ERROR CORRECTION IN GEOMETRIC METHOD OF 3D GAZE MEASUREMENT USING SINGULAR VALUE DECOMPOSITION

Sunu Wibirama*1,2 and Kazuhiko Hamamoto1

1Graduate School of Science and Technology, Tokai University, Tokyo, Japan
2Department of Electrical Engineering and Information Technology, Faculty of Engineering, Gadjah Mada University, Yogyakarta; E-mail: sunu@jeti.gadjahmada.edu

KEYWORDS: 3D Gaze Tracking, Geometric Method, Singular Value Decomposition

ABSTRACT: Utilizing gaze tracker as a tool for human-computer interaction or visual perception understanding has become increasingly popular. Gaze tracking is to compute the position of user’s gaze. Most researcher conduct gaze tracking studies on two-dimensional (2D) plane, such that X and Y position of the user gaze at the monitor is known. However, as the usage of stereoscopic display increases, the importance of understanding three-dimensional (3D) gaze of not only X, Y position, but also Z position, obtains bigger attention, particularly to understand how the user perceive depth of virtual object. One of commonly used method in 3D gaze tracking is by computing geometric relationship between eye position and screen position. In this paper, we propose a new approach to reduce error of 3D gaze measurement using Singular Value Decomposition (SVD). The error reduction method was simulated and the result is compared with the conventional geometric method. The result shows that the error correction approach is considered appropriate to be included in 3D gaze measurement as it reduces error more than 80%.

1. INTRODUCTION

Gaze tracking is estimating the position of user’s gaze through spatial position of pupil. Most gaze tracking research is conducted on 2D plane such that 2D coordinates of the gaze at the monitor are known. On the other side, immersive virtual environment (IVE) has been widely used to simulate real or imagined scenario, such as design and training processes in industrial and collaborative tasks (Kan, 2001; Monahan, 2008; Robert, 2001). The usage of 2D gaze tracking as interactive interface and visual search tools in IVE systems has been widely known (Hardiess, 2008; Lee, 2007; Murray, 2009; Pfeiffer, 2008).

To make the users feel more immersed, interested, and comfortable within these IVE systems, optimization of IVE content is needed. In IVE optimization, studying user’s depth perception measurement is very important to obtain and correct any discrepancy between the measured value and that specified in VE content (Nichols, 2000). In IVE system, binocular disparity is used to generate virtual 3D object, as shown in Figure 1(a). In practical implementation, the parallax angle ($\theta - \alpha$) of the user, which is mainly a trigger for virtual sickness, should be kept
between 0.5°-1° (3Dconsortium, 2010). Such study requires a gaze tracking system which is not only able to measure gaze in 2D plane (X and Y directions), but also in 3D direction (the Z-direction).

In our previous research, 3D eye and gaze positions measurement model has been proposed (Wibirama, 2012a; 2012b). We used gaze-to-screen mapping information to estimate user’s 3D gaze. 3D coordinate of pupil center is obtained by transforming and rotating pupil center with bottom left of screen as origin of world coordinate reference. A gaze depth fixation point is then obtained by crossing 3D line of sight from both eyes. However, 3D gaze measurement error in this geometric model is quite large such that the method can only estimate approximately 88.56% from the intended distance. In this paper, we propose an error correction method for this geometric model by using Singular Value Decomposition (SVD) to solve 2nd degree polynomial mapping in 3D direction. The proposed error correction was then simulated. The result shows that the error correction significantly increases the accuracy of 3D gaze measurement model.

2. METHOD

![Figure 2](attachment:image.png)

Figure 2. (a) Schematic drawing of eye coordinate system which relates to screen coordinate system  
(b) Computing intersection of two lines in 3D spaces

We use binocular vision to measure depth of virtual 3D object from screen as shown in Figure 1(b). The eye is assumed as a perfect sphere with known radius performing pure rotations around the center of the eye as shown in Figure 2(a) (Moore, 1996). There is also assumed that the subject’s head is fixed so that no head movement is produced during depth perception measurement. The eye coordinate system is based on modified right-handed Fick-order (Haslwanter, 1995): ye-axis (horizontal rotation), xe-axis (vertical rotation), and ze-axis (torsional rotation). The optical axis of the eye is initially assumed coinciding with the ze-axis. P1 and P3 are 3D coordinate of pupil center of left and right eye related to each eye coordinate system, respectively. The radius of the eyeball is 13.5mm as used in conventional gaze tracking research (Park, 1933). The distance between two eyeballs is assumed 6.5cm as used in (Jinjakam, 2011).

The origin of the world coordinate is located at the bottom left of the screen. While seeing the virtual 3D object at front or behind the screen, the eye gaze will be mapped on the screen as 3D point P2 and P4 for left and right eye, respectively. P2 and P4 are related to world coordinate system. The eye camera is a pinhole camera without any lens distortion. It is initially located at the front of the eyeball with 6cm distance from the center of pupil to center of camera coordinate. The optical axis of the camera coincides with optical axis of the eye. The focal length of the camera is 0.015pixel with resolution 1280 x 1024 pixels.

In this research, we use HoloStage® Christie Digital System configuration as used in Tokai University, Takanawa campus, Japan (Hewlett Packard, 2012). The screen size is 5.4m x 3m. The subject’s eye height is assumed 1.7m as in (Ruddle, 1997). The subject is positioned at 3m distance from the screen. To compute the gaze-to-screen mapping, homography method is used (Wribarama, 2012a). To compute gaze in Z-position, we use line-crossing method. Given two line segments formed by pupil and gaze position P1P2 and P3P4, joined by shortest line PaPb, a fixation point can be find by finding the midpoint Pm of PaPb as shown in Figure 2(b). Point Pa on the line P1P2 and point Pb on the line P3P4 are given by equations:

\[
P_a = P_1 + \mu( P_2 - P_1 )
\]

\[
P_b = P_3 + \eta( P_4 - P_3 )
\]

The 5th Indonesia-Japan Joint Scientific Symposium  
Poster  
IJJSS, 25-26 October 2012

555 Copyright IJJSS 2012
The shortest line between two crossing lines can be found by minimizing \(|P_b - P_a|\):
\[
P_b - P_a = P_3 - P_1 + \eta (P_2 - P_1) - \mu (P_2 - P_1)
\]
(3)

Since \(P_a P_b\) is perpendicular to line \(P_1 P_2\) and \(P_3 P_4\), the result of dot product operation between them are zero:
\[
(P_b - P_a) \cdot (P_2 - P_1) = 0
\]
(4)
\[
(P_b - P_a) \cdot (P_3 - P_1) = 0
\]
(5)

Substituting Equation 3 to Equations 4 and 5, we obtain:
\[
\begin{align*}
(P_3 - P_1 + \eta (P_2 - P_1) - \mu (P_2 - P_1)) &\cdot (P_2 - P_1) = 0 \\
(P_3 - P_1 + \eta (P_2 - P_1) - \mu (P_2 - P_1)) &\cdot (P_3 - P_1) = 0
\end{align*}
\]
(6)
(7)

Computing Equations 6 and 7 using known \(x, y,\) and \(z\) values of \(P_1 P_2\) and \(P_3 P_4\), we obtain \(\mu, \eta, P_a,\) and \(P_b\). The middle point of \(P_a P_b\), which is gaze in Z-position, can be computed as follows:
\[
z = \frac{P_a + P_b}{2}
\]
(8)

To correct the error in 3D gaze measurement, we conducted a 3D calibration. The 3D calibration involves three references point located in the difference \(Z\)-position. The computed 3D gaze for each reference point is obtained from the line-crossing method. If \(s_z\) is depth of the reference point and \(z\) is the estimated \(Z\)-position calculated by line-crossing method, second order non-linear mapping can be implemented as follows
\[
s_z = a_0 + a_1 z + a_2 z^2
\]
(9)

\[
\begin{bmatrix}
s_1 \\
s_2 \\
s_3 \\
\end{bmatrix} =
\begin{bmatrix}
1 & z_1 & z_1^2 \\
1 & z_2 & z_2^2 \\
1 & z_3 & z_3^2 \\
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
a_2 \\
\end{bmatrix}
\]
(10)

\[
b = A x
\]
(11)

To solve \(x\) in Equation 11, we used SVD method as follows:
First, any \(m \times n\) matrix \(A\) can be written as the product of three matrices
\[
A_{mn} = U_{mn} D_{mn} V_{mn}^T
\]
(12)

where matrix \(U\) (size \(m \times m\)) is left orthogonal matrix, matrix \(D\) (size \(m \times n\)) is diagonal matrix, and matrix \(V\) (size \(n \times n\)) right orthogonal matrix. The elements of matrix \(D\), which is \(\sigma_i\), is called singular values, such that \(\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0\). Next, to solve \(x\), we do the following computation:
\[
c = U^T b
\]
(13)
\[
y =
\begin{bmatrix}
c(1) \\
D(1,1) \\
c(2) \\
D(2,2) \\
\vdots \\
D(m,m) \\
\end{bmatrix}
\]
(14)
\[
x = y
\]
(15)

To perform validation, we simulate our geometric model using MATLAB® version R2011b on personal computer with Intel® i7-2600 3.4GHz processor, 4GB memory, and Windows XP® Service Pack 2 operating system. In this simulation, we assume that 1 MATLAB unit equals to 0.01 meter in real environment. A small red sphere with 0.05m radius is used as virtual 3D object. The first experiment was to validate the 2D gaze. In this experiment, we positioned the eye camera such that the optical axis of the camera performed 0° angle related to the optical axis of the eye by maintaining the 6cm distance between the eyeball and the camera center. We measured 2D gaze...
mapping on the screen and 3D gaze on virtual 3D object from screen. The 3D gaze measurement was conducted at
11 depth positions: 1, 0.8, 0.6, 0.4, 0.2, 0, -0.2, -0.4, -0.6, -0.8, and -1 m from screen. “+” and “-” signs indicate front
and behind screen condition, respectively. To perform 3D gaze error correction, we used three reference points
(0.4 m, 0 m, -0.4 m) and their corresponding Z-gaze positions. We then performed SVD to find the solution of matrix
X. From experimental result, we obtained parameters of Equation 9 as follows: 
\[ x = -15.695 + 0.9755z + 0.0002z^2. \]

3. EXPERIMENTAL RESULTS

This section shows our experimental results. Figure 3(a) shows the result of 2D gaze validation on vertical and
horizontal direction. The average error on horizontal and vertical directions are 0.65° and 0.75°. Total average error
is about 0.70°. Table 1 shows average error and standard deviation of uncorrected (conventional method) and
corrected (improved method) 3D gaze measurement. Figure 3(b) shows comparison result of uncorrected and
corrected 3D gaze data for each reference depth point.

<p>| Table 1. Comparison result of uncorrected and corrected 3D gaze error measured from screen |</p>
<table>
<thead>
<tr>
<th>3D Gaze Error (in meter)</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncorrected</td>
<td>0.1501</td>
<td>0.014</td>
</tr>
<tr>
<td>Corrected</td>
<td>0.0017</td>
<td>0.0017</td>
</tr>
</tbody>
</table>

![Figure 3. (a) Result of 2D gaze validation, compared with the validation target (in degree),
(b) Result of uncorrected and corrected 3D gaze error (in meter)](image)

4. DISCUSSIONS AND CONCLUSION

Given a known radius of eyeball and distance between the screen and the subject, the proposed geometric model
was found to be adequately describing the 2D gaze position since the gaze error of our geometric method (0.70°)
was smaller than average visual angle accuracy of human (1° - 2°) [Tobii Eye Tracker, 2012]. The error correction
method of 3D gaze measurement can reduce 3D gaze error up to 98% theoretically, which is a significant
improvement for its real implementation in the future. In real situation involving human subject, it was reported that
depth in virtual environments was often only 50%-80% of the intended distance [Pollock, 2012]. Therefore, our
improved geometric method is considered appropriate to be implemented since it has better depth measurement
accuracy compared to human depth perception. In the future, we would like to implement our geometric model to
measure user’s depth perception in real IVE system.
REFERENCE