

Gaze-Augmented Manual Interaction

Hans-Joachim Bieg

University of Konstanz, HCI Group
Box D73
78457 Konstanz, Germany
bieg@inf.uni-konstanz.de



Figure 1: Illustrating the research environment: An application is displayed on a 221" Powerwall display. Wearing a mobile eye-tracking system, the user selects and zooms an item of interest with the use of his gaze.

Abstract

This project will demonstrate a new approach to employing users' gaze in the context of human-computer interaction. This new approach uses gaze passively in order to improve the speed and precision of manually controlled pointing techniques. Designing such gaze-augmented manual techniques requires an understanding of the principles that govern the coordination of hand and eye. This coordination is influenced by situational parameters (task complexity, input device used, etc.), which this project will explore in controlled experiments.

Keywords

Gaze, Eye-tracking, Input devices, Fitts' Law

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation] User Interfaces – Input devices and strategies. H.1.2 [Models and Principles] User/Machine Systems – Human factors

Introduction

Fast and effortless control of computer systems with gaze as a pointing device has been a long-standing research goal in HCI. However, contemporary gaze-based systems are mainly marketed as an assistive technology for users with manual disabilities and have not been adopted by the general community. One reason for this

is the high cost, complexity and unreliability of eye-tracking hardware. Nonetheless, recent progress in the design of cheaper, robust, and comfortable systems [5] promises to bring eye-tracking out of the research laboratory to new applications in HCI: from built-in eye-trackers in laptop screens to wearable systems that locate the user's point of regard on large, wall-sized displays. In preparing for this, current research is faced with two additional challenges. The first is the inherent lack of precision of gaze data in revealing the user's intention. The second is the unnaturalness of using gaze for interaction tasks.

Although technological advances have increased tracking precision, they have not increased *inference precision*, that is, the precision of inferring the user's intention from gaze data. The assumption that users generally "... look at what they are working on" [9] may be too simplistic because different tasks require varying levels of acuity for completion. Some tasks may not even require foveation if the visual capability in the periphery suffices. Therefore, a user's point of regard is only a vague indicator of the actual point of information intake. At most, this point specifies a region, the *visual span* [8], where information is *likely* to be extracted from. Thus, in order to increase the inference precision of gaze-based systems, the task-dependency of eye movements must be considered.

The other principal issue is that users do not expect shifts of their gaze to produce a system reaction. The currently predominant user-interface metaphor is one of a model world. As in the real world, users expect to manipulate this world with their bare hands or tools (such as the mouse or a pen). It is therefore argued that gaze could be best used as a "passive input modal-

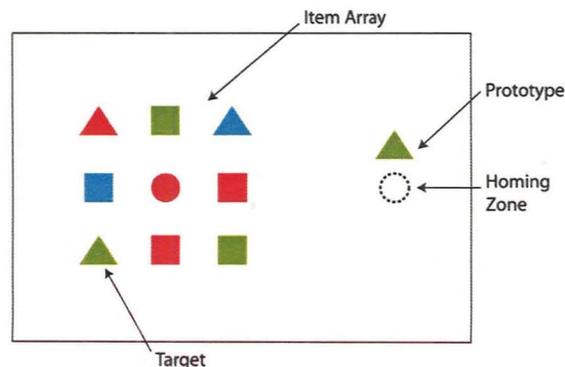
ity" [6], complementing the "active" manual input modality to create a natural interaction experience by hiding the effects of system changes that are due to interpretation of users' gaze. Moreover, studies of real-world tasks have shown that gaze and manual actions are closely related and that gaze often precedes a manual action [7]. Gaze data may provide clues about an intention even before the corresponding action is carried out. Thus, manual actions set a specific context for studying the task-dependency of eye movements in order to increase inference precision.

Research Objectives

The goal of this project is to design new gaze-augmented manual interaction techniques that incorporate gaze as a passive modality. For this, I intend to (a) explore the design space of these techniques and (b) study the coordination of manual actions and eye movements. A first approach in combining gaze and manual information was undertaken by Zhai, Morimoto, and Ihde [10]. Their MAGIC technique combined gaze and manual input to "warp" the mouse pointer to the vicinity of the gaze-fixated item, resulting in movement-time savings. The benefits of this technique can be formalized using Fitts' law [1]. The model predicts movement time according to the distance to the target, the width of the target, and limb- or device-specific determinants. The Fitts' model can therefore serve as a basis for an approach to exploring the savings induced by adding gaze information. The potential of this approach is apparent from the amount of research devoted to the design of pointing techniques that seek to optimize the determinants of Fitts' Law. For instance, Baudisch and colleagues invented drag-and-pop [2], a drag-and-drop technique for large displays that brings potential drop targets closer to the current position of

the pointer, reducing the movement distance. Others proposed target-expansion, a technique to increase the width of an item prior to its selection to facilitate the required movement [4]. Although these techniques proved to significantly reduce the pointing demands, their fundamental issue is scalability [1]. This means that in typical situations with more than just a small number of targets, the system must preselect the targets affected by the techniques in order not to obstruct the entire graphical representation. This scalability issue could be solved by using additional information from users' gaze to narrow the set of *potential targets*.

Figure 2: Select-and-drag task embedded in a conjunctive visual search task with two features, color and shape, so that eye movements are required for discrimination and targets do not "pop-out". Users point to the homing zone prior to the appearance of the array of items and the target prototype. When the items appear, the user acquires the target and drags it to the homing zone.



Progress

As a basis for designing such gaze-augmented manual techniques, patterns must be identified by studying the coordination of manual actions and eye movements. It is assumed that these patterns are determined by a variety of situational parameters (e.g. task complexity, type of input device, display size, etc.). As a basis for experiments that explore these parameters, a research environment was created during the first six months of the project. Foremost, the environment features a mobile eye-tracking system that was set-up in cooperation with colleagues from the Max-Planck Institute for Biological Cybernetics in Tübingen. This system allows for unconstrained tracking of users' point of regard on various screens, ranging from a large 221" high-resolution "Powerwall" to notebook displays. Additionally, the environment provides for easy access to a set of standard as well as novel input devices such as the mouse,

hand gestures, or a new laser pointer device [3]. Informal tests of gaze-based interaction techniques proved that the research environment was operational and suited for the desired purpose (Fig. 1).

Further Research

As a next step, a study is planned to compare the MAGIC technique against a conventional pointing technique with a mouse. The first goal of this study is to replicate the results of the original pilot study [10]. Although the original study demonstrated the advantage of the technique, the savings were small (70 ms). I expect larger savings in scenarios with large high-resolution displays because pointing movements on these displays have larger amplitudes on average. To investigate this, the planned study will use the 221" Powerwall display. The second goal is to collect ecologically valid data to explore the relationship between eye movements and manual actions. The study will therefore feature a pointing-and-dragging task that is embedded in a visual search task (Fig. 2). This task will show (a) if the MAGIC technique also works in a situation with more realistic perceptual and attentional demands and (b) if task-dependent parameters (e.g. visual search complexity) will affect its efficacy. One specific question that should be answered by the study is whether fixation duration can serve as an indicator of the intention to acquire a specific target with the manual device. This could enable an improvement of the efficiency and reliability of the MAGIC technique by making warping of the pointer contingent on fixation *duration* instead of fixation onset. Additionally, the results serve as a basis for exploring how interaction techniques that optimize the width component of Fitts' Law (e.g. target expansion [4]) could incorporate gaze input.

An important situational parameter to consider is the type of manual device used. While the first study will use the mouse, a relative input device, a second study is planned to explore gaze-augmented techniques for absolute devices (e.g. [3]). A primary goal of this second study is to investigate the principal differences in coordination of hand and eye between relative and absolute devices. It is assumed that absolute devices require less visual feedback because the pointer's location (if present) can be inferred from the position of the limb controlling the device. For instance, whereas users first have to locate the mouse pointer before carrying out a selection, they typically already know where their hands are before pointing somewhere on a touch surface! These differences also determine the type of gaze-augmented technique that could be realized with absolute devices.

Conclusion

This project employs a new approach that uses gaze information to improve traditional manual interaction techniques. Additionally, the project deepens the understanding of factors that govern the coordination of hand and eye in human-computer interaction tasks, which benefits the development of improved manual and gaze-controlled interaction techniques.

Acknowledgements

Thanks to Harald Reiterer, Lewis Chuang, Werner König and Jens Gerken for their feedback and support. This project is supported by the Baden-Württemberg Information Technology program (BW-FIT).

Citations

- [1] Balakrishnan, R. "Beating" Fitts' law: virtual enhancements for pointing facilitation. *Int. J. Hum-Comput. St.* 61, 6, (2004), 857-874.
- [2] 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 systems. *Proc. INTERACT*, IOS Press (2003), 57-64.
- [3] König, W. A., Bieg, H.-J., Schmidt, T., and Reiterer, H. Position-independent interaction for large high-resolution displays. *Proc. IADIS*, IADIS Press (2007), 117-125.
- [4] McGuffin, M., and Balakrishnan, R. Acquisition of expanding targets. *Proc. CHI*, ACM Press (2002), 57-64.
- [5] Morimoto, C. H., and Mimica, M. R. M. Eye gaze tracking techniques for interactive applications. *Comput. Vis. Image Und.* 98, 1, (2005), 4-24.
- [6] Oviatt, S. Multimodal Interfaces. In J. A. Jacko and A. Sears (Eds.), *The Human-Computer Interaction Handbook* (2nd ed.). Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2008.
- [7] Pelz, J., Hayhoe, M., and Loeber, R. The coordination of eye, head, and hand movements in a natural task. *Exp. Brain Res.* 139, 3, (2001), 266-277.
- [8] Pomplun, M., Reingold, E. M., and Shen, J. Investigating the visual span in comparative search: the effects of task difficulty and divided attention. *Cognition* 81, 2, (2001), 57-67.
- [9] Sibert, L. E., and Jacob, R. J. K. Evaluation of eye gaze interaction. *Proc. CHI*, ACM Press (2000), 281-288.
- [10] Zhai, S., Morimoto, C. H., and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. *Proc. CHI*, ACM Press (1999), 246-253.