GazeCast: Using Mobile Devices to Allow Gaze-based Interaction on Public Displays

Omar Namnakani
o.namnakani.1@research.gla.ac.uk
University of Glasgow
Glasgow, United Kingdom

Penpicha Sinrattanavong
chingpenpicha@gmail.com
University of Glasgow
Glasgow, United Kingdom

Yasmeen Abdrabou
yasmeen.essam@unibw.de
University of the Bundeswehr
Munich, Germany and
University of Glasgow, UK

Andreas Bulling
florian.alt@unibw.de
University of Stuttgart
Stuttgart, Germany

Florian Alt
florian.alt@unibw.de
University of the Bundeswehr Munich
Munich, Germany

Mohamed Khamis
mohamed.khamis@glasgow.ac.uk
University of Glasgow
Glasgow, United Kingdom

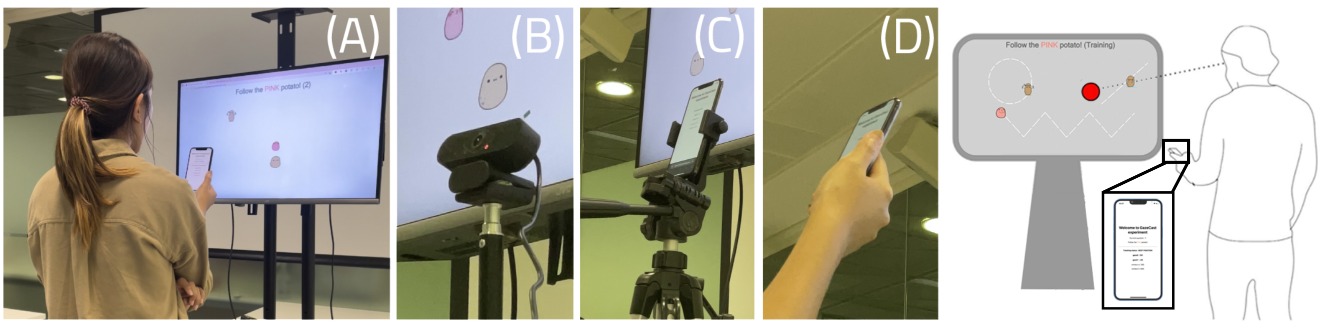


Figure 1: We propose GazeCast, a novel system that leverages users' personal hand-held mobile devices to enable gaze-based interaction with surrounding displays using Pursuits (A). In a user study, we evaluate GazeCast by experimenting with three tracking setups where participants' gaze was tracked: using a web camera (B), GazeCast with a mobile device attached to a stand (C), and GazeCast with a mobile device held by the participants (D). The illustration is adapted from GTmoPass by Khamis et al. [Khamis et al. 2017a].

ABSTRACT

Gaze is promising for natural and spontaneous interaction with public displays, but current gaze-enabled displays require movement-hindering stationary eye trackers or cumbersome head-mounted eye trackers. We propose and evaluate GazeCast – a novel system that leverages users' handheld mobile devices to allow gaze-based interaction with surrounding displays. In a user study (N = 20), we compared GazeCast to a standard webcam for gaze-based interaction using Pursuits. We found that while selection using GazeCast requires more time and physical demand, participants value GazeCast's high accuracy and flexible positioning. We conclude by discussing how mobile computing can facilitate the adoption of gaze interaction with pervasive displays.

ETRA '23, May 30-June 2, 2023, Tubingen, Germany

CCS CONCEPTS

• **Human-centered computing** → **Interaction design**; **Interaction techniques**.

KEYWORDS

Public Displays, Eye Tracking, Pursuits, Gaze Interaction, Mobile Devices

ACM Reference Format:
Omar Namnakani, Penpicha Sinrattanavong, Yasmeen Abdrabou, Andreas Bulling, Florian Alt, and Mohamed Khamis. 2023. GazeCast: Using Mobile Devices to Allow Gaze-based Interaction on Public Displays. In *2023 Symposium on Eye Tracking Research and Applications (ETRA '23), May 30-June 2, 2023, Tubingen, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3588015.3589663

1 INTRODUCTION

Gaze-based interaction is deemed particularly promising for public displays. Interacting using gaze is fast [Sibert and Jacob 2000] possible from a distance, hence offering users an attractive alternative to touch interaction that is also suitable for unreachable displays (e.g., displays behind glass windows [Davies et al. 2014]). So far, gaze interaction has been enabled for public displays by either augmenting the display using remote eye trackers and webcams or requiring the user to wear a head-mounted eye tracker [Khamis et al. 2016; Lander et al. 2015a]. While public displays require users to interact from different positions and distances to the display [Müller et al. 2012], stationary eye trackers have a limited range. This means that for users to interact with public displays using gaze, they need to position themselves in the display's "sweet spot" [Alt et al. 2015]. While head-mounted eye trackers allow for freedom of movement and interaction from a distance [Lander et al. 2015a], they require person-specific calibration and gaze mapping to each display. Although head-mounted trackers have recently become affordable [Kassner et al. 2014] and are envisioned to be integrated into daily Eyewear [Bulling and Kunze 2016], they are still not in wide-spread use and require augmentation of each user [Lander et al. 2015a]. At the same time, off-the-the-shelf mobile phones allow for accurate gaze estimation, which resulted in many mobile apps in which gaze outperforms touch [Khamis et al. 2022; Liu et al. 2015; Zhang et al. 2017a]. In this paper, we introduce GazeCast – a novel system that leverages users' handheld devices for gaze-based interaction on public displays. GazeCast uses the smartphone's front-facing camera to cast users' gaze onto the display. As such, GazeCast does not require special-purpose eye tracking equipment, does not restrict the user's positioning, can support calibration-free interaction as demonstrated in our study, and can be easily extended to support multiple users. Furthermore, collecting gaze data through public displays has privacy implications due to the sheer amount of sensitive information inferred from eye movements [Bozkir et al. 2021, 2020; David-John et al. 2022], GazeCast gives more control to the user on which gaze data is transferred to the display. In our implementation, we transfer the gaze estimates to the display for processing, but the concept allows gaze data to be processed locally on the user's phone.

We present the results of a user study in which 20 participants provided gaze input on a situated display using Pursuits [Vidal et al. 2013]. To evaluate GazeCast, we experimented with three tracking setups: participants' gaze was tracked using 1) a webcam (baseline), 2) GazeCast with the smartphone attached to a stand, GazeCast (on a stand), and 3) GazeCast with the smartphone held by the participant, GazeCast (handheld). Results show that while selection time and perceived workload increase with GazeCast, error counts are lower and subjective feedback indicates users value how GazeCast makes their task easier and more accurate and has the potential to make the experience more privacy-preserving. We discuss how mobile computing facilitates the adoption of gaze interaction on public displays.

2 RELATED WORK

We build on prior work on gaze-based interaction on public displays and gaze-enabled handheld mobile devices.

2.1 Gaze-based Interaction on Public Displays

Researchers have investigated the use of gaze input for public display applications, such as voting [Khamis et al. 2016], consuming news [Lander et al. 2015b], interacting with medical images data [Hatscher et al. 2017], gaming [Vidal et al. 2013; Zhang et al. 2013], measuring attention [Alt et al. 2016], and authentication [Khamis et al. 2017a, 2018c]. A recurring challenge in gaze interaction with public displays is that public displays expect users to interact from different positions and distances to the display [Müller et al. 2012]. However, current remote eye trackers limit users' mobility [San Agustin et al. 2010] and require users to keep their heads in a confined tracking box about 70 cm away from the screen [Khamis et al. 2017b]. Approaches to address this included the use of head-mounted eye trackers [Lander et al. 2015a], which require gaze mapping to each display and are cumbersome to wear. Another approach was to use active eye tracking, where the eye tracker physically moves depending on the user's position, as done in EyeScout [Khamis et al. 2017b]. In EyeScout, the eye tracker was mounted on a conveyor-belt-like rail but only covered a limited area. A second challenge is that most remote eye trackers are optimized for viewing angles that correspond to screens up to 24 inches, which means that they may not track gaze accurately on larger displays [Rajanna and Hammond 2018]. A third challenge is that public displays expect multiple users to interact simultaneously [Memarovic et al. 2014], but most remote eye-tracking setups support one user at a time. GazeCast works around these challenges as it does not pose positioning and display size requirements, and multiple users can connect to the same display using their handheld mobile devices.

2.2 Gaze-enabled Handheld Mobile Devices

Mobile devices are now equipped with high-resolution front-facing depth cameras and high-performance processors. These advances, alongside the recent developments in computational gaze estimation, have made eye tracking increasingly available on mobile devices [Khamis et al. 2018a]. This resulted in increased eye-tracking applications that run directly on smartphones and tablets. A survey on gaze-enabled handheld mobile devices classifies eye tracking applications on mobile devices to: 1) gaze behaviour analysis, where users' eye movements are silently tracked for later analysis, 2) implicit gaze-based interaction where the system reacts to the users' natural eye movements, and 3) explicit gaze-based interaction where the user deliberately moves their eyes to provide input [Khamis et al. 2018a]. We focus on the last as GazeCast is an application for explicit gaze-based interaction. Examples of mobile apps that feature real-time gaze-based interaction include GazeSpeak [Zhang et al. 2017a], which supports communication for people with motor disabilities, gaze-based authentication using GazeTouchPass [Khamis et al. 2022], and many other generic gaze interaction applications from research [Kong et al. 2021; Lee et al. 2017; Li et al. 2017; Miluzzo et al. 2010] and industry [Access 2019; SeeSo 2022].

Most relevant to our work are WorldGaze [Mayer et al. 2020] and GTmoPass [Khamis et al. 2017a]. WorldGaze estimates the head-pose using the front-facing camera and detects the gazed-at objects using the rear camera. This was then used to improve voice assistants e.g., looking at a store and asking "What time does this close?". GazeCast is different in that it detects eye gaze rather than head-pose, and relies on stimuli that are displayed on nearby displays. GTmoPass detects multimodal gaze and touch input on mobile devices to authenticate users on public displays. While GTmoPass only detects gaze gestures and only communicates

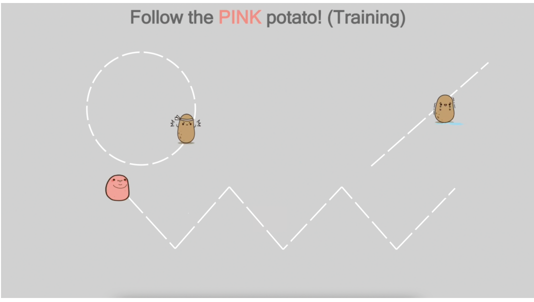


Figure 2: Three trajectories are shown on the screen. The circular movement is located at the upper left of the screen; the lower center portion of the screen is where the zigzag route is situated; the diagonal trajectory runs from the upper right corner of the screen to the middle. The white lines are shown for illustration purposes and were not shown to the participants.

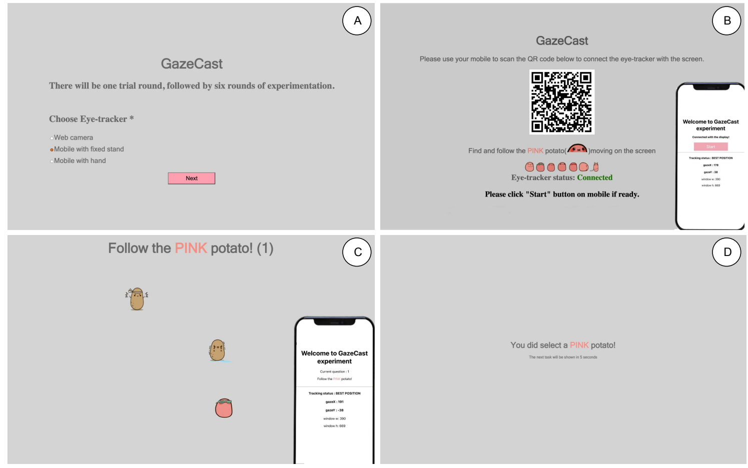


Figure 3: Study Screens A) Public display initial screen showing the different tracking setups, B) Connection status, C) Example trajectory on the public display with gaze data printed on the mobile device, D) Selection Screen.

to the display whether the authentication attempt was successful, GazeCast can detect smooth pursuit eye movements and is not limited to authentication applications.

3 GAZECAST CONCEPT AND IMPLEMENTATION

3.1 Concept

The idea behind GazeCast is to use one's smartphone's camera as an eye tracker to interact with public displays. There are different ways to connect the smartphone to the public display. In our work, a QR code is shown on the display, which users can scan, directing them to the web application with instructions allowing them to interact with the display. We implemented two web applications (server and client) that can be accessed using different internet browsers e.g., Google Chrome. The concept of GazeCast can be extended to allow more than one user to interact with the same display concurrently by connecting their devices to the public display and establishing the connection.

Building on previous work [Khamis et al. 2016; Vidal et al. 2013], our system uses Pursuits to enable calibration-free gaze-based interaction with public displays. Being a calibration-free interaction technique, Pursuits is particularly suitable for public displays which need to be "immediately usable" [Davies et al. 2014]. By adopting Pursuits, GazeCast allows spontaneous interaction with public displays eliminating the need for calibration. Pursuits checks for motion correlation between the user's eye movements and trajectories of on-screen moving targets [Esteves et al. 2015; Vidal et al. 2013]. The method's strength lies in its ability to determine which object the user is gazing at by studying their eye behaviour without the need for accurate gaze estimates.

3.2 Implementation

GazeCast has two main components: the web application user interface (Client) and the Server, implemented using Next.js. We used "socket.io-client" for two-way communication between the client and the server. The client can access the application via a browser, such as Google Chrome, and can connect to the application by sending a message to join the connection. The server receives the gaze data from the client and calculates the correlation every second (30 frames). After detecting the target, it sends the message back to the client to trigger the following view. The message can be sent directly to the client by the server or broadcast to all connected clients. Throughout the application, we continuously log the gaze data and the moving target's locations in x-, y- coordinates. Gaze data was collected using the SeeSo.io SDK [SeeSo 2022]. We used Pearson's product-moment correlation coefficient to compute the correlation between the user's gaze and the targets' movement as done in prior work [Drewes et al. 2018, 2019; Esteves et al. 2015; Velloso et al. 2017; Vidal et al. 2013]. A target is selected if it has the highest correlation to the user's eye movements as long as the correlation of the smallest of x and y is greater than the threshold value of 0.8 as suggested by the literature [Esteves et al. 2015]. The choice of the window size of 1 second (30 samples at a rate of 30 frames per second) is based on prior work [Esteves et al. 2020, 2015] and pilot testing with two participants.

While in principle, the calculation of the correlation can be done on the client's side (i.e., the personal mobile device), we decided to process the gaze data on the server. This was done to ensure a fair comparison with the baseline – this way, the data is processed by the same machine for all experimental conditions. Extending GazeCast to allow the processing of gaze data on the phone would have privacy benefits, as we discuss further in Section 6.

4 GAZECAST EVALUATION

4.1 Study Design

To evaluate the performance of the tracking setups for interaction with public displays, we designed a repeated measures lab study. The experiment has one independent variable: The tracking setups, which has three levels: (1) A webcam mounted on a stand (baseline) (Figure 1B), (2) GazeCast with the mobile device attached to a stand,

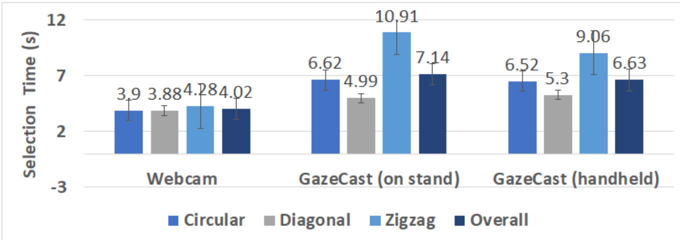


Figure 4: Webcam has the fastest selection time, and the zigzag trajectory led to a longer selection time in all setups. The error bars represent the SD.

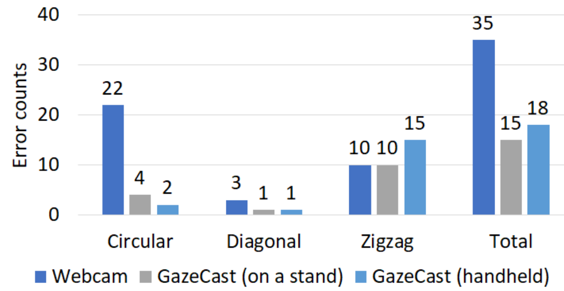


Figure 5: Webcam is more error-prone than GazeCast.

GazeCast (on a stand) (Figure 1C), and (3) GazeCast with the mobile device held by the participant, GazeCast (handheld) (Figure 1D). We included two GazeCast conditions to separate the impact of holding the mobile device from that of the GazeCast concept, as it is known that users hold phones differently, which impacts tracking accuracy [Khamis et al. 2018a]. To control the effect of the target trajectory types, we ensured that each target trajectory was used an equal number of times per tracking setup per participant. We measured: 1) Selection time, which is the time taken to perform a successful selection, 2) Error count, counted whenever a participant made a wrong selection (i.e., selected a brown rather than a pink target), 3) Number of timeouts, defined as the failure to perform either a correct or a wrong selection within 30 seconds, 4) User perceived cognitive load using NASA TLX (link), and 5) User preferences.

4.2 Participants and Apparatus

We recruited 20 participants (10 females), with an average age of 27.7 (SD = 3.96) and an average height of 168.58 cm. Nine participants wore glasses, and three wore contact lenses during the study. Two participants had previous experience with eye tracking, but none had any experience with Pursuits.

For our setup, we used a 32-inch display ($1920 \times 1080 pixels$) in a controlled light environment. We used a standard USB web camera (1080P) for the webcam condition. For GazeCast, we used an iPhone X running on iOS version 15.3.1 with a front-facing camera with a 7 MP sensor. Our web application was opened on a Microsoft Edge browser on an Apple MacBook Air (13", M1, 2020), which was connected to the display directly. To enable eye tracking on the mentioned devices, we used SeeSo SDK [SeeSo 2022], an eye-tracking library that provides real-time gaze data recorded as an (x,y) position in screen coordinates at 30 frames per second and a reported accuracy of 1.6°. We added a marker on the floor (60 cm away from the screen) indicating where the users should stand for a consistent distance between the participants and tracking setups.

4.3 Task and Procedure

We designed a game, "Follow the pink potato" inspired by EyeVote [Khamis et al. 2016]. In each task, the application displays three potatoes floating in different trajectories. Participants are asked to select the target highlighted in pink. Selecting the wrong target would result in an error. Each displayed target moves at a different speed in of the three trajectories: circular, zigzag, or diagonal [Khamis et al. 2016] (see Figure 2). Each target moves along its path back and forth until the system detects a selection or the task exceeds the time limit of 30 seconds [Dwivedi et al. 2017]. Once the system detects a selection or a timeout, a message is shown on the screen for seconds telling the user what their selection was, followed by a new round. Figure 3 shows the different screens for the study.

For the procedure, when the participants arrived at the lab, we explained the I am of the study and asked them to fill in the consent and demographics forms. Participants were asked to perform 6 selections per input technique using Pursuits. The order of the conditions was counterbalanced across participants by a Latin square. For each technique, participants started with one training session (1 selection) to familiarise themselves with the task and the eye-tracking device. After each technique, participants rated their experience through 5-point Likert-scale questions, responded to the NASA-TLX questionnaire, and open-ended questions assessing the ease of use, speed, and effort. At the end of the study, participants were required to answer a final questionnaire to rank the eye-tracking devices based on their preferences and performance. The experiment took an average of 60 minutes.

5 RESULTS

Our data showed normal distribution using Shapiro-Wilk test. hence we used one-way repeated measures ANOVA unless stated otherwise. We reported Greenhouse-Geisser-corrected degrees of freedom in cases where Mauchly's test showed a violation of sphericity. Bonferroni-correction to account for multiple comparisons was used when applicable.

5.1 Selection Time

Figure 4 shows the selection time. Our statistical test showed a significant effect of each tracking setup on selection time $F_{2,38} = 18.340$, $p < .001$. Post-hoc analyses showed significant differences between webcam ($M = 4024.11, SD = 2909.01$) and GazeCast (handheld) ($M = 6627.13, SD = 4284.71$) and between webcam and GazeCast (on a stand) ($M = 7137.47, SD = 4679.40$). As a result, using the webcam yielded the fastest selection time. However, no significant differences were found between both GazeCast conditions.

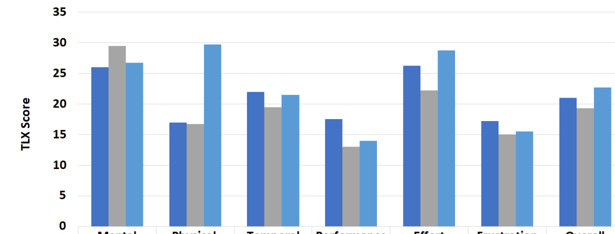
5.2 Error Counts and Timeout

The results revealed a significant effect of each tracking setup on errors $F_{2,38} = 6.961$, $p = .002$. Post-hoc analyses showed significant differences between webcam (35 errors out of 120 selections) and GazeCast (handheld) (18 errors out of 120) at $p = .039$, and between webcam and GazeCast (on a stand) (15 errors out of 120) at $p = .006$. Reflecting on the trajectories, the zigzag trajectory was the most error prune, then the circular and the diagonal trajectories. For the setups, we found that the webcam setup is the most error-pruned, then GazeCast (on a stand) and finally, GazeCast (handheld) (see Figure 5). For timeout, out of all trials by all participants (360 trials), six timeouts were found.

5.3 Perceived Workload and Subjective Rating

Figure 6 shows the average NASA-TLX scores out of 100. The statistical tests revealed that the tracking setup has a significant effect on physical demand, $F_{2,38} = 5.347$, $p < .05$. Post-hoc analyses showed significant differences between physical demand ($p < .05$) induced by GazeCast (handheld) ($M = 29.75, SD = 25.93$) and GazeCast (on a stand) ($M = 16.75, SD = 19.21$). Hence, the participants perceived GazeCast (handheld) as more strenuous than the webcam and GazeCast (on a stand), which is expected. No significant effects of the tracking setups on other workload dimensions were found.

Figure 7 shows participants' subjective ratings. No statistically significant difference was found by Freidman test for the tracking setup on any of the self-reported aspects: ease, speed, accuracy, eye tiredness, and willingness to use daily. However, participants reported that they could perform tasks most easily ($M = 4.4, SD = 0.99$) and accurately ($M = 3.95, SD = 1.1$) with GazeCast (on a stand). This aligns with the results of Error counts. Nonetheless, compared to both GazeCast conditions, participants perceived the webcam setup to be the fastest ($M = 4.25, SD = 1.02$), least tiring ($M = 2.2, SD = 1.15$), easiest to use ($M = 4.3, SD = 1.13$), and most likely to be used in daily life ($M = 4, SD = 1.12$).



	Mental	Physical	Temporal	Performance	Effort	Frustration	Overall
Webcam	26	17	22	17.5	26.25	17.25	21
GazeCast (on a stand)	29.5	16.75	19.5	13	22.25	15	19.3
GazeCast (handheld)	26.75	29.75	21.5	14	28.75	15.5	22.71

Figure 6: The mean responses for the Raw NASA TLX questionnaire.

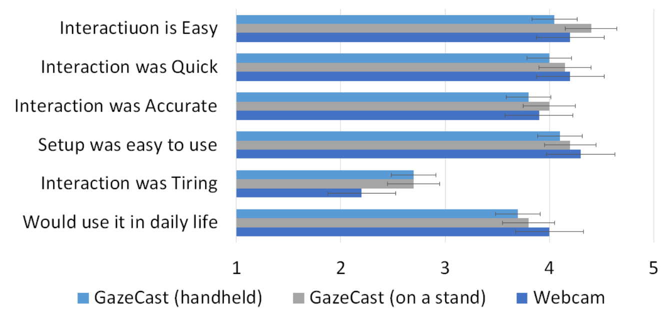


Figure 7: Participants rating for the six usability aspects. 1 "strongly disagree" - 5 "strongly agree." Error bars represent the SD.

5.4 Feedback and Ranking

For Usability and user experience, out of the 20 participants, 10 participants reported that they liked to use the mobile as an eye tracker because it was easy to use, *"using GazeCast (handheld) is easier than using a webcam"*(P17), *"using GazeCast (handheld) was the easiest, as I could hold the device at my comfort height"* (P19). For participants' preferences, ten participants reported that they would prefer a mobile phone as an eye tracker over a webcam for privacy and security reasons, as reported by four participants, *"Although it is more comfortable to go with a built-in eye-tracker, I prefer to use my phone as an eye-tracker for security reasons"*(P14). Finally, for the ranking of the tracking setups, after finishing all tasks, We asked the participants to rank the tracking setups based on their preference, we found that GazeCast(handheld), and webcam were preferred equally (8 out of 20), then GazeCast (on a stand) with four votes.

6 DISCUSSION

We investigated the potential of using a mobile device's front-facing camera as an eye tracker to interact with public displays. In the following, we discuss the implications of using GazeCast for gaze-based interaction with public displays.

6.1 Accuracy and Tracking Setups

Comparing our three tracking setups, the webcam one had the highest error and false detection rate (35 out of 360 selections), compared to 18 for GazeCast (handheld) and 15 for GazeCast (on a stand). Demonstrating that in our implementation, using a smartphone decreases inaccuracy by over 50% and outperforms using a webcam. This might be due to the distance between the user and the webcam (60 cm), which makes gaze estimation difficult, especially when the webcam has a lower resolution than the mobile device. Finally, due to the smaller distance when participants hold the phone with their hand (around 30 cm) the error rate was almost negligible. Another possible reason is that, because users are accustomed to apps that utilize the front-facing camera, they know how to hold their phone in a way that ensures their face is captured by the front-facing camera. On the downside, holding the mobile device for a long time results in a shaky setup, which in turn slows selection compared to the webcam setup. However, the interaction with public displays is normally short, around 8 seconds [Gentile et al. 2020; Müller et al. 2010] which won't require huge physical demand. Our findings show that utilizing a mobile device for gaze interaction while holding it works better for a shorter period of time. This aligns with our anticipated use cases for GazeCast, where users interact shortly with public displays to obtain information.

6.2 Ubiquitous and Privacy-preserving Gaze Interaction

The novel idea of GazeCast enables the ubiquitous use of eye-gaze interaction on different interfaces. Even if it is anticipated that eye trackers will soon be incorporated into most devices, this would require hardware integration which is costly. GazeCast allows gaze interaction with public displays with minimal costs for the display owner, and leveraging personal mobile devices, which are already ubiquitous. GazeCast would have promising applications in shopping centers and in airports where public info displays are common, and on ATMs.

However, with gaze interaction becoming ubiquitous, users' privacy is at risk. Privacy is particularly important in this context given the rich information content available in human eye movement [Kröger et al. 2020]. several user characteristics, including age, gender, mental illnesses, personality traits, and more, can be inferred from users' collected gaze data. On the other hand, the majority of users do not recognize the risks of sharing their gaze data, therefore they are willing to share their data with multiple parties [Steil et al. 2019]. In our implementation, the processing was done on the server to ensure a fair comparison with the webcam condition, but the GazeCast concept can enable privacy-preserving gaze-based interaction as processing could be done on the user's phone. This would prevent gaze data from being shared with a server or a third party. This shows that GazeCast is an effective and low-cost solution to enable gaze-based interaction on public displays while maintaining users' privacy.

6.3 Supporting Multiple Users

The GazeCast concept allows multiple users to interact with the same screen at the same time, which is a common requirement for pervasive public displays [Memarovic et al. 2014]. Although in our experiment one user interacted at a time, below we discuss how our implementation can be extended to support multiple users in terms of 1) connection methods and 2) feedback and screen indicators. For the connection methods, there are various strategies for connecting with the public display. The simplest method we used in our implementation was QR codes. However, since the majority of smartphones have NFC readers, it is also possible to utilize other types of technology, such as NFC tags or BLE beacons [Mäkelä et al. 2017]. To allow multiple users to connect, multiple sockets can be generated randomly and sent to various devices to establish connections. For the feedback and screen indicators, when multiple users interact with the same display, users need a way to distinguish their controls. Previous work visualized the users' gaze and employed different colours to differentiate users' gaze cursors [Zhang et al. 2017b]. Another approach could be to split the screen into multiple sections, one per user. Similarly, each section can be mapped to a colour that is mapped to a particular user. On the downside, this may impact the accuracy of the interface as the targets will have to be smaller and potentially closer to each other to fit into one display. Public displays researchers studied ways to support users in identifying their on-screen representation [Khamis et al. 2018b], but not for gaze interfaces which open future work directions.

7 LIMITATIONS

As the work aims to investigate GazeCast as a concept, we did not explore the effect of the holding posture, face visibility, and lighting conditions which might affect users' experiences. However, these are ongoing research topics in mobile eye tracking that are forecasted to be addressed in the near future, and thus we leave for future work. Also, we attempted to maintain consistency in the distance between participants and the camera by using the same distance for both the Webcam and GazeCast (on stand). In contrast, for GazeCast (handheld), we told the participants to hold the phone they would do naturally. However, it is important to acknowledge that the different camera resolutions may have introduced a potential confounding factor that could impact the accuracy of our results.

8 CONCLUSION AND FUTURE WORK

We presented GazeCast, a novel system that leverages users' personal handheld devices to enable gaze-based interaction with public displays. We used qualitative and quantitative measures to evaluate GazeCast. While the results show that the selection time and perceived workload increase with GazeCast, users' feedback suggested that they appreciate how GazeCast allows for flexible positioning and potential privacy benefits when interacting with public displays. Future work can explore how many concurrent users can use the display without affecting the detection accuracy. Another interesting direction is to investigate the different user distances and sweet spots in relation to the detection accuracy.

ACKNOWLEDGMENTS

This work has been funded by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which has been funded by the UK EPSRC under grant number EP/S035362/1, and an EPSRC New Investigator award (EP/V008870/1), the Islamic university of Madinah, the Royal Society of Edinburgh (RSE) grant number 1931. A. Bulling was funded by the European Research Council (ERC; grant agreement 801708).

REFERENCES

Hawkeye Access. 2019. Hawkeye Access - Control your iOS device with your eyes. Webpage. Retrieved November 4, 2022 from https://www.usehawkeye.com/accessibility

Florian Alt, Andreas Bulling, Gino Gravanis, and Daniel Buschek. 2015. GravitySpot: Guiding Users in Front of Public Displays Using On-Screen Visual Cues. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 47–56. https://doi.org/10.1145/2807442.2807490

Florian Alt, Andreas Bulling, Lukas Mecke, and Daniel Buschek. 2016. Attention, Please!: Comparing Features for Measuring Audience Attention Towards Pervasive Displays. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) *(DIS '16)*. ACM, New York, NY, USA, 823–828. https://doi.org/10.1145/2901790.2901897

Efe Bozkir, Onur Günlü, Wolfgang Fuhl, Rafael F. Schaefer, and Enkelejda Kasneci. 2021. Differential privacy for eye tracking with temporal correlations. *PLOS ONE* 16, 8 (08 2021), 1–22. https://doi.org/10.1371/journal.pone.0255979

Efe Bozkir, Ali Burak Ünal, Mete Akgün, Enkelejda Kasneci, and Nico Pfeifer. 2020. Privacy Preserving Gaze Estimation Using Synthetic Images via a Randomized Encoding Based Framework. In *ACM Symposium on Eye Tracking Research and Applications* (Stuttgart, Germany) *(ETRA '20 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 21, 5 pages. https://doi.org/10.1145/3379156.3391364

Andreas Bulling and Kai Kunze. 2016. Eyewear Computers for Human-computer Interaction. *interactions* 23, 3 (April 2016), 70–73. https://doi.org/10.1145/2912886

Brendan David-John, Kevin Butler, and Eakta Jain. 2022. For Your Eyes Only: Privacy-Preserving Eye-Tracking Datasets. In *2022 Symposium on Eye Tracking Research and Applications* (Seattle, WA, USA) *(ETRA '22)*. Association for Computing Machinery, New York, NY, USA, Article 10, 6 pages. https://doi.org/10.1145/3517031.3529618

Nigel Davies, Sarah Clinch, and Florian Alt. 2014. *Pervasive Displays: Understanding the Future of Digital Signage* (1st ed.). Morgan & Claypool Publishers.

Heiko Drewes, Mohamed Khamis, and Florian Alt. 2018. Smooth Pursuit Target Speeds and Trajectories. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia* (Cairo, Egypt) *(MUM 2018)*. Association for Computing Machinery, New York, NY, USA, 139–146. https://doi.org/10.1145/3282894.3282913

Heiko Drewes, Mohamed Khamis, and Florian Alt. 2019. DialPlates: Enabling Pursuits-Based User Interfaces with Large Target Numbers. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia* (Pisa, Italy) *(MUM '19)*. Association for Computing Machinery, New York, NY, USA, Article 10, 10 pages. https://doi.org/10.1145/3365610.3365626

Utkarsh Dwivedi, Karan Ahuja, Rahul Islam, Ferdous A. Barbhuiya, Seema Nagar, and Kuntal Dey. 2017. EyamKayo: Interactive Gaze and Facial Expression Captcha. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion* (Limassol, Cyprus) *(IUI '17 Companion)*. Association for Computing Machinery, New York, NY, USA, 53–56. https://doi.org/10.1145/3030024.3038266

Augusto Esteves, Yonghwan Shin, and Ian Oakley. 2020. Comparing selection mechanisms for gaze input techniques in head-mounted displays. *International Journal of Human-Computer Studies* 139 (2020), 102414.

Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 457–466. https://doi.org/10.1145/2807442.2807499

Vito Gentile, Mohamed Khamis, Fabrizio Milazzo, Salvatore Sorce, Alessio Malizia, and Florian Alt. 2020. Predicting mid-air gestural interaction with public displays based on audience behaviour. *International Journal of Human-Computer Studies* 144 (2020), 102497. https://doi.org/10.1016/j.ijhcs.2020.102497

Benjamin Hatscher, Maria Luz, Lennart E. Nacke, Norbert Elkmann, Veit Müller, and Christian Hansen. 2017. GazeTap: Towards Hands-Free Interaction in the Operating Room. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction* (Glasgow, UK) *(ICMI '17)*. Association for Computing Machinery, New York, NY, USA, 243–251. https://doi.org/10.1145/3136755.3136759

Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Adj. Proc. of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2014)*. 1151–1160. https://doi.org/10.1145/2638728.2641695

Mohamed Khamis, Florian Alt, and Andreas Bulling. 2018a. The Past, Present, and Future of Gaze-Enabled Handheld Mobile Devices: Survey and Lessons Learned. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Barcelona, Spain) *(MobileHCI '18)*. Association for Computing Machinery, New York, NY, USA, Article 38, 17 pages. https://doi.org/10.1145/3229434.3229452

Mohamed Khamis, Christian Becker, Andreas Bulling, and Florian Alt. 2018b. Which one is me? Identifying Oneself on Public Displays.. In *Proceedings of the 36th Annual ACM Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '18)*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3152832.3157813

Mohamed Khamis, Regina Hasholzner, Andreas Bulling, and Florian Alt. 2017a. GTmoPass: Two-Factor Authentication on Public Displays Using Gaze-Touch Passwords and Personal Mobile Devices. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays* (Lugano, Switzerland) *(PerDis '17)*. Association for Computing Machinery, New York, NY, USA, Article 8, 9 pages. https://doi.org/10.1145/3078810.3078815

Mohamed Khamis, Axel Hoesl, Alexander Klimczak, Martin Reiss, Florian Alt, and Andreas Bulling. 2017b. EyeScout: Active Eye Tracking for Position and Movement Independent Gaze Interaction with Large Public Displays. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 155–166. https://doi.org/10.1145/3126594.3126630

Mohamed Khamis, Karola Marky, Andreas Bulling, and Florian Alt. 2022. User-centred multimodal authentication: securing handheld mobile devices using gaze and touch input. *Behaviour & Information Technology* 41, 10 (2022), 2061–2083. https://doi.org/10.1080/0144929X.2022.2069597 arXiv:https://doi.org/10.1080/0144929X.2022.2069597

Mohamed Khamis, Ludwig Trotter, Ville Mäkelä, Emanuel von Zezschwitz, Jens Le, Andreas Bulling, and Florian Alt. 2018c. CueAuth: Comparing Touch, Mid-Air Gestures, and Gaze for Cue-Based Authentication on Situated Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 4, Article 174 (dec 2018), 22 pages. https://doi.org/10.1145/3287052

Mohamed Khamis, Ludwig Trotter, Markus Tessmann, Christina Dannhart, Andreas Bulling, and Florian Alt. 2016. EyeVote in the Wild: Do Users Bother Correcting System Errors on Public Displays?. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (Rovaniemi, Finland) *(MUM '16)*. Association for Computing Machinery, New York, NY, USA, 57–62. https://doi.org/10.1145/3012709.3012743

Andy Kong, Karan Ahuja, Mayank Goel, and Chris Harrison. 2021. EyeMU Interactions: Gaze+ IMU Gestures on Mobile Devices. In *Proceedings of the 2021 International Conference on Multimodal Interaction*. 577–585.

Jacob Leon Kröger, Otto Hans-Martin Lutz, and Florian Müller. 2020. What does your gaze reveal about you? On the privacy implications of eye tracking. *Privacy and Identity Management. Data for Better Living: AI and Privacy: 14th IFIP WG 9.2, 9.6/11.7, 11.6/SIG 9.2. 2 International Summer School, Windisch, Switzerland, August 19–23, 2019, Revised Selected Papers 14* (2020), 226–241.

Christian Lander, Sven Gehring, Antonio Krüger, Sebastian Boring, and Andreas Bulling. 2015a. GazeProjector: Accurate Gaze Estimation and Seamless Gaze Interaction Across Multiple Displays. In *Proc. of the 28th ACM Symposium on User Interface Software and Technology (UIST 2015)* (2015-11-01). https://perceptual.mpi-inf.mpg.de/files/2015/08/Lander_UIST15.pdf

Christian Lander, Marco Speicher, Denise Paradowski, Norine Coenen, Sebastian Biewer, and Antonio Krüger. 2015b. Collaborative Newspaper: Exploring an Adaptive Scrolling Algorithm in a Multi-User Reading Scenario. In *Proceedings of the 4th International Symposium on Pervasive Displays* (Saarbruecken, Germany) *(PerDis '15)*. Association for Computing Machinery, New York, NY, USA, 163–169. https://doi.org/10.1145/2757710.2757734

Jong-In Lee, Sunjun Kim, Masaaki Fukumoto, and Byungjoo Lee. 2017. Reflector: Distance-Independent, Private Pointing on a Reflective Screen. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 351–364. https://doi.org/10.1145/3126594.3126665

Yinghui Li, Zhichao Cao, and Jiliang Wang. 2017. Gazture: Design and Implementation of a Gaze Based Gesture Control System on Tablets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 74 (Sept. 2017), 17 pages. https://doi.org/10.1145/3130939

Dachuan Liu, Bo Dong, Xing Gao, and Haining Wang. 2015. *Exploiting Eye Tracking for Smartphone Authentication*. Springer International Publishing, Cham, 457–477. https://doi.org/10.1007/978-3-319-28166-7_22

Sven Mayer, Gierad Laput, and Chris Harrison. 2020. Enhancing Mobile Voice Assistants with WorldGaze. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3313831.3376479

Nemanja Memarovic, Marc Langheinrich, and Ava Fatah. 2014. Community is the Message: Viewing Networked Public Displays Through McLuhan's Lens of Figure and Ground. In *Proceedings of the 2Nd Media Architecture Biennale Conference: World Cities* (Aarhus, Denmark) *(MAB '14)*. ACM, New York, NY, USA, 30–33. https://doi.org/10.1145/2682884.2682891

Emiliano Miluzzo, Tianyu Wang, and Andrew T. Campbell. 2010. EyePhone: Activating Mobile Phones with Your Eyes. In *Proceedings of the Second ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds* (New Delhi, India) *(MobiHeld '10)*. Association for Computing Machinery, New York, NY, USA, 15–20. https://doi.org/10.1145/1851322.1851328

Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. 2010. Requirements and Design Space for Interactive Public Displays. In *Proceedings of the 18th ACM International Conference on Multimedia* (Firenze, Italy) *(MM '10)*. Association for Computing Machinery, New York, NY, USA, 1285–1294. https://doi.org/10.1145/1873951.1874203

Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. ACM, New York, NY, USA, 297–306. https://doi.org/10.1145/2207676.2207718

Ville Mäkelä, Jobin James, Tuuli Keskinen, Jaakko Hakulinen, and Markku Turunen. 2017. "It's Natural to Grab and Pull": Retrieving Content from Large Displays Using Mid-Air Gestures. *IEEE Pervasive Computing* 16, 3 (2017), 70–77. https://doi.org/10.1109/MPRV.2017.2940966

Vijay Rajanna and Tracy Hammond. 2018. A Fitts' Law Evaluation of Gaze Input on Large Displays Compared to Touch and Mouse Inputs. In *Proceedings of the Workshop on Communication by Gaze Interaction* (Warsaw, Poland) *(COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, Article 8, 5 pages. https://doi.org/10.1145/3206343.3206348

Javier San Agustin, John Paulin Hansen, and Martin Tall. 2010. Gaze-Based Interaction with Public Displays Using off-the-Shelf Components. In *Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct* (Copenhagen, Denmark) *(UbiComp '10 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 377–378. https://doi.org/10.1145/1864431.1864444

SeeSo. 2022. SeeSo : Eye Tracking Software. Webpage. Retrieved November 4, 2022 from https://seeso.io/

Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of Eye Gaze Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. ACM, New York, NY, USA, 281–288. https://doi.org/10.1145/332040.332445

Julian Steil, Inken Hagestedt, Michael Xuelin Huang, and Andreas Bulling. 2019. Privacy-Aware Eye Tracking Using Differential Privacy. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 27, 9 pages. https://doi.org/10.1145/3314111.3319915

Eduardo Velloso, Marcus Carter, Joshua Newn, Augusto Esteves, Christopher Clarke, and Hans Gellersen. 2017. Motion Correlation: Selecting Objects by Matching Their Movement. *ACM Trans. Comput.-Hum. Interact.* 24, 3, Article 22 (apr 2017), 35 pages. https://doi.org/10.1145/3064937

Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving

Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 439–448. https://doi.org/10.1145/2493432.2493477

Xiaoyi Zhang, Harish Kulkarni, and Meredith Ringel Morris. 2017a. Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 2878–2889. https://doi.org/10.1145/3025453.3025790

Yanxia Zhang, Andreas Bulling, and Hans Gellersen. 2013. SideWays: A Gaze Interface for Spontaneous Interaction with Situated Displays. In *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. 851–860. https://doi.org/10.1145/2470654.2470775

Yanxia Zhang, Ken Pfeuffer, Ming Ki Chong, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2017b. Look together: using gaze for assisting co-located collaborative search. *Personal and Ubiquitous Computing* 21, 1 (2017), 173–186.