A Novel Stereo Camera System by a Biprism

DooHyun Lee and InSo Kweon, *Member, IEEE*

***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 charge coupled device (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 setup 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 accurate for nearby objects in a small field of view. Since we use only a single lens, calibration of the system is greatly simplified. Given the parameters in the biprism-stereo camera system, we can reconstruct the three-dimensional structure using only the disparity between the corresponding points.

***Index Terms*—Biprism stereo, calibration, disparity, disparity map, Euclidean structure.**

I. Introduction

RESEARCH on the recovery and recognition of three-dimensional (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 the two images from two distinct viewpoints is a powerful cue to 3-D shapes and pose estimation [1], [2]. For the recovery of a 3-D scene from a pair of stereo image of the scene, it is required to establish correspondences [1]. The stereo process can be summarized by the following steps: 1) detection of features in each image; 2) matching of features between the images under certain geometric and other constraints; and 3) calculation of depth using the disparity values and the geometric parameters of the imaging configuration. While each of these steps is important in the stereo process, the matching of features known as the correspondence is generally thought to be the most difficult step and can easily become the most time consuming [1], [3], [4].

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. For stereo images acquired by the 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.

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. It is also easy to set up the biprism so that the corresponding features lie on the same scanline automatically. The basic concept of a biprism-stereo camera system is partially presented in [5] and [6], and we present the proposed stereo camera system in details.

This paper is organized as follows. In Section II, the previous work is introduced. In Section III, the proposed biprism-stereo system is introduced. In Section IV, the method of the calibration of the biprism-stereo camera system using the distance between the two reference points in 3-D space and the disparity in the image plane is presented. In Section V, we present some disparity map of objects. In Section VI, we present a method to compute the depth map using only the disparities. We discuss the error analysis with respect to the pixel noise, and conclusions are given in Section VII.

II. Related Researches

Stereo images are usually obtained either by displacing a single camera in the scene or using two cameras mounted on a platform separated by a small distance. By devising a more accurate calibration process and using higher-quality cameras, the unwanted geometric and intensity difference between stereo images can be reduced. Some difference, however, will still remain between the images.

Nishimoto and Shirai [7] proposed a single-lens camera system that can obtain stereo images as shown in Fig. 1(a). In the systems, a glass plate is placed in front of the camera, and images are obtained with the plate at two different rotational positions. When the glass plate is rotated, the optical axes of the camera shifts slightly, simulating two cameras with parallel optical axes. The obtained stereo images have very small disparities making the point correspondence easy. However, only coarse depth values can be obtained from the disparities.

Teoh and Zhang [8] described a single-lens stereo camera system with the geometry as shown in Fig. 1(b). Two mirrors,

Manuscript received August 31, 1999; revised February 8, 2000. This paper was recommended for publication by Associate Editor H. Zhuang and Editor V. Lumelsky upon evaluation of the reviewers' comments.

D. Lee is with the Department of Automation and Design Engineering, Korea Advanced Institute of Science and Technology, 207-43 Seoul, Korea (e-mail: doolee@cais.kaist.ac.kr).

I. Kweon is with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, 373-1 Taejon, Korea (e-mail: iskweon@cais.kaist.ac.kr).

Publisher Item Identifier S 1042-296X(00)08349-X.

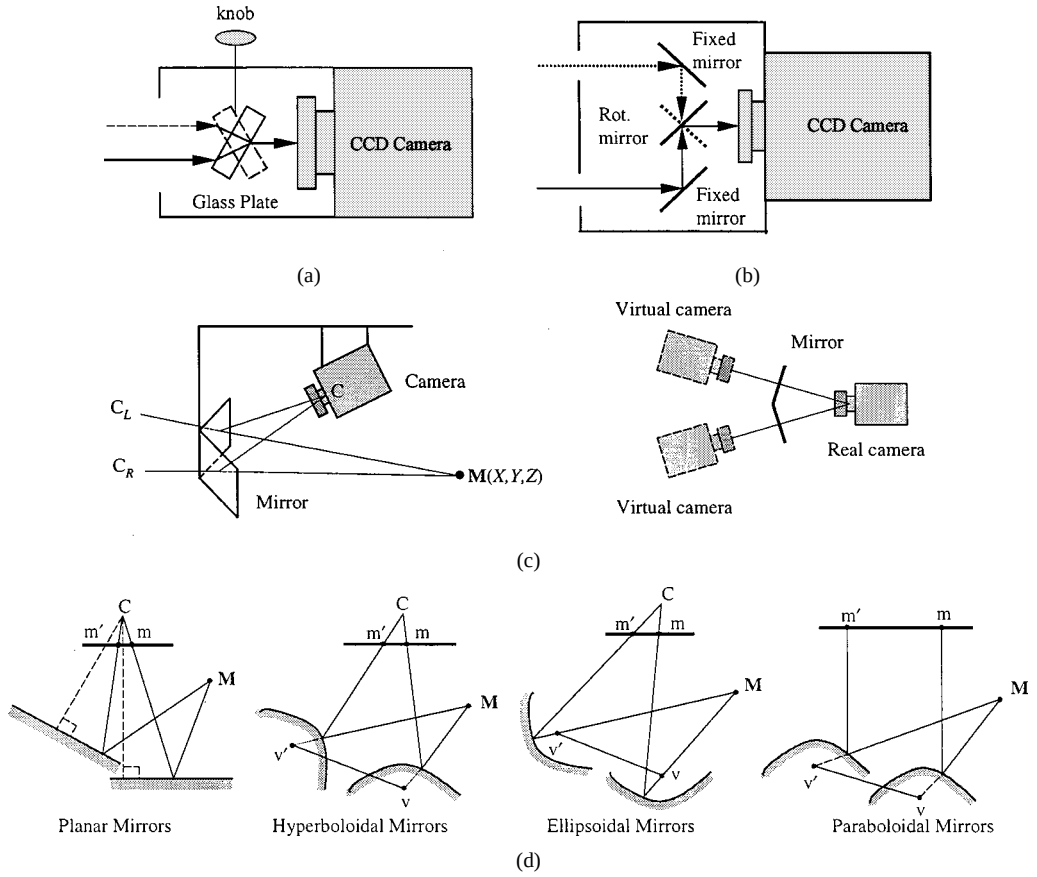


Fig. 1. Structure of stereo camera systems using single camera.

fixed to the body of the camera, make a 45° angle with the optical axes of camera. A third mirror that can rotate is placed directly in front of the lens. 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, Here also, although a single camera is used, the result is the same as using two cameras with parallel optical axes.

Both of these cameras considerably reduce unwanted geometric and intensity difference between stereo images. But the cameras have parts that should be rotated when obtaining a pair of images. Exact rotation of the parts is a major design issue in these systems, and two shots of a scene are required. Therefore, this camera system can be used only in a static scene.

Fig. 1(c) shows a single camera system that can obtain images in a single shot using a single lens proposed by Gosthasby and Gruver [9]. The obtained images are reflected about the image of the mirror's axis. This camera system can obtain images in a single shot and through a single camera. But, 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. The interreflection between the mirrors causes intensity difference between corresponding points in stereo images.

In their recent work, Nene and Nayar [10] proposed four stereo systems that use a single camera pointed toward planar, ellipsoidal, hyperboloidal, and paraboloidal mirrors as shown in Fig. 1(d). By use of nonplanar reflecting surfaces such as hyperboloids and paraboloids, a wide field of view (FOV) is easily obtained. However, their stereo system needs a complex mirror mechanism.

III. Biprism-Stereo System

A. Definition of Coordinates System

Figs. 2 and 3 show the geometry of a biprism and the associated coordinate system. We assume that the used charge coupled device (CCD) camera is modeled as the ideal pinhole camera model, which is a perspective projection from the world to the image plane. The coordinates (O, X, Y, Z) and (O_p, X_p, Y_p, Z_p) are the standard coordinate system of the camera and the biprism, respectively. The X- and Y-axis of the camera coordinates are parallel to the X_p- and Y_p-axis of the biprism coordinates, respectively, which means that the image plane of the real camera is parallel to the base plane Π of the biprism. The Z- and Z_p-axis coincides the optical axis of the camera. The distance t_z is the distance between the origin of the biprism and the optical center of the camera and α is the angle of the biprism.

B. The Principle of Biprism

The inclined planes Π_r and Π_l make the angle α with the base plane Π, respectively, as shown in Fig. 3. An arbitrary point in 3-D space is transformed into the two virtual points by the inclined biprism planes Π_r and Π_l, 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

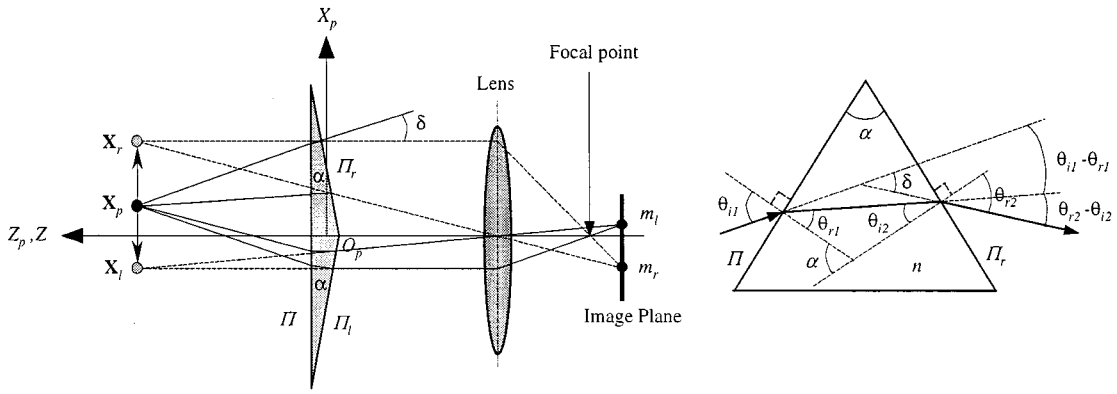
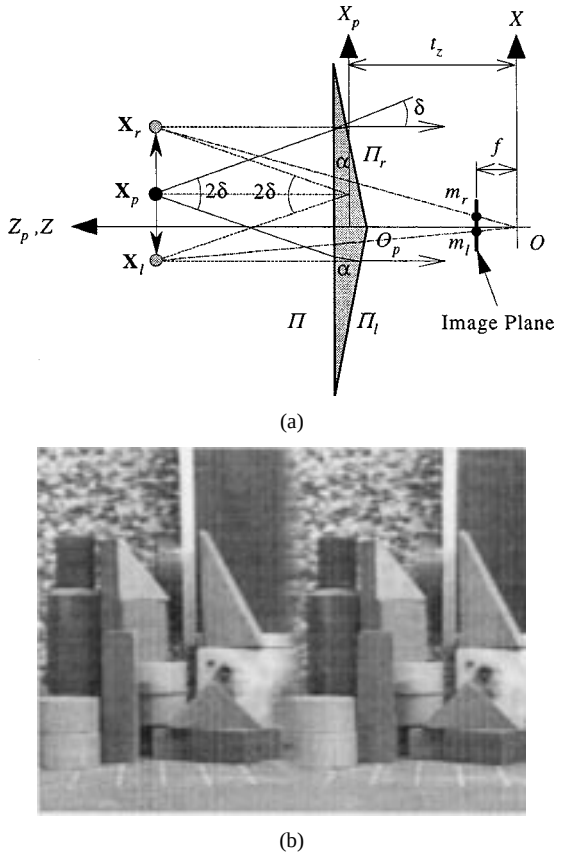


Fig. 2. Geometry of the biprism.



(a)

(b)

Fig. 3. Principle of the biprism. (a) Perspective projection of the two virtual points on the image plane. (b) An image captured by a biprism-stereo system.

index of refraction n of the biprism. From the geometry of the biprism as shown in Fig. 2, we obtain

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

In Fig. 2, a lens gathers all the light radiating from the two virtual points $\mathbf{X}_r$ and $\mathbf{X}_l$, and creates the two corresponding image points $\mathbf{m}_r$ and $\mathbf{m}_l$. From (1), given the angle α and the refractive index n, the basic relation for the biprism is given by [11]–[13]

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

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 FOV of the biprism.

In Fig. 3, the geometric relationship between an object point $\mathbf{X}_p = [X_p, Y_p, Z_p]^T$ and the two virtual 3-D points $\mathbf{X}_r = [X_{pr}, Y_{pr}, Z_{pr}]^T$ and $\mathbf{X}_l = [X_{pl}, Y_{pl}, Z_{pl}]^T$, created by the biprism, can be represented by simple transformations

$$\mathbf{X}_r = \mathbf{T}_r \mathbf{X}_p \qquad \mathbf{X}_l = \mathbf{T}_l \mathbf{X}_p \tag{3}$$

where

$$\mathbf{T}_r = \begin{bmatrix} 1 & 0 & \tan\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{T}_l = \begin{bmatrix} 1 & 0 & -\tan\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

The matrix $\mathbf{T}_l$ and $\mathbf{T}_r$ describe the transformation matrix that transform a real 3-D point into the virtual points by a biprism. From (3), the distance between the two virtual points can be obtained as

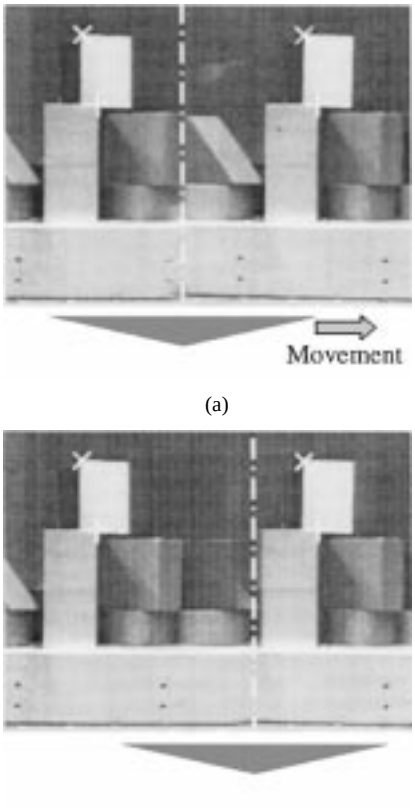
$$D = X_{pr} - X_{pl} = 2Z_p \tan\delta. \tag{5}$$

Equation (5) indicates that the distance between the two virtual points becomes larger as the 3-D point moves farther away from the biprism. The distance between the two virtual points increases with the increase of the distance from an arbitrary point on the object to the biprism.

C. 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 biprism coordinates system and the camera coordinates system. 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 the two virtual points $\mathbf{X}_r$ and $\mathbf{X}_l$. The overall imaging process for the two virtual cameras can be expressed as a single matrix multiplication in homogeneous coordinates [14], [15]

$$s\tilde{\mathbf{m}}_l = \mathbf{P}_l \tilde{\mathbf{X}}_p \qquad s\tilde{\mathbf{m}}_r = \mathbf{P}_r \tilde{\mathbf{X}}_p \tag{6}$$



(a)

(b)

Fig. 4. Pixel coordinates according to the position of biprism.

where

$$\mathbf{P}_l = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \quad \mathbf{t}] \begin{bmatrix} \mathbf{T}_l & \mathbf{0}_3^T \\ \mathbf{0}_3 & 1 \end{bmatrix}$$
$$\mathbf{P}_r = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \quad \mathbf{t}] \begin{bmatrix} \mathbf{T}_r & \mathbf{0}_3^T \\ \mathbf{0}_3 & 1 \end{bmatrix}. \tag{7}$$

In (7), the matrix $\mathbf{R}$ and the vector $\mathbf{t}$ describe the orientation and position of the camera with respect to the biprism coordinates system. The coordinate (u_0, v_0) is the coordinates of the principal point in the image plane, and parameters α_u and α_v are the focal length of the camera, measured in pixel units along horizontal and vertical directions, respectively.

D. Relationship Between Disparity and Depth

In (7), we consider the rotation matrix $\mathbf{R}$ between the biprism and the camera coordinates. In general case, the rotation matrix $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{8}$$

where $\mathbf{R}$ has nine components but there are only 3 degrees of freedom. A rotation matrix $\mathbf{R}$ must satisfy $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ and $\det(\mathbf{R}) = 1$. Next, we consider the translation vector $\mathbf{t}$. As shown in Fig. 4, the coordinates of the virtual points according to the translation along the X-axis do not change, which means that the pixel coordinates of its projected points onto the image plane are unchanged. Only the correspondence region will be changed as shown in Fig. 4. Therefore, installation of the biprism-stereo system is very easy. Table I shows the pixel coordinates at each position. For the further simplification, let the origin of the biprism be located on the optical axis of the camera. Accordingly, the translation vector can be represented as $\mathbf{t} = [0 \quad 0 \quad t_z]^T$.

TABLE I
PIXEL COORDINATES AND DISPARITY AT EACH POSITION

		$\mathbf{m}_l(u_l,v_l)$	$\mathbf{m}_r(u_r,v_r)$	Disparity
×	(a)	(121,47)	(472,47)	351
	(b)	(147,159)	(492,159)	351
+	(a)	(121,47)	(472,47)	345
	(b)	(147,159)	(492,159)	345

From (7), let the base plane Π of the biprism be parallel to the image plane of the camera, which means that the X- and Y-axis of the camera coordinates are parallel to the X_p- and Y_p-axis of the biprism coordinate, respectively, and the biprism does not rotate about the Z-axis. Thus, the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$ become

$$\mathbf{R} = \mathbf{I}_{3\times 3} \qquad \mathbf{t} = [0 \quad 0 \quad t_z]^T \tag{9}$$

where $\mathbf{I}$ is an identity matrix.

From (6), the relationship between an arbitrary point $\tilde{\mathbf{X}}_p = [X_p, Y_p, Z_p, 1]^T$ and its projected points $\tilde{\mathbf{m}}_l = [u_l, v_l, 1]^T$ and $\tilde{\mathbf{m}}_r = [u_r, v_r, 1]^T$ onto the left and right image plane, respectively, can be expressed as

$$\begin{bmatrix} u_l \\ v_l \end{bmatrix} = \frac{1}{Z_p + t_z} \begin{bmatrix} \alpha_u(X_p - Z_p \tan\delta) + u_0(Z_p + t_z) \\ \alpha_v Y_p + v_0(Z_p + t_z) \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} u_r \\ v_r \end{bmatrix} = \frac{1}{Z_p + t_z} \begin{bmatrix} \alpha_u(X_p + Z_p \tan\delta) + u_0(Z_p + t_z) \\ \alpha_v Y_p + v_0(Z_p + t_z) \end{bmatrix}. \tag{11}$$

From (10) and (11), 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 from (10) and (11). The relationship between the image disparity and the depth can be obtained as

$$d = u_r - u_l = \frac{2\alpha_u \tan\delta(Z - t_z)}{Z} \tag{12}$$

where $Z = Z_p + t_z$, which is the depth of the points on the object in the camera coordinates. The disparity increases with the increase of the angle α of the biprism and/or of the focal length of the camera. It is clear that the disparity approaches a constant value, when the depth approaches infinity. Equivalently, the relationship between the inverse disparity and the inverse depth can be expressed as

$$\frac{1}{d} = \frac{k_1}{Z_p} + k_2 \tag{13}$$

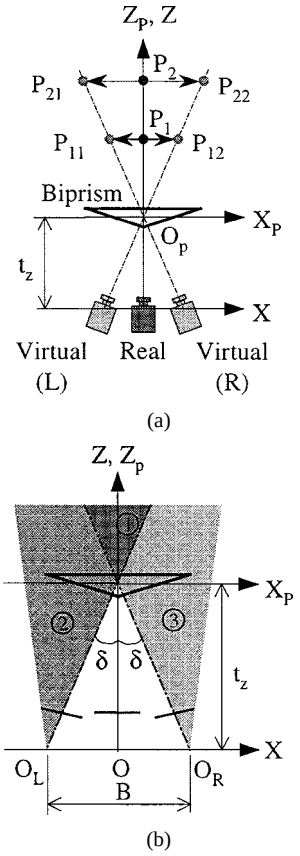


Fig. 5. Equivalent stereo system to the biprism stereo.

where

$$k_1 = k_2 t_z \quad \text{and} \quad k_2 = \frac{1}{2\alpha_u \tan\delta}.$$

The inverse disparity between the corresponding points is proportional to the inverse depth in the biprism coordinates as shown in (13). From (12), we observe that the disparity is constant independent of the location in the scene if $t_z = 0$.

E. Equivalent Stereo System

Fig. 5 shows the equivalent stereo camera system to a biprism-stereo system. The points P_1 and P_2 on the optical axis of the real camera are transformed into the virtual points P_{11}, P_{12} and P_{21}, P_{22} by the inclined biprism planes Π_r and Π_l, respectively. The virtual cameras are located on the virtual optical axis represented by the line $O_p \sim P_{21}$ and $O_p \sim P_{22}$, respectively. The optical center of the real camera is also transformed into the optical center of the virtual cameras, respectively. Accordingly, the optical center of the left and right virtual cameras is positioned at points O_L and O_R, respectively. The active image plane of the left and right virtual cameras are formed by the left and right halves of the image plane, respectively.

In Fig. 5(b), the effective baseline distance of the biprism-stereo camera system B is defined by

$$B = 2t_z \tan\delta = \frac{t_z}{k_2\alpha_u} = \frac{k_1}{k_2^2\alpha_u}. \tag{14}$$

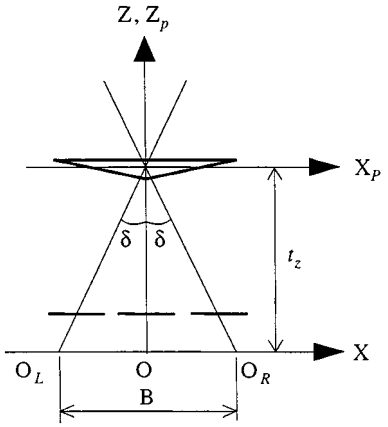


Fig. 6. Rectified images of the two virtual cameras.

The effective baseline distance increases with the increase of the distance t_z and/or of the angle α of the biprism. As illustrated in Fig. 6, the equivalent stereo images are coplanar and parallel to their baseline. In other words, the images are rectified. The left- and right-half images of the real camera corresponds to the image obtained from the left and right virtual cameras, respectively.

F. FOV

As shown in Fig. 5(b), the active image plane of the left and right virtual cameras are formed by the left and right halves of the image plane, respectively. In other words, regions ① and ② and regions ① and ③ comprise the FOV of the left and right virtual camera. The overlapped region between the two fields of view defines the FOV of the biprism as the correspondence region, which is related to the deviation δ as shown in Fig. 5(b)

$$-\delta \le \text{FOV} \le \delta \tag{15}$$

where $\delta = 2\sin^{-1}(n\sin\frac{\alpha}{2}) - \alpha$.

The biprism is also known to create the so-called interference area (IA). As shown in Fig. 7, the gray region denotes the IA in the biprism image. θ_b is defined as the angle between the base line and the line passing through the optical center of the virtual camera and the intersected point of the two inclined planes, which is defined by

$$\theta_b = \tan^{-1}\left(\frac{t_z - h_b/2}{t_z \tan\delta}\right) \tag{16}$$

where h_b is the thickness of the biprism. The IA in the biprism-stereo image is a function of h_b and t_z. In Fig. 7, the IA in the image plane can be obtained as

$$\text{IA} = \alpha_u\left(\frac{1}{\tan\theta_b} - \tan\delta\right). \tag{17}$$

We consider the IA in the image plane according to the parameters in (17). Fig. 8 shows the IA according to the distance t_z. For the simulation, the focal length of the camera lens is set to 16 mm, and the size of the used image and the angle of the

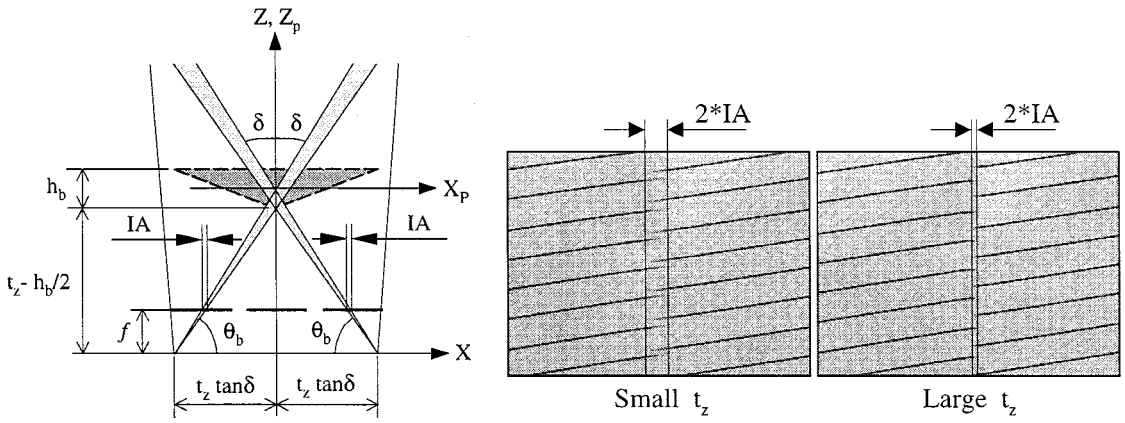


Fig. 7. IA in the biprism-stereo system.

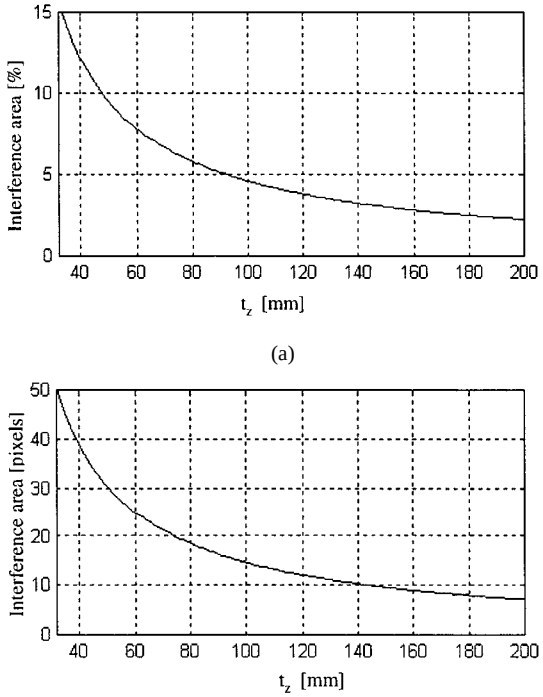


Fig. 8. IA according to t_z in the image plane: (a) the percentile IA and (b) the number of columns of the IA.

biprism are equal to 640(H) × 480(V) pixels and 12.4°, respectively. As shown in Fig. 8, it is easy to see that the IA decreases as the distance t_z increases. When $t_z = 150$ mm, the IA in the image plane is within 3% and 18 columns, which is small enough for many stereo applications.

In Fig. 5(b), a 3-D point in the region ① has the two corresponding transformed points, whereas any points out of the FOV do not have corresponding points. In Fig. 5(b), the transformed virtual region of region ① by the inclined planes Π_r and Π_l is region ①′ and ①″, respectively, as shown in Fig. 9. The transformed virtual region as the correspondence region can be obtained from the geometric relation. We consider an arbitrary point on the boundary of the FOV of the biprism. An arbitrary point is transformed into the two virtual points by the inclined planes Π_r and Π_l, respectively. Therefore, one of the transformed points lies on the Z_p-axis. As shown in Fig. 9, the

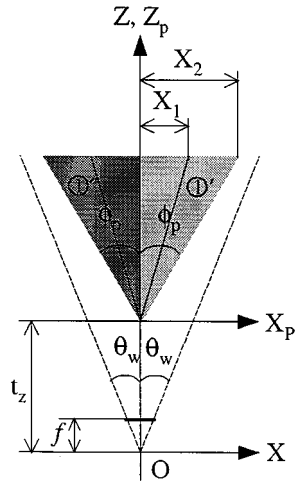


Fig. 9. Transformed region by the biprism.

distance between an arbitrary point in the biprism coordinates and the two virtual points, respectively, becomes

$$X_1 = Z_p \tan\delta \qquad X_2 = Z_p \tan\phi_p = 2X_1. \tag{18}$$

From (18), the transformed region ①′ and ①″ by the inclined planes Π_r and Π_l can be obtained as

$$\phi_p = \tan^{-1}(2\tan\delta). \tag{19}$$

Given the index of refraction n of the biprism, the transformed region ϕ_p by the biprism increases with the increase of the angle α of the biprism, which means that the distance between corresponding points increases as shown in (15).

We consider the effects according to the FOV of the camera for correspondence in the biprism-stereo system. The horizontal and vertical FOV of the camera, respectively, are defined by

$$\theta_w = \tan^{-1}\frac{w}{2f} \qquad \theta_h = \tan^{-1}\frac{h}{2f} \tag{20}$$

where w and h are the width and height of the CCD array of the camera, respectively, and f is the focal length of the camera lens.

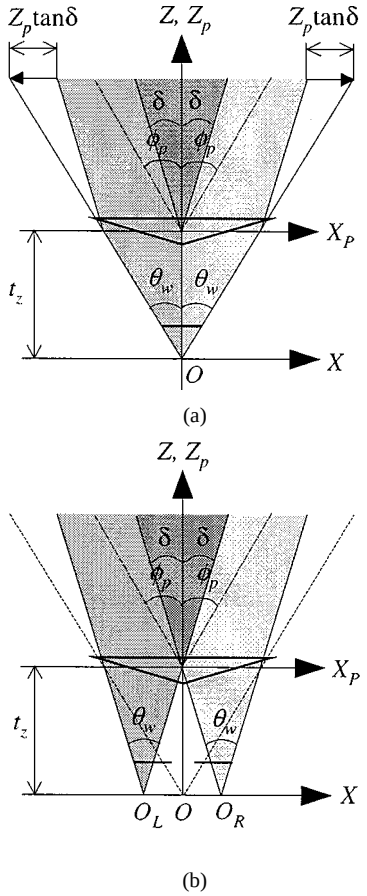


Fig. 10. FOV in case of $\phi_p \leq \theta_w$: (a) the FOV of the real camera and (b) the FOV of the virtual cameras.

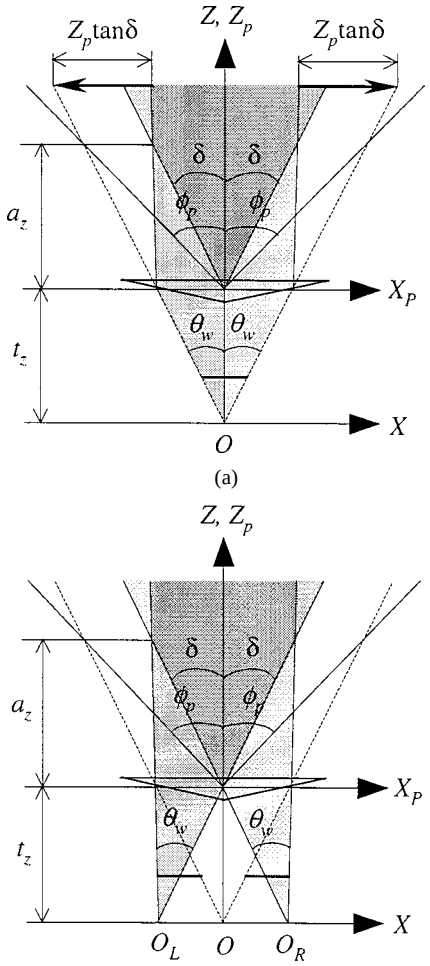


Fig. 11. FOV in the case of $\phi_p > \theta_w$: (a) real camera and (b) virtual cameras.

If the horizontal FOV θ_w of the camera must be greater than the transformed region ϕ_p by the biprism, all the transformed points by the biprism will be projected onto the image plane as shown in Fig. 10. Here, we assume that FOV of the real camera is equivalent to the FOV of the virtual cameras as shown in Fig. 10(a) and (b). If the transformed region ϕ_p is greater than the horizontal FOV θ_w of the camera, then all the transformed points by the biprism will not be projected onto the image plane. The transformed region ϕ_p by the biprism increases with the increase of the angle α of the biprism. From the geometric relation, the distance a_z can be obtained as

$$a_z = \frac{t_z \tan\theta_w}{\tan\phi_p - \tan\theta_w}. \tag{21}$$

The distance a_z increases with the increase of the distance between the origin of the biprism and the optical center of the camera and/or of the FOV of the camera. Given t_z and the parameters of the camera, the correspondence region decreases with the increase of the angle α of the biprism as shown in Fig. 11.

IV. Euclidean Structure and Calibration

This section presents a calibration method of the biprism-stereo camera system using the distance between the two reference points in 3-D space and the corresponding image disparity. It is straightforward to recover 3-D structure if the biprism-stereo camera is calibrated. To obtain the intrinsic parameters of the camera, we use a well-known calibration algorithm by Faugeras and Toscani [16].

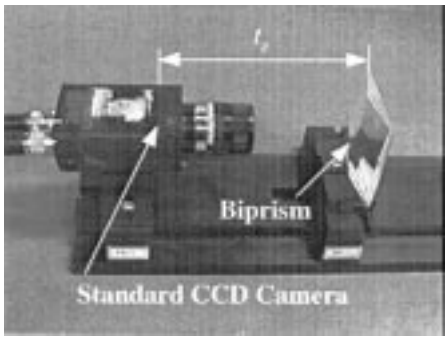
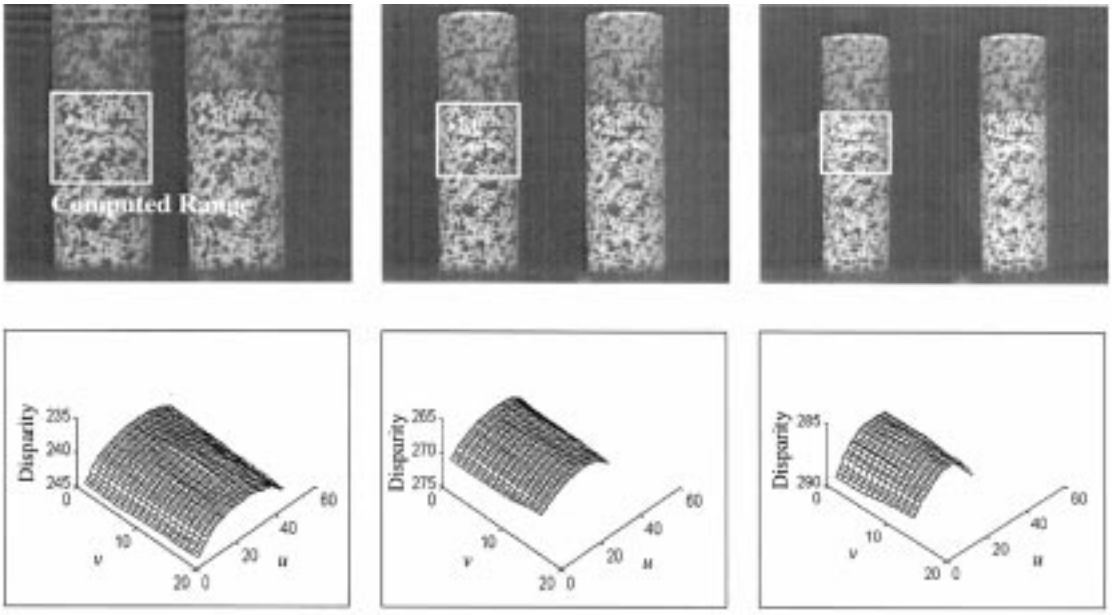


Fig. 12. Implementation of a biprism-stereo camera.

Suppose we know the Euclidean distance d^r between the two reference points in 3-D space. In (13), the biprism parameters can be obtained through the minimization of the following criterion:

$$\min \sum_{i=1}^{N} \|d_i^e(k_1, k_2) - d_i^r\|^2 \tag{22}$$



(a) (b) (c)

Fig. 13. Images and the computed disparity map in each position: (a) $Z_p = 150$ mm; (b) $Z_p = 200$ mm; and (c) $Z_p = 250$ mm.

where $d_i^c(k_1, k_2)$ can be easily derived from (10), (11), and (13) and N denotes the number of the corresponding points.

To solve (22), initial guesses for k_1 and k_2 are obtained from the design parameters of the biprism and the intrinsic parameters of the camera. With the calibrated biprism and camera, recovering 3-D structure using the disparity is straightforward. The distance between the origin of the biprism and the optical center of the real camera becomes

$$t_z = \frac{k_1}{k_2}. \tag{23}$$

Known k_1, 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. Given Z and the intrinsic parameters of the camera, we can compute the coordinates X and Y using

$$\begin{bmatrix} X \\ Y \end{bmatrix} = Z \begin{bmatrix} \dfrac{u_c - u_0}{\alpha_u} & \dfrac{v_c - v_0}{\alpha_v} \end{bmatrix}^T \tag{24}$$

where

$$Z = Z_p + t_z \quad \text{and} \quad \begin{bmatrix} u_c \\ v_c \end{bmatrix} = \begin{bmatrix} \dfrac{u_r + u_l}{2} & \dfrac{v_r + v_l}{2} \end{bmatrix}^T. \tag{25}$$

V. Disparity Map

This section presents some disparity maps of the objects, which are computed from correspondences found automatically in a biprism-stereo image. To obtain a biprism-stereo image, a biprism is placed in front of the camera as shown in Fig. 12. To align the biprism with respect to the image plane of the camera, they should be adjusted such that the corresponding points lie on the same scanline in the image plane.

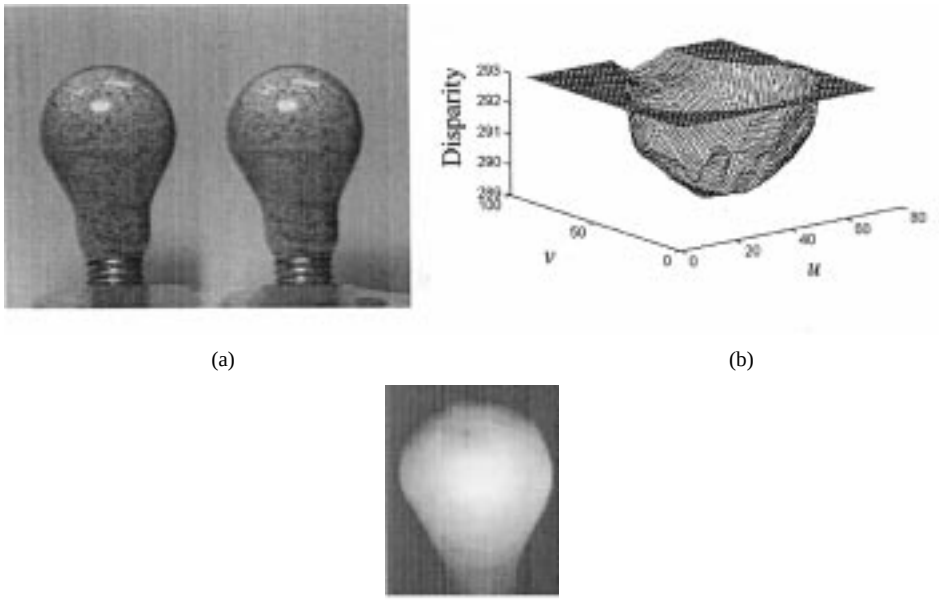
For this particular prototype, the distance between the optical center and the biprism t_z is about 150 mm and the biprism is designed to have the angle of 12.4 °. The thickness of the biprism in (16), h_b, is set to 8 mm. The focal length of the camera lens and the size of the used image are equal to 16 mm and 640(H) $\times$ 480(V), respectively.

The distance between the two corresponding points $\mathbf{m}_l(u_l, v_l)$ and $\mathbf{m}_r(u_r, v_r)$ in the image plane increases with increases of the distance between the two virtual points transformed by the biprism in the camera coordinate. To compute the disparity map, a standard sum of squared differences (SSD) technique is used for matching. For this particular experiment, we use a 25×25 window to compute the SSD. We consider the disparity according to the depth as the distance to the biprism. Fig. 13 shows the disparity for a cylindrical object with the radius of 11.6 mm. The distance from the biprism to the object in Fig. 13(a) is about 150 mm, and the case of Fig. 13(c) is about 250 mm. The disparity becomes larger as the object moves farther away from the biprism. Fig. 14 shows an input biprism image for a textured object and the computed disparity map. Fig. 15 shows an input biprism image for a block and the computed disparity map. Fig. 15(b) shows the extracted corner points, and we compute the disparity using the corner points as shown in Fig. 15(c) and (d).

VI. Euclidean Structure

A. Calibration of a Biprism-Stereo Camera System

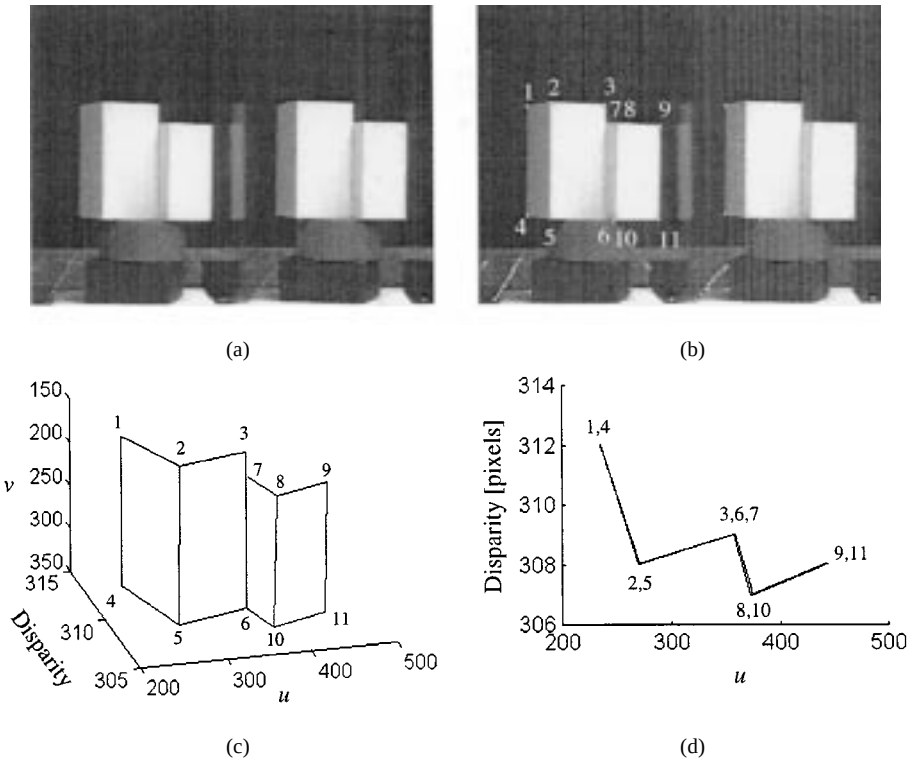
In the previous section, we presented a method to compute the depth map using only the disparity. Fig. 16 shows an image obtained and the dimension of a calibration box for the camera calibration. Let us assume that the world reference coordinate is centered at the lower-center corner



(a) (b)

(c)

Fig. 14. Images and the computed disparity map for a textured light bulb: (a) an input biprism image; (b) the computed disparity map; and (c) a gray-coded disparity image.

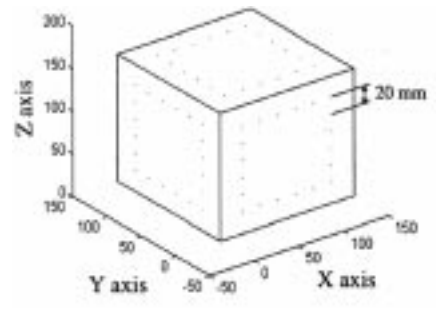


(a) (b)

(c) (d)

Fig. 15. Disparity map computed from the biprism-stereo image for a block: (a) an input biprism image; (b) The extracted corner points; (c) the computed disparity map; and (d) the top view of (c).

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 [16] is implemented to calibrate the camera. Table II shows the intrinsic parameters of the camera. Fig. 17(a) shows an image obtained by a biprism to estimate k_1 and k_2. The distance between any two points on the calibration box is 20 mm. Fig. 17(b) shows the computed disparity map. The calibrated parameters k_1 and k_2 using (22) are shown in Table III.



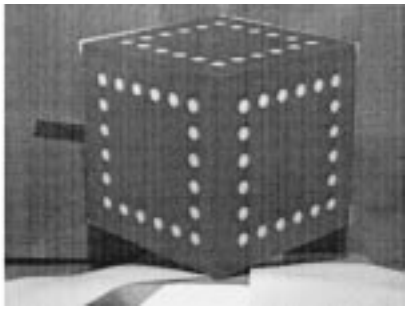


Fig. 16. Calibration box for the camera calibration (150 × 150 × 150).

TABLE II
CALIBRATED INTRINSIC PARAMETERS OF THE CAMERA

α_u	α_v	u_0	v_0
1657.412	1668.626	339.626	272.776

From the results of the calibration, t_z is 151.7692 mm, B is 35.2192 mm, and δ is 6.6167°. The perspective projection matrices for the biprism camera are obtained as

$$\mathbf{P}_l = \begin{bmatrix} 1657 & 0 & 147 & 51\,506 \\ 0 & 1669 & 273 & 41\,384 \\ 0 & 0 & 1 & 152 \end{bmatrix}$$
$$\mathbf{P}_r = \begin{bmatrix} 1657 & 0 & 532 & 51\,506 \\ 0 & 1669 & 273 & 41\,384 \\ 0 & 0 & 1 & 152 \end{bmatrix}.$$

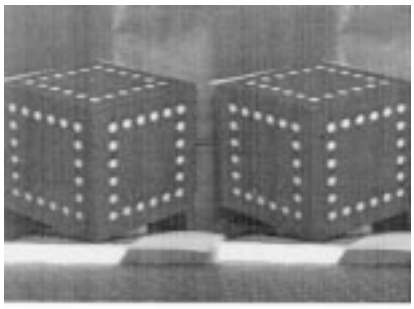
Fig. 18 shows epipolar lines obtained by the perspective projection matrices [17], [18]. The average residual is equal to 0.3214 pixels, which is mainly due to misalignment of the biprism with respect to the image plane of the camera. In the next section, we analyze the error analysis due to misalignment of the biprism.

B. Correspondence and IA

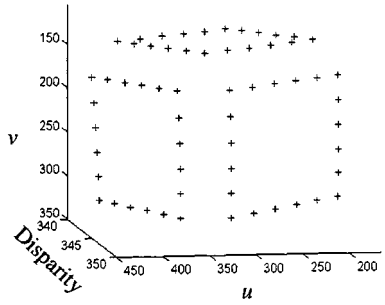
From the results of the calibration in the previous section, Fig. 19(a) shows the correspondence region that the corresponding points exist in the biprism-stereo image. The cross-hatching region represents the correspondence region as the FOV of a biprism-stereo camera system. From (19) and (20), the horizontal FOV of the camera θ_w and the transformed region ϕ_p are equal to 11.3° and 13.1°, respectively. The width and height of the CCD array of the used camera are 6.5 and 4.85 mm, respectively. The horizontal FOV of the camera is smaller than the transformed region by the biprism. From (21), the distance a_z is equal to 748.3 mm. As shown in Fig. 19(a), the correspondence region (FOV) can be obtained as follows:

$$\text{FOV} = \begin{cases} -6.6167° < \text{FOV} < 6.6167°, & \text{if } 0 < Z_p \le 748.3 \\ |X_p| < 0.0769 Z_p + 29.3[\text{mm}], & \text{if } 748.3 < Z_p. \end{cases} \quad (26)$$

From (17) and the results of the calibration, the IA in the image plane marked by two white lines in Fig. 19(b) is about 19 columns.



(a)



(b)

Fig. 17. Image and disparity map computed for a biprism calibration: (a) an input image with a biprism and (b) the computed disparity map.

TABLE III
CALIBRATED PARAMETERS k_1 AND k_2

k_1	k_2
0.3946	0.0026

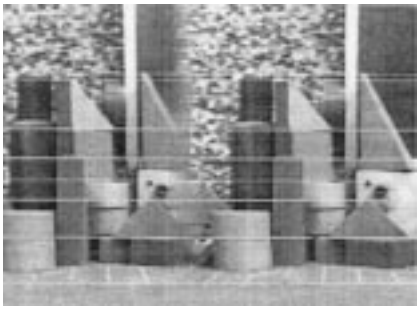
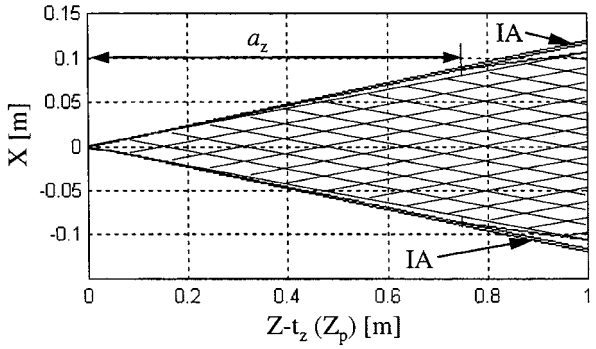
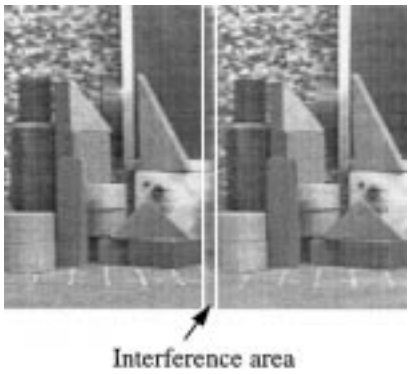


Fig. 18. Epipolar lines obtained by projection matrix.



(a)



(b)

Fig. 19. (a) Correspondence and IA. (b) IA in the biprism image.

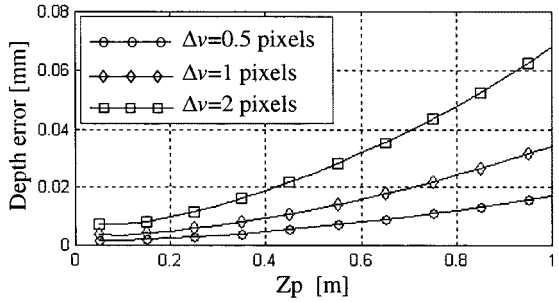


Fig. 20. Depth errors according to the distance Z_p.

C. Error Analysis

1) Depth Error Due to Misalignment: This section analyzes the depth error due to misalignment of the biprism with the image plane of the camera. To align the biprism with respect to the image plane of the camera, they should be adjusted so that the corresponding points lie on the same scanline in the image plane. But it is possible that the corresponding points from misalignment of the biprism may not lie on the same scanline. The image coordinate difference due to misalignment will be smaller than 1 pixel. From (12), the disparity between the corresponding points can be rewritten as

$$d = \sqrt{\Delta u^2 + \Delta v^2} \tag{27}$$

where $\Delta u = u_r - u_l$ and $\Delta v = v_r - v_l$.

We consider 3-D points at a distance of 500 mm. If the image coordinate difference is within 1 pixel, then the rotated angles of the biprism about the camera coordinates are smaller than 0.3°, respectively. Accordingly, we present the effect of the depth errors due to the image coordinate difference. From the results of the calibration in Tables II and III, Fig. 20 shows the depth error with respect to Δv at each distance Z_p. The ratio of Δv to Δu is very small. From the result of Fig. 20, we conclude that the effect of the depth error due to Δv is negligible. The average residual is 0.3214 pixels as shown in Fig. 18. The depth error by this error is within 0.01 mm. Therefore, we can define that the disparity by (12).

2) Depth Error Due to Mismatching:

Pixel Noise Model: We define the pixel noise as the degree of subpixel precision for matching. Also we assume that the noise model is the Gaussian distribution. Let (u_i, v_i) be the true and $(\tilde{u}_i, \tilde{v}_i)$ be noisy observation of (u_i, v_i), then we have

$$\tilde{u}_i = u_i + \xi_i \qquad \tilde{v}_i = v_i + \eta_i \tag{28}$$

where the noise terms ξ_i and η_i denote independently distributed noise terms having mean 0 and variance σ_i^2. Therefore

$$\begin{aligned} E[\xi_i] &= E[\eta_i] = 0 \\ V[\xi_i] &= V[\eta_i] = \sigma_i^2 \\ E[\xi_i \xi_j] &= \begin{cases} \sigma_0^2, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \\ E[\eta_i \eta_j] &= \begin{cases} \sigma_0^2, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \\ E[\xi_i \eta_j] &= 0. \end{aligned} \tag{29}$$

Disparity Variance: Let $(\tilde{u}_1, \tilde{v}_1)$ be the image coordinates of a point in the left image and $(\tilde{u}_2, \tilde{v}_2)$ be the image coordinates of its corresponding point in the right image. From these noisy measurements, we define the noisy disparity as follows:

$$\tilde{d} = \tilde{d}(\tilde{u}_1, \tilde{v}_1, \tilde{u}_2, \tilde{v}_2). \tag{30}$$

In order to determine the expected value and variance of $\tilde{d}$, we expand $\tilde{d}$ as a Taylor series at (u_1, v_1, u_2, v_2)

$$\begin{aligned} \tilde{d} &= d + \sum_{i=1}^{2} \left[(\tilde{u}_i - u_i) \frac{\partial \tilde{d}}{\partial \tilde{u}_i} + (\tilde{v}_i - v_i) \frac{\partial \tilde{d}}{\partial \tilde{v}_i} \right] \\ &= d + \sum_{i=1}^{2} \left[\xi_i \frac{\partial \tilde{d}}{\partial \tilde{u}_i} + \eta_i \frac{\partial \tilde{d}}{\partial \tilde{v}_i} \right]. \end{aligned} \tag{31}$$

Then, the variance becomes

$$E[(\tilde{d} - d)^2] = \sigma_0^2 \sum_{i=1}^{2} \left[\left(\frac{\partial \tilde{d}}{\partial \tilde{u}_i} \right)^2 + \left(\frac{\partial \tilde{d}}{\partial \tilde{v}_i} \right)^2 \right]. \tag{32}$$

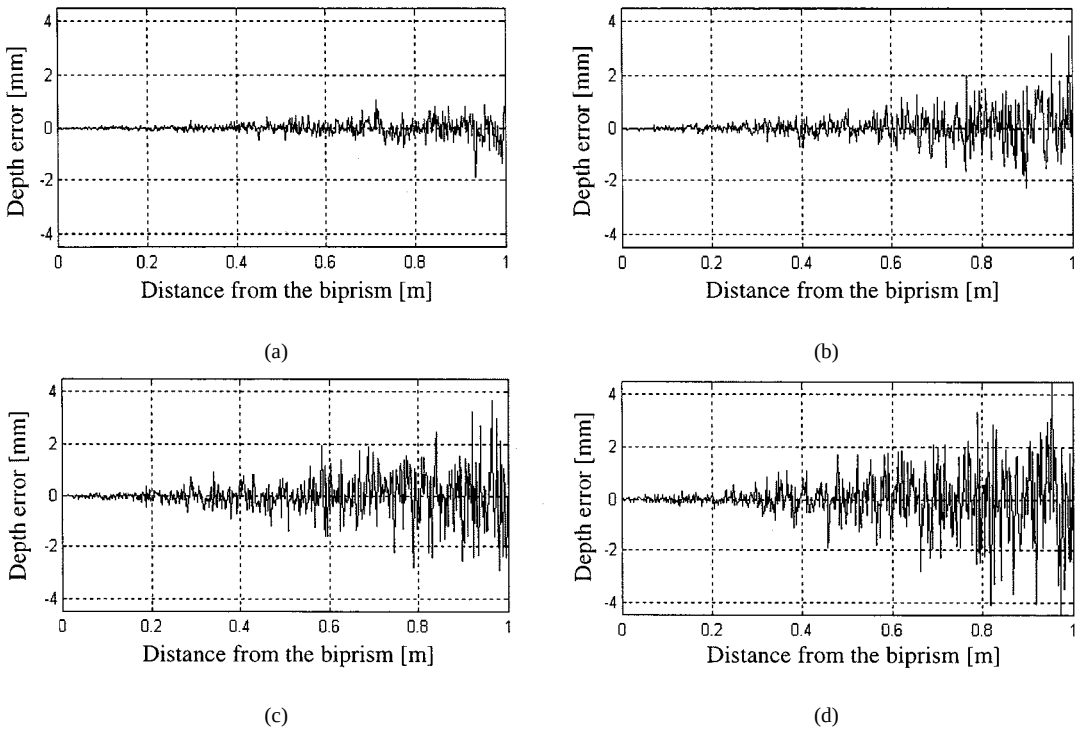


Fig. 21. Depth error according to the disparity variance: (a) $\mathrm{var}(\tilde{d}) = 0.25$; (b) $\mathrm{var}(\tilde{d}) = 0.5$; (c) $\mathrm{var}(\tilde{d}) = 0.75$; and (d) $\mathrm{var}(\tilde{d}) = 1$.

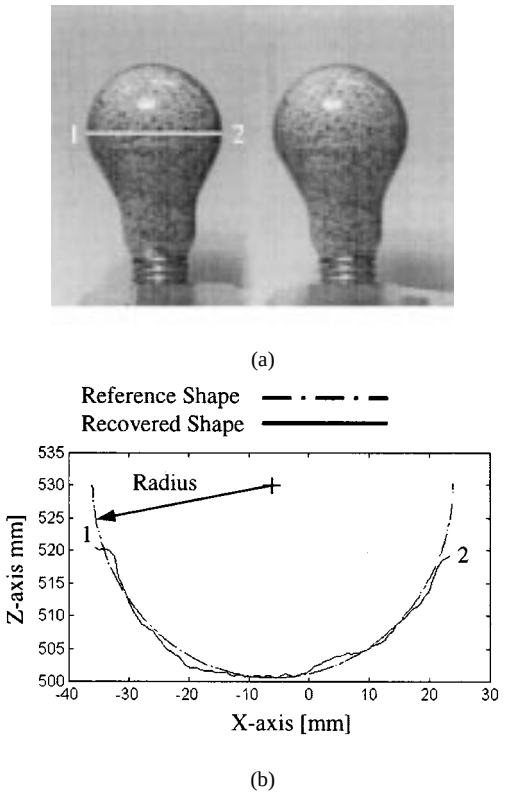


Fig. 22. Recovered shape with one scanline for a textured light bulb: (a) input biprism image and (b) recovered shape.

From (30), the disparity between the corresponding points in the biprism-stereo image becomes

$$\tilde{d} = \sqrt{(\tilde{u}_2 - \tilde{u}_1)^2 + (\tilde{v}_2 - \tilde{v}_1)^2}. \tag{33}$$

From (33), the partial derivatives in (32) can be represented as

$$\begin{aligned}
\left(\frac{\partial \tilde{d}}{\partial \tilde{u}_1}\right)^2 &= \frac{(\tilde{u}_2 - \tilde{u}_1)^2}{(\tilde{u}_2 - \tilde{u}_1)^2 + (\tilde{v}_2 - \tilde{v}_1)^2} \\
\left(\frac{\partial \tilde{d}}{\partial \tilde{u}_2}\right)^2 &= \frac{(\tilde{u}_2 - \tilde{u}_1)^2}{(\tilde{u}_2 - \tilde{u}_1)^2 + (\tilde{v}_2 - \tilde{v}_1)^2} \\
\left(\frac{\partial \tilde{d}}{\partial \tilde{v}_1}\right)^2 &= \frac{(\tilde{v}_2 - \tilde{v}_1)^2}{(\tilde{u}_2 - \tilde{u}_1)^2 + (\tilde{v}_2 - \tilde{v}_1)^2} \\
\left(\frac{\partial \tilde{d}}{\partial \tilde{v}_2}\right)^2 &= \frac{(\tilde{v}_2 - \tilde{v}_1)^2}{(\tilde{u}_2 - \tilde{u}_1)^2 + (\tilde{v}_2 - \tilde{v}_1)^2}.
\end{aligned} \tag{34}$$

From (31) and (34), the variance of the disparity can be obtained as

$$\mathrm{var}(\tilde{d}) = E[(\tilde{d} - d)^2] = 2\sigma_0^2. \tag{35}$$

From the above results, it is shown that the variance of the disparity is as twice large as the variance of the pixel noise.

Depth Error According to the Disparity Variance: From the results of the calibration in the previous section, we present the depth error according to the distance to the biprism with the disparity variance. Fig. 21 shows the depth error according to the disparity variance. The depth error increases with increase of the distance from the object to the biprism without considering the disparity variance. That is, it is readily seen that the depth error increases as the distance Z_p increases. Since the disparity difference becomes smaller as an object moves farther away from the biprism as shown in (12). Therefore, a biprism-stereo camera system is more accurate for nearby objects than for far ones.

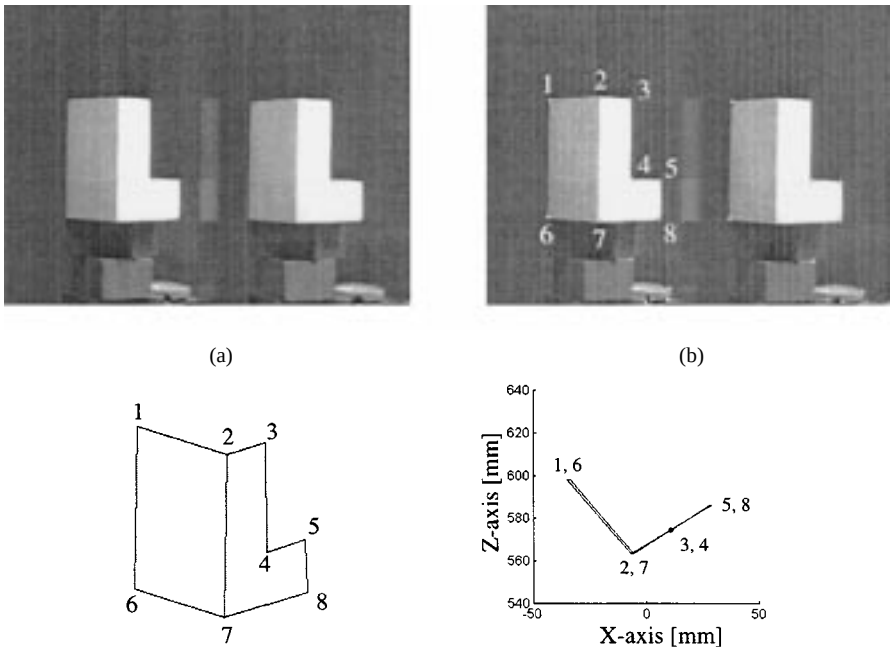


Fig. 23. Input image and recovered shape: (a) input biprism image; (b) extracted corner points; (c) recovered shape; and (d) top view of (c).

TABLE IV
RECOVERED LINE SEGMENT LENGTHS

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

From the results of the error analysis, it is evident that one major factor for the depth error is the degree of subpixel precision for matching.

D. Experimental Results

Fig. 22 shows an input biprism image and recovered shape with one scanline for a textured light bulb. To find the corresponding points, a simple cross-correlation technique is used. We use a 25 $\times$ 25 window to compute the SSD. A 3-D shape using the corresponding points can be reconstructed. The radius of a light bulb is 29.9 mm. The errors in radius between the recovered shape and the true shape of the light bulb is smaller than 2 mm. Fig. 23(a) shows an input biprism image and the recovered shape for a block. We have used the pixel coordinates of the vertexes of the block with the pairs of corresponding points. Fig. 23(b) shows the extracted corner points. Fig. 23(c) shows the recovered shape for the block. Table IV shows the error in length between the recovered shape and the true shape of the block.

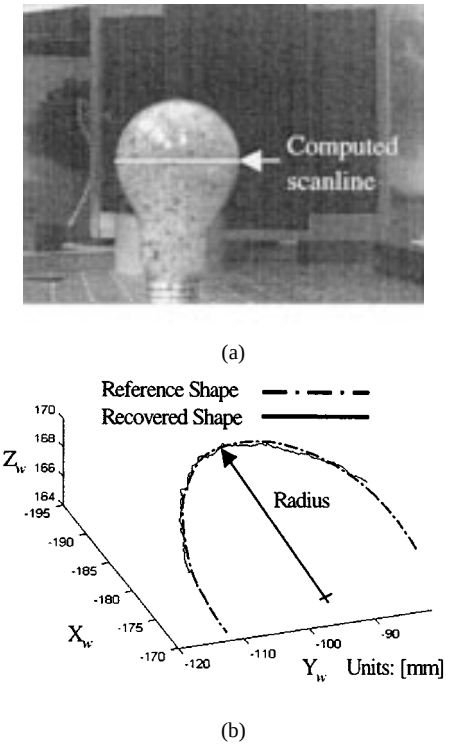


Fig. 24. Recovered shape with one scanline in the two cameras with parallel optical axis: (a) right image and (b) recovered shape.

We have also performed an experiment using a two-camera stereo system with parallel optical axis under the same experimental conditions as shown in Fig. 22. The coordinate (X_w, Y_w, Z_w) is the world coordinate system of the calibration box. In Fig. 24, we perform the 3-D reconstruction for the same

scanline. The maximum error in radius is 1.8 mm. From the above results, we conclude that the accuracy of the biprism stereo is comparable with that of the two-camera stereo system.

VII. Conclusions

In this paper, a single-lens camera system that can provide a pair of stereo images using a biprism has been introduced. A stereo image obtained by the proposed camera system is equivalent to the two images obtained from the two cameras with exactly the same optical properties, which greatly simplifies the calibration of the system. Also the corresponding points in the biprism-stereo image lying on the same scanline facilitate the correspondence process. A larger FOV using multiple two biprism-stereo images can be obtained. It has been also demonstrated that the 3-D map using the 3-D structure can be reconstructed from a sequence biprism-stereo images.

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 applications including dynamic scenes.

References

[1] S. T. Barnard and M. A. Fischler, "Computational stereo," *Comput. Surv.*, vol. 14, no. 4, pp. 553–572, 1982.

[2] U. R. Dhond and J. K. Aggarwal, "Structure from stereo: A review," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 6, pp. 1489–1510, 1989.

[3] *Handbook of Pattern Recognition and Image Processing: Computer Vision*, Academic, New York, 1994.

[4] V. S. Nalwa, *A Guided Tour of Computer Vision*. Reading, MA: Addison-Wesley, 1993.

[5] D. H. Lee, I. S. Kweon, and R. Cipolla, "Single lens stereo with a biprism," in *Proc. IAPR Int. Workshop Machine Vision and Applications (MVA'98)*, Japan, 1998, pp. 136–139.

[6] D. H. Lee, I. S. Kweon, and R. Cipolla, "A biprism stereo camera system," in *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition (CVPR'99)*, 1999, pp. 82–87.

[7] Y. Nishimoto and Y. Shirai, "A feature-based stereo model using small disparities," in *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition (CVPR'87)*, 1987, pp. 192–196.

[8] W. Teoh and X. D. Zhang, "An inexpensive stereoscopic vision system for robots," in *Proc. Int. Conf. Robotics*, 1984, pp. 186–189.

[9] A. Goshtasby and W. A. Gruver, "Design of a single-lens stereo camera system," *Pattern Recognit.*, vol. 26, pp. 923–936, 1993.

[10] S. Nene and S. Nayar, "Stereo with mirrors," in *Proc. Int. Conf. Computer Vision (ICCV'98)*, 1998, pp. 1087–1094.

[11] M. Bass and W. V. Stryland, *HandBook of Optics I, II*. New York: McGraw-Hill, 1995.

[12] E. Hecht, *Theory and Problems of Optics*, ser. Schaum's Outline Series. New York: McGraw-Hill, 1975.

[13] F. A. Henkins and H. E. White, *Fundamentals of Optics*. New York: McGraw-Hill, 1976.

[14] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MA: MIT Press, 1993.

[15] G. Xu and Z. Zhang, *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach*. Norwell, MA: Kluwer, 1996.

[16] O. Faugeras and G. Toscani, "The calibration problem for stereo," in *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition (CVPR'86)*, 1986, pp. 15–20.

[17] Z. Zhang, "Determining the epipolar geometry and its uncertainty: A review," INRIA, Res. Rep. 2927, 1996.

[18] R. I. Hartley, "In defence of the 8-point algorithm," in *Proc. Int. Conf. Computer Vision (ICCV'95)*, 1995, pp. 1064–1070.

DooHyun Lee received the B.S. degree in mechanical design and production engineering from Hanyang University, Seoul, Korea, in 1987, and the M.E. degree in production engineering from Korea Advanced Institute of Science and Technology (KAIST), Seoul, Korea, in 1989. He is currently working toward the Ph.D. degree in automation and design engineering at KAIST.

From 1989 to 1995, he worked as a Research Engineer for LG Electronics Company. His current research interests include geometric modeling, 3-D image processing, and analysis and machine vision.

InSo Kweon (M'95) received the B.S. and M.E. degrees in mechanical design and production engineering from Seoul National University, Seoul, Korea, in 1981 and 1983, respectively, and the Ph.D. degree in robotics from Carnegie Mellon University, Pittsburgh, PA, in 1990.

From 1991 and 1992, he was a Visiting Scientist in the Information Systems Laboratory at Toshiba Research and Development Center, where he worked on behavior-based mobile robots and motion vision research. Since 1992, he has been an Associate Professor of Electrical Engineering at Korea Advanced Institute of Science and Technology (KAIST). His current research interests include image sequence analysis, physics based vision, invariants and geometry, and 3-D range image analysis.

Dr. Kweon is a member of the IEEE Computer Society.