MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

3—24 Single Lens Stereo with a Biprism

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

* Department of Automation and Design Engineering, KAIST

207-43, Cheongryangri-dong, Dongdaemoon-gu, Seoul, KOREA

**Department of Engineering, University of Cambridge (On leave from KAIST)

^ Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, U.K.

Abstract

We present a novel and practical stereo vision system that uses only one camera and a biprism placed in front of the camera. The equivalent of a stereo pair of images are formed as the left and right halves of a single CCD image by using a biprism placed in front of the lens of a CCD camera. 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 automatically lie on the same scan line. There is also no need for the normalization of image intensities (needed for correlation matching); the calibration of intrinsic parameters of two cameras; nor rectification.

A prototype system has been built and tested. 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. Preliminary results are presented. Results are extremely promising in the area of cheap sensors for many robotic applications.

1. Introduction

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 3D-shape and pose estimation [1, 2].

For the recovery of a three dimensional scene from a pair of stereo image of the scene, it is required to establish correspondence – that is, finding a pair of image points corresponding to the same point in space. The establishment of correspondence is generally thought to be the most difficult step and can easily become the most time-consuming [1].

A correspondence algorithm can produce more reliable matches if the underlying images have smaller intensity and geometric difference. 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. If the scene has Lambertian surfaces, there would be no difference in the intensities of corresponding points in the images. Differences in the optical properties of the two cameras, however, 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.

For stereo images acquired by two cameras, the focal lengths and zoom levels of the cameras are often slightly different. Two cameras may have lenses that do not have exactly the same optical properties. These camera differences create unwanted geometric and intensity differences between stereo images, making the correspondence process more difficult

Nishimoto and Shirai [3] proposed a single-lens camera system that can obtain stereo images. Stereo images are obtained with the plate at two different rotational positions. This camera system requires two shots from a scene and, therefore, should be used only in static environments.

Teoh and Zhang [4] 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. This camera system requires two shots from a scene and, therefore, should be used only in a static scene.

Gosthasby and Gruver [5] 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 this paper, we propose a novel stereo camera system that can provide a pair of stereo images from a single shot of a single camera using a *biprism*. Accordingly, this camera system considerably reduces unwanted geometric and intensity differences between the stereo images. The biprism is placed in front of the camera. An arbitrary object point in three-dimensional space is transformed into two virtual points by the biprism. As in the conventional stereo system, the displacement between the two conjugate image points of the two virtual points is directly related to the depth of the object point.

2. Principles of the Biprism-Stereo

Fig. 1 shows the geometry of a biprism and the associated coordinate system. Both the inclined prism planes R_l and R_2 make the angle α with the base plane R, respectively. An arbitrary point in 3-D space $P(X_p, Y_p, Z_p)$ is transformed into the two virtual points $P_r(X_{pr}, Y_{pr}, Z_{pr})$ and $P_l(X_{pl}, Y_{pl}, Z_{pl})$ 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.

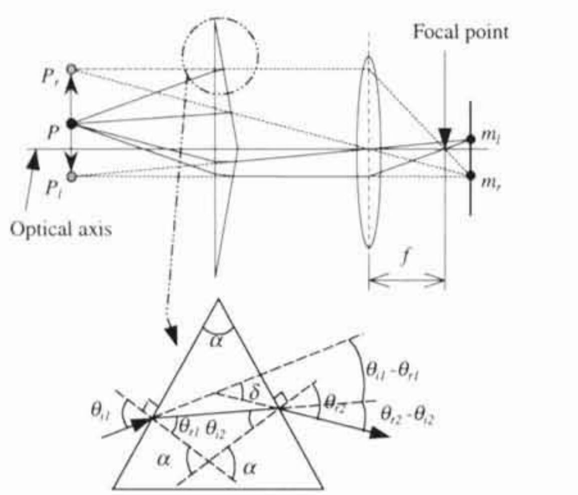
From the geometry of the biprism as shown in Fig. 1(a), we obtain

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

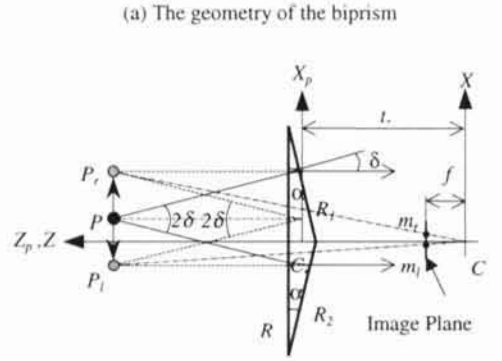
In Fig. 1, a lens gathers all the light radiating from the two virtual points P_r and P_l, and creates two corresponding image points m_r and m_l. From equation (1), given the angle α and the refractive index n, the basic relation for the biprism is given by [7, 8]

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

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.



(a) The geometry of the biprism



(b) Perspective projection of two virtual points on the image plane

Fig. 1 Principles of the biprism

In Fig. 1(b), 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 $P_r(X_{pr}, Y_{pr}, Z_{pr})$

$$\begin{bmatrix} X_{pr} \\ Y_{pr} \\ Z_{pr} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \tan\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} \tag{3}$$

2) For the virtual point $P_l(X_{pl}, Y_{pl}, Z_{pl})$

$$\begin{bmatrix} X_{pl} \\ Y_{pl} \\ Z_{pl} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\tan\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} \tag{4}$$

Let us define the disparity of the biprism as the translation distance between the two virtual points. From equation (3) and (4), we obtain

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

Equation (5) indicates that the distance between the two virtual points becomes larger as a 3-D point moves farther away from the biprism. Therefore, we can obtain the scene depth, given the distance between the two transformed points. Now, we define the distance between the two virtual points as a function of the disparity on the image plane. From perspective projection of the two virtual points, we obtain

$$u_r = \frac{\alpha_u}{Z} X_{pr} + u_0 \qquad u_l = \frac{\alpha_u}{Z} X_{pl} + u_0 \tag{6}$$

where $Z = Z_p + t_z$ and t_z is the distance between the biprism center and the optical center of camera.

The corresponding image disparity becomes

$$d = u_r - u_l = \frac{\alpha_u}{Z}(X_{pr} - X_{pl}) = \frac{2\alpha_u Z_p \tan\delta}{t_z + Z_p}\ . \tag{7}$$

Equation (7) can be rewritten as

$$d = \frac{2\alpha_u \tan\delta (Z - t_z)}{Z}\ . \tag{8}$$

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, the depth can be determined by equation (8). From equation (8), we observe that the disparity is constant independent of the location in the scene if $t_z = 0$.

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 $O_P - P_{22}$ and $O_P - P_{21}$. The effective baseline distance of the biprism-stereo, B, is defined by

$$B = 2t_z \tan\delta \tag{9}$$

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 biprism, which is related to the deviation δ:

$$-\delta \le FOV \le \delta\ . \tag{10}$$

A 3D point in the region ① has the two corresponding transformed points, whereas any points out of the FOV do not have the corresponding points though they have the transformed points.

Since the two virtual points have the same Y_p, Z_p in 3D space, epipolar lines in biprism-stereo images are parallel. As illustrated in Fg. 3, the equivalent stereo images are coplanar and parallel to their baseline. In other words, the images are rectified. An example biprism-stereo image pair is shown in Fig. 4.

The disparity with respect to the new image centers of the virtual cameras is proportional to the inverse of depth as for conventional stereo systems with two parallel cameras. For the biprism-stereo system, however, the disparity with respect to the original image center of the camera satisfies equation (8).

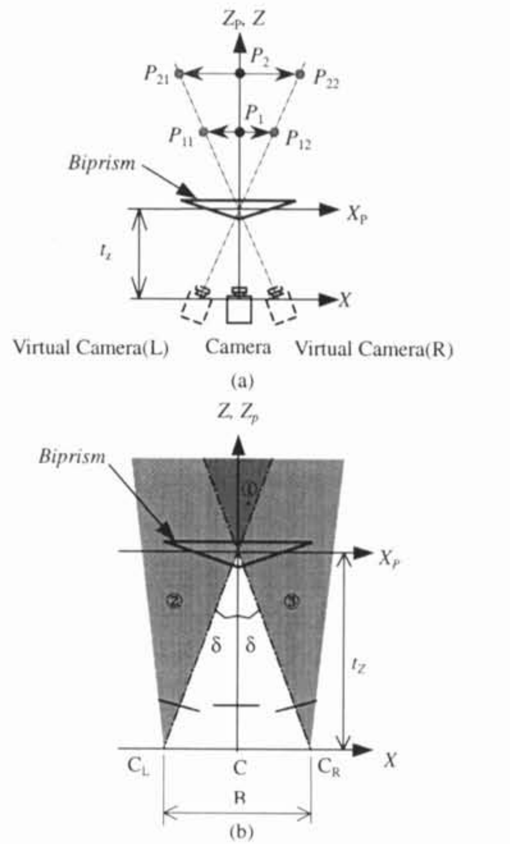


Fig. 2 The equivalent stereo system to the biprism stereo

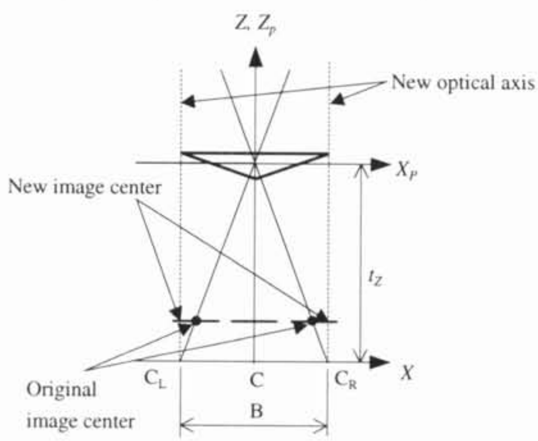


Fig. 3 The rectified images of two virtual camera

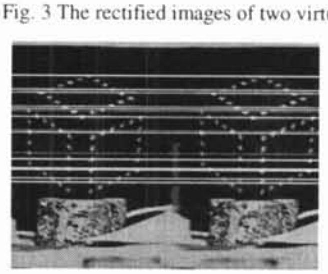


Fig. 4 Parallel epipolar lines in a pair of biprism-stereo image

3. Experimental Results

3-1 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. 5. For this particular prototype, distance between the optical center and the biprism, t_z, is 150 mm, and the biprism is designed to have an declination angle of 12.4 degrees. Therefore, the effective baseline distance is 38.8 mm.

As we explored in the previous section, the distance between two corresponding points $m_r(u_r, v_r)$ and $m_l(u_l, v_l)$ in the image plane is proportional to the distance between the two virtual points transformed by the biprism in the camera coordinate.

A simple cross-correlation technique is used for matching. For this particular experiment, we use a 25 x 25 window to compute the sum of squared differences. Fig. 6 and Fig. 7 show an input biprism-image and the computed disparity map.

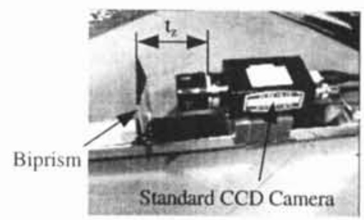
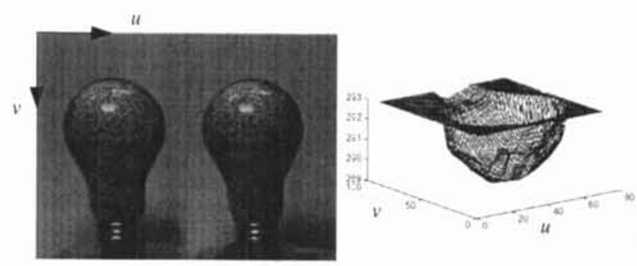
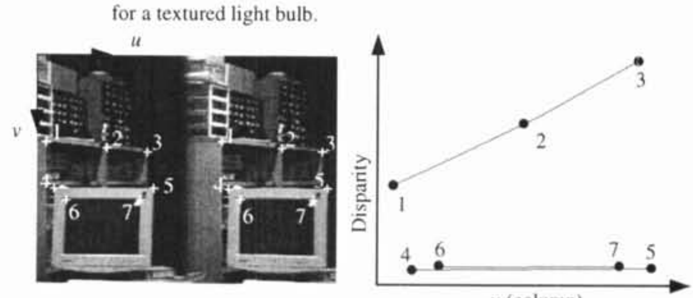


Fig. 5 An implementation of a biprism-stereo camera



(a) An input biprism image (b) Computed disparity map

Fig. 6 A disparity map computed from the biprism stereo image for a textured light bulb.

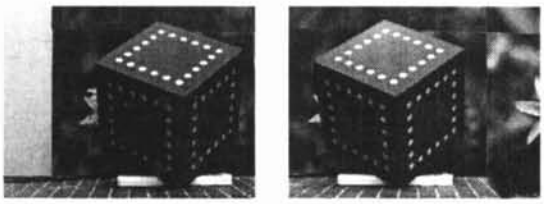


(a) An input biprism image (b) Computed disparity map

Fig. 7 A disparity map computed from the biprism stereo image

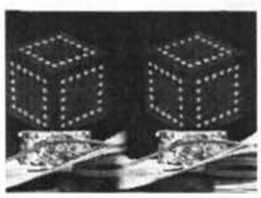
3-2 Euclidean Depth

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. 8 shows a stereo image pair and a biprism-stereo image of a calibration box, obtained by a conventional stereo camera system and the proposed biprism-stereo camera, respectively. The calibration box is 150 x 150 x 150 mm.



Left Image Right Image

(a) Stereo image pair (640x480)



(b) Biprism-stereo image (640 x 480)

Fig. 8 Images of a calibration box (150 x 150 x150 mm)

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 [6] is implemented to calibrate both a conventional two-camera stereo and the proposed biprism-stereo camera. The conventional stereo system consists of two CCD cameras with parallel camera geometry and a baseline distance of 109.6 mm. Note that the effective baseline distance for the biprism-stereo camera is 38.8 mm. As shown in Table 1, errors from the biprism-stereo camera are comparable with those of the conventional stereo system.

Table 1. Errors in the computed Euclidean depth for 60 reference points of the calibration box:

RMS error	Conventional stereo	Biprism stereo
Δ_x (mm)	0.0021	0.0366
Δ_y (mm)	0.0057	0.0103
Δ_z (mm)	0.0021	0.0154

4. Conclusions

A single camera system that can provide a pair of stereo images using a *biprism* was introduced. A biprism is placed in front of a CCD camera such that the base plane of the biprism is parallel to the image plane. An arbitrary point in three-dimensional spaces is transformed into two virtual points by the biprism. This biprism-stereo camera system has the following properties:

(1) A single image obtained by this camera is equivalent to two images obtained from two perfectly aligned cameras with exactly the same optical properties.

(2) The deviation δ of the biprism, as a function of α and n, is an angle when a three dimensional point is transformed into the two virtual points, and the distance between the two virtual points becomes larger as the three dimensional point moves farther away from the biprism. The distance between two virtual points is proportional to the distance to the biprism.

(3) Two virtual points lie on the same scanline. If the base plane of the biprism is parallel to the image plane, and the biprism does not rotate about the Z-axis, the corresponding points in the image lie on the same scanlines.

(4) If we know the disparity between corresponding points in the image, the relative distance between the points in three-dimensional spaces can be determined. Accordingly, we define the distance between the corresponding points as the *disparity*.

(5) The field of view of the biprism increases with the increase of the angle α of the biprism. Also, the disparity increases. But, in order to find corresponding points in the images, the field of view of the biprism must be smaller than the field of view of the camera.

(6) 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.

(7) This camera system can recover the relative depth using only the disparity between the corresponding points without knowledge of parameters of the camera system.

Reference

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

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

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

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

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

[6] 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.

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

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