Image Formation

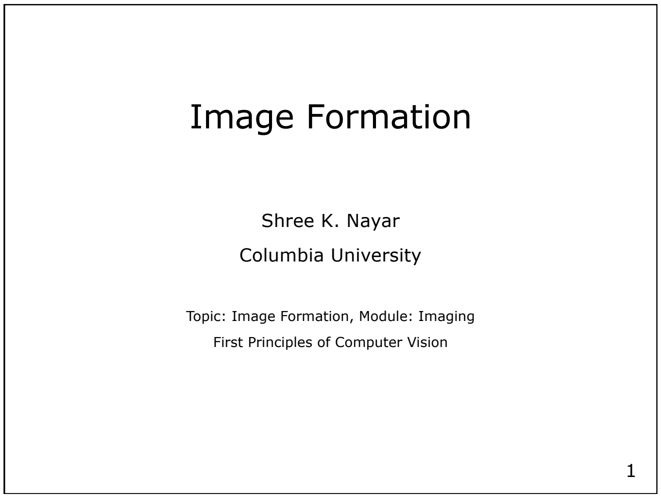
Shree K. Nayar

Monograph: FPCV-1-1
Module: Imaging
Series: First Principles of Computer Vision

Computer Science, Columbia University

February 15, 2022

FPCV Channel
FPCV Website



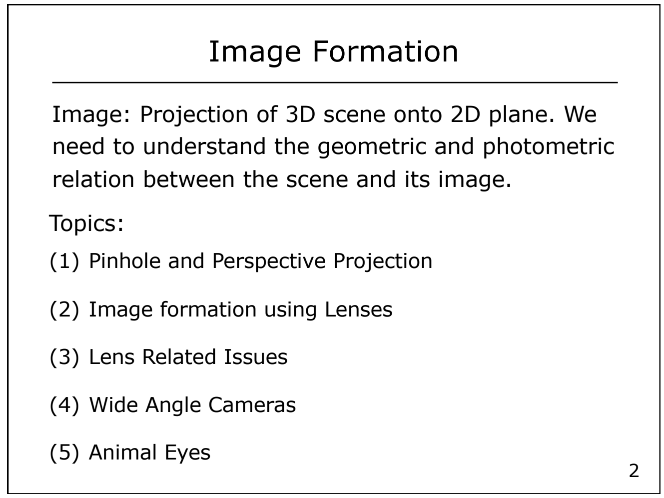


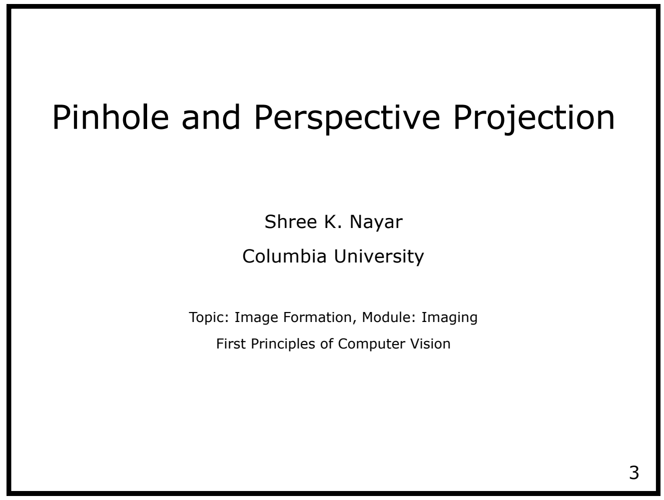
Image formation is the projection of a three-dimensional (3D) scene onto a two-dimensional (2D) plane. We seek to understand the geometric and photometric relation between the scene and its image. Given a point in the scene, the geometric relation tells us where it ends up in the image. Given the brightness of a point in the scene, the photometric relation determines what its brightness will be in the image.

We will begin with the concept of a pinhole camera. This is the simplest type of camera imaginable, and it has a very long history. The pinhole camera performs what is called perspective projection — this is one of the most important concepts in computer vision. We will derive the equations of perspective projection and discuss some of its visual manifestations. Then, we will show that, while the pinhole camera is great in terms of the clarity of images it produces, it simply does not gather enough light.

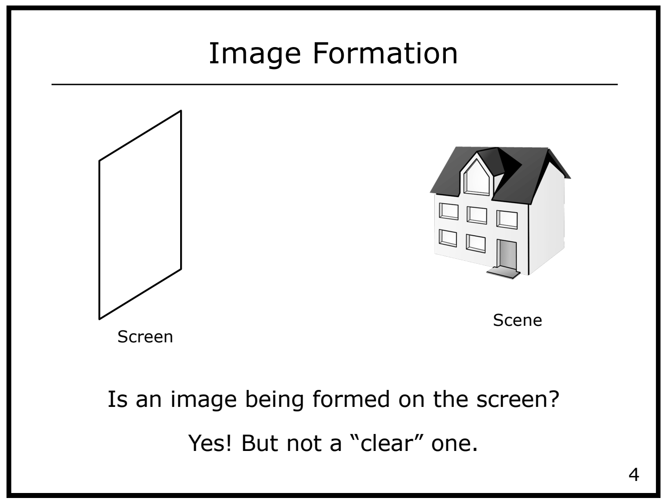
To address this limitation, we use lenses. We will discuss image formation using lenses in detail and describe various characteristics of a lens, including its focal length, defocus blur, f-number and depth of field. Next, we will describe various issues related to lenses. Even if a lens is perfectly manufactured, it can produce several undesirable effects in the image. We will talk about what these effects are and how we might be able to correct them.

Then, we will digress from perspective projection and look at the problem of imaging unusually large fields of view. A hemispherical field of view, for instance, cannot be captured using perspective projection. Instead, this can be done using a special type of lens called a fisheye lens, or by using a combination of mirrors and lenses. Finally, we will take a quick look at biological eyes. We will describe some fascinating designs that nature has come up with and then focus on the human eye and its remarkable characteristics.

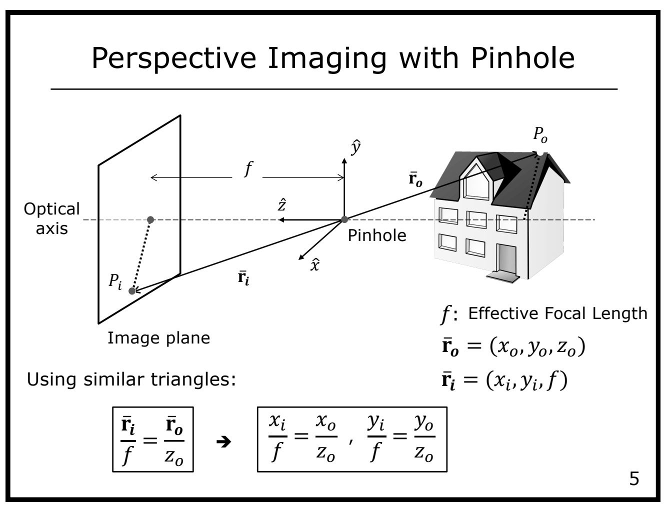
First, let us discuss the pinhole camera and the equations of perspective projection.



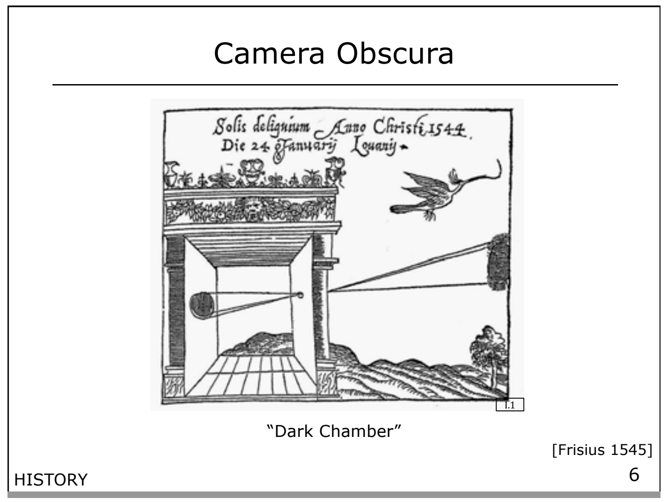
Here is a 3D scene with a house on the right and a 2D screen on the left. Is an image of the house being formed on the screen? If we consider any point on the screen, it does indeed receive light from many points on the house. So, in a sense, we may say that an image of the house is being formed on the screen. However, since each point on the screen receives light from many points on the house, the end result is a muddled or blurred image and not a clear one.



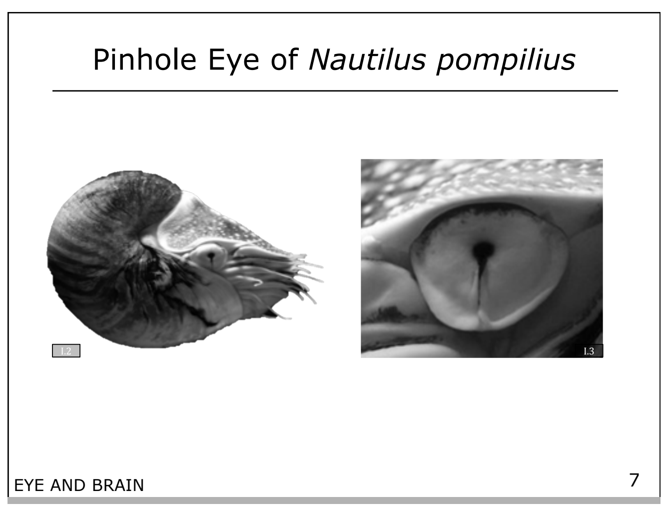
In order to create a clear, crisp image, we can place a pinhole—an opaque sheet with a tiny hole in it—between the scene and the image plane. We see that each point in the scene now projects onto a single point in the image. For instance, there is a single ray that travels from P_o on the house to P_i on the image plane. In order to understand the relationship between P_o and P_i, we erect a 3D coordinate frame and place it at the pinhole, with the z-axis ($\hat{z}$) pointing along the optical axis towards the image plane. The optical axis is perpendicular to the image plane and is shown here as a dotted line. The distance between the pinhole and the image plane is called the effective focal length, f.



We can express the point P_o as the vector $\bar{\mathbf{r}}_{\mathbf{o}}$ with coordinates x_o, y_o, and z_o. Its image can be denoted by the vector $\bar{\mathbf{r}}_{\mathbf{i}}$, which has coordinates x_i, y_i, and z_i, where z_i equals f irrespective of the location of P_o in the scene. Using similar triangles, we can write that the vector $\bar{\mathbf{r}}_{\mathbf{i}}$ divided by f is equal to the vector $\bar{\mathbf{r}}_{\mathbf{0}}$ divided by z_o. Note that z_o is the depth of the point P_o in 3D space. Since $\bar{\mathbf{r}}_{\mathbf{i}}$ and $\bar{\mathbf{r}}_{\mathbf{o}}$ are vectors, we can break them down into their components. As a result, we have x_i divided by f is equal to x_o divided by z_o, and y_i divided by f is equal to y_o divided by z_o. These are the equations of perspective projection. They are very simple equations, but nevertheless they produce several non-intuitive effects in images.

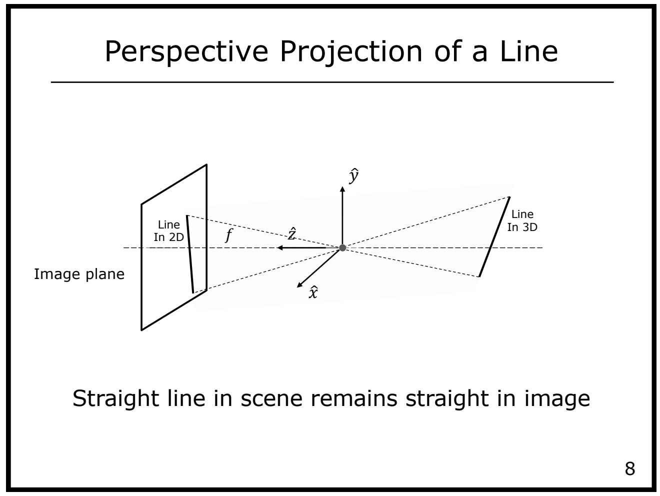


The idea of the pinhole camera dates back to 500 BC when Chinese philosophers were writing about the concept. Around 1000 AD, the Arab physicist Ibn Al-Haytham (known as Alhazen in the West) gave a detailed description of the pinhole camera in *Kitab al-Manazir* (Book of Optics), one of the first optics books. In the 16th century, the concept came to the West and became popular among artists. It was used by artists as a tool for rendering accurate depictions of 3D scenes. The sketch shown here is by the Dutch mathematician Gemma Frisius. The pinhole in one wall projects an image of the scene onto the second wall, enabling an artist to walk up to the wall and sketch a geometrically accurate representation of the scene. In the West, the pinhole camera was called *camera obscura*, which in Latin means "dark chamber."

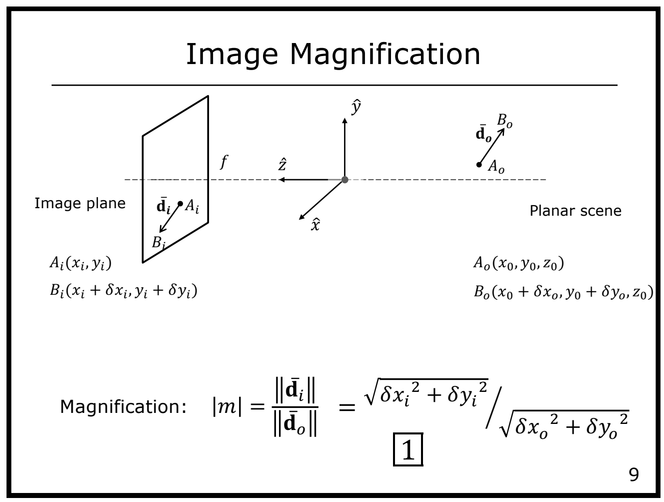


As with many scientific advancements, nature predates us in the creation of pinhole cameras. Shown here is a shelled sea creature called *Nautilus pompilius*. Its eye does not have a lens – it uses a large pinhole to create an image.

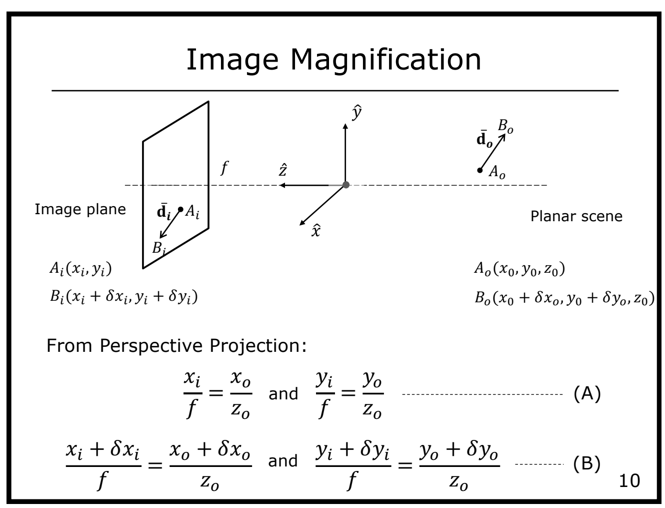
Now, let us discuss the perspective projection of a line in 3D space on to the image plane. We know that a line in 3D and the pinhole (a point) define a plane. All the rays of light that emanate from the line and pass through the pinhole also lie on that same plane. Thus, the image of the 3D line lies on the intersection of this plane and the image plane. In other words, the image of a line in the 3D scene has to be a line on the 2D image plane. That is why, as you may have noticed, straight lines in scenes map to straight lines in photographs.



Next, let us look at image magnification. Consider the segment from A_o to B_o in the scene given by the vector $\bar{\mathbf{d}}_\mathbf{o}$. A_o has coordinates x_o, y_o, and z_o, while B_0 has the same coordinates displaced by δx_o and δy_o. This segment lies on a plane in the scene which is parallel to the image plane. Its image is the vector $\bar{\mathbf{d}}_\mathbf{i}$. The ratio of the magnitude of $\bar{\mathbf{d}}_\mathbf{i}$ to that of $\bar{\mathbf{d}}_\mathbf{o}$ is called the magnification. This can be written in terms of the displacements in the scene and the image $\boxed{1}$.



We can substitute for the displacements using the equations of perspective projection that were discussed earlier. By applying perspective projection to the scene points A_o and B_o, we get the four equations given by (A) and (B).



Using these four equations, we end up with two expressions for the relationship between displacements in the image and displacements in the scene $\boxed{1}$. By substituting these expressions in the equation for magnification, we end up with a simple expression which reveals that the magnification m is simply the effective focal length divided by z_o – the distance (or "depth") of the scene plane on which the segment $\bar{\mathbf{d}}_{\mathbf{o}}$ lies from the pinhole. In other words, an object's magnification is inversely proportional to its distance from the camera. The sign of m will be positive if the image is upright and negative if the image is inverted. In the case of a pinhole camera, the image will be inverted, and hence m will be negative.

Image Magnification

From (A) and (B) we get:

$$\frac{\delta x_i}{f} = \frac{\delta x_o}{z_o} \quad \text{and} \quad \frac{\delta y_i}{f} = \frac{\delta y_o}{z_o} \qquad \boxed{1}$$

Magnification:

$$|m| = \frac{\|\bar{\mathbf{d}}_i\|}{\|\bar{\mathbf{d}}_o\|} = \sqrt{\delta x_i^{\,2} + \delta y_i^{\,2}} \Big/ \sqrt{\delta x_o^{\,2} + \delta y_o^{\,2}} = \left|\frac{f}{z_o}\right|$$

$\boxed{m = \frac{f}{z_o}}$ m is negative when image is inverted

11

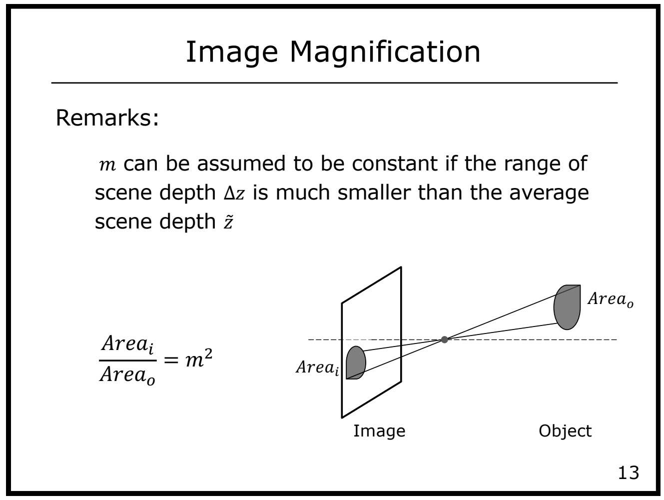
Now, let us take a look at some manifestations of image magnification. On the left is a pair of train tracks. Although the tracks are parallel in the scene, in the image they appear to meet as the distance between the tracks and the camera increases. This is a result of the inverse relationship between magnification and depth. Photographers often make use of this property. In the photo on the right, the man and woman are roughly the same height, yet the man looks small enough to be standing on the woman's palm.

Image Magnification

$\boxed{m = \frac{f}{z_o}}$ Image size inversely proportional to depth

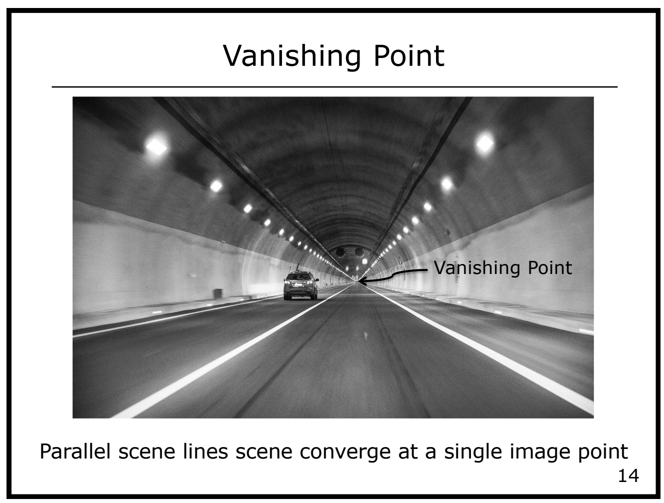
12

We can now make a couple of observations related to image magnification. First, if the size of an object is small compared to its distance from the camera, then we can assume that the entire object is subjected to the same magnification. Conversely, when the size of an object is significant compared to its distance from the camera, different parts of the object will be subjected to different magnifications. In the case of selfies, where the size of a person's head is significant compared to the distance of their head from the camera, the nose is more magnified than other parts of the head as it is closest to the camera. As a result, noses tend to appear disproportionately large in selfies!

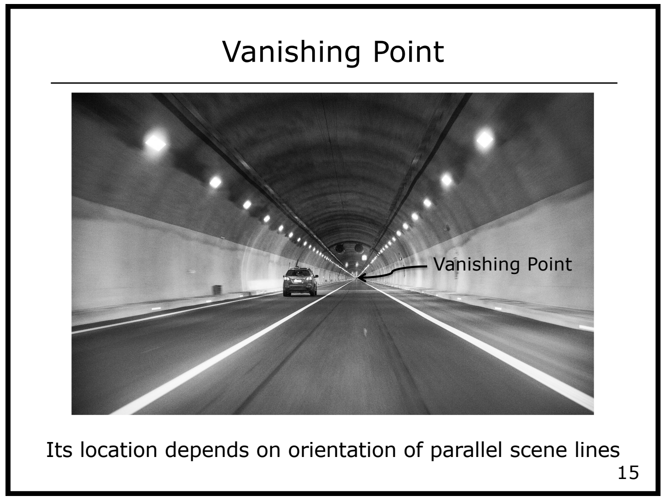


Finally, if the image magnification of a linear segment in the scene is *m*, we know that the magnification of an area in the scene would be m^2 as an area is, by definition, the product of two linear dimensions.

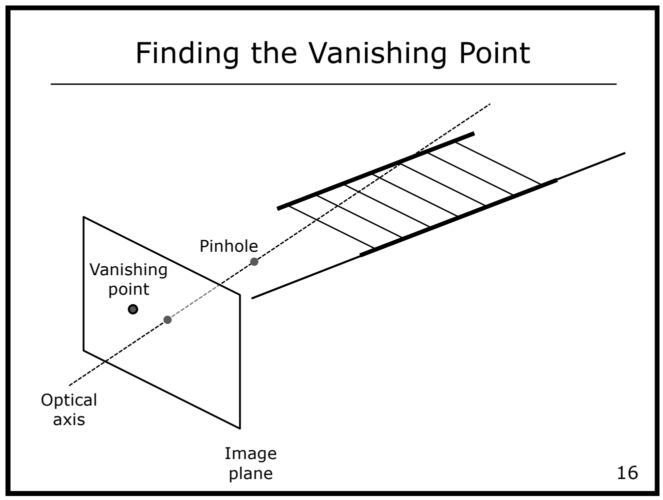
There is an interesting phenomenon produced by perspective projection called the vanishing point. In this photo, the tunnel is straight, and therefore the white lane lines and the lines on the side walls are all parallel to each other in the 3D scene. Yet, in the image they all seem to be emerging from a single point. We call this point the vanishing point. It is a point in the image where a set of any number of parallel lines in 3D appear to disappear.



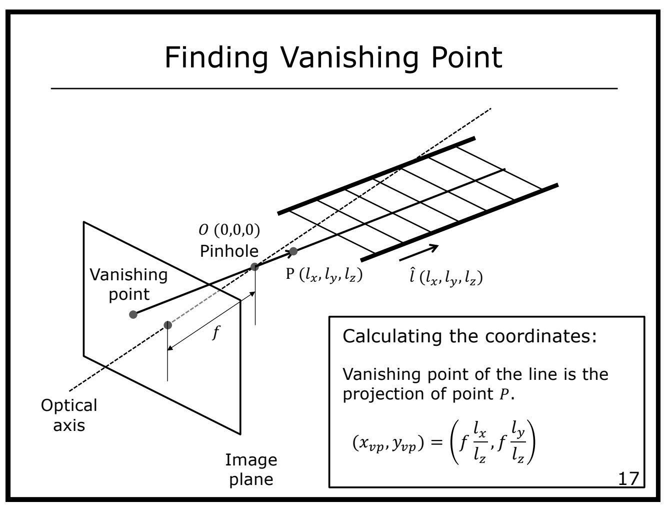
The exact location of the vanishing point in the image depends on the orientation of the parallel lines in 3D. Let us now look at how we might find the location of the vanishing point given a set of parallel lines.



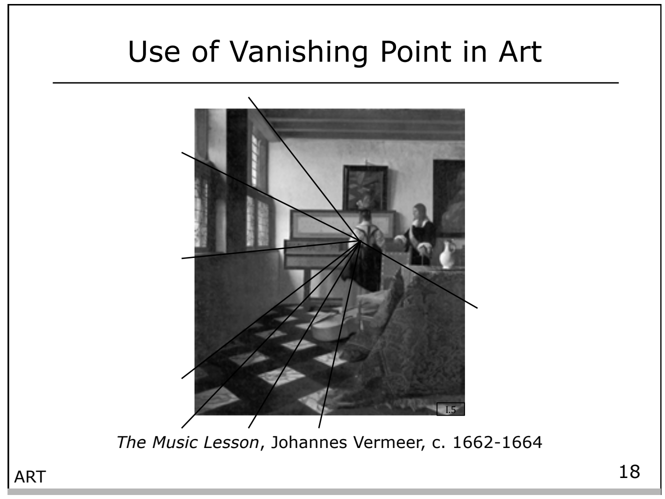
Consider the two parallel tracks shown here. Since they are parallel, they will produce a single vanishing point in the image. To visualize where this point would appear in the image, consider a line that is both parallel to the tracks and passes through the pinhole. We know that this line also will share the same vanishing point as the two tracks. Therefore, the point where this line pierces the image is the location of the vanishing point.



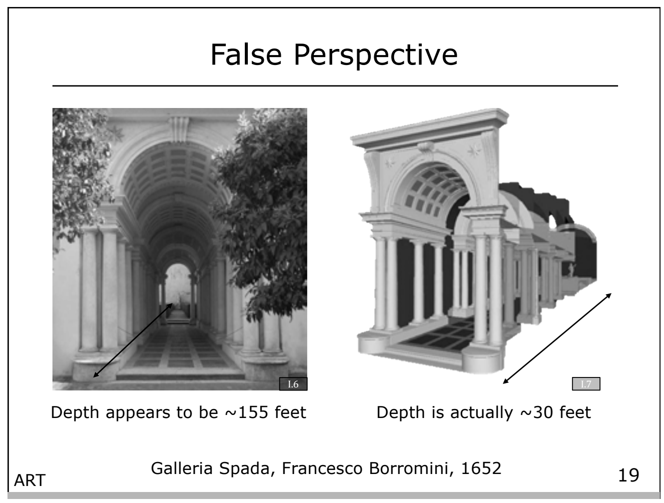
In order to find the coordinates of this point, let us define the direction of the parallel lines using the vector $\langle l_x, l_y, l_z \rangle$. Note that the origin of our coordinate frame lies at the pinhole. Therefore, the vector $\langle l_x, l_y, l_z \rangle$ corresponds to a point P that lies on the line that is parallel to the tracks and passes through the pinhole. We simply project the point P onto the image using our perspective projection equations. The resulting image point (x_{vp}, y_{vp}) is the vanishing point for the tracks.



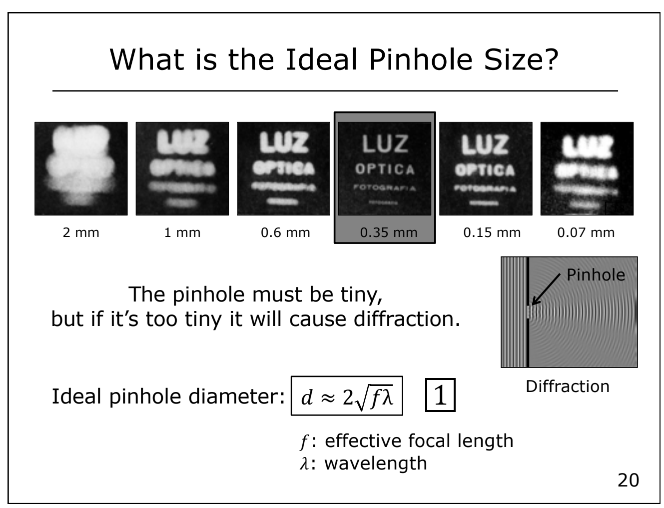
Artists have used the vanishing point to accentuate the most important subject or activity in their works. For example, consider *The Music Lesson* by the Dutch artist Johannes Vermeer. There are many sets of parallel lines in this painting, and overlaid on the painting are the lines that belong to one of these sets. Note that the vanishing point corresponding to this set of lines is located at the elbow of the piano player. The viewer's eye is naturally drawn to that region of the painting, emphasizing its centrality.



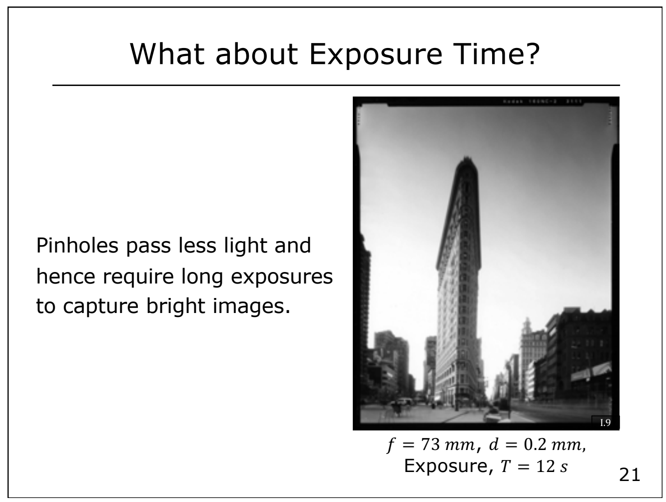
Here is another interesting effect related to perspective projection called false perspective. This is a gallery in Rome created by Francesco Borromini called Galleria Spada. Standing at the end of the hallway, one gets the impression that the sculpture at the far end is roughly 150 feet away. In reality, it is only 30 feet away from the viewer! This illusion is created by the tapered archway, in which the pillars become smaller with distance from the observer.



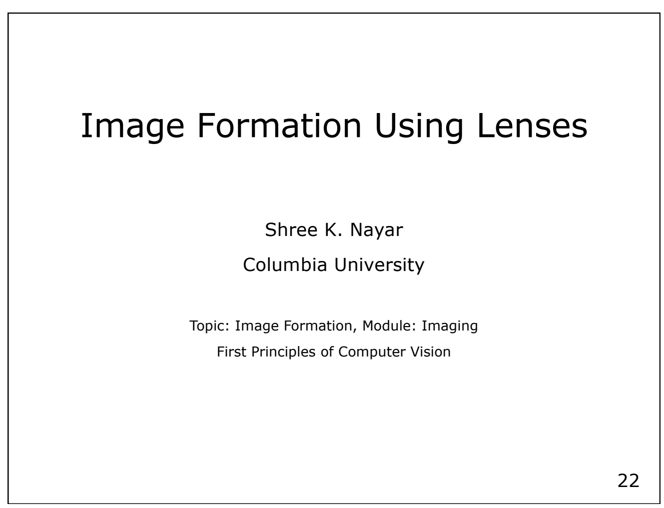
Returning to the pinhole camera, let us discuss the optimal size of the pinhole. At first, it would seem that the pinhole should be as tiny as possible, since then each ray that passes through it will be thin and well defined. Well, there is a bit more to the story. Shown here are images captured using different sizes of pinholes, the largest on the left and the smallest on the right. Note that the ideal size of the pinhole actually lies somewhere in the middle. When the pinhole is too large, it lets through a bundle of rays of light from each point in the scene, which results in a disk in the image over



which light is distributed. As a result, the image is blurred. As we would expect, as the pinhole size is reduced, the image sharpens. However, after a point, further reduction of the pinhole diameter causes the image to get blurry again. This is due to an effect in wave optics called diffraction. When light passes through an opening, the light waves are bent at the periphery of the opening. As the opening becomes smaller, the effect of the bending becomes more pronounced. In order to avoid this effect, the diameter of the pinhole should be roughly two times the square root of the product of the effective focal length and the wavelength of light [1]. In the case of visible light images, the wavelength would lie between 400 nanometers and 700 nanometers. Choosing the average, 550 nanometers, we can plug that in along with our focal length to find the optimal pinhole size.

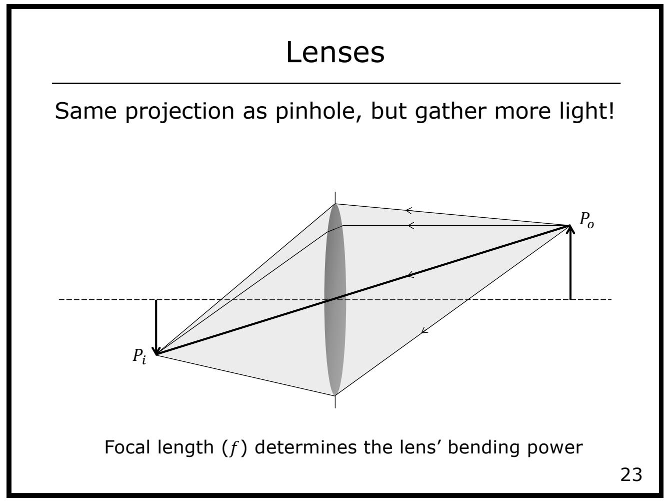


Using the optimal pinhole size, we can produce stunning images, such as this one of the Flatiron Building in New York City. The photographer used an effective focal length of 73 millimeters and a pinhole of 0.2 millimeters. Notice that the image is well-focused for the entire three-dimensional scene. This is what we sometimes refer to as an "all-focused" image. However, the downside here is that since pinholes capture very little light, the exposure times tend to be very long. In this case, the exposure time was 12 seconds, which is unacceptable for virtually any application of computer vision. That is where lenses come into play.



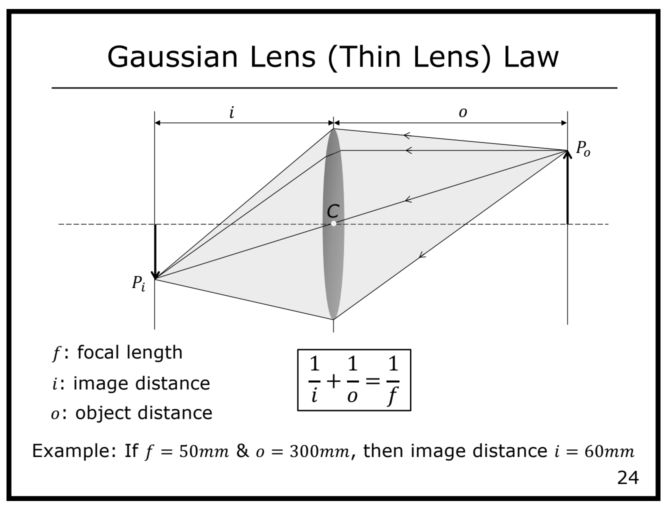
Let us take a look at how we can form an image using a lens. The lens performs perspective projection just as the pinhole does, but it is able to gather significantly more light.

As seen here, all of the rays of light received by the lens from the point P_o are refracted, or bent, to converge at the point P_i. That is, P_i is where P_o is focused behind the lens. The bending power of the lens is determined by its focal length, f.

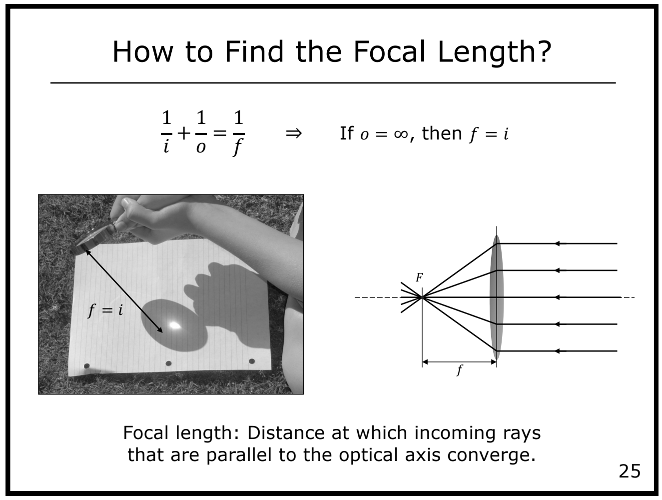


In the case of a pinhole camera, the only ray of light that would make it through the pinhole would be the bold line which passes through the center of the lens. Note that in the case of a lens, the perspective projection model remains the same, with the center of the lens playing the role of the pinhole. However, in contrast to the pinhole camera, the lens is able to gather a significantly greater amount of light.

Let us determine the relationship between the position of P_o in the scene and its image P_i. This is given by the Gaussian lens law. We denote the distance from P_o to the lens as o, and the distance from the lens to P_i as i. The Gaussian lens law states that 1 divided by i plus 1 divided by o is equal to 1 divided by f, where f is the focal length of the lens. Note that any point on a plane that is at a distance o from the lens will be focused on the image plane at a distance i behind the lens. Therefore, in any lens system there is a single plane in the scene that is perfectly focused on the image. We refer to that as the plane of focus. As an example, if a lens has a focal length of 50 millimeters and there is an object at a distance of 300 millimeters from the lens, then the image of the object will be formed at a distance of 60 millimeters behind the lens.

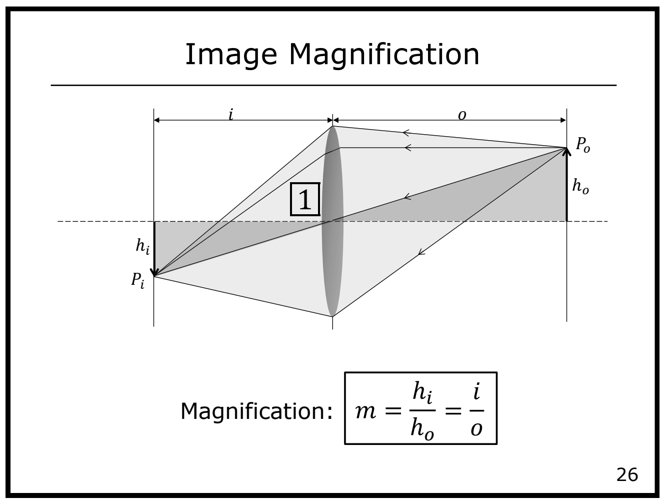


Given a lens, it is easy to find its focal length using the Gaussian Lens Law. If we use a point source in the scene that is really far from the lens (o equal to infinity), then we see that f will be equal to i. Thus, with a distant light source, such as a street lamp, we can measure the distance between where its image is formed (say, on a sheet of paper) and the lens — that is the focal length f.

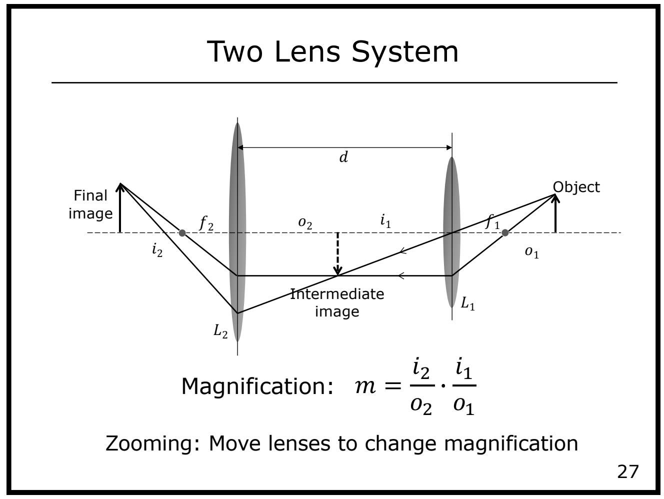


The focal length f of the lens is governed by two main factors. The first is the refractive index of the material the lens is made of. Lenses are typically made of glass or plastic, and in both cases the refractive index is significantly greater than that of air. The second factor that impacts the focal length of a lens is its shape. One or both of the surfaces of the lens are curved, and the radii of curvature of these surfaces affects the focal length.

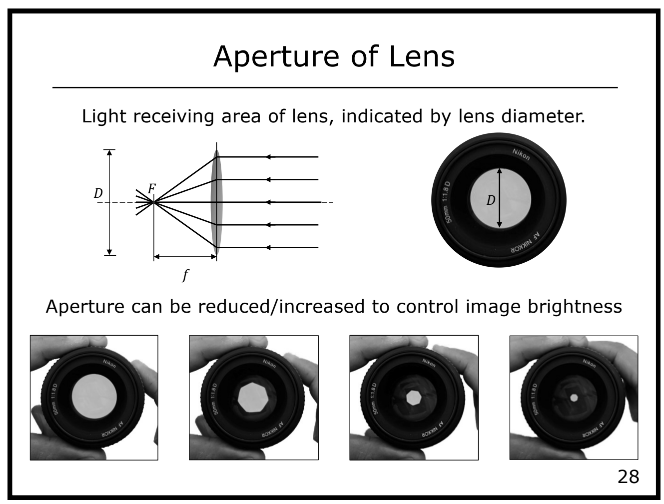
Now let us examine the magnification of a lens. Once again, the object distance is o and the image distance is i. Assume the height of the object to be h_o and the height of its image to be h_i. The magnification is defined as the image height divided by the object height. The two shaded triangles that share a vertex at the center of the lens are similar triangles $\boxed{1}$. Using these triangles, we can write that h_i divided by h_o is equal to i divided by o.



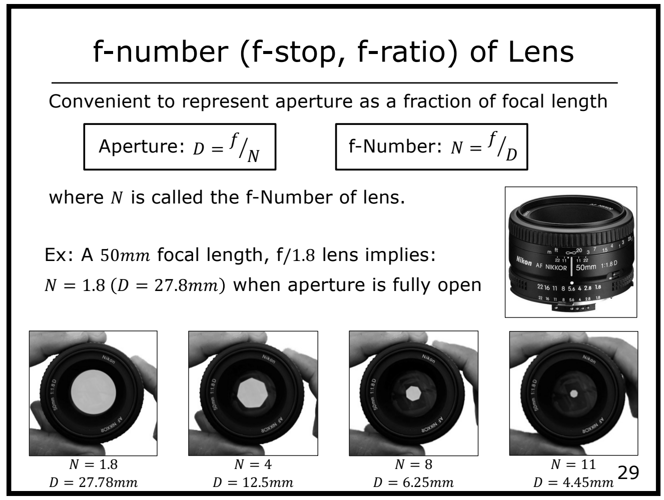
One can change the magnification of a lens camera by using multiple lenses. Consider this two-lens system, consisting of lenses L_1 and L_2, and an object at distance o_1 from L_1. The object is focused by lens L_1 to create an intermediate image between the two lenses. This intermediate image can be viewed as a new "object" that is then imaged by L_2 to obtain the final image. Thus, the magnification of the complete system is the magnification due to lens L_1 multiplied by the magnification due to lens L_2. That is, i_1 divided by o_1 multiplied by i_2 divided by o_2. Without changing the distance between the object and the image plane, the magnification of the entire system can be varied by shifting the positions of the two lenses. This process is referred to as "zooming."



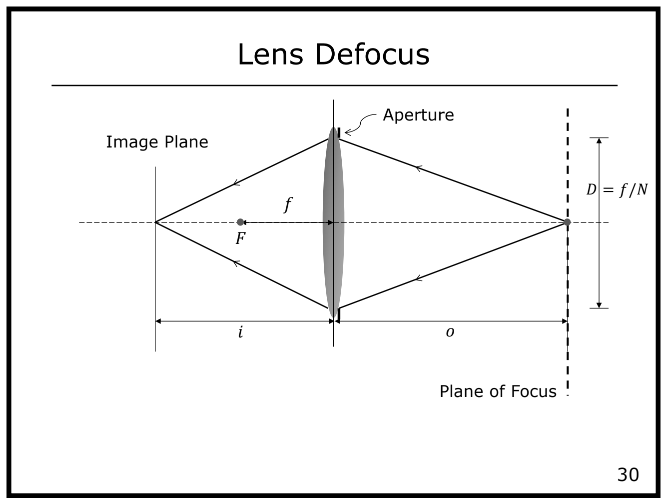
The aperture of the lens is the clear area of the lens that gathers light from points in the scene. In most lenses, the aperture is more or less a circular disk which we can represent using its diameter D. The aperture is generally implemented using a diaphragm, which regulates the passage of light. At the bottom, we see different settings of the aperture, going from fully open to a small hole in which case the lens essentially functions like a pinhole camera.



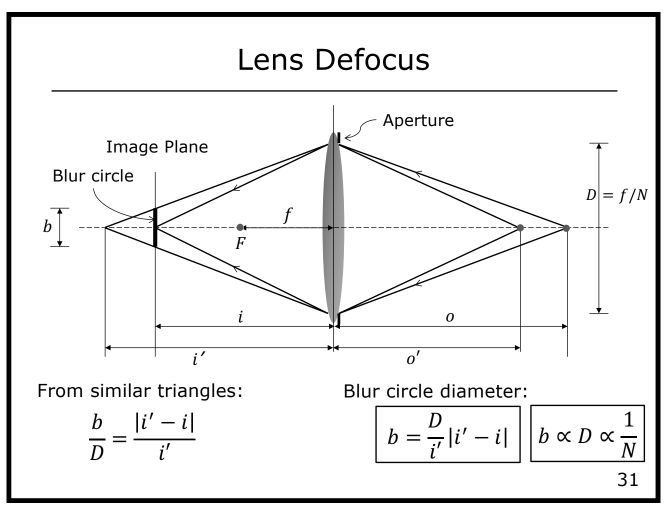
The ratio of the focal length to the diameter of the aperture is called the f-number. Thus, the diameter of the aperture can be defined as f divided by N, where f is the focal length and N is the f-number of the lens. For example, given a lens with a focal length of 15 millimeters and an f-number of 1.8, the diameter would be 27.8 millimeters. As shown at the bottom, as the aperture of a lens goes from open to closed, the f-number increases while the diameter decreases.



Using lenses comes at a price — there is only one plane in the scene that is perfectly focused onto the image plane by a lens. Consider a point at object distance o that is focused at distance i behind the lens. Let us place our image plane (where we record the image) at the distance i behind the lens such that it is parallel to the plane that the lens lies on. Then, only points on the plane that is at a distance o in front of the lens and parallel to the lens plane will be focused on the image plane. As mentioned before, this plane in the scene is called the plane of focus. Any object that lies outside of this plane will be out of focus in the image.

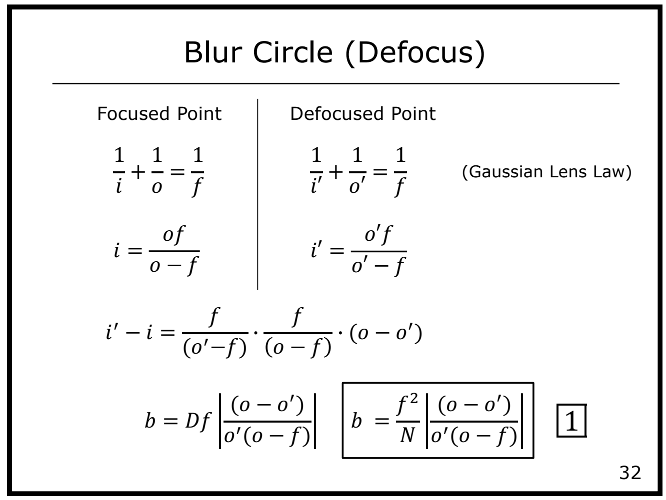


In this diagram, consider the point at a distance *o'* from the lens. Since it is closer to the lens than the focused point *o*, its image will be formed behind the image plane. In other words, the light that the lens receives from this point will not correspond to a single point on the image plane, but instead it will be distributed over a circular disk, which is referred to as the "blur circle." Let us use *b* to represent the diameter of the blur circle. To find the relationship between *b* and the object at position *o*, observe that the triangle with *b* as its base and the larger triangle with *D* as its base are similar triangles.

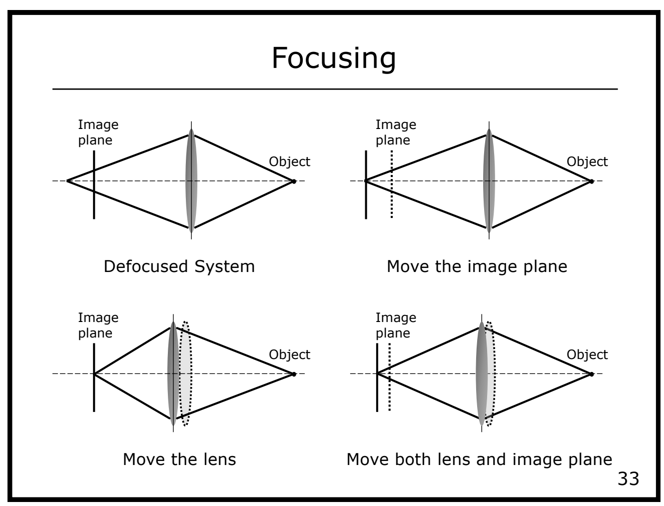


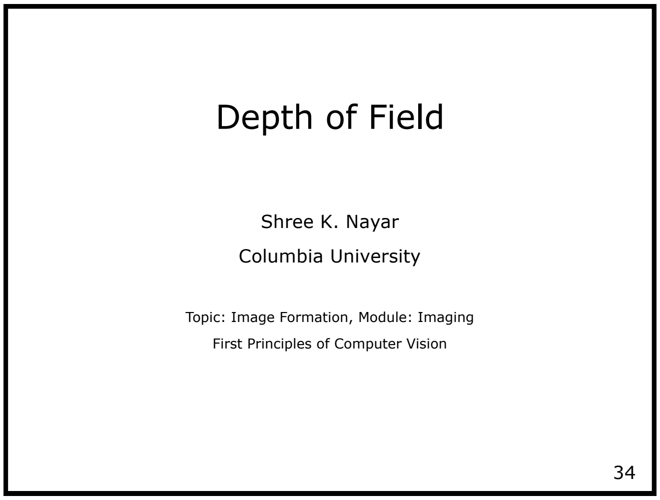
Thus, *b* divided by *D* is equal to |*i’*-*i*| divided by *i’*. We see that the blur circle diameter is proportional to the diameter of the aperture and therefore inversely proportional to the f-number of the lens. The larger the lens aperture, the more blur that will occur for points that lie outside the plane of focus.

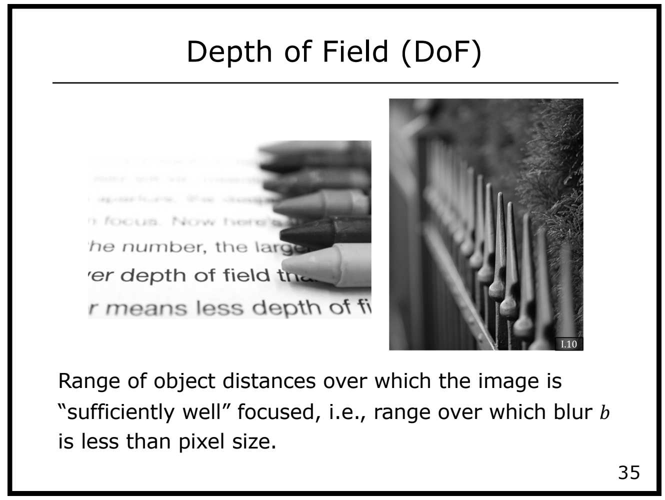
Now, we will express the blur circle diameter in terms of object distance in order to see what happens when we move an object away from the plane of focus in the scene. We use the Gaussian lens law to write the equations for the focused and defocused points. From these two expressions, we can obtain an expression for (*i’* - *i)*, in terms of (*o* - *o*’), and then substitute in our expression for the blur circle $\boxed{1}$. The diameter of the blur circle is proportional to *f* squared divided by *N*, where *f* is the focal length and *N* is the f-number of the lens. On the right side, the expression (*o* - *o*’) denotes the distance of the scene point from the plane of focus.



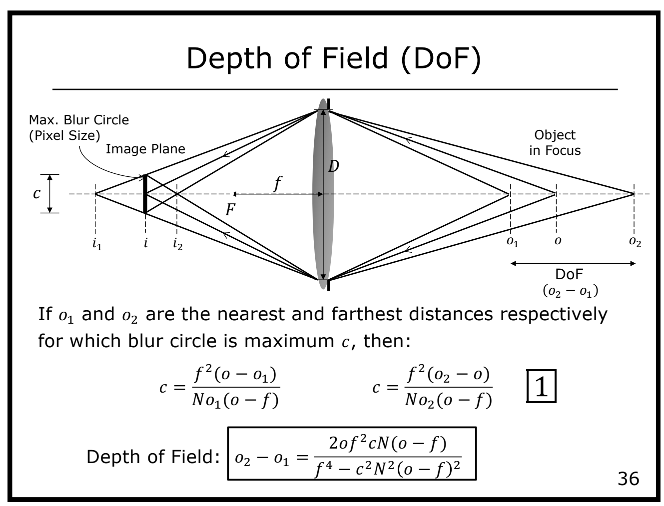
How do we focus an imaging system that is out of focus? In the top left corner is a defocused imaging system, in which the object is being imaged behind the image plane, creating a blur circle. To bring the object into focus, we can move the image plane away from the lens, move the lens itself, or move both the lens and the image plane. In each case, we are changing the object distance and/or the image distance in the Gaussian lens equation to bring the object into focus.





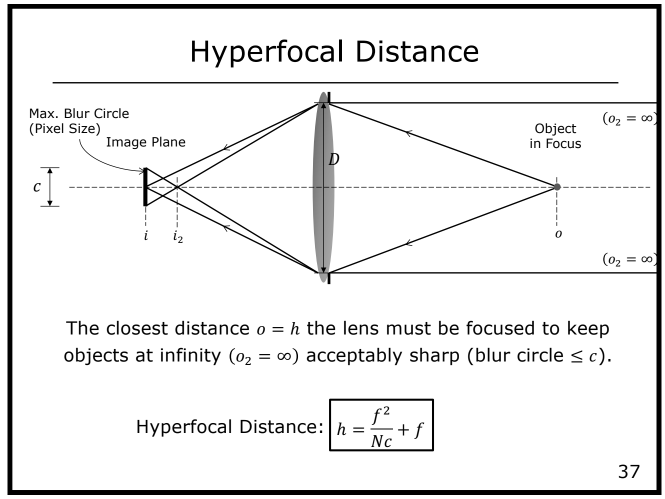


Since we understand defocus blur, we can now discuss the important concept of depth of field. For any given image plane (sensor) location, we know there is one plane in the scene that is perfectly focused — the plane of focus. The sensor used to record the image is made of pixels of finite size. As a result, all scene points with blur circles that are smaller than the size of a pixel are equally blurred (or equally focused) in the image. The range of object distances for which the image is equally well focused — that is, the range over which the blur circle diameter is less than the pixel size — is called the depth of field of the imaging system. Points become progressively more out of focus as they move away from the depth of field of the system. On the right are two images where the depths of field are easy to see.



Let us find the depth of field of a lens camera. To do so, we must first define the pixel size. Let us say that the width of each pixel is c. Now, we want to determine the range of distances of the object for which the blur circle is going to be smaller than c. Consider the point at distance o_1. This is the point for which the diameter of the blur circle exactly equals the width of the pixel c. In this case, because the point is closer to the lens than o, the image is going to be formed behind the image plane. There is another point at distance o_2, which is farther away from the lens than o, for which the image is going to be formed in front of the image plane. After the light rays converge at this point, however, they diverge again to create a blur circle of size c. We can find the depth of field by using our expression for the blur circle diameter in slide 32. We apply this equation to the scene points at distances o_1 and o_2, and for the blur diameter b we plug in c, the size of the pixel. This gives us these two equations $\boxed{1}$. Note that we know the focal length f, f-number N, pixel size c, and distance o of the plane of focus.

Using these two equations, we get an expression for the depth of field, which is o_2 minus o_1. Interestingly, the distance between o_2 and o is greater than the distance between o_1 and o. That is because, as the object approaches the lens, defocus increases more rapidly.

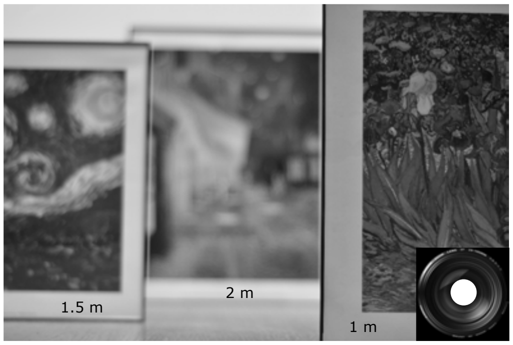


This brings us to the interesting concept of the hyperfocal distance *h*. It is the closest distance a lens needs to be focused at such that all points beyond that distance will be in focus (that is, within the depth of field of the camera). In other words, all points between *h* and infinity will appear focused in the image. In fact, since the plane of focus is at *h*, the image will also be focused between o_2 (which equals infinity) and o_1 which is closer to the lens than *h*. By setting o_2 equal to infinity in our expression for depth of field, we can obtain a simple expression for the hyperfocal distance *h*, where *h* is equal to f^2 divided by *Nc* plus *f*.

The hyperfocal distance is an important concept, because if an imaging system is designed to be focused on the hyperfocal distance, then we know that all points beyond *h* are going to be in focus. In the early days of smartphones, the cameras had fixed focal lengths and focus settings. The manufacturers would preset the focus at the hyperfocal distance *h* so that the user knew that when they took pictures, as long as the objects of interest lay beyond *h*, they would be focused in the image.

Remember that we are using a lens rather than a pinhole camera, because it creates brighter images. The price we pay when we use a lens is that it has a finite depth of field. So, let us take a look at the trade-off between the depth of field and the brightness of the image. In the scene shown below, there are three paintings — the first one at a distance of 1 meter from the camera, the second one at 1.5 meters, and the third one at 2 meters. We have used a lens with a focal length of 50 millimeters and a fully open aperture diameter of 25 millimeters, or an f-number of 2. The camera is focused on the painting at 1 meter. In the top left corner, the aperture is wide open. As expected, the painting at 1 meter is perfectly focused, but the one at 1.5 meters is out of focus, and the one at 2 meters is even more out of focus. Let us now increase the f-number, or, in other words, reduce the diameter of the aperture. Notice that the paintings that are out of focus get more in-focus, but the image gets darker because less light is being collected by the lens. When we go down to an f-number of 8, the image is even darker but sharper. With an f-number of 16, we are almost approaching a pinhole, resulting in an image that's nearly focused everywhere, except that it is very dark.

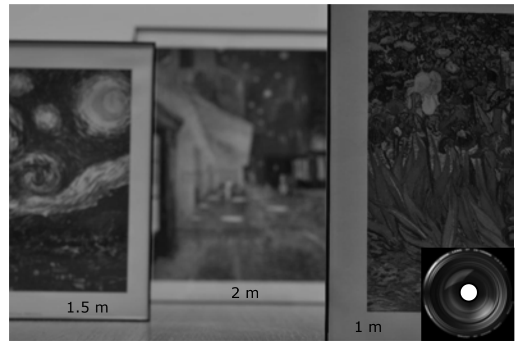
Aperture Size: DOF vs. Brightness



Focal Length 50 mm, Focus = 1 m, Aperture D = 25 mm, f-Number N = 2

38

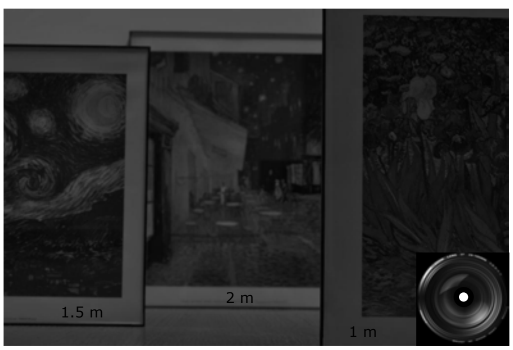
Aperture Size: DOF vs. Brightness



Focal Length 50 mm, Focus = 1 m, Aperture D = 12.5 mm, f-Number N = 4

39

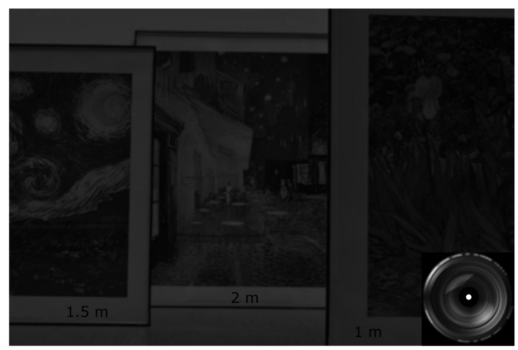
Aperture Size: DOF vs. Brightness



Focal Length 50 mm, Focus = 1 m, Aperture D = 6.25 mm, f-Number N = 8

40

Aperture Size: DOF vs. Brightness



Focal Length 50 mm, Focus = 1m, Aperture D = 3.125mm, f-Number N = 16

41

We can summarize this trade-off by saying that a large aperture (a small f-number), implies a bright image or, in a real-world application, a shorter exposure time. For a brighter image, or shorter exposure time, the price that we pay is that we have a shallower depth of field.

On the other hand, using a small aperture, which means a large f-number, results in a larger depth of field but a darker image or a longer exposure time.

Aperture Size: DOF vs. Brightness

Large Aperture (Small f-Number)

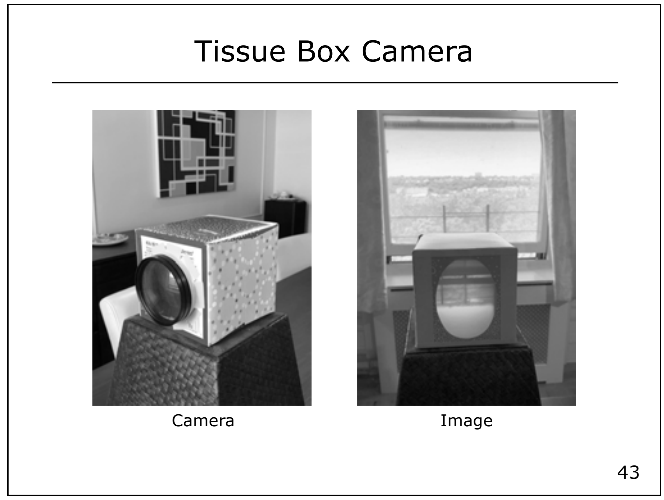
- Bright Image or Short Exposure Time
- Shallow Depth of Field

Small Aperture (Large f-Number)

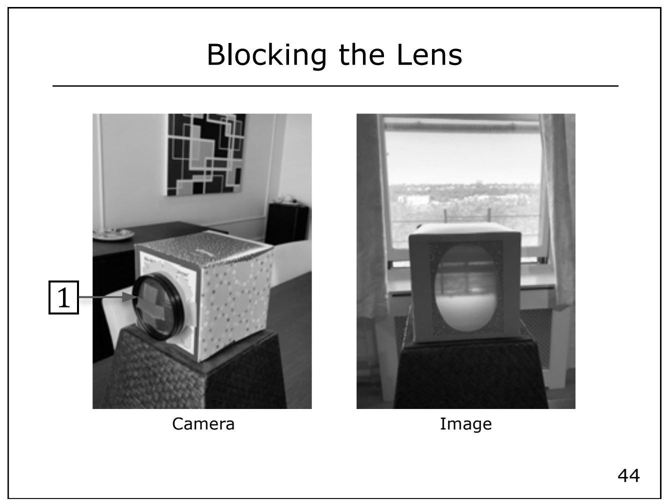
- Dark Image or Long Exposure Time
- Large Depth of Field

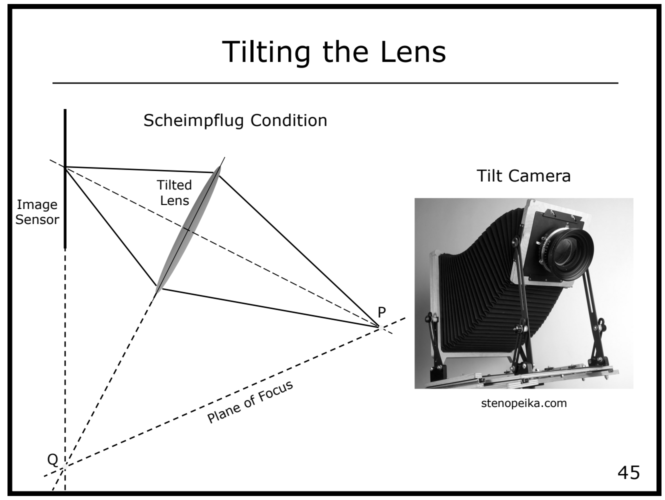
42

Now, let us demonstrate an important property of a lens with a tissue box camera—a tissue box with a disk cut out, where a lens has been attached. This lens is five diopters in power, which means it has a focal length of one over five meters or 200 millimeters. In the back, where the tissue box opening is located, I have attached a piece of translucent tracing paper to form the image. When this camera is pointed at a scene, a nice, clean image forms that is bright and in focus. But if I now block the large lens of the camera in some way, what will happen?



This is the same camera with a part of the lens blocked out by attaching pieces of tape in the shape of a cross $\boxed{1}$. The image produced is darker because less light is being passed through by the lens, but it is still in focus! This non-intuitive effect occurs because, regardless of the shape of the aperture, the part of the lens that is exposed to light rays arriving from the scene will focus those rays to create a sharp image. This is a really useful feature of a lens. It is for this reason that, in a real-world setting, reasonably well-focused images can be captured even when the lens has some dust particles or droplets on it.

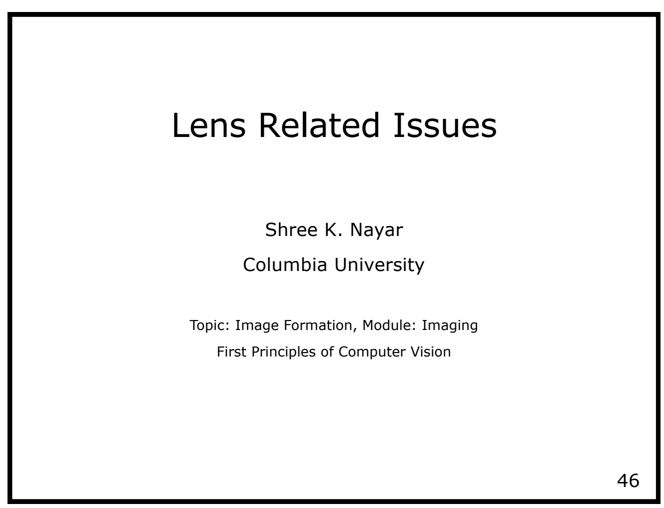




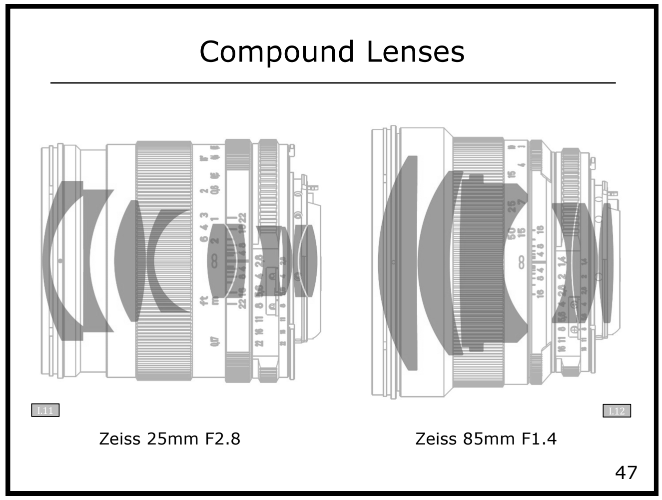
Let us now look at another interesting property of a lens. Here we have an image sensor on the left and a tilted lens on the right. Where is the plane of focus for this tilt-lens camera? Let us first consider the optical axis which is perpendicular to the lens, runs through the center of the lens, and pierces the image plane. The scene point P on the optical axis that is focused on the image sensor can, as before, be determined using the Gaussian lens law.

Because the image plane is tilted, we can guess that, in this case, the plane of focus is not going to be parallel to the image plane. We can determine the plane of focus using what is called the Scheimpflug condition. First, we extend the line that represents the image plane and the line that passes through the lens so that they intersect at the point Q. The line that passes through P and Q corresponds to the plane of focus of this imaging system with a tilted lens.

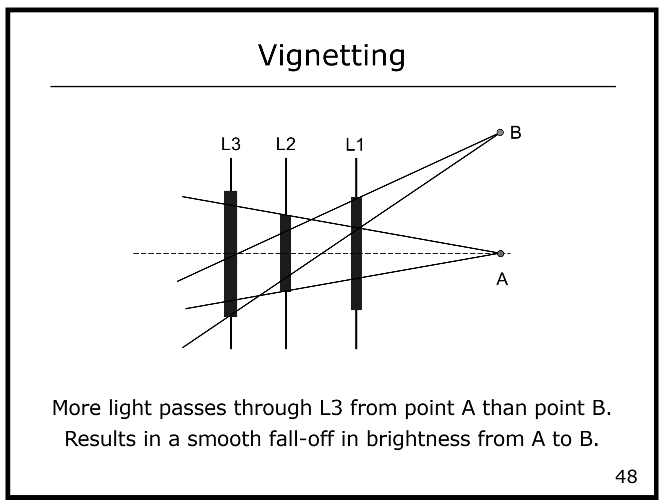
What is the utility of this system? There are applications in which a camera looks out at a scene and the region of interest in the scene is not a plane that is parallel to the image plane but rather a different plane. For example, consider a camera mounted on a car where the goal is to monitor the quality of the road ahead of the car. Photographers also use such a “tilt-camera” to capture interesting images with depths of field that are different from what a normal camera would provide.



Now, let us talk about some issues related to lenses. Even if a lens is perfectly manufactured, it turns out that it will still produce some aberrations, or undesirable effects.

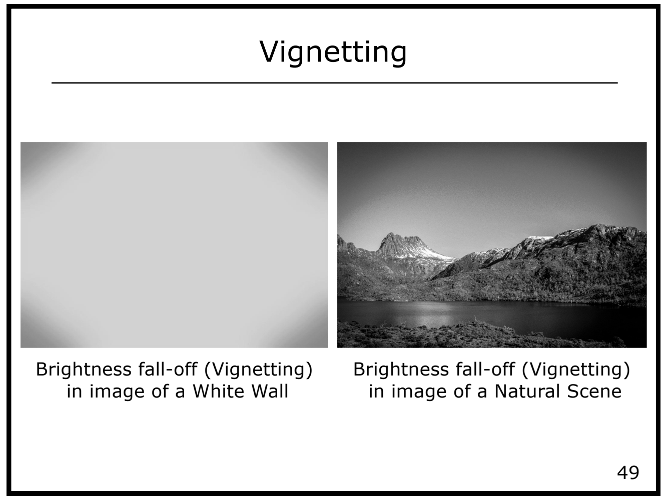


While we have used a single lens to describe most of our concepts, in practice, lenses seldom have a single lens inside them. Take a look at the two lenses shown here. They have somewhere between half a dozen to a dozen lenses within them. Such a lens is called a compound lens. This raises the question, why do we need all of these lenses when we seem to be able to do a lot with a single lens? It turns out that it is a real challenge to create an image that has the same quality across the entire image plane. Often a lens will produce higher image quality in the center of the image as compared to the periphery of the image. In order to produce images of high quality over the entire field of view, a series of lenses are needed. Lens design is where art meets science; experts use design recipes that combine a series of lenses of different shapes and materials (refractive indices) to compensate for the undesirable effects of each other. The end result is a high-quality image.

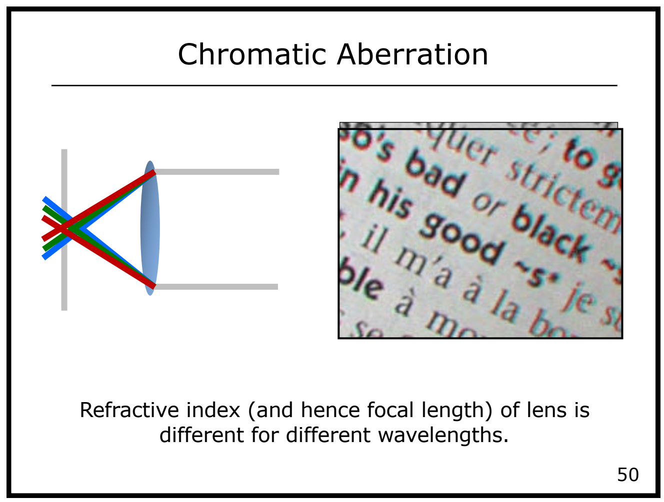


Let us now examine vignetting, one of the undesirable effects that compound lenses suffer from. Here, we have lenses L1, L2, and L3, and they all have different sizes, or openings. When we consider the point A on the optical axis, a lot of the light from this point that arrives at lens L1 manages to make its way through to the image sensor. However, if point A is moved to point B on the same plane of focus but away from the optical axis, we see that there is a greater chance that rays of light from point B are going to be blocked as they travel through the set of lenses to the image plane. For this simple reason, images often tend to be darker towards the periphery, an effect called vignetting.

On the left is an image of a perfectly flat white surface, but we can see that the corners are darker. On the right is another image where, again, the corners are darker. Given a lens, one can measure the vignetting introduced by the lens over the field of view and then compensate for vignetting in images taken using that lens.



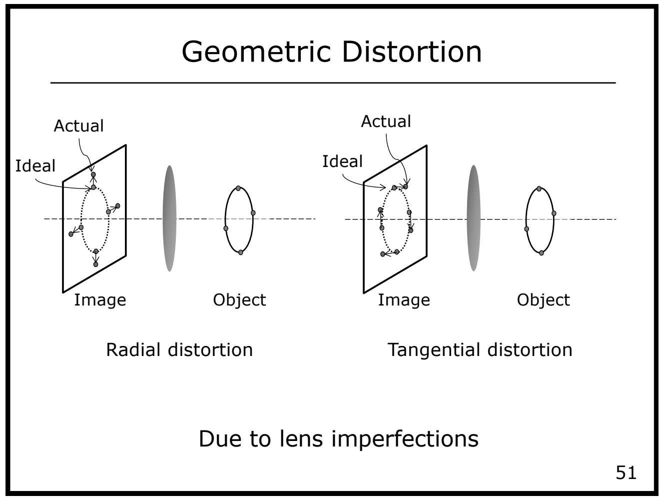
Here is another lens related effect called chromatic aberration. Remember that a lens is made of a certain material, such as glass or plastic, which has a certain refractive index — that is what gives it its bending power. The refractive index of the lens material is greater than the refractive index of air. It turns out that the refractive index is a function of the wavelength of light. We know that the wavelength of visible light goes from 400 nanometers to 700 nanometers, where 400 is blue light, 700 is red light, and green light is somewhere in between at about 550 nanometers. Since the refractive index depends on the wavelength, the focal length of a lens also depends on the wavelength. That is, the bending power of the lens depends on the wavelength.



We can see what happens when parallel rays of white light are imaged. Note that white light contains all wavelengths of light. Since the focal length depends on the wavelength, the red light gets bent the least, the green light gets bent more, and the blue light gets bent the most. This causes some shifts in color in the image. On the right is an image of a printed sheet of paper. Even though the sheet itself has no color (it has only shades of gray), its image has color effects around the edges of the printed letters. This is called chromatic aberration.

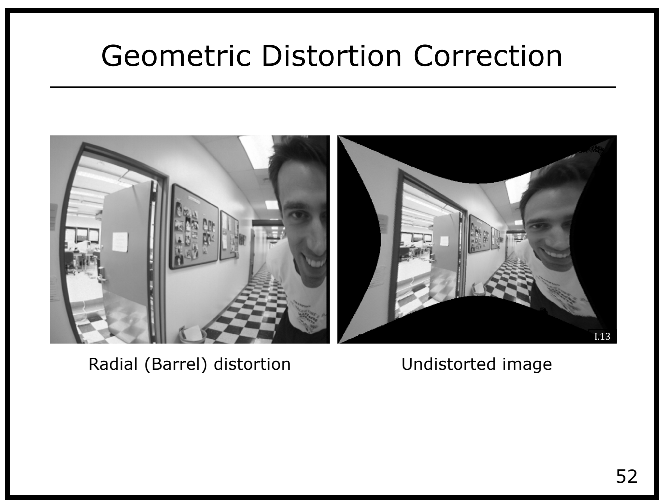
There are also geometric aberrations or distortions. Two well-known geometric distortions are radial distortion and tangential distortion. In the case of radial distortion, as we move away from the center of the image, points tend to get pushed out more and more. As a result, there is an apparent bulging of the image, also known as barrel distortion. If we know exactly what the barrel distortion is, we can correct for it.

In the case of tangential distortion, there is a slight twisting of the image. As we go farther away from the center of the image the twisting increases.

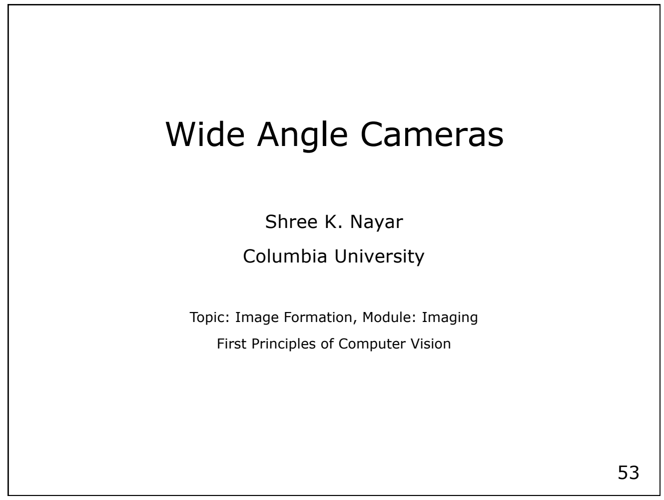


All of the lens effects we have discussed thus far tend to be more severe in the case of simple lens designs; the cheaper the lens is, typically speaking, the more visible the effects of vignetting, chromatic aberration, and radial and tangential distortion. A high-quality lens tends to be made of several individual lenses whose shapes and materials are optimized to minimize all of the above effects.

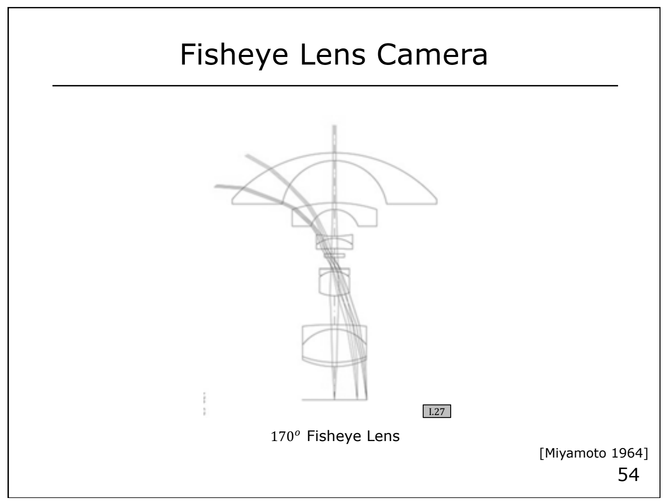
Barrel distortion is often found in images taken with wide-angle lenses. On the left is an image taken with one such lens. Note that straight lines in the scene no longer map to straight lines in the image. If we know a priori what the barrel distortion is, we can take this captured image and apply a simple mapping software to obtain a pure perspective image, such as the one on the right. In this image, all straight lines in the scene do end up as straight lines in the image. An interesting thing to note here is that the field of view of the corrected image is not rectangular because, due to barrel distortion, the field of view of the original image itself was not rectangular even though the image was formed on a rectangular sensor.



Consider the problem of capturing a hemispherical field of view. If one uses perspective projection, an infinite image plane would be needed to capture the image, which is clearly not practical. Even if this were possible, objects would appear severely stretched in the image. By designing a lens with geometric distortions, we take the complete hemispherical field of view and compress it down to a small area, where the image sensor can be placed. Let us take a look at how unusually large fields of view can be captured using lenses, as well as a combination of lenses and mirrors.

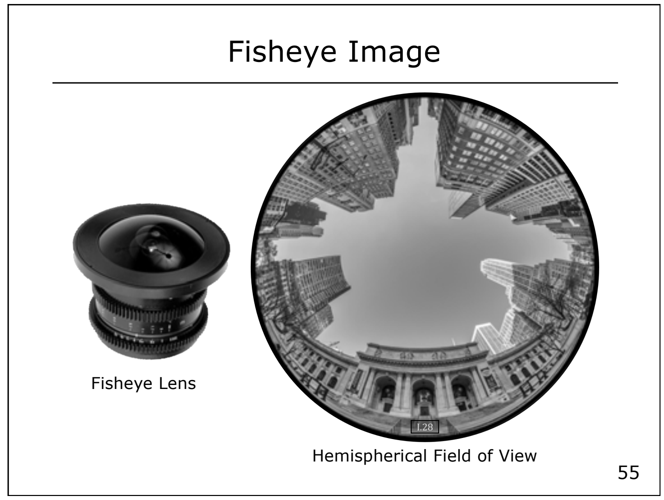


Shown here is the fisheye lens, which was introduced by Miyamoto in 1964. This lens was designed to have a very large field of view of 170 degrees. It uses a series of meniscus lenses, which are convex on one side and concave on the other. This allows the system to bend light rays severely, especially rays that are farther away from the optical axis. In the design shown here two meniscus lenses are combined with a series of other lenses to compress the field of view and project it onto a small image sensor.

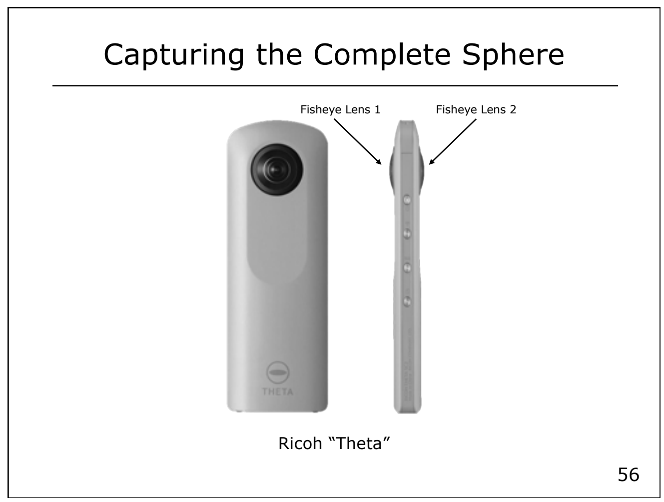


Such a lens may not have a single center of projection but rather at a locus of viewpoints. Generally, the locus of viewpoints is compact and hence can be approximated by a single point. In most vision applications, it is important for the camera to have a single viewpoint — if the world is viewed from disparate viewpoints, it is not possible to create a single perspective view of it without knowing the complete 3D structure of the scene.

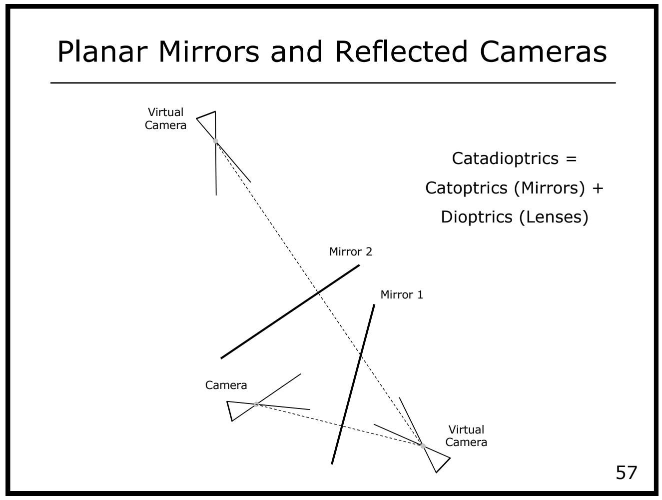
This fisheye lens has a 180-degree field of view, which means it captures a complete hemisphere. On the right is an image taken using this lens. Since it can be assumed to satisfy the single viewpoint constraint (that is, it has a compact locus of viewpoints), any part of this image can be mapped to a pure perspective image using software. In fact, one can create a "software camera" that allows the viewer to look around the scene in different directions, where all the views are computed from this single captured image.



One limitation of the fisheye lens is that it cannot capture much more than a hemisphere. The device shown here — Ricoh's "Theta" camera — makes uses of two fisheye lenses, each capturing one hemisphere. The two lenses are placed back-to-back and very close together; all the imaging hardware has been packed in between the two lenses so that the viewpoints of the lenses are close enough that the entire system can be assumed to have a single viewpoint. This allows the two captured hemispheres to be stitched together to create a complete spherical image. The spherical image then allows one to "look" at the scene in any direction from the viewpoint of the camera.

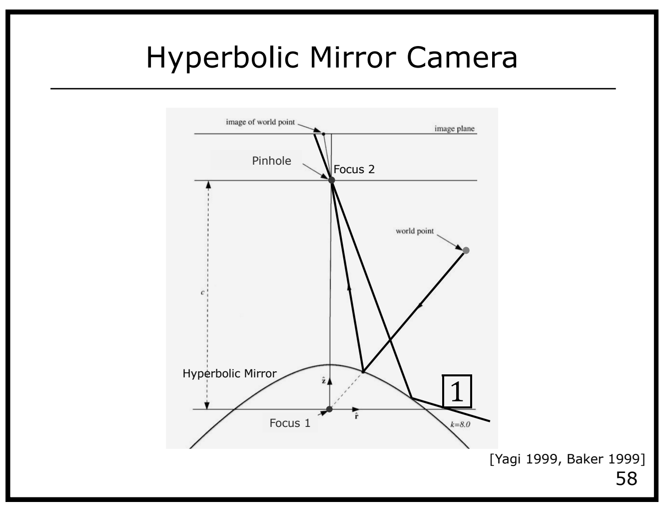


There are applications, such as video conferencing, where we would like to capture a 360-degree panorama. We may like the panorama to have a field of view that is, say, $\pm$ 45 degrees with respect to the equator. It is hard to design a lens that can achieve this. We can overcome this limitation by introducing mirrors into the imaging system.



Let us first consider planar mirrors. Shown here is a lens camera with its viewpoint and field of view. If we place a mirror in the field of view of this camera, the real camera is reflected by the mirror to create a virtual camera. Placing another mirror in the field of view of this virtual camera results in another virtual camera. Thus, planar mirrors can be used to manipulate the position and orientation of a real camera; this is called optical folding. The use of mirrors is called catoptrics and the use of lenses is called dioptrics. A system that uses both lenses and mirrors is known as a catadioptric system.

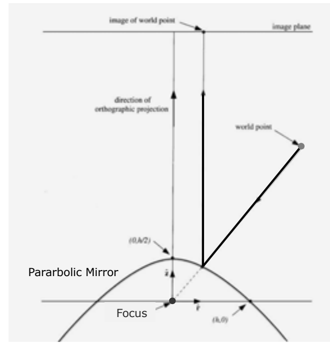
As seen above, by using planar mirrors we can change the position and the orientation of the field of view of a lens camera, but we cannot change the size of its field of view. The size of the virtual camera's field of view is exactly the same as that of the real camera. What we are interested in is wide-angle imaging. In order to enlarge the field of view of a lens camera we can use curved mirrors.



Here is a design that uses a hyperbolic mirror. The hyperbola has two foci. All the rays of light that come in the direction of the first focus, Focus 1, get reflected by the hyperbolic mirror towards the second focus, Focus 2. That is where we would place the center of projection, or effective pinhole, of the lens camera. The camera image is then the reflection of the world by the hyperbolic mirror, and it represents a view of the scene as seen from a single point, which is Focus 1. Note here that the use of a curved mirror allows a field of view that is much greater than a hemisphere to be captured. Consider this ray of light $\boxed{1}$ that corresponds to a point in the scene that lies beneath the equator, that is, beneath the horizontal line that passes through Focus 1. The parameters of the hyperbola can be chosen to achieve the field of view needed for the application at hand.

One can also use a parabolic mirror. The parabola has a single focus which sits within it. It takes all the rays of light that come in the direction of its focus and reflects these rays to be parallel to one another. In this case, a perspective lens is not used, but rather what's called an orthographic lens, or a telecentric lens, is used to capture the parallel reflected rays of light and then create a wide-angle image on the image plane.

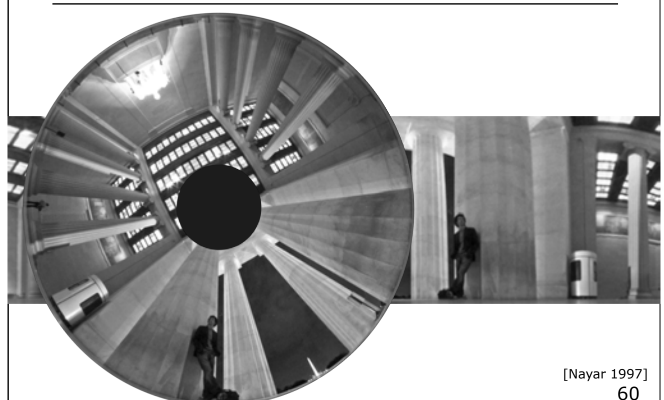
Parabolic Mirror Camera



[Baker 1999]
59

Here you see an image taken with a parabolic mirror camera. The black disc seen in the center represents one of the downsides of using mirror-based wide-angle cameras. This is where the mirror sees the lens itself; it is essentially the blind spot of the imaging system. On the positive side, a field of view is achieved here which goes well beyond 180 degrees. Since we have satisfied the single viewpoint constraint — the viewpoint of this imaging system lies at the focus of the parabola — software can be used to map any part of this image to a perspective image or to map the entire captured image to a 360-degree cylindrical panorama of the type shown here.

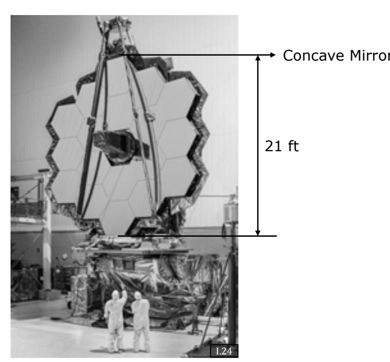
Parabolic Mirror Image



[Nayar 1997]
60

The use of curved mirrors is relatively new in the case of wide-angle imaging, but it has been used for another reason for centuries, namely the design of telescopes. In the case of a telescope, the goal is to view an object that is extremely far away, resulting in a narrow field of view. Since the object is far away, it is really dim, and we want to collect as much light as possible from it. For this it is necessary to use a wide aperture. That is where concave mirrors come in useful.

Concave Mirrors and Telescopes

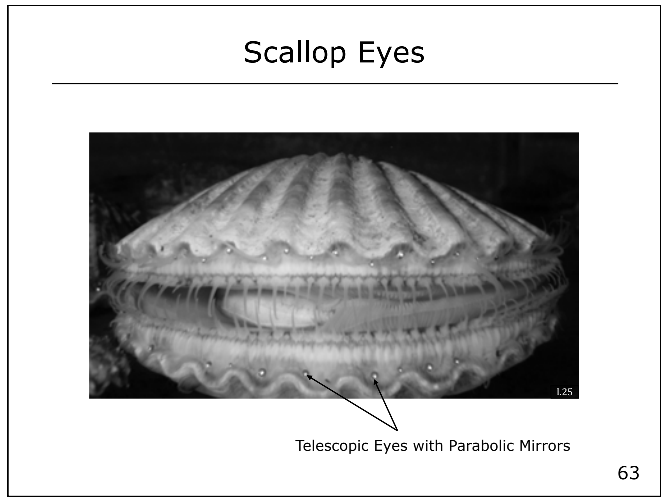


James Webb Space Telescope

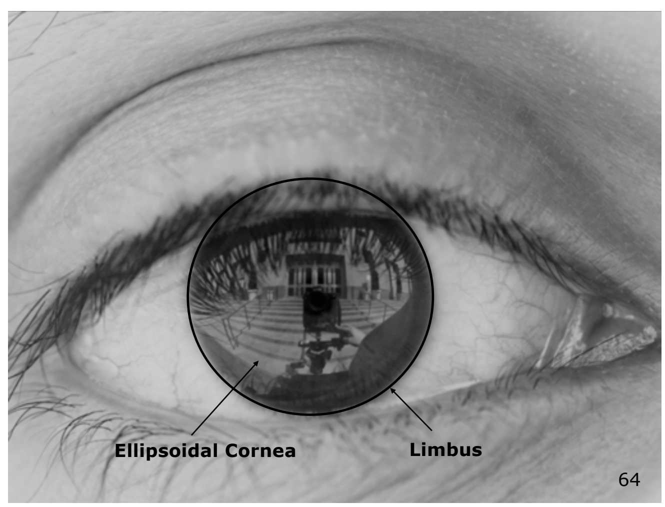
62

Here is the James Webb space telescope, which utilizes a concave parabola that is massive — 21 feet in diameter. There is no known method for manufacturing a single mirror of that size. Instead, the mirror is constructed by tiling segments where each one is a shallow curved mirror. The complete tiled mirror is a concave parabola that focuses the incoming light onto its focus, which is where the image sensor is placed.

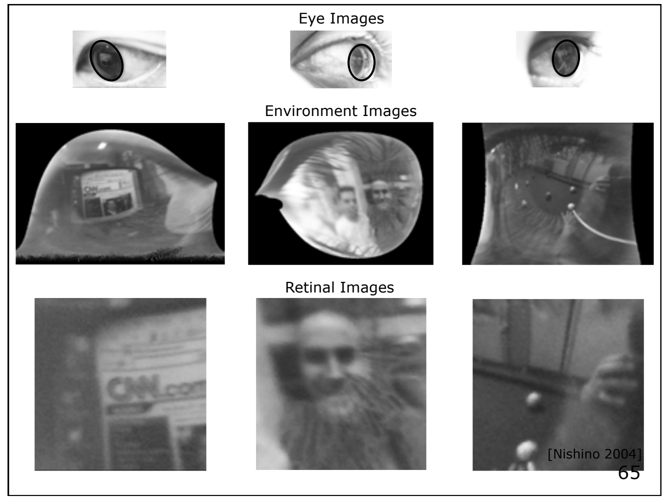
It is a remarkable fact that nature has created eyes that use curved mirrors. Here we see a scallop. Observe the tiny dots near the edge of the shell of the scallop — each of the dots is an eye. Each one of these has a concave parabolic mirror inside of it that takes light from a narrow field of view and then focuses it, like a telescope, onto a single receptor, a pixel. With hundreds of these eyes around its periphery, the scallop is able to get an idea of what the distribution of light is all around it.



Here is another example of a wide-angle imaging system. We call it the corneal imaging system. Shown here is a high-resolution image of an eye. The cornea of the eye has a thin film of tear on it, which makes it behave like a reflective surface. It is not a highly reflective mirror, but a mirror all the same. In the case of a normal adult cornea, its shape has been found to be ellipsoidal with fairly strict parameters. Observe the border, called the limbus, between the cornea and the white part of the eye, called the sclera. If we can find the limbus of an eye in an image, then we have found the position and orientation of the ellipsoidal mirror (the cornea) with respect to the camera that was used to capture the image. Now, the image information inside the cornea can be mapped into a wide-angle image of the world around the person the eye belongs to.

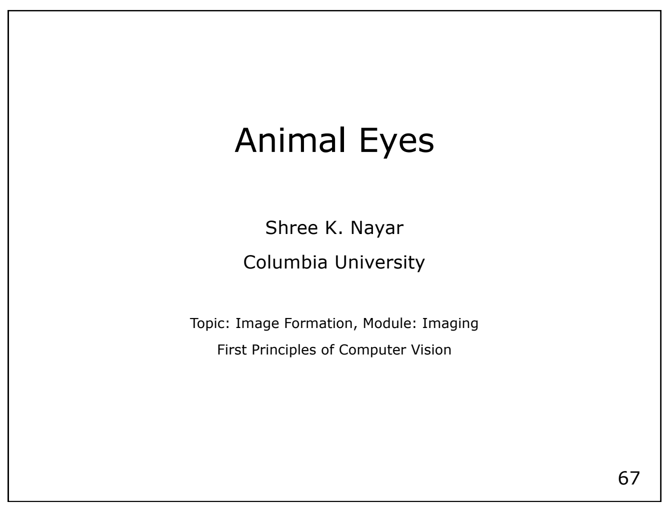


Here are eyes looking in different directions, and in each case we find the limbus. We can take the information inside the cornea and map it into a wide-angle image, or an environment image (middle row). We see that these images have different shapes. In other words, they have different fields of view. This is because the field of view of the corneal imaging system depends on the direction that the person is looking in with respect to the camera. That is, it depends on the orientation of the ellipsoidal mirror (cornea) with respect to the camera.

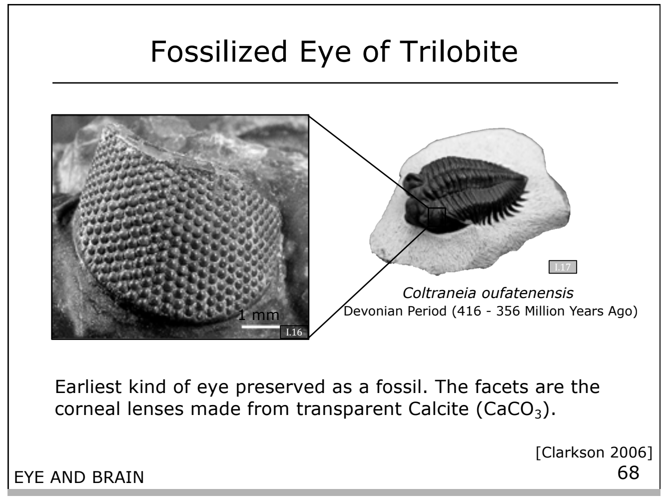


The orientation of the cornea can be determined from the parameters of the limbus, which is, in general, an ellipse. That information can be used, in conjunction with the wide-angle environment image, to get an estimate of the image that is falling on the fovea of the person's eye. These are called retinal images and are shown in the bottom row. In other words, without implanting a chip inside this person's eye, we are able to, just from a photograph of them, figure out what they are looking at.

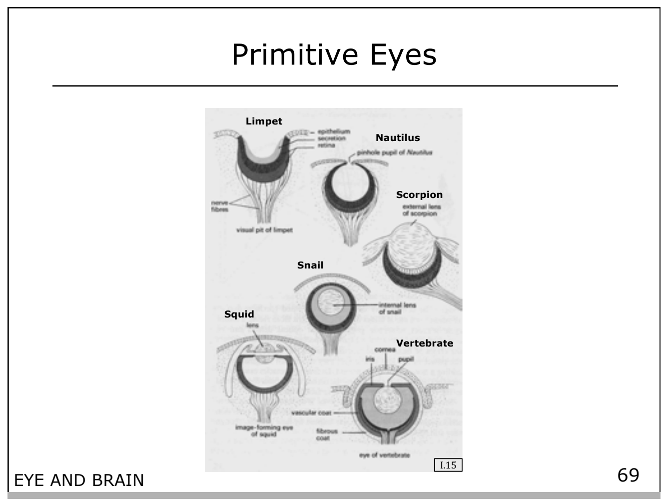
Evolution has created a stunning array of different types of eyes. Let us take a look at a few of these before focusing on the human eye.



Let us start with the trilobite. It is an arthropod, which is a type of insect, and it represents the oldest fossilized eye that we know of. The trilobite lived about 400 million years ago, and its eye is not like our eye. It is a compound eye with thousands of tiny little lenses. Each bump seen here, which is a small fraction of a millimeter, is a lens made of transparent calcite, and this lens focuses light on a single receptor.

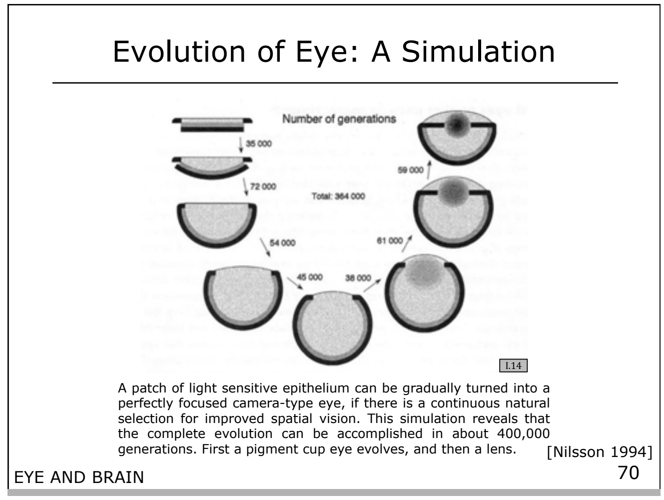


Here are a few more primitive eyes. Let us start with the eye of the limpet — it does not have a lens, but rather a curved image sensor. In the case of a flat image sensor, without a lens, all receptors essentially receive the same amount of light. But if the image sensor is curved, different points receive light from within different cones (fields of view). Therefore, the image measured by a curved sensor conveys some information regarding the spatial distribution of the light around the eye.



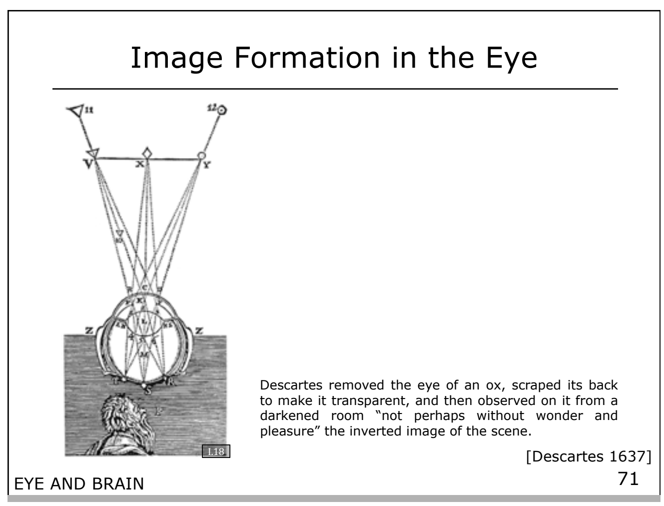
Also shown is the eye of the *Nautilus pompilius* — which we talked about earlier — which has a pinhole rather than a lens, the eye of the scorpion which has a very large external lens, the eye of the snail which has an internal lens and an external covering, and the eye of the squid which looks like a fully formed eye with an external layer. Finally, the eye of a vertebrate is a bit more sophisticated — it has a cornea, which acts as a protective layer but also partially as a lens. It also has an internal lens which can change shape, a pupil, and an iris.

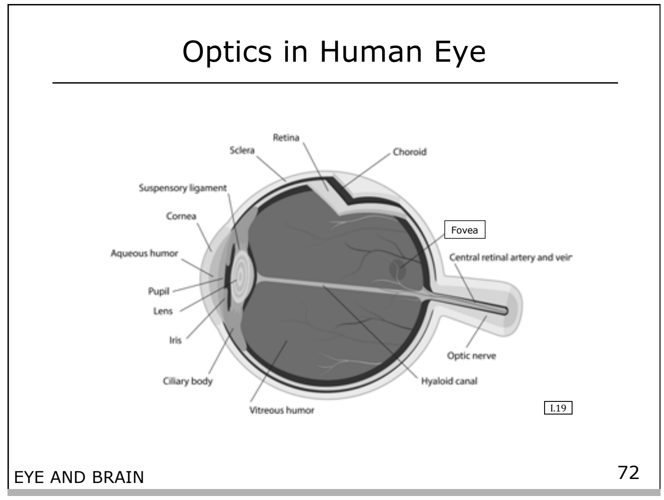
Here is an interesting simulation that was done by Nilsson and Pelger to try and understand how nature might have evolved the human eye. The simulation starts off with a flat layer of light-sensitive tissue and then attempts to modify the eye to make the image it captures brighter and sharper. The first thing that happens in the simulation is that the tissue begins to curve, so that it can gain some directional sensitivity to light. It gets more and more curved, but after a certain point, curving it further reduces the brightness of the image. At that point, a lens begins to form and finally we arrive at an eye that's more like the human eye.



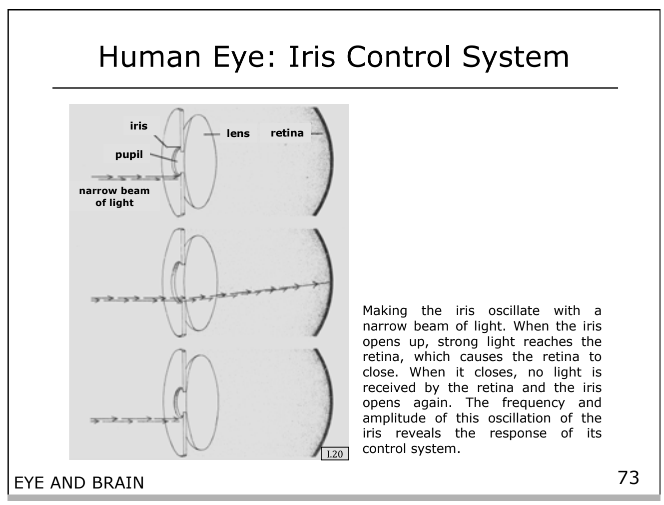
The numbers between consecutive stages of the simulation correspond to the number of generations that the simulation estimates it took to go from one stage to the next stage. To go from the flat tissue all the way to a fully formed eye, it is estimated that it may have taken roughly 400,000 generations.

It is not surprising that the eye has been of great interest to scientists and philosophers for many centuries. Some incredible experiments were done, in particular, by Keppler and by Descartes. Here we see an experiment by Descartes, who took the eye of an ox, scraped the backside, and then stuck this eye in a wall with the eye looking out at the scene. He then observed the back of the eye from within a dark room to see the inverted image formed by the eye. This is a remarkable experiment, given that it was done back in 1637. In fact, Descartes even noticed that if he squeezed the eye of the ox, the lens inside the eye changes in shape, thereby changing the focus of the image it forms!





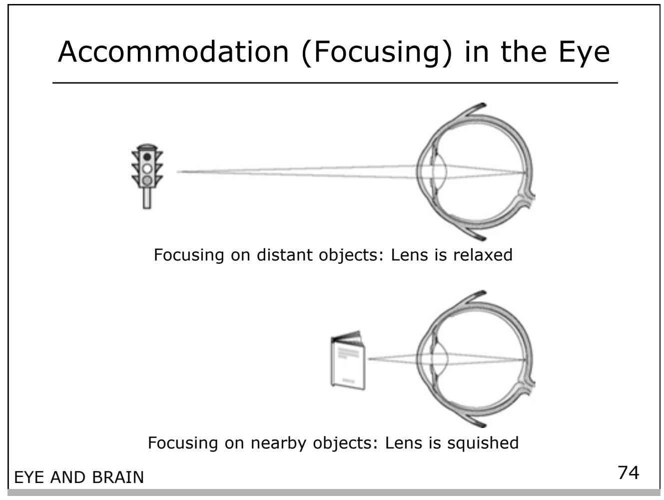
That brings us to the human eye. In this illustration we see the cornea, which is a transparent protective layer that also happens to have some bending power and therefore acts as a lens as well. Behind the cornea is the iris and the pupil, followed by the lens of the eye. The lens is a bag of fluid with gelatinous material inside of it. There are muscles around the lens that can apply forces to its periphery, so as to change the shape of the lens and therefore change its focal length, or its bending power. The image formed by the eye falls on the retina, which is a curved image sensor that has rods and cones — the pixels of the eye. The fovea is the part of the retina which has maximum resolution, with the resolution falling off towards the periphery of the retina. As such, when looking at something, that something is actually falling on the fovea. The image captured by the retina is passed through the optic nerve and makes its way to the visual cortex where it is analyzed.



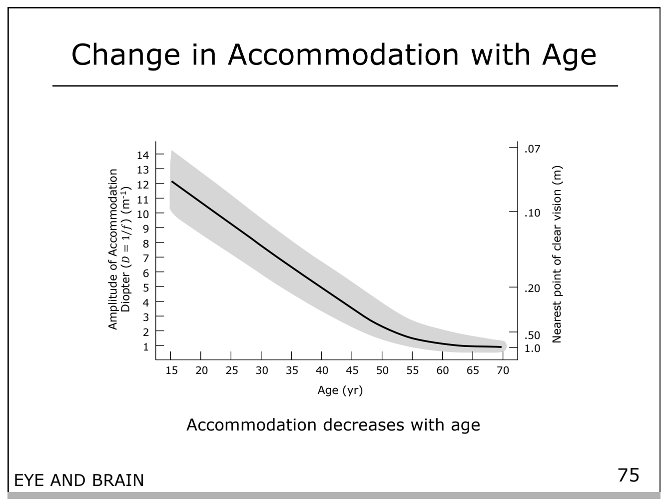
Let us focus on the region that includes the iris and the pupil. In the same way that the diaphragm of a camera's lens adjusts the aperture diameter, the iris modifies the size of the pupil based on how much light is entering the eye. When walking out in bright sunlight, the eye is flooded with light, and therefore the iris closes up and lets in less light. If one then walked into a dark room, the iris opens up so it can capture more light.

Research has been done to study the control system that drives the iris of the human eye. Shown here is the lens in the front and the retina in the back. In front of the lens is the iris that controls the pupil, or the aperture of the lens. In this experiment, a narrow beam of light is shone at the edge of the pupil on the iris. As long as it is just outside the pupil, the eye doesn't receive any light. Since the eye does not receive any light, the iris opens up to widen the pupil. At some point, the narrow light beam enters and floods the eye. This causes the iris to close again to make the pupil smaller. This simple setup makes the iris oscillate, and by studying the frequency and amplitude of the oscillation, researchers have characterized the control system that drives the iris.

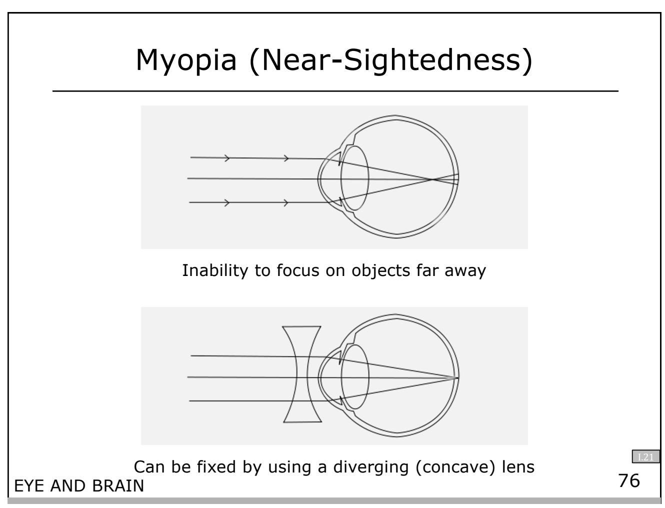
Now, let us consider the eye's ability to focus, also called accommodation. When focusing on a distant object, the lens is relaxed. We do not sense that our eye is relaxed, because these are very small changes that happen inside the eye. In order to focus on something close, like the page of a book, the lens is squished by the ciliary muscles, making it shorter in focal length and giving it more bending power.



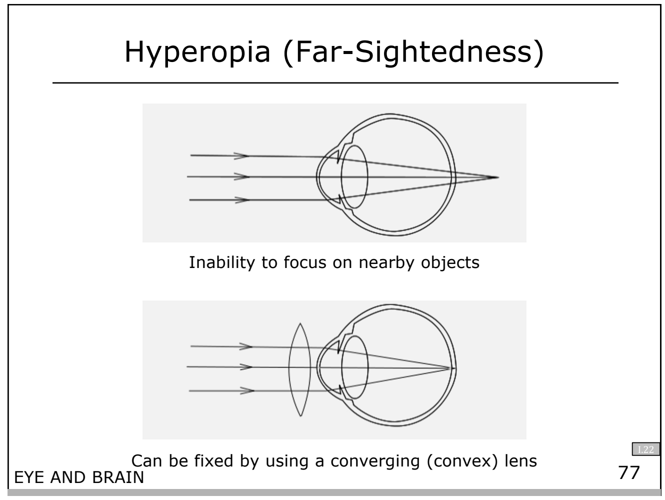
Our ability to accommodate (focus) goes down as a function of time. Shown here is age on the horizontal axis plotted against the shortest distance an average person can focus at. When a person is really young, they can focus on things that are as close as 7 to 10 centimeters away. The reason is that their liquid lens is malleable, allowing them better control over its shape. As a person gets older, the lens begins to harden and the shortest distance they can focus on begins to increase. Somewhere around 50 years of age, a person, on average, can only focus at a distance of about 100 centimeters. That is typically when one begins to need glasses.



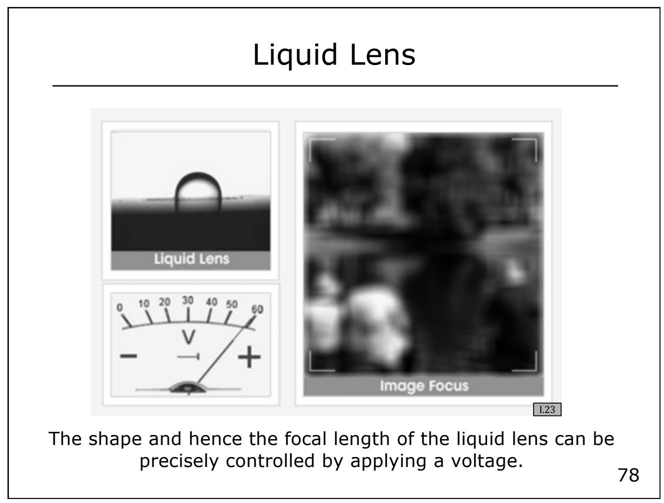
Here are some eye conditions that require correction. In the case of myopia, the lens of the eye has hardened into a shape which has too much bending power. In this case, when looking at something far away, the image is formed in front of the retina. This can be corrected using a concave lens, which diverges incoming rays of light before they go into the lens of the eye. The end result is an image that is focused on the retina.



The opposite of myopia is hyperopia or far-sightedness. In this case, when looking at something far away, the image is formed behind the retina. In other words, the lens does not have enough bending power. In this case, a convex lens is used to aid the lens of the eye so that the image is formed on the retina.



As we discussed, the lens of the eye is a bag of liquid with gelatinous material inside it. It has ciliary muscles that tug on the lens to change its shape. Most cameras, in contrast, use solid lenses made of glass or plastic and a series of such lenses in order to change the focus and magnification of the camera. Recently, scientists have been working on developing liquid lenses as well. Here we see a liquid lens developed by a company called Varioptic. It is based on a phenomenon called electrowetting. The curvature of the top surface of the liquid can be controlled by applying an electric field to the liquid. The curvature increases with the strength of the electric field. This in turn changes the focal length of the lens.



References: Textbooks

Computer Vision: Algorithms and Applications
Szeliski, R., Springer

Computer Vision: A Modern Approach
Forsyth, D and Ponce, J., Prentice Hall

Robot Vision
Horn, B. K. P., MIT Press

Animal Eyes
Land, M. and Nilsson, D., Oxford University Press

Medical Physiology, Vol. I
Mountcastle, V. B., C. V. Mosby Company

Eye and Brain
Gregory, R., Princeton University Press

79

References: Papers

[Aizenberg 2001] J. Aizenberg, A. Tkachenko, S. Weiner, L. Addadi and G. Hendler. "Calcitic microlenses as part of the photoreceptor system in brittlestars." Nature, 2001.

[Clarkson 2006] E. Clarkson, R. Levi-Setti, G. Horváth. "The eyes of trilobites: The oldest preserved visual system". Arthropod structure and development, 2006.

[Baker 1999] S. Baker and S. K. Nayar. "A theory of single-viewpoint catadioptric image formation". International journal of computer vision 35 (2), 175-196.

[Descartes 1637] R. Descartes. "La Dioptrique". 1637.

[Frisius 1545] Gemma-Frisius. "De Radio Astronomica Et Geometrico". 1545.

[Miyamoto 1964] K. Miyamoto. "Fish Eye Lens". JOSA, 1964.

[Nishino 2004] K. Nishino and S. K. Nayar. "The World in an Eye". IEEE CVPR, 2004.

80

References: Papers

[Nayar 1997] S. K. Nayar. "Catadioptric Omnidirectional Camera". IEEE CVPR, 1997.

[Nillson 1994] D-E. Nilsson and S. Pelger. "A pessimistic estimate of the time required for an eye to evolve". Proc of Royal Society, 1994.

[Yagi 1999] Y. Yagi. "Omnidirectional Sensing and Its Applications". IEICE Trans. Information and Systems, Vol.E82-D, No.3, pp.568-579, 1999.

81

Image Credits

I.1 Gemma-Frisius, 1545. Public Domain.
I.2 https://wikipedia.org/wiki/File:Nautilus_pompilius_3.jpg. © Hans Hillewaert. Licensed under CC BY-SA 4.0.
I.3 https://wikipedia.org/wiki/File:Nautilus_pompilius_(head).jpg. © Hans Hillewaert. Licensed under CC BY-SA 4.0.
I.4 Vanishing Tracks. Diesel Demon. Licensed under CC BY 2.0.
I.5 https://commons.wikimedia.org/wiki/File:Vermeer%27s_The_Music_Lesson.jpg. The Music Lesson. Vermeer. Public Domain.
I.6 Ian Norbury. Used with permission.
I.7 Camilo Trevisan. Used with permission.
I.8 N. Joel in Optics. Eugene Hecht. Addison-Wesley 1998.
I.9 gmpicket. Licensed under CC BY-NC-ND 2.0.
I.10 Jens Dahlin. Licensed under CC BY-NC-ND 2.0.
I.11 Zeiss 25mm F2.8. Carl Zeiss, Inc.
I.12 Zeiss 85mm F1.4. Carl Zeiss, Inc.
I.13 R. Swaminathan and S.K. Nayar.
I.14 Dan-Eric Nilsson. Used with permission.

82

Image Credits

I.15 Eye and Brain, 4th Edition. Richard L. Gregory. Princeton University Press. Used with permission.
I.16 Clarkson, Levi-Setti, and Horvath. Used with permission.
I.17 Clarkson, Levi-Setti, and Horvath. Used with permission.
I.18 La Dioptrique, 1637. Descartes. Public Domain.
I.19 Purchased from iStock by Getty Images.
I.20 Eye and Brain, 4th Edition. Richard L. Gregory. Princeton University Press. Used with permission.
I.21 https://en.wikipedia.org/wiki/File:Myopia.svg. Licensed under CC SA 1.0.
I.22 https://en.wikipedia.org/wiki/Far-sightedness. Licensed under CC SA 3.0.
I.23 Corning Varioptic.
I.24 https://en.wikipedia.org/wiki/Telescope . Public Domain.
I.25 https://en.wikipedia.org/wiki/Scallop. Public Domain.
I.26 http://cmp.felk.cvut.cz/~svoboda/Demos/Omnivis/Hyp2Img/ . Tomas Svoboda.
I.27 dpreview.com
I.28 Purchased from iStock by Getty Images.
I.29 Purchased from iStock by Getty Images.

83

Acknowledgements: Thanks to Nisha Aggarwal and Jenna Everard for their help with transcription, editing and proofreading.

References

[Szeliski 2022] Computer Vision: Algorithms and Applications, Szeliski, R., Springer, 2022.

[Forsyth and Ponce 2003] Computer Vision: A Modern Approach, Forsyth, D and Ponce, J., Prentice Hall, 2003

[Horn 1986] Robot Vision, Horn, B. K. P., MIT Press, 1986.

[Gregory 1966] Eye and Brain, Gregory, R., Princeton University Press, 1966.

[Land and Nilsson 2012] Animal Eyes, Land, M. and Nilsson, D., Oxford University Press, 2012.

[Mountcastle 1980] Medical Physiology (Vol. 2),(131h ed.), Mountcastle, V. B., St. Louis: The CV Mosby Company, 1980.

[Aizenberg 2001] J. Aizenberg, A. Tkachenko, S. Weiner, L. Addadi and G. Hendler. “Calcitic microlenses as part of the photoreceptor system in brittlestars.” Nature, 2001.

[Clarkson 2006] E. Clarkson, R. Levi-Setti, G. Horváth. ”The eyes of trilobites: The oldest preserved visual system”. Arthropod structure and development, 2006.

[Baker 1999] S. Baker and S. K. Nayar. “A theory of single-viewpoint catadioptric image formation”. International journal of computer vision 35 (2), 175-196.

[Descartes 1637] R. Descartes. “La Dioptrique”. 1637.

[Frisius 1545] Gemma-Frisius. “De Radio Astronomica Et Geometrico”. 1545.

[Miyamoto 1964] K. Miyamoto. “Fish Eye Lens”. JOSA, 1964.

[Nishino 2004] K. Nishino and S. K. Nayar. “The World in an Eye”. IEEE CVPR, 2004.

[Nayar 1997] S. K. Nayar. “Catadioptric Omnidirectional Camera”. IEEE CVPR, 1997.

[Nillson 1994] D-E. Nilsson and S. Pelger. “A pessimistic estimate of the time required for an eye to evolve”. Proc of Royal Society, 1994.

[Yagi 1999] Y. Yagi. “Omnidirectional Sensing and Its Applications”. IEICE Trans. Information and Systems, Vol.E82-D, No.3, pp.568-579, 1999.