Introductory Optics

Version 2.04

Norman Wittels
121 Clark Road
Brookline MA 02146
wittels@world.std.com

Introductory Optics

Introduction

These notes introduce the physical concepts and mathematical models used to analyze image-forming optical systems. An image is a spatial distribution of light that is geometrically similar to the pattern of light emitted or reflected from an object at some other place. An understanding of optics aids in understanding the relationship between objects in the three-dimensional world and the two-dimensional digital images used as the starting point in computer image analysis. It also help us to understand how our eyes work and how our brains use the information from our eyes to reconstruct the three-dimensional world we perceive.

Our emphasis is not on designing optical systems, which is a rather involved and complicated process, but on understanding how optical systems work. In particular, we will investigate several simple models for imaging systems, discuss the conditions under which each model is valid, and show, for each, how to calculate the image’s size, sharpness, and illuminance (how much light forms the image). These are the quantities of most use to people working with computer graphics, computer vision, and machine vision systems.

Pinhole Optics

Many useful results can be obtained by using the simplest of all optical models – the *pinhole optics* model. Image formation in this model is described by the equations describing the pinhole camera, in which an inverted image is formed on an image plane by a small hole in an opaque card, Figure 1.

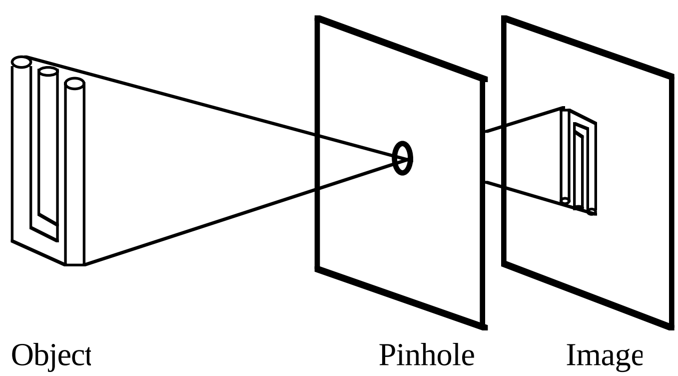


Figure 1. A Simplified Pinhole Camera

The image is a consequence of the fact that light travels in straight lines. The only light striking a point in the image plane comes from a single point on the object – the point colinear with both the image point and the pinhole. As we will show, this is sufficient to form an image that is geometrically similar to the object.

Construct a cylindrical coordinate system r,z,θ with its origin at the center of the pinhole, Figure 2. The z-axis is also called the *optical axis*. The direction of the optical axis arbitrary and it is usually selected to exploit some symmetry in the problem or to make some particular object point or image point lie on the z axis.

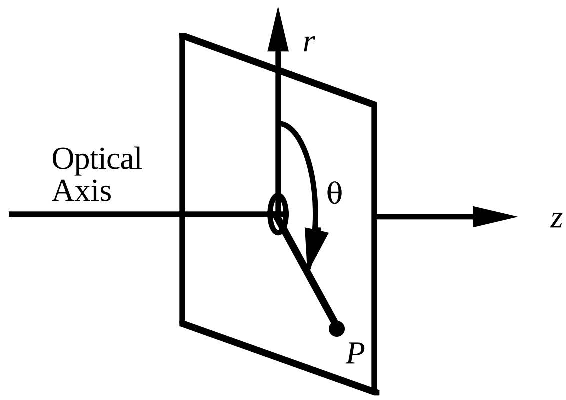


Figure 2. A Pinhole

Assume, for now, that the object lies in an object plane perpendicular to the optical axis. Light rays emanate from it along straight lines in all directions. Some of the rays will pass through the pinhole and illuminate the image plane, also assumed to be perpendicular to the optical axis. This last assumption is not essential but it is convenient so it is usually made in introductory optics. We will follow the light rays emanating from an arbitrary object point P_O that is not on the optical axis and see where they strike the image plane. The object point and the optical axis define a unique r-z plane in which the angle, θ, is constant: $\theta=\theta_O$. The ray from P_O through the center of the pinhole is a straight line connecting two points in this plane; hence, all points on the ray lie in the plane. Thus all points along the ray, including the point P_I where it intersects the image plane, will have the same value of θ. We can therefore confine our analysis to the plane $\theta=\theta_O$, Figure 3, and the ray can be described as a function of r and z only: $r=r(z)$.

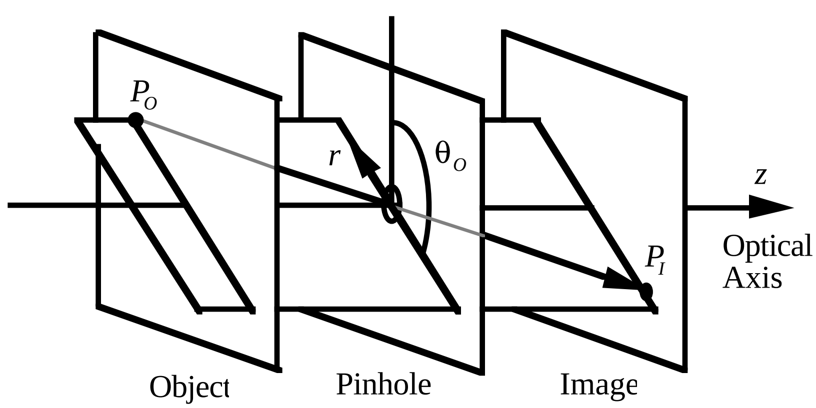


Figure 3. A light ray that passes through the pinhole

In the $\theta=\theta_O$ plane the equation of the ray is

$$r' = \frac{dr}{dz} = -\frac{R_O}{S_O} \qquad (1)$$

where R_O is the radius of the object point in the object plane located at $z=-S_O$. The sign is negative because the ray has to be moving towards the axis so that it can pass through the pinhole, as shown in Figure 3. The solution of Equation (1) is:

$$r = -\frac{R_O}{S_O} z + C = -\frac{R_O}{S_O} z \qquad (2)$$

where the constant of integration is $C=0$ since the ray passes through the pinhole at $r=z=0$. This ray strikes the image plane, $z=S_I$, at the radius

$$r = R_I = -\frac{R_O}{S_O} S_I \qquad (3)$$

This can be obtained from equation (1) or it can be derived from the ratio of sides of similar triangles in the object and image spaces, Figure 4. Remember that this Figure only shown what is happening in the $\theta=\theta_O$ plane, as shown in Figure 3.

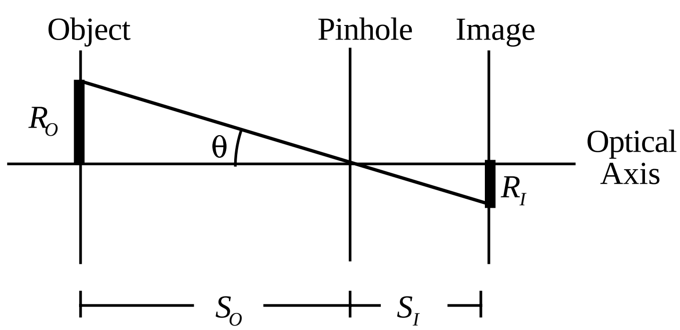


Figure 4 Geometry of a Pinhole Image

The linear magnification, the ratio of image size to object size, depends only on the ratio of image distance to object distance

$$M = \frac{R_I}{R_O} = -\frac{S_I}{S_O} \qquad (4)$$

The minus sign in Equation (4) tells us that the pinhole camera produces an inverted image. In the optics literature the minus signs are frequently omitted from equations (3) and (4), making *r* positive in both the object and image planes. In that case the image inversion is implicit. The reason for the sign uncertainty is that our analysis began with a polar coordinate system, in which the r coordinate is always positive However, when we confined our analysis to a plane of constant θ, we began using the cartesian coordinates of that plane in which negative values of r are both proper and necessary. Be careful when you are reading the optics literature, this distinction can cause confusion.

Each point in the object plane images into a point in the image plane that has the same θ coordinate and an r coordinate that is scaled by the ratio $-S_I/S_O$. Thus images are geometrically similar to objects; the pinhole images perfectly, without geometric distortion. The area of the image is therefore M^2 times that of the object. Similarity is guaranteed only when the object and image planes are parallel as we assumed. When they are not parallel, the object and image are no longer necessarily similar. We will discuss the general case later.

A consequence of perfect imaging is that off-axis objects form perfect off-axis images, Figure 5.

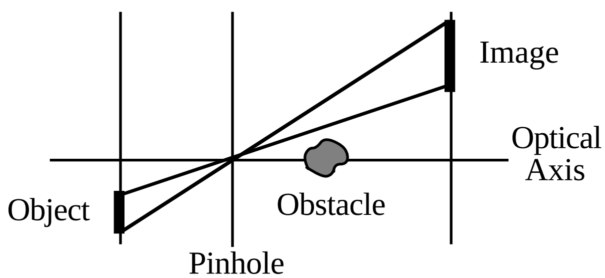


Figure 5. An off-axis object forms an off-axis image.

Thus, a lateral shift of the pinhole allows imaging without geometric distortion even though the camera cannot be placed directly in front of the object, Figure 6. This optical arrangement is often used in applied machine vision applications when it is necessary to fit a lot of equipment into a cramped space. It is also useful for making perfect (non-distorted) images of large objects that cannot be centered on the optical axis.

The focal properties above were derived for *meridional* rays. Those are rays which travel in planes of constant θ. All meridional planes include the optical axis. Rays that do not travel in planes of constant θ are called *skew rays.* They do not intersect the optical axis, cannot pass through the pinhole, and therefore do not contribute to the image in a pinhole camera.

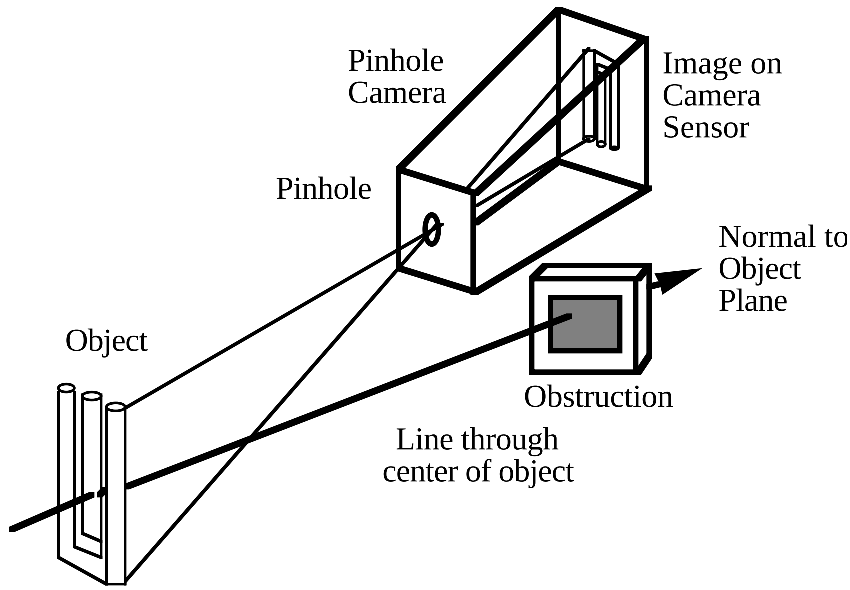


Figure 6. An example of off-axis imaging.

Perfect imaging implies perfect geometric similarity between objects and images, not perfectly sharp images. If the pinhole diameter is d, the light rays from an object point on the optical axis illuminate a circular area of diameter

$$d_b = d\frac{S_O + S_I}{S_O} = d(1 - M) \tag{5}$$

The image is unsharp and d_b is the on-axis blur circle size, which is always larger than the pinhole diameter, d; see Figure 7. An off-axis object point has a blur ellipse with major axis $d(1–M)$ and minor axis $d(1–M)\cos\theta$. This blur is the practical reason that pinholes are almost never used to form images for computer or machine vision applications.

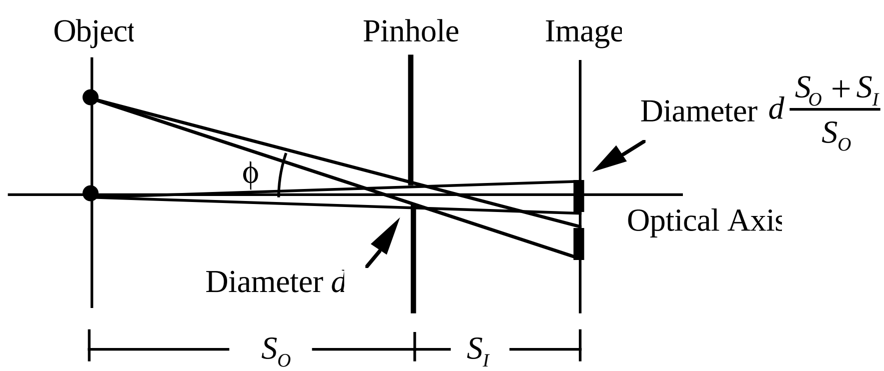


Figure 7 Pinhole Camera Blur

Pinhole Optics Depth of Field

With pinhole optics, the image can be formed at any distance along the optical axis, $z=S_I>0$. The magnification and image blur are determined uniquely by the relative sizes of the object and image distances, S_O to S_I

Also, since the image blur does not vary rapidly with small shifts in object or image location, the assumption that the object is planar can be relaxed. Notice, too, that the object and image are totally equivalent; they can be interchanged.

Pinhole Optics Illuminance

The axial image illuminance for the pinhole camera, E_I, is the total flux into an image region, such as the blur circle, divided by the area of that region,

$$A_I = \frac{\pi d^2 \left(S_O + S_I\right)^2}{4S_O^2} \tag{6a}$$

The total flux is the product of three terms: the object's axial luminance L_O, the area of the object that emits light into an image point,

$$A_O = \frac{\pi d^2 \left(S_O + S_I\right)^2}{4S_I^2} \tag{6b}$$

and the solid angle subtended by the pinhole at the object

$$\Omega_P = \frac{\pi d^2}{4S_O^2} \tag{6c}$$

Note that the similarity between Equations (6a) and (6b) is a direct consequence of the equivalence and interchangeability of the object and image. Combining Equations (6), the illuminance of an on-axis image point is

$$E_I = \frac{L_O A_O \Omega_P}{A_I} = \frac{\pi d^2 L_O}{4S_I^2} \tag{7}$$

Now we consider the illuminance of an off-axis image point due to the light received from the corresponding off-axis object point, as shown in Figure 6. Denote the object luminance in the direction of the pinhole by L(ϕ). Figure 6 shows that the pinhole's projected area decreases by a factor of cosϕ, the distance from the object point to the pinhole increases from S_O to $S_O/\cos\phi$, the distance from the pinhole to the image point similarly increases from S_I to $S_I/\cos\phi$, and the pinhole solid angle is

$$\Omega_P = \frac{\pi d^2 \cos\phi}{4\left(S_O^2 / \cos^2\phi\right)} \tag{8}$$

The image illuminance is

$$E_I = \frac{L(\phi) A_O \Omega_P}{A_I} = \frac{\pi d^2 L(\phi) \cos^3\phi}{4 S_I^2} \tag{9}$$

Note that the areas A_O and A_I scale by the same geometric factor so their ratio remains the same as in the axial case – the difference between Equations (7) and (9) is due completely to the change in the solid angle subtended by the pinhole, Equation (8).

If the object is a diffuse emitter, then it's luminance is described by, $L(\phi)=L_O\cos\phi$ (Lambert's law). The image illuminance becomes

$$E_I = \frac{\pi d^2 L_O \cos^4\phi}{4 S_I^2} \tag{10}$$

Because the angles are usually small, the factors of $\cos^3\phi$ and $\cos^4\phi$ in Equations (9) and (10) can often be neglected in machine vision optical calculations.

Thin Lens Optics

The pinhole optical system has the advantages of simplicity, lack of geometric distortion, and excellent depth-of-field (the blur changes slowly with changes in object or image distance). The major, and usually limiting, disadvantages are low image illuminance – a consequence of requiring that all image rays pass through a small pinhole, and a large blur circle – which can never be smaller than the pinhole. Enlarging the pinhole to increase the illuminance significantly increases the blur size. The usual solution is to replace the pinhole with a *lens*, an optical element with both a larger solid angle (greater illuminance) and a smaller blur circle than the pinhole.

As with the pinhole camera, we begin with a mathematical model for the lens and examine how to use the model. In lens models, we make an assumption that adequately describes systems made of simple glass lenses in air: light rays travel in straight lines except for when they are passing through lenses. Real lenses come in a variety of shapes and and a variety of lens models are used to analyze them. Where possible we prefer to use the mathematically simplest model, the rotationally symmetric *thin lens*, which is appropriate for round lenses whose thickness is small compared with the overall length of the optical system.

The thin lens is represented by a single plane perpendicular to the optical axis. In this plane, called the *principal plane*, a light ray is abruptly deviated such that its path is continuous but its slope r' changes by an amount proportional to the radius at which it strikes the plane. The constant of proportionality is $-1/f$, where f is a physical property of the lens called its *focal length*,

$$\Delta r' = -\frac{r}{f} \tag{11}$$

When f is positive the lens converges meaning that light rays are deviated toward the optical axis; a negative lens diverges the rays, Figure 8.

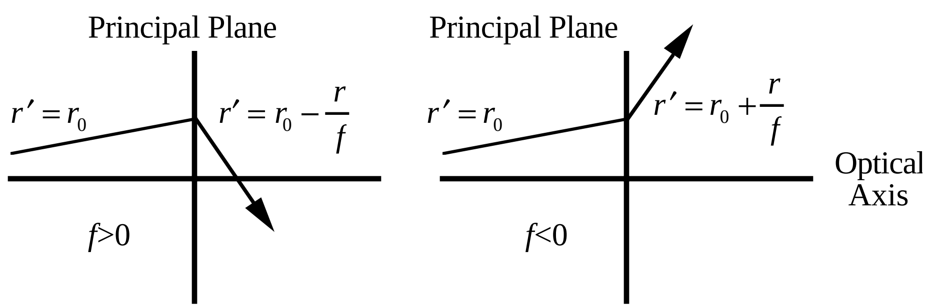


Figure 8 Converging and Diverging Lenses

We begin the analysis of the thin lens by selecting a coordinate system with its origin at the center of the principal plane. *Object space* is defined as the region $z<0$ which contains the object; the image is in the *image space* region $z>0$. In analyzing the optical properties of the thin lens, we will consider meridional rays first. Recall that in the thin lens model, rays travel in straight lines in both object and image space. We can represent them by linear equations in both regions.

$$\begin{aligned} r_O(z) &= A_O + B_O z, \qquad z<0 \\ r_I(z) &= A_I + B_I z, \qquad z>0 \end{aligned} \tag{12}$$

Using Equations (11) and (12) and requiring that r be continuous at z=0, we obtain the equations

$$
\begin{aligned}
&r_I(0) = r_O(0) = r(0) \Rightarrow A_I = A_O \\
&r_I'(0) = r_O'(0) - \frac{r(0)}{f} \Rightarrow B_I = B_O - \frac{A_O}{f}
\end{aligned}
\tag{13}
$$

Thus, if we know the values A_O and B_O, which describe the ray in object space, the values of A_I and B_I, which describe the ray in image space can be easily calculated using Equations (13). Values for the A_O and B_O are usually found from the ray's *initial conditions* in the object plane, $z=-S_O$.

The ray equations are linear so superposition applies: if $r_\alpha(z)$ and $r_\beta(z)$ are two rays with equations $r_{\alpha O}(z)$ and $r_{\beta O}(z)$ in object space and $r_{\alpha I}(z)$ and $r_{\beta I}(z)$ in image space, respectively, then the ray $r_O=r_{\alpha O}+r_{\beta O}$ in object space has as its image space continuation the ray $r_I=r_{\alpha I}+r_{\beta I}$. Note that although both r_α and r_β are meridional rays lying in planes of constant ϕ, they need not lie in the *same* ϕ plane.

From Equation (11), the ray deviation is proportional to radius. This causes all rays which enter the lens parallel to the optical axis to pass through a single point, the *focus*, located on the optical axis a distance f from the principal plane, Figure 9.

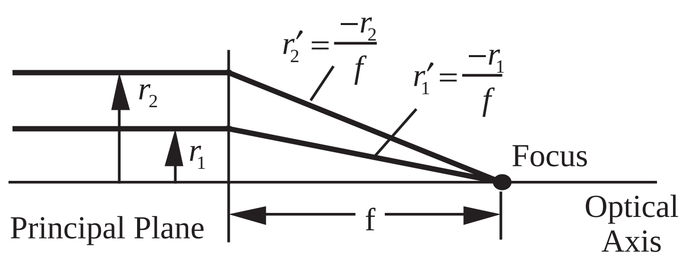


Figure 9 The Thin Lens Focal Point

Similarly, a collection of parallel rays striking the lens at some angle ϕ_O with respect to the optical axis will be focused at a point with radius $r=f\tan\phi_O$ in the plane $z=f$. If the angle ϕ_O is varied, the loci of all resulting points of focus describe a plane perpendicular to the optical axis called a *focal plane*, Figure 10 Since lens are symmetrical, each has two such focal planes symmetrically disposed about the principal plane.

One consequence of linearity, Equations (12), is that *any* two non-parallel rays form a basis set and any arbitrary ray can be formed from a linear combination of the two. We will examine the basis set formed by meridional rays, called *principal rays*. Consider an object located at $z=-S_O$ in an object plane perpendicular to the optical axis, Figure 11. One principal ray strikes the principal plane parallel to the optical axis and then passes through the image space focal point. The other passes through the object space focal point and then leaves the lens parallel to the optical axis. The intersection of these two rays forms an image point with radius $r = -R_I$ at a distance $z = S_I$ from the lens. Any ray leaving the object point that strikes the principal plane can be expressed as a linear combination of these two rays, so all rays from the object point are focussed at the image point. For each point in the object plane there is a corresponding point in the image plane where its rays are focussed. The lens is perfectly symmetric so the same results are obtained if object and image are interchanged.

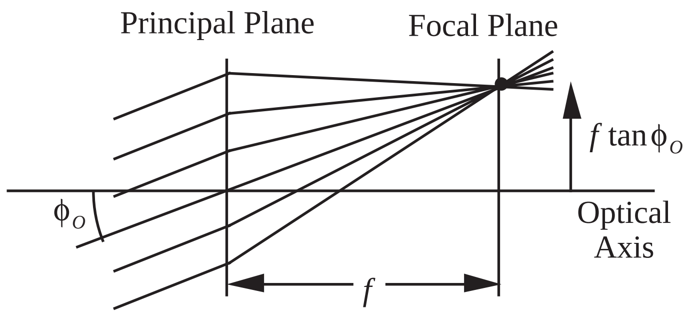


Figure 10. Rays that are parallel entering the lens converge in the focal plane.

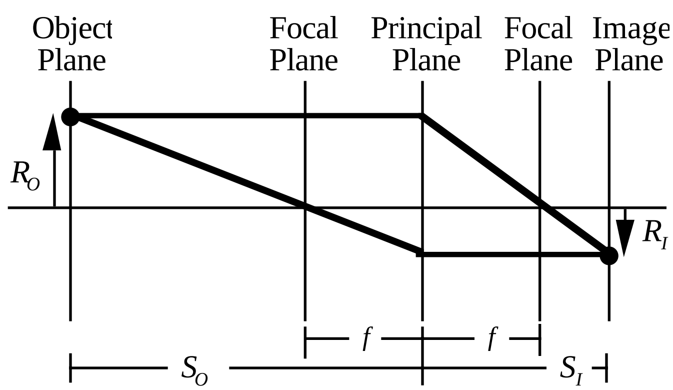


Figure 11 Image Formation in the Thin Lens

Example 1

Consider the example shown in Figure 10. Parallel rays strike the principal plane at the angle $\phi = \phi_O$. The ray that intersects the principal plane at the axis has radius zero at the principal plane so it is not, Equation 11. Thus, it continues on a further distance f along the z axis until it strikes the focal plane at a radius $r = f \tan\phi_O$ and slope $r' = \tan\phi_O$, as shown. Now any other of the parallel incident rays can be formed from the linear combination of this central ray and a second ray which is the principal ray that strikes the principal plane parallel to the optical axis and passes through the axis at the object side focal plane.

For example, the ray that strikes the principal plane at radius $r=r_O$ and can be decomposed into the two rays $r_1(z)$ and $r_2(z)$ whose values before striking the principal plane are:

$$r_1(z) = z\tan\phi_O, \qquad z < 0$$
$$r_2(z) = r(0) = r_O, \qquad z < 0$$

Using Equation (11) and requiring that the rays be continuous at z=0, we obtain the ray values after exiting the principal plane:

$$r_1(z) = z\tan\phi_O, \qquad z > 0$$
$$r_2(z) = r_O - z\frac{r_O}{f}, \qquad z > 0$$

Plotting the sum of these two functions $r(z)=r_1(z)+r_2(z)$ for various values of r_O produces the bundle of rays shown in Figure 10.

The focal properties of the thin lens were derived for meridional rays, which travel in planes of constant θ. The lens has rotational symmetry so this analysis applies to rays that travel in all such planes. A skew ray can be constructed as a linear combination of meridional rays. Therefore all rays from the object, including skew rays, are focussed at the image point. The assumption of constant θ can therefore be relaxed. With no loss of generality we will continue to show ray diagrams of meridional rays since they are much easier to draw and understand.

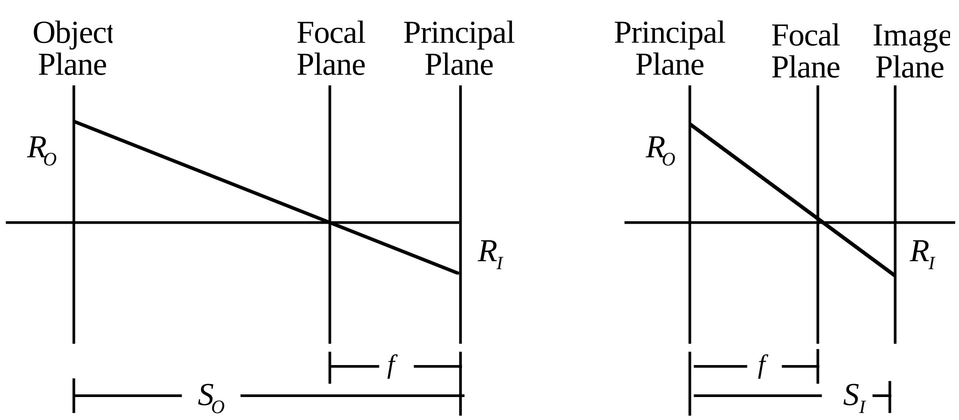


Figure 12. Triangles in Object and Image Spaces.

The principal rays form two sets of similar triangles, one each in object and image spaces, Figure 12.

Using the ratios of sides of similar triangles, the linear magnification, M, and the image distance, S_I, can be calculated as functions of the object distance, S_O, and lens focal length, f

$$M = \frac{R_I}{R_O} = \frac{f}{(S_O - f)} = \frac{(S_I - f)}{f} \tag{14}$$

$$\frac{1}{f} = \frac{1}{S_O} + \frac{1}{S_I} \tag{15}$$

Equation (15), called the *lens equation*, is the basis of all optical system design.

Note that the ray that passes through the center of the lens in Figure 10 is undeviated: since r=0 at the principal plan, $\Delta r'$=0 by Equation (11). Thus the magnification M for the thin lens is identical to the pinhole, Equation (3). This can be verified using Equations (14) and (15). Note that the magnification is again negative because R_O and R_I will have opposite signs, as will S_O and S_I. By the same arguments used in the pinhole optics case, objects and their images are geometrically similar in the thin lens case, as long as the object and image planes are parallel. Also, the lateral shift arrangement, Figures 6 and 7, can be used to provide distortion-free imaging with thin lenses. However, unlike pinhole optics, a lens does not allow arbitrary choice of both S_O and S_I only when they satisfy the lens equation, Equation (15), will the image be sharply focussed.

Most optical quantities can be calculated from Equations (14) and(15) so the optical properties of the thin lens are completely determined by the locations of three planes: the two focal planes and the principal plane. These can be measured by simple experiments. Rays parallel to the optical axis are sent into the lens from both directions. The planes in which they focus are the focal planes and the principal plane lies halfway between. In practice, the principal plane is usually near the geometric center of a lens.

Example 2

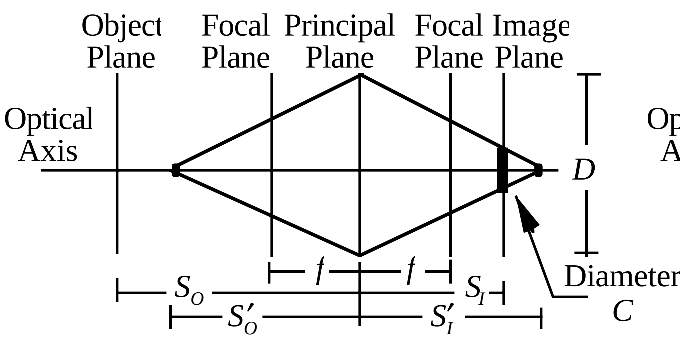
Consider the case of a 25mm lens that produces a focussed image of an object at infinity (very far away from the camera). The image is located at a distance f (25mm) from the principal plane. If, instead, the lens is used to form an image of an object placed 100mm from the principal plane, Equation (15) tells us that the image location will be:

$$S_I = \frac{1}{\frac{1}{f} - \frac{1}{S_O}} = \frac{1}{\frac{1}{0.025} - \frac{1}{0.1}} = \frac{1}{30} = 33.3\text{mm}$$

The image plane has moved 8.3mm (33.3mm – 25mm) farther from the lens. Extender tubes are often inserted between the lens and camera body to allow lenses to focus on nearby objects. The linear magnification of this optical arrangement will be, from Equation (3), M=–0.33. The minus sign means that the image is inverted.

Thin Lens Depth of Field

In the pinhole lens, the image plane is merely a tool for calculating optical properties: the image blur does not change greatly as the image plane is moved. In the thin lens model, the image is formed by the intersection of non-parallel rays so the blur can vary significantly with small movements of the image plane. The focal length is a constant for a lens. Thus, from Equation (15), an object placed on the optical axis at a point $z=-S_O'$, closer to the lens than S_O, forms an image in the plane at S_I' which is farther from the principal plane than S_I; if $|S_O'| < |S_O|$ then $|S_I'| > |S_I|$. The rays that converge at the plane S_I' have a diameter C when passing through the plane S_I. Similarly, if $|S_O'| > |S_O|$ then $|S_I'| < |S_I|$ and light rays from a point on the optical axis at S_O' form an image at S_I' and begin to diverge again, forming a circle of diameter C at the image plane, see Figures 13a and 13b.



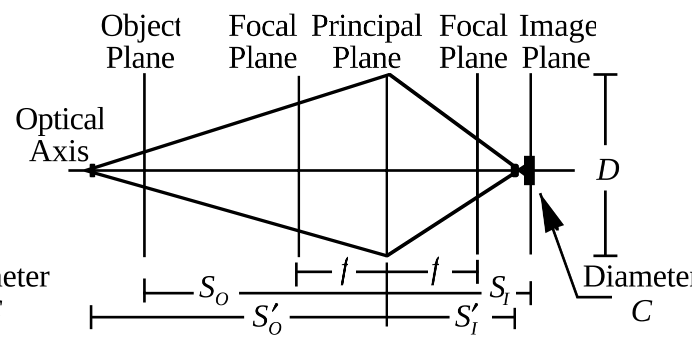


Figure 13a). Blur from an Object too Close to the Lens.
Figure 13b) Blur from an Object too Far from the Lens.

In these Figures, the lens diameter, D has been introduced. The blur size will be seen to be a function of this diameter. In most lenses, this is just the diameter of the glass lens elements or of the variable diameter *aperture* built into many lenses. The aperture size is often expressed as a dimensionless *F number*[1],

[1] Instead of F number, the numerical aperture is sometimes used in the optics literature: $n\sin\phi$, where n is the

$$N = \frac{D}{f} \tag{16}$$

Thus, a variable aperture lens that is set to *f*/5.6 (the arrow points to 5.6 on the barrel of the lens) has *F* number *N*=1/5.6=0.179.

The diameter of the blur circle can be found using similar triangles:

$$C = \frac{D(S_I' - S_I)}{S_I'}, \;\; S_O' < S_O \text{ and } S_I' > S_I$$
$$C = \frac{D(S_I - S_I')}{S_I'}, \;\; S_O' > S_O \text{ and } S_I' < S_I$$

These can be combined into a single equation,

$$S_I = S_I'\left(1 \pm \frac{C}{D}\right) = S_I' \pm \Delta S_I' \;\text{ or }\; S_I' = S_I \mp \Delta S_I' \tag{17}$$

where

$$\Delta S_I = S_I' \frac{C}{D} \approx S_I \frac{C}{D}$$

and where the + and – signs represent the cases of the object too close and too far, respectively.

From Equation (17) we can locate the corresponding object planes that bound the region within which the object will form an image with blur diameter less than *C*,

$$S_{O_{\text{min,max}}} = S_O \mp \Delta S_O \tag{18a}$$

where

$$\Delta S_O = \frac{-\Delta S_I}{M} \tag{18b}$$

Equations (18) apply only when *M* can be considered to be constant. This occurs in the limit

---

index of refraction outside the lens and ϕ is the cone angle of the ray that strikes an outermost edge of the lens. The reasons for having two measures are beyond the scope of these introductory notes.

$$\frac{\Delta S_I}{S_I} << 1 \text{ or } \frac{C}{D} << 1$$

This approximation is usually satisfied in practical computer or machine vision systems.

Example 3

As an example, consider a 50mm lens with *F* number *N*=*D*/*f*=1/2.8 used to form an image with magnification *M*=-0.1 on a camera sensor whose width is $2R_I$ =8mm wide and which has 400 pixels per row. The field of view of the camera is $2R_{O,max}=2R_{I,max}/M$=80mm, the pixel width is 8mm/400=20µm, and the lens diameter is *D*=*fN*=17.9mm. Thus if we select a design which produces a blur circle with the same diameter, then *C*/*d*=20µm/17.9mm=.0011, validating the approximation *C*/*d*<<1. The corresponding range of the object distance is

$$\begin{aligned} S_O' &= S_O\left(1 \pm \frac{C}{DM}\right) = f\left(\frac{M-1}{M}\right)\left(1 \pm \frac{C}{fNM}\right) \\ &= 550 \pm 6.2\text{mm} \end{aligned}$$

This 12.4mm band within which everything is considered to be in focus is called the *depth of field*. It is linearly proportional to the blur circle size, *C*, and inversely proportional to the lens *F* number, *N*.

If an exact solution for S_O' is required, rewrite Equation (15), noting that *f* is constant:

$$\frac{1}{f} = \frac{1}{S_O'} + \frac{1}{S_I'} = \frac{1}{S_O} + \frac{1}{S_I}$$

This equation can be solved for two values of S_O':

$$\frac{1}{S_O'} = \frac{1}{S_O M_0}\left[1 - \left(1 - \frac{1}{1 \pm \frac{C}{D}}\right)\right] \qquad (19a)$$

where

$$M_0 = \frac{-S_I}{S_O} \qquad (19b)$$

is the linear magnification at perfect focus. In the limit $C/D<<1$, Equations (19) reduce to Equations (18).

Parallel and Tilted Object and Image Planes

So far we have only considered the case where the object and image planes are parallel to the principal plane and to each other. In fact, this parallel plane assumption is not even necessary – we just used it to establish that there are two complementary planes, the object and image planes, with the property that every point in one images into a unique point in the other. In this section we will establish that these complementary planes exist, even when not parallel and we will examine the geometric nature of the imaging relationship between pairs of planes.

A point is specified by its (r,θ,z) coordinates. We showed previously that each point in object space (R_O,θ_O,S_O) images into its unique point inimage space (R_I,θ_O,S_I) where the variables are related by Equations (4), (14), and (15). The two points will have the same angular coordinate θ_O because that is the only condition that guarantees there will always be a meridional ray connecting the object and image point. If we restrict our analysis to only those object points lying in the plane of constant S_O, we can see that all of the corresponding image points lie in the plane of constant S_I – this is just the case of parallel object and image planes. Thus each object point (R_O,θ_O,S_O) images into the image point (MR_O,θ_O,MS_O) where M, the linear magnification, is constant for all points since they all share the same S_O and S_I values. This is sufficient to establish that the object and its image are geometrically similar: planar polygons image into similar planar polygons, parallel lines in the object form parallel lines in the image, lines in the image intersect if and only if they intersect in the object, ordering of vertices and edges is invariant between object and image, etc.

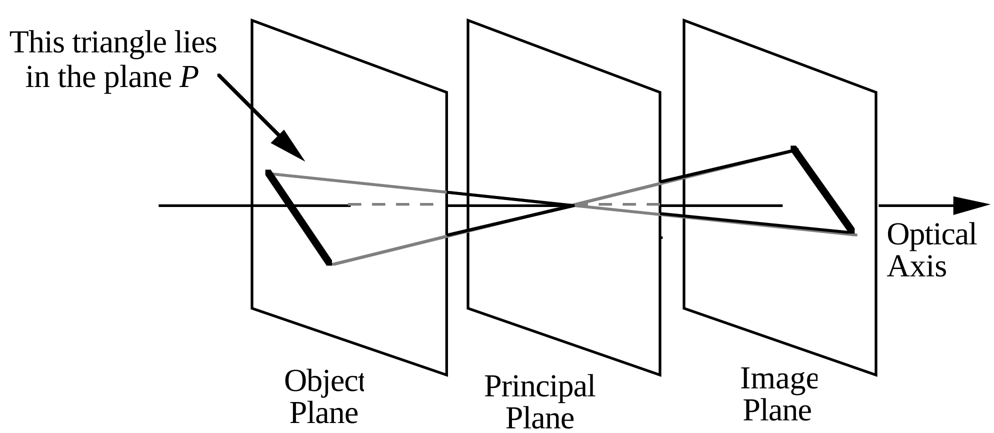


Figure 14. Lines in the object plane image into lines in the image plane.

You may wonder how we can be sure that object lines map into lines in the image and not, for example, points or curves. This can be seen by considering the object to be a line in the object plane and by looking only at those rays which pass through the origin at the center of the lens' principal plane, Figure 14.

The object line and the origin in the principal plane define a plane P in space. Every object point lies in P. But every image point lies on the ray connecting the corresponding object point with the origin and a line connecting two points in a plane lies within the plane so every image point lies in P. Thus every image point lies in the intersection of two planes, P and the image plane, and therefore the image is a line.

Figure 14 can also be used to show that any two object points plus the origin form a triangle that is similar to the triangle formed by the corresponding two image points and the origin. This is an alternative way to establish that objects and images are geometrically similar when the object and image planes are parallel.

Now consider the case of a object plane tilted at some arbitrary angle. Figure 15 shows a view of the plane containing both the optical axis and the normal to the object plane. In general, it is necessary to rotate the entire system about the optical axis to put it into this configuration in which the object plane projects into a line, as shown. Again, we will examine the behavior of meridional rays in this plane and use the fact that skew rays can be constructed from meridional rays to generalize the results to three dimensions. Look at the ray from a object point P_O which travels *in* the object plane toward the point P_P where it intersects the principal plane and deflects to pass through the image point P_I. The ray from any other object point that travels in the object plane also passes through the points P_O and P_P and, thus, through P_I as well. Thus every point in the tilted object plane images into a plane containing the points P_P and P_I This is sufficient to establish that a tilted object plane images into a tilted image plane and that the object, principal, and image planes intersect in a common line perpendicular to the optical axis. Thus the normal to theimage plane also lies in the same plane as the optical axis and normal to the object plane. Tilting the object and image planes so they both intersect the principal plane in the same line is sufficient to assure perfect focus from every point in the object plane into a point in the image plane.

This last condition, called the *Scheimpflug Condition*, is used in computer and machine vision applications to produce a sharply focussed image of a surface even when the optical axis cannot be arranged normal to the surface, Figure 16. Typically the camera is placed as close a possible to the normal of the surface being observed. Then the sensor and lens are tilted until the sensor plane, the lens plane, and the surface intersect in a single line.

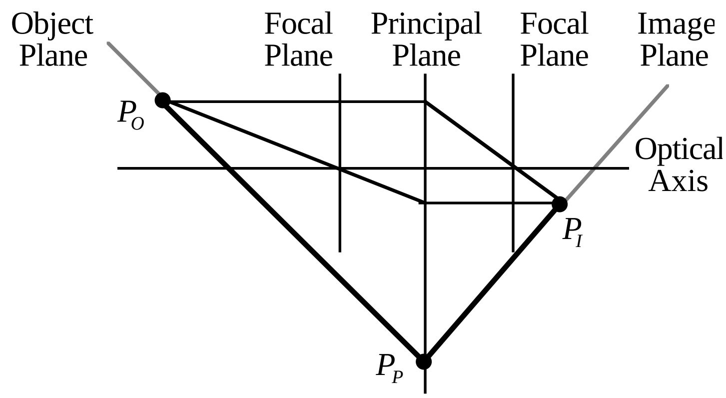


Figure 15. Titled object and image planes.

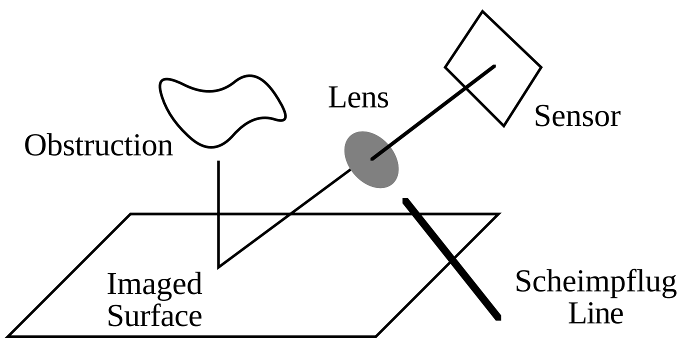


Figure 16. Tilted optics used to image perfectly using a tilted camera.

While the titled optics method can provide a perfectly focussed image of a tilted object plane, that image will not necessarily be geometrically similar to the object. From Figure 14 and the text explaining it, we can see that while lines still image into lines, geometric similarity is assured only when the object and images are parallel. As a final comment, note that the Schempflug condition is completely specified by the location of a single object point (the on-axis point is usually used) and the line of intersection in the principal plane (the point on the line which passes closest to the axis is usually used).

Three Dimensional Objects and Images

So far we have considered objects and images to be two-dimensional objects, although the planes in which they lie need not be parallel. This restriction is reasonable when applied to images since all image sensors, including the eye, are two-dimensional devices. Objects, however, are generally three-dimensional so it will be useful to consider what happens when a three-dimensional object is imaged by a thin lens.

At this point it will be useful to introduce a cartesian coordinate system with its z axis aligned with the optical axis, its origin in the center of the principal plane, and the directions chosen so that z<- in object space and z>0 in image space. The direction of the x axis is arbitray. The ray from a point P_O in object space can be decomposed into meridional rays in the x-z and y-z planes so the above analysis applies directly. The point P_O thus images with sharp focus into the point P_I

$$P_O = (x, y, z) \tag{20a}$$

$$P_I = \left(Mx, My, Mz\right) \tag{20b}$$

where

$$M = \frac{f}{z - f} = \frac{1}{\frac{1}{f}z - 1} \tag{21}$$

Three-dimensional objects thus transform into sharply focussed three-dimensional image points, forming a three-dimensional image. This is only a virtual image – it doesn't exist and we can't see anything unless we put some surface into image space to intercept the rays. The resulting real image on that two-dimensional surface is just a standard planar image so we know what to expect – mostly blurred rays with a few sharply focussed ones that happen to have originated in the corresponding object plane.

Equations (20) and (21) are the basis of a popular image analysis technique. The light hitting a known point P_I=(x,y,z) in a known image plane must have originated from an object point P_O=(αx,αy,αz). If one of the coordinates of the object point is known or can be guessed, the other two can be calculated. This analysis technique, although sometimes useful, is quite limiting. We will look at a more powerful technique, the concept of a virtual three-dimensional image.

Equations (21) show that the object and image spaces can be related by a transformation of coordinates that is affine and expressible using homogeneous coordinates:

$$\mathbf{P_I} = \mathbf{T}\mathbf{P_O}, \quad \mathbf{P_O} = \mathbf{T}^{-1}\mathbf{P_I} \tag{22a}$$

where the points are column m atrices of the form

$$\mathbf{P} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

and **T** is the transformation matrix

$$\mathbf{T} = \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & -1 \end{bmatrix} \tag{22b}$$

The fact that **T** is its own inverse is a consequence of the symmetrical nature of the thin lens.

Example 4

The location and orientation of a lens are fixed. For example, the lens is mounted in a port in the side wall of an environmental chamber, Figure 16a. Our goal is to form an image on a camera sensor of the inside of the chamber in which three arbitrary points, *A*, *B*, and *C* are in sharp focus. In this example we show how to select the lens focal lengths and the sensor orientation and location to achieve this goal.

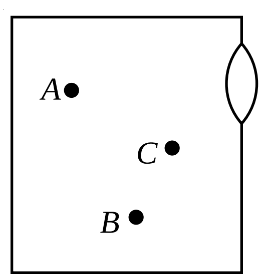


Figure 16a. Schematic drawing of an enclosuse with a fixed lens.

The first step is to select a coordinate system. As usual, place the origin at the center of the lens with the *z* axis perpendicular to the lens plane (the chamber wall, in this case). Next, find the object plane containing the points *A*, *B*, and *C*. Represent each point by the vector to it from the origin

$$\mathbf{A} = (x_A, y_A, z_A), \quad \mathbf{B} = (x_B, y_B, z_B), \quad \mathbf{C} = (x_C, y_C, z_C)$$

Calculate the normal to the plane

$$\mathbf{N_O} = (\mathbf{A} \times \mathbf{B}) + (\mathbf{B} \times \mathbf{C}) + (\mathbf{C} \times \mathbf{A}) \tag{23a}$$

Note, if this normal points in the *z* direction, then the object and image planes are parallel to each other and to the principal plane, the case we already solved above. If the normal is not directed along the *z* asis we next calculate the angle about the optical axis of the projection of the normal into the principal plane

$$\theta_z = \tan^{-1} \frac{\mathbf{N_O} \bullet y_{\mathbf{N_O}}}{\mathbf{N_O} \bullet x_{\mathbf{N_O}}} \tag{23b}$$

It is important that this be done as a four-quadrant calculation: $0 \le \theta_z \le 2\pi$.

Now rotate everything in the problem by multiplying each point and vector by the rotation matrix

$$\mathbf{R_z}(-\theta_z) = \begin{bmatrix} \cos\theta_z & \sin\theta_z & 0 & 0 \\ -\sin\theta_z & \cos\theta_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{23c}$$

This has the effect of rotating the optics into the configuration of Figure 15. We will assume this has been done but will not change from the original notation for clarity.

The the equation for any point **P** in the plane is

$$(\mathbf{P} - \mathbf{P_O}) \bullet \mathbf{N_O} = 0 \tag{24a}$$

where $\mathbf{P_O}$ is any known point in the plane (**A**, **B**, or **C**) and rotated values are used for all vectors.

The equation for any point in the lens principal plane is

$$\mathbf{P} \bullet \hat{\mathbf{z}} = 0 \Rightarrow z_{\mathbf{P}} = 0 \tag{24b}$$

where $\hat{\mathbf{z}}$ is the unit vector in the *z* direction, the direction of the optical axis. We also know that normal has no *x* component and the intersection line has to point in the *x* direction as a consequence of the rotation.

Combining Equations (24), we obtain the equation for the line of intersection

$$y_{\mathbf{P}} y_{\mathbf{N_O}} = y_{\mathbf{P}}\left(\hat{\mathbf{y}} \bullet \mathbf{N_O}\right) = \mathbf{A} \bullet \mathbf{N_O} = \mathbf{A} \bullet (\mathbf{B} \times \mathbf{C})$$

The intersection line has constant y value

$$y_{\mathbf{P}} = \frac{\mathbf{A} \bullet (\mathbf{B} \times \mathbf{C})}{\left(\hat{\mathbf{y}} \bullet \mathbf{N_O}\right)} \tag{25a}$$

and the point of closest approach to the optical axis is

$$\mathbf{P_P} = (0, y_{\mathbf{P}}, 0) \tag{25b}$$

The vector equation for the intersection line is

$$\left(\mathbf{P} - \mathbf{P_P}\right) \times \hat{\mathbf{z}} = 0 \tag{25c}$$

The on-axis object and image points can be found from Equations (15) and (24a)

$$\mathbf{S_O} = \left(0, 0, S_O\right), \quad \mathbf{S_I} = \left(0, 0, S_I\right) \tag{25d}$$

where

$$S_O = \frac{\mathbf{A} \bullet \mathbf{N_O}}{\hat{\mathbf{y}} \bullet \mathbf{N_O}}, \quad S_I = \frac{f S_O}{S_O - f} \tag{25e}$$

Sharp focus is obtained by placing the camera sensor in the plane containing the points $\mathbf{P_P}$ and $\mathbf{S_I}$. As a final part of this example, let's calculate the affine transformation between the tilted object and image planes.

A three step process is used to transform a point $\mathbf{P_O}$ expressed in the object plane coordiantes (s,t) into an point $\mathbf{P_I}$ expressed in the image plane coordinates (u,v). First, the transformation is calculated between s,t coordinates and x,y,z coordinates

$$\mathbf{T_O} = \mathbf{R_x}\left(\tan^{-1}\frac{y_\mathbf{P}}{S_O}\right)\mathbf{S_z}(S_O) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{S_O}{\sqrt{y_\mathbf{P}^2 + S_O^2}} & \dfrac{-y_\mathbf{P}}{\sqrt{y_\mathbf{P}^2 + S_O^2}} & 0 \\ 0 & \dfrac{y_\mathbf{P}}{\sqrt{y_\mathbf{P}^2 + S_O^2}} & \dfrac{S_O}{\sqrt{y_\mathbf{P}^2 + S_O^2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & S_O \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{26a}$$

Next, the transformation is calculated between *x,y,z* coordinates and *u,v* coordinates

$$\mathbf{T}_I = \mathbf{S_z}(-S_I)\mathbf{R_x}\left(\tan^{-1}\frac{-y_\mathbf{P}}{S_I}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -S_I \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{-S_I}{\sqrt{y_\mathbf{P}^2 + S_I^2}} & \dfrac{y_\mathbf{P}}{\sqrt{y_\mathbf{P}^2 + S_I^2}} & 0 \\ 0 & \dfrac{-y_\mathbf{P}}{\sqrt{y_\mathbf{P}^2 + S_I^2}} & \dfrac{-S_I}{\sqrt{y_\mathbf{P}^2 + S_I^2}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{26b}$$

Combining these results with Equation (22b), we obtain the complete transformation

$$\mathbf{P_I} = \mathbf{T_I}\mathbf{T}\mathbf{T_O}\mathbf{P_O}, \quad \mathbf{P_O} = \begin{bmatrix} s \\ t \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{P_I} = \begin{bmatrix} u \\ v \\ 0 \\ 1 \end{bmatrix} \tag{26c}$$

This transformation can be inverted to calculate the transformation between image and object spaces.

Thin Lens Illuminance

We stated above that lenses are used instead of pinholes because they form sharper images. There is another reason, too. The illumination in the image comprises only that portion of the light emitted from the object that is directed toward the open area of the pinhole. By using a lens which images all of the light emitted into the much larger area of the lens, we can predict that the image illuminance will increase dramatically. In this section we show how to calculate the image illuminance and will validate that prediction.

We begin by calculating the axial image illuminance. An object region of area A_O emits rays that are focussed into an image region with area A_I=M^2A_O, Figure 17.

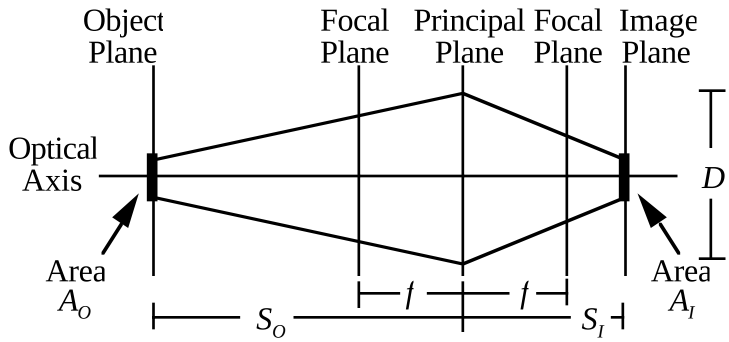


Figure 17 Image Illuminance

The image illuminance E_I is the total flux divided by the image area. The flux is the product of: the object's axial luminance, L_O, the object area, A_O, and, the solid angle subtended by the lens,

$$\Omega_L = \frac{\pi D^2}{4S_O^2}$$

Thus the image illuminance is

$$E_I = \frac{L_O A_O \Omega_L}{A_I} = \frac{\pi D^2 L_O}{4M^2 S_O^2} = \frac{\pi D^2 L_O}{4S_I^2} \tag{27}$$

Notice that this equation is similar to Equation (7) except that the pinhole diameter has been replaced by the lens diameter, showing that the increased illuminance is a result of being able to use larger lenses than pinholes, because focussing eliminates image blur.

Using Equations (14), (15) and (16), we can transform Equation (26) into a more useful form which incorporates the lens *F* number *N*

$$E_I = \frac{\pi D^2 L_O}{4 f^2 (1-M)^2} = \frac{\pi N^2 L_O}{4(1-M)^2} \tag{28}$$

Image illuminance is proportional to the object's axial luminance. The proportionality factor is a function of the lens *F* number, which can usually be read off the side of the lens barrel, and the magnification, which can be calculated by dividing the image size by the object size. In computer or machine vision applications incorporating solid state cameras, the image size is usually known in pixels, whose physical size is part of the camera specification.

Example 5
In this example we extend the calculations used in Example 3 to compare the performance of a thin lens with a pinhole lens. We begin with image illuminance. We cannot use Equation (28) directly for the pinhole camera which has no focal length f or F number N so we will compare the results of Equations (7) and (27) instead.

Assume the blur circle is equal to the size of a sensor pixel in a solid state camera,20μm, and that the object distance remains S_O=550mm. For the pinhole camera, the image illuminance is

$$E_{\text{pinhole}} = \frac{\pi d^2 L_O}{4M^2 S_O^2} = \frac{\pi(20\mu\text{m})^2 L_O}{4(-0.1)^2(550MM)^2} = 1.04 \times 10^{-7} L_O$$

For the thin lens, the image illuminance is

$$E_{\text{thin lens}} = \frac{\pi D^2 L_O}{4M^2 S_O^2} = \frac{\pi(19.7\text{mm})^2 L_O}{4(-0.1)^2(550MM)^2} = 0.083 L_O$$

Thus the increase in the image illuminance from the thin lens over that of the pinhole lens is

$$\left[\frac{D_{\text{thin lens}}}{d_{\text{pinhole}}}\right]^2 \approx 800{,}000$$

That is why lenses are used.

Now look at the thin lens image illuminance due to an off-axis object point. Note that no assumptions were made about the nature of the imaging system during the derivation of Equations (9) and (10); they are a geometrical consequence of Lambert's law and of requiring all of the imaging rays to pass through a hole (a pinhole or a lens aperture). Thus the cos $^3\theta$ falloff applies to the thin lens as well,

$$E_I = \frac{\pi N^2 L(\phi)\cos^3\phi}{4(1-M)^2} \qquad (29a)$$

In the case of a diffuse object, the image illuminance is

$$E_I = \frac{\pi N^2 L_O \cos^4 \phi}{4(1-M)^2} \tag{29b}$$

In using either Equations (28) or (29), one should note that in most computer or machine vision applications the magnification is small, $|M|<1$, so the factor of $1/(1-M)^2$ can often be ignored. Then, since $E_O=\pi L_O$ for a diffuse object, Equations (29) reduce to the following approximations:

$$E_I = \frac{\pi N^2 L(\phi) \cos^3 \phi}{4} \tag{30a}$$

in general, or,

$$E_I = \frac{N^2 E_O \cos^4 \phi}{4} \tag{30b}$$

for a diffuse object.

For small angles the factors of $\cos^3\theta$ and $\cos^4\theta$ are usually close enough to one to be ignored..

Example 6
Assume that an object with 50% diffuse reflectivity is illuminated with 1000lux, typical room lighting. The object's luminance is

$$E_O=(0.5)(1000\text{lux}) = 500\text{lux}$$

If a camera with an f/2.8 lens observes this object, the sensor illuminance is

$$E_{\text{sensor}} = \frac{\pi L_O N^2}{4} = 15.9\text{lux}$$

assuming M=0, which corresponds to a distant object. This illuminance value is well within the sensitivity range of most solid state cameras.

Thick Lens Optics

The thin lens model is adequate for understanding the first-order optical and illumination properties of lenses. However, it is not particularly good when used to describe most practical video camera lenses which typically will contain several glass elements. The model used to describe the first-order properties of these lenses is called the *thick lens.*

The thin lens model is extended to include two principal planes, one associated with the object space focal plane and one associated with the image space focal plane. The principal planes are separated by a distance Δ, which can be positive or negative, Figure 18.

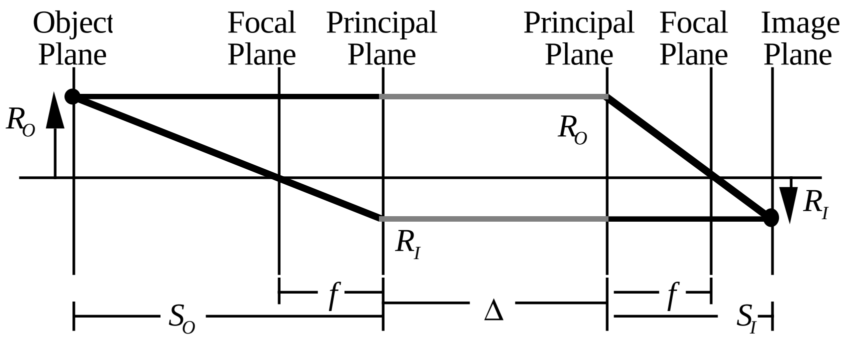


Figure 18. Image Formation in the Thick Lens

In this model, rays travel parallel to the optical axis between principal planes so the geometry of Figures 11 and 12 and Equations (14) and (15), which were derived from Figure 12, apply to the thick lens. In fact, the only difference between the thin and thick lenses is that the separation between two focal planes is not necessarily 2f as it always is in the thin lens model.

Many practical lens designs make use of the ability to control the principal plane separation, Δ. When the object principal plane is closer to the image focal plane (Δ<0), the lens has crossed principal planes. The optical advantage of this arrangement is that the overall lens length (glass to image distance) is less than for an equivalent thin lens. These compact lenses are called *telefocal lenses*. Some lenses even allow Δ to be varied without changing the focal length or image plane. They can be focussed over a large range of object distances and are the most widely used lenses in machine vision.

In the thick lens, the locations of four planes totally characterize the optical properties: two focal planes and two principal planes. The focal plane locations are measured in the same way as they were for the thin lens – parallel incident rays are used to find the planes of focus. Conceptually, the focal length, from which the principal plane locations are determined, is measured by dividing the radius at which a ray enters the lens parallel to the optical axis by its slope when it passes through the focal plane, see Figure 9.

Thick lenses are usually made from combinations of thin lenses. The most straightforward method of analyzing these combinations is to calculate the image from the first lens, use that image as the object for the second lens, etc. For example, consider the compound lens formed of two thin lenses with focal lengths f_1 and f_2 whose principal planes are a distance L apart, Figure 19. The diagram shows the case where L is large ($L > f_1 + f_2$). Calculate the effective focal length, f, by following a ray that enters the lens at a radius R_O parallel to the axis.

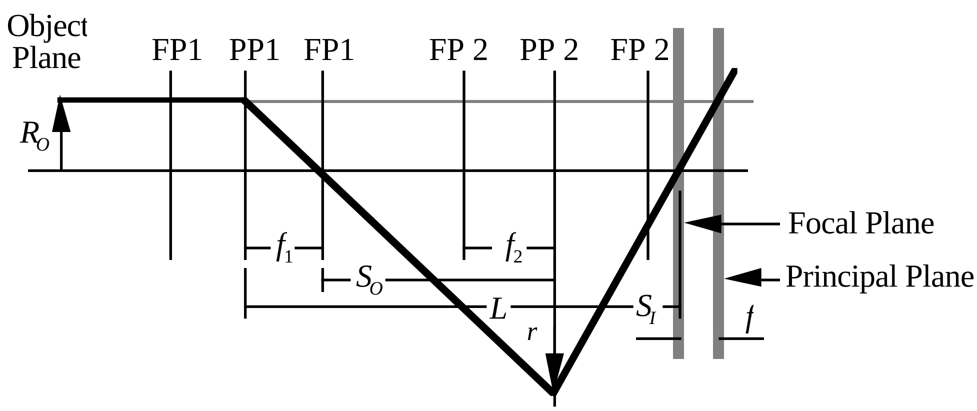


Figure 19 Focal Length of a Compound Lens

The ray leaves the first lens' principal plane with slope

$$r'_O = \frac{-R_O}{f_1} \tag{31}$$

The ray crosses the optical axis in the focal plane of the first lens and strikes the principal plane of the second lens at radius

$$r = -R_O \frac{L - f_1}{f_1} \tag{32}$$

The image of the first axis crossing at $S_O = L - f_1$ is another axis crossing at a distance S_I beyond the principal plane of the second lens. S_I is related to S_O by Equation (15). The slope at the second axis crossing is proportional to the radius r,

$$r'_I = \frac{-r}{S_I} \tag{33}$$

The effective focal length of the combined lenses is the object radius divided by the exit slope,

$$f = \frac{-R_O}{r'_I} \tag{34}$$

Combining equations (15), (32), (33), and (34), the focal length can be calculated

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{L}{f_1 f_2} \tag{35}$$

This expression can be shown to be valid for all values of L greater than zero, even for the case where the thin lenses are very close together. For the case shown in Figure 19 ($L > f_1 + f_2$) this lens diverges, $f<0$, which corresponds to the focal plane of the equivalent lens lying to the left of its principal plane. The focal plane is the plane in which the light ray crosses the axis for the second time since the ray was initially parallel to the axes. The principal plane of the equivalent lens lies at the intersection of the incident ray and the final ray, a distance f to the right of the second axis crossing. These planes are shown in hatched lines in Figure 19. When the lenses are close together ($L < f_1 + f_2$), Equation (35) is still correct, but $f>0$, the lens converges, and the principal plane is to the left of the focal plane. The other focal plane and principal plane can be found by tracing an analogous ray that travels through the system in the opposite direction. Once the two principal planes and focal planes are known, the compound lens can be considered to be a single thick lens that obeys Equations (14) and (15). Tilted object and image planes obey the extension of the thin lens case, Figure 15. The object plane and the object-side principal plane intersect in a line that is parallel to and at the same radius as the line where the image plane and image-side principal plane intersect.

The technique of using the image formed by the first lens (even if it is only a virtual image - one that is never actually formed because an optical component gets in the way) as the object of the second lens is the general method used for solving problems involving multiple lenses.

Once the thick lens has been replaced by the equivalent thin lens, illuminance calculations are done as described for the thin lens above.

References

There are many excellent references on the subject of geometrical optics - Goodman has prepared a bibliography[2] that lists over one thousand references! There are many optics text books that contain useful information. Books that have proven to be particularly valuable include those by Ray[3], Franke[4], Goodman[5], and Fincham and Freeman[6]. Handbooks,

[2] DS Goodman, **Bibliography of Classical Optics**, Unpublished Report, prepared at the IBM Research Center, Yorktown Heights NY 10598.

[3] SF Ray, **The Photographic Lens** (1979, Focal Press, London). This book is designed for users of lenses, not optics experts. It contains much valuable information and its explanations are easy to understand.

[4] G Franke, **Physical Optics in Photography** (1966, Focal Press, London).

[5] J Goodman, **Introduction to Fourier Optics** (). This book deals with the wave theory of optics, not the

such as the one published by the Optical Society of America[7], contain useful results, but not necessarily detailed derivations. Optics material can also be obtained from manufacturers of lens and other optical equipment. Finally, Newton's wonderful book on optics[8] is recommended to anybody who works with light.

---

geometric theory used in this paper, but it is a valuable reference book, particularly if diffraction effects and aberrations become important.

[6] WHA Fincham and MH Freeman, **Optics**, 9th ed. (1980, Butterworths, London).

[7] WG Driscoll, ed., **Handbook of Optics** (1978, McGraw-Hill, NY).

[8] I Newton, **Opticks**, 4th ed., 1730 (reprinted 1952, Dover, NY).