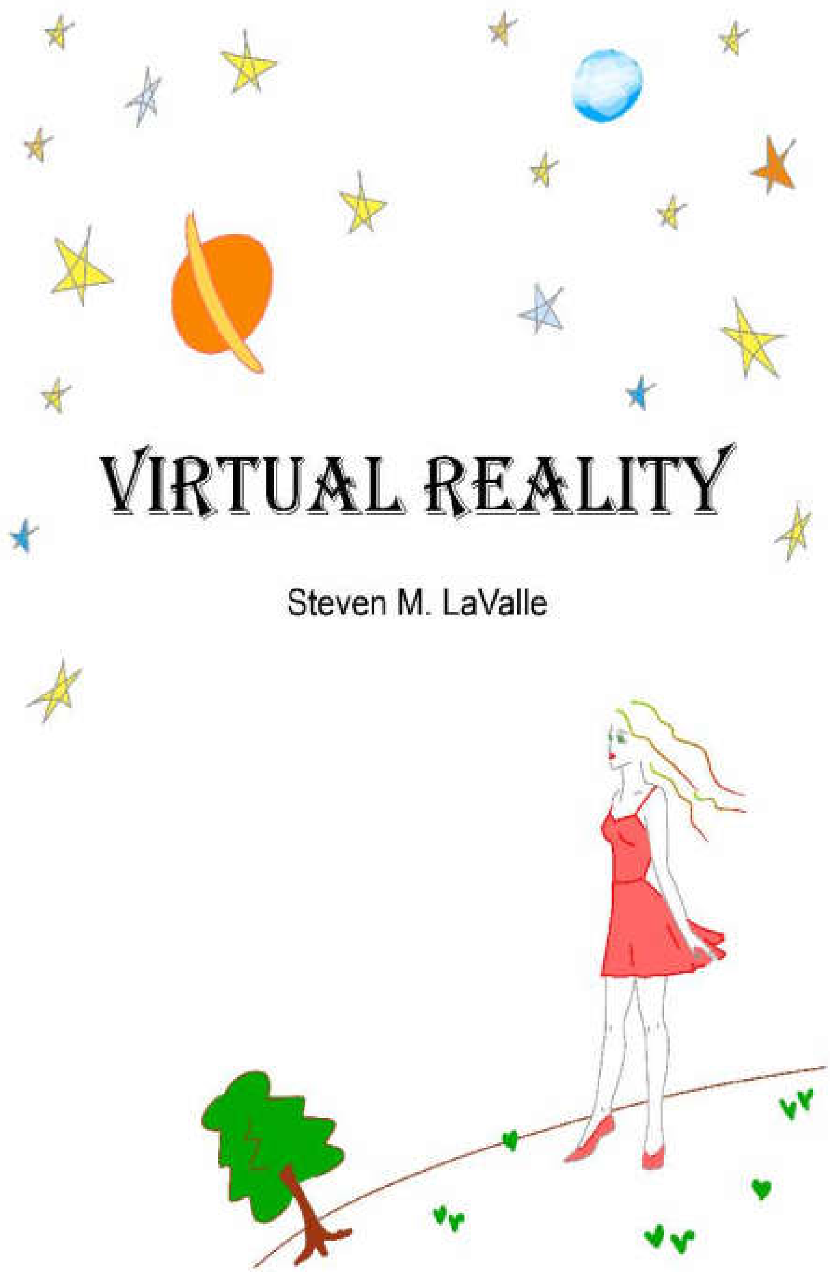


Aleksandra Klašnja-Milićević
Faculty of Scinces, Novi Sad, Serbia
Department of Mathematics and Informatics

Source: VIRTUAL REALITY book, Steven M. LaValle. Cambridge University Press. 2017. http://vr.cs.uiuc.edu/

Virtual reality

- **Virtual reality (VR) is a powerful technology that promises to change our lives unlike any other.**
- By artificially stimulating our senses, our bodies become tricked into accepting another version of reality.
- VR is like a waking dream that could take place in a magical cartoon-like world, or could transport us to another part of the Earth or universe.
- It is the next step along a path that includes many familiar media, from paintings to movies to video games.
- We can even socialize with people inside of new worlds, either of which could be real or artificial.



VIRTUAL REALITY

Steven M. LaValle

University of Illinois

CAMBRIDGE
UNIVERSITY PRESS

Virtual reality

- The book is growing out of material for a popular undergraduate course on VR that author introduced at the University of Illinois in 2015 (with hardware support from Oculus/Facebook).

The author:

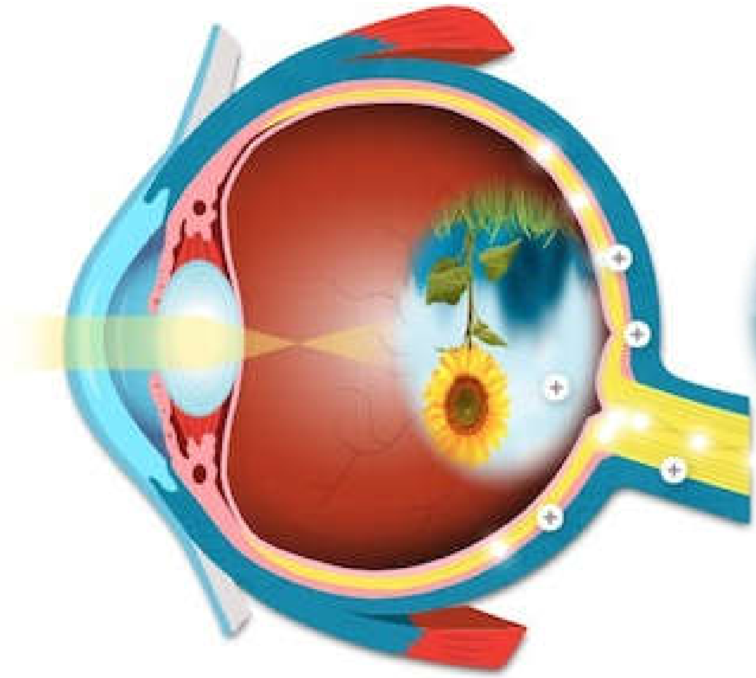
- „I have never in my life seen students so excited to take a Course.

  We cannot offer enough slots to come even close to meeting the demand“.

Virtual reality

- Therefore, the primary target of the book is undergraduate students around the world.
- The book would be an ideal source for starting similar VR courses at other universities.
- Although most of the interested students have been **computer scientists**, the course at Illinois has attracted **students from many disciplines**, such as psychology, music, kinesiology, engineering, medicine, and economics.
- Students in these other fields come with the most exciting project ideas because they can see how VR has the potential to radically alter their discipline.

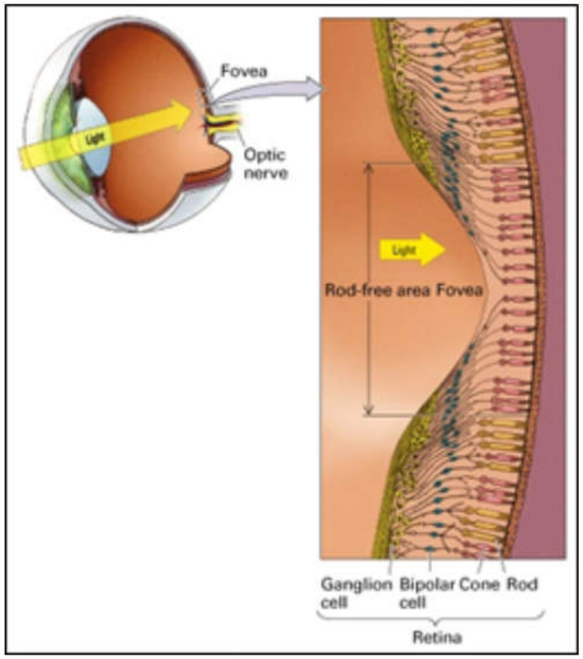
Physiology of Human Vision and Visual Perception

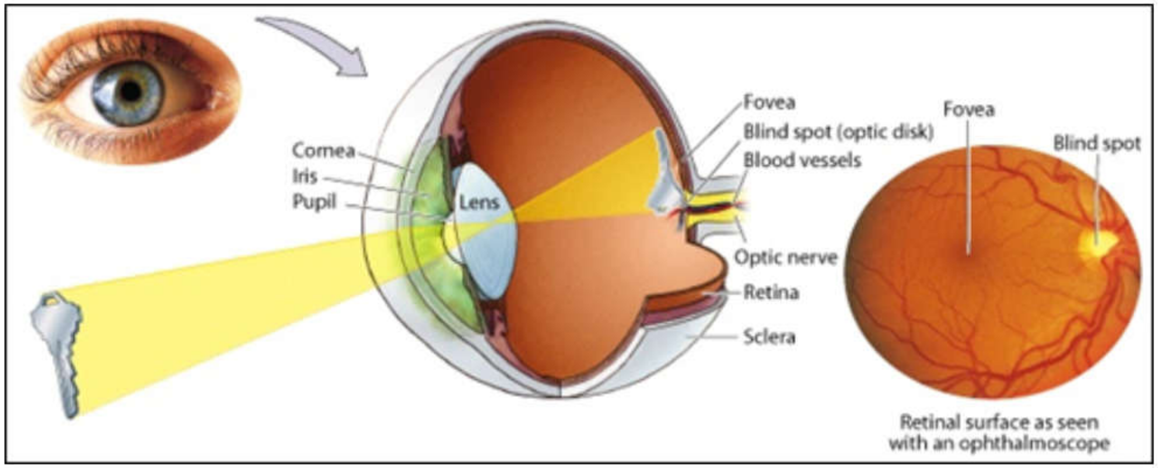


Aleksandra Klašnja-Milićević
Faculty of Scinces, Novi Sad, Serbia
Department of Mathematics and Informatics

Contents

- From the Cornea to Photoreceptors
- From Photoreceptors to the Visual Cortex
- Eye Movements
- Implications for VR

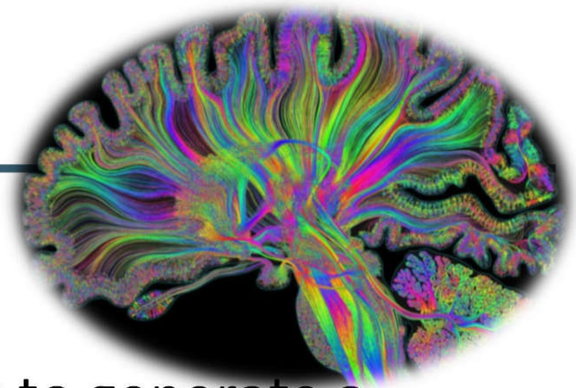




Introduction

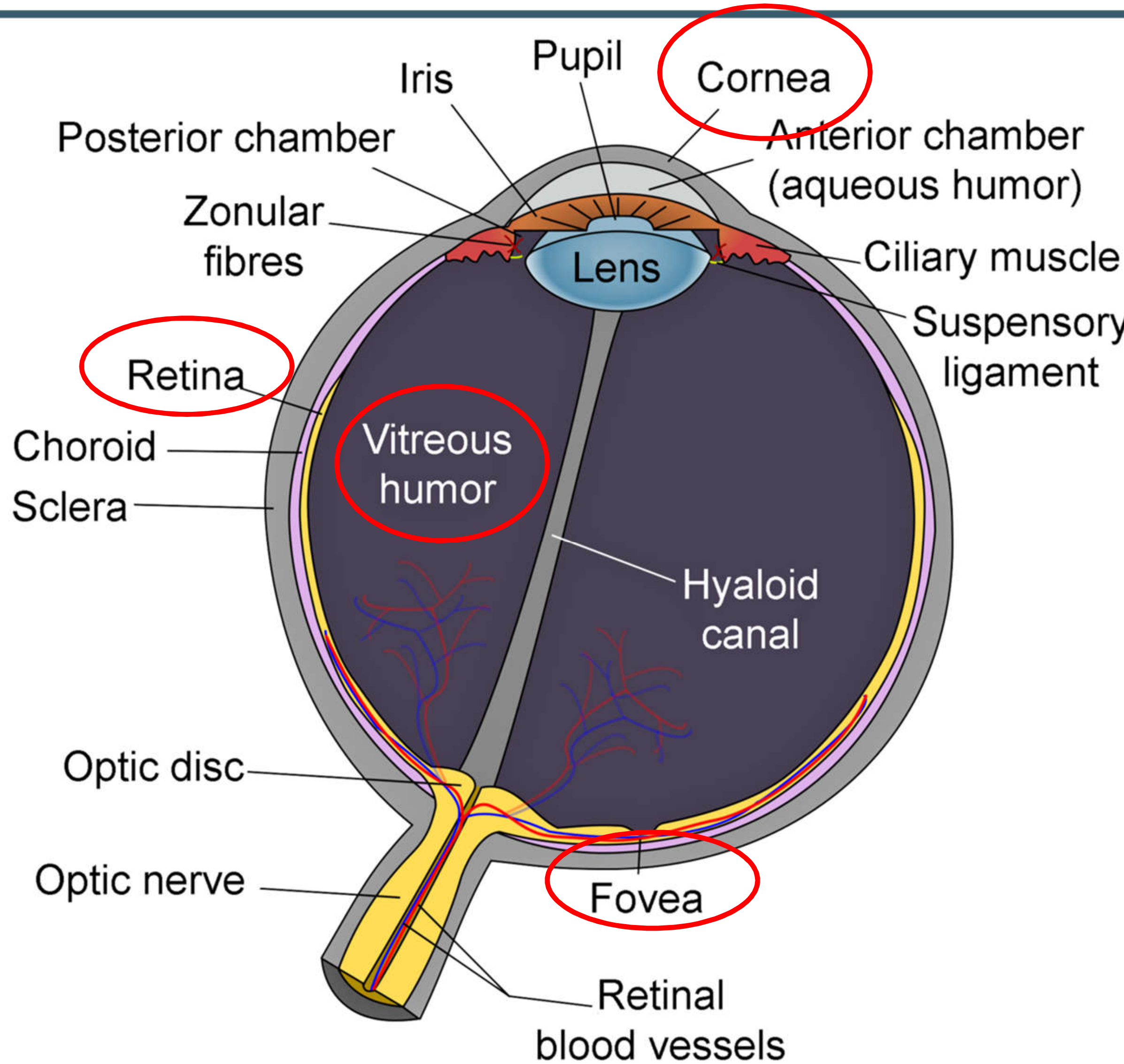
- What we perceive about the world around us is **"all in our head"**.
- The light around us - images on our retinas that capture colors, motions, and spatial relationships in the physical world.
- Normal vision - captured images may appear to have perfect clarity, speed, accuracy, and resolution, while being distributed over a large field of view.
- **However, we are being fooled.**

Introduction



- Perfection of our vision is mostly an illusion.
- **Neural structures** are filling in plausible details to generate a coherent picture in our heads that is consistent with our life experiences.
- When building Virtual Reality (VR) technology that co-opts these processes, it is important to understand how they work.
- They were designed to do more with less, and fooling these processes with VR produces many unexpected side effects:
- The display technology **is not a perfect replica** of the surrounding world.

Physiology of the human eye



From Wikipedia. This viewpoint shows how the right eye would appear if sliced horizontally (the nose would be to the left).

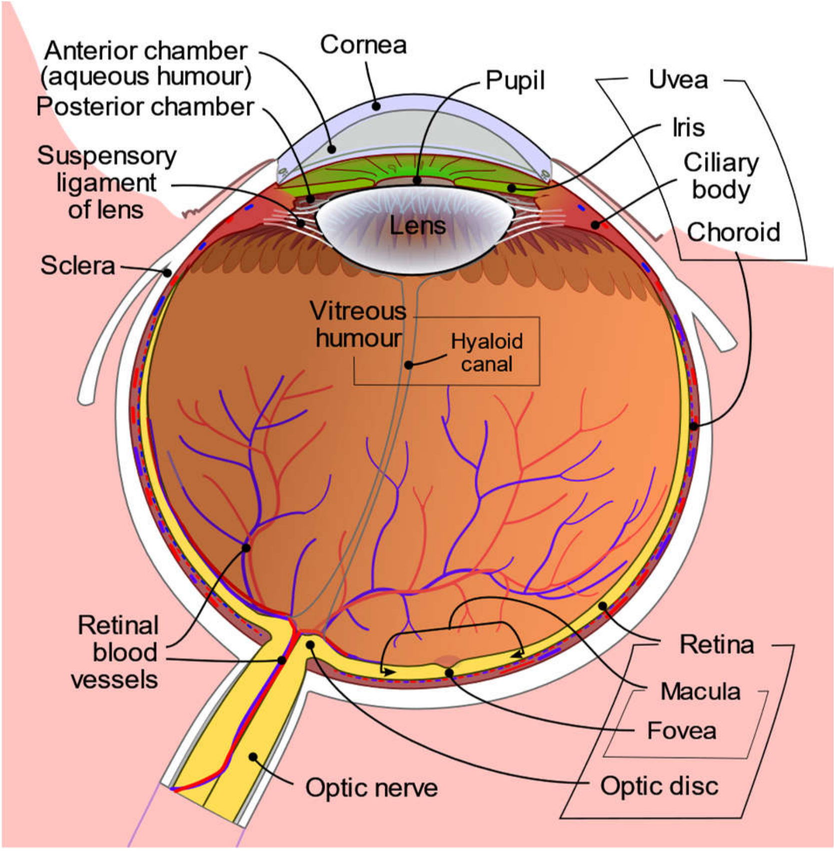
Physiology of the human eye

The shape is approximately spherical, with a diameter of around 24mm and only slight variation among people.

- The cornea is a hard, transparent surface through which light enters and provides the greatest optical.
- The rest of the outer surface of the eye is protected by a hard, white layer called the sclera.

Most of the eye interior consists of vitreous humor, which is a transparent, gelatinous mass that allows light rays to penetrate with little distortion or attenuation.

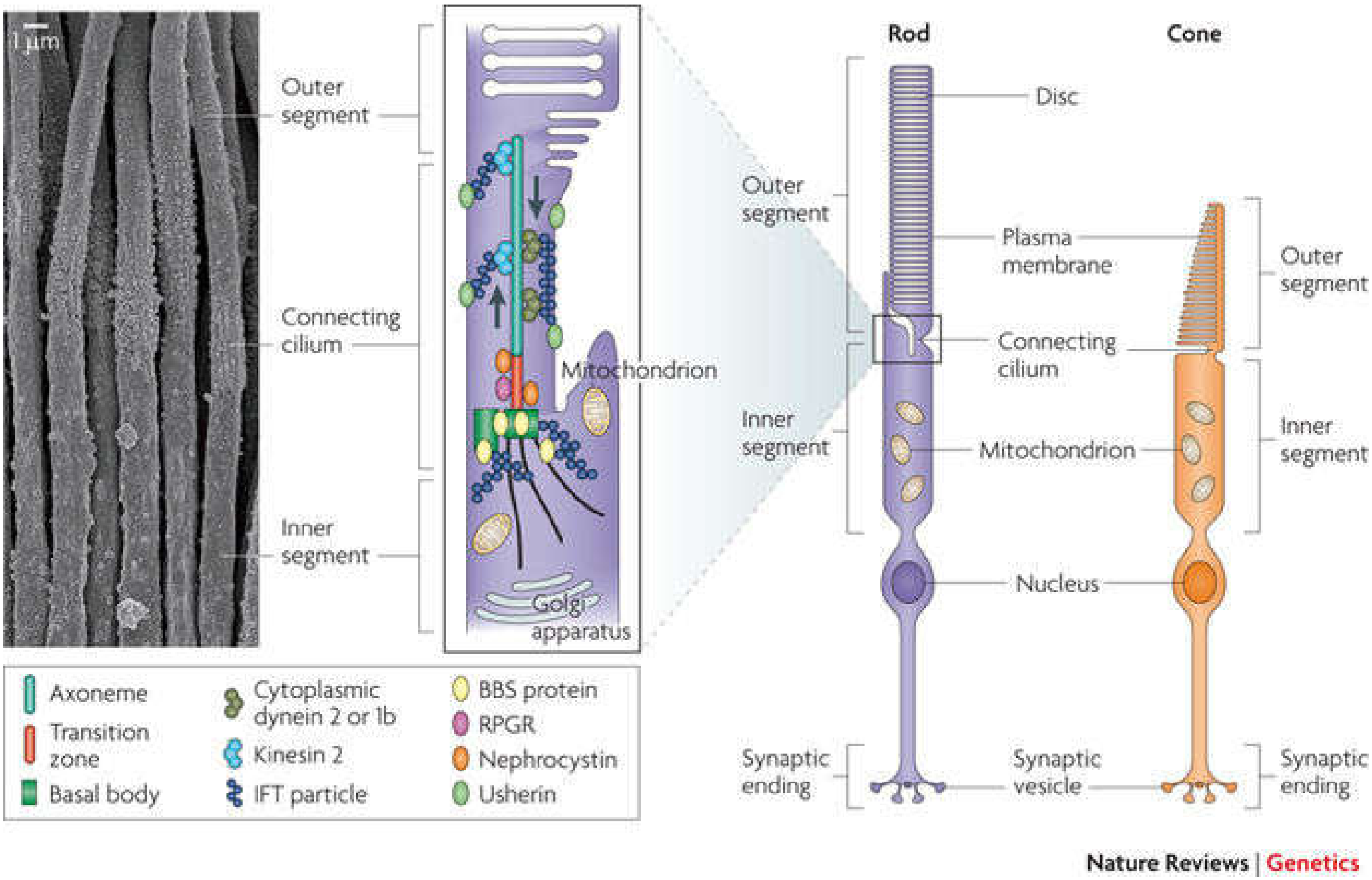
- As light rays cross the cornea, they pass through a small chamber containing aqueous humor, which is another transparent, gelatinous mass.



Photoreceptors

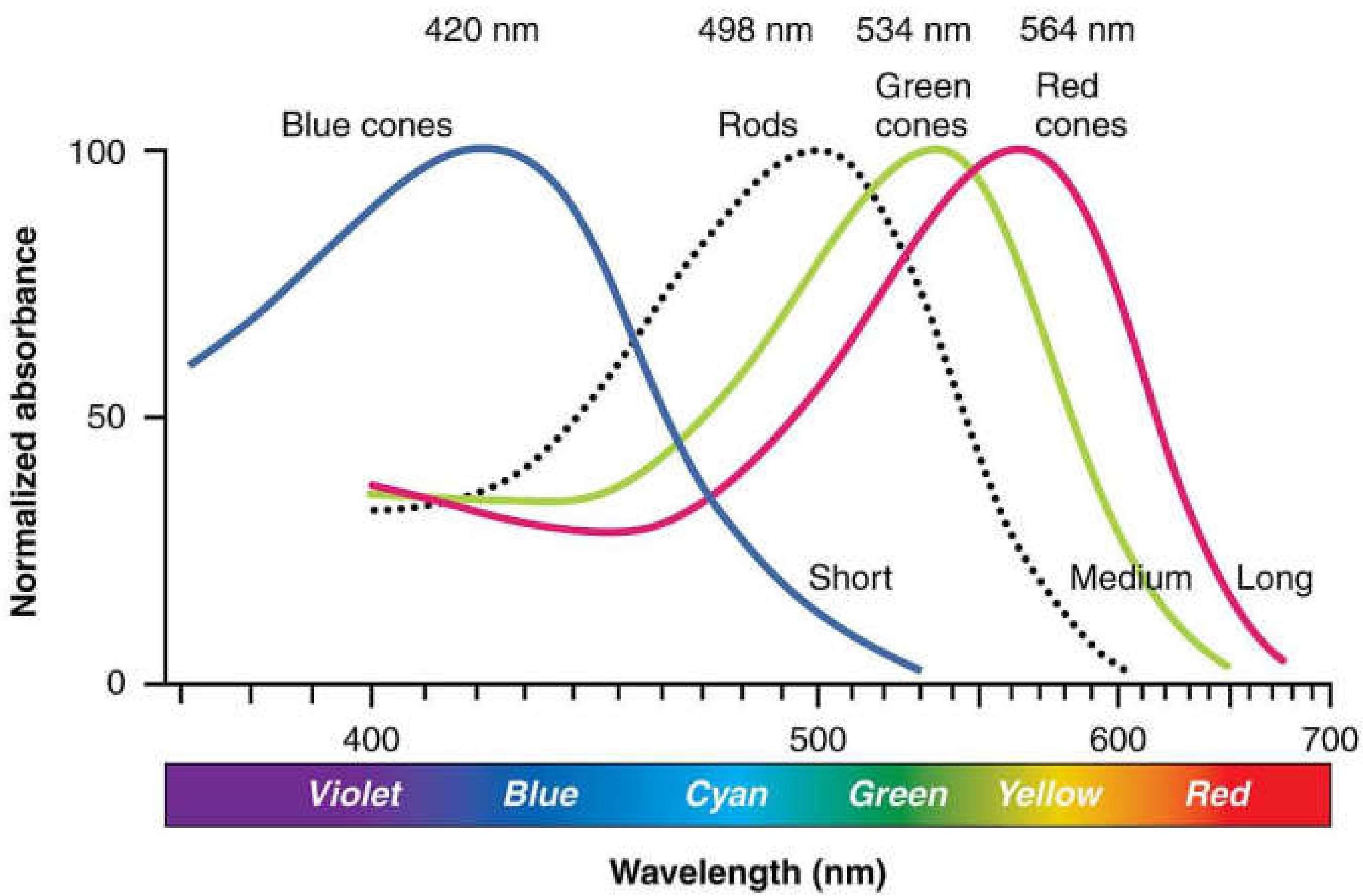
- The retina contains two kinds of photoreceptors for vision:
  - **rods**, which are triggered by very low levels of light, and
  - **cones**, which require more light and are designed to distinguish between colors.
- The width of the smallest cones is around 1000nm.
- This is quite close to the wavelength of visible light, implying that photoreceptors need not be much smaller.
- Each human retina contains about 120 million rods and 6 million cones that are densely packed along the retina.

Photoreceptors



On the left is an electron micrograph image of photoreceptors. The right shows the structure and components of rods and cones. The outer segments contain photopigments that electrochemically respond when attacked by photons.

The sensitivity of rods and cones as a function of wavelength



The detection capabilities of each photoreceptor type. Rod sensitivity peaks at 498nm, between blue and green in the spectrum. Three categories of cones exist, based on whether they are designed to sense blue, green, or red light.

Photoreceptors

- **Three categories of cones** exist, based on whether they are designed to sense blue, green, or red light.

- Photoreceptors respond to light levels over a large dynamic range. The **luminance** is measured in SI units of candelas per square meter, which corresponds directly to the amount of light power per area.

Light source	**Luminance (cd/m^2)**	**Photons per receptor**
Paper in starlight	0.0003	0.01
Paper in moonlight	0.2	1
Computer monitor	63	100
Room light	316	1000
Blue sky	2500	10,000
Paper in sunlight	40,000	100,000

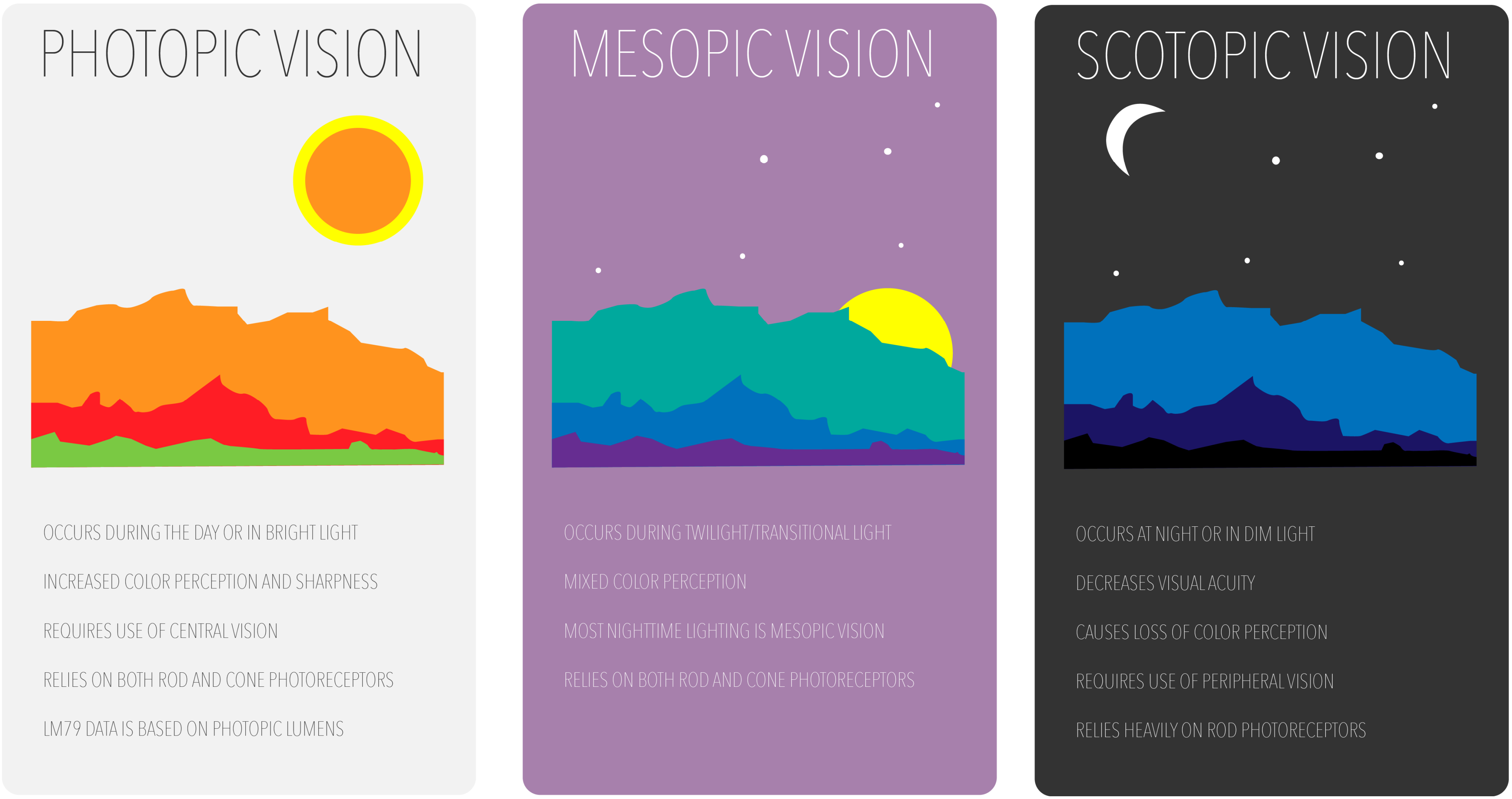
Figure 5.4: Several familiar settings and the approximate number of photons per second hitting a photoreceptor.

Photoreceptors - scotopic vision.

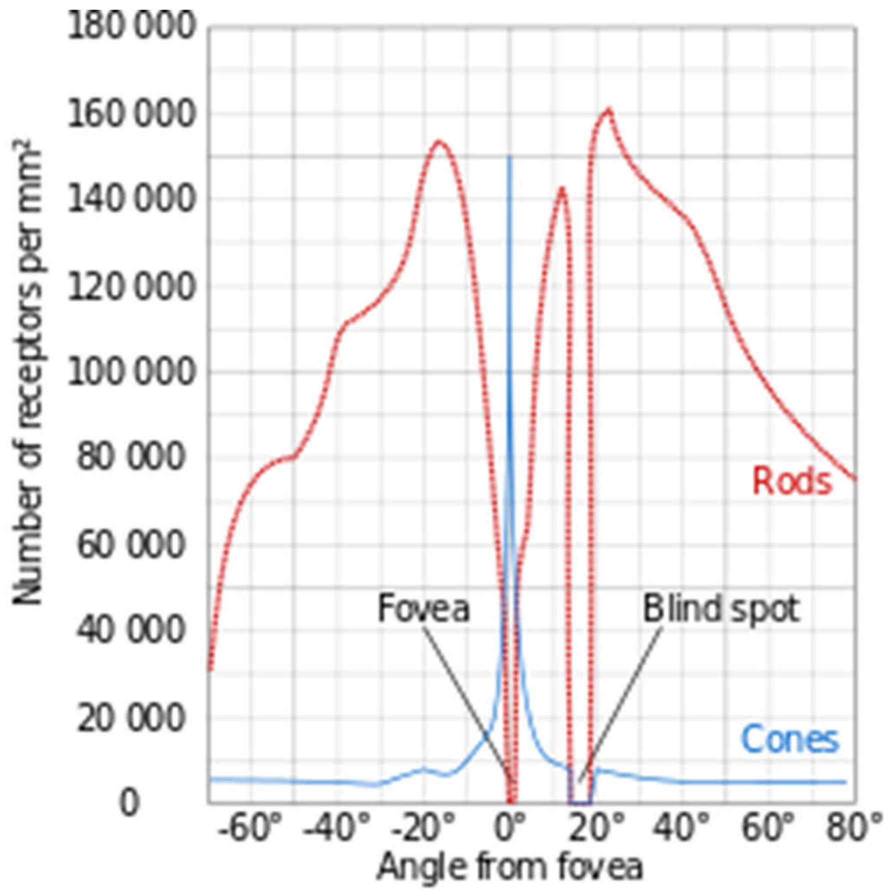
- The range spans seven orders of magnitude, from 1 photon hitting a photoreceptor every 100 seconds up to 100, 000 photons per receptor per second.
- At low light levels, only rods are triggered.
- Our inability to distinguish colors at night is caused by the inability of rods to distinguish colors.
- Our eyes may take up to 35 minutes to fully adapt to low light, resulting in a monochromatic mode called **scotopic vision**.

Photoreceptors - photopic vision

- By contrast, our cones become active in brighter light.
- Adaptation to this trichromatic mode, called **photopic vision.**



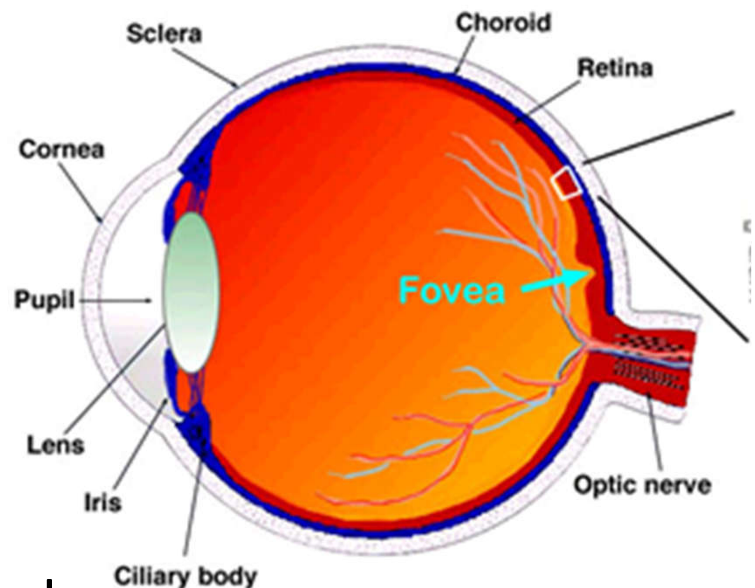
Photoreceptor density



With 20/20 vision, we perceive the world as if our eyes are capturing a sharp, colorful image over a huge angular range. This seems impossible, however, because we can only sense sharp, colored images in a narrow range. Furthermore, the blind spot should place a black hole in our image. Surprisingly, our perceptual processes produce an illusion that a complete image is being captured. This is accomplished by filling in the missing details using contextual information.

Photoreceptor density

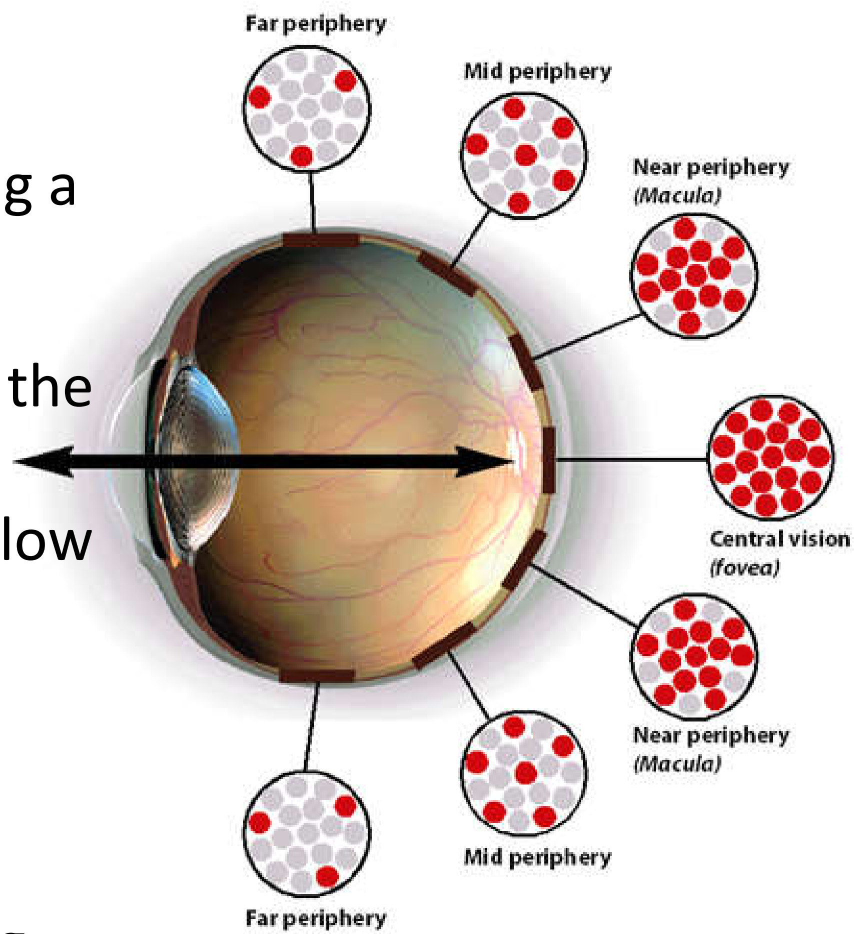
- The density of photoreceptors across the retina varies greatly, as plotted.
- The most interesting region is the **fovea**, which has the greatest concentration of photoreceptors.
- The innermost part of the fovea has a diameter of only 0.5mm or an angular range of ±0.85 degrees, and contains almost entirely cones.



- This implies that the eye must be pointed straight at a target to perceive a sharp, colored image.

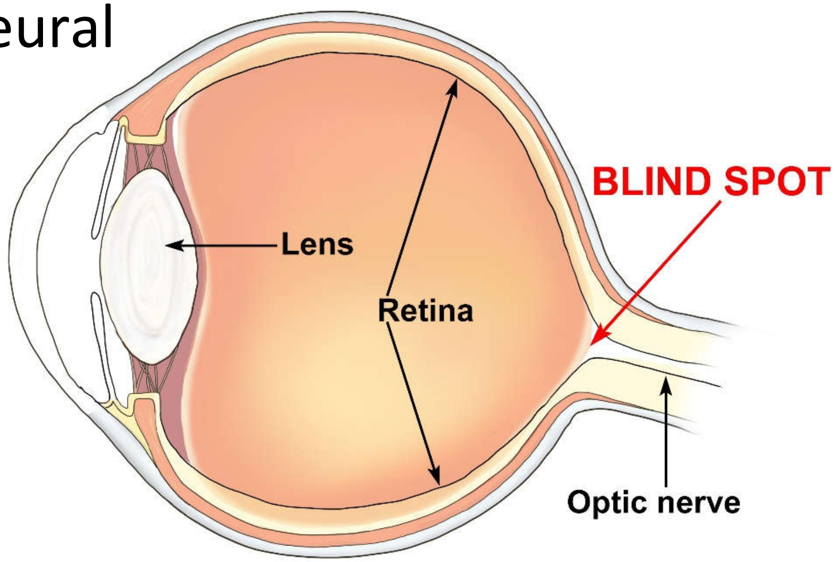
Photoreceptor density

- The entire fovea has diameter 1.5mm (±2.6 degrees angular range), with the outer ring having a dominant concentration of rods.
- Rays that enter the cornea from the sides land on parts of the retina with lower rod density and very low cone density.
- This corresponds to the case of **peripheral vision**.
- We are much better at detecting movement in our periphery, but cannot distinguish colors effectively.



Photoreceptor density

- Peripheral movement detection may have helped our ancestors from being eaten by predators.
- Finally, the most intriguing part of the plot is the **blind spot**, where there are no photoreceptors.
- This is due to our retinas being inside-out and having no other way to route the neural signals to the brain.
- The photoreceptor densities shown leave us with a conundrum.



Photoreceptor density

- With 20/20 vision, we perceive the world as if our eyes are capturing a sharp, colorful image over a huge angular range.
- This seems impossible, however, because we can only sense sharp, colored images in a narrow range.
- Furthermore, the blind spot should place a black hole in our image.
- Surprisingly, our perceptual processes produce an illusion that a complete image is being captured.
- This is accomplished by filling in the missing details using contextual information, and by frequent eye movements.

Blind spot experiment

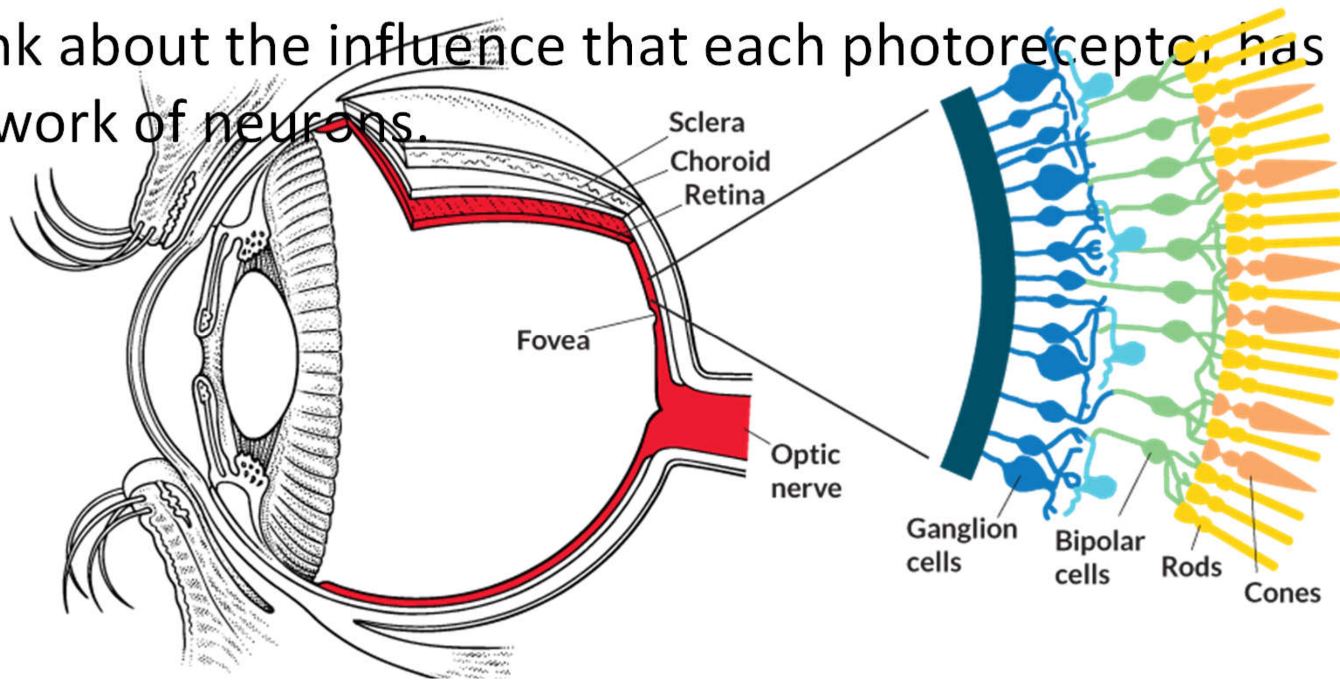
Figure 5.6: An experiment that reveals your blind spot. Close your right eye and look directly at the "X". Vary the distance of the paper (or screen) from your eye. Over some range, the dot should appear to vanish. You can carry this experiment one step further by writing an "X" and dot on a textured surface, such as graph paper. In that case, the dot disappears and you might notice the surface texture perfectly repeating in the place where the dot once existed. This is caused by your brain filling in the expected texture over the blind spot!

From Photoreceptors to the Visual Cortex

- Photoreceptors are transducers that convert the light-energy stimulus into an electrical signal called a neural impulse, thereby inserting information about the outside world into our neural structures.

- Signals are propagated upward in a hierarchical manner, from photoreceptors to the visual cortex.

- Think about the influence that each photoreceptor has on the network of neurons.



From Photoreceptors to the Visual Cortex

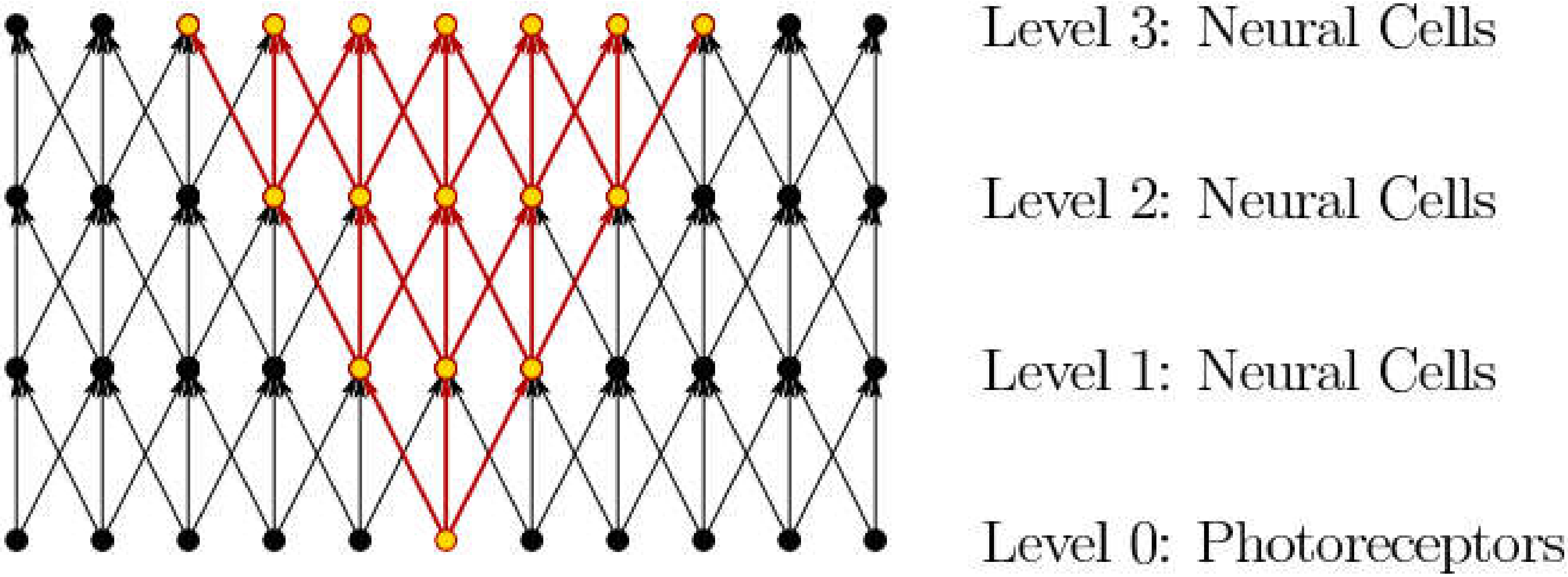


Figure 5.7: Four levels in a simple hierarchy are shown. Each disk corresponds to a neural cell or photoreceptor, and the arrows indicate the flow of information. Photoreceptors generate information at Level 0. In this extremely simplified and idealized view, each photoreceptor and neuron connects to exactly three others at the next level. The red and gold part highlights the growing zone of influence that a single photoreceptor can have as the levels increase.

From Photoreceptors to the Visual Cortex

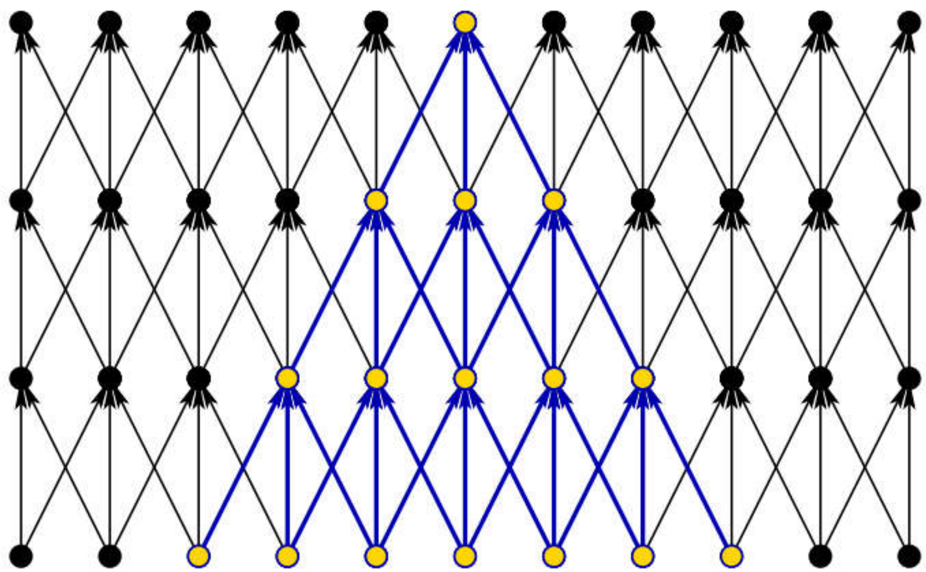
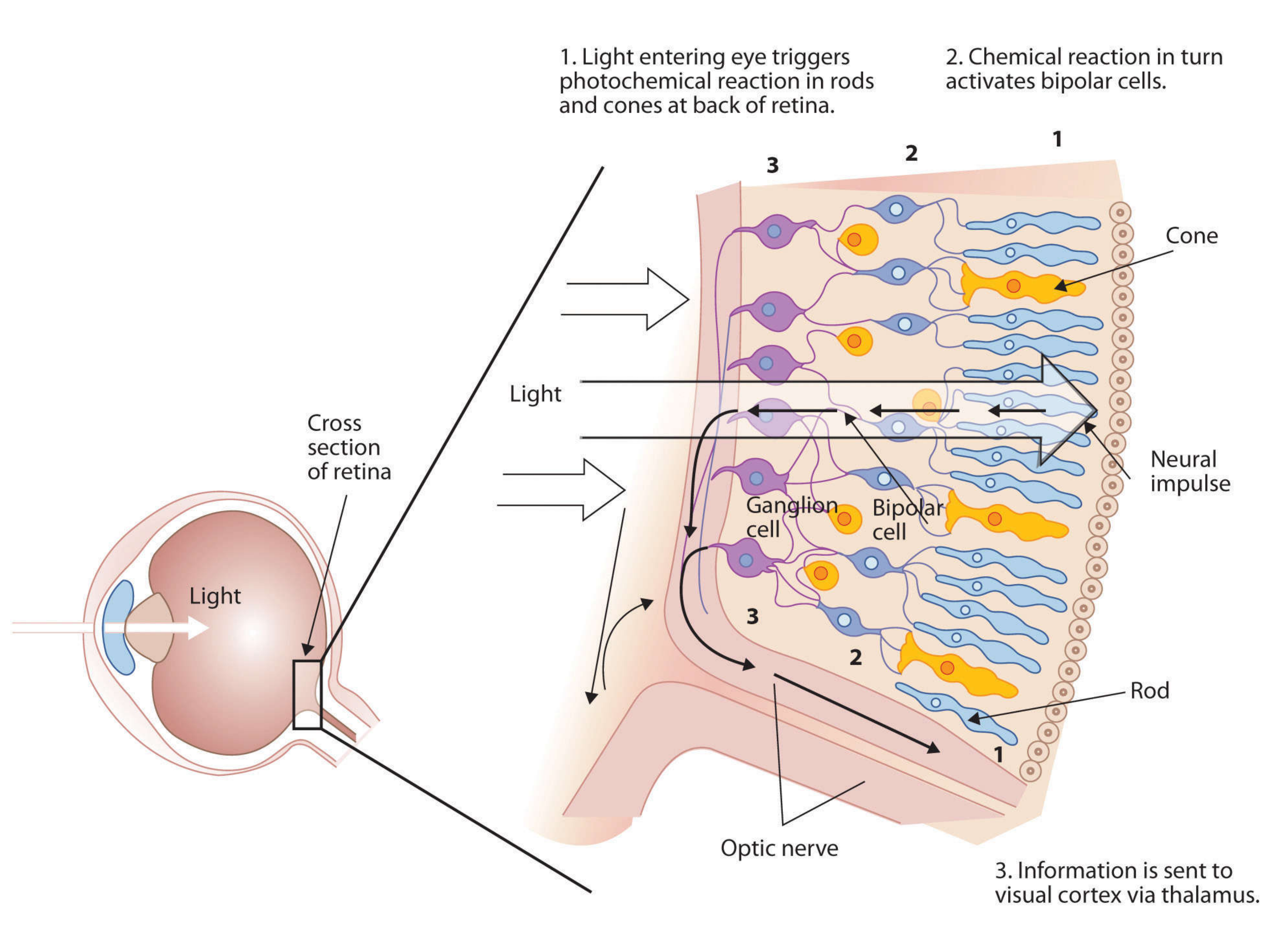
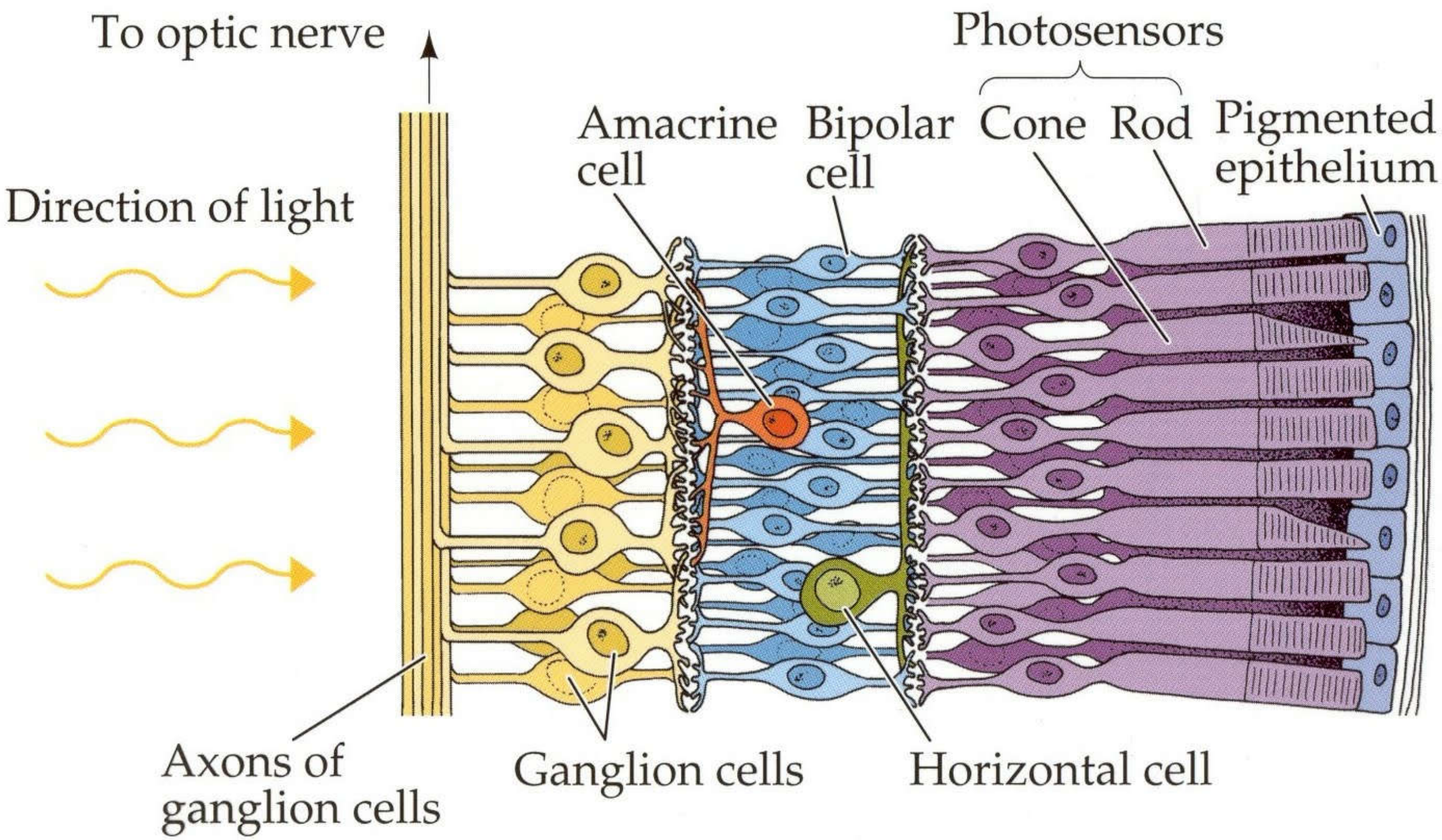


Figure 5.8: This diagram is the same as Figure 5.7 except that the information feeding into a single neuron is highlighted. Consider the set of photoreceptors involved in the reaction of a single neural cell. This is called the *receptive field*. As the level increases, the receptive field size grows dramatically. Due to the spatial arrangement of the photoreceptors, this will imply that each neuron responds to a growing patch in the image on the retina. The patch increases in size at higher levels.


1. Light entering eye triggers photochemical reaction in rods and cones at back of retina.
2. Chemical reaction in turn activates bipolar cells.
3. Information is sent to visual cortex via thalamus.
1
2
3
Cone
Light
Cross section of retina
Light
Neural impulse
Ganglion cell
Bipolar cell
3
2
1
Rod
Optic nerve

From Photoreceptors to the Visual Cortex



The information is sent from right to left, passing from the rods and cones to the bipolar, amacrine, and horizontal cells. These three types of cells are in the inner nuclear layer. From there, the signals reach the ganglion cells, which form the ganglion cell layer. Note that the light appears to be entering from the wrong direction: It passes over these neural cells before reaching the photoreceptors.

From Photoreceptors to the Visual Cortex

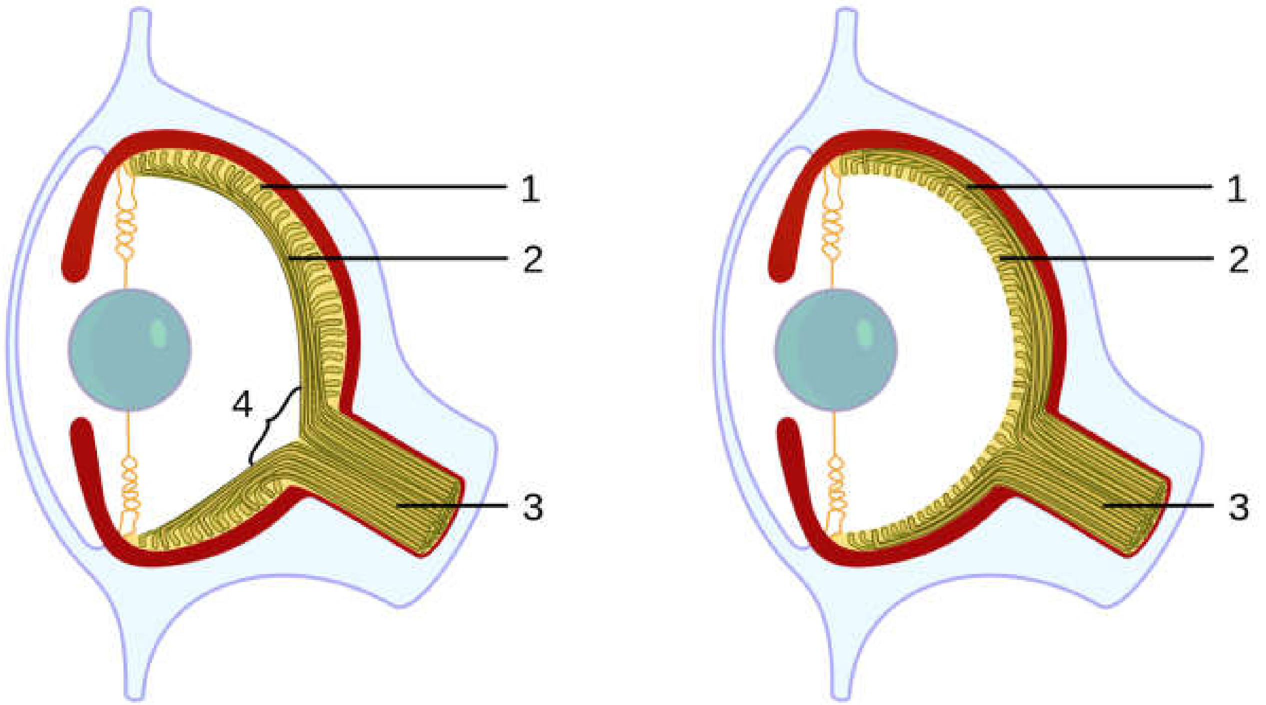
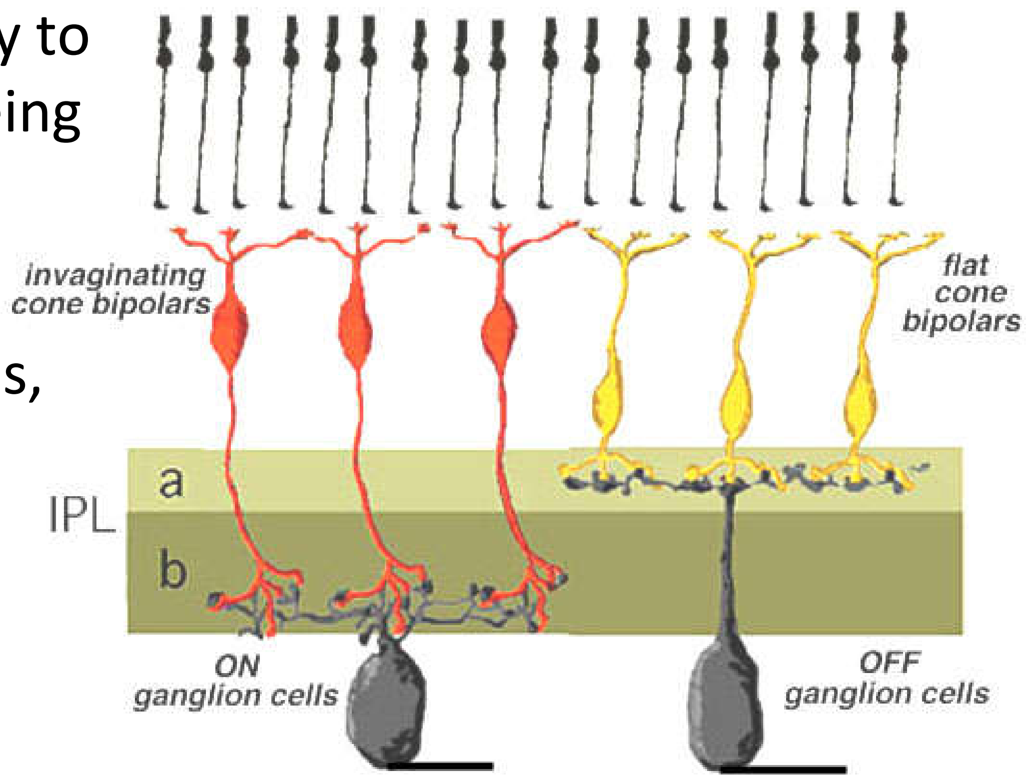


Figure 5.10: Vertebrates (including humans) have inside-out retinas, which lead to a blind spot and photoreceptors aimed away from the incoming light. The left shows a vertebrate eye, and the right shows a cephalopod eye, for which nature got it right: The photoreceptors face the light and there is no blind spot. (Figure by Jerry Crimson Mann.)

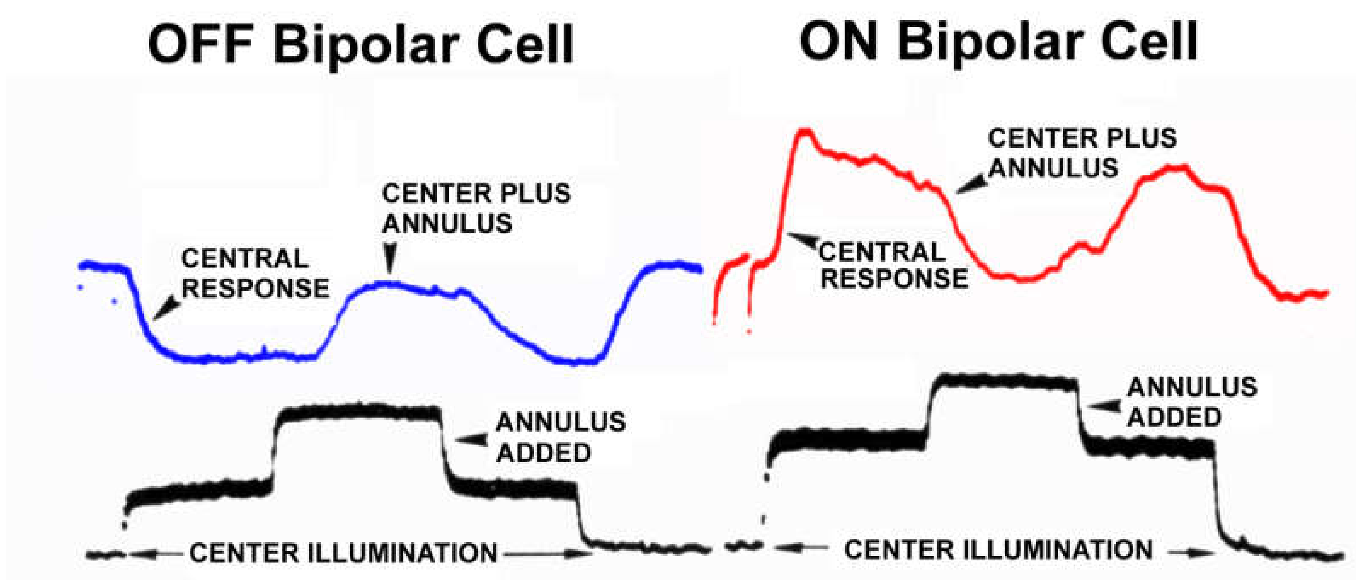
From Photoreceptors to the Visual Cortex

- It becomes clear that the neural cells are not arranged in the ideal way. The bipolar cells transmit signals from the photoreceptors to the ganglion cells.
- Some bipolars connect only to cones, with the number being between cones 1 and 10 per bipolar.
- Others connect only to rods, with about 30 to 50 rods per bipolar.



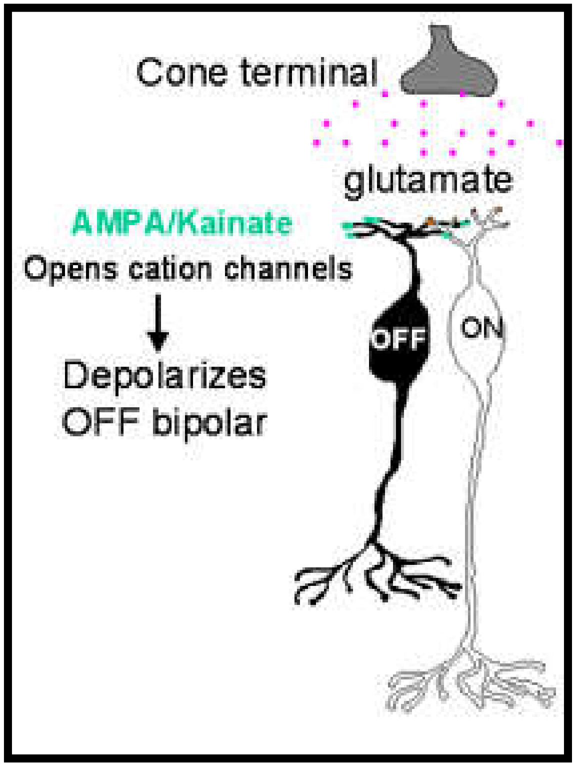
From Photoreceptors to the Visual Cortex

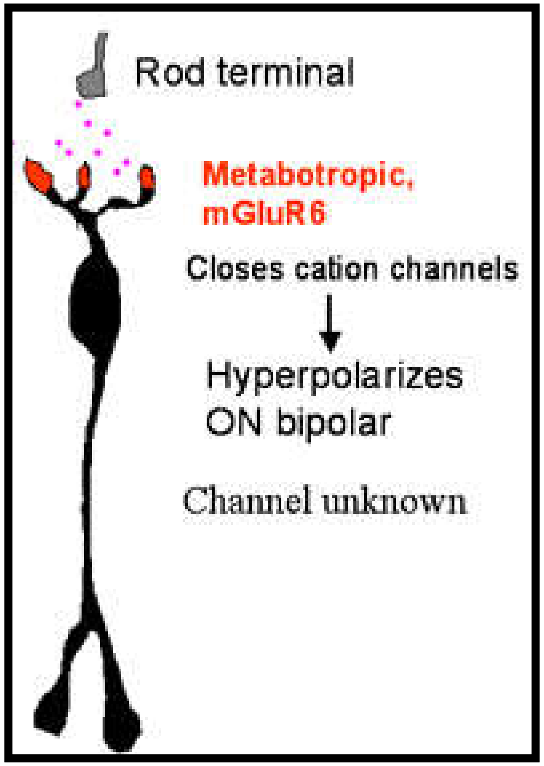
- There are two types of bipolar cells based on their function.
  - An **ON bipolar** activates when the rate of photon absorption in its connected photoreceptors increases.
  - An **OFF bipolar** activates for decreasing photon absorption.



From Photoreceptors to the Visual Cortex

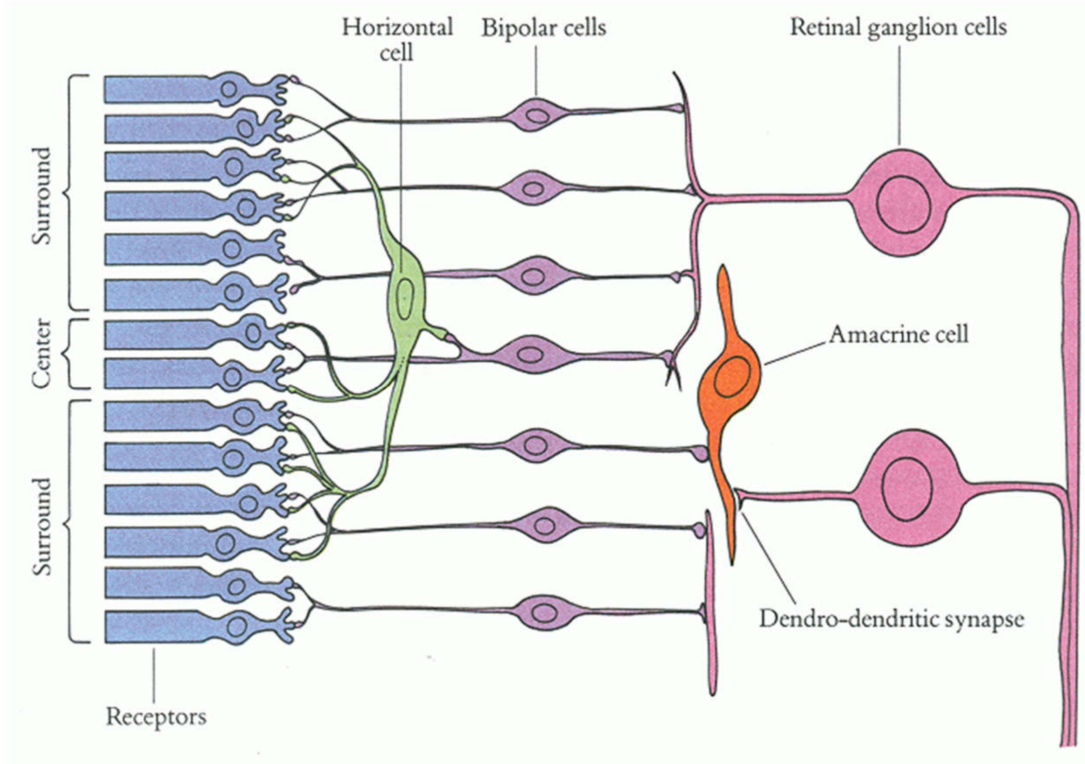
- The bipolars connected to cones have both kinds; however, the bipolars for rods have only ON bipolars.
- The bipolar connections are considered to be vertical because they connect directly from photoreceptors to the ganglion cells.





From Photoreceptors to the Visual Cortex

- **The horizontal cells** are connected by inputs (dendrites) to photoreceptors and bipolar cells within a radius of up to 1mm.
- Their output (axon) is fed into photoreceptors, causing lateral inhibition, which means that the activation of one photoreceptor tends to decrease the activation of its neighbors.

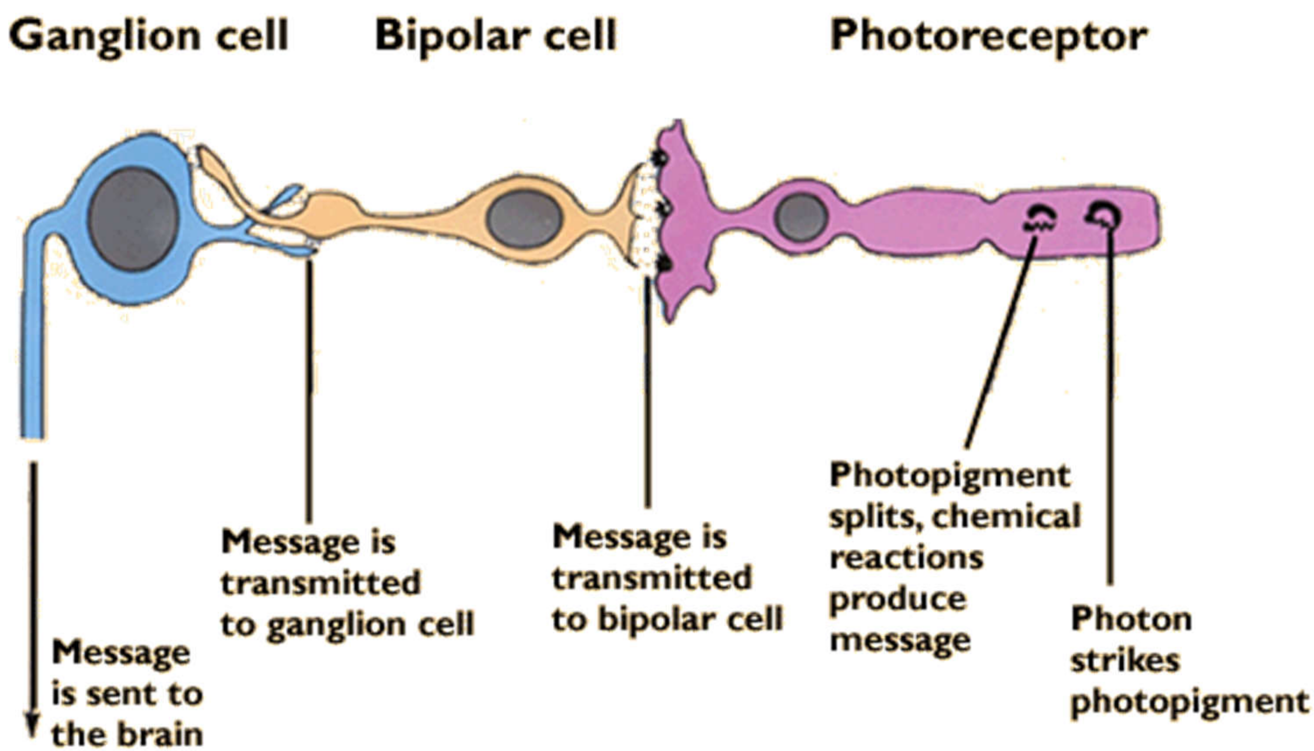


From Photoreceptors to the Visual Cortex

- Finally, amacrine cells connect horizontally between bipolar cells, other amacrine cells, and vertically to ganglion cells.
- There are dozens of types, and their function is not well understood. Thus, scientists do not have a complete understanding of human vision, even at the lowest layers.
- Nevertheless, the well understood parts contribute greatly to our ability to design effective VR systems and predict other human responses to visual stimuli.

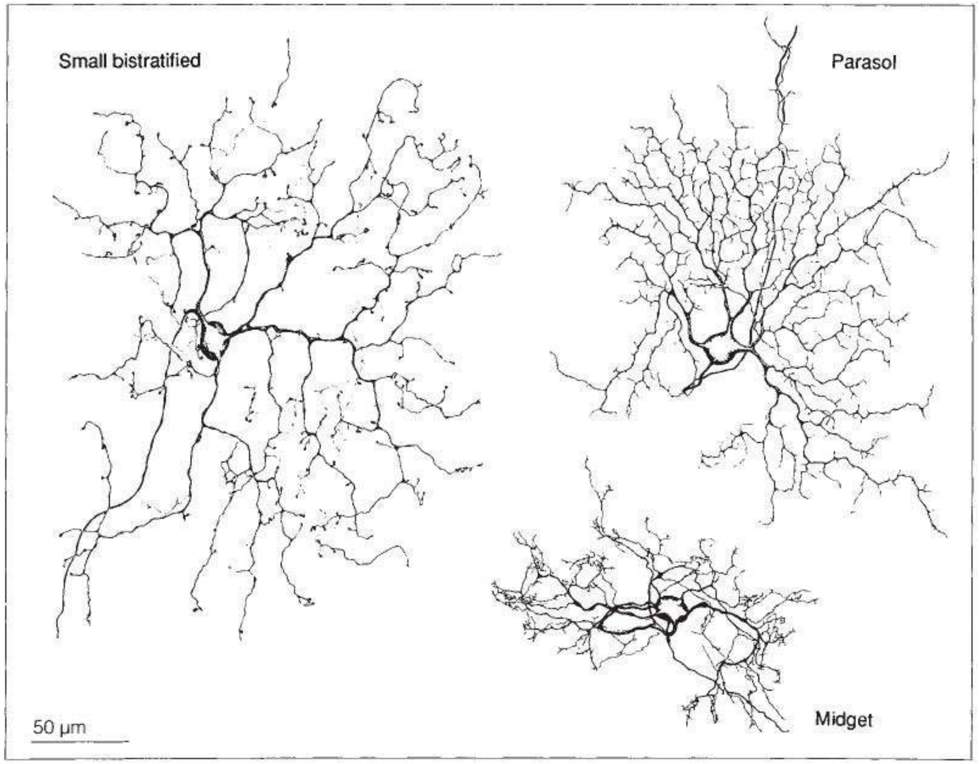
From Photoreceptors to the Visual Cortex

- At the ganglion cell layer, each ganglion cell has a large receptive field, which corresponds to the photoreceptors that contribute to its activation.



From Photoreceptors to the Visual Cortex

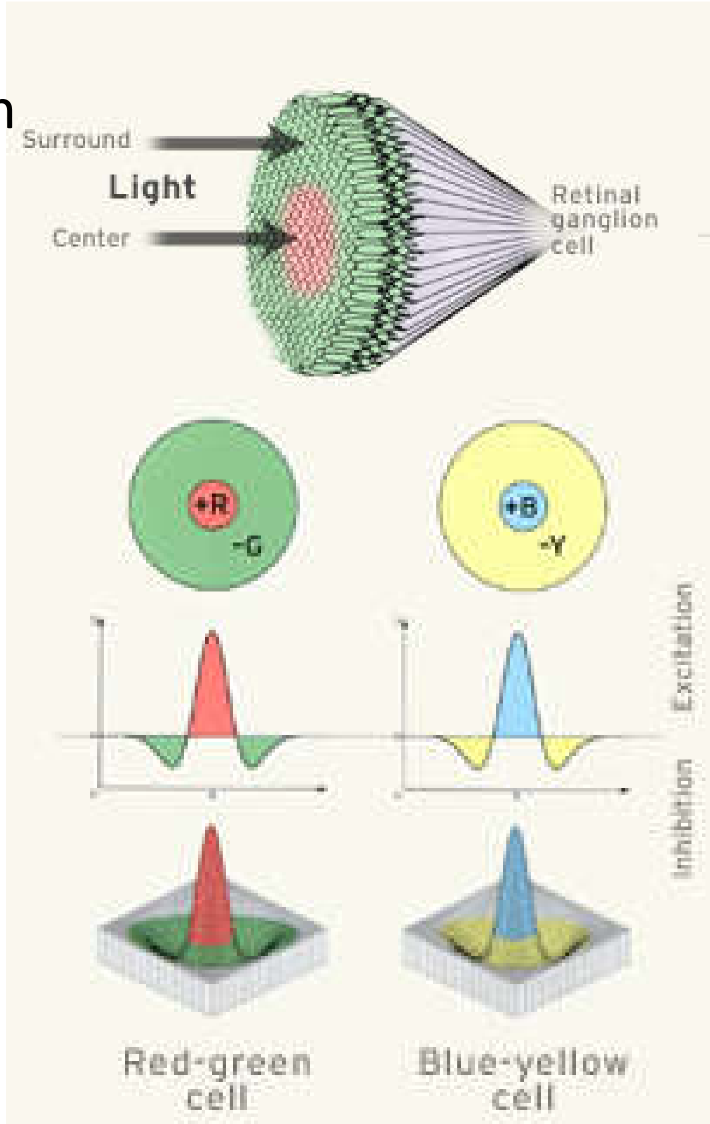
- The three most common and well understood types of ganglion cells are called:
  - midget,
  - parasol, and
  - bistratified.



- They perform simple filtering operations over their receptive fields based on spatial, temporal, and spectral (color) variations in the stimulus across the photoreceptors.

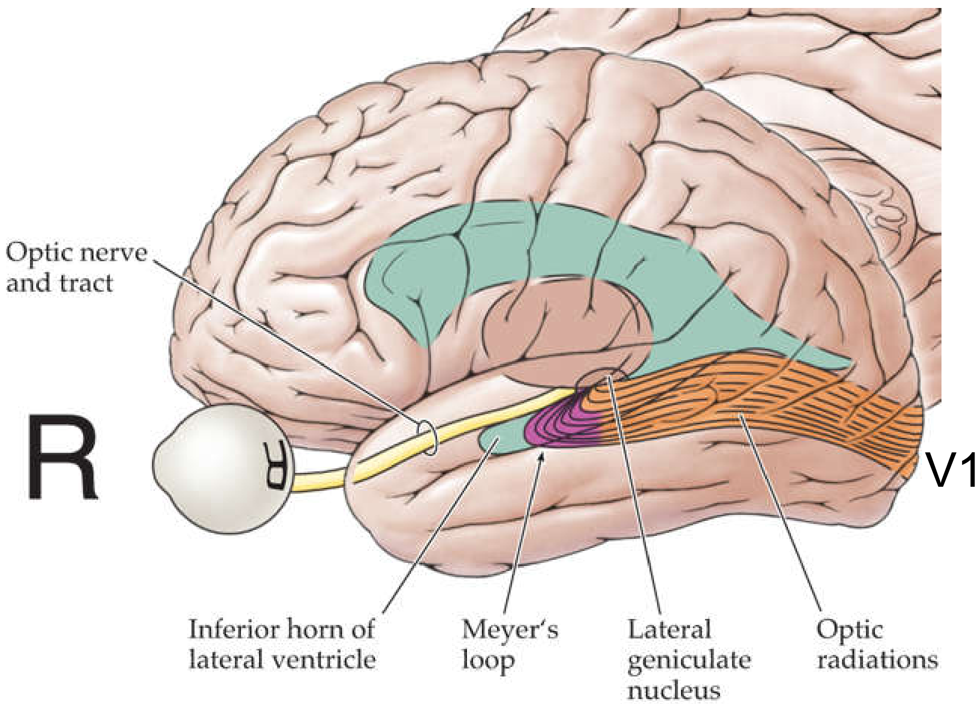
From Photoreceptors to the Visual Cortex

- Ganglion cell is triggered when red is detected in the center but not green in the surrounding area - **spatial opponency**.
- Neural structures are designed to detect local image variations.
- Thus, consider ganglion cells as tiny image processing units that can pick out local changes in time, space, and/or color.
- Once the ganglion axons leave the eye through the optic nerve, a significant amount of image processing has already been performed to aid in visual perception.



From Photoreceptors to the Visual Cortex

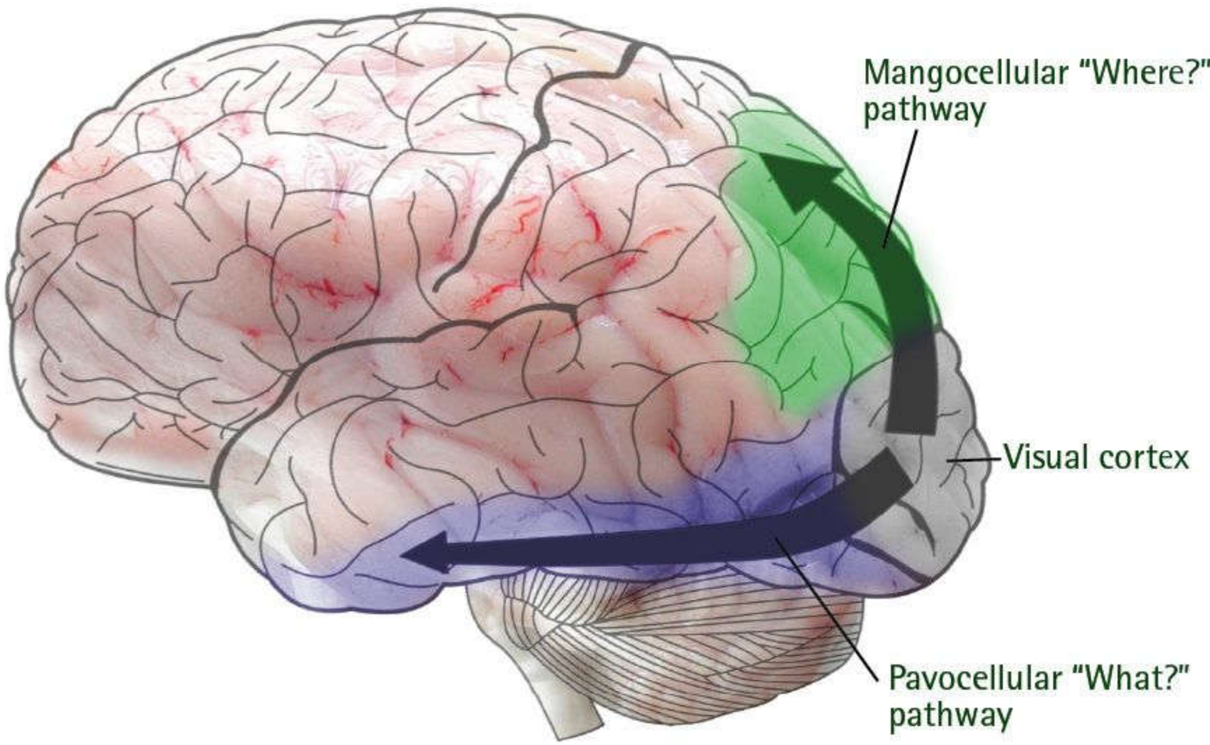
- The optic nerve connects to a part of the thalamus called **the lateral geniculate nucleus (LGN)**.
- The LGN mainly serves as a router that sends signals from the senses to the brain, but also performs some processing.
- The LGN sends image information to the primary visual cortex (V1), which is located at the back of the brain.



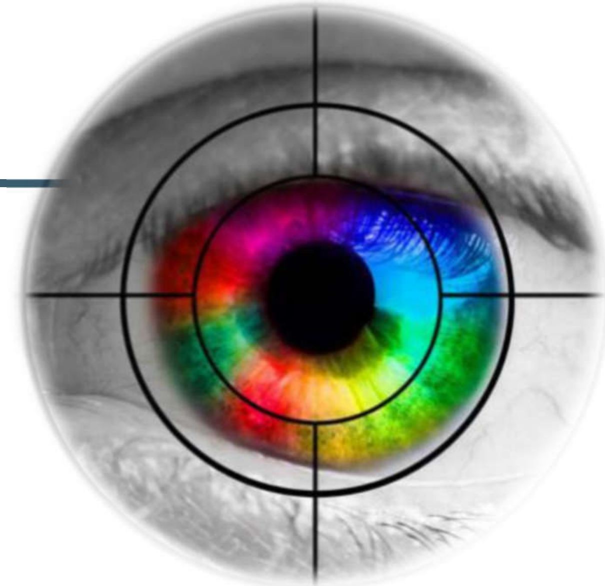
The visual pathway from the eyes to the LGN to the visual cortex

- The visual cortex is located in the back of the head.

- The visual cortex contains several interconnected areas that each perform specialized functions.



Eye Movements



- Eye rotations are a complicated and integral part of human vision.
- They occur both:
  - voluntarily and
  - involuntarily,
- and allow a person to fixate on features in the world, even as his head or target features are moving.
- One of the main reasons for eye movement is to position the feature of interest on the fovea.
- Only the fovea can sense dense, color images, and it unfortunately spans a very narrow field of view.
- To gain a coherent, detailed view of a large object, the eyes rapidly scan over it while fixating on points of interest.

Eye Movements

- Another reason for eye movement is that our photoreceptors are slow to respond to stimuli due to their chemical nature.
- They take up to 10ms to fully respond to stimuli and produce a response for up to 100ms.
- Eye movements help keep the image fixed on the same set of photoreceptors so that they can fully charge.
- This is similar to the image blurring problem that occurs in cameras at low light levels and slow shutter speeds.

Eye Movements

- Additional reasons for eye movement are to maintain a stereoscopic view and to prevent adaptation to a constant stimulation.
- To support the last claim, it has been shown experimentally that when eye motions are completely suppressed, visual perception disappears completely.
- As movements combine to build a coherent view, it is difficult for scientists to predict and explain how people interpret some stimuli.

Eye Movements

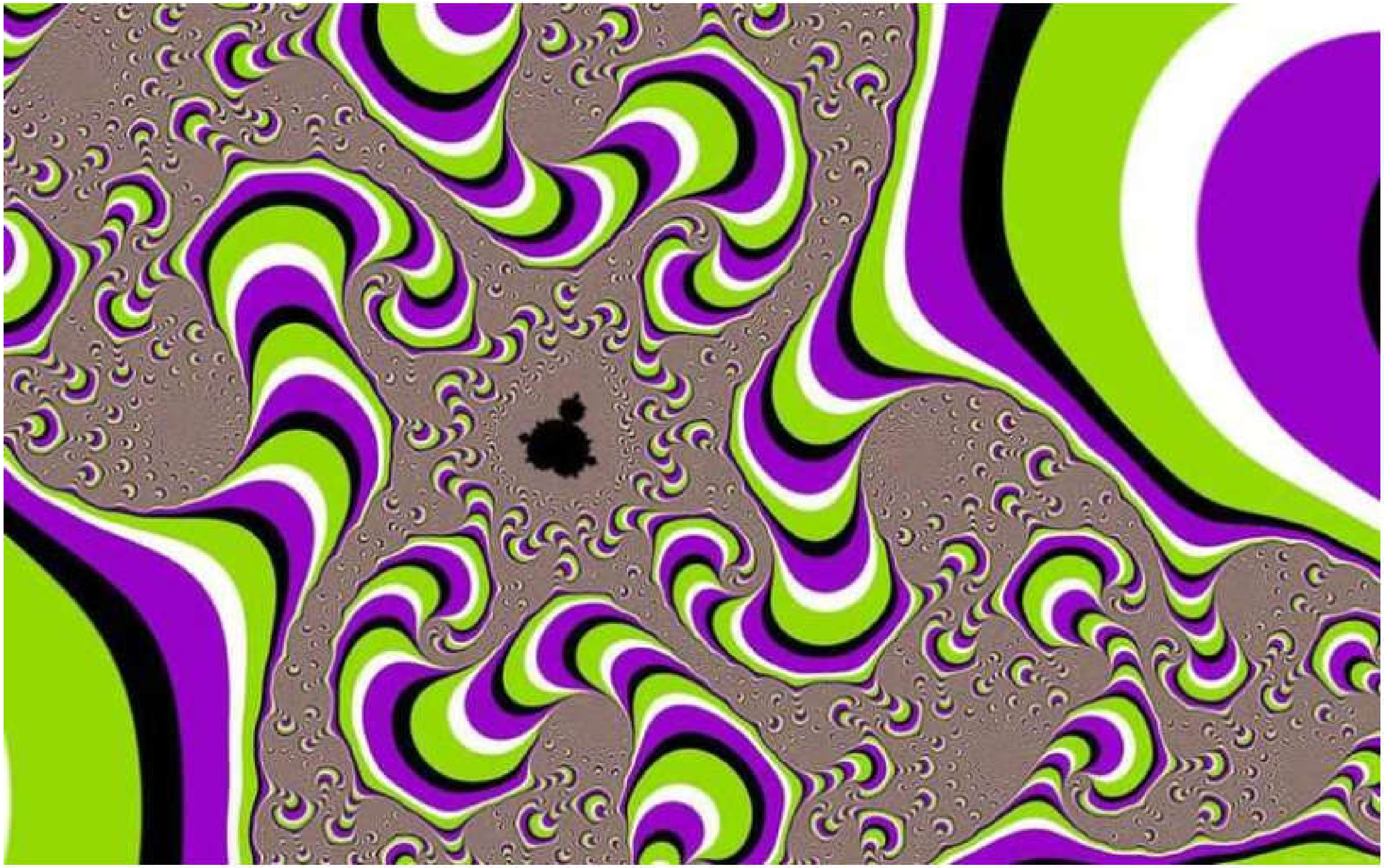


Figure 5.16: The fractal appears to be moving until you carefully fixate on a single part to verify that it is not.

Eye movement

- The rotation of each eye is controlled by six muscles that are each attached to the sclera (outer eyeball surface) by a tendon.

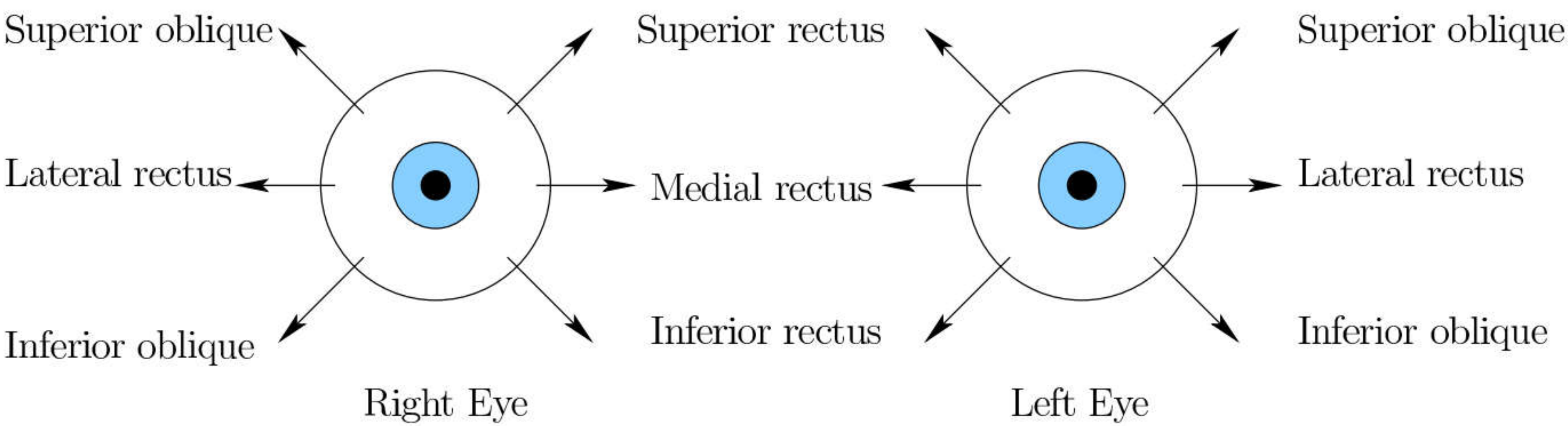


Figure 5.17: There are six muscles per eye, each of which is capable of pulling the pupil toward its location.

Eye movement

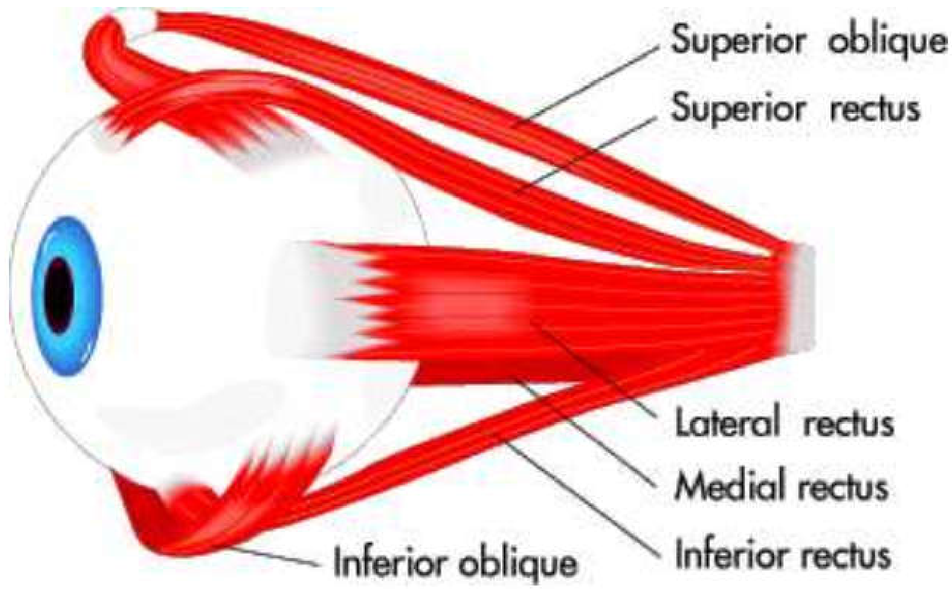
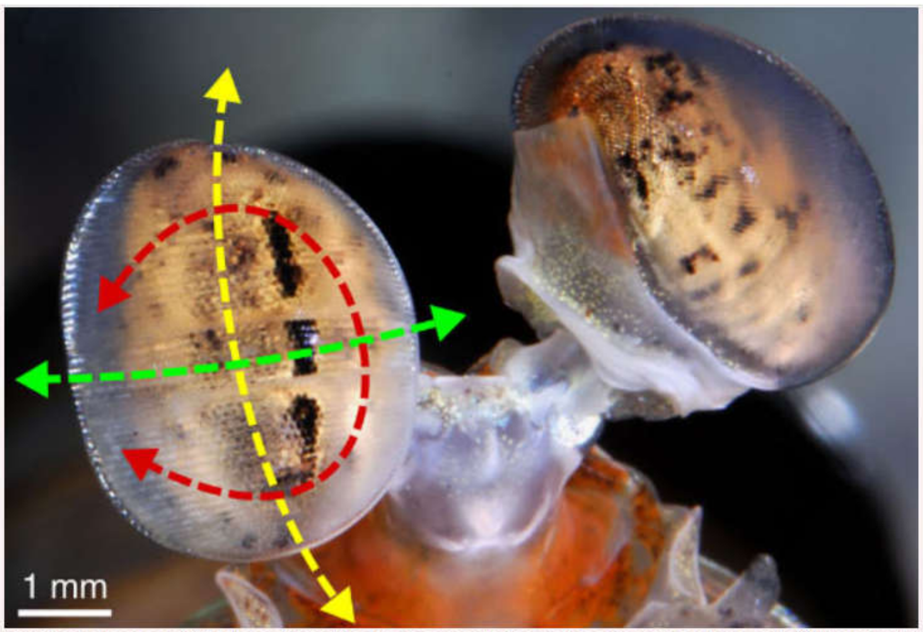


Figure 5.18: The six muscle tendons attach to the eye so that yaw, pitch, and a small amount of roll become possible.

Eye movement

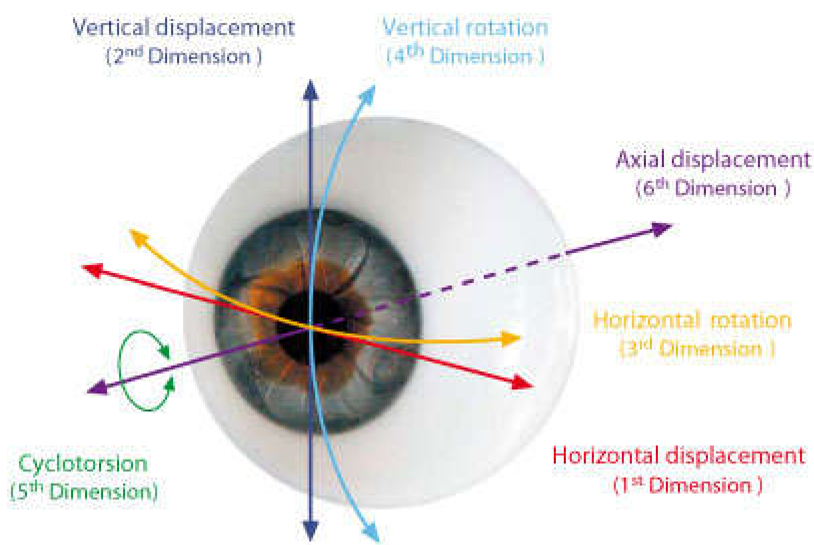
- For example, to perform a yaw (side-to-side) rotation, the tensions on the medial rectus and lateral rectus are varied while the other muscles are largely unaffected.
- To cause a pitch motion, four muscles per eye become involved.
- All six are involved to perform both a pitch and yaw, for example, looking upward and to the right.



Rotational degrees of freedom of mantis shrimp eyes relative to the external environment. Yellow arrows – pitch (up–down); green arrows – yaw (side-to-side); red arrows – torsional (roll) rotations. Image

Eye movement

- A small amount of roll can be generated; however, our eyes are generally not designed for much roll motion.
- Imagine if you could turn your eyeballs upside-down inside of their sockets!
- Thus, it is reasonable in most cases to approximate eye rotations as a 2D set that includes only yaw and pitch, rather than the full 3 DOFs obtained for rigid body rotations.

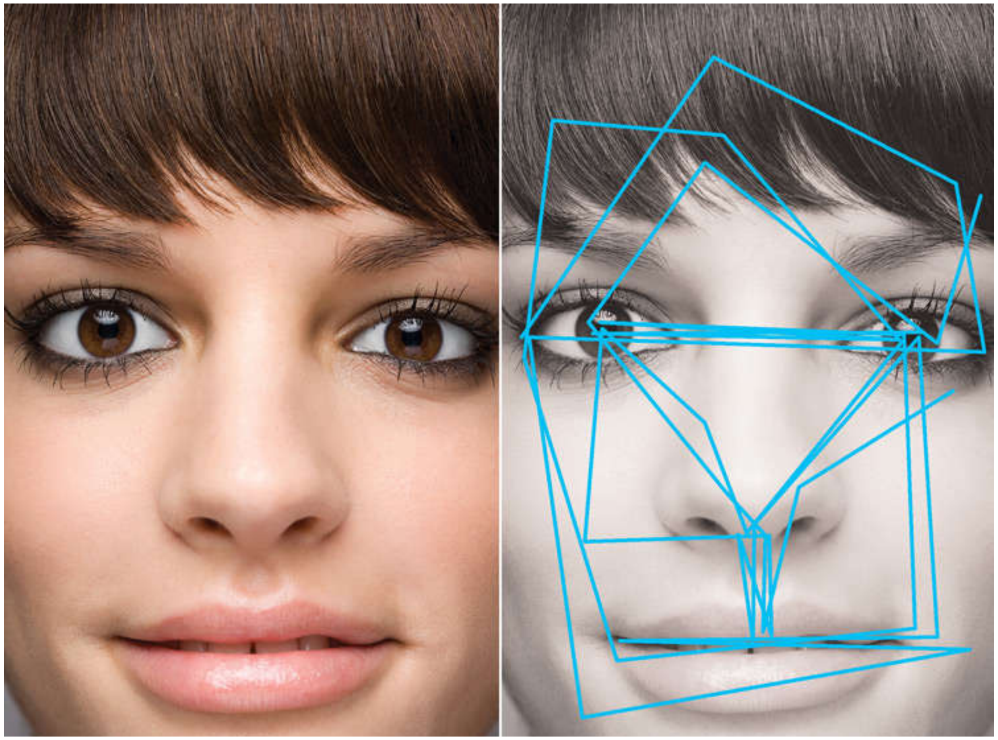


Types of movements

- Types of movements - based on their purpose:
  1) **saccades,**
  2) **smooth pursuit,**
  3) **vestibulo-ocular reflex,**
  4) **optokinetic reflex,**
  5) **vergence,**
  6) **micro saccades.**
- All of these motions cause both eyes to rotate approximately the same way, except for vergence, which causes the eyes to rotate in opposite directions.
- We will skip a seventh category of motion, called rapid eye movements (REMs), because they only occur while we are sleeping and therefore do not contribute to a VR experience.

1. Saccades

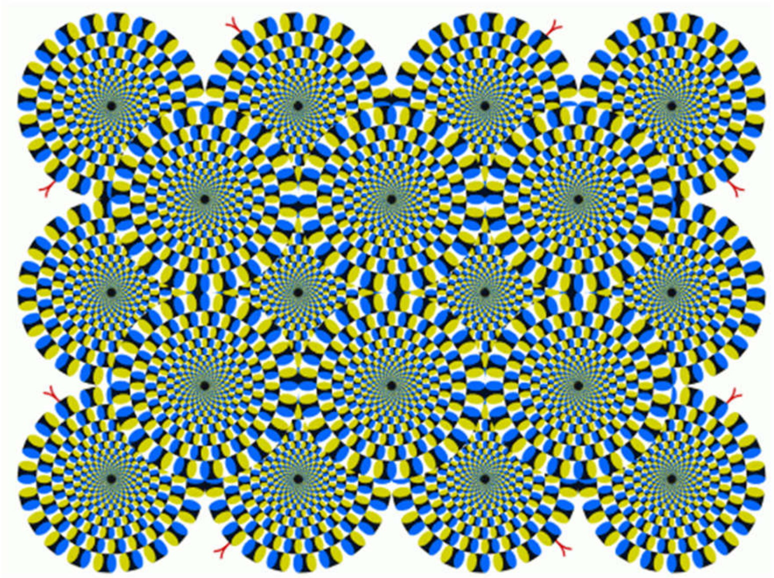
- The eye can move in a rapid motion called a saccade, which lasts less than 45ms with rotations of about 900◦ per second.
- The purpose is to quickly relocate the fovea so that important features in a scene are sensed with highest visual acuity.



The trace of scanning a face using saccades

1. Saccades

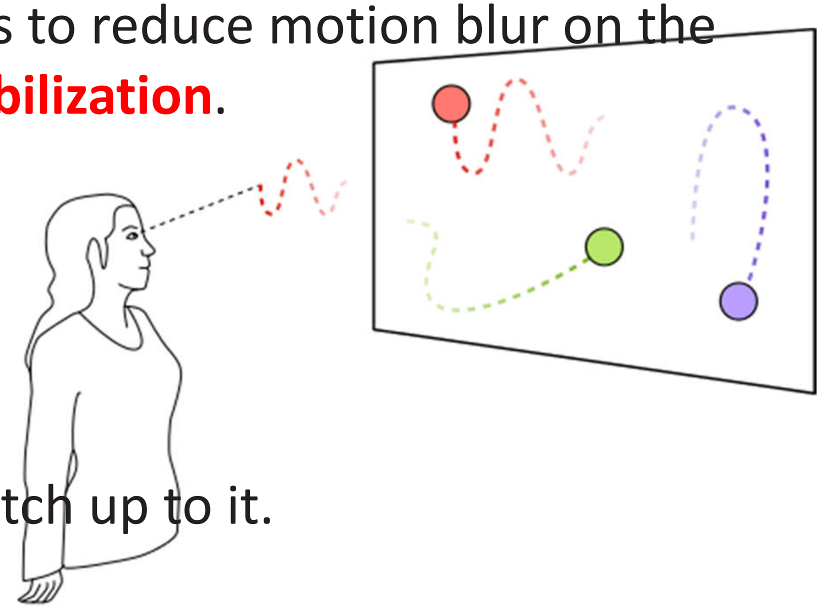
- Interestingly, our brains use ***saccadic masking*** to hide the intervals of time over which saccades occur from our memory.
- This results in distorted time perception, as in the case when second hands click into position on an analog clock.
- The result of saccades is that we obtain the illusion of high acuity over a large angular range.



- Although saccades frequently occur while we have little or no awareness of them, we have the ability to consciously control them as we choose features for fixation.

2. Smooth pursuit

- In the case of **smooth pursuit**, the eye slowly rotates to track a moving target feature.
- Examples are a car, a tennis ball, or a person walking by.
- The rate of rotation is usually less than 30◦ per second, which is much slower than for saccades.
- The main function of smooth pursuit is to reduce motion blur on the retina; this is also known as **image stabilization**.



- If the target is moving too fast, then saccades may be intermittently inserted into the pursuit motions to catch up to it.

3. Vestibulo-ocular reflex

- One of the most important motions to understand for VR is the **vestibulo-ocular reflex or VOR**.
- Hold your finger at a comfortable distance in front of your face and fixate on it.
- Next, yaw your head back and forth (like you are nodding “no”), turning about 20 or 30 degrees to the left and right sides each time.
- You may notice that your eyes are effortlessly rotating to counteract the rotation of your head so that your finger remains in view.
- The eye motion is involuntary.
- If you do not believe it, then try to avoid rotating your eyes while paying attention to your finger and rotating your head.
- It is called a **reflex** because the motion control bypasses higher brain functions.

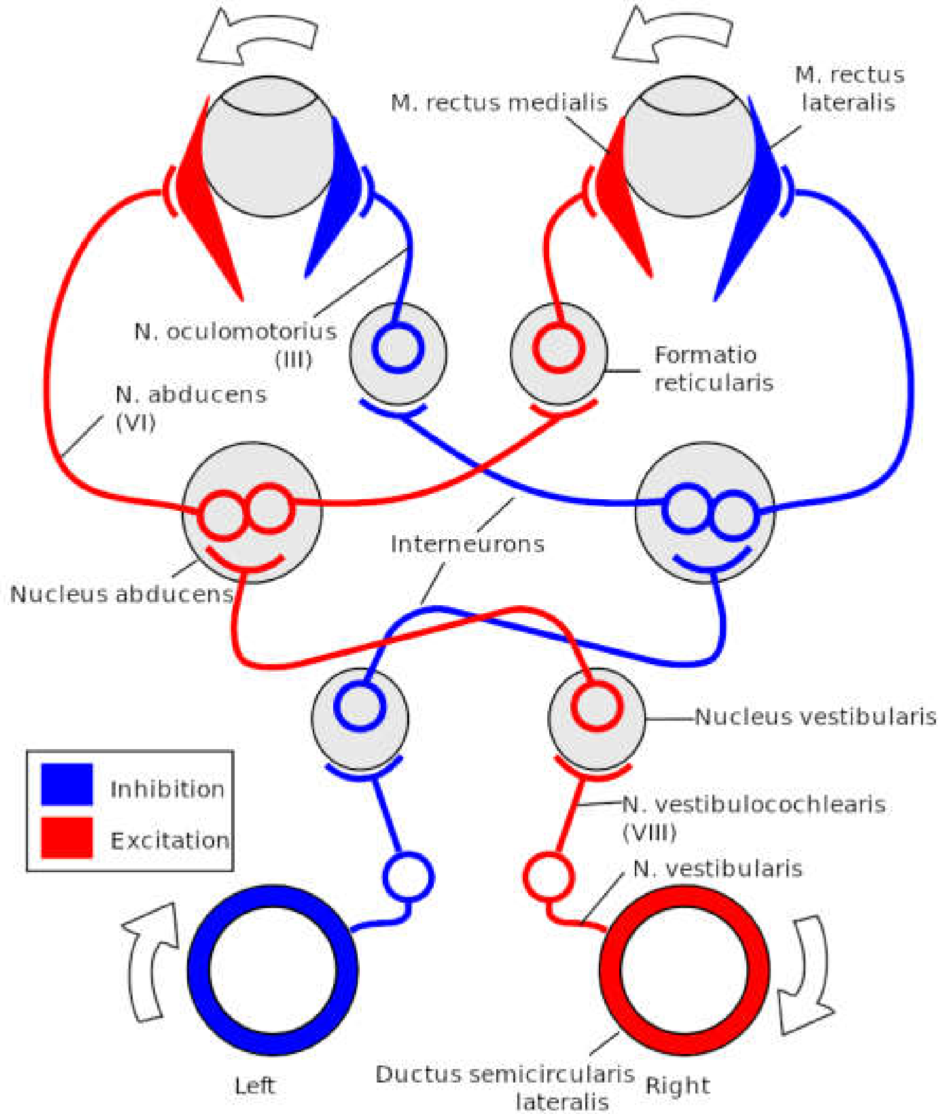
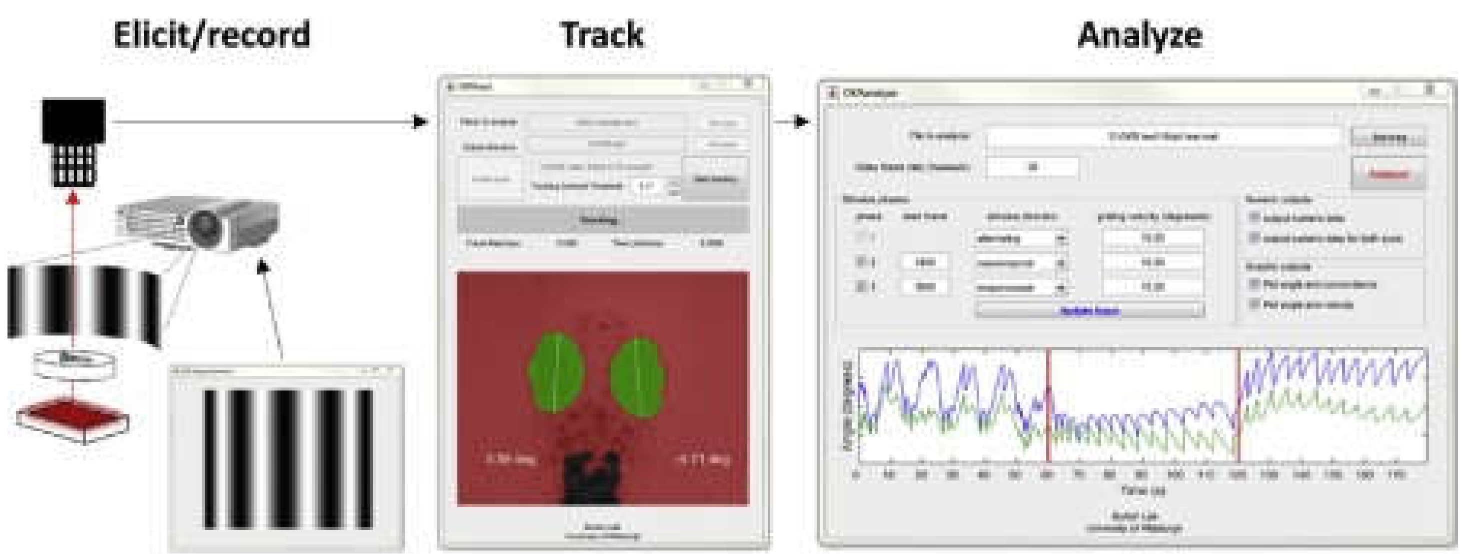


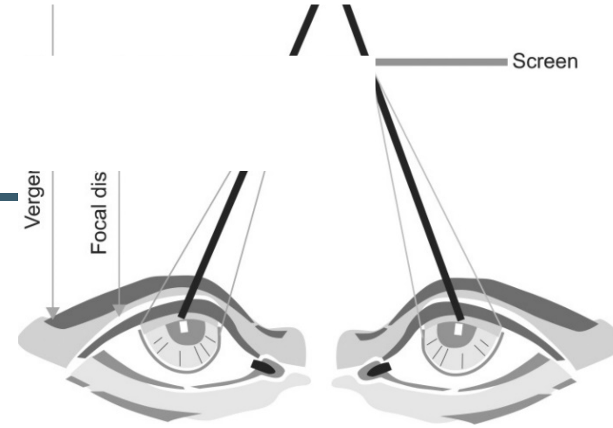
Figure 5.19: The vestibulo-ocular reflex (VOR). The eye muscles are wired to angular accelerometers in the vestibular organ to counter head movement with the opposite eye movement with less than 10ms of latency. The connection between the eyes and the vestibular organ is provided by specialized vestibular and extraocular motor nuclei, thereby bypassing higher brain functions.

4. Optokinetic reflex

- The optokinetic reflex occurs when a fast object speeds along.
- This occurs when watching a fast-moving train while standing nearby on fixed ground.
- The eyes rapidly and involuntarily choose features for tracking on the object, while alternating between smooth pursuit and saccade motions.



5. Vergence



- Stereopsis refers to the case in which both eyes are fixated on the same object, resulting in a single perceived image.
- Two kinds of vergence motions occur to align the eyes with an object.
- If the object is closer than a previous fixation, then a **convergence** motion occurs.
- This means that the eyes are rotating so that the pupils are becoming closer.
- If the object is further, then **divergence** motion occurs, which causes the pupils to move further apart.
- The eye orientations resulting from vergence motions provide important information about the distance of objects.

5. Vergence

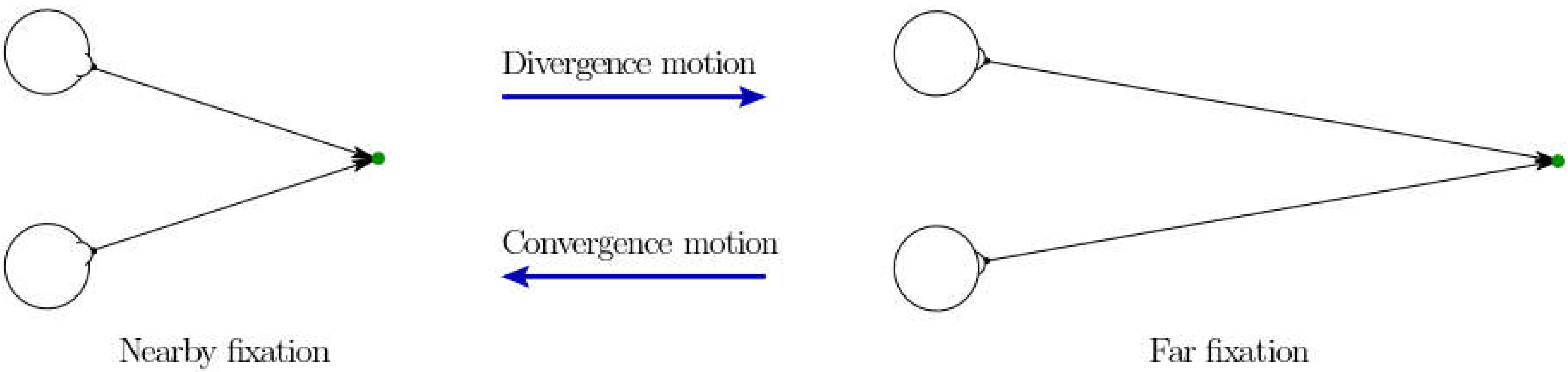
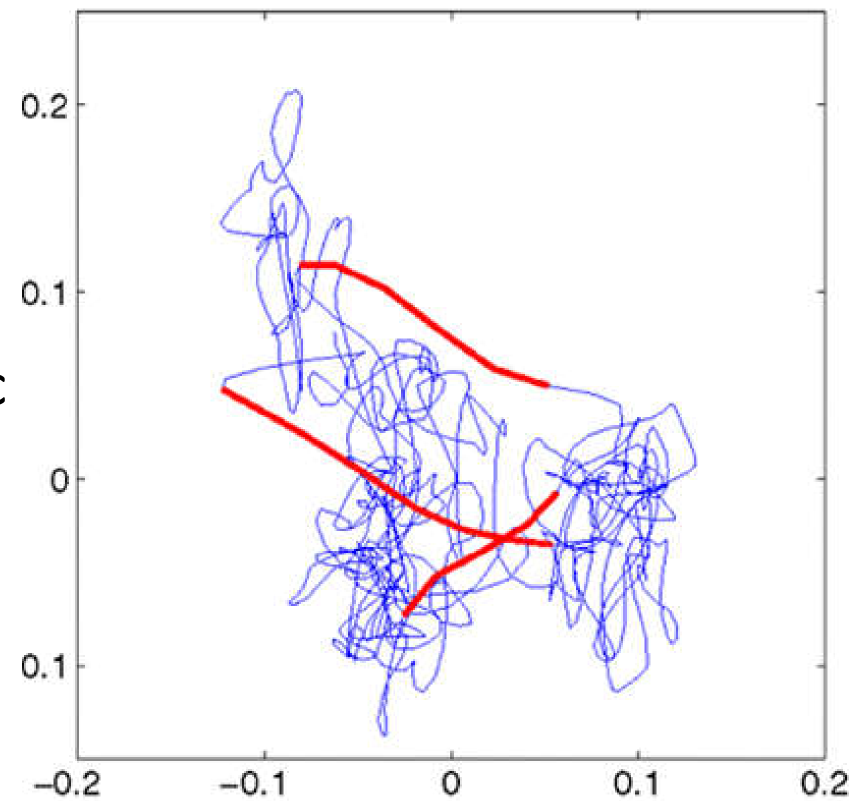


Figure 5.20: In the process of stereopsis, both eyes are fixated on the same feature in the world. To transition from a close to far feature, a divergence motion occurs. A convergence motion happens for the opposite transition.

6. Microsaccades

- Microsaccades are small, involuntary jerks of less than one degree that trace out an erratic path.
- Microsaccades are an active topic of research in perceptual psychology, biology, and neuroscience.

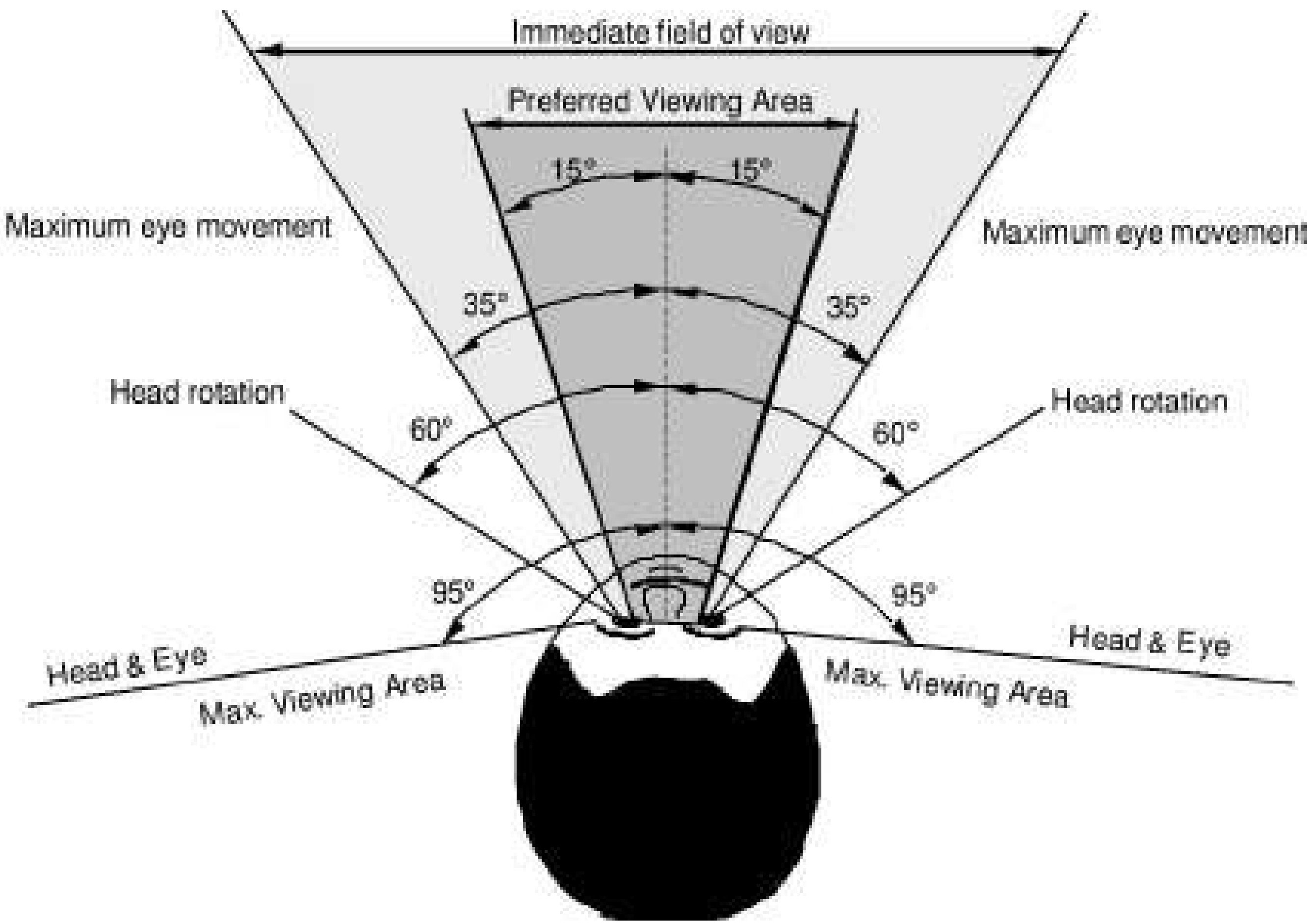


Slow movements (blue) are highly erratic, whereas microsaccades (red) are ballistic, small-amplitude epochs with a more linear trajectory.

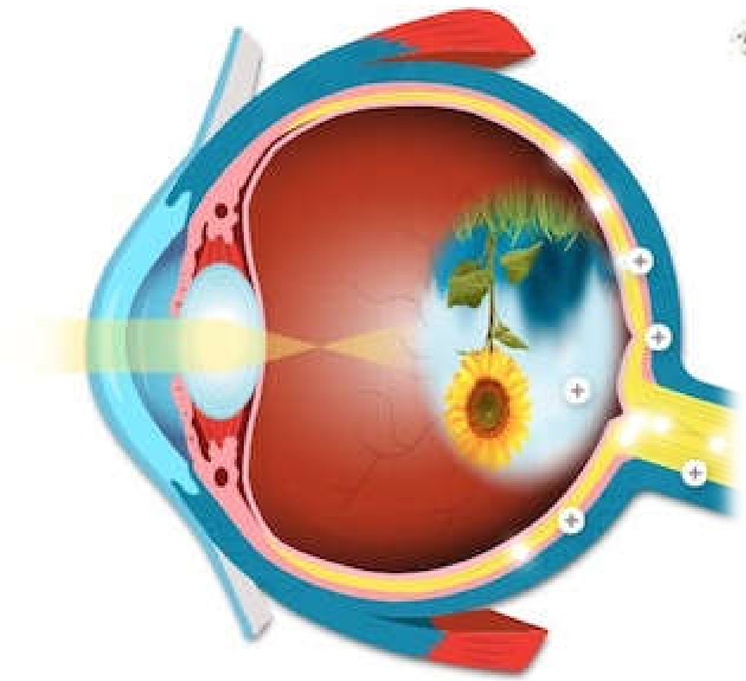
Eye and head movements together

- Most of the time the eyes and head are moving together.
- Although eye yaw is symmetric by allowing 35◦ to the left or right, pitching of the eyes is not.
- Human eyes can pitch 20◦ upward and 25◦ downward, which suggests that it might be optimal to center a VR display slightly below the pupils when the eyes are looking directly forward.
- In the case of VOR (Vestibulo-ocular reflex), eye rotation is controlled to counteract head motion.
- In the case of smooth pursuit, the head and eyes may move together to keep a moving target in the preferred viewing area.

Eye and head movements together



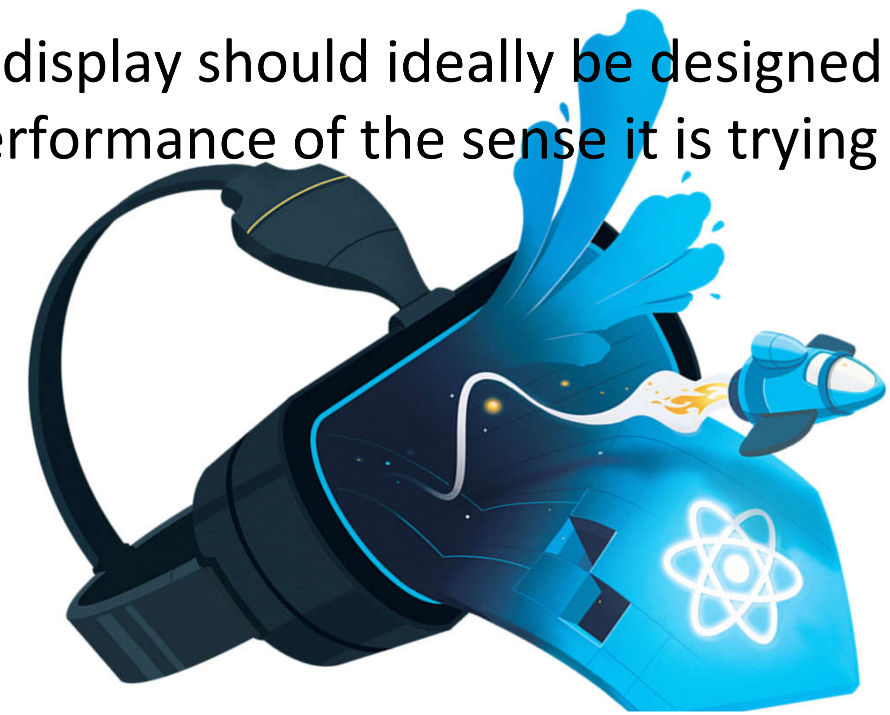
Implication for VR



Aleksandra Klašnja-Milićević
Faculty of Scinces, Novi Sad, Serbia
Department of Mathematics and Informatics

Implications for VR

- Basic physiological properties, such as photoreceptor density or VOR circuitry directly impact the engineering requirements for visual display hardware.
- The engineered systems must be good enough to adequately fool our senses, but they need not have levels of quality that are well beyond the limits of our receptors.
- Thus, the VR display should ideally be designed to perfectly match the performance of the sense it is trying to fool.

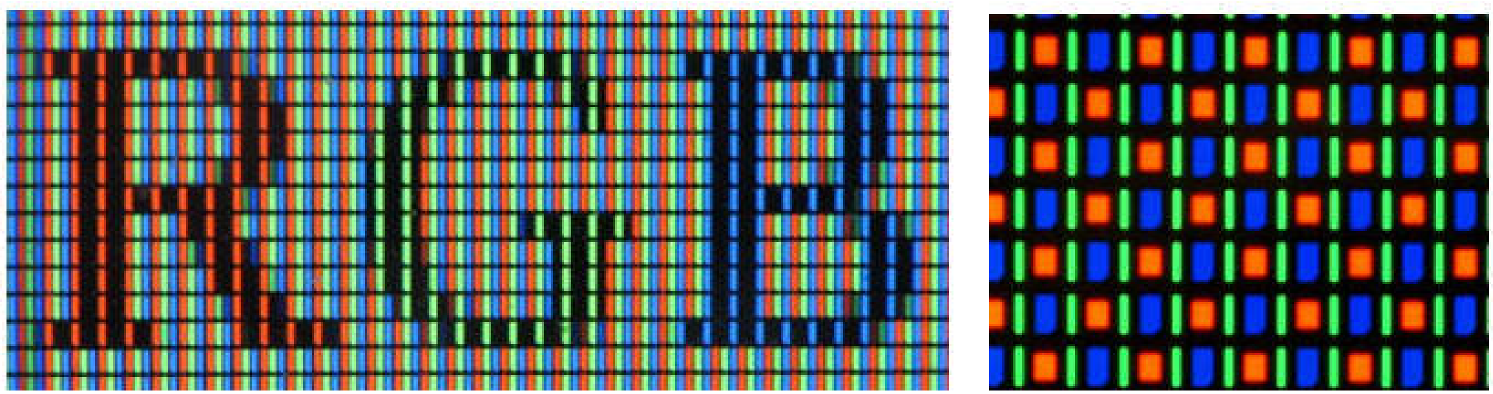


How good does the VR visual display need to be?

- Three crucial factors for the display are:
  - **Spatial resolution:** How many pixels per square area are needed?
  - **Intensity resolution and range:** How many intensity values can be produced, and what are the minimum and maximum intensity values?
  - **Temporal resolution:** How fast do displays need to change their pixels?

How good does the VR visual display need to be?

- How much pixel density is enough?
- Insights into the required spatial resolution are obtained from the photoreceptor densities. We see individual lights when a display is highly magnified.



(a) (b)

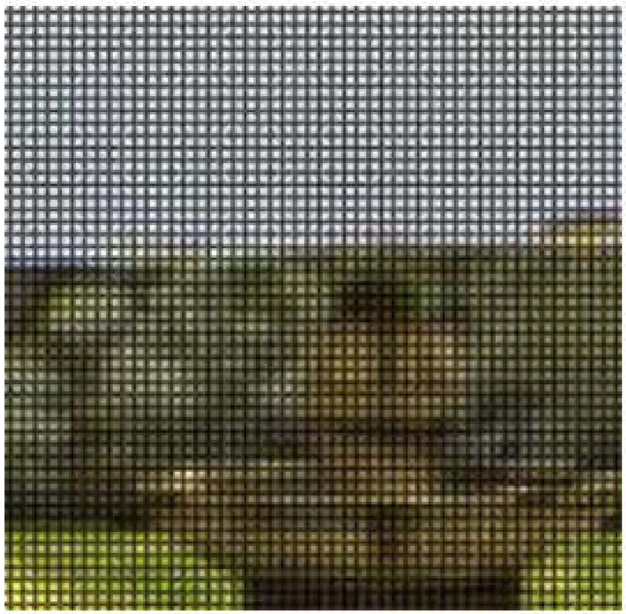
Figure 5.22: In displays, the pixels break into subpixels, much in the same way that photoreceptors break into red, blue, and green components. (a) An LCD display. (Photo by Luis Flavio Loureiro dos Santos.) (b) An AMOLED PenTile display from the Nexus One smartphone. (Photo by Matthew Rollings.)

How good does the VR visual display need to be?

- As it is zoomed out, we may still perceive sharp diagonal lines as being jagged, this phenomenon is known as **aliasing**.

How good does the VR visual display need to be?

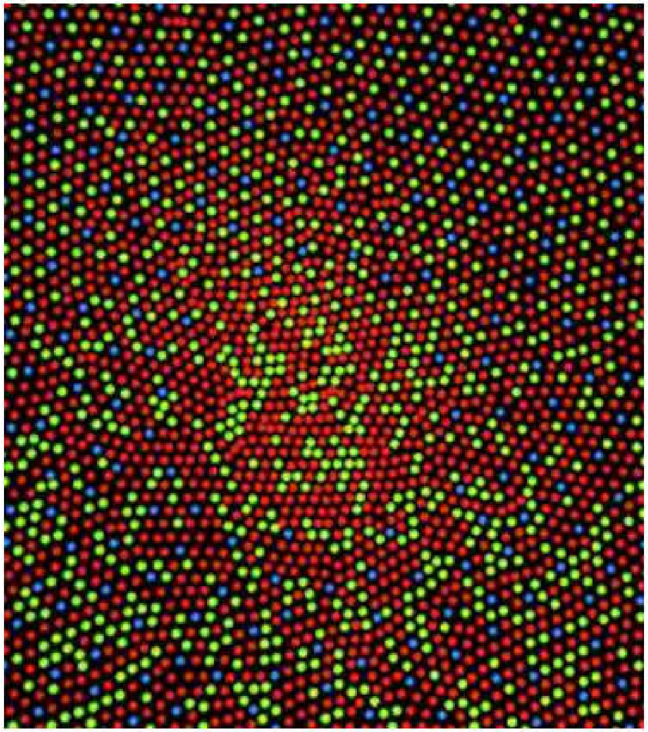
- Another artifact is **the screen-door effect**, this is commonly noticed in an image produced by a digital LCD projector.

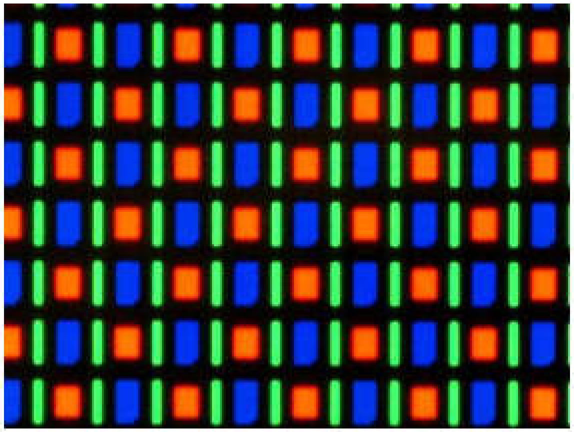


- What does the display pixel density need to be so that we do not perceive individual pixels?
- In 2010, Steve Jobs of Apple Inc. claimed that **326** pixels per linear inch (PPI) is enough, achieving what they called a retina display.

How good does the VR visual display need to be?

- Assume that the fovea is pointed directly at the display to provide the best sensing possible.
- The first issue is that red, green, and blue cones are arranged in a mosaic.





How good does the VR visual display need to be?

- Vision scientists and neurobiologists have studied the effective or perceived input resolution through measures of visual acuity.
- Subjects in a study are usually asked to indicate whether they can detect or recognize a particular target.
- In the case of detection,

  for example, scientists might

  like to know the smallest

  dot that can be perceived

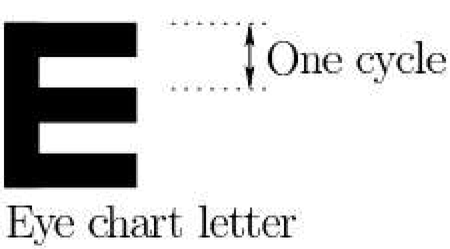
  when printed onto a surface.

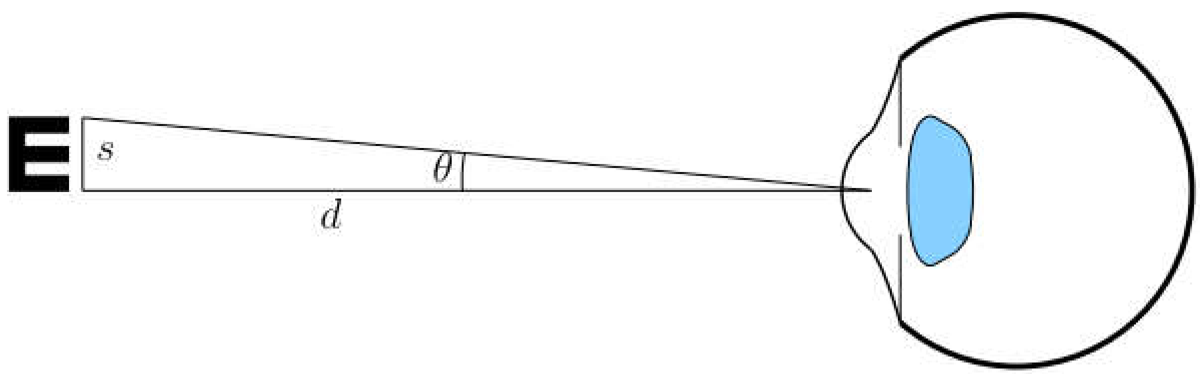
How good does the VR visual display need to be?

- In terms of displays, a similar question is: How small do pixels need to be so that a single white pixel against a black background is not detectable?
- In the case of recognition, a familiar example is attempting to read an eye chart, which displays arbitrary letters of various sizes.
- In terms of displays, this could correspond to trying to read text under various sizes, resolutions, and fonts.
- Many factors contribute to acuity tasks, such as brightness, contrast, eye movements, time exposure, and the part of the retina that is stimulated.

How good does the VR visual display need to be?

- The Snellen eye chart, is designed so that patients attempt to recognize printed letters from 20 feet away (or 6 meters).
- A person with “normal” 20/20 (or 6/6 in metric) vision is expected to barely make out the horizontal stripes in the letter “E”.
- This assumes he is looking directly at the letters, using the photoreceptors in the central fovea.
- The 20/20 line on the chart is designed so that letter height corresponds to 30 cycles per degree when the eye is 20 feet away.

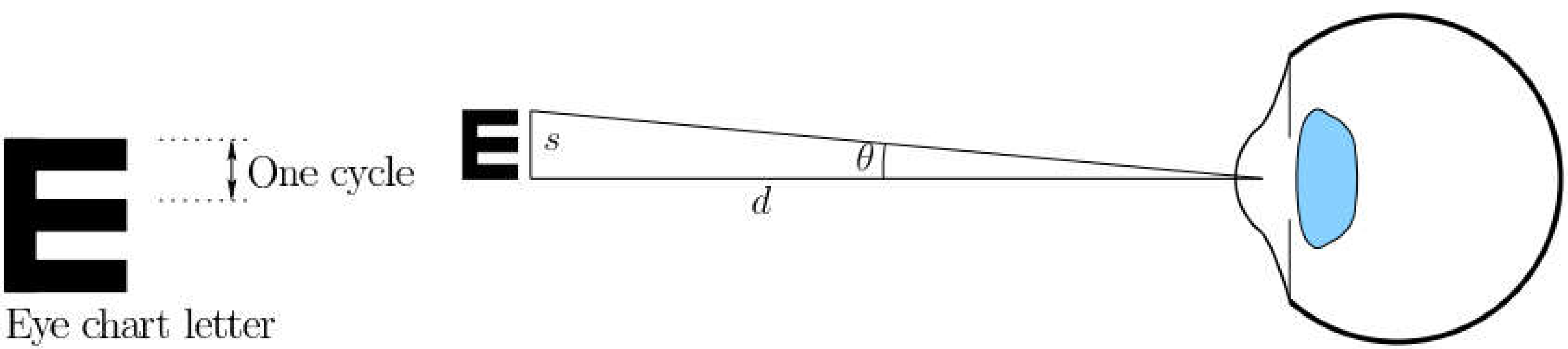




How good does the VR visual display need to be?

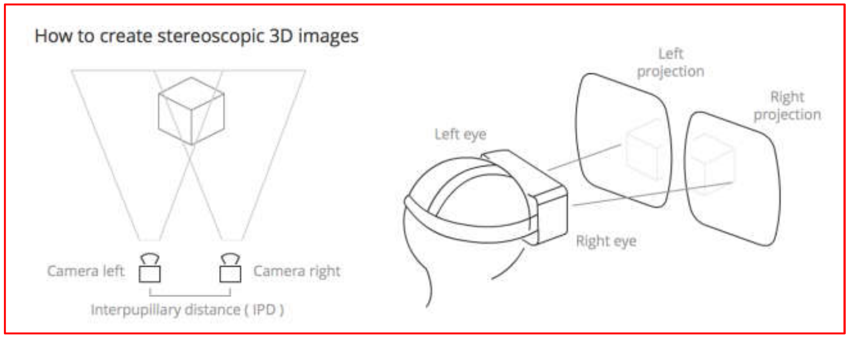
- The total height of the "E" is 1/6 of a degree.
- Each stripe is half of a cycle. What happens if the subject stands only 10 feet away from the eye chart?
- The letters should roughly appear to twice as large.
- Using simple trigonometry,

$$s = d \tan \theta$$



How good does the VR visual display need to be?

- Suppose that a smartphone screen is placed 12 inches from the user’s eye. In this case, $s = 12 * \tan 1^\circ = 0.209$in.
- This requires that the screen have at least $60/0.209 = 286.4$ PPI, which was satisfied by the 326 PPI originally claimed by Apple.
- In the case of VR, the user is not looking directly at the screen as in the case of smartphones.
- By inserting a lens for magnification, the display can be brought even closer to the eye.



How good does the VR visual display need to be?

- This is commonly done for VR headsets.

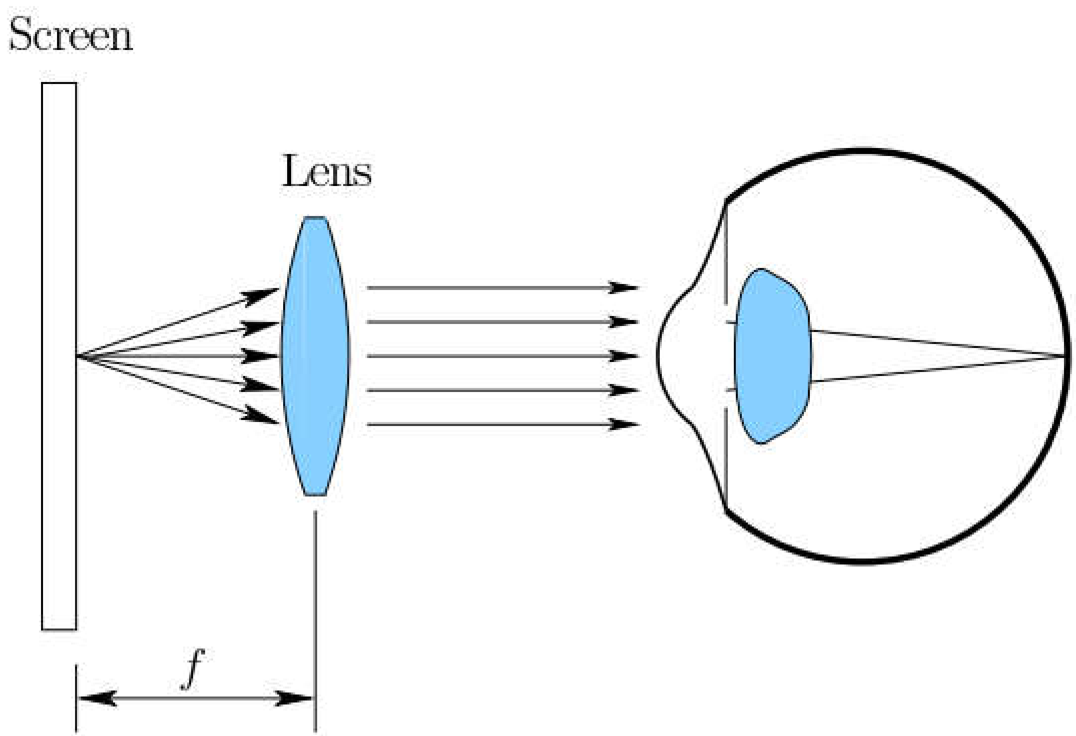


Figure 4.30: In VR headsets, the lens is placed so that the screen appears to be infinitely far away.

How good does the VR visual display need to be?

- Suppose that the lens is positioned at its focal distance away from the screen, which for the sake of example is only 1.5in (this is comparable to current VR headsets).
- In this case, $s = 1 * \tan 1^{\circ} = 0.0261$in, and the display must have at least 2291.6 PPI to achieve 60 cycles per degree!
- The highest-density smartphone display available today is the Super AMOLED 1440x2560 5.1 inch screen on the Samsung S6, S7, S8, S9, which is used in the Gear VR system.
- It has only 577 PPI, which means that the

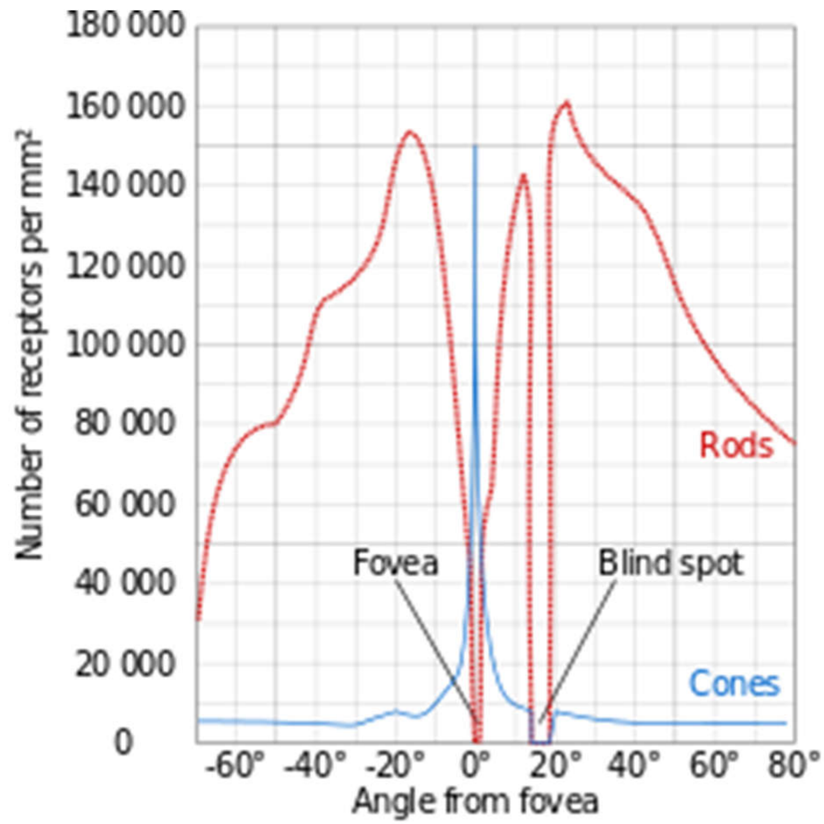
PPI needs to increase by roughly a factor

of four to obtain retina display resolution

for VR headsets.

How much field of view is enough?

- What if the screen is brought even closer to the eye to fill more of the field of view?
- Based on the photoreceptor density plot and the limits of eye rotations, the maximum field of view seems to be **around 270◦**, which is larger than what could be provided by a flat screen (less than 180◦).



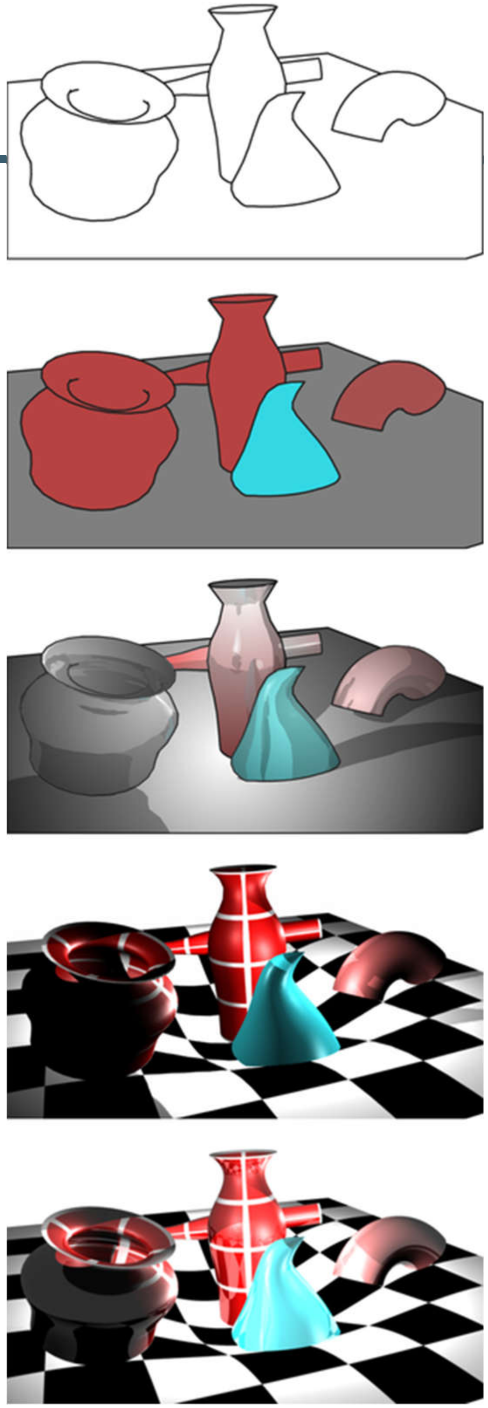
- Increasing the field of view by bringing the screen closer would require even higher pixel density, but lens aberrations at the periphery may limit the effective field of view.

How much field of view is enough?

- Furthermore, if the lens is too thick and too close to the eye, then the eyelashes may scrape it;
- Fresnel lenses may provide a thin alternative, but introduce artifacts.
- Thus, the quest for a VR retina display may end with a balance between optical system quality and limitations of the human eye.
- Curved screens may help alleviate some of the problems.

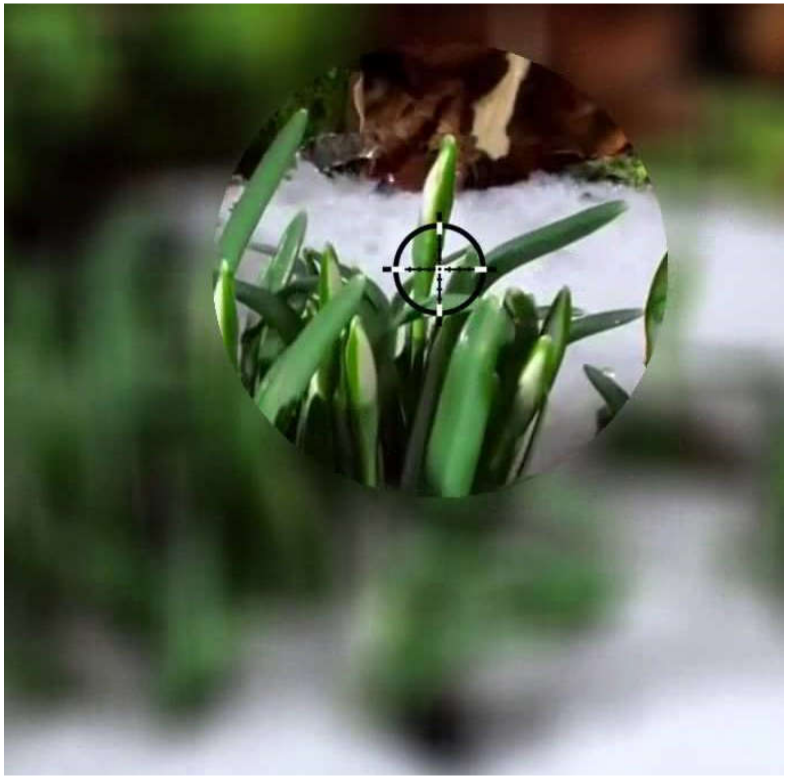
Foveated rendering

- We have not been able to exploit that fact that photoreceptor density decreases away from the fovea.
- We had to keep the pixel density high everywhere - we have no control over which part of the display the user will be look at.
- If we could track where the eye is looking and have a tiny, movable display that is always positioned in front of the pupil, with zero delay, then much fewer pixels would be needed.
- This would greatly decrease computational burdens on graphical rendering systems.



Foveated rendering

- Instead of moving a tiny screen, the process can be simulated by keeping the fixed display but focusing the graphical rendering only in the spot where the eye is looking.
- This is called **foveated rendering**, which has been shown to work, but is currently too costly and there is too much delay and other discrepancies between the eye movements and the display updates.



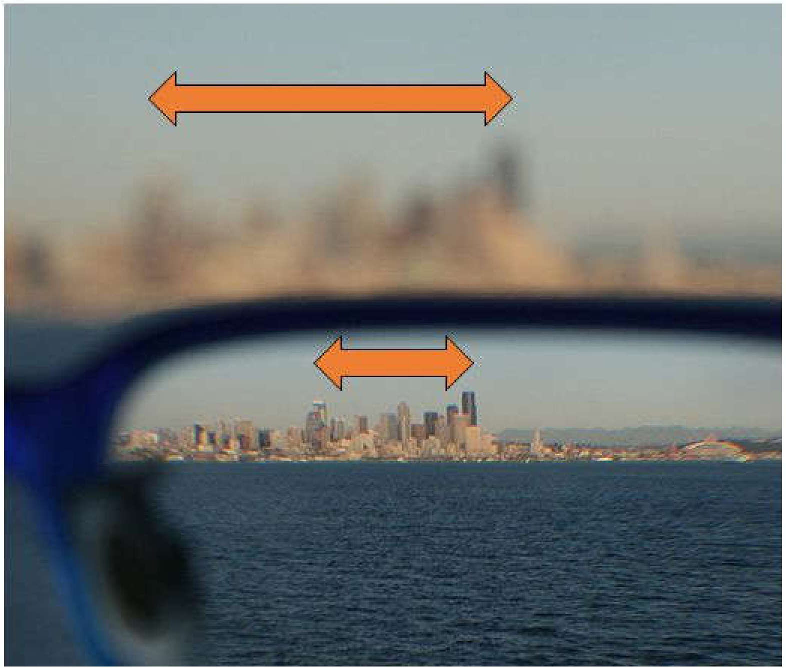
In the near future, it may become an effective approach for the mass market

VOR gain adaptation

- Vestibulo-Optical Reflex (VOR) is the adaptation we have which helps keep our eyes fixed on an object even while our head moves.
- It's a smooth motion in our eye muscles controlled by our ears sensitivity to rotation – it's involuntary, happens whether we are seeing anything or not (eyes closed or in the dark) and usually gives a 1:1 compensation between head rotation and eye motion.
- Because head motion has six degrees of freedom (DOFs), it is appropriate to break the gain into six components.

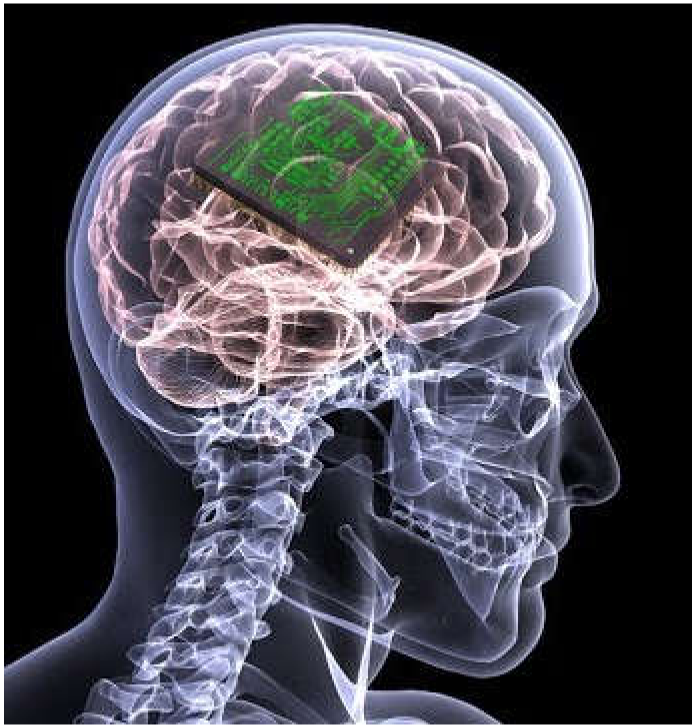
VOR gain adaptation

- Adaptation is a universal feature of our sensory systems.
- VOR gain is no exception.
- For those who wear eyeglasses, the VOR gain must adapt due to the optical transformations.
- Lenses affect the field of view and perceived size and distance of objects.
- The VOR comfortably adapts to this problem by changing the gain.



VOR gain adaptation

- Now suppose that you are wearing a VR headset that may suffer from flaws such as an imperfect optical system, tracking latency, and incorrectly rendered objects on the screen.
- In this case, adaptation may occur as the brain attempts to adapt its perception of stationarity to compensate for the flaws.
- In this case, your visual system could convince your brain that the headset is functioning correctly, and then your perception of stationarity in the real world would become distorted until you readapt.



VOR gain adaptation

- For example, after a flawed VR experience, you might yaw your head in the real world and have the sensation that truly stationary objects are sliding back and forth.

Display scanout

- Cameras have either a rolling or global shutter based on whether the sensing elements are scanned line-by-line or in parallel.
- Displays work the same way, but whereas cameras are an input device, displays are the output analog.
- Most displays today have a rolling scanout (called raster scan), rather than global scanout.
- This implies that the pixels are updated line by line.

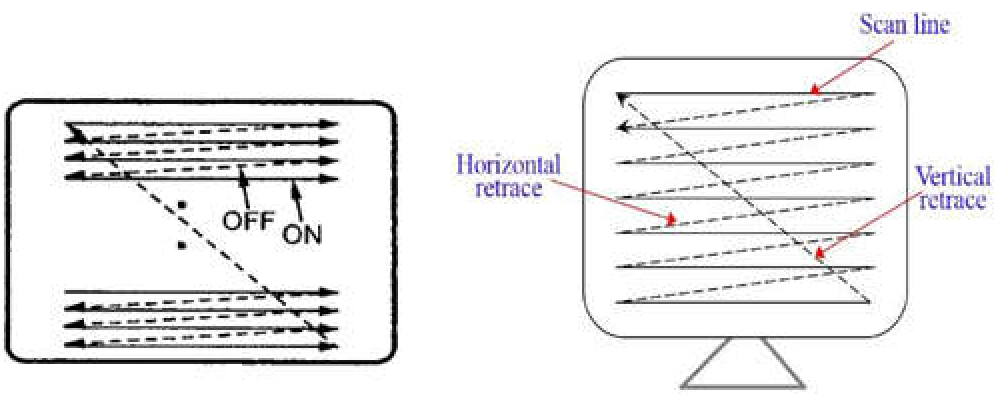


fig. Raster Scan CRT

Display scanout

- This procedure is an artifact of old TV sets and monitors, which each had a cathode ray tube (CRT) with phosphor elements on the screen.
- An electron beam was bent by electromagnets so that it would repeatedly strike and refresh the glowing phosphors.

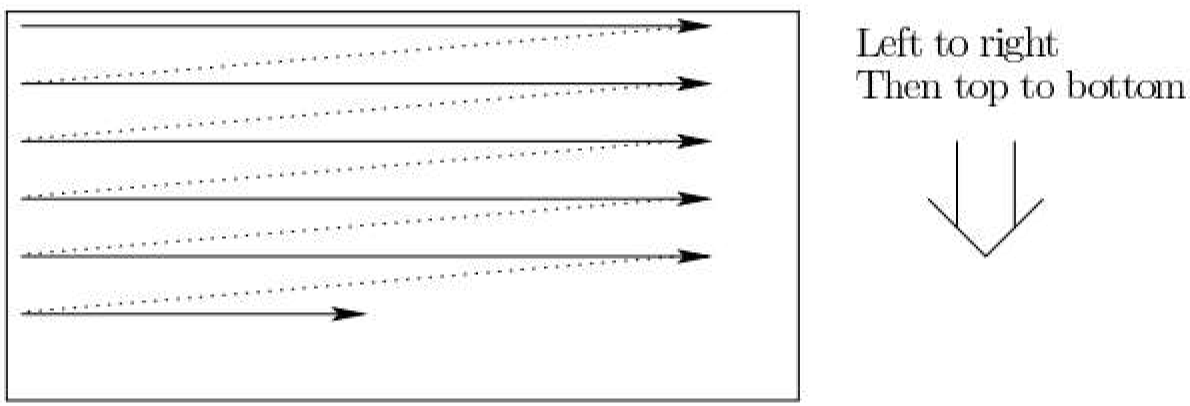
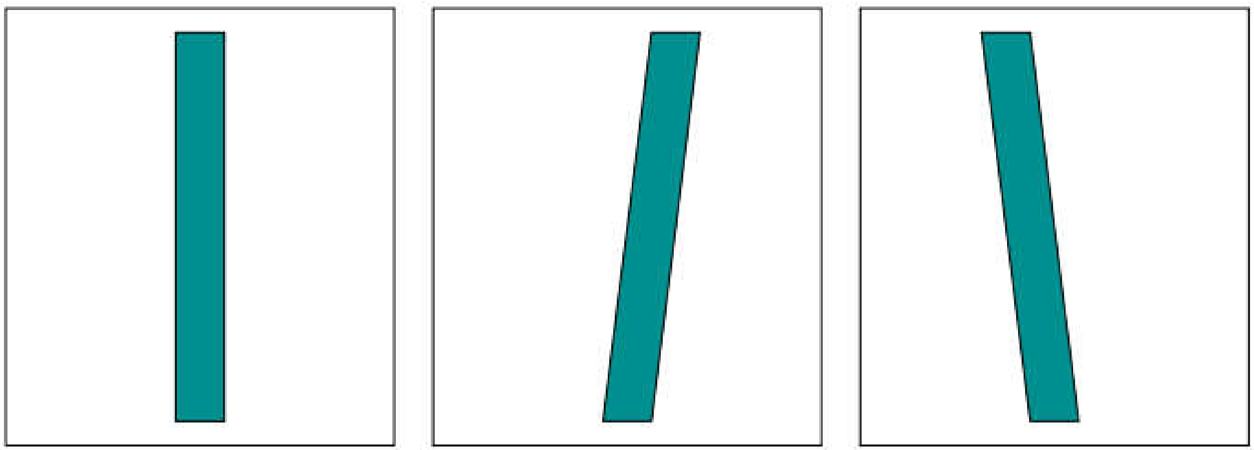


Figure 5.26: Most displays still work in the way as old TV sets and CRT monitors: By updating pixels line-by-line. For a display that has 60 FPS (frames per second), this could take up to 16.67ms.

Display scanout

- You can also achieve this effect by repeatedly drawing a horizontal line with a pencil while using the other hand to gently pull the paper in a particular direction.
- The paper in this analogy is the retina and the pencil corresponds to light rays attempting to charge photoreceptors.

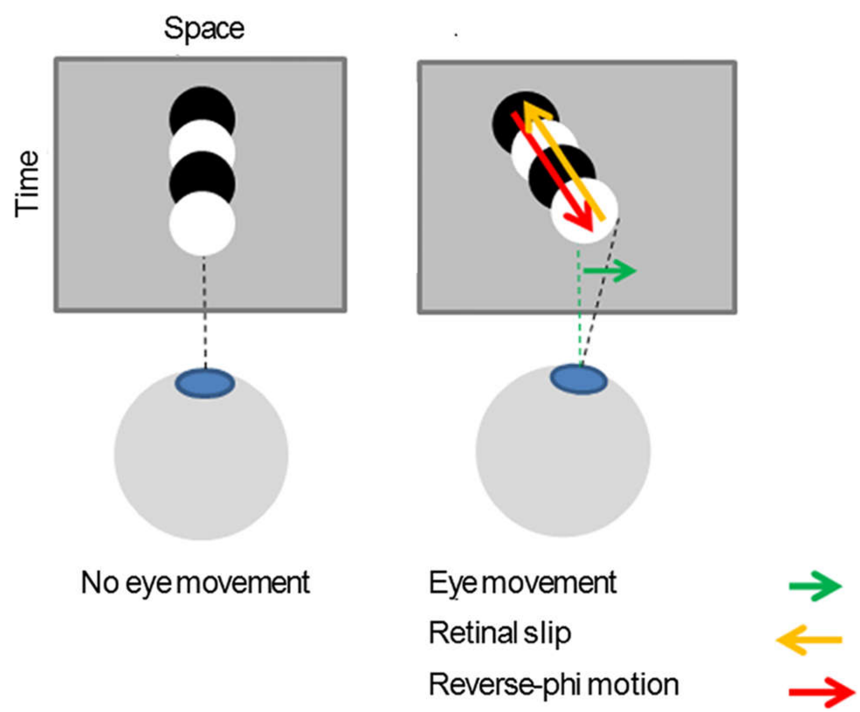


(a) (b) (c)

Figure 5.27: Artifacts due to display scanout: (a) A vertical rectangle in the scene. (b) How it may distort during smooth pursuit while the rectangle moves to the right in the virtual world. (c) How a stationary rectangle may distort when rotating the head to the right while using the VOR to compensate. The cases of (b) are (c) are swapped if the direction of motion is reversed in each case.

Retinal image slip

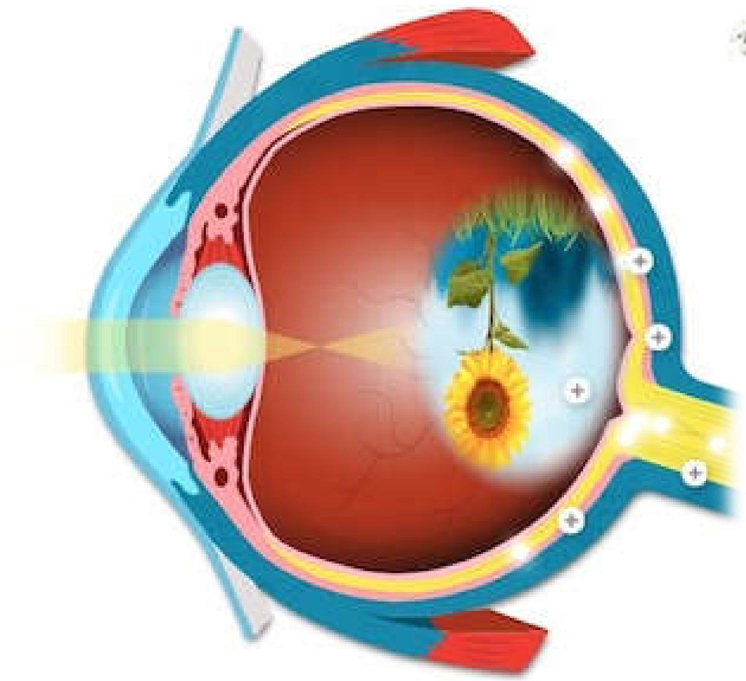
- Eye movements contribute both to maintaining a target in a fixed location on the retina (smooth pursuit, VOR) and also to changing its location slightly to reduce perceptual fading (microsaccades).
- During ordinary activities (not VR), the eyes move and the image of a feature may move slightly on the retina due to motions and optical distortions. This is called **retinal image slip.**



Retinal image slip

- Once a VR headset is used, the motions of image features on the retina might not match what would happen in the real world.
- This is due to many factors such as: optical distortions, tracking latency, and display scanout.
- Thus, the retinal image slip due to VR artifacts does not match the retinal image slip encountered in the real world.
- The consequences of this have barely been identified, much less characterized scientifically.
- They are likely to contribute to fatigue, and possibly VR sickness.

Modern VR experience



Aleksandra Klašnja-Milićević
Faculty of Scinces, Novi Sad, Serbia
Department of Mathematics and Informatics

Modern VR Experiences

- The modern era of VR was brought about by advances in display, sensing, and computing technology from the smartphone industry.
- From Palmer Luckey’s 2012 Oculus Rift design to simply building a case for smart phones, the world has quickly changed as VR headsets are mass produced and placed into the hands of a wide variety of people.
- As more people have access to the technology, the set of things they do with it substantially broadens.

Modern VR Experiences - Video games

- People have dreamed of entering their video game worlds for decades.
- By 1982, this concept was already popularized by the Disney movie Tron.

(a)

(b)

(c)

(d)

Figure 1.5: (a) Valve's Portal 2 demo for the HTC Vive headset is a puzzle-solving experience in a virtual world. (b) The Virtuix Omni treadmill for walking through first-person shooter games. (c) Lucky's Tale for the Oculus Rift maintains a third-person perspective as the player floats above his character. (d) In the Dumpy game from DePaul University, the player appears to have a large elephant trunk. The purpose of the game is to enjoy this unusual embodiment by knocking things down with a swinging trunk.

Modern VR Experiences - Immersive cinema

- Hollywood movies continue to offer increasing degrees of realism.
- Why not make the viewers feel like they are part of the scene?
- In VR, viewers can look in any direction, and perhaps even walk through the scene.
- What should they be allowed to do?
- How do you make sure they do not miss part of the story?
- Should the story be linear, or should it adapt to the viewer's actions?

Modern VR Experiences - Immersive cinema

- Should the viewer be a first-person character in the film, or a third-person observer who is invisible to the other characters?
- How can a group of friends experience a VR film together?
- When are animations more appropriate versus the capture of real scenes?
- It will take many years to resolve these questions and countless more that will arise. In the meantime, VR can also be used as a kind of "wrapper" around existing movies.

Modern VR Experiences – Immersive cinema

Figure 1.6: Oculus Story Studio produced Emmy-winning *Henry*, an immersive short story about an unloved hedgehog who hopes to make a new friend, the viewer.

Modern VR Experiences – Telepresence

- The first step toward feeling like we are somewhere else is capturing a panoramic view of the remote environment.
- Google’s Street View and Earth apps already rely on the captured panoramic images from millions of locations around the world.

(a) (b)

Figure 1.8: An important component for achieving telepresence is to capture a panoramic view: (a) A car with cameras and depth sensors on top, used by Google to make Street View. (b) Bublcam is a cheap, portable way to capture and stream omnidirectional videos.

Modern VR Experiences – Telepresence

- Simple VR apps that query the Street View server directly enable to user to feel like he is standing in each of these locations, while easily being able to transition between nearby locations.
- Panoramic video capture is even more compelling.

Modern VR Experiences – Telepresence

- Simple VR apps that query the Street View server directly enable to user to feel like he is standing in each of these locations, while easily being able to transition between nearby locations.
- Panoramic video capture is even more compelling.

(a) (b)

Figure 1.9: A simple VR experience that presents Google Street View images through a VR headset: (a) A familiar scene in Paris. (b) Left and right eye views are created inside the headset, while also taking into account the user's looking direction.

Modern VR Experiences – Telepresence

- Live panoramic video interfaces, through which people can attend sporting events and concerts.
- Through a live interface, interaction is possible. People can take video conferencing to the next level by feeling present at the remote location.
- By connecting panoramic cameras to robots, the user is even allowed to move around in the remote environment (Figure 1.11).

(a) (b)

Figure 1.11: Examples of robotic avatars: (a) The DORA robot from the University of Pennsylvania mimics the users head motions, allowing him to look around in a remote world while maintaining a stereo view (panoramas are monoscopic). (b) The Plexidrone, a low-cost flying robot that is designed for streaming panoramic video.

Modern VR Experiences – Telepresence

- Current VR technology allows us to virtually visit far away places and interact in most of the ways that were previously possible only while physically present.
- This could ultimately help reverse the urbanization trend sparked by the 19th-century industrial revolution, leading to deurbanization as we distribute more uniformly around the Earth.

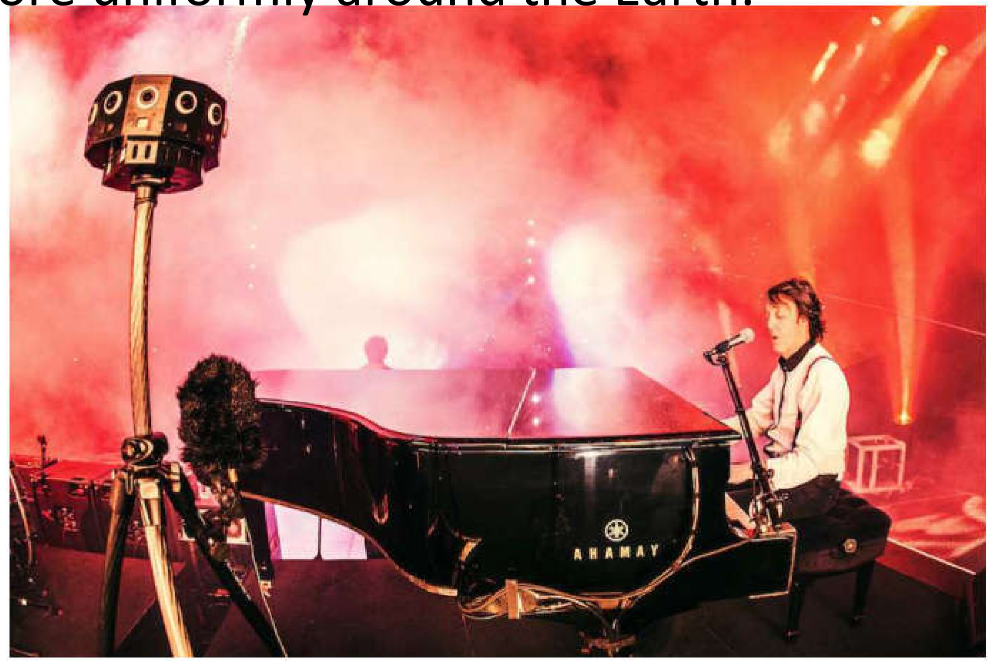


Figure 1.10: Jaunt captured a panoramic video of Paul McCartney performing Live and Let Die, which provides a VR experience where users felt like they were on stage with the rock star.

Modern VR Experiences – Virtual societies

- VR allows us to form entire societies that remind us of the physical world, but are synthetic worlds that contain avatars connected to real people.
- Groups of people could spend time together in these spaces for a variety of reasons, including common special interests, educational goals, or simply an escape from ordinary life.

Figure 1.12: Virtual societies develop through interacting avatars that meet in virtual worlds that are maintained on a common server. A snapshot from Second Life is shown here.

Modern VR Experiences – Empathy

- The first-person perspective provided by VR is a powerful tool for causing people to feel empathy for someone else's situation.
- The world continues to struggle with acceptance and equality for others of different race, religion, age, gender, sexuality, social status, and education, while the greatest barrier to progress is that most people cannot fathom what it is like to have a different identity.

Figure 1.13: In Clouds Over Sidra, film producer Chris Milk offers a first-person perspective on the suffering of Syrian refugees.

Modern VR Experiences – Empathy

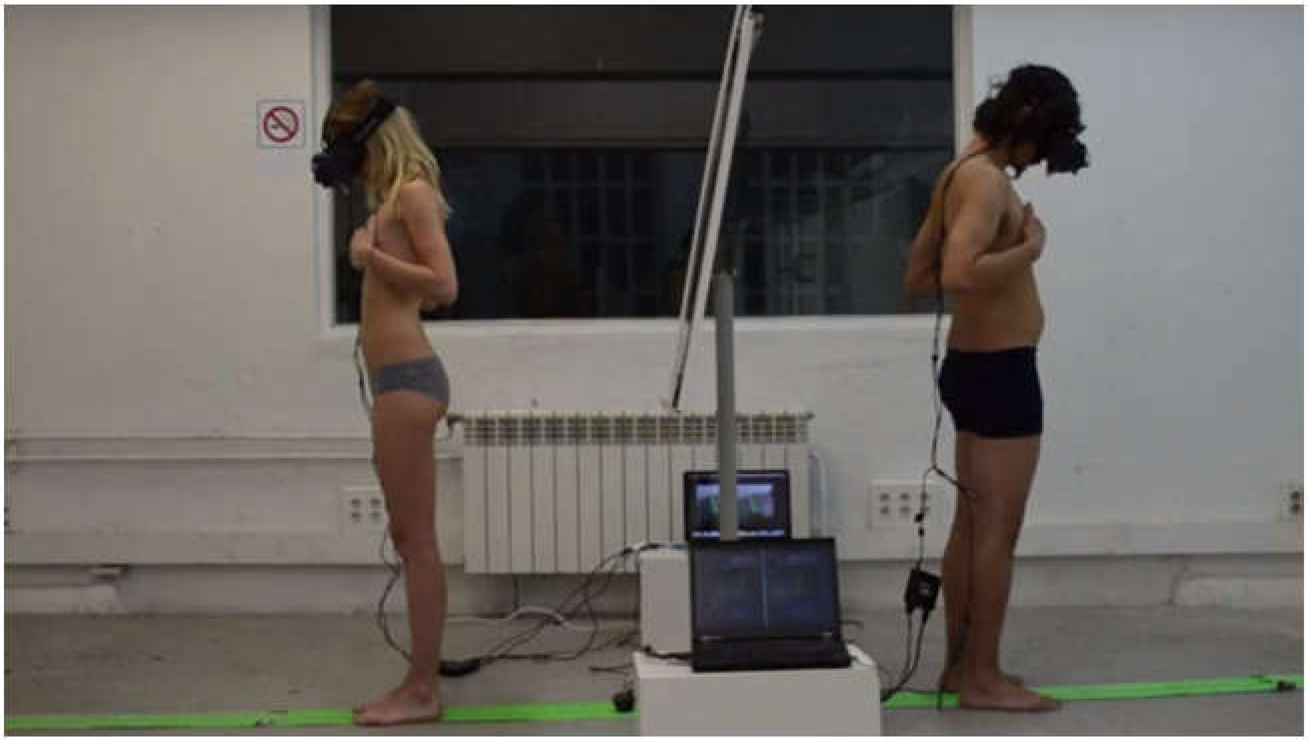


Figure 1.14: Students in Barcelona made an experience where you can swap bodies with the other gender. Each person wears a VR headset that has cameras mounted on its front. Each therefore sees the world from the approximate viewpoint of the other person. They were asked to move their hands in coordinated motions so that they see their new body moving appropriately.

Modern VR Experiences – Empathy

- Through virtual societies, many more possibilities can be explored.
- What if you were 10cm shorter than everyone else?
- What if you teach your course with a different gender?
- What if you were the victim of racial discrimination by the police?
- Using VR, we can imagine many “games of life” where you might not get as far without being in the “proper” group.

Modern VR Experiences – Education

- The first-person perspective could revolutionize many areas of education.
- In engineering, mathematics, and the sciences, VR offers the chance to visualize geometric relationships in difficult concepts or data that are hard to interpret.
- Furthermore, VR is naturally suited for practical training because skills developed in a realistic virtual environment may transfer naturally to the real environment.
- The motivation is particularly high if the real environment is costly to provide or poses health risks.

Modern VR Experiences – Education

- One of the earliest and most common examples of training in VR is flight simulation.
- Other examples include firefighting, nuclear power plant safety, search-and-rescue, military operations, and medical procedures

Figure 1.15: A flight simulator used by the US Air Force (photo by Javier Garcia). The user sits in a physical cockpit while being surrounded by displays that show the environment.

Modern VR Experiences – Education

- Perhaps the greatest opportunities for VR education lie in the humanities, including history, anthropology, and foreign language acquisition.
- Consider the difference between reading a book on the Victorian era in England and being able to roam the streets of 19th-century London, in a simulation that has been painstakingly constructed by historians.
- We could even visit an ancient city that has been reconstructed from ruins.
- Fascinating possibilities exist for either touring physical museums through a VR interface or scanning and exhibiting artifacts directly in virtual museums.
- These examples fall under the heading of digital heritage

Modern VR Experiences – Education

Figure 1.16: A tour of the Nimrud palace of Assyrian King Ashurnasirpal II, a VR experience developed by Learning Sites Inc. and the University of Illinois.

Modern VR Experiences - Virtual prototyping

- In the real world, we build prototypes to understand how a proposed design feels or functions.
- Thanks to 3D printing and related technologies, this is easier than ever.
- At the same time, virtual prototyping enables designers to inhabit a virtual world that contains their prototype (Figure 1.17).

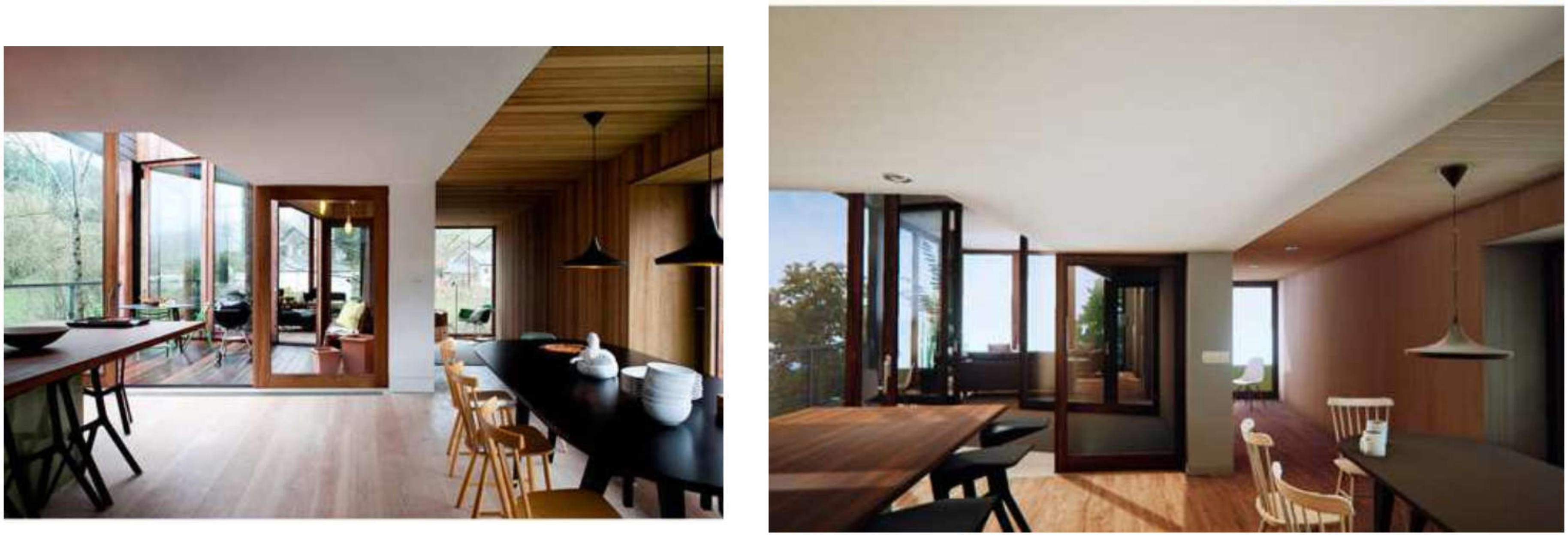


Figure 1.17: Architecture is a prime example of where a virtual prototype is invaluable. This demo, called Ty Hedfan, was created by designer Olivier Demangel. The real kitchen is above and the virtual kitchen is below.

Modern VR Experiences - Virtual prototyping

Figure 1.17: Architecture is a prime example of where a virtual prototype is invaluable. This demo, called Ty Hedfan, was created by designer Olivier Demangel. The real kitchen is above and the virtual kitchen is below.

Modern VR Experiences - Virtual prototyping

- They can quickly interact with it and make modifications.
- They also have opportunities to bring clients into their virtual world so that they can communicate their ideas.
- Imagine you want to remodel your kitchen. You could construct a model in VR and then explain to a contractor exactly how it should look.
- Virtual prototyping in VR has important uses in many businesses, including real estate, architecture, and the design of aircraft, spacecraft, cars, furniture, clothing, and medical instruments.

Modern VR Experiences - Health care

- Although health and safety are challenging VR issues, the technology can also help to improve our health.
- There is an increasing trend toward distributed medicine, in which doctors train people to perform routine medical procedures in remote communities around the world.

Modern VR Experiences - Health care

- Doctors can provide guidance through telepresence, and also use VR technology for training.
- In another use of VR, doctors can immerse themselves in 3D organ models that were generated from medical scan data.

Figure 1.18: A heart visualization system based on images of a real human heart. This was developed by the Jump Trading Simulation and Education Center and the University of Illinois.

Modern VR Experiences - Health care

- This enables them to better plan and prepare for a medical procedure by studying the patient's body shortly before an operation.
- They can also explain medical options to the patient or his family so that they may make more informed decisions.
- In yet another use, VR can directly provide therapy to help patients.

Modern VR Experiences - Health care

- Examples include overcoming phobias and stress disorders through repeated exposure, improving or maintaining cognitive skills in spite of aging, and improving motor skills to overcome balance, muscular, or nervous system disorders.
- VR systems could also one day improve longevity by enabling aging people to virtually travel, engage in fun physical therapy, and overcome loneliness by connecting with family and friends through an interface that makes them feel present and included in remote activities.

Augmented and mixed reality

- In many applications, it is advantageous for users to see the live, real world with some additional graphics superimposed to enhance its appearance.
- This has been referred to as augmented reality or mixed reality.

Augmented and mixed reality

- By placing text, icons, and other graphics into the real world, the user could leverage the power of the Internet to help with many operations such as navigation, social interaction, and mechanical maintenance.
- Many applications to date are targeted at helping businesses to conduct operations more efficiently.
- Imagine a factory environment in which workers see identifying labels above parts that need to assembled, or they can look directly inside of a machine to determine potential replacement parts.

Augmented and mixed reality

- These applications rely heavily on advanced computer vision techniques, which must identify objects, reconstruct shapes, and identify lighting sources in the real world before determining how to draw virtual objects that appear to be naturally embedded.
- Achieving a high degree of reliability becomes a challenge because vision algorithms make frequent errors in unforeseen environments.

Augmented and mixed reality

- The real world lighting conditions must be estimated to determine how to draw the virtual objects and any shadows they might cast onto real parts of the environment and other virtual objects.
- Furthermore, the real and virtual objects may need to be perfectly aligned in some use cases, which places strong burdens on both tracking and computer vision systems.

Augmented and mixed reality

- Several possibilities exist for visual displays. A fixed screen should show images that are enhanced through 3D glasses.
- A digital projector could augment the environment by shining light onto objects, giving them new colors and textures, or by placing text into the real world.

Figure 1.20: Nintendo Pokemon Go is a networked games that allows users to imagine a virtual world that is superimposed on to the real world. They can see Pokemon characters only by looking “through” their smartphone screen.

Augmented and mixed reality

- Two main approaches exist.
- In a see-through display,, the users see most of the real world by simply looking through a transparent material, while the virtual objects appear on the display to disrupt part of the view.
- Recent prototype headsets with advanced see-through display technology include Google Glass, Microsoft Hololens, and Magic Leap.
- Achieving high resolution,

wide field of view, and the ability to block out incoming light remain significant challenges for affordable consumer-grade devices; however, it may become well-solved within a few years.

Augmented and mixed reality

- An alternative is a pass-through display, which sends images from an outward-facing camera to a standard screen inside of the headset.
- Pass-through displays overcome current see-through display problems, but instead suffer from latency, optical distortion, color distortion, and limited dynamic range.