TextPursuits: Using Text for Pursuits-Based Interaction and Calibration on Public Displays

Mohamed Khamis[1], **Ozan Saltuk**[1], **Alina Hang**[2], **Katharina Stolz**[1], **Andreas Bulling**[3], **Florian Alt**[1]
[1] LMU Munich, Germany
[2] Sonix Group, Munich, Germany
[3] Max Planck Institute for Informatics, Saarbrücken, Germany
{mohamed.khamis, florian.alt}@ifi.lmu.de, bulling@mpi-inf.mpg.de

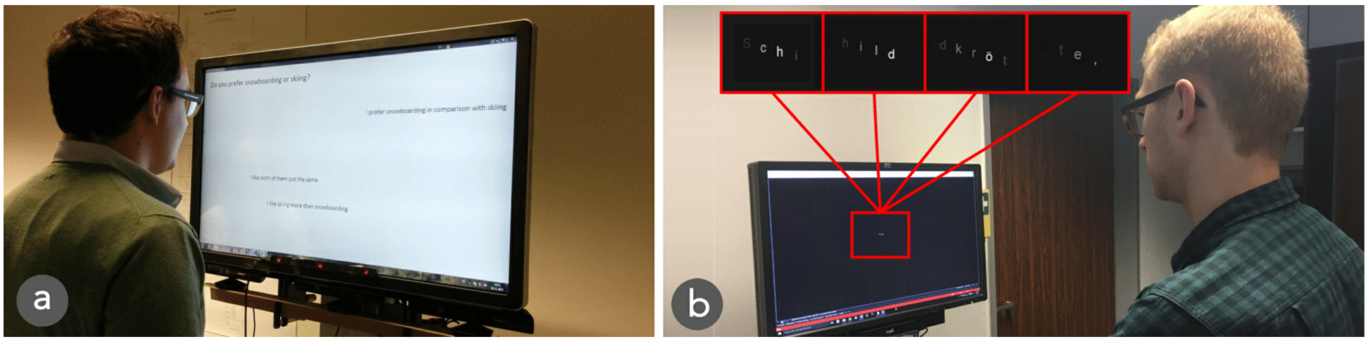


Figure 1. We explore the use of text-based stimuli to enable gaze interaction with public displays using the Pursuits technique [47]. Motivated by the fact that much of the content on large displays is text, we investigate two use cases: (a) Users can spontaneously *interact* with text-based content without calibration. A sample application could be a survey where answers in the form of text are selected by reading them (left). (b) An eye tracker can be *calibrated* implicitly as users read text on the screen (right). After calibration, fine-grained information on the user's gaze point are obtained.

ABSTRACT
In this paper we show how reading text on large display can be used to enable gaze interaction in public space. Our research is motivated by the fact that much of the content on public displays includes text. Hence, researchers and practitioners could greatly benefit from users being able to spontaneously *interact* as well as to implicitly *calibrate* an eye tracker while simply reading this text. In particular, we adapt Pursuits, a technique that correlates users' eye movements with moving on-screen targets. While prior work used abstract objects or dots as targets, we explore the use of Pursuits with text (read-and-pursue). Thereby we address the challenge that eye movements performed while reading interfere with the pursuit movements. Results from two user studies (N=37) show that Pursuits with text is feasible and can achieve similar accuracy as non text-based pursuit approaches. While calibration is less accurate, it integrates smoothly with reading and allows areas of the display the user is looking at to be identified.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User Interfaces — Input Devices and Strategies

Ubicomp'16, September 12-16, 2016, Heidelberg, Germany

DOI: http://dx.doi.org/10.1145/2971648.2971679

Author Keywords
Public Displays; Smooth Pursuit; Gaze Interaction; Text

INTRODUCTION
As they are becoming ubiquitous and cheap to deploy, displays can be found in public spaces such as airports [4], shopping centers [9] and train stations [8]. At the same time, sensing technologies are becoming increasingly available for easy and low cost integration with public displays, supporting different ways of interaction. Common interaction modalities for displays include touch [10], smart phone interaction [3, 12], mid-air gestures [34], and recently also gaze [20, 48, 55].

Gaze holds particular promise for public displays [22]. It is intuitive [46], natural to use [47], indicates visual attention, and usually precedes action [31]. However, a drawback is that eye trackers require calibration, which is time-consuming and cumbersome [31]. While devoting time for calibration is acceptable for desktop settings, public displays require *immediate usability* [33] as interaction times are usually short [34]. Hence, calibration has been identified as one of core challenges of gaze-enabled public displays [21]. Prior work investigated alternative techniques [48, 55].

A popular approach is Pursuits [47, 49], which relies on correlating movements of dynamic objects on the display with the smooth pursuit eye movement performed when following a moving object. Pursuits was successfully deployed for multiple public display installations, where it was used for both gaze interaction [20, 48] and eye tracker calibration [6, 36].

Meanwhile, one of the most prominent content types on public displays is text. For example, displays are utilized for opinion gathering and sharing in public areas [17, 24]. In many applications passersby read and select from a set of text-based options [15, 16, 35, 42, 44, 50]. And also (pervasive) advertising on public displays often heavily relies on text [1].

Nevertheless, little is known about whether and how Pursuits can be used with text. To date, Pursuits has been studied with moving dot-like stimuli, for which the user gazes at a single, spatially clearly defined target. On the other hand, the use of Pursuits with textual stimuli is not straight forward: reading is not spatially confined and overlays the smooth pursuit movement, which could result in difficulty in correlating eye movements and the trajectory of text-based stimuli. Also, due to the Midas effect, gaze-based systems need to distinguish users reading textual content from interacting with it.

We investigate the use of text as stimulus for Pursuits. We see two main use cases for public displays: (1) It can be used for calibration-free gaze *interaction* [20, 48]; by displaying moving objects, users can "pursue" the object they want to select. The system then determines which object the user is looking at by correlating eye movements and movements of the objects. (2) Pursuits can be used for easier and less tedious *calibration* of eye trackers [6, 36]; by following a moving stimulus, mappings between movements of the user's eyes and the stimulus can be collected and used for calibration.

To provide a proof-of-concept we implemented two applications: EyeVote is a survey application for public displays that enables a user to select an answer from a set of text-based options. Read2Calibrate shows animated text on the screen and feeds the gaze data to an algorithm that gradually enhances the calibration of the gaze tracking. We used both systems in two lab studies with the goal to assess accuracy based on different configurations (text inclination, text speed, text length, and trajectory) and to obtain early insights on the users' view. The results show that text-based stimuli can be used for Pursuits-based gaze *interaction* but that designers need to be careful about the text trajectories and text length in order to minimize detection errors. Text-based *calibration* is in general less precise than state-of-the-art calibration procedures. However, the accuracy is sufficient to identify objects a user is looking at on the screen. We found text inclination to have a strong influence on the calibration quality. Both for interaction and calibration designers may need to make a tradeoff between the configuration leading to the most accurate results and the users' preferred configuration.

The contribution of this work is threefold: (1) We present a prototype application, EyeVote, that allows text to be selected using Pursuits and report on a user study with 19 participants, assessing the accuracy and error rate based on different configurations (text length, trajectory). (2) We introduce Read2-Calibrate, a system for calibrating eye trackers for displays by utilizing smooth pursuit eye movements and a text-based stimulus. Again, we evaluated the system with regard to accuracy based on different configurations (text inclination, text speed). (3) We derive a set of guidelines and recommendations for using text with Pursuits for both interaction and calibration.

RELATED WORK

We build on two main strands of previous work: gaze interaction with public displays and interaction with text via gaze.

Gaze-based Interaction with Public Displays

Due to the benefits of gaze for public displays, two research directions have emerged to counter the problems associated with calibration on public displays: (1) enabling calibration-free gaze interaction for displays, and (2) making gaze calibration on public displays less tedious.

Calibration-free Gaze Interaction with Public Displays

Acknowledging the unsuitability of classic calibration for public displays, multiple systems were built to provide calibration-free gaze-based interaction. SideWays [53] and GazeHorizon [55, 56] use the pupil-canthi-ratio [54] to estimate horizontal gaze direction without calibration.

Pursuits [49, 47] can also be used for calibration-free gaze interaction. The technique requires displaying a dynamic interface [48], where "pursuitable" objects move. Eye movements are then correlated to the movements of the objects. The object whose movement correlates the most with that of the eyes is then assumed to be the one the user is looking at. Since its introduction, Pursuits has been used in a variety of applications including text entry [29], PIN code entry [11, 28] and entertainment applications [47, 49]. Pursuits has also been used for interaction with smart watches [13, 14] and interaction in smart environments [45]. The technique was shown to be intuitive and positively perceived by users [20].

Eye Tracker Calibration for Public Displays

Previous work aimed to reduce the effort needed for calibration. For example GazeProjector [27] allows gaze-based interaction across multiple displays using one time calibration and a mobile eye tracker. While mobile eye trackers have several advantages for interaction, public display users cannot be expected to wear them, unless trackers integrated with eye wear become commonplace. Hence remote eye trackers are currently more suited for that domain. Xiong et al. [52] used a remote RGB-D camera that requires one-time calibration.

Work by Pfeuffer et al. [36] uses the eye's smooth pursuit movement to facilitate calibration. The approach relies on showing a moving object, which acts as a stimulus for the eyes to perform the smooth pursuit movement. Mappings between eye movements and positions of the stimulus are then collected and used to calibrate the eye tracker.

Pfeuffer et al. used a floating "Please wait" label to calibrate eye trackers. Rather than reading a label and keeping fixating it, our approach for calibration relies on gradually revealing text, which intrigues the user to fixate at the gradually revealed letters to understand the statement. Moreover, our work on interaction with text via Pursuits investigates a different aspect, namely we study how users can select from a set of text-based options using Pursuits.

Interacting with Text via Gaze

In gaze-based systems, the "Midas touch" effect [18] occurs when the system mistakes a user perceiving content for selecting content. This effect is amplified in the case of text as reading requires time to perceive and understand the text. This challenge has been traditionally addressed by using dwell times – the system would require fixating the action element for a longer period of time (e.g. 1000 ms [30]).

Another approach to overcome the Midas touch is to use another modality in addition to gaze. Users of EyePoint [25] gaze at text, press and hold a keyboard button to magnify the area, refine the selection, and then release the button to select text. Stellmach et al. [40, 41] employed a similar approach by combining gaze and touch input. Although this approach was not used for text selection in particular, it is deemed suitable for the task. Kishi and Hayashi [23] combined gaze with on-screen buttons to enable users to select text. Chatterjee et al. [7] developed a text editor where users can move a text cursor by using gaze and pinch gestures.

A third approach is to use gaze gestures. In work by Toyama et al. [43], text selection was done either by repeatedly gazing at the beginning and the end of the text to be translated, or by gazing gradually from the beginning till the end of the text.

Sharmin et al. [38] introduced an automatic scrolling technique that is based on the user's gaze while reading text. Text 2.0 [5] exploits gaze by, for example, revealing content based on the words the user is currently reading.

Pursuits has the potential to cope with the Midas touch effect. Reading overlays the smooth pursuit eye movement, making false selections while reading less likely. Moreover, the Pursuits algorithm requires setting a window size, which is a time frame after which the correlation is calculated. This gives users the chance to perceive and read the text.

INTERACTING WITH TEXT USING PURSUITS

The use of text for interaction via Pursuits has not been investigated in detail before. With our work we close this gap and support designers and developers when it comes to creating text-based content suitable for interaction using Pursuits. In particular, the following section introduces a prototype application that allowed us important aspects of using text for pursuit interaction to be investigated.

Concept and Implementation

We implemented a voting system called EyeVote, that uses Pursuits as its only input mechanism for selecting one of several floating textual answers (see Figures 1A and 2). Once the system detects a selection, a confirmation message is shown on the screen, telling the user which answer was recognized. The message is kept for some seconds, followed by the next question and its options.

In the following we describe our implementation of Pursuits, and the experimental variables that we used in the study.

Text Selection via Pursuits

Pursuits works by correlating eye movements with those of the selectable options. Prior work utilized the Pearson's product-moment coefficient to calculate the correlation. Based on pilot experiments and previous work [14, 20, 47, 49], we used the same correlation function with a threshold of 0.9 and a window size of 2000 ms. This means that every 2 seconds, the system computes Pearson's correlation. The floating answer whose movement correlates the most with the eye movement, is deemed to be the object the user is looking at, as long as the correlation is more than 90%.

To account for reading time and overcome the midas effect, the used window size value is higher than those used in other implementations (e.g. previous work used 500 ms [20, 47] and 1000 ms [14]).

Trajectories and Text Representations

We investigate how already established trajectory movements perform with respect to text selection. In particular, the following trajectories were used in our experiments.

1. Circular trajectory [13, 14, 20, 47] (Figure 2 top left).
2. Linear trajectory [20, 47, 49] (Figure 2 rop right).
3. Rectangular trajectory [36] (Figure 2 bottom left).
4. Zigzag trajectory [47] (Figure 2 bottom right).
5. Mixed trajectory (each object follows one of the above trajectories).

We supported different text representations for the answers:

1. Short answers (<25 characters).
2. Two-lined answers.
3. Long answers (25+ characters).

Evaluating Text Selection Using Pursuits

The main goal of this experiment was to understand the influence of different text characteristics on the accuracy of selection via Pursuits. In particular, we compared the effect of the aforementioned trajectory types and text lengths on detection errors. In addition, we assessed the effect of the different trajectory types on the perceived workload. To minimize any external influences, we conducted the study in the lab [2].

Design

The study was designed as a repeated measures experiment. Each participant performed five blocks with each block covering one of the five trajectory types. In every block, participants performed 4 selections × 3 text representations = 12 text selections using Pursuits. The order was counter-balanced across participants using a Latin-square.

The theme of the study was a voting application, where participants had to answer questions by selecting one of three possible floating text-based answers via Pursuits (see examples in Figure 2). In total, every participant answered 60 questions: 5 trajectory types × 3 text representations × 4 selections.

Apparatus

We deployed the EyeVote system on a 42-inch display (3810 × 2160 pixels) in our lab (see Figure 1A). The display was equipped with a Tobii EyeX Controller (30Hz). Participants stood at a distance of roughly 60 cm from the eye tracker.

Participants

We recruited 19 participants (10 females) between 20 and 60 years through mailing lists and social networks. Four of them had previous experience with eye trackers but with Pursuits. All participants had normal or corrected-to-normal vision.

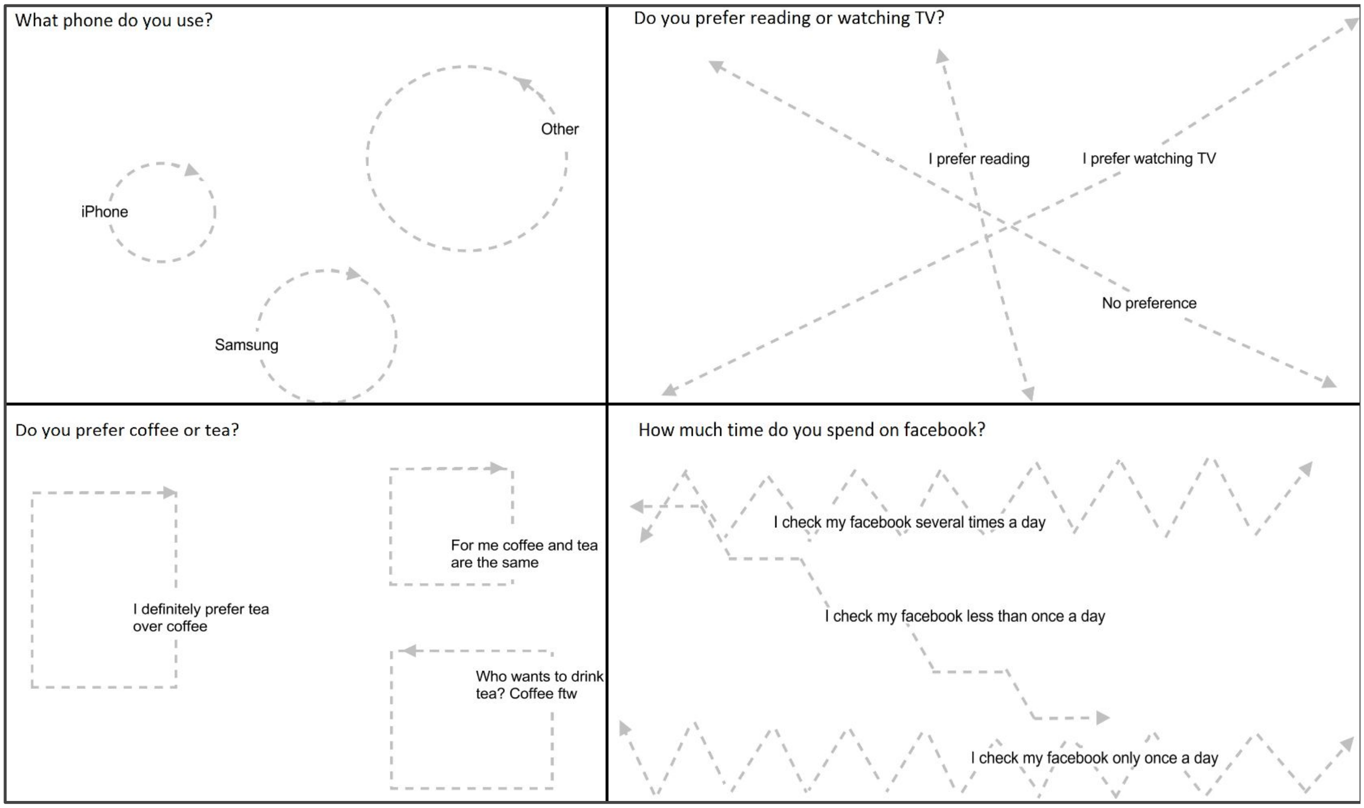


Figure 2. We manipulated two experimental variables: (1) Trajectory type: circular, linear, rectangular and zigzag trajectories shown above, in addition to a fifth condition "Mixed trajectory", where each answer followed a different trajectory (arrows are only for illustration and were not shown to participants). (2) Text representation: short answers (top), two-lined answers (bottom left) and long answers (bottom right).

Procedure

The experimenter began by explaining the study and asking the participant to sign a consent form. The experimenter then started the first block of 12 questions. After each successful Pursuit selection and before showing the following question, the system showed the user which answer was recognized. At that point the participant was asked to confirm whether or not the system detected the intended answer. In case of false detection, the participant was further asked to specify whether (a) the system detected a selection prematurely (i.e. the participant was still reading it) or (b) the participant was trying to select a different answer. To assess the perceived workload associated with text-based selection of every trajectory type, participants filled in a Nasa TLX questionnaire after each block.

Results

We logged the time taken to answer each question as well as the false detections by the system. In total, we recorded 1140 selections (19 participants × 60 selections).

We classify errors as (a) *early detection errors*, that are errors due to a premature recognition of an option while the participant is still reading, and (b) *false detection errors*, that are errors due to the system recognizing a selection other than the one the participant intended. Out of 1140 selections, there were 124 errors (≈10.9%): 88 of them were early detections (≈7.7%), while 36 were false detections (≈3.2%).

Significantly different pair		p-value
Circular (10.2%)	Linear (36.4%)	($p < 0.005$)
Circular (10.2%)	Zigzag (35.2%)	($p < 0.05$)
Rectangular (8.0%)	Linear (36.4%)	($p < 0.001$)
Rectangular (8.0%)	Zigzag (35.2%)	($p < 0.005$)
Mixed (10.2%)	Linear (36.4%)	($p < 0.001$)
Mixed (10.2%)	Zigzag (35.2%)	($p < 0.05$)

Table 1. Trajectory type has a significant main effect on early detection errors, which are cases where the system recognized a selection while the participant is still reading. The table above summarizes the results of post-hoc analyses using Bonferroni correction, which revealed significant differences between multiple pairs. The numbers between brackets denote the percentage of early detection errors caused by the corresponding trajectory type out of all 88 early detection errors. The results indicate that circular, mixed and rectangular trajectories result in fewer early detection errors compared to linear and zigzag trajectories.

We classify errors as (a) *early detection errors*, that are errors due to a premature recognition of an option while the participant is still reading it, and (b) *false detection errors*, that are errors due to the system recognizing a selection other than the one the participant intended. Out of the 1140 selections, there were 124 errors (≈10.9%): 88 of them were early detections (≈7.7%), while the remaining 36 were false detections (≈3.2%).

Early Detection Errors

A repeated measures ANOVA showed significant effects for trajectory type on early detection errors $F_{4,72} = 15.353$, $p < 0.001$. Table 1 summarizes the results of post-hoc anal-

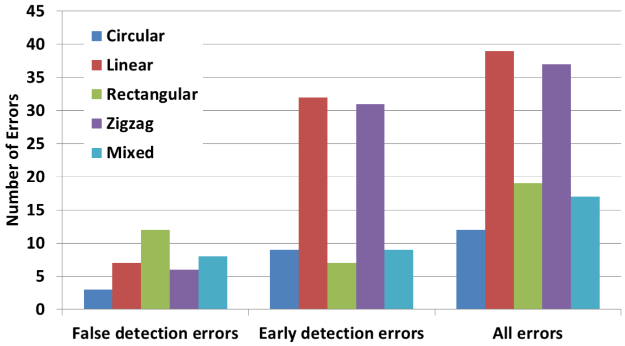


Figure 3. Circular, rectangular and mixed trajectories result in significantly fewer early detection errors compared to linear and zigzag ones. False detection errors are also fewer in the case of circular trajectories.

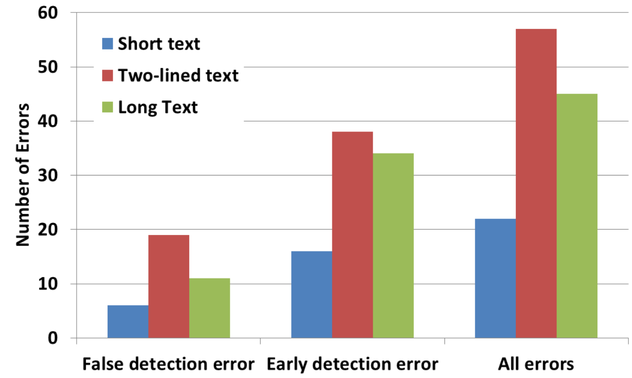


Figure 5. Fewer errors occurred when using short text compared to two-lined and long text.

yses using Bonferroni correction, which revealed significant differences between multiple pairs. Circular, rectangular and mixed trajectories result in significantly fewer early detection errors compared to linear and zigzag trajectories.

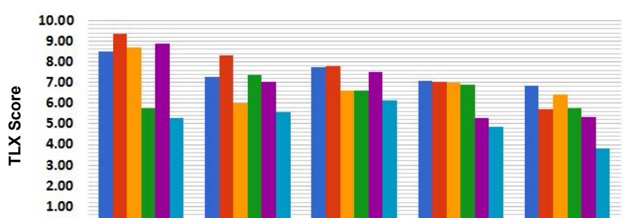
No significant effects were found for text representation on early detection errors. However the results suggest that fewer early detections occurred in the case of short text (Figure 5).

False Detection Errors

No significant effects for trajectory type on false detection errors were found ($p > 0.05$). Although a repeated measures ANOVA showed significant effects for text representation on false detection errors $F_{4,36} = 3.916$, $p < 0.05$, no significant differences were found between pairs. This is likely due to the low number of false detection errors (36 out of 1140 selections). Figures 3 and 5 indicate a tendency for fewer false detections in the case of Circular trajectories and short text.

Perceived Workload

Figure 4 summarises the mean values for each subscale. A repeated measures ANOVA showed significant effects for trajectory type on *physical demand* $F_{4,84} = 4.631$, $p < 0.005$ and *effort* $F_{4,84} = 4.334$, $p < 0.005$. Post-hoc analyses using Bonferroni correction showed that there are significant differences between *physical demand* ($p < 0.05$) induced by circular ($M = 9.4$, $SD = 5.1$) and mixed trajectories ($M = 5.7$, $SD = 3.8$). Significant differences were also found between *effort* ($p < 0.05$) to select circular ($M = 8.9$, $SD = 5.5$) and mixed trajectories ($M = 5.3$, $SD = 3.7$).



	Circular	Linear	Rectangular	Zigzag	Mixed
Mental	8.50	7.27	7.73	7.09	6.86
Physical	9.36	8.32	7.77	7.05	5.73
Temporal	8.68	6.00	6.59	7.00	6.41
Performance	5.77	7.36	6.59	6.91	5.77
Effort	8.86	7.05	7.50	5.27	5.32
Frustration	5.27	5.59	6.14	4.86	3.82

Figure 4. Although circular trajectories were perceived to have equal performance to mixed trajectories, participants perceived mixed trajectories to be less demanding in all other aspects.

In summary, although circular trajectories were perceived to perform similar to mixed trajectories, participants perceived mixed trajectories as less demanding in all other aspects.

Summary

The previous study showed that it is feasible to use Pursuits for text selection, with the selectable text itself being the moving stimulus. Our results confirm that by using certain text representations and trajectory types, text can be used as a stimulus for one of the major uses of Pursuits, that is interaction. Next we investigate the second major usage of Pursuits, which is eye tracker calibration, with text used as stimulus.

PURSUIT CALIBRATION USING TEXTUAL STIMULI

One of the major strengths of using text for Pursuit calibration is that it allows for seamlessly calibrating an eye tracker on a public display as passersby simply read text. Previous work by Pfeuffer et al. [36] on Pursuit calibration also included a text label "Please wait", floating from one corner of the display to another for 13 seconds. However, this required users to fixate the floating word for a long period even after reading, which might not be intuitive without prior instructions. To address this, we reveal text gradually to intrigue users to *pursue* the revealing text till its end (see Figure 1B).

In this section we present the implementation and evaluation of our prototype system called Read2Calibrate.

Concept and Implementation

In our implementation of Read2Calibrate, we developed a method that reveals the text gradually at an inclined angle. As new parts of the text gradually appear, the preceding parts disappear at the same rate (see Figure 1B). We opted for this representation in order to (1) ensure that users follow the stimulus (i.e. the text) till the end of the calibration session, (2) control the user's reading speed, and (3) calibrate for as much area of the display as possible. In the following, we explain the rationale behind these three motives.

(A) (B)

Figure 6. In the calibration session, participants read text gradually appearing across the display (as in Figure 1B). (A) and (B) show a sample text displayed at inclination angles of 15° and 315°. The angle and dotted lines are only shown for illustration and were not displayed to participants.

Revealing Speed

To correlate eye movements to movements of a stimulus, the speed of the moving stimulus needs to be known beforehand. Reading speeds are different across users [26], which makes it difficult to predict which part of the text a user is looking at, in particular in a public display setting with no prior knowledge about the user[1]. However, revealing the text gradually ensures a controlled reading speed, that is, if the user is reading any part of the text, we can expect that the user is looking at the letters being revealed at that moment. To ensure the user is performing a smooth pursuit eye movement rather than a series of saccadic jumps, new parts of the text are gradually revealed. As newer parts of the text appear, preceding parts disappear gradually, to reduce the chances of backward saccadic jumps (see Figure 1B).

The speed of revealing text is an important variable. If the revealing speed is too fast, users might not have enough time to comprehend the text. Additionally, because eye trackers are typically limited to a certain sampling frequency, the faster text is revealed, the less mappings between eye movements and points on the display are collected. On the other hand, very slow revealing speed could also result in difficulty in understanding the text; as the time difference between revealing the first and last letters of a word is larger, the more difficult it becomes to understand the word. There is also a risk of the users losing interest and looking at other parts on the screen, which would negatively influence the calibration.

Based on prior work [36] and pilot tests, we introduced pauses of 500 ms, 350 ms and 200 ms in-between revealing each letter, which in visual angles equates to speeds of 1°/s, 2.1°/s, and 4°/s. Higher and lower revealing speeds were found to be very difficult to read. We refer to these speeds as the slow, medium and fast speed, respectively.

Inclination Angle

Read2Calibrate needs to collect mappings between gaze points and points on the display. To cover as large an area of the display as possible, previous work used a stimulus that moved across the display in diagonal, circular, or rectangular trajectories [36, 37]. Circular and rectangular trajectories are unnatural for gradually revealing text, while limiting stimuli to a horizontal line would calibrate with respect to the x-axis only. Hence, we chose to reveal the text in diagonal shapes.

However, since there has been no previous work about reading inclined text that is gradually appearing, we experimented with multiple inclination angles. Latin script is read from left to right, hence the text could be shown in two ways: starting from the upper-left part and ending in the lower-right part of the screen, or starting from the lower-left part and ending in the upper-right part of the screen. This translates to inclination angles between 270°– 360° and between 0°– 90 degrees°.

Taking into consideration the need to move the stimulus with respect to both axes, we experimented with six angles: 15°, 45°, 75°, 285°, 315°, and 345°. Figures 6A and 6B show sample text displayed at inclination angles of 15° and 315°.

Calibration Correction

A prerequisite for calibration is to gaze at the stimulus. To exclude cases where users are not looking at the stimulus, our system calibrates after a certain correlation has been reached. We used the Pearson's product-moment coefficient with a correlation threshold of 0.6, that is, the user is assumed to follow the stimulus if the correlation between its movement and the users' eye-movements is ≥ 60%. We selected this value based on pilot testing and experience from prior work [36, 47].

In a calibration session, the letters are placed on the screen according to their angle. As the letters start to appear, pairs of gaze points and the revealed letter's coordinates are collected.

To calculate the correction offset, for every gaze point (G) recorded by the eye tracker, we measured the Euclidean distance between G and the center of the currently revealed letter (L). After the calibration phase ends, the sum of these distances is divided by the total number of gaze points (N) detected in that time frame. The resulting average distance value is then used as the correction offset (see Equation 1).

$$Offset = \frac{\sum_{k=1}^{N} L_k - G_k}{N} \quad (1)$$

[1]Note, that a future implementation could try to automatically assess the reading speed. This, however, would prolong the process.

Figure 7. After each calibration session, participants proceeded to a testing session where they fixated at each shown point. During the study, one stimulus was shown at a time and blinked for 3 seconds, before transitioning to the following one.

Evaluation of Pursuit Calibration Using Text

The goal of this study was to evaluate the effectiveness of Read2Calibrate, as well as to understand the influence of the different variables (inclination angle of the text and revealing speed). We studied the influence of these variables on the calibration quality, in addition to how users perceive them.

Apparatus

A 24 inch display (1920×1080 pixels) was equipped with a Tobii REX eye tracker (30Hz) and deployed in our lab (see Figure 1B). We invited 18 participants (10 females) aged 18 – 42 years ($M = 26.2$, $SD = 5.3$) through mailing lists. All participants had normal or corrected to normal vision.

Design

In a repeated measures experiment, every participant performed one calibration session per condition (6 angles × 3 speeds = 18 conditions). Each calibration session was followed by a testing session, where the participant was asked to gaze at a stationary point that appeared at nine positions on the screen (see Figure 7). The point blinked at each of the nine positions for 3 seconds. The order of the conditions was counter balanced across participants using a Latin-square.

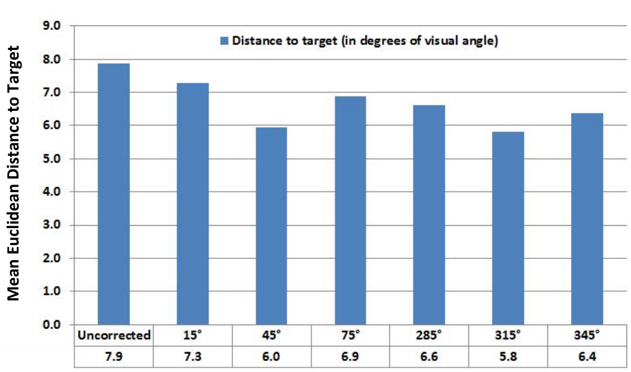


Figure 8. The graph shows the mean Euclidean distances between the gaze point and the target point after calibration by revealing text at different inclination angles, and in case of not applying any corrections. Lowest distance is achieved when text is revealed at an angle of 315°, followed by 45°.

Procedure

The experimenter started by explaining the study and the participant filled in a consent-form. The participant was then asked to stand at a distance of 60 cm from the eye tracker. The font-size was set to be easily readable at that distance. In each calibration session, the participant read a piece of a story that was gradually revealing across the display according to the condition's angle and speed. The participant then proceeded to a testing session, where we logged the gaze points, as detected by the eye tracker, the coordinates of the revealed text, as well as the corrections by our algorithm. After gazing at all nine stimuli, the calibration was reset and the participant proceeded to the next calibration session.

To reduce possible eye fatigue, participants were optionally allowed to take a break after every 3 calibration sessions. No visual feedback was shown during the whole experiment to avoid influencing the participant's gaze behavior.

We used a fable as a source of the revealing text. To ensure the validity of our analysis, it was crucial to make sure participants paid attention to the text. Hence, we asked participants three questions about the fable at the end. In addition to the compensation for participation, participants were encouraged to pay attention to the story by promising them an additional monetary incentive (50 Euro cents) for each correct answer they provide to the questions. All participants were aware of the rewarding mechanism before taking part in the study.

We concluded the study with a questionnaire and a semi-structured interview.

Results

To evaluate the effectiveness of Read2Calibrate, we based our calculations on

1. the *target point* (the center of the stimulus which was gazed at during the testing session),
2. the *uncalibrated gaze point* (the gaze point as detected by the uncalibrated eye tracker), and
3. the *calibrated gaze point* (the gaze point after correction by Read2Calibrate).

For each stimuli shown in the testing sessions, we measured the *mean Euclidean distance* (1) between the uncalibrated gaze points and the target point and (2) between the calibrated points and the target point. Moreover, we measured the *positive correction rate*, which we define as the number of times the calibrated gaze point was closer to the target compared to the uncalibrated one.

Quantitative Results

Figure 8 shows that the mean Euclidean distance is shorter when text is revealed at angles of 315° and 45°. Thus, these angles result in better correction compared to others.

A repeated measures ANOVA showed significant main effects for angle $F_{3.2,51.3} = 5.2$, $p < 0.005$ and speed $F_{2,34} = 4.8$, $p < 0.05$ on positive correction rate. Post-hoc analysis using Bonferroni correction showed a significant difference ($p < 0.05$) in positive correction rate for an inclination angle of 315° ($M = 65\%$, $SD = 0.06\%$) compared to 15° ($M = 39.7\%$,

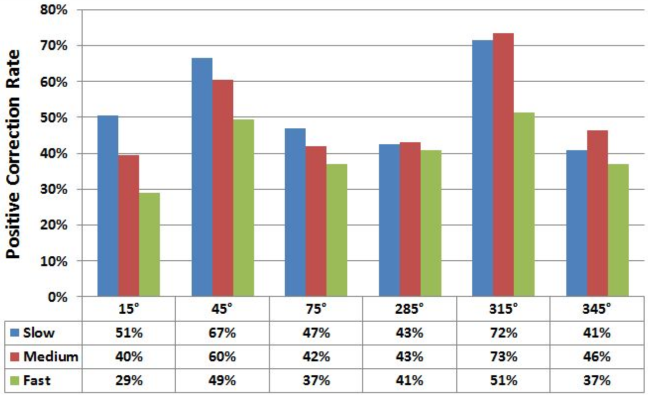


Figure 9. Revealing the text in a 315°inclination resulted in the highest number of positive corrections, e.g. 73% of the corrections by the Read-2Calibrate brought the gaze point closer to the target.

SD = 0.05%), and also for 315° (M = 65%, SD = 0.06%) compared to 345° (M = 41.4%, SD = 0.06%). There were also significant differences in positive correction rate for fast revealing speed (M = 40.7%, SD = 0.05%) compared to slow revealing speed (M = 53.2%, SD = 0.045%). Figure 9 shows that angles of 315° and 45° resulted in more positive corrections compared to other angles. The figure also shows that fast revealing speeds result in less positive corrections.

Qualitative Feedback

When asked how easy it is to read the text at the different angles (5-point Likert scale; 1=Strongly disagree; 5=Strongly agree), participants indicated that most angles were easy to read (see Table 2). As for the revealing speeds, participants found the medium speed (Med = 5, SD = 0.6) to be easier to follow compared to slow (Med = 4, SD = 1.3) and fast speeds (Med = 4, SD = 1.2).

When asked about their preference, participants pointed that they preferred angles that are closer to a horizontal line (i.e. 15° and 345°) as they felt more natural. However as indicated in the questionnaire, other angles are also easy to follow. On the other hand, multiple participants indicated that it felt unnatural to read the slow revealing text, P6 noted that "*I felt I was following the letters without really understanding the words*". According to the participants, fast text is easy to follow, but difficult to comprehend.

DISCUSSION & DESIGN RECOMMENDATIONS

Overall, the results of both studies suggest that text can be used as a stimulus to support interaction and calibration using Pursuits.

The text selection study showed that when using circular and mixed trajectories, shorter text can be selected with high accuracy using Pursuits while selecting longer pieces of text is more difficult. The text-based Pursuit calibration study showed that text can be an effective stimulus for seamlessly integrating eye tracker calibration as users read text. More specifically, gradually revealing text inclined at 45° or 315° at a speed of 2.1 visual degree angles per second highly improves the accuracy of the gaze point. The results also indicated that although participants preferred flat text, revealing text at inclined angles is easily readable and can be used for calibration.

Text Inclination Angle	15°	45°	75°	285°	315°	345°
Median Score	5	4	4	4	4	4.5
Standard Deviation	0.43	0.67	0.92	0.96	0.83	0.70

Table 2. The table summarizes responses of participants when asked if it is easy to read text inclined at the corresponding angles (5-point Likert scale; 1=Strongly disagree; 5=Strongly agree). This indicates that the closer the text's inclination to a horizontal line (i.e. 15° and 345°), the easier it is to read. However, the average scores of other angles as well as feedback from participants indicate that although other angles are less preferred, they are still readable.

Text Selection via Pursuits

Overall, there was a low number of false detection errors – 36 false detection errors out of 1140 selections. Figure 3 shows that circular trajectories tend to be associated with less false detection errors. The results of the text selection study show that circular, rectangular, and mixed trajectories result in fewer early detection errors compared to linear and zigzag ones (see Figure 5 and Table 1). This means that reading text moving in linear and zigzag trajectories results in high correlation between eye movements and text movement, making the system confuse reading with selection.

Reading involves repeated saccades across a line of text as well as back to the beginning of the next line. Performing these saccades while pursuing text moving in a circular or rectangular trajectory distinguishes the eye movements from the text's trajectory. This reading overlay makes the gaze trajectory less likely to correlate with that of the moving text, giving the user a chance to read and comprehend the text. On the other hand, reading text moving in linear and zigzag motion can be hardly distinguished from a Pursuit selection, resulting in a high correlation while reading, which in turn results in many early detection errors.

Selecting Long Pieces of Text

Our motivation behind the use of different text representations and trajectories was to study how the Pursuits method can cope with a read-and-pursue eye movement. Our main finding is that Pursuits is indeed suitable for text selection, but only with shorter pieces of text. In cases where it is essential to show passersby longer pieces of text to select from, we recommend using a different stimulus.

R1: *Use Pursuits for selection of short pieces of text; for longer pieces of text (25+ letters) use non-textual stimuli.*

In case longer textual descriptions are needed, a display could be divided into two regions: a non-interactive region and an interactive region. In the case of the voting application, the non-interactive region could display the detailed answer options, each with a short but meaningful headline. The interactive region could then display the moving headlines only from which users can then make their selection. Alternatively, answers in the non-interactive region could be associated with colors or shapes. However, this may result in a higher cognitive load, since users need to associate colors or forms with the correct answer option.

Choosing the Right Trajectory Type

Our analyses of the text selection experiment results show that circular trajectories are safer to use, as they result in significantly fewer errors. However circular trajectories are perceived to be highly demanding (see Figure 4). Mixed trajectories result in slightly more errors than circular ones. However, mixed trajectories were perceived to be significantly less demanding compared to other trajectories. This indicates a trade-off between user preference and accuracy of the system.

R2: *For peak performance when performing text selection using Pursuits, move text in circular trajectories. To increase user experience, a mixture of trajectories can be used together with an undo option.*

We conclude that if high accuracy when selecting text using Pursuits is required, designers should opt for circular trajectories. An example could be a situation where users are encountered in a passing-by situation or in a situation where they have only too little time to undo their selections, such as asking a question on customer satisfaction near a cashier. On the other hand, in cases where user experience should be maximized, mixed trajectories may be used that are slightly more error prone. In these cases, a floating "undo" label could be shown after the system detected a selection. An example could be users filling in a longer survey in return for an incentive, such as a gift voucher. Here it may be acceptable to occasionally 'correct' an answer while at the same time having a less demanding experience with the system.

Text-based Pursuit Calibration

In general, participants' feedback indicates that Read2-Calibrate is positively perceived. The results of its evaluation show that inclination angles that result in diagonal-like orientation of the revealing text, such as 45° and 315°, significantly improve the accuracy of the gaze point. This is due to the fact that these angles result in the text covering larger areas of the screen. Inclination angles that bring the text closer to a horizontal line are preferred by users (15° and 345°), as they are more similar to flat text which users are acquainted to read. However, at these angles the text covers relatively less area with respect to the y-axis, resulting in poor calibration.

Slow revealing speeds result in users focusing on letters and losing track of the words they read. By analyzing the data, it was found that when using slower speeds participants were more likely to lose interest and look at other parts of the screen, presumably out of boredom. Fast speeds result in less data collected for the correlation check, which in turn results in lower calibration accuracy. Moreover, participants reported that revealing the text too fast makes it harder for them to understand what they read. Medium revealing speed turned out to be a good compromise: it is preferred by users and also results in a good calibration quality.

It should be noted, however, that the accuracy achieved by Read2Calibrate is lower than that of previous approaches that use Pursuits for calibration as well as of explicit calibration methods commonly known from eye tracking in desktop settings. At the same time, the major advantage of text-based Pursuits calibration is the seamless integration with users simply reading content on the public display. As a result, the calibration can be performed and used even in cases where the reader is not being consciously aware of it. Gaze information can then be used to enhance the user interface. For example, one may show a description of different sights next to a map of a city, like museums, churches, or historic buildings. As the system is calibrated, dwell time towards different sights could be used to determine what the reader is most interested in and additional information on how to reach a particular sight together with a discount coupon could be presented.

R3: *For moderate eye tracker calibration (accuracy of 5.8° of visual angle), text-based Pursuit calibration is recommended as it results in better user experience. If accuracy is crucial, classical Pursuit calibration should be used.*

The trade-off between accuracy and user experience can also be found when determining the angles at which the revealed text is inclined in Read2Calibrate. While bringing the text closer to a horizontal line makes reading feel more natural, revealing the text in a diagonal-like path results in the highest accuracy. Very steep text (e.g. 75° and 285°) result in both low accuracy and worse user experience and should hence be avoided.

R4: *Use diagonally-shaped paths, at an inclination of 315° or 45°, when revealing text to achieve highest accuracy with Read2Calibrate. For better user experience at the expense of calibration accuracy, reveal text in flatter shaped paths.*

A clear recommendation with regard to revealing speed can be provided. Here, accuracy is highest for medium revealing speed (2.1°/s) and this it in line with the users' preference.

R5: *For text-based Pursuit calibration, an average revealing speeds of about 2.1° of visual angles per second should be used.*

The more comfortable a participant is with the revealing speed, the more accurate gaze-to-display mappings are collected and hence the more accurate the calibration is. Faster speeds result in fewer mappings, while slower ones distract the user.

Use Cases and Integration With Interactive Applications

As a sample applications that can be explicitly controlled using gaze, we implemented EyeVote, a voting system. Civic discourse is a popular use case for public displays [15, 16, 35, 39], where passersby select from a set of text-based options to express their opinions. Given the advantages of gaze for public displays [22], gaze-based voting systems can utilize Pursuits for selection of textual answers. Similarly, Pursuits can be used to answer on-screen quizzes. Selection of text via Pursuits can be useful in various other contexts, for example, users can select from a set of text-based options displayed at a museum to learn more about particular topics. In train-stations and airports, Pursuits can be employed to set fare preferences or pick products where possible options are displayed as text.

The second major use case is the implicit use of gaze data, either for analysis or for adaptive interfaces. Therefore, text-based Pursuit calibration can be integrated into public display applications in several ways. For example, a common practice to tackle interaction blindness on public displays is to use call-to-action labels [32]. Such labels could serve as stimuli to calibrate an eye-tracker via Read2Calibrate. Further stimuli could be welcome messages or brief instructions on how to use a system or how to play a game. While the aforementioned examples utilize Read2Calibrate at the beginning of the interaction process, revealing text can also be shown amidst interaction. For example, a short text could hint at hidden or yet undiscovered features. Such a *calibration while interacting* may be useful for displays featuring multiple applications. Here, a game that is instantly usable may serve as an entry point. As the user finishes playing the game in the course of which the eye-tracker was calibrated using in-game textual content, the display could present further content that could benefit from knowledge about the user's gaze behavior. Note, that after the calibration, fine-grained gaze points can be collected and, hence, also other types of eye movements, such as fixations and saccades can be detected. As a result, an application may determine interest towards a particular content – this may be of particular interest for advertisers – as well as identify difficulties of users in perceiving content, for example, as they read text over and over again.

Limitations and Future Work

Firstly, our evaluations so far were conducted in the lab. While this controlled setting was necessary to maximize internal validity and obtain comparable results, future work could employ text-based stimuli for Pursuits in an in-the-wild setting. Apart from verifying the results with regard to accuracy and errors, this may yield further insights on audience behavior and acceptance.

Secondly, participants of the text selection study answered 60 questions using Pursuits, whereas participants of the Read2-Calibrate study performed 18 calibration sessions. In a real-world situation, it is unlikely that users would perform such a high number of selections and users would not be required to verify the accuracy of the calibration. As a result, we expect the study to have caused a higher level of eye fatigue as an in-the-wild exposure to any of the systems would have done. Hence, participants may have been overly critical during their assessment of the system. Future work could capture in-situ feedback to verify the impact of our approach on the experience users have during use of our system.

Recent work explored feedback methods for Pursuit selections. Kangas et al. [19] compared different feedback modalities for Pursuits and found that haptic feedback is preferred by users compared to visual and auditory feedback. Špakov et al. [51] compared two smooth pursuit widgets to find that circular widgets exhibit higher performance. An additional direction for future work is to enable feedback methods to improve the user experience when using EyeVote and Read2-Calibrate. For example, in text-selection tasks, visual cues can be used to incrementally highlight the text whose trajectory correlates the most with eye movements, depending on the current correlation value.

Another interesting direction for future work would be to try different scripts. For example, Arabic and Hebrew are read from right to left, while Chinese, Japanese and Korean can also be read vertically. In our implementation of Read2-Calibrate, text was revealed in a diagonal path. The flexibility of some east Asian scripts makes it possible to experiment with revealing text in different paths (e.g., a rectangular path).

CONCLUSION

In this work we investigated the use of text as a stimulus for Pursuits, to enable gaze-based interaction and eye tracker calibration on public displays. Our results show that text can be a powerful stimulus for both tasks. Shorter pieces of text can be robustly selected using Pursuits, and text-based calibration improves gaze point accuracy. We found that Pursuits-based text selection is less error-prone when text follows circular trajectories. We also found that the use of different trajectories simultaneously (mixed trajectories) is better perceived by users and results in relatively few errors. Read2Calibrate was shown to improve the accuracy of the gaze point, in particular when using text that is gradually revealing at a speed of 2.1°/s and inclined at a 315° or 45° angle.

REFERENCES

1. Florian Alt, Jörg Müller, and Albrecht Schmidt. 2012a. Advertising on Public Display Networks. *Computer* 45, 5 (2012), 50–56.
2. Florian Alt, Stefan Schneegaß, Albrecht Schmidt, Jörg Müller, and Nemanja Memarovic. 2012b. How to Evaluate Public Displays. In *Proc. of PerDis '12*. ACM, New York, NY, USA.
3. Florian Alt, Alireza Sahami Shirazi, Thomas Kubitza, and Albrecht Schmidt. 2013. Interaction Techniques for Creating and Exchanging Content with Public Displays. In *Proc. of CHI '13*. ACM, New York, NY, USA, 1709–1718.
4. Florian Alt and Julia Vehns. 2016. Opportunistic Deployments: Challenges and Opportunities of Conducting Public Display Research at an Airport. In *Proc. of PerDis'16*. ACM, New York, NY, USA.
5. Ralf Biedert, Georg Buscher, Sven Schwarz, Jörn Hees, and Andreas Dengel. 2010. Text 2.0. In *Proc. of CHI EA '10*. ACM, New York, NY, USA, 4003–4008.
6. Feridun M. Celebi, Elizabeth S. Kim, Quan Wang, Carla A. Wall, and Frederick Shic. 2014. A Smooth Pursuit Calibration Technique. In *Proc. of ETRA '14*. ACM, New York, NY, USA, 377–378.
7. Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proc. of ICMI '15*. ACM, New York, NY, USA, 131–138.
8. Jorgos Coenen, Niels Wouters, and Andrew Vande Moere. 2016. Synchronized Signage on Multiple Consecutively Situated Public Displays. In *Proc. of PerDis '16*. ACM, New York, NY, USA.
9. Ashley Colley, Leena Ventä-Olkkonen, Florian Alt, and Jonna Häkkilä. 2015a. Insights from Deploying See-Through Augmented Reality Signage in the Wild. In *Proc. of PerDis '15*. ACM, New York, NY, USA, 179–185.

10. Ashley Colley, Lasse Virtanen, Jani Väyrynen, and Jonna Häkkilä. 2015b. Physically Guiding Touch Screen Interaction with Public Displays. In *Proc. of PerDis '15*. ACM, New York, NY, USA, 255–256.

11. Dietlind Helene Cymek, Antje Christine Venjakob, Stefan Ruff, Otto Hans-Martin Lutzand Simon Hofmann, and Matthias Roetting. 2014. Entering PIN codes by smooth pursuit eye movements. *Journal of Eye Movement Research* 7(4):1 (2014), 1–11.

12. Tilman Dingler, Tobias Bagg, Yves Grau, Niels Henze, and Albrecht Schmidt. 2015. uCanvas: A Web Framework for Spontaneous Smartphone Interaction with Ubiquitous Displays. In *Human-Computer Interaction–INTERACT 2015*. Springer, 402–409.

13. Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015a. Orbits: Enabling Gaze Interaction in Smart Watches Using Moving Targets. In *Proc. of UbiComp/ISWC'15 Adjunct*. ACM, New York, NY, USA, 419–422.

14. Augusto Esteves, Eduardo Velloso, Andreas Bulling, and Hans Gellersen. 2015b. Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements. In *Proc. of UIST '15*. ACM, New York, NY, USA, 457–466.

15. Luke Hespanhol, Martin Tomitsch, Ian McArthur, Joel Fredericks, Ronald Schroeter, and Marcus Foth. 2015a. Situated Interfaces for Engaging Citizens on the Go. *interactions* 23, 1 (Dec. 2015), 40–45.

16. Luke Hespanhol, Martin Tomitsch, Ian McArthur, Joel Fredericks, Ronald Schroeter, and Marcus Foth. 2015b. Vote as you go: Blending interfaces for community engagement into the urban space. In *7th International Conference on Communities and Technologies (C&T)*. ACM, University of Limerick, Limerick, Ireland, 29–37.

17. Simo Hosio, Jorge Goncalves, Vassilis Kostakos, and Jukka Riekki. 2014. Exploring Civic Engagement on Public Displays. In *User-Centric Technology Design for Nonprofit and Civic Engagements*. Springer, 91–111.

18. Robert J. K. Jacob. 1991. The Use of Eye Movements in Human-computer Interaction Techniques: What You Look at is What You Get. *ACM Trans. Inf. Syst.* 9, 2 (April 1991), 152–169.

19. Jari Kangas, Oleg Špakov, Poika Isokoski, Deepak Akkil, Jussi Rantala, and Roope Raisamo. 2016. Feedback for Smooth Pursuit Gaze Tracking Based Control. In *Proc. of AH '16*. ACM, New York, NY, USA, Article 6, 8 pages.

20. Mohamed Khamis, Florian Alt, and Andreas Bulling. 2015. A Field Study on Spontaneous Gaze-based Interaction with a Public Display Using Pursuits. In *Adjunct Proc. of UbiComp '15*. ACM, New York, NY, USA, 863–872.

21. Mohamed Khamis, Florian Alt, and Andreas Bulling. 2016. Challenges and Design Space of Gaze-enabled Public Displays. In *Ajunct Proc. of UbiComp '16*. ACM, New York, NY, USA, 10.

22. Mohamed Khamis, Andreas Bulling, and Florian Alt. 2015. Tackling Challenges of Interactive Public Displays Using Gaze. In *Adjunct Proc. of UbiComp '15*. ACM, New York, NY, USA, 763–766.

23. R. Kishi and T. Hayashi. 2015. Effective gazewriting with support of text copy and paste. In *Computer and Information Science (ICIS), 2015 IEEE/ACIS 14th International Conference on*. 125–130.

24. Lisa Koeman, Vaiva Kalnikaité, and Yvonne Rogers. 2015. "Everyone Is Talking About It!": A Distributed Approach to Urban Voting Technology and Visualisations. In *Proc. of CHI '15*. ACM, New York, NY, USA, 3127–3136.

25. Manu Kumar, Andreas Paepcke, and Terry Winograd. 2007. EyePoint: Practical Pointing and Selection Using Gaze and Keyboard. In *Proc. of CHI '07*. ACM, New York, NY, USA, 421–430.

26. Kai Kunze, Katsutoshi Masai, Masahiko Inami, Ömer Sacakli, Marcus Liwicki, Andreas Dengel, Shoya Ishimaru, and Koichi Kise. 2015. Quantifying Reading Habits: Counting How Many Words You Read. In *Proc. of UbiComp '15*. ACM, New York, NY, USA, 87–96.

27. Christian Lander, Sven Gehring, Antonio Krüger, Sebastian Boring, and Andreas Bulling. 2015. GazeProjector: Accurate Gaze Estimation and Seamless Gaze Interaction Across Multiple Displays. In *Proc. of UIST '15*. ACM, New York, NY, USA, 395–404.

28. Dachuan Liu, Bo Dong, Xing Gao, and Haining Wang. 2015. Exploiting Eye Tracking for Smartphone Authentication. In *Proc. of ACNS '15*. 20.

29. Otto Hans-Martin Lutz, Antje Christine Venjakob, and Stefan Ruff. 2015. SMOOVS: Towards calibration-free text entry by gaze using smooth pursuit movements. *Journal of Eye Movement Research* 8(1):2 (2015), 1–11.

30. Päivi Majaranta. 2012. Communication and Text Entry by Gaze. In *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies*. IGI Global, 63–77.

31. Päivi Majaranta and Andreas Bulling. 2014. Eye Tracking and Eye-Based Human-Computer Interaction. In *Advances in Physiological Computing*. Springer, 39–65.

32. Paul Marshall, Richard Morris, Yvonne Rogers, Stefan Kreitmayer, and Matt Davies. 2011. Rethinking 'Multi-user': An In-the-wild Study of How Groups Approach a Walk-up-and-use Tabletop Interface. In *Proc. of CHI '11*. ACM, New York, NY, USA, 3033–3042.

33. Jörg Müller, Florian Alt, Daniel Michelis, and Albrecht Schmidt. 2010. Requirements and Design Space for Interactive Public Displays. In *Proc. of MM '10*. ACM, New York, NY, USA, 1285–1294.

34. Jörg Müller, Robert Walter, Gilles Bailly, Michael Nischt, and Florian Alt. 2012. Looking Glass: A Field Study on Noticing Interactivity of a Shop Window. In *Proc. of CHI '12*. ACM, New York, NY, USA, 297–306.

35. Erica C. Ostermann, Long Ma, Daniel Sussman, and Susan R. Fussell. 2015. CommunityConnect: An Interactive Display for Educational Residential Settings. In *Proc. of CSCW '15 Companion*. ACM, New York, NY, USA, 175–178.
36. Ken Pfeuffer, Mélodie Vidal, Jayson Turner, Andreas Bulling, and Hans Gellersen. 2013. Pursuit Calibration: Making Gaze Calibration Less Tedious and More Flexible. In *Proc. of UIST '13*. ACM, New York, NY, USA, 261–270.
37. Ken Pfeuffer, Yanxia Zhang, and Hans Gellersen. 2015. A Collaborative Gaze Aware Information Display. In *Adjunct Proc. of UbiComp '15*. ACM, New York, NY, USA, 389–391.
38. Selina Sharmin, Oleg Špakov, and Kari-Jouko Räihä. 2013. Reading On-screen Text with Gaze-based Auto-scrolling. In *Proc. of ETSA '13*. ACM, New York, NY, USA, 24–31.
39. Fabius Steinberger, Marcus Foth, and Florian Alt. 2014. Vote With Your Feet: Local Community Polling on Urban Screens. In *Proc. of PerDis '14*. ACM, New York, NY, USA, 44:44–44:49.
40. Sophie Stellmach and Raimund Dachselt. 2012. Look & Touch: Gaze-supported Target Acquisition. In *Proc. of CHI '12*. ACM, New York, NY, USA, 2981–2990.
41. Sophie Stellmach and Raimund Dachselt. 2013. Still Looking: Investigating Seamless Gaze-supported Selection, Positioning, and Manipulation of Distant Targets. In *Proc. of CHI '13*. ACM, New York, NY, USA, 285–294.
42. Nick Taylor, Justin Marshall, Alicia Blum-Ross, John Mills, Jon Rogers, Paul Egglestone, David M. Frohlich, Peter Wright, and Patrick Olivier. 2012. Viewpoint: Empowering Communities with Situated Voting Devices. In *Proc. of CHI '12*. ACM, New York, NY, USA, 1361–1370.
43. Takumi Toyama, Daniel Sonntag, Andreas Dengel, Takahiro Matsuda, Masakazu Iwamura, and Koichi Kise. 2014. A Mixed Reality Head-mounted Text Translation System Using Eye Gaze Input. In *Proc. of IUI '14*. ACM, New York, NY, USA, 329–334.
44. Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. 2014. MyPosition: Sparking Civic Discourse by a Public Interactive Poll Visualization. In *Proc. of CSCW '14*. ACM, New York, NY, USA, 1323–1332.
45. Eduardo Velloso, Markus Wirth, Christian Weichel, Augusto Esteves, and Hans Gellersen. 2016. AmbiGaze: Direct Control of Ambient Devices by Gaze. In *Proc. DIS '16*. ACM, New York, NY, USA, 812–817.
46. Roel Vertegaal. 2003. Attentive user interfaces. *Commun. ACM* 46, 3 (2003), 30–33.
47. Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proc. of UbiComp '13*. ACM, New York, NY, USA, 439–448.
48. Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2015. Pursuits: Spontaneous Eye-Based Interaction for Dynamic Interfaces. *GetMobile: Mobile Comp. and Comm.* 18, 4 (Jan. 2015), 8–10.
49. Mélodie Vidal, Ken Pfeuffer, Andreas Bulling, and Hans W. Gellersen. 2013. Pursuits: Eye-based Interaction with Moving Targets. In *Proc. of CHI EA '13*. ACM, New York, NY, USA, 3147–3150.
50. Vasilis Vlachokyriakos, Rob Comber, Karim Ladha, Nick Taylor, Paul Dunphy, Patrick McCorry, and Patrick Olivier. 2014. PosterVote: Expanding the Action Repertoire for Local Political Activism. In *Proc. of DIS '14*. ACM, New York, NY, USA, 795–804.
51. Oleg Špakov, Poika Isokoski, Jari Kangas, Deepak Akkil, and Päivi Majaranta. 2016. PursuitAdjuster: An Exploration into the Design Space of Smooth Pursuit –based Widgets. In *Proc. of ETRA '16*. ACM, New York, NY, USA, 287–290.
52. Xuehan Xiong, Zicheng Liu, Qin Cai, and Zhengyou Zhang. 2014. Eye Gaze Tracking Using an RGBD Camera: A Comparison with a RGB Solution. In *Adjunct Proc. of UbiComp '14*. ACM, New York, NY, USA, 1113–1121.
53. Yanxia Zhang, Andreas Bulling, and Hans Gellersen. 2013. SideWays: A Gaze Interface for Spontaneous Interaction with Situated Displays. In *Proc. of CHI '13*. ACM, New York, NY, USA, 851–860.
54. Yanxia Zhang, Andreas Bulling, and Hans Gellersen. 2014. Pupil-canthi-ratio: a calibration-free method for tracking horizontal gaze direction. In *Proc. of AVI '14*. ACM, New York, NY, USA, 129–132.
55. Yanxia Zhang, MingKi Chong, Jörg Müller, Andreas Bulling, and Hans Gellersen. 2015. Eye tracking for public displays in the wild. *Personal and Ubiquitous Computing* 19, 5-6 (2015), 967–981.
56. Yanxia Zhang, Jörg Müller, Ming Ki Chong, Andreas Bulling, and Hans Gellersen. 2014. GazeHorizon: Enabling Passers-by to Interact with Public Displays by Gaze. In *Proc. of UbiComp '14*. ACM, New York, NY, USA, 559–563.