Gaze-Based Interaction with Public Displays
Using Off-the-Shelf Components

Javier San Agustin  
IT University of Copenhagen  
javier@itu.dk

John Paulin Hansen  
IT University of Copenhagen  
paulin@itu.dk

Martin Tall  
Stanford University  
tall@stanford.edu

ABSTRACT
Eye gaze can be used to interact with high-density information presented on large displays. We have built a system employing off-the-shelf hardware components and open-source gaze tracking software that enables users to interact with an interface displayed on a 55” screen using their eye movements. The system works at a viewing distance of 1 to 1.5 meters and requires a 30 second calibration procedure for every user. We demonstrate how it can be used to navigate a digital bulletin board display with several notes on top of each other. There are some technical challenges detecting the eyes when people are wearing glasses and when external light sources are present.

Author Keywords
Gaze interaction, Public displays, Low-cost eye tracking, Off-the-shelf components

ACM Classification Keywords
H.5.2 [Input Devices and Strategies]: Evaluation/Methodology.

General Terms
Human Factors

INTRODUCTION
The use of large (e.g., 55 inch) displays in public areas is becoming increasingly popular as the prices of large-size LCD and LED panels decreases. However, these panels do not offer any simple way for the user to interact with the interface.

Information presented on public displays is usually one-to-many, that is, the same information is displayed for many users, limiting the scope to static, general information, such as, for instance, maps and advertisements.

Vogel and Balakrishnan [3] presented a framework for interactive displays in public spaces. Depending on the user’s actions (e.g., approaching the panel, walking away, etc.), the state of the display changes from showing public information, to a personal interaction state. Positioning technologies offer some possibilities of displaying information relevant to the user that is currently next to the display. For example, a user can be detected by the Bluetooth MAC address of his or her mobile phone. Then only information relevant to this individual will be shown on the display.

Although positioning systems may improve communication from display to user by taking into account the user identity, large displays in public areas seldom incorporate an input device. Faces and eyes can be tracked with standard video cameras. Most people can control the movements of their head or eyes well enough for the tracking systems to work as a simple pointing device – similar to a mouse-pointer.

In this paper we present a gaze tracking system built from off-the-shelf components that enables users to control an interface displayed on a large display while standing at a distance in front of the display. The gaze tracking software is based on the ITU Gaze Tracker [1, 2], and has been designed to detect and track eyes at a distance of around 1 to 1.5m. The hardware employed is a video camera equipped with night vision mode. The total cost of the hardware is below $1000 (camera + infrared light sources).

GAZE TRACKER
The ITU Gaze Tracker is an open-source eye tracking software that allows users to try out gaze interaction for a low price. The system described in [1] and [2] employed a webcam mounted close to the user’s eye. The software has been modified and extended to make it work with a video camera placed below the large display. Therefore, the user does not carry any gear on his or her body.

Hardware
The system includes the hardware required in most remote video-based gaze tracking systems. An off-the-shelf video camera (Sony HDR-HC5) equipped with night shot mode is placed below the display and records the user’s face and eyes. Two IR light sources are employed to improve the illumination conditions and to estimate gaze. The software runs on a MacMini computer running Windows 7. The computer is connected to a 55” Samsung LED display.

Gaze Tracking
The gaze tracking software employs image analysis techniques to detect the eye region and the pupils and corneal reflections produced by the IR light sources. The parameters for the detection of eye features are calculated automatically by the software.

The system needs to be calibrated through a 9-point calibration procedure that takes about 30 seconds to complete. The accuracy obtained is around 2 to 4° of visual angle. Due to the small size of the pupil and corneal reflections at long distances, a rather large amount of jitter appears on the estimated gaze coordinates. For this reason, a smoothing stage is applied in order to remove this noise and stabilize the cursor position.
HOW IT WORKS
In order to improve the accuracy of the estimated gaze coordinates, the field of view of the camera is rather limited. For this reason, the user is asked to stand right in front of the screen, and he or she will need to adjust his or her position by moving back and forth so that the eyes are visible on the camera. On average, users stand at around 1 to 1.5m away from the screen.

When no eyes are detected on the image for some time, the system enters a low-energy mode in which the frame rate is reduced. If a pair of eyes is detected, the system begins tracking them, and launches the calibration procedure automatically. Once this is completed successfully, the main interface is shown, and the user can interact with it using the eyes only.

Figure 1 shows the interface of a digital bulletin board running at our campus. Users can interact with the interface using their eye movements. If messages are lying on top of each other, the user can look at them and the system separates them. If the user looks at a message being moved, this will be placed on an empty area of the screen to be viewed in solitude. In this way the display can contain numerous of messages overlaying each other.

Figure 1. Messages on a digital bulletin board. When people look at one of them it gets separated from the others.

TECHNICAL ISSUES
Eye tracking systems are usually employed on a desktop, with the user sitting around 60cm away from the screen and the camera. The accuracy obtained with commercial systems is around 1° of visual angle. In our setup, the user stands 1 to 1.5 meters away from the camera, and therefore the size of the eye region can become rather small on the images captured, which can affect the accuracy and the noise in the estimated gaze coordinates. For this reason, the interfaces should be tolerant to inaccuracies and the interactive elements should be big enough to be selected by the user.

As with most gaze tracking systems, the proposed system has issues with changes in illumination conditions and with glasses. The current setup makes use of IR light in order to stabilize the light conditions. However, light coming from other light sources like windows or lamps can greatly affect the detection of the eye features, and in some cases make eye tracking very challenging. To improve the performance and help in the detection of eye features, the system should be placed in a rather dimmed area where no undesired light sources are present.

The current system does not handle glasses well. The light emitted by the light sources is reflected on glasses, producing big bright areas that often cover the eye region.

When the user moves away from the screen, the system enters the low-energy mode and expects a new user. Even if the same user steps in, the calibration procedure will run again, since the system does not incorporate a mechanism to identify users.

The current version of the system does not offer walk-up-and-use yet. However, with improvements in next generation of standard video cameras and with more experiences on how to best track the eyes in public spaces, we expect that gaze interaction with large displays may find its way into for instance art exhibitions and entertainment. The ITU gaze Gaze Tracker offers a low-cost possibility for all who would like to do research and design within this area.

REFERENCES
