A Biprism-Stereo Camera System

Doo Hyun Lee*, In So Kweon**, Roberto Cipolla^

* Dept. of Automation and Design Engineering
** Dept. of Electrical Engineering, KAIST
373-1 Kusung-Dong Yusong-Ku
Taejon 305-701, Korea
{doolee,iskweon}@cais.kaist.ac.kr

^ Dept. of Engineering, University of Cambridge
Trumpington Street
Cambridge, CB2 1PZ, U.K.
cipolla@eng.cam.ac.uk

Abstract

In this paper, we propose a novel and practical stereo camera system that uses only one camera and a biprism placed in front of the camera. The equivalent of a stereo pair of images is formed as the left and right halves of a single CCD image using a biprism. The system is therefore cheap and extremely easy to calibrate since it requires only one CCD camera. An additional advantage of the geometrical set-up is that corresponding features lie on the same scanline automatically.

The single camera and biprism have led to a simple stereo system for which correspondence is very easy and which is accurate for nearby objects in a small field of view. Since we use only a single lens, calibration of the system is greatly simplified. This is due to the fact that we need to estimate only one focal length and one center of projection. Given the parameters in the biprism-stereo camera system, we can recover the depth of the object using only the disparity between the corresponding points.

1 Introduction

Research on the recovery and recognition of 3-D shapes has been undertaken using a monocular image and multiple views. Depth perception by stereo disparity has been studied extensively in computer vision. The stereo disparity between two images from two distinct viewpoints is a powerful cue to 3-D shapes and pose estimation [3, 4]. For the recovery of a 3-D scene from a pair of stereo image of the scene, it is required to establish correspondence [3].
A correspondence algorithm can produce more reliable matches if the underlying images have smaller intensity and geometric differences. Some geometric difference between stereo images is unavoidable, for it is actually the local geometric difference between stereo images that results in the perception of depth. For stereo images acquired by two cameras, the focal lengths and zoom levels of the cameras are often slightly different. Differences in the optical properties of the two cameras cause intensity differences between corresponding points in stereo images. These unwanted geometric and intensity differences should be reduced as much as possible to increase the ability to find correspondences reliably.

Nishimoto and Shirai [5] proposed a single-lens camera system that can obtain stereo images. Stereo images are obtained with a mirror at two different rotational positions. Teoh and Zhang [6] proposed a single-lens stereo camera system. The rotating mirror is made parallel to one of the fixed mirrors and an image is obtained. Then it is made parallel to the other fixed mirror and another image is obtained. Gosthasby and Gruver [7] proposed a single camera system that can obtain images in a single shot and through a single lens. The reversed image should be transformed to appear as if obtained by cameras with parallel optical axes, before carrying out the correspondence and measuring the depth values from the correspondence. In their recent work, Nene and Nayar [11] proposed four stereo systems that use a single camera pointed towards planar, ellipsoidal, hyperboloidal, and paraboloidal mirrors. By use of non-planar reflecting surfaces such as hyperboloids and paraboloids, a wide field of view is easily obtained. However, their stereo system needs a complex mirror mechanism.

In this paper, we propose a novel stereo camera system that can provide a pair of stereo images using a biprism. This camera system also has the advantage that unwanted geometric and intensity differences between the stereo images are reduced. An arbitrary object point in 3-D space is transformed into two virtual points by the biprism. As in the conventional stereo system, the image disparity between the two image points of the two virtual points is

0-7695-0149-4/99 $10.00 © 1999 IEEE

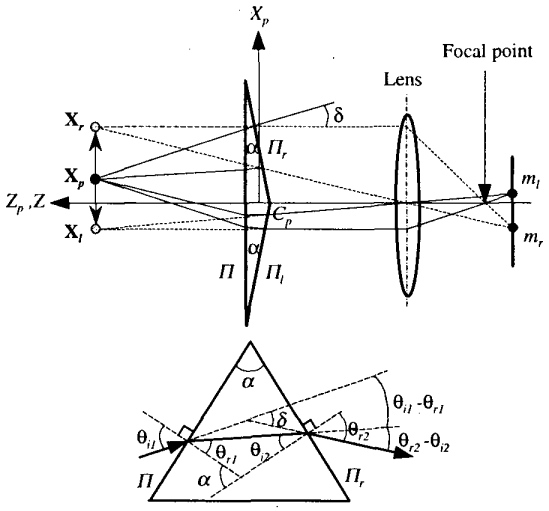
directly related to the depth of the object point.

2 Biprism Stereo System

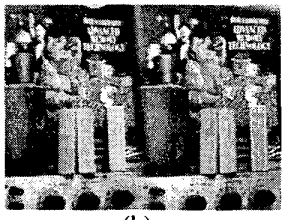
2.1 Principles of biprism

Fig. 1(a) shows the geometry of a biprism and the associated coordinate system. Both the inclined prism planes Π_l and Π_r make the angle α with the base plane Π, respectively. An arbitrary point in 3-D space $\mathbf{X}_p(X_p, Y_p, Z_p)$ is transformed into the two virtual points $\mathbf{X}_r(X_{pr}, Y_{pr}, Z_{pr})$ and $\mathbf{X}_l(X_{pl}, Y_{pl}, Z_{pl})$ by the inclined prism planes, respectively. That is, an object point in 3-D space is transformed into the two virtual points by the deviation δ, which is determined by the angle α and the index of refraction n of the biprism. From the geometry of the biprism as shown in Fig. 1(a), we obtain

$$\alpha = \theta_{r1} + \theta_{i2}, \delta = \theta_{i1} + \theta_{r2} - \alpha . \quad (1)$$



(a)



(b)

Fig. 1 Principles of the biprism: (a) The geometry of the biprism, (b) An image captured by a biprism-stereo system.

In Fig. 1(a), a lens gathers all the light radiating from the two virtual points $\mathbf{X}_r$ and $\mathbf{X}_l$, and creates two corresponding image points m_r and m_l. From equation (1), given the angle α and the refractive index n, the basic relation for the biprism is given by [9, 10]

$$n = \frac{\sin((\alpha + \delta)/2)}{\sin(\alpha/2)} \quad (2)$$

where δ is the angle between a 3-D point and one of the two virtual points. The deviation of biprism δ, as a function of α and n, defines the field of view of the biprism. In Fig. 1(a), the geometric relationship between the 3-D point (X_p, Y_p, Z_p) and the two virtual 3-D points, created by the biprism, can be represented by simple transformations:

1) For the virtual point $\mathbf{X}_r = (X_{pr}, Y_{pr}, Z_{pr})$

$$\begin{bmatrix} X_{pr} \\ Y_{pr} \\ Z_{pr} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \tan\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} = \mathbf{T}_r \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} \quad (3)$$

2) For the virtual point $\mathbf{X}_l = (X_{pl}, Y_{pl}, Z_{pl})$

$$\begin{bmatrix} X_{pl} \\ Y_{pl} \\ Z_{pl} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\tan\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} = \mathbf{T}_l \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} \quad (4)$$

where the matrix T_l and T_r describe the transformation matrix that transformed into the virtual points by a biprism. From equations (3) and (4), we obtain

$$D = X_{pr} - X_{pl} = 2Z_p \tan\delta . \quad (5)$$

Equation (5) indicates that the distance between the two virtual points becomes larger as a 3-D point moves farther away from the biprism.

2.2 Biprism-stereo projection matrix

We derive the camera projection matrix for the biprism-stereo camera in homogeneous coordinates. Let the coordinates of the biprism be the world coordinate system. There is a rigid body transformation between the world coordinates $\tilde{\mathbf{x}}_p$ and the camera-centered coordinates $\tilde{\mathbf{x}}_c$. Perspective projection and pixel coordinates conversions are both linear transformations in homogeneous coordinates. In addition, there is another linear transformation involved for the biprism, which accounts for generation of two virtual points $\mathbf{X}_r$ and $\mathbf{X}_l$. The overall imaging process for the biprism-stereo camera can be expressed as a single matrix multiplication in homogeneous coordinates:

$$s\tilde{\mathbf{m}}_l = \mathbf{P}_l\tilde{\mathbf{X}}_p, \quad s\tilde{\mathbf{m}}_r = \mathbf{P}_r\tilde{\mathbf{X}}_p \quad (6)$$

where $\tilde{\mathbf{m}}_l$ is the image coordinates in the left image plane, $\tilde{\mathbf{m}}_r$ is the image coordinates of its corresponding points in the right image plane. If the base plane of the biprism is parallel to the image plane and the biprism does not rotate about the Z-axis, the rotation matrix $\mathbf{R}$ is an identity matrix and $\mathbf{t}=[0, 0, t_z]^T$. The projection matrix in the biprism stereo camera system is defined as

$$\mathbf{P}_l=\begin{bmatrix}\alpha_u & 0 & u_0\\ 0 & \alpha_v & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\begin{bmatrix}\mathbf{T}_l & \mathbf{0}_3^T\\ \mathbf{0}_3 & 1\end{bmatrix}=\begin{bmatrix}\alpha_u & 0 & u_0-\alpha_u\tan\delta & u_0t_z\\ 0 & \alpha_v & v_0 & v_0t_z\\ 0 & 0 & 1 & t_z\end{bmatrix},$$

$$\mathbf{P}_r=\begin{bmatrix}\alpha_u & 0 & u_0\\ 0 & \alpha_v & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\begin{bmatrix}\mathbf{T}_r & \mathbf{0}_3^T\\ \mathbf{0}_3 & 1\end{bmatrix}=\begin{bmatrix}\alpha_u & 0 & u_0+\alpha_u\tan\delta & u_0t_z\\ 0 & \alpha_v & v_0 & v_0t_z\\ 0 & 0 & 1 & t_z\end{bmatrix}. \quad (7)$$

2.3 Relationship between disparity and depth

From equations (6) and (7), the relationship between $\tilde{\mathbf{X}}_p$ and $\tilde{\mathbf{m}}_l$ and $\tilde{\mathbf{m}}_r$ can be expressed as

$$\begin{bmatrix}u_l\\ v_l\end{bmatrix}=\frac{1}{Z_p+t_z}\begin{bmatrix}\alpha_u(X_p-Z_p\tan\delta)+u_0(Z_p+t_z)\\ \alpha_vY_p+v_0(Z_p+t_z)\end{bmatrix}, \quad (8)$$

$$\begin{bmatrix}u_r\\ v_r\end{bmatrix}=\frac{1}{Z_p+t_z}\begin{bmatrix}\alpha_u(X_p+Z_p\tan\delta)+u_0(Z_p+t_z)\\ \alpha_vY_p+v_0(Z_p+t_z)\end{bmatrix}. \quad (9)$$

From equations (8) and (9), it is easy to see that the corresponding points lie on the same scanline. Equivalently, epipolar lines in biprism-stereo images are parallel. We can derive a simple formula between the disparity and the depth. The relationship between the image disparity and the depth can be expressed as

$$d=u_r-u_l=\frac{2\alpha_u\tan\delta(Z-t_z)}{Z} \quad (10)$$

where $Z=Z_p+t_z$ and t_z is the distance between the origin of the biprism and the optical center of the camera. Equivalently, the relationship between the inverse disparity and the inverse depth can be expressed as

$$\frac{1}{d}=\frac{k_1}{Z_p}+k_2 \quad (11)$$

where $k_1=\dfrac{t_z}{2\alpha_u\tan\delta}$ and $k_2=\dfrac{1}{2\alpha_u\tan\delta}$.

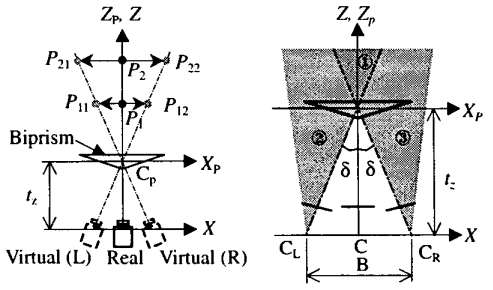
Once the corresponding points in the left and the right image of a biprism-image are found and the internal parameters of camera and biprism are known, we can determine the depth of points in the scene. From equation (10), we observe that the disparity is constant independent of the location in the scene if $t_z=0$.

2.4 Equivalent stereo system

Fig. 2 shows the equivalent stereo camera system to a biprism stereo system. The virtual cameras are located on the virtual optical axis represented by the line C_P~P_{22} and C_P~P_{21}. The optical center of two virtual cameras is positioned at points C_L and C_R, respectively. Accordingly, the effective baseline distance of the biprism-stereo, B, is defined by

$$B=2t_z\tan\delta=\frac{t_z}{k_2\alpha_u}=\frac{k_1}{k_2^2\alpha_u} \quad (12)$$

As illustrated in Fig. 3, the equivalent stereo images are coplanar and parallel to their baseline. In other words, the images are rectified.



(a) (b)

Fig. 2 The equivalent stereo system to the biprism stereo

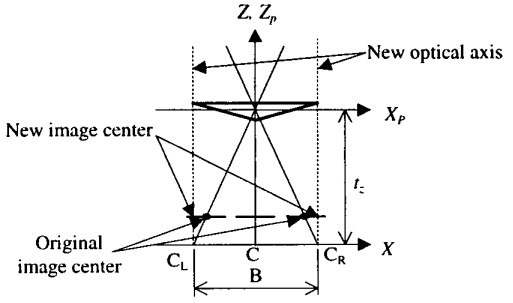


Fig. 3 The rectified images of two virtual cameras

2.5 Field of view

As shown in Fig. 2(b), the active image plane of the left and right virtual camera is formed by the left and right halves of the image plane, respectively. In other words, regions ① and ②, ① and ③ comprise the field of view of the left and right virtual camera. Overlap between the two fields of view defines the field of view (FOV) of the biprism, which is related to the deviation δ:

$$-\delta \le FOV \le \delta \quad (13)$$

A 3-D point in the region ① has two corresponding transformed points, whereas any points out of the FOV do not have corresponding points though they have the transformed points. The field of view of the biprism increases with the increase of the angle α of the biprism.

3 3-D reconstruction and calibration

It is straightforward to recover 3-D structure if the biprism-stereo camera is calibrated. For the calibration of the camera, we use a well-known calibration algorithm by Faugeras and Toscani [8].

Suppose we know the Euclidean distance d' between two reference points in 3D space. The biprism parameters can be obtained through the minimization of the following criterion:

$$\min \sum_{i=1}^{N} \left\| d_i^c(k_1, k_2) - d_i^r \right\|^2 \quad (14)$$

where $d_i^c(k_1, k_2)$ can be easily derived from equations (8), (9), and (10) and N denotes the number of the corresponding points.

To solve equation (14), initial guesses for k_1 and k_2 are obtained from the design parameters of the biprism. With the calibrated biprism and camera, recovering 3-D structure is straightforward. From equation (11), the distance between the origin of the biprism and the optical center of the camera becomes

$$t_z = \frac{k_1}{k_2}. \quad (15)$$

The depth of the object points in the camera coordinate system is given by:

$$Z = Z_p + t_z \quad (16)$$

Given Z and the intrinsic parameters of the camera, we can compute the coordinates X and Y using

$$\begin{bmatrix} X \\ Y \end{bmatrix} = Z \begin{bmatrix} \frac{u_c - u_0}{\alpha_u} & \frac{v_c - v_0}{\alpha_v} \end{bmatrix}^T \quad (17)$$

where

$$\begin{bmatrix} u_c \\ v_c \end{bmatrix} = \begin{bmatrix} \frac{u_r + u_l}{2} & \frac{v_r + v_l}{2} \end{bmatrix}^T.$$

Known k_1 and k_2 and the intrinsic parameters of the camera, 3-D structure in the camera coordinates can be recovered using the disparity and the center coordinates (u_c, v_c) of the corresponding points in the biprism stereo image.

4 Experimental Results

4.1 Disparity Map

This section presents some disparity maps of real objects, which are computed from correspondences found automatically in a biprism-image. To obtain a biprism-image, a biprism is placed in front of a camera as shown in Fig. 4.

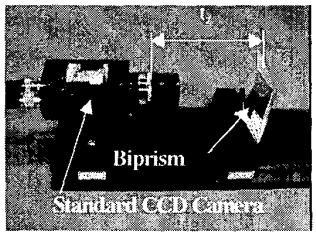


Fig. 4 An implementation of a biprism-stereo camera

For this particular prototype, the distance between the optical center and the biprism, t_z, is about 150 mm, and the biprism is designed to have a declination angle of 12.4 degrees. Therefore, the effective baseline distance is approximately 38.8 mm. As we explored in the previous section, the distance between two corresponding points $m_l(u_l, v_l)$ and $m_r(u_r, v_r)$ in the image plane is proportional to the distance between the two virtual points transformed by the biprism in the camera coordinates. A simple cross-correlation technique is used for matching. For this particular experiment, we use a 25×25 window to compute the SSD (Sum of Squared Differences). Fig. 5 and Fig. 6 show an input biprism image and the computed disparity map.

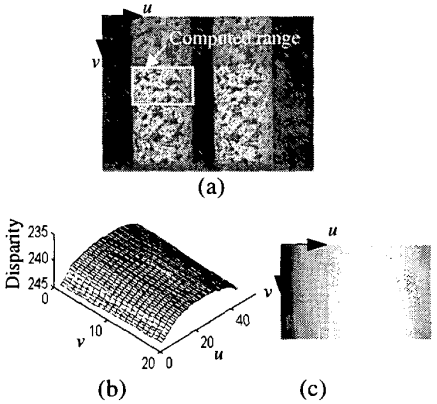


Fig. 5 A disparity map computed from the biprism stereo image for a cylindrical object: (a) An input biprism image, (b) The computed disparity map, (c) A gray-coded disparity image.

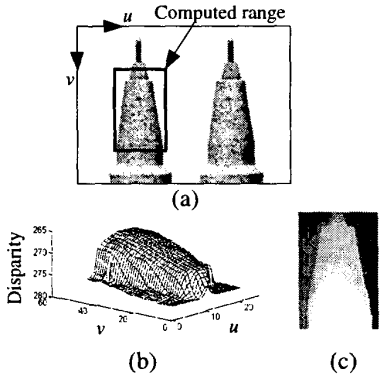


Fig. 6 A disparity map computed from the biprism stereo image for a pen: (a) An input biprism image, (b) The computed disparity map, (c) A gray-coded disparity image.

4.2 3-D Reconstruction

In the previous section, we presented a method to compute the depth map using only the disparity. To obtain Euclidean depth, however, it is necessary to know the intrinsic parameters of the camera and biprism. Fig. 7 shows an image obtained and the dimension of the cali-

bration box for the camera calibration. Let us assume that the world reference coordinate is centered at the lower-center corner of the calibration box. Using 60 reference markings on the calibration box as the reference points with known world coordinates, a well-known calibration algorithm by Faugeras and Toscani [8] is implemented to calibrate the camera. Table 4.1 shows the intrinsic parameters of the camera.

Fig. 8(a) shows an image obtained by a biprism to estimate k_1 and k_2. The distance between two points on the calibration box is 20 mm. Fig. 8(b) shows the computed disparity map. In the previous section, we presented a method to compute the depth map using only the disparity. The parameters k_1 and k_2 using the information of the distance between two points on the calibration box and the camera intrinsic parameters are estimated. Table 4.2 shows the results of the calibrated parameters k_1 and k_2.

From the results of the calibration, t_z is 151.7692 mm, B is 35.2192 mm, and δ is 6.6167°.

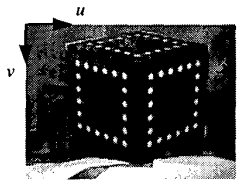
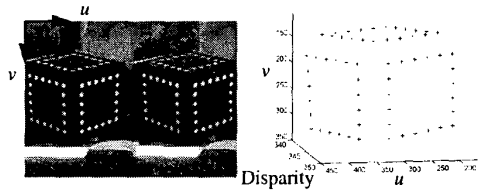


Fig. 7 A calibration box used for the camera calibration (150×150×150 mm)

Table 4.1 The calibrated intrinsic parameters of the camera

α_u	α_v	u_0	v_0
1657.412	1668.626	339.626	272.776



(a) (b)

Fig. 8 Image and disparity map computed for a biprism calibration: (a) An input image with a biprism, (b) The computed disparity map.

Table 4.2 The computed parameters k_1 and k_2

k_1	k_2
0.3946	0.0026

From the above results, if the base plane of the biprism is parallel to the image plane and the biprism does not rotate about the Z-axis, the perspective projection matrix of the left and right half image plane of real camera, respectively, can be obtained as

$$P_l = \begin{bmatrix} 1657 & 0 & 147 & 51506 \\ 0 & 1669 & 273 & 41384 \\ 0 & 0 & 1 & 152 \end{bmatrix}, P_r = \begin{bmatrix} 1657 & 0 & 532 & 51506 \\ 0 & 1669 & 273 & 41384 \\ 0 & 0 & 1 & 152 \end{bmatrix}.$$

Fig. 9 shows epipolar lines obtained by the perspective projection matrix. The average residual is 0.32 pixels.

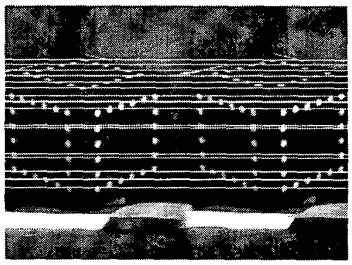
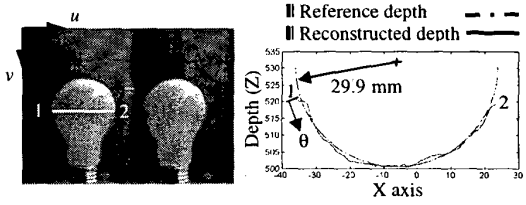


Fig. 9 Epipolar lines of the biprism stereo system

Fig. 10 shows an input biprism image and the recovered shape along one scanline for a textured light bulb. In the same manner, a simple cross-correlation technique is used for matching. We use a 25×25 window to compute the SSD(Sum of Squared Differences). The radius of a light bulb is 29.9 mm. The error in radius between the recovered shape and the true shape of the light bulb is smaller than 2 mm, and the percentile error is within 5%. Fig. 11 shows an input biprism image and the recovered structure for a block. We have used the pixel coordinates of the vertexes of the block with the pairs of corresponding points. Fig. 11(b) shows the extracted corner points. Fig. 11(c) shows the recovered shape for the block. Table 4.3 shows the error in length between the recovered shape and the true shape of the box. The maximum error is 2.5 mm.



(a) (b)

Fig. 10 Recovered shape with one scanline for a textured light bulb: (a) An input biprism image, (b) Recovered shape.

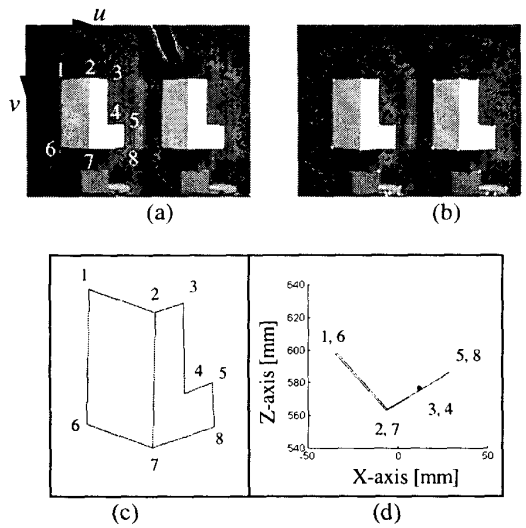


Fig. 11 Input image and recovered shape: (a) An input biprism image, (b) The extracted corner points, (c) The recovered shape, (d) Top view of (c).

Table 4.3 The length of the recovered line segments

Line segment	Recovered Value	True value	Error (recovery-true)
1 - 2	44.3981	44.8	-0.4019
2 -3	19.9069	22.0	-2.0931
1 - 6	65.7794	66.8	-1.0206
2 - 7	65.5072	66.8	-1.2928
3 - 4	44.0680	45.1	-1.0320
4 - 5	20.4130	22.2	-1.7870
6 - 7	44.6866	44.8	-0.1134
7 - 8	42.5620	44.7	-2.1380
5 - 8	21.6107	21.7	-0.0893

[Units: mm]

5 Conclusions

A single-lens camera system that can provide a pair of stereo images using a biprism was introduced. A stereo image obtained by this camera system is equivalent to two images obtained from two perfectly aligned cameras with exactly the same optical properties, which greatly simplifies the calibration of the system. Since the corresponding points in the image lie on the same scanline, correspondence problem becomes trivial.

This camera system can obtain a stereo pair from a single shot and through a single camera. Since the two images of a stereo pair are obtained at the same time, this camera can be used in many vision applications such as dynamic scene analysis.

Work is currently underway to improve the accuracy of structure recovery by taking into account the relative rotation of the biprism with respect to the camera coordinate system.

Acknowledgments

This paper was accomplished in part with Research Fund provided by Korea Research Foundation, Support for faculty Research Abroad.

Reference

[1] D.H. Lee, I.S. Kweon, A New Stereo Camera System by a Biprism, 10th Workshop on Image Processing and Understanding, JeJu, Korea, , pp. 291-296, 1998.

[2] D.H. Lee, I.S. Kweon, R. Cipolla, Single Lens Stereo with a Biprism, IAPR Workshop on Machine Vision Applications (MVA'98), Japan, 1998.

[3] S. T. Barnard and M. A. Fischler, Computational Stereo, Computing Surveys, 14(4), 1982, pp. 553-572.

[4] U. R. Dhond and J. K. Aggarwal, Structure from Stereo: A Review, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(6), pp. 1489-1510, 1989.

[5] Y. Nishimoto and Y. Shirai, A feature-based stereo model using small disparties, Proc. Computer Vision and Pattern Recognition. pp. 192-196, 1987.

[6] W. Teoh and X. D. Zhang, An inexpensive stereoscopic vision system for robots. Proc. Int. Conf. Robotics, pp. 186-189, 1984.

[7] A. Goshtasby and W. A. Gruver, Design of a Single-Lens Stereo Camera Sstem, Pattern Recognition, vol. 26, pp. 923-936, 1993.

[8] O. Faugeras and G. Toscani, The caibration problem for stereo, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'86, pp. 15-20, 1986.

[9] M. Bass and W. V. Stryland, HandBook of Optics I, II, McGRAW-HILL, INC. 1995.

[10] E. Hecht, Theory and Problems of Optics, Schaum's Outline Series, 1975.

[11] S. Nene and S. Nayar, Stereo with Mirrors, ICCV98, pp. 1087-1094, 1998.