A Smooth Pursuit Calibration Technique

Conference Paper · March 2014

DOI: 10.1145/2578153.2583042

5 authors, including:

Feridun Mert Celebi
Yale University

Elizabeth S. Kim
Children's Hospital of Philadelphia

Quan Wang
Yale University

Carla Anne Wall
University of South Carolina

A smooth pursuit calibration technique

Feridun M. Celebi[1, 2], Elizabeth S. Kim[1], Quan Wang[1], Carla A. Wall[1], Frederick Shic[1, 2]
[1]Child Study Center, Yale University School of Medicine, New Haven, CT, USA
[2]Yale University Department of Computer Science, New Haven, CT, USA
{feridun.celebi, elizabeth.kim, quan.wang, carla.wall, frederick.shic}@yale.edu

1 Background and Motivation

Many different eye-tracking calibration techniques have been developed [e.g. see Talmi and Liu 1999; Zhu and Ji 2007]. A community standard is a 9-point-sparse calibration that relies on sequential presentation of known scene targets. However, fixating different points has been described as tedious, dull and tiring for the eye [Bulling, Gellersen, Pfeuffer, Turner and Vidal 2013].

As an alternative, some research groups have proposed using smooth pursuit for eye-tracking calibration. E.g., [Bulling, Gellersen, Pfeuffer, Turner and Vidal 2013] tested rectangular constant velocity paths for calibration, and [Blake, Cipolla and Williams 2006] developed a Bayesian-Gaussian regression smooth pursuit approach. These smooth-pursuit approaches can obtain a greater amount of unique pupil-scene mapping information per unit time, potentially increasing the robustness and accuracy of eye-tracking calibration.

2 Problem Statement

We aimed to advance alternative calibration techniques relying on smooth pursuit movements rather than saccades. To this end, we endeavored to preserve the efficiency of calibration in terms of time while increasing the accuracy, reliability and stability of calibration.

In comparison to previous work, our approach differs in two ways: our smooth pursuit pattern and our regression technique. Many different designs would enforce smooth pursuit movements, but a good pattern should be predictable, have good spatial coverage, and little redundancy. Hence, we select an Archimedean spiral trajectory with constant linear velocity (6.4°/sec). In contrast, the corners of a rectangular path [Bulling et al., 2013] may lead to instabilities in pursuit following, and tracing the border alone provides little coverage of the rectangle's interior. As compared to the Bayesian-Gaussian regression technique [Blake, Cipolla and Williams 2006], we utilize a simple-to-implement error corrected regression technique (lag correction and outlier rejection) that achieves good performance. While the Bayesian-Gaussian regression technique is more general, we believe that our approach is highly accessible and effective, which should facilitate the adoption of the technique more broadly.

3 Approach and Method

3.1 Stimuli

Two different calibration techniques are discussed in this paper: a smooth pursuit calibration using an Archimedean spiral with constant linear velocity (6.4 deg/sec), and a standard 9-point-sparse calibration. The accuracy of both techniques is validated by a 7x7 validation grid. Experiments were developed in Windows 7 using MATLAB (Windows, 32-bit) Eyelink Toolbox and data were collected with an SR Research Eyelink 1000 500hz desktop-mounted monocular eye-tracker. Smooth pursuit calibration took 27 seconds, during which up to 1600 data points were recorded. Presentation of scene targets in 9-point calibration (took 23 seconds) and validation was randomized. In order to examine comparable durations in smooth pursuit as compared to sparse calibration, we also examined truncating the tail of our smooth pursuit data to 1350 data points.

3.2 Calibration Error

Our MATLAB script produces two different matrices for each run of smooth pursuit, 9-point-sparse calibration and 49-point validation. One of these matrices includes the actual x and y coordinates of the points presented on the screen in pixels. The other matrix includes x and y coordinates of the user's point of regard (POR) on the screen in pixels. A quadratic regression is applied to these two matrices to produce a transformation matrix that is used to project eye-tracking data to actual screen locations. This transformation matrix is then applied to the eye-position matrix of 49-point validation to create an expected position matrix. This transformation matrix acts as a mapping function that converts the eye-tracker data to screen coordinates [Stampe 1993]. This expected position matrix is compared to the actual coordinates matrix of the 49-point validation, and the errors represented in pixels are converted to degrees. These steps are applied to determine the final calibration error for corresponding calibration techniques. Finally, all error matrices go through the error correction routine that includes outlier rejection and lag correction. The specific lag between the smooth pursuit motion and the true position of the presented target is detected and corrected by minimization of the fit error under multiple time lags (Lag_Points={0..20}, i.e. 0 to 338 ms) in the smooth pursuit reported values and the outliers are removed using a simple fit residual rejection criterion (points above 95th percentile error removed) applied 3 times.

3.3 Protocol

Ten healthy adults (n = 10, 7 females, 3 males; ages 20 to 39 years) participated in the experiment. Four participants had corrected vision (glasses or contact lenses). Initially, a standard 9-point SR Research calibration routine was used to calibrate the system and assure the flow of reliable eye-tracking data. For each participant, it was assured that the error in the system calibration was at most 1 degree. Then, the stimuli presentation was initiated. Two calibration techniques were presented in alternating order, each followed by a 7x7 validation grid that was used to produce a calibration error measured

ETRA 2014, March 26 – 28, 2014, Safety Harbor, Florida, USA.

ACM 978-1-4503-2751-0/14/03

in degrees for each technique. The alternation was repeated four times. Total experiment run-time was 9 minutes.

4 Results to Date

After applying our fitting routines, the rms error values were determined in the horizontal (x) and vertical (y) axes as well as under a Euclidean norm (combined error) for the validation grid under smooth pursuit calibration and 9-point-sparse calibration. The results of these calculations are presented in Table 1. All errors are reported in degrees. The errors are separated into three different categories: mean errors and standard deviations in x-coordinates, y-coordinates, and combined Euclidean (x-y) error and standard deviation.

Trial Type / Coordinate Space	Smooth Pursuit (Not Truncated)	Smooth Pursuit (Truncated)	Sparse
	Mean (Standard Deviation) - Degree (Degree)		
X	0.429 (0.153)	0.452 (0.168)	0.679 (0.641)
Y	0.611 (0.303)	0.698 (0.307)	1.062 (0.787)
Combined	0.838 (0.278)	0.913 (0.272)	1.388 (0.963)

Table 1. Root Mean Square (RMS) Calibration Errors associated with Different Calibration Techniques

To examine differences in calibration quality between our three calibration methods (smooth pursuit, truncated smooth pursuit, and 9-point calibration), w conducted a linear mixed model analysis with factorial fixed effects of calibration method and trial number and participant intercept random effects. For combined error a main effect of calibration method was found $F(2,18.0) = 3.7$, p=.04, with significant differences observed between 9-point and both smooth (p=.02) and truncated smooth (p=.04) calibration methods. No significant trial or calibration method x trial interactions were observed. Differences in combined error between 9-point calibration and other methods were driven by y-axis error, where similar omnibus results were observed, $F(2,19.4) = 3.8$, p=.04. By contrast, these effects were not observed along the x-axis, $F(2,19.0) = 1.6$, p=.22.

5 Current Status

The goal of the primary author (FC, a 2nd year undergraduate researcher) is to improve calibration techniques for use in gaze-adaptive experiments in toddlers and children with and without developmental disabilities. This work stems from a need to maintain accurate calibration over a larger spatial extent than traditional 5 point calibration techniques can typically provide (with 9 point calibration being even less feasible in our target population). However, the study team is interested in calibration research in its own right, and is currently seeking advice as to how to both improve the methods described currently as well as to develop the research along more sophisticated lines that would be of greater interest to the eye-tracking research community.

6 Conclusions and Future Directions

Given the potential trade-offs of smooth pursuit calibration over community standard 9-point-sparse calibration, the choice of the calibration routine should be made based on the needs of the experimenter. Although collecting up to 1600 data points provides greater flexibility in correcting distortions in the eye-tracking data, it might also increase the spatial noise [Stampe 1993]. As evidenced by the data provided in Table 1, we see that smooth pursuit calibration delivers several improvements to the accuracy and stability of the data, indexed by reported mean errors and standard deviations. Given the increased need for stability in the use of desktop-mounted systems, such improvements become even more crucial as participants are allowed to move their heads more freely.

Whether smooth pursuit calibration is universally better than sparse calibration or not is an important issue to be discussed. It is a limitation on our part that our experiment uses a specific eye tracker (SR Research Eyelink 1000). However, current 9-point-sparse calibration routines are also run on similar systems. So, we argue that if the smooth pursuit calibration works better on our particular setup, it is likely that it will work better on other setups. Theoretically, smooth pursuit calibration should be universally better, as long as the participant can maintain the smooth pursuit movement.

However, it is likely that, as with typical point-based calibration, not all participants can maintain smooth pursuit movements. In our study, participants were 10 healthy adults. It remains to be seen how well smooth pursuit calibration would perform against 9-point-sparse calibration in broader populations. Given that adults may be more willing to cooperate as opposed to children, it would be interesting to see whether smooth pursuit calibration routine might reproduce its stability and accuracy in an experiment with children. In addition, individuals with oculomotor problems and infants might have difficulty maintaining the smooth pursuit movement. Our results thus far use an elementary model of error corrected regression. More robust fitting algorithms might be implemented to compensate for the inability to track smoothly (as seen in infants) or loss of concentration during smooth pursuits. Thus, improvements to the algorithm that handles the regression are a major area of future work.

Finally, diverging from the calibration aspect of the project, examining the properties of smooth pursuit in subjects participating in experimental studies may provide additional insight about those participants. For example, studies have shown smooth pursuit properties such as gain and lag are associated with neuropsychiatric conditions [Ross, Olincy, Harris, Sulliva, and Radant, 2000] and show robust psychometric properties likely related to attentional control [Maruta, Heaton, Kryskow, Maule, and Ghajar, 2013]. Examining the interplay among smooth pursuit properties, individual differences, and systematic deviations from POR expectations may provide a new and exciting area for future clinical and technical advancement.

References

BLAKE A., CIPOLLA R., WILLIAMS O. 2006. *Sparse and Semi-supervised Visual Mapping with the S³GP*. CVPR.

BULLING A., GELLERSEN H., PFEUFFER K., TURNER J. VIDAL M. 2013. *Pursuit calibration: making gaze calibration less tedious and more flexible.* ACM UIST.

CORNELISSEN, F.W., PETERS. E. AND PALMER, J. 2002. *The Eyelink Toolbox: Eye Tracking With MATLAB And The Psychophysics Toolbox*. Behavior Research Methods, Instruments & Computers.

MARUTA J., HEATON K.J., KRYSKOW E.M., MAULE A.L., AND GHAJAR J. 2013. *Dynamic visuomotor synchronization: Quantification of predictive timing.* Behavior Research, 45, 1, pp. 289-300.

ROSS R. G., OLINCY A., HARRIS J. G., SULLIVAN B., AND RADANT A. 2000. *Smooth pursuit eye movements in schizophrenia and attentional dysfunction: adults with schizophrenia, ADHD, and a normal comparison group.* Biological Psychiatry, 48, 3, pp. 197–203.

STAMPE, DAVE M. .1993. *Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems*. Behavior Research Methods, Instruments, & Computers Volume 25, Issue 2, pp. 137-142.

TALMI K., LIU J. 1999. *Eye and gaze tracking for visually controlled interactive stereoscopic displays*. Signal Processing: Image Communication, 14, pp. 799-810.

ZHU ZHIWEI AND JI QIANG. 2007. *Novel Eye Gaze Tracking Techniques Under Natural Head Movement*. IEEE Transactions on Biomedical Engineering, 54, 12, pp. 2246-60.