Appeared in the Proceedings of the 1997 DARPA Image Understanding Workshop, May, 1997

Catadioptric Image Formation *

Shree K. Nayar and Simon Baker
Department of Computer Science, Columbia University
New York, New York 10027
Email: {nayar,simonb}@cs.columbia.edu

Abstract

Conventional video cameras have limited fields of view that make them restrictive in certain vision applications. A catadioptric sensor uses a combination of lenses and mirrors placed in a carefully designed configuration to capture a much wider field of view. In particular, the shape of the mirror must be selected to ensure that the complete catadioptric system has a single effective viewpoint, which is a requirement for the generation of pure perspective images from the sensed image. In this paper, we derive and analyze the complete class of single-lens single-mirror catadioptric sensors which satisfy the fixed viewpoint constraint. Some solutions turn out to be degenerate with no practical value while other solutions lead to realizable sensors.

1 Introduction

Conventional imaging systems are limited in their fields of view. An effective way to enhance the field of view is to use mirrors in conjunction with lenses. This approach to image formation is fast gaining in popularity (see [Nayar, 1988], [Yagi and Kawato, 1990], [Hong, 1991], [Goshtasby and Gruver, 1993], [Yamazawa *et al.*, 1993], [Nalwa, 1996] [Nayar, 1997]). We refer to the general approach of incorporating mirrors into conventional imaging systems as *catadioptric*[1] image formation. Our recent work in this context has led to the development of a truly omnidirectional video camera with a spherical field of view [Nayar, 1997].

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

As recently noted in [Yamazawa *et al.*, 1993], [Nalwa, 1996] and [Nayar, 1997], it is highly desirable that the catadioptric system (or any imaging system, for that matter) have a single center of projection (viewpoint). A single viewpoint permits the creation of pure perspective images from the image sensed by the catadioptric system. This is done by mapping sensed brightness values onto a plane placed at any distance (effective focal length) from the viewpoint. Any image computed in this manner preserves linear perspective geometry. For instance, straight lines in the scene produce straight lines in the computed image. Images that adhere to perspective projection are desirable from two standpoints; they are consistent with the way we are used to seeing images, and they lend themselves to further processing by the large body of work in computational vision that assumes linear perspective projection. When the catadioptric system is omnidirectional in its field of view, the single viewpoint permits the construction of not only perspective but also panoramic images.

In this paper, we derive the complete set of catadioptric systems with a single effective viewpoint and which are constructed from a single conventional lens and a single mirror. As we will show, the class 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 realizable sensors. During our analysis we will stop at many points to evaluate the merits of the solutions as well as the merits of catadioptric sensors proposed in the literature.

2 General Solution

Let the final (dioptric) stage of our sensor be a conventional perspective lens. In Figure 1, the effective pinhole of the lens is $\mathbf{p}$. We formulate the catadioptic image formation problem as follows:

Find the class of reflecting surfaces that, when used in conjunction with a perspective lens, produce an image of the world as seen from a fixed viewpoint. Let us assume that the fixed viewpoint $\mathbf{v}$ is at the origin of the coordinate frame (see Figure 1) and the center $\mathbf{p}$ of the perspective lens is located on the vertical axis at a distance c from $\mathbf{v}$.

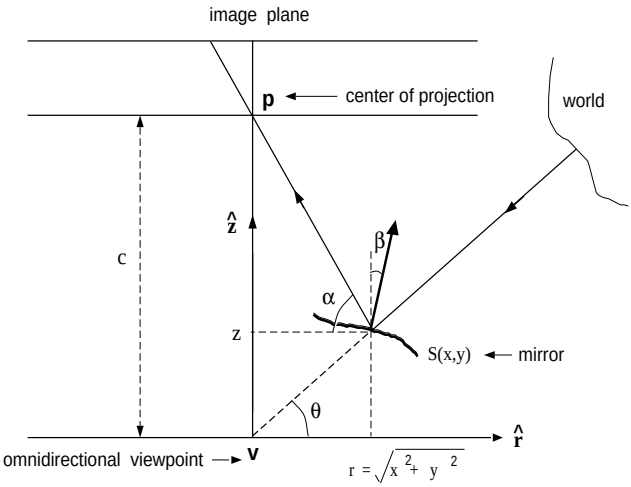

Figure 1: Geometry used to derive the reflecting surface that produces an image of the world as seen from a fixed viewpoint $\mathbf{v}$. This image is captured using a conventional perspective camera with an effective pinhole $\mathbf{p}$.

For the fixed viewpoint constraint to hold, each world point seen from $\mathbf{v}$ must be reflected by a point on the mirror surface $S(x, y)$ towards $\mathbf{p}$. Note that since perspective projection is rotationally symmetric about the optical axis $\hat{\mathbf{z}}$, the mirror can be assumed to be a surface of revolution around $\hat{\mathbf{z}}$. Therefore, it suffices to find the one-dimensional profile $z(r) = S(x, y)$, where $r = \sqrt{x^2 + y^2}$.

In fact, the viewpoint $\mathbf{v}$ and the determined profile $z(r)$ are in no way restricted to the optical axis. Since perspective projection is rotationally symmetric with respect to any ray that passes through the pinhole $\mathbf{p}$, the viewpoint and the profile could be moved from the optical axis by keeping the distance c the same and aligning the symmetry axis of the profile with the ray that passes through the viewpoint and the pinhole. Of course, in this case, the image plane shown in Figure 1 would be non-frontal. This does not pose any additional ambiguity as the mapping from any non-frontal image plane to a frontal image plane is one-to-one.

With the above generalizations in place, we are ready to derive the profile of the reflecting surface. The relation between the angle θ of the incoming ray and the reflecting surface is

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

The angle α made by the reflected ray with the horizontal axis is given by

$$\tan\alpha \;=\; \frac{(c-z)}{r}\,. \tag{2}$$

Let the surface slope at the point of reflection be defined by the angle β made by the surface normal with the vertical axis:

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

This allows us to write

$$\tan 2\beta \;=\; \frac{2\tan\beta}{1-\tan^2\beta} \;=\; \frac{-2\frac{dz}{dr}}{1-(\frac{dz}{dr})^2}\,. \tag{4}$$

Since the surface is *specular*, the angles of incidence and reflection are equal. Consequently,

$$\beta \;=\; \frac{(\alpha-\theta)}{2}\,, \tag{5}$$

which gives us

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

In the right hand side of the above expression, we substitute (1) and (2) and equate with the right hand side of (4) to get

$$\frac{-2\frac{dz}{dr}}{1-(\frac{dz}{dr})^2} \;=\; \frac{(c-2z)\,r}{(r^2+c\,z-z^2)}\,. \tag{7}$$

Thus, we find that the reflecting surface must satisfy the above quadratic first-order differential equation. It is straightforward to solve the quadratic for surface slope:

$$\frac{dz}{dr} \;=\; \frac{(z^2-r^2-c\,z)\pm\sqrt{r^2c^2+(z^2+r^2-cz)^2}}{r\,(2z-c)} \tag{8}$$

Next, we substitute $y = z - c/2$ and set $b = c/2$ which yields

$$\frac{dy}{dr} = \frac{(y^2-r^2-b^2)\pm\sqrt{4r^2b^2+(y^2+r^2-b^2)^2}}{2ry}. \tag{9}$$

Then, substituting $2rx = y^2 + r^2 - b^2$, we get

$$\frac{1}{\sqrt{b^2 + x^2}} \frac{dx}{dr} \;=\; \pm \frac{1}{r}. \qquad (10)$$

Integrating both sides results in

$$\ln\left(x + \sqrt{b^2 + x^2}\right) \;=\; \pm \ln r + C \qquad (11)$$

where, C is the constant of integration. Hence,

$$x + \sqrt{b^2 + x^2} \;=\; \frac{k}{2} r^{\pm 1} \qquad (12)$$

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

$$\left(z - \frac{c}{2}\right)^2 + r^2 \left(1 - \frac{k}{2}\right) \;=\; \frac{c^2}{4}\left(\frac{k-2}{k}\right) \;, \qquad (13)$$

$$\left(z - \frac{c}{2}\right)^2 + r^2 \left(1 + \frac{c^2}{2k}\right) \;=\; \left(\frac{2k + c^2}{4}\right). \qquad (14)$$

Together, these two expressions represent the entire class of mirrors that satisfy the fixed viewpoint constraint. Again, since perspective projection is symmetric about any ray that passes through the pinhole, the viewpoint $\mathbf{v}$ and the corresponding mirror are in no way restricted to the optical axis.

3 Specific Mirrors

A quick glance at the forms of equations (13) and (14) reveals that the mirror profiles are conic sections. However, each conic section must be placed at a specific distance from the pinhole. As we shall see, though all our conic sections are theoretically valid, many prove to be impractical and only a few lead to useful solutions. We are now in a position to evaluate several specific cases.

3.1 Planes

In solution (13), if we set $k = 2$, we get the cross-section of a planar mirror:

$$z \;=\; \frac{c}{2} \;. \qquad (15)$$

As shown in Figure 2, the plane bisects the line segment joining the pinhole and the viewpoint. This result is easily generalized to arbitrary planes or viewpoints. For any plane with unit normal $\hat{\mathbf{n}}$ and any point $\mathbf{q}$ on it, the viewpoint is simply the reflection of the pinhole

$$\mathbf{v} \;=\; \mathbf{p} \;-\; 2\,(\,(\mathbf{p} - \mathbf{q}) \cdot \hat{\mathbf{n}}\,)\,\hat{\mathbf{n}} \;. \qquad (16)$$

Equivalently, for any desired viewpoint, points $\mathbf{x}$ on the planar mirror are given by

$$\left(\mathbf{x} - \frac{(\mathbf{p} + \mathbf{v})}{2}\right) \cdot (\mathbf{p} - \mathbf{v}) \;=\; 0 \;. \qquad (17)$$

These expressions lead us to a simple but unfortunate theorem: For a single fixed pinhole, no two planar mirrors can share the same viewpoint, and equivalently, two different viewpoints cannot be generated by the same planar mirror. It is clear from Figure 2 that a single planar mirror does not enhance the field of view of the imaging system. At the same time, the above theorem makes it impossible to increase the field of view by packing a large number of planar mirrors (pointing in different directions) in front of a conventional imaging system. On the brighter side, the two views of a scene needed for stereo can be captured by a single lens and two planar mirrors, as shown in [Goshtasby and Gruver, 1993].

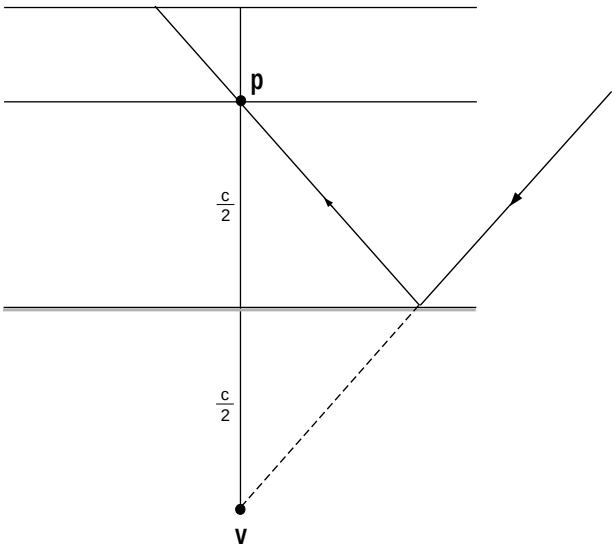

Figure 2: A planar mirror must bisect the segment joining the pinhole and the desired viewpoint. Since two planar mirrors cannot generate the same viewpoint, multiple planar mirrors cannot be used to enhance the field of view of a conventional imaging system.

To ensure a single viewpoint while using multiple planar mirrors, Nalwa [Nalwa, 1996] has ar-

rived at a clever design that includes four planar mirrors that form the faces of a pyramid. Four separate imaging systems are used, each one placed above one of the faces of the pyramid. The optical axes of the imaging systems and the angles made by the four planar faces are adjusted so that the four viewpoints produced by the planar mirrors coincide. The result is a sensor that has a single viewpoint and a panoramic field of view of approximately $360° \times 50°$. The panoramic image is of relatively high resolution as it is a concatenation of four images provided by four non-overlapping imaging systems. Nalwa's sensor is easy enough to implement but requires the use of four of each component (cameras, lenses, and digitizers).

3.2 Cones

In solution (13), if we set $c = 0$, the result is a conical mirror with cross-section

$$z = \sqrt{\frac{k-2}{2} r^2}\ , \qquad k \geq 2\ . \qquad (18)$$

The angle at the apex of the cone varies with k. At first glance, this may seem like a reasonable solution. However, since $c = 0$, the apex of the cone must be at the pinhole. This implies that the rays of light entering the pinhole can only graze the cone and do not represent reflections of the world (see Figure 3). Hence, we have a degenerate solution that is of no practical value.

Indeed, the cone has been used for wide-angle imaging, in particular, for autonomous navigation [Yagi and Kawato, 1990]. In these implementations, the apex of the cone was placed at a distance from the pinhole. In such cases, it is easy to show that the viewpoint is no longer a single point but rather a locus [Nalwa, 1996]. If the axis of the cone points in the direction of the pinhole, the locus is a circle that hangs like a halo around the cone.

3.3 Spheres

In solution (14), if we set $c = 0$, we get a spherical mirror with cross-section

$$z^2 + r^2 = \frac{k}{2}\ , \qquad k > 0\ . \qquad (19)$$

Like the cone, this proves to be a solution of little value; since the viewpoint and pinhole must coincide, the observer sees itself and nothing else, as shown in Figure 4.

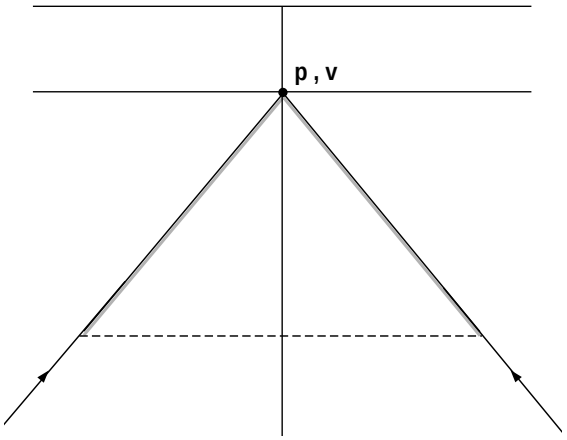

Figure 3: The conical mirror has its apex at the pinhole. This degenerate solution is of little practical value. If the apex is moved away from the pinhole, the viewpoint is no longer single but rather lies on a circular locus.

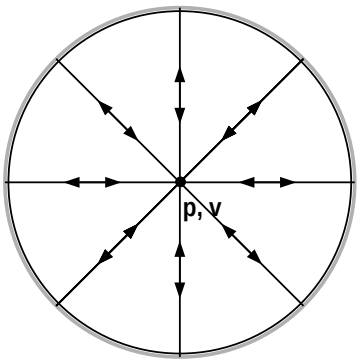

Figure 4: A spherical mirror produces a single viewpoint only when the pinhole lies at its center. This, again, is a solution of little use as the observer sees itself and nothing else.

In [Hong, 1991], a wide-angle implementation with a sphere is described that was used for landmark navigation. In this case, the sphere was placed at a distance from the effective pinhole of the camera. As with the cone, the result is a locus of viewpoints rather than a single viewpoint. The locus in the case of the sphere can turn out to be a surface of large extent, depending on the distance between the center of the sphere and the pinhole. In [Nayar, 1988], a stereo system is proposed that uses a single image of two specular spheres to compute depth. In this case, the sin-

gle viewpoint constraint is not critical as stereo requires multiple viewpoints.

3.4 Ellipsoids

In solution (14), when $c > 0$ and $k > 0$, we get an ellipsoid with cross-section

$$\frac{(z - \frac{c}{2})^2}{a^2} + \frac{r^2}{b^2} = 1 , \qquad (20)$$

where

$$a = \sqrt{\frac{2k + c^2}{4}} , \quad b = \sqrt{\frac{2k}{4}} . \qquad (21)$$

We have now arrived at a solution that can be used to enhance the field of view. As shown in Figure 5, the viewpoint **v** and pinhole **p** are located at the two foci of the ellipse, respectively. If, for instance, the section of the ellipse that lies beneath the viewpoint is used, the effective field of view (ignoring self-occlusion by the lens) corresponds to the upper hemisphere. It is easy to see that terminating the ellipse below the viewpoint does not enhance the field of view. Unfortunately, extending the mirror above the viewpoint does to help either; in this case, rays of light entering the pinhole would have undergone multiple reflections by the mirror. Yet, the ellipse does represent our first useful solution. It is similar in nature to our next solution, the hyperboloid. Hence, we shall defer our discussion on implementation issues related to the ellipsoid.

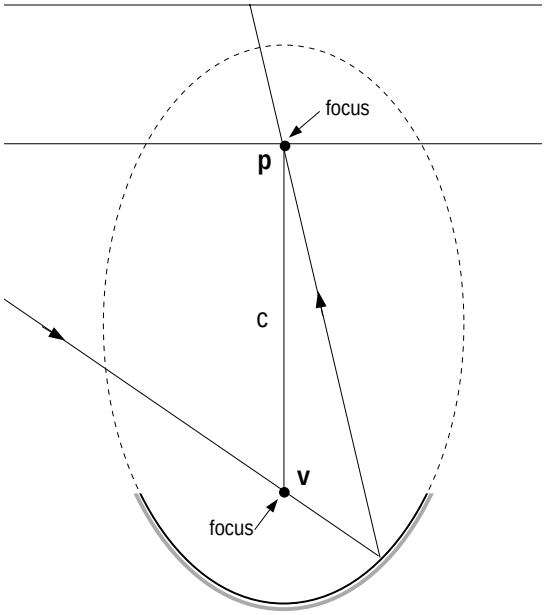

Figure 5: A ellipsoidal mirror is a viable solution when the pinhole and the viewpoint are located at its two foci, respectively. If the ellipsoid is terminated by a plane passing through the viewpoint, the field of view corresponds a hemisphere.

3.5 Hyperboloids

In solution (13), when $c > 0$ and $k > 2$, we get a hyperboloid[2] with cross-section

$$\frac{(z - \frac{c}{2})^2}{a^2} - \frac{r^2}{b^2} = 1 , \qquad (22)$$

where

$$a = \frac{c}{2}\sqrt{\frac{k-2}{k}} , \quad b = \frac{c}{2}\sqrt{\frac{2}{k}} . \qquad (23)$$

As we see in Figure 6, in the limit $k \to 2$, the hyperboloid flattens to yield the planar solution of section 3.1. As k increases, the curvature of the hyperboloid increases, and hence also the field of view of the catadioptric system. The two foci of

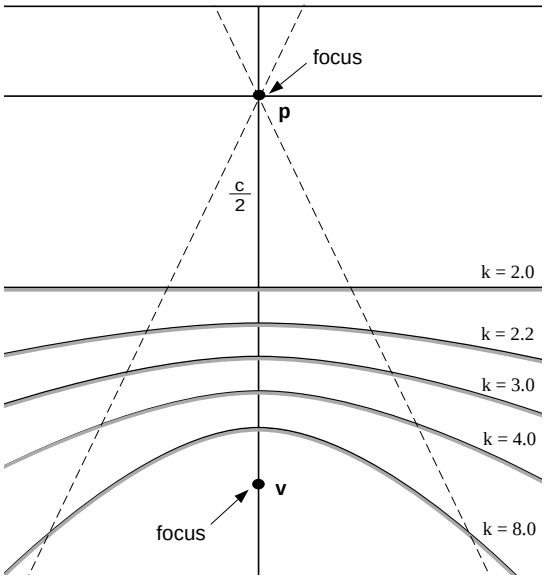

Figure 6: The hyperboloidal mirror can produce the desired increase in field of view. The pinhole and the viewpoint are located at the two hyperboloidal foci.

the hyperboloid remain fixed, one at the pinhole $\mathbf{p}$ and the other at the viewpoint $\mathbf{v}$.

This solution provides a practical approach to wide-angle imaging. Yamazawa et al. [Yamazawa *et al.*, 1993] recognized that the hyperboloid, if chosen and positioned carefully, would produce a single viewpoint. They implemented a sensor for autonomous navigation and demonstrated the construction of perspective images from hyperboloidal ones.

While this solution is both interesting and feasible, it must be implemented with care. As can be seen from Figure 6, for any chosen value of k, if the viewpoint is distant from the pinhole, the mirror must be large. As the viewpoint approaches the pinhole, the mirror reduces in size but the curvatures at all points on the mirror increase. This increases the optical effects of coma and astigmatism that are known to produce blurring [Hecht and Zajac, 1974]. Furthermore, it is hard to configure an imaging system with a large enough depth of field close to the pinhole that would allow the hyperboloidal mirror to be placed in close proximity. These trade-offs imply that the distance of the viewpoint must be chosen with care. Also, the axis of the hyperboloid must pass through the pinhole. For these reasons, careful implementation is required to achieve the desired optical properties and precise calibration is needed to establish the mapping between an incoming principle ray and its image coordinates. It is easy to see that all of the above implementational issues also apply to the ellipsoidal solution.

3.6 Paraboloids

If image projection is orthographic rather than perspective, the geometrical mappings between the image, the mirror and the world are invariant to translations of the mirror with respect to the imaging system. Consequently, both calibration as well as the computation of perspective images are greatly simplified. There are simple ways to achieve pure orthographic projection, as described in [Nayar, 1997].

The shape of the mirror in this case can be derived by assuming orthographic projection rather than perspective projection in the dioptric stage of image formation. This derivation is given in [Nayar, 1997] and the mirror is shown to be paraboloidal (see Figure 7). Paraboloidal mirrors are frequently used to converge an incoming set of parallel rays at a single point (the focus), or to generate a collimated light source from a point source (placed at the focus). In both these cases, the paraboloid is a concave mirror that is reflective on its inner surface. In our case, the paraboloid is reflective on its outer surface (convex mirror); all incoming principle rays are orthographically reflected by the mirror. Further, the incoming rays can be extended to intersect at the focus of the paraboloid, which serves as the viewpoint.

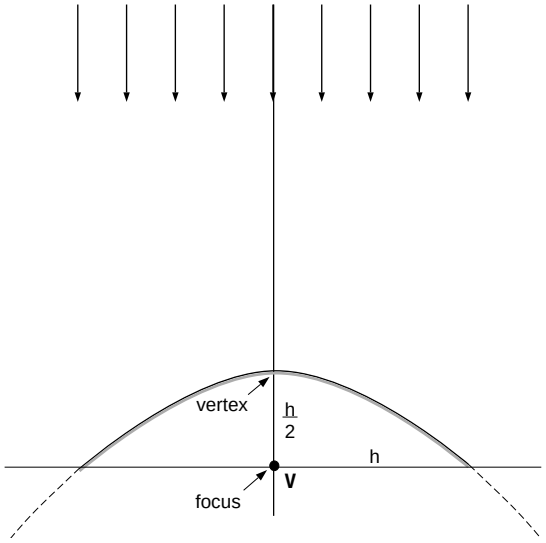

Figure 7: For orthographic projection, the solution is a paraboloid with the viewpoint located at the focus. Orthographic projection makes the geometric mappings between the image, the paraboloidal mirror and the world invariant to translations of the mirror. This greatly simplifies calibration and the computation of perspective images from paraboloidal ones.

Alternatively, the same solution can be derived from our general solution for catadioptric imaging. We know that orthographic projection is a limiting case of perspective, where the distance between the pinhole and viewpoint approaches infinity. Equation (13) can be rewritten as:

$$\frac{z^2}{k} - \frac{zc}{k} + \frac{r^2}{k} - \frac{1}{2}r^2 \;=\; -\frac{c^2}{2k^2}. \qquad (24)$$

Then, in the limit $c \to \infty$, $k \to \infty$, while keeping $c/k = h$ a constant, we have

$$z \;=\; \frac{h^2 - r^2}{2h}. \qquad (25)$$

[2]Note that $k < 2$ is not possible in solution (13) since this yields an imaginary surface.

The parameter h of the paraboloid is its radius at $z = 0$. The distance between the vertex and the focus is $h/2$. Therefore, h determines the size of the paraboloid that, for any given orthographic lens system, can be chosen to maximize resolution. If the paraboloid is terminated at its focus, the imaging system yields a hemispherical field of view. As shown in [Nayar, 1997], two such imaging systems can be placed back-to-back to achieve a spherical field of view.

4 Resolution

Here, we define resolution as the solid angle subtended from the viewpoint by a pixel in the sensed image. Let us assume that the area projected by a pixel along its line of sight is da, as shown in Figure 8. Note that for orthographic projection da is a constant, while for perspective projection it is easily computed from the distance of the corresponding point on the mirror and the focal length of the imaging system. Given that all reflections are specular, the reflecting surface area occupied by da is $ds = da/\cos\phi$. The foreshortened surface area as seen by the viewpoint $\mathbf{v}$ is $(da/\cos\phi)\cos\phi = da$. The solid angle subtended by the reflecting surface element is $d\omega = da/t^2$, where t is the distance of the surface element from the viewpoint. Hence, the spatial resolution for any catadioptric sensor can be written as

$$\frac{da}{d\omega} = t^2 = \sqrt{z^2 + r^2}\,. \qquad (26)$$

For instance, in the case of a paraboloidal mirror [Nayar, 1997], the resolution increases by a factor of 4 from the vertex ($t = h/2$) of the paraboloid to the fringe ($t = h$). With (26), it is easy to see that a variation in spatial resolution occurs not only in the case of curved mirrors but also planar ones. In principle, it is of course possible to use image detectors with non-uniform resolution to compensate for the above variation.

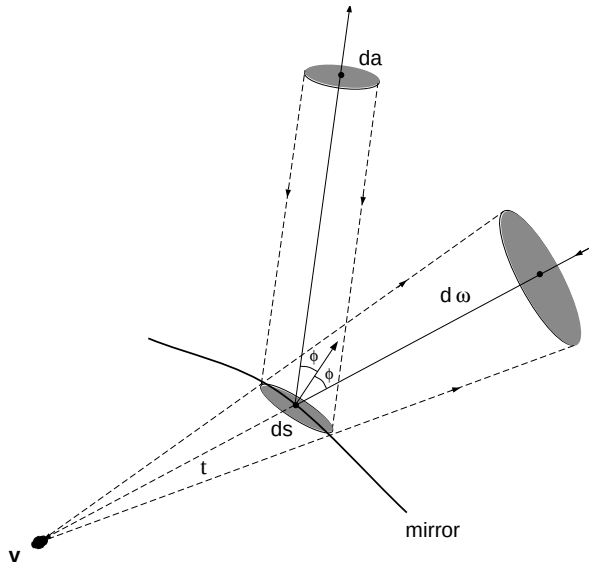

Figure 8: Geometry used to derive spatial resolution ($da/d\omega$) for any catadioptric sensor.

References

[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, Reading, Massachusetts, 1974.

[Hong, 1991] J. Hong. Image Based Homing. *Proc. of IEEE International Conference on Robotics and Automation*, May 1991.

[Nalwa, 1996] V. Nalwa. A True Omnidirectional Viewer. Technical report, Bell Laboratories, Holmdel, NJ 07733, U.S.A., February 1996.

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

[Nayar, 1997] S. K. Nayar. Omnidirectional Video Camera. *Proc. of DARPA Image Understanding Workshop*, May 1997.

[Yagi and Kawato, 1990] Y. Yagi and S. Kawato. Panoramic Scene Analysis with Conic Projection. *Proc. of International Conference on Robots and Systems (IROS)*, 1990.

[Yamazawa *et al.*, 1993] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional Imaging with Hyperboloidal Projection. *Proc. of International Conference on Robots and Systems (IROS)*, 1993.