The Eighth International
Symposium of Robotics Research
Hayama, Japan, October 3-7, 1997

Omnidirectional Vision

Shree K. Nayar
Dept. of Computer Science, Columbia University
New York City, New York 10027, (U.S.A.)
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 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.

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, video surveillance, video conferencing, virtual reality and site modelling.

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 (Nayar 1997). There are a few existing implementations that are based on this approach to image sensing (see (Rees 1970), (Charles 1987), (Nayar 1988), (Yagi and Kawato 1990), (Hong 1991), (Goshtasby and Gruver 1993), (Yamazawa, Yagi, and Yachida 1995), (Bogner 1995), (Murphy 1995), (Nalwa 1996), (Southwell, Basu, Fiala, and Reyda 1996), (Nayar 1997) and (Nene and Nayar 1998)). As noted in (Rees 1970), (Yamazawa, Yagi, and Yachida 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) (Baker and Nayar 1998), 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 (Nayar 1997) 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. More recently, the proposed omnidirectional camera has been used for robust computation of egomotion (Gluckman and Nayar 1998). Also, planar and curved mirrors have been used to develop compact binocular stereo systems (Nene and Nayar 1998).

2 Single Viewpoint

The merits of having a single center of projection (viewpoint) have been emphasized by Rees (Rees 1970), Yamazawa et al. (Yamazawa, Yagi, and Yachida 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

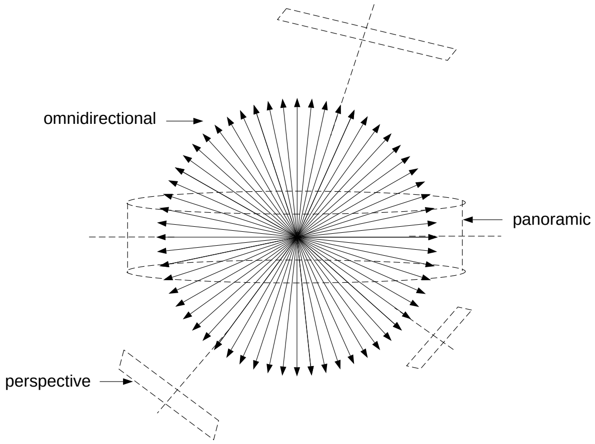


Figure 1: A truly omnidirectional image sensor views the world through an entire "sphere of view" as seen from its center of projection. The single viewpoint permits the construction of pure perspective images (computed by planar projection) or a panoramic image (computed by cylindrical projection).

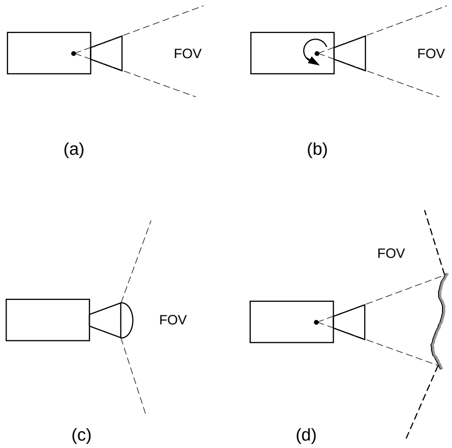


Figure 2: (a) A conventional imaging system and its limited field of view. A larger field of view may be obtained by (b) rotating the imaging system about its center of projection, (c) appending a fish-eye lens to the imaging system, and (d) imaging the scene through a mirror.

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.

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 overview 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 limited 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 proposed by several investigators (see (Chen 1995), (McMillan and Bishop 1995), (Krishnan and Ahuja 1996), (Zheng and Tsuji 1990), for examples). Of these the most novel is the system developed by Krishnan and Ahuja (Krishnan and Ahuja 1996) which uses a camera with a nonfrontal 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), (Slater 1932), (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 a hemisphere or more (see Figure 2(c)). The use of fish-eye lenses for wide-angle video has been advocated in (Oh and Hall 1987), (Kuban, Martin, Zimmermann, and Busico 1994) and (Xiong and Turkowski 1997), 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. Furthermore, in our quest for a truly omnidirectional sensor, we are physically restricted in placing two fish-eye imaging systems back-to-back to image the complete sphere of view; the two viewpoint loci reside inside their respective lenses and hence cannot be brought close to one another. However, in applications where a single viewpoint is not critical, a back-to-back configuration such as the one implemented by Slater (Slater 1996) can be used.

3.4 Catadioptric Systems

As shown in Figure 2(d), a catadioptric imaging system uses a reflecting surface to enhance the field of view. 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 harder to keep the effective viewpoint fixed in space. Examples of catadioptric image sensors can be found in (Rees 1970), (Charles 1987), (Yagi and Kawato 1990), (Hong 1991), (Yamazawa, Yagi, and Yachida 1995), (Bogner 1995), (Murphy 1995) and (Nalwa 1996). A recent theoretical result (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 (Baker and Nayar 1998).

Here, we will briefly review previous approaches. In (Yagi and Kawato 1990) and (Bogner 1995), 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), (Bogner 1995) and (Murphy 1995), a spherical mirror was used with a perspective lens. Again, the result is a large locus of viewpoints rather than a single point. Hyperboloidal, paraboloidal and ellipsoidal mirrors have been used in the implementation of "all-sky" photographic cameras dating back to the late 1950's (examples can be found in (Charles 1987)). In most of these implementations, the single viewpoint constraint does not seem to have been a major consideration.

Rees (Rees 1970) appears to have been the first to use a hyperboloidal mirror with a perspective lens placed at its external focus to achieve a single viewpoint video camera system. A similar implementation was recently proposed in (Yamazawa, Yagi, and Yachida 1995). The hyperboloidal solution is a useful one. However, the sensor must be implemented and calibrated with care. More recently, Nalwa (Nalwa 1996) has proposed a panoramic sensor 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 $360° \times 50°$.

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. Most of these are described in (Nayar 1997). The one that we have adopted in many of our implementations is the use of an inexpensive relay lens in front of 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.

Since orthographic projection is rotationally symmetric about the optical axis, 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. A detailed derivation of the mirror shape for orthographic projection is given in (Nayar 1997). Not surprisingly, the mirror that guarantees a single viewpoint is a paraboloid with cross-section:

$$z = \frac{h^2 - r^2}{2h} , \qquad (1)$$

where, $h > 0$ is the *parameter* of 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 solution is less desirable to us since incoming rays with large angles of incidence could be self-occluded by the mirror.

As shown in Figure 3, 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.

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

$$\left| \frac{dz}{dr} \right|\Big|_{z=0} = 1 . \qquad (2)$$

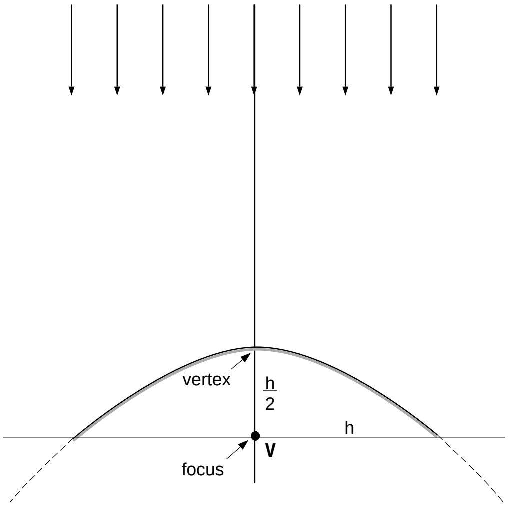


Figure 3: For orthographic projection, the solution is a paraboloid with the viewpoint located at the focus. Orthographic projection makes the geometric mappings between the image, the paraboloidal mirror and the world invariant to translations of the mirror. This greatly simplifies calibration and the computation of perspective images from paraboloidal ones.

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.

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 4. Since the paraboloid is termінted 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.

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 5 shows and details the different sensors and their components. The basic components of all

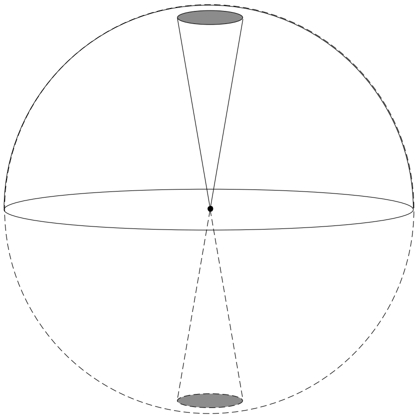


Figure 4: If the paraboloid is cut by the horizontal plane that passes through its focus, the field of view of the catadioptric system exactly equals the upper hemisphere. This allows us to place two catadioptric sensors back-to-back such that their foci (viewpoints) coincide. The result is a truly omnidirectional sensor that can acquire the entire sphere of view. The shaded regions are parts of the field of view where the sensor sees itself.

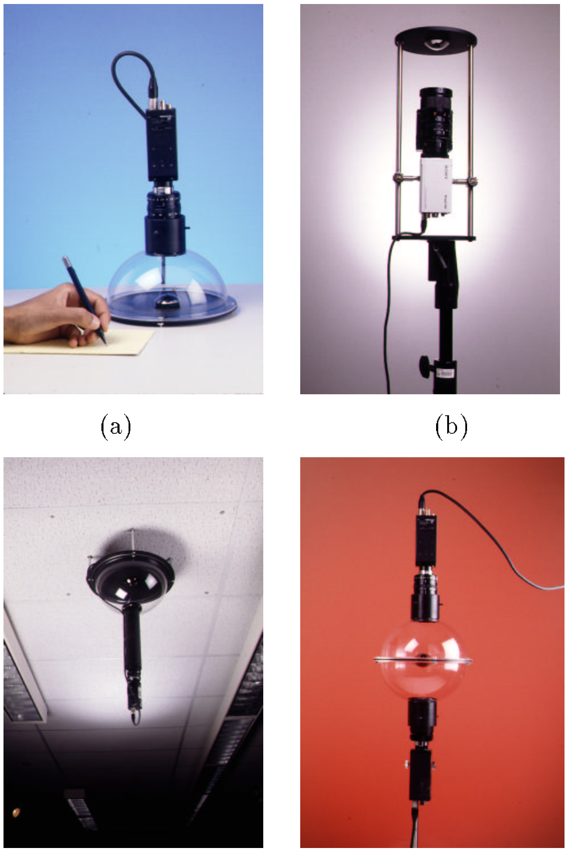
the sensors are the same; each one includes a paraboloidal mirror, an orthographic lens system and a CCD video camera. The sensors differ primarily in their mechanical designs and their attachments. Figure 5(d) shows a back-to-back implementation that is capable of acquiring the complete sphere of view.

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:

$$\theta \;=\; \cos^{-1}\frac{z_p}{\sqrt{{x_p}^2+{y_p}^2+{z_p}^2}}\;, \qquad \phi \;=\; \tan^{-1}\frac{y_p}{x_p}\;. \tag{3}$$

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:

$$\rho \;=\; \frac{h}{(1\,+\,\cos\theta)}\;. \tag{4}$$



(a) (b)

(c) (d)

Figure 5: 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).

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

$$x_i = \rho \sin\theta \cos\phi \ , \qquad y_i = \rho \sin\theta \sin\phi \ . \quad (5)$$

The above computation is repeated for all points in the desired perspective image. Figure 6 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. A video-rate version of the above described image generation is detailed in (Peri and Nayar 1997).

7 Resolution

Several factors govern the resolution of a catadioptric sensor. The most obvious of these is 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 many applications, this turns out to be a benefit of using a curved mirror instead of a fish-eye lens; often, the panorama is of greater interest than the rest of the field of view. If a uniform resolution over the entire field of view is desired, 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 field curvature, 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, a suitable set of deblurring filters could be explored. Alternatively, these effects can be significantly reduced using inexpensive corrective lenses. We have recently used such corrective lenses to obtain excellent results.

Figure 6: 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.

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. This project was supported in parts by an NSF National Young Investigator Award, an ONR MURI Grant (No. N00014-95-1-0601), and a David and Lucile Packard Fellowship.

References

Baker, S. and S. K. Nayar (1998, January). Catadioptric Image Formation. *Proc. of International Conference on Computer Vision.*

Bogner, S. (1995, October). Introduction to Panoramic Imaging. *Proc. of IEEE SMC Conference*, 3100–3106.

Born, M. and E. Wolf (1965). *Principles of Optics.* London:Permagon.

Charles, J. R. (1987, April). How to Build and Use an All-Sky Camera. *Astronomy Magazine.*

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

Gluckman, J. and S. K. Nayar (1998, January). Egomotion and Omnidirectional Cameras. *Proc. of International Conference on Computer Vision.*

Goshtasby, A. and W. A. Gruver (1993). Design of a Single-Lens Stereo Camera System. *Pattern Recognition 26*(6), 923–937.

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

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

Kuban, D. P., H. L. Martin, S. D. Zimmermann, and N. Busico (1994, October). Omniview Motionless Camera Surveillance System. *United States Patent* (5,359,363).

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

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

Murphy, J. R. (1995, October). Application of Panoramic Imaging to a Teleoperated Lunar Rover. *Proc. of IEEE SMC Conference*, 3117–3121.

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

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

Nayar, S. K. (1997, June). Catadioptric Omnidirectional Camera. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition.*

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

Nene, S. A. and S. K. Nayar (1998, January). Stereo Using Mirrors. *Proc. of International Conference on Computer Vision.*

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

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

Rees, D. W. (1970, April). Panoramic Television Viewing System. *United States Patent* (3,505,465).

Slater, D. (1996). Panoramic photography with fish eye lenses. *Panorama, International Association of Panoramic Photograpbers 13*(2,3).

Slater, J. (1932, October). Photography with the whole sky lens. *American Photographer.*

Southwell, D., A. Basu, M. Fiala, and J. Reyda (1996, August). Panoramic Stereo. *Proc. of International Conference on Pattern Recognition.*

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

Xiong, Y. and K. Turkowski (1997, June). Creating Image-Based VR Using a Self Calibrating Fisheye Lens. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition.*

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

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

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