Portions of this paper appeared in the Proceedings of the IEEE Workshop on Omnidirectional Vision 2002 held in conjunction with ECCV 2002, Copenhagen, Denmark.

Multiview Panoramic Cameras Using a Mirror Pyramid

Kar-Han Tan
tankh@vision.ai.uiuc.edu

Hong Hua
honghua@uiuc.edu

Narendra Ahuja
ahuja@vision.ai.uiuc.edu

Computer Vision and Robotics Laboratory
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
http://vision.ai.uiuc.edu/~tankh/Camera

Abstract

A mirror pyramid consists of a set of flat mirror faces arranged around an axis of symmetry, inclined to form a pyramid. By strategically positioning a number of conventional cameras around a mirror pyramid, the viewpoints for the individual cameras' mirror images can be colocated at a single point within the pyramid, effectively forming a virtual camera with a wide field of view. Mirror pyramid-based panoramic cameras have a number of attractive properties, including single-viewpoint imaging, high resolution and video rate capture. Currently existing designs realize a single viewpoint within each mirror pyramid. In order to capture panoramas from multiple viewpoints with these designs, the entire physical setup would need to be relocated or duplicated. The former solution lacks the capability of video rate imaging, and the latter leads to bulky designs due to the multiple mirror pyramids. In this paper we propose a method for generalizing existing designs such that multiple viewpoints can be created in a single mirror pyramid. This enables simultaneous multiview panoramic video rate imaging with a compact design. We describe the generalized design, the imaging geometry and the cameras' field of view(FOV), using simulations to illustrate the concepts. We also show the results obtained from an experimental prototype for two viewpoints.

1 Introduction

Panoramic images and video are useful in many applications such as special effects, immersive virtual reality environments, and video games. In recent years, the subject has been actively investigated by a number of researchers [5]. Among the numerous devices proposed for capturing panoramas, mirror pyramid-based camera systems are a promising approach for video rate capture, as they offer single-viewpoint imaging, and use only flat mirrors that are easier to produce than curved mirrors. To date, the designs proposed typically capture panoramas from a single viewpoint. In order to capture panoramas from multiple viewpoints, it would be necessary to do so sequentially using a single camera system or have multiple systems located at all viewpoints to operate in parallel. Obviously the sequential solution will not be able to capture consistent panoramas when the scene is not static. On the other hand, the parallel solution results in bulky designs as there would need to be one mirror pyramid per viewpoint, but with adjacent viewpoints separated by a sufficiently large distance. In this paper, we show that it is possible to create multiple viewpoints with only a single mirror pyramid while retaining the ability to capture panoramas at video rates.

2 Previous Work

Techniques for constructing panoramic cameras can be classified into two categories: dioptric methods, where only refractive elements (such as lenses) are employed, and catadioptric methods, where reflective components (such as mirrors) are used in combination with refractive elements. Dioptric systems include camera clusters [4, 26, 27], fish eye lens-based systems [18, 28, 32], and rotating cameras [6, 15, 17, 16, 1, 3, 2, 23]. Catadioptric systems include sensors that use curved mirrors and a single camera [7, 8, 10, 11, 13, 20, 22, 29, 31, 24], and sensors that employ planar mirrors and multiple cameras described in the next subsection.

Dioptric camera clusters, in which multiple cameras point in different directions to achieve a large FOV, are capable of achieving high resolution panoramic video rate capture. However, cameras in these clusters typically do not share a unique viewpoint due to physical constraints, which makes it impossible to mosaic individual images to form a true panoramic view. Although apparent continuity across images may be achieved by ad hoc image blending, panoramas produced as such are not suitable for some machine vision tasks that need images to be captured through a single viewpoint. Systems using a fisheye lens are able to deliver large FOV images at video rate, but have limited sensor resolution as the entire FOV is covered by a single sensor. Fisheye lenses also introduce irreversible distortion for close-by objects and may have different viewpoints for different portions of the FOV. Rotating cameras, in which a conventional camera rotates about its viewpoint to acquire panoramic images, deliver high-resolution wide FOV but are not capable of video rate panoramic capture.

Catadioptric systems that use a curved mirror to map a panoramic view onto a single sensor are able to achieve a single viewpoint at video rate, but have the same limitation on sensor resolution as fisheye lens-based systems. Furthermore, the resolution varies with the viewing direction across the FOV. Similar to the dioptric case, this resolution limitation can be overcome at the expense of video rate capture capability by panning the camera system [9, 21, 25] .

A mirror pyramid camera system, first described in [19], consists of a number of flat mirror surfaces arranged in the form of a pyramid together with a set of conventional cameras each associated with a face on the mirror pyramid. These cameras are strategically positioned such that the mirror images of their viewpoints are colocated at a single point within the mirror pyramid. Effectively this creates a virtual camera with a wide FOV that is capable of capturing panoramas at video rates. The first mirror pyramid camera design colocates the viewpoints for the conventional cameras at a point on the main axis of the pyramid, between the apex and the base plane. A recently-proposed camera system [12] uses a double mirror pyramid (two mirror pyramids sharing a common base plane), and colocates the viewpoints at the intersection point of the main axis and the base plane. This doubles the vertical FOV. An attempt at creating a stereoscopic (two-view) mirror pyramid camera uses two vertically-stacked mirror pyramids, and colocate two viewpoints, one in each pyramid, effectively duplicating the arrangement in the first design [14]. Although this creates two panoramic viewpoints, the two views are vertically displaced, i.e., displaced in a direction orthogonal to the panoramic strip. In many applications, it is more useful to have the camera displacement be aligned with the direction of the panoramic strip, i.e., in the commonly encountered mode of stereo vision.

There do not appear to be any existing mirror pyramid camera designs that allow the capture of horizontal panoramas from multiple horizontally displaced viewpoints. One possible solution would be to have two single-viewpoint mirror pyramid cameras located side-by-side. Alternatively, a single mirror pyramid camera could be relocated to sequentially capture the panoramas at each viewpoint if the scene is stationary. Obviously, the second solution will not be capable of video rate capture, and the first would result in bulkier designs since there would be two mirror pyramids next to each other. More importantly, each mirror pyramid would occlude a part of the other mirror pyramid's FOV.

3 Multiview Mirror Pyramid Cameras

This paper proposes a mirror pyramid camera design that allows two or more more viewpoints to be colocated horizontally within one mirror pyramid. In addition, the viewpoints can be placed in arbitrary spatial configurations within the mirror pyramid so that, for example, three viewpoints can be configured such that they lie in a plane inclined at an arbitrary angle to the base plane, or four viewpoints can be configured in an irregular tetrahedron with arbitrary orientation. Of course, it is also possible to place two viewpoints horizontally within a single pyramid. Essentially, each viewpoint position within a mirror pyramid dictates the position of a set of conventional cameras around the pyramid. A designer can thus start with the desired spatial configuration of the viewpoints and work out the required configuration of the set of conventional cameras. Video rate imaging can then be achieved from all viewpoints.

In the following sections, we will describe the proposed class of mirror pyramid cameras, starting with a characterization of mirror pyramids. We then examine the relation between a viewpoint inside a mirror pyramid and the position of its corresponding set of conventional cameras around the pyramid. Subsequent sections show how, for each conventional camera, the focal length and orientation can be chosen to maximize utilization of each camera's optical sensor. We then discuss the tradeoffs and limitations and show the results obtained from an experimental prototype that uses four conventional cameras to form two viewpoints.

3.1 Properties of Mirror Pyramids

We now describe the class of symmetric mirror pyramids considered in this paper, and used either alone or as a part of double mirror pyramids. Any such mirror pyramid can be fully characterized by the following parameters: radius, tilt angle, height, and the number of faces. Radius refers to the perpendicular distance from the main axis to the line of intersection of each planar mirror face with the base of the pyramid. Tilt angle refers to the angle between each mirror face plane and the base plane. If the pyramid is

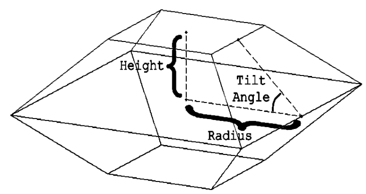


Figure 1: The geometry of a mirror pyramid.

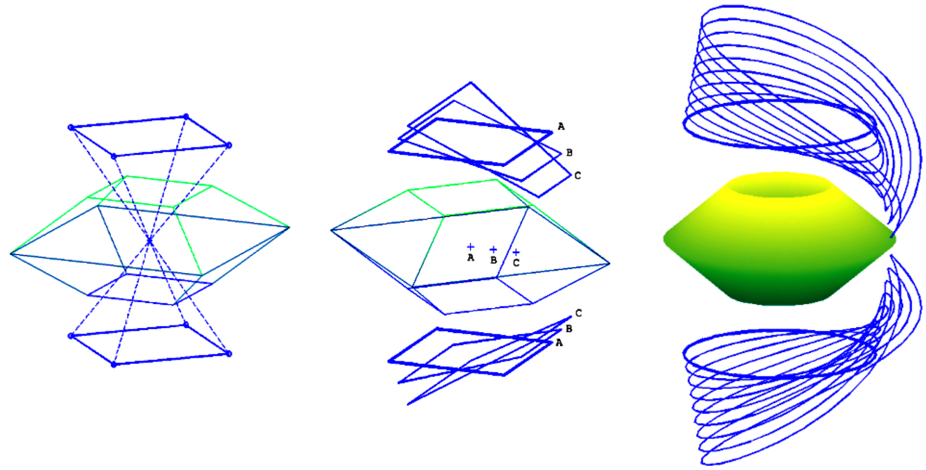
not truncated, all the mirror faces will intersect at the apex of the pyramid. If the pyramid is truncated, then the distance between the truncation plane and the base plane is called its height. Finally, the number of faces refers to the number of mirror faces in a single pyramid (twice as many in a corresponding double pyramid). Figure 1 illustrates the geometry involved.

3.2 Individual Viewpoint Placement

As mentioned earlier, previous designs of mirror pyramid cameras locate the viewpoint on the axis of symmetry of the pyramid. This viewpoint is placed at the base of the pyramid in a double mirror pyramid, and at a distance away from the base in a single mirror pyramid. In this section we will show how a viewpoint can be placed at an arbitrary location within the pyramid.

This can be easily accomplished by starting with a viewpoint and projecting its image into the physical world by finding the reflections of the viewpoint in the planes containing each respective mirror face. Each such projection is the location of the viewpoint of the physical camera associated with the corresponding pyramid face. An example is shown in Figure 2(a), in which a four-sided double mirror pyramid is used to create a viewpoint at its center. In the figure, dotted lines join the viewpoint and its corresponding physical camera positions for each mirror face.

When the viewpoint is on the main axis of the pyramid, symmetry causes the positions of the cameras for each of the upper and lower pyramids to form the vertices of a regular polygon. As the viewpoint is shifted away from the center, the polygonal shape changes. Figure 2(b) illustrates the effect with the same mirror pyramid as in Figure 2(a), and Figure 2(c) illustrates the effect of a shift in the viewpoint for a pyramid with a very large number of faces, and showing how the shape deforms as the viewpoint approaches the outer edge of the pyramid. It can be seen from this last diagram that



(a) (b) (c)

Figure 2: Variation in the physical camera position with viewpoint position. (a) Viewpoint is centered in four-sided pyramid, shown with the corresponding eight camera positions. (b) Translated viewpoints marked A, B, and C are shown with correspondingly marked physical camera positions. (c) Same as (b), but with a mirror pyramid with a large number of faces to show how the shape changes as the viewpoint translates.

the initial almost circular (approaching a circle for an arbitrarily large number of faces) shape smoothly deform into an irregular non-planar shape as the viewpoint shifts away from the center. A similar deformation will occur for the case of a small number of faces. The practical implication of this for camera designers is that if it is necessary for a mirror pyramid camera to change the position of a viewpoint on-the-fly, the camera mounting mechanism would have to take into account this irregular deformation.

3.3 Physical Camera FOV Determination

After placing the physical cameras at the locations dictated by the viewpoint in a given mirror pyramid, we need to determine the orientation and focal length of each physical camera, which together with the size of the camera CCD sensor determines the field of view (FOV) of each camera. The minimum requirement here is for each camera to be able to capture, on its sensor, a complete image of its corresponding mirror face. These mirror face images from the cameras can then be mosaiced to form a panoramic

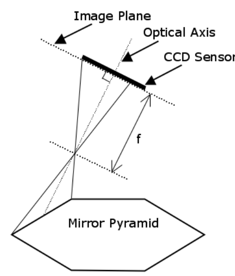


Figure 3: The projection geometry of mirror faces onto optical sensors.

image. This requirement may not ensure optimal use of the available camera sensors. For example, the image of a mirror face on a sensor may be smaller than the sensor, and thus not fully utilize the sensor. In this section, we present a method for finding an optimal orientation, and the maximum allowed focal length for each camera. The procedure first determinines the camera orientation and then uses the orientation found to determine the maximum allowed focal length. Before we proceed, it is important to note that in this section we are concerned only with the shape and size of the mirrors, and their projection onto camera sensors under perspective viewing transformation. Therefore when we refer to a 'mirror face' and its image on the camera sensor, we are refering only to the shapes and sizes of the mirrors, and not the photometric variations captured on the camera sensors.

3.3.1 Focal Length

Figure 3 illustrates the projection geometry of a mirror face onto the sensor plane of a physical camera. The CCD sensor is shown to be a thick black line that is perpendicular to the camera optical axis and parallel to the image plane which coincides with the sensor plane. It can be seen that the sensor captures only a portion of the infinite image plane. In commercially available cameras, the sensor typically covers a rectangular region that is approximately centered at the point where the optical axis intersects the image plane. The focal length, together with the size of the sensor, then determines the effective field of view of the camera. Given a particular orientation and position of a camera, we can then find the largest focal length such that the mirror face image is contained in the sensors capture area. Assuming that the sensor is rectangular and its sides are aligned with the axes of the frame of reference, the procedure to determine the focal

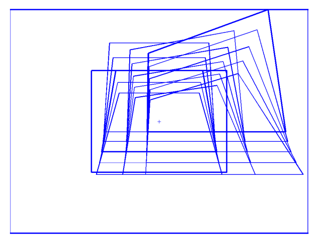


Figure 4: Variations in the mirror face image as the camera orientation changes.

length is as follows:

1. Given a camera orientation, find the image of the corresponding mirror face on the image plane by projecting the four vertices of the mirror face.

2. Find the smallest axis-aligned rectangle on the image plane centered at the optical axis that contains the mirror face image.

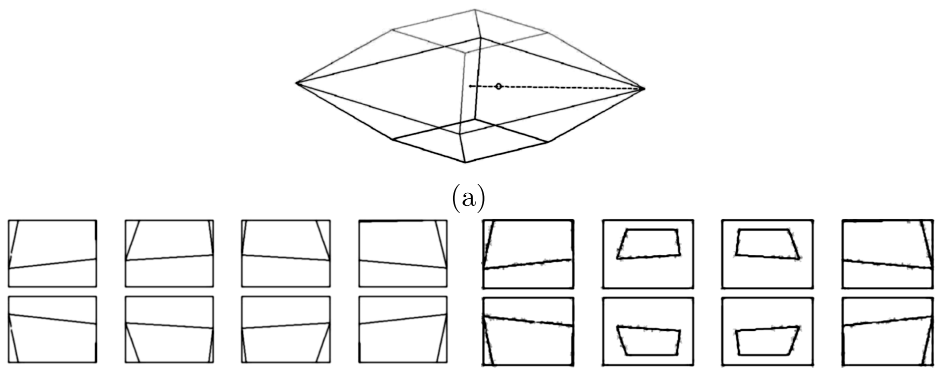
3. Find the smallest axis-aligned rectangle with the same aspect ratio as the sensor that contains the rectangle found from the previous step.

4. Find the focal length that produces this rectangle for the given size of the sensor.

Figure 4 shows an example of the range of mirror face images projected on the sensor plane as a camera's orientation varies. The figure also shows two of the bounding rectangles corresponding to the maximum allowed focal length for the smallest and largest mirror face image. It can be seen that changes in orientation affect the shape, size, and location of the face image within the sensor capture area.

3.3.2 Orientation

It should be noted that the solution for the focal length is uniquely determined for each given orientation and position of the camera, and the size of the sensor. Thus for each orientation and position, we can effectively estimate the maximum possible usage of the sensor by finding the image of the mirror face on the sensor using the maximum possible focal length. Given



(a)

(b) (c)

Figure 5: Optimizing the FOV of each camera. (a)Mirror pyramid and viewpoint. (b) FOV optimized for each camera. (c) Enforcing the uniform-resolution constraint by using a constant focal length for all cameras.

that each mirror face is a quadrilateral, we can find the projection of the four vertices of each mirror face and compute its area. Now we can search for a set of orientation parameters that makes maximum utilization of the sensor in each camera. We define utilization as

$$Utilization = \frac{F}{S}.$$

where
F = area of mirror face image on sensor
S = area of sensor

We then find the orientation for each camera that maximizes its utilization. The results of this optimization is illustrated in Figure 5. In Figure 5(a), a mirror pyramid is shown with a viewpoint that is shifted from the center. Figure 5(b) shows the results of optimizing the FOV for each camera : each rectangle represents the effective sensor capture area, and the quadrilaterals are the images of the respective mirror faces.

3.3.3 The Uniform-Resolution Constraint

While the sensor utilization in each camera can be maximized by varying the focal length, it is not always desired. One of the drawbacks of optimizing the focal length is that the sensor resolution per unit solid angle now varies among cameras. When the uniform-resolution property is desired, we can simply find the minimum among all the optimal focal lengths found, and use the same focal length for all cameras. The result of imposing this uniform-resolution constraint is shown in Figure 5(c). It can be seen that some of the sensors are not fully utilized. However, now the resolution is constant across the entired panoramic image captured.

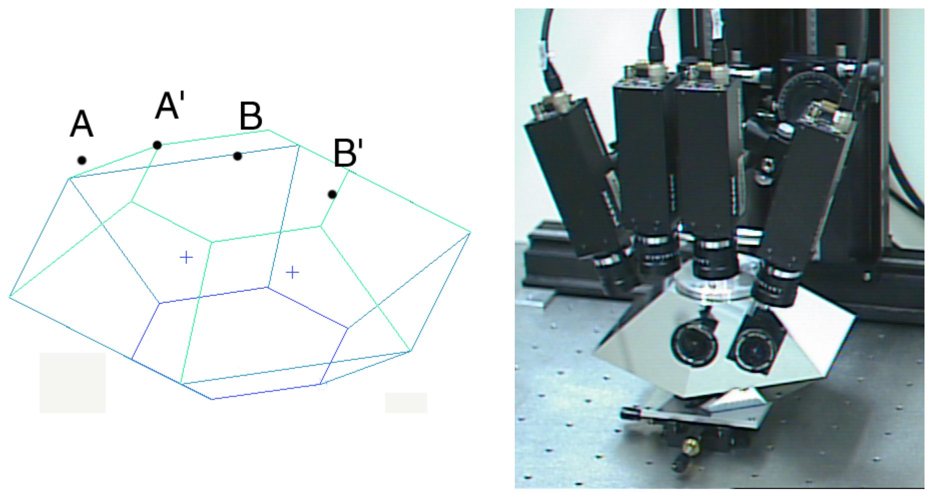
3.4 Multiview Setup Considerations

Having described the method for placing individual viewpoints at arbitrary locations within a mirror pyramid, we now discuss the issues involved in the placement of multiple viewpoints within the same pyramid. Each additional viewpoint adds a new set of physical cameras configured by the method discussed in the previous section. A fact of interest to the camera designer is that the locations of the physical cameras corresponding to different viewpoints and a given mirror face are simply the mirror images of the locations of the viewpoints in the face. This means that the relative distances and angles among the physical camera positions for a given face are replicated for the other faces. Further, if a set of viewpoints within the pyramid undergos a rigid transformation the corresponding camera configuration also undergo a rigid transformation which is given by reflections of the viewpoints into the corresponding faces. It should be noted, however, that this invariance property applies only to the camera positions and not the FOV-maximizing orientations.

The number of viewpoints and their spatial configuration will also have another constraint arising from the need to place the physical cameras around the mirror pyramid, which of course will depend on the actual physical size, shape, and orientations, and locations of the cameras, and the size and shape of the mirror pyramid.

4 Implementation

We implemented a stereo (two-view) mirror pyramid camera system that utilizes two conventional monochrome cameras for each viewpoint, as shown in Figure 6. We used an experimental setup that does not allow all the degrees of freedom required for optimized performance, as discussed in previous sections. We had limitations on the achievable camera orientations, and also employed lenses with equal focal lengths on all the cameras. The most significant implication of these limitations is that the usage of the individual sensors may not be optimized, as described in the previous sections. However, the setup proves that it is possible to create a mirror pyramid camera with more than a single viewpoint, each located at an arbitrary position



(a) (b)

Figure 6: (a) A design showing a 6-face pyramid and only two pairs of cameras (AA', BB'). Each pair is associated with one face and the two viewpoints shown as crosses(+). (b) The experimental setup implementing the design in (a).

within the mirror pyramid. In contrast, all previous mirror pyramid camera designs have only one viewpoint and need to place the viewpoint on the axis of symmetry. As a result, the techniques described in this paper enables the design of an entirely new class of multiview mirror pyramid cameras.

To see that we have achieved a stereo configuration in Figure 6, note that the images of the four cameras in the figure form a pair of cameras pointing outwards from within the mirror pyramid. In our experiments, the intrinsic parameters and the radial distortion are estimated and compensated for using the camera calibration software described in [30]. Figure 7 shows the images captured by each individual camera (after radial distortion compensation). The mosaiced images are shown in Figure 8. It may appear from the experimental results that the setup shown makes suboptimal utilization of the sensors and it might even be possible to obtain the same results using a pair of conventional cameras. However one should remember that the experimental setup utilizes only two faces of the mirror pyramid and is not fully optimized in the manner described in section 3.3. If a full set of cameras were used, even this limited setup would still be able to capture 360-degree panoramas, which is beyond the capabilites of a conventional camera.

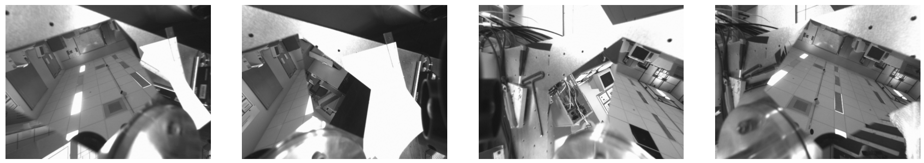


Figure 7: The raw images captured by the four cameras, after correcting for radial distortion.

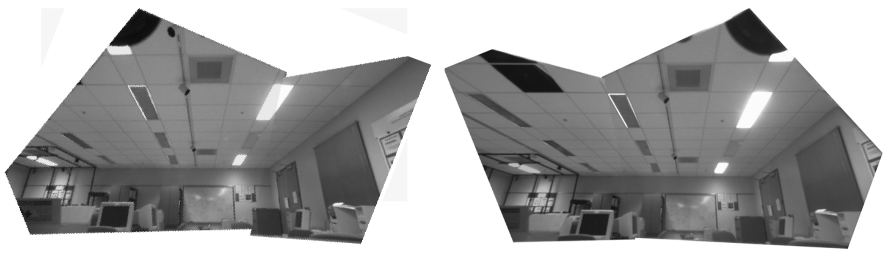


Figure 8: The stereo pair of panoramic views captured with the experimental setup.

5 Conclusion and Future Work

By observing that it is possible to place viewpoints inside a mirror pyramid in arbitrary positions, we have shown how a new class of multiview mirror pyramid cameras can be designed. We have studied the impact of viewpoint shifting on the placement of the conventional cameras around the pyramid, and experimentally demonstrated the feasibility of a two-view mirror pyramid camera. In our ongoing efforts, we are investigating the use of these multiview mirror pyramid cameras in areas like robot navigation and immersive telepresence. In addition, an optical artifact noted in [12] also needs to be addressed.

References

[1] M. Aggarwal and N. Ahuja. On generating seamless mosaics with large depth of field. In *International Conference on Pattern Recognition*, 2000.

[2] M. Aggarwal and N. Ahuja. High dynamic range panoramic imaging. In *IEEE International Conference on Computer Vision*, 2001.

[3] M. Aggarwal and N. Ahuja. A new imaging model. In *IEEE International Conference on Computer Vision*, 2001.

[4] P. I. Anderson. From telepresence to true immersive imaging: into real-life video-now! *Advanced Imaging*, 10(7):48–50, 1995.

[5] R. Benosman, S. B. Kang, and O. Faugeras, editors. *Panoramic Vision*. Springer-Verlag, 2001.

[6] A. Castano and N. Ahuja. Omnifocused 3d display using the nonfrontal imaging camera. In *Workshop on Computer Vision for Virtual Reality Based Human Comms.*, 1998.

[7] J. S. Chahl and M. V. Srinivasan. Reflective surfaces for panoramic imaging. *Applied Optics*, 36(31):8275–8285, 1997.

[8] J. S. Chahl and M. V. Srinivasan. A complete panoramic vision system, incorporating imaging, ranging, and three dimensional navigation. In *Workshop on Omnidirectional Vision*, pages 104–111, 2000.

[9] S. Coorg, N. Master, and S. Teller. Acquisition of a large pose-mosaic dataset. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998.

[10] R. A. Hicks. Reflective surfaces as computational sensors. In *Workshop on Perception for Mobile Agents*, 1999.

[11] R. A. Hicks and R. Bajcsy. Catadioptric sensors that approximate wide-angle perspective projections. In *Workshop on Omnidirectional Vision*, pages 97–103, 2000.

[12] H. Hua and N. Ahuja. A uniform-resolution panoramic camera. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.

[13] H. Ishiguro, M. Yamamoto, and S. Tsuji. Omni-directional stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):257–262, 1992.

[14] T. Kawanishi and e. a. Kazumasa Yamazawa. Generation of high resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors. In *International Conference on Pattern Recognition*, 1998.

[15] A. Krishnan and N. Ahuja. Range estimation from focus using a non-frontal imaging camera. In *National Conference on Artificial Intelligence*, 1993.

[16] A. Krishnan and N. Ahuja. Panoramic image acquisition. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1996.

[17] A. Krishnan and N. Ahuja. Range estimation from focus using a nonfrontal imaging camera. *International Journal of Computer Vision*, 20(3):169–185, 1996.

[18] K. Miyamoto. Fish eye lens. *Journal of Optical Society of America*, 64:1060–1061, 1964.

[19] V. Nalwa. A true omnidirectional viewer. *Bell Laboratories Technical Report*, 1996.

[20] S. K. Nayar. Catadioptric omnidirectional camera. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997.

[21] S. K. Nayar and A. Karmarkar. 360x360 mosaics. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2000.

[22] S. K. Nayar and V. Peri. Folded catadioptric cameras. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1999.

[23] S. Peleg. Panoramic mosaics by manifold projection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997.

[24] S. Peleg, M. Ben-Ezra, and Y. Pritch. Omnistereo: Panoramic stereo imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):279–290, 2001.

[25] H.-Y. Shum and R. Szeliski. Panoramic image mosaics. *Microsoft Research Technical Report*, MSR-TR-97-23, 1997.

[26] R. Swaminathan and S. K. Nayar. Polycameras: Camera clusters for wide angle imaging. *Columbia University Department of Computer Science Technical Report*, CUCS-013-99, 1999.

[27] R. Swaminathan and S. K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10), 2000.

[28] Y. Xiong and K. Turkowski. Creating image-based virtual reality using a self-calibrating fisheye lens. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997.

[29] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional imaging with hyperboloidal projection. In *International Conference on Intelligent Robots and Systems*, 1993.

[30] Z. Zhang. A flexible new technique for camera calibration. *Microsoft Research Technical Report*, MSR-TR-98-71, 1998.

[31] Z. Zhu, K. D. Rajasekar, E. M. Riseman, and A. R. Hanson. Panoramic virtual stereo vision of cooperative mobile robots for localizing 3d moving objects. In *Workshop on Omnidirectional Vision*, 2000.

[32] S. Zimmerman and D. Kuban. A video pan/tilt/magnify/rotate system with no moving parts. In *IEEE/AIAA Digital Avionics Systems Conference*, 1992.