

# UC Santa Cruz

## UC Santa Cruz Previously Published Works

### Title

A Functional Usability Analysis of Appearance-Based Gaze Tracking for Accessibility

### Permalink

<https://escholarship.org/uc/item/2pf3t7p7>

### Authors

Park, Youn Soo

Manduchi, Roberto

### Publication Date

2024-04-29

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# A Functional Usability Analysis of Appearance-Based Gaze Tracking for Accessibility

Youn Soo Park  
University of California, Santa Cruz  
Santa Cruz, CA, USA  
ypark43@ucsc.edu

Roberto Manduchi  
University of California, Santa Cruz  
Santa Cruz, CA, USA  
manduchi@ucsc.edu

## ABSTRACT

Appearance-based gaze tracking algorithms, which compute gaze direction from user face images, are an attractive alternative to infrared-based external devices. Their accuracy has greatly benefited by using powerful machine-learning techniques. The performance of appearance-based algorithms is normally evaluated on standard benchmarks typically involving users fixating at points on the screen. However, these metrics do not easily translate into functional usability characteristics. In this work, we evaluate a state-of-the-art algorithm, FAZE, in a number of tasks of interest to the human-computer interaction community. Specifically, we study how gaze measured by FAZE could be used for dwell-based selection and reading progression (line identification and progression along a line) – key functionalities for users facing motor and visual impairments. We compared the gaze data quality from 7 participants using FAZE against that from an infrared tracker (Tobii Pro Spark). Our analysis highlights the usability of appearance-based gaze tracking for such applications.

## CCS CONCEPTS

• **Computing methodologies** → **Tracking; Computer vision.**

## KEYWORDS

Gaze Estimation, Text Reading, Dwelling

### ACM Reference Format:

Youn Soo Park and Roberto Manduchi. 2024. A Functional Usability Analysis of Appearance-Based Gaze Tracking for Accessibility. In *2024 Symposium on Eye Tracking Research and Applications (ETRA '24)*, June 4–7, 2024, Glasgow, United Kingdom. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3649902.3656363>

## 1 INTRODUCTION

Eye gaze tracking has been used extensively as a human-computer interface modality (e.g. pointer control [Drewes et al. 2007; Sibert and Jacob 2000], magnification control [Ashmore et al. 2005; Manduchi and Chung 2022]), to measure the user’s attention (e.g. when driving a vehicle [Vicente et al. 2015] or visiting a web site [Pan et al. 2004]), to study reading behaviors [Rajendran et al. 2018; Vo et al. 2010], and to identify specific conditions such as autism

spectrum disorders [Murias et al. 2018], ADHD [De Silva et al. 2019], or dyslexia [Raatikainen et al. 2021; Rayner 1998; Wang et al. 2024]. Measurements of the user’s gaze point (point of regard on the screen) are usually obtained through an external device (a gaze tracker) that uses an infrared illuminator and one or more cameras to compute the visual axis [Guestrin and Eizenman 2006] (the line from the point of regard to the center of the fovea through the pupil). Modern commercial gaze trackers can be rather accurate (with errors of a fraction of a degree) while allowing users to move their heads within a certain volume of space [Tobii [n. d.]].

In recent years, there has been increasing interest in software systems that leverage modern machine learning to estimate a person’s gaze direction from an image of their face, taken e.g. from a screen camera. The practical advantages of “appearance-based” tracking software are apparent, both in terms of convenience (no need for an external device to connect) and cost (infrared-based trackers are still quite expensive). However, the accuracy of appearance-based trackers still lags behind that of infrared trackers [Zhang et al. 2019]. This article presents a functional usability analysis of a state-of-the-art appearance-based tracker. While the performance of gaze trackers is normally expressed in quantities such as angular errors, typically computed in specific settings (e.g. with users looking at a target on the screen), these quantities do not easily translate into desired usability parameters. Therefore, we investigate whether state-of-the-art appearance-based trackers can serve as potential substitutes for infrared-based systems, especially in the following applications tailored for users with disabilities:

*Dwell-Based Selection.* This is a standard technique for users who are unable to trigger a click event using a mouse or a switch [Jacob 1991; Müller-Tomfelde 2007; Paulus and Remijn 2021; Sibert and Jacob 2000; Zhang and MacKenzie 2007]. While other approaches have been considered (e.g., blink-based [Huckauf and Urbina 2008; Lu et al. 2020]), dwell-based selection remains a popular choice and is implemented in commercial devices such as the Tobii Dynavox communication system [Menges et al. 2019] enhancing accessibility for those with physical limitations.

*Reading Progression Tracking.* Measuring progression when reading a document can be useful to assess one’s cognitive skills of reading [Huck 2016; Patterson and Ralph 1999] or to provide gaze-contingent reading support (e.g., highlighting the line currently being read [Rosenberg 2008], controlling the speed of auto-scrolling [Kumar et al. 2007; Sharmin et al. 2013] or of text-to-speech [Schiavo et al. 2015], magnifying the text being gazed at [Ashmore et al. 2005; Manduchi and Chung 2022; Maus et al. 2020], or detecting reading difficulties and augmenting text [Biedert et al. 2009; Bottos and Balasingam 2020; Lunte and Boll 2020]), aiding those with dyslexia or low vision [Wang et al. 2024]. We are interested in *reading line*

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ETRA '24, June 4–7, 2024, Glasgow, United Kingdom

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0607-3/24/06

<https://doi.org/10.1145/3649902.3656363>

*identification* (detecting which text line in the document is currently being read [Bottos and Balasingam 2020; Sun and Balasingam 2021; Wang et al. 2024]) as well as in tracking *progression along a line* by measuring fixation scanpaths [Deng et al. 2023; Reichle et al. 2003].

We selected FAZE [Park et al. 2019] as our reference appearance-based gaze tracking algorithm (described in Sec. 3.1.3), showing to achieve accuracy of about  $3^\circ$  on multiple standard benchmark data sets. One important feature of FAZE is that it adapts to the appearance characteristics of a new user from just a few calibration images. An open-source implementation of FAZE was made available by the authors<sup>1</sup>. On a Lambda Tensorbook, FAZE produces gaze data at a rate of 6 fps.

In order to evaluate the feasibility of FAZE for the considered applications, we conducted a small study with 7 participants, who underwent two tasks: a fixation task (representative of dwell-based selection), and a reading task. Images of the participants during these tasks were taken by a computer camera. In addition, we used an infrared-based gaze tracker (Tobii Pro Spark) to capture their gaze direction. The Tobii tracker is used as a reference against which to compare FAZE data. We define specific metrics for each task, and evaluate FAZE and Tobii data comparatively against these metrics. Our results give a detailed picture of the type of errors that can be expected when using FAZE, providing insights for designers integrating appearance-based gaze tracking in applications for individuals with disabilities.

## 2 RELATED WORK

Hohlfeld et al. [Hohlfeld et al. 2015] presented an analysis of the applicability of computer vision-based gaze tracking for mobile scenarios that is germane to our work. Here are the main differences between this contribution and [Hohlfeld et al. 2015]. *1. Appearance-based algorithm:* Hohlfeld et al. used EyeTab [Wood and Bulling 2014], a model-based tracker whose accuracy (errors of  $7^\circ$  in ideal conditions) is substantially inferior to learning-based algorithms such as FAZE; *2. Tasks set:* the following tasks were considered in [Hohlfeld et al. 2015]: Focus on Device (determining whether the user was looking at a tablet computer or behind it); Line Progression: Line Test (finding regressions when following a moving dot); Word Fixation: Point Test (finding fixation times). Our tasks (dwell-based selection, reading line identification, progression along a line) are substantially different than those in [Hohlfeld et al. 2015]. *3. Infrared gaze tracker as reference.* We use a commercial-grade infrared gaze tracker to produce a reliable baseline against which to compare the data from appearance-based tracking. Comparison between the two trackers is important to establish whether an appearance-based tracker can substitute for an infrared-based tracker, which is the main research question motivating our work.

Zhang et al. [Zhang et al. 2019] presented a comparative evaluation of two appearance-based gaze tracking algorithms (MPI-IFaceGaze [Zhang et al. 2017] and GazeML [Park et al. 2018]) against a consumer-grade infrared-based device (Tobii EyeX). This work was concerned with the range of viewing distances for which gaze could be reliably computed, the required number of calibration samples, the systems' robustness to varying illumination (indoor vs. outdoor), and their ability to measure gaze for users wearing

glasses. While very valuable, these tasks are very different from the tasks considered in our contribution.

Wang et al. [Wang et al. 2024] developed GazePrompt to improve digital reading for low-vision users by providing line-switching and difficult-word recognition features, utilizing an infrared-based tracker. This innovation highlights the necessity of investigating appearance-based gaze tracking as a means to enhance usability and accessibility. Such exploration could lead to significant advancements in assistive reading technologies.

## 3 METHOD

### 3.1 Apparatus

*3.1.1 Computer.* We used a Lambda Tensorbook (equipped with an NVIDIA RTX 2080 GPU and 8-core Intel i7-10875H at 2.30 GHz, running Ubuntu 20.04.6) for our tests. The screen size (active pixel area) was 349 mm by 195 mm, for a resolution of 1920 by 1080 pixels. A 1080p webcam was located on the top edge of the screen.

*3.1.2 Infrared Gaze Tracker.* We used a Tobii Pro Spark gaze tracker for baseline measurements. This is a moderately priced model that produces binocular measurements at 60 Hz. In ideal conditions, its nominal accuracy (mean angular error) is of  $0.45^\circ$ , while its precision (standard deviation of the error) is of  $0.26^\circ$  [Tobii [n. d.]]. For a person looking at the TensorBook's screen from a distance of 500 mm, these values translate to 20 and 11 pixels, respectively. The tracker can measure gaze from a user located between 450 mm and 950 mm from the screen, with a nominal freedom of head movement of  $350 \times 350$  mm. The tracker was placed at the bottom of the TensorBook's screen and was calibrated for each participant using the Tobii Pro Eye Tracker Manager utility (9 targets).

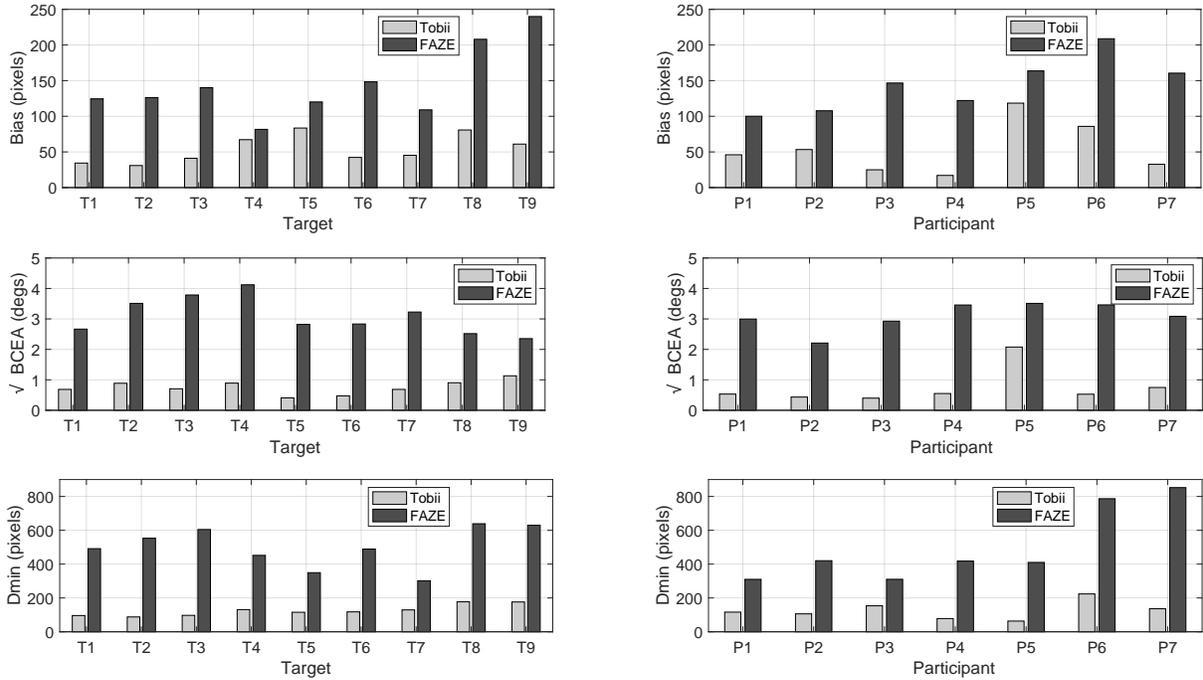
*3.1.3 Appearance-Based Gaze Detection.* FAZE (Few-shot Adaptive GaZE Estimation) is a state-of-the-art appearance-based gaze tracking algorithm. It incorporates several few-shot learning paradigms, most notably Model-Agnostic Meta-Learning (MAML). At the core of FAZE is an encoder-decoder architecture that captures latent representations related to appearance, gaze direction, and head pose from eye region imagery. After the initial learning phase of these latent features, FAZE undergoes fine-tuning with a minimal set of calibration samples from individual users. The use of MAML significantly reduces over-fitting, thereby facilitating rapid and person-specific model fine-tuning. The average angular error of FAZE is of  $3.14^\circ$  [Park et al. 2019].

In our tests, we noted that data from FAZE sometimes exhibits a consistent location bias, even after calibration. To remedy this, we considered an additional geometric calibration. Specifically, for each participant, we recorded the barycenter of the gaze points produced by FAZE while the participant fixated each of the 9 points in a pattern (Sec. 3.3), then regressed the parameters of an affine transform minimizing the squared norm of the location error. This affine transform was then applied on the gaze points returned by FAZE for that participant.

### 3.2 Population

We recruited 7 participants (3 female, 4 male; age min: 22; max: 58; mean: 33.7) for this test. Three participants (P5, P6, P7) wore glasses during the test. The study was conducted following a Human

<sup>1</sup>[https://github.com/NVlabs/few\\_shot\\_gaze](https://github.com/NVlabs/few_shot_gaze)



**Figure 1: Dwelling experiment. Top: Bias; Middle: Dispersion (square root of BCEA); Bottom:  $D_{min}$ . Left: values per target, averaged over participants. Right: Values per participant, averaged over targets.**

Subject protocol approved by the Institutional Review Board at our school.

### 3.3 Procedure

Participants were asked to sit in front of the computer, which was placed on a tabletop. The experimenter ensured that they sat at a distance from the screen that was within the admissible range for the Tobii tracker. The average distance of each participant to the screen was recorded by the tracker (min: 533 mm; max: 717 mm; mean: 632 mm). They first completed the procedure for calibration of the Tobii gaze tracker. Then, they completed the calibration procedure for the FAZE algorithm. At this point, the data acquisition part started. This comprised two tasks.

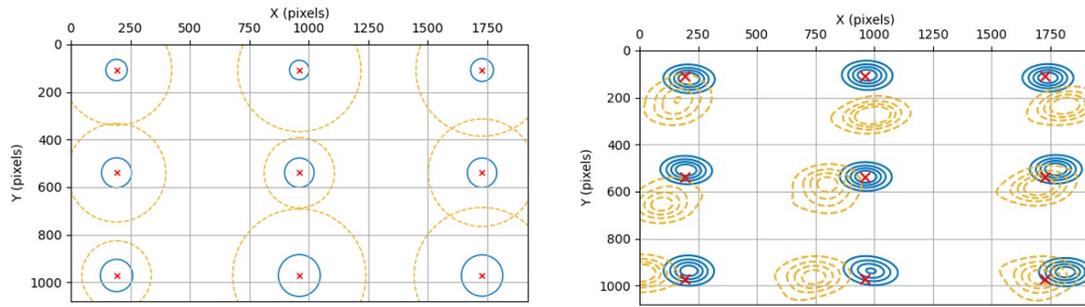
**Task 1:** Participants were asked to stare at a target (a small blue disk of 16 pixels in diameter) appearing in a sequence of 9 locations on the screen (see Fig. 1), and remaining in the same locations for 6 seconds before moving to the next one. (This amount of time is consistent with other experiments on fixation stability [Fragiotta et al. 2018].)

**Task 2:** Participants were presented with a text document (extracted from Carroll’s Alice in Wonderland). The text document was formatted using Times New Roman font at 11pt, consisting of 15 lines with an interline distance of 18pt (24 pixels), and they were asked to read it in its entirety. In addition, participants were asked to press a button on the keyboard when they started a new line and to press another button when they ended that line. In this way, we were able to record the in-line time intervals. Participants were at liberty to read the text aloud or silently (only P1 read it aloud).

Timestamped images of the participants were recorded from the computer camera at a rate of 10 fps for offline processing. Timestamped gaze points from the Tobii tracker were recorded by a Python application built on the Tobii Pro SDK.

### 3.4 Measurements

**3.4.1 Dwelling.** Selection by dwelling mechanisms [Müller-Tomfelde 2007] typically defines an area (e.g., a circle with diameter  $D$ ) around a certain target (e.g., a button to be clicked). When the gaze point is located within this area continuously for a period of time  $T$ , the selection is triggered. We are interested in evaluating how errors in gaze measurements affect selection by dwelling, and how to properly design a system that accounts for these errors. We are not considering here the dynamic aspects of this task, which can be described using variants of Fitt’s law [Zhang et al. 2010, 2011]. Rather, we look for the minimum diameter  $D_{min}$  of a circle around the target that ensures, with a certain probability  $P$ , that selection is triggered when the user is fixating the target for a period of time  $T$ . Intuitively,  $D$  will need to be larger for noisy measurements, as noise may push measurements away from the point of fixation. In our experiments, we set  $T$ , the dwelling time, to 700 ms, as this was found to be appropriate for simple tasks in prior research [Stampe and Reingold 1995; Zhang et al. 2011]. We set  $P$  to 0.9. To find  $D_{min}$ , we first considered the interval of time  $[t_{in}(i), t_{end}(i)]$  (approximately 6 seconds long) during which participants fixated the  $i$ -th target in Task 1. We defined a sequence of finely spaced values for  $D$ , and for each such value, we slid a time window of duration  $T$  through  $[t_{in}(i), t_{end}(i)]$ . For each window location, we checked



**Figure 2: Dwelling experiment.** Left: A visualization of the dwelling circles with a diameter equal to the mean of  $D_{\min}$  across all participants. Right: Contour plots of probability density functions fitted to the recorded data for P7. Contour levels were set at 20,40,60, and 80 percentile. Solid blue line: Tobii data. Yellow dashed line: FAZE data.

whether or not all gaze points measured in that time window were within a distance of  $D/2$  from the center of the target. The proportion of window locations for which this was the case represents the probability  $P_D(i)$  that, when staring at the  $i$ -th target, selection would be triggered for a dwelling circle of diameter  $D$ . Finally, we defined  $D_{\min}$  as the smallest value of  $D$  for which  $P_D(i) \geq 0.9$ .

In addition, we provide measures of bias and dispersion. Bias is defined as the distance, for each target, between the barycenters of the gaze points measured from Tobii or FAZE and the actual target location. Dispersion is measured as the square root of BCEA (bivariate contour ellipse area). BCEA, a metric commonly used for fixation studies (e.g. [Blignaut and Beelders 2012; Niehorster et al. 2020]), represents the area of the ellipse containing 63% of the gaze values, which are modeled as normally distributed. Noisy measurements are typically characterized by large BCEA values. We used all the data within each period  $[t_{in}(i), t_{end}(i)]$  to measure bias and BCEA at each target.

It is important to note that both  $D_{\min}$  and dispersion are affected by measurement noise *as well as* by any fixation instability of the viewer. BCEA is unaffected by bias (constant error terms).

To determine the intervals  $[t_{in}(i), t_{end}(i)]$ , we define a circle of radius 4 pixels around each marker.  $t_{in}(i)$  and  $t_{out}(i)$  are the times at which gaze as measured by the reference Tobii tracker enters and exits the circle defined at the  $i$ -th marker.

**3.4.2 Text Reading - Line Identification.** The ability to identify which text line in an onscreen document one is currently reading hinges on the measured gaze being located within a narrow area containing the line. We are only concerned with in-line reading here, and neglect retracing time (*return sweeps* [Rayner and Pollatsek 2006]). We do not consider a specific vertical coordinate as a reference (e.g., the midline of the text) since the user’s gaze is not constrained to such a line while reading. Instead, we take the Tobii data as a reference, against which to compare FAZE data.

For the  $i$ -th text line, we measure, for both Tobii and FAZE data, the mean  $\mu_y(i)$  and standard deviation  $\sigma_y(i)$  of the Y coordinate of gaze points.  $\sigma_y(i)$  measures the vertical dispersion; it provides an indication of the minimum interline distance for reliable line identification. The differences of the means  $\mu_y(i)$  between FAZE data and the reference Tobii data represent the residual vertical bias.

**3.4.3 Text Reading - Progression Along a Line.** During reading, one’s eyes are not gliding smoothly along a text line; rather, gaze proceeds as a sequence of fixations (during which gaze is relatively static) and saccades, which are rapid movements forward in the line, or, occasionally, backward (*regressions*) [Rayner and Pollatsek 2006]. For our measure of progression along a line, we consider all fixations detected from Tobii data during line reading. Fixation detection is a relatively straightforward operation, and the accuracy of infrared trackers such as Tobii Spark is adequate for this purpose [Olsen 2012]. For this purpose, we use a simple velocity-based algorithm inspired by the Tobii I-VT fixation filter [Olsen 2012]. For the  $i$ -th fixation period, we compute the average value  $\mu_{x,f}(i)$  and the standard deviation  $\sigma_{x,f}(i)$  of the X coordinate for both Tobii and FAZE data. The difference between  $\mu_{x,f}(i)$  values in the two cases is an indication of how accurately the reading location along a line can be tracked using an algorithm like FAZE.

In addition, we computed the standard deviation  $\sigma_{x,s}(i)$  of the X coordinate of gaze point for both Tobii and FAZE data in the periods outside fixations (saccades [Rayner and Pollatsek 2006]). Comparison of  $\sigma_{x,s}(i)$  against  $\sigma_{x,f}(i)$  provides an indication of the relative dispersion during fixations (periods with low gaze point variance) and during saccades (when variance is large due to fast motion).

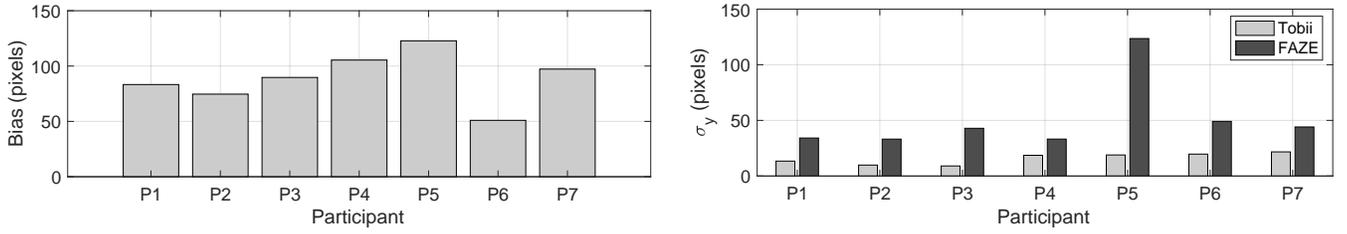
## 4 RESULTS

We present the results of our experiments in the following. All statistical tests were conducted at 5% significance level. In order to visually highlight any dependencies of the recorded values on the participants’ distance to the screen, participant indices were sorted according to increasing distance to the screen.

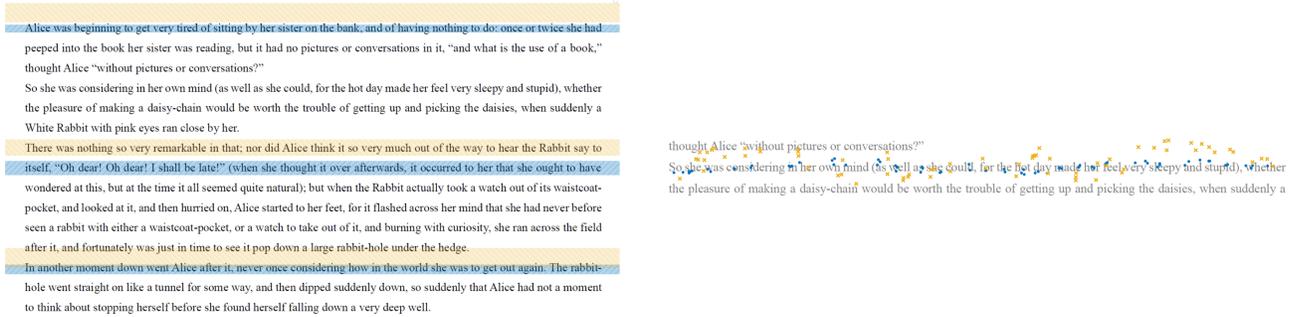
### 4.1 Fixation

Recorded values of bias, dispersion, and  $D_{\min}$  are shown in Fig. 1. Specifically, we report, for both Tobii and FAZE, the values averaged across participants for each target, as well as the values averaged across targets for each participant. As expected, FAZE data have significantly larger bias and dispersion than Tobii data (as revealed by paired t-test).

Total means for Tobii data were: bias: 54.0 pixels; the square root of BCEA: 1.13°;  $D_{\min}$ : 125.3 pixels. For FAZE data, the total



**Figure 3: Text Reading - Line Identification experiment. (a) Bias (RMSE of the difference of  $\mu_y(i)$  between Tobii and FAZE data). (b)  $\sigma_y$  averaged over all text lines.**



**Figure 4: Left: Strips showing for  $\mu_y(i) \pm \sigma_y(i)$  for data recorded for participant P4, shown for three text lines. Right: An example of gaze points recorded for a text line (P6). Data from Tobii was subsampled to 10 Hz for comparison with FAZE data. Blue forward slash or circular symbol: Tobii data. Yellow backward slash or cross symbol: FAZE data.**

means were: bias: 144.2 pixels; the square root of BCEA:  $10.45^\circ$ ;  $D_{\min}$ : 501 pixels. (For references, values of the square root of BCEA reported in the literature, measured using accurate micropenimeter instruments, varied from  $0.08^\circ$  to  $0.4^\circ$  [Kumar and Chung 2014].) Note from Fig. 1 that the BCEA value for P5 was substantially higher than for other participants, though this did not translate into a larger  $D_{\min}$  value.

Two-way analysis of variance revealed a significant effect of participants on both  $D_{\min}$  and square root of BCEA, for both Tobii and FAZE data. A significant effect of target was found for Tobii data only, on both  $D_{\min}$  and square root of BCEA. A significant correlation between distance and both bias and  $D_{\min}$  was found for FAZE data only ( $\rho = 0.85$  in both cases). A graphical representation of  $D_{\min}$  for each target (averaged over all participants) for both Tobii and FAZE is shown in Fig. 2, left. An example of data collected with the two modalities for a single participant (P7) is presented in Fig. 2, right, which shows contours at the same percentile levels of the probability density functions fitted to the recorded samples.

## 4.2 Text Reading - Line Identification

Relevant data from the experiment is shown in Fig. 3. The text line index was not shown to have a significant effect on either bias (RMSE of the difference of the means  $\mu_y$  measured for each line for FAZE or Tobii), nor on  $\sigma_y$  for either FAZE or Tobii data. Participant index had a significant effect on both bias and  $\sigma_y$ .  $\sigma_y$  was found to be significantly larger for FAZE than for Tobii. For Tobii data only,

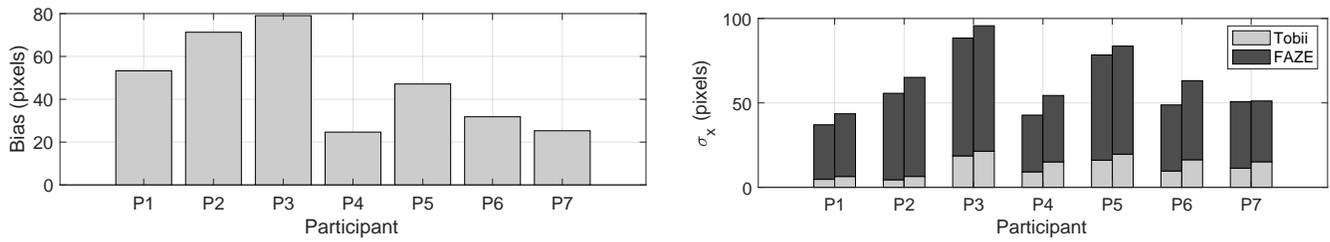
$\sigma_y$  was found to be correlated with distance to the screen ( $\rho = 0.78$ ). The total mean of the bias was 91.7 pixels, while the total mean of  $\sigma_y$  was 15.9 pixels for Tobii data and 51.5 for FAZE data. From Fig. 3, it is seen that P5 had a much larger value of  $\sigma_y$  (averaged across lines) than the others. An example of strips containing gaze data at  $\mu_y(i) \pm \sigma_y(i)$  is shown in Fig. 4, left.

## 4.3 Text Reading - Progression Along a Line

We computed all fixation times (during in-line reading intervals) on the Tobii data, then, as explained in Sec. 3.4.3, we computed the RMSE of the difference of the mean values  $\mu_{x,f}(i)$  of the X coordinate of measurements from Tobii and FAZE. The resulting bias value is shown in Fig. 5, left. The mean of RMSE across participants was 89.2 pixels. For both Tobii and FAZE data, we also computed the standard deviation  $\sigma_{x,f}(i)$  and  $\sigma_{x,s}(i)$  of the X coordinate of gaze for all periods identified as fixations and saccades, respectively, based on the Tobii data. The mean values are shown in Fig. 5. For both Tobii and FAZE data, paired t-test rejected the null hypothesis of equal mean of  $\sigma_{x,f}$  and of  $\sigma_{x,s}$ . An example of gaze data on a text line is shown in Fig. 4, right.

## 5 DISCUSSION AND CONCLUSIONS

Appearance-based gaze tracking algorithms hold the promise to “democratize” gaze-based interactions and analysis by removing the need to purchase dedicated devices. However, it is critical that these systems be tested in realistic applications, in order to assess their



**Figure 5: Text Reading - Progression Along a Line experiment. (a) Bias (RMSE of the difference of  $\mu_{x,f}(i)$  for Tobii and FAZE data). (b)  $\sigma_x$  averaged over all fixations (left bar in bar groups) and saccades (right bar in bar groups).**

practical usability [Hohlfeld et al. 2015; Zhang et al. 2019]. In this paper, we proposed a number of metrics associated with specific applications of interest, and compared measurements taken with a state-of-the-art appearance-based tracker against those taken with an infrared gaze tracker.

Our first experiment showed that dwelling-based selection is possible with FAZE, but the dwelling areas must be substantially larger than those afforded by an infrared tracker for equal effectiveness (Fig. 2, left). In our measurements, the ratio of the diameters  $D_{\min}$  found for FAZE to those found with Tobii (averaged over all participants) varied from 2.3 to 6.3. Our text reading - line identification experiment showed the dispersion across the Y coordinate of FAZE data to be more than 3 times larger than that of Tobii data. This suggests that the minimum interline distance needs to be larger by at least that same amount, in order to ensure reliable text line identification. This is compounded by the effect of bias, which measures the difference between the Y coordinate of the values measured by Tobii and FAZE in the same line, and that was found to be 92 pixels on average in our experiment. This is almost 4 times the interline distance used in the text document considered for our experiment (see Fig. 4, left). Our text reading - progression along a line experiment showed an RMSE value of the difference of X coordinates during fixations of almost 90 pixels. Considering that in our document the width of a character was about 13.5 pixels on average, this bias translates to an expected error of about 7 characters. Interestingly, we found a significant difference in the mean of the standard deviation of FAZE measured during fixation and saccade intervals (where these intervals were computed based on our reference Tobii data). This suggests that it may be possible to identify fixations on FAZE data using appropriate local analysis.

In most of the cases, measurements on the FAZE data were found to correlate positively with the distance to the screen. This should not be surprising, considering that gaze tracking algorithms measure the direction of the visual axis, and the effect of an angular error on the location of the gaze point increases linearly with the distance.

Our study considered a relatively small population sample (7 participants), and we are planning for a larger study in the near future, which will include different illumination types (which can affect the quality of FAZE data) and a larger range of viewing distances. Another limitation of this work is that the image data was processed offline. In future experiments, we will run FAZE online. Besides a reduced frame rate (6 frames/second on our TensorBook),

latency (delay) should be expected, and its effect on specific tasks (e.g., dwelling) will be analyzed.

## ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Eye Institute of the National Institutes of Health under award number R01EY030952-01A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank the participants who volunteered for this study.

## REFERENCES

- Michael Ashmore, Andrew T Duchowski, and Garth Shoemaker. 2005. Efficient eye pointing with a fisheye lens. In *Proceedings of Graphics interface 2005*. 203–210.
- Ralf Biedert, Georg Buscher, and Andreas Dengel. 2009. The eye book. *Informatik-Spektrum* 33, 3 (2009), 272–281.
- Pieter Blignaut and Tanya Beelders. 2012. The precision of eye-trackers: a case for a new measure. In *Proceedings of the symposium on eye tracking research and applications*. 289–292.
- Stephen Bottos and Balakumar Balasingam. 2020. Tracking the progression of reading using eye-gaze point measurements and hidden markov models. *IEEE Transactions on Instrumentation and Measurement* 69, 10 (2020), 7857–7868.
- Senuri De Silva, Sanuwani Dayarathna, Gangani Ariyaratne, Dulani Meedeniya, Sampath Jayarathna, Anne MP Michalek, and Gavindya Jayawardena. 2019. A rule-based system for ADHD identification using eye movement data. In *2019 Moratuwa Engineering Research Conference (MERCon)*. IEEE, 538–543.
- Shuwen Deng, David R Reich, Paul Prasse, Patrick Haller, Tobias Scheffer, and Lena A Jäger. 2023. Eyettention: An Attention-based Dual-Sequence Model for Predicting Human Scanpaths during Reading. *Proceedings of the ACM on Human-Computer Interaction* 7, ETRA (2023), 1–24.
- Heiko Drewes, Alexander De Luca, and Albrecht Schmidt. 2007. Eye-gaze interaction for mobile phones. In *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*. 364–371.
- Serena Fragiotta, Carmela Carnevale, Alessandro Cutini, Erika Rigoni, Pier Luigi Grenga, and Enzo Maria Vingolo. 2018. Factors influencing fixation stability area: a comparison of two methods of recording. *Optometry and Vision Science* 95, 4 (2018), 384–390.
- Elias Daniel Guestrin and Moshe Eizenman. 2006. General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on biomedical engineering* 53, 6 (2006), 1124–1133.
- Oliver Hohlfeld, André Pomp, Jö Ägila Bitsch Link, and Dennis Guse. 2015. On the applicability of computer vision based gaze tracking in mobile scenarios. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 427–434.
- Anneline Huck. 2016. *An eye tracking study of sentence reading in aphasia: influences of frequency and context*. Ph.D. Dissertation. City University London.
- Anke Huckauf and Mario H Urbina. 2008. On object selection in gaze controlled environments. *Journal of Eye Movement Research* 2, 4 (2008).
- Robert JK Jacob. 1991. The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS)* 9, 2 (1991), 152–169.
- Girish Kumar and Susana TL Chung. 2014. Characteristics of fixational eye movements in people with macular disease. *Investigative ophthalmology & visual science* 55, 8 (2014), 5125–5133.

- Manu Kumar, Terry Winograd, and Andreas Paepcke. 2007. Gaze-enhanced scrolling techniques. In *CHI'07 Extended Abstracts on Human Factors in Computing Systems*. 2531–2536.
- Xueshi Lu, Difeng Yu, Hai-Ning Liang, Wenge Xu, Yuzheng Chen, Xiang Li, and Khalad Hasan. 2020. Exploration of hands-free text entry techniques for virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 344–349.
- Tobias Lunte and Susanne Boll. 2020. Towards a gaze-contingent reading assistance for children with difficulties in reading. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–4.
- Roberto Manduchi and Susana Chung. 2022. Gaze-Contingent Screen Magnification Control: A Preliminary Study. In *International Conference on Computers Helping People with Special Needs*. Springer, 380–387.
- Natalie Maus, Dalton Rutledge, Sedeeq Al-Khazraji, Reynold Bailey, Cecilia Ovesdotter Alm, and Kristen Shinohara. 2020. Gaze-guided magnification for individuals with vision impairments. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
- Raphael Menges, Chandan Kumar, and Steffen Staab. 2019. Improving user experience of eye tracking-based interaction: Introspecting and adapting interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 6 (2019), 1–46.
- Christian Müller-Tomfelde. 2007. Dwell-based pointing in applications of human computer interaction. In *Human-Computer Interaction—INTERACT 2007: 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part I 11*. Springer, 560–573.
- Michael Murias, Samantha Major, Katherine Davlantis, Lauren Franz, Adrienne Harris, Benjamin Rardin, Maura Sabatos-DeVito, and Geraldine Dawson. 2018. Validation of eye-tracking measures of social attention as a potential biomarker for autism clinical trials. *Autism Research* 11, 1 (2018), 166–174.
- Diederick C Niehorster, Raimondas Zembyls, Tanya Beelders, and Kenneth Holmqvist. 2020. Characterizing gaze position signals and synthesizing noise during fixations in eye-tracking data. *Behavior Research Methods* 52 (2020), 2515–2534.
- Anneli Olsen. 2012. The Tobii I-VT fixation filter. *Tobii Technology* 21 (2012), 4–19.
- Bing Pan, Helene A Hembrooke, Geri K Gay, Laura A Granka, Matthew K Feusner, and Jill K Newman. 2004. The determinants of web page viewing behavior: an eye-tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications*. 147–154.
- Seonwook Park, Shalini De Mello, Pavlo Molchanov, Umar Iqbal, Otmar Hilliges, and Jan Kautz. 2019. Few-shot adaptive gaze estimation. In *Proceedings of the IEEE/CVF international conference on computer vision*. 9368–9377.
- Seonwook Park, Xucong Zhang, Andreas Bulling, and Otmar Hilliges. 2018. Learning to find eye region landmarks for remote gaze estimation in unconstrained settings. In *Proceedings of the 2018 ACM symposium on eye tracking research & applications*. 1–10.
- Karalyn Patterson and Matthew A Lambon Ralph. 1999. Selective disorders of reading? *Current opinion in neurobiology* 9, 2 (1999), 235–239.
- Yesaya Tommy Paulus and Gerard Bastiaan Remijn. 2021. Usability of various dwell times for eye-gaze-based object selection with eye tracking. *Displays* 67 (2021), 101997.
- Peter Raatikainen, Jarkko Hautala, Otto Loberg, Tommi Kärkkäinen, Paavo Leppänen, and Paavo Nieminen. 2021. Detection of developmental dyslexia with machine learning using eye movement data. *Array* 12 (2021), 100087.
- Ramkumar Rajendran, Anurag Kumar, Kelly E Carter, Daniel T Levin, and Gautam Biswas. 2018. Predicting Learning by Analyzing Eye-Gaze Data of Reading Behavior. *International Educational Data Mining Society* (2018).
- Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (1998), 372.
- Keith Rayner and Alexander Pollatsek. 2006. Eye-movement control in reading. In *Handbook of psycholinguistics*. Elsevier, 613–657.
- Erik D Reichle, Keith Rayner, and Alexander Pollatsek. 2003. The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences* 26, 4 (2003), 445–476.
- Louis B Rosenberg. 2008. Gaze-responsive interface to enhance on-screen user reading tasks. US Patent 7,429,108.
- Gianluca Schiavo, Simonetta Osler, Nadia Mana, and Ornella Mich. 2015. Gary: Combining speech synthesis and eye tracking to support struggling readers. In *Proceedings of the 14th international conference on mobile and ubiquitous multimedia*. 417–421.
- Selina Sharmin, Oleg Špakov, and Kari-Jouko Rähkä. 2013. Reading on-screen text with gaze-based auto-scrolling. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*. 24–31.
- Linda E Sibert and Robert JK Jacob. 2000. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 281–288.
- Dave M Stampe and Eyal M Reingold. 1995. Selection by looking: A novel computer interface and its application to psychological research. In *Studies in visual information processing*. Vol. 6. Elsevier, 467–478.
- Xiaohao Sun and Balakumar Balasingam. 2021. Reading line classification using eye-trackers. *IEEE Transactions on Instrumentation and Measurement* 70 (2021), 1–10.
- Tobii. [n. d.]. Tobii Pro Spark: Enter the world of eye tracking. <https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-spark>. Online; accessed Jan. 8, 2024.
- Francisco Vicente, Zehua Huang, Xuehan Xiong, Fernando De la Torre, Wende Zhang, and Dan Levi. 2015. Driver gaze tracking and eyes off the road detection system. *IEEE Transactions on Intelligent Transportation Systems* 16, 4 (2015), 2014–2027.
- Tan Vo, B Sumudu U Mendis, and Tom Gedeon. 2010. Gaze pattern and reading comprehension. In *Neural Information Processing, Models and Applications: 17th International Conference, ICONIP 2010, Sydney, Australia, November 22-25, 2010, Proceedings, Part II 17*. Springer, 124–131.
- Ru Wang, Zach Potter, Yun Ho, Daniel Killough, Linxiu Zeng, Sanbrita Mondal, and Yuhang Zhao. 2024. GazePrompt: Enhancing Low Vision People's Reading Experience with Gaze-Aware Augmentations. *arXiv preprint arXiv:2402.12772* (2024).
- Erroll Wood and Andreas Bulling. 2014. Eytetab: Model-based gaze estimation on unmodified tablet computers. In *Proceedings of the symposium on eye tracking research and applications*. 207–210.
- Xuan Zhang and I Scott MacKenzie. 2007. Evaluating eye tracking with ISO 9241-Part 9. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments: 12th International Conference, HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part III 12*. Springer, 779–788.
- Xinyong Zhang, Xiangshi Ren, and Hongbin Zha. 2010. Modeling dwell-based eye pointing target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2083–2092.
- Xucong Zhang, Yusuke Sugano, and Andreas Bulling. 2019. Evaluation of appearance-based methods and implications for gaze-based applications. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. 2017. It's written all over your face: Full-face appearance-based gaze estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 51–60.
- Xinyong Zhang, Pianpian Xu, Qing Zhang, and Hongbin Zha. 2011. Speed-accuracy trade-off in dwell-based eye pointing tasks at different cognitive levels. In *Proceedings of the 1st international workshop on pervasive eye tracking & mobile eye-based interaction*. 37–42.