35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

3D Gaze Tracking System for NVidia 3D Vision®

Sunu Wibirama* and Kazuhiko Hamamoto, *Member*, *IEEE*

Abstract— **Inappropriate parallax setting in stereoscopic content generally causes visual fatigue and visual discomfort. To optimize three dimensional (3D) effects in stereoscopic content by taking into account health issue, understanding how user gazes at 3D direction in virtual space is currently an important research topic. In this paper, we report the study of developing a novel 3D gaze tracking system for Nvidia 3D Vision® to be used in desktop stereoscopic display. We suggest an optimized geometric method to accurately measure the position of virtual 3D object. Our experimental result shows that the proposed system achieved better accuracy compared to conventional geometric method by average errors 0.83 cm, 0.87 cm, and 1.06 cm in X, Y, and Z dimensions, respectively.**

I. Introduction

Nvidia 3D Vision® (N3DV) is one of well-known active three dimensional (3D) technology which is commonly used for visualization, assistive technology, and therapeutic device. There are so many medical applications that utilize N3DV, including augmented-reality for surgery [1], interactive medical visualization [2], navigation for visually impaired people [3], medical image analysis [4], and visual rehabilitation device for children with amblyopia [5].

As the usage of active 3D technology is increasing rapidly, concerns for 3D technology safety have also increased [6]. Visual fatigue and visual discomfort are two important health issues related with research in stereoscopic display technology [7]. One major issue that is under extensive investigation is inappropriate parallax setting in stereoscopic content. Thus, to optimize 3D effect in stereoscopic content by taking into account health issue, it is important to measure depth of virtual 3D object in stereoscopic display. Gaze tracking can be used to objectively measure the position of virtual 3D object.

There are two approaches for measuring 3D gaze in virtual space: neural network and geometric method. Essig *et al.* [8] proposed the usage of parameterized self-organizing map (PSOM) to heuristically translate 2D gaze on screen to 3D position in virtual space using 27 calibration points. Lee *et al.* [9] suggested the usage of multi-layered perceptron (MLP) to measure gaze in Z-direction based on dynamic characteristic of pupil accommodation using 45 calibration points. Although the previously proposed neural network methods yield rigorous results, the user normally has to gaze to more than three calibration points in 3D gaze calibration session to train the algorithm, which is time consuming and exhausting.

*Corresponding author
Sunu Wibirama and Kazuhiko Hamamoto are with Graduate School of Science and Technology, Tokai University, Takanawa, Tokyo, Japan (e-mail:sunu@jteti.gadjahmada.edu, hama@keyaki.cc.u-tokai.ac.jp).

The second approach is geometric approach. This method measures 3D gaze position by using virtual rays extending from the eyes to its fixation on a 2D plane [10-11]. The 3D gaze can be computed by intersecting virtual rays in 3D space. Although this method does not implement exhausting 3D gaze calibration session, the 3D gaze measurement result is not accurate since intersection of virtual rays in 3D space is unlikely. Furthermore, human dependent parameters, such as interpupillary distance and eye diameter, are assumed same for all users.

In case the user has to wear N3DV glasses, an appropriate method should be developed. The N3DV glasses should be properly installed and used simultaneously with the gaze tracking system. The eye tracker camera should be properly aligned to reduce inaccuracy in 3D gaze measurement [12]. The system should provide simpler 3D gaze calibration technique to avoid user from fatigue.

In this paper, we suggest a novel approach to measure 3D gaze while the user seeing virtual object using N3DV glasses. We have developed a head-mounted gaze tracking goggle that can be used properly with N3DV glasses. We also optimize the conventional geometric method by providing user dependent parameters using computer vision technique and incorporating 3D gaze calibration session using only three calibration points. We performed experimental validation to confirm accuracy of our 3D gaze tracking system.

II. Materials and Methods

A. Fundamental procedures

Fig. 1(a) shows the proposed 3D gaze tracking system. The eye is assumed as perfect sphere with known radius performing pure rotations around the center of the eye. The geometric model uses user coordinate system. The origin of user coordinate system is located between left and right eyes. The left and right eyeballs are separated by interpupillary distance. The head of the user is stabilized using chinrest, so that there is no head movement during 3D gaze measurement process.

Before 3D gaze measurement is conducted, several initial data are measured: E_{height} (height of eye measured from table), E_{ipd} (interpupillary distance), E_{rad} (radius of eyeball), $D_{\text{screen-eye}}$ (distance between screen and eye), S_{height} (height of screen measured from table), S_{width} (width of screen). E_{ipd} and E_{rad} are digitally measured by computing Euclidean distance in pixel unit from left and right eye cameras. The resulted E_{ipd} and E_{rad} in pixel unit are then converted to cm unit by using *pixel-to-cm* conversion value stored in the gaze tracking system's database.

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

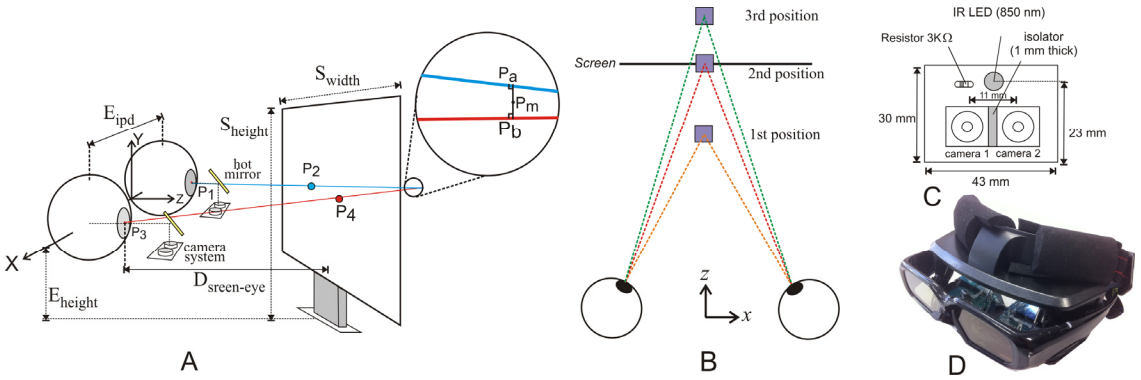


Figure 1. The proposed 3D gaze tracking system: (a) Schematic drawing of 3D gaze measurement; (b) 3D calibration to improve accuracy in *Z* direction is conducted by mapping the computed *Z* element of 3D gaze and *Z* position of calibration point. In this case, three calibration points are positioned in the front of screen, right on the screen, and behind the screen. (c) configuration of camera system; (d) N3DV glasses is installed on our head-mounted goggle

To measure 3D gaze in virtual space, there are four main steps:

Step 1: mapping gaze to screen with 2D calibration

The user is asked to gaze at several 2D calibration points. Given n calibration points, second order polynomial calibration function is used to map 2D pupil coordinate (X, Y) and screen coordinates (X_S, Y_S):

$$X_S = f(X,Y) = a_0 + a_1X + a_2Y + a_3XY + a_4X^2 + a_5Y^2 \quad (1)$$

$$Y_S = g(X,Y) = b_0 + b_1X + b_2Y + b_3XY + b_4X^2 + b_5Y^2 \quad (2)$$

The left and right gazes on screen after 2D calibration are denoted by (X_L, Y_L) and (X_R, Y_R), respective to the top left corner of the screen.

Step 2: extracting coordinate of binocular line-of-sight

P_1 and P_3 are 3D coordinates of pupil center of left and right eyes, respectively. Given E_{ipd} and E_{rad}, the 3D coordinate of pupil respective to user coordinate system is computed by using a computer vision technique named Direct Linear Transformation (DLT), as used in our previous research [13-14]. While seeing virtual 3D object, the gaze is mapped on the screen as 3D point P_2 and P_4:

$$P_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} X_L - \frac{S_{width}}{2} \\ S_{height} - E_{height} - Y_L \\ D_{screen-eye} + E_{radius} \end{bmatrix} \quad (3)$$

$$P_4 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} X_R - \frac{S_{width}}{2} \\ S_{height} - E_{height} - Y_R \\ D_{screen-eye} + E_{radius} \end{bmatrix} \quad (4)$$

Step 3: computing intersection of line-of-sight in 3D space

Exact line-of-sight intersection is unlikely, thus the nearest point of two line-of-sights is computed. Given two line-of-sights formed by P_1P_2 and P_3P_4, joined by shortest line P_aP_b, a fixation point is found by finding the midpoint P_m of P_aP_b as shown in Fig. 1(a). Point P_a at line P_1P_2 and point P_b at line P_3P_4 are given by equations:

$$P_a = P_1 + \mu(P_2 - P_1) \quad (5)$$

$$P_b = P_3 + \eta(P_4 - P_3) \quad (6)$$

The shortest line between two crossing lines is found by minimizing $|P_b - P_a|$:

$$P_b - P_a = P_3 - P_1 + \eta(P_4 - P_3) - \mu(P_2 - P_1) \quad (7)$$

Since P_aP_b is perpendicular to line P_1P_2 and P_3P_4, the result of dot product operation between them are zero:

$$(P_b - P_a).(P_2 - P_1) = 0 \quad (8)$$

$$(P_b - P_a).(P_4 - P_3) = 0 \quad (9)$$

Substituting (7) to (8) and (9), resulting:

$$[P_3 - P_1 + \eta(P_4 - P_3) - \mu(P_2 - P_1)](P_2 - P_1) = 0 \quad (10)$$

$$[P_3 - P_1 + \eta(P_4 - P_3) - \mu(P_2 - P_1)](P_4 - P_3) = 0 \quad (11)$$

Computing (10) and (11) using known x, y, and z values of P_1P_2 and P_3P_4, we can compute μ, η, P_a, and P_b. The middle point of P_aP_b, is computed as follows:

$$P_m = \frac{P_a + P_b}{2} \quad (12)$$

Step 4: 3D gaze estimation with 3D gaze calibration

Since $P_m(x,y,z)$ is the nearest point of two line-of-sights intersection, to obtain precise depth measurement, the accuracy of Z element is improved by 3D gaze calibration. Three calibration points positioned at difference depth are used, as shown in Fig. 1(b). If s_z is depth of the reference points and z is the Z element of P_m, polynomial mapping is described as follows:

$$s_z = a_0 + a_1z + a_2z^2 \quad (13)$$

or expressed as matrices:

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \quad (14)$$

$$b = A\ x \quad (15)$$

To solve vector x in (15), Singular Value Decomposition (SVD) is used. (13) is used to correct the Z element of 3D gaze point P_m. Our previous simulation confirmed the effectiveness of 3D gaze calibration [15].

B. Hardware configuration

The N3DV glasses are designed to be used appropriately with prescription glasses. We consider the size of the gaze tracking goggle to not exceed the maximum width range of N3DV glasses. For such purpose, a modified Famicom gaming goggle (Nintendo Co. Ltd., Kyoto, Japan) is used in this research. A "Y"-shaped shutter glasses holder was designed and added at the bottom part of the goggle. The goggle consists of two infra-red mirrors and two camera system for both left and right eyes. The camera system consists of two mini CCD cameras (Analog Technologies, Inc., Santa Clara, United States) (Fig. 1(c)). The camera size is 10.7 x10.7 mm. The focal length of the lens is 3.9 mm.

Fig. 1(d) shows how the N3DV glasses are properly installed on our proposed head-mounted goggle. The goggle is used to capture real-time images sequence of the eye using a resolution of 320x240 pixels with 15 Hz sampling rate. The gaze tracking system is developed by using Microsoft Visual C++ 2010. We use personal computer with Intel® i7-2600 3.4GHz processor, 4GB memory, and Windows XP® operating system.

C. Design of experiment

To confirm our gaze tracking accuracy, 9 participants (7 male, 2 female) with normal or corrected eyes were positioned 63 cm from the monitor. The average age of participants was 22.7 years old. An I-O Data LCD-3D231XBR-S 23 inches monitor with 1920x1080 pixels resolution was used in this experiment. The participants were positioned such that the center of the screen was vertically and horizontally aligned at the middle point of both eyes.

After setting up the gaze tracking hardware, a fast 2D gaze calibration procedure was performed. In this calibration session, the participants were asked to gaze to 9 reference points on 27x27 cm sized square. Following the 2D gaze calibration, 3D gaze calibration was performed. A 3.6x3.6 cm 3D rotating square was used as reference target [16]. The 3D cube was positioned vertically and horizontally at the center of the screen. The *Z* distances of the 3D cube were set to 4 cm in front of, 0 cm from, and 4 cm behind the screen. The participants were asked to gaze to the target and confirmed the 3D gaze position on the target by pressing space bar.

Afterwards, 3D gaze validation session using same rotating 3D cube was carried out. We provided four virtual planes as shown in Fig. 2. The planes were positioned at 11 and 3.67 cm in front of, and 3.67 and 11 cm behind the screen as suggested in [8]. The sequence of the virtual planes was first, second, third, and fourth as seen from the subject point of view with the first plane as the closest plane from the user.

Sixteen targets on each of four virtual planes were presented to the participants to observe the accuracy of 3D gaze-point measurement. The positions of these 64 points were not identical to any of 9 points in 2D gaze calibration or 3 points in 3D gaze calibration.

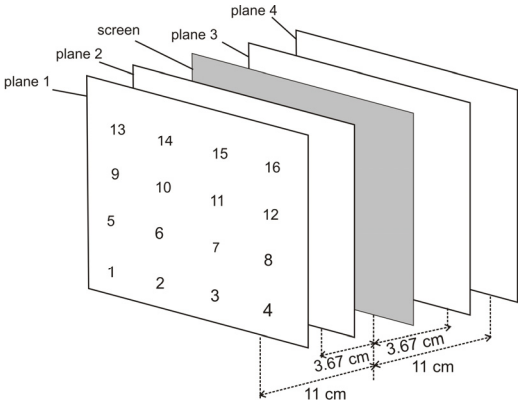
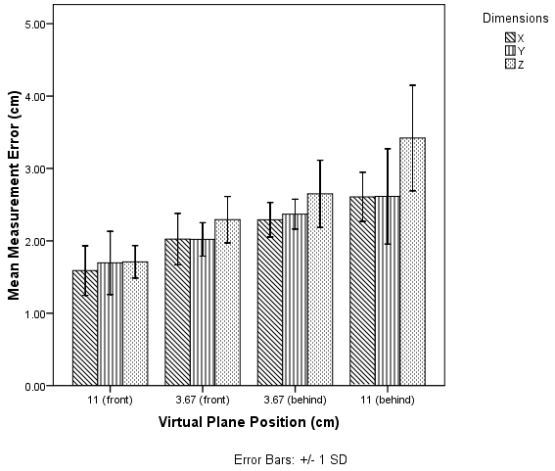
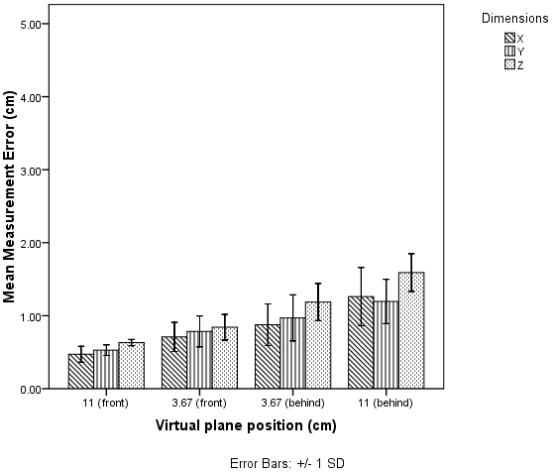


Figure 2. Four virtual planes used in experimental validation



(a) Conventional geometric method



(b) Optimized geometric method

Figure 3. Comparison of average error in X, Y, Z dimensions between conventional and optimized geometric method on four virtual planes

TABLE I. AVERAGE ERRORS FOR INDIVIDUAL COORDINATES

Methods	Measurement Error (cm)		
	X	***Y***	***Z***
Conventional Geometry	2.13 ± 0.48	2.17 ± 0.53	2.52 ± 0.76
Optimized Geometry	0.83 ± 0.38	0.87 ± 0.34	1.06 ± 0.41

To compare the accuracy of our system, we also recorded the result of conventional geometric method simultaneously in 3D gaze validation session. In conventional geometric method, the interpupillary distance and the eye radius was defined using standard parameters: 6.5 cm [17] and 1.35 cm [18], respectively.

III. Result and discussion

The result of 3D gaze validation is shown in Table 1. The average errors of our proposed method in X, Y, and Z dimensions are 0.83 (S.D=0.38), 0.87 (S.D=0.34), and 1.06 (S.D=0.41) cm, respectively. Average errors of Z position, for conventional and optimized geometric methods, are higher than errors of X and Y positions because Z position relies on accuracy of 2D gaze position. It can be seen that our proposed algorithm is superior compared to conventional geometric method. Conventional geometric method is designed to fit average user dependent parameters, while we provide customized user dependent parameters and 3D calibration session that lead to improvement of tracking accuracy.

Fig. 3 shows performance comparison of conventional and optimized geometric methods on four virtual planes. For both methods, the accuracy of 3D gaze point calculation decreases as the distance of the virtual 3D object increases. The explanation is related with the magnitude of vergence angle: if 3D object is positioned closer to the user, the change of vergence angle to adjust binocular gaze is greater, compared to further 3D object. It is said that vergence angle is effective cues of 3D gaze within 1m depth range [19]. This result also confirms finding from previous work [16] that suggests closer virtual object placement from the user to obtain effective distance perception.

We analyzed effect of two independent variables: methods and virtual planes, and one dependent variable: measurement error, by using two-way ANOVA. We identified significance of the main effects of the method ($F(1,8)=235.52$, $p<0.005$) and virtual planes ($F(3,24)=54.58$, $p<0.005$). Interaction between methods and virtual planes also shows significant result ($F(3,24)=5.28$, $p<0.01$).

IV. Conclusion and future works

We demonstrate a new approach of 3D gaze tracking system for Nvidia 3D Vision®. The gaze tracking goggle is designed by considering physical characteristics of Nvidia 3D Vision® glasses. The algorithm is designed to optimize the 3D gaze measurement by including human dependent parameters and 3D gaze calibration with only three calibration points. Our proposed method shows better performance compared to conventional geometric method with 0.92 (S.D=0.38) cm of total average errors in X, Y, and Z dimensions. In the future, we will use our system as an objective measurement device for accompanying subjective measurement of human distance perception in virtual space.

References

[1] H. Liao, "3D Medical Imaging and Augmented Reality for Image-Guided Surgery," in *Handbook of Augmented Reality*, B. Furht, Ed.: Springer New York, 2011, pp. 589-602.

[2] D. Meziat, M. H. Van Horn, S. Weeks, and E. Bullitt, "3D stereo interactive medical visualization," *IEEE Computer Graphics and Applications*, vol. 25, 2005, pp. 67-71.

[3] S. Vitek, M. Klima, L. Husnik, and D. Spirk, "New possibilities for blind people navigation," in *International Conference on Applied Electronics*, 2011, pp. 1-4.

[4] S. Muraki and Y. Kita, "A survey of medical applications of 3D image analysis and computer graphics," *Systems and Computers in Japan*, vol. 37, 2006, pp. 13-46.

[5] A. Gargantini, M. Bana, and F. Fabiani, "Using 3D for rebalancing the visual system of amblyopic children," in *International Conference on Virtual Rehabilitation* 2011, 2011, pp. 1-7.

[6] 3DConsortium. (2010, 5 September). 3DC Safety Guidelines for Dissemination of Human-friendly 3D. Available: http://3dc.gr.jp/english/scmt_wg_rep/3dc_guideE_20111031.pdf

[7] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: a review," *Journal of Imaging Science*, vol. 53, 2009, pp. 30201-(14).

[8] Kai Essig, M. Pomplun, and H. Ritter, "A neural network for 3D gaze recording with binocular eye trackers," *The International Journal of Parallel, Emergent and Distributed Systems*, vol. 21, 2006, pp.79-95.

[9] J. W. Lee, C. W. Cho, K. Y. Shin, E. C. Lee, and K. R. Park, "3D gaze tracking method using Purkinje images on eye optical model and pupil," *Optics and Lasers in Engineering*, vol. 50, 2012, pp.736-751.

[10] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 2nd ed.: Springer New York, 2007.

[11] B. C. Daugherty, A. T. Duchowski, D. H. House, and C. Ramasamy, "Measuring vergence over stereoscopic video with a remote eye tracker," in *The 2010 Symposium on Eye-Tracking Research and Applications*, Austin, Texas, 2010.

[12] S. Wibirama and K. Hamamoto, "A Geometric Model for Measuring Depth Perception in Immersive Virtual Environment," in *The 10th Asia Pacific Conference on Computer Human Interaction*, Matsue, Japan, 2012, pp. 325-330.

[13] S. Wibirama, "Real-Time 3D Eye Movements Tracking and Visualization using Dual Cameras Acquisition," *Master Thesis,* Department of Electronics, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, 2010.

[14] S. Wibirama, S. Tungjitkusolmun, and C. Pintavirooj, "Dual-Camera Acquisition for Accurate Measurement of Three-Dimensional Eye Movements," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 8(3), 2013, pp.238-246.

[15] S. Wibirama and K. Hamamoto, "Error Correction in Geometric Method of 3D Gaze Measurement using Singular Value Decomposition," in *The 5th Indonesia Japan Joint Scientific Symposium*, Chiba, Japan, 2012, pp. 554-558.

[16] M. Poyade, A. Reyes-Lecuona, and R. Viciana-Abad, "Influence of Binocular Disparity in Depth Perception Mechanisms in Virtual Environments," *New Trends on Human–Computer Interaction*, J. A. Macías, et al., Eds.: Springer London, 2009, pp. 13-22.

[17] C. Jinjakam and K. Hamamoto, "Study on Parallax Affect on Simulator Sickness in One-screen and Three-screen Immersive Virtual Environment", *Proceedings of The School of Information and Telecommunication Engineering Tokai University*, 2011, pp. 34-39.

[18] R. S. Park and G. E. Park, "The center of ocular rotation in the horizontal plane," *American Journal of Physiology--Legacy Content*, vol. 104, 1933, pp. 545-552.

[19] I. P. Howard and B. J. Rogers, "Geometry of Binocular Images," in *Seeing in Depth: Depth Perception*. vol. 2, Toronto: I Porteous, 2002, pp. 1-9.83