Multimodal Technologies and Interaction

MDPI

Article

EyeSpot: Leveraging Gaze to Protect Private Text Content on Mobile Devices from Shoulder Surfing

Mohamed Khamis [1,2,*], Malin Eiband [1], Martin Zürn [1] and Heinrich Hussmann [1]

[1] Media Informatics Group, Ludwig Maximilian University of Munich, 80337 München, Germany; malin.eiband@ifi.lmu.de (M.E.); martin.zuern@campus.lmu.de (M.Z.); heinrich.hussmann@ifi.lmu.de (H.H.)

[2] Glasgow Interactive Systems Section, School of Computing Science, University of Glasgow, Glasgow, G12 8RZ, UK

* Correspondence: mohamed.khamis@ifi.lmu.de; Tel.: +49-163-485-6721

Received: 10 June 2018; Accepted: 25 July 2018; Published: 9 August 2018

Abstract: As mobile devices allow access to an increasing amount of private data, using them in public can potentially leak sensitive information through shoulder surfing. This includes personal private data (e.g., in chat conversations) and business-related content (e.g., in emails). Leaking the former might infringe on users' privacy, while leaking the latter is considered a breach of the EU's General Data Protection Regulation as of May 2018. This creates a need for systems that protect sensitive data in public. We introduce EyeSpot, a technique that displays content through a spot that follows the user's gaze while hiding the rest of the screen from an observer's view through overlaid masks. We explore different configurations for EyeSpot in a user study in terms of users' reading speed, text comprehension, and perceived workload. While our system is a proof of concept, we identify crystallized masks as a promising design candidate for further evaluation with regard to the security of the system in a shoulder surfing scenario.

Keywords: mobile devices; privacy; gaze; eye tracking; security

1. Introduction

Users interact with their mobile devices in different contexts, ranging from private spaces, such as homes and offices, to public areas such as public transport, transit areas, and workplaces. However, the convenience of being able to access sensitive data anywhere comes with the risk of exposing private information to bystanders through shoulder surfing. A survey by Eiband et al. revealed that interactions on mobile devices are often observed by bystanders which may leak sensitive private information about, for example, the user's personal relationships, interests and plans [1]. These problems may potentially become more prominent with the increased popularity of larger screens of smartphones and tablets where the content that users are not interacting with is unnecessarily exposed. Leaking business-related third-person data when working on the go, such as information about customers and employees, is even deemed a breach of the EU's General Data Protection Regulation as of May 2018 [2]. Furthermore, with growing tendency and expectations by employers and society to be always accessible, employees are more likely to engage in acts that compromise privacy, such as reading sensitive emails on the train.

This creates a need for systems that protect users' "visual privacy" [3,4] in public. While prior work presented various approaches to mitigate the shoulder surfing of credentials [5–7], the protection of other data is relatively underexplored with the exception of few works that we discuss further in Section 2 [4,8]. This work focuses on protecting sensitive text that is normally shown on screens, and hence, visible to bystanders.

We introduce EyeSpot, a technique inspired by prior work on visual privacy protection [4,9] and privacy protection with eye tracking [8,10,11]. EyeSpot utilizes the user's gaze to protect their visual privacy when interacting with a handheld mobile device. As illustrated in Figure 1, the on-screen content is hidden through overlaid masks while revealing the "spot" the user is gazing at. This allows the user to read content on the display while hiding the surrounding content from potential observers. We explored different configurations of EyeSpot in a user study. Namely, we compared blackout, crystallized, and fake text masks for hiding content (shown in Figure 1) using two different spot sizes (small and large) to a baseline (no-mask). We identify the crystallized mask as a promising candidate for further evaluation. It distorts the content while still allowing the user to read with higher speed and comprehension and a lower perceived workload compared to the other masks.

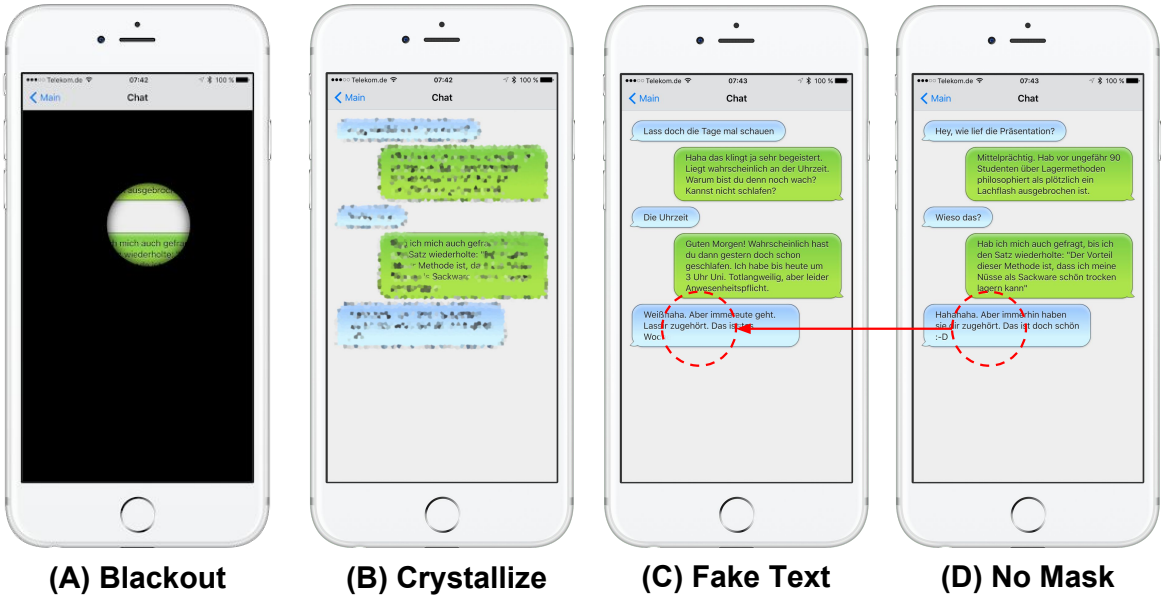


Figure 1. EyeSpot allows the user to perceive the display's content at the spot being gazed at, while hiding the remaining content to prevent leaking private data to bystanders. We experimented with different masks and spot sizes. The blackout mask (**A**) overlays the area around the user's gaze point with a black filter. crystallized mask (**B**) maintains the context (e.g., chat bubbles) but distorts the text. fake text mask (**C**) replaces the content with fake text (cf., no-mask in (**D**)). The circular markers illustrate how fake text differs from no-mask, but were not shown to the user.

Contribution Statement

The contribution of this work is two-fold: (1) We introduce the concept and implementation of EyeSpot that, when activated, utilizes the user's gaze point to preserve privacy on handheld mobile devices. (2) We report on the results of an evaluation of different configurations of EyeSpot, and conclude with recommendations based on these results.

2. Related Work

Our work builds on prior work in the areas of (1) visual privacy protection and (2) privacy protection using eye tracking.

2.1. Visual Privacy Protection

Physical privacy protectors are widely sold on the market. They utilize dark filters to reduce the viewing angles of screens and can be attached on mobile devices. However, while they indeed make observations harder, shoulder surfers can still peek at content from certain angles [12]. Brudy et al. introduced several concepts for protecting users from shoulder surfing on large public displays [13].

For example, they proposed blacking out sensitive content and dimming the entire screen except for the parts shielded by the user's body. In a wizard of Oz study, Zhou et al. explored concepts to inform users that they are being shoulder surfed [8,14]. They also collected subjective feedback about selectively hiding sensitive information, selectively showing part of the screen, or blacking out the screen. Although these privacy protection ideas are promising, to date there have been no formal user studies to investigate the impact of these methods on the usability of the system, apart from the subjective feedback collected by Zhou et al. [8,14]. In contrast, we evaluated an actual implementation of EyeSpot for handheld mobile devices and quantified the impacts of different configurations on reading speed, text comprehension, and perceived workload.

Other works explored protecting private content by distorting it. For example, von Zezschwitz et al. blurred photos in galleries of mobile devices to protect privacy [9]. They exploited the human ability to recognize distorted versions of photos they have seen before [15], while they remain ambiguous to observers who have not seen the original undistorted version. Several works have used the same concept for graphical password authentication [16–18]. Eiband et al. leveraged the fact that handwritten text is easier to read by the writer compared to others and accordingly, distorted on-screen text by displaying it in the user's handwriting [4]. They found that it is easier for users to read text shown in their own handwriting than in another person's handwriting. These approaches inspired us to experiment with different masks for EyeSpot. A key difference to the work of von Zezschwitz et al. [9] is that EyeSpot deals with content that has not been seen by the user before, and instead of distorting the entire view [4,9], we only distort parts of the display that are not being gazed at.

2.2. Privacy Protection Using Eye Tracking

Eye movements are generally subtle and difficult to observe [19]. This motivated the use of eye tracking in different ways for privacy protection. For example, prior work utilized gaze for explicit authentication [10,20–25], for implicitly improving graphical password selection [26,27], and for biometric authentication [28,29].

Recently, eye tracking has become feasible on handheld mobile devices [30]. For example, Liu et al. leveraged gaze gestures for authentication on mobile devices [11], while Khamis et al. combined gaze and touch for multimodal authentication on such devices [31]. Song et al. introduced biometric gaze-based authentication on mobile devices [32].

Apart from authentication, eye tracking has also been used for visual privacy protection. For example, Lian et al. detected eyes in video frames to determine whether anyone is shoulder surfing the user; if eyes are detected, the brightness of the screen is adapted accordingly [33]. Brudy et al. used a Kinect to estimate the gaze direction of passers-by in front of a large public display and visualized it to users of the display to make them aware of shoulder surfers [13]. While this approach works well for public displays, detecting the gaze direction of bystanders with the front-facing camera of a user's mobile device is often unfeasible due to its narrow-angle lens [3,34].

3. Concept

Two key factors influenced the design of EyeSpot: the *type of screen mask* used to hide the content that is not currently being gazed at and the *size of the spot* that follows the user's gaze.

3.1. Screen Masks

There are many possible ways to hide private content on screens of mobile devices. In this work, we investigated three masks that were motivated by related work.

3.1.1. Blackout Mask

The most intuitive way to protect private content is by completely hiding it. Hence, we experimented with a mask that blacks out the entire screen except for the region that the user is

gazing at. This was motivated by previously reported reactions to shoulder surfing in which users turned the screen off [1] or used privacy screens [12].

While this mask provides basic protection of the screen content, it has two disadvantages. First, it hides the layout of the interface. While hiding the activity itself has its security merits, hiding contextual elements that are normally visible in the periphery and useful for the legitimate user could impede usability. Second, it is straightforward to identify the spot's position (see Figure 1A).

3.1.2. Crystallized Mask

Research about "unconscious inference" confirms that humans can recognize objects that are distorted [35]. Previous work utilized this perception property of humans to improve the security of graphical passwords [17,18,36]. Similarly, von Zezschwitz et al. [9] distorted images using different kinds of filters (e.g., oil paint, crystallize, and pixelate) to protect privacy while browsing pictures and found that crystallized filters achieve a balance between usability and security.

These positive results motivated us to experiment with a crystallized mask to hide private information. Unlike the blackout mask, the crystallized mask maintain the context that would help users orient themselves on the interface. For example, users can still perceive the layout of the interface in the periphery (see Figure 1B). While this is an advantage from a usability perspective, the revealed context could also leak private information to attackers. For example, observers might still recognize emojis or see who is more talkative in a chat conversation, or they could identify the number of digits in a bank balance shown in an online banking app which could help them to guess the user's balance. In addition to providing context to the user, a further advantage from the security perspective is that the position of the revealed spot is not as clear to the observer compared to the blackout mask (compare Figure 1A,B).

3.1.3. Fake Text

Steganography has been anciently used to obfuscate text [37,38]. One way to do this is by using fake text to cover the original text [39]. This motivated us to experiment with the impact of using fake, yet meaningful, overlaid text (see Figure 1C) to mask the actual content on the screen. Unlike the other masks, fake text requires prior knowledge about the User Interface structure to create a similar one and replace the text with fake content. Fake text can be loaded from online random text generators (e.g., the RandomText API [40] and Spam Mimic [41]).

From a security perspective, the advantage of this method is that it does not only conceal private information, it also makes it less obvious to the observer that a protection mechanism is in place. Additionally, as seen in Figure 1C, it is more difficult for observers to identify the spot due to the lack of obvious edges. Similar to the crystallized mask, this mask maintains the layout and structure of the interface. While this mask has obvious privacy protection advantages, we were interested in investigating whether it confuses the user or reduces reading speed.

3.2. Spot Size

The second main design factor is the size of the spot through which the content is revealed. Again, the choice of the spot size is likely to influence the trade-off between usability and security; a very large spot could be less effective at protecting privacy, while a very small one could hinder usability.

Humans can perceive everything inside the fovea centralis with high acuity [42]. The diameter of the fovea is approximately 2°, which is about the size of a thumbnail when when looked at with the arm outstretched [43]. These facts influenced our design of the spot size. The two degrees of visual angle corresponded to 1.05 cm in our setup at a distance of 30 cm from the screen, which is the typical distance between the user and the smartphone [44,45]. Pilot testing indicated that reading through a spot of diameter 1.05 cm is difficult. Hence, we chose two larger sizes to experiment with: 1.9 cm and 2.85 cm (which correspond to visual angles of 3.6° and 5.4° at a 30 cm reading distance) that we call the *small spot* and *large spot*, respectively, in the remainder of this article.

4. Implementation

EyeSpot is an iOS application. In the following text, we explain our implementation of the masks and moving spot.

4.1. Interface and Masks

The masks are implemented using built-in features of iOS. *Blackout* is implemented by overlaying an empty view that is then set to black. The *Crystallized* mask uses the CICrystallize function in iOS's image processing library [46]. *Fake text* is realized by overlaying a view in a manner similar to *blackout*. The view loads an HTML template based on the shown interface and fills it with fake text. In contrast to the previous masks, fake text requires prior knowledge about the structure of the current UI. When using fake text, it is important that the fake content looks similar to the original content in terms of length. We used predetermined templates in our prototype to evaluate fake text, in which we manually selected random text to match the alignment of the original text (i.e., to match the sentence length and paragraphs). This is a limitation of our current implementation of the fake text mask; future versions of EyeSpot will dynamically generate fake content that fits the displayed UI. This can be done using existing approaches from linguistic steganography [37] that generate meaningful random text that matches the format of the original text (e.g., Spam Mimic [41]). However, our manual adaptation allowed us to get early insights into the performance of the mask, assuming an ideal fitting algorithm and text generation are already in place.

4.2. Eye Tracking and the Moving Spot

Despite recent advancements in front-facing cameras that allow eye tracking on mobile devices, achieving highly accurate tracking in real-time is still a challenge. Wood and Bulling [47] achieved a maximum accuracy of 6.88°, which is insufficient for EyeSpot. On the other hand, recent works have achieved higher accuracy levels using appearance-based gaze estimation (e.g., 2.58° [48]), yet such methods do not run in real-time. Hence, in this work, we opted for the PUPIL eye tracker [49], an external mobile eye tracker that is worn by the user (see Figure 2A). There was no constant delay, i.e., the spot was moving in real time in response to the user's gaze. However, in some cases, delays did happen and negatively influenced the study participants' experience. The delays occurred due to loose cables and network congestion. We envision that accurate and real-time eye tracking will be feasible on unmodified mobile devices in the near future, and hence, our technique will be achievable on commodity devices.

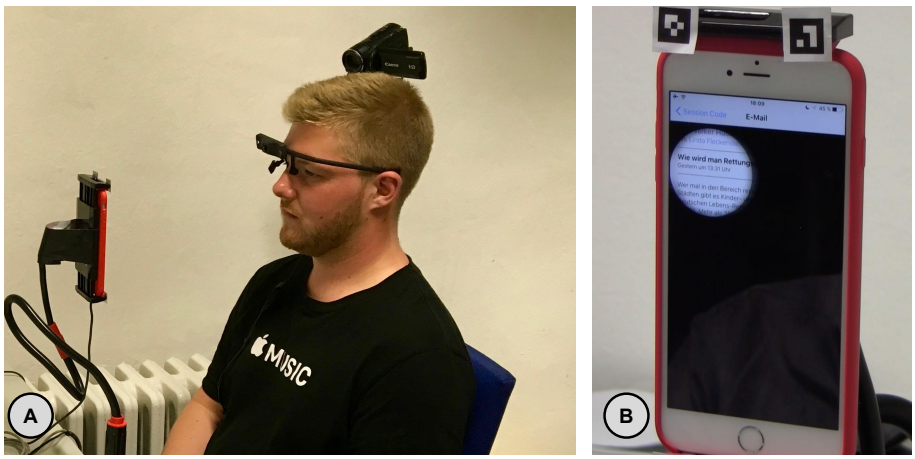


Figure 2. The figure shows the usability study's setup. The iPhone was mounted on a holder such that it was 30 cm away from the user's eyes (**A**). From the user's perspective, they could see the phone with the spot moving according to the user's gaze (**B**).

By detecting four markers at the corners of the phone (see Figure 2B) from the world-view camera of the eye tracker, we were able to accurately map the user's gaze to a point on the phone's screen. The gaze points were communicated to the mobile device via WiFi through a proxy server, which was running on a laptop connected to the eye tracker. The iOS app then moved the spot based on the user's gaze point in real-time (i.e., saccade speed). We applied smoothing to the edges of the spot to avoid sharp visible edges.

5. User Study

The goal of this study was to collect usage data to analyze the reading speed, reading comprehension, and perceived workload when using EyeSpot with different masks and spot sizes.

5.1. Participants

We invited 14 participants (3 females) aged between 18 and 28 years ($M = 24.2$, $SD = 1.8$) through mailing lists and social networks. Participants were compensated with 20 Euro e-shop vouchers. All participants had normal or corrected to normal sight. The only prerequisite was being a native speaker of the language used for the text displayed during the study.

5.2. Setup and Design

The EyeSpot app ran on an iPhone 6s Plus (5.5″, 1920 × 1080 pixels at 401 ppi). The iPhone was mounted on a holder that was adjusted based on the participant's height to be at a distance of 30 cm from the participant's eyes. This distance was determined based on previous work [44,45]. The study ran for two hours per participant. We adjusted the angle of the holder so that participants would not have to bend their neck while reading (see Figure 2).

We experimentally manipulated two independent variables: (IV1) **Mask type**, i.e., (a) blackout, (b) crystallized, (c) fake text masks, and (d) a baseline no-mask; and (IV2) **Spot size**, i.e., the aforementioned small spot and large spot.

Participants consumed two types of text—we experimented with *Chat* conversations to represent dynamic short snippets of text (as an example used in a private context) and *Email* messages to represent static larger pieces of text (as an example used in a business context). Instant messaging was identified as the most observed content type in public space, while email messages are among the most observed text-based contents [1,50]. We used chat conversations from the WhatsApp-Column [51]. As for emails, we used essays from Level 4 German Practice Exams by the American Association of Teachers of German [52]. This means that the texts were equally difficult according to language experts and that the comprehension questions of these exams could be used to assess participants' text comprehension.

In a repeated measures experiment, all participants went through all conditions in addition to a baseline (no-mask), in which the interface was not modified. This means that each participant performed 14 reading sessions: 2 content types × ((3 masks × 2 spot sizes) + no-mask). The order of conditions was counterbalanced using Latin square. We distributed the 14 texts such that each participant would read a different piece of text at each condition, e.g., if a participant read email text TEXT1 in the blackout small spot condition, then no other participant read that same text in the blackout small spot condition. This was done to reduce any possible influences of the text on our evaluation of a condition's performance.

5.3. Measures

We measured the influence on three dependent variables. First, **reading speed** was measured in words per minute (wpm) from the moment participants read the first words until they finished reading the last word in the text. Second, we evaluated **text comprehension** by asking the participants comprehension questions after each reading session. For the *chat* conversations, we asked about the sensitive dynamics of the conversation that previous work has indicated observers can infer from

shoulder surfing [1]. Namely, we asked 5-point Likert scale questions (1 = strongly agree; 5 = strongly disagree) about whether (1) the two chatting parties were in a close relationship; (2) the conversation was emotional; and (3) the amount of text written by each party was balanced. We also asked multiple choice questions about the content of the conversation, e.g., "What did the person talk about in the chat: (a) storage methods; (b) research results; (c) financial accounting; (d) computer graphics". For the *emails*, we asked the comprehension questions used in the practice exams from which we extracted the text [52]. Third, **perceived mental workload** was collected through a NASA TLX (Task Load Index) questionnaire that was filled out for each condition. Finally, we asked participants to indicate their **subjective difficulty rating** of each condition using a 5-point Likert scale.

5.4. Procedure

Participants were explained the purpose of the study and then signed a consent form. The experimenter launched the system, adjusted the holder's height, and helped the participant wear and calibrate the eye tracker. For each condition, participants were asked to read the text at a pace at which they would normally read. Afterwards, participants verbally answered the comprehension questions, filled in the NASA TLX questionnaire, and provided the subjective difficulty rating. Then, they proceeded to the next condition. We concluded the study with a semi-structured interview. The study took two hours per participant; to reduce effects of fatigue, participants were allowed to take breaks whenever they needed.

6. Limitations

While eye tracking has become feasible using front-facing cameras of mobile devices, there are still open challenges [30,34]. Therefore, we opted for using a mobile eye tracker and markers on the corners of the phone's screen. This means that the phone's position was fixed and could not be moved by the user. While this limits the ecological validity, it ensures highly accurate tracking. However, we expect that advances in eye tracking and front-facing cameras will make this concept feasible in everyday scenarios in the future. Finally, another limitation is that mobile eye tracking systems, including the one used in this work, are generally inaccurate and do not represent the user's exact gaze point due to technological and anatomical reasons [42].

6.1. Results

We analyzed 196 reading sessions (14 texts $\times$ 14 participants).

6.1.1. Reading Speed

A repeated measures ANOVA with Greenhouse–Geisser correction showed that the masks had a significant effect on reading speed $F_{1.45,18.86} = 45.99$, $p < 0.001$. Post-hoc analysis with Bonferroni correction revealed a significant difference in reading speed between no-mask $(M = 244.24, SD = 78.16)$ and (1) blackout $(M = 118.32, SD = 47.25)$, $p < 0.001$; (2) crystallized $(M = 136.23, SD = 55.77)$, $p < 0.01$; and (3) fake text $(M = 106.74, SD = 63.62)$, $p < 0.001$. No significant differences were found between the other pairs ($p > 0.05$). This means that users were fastest when using no-mask; however, we did not find any evidence that either crystallized, blackout, or fake text masks resulted in significantly faster reading speed than the other test conditions. A repeated measures ANOVA showed a significant effect of the spot size on the reading speed as well $F_{1,13} = 29.25$, $p < 0.001$. Post-hoc analysis with Bonferroni correction indicated that reading with a bigger spot $(M = 176, SD = 73.69)$ was significantly faster than reading with a smaller spot $(M = 126.57, SD = 83.67)$, $p < 0.001$. The results are summarized in Figure 3.

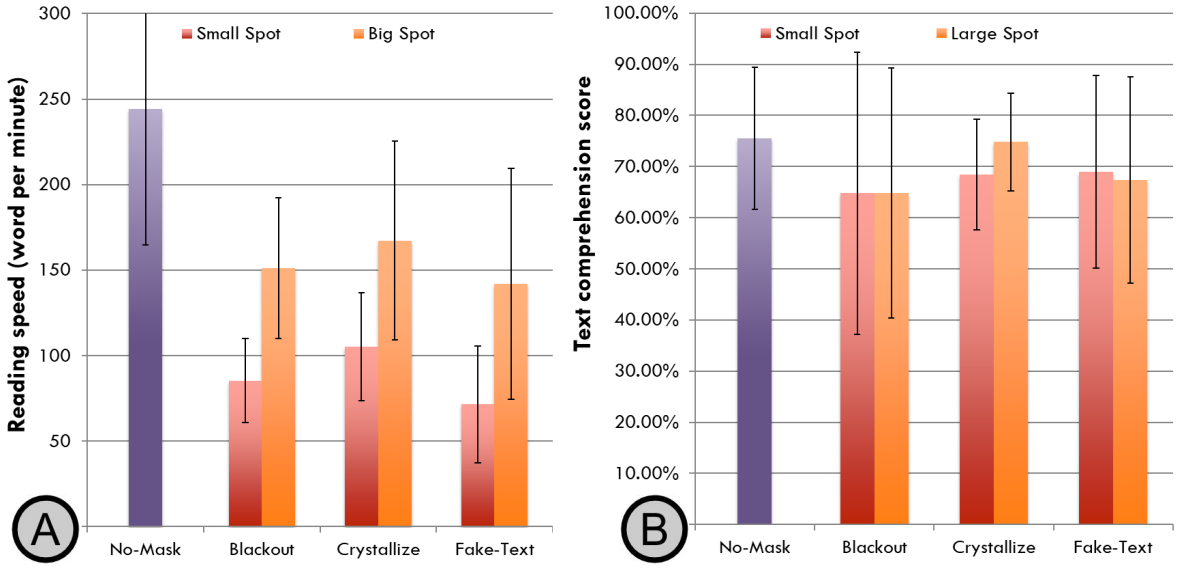


Figure 3. The masks resulted in slower reading speeds at the expense of reducing the amount of exposed content to shoulder surfers. While statistical testing only confirmed that no-mask resulted in a significantly faster reading speed, the figure suggests that reading was faster with larger spots and that users were faster when using the crystallized mask, followed by the blackout mask, and slowest when using fake text mask (**A**). As for comprehension, we could not find significant effects of masks and spot sizes on text comprehension. The differences were negligible, suggesting that EyeSpot does not have a strong impact on understanding the text (**B**).

We also found a statistically significant interaction $F_{3,39} = 6.6$, $p < 0.005$ between the mask type and the spot size ($p < 0.005$). Therefore, we compared further differences for each spot size and found that with a small spot size, reading with the crystallized mask ($M = 105.17$, $SD = 31.58$) was significantly faster than reading with the fake text mask ($M = 71.48$, $SD = 34.26$), $p < 0.05$.

6.1.2. Text Comprehension

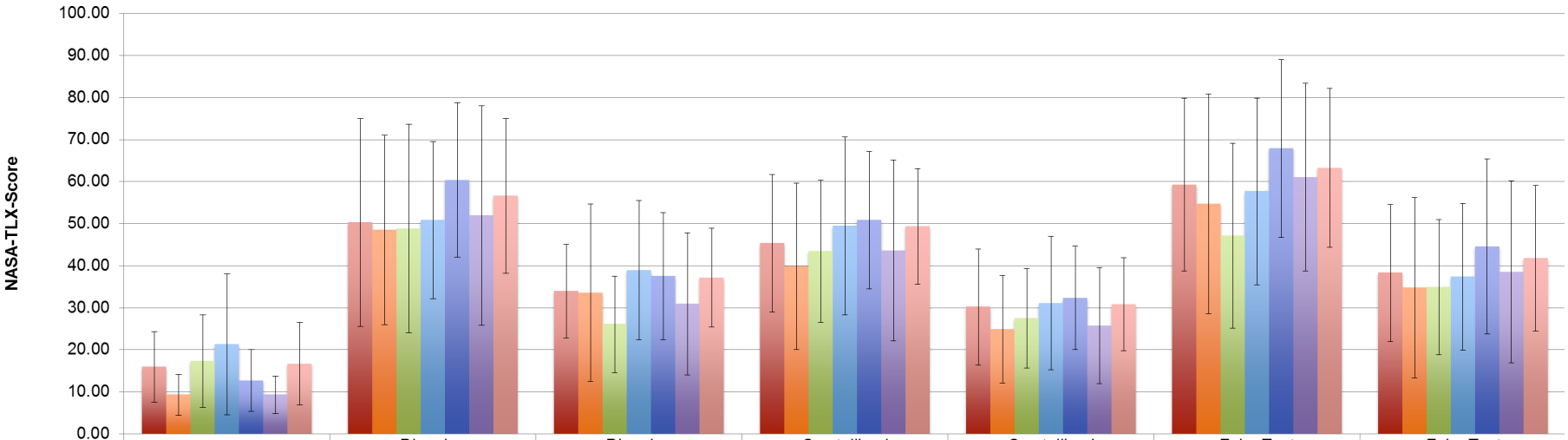
We scored participants' answers to the comprehension questions out of 100. The mean scores for each mask and spot size are shown in Figure 3. We could not find any statistically significant effect of masks nor spot sizes on text comprehension ($p > 0.05$). No interaction effects were found either ($p > 0.05$). This means there is no evidence that one mask results in lower comprehension compared to the other masks or the no-mask. Visual inspection of Figure 3 suggests that participants answered comprehension questions correctly more often with no-mask, followed by with the crystallized mask when using the large spot. Correct answers when using the other masks as well as the crystallized mask along with the small spot were all below 70%.

6.1.3. Mental Workload

While we found no significant differences ($p > 0.05$), Figure 4 suggests that the fake text mask is associated with the highest demand and that the least demanding condition is the crystallized mask when used with the large spot. The blackout mask is less demanding than fake text, but more demanding than the crystallized mask.

6.1.4. Subjective Difficulty Rating

When asked which masks and spot sizes caused the most confusion, participants agreed that the large spot and the crystallized mask were the least confusing. The blackout mask received a middle score, while fake text and small spot were perceived to be the most confusing (Table 1).



	Baseline	Dimming Small Spot	Dimming Large Spot	Crystallized Small Spot	Crystallized Large Spot	Fake Text Small Spot	Fake Text Large Spot
Mental	15.90	50.30	33.93	45.36	30.18	59.29	38.28
Physical	9.29	48.53	33.57	39.83	24.82	54.64	34.80
Temporal	17.32	48.85	26.08	43.39	27.50	47.14	34.88
Performance	21.25	50.83	38.93	49.46	31.07	57.68	37.39
Effort	12.68	60.39	37.50	50.90	32.32	67.86	44.53
Frustration	9.29	51.97	30.89	43.57	25.72	61.07	38.52
Weighted	16.69	56.60	37.17	49.37	30.84	63.23	41.71

Figure 4. The figure shows the weighted results of the NASA TLX questionnaire administered using an online tool (https://www.keithv.com/software/nasatlx/nasatlx.html). No significant differences were found, which means there is no evidence that one mask is more demanding than the other. However, visual inspection of the figure suggests a trend whereby using a large spot and a crystallized mask is associated with the least demand, while fake text mask is highly demanding with both spot sizes.

Table 1. When asked if "Reading with . . . was confusing" on a 5-point Likert scale (1 = strongly agree; 5 = strongly disagree), participants indicated that the crystallized mask was the least confusing, while fake text was the most confusing. Participants also preferred a larger spot over a smaller one.

	Spot Size		Mask			
	Small Spot	**Large Spot**	**No-Mask**	**Blackout**	**Crystallize**	**Fake Text**
Median	2.00	4.00	5.00	3.00	4.00	1.00
SD	0.92	1.00	0.00	0.19	0.92	0.69

6.1.5. Qualitative Feedback

When asked about feedback on EyeSpot, all participants mentioned they believe that the system would mitigate the risk of shoulder surfing, in particular, in crowded environments like public transport. Some felt that using the system was tiring during the study—this concerned the small spot in particular as well as the fake text condition, where participants experienced difficulty orienting themselves in the text. One participant stated that using the smartphone at eye level felt strange. We expect that as lenses of front-facing cameras advance, there will be no need to hold the mobile device in a particular way to enable eye tracking. In a real scenario, users would not spend as much time as they spent in our study (2 h) reading text using EyeSpot; therefore, we expect eye fatigue to be less prominent in real scenarios. In cases where users read for extended periods of time, the system should be capable of noticing when the user is vulnerable to shoulder surfing (e.g., by detecting potential shoulder surfers) and accordingly activate or deactivate EyeSpot). Three participants said that the gaze-based interaction in general was difficult and unfamiliar. Since gaze interaction is a new and unfamiliar concept, it is expected that users will need time to get used to it.

7. Discussion and Future Work

EyeSpot protects the user's privacy by obfuscating content that is currently not being gazed at by the user. We experimented with multiple masks and spot sizes. Visually inspecting the figures showed that the crystallized mask tends to be associated with the fastest reading time (Figure 3) and the lowest perceived workload (Figure 4) and associated level of confusion (Table 1). However, since no significant differences were found between the different masks, there is no statistical evidence that one of them is better than the other in terms of reading speed, text comprehension, and workload. While we present the crystallized mask as a promising mask, confirmation that it performs significantly better than the other masks requires further research.

7.1. Threat Model

According to a survey by Eiband et al. [1], most shoulder surfing cases involve the observer uncovering private information about the user in a passive, rather than an active, manner. That is, shoulder surfers are often opportunistic, acting out of boredom and curiosity [1]. Since this constitutes the majority of shoulder surfing situations, EyeSpot is intended to protect against this type of casual shoulder surfing attack, where an observer occasionally glances at the user's screen. EyeSpot does this by limiting the visible part of the screen. That being said, EyeSpot is not optimized for cases of intentional shoulder surfing where the attacker intentionally seeks to invade the user's privacy. In fact, if the user is observed actively and "continously" while using EyeSpot for an extended period of time, attackers at a close distance are expected to be able to follow the spot and uncover the information they are after. Furthermore, knowing what the user is attending to by observing their gaze point could potentially leak even more private information.

A simple solution could be to display multiple spots instead of one, with only one of them following the user's gaze while the others display fake content (similar to fake cursors introduced by De Luca et al. [53]). The shoulder surfer would then not know which one to follow and would more

likely observe fake content. The additional spots would not follow the user's gaze to avoid confusing the legitimate user, but could still impact usability due to the cluttered peripheral view.

7.2. Finding the Spot

Finding the moving spot when the blackout mask is employed is relatively easy, since the contrast between the spot and the surrounding (black) content is obvious (see Figure 1). On the other hand, the crystallized mask makes this distinction less clear, while the fake text mask makes it even less prominent.

7.3. Usability–Security Trade-Off

The trade-off between usability and security has been discussed in prior work [9,10,31]. The crystallized mask shows some interface features that could leak meta information about the content (e.g., who is writing more messages, how many digits make up a certain on-screen number, etc.). While the fake text and blackout masks hide contextual information, they tend to reduce reading speed and are more demanding. For example, by observing content that is masked with a crystallized mask, the shoulder surfer might be able to estimate how many digits are in a price tag or a shopping cart, and, accordingly, learn how expensive a purchase is. Similarly, several emojis can be guessed even if they are distorted, and attackers can learn which party is sending more messages (see Figure 1). In contrast, the blackout and fake text masks completely hide the screen content, hence resisting such attacks. On the downside, these masks eliminate contextual information that users often rely on to find content. This was echoed in our quantitative and qualitative results that showed that users were slower and under higher workload when using blackout and fake text masks, as opposed to the crystallized mask. The same applies for the spot size—the smaller the spot, the less content the attacker can see, at the expense of a less usable experience for the legitimate user.

Since the crystallized mask already hides content to a large degree and tends to have less impact on reading speed than other masks, we recommend it for more frequent use. The other masks could be available when additional obfuscation is needed at the expense of reading speed, e.g., when the user accesses private or sensitive content in public settings.

7.4. Applicability to other Content Types

We focused on text-based content. Namely, we experimented with chat conversations and email messages because they were identified as two content types that are often shoulder surfed [1,50] and serve as examples for a private and a business context, respectively. We did not find any significant differences across the different content types in our study. We expect that similar results would be found when using different contexts in which text is displayed on a mobile device, such as online banking apps, online news articles, and social networking apps. However, using EyeSpot when viewing photos or videos might negatively influence the user's experience since users would want to perceive the entire content and swiftly switch focus to the periphery. Therefore, this needs to be investigated in future work.

7.5. Technical Realization

Although we used an external eye tracker, advances in computing power and front-facing cameras (e.g., depth camera in iPhone X) promise highly accurate gaze estimation on commodity mobile devices [30]. A challenge that was identified in prior work is that users do not always hold smartphones in a way such that the front-facing camera can detect their faces [34]. However, this is expected to improve over time, as more gaze estimation methods continue to develop that do not rely on an entirely visible face, but rather only visible eyes [54]. This problem will further become less prominent as angles of front-facing camera lenses become wider. Another challenge is how EyeSpot can adapt to shaky environments (e.g., looking at the phone while walking or in a car). Ideally, the system would detect these situations and either enlarge the spot or disable EyeSpot completely. For

example, previous work adapted brightness if a shoulder surfer was gazing at the user's device [33]. EyeSpot can be enabled in these situations and disabled in situations in which the inertial sensors suggest that the user is walking.

8. Conclusions

We presented EyeSpot, a technique for protecting user privacy with eye tracking by hiding content that is not being gazed at with overlaid masks. We experimented with three mask types and two spot sizes. We found that using a crystallized mask along with a large spot is a promising design candidate in terms of reading speed, comprehension, and mental workload. In future work, we plan to investigate EyeSpot further and to extent the threat model presented in this work in order to optimize against threats where attackers continuously observe the user's screen. We provide a video summarizing the article, see the Supplementary Materials.

Supplementary Materials: The following are available online at http://www.mdpi.com/2414-4088/2/3/45/s1.

Author Contributions: Conceptualization, M.K., M.E. and M.Z.; Data curation, M.Z.; Formal analysis, M.K. and M.Z.; Methodology, M.K. and M.E.; Resources, H.H.; Software, M.Z.; Supervision, M.K., M.E. and H.H.; Writing – original draft, M.K. and M.E.; Writing-Review & Editing, M.K. and M.E.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eiband, M.; Khamis, M.; von Zezschwitz, E.; Hussmann, H.; Alt, F. Understanding Shoulder Surfing in the Wild: Stories from Users and Observers. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; ACM: New York, NY, USA, 2017.
2. General Data Protection Regulation. 2016. Avaliable online: ec.europa.eu/justice/data-protection/ (accessed on 8 February 2018).
3. Ali, M.E.; Anwar, A.; Ahmed, I.; Hashem, T.; Kulik, L.; Tanin, E. Protecting Mobile Users from Visual Privacy Attacks. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, WA, USA, 13–17 September 2014; ACM: New York, NY, USA, 2014; pp. 1–4.
4. Eiband, M.; von Zezschwitz, E.; Buschek, D.; Hußmann, H. My Scrawl Hides It All: Protecting Text Messages Against Shoulder Surfing With Handwritten Fonts. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; ACM: New York, NY, USA, 2016; pp. 2041–2048.
5. De Luca, A.; Harbach, M.; von Zezschwitz, E.; Maurer, M.E.; Slawik, B.E.; Hussmann, H.; Smith, M. Now You See Me, Now You Don'T: Protecting Smartphone Authentication from Shoulder Surfers. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; ACM: New York, NY, USA, 2014; pp. 2937–2946.
6. von Zezschwitz, E.; De Luca, A.; Brunkow, B.; Hussmann, H. SwiPIN: Fast and Secure PIN-Entry on Smartphones. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; ACM: New York, NY, USA, 2015; pp. 1403–1406.
7. Yang, Y.; Clark, G.D.; Lindqvist, J.; Oulasvirta, A. Free-Form Gesture Authentication in the Wild. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; ACM: New York, NY, USA, 2016; pp. 3722–3735.
8. Zhou, H.; Tearo, K.; Waje, A.; Alghamdi, E.; Alves, T.; Ferreira, V.; Hawkey, K.; Reilly, D. Enhancing Mobile Content Privacy with Proxemics Aware Notifications and Protection. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; ACM: New York, NY, USA, 2016; pp. 1362–1373.
9. von Zezschwitz, E.; Ebbinghaus, S.; Hussmann, H.; De Luca, A. You Can't Watch This!: Privacy-Respectful Photo Browsing on Smartphones. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; ACM: New York, NY, USA, 2016; pp. 4320–4324.

10. De Luca, A.; Denzel, M.; Hussmann, H. Look into My Eyes!: Can You Guess My Password? In Proceedings of the 5th Symposium on Usable Privacy and Security, Mountain View, CA, USA, 15–17 July 2009; ACM: New York, NY, USA, 2009; pp. 7:1–7:12.
11. Liu, D.; Dong, B.; Gao, X.; Wang, H. Exploiting Eye Tracking for Smartphone Authentication. In Proceedings of the 13th International Conference on Applied Cryptography and Network Security, New York, NY, USA, 2–5 June 2015.
12. Probst, G. Analysis of the Effects of Privacy Filter Use on Horizontal Deviations in Posture of VDT Operators. Master's Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, 2000.
13. Brudy, F.; Ledo, D.; Greenberg, S.; Butz, A. Is Anyone Looking? Mitigating Shoulder Surfing on Public Displays Through Awareness and Protection. In Proceedings of the International Symposium on Pervasive Displays, Copenhagen, Denmark, 5–6 June 2014; ACM: New York, NY, USA, 2014.
14. Zhou, H.; Ferreira, V.; Alves, T.; Hawkey, K.; Reilly, D. Somebody Is Peeking!: A Proximity and Privacy Aware Tablet Interface. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; ACM: New York, NY, USA, 2015; pp. 1971–1976.
15. Kinjo, H.; Snodgrass, J.G. Does the Generation Effect Occur for Pictures? *Am. J. Psychol.* **2000**, *113*, 95–121. [CrossRef] [PubMed]
16. Denning, T.; Bowers, K.; van Dijk, M.; Juels, A. Exploring Implicit Memory for Painless Password Recovery. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; ACM: New York, NY, USA, 2011; pp. 2615–2618.
17. Harada, A.; Isarida, T.; Mizuno, T.; Nishigaki, M., A User Authentication System Using Schema of Visual Memory. In Proceedings of the Second International Workshop on Biologically Inspired Approaches to Advanced Information Technology (BioADIT 2006), Osaka, Japan, 26–27 January 2006; Ijspeert, A.J., Masuzawa, T., Kusumoto, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 338–345.
18. Hayashi, E.; Dhamija, R.; Christin, N.; Perrig, A. Use Your Illusion: Secure Authentication Usable Anywhere. In Proceedings of the 4th Symposium on Usable Privacy and Security, Pittsburgh, PA, USA, 23–25 July 2008; ACM: New York, NY, USA, 2008; pp. 35–45.
19. Khamis, M.; Bulling, A.; Alt, F. Tackling Challenges of Interactive Public Displays Using Gaze. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*; ACM: New York, NY, USA, 2015; pp. 763–766.
20. Best, D.S.; Duchowski, A.T. A Rotary Dial for Gaze-based PIN Entry. In Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications, Charleston, SC, USA, 14–17 March 2016; ACM: New York, NY, USA, 2016; pp. 69–76.
21. Cymek, D.H.; Venjakob, A.C.; Ruff, S.; Lutz, O.H.M.; Hofmann, S.; Roetting, M. Entering PIN codes by smooth pursuit eye movements. *J. Eye Mov. Res.* **2014**, *7*. [CrossRef]
22. De Luca, A.; Weiss, R.; Drewes, H. Evaluation of Eye-gaze Interaction Methods for Security Enhanced PIN-entry. In Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces, Adelaide, Australia, 28–30 November 2007; ACM: New York, NY, USA, 2007; pp. 199–202.
23. Forget, A.; Chiasson, S.; Biddle, R. Shoulder-surfing Resistance with Eye-gaze Entry in Cued-recall Graphical Passwords. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; ACM: New York, NY, USA, 2010; pp. 1107–1110.
24. Kumar, M.; Garfinkel, T.; Boneh, D.; Winograd, T. Reducing Shoulder-surfing by Using Gaze-based Password Entry. In Proceedings of the 3rd Symposium on Usable Privacy and Security, Pittsburgh, PA, USA, 18–20 July 2007; ACM: New York, NY, USA, 2007; pp. 13–19.
25. Sakai, D.; Yamamoto, M.; Nagamatsu, T.; Fukumori, S. Enter Your PIN Code Securely!: Utilization of Personal Difference of Angle Kappa. In Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications, Charleston, SC, USA, 14–17 March 2016; ACM: New York, NY, USA, 2016; pp. 317–318.
26. Alt, F.; Mikusz, M.; Schneegass, S.; Bulling, A. Memorability of Cued-Recall Graphical Passwords with Saliency Masks. In Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia, Rovaniemi, Finland, 12–15 December 2016; ACM: New York, NY, USA, 2016.
27. Bulling, A.; Alt, F.; Schmidt, A. Increasing the Security of Gaze-based Cued-recall Graphical Passwords Using Saliency Masks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; ACM: New York, NY, USA, 2012; pp. 3011–3020.

28. Kinnunen, T.; Sedlak, F.; Bednarik, R. Towards Task-independent Person Authentication Using Eye Movement Signals. In Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, Austin, TX, USA, 22–24 March 2010; ACM: New York, NY, USA, 2010; pp. 187–190.
29. Zhang, Y.; Hu, W.; Xu, W.; Chou, C.T.; Hu, J. Continuous Authentication Using Eye Movement Response of Implicit Visual Stimuli. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*; ACM: New York, NY, USA, 2018.
30. Khamis, M.; Alt, F.; Bulling, A. The Past, Present, and Future of Gaze-enabled Handheld Mobile Devices: Survey and Lessons Learned. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, Barcelona, Spain, 3–6 September 2018; ACM: New York, NY, USA, 2018.
31. Khamis, M.; Alt, F.; Hassib, M.; von Zezschwitz, E.; Hasholzner, R.; Bulling, A. GazeTouchPass: Multimodal Authentication Using Gaze and Touch on Mobile Devices. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; ACM: New York, NY, USA, 2016; pp. 2156–2164.
32. Song, C.; Wang, A.; Ren, K.; Xu, W. EyeVeri: A Secure and Usable Approach for Smartphone User Authentication In Proceedings of the IEEE International Conference on Computer Communication (INFOCOM'16), San Francisco, CA, USA, 10–15 April 2016; pp. 1 – 9.
33. Lian, S.; Hu, W.; Song, X.; Liu, Z. Smart privacy-preserving screen based on multiple sensor fusion. *IEEE Trans. Consum. Electron.* **2013**, *59*, 136–143. [CrossRef]
34. Khamis, M.; Baier, A.; Henze, N.; Alt, F.; Bulling, A. Understanding Face and Eye Visibility in Front-Facing Cameras of Smartphones used in the Wild. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018.
35. Gregory, R.L. Knowledge in perception and illusion. *Phil. Trans. R. Soc. Lond. B* **1997**, *352*, 1121–1127. [CrossRef] [PubMed]
36. Wang, Z.; Jing, J.; Li, L. Time Evolving Graphical Password for Securing Mobile Devices. In Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security, Berlin, Germany, 4–8 November 2013; ACM: New York, NY, USA, 2013; pp. 347–352.
37. Bennett, K. *Linguistic Steganography: Survey, Analysis, and Robustness Concerns for Hiding Information in Text*; Purdue University: West Lafayette, IN, USA, 2004.
38. Shirali-Shahreza, M.H.; Shirali-Shahreza, M. A New Approach to Persian/Arabic Text Steganography. In Proceedings of the 5th IEEE/ACIS International Conference on Computer and Information Science and 1st IEEE/ACIS International Workshop on Component-Based Software Engineering, Software Architecture and Reuse (ICIS-COMSAR'06), Honolulu, HI, USA, 10–12 July 2006; pp. 310–315.
39. Petitcolas, F.A.P.; Anderson, R.J.; Kuhn, M.G. Information Hiding–A Survey. *Proc. IEEE* **1999**, *87*, 1062–1078. [CrossRef]
40. Davies, D. RandomText API. 2014. Avaliable online: http://www.randomtext.me/ (accessed on 8 December 2016).
41. Wayner, P. Spam Mimic. 2000. Avaliable online: http://www.spammimic.com/ (accessed on 8 December 2016).
42. Majaranta, P.; Bulling, A. Eye Tracking and Eye-Based Human—Computer Interaction. In *Advances in Physiological Computing*; Fairclough, S.H., Gilleade, K., Eds.; Springer: London, UK, 2014; pp. 39–65.
43. Duchowski, A.; Vertegaal, R. *Eye-Based Interaction in Graphical Systems: Theory & Practice*; ACM: New York, NY, USA, 2000.
44. Ho, J.; Pointner, R.; Shih, H.C.; Lin, Y.C.; Chen, H.Y.; Tseng, W.L.; Chen, M.Y. EyeProtector: Encouraging a Healthy Viewing Distance when Using Smartphones. In Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, Copenhagen, Denmark, 24–27 August 2015; ACM: New York, NY, USA, 2015; pp. 77–85.
45. Huang, M.X.; Li, J.; Ngai, G.; Leong, H.V. ScreenGlint: Practical, In-situ Gaze Estimation on Smartphones. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; ACM: New York, NY, USA, 2017; pp. 2546–2557.
46. Core Image Filter Reference. 2018. Avaliable online: https://developer.apple.com/library/content/documentation/GraphicsImaging/Reference/CoreImageFilterReference/index.html (accessed on 29 January 2018).

47. Wood, E.; Bulling, A. EyeTab: Model-based Gaze Estimation on Unmodified Tablet Computers. In Proceedings of the Symposium on Eye Tracking Research and Applications, Safety Harbor, FL, USA, 26–28 March 2014; ACM: New York, NY, USA, 2014; pp. 207–210.
48. Krafka, K.; Khosla, A.; Kellnhofer, P.; Kannan, H.; Bhandarkar, S.; Matusik, W.; Torralba, A. Eye Tracking for Everyone. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 2176–2184.
49. Kassner, M.; Patera, W.; Bulling, A. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, WA, USA, 13–17 September 2014; ACM: New York, NY, USA, 2014; pp. 1151–1160.
50. Ponemon Institute. *Global Visual Hacking Experimental Study: Analysis*; Ponemon Institute: Traverse City, MI, USA, 2016.
51. Süddeutschen-Zeitung Jetzt.de—WhatsApp Column. 2018. Avaliable online: http://www.jetzt.de/tag/whatsapp-kolumne (accessed on 29 January 2018).
52. American Association of Teachers of German. Practice Exams. 2016. Avaliable online: http://www.aatg.org/?page=NGEPracExams (accessed on 29 January 2018).
53. De Luca, A.; von Zezschwitz, E.; Pichler, L.; Hussmann, H. Using Fake Cursors to Secure On-screen Password Entry. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; ACM: New York, NY, USA, 2013; pp. 2399–2402.
54. Bérard, P.; Bradley, D.; Nitti, M.; Beeler, T.; Gross, M. High-quality Capture of Eyes. *ACM Trans. Graph.* **2014**, *33*, 223. [CrossRef]