In "Panoramic Imaging: Sensors, Theory, and Applications," Ed. by Benosman and Kang.

Single Viewpoint Catadioptric Cameras

Simon Baker
The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Shree K. Nayar
Department of Computer Science
Columbia University
New York, NY 10027

Abstract

Conventional video cameras have limited fields of view which make them restrictive for certain applications in computational vision. A catadioptric camera uses a combination of lenses and mirrors placed in a carefully arranged configuration to capture a much wider field of view. One important design goal for catadioptric cameras is choosing the shapes of the mirrors in a way that ensures that the complete catadioptric system has a single effective viewpoint. The reason a single viewpoint is so desirable is that it is a requirement for the generation of pure perspective images from the sensed images. In this chapter, we derive the complete class of single-lens single-mirror catadioptric cameras that have a single viewpoint. We describe all of the solutions in detail, including the degenerate ones, with reference to many of the catadioptric systems that have been proposed in the literature. In addition, we derive a simple expression for the spatial resolution of a catadioptric camera in terms of the resolution of the cameras used to construct it, and present an analysis of the defocus blur caused by the use of a curved mirror in a catadioptric camera. We end with a case study describing several implementations of one of the possible designs.

1 Introduction

Many applications in computational vision require that a large field of view is imaged. Examples include surveillance, teleconferencing, and model acquisition for virtual reality. A number of other applications, such as ego-motion estimation and tracking, would also benefit from enhanced fields of view. Unfortunately, conventional imaging systems are severely limited in their fields of view. Both researchers and practitioners have therefore had to resort to using either multiple or rotating cameras in order to image the entire scene.

One effective way to enhance the field of view is to use mirrors in conjunction with lenses. See, for example, [Rees, 1970], [Charles *et al.*, 1987], [Nayar, 1988], [Yagi and Kawato, 1990], [Hong, 1991], [Goshtasby and Gruver, 1993], [Yamazawa *et al.*, 1993], [Bogner, 1995], [Nalwa, 1996], [Nayar, 1997a], and [Chahl and Srinivassan, 1997]. We refer to the approach of using mirrors in combination with conventional imaging systems as *catadioptric* image formation. *Dioptrics* is the science of refracting elements (lenses) whereas *catoptrics* is the science of reflecting surfaces (mirrors) [Hecht and Zajac, 1974]. The combination of refracting and reflecting elements is therefore referred to as catadioptrics.

As noted in [Rees, 1970], [Yamazawa *et al.*, 1995], [Nalwa, 1996], and [Nayar and Baker, 1997], it is highly desirable that a catadioptric system (or, in fact, any imaging system) have a single viewpoint (center of projection). The reason a single viewpoint is so desirable is that it permits the generation of geometrically correct perspective images from the images captured by the catadioptric cameras. This is possible because, under the single viewpoint constraint, every pixel in the sensed images measures the irradiance of the light passing through the viewpoint in one particular direction. Since we know the geometry of the catadioptric system, we can precompute this direction for each pixel. Therefore, we can map the irradiance value measured by each pixel onto a plane at any distance from the viewpoint to form a planar perspective image. These perspective images can subsequently be processed using the vast array of techniques developed in the field of computational vision that assume perspective projection. Moreover, if the image is to be presented to a human, as in [Peri

and Nayar, 1997], it needs to be a perspective image not to appear distorted. Naturally, when the catadioptric imaging system is omnidirectional in its field of view, a single effective viewpoint permits the construction of panoramic images as well as perspective ones.

In this chapter, we take the view that having a single viewpoint is the primary design goal for the catadioptric camera and restrict attention to catadioptric cameras with a single effective viewpoint [Baker and Nayar, 1998] [Baker and Nayar, 1999]. However, for many applications, such as robot navigation, having a single viewpoint may not be a strict requirement [Yagi *et al.*, 1994]. In these cases, cameras that do not obey the single viewpoint requirement can also be used. Then, other design issues become more important, such as spatial resolution, camera size, and the ease of mapping between the catadioptric images and the scene [Yamazawa *et al.*, 1995]. Naturally, it is also possible to investigate these other design issues. For example, Chahl and Srinivassan recently studied the class of mirror shapes that yield a linear relationship between the angle of incidence onto the mirror surface and the angle of reflection into the camera [Chahl and Srinivassan, 1997].

We begin this chapter by deriving the entire class of catadioptric systems with a single effective viewpoint, and which can be constructed using just a single conventional lens and a single mirror. As we will show, the 2-parameter family of mirrors that can be used is exactly the class of rotated (swept) conic sections. Within this class of solutions, several swept conics are degenerate solutions that cannot, in fact, be used to construct cameras with a single effective viewpoint. Many of these solutions have, however, been used to construct wide field of view cameras with non-constant viewpoints. Some of the non-degenerate solutions have also been used in cameras proposed in the literature. In both cases, we mention all of the designs that we are aware of. A different derivation of the fact that only swept conic sections yield a single effective viewpoint was recently suggested by Drucker and Locke [1996].

A very important property of a camera that images a large field of view is its resolution. The resolution of a catadioptric camera is not, in general, the same as that of any of the cameras used to construct it. In Section 3, we study why this is the case, and derive a simple expression for the relationship between the resolution of a conventional imaging

system and the resolution of a derived single-viewpoint catadioptric camera. We specialize this result to the mirror shapes derived in the previous section. This expression should be carefully considered when constructing a catadioptric imaging system in order to ensure that the final camera has sufficient resolution. Another use of the relationship is to design conventional cameras with non-uniform resolution, which when used in an appropriate catadioptric system have a specified (e.g. uniform) resolution.

Another optical property which is affected by the use of a catadioptric system is focusing. It is well known that a curved mirror increases image blur [Hecht and Zajac, 1974]. In Section 4, we analyze this effect for catadioptric cameras. Two factors combine to cause additional blur in catadioptric systems: (1) the finite size of the lens aperture, and (2) the curvature of the mirror. We first analyze how the interaction of these two factors causes defocus blur and then present numerical results for three different mirror shapes: the hyperboloid, the ellipsoid, and the paraboloid. The results show that the focal setting of a catadioptric camera using a curved mirror may be substantially different from that needed in a conventional camera. Moreover, even for a scene of constant depth, significantly different focal settings may be needed for different points in the scene. This effect, known as *field curvature*, can be partially corrected using additional lenses [Hecht and Zajac, 1974].

As a case study, in Section 5 we describe several implementations of single viewpoint catadioptric cameras using paraboloid mirrors. We outline the construction of the cameras, their calibration, and the real-time software that we developed to unwarp the catadioptric images to give perspective images. We conclude this chapter with a discussion of the design issues involved when building single viewpoint catadioptric cameras.

2 The Fixed Viewpoint Constraint

The fixed viewpoint constraint is the requirement that a catadioptric camera only measure the intensity of light passing through a single point in 3-D space. The direction of the light passing through this point may vary, but that is all. In other words, the catadioptric camera

must sample the 5-D plenoptic function [Adelson and Bergen, 1991] [Gortler *et al.*, 1996] at a single point in 3-D space. The fixed 3-D point at which a catadioptric camera samples the plenoptic function is known as the *effective viewpoint*.

Suppose we use a single conventional camera as the only sensing element and a single mirror as the only reflecting surface. If the camera is an ideal perspective camera and we ignore defocus blur, it can be modeled by the point through which the perspective projection is performed; i.e. the *effective pinhole*. Then, the fixed viewpoint constraint requires that each ray of light passing through the effective pinhole of the camera (that was reflected by the mirror) would have passed through the effective viewpoint if it had not been reflected by the mirror. We now derive this constraint algebraically.

2.1 Derivation of the Fixed Viewpoint Constraint Equation

Without loss of generality we can assume that the effective viewpoint $\mathbf{v}$ of the catadioptric system lies at the origin of a Cartesian coordinate system. Suppose that the effective pinhole is located at the point $\mathbf{p}$. Then, again without loss of generality, we can assume that the z-axis $\hat{\mathbf{z}}$ lies in the direction $\vec{\mathbf{vp}}$. Moreover, since perspective projection is rotationally symmetric about any line through $\mathbf{p}$, the mirror can be assumed to be a surface of revolution about the z-axis $\hat{\mathbf{z}}$. Therefore, we work in the 2-D Cartesian frame $(\mathbf{v}, \hat{\mathbf{r}}, \hat{\mathbf{z}})$ where $\hat{\mathbf{r}}$ is a unit vector orthogonal to $\hat{\mathbf{z}}$, and try to find the 2-dimensional profile of the mirror $z(r) = z(x, y)$ where $r = \sqrt{x^2 + y^2}$. Finally, if the distance from $\mathbf{v}$ to $\mathbf{p}$ is denoted by the parameter c, we have $\hat{\mathbf{v}} = (0, 0)$ and $\hat{\mathbf{p}} = (0, c)$. See Figure 1 for an illustration[1] of the coordinate frame.

We begin the translation of the fixed viewpoint constraint into symbols by denoting

[1] In Figure 1 we have drawn the image plane as though it were orthogonal to the z-axis $\hat{\mathbf{z}}$ indicating that the optical axis of the camera is (anti) parallel to $\hat{\mathbf{z}}$. In fact, the effective viewpoint $\mathbf{v}$ and the axis of symmetry of the mirror profile $z(r)$ need not necessarily lie on the optical axis. Since perspective projection is rotationally symmetric with respect to any ray that passes through the pinhole $\mathbf{p}$, the camera could be rotated about $\mathbf{p}$ so that the optical axis is not parallel to the z-axis. Moreover, the image plane can be rotated independently so that it is no longer orthogonal to $\hat{\mathbf{z}}$.

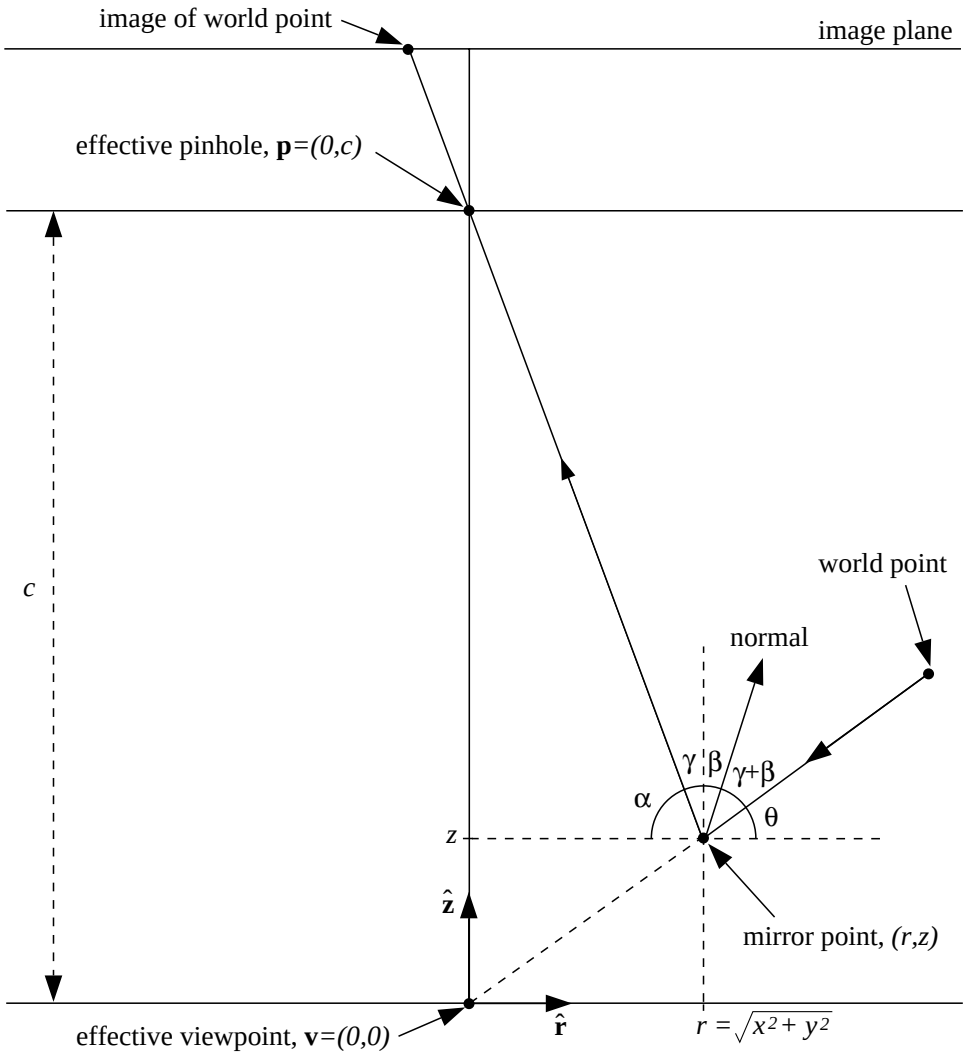


Figure 1: The geometry used to derive the fixed viewpoint constraint equation. The viewpoint $\mathbf{v} = (0, 0)$ is located at the origin of a 2-D coordinate frame $(\mathbf{v}, \hat{\mathbf{r}}, \hat{\mathbf{z}})$, and the pinhole of the camera $\mathbf{p} = (0, c)$ is located at a distance c from $\mathbf{v}$ along the z-axis $\hat{\mathbf{z}}$. If a ray of light, which was about to pass through $\mathbf{v}$, is reflected at the mirror point (r, z), the angle between the ray of light and $\hat{\mathbf{r}}$ is $\theta = \tan^{-1}\frac{z}{r}$. If the ray is then reflected and passes through the pinhole $\mathbf{p}$, the angle it makes with $\hat{\mathbf{r}}$ is $\alpha = \tan^{-1}\frac{c-z}{r}$, and the angle it makes with $\hat{\mathbf{z}}$ is $\gamma = 90° - \alpha$. Finally, if $\beta = \tan^{-1}\left(-\frac{\mathrm{d}z}{\mathrm{d}r}\right)$ is the angle between the normal to the mirror at (r, z) and $\hat{\mathbf{z}}$, then by the fact that the angle of incidence equals the angle of reflection, we have the constraint that $\alpha + \theta + 2\gamma + 2\beta = 180°$.

the angle between an incoming ray from a world point and the r-axis by θ. Suppose that this ray intersects the mirror at the point (z, r). Then, since we assume that it also passes through the origin $\mathbf{v} = (0, 0)$ we have the relationship:

$$\tan\theta \;=\; \frac{z}{r}. \tag{1}$$

If we denote the angle between the reflected ray and the (negative) r-axis by α, we also have:

$$\tan\alpha \;=\; \frac{c - z}{r} \tag{2}$$

since the reflected ray must pass through the pinhole $\mathbf{p} = (0, c)$. Next, if β is the angle between the z-axis and the normal to the mirror at the point (r, z), we have:

$$\frac{\mathrm{d}z}{\mathrm{d}r} \;=\; -\tan\beta. \tag{3}$$

Our final geometric relationship is due to the fact that we can assume the mirror to be specular. This means that the angle of incidence must equal the angle of reflection. So, if γ is the angle between the reflected ray and the z-axis, we have $\gamma = 90^{\mathrm{o}} - \alpha$ and $\theta + \alpha + 2\beta + 2\gamma = 180^{\mathrm{o}}$. (See Figure 1 for an illustration of this constraint.) Eliminating γ from these two expressions and rearranging gives:

$$2\beta \;=\; \alpha - \theta. \tag{4}$$

Then, taking the tangent of both sides and using the standard rules for expanding the tangent of a sum:

$$\tan(\mathrm{A} \pm \mathrm{B}) \;=\; \frac{\tan A \pm \tan B}{1 \mp \tan A \tan B} \tag{5}$$

we have:

$$\frac{2\tan\beta}{1 - \tan^2\beta} \;=\; \frac{\tan\alpha - \tan\theta}{1 + \tan\alpha\tan\theta}. \tag{6}$$

Substituting from Equations (1), (2), and (3) yields the *fixed viewpoint constraint* equation:

$$\frac{-2\frac{\mathrm{d}z}{\mathrm{d}r}}{1 - \left(\frac{\mathrm{d}z}{\mathrm{d}r}\right)^2} \;=\; \frac{(c - 2z)r}{r^2 + cz - z^2} \tag{7}$$

which when rearranged is seen to be a quadratic first-order ordinary differential equation:

$$r(c - 2z)\left(\frac{\mathrm{d}z}{\mathrm{d}r}\right)^2 - 2(r^2 + cz - z^2)\frac{\mathrm{d}z}{\mathrm{d}r} + r(2z - c) \;=\; 0. \tag{8}$$

2.2 General Solution of the Constraint Equation

The first step in the solution of the fixed viewpoint constraint equation is to solve it as a quadratic to yield an expression for the surface slope:

$$\frac{\mathrm{d}z}{\mathrm{d}r} = \frac{(z^2 - r^2 - cz) \pm \sqrt{r^2c^2 + (z^2 + r^2 - cz)^2}}{r(2z - c)}. \tag{9}$$

The next step is to substitute $y = z - \frac{c}{2}$ and set $b = \frac{c}{2}$ which yields:

$$\frac{\mathrm{d}y}{\mathrm{d}r} = \frac{(y^2 - r^2 - b^2) \pm \sqrt{4r^2b^2 + (y^2 + r^2 - b^2)^2}}{2ry}. \tag{10}$$

Then, we substitute $2rx = y^2 + r^2 - b^2$, which when differentiated gives:

$$2y\frac{\mathrm{d}y}{\mathrm{d}r} = 2x + 2r\frac{\mathrm{d}x}{\mathrm{d}r} - 2r \tag{11}$$

and so we have:

$$2x + 2r\frac{\mathrm{d}x}{\mathrm{d}r} - 2r = \frac{2rx - 2r^2 \pm \sqrt{4r^2b^2 + 4r^2x^2}}{r}. \tag{12}$$

Rearranging this equation yields:

$$\frac{1}{\sqrt{b^2 + x^2}}\frac{\mathrm{d}x}{\mathrm{d}r} = \pm\frac{1}{r}. \tag{13}$$

Integrating both sides with respect to r results in:

$$\ln\left(x + \sqrt{b^2 + x^2}\right) = \pm \ln r + C \tag{14}$$

where C is the constant of integration. Hence,

$$x + \sqrt{b^2 + x^2} = \frac{k}{2}r^{\pm 1} \tag{15}$$

where $k = 2e^C > 0$ is a constant. By back substituting, rearranging, and simplifying we arrive at the two equations which comprise the general solution of the fixed viewpoint constraint equation:

$$\left(z - \frac{c}{2}\right)^2 - r^2\left(\frac{k}{2} - 1\right) = \frac{c^2}{4}\left(\frac{k-2}{k}\right) \qquad (k \geq 2) \tag{16}$$

$$\left(z - \frac{c}{2}\right)^2 + r^2\left(1 + \frac{c^2}{2k}\right) = \left(\frac{2k + c^2}{4}\right) \qquad (k > 0). \tag{17}$$

In the first of these two equations, the constant parameter k is constrained by $k \geq 2$ (rather than $k > 0$) since $0 < k < 2$ leads to complex solutions.

2.3 Specific Solutions of the Constraint Equation

Together, Equations (16) and (17) define the complete class of mirrors that satisfy the fixed viewpoint constraint. A quick glance at the form of these equations reveals that the mirror profiles form a 2-parameter (c and k) family of conic sections. Hence, the shapes of the 3-D mirrors are all swept conic sections. As we shall see, however, although every conic section is theoretically a solution of one of the two equations, a number of the solutions are degenerate and cannot be used to construct real cameras with a single effective viewpoint.

We will now describe the solutions in detail. For each solution, we demonstrate whether it is degenerate or not. Some of the non-degenerate solutions have actually been used in real cameras. For these solutions, we mention all of the existing designs that we are aware of which use that mirror shape. Several of the degenerate solutions have also been used to construct cameras with a wide field of view, but with no fixed viewpoint.

2.3.1 Planar Mirrors

In Equation (16), if we set $k = 2$ and $c > 0$, we get the cross-section of a planar mirror:

$$z \;=\; \frac{c}{2}. \tag{18}$$

As shown in Figure 2, this plane is the one which bisects the line segment $\vec{\mathbf{vp}}$ joining the viewpoint and the pinhole.

The converse of this result is that for a fixed viewpoint $\mathbf{v}$ and pinhole $\mathbf{p}$, there is only one planar solution of the fixed viewpoint constraint equation. The unique solution is the perpendicular bisector of the line joining the pinhole to the viewpoint:

$$\left[\mathbf{x} - \left(\frac{\mathbf{p}+\mathbf{v}}{2}\right)\right] \cdot (\mathbf{p}-\mathbf{v}) \;=\; 0. \tag{19}$$

To prove this, it is sufficient to consider a fixed pinhole $\mathbf{p}$, a planar mirror with unit normal $\hat{\mathbf{n}}$, and a point $\mathbf{q}$ on the mirror. Then, the fact that the plane is a solution of the fixed viewpoint constraint implies that there is a single effective viewpoint $\mathbf{v} = \mathbf{v}(\hat{\mathbf{n}}, \mathbf{q})$. To be

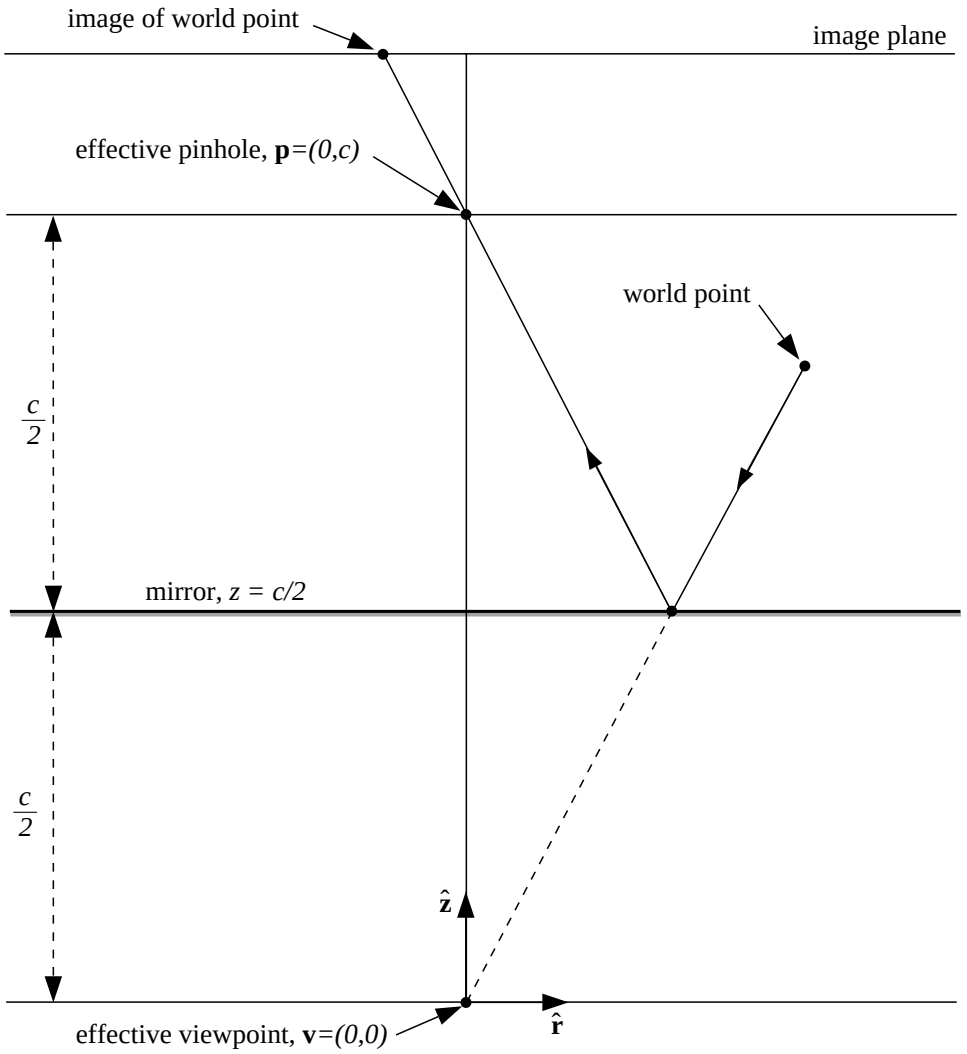


Figure 2: The plane $z = \frac{c}{2}$ is a solution of the fixed viewpoint constraint equation. Conversely, it is possible to show that, given a fixed viewpoint and pinhole, the only planar solution is the perpendicular bisector of the line joining the pinhole to the viewpoint. Hence, for a fixed pinhole, two different planar mirrors cannot share the same effective viewpoint. For each such plane the effective viewpoint is the reflection of the pinhole in the plane. This means that it is impossible to enhance the field of view using a single perspective camera and an *arbitrary number* of planar mirrors, while still respecting the fixed viewpoint constraint. If multiple cameras are used then solutions using multiple planar mirrors are possible [Nalwa, 1996].

more precise, the effective viewpoint is the reflection of the pinhole $\mathbf{p}$ in the mirror; i.e. the single effective viewpoint is:

$$\mathbf{v}(\hat{\mathbf{n}}, \mathbf{q}) \;=\; \mathbf{p} - 2\left[(\mathbf{p} - \mathbf{q}) \cdot \hat{\mathbf{n}}\right]\hat{\mathbf{n}}. \tag{20}$$

Since the reflection of a single point in two different planes is always two different points, the perpendicular bisector is the unique planar solution.

An immediate corollary of this result is that for a single fixed pinhole, no two different planar mirrors can share the same viewpoint. Unfortunately, a single planar mirror does not enhance the field of view, since, discounting occlusions, the same camera moved from $\mathbf{p}$ to $\mathbf{v}$ and reflected in the mirror would have exactly the same field of view. It follows that it is impossible to increase the field of view by packing an *arbitrary number* of planar mirrors (pointing in different directions) in front of a conventional imaging system, while still respecting the fixed viewpoint constraint. On the other hand, in applications such as stereo where multiple viewpoints are a necessary requirement, the multiple views of a scene can be captured by a single camera using multiple planar mirrors. See, for example, [Goshtasby and Gruver, 1993], [Inaba *et al.*, 1993], and [Nene and Nayar, 1998].

This brings us to the panoramic camera proposed by Nalwa [1996]. To ensure a single viewpoint while using multiple planar mirrors, Nalwa [1996] arrived at a design that uses four separate imaging systems. Four planar mirrors are arranged in a square-based pyramid, and each of the four cameras is placed above one of the faces of the pyramid. The effective pinholes of the cameras are moved until the four effective viewpoints (i.e. the reflections of the pinholes in the mirrors) coincide. The result is a camera that has a single effective viewpoint and a panoramic field of view of approximately $360^\circ \times 50^\circ$. The panoramic image is of relatively high resolution since it is generated from the four images captured by the four cameras. This camera is straightforward to implement, but requires four of each component: i.e. four cameras, four lenses, and four digitizers. (It is, of course, possible to use only one digitizer but at a reduced frame rate.)

2.3.2 Conical Mirrors

In Equation (16), if we set $c = 0$ and $k \geq 2$, we get a conical mirror with circular cross section:

$$z = \sqrt{\frac{k-2}{2}r^2}. \tag{21}$$

See Figure 3 for an illustration of this solution. The angle at the apex of the cone is 2τ where:

$$\tan\tau = \sqrt{\frac{2}{k-2}}. \tag{22}$$

This might seem like a reasonable solution, but since $c = 0$ the pinhole of the camera must be at the apex of the cone. This implies that the only rays of light entering the pinhole from the mirror are the ones which graze the cone and so do not originate from objects in the world (see Figure 3.) Hence, the cone with the pinhole at the vertex is a degenerate solution that cannot be used to construct a wide field of view camera with a single effective viewpoint.

In spite of this fact, the cone has been used in wide-angle imaging systems several times [Yagi and Kawato, 1990] [Yagi and Yachida, 1991] [Bogner, 1995]. In these implementations the pinhole is placed some distance from the apex of the cone. It is easy to show that in such cases the viewpoint is no longer a single point [Nalwa, 1996] [Baker and Nayar, 1999]. In some applications such as robot navigation, the single viewpoint constraint is not vital. Conical mirrors can be used to build practical cameras for such applications. See, for example, the designs in [Yagi *et al.*, 1994] and [Bogner, 1995].

2.3.3 Spherical Mirrors

In Equation (17), if we set $c = 0$ and $k > 0$, we get the spherical mirror:

$$z^2 + r^2 = \frac{k}{2}. \tag{23}$$

Like the cone, this is a degenerate solution which cannot be used to construct a wide field of view camera with a single viewpoint. Since the viewpoint and pinhole coincide at the center

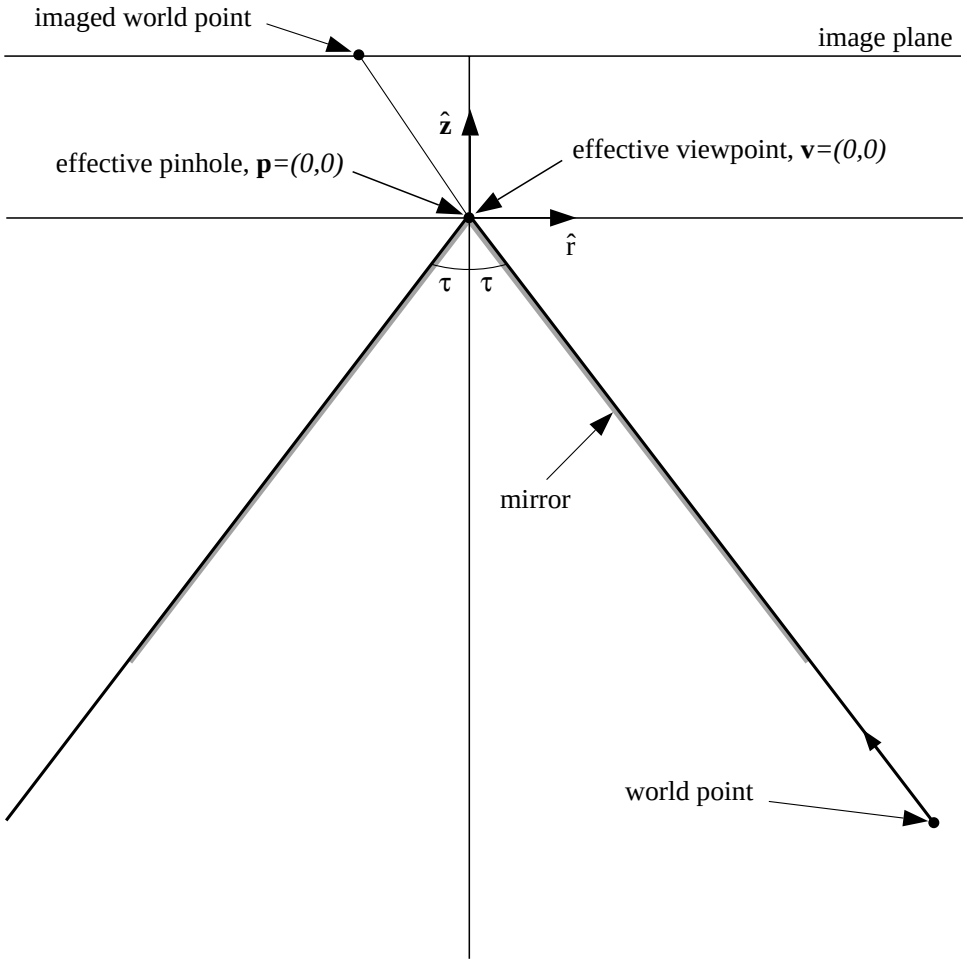


Figure 3: The conical mirror is a solution of the fixed viewpoint constraint equation. Since the pinhole is located at the apex of the cone, this is a degenerate solution that cannot be used to construct a wide field of view camera with a single viewpoint. If the pinhole is moved away from the apex of the cone (along the axis of the cone), the viewpoint is no longer a single point but rather lies on a circular locus [Nalwa, 1996] [Baker and Nayar, 1999].

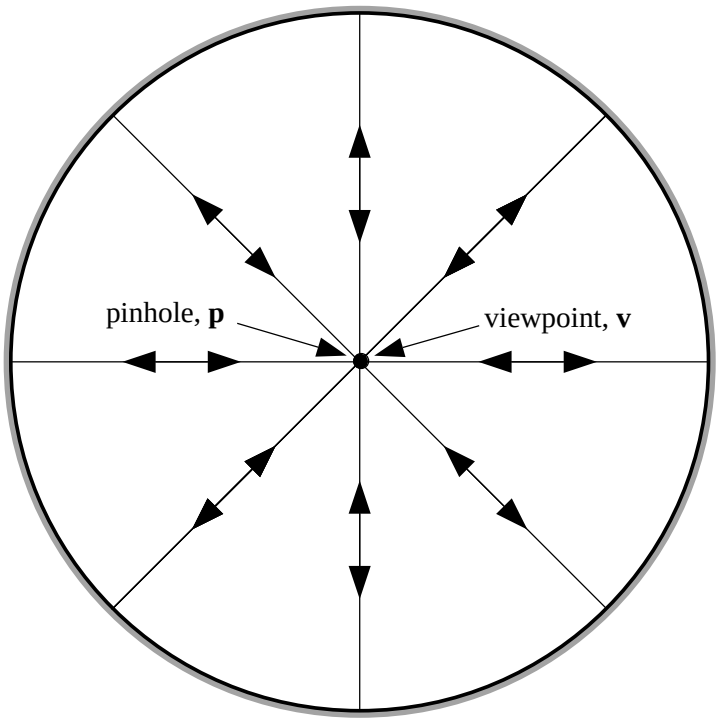


Figure 4: The spherical mirror satisfies the fixed viewpoint constraint when the pinhole lies at the center of the sphere. (Since $c = 0$ the viewpoint also lies at the center of the sphere.) Like the conical mirror, the sphere cannot actually be used to construct a wide field of view camera with a single viewpoint because the observer can only see itself; rays of light emitted from the center of the sphere are reflected back at the surface of the sphere directly towards the center of the sphere.

of the sphere, the observer would see itself and nothing else, as is illustrated in Figure 4.

The sphere has also been used to build wide field of view cameras several times [Hong, 1991] [Bogner, 1995] [Murphy, 1995]. In these implementations, the pinhole is placed outside the sphere and so there is no single effective viewpoint. The locus of the effective viewpoint can be computed in a straightforward manner using a symbolic mathematics package [Baker and Nayar, 1999]. Like multiple planes, spheres have also been used to construct stereo rigs [Nayar, 1988] [Nene and Nayar, 1998], but as described before, multiple viewpoints are a requirement for stereo.

2.3.4 Ellipsoidal Mirrors

In Equation (17), when $k > 0$ and $c > 0$, we get the ellipsoidal mirror:

$$\frac{1}{a_e^2}\left(z - \frac{c}{2}\right)^2 + \frac{1}{b_e^2}r^2 \;=\; 1 \tag{24}$$

where:

$$a_e = \sqrt{\frac{2k + c^2}{4}} \quad \text{and} \quad b_e = \sqrt{\frac{k}{2}}. \tag{25}$$

The ellipsoid is the first solution that can actually be used to enhance the field of view of a camera while retaining a single effective viewpoint. As shown in Figure 5, if the viewpoint and pinhole are at the foci of the ellipsoid and the mirror is taken to be the section of the ellipsoid that lies below the viewpoint (i.e. $z < 0$), the effective field of view is the entire upper hemisphere $z \geq 0$.

2.3.5 Hyperboloidal Mirrors

In Equation (16), when $k > 2$ and $c > 0$, we get the hyperboloidal mirror:

$$\frac{1}{a_h^2}\left(z - \frac{c}{2}\right)^2 - \frac{1}{b_h^2}r^2 \;=\; 1 \tag{26}$$

where:

$$a_h = \frac{c}{2}\sqrt{\frac{k-2}{k}} \quad \text{and} \quad b_h = \frac{c}{2}\sqrt{\frac{2}{k}}. \tag{27}$$

As seen in Figure 6, the hyperboloid also yields a realizable solution. The curvature of the mirror and the field of view both increase with k. In the other direction (in the limit $k \to 2$) the hyperboloid flattens out to the planar mirror of Section 2.3.1.

Rees [1970] appears to have been first to use a hyperboloidal mirror with a perspective lens to achieve a large field of view camera system with a single viewpoint. Later, Yamazawa *et al.* [1993] [1995] also recognized that the hyperboloid is indeed a practical solution and implemented a camera designed for autonomous navigation.

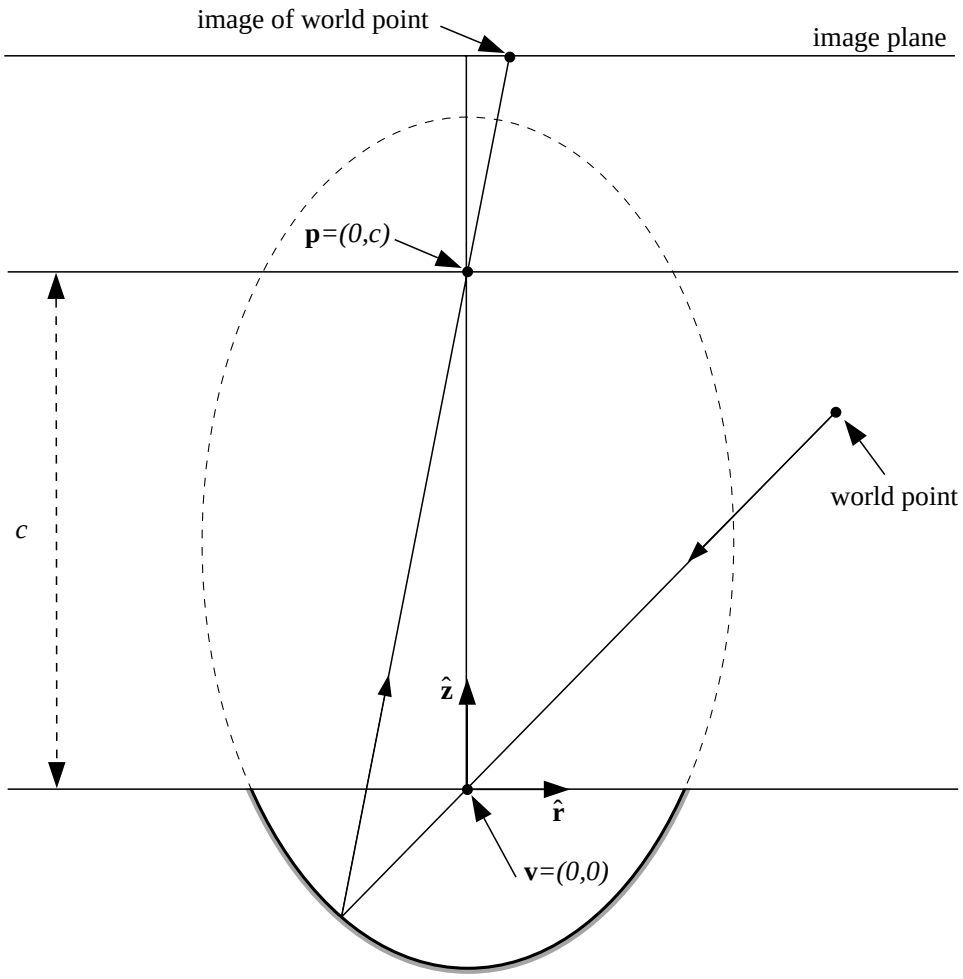


Figure 5: The ellipsoidal mirror satisfies the fixed viewpoint constraint when the pinhole and viewpoint are located at the two foci of the ellipsoid. If the ellipsoid is terminated by the horizontal plane passing through the viewpoint $z = 0$, the field of view is the entire upper hemisphere $z > 0$. It is also possible to cut the ellipsoid with other planes passing through $\mathbf{v}$.

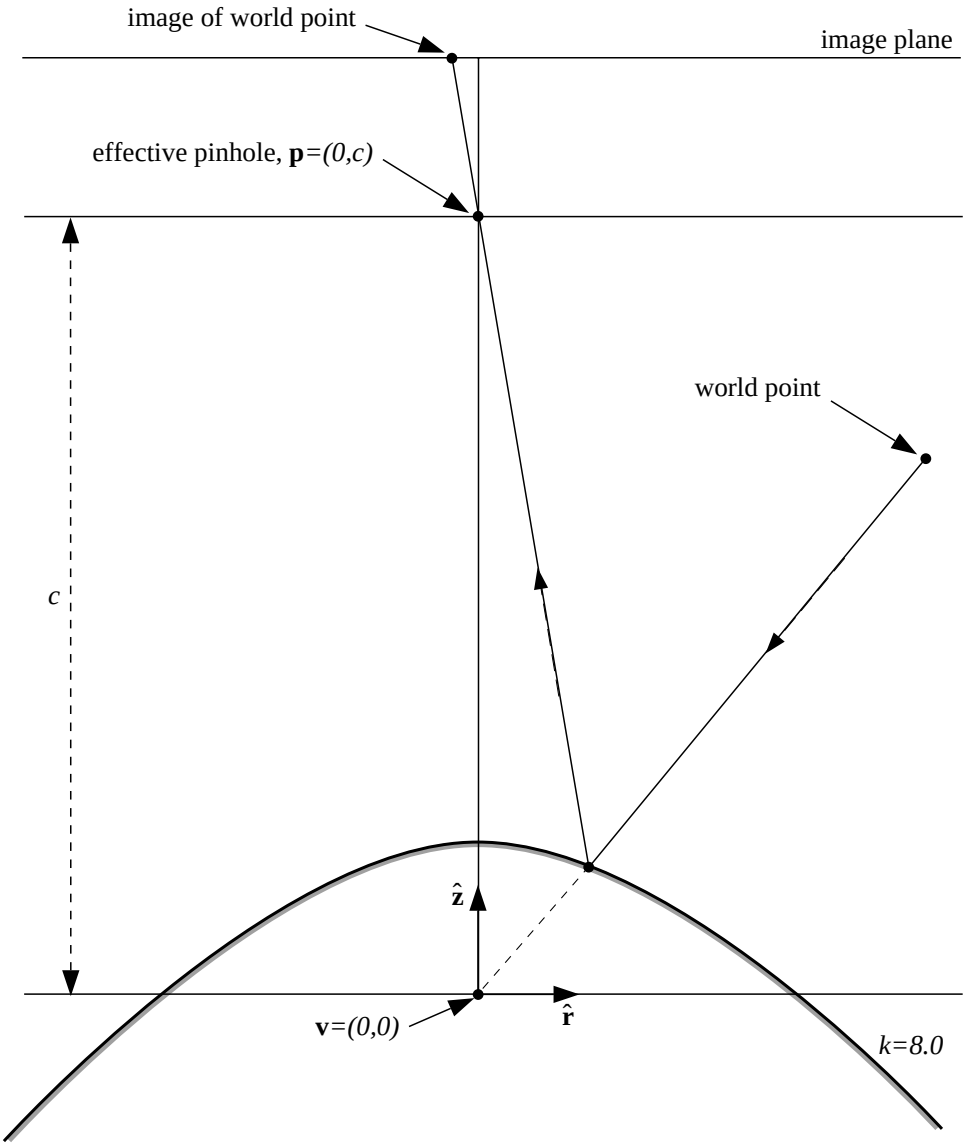


Figure 6: The hyperboloidal mirror satisfies the fixed viewpoint constraint when the pinhole and the viewpoint are located at the two foci of the hyperboloid. This solution does produce the desired increase in field of view. The curvature of the mirror and hence the field of view increase with k. In the limit $k \to 2$, the hyperboloid flattens to the planar mirror of Section 2.3.1.

2.4 The Orthographic Case: Paraboloidal Mirrors

There is one conic section that we have not mentioned: the parabola. Although the parabola is not a solution of either equation for finite values of c and k, it is a solution of Equation (16) in the limit that $c \to \infty$, $k \to \infty$, and $\frac{c}{k} = h$, a constant. These limiting conditions correspond to orthographic projection.

Orthographic projection can be modeled by setting $\alpha = 90°$ in Figure 1; the direction of projection is then along the axis of symmetry $\hat{\mathbf{z}}$. Equation (6) then simplifies to:

$$\frac{2\tan\beta}{1-\tan^2\beta} = \frac{1}{\tan\theta}. \tag{28}$$

The *fixed viewpoint constraint* equation for orthographic projection is therefore:

$$\frac{-2\frac{dz}{dr}}{1-(\frac{dz}{dr})^2} = \frac{r}{z}\ . \tag{29}$$

As above, the first step in determining the shape of the mirror is to solve this quadratic equation for the surface slope:

$$\frac{dz}{dr} = \frac{z}{r} \mp \sqrt{1+(\frac{r}{z})^2}\ . \tag{30}$$

This first-order differential equation can be solved using similar transformations to those used above to obtain the following expression for the reflecting surface:

$$z = \pm\ \frac{h^2 - r^2}{2h}\ , \tag{31}$$

where, $h > 0$ is the constant of integration.

Not surprisingly, the mirror that guarantees a single viewpoint for orthographic projection is a paraboloid. Paraboloidal mirrors are frequently used to converge an incoming set of parallel rays at a single point (the focus), or to generate a collimated light source from a point source (placed at the focus). In both these cases, the paraboloid is a concave mirror that is reflective on its inner surface. This corresponds to the negative solution in Equation (31). In our case, the more natural solution to use is the positive one. Here, the paraboloid is reflective on its outer surface (a convex mirror) as is shown in Figure 7.

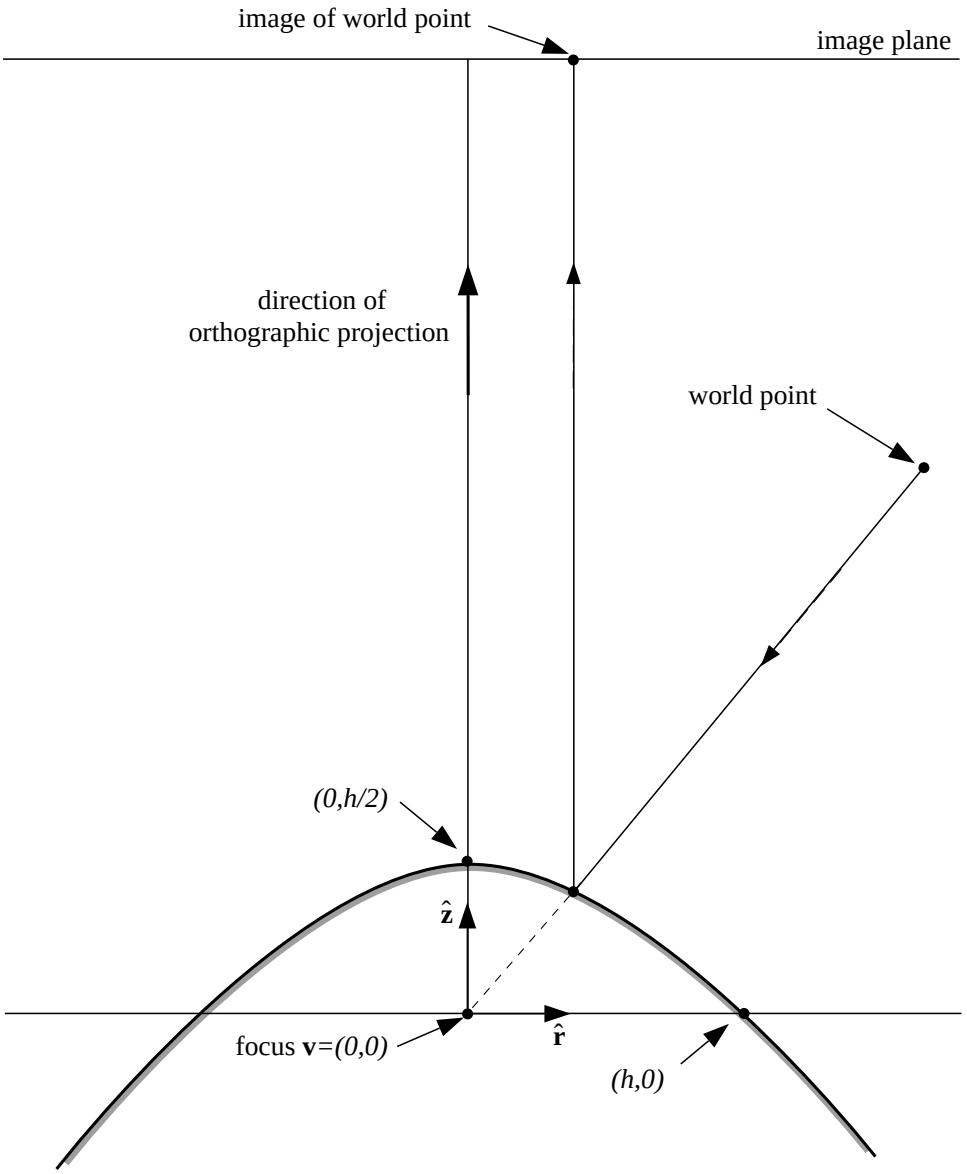


Figure 7: Under orthographic projection the only solution is a paraboloid with the effective viewpoint at the focus. Here we show one of the two possible paraboloids; the other is similar to the ellipse shown in Figure 5 and is the reflection of the one shown in the plane $z = 0$.

3 Resolution of a Catadioptric Camera

In this section, we assume that the conventional camera used in the catadioptric camera has a frontal image plane located at a distance u from the pinhole, and that the optical axis of the camera is aligned with the axis of symmetry of the mirror. See Figure 8 for an illustration of this scenario. Then, the definition of resolution that we will use is the following. Consider an infinitesimal area $\mathrm{d}A$ on the image plane. If this infinitesimal pixel images an infinitesimal solid angle $\mathrm{d}\nu$ of the world, the *resolution* of the camera as a function of the point on the image plane at the center of the infinitesimal area $\mathrm{d}A$ is:

$$\frac{\mathrm{d}A}{\mathrm{d}\nu}. \tag{32}$$

If ψ is the angle made between the optical axis and the line joining the pinhole to the center of the infinitesimal area $\mathrm{d}A$ (see Figure 8), the solid angle subtended by the infinitesimal area $\mathrm{d}A$ at the pinhole is:

$$\mathrm{d}\omega \;=\; \frac{\mathrm{d}A \cdot \cos\psi}{u^2/\cos^2\psi} \;=\; \frac{\mathrm{d}A \cdot \cos^3\psi}{u^2}. \tag{33}$$

Therefore, the resolution of the conventional camera is:

$$\frac{\mathrm{d}A}{\mathrm{d}\omega} \;=\; \frac{u^2}{\cos^3\psi}. \tag{34}$$

Then, the area of the mirror imaged by the infinitesimal area $\mathrm{d}A$ is:

$$\mathrm{d}S \;=\; \frac{\mathrm{d}\omega \cdot (c-z)^2}{\cos\phi\cos^2\psi} \;=\; \frac{\mathrm{d}A \cdot (c-z)^2 \cdot \cos\psi}{u^2\cos\phi} \tag{35}$$

where ϕ is the angle between the normal to the mirror at (r,z) and the line joining the pinhole to the mirror point (r,z). Since reflection at the mirror is specular, the solid angle of the world imaged by the catadioptric camera is:

$$\mathrm{d}\nu \;=\; \frac{\mathrm{d}S \cdot \cos\phi}{r^2+z^2} \;=\; \frac{\mathrm{d}A \cdot (c-z)^2 \cdot \cos\psi}{u^2(r^2+z^2)}. \tag{36}$$

Therefore, the resolution of the catadioptric camera is:

$$\frac{\mathrm{d}A}{\mathrm{d}\nu} \;=\; \frac{u^2(r^2+z^2)}{(c-z)^2 \cdot \cos\psi} \;=\; \left[\frac{(r^2+z^2)\cos^2\psi}{(c-z)^2}\right]\frac{\mathrm{d}A}{\mathrm{d}\omega}. \tag{37}$$

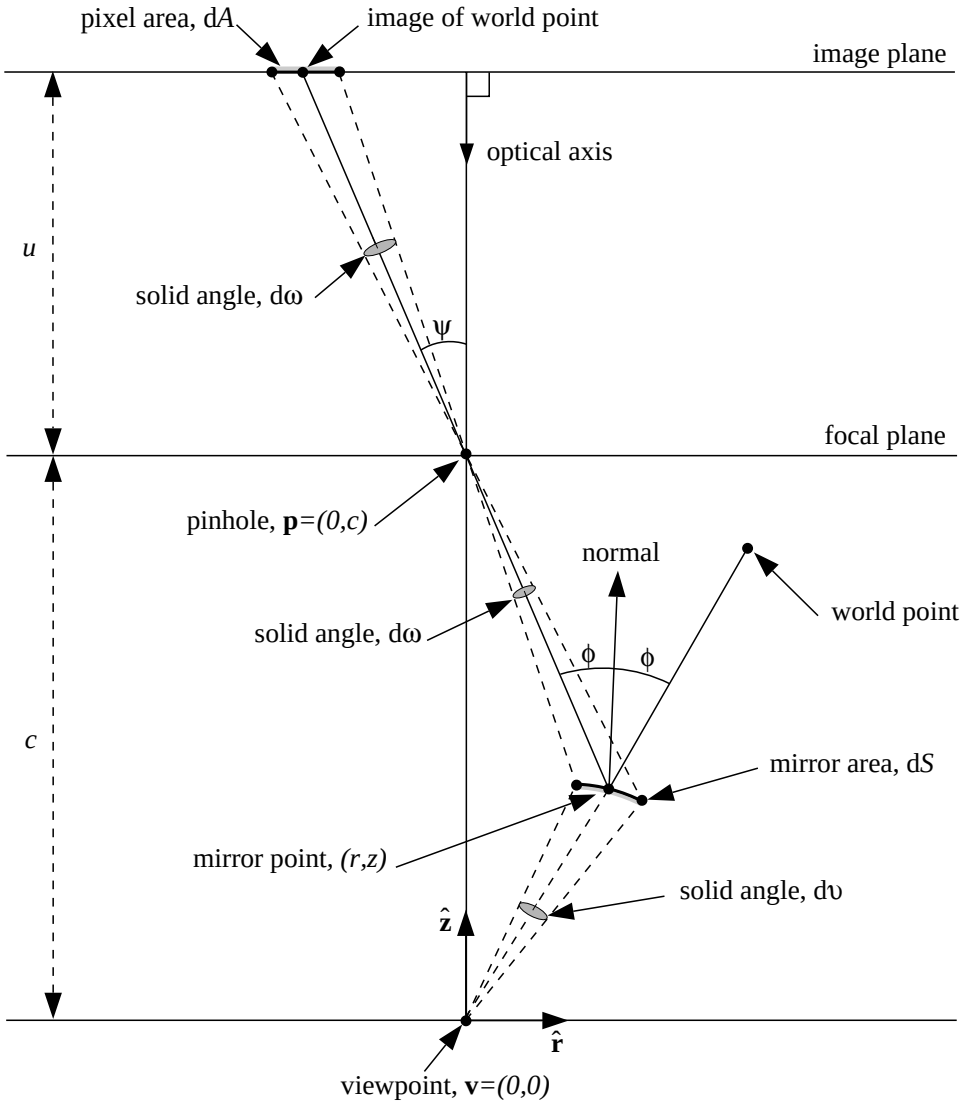


Figure 8: The geometry used to derive the spatial resolution of a catadioptric camera. Assuming the conventional camera has a frontal image plane which is located at a distance u from the pinhole and the optical axis is aligned with the z-axis $\hat{\mathbf{z}}$, the spatial resolution of the conventional camera is $\frac{\mathrm{d}A}{\mathrm{d}\omega} = \frac{u^2}{\cos^3\psi}$. Therefore the area of the mirror imaged by the infinitesimal image plane area $\mathrm{d}A$ is $\mathrm{d}S = \frac{(c-z)^2 \cdot \cos\psi}{u^2 \cos\phi} \cdot \mathrm{d}A$. So, the solid angle of the world imaged by the infinitesimal area $\mathrm{d}A$ on the image plane is $\mathrm{d}\nu = \frac{(c-z)^2 \cdot \cos\psi}{u^2(r^2+z^2)} \cdot \mathrm{d}A$. Hence, the spatial resolution of the catadioptric camera is $\frac{\mathrm{d}A}{\mathrm{d}\nu} = \frac{u^2(r^2+z^2)}{(c-z)^2 \cdot \cos\psi} = \frac{r^2+z^2}{r^2+(c-z)^2} \cdot \frac{\mathrm{d}A}{\mathrm{d}\omega}$ since $\cos^2\psi = \frac{(c-z)^2}{(c-z)^2+r^2}$.

But, since:

$$\cos^2 \psi \;=\; \frac{(c-z)^2}{(c-z)^2 + r^2} \tag{38}$$

we have:

$$\frac{\mathrm{d}A}{\mathrm{d}\nu} \;=\; \left[\frac{r^2 + z^2}{(c-z)^2 + r^2}\right] \frac{\mathrm{d}A}{\mathrm{d}\omega}. \tag{39}$$

Hence, the resolution of the catadioptric camera is the resolution of the conventional camera used to construct it multiplied by a factor of:

$$\frac{r^2 + z^2}{(c-z)^2 + r^2} \tag{40}$$

where (r, z) is the point on the mirror being imaged.

The first thing to note from Equation (39) is that for the planar mirror $z = \frac{c}{2}$, the resolution of the catadioptric camera is the same as that of the conventional camera used to construct it. This is as expected by symmetry. Secondly, note that the factor in Equation (40) is the square of the distance from the point (r, z) to the effective viewpoint $\mathbf{v} = (0, 0)$, divided by the square of the distance to the pinhole $\mathbf{p} = (0, c)$. Let $d_{\mathbf{v}}$ denote the distance from the viewpoint to (r, z) and $d_{\mathbf{p}}$ the distance of (r, z) from the pinhole. Then, the factor in Equation (40) is $\frac{d_{\mathbf{v}}^2}{d_{\mathbf{p}}^2}$. For the ellipsoid, $d_{\mathbf{p}} + d_{\mathbf{v}} = K_e$ for some constant $K_e > d_{\mathbf{p}}$. Therefore, for the ellipsoid the factor is:

$$\left(\frac{K_e}{d_{\mathbf{p}}} - 1\right)^2 \tag{41}$$

which increases as $d_{\mathbf{p}}$ decreases and $d_{\mathbf{v}}$ increases. For the hyperboloid, $d_{\mathbf{p}} - d_{\mathbf{v}} = K_h$ for some constant $0 < K_h < d_{\mathbf{p}}$. Therefore, for the hyperboloid the factor is:

$$\left(1 - \frac{K_h}{d_{\mathbf{p}}}\right)^2 \tag{42}$$

which increases as $d_{\mathbf{p}}$ increases and $d_{\mathbf{v}}$ increases. So, for both ellipsoids and hyperboloids, the factor in Equation (40) increases with r. Hence, both hyperboloidal and ellipsoidal catadioptric cameras constructed with a uniform resolution camera will have their highest resolution around the periphery, a useful property for applications such as teleconferencing.

3.1 The Orthographic Case

The orthographic case is slightly simpler than the projective case and is illustrated in Figure 9. Again, we assume that the image plane is frontal; i.e. perpendicular to the direction of orthographic projection. Then, the resolution of the conventional orthographic camera is:

$$\frac{\mathrm{d}A}{\mathrm{d}\omega} = M^2 \tag{43}$$

where the constant M is the linear magnification of the camera. If the solid angle $\mathrm{d}\omega$ images the area $\mathrm{d}S$ of the mirror and ϕ is the angle between the mirror normal and the direction of orthographic projection, we have:

$$\mathrm{d}\omega = \cos\phi \cdot \mathrm{d}S. \tag{44}$$

Combining Equations (36), (43), and (44) yields:

$$\frac{\mathrm{d}A}{\mathrm{d}\nu} = \left[r^2 + z^2\right] \frac{\mathrm{d}A}{\mathrm{d}\omega}. \tag{45}$$

For the paraboloid $z = \frac{h^2 - r^2}{2h}$, the multiplicative factor $r^2 + z^2$ simplifies to:

$$\left[\frac{h^2 + r^2}{2h}\right]^2. \tag{46}$$

Hence, as for both the ellipsoid and the hyperboloid, the resolution of paraboloid based catadioptric cameras increases with r, the distance from the center of the mirror.

4 Defocus Blur of a Catadioptric Camera

In addition to the normal causes present in conventional dioptric systems, such as diffraction and lens aberrations, two factors combine to cause defocus blur in catadioptric sensors. They are: (1) the finite size of the lens aperture, and (2) the curvature of the mirror. To analyze how these two factors cause defocus blur, we first consider a fixed point in the world and a fixed point in the lens aperture. We then find the point on the mirror which reflects a ray of

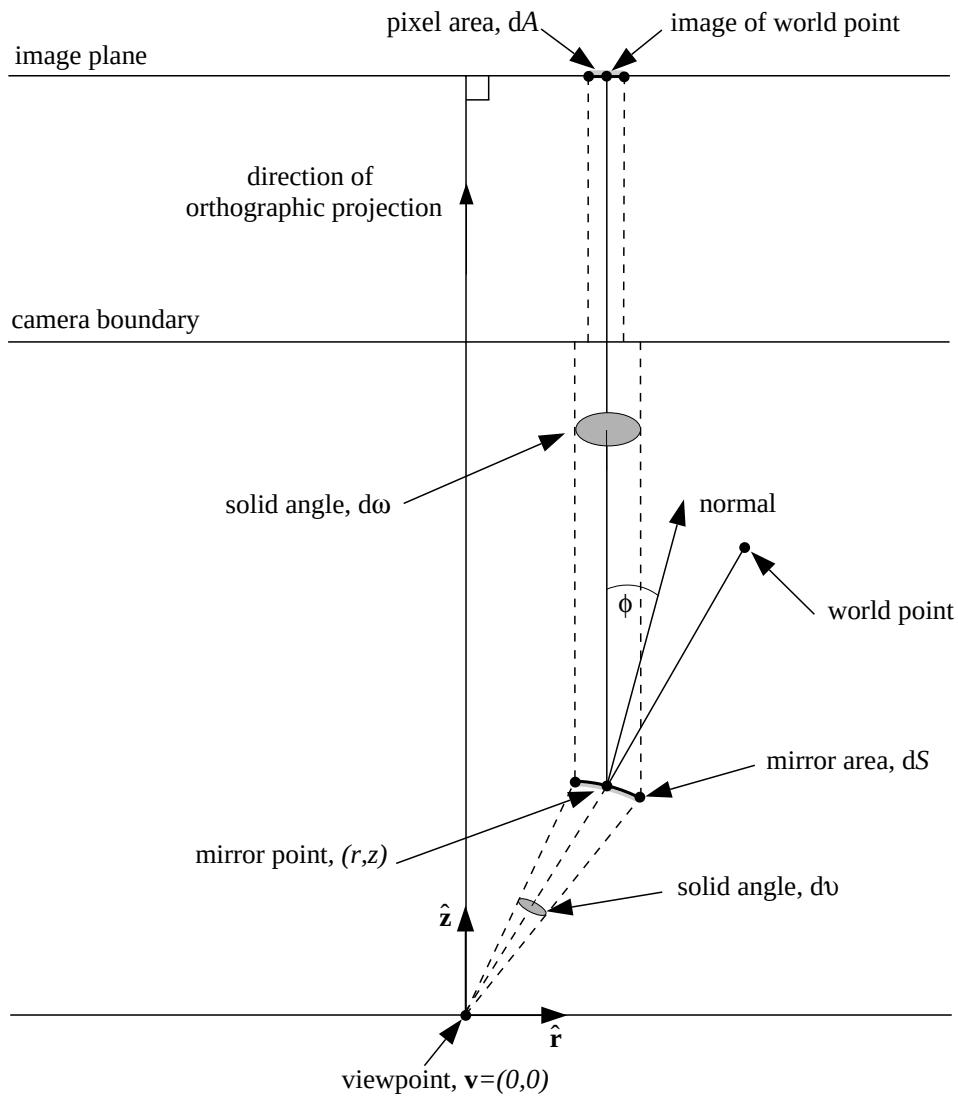


Figure 9: The geometry used to derive the spatial resolution of a catadioptric camera in the orthographic case. Again, assuming that the image plane is frontal and the conventional orthographic camera has a linear magnification M, its spatial resolution is $\frac{\mathrm{d}A}{\mathrm{d}\omega} = M^2$. The solid angle $\mathrm{d}\omega$ equals $\cos\phi \cdot \mathrm{d}S$, where $\mathrm{d}S$ is the area of the mirror imaged and ϕ is the angle between the mirror normal and the direction of orthographic projection. Combining this information with Equation (36) yields the spatial resolution of the orthographic catadioptric camera as $\frac{\mathrm{d}A}{\mathrm{d}\nu} = [r^2 + z^2]\,\frac{\mathrm{d}A}{\mathrm{d}\omega}$.

light from the world point through that lens point. Next, we compute where on the image plane this mirror point is imaged. By considering the locus of imaged mirror points as the lens point varies, we can compute the area of the image plane onto which a fixed world point is imaged. In Section 4.1, we derive the constraints on the mirror point at which the light is reflected, and show how it can be projected onto the image plane. In Section 4.2, we extend the analysis to the orthographic case. Finally, in Section 4.3, we present numerical results for hyperboloid.

4.1 Analysis of Defocus Blur

To analyze defocus blur, we need to work in 3-D. We use the 3-D cartesian frame $(\mathbf{v}, \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ where $\mathbf{v}$ is the location of the effective viewpoint, $\mathbf{p}$ is the location of the effective pinhole, $\hat{\mathbf{z}}$ is a unit vector in the direction $\vec{\mathbf{vp}}$, the effective pinhole is located at a distance c from the effective viewpoint, and the vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal unit vectors in the plane $z = 0$. As in Section 3, we also assume that the conventional camera used in the catadioptric sensor has a frontal image plane located at a distance u from the pinhole and that the optical axis of the camera is aligned with the z-axis. In addition to the previous assumptions, we assume that the effective pinhole of the lens is located at the center of the lens, and that the lens has a circular aperture. See Figure 10 for an illustration of this configuration.

Consider a point $\mathbf{m} = (x, y, z)$ on the mirror and a point $\mathbf{w} = \frac{l}{\|\mathbf{m}\|}(x, y, z)$ in the world, where $l > \|\mathbf{m}\|$. Then, since the hyperboloid mirror satisfies the fixed viewpoint constraint, a ray of light from $\mathbf{w}$ which is reflected by the mirror at $\mathbf{m}$ passes directly through the center of the lens (i.e. the effective pinhole.) This ray of light is known as the *principal ray* [Hecht and Zajac, 1974]. Next, suppose a ray of light from the world point $\mathbf{w}$ is reflected at the point $\mathbf{m}_1 = (x_1, y_1, z_1)$ on the mirror and then passes through the lens aperture point $\mathbf{l} = (d \cdot \cos\lambda, d \cdot \sin\lambda, c)$. In general, this ray of light will not be imaged at the same point on the image plane as the principal ray. When this happens there is defocus blur. The locus of the intersection of the incoming rays through $\mathbf{l}$ and the image plane as $\mathbf{l}$ varies over the lens

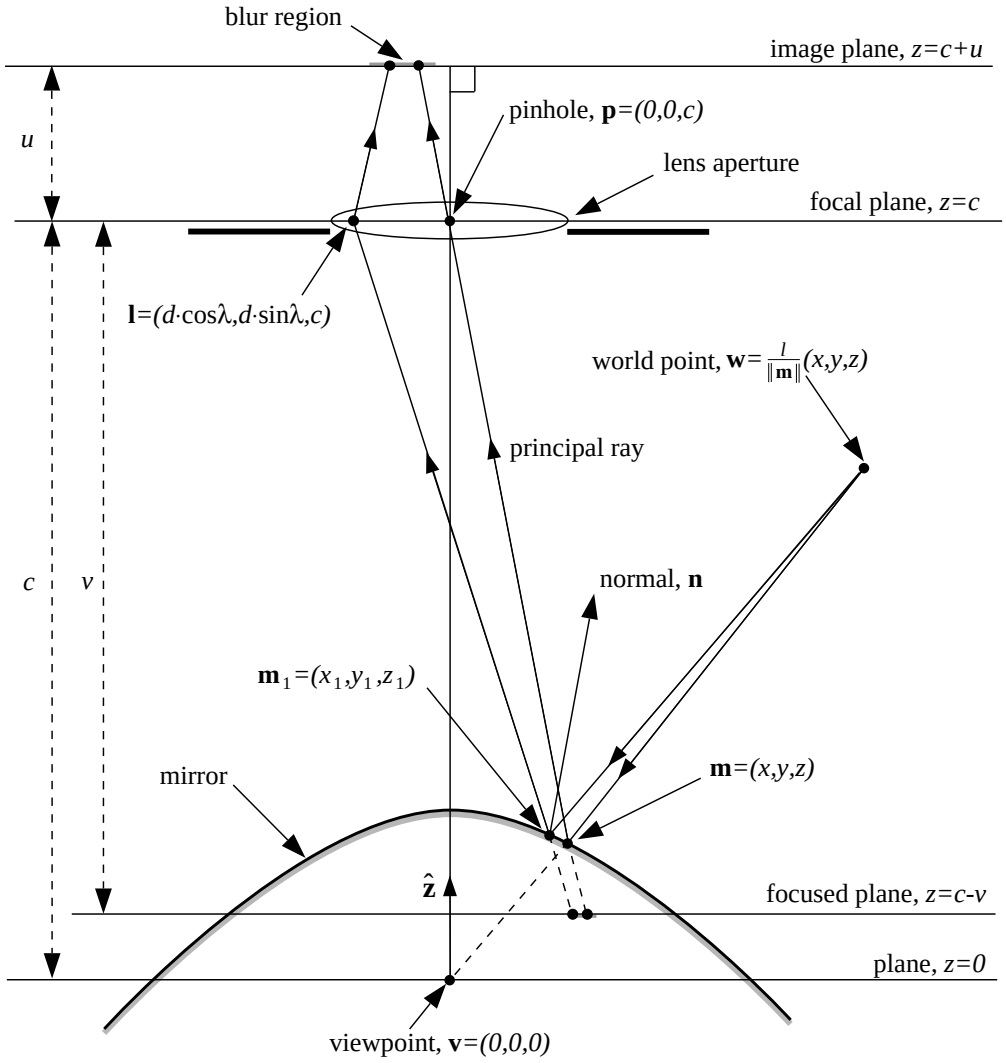


Figure 10: The geometry used to analyze the defocus blur. We work in the 3-D cartesian frame $(\mathbf{v}, \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal unit vectors in the plane $z = 0$. In addition to the assumptions of Section 3, we also assume that the effective pinhole is located at the center of the lens and that the lens has a circular aperture. If a ray of light from the world point $\mathbf{w} = \frac{l}{\|\mathbf{m}\|}(x, y, z)$ is reflected at the mirror point $\mathbf{m}_1 = (x_1, y_1, z_1)$ and then passes through the lens point $\mathbf{l} = (d \cdot \cos\lambda, d \cdot \sin\lambda, c)$, there are three constraints on $\mathbf{m}_1$: (1) it must lie on the mirror, (2) the angle of incidence must equal the angle of reflection, and (3) the normal $\mathbf{n}$ to the mirror at $\mathbf{m}_1$, and the two vectors $\mathbf{l} - \mathbf{m}_1$ and $\mathbf{w} - \mathbf{m}_1$ must be coplanar.

aperture is known as the *blur region* or *region of confusion* [Hecht and Zajac, 1974]. For an ideal thin lens in isolation, the blur region is circular and so is often referred to as the *blur circle* [Hecht and Zajac, 1974].

If we know the points $\mathbf{m}_1$ and $\mathbf{l}$, we can find the point on the image plane where the ray of light through these points is imaged. First, the line through $\mathbf{m}_1$ in the direction $\vec{\mathbf{lm}}_1$ is extended to intersect the *focused plane.* By the thin lens law [Hecht and Zajac, 1974] the focused plane is:

$$z \;=\; c - v \;=\; c - \frac{f \cdot u}{u - f} \tag{47}$$

where f is the focal length of the lens and u is the distance from the focal plane to the image plane. Since all points on the focused plane are perfectly focused, the point of intersection on the focused plane can be mapped onto the image plane using perspective projection. Hence, the x and y coordinates of the intersection of the ray through $\mathbf{l}$ and the image plane are the x and y coordinates of:

$$-\frac{u}{v}\left(\mathbf{l} + \frac{v}{c - z_1}(\mathbf{m}_1 - \mathbf{l})\right) \tag{48}$$

and the z coordinate is the z coordinate of the image plane $c + u$.

Given the lens point $\mathbf{l} = (d \cdot \cos\lambda, d \cdot \sin\lambda, c)$ and the world point $\mathbf{w} = \frac{l}{\|\mathbf{m}\|}(x, y, z)$, there are three constraints on the point $\mathbf{m}_1 = (x_1, y_1, z_1)$. First, $\mathbf{m}_1$ must lie on the mirror and so (for the hyperboloid) we have:

$$\left(z_1 - \frac{c}{2}\right)^2 - \left(x_1^2 + y_1^2\right)\left(\frac{k}{2} - 1\right) \;=\; \frac{c^2}{4}\left(\frac{k-2}{k}\right). \tag{49}$$

Secondly, the incident ray $(\mathbf{w} - \mathbf{m}_1)$, the reflected ray $(\mathbf{m}_1 - \mathbf{l})$, and the normal to the mirror at $\mathbf{m}_1$ must lie in the same plane. The normal to the mirror at $\mathbf{m}_1$ lies in the direction:

$$\mathbf{n} \;=\; ([k-2]x_1, [k-2]y_1, c - 2z_1) \tag{50}$$

for the hyperboloid. Hence, the second constraint is:

$$\mathbf{n} \cdot (\mathbf{w} - \mathbf{m}_1) \wedge (\mathbf{l} - \mathbf{m}_1) \;=\; 0. \tag{51}$$

Finally, the angle of incidence must equal the angle of reflection and so the third constraint on the point $\mathbf{m}_1$ is:

$$\frac{\mathbf{n} \cdot (\mathbf{w} - \mathbf{m}_1)}{\|\mathbf{w} - \mathbf{m}_1\|} = \frac{\mathbf{n} \cdot (\mathbf{l} - \mathbf{m}_1)}{\|\mathbf{l} - \mathbf{m}_1\|}. \tag{52}$$

These three constraints on $\mathbf{m}_1$ are all multivariate polynomials in x_1, y_1, and z_1: Equation (49) and Equation (51) are both of order 2, and Equation (52) is of order 5. We were unable to find a closed form solution to these three equations (Equation (52) has 25 terms in general and so it is probable that none exists) but we did investigate numerical solutions. Before we present the results, we briefly describe the orthographic case.

4.2 Defocus Blur in the Orthographic Case

The orthographic case is slightly different, as is illustrated in Figure 11. One way to convert a thin lens to produce orthographic projection is to place an aperture at the focal point behind the lens [Nayar, 1997b]. Then, the only rays of light that reach the image plane are those that are (approximately) parallel to the optical axis. For the orthographic case, there is therefore only one difference to the analysis. When estimating the blur region, we need to check that the ray of light actually passes through the (circular) aperture at the rear focal point. This task is straightforward. The intersection of the ray of light with the rear focal plane is computed using linear interpolation of the lens point and the point where the mirror point is imaged on the image plane. It is then checked whether this point lies close enough to the optical axis to pass through the aperture.

4.3 Numerical Results

It is possible to use the constraints derived in the previous two sections to investigate how the shape and size of the blur areas varies with the focal setting. For lack of space, however, we are unable to present these results. The reader is referred to [Baker and Nayar, 1999] for the full details. Instead, we investigate how the focus setting that minimizes the area of

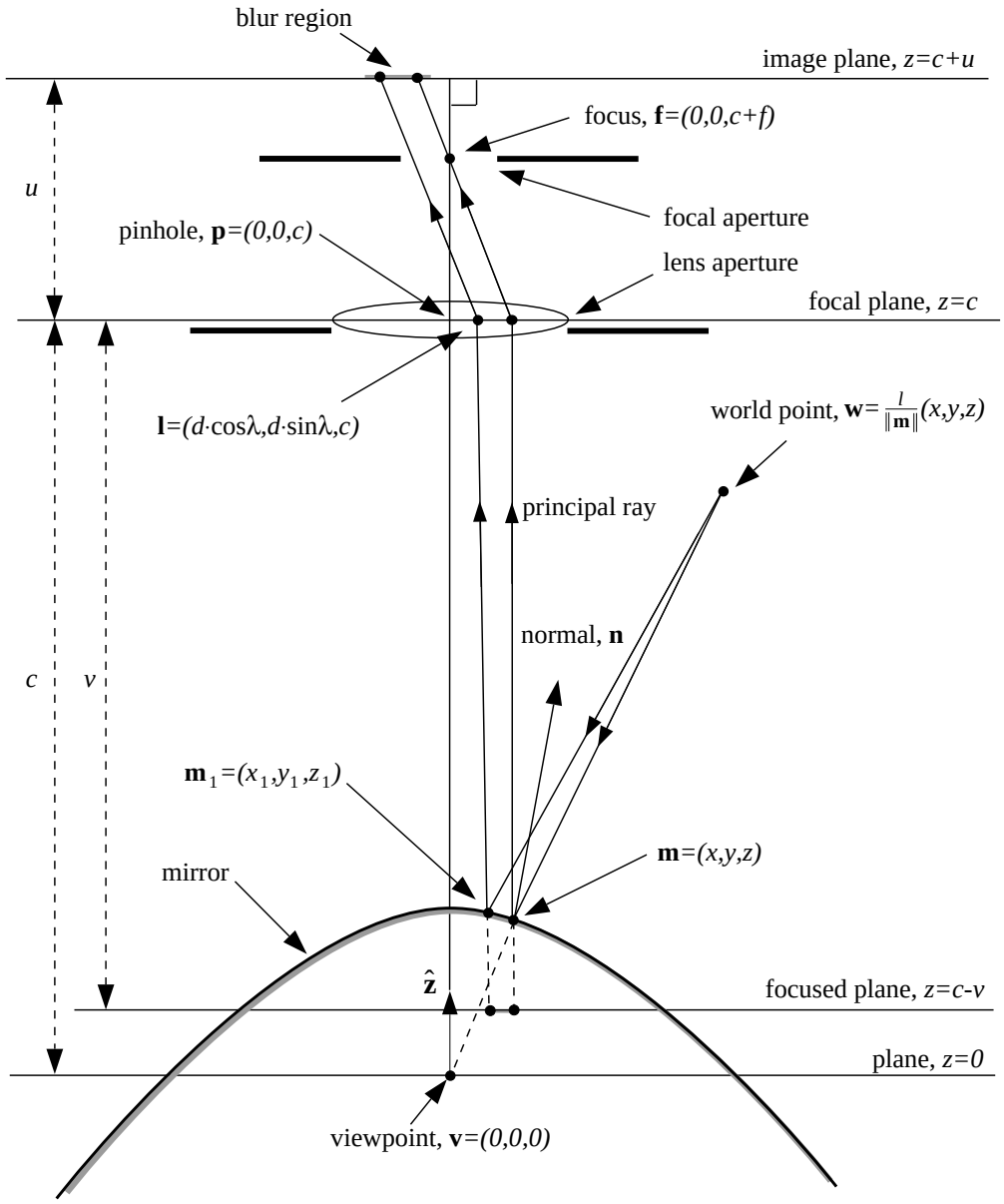


Figure 11: The geometry used to analyze defocus blur in the orthographic case. One way to create orthographic projection is to add a (circular) aperture at the rear focal point (the one behind the lens) [Nayar, 1997b]. Then, the only rays of light that reach the image plane are those which are (approximately) parallel to the optical axis. The analysis of defocus blur is then essentially the same as in the perspective case except that we need to check whether each ray of light passes through this aperture when computing the blur region.

the blur region for points a fixed distance away in the world varies with the angle which the world point $\mathbf{w}$ makes with the plane $z = 0$. The results are presented in Figures 12–14.

In our numerical experiments we set the distance between the effective viewpoint and the pinhole to be $c = 1$ meter, and the distance from the viewpoint to the world point $\mathbf{w}$ to be $l = 5$ meters. For the hyperboloidal and ellipsoidal mirrors, we set the radius of the lens aperture to be 10 mm. For the paraboloidal mirror, the limiting aperture is the one at the focal point. We chose the size of this aperture so that it lets through exactly the same rays of light that the front 10 mm one would for a point 1 meter away on the optical axis. We assumed the focal length to be 10 cm and therefore set the aperture to be 1 mm. With these settings, the F-stop for the paraboloidal mirror is $2 \times 10/100 = 1/5$. The results for the other two mirrors are independent of the focal length, and hence the F-stop.

To allow the three mirror shapes to be compared on an equal basis, we used values for k and h that correspond to the same mirror radii. The radius of the mirror is taken to be the radius of the mirror cut off by the plane $z = 0$; i.e. the mirrors are all taken to image the entire upper hemisphere. Some values of k and h are plotted in Table 1 against the corresponding mirror radius, for $c = 1$ meter.

Table 1: The mirror radius as a function of the mirror parameters (k and h) for $c = 1$ meter.

Mirror Radius	Hyperboloid (k)	Ellipsoid (k)	Paraboloid (h)
20 cm	6.1	0.24	0.2
10 cm	11.0	0.11	0.1
5 cm	21.0	0.05	0.05
2 cm	51.0	0.02	0.02

From Figures 12–14, we see that the best focus setting varies considerably across the mirror for all of the mirror shapes. Moreover, the variation is roughly comparable for all three mirrors (of equal sizes.) In practice, these results, often referred to as "field curvature" [Hecht and Zajac, 1974], mean that it can sometimes be difficult to focus the entire scene

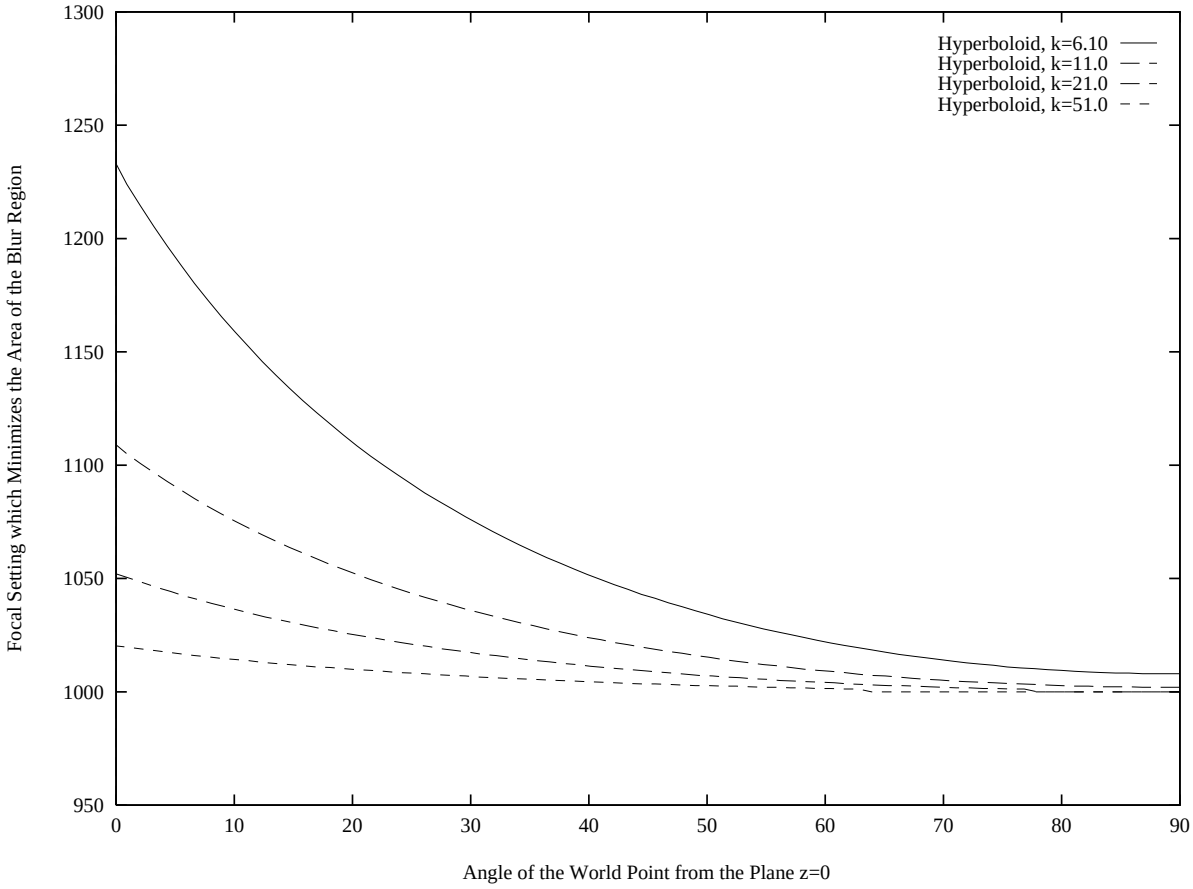


Figure 12: The focus setting which minimizes the area of the blur region plotted against the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. Four separate curves are plotted for different values of the parameter k. See Table 1 for the corresponding radii of the mirrors. We see that the best focus setting for $\mathbf{w}$ varies considerably across the mirror. In practice, these results mean that it can sometimes be difficult to focus the entire scene at the same time, unless additional compensating lenses are used to compensate for the field curvature [Hecht and Zajac, 1974]. Also, note that this effect becomes less important as k increases and the mirror gets smaller.

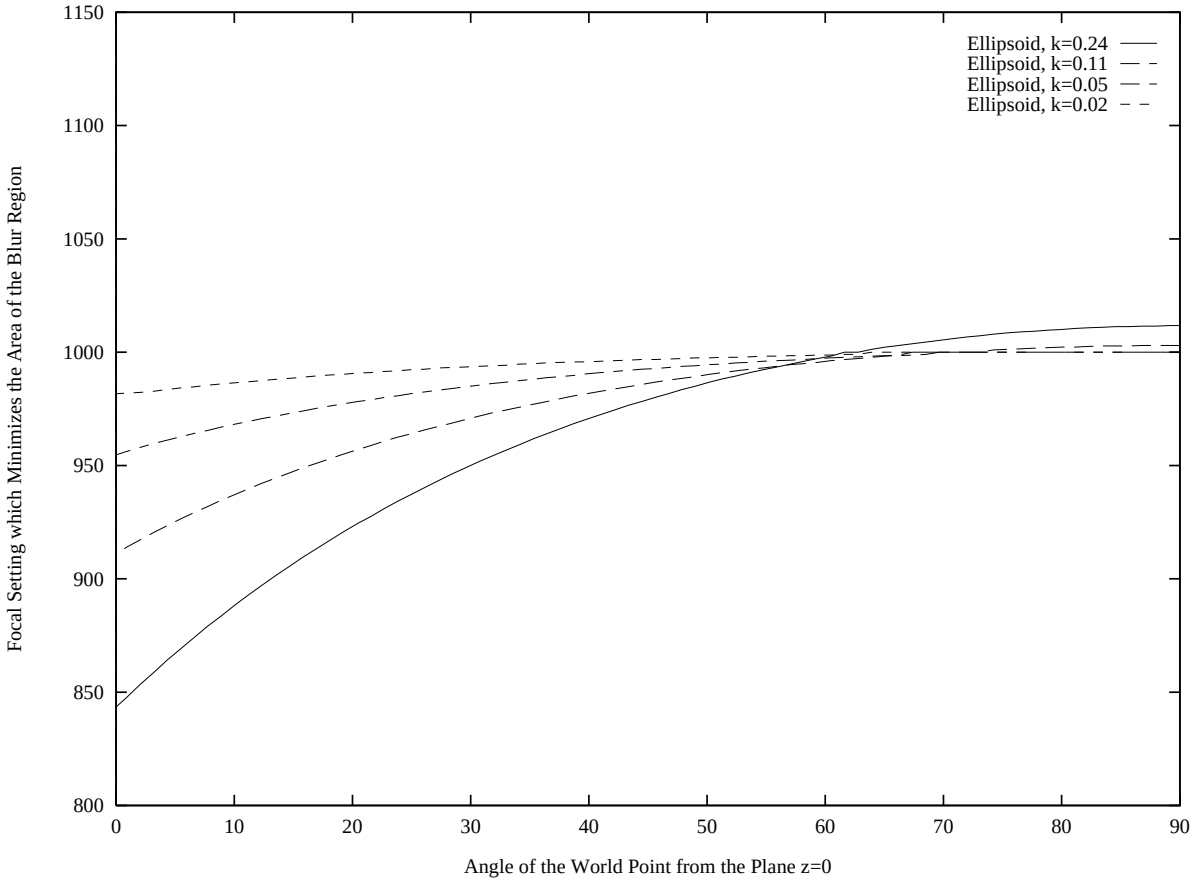


Figure 13: The focus setting which minimizes the area of the blur region plotted against the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. Four separate curves are plotted for different values of the parameter k. See Table 1 for the corresponding radii of the mirrors. The field curvature for the ellipsoidal mirror is roughly comparable to that for the hyperboloidal, and also decreases rapidly as the mirror is made smaller.

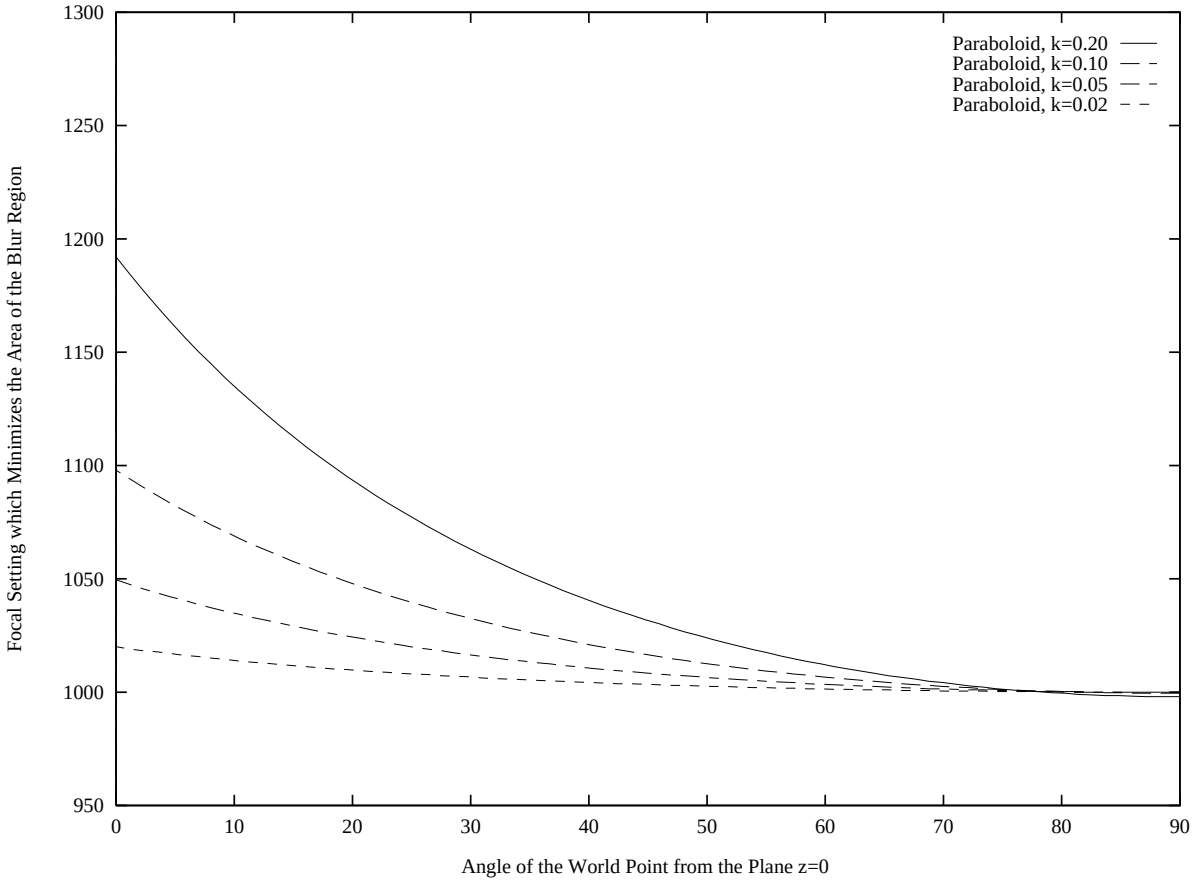


Figure 14: The focus setting which minimizes the area of the blur region plotted against the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. Four separate curves are plotted for different values of the parameter h. See Table 1 for the corresponding radii of the mirrors. The field curvature for the paraboloidal mirror is roughly comparable to that for the hyperboloidal, and also decreases rapidly as the mirror is made smaller.

at the same time. Either the center of the mirror is well focused or the points around the periphery are focused, but not both. Fortunately, it is possible to introduce additional lenses which compensate for the field curvature [Hecht and Zajac, 1974]. Also note that as the mirrors become smaller in size (k increases for the hyperboloid, k decreases for ellipsoid, and h decreases for the paraboloid) the effect becomes significantly less pronounced.

5 Case Study: Parabolic Omnidirectional Cameras

As a case study, we describe the design and implementation of a parabolic omnidirectional camera [Nayar, 1997a]. As described above, such a camera requires an orthographic camera. There are several ways to achieve orthographic projection. The most obvious of these is to use commercially available telecentric lenses [Edmund Scientific, 1996] that are designed to be orthographic. It has also been shown [Watanabe and Nayar, 1996] that precise orthography can be achieved by simply placing an aperture [Kingslake, 1983] at the back focal plane of an off-the-shelf lens. Further, several zoom lenses can be adjusted to produce orthographic projection. Yet another approach is to mount an inexpensive relay lens onto an off-the-shelf perspective lens. The relay lens not only converts the imaging system to an orthographic one but can also be used to reduce more subtle optical effects such as coma and astigmatism [Born and Wolf, 1965] produced by curved mirrors. In short, the implementation of pure orthographic projection is viable and easy to implement.

One advantage of using an orthographic camera is that it can make the calibration of the catadioptric system far easier. Calibration is simpler because, so long as the direction of orthographic projection remains parallel to the axis of the paraboloid, any size of paraboloid is a solution. The paraboloid constant and physical size of the mirror therefore do not need to be determined during calibration. Moreover, the mirror can be translated arbitrarily and still remain a solution. Implementation of the camera is therefore also much easier because the camera does not need to be positioned precisely. By the same token, the fact that the mirror may be translated arbitrarily can be used to set up simple configurations

where the camera zooms in on part of the paraboloid mirror to achieve higher resolution (with a reduced field of view), but without the complication of having to compensate for the additional non-linear distortion caused by the rotation of the camera that would be needed to achieve the same effect in the perspective case.

5.1 Selection of the Field of View

As the extent of the paraboloid increases, so does the field of view of the catadioptric camera. It is not possible, however, to acquire the entire sphere of view since the paraboloid itself must occlude the world beneath it. This brings us to an interesting practical consideration: Where should the paraboloid be terminated? Note that

$$\left| \frac{dz}{dr} \right| \Big|_{z=0} = 1 \ . \tag{53}$$

Hence, if we cut the paraboloid at the plane $z = 0$, the field of view exactly equals the upper hemisphere (minus the solid angle subtended by the imaging system itself). If a field of view greater than a hemisphere is desired, the paraboloid can be terminated below the $z = 0$ plane. If only a panorama is of interest, an annular section of the paraboloid may be obtained by truncating it below and above the $z = 0$ plane. For that matter, given any desired field of view, the the corresponding section of the parabola can be used and the entire resolution of the imaging device can be dedicated to that section. In other words, for an orthographic imaging system of given magnification, the parabolic mirror can be resized and translated horizontally to obtain any desired field of view. Note that the resulting imaging system also adheres to the single viewpoint constraint.

For the prototypes we present here, we have chosen to terminate the parabola at the $z = 0$ plane. This proves advantageous in applications in which the complete sphere of view is desired, as shown in Figure 15. Since the paraboloid is terminated at the focus, it is possible to place two identical catadioptric cameras back-to-back such that their foci (viewpoints) coincide. The shaded regions represents a small part of the field that is lost due

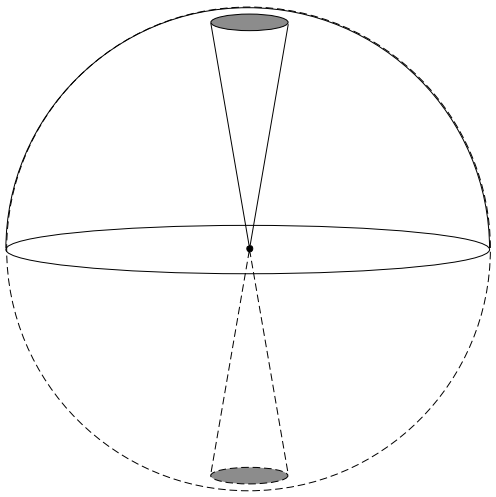
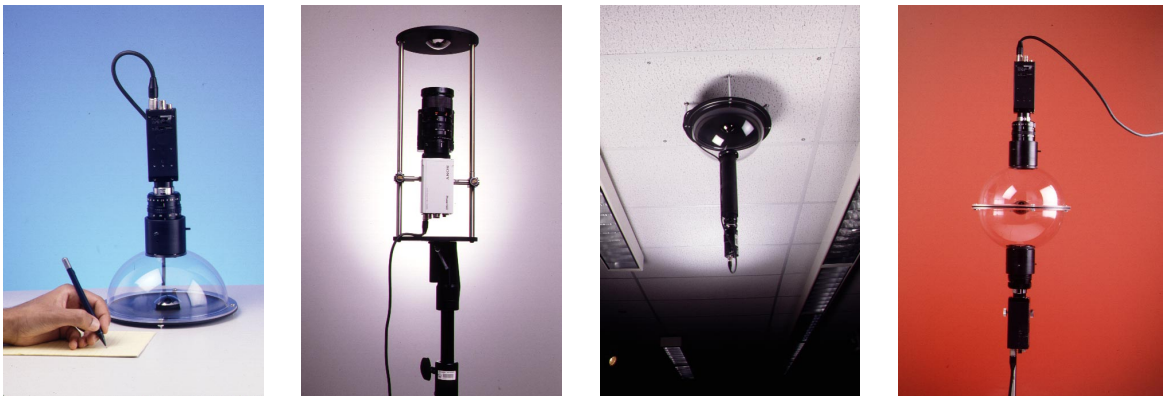


Figure 15: If the paraboloid is cut by the horizontal plane that passes through its focus, the field of view of the catadioptric system exactly equals the upper hemisphere. This allows us to place two catadioptric cameras back-to-back such that their foci (viewpoints) coincide. The result is a truly omnidirectional camera that can acquire the entire sphere of view. The shaded regions are parts of the field of view where the camera sees itself.

to obstruction by the imaging system itself. Thus, we have a truly omnidirectional camera, one that is capable of acquiring an entire sphere of view at video rate.

5.2 Implementations of Parabolic Systems

Several versions of the catadioptric design based on the paraboloidal mirror have been implemented at Columbia University [Nayar, 1997a]. These sensors were designed keeping specific specific applications in mind. The applications we have in mind include video teleconferencing, remote surveillance and autonomous navigation. Figure 16 shows and details the different cameras and their components. The basic components of all the cameras are the same; each one includes a paraboloidal mirror, an orthographic lens system and a CCD video camera. The cameras differ primarily in the their mechanical designs and their attachments. For instance, the cameras in Figures 16(a) and 16(c) have transparent spherical domes that minimize self-obstruction of their hemispherical fields of view. Figure 16(d) shows a back-



(a) (b) (c) (d)

Figure 16: Four implementations of catadioptric omnidirectional video cameras that use paraboloidal mirrors. (a) This compact camera for *teleconferencing* uses a 1.1 inch diameter paraboloidal mirror, a Panasonic GP-KR222 color camera, and Cosmicar/Pentax C6Z1218 zoom and close-up lenses to achieve orthography. The transparent spherical dome minimizes self-obstruction of the field of view. (b) This camera for *navigation* uses a 2.2 inch diameter mirror, a DXC-950 Sony color camera, and a Fujinon CVL-713 zoom lens. The base plate has an attachment that facilitates easy mounting on mobile platforms. (c) This camera for *surveillance* uses a 1.6 inch diameter mirror, an Edmund Scientific 55mm F/2.8 telecentric (orthographic) lens and a Sony XR-77 black and white camera. The camera is lightweight and suitable for mounting on ceilings and walls. (d) This camera is a back-to-back configuration that enables it to sense the entire sphere of view. Each of its two units is identical to the camera in (a).

to-back implementation that is capable of acquiring the complete sphere of view.

The use of paraboloidal mirrors virtually obviates calibration. All that is needed are the image coordinates of the center of the paraboloid and its radius h. Both these quantities are measured in pixels from a single omnidirectional image. We have implemented software for the generation of perspective images. First, the user specifies the viewing direction, the image size and effective focal length (zoom) of the desired perspective image. Again, all these quantities are specified in pixels. For each three-dimensional pixel location (x_p,y_p,z_p) on the desired perspective image plane, its line of sight with respect to the viewpoint is computed in terms of its polar and azimuthal angles:

$$\theta \;=\; \cos^{-1} \frac{z_p}{\sqrt{{x_p}^2 + {y_p}^2 + {z_p}^2}} \;, \quad \phi \;=\; \tan^{-1} \frac{y_p}{x_p} \;. \tag{54}$$

This line of sight intersects the paraboloid at a distance ρ from its focus (origin), which is computed using the following spherical expression for the paraboloid:

$$\rho \;=\; \frac{h}{(1 \,+\, \cos\theta)} \;. \tag{55}$$

The brightness (or color) at the perspective image point (x_p,y_p,z_p) is then the same as that at the omnidirectional image point

$$x_i \;=\; \rho \sin\theta \cos\phi \;, \qquad y_i \;=\; \rho \sin\theta \sin\phi \;. \tag{56}$$

The above computation is repeated for all points in the desired perspective image. Figure 17 shows an omnidirectional image (512x480 pixels) and several perspective images (200x200 pixels each) computed from it. It is worth noting that perspective projection is indeed preserved. For instance, straight lines in the scene map to straight lines in the perspective images while they appear as curved lines in the omnidirectional image. A video-rate version of the above described image generation was developed as an interactive software system called OmniVideo [Peri and Nayar, 1997]. This system can generate about a dozen perspective image streams at 30 Hz using no more than a standard PC.

(a) (b)

Figure 17: Software generation of (b) perspective images from an (a) omnidirectional image. Each perspective image is generated using user-selected parameters, including, viewing direction (line of sight from the viewpoint to the center of the desired image), effective focal length (distance of the perspective image plane from the viewpoint of the camera), and image size (number of desired pixels in each of the two dimensions). It is clear that the computed images are indeed perspective; for instance, straight lines are seen to appear as straight lines even though they appear as curved lines in the omnidirectional image.

6 Conclusion

In this chapter, we have studied three design criteria for catadioptric cameras: (1) the shape of the mirrors, (2) the resolution of the cameras, and (3) the focus settings of the cameras. In particular, we have derived the complete class of mirrors that can be used with one camera to give a single effective viewpoint, found an expression for the resolution of a catadioptric camera in terms of the resolution of the conventional camera used to construct it, and presented detailed analysis of the defocus blur caused by the use of a curved mirror.

We have described a large number of mirror shapes in this chapter, including cones, spheres, planes, hyperboloids, ellipsoids, and paraboloids. Practical catadioptric cameras have been constructed using most of these mirror shapes. See, for example, [Rees, 1970], [Charles *et al.*, 1987], [Nayar, 1988], [Yagi and Kawato, 1990], [Hong, 1991], [Goshtasby and Gruver, 1993], [Yamazawa *et al.*, 1993], [Bogner, 1995], [Nalwa, 1996], and [Nayar, 1997a]. As described in [Chahl and Srinivassan, 1997], even more mirror shapes are possible if we relax the single-viewpoint constraint. Which then is the "best" mirror shape to use?

Unfortunately, there is no simple answer to this question. If the application requires exact perspective projection, there are three alternatives: (1) the ellipsoid, (2) the hyperboloid, and (3) the paraboloid. The major limitation of the ellipsoid is that only a hemisphere can be imaged. As far as the choice between the paraboloid and the hyperboloid goes, using an orthographic imaging system does require extra effort on behalf of the optical designer, but makes construction and calibration of the entire catadioptric system far easier. No careful positioning of the camera relative to the mirror is needed. Moreover, all that is required to calibrate the camera is the image of the circle formed by the circumference of the mirror; no physical distances or other parameters are needed to obtain accurate perspective images.

Acknowledgments

This work was conducted at the Computer Vision Laboratory (CAVE) at Columbia University. It was supported in parts by an ONR/DARPA MURI grant under ONR contract No. N00014-97-1-0553, an NSF National Young Invenstigator Award and a David and Lucile Packard Fellowship. Parts of this chapter were previously published in the International Journal of Computer Vision and the proceedings of the 1997 IEEE Computer Vision and Pattern Recognition Conference (CVPR).

References

[Adelson and Bergen, 1991] E.H. Adelson and J.R. Bergen. The plenoptic function and elements of early vision. In Landy and Movshon, editors, *Computational Models of Visual Processing*, chapter 1. MIT Press, 1991.

[Baker and Nayar, 1998] S. Baker and S.K. Nayar. A theory of catadioptric image formation. In *Proceedings of the 6th International Conference on Computer Vision*, pages 35–42, Bombay, India, January 1998.

[Baker and Nayar, 1999] S. Baker and S.K. Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2):175–196, November 1999.

[Bogner, 1995] S. Bogner. Introduction to panoramic imaging. In *Proceedings of the IEEE SMC Conference*, pages 3100–3106, October 1995.

[Born and Wolf, 1965] M. Born and E. Wolf. *Principles of Optics*. London:Permagon, 1965.

[Chahl and Srinivassan, 1997] J.S. Chahl and M.V. Srinivassan. Reflective surfaces for panoramic imaging. *Applied Optics*, 36(31):8275–8285, November 1997.

[Charles *et al.*, 1987] J.R. Charles, R. Reeves, and C. Schur. How to build and use an all-sky camera. *Astronomy Magazine*, April 1987.

[Drucker and Locke, 1996] D. Drucker and P. Locke. A natural classification of curves and surfaces with reflection properties. *Mathematics Magazine*, 69(4):249–256, October 1996.

[Edmund Scientific, 1996] *1996 Optics and Optical Components Catalog*, volume 16N1. Edmund Scientific Company, New Jersey, 1996.

[Gortler *et al.*, 1996] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The lumigraph. In *Computer Graphics Proceedings, Annual Conference Series*, pages 43–54. ACM SIGGRAPH, 1996.

[Goshtasby and Gruver, 1993] A. Goshtasby and W.A. Gruver. Design of a single-lens stereo camera system. *Pattern Recognition*, 26(6):923–937, 1993.

[Hecht and Zajac, 1974] E. Hecht and A. Zajac. *Optics*. Addison-Wesley, 1974.

[Hong, 1991] J. Hong. Image based homing. In *Proceedings of the IEEE International Conference on Robotics and Automation*, May 1991.

[Inaba *et al.*, 1993] M. Inaba, T. Hara, and H. Inoue. A stereo viewer based on a single camera with view-control mechanism. In *Proceedings of the International Conference on Robots and Systems*, July 1993.

[Kingslake, 1983] R. Kingslake. *Optical System Design*. Academic Press, 1983.

[Murphy, 1995] J.R. Murphy. Application of panoramic imaging to a teleoperated lunar rover. In *Proceedings of the IEEE SMC Conference*, pages 3117–3121, October 1995.

[Nalwa, 1996] V.S. Nalwa. A true omnidirectional viewer. Technical report, Bell Laboratories, Holmdel, NJ 07733, USA, February 1996.

[Nayar and Baker, 1997] S.K. Nayar and S. Baker. Catadioptric image formation. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, pages 1431–1437, New Orleans, Louisiana, May 1997.

[Nayar, 1988] S.K. Nayar. Sphereo: Recovering depth using a single camera and two specular spheres. In *Proceedings of SPIE: Optics, Illumination, and Image Sensing for Machine Vision II*, November 1988.

[Nayar, 1997a] S. K. Nayar. Catadioptric Omnidirectional Camera. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, June 1997.

[Nayar, 1997b] S.K. Nayar. Omnidirectional video camera. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, May 1997.

[Nene and Nayar, 1998] S.A. Nene and S.K. Nayar. Stereo with mirrors. In *Proceedings of the 6th Internation Conference on Computer Vision*, Bombay, India, January 1998. IEEE Computer Society.

[Peri and Nayar, 1997] V. Peri and S. K. Nayar. Generation of Perspective and Panoramic Video from Omnidirectional Video. *Proc. of DARPA Image Understanding Workshop*, May 1997.

[Rees, 1970] D.W. Rees. Panoramic television viewing system. United States Patent No. 3,505,465, April 1970.

[Watanabe and Nayar, 1996] M. Watanabe and S. K. Nayar. Telecentric optics for computational vision. *Proc. of European Conference on Computer Vision*, April 1996.

[Yagi and Kawato, 1990] Y. Yagi and S. Kawato. Panoramic scene analysis with conic projection. In *Proceedings of the International Conference on Robots and Systems*, 1990.

[Yagi and Yachida, 1991] Y. Yagi and M. Yachida. Real-time generation of environmental map and obstacle avoidance using omnidirectional image sensor with conic mirror. In *Proceedings of the 1991 Conference on Computer Vision and Pattern Recognition*, pages 160–165, June 1991.

[Yagi *et al.*, 1994] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (COPIS) for vision-guided navigation. *IEEE Transactions on Robotics and Automation*, 10(1):11–22, February 1994.

[Yamazawa *et al.*, 1993] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional imaging with hyperboloidal projection. In *Proceedings of the International Conference on Robots and Systems*, 1993.

[Yamazawa *et al.*, 1995] K. Yamazawa, Y. Yagi, and M. Yachida. Obstacle avoidance with omnidirectional image sensor HyperOmni Vision. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1062–1067, May 1995.