Catadioptric Omnidirectional Camera *

Shree K. Nayar

Department of Computer Science, Columbia University

New York, New York 10027

Email: nayar@cs.columbia.edu

Abstract

Conventional video cameras have limited fields of view that make them restrictive in a variety of vision applications. There are several ways to enhance the field of view of an imaging system. However, the entire imaging system must have a single effective viewpoint to enable the generation of pure perspective images from a sensed image. A new camera with a hemispherical field of view is presented. Two such cameras can be placed back-to-back, without violating the single viewpoint constraint, to arrive at a truly omnidirectional sensor. Results are presented on the software generation of pure perspective images from an omnidirectional image, given any user-selected viewing direction and magnification. The paper concludes with a discussion on the spatial resolution of the proposed camera.

1 Introduction

Conventional imaging systems are quite limited in their field of view. Is it feasible to devise a video camera that can, at any instant in time, "see" in all directions? Such an *omnidirectional* camera would have an impact on a variety of applications, including autonomous navigation, remote surveillance, video conferencing, and scene recovery.

Our approach to omnidirectional image sensing is to incorporate reflecting surfaces (mirrors) into conventional imaging systems. This is what we refer to as *catadioptric* image formation. There are a few existing implementations that are based on this approach to image sensing (see [Nayar-1988], [Yagi and Kawato-1990], [Hong-1991], [Goshtasby and Gruver-1993], [Yamazawa *et al.*-1995], [Nalwa-1996]). As noted in [Yamazawa *et al.*-1995] and [Nalwa-1996], in order to compute pure perspective images from a wide-angle image, the catadioptric imaging system must have a single center of projection (viewpoint). In [Nayar and Baker-1997], the complete class of catadioptric systems that satisfy the single viewpoint constraint is derived. Since we are interested in the development of a practical omnidirectional camera, two additional conditions are imposed. First, the camera should be easy to implement and calibrate. Second, the mapping from world coordinates to image coordinates must be simple enough to permit fast computation of perspective and panoramic images.

We begin by reviewing the state-of-the-art in wide-angle imaging and discuss the merits and drawbacks of existing approaches. Next, we present an omnidirectional video camera that satisfies the single viewpoint constraint, is easy to implement, and produces images that are efficient to manipulate. We have implemented several prototypes of the proposed camera, each one designed to meet the requirements of a specific application. Results on the mapping of omnidirectional images to perspective ones are presented. In [Peri and Nayar-1997], a software system is described that generates a large number of perspective and panoramic video streams from an omnidirectional video input. We conclude with a discussion on the resolution of the proposed camera.

2 Omnidirectional Viewpoint

It is worth describing why it is desirable that any imaging system have a single *center of projection.* Strong cases in favor of a single viewpoint have also been made by Yamazawa et al. [Yamazawa *et al.*-1995] and Nalwa [Nalwa-1996]. Consider an image acquired by a sensor that can view the world in all directions from a single effective pinhole (see Figure 1). From such an omnidirectional image, pure perspective images can be constructed by mapping sensed brightness values onto a plane placed at any distance (effective focal length) from the viewpoint, as shown in Figure 1. Any image computed in this manner preserves linear perspective geometry. 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.

*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.

1063-6919/97 $10.00 © 1997 IEEE

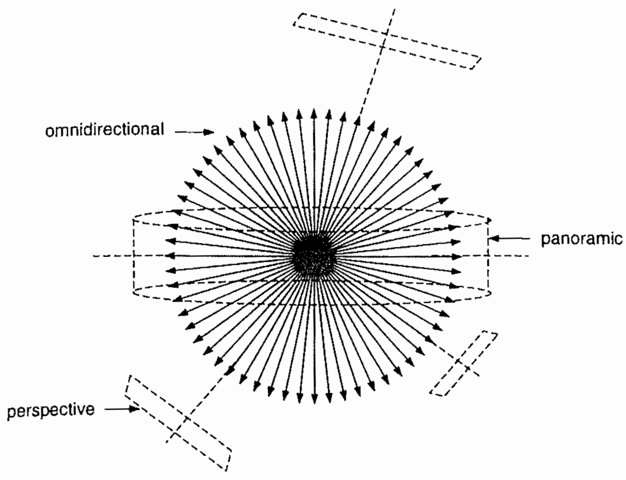


Figure 1: A truly omnidirectional image sensor views the world through an entire "sphere of view" as seen from its center of projection. The single viewpoint permits the construction of pure perspective images (computed by planar projection) or a panoramic image (computed by cylindrical projection). Panoramic sensors are not equivalent to omnidirectional sensors as they are omnidirectional only in one of the two angular dimensions.

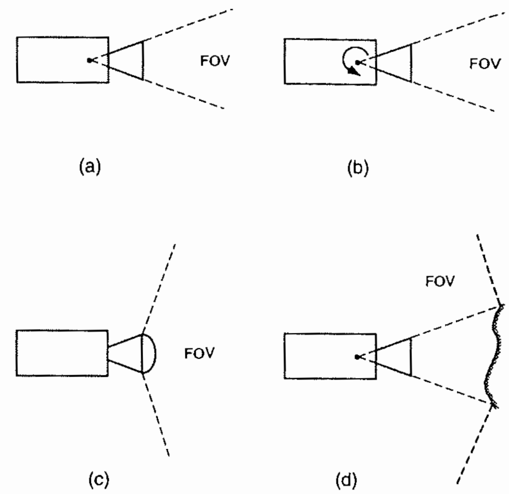


Figure 2: (a) A conventional imaging system and its limited field of view. A larger field of view may be obtained by (b) rotating the imaging system about its center of projection, (c) appending a fish-eye lens to the imaging system, and (d) imaging the scene through a mirror.

3 State of the Art

Before we present our omnidirectional camera, a review of existing imaging systems that seek to achieve wide fields of view is in order. An excellent review of some of the previous work can be found in [Nalwa-1996].

3.1 Traditional Imaging Systems

Most imaging systems in use today comprise of a video camera, or a photographic film camera, attached to a lens. The image projection model for most camera lenses is perspective with a single center of projection. Since the imaging device (CCD array, for instance) is of finite size and the camera lens occludes itself while receiving incoming rays, the lens typically has a small field of view that corresponds to a small cone rather than a hemisphere (see Figure 2(a)). At first thought, it may appear that a large field can be sensed by packing together a number of cameras, each one pointing in a different direction. However, since the centers of projection reside inside their respective lenses, such a configuration proves infeasible.

3.2 Rotating Imaging Systems

An obvious solution is to rotate the entire imaging system about its center of projection, as shown in Figure 2(b). The sequence of images acquired by rotation are "stitched" together to obtain a panoramic view of the scene. Such an approach has been recently proposed by several investigators (see [Chen-1995], [McMillan and Bishop-1995], [Krishnan and Ahuja-1996], [Zheng and Tsuji-1990]). Of these the most novel is the system developed by Krishnan and Ahuja [Krishnan and Ahuja-1996] which uses a camera with a non-frontal image detector to scan the world.

The first disadvantage of any rotating imaging system is that it requires the use of moving parts and precise positioning. A more serious drawback lies in the total time required to obtain an image with enhanced field of view. This restricts the use of rotating systems to static scenes and non-real-time applications.

3.3 Fish-Eye Lenses

An interesting approach to wide-angle imaging is based on the fish-eye lens (see [Wood-1906], [Miyamoto-1964]). Such a lens is used in place of a conventional camera lens and has a very short focal length that enables the camera to view objects within as much as a hemisphere (see Figure 2(c)). The use of fish-eye lenses for wide-angle imaging has been advocated in [Oh and Hall-1987] and [Kuban *et al.*-1994], among others.

It turns out that it is difficult to design a fish-eye lens that ensures that all incoming principal rays intersect at a single point to yield a fixed viewpoint (see [Nalwa-1996] for details). This is indeed a problem with commercial fish-eye lenses, including, Nikon's Fisheye-Nikkor 8mm f/2.8 lens. In short, the acquired image does not permit the construction of distortion-free perspective images of the viewed scene (though constructed images may prove good enough for some visualization applications). In addition, to capture a hemispherical view, the fish-eye lens must be quite complex and large, and hence expensive.

3.4 Catadioptric Systems

As shown in Figure 2(d), a catadioptric imaging system uses a reflecting surface to enhance the field of view. The rear-view mirror in a car is used exactly in this fashion. However, the shape, position, and orientation of the reflecting surface are related to the viewpoint and the field of view in a complex manner. While it is easy to construct a configuration which includes one or more mirrors that dramatically increase the field of view of the imaging system, it is hard to keep the effective viewpoint fixed in space. Examples of catadioptric image sensors can be found in [Yagi and Kawato-1990], [Hong-1991], [Yamazawa *et al.*-1995], and [Nalwa-1996]. A recent theoretical result (see [Nayar and Baker-1997]) reveals the complete class of catadioptric imaging systems that satisfy the single viewpoint constraint. This general solution has enabled us to evaluate the merits and drawbacks of previous implementations as well as suggest new ones [Nayar and Baker-1997].

Here, we will briefly summarize previous approaches. In [Yagi and Kawato-1990], a conical mirror is used in conjunction with a perspective lens. Though this provides a panoramic view, the single viewpoint constraint is not satisfied. The result is a viewpoint locus that hangs like a halo over the mirror. In [Hong-1991], a spherical mirror was used with a perspective lens. Again, the result is a large locus of viewpoints rather than a single point. In [Yamazawa *et al.*-1995], a hyperboloidal mirror used with a perspective lens is shown to satisfy the single viewpoint constraint. This solution is a useful one. However, the sensor must be implemented and calibrated with care. More recently, in [Nalwa-1996], a novel panoramic sensor has been proposed 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 $360° \times 50°$. Again, careful alignment and calibration are needed during implementation.

4 Omnidirectional Camera

While all of the above approaches use mirrors placed in the view of perspective lenses, we approach the problem using an orthographic lens. It is easy to see that 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 is greatly simplified.

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

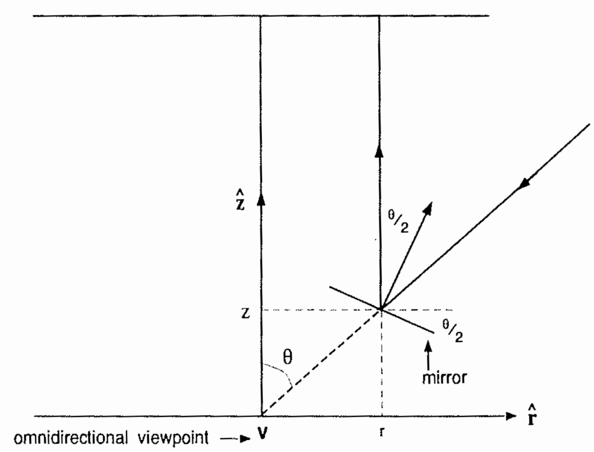


Figure 3: Geometry used to derive the reflecting surface that produces an image of the world as seen from a fixed viewpoint v. This image is captured using an orthographic (telecentric) imaging lens.

We are now ready to derive the shape of the reflecting surface. Since orthographic projection is rotationally symmetric, all we need to determine is the cross-section $z(r)$ of the reflecting surface. The mirror is then the solid of revolution obtained by sweeping the cross-section about the axis of orthographic projection. As illustrated in Figure 3, each ray of light from the world heading in the direction of the viewpoint v must be reflected by the mirror in the direction of orthographic projection. The relation between the angle θ of the incoming ray and the profile $z(r)$ of the reflecting surface is

$$\tan\theta \;=\; \frac{r}{z}\,. \qquad (1)$$

Since the surface is specular, the angles of incidence and reflectance are equal to $\theta/2$. Hence, the slope at the point of reflection can be expressed as

$$\frac{dz}{dr} \;=\; -\,\tan\frac{\theta}{2}\,. \qquad (2)$$

Now, we use the trignometric identity

$$\tan\theta \;=\; \frac{2\tan\frac{\theta}{2}}{1\,-\,\tan^2\frac{\theta}{2}}\,. \qquad (3)$$

Substituting (1) and (2) in the above expression, we obtain

$$\frac{-\,2\,\frac{dz}{dr}}{1\,-\,(\frac{dz}{dr})^2} \;=\; \frac{r}{z}\,. \qquad (4)$$

Thus, we find that the reflecting surface must satisfy a quadratic first-order differential equation. The first step is to solve the quadratic expression for surface slope. This gives us two solutions of which only one is valid since the slope of the surface in the first quadrant is assumed to be negative (see Figure 3):

$$\frac{dz}{dr} \;=\; \frac{z}{r} \,-\, \sqrt{1\,+\,(\frac{r}{z})^2}\,. \qquad (5)$$

This first-order differential equation can be solved to obtain the following expression for the reflecting surface:

$$z \;=\; \frac{h^2\,-\,r^2}{2h}\,, \qquad (6)$$

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

Not surprisingly, the mirror that guarantees a single viewpoint for orthographic projection is a paraboloid. Paraboloidal mirrors are frequently used to converge an incoming set of parallel rays at a single point (the focus), or to generate a collimated light source from a point source (placed at the focus). In both these cases, the paraboloid is a concave mirror that is reflective on its inner surface. In our case, the paraboloid is reflective on its outer surface (convex mirror); all incoming principle rays are orthographically reflected by the mirror but can be extended to intersect at its focus, which serves as the viewpoint. Note that a concave paraboloidal mirror can also be used (this corresponds to the second solution we would get from equation (4) if the slope of the mirror in the first quadrant is assumed to be positive). This solution is less desirable to us since incoming rays with large angles of incidence θ would be self-occluded by the mirror.

As shown in Figure 4, 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. Shortly, the issue of resolution will be addressed in more detail.

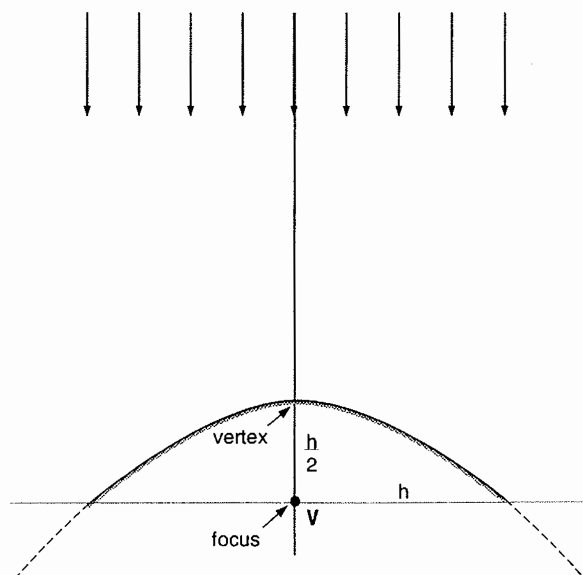


Figure 4: For orthographic projection, the solution is a paraboloid with the viewpoint located at the focus. Orthographic projection makes the geometric mappings between the image, the paraboloidal mirror and the world invariant to translations of the mirror. This greatly simplifies calibration and the computation of perspective images from paraboloidal ones.

5 Field of View

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

$$| \frac{dz}{dr} | \Big|_{z=0} = 1 . \quad (7)$$

Hence, if we cut the paraboloid at the plane $z = 0$, the field of view exactly equals the upper hemisphere (minus the solid angle subtended by the imaging system itself). If a field of view greater than a hemisphere is desired, the paraboloid can be terminated below the $z = 0$ plane. If only a panorama is of interest, an annular section of the paraboloid may be obtained by truncating it below and above the $z = 0$ plane. For that matter, given any desired field of view, the corresponding section of the parabola can be used and the entire resolution of the imaging device can be dedicated to that section's projection in the image.

In our prototypes, we have chosen to terminate the parabola at the $z = 0$ plane. This proves advantageous in applications in which the complete sphere of view is desired, as shown in Figure 5. Since the paraboloid is terminted at the focus, it is possible to place two identical catadioptric cameras back-to-back such that their foci (viewpoints) coincide. Thus, we have a truly omnidirectional sensor, one that is capable of acquiring an entire sphere of view at video rate.

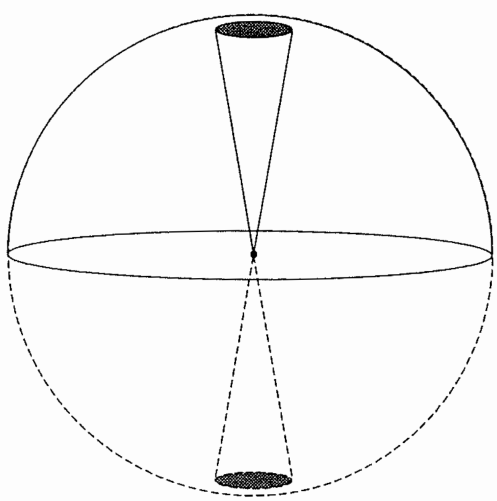


Figure 5: If the paraboloid is cut by the horizontal plane that passes through its focus, the field of view of the catadioptric system exactly equals the upper hemisphere. This allows us to place two catadioptric sensors back-to-back such that their foci (viewpoints) coincide. The result is a truly omnidirectional sensor that can acquire the entire sphere of view. The shaded regions are parts of the field of view where the sensor sees itself.

6 Implementation

Several versions of the proposed omnidirectional sensor have been built, each one geared towards a specific application. The applications we have in mind include video teleconferencing, remote surveillance and autonomous navigation. Figure 6 shows and details the different sensors and their components. The basic components of all the sensors are the same; each one includes a paraboloidal mirror, an orthographic lens system and a CCD video camera. The sensors differ primarily in the their mechanical designs and their attachments. For instance, the sensors in Figures 6(a) and 6(c) have transparent spherical domes that minimize self-obstruction of their hemispherical fields of view. Figure 6(d) shows a back-to-back implementation that is capable of acquiring the complete sphere of view.

The use of paraboloidal mirrors virtually obviates calibration. All that is needed are the image coordinates of the center of the paraboloid and its radius h. Both these quantities are measured in pixels from a single omnidirectional image. We have implemented software for the generation of perspective images. First, the user specifies the viewing direction, the image size and effective focal length (zoom) of the desired perspective image (see Figure 1). Again, all these quantities are specified in pixels. For each three-dimensional pixel location (x_p, y_p, z_p) on the desired perspective image plane, its line of sight with respect to the viewpoint is computed in terms of its polar and azimuthal angles:

$$\theta = \cos^{-1} \frac{z_p}{\sqrt{{x_p}^2 + {y_p}^2 + {z_p}^2}} , \quad \phi = \tan^{-1} \frac{y_p}{x_p} . \quad (8)$$

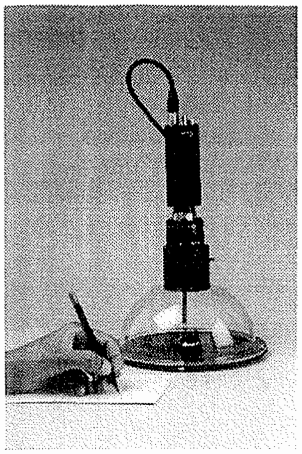
This line of sight intersects the paraboloid at a distance ρ from its focus (origin), which is computed using the following spherical expression for the paraboloid:

$$\rho = \frac{h}{(1 + \cos\theta)} . \quad (9)$$

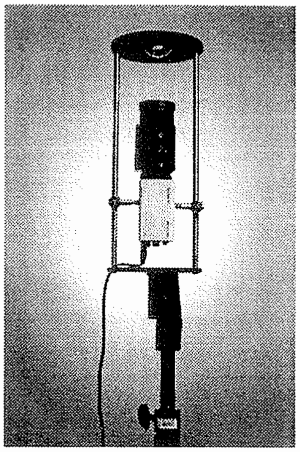
The brightness (or color) at the perspective image point (x_p, y_p, z_p) is then the same as that at the omnidirectional image point

$$x_i = \rho \sin\theta \cos\phi , \quad y_i = \rho \sin\theta \sin\phi . \quad (10)$$

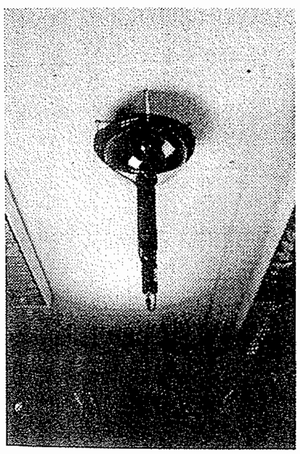
The above computation is repeated for all points in the desired perspective image. Figure 7 shows an omnidirectional image (512x480 pixels) and several perspective images (200x200 pixels each) computed from it. It is worth noting that perspective projection is indeed preserved. For instance, straight lines in the scene map to straight lines in the perspective images while they appear as curved lines in the omnidirectional image. Recently, a video-rate version of the above described image generation has been developed as an interactive software system called OmniVideo [Peri and Nayar-1997].



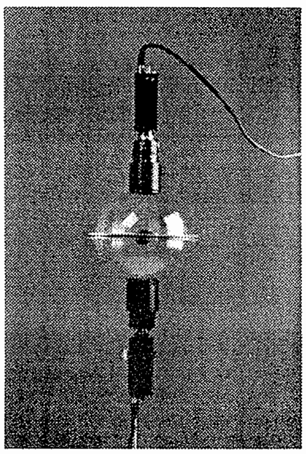
(a)



(b)



(c)



(d)

Figure 6: Four implementations of catadioptric omnidirectional video cameras that use paraboloidal mirrors. (a) This compact sensor for *teleconferencing* uses a 1.1 inch diameter paraboloidal mirror, a Panasonic GP-KR222 color camera, and Cosmicar/Pentax C6Z1218 zoom and close-up lenses to achieve orthography. The transparent spherical dome minimizes self-obstruction of the field of view. (b) This camera for *navigation* uses a 2.2 inch diameter mirror, a DXC-950 Sony color camera, and a Fujinon CVL-713 zoom lens. The base plate has an attachment that facilitates easy mounting on mobile platforms. (c) This sensor for *surveillance* uses a 1.6 inch diameter mirror, an Edmund Scientific 55mm F/2.8 telecentric (orthographic) lens and a Sony XR-77 black and white camera. The sensor is lightweight and suitable for mounting on ceilings and walls. (d) This sensor is a back-to-back configuration that enables it to sense the entire sphere of view. Each of its two units is identical to the sensor in (a).

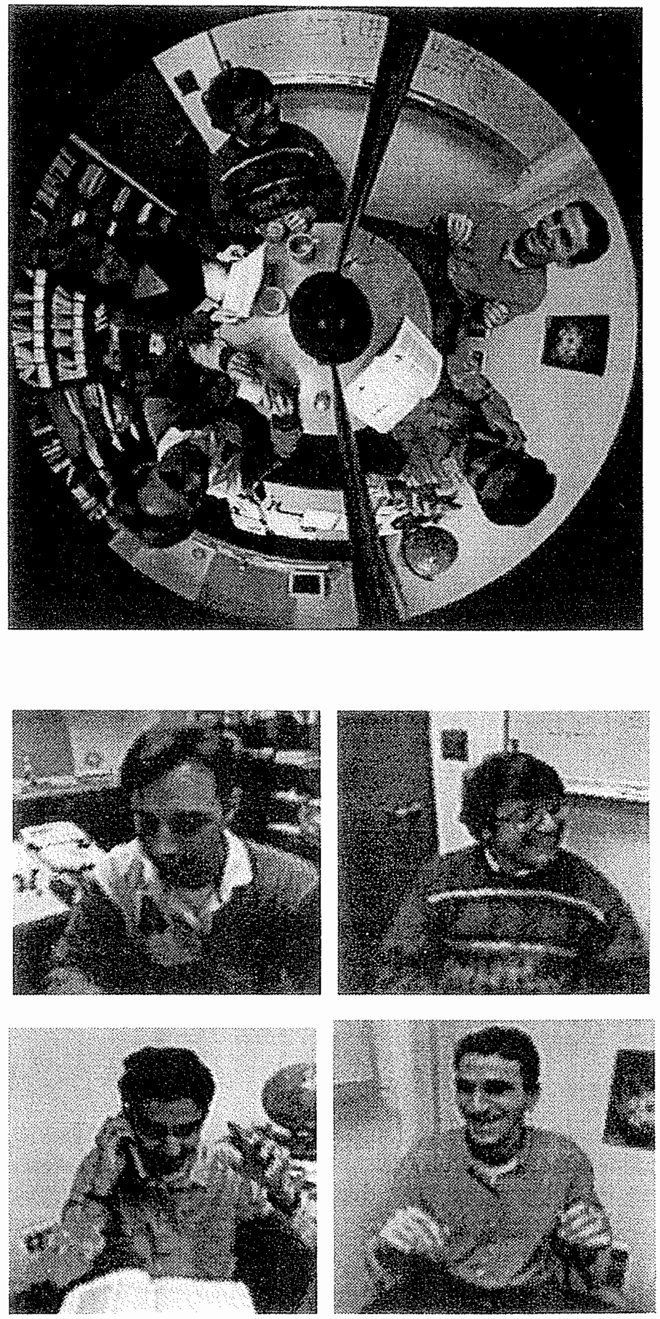


Figure 7: Software generation of perspective images (bottom) from an omnidirectional image (top). Each perspective image is generated using user-selected parameters, including, viewing direction (line of sight from the viewpoint to the center of the desired image), effective focal length (distance of the perspective image plane from the viewpoint of the sensor), and image size (number of desired pixels in each of the two dimensions). It is clear that the computed images are indeed perspective; for instance, straight lines are seen to appear as straight lines though they appear as curved lines in the omnidirectional image.

7 Resolution

Several factors govern the resolution of a catadioptric sensor. Let us begin with the most obvious of these, the spatial resolution due to finite pixel size. In [Nayar and Baker-1997], we have derived a general expression for the spatial resolution of any catadioptric camera. In the case of our paraboloidal mirror, the resolution increases by a factor of 4 from the vertex ($r = 0$) of the paraboloid to the fringe ($r = h$). In principle, it is of course possible to use image detectors with non-uniform resolution to compensate for the above variation. It should also be mentioned that while all our implementations use CCD arrays with 512x480 pixels, nothing precludes us from using detectors with 1024x1024 or 2048x2048 pixels that are commercially available at a higher cost.

More intriguing are the blurring effects of coma and astigmatism that arise due to the aspherical nature of the reflecting surface [Born and Wolf-1965]. Since these effects are linear but shift-variant [Robbins and Huang-1972], a suitable set of deblurring filters need to be explored. Alternatively, these effects can be significantly reduced using inexpensive corrective lenses.

Acknowledgements

This work was inspired by the prior work of Vic Nalwa of Lucent Technologies. I have benefitted greatly from discussions with him. I thank Simon Baker and Venkata Peri of Columbia University for their valuable comments on various drafts of this paper.

References

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

[Chen, 1995] S. E. Chen. QuickTime VR - An Image Based Approach to Virtual Environment Navigation. *Computer Graphics: Proc. of SIGGRAPH 95*, pages 29–38, August 1995.

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

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

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

[Krishnan and Ahuja, 1996] A. Krishnan and N. Ahuja. Panoramic Image Acquisition. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR-96)*, pages 379–384, June 1996.

[Kuban *et al.*, 1994] D. P. Kuban, H. L. Martin, S. D. Zimmermann, and N. Busico. Omniview Motionless Camera Surveillance System. *United States Patent No. 5,359,363*, October 1994.

[McMillan and Bishop, 1995] L. McMillan and G. Bishop. Plenoptic Modeling: An Image-Based Rendering System. *Computer Graphics: Proc. of SIGGRAPH 95*, pages 39–46, August 1995.

[Miyamoto, 1964] K. Miyamoto. Fish Eye Lens. *Journal of Optical Society of America*, 54(8):1060–1061, August 1964.

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

[Nayar and Baker, 1997] S. K. Nayar and S. Baker. Catadioptric Image Formation. *Proc. of DARPA Image Understanding Workshop*, May 1997.

[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.

[Oh and Hall, 1987] S. J. Oh and E. L. Hall. Guidance of a Mobile Robot using an Omnidirectional Vision Navigation System. *Proc. of the Society of Photo-Optical Instrumentation Engineers, SPIE*, 852:288–300, November 1987.

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

[Robbins and Huang, 1972] G. M. Robbins and T. S. Huang. Inverse Filtering for Linear Shift-Variant Imaging Systems. *Proceedings of the IEEE*, 60(7):862–872, July 1972.

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

[Wood, 1906] R. W. Wood. Fish-eye views, and vision under water. *Philosophical Magazine*, 12(Series 6):159–162, 1906.

[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.*, 1995] K. Yamazawa, Y. Yagi and M. Yachida. Obstacle Avoidance with Omnidirectional Image Sensor HyperOmni Vision. *Proc. of IEEE International Conference on Robotics and Automation*, pages 1062–1067, May 1995.

[Zheng and Tsuji, 1990] J. Y. Zheng and S. Tsuji. Panoramic Representation of Scenes for Route Understanding. *Proc. of the Tenth International Conference on Pattern Recognition*, 1:161–167, June 1990.