

3D Gaze Estimation and Interaction to Stereo Display



Yong-Moo Kwon, Kyeong-Won Jeon, Jeongseok Ki,
Qonita M. Shahab, Sangwoo Jo and Sung-Kyu Kim

Abstract— There are several researches on 2D gaze tracking techniques to the 2D screen for the Human-Computer Interaction. However, the researches for the gaze-based interaction to the stereo images or 3D contents are not reported. The stereo display techniques are emerging now for the reality service. Moreover, the 3D interaction techniques are needed in the 3D contents service environments. This paper presents 3D gaze estimation technique and its application to gaze-based interaction in the parallax barrier stereo display.

Index Terms— 3D Gaze Estimation, 3D Interaction, Stereo Display, 3D Computer Mouse

I. INTRODUCTION

Many researches on gaze estimation technique for inferring a fixation point on the 2D screen have been reported [1-4]. One of the typical methods is PCCR (Pupil Center & Corneal Reflection)-based approach [1]. The model-based approach is also reported in which gaze direction is estimated from monocular or stereo camera while using 3D face model [2-3]. The gaze direction estimation technique is proposed while using stereo vision technique for the estimation of 3D eye position and gaze direction [4]. However, all of these techniques address the estimation of gaze fixation point on the 2D screen.

In this paper, we address 3D gaze direction and depth estimation technique and its application to the gaze-based interaction in the stereo display. There are several researches on eye vergence and movement, although their main objectives are not for the gaze computer interaction [5-6]. Recently, a research on 3D gaze estimation is reported while using anaglyph-based HMD and commercial binocular eye tracking system in the context of the investigation of human visual system in 3D scene [7].

Fig. 1 shows the difference between 2D gaze and 3D gaze concept. As shown in Fig. 1, the 2D gaze to a fixation point on the 2D screen can be estimated with single eye. However, the

3D gaze can be estimated with both eyes, because 3D gaze needs not only gaze direction but also gaze depth. Until now, the most researches focus on only single-eye based gaze direction estimation for the application to the 2D display.

In this paper, we address the estimation of 3D gaze using the gaze direction and the gaze depth from both eyes. The estimation of gaze direction and gaze depth from both eyes is a new important research topic for 3D gaze interaction. The contribution of this paper can be summarized as follows. This paper presents an algorithm for estimating the 3D gaze, which only needs one monocular camera and IR LEDs. The proposed algorithm is implemented and applied to the parallax barrier type stereo display while using 3D contents. Moreover, our technique is applied to the 3D gaze-based interaction to object in 3D virtual space and evaluated.

We first propose our algorithm for the 3D gaze estimation that is composed of the gaze direction estimation and the gaze depth estimation. Second, we describe the implementation of our algorithm and its application to the parallax-barrier stereo display system with 3D contents. The evaluation of our system is also provided.

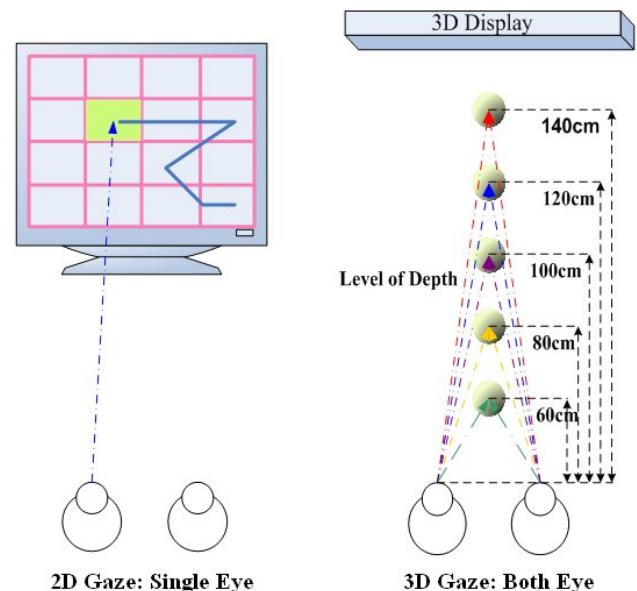


Fig. 1. The concept of 2D and 3D gaze

II. RELATED WORK

The researches on gaze estimation and its application have been reported. However, most of them address the gaze estimation of fixation point on the 2D screen. Although the research objectives are not the gaze estimation, there are

Manuscript Received on September 25, 2006.

Yong-Moo Kwon, Korea Institute of Science & Technology

Kyeong-Won Jeon, Korea Institute of Science & Technology

Jeongseok Ki, Korea Institute of Science & Technology

Qonita M. Shahab, Korea Institute of Science & Technology

Sangwoo Jo, Korea Institute of Science & Technology

Sung-Kyu Kim, Korea Institute of Science & Technology ,

E-mail:ymk@kistmail.kist.re.kr

several studies on eye vergence and movement. Recently, the research on 3D gaze estimation and recording is reported.

A. Gaze Direction Estimation

This category addresses the gaze direction estimation for inferring a fixation point on the 2D screen. The one of the typical method is PCCR technique that uses the relation between the position of pupil and the first Purkinje image - the glints [1]. When infrared light is shone into the user's eye, the reflection occurs on the surface of cornea. According to the seeing object position, the distance between the pupil center and the glint changes. The reason is the change of pupil center according to the seeing object position while the glint position is not changed under the condition there is no head movement.

Another gaze estimation method is a model-based 3D gaze tracking technique that uses a 3D model of some facial feature points to provide 3D information of the line of sight from a monocular or stereo image [2,3].

As an approach using 3D position of eye and gaze direction, the estimation method of 3D line-of-sight (LOS) of a user has been proposed. This method employs stereo cameras to determine the cornea center and the gaze direction using linear solutions [4].

B. Eye Vergence

It is well known that there is a mismatch between accommodation and vergence in stereo display system. Many researches on this mismatch issue have been reported. A HMD type vergence monitoring system is developed for the better understanding about accommodation and vergence mismatch to symptoms and visual problems associated with viewing stereoscopic imagery [5].

There is a research on the relationship between vergence eye movement pattern (time-vergence angle) and autostereogram image granularity (fine or coarse) [6].

C. 3D Gaze Estimation

A research on a neural-net based 3D gaze estimation and recording is introduced while using anaglyph-based HMD for the stereo display and the commercial binocular eye tracker for the acquisition of subject's eye position data from each eye. This paper also addresses the importance of 3D gaze estimation in view of investigation of human visual system in 3D scene [7].

III. 3D GAZE ESTIMATION

A. 2D Gaze vs 3D Gaze

Fig. 2 shows the basic concept of gaze direction and gaze depth while comparing the basic concept of 2D gaze and 3D gaze. As shown in Fig. 2, the gaze depth is only needed for 2D gaze, because the depth of screen is fixed. So, we can decide the gaze fixation point on the 2D screen only using gaze direction from single eye. In case of 3D gaze, we need both of gaze direction and gaze depth using both eyes.

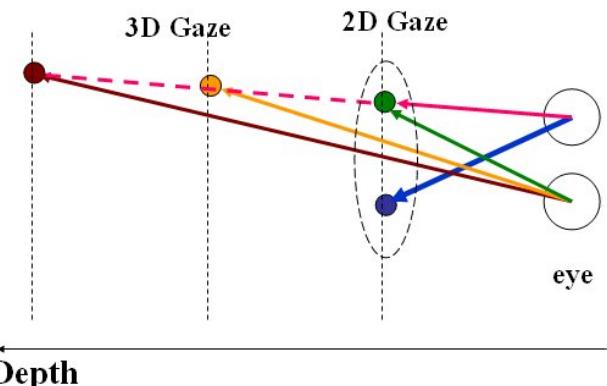


Fig. 2. Gaze direction and gaze depth

The gaze direction can be estimated based on the degree of eye rotation according to the gaze point. More exactly, the center of pupil is changed when the gaze point is varied.

By the way, it should be noted that there are several gaze points in the same gaze direction in case of 3D space. That is the reason why we need information from both eyes. It is also important to note that the distance between two pupil centers varies according to the target gaze point, i.e., the deeper target point, the longer the distance between two pupil centers of both eyes.

B. Gaze Direction Estimation

We use the relationship between pupil center and gaze point on the screen. It can be understood easily that the pupil center is changed according to the gaze point on the target point of the screen. Here we use Purkinje image for eye, which includes the pupil and glint from IR LED. Using image processing, the pupil center (PC) and glint center (GC) are extracted. The GC is used for the reference to compute the relative PC position according to gaze movement. To estimate the gaze point on the screen, we use distance values among three feature points that are pupil center P and two glints' centers, g1, g2, shown in Fig. 5. The perpendicular foot onto the line connected between g1 and g2 is M. The value of h is a height of line connected between P and M, and c and d are the distances of line between g1, M and M, g2 respectively.

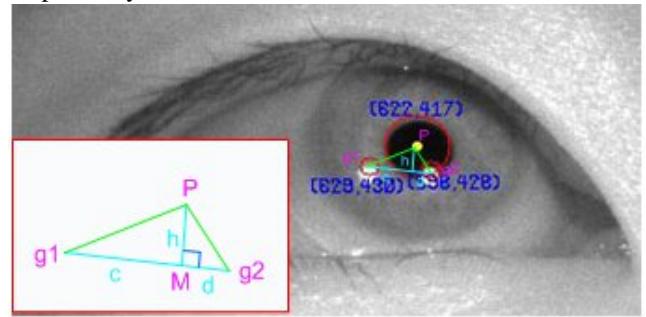


Fig. 3. Feature relations of gaze tracking algorithm

Assumed that the linear relations between the distance values among three feature points and the gaze point, the values of c and h are used as the corresponding information to the gaze point on the screen in the x-axis and the y-axis, respectively. By

the way, because field of view is different about each user, we perform the personal calibration. Through the calibration, the minimum and maximum values of c and h are extracted and stored when user is looking at the left-top and right-bottom corner of the screen. While using these stored values in the personal calibration as the personal reference data, the system calculates the gaze point on the screen using the extracted value of c and h according to the change of pupil center.

C. Gaze Depth Estimation

We use the concept of Pupil Center Distance (PCD) between both eye pupil centers for finding the depth. The idea is that the PCD changes according to depth. It should be noted that the PCD increases and decreases when the seeing object appears far and near from eye, respectively. Fig. 4 shows the relationship between the gaze depth and PCD.

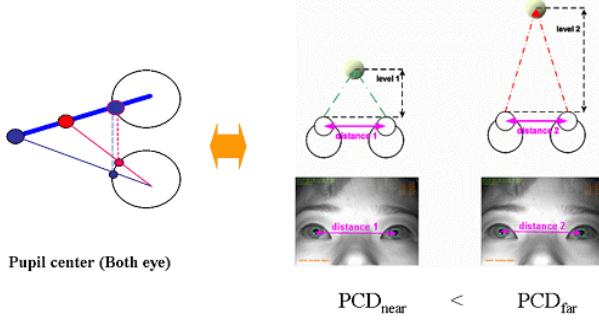


Fig. 4. The relationship between gaze depth and PCD

Based on the concept shown in Fig. 4, we capture the image for both eyes and extract the PCD value. Then we can estimate the gaze depth using PCD value. It should be mentioned that we need preliminary calibration for the personal relationship between gaze depth and PCD.

D. 3D Gaze Estimation Algorithm

The principal algorithm for 3D gaze estimation is shown in Fig. 5. There are two steps for 3D gaze estimation, i.e., first estimate the gaze direction using triangle method described in Section III(B) and the gaze direction using PCD described in Section III(C).

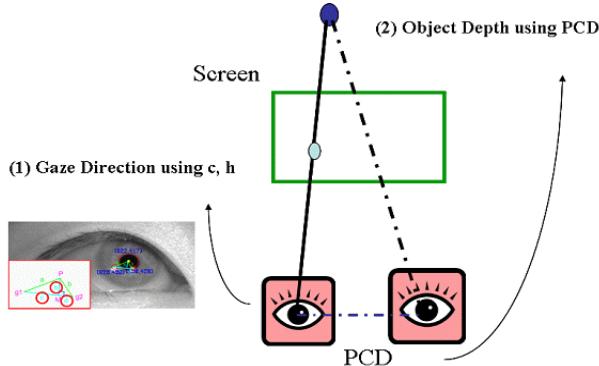


Fig. 5. Overview of 3D Gaze Estimation Algorithm

Fig. 6 shows the concept for finding 3D gaze target point or region from the gaze direction and the gaze depth. As shown in Fig. 6, we first estimate the gaze direction using c and h value

and then estimate gaze depth plane using PCD value and finally estimate the target point or region by intersecting the gaze direction and gaze depth plane.

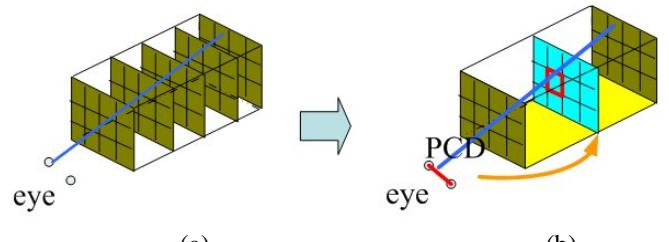


Fig. 6. Find 3D gaze target from the gaze direction and the gaze depth

IV. IMPLEMENTATION OF 3D GAZE ESTIMATION SYSTEM

A. System Overview

We use two IR LEDs and one monocular 1394 camera with infrared filter (LP830-27). We also use chin-rest for fixing head. For 3D display device we use parallax barrier type stereo display. (Model No PAVONINE PA3D-17EXN).

B. Gaze Direction

Our algorithm for the estimation of 2D gaze point on screen is based on Purkinje image. Our system has one monocular camera and two infrared light sources. These two light sources make two glints on the surface of user's eye called Purkinje image. Briefly our gaze tracking procedure can be divided into 4 steps. The system captures image of user's eye and estimates pupil and glints center using image processing techniques. Then the system calculates distance values among three feature points which are pupil center and two glints center and estimates the gaze point by mapping feature points relation onto the screen.

C. Gaze Depth

In view of gaze depth, the pupil center movement is first simulated using simple eye ball model while changing the depth of target point on the screen. Using this simulation, we can estimate the PCD variations theoretically according to the variation of target point in the 3D space. Fig. 7 shows our simulation result for estimating the relationship between PCD and target depth. It should be noted that the target depth is not linear compared to the steps of PCD value increase. So, we estimate the target depth values according to the linear increase of PCD values. Fig. 7 shows the corresponding depth according to the linear increase of PCD.

V. GAZE-BASED 3D INTERACTION

We apply our 3D gaze estimation algorithm to the gaze-based 3D interaction on the parallax barrier stereo display. We developed 3D contents with OpenGL Performer while the contents are displayed on the parallax barrier stereo display.

The 3D contents are composed with dart and arrows, in which the dart is located far and the arrows are located in near in 3D

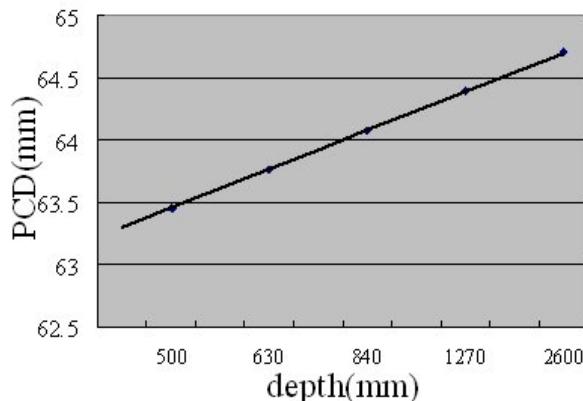


Fig. 7. Relationship between depth and PCD (Simulation)

space while each depth of the arrow is different. The distance between user and the parallax barrier stereo display is 840mm.

Fig. 8 shows our demo scenario for gaze-based 3D interaction. There are three steps, first select the arrow with the 3D gaze, second pointing a target point on dart with the 3D gaze, then the arrow will be drawn to the target point on the dart.

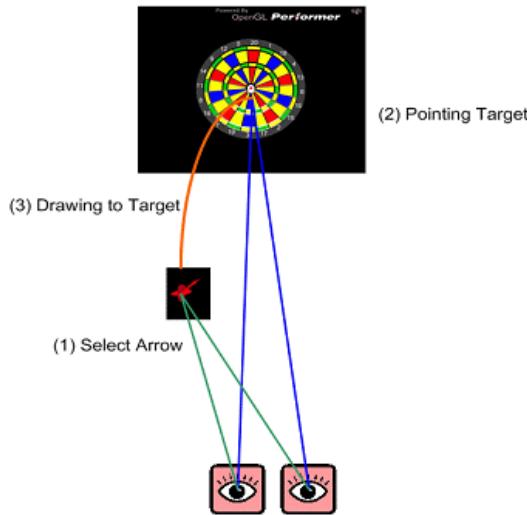


Fig. 8 Scenario of 3D gaze interaction to stereo contents



Fig. 9. Implemented demo system. See Color Plate 14.

Fig. 9 shows a screenshot of the demonstration of 3D gaze interaction using our 3D dart-arrow stereo contents.

Fig. 10 shows the measured experimental data of c, h, and PCD values from two subjects for our 3D gaze-interaction

demo. In the Fig. 10, the value of Gaze X and Y are measured c and h values at the given depth, respectively.

	Subject 1	Subject 2	Index	Depth (mm)	PCD (pixel)
Gaze X (840)	3~23	3~23	1	500	750 ~ 752
Gaze X (2600)	9~15	9~16	2	630	753 ~ 755
Gaze Y (840)	8~22	8~23	3	840	756 ~ 758
Gaze Y (2600)	14~18	13~18	4	1270	759
PCD (840)	756~758	763~764	5	2600	762
PCD (2600)	762	767			

Index	Depth (mm)	PCD (pixel)
1	500	758 ~ 750
2	630	761 ~ 762
3	840	763 ~ 764
4	1270	765
5	2600	767

Fig. 10. Experimental Data of Gaze Direction and Gaze Depth

VI. EVALUATION

A. Gaze Direction

The gaze point accuracy of the proposed system is evaluated by the experiment of looking at sixteen fixed points on the screen for evaluation.

Four subjects were asked to stare at sixteen points, three times in order. Fig. 11 shows the average error for each target of four users, while evaluating x and y direction. There are no resultant error larger than 0.60 degree and the result is almost uniformly distributed around 0.40 ~ 0.50 degree

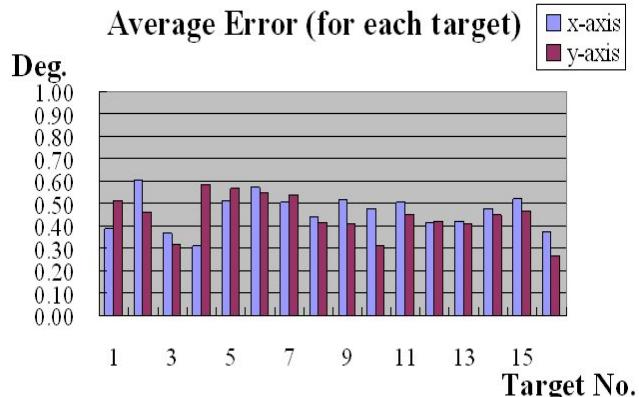


Fig. 11. Accuracy evaluation of our 2D gaze estimation algorithm

B. 3D gaze direction and depth

We also evaluate the 3D gaze using 3D gaze direction and gaze depth in our 3D dart-arrow demo system. The 3D stereo image space is divided by the five depths from 500mm to 2600 mm and 4x3 regions in each depth plane. The arrow is displayed at a position in the 3D image space randomly and then the 3D gaze accuracy is evaluated whether user can select the displayed arrow with the estimated c, h and PCD values in our 3D gaze system. The results are shown in Table 1. It should be mentioned that we evaluate the accuracy under the condition our system can get the correct 3D gaze information within 1

second. As shown in Table 1, we have at least 93 % accuracy under the condition within 1 second.

TABLE 1: ACCURACY EVALUATION OF 3D GAZE ESTIMATION

Subject	accuracy
Subject 1	97% (29/30)
Subject 2	93% (28/30)
Subject 3	93% (28/30)

VII. CONCLUSION

This paper addresses the gaze-based 3D interaction techniques to 3D contents which are displayed on the parallax barrier stereo display. This paper presents 3D gaze estimation algorithm. We also show the 3D gaze interaction technique to 3D contents on the stereo display. Our current research result shows that the proposed 3D gaze estimation technique can be used as new interaction schemes for stereo display, 3D game, and virtual reality and so on. Especially, our research result can be also used to record and analyze 3D gaze movement in view of human factor study in 3D display and VR research areas.

REFERENCES

- [1] T. Ohno, N. Mukawa and A. Yoshiokawa. FREEGAZE: a gaze tracking system for everyday dazes interaction, *Proceedings of the Symposium on ETRA 2002: Eye Tracking Research and Applications Symposium*, pp. 125-132, 2002.
- [2] Rhys Newman and Alexander Zelinsky. Error analysis of head poses and gaze direction from stereo vision, *Proceedings of the Australian Conf. on Robotics and Automation*, pp. 114-118, Mar. 1999.
- [3] <http://www.seeingmachines.com/facelab.htm>
- [4] S.W. Shih et al. A novel Approach to 3-D Gaze Tracking sing Stereo Cameras, *IEEE Trans. on Systems, Man, and Cybernetics part B: Cybernetics*, vol. 34, no. 1, Feb, 2004.
- [5] Shelly L. Ames and Neville A. McBrien. Development of a miniaturized system for monitoring vergence during viewing of stereoscopic imagery using a head mounted display, *Proceedings Of SPIE IS&T Electronics Imaging*, SPIE vol. 5291, pp. 25-35, 2004.
- [6] Kai Essig, Marc Pomplun and Helge Ritter. Application of a novel neural approach to 3D gaze tracking: vergence eye-movements in autostereograms, *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago, USA, Aug. 2004.
- [7] Kai Essig, Marc Pomplun and Helge Ritter. A neural network for 3D gaze recording with binocular eye trackers, *Int'l Journal of Parallel, Emergent and Distributed Systems*, vol. 21, no. 2, pp. 79-95, Apr. 2006.



Kyeong-Won Jeon received BS degree in Electronic Engineering from Korea University, Korea in 2006. Currently he is a research assistant at Imaging Media Research Center (IMRC), Korea Institute of Science and Technology (KIST) and also in the master's course at Korea University. His research mainly focuses on 3D gaze tracking, 3D modeling of indoor space and interaction technique.



Jeongseok Ki received BS degree in Avionics engineering from Hankuk Aviation University, Korea in 2003. Currently he is a research assistant at Imaging Media Research Center (IMRC), Korea Institute of Science and Technology (KIST) and also in the master's course at Yonsei University. His research mainly focuses on 3D gaze tracking, 3D modeling of indoor space and interaction technique.



Qonita M. Shahab received BS degree in Computer Science from University of Indonesia in 2002, and an MS in Multimedia Processing from Information and Communications University, Korea in 2005. She has been working as a lecturer at department of Computer Science, University of Indonesia. Currently she is a visiting researcher at Imaging Media Research Center (IMRC), Korea Institute of Science and Technology (KIST). Her research mainly focuses on interaction systems and multimedia communications



Sangwoo Jo received BS degree in Electronic Engineering from Sogang University, Korea in 2005. Currently he is a research assistant at Imaging Media Research Center (IMRC), Korea Institute of Science and Technology (KIST) and also in the master's course at Korea University. His research mainly focuses on 3D gaze tracking, 3D modeling of indoor space and interaction techqnue.



Sung-Kyu Kim received the BS, MS, and PhD degrees from the Quantum Optics Group of Physics, Korea University, in 1989, 1991, and 2000, respectively. Then he spent two years as an invited research scientist at the 3-D TV Group of Telecommunications Advancement Organization in Japan. In 2001 he was appointed senior research scientist at the Korea Institute of Science and Technology. His research interests include optical design of 3-D display systems, super multi-view display, multi-view image processing, digital holography, holographic optical elements, and multi-focus 3D display. He is a member of the Physical Society of Korea, the Optical Society of Korea and SPIE



Yong-Moo Kwon is a principal researcher of Imaging Media Research Center at KIST. He received Ph.D. in Dept. of Electronic Eng. from Hanyang University, Korea. He had researches in the area of image processing technology, multimedia databases, virtual heritage technology and 3D media technology. Currently, he is now investigating the following three projects entitled with "Gaze-based Interaction Technology," "Virtual Ubiquitous Robotic Space based on Ubiquitous Sensor Network (USN)," and "Interactive Authoring of Intelligent Tangible Space."