Multi-user Gaze-based Interaction Techniques on Collaborative Touchscreens

Ken Pfeuffer
Bundeswehr University Munich
Germany
ken.pfeuffer@unibw.de

Jason Alexander
University of Bath
UK
jma73@bath.ac.uk

Hans Gellersen
Aarhus University
Denmark
hwg@cs.au.dk

ABSTRACT

Eye-gaze is a technology for implicit, fast, and hands-free input for a variety of use cases, with the majority of techniques focusing on single-user contexts. In this work, we present an exploration into gaze techniques of users interacting together on the same surface. We explore interaction concepts that exploit two states in an interactive system: 1) users visually attending to the same object in the UI, or 2) users focusing on separate targets. Interfaces can exploit these states with increasing availability of eye-tracking. For example, to dynamically personalise content on the UI to each user, and to provide a merged or compromised view on an object when both users' gaze are falling upon it. These concepts are explored with a prototype horizontal interface that tracks gaze of two users facing each other. We build three applications that illustrate different mappings of gaze to multi-user support: an indoor map with gaze-highlighted information, an interactive tree-of-life visualisation that dynamically expands on users' gaze, and a worldmap application with gaze-aware fisheye zooming. We conclude with insights from a public deployment of this system, pointing toward the engaging and seamless ways how eye based input integrates into collaborative interaction.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Pointing**; **Collaborative interaction**.

KEYWORDS

gaze input, eye-tracking, collaboration, multi-user interaction, shared user interface

ACM Reference Format:
Ken Pfeuffer, Jason Alexander, and Hans Gellersen. 2021. Multi-user Gaze-based Interaction Techniques on Collaborative Touchscreens. In *ETRA '21: 2021 Symposium on Eye Tracking Research and Applications (ETRA '21 Short Papers), May 25–27, 2021, Virtual Event, Germany*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3448018.3458016

1 INTRODUCTION

Eye gaze input is long established in HCI [Bolt 1981]. The basic technique enables to simply look at a target of interest to enter a command, and to adapt the content to one's visual interest [Bolt 1981]. The technology assists the user with hands-free, covert and hygienic input with low effort [Pfeuffer and Gellersen 2016; Zhai et al. 1999], that can be useful for instance in public settings [Zhang et al. 2014]. However, while gaze has been widely studied as an explicit interaction modality, this has mostly focused on single-user interfaces.

For manual interfaces, there exists a body of literature on collaborative interaction techniques [Hinckley 2003; Hinckley et al. 1997; Morris et al. 2006, 2004]. Typically, an interactive object in the UI is designed to map to a single UI command, independent of how many users are providing input toward it. In contrast, these works show that interactive systems can be enriched when considering a distinction between input by a single or multiple users. For example, objects of the UI can adapt to the users' combined input in order to enable more expressive command invocation or to provide personalised views on the content.

Past work on multi-user eye-tracking focused on the implicit gaze awareness technique between users [D'Angelo and Gergle 2018; Newn et al. 2017; Siirtola et al. 2019; Steptoe et al. 2009; Zhang et al. 2017]. This offers, for instance, inferences of the intention and current activity of the collaborators [Newn et al. 2017; Zhang et al. 2017]. However, this leaves out the opportunities for more interactive use, where users explicitly direct their gaze, and the UI adapts according to single- or multi-user input. For gaze, only little work exists with regards to collaborative interaction techniques. Early work suggested theoretical concepts [Holman 2007]. A prior work of us focused on a game application [Pfeuffer et al. 2016], which we extend to a more general consideration of gaze interaction for multi-user UI.

In this paper, we investigate interaction techniques that exploit multiple users' gaze to enhance the interaction on a shared display. Figure 1 illustrates three cases where the system can adapt to a particular gaze interaction to the advantage of the users. First, targets can be personalised to users when each user views separate

---

ETRA '21 Short Papers, May 25–27, 2021, Virtual Event, Germany

ACM ISBN 978-1-4503-8345-5/21/05…$15.00
https://doi.org/10.1145/3448018.3458016

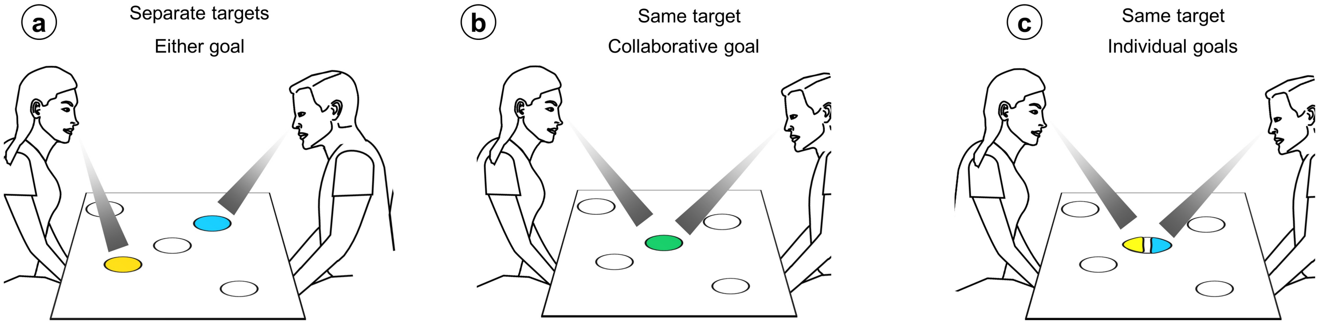

Figure 1: Collaboration surfaces enhanced with eye-tracking offer new interaction states based on gaze, exploitable for UI adaptation beneficial to the users. For example, to personalise content and automatically orient text to the viewer (a). It also affords adaptive object representations when gazes converge, to enable a collaborative interaction technique (b) or compromised views (c).

targets on the same UI. For example, the text can be oriented to each user, and the appropriate language could be displayed. Second, when users view the same target, the system can enable a merged view. Thus, the detection of shared attention toward a target can be used for, e.g., an advanced system state that is in line with both users' goals. Lastly, viewing the same target can as well lead to a compromised view – the visuals could be split between the users.

We explore these types of interaction states on a shared horizontal touchscreen. Each users' gaze can be sensed through eye-trackers placed in front of the users that are located at distinct positions around the display. We explore the interaction states through three applications that demonstrate novel use cases for multi-user gaze interaction. Each adapts the target toward single or multiple viewers. For example, a map application adapts the annotations to each users' gaze, and an educational application unfolds a visualisation gradually based on both users' attention. We evaluate the applications in a public setting to get insights into feasibility, learnability, and collaboration aspects.

Our work makes the following contributions. First, we present three concepts on how multi-user gaze input can provide distinct interaction states when interacting on the same surface. Second, we present three example applications that demonstrate the concepts in example scenarios on an interactive interface for two users. Third, we present an informal feedback from a multi-user eye-tracking study in a public setting, gathering insights into learnability and usability of multi-user gaze interaction. Collectively, our work opens new perspectives to consider where eye gaze interaction in a much more integrated way into the UI, which can in principle be applied to many situations where multiple users interact on the same computer system.

2 RELATED WORK

Bolt's seminal work presented an early example gaze interaction where a set of video clips was displayed but only the one visually attended was active [Bolt 1981]. This aligns with non-command-based interfaces, where "the computer passively observes the user and provides appropriate responses" [Jacob 1993]. Zhai's MAGIC technique uses gaze to implicitly eliminate cursor movement [Zhai et al. 1999]. Gaze-contingent displays adapt information density to the user's visual focus [Duchowski 2002], and Attentive User Interfaces adapt display content to fit the user's attentive capabilities [Vertegaal 2003]. We investigate how gaze can be implicitly used in multi user scenarios. Gaze as explicit pointer is an alternative that has been long researched in HCI [Duchowski 2002; Jacob 1990; Zhai et al. 1999]. Jacob's early work on gaze interaction revealed the "Midas Touch" problem, where gaze-only methods are prone to accidental activations [Jacob 1990], which needs to be overcome by using explicit triggers such as dwell-time [Majaranta et al. 2006], mouse click [Jacob 1990; Zhai et al. 1999], or touch [Pfeuffer et al. 2014; Stellmach and Dachselt 2012].

Multimodal gaze and touch interaction received increased attention in recent times. Pfeuffer et al. for instance have explored how gaze extends multi-touch gestures [Pfeuffer et al. 2014], digital pens [Pfeuffer et al. 2015], and mobile tablets [Pfeuffer and Gellersen 2016]. Newn et al. devised techniques to reach out-of-reach targets on large tabletops [Newn et al. 2016]. Serim and Jacucci utilised gaze on touchscreens to support users when providing input at varying degrees of visual attention [Serim and Jacucci 2016]. Rivu et al. assessed how the touch button concept can be enhanced by gaze input [Rivu et al. 2019], and how text selection can be improved through gaze pointing [Rivu et al. 2020].

Eye-tracking research for multiple users focused on gaze awareness for remote collaborations. Gaze awareness depicts the user's knowledge about where other users look at and can increase workspace awareness and group coordination [Gutwin and Greenberg 2001; Ishii and Kobayashi 1992]. Particular advantages were found for video conferencing, problem solving in software programming, visual search, games, and travel planning [Brennan et al. 2008; Newn et al. 2018; Stein and Brennan 2004; Vertegaal 1999]. There is only little work focusing on multi-user interaction. Serim et al. utilised gaze (by head direction) to enable multiple users to manage information access on a vertical surface [Serim et al. 2018]. Pfeuffer et al.'s work explores multi-user gaze specifically in a shooter game [Pfeuffer et al. 2016]. Our work extends their work by a broader consideration of interaction states and exploration across use cases.

Research on interactive surfaces investigated techniques that incorporate multi-user input at the same objects, involving three cases. *First*, a conflict can arise when each user interacts on the same target. Zanella and Greenberg proposed transparent objects to enable synchronous interaction without occlusion [Zanella and Greenberg 2001]. Morris et al. developed multi-user coordination policies that approach the different types of conflict that can occur; e.g. a voting between all users can resolute a conflict [Morris et al. 2004]. *Second*, input on the same target can enable different semantics than its single-user counterpart. This can be helpful in a file-sharing scenario for collaborative techniques [Ringel et al. 2004]; e.g. when the owner and the receiver of a file both touch on the same target's center, the ownership is automatically transferred to the receiver. Collaborative Gestures [Morris et al. 2004] provide different gesture semantic when performed together, e.g. when all users drag the corner of the same image, it enlarges to the screen's background [Morris et al. 2004]. *Third*, interaction on separate targets can offer personalised views, such as private display devices for private windows [Wigdor et al. 2009], or touch actions to create personal windows on-surface [Schmidt et al. 2010; Valdes et al. 2012]. Implicit methods overlay the user's view by a special dual-view screen where different user perspectives map to different views [Karnik et al. 2012; Kim et al. 2012], or shutter glasses that augment the whole vision of the user [Agrawala et al. 1997; Lissermann et al. 2014]. In our work, we explore these three cases from the perspective of the eye gaze modality.

3 PROTOTYPE SYSTEM

We implemented a prototype system to explore example techniques and applications, and as a basis for a user study. Our prototype consists of a table-mounted 27" Acer T272 touchscreen (60x32.5cm, 1920x1080px), and two Tobii EyeX eye trackers that come with an accuracy of about 1-2° at 30Hz (Figure 2a). Each eye tracker is connected to a separate computer, of which one serves as the server and displays the content on the interface. The software is written in Java with MT4J. The applications are designed to fit potential technical limitations such as tracking range and accuracy [Newn et al. 2016], in order to be able to provide first insights into UI issues and user experience.

We adapt the Pursuit Calibration [Pfeuffer et al. 2013] method to multiple users. This enables us to calibrate users ad-hoc, important for our walk-up multi-user scenario in our evaluation. The calibration takes 10 seconds and the procedure is performed separately for each user. At any time during use, the calibration for either user can be triggered by a button press of the experimenter. This will instance a moving target to pursue with the eyes to collect eye data to establish a calibration mapping.

To avoid jittery cursor movement, we used a simplistic fixation detection where gaze data is sampled for 150ms for fast eye movements (>75px or 2.3cm between two frames at 30 Hz), while raw data is used when users rapidly move across the screen. In addition, applications that are based on discrete objects use target assistance/snapping (as in [Turner et al. 2014]). Thus, when users look closely enough to a target (≈100px or 3.1cm), the system automatically locks on to it; the closer target is selected when multiple objects intersect.

4 APPLICATIONS

We now present example applications. Each demonstrates multiple types of the multi-user states presented in Figure 1.

4.1 Gaze-aware Map

This application demonstrates how implicit gaze interaction can aid users with individual goals. A typical map has a fixed orientation, requiring users to be located side by side, or to perform frequent reorientation. In this gaze-aware map application, the content is implicitly adapted towards each user when they view the same or view separate targets. We envision such an application in foyers of buildings such as a hospital, university, or a museum to provide users with information about building structure and organisation. The whole table surface displays a two-dimensional map of the building (Figure 2a). The map consists of many rooms, such as a laboratory, radiology, or cafeteria. These rooms can have additional information, such as the name, description, people involved, historical events, etc. When no users gaze upon the screen, no textual information is shown on the interface, only the overall map is displayed with dots indicating 19 gaze-aware targets. Each target is in the center of a room, and has the following functionality.

4.1.1 Personalisation (Separate target). When a user looks at a room, the name of its function is displayed. The system also provides more information over time. Each two seconds users gaze at the location, new text is given to the user. The previous displayed text is gradually faded out while new text is faded in. The texts are fictional information about the room, including the room's name, the built date, or people who work in the location.

We use two methods to personalise the content toward the user. First, information about the room is automatically oriented toward the viewer. The system fades in name and other information about the room on-gaze; all of these information are adjusted to the user's view direction. This is beneficial as it makes explicit functionality implicit, by automating an explicit rotation gesture (Figure 2b). Second, the text is displayed in the language the user prefers. This can be beneficial in a museum which is frequented by tourists speaking different languages; other application areas might be airport maps. Users can choose between four languages with a direct touchable menu close to their seating position.

4.1.2 Shared View: Same target × Compromised Effect. In case multiple users look at the same location, users are provided with a merged view (Figure 2b (right). Information about the location is copied, shown in both directions, and adapted to both users' language setting. Copied information is offset to not interfere between the users, but still close to the location so the association of the text to the location is obvious.

4.1.3 Implementation Considerations and Issues. Inspired by prior work that use automatic content orientation forming like a 'circle' around the user [Vernier et al. 2002], we also oriented content exactly to the user's seat position. We quickly noticed that this is unnatural: only when users move their head, exact orientation toward user was correct; but for reading activity, users mainly move their eyes, so that the actual orientation needs to keep horizontal toward the user.

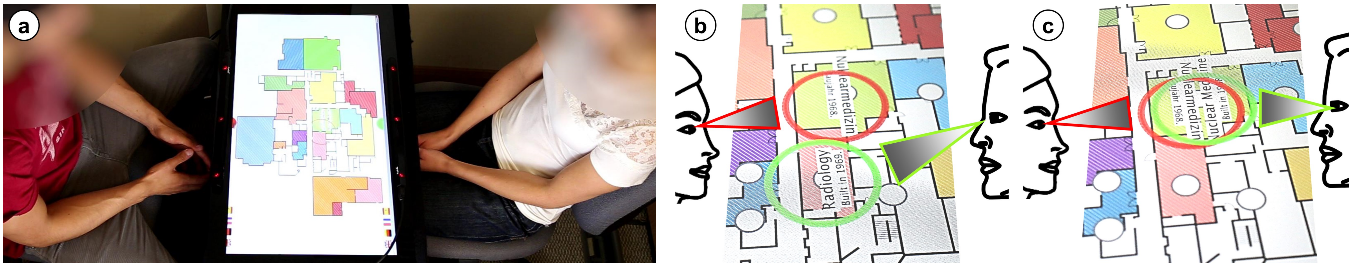

Figure 2: (a) The setup with a touchscreen and two eye trackers facing the users. (b) The gaze-aware map where content is adapted to the user, for separate attention (each directed toward user) and (c) shared attention (information is split) .

We also considered to constantly orient all targets, even when users do not look at it. A target would then always be oriented toward the user who is looking closer to it. This initially seemed more convenient: when users view several close targets in sequence, the targets would already be correctly aligned toward the user. In practice however, users frequently glance over the whole screen. With two users doing this, all targets constantly changed their orientation, which was confusing. For this reason, we kept adaptation to single targets, and only when users looked at them.

4.2 Tree-of-Life Exploration

This application shows how implicit gaze interaction can support collaboration of the users. Previous work showed successful deployments of tree-of-life explorations in public museums [Horn et al. 2012; Hornecker 2008]. We implemented a variation with eye trackers where collaborative viewing is used to gradually unfold the tree of life. The goal of the application is to implicitly guide the user through the various stages of the evolution of species and their subspecies. However, the application's concept can be applied to other exploratory and didactic material that is based on progressive information as well, such as an animation that is based on a sequence of scenes, a presentation showing slides one by one, or the operation of a computer that consists of multiple components.

4.2.1 Gradually-revealing Exploration: Same target × Collaborative Effect. The exploration begins with the origin of life, and guides the user sequentially from the first species, their subspecies, and so on. The users' gaze is implicitly used to proceed to the next elements. For example, both users look at the 'origin of life' (Figure 3a), which then unfolds to four subspecies (b). Users then look at the 'animals' subspecies (b) which again unfolds four subspecies of animals (c). The next element, 'sponges' is final and thus unfolds several example images when looked (d). The whole exploration ends when all 22 species were unfolded (Figure 3). Thus when both users looked at the element, the system proceeds to the next item. The progression is on hold until the users viewed the object, making sure users do not miss important steps of the sequence.

4.2.2 Image and Text Adaptation (Separate target). The name of the species is displayed on each element. This name is automatically oriented towards the user. For the conflict of both users looking at the same target, we implemented a first-come-first-serve mechanism. The first users who looks keeps the orientation until looked away. The species without further subspecies present several example life forms in form of images. These images have the same orientation functionality as the text. In addition, the size of an image increases with increasing time spent looking at the image (Figure 3).

4.2.3 Implementation Considerations and Issues. Activation of next elements occurs if any gaze data of both users is sensed on the same object. Unfolding is triggered which in itself is a 1 second process, during which the next elements are not gaze-activatable yet. The image up-scaling is based on gaze dwell-time (1 second) on the target. Image size increases to 150-300% when users look, and decreases when not looked. First user tests showed that this often leads to confusing behaviour as users often shortly glanced away, because they got distracted by or wanted to get the attention of the second user. We therefore implemented a hold mechanism, that keeps the current scale for two seconds after a user looked away, and only then begins scaling down.

4.3 Shared Zooming Map

This application shows how explicit gaze interaction supports users with individual goals. Enabling multi-user navigation on a shared map is challenging, as zooming operations typically affect the whole map. Researchers proposed the use of local zooming windows for this issue, where each user zooms where they perform a direct pinching gesture [Forlines et al. 2006]. Direct gestures, however, can be difficult to perform when other users occlude the area of interest with their own zooming operation, or when multiple user's hands and arms spatially interfere. As one potential approach, we leverage gaze-touch to perform the zooming operation, which renders the gesture indirect. The application also provides an additional functionality. Users can set a different zooming filter. Users can use the close-by menu to either select a fisheye or a magnification lens.

4.3.1 Multi-user Zoom: Separate target. The gaze-touch techniques is used to enable pinch-to-zoom into the user's gaze position. Each user can look at a target, and touch down two fingers to create a zooming window (Figure 4a). By pinching these fingers, the window changes size and zooming factors (4b). Users can either begin with a short finger distance, to gradually increase the zoom into the looked region. Or, users can directly begin with a larger distance between both fingers, in order to instantly view a magnified area. When users touched down, they can also move their finger position on the surface to indirectly drag the zooming window.

4.3.2 Zoom Conflict: Same target × Compromised Effect. When both users perform the gaze-touch zooming into the same position,

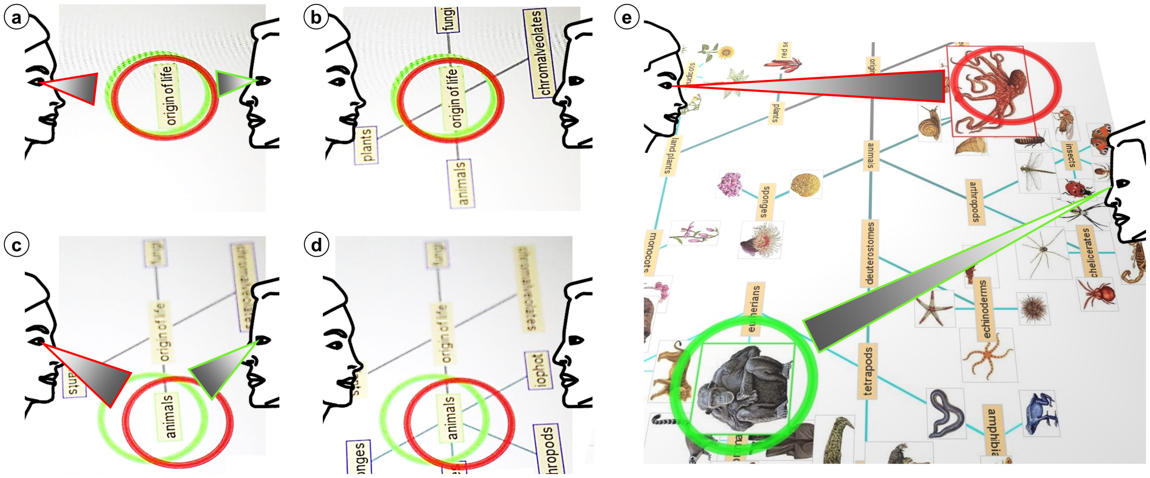

Figure 3: Tree-of-Life: The tree begins with the 'origin of life', and when both users looked at it (a), the successor species are faded in (b). Each further species is gaze-aware, and a joint gaze (c) unfolds its subspecies (d). The whole tree comprises 22 life species with additional image examples, that can be individually viewed in corrected orientation on-gaze (e).

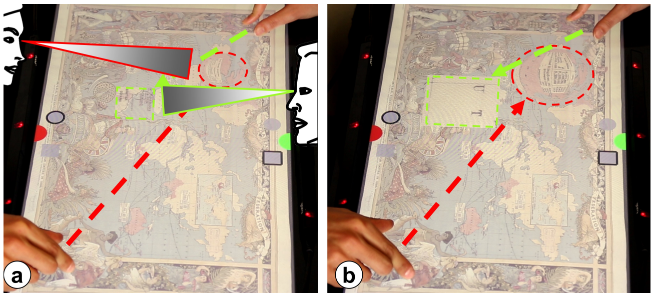

Figure 4: Zooming application: multiple users can perform zooming operations on a single shared map without interference or occlusion issues. To select the zooming position, users look at the remote target and touch down two fingers at a local position (a). Pinching increases the zooming focus (b). Users can select the preferred zooming mode, i.e. a magnifying glass (green) or a fisheye zoom (red).

a conflict occurs. This conflict occurs when users want different zooming scales, or when they selected a distinct zooming type (fisheye versus magnification lens). We use a first-come-first-serve mechanism here to avoid conflicting behaviour. The first user who touched down obtains precedence.

5 INFORMAL USER FEEDBACK IN PUBLIC DEPLOYMENT

We evaluate the Gaze-aware Map and Tree-of-Life Exploration, focusing on the user experience, particularly assessing system issues, learning and behaviour observations.

We use an uncontrolled, public setting to evaluate our concepts with spontaneous users without preparation, representing a realistic multi-user scenario. This can involve rapid group changing dynamics and unexpected behaviours, often experienced in real-world rather than in laboratory settings. We set the system in a public cafe during a student event at Lancaster University. The system was set in the center of the cafe on a high table, enabling spontaneous interactions with the system in standing position. The whole study session lasted 7 hours (11am to 6pm). Notes taken from interviews and impromptu feedback from the users were included in the post-hoc analysis.

Thus, users either spontaneously approached the system, or got interested from comrades telling them about the system. The experimenter briefly explained the system at the beginning, and assisted users in interacting with the system when needed. The explanation involved where users had to stand (in front of eye tracker), to look at the surface, and that it is gaze-aware. During interactions, users were further introduced to features of the application. We switched between the applications for the users so that many users can experience the applications.

60 users (21 female, 24 children), in 24 groups of sizes of 3 to 4 users participated. Users interacted on average 5.85 minutes (SD=3.85 minutes, a session lasts as long as users interact with system). A group typically consisted of adults (students, or random visitors) or a family. 37 users interacted with the campus map, 47 with the tree-of-life exploration.

5.1 System Challenges and Feedback

Regarding eye-tracking, the calibration went well as the majority of users successfully used the applications pre-calibrated. This was still the case for varying user age (children, adult, seniors) or eye-correction (normal sight, glasses, contact-lenses). Only a few users needed recalibration as the gaze cursor was reported as very offset. We did not conduct an accuracy test to measure the quality, but observed that for most users the quality was sufficient to use the apps. Our system was initially not designed for young children. As their height disallowed to be in tracking range of the table-mounted tracker, a chair was added on which they could stand

and successfully interact with the system. Regarding UI design, for the Map application, a few users stated that the gaze-based text appearance is too quick, and therefore becomes confusing when gazing around. With the Exploration application, some users stated inaccurate selection of images, caused by tracking errors. The gaze cursor that was displayed on the screen was reportedly often not precisely at the user's gaze, but did not really harm the user experience which we account mostly to the smoothing.

5.2 Learning and Behaviour Observations

Users immediately grasped how the system implicitly adapts content toward their gaze on the map and exploration applications, as only visual inspection was required. While the Map application was the simplest as it only highlighted text per gaze-location, the Exploration was similarly learned. The feature that users look at the same item to unfold its successors seemed difficult before trial; however after practical trial, users immediately understood and focused on exploring elements of the tree of life. The other feature, that images scale up when looking longer, did not require explanation by the experimenter as it became apparent in trial.

With regards to the map application, part of the users did not immediately understand the purpose of the application, only after the experimenter explained the concept of having personalised content towards each user (here: orientation). The Exploration application was found useful from adults for didactic and exploratory material, as an additional method that can aid learning activities. Children were positive about this application for two reasons. First, that the system reacted to them looking at images of animals and insects; even after unfolding the whole tree, they were still interested in seeing different example species. When a particularly interesting animal was found, they drew the attention of the partner to look as well. Thus, the UI rendered it easy for users to quickly grasp where the other user was focusing on and to direct attention.

6 CONCLUSION

In this paper, we presented an exploration into collaborative gaze and touch interactions. We designed three applications and interaction techniques that specifically exploit the user's gaze that is sensed on the same user interface. The deployment of our multi-user system and the positive feedback gained from the broad range of users that interacted with our system showed the high potential of gaze for multi-user scenarios. Gaze input did not hamper communication between users, instead users interacted with our system and at the same time synchronised their actions through verbal and gestural cues. Challenges such as calibration remain, but with current attachable eye trackers collaborative gaze interaction is becoming increasingly feasible. We have shown how it can become an integrative part in the user interface for several application cases, and that it can enrich the user's collaborative experience with a public interactive system.

REFERENCES

Maneesh Agrawala, Andrew C. Beers, Ian McDowall, Bernd Fröhlich, Mark Bolas, and Pat Hanrahan. 1997. The Two-user Responsive Workbench: Support for Collaboration Through Individual Views of a Shared Space. In *SIGGRAPH*. ACM, 327–332.

Richard A. Bolt. 1981. Gaze-orchestrated Dynamic Windows. *SIGGRAPH Comput. Graph.* 15 (1981), 109–119.

Susan E Brennan, Xin Chen, Christopher A Dickinson, Mark B Neider, and Gregory J Zelinsky. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition* 106, 3 (2008), 1465–1477.

Sarah D'Angelo and Darren Gergle. 2018. *An Eye For Design: Gaze Visualizations for Remote Collaborative Work.* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173923

Andrew T. Duchowski. 2002. A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers* 34, 4 (2002), 455–470.

Clifton Forlines, Alan Esenther, Chia Shen, Daniel Wigdor, and Kathy Ryall. 2006. Multi-user, Multi-display Interaction with a Single-user, Single-display Geospatial Application. In *UIST*. ACM, 273–276.

Carl Gutwin and Saul Greenberg. 2001. The Importance of Awareness for Team Cognition in Distributed Collaboration. In *Team Cognition: Understanding the Factors That Drive Process and Performance*. Press, 177–201.

Ken Hinckley. 2003. Synchronous Gestures for Multiple Persons and Computers. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (Vancouver, Canada) *(UIST '03)*. Association for Computing Machinery, New York, NY, USA, 149–158. https://doi.org/10.1145/964696.964713

Ken Hinckley, Randy Pausch, Dennis Proffitt, James Patten, and Neal Kassell. 1997. Cooperative Bimanual Action. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 27–34. https://doi.org/10.1145/258549.258571

David Holman. 2007. Gazetop: Interaction Techniques for Gaze-aware Tabletops. In *CHI EA*. ACM, 1657–1660.

Michael Horn, Zeina Atrash Leong, Florian Block, Judy Diamond, E. Margaret Evans, Brenda Phillips, and Chia Shen. 2012. Of BATs and APEs: An Interactive Tabletop Game for Natural History Museums. In *CHI*. ACM, 2059–2068.

E. Hornecker. 2008. "I don't understand it either, but it is cool" - visitor interactions with a multi-touch table in a museum. In *TABLETOP, 3rd IEEE International Workshop*. 113–120.

Hiroshi Ishii and Minoru Kobayashi. 1992. ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact. In *CHI*. ACM, 525–532.

Robert J. K. Jacob. 1990. What you look at is what you get: eye movement-based interaction techniques. In *CHI*. ACM, 11–18.

Robert J. K. Jacob. 1993. Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces. In *Advances in Human-Computer Interaction*. Vol. 4, Ablex Publishing, 151–190.

Abhijit Karnik, Diego Martinez Plasencia, Walterio Mayol-Cuevas, and Sriram Subramanian. 2012. PiVOT: Personalized View-overlays for Tabletops. In *UIST*. ACM, 271–280.

Seokhwan Kim, Xiang Cao, Haimo Zhang, and Desney Tan. 2012. Enabling Concurrent Dual Views on Common LCD Screens. In *CHI*. ACM, 2175–2184.

Roman Lissermann, Jochen Huber, Martin Schmitz, Jürgen Steimle, and Max Mühlhäuser. 2014. Permulin: Mixed-focus Collaboration on Multi-view Tabletops. In *CHI*. ACM, 3191–3200.

Päivi Majaranta, I Scott MacKenzie, Anne Aula, and Kari-Jouko Räihä. 2006. Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society* 5, 2 (2006), 199–208.

Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. 2006. Cooperative Gestures: Multi-user Gestural Interactions for Co-located Groupware. In *CHI*. ACM, 1201–1210.

Meredith Ringel Morris, Kathy Ryall, Chia Shen, Clifton Forlines, and Frederic Vernier. 2004. Beyond "Social Protocols": Multi-user Coordination Policies for Co-located Groupware. In *CSCW*. ACM, 262–265.

Joshua Newn, Fraser Allison, Eduardo Velloso, and Frank Vetere. 2018. *Looks Can Be Deceiving: Using Gaze Visualisation to Predict and Mislead Opponents in Strategic Gameplay*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173835

Joshua Newn, Eduardo Velloso, Fraser Allison, Yomna Abdelrahman, and Frank Vetere. 2017. Evaluating Real-Time Gaze Representations to Infer Intentions in Competitive Turn-Based Strategy Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 541–552. https://doi.org/10.1145/3116595.3116624

Joshua Newn, Eduardo Velloso, Marcus Carter, and Frank Vetere. 2016. Multimodal Segmentation on a Large Interactive Tabletop: Extending Interaction on Horizontal Surfaces with Gaze. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces* (Niagara Falls, Ontario, Canada) *(ISS '16)*. Association for Computing Machinery, New York, NY, USA, 251–260. https://doi.org/10.1145/2992154.2992179

Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: Combining Gaze with Multi-touch for Interaction on the Same Surface. In *UIST*. ACM, 509–518.

Ken Pfeuffer, Jason Alexander, Ming Ki Chong, Yanxia Zhang, and Hans Gellersen. 2015. Gaze-Shifting: Direct-Indirect Input with Pen and Touch Modulated by Gaze. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. ACM, New York, NY, USA, 373–383.

https://doi.org/10.1145/2807442.2807460

Ken Pfeuffer, Jason Alexander, and Hans Gellersen. 2016. GazeArchers: Playing with Individual and Shared Attention in a Two-player Look&Shoot Tabletop Game. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (Rovaniemi, Finland) *(MUM '16)*. ACM, New York, NY, USA, 213–216. https://doi.org/10.1145/3012709.3012717

Ken Pfeuffer and Hans Gellersen. 2016. Gaze and Touch Interaction on Tablets. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. ACM, New York, NY, USA, 301–311. https://doi.org/10.1145/2984511.2984514

Ken Pfeuffer, Melodie Vidal, Jayson Turner, Andreas Bulling, and Hans Gellersen. 2013. Pursuit Calibration: Making Gaze Calibration Less Tedious and More Flexible. In *UIST*. ACM, 261–270.

Meredith Ringel, Kathy Ryall, Chia Shen, Clifton Forlines, and Frederic Vernier. 2004. Release, Relocate, Reorient, Resize: Fluid Techniques for Document Sharing on Multi-user Interactive Tables. In *CHI EA*. ACM, 1441–1444.

Radiah Rivu, Yasmeen Abdrabou, Ken Pfeuffer, Mariam Hassib, and Florian Alt. 2020. Gaze'N'Touch: Enhancing Text Selection on Mobile Devices Using Gaze. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382802

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

Dominik Schmidt, Ming Ki Chong, and Hans Gellersen. 2010. IdLenses: Dynamic Personal Areas on Shared Surfaces. In *ITS*. ACM, New York, NY, USA, 131–134.

Baris Serim and Giulio Jacucci. 2016. *Pointing While Looking Elsewhere: Designing for Varying Degrees of Visual Guidance during Manual Input.* Association for Computing Machinery, New York, NY, USA, 5789–5800. https://doi.org/10.1145/2858036.2858480

Bariş Serim, Ken Pfeuffer, Hans Gellersen, and Giulio Jacucci. 2018. Visual Attention-Based Access: Granting Access Based on Users' Joint Attention on Shared Workspaces. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 133 (Sept. 2018), 23 pages. https://doi.org/10.1145/3264943

Harri Siirtola, Oleg Špakov, Howell Istance, and Kari-Jouko Räihä. 2019. Shared gaze in collaborative visual search. *International Journal of Human–Computer Interaction* 35, 18 (2019), 1693–1705.

Randy Stein and Susan E. Brennan. 2004. Another Person's Eye Gaze As a Cue in Solving Programming Problems. In *ICMI*. ACM, 9–15.

Sophie Stellmach and Raimund Dachselt. 2012. Look & Touch: Gaze-supported Target Acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. ACM, New York, NY, USA, 2981–2990. https://doi.org/10.1145/2207676.2208709

W. Steptoe, O. Oyekoya, A. Murgia, R. Wolff, J. Rae, E. Guimaraes, D. Roberts, and A. Steed. 2009. Eye Tracking for Avatar Eye Gaze Control During Object-Focused Multiparty Interaction in Immersive Collaborative Virtual Environments. In *2009 IEEE Virtual Reality Conference*. 83–90. https://doi.org/10.1109/VR.2009.4811003

Jayson Turner, Andreas Bulling, Jason Alexander, and Hans Gellersen. 2014. Cross-device Gaze-supported Point-to-point Content Transfer. In *ETRA*. ACM, 19–26.

Consuelo Valdes, Michelle Ferreirae, Taili Feng, Heidi Wang, Kelsey Tempel, Sirui Liu, and Orit Shaer. 2012. A Collaborative Environment for Engaging Novices in Scientific Inquiry. In *ITS*. ACM, 109–118.

Frédéric Vernier, Neal Lesh, and Chia Shen. 2002. Visualization Techniques for Circular Tabletop Interfaces. In *AVI*. ACM, 257–265.

Roel Vertegaal. 1999. The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration. In *CHI*. ACM, 294–301.

Roel Vertegaal. 2003. Attentive user Interfaces. *Commun. ACM* 46, 3 (2003), 30–33.

Daniel Wigdor, Hao Jiang, Clifton Forlines, Michelle Borkin, and Chia Shen. 2009. WeSpace: The Design Development and Deployment of a Walk-up and Share Multi-surface Visual Collaboration System. In *CHI*. ACM, 1237–1246.

Ana Zanella and Saul Greenberg. 2001. Reducing Interference in Single Display Groupware Through Transparency. In *ECSCW*. Kluwer Academic Publishers, 339–358.

Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and gaze input cascaded (MAGIC) pointing. In *CHI*. ACM, 246–253.

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 Ubiquitous Comput.* 21, 1 (Feb. 2017), 173–186. https://doi.org/10.1007/s00779-016-0969-x