GazeTop: Interaction Techniques for Gaze-Aware Tabletops

David Holman
Media Computing Group
RWTH Aachen University, Germany
holman@cs.rwth-aachen.de

Abstract
GazeTop is a tabletop system that tracks multi-user eye movement in a co-located setting. Knowledge of eye movement is highly relevant to tabletop interaction: eyes can point to distant targets on large tables, address usability issues imposed by rotation sensitive objects, such as menu and text, and facilitate new types of multimodal interactions. This research will evaluate a set of novel eye-controlled interactions and explore the design space of gaze-aware tabletop systems.

Keywords
Tabletops, Multi-touch, Eye tracking, Gaze-based Interaction.

ACM Classification Keywords
H.5.2. [Information interfaces and presentation]: User Interfaces.

Introduction
Although research in tabletop displays is highly active in HCI, tables that afford eye-controlled input have not been explored. Gaze-based interaction offers many benefits for interactive systems**.** Eye-based input is *natural*; humans frequently use gaze to establish turn taking protocols during group communication [15]. Eye movements are also fast, far reaching, and require a low amount of effort from the user — we make thousands of eye movements daily and experience little fatigue [13]. Eyes also provide context that can simplify interaction. For example, in the AuraLamp system [11], the user looks at a lamp and issues the speech command "On", instead of the cumbersome "Lamp on".

In my dissertation, I will focus on the integration of gaze-based interfaces and tabletop interactions. Motivated by previous work in gazed-based interfaces [2,5,9,13], this research will: (a) examine a set of interaction techniques that explore the design space of gaze-aware tabletops, (b) evaluate gaze-based interactions in this domain, and (c) offer a software framework that supports rapid development of gazed-based tabletop interactions.

Current Progress
As part of an interactive museum exhibit [8], I constructed both the hardware and software for large a multi-touch table based on [3]. During a two-week trial, I completed silent observations and structured interviews with ten participants, ranging in both age, twenty to fifty years old, and technological expertise.

CHI 2007, April 28–May 3, 2007, San Jose, California, USA.
ACM 978-1-59593-642-4/07/0004.

Figure 1. The envisioned GazeTop system: infrared (IR) cameras are placed at each corner of the table. IR LEDs are distributed around the perimeters of the table and act as point light sources. Multiple cameras, aware of the LEDs positions, observe the IR reflections on the user's eye and compute the gaze path [12]. This removes the need for a user calibration step. Unfortunately, a full discussion of this approach is outside the scope of this proposal.

From these observations and a review of existing literature [9,13,15], maintaining fluid tabletop interactions would require a gaze-aware tabletop to:

- Enhance multi-user interaction: the table must support a variable number of users on each side. It should not negatively impact key activities such as face-to-face communication, group members working adjacent, or user awareness through gaze monitoring [14].
- Be calibration-free: Current vision based eye tracking systems require the user to follow a calibration procedure: users track a series of points on a display before engaging the system. Although functional, the calibrated system limits head movement and is unique to each user. This is particularly restrictive for a tabletop scenario, as it would limit communication and interaction. To address this, I will adapt a calibration free-technique for vertical displays [12] (see Figure 1).

Gaze-Aware Tabletops

To explore the design space of gaze-based interaction, I will investigate three broad scenarios that incorporate gaze-based interaction on a tabletop surface.

Selection Of Distant Targets

As the size of horizontal display surfaces increase, selection of distant targets is increasingly difficult [1]. Although multi-user scenarios offer a social means to solve this problem, the single user is limited.

TractorBeam [7] address this problem using a stylus and a 6 degree-of-freedom tracking system. The user extends their reach and transfers between touch for local targets and stylus for distant targets. Baudisch et al. [1] present Drag-and-Pop, an interaction technique that moves potential targets closer to the user's cursor, and Drag-and-Pick, a similar interaction that allows target selection.

Although these approaches make distance objects more accessible, they do not account for user eye movement. Eye movements are often issued before other commands [13]. For example, when using a mouse, users typically look at a target before moving the cursor to it. To simplify the effort of selecting distant items, we will make cursor movement implicit in user eye movement. Although the naive approach suggests simply selecting what the user is looking at, Jacob's work in gaze-based interaction shows this is problematic [5]. This approach leads to the "Midas Touch" effect, a situation where the brain is overloaded as an input and output device. He suggests adding an additional step in the selection process. To confirm selection, users issue an auxiliary key activation, speech command, or dwell on the target.

Fono et al. [2] evaluated eye-controlled focus selection techniques for multiple windows on vertical displays. Eye-controlled selection with key activation performed 72% faster than mouse based selection and was preferred by most users. I will study this selection technique on horizontal displays and compare it against existing techniques in a multidirectional target selection task, similar to the experimental design of TractorBeam [7].

Text Orientation

Tabletops are a frequent meeting point for group collaboration. However, in a multi-user tabletop scenario, displaying information that is properly oriented for each user is challenging. Rotation-sensitive components, such as menus and text, work well for vertical displays but are problematic for horizontal displays.

To address this, text can be automatically oriented on the angle that is most readable to the user [9]. The initial attempts are then improved through manual interaction. Although this reduces user effort, it assumes that each user occupies a fixed location at the table. As users move, a common activity during collaboration, this approach becomes less precise. To account for user drift, adaptive methods have used neural nets to predict the location of the user [4]. Although this successful at orienting text after a position change, it relies on the use of a stylus to identify the user. If the user interacts manually, the neural net fails to account for this. To avoid these location issues, the table can be divided into independent areas that oriented for each user. The Lumisight [6] table renders both a local views and a shared central area. Although this approach solves the orientation problem, it sacrifices screen space.

Knowledge of user gaze position can address the design challenges of rotation-sensitive items. Since a gaze-aware tabletop understands the vector between the user's eye and the table surface, the correct orientation angle is easily calculated. This work will focus on what types of orientation techniques are best suited for this scenario.

We must also consider what orientation techniques are appropriate when multiple users select the same item. Gaze-interaction affords divergent approaches. While gazing at an item, a user can remain its owner. Control can be transferred when the user is finished or through explicit negotiation with other group members. Alternatively, a democratic approach can calculate the "best" angle that makes the item equally readable by all. It is not clear which of these approaches is more appropriate. To determine this, we will evaluate user performance in a multidirectional text orientation task and its influence on behavior and communication between group members [14].

Multimodal Interactions

This work will explore multimodal interactions that combine eye and speech gestures. Tse [14] supports expressive interaction in co-located gaming scenario that combines speech and bi-manual gesturing. For example, in the popular game “the Sims”, users can select game characters with a speech command and manually point to their new location. Likewise, users can select multiple items with a bi-manual gesture and utter a delete command.

Gaze-based interaction extends these techniques and uses eye-control to support interactions that feel more natural to the user. Speech driven commands that require multiple pointing gestures, such as moving multiple game characters to disjoint locations, are tedious and may distract from game play. Since the eyes are well suited for rapid movement, eye input can speed up this process. In general, this work will study how eye input can extend other multimodal interactions, such as the gestures Tse supports [14].

Eye-based selection may also improve game immersion. Smith et al. [13] notes this effect and contributes it to the continuous nature of eye-based control as target selection occurs implicitly when observing a game object. This offers a noticeable increase in feedback and leads to a stronger feeling of being in the game. I will study this effect in gaze aware tabletops.

Additional Directions

Gaze-based interaction offers additional areas of research I will explore. Wu et al. [15] present multi-finger and whole hand gestures for multi-user tabletop displays. We will extend the multimodal gaming work to Wu’s more general interaction techniques and examine how eye-based target selection can be incorporated into multi-touch interaction scenarios.

Eye contact is also an important communication cue during group interaction. Vertegaal [15] shows that visualizing who is speaking and where participants are looking during a multi-party videoconference can significantly improve awareness and communication. Visualizing gaze-awareness might also benefit communication and collaboration in distributed tabletop groupware.

So far, this work has focused on tabletop scenarios. Interaction also takes place *above* and *around* the table. Increasing camera count might allow us to measure gaze paths between users and in the environment. Understanding this would support between-user gesturing and more fluid interaction in a multi-display environment

Conclusion

This work presents a set of novel interaction techniques that incorporate eye-based control on tabletop displays. This research will evaluate these techniques and explore the design space of gaze-aware tabletop systems.

References

[1] Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A., Drag-and-Pop and Drag-and-Pick: Techniques for accessing remote screen content on touch- and pen-operated system. In INTERACT '03. Zurich, Switzerland, 2003.

[2] Fono, D., and Vertegaal, R. EyeWindows: Evaluation of Eye-Controlled Zooming Windows for Focus Selection. In Proc. of CHI '05, Portland, OR, 2005.

[3] Han, J. Y. Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. In Proc. Of UIST '05. Seattle, Washington, 2005.

[4] Hancock, M.S. A Feed Forward Neural Network fro Determining a User's Location. Simon Fraser University Techincal Report TR 2001-2. December, 2001.

[5] Jacob, R. J. What You Look at is What You Get: Eye Movement-Based Interaction Techniques. In CHI '90 Proceedings, 1990, pp. 11–18.

[6] Matushita, M., and Iida, M., Ohguro, T., Shirai, Y., Kakehi, Y., and Naemura, T. Lumisight Table: A face-to-face Collaboration Support System That Optimizes Direction of Projected Information to Each Stakeholder. In Proc. of CSCW '04, Chicago, USA, 2004.

[7] Parker, K., and Mandryk, R., and Inkpen, K. Integrating Point and Touch for Interaction with Digital Tabletop Displays. In IEEE Computer Graphics and Applications, Vol. 26, No., 5, 2006, pp. 28-35.

[8] Regensburg Experience. http://www.rex-regensburg.de/.

[9] Rekimoto, J. and Saitoh, M. Augemented Surfaces: A Spatially Continuous Work Space for Hyrbrid Computing Environments. In Proc. of CHI '99. Pittsburgh, USA, 1999.

[10] Scott, S.D., Grant., & Mandryk, R.L. (2003). System Guidelines for Co-located Collaborative Work on a Tabletop Display. In Proc. of ECSCW '03. Helsinki, Finland, 2003.

[11] Shell, J. S., Vertegaal, R., Cheng, D., Skaburskis, A.W., Sohn, C., Stewart, A. J., Aoudeh, O., Dickie, C. ECSGlasses and EyePliances: Using Attention to open Sociable Windows of Interaction. In Proc. of ETRA '04. San Antonio, USA, 2004.

[12] Shih, S., Wu, Y., and Liu, J. A Calibration-Free Gaze Tracking Technique. In the IEEE International Conference of Pattern Recognition '00. Barcelona, Spain, 2000.

[13] Smith, J.D. and Graham, T.C. Use of eye movements for video game control. In Proc. Of ACE '06. Hollywood, USA, 2006.

[14] Tse, E., Greenberg, S., Shen, C., and Forlines, C. Multimodal Multiplayer Tabletop Gaming. In Proc. Of PerGames '06. Dublin, Ireland, 2006.

[15] Vertegaal, R. The GAZE Groupware System: Mediating joint Attention in Multiparty Communication and Collaboration. In Pro. of CHI'99. Pittsburgh, USA, 1999.

[16] Wu, M., and Balakrishnan, R. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. In Proc. of UIST '03. Vancouver, Canada, 2003.