Analysing the Potential of Adapting Head-Mounted Eye Tracker Calibration to a New User

Benedict Fehringer [*]
DFKI Saarbrücken

Andreas Bulling [†]
University of Cambridge
Lancaster University

Antonio Krüger [‡]
DFKI Saarbrücken

Abstract

A key issue with state-of-the-art mobile eye trackers, particularly during long-term recordings in daily life, is the need for cumbersome and time consuming (re)calibration. To reduce this burden, in this paper we investigate the feasibility of adapting the calibration obtained for one user to another. Calibration adaptation is automatically performed using a light-weight linear translation. We compare three different methods to compute the translation: "multi-point", where all calibration-points are used, "1-point", and "0-point" that uses only an external parameter. We evaluate these methods in a 6-participant user study in a controlled laboratory setting by measuring the error in visual angle between the predicted gaze point and the true gaze point. Our results show that, averaged across all participants, the best adapted calibration is only 0.8° (mean) off the calibration obtained for that specific user. We also show the potential of the 1-point and 0-point methods compared to the time-consuming multi-point computation.

CR Categories: H.1.2 [Information Systems]: Models and Principles—User/Machine Systems; I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Camera calibration;

Keywords: Eye Tracking, Calibration, Adaptation

1 Introduction

Eye tracker calibration refers to finding the parameters of the eye model used to determine the point of gaze (POG), i.e. the offset between visual and optical axis. Good calibration is crucial to achieve high gaze tracking accuracy that means a low deviation between the predicted and the actual POG. Despite considerable advances in calibration techniques, accurate eye tracker calibration is still challenging to perform, time-consuming, and error-prone.

While the burden caused by calibration may be acceptable in laboratory settings using remote eye trackers, field studies with mobile systems, e.g. in a supermarket, are more critical. In these settings, participants may be recruited on-the-fly and may be more willing to participate if the preparation time is short and they could start without a calibration procedure. Additionally, in contrast to stationary systems, the initial calibration obtained for a mobile system may be hampered due to body movements, which requires additional recalibrations.

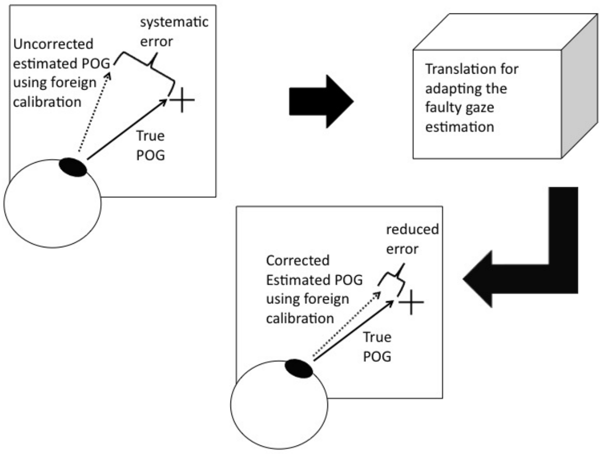

Figure 1: *Overview of the adaptation approach. The error between the estimated point of gaze (POG) using a foreign calibration and the true POG is corrected using a translation.*

This work is part of an ongoing effort to develop a calibration-free monocular eye tracking system. As a first step towards this vision, in this paper we investigate the feasibility of adapting the calibration obtained for one user (the "foreign calibration") to a new user. To perform this adaptation we propose a light-weight approach that uses a constant linear geometric transformation. In a user study with six participants we compare three methods for computing the main parameter of this linear transformation, namely "multi-point", "1-point", and "0-point" computation.

2 Related Work

The development of calibration-free eye tracking systems has seen increasing research interests over the last years. Considerable advances have been achieved in the calibration of stationary eye trackers, e.g. with stereo cameras [Model and Eizenman 2010; Nagamatsu et al. 2010]. These approaches are promising for developing calibration-free eye trackers but they require binocular tracking, which is based on the assumption that both eyes look at the same point. This assumption can be questioned [Dell'Osso 1994] and doesn't always hold in concrete cases, e.g. if the user has lost or is blind on one eye or if the eyes are not properly aligned with each other. Other approaches model the eye's anatomy and movements in a more elaborate manner (e.g. [Hansen et al. 2010]). While providing good gaze estimation accuracy these approaches require detailed measurement information for calculating the eye model that is not available on off-the-shelf eye trackers. Our aim is to find a simple and fast calibration procedure without any model assumptions or additional information.

---

[*] e-mail: benedict.fehringer@dfki.de
[†] e-mail: andreas.bulling@acm.org
[‡] e-mail: antonio.krueger@dfki.de

3 Method

The goal of the current work is to investigate the feasibility of adapting a foreign calibration to a new user, thus rendering an explicit second calibration obsolete. Such an adaptation leads to a systematic error in the estimated point of gaze because of wrong calibration parameter values. We aim to correct this systematic error with a constant translation of all estimated POGs (from the scene image of the eye tracker) using the following formula: $\forall i : x'_i = x_i + c$ (cf. Figure 1). x_i is the uncorrected estimated POG, c is the constant to adapt the uncorrected POG, and x'_i is the corrected POG that - if the translation works well - is close to the true POG. c is equal for all POG. In our setting the true POGs are all in a plane, so we do this correction for both coordinates, each with its own constant.

The accuracy of this adaptation - the differences between the estimated and the true POG - strongly depends on the visual behaviour of the user. To illustrate this dependency let us consider two hypothetical extreme cases. If only the head is moving and the eyes are not moving at all, the eye orientation relative to the eye tracker remains the same. That means the deviation between the uncorrected, estimated and the true POG is always the same, independently how the head is oriented. If the constant is set to this deviation the corrected POG's error should be close to zero, beside of "noise". In contrast, if only the eyes are moving and the head remains stable, gaze can cover the whole glasses' view field. In this case the constant changes (probably) depending on the eyes' orientation relative to the eye tracker and results in a corrected POG's error that is larger than zero.

This translation constant can be computed using three basic approaches: Firstly, to use different view angels and as kind of an optimal correction, many measured points from the view field can be used ("multi-point"). Secondly, the constant can be computed with only one point ("1-point") and thirdly it can be computed by only using an external parameter, that can be measured without any help by the subject ("0-point").

4 User study

We designed a user study to investigate the feasibility of adapting calibration to a new user and to evaluate the accuracy of the proposed adaptation approach. The experimental design contained one factor with three levels. In the first condition participants were asked to not move their head while looking at the different points. In the second condition they were asked for the opposite, namely to move their head but not their eyes. In the third condition, the baseline, participants were asked to look at the points in a "natural way", i.e. to move their eyes and head as they would in daily life. The last condition is only to compare to the other ones, so that we can be sure that the first two conditions are the extremes.

4.1 Participants

Five participants took part in the study (two female, three male), aged between 15 and 33 years ($mean = 24.5$, $sd = 5.79$). One participant normally needs glasses but did not wear them during the experiment. One participant wore contact lenses during the experiment. The calibration quality of the Tobii Eye Tracker Glasses ranged from two to five on the six-point-scale (with zero indicating the worst and five the best calibration quality).

We had one additional participant from whom we took the calibration as foreign calibration of the others. He was male, 30 years old, has a height of $1.86m$ and eye height of $1.71m$. He was not wearing glasses during the experiment. His calibration quality was five on the Tobii Eye Tracker Glasses calibration scale.

4.2 Apparatus

For recording gaze data we used the Tobii Glasses. The Tobii Glasses is a mobile video-based eye tracker recording monocular gaze data from the right eye at a sampling rate of 30 Hz. The system has a camera to record a scene video with a resolution of 640x480 pixels. The maximum recording angles are 56 degrees of visual angle in horizontal and 40 degrees of visual angle in vertical direction. With this information we approximately computed the relation between the pixels and the visual angle based on proportionality to $1^\circ \widehat{=} 640 pixel/56 = 11.4 pixel$ in horizontal direction (x direction) and $1^\circ \widehat{=} 480 pixel/40 = 12 pixel$ in vertical direction (y direction).

To map the estimated POG the Tobii system uses several infra-red markers (IR-markers) to be placed in the environment. All markers send a unique ID that can be recognised by the glasses, so that the glasses "know" the position of each detected marker in each video frame. Based on this information the points in the video can be marked and the relative position to the markers can be stored. In later frames the position of the points in the video can be recomputed by the detected markers and the stored relative position to them. For recording movements of the eye tracking glasses we used the OptiTrack Arena motion tracking system. The Arena system consists of a set of infra-red cameras that visually detect and track objects in 3D using reflecting markers attached to the object. For our purpose we measured the rotation angle in all three Cartesian axis. The accuracy is not worse than 0.1° and the sampling rate is at $100Hz$.

4.3 Correction methods

To correct the deviation of the foreign calibration we used three methods: Firstly we wanted to approximate the best gaze prediction that can be reached with a constant translation. To this end we used all measured points ("multi-point") to compute the translation constant. The constant was simply the mean deviation of all estimated POGs from the corresponding true POGs. This computation procedure is similar to the one in the 1-point method where the deviation only of the central estimated POG from the true POG was used as the constant. The central point is in the intersection of the 3^{th} row and the 3^{th} column (see Figure 2).

For the last correction (0-point) for every subject the relative pupil's position of the right eye to the glass' frame was measured based on the subject's "front view". Then the distance between the pupil's position of the i^{th} person who is using the foreign calibration and pupil's position of the person who is given the foreign calibration was computed (Δd_i). Now, we wanted to bring this distance in relation to the optimal translation constant. As approximation for this optimal constant, the constant from the multi-point method was taken. After that the constants of each person i ($c_{i;multi-point}$) were divided by the measured distance ($f_i = c_{i;multi-point}/\Delta d_i$). If our assumptions would be as simple as we have done all f_i's would be equal. For our further computation we took the mean of all f_i's, labelled as $\bar{f}$. For each participant i the translation constant was computed by $c_{i;eyePos} = \bar{f} * \Delta d_i$. The idea is that after $\bar{f}$ is fixed, for a new user we only need the relative pupil's position and then immediately, we can compute the translation constant.

4.4 Setup and procedure

After arriving in the lab participants were first introduced to the equipment as well as the purpose and the procedure of the experiment. Afterwards, participants put on the glasses and performed a standard nine point calibration. Then, before each condition, the participants were asked to stand in front of a projected screen at

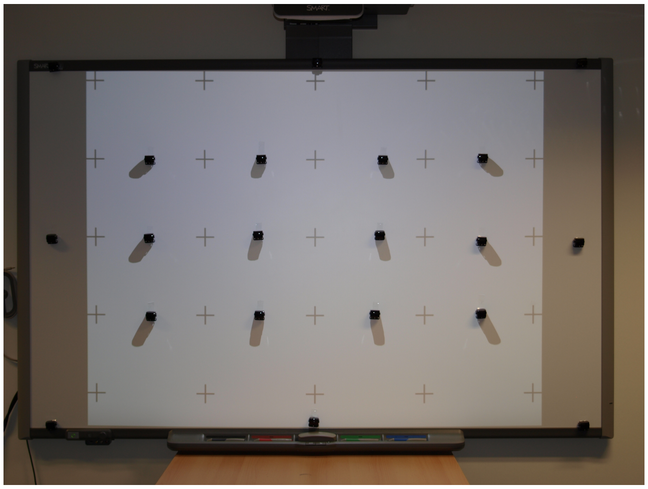

Figure 2: *Screen showing the 25 calibration points (indicated as black crosses) as well as the attached Tobii eye tracking markers.*

a distance of about two meters. They were asked to look at a sequence of 25 targets according to the current condition (“eye only”, “head only” or “normal”). These targets were black points shown on the screen at predefined locations given by a grid (see Figure 2). Before each of these targets was shown on the screen, a series of three coloured crosses (red, yellow, and green) was presented at the same location to “prepare” the participants for the upcoming point. Each cross was shown for $1s$ and each target for $5s$. If there are no missing values we got 150 measuring points ($30Hz$ glasses’ rate with a $5s$ duration). The order of the presented targets was from top left to bottom right, row by row. After the three conditions, the participants completed the questionnaire with questions on age, sex, as well as the right eye’s dioptre.

5 Results

At first we checked our data set and excluded every point if less than 10 frames were recorded, e.g. if the gaze could not be computed, no IR-markers were detected or the POG was out of the camera’s view-field. Hence, for some points the recorded frames were more than for others. To avoid any potential bias, all computations were only done using the means for each point. Also, for some participants in some conditions not all 25 points are recorded, so in the multi-point method the number of used points varies. In condition “head only” only participant 2 has less than 25 points (21 points) for the foreign calibration. For condition “eye only” the number of points goes from 12 to 22. Afterwards we randomly chose one participant with the best calibration quality (five points on the scale) to deliver his calibration result as foreign calibration for the other participants.

5.1 Analysis of eye and head movements

To check how good participants followed our instructions we compared the eye movement relative to the glasses and the head movement for all three conditions. The eye movement was measured by standard deviation (sd) of the x- and y-coordinates (measured as angles, like described above) from the eye tracker glasses and the head movement by the sd of the rotation values delivered by the OptiTrack Arena system.

The computed two two-factor ANOVAs have as first factor the condition and as second factor either the coordinates’ sd or the rotation values’ sd. Each of the two ANOVAs delivered significant Pillai-Traces [Pillai 1955] as overall result, $F(1,4) = 127.675, p < 0.001$ for the first ANOVA and $F(1,4) = 32.023, p < 0.01$ for the second one. The differences between the first two conditions and the first and the second condition were also significant in both ANOVAs. How we expected: The most eye movement was in condition one (mean from both coordinates: $12.00°$), the fewest in condition two ($3.27°$) and vice versa for the head movement: $1.00°$ (mean from all three rotation values) for condition one and $11.06°$ for condition two. The eye movement for the third condition was in between ($6.88°$) and the head movement quite similar to second condition ($10.78°$) with no significant difference.

The results show that participants were able to follow the instruction in the first two conditions and that the head’s and eye’s movements in the natural condition were in between. This results confirms that the first two conditions can be considered as extremes concerning visual behaviour.

5.2 Analysis of tracking accuracy

The error between the true POG (point where the subjects had to look at) and the estimated POG for the own and foreign calibration is shown for condition one (Table 1) and two (Table 2). Independent of the condition the foreign calibration results in bigger errors than the own calibration. Independent of the method the adaption works better for “head only” than “eye only”. The differences of the means between both calibrations range from $3.8°$ (condition one, y-coordinate) to $> 8°$ (two, x). The multi-point correction improves the estimation in nearly all cases: For the own calibration the improvement was around $0.6°$ for the x-coordinate in condition two. The mean improvement for the foreign calibration ranges from around $4.4°$ (one, x) to $8.5°$ (two, x). So, the differences of the means between the corrected own calibration and the corrected foreign calibration go from $0.2°$ (two, y) to $< 1.4°$ (one, x).

The 1-point correction method shows slightly worse mean errors compared to the multi-point method (not worse than $0.3°$) beside in condition one, x-coordinate, where the error is $\approx 0.8°$ worse. The 0-point method leads to worse mean errors than the multi-point method with differences between $0.4°$ (one, x) and $3.1°$ (two, y).

From the single participants’ results it can be seen that participant 3 produces outlier data with clear higher deviations than the other participants in condition one. This is obvious for the corrected results, where the minimal deviation is more than $6.4°$ compared to the maximal deviation of the means of around $4.5°$ (Table 1). This participant has the worst calibration value of 2 on the Tobii calibration scale. Participant 1 (calibration quality 3) produces specific outliers for the y-coordinate with the foreign calibration in all conditions if the 0-point method is used. In these two cases the corrected values are worse than the uncorrected ones: $9.01°$ to $1.33°$ for condition one and $7.34°$ to $2.34°$ in condition two (Table 2).

6 Discussion

Our results have shown that the use of a foreign calibration corrected by a constant translation leads to acceptable results with an error that is not higher than $1.4°$ compared to the corrected own calibration. The 1-point method can be seen as a real alternative because the additional error relative to the multi-point method is very small and for the worst case only $\approx 0.8°$. For the 0-point method the results are little worse. Generally, the deviation could reduced, beside of subject 1 but the improvements are different between the participants. This indicates that there is at least one other factor which has to be controlled. However, our study shows that with little calibration effort (1-point- and/or 0-point) good results can be achieved, although the basic assumptions are very simple.

Participant	Method	X [°] own	X [°] foreign	Y [°] own	Y [°] foreign
1	uncor	1.32	7.81	0.74	1.33
	multi-point	1.31	2.77	0.76	1.09
	1-point		2.80		1.16
	0-point		3.14		9.01
2	uncor	1.49	6.96	1.07	8.80
	multi-point	1.49	2.15	1.01	1.21
	1-point		3.36		1.87
	0-point		2.27		1.38
3	uncor	5.32	9.81	3.28	10.85
	multi-point	5.28	6.42	3.30	7.39
	1-point		8.77		7.62
	0-point		7.79		8.17
4	uncor	1.26	8.36	0.93	3.46
	multi-point	0.93	2.43	0.50	1.03
	1-point		2.57		1.06
	0-point		2.40		2.25
5	uncor	0.74	5.45	0.52	1.29
	multi-point	0.74	2.72	0.47	1.29
	1-point		2.80		1.41
	0-point		3.18		1.72
mean	uncor	2.02	7.68	1.31	5.15
	multi-point	1.95	3.30	1.21	2.40
	1-point		4.06		2.62
	0-point		3.76		4.51

Table 1: *Absolute deviations between the predicted and the true POG for* own *and* foreign *calibration in the* ***first condition*** *(only eyes should be moved). There are the* uncor*rected errors and the errors for the* multi-point, 1-point *and* 0-point *method.*

Participant	Method	X [°] own	X [°] foreign	Y [°] own	Y [°] foreign
1	uncor	1.88	9.55	0.50	2.34
	multi-point	0.43	0.65	0.43	0.48
	1-point		0.65		0.47
	0-point		0.66		7.34
2	uncor	1.87	10.18	1.21	10.73
	multi-point	0.43	0.62	0.21	0.30
	1-point		0.62		0.29
	0-point		4.27		1.05
3	uncor	0.77	11.85	0.35	9.60
	multi-point	0.71	1.53	0.34	0.92
	1-point		1.57		0.93
	0-point		2.07		4.76
4	uncor	0.79	9.47	0.85	3.50
	multi-point	0.29	0.63	0.28	0.36
	1-point		1.15		0.59
	0-point		1.80		2.29
5	uncor	0.56	6.39	0.64	1.46
	multi-point	0.39	1.15	0.31	0.55
	1-point		1.20		0.77
	0-point		1.23		2.67
mean	uncor	1.17	9.49	0.71	5.53
	multi-point	0.45	0.92	0.32	0.52
	1-point		1.04		0.61
	0-point		2.00		3.62

Table 2: *Absolute deviations between the predicted and the true POG for* own *and* foreign *calibration in the* ***second condition*** *(only head should be moved). There are the* uncor*rected errors and the errors for the* multi-point, 1-point *and* 0-point *method.*

Depending on the research question these errors are passable. Further research could focus on using more complex algorithms and on identifying additional external parameters that influence the translation algorithm. The better adaption results for "head only" can explained by the more or less constant pupil's position relative to the glasses how it was described in Section 3.

One limitation of our study is the rather small number of participants and therefore only little different calibration values. So, we were not allowed to work with more complex algorithms because of the danger of overfitting. Nevertheless our data is not meaningless because we recorded large data sets for each participant (up to $25 points * 150 frames/point = 3750$ data points per condition). A second limitation is that we did not consider all potential error sources, e.g. the point tracking. As mentioned before, the point tracking is based on the IR-markers of which the position could not be detected perfectly by the glasses' camera. This limitation leads to a contamination of the true data with random error. We plan to address both limitations in a future study.

7 Conclusion

In this paper we demonstrated the feasibility of using a foreign calibration from one user and to adapt it to a new user. We proposed a novel approach to calibration adaptation for monocular eye trackers that only uses very simple algorithmic assumptions and in particular does not require any eye model. We presented three methods to compute the adaptation that can be selected depending on the gaze estimation accuracy required by a particular application. Initial results from a user study show a clear improvement in gaze prediction error while requiring less effort than a full 9-point calibration. In combination with the fact that the approach is generic these results are very promising for the development of future calibration-free monocular eye trackers.

References

Dell'Osso, L. F. 1994. Evidence suggesting individual ocular motor control of each eye (muscle). *J Vestib Res 4*, 5, 335–45.

Hansen, D. W., Agustin, J. S., and Villanueva, A. 2010. Homography normalization for robust gaze estimation in uncalibrated setups. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 13–20.

Model, D., and Eizenman, M. 2010. User-calibration-free remote gaze estimation system. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 29–36.

Nagamatsu, T., Sugano, R., Iwamoto, Y., Kamahara, J., and Tanaka, N. 2010. User-calibration-free gaze tracking with estimation of the horizontal angles between the visual and the optical axes of both eyes. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 251–254.

Pillai, K. 1955. Some new test criteria in multivariate analysis. *Annaly of the mathematical statistics 26*, 117.