Image Invisibility Assessment for Visually Encrypted Images

Akira UCHIDA, Masaaki FUJIYOSHI, Sayaka SHIOTA, and Hitoshi KIYA
Department of Information and Communication Systems, Tokyo Metropolitan University, Hino, Tokyo 191–0065, Japan
Email: uchida-akira@ed.tmu.ac.jp, mfujiyoshi@ieee.org, sayaka@tmu.ac.jp, kiya@sd.tmu.ac.jp

***Abstract*—A variety of perceptual encryption schemes have been studied to generate visually encrypted images for protecting visual secrecy and privacy while supporting ongoing innovation and growth in the applications of digital imagery. However, typical image quality metrics such as PSNR and SSIM are not suitable to evaluate the invisibility of visually encrypted images. To overcome this issue, this paper proposes a novel full-reference invisibility index, referred to as the Triple feature similarity index (TFSI), for visually encrypted images. The TFSI is computed through four steps: 1) edge extraction as preprocessing, 2) extracting the Triple feature based on the Trace transform, 3) normalization of resolution, and 4) TFSI is computed between the reference and visually encrypted images. Experimental results show that the TFSI measures the invisibility of visually encrypted images.**

Index Terms—MPEG–7, Biometrics, Privacy protection

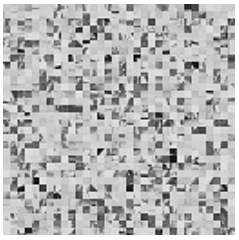
I. Introduction

Digital image is now ubiquitous, being used by governments, corporations, and individuals. While digital image is widespread, information leakage has been of concern. To overcome this issue, perceptual encryption [1]–[3] has been used to protect visual secrecy in images in various applications, such as face recognition [4]–[7], fingerprint matching [8]–[11], image trading system [12], [13], and so on. Perceptual encryption schemes generate visually encrypted images where the image provides no sufficient visual information of the original image.

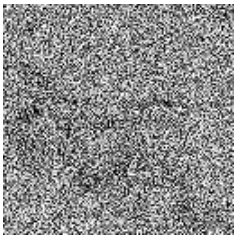
There are various types of perceptual encryption schemes, such as block scrambling [14], adding noise [15], phase scrambling [8], [10], [16]–[18], and so on, and the invisibility differs depending on the scheme. In addition, the required invisibility differs according to applications, e.g., a part of information is desired to be recognizable in image trading systems [12], [13]. Therefore, evaluation of the invisibility is required to confirm whether a specific visually encrypted image is applicable to the specified application.

Objective image quality metrics [19] can play a variety of roles in image processing applications. Various metrics, such as peak signal-to-noise ratio (PSNR), structural similarity (SSIM) [20], CIE2000 [21], and visible difference predictor (VDP) [22] are widely used. These are full-reference metrics in which a complete reference image is assumed to be known to compute the quality between the image and a distorted image. These metrics, however, are not suitable to evaluate the invisibility of visually encrypted images and there is no invisibility index to the best of the authors' knowledge.

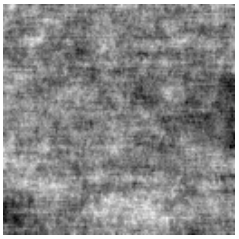
(a) Original image.



(b) Block scrambling ($N_x = N_y = 32$).



(c) Adding noise ($s = 500$).



(d) Phase scrambling (50%).

Fig. 1. Visually encrypted images.

This paper proposes a new paradigm for invisibility assessment based on the Triple feature [23], which is used in MPEG-7 for an image retrieval. This paper focuses on full-reference invisibility assessment. In the proposed invisibility index model, a feature is extracted from each image based on the Trace transform. The proposed invisibility index, referred to as the Triple feature similarity index (TFSI), is computed by comparing the feature of images. It is confirmed by comparing with other quality assessment models using various distorted images that the TFSI is suitable to evaluate the invisibility of visually encrypted images.

II. Preliminaries

This section introduces examples of perceptual encryptions, some well-known full-reference image quality metrics, and the Triple feature which is used in MPEG-7 for an image retrieval. It also summarizes the problems of conventional metrics and the goal of this paper.

(a) Original image. (b) Rotated. (b) Shifted.

(d) Visually encrypted (Scheme 1). (e) Visually encrypted (Scheme 2). (f) Visually encrypted (Scheme 3).

Fig. 2. Original image and distorted images.

2.1. Visually Encrypted Images

Visually encrypted images are required that the quality of visual information is partially degraded for secrecy protection, as shown in Fig.1, and three visual encryption schemes are given here.

Scheme 1 (Block Scrambling) [14]: An original image is divided into $N_x \times N_y$ of blocks and blocks are randomly permuted. The more the number of blocks becomes, the more invisible the image becomes, and the pixel scrambling is done by scrambling an image on a pixel-by-pixel basis. Figure 1 (b) shows an example of block scrambling where $N_x = N_y = 32$.

Scheme 2 (Adding Noise) [15]: A pseudo random noise matrix of range $[-s, s]$ is added to an original image. After adding the noise, the image is clipped at [0 .. 255] to make it difficult to restore the original image. Figure 1 (c) shows an example of adding noise where $s = 500$.

Scheme 3 (Phase Scrambling) [8], [10]: The phase component of discrete Fourier transformed (DFTed) coefficients of an original image are multiplied by a sign key consisting of random positive and negative signs, and applying inverse DFT to the processed coefficients gives the phase scrambled image. Figure 1 (d) shows an example of phase scrambling using the key with 50% of the negative signs. It is noted that discrete cosine transformation can be used instead of DFT.

2.2. Image Quality Metrics

There are many objective image quality metrics [19], such as PSNR, SSIM [20], CIE2000 [21], VDP [22], and so on, and each metric measures the quality of a distorted image by computing the similarity between a reference image and the distorted image. For example, the SSIM index of two images A and B is given as

$$SSIM(A,B) = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{A,B} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\sigma_A^2 + \sigma_B^2 + C_2)}, \quad (1)$$

TABLE I. PSNR, SSIM, and the proposed index (TFSI) between the image shown as Fig. 2 (a) and images shown in Fig. 2.

Image	(a)	(b)	(c)	(d)	(e)	(f)
PSNR	∞	11.87	15.85	10.79	13.35	15.73
MSSIM	1.00	0.30	0.37	0.30	0.081	0.55
TFSI	0.00	1.37	1.41	5.78	2.74	7.99

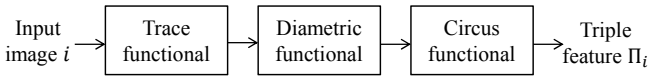


Fig. 3. The process of Triple feature extraction.

TABLE II. Triple feature of images shown in Fig. 2.

Image	(a)	(b)	(c)	(d)	(e)	(f)
Π_i	3.4E+05	4.9E+05	2.5E+05	6.4E+05	9.3E+06	4.8E+05

where μ_A, μ_B, σ_A^2, σ_B^2, and $\sigma_{A,B}$ are the mean of A, the mean of B, the variance of A, the variance of B, and the covariance of A and B, respectively, and $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$. In this paper, $k_1 = 0.01$, $k_2 = 0.03$, and $L = 8$. The SSIM index is a decimal value between -1 and 1, and higher value indicates higher similarity, i.e., better quality.

Between the original image (Fig. 2 (a)) and other images shown in Fig.2, the PSNR, SSIM, and the proposed index (TFSI) which will be proposed in Section III, are given as TableI. It is confirmed from TableI that PSNR and SSIM give low values regardless of the invisibility of evaluated images, viz., the PSNR and SSIM are not suitable to evaluate the invisibility of visually encrypted images. For the same reason, other ordinary metrics can not evaluate the invisibility of visually encrypted images.

2.3. Trace Transform

In contrast to that ordinary image quality metrics using the similarity between images are strongly affected by geometric transformations as shown in Section 2.3, the Triple feature used in MPEG-7 for an image retrieval is robust to geometric transformations thanks to the Trace transform [23]. Figure 3 shows the process of feature extraction using the Trace transform where three transformations, namely, trace functional, diametric functional, and circus functional, are performed on image i, and Triple feature Π_i is extracted.

In this paper, the following functionals are used.

$$\text{Trace functional} : \int f(x)dx, \quad (2)$$

$$\text{Diametric functional} : \int |f'(x)|\,dx, \quad (3)$$

$$\text{Circus functional} : \int f(x)dx. \quad (4)$$

Using these functionals, Triple feature Π_i becomes the value reflecting the gradient of the image. The Triple feature of the images shown in Fig. 2 is given in Table II. It is confirmed from Table II that the Triple feature based on the Trace transform is not sensitive to geometric transformation. It, however, does not reflect the invisibility

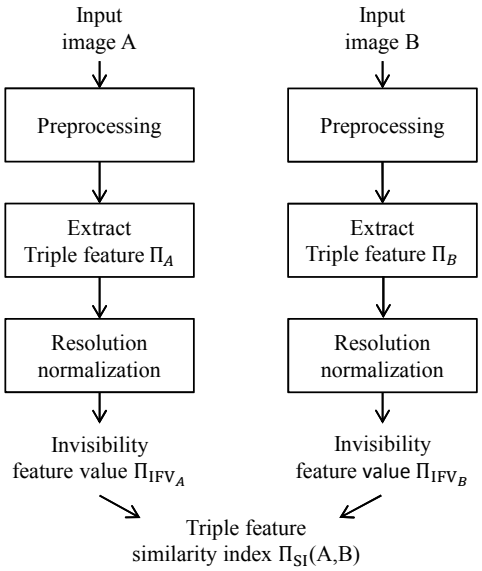


Fig. 4. The process of extracting invisibility feature value and the Triple feature similarity index.

of visually encrypted images into its value as no significant difference exits among Triple feature of geometric transformed images and visually encrypted images.

2.4. Summary of the Problems and the Goal of This Paper

Existing image quality metrics are sensitive to visually encrypted images as well as to geometric transformed images. Therefore, these metrics are not suitable to evaluate the invisibility of visually encrypted images. The goal of this paper is to propose a new metric to evaluate the invisibility of visually encrypted images regardless of geometric transformations.

III. Proposed Method

This section proposes an invisibility index based on the Triple feature for visually encrypted images. Figure 4 shows the block diagram to compute the proposed index, referred to as the Triple feature similarity index (TFSI), where three steps for deriving the invisibility feature value and one step for calculating the TFSI between the reference and visually encrypted images exist.

3.1. Invisibility Feature Value

This section describes three steps to derive invisibility feature value Π_{IFV}, namely, preprocessing, extracting the Triple feature, and resolution normalization.

3.1.1. Preprocessing: As the preprocessing, 1) noise suppression and 2) edge extraction are applied to an input image. First, an input image is down-sampled using the mean of non-overlapped $X_B \times Y_B$-sized blocks to suppress visibility-unrelated noise in the image, where $X_B = Y_B = 2$ in this paper. After that, the Gabor filter is performed to the image for emphasizing visibility-related edge parts. The Gabor filter is given as

$$g\left(x, y; \lambda, \theta, \psi, \gamma\right) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right), \tag{5}$$

where

$$x' = x\cos\theta + y\sin\theta, \tag{6}$$
$$y' = -x\sin\theta + y\cos\theta, \tag{7}$$

(x, y) are the filter coordinates, λ is the wave length, θ is the orientation of the normal to the parallel stripes of a Gabor function, ψ is the phase offset, γ is the spatial aspect ratio, and σ is the standard deviation of the Gaussian envelope. In this paper, It is set that $\lambda = 8$, $\theta = 0$, $\psi = [0, \pi/2]$, and $\gamma = 1.0$.

3.1.2. Extracting the Triple Feature: After preprocessing, Triple feature Π is computed by using the Trace transform described in Section II.

3.1.3. Resolution Normalization: Triple feature Π is proportional to the resolution, c.f., Eqs. (2), (3), and (4). Therefore, the resolution normalization of Triple feature Π considering the difference of the resolution needs to be done and is proposed as an invisibility feature value. The invisibility feature value is denoted by Π_{IFV} and is computed as

$$\Pi_{\mathrm{IFV}} = \frac{\Pi}{W \times H \times D}, \tag{8}$$

where W is the width of the image, H is the height of the image, and D is the bit depth of a pixel.

3.2. Triple Feature Similarity Index

The TFSI between reference image A and visually encrypted image B is denoted by $\Pi_{\mathrm{SI}}(A, B)$ and is computed as

$$\Pi_{\mathrm{SI}}(A, B) = 20\left|\log\frac{\Pi_{\mathrm{IFV}_B}}{\Pi_{\mathrm{IFV}_A}}\right|, \tag{9}$$

where Π_{IFV_A} and Π_{IFV_B} are the invisibility feature values derived from images A and B, respectively. TFSI $\Pi_{\mathrm{SI}}(A, B)$ equals to 0 when B is identical to A. As TFSI $\Pi_{\mathrm{SI}}(A, B)$ is higher, visually encrypted image B is more invisible.

IV. Experimental Results

4.1. Conditions

To evaluate the performance of the TFSI, thirty test images including facial images (192×256 pixels and 8 bits per pixel, i.e., $W = 192$, $H = 256$, and $D = 8$) and scenery images (256×256 pixels and 8 bits per pixel, i.e., $W = 256$, $H = 256$, and $D = 8$) were used. Each image is visually encrypted by three visual encryption schemes introduced in Section II. In addition, a rotation is applied to the image before it is encrypted by scheme 2 (adding noise [15]). PSNR, SSIM, and TFSI are computed between the original image and visually encrypted images.

TABLE III. THE AVERAGED RESULT FOR THIRTY TEST IMAGES.

(a) For visual encryption scheme 1 (block scrambling [14]).

Number of blocks	1 × 1 (original)	2 × 2	4 × 4	8 × 8	24 × 24	$N_x \times N_y$
PSNR	∞	12.71	11.81	11.39	11.39	11.40
MSSIM	1.00	0.41	0.29	0.22	0.087	0.023
TFSI	0.00	1.75	2.27	2.79	6.23	6.76
TFSI (w/o preprocessing)	0.00	2.72	3.24	3.48	3.89	5.69

(b) For visual encryption scheme 2 (adding noise [15]).

s	0 (original)	50	100	300	500
PSNR	∞	18.98	13.36	7.33	6.42
MSSIM	1.00	0.27	0.12	0.023	0.013
TFSI	0.00	1.60	2.15	6.20	6.43
TFSI (w/o preprocessing)	0.00	3.05	5.05	9.17	9.81

(c) For visual encryption scheme 3 (phase scrambling [8], [10].

Negative sign rate	0% (original)	10%	20%	30%	40%	50%
PSNR	∞	17.72	14.80	14.24	12.68	12.25
MSSIM	1.00	0.68	0.53	0.43	0.32	0.24
TFSI	0.00	2.72	4.43	5.63	6.01	6.18
TFSI (w/o preprocessing)	0.00	2.56	3.21	3.22	3.61	3.27

(d) For rotated and visually encryption scheme 2 (adding noise [15])

s	non-rotated (original)	0	50	100	300	500
PSNR	∞	15.87	14.34	11.83	7.29	6.50
MSSIM	1.00	0.34	0.060	0.022	0.0056	0.0046
TFSI	0.00	2.37	2.34	2.73	5.56	6.50
TFSI (w/o preprocessing)	0.00	5.44	2.85	4.11	8.64	9.66

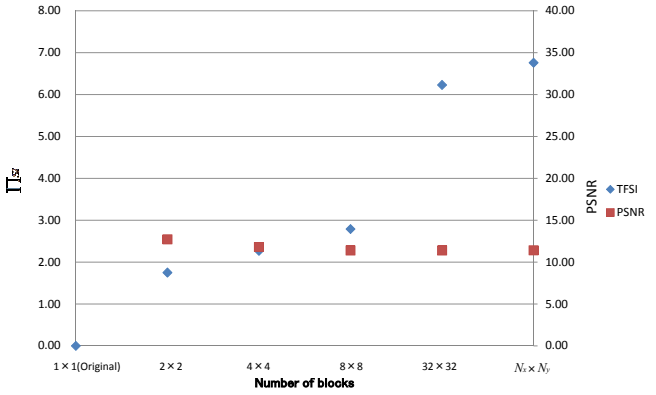
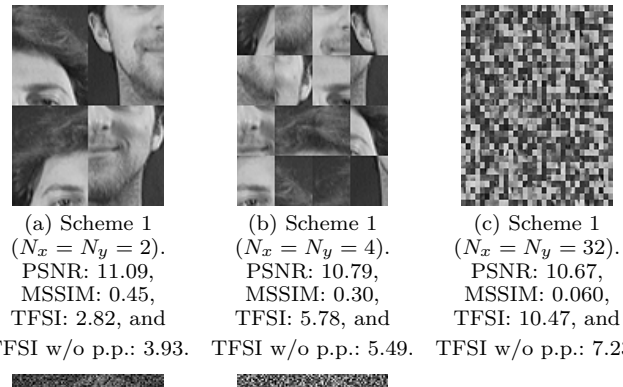


Fig. 5. The proposed index (TFSI) and PSNR averaged over thirty images encrypted by scheme 1 (block scrambling [14]).

4.2. Performance Evaluation

4.2.1. Comparison between TFSI and Other Metrics: Table III shows the metrics and index values averaged over thirty test images, and Fig. 5 shows the TFSI and PSNR for the image encrypted by scheme 1 (block scrambling [14]). Figure 6 shows some example of visually encrypted images and metrics and index values for images. It is confirmed that ordinary objective image quality metrics like PSNR and SSIM are not suitable to evaluate the invisibility of visually encrypted images, whereas the proposed TFSI reflects the invisibility of the images into its values.

4.2.2. The Effectiveness of the Preprocessing: It is noted that the preprocessing described in Section III helps the TFSI to reflect the invisibility of images into its values. Without the preprocessing, the TFSI indicates that images encrypted by scheme 2 (Figs. 6 (d) and (e)) are much invisible than images encrypted by scheme 1 (Figs. 6 (a), (b), and (c)), whereas Fig. 6 (d) is visible as much as Figs. 6 (a) and (b). With the preprocessing, the TFSI adequately reflects the invisibility of images into its values.



(a) Scheme 1 ($N_x = N_y = 2$). PSNR: 11.09, MSSIM: 0.45, TFSI: 2.82, and TFSI w/o p.p.: 3.93.

(b) Scheme 1 ($N_x = N_y = 4$). PSNR: 10.79, MSSIM: 0.30, TFSI: 5.78, and TFSI w/o p.p.: 5.49.

(c) Scheme 1 ($N_x = N_y = 32$). PSNR: 10.67, MSSIM: 0.060, TFSI: 10.47, and TFSI w/o p.p.: 7.23.

(d) Scheme 2 ($s = 100$). PSNR: 13.35, MSSIM: 0.081, TFSI: 2.74, and TFSI w/o p.p.: 8.65.

(e) Scheme 2 ($s = 300$). PSNR: 7.34, MSSIM: 0.017, TFSI: 8.88, and TFSI w/o p.p.: 14.18.

Fig. 6. Examples of visually encrypted images and metrics and index values for images. (TFSI w/o p.p. is the TFSI without preprocessing).

V. CONCLUSIONS

This paper has proposed a novel invisibility index based on Triple feature for visually encrypted images. It is confirmed that the proposed index, TFSI, indicates the invisibility of visually encrypted images whereas ordinary objective image quality metrics like PSNR and SSIM are not suitable to evaluate the invisibility of visually encrypted images. In order to confirm the effectiveness of the TFSI, we will perform the subjective evaluation of invisibility in the future.

REFERENCES

[1] X.Zhang, "Separable reversible data hiding in encrypted image," *IEEE Trans. Inf. Forensics Security*, vol.7, pp.826–832, Apr. 2012.

[2] H.Kiya, S.Imaizumi, and O.Watanabe, "Partial-scrambling of images encoded using JPEG2000 without generating marker codes," in *Proc. IEEE ICIP*, pp.III-205–III-208, 2003.

[3] O.Watanabe, A.Nakazaki, and H.Kiya, "A fast image-scramble method using public-key encryption allowing backward compatibility with JPEG2000," in *Proc. IEEE ICIP*, pp.3435–3438, 2004.

[4] A.B.J.Teoh, A.Goh, and D.C.L.Ngo, "Random multispace quantization as analytic mechanism biohashing of bio metric and random identity inputs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28, pp.1892–1901, Dec. 2006.

[5] A.B.J.Teoh and C.T.Yuang, "Cancelable biometrics realization with multispace random projections," *IEEE Syst., Man, Cybern. B*, vol.37, pp.1096–1106, Oct. 2007.

[6] A.Ross and A.Othmen, "Visual cryptography for biometric privacy," *IEEE Trans. Inf. Forensics Security*, vol.6, pp.70–81, Mar. 2011.

[7] S.Z.Li and A.K.Jain, Eds., *Handbook of Face Recognition*. New York: Springer Verlag, 2011.

[8] H.Kiya and I.Ito, "Image matching between scrambled images for secure data management," in *Proc. EURASIP EUSIPCO*, Aug. 2008.

[9] I.Ito and H.Kiya, "One-time key based phase scrambling for phase-only correlation between visually protected images," *EURASIP J. Inf. Security*, vol.2009, Jan. 2009.

[10] I.Ito and H.Kiya, "Phase scramble for blind image matching," in *Proc. IEEE ICASSP*, pp.1521–1524, 2009.

[11] I.Ito and H.Kiya, "A new class of image registration for guaranteeing secure data management," in *Proc. IEEE ICIP*, pp.269–272, 2008

[12] S.Liu, M.Fujiyoshi, and H.Kiya, "An image trading system using amplitude-only images for privacy- and copyright-protection," *IEICE Trans. Fundamentals*, vol.E96-A, pp.1245–1252, Jun. 2013.

[13] W.Sae-tang, S.Liu, M.Fujiyoshi and H.Kiya, "1D frequency transformation-based amplitude-only images for copyright- and privacy-protection in image trading systems," *ECTI Trans. Comput. and Inf. Tech.* vol.8, no.2, pp.98–107, Nov. 2014.

[14] N.K.Ratha, J.H.Connell, and R.M.Bolle, "Enhancing security and privacy in biometrics-based authentication systems," *IBM Syst. J.*, vol.40, pp.614–634, Mar. 2001.

[15] M.Furukawa, Y.Muraki, M.Fujisyoshi, and H.Kiya, " A secure face recognition scheme using noisy images based on kernel sparse representation" in *Proc. APSIPA ASC*, 2013.

[16] K.Ando, O.Watanabe and H.Kiya, "Partial-scrambling of images encoded by JPEG2000," *IEICE Trans. Inf. and Sys.*, vol.E85-D(2), pp.439, Feb. 2002

[17] O.Watanabe, A.Nakazaki and H.Kiya, "A scalable encryption method allowing backward compatibility with JPEG2000 images," in *Proc. IEEE ISCAS*, pp.6324–6327, 2005.

[18] K.Kurokawa, M.Fujiyoshi and H.Kiya, "Codestream domain scrambling of moving objects based on DCT sign-only correlation for motion JPEG movies," in *Proc. ICIP*, pp.157–160, 2007.

[19] T.Liu, Y.-C.Lin, and C.-C.J.Kuo, "Visual quality assessment; recent developments, coding applications and future trends," *APSIPA Trans. Signal Inf. Process.*, vol.2, Jan. 2013.

[20] Z.Wang, A.Bovik, H.Sherikh, and E.Simoncelli,"Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol.13, pp.600–612, Apr. 2004.

[21] G.Sharma, W.Wu, and E.N.Dalal, "The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations," *Color Reserach and application*, vol.30, pp.21–30, Feb. 2005.

[22] S.Daly, "The visible difference predictor; an algorithm for the assessment of image fidelity," in *Digital Images and human Vision*, A.B.Watson, Ed. Cambridge, Ma: MIT Press, 1993.

[23] A.Kadyrov and M.Petrou, "The Trace transform and its application," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.23, pp.811–828, Aug. 2001.