Depth Compensation Model for Gaze Estimation in Sport Analysis

Fabricio Batista Narcizo and Dan Witzner Hansen
IT University of Copenhagen
Rued Langgaards Vej 7, 2300 København S., Danmark
{fabn,witzner}@itu.dk

Abstract

A depth compensation model is presented as a novel approach to reduce the effects of parallax error for head-mounted eye trackers. The method can reduce the parallax error when the distance between the user and the target is prior known. The model is geometrically presented and its performance is tested in a totally controlled environment with aim to check the influences of eye tracker parameters and ocular biometric parameters on its behavior. We also present a gaze estimation method based on epipolar geometry for binocular eye tracking setups. The depth compensation model has shown very promising to the field of eye tracking. It can reduce 10 times less the influence of parallax error in multiple depth planes.

1. Introduction

Eye tracking is an expressive tool for analyzing human interest and intend [13]. Its potential for detailed and objective performance analysis in sports has been shown in various experiments [1, 12, 17] and the collected information from eye tracking can potentially help athletes to become more effective in their daily training.

Head mounted eye tracking is the most obvious type of eye tracker to use for analysis of athletes during daily training [8] and with current technological developments it could be implemented without significantly disturbing the athletes. However, head mounted eye tracking analysis are typically challenged by the parallax error, which happens as a consequence of the spatial offset between the eye and the camera observing the scene [15]. This causes significant gaze estimation errors (in the scene view) when the apparent objects are located at different depths than during calibration. In addition to the inherent inaccuracy of gaze estimation (0.5 degrees or more), the parallax error will effectively make it hard, if not impossible, to analyze the eye tracking data reliably. Despite of this, many research results are based on manual inspection of video data with overlaid gaze data and the depth compensated point of regard is done based on human estimation when the objects move in space. The parallax error is a practical problem for gaze-analysis in sports where the target (say ball, stone or person) constantly moves in depth relative to the athlete. So far there is no commercial eye tracker that can account for the parallax error and therefore research results are often based on human inspection and estimates on the location the point of regard (PoR) when the target moves in space [13].

In this paper we will present a method that uses the depth as a prior to compensate for the parallax error. Having the object depth can be made feasible via visual tracking or through other sensors. Even without an accurate estimate the method can be used to discern between which of two objects the person is most likely to look at. In Section 2 we describe related work and in Section 3 we describe the parallax error in more detail. The parallax error compensation model is described in Section 4.

The proposed depth compensation model is shown to consistently improve the accuracy level of gaze estimation process when the target is viewed on both calibration plane (Section 5.2) and different depth planes (Section 5.3). Through this paper we intend to show that it is possible to estimate the athletes' gaze actively in given sports situations and thus overcome some of the problems relate to gaze estimation in depth using head-mounted eye trackers. An overview of eye and gaze tracking models is reviewed by Hansen and Ji [8].

2. Related Work

Eye tracking has been used for sports analysis but mostly using head mounted eye tracking [1, 10, 12, 14, 17, 18]. Most eye tracking results are of psychological nature but gaze estimation data collected during daily training where it can be used for analyzing the athletes' performance, such as what happens when the athlete perform specific actions (e.g., shoot, catch, throw)? Are there ocular differences between novice and experts [12, 14, 17]? Which strategies can be used to improve the novice athletes' performance based on knowledge of eye movements patterns collected from training activities of expert athletes [1]? For example, Paeglis *et al*. [17] presented a study for analyzing eye movements from elite junior basketball players during shots practice training. After only a year of training, their free shot rates significantly improved as a consequence of the use of eye tracking in their training. They concluded among other important results that if compared directly with expert basketball players, novice basketball players need more time for quality decisions before making their shots but during free shots, expert players spend more time to do this action than novice players.

Hüttermann *et al*. [12] presented a study for verifying the ability to devote attention simultaneously to multiple visual objects into a athlete's field of view and important concept in team sports, for example. Hüttermann *et al*. [12] showed that athletes present better attention performance when focusing attention simultaneously on two stimuli (i.e., when the athlete fixates between two stimuli). This could mean, in the future, it will be possible to improve training athletes to better focus through of subsequent training studies for improving fixation strategies of a determinate athlete.

Eye trackers for sports analysis are based on the same fundamental eye and gaze tracking models as described in the overview [8]. The parallax error in head-mounted eye trackers (HMET) can be minimized through hardware. Velez and Borah [20] presented an eye tracker that uses a hot-mirror glass in front of the user's eyes for controlling the distances and angular relationships of eye camera, scene camera and user's eyes. The hot-mirror is positioned in the eye's optic axis with aim to reflect images from the eyes and environment toward their respective cameras. This setup removes the parallax error and ensures a wide-angle scene viewing over multiple depth planes. Not all HMET have a physical structure that allows hot-mirror glasses to be used. Mardanbegi and Hansen [15] proposed a study to identify the main sources of parallax error in head mounted eye trackers. They analyzed the influence of scene camera positions, the calibration and fixation distances on the parallax error. They showed that the angle *kappa* (the difference between visual and optical axes) does not have a significant effect on the parallax error [15].

3. Parallax Error in Gaze Estimation

We are going to explain the parallax error based on a HMET. The most of current HMET cannot estimate high accuracy gaze due to the parallax error. Parallax error is a geometrical problem due to the projection center of HMET scene camera and the user's eyeball center are not co-axial. Since the scene camera cannot be placed at the same axis of the user's eye, it is necessary to find out a solution for compensating the parallax error for general HMET.

HMET usually have two (monocular) or three (binocular) cameras attached in their physical structure. For example, a binocular HMET has two eye cameras located slightly close to each eye and one scene camera on the head. The scene camera is used for capturing images from the user's field of view. However, the scene camera is usually not coaxial with any user's eye. In this case, the gaze estimation includes the parallax error on it. Figure 1 shows an example of the parallax error when a person uses an HMET like that.

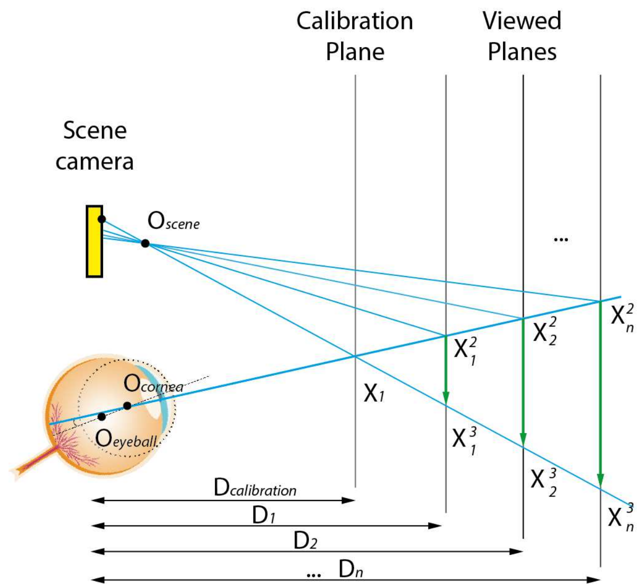

Figure 1. Geometry of parallax error in a HMET. The head mounted eye tracking system is calibrated on the calibration plane. Targets on calibration plane can be estimated with high accuracy. However, when the user looks at to a target in the same position X_1 but on a different plane D_i, the gaze will be estimated on position X_i^3 instead of the position X_i^2. The green arrows represent the parallax error corresponding to the vector $||X_i^2 X_i^3||$ on the viewed plane and a vector $||x^1 x_i^2||$ on the scene camera plane.

The first parameter to be analyzed is the user's eye. According to *Gullstrand-Le Grand Eye Model*, the simplified model for representing the human visual system is formed by two spheres with distinct sizes for representing the eyeball and the cornea surface [4, 5, 11]. The center of rotation of these spheres is around a fixed point $O_{eyeball}$ and there is a small angular difference between the optical and visual axes, which is user dependent and they intersect in the point O_{cornea}. The second parameter is the scene camera, in which is not co-axial with the user's eye. The scene camera is represented as a pinhole camera with a vertical image plane. The last parameter to be analyzed are the planes viewed by the user during the eye tracking session.

The eye tracking system is calibrated in a given distance $D_{calibration}$ from the user to the calibration plane. All points on the calibration plane can be estimated with high accuracy level. However, what happen when the user fixates his/her gaze so far away from the calibration plane (at

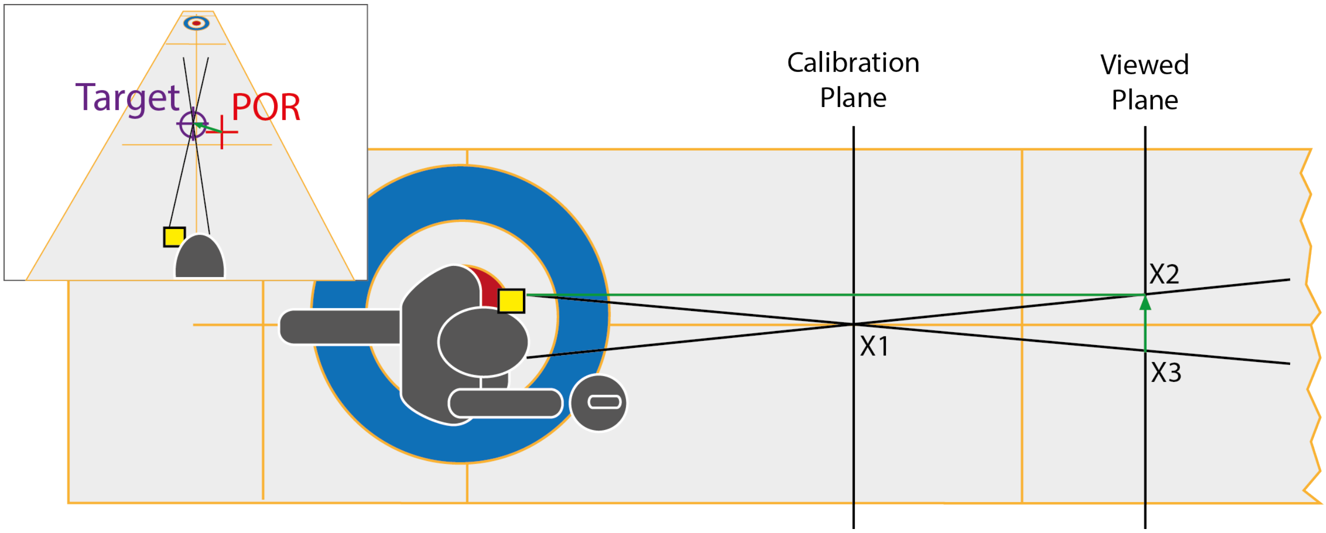

Figure 2. Example of parallax error during a curling daily training session. The system is calibrated on the calibration plane and there is no influence of parallax error on that. On the other hand, when the stone goes to position X_2 on the viewed plane, the gaze will be estimated on position X_3. When the distance between the curling athlete and the viewed stone is prior known, the proposed depth compensation model is able to correct the estimated PoR for the correct position in the scene image, as shown in the small upper picture.

the distance D_i)? The user's visual axis intersects the calibration plane exactly at the point X^1 and the multiple depth planes at the point X_i^2 ($1 \leq i \leq n$). When the user looks directly at the point X^1, the gaze is estimated correctly as the point x^1 on the image plane. On the other hand, when the user looks at to any depth plane at the point X_i^2, the gaze is going to be estimated as the same point x^1 instead of the point x_i^2 on the image plane. According to Mardanbegi and Hansen (2012), the parallax error is defined as a vector $||x^1 x_i^2||$ on the image plane, which is corresponding to the vector $||X_i^2 X_i^3||$ on the viewed plane [15]. Figure 2 shows a practical example of using an eye tracking system for a curling daily training. The parallax error appears in different planes on the curling sheet when the athlete look at to the stone (target) far way from the calibration plane.

4. Depth Compensation Model

The phenomenon of parallax is related by the geometry of the HMET, in which it can be described by *epipolar geometry* in a stereo vision system [2, 15]. In this case, the epipolar geometry is expressed by the point p_{eye} on the eye plane and the point p_{scene} on scene camera plane that must lie on a line called *epipolar line*. As shown in Figure 3, if the point P in the athlete's field of view moves along the line formed by the optical center of the scene camera O_{scene} and the point p_{scene}, its projection on the scene plane will not change but the projection on the eye plane will change. This movement traces out the epipolar line D_{eye} [2].

When the athlete focuses on objects at different planes (see Figure 1), the PoR projection will move along an epipolar line in the image plane. Based on epipolar geometry, all epipolar lines intersect at a common point called *epipole*. In our context, the epipole is placed in the optical center $O_{eyeball}$ into eye plane. Each epipolar line can be estimated through an algebraic representation called *fundamental matrix* (F). F can be estimated given at least seven point correspondences in both image and scene camera planes. These correspondences represent the geometric information about the intrinsic and extrinsic parameters of the cameras.

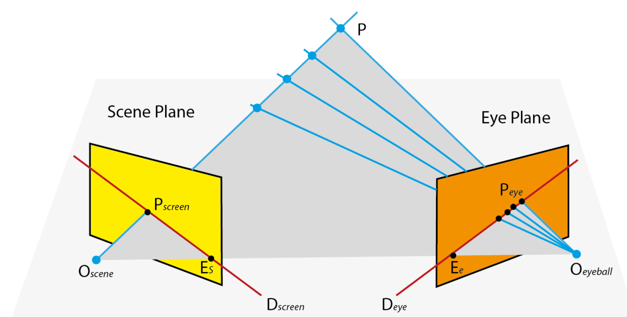

Figure 3. Geometry of parallax error in a HMET.

The fundamental matrix F encapsulates the intrinsic camera geometry and it is independent of scene structure. Given F as a 3×3 matrix, it is possible to calculate the corresponding epipolar line D_{screen} for every point p_{eye} in the eye image plane by Equation 1:

$$D_{screen} = F \times p_{eye}. \tag{1}$$

If any point P is imaged as p_{eye} in the eye camera and

p_{screen} in the scene camera, then the following relation is equivalent to the corresponding epipolar line and constrains the matching of points through Equation 2:

$$p_{screen}^T \times F \times p_{eye} = 0, \qquad (2)$$

for all corresponding points $p_{screen} \leftrightarrow p_{eye}$.

In a binocular HMET, each pair of eye camera and scene camera forms a structure similar to a stereo vision system. We uses a binocular eye tracking approach to define two epipolar lines in the scene camera plane. As the PoR moves along each epipolar line, the intersection between these epipolar lines will be very close to the real athlete's gaze. The biggest problem is when the target moves so far away from the calibration plane, because the parallax error will drastically reduce the precision of the gaze estimation process. Our depth compensation model is based on pure translation motion, i.e. a planar motion case where there is no rotation [9, 19]. According to Hartley and Zisserman [9], the "*pure*" definition means that there is no change in the internal parameters of the cameras.

After estimating the gaze, it is necessary to correct the parallax error based on the information about the depth distance between the viewed target and the calibration plane. As the cameras are stationary in the HMET and we consider that only the targets undergoes a translation $-t$. In this case, the three-dimensional points move on straight lines parallel in the direction of t. One may assume that the calibration plane and viewed plane are respectively $P_{calibration} = K[I|0]$ and $P_{viewed} = K'[I|t]$. Equation 3 is used for calculating the fundamental matrix when there is no rotation ($R = I$) and both camera matrices are the same ($K = K'$):

$$F = [e']_{\times} K' K^{-1} = [e']_{\times} \qquad (3)$$

in which the notation $[e']_{\times}$ is a rank 2 skew-symmetric 3×3 matrix. If the target plane translation is parallel to the calibration plane z-axis, then $e' = (0, 0, 1)^T$. In this case, the fundamental matrix can be represented by the Equation 4:

$$F = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad (4)$$

If a point x in the calibration plane is normalized as $x = (x, y, 1)^T$, then from $x = P_{calibration}X = K[I|0]X$, the space point's coordinates are $(X, Y, Z)^T = ZK^{-1}x$, where Z is the depth of the point X from the viewed plane along the principal axis of the calibration plane. It then follows from $x' = P'X = K[I|t]X$, Equation 5 corrects the estimated gaze point x to the real gaze point x' viewed by the athlete without parallax error:

$$x' = x + Kt/Z \qquad (5)$$

in which depends on the magnitude value of the translation t and the inverse depth Z [9].

5. Assessment on Simulated Data

Simulated eye tracking data were used for assessing the proposed gaze estimation approach (using epipolar geometry) and depth compensation model (using pure translation) in a totally controlled environment. We have used a MATLAB eye tracker simulator in which it is possible to control the eye tracker parameters and the ocular biometric parameters [3]. Therefore, it was possible to evaluate the noise effects of each parameter in the gaze estimation process. The evaluation process was divided according to when the user visualizes targets on the calibration plane (Subsection 5.2) and on the multiple depth planes (Subsection 5.3).

The assessment on the calibration plane has evaluated the accuracy of the proposed gaze estimation approach based on epipolar geometry, and the assessment on the multiple depth planes has evaluated the parallax error rectification of the proposed depth compensation model based on pure translation. We have evaluated the following aspects during our assessment process: (1) refractive index of aqueous humor [α]; (2) number of calibration targets [N]; (3) horizontal [γ] and vertical [β] angle offset between optical and visual axes [a.k.a. *angle kappa*]; (4) the influence of noise in the eye features detection process [$Pc + \lambda$]; and (5) depth movements along to the calibration plane z-axis.

5.1. Setup

The simulated eye tracker device was setup as a binocular HMET. In this case, it had two eye cameras (one for each user's eye) located slightly close to the calibration plane and one scene camera (to get images from the user's field of view) slightly close to the user's head. The calibration plane was adjusted to 55 cm distance from the user. During each test, it was estimated the gaze error from 4,096 targets distributed in a 64×64 matrix over the viewed plane. For all simulated tests, we have used two eye models with different angle kappa offsets [7], namely: $E_0[\beta = \gamma = 0°]$ (a physically infeasible setup only to avoid some eye specific biases) and $E_1[\beta = 1.5°, \gamma = 4.5°]$ (a more realistic ocular biometric setting). For standard, it has used the minimum number of calibration targets necessary to create the fundamental matrix ($N = 8$) during the user calibration process.

5.2. Tests on the Calibration Plane

Refractive Index of Aqueous Humor The first test evaluated the influence of refraction index in the gaze estimation process. According to Hansen and Ji [8], the refractive index of aqueous humor has a constant value around 1.336. It can add some noise or have some directly influence to the gaze estimation process. Table 1 presents the influence of the refractive index of aqueous humor when it is included and when it is not. We concluded that there is no influence of the refractive index in the gaze estimation process. On the other hand, changes in the angle kappa offset present a notable difference for a similar test.

Model	Refraction	Maximum Error	Mean Error
E_0	No	0.0000231°	0.0000037°
E_0	Yes	0.0000231°	0.0000037°
E_1	No	0.0791455°	0.0088679°
E_1	Yes	0.0791455°	0.0088679°

Table 1. The influence of refractive index of aqueous humors [1.336] to the gaze estimation using E_0 $[\beta = \gamma = 0°]$ and E_1 $[\beta = 1.5°, \gamma = 4.5°]$ eye models.

Number of Calibration Targets The second test evaluated how the number of calibration targets ($8 \leq N \leq 25$) influences the gaze estimation process. The user calibration process is performed in the beginning of an eye tracking session. The user needs to look at N targets on the calibration plane for creating a mapping used by the gaze estimation process. Figure 4 shows the accuracy of the gaze estimation as a function of the number of calibration targets. For both eye models the minimum number of calibration targets ($N = 8$) has achieved a good accuracy. In opposite to the classical gaze estimation methods, in a more realistic ocular biometric setting this approach does not improve its accuracy when the number of the calibration targets increase during the user calibration process (see the blue graphic).

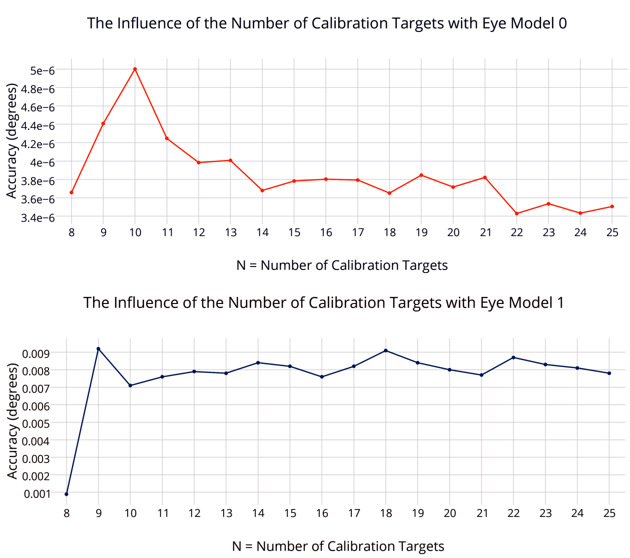

Figure 4. The influence of the number of calibration targets N to the gaze estimation using eye model (up) $E_0[\beta = \gamma = 0°]$ and (down) eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$.

Angle Kappa Offset The third test showed that there is a huge difference among the tests performed with different angle kappa offsets. Figure 5 shows the influence of different angle kappa offsets within a range of angular horizontal offsets ($-4.5° \leq \gamma \leq 4.5°$) and a fixed angular vertical offset ($\beta = 0°$). We conclude that the gaze estimation based on epipolar geometry does not model the angle kappa with high accuracy. The accuracy linearly decreases according to angle kappa, i.e. the bigger the angular difference among visual and optical axes the lower will be the accuracy.

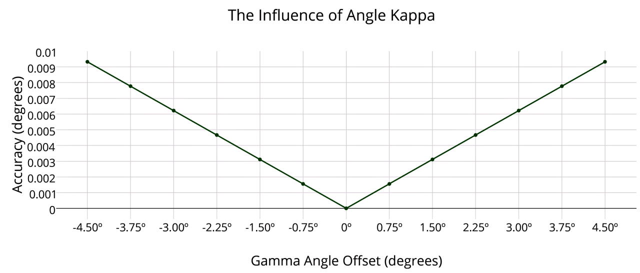

Figure 5. The influence of the angle kappa offset to the gaze estimation. Angle kappa has two angles offsets, i.e. horizontal (γ) and vertical (β). We observed the influence of angle kappa with $-4.5° \leq \gamma \leq 4.5°$ and $\beta = 0°$.

Noise During the aforementioned tests, the gaze estimation approach based on the intersection of multiple epipolar lines showed very promising ($error < 0.01°$). However, in a real application will this approach achieve the same accuracy degree? With aim to answer this question, we have performed a fourth test add a controlled noise in the pupil center coordinate before calculate the epipolar line, i.e. $line_{left} = F_{left} \times (Pc_{left} + \lambda)$ and $line_{right} = F_{right} \times (Pc_{right} + \lambda)$. Figure 6 shows a two-dimensional view of the noise tests with different values to horizontal coordinates and a fixed vertical coordinate $Pc = (x + \lambda, y)$.

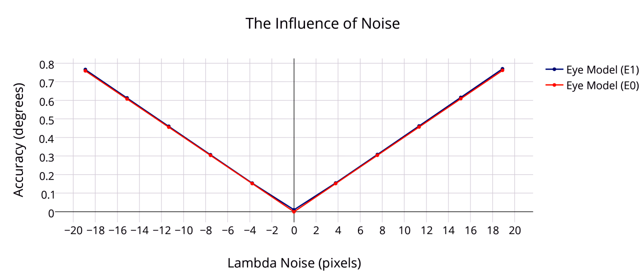

Figure 6. Two-dimensional view of the influence of noise added to the pupil center coordinate to the gaze estimation process using eye model (up) E_0 $[\beta = \gamma = 0°]$ and (down) eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$. The noise ($\lambda$) was added to $P_{center} = (x, y)$ in the following range $-18.90 \leq \lambda \leq 18.90$ pixels.

We concluded that this gaze estimation approach only calculate the PoR with high accuracy level when there is no noise in the eye features detection process. In a real eye tracking application, this approach is going present the same accuracy degree as the classical gaze estimation methods (e.g. homography, cross-ratio, polynomial, among others [5, 7, 16, 23]). Figure 7 shows a three-dimensional views of the influence of noise in the gaze estimation process. The noise was added to each (x, y) coordinate of the pupil center, in the following range $-18.90 \leq \lambda \leq 18.90$ pixels (± 5 mm) on the calibration plane.

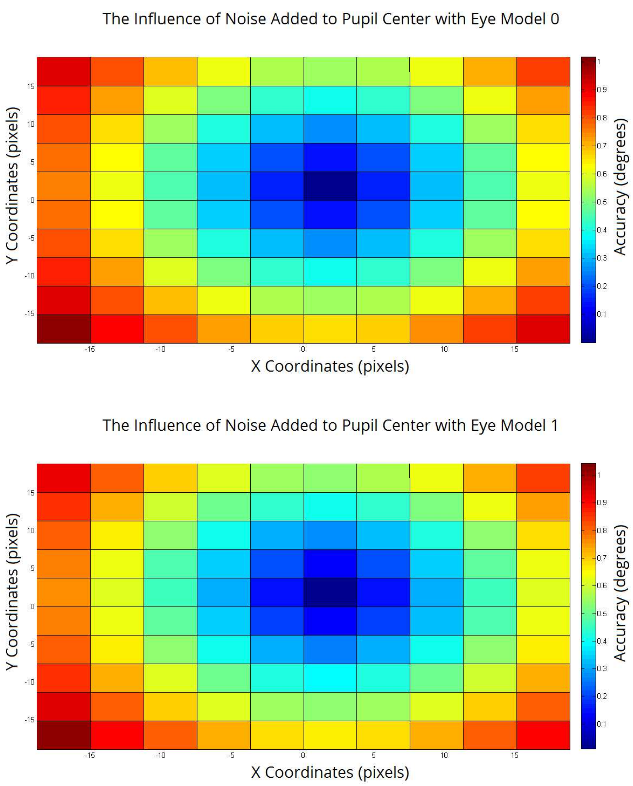

Figure 7. Three-dimensional view of the influence of noise added to the pupil center coordinate to the gaze estimation process using eye model (up) E_0 $[\beta = \gamma = 0°]$ and (down) eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$. The noise ($\lambda$) was added to $P_{center} = (x, y)$ in the following range $-18.90 \leq \lambda \leq 18.90$ pixels.

There is a huge difference only when the noise is $\lambda = 0$. In this case, E_0 presents an accuracy degree around $0.0106921°$ and E_1 is around $0.0000039°$. For others noise level, the accuracy degree is basically the same for E_0 and E_1, i.e. the difference mean is $\pm 0.01°$.

5.3. Tests on the Depth Planes

The last simulated test was performed with aim to evaluate the depth compensation model proposed in this paper. During this test, the HMET hardware components and the user are still while the targets moves along to the calibration plane z-axis. The calibration plane is 55 cm far away from the user and the viewed plane moves in a range 35-105 cm from the user (step of 10 cm). Figure 8 and 9 show the influence of parallax error to the gaze estimation. For each 10 cm far away from the calibration plane, the parallax error adds a gaze error around $\pm 0.23°$.

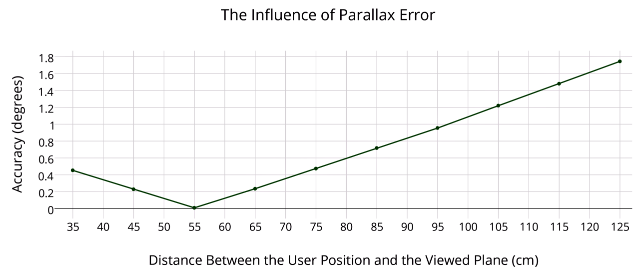

Figure 8. The influence of parallax error to the gaze estimation. The viewed plane was moved to 10 different distances far away from the user position, i.e. in a range 35-125 cm. The accuracy level decrease because the parallax error.

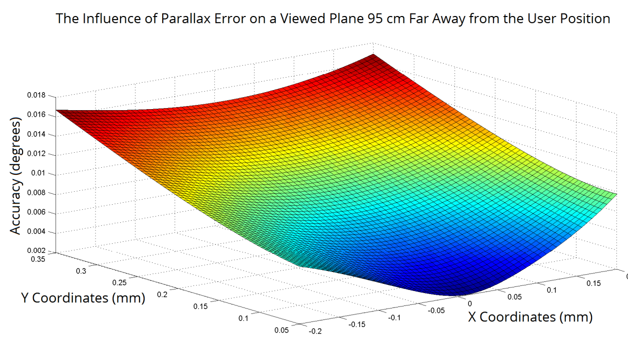

Figure 9. The influence of parallax error to the gaze estimation. After the user calibration process, the viewed plane was moved to 95 cm far away from the user position. The gaze estimation presented an accuracy level around $0.95°$.

Figure 10 and 11 show the influence of proposed depth compensation model to the gaze estimation process. The depth compensation model was able to correct the parallax error of simulated eye tracking data. For each 10 cm far away from the calibration plane, the parallax error adds only a gaze error around $\pm 0.02°$ (i.e. 10 times less). At this point, we concluded that this depth compensation model is very promising to the field of eye tracking.

6. Conclusions

This paper has presented a novel depth compensation model used for correcting the parallax error in HMET. The proposed model is robust to large depth planes when the distance between the user and the target is prior known. The distance is used for compensating the parallax error using the pure translation approach. This paper has also described a gaze estimation method based on epipolar geometry. This method has presented high accuracy degree with simulated date. However, it has shown very sensitive to intrinsic and

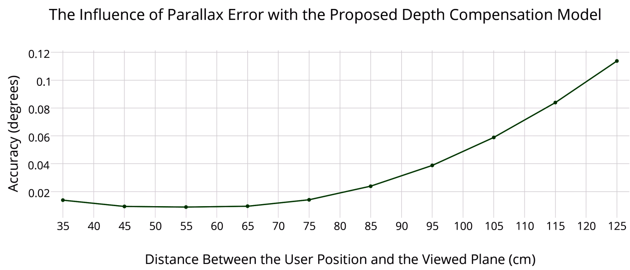

Figure 10. The influence of depth compensation model to the gaze estimation. The viewed plane was moved to 10 different distances far away from the user position, i.e. in a range 35-125 cm. The gaze estimation presented a good accuracy level despite the parallax error.

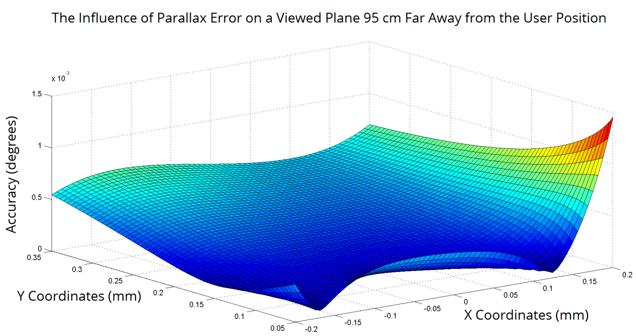

Figure 11. The influence of depth compensation model to the gaze estimation. Using the depth compensation model, the gaze estimation achieved an accuracy level around 0.04° when the viewed plane was moved to 95 cm far away from the user position.

extrinsic noise and it accuracy is similar to others classical gaze estimation methods (e.g. homography, cross-ratio, polynomial, among others).

The proposed model was developed to be used in a bigger project with elite sport athletes (shooting and curling). The main expected contributions by this research project is to develop flexible eye tracking models that can be used for elite sport athletes in their daily training. Eye tracking has been used for sports and has already shown some promise. However, eye trackers used for sports analysis are general purpose, expensive and not adapted to be used actively in sports situations.

Eye tracking can used for collecting information about the pattern of ocular activities of experts athletes and let other novices athletes observe their eye movements. The use of eye tracking in sport can go further, e.g. to auxiliary the hawk-eye technology for evaluating information that has raised doubts during a match [22], to activate resources of a vehicle cockpit through fixations [21], to find the better alternative to view multiple targets during an action of attack or defenses [12] and to identify external points of distraction presents during an eye tracking session [6]. Eye tracking data and tools will allow the athletes and trainers to get much deeper insight into thoughts and strategies used by the athletes, and adapt the training correspondingly thus improving their performance in stressful and time critical situations. While the focus of this project is on sports training, it is evident that progress made within this project on eye tracking and supporting tools for sports activities could have a direct impact on other areas that use eye tracking.

Acknowledgments

The authors would like to thank CNPq – Brazilian National Council for Scientific and Technological Development – for the financial support to the first author in his PhD research project.

References

[1] J. Afonso, J. Garganta, A. Mcrobert, A. M. Williams, and I. Mesquita. The perceptual cognitive processes underpinning skilled performance in volleyball: Evidence from eye-movements and verbal reports of thinking involving an in situ representative task. *Journal of Sports Science and Medicine*, 11(1):339–345, jun. 2012.

[2] S. Bernet, P. Sturm, C. Cudel, and M. Basset. Study on the interest of hybrid fundamental matrix for head mounted eye tracker modeling. In *Proceedings of the 2011 British Machine Vision Conference*, BMVC '11, pages 15.1–15.10, South Road, Durham, UK, aug. 2011. BMVA Press.

[3] M. Böhme, M. Dorr, M. Graw, T. Martinetz, and E. Barth. A software framework for simulating eye trackers. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, pages 251–258, New York, NY, USA, mar. 2008. ACM.

[4] K. Cen and M. Che. Research on eye gaze estimation technique base on 3d model. In *Proceedings of the 2011 International Conference on Electronics, Communications and Control*, ICECC '11, pages 1623–1626, Washington, DC, USA, sep. 2011. IEEE Computer Society.

[5] F. L. Coutinho and C. H. Morimoto. Improving head movement tolerance of cross-ratio based eye trackers. *International Journal of Computer Vision*, pages 1–23, 2012.

[6] T. M. Gable, B. N. Walker, H. R. Moses, and R. D. Chitloor. Advanced auditory cues on mobile phones help keep drivers' eyes on the road. In *Proceedings of the 5^{th} International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '13, pages 66–73, New York, NY, USA, oct. 2013. ACM.

[7] D. W. Hansen, J. S. Agustin, and A. Villanueva. Homography normalization for robust gaze estimation in uncalibrated setups. In *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications*, ETRA '10, pages 13–20, New York, NY, USA, mar. 2010. ACM.

[8] D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):478–500, mar. 2010.

[9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, England, 2^{nd} edition edition, mar. 2004.

[10] M. M. Hayhoe, T. McKinney, K. Chajka, and J. B. Pelz. Predictive eye movements in natural vision. *Experimental Brain Research*, 217(1):125–136, mar. 2012.

[11] C. Hennessey and P. D. Lawrence. Improving the accuracy and reliability of remote system-calibration-free eye-gaze tracking. *IEEE Transactions on Biomedical Engineering*, 56(7):1891–1900, jul. 2009.

[12] S. Hüttermann, D. Memmert, D. J. Simons, and O. Bock. Fixation strategy influences the ability to focus attention on two spatially separate objects. *PLoS ONE*, 8(6):1–8, jun. 2013.

[13] M. Land and B. Tatler. *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. OUP Oxford, 2009.

[14] D. L. Mann, W. Spratford, and B. Abernethy. The head tracks and gaze predicts: how the world's best batters hit a ball. *PLoS ONE*, 8(3):1–11, mar. 2013.

[15] D. Mardanbegi and D. W. Hansen. Parallax error in the monocular head-mounted eye trackers. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 689–694, New York, NY, USA, sep. 2012. ACM.

[16] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner. Frame-rate pupil detector and gaze tracker. In *Proceedings of the $7^t h$ IEEE International Conference on Computer Vision*, ICCV '99, Washington, DC, USA, sep. 1999. IEEE Computer Society.

[17] R. Paeglis, K. Bluss, A. Rudzitis, A. Spunde, T. Brice, and E. Nitiss. Nir tracking assists sports medicine in junior basketball training. In *Proceedings of the International Society for Optical Engineering*, SPIE '11, pages 1–6, Bellingham, WA, USA, jun. 2011. SPIE.

[18] B. R. Pires, M. Hwangbo, M. Devyver, and T. Kanade. Visible-spectrum gaze tracking for sports. In *Proceedings of the $1^s t$ IEEE International Workshop on Computer Vision in Sports*, CVsports '13, pages 1–6, Washington, DC, USA, jun. 2013. IEEE Computer Society.

[19] S. J. D. Prince. *Computer Vision: Models, Learning, and Inference*. Cambridge University Press, Cambridge, UK, England, jul. 2012.

[20] J. Velez and J. D. Borah. Visor and camera providing a parallax-free field-of-view image for a head-mounted eye movement measurement system, aug. 1989. U.S. Patent 4,852.988.

[21] H. Vrzakova and R. Bednarik. Hard lessons learned: mobile eye-tracking in cockpits. In *Proceedings of the 4^{th} Workshop on Eye Gaze in Intelligent Human Machine Interaction*, Gaze-In '12, pages 7:1–7:6, New York, NY, USA, oct. 2012. ACM.

[22] X. Wei, P. Lucey, S. Morgan, P. Carr, M. Reid, and S. Sridharan. Predicting serves in tennis using style priors. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, SIGKDD '15, pages 2207–2215, New York, NY, USA, aug. 2015. ACM.

[23] D. H. Yoo and M. J. Chung. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Computer Vision and Image Understanding*, 98(1):25–51, apr. 2005.