Entering PIN Codes by Smooth Pursuit Eye Movements

Dietlind Helene Cymek
Technische Universität Berlin

Antje Christine Venjakob
Technische Universität Berlin

Stefan Ruff
Technische Universität Berlin

Otto Hans-Martin Lutz
Technische Universität Berlin

Simon Hofmann
Technische Universität Berlin

Matthias Roetting
Technische Universität Berlin

Despite its potential gaze interaction is still not a widely-used interaction concept. Major drawbacks as the calibration, strain of the eyes and the high number of false alarms are associated with gaze based interaction and limit its practicability for every-day human computer interaction. In this paper two experiments are described which use smooth pursuit eye movements on moving display buttons. The first experiment was conducted to extract an easy and fast interaction concept and at the same time to collect data to develop a specific but robust algorithm. In a follow-up experiment, twelve conventionally calibrated participants interacted successfully with the system. For another group of twelve people the eye tracker was not calibrated individually, but on a third person. Results show that for both groups interaction was possible without false alarms. Both groups rated the user experience of the system as positive.

Keywords: gaze interaction, gaze control, eye tracking, smooth pursuit

Introduction

Gaze Interaction Concepts

Gaze interaction is an innovative form of human machine interaction. Having its field of application originally in providing physically disabled users with a possibility to communicate (Majaranta & Räihä, 2002), more and more research aims at applying gaze interaction to every-day human computer interaction (Drewes & Schmidt, 2009). The often cited advantages of gaze interaction are freeing hands for other tasks and increasing hygiene due to contactless interaction. In an operating theater, surgeons may e.g. benefit from using their hands entirely for the primary task, which is navigating instruments, while controlling monitors by gaze.

Gaze interaction may not only comprise pointing, thus substituting the mouse cursor by gaze, but also selecting elements of interest, equivalent to clicking onto them by mouse button. For the selection of objects three different mechanisms have been applied in the past: dwell based, blink based and gesture based interaction.

Dwell based gaze interaction. Selecting screen elements by gazing at them for a prolonged amount of time is the most established implementation of gaze interaction and has been suggested as early as in 1982 (Friedman, Kiliany, Dzmura & Anderson, 1982). Until today, dwell based interaction is probably the most widely applied gaze interaction concept. The threshold for activation has varied substantially between experiments, spanning from 100 ms up to 1000 ms. Short dwells are reported to be associated with the advantage of fast interaction at the cost of frequent errors, whereas longer dwells are more time consuming and supposedly

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

strain the eye, yet reducing the number of involuntary errors (Majaranta, MacKenzie, Aula & Räihä, 2006). To circumvent these drawbacks, individualized dwell times have been developed (Spakov & Miniotas, 2004).

Blink based gaze interaction. Voluntary eye blinks have also been used to select screen elements. Here typically a prolonged interval of gaze absence is detected. The last fixation location before the disappearance of the gaze is then identified and selected. Compared to dwell based selection methods, blinks have attracted less research interest (Heikkilä & Räihä, 2012). This may be due to unease by the user because visual intake is interrupted during the selection process (Koesling, Zöllner, Sichelschmidt & Ritter, 2009).

Gesture based gaze interaction. Gaze gestures are a relatively recent interaction technique. The concept of gaze gestures comprises a number of divers approaches. In many implementations the gestures are used as symbols representing distinct operations (Heikkilä & Räihä, 2009) such as opening a new tab or closing an application. Other experiments used gaze gestures to draw letters or numbers with the eyes (De Luca, Weiss & Drewes, 2007). This application is thus closer to handwriting than to performing commands. Gaze gestures were often performed as eye movements between markings on a template. However, in more recent implementations, gaze gestures are trained and then performed from memory without a template (Drewes & Schmidt, 2007). Studies have shown that this may be just as reliable as gaze gestures performed with the help of a template, while time to complete a gesture was lower without template (Møllenbach, Hansen & Lillholm, 2013). Gestures can be compositions of multiple saccades and fixations, or in the simplest cases, single movements, called strokes (Møllenbach, Hansen, Lillholm, & Gale, 2009; Heikkilä & Räihä, 2012). Single stroke gestures naturally limit the number of commands that can be performed. However, performing them takes less time compared to dwell or blink based interaction.

Problems in Gaze Interaction

Despite its advantages, gaze interaction in its classical form has so far suffered from a number of major drawbacks. Among these are a high number of false alarms, the need to calibrate the eye tracker to the individual user to achieve high spatial accuracy and low user acceptance.

Midas Touch Problem

In the literature, the high number of false alarms is known as *Midas touch problem* and refers to the legendary king Midas, who wished that everything he touched would be turned into gold, but then finds himself trapped in his wish (Jacob, 1991). Similarly, users of dwell based gaze interaction often report that everything that attracts their interest, and is thus fixated with their eyes, is quickly selected without the option of exploring it sufficiently to make a conscious decision whether or not to select an item. The problem is most notable in connection with dwell based selection, where every item that is visually scrutinized for a longer period, is selected. The problem is less acute in blink based selection, as this mechanism is less directly linked to visual exploration and the execution relies on a voluntarily controlled action. Though gaze gestures do not entirely eliminate the high number of false activations associated with dwell based interaction, the number of false alarms is reduced by multiple stroke gestures that would not be performed in natural looking.

Individual calibration

A high need for accuracy is most common in dwell and blink based interaction. Both concepts rely on the identification of the exact gaze position to evaluate if the gaze dwells within a margin that is predefined to select a specific button. It is thus indispensable to calibrate the eye tracker individually to the user. However, calibrations suffer from low user acceptance. Particularly in cases where a recalibration is needed, users often report strain resulting from the calibration (Villanueva, Cabeza & Porta, 2004; Pfeuffer, Vidal, Turner, Bulling & Gellersen, 2013). Furthermore, the time needed for the calibration makes it impractical in settings where a quick interaction is required. Gaze gestures could possibly solve this problem: As gaze gestures do not aim at selecting a button, but at calling a specific command, e.g. closing a browser, spatial accuracy is less of a problem: It has even been suggested that they can be performed without calibrating the system to the individual user (Drewes, Hußmann & Schmidt, 2007).

Low user acceptance

Gaze interaction has been described as intuitive and natural (Jacob, 1991; Majaranta & Räihä, 2002; Sibert & Jacob, 2000). However, in our lab we often observed that gaze interaction suffers from low user acceptance. Our

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

idea is that the eye is naturally used for the exploration of the environment, but not for the manipulation of objects. Particularly when employing blink based interaction concepts or gaze gestures, users perform voluntary commands with their eyes, which are naturally associated to manual control rather than gaze. In this regard Jacob (1993) and Nielsen (1993) emphasized the potential of "non-command" interfaces at an early stage, which enable a selection of objects without consciously manipulating them with the eyes.

To develop a genuinely natural form of gaze interaction that enjoys high user acceptance, it is thus important to identify an interaction concept that does not require the performance of a conscious command. At the same time it should not trigger the *Midas touch problem* as dwell based selection does. To achieve optimal results, the concept should not depend on high accuracy, i.e. it should require no individual calibration.

Smooth Pursuit Eye Movements in Gaze Interaction

We assume that the aforementioned requirements for natural and reliable gaze interaction are met by making use of a specific type of eye movement: the smooth pursuit. Smooth pursuit eye movements are relatively slow and regular ("smooth") movements of the eye that occur when a moving object is followed by the gaze (Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka & Van de Weijer, 2011). The implementation of smooth-pursuit eye movements in gaze interaction requires an innovative display on which selectable objects move around. The approach allows for longer exploration times than those employed in dwell based concepts. At the same time no learning phase is required, as it is needed for gaze gestures. The object of the interaction is visible at all times but does not need to be manipulated. As the concept is based on the identification of the gaze relative to the movement of the displayed objects rather than the absolute gaze position, high accuracy is not required and an individual calibration is not needed.

Gaze based text-entry systems like Dasher (Ward & MacKay, 2002) and StarGazer (Hansen & Hansen, 2006) use moving display elements to guide attention. However, the text input detection is not explicitly based on smooth pursuit eye movements. To the authors knowledge Vidal and colleagues were the first that have shown that it is possible to identify smooth pursuit eye movements and match them to the course of a moving object (Vidal, Bulling & Gellerson, 2013a; Vidal, Pfeuffer, Bulling & Gellersen, 2013b). To do so they correlated the movement of the eye with the path of different objects on the screen. In a laboratory experiment Vidal et al. achieved 89 % correct identifications. Classification of a gaze path occurred after a mean correlation time of 1.88 seconds. These are promising results that lead us to the conclusion that smooth pursuit eye movements are a suitable option for gaze interaction.

The aim of the two experiments that are presented in this paper is the development of a robust and user-friendly form of gaze interaction based on smooth pursuit eye movements. To achieve high robustness, i.e. an interaction that neither suffers from high numbers of involuntary activations nor from failures to identify a movement, we suspect that it may be necessary to develop a classification that is bespoke to the specific implementation of the graphical user interface. Based on this assumption an algorithm that is uniquely adapted to the employed graphical user interface will be designed.

As an exemplary application the input on a PIN pad is chosen. The process of entering ones PIN by gaze increases security because the risk of "shoulder surfing", which is the observation of the PIN number by a third person when it is entered, is reduced (Kumar, Garfinkel, Boneh & Winograd, 2007; De Luca et al., 2007; Bulling, Alt & Schmidt, 2012). We aim at demonstrating that a robust implementation is possible even without the performance of an individual calibration. These improvements should lead to high user acceptance.

Experiment 1

The aim of our first study was to learn about smooth pursuit characteristics on moving display targets to extract a suitable algorithm that matches the eye movements to the trajectory of the moving numbers. Based on this we aimed at identifying an easy and comfortable to follow target movement for the gaze-based PIN entry. For this purpose no gaze interaction was implemented at this stage. Instead we inspected the gaze data of participants to derive a suitable algorithm. Moreover, the first experiment aimed at exploring whether an interaction without an individual calibration of the eye tracking system is generally possible.

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

Materials and apparatus

We used the SMI iViewRED250 eye tracker sampling at 60 Hz, which was attached to a 20" monitor with a display resolution of 1680*1050 px.

The moving PIN pad was implemented in Microsoft PowerPoint 2010 using animated slides. The target buttons could only move in vertical and horizontal direction. Each button movement was divided into three time segments. For each time segment the button moved either upwards, downwards, to the left or to the right. Each number had a unique movement consisting of a combination of three single movements. To enter number "1", the associated display object e.g. moved upwards, to the left and then downwards. The movement was constant in speed. Figure 1 visualizes the PIN interface with exemplary movements of three numbers. All 16 display elements were static for a short moment so that the user could fixate the number of interest before the movement of all objects started simultaneously. To enter a four-digit PIN code the procedure had to be iterated three times.
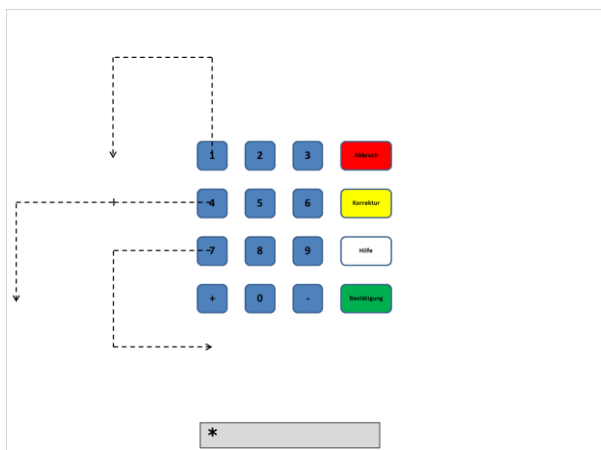

Figure 1. PIN interface with exemplary button movements (dashed lines).

Experimental Design

To identify an easy and comfortable implementation of smooth pursuit movements, the speed (within-subject factor 1) and the density of the moving targets on the screen (within-subject factor 2) were varied. Three different levels of speed (436 (fast); 218 (medium) and 145 (slow) px/s) and two variations of minimal object distance throughout the movement (4 (small) and 39 px (large)) were tested, resulting in six different variations of

moving objects on the display. Each participant entered one four-digit PIN code with each variation.

Note that with minimal object distance the length of one single button movement varied accordingly. In the conditions of larger button distances the length was 218 px and for the smaller button distances 137 px. Table 1 shows the single button movement characteristics of each variation.

Table 1. Resulting single button movement characteristics.

Minimal button distance *(within-subject factor 2)*	Speed *(within-subject factor 1)*		
	Fast (436 px/s)	**Medium** (218 px/s)	**Slow** (145 px/s)
Large (39 px)	218 px in 0.5 s	218 px in 1.0 s	218 px in 1.5 s
Small (4 px)	137 px in 0.3 s	137 px in 0.6 s	137 px in 0.9 s

User-friendliness of the implementations was measured by asking the participants to rate three criteria: the ease of performing the pursuit movement (*ease*), the strenuousness for the eyes (*strenuousness*) and the difficulty to maintain the fixation on the target object without being distracted by the other object movements (*distraction*). The participants rated the three criteria by setting a mark on a continuous 10 cm scale with a zero-point in the middle of the scale (-5 cm = very negative, to +5 cm = very positive).

Task and Procedure

Upon arrival, participants were seated at a distance of approximately 70 cm to the eye tracker and the screen (Figure 2).
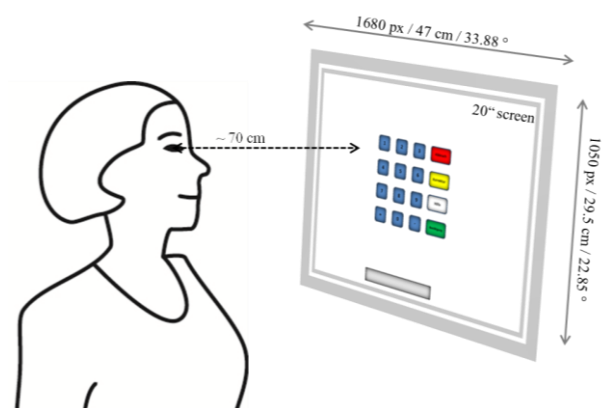

Figure 2. Setup of experiment 1.

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

Participants were made believe that by following a number with their gaze they indeed made a selection. In fact the system ran autonomously, assuming that the participants followed the number that was indicated. The procedure was trained for one four-digit PIN code. After that the system was calibrated using a five-point calibration. For twelve participants the eye tracker was calibrated in the conventional way (*individual calibration*), for the remaining six participants gaze data was collected with the system being calibrated on the investigator's eyes (*external calibration*).

After calibration and training, participants were asked to enter six different PIN codes by pursuing each number successively. The PIN codes that should be entered were presented before the start of the trial. During the trial the code was visible in the left upper corner of the screen to exclude incorrect gaze behavior because the participants did not remember the code correctly. After each PIN the participant was asked to rate three criteria of user-friendliness: ease, strenuousness and distraction. In total, the experiment lasted approximately 20 minutes.

Participants

Eighteen participants, nine males and nine females, completed the experiment. Their ages ranged from 22 to 45 years (mean 27.3 years). None of the participants wore glasses during the experiment. Almost all of the participants (17 participants) had practical experience with eye tracking; twelve were experienced in gaze interaction. For participating they were rewarded with either five Euros or a certification of student experimental hours.

Results

As displayed in Figure 3, the mean rating for the medium speed condition and the large distances between the moving targets was comparatively high regarding the three criteria ease, strenuousness and distraction. Ease was rated with an average of $M = 3.81$ ($SD = 0.93$), strenuousness with a mean of $M = 3.18$ ($SD = 1.86$) and distraction with an average of $M = 3.49$ ($SD = 1.21$).
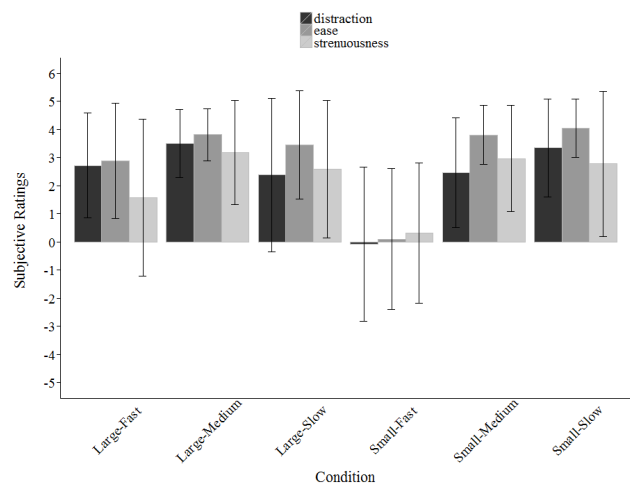

Figure 3. Mean subjective rating results of user friendliness (-5 to +5 rating scale, error bars indicate standard deviations).

The most favorable results with regard to user friendliness were obtained in the medium speed/large distance condition. As we assume that a robust classification can be achieved most easily when based on a specific implementation, we proceeded with the development of a suitable gaze interaction algorithm taking only the data of the medium speed/large distance condition into account.

Algorithmic Classification

In order to set up an interactive system we aimed to develop an easy to implement and robust algorithm to map the observed eye gaze behavior to the movements of the PIN pad buttons. For the development of such an interaction we used the medium speed/large distance eye-tracking raw data collected during the first experiment. Each data sample provides information about the gaze position on the display (Point Of Regard - POR) and is defined by x/y coordinates. As we sampled at 60 Hz and a single movement in one direction took one second, a single pursuit is represented by approximately 60 data points. An exemplary gaze track consisting of three single pursuits is shown below in Figure 4.
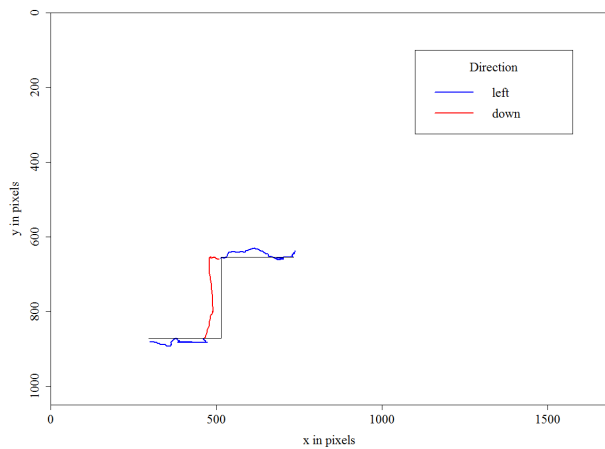
5

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

Figure 4. Resulting gaze track (colored line) of a pursued button movement (black line) on the display.

After visual scrutiny of the gaze data we decided for a simple arithmetic two-stage classification algorithm. Stage one identifies the direction of a single eye movement made while following a single button movement. Stage two detects the entered number by combining three single eye movements and comparing it with the movement patterns of the numbers.

Stage 1: To classify the direction of a single eye movement (left, right, down, up) we used relative measures to compensate for possible imprecision of the eye tracker or a less accurate calibration (e.g. external calibration). For each button movement we calculated the differences between the first and the last POR for the x- and y-coordinates, respectively. When no movement occurred in one direction the difference is expected to be around 0 px, whereas the difference in case of a movement is presumed to be approximately 218 px. To account for individual variability a range was defined. This range was set to comprise 95 percent of all upward movements that occurred in Experiment 1. Upward movements were chosen as a reference, since they are reported to be most variable in humans (Holmqvist et al., 2011). The resulting ranges are displayed in Table 2.

Table 2. Ranges of differences between first and last POR (x- and y-PORs).

Direction	Range of movement distance in pixels	
	Xend - Xstart	Yend - Ystart
down	-80 to 80 (0)	140 to 300 (218)
up	-80 to 80 (0)	-140 to -300 (-218)
right	140 to 300 (218)	-80 to 80 (0)
left	-140 to -300 (-218)	-80 to 80 (0)

In case that the eye movement data fell outside the defined ranges, we implemented a second step. Since the eyes should normally follow the single button movement they are expected to only move into one direction. Therefore the majority of differences between consecutive PORs should have the same orientation on the x- or y-axis. For a downward movement for example the differences between the x-coordinates of consecutive data points are expected to be equally balanced around zero. For the y-coordinates, however, the majority of differences between two samples should be considerably greater than zero. Based on this we calculated the cumulative number of positive and negative differences between successive PORs within a single eye movement. Depending on the distribution of positive and negative x- and y- POR differences a decision for a direction was taken. This two-stepped procedure led to a correct classification of 99.54 % of the single movements (Table 3).

Table 3. Percentage of single movements correctly classified.

Classification step	Calibration type		
	Individual	External	Both
Step 1	**95.14 %** (137/144)	**94.44 %** (68/72)	**94.91 %** (205/216)
Step 2	**85.71 %** (6/7)	**100 %** (4/4)	**90.91 %** (10/11)
Total	**99.31 %** (143/144)	**100 %** (72/72)	**99.54 %** (215/216)

Note. In brackets: Total number of correctly classified single eye movements/ all single eye movements.

Stage 2: In a second stage we combined the detected single movements obtained from the first stage of the algorithm to identify the entered number. Each complete movement of a number on the PIN pad is a distinct combination of three movements. Overall, the algorithm

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

correctly detected 98.61 % of the entered numbers. The remaining 1.39 % of eye movements could not be related to a movement of a number on the PIN pad. It should, however, be noted that the visual inspection of these patterns could not link these movements to a specific pattern either. This implies that in these cases no number had been pursued by the participant. In summary, it can be stated that the described algorithm identified nearly all numbers entered, independent of whether the individual or the external calibration was used.

Experiment 2

Our second experiment aimed to validate the developed algorithm in a truly interactive setup. A further aim was to learn about the user experience of this way of interacting.

Materials and Apparatus

The set-up of experiment two is very similar to the first experiment. This time an actual interaction was possible, but hardware and interface design were not changed. To realize the interaction a mouse emulator software (Mousey [1]) was used as an interconnection between gaze and mouse. The developed algorithm was implemented as Visual Basic for Applications 2010 scripts in the presentation. Gaze input was derived from the emulated mouse position which was sampled with a frequency of 30 Hz. For reasons of simplicity the functionality of the PIN pad was limited to the recognition of the ten numbers (0-9) only. Buttons like "clear" and "cancel" were visible on the screen, but could not be activated.

Experimental Design

To validate the algorithm all participants were asked to enter three four-digit PIN codes. Twelve of the participants were conventionally calibrated (*individual calibration*) and twelve interacted with the system being calibrated on the investigator's eyes (*external calibration*).

---

[1] A multi-purpose eye-tracking and gaze-interaction interface developed at the Chair of Human-Machine Systems at Technische Universität Berlin (http://www.mms.tu-berlin.de/fileadmin/fg268/Mitarbeiter/Mousey2_Documentation.pdf)

Task and Procedure

The procedure is similar to the one in the first experiment. Before the experiment started the procedure was explained and participants were given the opportunity to practice. Additionally participants were informed that whenever an error message occurred they should repeat the input. A five-point calibration was used. After the calibration phase the participant entered three different four-digit PIN codes that were shown to them before each trial. Afterwards they filled out the *User Experience Questionnaire* (UEQ) by Laugwitz, Held and Schrepp (2008) featuring six scales (attractiveness, perspicuity, efficiency, dependability, stimulation and originality) and using a seven staged semantic differential format (-3 to +3). The UEQ covers a comprehensive impression of user experience. *Attractiveness* measures the general attitude towards the product; *efficiency*, *perspicuity*, *dependability* measure classical usability aspects and *originality* and *stimulation* are related to user experience aspects. In total, each experiment lasted approximately 20 minutes.

Participants

Twenty-four participants, twelve males and twelve females, completed the experiment. Their ages ranged from 23 to 54 years (mean 30.3 years). None of the participants wore glasses during the experiment. Fifteen participants had practical experience with eye tracking; seven were experienced in gaze interaction. Participants' reward was equal to the first experiment.

Results

Classification Rate. With the algorithm applied, a total of 99.07 % of all single eye movements and accordingly 97.57 % of all numbers were correctly identified on first input (Figure 5, Table 4). The remaining 2.43 % of numbers needed repeated input. In total, no false number or PIN was selected. The first classification step identified 88.89 % of the single movements correctly. No single movement was classified incorrectly in this first step. 11.11 % of the single movements were not identified and therefore referred to the second step, which classified 91.67 % of the remaining strokes correctly.
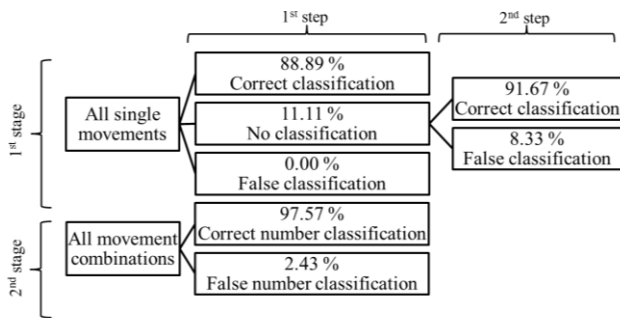
Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

Figure 5. Stages and steps of the classification.

Table 4. Single eye movements correctly classified in stage 1.

Classification step	Calibration type		
	Individual	External	Both
Step 1	**94.21 %**	**83.56 %**	**88.89 %**
	(407/432)	(361/432)	(768/864)
Step 2	**92 %**	**91.55 %**	**91.67 %**
	(23/25)	(65/71)	(88/96)
Total	**99.54 %**	**98.61 %**	**99.07 %**
	(430/432)	(426/432)	(856/864)

Note. In brackets: Total number of correctly classified single eye movements/ all single eye movements.

Questionnaire results. The result for *efficiency* is within the neutral range (-0.8 to 0.8), the remaining dimensions were all rated positive (> 0.8 to 3). Figure 6 contains the detailed results. There were no group differences between individually and externally calibrated participants for any dimension.
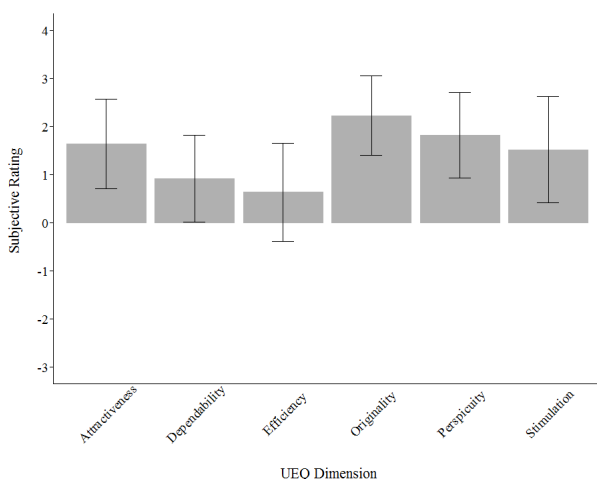

Figure 6. Mean UEQ results (-3 to 3 rating scale, error bars indicate standard deviations).

Discussion

The aim of this study was to develop a robust form of gaze interaction based on smooth pursuit eye movements on moving display buttons by utilizing a gaze path classification. The classification is uniquely adapted to the specific graphical PIN pad interface to achieve a higher rate of correct identifications.

Results of our experiments could demonstrate the robustness of the approach. There was no case where a number had been entered falsely. With no involuntary activation of buttons throughout the second experiment and a correct identification of 97.57 % of the entered numbers, the interaction can be termed highly reliable and a good solution to the *Midas touch problem*. The classification rate in the first experiment was 98.61 %. This marginal difference between the two experiments may be due to the lower sampling rate in the second experiment. The absence of false alarms may in part be due to the low number of paths that are implemented. However, the overall low error rate suggests that this is not the only reason. One reason for the low false alarm rate seems to be the prolonged interval for selection. As this prolongation is not achieved within one fixation, it does not strain the eye.

Furthermore the omission of an individual calibration was tested. In total 98.61 % of the movements were correctly identified in the externally calibrated group. These results are only slightly inferior to those of the individually calibrated (99.54 %). Thus a spontaneous interaction without calibrating the system to the individual user is clearly possible and a standardized calibration appears feasible. It is surprising that the first step of the first stage of the algorithm is activated less often in the externally calibrated group. This suggests that the increased imprecision of the external calibration may more often result in single eye movements that surpass the ranges set for this first algorithm step. However, by applying the second step an overall high classification rate can be sustained.

It should be noted that for technical reasons the eye tracker was calibrated to the experimenter in the 'external calibration' condition. This still represents a calibration though not to the user itself. However, modern eye trackers often feature a 'standard calibration' that is based on average facial, respectively eye features. This template can be invoked instead of the usual calibration procedure.

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

The use of such a system would eliminate the need to calibrate the system to a third person. Gaze interaction systems that do not require any calibration have already been realized by Shell, Selker & Vertegaal (2003) Zhang, Bulling & Gellersen (2013) and Vidal et al. (2013a, 2013b).

In the experiments, we made use of a high-end eye tracking system with good tracking accuracy. However, as the algorithm uses relational movements rather than the exact eye positions on the monitor, it seems likely that the application would work with less precise eye trackers as well.

Another aim was to learn about the user experience. Participants rated the way of interacting positive. Only the efficiency was rated neutral, which is not surprising since the entry of a PIN by gaze takes 25 seconds provided that no number is entered wrongly. A faster input could be possible by using button movements composed of two instead of three strokes. In this case the risk of faulty insertions would rise due to the fact that almost all possible stroke combinations are allocated to a button (16 (4^2) possible combinations). Another solution would be a faster button movement. However, it should be considered that the faster speed level was clearly evaluated inferior in the first experiment. On the contrary, the desire for a faster movement may arise with higher interaction experience. The overall good assessment of the user experience can probably be traced back to multiple factors – such as the novelty of the approach, the low level of false activations and the user-centered extraction of minimally distracting button movement in the course of the first experiment. It is unclear if the fact that most participants had some experience with gaze interaction contributed to these results. It is possible that users internally compared the smooth pursuit-based action to other forms of gaze interaction they know. It would therefore be of interest to test the system again with users who have never used gaze interaction before, or to directly compare smooth pursuit based interaction to other gaze based input methods.

The algorithm uniquely adapted to the PIN pad interface performed well but is for the same reason limited to the selection of horizontally and vertically moving display objects. Non-linear and diagonal smooth pursuit movements cannot be identified. In addition the classification only works under the precondition that the users follow the instructions. If a user does not follow a button movement, an error message or a false selection will be the result. The implementation is thus a "best guess" rather than an identification of the actual eye movement.

Up to now, moving display buttons are rarely used for interaction. Consideration is needed on how to implement and design dynamic interfaces for gaze interaction. For instance the number of moving objects on the screen is limited in this approach a) to avoid clutter and b) because there would be a substantial overlap between the path ways if considerably more buttons were to move. The design should also allow the user's eye to find and fixate on an inactive button before the movements start because finding a target button within several moving ones is difficult and potentially leads to a faulty selection. This could also be facilitated by the use of familiar and clearly arranged interfaces.

The obtained results are of course preliminary, and the algorithm needs further validation with different participant populations. All our participants were rather young and can therefore be assumed to have high levels of experience in the interaction with technology. Additionally, we purposefully excluded participants wearing glasses, as we wanted to validate our algorithm rather than the robustness of the eye tracking hard- and software. This demonstrates that the implementation of smooth pursuit based gaze interaction in real world applications is not yet within reach. Nevertheless, we believe that our results are encouraging and justify the further investigation of smooth pursuit based gaze interaction. Use cases of this form of interaction are not limited to PIN entry and could be expanded to password entry or general typing. In case of entering a PIN the risk of shoulder surfing can be diminished.

Conclusion

Classifying smooth pursuit eye movements for gaze interaction proved to be a robust approach. The implementation not only led to a high percentage of correctly identified entries and eliminated the typically high false alarm rate often associated with gaze interaction, but also allows for the omission of an individual calibration. At the same time gaze interaction based on pursuing the movement is accompanied by high user acceptance ratings.

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

References

Bulling, A., Alt, F., & Schmidt, A. (2012). Increasing the Security of Gaze-Based Cued-Recall Graphical Passwords Using Saliency Masks. *Proc. of the 30th SIGCHI International Conference on Human Factors in Computing Systems (CHI 2012)* (3011-3020). Austin, Texas, USA: ACM.

De Luca, A., Weiss, R., & Drewes, H. (2007). Evaluation of Eye-Gaze Interaction Methods for Security Enhanced PIN-Entry. *Proceedings of the 19th australasian conference on computer-human interaction: entertaining user interfaces* (199-202). Adelaide, Australia: ACM.

Drewes, H., & Schmidt, A. (2007). Interacting with the Computer using Gaze Gestures. *INTERACT 2007, 11th IFIP TC 13 Int. Conference, Part II* (475-488). Springer.

Drewes, H., Hußmann, H., & Schmidt, A. (2007). Blickgesten als Fernbedienung. *Mensch und Computer 2007: Interaktion im Plural*, 79-88.

Friedman, M. B., Kiliany, G., Dzmura, M. & Anderson, D. (1982). The eyetracker communication system. *Johns Hopkins APL Technical Digest 3(3)*: 250-252.

Heikkilä, H. & Räihä, K.-J. (2009). Speed and accuracy of gaze gestures. *Journal of Eye Movement Research, 3(2)*, 1-14.

Hansen, D. W., & Hansen, J. P. Eye typing with common cameras. In K.-J. Räihä & A. T. Duchowski (Eds.), *Proceedings of the 2006 symposium on Eye Tracking Research & Applications (ETRA '06) (55-55)*, New York, NY: ACM.

Heikkilä, H., & Räihä, K.-J. (2012). Simple gaze gestures and the closure of the eyes as an interaction technique. *Proceedings of the symposium on eye tracking research and applications* (147-154). New York: ACM.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye Tracking - A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Jacob, R. J. (April 1991). The use of eye movements in human-computer interaction techniques: What You Look At is What You Get. *ACM Transaction on Information Systems, 9(3)*, 152-169.

Jacob, R. J. (1993). Eye movement-based human-computer interaction techniques: Toward non-command interfaces. In *Advances in Human-Computer Interaction*, *4*, 151-190.

Koesling, H., Zöllner, M., Sichelschmidt, L. & Ritter, H. (2009). With a flicker of the eye: Assessing gaze-controlled human-computer interaction. *Cognitive Systems Monographs, Human Centered Robot Systems: Cognition, Interaction, Technology* (83-92). Berlin: Springer.

Kumar, M., Garfinkel, T., Boneh, D., & Winograd, T. (2007). Reducing shoulder-surfing by using gaze-based password entry. *Proceedings of the 3rd symposium on usable privacy and security (SOUPS '07)*.

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 (ETRA '02)* (15-22). New Orleans, Louisiana: ACM.

Majaranta, P., MacKenzie, I., 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.

Møllenbach, E., Hansen, J. P., Lillholm, M., & Gale, A.G. (2009). Single stroke gaze gestures. In *Proceedings of the 27th international Conference Extended Abstracts on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, 4555–4560.

Møllenbach, E., Hansen, J. P., Lillholm, & Martin. (2013). Eye Movements in Gaze Interaction. *Journal of Eye Movement Research, 6(2)*, 1-15.

Nielsen, J. (1993). Noncommand user interfaces. *Comm. ACM 36*, 4 (April 1993), 83-99.

Pfeuffer, K., Vidal, M., Turner, J., Bulling, A., & Gellersen, H. (in press). Pursuit Calibration: Making Gaze Calibration Less Tedious and More Flexible. *Proceedings of UIST 2013*.

Shell, J.S., Selker, T., & Vertegaal, R. (2013). Interacting with groups of computers. *Communications of the ACM. 46(3)*, 40-46.

Špakov, O. & Miniotas, D. (2004). On-line adjustment of dwell time for target selection by gaze. *Proceedings of the third Nordic Conference on Human-computer interaction (NordiCHI)* (203-206). New York: ACM.

Journal of Eye Movement Research
7(4):1, 1-11

Cymek, D.H., Venjakob, A.C., Ruff, S., Lutz, O. H., Hofmann, S. & Rötting, M. (2014)
Entering PIN codes by smooth pursuit eye movements

Sibert, L. E., & Jacob, R. J. (2000). Evaluation of eye gaze interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI' 00)* (281-288). New York: ACM.

Vidal, M., Bulling, A., & Gellersen, H. (2013a). Pursuits: Spontaneous Interaction with Displays based on Smooth Pursuit Eye Movement and Moving Targets. *Proceedings of UbiComp' 13*. Zürich.

Vidal, M., Pfeuffer, K., Bulling, A., & Gellersen, H. (2013b). Pursuits: Eye-Based Interaction with Moving Targets. *CHI 2013 Extended Abstracts* (3147-3150). Paris: ACM.

Villanueva, A., Cabeza, R., & Porta, S. (2004). Eye tracking system model with easy calibration. *Proceedings of the 2004 symposium on Eye tracking research & applications (ETRA '04)* (55-55). New York: ACM.

Ward, D. J., & MacKay, David J C. (2002). Artificial intelligence: fast hands-free writing by gaze direction. *Nature*, *418(6900)*, 838. doi:10.1038/418838a