AN ENCRYPTION-THEN-COMPRESSION SYSTEM FOR JPEG 2000 STANDARD

Osamu WATANABE†, Akira UCHIDA‡, Takahiro FUKUHARA and Hitoshi KIYA‡*

†Takushoku University, Dpet. of Electronics & Computer Systems, 815–1, Tatemachi, Hachioji-shi, Tokyo, JAPAN
‡Tokyo Metropolitan University, Faculty of Info. and Commun. Systems, 6–6 Asahigaoka, Hino-shi, Tokyo, JAPAN
* Sony Corporation, CPDG PSG, 4–14–1, Asahi-cho, Atsugi-shi, Kanagawa, JAPAN

ABSTRACT

A new Encryption-then-Compression (ETC) system for the JPEG 2000 standard is proposed in this paper. An ETC system is known as a system that makes image communication secure and efficient by using perceptual encryption and image compression. The proposed system uses the sign-scrambling and block-shuffling of discrete wavelet transform (DWT) coefficients as perceptual encryption. Unlike conventional ETC systems, the proposed system is compatible with the JPEG 2000 standard because the perceptually encrypted coefficients can be efficiently compressed by the JPEG 2000. The experimental results demonstrated that the proposed system achieved both acceptable compression performance and enough key-space for secure image communication while remaining compatible with the JPEG 2000 standard.

Index Terms— Encryption-then-compression, perceptual encryption, JPEG 2000

1. INTRODUCTION

Many methods of protecting compressed multimedia content have been reported with the wide/rapid spread of distributed systems for information processing, such as cloud computing and social networks. Communication and processing for multimedia content is performed over the Internet in these systems. In other words, the content is transmitted/received via non-secure telecommunication channels with restricted bandwidth. Therefore, both encryption and compression are necessary to make the communication secure and efficient. In the meantime, the JPEG committee has started to standardize a new work item, which is referred to as *JPEG Privacy* [1, 2]. Secure transmission between network servers that are used in cloud computing and social networks is supposed to be one of the technical requirements of JPEG Privacy.

Image encryption has to be performed anterior to image compression in certain practical scenarios for such distributed systems, e.g., image communication with security/privacy considerations. This framework is known as the *Encryption-then-Compression (ETC)* system [3, 4]. ETC systems have several advantages over traditional solutions called *Compression-then-Encryption (CTE)* systems: one is that image compression can be done more efficiently, another is that it is not necessary for an owner of images to disclose them to network providers. In most cases, such ETC systems adopt a perceptual encryption scheme that is known as an operation for making it difficult to understand images visually and is performed in both spatial and frequency-transformed domains [5–9]. This is although most studies on ETC systems have assumed the use of their own compression schemes that had no compatibility with international standards such as JPEG or JPEG 2000 [10–14]. The use of an international standard for image compression is one of the essential factors in its wide acceptance considering practical application scenarios for ETC systems.

The JPEG 2000 [15] is well known as the international standard for image compression. There have been many studies related to the encryption of JPEG 2000 images [16–18]; however, most of these were designed for CTE systems. A new ETC system with the assumption of using the JPEG 2000 is proposed in this paper. JPEG 2000 compliant perceptual encryption schemes have also been proposed in the proposed system. Perceptual encryption is achieved by scrambling the sign of discrete wavelet transform (DWT) coefficients and/or shuffling blocks defined in these coefficients. Sign-scrambling and block-shuffling are performed based on random permutation with secret keys. The experimental results demonstrated that the proposed system achieved smaller distortion caused by the JPEG 2000 compression than conventional ETC systems that had no compatibility with the JPEG 2000. Moreover, the results revealed that the parameters for perceptual encryption for the JPEG 2000 compliant ETC system should be carefully determined to enable acceptable compression.

2. SECURE IMAGE COMMUNICATION SYSTEMS WITH IMAGE COMPRESSION

Let us suppose a scenario in which content owner Alice wants to securely and efficiently transmit image I to recipient Bob over insecure and band-restricted communication channel provider Charlie. CTE and ETC systems are described in what follows.

2.1. Image communication with CTE system

Image communication considering security with a CTE system is summarized in what follows. First, I is compressed into I_c and then I_c is encrypted into I_{ce} using encryption function $E(\cdot)_k$ with secret key k. I_{ce} is transmitted to Bob via the channel provided by Charlie. The received I_{ce} is sequentially decrypted and decompressed, and then Bob obtains a reconstructed image $\hat{I}$.

This work was supported by JSPS KAKENHI Grant Number 25730073.

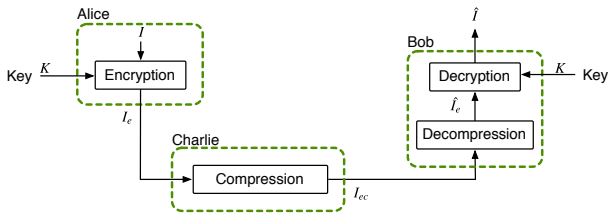


Fig. 1: Image communication with Encryption-then-Compression (ETC) system

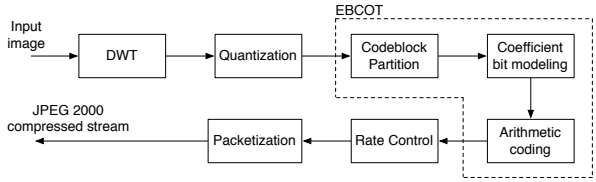


Fig. 2: Block diagram of JPEG 2000 encoder

The three problems with this CTE system are summarized as:

- Image I should be disclosed to Charlie. This conflicts with the demands of Alice.
- If channel error is added to I_{ce}, Bob cannot decrypt I_{ce} because $E(\cdot)_k$ is generally a number theory based cipher, such as advanced encryption standard (AES) or data encryption standard (DES).
- Charlie cannot control the coding rate of I_{ce} to meet the demands of the channel, e.g., efficient usage of the network bandwidth.

2.2. Image communication with ETC system

Image communication with an ETC system is outlined in Figure 1. Alice encrypts I into I_e then I_e is transmitted to Bob over the channel provided by Charlie. Charlie compresses I_e into I_{ec} according to the band-restrictions of the channel. The received I_{ec} is decompressed and then decrypted. Finally, Bob gets reconstructed image $\hat{I}$. A method of perceptual encryption in an ETC system is used as encryption function $E(\cdot)_k$ instead of an AES-like cipher. Perceptual encryption makes image I difficult to understand visually so that it is possible to decrypt I_{ec} with communication error.

An ETC system using perceptual encryption has three main advantages over a CTE system:

- Alice does not need to disclose image I to Charlie.
- Decryption of I_{ec} with communication error is possible.
- It is possible for Charlie to control the coding rate to maximize network utilization.

2.3. JPEG 2000 coding

Figure 2 is a block diagram of a JPEG 2000 encoder and Figure 3 is an example of an image analyzed with DWT. The R denotes three DWT resolution levels in Fig. 3, and r = 1, 2, ..., R is the index of the resolution level. The b denotes the indices of subbands, where 1, 2, 3, and 4 correspond to LL, HL, LH, and HH. The subbands are divided into $c_x \times c_y$ -sized code-blocks. The default value for c_x and c_y is 64. DWT coefficients in the code-blocks are quantized and then are encoded by embedded block coding with optimized truncation (EBCOT). Each quantized coefficient in EBCOT is separated into its sign and absolute magnitude, where the absolute magnitudes are factorized as bit-planes from the most significant bit (MSB) to the least significant bit (LSB). All the samples in the bit-planes are either zero or one. Then, bit-modeling and arithmetic coding is performed followed by rate control operations. Rate control is used to make the bitstream conform to a target size. Finally, a JPEG 2000 compliant bitstream is generated by adding packet headers, a main header, and other control codes.

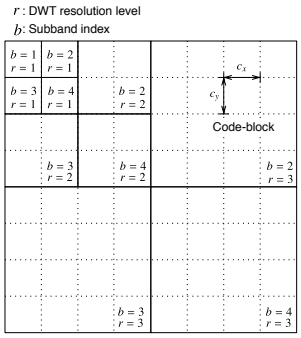


Fig. 3: Definition of code-blocks, subbands, and resolution levels on DWT coefficients (decomposition level R = 3)

3. PROPOSED ETC SYSTEM FOR JPEG 2000

Here, a new ETC system for the JPEG 2000 is proposed, which includes three types of perceptual encryption schemes. Each type of perceptual encryption can be changed according to the demands of an image owner. First, the overall procedure for the proposed system is described, and then the perceptual encryption schemes are explained.

3.1. Procedure for generating perceptual encrypted images

The procedure to generate the perceptual encrypted images is outlined in Fig. 4 and is summarized in four steps (A-D). Note that input image I is supposed to be $X \times Y$ in size.

(A) The JPEG 2000 compliant DWT is applied to I. Then, the DWT coefficients, C, which are $X \times Y$ in size, are obtained.

(B) DWT coefficients C are perceptually encrypted into C_e with two types of encryption schemes based on a pseudo-random number generator (PRNG) with a secret key. Detailed descriptions of each of the schemes are described in the next subsection.

(C) Encrypted DWT coefficients C_e are transformed to a spatial image by using inverse DWT.

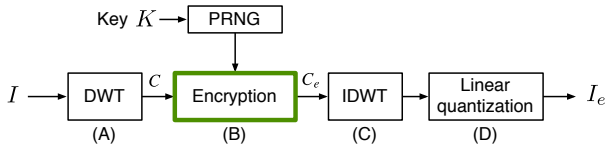
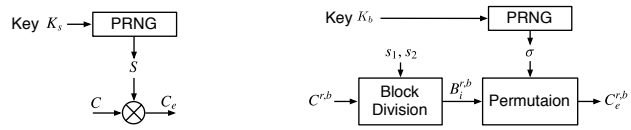


Fig. 4: Procedure of generating a perceptual encrypted image



(a) Sign-scrambling (b) Block-shuffling

Fig. 5: Procedures of perceptual encryption schemes

(D) The spatial image is linear-quantized to maintain the signal range of the input image, and then perceptually encrypted image I_e is obtained.

3.2. Methods of perceptual encryption

3.2.1. Sign-scrambling of DWT coefficients

The sign-scrambling of DWT coefficients is described and the procedure for this is outlined in Fig. 5(a) as a perceptual encryption scheme for the proposed ETC system. Let K_s denote a secret key for encryption and S denote a pseudo-random matrix. The S is generated by using K_s as its seed and $S(i, j)$ with $(1 \le i \le X, 1 \le j \le Y)$, which means an element of S, consists of "1" or "−1". Equation (1) has an example of S as:

$$S = \begin{pmatrix} -1 & 1 & \cdots & 1 \\ 1 & -1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & \cdots & 1 \end{pmatrix}. \quad (1)$$

The DWT coefficients $C(i, j)$ are scrambled by:

$$C_e(i, j) = C(i, j)S(i, j). \quad (2)$$

Note that the ratio of "−1" to "1" can be controlled. If Alice wants to get a fully scrambled image, the ratio should be 1:1.

3.2.2. Block-shuffling of DWT coefficients

The block-shuffling of DWT coefficients is described as another perceptual encryption scheme. The procedure for this is given in Fig. 5(b). First, DWT coefficients C are divided into blocks that are $s_1 \times s_2$ in size. Suppose i is the index of the blocks in subband and a $B_i^{r,b}$ is the i-th block in subband b of DWT resolution level r. The permutation, σ, for the block-shuffling is defined as a two-line notation:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & i & \cdots & n \\ x_1 & x_2 & \cdots & x_i & \cdots & x_n \end{pmatrix}, \quad (3)$$

where the first row represents the elements of i and the second row, which is $(x_1, x_2, \ldots, x_n)$, is a sequence generated by sequentially taking a pseudo-random number between 1 and n, ensuring that there are no repetitions. A block, $B_i^{r,b}$, is set to the position having its index equal to x_i in the block-shuffled subband. Note that a secret key, K_b, is used as a seed for generating the pseudo-random numbers.

3.2.3. Combination of sign-scrambling and block-shuffling

Combined scrambling is easily achieved by sign-scrambling followed by the block-shuffling. Let $\hat{C}_e^{r,b}$ denote the output of sign-scrambling in subband b of DWT resolution level r. The $\hat{C}_e^{r,b}$ is input into the following the block-shuffling. Then, the block-shuffled sign-scrambling output, $C_e^{r,b}$, is obtained.

3.3. Security analysis

There are several kinds of attacks on encryption, such as brute-force, differential, and statistical attacks. The key spaces of encryption schemes should be large enough to make such attacks infeasible. We evaluated the safety of the proposed encryption schemes with these key spaces assuming brute-force attacks.

The sign of a DWT coefficient in the proposed sign-scrambling is scrambled into "1" or "-1". Thus, the size of key-space N_{sign}in sign-scrambling is written as:

$$N_{sign} = 2^{XY}, \quad (4)$$

where X and Y are the horizontal and the vertical sizes of the original image. Let us suppose that I has a size of 1024×768. In this case, $N_{sign} = 2^{1024\times768}$. This key-space size is larger than that of an AES-like cipher with a 256-bit key.

The number of blocks n in block-shuffling is approximately equal to $\lfloor \frac{X}{s_1} \rfloor \times \lfloor \frac{Y}{s_2} \rfloor$. The number of N_{block} permutations is written as:

$$N_{block} = n! \,. \quad (5)$$

If I has a size of 1024×768 and $s_1 \times s_2 = 64 \times 64$, the number of blocks $n = 192$. Therefore, it is clear that the key-space for block-shuffling is large enough because $192! > 2^{256}$.

The size of key-space N_c in the combined scramble is

$$N_c = N_{sign} N_{block} = 2^{XY} n! \,. \quad (6)$$

4. EXPERIMENTAL RESULTS

This section presents the experimental results obtained from evaluating the performance of the proposed ETC system.

4.1. Conditions for experiments

Kakadu version 6.3 [19] was used as a codec for the JPEG 2000 standard. The test images were taken from standard evaluation material (StEM) test data [20]. Each of the test images consisted of RGB color space, and had a resolution of 4096×1714. The bit-depth of single color components was 12 bits/pixel. Five different frames (#2254, #3387, #6482, #8132, and #9924) were used. Four cases for the proposed encryption schemes were investigated in the experiments.

Sign: Sign-scrambling with "1"-to-"-1" ratio of 1:1.

BS $n \times n$: Block-shuffling with $s_1 = s_2 = n$. Two, eight, and ten were used as values for n.

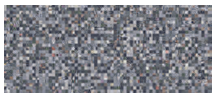
Table 1: FSIM index of encrypted image I_e

Frame #	Sign	BS 2x2	BS 8x8	BS 10x10	BS (LL)	Mixed
2254	0.449	0.451	0.489	0.507	0.455	0.428
3387	0.564	0.511	0.504	0.520	0.509	0.493
6482	0.569	0.504	0.530	0.525	0.512	0.485
8132	0.373	0.390	0.496	0.518	0.383	0.347
9924	0.529	0.518	0.569	0.571	0.515	0.467
Average	0.497	0.475	0.517	0.528	0.475	0.444

(a) Original

(b) Sign

(c) BS8x8



(d) BS(LL)

(e) Mixed

Fig. 6: Example of encrypted images: (a) Original image from StEM (Frame #2254) (b) Sign-scrambling with "1"-to-"-1" ratio of 1:1. (c) Block-shuffling with $s_1 = s_2 = n$. As values of n, 2, 8, and 10 were used. (d) For the lowest subband LL, block-shuffling with $s_1 = s_2 = 2$. For the others, block-shuffling with $s_1 = s_2 = 8$. (e) Combination of sign-scrambling and block-shuffling with $s_1 = s_2 = 8$.

BS (LL): Block-shuffling with $s_1 = s_2 = 2$ was done for lowest subband LL. Block-shuffling with $s_1 = s_2 = 8$ was done for the others.

Mixed: Combination of **Sign** and **BS** 8×8.

4.2. Results and remarks

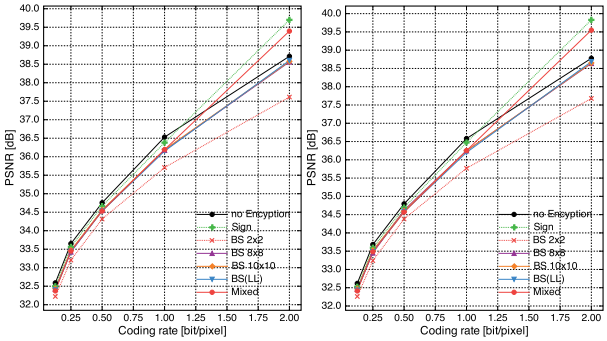
4.2.1. Efficiency of encryption

Perceptual encryption makes an image difficult to understand visually. The feature similarity (FSIM) index [21] between I and I_e was calculated for all these cases to confirm this property for perceptual encryption we propose. The FSIM indices ranged from zero to one. If a processed image had an FSIM index of less than 0.5, it was very hard to identify the processed image to be the same as the original because an FSIM index of less than 0.5 provided Zhang et al. a very small subjective MOS score that was close to zero [21].

The calculated FSIM indices are summarized in Table 1 and examples of I_e are given in Fig.6. This table and figure confirmed that perceptual encryption we propose is sufficiently capable of making an image unrecognizable.

4.2.2. JPEG 2000 compression

Figure 7 plots the results for compression by using the proposed ETC system. Two different code-block sizes, that were 32×32 and 64×64, were used and the results obtained from a PSNR evaluation of the average test images between reconstructed image $\hat{I}$ and original image I are given in Figs. 7(a) and 7(b). The *BS* 2*x*2 performed the worst in Figs 7(a) and 7(b); however, better quality in reconstructed image $\hat{I}$



(a) Codeblock size 32x32 (b) Codeblock size 64x64

Fig. 7: Comparison of the coding rate-PSNR performance: Each of curves is an average of PSNR values for all test images. There is no significant difference between *no-Encryption* and the others in terms of the value of PSNR without *BS 2x2* case.

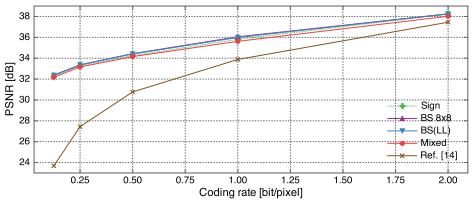


Fig. 8: The average PSNR comparison between the proposed method and the conventional method of Ref. [14].

could be obtained by using larger sizes for s_1 and s_2. We also tested and confirmed that there were no significant differences between *Sign* and the others in terms of the value of PSNR. Using *Mixed* was a good choice because the combination of sign-scrambling and block-shuffling was more secure against brute-force attacks. It was also confirmed that the scalability function of the JPEG 2000 standard, which is known to be an additional feature of the standard, was retained when using the proposed ETC system.

Figure 8 plots the average PSNR comparison between the proposed method and the conventional method of Ref. [14]. To fit the distributed code of the conventional method, test images were converted grayscale and were quantized into 8 bits/pixel. The reconstructed images obtained by using the proposed method provided better quality than that of the conventional method.

5. CONCLUSIONS

This paper proposed a new Encryption-then-Compression (ETC) system for the JPEG 2000 standard. Sign-scrambling and block-shuffling that are compatible with the JPEG 2000 standard were used as the perceptual encryption schemes in the proposed system. The experimental results demonstrated that the proposed system achieved both acceptable compression and sufficient security for secure image communication while maintaining the compatibility with the JPEG 2000 standard. We intend to extend the proposed ETC system to lossless coding of the JPEG 2000 standard in future work.

6. REFERENCES

[1] "Use cases and requirements for JPEG Privacy," ISO/IEC JTC 1/SC 29/WG 1 N6402, Jul. 2013.

[2] P. Schelkens, "Image security tools for JPEG standards," in *Proceedings of the 2nd ACM workshop on Information hiding and multimedia security - IH&MMSec '14*. ACM Press, Jun. 2014, pp. 1–1.

[3] Z. Erkin, A. Piva, S. Katzenbeisser, R. L. Lagendijk, J. Shokrollahi, G. Neven, and M. Barni, "Protection and retrieval of encrypted multimedia content: When cryptography meets signal processing," *EURASIP Journal on Information Security*, vol. 2007, no. 3, 2007.

[4] N. Kalyani G. and S. Milind V., "Article: A survey based on designing an efficient image encryption-then-compression system," *IJCA Proceedings on National Level Technical Conference X-PLORE 2014*, vol. XPLORE2014, pp. 6–8, May 2014, full text available.

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

[6] H. Kiya and I. Ito, "Image matching between scrambled images for secure data management," in *Proc. 16th European Signal Process. Conf*, 2008.

[7] I. Ito and H. Kiya, "One-time key based phase scrambling for phase-only correlation between visually protected images," *EURASIP Journal on Information Security*, vol. 2009, p. 3, 2009.

[8] W. Zeng and S. Lei, "Efficient frequency domain selective scrambling of digital video," *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 118–129, Mar. 2003.

[9] Z. Tang, X. Zhang, and W. Lan, "Efficient image encryption with block shuffling and chaotic map," *Multimedia Tools and Applications*, pp. 1–20, 2014.

[10] M. Johnson, P. Ishwar, V. Prabhakaran, D. Schonberg, and K. Ramchandran, "On compressing encrypted data," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2992–3006, 2004.

[11] W. Liu, W. Zeng, L. Dong, and Q. Yao, "Efficient compression of encrypted grayscale images," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 1097–1102, 2010.

[12] X. Zhang, "Lossy compression and iterative reconstruction for encrypted image," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 53–58, 2011.

[13] R. Hu, X. Li, and B. Yang, "A new lossy compression scheme for encrypted gray-scale images," in *Proc. ICASSP 2014*. IEEE, 2014, pp. 7387–7390.

[14] J. Zhou, X. Liu, O. C. Au, and Y. Y. Tang, "Designing an Efficient Image Encryption-Then-Compression System via Prediction Error Clustering and Random Permutation," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 1, pp. 39–50, Jan. 2014.

[15] "Information technology — JPEG 2000 image coding system – Part 1: Core coding system," International Standard ISO/IEC IS-15444-1, Dec. 2000.

[16] D. Engel, T. Stütz, and A. Uhl, "A survey on JPEG2000 encryption," *Multimedia Systems*, vol. 15, no. 4, pp. 243–270, Jan. 2009.

[17] H. Kiya, S. Imaizumi, and O. Watanabe, "Partial-scrambling of images encoded using JPEG2000 without generating marker codes," in *Proc. ICIP 2003*, vol. 2. IEEE, Sep. 2003, pp. III–205–8.

[18] O. Watanabe, T. Iida, T. Fukuhara, and H. Kiya, "Identification of JPEG 2000 images in encrypted domain for digital cinema," in *Proc. ICIP 2009*. IEEE, Nov. 2009, pp. 2065–2068.

[19] kakadu software. [Online]. Available: http://www.kakadusoftware.com

[20] Digital Cinema Initiatives, LLC Technology Committee. (2010, Sep.) StEM Access Procedures. [Online]. Available: http://www.dcimovies.com/StEM/

[21] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment." *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–86, Aug. 2011.