Gaze-supported Foot Interaction in Zoomable Information Spaces.

Conference Paper · April 2013

DOI: 10.1145/2468356.2479610

Raimund Dachselt
Technische Universität Dresden

Fabian Göbel
ETH Zurich

Gaze-supported Foot Interaction in Zoomable Information Spaces

Fabian Göbel
TU Dresden
Dresden, Germany
fabian.goebel@mailbox.tu-dresden.de

Stefan Vogt
TU Dresden
Dresden, Germany
stefan.vogt1@mailbox.tu-dresden.de

Konstantin Klamka
TU Dresden
Dresden, Germany
konstantin.klamka@mailbox.tu-dresden.de

Sophie Stellmach
Interactive Media Lab
TU Dresden
Dresden, Germany
stellmach@acm.org

Andreas Siegel
TU Dresden
Dresden, Germany
andreas.siegel@mailbox.tu-dresden.de

Raimund Dachselt
Interactive Media Lab
TU Dresden
Dresden, Germany
dachselt@acm.org

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
When working with zoomable information spaces, we can distinguish complex tasks into primary and secondary tasks (e.g., pan and zoom). In this context, a multimodal combination of gaze and foot input is highly promising for supporting manual interactions, for example, using mouse and keyboard. Motivated by this, we present several alternatives for multimodal gaze-supported foot interaction in a computer desktop setup for pan and zoom. While our eye gaze is ideal to indicate a user's current point of interest and where to zoom in, foot interaction is well suited for parallel input controls, for example, to specify the zooming speed. Our investigation focuses on varied foot input devices differing in their degree of freedom (e.g., one- and two-directional foot pedals) that can be seamlessly combined with gaze input.

Author Keywords
Multimodal; foot; gaze; eye tracking; interaction; navigation; pan; zoom

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

Introduction

Zoomable information spaces are very common in various application areas. A popular example are Geographic Information Systems (GIS) that require a high degree of expertise from their users [1]. Such systems include a range of diverse interaction tasks (e.g., the exploration, selection, and modification of geospatial data) which can be distinguished into primary and secondary tasks. For example, while manipulating virtual content, the user can move around the virtual camera to get a better view. Despite this diversity of interaction tasks, the interaction with GIS (and zoomable information spaces in general) is often limited to mouse and keyboard input. This prevents a fluent interaction due to the necessity of mode switches that are often performed via toolbars. To leave the hands free for primary interaction tasks, we propose a multimodal combination of gaze and foot input for simultaneously performing secondary tasks, such as seamless pan and zoom.

Figure 1: Overview over setting with *foot devices* & *eye tracker*

While people are familiar with using their feet for the interaction in various application areas (e.g., driving a car, playing music instruments, operating medical equipment), it is scarcely used so far for facilitating the interaction with zoomable information spaces [1, 4]. Foot input serves as an interesting input channel to support manual interactions, but it is not well suited for pointing tasks. In this context, gaze input shows high potential as a fast, implicit and coarse pointing modality that requires low effort from the user. Motivated by this, we contribute several novel approaches for multimodal interaction with zoomable information spaces (in a computer desktop setup) integrating gaze and foot input. We present three prototypes for gaze-supported foot interaction to pan and zoom, leaving the hands free for primary tasks. This includes a combination of different foot input devices, including commercial high-quality gaming pedals and two custom-made foot input devices.

Related Work

Already in the 1980s, Pearson and Weiser [3] propose using foot input as supporting input to overcome interruptions of the user's workflow due to shifting the hand between mouse and keyboard. Pakkanen and Raisamo [2] investigate the appropriateness of foot input for non-accurate spatial tasks and propose to assign supporting tasks like scrolling, moving or resizing objects to the feet. They conclude that *"feet are suitable for the secondary tasks that do not require high accuracy or execution times"* [2]. Several works address multimodal foot input, e.g., with multi-touch gestures (e.g., [1, 4]), but none with gaze input. While Pearson and Weiser [3] have already considered the promising potential for combining gaze and foot input, further investigations have been hindered by the extensive costs and inconvenience of eye tracking devices so far. Thereby, our eye gaze is one of the fastest possible pointing methods, since our eye gaze reaches a target prior to a manual pointer without even thinking about it [5, 7]. Several multimodal gaze-supported pan and zoom approaches exist, e.g., with a touch-sensitive handheld [5, 6]. Stellmach and Dachselt [5, 6] use gaze data to indicate where to zoom in. Based on a user study, users assessed this implicit use of gaze data very positively. Quick panning motions via gaze input should be avoided as this may cause disorientation and motion sickness for the user [5].

Gaze-supported Foot Interaction Concepts

We investigate a combination of gaze and foot input for panning and zooming. For this, we assume a single user working at a computer desk which implies a seated position. For the design of our novel gaze-supported foot

Figure 2: *Three pedals* setup.

Figure 3: *Foot-rocker* device.

Figure 4: *Foot-joystick & Foot-rocker* setup

interaction, we pursued the following design goals:

Comfort: Natural, precise and low fatigue foot input (as far as possible)

Unobtrusiveness: No need for attaching additional equipment to the user's foot

Robustness: The foot device should be sufficiently strong to withstand the weight and force of a human foot

Precision & Customization: High-resolution input for precise movement control and highly adjustable configuration for hardware and software parameters

Intentions: Specific foot-based starting conditions (e.g., a minimum amount a foot pedal has to be pressed down) to avoid involuntarily issuing an action

Prototype Setup

For our software prototype, we use *Google Earth* as a popular representative of a zoomable information space. For tracking gaze data, we use a Tobii TX300. This is a table-mounted, high frequency (up to 300 Hz) binocular eye tracker integrated in a removable 23" HD monitor. For foot input, we use three different input devices that we have combined in different ways. First, we briefly describe the three foot input devices and then discuss the respective input combinations.

Foot pedals. On the one hand, we use the Fanatec CSR Elite[1] pedals as imitation of common car pedals (see Figure 2). The aluminum pedals are robust and offer various configuration possibilities, for example, the angle, order, orientation, as well as horizontal and vertical position of pedals can be adjusted. The setup contains one pressure-sensitive pedal (originally brake) and two pedals depending on angular displacement (originally gas and clutch). All pedals deliver values with a maximum resolution of 10 bits.

Foot joystick and Foot rocker. As an alternative to common car pedals, we designed and manufactured two custom-made foot input devices: *Foot-joystick* (see Figure 4) with two axes and *Foot-rocker* (see Figure 3). The *Foot-joystick* is a ball joint mounted pedal construction containing a cardan joint from a regular gaming-joystick, giving the device two degrees of freedom, a deviation angle of 20° in each direction. The *Foot-rocker* is a pedal we built with a fixed centered single axis that can be tilted about 20° forward and backward. With the help of a spring mechanism both custom pedals will return to their neutral position from any angular displacement after lifting up the foot. All devices can be connected to a computer via USB using DirectInput. Finally, please note that the prototypes shown in Figure 3 and Figure 4 have not yet been optimized in size for a more convenient use.

The user is sitting ca. 60 cm in front of the eye tracker at a standard desk with the foot input device underneath it. The foot devices are not fixed to the ground and can be placed according to personal preferences. We distinguish three particular setup configurations for further investigations that are described in the following.

Prototype 1: **Three pedals**

For the first prototype, we use three foot pedals as they are common for controlling a car. With this similarity, we anticipate that the interaction should be easy to learn and use. Based on our previous description of the Fanatec CSR Elite pedals, the prototype includes one pressure-sensitive pedal that is installed on the left in a standing position (attached from below) while the others are in a hanging

[1] For further information see http://eu.fanatec.com/

Figure 5: *Foot-joystick*: Two-axes foot-based tilting.

Figure 6: *Foot-rocker*: Two-directional foot pedal.

position (see Figure 2). A user can pan into the direction, he/she is currently looking at by pressing the left pedal to increase panning speed. Similarly, a user can zoom towards a currently looked at area by pushing the middle/right pedal to zoom in/out (see [6]). However, this setup has the disadvantage that the mapping of zooming directions is inconsistent: both zoom pedals have the same operation direction (i.e., a pedal has to be pushed forward), but are associated to contrary virtual movement directions (i.e., zoom in and out). This may confuse users and impede a fluent interaction.

Prototype 2: **Single pedal & Foot-rocker**
To allow for a more intuitive mapping of zooming directions, we combine a single foot pedal (from Fanatec CSR Elite) with a two-directional foot pedal (i.e., the *Foot-rocker*) for our second prototype. Analogous to *Prototype 1*, the single pedal is used to control the panning speed towards the current gaze position. Zoom in/out can be performed by tilting the *Foot-rocker* forward/backward. Again the user can indicate where to zoom in/out by looking at a respective location.

Prototype 3: **Foot-joystick & Foot-rocker**
Both *Prototype 1* and *2* use an approach for which the user controls the panning direction via eye gaze. However, fast gaze-based panning across large distances should be avoided as this may lead to disorientation and nausea [6]. To address this issue, we combine both *Foot-rocker* and *Foot-joystick* for our third prototype. The *Foot-joystick* allows for controlling the pan direction in all directions without the need for additional gaze input. Analogous to *Prototype 2*, the *Foot-rocker* is used for controlling the zooming speed, whereby the zooming pivot is based on what the user's currently looking at. This setting enables users to intuitively zoom towards viewed targets and to quickly pan via foot input.

Conclusion

The multimodal combination of gaze and foot input to support secondary interaction tasks is very promising. In our initial investigations, we have developed novel ways for pan and zoom control benefiting from implicit gaze input with explicit foot controls. This enables users to perform secondary tasks in a non-fatiguing way without interrupting the primary task. We presented three prototypes for gaze-supported foot interaction that we will test and evaluate more thoroughly in the future. For this, we put a high emphasis on a flexible setup for easy and quick adjustments for further investigations.

References

[1] Daiber, F., Schöning, J., and Krüger, A. Whole body interaction with geospatial data. In *Smart Graphics*, vol. 5531 of *Lect. Notes in CS*, Springer (2009), 81–92.

[2] Pakkanen, T., and Raisamo, R. Appropriateness of foot interaction for non-accurate spatial tasks. In *Proc. of CHI EA '04*, ACM (2004), 1123–1126.

[3] Pearson, G., and Weiser, M. Of moles and men: the design of foot controls for workstations. *SIGCHI Bull. 17*, 4 (Apr. 1986), 333–339.

[4] Schöning, J., Daiber, F., Krüger, A., and Rohs, M. Using hands and feet to navigate and manipulate spatial data. In *CHI EA '09*, ACM (2009), 4663–4668.

[5] Stellmach, S., and Dachselt, R. Investigating gaze-supported multimodal pan and zoom. In *Proc. of ETRA '12*, ACM (2012), 357–360.

[6] Stellmach, S., Stober, S., Nürnberger, A., and Dachselt, R. Designing gaze-supported multimodal interactions for the exploration of large image collections. In *Proc. of NGCA '11*, ACM (2011), 1:1–1:8.

[7] Vertegaal, R. A fitts law comparison of eye tracking and manual input in the selection of visual targets. In *Proc. of ICMI '08*, ACM (2008), 241–248.