

Chapter 4

Light and Optics

Steven M. LaValle

University of Oulu

Copyright Steven M. LaValle 2019

Available for downloading at <http://vr.cs.uiuc.edu/>

Chapter 4

Light and Optics

Knowing how light propagates in the physical world is crucial to understanding VR. One reason is the interface between visual displays and our eyes. Light is emitted from displays and arrives on our retinas in a way that convincingly reproduces how light arrives through normal vision in the physical world. In the current generation of VR headsets, a system of both engineered and natural lenses (parts of our eyes) guide the light. Another reason to study light propagation is the construction of virtual worlds. Chapter 3 covered purely *geometric* aspects of modeling. The next logical step is to model the *physics* of light propagation through virtual worlds; this will be continued in Chapter 7, which describes what should be rendered on the visual display. Finally, light propagation is also helpful to understanding how cameras work, which provides another way present a virtual world: Through panoramic videos. Cameras are also important for tracking, which will be discussed in Section 9.3.

Section 4.1 covers basic physical properties of light, including its interaction with materials and its spectral properties. Section 4.2 provides idealized models of how lenses work. Section 4.3 then shows many ways that lens behavior deviates from the ideal model, thereby degrading VR experiences. Section 4.4 introduces the human eye as an optical system of lenses, before eyes and human vision are covered in much more detail in Chapter 5. Cameras, which can be considered as engineered eyes, are introduced in Section 4.5. Finally, Section 4.6 briefly covers visual display technologies, which emit light that is intended for consumption by the human eyes.

4.1 Basic Behavior of Light

Light can be described in three ways that appear to be mutually incompatible:

1. Photons: Tiny particles of energy moving through space at high speeds (no need for quantum mechanics in this book!). This interpretation is helpful when considering the amount of light received by a sensor or receptor.

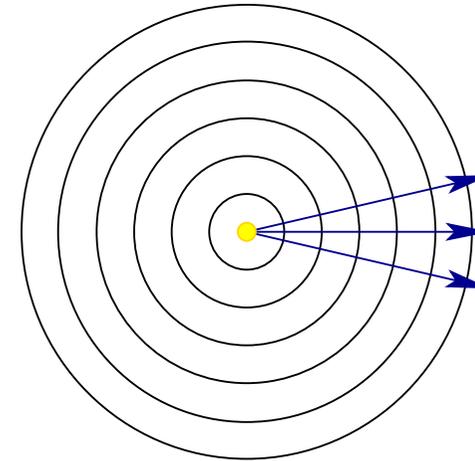


Figure 4.1: Waves and visibility rays emanating from a point light source.

2. Waves: Ripples through space that are similar to waves propagating on the surface of water, but are 3D. The *wavelength* is the distance between peaks. This interpretation is helpful when considering the spectrum of colors.
3. Rays: A ray traces the motion of a single hypothetical photon. The direction is perpendicular to the wavefronts (see Figure 4.1). This interpretation is helpful when explaining lenses and defining the concept of visibility.

Fortunately, modern physics has explained how these interpretations are in fact compatible; each is useful in this book.

Spreading waves Figure 4.1 shows how waves would propagate from a hypothetical point light source. The density would be the same in all directions (radial symmetry), but would decrease as the light source becomes more distant. Recall that the surface area of a sphere with radius r is $4\pi r^2$. Consider centering a spherical screen around the light source. The total number of photons per second hitting a screen of radius 1 should be the same as for a screen of radius 2; however, the density (photons per second per area) should decrease by a factor of $1/4$ because they are distributed over 4 times the area. Thus, photon density decreases quadratically as a function of distance from a point light source.

The curvature of the wavefronts also decreases as the point light source becomes further away. If the waves were to propagate infinitely far away, then they would completely flatten as shown in Figure 4.2. This results in the important case of *parallel wavefronts*. Without the help of lenses or mirrors, it is impossible to actually obtain this case from a tiny light source in the physical world because it cannot be so far away; however, it serves as both a useful approximation for

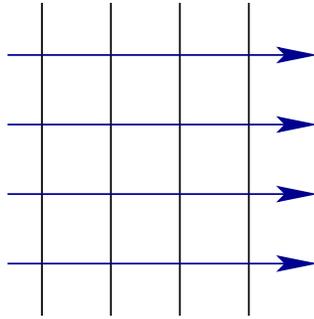


Figure 4.2: If the point light source were “infinitely far” away, then parallel wavefronts would be obtained. Other names for this setting are: Collimated light, parallel rays, rays from infinity, rays to infinity, and zero vergence.

distant light sources and as an ideal way to describe lenses mathematically. Keep in mind that at any finite distance from a point light source, the rays of light always diverge; it is impossible to make them converge without the help of lenses or mirrors.

Interactions with materials As light strikes the surface of a material, one of three behaviors might occur, as shown in Figure 4.3. In the case of *transmission*, the energy travels through the material and exits the other side. For a transparent material, such as glass, the transmitted light rays are slowed down and bend according to Snell’s law, which will be covered in Section 4.2. For a translucent material that is not transparent, the rays scatter into various directions before exiting. In the case of *absorption*, energy is absorbed by the material as the light becomes trapped. The third case is *reflection*, in which the light is deflected from the surface. Along a perfectly smooth or polished surface, the rays reflect in the same way: The exit angle is equal to the entry angle. See Figure 4.4. This case is called *specular reflection*, in contrast to *diffuse reflection*, in which the reflected rays scatter in arbitrary directions. Usually, all three cases of transmission, absorption, and reflection occur simultaneously. The amount of energy divided between the cases depends on many factors, such as the angle of approach, the wavelength, and differences between the two adjacent materials or media.

A jumble of wavelengths Figure 4.1 presented an oversimplified view that will make it easy to understand idealized lenses in Section 4.2. Unfortunately, it misses many details that become important in other settings, such as understanding lens aberrations (Section 4.3) or how light interacts with materials in the physical world. The remainder of this section therefore considers various realistic complications that arise.

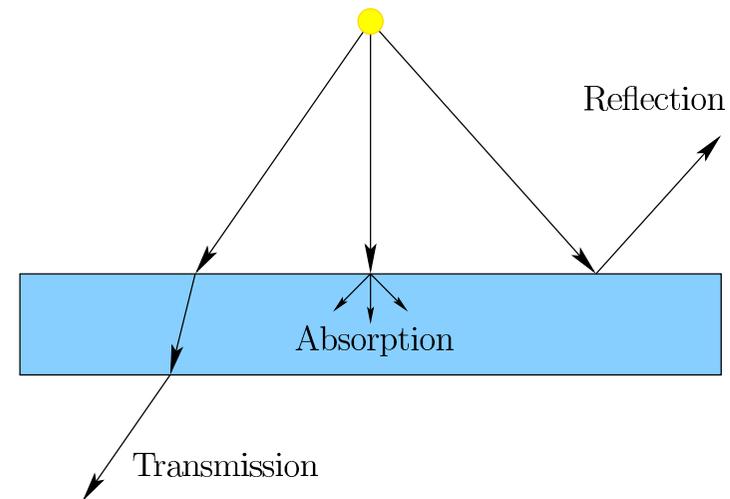


Figure 4.3: As light energy hits the boundary of a different medium, there are three possibilities: transmission, absorption, and reflection.

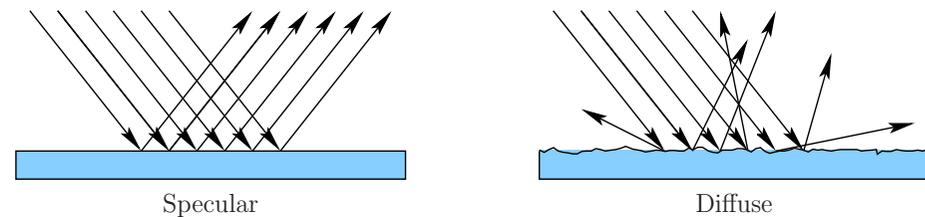


Figure 4.4: Two extreme modes of reflection are shown. Specular reflection means that all rays reflect at the same angle at which they approached. Diffuse reflection means that the rays scatter in a way that could be independent of their approach angle. Specular reflection is common for a polished surface, such as a mirror, whereas diffuse reflection corresponds to a rough surface.

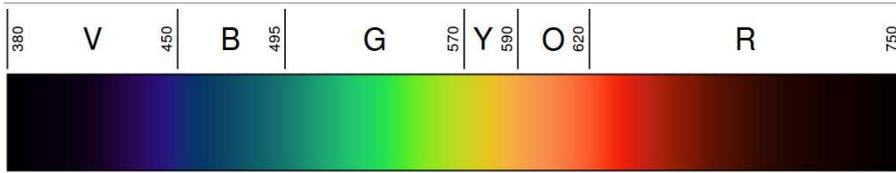


Figure 4.5: Visible light spectrum corresponds to the range of electromagnetic waves that have wavelengths between 400nm and 700nm. (Figure by David Eccles for Wikipedia.)

Coherent versus jumbled light The first complication is that light sources usually do not emit *coherent light*, a term that means the wavefronts are perfectly aligned in time and space. A laser is an exceptional case that indeed produces coherent light. It emits parallel waves of a constant wavelength that are also synchronized in time so that their peaks align as they propagate. Common light sources, such as light bulbs and the sun, instead emit a jumble of waves that have various wavelengths and do not have their peaks aligned.

Wavelengths and colors To make sense out of the jumble of waves, we will describe how they are distributed in terms of wavelengths. Figure 4.5 shows the range of wavelengths that are visible to humans. Each wavelength corresponds to a *spectral color*, which is what we would perceive with a coherent light source fixed at that wavelength alone. Wavelengths between 700 and 1000nm are called *infrared*, which are not visible to us, but some cameras can sense them (see Section 9.3). Wavelengths between 100 and 400nm are called *ultraviolet*; they are not part of our visible spectrum, but some birds, insects, and fish can perceive ultraviolet wavelengths over 300nm. Thus, our notion of visible light is already tied to *human* perception.

Spectral power Figure 4.6 shows how the wavelengths are distributed for common light sources. An ideal light source would have all visible wavelengths represented with equal energy, leading to idealized *white* light. The opposite is total darkness, which is *black*. We usually do not allow a light source to propagate light directly onto our retinas (don't stare at the sun!). Instead, we observe light that is reflected from objects all around us, causing us to perceive their color. Each surface has its own distribution of wavelengths that it *reflects*. The fraction of light energy that is reflected back depends on the wavelength, leading to the plots shown in Figure 4.7. For us to perceive an object surface as red, the red wavelengths must be included in the light source *and* the surface must strongly reflect red wavelengths. Other wavelengths must also be suppressed. For example, the light source could be white (containing all wavelengths) and the object could strongly reflect all wavelengths, causing the surface to appear white, not

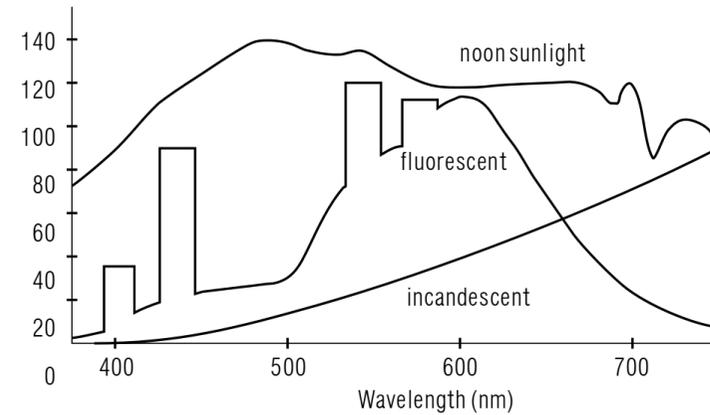


Figure 4.6: The *spectral power distribution* for some common light sources. (Figure from [14]).

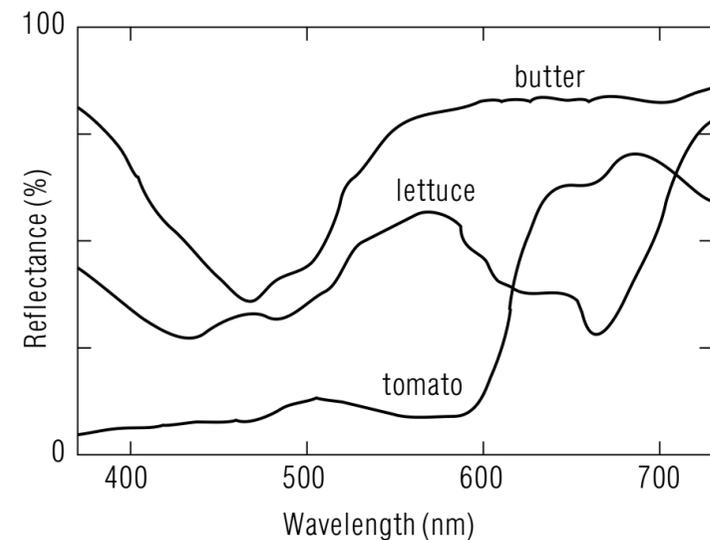


Figure 4.7: The *spectral reflection function* of some common familiar materials. (Figure from [14]).

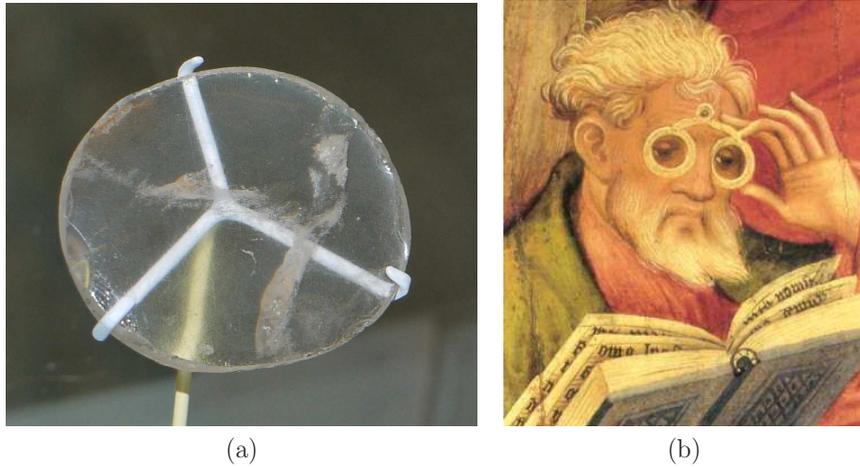


Figure 4.8: (a) The earliest known artificially constructed lens, which was made between 750 and 710 BC in ancient Assyrian Nimrud. It is not known whether this artifact was purely ornamental or used to produce focused images. Picture from the British Museum. (b) A painting by Conrad von Soest from 1403, which shows the use of reading glasses for an elderly male.

red. Section 6.3 will provide more details on color perception.

Frequency Often times, it is useful to talk about frequency instead of wavelength. The *frequency* is the number of times per second that wave peaks pass through a fixed location. Using both the wavelength λ and the speed s , the frequency f is calculated as:

$$f = \frac{s}{\lambda}. \quad (4.1)$$

The speed of light in a vacuum is a universal constant c with value approximately equal to 3×10^8 m/s. In this case, $s = c$ in (4.1). Light propagates roughly 0.03 percent faster in a vacuum than in air, causing the difference to be neglected in most engineering calculations. Visible light in air has a frequency range of roughly 400 to 800 terahertz, which is obtained by applying (4.1). As light propagates through denser media, such as water or lenses, s is significantly smaller; that difference is the basis of optical systems, which are covered next.

4.2 Lenses

Lenses have been made for thousands of years, with the oldest known artifact shown in Figure 4.8(a). It was constructed before 700 BC in Assyrian Nimrud.

Whether constructed from transparent materials or from polished surfaces that act as mirrors, lenses bend rays of light so that a focused image is formed. Over the centuries, their uses have given rise to several well-known devices, such as eyeglasses (Figure 4.8(b)), telescopes, magnifying glasses, binoculars, cameras, and microscopes. Optical engineering is therefore filled with design patterns that indicate how to optimize the designs of these well-understood devices. VR headsets are unlike classical optical devices, leading to many new challenges that are outside of standard patterns that have existed for centuries. Thus, the lens design patterns for VR are still being written. The first step toward addressing the current challenges is to understand how simple lenses work.

Snell's Law Lenses work because of *Snell's law*, which expresses how much rays of light bend when entering or exiting a transparent material. Recall that the speed of light in a medium is less than the speed c in a vacuum. For a given material, let its *refractive index* be defined as

$$n = \frac{c}{s}, \quad (4.2)$$

in which s is the speed of light in the medium. For example, $n = 2$ means that light takes twice as long to traverse the medium than through a vacuum. For some common examples, $n = 1.000293$ for air, $n = 1.33$ for water, and $n = 1.523$ for crown glass.

Figure 4.9 shows what happens to incoming light waves and rays. Suppose in this example that the light is traveling from air into glass, so that $n_1 < n_2$. Let θ_1 represent the incoming angle with respect to the surface normal, and let θ_2 represent the resulting angle as it passes through the material. Snell's law relates the four quantities as

$$n_1 \sin \theta_1 = n_2 \sin \theta_2. \quad (4.3)$$

Typically, n_1/n_2 and θ_1 are given, so that (4.3) is solved for θ_2 to obtain

$$\theta_2 = \sin^{-1} \left(\frac{n_1 \sin \theta_1}{n_2} \right). \quad (4.4)$$

If $n_1 < n_2$, then θ_2 is closer to perpendicular than θ_1 . If $n_1 > n_2$, then θ_2 is further from perpendicular. The case of $n_1 > n_2$ is also interesting in that light may not penetrate the surface if the incoming angle θ_1 is too large. The range of \sin^{-1} is 0 to 1, which implies that (4.4) provides a solution for θ_2 only if

$$(n_1/n_2) \sin \theta_1 \leq 1. \quad (4.5)$$

If the condition above does not hold, then the light rays reflect from the surface. This situation occurs while under water and looking up at the surface. Rather than being able to see the world above, a swimmer might instead see a reflection, depending on the viewing angle.

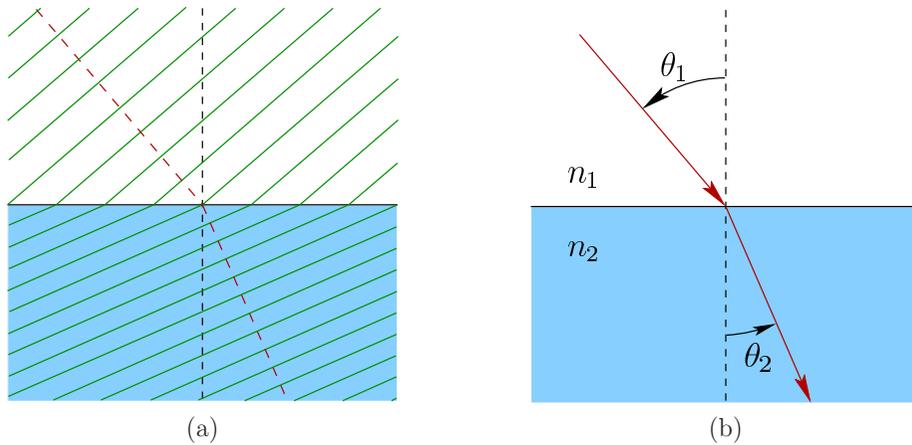


Figure 4.9: Propagating wavefronts from a medium with low refractive index (such as air) to one with a higher index (such as glass). (a) The effect of slower propagation on the wavefronts is shown as they enter the lower medium. (b) This shows the resulting bending of a light ray, which is always perpendicular to the wavefronts. Snell's Law relates the refractive indices and angles as $n_1 \sin \theta_1 = n_2 \sin \theta_2$.

Prisms Imagine shining a laser beam through a prism, as shown in Figure 4.10. Snell's Law can be applied to calculate how the light ray bends after it enters and exits the prism. Note that for the upright prism, a ray pointing slightly upward becomes bent downward. Recall that a larger refractive index inside the prism would cause greater bending. By placing the prism upside down, rays pointing slightly downward are bent upward. Once the refractive index is fixed, the bending depends only on the angles at which the rays enter and exit the surface, rather than on the thickness of the prism. To construct a lens, we will exploit this principle and construct a kind of curved version of Figure 4.10.

Simple convex lens Figure 4.11 shows a simple convex lens, which should remind you of the prisms in Figure 4.10. Instead of making a diamond shape, the lens surface is spherically curved so that incoming, parallel, horizontal rays of light converge to a point on the other side of the lens. This special place of convergence is called the *focal point*. Its distance from the lens center is called the *focal depth* or *focal length*.

The incoming rays in Figure 4.11 are special in two ways: 1) They are parallel, thereby corresponding to a source that is infinitely far away, and 2) they are perpendicular to the plane in which the lens is centered. If the rays are parallel but not perpendicular to the lens plane, then the focal point shifts accordingly, as shown in Figure 4.12. In this case, the focal point is not on the optical axis. There are two DOFs of incoming ray directions, leading to a *focal plane* that contains all

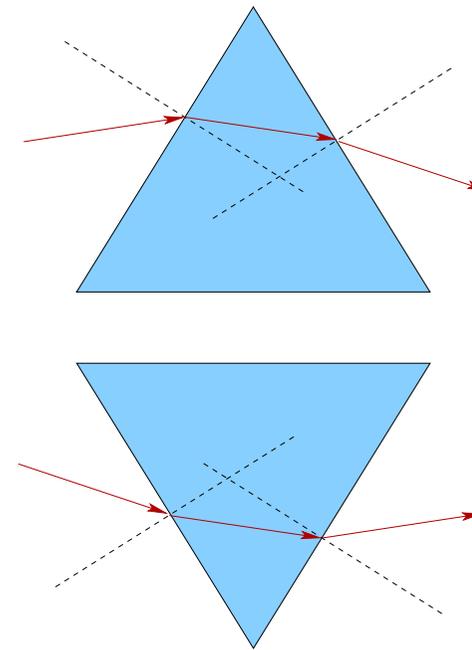


Figure 4.10: The upper part shows how a simple prism bends ascending rays into descending rays, provided that the incoming ray slope is not too high. This was achieved by applying Snell's law at the incoming and outgoing boundaries. Placing the prism upside down causes descending rays to become ascending. Putting both of these together, we will see that a lens is like a stack of prisms that force diverging rays to converge through the power of refraction.

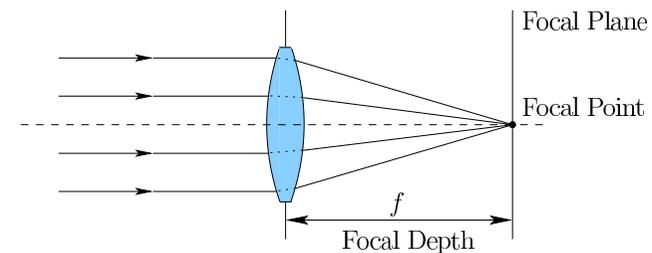


Figure 4.11: A simple convex lens causes parallel rays to converge at the focal point. The dashed line is the *optical axis*, which is perpendicular to the lens and pokes through its center.

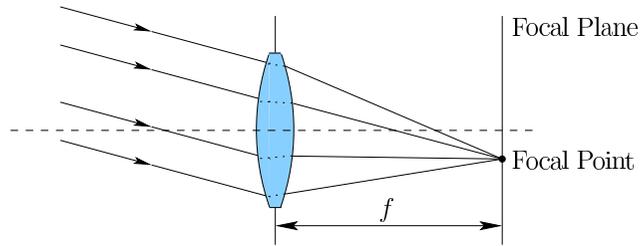


Figure 4.12: If the rays are not perpendicular to the lens, then the focal point is shifted away from the optical axis.

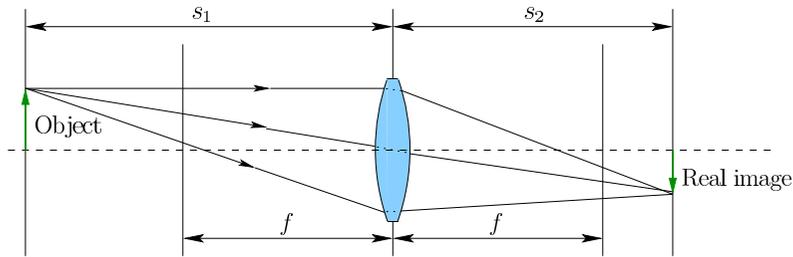


Figure 4.13: In the real world, an object is not infinitely far away. When placed at distance s_1 from the lens, a real image forms in a focal plane at distance $s_2 > f$ behind the lens, as calculated using (4.6).

of the focal points. Unfortunately, this planarity is just an approximation; Section 4.3 explains what really happens. In this idealized setting, a *real image* is formed in the image plane, as if it were a projection screen that is showing how the world looks in front of the lens (assuming everything in the world is very far away).

If the rays are not parallel, then it may still be possible to focus them into a real image, as shown in Figure 4.13. Suppose that a lens is given that has focal length f . If the light source is placed at distance s_1 from the lens, then the rays from that will be in focus if and only if the following equation is satisfied (which is derived from Snell's law):

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f}. \quad (4.6)$$

Figure 4.11 corresponds to the idealized case in which $s_1 = \infty$, for which solving (4.6) yields $s_2 = f$. What if the object being viewed is not completely flat and lying in a plane perpendicular to the lens? In this case, there does not exist a single plane behind the lens that would bring the entire object into focus. We must tolerate the fact that most of it will be approximately in focus. Unfortunately, this is the situation almost always encountered in the real world, including the focus provided by our own eyes (see Section 4.4).

If the light source is placed too close to the lens, then the outgoing rays might

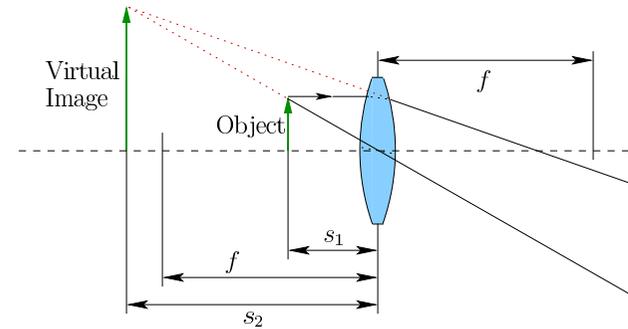


Figure 4.14: If the object is very close to the lens, then the lens cannot force its outgoing light rays to converge to a focal point. In this case, however, a virtual image appears and the lens works as a magnifying glass. This is the way lenses are commonly used for VR headsets.

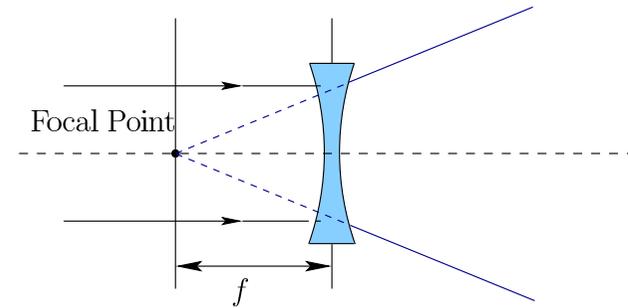


Figure 4.15: In the case of a concave lens, parallel rays are forced to diverge. The rays can be extended backward through the lens to arrive at a focal point on the left side. The usual sign convention is that $f < 0$ for concave lenses.

be diverging so much that the lens cannot force them to converge. If $s_1 = f$, then the outgoing rays would be parallel ($s_2 = \infty$). If $s_1 < f$, then (4.6) yields $s_2 < 0$. In this case, a real image is not formed; however, something interesting happens: The phenomenon of *magnification*. A *virtual image* appears when looking into the lens, as shown in Figure 4.14. This exactly what happens in the case of the View-Master and the VR headsets that were shown in Figure 2.11. The screen is placed so that it appears magnified. To the user viewing looking through the lenses, it appears as if the screen is infinitely far away (and quite enormous!).

Lensmaker's equation For a given simple lens, the focal length f can be calculated using the Lensmaker's Equation (also derived from Snell's law):

$$(n_2 - n_1) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) = \frac{1}{f}. \quad (4.7)$$

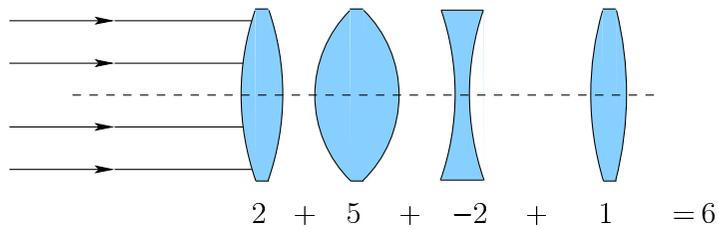


Figure 4.16: To calculate the combined optical power of a chain of lenses, the algebra is simple: Add their diopeters. This arrangement of four lenses is equivalent to a 6-diopter lens, which has a focal length of 0.1667m.

The parameters r_1 and r_2 represent the radius of curvature of each of the two lens surfaces (front and back). This version assumes a *thin lens approximation*, which means that the lens thickness is small relative to r_1 and r_2 . Also, it is typically assumed that $n_1 = 1$, which is approximately true for air.

Concave lenses For the sake of completeness, we include the case of a *concave* simple lens, shown in Figure 4.15. Parallel rays are forced to diverge, rather than converge; however, a meaningful notion of negative focal length exists by tracing the diverging rays backwards through the lens. The Lensmaker's Equation (4.7) can be slightly adapted to calculate negative f in this case [5].

Diopters For optical systems used in VR, several lenses will be combined in succession. What is the effect of the combination? A convenient method to answer this question with simple arithmetic was invented by ophthalmologists. The idea is to define a *diopter*, which is $D = 1/f$. Thus, it is the reciprocal of the focal length. If a lens focuses parallel rays at a distance of 0.2m in behind the lens, then $D = 5$. A larger diopter D means greater converging power. Likewise, a concave lens yields $D < 0$, with a lower number implying greater divergence. To combine several nearby lenses in succession, we simply add their diopeters to determine their equivalent power as a single, simple lens. Figure 4.16 shows a simple example.

4.3 Optical Aberrations

If lenses in the real world behaved *exactly* as described in Section 4.2, then VR systems would be much simpler and more impressive than they are today. Unfortunately, numerous imperfections, called *aberrations*, degrade the images formed by lenses. Because these problems are perceptible in everyday uses, such as viewing content through VR headsets or images from cameras, they are important to understand so that some compensation for them can be designed into the VR system.

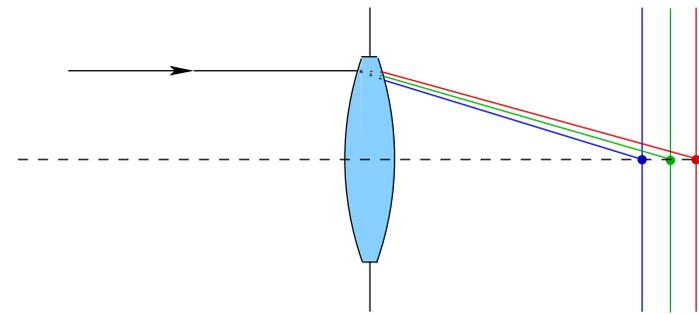


Figure 4.17: Chromatic aberration is caused by longer wavelengths traveling more quickly through the lens. The unfortunate result is a different focal plane for each wavelength or color.

Chromatic aberration Recall from Section 4.1 that light energy is usually a jumble of waves with a spectrum of wavelengths. You have probably seen that the colors of the entire visible spectrum nicely separate when white light is shined through a prism. This is a beautiful phenomenon, but for lenses it is terrible annoyance because it separates the focused image based on color. This problem is called *chromatic aberration*.

The problem is that the speed of light through a medium depends on the wavelength. We should therefore write a material's refractive index as $n(\lambda)$ to indicate that it is a function of λ . Figure 4.17 shows the effect on a simple convex lens. The focal depth becomes a function of wavelength. If we shine red, green, and blue lasers directly into the lens along the same ray, then each color would cross the optical axis in a different place, resulting in red, green, and blue focal points. Recall the spectral power distribution and reflection functions from Section 4.1. For common light sources and materials, the light passing through a lens results in a whole continuum of focal points. Figure 4.18 shows an image with chromatic aberration artifacts. Chromatic aberration can be reduced at greater expense by combining convex and concave lenses of different materials so that the spreading rays are partly coerced into converging [17].

Spherical aberration Figure 4.19 shows *spherical aberration*, which is caused by rays further away from the lens center being refracted more than rays near the center. The result is similar to that of chromatic aberration, but this phenomenon is a *monochromatic* aberration because it is independent of the light wavelength. Incoming parallel rays are focused at varying depths, rather than being concentrated at a single point. The result is some blur that cannot be compensated for by moving the object, lens, or image plane. Alternatively, the image might instead focus onto a curved surface, called the *Petzval surface*, rather than the image plane. This aberration arises due to the spherical shape of the lens. An



Figure 4.18: The upper image is properly focused whereas the lower image suffers from chromatic aberration. (Figure by Stan Zurek, license CC-BY-SA-2.5.)

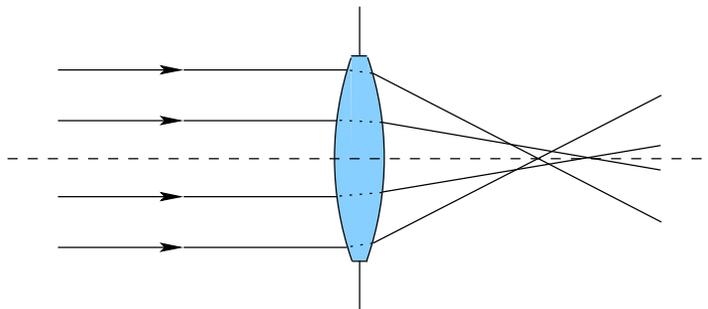


Figure 4.19: Spherical aberration causes imperfect focus because rays away from the optical axis are refracted more than those at the periphery.

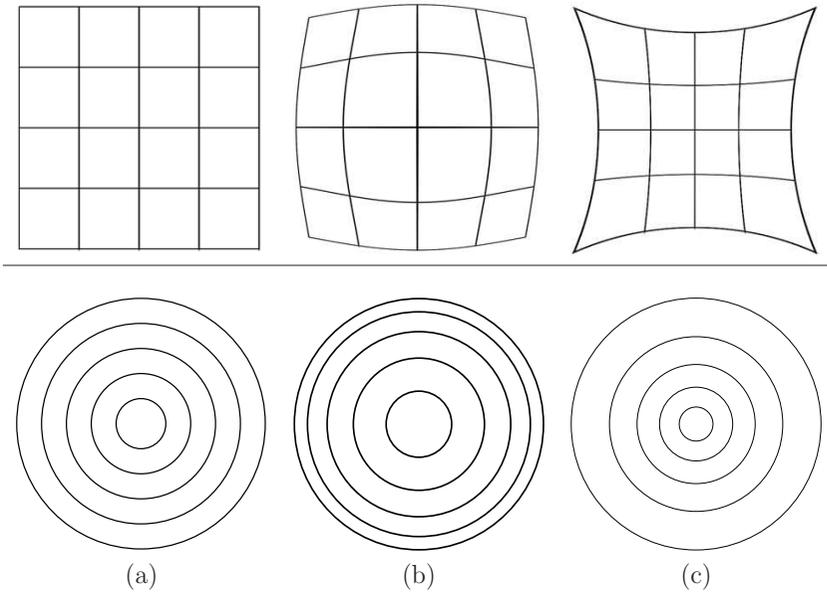


Figure 4.20: Common optical distortions. (a) Original images. (b) Barrel distortion. (c) Pincushion distortion. For the upper row, the grid becomes nonlinearly distorted. For lower row illustrates how circular symmetry is nevertheless maintained.

aspheric lens is more complex and has non-spherical surfaces that are designed to specifically eliminate the spherical aberration and reduce other aberrations.

Optical distortion Even if the image itself projects onto the image plane it might be distorted at the periphery. Assuming that the lens is radially symmetric, the distortion can be described as a stretching or compression of the image that becomes increasingly severe away from the optical axis. Figure 4.20 shows how this effects the image for two opposite cases: *barrel distortion* and *pincushion distortion*. For lenses that have a wide field-of-view, the distortion is stronger, especially in the extreme case of a *fish-eyed lens*. Figure 4.21 shows an image that has strong barrel distortion. Correcting this distortion is crucial for current VR headsets that have a wide field-of-view; otherwise, the virtual world would appear to be warped.

Astigmatism Figure 4.22 depicts *astigmatism*, which is a lens aberration that occurs for incoming rays that are not perpendicular to the lens. Up until now, our lens drawings have been 2D; however, a third dimension is needed to understand this new aberration. The rays can be off-axis in one dimension, but aligned in



Figure 4.21: An image with barrel distortion, taken by a fish-eyed lens. (Image by Wikipedia user Ilveon.)

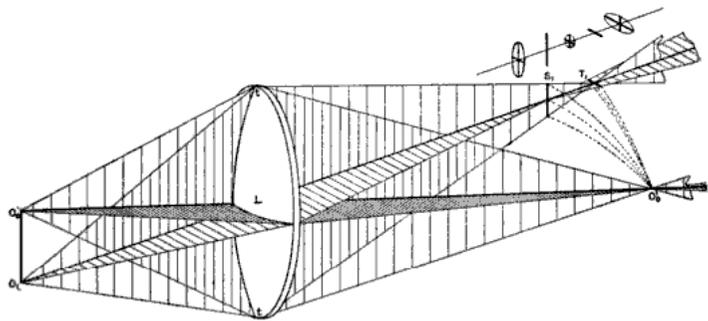


Figure 4.22: Astigmatism is primarily caused by incoming rays being off-axis in one plane, but close to perpendicular in another. (Figure from [20].)

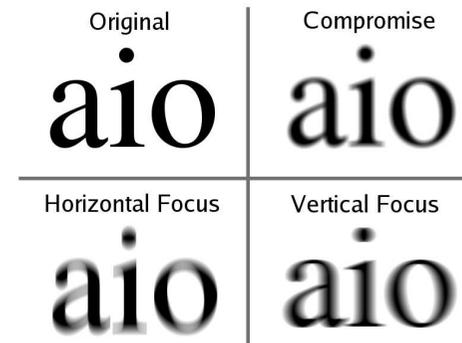


Figure 4.23: Due to astigmatism, it becomes impossible to bring the image perfectly into focus. At one depth, it might be focused horizontally, while at another it is focused vertically. We are forced to choose a compromise.

another. By moving the image plane along the optical axis, it becomes impossible to bring the image into focus. Instead, horizontal and vertical focal depths appear, as shown in Figure 4.23.

Coma and flare Finally, *coma* is yet another aberration. In this case, the image magnification varies dramatically as the rays are far from perpendicular to the lens. The result is a “comet” pattern in the image plane. Another phenomenon is *lens flare*, in which rays from very bright light scatter through the lens and often show circular patterns. This is often seen in movies as the viewpoint passes by the sun or stars, and is sometimes added artificially.

All of the aberrations of this section complicate the system or degrade the experience in a VR headset; therefore, substantial engineering effort is spent on mitigating these problems.

4.4 The Human Eye

We have covered enough concepts in this chapter to describe the basic operation of the human eye, which is clearly an important component in any VR system. Here it will be considered as part of an optical system of lenses and images. The physiological and perceptual parts of human vision are deferred until Chapter 5.

Figure 4.24 shows a cross section of the human eye facing left. Parallel light rays are shown entering from the left; compare to Figure 4.11, which showed a similar situation for an engineered convex lens. Although the eye operation is similar to the engineered setting, several important differences arise at this stage. The focal plane is replaced by a spherically curved surface called the *retina*. The retina contains *photoreceptors* that convert the light into neural pulses; this is

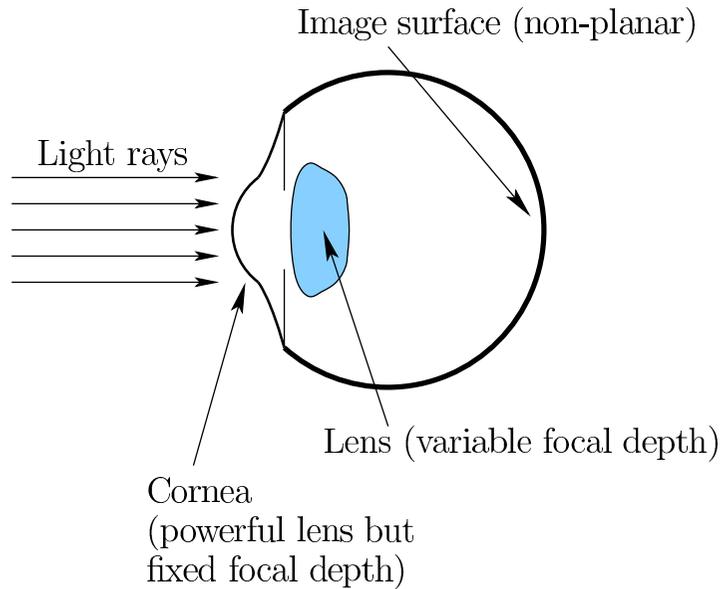


Figure 4.24: A simplified view of the human eye as an optical system.

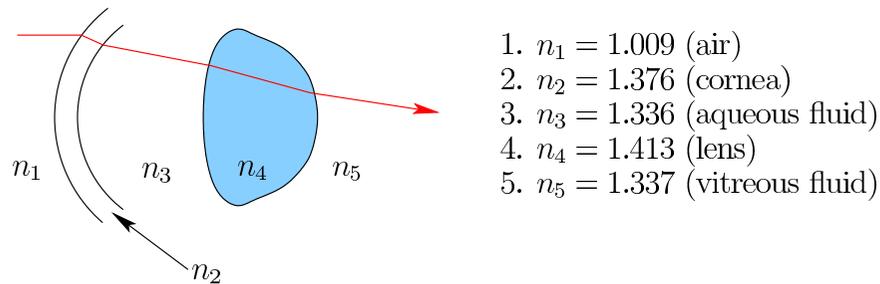


Figure 4.25: A ray of light travels through five media before hitting the retina. The indices of refraction are indicated. Considering Snell's law, the greatest bending occurs due to the transition from air to the cornea. Note that once the ray enters the eye, it passes through only liquid or solid materials.

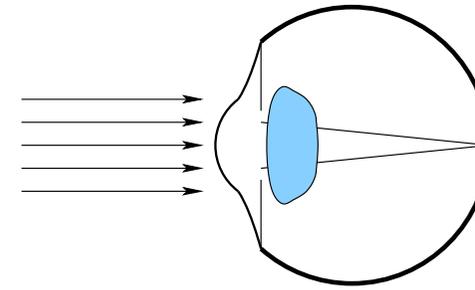


Figure 4.26: Normal eye operation, with relaxed lens.

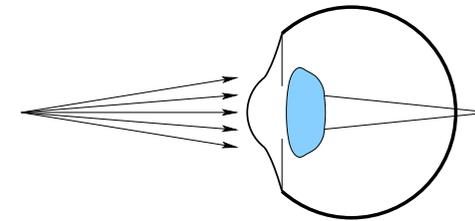


Figure 4.27: A closer object yields diverging rays, but with a relaxed lens, the image is blurry on the retina.

covered in Sections 5.1 and 5.2. The interior of the eyeball is actually liquid, as opposed to air. The refractive indices of materials along the path from the outside air to the retina are shown in Figure 4.25.

The optical power of the eye The outer diameter of the eyeball is roughly 24mm, which implies that a lens of at least 40D would be required to cause convergence of parallel rays onto the retina center inside of the eye (recall diopters from Section 4.2). There are effectively two convex lenses: The *cornea* and the *lens*. The cornea is the outermost part of the eye where the light first enters and has the greatest optical power, approximately 40D. The eye lens is less powerful and provides an additional 20D. By adding diopters, the combined power of the cornea and lens is 60D, which means that parallel rays are focused onto the retina at a distance of roughly 17mm from the outer cornea. Figure 4.26 shows how this system acts on parallel rays for a human with normal vision. Images of far away objects are thereby focused onto the retina.

Accommodation What happens when we want to focus on a nearby object, rather than one “infinitely far” away? Without any changes to the optical system, the image would be blurry on the retina, as shown in Figure 4.27. Fortunately, and miraculously, the lens changes its diopter to accommodate the closer distance.

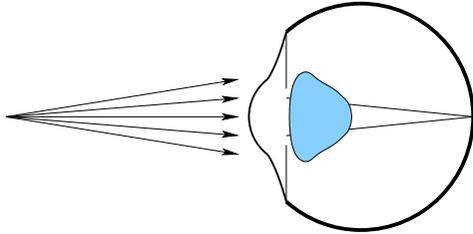


Figure 4.28: The process of accommodation: The eye muscles pull on the lens, causing it to increase the total optical power and focus the image on the retina.

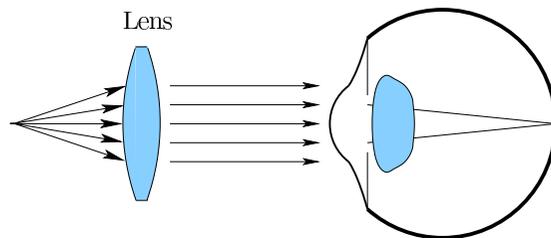


Figure 4.29: Placing a convex lens in front of the eye is another way to increase the optical power so that nearby objects can be brought into focus by the eye. This is the principle of reading glasses.

This process is appropriately called *accommodation*, as is depicted in Figure 4.28. The diopter change is effected through muscles that pull on the lens to change its shape. In young children, the lens can increase its power by an additional 15 to 20D, which explains why a child might hold something right in front of your face and expect you to focus on it; they can! At 20D, this corresponds to focusing on an object that is only 5cm from the cornea. Young adults already lose this ability and can accommodate up to about 10D. Thus, with normal vision they can read a book down to a distance of about 10cm (with some eye strain). Once adults reach 50 years old, little or no accommodation ability remains. This condition is called *presbyopia*. Figure 4.29 shows the most common treatment, which is to place reading glasses in front of the eye.

Vision abnormalities The situations presented so far represent normal vision throughout a person's lifetime. One problem could be that the optical system simply does not have enough optical power to converge parallel rays onto the retina. This condition is called *hyperopia* or *farsightedness*. Eyeglasses come to the rescue. The simple fix is to place a convex lens (positive diopter) in front of the eye, as in the case of reading glasses. In the opposite direction, some eyes have too much optical power. This case is called *myopia* or *nearsightedness*, and a concave lens (negative diopter) is placed in front of the eye to reduce the optical power appropriately. Recall that we have two eyes, not one. This allows the possibility for each eye to have a different problem, resulting in different lens diopters per eye. Other vision problems may exist beyond optical power. The most common is astigmatism, which was covered in Section 4.3. In human eyes this is caused by the cornea having an excessively elliptical shape, rather than being radially symmetric. Special, non-simple lenses are needed to correct this condition. You might also wonder whether the aberrations from Section 4.3, such as chromatic aberration, occur in the human eye. They do, however they are corrected automatically by our brains because we have learned to interpret such flawed images our entire lives!

A simple VR headset Now suppose we are constructing a VR headset by placing a screen very close to the eyes. Young adults would already be unable to bring it into focus if it were closer than 10cm. We want to bring it close so that it fills the view of the user. Therefore, the optical power is increased by using a convex lens, functioning in the same way as reading glasses. See Figure 4.30. This is also the process of magnification, from Section 4.2. The lens is usually placed at the distance of its focal depth. Using (4.6), this implies that $s_2 = -f$, resulting in $s_1 = \infty$. The screen appears as an enormous virtual image that is infinitely far away. Note, however, that a real image is nevertheless projected onto the retina. We do not perceive the world around us unless real images are formed on our retinas.

To account for people with vision problems, a focusing knob may be appear

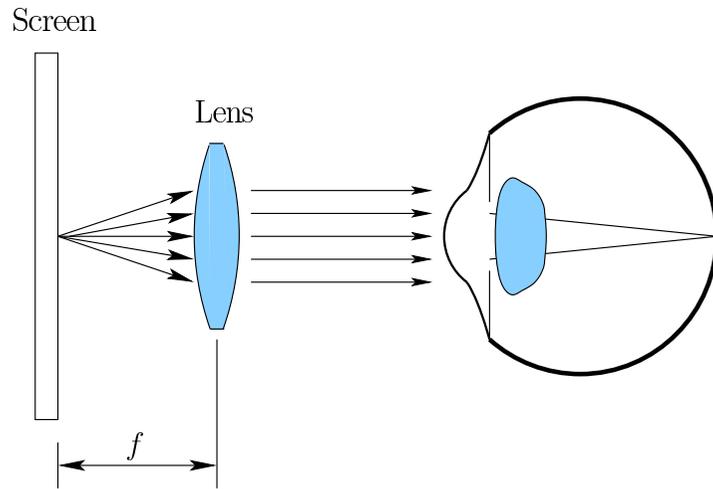


Figure 4.30: In VR headsets, the lens is placed so that the screen appears to be infinitely far away.

on the headset, which varies the distance between the lens and the screen. This adjusts the optical power so that the rays between the lens and the cornea are no longer parallel. They can be made to converge, which helps people with hyperopia. Alternatively, they can be made to diverge, which helps people with myopia. Thus, they can focus sharply on the screen without placing their eyeglasses in front of the lens. However, if each eye requires a different diopter, then a focusing knob would be required for *each* eye. Furthermore, if they have astigmatism, then it cannot be corrected. Placing eyeglasses inside of the headset may be the only remaining solution, but it may be uncomfortable and could reduce the field of view.

Many details have been skipped or dramatically simplified in this section. One important detail for a VR headset is each lens should be centered perfectly in front of the cornea. If the distance between the two lenses is permanently fixed, then this is impossible to achieve for everyone who uses the headset. The *interpupillary distance*, or IPD, is the distance between human eye centers. The average among humans is around 64mm, but it varies greatly by race, gender, and age (in the case of children). To be able to center the lenses for everyone, the distance between lens centers should be adjustable from around 55 to 75mm. This is a common range for binoculars. Unfortunately, the situation is not even this simple because our eyes also rotate within their sockets, which changes the position and orientation of the cornea with respect to the lens. This amplifies optical aberration problems that were covered in Section 4.3. Eye movements will be covered in Section 5.3. Another important detail is the fidelity of our vision: What pixel density is needed for the screen that is placed in front of our eyes so that we do not notice the pixels?

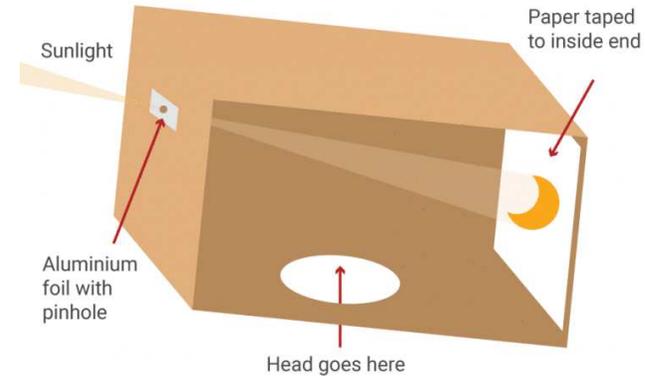


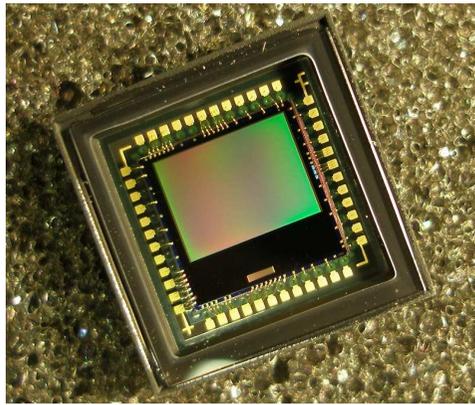
Figure 4.31: A pinhole camera that is recommended for viewing a solar eclipse. (Figure from <https://www.timeanddate.com/>.)

A similar question is how many dots-per-inch (DPI) are needed on a printed piece of paper so that we do not see the dots, even when viewed under a magnifying glass? We return to this question in Section 5.1.

4.5 Cameras

Now that we have covered the human eye, it seems natural to describe an engineered eye, otherwise known as a camera. People have built and used cameras for hundreds of years, starting with a *camera obscura* that allows light to pass through a *pinhole* and onto a surface that contains the real image. Figure 4.31 shows an example that you might have constructed to view a solar eclipse. (Recall the perspective transformation math from Section 3.4.) Eighteenth-century artists incorporated a mirror and tracing paper to un-invert the image and allow it to be perfectly copied. Across the 19th century, various chemically based technologies were developed to etch the image automatically from the photons hitting the imaging surface. Across the 20th century, film was in widespread use, until *digital cameras* avoided the etching process altogether by electronically capturing the image using a sensor. Two popular technologies have been a *Charge-Coupled Device (CCD)* array and a *CMOS active-pixel image sensor*, which is shown in Figure 4.32(a). Such digital technologies record the amount of light hitting each pixel location along the image, which directly produces a captured image. The costs of these devices has plummeted in recent years, allowing hobbyists to buy a camera module such as the one shown in Figure 4.32(b) for under \$30 US.

Shutters Several practical issues arise when capturing digital images. The image is an 2D array of *pixels*, each of which having red (R), green (G), and blue



(a)



(b)

Figure 4.32: (a) A CMOS active-pixel image sensor. (b) A low-cost CMOS camera module (SEN-11745), ready for hobbyist projects (picture by SparkFun).

(B) values that typically range from 0 to 255. Consider the total amount of light energy that hits the image plane. For a higher-resolution camera, there will generally be less photons per pixel because the pixels are smaller. Each sensing element (one per color per pixel) can be imagined as a bucket that collects photons, much like drops of rain. To control the amount of photons, a *shutter* blocks all the light, opens for a fixed interval of time, and then closes again. For a long interval (low *shutter speed*), more light is collected; however, the drawbacks are that moving objects in the scene will become blurry and that the sensing elements could become saturated with too much light. Photographers must strike a balance when determining the shutter speed to account for the amount of light in the scene, the sensitivity of the sensing elements, and the motion of the camera and objects in the scene.

Also relating to shutters, CMOS sensors unfortunately work by sending out the image information sequentially, line-by-line. The sensor is therefore coupled with a *rolling shutter*, which allows light to enter for each line, just before the information is sent. This means that the capture is not synchronized over the entire image, which leads to odd artifacts, such as the one shown in Figure 4.33. Image processing algorithms that work with rolling shutters and motion typically transform the image to correct for this problem. CCD sensors grab and send the entire image at once, resulting in a *global shutter*. CCDs have historically been more expensive than CMOS sensors, which resulted in widespread appearance of rolling shutter cameras in smartphones; however, the cost of global shutter



Figure 4.33: The wings of a flying helicopter are apparently bent backwards due to the rolling shutter effect.

cameras is rapidly decreasing.

Aperture The optical system also impacts the amount of light that arrives to the sensor. Using a pinhole, as shown in Figure 4.31, light would fall onto the image sensor, but it would not be bright enough for most purposes (other than viewing a solar eclipse). Therefore, a convex lens is used instead so that multiple rays are converged to the same point in the image plane; recall Figure 4.11. This generates more photons per sensing element. The main drawback is that the lens sharply focuses objects at a single depth, while blurring others; recall (4.6). In the pinhole case, all depths are essentially “in focus”, but there might not be enough light. Photographers therefore want to tune the optical system to behave more like a pinhole or more like a full lens, depending on the desired outcome. The result is a controllable *aperture* (Figure 4.34), which appears behind the lens and sets the size of the hole through which the light rays enter. A small radius mimics a pinhole by blocking all but the center of the lens. A large radius allows light to pass through the entire lens. Our eyes control the light levels in a similar manner by contracting or dilating our pupils. Finally, note that the larger the aperture, the more that the aberrations covered in Section 4.3 interfere with the imaging process.

4.6 Displays

Section 2.1 introduced displays as devices that stimulate a sense organ, which in this chapter is the human eye. What should we consider to be the first visual displays? Paintings from Section 1.3 as early instances of displays, but they have the unfortunate limitation that they display a single image in their lifetime. The

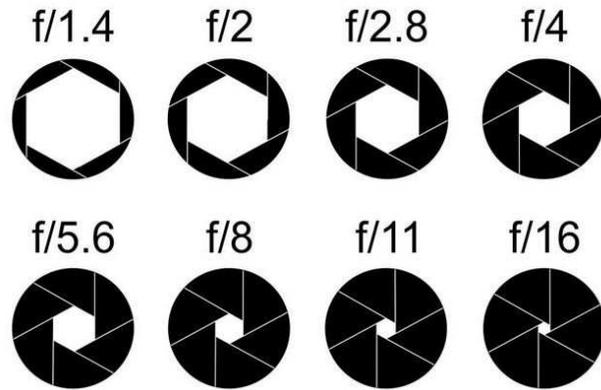


Figure 4.34: A spectrum of aperture settings, which control the amount of light that enters the lens. The values shown are called the *focal ratio* or *f-stop*.

ability to change images, ultimately leading to motion pictures, was first enabled by projection technologies. The camera obscura principle was refined into a *magic lantern*, which was invented in the 17th century by Christiaan Huygens. Popular until the 19th century, it allowed small, painted images to be projected onto a wall through the use of lenses and a bright light source. By the end of the 19th century, motion pictures could be shown on a large screen by a mechanical projector that cycled quickly through the frames.

Cathode ray tubes The most important technological leap was the *cathode ray tube* or *CRT*, which gave birth to electronic displays, launched the era of television broadcasting, and helped shape many concepts and terms that persist in modern displays today. Figure 4.35 shows the basic principles. The CRT enabled videos to be rendered to a screen, frame by frame. Each frame was scanned out line by line due to the physical limitations of the hardware. The scanning needed to repeat frequently, known as refreshing the phosphor elements. Each light in each position would persist for less than a millisecond. The scanout behavior and timing remains today for modern smartphone displays because of memory and computation architectures, but it is not ideal for VR usage. Section 6.2.2 will explain how motion is perceived when a sequence of frames is rendered to a display, and what goes wrong with VR systems.

The next major advance was to enable each picture element, or pixel, to be directly and persistently lit. Various technologies have been used to produce *flat-panel displays*, the output of which is illustrated in Figure 4.36. *Liquid crystal displays* (*LCD* displays) became widely available in calculators in the 1970s, and progressed into larger, colorful screens by the 1990s. The liquid crystals themselves do not emit light, but most commonly a *backlight* shines from behind to illuminate

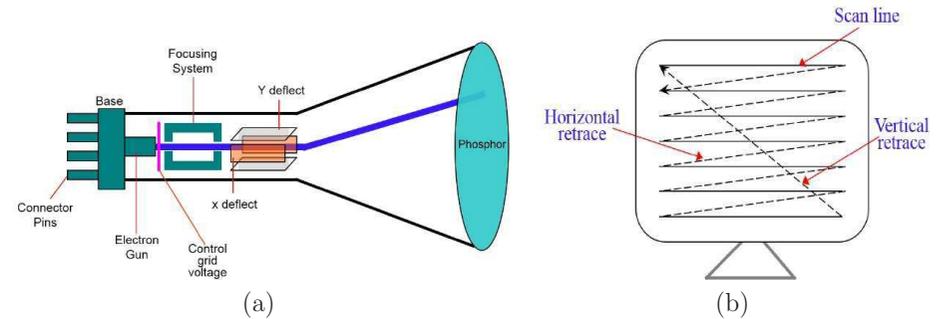


Figure 4.35: (a) In a cathode ray tube (CRT), an electron gun emits a beam of electrons, which are deflected in both X and Y directions according to an analog signal. When the beam hits the screen, an element of phosphor briefly lights up. (b) A special scanning pattern was used to draw out each video frame, line by line.

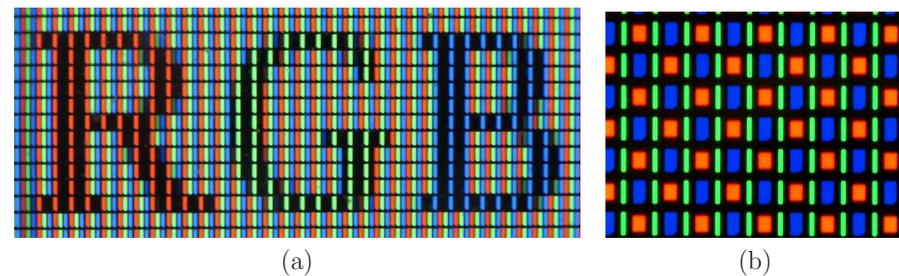


Figure 4.36: In displays, the pixels break into subpixels, much in the same way that photoreceptors break into red, blue, and green components. (a) An LCD display. (Photo by Luis Flavio Loureiro dos Santos.) (b) An AMOLED PenTile display from the Nexus One smartphone. (Photo by Matthew Rollings.)

the whole screen. Currently, the vast majority of flat-panel displays are on either LCDs or light emitting diodes (LEDs). In the case of LEDs, each pixel is able to be directly lit. The consumer market for flat-panel displays was first driven by the need for flat, big-screen televisions and computer monitors. With the advancement of smartphones, miniaturized versions of these displays have been available with low cost, low power, and extremely high resolution. This enabled low-cost VR headset solutions by putting a lens in front of a smartphone screen, as was shown in Figure 4.30.

Toward custom VR displays The first step toward thinking about displays for VR is to consider the distance from the eyes. If it is meant to be viewed from far away, then it is called a *naked-eye display*. For a person with normal vision (or while wearing prescription glasses), the display should appear sharp without any addition help. If it is close enough so that lenses are needed to bring it into focus, then it is called a *near-eye display*. This is the common case in current VR headsets because the display needs to be placed very close to the eyes. It remains an active area of research to develop better near-eye display technologies, with a key challenge being whether the solutions are manufacturable on a large scale.

An important family of near-eye displays is based on a *microdisplay* and *waveguide*. The microdisplay is typically based on *liquid crystal on silicon* (or *LCoS*), which is a critical component in overhead projectors; microdisplays based on organic LEDs (*OLEDs*) are also gaining popularity. The size of the microdisplay is typically a few millimeters, and its emitted light is transported to the eyes through the use of reflective structures called a *waveguide*; see Figure 4.37. The Microsoft HoloLens, Google Glass, and Magic Leap One are some well-known devices that were based on waveguides. The current engineering challenges are limited field of view, overall weight, difficult or costly manufacturing, and power loss and picture degradation as the waves travel through the waveguide.

A promising device for future VR display technologies is the *virtual retinal display* [19]. It works by a scanning beam principle similar to the CRT, but instead draws the image directly onto the human retina; see Figure ?? . A low-power laser can be pointed into a micromirror that can be rapidly rotated so that full images are quickly drawn onto the retina. Current engineering challenges are eye safety (do not shine an ordinary laser into your eyes!), mirror rotation frequency, and expanding the so-called *eye box* so that the images are drawn onto the retina regardless of where the eye is rotated.

To maximize human comfort, a display should ideally reproduce the conditions that occur from the propagation of light in a natural environment, which would allow the eyes to focus on objects at various distances in the usual way. The previously mentioned displays are known to cause vergence-accommodation mismatch (see Section 5.4), which is known to cause discomfort to human viewers. For this reason, researchers are actively prototyping displays that overcome this limitation. Two categories of research are *light-field displays* [4, 10, 12] and

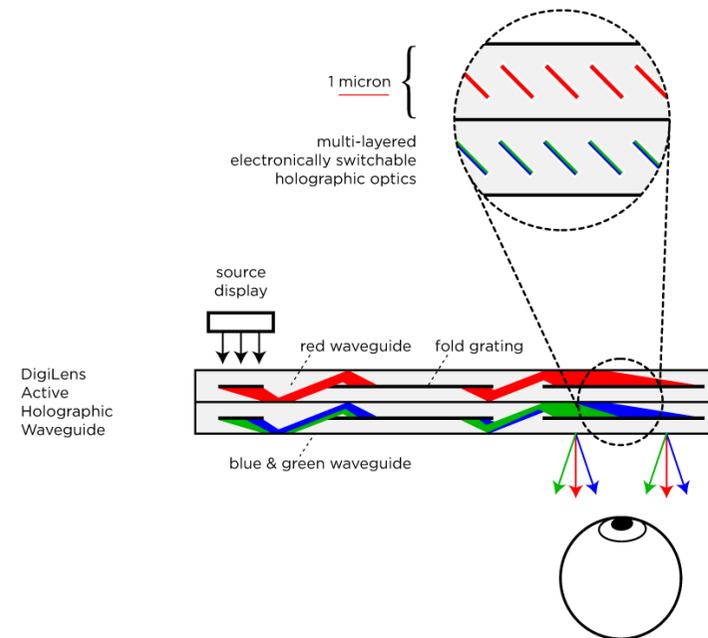


Figure 4.37: An illustration of how a DigiLens waveguide operates, as light is propagated from a small source display to the human eye. (Figure by Christopher Grayson; uploadvr.com/waveguides-smartglasses/)

varifocal displays [1, 3, 8, 11, 13].

Further Reading

Most of the basic lens and optical system concepts are covered in introductory university physics texts. For more advanced coverage, especially lens aberrations, see the classic optical engineering text: [17]. A convenient guide that quickly covers the geometry of optics is [5]. Thorough coverage of optical systems that utilize electronics, lasers, and MEMS, is given in [9]. This provides a basis for understanding next-generation visual display technologies. An excellent book that considers the human eye in combination with engineered optical components is [15]. Cameras are covered from many different perspectives, including computer vision [6, 18], camera engineering [7], and photography [16]. Mathematical foundations of imaging are thoroughly covered in [2].

Bibliography

- [1] K. Akeley, S. J. Watt, A. Reza Girschick, and M. S. Banks. A stereo display prototype with multiple focal distances. *ACM Transactions on Graphics*, 23(3), 2004.
- [2] H. H. Barrett and K. J. Myers. *Foundations of Image Science*. Wiley, Hoboken, NJ, 2004.
- [3] W. Cui and L. Gao. Optical mapping near-eye three-dimensional display with correct focus cues. *Optics Letters*, 42(13):2475–2478, 2017.
- [4] D. Fattal, Z. Peng, T. Tran, S. Vo, M. Fiorentino, J. Brug, and R. G. Beausoleil. A multi-directional backlight for a wide-angle, glasses-free three-dimensional display. *Nature*, 495:348–351, 2013.
- [5] J. E. Greivenkamp. *Field Guide to Geometrical Optics*. SPIE Press, Bellingham, WA, 2004.
- [6] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision, 2nd Ed.* Cambridge University Press, Cambridge, U.K., 2004.
- [7] G. C. Holst and T. S. Lomheim. *CMOS/CCD Sensors and Camera Systems*. SPIE Press, Bellingham, WA, 2011.
- [8] X. Hu and H. Hua. Design and assessment of a depth-fused multi-focal-plane display prototype. *Journal of Display Technology*, 10(4):308–316, 2014.
- [9] B. C. Kress and P. Meyrueis. *Applied Digital Optics: From Micro-optics to Nanophotonics*. Wiley, Hoboken, NJ, 2009.
- [10] D. Lanman and D. Luebke. Near-eye light field displays. *ACM Transactions on Graphics*, 32(6), 2013.
- [11] G. D. Love, D. M. Hoffman, P. J. H. Hands, J. Gao, A. K. Kirby, and M. S. Banks. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics Express*, 17(18):15716–15725, 2009.
- [12] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs. Pinlight displays: Wide field of view augmented-reality eyeglasses using defocused point light sources. *ACM Transactions on Graphics*, 33(4), 2014.
- [13] N. Matsuda, A. Fix, and D. Lanman. Focal surface displays. *ACM Transactions on Graphics*, 36(4), 2017.
- [14] P. Signell. Predicting and specifying the perceived colors of reflective objects. Technical Report MISN-0-270, Michigan State University, East Lansing, MI, 2000. Available at <http://www.physnet.org/>.
- [15] G. Smith and D. A. Atchison. *The Eye and Visual Optical Instruments*. Cambridge University Press, Cambridge, U.K., 1997.
- [16] R. Sawdon Smith and A. Fox. *Langford's Basic Photography, 10th Ed.* Focal Press, Oxford, U.K., 2016.
- [17] W. J. Smith. *Modern Optical Engineering, 4th Ed.* SPIE Press, Bellingham, WA, 2008.
- [18] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag, Berlin, 2010.
- [19] E. S. Viirre, H. Pryor, S. Nagata, and T. Furness. The virtual retinal display: A new technology for virtual reality and augmented vision in medicine. *Studies in Health Technology and Informatics*, 50:252–257, 1998.
- [20] F. E. Wright. *The Methods of Petrographic-microscopic Research*. Carnegie Institution of Washington, 1911.