Visual Quality Optimization for Privacy Protection Bar-based Secure Image Display Technique

Sanghyun Park[1] and Sang-ug Kang[2]
[1] Department of Computer Science, Sangmyung University
Seoul , 20 Hongjimun 2-gil Jongno-gu - Republic of Korea
[e-mail: commam6@gmail.com]
[2] Department of Computer Science, Sangmyung University
Seoul , 20 Hongjimun 2-gil Jongno-gu - Republic of Korea
[e-mail: sukang@smu.ac.kr]
*Corresponding author: Sang-ug Kang

Received December 14, 2016; revised March 30, 2017; accepted April 18, 2017; published July 31, 2017

Abstract

Abrupt scene changes generally incur the afterimage effect. So, the unblocked image portion is still viewed by human eyes just after it is blocked by some pattern. Yovo's secure display method utilized this phenomenon and it is systematically analyzed using computational afterimage modeling by replacing the complex afterimage effect via simple low-pass filtering. With this approach, realistic images perceived by the human eye can be computationally generated at every single moment, especially reflecting the afterimage effect. The generated images are compared with the original images to determine the factors that affect the image quality of the secure display method. The simulation results demonstrate that the ratio of the unblocked portion to the blocked portion of an image and the playback rate are two primary factors related to the recognized image quality. We also found that the two factors are still effective for generalized secure display techniques

***Keywords*:** capture prevention, analog hole, secure image display, afterimage, privacy

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02037777)

https://doi.org/10.3837/tiis.2017.07.020 **ISSN : 1976-7277**

1. Introduction

Social network service (SNS) is a popular communication method nowadays. Numerous and various forms of digital content are shared and published on SNS platforms. However, as the number of SNS users increases dramatically and the content of the posts becomes more diverse, privacy issues arise. Since photos are more impactful means of communicating than text or video, unwanted circulation of the Internet has a great impact on privacy. Thus, some SNS platforms provide screen capture prevention functionality because an image can be easily obtained using built-in capture function of a computing device and it is distributed throughout the Internet immediately.

Many capture prevention technologies have been researched. Lee [1] proposed a message hooking-based method wherein 'print screen' or 'ctrl+c' keystrokes are detected by a background process. The current content of the clipboard is cleared as soon as something new is captured and copied to the clipboard. Stamp [2] introduced a low-level utility that intercepts system calls related to screen capture. With this approach, a digital rights management (DRM) system can distinguish between legitimate screen capture operations and unauthorized operations that attempt to access screen memory currently in use by protected applications. The common point of these methods [1][2] is that the application requires kernel level privilege, which may involve security vulnerabilities. Posteriori control methods, such as capture notification [3] and circulation monitoring [4], have been studied. In [3], the SNS application has a capture detection function that informs the photo owner that their image has been captured. However, this method can be circumvented by setting the device to airplane mode when taking a screenshot to prevent the notification message from being sent. In [4], an image illegally captured from the Internet is monitored or traced using digital fingerprints extracted from that image. For example, web hard sites install fingerprint matching filters and those are operated during file uploading. However, content owners and individuals with privacy concerns want more active and a priori countermeasures.

Since an attacker can use a camera to take a picture of the screen, all previously described methods may be compromised. Using a camera to capture an image is called an "analog hole" and this is an inherent weakness in any DRM system. Therefore, Yamamoto et al. [5] proposed a secure information display method based on multi-color visual cryptography. In that method, a pair of images and a single decoding mask share the visual information. The correct image can be seen only when the displayed image is viewed using the corresponding decoding mask. Because the displayed image is a random pattern, taking a screenshot, even with a camera, is ineffective. Hou et al. [6] utilized a secure image display technique to display a static binary image. The binary image is grouped into two parts: background and content. A series of randomly patterned images are sequentially displayed on a monitor. The random pattern is designed as follows. The content pattern varies at a slow frequency rate, and the background pattern varies at a rate sufficiently fast to utilize the temporal summation effect of the human eye [6]. As a result, the content is viewed as a distinct randomly patterned area and the background is perceived as a smooth gray area owing to the afterimage effect. Thus, the human visual system can recognize the content of binary images. Secure display methods [5][6] have relied on the characteristics of the human visual system rather than sophisticated digital algorithms. Yovo [7] is a kind of SNS platform that provides a picture capture prevention function using color image cryptography. If a user enables the "virtual fence" mode, the application places moving bars that flicker at a fast rate over the picture. When someone

attempts to take a screenshot of the picture, the screen shot is obscured by a virtual fence. However, when a person views the picture, he/she can see a little blurred image due to the afterimage effect rather than the image hidden by the bars. Here, the temporal summation or afterimage effect is exploited for secure image display. Even though the method used in [7] prevents both unwanted still image capturing and analog holes, the image quality is not as high as the original picture due to the virtual fences. Since image quality is the most important problem in this method, we investigate the factors affecting image quality and find out the correlation between these factors and image quality.

The remainder of this paper is organized as follows. We analyze the factors that affect image quality and identify optimal image quality conditions in Section 2. Simulation results are presented in Section 3, and conclusions are presented in Section 4.

2. Analysis of Yovo's capture protection method

In this section, Yovo's capture protection method [7] is analyzed from the perspective of obtaining optimal visual quality. Then, computational model of the afterimage phenomenon of the human eye is introduced.

2.1 Overall procedure of the capture protection method

The method presented in [7] is decomposed into fundamental elements to analyze and clarify the working mechanism. The decomposition analysis of the screenshot protection scheme in [7] is illustrated in **Fig. 1**. To simplify the analysis, the virtual fence is replaced with black privacy protection bar (PPB). This analysis is divided into two stages, i.e., GI generation and playback stages. At the GI generation stage, we need to generate PN different binary images (BIs) with zeros representing bars and ones representing free spaces, and the BIs overlap with the original image I with width I_w and height I_h to generate PN different grille images (GIs). Finally, all GIs are sequentially and repeatedly displayed, like continuous frames in a video sequence, at the playback stage. When $GI_1, GI_2, \ldots, GI_N, GI_1, \ldots$ are displayed at a certain frame rate, the viewer can see an image that is similar to the original image I. This scheme incapacitates the built-in capture function provided by an operating system or application because the captured image at one instance is obscured by PPBs.

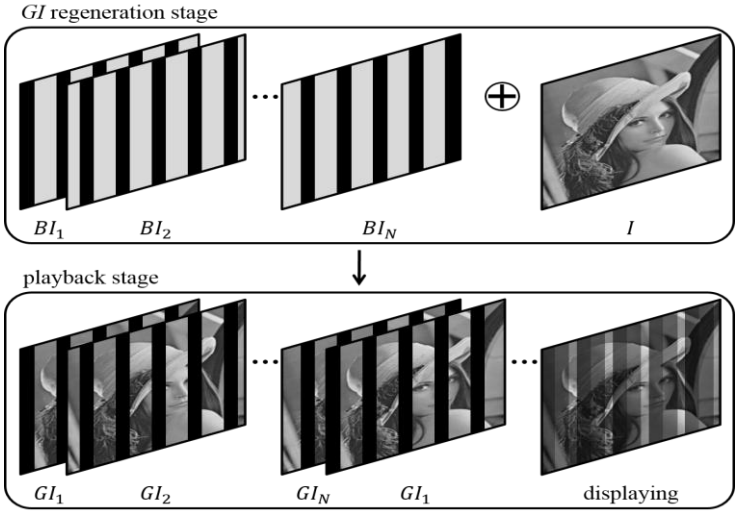
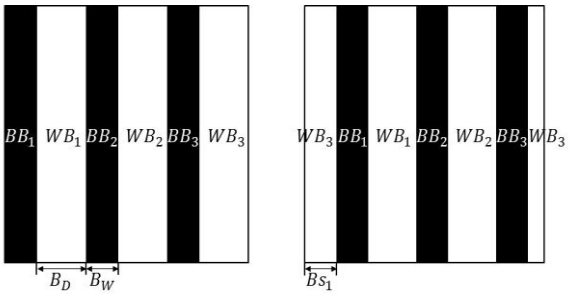


Fig. 1. The decomposition analysis of screenshot protection mechanism in [7].



(a) BI_1 (b) BI_2

(c) I (d) $\boldsymbol{GI_1}$

Fig. 2. (a) (b) Two consecutive BI frames with PPBs inside (c) the original image to be protected (d) an example of GI

At the GI generation stage, PN different BIs with the same height and width as I are produced. Each BI is denoted by BI_i, where $i = 1, \dots, PN$. As shown in **Fig. 2**, BI_i consists of multiple black bars BB_k and white bars WB_k, with the pixel values of each bar being 0 and 1, respectively. Let's assume that every BB_k has the same width B_W, and B_D is the distance between two black bars. The previously defined PN is referred to as the playback cycle, and the minimum value should be 2 for continuous display. Consecutive BI_i are generated by circularly shifting all BB_ks and WB_ks of the current BI_i by $Bs_i \geq 0$. For example, BI_2 is generated by shifting all BB_ks and WB_ks of BI_1 by Bs_1, and BI_3 is generated by shifting all BB_ks and WB_ks of BI_2 by $Bs_1 + Bs_2$ as shown in **Fig. 2**(a)(b). After shifting by $Bs_1 + \cdots + Bs_{PN}$, BI_1 is generated again. The statements can be expressed as Equation (1).

$$\begin{cases} BI_1 \neq BI_2 \neq \cdots \neq BI_{PN} \\ BI_i = BI_{i+jPN} \quad , j = 0, \cdots, n \end{cases} \tag{1}$$

GI_i is generated by adding BB_ks in BI_i and the portion of I corresponding to WB_ks together, as illustrated in **Fig. 2**(d). Therefore, similar to Equation (1), the statement can be expressed as Equation (2).

$$\begin{cases} GI_1 \neq GI_2 \neq \cdots \neq GI_{PN} \\ GI_i = GI_{i+jPN} \quad , j = 0, \cdots, n \end{cases} \tag{2}$$

At the playback stage, PN GIs are sequentially displayed from GI_1 to GI_{PN} on a screen and then the playback is repeated again and again. Due to the temporal summation effect [6], sequential and repeated playback results in overlapped images being perceived as the flickering single original image I with some distortion on the PPB portions. The playback rate

PR, which is strongly related to the afterimage effect, is measured in frames per second (fps) and is not dependent on PN.

2.2 Computational model of afterimages

Afterimages are optical illusions, and some attempts have been made to computationally simulate the illusion mechanism. The basic theme associated with the afterimage effect is the temporal resolution of the human eye depicted in [8]. Human eyes constantly sample information regarding images projected onto the retina. This information is accumulated so that objects around us appear to be stable and move smoothly. Temporal resolution is relatively low because it takes time for the eyes and brain to collect and process information. Consequently, if the display rate is sufficiently fast, humans do not perceive intermittent stimuli as separate, sometimes with flickering. Beyond the critical flicker frequency (CFF), flickering ceases and the stimuli appear as continuous light. The CFF is influenced by several factors, such as critical duration, Bloch's law, the Broca–Sulzer effect, the Ferry–Porter law, retinal position, and the size of the test field and so on. The whole mechanism is too complex and yet not fully understood. In [6], the CFF is first determined by Bloch's law and the Broca–Sulzer effect, and black and white speckles are spread throughout the whole image. Then, the background part of the image is changed over the CFF and the content part under the CFF such that the background is perceived as gray pixels and the content part is still viewed as speckles due to temporal summation. Ritschel et al. [9] proposed a model to compute afterimages that allows us to simulate their temporal, color, and time–frequency behaviors. The model used the simulations of retinal processes based on the local bleaching of photoreceptors. On top of that, several human visual system phenomena, such as retinal kinetics, diffusion, and chromatic effect, are included in the model to simulate more realistic afterimages. However, this approach ignored some characteristics of the human eye, such as saccade, rod photoreceptors, and retinal kinetics for six different chemicals. This limitation led to inaccurate simulation results and restricted the simulation environment as well.

Brettel et al. [10] proposed a subjective modeling method wherein people's observations were used to determine parameters of the impulse response of a filter. This approach assumes that the temporal integration property of the human eye can be represented as a hypothetical bi-phasic linear filter. Filter parameters are adjusted relative to the observation results and the spatial frequencies of the vertical and horizontal patterns. The vertical pattern image is presented for 108ms, followed by a blank interval of variable duration. Then, the horizontal pattern image is finally presented. Viewers compare the observed image with many comparison images and select the most similar one. By changing the blank interval, Brettel et al. could derive the afterimage effect of human eyes and finally the parameter values of the filter. The impulse response of filter, which is suggested in [10], is rewritten below.

$$h(t) = (t/\tau 1)^{n1-1} e^{-t/\tau 1} - \eta (t/\tau 2)^{n2-1} e^{-t/\tau 2} \tag{3}$$

,where $\tau 1$ and $\tau 2$ are time constants, and n_1 and n_2 are the number of stages in the excitatory and inhibitory parts of the filter, and η is the amplitude ratio between the two parts of the filter.

3. Factors and conditions for optimal image quality

3.1 Design of privacy protection bars

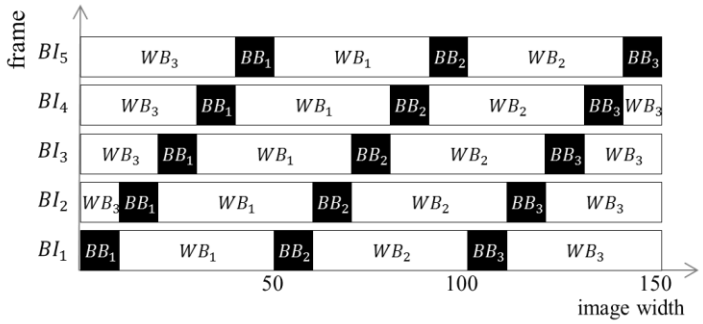
There are two important elements for generating BIs, i.e., heterogeneity between the captured image by an attacker and I, and homogeneity between the human perception of sequentially displaying grille images, or perceptual GI, and I. The contextual meaning of heterogeneity is that an image interrupted by opaque vertical bars, which is obtained by screen capture or by clicking a photo with another camera, should sufficiently differ from I by widening B_W and/or increasing the number of BB_is. The security purpose of this secure image display scheme is obtained by heterogeneity. Homogeneity is important for perceptual image quality, and the visual quality degradation of perceptual GI, compared with I, should be quantitatively measured. The homogeneity primarily depends on the PPB design. The fundamental PPB design principles are listed below.

[Principle 1, P1] $255 \times (\bigcup_{i=1}^{PN} W(BI_i)) = WI$
[Principle 2, P2] $255 \times W(BI_i) \neq WI$
[Principle 3, P3] $B(BI_i) \cap B(BI_{i+1})=\emptyset$
[Principle 4, P4] $M(B_W + B_D) = I_W,\ M = 1,2,\cdots,n$
[Principle 5, P5] $PN = 2,3,\cdots,n$

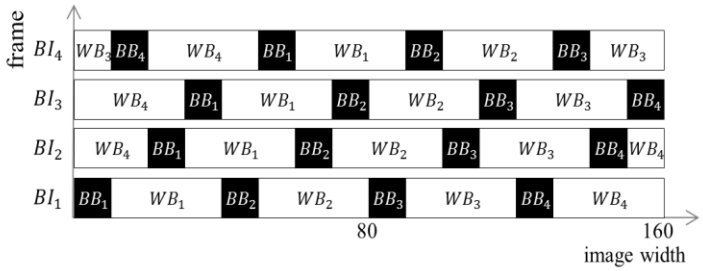
P1 states that the union of white bars of all BI_is multiplied by 255 (maximum pixel value) should be WI, where $W(\cdot)$ denotes the one-value area of an image and WI represents a white image with the same size as I. Without P1, a portion of an image is never displayed during playback, i.e., some portion of I is always occluded by the PPBs in every BI_i. Given P1, the viewer can see every part of the image at least once during playback. However, P1 does not mean that the union of PPB areas of all BI_is should result in a black image. In other words, it is possible for a part of the image to be displayed without being hidden by the PPBs.

With principle P2, the one-value area of one BI_i cannot correspond to WI. It means that some area of BI_i would be blocked by a PPB. P3 means that the zero-value area of BI_i does not overlap with the zero-value area of BI_{i+1}, which makes us to take full advantage of the afterimage effect. Similarly, $B(\cdot)$ denotes the zero-value area of an image. Both P4 and P5 are ancillary principles that reduce design complexity. P4 indicates that $B_W + B_D$ divides I_W and M is a positive integer. M is a scaling factor to determine the width and number of PPBs. For example, $B_W = I_W/2$ and the number of PPB is one in case of $M = 1$ assuming $B_W = B_D$. Likewise, $B_W = I_W/4$ and the number of PPB is two in case of $M = 2$. P5 is mentioned earlier and included as a principle. By considering these five principles, we deduced the following PPB design procedures.

[Step 1] Obtain I_W from I and then arbitrary determine M in P4.
[Step 2] Determine B_W and B_D by P4. In addition, $B_D \geq 1$ by P1, $B_W \geq 1$ by P2, and $B_W \leq B_D$ by P3 should hold.
[Step 3] Determine each Bs_i such that $B_W + \sum_{j=1}^{i-1} Bs_j \leq Bs_i \leq B_D + \sum_{j=1}^{i-1} Bs_j$, $i = 1,2,\cdots,n$ by P3.
[Step 4] Determine PN such that $\sum_{i=1}^{PN} Bs_i \equiv 0\ (mod\ T_W)$ by P4 and P5, where $Bs_i = Bs_{i+jPN},\ j = 1,2,\cdots,n$ and $T_W = B_W + B_D$.
[Step 5] Generate BI_is using B_W, B_D and Bs_i.
[Step 6] Determine PR.



(a) $\boldsymbol{I_W = 150, M = 3, B_W = 10, B_D = 40,\ Bs_i = 10}$



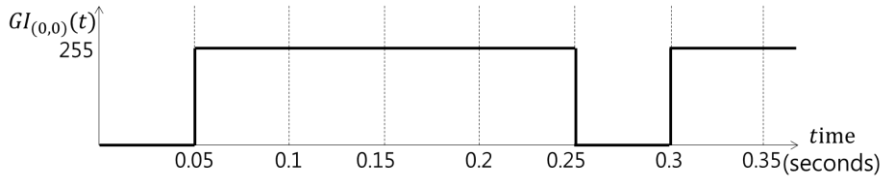
(b) $\boldsymbol{I_W = 160, M = 4, B_W = 10, B_D = 30,\ Bs_1 = Bs_3 = 20,\ Bs_2 = 10}$

Fig. 3. Two PPB design examples

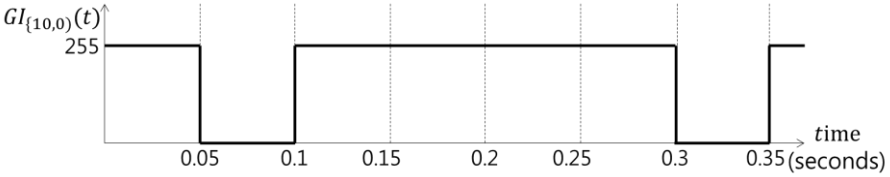
Fig. 3 illustrates two PPB design examples that adhere to the deduced PPB design procedures. For an efficient representation, the whole image is expressed as horizontal bands with reduced height. The design shown in **Fig. 3** (a) proceeds as follows:

[Step 1] Obtain $I_W = 150$ from the image size of 150 × 150 and then arbitrary determine $M = 3$.
[Step 2] Since $3(B_W + B_D) = 150$ and $B_D \geq B_W \geq 1$, we can set $B_W = 10$ and $B_D = 40$.
[Step 3] We set $Bs_1 \sim Bs_5 = B_W$.
[Step 4] To satisfy $\sum_{i=1}^{PN} Bs_i \equiv 0\ (mod\ 50)$, we can set PN to 5.
[Step 5] We can generate five binary images $BI_1 \sim BI_5$, as shown in **Fig. 3** (a).
[Step 6] We set $PR = 20\ fps$. Thus, $GI_1 \sim GI_5$ are repeatedly displayed on the screen at the speed of 20 frames per second.

The design shown in **Fig. 3** (b) proceeds as follows. In Steps 1 and 2, set $I_W = 160$, $M = 4$, $B_W = 10$, and $B_D = 30$. Then, in Steps 3~5, we determine that $Bs_1 = Bs_3 = 20$, $Bs_2 = 10$, $Bs_4 = 30$, and $PN = 4$. Finally, we set $PR = 30\ fps$, and then $GI_1 \sim GI_4$ are displayed repeatedly.



(a) The signal GI at (0,0) along the time axis t



(b) The signal GI at (10,0) along the time axis t

Fig. 4. Two examples of GI in Figure 3(a) at $PR = 20$ fps

3.2 Optimal image quality

A recognized image (RI) is defined as a subjective image recognized by a viewer when GI_is are displayed continuously at a fixed frame rate. $RI_{(x,y)}(t)$ is an RI at position (x, y) and time t. The afterimage effect is reflected in RI so that it looks different from GI and $RI_{(x,y)}(t) = I_{(x,y)}$ is the ideal case. $GI_{(x,y)}(t)$ is the displayed image on a monitor screen at position (x, y) and time t, so no human visual characteristics are reflected. The RI can be obtained by convolving the displayed image and a low-pass filter $h_{low}(t)$ based on Brettel et al.'s observation [10], as shown in Equation (4).

$$RI_{(x,y)}(t) = GI_{(x,y)}(t) * h_{low}(t) \tag{4}$$

For simplicity, let's assume that $GI_{(x,y)}(t)$ is made of overlapping BI_is with a white image WI, where BI_is are generated using the parameters employed in **Fig. 3**(a). If the original image is a natural one, there may be many kinds of GIs, but only five kinds of GIs can exist in this case; $GI_{(0,0)}(t)$, $GI_{(10,0)}(t)$, $GI_{(20,0)}(t)$, $GI_{(30,0)}(t)$, and $GI_{(40,0)}(t)$. Those five signals have the same shape except that they are right-shifted by 50ms on the time axis, sequentially. For example, $GI_{(0,0)}(t)$ and $GI_{(10,0)}(t)$ signals illustrated in **Fig. 4**(a) and 4(b) have the same shape and $GI_{(0,0)}(t)$ is right-shifted by 50ms to be $GI_{(10,0)}(t)$. Naturally, the image quality of the RI, denoted by Q_{RI}, increases as $RI_{(x,y)}(t)$ approaches the original image. In this case, the original image is WI and $RI_{(x,y)}(t) = WI$ is the ideal situation in terms of image quality. Therefore, we need to find out conditions that satisfies the following equation.

$$arg\ min \sum_{x=0}^{I_W} \sum_{y=0}^{I_H} \left| GI_{(x,y)}(t) * h_{low}(t) - 255u(t) \right| \tag{5}$$

, where $u(t)$ is the unit function. It is possible to analyze Q_{RI} with pixels in one horizontal line rather than the whole image because all horizontal lines are identical. So, Equation (5) can be replaced with Equation (6).

$$arg\ min \sum_{x=0}^{I_W} \left| GI_{(x,0)}(t) * h_{low}(t) - 255u(t) \right| \tag{6}$$

It should be noticed that only $GI_{(x,0)}(t)$ is changeable in Equation (6), which can affect to the recognized image quality. B_W and B_D determines how many different GI signals can be generated and how much the original image is displayed. Since the pixel value at a position (x, y) of GI_i depends on these two parameters, B_W and B_D are determinants of Q_{RI}. Using the inequality relation of the convolution operation, Equation (6) is deduced.

$$arg\,min \sum_{x=0}^{I_w} \left| \left\| GI_{(x,0)}(t) \right\| \left\| h_{low}(t) \right\| - 255u(t) \right| \tag{7}$$

Let's assume that the $h_{low}(t)$ is a sync function, that is an ideal low pass filter. The convolution operation is removed in Equation (7) and we can intuitively know that the magnitude of $GI_{(x,y)}(t)$ is the main factor to minimize the difference. By changing PR, frequency of $GI_{(x,y)}(t)$ changes as well. If the frequency of a square wave function increases and the low pass filtering effect also increases, which results in a zero-valued portion of $GI_{(x,y)}(t)$ decreasing. Therefore, the frequency of $GI_{(x,y)}(t)$ must be maximized to take advantage of the filtering effect. The zero-valued portion of $GI_{(x,y)}(t)$ is proportional to the ratio of B_D/B_W. It is also natural that the effect of low-pass filtering is dependent on PR. In the next section, we verify the effect of B_D/B_W and PR through computational simulation.

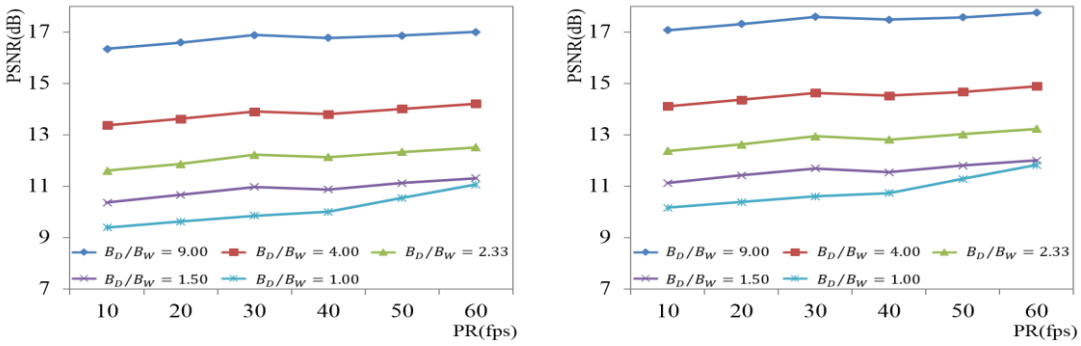
4 Simulation Results

4.1 Simulation of image quality with PPBs

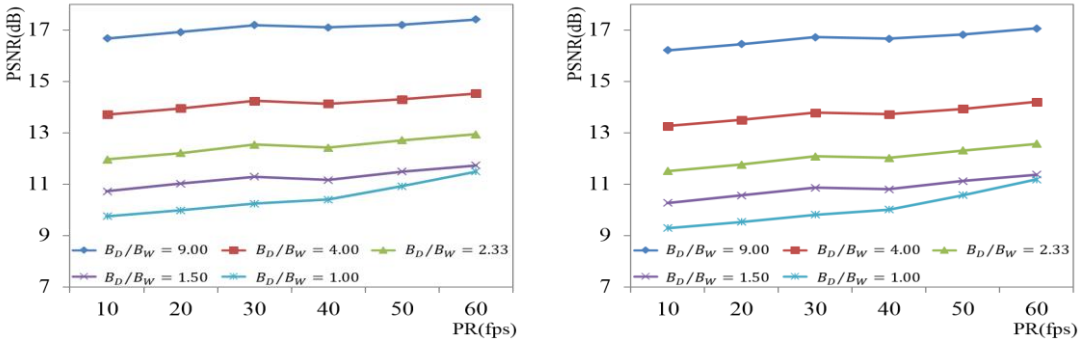
The overall simulation was performed using MATLAB R2013a version with 150 × 150 grayscale images that were resized, for simulation convenience, from 256 × 256 Lena, Goldhill, Barbara, and Baboon test images. The purpose of the simulation is to verify the effect of B_D/B_W and PR on visual quality. This experiment does not calculate how the image actually looks in the human eye, but it does calculate how the person perceives the afterimage effect. The afterimages were considered by replacing $h_{low}(t)$ in Equation (4) with Equation (3). All other elements that may affect the RI quality are ignored. These elements include monitor brightness, simulation room brightness, saccade, rod photoreceptors, and retinal kinetics for six different chemicals. Therefore, $RI_{(x,y)}(t)$ is not the same as the real image perceived by the human visual system. The simulation is performed to obtain recognized images at any time t so that we can measure the similarity between the original image and the recognized image at any time. $RI_{(x,y)}(t)$ is obtained using Equation (3) and (4), and the resulting pixel values should be normalized to 0 through 255. The image quality is objectively measured by changing the exposure rate B_D/B_W and PR based on peak-signal-to-noise ratio (PSNR) criteria. In addition, image quality is also measured by changing Bs_i but not changing B_D/B_W and PR to prove that Bs_i is not an effective factor.

Table 1. Comparison of image quality by changing Bs_i ($B_W = 10, B_D = 40,\ PR = 30$)

Bs_i	PN	PSNR(dB)			
		Lena	Goldhill	Barbara	Baboon
20,10,10,20,40	5	13.94	14.64	14.24	13.78
15,15,15,15,15,15,10	7	13.97	14.71	14.30	13.81
27,33,21,39,20,10	6	13.90	14.77	14.23	13.87
12,10,12,12,10,12,10,12,10	9	13.96	14.66	14.25	13.80

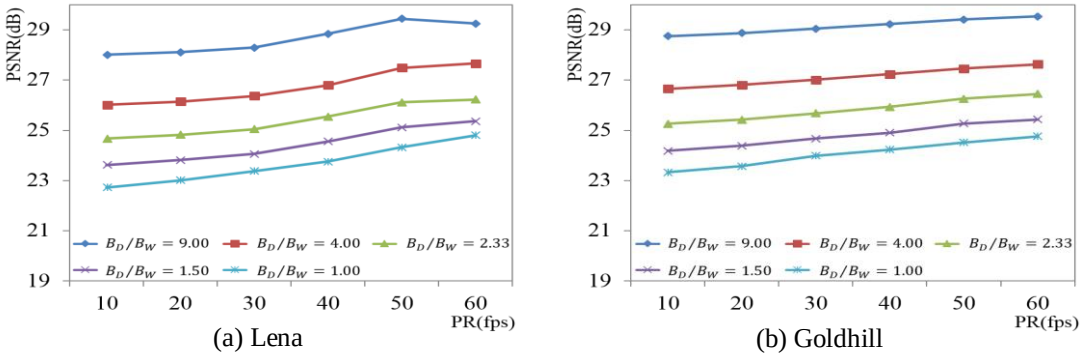


(a) Lena (b) Goldhill

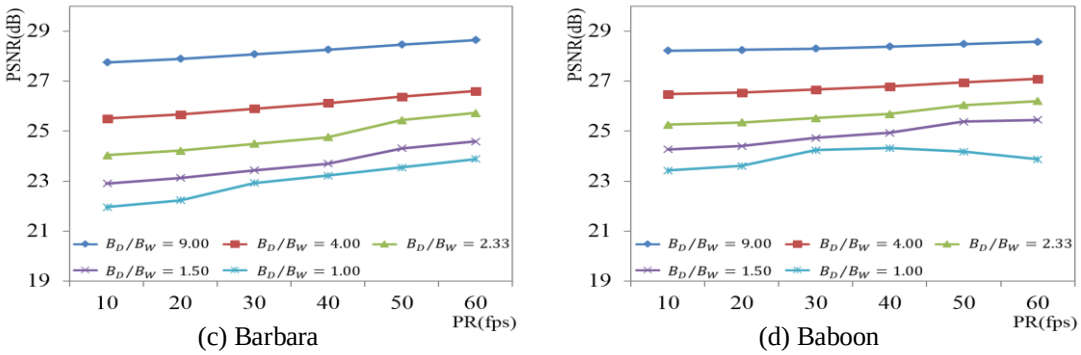


(c) Barbara (d) Baboon

Fig. 5. Average PSNR variation of $\boldsymbol{RI}_{(x,y)}(\boldsymbol{t})$ with respect to exposure rate and *PR* (black bar)



(a) Lena (b) Goldhill

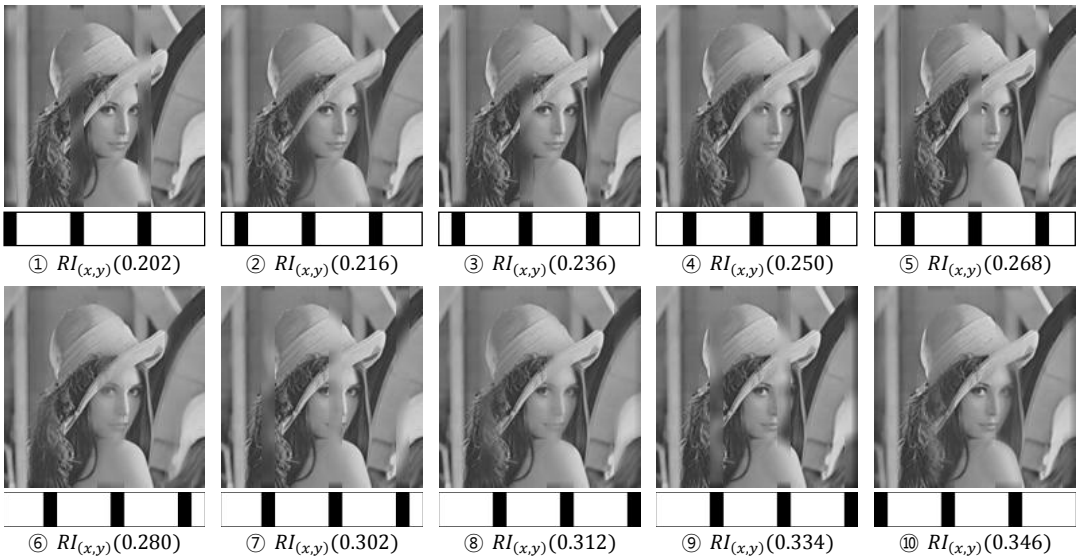


(c) Barbara (d) Baboon

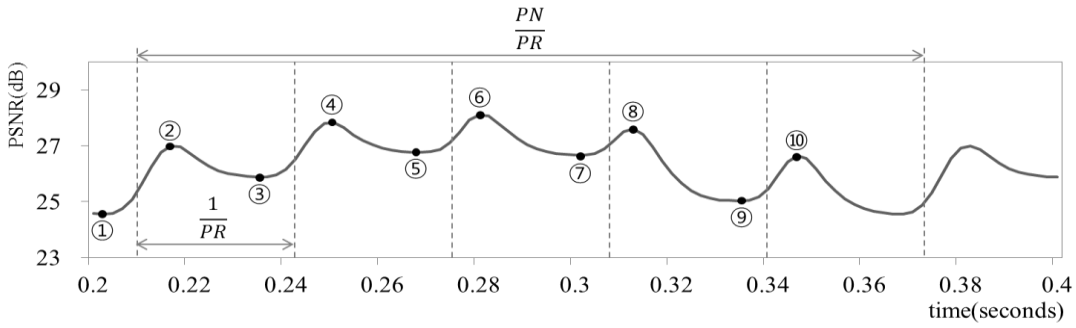
Fig. 6. Average PSNR variation of $\boldsymbol{RI}_{(x,y)}(\boldsymbol{t})$ with respect to the exposure rate and *PR* (blur bar)

The parameters for the first experiment were set to $T_w = 50$ and $Bs_i = B_w$ while varying B_D/B_W and PR. Many recognized images are produced between $t = 0$ to 1 second with a time interval of 0.002 seconds. The PSNR is calculated for each *RI*. The average PSNR is shown in **Fig. 5**. All results indicate that image quality is proportional to both B_D/B_W and PR. For example, at $PR = 20$, the image quality of Lena is 16.59, 13.62, 11.86, 10.66, and 9.62 dB at exposure rates of 9, 4, 2.33, 1.5, and 1, respectively (**Fig. 5**(a)). In addition, at $B_D/B_W = 9$, the image quality is 16.59, 16.78, and 17 dB at a playback rate of 20, 40, and 60 fps, respectively. It is evident that image quality increases as the exposure and playback rates increase.

On the contrary, as seen in **Table 1**, Bs_i has little effect on image quality. With a fixed exposure rate of 4.0 and a playback rate of 30 fps, the four variations of Bs_i did not significantly change the image quality for any test image. Therefore, it can be seen that how much one frame moves to make the next frame does not affect the image quality.



(a) Sample *RI*s at various time



(b) Image quality vs. time

Fig. 7. The sample *RI*s and image quality based on *BI*s with blur bars shown on the time axis.

4.2 Simulation for better image quality and security

The image quality measured in the previous section is quite low because the black bars caused significant distortion. As mentioned above, black bars are used to provide accurate analysis; those cannot not be used in an actual application. To improve image quality, the black bars are replaced with blur bars [7] that can be generated using a 9 × 9 average filter. Specifically, the pixel values of the blur bar are obtained by averaging the pixel values of the original image *I*. Under these conditions, image quality is improved by more than 10 dB for all test images, and the overall simulation result is consistent with the previous result, as illustrated in **Fig. 6**. For example, the PSNR of Lena increases from 16.19 to 28.11 dB at $PR = 30$ fps and $B_D/B_W = 9$, which supports the previous conclusion that the exposure rate and playback rate are proportional to the image quality. However, the playback rate has some limitations for improving image quality, particularly for Lena and Baboon at high *PR*. Ten sample *RI*s with the positions of blur bars and sampling time are shown in **Fig. 7**(a). The positions of blur bars are clearly visible in black just under the image. Each sample *RI* is taken at the times marked with circle number in **Fig. 7**(b). Also, the change in PSNR over time is also plotted in **Fig. 7**(b). It is important to note that the PSNR increases after the frame change and gradually decreases thereafter.

With regard to security, if an attacker captures PN screenshots of every different GI, then the original image I is easily reconstructed. Since the optimal capture condition for this multiple capture attack is easily calculated when this scheme is used, a randomly generated pattern can replace the blur bars. In this simulation, squares blurred with the same 9 × 9 2-dimensional average filter are used with exposure rate and playback rate parameters.

Since Bs_i has no effect on the image quality, the presence of a randomly placed 15 × 15 blur squares throughout the image will protect against unwanted screen capture. The exposure rate is redefined as the area uncovered by blur squares over the area covered by those. PN cannot be defined in this case because positions of blur squares are random on every GI. So, the corresponding limitation is that the previously blocked area should not be blocked in the current GI to avoid occlusion, or visual outage. The PR remains the same. As shown in **Table 2**, the simulation results for random blur squares are nearly the same as those for the blur bar, which means that the pattern shape and periodicity do not affect image quality. The only factors that affect image quality significantly are PR and exposure rate. Five recognized images at $t = 0.33$, marked with a circle number, are shown in **Fig. 8** and those are corresponding to the PSNR values next to the circle number in **Table 2**.

Table 2. Simulation results with random blur block compared to blur bar

Exposure Rate	Protection Type	PSNR(dB)			
		Lena	Goldhill	Barbara	Baboon
9.00	Random blur block	28.73 ①	29.47	28.07	28.23
	Blur bar	28.84	29.23	28.26	28.38
4.00	Random blur block	26.81 ②	27.24	26.07	26.52
	Blur bar	26.80	27.23	26.12	26.79
2.33	Random blur block	25.36 ③	25.95	24.88	25.31
	Blur bar	25.54	25.93	24.76	25.69
1.50	Random blur block	24.52 ④	25.03	24.03	24.35
	Blur bar	24.56	24.91	23.69	24.93
1.00	Random blur block	23.75 ⑤	24.23	22.96	23.70
	Blur bar	23.75	24.23	23.22	24.32

① ② ③ ④ ⑤

Fig. 8. $RI_{(x,y)}(\mathrm{t} = 0.33)$ at the exposure rate of ① 9.00, ② 4.00, ③ 2.33, ④ 1.50, ⑤ 1.00 for the simulation results in Table 2.

5 Conclusion

The secure image-display technique used in [7] has been analyzed in terms of image quality because psychological acceptability is an important secure-system design principle. Capture protection is a difficult but important task, particularly when the analog hole issue is considered. At the same time, the *RI* should be acceptable under all circumstances. Thus, fundamental design principles and procedures have been suggested to systematically control image quality. In addition, we proposed a computational model for the afterimage effect to quantitatively measure image quality and found that only exposure rate and playback rate are proportional to *RI* quality. To verify the theoretical results and illustrate how the afterimages are seen by the human eye, *RI*s were generated at specific times using MATLAB. To achieve better image quality and security, blur bars and random blur squares are used rather than black bars. Random blur blocks showed the best results in terms of both image quality and security because it is harder for an attacker to collect all the captured images required to reconstruct the original image. The resulting PSNR values are approximately between 23dB and 29dB according to the exposure rate and playback rate. It should be understood that the degradation is a price of preserving privacy. In actual use, high brightness of display monitor may help to increase the subjective image quality because high luminance doubles afterimage effect.

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02037777).

References

[1] J. Lee, "Implementation of anti-screen capture modules for privacy protection," *Journal of the Korea Institute of Information and Communication Engineering,* vol. 18, no. 1, pp. 91-96, 2014. Article (CrossRef Link)

[2] M. Stamp, "Digital Rights Management: The Technology Behind the Hype," *J. Electron. Commer. Res.,* vol. 4, no. 3, pp. 102-112, 2003.

[3] J. Charteris, S. Gregory, and Y. Masters, "Snapchat 'selfies': The case of disappearing data," *eds.) Hegarty, B., McDonald, j., & Loke, S. K., Rhetoric and Reality: Crit. Perspect. Educ. Tehnol.,* pp. 389-393, 2014.

[4] W. Kim, S. Lee, and Y. Seo, "Image fingerprinting scheme for print-and-capture model," *Pacific-Rim Conference on Multimedia*. Springer Berlin Heidelberg, pp. 106-113, 2006. Article (CrossRef Link)

[5] H. Yamamoto, Y. Hayasaki, and N. Nishida, "Secure information display with limited viewing zone by use of multi-color visual cryptography," *Opt. Express,* vol. 12, no. 7, pp. 1258-1270, 2004. Article (CrossRef Link)
[6] J. Hou, D. Kim, H. Song, and H. Lee, "Secure Image Display through Visual Cryptography: Exploiting Temporal Responsibilities of the Human Eye," in *Proc. of the 4th ACM Workshop on Information Hiding and Multimedia Security*, ACM, pp. 169–174, 2016. Article (CrossRef Link)
[7] Yovo [Internet] Available at:http://yovo.me. [Accessed 14 Aug 2016].
[8] H. Kolb, E. Fernandez, and R. Nelson, "Webvision: the organization of the retina and visual system," *Salt Lake City (UT): University of Utah Health Sciences Center*, 1995.
[9] T. Ritschel, and E. Elmar, "A computational model of afterimages," in *Proc. of Computer Graphics Forum*. Blackwell Publishing Ltd, vol. 31. no. 2pt3, pp. 529-534, 2012. Article (CrossRef Link)
[10] H. Brettel, L. Shi, and H. Strasburger, "Temporal image fusion in human vision," *Vision Res.,* vol. 46, no. 6, pp. 774-781, 2006. Article (CrossRef Link)

Sang-ug Kang is a Professor in Computer Science at Sangmyung University, Korea Rep. of. He received his M.S. degree in Electrical Engineering in 1995 from University of Southern California and Ph.D. degree in Information Security in 2011 from Korea University. His research interests include multimedia security, information security, and secure ICT services.

Sanghyun Park received her B.S. degree in Computer Science in 2015 from Sangmyung University, Korea Rep. of. He is currently a Master degree candidate in the Department of Computer Science, Sangmyung University, Seoul Korea. Her research interests include multimedia security and secure image display.