Journal of the Optical Society of Korea, Vol. 10, No. 1, March 2006, pp. 37-41

A Method for Precise Depth Detection in Stereoscopic Display

Seung Woo Lee* and Nam Kim

School of Electronic and Computer Engineering, ChungBuk National University 12 Gaeshin-dong, Heungduk-gu, Chongju, Chungbuk 361-763, KOREA

(Received October 14, 2005 : revised January 23, 2006)

This paper discusses a method for precise depth detection in stereoscopic images. The geometry of a stereoscopic camera and a display system is presented. It was found that there exists a difference between the calculated depth and the perceived depth of objects. For precise depth detection, the proposed registration method is investigated and proved by experiment in stable condition.

OCIS codes : 100.0100, 100.6890

I. INTRODUCTION

The purpose of a stereoscopic display is to create 3D images from 2D images in order to give a greater feeling of reality when displaying images. It is now possible to watch images that are displayed as 3D images at home, at an exhibition place, during events and even in theme parks. The 3D images have different effects and compare 3D images such as pictures, photographs and real-time video. For example, they can be so life-like that viewers react with their arms and appear to dodge from the moving stereoscopic images in front of their eyes as they watch the special images. The combination virtual space and real space is one of the most important subjects in the field of virtual reality. When watching the stereoscopic images, that is, the virtual space used by the stereoscopic display system, it would be useful to know that the object is located in a precise position compared to a real space. In cases such as these, it is necessary to perceive that a virtual object in a 3D system is located in a precise or imprecise position. In order to facilitate this, it is also necessary to detect the depth value from the stereoscopic images. This is one of the most important aspects of the stereoscopic images.[1-5]

Recently, many researchers in the area of stereoscopic image acquisition have been investigating the precise depth value detection by means of various methods in order to obtain clearer and more accurate images. It is necessary to define various parameters concerning the stereoscopic camera and display system for a precise depth value detection. A mathematical expression is needed concerning the connection of the CCD camera and the object in the stereoscopic camera system; it is also necessary to understand the additional expressions concerning the connection between the human eye and the stereoscopic display system. However, this mathematical expression is very complex; additionally if other factors such as environmental factors, human psychological factors, or distortions of the lens are considered, the numerical formula becomes even more multifaceted.[6] Hence, in this study, the distortion which can occur in the CCD or the lens as well as the psychology of a person's vision is not considered.

In Section 2 and 3, the stereoscopic camera system and stereoscopic display system is illustrated, and a procedure to determine the optimal values of the parameters is presented. In Section 4, the effectiveness of the proposed method is discussed from several aspects and the experimental results are presented. Finally, the experiment results and the computer simulations are given in the conclusion.

II. STEREOSCOPIC CAMERA SYSTEM CONFIGURATION

Fig. 1 illustrates the stereoscopic camera system. The stereoscopic camera system consists of a pair of CCD camera mounted side-by-side to obtain left and right images. There are three basic parameters that compose the stereoscopic camera system. Fig. 1 illustrates these parameters. The camera system configuration is determined by the convergence distance, the field of view and the camera separation. The convertgence distance is the distance from the lens to where the optical axes of the camera intersect, and the field of view of the cameras is determined by the CCD format size and the lens focal length. The camera separation

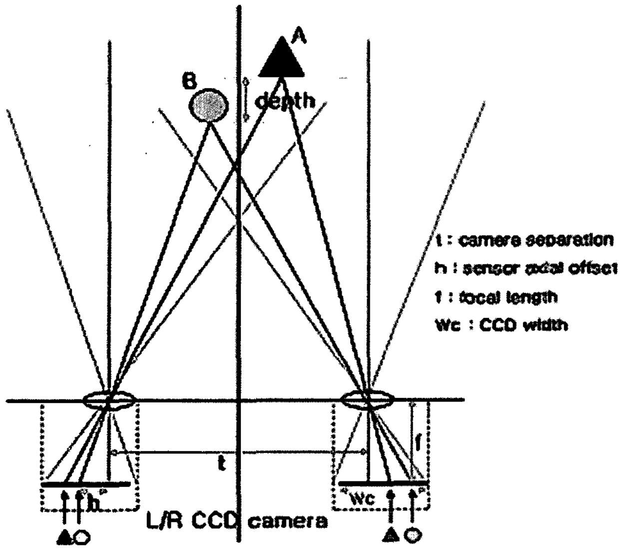

FIG. 1. The stereoscopic camera system.

is the distance between the two cameras.

In this system, a parallel system of the camera configurations is considered. One of the most important advantages in this system is that it avoids the distortions such as the keystone distortion and the depth-plane curvature distortion. However, the most important point is that the data processing of the parallel camera configuration is more convenient and easier to use than the toed-in camera configuration. If the parallel camera system is implemented without an image shift, the convergence, or zero parallax distance, will be at infinity and all images will be cast in front of the display screen. In most cases, the parallel camera configuration should be implemented with some horizontal image shift so that a more suitable convergence or zero parallax distance is defined.[7,8]

The geometry of the stereoscopic camera system can be determined by considering the imaging and display process as three separate coordinate transforms. Firstly X, Y, Z coordinates in real object space are transformed to X,Y positions on the two camera imaging sensors (left and right CCD). Secondly, the two sets of CCD coordinates are transformed to X, Y positions of the left and right images on the stereoscopic display device. Finally, the set of X, Y ,Z coordinates in the image space is transformed to the viewer space. The first coordinate transforms are expressed in Eqs. (1), (2) and (3).

$$X_{cl} = f\left(\frac{t+2X_o}{2Z_o}\right) - h \tag{1}$$

In Eq. (1), X_{cl} is the X coordinate of the left CCD from the object point (X_o, Y_o, Z_o), f is the focal length of the camera lens, is the distance between the first nodal points of the two camera lenses, and is the sensor axial offset. In the parallel camera configuration with image shift, is the distance by which the center of each CCD has been moved away from the optical axis of the lens to achieve convergence.

$$Y_{cl} = \frac{Y_o f}{Z_o} \tag{2}$$

In Eq. (2), is the Y coordinate of the left CCD. Eqs. (1) and (2) express the left CCD coordinate.

For the right CCD, the transformation is the same process as shown in Eqs. (1) and (2). Eq. (3) expresses the right CCD coordinates.

$$X_{cr} = f\left(\frac{t-2X_o}{2Z_o}\right) + h \ , \ Y_{cr} = \frac{Y_o f}{Z_o} \tag{3}$$

The equations are established by a trigonometric function of a geometric system. Many researchers have recognized these expressions, but here they are reconstructed in simple expressions.

III. STEREOSCOPIC DISPLAY SYSTEM CONFIGURATION

In this section, the stereoscopic display system for watching the 3D images is illustrated. In reality, the stereoscopic display system consists of a single display surface on which left and right images are displayed and separated by coding methods such as the time, color or polarization. Fig. 2 shows the concept of the 3D display device. When the screen is watched, it is possible to see the stereoscopic objects and 'feel' the depth by the screen. This system also needs three basic parameters. The display system configuration is determined by the screen size, the eye separation and the viewing distance. The size of the screen is measured by its horizontal width and the eye separation means the distance between the viewer's eyes. The viewing distance is the distance from observer's eyes to the screen.

There is a secondary transformation, done when the coordinates of the objects on the CCD camera are obtained. At that point, the CCD coordinates are transformed into the screen coordinates. The transformation from CCD coordinates to screen coordinates is achieved by multiplying by the screen magnification factor M. Eqs. (4) and (5) are expressed the transformations of left/right CCDs.

$$X_{sl} = MX_{cl} \ , \ Y_{sl} = MY_{cl} \tag{4}$$
$$X_{sr} = MX_{cr} \ , \ Y_{sr} = MY_{cr} \tag{5}$$

The final transform from the screen coordinates to the image space coordinates is shown in Eqs. (6), (7) and (8). When one watches the screen of the stereoscopic device, the brain combines the two pieces of the

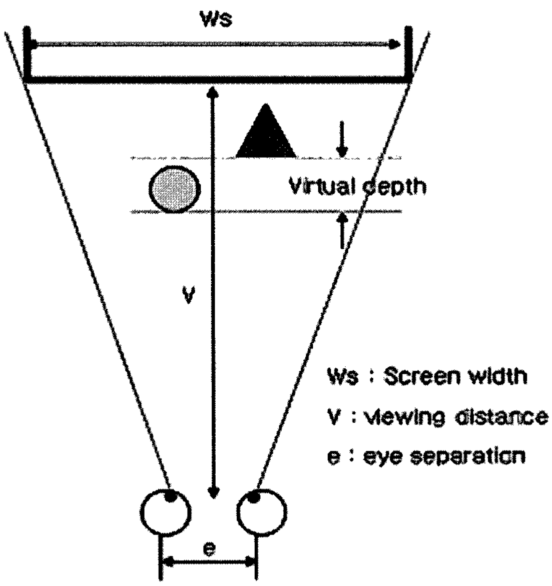

FIG. 2. Stereoscopic display system.

picture and creates the 3D image.

$$X_i = \frac{4Mfe\frac{X_o}{Z_o}}{e-2Mh+Mf\frac{t}{Z_o}} \quad (6)$$

$$Y_i = \frac{Y_s e}{e-2Mh+Mf\frac{t}{Z_o}} \quad (7)$$

$$Z_i = \frac{Ve}{e-2Mh+Mf\frac{t}{Z_o}} \quad (8)$$

Z_i is the depth value on the virtual space, and V is the viewing distance from the viewer's eyes to the screen.

These equations apply only to a parallel camera system. However, it should be noted that these equations do not contain any small-angle approximations. It has been found that small-angle approximations can cause some stereoscopic distortions to be uncertain. In this paper, these distortions are not considered.

IV. EXPERIMENTS

First, the stereoscopic camera system was constructed. Fig. 3 shows this system. There are two cameras on the table and the objects are located in front of the cameras. In addition, an optical table was used, which was flat and had non-vibration properties that would mitigate errors.

Fig. 3 shows the parallel camera system which was created for the experiment. For the parameters of the system, the camera separation was 75 mm between two cameras, and the focal length of the camera was 6.5 mm. The CCD size was 2/3 inch. The horizontal width for the 1/2 and 2/3 inch CCDs were 6.4 and 8.8 mm, respectively. The CCD resolution was 640×480 pixels. The fundamental distance from the CCD lenses to the object was 350 mm. The two cameras were connected to a computer with a frame grabber. A frame grabber is a device which can transmit images from the camera to memory through an A/D converter.

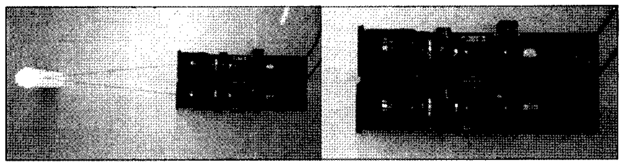

FIG. 3. The real stereoscopic camera system.

The present purpose is to obtain precise depth values on the stereoscopic images. If the precise depth value of objects in random positions could be found, this would fulfill the present purpose. Accordingly, four types of cases were suggested.

First, Object A was located at the center of the two cameras and 350 mm in front of the cameras. Object B was located 100 mm away on the right side from Object A. Second, Object A was located at the center of the two cameras and 350 mm in front of the cameras. Object B was located 100 mm from the back of Object A. Third, Object A was located at the center of the two cameras and 350 mm in front of the cameras. Object B was located at both 100 mm on the right and the back side from Object A. Finally, Object A was located at the center of the two cameras and 350 mm in front of them. Object B was located 100 mm on the left side and 100 mm from the back side far of Object A. Each case in an actual photo is shown in Figs. 4 (a), 5 (a), 6 (a) and 7 (a). Moreover, pictures are shown which were obtained from the left camera in Figs. 4 (b), 5 (b), 6 (b) and 7 (b) and from the right camera in Figs. 4 (c), 5 (c), 6 (c) and 7 (c).

In this experiment, it was necessary to understand that the images must satisfy a camera condition which was intended to concentrate on the objects clearly, as the distinct images created the precise depth value. Pictures from each case were obtained and the depths were determined by computer calculations. Fig. 8 shows the computer simulations of each case.

Object A, which is a red color circle in the programming result is a cylinder in the picture; correspondingly, Object B is a blue color square and a square pillar. For a more precise result, the experiments were conducted in random positions. Fig. 9 shows the other computer simulations.

Upon seeing the graphs, it can be noted that an error exists in each case. In the case of Object A, the error is nearly zero, but Object B has an error which is

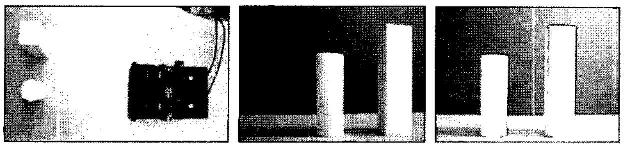

FIG. 4. Case 1: the objects are located at (0,350), (100,350). (a) the object's position (b) left image on the left CCD (c) right image on the right CCD.

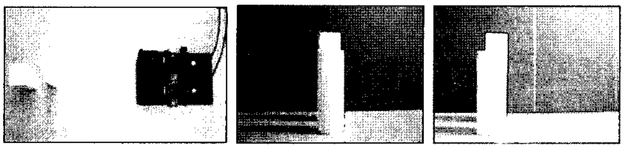

FIG. 5. Case 2: the objects are located at (0,350), (0,450). (a) the object's position (b) left image on the left CCD (c) right image on the right CCD.

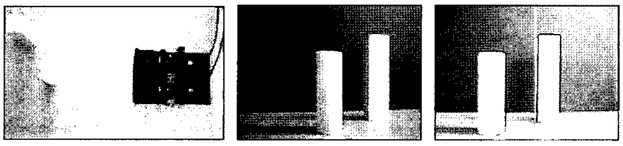

FIG. 6. Case 3: the objects are located at (0,350), (100,450). (a) the object's position (b)left image on the left CCD (c) right image on the right CCD.

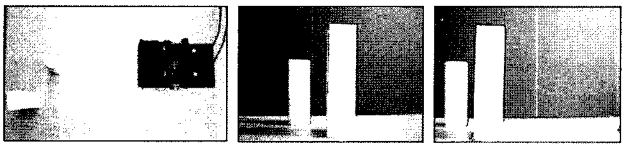

FIG. 7. Case 4: the objects are located at (0,350), (-100,450). (a) the objects position (b) left image on the left CCD (c) right image on the right CCD.

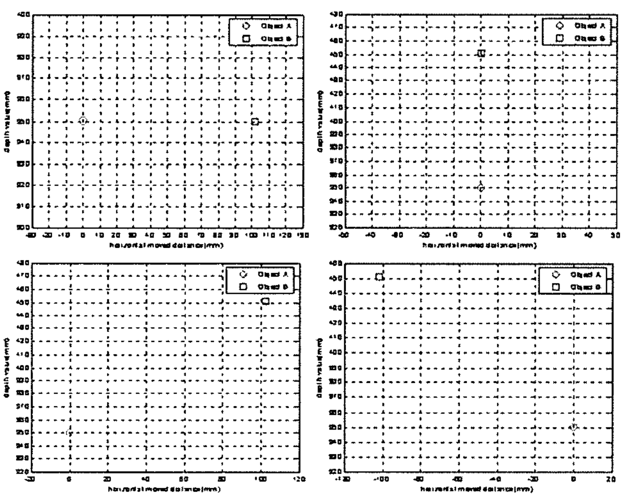

FIG. 8. The computer simulation results about the pictures for the case 1 to 4. Object A is the red color circle, and object B is the blue color square.

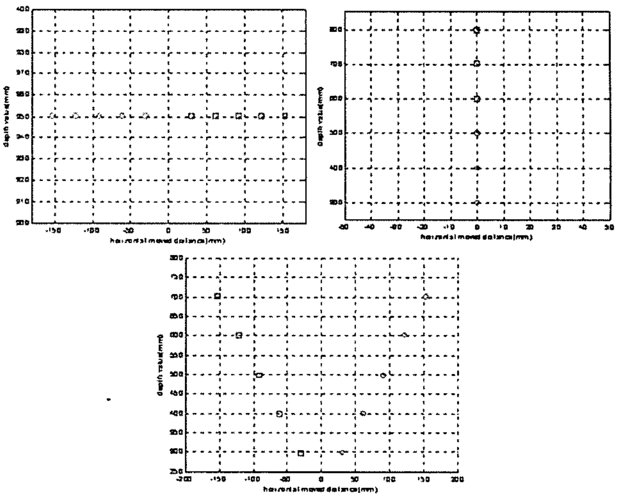

FIG. 9. The other cases of computer simulation results in which the objects are located in random positions.

within 1~2 mm. The reason for these errors is the precise focusing concerning Object B and the lens distortions. The experiment was conducted several times, but it was still too difficult to take precisely focused pictures. In Section 2, the advantages of the parallel configuration of the camera are explained. After reviewing these results, it is now known that if the parallel camera method does not concentrate clearly, it is not possible to obtain the precise depth value.

V. CONCLUSION

In this paper, the precise depth detection is sought in stereoscopic images using the transformation from actual objects to a stereoscopic screen through two CCD cameras. A camera configuration consisting of a parallel system is proposed. The parallel configuration has many advantages. This configuration avoids distortions such as keystone distortion and depth-plane curvature distortion. However, it is necessary to understand that the parallel configuration system must clearly focus on the objects. As seen by computer simulations, a graph created from the distance and depth of objects was obtained. It has an error, which is less than 1~2mm, in the comparison between the real space and image space. However, the depth value in the computer calculations can be verified. The result of the first case is 0 mm in the X-axis and 350 mm in the Z-axis for Object A, and 102 mm in the X-axis and 350 mm in the Z-axis for Object B. In the second case, the results for Object A are 0 mm in the X-axis and 350 mm in the Z-axis. The results for Object B are 0 mm in the X-axis and 451 mm in the Z-axis. In the third case, the results for Object A are 0 mm in the X-axis and 350 mm in the Z-axis and the results for Object B is 102 mm in the X-axis and 451 mm in the Z-axis. In the last case, the result for Object A is 0 mm the X-axis and 350 mm in the Z-axis and the results for Object B is 102 mm in X-axis and 451 mm in Z-axis.

It is believed that the reason for the error was caused by imprecise focusing, as when the 3D images were created using the non-focused 2D images, the displayed images were not focused. Therefore, a precise focusing is a very important point in this experiment. In conclusion, it was seen that the precise depth detection technique is useful and applicable for checking the size of microscopic objects.

ACKNOWLEDGEMENT

This work was supported by the Regional Research Centers Program of the Ministry of Education & by Human Resources Development in Korea and by the Program for Cultivating Graduate Students in Regional Strategic Industry.

*Corresponding author : swlee@osp.chungbuk.ac.kr

REFERENCE

[1] S. Yoshida, S. Miyazaki, T. Hoshino, J. Hasegawa, T. Ozeki, T. Yasuda, and S. Yokoi, "A technique for precise depth representation in stereoscopic display," *Computer Graphics International*, pp. 80-84, 1999.

[2] S. Nagata, "The binocular fusion of human vision on stereoscopic displays - field of view and environment effect," *Ergonomics*, vol. 39, no. 11, pp. 1273-1284, 1996.

[3] M. Deering, "High resolution virtual reality," *Computer Graphics*, vol. 26, no. 2, pp. 195-202, 1992.

[4] S. Ohtsuka, Y. Ishigure, Y. Kanatsugu, T. Yoshida, and S. Usui, "Virtual window: a technique for correcting depth-perception distortion in stereoscopic displays," *The Society for Information Display Symposium Digest*, vol. 27, pp. 893, 1996.

[5] S. Yoshida, "An improvement method of depth perception in stereoscopic display," *International Archives of Photogrammetry and Remote Sensing*, vol. XXXII, part 5, pp. 386-391, 1998.

[6] A. Woods, "Image distortions in stereoscopic video systems," *Stereoscopic Displays and Applications IV, Proc. SPIE*, vol. 1915, pp. 36-48, 1993.

[7] A. Woods, M. Bolas, J. Merritt, and S. Benton, "Stereoscopic displays and virtual reality systems IX," *Proc. SPIE*, vol. 4660, 2002.

[8] Eunsoo Kim and Seunghyun Lee, *The Fundamentals of 3D Images* (Kidari, Seoul, Korea, 1998), pp. 73-87.