Stealth Updates of Visual Information by Leveraging Change Blindness and Computational Visual Morphing

SHUNICHI KASAHARA and KAZUMA TAKADA, Sony Computer Science Laboratories, Inc., Japan

We present an approach for covert visual updates by leveraging change blindness with computationally generated morphed images. To clarify the design parameters for intentionally suppressing change detection with morphing visuals, we investigated the visual change detection in three temporal behaviors: visual blank, eye-blink, and step-sequential changes. The results showed a robust trend of change blindness with a blank of more than 33.3 ms and with eye blink. Our sequential change study revealed that participants did not recognize changes until an average of 57% morphing toward another face in small change steps. In addition, changes went unnoticed until the end of morphing in more than 10% of all trials. Our findings should contribute to the design of covert visual updates without consuming users' attention by leveraging change blindness with computational visual morphing.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; *Interaction devices*; *Displays and imagers;*

Additional Key Words and Phrases: Change blindness, eye tracking, blink detection, morphing, generative adversarial network

ACM Reference format:
Shunichi Kasahara and Kazuma Takada. 2021. Stealth Updates of Visual Information by Leveraging Change Blindness and Computational Visual Morphing. *ACM Trans. Appl. Percept.* 18, 4, Article 23 (October 2021), 17 pages.
https://doi.org/10.1145/3486581

1 INTRODUCTION

Pervasive displays and wearable displays such as AR glasses are effective for keeping information updated at all times. However, it is necessary to constantly refresh the display to keep it updated or for the users to be constantly context-aware. Updating the visual presentation after it appears on the display will always cause a change in perception and draw the user's attention, which will distract users from the task or action they were originally focused on [1]. In other words, contextual information might make changes more apparent. How then can we update visual information without drawing attention?

In this work, we propose a covert visual information update that does not draw attention by leveraging visual change blindness to suppress the change perception. Change blindness is a well-known phenomenon where humans fail to detect obvious changes even if they are on display. This occurs when a visual transient is suppressed or obscured in such a way that the user's attention cannot be drawn to it [29, 32, 36]. Since change detection

23

This work was supported by JST Moonshot R&D Program Grant No. JPMJMS2013.
Authors' address: S. Kasahara and K. Takada, Sony Computer Science Laboratories, Inc., 3-14-13, Higashigotanda, Shinagawa-ku, Tokyo, Japan 141-0022; emails: {kasahara, k-takada}@csl.sony.co.jp.

1544-3558/2021/10-ART23 $15.00
https://doi.org/10.1145/3486581

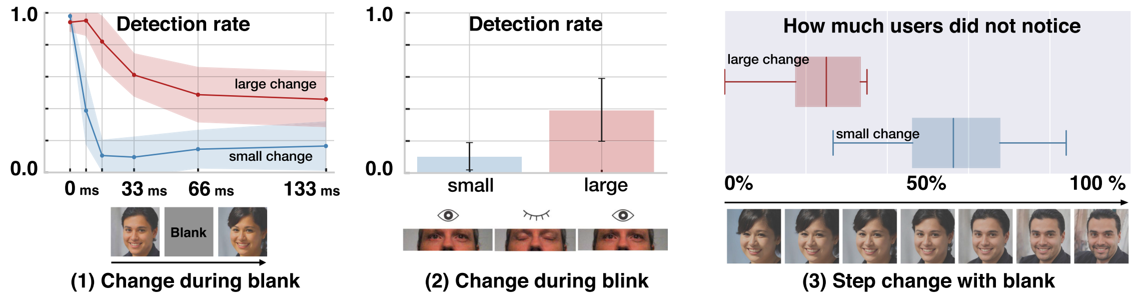

Fig. 1. Highlight of our investigation about change detection in visual blank, eye-blink, and step changes. The results showed a robust trend of change blindness with (a) a blank of more than 33.3 ms and (b) eye blink with small morphed image change. Moreover, users did not notice the morphing changes until (c) more than halfway through morphing toward another face.

without a visual transient depends on a comparison of visual impression and visual short-term memory, previous studies have shown that change blindness does not occur with large image changes [2]. In our approach, we computationally generate a series of interpolated images and gradually apply the changes to update the visual information while suppressing the perception of changes between before and after the visual update.

To design such a stealth visual information update, we first need to clarify how the amount of visual change and the temporal design affect the user's detectability of visual change. Therefore, the purpose of this article is to acquire the design parameters for intentionally suppressing change detection through a series of change blindness experiments. We performed computational morphing to generate gradual visual changes in three types of visual information (face, natural images, and text) and investigated the occurrence of change blindness (i.e., user's detectability of visual change) in three temporal characteristics: visual blank, blink, and sequential changes.

In the first experiment, we investigated the change detection with different visual blank durations and amounts of visual change in a one-shot change detection task. The results showed that the rate of change blindness stabilized for visual blanks of 33 ms or longer for all visual types. As the blank time of 33 ms is shorter than the typical duration of a human blink, we performed a similar change detection task for eye blink in our second experiment and found similar trends to the visual blank. On the basis of these results, we conducted a sequential change experiment to investigate a situation where interpolated visual changes are applied sequentially and gradually. As one of the highlighted results, we found that by applying interpolated facial image change during blanks, we could perform up to an average of 57% facial morphing to another face without the user noticing, and in more than 10% of the trials, participants did not recognize the change even after the face was completely altered. Our results and findings may suggest new methods for updating information without attracting users' attention, by leveraging visual change blindness with computational visual information morphing.

2 RELATED WORK

2.1 Change Blindness

When a visual object on a display suddenly changes (i.e., when updating text on a UI, displaying notifications, or changing pictures in ads), users typically notice it, because the change produces a sudden perceptual transient that draws attention to itself [29]. However, when transients do not occur or are obscured, visual changes can be remarkably difficult to detect, which causes change blindness [29, 32, 36]. Change blindness has been studied in various experimental paradigms to occlude transient changes [14, 18, 26, 32]. For example, in the flicker paradigm [32], participants view a cycle of two versions of images sequentially with a blank screen inserted in between each image. By manipulating the image presentation time and various blank times (40, 80, 160, and 320 ms), Rensink et al. showed that change blindness is affected by the volatility of the visual processing [31] To separate the effects of long-term memory from the repeated presentation in the flicker paradigm, they also introduced a one-shot paradigm where an image is briefly presented, followed by a brief blank or mask, and then followed by a second image [30].

In addition to manipulation of the visual sequence, Grimes found that when a portion of an image was clearly changed in synchronization with the timing of the saccade (large eye-movement), 50% of observers failed to notice the change [14]. Moreover, O'Regan et al. reported that when the image was changed at the moment of the blink, observers failed to noticed changes even when they were gazing at the location of the change [18].

As display technology continues to evolve, the frame rate (frequency of visual updates) has increased [28], along with high-speed eye-tracking, which enables the production of short visual interruption even when synchronized with eye movement. Since off-the-shelf devices allow us to control high-frame-rate displays, change blindness is no longer merely an experimental phenomenon in a lab but can be deployed in our daily environment. However, it is unclear how the length of a visual interruption influences the detectability of changes in the image as the drawing speed of displays increases, which is one of our interests in this study.

2.2 Applications of Change Blindness

Since computer systems often rely on visual displays to convey information to users, the prevention of change blindness is important for human-computer interface design [23, 39]. In process monitoring and control systems, where it is important to detect changes by monitoring visual information, change blindness must be considered [12] in the context of tasks performed by naval command and control system personnel [11]. Davies et al. reported the change blindness phenomenon in handheld mobile devices and its impact on mobile UI design [8]. Brock et al. investigated change blindness in a proximity-aware mobile interface and discussed the merits of either exploiting change blindness or mitigating it for mobile proximity-aware interfaces [6]. While change blindness is typically considered a phenomenon that should be mitigated to avoid potential failures in human-computer interaction, there is also research that seeks to take advantage of un-noticed changes. In a virtual environment, for example, redirected walking with head motion-induced change blindness [37], blink-induced change blindness [21], and saccade-based change blindness [4] have been proposed. Sebastian et al. [24] proposed covert changes in a virtual reality scene by using eye-tracking to determine when and where an image can be changed. There is also a remarkable example of exploiting change blindness as well as inattention blindness [35]. By exploiting change blindness in saccades, auto page-turning enables continuous reading without manual control of the scrolling [42]. Goddard et al. reported that smooth illuminating color changes occur in the absence of transient signals, which means that color changes can go unnoticed without visual disruption and the change is not required to be very gradual [13].

For covert updates in visual information systems, it is necessary to design visual change based on the nature of the change perception. Ma et al. proposed a computational model to quantify the degree of blindness between an image pair by considering the saliency of the image [22]. Additionally, even if a visual transient is interrupted, a user may notice large changes based on visual impression and visual short-term memory [2]. Hollingworth et al. reported that semantically inconsistent changes can be detected faster than consistent ones in the change detection task [15]. This suggests that the semantic properties of the morphed images should be maintained smoothly enough to be consistent during the interpolation of visual information.

3 GENERATION OF MORPHING VISUAL INFORMATION

In our study, considering the information we usually see on screen media, we focused on three categories of visual information: face, non-face (natural landscapes, animals, food, and everyday objects), and text. To generate a morphing visual sequence for each category, we used StyleGAN2 [17] for the face image generation, BigGAN [5] for the non-face images, and BERT for text [10]. Figure 2 shows examples of generated images in each category. Here, we describe how the intermediate visuals were generated.

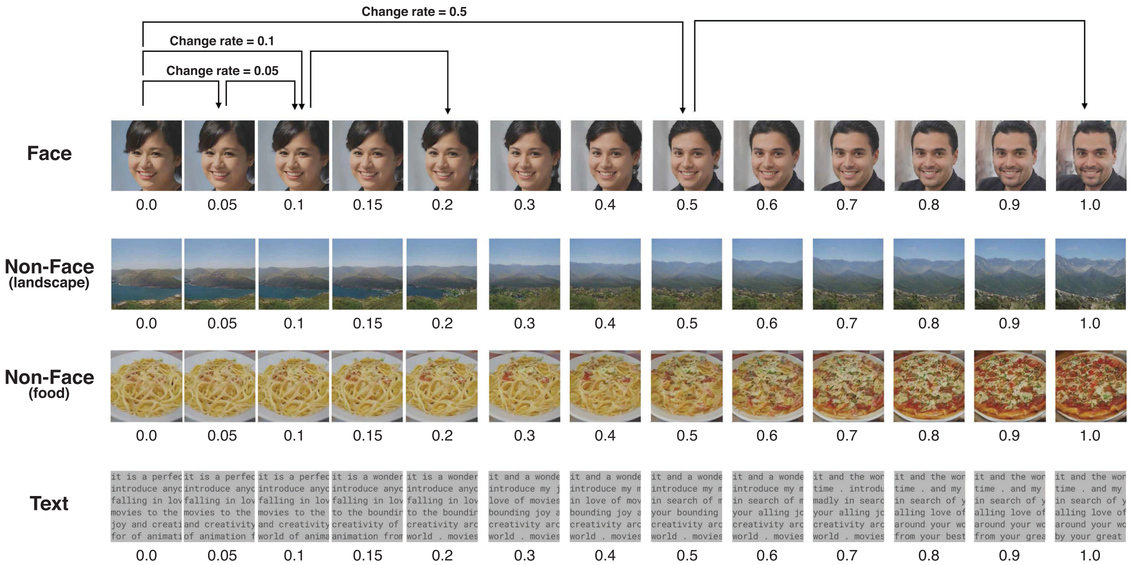

Fig. 2. Examples of image morphing in face, non-face, and text categories. Note that images of text have been cropped. For the actual image, see Figure 3.

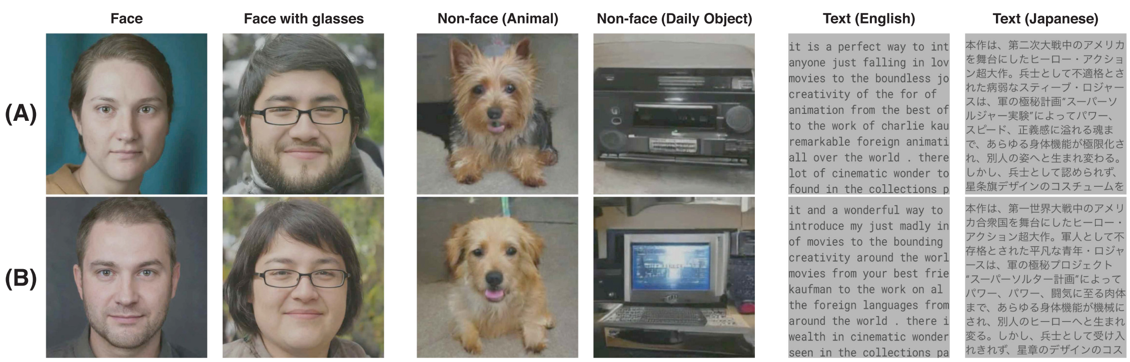

Fig. 3. Example of generated images for face, non-face, and text categories. (A) morphing rate 0.0; (B) morphing rate 0.5.

3.1 Face Morphing

By utilizing latent space in a **generative adversarial network (GAN)**, morphing faces even with high-level semantic features can be achieved [16]. In our experiments, we used this interpolation of faces in the latent space to generate the morphing of faces. To consider the cross-race effect [38], we chose a pre-learned model of StyleGAN that was trained with the **Flickr-Faces-HQ Dataset (FFHQ)**.[1] This model can generate a wide variety of faces of any race, gender, or age. In our experiments, all presented faces were unique and new for participants.

With the pre-trained StyleGAN2 model [17], we generated image morphing by interpolating two vectors in the latent space, which we randomly generated with a standard normal distribution. To avoid unexpected

[1] https://github.com/NVlabs/ffhq-dataset.

visual clues in the experiments, we excluded obvious fragmentation or obstacles on the face, children, images containing other faces, and extremely conspicuous backgrounds. We then randomly selected two vectors from these sets of vectors and generated images along with the interpolation steps (defined as the change rate) of 0.05, 0.1, and 0.5; therefore, the number of steps in morphing was 20, 10, and 2, respectively (Figure 2).

To clarify the distribution of images, we investigated the Euclidean distance of pairs of generated images in the latent space. The average distance between the "From" and "To" images, which were randomly determined with a normal distribution, was 31.7 (SD = 0.93). In addition, the average distance of interpolated images in change rates 0.05, 0.1, and 0.5 was 2.99 (SD = 0.38), 5.97 (SD = 0.77), and 29.87 (SD = 3.84), respectively. Since these are small SD values, we can roughly observe the correlation between the interpolation change rate and the distances in the latent space. These numbers might reflect visual impressions between each interpolation rate; for example, with 0.05, 0.1 interpolation change rates, two faces look similar and identical, respectively, but with the 0.5 interpolation change rate, two faces look like a different person.

3.2 Non-face Image Morphing

To investigate visual scenes and objects other than faces, we chose the pre-trained model of BigGAN [5] (256 × 256 pixels), which was trained using Imagenet [9]. To generate interpolated images for our experiments, we selected approximately 15 labels, each corresponding to landscapes, animals, food, and everyday objects, for a total of 60 labels, and then chose interpolation pairs between similar label categories to avoid image artifacts. Then, we randomly selected pairs of two vectors in the latent space and generated images along with the interpolation change rate with 0.05, 0.1, and 0.5 (Figure 2). We also investigated the Euclidean distance of pairs of generated images in the latent space and found that between "From" image and "To" image it was 1.715 (SD, 0.0897) and between interpolated images in 0.05, 0.1, and 0.5 it was 0.089 (SD = 0.0043), 0.190 (SD = 0.0099), and 1.694 (SD = 0.114), respectively.

3.3 Text Morphing

Unlike faces and natural images, text interpolation cannot be produced as pixel interpolation. Therefore, we implemented word-based text interpolation by using **Bidirectional Encoder Representations from Transformers (BERT)**, a neural network that is often used for natural language processing [10]. Using masked language modeling, we extracted nouns, adjectives, and verbs from a sentence by morphological analysis and replaced these original words with the estimated words so that the original sentences could be slightly updated. We defined a 100% morphing change as a sentence in which all words were replaced, and then performed text morphing by gradually replacing the replaceable words in an order-randomized manner. Figure 2 shows examples of the text morphing. In the following experiments, we used a pre-trained model trained in the native language of the participants, and the source text for morphing was randomly extracted from the text datasets of open-source movie review articles with 150–300 characters. We generated the series of 60-step morphed sentences and then converted them into images with a monospaced font as visual stimulus. We note here, Fig 3 includes an example of text image in English. In the native language of the participants, line breaks occur character by character. In the case of English, the stimulus image is cropped so that each line starts with the first letter of the word.

4 STUDY 1: ONE-SHOT CHANGE DETECTION WITH BLANK

In the first experiment, our aim was to investigate the detection rate in different durations of visual blank occlusion with different degrees of visual change. We implemented a change detection task in a one-shot paradigm [30] where participants were instructed to report whether they noticed the change.

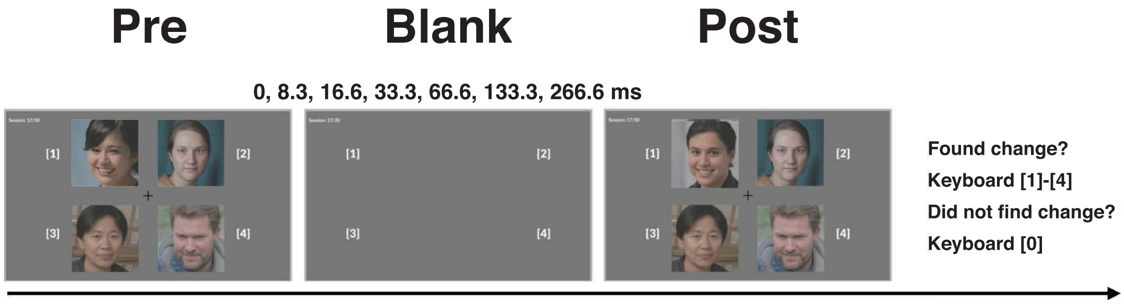

Fig. 4. Study 1 paradigm. One-shot change detection with visual blank.

4.1 Apparatus

We conducted the study using a keyboard and a high-frame-rate display[2] running at a resolution of 1,920 × 1,080 pixels with a refresh rate of 240 Hz. Considering inconsistent frame intervals when rendering the frame drop, we measured the frame drop at the moment of blanks and excluded data when the frame drop occurred. Note that the occurrence of a frame drop was 0.57 % in the total data of study 1. These devices were connected to a Windows 10 computer with an Nvidia GPU (Geforce RTX2070 SUPER) and CPU (AMD Ryzen 3960X 3.8 GHz). The study took place in a semi-lit room. Participants viewed the monitor from a distance of 70 cm, with a 40-degree horizontal field of view for the full screen.

4.2 Task

Participants started a trial by pressing "9" on the keyboard, and then four images (PRE) were displayed on the screen with a 50% gray color background. The size of each image was 10.3 deg with a 70-cm display distance. The 50% gray blank was inserted with six different durations or no blank, followed by the display of four images, but one of the four was changed (POST). Participants were instructed to press the corresponding key "1"–"4" when they detected a change. They were also asked to look at the fixation throughout all trials. Here, our main focus here was the explicit detection of change. Considering that previous studies have shown that undetectable stimulus can influence subsequent decisions in forced choice task [20]. To avoid any subliminal influence of blinded images, we also asked participants to press "0" when they could not detect any change among the four images.

4.3 Materials

For generating two levels of visual change rate (CR = 0.1 and 0.5) from the generated morphing sequences described in Section 3, we extracted one image for PRE with the randomly selected morphing rate $MR_{pre} = random(0.0–0.5)$ and another for POST with the morphing rate $MR_{post} = MR_{pre} + CR$. To avoid the effect of the participants' memory, three other un-changed images were also extracted from the morphing image sequences with a randomly selected morphing point. In our experiment, we compressed the gradation of each RGB channel of the color image to between 20% and 80% and set the blank frame color as 50% gray to make sure we could induce a reliable visual blank. By using a high-frame-rate camera (1,000 Hz), we ensured that images became invisible in all grade-compressed color channels in the minimum blank duration (8.3 ms).

4.4 Procedure

Participants gave their informed consent prior to the experiment and were provided with detailed instructions on how to perform the experimental task. To avoid experiment artifacts stemming from a low engagement or

[2]ASUS ROG SWIFT PG258Q (ASUS, Taipei, Taiwan, 24.5-in.).

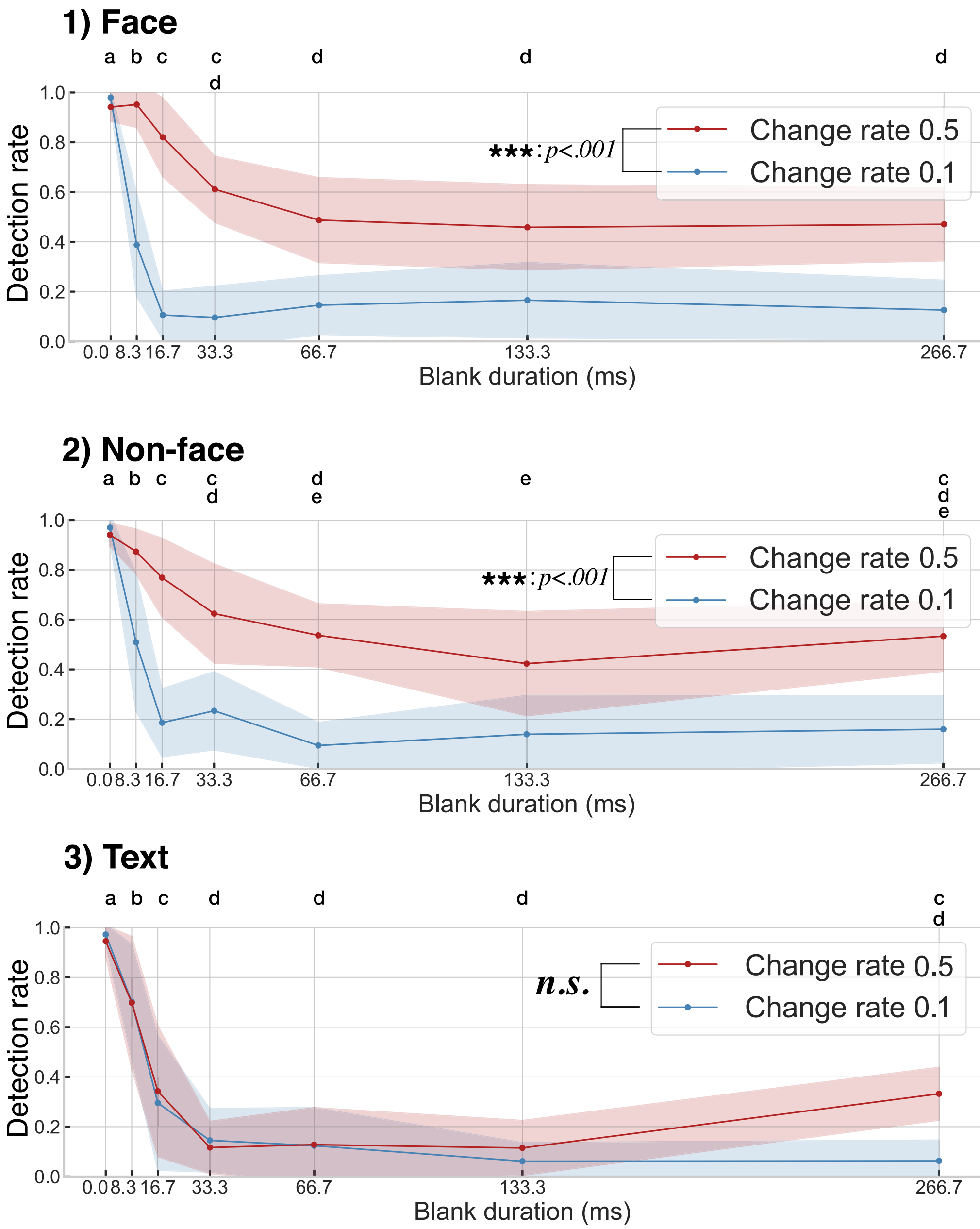

Fig. 5. Results for study 1: One-shot change detection with visual blank. We investigated two levels of visual change rate and seven different lengths of blank duration [0.0 (no blank), 8.3, 16.6, 33.3, 66.6, 133.3, and 266.6 ms]. Compact letter displays (a–e) letters show which blank groups are not significantly different in the post hoc analysis.

automatic response, we also added dummy trials where one of the four images was changed to an obviously different object (e.g., a face to a banana). We then confirmed that all participants responded to them correctly. We designed a within-subjects experiment with two independent variables featuring seven different lengths of blank duration [0.0 (no blank), 8.3, 16.6, 33.3, 66.6, 133.3, and 266.6 ms] and two levels in the change rate

[0.1 and 0.5] of visual change. Participants performed the tasks in three blocks (face, non-face, and text image). In each of the three image categories, 12 trials with two levels of change rate featuring seven blank durations and five dummy trials were performed in randomized order. The order of each category was counter-balanced. The total time per participant, including instructions, experiments, breaks, post-questionnaires, and debriefing, was 30 min.

4.5 Participants

We recruited 12 participants (two self-identified as female, ten as male) from a local institute. Their ages ranged from 19 to 33 (mean = 25.8, SD = 5.0) and all had normal or corrected to normal vision.

4.6 Results

We collected participants' responses for change detection (we treated wrong answers and "0" (non-detected) as incorrect) for each trial. In 14 possible combinations (of seven blank durations and two change rates of morphing), we acquired 12 samples of data points and then calculated the change detection rate [0.0–1.0]. Figure 5 shows the results for change detection rate in different blank durations for both levels of change rate. To investigate the significant difference in seven blank durations (0 ms, 8.3–266.7 ms) and the two levels of change rate, we conducted two-way repeated ANOVA analysis on the change detection rate in the three visual categories. In the face image category, we found a significant effect of both the change rate ($F(1, 154) = 6.40$, $p < 0.001$) and the blank duration ($F(6, 154) = 1.48$, $p < 0.001$), and an interaction ($F(6, 154) = 0.35$, $p < 0.001$) in the change detection rate. The non-face image category had similar results: significant effects of both variables were found (change rate: $F(1, 154) = 4.97$, $p < 0.001$; blank duration: $F(6, 154) = 1.42$, $p < 0.001$), as well as an interaction ($F(6, 154) = 0.212$, $p < 0.001$). In contrast, in the text category, while the effect of the blank duration rate was significant ($F(6, 154) = 2.73$, $p < 0.001$), the effect of the change rate was not significant ($F(1, 154) = 0.08$, $p = 0.093$), but significant effects were found in the interaction ($F(6, 154) = 0.065$, $p = 0.047$). These results demonstrate that the duration of the visual blank affects the detection rate in all image categories and that the change rate matters in (1) face and (2) non-face images but not in (3) the text category. Subsequent TukeyHSD post hoc tests revealed statistical significance in the blank durations. Figure 5 shows the details with **compact letter displays (CLDs)** indicating which blank duration groups are not significantly different by TukeyHSD post hoc test.

4.7 Discussion

As we see in Figure 5, there was a sudden drop in the detection rate around the blank duration of >16 ms in all three categories with change rate 0.1. We assume that some level of visual transients are still perceptible in visual blanks of less than 16 ms, so participants easily detected the changes without any cognitive process. In contrast, in blanks of more than 33 ms, we observed a static trend of the detection rate in the face and non-face image categories. More interestingly, the detection rates in change rates 0.1 and 0.5 were converged to a different values, which suggests that the amount of change matters in terms of the cognitive comparison (Figure 5). This converged trend of the detection rate provides an important implication for designing stealth updates; namely, in the case of applying covert changes for a non-focused visual object, it is not necessary to apply a longer visual blank. For example, a visual blank of 16 ms (one frame on a 60-Hz display) was sufficient for a small change, which is equivalent to the 0.1 change rate in our stimuli (Figure 2). This indicates that we can perform stealth updating without adding a longer distractive visual blank than necessary. In the text visual category (Figure 2—Text), the trend of detection rates in different blanks did not reveal any significant difference between the two levels of change rate. We presume that even if the number of changes in the sentence became larger, the amount of visual change to the image was not large.

In study 1, we have clarified the trends of detection rate corresponding to the blank duration and showed that the occurrence of change blindness is robust with blanks of more than 33.3 ms. The fact that 33 ms is shorter

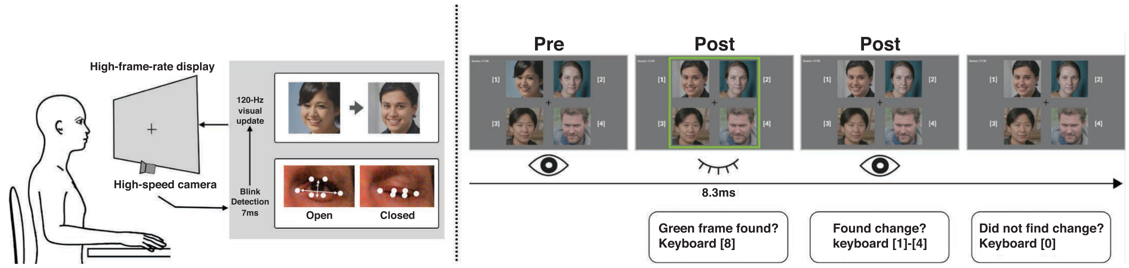

Fig. 6. Study 2 setup and paradigm.

than the typical duration of a human blink [25] led us to consider how these findings might be compatible to change blindness with eye blink, which does not require an unnecessary perceptible visual distraction.

5 STUDY 2: ONE-SHOT CHANGE DETECTION WITH BLINK

The main objective of this study was to investigate whether the trend of change blindness in eye blink was consistent with visual blank (study 1). In study 2, we utilized the same apparatus as study 1 except for the eye-blink detection, which is synchronized with a high-frame-rate display. Participants in study 2 were the same as those in study 1. We also used the same material for visual stimulus.

5.1 Blink Detection

Eye blinks can be classified as voluntary blinks [19], reflex blinks, and spontaneous blinks, which are done without conscious control. We mainly focus on spontaneous blinks in this study. The average duration of the down-phase closing and the up-phase opening in a spontaneous blink is 100 ms and 220 ms, respectively [25]. During blinks, visual perception is suppressed in a similar way to saccadic suppression. This suppression of visual input lasts for a maximum of 30–40 ms before the upper eyelid covers the eyes and gradually resumes the sensitivity at 100–200 ms after the beginning of an eye blink [33, 40, 41]. To leverage the maximum suppression during spontaneous blinks, we need to detect blinks and update the visual as quickly as possible.

Therefore, we built a low-latency blink detection system synchronized with a 120-Hz display frame rate to update visuals during blinking with low latency. The high-frame-rate camera (Ximea MQ013CG-ON) captures a color image at 140 fps with an exposure of 5 microsecond. This camera is connected to the computer via USB-3.0. The system tracks facial landmark points using dlib (68 facial landmarks[3]) and then calculates the **eye aspect ratio (EAR)**, which describes the level of the eye opening. By identifying a rapid drop of the EAR value, the system can detect blinking in real time [7]. These recognition processes are completed within 7.0 ms and are designed to be synchronized with the 120-fps drawing process.

5.2 Procedure and Task

Participants followed the same procedure as study 1. After pressing "9" to start, they looked at the fixation and four images (PRE) that were displayed, the same as in study 1. When the blink of a participant was detected, one of the four images was changed (POST). Participants were instructed to respond to the detected change by pressing "1"–"4." If there was no response from a participant 2 s after the PRE-POST changes during the blink of an eye, then the text dialog "[0] for not found" was displayed and the participant was prompted to answer whether or not the change had been detected. We designed it this way because participants did not know when to respond if change blindness occurred.

[3]http://dlib.net/.

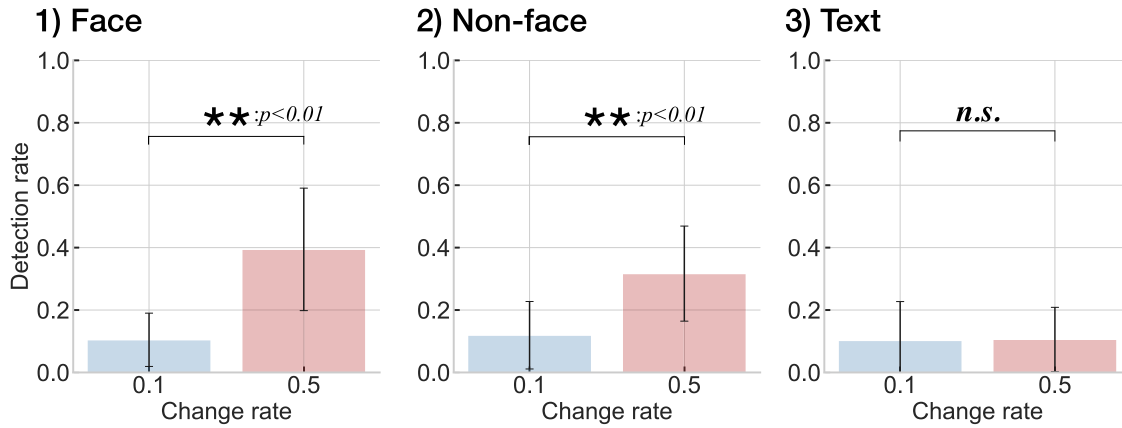

Fig. 7. Results for study 2: One-shot change detection with human blink.

Due to individual differences in the shape of the eyes or the manner of blinking, there is a possibility of confusing whether a participant's response to detecting a change is due to the participant's ability or to the system accuracy of blink detection. Therefore, to exclude this confusion, we added a single rendering frame (8.3 ms, one frame in 120 Hz) green color visual stimulus only at the first rendering frame of the POST image, as shown in Figure 6. Participants were instructed to "press [8] whenever you see green," without knowing the purpose of the question or what caused the green stimuli to appear. We then ignored any responses with this [8] key response. The detection rate of the green frame was approximately 12% throughout the experiments. Note that we confirmed in a preliminary experiment that we can usually detect a single frame green visual stimulus in 120-Hz display rendering.

Participants performed the tasks in three blocks (face, non-face, and text image). The order of each category was counter-balanced.

5.3 Result

We collected the participant responses in the two levels of change rate of morphing, acquired 15 samples of data points, and then calculated the change detection rate [0.0–1.0]. Figure 7 shows the results for the change detection rate in eye blink for both levels of change rate in three visual categories. The average of the change detection rate was 0.39 (SD = 0.20) and 0.10 (SD = 0.09) in face, 0.32 (SD = 0.15) and 0.12 (SD = 0.11) in non-face, and 0.11 (SD = 0.10) and 0.10 (SD = 0.12) in text, at change rates 0.5 and 0.1, respectively.

For the analysis, all data were assessed for normality using the Shapiro-Wilk test and we found that the detection rate in one of the two conditions or both conditions was not normally distributed. Wilcoxon signed-rank tests between the two levels of change rate (0.1 and 0.5) indicated significant differences in the detection rate for the face ($p < 0.01$) and non-face ($p < 0.01$) visual categories, but not for the text ($p = 0.674$). Our statistical analysis shows that the change rate of visual stimulus affects the detection rate in (a) face and (b) non-face images but not in (c) the text category.

5.4 Discussion

Similar to the results in study 1, the detection rate with change rate 0.1 was quite low (less than 10%) in general. Participants stated that they did not notice when the image was changed, even if they were closely concentrating on the task. This demonstrates the robustness of the change blindness with our frame-synchronized blink detection. The trends of the detection rate were similar to the converged detection rate where the duration of the blank was more than 33 ms (Figure 5). While it is not appropriate to compare these two results statistically

Table 1. Experimental Design of Study3, Sequential Change Detection Task

(1) step change			(2) gradual change	
(a) face	(b) non-face	(c) text	(a) face	(b) non-face
change rate = 0.05			duration = 50 s	
change rate = 0.1			duration = 25 s	
change rate = 0.5			—	

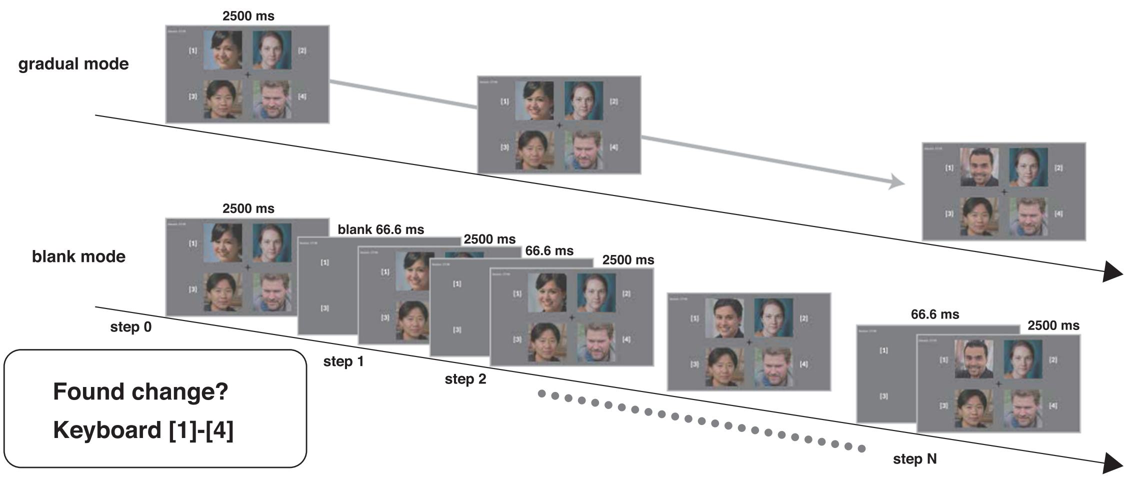

Fig. 8. Study 3 paradigm: Step change and gradual change.

due to their slightly different paradigms, we can expect similar behaviors of the blink-induced change blindness to visual blanks of more than 33.3 ms.

The similarity of the results in studies 1 and 2 indicates that we can apply the blank paradigm to simulate the phenomena that can occur in eye blink for more complex situations. The results of both studies have shown that an average 10% of changes can be detected with small visual changes (morphing rate of 10%). Here, a new question arises: in the case of a change applied continuously in steps, not just a one-time change, to what extent can visual changes be stealthily applied through a series of changes?

6 STUDY 3: SEQUENTIAL CHANGE DETECTION TASK WITH BLANK

To investigate how covert updates can be applied continuously, we conducted a change detection study examining sequential visual updates with different levels of change step (or speed) and gradual change.

6.1 Task

We designed a within-subjects experiment with one independent variable (change rate) featuring three levels in (1) step change and two levels in (2) gradual change.

For step change, we changed one of the four visual stimuli every blank with the designated change step, from the initial image to the target image, step by step throughout each blank. Blanks occurred periodically in each trial sequence. This was a modified version of the flicker paradigm [32], where the target visual sequentially changes over blanks. Participants were instructed to press the corresponding key “1”–“4” whenever they noticed which pictures had been changed. We set the blank duration to 66.6 ms, since study 1 revealed that the converged detection rate for change blindness with a blank of more than 33.3 ms. To enable closer estimation of a blink-induced change blindness, the blanks were scheduled to occur once every 2.5 s. This assumes the timing of a natural human blink, which occurs approximately 24 times per minute [34]. For step change, we compared three levels of the change rate CR: 0.05, 0.1, and 0.5, with 1.0 as the 100% change. Therefore, in each condition, 20 (50 s), 10 (25 s), and 2 (5 s) changes are required for 100% completion of a change. Note that we added one more change rate in addition to the conditions in studies 1 and 2, to investigate how a smaller change rate would affect the detection of change in sequential changes.

In addition to the step changes, we investigated the gradual change of images [36] as another option for continuous change. For a better comparison with step change, we set two gradual speeds, 25 and 50 s for a 100% completion of change, corresponding to the 0.1 and 0.05 change rate conditions in the step change. Note that we exclude the 5-s condition (corresponding to a 0.5 change rate), since the visual motion was clearly obvious. We generated a gradually changing video with 60 keyframes displayed at 60 fps, with pixel interpolation performed between keyframes.

The main aim of this study is to investigate the extent to which visual morphing updates go unnoticed. We defined a "stealth score" metric for this evaluation, as follows. In step change mode, the progress of change toward its completion is described as $CR \times N$, where N is the number of steps that have been taken. In cases where a participant detects the change at step n in a trial, we can assume the participant was unaware of changes until step $n-1$. Therefore, the 'stealth score' to indicate the extent to which a user is unaware of changes is defined as $CR \times (n-1)$. For example, when a participant detects the change at step 6 in the 0.1 change rate (CR) condition, the stealth score is $0.1 \times (6-1) = 0.5$. In cases where a participant does not find changes throughout the whole trial, the stealth score is 1.0, as a stealth update has been completed without change detection.

For the gradual change mode, similarly, the "stealth score" is described as a value of the time of detection divided by the length of the gradually changing movie. For example, when a participant detects the change 20 s after the task started (and the gradual video starts to proceed) in the 50-s speed condition, the stealth score is calculated as 20 s/50 s = 0.4. Again, this value is aligned with the one in step change.

6.2 Participants

We recruited 18 participants online. Their ages ranged from 22 to 27 (mean = 23.7, SD = 1.3), with four self-identified as female and 14 as male. None of them had participated in study 1 or 2. All had normal or corrected to normal vision.

6.3 Setup for Online Experiment

Participants were instructed to visit a designated web site and run the experimental program implemented in p5js[4] on a web browser. We had them set up a normal lighting environment in the room. At the beginning of the experiment, participants distanced themselves from the screen as specified so that the task window was at a 40-degree viewing angle. In the experimental program, we confirmed that the graphic rendering was executed at 60 Hz. The blank duration in each trial was measured, and if we observed any blank duration of 66.6 ms +/− 10 ms or more, we excluded the corresponding trial data.

6.4 Procedure

All sessions began with four practice trials to demonstrate the procedure with one highly noticeable change in both the step and gradual change. At the beginning of each trial, participants were instructed to focus on the fixation in the first blank in the step change mode and the first 2,500 ms in the gradual mode with no change. This was to make sure the participants could concentrate. After the fixation disappeared, participants were allowed to move their eyes as they pleased during the trial.

[4] https://p5js.org/.

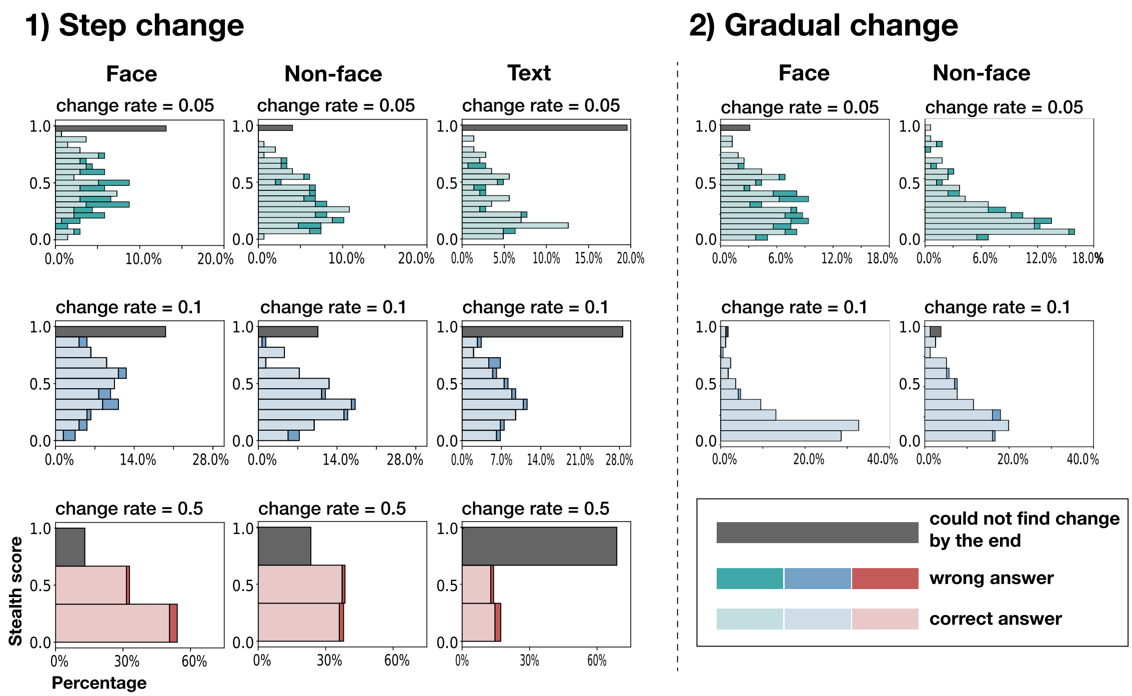

Fig. 9. Response distribution of stealth score of all participants in study 3. Pale color represents correct detection of change and solid color represents wrong detection. Gray indicates that participants never detected anything.

We conducted five separate blocks for each participant: (1-a) face, (1-b) non-face, and (1-c) text category in blank mode, and (2-a) face, (2-b) non-face in gradual mode (see Table 1). For the three blocks in step change, ten trials with three levels of change rate (0.05, 0.1, and 0.5) were performed in randomized order, for a total of 30 trials. In the two blocks for gradual change, ten trials with two levels of change rate (25 and 50 s) were performed in randomized order. The order of each mode and category was counter-balanced.

6.5 Result

Figure 9 depicts the raw data for each condition with all participants' responses. Note that the stealth score in the step change is plotted with intervals of the change rate. We excluded 7% of the collected data due to frame drop. We also excluded the data from three (in step change) and one (in gradual change) of the participants, because it had a high frequency of frame drop (more than 50%). In the end, we acquired data of 15 participants for the step change and 17 participants for the gradual change. We did not observe the effects of learning in the trials, as we had found no correlation between the trial index and participants' responses in stealth score ($R^2 < 0.001$).

6.5.1 Step Change. For the step change, we calculated the average stealth score of each participant. Figure 10 shows the distribution of average stealth scores from each participant's response in each condition. In the face category, the mean of the average stealth score of each participant was 0.560. (SD = 0.233), 0.575 (SD = 0.160), and 0.295 (SD = 0.227) at the change rate of 0.05, 0.1, and 0.5, respectively. Pairwise Friedman's tests ($p < 0.05$) were conducted to compare the effect of the change rate CR on the stealth score in the change rate 0.05, 0.1, and 0.5 conditions. There was a significant difference among the distributions of the three categories of images (based on Friedman's test, $\chi^2(2) = 12.4$, $p < 0.01$). Subsequent post hoc Steel-Dwass tests revealed a significant difference (p < 0.05) between the conditions of [0.05, 0.5] and [0.1, 0.5] but not in [0.05, 0.1] in the face visual. In

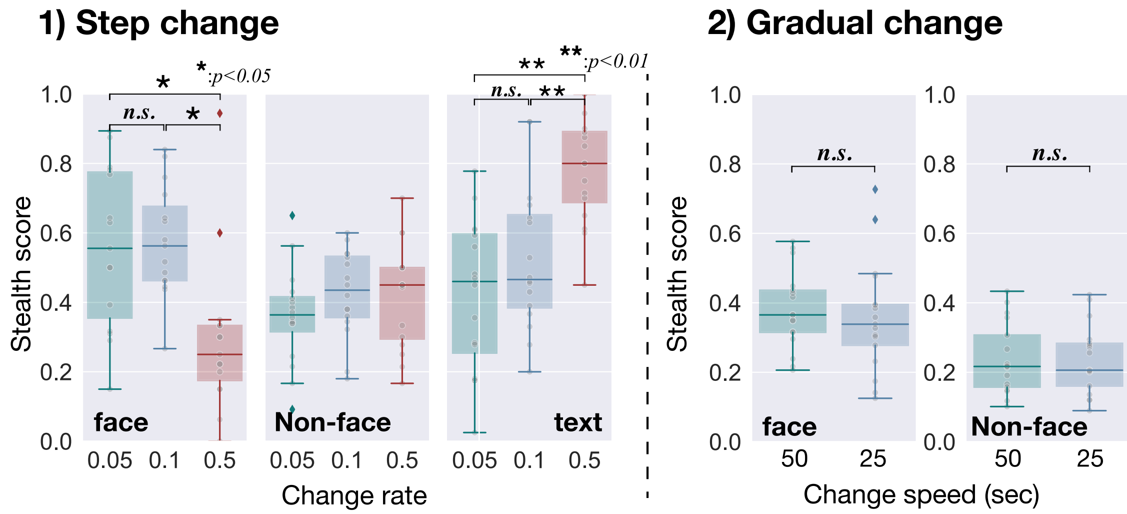

Fig. 10. Average stealth score (how much users did not notice) in each condition in blank mode. For example, stealth score = 0.5 means participants did not notice a change until a 50 % change in sequential morphing.

the non-face category, the stealth score of each participant was 0.360 (SD = 0.139), 0.425 (SD = 0.126), and 0.421 (SD = 0.151) at the change rate of 0.05, 0.1, and 0.5, respectively. A one-way repeated measures ANOVA was conducted to compare the effect of the change rate on the stealth score in the 0.05, 0.1, and 0.5 conditions. There was no significant effect of the change rate on the stealth score at the $p < 0.05$ level for any of the three conditions [$F(2, 26) = 1.4785$, $p = 0.2465$] in the non-face visual. Interestingly, in the text visual category, the average values of the stealth score showed opposite trends: 0.420 (SD = 0.218), 0.525 (SD = 0.209), and 0.783 (SD = 0.156) for the change rate of 0.05, 0.1, and 0.5, respectively. A one-way repeated measures ANOVA was conducted to compare the effect of the change rate on the stealth score in the change rate 0.05, 0.1, and 0.5 conditions. There was a significant effect of the change rate on the stealth score at the $p < 0.05$ level for all three conditions [$F(2, 26) = 27.0615$, $p < 0.01$]. Subsequent Tukey-Kramer post hoc tests revealed statistical significance in the [0.05, 0.5] and [0.1, 0.5] pair-wise conditions ($p < 0.01$) but not in the [0.05, 0.1] conditions in the text visual.

6.5.2 Gradual Change. A paired-samples t-test was conducted to compare the average stealth scores in the 25- and 50-s speed conditions for both image categories. There were no significant differences in the 25-s condition (M = 0.358, SD = 0.161) or 50-s condition (M = 0.378, SD = 0.113); $t(16)) = 0.514$), $p = 0.613$) in the face category, nor for the non-face category (25-s condition: M = 0.236, SD = 0.100; 50-s condition: M = 0.237, SD = 0.102, $t(16)) = 0.046$, $p = 0.964$).

6.6 Discussion

As we found in the statistical analysis, in each visual category, we observed different characteristics. In the face category, the trends of the small change rate (0.05 and 0.1) and large change rate (0.5) were consistent with those in study 1. While this would simply seem to suggest that larger visual changes are more likely to be detected earlier, it is notable that there was no difference between the change rates of 0.05 and 0.1. We presume this is because the stealth score also depends on the number of changes required to obtain a certain morphing progress, i.e., more change steps happen and there is a greater probability that participants would be able to detect a changing visual object. From this point of view, the change rate for each step should be designed with a consideration of this trade-off.

In the non-face category, we did not find any statistically consistent trends among the three levels of change rate. To understand why, we conducted post-interviews to determine what kind of strategy the participants took. Their answers clarified several strategies used to find changes for different types of images, such as finding added or removed objects, motion with high-frequency images (pasta, pizza), and the border between objects and background. As P3 stated, *"Once I found a change in a food image, I knew what to focus on when finding [changes] in similar images."* We observed analogous comments from other participants, indicating that once participants confidently found changes in an image, they established a strategy to find changes in similar images. As we can observe in Figure 9, participants were unaware of changes by the end of morphing in more than 12 % of the trials. This suggests that even if they could find cues or strategies, finding changes was still difficult even when participants were intentionally looking for them. We also conjecture that updating visual representations in user's visual memory eventually kept users unaware. Previous literature showed that our lack the ability to simultaneously represent two complete visual representations as one of the mechanism of change blindness [3]. Since participants were allowed to look for changes while moving their eyes, four visuals (including the updated one) were always exposed to them. This might also overwrite the visual representations in visual memory along with the visual updating, which means participants did not acquire cues for detection even when the total change was quite large.

Particularly in the non-face category, three participants mentioned they were aware of certain visual artifacts that are typical visual errors in machine learning generation. Since we utilized the pre-trained model of BigGAN [5], which generates images of relatively low quality, those visual errors were useful clues for detection.

As another interesting finding, in the text category, we observed that, in direct contrast to the trend in the face category, the average stealth scores were higher even in the large change rate. As we found in studies 1 and 2 that the change rate did not have any effect on the detection rate in text visuals, we interpret this trend in study 3 as stemming from the effect of the number of change steps; in other words, the number of changes will influence the stealth score.

Although it is not feasible to statistically compare the blank and gradual modes, we observed lower trends of stealth score for the gradual changes in both visual categories. This is supported by comments from all participants like *"Gradual mode was easier than blank mode."* As we designed the speed of the gradual change to correspond to the speed of the step changes, this suggests that observers are less sensitive to change with blanks and blinks as compared to gradual changes.

6.7 Limitation

We here note that there may be some bias in our experiment, in terms of the diversity of participants [27]. Even, to the best of our knowledge, no gender differences in change blindness have been reported, it is necessary to consider the diversity of participants not only in change blindness but also to investigate more universal properties including memory and attention.

7 CONCLUSION

We have presented an approach for unaware visual updates by leveraging change blindness with computationally generated morphed images. To investigate the design parameters for stealthy updates of visual information, we conducted three change detection experiments for visual blank, eye-blink, and step-sequential changes with a generated morphed visual. Our findings clarified the fundamental nature of stealth updates. Designers and researchers should be able to use our findings as a basis for designing more practical visual information layouts for covert visual updates while taking attention and gaze into account. We envision that this study will contribute to the design of covert visual updates without consuming users' attention by leveraging change blindness with computational visual morphing.

ACKNOWLEDGMENT

We thank Daisuke Tajima for his advice on the statistical analysis.

REFERENCES

[1] Piotr D. Adamczyk and Brian P. Bailey. 2004. If not now, when? The effects of interruption at different moments within task execution. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'04)*. Association for Computing Machinery, New York, NY, 271–278. https://doi.org/10.1145/985692.985727

[2] Nora Andermane, Jenny M. Bosten, Anil K. Seth, and Jamie Ward. 2019. Individual differences in change blindness are predicted by the strength and stability of visual representations. *Neurosci. Conscious.* 2019, 1 (Jan. 2019), niy010.

[3] M. W. Becker, H. Pashler, and S. M. Anstis. 2000. The role of iconic memory in change-detection tasks. *Perception* 29, 3 (2000), 273–286.

[4] Benjamin Bolte and Markus Lappe. 2015. Subliminal reorientation and repositioning in immersive virtual environments using saccadic suppression. *IEEE Trans. Vis. Comput. Graph.* 21, 4 (Apr. 2015), 545–552.

[5] Andrew Brock, Jeff Donahue, and Karen Simonyan. 2018. Large scale GAN training for high fidelity natural image synthesis. Retrieved from https://arxiv.org/abs/1809.11096.

[6] Michael Brock, Aaron Quigley, and Per Ola Kristensson. 2018. Change blindness in proximity-aware mobile interfaces. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'18)*. ACM Press, New York, New York, 1–7.

[7] J. Cech and T. Soukupova. 2016. Real-time eye blink detection using facial landmarks. In *Proceedings of the 21st Computer Vision Winter Workshop*.

[8] Thomas Davies and Ashweeni Beeharee. 2012. The case of the missed icon: Change blindness on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'12)*. Association for Computing Machinery, New York, NY, 1451–1460.

[9] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 248–255.

[10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. Retrieved from arXiv:1810.04805.

[11] Joseph DiVita, Richard Obermayer, William Nugent, and James M. Linville. 2004. Verification of the change blindness phenomenon while managing critical events on a combat information display. *Hum. Factors* 46, 2 (2004), 205–218.

[12] Paula J. Durlach. 2004. Change blindness and its implications for complex monitoring and control systems design and operator training. *Human–Computer Interaction* 19, 4 (Dec. 2004), 423–451.

[13] Erin Goddard and Colin W. G. Clifford. 2013. A new type of change blindness: Smooth, isoluminant color changes are monitored on a coarse spatial scale. *J. Vis.* 13, 5 (Apr. 2013), 20.

[14] John Grimes. 1996. On the failure to detect changes in scenes across saccades. *Perception* 339 (1996), 89–110.

[15] Andrew Hollingworth and John M. Henderson. 2000. Semantic informativeness mediates the detection of changes in natural scenes. *Vis. Cogn.* 7, 1–3 (Jan. 2000), 213–235.

[16] Tero Karras, Samuli Laine, and Timo Aila. 2020. A style-based generator architecture for generative adversarial networks. *IEEE Trans. Pattern Anal. Mach. Intell.* PP (January 2020). https://doi.org/10.1109/TPAMI.2020.2970919

[17] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2019. Analyzing and improving the image quality of StyleGAN. Retrieved from https://arxiv.org/abs/1912.04958.

[18] J. Kevin O'Regan, Heiner Deubel, James J. Clark, and Ronald A. Rensink. 2000. Picture changes during blinks: Looking without seeing and seeing without looking. *Vis. Cogn.* 7, 1–3 (Jan. 2000), 191–211.

[19] Kyung-Ah Kwon, Rebecca J. Shipley, Mohan Edirisinghe, Daniel G. Ezra, Geoff Rose, Serena M. Best, and Ruth E. Cameron. 2013. High-speed camera characterization of voluntary eye blinking kinematics. *J. R. Soc. Interface* 10, 85 (Aug. 2013), 20130227.

[20] Cédric Laloyaux, Christel Devue, Stéphane Doyen, Elodie David, and Axel Cleeremans. 2008. Undetected changes in visible stimuli influence subsequent decisions. *Conscious. Cogn.* 17, 3 (Sept. 2008), 646–656.

[21] Eike Langbehn, Frank Steinicke, Markus Lappe, Gregory F. Welch, and Gerd Bruder. 2018. In the blink of an eye: Leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. *ACM Trans. Graph.* 37, 4 (Aug. 2018), 1–11.

[22] Li-Qian Ma, Kun Xu, Tien-Tsin Wong, Bi-Ye Jiang, and Shi-Min Hu. 2013. Change blindness images. *IEEE Trans. Vis. Comput. Graph.* 19, 11 (Nov. 2013), 1808–1819.

[23] Gabriela Mancero, William Wong, and Paola Amaldi. 2007. Looking but not seeing: Implications for HCI. In *Proceedings of the 14th European Conference on Cognitive Ergonomics Invent! Explore! (ECCE'07)*. ACM Press, New York, New York, 167.

[24] Sebastian Marwecki, Andrew D. Wilson, Eyal Ofek, Mar Gonzalez Franco, and Christian Holz. 2019. Mise-unseen: Using eye tracking to hide virtual reality scene changes in plain sight. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST'19)*. Association for Computing Machinery, New York, NY, 777–789.

[25] Yoshiko Nakamura, Jumpei Matsuda, Kazutaka Suzuki, Haruyoshi Toyoda, Naotoshi Hakamata, Takasumi Shimamoto, and Shigeru Kinoshita. 2008. [Measurement of spontaneous blinks with a high-speed blink analyzing system]. *Nihon Ganka Gakkai Zasshi* 112, 12 (Dec. 2008), 1059–1067.

[26] J. K. O'Regan, R. A. Rensink, and J. J. Clark. 1999. Change-blindness as a result of "mudsplashes." *Nature* 398, 6722 (March 1999), 34.

[27] Tabitha C. Peck, Laura E. Sockol, and Sarah M. Hancock. 2020. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE Trans. Visual. Comput. Graph.* 26, 5 (2020), 1945–1954. https://doi.org/10.1109/TVCG.2020.2973498

[28] Christian H. Poth, Rebecca M. Foerster, Christian Behler, Ulrich Schwanecke, Werner X. Schneider, and Mario Botsch. 2018. Ultrahigh temporal resolution of visual presentation using gaming monitors and G-Sync. *Behav. Res. Methods* 50, 1 (Feb. 2018), 26–38.

[29] Ronald A. Rensink. 2002. Change detection. *Annu. Rev. Psychol.* 53 (2002), 245–277.

[30] Ronald A. Rensink. 2005. Change blindness. In *Neurobiology of Attention*, Laurent Itti, Geraint Rees, and John K. Tsotsos (Eds.). Academic Press, Burlington, 76–81.

[31] Ronald A. Rensink, J. Kevin O'Regan, and James J. Clark. 2000. On the failure to detect changes in scenes across brief interruptions. *Vis. Cogn.* 7, 1–3 (Jan. 2000), 127–145.

[32] Ronald A. Rensink, J. Kevin O'Regan, and James J. Clark. 1997. To see or not to see: The need for attention to perceive changes in scenes. *Psychol. Sci.* 8, 5 (Sept. 1997), 368–373.

[33] W. H. Ridder, 3rd and A. Tomlinson. 1997. A comparison of saccadic and blink suppression in normal observers. *Vision Res.* 37, 22 (Nov. 1997), 3171–3179.

[34] Harvey Richard Schiffman. 1990. *Sensation and Perception: An Integrated Approach*, 3rd ed. Wiley, 555.

[35] D. J. Simons and C. F. Chabris. 1999. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception* 28, 9 (1999), 1059–1074.

[36] D. J. Simons, S. L. Franconeri, and R. L. Reimer. 2000. Change blindness in the absence of a visual disruption. *Perception* 29, 10 (2000), 1143–1154.

[37] E. A. Suma, S. Clark, D. Krum, S. Finkelstein, M. Bolas, and Z. Warte. 2011. Leveraging change blindness for redirection in virtual environments. In *Proceedings of the IEEE Virtual Reality Conference*. 159–166.

[38] James W. Tanaka, Markus Kiefer, and Cindy M. Bukach. 2004. A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition* 93, 1 (Aug. 2004), B1–9.

[39] D. Alexander Varakin, Daniel T. Levin, and Roger Fidler. 2008. Unseen and unaware: Implications of recent research on failures of visual awareness for human-computer interface design. *Hum.-Comput. Interact.* 19, 4 (Dec. 2008), 389–422.

[40] F. C. Volkmann. 1986. Human visual suppression. *Vision Res.* 26, 9 (1986), 1401–1416.

[41] F. C. Volkmann, L. A. Riggs, and R. K. Moore. 1980. Eyeblinks and visual suppression. *Science* 207, 4433 (Feb. 1980), 900–902.

[42] Andrew D. Wilson and Shane Williams. 2018. Autopager: Exploiting change blindness for gaze-assisted reading. In *Proceedings of the ACM Symposium on Eye Tracking Research and Applications*, Vol. 2. ACM, New York, NY, 1–5.

Received August 2021; accepted August 2021