DETECTION OF PERIODIC PILOT SIGNAL IN IMAGE WATERMARKING

Rinka Kawano and Masaki Kawamura
Yamaguchi University, Yamaguchi, Japan
E-mail: {m.kawamu,b027vbv}@yamaguchi-u.ac.jp Tel: +81-83-933-5701

***Abstract*—A watermarking method that has robustness against geometric and non-geometric attacks is desired. Conventional methods extract watermarks regardless of the type of attacks. For example, error correction codes have been used to correct errors in watermarks, but these methods have limitations. If we could identify the type and strength of attacks, we might be able to achieve fewer errors. To detect the attacks, a method that embeds a template in the discrete Fourier transform (DFT) domain has been proposed. This method has issues in that template matching takes a very long time and it does not support clipping since it uses the DFT domain. Therefore, we propose a method to estimate the magnification ratio by embedding a pilot signal in a non-transformed domain. A watermark and the pilot signal are embedded in the Y and U components, respectively. In this study, a grid pattern was used as the pilot signal. The magnification ratio was estimated by extracting the periodicity of the grid pattern. As a result, for uncompressed images, the estimations were almost accurate. For compressed images, the results were mostly accurate as long as the Q factors were above 60.**

I. INTRODUCTION

It is now common for ordinary users to post images and photographs on social networking services (SNS) such as Twitter and Facebook. Anyone can easily upload images on SNS. At the same time, others can easily download the posted images or take screenshots of them. The unauthorized use of images has become a problem. Digital watermarking is one of the solutions to this problem [1] where additional information (message) is secretly embedded in digital contents such as images and photographs. The embedded message, which is encoded for error correction, is called a watermark. Images embedded with watermarks are called stego-images. By embedding and extracting the watermarks, it can be used to prevent unauthorized use of contents and manage digital contents.

Images misused on SNS are processed and tampered with. In digital watermarking, these processes are called attacks. Attacks can be generally categorized into geometric and non-geometric attacks. Non-geometric attacks change the pixel values, such as JPEG compression and noise addition. Geometric attacks change the position of pixels, such as scaling, rotation, and clipping. Watermarking methods require resistance to these attacks to extract messages with high accuracy. To correctly extract the message from an attacked stego-image, error correction codes [2]–[4] or spread spectrum methods [5], [6] have been used to reduce the error during decoding. By using the scale-invariant feature transform (SIFT) feature points of an image [7], it is possible to resist geometric attacks [8]–[10]. Identical SIFT feature points tend to be extracted even when the magnification ratio and rotation angle are unknown. However, in this case, the extracted watermark contains many errors. Therefore, if we could estimate parameters of attacks applied to the stego-image, the error rate could be reduced.

Methods for embedding watermarks in the discrete Fourier transform (DFT) domain have been proposed [11], [12]. DFT is invariant to scaling but has a property in which the positions of the DFT coefficients are rotated with the angle of rotation. It should be noted that the value of the DFT coefficient changes in proportion to the magnification ratio. Therefore, conventional methods require the original image. If the magnification ratio can be estimated, watermarks can be accurately extracted without the original image. A method to estimate attack parameters, e.g., magnification ratio and angle of rotation, includes template matching in the DFT domain [13]. The type and parameters of attacks can be estimated by embedding templates. However, the problem is that the template matching takes a huge amount of time. In this paper, we propose a watermarking method that embeds a pilot signal as a template, detects the pilot signal from degraded images, and estimates the magnification ratio as an attack parameter.

This paper is organized as follows: Section II describes the method for estimating the magnification ratio using the pilot signal. Section III presents the results of computer simulations. We conclude our study in Section IV.

II. ESTIMATION OF MAGNIFICATION RATIO USING PILOT SIGNAL

In the proposed method, a pilot signal is embedded to estimate the magnification ratio. To avoid interference with watermarks, the pilot signal is embedded in a different color component from one of the watermarks.

A. Embedding a pilot signal

In this method, an image with RGB color space is converted to one with YUV color space without Chroma subsampling. That is, each of the YUV components has the same sample rate, 4:4:4. The watermarks are embedded in the Y component (luminance). To avoid affecting the watermarks, the pilot signal is embedded in the U or V component. Here, the U component is selected. Let the image size of the U component be $M \times N$ pixels. Let $U(i,j)$ be the intensity at position (i,j). Now, the intensity $U(i,j)$ can be decomposed into the

978-988-14768-9-0/21/$31.00 ©2021 APSIPA

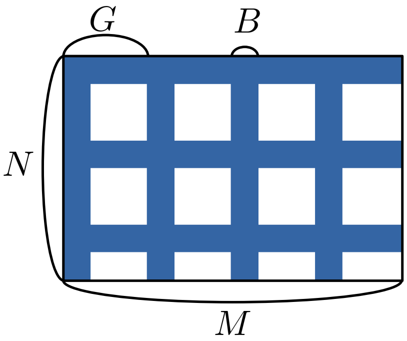


Fig. 1. Layout of pilot signal

bit representation $\{b_7(i,j), \cdots, b_2(i,j), b_1(i,j)\}$. That is,

$$U(i,j) = \sum_{l=0}^{7} 2^l b_l(i,j), \tag{1}$$

where $b_7(i,j)$ is the most significant bit (MSB) and $b_0(i,j)$ is the least significant bit (LSB). The image consisting of the l-th bit representation $b_l(i,j)$ is called the l-th bitplane. We will embed a pilot signal in the $l = 2$-nd bitplane. As shown in Figure 1, the pilot signal replaces the value of the bitplane with 1 in a grid pattern. In other words, the values in the blue region take the value of 1, and those in the white region are unchanged. The interval and width of the grid is $G = 50$ and $B = 5$ pixels, respectively. To avoid long search times for template matching [13], the grid pattern is used in this paper. Thereby, the magnification ratio can be estimated easily.

B. Estimation of magnification ratio

We assume that the stego-image has been processed with scaling and JPEG compression. The size of an attacked image is $M' \times N'$ pixels and the original size $M \times N$ is unknown. Under this assumption, the magnification ratio is estimated by using the pilot signal. Since the same processing is performed in the vertical and horizontal directions for the attacked image, we will describe the vertical direction here.

(1) Calculate the sum of the bits vertically in the l-th bitplane. For the U component of the attacked image, the sum of the bits $b_l(i,j)$ in the i-th column, S_i, is given by

$$S_i = \sum_{j=0}^{N'-1} b_l(i,j). \tag{2}$$

Let the series of sums be $\boldsymbol{S} = \{S_0, S_1, ..., S_{M'-1}\}$.

(2) Calculate the autocorrelation of the series $\boldsymbol{S}$. Since the values on the grid are 1, the series $\boldsymbol{S}$ takes large values periodically at each grid interval. When compression is applied, this periodicity becomes harder to detect. By calculating the autocorrelation, it is easier to detect the period.

(3) Perform the DFT of the autocorrelation coefficients to estimate the period. The period is calculated as the inverse of the main frequency obtained from the DFT coefficients. When scaling or compression is applied, the possibility of detecting the wrong period is increased. Since the frequency related to data length of the autocorrelation also appears, the DFT is performed after applying a window function. In this study, the flat top window is applied.

Since the grid interval G is constant, the autocorrelation has peaks at multiples of the interval. Therefore, in the DFT domain, peaks also appear at multiples of the frequency $f_0, 2f_0, 3f_0, \cdots$, where $f_0 = 1/G$. This cyclic property is used for the period detection to avoid detecting the wrong period, Specifically, when integer multiples of frequencies such as $f_0, 2f_0, 3f_0$ are detected, the grid interval G is estimated as $G = 1/f_0$ by the smallest frequency f_0. In the case in which more than one period is detected, we treat it as a detection failure since the period cannot be identified automatically.

(4) Letting the estimated grid interval be $\widehat{G}$ pixels, the estimated magnification ratio $\hat{m}$ can be derived by

$$\hat{m} = \frac{\widehat{G}}{G}. \tag{3}$$

C. Detection failure

As previously mentioned, detection failures sometimes occur when the images are attacked. If no frequency f_0 is found or multiple frequencies are found, the estimation is treated as a detection failure. If an image has periodic features, the frequency may be incorrectly detected as a pilot signal.

Two ratios $\hat{m}$ can be obtained from the vertical and horizontal directions. For scaling with an equal aspect ratio, the average of the two ratios $\hat{m}$ can be used as the estimated ratio. Moreover, if one of the two is a detection failure, the other can be used as the estimated ratio.

D. Application to watermarking

Once the estimated magnification ratio $\hat{m}$ is obtained, the watermark can be extracted with high accuracy by inverse transformation. As an example, we will discuss the application of the DFT-based watermarking method. For DFT on an image $I(x,y)$ with $M \times N$ pixels, the spatial frequency $F(k_1,k_2)$ is expressed as

$$F(k_1,k_2) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y) e^{-2\pi j\left(\frac{k_1 x}{M} + \frac{k_2 y}{N}\right)}, \tag{4}$$

where k_1, k_2 denote the frequencies of the x, y components, respectively, and j is the imaginary unit. When the image is magnified by the magnification ratio m, the following equation is satisfied.

$$\mathrm{DFT}\left[I(mx,my)\right] = \frac{1}{m} F\left(\frac{k_1}{m}, \frac{k_2}{m}\right), \tag{5}$$

Therefore, when extracting a watermark from a rescaled stego-image, the original DFT coefficients can be obtained by multiplying the calculated DFT coefficients by $1/\hat{m}$.

III. Computer simulations

Assuming scaling and JPEG compression as attacks, we estimate the magnification ratio. In this paper, to focus on the extraction of the pilot signal, no watermark is embedded.

A. Evaluation for scaling

First, the estimated magnification ratio $\hat{m}$ is calculated for the case where only scaling without JPEG compression is applied. Twelve original images with 1920×1080 pixels are used as evaluation images. The pilot signal is embedded in the U component of the original images. The magnified images are generated by scaling the original images from 0.7 to 1.3 times. The estimated magnification factor $\hat{m}$ will be calculated from the magnified images.

The pilot signal is detected as follows.

(I) For the $l = 2$-nd bitplane, the totals of each row and column are calculated. We will describe the vertical direction here. A histogram of the bit totals is shown in Fig. 2 (a). An abscissa shows the index of a column and an ordinate represents the total of bits. We can see peaks appearing periodically.

(II) To find the period of these peaks, the autocorrelation of the histogram is calculated as shown in Fig. 3 (a). We can see that a peak appears at every grid interval $G = 50$ pixels.

(III) The period of this peak can be determined by DFT. The power spectrum of the autocorrelation is shown in Fig. 4 (a). An abscissa shows the frequency and an ordinate represents power spectral density (PSD). From the frequencies with large PSDs, the estimated grid interval $\hat{G}$ can be estimated.

(IV) Hence, from (3), the estimated magnification ratio $\hat{m}$ can be determined using the estimated grid interval $\hat{G}$ and the original grid interval G.

Now, we extracted the pilot signal from magnified images and calculated the estimated grid intervals $\hat{G}$. We define the relative error R as a measure of estimation accuracy, and it is defined by using the estimated magnification ratio $\hat{m}$ to the true magnification ratio m as

$$R = \frac{|\hat{m} - m|}{m}. \tag{6}$$

When the estimation is excellent, $R = 0$ holds. Figure 5 shows the error R. The abscissa shows the true magnification ratio m from 0.7 to 1.3, and the ordinate shows the relative error R by a box-and-whisker plot. The upper and lower whiskers represent the maximum and minimum values, respectively. The upper and lower regions of the box represent the first and third quartiles, respectively. The orange line inside each box represents the median and the blue line represents the average error. As shown in Fig. 5, the errors R were approximately under 0.05 for all magnification ratios. Note that three cases out of $12 \times 7 = 84$ images were excluded due to detection failures (DF). The number of DF at each magnification ratio are shown in Table I. All DF occurred when the images were shrunk. We found that the proposed method could estimate the estimated magnification ratio accurately in most cases, and that it tends to be difficult to estimate the ratio when the images were shrunk.

Next, let us consider image quality. Figure 6 shows (a) a part of stego-image and (b) the pilot signal. In the pilot signal, white pixels represent the area where values of 1 are embedded, and black pixels represent the area where the values of the bitplane are not changed. As far as a visual inspection of the stego-image was concerned, the embedded pilot signal could not be detected. The image quality for images embedded with a pilot signal is evaluated by the peak signal-to-noise ratio (PSNR). The average PSNR was 44.8 dB.

TABLE I
NUMBERS OF DF FOR NO COMPRESSION.

	magnification ratio m							total
	0.7	0.8	0.9	1.0	1.1	1.2	1.3	
number of DF	2	1	0	0	0	0	0	3

B. Evaluation for scaling and JPEG compression

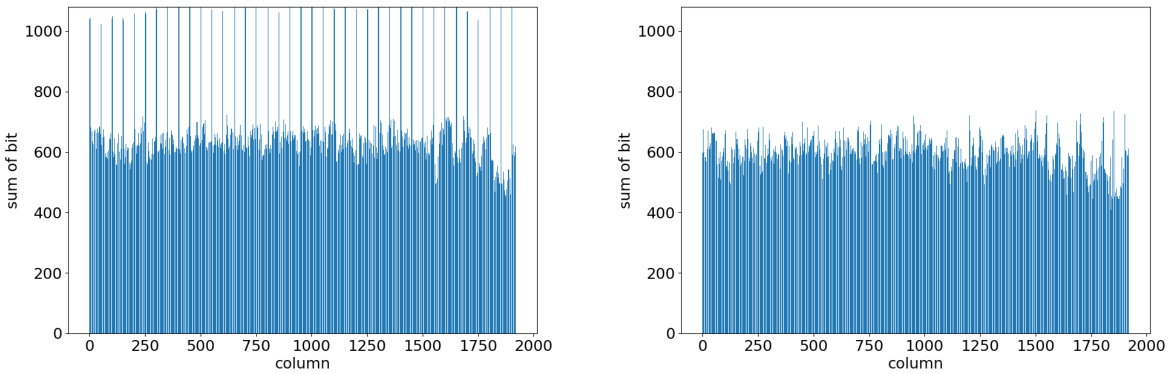
Let us evaluate the estimated magnification ratio $\hat{m}$ for the case where scaling and JPEG compression are applied. Twelve images were magnified with a magnification ratio $m = 0.7, 1.0, 1.3$ and were saved in the JPEG format, where the Q factors (QF) for JPEG compression were set from 10 to 100.

A histogram of the bit total in the 2nd bitplane of the compressed image is shown in Fig. 2 (b), where $QF = 55$. Since the values of the bitplane were changed by the compression, it was difficult to detect the pilot signal. The autocorrelation of the histogram was shown in Fig. 3 (b). Compared with the uncompressed case of (a), the peaks were smaller. We could see that components other than the pilot signal were included in the autocorrelation. To obtain the period from the autocorrelation, the DFT was performed as shown in Fig. 4 (b). As described in II-B, the grid interval $\hat{G}$ was obtained from the DFT coefficients, and then the magnification ratio (MR) $\hat{m}$ was estimated. Figure 7 shows the relative error R. The abscissa is Q factors QF and the ordinate is the error R with a box-and-whisker plot. We found that the estimated magnification ratio $\hat{m}$ could be correctly estimated for $QF \geq 60$. Note that a number of DF were excluded. Therefore, the error R was plotted when it could be estimated from at least one image. The number of DF for each QF are shown in Table II. There were 375 DF out of 684 cases. Most DF occurred at small QFs. In other words, the more the images were compressed, the more the detection failed.

IV. CONCLUSION

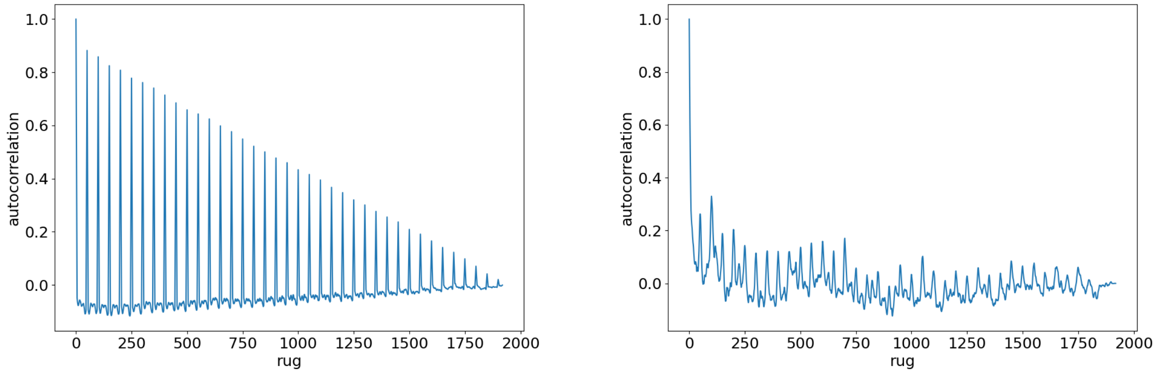
We proposed a method to estimate the magnification ratio by embedding a pilot signal. To counter scaling and compression attacks, the pilot signal was embedded in the second bitplane of the U component. Since the proposed method used a bitplane, it was vulnerable to JPEG compression. As a result of estimating the magnification ratio on the basis of the periodicity of the pilot signal, the method could estimate the magnification ratio almost accurately in the case of scaling only, and achieved 44.8 dB of image quality on average. No visual differences were observed. In the case of scaling and JPEG compression, the magnification ratio could be estimated almost accurately with QFs above 60.

The estimations of the magnification ratio failed especially in the small Q factor range. There were three causes for



(a) no compression (b) compressed with $QF = 55$

Fig. 2. Histogram of total of bits.



(a) no compression (b) compressed with $QF = 55$

Fig. 3. Autocorrelation of the histogram of Fig. 2

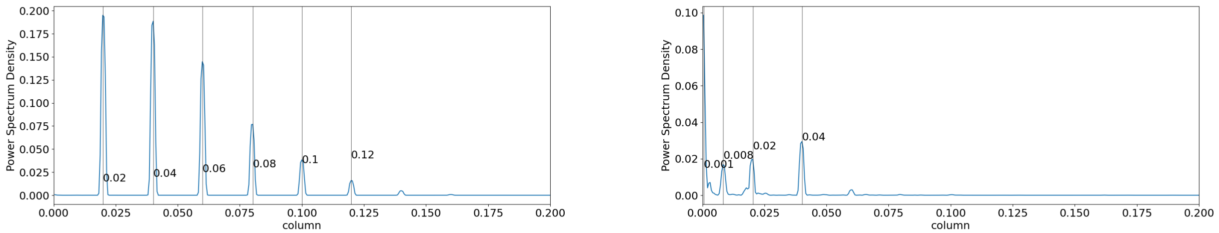
TABLE II
NUMBER OF DF FOR COMPRESSED IMAGES

		QF																			
		10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	total
MR	0.7	12	12	12	12	12	12	12	11	12	12	10	10	6	6	5	3	3	3	2	167
	1.0	12	12	11	10	11	10	12	7	6	4	3	3	3	3	3	3	3	2	0	118
	1.3	11	9	9	7	5	4	4	3	3	3	3	3	3	3	3	3	2	2	1	90

this failure. The first was the case where an image itself had periodicity. When there was a tile-like pattern in the image, the method detected its period. Figure 8 shows an example of a tile-like pattern, where (a) and (b) show a stego-image and its second bitplane, respectively. Since the lines of the building appeared periodically, a period other than that of the pilot signal might be detected. Therefore, the method resulted in an incorrect period detection.

The second was image degradation caused by scaling and JPEG compression. When images were shrunk, the grid width of the pilot signal became smaller. For example, when the magnification ratio was 70%, the grid width would be approximately $B' = 3.5$ pixels. Therefore, it became difficult to detect the pilot signal. For small QFs, the pilot signal was eliminated. In this case, since the image was also significantly degraded, this was not a serious problem.

The third was in the color space conversion. A RGB image was converted to a YUV image, and the watermark and the pilot signal were embedded. The image was then converted back to an RGB image. Since the pixel values were quantized, the pilot signal was corrupted at this time. Figure 9 shows an example corrupted by the conversion and (a) and (b) show the bitplanes immediately after embedding the pilot signal and after converting the color space from YUV to RGB, and then



(a) no compression

(b) compressed with $QF = 55$

Fig. 4. Power spectrum of the autocorrelation of Fig. 3

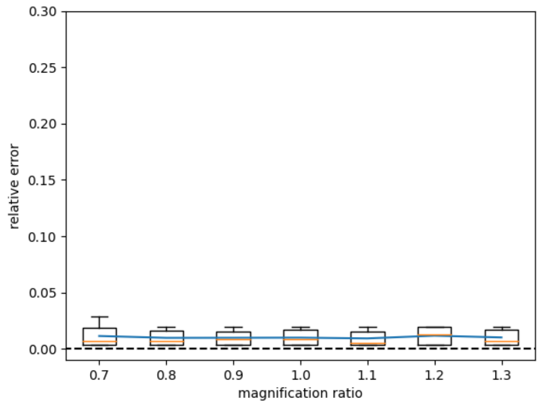


Fig. 5. Relative error R.

from RGB to YUV again. The pilot signal disappeared in the upper part of the figure. Future work includes a design of the pilot signal shape.

V. Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 20K11973. The computation was carried out using the super computer facilities at the Research Institute for Information Technology, Kyushu University.

References

[1] F.A.P. Petitcolas, R.J. Anderson, M.G. Kuhn, “Information Hiding-A Survey,” Proc. IEEE, vol. 87, no. 7, pp. 1062–1078, 1999

[2] H. Kang, K. Iwamura, “Watermarking based on the difference of discrete cosine transform coefficients and an error-correcting code,” Proc. of IWIHC, pp. 9–17, 2014

[3] N. Hirata, M. Kawamura, “Watermarking method using concatenated code for scaling and rotation attacks,” LNCS, Digital-Forensics and Watermarking, Springer, vol. 9569, pp. 259–270, 2016

[4] N. Hirata, T. Nozaki, M. Kawamura, “Image Watermarking Method Achieving IHC by Using PEG LDPC Code,” IEICE Trans. Inform. Syst., vol. E100-D, no. 1, pp. 13–23, 2017

[5] J.J.K.O. Ruanaidh, T. Pun, “Rotation, scale and translation invariant spread spectrum digital image watermarking,” Signal Process., vol. 66, Issue 3, pp. 303–317, 1998

[6] T. Yamamoto, M. Kawamura, “Method of spread spectrum watermarking using quantization index modulation for cropped images,” IEICE Trans. Inform. Syst., vol. E98-D, no. 7, pp. 1306–1315, 2015

[7] D.G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” Int. J. Comput. Vision, vol. 60, no. 2, pp. 381–389, 2017

[8] H. Luo, X. Sun, H. Yang, Z. Xia, “A robust image watermarking based on image restoration Using SIFT,” Radioengineering, vol. 20, no. 2, pp. 525–532, 2011

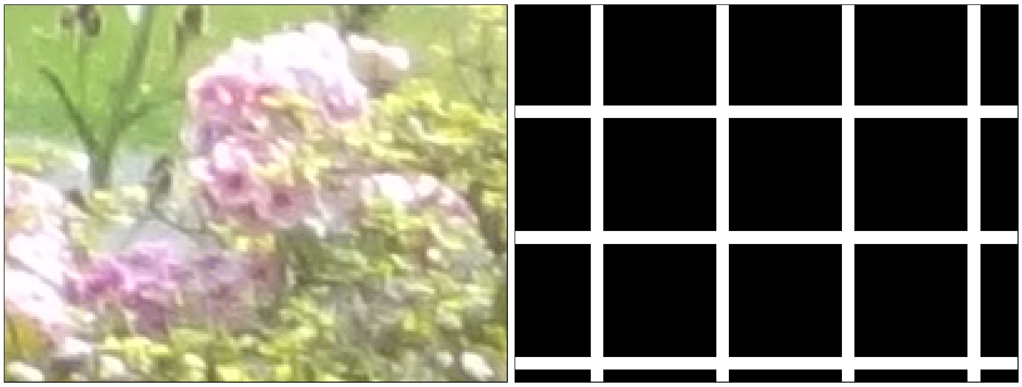
[9] L. Li, B. Guo, J. Pan, “Feature-based image watermarking resisting geometric attacks,” 3rd International Conference on Innovative Computing Information and Control, Dalian, Liaoning, p.18, 2008

[10] M. Kawamura, K. Uchida, “SIFT feature-based watermarking method aimed at achieving IHC Ver.5,” IIH-MSP 2017, Smart Innovation, Systems and Technologies, vol. 81. Springer, Cham, 2018

[11] V. Solachidis, I. Pitas, “Circularly symmetric watermark embedding in 2-D DFT domain,” IEEE Trans. Image Process, vol. 10, no. 11, pp. 1741–1752, 2001

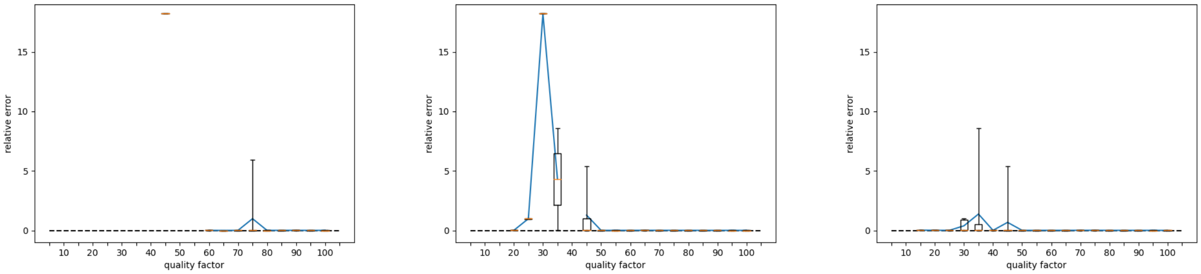
[12] J.S. Seo, C.D. Yoo, “Localized image watermarking based on feature points of scale-space representation,” Pattern Recognition, vol. 37, Issue 7, pp. 1365–1375, 2004

[13] S. Pereira, T. Pun, “Fast robust template matching for Affine resistant image watermarks,” Proc. 3rd Int. Information Hiding Workshop, pp. 207–218, 1999



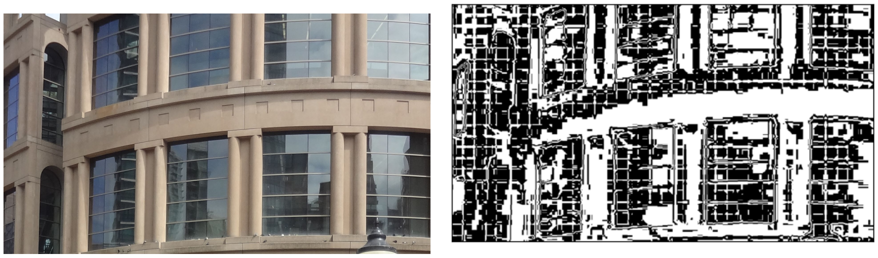
(a) stego-image (b) pilot signal

Fig. 6. Part of a stego-image and pilot signal.



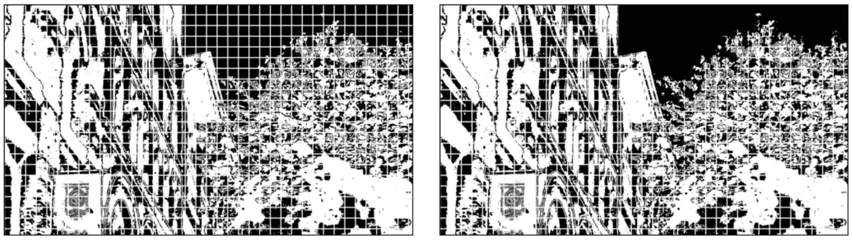
(a) $m = 0.7$ (b) $m = 1.0$ (c) $m = 1.3$

Fig. 7. Relative error R for compressed images.



(a) stego-image (b) bitplane

Fig. 8. Bitplane in the case of detection failure



(a) bitplane immediately after embedding (b) bitplane after converting color space

Fig. 9. Bitplane in the case of detection failure