International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 24 (2017) pp. 14652-14657

A Study towards Improving Eye Tracking Calibration Technique Using Support Vector Regression

Farah Nadia Ibrahim[1], Norazlin Ibrahim[2] and Zalhan Mohd Zin[3]

Universiti Kuala Lumpur Malaysia France Institute (UniKL MFI), Section 14, Jalan Teras Jernang, 43650 Bandar Baru Bangi, Selangor, Malaysia.

Abstract

Visual attention and movement have evolved by many ways in eye tracking system. With the mechanism and functionality of eye movement, tracking human eye with high accuracy is possible. In eye tracking system, calibration technique can be considered as the first crucial step to be taken before eye can be tracked. The technique highly depends on a number of calibration points used and the se- lection of these points influences the overall performance of eye tracking system. Knowing the possibility, there are various methods of eye tracking with a good calibration process. However, determining the optimum number of calibration points is a huge challenge when developing eye tracking system. The research focus is then oriented towards determining an optimum number of calibration points by integrating the method of Support Vector Regression (SVR) with the conventional calibration technique. The proposed method should be able to reduce processing time while having high eye tracking accuracy. This paper describes the calibration process and analyses the possible optimum number of calibration points which will be able to reduce processing time and improve accuracy of eye tracking system.

Keywords: Eye Tracking; Calibration System; Processing Time; Accuracy; Support Vector Regression.

INTRODUCTION

Overview of vision

Computer vision and image processing are the fields based on the input of an image or signal for acquiring, analyzing, measuring and processing the images into an output desires. What are the differences between computer vision and image processing? The answer is that, there is no major distinction between the fields since both are using the same input of images. But the aims of both fields are a bit different.

For computer vision, it employs the image processing and ma- chine learning to extract and to understand the information of the image data and produce the knowledge of the image. But image processing only focuses on obtaining this information from raw image data without learning or giving feedback to the image. The image will only be processed by going through several transformation or filtering methods and produce a processed image.

Eye Detection and Tracking

Eye tracking is one of the vision system that become widely known nowadays in the world of vision technologies. Its attraction is the usage of eye as the input controller. Modern world these days are equipped with various high-end technologies and applications such as smartphones, tablets and laptops. These gadgets are implemented with touch sensor, where the human can use and control their technology by using only the touch of the finger. As for eye tracking technology, many researchers and developers are eager in producing an efficient and accurate eye tracking system that is robust and user friendly. But to produce such an efficient system, the calibration process is considered as the major element to be focused on.

Eye Movement

Eye movement is the voluntary and involuntary movement of the eyes functionally in acquiring, fixating and tracking the visual stimuli. Two major functions of the eye movements are the fixa- tion and tracking. Fixation is the maintaining of visual gaze on a single location. It allows human eyes to maximize the focus at a single object. While as tracking is the ability to control the move- ment of the eyes. There are four other basic types of eye move- ments which are saccades, smooth pursuit movements, vergence movements, and vestibulo-ocular movements.

Saccades are quick, rapid eye movements that abruptly change the point of fixation. Saccade is very fast that the eye cannot capture the images and the speed of movement during each saccade cannot be controlled.

Smooth pursuit movement is used to follow a moving object. This movement is smooth and continuous but requires a constant feed- back to track the moving object. It is slower than saccade and less accurate than the vestibulo-ocular movement.

Vergence movement is to select the distance of the object that the human want to observe. To look at an object close by, the eyes rotate 'towards each other' (convergence), while for an object far- ther away they rotate 'away from each other' (divergence).

Vestibulo-ocular movement is a reflex eye movement that stabi- lizes images on the retina during head movement by producing an eye movement in the direction opposite to head movement, thus preserving the image on the center of the visual field. This movement will move the eyes in the opposite direction of the head to keep the eyes focused on the same point at all times.

RELATED WORKS

Tracking the Human Eye

There are many eye tracking methods and studies have been discussed in the literature review such as by using Artificial Neural Network (ANN) [1], Electro-Oculography (EOG) [2], [3], [4], pupil detection [5], [6], [7], [8], [9], [10] and also eye and hand gestures [11], [12], [13], [14]. Through different methods, the outputs of the tracking system are all almost the same, aiming for an efficient, less complicated system and good accuracy of eye detection.

Calibration-free

Calibration is the very important element to measure and detect the eye accurately. This process is far from perfect. Thus, there are various comparative studies regarding this element from improving the calibration technique to the calibration-free, where the researchers aiming on removing the calibration process.

Dmitri Model and Moshe Eizenman [15] have proposed general method by extending the tracking range for the remote eye gaze tracking system. By using a stereo pair of cameras, the overlap- ping field of view is used to estimate the user eye parameters. But to compute the Point-of-Gaze with the estimation of pupil center and corneal reflection, the two-camera stereo mode must be separated into the split mode (one camera system) since the split mode has larger range of gaze direction (see Figure 2.1).

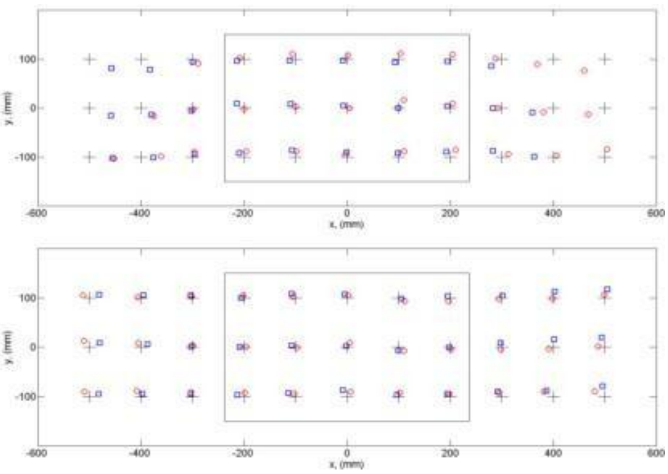

Figure 2.1: Stereo mode system (top) and split mode system (bottom).

There is another approach of gaze estimation without personal calibration which 3D gaze estimation as shown in Figure 2.2. Jixu Chen and QiangJi [16] have implemented 3D gaze estimation method by combining it with a visual saliency map. In order to receive the estimation point, the user only needs to gaze at the image given without looking at any specific point. By using novel incremental learning framework, this system does not need any training data beforehand and allows natural head movement com- pared to system as described by Y. Sugano, Y. Matsushita, and Y. Sato in [17].

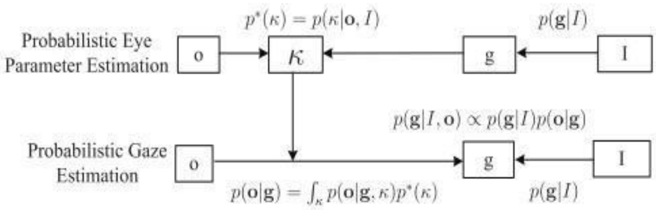

Figure 2.2: Diagram of proposed probability gaze estimation.

In a flexible uncalibrated setup with a web camera and no chin rest, F. Alnajar, T. Gevers, R. Valenti, and S. Ghebreab [18] present a method where the gaze points are estimated after looking at a stimulus for a few seconds. This method uses gaze patterns of individuals to estimate the gaze points for new viewers without active calibration. However, by using the human gaze patterns, evaluation on several users and images are needed to test the accuracy of the system.

Another approach is by using only an off-the-shelf camera to track the user gaze horizontally [19]. Y. Zhang, A. Bulling, and H. Gellersen introduced the Pupil-Canthi-Ratio (PCR) as a novel measure for estimating the gaze directions (see Figure 2.3). This system extracted the eye feature points by using image processing technique. The coordinate of inner eye corner and pupil center are denoted and being defined together with the horizontal distance of each eye to be defined as the distance ratio called Pupil-Canthi- Ratio or PCR.

For data analysis, Gaussian process regression is used to establish the mapping between the PCR and gaze direction on a display. A single webcam is used to record the eye images under normal lightning condition such as in the office environment. According to [19], PCR can be considered as a robust measurement ratio and it is able to achieve an average accuracy, but the head tilting influences a lot the performance of PCR.

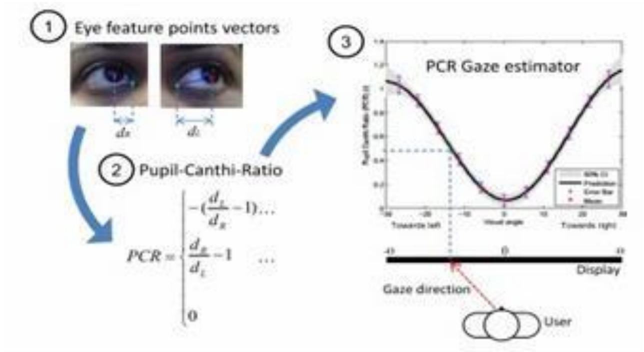

Figure 2.3: Overview of proposed gaze estimation using (PCR).

Calibration Technique

Calibration-free is indeed a great approach in order to eliminate active calibration processes that are thought to be tedious and cumbersome. Nevertheless, there are still many researchers aiming to improve the current calibration techniques respecting to its necessity for high accuracy of eye tracking system.

K. Harezlak, P. Kasprowski, and M. Stasch [20] have focused on the process of calibration, analyses of the possible steps and how to simplify this process. They have compared the results taken from three regression methods, Classic Polynomial, Artificial Neural Networks (ANN) and Support Vector Regression (SVR) to build calibration models with sets of calibration points.

The calibration procedure has been done by using a set of 29 points. The results revealed that ANN and SVR methods are better that Polynomial (Figure 2.4) method in handling bad samples. But Polynomial method has been able to come out with low error models using less number of points and higher accuracy. The out- come from good layout of calibration point is another important conclusion achieved in providing lesser number of calibration points without affect the accuracy level.

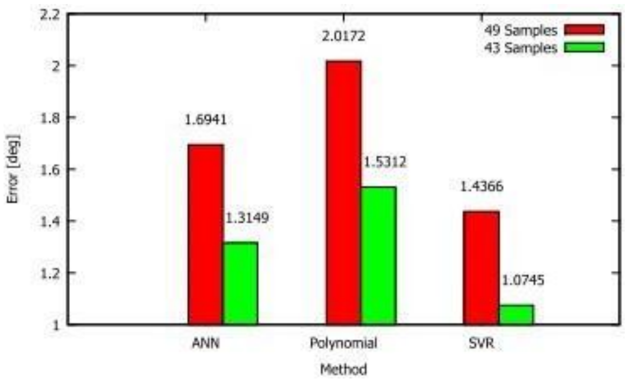

Figure 2.4: Comparison of error values before & after removing incorrect samples.

Nuri Murat, Hua Gao and Jean-Philippe Thiran [21] have proposed a cross-ratio based automatic gaze estimation system that accurately works under natural head movements. They used a subject- specific calibration method based on regularized Least-Squares Regression (LSR) for achieving higher accuracy (see Figure 2.5). The system does not require high resolution eye data enables the system to output the Point-Of-Regions from each eye simultaneously. In their work in [21], the overall estimation of the accuracy has been improved.

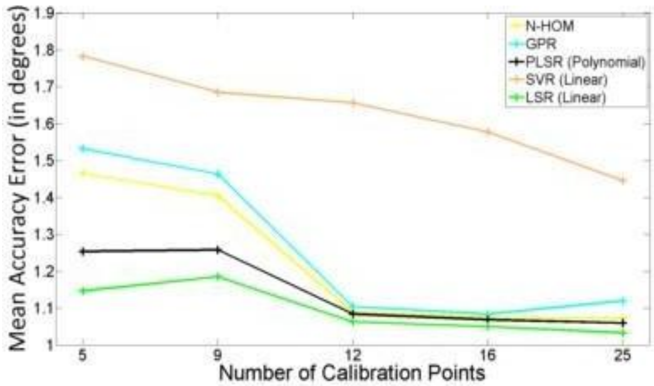

Figure 2.5: Comparison with different regression techniques (learning calibration models).

By using one camera and one infrared source, a simple video tracking system was developed to compare the accuracy of various mapping models. Pieter Blignaut and Daniël Wium [22] compared model-based and regression-based gaze estimation to determine the user Point-of-Regard (POR). The gaze distance for the calibration target arrangements have to be controlled in order to compare the accuracy of the mapping models.

Nine different sets of calibration targets were divided into a 15x9 grid (see Figure 2.6) to have the same width: height ratio as that of the calibration display area (1200×720 pixels). Smaller deviation values produce good accuracies from mapping model. Pieter Blignaut and Daniël Wium found out that the calibrations are not robust with larger backward (>5cm) and forward movements of the head.

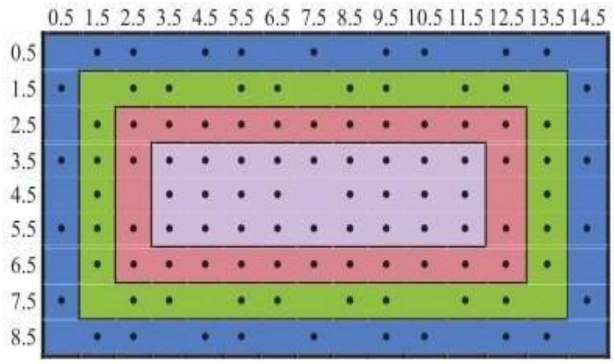

Figure 2.6: Targets for calibration in a 15x9 grid.

K. Pfeuffer, M. Vidal, J. Turner,A. Bulling and H.Gellersen [23] present a novel approach to detect the user attention moving cali- bration targets. This work focused on smooth pursuit eye move- ment in order to capture the user's attention and making it less tedious on calibration process. This system collects sample of gaze points by displaying a smooth moving target to the user. Three realistic applications (see Figure 2.7) have been designed with different abilities in each. Automated Teller Machine (ATM) is created by selecting a few digits with an accurate eye tracking.

Figure 2.7: (a) moving target, (b) ATM, (c) stargazing and (d) waiting screen.

This application is also included with a distraction system where the calibration process will be "paused" when the user is distracted and it will be able to continue the calibration without re-calibrate from the beginning. Second application is the stargazing. A shooting stars will appear at a random four trajectories and are able to attract the user attention naturally.

Lastly is the waiting screen application where the user will be unaware of the calibration process. Four floating words and one word is fixated in the middle with different languages are moving smoothly while the user unconsciously read and follows the word- ing without been know that it is part of the calibration process. However, this study showed that too slow and too quick effects the accuracy of the system. The size of the calibration target is also one of the elements that can influence calibration accuracy.

PROPOSED METHOD

Research System Process and Framework

The eye tracking system will be processed by following the sequences of this methodology as shown in Figure 3.1.

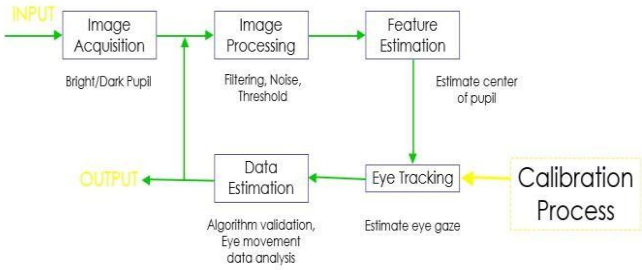

Figure 3.1: Eye tracking system process.

As for the proposed system framework, the processes are illustrated in Figure 3.2.

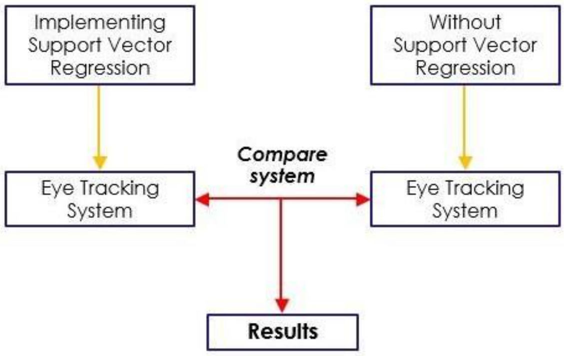

Figure 3.2: Proposed eye tacking system framework.

By studying the previous related work, the focus will be given on the usage of a regression method by Support Vector Regression (SVR) instead of using the traditional method of polynomial regression. It has the purpose to enhance the calibration technique thus improving overall eye tracing system performance. Since the proposed calibration process is aiming on investigating the opti- mum value that will be few number in samples and a nonlinear data, SVR can be considered as the most suitable method to be applied in this case.

Based on the research by K. Harezlak, P. Kasprowski, and M. Stasch [20], they have compared three type of regression method. Polynomial method is able to produce low error and high accuracy in the training data, but the SVR method able to detect the bad error with few number of sample and also led to a good accuracy. The element that can be improved from this study is that by using SVR, this research will improve the accuracy higher than the polynomial method by comparing two types of data, that is the training data and testing data.

These research objectives are to identify the optimum number of calibration points that will reduce calibration processing time and to improve the accuracy of the eye tracking system. By improvise the limitation of the aforementioned; this research can achieve better gaze calibration together with the implementation of SVR technique that can reduce bad error of sampling and at the same time high accuracy. This system will be developed and tested by using Python programming language.

Data Collection and Analysis

The data collection for this research will be conducted by using a webcam, Logitech HD Pro Webcam C920 which is a full 1080p high definition. This webcam is equipped with the application of video and photo capture, face tracking and also motion detection. The specification of able to control by pan, tilt and zooming are beneficial in conducting this research.

This research will train and test the data with different number of calibration points. The minimum points will be 4 points and the maximum points will be 9 points. The points will be distributed evenly towards the screen and will be show to the user for training and testing.

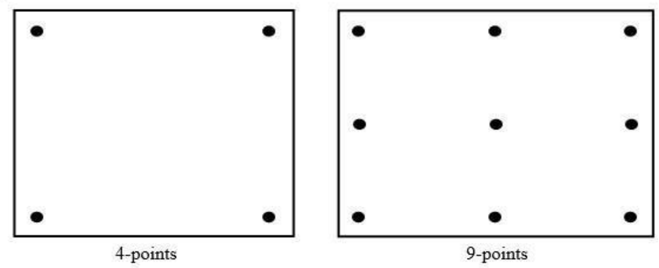

Figure 3.3: 4-points and 9-points of calibration points.

For the data analysis, the regression method will be used to map the training data with the user's gaze estimation towards the posi- tion of the calibration points. Non-linear SVR and kernel functions will be used and they are defined as in Figure 3.4.

Non-linear SVR:

$$y = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \cdot \langle \varphi(x_i), \varphi(x) \rangle + b$$

$$y = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) \cdot K(x_i, x) + b$$

Kernel Functions:

Polynomial

$$k(\mathrm{x}_i, \mathrm{x}_j) = (\mathrm{x}_i . \mathrm{x}_j)^d$$

Gaussian Radial Basis function

$$k(\mathrm{x}_i, \mathrm{x}_j) = \exp\left(- \frac{\left\| \mathrm{x}_i - \mathrm{x}_j \right\|^2}{2\sigma^2} \right)$$

Figure 3.4: Support Vector Regression Functions.

As for the process of the algorithm testing, to evaluate the perfor- mance of the algorithm, the function of Mean Square Error (MSE) will be used. MSE is defined as in Figure 3.5.

$$\sum_{i=1}^{N} \sqrt{\left(x_{oi}^{'} - x_{oi}\right)^2 + \left(y_{oi}^{'} - y_{oi}\right)^2} \Big/ N$$

Figure 3.5: Mean absolute error, MSE.

where,

(xo', yo') is the estimated gaze position coordinates, and (xo,yo) is the actual gaze position coordinates.

CONCLUSION

The technology of eye tracking system has become widely used and useful in many integrated applications particularly in vision technology. This research reviews some related work to this appli- cation. It should be able to help us to have more understanding on eye tracking system and also image processing while improvised the vision technology in our country. Besides that, the improved number of research regarding eye tracking system will benefits a lot of references and studies to the future researchers. Therefore, this integrated system of eye tracking and image techniques will be essential for efficient system in the future onwards.

ACKNOWLEDGEMENT

The authors would like to thank Ministry of Education Malaysia and Universiti Kuala Lumpur (UniKL) for funding this research through research Grants (FRGS/1/2016/ICT02/UNIKL/02/2) and (UniKL STR15142).

REFERENCES

[1] E. De mjén, V. Aboši, and Z. Tomori, "Eye tracking using artific ia l neural networks for hu man co mputer interaction.," *Physiol. Res.*, vol. 60, no. 5, pp. 841–4, 2011.

[2] N. Steinhausen, R. Prance, and H. Prance, "A three sensor eye tracking system based on electrooculography," *IEEE SENSORS 2014 Proc.*, pp. 1084–1087, 2014.

[3] P. Aqueveque and E. J. Pino, "Eye Tracking Capabilities of Low- Cost EOG System," *36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 610–613, 2014.

[4] H. Manabe, M. Fuku moto, and T. Yagi, "Automatic drift ca libra - tion for EOG-based gaze input interface," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 53–56, 2013.

[5] R. Shimata, Y. Mitani, and T. Ochiai, "A Study of Pupil Detection and Tracking by Image Processing Techniques for a Hu man Eye – Computer Interaction System," pp. 3–6, 2015.

[6] I. I. Conference and C. October, "Development of the Eye Input Device Using Eye Movement Obtained by Measuring the Center Position of the Pupil," pp. 2948–2952, 2012.

[7] A. B. Roig, M. Morales, J. Espinosa, J. Pere z, D. Mas, and C. Illueca, "Pupil detection and tracking for analysis

of fixational eye micromovements," *Opt. - Int. J. Light Electron Opt.*, vol. 123, no. 1, pp. 11–15, 2012.

[8] F. Timm and E. Barth, "Accurate eye centre localisation by means of gradients," *Int. Conf. Comput. Theory Appl.*, pp. 125–130, 2011.

[9] F. U. Robert G. Lupu, Radu G. Bozo mitu and V. Cehan, "Eye Tracking Based Communication System for Patient with Major Neoroloco motor Disabilites," *Syst. Theory, Control. Comput.*, pp. 1– 5, 2011.

[10] B. Yan, X. Zhang, and L. Gao, "Improvement on Pupil Positioning Algorithm in Eye Tracking Technique," *2009 Int. Conf. Inf. Eng. Comput. Sci.*, pp. 1–4, 2009.

[11] N. K. Chuan and A. Sivaji, "Combining Eye Gaze and Hand Tracking for Pointer Control in HCI Developing a more robust and accurate interaction system for pointer positioning and," no. Chuser, pp. 172–176, 2012.

[12] Y. Y. Pang, N. A. Ismail, P. Leong, and S. Gilbert, "A eal-Time Vision-Based Hand Gesture Interaction," 2010.

[13] Y. C. Ha m, "Developing a Smart Camera for Gesture Recognition in HCI Applications," pp. 994–998, 2009.

[14] G. Shin and J. Chun, "Vision-based Multimodal Hu man Computer Interface based on Parallel Tracking of Eye and Hand Motion," pp. 2443–2448, 2007.

[15] D. Model and M.Eizen man, "User-calibration-free re mote eye- gaze tracking system with extended tracking range," *Can. Conf. Electr. Comput. Eng.*, pp. 001268–001271, 2011.

[16] J. Chen and Q. Ji, "Probabilistic gaze estimation without active personal calibration," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 609–616, 2011.

[17] Y. Sugano, Y. Matsushita, and Y. Sato, "Calibration-free gaze sensing using saliency maps," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2667–2674, 2010.

[18] F. Alnajar, T. Gevers, R. Valenti, and S. Ghebreab, "Ca libration - Free Gaze Estimation Using Human Gaze Patterns," *2013 IEEE Int. Conf. Comput. Vis.*, pp. 137–144, 2013.

[19] Y. Zhang, A. Bulling, and H. Gellersen, "Pupil-canthi-ratio: a calibration-free method for tracking horizontal gaze direct ion," *Proc. AVI*, pp. 129–132, 2014.

[20] K. Ha rezla k, P. Kasprowski, and M. Stasch, "Towards accurate eye tracker calibration – methods and procedures," *Procedia - Procedia Comput. Sci.*, vol. 35, pp. 1073–1081, 2014.

[21] N. M. Arar, H. Gao, and J.-P. Thiran, "Towards Convenient Calibration for Cross -Ratio Based Ga ze Estimation," *2015 IEEE Winter Conf. Appl. Comput. Vis.*, pp. 642–648, 2015.

[22] P. Blignaut and D. Wium, "The • Effect • of • Mapping • Function • on • the • Accuracy • of • a • Video - Based • Eye • Tracker •," vol. 1, no. August, pp. 29–31, 2013.

[23] K. Pfeuffer, M. Vidal, and J. Turner, "Pursuit calibration: making gaze calibration less tedious and more flexible," *Proc. 26th* , pp. 261–269, 2013