DTU Library

Eye Movements in Gaze Interaction

Møllenbach, Emilie; Hansen, John Paulin; Lillholm, Martin

[Link back to DTU Orbit](#)

Eye Movements in Gaze Interaction

Emilie Møllenbach
University of Copenhagen

John Paulin Hansen
IT University of Copenhagen

Martin Lillholm
Biomediq A/S

Gaze, as a sole input modality must support complex navigation and selection tasks. Gaze interaction combines specific eye movements and graphic display objects (GDOs). This paper suggests a unifying taxonomy of gaze interaction principles. The taxonomy deals with three types of eye movements: fixations, saccades and smooth pursuits and three types of GDOs: static, dynamic, or absent. This taxonomy is qualified through related research and is the first main contribution of this paper. The second part of the paper offers an experimental exploration of single stroke gaze gestures (SSGG). The main findings suggest (1) that different lengths of SSGG can be used for interaction, (2) that GDOs are not necessary for successful completion, and (3) that SSGG are comparable to dwell time selection.

Introduction

Gaze interaction entails using eye tracking technology to place direction of gaze in relation to a screen or other objects of interest. This can enable control of a computer and has been used by people with severe motor disabilities for decades (Majaranta & Räihä, 2002).

Research suggests that it may also become part of everyday interaction with TV, tablets, smartphones, and head mounted displays (Drewes & Schmidt, 2009). When implemented as mono-modal input, the main challenge becomes to determine when and what the user wants to interact with. When implemented in a multi-modal context the challenge becomes identifying appropriate means of control and how to fuse gaze modality with other input streams.

Eye tracking technology has yet to reach equilibrium: being consistent, stable and precise, while affordable and applied to mainstream computing. As a consequence there is a need to continuously define and explore both the fundamental and changing principles of interaction related to gaze.

The eyes as input are always 'on'. In a mono-modal context, this is referred to as the Midas touch problem (Bederson, Meyer, & Good, 2000; Jacob, 1991a). The mythical figure of King Midas was granted a wish for doing service to a god. He wished that all he touched would turn to gold. When this was granted, he quickly found a flaw in his plan, as all he touched, including food and drink, indeed turned to gold and he nearly starved to death. In gaze interaction, the story is synonymous with the user unintentionally selecting everything he looks at unless precautionary measures are taken. This is the essential design problem of mono-modal gaze interaction and a main focus of this paper.

The paper is divided into three main parts. First, the related work section introduces the basic components of gaze interaction: the physiology of the eye, eye movements, common contexts of use, and selection methods in gaze interaction. A taxonomy is subsequently introduced; this is intended to provide an understanding of how implementation of gaze interaction can be approached. Finally, an experiment concerning single stroke gaze gestures is presented. The basic rationale behind this selection strategy was to mimic the single finger swipe movement that has become so common in dextral gesture interaction.

Related Work

Physiology of The Eye

The eye funnels light through the cornea, pupil, lens, and vitreous body to the retina (Sekuler & Blake, 1994). The process of receiving and transforming light waves into visual information begins in the retina. There are approximately 130 million light receptive cells in each retina; these can be sub-divided into central and peripheral portions (Hubel, 1963) and are split into two types: cone and rod cells. Cone cells are responsible for our central daylight vision and are placed mainly in the macula, an area that covers a 5˚ angle. In the fovea, a small pit in the macula, cone cells are placed even more densely and are responsible for 1˚ of visual angle of high definition viewing or foveal vision. Rod cells cover the remaining retina and are responsible for peripheral vision and are active in situations with low illumination, such as nighttime.

The eyes move in a limited number of ways broadly defined as: fixations, saccades and smooth pursuits (Michael & Benjamin, 2009). The term fixation conveys the state in which the eye is relatively still and 'fixed' on a feature of interest. As the eye is constantly moving true fixation does not occur. The movements that happen during fixations are called micro-saccades, tremors, and drifts (Carpenter, 1988) and they scatter around 1˚ of the visual angle (Young & Sheena, 1975). These micro-movements ensure that new information constantly passes through the retina.

Fixations are separated by saccades; movements of the fovea from one fixation point to the next. Saccades are reactions, forced or unforced, to visual stimuli; physiologically it is the eyes' way of making up for not being entirely comprised of cone cells (Zeki & Marg, 1993); it needs to move from high resolution point to high resolution point. Saccades last between 30-120ms (milliseconds), and can reach speeds of up to 700˚/s for large amplitudes (Carpenter, 1991). Saccadic duration, average velocity, and peak velocity all increase as the size of the saccade increases (Bahill, Clark, & Stark, 1975). In other words the higher the amplitude of the saccade the longer but also faster the saccade (Carpenter, 1988; Duchowski, 2007; Robinson, 1964).

Where saccades move the eye from fixation point to fixation point – smooth pursuits can be described as a fixation-in-motion and happens when the eyes are following a moving target that should not exceed 100-200 deg/sec (Yarbus, 1967). Smooth pursuits require a target

in motion and a combination of saccades and fixations. The fixation stabilizes the object on the retina in order to retrieve stabile information, the saccadic movements perform minor corrections (Rashbass, 1961). Imagine throwing a ball and following it with your eyes. Usually smooth pursuits are split into two stages. First there is the open loop pursuit which is the initial motion to the target, this is saccade-like and usually lasts less than 100ms; after this comes the closed loop pursuit where the visual system is constantly trying to adjust the velocity of the retinal movement to that of the moving target (Duchowski, 2007; Krauzlis & Lisberger, 1994).

Fixations, saccades, and smooth pursuits are the foundation for the different interaction principles that will be explained in further detail in a subsequent section.

Fixation Based Interaction

One way of addressing the Midas's Touch problem has been by implementing onscreen dwell-buttons. The user maintains fixation on a graphic object on the screen for a prolonged period of time in order to complete a selection. Dwell-time activations can be set anywhere from 200ms to 1000ms (Jacob, 1991a; P. Majaranta, Ahola, & Špakov, 2009; Ware & Mikaelian, 1986) (Qvarfordt & Zhai, 2005), or more, depending on the abilities and preferences of the user. One of the first evaluations of eye tracking and dwell-time came from Ware in 1986 (Ware & Mikaelian, 1986). A few years later Jacob presented a study in 1991 where dwell-time was compared with eye pointing and button selection (Jacob, 1991b). While dwell-time affords gaze selections to be separate from inspection, it has some limitations. To compensate for the inaccuracy caused by micro-movements, tremors (Duchowski, 2007) and inaccuracy of the gaze tracking equipment, on-screen dwell-buttons usually take up quite a lot of screen space, limiting the amount of displayable information. Dwell selection sets a time limit on both the user's ability to inspect/navigate an interface and on the speed by which selections can be made.

Research into dwell-time based interaction is often done in conjunction with solutions to specific tasks, such as type-to-talk applications (Hansen & Hansen, 2006; Majaranta & Räihä, 2002), environmental control (Shi, Gale, & Mollenbach, 2008; Shi, Gale, & Purdy, 2006) and computer gaming (H. Istance, Bates, Hyrskykari, & Vickers, 2008; H. Istance, Hyrskykari, Immonen, Mansikkamaa, & Vickers, 2010; H. Istance, Vickers, & Hyrskykari, 2009). The main beneficiaries and target group of this research has been people with motor-skill

impairments, and the intention has been to provide usable software wherever possible.

Saccade Based Interaction

Gaze gestures based on saccadic movements are a more recent addition to the interaction vocabulary of gaze. Saccadic movements as strokes are the foundations of gaze gestures. A stroke is the motion between two intended fixations. It differs from a saccade, which can be defined as an eye movement between any two fixations. The concept of an anti-saccade gesture has also been explored in (Huckauf & Urbina, 2011). Here the goal is to train the eyes to move in the opposite direction of the visual stimuli. As this is a counterintuitive action it is easily distinguishable from regular search patterns. For the purpose of the research presented here, saccadic interaction is based on pro-saccades.

As with dwell-time there are both advantages and disadvantages with gaze gestures. Some of the advantages are:

Speed: gaze gestures can potentially be very fast (saccades can cover a 1° and 40° visual angle and last between 30-120ms (Duchowski, 2007).

Screen real estate: gaze gestures need not take up much screen space. The initiation and completion fields of the gaze gestures could be transparent, allowing for more on-screen information that is unaffected by gaze.

Midas' touch problem: Gaze gestures can be implemented so that they are distinguishable from navigational eye movements. Being able to distinguish between navigational eye movements and selections is key to addressing the Midas' Touch issue.

Some of the disadvantages are:

Accidental gesture completion (AGC): the potential overlap between natural search patterns and the eye movement patterns needed to complete a given gaze gesture. Avoiding this is fundamental when designing gaze gesture controlled interactions and interfaces.

Actual screen real estate: gaze gestures have mainly been implemented with semi-transparent or opaque graphic onscreen objects that still infringe on screen real-estate (H. Istance et al., 2010; H. O. Istance, Spinner, & Howarth, 1996; Vickers, Istance, Hyrskykari, Ali, & Bates, 2008; Wobbrock, Rubinstein, Sawyer, & Duchowski, 2008).

Saccadic fatigue: intentional pattern control of saccadic eye movements is not natural; as such it can quickly become a tiresome activity.

As mentioned, gaze gestures have had a shorter history than dwell-time. Istance et al. (H. Istance et al., 2010) proposed a definition of gaze gestures:

"*A definable pattern of eye movements performed within a limited time period, which may or may not be constrained to a particular range or area, which can be identified in real time, and used to signify a particular command or intent*"

Smooth Pursuit Based Interaction

Recently smooth pursuit is beginning to be explored distinctly as a selection strategy in a smooth pursuit text input system (Lorenceau, 2012). The interaction requires some training. However, the results show that participants gained control of smooth pursuit eye movements at varying rates. The authors also suggest that smooth pursuit can be used for other purposes such as drawing applications. However, there is still much room for research on the characteristics of smooth pursuit interaction.

Table 1
Overview of the interaction between eye movements and graphic display object.

	Fixations	Saccades	Smooth Pursuits	Characteristics
No Graphic Display Object	–	Single and Complex Gaze Gestures	–	Does not use screen real estate Does not require a screen Does not provide feedback in the selection process Limited vocabulary
Static Graphic Display Object	Static Dwell Objects	Single and Complex Gaze Gestures	–	Requires screen real estate Requires screen Provides selection feedback Limitless Vocabulary
Dynamic Graphic Display Object	Dynamic Dwell Interfaces	Dynamically Defined Gaze Gestures	Moving Targets	Requires screen real estate Requires screen Provides feedback Limitless Vocabulary

Taxonomy for Selection and Visualization in Gaze Interaction

In the following, a series of gaze selection strategies will be described as a combination of eye movements and graphic display objects (GDOs) (overview is shown in Table 1). The eye movement distinctions are the same as previously explained; fixations, saccades, and smooth pursuits. Graphic display objects (GDO) are divided as follows: no GDO (e.g., none of the onscreen GDO's are related to the selection principle used), static GDO (e.g., the graphic objects on the screen are static and used to direct gaze towards targets and provide selection feedback) and dynamic GDO (e.g., the graphic objects on the screen dynamically adapt, guide gaze direction, and supply visual feedback). The goal of this taxonomy is to categorize and exemplify different selection strategies in mono-modal gaze interaction. Some strategies have been extensively explored. However, the taxonomy also reveals potential selection strategies that could be the subject of future research.

Fixation Based Interaction

FIXATIONS AND NO GDOS

Fixation selections require a GDO order to be sustained. Therefore, fixation-based interaction with no GDO is not a viable option for gaze interaction purposes. In (Majaranta, MacKenzie, Aula, & Räihä, 2006) visual and auditory feedback strategies for dwell time interaction. Fixation without a specific GDO is the closest that has been explored. In the EyeDraw system, Hornof and Cavender (Hornof & Cavender, 2005) allowed drawing points to start anywhere on a blank canvas. However, a grid of dots (i.e. non-specific targets) was available to aid fixations. Fixation based interaction is always accompanied by either specific or non-specific GDOs.

FIXATIONS AND STATIC GDOS

Static GDO is the 'original' approach to handling the Midas' Touch issue. Static onscreen GDOs are presented to the user and fixating for a specific duration of time triggers a selection. This approach has been widely explored in type-to-talk systems that allow people with severe motor disabilities to communicate using onscreen keyboards (Majaranta & Räihä, 2002). GazeTalk is an example of a communication system built on fixation and static GDOs (Hansen & Hansen, 2006). Fixation based interaction with static GDOs is standard with most commercial interfaces.

FIXATIONS AND DYNAMIC GDOS

In this type of interface dwelling causes a dynamic shift in the GDOs. An example of dynamic GDO interface is NeoVisus (Tall, 2008) (Figure 1). The initial fixation causes the target to expand and reveal a second target that must subsequently be dwelled upon in order to complete a selection. The intention was to limit potential erroneous activations by requiring two subsequent fixations through dynamic target visualization.
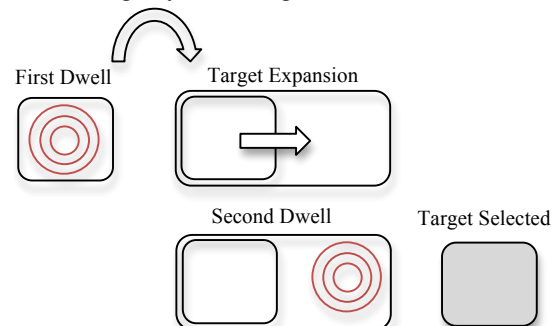

Figure 1: Adapted from NeoVisus, gaze selection components. A short dwell duration leads to the expansion of a second target. By dwelling on the second target the selection can be completed (Tall, 2008)

Saccade Based Interaction

SACCADES AND NO GDOS

An assumption has been that GDOs are needed in gaze interaction. However, the experimental research subsequently presented in this paper shows that gaze gestures can be completed without visual cues. The concept of implementing gaze interaction without visualization is important for three reasons: 1) it frees up screen real estate to be used for other things (i.e., a user could be browsing the web and use gaze gestures to navigate back and forth between pages, without the pages themselves being gaze contingent in any other way). 2) It introduces the possibility of combining dwell-time and gaze gestures in the same interface (i.e. a user could have several desktops with dwell GDOs for different purposes and use gestures to switch between them). And 3) it takes away the necessity of having a screen in front of the user (i.e., a camera system without a screen could be mounted on a wheelchair and looking rapidly from one camera to another could complete activation).

SACCADES AND STATIC GDOS

In designing saccadic gaze interaction static GDOs have been used to present patterns or grids in which specific patterns of eye movements can be completed. These completed patterns of eye movements are referred to as gestures (Heikkilä & Räihä, 2012; Mollenbach, Hansen, Lillholm, & Gale, 2009; Mollenbach, Lillholm, Gail, & Hansen, 2010).

In 2000, Isokoski et al. presented the MDITIM text input system (Poika Isokoski, 2000), where gestures were completed between large off-screen targets, (Figure 2, 1). Woobrock et al. (Wobbrock, Myers, & Kembel, 2003) presented EyeWrite, a gaze controlled text-input system, based on a text input system for PDAs called EdgeWrite, (Figure 2, 2). The user could map out letters by combining the four corners of a square in various ways (Wobbrock et al., 2003). Similarly, Porta and Turina (Porta & Turina, 2008) developed the EyeS system that had users enter letters by looking at fix-points in various sequences. Here the gaze patterns were designed to resemble the shape of the character that was being completed, (Figure 2, 3). They also made suggestions of using these types of gaze gestures for shortcuts in a windows interface (Porta & Turina, 2008).
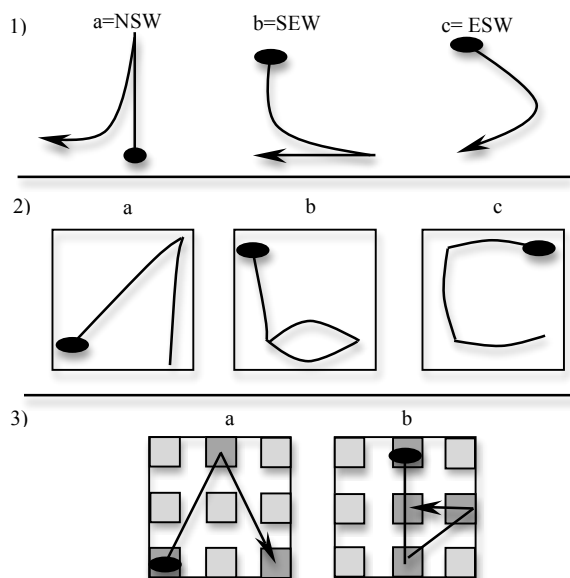

Figure 2: 1) Adapted from MDITIM gaze gestures (P. Isokoski, 2000). 2) Adapted from the eyewrite system (Wobbrock et al., 2008) 3) Adapted from EyeS gesture completion based on fix-points (Porta & Turina, 2008)

Drewes and Schmidt explored abstract path based gaze gestures where the completed gestures where not symbolic (H. Drewes & Schmidt, 2007) (Figure 3, 1). The main hypothesis in this work was that accuracy becomes less of an issue when gaze is not used for pointing.

Another example of saccadic interaction with static visualization is the typing interface developed by Morimoto and Amir. (Morimoto & Amir, 2010) (Figure 3,2). Two complete QWERTY keyboards where

presented on the screen. Two different eye movements where used to focus and select a target. A short dwell duration (150ms) informed the system of which target was in focus and a saccade across the opposite keyboard completed the selection. They called this selection strategy context switching. This is an example of combining gestures and dwell time in the same interaction.
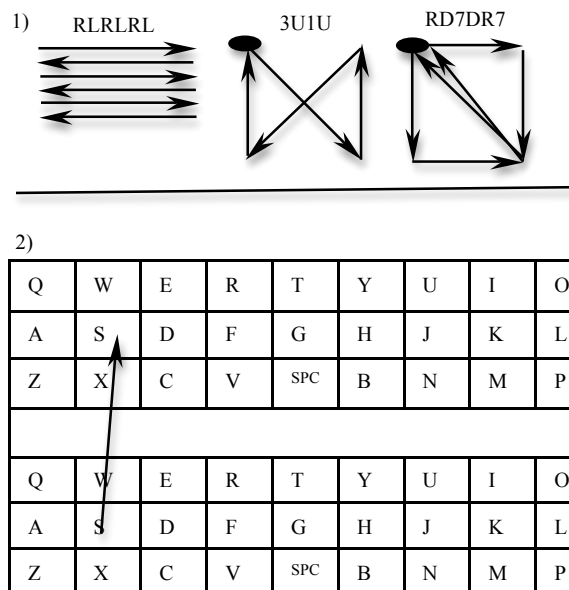

Figure 3: Gestures adapted from (H. Drewes & Schmidt, 2007) Context switching keyboard adapted from (Morimoto & Amir, 2010)

SACCADES AND DYNAMIC GDOS

A characteristic of dynamic GDOs is that content dynamically changes as the user interacts with the interface, an example of this is presented by Bee et al. (Bee & André, 2008). Inspired by Perlin's QuickWriting system, they explore text entry based on continuous input and argue that this is a well-suited text entry method for gaze. The user looks at the center of the dynamic GDO that contains several groups of letters and characters. The letters and characters contained in that group become highlighted and an individual character can be selected. (Bee & André, 2008) (Figure 4, 1).

In pEYEdit, suggested by Urbina et al. [48], the dynamic GDOs come in the form of expanding menus controlled by gaze gestures. Each slice contained a group of letters – when selected, by crossing the outer border of the slice, the group expands into a new pie where each slice had only one letter, which could then be selected based on the same selection principle (Figure 4, 2). Saccadic interaction is the only type of interaction that

can be completed within all three GDO principles, cf. Table 1.

Figure 4: 1) Example of a gesture in quickwriting adapted from (Bee & André, 2008). 2) The modified pEYE interface with third predictive expansion. (Urbina & Huckauf, 2010)

Smooth Pursuit Interaction

Smooth pursuit has not been investigated intentionally as a separate gaze interaction principle. However, we argue that smooth pursuit interaction has indeed been implemented; it has just not been defined as such. There is therefore much room for investigation into the affordances and constraints of this eye movement pattern. The only criterion that needs to be fulfilled is that the target being viewed is in motion during the fixation process. Because of this smooth pursuit selection cannot be completed in the no GDO or static GDO conditions.

SMOOTH PURSUIT AND DYNAMIC VISUALIZATION

An example of smooth pursuit as a selection principle can be found in StarGazer (Hansen, Skovsgaard, Hansen, & Mollenbach, 2008) (Figure 5). Pan and zoom were used to navigate the interface. The idea was to allow the user to navigate through graph structured data. The reason this can be considered a smooth pursuit selection is the fact that the target moves

while fixation is maintained. A target is chosen in one of the outer rims and the user then fixates on this target as it moves toward the center of the screen. When at the center fixation must be maintained in order to select the target. The system returns to a default view of the rims and a new target can be selected. Another example that has previously not been classified as a smooth pursuit selection is Dasher (Ward, Blackwell, & MacKay, 2000). Initially 27 characters are placed in a column to the right of the screen. To select the letter 't' the user looks at the letter 't' in the right column. The size of the letter begins to increase and move towards the left. Once the letter crosses the line dividing the screen, it is 'selected'. The user can reverse and 'deselect' a chosen letter at any time by looking at the left side of the centre line. The criterion for smooth pursuit selection is achieved by having the target being selected moving while fixation is maintained.
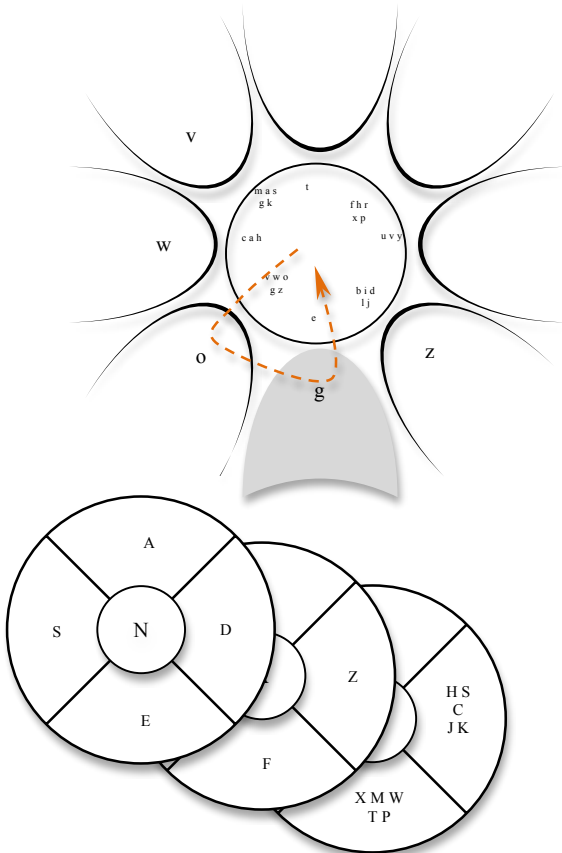
Figure 5: Adapted version of the Stargazer Interface, the cursor indicates the direction of gaze. (Hansen et al., 2008)

Another way of using smooth pursuit interactively has been explored in gaze controlled driving. In (Tall et al., 2009), smooth pursuit was used to guide a robotic vehicle. Direction was determined by the user maintaining focus on a self-determined way-point or target area while the robot moved toward it. As the robot moved the view presented to the user also moved. Target areas or waypoints were continuously redefined. In effect this meant that the robot was being controlled by a series of smooth pursuit loops. The system also used dwell selection when turning the robot on the spot; this was done by dwelling near the edges of the screen; another example of combining different selection strategies.

The calibration-free gaze interactive principle "Pursuits" (Vidal et al., 2013) was presented at the 2013 CHI

conference. Here people would select objects on the screen by following movement trajectories. Avoiding calibration procedures holds great potential for gaze interaction, for instance in information kiosks that require people to be able to just walk-up-and-use them. Pursuit interaction could also be used in gaze interactive systems for people with limited cognitive abilities who are not able to understand calibration instructions.

A final example of smooth pursuit selection was found in US patent 6,243,076 B1 (Hatfield, 2001). Objects, described as "comets", are imagined being moved in circular, elliptic or e.g. straight lines. Activations will then be determined on basis of the gaze-vector created by fixating the one comet that is the target for selection.

As mentioned in the beginning smooth pursuit has generally not been defined explicitly as a gaze interaction principle. However, the argument for including this when structuring gaze interaction principles is: As long as a target is moving while being fixated upon smooth pursuit interaction is occurring. A topic for future research will be to look at smooth pursuit interaction individually and determine the most appropriate use for this selection principle. An interesting feature of this interaction principle might be that the gaze tracker need not be accurately calibrated, since selection would happen on the basis of speed and direction of the GDO movement, and not the exact position of it.

Exploring Single Stroke Gaze Gestures

Gestures in regular HCI

Dextral gestures have become part of mainstream device navigation with the popularization of smartphones. Raskin defines a gesture as: 'a gesture is a sequence of human actions completed automatically once set in motion' (Raskin, 2000)

A gesture can, as a consequence, consist of any repeatable and recognizable bodily motion that can be robustly recognized as separate from noise related physical activity.

The introduction of single and multi touch dextral gestures for mobile devices, laptops, and tablets has propelled this form of interaction into the common sphere. Some of the most common dextral gestures are (a) tapping, (b) swiping and (c) pinching. Tapping resembles a button press and as with mouse interaction both single and double taps mean different things.

Swiping is often used to instigate a scrolling action and converse/inverse pinching allows the user to zoom in and out of the interface. The single stroke gaze gesture that will subsequently be explored in this section can be seen as the gaze equivalent of a single swipe finger gesture. The vocabulary of dextral gestures is increasing rapidly. Two-, three-, four- and five-finger gestures are gaining impasse (Buxton, 2007).

Stylus based text entry is another variation on this theme. The main similarity between gaze and stylus based interaction is that they both consist of one pointer connecting with the interface. Unistrokes (Goldberg & Richardson, 1993) was a gesture alphabet designed for the stylus. Other examples of this type of stylus alphabets are Cirrin (Mankoff & Abowd, 1998) and Grafitti (Blickenstorfer, 1995). Unsurprisingly, stylus based text entry has served as an inspiration for gaze gesture based text entry.

Various other types of physical gestures have also been explored when considering users with motor-skill impairments because they allow for individually adapted interaction. For example, head gestures consisting of simple horizontal and vertical motions have been used as switch control gestures (Keates & Robinson, 1998).

What are single stroke gaze gestures?

A distinction that needs to be made is between: single stroke gaze gestures (SSGG) and complex gaze gesture (Mollenbach et al., 2009, 2010). A SSGG can be defined as the motion between two intended fixations to complete activation. The complex gaze gesture can be defined as the motion between three or more intended fixation points.

Complex gaze gestures have the advantage of greatly increasing the interaction vocabulary of gaze. However, this brings with it both cognitive and physiological difficulties. Cognitively it is difficult to remember a large number of gestures and physiologically it is difficult to create and complete them (Porta & Turina, 2008). Like dwell selection, SSGG represent one action and are easily learnt.

Single stroke gaze gestures are intended for basic limited navigation (i.e., switching from a communication application to an environmental control application). This would be equivalent to toggling in Windows that is performed by the 'Ctrl+tab' function or the 'Spaces' functionality on the Mac that allows users to switch between various desktop environments.

In subsequent sections research on the SSGG will be presented. This will contribute to the understanding of whether or not SSGG can be employed as a selection method and what the innate properties of this selection strategy are.

Experiment

The experiment presented here examined three aspects concerning SSGG. First, the difference between SSGG with and without GDOs. If SSGG can be completed without GDOs they can be used as basic navigation for a variety of interfaces without having to modify them specifically for gaze. Secondly, comparing two different lengths of SSGG, this was done to investigate the assumption that selection patterns close to the centre of the screen would be more prone to correlation with search patterns. Finally, a comparison between SSGG and dwell selection was done in order to explore the usability of SSGG relative to the main selections strategy used in gaze interaction.

Eighteen participants (9 female) completed the experiment. Seven had used an eye tracker before. Each participant was given an introduction to the experiment. The gesture and dwell selection strategies were explained, as well as the task.. A demonstration of each selection strategy was given and the participant practiced each of the selection strategies for a few minutes before initiating the experiment.

The experiment was balanced by having participants complete three overall conditions (long SSGG, short SSGG and dwell selection). Furthermore, the dwell selection condition switched between beginning with 25ms and beginning with 400ms. All of the SSGG conditions started by having visualization on and subsequently off. This was based on the notion that gaze gestures without feedback need to be learnt.

Design

The application was written in Java and the experiment was completed on a LC technologies eye follower system on a 60 Hz 17" monitor for each eye, set for 120 Hz frame rate.

The single stroke gaze gestures were implemented as the motion between an initiation field and a completion field on the screen. This had to be done within 1000ms or the event would reset. If a SSGG selection was initiated but gaze entered one of the adjacent fields the selection process would be cancelled and the system would reset. The timer event was fired when the boundary of an initiation field was crossed and ended when gaze moved across the boundary of a completion field on the opposite side. Figure 6 (1,2,3) shows the three basic states of the interface. There were four basic SSGG selections: left-to-right, right-to-left, top-to-bottom, bottom-to-top and four dwell buttons: top, bottom, right and left.

A series of arrowshaped targets descended the screen at a constant speed. The direction of the arrow indicated which way the gestures should be completed or which dwell button to select. The bottom most target was the object for selection (Figure 6, 4).
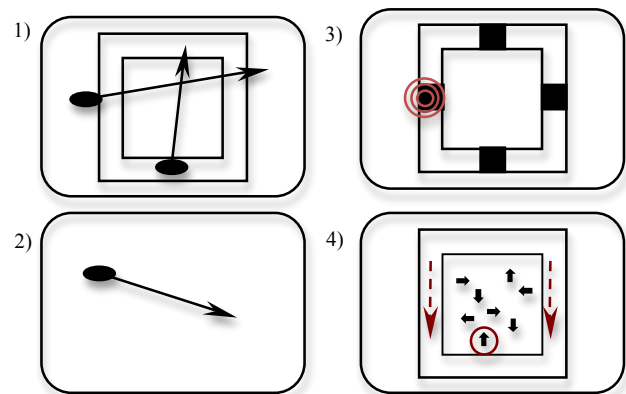

Figure 6: 1) Long and short gestures, 2) Without GDOs, and 3) Dwell GDO Interface, 4) Task targets. The small black arrows descend the screen; the bottommost arrow is the target the user must select..

Feedback was given in three ways depending on the three overall conditions. Firstly, in SSGG the selection fields were visible and shifted to a light grey colour when looked upon. Secondly, in the no- visualization SSGG conditions the target would disappear after selection. Finally, in the dwell selection condition feedback was given on the dwell button.

A 17'' monitor being viewed at 50 cm distance equates to ~ 37° visual angle. Short SSGG required the user to cover 30% ~ 11,5° visual angle. Long SSGG required the user to cover 60% of the screen ~ 11,5° [4] visual angle.

In the dwell condition the inner border of the dwell buttons were fitted to the borders of the short SSGG condition and the outer borders were fitted to the borders of the long SSGG. This way the center of the dwell button was exactly in the middle of the short and long SSGG conditions.

INDEPENDENT VARIABLES:

Selection method: Overall there were nine different selection methods, based on four separate parameters;

first, the overall selection methods of SSGG and dwell selection; secondly, for SSGG, selection length (long and short SSGG); thirdly, for SSGG, with and without visualization and finally, for dwell selection, five increments of fixation duration time.

The nine levels of selection method were therefore: Long SSGG with visualization, long SSGG without visualization, short SSGG with visualization, short SSGG without visualization and dwell selection: 25 ms, 50ms, 100ms, 200ms and 400ms.

Each participant had to complete 20 successful selections per level of selection method.

DEPENDENT VARIABLES:

Selection Time: the time from when the user exits the initiation field and enters the opposite field measured in milliseconds (ms) – only measure for the SSGG conditions

Task completion Time: the time elapsed between each successful selection. Both accumulated and individual task times were recorded and measured in milliseconds (ms).

Selection Error: A full completed selection which does not respond to the current target.

Results

LONG AND SHORT SSGGS

The goal of the comparison between long and short SSGGs was to examine whether there would be a difference in selection and task completion time between the two conditions and whether shorter SSGGs would be more error prone.

The data were explored and Levene's test for homogeneity of variance was found to be significant, so the variance between the long and short SSGG datasets respectively was too great to employ parametric statistics. Also the data were not normally distributed. After ranking the scores a Wilcoxon signed ranks test was used to compare the means.

Selection Completion Time: There was a significant difference in the selection completion time between long and short SSGG. Short SSGG were significantly faster. SSGG $Z = -12,915$; $p < .001$.

Task Completion Time: There was a significant difference in task completion time between long and short SSGG. Short SSGGs were had faster task completion times. $Z = -3,671$; $p < .001$.

Selection Error: There was no significant difference between the number of selection errors for long and short SSGG. $Z = -0,306$; $p = 0,759$.

Table 2
Descriptive statistics for Long and Short Gestures.

Dependent Variable	Independent Variable	N	Mean	Std. Deviation
Selection Completion Time	Short SSGG	720	118,68	134,19
	Long SSGG	720	186,50	148,89
Task Completion Time	Short SSGG	720	1747,14	1035,66
	Long SSGG	720	2151,30	2183,07
Selection Error	Short SSGG	36	3,89	2,11
	Long SSGG	36	4,72	4,54

SSGGS WITH AND WITHOUT GDOS

As mentioned earlier one of the issues in gaze interaction has been the requirement of GDOs to guide and provide feedback in gaze contingent interfaces. The comparison between SSGGs with and without visual feedback was to determine whether or not it would be possible to complete SSGGs without visual feedback.

For selection completion times the data were explored and the assumptions for the use of parametric statistics were not met, so a Wilcoxon signed ranks test was used to compare the means. For task completion times Levene's test for homogeneity of variance was found to be insignificant $p = 0,145$. By squaring the data a normal distribution was achieved. For selection error Levene's test for homogeneity of variance was found to be insignificant $p = 0,602$. After logarithmically transforming the data a normal distribution was achieved and parametric statistics could therefore be preformed

Selection Completion Time: There was a significant difference between the selection completion times for SSGG with and without Graphic Display Objects. SSGGs without GDOs were significantly faster. $Z = -11,401$; $p < 0,001$.

Task Completion Time: There was no significant difference in task completion time for SSGG with or without visual feedback t (719) = - 0,199; p = 0,843.

Selection Error: There was no significant difference in the number of selection errors for SSGG with or without visual feedback t (35) = -0,626; p = 0,535

Table 3
Descriptive statistics for Gestures With and Without Graphic Display Objects

Dependent Variable	Independent Variable	N	Mean	Std. Deviation
Selection Completion Time	Visualization On	720	190,07	171,91
	Visualization Off	720	115,11	100,58
Task Completion Time	Visualization On	720	1940,39	1612,93
	Visualization Off	720	1958,05	1821,67
Selection Error	Visualization On	36	3,89	2,11
	Visualization Off	36	4,72	4,54

DWELL SELECTION

The reason for exploring a dwell condition in conjunction with the SSGG condition was to determine what dwell durations SSGGs could be compared to and whether they would be more error prone. Below are the results of the dwell condition. The data for dwell were explored. Mauchly's test of sphericity was significant and the data were skewed, so the assumptions that need to be fulfilled in order to conduct parametric statistics were not met. Friedman's non-parametric ANOVA was used to compare the means and a post hoc analysis was used to determine where the differences were.

Table 4
Descriptive Statistics for Task Completion Times for the Dwell Condition.

Dependent Variable	Independent Variable	N	Mean	Std. Deviation
Task Completion Times, Dwell	25 ms	360	1550,30ms	451,29
	50 ms	360	1547,26ms	321,52
	100 ms	360	1538,45ms	303,26
	200 ms	360	1584,12ms	570,24
	400 ms	360	2532,30ms	3973,94

Selection Error, Dwell	25 ms	17	11,06	3,44
	50 ms	17	7,06	2,82
	100 ms	17	8,06	3,41
	200 ms	17	3,78	1,13
	400 ms	17	1,88	0,88

Task Completion Time: There was an overall significant difference in task completion time for the different dwell conditions t (719) = - 0,199; p = 0,843. The 400ms condition had significantly longer task completion times.

Selection Error: There was no overall significant effect between the number of selection errors for all of the dwell conditions. X ($\_r^2$)= 6,349 (4, n=18), p < 0,175.

OVERALL COMPARISON OF TASK COMPLETION TIMES AND SELECTION ERROR

The task in this experiment was the same for all conditions. In order to examine how SSGGs compared to dwell an overall comparison of the task times was conducted.

Table 5
Descriptive Statistics for Overall Task Completion and Selection Errors

Dependent Variable	Independent Variable	N	Mean	Std. Deviation
Task Completion Times, Overall Comparison	Short SSGG GDO On	360	1739,32	741,38
	Short SSGG GDO Off	360	1754,95	1264,28
	Long SSGG GDO On	360	2141,45	2140,00
	Long SSGG GDO Off	360	2161,15	2228,24
	25 ms	360	1550,39	451,29
	50 ms	360	1547,26	321,52
	100 ms	360	1538,45	303,26
	200 ms	360	1584,12	570,24
	400 ms	360	2532,30	3973,94
Selection Error, Dwell	Short SSGG GDO On	17	3,82	2,19
	Short SSGG GDO Off	17	4,06	2,08
	Long SSGG GDO On	17	4,18	4,23
	Long SSGG GDO Off	17	5,12	5,09
	25 ms	17	16,18	23,63
	50 ms	17	11,94	21,86

100 ms	17	8,33	14,75
200 ms	17	4,00	4,84
400 ms	17	5,88	16,41

The data were explored. Sphericity could not be assumed so non parametric statistics were used. Friedman's non parametric ANOVA was used to compare the means, where selection method was the independent variable with nine levels (short SSGG with visualization, short SSGG without visualization, long SSGG with visualization, long SSGG without visualization, 25ms dwell, 50ms dwell, 100ms dwell, 200ms dwell, 400ms dwell).

Task Completion Time: There was an overall significant effect between the task completion times for all the conditions 43,568 (8, n=360), p < 0,01. 400ms dwell was significantly slower than all conditions, except both the long SSGG conditions. The long SSGG conditions had significantly longer task times than all conditions except the 400ms dwell and the short SSGG without visual feedback. None of the other conditions were significantly different to one another.

Selection Error: There was no overall significant effect between the number of selection error for any of the conditions 11,712 (8, n=17), p = 0,165.

Discussion

LONG AND SHORT SSGG

The selection completion time for long and short SSGG showed a significant difference, with short being significantly faster than long SSGG. The main problem with this result, as mentioned earlier, is the lacking systematic approach in testing several different lengths and establishing intervals that might be used as robustly recognisable input patterns. The results are not that surprising; there is evidence from the field of eye movements during reading that show a difference in the completion time of long and short saccades. (Abrams, Meyer, & Kornblum, 1989; Collewijn, Erkelens, & Steinman, 1988a, 1988b). However, the research presented here indicates that length of eye movements could be used as a factor in interface design. The fast selection time also effects the task completion time, which is also significantly different. However, in reality this difference would be inconsequential. The main argument for further investigation into this area is not one of efficiency, but one of flexibility. Having an interface

that is sensitive to various lengths of saccades means that an SSGG in a specific direction could have different consequences depending on length. Figure 7 illustrates a bull's eye type interface this interaction principle could result in.
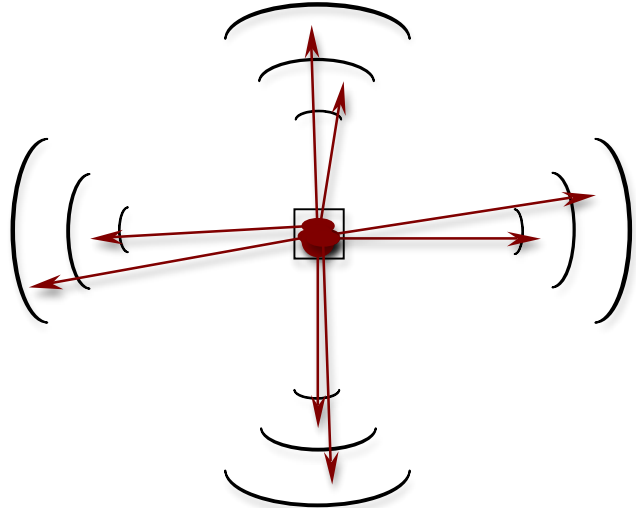

Figure 7: A bull's eye interface for long and short gestures

The surprising result regarding long and short SSGG was the lack of significance in the number of selection errors. It had been expected that the short SSGG would result in a higher degree of accidental gesture completion, because the initiation and completion borders were closer to the target field being inspected. However, further investigation with varying display content is required to substatiate this. However, the results show the plausibility of using length as a variable in saccadic controlled interfaces.

SSGG WITH AND WITHOUT GDOS

To the knowledge of the authors, no one else has tried to look at gaze based selection strategies without GDOs. Probably because it seems counter intuitive. Saccades are designed to allow the eye to move from fix point to fix point, not for arbitrary motion. The argument presented here is that arbitrary motions are possible, recognisable and simple to do. The potential of no visualization means that interaction can occur separately from the on screen content. The most surprising result in the context of this experiment was that selection times for the no GDO condition were significantly lower than for the SSGG condition with GDOs. The reason for this could not be related to the fact that it takes time to perceive the changing graphics. Because, as explained earlier, selection completion time was measured as the crossing of two borders this was done in order to isolate the actual selection from navigational tasks. A reason could be that people were using the edges of the screen as

self-determined reference points; this would of course have meant that they were doing much larger saccades but the larger amplitude would have meant higher velocities resulting in shorter selection times for the no visualization selection times. This final point is further validated by the fact that there was no difference in task time for the two conditions. Again the argument for using no GDO SSGG is not efficiency, but flexibility. The fact that there were no substantial differences in the number of selection errors between the two conditions further supports the usability of no visualization SSGG.

It is the potential of no visualization SSGG that is the main experimental contribution of this research. This type of selection could be used for simple navigation in a web browser, as basic channel switching of a TV or even as the play, pause, fast-forward, fast-backward functions on a DVD or similar device, as shown in Figure 8.
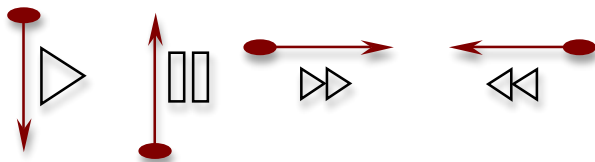

Figure 8: A video control system using gestures without visual representation

DWELL CONDITION

The dwell condition was implemented to enable a comparison between dwell and SSGG within the same type of task. The results regarding the different dwell completion times are based on movements in all four directions. This therefore balances out any directional bias that might have occurred. This experiment was designed with the exploration of single stroke gaze gestures in mind. However, two interesting findings were made in this condition. The first is the insignificance of dwell times under 200ms on task completion times. The reason is that the amount of time it takes to perceive the task and navigate to the correct selection evens out any immediate benefits of fast selection. At 400 milliseconds selection time the task completion time became significantly different. This indicates that in terms of efficiency 200 milliseconds could be a significant boundary. This should be substantiated through dedicated experiments in different types of layout. There is no doubt that the layout played a big difference in the number of selection errors. Under many other circumstances a dwell time of 25ms would have resulted in a high number of errors. However, here there was no significant difference between any of the conditions. This was due to the layout and nature of the task. Each

selection path was separate from all the others. The participant would look at the centre of the screen and glance outward in one of four directions in order to complete a selection. The task target indicated the direction, which meant that the cognitive load was very low. The combination of path separation and low cognitive load enabled very fast selection completion times as well as a low error rate. These factors should therefore be taken into consideration when designing dwell time interfaces.

OVERALL COMPARISON

One common denominator for all of the conditions was task completion time. The results here show that there is great overlap between all the conditions.

400ms dwell and both the long SSGG conditions were in one group that had significantly slower task completion times. Short SSGG and the 25-200ms dwell conditions were in another group that had significantly faster task completion times.

The issues regarding task completion times have already been discussed for the SSGG and dwell conditions separately. This overall comparison was simply done in order to examine whether one of the overall selection strategies, dwell or SSGG, stood out, which they did not. This was a further indication that SSGG and dwell are not that far apart in terms of being effective strategies for solving similar tasks. This is further indicated by the lack of difference in error rate between the conditions.

Conclusion

There are two main contributions of this research: 1) the taxonomy of gaze interaction based on eye movement characteristics and visualization and 2) the introduction of SSGG as an effective and efficient scheme for limited interaction. The distinction between fixation -, saccade- and smooth pursuit-based interaction with and without visualization will make it more manageable to define, create and compare existing and new selection strategies. In regard to single stroke gaze gestures the conclusion is that they could be well suited for simple change-of-context tasks and that they may become a useful addition to the toolbox being developed for gaze interaction. The toolbox consists of a range of physiological affordances (fixations, saccades and smooth pursuits) and visualization principles (no visualization, static and dynamic visualization). The multitude of ways that these parameters can be combined is the foundation of gaze interaction. Paring the

consequences of actions with the appropriate methods and visualization constitutes one of the fundamental challenges in interface and interaction design.

The key to implementing gaze selection strategies is combine the different selection strategies that have been explored over the past few decades in order to develop appropriate interaction patterns for control of gaze controlled interfaces. Designing interfaces that rely on mixtures of dwell time, gesture selection and/or smooth pursuit. The taxonomy, distinctions and methods presented in this work will aid in ensuring this.

References

Abrams, R. A., Meyer, D. E., & Kornblum, S. (1989). Speed and accuracy of saccadic eye movements: Characteristics of impulse variability in the oculomotor system. *Journal of Experimental Psychology: Human Perception and Performance, 15(3)*, 529.

Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences, 24,* 191–204.

Bederson, B. B., Meyer, J., & Good, L. (2000). Jazz: an extensible zoomable user interface graphics toolkit in Java. In *Proceedings of the 13th annual ACM symposium on User interface software and technology* (pp. 171–180).

Bee, N., & André, E. (2008). Writing with your eye: A dwell time free writing system adapted to the nature of human eye gaze. *Proceedings of the 4th IEEE tutorial and research workshop on Perception and Interactive Technologies for Speech-Based Systems: Perception in Multimodal Dialogue Systems*, 111–122.

Blickenstorfer, C. H. (1995). Graffiti: wow!. *Pen Computing Magazine, 1*, 30–31.

Buxton, B. (2007). Multi-touch systems that I have known and loved. Microsoft Research. Retrieved http://www.billbuxton.com/multitouchOverview.html (retrieved April 4, 2013)

Carpenter, R. H. S. (1988). *Movements of the eyes*. London: Pion.

Carpenter, R. H. S. (1991). The visual origins of ocular motility. *Vision and visual function, 8*, 1–10.

Collewijn, H., Erkelens, C. J., & Steinman, R. (1988a). Binocular co-ordination of human vertical saccadic eye movements. *Journal of Physiology, 404(1)*, 183–197.

Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988b). Binocular co-ordination of human horizontal saccadic eye movements. *Journal of Physiology, 404(1)*, 157–182.

Drewes, H., & Schmidt, A. (2007). Interacting with the computer using gaze gestures. *Proceedings of the 11th IFIP TC 13 international conference on Human-computer interaction-Volume Part II* (pp. 475–488).

Drewes, Heiko, & Schmidt, A. (2009). Interacting with the Computer Using Gaze Gestures. *Human-Computer Interaction – INTERACT 2007* (pp. 475–488). Retrieved from http://dx.doi.org/10.1007/978-3-540-74800-7_43

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice.* New York Springer.

Goldberg, D., & Richardson, C. (1993). Touch-typing with a stylus. *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 80–87).

Hansen, D. W., & Hansen, J. P. (2006). Eye typing with common cameras. *Proceedings of the 2006 symposium on Eye tracking research & applications* (p. 55).

Hansen, D. W., Skovsgaard, H. H. ., Hansen, J. P., & Mollenbach, E. (2008). Noise tolerant selection by gaze-controlled pan and zoom in 3D. *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 205–212).

Hatfield, F. (2001). System and method for controlling host system interface with point-of-interest data. *United States Patent US 6,243,076 B1*.

Heikkilä, H., & Räihä, K. J. (2012). Simple gaze gestures and the closure of the eyes as an interaction technique. In Proceedings of the Symposium on Eye Tracking Research and Applications (pp. 147–154). Retrieved from http://dl.acm.org/citation.cfm?id=2168579 (retrieved April 4, 2013)

Hornof, A. J., & Cavender, A. (2005). EyeDraw: enabling children with severe motor impairments to draw with their eyes. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 161–170).

Hubel, D. H. (1963). The Visual Cortex of the Brain. *Scientific American, 209*, 2–10.

Huckauf, A., & Urbina, M. H. (2011). Object selection in gaze controlled systems: What you don't look at is what you get. *ACM Transactions on Applied Perception, 8(2)*, 13:1–13:14. doi:10.1145/1870076.1870081

Isokoski, P. (2000). Text input methods for eye trackers using off-screen targets. *Proceedings of the 2000 symposium on Eye tracking research & applications* (p. 21).

Isokoski, Poika. (2000). Text input methods for eye trackers using off-screen targets. *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 15–21). Palm Beach Gardens, Florida, United States: ACM. doi:10.1145/355017.355020

Istance, H., Bates, R., Hyrskykari, A., & Vickers, S. (2008). Snap clutch, a moded approach to solving the Midas touch problem. *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 221–228).

Istance, H., Hyrskykari, A., Immonen, L., Mansikkamaa, S., & Vickers, S. (2010). Designing gaze gestures for gaming: an investigation of performance. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 323–330).

Istance, H. O., Spinner, C., & Howarth, P. A. (1996). Providing motor impaired users with access to standard Graphical User Interface (GUI) software via eye-based interaction. *Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies (ECDVRAT'96)* (pp. 109–116).

Istance, H., Vickers, S., & Hyrskykari, A. (2009). Gaze-based interaction with massively multiplayer on-line games. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4381–4386).

Jacob, R. J. . (1991a). The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS), 9(2)*, 152–169.

Jacob, R. J. . (1991b). The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS), 9(2)*, 152–169.

Keates, S., & Robinson, P. (1998). The use of gestures in multimodal input. *Proceedings of the third international ACM conference on Assistive technologies* (pp. 35–42).

Krauzlis, R. J., & Lisberger, S. G. (1994). Temporal properties of visual motion signals for the initiation of smooth pursuit eye movements in monkeys. *Journal of Neurophysiology, 72(1)*, 150.

Lorenceau, J. (2012). Cursive Writing with Smooth Pursuit Eye Movements. *Current Biology, 22(16)*, 1506–1509. doi:10.1016/j.cub.2012.06.026

Majaranta, P., Ahola, U. K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 357–360).

Majaranta, P., & Räihä, K. J. (2002). Twenty years of eye typing: systems and design issues. *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 15–22).

Majaranta, Päivi, MacKenzie, I. S., Aula, A., & Räihä, K.-J. (2006). Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society, 5(2)*, 199–208.

Mankoff, J., & Abowd, G. D. (1998). Cirrin: A word-level unistroke keyboard for pen input. *Proceedings of the 11th annual ACM symposium on User interface software and technology* (pp. 213–214).

Michael, L., & Benjamin, T. (2009). *Looking and acting: Vision and eye movements in natural behaviour.* Oxford: Oxford University Press.

Mollenbach, E., Hansen, J. P., Lillholm, M., & Gale, A. G. (2009). Single stroke gaze gestures. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4555–4560).

Mollenbach, E., Lillholm, M., Gail, A., & Hansen, J. P. (2010). Single gaze gestures. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 177–180).

Morimoto, C. H., & Amir, A. (2010). Context switching for fast key selection in text entry applications. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 271–274).

Porta, M., & Turina, M. (2008). Eye-S: a full-screen input modality for pure eye-based communication. *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 27–34).

Qvarfordt, P., & Zhai, S. (2005). Conversing with the user based on eye-gaze patterns. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 221–230).

Rashbass, C. (1961). The relationship between saccadic and smooth tracking eye movements. *Journal of Physiology, 159(2)*, 326.

Raskin, J. (2000). *The humane interface: new directions for designing interactive systems.* New York: ACM Press/Addison-Wesley.

Robinson, D. A. (1964). The mechanics of human saccadic eye movement. *Journal of Physiology, 174(2)*, 245.

Sekuler, R., & Blake, R. (1994). *Perception,* 3rd edition. New York: McGraw-Hill.

Shi, F., Gale, A., & Mollenbach, E. (2008). Eye, Me and the Environment. Computers Helping People with Special Needs, *Lecture Notes in Computer Science, 5105,* 1030–1033. New York: ACM-Press.

Shi, F., Gale, A., & Purdy, K. (2006). Helping people with ICT device control by eye gaze. Computers Helping People with Special Needs, *Lecture Notes in Computer Science, 4061,* 480–487. New York: ACM-Press.

Tall, M. (2008). NeoVisus:Gaze Driven Interface Components. Presentation at the 4th Conference on Communication by Gaze Interaction (COGAIN 2008).

Tall, M., Alapetite, A., San Agustin, J., Skovsgaard, H. H., Hansen, J. P., Hansen, D. W., & Mollenbach, E. (2009). Gaze-controlled driving. *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems* (pp. 4387–4392).

Urbina, M. H., & Huckauf, A. (2010). Alternatives to single character entry and dwell time selection on eye typing. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 315–322).

Vickers, S., Istance, H. O., Hyrskykari, A., Ali, N., & Bates, R. (2008). Keeping an eye on the game: eye gaze interaction with Massively Multiplayer Online Games and virtual communities for motor impaired users. *Proceedings of the Seventh International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT)*

Vidal, M., Pfeuffer, K., Bulling, A., and Gellersen, H.W. (2013). Pursuits: eye-based interaction with moving targets. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '13) (3147-3150).

Ward, D. J., Blackwell, A. F., & MacKay, D. J. (2000). Dasher—a data entry interface using continuous gestures and language models. *Proceedings of the 13th annual ACM symposium on User interface software and technology* (pp. 129–137).

Ware, C., & Mikaelian, H. H. (1986). An evaluation of an eye tracker as a device for computer input2. *ACM SIGCHI Bulletin, 17(SI)*, 183–188.

Wobbrock, J. O., Myers, B. A., & Kembel, J. A. (2003). EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion. *Proceedings of the 16th annual ACM symposium on User interface software and technology* (pp. 61–70).

Wobbrock, J. O., Rubinstein, J., Sawyer, M. W., & Duchowski, A. T. (2008). Longitudinal evaluation of discrete consecutive gaze gestures for text entry. *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 11–18).

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Pess.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods and Instrumentation, 7(5)*, 397–429.

Zeki, S., (1993). *A Vision of the brain*. Oxford: Blackwell.