

Easy Post-Hoc Spatial Recalibration of Eye Tracking Data

Yunfeng Zhang and Anthony J. Hornof
Computer and Information Science
University of Oregon
Eugene, OR 97403 USA
{zywind, hornof}@cs.uoregon.edu

Abstract

The gaze locations reported by eye trackers often contain error resulting from a variety of sources. Such error is of increasing concern to eye tracking researchers, and several techniques have been introduced to clean up the error. These methods, however, either compensate only for error caused by a particular source (such as pupil dilation) or require the error to be somewhat constant across space and time. This paper introduces a method that is applicable to error generated from a variety of sources and that is resilient to the change in error across the display. A study shows that, at least in some cases, although the change in error across the display appears to be random it in fact follows a consistent pattern which can be modeled using quadratic equations. The parameters of these equations can be estimated using linear regression on the error vectors between recorded fixations and possible target locations. The resulting equations can then be used to clean up the error. This regression-based approach is much easier to apply than some of the previously published methods. The method is applied to the data of a visual search experiment, and the results show that the regression-based error correction works very well.

CR Categories: H.5.2 [User Interfaces]: Theory and Methods;

Keywords: eye tracking, error correction.

1 Introduction

The last decade has seen an increasing number of experiments that use eye movement data to study human performance. These experiments require accurate eye tracking data. The dominant eye tracking systems used today are video-based pupil-center corneal-glint-reflection eye trackers, and such eye trackers can have varying degrees of error depending on the properties of the hardware (cameras resolution, sampling frequency, etc.) as well as the conditions in which the data are recorded (e.g., stimuli luminance and head movements) [Holmqvist et al., 2011]. For scientific studies, even a small amount of eye tracking error can have a detrimental effect. For example, Holmqvist et al. [2012] showed that even a 0.5° offset of gaze locations can lead to very different conclusions about dwell time on areas of interest (AOI). Running a rigorous eye tracking experiment requires efforts to minimize error.

Several methods have been proposed to correct error in video-based eye tracking systems. These methods can be roughly classified into two categories: source-centric and data-centric. The source-centric methods deal with error generated by a particular cause such as head movements or changes in pupil size. Cerrolaza et al. [2012], for example, showed that head movements, particularly movements in depth (perpendicular to the screen), can cause error in gaze location estimates. The authors proposed a new calibration procedure and a new gaze estimation function (which translates the pupil-center corneal-glint vector into x and y screen coordinates) that incorporates the eye-to-screen distance to compensate for the error. The new procedure, however, requires lengthier calibration because it repeats traditional calibration two more times, one at a shorter eye-to-screen distance and one at a longer distance. Another example of a source-centric error correction method is proposed by Drewes et al. [2012] to address error caused by changes in pupil size. The method is very similar in spirit to Cerrolaza et al. [2012]: A two-pass calibration is needed, one using a dark background (causing pupil dilation), and one using a bright background (causing pupil constriction). Then, a new gaze estimation function that incorporates pupil size is used to generate better gaze location estimates. Both methods were shown to work very well for error that is solely caused by one particular source. However, it might be impractical to combine the two methods to reduce both types of error because a straightforward combination would require a six-pass calibration. Clearly, the field needs better ways for dealing with error from a variety of sources.

The other category of error correction methods, data-centric methods, works by extrapolating the pattern of the error from the data and then removing the error based on the extracted pattern. These are typically post hoc methods in that they do not require modifications to the calibration procedure or the gaze estimation function, but only need to process the collected eye tracking data. Though in order to identify the error the experimental design may need to incorporate some *required fixation locations* (RFLs) [Hornof and Halverson, 2002]. These are points on the screen that the analyst can be relatively certain that a participant fixated at a specific point in time, provided that the participant completed the trial accurately. For example, when moving a mouse cursor to click on a button, the participant normally needs to fixate the button to ensure an accurate click; this particular behavior can be further reinforced by implementing a *point-completion deadline* [Hornof, 2001], which financially rewards a participant if the click on the button occurs within a short period of time after the mouse cursor starts to move. With RFLs, the eye tracking error can be identified as the disparity between the RFLs and the fixations that were recorded at the time that the RFLs were fixated.

One problem with previous data-centric error correction methods that use RFLs is that these methods assume that the eye tracking error remains constant across time and locations on the display. This assumption, however, does not always hold. Instead, the

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ETRA 2014, March 26 – 28, 2014, Safety Harbor, Florida, USA.
Copyright © ACM 978-1-4503-2751-0/14/03 \$15.00

error can vary substantially across the display, such as perhaps around the edges of the screen when the display area approaches the maximum tracking zone of the eye tracker. This paper presents a data-centric error correction method that recalibrates eye tracking data by identifying how the error changes across the display. The method uses both RFLs and a relaxed version of RFLs—probable fixation locations (PFLs). The method does not require additional calibrations with the participant and uses simple linear regression to correct the data.

2 Changes in Error Across the Display

Figure 1 shows how error can change across the display. The data were collected from one session of our replication of the Williams [1966] visual search experiment, with a point-completion deadline enforced to obtain an accurate measurement of the eye tracking error. The circles in Figure 1 represent the target location from every trial in a session, and the vectors, which are used to assess the eye tracking error, connect each target to the fixation that was recorded when the target was clicked. Targets and distractors were randomly distributed across the display, which was useful for assessing the error. As can be seen in Figure 1, the error varied in direction and magnitude across the display. In the top half of the display, the error tended to shift the recorded fixation downward and was about 1° (equal to 38.7 pixels). In the bottom half, the error was much smaller and tended to shift the recorded fixation upward.

The error shown in Figure 1 was not likely caused by improper or careless setup of the eye tracking equipment. Many steps, discussed in more detail in Section 4, were taken to reduce error in our eye tracking data. The only aspect of how the eye tracker was configured that may have contributed to the error was that it was used with a display that was 40° wide, which is right at the limit of the maximum tracking zone for the eye tracker (a 120 Hz binocular LC Technologies Eyegaze Analysis System). Our average error of 0.74° across all sessions is in line with the error recorded in other experiments for which eye tracking error was reported with other pupil-center corneal-glint eye trackers (e.g., [Cerrolaza et al., 2012]).

One approach to addressing how eye tracking error changes across a display might be to calculate the average error separately

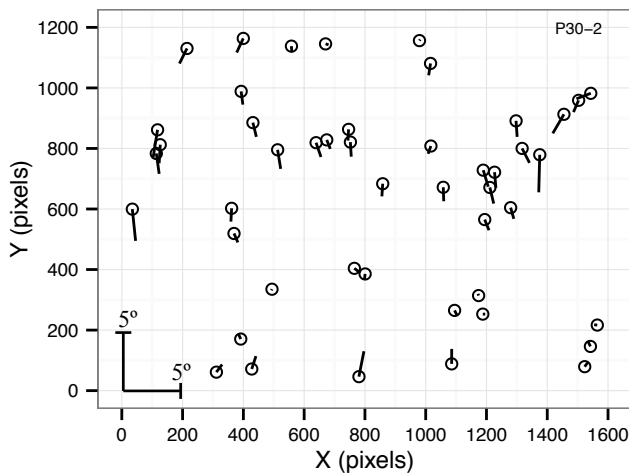


Figure 1: The error vectors between the search targets (circles) and the fixations reported by the eye tracker when clicking on each target, from one session of the experiment. 1° of visual angle = 38.7 pixels

for the top half and for the bottom half of the screen based on the RFL error vectors. This solution, however, would be problematic because it would require inspecting visualizations such as Figure 1 for every session of the experiment for every participant to determine how the screen should be best divided. This specific approach would also be susceptible to noise (random variation) in the error because as the screen is divided into more regions, there are fewer error vectors in each region, and thus the noise would be less likely to be reduced by averaging all of the error vectors.

Figure 2 provides another view into how the error changes across locations for the same set of data as in Figure 1. The top panel shows how the vertical component of the error changes in relation to the fixations' x coordinate, and the bottom panel how the vertical component changes in relation to the fixations' y coordinate. (The horizontal component of the error has similar but less apparent trends, and is not shown here.) The solid curves are the best-fitting quadratic curves for the data. The graphs show that there is a gradual transition in the vertical component of the error from left to right (the x coordinate), and from top to bottom (the y coordinate). Further, this transition seems to be largely accounted for by the quadratic curves. These results indicate that the change in error is not completely random, but that it follows a smooth trend that can be characterized mathematically. Thus, there is no need to discretize the changes using a region-based RFL technique such as proposed earlier. Instead, we can simply find the equations that most accurately describe the changes across the entire display.

3 Error Correction by Spatial Recalibration

Based on Figure 2, we postulate that the true gaze locations and

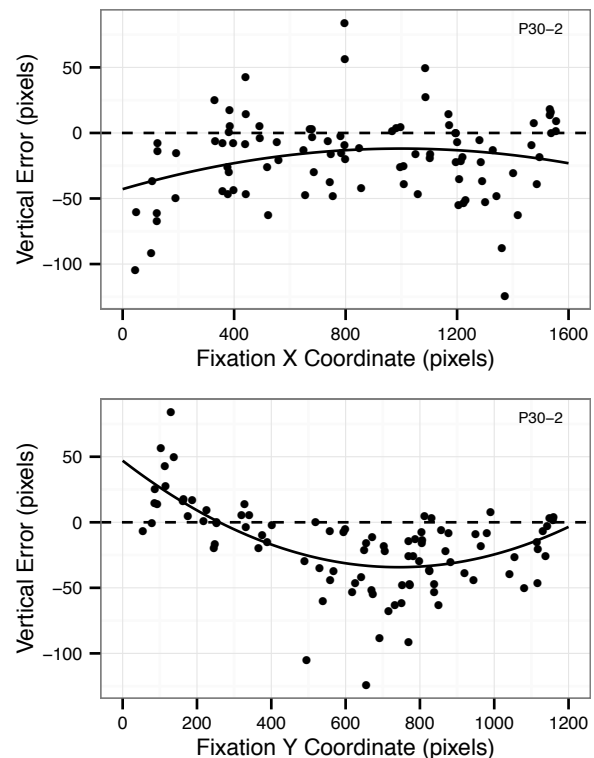


Figure 2: The vertical component of the error vectors shown in Figure 1 as a function of each fixation's reported x coordinate (the top panel) and y coordinate (the bottom panel). The curves are the best fitting quadratic functions.

the reported gaze locations have the following relations:

$$\begin{aligned}x' &= a_0 + a_1x + a_2x^2 + a_3y + a_4y^2 + a_5xy \\y' &= b_0 + b_1x + b_2x^2 + b_3y + b_4y^2 + b_5xy\end{aligned}$$

where x' and y' are the true gaze coordinates, and x and y are the reported coordinates. The terms a_0, \dots, a_5 and b_0, \dots, b_5 are coefficients that might change for each calibration, and need to be determined from the data. This equation not only has the quadratic components needed to generate the fitting curves in Figure 2, but it also has the interaction component, xy , for capturing changes that depend on both coordinates. We call this the “recalibration” equation because it essentially compensates for the loss of accuracy inherent in the original calibration.

To determine the coefficients of the recalibration equation, straightforward linear regression can fit the equation to the fixation-to-RFL error vectors such as those shown in Figure 1. The RFLs serve as the true gaze coordinates, and the fixation-to-RFL vectors are used to parameterize the equation to correct fixations at different screen coordinates. Provided that there are a sufficient number of RFLs distributed across the display, the method will find appropriate parameters to reliably remove the error.

When the experimental design does not provide enough RFLs to cover the entire display, *probable* rather than *required* fixation locations can be used to provide more data for fitting the recalibration equation parameters. Probable fixation locations (PFLs) are visual objects that need to be fixated in order to successfully complete a trial, but unlike RFLs, the precise time when each PFL is looked at cannot be inferred from the experimental design or participant performance. For example, small text that needs to be read at an arbitrary point in a task could be used as a PFL because the participant would need to look directly at the text to recognize it (whereas other visual properties such as color could be recognized in the periphery), but the participant could do so at various different times in a trial, which would make it difficult to determine which exact fixation should be mapped to the PFL. Zhang and Hornof’s [2011] solution to this problem was to map each fixation to its nearest PFL. The authors showed that with a noise-resilient algorithm (in their case, the somewhat complicated mean-shift algorithm), the method will reliably identify the pattern of error provided that the RFLs are not arranged in an evenly-spaced grid pattern and that the error is constant across the display. This study extends the use of PFLs to identify eye tracking error that changes across the display.

For the visual search experiment presented here, a combination of RFLs and PFLs were used to estimate the equation parameters. To minimize the effect of potentially incorrect fixation-to-PFL and fixations-to-RFL error vectors, we used the *robust* linear regression method [Hampel et al., 1986], which has been shown to be resilient to outliers when used to estimate equation parameters. In addition, to reflect their different levels of reliability, the RFL error vectors were weighted ten times higher than the PFL error vectors when submitted to regression.

4 Validation of the Error Correction Method

This post hoc regression-based spatial recalibration method was applied to the eye movement data of our visual search experiment. Figure 3 shows a sample search field used in the experiment, with fixation scanpaths superimposed onto the display. The task was to search for a target within a grid of 75 objects that have different colors, shapes, and sizes. Each object had a unique two-digit

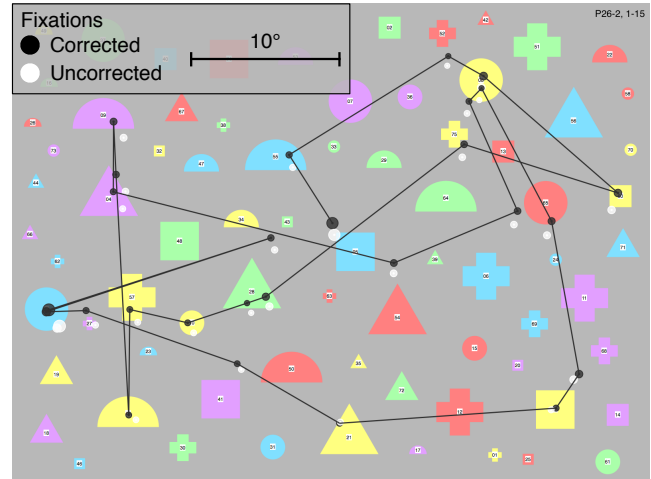


Figure 3: An eye movement data visualization of one trial. White circles represent the uncorrected fixations, and black the corrected fixations. Larger circles indicate longer fixations. The participant was looking for a large circle with number 3.

number in the center. The search precues were shown before each trial and included the number of the target object and, depending on the experimental condition, some combinations of the target’s color, size, and shape. Twenty-two participants completed the experiment. Each participant was given a total of 96 trials in two sessions.

A binocular 120 Hz LC Technologies Eyegaze Analysis System was used to collect the eye movement data. To reduce eye tracking error from head movements, a chinrest was used. The eye tracker was calibrated at the beginning of each of the two 10-minute long sessions. To reduce changes in pupil diameter due to changes in brightness, the calibration screen used the same background as the search task. The raw gaze samples were grouped into fixations using a dispersion-based algorithm with a maximum dispersion window size of 0.7° and a minimum fixation duration of 60 ms.

As mentioned earlier, the error correction for this experiment used a combination of RFLs and PFLs to find the appropriate parameters for the recalibration equation. The RFLs were the search targets, and were mapped to the fixations that occurred when the target was clicked. The PFLs were the object numbers in the experimental condition in which no features other than the target’s two-digit number were provided in the precue. Because the number was small (0.26° in height), the participants had to look directly at the numbers to find the target. Similar to Zhang and Hornof [2011], fixations in this condition were mapped to their nearest PFLs. As mentioned earlier, the RFLs were given a larger weight than the PFLs in the *robust* linear regression computation.

4.1 Results

The post hoc spatial recalibration cleaned up the error very well. Before error correction, 43% of the RFL-associated fixations were within 0.5° of the target, 81% were within 1.0° , and the mean deviation was 0.74° ($SD = 1.0^\circ$). After error correction, 77% of the RFL-associated fixations were within 0.5° of the target, 96% were within 1.0° , and the mean deviation was 0.45° ($SD = 0.9^\circ$). Figure 4 shows the vertical component of the error (measured using the

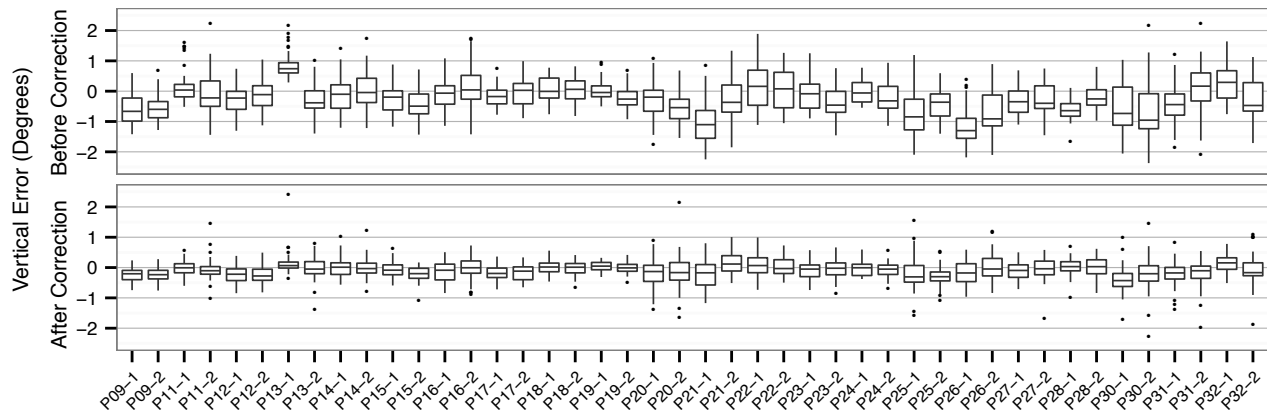


Figure 4: Distribution of the vertical component of the eye tracking error across sessions, before and after error correction.

RFLs) before and after error correction. After error correction, the medians were better aligned at zero and the error distributions were smaller.

The scanpath in Figure 3 shows the results of applying the error correction technique, including how the technique can model different error for different regions of the display. The visualization shows the corrected fixations as black circles and the uncorrected fixations as white circles. Note that the black circles are much closer to the objects than the white circles. Many of the uncorrected (white) fixations were between objects, but after correction these fixations were at much more plausible locations such as on the object labels. The visualization shows how the degree of correction differs for different screen locations. For example, the correction for the fixations in the top half of the display are generally larger than the corrections for the fixations in the bottom half. This illustrates how the quadratic recalibration equation can adjust differentially across the display.

5 Conclusion

This paper presents a post hoc error correction technique for cleaning up eye tracking error. The technique uses RFLs [Hornof and Halverson, 2002] and PFLs [Zhang and Hornof, 2011] to measure eye tracking error, and uses quadratic equations to characterize the change in error across the display. Example data from an eye tracking experiment demonstrate that eye tracking error that initially appears to be somewhat random can actually follow a smooth quadratic curve. Though the technique presented here was developed with data from one particular eye tracker, the LC Technologies Eyegaze system, a similar quadratic error trend was also reported in Cerrolaza et al. [2012] with eye movement data collected with a different video-based eye tracker, which suggests that this pattern of error and the efficacy of our technique will persist across video-based systems.

The advantages of the error correction method presented here over previously published methods are threefold: (1) The method does not complicate the calibration procedure. Experimenters often experience frustration with eye tracking calibration because, for many video-based cameras, the conditions (e.g., lighting, camera focus, and eye-to-camera distance) have to be just right to successfully calibrate a participant. It is somewhat impractical to extend the calibration procedure as required by error correction methods such as Cerrolaza et al. [2012] and Drewes et al. [2012]. (2) The current method can accurately and automatically deal with

changes in error across the display, which improves on previous post hoc error correction methods that require analytic decisions by the analysts. (3) The method is easy to apply because it uses linear regression, which is readily available in many software packages, unlike more sophisticated machine-learning algorithms used by other methods such as John et al. [2012].

References

- CERROLAZA, J. J., VILLANUEVA, A., VILLANUEVA, M., AND CABEZA, R. 2012. Error characterization and compensation in eye tracking systems. In *ETRA '12*, 205–208.
- DREWES, J., MASSON, G. S., AND MONTAGNINI, A. 2012. Shifts in reported gaze position due to changes in pupil size: ground truth and compensation. In *ETRA '12*, 209–212.
- HAMPEL, F. R., RONCHETTI, E. M., ROUSSEEUW, P. J., AND STAHEL, W. A. 1986. *Robust statistics: the approach based on influence functions*. Wiley.
- HOLMQVIST, K., NYSTROM, M., AND MULVEY, F. 2012. Eye tracker data quality: What it is and how to measure it. In *ETRA '12*, 45–52.
- HOLMQVIST, K., NYSTRÖM, M., ANDERSSON, R., DEWHURST, R., JARODZKA, H., AND VAN DE WEIJER, J. 2011. *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- HORNOF, A. J. 2001. Visual search and mouse-pointing in labeled versus unlabeled two-dimensional visual hierarchies. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 8(3), 171–197.
- HORNOF, A. AND HALVERSON, T. 2002. Cleaning up systematic error in eye-tracking data by using required fixation locations. *Behavior Research Methods*, 34, 592–604.
- JOHN, S., WEITNAUER, E., AND KOESLING, H. 2012. Entropy-based correction of eye tracking data for static scenes. In *ETRA '12*, 297–300.
- WILLIAMS, L. G. 1966. The effect of target specification on objects fixated during visual search. *Perception & Psychophysics*, 1, 315–318.
- ZHANG, Y. AND HORNOF, A. J. 2011. Mode-of-disparities error correction of eye-tracking data. *Behavior Research Methods*, 43(3), 834–842.