Parallax error in the monocular head-mounted eye trackers

Diako Mardanbegi
IT University of Copenhagen
Rued Langgaards Vej 7,
DK-2300 Copenhagen S
dima@itu.dk

Dan Witzner Hansen
IT University of Copenhagen
Rued Langgaards Vej 7,
DK-2300 Copenhagen S
Witzner@itu.dk

ABSTRACT
This paper investigates the parallax error, which is a common problem of many video-based monocular mobile gaze trackers. The parallax error is defined and described using the epipolar geometry in a stereo camera setup. The main parameters that change the error are introduced and it is shown how each parameter affects the error. The optimum distribution of the error (magnitude and direction) in the field of view varies for different applications. However, the results can be used for finding the optimum parameters that are needed for designing a head-mounted gaze tracker. It has been shown that the difference between the visual and optical axes does not have a significant effect on the parallax error, and the epipolar geometry can be used for describing the parallax error in the HMGT.

Author Keywords
Head-mounted gaze tracker, Parallax error, Mobile gaze tracker, epipolar geometry

ACM Classification Keywords
I.4.1 Image processing and computer vision: Digitization and Image Capture

General Terms
Measurement, Performance

INTRODUCTION
Head mounted gaze trackers (HMGT) are used for estimating the PoR in the user's field of view and are widely used for diagnostic applications. They have also been used for interaction in virtual [4, 5] or real [3, 6] environments. Head mounted gaze trackers have a scene camera for capturing the scene and another camera for capturing the eye image. HMGT is also called mobile gaze tracker because it is mounted on the user's head and can be used when the user is fully mobile. HMGT can potentially obtain a high degree of flexibility and mobility. However, most of the HMGT systems do not still allow for estimating the gaze point accurately in wide range of distances. A common problem with Head-mounted gaze trackers is that they introduce gaze estimation errors (a.k.a. *parallax error*) when the distance between the point of regard and the user (fixation distance) is different than when the system was calibrated. This error is due to the scene camera and the eye are not co-axial. Parallax error limits the use of head-mounted gaze trackers into a certain range of depth (a.k.a. *effective depth*).

There is a physical solution for removing the parallax between the scene camera and the eye. When the projection center of the scene camera coincides with the eyeball center, there is no parallax error. This can be done by using a visor (half mirror) in front of the eye and transferring the field of view of the eye to the scene camera. Head-mounted gaze trackers that do not have the scene camera mounted co-axial with the eye, require an indirect way of compensating for the parallax error. In order to be able to compensate for the parallax error, it important to know more about the error behavior and the main parameters that change the error. There are two main questions here: first, how do the scene camera orientation and position influence the parallax error? And second, with a fixed camera configuration, how does changing the calibration and fixation distances change the error? This paper investigates the answers of these two questions. The answer of the first question helps for having a better and optimum design for the head-mounted gaze trackers that have less error in gaze estimation. The answer of the second question helps for understanding the behavior of the parallax error when the fixation distance changes. It may help for estimating a function that calculates the parallax error given the fixation distance, which can be used for compensating for the error.

The remainder of this paper is organized as follows. Some related works about the parallax error in HMGT are briefly mentioned in the next section. We then introduce

UbiComp' 12, Sep 5 – Sep 8, 2012, Pittsburgh, USA.

the parallax error and describe the HMGT setup as a stereo camera setup. Then we describe in details how to calculate the parallax error in HMGT. Then we present the results of calculating the parallax error in the image and fixation planes with a summary.

PREVIOUS WORK

Velez et.al [7] at 1988 introduced a method for direct compensating for parallax error in the head-mounted eye trackers, using a transparent visor in front of the eye, which reflects the eye image towards the eye camera and the scene image towards the scene camera making a parallax free scene camera configuration. This method is a direct way for eliminating the parallax error and the eye tracker works quite accurate for different depths using only one time calibration. However, not all the head-mounted gaze trackers have such design today. Li [2], investigated the parallax error behavior in a simplified model of a HMGT, where the scene camera is mounted above the eye (only a vertical displacement). The angle between the visual and optical axis was not also considered in the analysis.

PARALLAX ERROR

The problem of parallax error can be simplified into two dimensions, which is visualized in Figure 1.

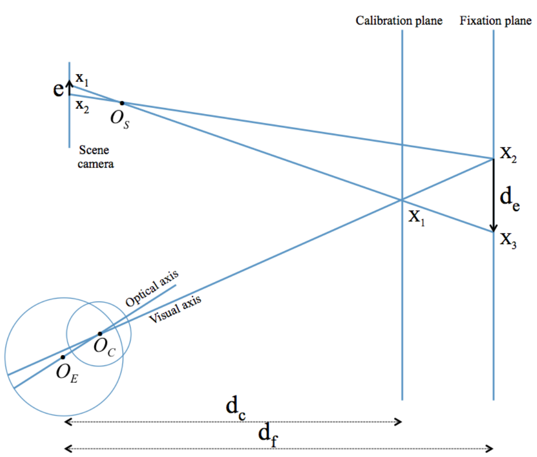

Figure 1: 2D parallax error

Suppose that the center of rotation of the eye is the point O_E, and the point O_C, is the center of the cornea where the visual and optical axes of the eye intersect. The scene camera is shown as a pinhole camera with a vertical image plane. Suppose that the system is calibrated for a plane (calibration plane) at a given distance d_c and the user fixates on a plane (fixation plane) at a further distance d_f. The visual axis of the eye intersects the calibration plane at the point X_1 and the fixation plane at the point X_2. The projections of these two points are not coincident on the image plane. When the user is looking at the point X_1 in the calibration plane, the estimated gaze point on the scene image would be the point x_1. When the user is looking at the point X_2, the visual axes and subsequently the eye image would be the same as for point X_1 and therefore the estimated gaze point would be the same point as x_1. The projection of the gaze point X_2, is the point x_2, however since the eye image has not been changed[1], the gaze tracker cannot compensate for this error. The parallax error can be defined as a vector in the scene image (x_1-x_2), which is corresponding to the vector X_3-X_2 in the fixation plane.

The relationship between the parallax error and the geometry of the system can be described in the general condition by epipolar geometry in a stereo camera system. Figure 2 shows a scene camera mounted on the head modeled as a pinhole camera with an optical center located at the point O_S and the focal length of f. The general transformation matrix $[R|t]$ represents the translation (t) and orientation (R) of the camera coordinate system relative to the fixed coordinate system. The fixed head coordinate system (X_E, Y_E, Z_E) is a right-handed 3D cartesian coordinate system located at the center of the eyeball (O_E), such that the Z_E axis is pointing forward, X_E is pointing to the left and Y_E is upward. This coordinates system, is considered as the fixed world coordinates system. Both calibration and fixation planes are assumed to be two planar surfaces in front of the head and parallel to the anatomical frontal plane of the body (X_E-Y_E plane).

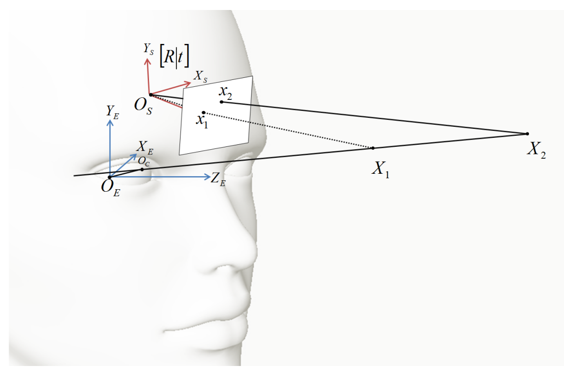

Figure 2: General configuration

Transformation from the world coordinate into the camera coordinate system can be done by the scene camera matrix which can be defined as the matrix $C=K[R|t]$ where R and t are the external parameters, and the matrix K is the internal parameters of the camera. For a normal CCD camera with the focal length of f (in meter) and the principal point at the center of the image, the matrix K can be described by:

$$K = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (1)$$

[1] The lens thickness would be different for points X_1 and X_2 as its focusing distance varies. However, it cannot be observed by a regular camera.

Assuming that the scene camera has a translation of ${}^{E}_{S}\mathrm{Tr} = (tx, ty, tz)$ relative to the fixed coordinate frame and three rotations around the fixed axes, the scene camera can be described by:

$$ {}^{E}_{S}T = \begin{bmatrix} {}^{E}_{S}\mathrm{R} & {}^{E}_{S}\mathrm{Tr} \\ 0 & 1 \end{bmatrix} \quad (2)$$

Where ${}^{E}_{S}\mathrm{R}$ and ${}^{E}_{S}\mathrm{T}$ define the camera coordinate frame (S) relative to the fixed coordinate frame (E) in the homogeneous form. The rotation matrix ${}^{E}_{S}\mathrm{R}$ is the multiplication of three rotations:

$$ {}^{E}_{S}\mathrm{R} = \mathrm{RZ}(\gamma_z)\mathrm{RY}(\gamma_y)\mathrm{RX}(\gamma_x) \quad (3)$$

where $\mathrm{RX}(\gamma_x)$ is the rotation by γ_x around the X_E axis, $\mathrm{RY}(\gamma_y)$ is the rotation by γ_y around the Y_E axis, and $\mathrm{RZ}(\gamma_z)$ is the rotation by γ_z around the Z_E axis. For simplicity, in the following of this paper the orientation of the camera is shown by the angles as $\mathrm{R}= (\gamma_x, \gamma_y, \gamma_z)$. The external parameters of the camera can be calculated by:

$$ [\mathrm{R}|\mathrm{t}] = {}^{E}_{S}T^{-1} = \begin{bmatrix} {}^{E}_{S}\mathrm{R}^{\mathrm{T}} & -{}^{E}_{S}\mathrm{R}^{\mathrm{T}}\,{}^{E}_{S}\mathrm{Tr} \\ 0 & 1 \end{bmatrix} \quad (4)$$

Knowing the camera matrix (C=K[R|t]), we can project any point in the field of view (X) into the camera image by multiplying the point by the camera matrix (x=CX). The parallax error is defined as the vector x_1-x_2 which is the projection of vector X_1-X_2. The parallax error may be different for each point in the fixation plane, and for each point the error can be considered as a function of the calibration distance (dc), geometrical parameters of the camera (R, Tr, f) and the coordinates of that point in the fixation distance (xf,yf,df).

When the fixation point (X) goes further away from the calibration plane, the projection (x) moves along a line in the scene image called epipolar line. If we assume that the point of regard is along the optical axis, then the eye and scene camera can be considered as a stereo setup. The projections of the points of an optical axis onto the scene image are all along a line called epipolar line. Changing the angle of the optical axis change the epipolar line, however, all epipolar lines intersect at a point called the epipole, which is the projection of the eyeball center into image plane. Considering the difference between the optical and visual axes and the fact that the point of regard is along the visual axis, the result would be slightly different.

In this paper, the displacement of the fovea from the optical axis is taken into account and the visual axis has been used instead of the optical axis.

CALCULATING THE PARALLAX ERROR

For calculating the parallax error in the image plane, we choose a point in the scene image (e.g. x_2), and find the visual axis passes through the point X_2 of the fixation plane. Then, the error vector can be calculated by having the projection of the point X_1, which is the intersection of the visual axis and the calibration plane. In order to find the visual axis, first the selected point x_2 on the image is back-projected on the fixation plane:

$$ X_2 = \begin{bmatrix} x_f \\ y_f \\ df \end{bmatrix} = C^{-1}x_2 \quad (5)$$

Then the visual axis can be calculated as a line that passes through the points O_C and X_2. Figure 3 shows the points X_1 and X_2 in the fixed coordinate system. The eye rotates around the center of the eyeball (O_E) and it changes the direction of the optical axis (O_E-O_C). The visual axis intersects the optical axis at the center of the cornea (O_C), which is also the nodal point of the eye. The orientation of the optical axis can be described by the horizontal (pan) angle θ and the vertical (tilt) angle ϕ. The point O_C can be described by these angles as below:

$$ O_C = d \begin{bmatrix} cos\varphi\ sin\theta \\ sin\varphi \\ cos\varphi\ cos\theta \end{bmatrix} \quad (6)$$

where the parameter d is the distance between the center of cornea and the center of eyeball (O_E).

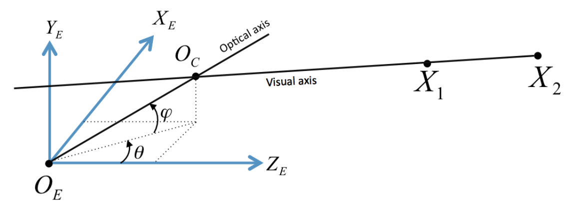

Figure 3: Showing the visual and optical axes in the fixed coordinate frame

The orientation of the visual axis can be expressed by the pan angle $\theta+\alpha$ and the tilt angle $\varphi+\beta$ where the α and β are the horizontal and vertical angles between the visual and optical axes. Therefore, any point on the visual axis can be expressed by:

$$ \mathrm{X} = O_c + k \begin{bmatrix} \cos(\varphi+\beta)\sin(\theta+\alpha) \\ sin(\varphi+\beta) \\ cos(\varphi+\beta)\ cos(\theta+\alpha) \end{bmatrix} \quad (7)$$

Where the scalar k defines the distance from the point O_C.

Given the known point X_2, the three unknown parameters (φ, θ, and k) of the equation (7) can be obtained, and by knowing these parameters, we can calculate the point X_1, which is on the calibration distance (dc). Finally, the point x_1 can be obtained by projecting the intersection of the visual axis and the calibration plane (X_1).

The parallax error can be both represented as a vector in the scene image (x_1-x_2), or as a vector in the fixation plane (X_3-X_1). The error in the fixation plane can be obtained by having the point X_1 and the point X_3 (figure 1) which can be obtained by back-projecting the point x_1 onto the fixation plane.

In the next two sections, the parallax error has been calculated for different camera positions and distances and the simulation has been performed based on the equations above.

ERROR IN THE IMAGE PLANE

In this section, we measure the parallax error for different points of the scene image and it is shown how the angle and magnitude of the error vector will be influenced by changing the calibration and fixation distances, and also the camera position and orientation. It has been observed that by considering the visual axis, the epipolar lines do not intersect in exactly one point, however, there is not a significant difference in the overall distribution of the error directions in the image. In order to provide a better understanding of distribution of the error in the scene image, the error is measured in meter in the image plane instead of visual angle. However, in the next section when the error is presented in the fixation plane, it has also been measured in visual degree. Therefore, the unit meter is used for the focal length (f) in this section, and wherever the error is measured in the image, it will be in meter. If the focal length of the camera is known in pixel (f_{Pixel}), the error can be obtained in pixel by multiplying the error value to f_{Pixel}/f.

We start with the vertical translation of the camera (*ty*), and show the error changes by changing the fixation and calibration distances. Then we investigate the other transformations of the camera. The typical values of the eye parameters (α=±5°, β=1.5°[1], d=5.3mm[9]) have been used for calculation in the following, and all the calculations are done for the right eye (α=+5°). The range of the eyeball rotation ~70°×~70°[8] has been used to define the user's field of view and the size of the fixation plane in different distances, however, this angle is not used in practice, and the actual range of the eye movements is less than 50°×50°.

Figure 4 shows the magnitude of the parallax error in the center of the scene image for three different calibration distances (dc= 1.5, 3 and 5m), and the fixation distance from 0.4m to 10m. The camera parameters are R=(0,0,0), Tr=(0,0.05m,0) and f=0.005m. The field of view of the camera is considered to be 50°.

It can be seen that the parallax error is zero when the fixated and calibrated distances are equal and then increases as they diverge. The parallax error is larger for the closer distances and rises a bit faster as the fixation distance falls behind the calibration distance.

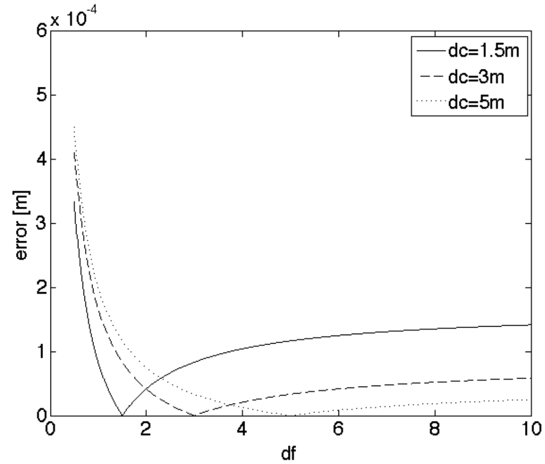

Figure 4: Changes in parallax error by changing the fixation distance

Figure 4, can give us an idea about how to choose the calibration distance when the gaze tracker is supposed to be used in a certain distances. For example for the range of 2m-10m, the calibration distance around 5m results less average error in the range of use. The error shown in the figure 4 is almost the same for all points in the scene image. The small variance has been observed for different points which is because of the angle between the visual and optical axes and is not significant. Generally, in a stereo setup, when the camera image is parallel to the fixation plane, the epipole in the image plane is in infinity and the parallax error (magnitude and direction) is the same for all the points in the image. Therefore, by translating the camera horizontally or vertically (*tx,ty*) or rotating the camera around its optical axis the parallax error would still be the same for all the points in the image and can be described by one vector.

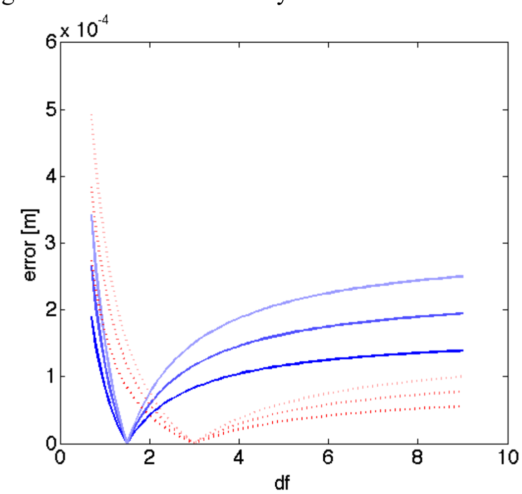

Figure 5: The effects of vertical and horizontal translations of the camera on the parallax error

Increasing the vertical distance between the camera and the eye, increase the level of the error curve shown in figure 4. Figure 5 shows these changes for two calibration distances 1.5m (blue curves) and 3m (red dotted curves).

Three different curves can be seen for each calibration distance. The curves with the lower levels are for the ty=0.05m, the curves in the middle are for the ty=0.07m and the upper curves are for ty=0.09m.

In general, the parallax error can be shown as a function of both calibration distance and fixation distance (figure 6), when the image plane in parallel to the fixation plane and the error is the same for all points in the image.

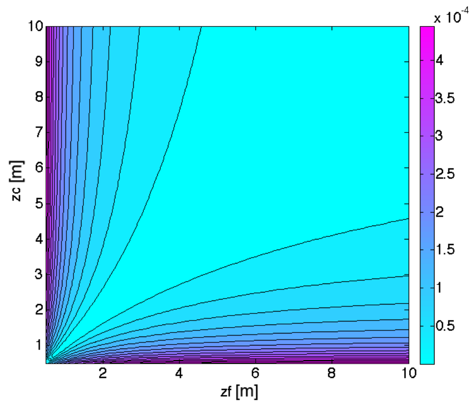

Figure 6: 2D error diagram for showing the error changes by changing the calibration and fixation distances

The results shown above for the magnitude of the parallax error, are the same when the camera translation is along the X_E axis instead of Y_E. The only difference is the changes in the direction of the error vectors. However, difference between the visual and optical axes, makes small differences in the magnitude of the error within the image, but is not significant. Figure 7, shows the vector field of the parallax error in the scene image for camera translations of Tr={(0,0.05m,0), (-0.05m,0,0), (0.05m,0,0), (0.05m,0.05m,0)} with the calibration distance of *dc*=2m and fixation distance of *df*=0.5m, without considering the difference between optical and visual axes.

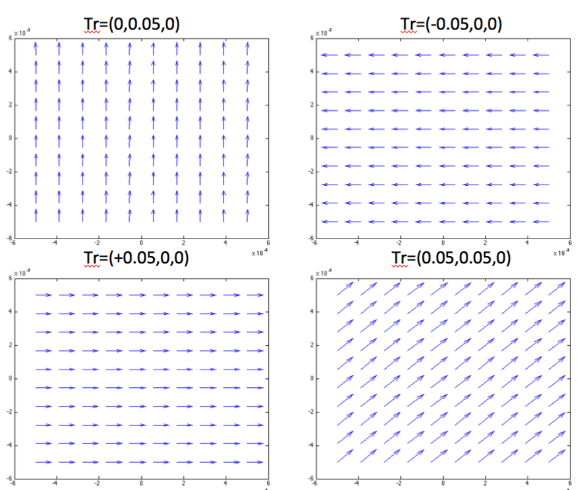

Figure 7: The vector field of the error in the scene image when the camera has the vertical and horizontal translations

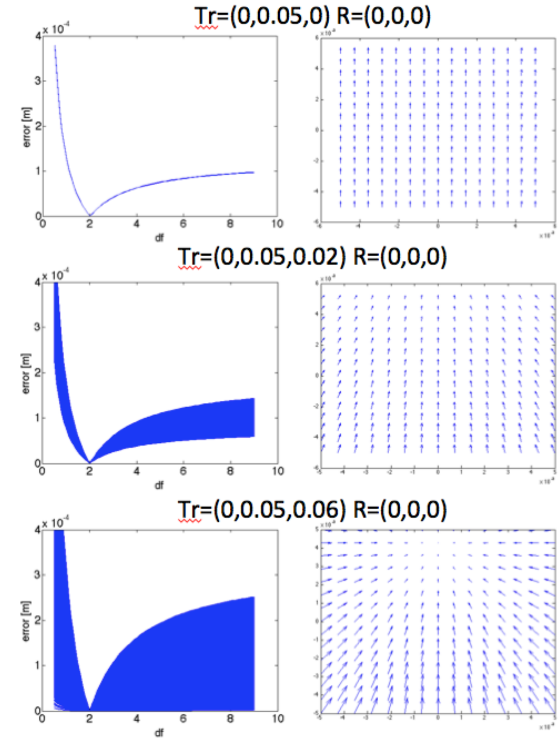

Figure 8: The effects of moving the camera along the Z-axis on the error

Translating the camera in the *Z* direction, moves the epipoles from the infinity toward the center of the image, and it changes the uniformity of the error within the image. It increases the error in some points and decreases the error in some other points. When the epipole is inside the image, the parallax error is zero for the epipole. Generally when the epipole is not at infinity, the *tx* and *ty* move the epipole horizontally and vertically respectively.

Figure 8 shows the parallax error for three different camera positions of Tr={(0,0.05m,0), (0,0.05m,0.02m), (0,0.05m,0.06m)} when the calibration distance is 2m. The graphs on the left side, show the changes in error magnitude for the different points of the image when the fixation distance is changing. It can be seen in this figure that how moving the camera in the Z direction changes the upper and lower bound of the error in the image. The vector field of the error for the fixation distance of 0.8m has been also shown in the right side.

Regarding the camera rotation, we show the effect of two important rotations pan and tilt. Usually the scene cameras do not have the roll rotation. When the roll angle is zero the pan (horizontal) rotation and tilt (vertical) rotation move the epipole horizontally and vertically respectively. It means that for example when the translation tz moves the epipole from the infinity to the center, the rotation Rx can translate it back again to the infinity. Therefore, direction of the error vectors would be the same but their magnitudes are different. Figure 9 shows the error for the camera with a vertical rotation of $\gamma_x = 20°$ and a translation Tr=(0,0.05m,0.02m). It can be compared to the second row of the figure 8.

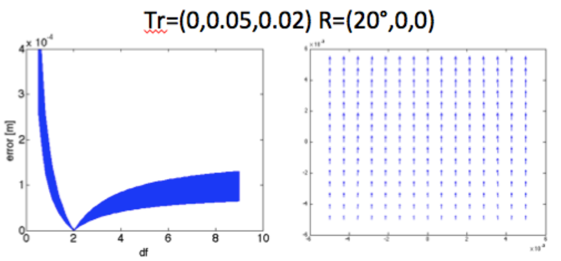

Figure 9: Moving the epipole to infinity by tilting the camera

As it can be seen in figure 9, the direction of the error vectors is uniform in the image but the range of the error size has not been changed too much after the rotation.

ERROR IN THE FIXATION PLANE

Sometimes it is useful to know the size of the corresponding error in the fixation plane. Figure 10 shows the magnitude of the parallax error in the fixation plane with the same configuration as for figure 4 and three different calibration distances (dc= 1.5, 3 and 5m). Figure 11 shows the error in visual angle for different points in the fixation plane when dc=3 and df=6.

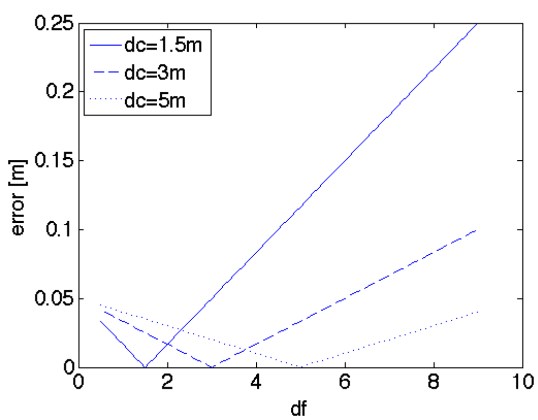

Figure 10: The actual error size in the fixation plane

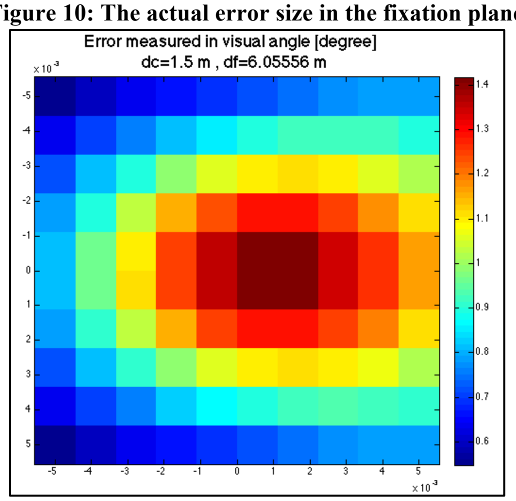

Figure 11: Error in a fixation plane in visual angle

CONCLUSION

In this paper, the parallax error in the head-mounted gaze trackers has been defined and described using the epipolar geometry in a stereo camera setup. The effect of changing the calibration and fixation distances on the parallax error has been investigated. It has been shown that the effective range of the gaze estimation with less parallax error is larger when the distance between the user and calibration plane is larger. The changes in the parallax error for different positions of the scene camera have been investigated. Camera translation and rotations relative to the eye, change the distribution of the error size and the direction of the error vectors in the image. The optimum configuration can be chosen based on the method that will be applied for compensating for the parallax error. It has also been shown that the difference between the visual and optical axes does not have a significant effect on the parallax error.

REFERENCES

1. Gale, A. G. "A note on the remote oculometer technique for recording eye movements," *Vis. Res.*, vol. 22, no. 1, pp. 201–202, 1982.
2. Li., D. Low-Cost Eye-Tracking for Human Computer Interaction. Master's thesis, Iowa State University, Ames, IA., Techreport TAMU-88-010, 2006.
3. Mardanbegi, D., and Hansen, D.W. Mobile gaze-based screen interaction in 3D environments. In Proc. Novel Gaze-Controlled Applications (NGCA '11). Blekinge Institute of Technology, Karlskrona, Sweden, 2011.
4. Nilsson, S., Gustafsson, T., Carleberg, P.: Hands Free Interaction with Virtual Information in a Real Environment. In: Proc. COGAIN 2007, Leicester, UK, pp. 53–57, 2007.
5. Park, H. M., Lee, S. H., Choi, J.S. Wearable augmented reality system using gaze interaction, Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, p.175-176, September 15-18, 2008.
6. Shi, F., Gale, A.G. & Purdy, K.J. Eye-centric ICT control. In Bust P.D. & McCabe P.T. (Eds.) Contemporary Ergonomics, 215-218, 2006.
7. Valez, J., Borah, J. D. "Visor and camera providing a parallax-free field of view image for a head-mounted eye movement measurement system." U.S. Patent 4 852 988, Aug. 1, 1989.
8. Wandell, B. A. *Foundations of vision.* Sinauer Associates Inc; USA, 1 edition, 1995.
9. Young, L. R., and Sheena, D. "Methods and designs—survey of eye movement recording methods," *Behav. Res. Meth. Instrum.*, vol. 7, no. 5, pp. 397–429, 1975.