Single Stroke Gaze Gestures

Emilie Mollenbach
Dept. of Computer Science
Loughborough University
Holywell Park
Loughborough
LE11 3TU, UK
e.mollenbach@lboro.co.uc

Alastair Gale
Dept. of Computer Science
Loughborough University
Holywell Park
Loughborough
LE11 3TU, UK
A.G.Gale@lboro.ac.uk

John Paulin Hansen
IT University of Copenhagen
Rued Langgaards Vej 7,
2300 Copenhagen
Denmark
paulin@itu.dk

Martin Lillholm
Dept. of Computer Science
University College London
Malet Place, London
WC1E 6BT, UK
m.lillholm@cs.ucl.uk

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.
ACM 978-1-60558-247-4/09/04.

Abstract
This paper introduces and explains the concept of single stroke gaze gestures. Some preliminary results are presented which indicate the potential efficiency of this interaction method and we show how the method could be implemented for the benefit of disabled users and generally how it could be integrated with gaze dwell to create a new dimension in gaze controlled interfaces.

Keywords
Eye tracking, Gaze Based Interaction, Gaze Gestures

ACM Classification Keywords
H5.2. User Interfaces: Input devices and strategies (e.g., mouse, touch-screen), Interaction styles (e.g., commands, menus, forms, direct manipulation), User-centered design.

Introduction
The main motivation behind this work is to give users with severe motor-skill impairments more flexibility when using gaze as sole input - by increasing the interaction possibilities available.

Many specific gaze applications have been developed for individual tasks, *e.g.* various type-to-talk implementations [3,5,7], environmental control systems [1,2], wheelchair controls [6] and applications giving disabled users the opportunity to participate in online communities, such as 'Second Life' [4]. As the list of applications available for eye tracking users is rapidly growing, the system requirements are increasing. The goal of this research is to add to the 'vocabulary' that can be used when interacting and completing selections in gaze only systems, i.e. systems which are controlled solely by eye movements.

The Concept of Single Stroke Gaze Gestures

Several gaze gesture based approaches have and are being explored either in conjunction with dwell time or not. A gesture is a controlled saccade or a series of controlled saccades that cause an action when completed.

Compound gaze gestures are complex eye movements consisting of several strokes in order to complete a selection. An advantage of this is that they greatly increase the interaction 'vocabulary' of gaze. However, cognitively it is difficult to remember a large number of complex gestures and physiologically it can be difficult to complete them.

Single stroke gaze gestures are simple and lack the broad scope of compound gestures. The advantage of single stroke gaze gestures is that they are easily learnt and sustained and may require less effort by the user than both dwell and compound gestures. See figure 1

Four different single stroke gaze gestures were used; from left to right; right to left; top to bottom; and bottom to top. Using diagonal gestures could have been equally simple to implement and potentially future single stroke gaze gesture structures could incorporate vertical, horizontal as well as diagonal eye movements.

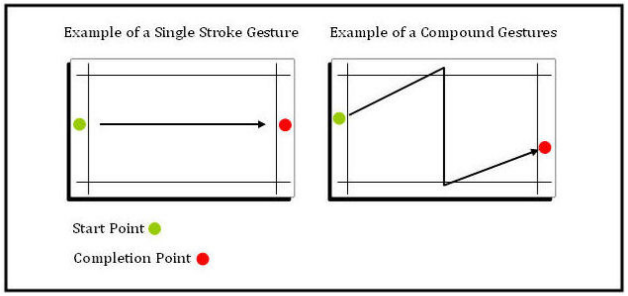

figure 1. Examples of Single Stroke and Compound Gestures

In order to limit the potential overlap between natural eye movements and selections, only the display borders were used to initiate and complete the gestures, i.e. looking from the left edge of the screen to the right edge would initiate a selection, see Figure 2.

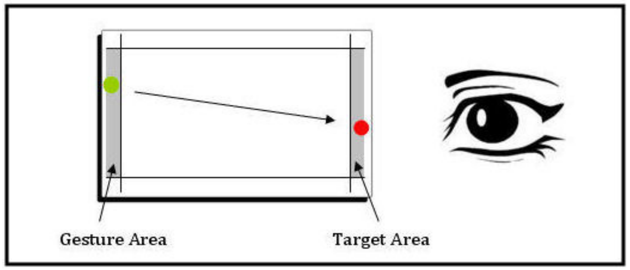

figure 2. The Single Gaze Gesture Selection Process

The stroke needed to be completed within a certain time frame (1000ms) or else the gesture areas would be reset. This function was also useful if the user had initiated a selection that he/she did not wish to complete, i.e. by looking anywhere in the centre of the screen and allowing the system to time out it would reset. See Figure 3.

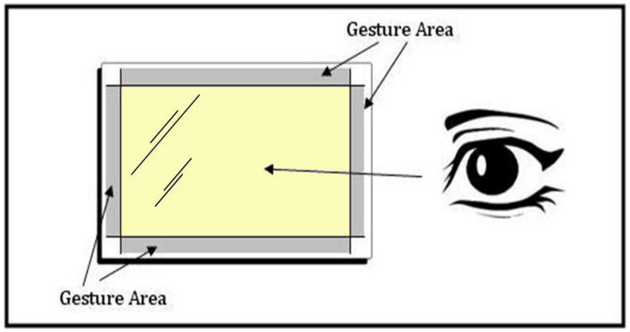

figure 3. The Resetting of Gaze Gesture Areas in the Middle of the Screen

Likewise, if the user entered another gesture area, which was not the adjacent target area, the selection would be cancelled and a new selection process would start from the new gaze gesture area. See figure 4.

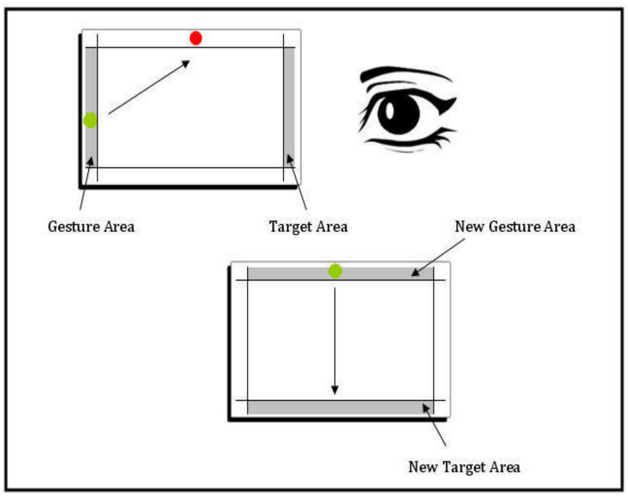

figure 4. Changing Gesture and Target Area in the Middle of a Selection.

Preliminary Findings

12 participants (5 female) had to make eight selections of four different colours in a specific order. The experiment was conducted using a QuickGlance 3 eye tracker with a 1024x768 display. The participant where placed approximately 50 cm from the eye tracker.

Eight small squares on the screen represented the order and the four colours were red, blue, green, and yellow. A colour indicator was placed at all four edges of the screen, these represented starting points, i.e. looking at the blue indicator (the left edge) would start a blue selection which would then be completed by moving the point of gaze to the opposite side of the screen. See Figure 5. The task of eight selections was completed 15 times by each participant with an extra five times for practice, resulting in 120 usable selections per participant.

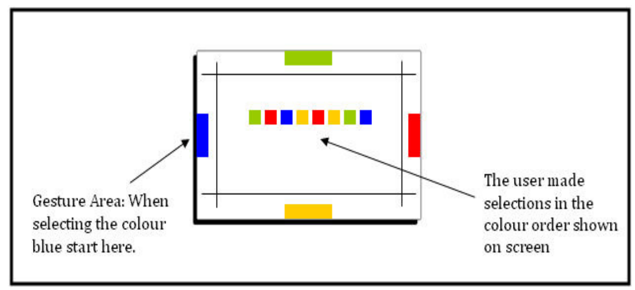

figure 5. The Basic Structure of the Experiment

There are varying reports as to what constitutes standard dwell time when dealing with gaze input, a dwell threshold of 450ms was adopted here, based on the literature [8]. In future research the task dependent threshold will be established experimentally. The results are shown in figure 6. The mean (μ) value was 334ms with a median value of 267ms which both fall significantly below the adopted 450 ms gaze threshold. The 25th percentile was 250ms and the 75th percentile was 391ms. Overall, 83% of all selection times fell below the chosen 450ms threshold.

The histogram in figure 6 shows that the timing data distribution is skewed, which makes the median value of 267ms the more reliable one. The outliers make for interesting observations. The fastest selection time was 124 ms and the slowest was 996 ms.

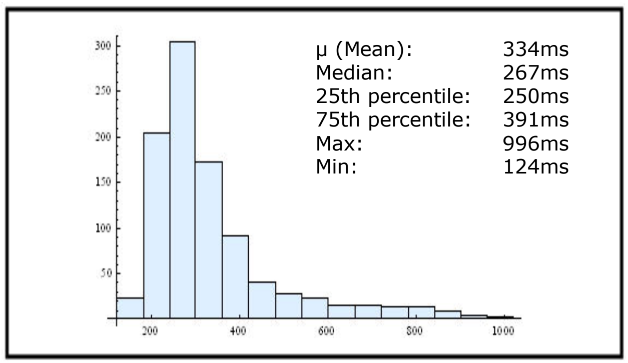

figure 6. Histogram of resultant times and Key data.

Furthermore, the histogram indicates that the cut off point of 1000ms does not have great influence. However, two questions could be posed: First of all, if this threshold had been higher would there have been instances of longer selection times? – Which leads to the more interesting question: What is causing these long selection times, when we know that a selection can take as little as 124ms and has a median time of 267ms? Future experiments will look to minimize selection times through design choices.

Combining Gestures and Dwell

The main motivation of this research is to develop gaze interaction which is efficient for motor-skill impaired users. Therefore, it is especially important that the burden of complexity be shouldered by the system, rather than the user. A way of achieving this would be to use both dwell time buttons and gaze gestures together. This addition of gestures, both single stroke and compound, gives the mono-modality of gaze input a new dimension. Thereby, potentially enabling greater amounts of information to be more accessible and condensed, one of the issues with many gaze controlled systems is that they have elaborate and complex menu structures which the user must navigate with few available means. There are three main areas which need addressing when developing systems for users with physical impairments: Communication, Environmental Control and Mobility.

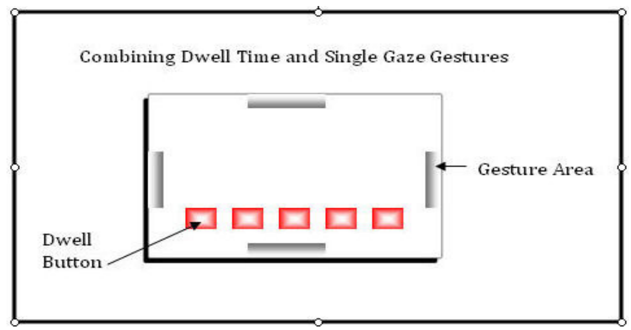

figure 7. A Potential way of combining Single Stroke Gaze Gestures and Dwell Time Buttons

We intend to integrate dwell time and single stroke gaze gestures into a full system which enables real world interaction.

The idea is to allow the users to interact directly with objects and paths in their environment. This approach is intended to facilitate and unify communication, environmental control and potentially wheelchair mobility in one single application. Communication will be implemented by allowing the user to 'point' with a wheelchair mounted camera at desired inanimate objects, rather than having to spell them out in a type-to-talk system or find a symbol in an intricate grid based structure. The intention is not to substitute these, but to supplement them. Environmental control could benefit greatly from this approach - the real world is to function as a super-menu allowing the user to access sub-menus by looking at an actual object, rather than having to flick through several layers of menus and information. Safe wheelchair mobility is very high on the wish-list of most impaired users, creating a gaze only implementation is part of the goal.

Combining these three elements into a fluid and flexible system, where all aspects are catered for within the same construct, is the ultimate goal of this research. However, this type of complexity requires a lot of a mono-modal input device such as the eye tracker, which is why sustainable forms of interaction are needed. An example of some basic functions which could benefit from the integration of dwell and gestures would be camera control by pan and zoom, which allows the user 360° of visual freedom. In figure 8 a potential implementation of this is shown. Zooming in and out is accomplished by single stroke gaze gestures and panning is done by using dwell time on specific target areas. All target areas are placed at the edge of the screen, allowing the user to freely search the screen content itself.

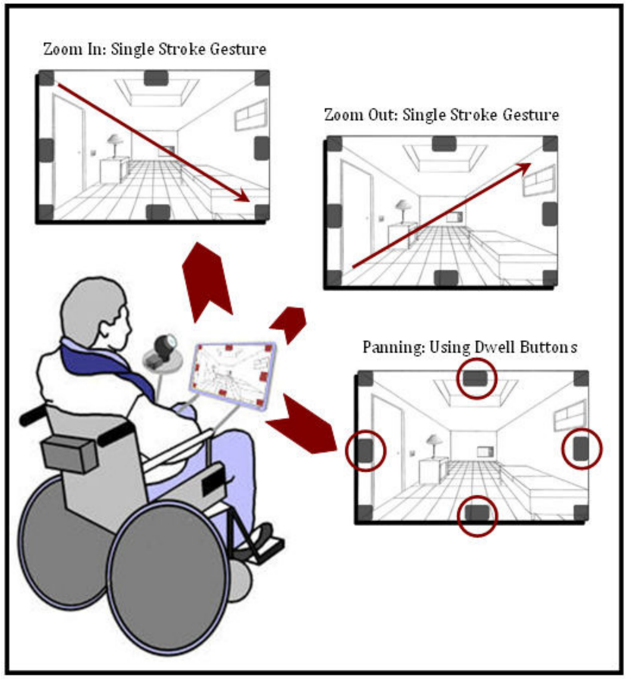

figure 8. A Potential Integration of Gestures and Dwell allowing the User to Navigate their Real World Environment

Conclusion

Single stroke gaze gestures as presented here could be used for various global commands, e.g. controlling zoom interfaces, flicking through websites etc. If gesture based interaction is to be implemented successfully, the way we structure and visualize information needs to be rethought. Applying single stroke gaze gestures to appropriate applications and tasks are subjects for the future. This work simply shows that the interaction method has potential.

Acknowledgements

This research was funded by the EPSRC and the COGAIN Network of excellence.

Example citations

[1] Figueiredo, L., Nunes, T., Caetano, F., and Gomes, A., Magic Environment, *The 4th COGAIN 2008, Proc. 15- 18*

[2] Gale, A. & Shi, F., Look In, Turn On, Drop Out. *Contemporary Ergonomics 2007*, 348

[3] Hansen, D.W., Skovsgaard, H.H.T., Hansen, J.P., Mollenbach, E. 2008. Noise tolerant selection by gaze-controlled pan and zoom in 3D. *Proc. 2008 ETRA*, 205-212.

[4] Istance, H., Hyrskykari,A., Vickers, S., Ali, N. 2008. User Performance of Gaze Based Interaction with On-line Virtual Communities. *The 4th COGAIN 2008, Proc. 28-32*

[5] Itoh, K., Aoki, H. & Hansen, J.P., 2006. A comparative usability study of two Japanese gaze typing systems. *Proc. 2006 ETRA*, 59-66.

[6] Novák, P., Krajnik, T., Preucil, L., Fejtova, M., Stepankova, O. 2008 AI support for a Gaze Controlled Wheelchair *The 4th COGAIN 2008, Proc. 19-22*

[7] Ward, D.J., Blackwell, A.F. & MacKay, D.J.C., 2002. Dasher: A Gesture-Driven Data Entry Interface for Mobile Computing. *Human-Computer Interaction*, 17(2 & 3), 199-228.

[8] Wobbrock, J.O., Rubinstein, J., Sawyer, M.W., Duchowski, A.T. 2008. Longitudinal evaluation of discrete consecutive gaze gestures for text entry. *Proc. 2008 ETRA* 11-18.