

A Theory of Catadioptric Image Formation *

Simon Baker and Shree K. Nayar

Department of Computer Science
Columbia University
New York, NY 10027

Abstract

Conventional video cameras have limited fields of view which make them restrictive for certain applications in computational vision. A catadioptric 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. In this paper, we derive the complete class of single-lens single-mirror catadioptric sensors which have a single viewpoint and an expression for the spatial resolution of a catadioptric sensor in terms of the resolution of the camera used to construct it. We also include a preliminary analysis of the defocus blur caused by the use of a curved mirror.

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.

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], and [Nayar, 1997]). We refer to the general approach of using mirrors in combination with conventional imaging systems as *catadioptric*¹ image formation. As noted in [Rees, 1970], [Yamazawa *et al.*, 1995], [Nalwa, 1996], and [Nayar and Baker, 1997a], 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 desir-

able is that it permits the generation of geometrically correct perspective images from the image(s) captured by the catadioptric cameras. These perspective images 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.

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.

Another optical property which is modified by the use of a catadioptric system is focusing. It is well known that a curved mirror increases image blur [Hecht and Zajac, 1974], and so in Section 4 we analyze this effect for catadioptric sensors. Two factors combine to cause 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 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

*This work was supported in parts by the DARPA/ONR MURI Grant N00014-95-1-0601, an NSF National Young Investigator Award, and a David and Lucile Packard Fellowship.

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].

settings.

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. In other words, the catadioptric sensor can only sample the 5-D plenoptic function [Adelson and Bergen, 1991] at a single point. The fixed 3-D point at which a catadioptric sensor samples the plenoptic function will be referred to 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 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 $\mathbf{v}\mathbf{p}$. 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 illustration² 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:

$$\tan \theta = \frac{z}{r}. \quad (1)$$

²In 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 perpendicular to $\hat{\mathbf{z}}$.

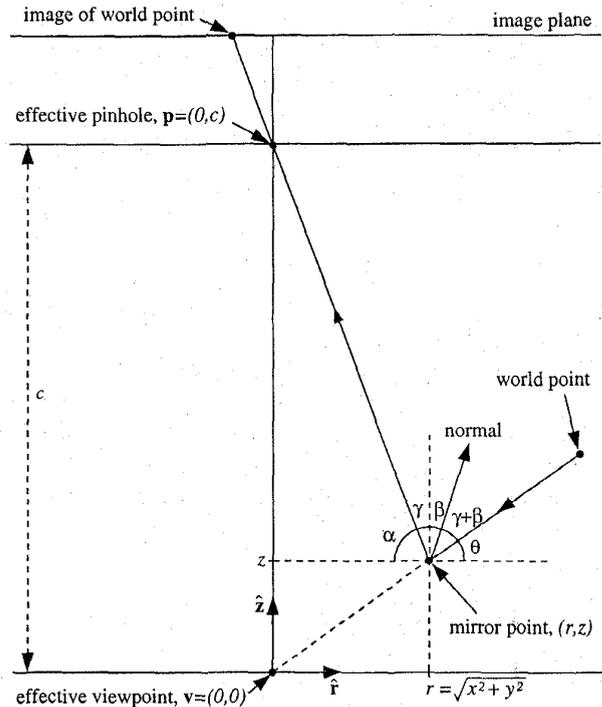


Figure 1: The geometry used to derive the fixed viewpoint constraint equation. The viewpoint $\mathbf{v} = (0, 0)$ is located at the origin of a 2-D coordinate frame $(\mathbf{v}, \hat{\mathbf{r}}, \hat{\mathbf{z}})$, and the pinhole of the camera $\mathbf{p} = (0, c)$ is located at a distance c from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

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

$$\tan \alpha = \frac{c-z}{r} \quad (2)$$

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:

$$\frac{dz}{dr} = -\tan \beta. \quad (3)$$

Our final geometric relationship is due to the fact that we 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 Figure 1 for an illustration of this constraint.) Eliminating γ from these two expressions and rearranging gives $2\beta = \alpha - \theta$. Then, taking the tangent of both sides yields:

$$\frac{2 \tan \beta}{1 - \tan^2 \beta} = \frac{\tan \alpha - \tan \theta}{1 + \tan \alpha \tan \theta}. \quad (4)$$

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

$$r(c-2z) \left(\frac{dz}{dr} \right)^2 - 2(r^2 + cz - z^2) \frac{dz}{dr} + r(2z - c) = 0. \quad (5)$$

2.2 General Solution of the Constraint

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:

$$\frac{dz}{dr} = \frac{(z^2 - r^2 - cz) \pm \sqrt{r^2 c^2 + (z^2 + r^2 - cz)^2}}{r(2z - c)}. \quad (6)$$

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

$$\frac{dy}{dr} = \frac{(y^2 - r^2 - b^2) \pm \sqrt{4r^2 b^2 + (y^2 + r^2 - b^2)^2}}{2ry}. \quad (7)$$

Then, after substituting $2rx = y^2 + r^2 - b^2$ and rearranging we have:

$$\frac{1}{\sqrt{b^2 + x^2}} \frac{dx}{dr} = \pm \frac{1}{r}. \quad (8)$$

Integrating both sides with respect to r results in:

$$\ln \left(x + \sqrt{b^2 + x^2} \right) = \pm \ln r + C \quad (9)$$

where C is the constant of integration. Hence,

$$x + \sqrt{b^2 + x^2} = \frac{k}{2} r^{\pm 1} \quad (10)$$

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:

$$\left(z - \frac{c}{2} \right)^2 - r^2 \left(\frac{k}{2} - 1 \right) = \frac{c^2}{4} \left(\frac{k-2}{k} \right) \quad (k \geq 2) \quad (11)$$

$$\left(z - \frac{c}{2} \right)^2 + r^2 \left(1 + \frac{c^2}{2k} \right) = \left(\frac{2k + c^2}{4} \right) \quad (k > 0). \quad (12)$$

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

Together, the two relations in Equations (11) and (12) 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 every³ conic section is theoretically a solution of one of these 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.

2.3.1 Planar Mirrors

In Solution (11), if we set $k = 2$ and $c > 0$, we get the cross-section of a planar mirror:

$$z = \frac{c}{2}. \quad (13)$$

This plane is the one which bisects the line segment $\vec{v}\vec{p}$ joining the viewpoint and the pinhole. Moreover, by examining the other solutions it follows that the perpendicular bisector of $\vec{v}\vec{p}$ is the only 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 single conventional imaging system, while still obeying 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 [1996]. To ensure a single viewpoint while using multiple planar mirrors, 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 $360^\circ \times 50^\circ$. The panoramic image is of relatively high resolution since it is generated from the four images captured by the four cameras. Nalwa's sensor is

³There is one conic section which is an exception: the parabola. Although the parabola is not a solution of either Equation (11) or Equation (12) for finite values of c and k , it is a solution of Equation (11) in the limit that $c \rightarrow \infty$, $k \rightarrow \infty$, and $c/k = h$, a constant. As shown in [Nayar and Baker, 1997a], 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 [Nayar, 1997].

straightforward to implement, but requires four of each component: i.e. four cameras and four lenses.

2.3.2 Conical Mirrors

In Solution (11), if we set $c = 0$ and $k \geq 2$, we get a conical mirror with circular cross section:

$$z = \sqrt{\frac{k-2}{2}} r^2. \quad (14)$$

See Figure 2 for an illustration of this solution. The angle at the apex of the cone is 2τ where $\tan \tau = \sqrt{2/(k-2)}$. 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 2.) Hence, the cone (with the pinhole at the vertex) is a degenerate solution of no practical value.

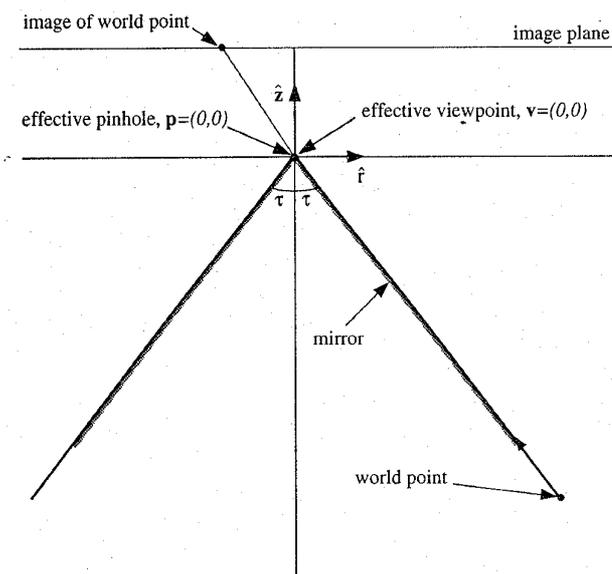


Figure 2: The conical mirror is a solution of the fixed viewpoint constraint equation. Since the pinhole is located at the apex of the cone, this is a degenerate solution of little practical value. If the pinhole is moved away from the apex of the cone (along the axis of the cone), the viewpoint is no 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 $\tau > 60^\circ$, 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.

The cone has been used for wide-angle imaging a number of times [Yagi and Kawato, 1990] [Yagi and Yachida, 1991] [Bogner, 1995]. 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 $e \cdot \cos 2\tau$. 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 (12), if we set $c = 0$ and $k > 0$, we get the spherical mirror:

$$z^2 + r^2 = \frac{k}{2}. \quad (15)$$

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 only sees itself.

The sphere has also been used to enhance the field of view several times [Hong, 1991] [Bogner, 1995] [Murphy, 1995]. In these implementations, the pinhole is placed outside the sphere and so there is no single effective viewpoint. The locus of the effective viewpoint can be computed in a straightforward manner using a symbolic mathematics package, but it is a quite complex function of the distance between the center of the sphere and the pinhole. Spheres have also been used in stereo applications [Nayar, 1988], but as described before multiple viewpoints are a requirement for stereo.

2.3.4 Ellipsoidal Mirrors

In Solution (12), when $k > 0$ and $c > 0$, we get the ellipsoidal mirror:

$$\frac{1}{a_e^2} \left(z - \frac{c}{2} \right)^2 + \frac{1}{b_e^2} r^2 = 1 \quad (16)$$

where:

$$a_e = \sqrt{\frac{2k+c^2}{4}} \quad \text{and} \quad b_e = \sqrt{\frac{k}{2}}. \quad (17)$$

The ellipsoid is the first solution that can actually be used to enhance the field of view. As shown in Figure 3, 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 $z \geq 0$. It is also possible to cut the ellipsoid with other planes passing through \mathbf{v} , but it appears there is little to be gained by doing so.

2.3.5 Hyperboloidal Mirrors

In Solution (11), when $k > 2$ and $c > 0$, we get the hyperboloidal mirror:

$$\frac{1}{a_h^2} \left(z - \frac{c}{2} \right)^2 - \frac{1}{b_h^2} r^2 = 1 \quad (18)$$

where:

$$a_h = \frac{c}{2} \sqrt{\frac{k-2}{k}} \quad \text{and} \quad b_h = \frac{c}{2} \sqrt{\frac{2}{k}}. \quad (19)$$

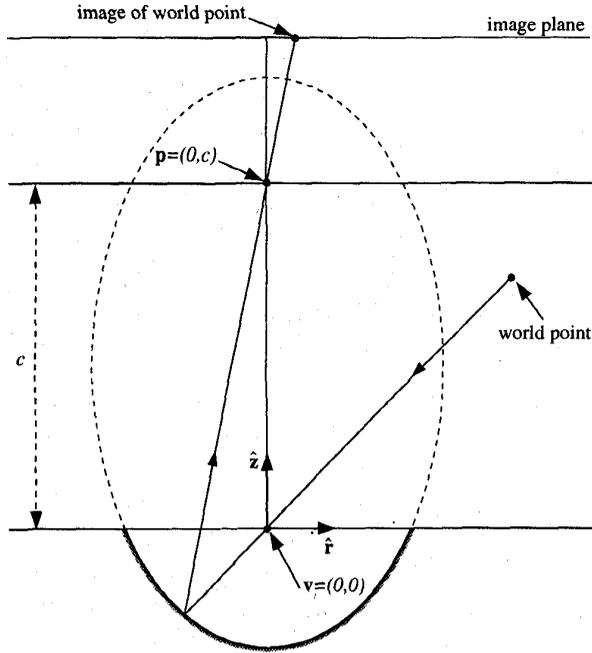


Figure 3: The ellipsoidal mirror satisfies the fixed viewpoint constraint when the pinhole and viewpoint are located at the two foci of the ellipsoid. If the ellipsoid is terminated by the horizontal plane passing through the viewpoint $z = 0$, the field of view is the entire upper hemisphere $z > 0$.

As seen in Figure 4, 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 $k \rightarrow 2$, 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.

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 5 for an illustration of this scenario. Then, the definition of resolution which 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

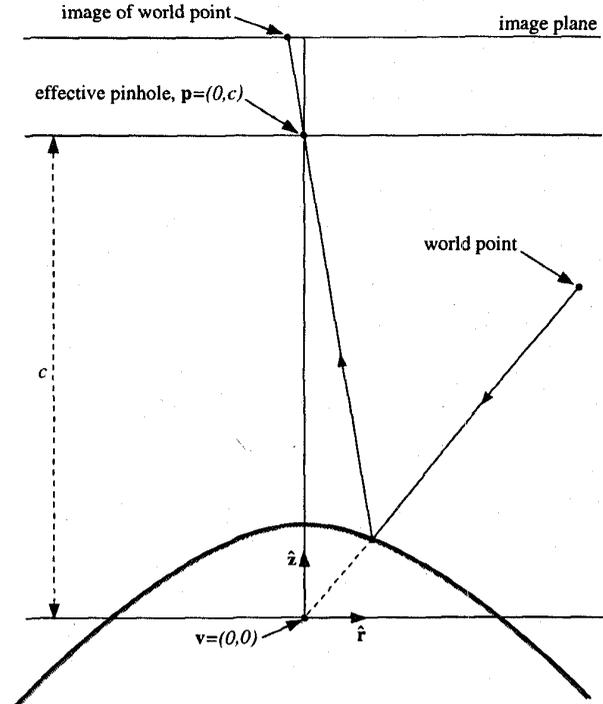


Figure 4: The hyperboloidal mirror satisfies the fixed viewpoint constraint when the pinhole and the viewpoint are located at the two foci of the hyperboloid. This solution does produce the desired increase in field of view.

dA) is:

$$\frac{dA}{d\nu} \quad (20)$$

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 Figure 5), the solid angle subtended by the infinitesimal area dA at the pinhole is:

$$d\omega = \frac{dA \cdot \cos \psi}{u^2 / \cos^2 \psi} = \frac{dA \cdot \cos^3 \psi}{u^2} \quad (21)$$

Therefore, the resolution of the conventional camera is:

$$\frac{dA}{d\omega} = \frac{u^2}{\cos^3 \psi} \quad (22)$$

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

$$dS = \frac{d\omega \cdot (c - z)^2}{\cos \phi \cos^2 \psi} = \frac{dA \cdot (c - z)^2 \cdot \cos \psi}{u^2 \cos \phi} \quad (23)$$

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,

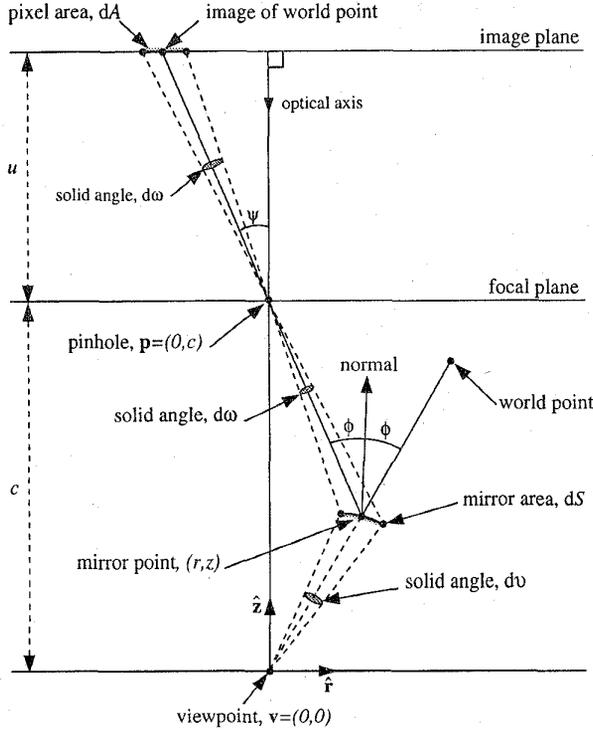


Figure 5: The geometry used to derive the spatial resolution of a catadioptric sensor. We assume that the conventional sensor has a frontal image plane which is located at a distance u from the pinhole and that the optical axis is aligned with the z -axis \hat{z} .

the solid angle of the world imaged by the catadioptric camera is:

$$d\nu = \frac{dS \cdot \cos \phi}{r^2 + z^2} = \frac{dA \cdot (c-z)^2 \cdot \cos \psi}{u^2(r^2 + z^2)}. \quad (24)$$

Hence, the resolution of the catadioptric camera is:

$$\frac{dA}{d\nu} = \frac{u^2(r^2 + z^2)}{(c-z)^2 \cdot \cos \psi} = \left[\frac{(r^2 + z^2) \cos^2 \psi}{(c-z)^2} \right] \frac{dA}{d\omega}. \quad (25)$$

But, since:

$$\cos^2 \psi = \frac{(c-z)^2}{(c-z)^2 + r^2} \quad (26)$$

we have:

$$\frac{dA}{d\nu} = \left[\frac{r^2 + z^2}{(c-z)^2 + r^2} \right] \frac{dA}{d\omega}. \quad (27)$$

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

$$\frac{r^2 + z^2}{(c-z)^2 + r^2} \quad (28)$$

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

The first thing to note from Equation (27) 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 Equation (28) is the square of the distance from the point (r, z) to the effective viewpoint \mathbf{v} divided by the square of the distance to the pinhole \mathbf{p} . Using simple properties of ellipsoids and hyperboloids it follows that for these two mirror shapes, the factor in Equation (28) increases with r . Hence both hyperboloidal and ellipsoidal catadioptric sensors will have highest resolution around the periphery, a useful property for certain applications such as teleconferencing.

4 Defocus Blur of a Catadioptric Sensor

Two factors combine to cause defocus blur in catadioptric imaging systems (that is, in addition to the normal causes in conventional dioptric systems, such as diffraction and lens aberrations): (1) the finite size of the lens aperture, 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 $(\mathbf{v}, \hat{x}, \hat{y}, \hat{z})$ where \mathbf{v} is the location of the effective viewpoint, \mathbf{p} is the location of the effective pinhole, \hat{z} is a unit vector in the direction $\mathbf{v}\mathbf{p}$, and the vectors \hat{x} and \hat{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 6 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

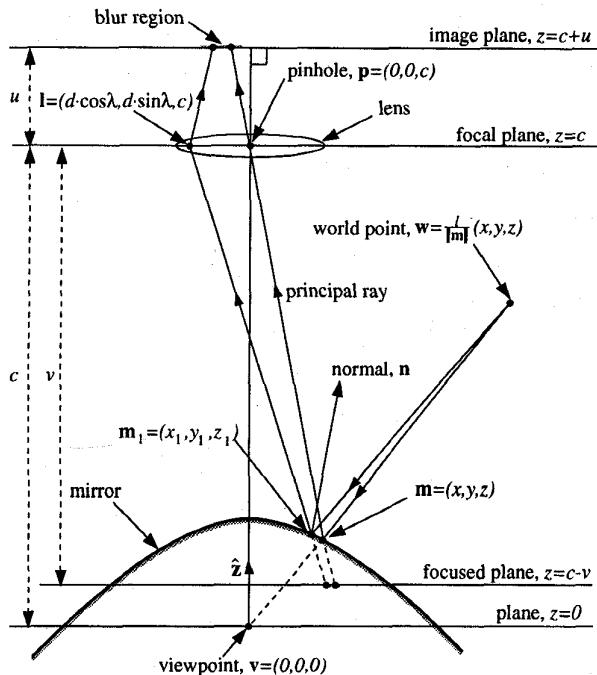


Figure 6: The geometry used to analyze the defocus blur. We work in the 3-D cartesian frame $(v, \hat{x}, \hat{y}, \hat{z})$ where \hat{x} and \hat{y} are orthogonal unit vectors in the plane $z = 0$. In addition to the assumptions of Section 3, we also assume that the effective pinhole is located at the center of the lens and that the lens has a circular aperture.

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 l and the image plane as l varies over the lens is known as the *blur region* or *region of confusion* [Hecht and Zajac, 1974]. 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 is shown in [Nayar and Baker, 1997b], for a catadioptric sensor the shape of the blur region is not, in general, circular.

If we know the points m_1 and l , we can find the point on the image plane where the ray of light through these points is imaged. First, the line through m_1 in the direction $l - m_1$ is extended to intersect the *focused plane*. By the thin lens law [Hecht and Zajac, 1974] the focused plane is:

$$z = c - v = c - \frac{f \cdot u}{u - f} \quad (29)$$

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 l and the image plane are the x and

y coordinates of:

$$-\frac{u}{v} \left(1 + \frac{v}{c - z_1} (\mathbf{m}_1 - 1) \right) \quad (30)$$

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

Given the lens point $l = (d \cdot \cos \lambda, d \cdot \sin \lambda, c)$ and the world point $w = \frac{l}{\|m_1\|} (x, y, z)$, there are three constraints on the point $m_1 = (x_1, y_1, z_1)$. First, m_1 must lie on the mirror and so we have:

$$\left(z_1 - \frac{c}{2} \right)^2 - (x_1^2 + y_1^2) \left(\frac{k}{2} - 1 \right) = \frac{c^2}{4} \left(\frac{k - 2}{k} \right). \quad (31)$$

Secondly, the incident ray $(w - m_1)$, the reflected ray $(m_1 - l)$, and the normal to the mirror at m_1 must lie in the same plane. The normal to the mirror at m_1 lies in the direction:

$$\mathbf{n} = ([k - 2]x_1, [k - 2]y_1, c - 2z_1). \quad (32)$$

Hence, the second constraint is:

$$\mathbf{n} \cdot (w - m_1) \wedge (l - m_1) = 0. \quad (33)$$

Finally, the angle of incidence must equal the angle of reflection and so the third constraint on the point m_1 is:

$$\frac{\mathbf{n} \cdot (w - m_1)}{\|w - m_1\|} = \frac{\mathbf{n} \cdot (l - m_1)}{\|l - m_1\|}. \quad (34)$$

These three constraints on m_1 are all multivariate polynomials in x_1 , y_1 , and z_1 : Equation (31) and Equation (33) are both of order 2, and Equation (34) is of order 5. We were unable to find a closed form solution to these three equations (Equation (34) has 25 terms in general and so it is probable that none exists) but we did investigate numerical solutions. Some of the results are presented in Figure 7. (The interested reader is referred to [Nayar and Baker, 1997b] for a more complete presentation of the results, including an explanation of the 2 local minima in Figure 7.)

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 $m = (0.125, 0.0, 0.125)$ on the mirror and set the distance from the viewpoint to the world point w to be $l = 5$ meters. In Figure 7 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 7 that the area never reaches exactly 0, and so an image formed using this catadioptric sensor can never be perfectly focused. 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 world point. Moreover, it is possible to use additional corrective lenses to compensate for most of this effect [Hecht and Zajac, 1974].

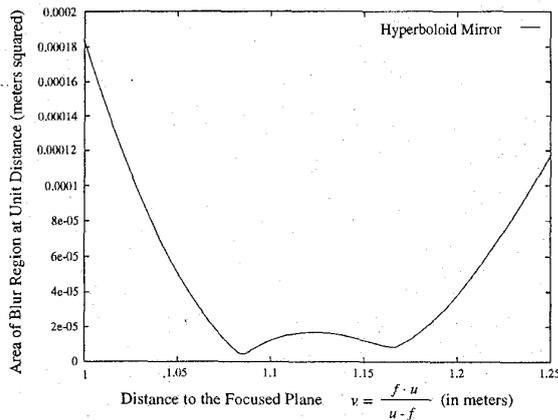


Figure 7: The area of the blur region plotted against the distance to the focused plane $v = \frac{f \cdot u}{u-f}$ for a point $m = (0.125, 0.0, 0.125)$ on the hyperboloid mirror with $k = 4$. In this example, we have $c = 1$ meter, the radius of the lens 5 centimeters, and the distance from the viewpoint 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 world point is not a problem in practice.

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 phenomenon is that the mirror is convex and so tends to increase the divergence of rays emanating from the world point.

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.
- [Bogner, 1995] S. Bogner. Introduction to panoramic imaging. In *Proceedings of the IEEE SMC Conference*, pages 3100–3106, October 1995.
- [Charles et al., 1987] J.R. Charles, R. Reeves, and C. Schur. How to build and use an all-sky camera. *Astronomy Magazine*, April 1987.
- [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.
- [Murphy, 1995] J.R. Murphy. Application of panoramic imaging to a teleoperated lunar rover. In *Proceedings of the IEEE SMC Conference*, October 1995.
- [Nalwa, 1996] V.S. Nalwa. A true omnidirectional viewer. Technical report, Bell Laboratories, Holmdel, NJ 07733, USA, February 1996.
- [Nayar and Baker, 1997a] S.K. Nayar and S. Baker. Catadioptric image formation. In *Proceedings of the 1997 DARPA Image Understanding Workshop*, pages 1431–1437, New Orleans, May 1997.
- [Nayar and Baker, 1997b] S.K. Nayar and S. Baker. A theory of catadioptric image formation. Technical Report CUCS-015-97, Department of Computer Science, Columbia University, USA, April 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. Catadioptric omnidirectional camera. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*, pages 482–488, June 1997.
- [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.
- [Rees, 1970] D.W. Rees. Panoramic television viewing system. United States Patent No. 3,505,465, April 1970.
- [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.