



## The Raxel Imaging Model and Ray-Based Calibration

MICHAEL D. GROSSBERG AND SHREE K. NAYAR

*Department of Computer Science, Columbia University, New York, NY 10027, USA*

mdog@cs.columbia.edu

nayar@cs.columbia.edu

*Received February 25, 2002; Revised March 25, 2004; Accepted March 25, 2004*

*First online version published in October, 2004*

**Abstract.** An imaging model provides a mathematical description of correspondence between points in a scene and in an image. The dominant imaging model, perspective projection, has long been used to describe traditional cameras as well as the human eye. We propose an imaging model which is flexible enough to represent an arbitrary imaging system. For example using this model we can describe systems using fisheye lenses or compound insect eyes, which violate the assumptions of perspective projection. By relaxing the requirements of perspective projection, we give imaging system designers greater freedom to explore systems which meet other requirements such as compact size and wide field of view. We formulate our model by noting that all imaging systems perform a mapping from incoming scene rays to photosensitive elements on the image detector. This mapping can be conveniently described using a set of virtual sensing elements called *raxels*. Raxels include geometric, radiometric and optical properties. We present a novel ray based calibration method that uses structured light patterns to extract the raxel parameters of an arbitrary imaging system. Experimental results for perspective as well as non-perspective imaging systems are included.

**Keywords:** calibration, caustic, catacaustic, catadioptric, diacaustic, imaging model, light-field, non-single viewpoint, non-perspective, perspective, plenoptic, raxel, ray surface, ray-based, viewpoint locus

### 1. Introduction

Central to computer vision is the interpretation of a scene from one or more images. To accomplish this, we must have a model of the imaging process which relates the location and intensities of intensity values in the image to radiance in the scene. Through the widespread use of cameras whose optics were designed for conventional photography it has been possible to use the simple and well understood perspective projection model. This model is based on three essential assumptions:

1. Each point in the image corresponds to a ray entering the imaging system.
2. All rays entering the imaging systems intersect at

a single point called the center of projection (COP) or viewpoint.

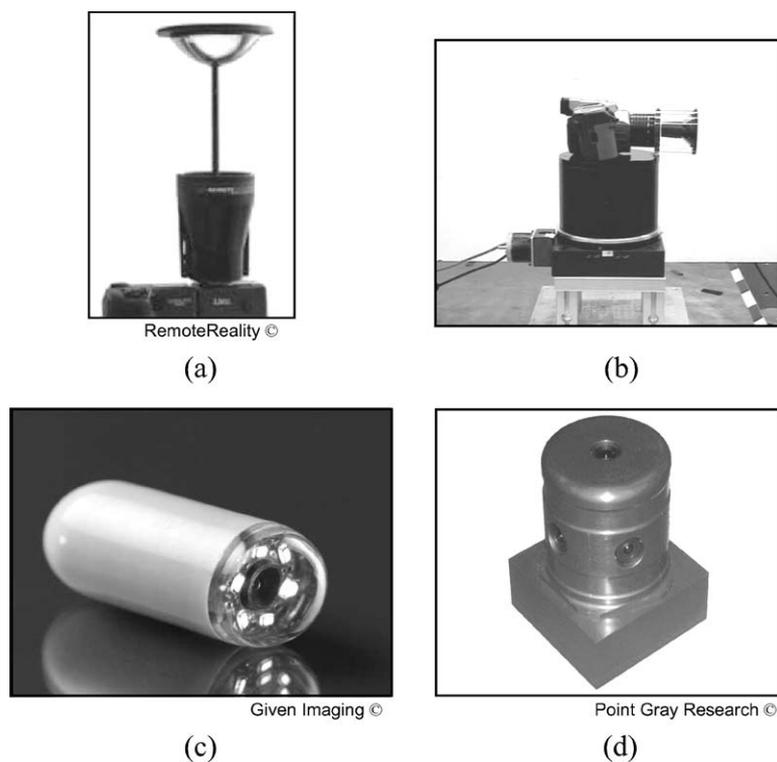
3. The correspondence between rays passing through the viewpoint and image is given by intersection of the rays with a plane.

Perspective projection closely corresponds to the original device engineered to create an image, the camera obscura. The camera obscura has been known since the renaissance if not much earlier (Hammond, 1981). In a camera obscura rays of light pass through a hole in one wall of a room and fall on the opposite wall forming an image. Originally an artist was required to record the image. The artist has been replaced first by the advent of photographic film, and more recently by solid state devices which can record the image

virtually instantaneously. More importantly, the hole has been replaced by complex optics. These optics permit the aperture of imaging device to be large, creating a bright image, while essentially still following the perspective imaging model. Geometers have studied the properties of this model and derived a large suite of projective invariants that provide insights into the relationship between configurations of points in a scene with their projection in the image (Faugeras, 1993).

Despite its great relevance, the perspective model no longer suffices to describe all imaging systems used for computer vision. In recent years, the notion of a “vision sensor” has taken on a much broader meaning. Researchers developed a variety of devices that sample the light field (Gershun, 1939) or the plenoptic function (Adelson and Bergen, 1991) associated with a scene in interesting and useful non-perspective ways. Figure 1 shows some examples currently used. Figure 1(a) shows a commercially available system consisting of

a parabolic mirror and a tele-centric lens. Systems combining lenses and mirrors are called catadioptric. In this case the system has a single viewpoint. The image formed cannot, however, be described by the intersection of rays from that viewpoint with a plane, thus violating the assumptions of perspective projection (Baker and Nayar, 1999; Nayar and Baker, 1997; Nayar and Peri, 1999; Yagi and Kawato, 1990). The properties of systems consisting of perspective cameras with quadratic surface mirrors have been explored in Swaminathan et al. (2001). For many of these systems there is no single viewpoint through which the incoming rays pass. It is interesting to note that certain applications, such as panoramic stereo, require an imaging system to have a locus of viewpoints (Peleg et al., 2000). Figure 1(b) shows a conical mirror attached to a perspective camera. This rotating system, designed for capturing data to create stereo panoramas, does not have a single viewpoint (Nayar and Karmarkar, 2000; Pajdla, 2002).



*Figure 1.* Examples of non-perspective imaging systems: (a) A catadioptric system consisting of a camera with a tele-centric lens and parabolic mirror (RemoteReality). This system has a single viewpoint but is not perspective. (b) A system consisting of a perspective lens and a conical mirror, which is used to collect data for a stereo panorama. The design requirements for this application dictated the use of a non-single viewpoint device. (d) The M2A camera pill from Given Imaging. Once swallowed the system provides imaging for medical diagnosis. The imaging system’s fisheye lens does not have a single viewpoint. (e) Ladybug imaging system from Point Gray Research for producing spherical panoramas. The system is a camera cluster having multiple viewpoints. None of these systems can be described by the perspective projection model.

Similarly, commercially available wide-angle lens systems (Miyamoto, 1994) create severe distortions and do not have a single viewpoint. For example the commercial system shown in Fig. 1(c) incorporates a camera with a  $140^\circ$  field-of-view wide-angle lens into a pill. The distorted images this camera produces when swallowed are used by medical specialists to diagnose a range of gastrointestinal illnesses.

Clusters of cameras like the commercial system shown in Fig. 1(d) have become popular for creating high resolution mosaics (McCutchen, 1991; Swaminathan and Nayar, 2000). Camera clusters have been used to recover ego-motion (Baker et al., 2001; Neumann et al., 2003). Planar mirrors create virtual camera clusters, which have been applied to stereo recovery (Gluckman and Nayar, 1999). Such systems include multiple viewpoints, each one associated with one of the cameras in the cluster. In the case of insects, nature has evolved eyes that have compound lenses (Dawkins, 1996; Gaten, 1994). These eyes are composed of thousands of “ommatidia”, each ommatidium including a receptor and lens. It is only a matter of time before we see solid-state cameras with flexible imaging surfaces that include a large number of such ommatidia. More generally, flexible imaging surfaces that can change over time have already been used in adaptive optics (Tyson, 1998). We may accommodate some imaging systems with extensions to perspective imaging models such as by considering radial and tangential distortions (Brown, 1966; Conrady, 1919; Goshtasby, 1989; Tsai, 1987; Weng et al., 1992). However, to create a unified model covering all above examples must entirely rethink our imaging model.

For many vision applications, such as structure from motion, or shape from shading the exact nature of the imaging process may make little difference to the application. If the details of the imaging process do not matter we should be able to treat the imaging process as a black box which takes a scene as input and outputs an image. A general imaging model should capture the relevant information for the black box by relating radiance at scene points and their location in space to points in the acquired images. In this paper we address two questions that we believe are fundamental to imaging:

- Is there an imaging model that is general enough to represent an arbitrary imaging system?
- Given an unknown imaging system (a black box), is there a simple calibration method that can estimate the parameters of the imaging model?

Our approach to formulating a general imaging model to exploit the fact that all imaging systems perform a mapping from incoming scene rays to a set of photosensitive elements. The most general imaging model we consider is simply a discrete representation of this map. The smallest element of our imaging model is a virtual photosensitive element that measures light in essentially a single direction. We refer to these virtual elements as ray pixels, or *raxels*. For the class of imaging systems where we can assume a 2-dimensional image plane and a continuous mapping, the model can be conveniently represented by a *ray surface* which is a surface in 3-dimensional space from which the rays are measured in various directions.

While there are many possible choices for a ray surface we show that the *caustic* of the imaging system provides an essentially unique and natural choice. Thus the caustic gives a natural geometric parametrization of the imaging system raxels. In addition to its geometric parameters, each raxel may have its own radiometric response function and local point spread function. To accommodate adaptive optics, for example, the parameters of our model may also be made time dependent. We note that since we first proposed a model of this type in Grossberg and Nayar (2001), some of the machinery from the perspective model has been generalized to apply to the general imaging model. These include an analysis of epi-polar constraints, ego-motion and structure from motion (Pless, 2002, 2003; Neumann et al., 2003).

After describing the general imaging model and its properties, we present a simple method for finding the parameters of the model for any arbitrary imaging system. We will describe our calibration method for a still camera with fixed components. Were the components to change, as would be the case for adaptive optics, the calibration could be repeated very rapidly, in principle. It is also important to note that, given the non-perspective nature of a general device, conventional calibration methods (Tsai, 1987; Faugeras, 1992; Hartley, 1993), cannot be directly applied. Since we are interested recovering the mapping from rays to image points, we need a dense ray-based calibration method. We describe a simple and effective ray-based approach that uses binary-coded light patterns like those described in Sato and Inokuchi (1985). This method allows a user to obtain the geometric, radiometric, and optical parameters of an arbitrarily complex imaging system in a matter of minutes.

## 2. General Imaging Model: Geometry

As shown in Fig. 2, an arbitrary imaging system has two essential components: imaging optics, which direct and focus light; and photosensitive elements, which respond to the intensity of that light. These photosensitive elements, corresponding to pixels, could be elements of an electronic chip (CMOS or CCD), film, or any other light sensitive device. They may be arranged on a plane, or a curved surface. The elements may be uniformly spaced, log-polar spaced to create a foveal region, or randomly arranged according to some distribution, as in the human retina (Moini, 2000).

The imaging optics directing the light to the photosensitive elements typically include several parts. Even a relatively simple optical component has about five individual lenses within it. In our arbitrary system, there may be additional optical elements such as mirrors, prisms, or beam-splitters. In fact, the system could be comprised of multiple individual imaging systems, each with its own imaging optics and image detector.

Irrespective of its specific design, the purpose of an imaging system is to map incoming rays of light from the scene onto pixels of the detector. Each pixel collects light energy from a bundle of closely packed rays in any system that has a non-zero aperture size. We can, however, represent this bundle by a single chief (or principle) ray when studying the geometric properties of the imaging system. As shown in Fig. 2, the system maps the ray  $\mathbf{P}_i$  to the pixel  $i$ . Even in a perspective system

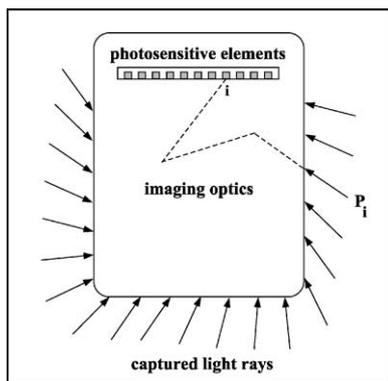


Figure 2. An arbitrary imaging system consisting of imaging optics, which direct incoming light rays, and photosensitive elements (pixels). Each pixel collects light from a bundle of rays that pass through the finite aperture of the imaging system. Nevertheless, we can assume a correspondence between each individual detector element  $i$  and a specific ray  $\mathbf{P}_i$ , representing a narrow bundle of light rays entering the system.

the actual path that a single incoming ray traverses to the pixel can be arbitrarily complex.

If the imaging system is perspective, all the incoming light rays converge through a single viewpoint, namely, the center of projection of the perspective system. This is not true in an arbitrary imaging system. For instance, in Fig. 2, if we extend the captured rays, they do not meet at a single viewpoint. The goal of this section is to present a geometrical model that can represent arbitrary imaging systems.

### 2.1. Raxels

It is convenient to represent the mapping from scene rays to pixels in a form that easily lends itself to manipulation and analysis. Once we treat our imaging system as a black box, the physical location of our photosensitive elements *do not matter*. If pixel  $i$  responds to the intensity of light associated with the chief ray  $\mathbf{P}_i$  as in Fig. 2, we may as well assume that the photosensitive element  $i$  is actually located along the ray  $\mathbf{P}_i$ . Thus, we replace our physical pixels with an abstract mathematical equivalent we refer to as a ray pixel or *raxel*. A raxel is a virtual photosensitive element that measures the light energy of a compact bundle of rays represented by a single principle incoming ray.<sup>1</sup>

The abstract optical model of our virtual raxel is shown in Fig. 3. Each raxel includes a pixel that measures light energy and imaging optics (a lens) that collects the bundle of rays around an incoming principal ray. In this section, we will focus on the geometric properties (locations and orientations) of raxels. Each raxel, however, can possess a radiometric (brightness

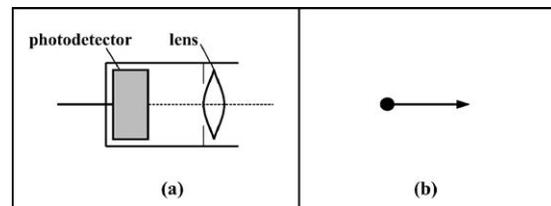


Figure 3. (a) A raxel, the virtual replacement for a real photosensitive element combining a ray and a pixel. The physical location of the photosensitive element may not lie along the principle ray of light associated to it. Nevertheless, we are treating our imaging system as a black box so in our model we *may as well assume* that the photosensitive element lies along the principle ray it detects. Thus a raxel has both location and orientation. In addition, a raxel may have radiometric and optical parameters. (b) The notation for a raxel used in this paper.

and wavelength) response as well as optical (point spread) properties.

## 2.2. The Plenoptic Function and the Discrete Raxel Model

What does an imaging system see? The input to an imaging system is the intensity of light along rays in space. This information is called the *light field* and is encoded in the plenoptic function (Gershun, 1939; Adelson and Bergen, 1991). The plenoptic function  $\Phi(\mathbf{p}, \mathbf{q}, t, \lambda)$  gives the intensity of light at each point  $\mathbf{p}$  in space, from direction  $\mathbf{q}$ , at an instant of time  $t$  and wavelength  $\lambda$ . We can specify position by  $(p_X, p_Y, p_Z)$  and direction by the vector  $(q_X, q_Y, q_Z)$ .<sup>2</sup> Still images represent an integration of light energy over a short time period, given by the effective shutter speed. Further, each photosensitive element will average the plenoptic function across a range of wavelengths. Although adding time and wavelength would not introduce any new theoretical complications to our model, it would require our notation to become more complex. Thus to keep the exposition simpler we set aside time and wavelength by considering monochromatic still imaging. In this way we consider the plenoptic function as a function on  $\mathbf{R}^6$ , depending on only position and direction:  $\Phi(\mathbf{p}, \mathbf{q})$ .

If the plenoptic function is the right mathematical abstraction to describe the input to the system, then we may ask, what abstraction describes a raxel? Assume, for the moment, that our photosensitive element responds linearly to intensity. Since a raxel samples  $\Phi$  at  $(\mathbf{p}_0, \mathbf{q}_0)$ , a raxel is just a constant times the delta function  $\delta_{\mathbf{p}_0, \mathbf{q}_0}$  in  $\mathbf{R}^6$ .<sup>3</sup> The parameters for the raxel are thus just position  $\mathbf{p}_0$  and direction  $\mathbf{q}_0$  as well as the multiplicative constant of the delta function. Hence, we can define a geometric model of the camera as follows:

*Geometric Model 1 (Discrete Raxel Model).* The geometry of an imaging system is specified by a sequence of positions and directions indexed by the  $N$  image pixels:

$$\{(p_1, q_1), \dots, (p_N, q_N)\} \quad (1)$$

The advantages of this representation is the generality it affords. Since no assumptions are made about the source of the image pixels this covers camera clusters or even a sequence of images in a video. One potential problem is that our assumption that a raxel is a delta

function and hence responds linearly to input. Photosensitive elements of film, for example, do not usually respond linearly. In that case, the raxel is a delta function only after the response is linearized. We easily address this by attaching other parameters to a raxel such as a response function.

One price of generality is that we lose the simple geometry afforded in the perspective model. A second cost is that the representation is not unique to a given imaging system. We have the freedom to change the position of  $p_i$  along the ray associated to the raxel.<sup>4</sup> We can address these issues but only by making further assumptions on the imaging system.

## 2.3. Pencils of Rays and Ray Surfaces

The raxel model we described in Section 2.2 can be interpreted as a discrete mapping from the set of pixels to a selection of rays in space. For many imaging systems this mapping is really a discrete sampling of a continuous mapping. To see this, consider the case of the perspective model. The viewpoint and image plane together determine a continuous mapping from points in the image plane to the pencil of rays passing through the single viewpoint. The discrete raxel model selects one ray for each pixel in the image plane, discretely sampling this mapping.

For many non-perspective imaging systems we can make two assumptions. Firstly we assume that the image plane is 2-dimensional. Secondly we assume that the mapping from the image plane to rays in space is piecewise continuous. These two assumptions cover all the examples shown in Fig. 1. For camera clusters or image sequences we may arrange all the images in a single plane. Since we only assume the mapping to rays to be piecewise continuous it can be described by a single map.<sup>5</sup>

Generalizing from continuous to piecewise continuous mappings amounts to treating each piece independently. Thus for clarity of our presentation we will assume the mapping of the image plane to rays is actually continuous. The continuous mapping selects a pencil of rays which represent locations and directions in space that the imaging system sees. For the perspective imaging system, this pencil of rays is compactly characterized by the single viewpoint. How do we generalize this?

While we may not be able to organize our raxels to reside at a single viewpoint, we can organize the raxels on a surface we call a *ray surface* (Grossberg and Nayar,

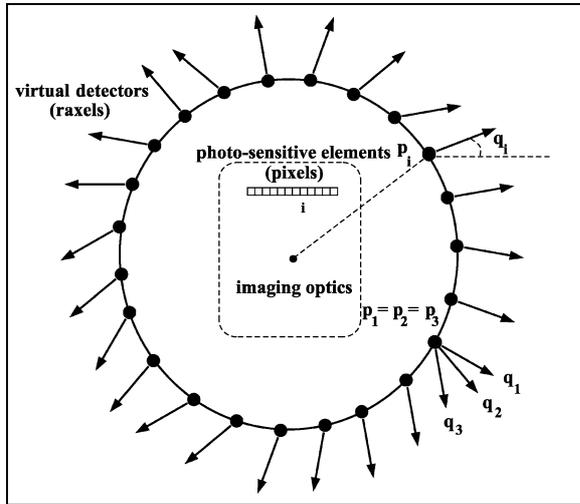


Figure 4. An imaging system may be modeled as a set of raxels on a sphere surrounding the imaging system. Each raxel  $i$  has a position  $\mathbf{p}_i$  on the sphere, and an orientation  $\mathbf{q}_i$  aligned with an incoming ray. Multiple raxels may be located at the same point ( $\mathbf{p}_1 = \mathbf{p}_2 = \mathbf{p}_3$ ), but have different directions.

2001). For example, consider a sphere enclosing our imaging system, as shown in Fig. 4. The ray associated with a photosensitive element  $i$  must pass through the sphere at some point  $\mathbf{p}_i$  and in some direction  $\mathbf{q}_i$ . The sphere determines the placement of a raxel for each ray. Classifying the positions and directions determined by a light source, a related concept, was explored in Langer and Zucker (1997). We also note that our concept of a ray surface is similar to the concept of a “generator” introduced in Seitz and Kim (2001) to classify the space of stereo images.

A choice of ray surface fixes the location of a raxel for each ray, but the surface need not parameterize the rays. There may be several rays that enter the sphere at the same point but with different directions (see  $\mathbf{q}_1$ ,  $\mathbf{q}_2$  and  $\mathbf{q}_3$  in Fig. 4). For example, if the imaging device is perspective and the single viewpoint lies on the sphere, all the rays pass through the same point. Thus the direction  $\mathbf{q}$  is not, in general, a function of  $\mathbf{p}$ , even for a sphere.

The sphere is convenient because it works even when an imaging device has a wide field of view. The choice of intersecting the incoming rays with a sphere is arbitrary. In Gortler et al. (1996) and Levoy and Hanrahan (1996), it was suggested that the plenoptic function can be conveniently restricted to a plane. The important thing is to choose some ray surface so that an incoming ray intersects this surface at a unique point. This

leads to a model for imaging systems with continuous raxels.

*Geometric Model 2 (Continuous Raxel Model).* The geometry of an imaging system is specified by a mapping from the image plane to a set of points, called the ray surface,<sup>6</sup> and a set of direction vectors in space from that surface.

Note that the ray surface allows us to arrange the geometric parameters of the raxels. As delta functions the raxels continue to be defined in  $\mathbf{R}^6$ . Setting our notation, if the incoming rays are parameterized by image coordinates  $(x, y)$ , each ray will intersect a ray surface at one point  $\mathbf{p}(x, y)$ . We can write the ray surface as a function of  $(x, y)$  as:

$$\mathbf{s}(x, y) = (\mathbf{p}(x, y), \mathbf{q}(x, y)). \quad (2)$$

We can express the position of any point along the ray as  $\mathbf{p}(x, y, r) = \mathbf{p}(x, y) + r\mathbf{q}(x, y)$ . This allows us to express the domain of the plenoptic function which the imaging system can see as the range of

$$\mathcal{L}(x, y, r) = (\mathbf{p} + r\mathbf{q}, \mathbf{q}). \quad (3)$$

In the case of a known imaging system we are able to compute  $\mathbf{s}(x, y)$  a priori. We trace the rays through each optical component to establish a ray surface for the incoming rays captured. Using Eq. (3) we may express one ray surface in terms of another ray surface such as a sphere or a plane. In the case of an unknown imaging system we must measure  $\mathbf{s}(x, y)$  along some ray surface via calibration, as we will describe. In any case a ray surface provides a compact and convenient representation of the geometry of the raxels.

### 3. Caustics

The primary weakness of the continuous raxel model described above is that the representation is not unique. For comparison, consider the perspective model. In that model the single viewpoint is determined by the geometry of rays and is unique. This makes it very easy to compare imaging systems. Suppose two imaging systems have the same single viewpoint. Suppose also that mapping from the image plane to rays passing through the single viewpoint is the same corresponding to identical internal camera parameters. Even though the two

imaging systems may have radically different optics, we can consider these systems equivalent.

It is difficult to determine when two general imaging systems are equivalent if we choose arbitrary ray surfaces to characterize them. We would like to find a natural ray surface which is uniquely determined from the geometry of the rays. We motivate our choice of such a surface by exploring in greater detail how the single viewpoint is picked out by the geometry of rays in a perspective system.

To understand this geometry, we consider all rays in 2-dimensions where position is given by  $(X, Y)$  and direction of a ray is given by  $\theta$ . Figure 5(a) shows points on the rays of a perspective system in 2-dimensions on the bottom. At the top of the figure, the third dimension is used to indicate the directions of the rays. In this space of positions *and* directions, the rays do not intersect. In algebraic geometry this is called a “blow up” (Griffiths and Harris, 1978). When projected back to the plane the lines intersect. The mapping from positions and directions to only positions is singular at the intersection as the mapping becomes many to one. We may also understand the singularity by seeing how the mapping effects area. If we mark a rectangular area in the space of positions and directions as in Fig. 5(a), away from where positions converge, we see it projects in the plane to a similar area. Near where the points converge the rectangular area will collapse, that is, the area after

projection will be dramatically reduced indicating the singularity.

There can be a singularity even if the rays do not converge to a point. In Fig. 5(b), we see rays resulting from reflection of parallel rays off a circular reflector. In this example the rays do not converge to a single point in the space of positions and directions, or in the plane after projection. Nevertheless, from looking at small rectangles of area under projection, we see that there is also a singularity in the rays. In this case the singularity occurs along the envelope of the rays in the plane.

### 3.1. Caustic Raxel Model

Singularities in the set of rays of an imaging system are called *caustics*. In Section 2.3, we described the set of points and directions that can be detected by the imaging system. We parameterized this set using Eq. (3) as the range of the function  $\mathcal{L}(x, y, r)$ . The position component functions for  $\mathcal{L}$  are  $X = p_X(x, y, r)$ ,  $Y = p_Y(x, y, r)$ , and  $Z = p_Z(x, y, r)$ . The mapping from  $(x, y, r)$  to  $(X, Y, Z)$  can be viewed as a change of coordinates.

*Definition.* The *caustic* is defined as the locus of points in  $X, Y, Z$  space where this change of coordinates is singular.<sup>7</sup>

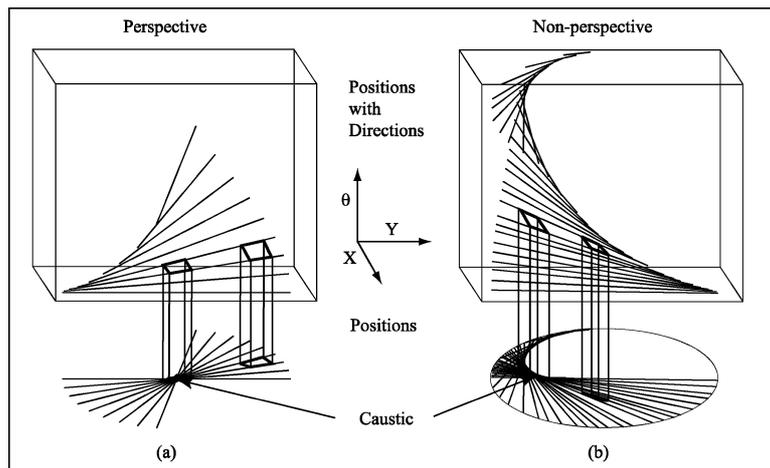


Figure 5. Rays in 2-dimensions  $(X, Y)$  shown with their directions  $\theta$  plotted in the third dimension. Below, points on rays are projected back to the plane of positions. (a) The perspective case. Rays pass through a single point in 2-dimensions. The singularity of the map, the caustic, can be seen where the projection is many-to-one. It can also be located by noting where the area of an infinitesimal rectangle collapses under projection. (b) A non-perspective case. Shown are rays resulting from parallel rays reflecting off a circular reflector. The caustic can be found where area collapses under projection. The caustic is the envelope of the rays in the plane.

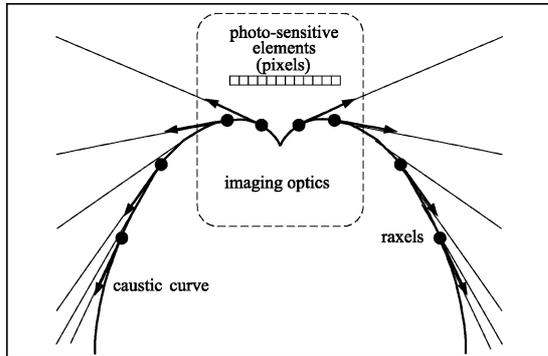


Figure 6. The caustic is a good candidate for the ray surface of an imaging system as it is closely related to the geometry of the incoming rays; the incoming ray directions are tangent to the caustic.

Caustics are important in understanding the properties of optical components and have been important for analyzing telescopes as well (Born and Wolf, 1965; Burkhard and Shealy, 1973). For optical systems the caustics characterize the geometry of the rays. Thus, we use the caustic to define a geometric model of our imaging system (see Fig. 6).

*Geometric Model 3 (Caustic Raxel Model).* The geometry of an imaging system is specified by a piecewise differentiable mapping from the image plane to the caustic of the imaging system's incoming rays.

We must assume the mapping from the image plane to the caustic is piecewise differentiable to compute the caustic. In principal, our assumption that the mapping is piecewise differentiable, and that the caustic characterizes the geometry of rays further restricts our model. In practice the all imaging systems in Fig. 1 satisfy these assumptions.

We also note the term caustic usually refers to the loci of positions of the singularity. Each ray passing through the caustic determines both a position *and* a direction on the caustic. Thus the caustic raxel model is a special case of the continuous raxel model where take the caustic to represent both the positions and directions of the raxels associated to a points in the image plane.

Consider rays in 2-dimensions. When the singularity is a curve rather than a point, the unit tangents along the curve give the directions of the ray surface as in Fig. 6. For rays in 3-dimensions, caustics can be 0-, 1-, or 2-dimensional. When the caustic is 2-dimensional the directions of the rays will be tangent to the surface. Each point on a surface has a plane of tangent vectors.

For the sphere ray surface, in Section 2.3, we needed two parameters to specify direction. In the case of a caustic surface we only need to specify one parameter: a single direction for each tangent plane. On caustic surfaces, the direction  $\mathbf{q}$  is really a function of position  $\mathbf{p}$ , and the incoming rays are tangent to the surface.

Emphasizing direction  $\mathbf{q}$  as a function of position  $\mathbf{p}$  is misleading. An imaging model must not only describe the rays in space but their relationship to points in the image. Thus  $\mathbf{p}$  and  $\mathbf{q}$  should be functions of  $(x, y)$  in the image plane. Some simple examples illustrate the point. In Fig. 7(a), the caustic of the perspective system is a single point, and the relationship to the image is given by the intersection with the image plane. In Fig. 7(b) a camera has a tele-centric lens. All the incoming rays to the camera are parallel. The parallel rays are allowed to reflect off a parabolic mirror whose axis of symmetry is aligned with the direction of the rays. Tracing the incoming rays, we find that they all meet at the focus of the parabola (Cornbleet, 1994). Thus this system also has a single point caustic. Nevertheless, the system is not perspective. The mapping of rays to image points is not given by intersection with a plane but is given by a more complex mapping. Finally, in Fig. 7(c), when the rays entering a perspective camera are first allowed to reflect off a parabolic mirror, they form a 2-dimensional caustic.

The singular nature of caustics is best illustrated by considering rays of illumination. A caustic formed by rays of illumination generates a very bright region when it intersects a surface (Mitchell and Hanrahan, 1992). For example, when light refracts through shallow water of a pool, bright curves can be seen where the caustics intersect the bottom (Watt, 1990). In the context of illumination we also note that since projectors can be thought of a dual to cameras, our raxel model can be applied to projection systems as well.

### 3.2. Computing Caustics

Determining the geometric parameters of the caustic raxel model for a given imaging system amounts to computing the caustics of the incoming rays. When the optical components and parameters of a system are known a-priori we can derive the mapping from the image plane to the incoming rays analytically or compute it numerically. For a black box imaging system we can measure the mapping of pixels to incoming rays via a calibration procedure. Whatever the case, our goal is to compute the caustic surface from the given mapping.

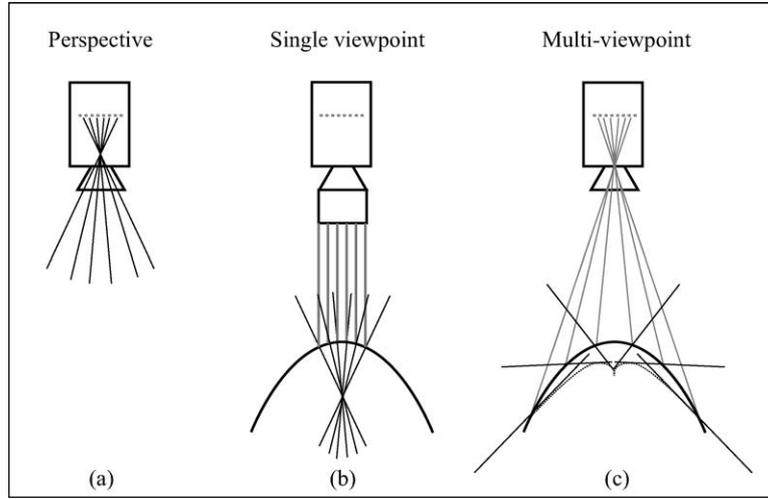


Figure 7. (a) The perspective system, with caustic given by a point. The correspondence between image points and rays is given simply by intersection with a plane. (b) A catadioptric system consisting of a camera with tele-centric lens, and a parabolic mirror. Incoming rays intersect in a single point. The mapping from rays to image points, however, is more complex than intersection with plane. (c) A catadioptric system consisting of a perspective camera and a parabolic mirror. Here the caustic is a surface. The mapping from the image plane, to the caustic curve and its tangent vectors, completely describes the geometry of system.

When the mapping is given numerically, a host of methods (Jensen, 1997; Mitchell and Hanrahan, 1992; Watt, 1990), may be used. When this mapping is known in closed form, analytic methods can be used to derive the caustic surface (Cornbleet, 1994; Born and Wolf, 1965; Burkhard and Shealy, 1973). The method we use assumes regularity of the mapping and computes the caustic by finding all the points where the change in coordinates described above is singular (Born and Wolf, 1965; Burkhard and Shealy, 1973).

Equation (3) expresses  $\mathcal{L}$  in terms of a known or measured ray surface  $\mathbf{s}(x, y)$ . The caustic is defined as the singularities in the change from  $(x, y, r)$  coordinates to  $(X, Y, Z)$  coordinates given by  $\mathbf{p}$ . Singularities arise at those points  $(X, Y, Z)$  where the Jacobian matrix  $J$  of the transformation does not have full rank. We find those points by computing the determinant of the Jacobian.

$$\det(J) = \begin{vmatrix} \frac{\partial p_x}{\partial x} + r \frac{\partial q_x}{\partial x} & \frac{\partial p_x}{\partial y} + r \frac{\partial q_x}{\partial y} & q_x \\ \frac{\partial p_y}{\partial x} + r \frac{\partial q_y}{\partial x} & \frac{\partial p_y}{\partial y} + r \frac{\partial q_y}{\partial y} & q_y \\ \frac{\partial p_z}{\partial x} + r \frac{\partial q_z}{\partial x} & \frac{\partial p_z}{\partial y} + r \frac{\partial q_z}{\partial y} & q_z \end{vmatrix}, \quad (4)$$

and setting it equal to zero. Since this is quadratic in  $r$  we can solve for  $r$  explicitly in terms of  $\mathbf{p}$ ,  $\mathbf{q}$ , and their first derivatives with respect to  $x$  and  $y$ . Plugging this back into  $\mathcal{L}$  gives us an expression for the caustic ray

surface parameterized by  $(x, y)$  as in Eq. (2). In general the caustic has two solutions with either acceptable as a ray surface.

If the optical system has translational or rotational symmetry then we need only consider one parameter, for example  $x$ , in the image plane. In this case the determinant Jacobian becomes linear in  $r$  and the solution simplifies to:

$$r = \frac{(q_x \frac{dp_y}{dx} - q_y \frac{dp_x}{dx})}{(q_y \frac{dq_x}{dx} - q_x \frac{dq_y}{dx})}. \quad (5)$$

Note that the determinant of the Jacobian represents the infinitesimal change volume between two coordinate systems in 3-dimensions, and in 2-dimensions the infinitesimal change in area. This is why the collapse of areas in the projections of Fig. 5(a) and (b) determine the caustic.

### 3.3. Derived Properties: Field of View

Some parameters used to specify a perspective camera model are derived from the ray surface representation in our general imaging model. Other parameters depend on the perspective assumption and are ill-defined in our model. For example, field of view presents an ambiguity, since in the non-perspective case the rays

may no longer form a simple rectangular cone. One candidate for a field of view is the range of  $\mathbf{q}(x, y)$  over the image. This is the same as the range of the Gauss mapping (Horn, 1986). The range of the Gauss mapping is a good approximation to the field of view when the scene points are distant relative to the size of the imaging system. If a single number is needed then the solid angle coming from this range may be used. Other geometric parameters of a perspective imaging system, such as aspect ratio and the distortion (Tsai, 1987), are simply ill-defined in the general imaging model.

#### 4. Non-Geometric Raxel Parameters

A raxel is just a very narrow field of view, perspective imaging system. Many of the conventional parameters associated with a perspective or near-perspective systems may be attributed to a raxel.

##### 4.1. Point Spread

Imagine our scene point as single impulse or point source of light. Suppose it is a distance  $r$  along the incoming ray from our raxel, as in Fig. 8(a). When only one raxel detects the light, that scene point is in perfect focus. More likely, nearby raxels will also detect some light from the impulse. The function that describes the response of the raxels to a single impulse at depth  $r$  is the point spread function.

Approximations to the point spread may be interpreted in terms of various aberrations such as spherical aberration, coma, and astigmatism, as well as higher order effects (Hecht, 1998; Guillemin and Sternberg, 1977). A complete description of aberrations associated with a general imaging system is a worthy subject in and of itself. Here we will satisfy ourselves by naively approximating the point spread by elliptical Gaussian as Fig. 8(b). For each depth the elliptical Gaussian has a major and a minor axis (assuming they are different). This in turn leads to a standard deviation,  $\sigma_M(r)$ ,  $\sigma_m(r)$  as well as an angle  $\psi(r)$  the major axis make with the  $x$  axis in the image plane. This is enough to capture effects such as local spherical aberration and astigmatism, but not coma.

In Section 2.2 we pointed out that a raxel with a linear response could be modeled as a delta function. Point spread smears the delta function out. Thus a raxel is really a density on positions and directions. When we integrate against the plenoptic function of a scene, the number obtained represents the measurement at that raxel.

##### 4.2. Radiometry

Recall that in Section 2 we isolated two essential components for an imaging device: photosensitive elements and imaging optics. We have seen that these components choose the rays which effect a raxel. These components also effect how a raxel responds to intensities along that ray.

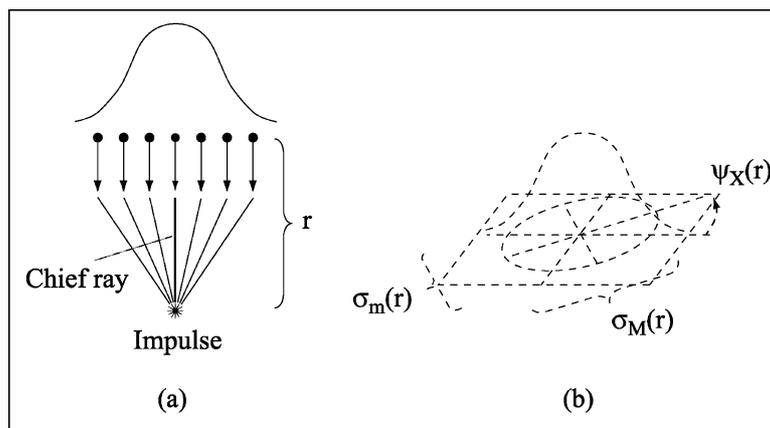


Figure 8. (a) The point spread of a raxel. Assume an impulse at distance  $r$  lies on a chief ray of a raxel. Light from that impulse may be measured at other raxels perceived as blur in the image. Here it is modeled, naively with an Gaussian. (b) An elliptical Gaussian point spread function. The ellipse has a major and minor axis, when their lengths,  $\sigma_M$ ,  $\sigma_m$ , are distinct, as well as the angle  $\psi$  the major axis makes with the  $x$  axis in the image. The parameters of the point spread here are functions of  $r$ , distance to the impulse.

As light passes through the imaging optics some light is lost. Moreover, since an imaging device has a finite aperture, our idealized chief ray really represents a bundle of rays. The size of that bundle, along with the intensities of light, determines the irradiance at the photosensitive elements. The size of the bundle of rays generally varies across the image. At any point  $(x, y)$  in the image, we can model the attenuation of the scene radiance due to the optics by a single factor  $h(x, y)$ , which we will call the *fall-off function*. This accounts for such effects as vignetting. For a simple lens, one can derive that this function should be proportional to  $\cos^4 \phi$  (Horn, 1986). It is important to note that systems like those shown in Fig. 7(b) and (c) have varying resolution. As resolution changes, the size of the bundle of rays associated to a chief ray will change too, effecting the fall-off. When the fall-off is normalized so its maximum is one, we call it the *normalized fall-off*.

While the effect on the intensities due to the optics may be described by fall-off, which is a linear factor for each point in the image, the response  $g(e)$  of the photosensitive elements need not be linear in the intensity  $e$ . In fact both for film and in electronic imaging, it is often chosen to be non-linear. Sometimes this is done to enhance or decrease contrast or to capture a wider range of intensities. We may assume that  $g$ , called *radiometric response function*, is smooth and monotonic.

It may be approximated by a gamma function, a polynomial, or discretely with a regularity constraint (Mann and Picard, 1995; Debevec and Malik, 1997; Mitsunaga and Nayar, 1999). If one can compute the radiometric response function of each raxel, one can linearize the response with respect to scene radiance, assuming the response is monotonic. It is only after linearizing the response that we can treat a raxel as a delta function times a constant,  $h(x, y)$ .

#### 4.3. Complete Imaging Model

The general imaging model consists of a set of raxels parameterized by  $x$  and  $y$  in pixel coordinates. The parameters associated with these raxels (see Fig. 9), are (a) position and direction, that describe the ray surface of the caustic, (b) major and minor standard deviations as well as an elliptical orientation, each a function of depth, (c) a radiometric response function, and (d) a fall-off constant for each pixel.

Camera parameters are separated into external and internal parameters. A coordinate frame specifies the

Index		Geometry		Point Spread			Response	Fall-off
$x$	$y$	$p$	$q$	$\sigma_M(r)$	$\sigma_m(r)$	$\Psi(r)$	$g(e)$	$h$

Figure 9. Each raxel has the above parameters of image coordinates, position and direction in space, major and minor standard deviation, elliptical orientation, radiometric response, and fall-off factor. These parameters are measured with respect to a coordinate frame fixed to the imaging system.

external parameters. The internal parameters (Fig. 9) are specified with respect to that coordinate frame. In particular, for each raxel  $i$ , the parameters  $\mathbf{p}_i$ ,  $\mathbf{q}_i$  are measured with respect to a single coordinate frame fixed to the imaging system. If the system is rotated or translated, these parameters will not change but the coordinate frame will.

In the case of perspective projection, the essential (Faugeras, 1992) or fundamental (Hartley, 1993) matrix provides the relationship between points in one image and lines in another image (of the same scene). Some progress has been made in generalizing the machinery of projective imaging to the general imaging model (Pless, 2002, 2003; Neumann et al., 2003).

## 5. Finding the Model Parameters

Determining the general imaging model parameters for a given imaging system requires determining the geometric and non-geometric parameters of all the raxels. We begin by discussing the determination of the geometric raxel parameters. The geometric raxel model, whether discrete, continuous, or caustic, describes the correspondence between the location of points in image plane and rays in space. When the components of the imaging system are known this may be computed. When the components of the imaging system are not known this correspondence must be determined experimentally. In order to determine this correspondence we construct a calibration environment where the geometric and radiometric parameters can be efficiently estimated. It is important to note that given the non-perspective nature of a general device, conventional geometric calibration methods are not directly applicable (Faugeras, 1992; Hartley, 1993; Tsai, 1987).

For this calibration assume we have a known plane in the scene. Suppose that by imaging this plane we can establish a correspondence between each pixel  $i$  of the image and a point  $p_f$  on the plane whose location is known (Fig. 10). This determines that the incoming

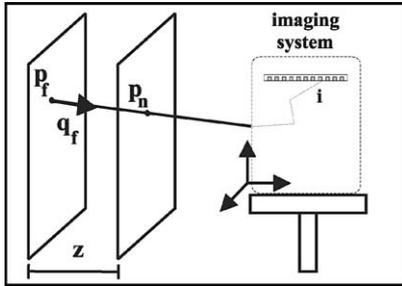


Figure 10. A ray corresponding to a pixel  $i$  intersects two planes, separated by a known distance  $z$ , at points  $\mathbf{p}_n$  and  $\mathbf{p}_f$ . If these positions are known, the direction of the ray  $\mathbf{q}_f$  may be determined for each pixel. From this we compute the raxel parameters. To determine from the image the positions of  $\mathbf{p}_n$  and  $\mathbf{p}_f$  we place an active display at the planes and use binary coding patterns. By rotating the image device we may perform ray-based calibration for imaging systems with an arbitrary field of view.

ray associated to pixel  $i$  passes through the point  $p_f$ . We now move the plane by a known distance  $z$ . Again we establish the point  $p_n$  corresponding to the pixel  $i$ . The two points  $p_n, p_f$  determine the ray and thus the direction of the ray  $\mathbf{q}_f$ .

If we assume our raxels to be located at the first plane then the pairs  $(p_f, q_f)$  determine the discrete raxel model. Suppose that the geometry of imaging system may be assumed continuous. Given dense measurements of all the pixels, we fit a continuous mapping to the discrete samples to obtain the continuous raxel model. This model yields a ray surface where  $p_f$  and  $q_f$  are parameterized by image plane coordinates. If we assume this mapping is smooth we can determine the caustic for the caustic raxel model by the Jacobian method described in Section 3.2. We emphasize that for a black box imaging system a dense calibration is required.<sup>8</sup> The two-plane method densely recovers the image plane to scene ray correspondence for systems with a limited field of view. We may calibrate wide-field of view systems by placing them on a turntable and rotating the coordinate frame of the imaging system.

### 5.1. Caustic Raxel Model: Simulation

We simulated the recovery of the caustic for three imaging systems. The first system is a catadioptric system consisting of a perspective camera pointed at a spherical reflector. The system is assumed to be rotationally symmetric about the optical axis. Due to symmetry we can solve for the caustic along a cross section of the system as shown in Fig. 11(a). The sphere was as-

sumed to have a radius of  $\sqrt{2}/8$  units and the center of the projection was at .25 units from the sphere center. The parameters were chosen so that with a 60 degree field of view of the perspective camera, the spherical system would have a 120 degree field of view. We assumed that we had one plane at a distance .5 units along the optical axis and one plane at 1 unit along the axis.

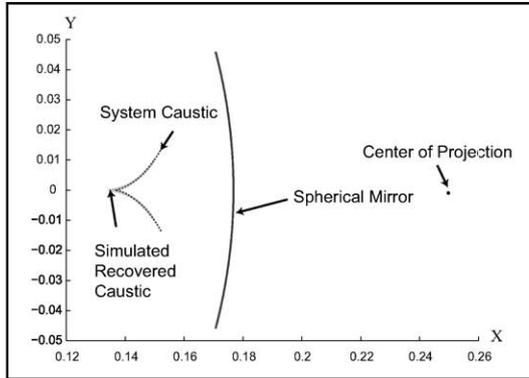
Along the cross section of the planes we created a uniform distribution of reference points perpendicular to the axis of symmetry. The positions of these points were assumed known precisely. In practice the accuracy of the locations of the world points is not a limiting factor. The limiting factor on accuracy is the resolution of the imaging system. We simulated this by assuming a finite spatial resolution in the imaging plane and adding half a pixel of uniformly distributed noise to the measured locations in the image plane.

The recovery of the caustic involved the following steps:

1. We smoothed the simulated noisy measurements of the image plane locations using a local linear regression.
2. We then fit a cubic spline to the data to obtain a smooth correspondence between the scene planes and the image plane.
3. From the difference of the spine mappings we were able to obtain a mapping from the image plane to directions of the corresponding rays.
4. We used Eq. (5) to compute the caustic.

In Fig. 11(a), the curve labelled ‘‘System Caustic’’ was the recovered caustic without noise while ‘‘Simulated Recovered Caustic’’ was recovered with noise. The resolution was assumed to be 1000 along the cross section for the spline shown in the figure. Figure 11(b) shows the mean RMSE for 10 simulated recoveries with the same parameters at different resolutions. The standard deviation of the RMSE is also shown. The recovered curve has the correct shape and is localized fairly accurately.

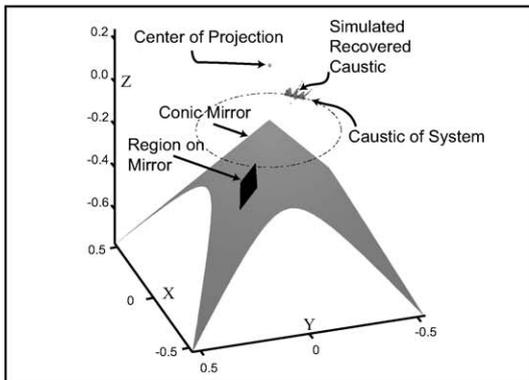
For our second example consisted of a perspective camera viewing a conic mirror as shown in Fig. 11(c). The system is rotationally symmetrical but we performed the recovery in 3D. The perspective camera’s single viewpoint was again located at a distance of .25 units from the origin and the equation of the cone was  $4x^2 + 4y^2 = 4z^2$ . We assumed that the scene points imaged were uniformly spaced for planes at  $z = 1$  and  $z = 5$ . Again, these points were assumed to be known



(a)

Spherical Reflector System		
Effective Resolution	Mean RMSE	Std RMSE
1000 <sup>2</sup> pixels	0.0016178	0.0000372
2500 <sup>2</sup> pixels	0.0016175	0.0000148
5000 <sup>2</sup> pixels	0.0016130	0.0000081

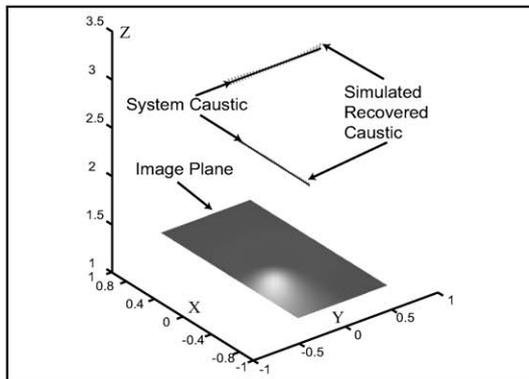
(b)



(c)

Conic Reflector System		
Effective Resolution	Mean RMSE	Std RMSE
1000 <sup>2</sup> pixels	0.0075317	0.00072648
2500 <sup>2</sup> pixels	0.0071549	0.00097787
5000 <sup>2</sup> pixels	0.0069886	0.00075699

(d)



(e)

Cross Slit System			
Effective Resolution	Slit	Mean RMSE	Std RMSE
1000 <sup>2</sup> pixels	Near	0.0063847	0.0010702
	Far	0.013916	0.0021191
2500 <sup>2</sup> pixels	Near	0.0027834	0.00076479
	Far	0.0054602	0.0014079
5000 <sup>2</sup> pixels	Near	0.001717	0.00024404
	Far	0.0036491	0.00065923

(f)

Figure 11. Simulated recovery of the caustic raxel model for three imaging systems. In (a) the system consists of a spherical mirror and perspective camera. Because the system is rotationally symmetric the caustic was recovered from a cross section. Localization of position in the image plane was corrupted by 1/2 pixel of uniform noise. The mean errors and standard deviations of the RMSE errors for 10 simulated recoveries are shown in (b). In (c) the imaging system consists of a perspective camera and conical mirror. Here a full 3D recovery is performed with the quantitative results shown in (d). In (e) a simulated crossed slit imaging system is calibrated with the errors shown in (f).

precisely but half a pixel of uniform noise was added to the simulated measured positions in the image pane.

As before, the simulated recovered data was first smoothed with a local linear regression. The image to

plane mappings were fit with 5th order tensor product splines. To avoid extrapolation with the spline we chose a square sub-patch on the scene plane which corresponds to the dark blue patch on the mirror surface

shown in Fig. 11(c). To obtain the entire caustic we would repeat this procedure for a set of square patches covering the field of view. The caustic was recovered using Eq. (4). The expected caustic should be a circle. The noise free computed caustic did give points along a circle. The simulated recovery was exceptionally sensitive to noise as can be seen from the table of errors in Fig. 11(d). Even with smoothing the results were essentially a point cloud around the predicted caustic.

Our third example system is a cross slit system (Feldman et al., 2002). The rays in this system pass through two fixed orthogonal lines before intersecting the image plane as shown in Fig. 11(e). The caustic was recovered using as in the second example. The table of errors is shown in Fig. 11(f). While the caustic is localized properly the quality of the results were again sensitive to resolution.

From our simulation we found that when symmetry of the imaging system can be exploited a reasonably accurate recovery of both the shape and location of the caustic is possible using Eq. (5). Accurate recovery of the caustic of a completely black box system is possible using Eq. (4) but requires very high resolution. Even then smoothing may be required and the results are likely to be very sensitive to the choice of regularization. Nevertheless even a point cloud caustic can give useful information such as an estimate of the size of the caustic. For a small caustic parallax effects may be neglected and the system can be assumed to have an effective single viewpoint. The sensitivity to resolution means that without a-priori knowledge of the imaging system we can not calibrate the system from sparse features.

### 5.2. Calibration Using an Active Display

To create distinguishable features for each pixel  $i$  we place an active display at the plane and take multiple images. If a display has  $N$  locations, we can make each point distinct in  $\log N$  images using simple gray coding or bit coding like those described in Sato and Inokuchi (1985). A million pixel display requires only twenty images to encode every point. A calibrated active display also permits the calibration of the non-geometric parameters of the model. We display a linearly increasing constant brightness sequence to our imaging system. First, we calibrate the radiometric response function for a representative point in the image from the known input brightness. We may then compute the fall-off function across all the points.

We used twenty binary coded images for each plane for ray-based calibration (assuming a one mega-pixel display). We compute both the radiometric response function and the fall-off from seventeen uniform brightness levels. Thus, with roughly 60 images we have measured the parameters of our general imaging model, with the exception of point spread. To make those 60 images robust to noise each image was an average of 30 captured images of the same display input.

### 5.3. Experimental Apparatus

We calibrated two imaging systems to demonstrate the generality of our method. One imaging system was a perspective camera. The second imaging system was a non-perspective catadioptric system consisting of a perspective camera and a parabolic mirror (Swaminathan et al., 2001). In both experiments the camera was a Cannon Optura digital video camera. Bit patterns were displayed on a laptop with a 14.1 inch LCD screen with resolution  $1024 \times 768$  pixels. Rather than move the display, the camera was mounted on a stage, which was translated 60 mm in the direction normal to the screen.

The perspective imaging system, consisting of just the camera itself, can be seen in Fig. 12(a). Figure 12(b) shows an image of a pattern of vertical bars from the perspective camera. In Fig. 13(a) the parabolic catadioptric system is shown. This is the same system we described in Fig. 7(c). Not only is the system not perspective but also the caustic of the system can be seen to form a curve in the plane, or a surface when the figure is rotated about the axis of symmetry. The parabolic mirror we used had an outer diameter of 100 mm and an inner diameter of 3 mm. The focus of the parabola was 25 mm from the base. The laptop was oriented so as to give the maximum screen resolution along the axis of symmetry. Figure 13(b) shows a sample binary pattern as seen from the parabolic catadioptric system.

### 5.4. Experimental Results: Geometric Parameters

Even though the raw binary patterns were averaged to reduce noise, pixels imaged on the edge, between black and white regions, must be ignored as they give an ambiguous response. When the information for the patterns is combined this creates holes for which we do

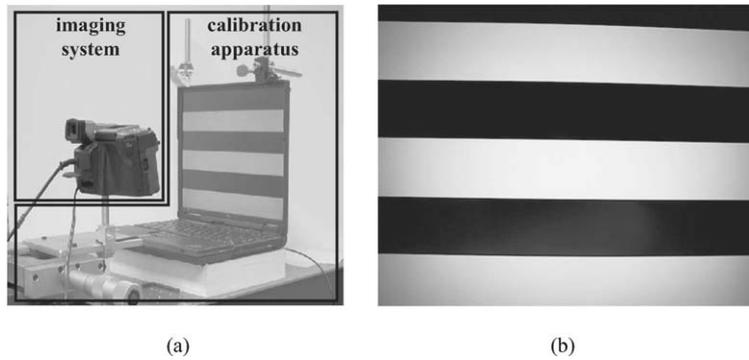


Figure 12. (a) A calibration apparatus, which consists of a laptop and a translating stage, and perspective imaging system consisting of a Canon Optura video camera. The axis of perspective camera is normal to the plane of the screen. The laptop displays 26 patterns. Images were taken at two positions of the translating stage. (b) A sample bit pattern as seen through the perspective system.

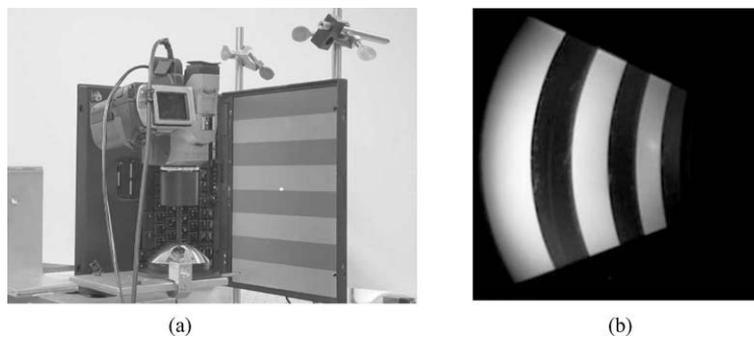


Figure 13. (a) The same calibration apparatus that was used for the perspective system, here with a non-perspective imaging system. The non-perspective catadioptric system consists of a perspective camera a parabolic mirror. (b) A sample bit pattern as seen through the parabolic catadioptric system.

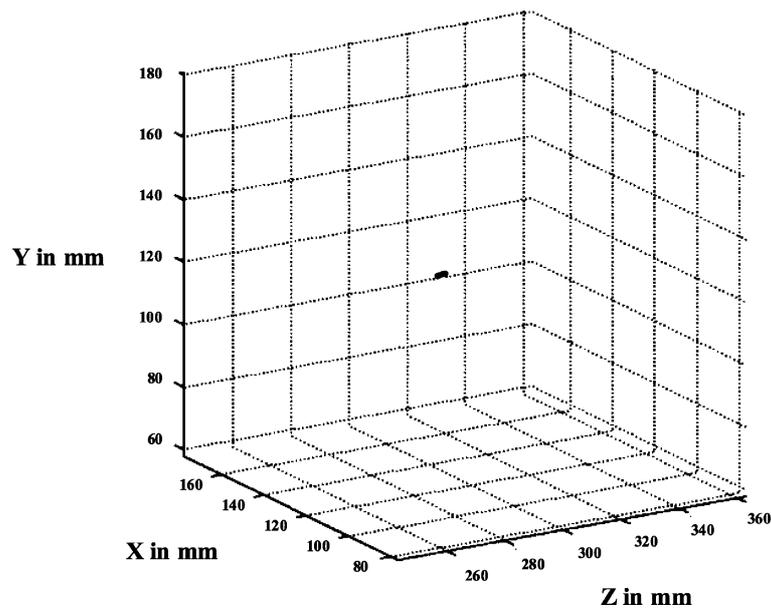


Figure 14. The perspective caustic is a small cluster of points corresponding to the single viewpoint.

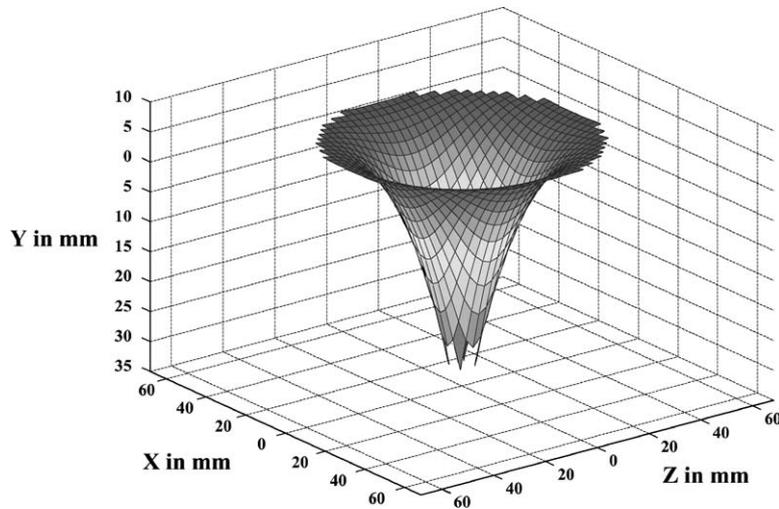


Figure 15. The caustic recovered for a parabolic catadioptric system. A cross section of the caustic of the system is similar shown in Fig. 7(c).

not have data. In the perspective case we fit a polynomial to our data. We found a second-degree polynomial produced a good fit. Once we had parameterized the positions on the near and far planes in image coordinates, we were able to find directions at the far plane. This gave us a ray surface for the far plane. Using this we solved for the zeros of the 3-dimensional Jacobian of Eq. (4) to recover the caustic scatter plot as seen in Fig. 14. Since the caustic of a perspective camera is the single viewpoint, the point-like measured caustic agrees with expectation.

Figure 15 shows the recovery of the caustic for the non-perspective catadioptric system using ray-based calibration. The system is rotationally symmetric so we used the 2-dimensional recovery method which is less sensitive to limited resolution. The caustic was computed using Eq. (5). Since there is no data along the axis of symmetry, its position was estimated. A cross section of the system is shown in Fig. 7(c). The caustic is shown as the envelope of rays (see Swaminathan et al. (2001) for details). The recovered caustic of Fig. 15 matches that part of the curve near the cusp. We only see this part of the caustic because the outer field of view of the imaging system ends where the rays are nearly normal to the mirror's axis of symmetry. Near the axis of symmetry the radial partial derivative increases faster. This means that near the axis, smaller changes in position on the caustic yield larger changes in angle. Thus, there is a drop in resolution near the axis of symmetry.

### 5.5. Experimental Results: Non-Geometric Parameters

Figure 16(a) shows the normalized radiometric response function. As is the case with higher quality digital imaging devices, the response is close to linear over much of its dynamic range. The curve shown is a polynomial fit with the non-linear constraint that the curve must have a positive derivative.

For the perspective system, we might expect a  $\cos^4 \alpha$  type fall-off (Horn, 1986). However, many lens systems have been designed to remove this effect. Indeed, we

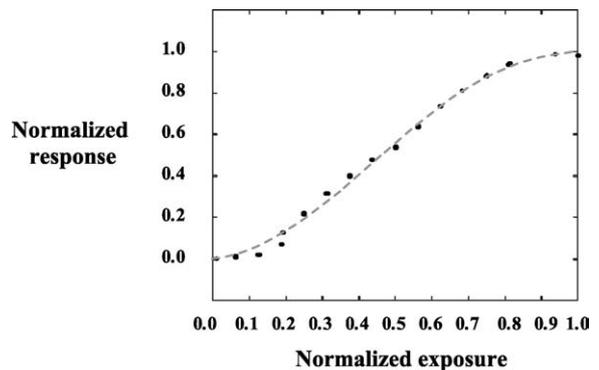


Figure 16. The normalized radiometric response is calculated from images of 17 uniform screens. The curve shown is a polynomial fit with the endpoints constrained and the first derivative required to be positive.

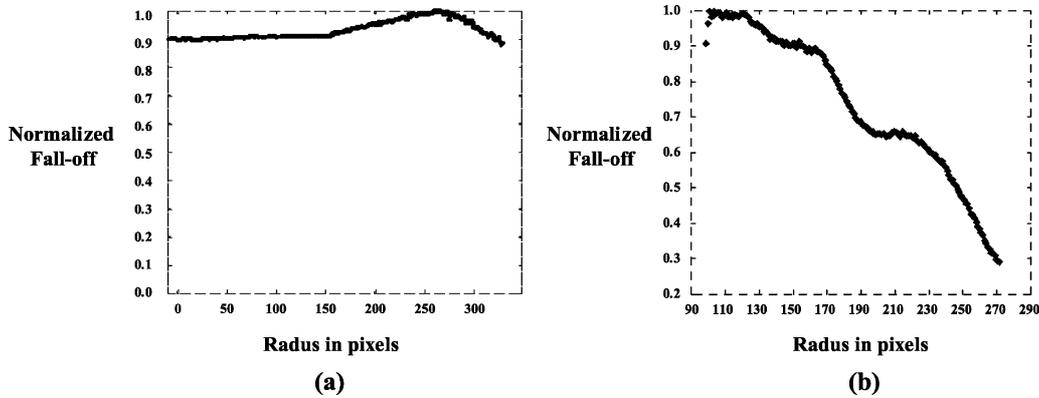


Figure 17. (a) A plot of the radial fall-off function for the perspective system. It is nearly constant because the optics and electronics were designed to minimize this effect. (b) The radial fall-off function of the parabolic catadioptric system. Because resolution is higher at greater radii, the amount of light collected there drops.

find no significant fall-off as the function is nearly constant. This can be seen in the graph of a radial slice 16(b).

We linearized the response and calculated the radiometric fall-off for the non-perspective case (Fig. 17). The fall-off is normalized so the maximum is unity. The directionality of the LCD irradiance has been measured and normalized to be uniform with respect to direction. The pronounced fall-off away from the axis of symmetry is due to the higher resolution of non-perspective system there. Image pixels away from the axis of symmetry see a smaller object area. Hence, they gather less light from a uniform radiator. To conclude, this simple calibration procedure allows us to compute the parameters of our general and flexible model.

## 6. Conclusion

We have provided a general model for imaging systems. This provides a unified model for growing number of imaging systems which do not satisfy the perspective imaging model. For finite resolution systems this geometry can always be represented as a list of raxels with geometric and radiometric parameters. By assuming the correspondence is continuous we can represent the set of raxels in terms of a ray surface. A special and essentially unique ray surface called a caustic provides a generalization for the single viewpoint in the perspective model. Since an early version of this model was proposed in Grossberg and Nayar (2001), it has been shown that the epi-polar constraints, ego-motion and structure from motion can be generalized

for this model (Pless, 2002, 2003; Neumann et al., 2003). We have also shown that we can attach non-geometric parameters to raxels to provide a complete imaging model.

We have shown that our calibration technique yields excellent results for symmetric systems and gives acceptable results for general systems. Recovery of the caustic for general system is sensitive to the limits of image resolution and noise. Thus there is room for improvement in finding the ideal regularization to recover the caustic in the general case. This model provides a framework for camera designers to explore a wide range of new camera designs. We also note that because cameras and projectors are in some sense dual our model can also be applied to projection systems.

## Acknowledgments

This work was supported in parts by a National Science Foundation ITR Award, IIS-00-85864, and a DARPA/ONR MURI Grant, N00014-95-1-0601. The authors would also like to thank Rahul Swaminathan, Yoav Schechner, Srinivasa Narasimhan, and Joshua Gluckman for their comments. Parts of these results appeared in the proceedings of ICCV'01.

## Notes

1. This is similar to the observation of Mann, who noted that a perspective system could be considered a bundle of light detecting elements he called a pencigraph (Mann, 1996).

2. We really only require two angles  $\phi$ ,  $\theta$  to specify direction but it is more convenient to work with vectors. The plenoptic function is simply assumed constant as  $|q|$  varies.
3. The delta functions are not actually functions. They are dual to functions in that they take functions as input and produce a number as output. Delta functions are a special case of a *distribution* (Zemanian, 1965). Since we have assumed a raxel to be localized at a point and to view along a single ray we have not considered more general distributions.
4. We may place a raxel anywhere along a ray. This is because intensities usually do not change much along a ray (particularly when the medium is air) provided the displacement is small with respect to the total length of the ray.
5. For a camera cluster it is usually more convenient to index each camera by an independent index. This can be thought of as a hybrid between the case of a continuous mapping and the discrete mapping used for the discrete raxel model. Rather than listing individual raxels, we list the continuous mappings corresponding to the cameras.
6. Generically this must be a surface. For certain imaging systems however, it is possible to choose a ray curve or even point. For example in the perspective model the ray “surface” can be chosen to be the single viewpoint.
7. We have explicitly assumed that our image plane is 2-dimensional. An  $b$ -dimensional image plane would lead to a mapping from a  $b + 1$  parametrization of the rays to 3-dimensional space. In this case, the caustic is still the locus of singularities of corresponding map. It is not clear in this case whether the geometry of the imaging system is completely determined by the caustic.
8. If enough is known about a particular non-perspective system, calibration can also be accomplished using the correspondence of known 3D points under known motion. For example in Swaminathan et al. (2001), a non-perspective system was assumed to belong to a family of models specified by a small number of parameters. In this case the parameters can be estimated from sparse feature correspondence under known motion.

## References

- Adelson, E.H. and Bergen, J.R. 1991. *The Plenoptic Function and the Elements of Early Vision*. The MIT Press.
- Baker, S. and Nayar, S.K. 1999. A theory of single-viewpoint catadioptric image formation. *IJCV*, 35(2):1–22.
- Baker, P., Fermüller, C., Aloimonos, Y., and Pless, R. 2001. A spherical eye from multiple cameras (makes better models of the world). In *CVPR01*, pp. I:576–583.
- Born, M. and Wolf, E. 1965. *Principles of Optics*. Pergamon Press.
- Brown, D.C. 1966. Decentering distortion of lenses. *Photogrammetric Engineering*, 32(3):444–462.
- Burkhard, D.G. and Shealy, D.L. 1973. Flux density for ray propagation in geometrical optics. *Journal of the Optical Society of America*, 63(3):299–304.
- Conrady, A. 1919. Decentering lens systems. *Monthly Notices of the Royal Astronomical Society*, 79:384–390.
- Cornbleet, S. 1994. *Microwave and Geometrical Optics*. Academic Press: London.
- Dawkins, R. 1996. *Climbing Mount Improbable*. Norton: New York.
- Debevec, P.E. and Malik, J. 1997. Recovering high dynamic range radiance maps from photographs. *SIGGRAPH*, 31:369–378.
- Faugeras, O.D. 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. ECCV*, pp. 563–578.
- Faugeras. 1993. *Three-Dimensional Computer Vision*. MIT Press.
- Feldman, D., Zomet, A., Peleg, S., and Weinshall, D. 2002. Video synthesis made simple with the x-slits projection. In *Motion02*, pp. 195–200.
- Gaten, E. 1994. Geometrical optics of a galatheid compound eye. *Journal Comparative Physiology A*, 175:749–759.
- Gershun, A. 1939. Svetovoe pole (the light field, in English). *Journal of Mathematics and Physics*, XVIII:51–151.
- Gluckman, J. and Nayar, S.K. 1999. Planar catadioptric stereo: Geometry and calibration. In *CVPR99*, pp. I:22–28.
- Gortler, S.J., Grzeszczuk, R., Szeliski, R., and Cohen, M. 1996. The lumigraph. In *Computer Graphics, Proc. SIGGRAPH*, p. 43.
- Goshtasby, A. 1989. Correction of image deformation from lens distortion using bezier patches. *Computer Vision, Graphics, and Image Processing*, 47:385–394.
- Griffiths, P. and Harris, J. 1978. *Principles of Algebraic Geometry*. Wiley Interscience.
- Grossberg, M.D. and Nayar, S.K. 2001. A general imaging model and a method for finding its parameters. In *ICCV01*, pp. II:108–115.
- Guillemin, V. and Sternberg, S. 1977. *Geometric Asymptotics*. American Mathematical Society.
- Hammond, J.H. 1981. *The Camera Obscura. A Chronicle*. Bristol: Adam Hilger Ltd.
- Hartley, R. 1993. Estimation of relative camera positions for uncalibrated cameras. In *Proc. ECCV*, pp. 579–587.
- Hecht, E. 1998. *Optics*, 3rd edition. Addison-Wesley Publishing.
- Horn, B. 1986. *Robot Vision*. The MIT Press.
- Jensen, H.W. 1997. Rendering caustics on non-lambertian surfaces. *Computer Graphics Forum*, 16(1):57–64.
- Langer, M.J. and Zucker, S.W. 1997. What is a light source? In *Proc. CVPR*, pp. 172–178.
- Levoy, M. and Hanrahan, P. 1996. Light field rendering. In *Computer Graphics, Proc. SIGGRAPH*, p. 31.
- Mann, S. and Picard, R. 1995. Being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proceedings of IS&T, 46th Annual Conference*, pp. 422–428.
- Mann, S. 1996. Pencigraphy with AGC: Joint parameter estimation in both domain and range of functions in same orbit of the projective-wyckoff group. <http://citeseer.nj.nec.com/36900.html>.
- McCutchen, D. 1991. Method and apparatus for dodecahedral imaging system. U. S. Patent 5,023,725, June 11.
- Mitchell, D. and Hanrahan, P. 1992. Illumination from curved reflectors. *Computer Graphics*, 26(4):283–291.
- Mitsunaga, T. and Nayar, S.K. 1999. Radiometric self calibration. In *Proc CVPR*, vol. 2, pp. 374–380.
- Miyamoto, K. 1994. Fish eye lens. *Journal of Optical Society of America*, 54(8):1060–1061.
- Moini, A. 2000. *Vision Chips*. Kluwer Academic Publishers.
- Nayar, S.K. and Baker, S. 1997. Catadioptric image formation. In *DARPA97*, pp. 1431–1438.
- Nayar, S. and Karmarkar, A. 2000. 360 × 360 mosaics.
- Nayar, S.K. and Peri, V. 1999. Folded catadioptric cameras. In *CVPR99*, pp. II:217–223.
- Neumann, J., Fermüller, C., and Aloimonos, Y. 2003. Polydioptric camera design and 3d motion estimation. In *CVPR03*, pp. II:294–301.

- Pajdla, T. 2002. Stereo with oblique cameras. *IJCV*, 47(1–3):161–170.
- Peleg, S., Pritch, Y., and Ben-Ezra, M. 2000. Cameras for stereo panoramic imaging. In *Proc. CVPR*, IEEE Computer Society, vol. 2, pp. 201–214.
- Pless, R. 2002. Discrete and differential two-view constraints for general imaging systems. In *OMNIVIS02*, pp. 53–59.
- Pless, R. 2003. Using many cameras as one. In *CVPR03*, pp. II:587–593.
- Sato, K. and Inokuchi, S. 1985. Three-dimensional surface measurement by space encoding range imaging. *Journal of Robotic Systems*, 2:27–39.
- Seitz, S.M. and Kim, J. 2001. The space of all stereo images. In *ICCV01*, pp. I:26–33.
- Swaminathan, R. and Nayar, S.K. 2000. Nonmetric calibration of wide-angle lenses and polycameras. *PAMI*, 22(10):1172–1178.
- Swaminathan, R., Grossberg, M.D., and Nayar, S.K. 2001. Caustics of Catadioptric Cameras. In *Proc. ICCV*.
- Tsai, R.Y. 1987. A versatile camera calibration technique for high-accuracy 3d machine vision. *International Journal of Robotics and Automation*, 3(4):323–344.
- Tyson, R. 1998. *Principles of Adaptive Optics*. Academic Press.
- Watt, M. 1990. Light-water interaction using backward beam tracing. *Computer Graphics*, 24(4):377–385.
- Weng, J., Cohen, P., and Herniou, M. 1992. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10): 965–980.
- Yagi, Y. and Kawato, S. 1990. Panoramic scene analysis with conic projection. In *IROS90*.
- Zemanian, A.H. 1965. *Distribution Theory and Transform Analysis*. Dover.