Chapter 9
Computer Control by Gaze

Henrik Skovsgaard  
*IT University of Copenhagen, Denmark*

Kari-Jouko Räihä  
*University of Tampere, Finland*

Martin Tall  
*Duke University, USA*

ABSTRACT

This chapter provides an overview of gaze-based interaction techniques. We will first explore specific techniques intended to make target selection easier and to avoid the Midas touch problem. We will then take a look at techniques that do not require the use of special widgets in the interface but instead manipulate the rendering on the basis of eye gaze to facilitate the selection of small targets. Dwell-based interaction makes use of fixations; recent research has looked into the other option, using saccades as the basis for eye gestures. We will also discuss examples of how eye gaze has been used with other input modalities (blinks and winks, keyboard and mouse, facial gestures, head movements, and speech) to speed up interaction. Finally, we will discuss examples of interaction techniques in the context of a specific area of application: navigating information spaces.

INTRODUCTION

Users of a standard PC need to be able to perform point-and-select operations to interact with the modern graphical user interface. Interaction with interface elements such as icons and buttons is often performed with a conventional mouse, a task that has been mastered by most able-bodied users. The advantage with the conventional mouse is that with this device there is virtually no noise between the physical movement of the device and the movement of the on-screen cursor. This allows selection of very small targets in a windowed environment down to the finest levels, to the pixel.

Users with motor disabilities who are not able to use and control a conventional mouse...
need alternative input devices for performing point-and-select operations to gain access to the graphical user interface. Eye trackers are feasible input devices for users who retain control of their eye movements. Eye gaze has several desirable characteristics, such as natural and fast pointing (Sibert & Jacob, 2000). However, most graphical user interfaces are not designed for use with these alternative input devices, which often have limited accuracy or may require unnatural selection techniques that interfere with access to mainstream GUIs.

Gaze tracking is well suited to pointing because humans naturally tend to direct their eyes in the direction of the target of interest. On the other hand, gaze tracking with no additional input modalities is not very suitable for selection, since humans tend to look at objects of interest to explore them independently of their intention to select them (Jacob, 1991, 1993). It therefore cannot be assumed that the user wants to perform an operation on every object that has been looked at. The speed of eye movements can, in fact, turn into a disadvantage: while it is extremely fast to turn attention to the target of interest, perceiving that item cognitively may well take enough time that the system interprets the lack of activity as an indication of expected system action. Finding a proper balance here is one of the main themes of research in this field.

The methods for computer control by gaze interaction can be divided into two main categories: either the eye tracker is simply used to control the mouse in the normal graphical user interface or a custom interface is constructed. In the first case of mouse control (the so-called eye mouse; see Bates & Istance, 2003), the main problem is that there is no universal method for issuing mouse clicks. The most common method to distinguish inspections from selections is to set a time threshold (i.e., dwell time), with a click issued after the duration of the fixation exceeds a specified amount of time. The use of dwell time may lead to unintentional activations resulting from fixations used for inspection being confused for a selection. This issue is referred to as the Midas touch problem (Jacob, 1991). Increasing the dwell time leads to slow and unnatural interaction, whereas a short dwell time leads to an increase in unintentional activations, which may cause frustration. Dwell-based activation therefore typically faces the classic speed-accuracy trade-off: the faster the interaction, the higher the number of erroneous actions.

These problems have led to the creation of several customised interfaces that are built to accommodate the special needs of target acquisition by eye gaze. The limited precision of gaze interaction restricts the possibilities for target selection per se and has resulted in a number of research projects exploring this issue. Most of these have addressed the problem by means of signal smoothing and effectively manipulating the target area with so-called zooming and distortion interfaces.

In this chapter, we will first explore specific techniques for gaze-based interaction, intended to make target selection easier and to avoid the Midas touch problem. We will then take a look at techniques that do not require the use of special widgets in the interface but instead manipulate the rendering on the basis of eye gaze (so-called gaze-contingent interfaces) to facilitate the selection of small targets. Dwell-based interaction makes use of fixations; recent research has looked into the other option, using saccades as the basis for eye gestures. We will also discuss examples of how eye gaze has been used with other input modalities (blinks and winks, keyboard and mouse, facial gestures, head movements, and speech) to speed up interaction. Finally, we will discuss examples of interaction techniques in the context of a specific area of application: navigating information spaces.

**DWELL-BASED SELECTION**

In dwell-based interaction, fixating for a prespecified time threshold on a certain location will make the system issue an activation at that