A New Interaction Technique Involving Eye Gaze Tracker and Scanning System

Pradipta Biswas   Pat Langdon

University of Cambridge
Cambridge, UK
{pb400, pml24}@cam.ac.uk

ABSTRACT

This paper presents a new input interaction system for people with severe disabilities by combining eye gaze tracking and single switch scanning interaction techniques. The system is faster than only scanning based systems while more comfortable to use than existing eye gaze tracking based systems. We reported results from a couple of user studies that show the new system is equally fast compared to existing eye tracking systems that does not involve scanning, participants with no prior experience with eye tracking based system could learn using this new system successfully within 10 minutes but it demands higher mental effort in comparison to another new modality of interaction- a gesture based system.

Categories and Subject Descriptors

D.2.2 [Software Engineering]: Design Tools and Techniques – *user interfaces;* **K.4.2 [Computers and Society]:** Social Issues – *assistive technologies for persons with disabilities*

General Terms

Algorithms, Experimentation, Human Factors, Measurement

Keywords

Eye Gaze Tracker, Assistive Technology, Single switch Scanning.

1. INTRODUCTION

Research on eye tracking dated back to late 18th century when Louis Émile Javal investigated saccadic movements in a reading task [14]. Edmund Huey pioneered in building the first eye tracker which was a contact lens connected with an aluminium pointer [14]. At present time eye tracking research [10], [17] can be classified according to the following diagram (Figure 1).

Research on developing eye tracker investigates on reducing the cost of existing infra-red based trackers (e.g. Tobii [19] or FaceLab [11] Tracker) as well as increasing their accuracy. Researchers also worked on developing customized eye trackers for tasks that require only binary input from the tracker [9], [22].

On a different set of applications, eye trackers often help to design better billboards, traffic signs and advertising posters through analysis of users' eye gaze patterns [8], [10], [17]. This analysis even leads to developing models to simulate eye gaze
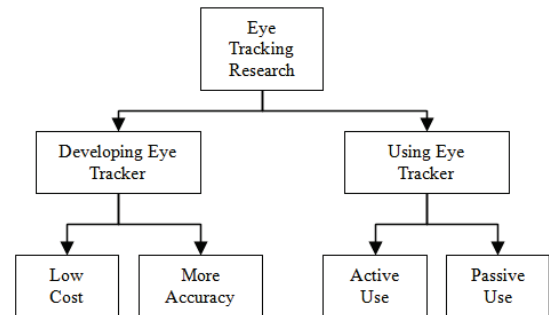
Figure 1. Taxonomy of Eye Tracking Research

movements including people with various visual impairments [4], [6]. Eye trackers can also be used as an interaction device to control electronic devices and applications [1], [2], [10], [16], [20]. This role of eye trackers becomes more significant for certain types of users with physical impairment [20] (e.g. ALS, severe spasticity, cerebral palsy and so on) and situation impairment (e.g. aircraft pilot operating under high G-force). This paper presents a system to use an eye tracker as a pointing device. The system is primarily aimed to users with age-related or physical impairment but can be modified for situational impairment as well. We have reported results from a couple of user studies conducted on the system. We have found that the system is less strenuous to use than existing eye tracking based interface though require more cognitive load and learning time than a gesture based interface.

2. THE SYSTEM

Many eye tracking based interfaces for people with disabilities use the eye gaze as a binary input like a switch press input through a blink [8]. But the resulting system remain as slow as the scanning system. A better solution may be to use the eye gaze to directly control the pointer position in the screen. Zhai [21] and Jacob [15] present a detailed list of advantages and disadvantages of using eye gaze based pointing devices. In short, using the eye gaze for controlling the cursor position pose several challenges as follows

Strain: It is quite strenuous to control the cursor through eye gaze for long time as the eye muscles soon become fatigue. Fejtova and colleagues [12] reported eye strain in six out of ten able bodied participants in their study.

Accuracy: The eye gaze tracker does not always work accurately, even the best eye trackers used to provide accuracy of 0.5° of visual angle. It often makes clicking on small target difficult. Donegan and colleagues [8] also reported problems in precision and speed of an eye gaze based system. So existing systems often change the screen layout and enlarge screen items for Augmentative and Alternative Communication (AAC) systems based on eye gaze [20]. However interface layout of any system cannot be always accessed and changed and cannot always be enlarged especially for small screen based systems - so surely it is not a scalable solution.

Clicking: Clicking or selecting a target using only eye gaze is also a problem. It is generally performed through either dwell time or blinking or both. But either solution increases the chance of false positives or missed clicks.

We tried to solve these problems by combining eye gaze tracking and a scanning system in the following way. Any pointing movement has two phases [13]

- An initial ballistic phase, which brings one near the target.

- A homing phase, which is one or more precise sub movements to home on the target.

We used the eye gaze tracking for the initial ballistic phase and switch to scanning system for the homing phase and clicking. The approach is similar to the MAGIC system [21] though it replaces the regular pointing device with the scanning system. This eye tracking system combines three different assistive technologies together in a single system giving the end users maximum flexibility. It combines

- A basic system of moving a pointer in screen using eye gaze
- A mouse-emulator scanning system sensitive to eye blink
- A mouse-emulator scanning system sensitive to a physical switch

Initially, the system moves the pointer across the screen based on the eye gaze of the user. The user sees an icon moving across the screen and the icon is placed approximately where they are looking at the screen. We extract the eye gaze position by using the Tobii SDK [19].

2.1 Smoothing

Eye movement is different from a mouse movement as it does not smoothly move in a continuous space as a mouse. The eye gaze rather follows a spotlight metaphor and the saccade focuses at the regions of interest. The raw reading from the eye tracker results in jerky movement of the pointer in the screen and it never stays steady at a single location, which makes it difficult to make a selection. We record the eye gaze positions continuously and average the pixel locations in every 400 msec to estimate the region of interest or saccadic focus points. We simulate the eye movement using a Beizer curve [18] that smooths the cursor movement between two focus points. We push the focus points into a stack and the Beizer curve algorithm interpolates points in between two focus points.

2.2 Selection

To make a selection, the user needs to blink his eyes. The system can classify between intentional and non-intentional eye blinks and only recognizes the intentional ones through dwell time adjustment. After the first blink the system waits for a certain configurable interval while the user can make another blink to make a selection (left click, double click or right click). If the user does not make a selection the system switches to a scanning system which can be operated either by a blink or by a physical switch.

2.3 Scanning

Scanning is a technique of successively highlighting interface elements and selecting an element through a binary input channel when the intended element is highlighted. We have used a particular type of scanning system, known as eight directional scanning [3] to navigate across the screen. In eight-directional scanning technique the pointer icon is changed at regular time intervals to show one of eight directions (Up, Up-Left, Left, Left-Down, Down, Down-Right, Right, Right-Up). The user can choose a direction by pressing the switch when the pointer icon shows the required direction. After getting the direction choice, the pointer starts moving. When the pointer

reaches the desired point in the screen, the user has to make another key press to stop the pointer movement and make a click. The technique is less strenuous than the only eye gaze based interfaces because users can switch back and forth between eye gaze tracking and scanning which gives rest to the eye muscles. Additionally, since they need not to home on a target using eye gaze, they are relieved from looking at a target for a long time to home and click on it. Finally, this technique does not depend on the accuracy of the eye tracker as eye tracking is only used to bring the cursor near the target (as opposed to on the target), so it can be used with low cost and low accuracy web cam based eye trackers. A complete state chart diagram (Figure 2) of the system is as follows.
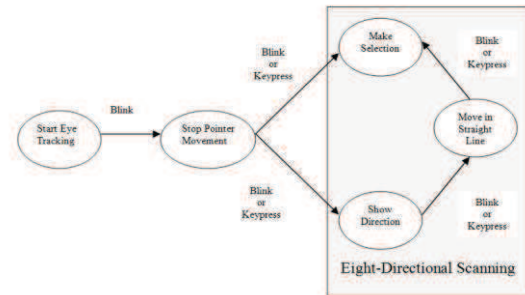

Figure 2. State Transition Diagram of the whole system

3. USER STUDY

However this technique may seem slower than only eye gaze based interface as users need to switch back to the slower scanning technique for each pointing task. So we conducted a user study [7] to compare the speed of our system with respect to only eye gaze based pointing. We conducted a comparative study with eight users involving only eye tracking system and our system combining eye tracking and scanning. Participants rated the new system easier and less strenuous to use than only eye gaze tracking based system while the new system did not slower the speed of interaction.

Eye gaze based interfaces often turns the only possible interaction modality for people with severe disabilities. However for users with situational impairment, eye tracking interface may not be the only option rather part of a multimodal suite. For example, fighter-aircraft pilots use head gestures, direct voice input, and pointing device, but for certain tasks eye gaze based interfaces may be a useful alternative. Similar case studies can be found for users interacting with a smartphone in a moving vehicle. Samsung S4 smartphone already allows pausing and resuming a video by tracking users' eye gaze, though the same task can be performed by the usual touch pad as well. For these types of users and use-case scenarios, it is important to compare eye tracking interface with other novel types of interaction.

We conducted another user study to find out how this new technique performed with respect to another new modality of interaction. We choose a gesture recognition system as it is now pretty common in tablet, smartphone and digital TV based systems. The aim of the study was to investigate whether novice users can learn this system and what should be our future goal in terms of cognitive load with respect to gesture based system.

3.1 Participants

We collected data from 6 users (age between 20 and 30, 3 males, 3 females). 3 out of 6 users used spectacles during the study. None of the participants ever used a gesture recognition or eye tracking system before.

3.2 Procedure

The study was an exploratory one, participants were not given a particular task to complete. They were initially demonstrated the gesture recognition and eye tracking system (Figure 3). We used a gesture recognition system developed in the EU GUIDE (http://www.guide-project.eu) project. This system allows user to interact with a digital TV by pointing at the screen. Selection is done through dwell time – the user needed to hold the pointer still for 2 seconds to make a selection. The study involved a smart home application developed in the GUIDE project. The participants were instructed to use the smart home application for 10 minutes using both the gesture recognition and eye tracking system. The order of the two systems was altered among participants. After each session, participants were instructed to fill up the NASA TLX scoring sheet and also asked about their general comments about the system.

3.3 Material

We used Microsoft Kinect for gesture recognition and Tobii X120 Eye Tracker for eye tracking interface. The study was conducted in 40" LG television screen.

3.4 Results

Figure 4 shows a comparison of different cognitive parameters of the TLX scores. Four out of 6 participants found the gesture based system easier though all of them can operate the eye tracking system as well within the 10 minutes interval.
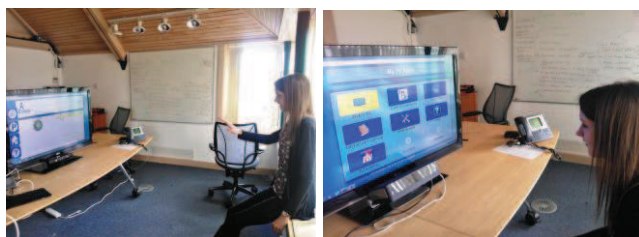
Gesture interaction   **Eye Tracking interaction**

Figure 3. Experimental Setup for User Study 2

Five out of 6 participants can complete their intended task involving both system, however one participant find it difficult to operate the eye tracking system as the eye tracker was losing signal quite frequently for him.
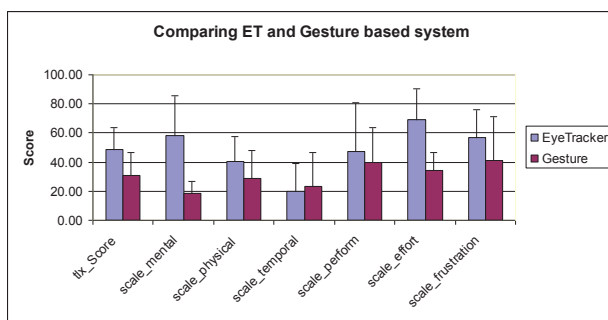

Figure 4. Comparing TLX scores between Eye Tracking and Gesture based system

3.5 Discussion

The results show that the eye tracking system demands more cognitive load than the gesture based system. The highest difference is observed for mental demand and effort. Learning an eye tracking based system often takes a longer duration of use as Kristensson [16] reported that users attained higher typing speed after 40 minutes training. In our experimental set up, participants reported that the eye tracker often lost signal and they found it difficult to get control of the system. They also found it difficult to point at the edge of the screen. These

problems can be solved by using a spectacle or eyeglass based tracker which will allow more head and body movement. Participants did not report issues with the scanning system itself, rather one participant found the scanning system easier to control than the eye gaze based pointer movement. We also observed that the temporal and performance scales are not significantly different ($p<0.05$), indicating the participants did not feel rushed and performed equally well in both systems. We noted that all participants could successfully point and click within the 10 minutes interval. The TLX scores also provide us a future goal and reference to achieve with the eye tracking system.

4. NOVELTY OF THE SYSTEM

Portability: When we use a physical pointer or electronic one like a mouse or trackball, our motor and visual feedback channels are physically separated. While we try to control an on-screen pointer using eye gaze, it often becomes strenuous and confusing as we are using eyes for both visual feedback and motor control. It results in less accurate control of a pointer than a traditional computer mouse even with an accurate eye tracker. Additionally eye trackers may also become less accurate due to lighting condition, technology of the system (webcam vs. infra-red based) or even the user (e.g. presence of hair frills on forehead may reduce the accuracy as the infra-red cast shadow of the hair on eyes). Existing systems try to solve these problems by designing special interfaces with big on-screen items (e.g. virtual keyboards in AAC systems) to counteract reduced accuracy. However this is not a scalable approach as a user should be able to use any general purpose interface with an eye tracking based system. Our system combines a mouse emulator with an eye tracking system to solve this issue. The mouse emulator can reach every single pixel in screen and can be operated through eye blink. It enables the user to interact with any system without any special adaptation of the screen layout for the eye tracking system. Our user trials confirmed that the combined system does not sacrifice speed of interaction by introducing the scanning system.

Midas-Touch Problem: Eye tracking systems often find it difficult to make a selection. Many systems achieve it through adjusting the dwell time of a blink. However we find that keeping a pointer steady for adequate time duration is difficult even with the best quality of eye tracker. As a result when the user finishes the blink action the pointer has already moved out of the target. Here we solved the issue by integrating the scanning system. After the user makes a blink the system does not immediately makes a selection rather wait for the user to make another blink. During this time the eye gaze tracker stops pointer movement. If the user finds the pointer off target, he can bring it in using the eight-directional scanning system which is easier to control for closely spaced screen items like a pop-up menu.

User Centred Design: The development of the Eye Tracking system follows a user centred design process and we constantly collect feedback from users to improve the system. This paper reports studies involving 14 users unlike other work that even reports result on the performance of a single user [9].

5. CONCLUSIONS

This paper presents a new type of eye gaze tracking based interaction that includes a mouse emulator scanning system. The system does not require adaptation of screen layout and also solves the 'Midas-Touch' problem with eye tracking based interface. We conducted two user trials with the system. The trials show that users with no prior experience with eye tracking based interface can learn the system within 10 minutes and the system is less strenuous than existing eye tracking system without and scanning interface. We also reported result on

comparing the system with a gesture based system considering users with situational impairment (like fighter aircraft pilots) and found the eye tracking system requires more mental effort but almost equal temporal demand.

6. REFERENCES

[1] Adjouadi, M. et. al. 2004. Remote Eye Gaze Tracking System as a Computer Interface for Persons with Severe Motor Disability. ICCHP 2004, LNCS 3118 2004. 761-769.

[2] Bates, R. 1999. Multimodal Eye-Based Interaction for Zoomed Target Selection on a Standard Graphical User Interface. INTERACT 1999.

[3] Biswas, P. and Robinson, P. 2008. A New Screen Scanning System based on Clustering Screen Objects, *Journal of Assistive Technologies*, 2, 3 (September 2008), 24-31, ISSN: 1754-9450

[4] Biswas, P. and Robinson, P. 2009. Modelling Perception using Image Processing Algorithms, 23rd British Computer Society Conference on Human-Computer Interaction (HCI 09)

[5] Biswas, P. and Robinson, P.2010. The effects of hand strength on pointing performance, Designing Inclusive Interactions, Springer-Verlag, 3-12, ISBN: 978-1-84996-165-3.

[6] Biswas, P. and Robinson, P. 2008. Modelling user interfaces for special needs, Accessible Design in the Digital World (ADDW).

[7] Biswas, P. and Langdon, P. 2011. A new input system for disabled users involving eye gaze tracker and scanning interface, *Journal of Assistive Technologies*, 5, 2 (June 2011), ISSN: 1754-9450

[8] Donegan, M. et. al. 2009. Understanding users and their needs, *Universal Access in the Information Society,* 8 (2009), 259-275

[9] Dostal, J., Kristensson, P. O. and Quigley A. 2013. Subtle Gaze-Dependent Techniques for Visualising Display Changes in Multi-Display Environments, ACM International Conference of Intelligent User Interfaces (IUI).

[10] Duchowski, A. T. 2007. Eye Tracking Methodology. Springer.

[11] Facelab Eye Tracker, Available at: http://www.seeingmachines.com /product/facelab/, Accessed on: 29/04/13

[12] Fejtova, M. et. al. 2009. Hands-free interaction with a computer and other technologies, *Universal Access in the Information Society,* 8.

[13] Fitts, P.M. 1954. The Information Capacity of The Human Motor System In Controlling The Amplitude of Movement. *Journal of Experimental Psychology,* 47 (1954), 381-391.

[14] Huey, E. 1968. The Psychology and Pedagogy of Reading (Reprint). MIT Press 1968 (originally published 1908)

[15] Jacob, R. 1993. Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces, Advances In Human-Computer Interaction.

[16] Kristensson, P. O. and Vertaneny, K. 2012. The Potential of Dwell-Free Eye-Typing for Fast Assistive Gaze Communication, in *Proceedings of Symposium on Eye Tracking Research and Applications (ETRA)*.

[17] Majaranta, P. and Raiha, K. 2002. Twenty Years of Eye Typing: Systems and Design Issues. Eye Tracking Research & Application.

[18] Salomon, D. 2005. Curves and Surfaces for Computer Graphics, Springer Verlag. ISBN 0-387-24196-5

[19] Tobii Eye Tracker, URL: http://www.tobii.com Accessed on 12th December 2012

[20] Ward, D., Dasher with an eye-tracker, URL: http://www.inference.phy.cam.ac.uk/djw30/dasher/eye.html, Accessed on 19th August 2010.

[21] Zhai, S., Morimoto, C. and Ihde, S. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. ACM SIGCHI Conference on Human Factors in Computing System (CHI).

[22] Zhang, Y., Bulling, A. and Gellersen, H. 2013. Subtle Gaze-Dependent Techniques for Visualising Display Changes in Multi-Display Environments, CHI 2013.