Gaze-based Interaction on Handheld Mobile Devices

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

ABSTRACT

With the advancement of smartphone technology, it is now possible for smartphones to run eye-tracking using the front-facing camera, enabling hands-free interaction by empowering mobile users with novel gaze-based input techniques. While several gaze-based interaction techniques have been proposed in the literature, these techniques were deployed in settings different from daily gaze interaction with mobile devices, posing several unique challenges. The user's holding posture may hinder the camera's view of their face during the interaction, the front-facing camera may be obstructed by the user's clothing or hands, or the environment is shaky due to the user's movements and the dynamic environment. This PhD research investigates the usability of state-of-the-art gaze-based input techniques in mobile settings, develops a novel concept of combining multiple gaze-based techniques, and addresses the challenges imposed by the unique aspects of these devices.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Interaction techniques**; **Mobile devices**.

KEYWORDS

Eye tracking, Gaze interaction, Mobile devices

ACM Reference Format:

Omar Namnakani. 2023. Gaze-based Interaction on Handheld Mobile Devices. In *2023 Symposium on Eye Tracking Research and Applications (ETRA '23), May 30–June 02, 2023, Tubingen, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3588015.3589540

1 RESEARCH OBJECTIVES

The recent advancement in smartphone manufacturing has led to fast processors and more powerful front-facing cameras, enabling eye-tracking applications on mobile devices [Dybdal et al. 2012; Khamis et al. 2018a]. This includes providing novel gaze-based input techniques to mobile consumers that allow for hands-free interaction.

Input using gaze has much potential for fast interaction [Majaranta and Bulling 2014; Ramirez Gomez et al. 2021; Sibert and Jacob 2000], requires no training as it is natural to look at what we are interested in [Majaranta and Bulling 2014; Plopski et al. 2022; Zhai 2003], and also requires less energy and efforts compared to input with heads or hands [Ramirez Gomez et al. 2021; Sidenmark and Gellersen 2019]. Depending on the user's abilities, context, and preference, gaze input can be a feasible alternative input technique. [Kumar et al. 2007]. It is compelling for a wide range of contexts, such as while users are holding something in their hands, wearing gloves, or when it is raining, and touch is not viable [Kong et al. 2021; Lutteroth et al. 2015; Majaranta and Bulling 2014; Ramirez Gomez et al. 2021; Sibert and Jacob 2000].

There are two main objectives for this PhD. First, assessing the usability and suitability of state-of-the-art gaze-based techniques in mobile settings and developing novel ones that combine multiple techniques. This will include exploring ways to design novel User interface elements optimised for mobile gaze interaction to improve its general usability. Second, for eye tracking to become a ubiquitous part of our everyday mobile interactions, we will identify and address the novel challenges of gaze-based interaction techniques on mobile devices to support the design of highly usable interfaces for mobile contexts.

2 PROBLEM STATEMENT AND HYPOTHESIS

Most HCI research studies on gaze interaction were deployed in settings that are different from daily gaze interaction with mobile devices. For example, in controlled and static setups such as on interactive surfaces [Zhang et al. 2017], desktop machines [Rajanna et al. 2018], public displays [Khamis et al. 2017, 2016], mobile devices held by a mount [Dybdal et al. 2012; Khamis et al. 2018a], wearables [Esteves et al. 2015], and head-mounted displays [Esteves et al. 2020; Hirzle et al. 2019; Khamis et al. 2018c]. However, the study of gaze tracking on mobile devices introduces various challenges absent in other eye-tracking methods. The front-facing camera of handheld mobile devices, which is used to track eye movements, may be obstructed by clothing or hands, [Huang et al. 2017b; Khamis et al. 2018b]. Additionally, the instability of the device and the dynamic environment, such as when a user is walking or in a moving vehicle, can also affect gaze tracking accuracy. Studies have also shown that users hold their mobile devices in various ways and postures, and full-face visibility is not always present for effective gaze estimation [Huang et al. 2017a; Khamis et al. 2018b]. This PhD focuses on answering the following research questions:

- **RQ1** How do gaze-based interaction techniques compare to each other in mobile settings?
- **RQ2** How can we improve the accuracy of gaze-based input techniques?
- **RQ3** How can we address the imposed challenges brought by natural mobile settings when using gaze-based input techniques?
- **RQ4** What are the best practices for mobile gaze interface design?

ETRA '23, May 30–June 02, 2023, Tubingen, Germany

ACM ISBN 979-8-4007-0150-4/23/05…$15.00
https://doi.org/10.1145/3588015.3589540

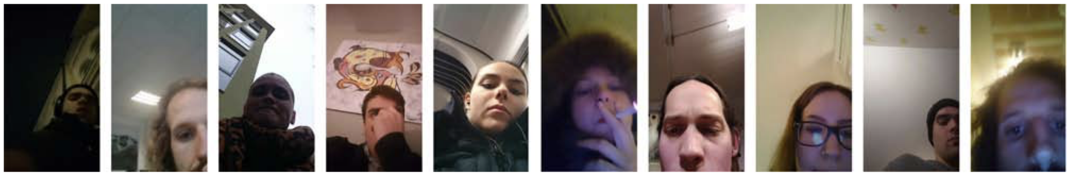

Figure 1: In a dataset collected in the wild using smartphone front-facing cameras, the face was visible in only 29% of the cases, while the eyes, but not the whole face, were visible in 48% of the cases [Khamis et al. 2018b].

3 APPROACH AND METHOD

This PhD will involve designing, conducting, analysing, and reporting experiments to demonstrate use cases and to aid the tasks at hand. I break down my PhD research into three main components.

The first component will explore how the current gaze-based interaction techniques perform in mobile settings and present novel concepts of combining multiple gaze-based techniques to enhance the interaction. To assess the usability of the interaction techniques in mobile settings, we selected three of the most widely used gaze-based interaction techniques; Dwell time, Pursuits, and Gaze gestures. While other interaction techniques are available in the literature, these techniques were argued to be promising for mobile context [Khamis et al. 2018a]. Through an experiment, we evaluated the performance of the techniques directly on mobile devices, using front-facing cameras while the users are stationary and on the move (See Section 4.1). The results of this study contribute to research question **RQ1**. On the other hand, informed by design guidelines resulting from the first experiment, we will present and develop a novel concept of combining multiple gaze input techniques to enhance gaze-based interaction—for example, we will investigate the acceptability and performance of using Pursuits for selection and using Gaze gestures for swiping left and right compared to using Pursuits only, or using Dwell time when the context is walking and switching to Pursuits when the user is stationary. This will also include comparisons with baseline (Dwell time or all techniques when used independently) to understand users' preferences better. All this contributes to answering the research question **RQ2**.

The second component's primary focus will be addressing unique challenges when enabling eye tracking on mobile devices. Since Eye tracking on mobile devices uses front-facing cameras and assume that users' face and eyes are visible, an ideal holding posture is required (See Figure 1). In a user study, we will evaluate different guiding methods to mitigate the impact of the holding posture challenge when interacting with mobile via gaze. This study is motivated by prior work [Alt et al. 2015; Khamis et al. 2018a,b; Zhang et al. 2014]. On the other hand, prior work also suggested a lack of uniformity in phone-holding posture [Huang et al. 2017a]. These changes affect the performance of gaze input techniques, mainly Dwell time, as it relies on calibration and accurate gaze estimates [Khamis et al. 2018a]. Even if the user holds the device in the same way, the inevitable shaking of the device will make the calibration data obsolete and re-calibration necessary. In a study conducted in the wild, we will attempt to mitigate such issues by automatically compensating for the changed posture using the internal sensing data or camera to update calibration parameters without re-calibration. Results from studies will contribute to research questions **RQ3**.

Based on our experiences in implementing and evaluating gaze-based interaction techniques, the last component will focus on formulating design guidelines for supporting effective gaze-based interaction. Moreover, we will investigate how to replicate existing touch-based gestures to be entered by gaze [app 2022]. We also plan to explore the design of novel UI elements optimised for mobile gaze interaction to improve its general usability. Such interfaces should minimise the learning process as much as possible with dedicated UI elements to be displayed and by unifying the interfaces for gaze input, such as menus and buttons, considering how touchscreen devices are typically used [Huckauf and Urbina 2008; Rivu et al. 2019; Skovsgaard et al. 2010]. All this contributes to answering the research question **RQ4**.

4 CURRENT STATUS AND FUTURE WORK

4.1 Current Status

We created a prototype application and conducted a user study to evaluate the performance of Dwell time, Pursuits, and Gaze gestures in mobile settings [Namnakani et al. 2023]. We recruited 24 participants, and their task was to select one of 2, 4, 9, 12 and 32 targets via gaze while sitting and walking using all three techniques. The results showed that participants demonstrated a preference for using Pursuits when they were seated. Pursuits was also significantly faster than the other two techniques, with an average selection time of 1.36 sec, compared to 2.33 sec for Dwell time and 5.17 sec for Gaze Gestures. On the other hand, when on the move, participants preferred the Dwell time technique, with an average selection time of 2.76 sec. Although Dwell time was slightly slower than Pursuits (2.14 sec), it was still significantly faster than Gaze Gestures, which had an average selection time of 6.68 sec. Based on the results, we concluded the paper with guidelines for gaze-based interaction. We documented the result of the study in a paper and submitted it to CHI 2023, which was accepted for publishing.

4.2 Future Work

Based on the first experiment, we observed that tracking accuracy decreases towards the screen's left edge. This could be due to the camera's position on the smartphone, which is on the center-right of the top of the device in our experiment. Prior work on gaze-based interaction showed that experiments were conducted with various phone models with varying front-facing camera positions or with external cameras positioned in varying positions [Dybdal et al. 2012; Huang et al. 2017a; Nagamatsu et al. 2010; Rozado et al. 2015;

Wood and Bulling 2014]. Hohlfeld et al. evaluated the performance of EyeTab, a gaze estimation algorithm, on a Nexus 7 tablet with a front-facing camera mounted at the bottom. They found that the tracking accuracy decreases from the top to the bottom of the screen [Hohlfeld et al. 2015]. In a user study, we plan to systematically investigate how the camera position can impact the interface design regarding target positioning. The finding of this study will inform our decision when conducting the following studies on using a combination of techniques by determining the appropriate target positions based on the camera position of the device. Then we will attempt to address some unique challenges when enabling eye tracking on mobile devices. Finally, we will conclude with practical guidelines for enabling gaze interactions on handheld mobile devices.

REFERENCES

2022. Touchscreen gestures - Human Interface Guidelines. Webpage. https://developer.apple.com/design/human-interface-guidelines/inputs/touchscreen-gestures accessed 2 October 2022.

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

Morten Lund Dybdal, Javier San Agustin, and John Paulin Hansen. 2012. Gaze Input for Mobile Devices by Dwell and Gestures. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 225–228. https://doi.org/10.1145/2168556.2168601

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

Teresa Hirzle, Jan Gugenheimer, Florian Geiselhart, Andreas Bulling, and Enrico Rukzio. 2019. A Design Space for Gaze Interaction on Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300855

Oliver Hohlfeld, André Pomp, Jó Ágila 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* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 427–434. https://doi.org/10.1145/2785830.2785869

Michael Xuelin Huang, Jiajia Li, Grace Ngai, and Hong Va Leong. 2017a. ScreenGlint: Practical, In-Situ Gaze Estimation on Smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2546–2557. https://doi.org/10.1145/3025453.3025794

Qiong Huang, Ashok Veeraraghavan, and Ashutosh Sabharwal. 2017b. TabletGaze: dataset and analysis for unconstrained appearance-based gaze estimation in mobile tablets. *Machine Vision and Applications* 28, 5 (2017), 445–461.

Anke Huckauf and Mario H. Urbina. 2008. Gazing with PEYEs: Towards a Universal Input for Various Applications. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications* (Savannah, Georgia) *(ETRA '08)*. Association for Computing Machinery, New York, NY, USA, 51–54. https://doi.org/10.1145/1344471.1344483

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*. 1–17.

Mohamed Khamis, Anita Baier, Niels Henze, Florian Alt, and Andreas Bulling. 2018b. Understanding Face and Eye Visibility in Front-Facing Cameras of Smartphones Used in the Wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173854

Mohamed Khamis, Alexander Klimczak, Martin Reiss, Florian Alt, and Andreas Bulling. 2017. 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 & Technology* (Quebec City, QC, Canada) *(UIST '17)*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3126594.3126630

Mohamed Khamis, Carl Oechsner, Florian Alt, and Andreas Bulling. 2018c. VRpursuits: Interaction in Virtual Reality Using Smooth Pursuit Eye Movements. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces* (Castiglione della Pescaia, Grosseto, Italy) *(AVI '18)*. Association for Computing Machinery, New York, NY, USA, Article 18, 8 pages. https://doi.org/10.1145/3206505.3206522

Mohamed Khamis, Ozan Saltuk, Alina Hang, Katharina Stolz, Andreas Bulling, and Florian Alt. 2016. TextPursuits: Using Text for Pursuits-Based Interaction and Calibration on Public Displays. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 274–285. https://doi.org/10.1145/2971648.2971679

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.

Manu Kumar, Andreas Paepcke, and Terry Winograd. 2007. EyePoint: Practical Pointing and Selection Using Gaze and Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 421–430. https://doi.org/10.1145/1240624.1240692

Christof Lutteroth, Moiz Penkar, and Gerald Weber. 2015. Gaze vs. Mouse: A Fast and Accurate Gaze-Only Click Alternative. 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, 385–394. https://doi.org/10.1145/2807442.2807461

Päivi Majaranta and Andreas Bulling. 2014. Eye tracking and eye-based human–computer interaction. In *Advances in physiological computing*. Springer, 39–65.

Takashi Nagamatsu, Michiya Yamamoto, and Hiroshi Sato. 2010. MobiGaze: Development of a Gaze Interface for Handheld Mobile Devices. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI EA '10)*. Association for Computing Machinery, New York, NY, USA, 3349–3354. https://doi.org/10.1145/1753846.1753983

Omar Namnakani, Yasmeen Abdrabou, Jonathan Grizou, Augusto Esteves, and Mohamed Khamis. 2023. Comparing Dwell time, Pursuits and Gaze Gestures for Gaze Interaction on Handheld Mobile Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3544548.3580871

Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. 2022. The Eye in Extended Reality: A Survey on Gaze Interaction and Eye Tracking in Head-Worn Extended Reality. *ACM Comput. Surv.* 55, 3, Article 53 (mar 2022), 39 pages. https://doi.org/10.1145/3491207

Vijay Rajanna, Adil Hamid Malla, Rahul Ashok Bhagat, and Tracy Hammond. 2018. DyGazePass: A gaze gesture-based dynamic authentication system to counter shoulder surfing and video analysis attacks. In *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)*. 1–8. https://doi.org/10.1109/ISBA.2018.8311458

Argenis Ramirez Ramirez Gomez, Christopher Clarke, Ludwig Sidenmark, and Hans Gellersen. 2021. Gaze+Hold: Eyes-Only Direct Manipulation with Continuous Gaze Modulated by Closure of One Eye. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Full Papers)*. Association for Computing Machinery, New York, NY, USA, Article 10, 12 pages. https://doi.org/10.1145/3448017.3457381

Sheikh Rivu, Yasmeen Abdrabou, Thomas Mayer, Ken Pfeuffer, and Florian Alt. 2019. GazeButton: Enhancing Buttons with Eye Gaze Interactions. 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 73, 7 pages. https://doi.org/10.1145/3317956.3318154

David Rozado, T Moreno, Javier San Agustin, FB Rodriguez, and Pablo Varona. 2015. Controlling a smartphone using gaze gestures as the input mechanism. *Human–Computer Interaction* 30, 1 (2015), 34–63.

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)*. Association for Computing Machinery, New York, NY, USA, 281–288. https://doi.org/10.1145/332040.332445

Ludwig Sidenmark and Hans Gellersen. 2019. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.* 27, 1, Article 4 (dec 2019), 40 pages. https://doi.org/10.1145/3361218

Henrik Skovsgaard, Julio C. Mateo, John M. Flach, and John Paulin Hansen. 2010. Small-Target Selection with Gaze Alone. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (Austin, Texas) *(ETRA '10)*. Association for Computing Machinery, New York, NY, USA, 145–148. https://doi.org/10.1145/1743666.1743702

Erroll Wood and Andreas Bulling. 2014. EyeTab: Model-Based Gaze Estimation on Unmodified Tablet Computers. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Safety Harbor, Florida) *(ETRA '14)*. Association for Computing Machinery, New York, NY, USA, 207–210. https://doi.org/10.1145/

2578153.2578185

Shumin Zhai. 2003. What's in the Eyes for Attentive Input. *Commun. ACM* 46, 3 (mar 2003), 34–39. https://doi.org/10.1145/636772.636795

Yanxia Zhang, Jörg Müller, Ming Ki Chong, Andreas Bulling, and Hans Gellersen. 2014. GazeHorizon: Enabling Passers-by to Interact with Public Displays by Gaze. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Seattle, Washington) *(UbiComp '14)*. Association for Computing Machinery, New York, NY, USA, 559–563. https://doi.org/10.1145/2632048.2636071

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