Single-camera panoramic stereo system with single-viewpoint optics

Gijeong Jang, Sungho Kim, and Inso Kweon

Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology,
373-1, Guseong-dong, Yuseong-gu, Daejeon, Korea

A simple and successful design for a panoramic stereo system that uses a single camera and multiple mirrors is presented. A wide-baseline stereo enhances the three-dimensional reconstruction accuracy even with a single camera. The feasibility of the system as a practical stereo sensor has been demonstrated with experiments in an indoor environment. © 2006 Optical Society of America
OCIS codes: 150.6910, 100.6890.

With stereo vision, one accomplishes depth perception by establishing stereo disparity between the two images from two distinct viewpoints. Many possible types of stereo vision system have been proposed. Among them, the single-camera stereo is especially attractive. In general, the characteristics of the stereo cameras are slightly different. For stereo images acquired by the two cameras, the focal lengths of the cameras are not exactly the same, and the alignment of the imaging sensors is not accurate. Moreover, differences in the characteristics of the two imaging sensors cause intensity differences between the corresponding points in stereo images. Using the single-camera stereo system, one can eliminate these unwanted geometric and chromatic differences to increase the ability to find correspondence reliably. A single-camera stereo system that uses a couple of planar mirrors has been suggested.[1] This stereo system has been modified such that the corresponding features lie on the same scan line automatically.[2] The system's complexity has been largely reduced by use of an image-splitting prism, which is called a biprism.[3] However, this system suffers from a short baseline, a narrow field of view, and chromatic aberration.

The panoramic stereo is another attractive method in that it provides a full horizontal field of view (FOV). A catadioptric vision system that uses mirrors has been a popular means to get panoramic images.[4] In general, this system consists of coaxially aligned conic mirrors and cameras. A stereo system that coaxially combines a couple of catadioptric vision systems was proposed.[5] However, it is bulky, its two cameras have different physical characteristics, and there is a slight misalignment between the two systems. One partial solution to these problems uses double conic mirrors positioned in front of a single camera.[6] This system reduces the number of cameras required, but suffers from such problems as difficult calibration and a complex stereo matching process. Another partial solution is a catadioptric stereo sensor, which uses a cone mirror and a beam splitter.[7] This system reduces the number of mirrors by using a beam splitter, but the two orthogonally positioned cameras still cause the system to suffer from bulky size, camera misalignment, and different characteristics of the cameras. A novel solution to these problems is to use a single camera and a double-lobed mirror,[8–10] which is a combined coaxial mirror pair. However the critical deficiency of that system is a short baseline, from which an accurate three-dimensional (3D) reconstruction is impossible.

In this Letter we propose a compact, wide-baseline panoramic stereo system, for which the camera alignment procedure and the complex search procedure for the epipolar line are not necessary. Additionally, the system has all the advantages of the single-camera stereo system.

For a wide FOV, conic mirrors are useful. In particular, hyperboloidal mirrors, which have two foci, satisfy the single-viewpoint constraint, with which the entire light ray directed to the focus inside the mirror (the effective viewpoint) reflects and converges at the other focus (the effective pinhole).[4] Figure 1 shows the proposed system structure, which is composed of a single perspective camera, two hyperboloidal mirrors, and one planar mirror. For the generation of the upper view, the light ray from the world point going to the first focus (F_1) of the first
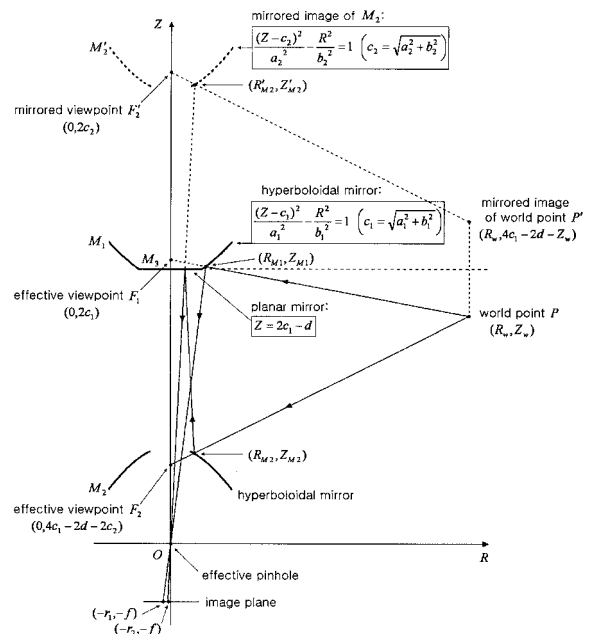

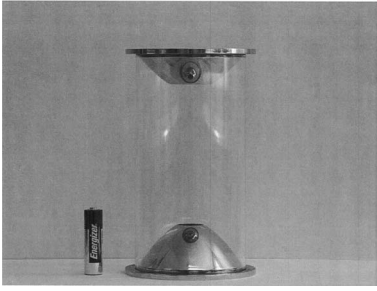Fig. 1. Schematic of the proposed single-camera panoramic stereo system.

Fig. 2.   Experimental single-camera panoramic stereo system.

mirror (M_1) is reflected to the second focus, which is also an effective pinhole (O), and then projected onto the outer rim of the imaging circle. For generation of the lower view, the light ray from the world point going to the mirrored focus (F_2) of the second mirror (M_2) is reflected and then folded by planar mirror M_3 to the second focus (O) and then projected onto the central region of the imaging circle. Because the projection scheme of the lower view includes the folding of the light ray by a planar mirror, one can simplify the relationship between the world point and its projection by unfolding the reflection mechanism.

Given the parameters of the hyperboloids $(a_1,b_1,c_1$ and $a_2,b_2,c_2)$, the focal length of the camera (f), and the offset of the planar mirror (d), the relationship between the world point (P) and its projection onto the image plane (r_1,r_2) can be established by

$$r_1 = f\frac{R_{M1}}{Z_{M1}}, \quad r_2 = f\frac{R_{M2}'}{Z_{M2}'} \qquad (1)$$

where

$$R_{M1} = \frac{m_1 c_1 + a_1(1 + m_1{}^2)^{1/2}}{a_1{}^2/b_1{}^2 - m_1{}^2},$$

$$R_{M2}' = \frac{m_2 c_2 + a_2(1 + m_2{}^2)^{1/2}}{a_2{}^2/b_2{}^2 - m_2{}^2}, \qquad (2)$$

$$Z_{M1} = m_1 R_{M1} + 2c_1, \quad Z_{M2}' = m_2 R_{M2}' + 2c_2. \qquad (3)$$

Here m_1 is the slope of the line from the world point to the focus of mirror M_1 and m_2 is the slope of the line from the mirrored image of the world point to the focus of mirror M_2':

$$m_1 = (Z_w - 2c_1)/R_w,$$

$$m_2 = [(4c_1 - 2d - Z_w) - 2c_2]/R_w \quad (R_w > 0), \qquad (4)$$

where the possible range of m_1 and m_2 means the vertical FOV for each view, which is determined by the mirror design.

The proposed catadioptric stereo system is designed to keep the single-viewpoint constraint because it permits easy conversion from the catadioptric omnidirectional image to the panoramic image seen from a single effective viewpoint. If we assume that a virtual cylindrical plane is located at each vir-

tual viewpoint and aligned to the axis of the system, one can find correspondence between the point on the virtual cylindrical image plane and its projection onto the planar image plane by applying Eqs. (1)–(4). Then reprojection from the catadioptric image to the virtual cylindrical image pair can be accomplished by sampling intensity from the catadioptric image for all the pixels on the virtual cylindrical image, and these cylindrical projections will finally produce rectified panoramic stereo images. This rectification simplifies a stereo matching problem into a simple one-dimensional search along the vertically collinear scan line. Once the point correspondences between stereo images have been established, a 3D estimation of the world point can be made by simple triangulation.[5]

However, the proposed system has some problems to be considered. The first problem is defocusing. Because the system is a combination of two different catadioptric systems, the focusing ranges of the two systems are largely different. Even for a single view, the focusing range varies radially. A partial solution is to increase the depth of field by reducing the aperture size at the cost of reducing shutter speed. The second problem is the difference in image resolution between generated panoramic stereo images. It reduces the stereo matching accuracy if the resolution of the input image is not so high. A partial solution is to match the resolution by applying Gaussian smoothing with different kernel shapes for the corresponding coordinates of the stereo images.

To demonstrate the feasibility of the system, we built the experimental system setup shown in Fig. 2. The length of the baseline was set to 120 mm, which is adequate for indoor use. The diameter of the mirror was set to 80 mm to satisfy the baseline and camera's FOV. By adjusting the mirror parameters, we set the vertical FOV to slightly more than 16° for both upward and downward views, and this FOV was applied for both upper and lower views to guarantee a common FOV for stereo imaging. A 1/1.8 in. (1 in. =2.54 cm) CCD camera with 1200 horizontal scan
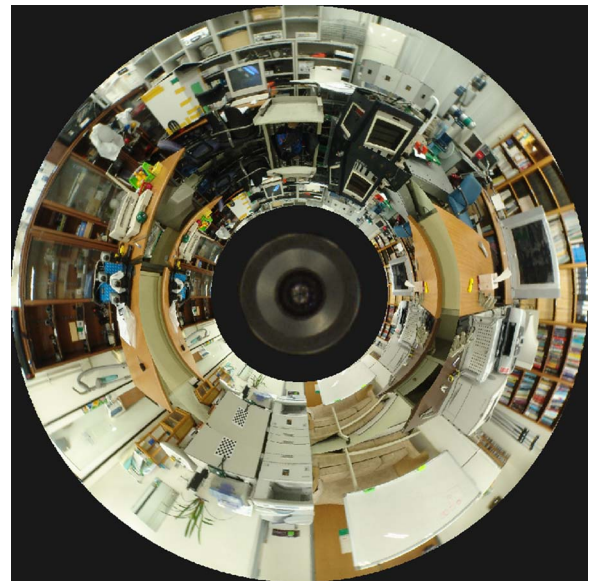

Fig. 3.   Stereo image captured by the proposed system.
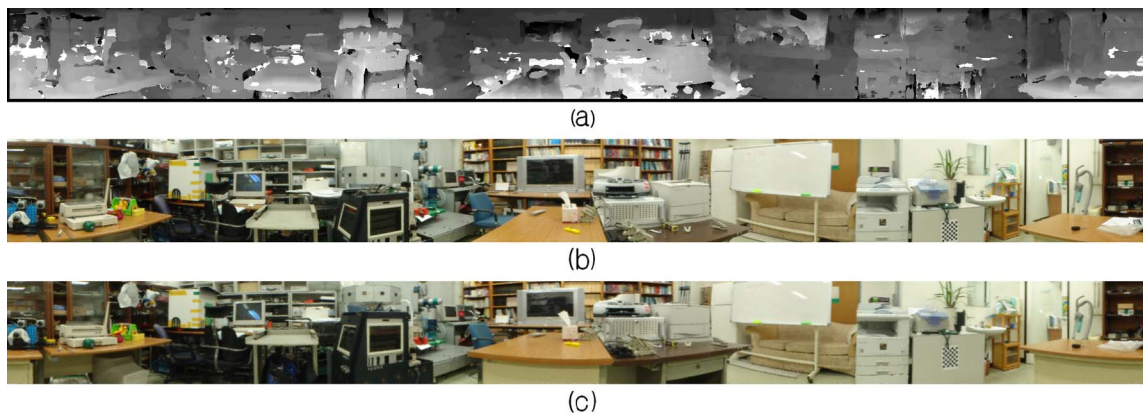
Fig. 4. (b), (c) Panoramic stereo images rectified from Fig. 3 and (a) the corresponding disparity map.

Fig. 5. Reconstructed 3D map generated from Fig. 4 and the corresponding 3D sensation as the viewpoint changes.

lines and a lens with 12 mm focal length were used to capture the image.

Figure 3 shows a catadioptric stereo image of our laboratory captured by the given experimental setup. All the corresponding points of each view are positioned radially collinearly. Figures 4(b) and 4(c) constitute a rectified panoramic stereo pair generated from Fig. 3. Stereo matching was achieved with a simple window-based correlation search. Figure 4(a) is the resultant disparity map, in which nearby objects such as tables and a robot appear brighter than distant objects such as bookshelves and racks. This map estimates the overall structure of the environment in spite of some false matches caused by self-occlusion, specularity, homogeneity, and repetitive pattern. Figure 5 is the result of the 3D reconstruction with the panoramic images and the disparity map from Fig. 4. We can observe the 3D sensation as the viewing direction changes.

The result demonstrates that the proposed compact, wide-baseline panoramic stereo system with a single camera provides sufficient accuracy for various machine vision applications in which panoramic range measurements are important. This system can be widely used for robotic applications (environmental recognition, path planning, and avoidance of obstacles), virtual reality, surveillance systems, and military applications.

References

1. J. Gluckman and S. K. Nayar, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, 1999), p. 28.
2. J. Gluckman and S. K. Nayar, IEEE Trans. Pattern Anal. Mach. Intell. **24**, 224 (2002).
3. D. Lee and I. Kweon, IEEE Trans. Rob. Autom. **16**, 528 (2000).
4. S. Baker and S. Nayar, Int. J. Comput. Vis. **35**, 175 (1999).
5. J. Gluckman, S. K. Nayar, and K. J. Thoresz, "Real-time omnidirectional and panoramic stereo," in *Proceedings of the Image Understanding Workshop* (Defense Advanced Research Projects Agency, 1998), p. 299.
6. S. A. Nene and S. Nayar, in *Proceedings of the IEEE International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, 1998), p. 1087.
7. S. Lin and R. Bajcsy, in *Proceedings of the IEEE International Conference on Robotics and Automation* (Institute of Electrical and Electronics Engineers, 2003), p. 1694.
8. D. Southwell, A. Basu, M. Fiala, and J. Reyda, in *Proceedings of the International Conference on Pattern Recognition* (International Association for Pattern Recognition/Institute of Electrical and Electronics Engineers, 1996), p. 378.
9. M. Fiala and A. Basu, in *Proceedings of the International Conference on Pattern Recognition* (International Association for Pattern Recognition/Institute of Electrical and Electronics Engineers, 2002), p. 11.
10. E. L. L. Cabral, J. C. de Souza, Jr., and M. C. Hunold, in *Proceedings of the International Conference on Pattern Recognition* (International Association for Pattern Recognition/Institute of Electrical and Electronics Engineers, 2004), p. 1.