A Theory of Catadioptric Image Formation

Shree K. Nayar and Simon Baker

(Technical Report CUCS-015-97)

Department of Computer Science

Columbia University

New York, NY 10027

Email: {nayar,simonb}@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. When designing a catadioptric sensor, the shape of the mirror(s) should ideally be selected to ensure 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 image(s). In this paper, we derive and analyze the complete class of single-lens single-mirror catadioptric sensors which satisfy the fixed viewpoint constraint. Some of the solutions turn out to be degenerate with no practical value, while other solutions lead to realizable sensors. We also derive an expression for the spatial resolution of a catadioptric sensor, and include a preliminary analysis of the defocus blur caused by the use of a curved mirror.

Keywords: Image formation, sensor design, sensor resolution, defocus blur.

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. Moreover, a number of other applications, such as ego-motion estimation and tracking, could benefit from enhanced fields of view. Unfortunately, conventional imaging systems are severely limited in their fields of view and so both researchers and practitioners have had to resort to using either multiple or rotating cameras in order to image the entire scene.

A recently developed and effective way to enhance the field of view is to use mirrors in conjunction with lenses. This approach to image capture is fast gaining in popularity (see, for example, [Nayar, 1988], [Yagi and Kawato, 1990], [Hong, 1991], [Goshtasby and Gruver, 1993], [Yamazawa *et al.*, 1993], [Nalwa, 1996], [Nayar and Baker, 1997], and [Nayar, 1997]). We refer to the general approach of using mirrors in combination with conventional imaging systems as *catadioptric*[1] image formation. Recent work in this area has led to the development of a truly omnidirectional video camera with a spherical field of view [Nayar, 1997].

As noted in [Yamazawa *et al.*, 1995], [Nalwa, 1996], and [Nayar, 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 image(s) sensed by the catadioptric cameras. This is possible because under the single viewpoint constraint, every pixel in the sensed image(s) measures the irradiance of the light passing through the viewpoint in one particular direction. Since we know the geometry of the catadioptric sys-

[1] *Dioptrics* is the science of refracting elements (lenses) whereas *catoptrics* is the optics of reflecting surfaces (mirrors). The combination of refracting and reflecting elements is therefore referred to as *catadioptrics* [Hecht and Zajac, 1974].

tem, 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 perspective image. This perspective image can subsequently be processed using the vast array of techniques developed in the field of computational vision which 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 in order to not appear distorted. Naturally, when the catadioptric imaging system is omnidirectional in its field of view, a single viewpoint permits the construction of panoramic images in addition to perspective ones.

In this paper, we begin in Section 2 by deriving the entire class of catadioptric systems with a single effective viewpoint and which are constructed just using a single conventional lens and a single mirror. As we will show, the 2-parameter family of mirrors which can be used is exactly the class of rotated (swept) conic sections. Within this class of solutions, several swept conics prove to be degenerate solutions and hence impractical, while others lead to practical sensors. During our analysis we will stop at many points to evaluate the merits of the solutions as well as the merits of closely related catadioptric sensors proposed in the literature.

An 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 the sensors used to construct it. In Section 3 we study why this is the case, and derive an expression for the relationship between the resolution of a conventional imaging system and the resolution of a derived catadioptric sensor. This expression should be carefully considered when constructing a catadioptric imaging system in order to ensure that the final sensor has sufficient resolution. It could also be used 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 modified by the use of a catadioptric system is focusing. It is a well known fact that a curved mirror increases image blur [Hecht and Zajac, 1974], and so in Section 4 we analyze the effect which the use of a curved mirror has on defocus blur. Two factors combine to cause blur in catadioptric systems: (1) the finite size of the lens, and (2) the curvature of the mirror. We first analyze how the interaction of these two factors causes defocus blur and then present some preliminary numerical results for one specific mirror shape: the hyperboloid. The results show that the focal setting of a catadioptric sensor using a curved mirror may be substantially different to that needed in a conventional sensor. Moreover, our analysis illustrates a numerical method of computing the relationship between the two focus settings which could be used to precompute a focus lookup table for a specific catadioptric sensor.

2 The Fixed Viewpoint Constraint

The fixed viewpoint constraint is a 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 can only 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 (which was reflected by the mirror) would have passed through the effective viewpoint if it had not been reflected by the mirror.

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, also 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[2] 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 it also passes through the origin $\mathbf{v} = (0, 0)$ we have the relationship:

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

[2]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, and 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.

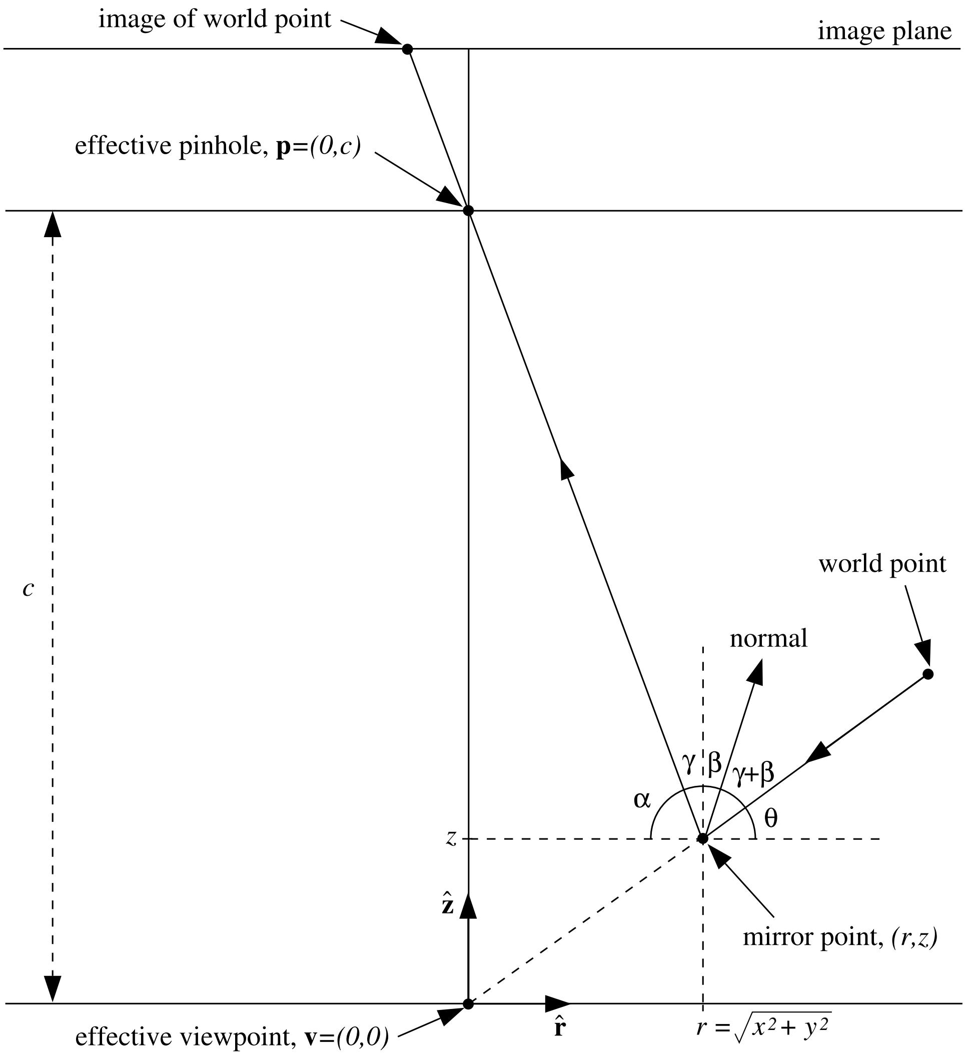


Figure 1: The geometry used to derive the fixed viewpoint constraint equation. The viewpoint $\mathbf{v} = (0, 0)$ is located at the origin of a 2-D coordinate frame $(\mathbf{v}, \hat{\mathbf{r}}, \hat{\mathbf{z}})$, and the pinhole of the camera $\mathbf{p} = (0, c)$ is located at a distance c from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

If we denote by α the angle between the reflected ray and the (negative) r-axis, 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, the two relations in Equations (16) and (17) represent the entire class of mirrors that satisfy the fixed viewpoint constraint. A quick glance at the form of these equations reveals that the mirror profiles are all conic sections. Hence, the 3-D mirrors themselves are swept conic sections. However, as we shall see, although every conic section is theoretically a solution of one of the two equations, a number of them prove to be impractical and only some lead to realizable sensors. We will now describe each of the solutions in detail. Since the class of solutions forms a 2-parameter (c and k) family, we consider them 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 $c > 0$.

There is one conic section which we have not mentioned: the parabola. Although, the parabola is not a solution of either Equation (16) or Equation (17) for finite values of c and k, it is a solution of Equation (16) in the limit that $c \rightarrow \infty$, $k \rightarrow \infty$, and $\frac{c}{k} = h$ a constant. Under these limiting conditions, Equation (16) tends to:

$$z = \frac{h^2 - r^2}{2h}. \qquad (18)$$

As shown in [Nayar, 1997], this limiting case corresponds to orthographic projection. Moreover, in that setting the parabola does yield a practical omnidirectional sensor with a number of advantageous properties. In this paper we restrict attention to the perspective case and refer the reader to [Nayar, 1997] for a discussion of the orthographic case. However, most of the results can either be extended to or applied directly to the orthographic case.

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:

$$z \;=\; \frac{c}{2}. \tag{19}$$

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:

$$\left[\mathbf{x} - \left(\frac{\mathbf{p}+\mathbf{v}}{2}\right)\right] \cdot (\mathbf{p} - \mathbf{v}) \;=\; 0. \tag{20}$$

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 $\mathbf{v} = \mathbf{v}(\hat{\mathbf{n}}, \mathbf{q})$. Moreover, the effective viewpoint is the reflection of the pinhole $\mathbf{p}$ in the mirror; i.e. the single effective viewpoint is:

$$\mathbf{v}(\hat{\mathbf{n}}, \mathbf{q}) \;=\; \mathbf{p} - 2\left[(\mathbf{p} - \mathbf{q}) \cdot \hat{\mathbf{n}}\right] \hat{\mathbf{n}} \tag{21}$$

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

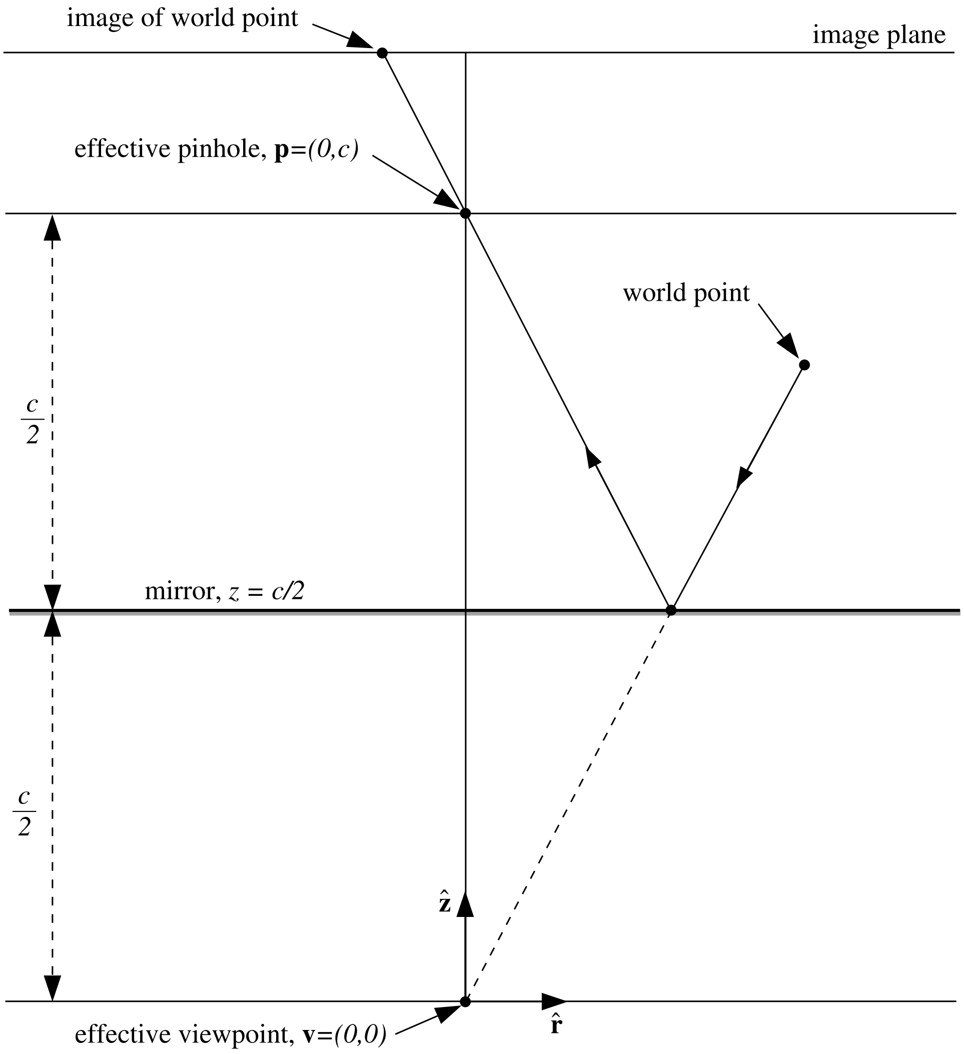


Figure 2: The plane $z = \frac{c}{2}$ is a solution of the fixed viewpoint constraint equation. Conversely, it is possible to show that, given a fixed viewpoint and pinhole, the only planar solution is the perpendicular bisector of the line joining the pinhole to the viewpoint. Hence, for a fixed pinhole,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

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].

This brings us to the panoramic camera proposed by Nalwa [Nalwa, 1996]. To ensure a single viewpoint while using multiple planar mirrors, Nalwa [Nalwa, 1996] has 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 $360° \times 50°$. The panoramic image is of relatively high resolution since it is generated from the four images captured by the four cameras. Nalwa's sensor is straightforward to implement, but requires four of each component: i.e. four cameras, four lenses, and four digitizers. (It is 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:

$$z = \sqrt{\frac{k-2}{2} r^2}. \tag{22}$$

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

$$\tan \tau = \sqrt{\frac{2}{k-2}}. \tag{23}$$

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 Figure 3.) Hence, the cone (with the pinhole at the vertex) is a degenerate solution of no practical value.

The cone has been used for wide-angle imaging [Yagi and Kawato, 1990] [Yagi and Yachida, 1991]. In these implementations the pinhole is placed quite 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:

$$e \cdot \cos 2\tau. \tag{24}$$

If $\tau > 60°$, the circular locus lies inside (below) the cone, if $\tau < 60°$ the circular locus lies outside (above) the cone, and if $\tau = 60°$ 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:

$$z^2 + r^2 = \frac{k}{2}. \tag{25}$$

Like the cone, this is a solution with little practical value. Since the viewpoint and pinhole coincide at the center of the sphere, the observer sees itself and nothing else, as is illustrated in Figure 4.

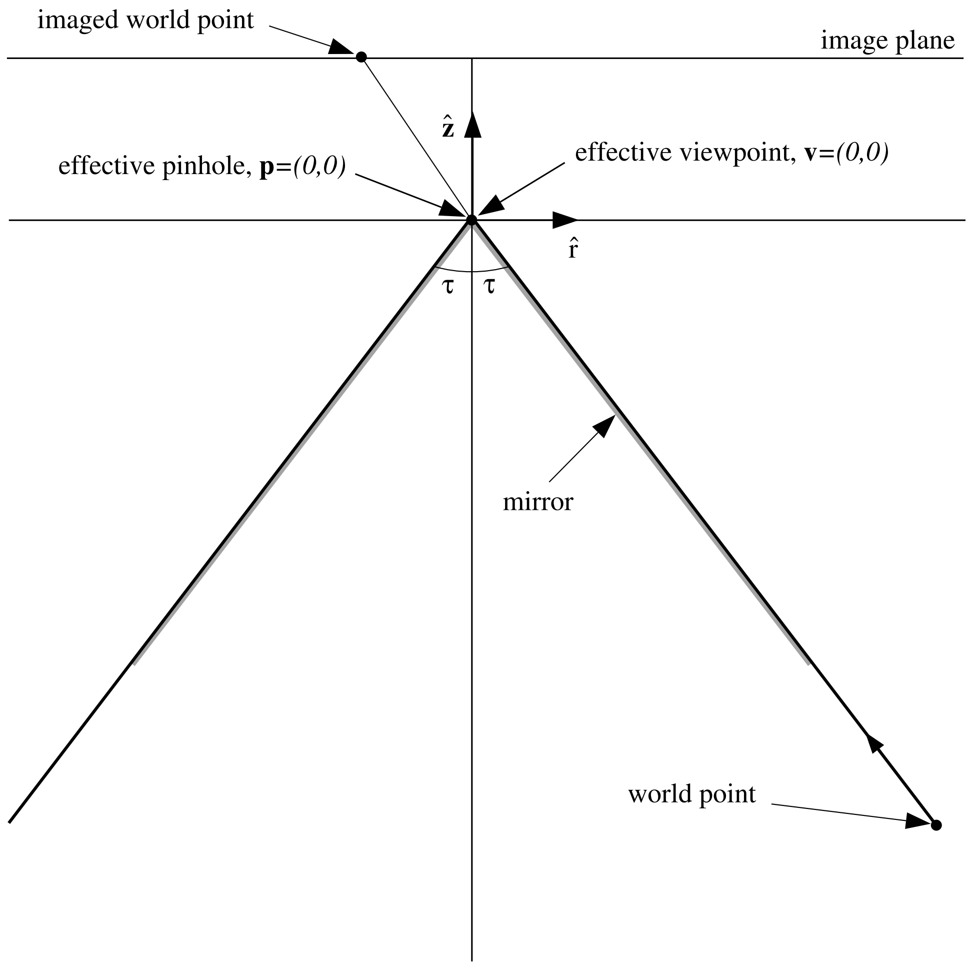


Figure 3: The conical mirror is a solution of the fixed viewpoint constraint equation. Since the pinhole is located at the apex of the cone, this is a degenerate solution of little practical value. If the pinhole is moved away from the apex of the cone (along the axis of the cone), the viewpoint is no 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. Further, if $\tau > 60°$, the circular locus lies inside (below) the cone, if $\tau < 60°$ the circular locus lies outside (above) the cone, and if $\tau = 60°$ the circular locus lies on the cone.

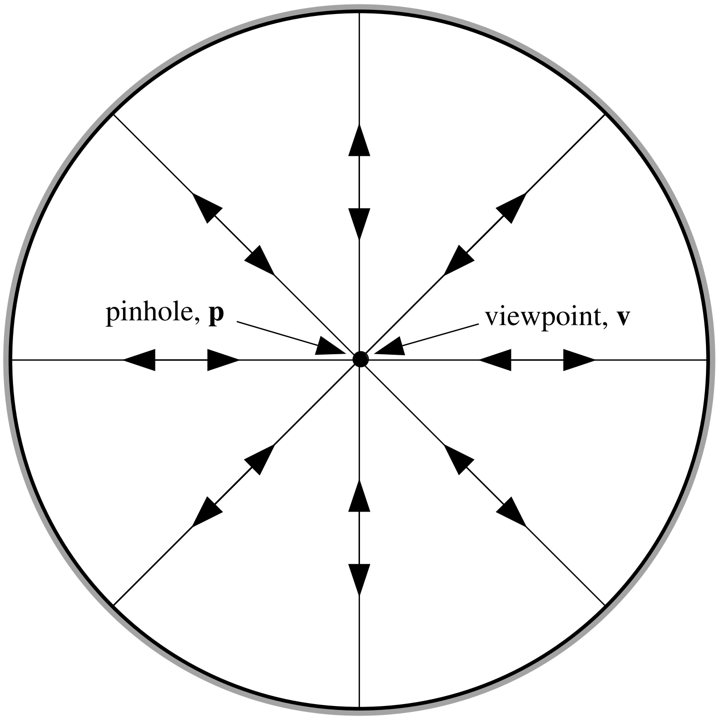


Figure 4: The spherical mirror satisfies the fixed viewpoint constraint when the pinhole lies at the center of the sphere. (Since $c = 0$ the viewpoint also lies at the center of the sphere.) Like the conical mirror, the sphere is of little practical value because the observer can only see itself; rays 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.

The sphere has been used to enhance the field of view for landmark navigation [Hong, 1991]. In this implementation, 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, but it is a quite complex function of the distance between the center of the sphere and the pinhole. The locus is of comparable size to the mirror. Spheres have also been used in stereo applications [Nayar, 1988], but as described before multiple viewpoints are a requirement for stereo and hence the fixed viewpoint constraint is not critical.

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. 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 $z \geq 0$. The ellipsoid has never before been proposed as a way of enhancing the field of view of a video camera.

2.3.5 Hyperboloidal Mirrors

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

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

where:

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

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. Yamazawa *et al.* [Yamazawa *et al.*, 1993] [Yamazawa *et al.*, 1995] recognized that the hyperboloid is indeed a practical solution. They have implemented a sensor designed for autonomous navigation and demonstrated the generation of perspective images.

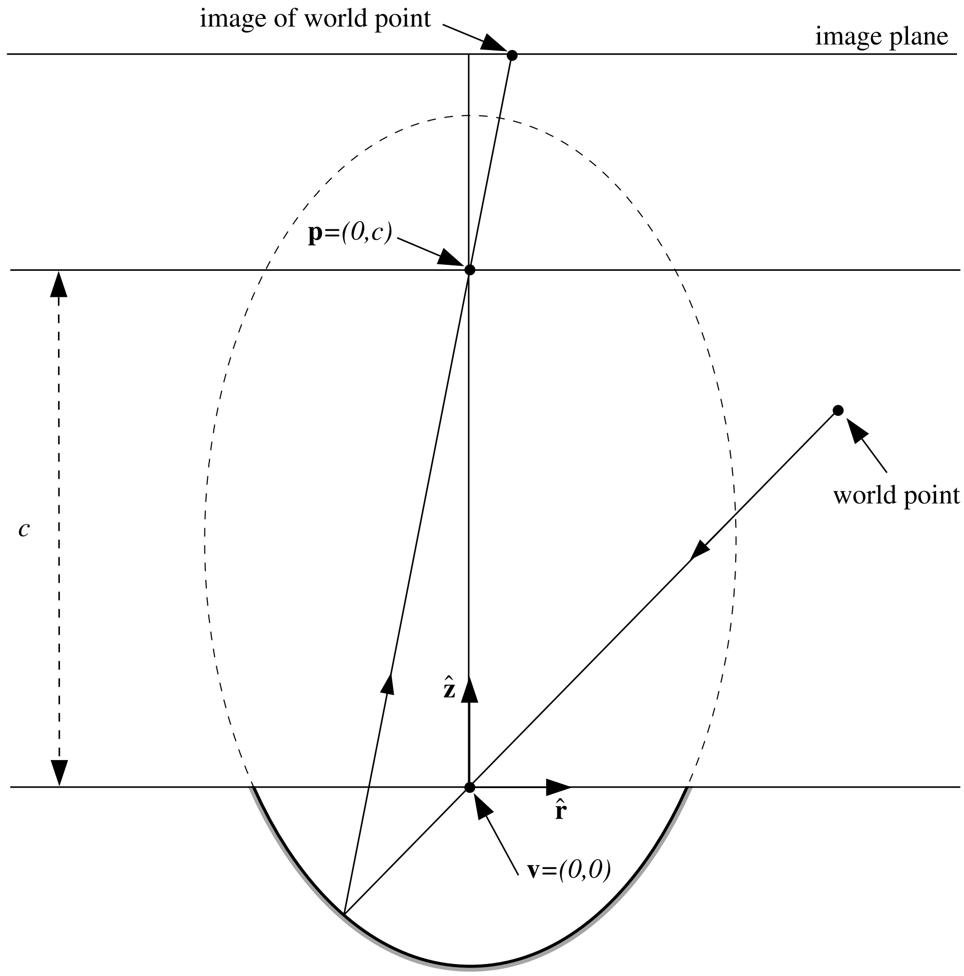


Figure 5: The ellipsoidal mirror satisfies the fixed viewpoint constraint when the pinhole and viewpoint are located at the two foci of the ellipsoid. If the ellipsoid is terminated by the horizontal plane passing through the viewpoint $z = 0$, the field of view is the entire upper hemisphere $z > 0$. (It is also possible to cut the ellipsoid with other planes passing through $\mathbf{v}$, but it appears there is to be little to be gained by doing so.) The ellipsoid has never before been proposed as a way of enhancing the field of view of a video camera.

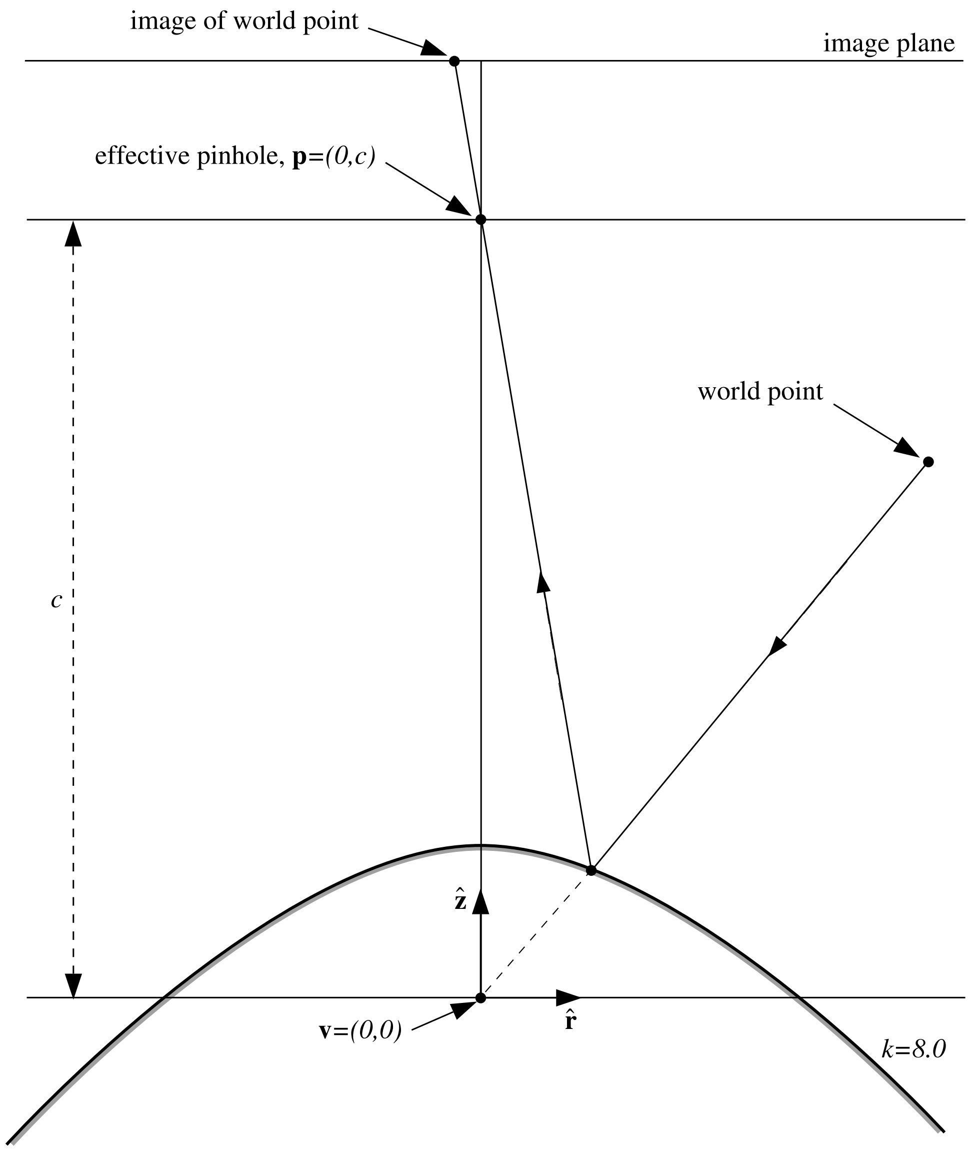


Figure 6: The hyperboloidal mirror satisfies the fixed viewpoint constraint when the pinhole and the viewpoint are located at the two foci of the hyperboloid. This solution does produce the desired increase in field of view. The curvature of the mirror and hence the field of view increase with k. In the limit $k \to 2$, the hyperboloid flattens to the planar mirror of Section 2.3.1. Yamazawa *et al.* [Yamazawa *et al.*, 1993] [Yamazawa *et al.*, 1995] recognized that the hyperboloid is indeed a practical solution and have implemented a sensor designed for autonomous navigation.

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 Figure 7 for an illustration of this scenario. Then, the definition of resolution which 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 sensor is:

$$\frac{\mathrm{d}A}{\mathrm{d}\nu}. \tag{30}$$

(Note that the resolution $\frac{\mathrm{d}A}{\mathrm{d}\nu}$ will vary as a function of the point on the image plane at the center of infinitesimal area $\mathrm{d}A$.)

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 7), the solid angle subtended by the infinitesimal area $\mathrm{d}A$ at the pinhole is:

$$\mathrm{d}\omega \;=\; \frac{\mathrm{d}A \cdot \cos\psi}{u^2 / \cos^2\psi} \;=\; \frac{\mathrm{d}A \cdot \cos^3\psi}{u^2}. \tag{31}$$

Therefore, the resolution of the conventional camera is:

$$\frac{\mathrm{d}A}{\mathrm{d}\omega} \;=\; \frac{u^2}{\cos^3\psi}. \tag{32}$$

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

$$\mathrm{d}S \;=\; \frac{\mathrm{d}\omega \cdot (c-z)^2}{\cos\phi \cos^2\psi} \;=\; \frac{\mathrm{d}A \cdot (c-z)^2 \cdot \cos\psi}{u^2 \cos\phi} \tag{33}$$

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:

$$\mathrm{d}\nu \;=\; \frac{\mathrm{d}S \cdot \cos\phi}{r^2 + z^2} \;=\; \frac{\mathrm{d}A \cdot (c-z)^2 \cdot \cos\psi}{u^2 (r^2 + z^2)}. \tag{34}$$

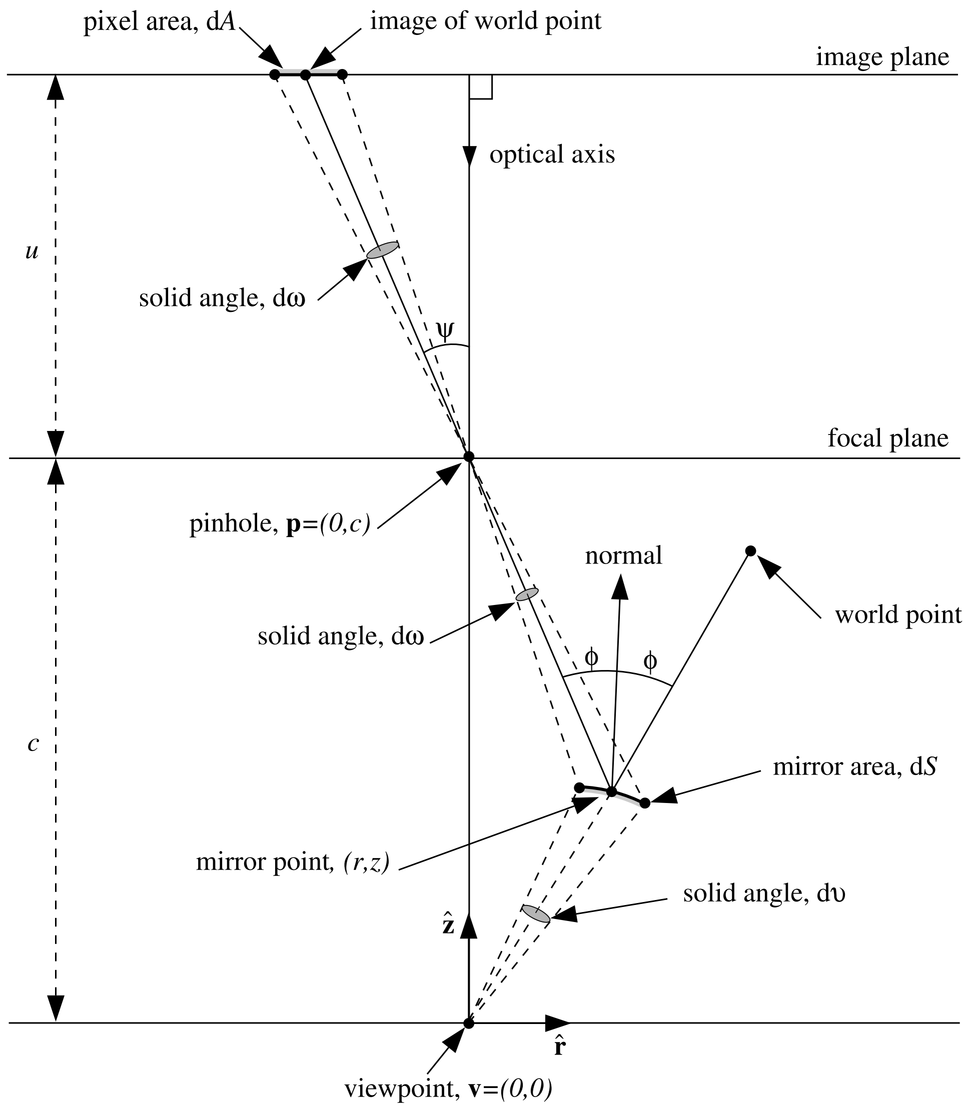


Figure 7: The geometry used to derive the spatial resolution of a catadioptric sensor. Assuming the conventional sensor has a frontal image plane which is located at a distance u from the pinhole and the optical axis is aligned with the z-axis $\hat{\mathbf{z}}$, the spatial resolution of the conventional sensor is $\frac{\mathrm{d}A}{\mathrm{d}\omega} = \frac{u^2}{\cos^3\psi}$. 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 sensor 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 $\cos^2\psi = \frac{(c-z)^2}{(c-z)^2+r^2}$.

Hence, the resolution of the catadioptric camera is:

$$\frac{\mathrm{d}A}{\mathrm{d}\nu} = \frac{u^2(r^2+z^2)}{(c-z)^2 \cdot \cos\psi} = \left[\frac{(r^2+z^2)\cos^2\psi}{(c-z)^2}\right]\frac{\mathrm{d}A}{\mathrm{d}\omega}. \tag{35}$$

But, since:

$$\cos^2\psi = \frac{(c-z)^2}{(c-z)^2+r^2} \tag{36}$$

we have:

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

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

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

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

The first thing to note from Equation (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. Naturally, the resolution in both cases does vary across the image plane like $\frac{1}{\cos^3\psi}$, as seen in Equation (32). Finally, Equations (32) and (37) can in principle be used to design image detectors with non-uniform resolution which compensate for the above variation and produce a catadioptric sensor with uniform resolution across the field of view. More generally, any specified variation in resolution could be achieved.

4 Defocus Blur of a Catadioptric Sensor

Two factors combine to cause defocus blur in catadioptric imaging systems: (1) the finite size of the lens, and (2) the curvature of the mirror. In this section we investigate this effect for the hyperboloid mirror. (Generalization to the other mirrors is straightforward.) We proceed by considering a fixed point in the world and a fixed point in the lens. We find the point on the curved mirror which reflects a ray of light from the world point through the lens point. Then, 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.

To perform this analysis 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 $\vec{\mathbf{vp}}$, and the vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal unit vectors in the plane $z = 0$. As before, we assume that the effective pinhole is located at a distance c from the effective viewpoint. Moreover, as in Section 3, 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 z-axis. Finally, 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 8 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 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 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 is known as the *blur region* or *region of confusion* [Hecht and Zajac, 1974].

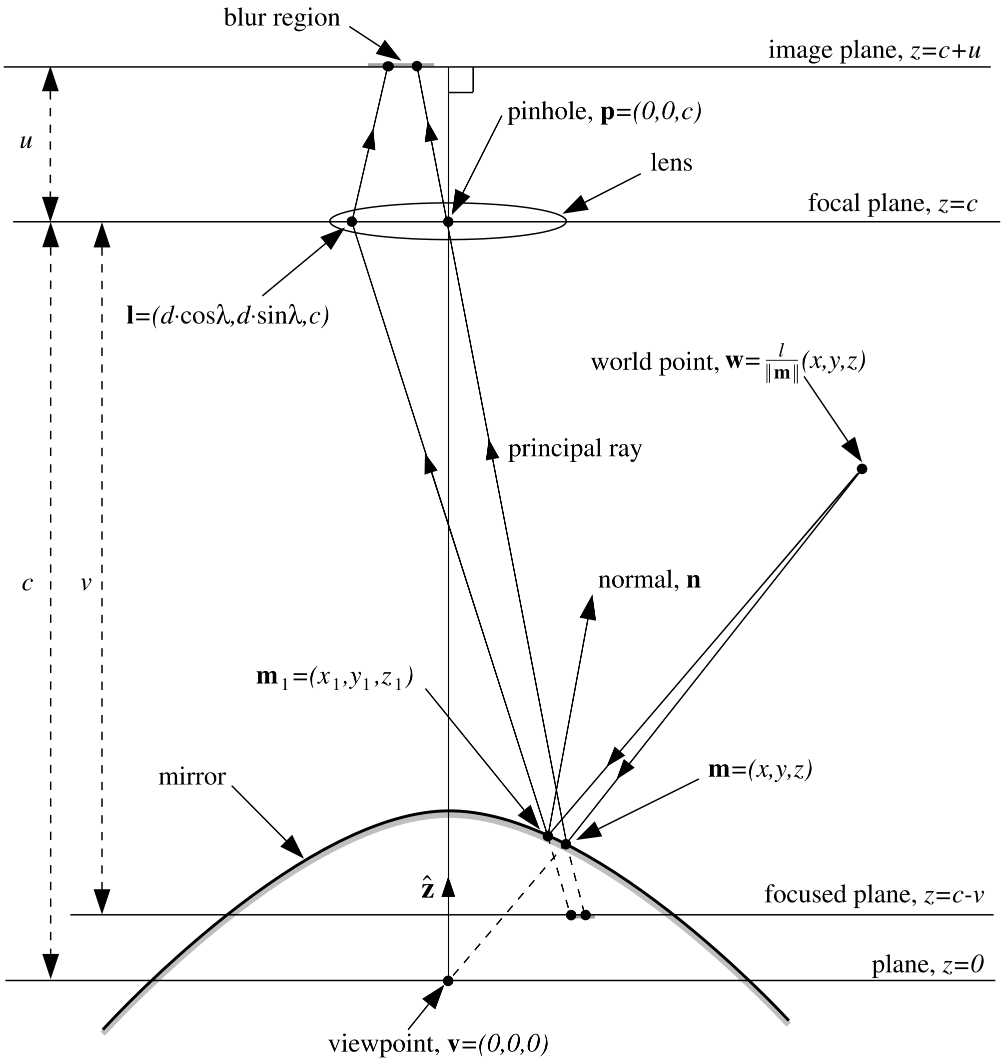


Figure 8: The geometry used to analyze the defocus blur. We work in the 3-D cartesian frame $(\mathbf{v}, \hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal unit vectors in the plane $z = 0$. In addition to the assumptions of Section 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.

For an ideal thin lens, the blur region is circular and so is often referred to as the *blur circle* [Hecht and Zajac, 1974]. As well shall see, for a catadioptric sensor using a curved mirror, the shape of the blur region is not, in general, circular.

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 the plane:

$$z \;=\; c - v \;=\; c - \frac{f \cdot u}{u - f} \tag{39}$$

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:

$$-\frac{u}{v}\left(\mathbf{l} + \frac{v}{c - z_1}(\mathbf{m}_1 - \mathbf{l})\right) \tag{40}$$

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 we have:

$$\left(z_1 - \frac{c}{2}\right)^2 - \left(x_1^2 + y_1^2\right)\left(\frac{k}{2} - 1\right) \;=\; \frac{c^2}{4}\left(\frac{k-2}{k}\right). \tag{41}$$

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:

$$\mathbf{n} \;=\; ([k-2]x_1, [k-2]y_1, c - 2z_1)\,. \tag{42}$$

Hence, the second constraint is:

$$\mathbf{n} \cdot (\mathbf{w} - \mathbf{m}_1) \wedge (\mathbf{l} - \mathbf{m}_1) \;=\; 0. \tag{43}$$

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

$$\frac{\mathbf{n} \cdot (\mathbf{w} - \mathbf{m}_1)}{\|\mathbf{w} - \mathbf{m}_1\|} = \frac{\mathbf{n} \cdot (\mathbf{l} - \mathbf{m}_1)}{\|\mathbf{l} - \mathbf{m}_1\|}. \tag{44}$$

These three constraints on $\mathbf{m}_1$ are all multivariate polynomials in x_1, y_1, and z_1: Equation (41) and Equation (43) are both of order 2, and Equation (44) is of order 5. We were unable to find a closed form solution to these three equations (Equation (44) has 25 terms in general and so it is probable that none exists) but we did investigate numerical solutions. The results are presented in Figures 9 and 10.

For the numerical solutions we set $c = 1$ meter, used the hyperboloid mirror with $k = 4$, and assumed the radius of the lens to be 5 centimeters. We considered the point $\mathbf{m} = (0.125, 0.0, 0.125)$ on the mirror and set the distance from the viewpoint to the world point $\mathbf{w}$ to be $l = 5$ meters. In Figure 9 we plot the area of the blur region (on the ordinate) against the distance to the focused plane v (on the abscissa). The smaller the area of the blur region, the better focused the image will be. We see from Figure 9 that the area never reaches exactly 0, and so an image formed using this catadioptric sensor can never be perfectly focused. (The same is true whenever the mirror is curved.) However, it should be emphasized that the minimum area is very small, and in practice there is no problem focusing the image for a single point.

It is interesting to note that the distance at which the image of the world point will be best focused (i.e. somewhere in the range 1.05–1.2 meters) is much less than the distance from the pinhole to the world point (approximately 1 meter from the pinhole to the mirror and then 5 meters from the mirror to the world point). The reason for this is that the mirror is convex and so tends to increase the divergence of rays coming from the world point.

Finally, we provide an explanation of the fact that there are two minima of the blur

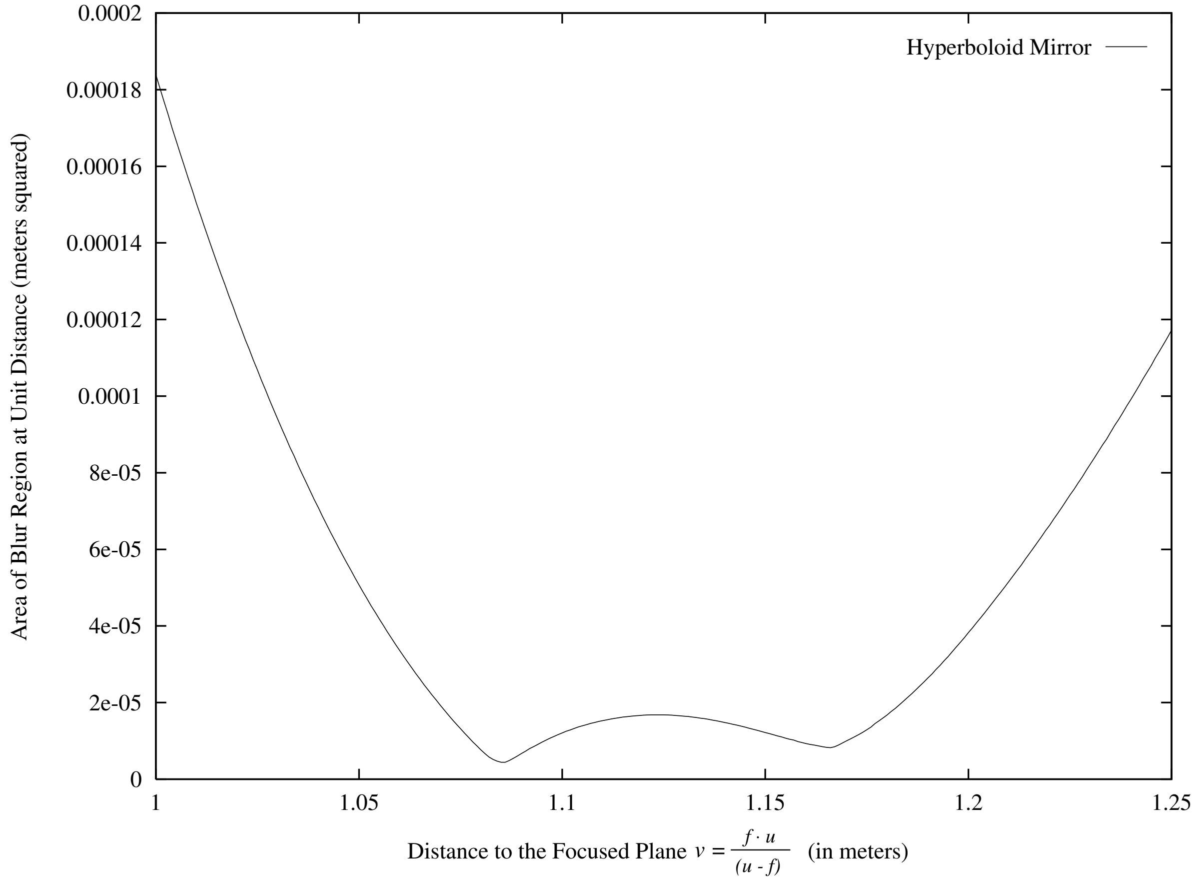
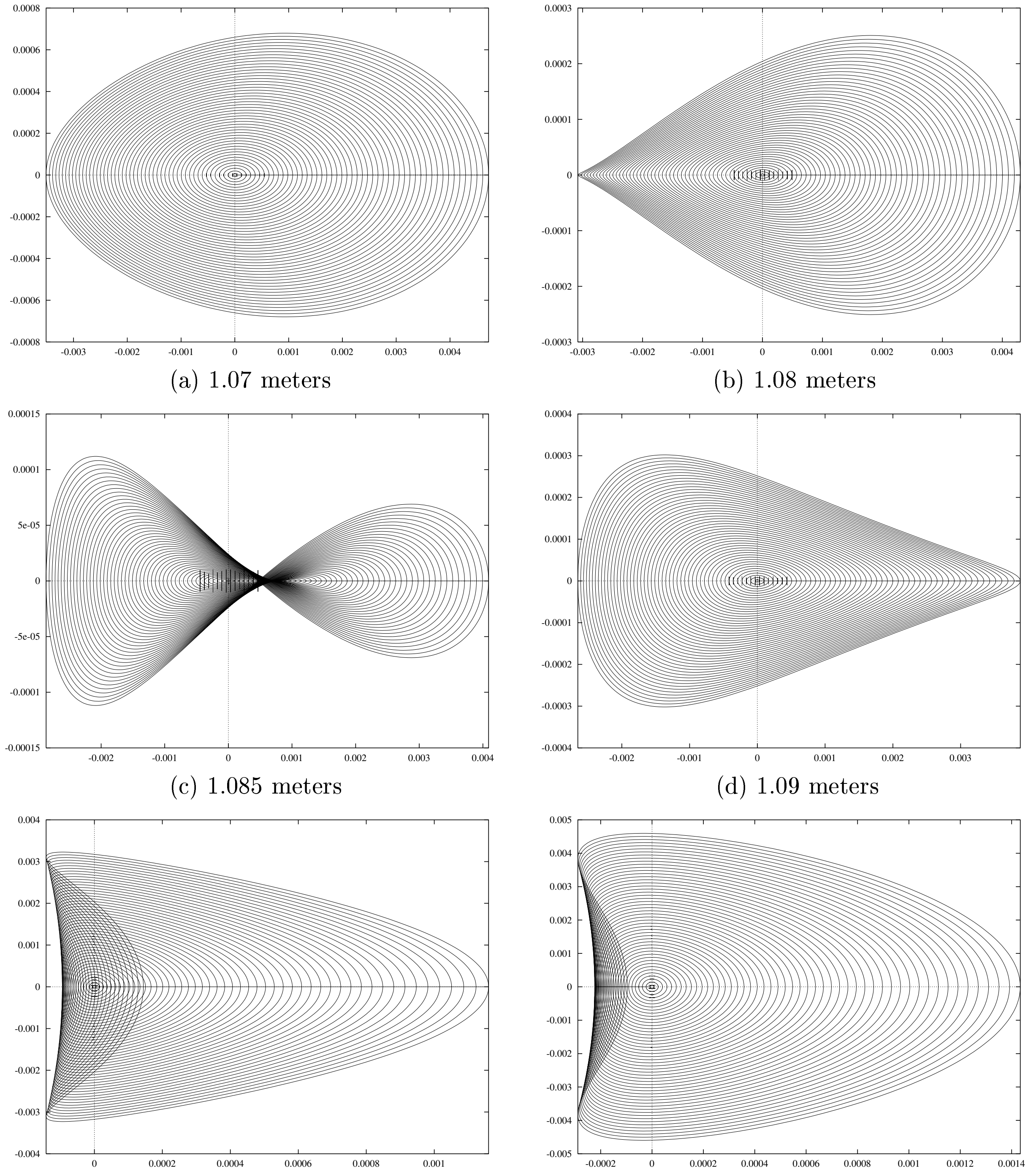


Figure 9: The area of the blur region plotted against the distance to the focused plane $v = fracf \cdot uu - f$ for a point $\mathbf{m} = (0.125, 0.0, 0.125)$ on the hyperboloid mirror with $k = 4$. In this example, we have $c = 1$ meter, the radius of the lens 5 centimeters, and the distance from the viewpoint to the world point $l = 5$ meters. The area never becomes exactly 0 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. There are two minima in the area which correspond to the two different foldings of the blur region illustrated in Figures 10. Also note that the distance at which the image will be best focused (1.05–1.2 meters) is much less than the distance from the pinhole to the world point (approximately 1 meter from the pinhole to the mirror and then 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 coming from the world point $\mathbf{w}$.

area in Figure 9. As mentioned before, for an isolated 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 shrinks until 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. On the other hand, in the catadioptric case the speed with which points move is dependent on their position in a more complex way. Moreover, the direction in which the points are moving is not a constant. This behavior is illustrated in Figures 10(a)–(f).

In Figure 10(a) the focused plane is 1.07 meters from the pinhole and the image is quite defocused. As the focused plane moves to 1.08 meters away in Figure 10(b), the points in the left half of Figure 10(a) are moving upwards more rapidly than those in the right half (the points in the right half are also moving upwards.) Further, the points in the left half are moving upwards more rapidly than they are moving towards the right. This effect continues in Figures 10(c) and (d). In Figure 10(d) all the points are still moving horizontally towards the center of the blur region, but vertically they are now moving away from the center. The points continue to move horizontally towards the center of the blur region, but with those in the left half again moving faster. This causes the blur region to overlap itself as seen in Figure 10(e). Finally, for the focused plane at 1.185 meters in Figure 10(f), all points are moving away from the center of the blur region in both directions. The blur region is expanding and the image becoming more defocused.

Using the numerical techniques of this section it would be simple to compute a lookup table for converting the focus settings of the conventional camera into ones for the catadioptric sensor. The most important property of such a table would be how quickly it changes across the mirror. If it changes too quickly, it may be difficult to focus a large region of the catadioptric image at one time without using sophisticated compensating



(a) 1.07 meters

(b) 1.08 meters

(c) 1.085 meters

(d) 1.09 meters

(e) 1.175 meters

(f) 1.185 meters

Figure 10: The variation in the shape of the blur region as the distance to the focused plane is varied. All of the blur regions in this figure are relatively well focused. Note that the scale of the 6 figures are all different and that the scale on the two axes is different in general.

lenses. Intuitively, the more curved the mirror, the more quickly the conversion table will vary. However, investigating to what extent this is the case is left for a future study.

5 Summary

In this paper we have studied three design criteria of catadioptric sensors: (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 a single camera, found an expression for the resolution of a catadioptric sensor in terms of the resolution of the conventional camera used to construct it, and presented a preliminary analysis of defocus blur.

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.

[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.

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

[Nayar and Baker, 1997] Shree K. Nayar and Simon Baker. Catadioptric image formation. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, New Orleans, 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, 1997] S.K. Nayar. Omnidirectional video camera. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, May 1997. (A related paper will appear in CVPR 97).

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

[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.

[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.