International Journal of Computer Vision 35(2), 175–196 (1999)

A Theory of Single-Viewpoint Catadioptric Image Formation

SIMON BAKER
The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213
simonb@cs.cmu.edu

SHREE K. NAYAR
Department of Computer Science, Columbia University, New York, NY 10027
nayar@cs.columbia.edu

Abstract. Conventional video cameras have limited fields of view which make them restrictive for certain applications in computational vision. A catadioptric sensor 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 sensors 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 paper, we derive the complete class of single-lens single-mirror catadioptric sensors 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 sensor in terms of the resolution of the cameras used to construct it. Moreover, we include detailed analysis of the defocus blur caused by the use of a curved mirror in a catadioptric sensor.

Keywords: image formation, sensor design, sensor resolution, defocus blur, omnidirectional imaging, panoramic imaging

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; 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; 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 so as 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 geometrically correct panoramic images as well as perspective ones.

In this paper, we take the view that having a single viewpoint is the primary design goal for the catadioptric sensor and restrict attention to catadioptric sensors with a single effective viewpoint (Baker and Nayar, 1998). 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, sensors that do not obey the single viewpoint requirement can also be used. Then, other design issues become more important, such as spatial resolution, sensor 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 paper in Section 2 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 sensors with a single effective viewpoint. Many of these solutions have, however, been used to construct wide field of view sensors with non-constant viewpoints. For these mirror shapes, we derive the loci of the viewpoint. Some, but not all, of the non-degenerate solutions have been used in sensors proposed in the literature. In these cases, we mention all of the designs that we are aware of. A different, coordinate free, 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 sensor that images a large field of view is its resolution. The resolution of a catadioptric sensor is not, in general, the same as that of any of the sensors 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 sensor. 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 sensor has sufficient resolution. Another use of the relationship is to design conventional sensors 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 sensors. 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 sensor using a curved mirror may be substantially different from that needed in a conventional sensor. 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).

2. The Fixed Viewpoint Constraint

The fixed viewpoint constraint is the requirement that a catadioptric sensor 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 sensor 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 sensor 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 Fig. 1 for an illustration[1] of the coordinate frame.

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

$$\frac{dz}{dr} = -\tan\beta. \tag{3}$$

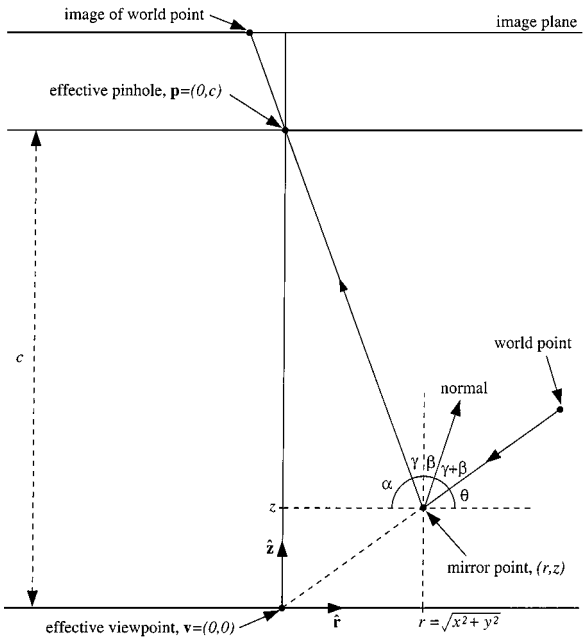

Figure 1. The geometry used to derive the fixed viewpoint constraint equation. The viewpoint $\mathbf{v} = (0, 0)$ is located at the origin of a 2-D coordinate frame $(\mathbf{v}, \hat{\mathbf{r}}, \hat{\mathbf{z}})$, and the pinhole of the camera $\mathbf{p} = (0, c)$ is located at a distance c from $\mathbf{v}$ 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 $\gamma = 90^\circ - \alpha$. Finally, if $\beta = \tan^{-1}(-\frac{dz}{dr})$ 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 $\alpha + \theta + 2\gamma + 2\beta = 180^\circ$.

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 $\gamma = 90^\circ - \alpha$ and $\theta + \alpha + 2\beta + 2\gamma = 180^\circ$. (See Fig. 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:

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

we have:

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

Substituting from Eqs. (1)–(3) yields the *fixed viewpoint constraint* equation:

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

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

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

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:

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

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

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

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

$$2y\frac{dy}{dr} = 2x + 2r\frac{dx}{dr} - 2r \tag{11}$$

and so we have:

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

Rearranging this equation yields:

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

Integrating both sides with respect to r results in:

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

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:

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

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

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, Eqs. (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 sensors with a single effective viewpoint. We will describe the solutions in detail in the following order:

Planar Solutions: Equation (16) with $k = 2$ and $c > 0$.
Conical Solutions: Equation (16) with $k \geq 2$ and $c = 0$.
Spherical Solutions: Equation (17) with $k > 0$ and $c = 0$.
Ellipsoidal Solutions: Equation (17) with $k > 0$ and $c > 0$.
Hyperboloidal Solutions: Equation (16) with $k > 2$ and $c > 0$.

For each solution, we demonstrate whether it is degenerate or not. Some of the non-degenerate solutions have actually been used in real sensors. 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 sensors with a wide field of view, but with no fixed viewpoint. In these cases we derive the loci of the viewpoint.

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 Eq. (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. We briefly discuss the orthographic case and the corresponding paraboloid solution in Section 2.4.

2.3.1. Planar Mirrors. In Solution (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 Fig. 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 more precise, the effective viewpoint is the reflection of the pinhole $\mathbf{p}$ in the mirror; i.e. the single effective viewpoint is:

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

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; Nene and Nayar, 1998).

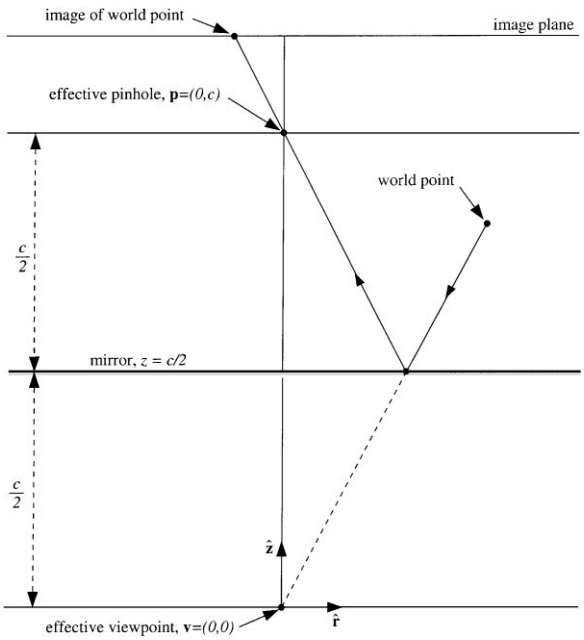

Figure 2. The plane $z = \frac{c}{2}$ is a solution of the fixed viewpoint constraint equation. Conversely, it is possible to show that, given a fixed viewpoint and pinhole, the only planar solution is the perpendicular bisector of the line joining the pinhole to the viewpoint. Hence, for a fixed pinhole, 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).

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 sensor 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 sensor 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 Solution (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 Fig. 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 (finite extent) objects in the world (see Fig. 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 sensor with a single 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). If the pinhole lies on the axis of the cone at a distance e from the apex of the cone, the locus of the effective viewpoint is a circle. The radius of the circle is easily seen to be:

$$e \cdot \cos 2\tau. \tag{23}$$

If $\tau > 60^\circ$, the circular locus lies inside (below) the cone, if $\tau < 60^\circ$ the circular locus lies outside (above) the cone, and if $\tau = 60^\circ$ the circular locus lies on the cone. In some applications such as robot navigation, the single viewpoint constraint is not vital. Conical mirrors can be used to build practical sensors for such applications. See, for example, the designs in (Yagi et al., 1994; Bogner, 1995).

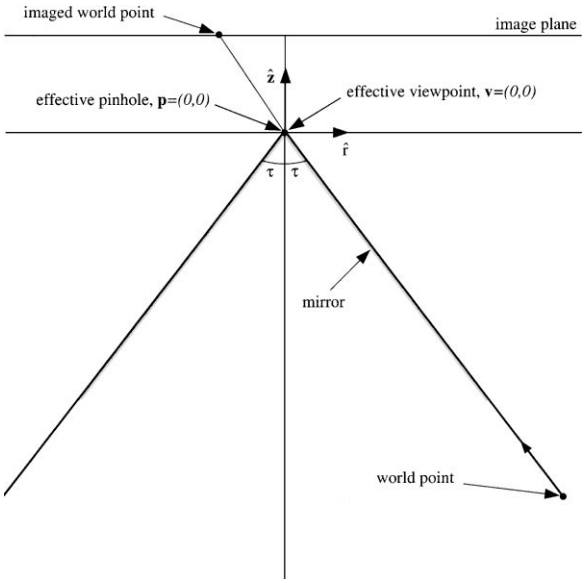

Figure 3. The conical mirror is a solution of the fixed viewpoint constraint equation. Since the pinhole is located at the apex of the cone, this is a degenerate solution that cannot be used to construct a wide field of view sensor with a single viewpoint. If the pinhole is moved away from the apex of the cone (along the axis of the cone), the viewpoint is no longer a single point but rather lies on a circular locus. If 2τ is the angle at the apex of the cone, the radius of the circular locus of the viewpoint is $e \cdot \cos 2\tau$, where e is the distance of the pinhole from the apex along the axis of the cone. If $\tau > 60^\circ$, the circular locus lies inside (below) the cone, if $\tau < 60^\circ$ the circular locus lies outside (above) the cone, and if $\tau = 60^\circ$ the circular locus lies on the cone.

2.3.3. Spherical Mirrors. In Solution (17), if we set $c = 0$ and $k > 0$, we get the spherical mirror:

$$z^2 + r^2 = \frac{k}{2}. \tag{24}$$

Like the cone, this is a degenerate solution which cannot be used to construct a wide field of view sensor with a single viewpoint. Since the viewpoint and pinhole coincide at the center of the sphere, the observer would see itself and nothing else, as is illustrated in Fig. 4.

The sphere has also been used to build wide field of view sensors several times (Hong, 1991; Bogner, 1995; Murphy, 1995). In these implementations, the

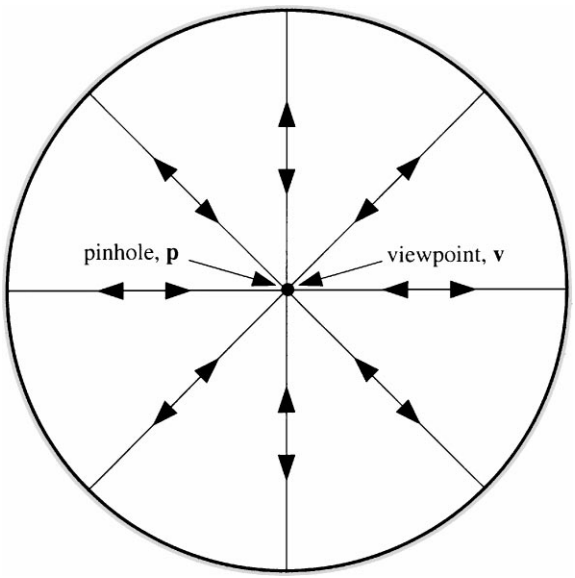

Figure 4. The spherical mirror satisfies the fixed viewpoint constraint when the pinhole lies at the center of the sphere. (Since $c = 0$ the viewpoint also lies at the center of the sphere.) Like the conical mirror, the sphere cannot actually be used to construct a wide field of view sensor with a single 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.

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. Without loss of generality, suppose that the radius of the mirror is 1.0. The first step is to compute the direction of the ray of light which would be reflected at the mirror point $(r, z) = (r, \sqrt{1-r^2})$ and then pass through the pinhole. This computation is then repeated for the neighboring mirror point $(r + dr, z + dz)$. Next, the intersection of these two rays is computed, and finally the limit $dr \to 0$ is taken while constraining dz by $(r + dr)^2 + (z + dz)^2 = 1$. The result of performing this derivation is that the locus of the effective viewpoint is:

$$\left(\frac{c[1 + c(1 + 2r^2)\sqrt{1-r^2}]}{1 + 2c^2 - 3c\sqrt{1-r^2}}, \frac{2c^2r^2}{1 + 2c^2 - 3c\sqrt{1-r^2}}\right) \tag{25}$$

as r varies from $-\sqrt{1-\frac{1}{c^2}}$ to $\sqrt{1-\frac{1}{c^2}}$. The locus of the effective viewpoint is plotted for various values of c in Fig. 5. As can be seen, for all values of c the locus

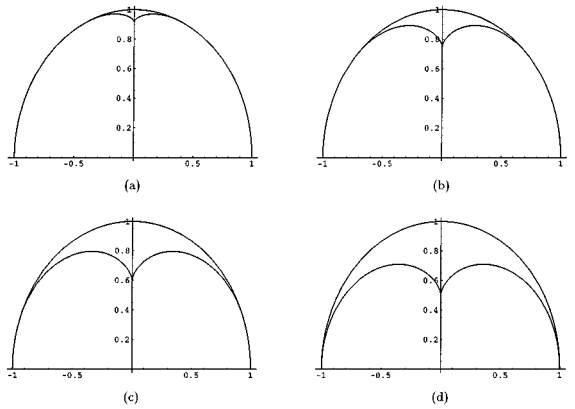

Figure 5. The locus of the effective viewpoint of a circular mirror of radius 1.0 (which is also shown) plotted for $c = 1.1$ (a), $c = 1.5$ (b), $c = 3.0$ (c), and $c = 100.0$ (d). For all values of c, the locus lies within the mirror and is of comparable size to the mirror.

lies within the mirror and is of comparable size to it. 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 Solution (17), when $k > 0$ and $c > 0$, we get the ellipsoidal mirror:

$$\frac{1}{a_e^2}\left(z - \frac{c}{2}\right)^2 + \frac{1}{b_e^2}r^2 = 1 \tag{26}$$

where:

$$a_e = \sqrt{\frac{2k + c^2}{4}} \quad \text{and} \quad b_e = \sqrt{\frac{k}{2}}. \tag{27}$$

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 Fig. 6, 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 Solution (16), when $k > 2$ and $c > 0$, we get the hyperboloidal mirror:

$$\frac{1}{a_h^2}\left(z - \frac{c}{2}\right)^2 - \frac{1}{b_h^2}r^2 = 1 \tag{28}$$

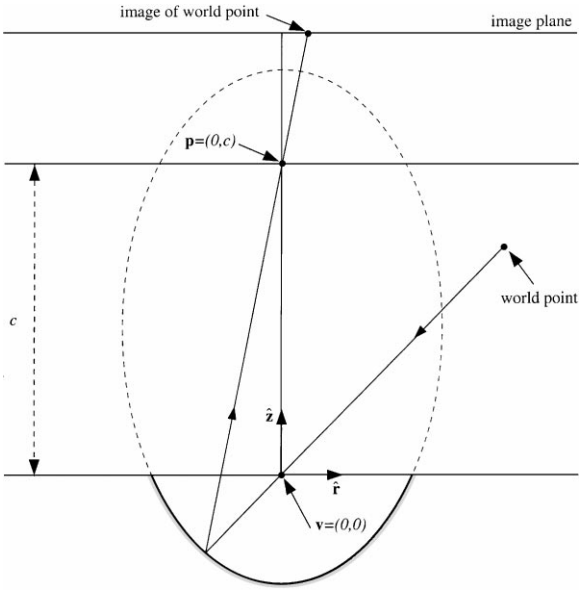

Figure 6. The ellipsoidal mirror satisfies the fixed viewpoint constraint when the pinhole and viewpoint are located at the two foci of the ellipsoid. If the ellipsoid is terminated by the horizontal plane passing through the viewpoint $z = 0$, the field of view is the entire upper hemisphere $z > 0$. It is also possible to cut the ellipsoid with other planes passing through **v**, but it appears there is little to be gained by doing so.

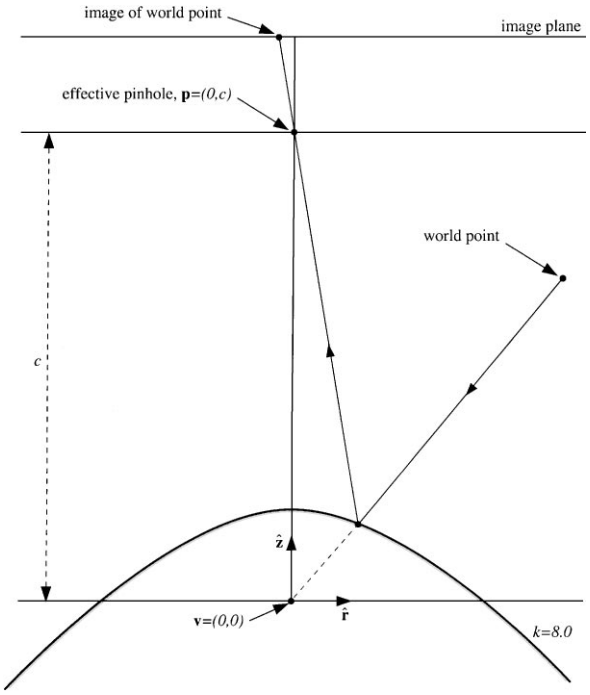

Figure 7. The hyperboloidal mirror satisfies the fixed viewpoint constraint when the pinhole and the viewpoint are located at the two foci of the hyperboloid. This solution does produce the desired increase in field of view. The curvature of the mirror and hence the field of view increase with k. In the limit $k \to 2$, the hyperboloid flattens to the planar mirror of Section 2.3.1.

where:

$$a_h = \frac{c}{2}\sqrt{\frac{k-2}{k}} \quad \text{and} \quad b_h = \frac{c}{2}\sqrt{\frac{2}{k}}. \qquad (29)$$

As seen in Fig. 7, 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 sensor designed for autonomous navigation.

2.4. The Orthographic Case: Paraboloidal Mirrors

Although the parabola is not a solution of the fixed viewpoint constraint equation for finite values of c and k, it is a solution of Eq. (16) in the limit that $c \to \infty$, $k \to \infty$, and $\frac{c}{k} = h$, a constant. Under these limiting conditions, Eq. (16) tends to:

$$z = \frac{h^2 - r^2}{2h}. \qquad (30)$$

As shown in (Nayar, 1997b) and Fig. 8, this limiting case corresponds to orthographic projection. Moreover, in that setting the paraboloid does yield a practical omnidirectional sensor with a number of advantageous properties (Nayar, 1997b).

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 sensor is therefore also much easier because

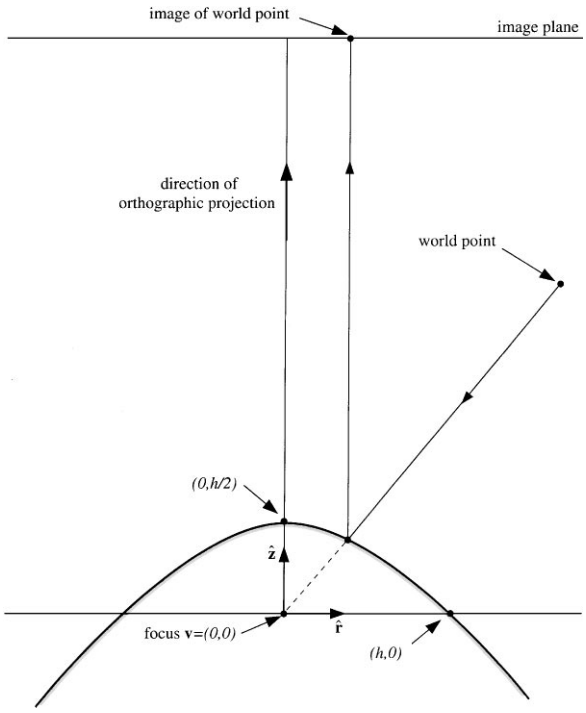

Figure 8. Under orthographic projection, the only solution is a paraboloid with the effective viewpoint at the focus of the paraboloid. One advantage of this solution is that the camera can be translated arbitrarily and remain a solution. This property can greatly simplify sensor calibration (Nayar, 1997b). The assumption of orthographic projection is not as restrictive a solution as it may sound since there are simple ways to convert a standard lens and camera from perspective projection to orthographic projection. See, for example, (Nayar, 1997b).

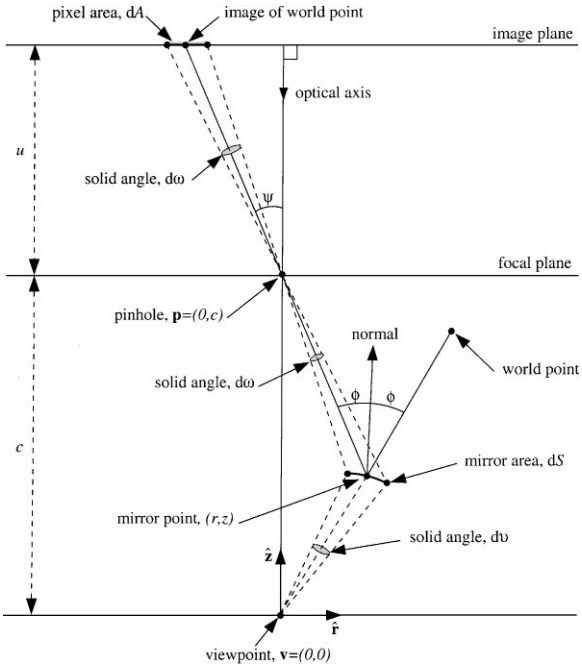

Figure 9. The geometry used to derive the spatial resolution of a catadioptric sensor. Assuming the conventional sensor has a frontal image plane which is located at a distance u from the pinhole and the optical axis is aligned with the z-axis $\hat{\mathbf{z}}$, the spatial resolution of the conventional sensor is $\frac{dA}{d\omega} = \frac{u^2}{\cos^3\psi}$. Therefore the area of the mirror imaged by the infinitesimal image plane area dA is $dS = \frac{(c-z)^2 \cdot \cos\psi}{u^2 \cos\phi} \cdot dA$. So, the solid angle of the world imaged by the infinitesimal area dA on the image plane is $d\nu = \frac{(c-z)^2 \cdot \cos\psi}{u^2 (r^2+z^2)} \cdot dA$. Hence, the spatial resolution of the catadioptric sensor is $\frac{dA}{d\nu} = \frac{u^2(r^2+z^2)}{(c-z)^2 \cdot \cos\psi} = \frac{r^2+z^2}{r^2+(c-z)^2} \cdot \frac{dA}{d\omega}$ since $\cos^2\psi = \frac{(c-z)^2}{(c-z)^2+r^2}$.

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.

3. Resolution of a Catadioptric Sensor

In this section, we 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 axis of symmetry of the mirror. See Fig. 9 for an illustration of this scenario. Then, the definition of resolution that we will use is the following. Consider an infinitesimal area dA on the image plane. If this infinitesimal pixel images an infinitesimal solid angle $d\nu$ of the world, the *resolution* of the sensor as a function of the point on the image plane at the center of the infinitesimal area dA is:

$$\frac{dA}{d\nu}. \tag{31}$$

If ψ is the angle made between the optical axis and the line joining the pinhole to the center of the infinitesimal area dA (see Fig. 9), the solid angle subtended by the infinitesimal area dA at the pinhole is:

$$d\omega = \frac{dA \cdot \cos\psi}{u^2/\cos^2\psi} = \frac{dA \cdot \cos^3\psi}{u^2}. \tag{32}$$

Therefore, the resolution of the conventional camera is:

$$\frac{dA}{d\omega} = \frac{u^2}{\cos^3\psi}. \tag{33}$$

Then, the area of the mirror imaged by the infinitesimal area dA is:

$$dS = \frac{d\omega \cdot (c-z)^2}{\cos\phi\cos^2\psi} = \frac{dA \cdot (c-z)^2 \cdot \cos\psi}{u^2\cos\phi} \tag{34}$$

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:

$$d\nu = \frac{dS \cdot \cos\phi}{r^2+z^2} = \frac{dA \cdot (c-z^2) \cdot \cos\psi}{u^2(r^2+z^2)}. \tag{35}$$

Therefore, the resolution of the catadioptric camera is:

$$\frac{dA}{d\nu} = \frac{u^2(r^2+z^2)}{(c-z)^2 \cdot \cos\psi} = \left[\frac{(r^2+z^2)\cos^2\psi}{(c-z)^2}\right]\frac{dA}{d\omega} \tag{36}$$

But, since:

$$\cos^2\psi = \frac{(c-z)^2}{(c-z)^2+r^2} \tag{37}$$

we have:

$$\frac{dA}{d\nu} = \left[\frac{r^2+z^2}{(c-z)^2+r^2}\right]\frac{dA}{d\omega}. \tag{38}$$

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

$$\frac{r^2+z^2}{(c-z)^2+r^2} \tag{39}$$

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

The first thing to note from Eq. (38) is that for the planar mirror $z = \frac{c}{2}$, the resolution of the catadioptric sensor is the same as that of the conventional sensor used to construct it. This is as expected by symmetry. Secondly, note that the factor in Eq. (39) 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 Eq. (39) is $d_\mathbf{v}^2/d_\mathbf{p}^2$. 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:

$$\left(\frac{K_e}{d_\mathbf{p}} - 1\right)^2 \tag{40}$$

which increases as d_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:

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

which increases as $d_\mathbf{p}$ increases and $d_\mathbf{v}$ increases. So, for both ellipsoids and hyperboloids, the factor in Eq. (39) increases with r. Hence, both hyperboloidal and ellipsoidal catadioptric sensors constructed with a uniform resolution conventional camera will have their highest resolution around the periphery, a useful property for certain applications such as teleconferencing.

3.1. The Orthographic Case

The orthographic case is slightly simpler than the projective case and is illustrated in Fig. 10. 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:

$$\frac{dA}{d\omega} = M^2 \tag{42}$$

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

$$d\omega = \cos\phi \cdot dS. \tag{43}$$

Combining Eqs. (35), (42), and (43) yields:

$$\frac{dA}{d\nu} = [r^2+z^2]\frac{dA}{d\omega}. \tag{44}$$

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

$$\left[\frac{h^2+r^2}{2h}\right]^2. \tag{45}$$

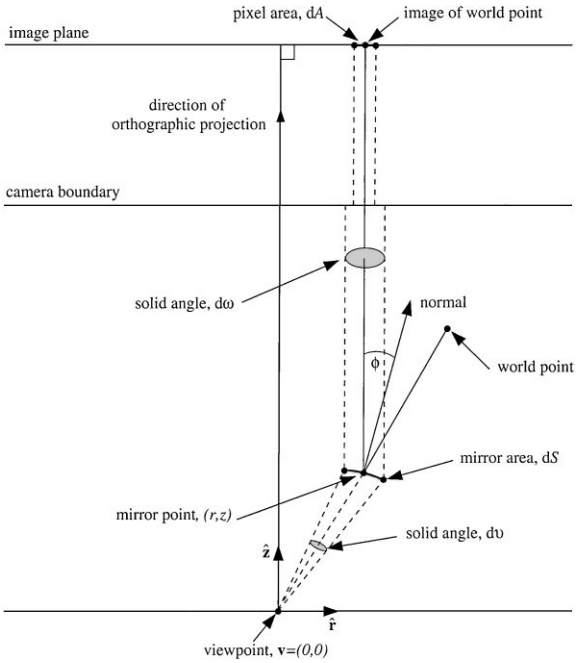

Figure 10. The geometry used to derive the spatial resolution of a catadioptric sensor in the orthographic case. Again, assuming that the image plane is frontal and the conventional orthographic camera has a linear magnification M, its spatial resolution is $\frac{dA}{d\omega} = M^2$. The solid angle $d\omega$ equals $\cos\phi \cdot dS$, where dS 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 Eq. (35) yields the spatial resolution of the orthographic catadioptric sensor as $\frac{dA}{dv} = [r^2 + z^2]\frac{dA}{d\omega}$.

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

4. Defocus Blur of a Catadioptric Sensor

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 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, ellipsoid, and paraboloid mirrors.

4.1. *Analysis of Defocus Blur*

To analyze defocus blur, we need to work in 3-D. We use the 3D 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 Fig. 11 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 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*

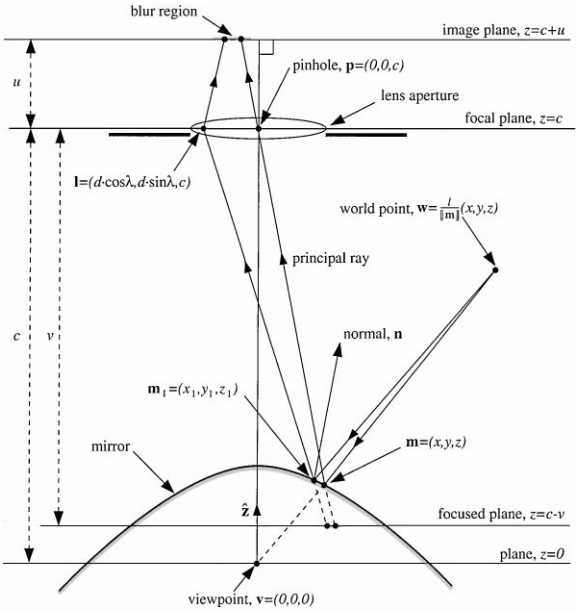

Figure 11. The geometry used to analyze the defocus blur. We work in the 3D cartesian frame $(\mathbf{v}, \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal unit vectors in the plane $z = 0$. In addition to the assumptions of Section 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.

plane. By the thin lens law (Hecht and Zajac, 1974) the focused plane is:

$$z = c - v = c - \frac{f \cdot u}{u - f} \tag{46}$$

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:

$$-\frac{u}{v}\left(\mathbf{l} + \frac{v}{c - z_1}(\mathbf{m}_1 - \mathbf{l})\right) \tag{47}$$

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:

$$\left(z_1 - \frac{c}{2}\right)^2 - \left(x_1^2 + y_1^2\right)\left(\frac{k}{2} - 1\right) = \frac{c^2}{4}\left(\frac{k-2}{k}\right). \tag{48}$$

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:

$$\mathbf{n} = ([k-2]x_1, [k-2]y_1, c - 2z_1) \tag{49}$$

for the hyperboloid. Hence, the second constraint is:

$$\mathbf{n} \cdot (\mathbf{w} - \mathbf{m}_1) \wedge (\mathbf{l} - \mathbf{m}_1) = 0. \tag{50}$$

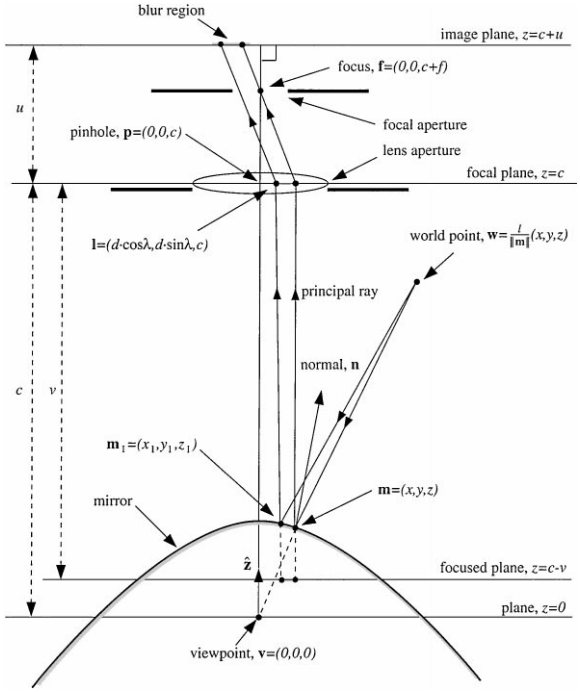

Figure 12. The geometry used to analyze defocus blur in the orthographic case. One way to create orthographic projection is to add a (circular) aperture at the rear focal point (the one behind the lens) (Nayar, 1997b). Then, the only rays of light that reach the image plane are those which are (approximately) 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.

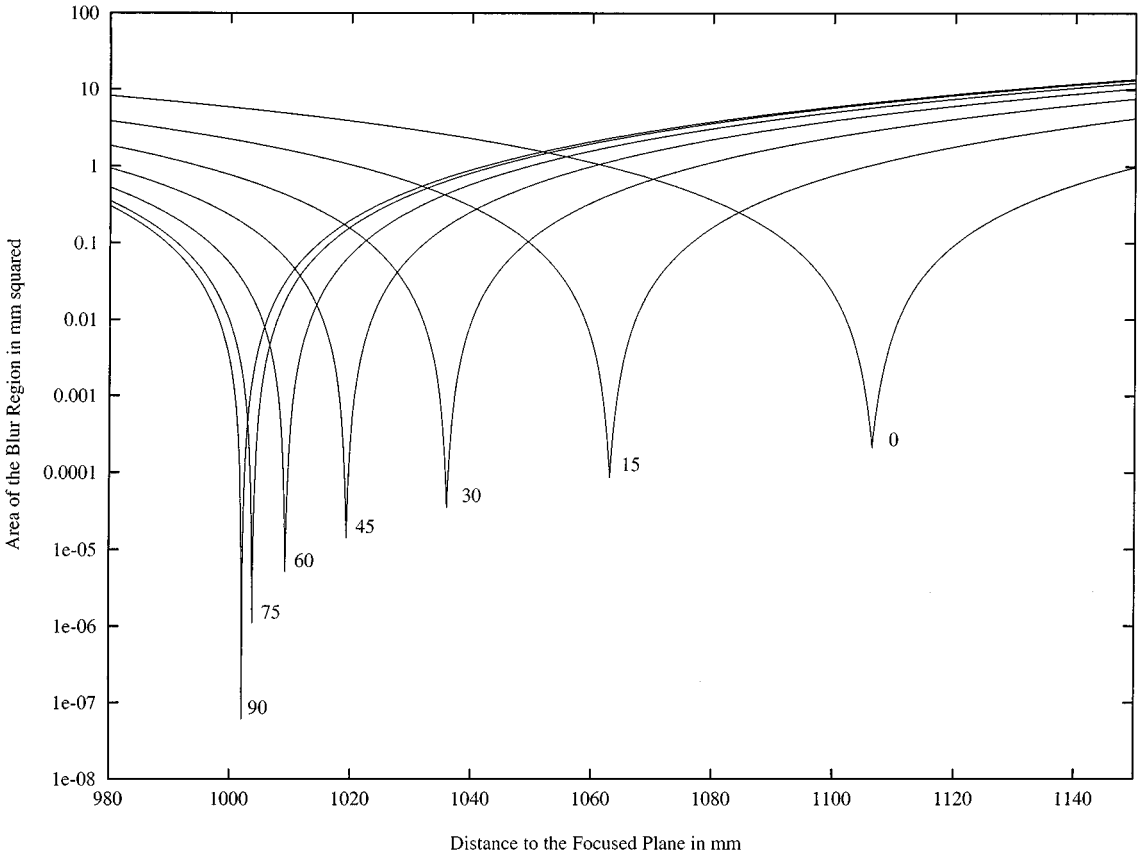

Figure 13. The area of the blur region plotted against the distance to the focused plane $v = \frac{f \cdot u}{u - f}$ for the hyperboloidal mirror with $k = 11.0$. In this example, we have $c = 1$ meter, the radius of the lens aperture 10 millimeters, and the distance from the viewpoint to the world point $l = 5$ meters. We plot curves for 7 different world points, at 7 different angles from the plane $z = 0$. The area of the blur region never becomes exactly zero and so the image can never be perfectly focused. However, the area does become very small and so focusing on a single point is not a problem in practice. Note that the distance at which the image will be best focused (around 1.0–1.15 meters) is much less than the distance from the pinhole to the world point (approximately 1 meter from the pinhole to the mirror plus 5 meters from the mirror to the world point.) The reason is that the mirror is convex and so tends to increase the divergence of rays of light.

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

$$\frac{\mathbf{n} \cdot (\mathbf{w} - \mathbf{m}_1)}{\|\mathbf{w} - \mathbf{m}_1\|} = \frac{\mathbf{n} \cdot (\mathbf{l} - \mathbf{m}_1)}{\|\mathbf{l} - \mathbf{m}_1\|}. \tag{51}$$

These three constraints on $\mathbf{m}_1$ are all multivariate polynomials in x_1, y_1, and z_1: Eqs. (48) and (50) are both of order 2, and Eq. (51) is of order 5. We were unable to find a closed form solution to these three equations (Eq. (51) has 25 terms in general and so it is probable that none exists) but we did investigate numericals solution. 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 Fig. 12. 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

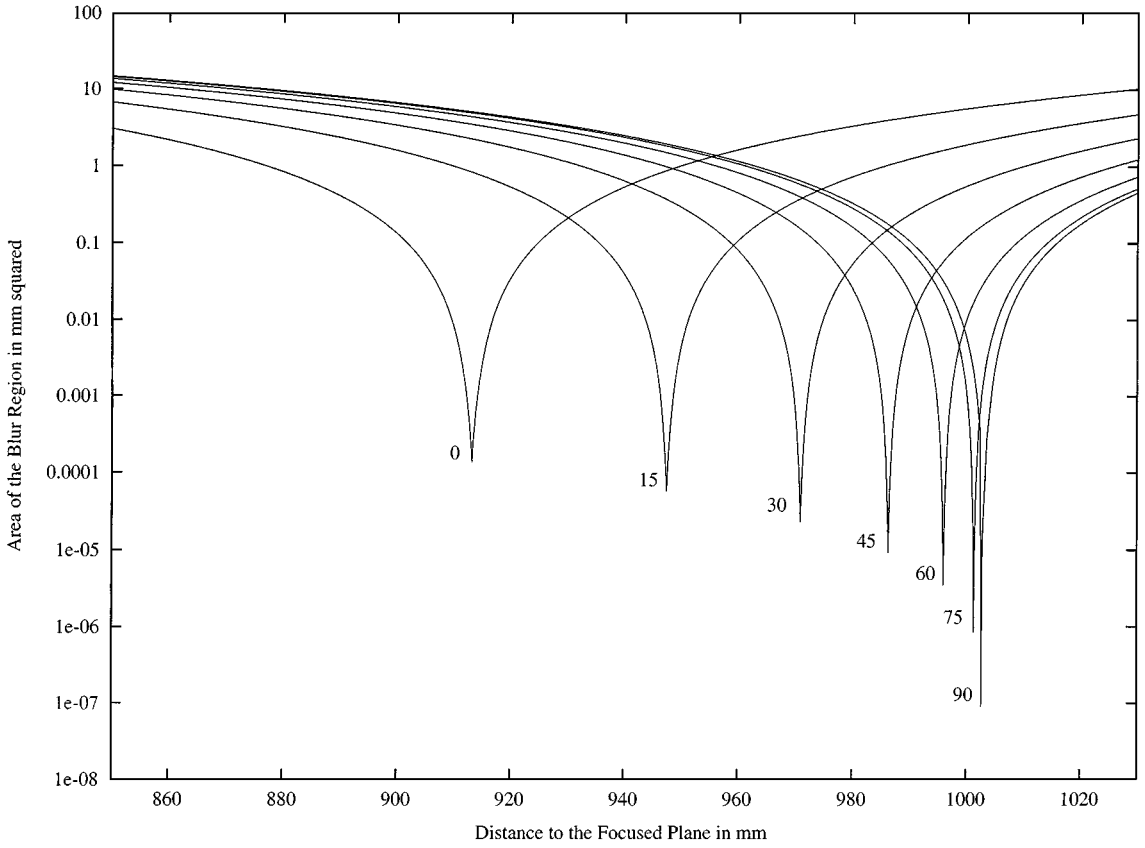

Figure 14. The area of the blur region plotted against the distance to the focused plane $v = \frac{f \cdot u}{u - f}$ for the ellipsoidal mirror with $k = 0.11$. The other settings are the same as for the hyperboloidal mirror in Fig. 13. Again, the distance to the focused plane is less than the distance to the point in the world, however the reason is different. For the concave ellipsoidal mirror, a virtual image is formed between the mirror and the lens. The lens needs to focus on this virtual image.

where the mirror point is imaged on the image plane. It is then checked whether this point lies close enough to the optical axis.

4.3. Numerical Results

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.

4.3.1. Area of the Blur Region. In Figs. 13–15, we plot the area of the blur region (on the ordinate) against the distance to the focused plane v (on the abscissa) for the hyperboloidal, ellipsoidal, and paraboloidal mirrors. In each figure, we plot separate curves for different world point directions. The angles are measures in degrees from the plane $z = 0$, and so the curve at 90° corresponds to the (impossible) world point directly

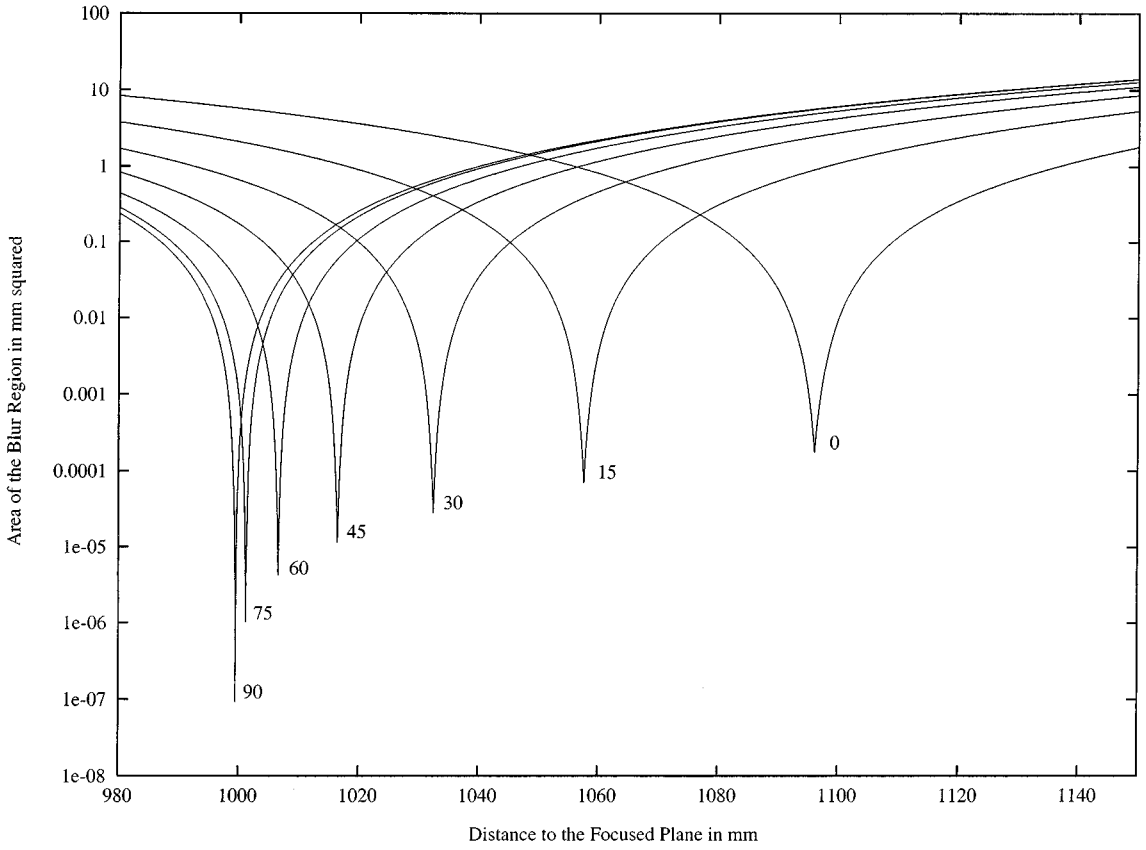

Figure 15. The area of the blur region plotted against the distance to the focused plane $v = \frac{f \cdot u}{u - f}$ for the paraboloidal mirror with $h = 0.1$. The settings are the same as for the hyperboloidal mirror, except the size of the apertures. The limiting aperture is the one at the focal point. It is chosen so that it lets through exactly the same rays of light that the 10 mm one does for the hyperboloidal mirror for a point 1 meter away on the optical axis. The results are qualitatively very similar to the hyperboloidal mirror.

upwards in the direction of the z-axis. For the hyperboloid we set $k = 11.0$, for the ellipsoid $k = 0.11$, and for the paraboloid $h = 0.1$. As can be seen in Table 1, these settings correspond to a mirror with radius 10 cm. Qualitatively similar results were obtained for the other radii. Section 4.3.3 contains related results for the other radii.

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

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

The smaller the area of the blur region, the better focused the image will be. We see from the figures that the area never reaches exactly zero, and so an image formed using these catadioptric sensors can never be perfectly focused. However, the minimum area is very small, and in practice there is no problem focusing the image for a single world point. Moreover, it is possible to use additional corrective lenses to compensate for most of this effect (Hecht and Zajac, 1974).

Note that the distance at which the image of the world point will be best focused (i.e. somewhere in the range 0.9–1.15 meters) is much less than the distance from the pinhole to the world point (approximately 1 meter from the pinhole to the mirror plus 5 meters from the mirror to the world point). The reason for this effect is that the mirror is curved. For the hyperboloidal and paraboloidal mirrors which are convex, the curvature tends to increase the divergence of rays coming from

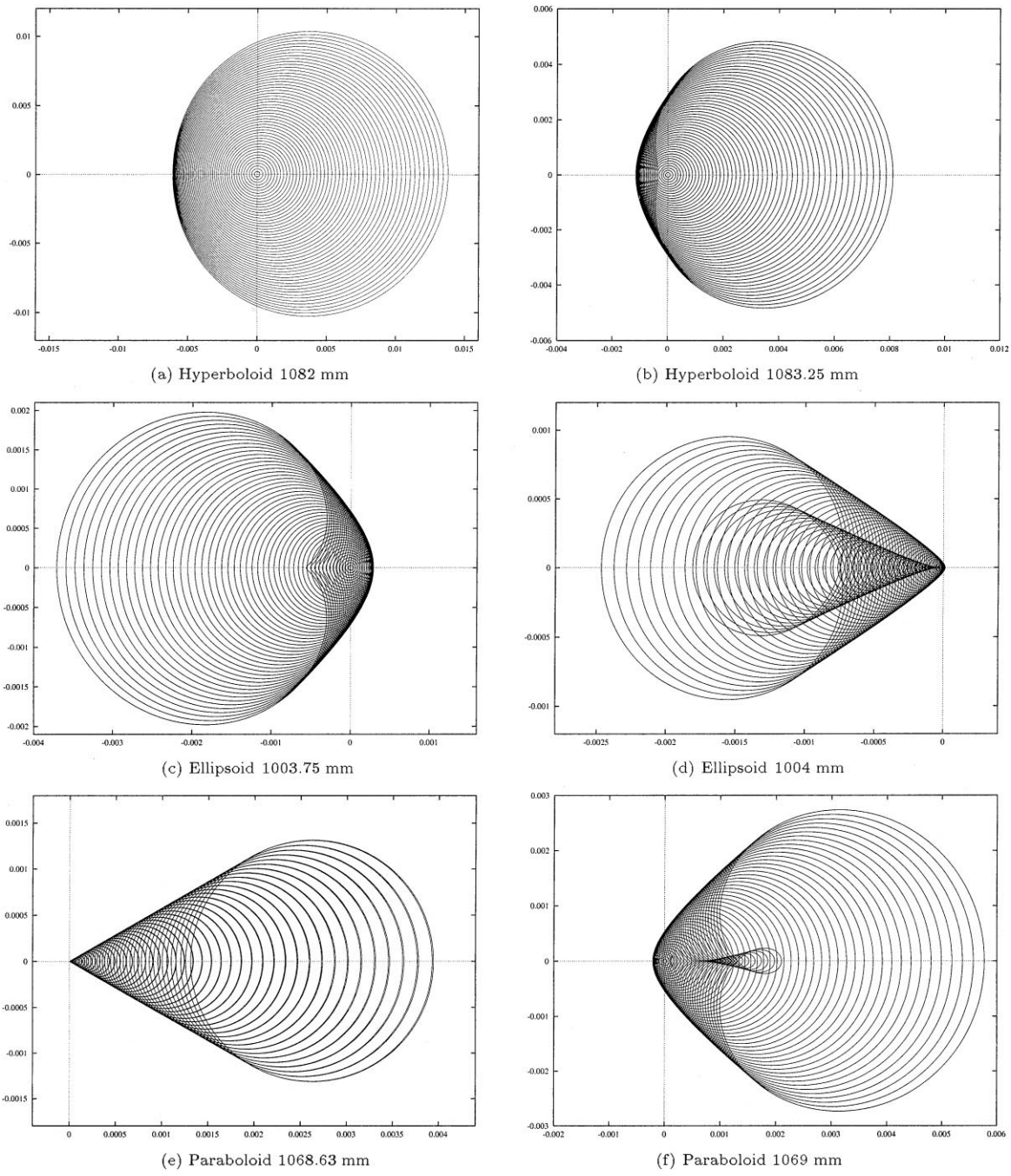

Figure 16. The variation in the shape of the blur region as the focus setting is varied. Note that all of the blur regions in this figure are relatively well focused. Also, note that the scale of the 6 figures are all different.

the world point. For these rays to be converged and the image focused, a larger distance to the image plane u is needed. A larger value of u corresponds to a smaller value of v, the distance to the focused plane. For the concave ellipsoidal mirror, the mirror converges the rays to the extent that a virtual image is formed between the mirror and the lens. The lens must be focused on this virtual image.

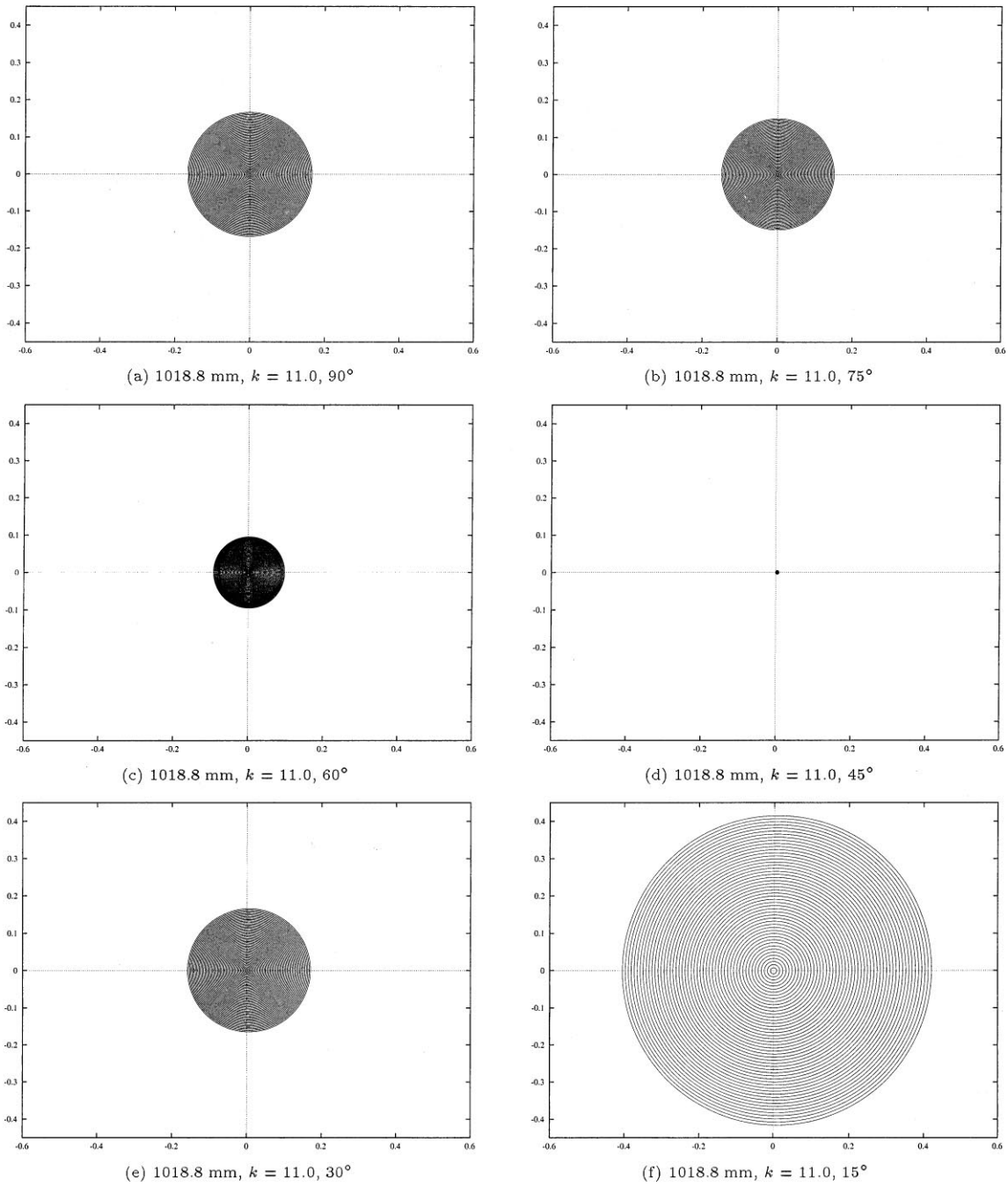

Figure 17. An example of the variation in the blur region as a function of the angle of the point in the world. In this example for the hyperboloid with $k = 11.0$, the point at 45° is in focus, but the points in the other directions are not.

4.3.2. Shape of the Blur Region. Next, we provide an explanation of the fact that the area of the blur region never exactly reaches zero. For a conventional lens, the blur region is a circle. In this case, as the focus setting is adjusted to focus the lens, all points on the blur circle move towards the center of the blur circle at a rate which is proportional to their distance from the center of the blur circle. Hence, the blur circle steadily

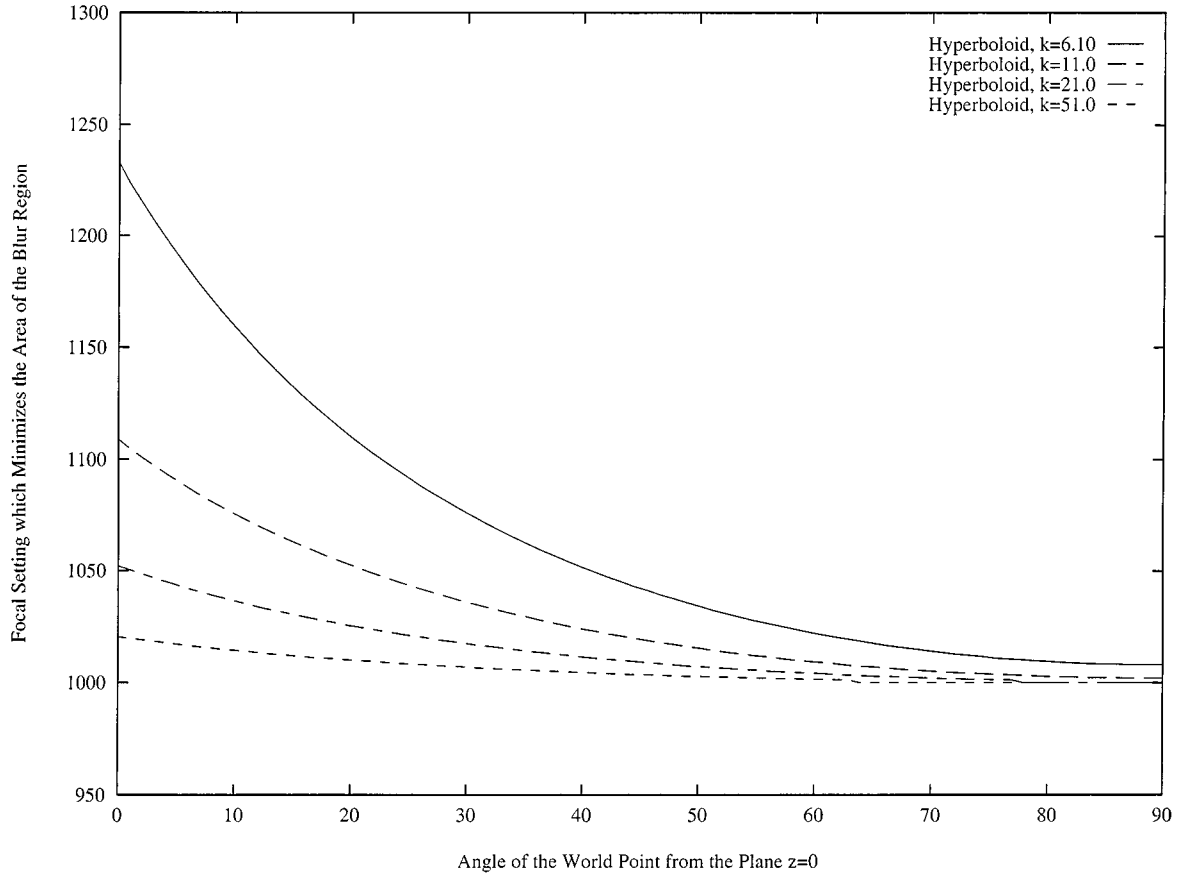

Figure 18. The focus setting which minimizes the area of the blur region in Fig. 13 plotted against the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. Four separate curves are plotted for different values of the parameter k. See Table 1 for the corresponding radii of the mirrors. 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.

shrinks until the blur region has area 0 and the lens is perfectly focused. If the focus setting is moved further in the same direction, the blur circle grows again as all the points on it move away from the center.

For a catadioptric sensor using a curved mirror, the blur region is only approximately a circle for all three of the mirror shapes. Moreover, as the image is focused, the speed with which points move towards the center of this circle is dependent on their position in a much more complex way than for a single lens. The behavior is qualitatively the same for all of the mirrors and is illustrated in Fig. 16. From Fig. 16(a) to (e), the blur region gets steadily smaller, and the image becomes more focused. In Fig. 16(f), the focus is beginning to get worse again. In Fig. 16(a) the blur region is roughly a circle, however as the focus gets better, the circle begins to overlap itself, as shown in Fig. 16(b). The degree of overlap increases in Figs. 16(c) and (d). (These 2 figures are for the ellipse and are shown to illustrate how similar the blur regions are for the 3 mirror shapes. The only difference is that the region has been reflected about a vertical axis since the ellipse is a concave mirror.) In Fig. 16(e), the image is as well focused as possible and the blur region completely overlaps itself. In Fig. 16(f), the overlapping has begun to unwind.

Finally, in Fig. 17, we illustrate how the blur regions vary with the angle of the point in the world, for a fixed focal setting. In this figure, which displays results for the hyperboloid with $k = 0.11$, the focal setting is chosen so that the point at 45° is in focus. As can be seen, for points in the other directions the blur region can be quite large and so points in those directions

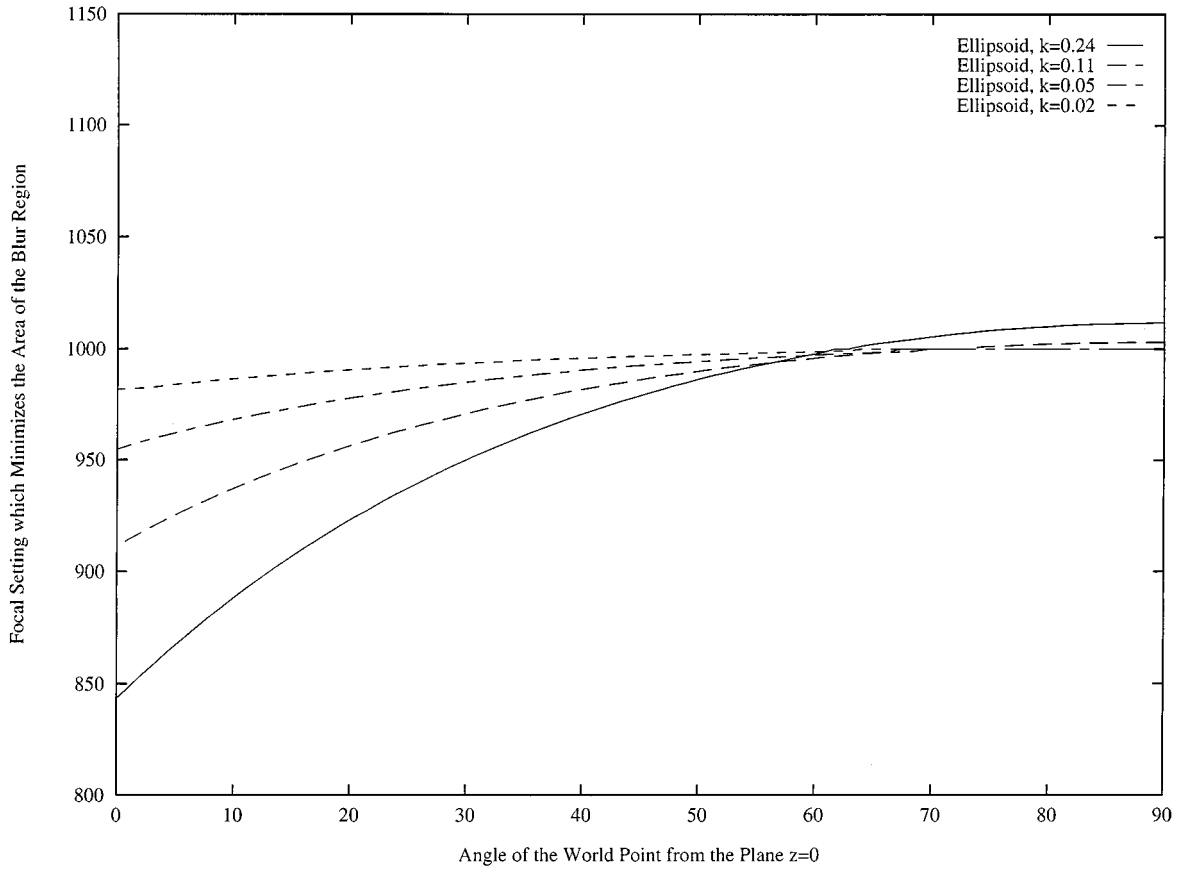

Figure 19. The focus setting which minimizes the area of the blur region in Fig. 14 plotted against the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. Four separate curves are plotted for different values of the parameter k. See Table 1 for the corresponding radii of the mirrors. 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.

are not focused. This effect, known as field curvature (Hecht and Zajac, 1974), is studied in more detail in the following section.

4.3.3. Focal Settings. Finally, we investigated how the focus setting that minimizes the area of the blur region (see Figs. 13–15) changes with the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. The results are presented in Figs. 18–20. As before, we set $c = 1$ meter, assumed the radius of the lens aperture to be 10 millimeters (1 millimeter for the paraboloid), and fixed the world point to be $l = 5$ meters from the effective viewpoint. 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 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). (See the discussion at the end of this paper for more details.) 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. Discussion

In this paper, we have studied three design criteria for catadioptric sensors: (1) the shape of the mirrors, (2) the resolution of the cameras, and (3) the focus settings

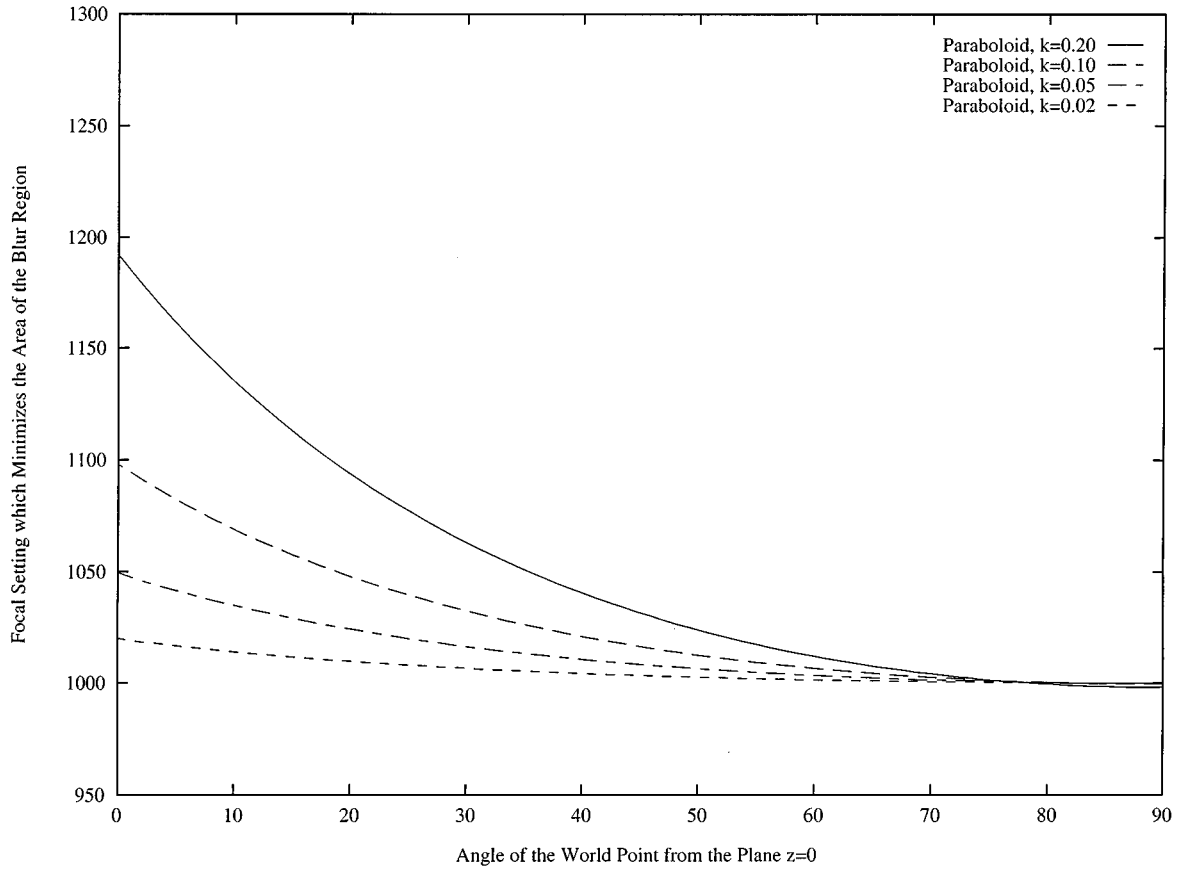

Figure 20. The focus setting which minimizes the area of the blur region in Fig. 15 plotted against the angle θ which the world point $\mathbf{w}$ makes with the plane $z = 0$. Four separate curves are plotted for different values of the parameter h. See Table 1 for the corresponding radii of the mirrors. 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.

of the cameras. In particular, we have derived the complete class of mirrors that can be used with a single camera to give a single viewpoint, found an expression for the resolution of a catadioptric sensor in terms of the resolution of the conventional camera(s) used to construct it, and presented detailed analysis of the defocus blur caused by the use of a curved mirror.

There are a number of possible uses for the (largely theoretical) results presented in this paper. Throughout the paper we have touched on many of their uses by a sensor designer. The results are also of interest to a user of a catadioptric sensor. We now briefly mention a few of the possible uses, both for sensor designers and users:

- For applications where a fixed viewpoint is not a requirement, we have derived the locus of the viewpoint for several mirror shapes. The shape and size of these loci may be useful for the user of such a sensor requiring the exact details of the geometry. For example, if the sensor is being used in an stereo rig, the epipolar geometry needs to be derived precisely.
- The expression for the resolution of the sensor could be used by someone applying image processing techniques to the output of the sensor. For example, many image enhancement algorithms require knowledge of the solid angles of the world integrated over by each pixel in sensor.
- Knowing the resolution function also allows a sensor designer to design a CCD with non-uniform resolution to get an imaging system with a known (for example uniform) resolution.
- The defocus analysis could be important to the user of a catadioptric sensor who wishes to apply various

image processing techniques, from deblurring to restoration and super-resolution.

- Knowing the defocus function also allows a sensor designer to compensate for the field curvature introduced by the use of a curved mirror. One method consists of introducing optical elements behind the imaging lens. For instance, a plano-concave lens placed flush with the CCD permits a good deal of field curvature correction. (Light rays at the periphery of the image travel through a greater distance within the plano-concave lens). Another method is to use a thick meniscus lens right next to the imaging lens (away from the CCD). The same effect is achieved. In both cases, the exact materials and curvatures of the lens surfaces are optimized using numerical simulations. Optical design is almost always done this way as analytical methods are far too cumbersome. See (Born and Wolf, 1965) for more details.

We have described a large number of mirror shapes in this paper, including cones, spheres, planes, hyperboloids, ellipsoids, and paraboloids. Practical catadioptric sensors 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; 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 may also make construction and calibration of the entire catadioptric system easier, as discussed in Section 2.4.

If the application at hand does not require a single viewpoint, many other practical issues may become more important, such as the size of the sensor, its reso-break lution variation across the field of view, and the ease of mapping between coordinate systems. In this paper we have restricted attention to single-viewpoint systems. The reader is referred to other papers proposing catadioptric sensors, such as (Yagi and Kawato, 1990; Yagi and Yachida, 1991; Hong, 1991; Bogner, 1995; Murphy, 1995; Chahl and Srinivassan, 1997), for discussion of the practical merits of catadioptric systems with extended viewpoints.

Acknowledgments

The research described in this paper was conducted while the first author was a Ph.D. student in the Department of Computer Science at Columbia University in the City of New York. This work was supported in parts by the VSAM effort of DARPA's Image Understanding Program and a MURI grant under ONR contract No. N00014-97-1-0553. The authors would also like to thank the anonymous reviewers for their comments which have greatly improved the paper.

Note

1. In Fig. 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 $\hat{\mathbf{z}}$. In this second case, the image plane would be non-frontal. This does not pose any additional problem since the mapping from a non-frontal image plane to a frontal image plane is one-to-one.

References

Adelson, E.H. and Bergen, J.R. 1991. The plenoptic function and elements of early vision. In *Computational Models of Visual Processing*, chap. 1, Landy and Movshon (Eds.), MIT Press.

Baker, S. and Nayar, S.K. 1998. A theory of catadioptric image formation. In *Proceedings of the 6th Internation Conference on Computer Vision*, Bombay, India, IEEE Computer Society, pp. 35–42.

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

Born, M. and Wolf, E. 1965. *Principles of Optics*. Permagon Press.

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

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

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

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

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

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

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

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

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

Nalwa, V.S. 1996. A true omnidirectional viewer. Technical Report, Bell Laboratories, Holmdel, NJ 07733, USA.

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

Nayar, S.K. 1997a. Catadioptric omnidirectional camera. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*, pp. 482–488.

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

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

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

Peri, V. and Nayar, S.K. 1997. Generation of perspective and panoramic video from omnidirectional video. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, New Orleans.

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

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

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

Yagi, Y. and Yachida, M. 1991. 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*, pp. 160–165.

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

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