A Realistic Camera Model for Computer Graphics

Craig Kolb
Computer Science Department
Princeton University

Don Mitchell
Advanced Technology Division
Microsoft

Pat Hanrahan
Computer Science Department
Stanford University

Abstract

Most recent rendering research has concentrated on two subproblems: modeling the reflection of light from materials, and calculating the direct and indirect illumination from light sources and other surfaces. Another key component of a rendering system is the camera model. Unfortunately, current camera models are not geometrically or radiometrically correct and thus are not sufficient for synthesizing images from physically-based rendering programs.

In this paper we describe a physically-based camera model for computer graphics. More precisely, a physically-based camera model accurately computes the irradiance on the film given the incoming radiance from the scene. In our model a camera is described as a lens system and film backplane. The lens system consists of a sequence of simple lens elements, stops and apertures. The camera simulation module computes the irradiance on the backplane from the scene radiances using distributed ray tracing. This is accomplished by a detailed simulation of the geometry of ray paths through the lens system, and by sampling the lens system such that the radiometry is computed accurately and efficiently. Because even the most complicated lenses have a relatively small number of elements, the simulation only increases the total rendering time slightly.

CR Categories and Subject Descriptors: I.3.3 [Computer Graphics]: Picture/Image Generation; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism.

Additional Key Words and Phrases: ray tracing, camera modeling, lens simulation, sampling.

1 Introduction

The challenge of producing realistic images of 3d scenes is often broken into three subproblems: modeling reflection to account for the interaction of light with different materials, deriving illumination algorithms to simulate the transport of light throughout the environment, and modeling a camera that simulates the process of image formation and recording. In the last several years the majority of the research in image synthesis has been concentrated on reflection models and illumination algorithms. Since the pioneering work by Cook *et al.*[2] on simulating depth of field and motion blur, there has been very little work on camera simulation.

Although current camera models are usually adequate for producing an image containing photographic-like effects, in general they are not suitable for approximating the behavior of a particular physical camera and lens system. For instance, current models usually do not correctly simulate the geometry of image formation, do not properly model the changes in geometry that occur during focusing, use an improper aperture in depth of field calculations, and assume ideal lens behavior. Current techniques also do not compute exposure correctly; in particular, exposure levels and variation of irradiance across the backplane are not accounted for.

There are many situations where accurate camera models are important:

- One trend in realistic computer graphics is towards physically-based rendering algorithms that quantitatively model the transport of light. The output of these programs is typically the radiance on each surface. A physically-based camera model is needed to simulate the process of image formation if accurate comparisons with empirical data are to be made.
- In many applications (special effects, augmented reality) it is necessary to seamlessly merge acquired imagery with synthetic imagery. In these situations it is important that the synthetic imagery be computed using a camera model similar to the real camera.
- In some machine vision and scientific applications it is necessary to simulate cameras and sensors accurately. For example, a vision system may want to test whether its internal model of the world matches what is being observed.
- Many users of 3d graphics systems are very familiar with cameras and how to use them. By using a camera metaphor the graphics system may be easier to use. Also, pedagogically it is helpful when explaining the principles of 3d graphics to be able to relate them to real cameras.

Perhaps the earliest introduction of a camera model in computer graphics was the *synthetic camera model* proposed in the CORE system[3]. This and later work used a camera metaphor to describe the process of synthesizing an image, but did not intend to reproduce photographic effects or provide photographic-like control over image formation. The next major breakthrough in camera modeling was the simulation of depth of field and motion blur[10][2][12]. Current methods for simulating these effects use idealized lens systems and thus cannot be used to simulate the behavior of a particular physical system. A number of researchers have shown how to perform non-linear camera projections, such as those for fisheye or OMNIMAX lenses[7][5]. These methods derive a transformation that maps image points to directions in 3D, and have the disadvantage that effects such as depth of field cannot be combined with these special-purpose projections.

radius	thick	n_d	V-no	ap
58.950	7.520	1.670	47.1	50.4
169.660	0.240			50.4
38.550	8.050	1.670	47.1	46.0
81.540	6.550	1.699	30.1	46.0
25.500	11.410			36.0
	9.000			34.2
-28.990	2.360	1.603	38.0	34.0
81.540	12.130	1.658	57.3	40.0
-40.770	0.380			40.0
874.130	6.440	1.717	48.0	40.0
-79.460	72.228			40.0

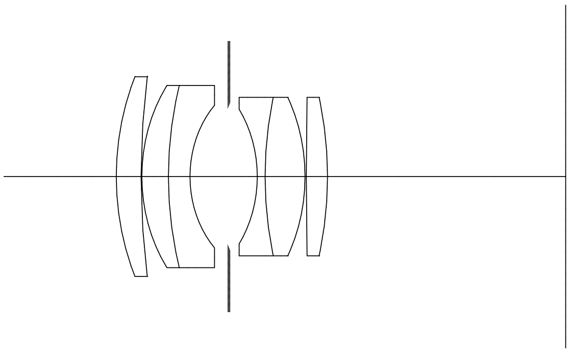


Figure 1: A tabular description and profile view of a double-Gauss lens. [14, page 312]. Each row in the table describes a surface of a lens element. Surfaces are listed in order from the front (nearest object space) to rear (nearest image space), with linear measurements given in millimeters. The first column gives the signed radius of curvature of a spherical element; if none is given, the surface is planar. A positive radius of curvature indicates a surface that is convex when viewed from the front of the lens, while a negative radius of curvature is concave. The next entry is thickness, which measures the distance from this surface to the next surface along the central axis. Following that is the index of refraction at the sodium d line (587.6 nm) of the material on the far side of the surface (if none is given, the material is assumed to be air). Next is the V-number of the material, characterizing the change of index of refraction with wavelength. The last entry is the diameter, or aperture, of each lens element. The row with a missing radius signifies an adjustable diaphragm; the diameter gives the size of the diaphragm when fully open. Note that if a surface separates two materials other than air, this indicates that two lenses have been cemented together as part of a "group." The lens as given has a focal length of approximately 100mm. The design may be changed to have any desired focal length by scaling each of the linear dimensions by the desired focal length divided by 100. The profile view on the right shows a 50mm version of this lens in relation to the diagonal of a piece of 35mm film.

This paper describes a physically-based camera model for computer graphics. The model is capable of simulating the image formation of a particular physical lens system described by the arrangement of simple lenses as specified by the manufacturer. Image formation is simulated by a modified distributed ray tracing algorithm that traces rays through the lens system in order to compute the exposure on the film plane. This algorithm is a hybrid of rendering techniques used by the computer graphics community and techniques used by lens makers to design camera lenses. Tracing rays through the lens system has the advantage that both the geometry and the radiometry of image formation can be accurately modeled. Moreover, we show that this simulation costs little more than previous algorithms.

For the purposes of this paper, our emphasis is on simulating the lens system, and as such the important effects caused by film response, shutter shape and movement, filters, and other parts of the camera will not be addressed here. We will further assume that the system is "aberration-limited," and so the effects of diffraction can be ignored.

The paper begins with a discussion of the construction of lenses and how they are modeled in our system. We then consider the various geometrical factors that effect image formation and how those factors can be accurately accounted for. The radiometry of image formation and its computation are then presented. Finally, results of an implementation of our model are shown and discussed.

2 Lens Systems

Lens systems are typically constructed from a series of individual spherical glass or plastic lenses and stops centered on a common axis. A stop is an opaque element with a roughly circular opening to permit the passage of light. The element that most limits the angular spread of the bundle of rays that will pass unobstructed through the system from the axial point on the image plane is termed the *aperture stop*. The size of the aperture stop in a camera is typically set by the photographer through the use of an adjustable diaphragm, and serves to provide control over the quantity of light striking the film plane and the depth of field in the image.

As shown in Figure 1, the construction of a lens is traditionally presented in a tabular format[1]. Our system reads tables like these and uses the information to model the behavior of the lenses they describe. Lens manufacturers are reluctant to release lens design data, but it is possible to find tables in patents that might cover a particular lens, or in collections of lens designs such as those given in the book by Smith[14].

There are two challenges to simulating a real lens system:

- *The geometry of image formation must be correctly computed.*

  Ideally, a lens will cause a point in object space to be imaged as a single point in image space, and will have constant magnification over the entire field of view. This is the assumption that is made in most rendering systems that use the pin-hole camera model or projective transformations. Unfortunately, no physical system is capable of ideal image formation. Real lenses exhibit deviations from the ideal in the form of *aberrations* such as coma or pin-cushion distortion[15].

- *The radiometry of image formation must be correctly computed.*

  The correct exposure must be computed given the lighting in the scene. In most rendering systems this computation is arbitrary, with little attention paid to units and their physical magnitudes. In a real camera, the exposure is controlled by a variety of factors and these must be correctly simulated if a physically-based rendering system is to produce realistic output. Moreover, while ideal lenses focus light energy evenly at

[1] In our figures, we follow the convention of drawing object space to the left of the lens system, image space to the right, with coordinates along the axis increasing from left to right. Distances in the lens system are signed quantities, with a distance measured from left to right being positive, and right to left negative. Unprimed variables are in object space, primed are in image space.

all points on the image plane, real lenses suffer from an uneven exposure across the backplane. Accurate computation is therefore more than a matter of simply computing a correct overall scale factor.

Abstractly, the purpose of our camera module is to transform the scene radiances computed by the lighting and shading modules into the response at a pixel. This may be modeled by the *measurement equation*[9] (in computer graphics sometimes called the *pixel equation*)

$$R = \iiiint L(T(x', \omega', \lambda); \lambda) S(x', t) P(x', \lambda) dx' \cdot d\omega' \, dt \, d\lambda \tag{1}$$

In this equation, x' represents a position vector on the backplane, ω' is a direction vector towards the lens system, t is time and λ is wavelength. L is the scene spectral radiance defined in object space. The function T models the geometry of image formation, in effect transforming from image space to object space (for generality, we assume this is a function of wavelength). S models the behavior of the shutter and is a function of time (more generally, the response of real shutters may also depend on position). P describes the sensor response characteristics and is a function of position within a pixel and wavelength.

The measurement equation provides the basis for quantifying the effects of the lens and other camera components on image formation. The rest of the paper discusses how we model the lens and evaluate the measurement equation.

3 Lens Geometry and Image Formation

In this section, we discuss the geometrical properties of lens systems. We describe how to trace rays through a lens system, how to derive a projective transformation that approximates the action of the lens system, how to accurately model the geometry of focusing, and finally how to derive the effective size of the aperture. These techniques allow us to use actual lens descriptions in rendering systems that use ray tracing, as well as those that use linear viewing transformations. They also allow us to model the depth of field and exposure due to real lens systems.

3.1 Tracing Rays Through Lens Systems

One robust and accurate method to predict how a lens will form an image is to trace rays of light through the system. Lens and optical system designers have employed ray tracing techniques to evaluate designs for more than a century, and thus the process is now quite well-understood. Typically, a random set of rays are traced from object space to image space and their positions on the film plane are recorded to form a spot diagram. Various statistics are derived from these diagrams to evaluate the quality of the lens. Surprisingly, to our knowledge, ray tracing is not used by lens designers to create synthetic imagery because of the perceived high cost of doing these calculations.

```
R = Ray(point on film plane, point on rear-most element)
For each lens element E_i, from rear to front,
        p = intersection of R and E_i
        If p is outside clear aperture of E_i
                ray is blocked
        Else if medium on far side of E_i ≠ medium on near side
                compute new direction for R using Snell's law
```

Figure 2: Basic algorithm for tracing a ray through a lens system.

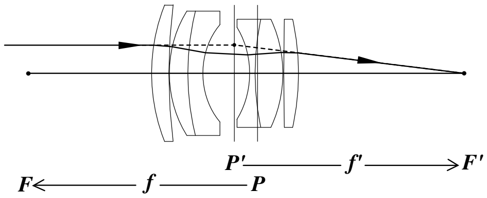


Figure 3: Finding a thick approximation to the lens in Figure 1. The actual path of an axis-parallel ray from object space is drawn as a solid line, and its idealized path is drawn as a dashed line.

The standard algorithm for tracing a ray through the lens is given in Figure 2. The propagation of a ray through a lens surface involves both finding the point of intersection between the ray and the surface and the refraction of the ray as it crosses the interface between the two media. The vast majority of lenses have spherical or planar surfaces, and therefore these computations are quite simple [16][17]. Although spherical surfaces are by far the most common, an object-oriented design of the lens software makes it possible to include elements of any shape for which intersection and normal-finding routines can be written.

Tracing rays through a lens system described in the tabular format is considerably faster than it would be if the lens were modeled as a collection of general objects for the ray tracer to render. This is because the exact visibility ordering of the surfaces is known *a priori*, and thus there is no search required to find the closest surface in a given direction. The main computational cost of tracing rays through spherical systems is two square roots per surface. This cost is fixed relative to scene complexity, and is usually small compared to the total cost of object intersection tests and other lighting calculations.

3.2 Thick Lens Approximation

In some situations the geometry of image formation may be approximated by treating the lens as an ideal *thick lens*. A thick lens forms perfect images; that is, each point in object space is imaged onto a single point in image space and all points in the plane of focus map onto the image plane with uniform magnification. We use thick lenses in our model to determine the exit pupil, as discussed in Section 3.4.

The behavior of a thick lens can be characterized by its focal points and principal planes, which are illustrated in Figure 3. Axis-parallel rays from a point at infinity in object space will enter the lens, be refracted through it, and emerge with a new direction and intersect the axis at the secondary focal point, F'. The point at which the incident ray and the emergent ray would intersect defines the secondary principal plane P'. P' is an imaginary surface normal to the axis at which we assume refraction to have occurred. Similarly, axis-parallel rays from image space intersect the axis at F, the primary focal point, and the intersection of the original and refracted rays define P, the primary principal plane. The signed distance from P' to F' is the effective focal length of the lens, f', and is equal to $-f$ when both object and image space are in the same medium.

The thick lens derives its name from the fact that, unlike the thin lens model usually used in computer graphics, the principal planes are not assumed to coincide. The distance from P to P' is the the lens' effective thickness, and may be negative, as for the lens in Figure 3. This additional parameter allows for a more general model of image formation. Although a thin lens approximation can be valid if the thickness is negligible, the thickness of photographic lenses is usually significant. The utility of both approximations is that their imaging properties can be modeled by a simple transformation.

To find a thick approximation to a given lens system, we apply the above definitions of focal points and principal planes directly. We trace rays through the lens system from each side and find the appropriate points of intersection to define P, F, P', and F'. An alternative way to find these values is by using the various thick lens formulas, which provide an analytical means for deriving a thick lens from a collection of simple lenses. The advantage of the first method is that it yields a more accurate approximation to the lens because typical lens systems are designed to exhibit ideal image formation even though the individual elements are less than ideal.

The geometry of image formation by a thick lens may be realized by a projective transformation defined by the focal points and principal planes[1]. Given a point in object space at a signed distance z along the axis from P, the conjugate equation holds that

$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f'} \tag{2}$$

where z' is the axial distance from P' to the point's image in image space. This equation and some simple geometry can be used to find the image of a point on either side of the lens. However, the resulting equations are inconvenient in that z and z' are measured from different origins. If the origin is assumed to be at P and both distances are measured from it, the same equations apply, except that z' must then be translated by $t = P' - P$, the thickness of the lens. The total transformation can be written as a 4x4 matrix:

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ W' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 + \frac{t}{f'} & t \\ 0 & 0 & \frac{1}{f'} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Thus the thick lens approximation may be used in conventional rendering systems that use 4x4 projective transformations to model the camera. Note that when t is zero, the above transformation is identical to the usual thin lens transformation used in computer graphics[11].

3.3 Focusing

In order to make the camera model easy to control, it should be possible to specify the distance from the film plane at which the camera is focused. Focusing physical systems involves moving one or more lens elements along the axis in order to change the distance at which points are focused. For simple lenses, the housing and all of the elements are moved together, while in more complicated internal focusing lenses, only a few elements move while the lens housing itself remains stationary.

Given a point located at an axial distance z from the film plane, we can use (2) to determine how far the lens must be moved in order to bring the point into focus. If the lens is focused at infinity, refocusing at z can be done by moving the lens a distance T away from the film plane, where T satisfies:

$$T^2 + T(2f' + t - z) + f'^2 = 0 \tag{3}$$

One solution to (3) corresponds to the lens being near the film and far from the object, the other to the lens being near the object and far from the film. In most situations, physical constraints on the distance the lens can move will make the latter solution unrealizable.

Moving the lens relative to the film plane has the additional effect of changing the field of view. As the distance at which the camera is focused is decreased, the distance of the lens from the film plane is increased and the field of view shrinks. This effect is not modeled in the standard camera model, which assumes that the film plane is always located at the focal point and that the lens can be focused at any arbitrary distance without any change of configuration.

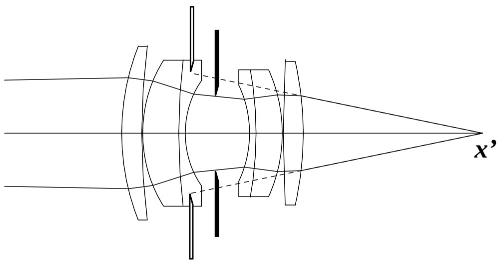


Figure 4: Illustration of the exit pupil for the double-Gauss lens of Figure 1. The diaphragm, drawn in solid black, acts as the aperture stop for the point x' on the axis at the film plane. The extent of the bundle of rays from x' that pass unobstructed through the lens is represented by the pair of solid lines on either side of the axis. The exit pupil, the image of the aperture stop through the rear-most two groups of elements, is drawn in outline. The exit pupil defines the cone of rays from x that pass unobstructed through the lens, as shown by the dashed lines.

3.4 The Exit Pupil

Recall that when looking through a lens system from a point on the backplane, there is a cone of rays within which the environment is visible, and that the aperture stop is the element limiting the extent of this cone. The *exit pupil* is defined to be the image of the aperture stop as viewed from image space (see Figure 4). Only rays directed from the film plane at the interior of the exit pupil will pass through the physical aperture stop, and so it is only these rays that we need consider when tracing rays through the system. Note the difference between this and directing rays at the aperture stop itself; this can produce incorrect results, because the image of the aperture may be larger than the aperture itself (as shown in Figure 4), and some rays that would pass through the system would not be generated. Note also the difference between this and firing rays at the lens element closest to the film plane. While this will produce correct results in the limit, it is wasteful because some of these rays may be blocked by the aperture stop. Using the correct exit pupil is critical if the depth of field and the exposure are to be computed consistently.

We find the exit pupil as follows: For each potential stop, we determine its apparent size and position from the axial point on the image plane. This is done by imaging the stop through those lens elements that fall between the stop and image space. We then determine which image disk subtends the smallest angle from the axial point on the image plane. This image is the exit pupil, and the stop corresponding to it is the aperture stop.

If we assume that each group of lens elements exhibits ideal image formation, the image of a given stop can be computed using a thick lens approximation to the appropriate subsystem of elements. In physical lenses, this is accurate only to the extent that the circular exit pupil is a reasonable approximation to the actual image of the aperture stop as viewed from off-axis points. In particular, some lenses distort the shape and position of the exit pupil when viewed from off-axis in order to increase or decrease exposure at points near the edge of the film[6]. We cannot validly use a thick approximation to find the exit pupil for such lenses in these cases. However, we can always ensure correct simulation by using the rear-most lens element as the exit pupil, at the cost of some loss of efficiency.

The exit pupil, rather than the aperture, should also be considered when using a thick lens in a ray tracer. Cook *et al.* described an algorithm for tracing rays through a thin lens by selecting a point on the aperture stop and tracing a ray from that point through the image of the current image plane point. As noted above, using the aperture stop rather than the exit pupil can lead to errors. The process of tracing a ray through a thick lens and exit pupil is shown in Figure 5.

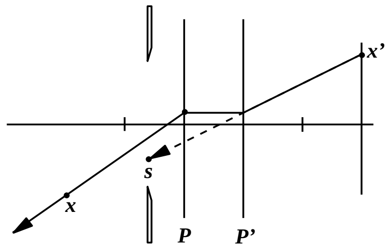


Figure 5: To trace a ray from x' through a thick lens, a point s on the exit pupil is chosen. The point of intersection of the ray from x' to s with P' is found, and is then translated parallel to the axis to P. The ray from this point through x, the image of x', is then used to sample the scene.

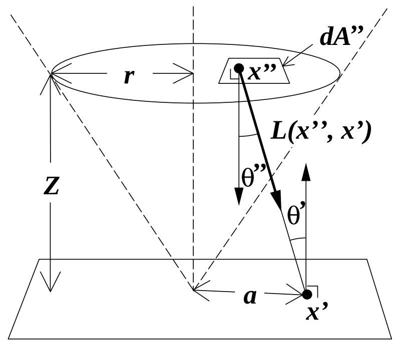


Figure 6: Geometry for computing the irradiance at a point on the film plane and the exact form factor.

4 Radiometry and Sampling

In this section we describe how we compute exposure on the film plane.

4.1 Exposure

Sensor response is a function of exposure, the integral of the irradiance at a point x' on the film plane over the time that the shutter is open. If we assume that irradiance is constant over the exposure period, and that exposure time is fixed,

$$H(x') = E(x')T \tag{4}$$

where $E(x')$ is the irradiance at x', T is the exposure duration, and $H(x')$ is the exposure at x'. This model is a simplification of the exposure process in physical systems, where the exposure at a point is dependent upon the shape and movement of the shutter.

In order to compute $E(x')$, we integrate the radiance at x' over the solid angle subtended by the exit pupil, which is represented as a disk, as shown in Figure 6.

$$E(x') = \int_{x'' \in D} L(x'', x') \frac{\cos\theta' \cos\theta''}{\|x'' - x'\|^2} dA'' \tag{5}$$

If the film plane is parallel to the disk, this can be rewritten as

$$E(x') = \frac{1}{Z^2} \int_{x'' \in D} L(x'', x') \cos^4\theta' dA'' \tag{6}$$

where Z is the axial distance from the film plane to the disk. This formula differs from that described by Cook *et al.*, which assumed each ray has the same weight. It is also important to perform the integral using a disc-shaped exit pupil, rather than a rectangular one. Using a rectangular pupil causes the depth of field to be computed incorrectly, since points not in focus will then have rectangular "circles" of confusion on the film plane.

The weighting in the irradiance integral leads to variation in irradiance across the film plane due to the lens system. There are two simple analytical ways to estimate this effect: the $\cos^4$ law and the differential form factor to a disk.

1. If the exit pupil subtends a small solid angle from x', θ' can be assumed to be constant and equal to the angle between x' and the center of the disk. This allows us to simplify (5) to:

$$E(x') = L\frac{A}{Z^2}\cos^4\theta' \tag{7}$$

where Z is the axial distance from the film plane to the disk, and A is the area of the disk. If Z is assumed to be the focal length, (7) can be written

$$E(x') = L\frac{\pi}{4}\frac{\cos^4\theta'}{n^2} \tag{8}$$

where n is the *f-number* of the lens. Equation (7) is the one most often found in optics texts, while (8) appears in many photographic texts. Note that both assume a small solid angle.

2. For larger solid angles, a more accurate way to estimate the variation in irradiance is to compute the differential form factor from a point on the film plane to a disk. This correctly accounts for the finite size of the disk, and the variation in angle as we integrate over the disk. This integral may be computed analytically[4] (an elegant derivation may be found in [8]).

$$F = \frac{1}{2}\left(1 - \frac{a^2 + Z^2 - r^2}{\sqrt{(a^2 + Z^2 + r^2)^2 - 4r^2a^2}}\right) \tag{9}$$

In real lens systems these analytical formulas overestimate the exposure. This is due to *vignetting*, the blocking of light by lens elements other than the aperture stop when a ray passes through the system at a large angle to the axis. Vignetting can be a significant effect in wide-angle lenses and when using a lens at full aperture. Fortunately, the ray tracing algorithm described in the last section accounts for this blockage, and hence computes the exposure correctly.

Figure 7 compares the irradiance computed by tracing rays through the lens system pointed at a uniform radiance field with

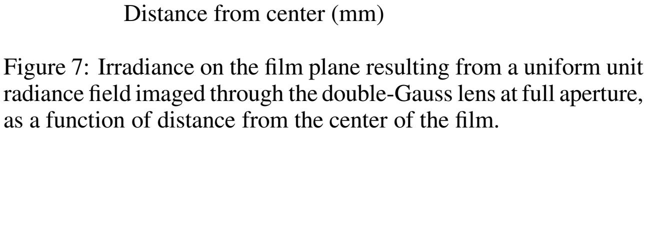


Figure 7: Irradiance on the film plane resulting from a uniform unit radiance field imaged through the double-Gauss lens at full aperture, as a function of distance from the center of the film.

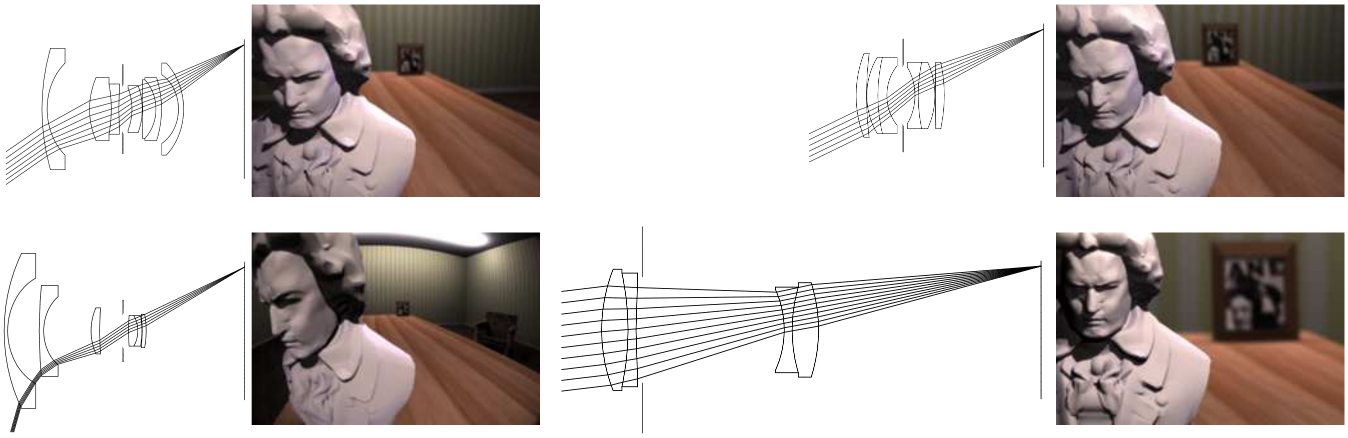


Figure 8: Four views of the same scene taken with a 16mm fisheye lens (bottom left), 35mm wide-angle lens (top left), 50mm double-Gauss lens (top right), and a 200mm telephoto lens (bottom right). A profile view of the lens system used to take each image is shown on the left. As with physical lenses, perspective is compressed with long focal lengths and expanded with short focal lengths. The fisheye image shows the lens' signature barrel distortion.

values computed using the usual computer graphics camera model (no weighting), the form factor, the $\cos^4$ approximation, and the full lens simulation. For this particular lens, the $\cos^4$ law and the form factor approximations do not differ significantly. However, vignetting reduces the true exposure near the edge of the film to nearly one third of its approximated value.

4.2 Sampling

In our model, a pixel's value is proportional to the radiant power falling on a hypothetical pixel-sized sensor in the image plane. The radiant power is given by the radiance integrated over the four-dimensional domain of pixel area and solid angle of incoming directions. This is estimated by sampling radiance over this domain (i.e., by casting rays from the pixel area toward the lens).

There are several ways to improve the efficiency of this calculation. First, we sample within the solid angle subtended by the exit pupil rather than sampling radiance over the entire hemisphere. Additional noise reduction might also be obtained by importance sampling, folding the factor of $\frac{\cos\theta' \cos\theta''}{\|x''-x'\|^2}$ into the distribution of rays over solid angle. Finally, efficiency can be improved by the use of good sampling patterns, which can reduce the amount of error in a pixel as well as affecting the overall distribution of noise in the final image. We have used stratified and quasirandom sampling patterns in our experiments.

Sampling without importance distribution is straightforward. Rays are cast from points in the pixel area toward points on the disk of the exit pupil, and the resulting values are weighted by $\frac{\cos\theta' \cos\theta''}{\|x''-x'\|^2}$. We can think of these pairs of points on the pixel area and exit pupil as being single points in the four-dimensional domain of integration.

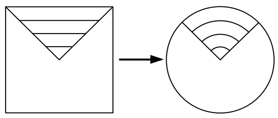
Rather than generating uniformly distributed points on the lens and weighting them, we can perform importance sampling by generating rays with a cosine-weighted distribution in solid angle and averaging the unweighted radiance values. We implemented an importance version of the third square-to-disk mapping described below. In the 35mm camera lenses that we tested, importance sampling reduced noise by only about one percent, because the $\cos^4\theta'$ weighting factor only varied by approximately twenty percent (see Figure 7). Since importance sampling adds a great deal of complexity and expense to the sampling operation, we believe it is not worth the effort in this particular application.

To generate these sample locations, it is usually necessary to start with some pattern of points defined in the hypercube $[0,1]^4$. Two of the dimensions are translated and scaled to the pixel area, and the other two dimensions are mapped to the disk of the exit pupil. The mapping from unit square to disk must be *measure preserving* (have a constant Jacobian) in order to avoid introducing a sampling bias. Thus, uniformly distributed points in the square map to uniformly distributed points on the disk. There are a number of such mappings. However, when mapping special sampling patterns such as stratified patterns it is good to choose a mapping that does not severely distort the shape of the strata. The obvious mapping,

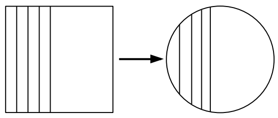
$$r = \sqrt{u},\ \theta = 2\pi v \tag{10}$$

is actually rather poor in this respect. A better mapping, used by Shirley[13], takes concentric squares to concentric circles. For example, in one wedge of the square, we have:

$$\begin{aligned} x' = 2x - 1,\ y' = 2y - 1 \\ r = y',\ \theta = \frac{x'}{y'} \end{aligned} \tag{11}$$



A third mapping we have implemented takes subrectangles $[0,x] \times [0,1]$ to a chord with area proportional to x, as illustrated below.



We have used two schemes to generate good sampling patterns in the hypercube. One is stratified sampling, dividing the dimensions of the hypercube into blocks and placing a sample randomly within each block. Given N samples, we could divide the hypercube into $N^{\frac{1}{4}}$ strata along each dimension. For typical values of N (it is unusual for a distributed ray tracer to cast more than a few hundred rays per pixel), this does not amount to many divisions of each dimension, and the benefits of stratification would be small. Instead, the pixel-area dimensions and the aperture-disk dimensions are strati-

Figure 9: Images synthesized with a 35mm wide-angle lens using, in order of decreasing accuracy, the full lens simulation (left), thick approximation (center), and the standard model (right). The top left arrow indicates the location of the scanline used in Figure 11.

fied separately as $\sqrt{N}$ by $\sqrt{N}$ grids on subsquares. To avoid systematic noise, the correlation of strata between pixel area and disk are randomly permuted.

We have found that the choice of square-to-disk mapping and stratification scheme strongly affect sampling efficiency and image quality. Using Shirley's mapping (11) yielded significantly lower RMS error (15 percent lower than (10) in typical experiments) as well as visibly improved image quality. Using the stratification method described above gives reasonable pixel antialiasing where edges are in sharp focus, and good distribution of noise in regions where depth of field causes blur.

5 Results

We have implemented our camera model as part of a ray tracer. The system supports rendering scenes using cameras constructed with different lenses and film formats. Figure 8 shows four images generated by the renderer and the lenses used in taking each of them. For each image, the camera was positioned in the scene so that the bust was imaged at approximately the same place and magnification on the film. As the focal length of the lens is increased, the relative size of the picture frame in the background grows, as expected. Darkening near the edge of the image due to vignetting is also apparent when using the fisheye and double-Gauss lens. These images typically required 90 minutes of CPU time to compute on a Silicon Graphics Indigo2 workstation at 16 rays per pixel. Approximately 10% of that time was spent tracing rays through the lens system, and thus the use of the full lens simulation is quite practical.

Figure 9 illustrates the differences in image geometry resulting from the use of different camera models. The standard camera model produces an image with too large a field of view, with both the bust and picture frame appearing smaller than in the full simulation image. The similarity of the full simulation and thick lens images illustrates the fact that using a thick lens can result in a good approximation if the actual lens forms nearly ideal images.

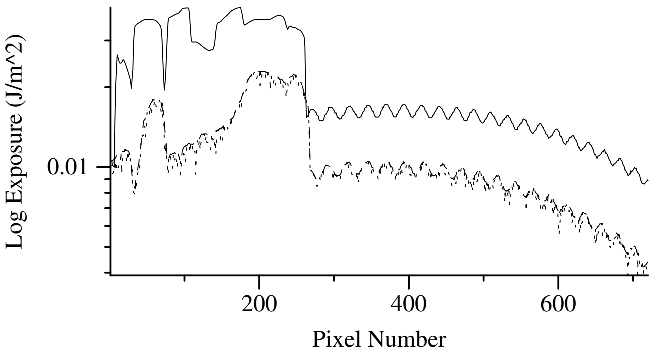


Figure 11: Comparison of exposure computed by the standard model (solid line), full simulation (dotted line), and thick approximation (dashed line), recorded at pixels along the indicated scanline of each image of Figure 9.

Figure 10 illustrates the change in field of view that occurs when the focus of a lens is changed. The figure shows the same scene as that in Figure 9, but with the lens focused on the picture frame in the background. Note that more of the front of the bust can be seen in the lower-left corner compared to the full simulation image in Figure 9. This increased field of view is caused by the movement of the lens towards the film plane when the lens is focused on a more distant object.

The new camera model also produces significant differences in exposure compared to the standard model. The exposure computed for a typical scene by the full simulation and the two approximations is shown in Figure 11. Exposure is generally overestimated in the standard model, and, as expected, the error tends to grow near the edge of the film.

An image taken with the fisheye lens is shown in Figure 12. Again, barrel distortion and darkening caused by vignetting are evident.

Figure 10: Camera focused on picture frame.

6 Summary and Discussion

The physically-based camera model that we have described draws upon techniques from both the lens design and computer graphics literature in order to simulate the geometry and radiometry of image formation. The lens system is described using standard lens construction information, and its behavior is characterized by tracing light rays through its various elements and weighting them properly. The primary added cost of using the model is finding the intersection with and refraction caused by each lens surface; for reasonably complex scenes the increase in rendering time is small, and the full

Figure 12:

simulation is very practical. Further, we show how the behavior of well-corrected lens systems can be approximated using a projective transformation derived from a thick lens approximation.

The new model is an improvement over standard models in a number of ways:

- The geometric relationships between the lens, object, and film plane are modeled properly by precise placement and movement of lens elements. This is necessary for accurate field of view and depth of field calculations.
- Image geometry is computed correctly by tracing the path of light through the system. The model is capable of simulating non-linear geometric transformations such as those produced by fisheye and anamorphic lenses, while simultaneously computing the correct exposure and depth of field.
- The image irradiance, or exposure, is computed properly because the model applies the correct weighting to rays traced through the lens system, and derives the correct exit pupil to control the limits of the integration. The model also correctly accounts for vignetting due to the blockage of rays.

Although our model is more accurate than previous camera models, there are many aspects of cameras and lens systems that we have not simulated. For example, our model assumes the shutter opens and closes instantaneously, which is not true. Our model also assumes that the lens transmittance is perfect, and that the properties of lens surfaces do not vary with position or time. We have also ignored many wavelength-dependent effects, in particular sensor sensitivity and response, and chromatic aberration due to the variation of index of refraction with wavelength. In the future we intend to experimentally verify our model by simulating particular lens systems and comparing the results with captured images. The goal of these experiments will be to find what level of detail of the camera must be simulated to match computer-generated and photographic images.

7 Acknowledgements

Thanks to Matt Pharr and Reid Gershbein for help with the figures and the rendering system, Viewpoint DataLabs for the Beethoven and Pentax models, and to the anonymous reviewers for their suggestions. This research was supported by equipment grants from Apple and Silicon Graphics Computer Systems, and a research grant from the National Science Foundation (CCR-9207966).

References

[1] Max Born and Emil Wolf. *Principles of Optics*. MacMillan, New York, second edition, 1964.

[2] Robert L. Cook, Thomas Porter, and Loren Carpenter. Distributed ray tracing. In *Computer Graphics (SIGGRAPH '84 Proceedings)*, volume 18, pages 137–145, July 1984.

[3] Status report of the graphics standards planning committee of the ACM/SIGGRAPH. *Computer Graphics*, 11:19+117, Fall 1977.

[4] P. Foote. Scientific paper 263. *Bulletin of the Bureau of Standards*, 12, 1915.

[5] Ned Greene and Paul S. Heckbert. Creating raster Omnimax images from multiple perspective views using the elliptical weighted average filter. *IEEE Computer Graphics and Applications*, 6(6):21–27, June 1986.

[6] Rudolph Kingslake. *Optics in Photography*. SPIE Optical Engineering Press, Bellingham, Washington, 1992.

[7] Nelson L. Max. Computer graphics distortion for IMAX and OMNIMAX projection. In *Nicograph '83 Proceedings*, pages 137–159, December 1983.

[8] Parry Moon and Domina Eberle Spencer. *The Photic Field*. The MIT Press, Cambridge, Massachusetts, 1981.

[9] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. Geometric considerations and nomenclature for reflectance. Monograph 161, National Bureau of Standards (US), October 1977.

[10] M. Potmesil and I. Chakravarty. A lens and aperture camera model for synthetic image generation. *Computer Graphics (SIGGRAPH '81 Proceedings)*, 15(3):297–305, August 1981.

[11] David F. Rogers and J. Alan Adams. *Mathematical Elements for Computer Graphics*. McGraw Hill, New York, second edition, 1990.

[12] Mikio Shinya. Post-filter for depth of field simulation with ray distribution buffer. In *Proceedings of Graphics Interface '94*, pages 59–66. Candian Human-Computer Communications Society, 1994.

[13] Peter Shirley. Discrepancy as a quality measure for sample distributions. *Eurographics '91 Proceedings*, pages 183–193, June 1991.

[14] Warren J. Smith. *Modern Lens Design*. McGraw Hill, New York, 1992.

[15] W. T. Welford. *Aberrations of the Symmetrical Optical System*. Academic Press, London, 1974.

[16] Turner Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 23(6):343–349, June 1980.

[17] Charles S. Williams and Orville A. Becklund. *Optics: A Short Course for Engineers and Scientists*. Wiley-Interscience, New York, 1972.