Encryption-Then-Compression Systems Using Grayscale-Based Image Encryption for JPEG Images

Tatsuya Chuman, Warit Sirichotedumrong, Student Member, IEEE, and Hitoshi Kiya, Fellow, IEEE

Abstract—A block scrambling-based encryption scheme is presented to enhance the security of Encryption-then-Compression (EtC) systems with JPEG compression, which allow us to securely transmit the images through an untrusted channel provider, such as social network service providers. The proposed scheme enables the use of a smaller block size and a larger number of blocks than the conventional scheme. Images encrypted using the proposed scheme include less color information due to the use of grayscale images even when the original image has three color channels. These features enhance security against various attacks such as jigsaw puzzle solver and brute-force attacks. In an experiment, the security against jigsaw puzzle solver attacks is evaluated. Encrypted images were uploaded to and then downloaded from Facebook and Twitter, and the results demonstrated that the proposed scheme is effective for EtC systems.

Index Terms—Image encryption, jigsaw puzzle, EtC system, JPEG.

I. INTRODUCTION

The use of images and video sequences has greatly increased because of the rapid growth of the Internet and widespread use of multimedia systems. While many studies on secure, efficient, and flexible communications have been reported [1]–[4], full encryption with provable security (like RSA and AES) is the most secure option for securing multimedia data. However, there is a trade-off between security and other requirements such as low processing demand, bitstream compliance, and signal processing in the encrypted domain. Several perceptual encryption schemes have been developed to achieve this trade-off [5]–[10].

Image encryption prior to image compression is required in certain practical scenarios such as secure image transmission through an untrusted channel provider. Encryption-then-Compression (EtC) systems [3], [11], [12] are used in such scenarios. In this paper, we focus on EtC systems although the traditional way of securely transmitting images is to use a compression-then-encryption (CtE) system. Most studies on EtC systems assumed the use of a proprietary compression scheme incompatible with international compression standards such as JPEG [3], [13]–[17], so they cannot be applied to social media. For example, a block scrambling-based approach and a singular value decomposition transformation are used to efficiently compress encrypted images in [3] and [17], respectively. Because of such a situation, block scrambling-based image encryption schemes, which are compatible with international standards, have been proposed for EtC systems [18]–[22]. However, the conventional block scrambling-based encryption schemes have a limitation on block size to prevent JPEG distortion.

In this paper, we present a novel block scrambling-based image encryption scheme for EtC systems that enhances security compared with the conventional scheme. Compared with the conventional schemes, for which robustness against several attacks such as jigsaw puzzle and brute-force attacks has been discussed [23], [24], the proposed one enables the use of a smaller block size and a larger number of blocks, which enhances both invisibility and security against several attacks. Furthermore, images encrypted by using the proposed scheme include less color information due to the use of grayscale images, which makes the EtC system more robust. Although EtC systems can be applied to social media by using JPEG images [25], there is a limitation on block size to prevent JPEG distortion due to recompression forced by social media. The proposed scheme relaxes this limitation.

An evaluation of the proposed encryption scheme in terms of security and compression showed that it enhances security against ciphertext-only attacks and that it is effective for EtC systems in terms of image quality.

The rest of this paper is organized as follows. Section II provides a review of conventional encryption schemes used in EtC systems. Section III presents the proposed grayscale-based encryption and its security enhancement. Extensive experimental results including robustness against jigsaw puzzle solver attacks are given in Section IV. Finally, Section V concludes this paper.

II. PREPARATION

In this section, after the conventional block scrambling-based image encryption scheme [19]–[22] is summarized, the security of the scheme against brute-force and jigsaw
A. Block Scrambling-Based Image Encryption

A block scrambling-based image encryption scheme was proposed for Etc systems [18]–[22], in which a user wants to securely transmit image \( I \) to an audience, via a Social Networking Service (SNS) provider, as illustrated in Fig. 1. Since the user does not give the secret key \( K \) to the SNS provider, the privacy of image to be shared is under control of the user even when the SNS provider recompresses image \( I \). Therefore, the user can ensure privacy by himself. In comparison, in CtE systems, the user has to disclose unencrypted images to recompress them.

In the scheme [19]–[22], an image with \( X \times Y \) pixels is first divided into non-overlapped blocks with \( B_x \times B_y \); then four block scrambling-based processing steps are applied to the divided image. Figure 2 illustrates the procedure of the scheme with \( B_x = B_y = 16 \). In this paper, \( B_x = B_y = 16 \) is used as well as in [19] and [20]. The procedure for performing image encryption to generate an encrypted image \( I_e \) is given as follows.

- **Step 1:** Divide image with \( X \times Y \) pixels into blocks, each with \( B_x \times B_y \) pixels, and permute randomly the divided blocks using a random integer generated by a secret key \( K_1 \), where \( K_1 \) is commonly used for all color components.

- **Step 2:** Rotate and invert randomly each block (see Fig. 3) by using a random integer generated by a key \( K_2 \), where \( K_2 \) is commonly used for all color components as well.

- **Step 3:** Apply negative-positive transformation to each block by using a random binary integer generated by a key \( K_3 \), where \( K_3 \) is commonly used for all color components. In this step, a transformed pixel value in the \( i \)th block \( B_i \), \( p' \), is computed using

\[
p' = \begin{cases} 
  p & (r(i) = 0) \\
  p \oplus (2^L - 1) & (r(i) = 1)
\end{cases}
\]

where \( r(i) \) is a random binary integer generated by \( K_3 \), and \( p \in B_i \) is the pixel value of the original image with \( L \) bit per pixel. In this paper, the value of occurrence probability \( P(r(i)) = 0.5 \) has been used to invert bits randomly.

- **Step 4:** Shuffle three color components in each block by using an integer randomly selected from six integers by a key \( K_4 \).

An example of an encrypted image \( (B_x = B_y = 16) \) is shown in Fig. 4(b); Fig. 4(a) shows the original one. In this paper, we focus on block scrambling-based image encryption for the following reasons.

(a) The encrypted images are compatible with the JPEG standard.

(b) The compression efficiency for the encrypted images is almost the same as that for the original ones under the JPEG standard.

(c) Robustness against various attacks has been demonstrated [23], [24].

The conventional encryption used in Etc systems has a limitation on block size, i.e. \( B_x = B_y = 16 \), to avoid the effect of color sub-sampling. If \( B_x = B_y = 8 \) is chosen as a block size, the compression performance decreases and some block distortion is generated in decompressed images. In the JPEG standard, when the color sub-sampling is applied to the chroma components \( C_b \) and \( C_r \) of a color image in an encoder, the sub-sampled chroma components are interpolated to reconstruct the same resolution as that of the original image in a decoder. If \( 4:2:0 \) color sub-sampling is applied to encrypted images, each \( 8 \times 8 \)-block in the sub-sampled chroma components consists of four \( 4 \times 4 \)-blocks from different \( 8 \times 8 \)-blocks, which have generally a low correlation among the blocks. Therefore, the compression performance of the encrypted images decreases, and moreover block distortion is generated due to the interpolation of the sub-sampled chroma components with discontinuous values.

B. Security Against Ciphertext-Only Attacks

1) Security: Security mostly refers to protection from adversarial forces. The proposed encryption scheme aims to protect visual information of images that allow us to identify puzzle solver attacks as ciphertext-only attacks (COA) is addressed.
an individual, a time and the location for taking a photograph. Untrusted providers and unauthorized users are assumed as
the adversary. In this paper, we consider security against not

2) Brute-Force Attack: If an image with \( X \times Y \) pixels
is divided into blocks with \( B_x \times B_y \) pixels, the number of
blocks \( n \) is given by

\[
n = \left\lfloor \frac{X}{B_x} \right\rfloor \times \left\lfloor \frac{Y}{B_y} \right\rfloor,
\]

where \( \lfloor \cdot \rfloor \) is a function that rounds down to the nearest integer.
The four block scrambling-based processing steps are then
applied to the divided image.

Fig. 4. Encrypted images with the conventional and proposed scheme.
(a) Original image \((X \times Y = 384 \times 512)\). (b) Encrypted image [19], [20]
\((B_x = B_y = 16, n = 768)\). (c) Encrypted grayscale-based image using proposed
scheme (Rectangular, \(B_x = B_y = 8, n = 9216)\). (d) Encrypted grayscale-based
image using proposed scheme (Square, \(B_x = B_y = 8, n = 9216)\).

The key space of the block scrambling (Step 1) \( N_S(n) \),
which is the number of permutation of \( n \) blocks, is given by

\[
N_S(n) = n!.
\]

Similarly, the key spaces of other encryption steps are given as

\[
\begin{align*}
N_R(n) &= 4^n, \quad N_I(n) = 4^n, \quad N_{R\&I}(n) = 8^n \\
N_N(n) &= 2^n, \quad N_C(n) = (3^P_3)^n = 6^n
\end{align*}
\]

where \( N_R(n) \) and \( N_I(n) \) are the key spaces of the block
rotation and block inversion, and \( N_{R\&I}(n) \) is the key space of
the encryption combining them (Step 2). Note that \( N_{R\&I} \) is
the key space considering the collision between block rotation
and inversion. Namely, rotating pieces 180 degrees is the
same operation as inverting them horizontally and vertically.
\( N_N(n) \) and \( N_C(n) \) are the key spaces of the negative-positive
transformation (Step3) and color component shuffling (Step 4)
respectively. Consequently, the key space of images encrypted
by using all the proposed encryption steps, \( N_A(n) \), is represen
ted by

\[
N_A(n) = N_S(n) \cdot N_{R\&I}(n) \cdot N_N(n) \cdot N_C(n)
= n! \cdot 8^n \cdot 2^n \cdot 6^n.
\]

Fig. 5. Assembled images by using the extended jigsaw puzzle solver [26].
(a) Original image \((X \times Y = 224 \times 160)\). (b) Encrypted image (Conventional
scheme, \(B_x = B_y = 16, n = 140)\). (c) Encrypted image (Conventional
scheme, \(B_x = B_y = 8, n = 560)\). (d) Assembled image \((B_x = B_y = 16)\). (e) Assembled
image \((B_x = B_y = 8)\).

3) Jigsaw Puzzle Solver Attack: To assemble encrypted images including
inverted, negative-positive transformed and color component shuffled blocks, extended jigsaw puzzle
solvers for block scrambling-based image encryption have been proposed [23], [24]. It has been shown that assembling
jigsaw puzzles becomes difficult when the encrypted images
are satisfied with under three conditions [23], [24], [26], [27].
(a) Number of blocks is large.
(b) Block size is small.
(c) Encrypted images include JPEG distortion.
Figures 5(d) and (e) are examples of images assembled from
Figs. 5(d) and (e) respectively; Fig. 5(a) shows the original one.
TABLE I
PROPERTIES OF BLOCK SCRAMBLING-BASED IMAGE ENCRYPTION SCHEMES

| Scheme  | Conventional [19], [20] | Proposed |
|---------|-------------------------|----------|
| Color channel | RGB | Grayscale |
| Minimum block size | 16 x 16 | 8 x 8 |
| Image size | X x Y | 3 x X x Y |
| Number of blocks | n | 3n |

- Robustness against jigsaw puzzle solver attacks: Robust, More robust
- Effect of color sub-sampling: Affected, Unaffected

* Calculated from Eq. 2

Compared with Figs. 5(d) and (e), the difficulty of assembling encrypted images strongly depends on the block size. In addition, most conventional jigsaw puzzle solvers also utilize color information to assemble puzzles. Thus, reducing the number of color channels in each pixel makes assembling encrypted images much more difficult. The novel scheme in this paper has a higher security level than that of the conventional scheme, because it provides a large number of blocks and the small block size.

Other attacking strategies such as known-plaintext attack (KPA) and chosen-plaintext attack (CPA) should be considered for the security. The block scrambling-based image encryption becomes robust against KPA through the assigning of a different key to each image for the encryption. In addition, the keys used for the encryption do not need to be disclosed because the encryption scheme is not public key cryptography. Therefore, the encryption can avoid the CPA, unlike public key cryptography.

C. Summary of Image Encryption for EtC Systems
The properties of the conventional encryption scheme [19], [20] that is only one conventional scheme for EtC systems with the JPEG standard is summarized in Table I. The proposed one enables the use of a smaller block size and a larger number of blocks, which enhances both invisibility and security against several attacks. Furthermore, images encrypted by using the proposed scheme include less color information due to the use of grayscale images, which makes the EtC system more robust. The properties of the proposed one allow us not only to enhance security against several attacks, but also to avoid the effect of color sub-sampling.

III. PROPOSED METHOD
In this section, the grayscale-based image encryption scheme, which has higher security than the conventional one, is proposed.

A. Procedure of Proposed Image Encryption
Although \( B_x = B_y = 16 \) is used as the smallest block size in the conventional block scrambling-based image encryption to avoid the effect of color sub-sampling in JPEG compression, the proposed method enables us to use \( B_x = B_y = 8 \) as a block size, which enhances robustness against ciphertext-only attacks. Moreover, applying EtC systems with the proposed scheme to social media performs better than with the conventional one as described below.

The procedure of the proposed scheme for an 8-bit RGB full-color image with \( X \times Y \) pixels is given as follows (See Fig. 6):

1) Split color image into three (RGB) channels.
2) Considering JPEG compression efficiency, RGB components are transformed into YCbCr color space by using the equation below, as in [28].

\[
Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \tag{7}
\]

\[
C_b = -0.1687 \times R - 0.3313 \times G + 0.5 \times B + 128 \tag{8}
\]

\[
C_r = 0.5 \times R - 0.4187 \times G - 0.0813 \times B + 128 \tag{9}
\]

3) Select \( B_x = B_y = 8 \) as a block size.
4) Combine YCbCr channels into one grayscale-based image. A grayscale image with \( 3 \times X \times Y \) pixels is thereby generated.
5) Apply block scrambling, rotation, inversion, and negative-positive transformation by using Steps 1 to 3 in Sec. II-A.

An example of an encrypted image \( (B_x = B_y = 8) \) is shown in Fig. 4(c), where YCbCr channels are combined horizontally. As shown in Fig. 4(d), which is an example of an encrypted image combined YCbCr channels to become square, the way of combining YCbCr channels has some freedom.

B. Compression of Grayscale-Based Image
In this paper, we focus on JPEG lossy compression, although the JPEG standard supports both lossless and lossy compressions, and the encryption schemes are applicable to lossless compression methods as discussed in [22]. This is because most JPEG compression applications, especially SNS providers and Cloud Photo Storage Services (CPSS), use lossy compression, and lossless compression does not generate any distortion.

Whereas RGB images are generally compressed by using two quantization tables, namely for luminance \( (Y) \) and for chrominance \( (C_b, C_r) \) respectively, grayscale images are compressed with only one quantization table in the JPEG standard. Images encrypted with the proposed scheme are grayscale-based, so one quantization table is used. The quantization table that is selected for the proposed scheme affects the compression performance. In the experiments, the relationship between quantization tables and compression performances is discussed.
color sub-sampling is not carried out. Therefore, $B_x = B_y = 8$ is selected as the smallest block size for grayscale images in the JPEG standard.

The encrypted images generated by the proposed scheme are grayscale-based ones. Therefore, $B_x = B_y = 8$ can be selected for the encryption, as shown in Fig. 4(c) and (d), even if the encrypted images are compressed using the JPEG standard.

2) Large Number of Blocks: When $B_x = B_y = 8$ is selected, the number of blocks increases fourfold, $B_x = B_y = 16$, from Eq. (2). Moreover, the proposed scheme generates encrypted images with $3 \times X \times Y$ pixels from an original image with $X \times Y$ pixels. As a result, the number of blocks is 12 times that with the conventional one, as shown in Figs. 4(b) and (c).

The running time to assemble encrypted images strongly depends on the number of pieces when jigsaw puzzle solver attacks are utilized. Furthermore, increasing the number of blocks makes assembling encrypted images much more difficult. Therefore, the proposed scheme enhances security.

3) Less Color Information in Blocks: The encrypted images include only one channel per pixel due to the use of grayscale-based images. This makes assembling jigsaw puzzles more difficult because almost all solvers utilize three color channels in each block to assemble puzzles correctly [29–32]. Therefore, robustness against jigsaw puzzle solver attacks is enhanced.

4) Key Space Expansion: Using the proposed image encryption makes the key space of encrypted images larger than with the conventional scheme due to the increased number of blocks. The key space of the proposed scheme $N_B(n)$ is given by

$$N_B(n) = N_S(3n) \cdot N_{RGB}(3n) \cdot N_N(3n)$$

$$= 3n! \cdot 8^{3n} \cdot 2^{2n} \gg N_A(n), \quad (10)$$

where $n$ is the number of blocks, calculated from an original image with $X \times Y$ pixels in accordance with Eq. (2). Although the proposed scheme does not apply color component shuffling, the number of blocks is larger, as shown in Eqs. (2) and (10). Thus, the proposed scheme enhances robustness against brute-force attacks.

E. Application to Social Media

Figure 1 illustrates the application of the EtC system to social media such as Facebook and Twitter, where the user wants to securely transmit image $I$ to an audience via social media. Since the user does not give the secret key $K$ to social media providers, the privacy of the image to be shared is under the user’s control, even if the social media provider decompresses the image. This means that, if an encrypted image saved on the provider’s server is leaked, third parties and general audiences cannot see the images visually unless they have the key.

Table II summarizes the relationship between uploaded and downloaded JPEG images in terms of sub-sampling ratios and JPEG quality factor $Q_f$. The previous work [25] discussed the relationship for color images, but Table II includes the
TABLE II
RELATIONSHIP BETWEEN UPLOADED JPEG FILES AND DOWNLOADED ONES IN TERMS OF SUB-SAMPLING RATIOS AND THE MAXIMUM RESOLUTIONS.

| SNS provider     | Uploaded JPEG file | Downloaded JPEG file |               |
|------------------|--------------------|----------------------|---------------|
|                  | Sub-sampling ratio | Qfu                  | Sub-sampling ratio | Qfd |
| Twitter (Up to 4096×4096 pixels) | 4:4:4 low          | 1.2...84             | high          | 4:2:0                                     |
|                  |                    |                      |               | No recompression                          |
|                  | 4:2:0              |                      | 85,86,...100   | 4:2:0                                     |
|                  |                    |                      |               | 85                                         |
| Facebook (HQ, Up to 2048×2048 pixels) | 4:4:4               | 1.2...100            | (Grayscale) | 4:2:0                                     |
| Facebook (LQ, Up to 960×960 pixels)  | 4:2:0              |                      | (Grayscale)   | 71.72...85                                |
| Tumblr (Up to 1280×1280 pixels)     | 4:4:4              |                      |               | No recompression                          |
| Google+ Flicker  | 4:2:0              |                      |               |                                           |

Fig. 8. Decrypted images downloaded from Facebook. The image in (a) includes some block distortion due to the effect of color sub-sampling. (a) Decrypted image with block artifact (Conventional scheme, PSNR = 31.4dB, sub-sampling ratio=4:2:0). (b) Decrypted image (Proposed scheme, PSNR = 36.3dB).

relationship for grayscale images. Qfu and Qfd indicate the quality factor of uploaded and downloaded images, respectively. Although it has been confirmed that the EtC system with the conventional scheme is applicable to social media, EtC systems using the proposed scheme have superior features to the conventional one as shown below.

1) Color Sub-Sampling: As indicated in Table II, SNS providers manipulate uploaded images by changing the sub-sampling ratio and Qfu. Thus, we have to consider the effect of sub-sampling ratios, when encrypted images are upload to SNS providers that recompress uploaded images. JPEG images with 4:2:0 sub-sampling ratio are interpolated to increase the spatial resolution for chroma components in the decoding process. This interpolation processing is carried out by using the relationship among blocks. Therefore, encrypted images with 4:2:0 sub-sampling ratio are affected by this interpolation. As shown in Fig. 8(a), the decrypted image includes block distortion due to the interpolation on Facebook. In comparison, the interpolation in the decoding process does not affect encrypted images with the proposed scheme, so encrypted images can avoid the distortion as shown in Fig. 8(b). Thus, EtC systems with the proposed scheme can avoid the effects of the interpolation, which is carried out on social media.

2) Recompression: Furthermore, images encrypted with the proposed scheme perform better in terms of image quality. The quality of images downloaded from social media is generally degraded due to recompression forced by SNS providers, when the operation of color sub-sampling is carried out. Images encrypted by the proposed scheme always avoid this operation, which enable users to transmit images of higher quality than using the conventional scheme. As well as SNS providers, EtC systems with the proposed scheme are applicable to Cloud Photo Services like iCloud and Google Photos.

3) Resizing: To apply EtC systems to social media, the resolution of encrypted images needs to be smaller than the maximum resolutions that each provider decides on as a resizing condition [25]. For example, Twitter does not resize uploaded images with the resolution of up to 4096×4096 pixels, as shown in Table II. Resizing the resolution of encrypted images makes the block size of the encrypted images smaller, although the JPEG compression is still carried out based on the size of 8×8. As a result, each 8×8-block in resized images includes pixels from originally different blocks, so the compression performance decreases and block distortion is generated in the decrypted images due to the discontinuity among pixels. Thus, we have to upload encrypted images within the maximum resolution as indicated in Table II. In this paper, the use of the image resolution that is not forcedly resized by SNS providers is assumed as well as the conventional EtC systems [25].

IV. Evaluation

In this section, we evaluated the effectiveness of the proposed scheme in a number of simulations. First, we evaluated the compression performance of images encrypted by the proposed scheme. Next, encrypted images were uploaded to SNS providers and then downloaded from the providers to confirm robustness against image manipulation on social media. Finally, the security enhancement of the EtC systems against jigsaw puzzle solver attacks [26] is shown from the aspect of the difficulty of assembling encrypted images.
To determine the compression performance of the proposed scheme, a lot of images encrypted by the conventional [20] and proposed scheme were compressed and decompressed. Then, we calculated peak signal-to-noise ratio (PSNR) between original images and images encrypted with the proposed method. Uncompressed Color Image Database (UCID) dataset, which contains 1338 color images with the sizes of 512 × 384 or 384 × 512, was used for the evaluation. The average PSNR values of all images per $Q_f$ in the dataset were used for the evaluation. Software by the Independent JPEG Group’s (IJG) [33] was used for encoding and decoding the images with following parameters.

- Sub-sampling ratio : 4:4:4, 4:2:0
- Quantization table : IJG standard table

Non-encrypted images were also compressed with the same parameters. As we mentioned before, grayscale-based images are generally compressed with only one quantization table, so the luminance and chrominance tables are used respectively for encoding images encrypted with the proposed scheme. In this paper, the quantization tables for luminance and chrominance designed by IJG were used.

Figure 9 shows rate-distortion (RD) curves of JPEG compressed images without any encryption and with encryption. As the bitrate increased, the encrypted images with the proposed scheme had better compression performance than non-encrypted images compressed with 4:2:0 sub-sampling ratio as shown in Fig. 9(a). From the results, the operation of color sub-sampling affected the compression performance as the bitrate increases. In terms of the relationship between the proposed scheme with chrominance table and with luminance table, the compression performance was very similar. Fig. 9(b) shows the result with 4:4:4 sub-sampling ratio. The compression performance of the proposed scheme was almost the same as the non-encrypted images compressed with 4:4:4 sub-sampling ratio. Therefore, it was determined that images encrypted by the proposed scheme have almost the same compression performance as both non-encrypted and conventional ones.

Next, other conventional encryption methods are considered in terms of compression performance. There are various encryption methods which can maintain an image format after encrypting as well as the proposed scheme. However, they are not suitable to EtC systems with JPEG compression, because they do not consider using JPEG compression. We numerically compared the encryption methods [34], [35] with the proposed scheme. Figures 10(c) and (d) indicate decrypted images after compressing and decompressing encrypted ones, where Fig. 10(a) is the original one. The image quality of decrypted images heavily decreased due to JPEG compression as shown in Figs. 10(c) and (d), because they do not consider using JPEG compression as well as most other conventional encryption methods. In contrast, the image encrypted with the proposed scheme had almost same quality as that of non-encrypted images. Note that Fig. 10(d) is the grayscale-based image [35], thus the PSNR value is not listed.

### B. Image Manipulation on Social Media

We uploaded images encrypted by using the proposed scheme to Facebook and Twitter, and then downloaded them, as well as images encrypted using the conventional scheme [19], [20] and non-encrypted images to confirm the effectiveness of the proposed scheme.
To compare PSNR values in step 7, original image $I$ was uploaded. For this reason, we uploaded images compressed with Q fu such as $Q_{fu} = 8$, $Q_{fu} = 16$, $Q_{fu} = 32$, $Q_{fu} = 64$, and $Q_{fu} = 128$. While Facebook recompresses all uploaded images by users, Twitter recompresses uploaded images under some conditions such as $Q_{fu} \geq 85$ for 4:2:0 sub-sampling ratio. Therefore, the PSNR values of the decrypted images were calculated.

1) Experimental Conditions: In our experiments, the same dataset and parameters for JPEG compression as in Sec.IV-A were utilized for the evaluation. The following procedure was carried out according to Fig.1.

1) Generate encrypted image $I_e$ from original image $I$.
2) Compress encrypted image $I_e$ with $Q_{fu}$.
3) Upload encrypted JPEG image $I_{ec}$ to SNS providers.
4) Download recompressed JPEG image $\hat{I}_{ec}$ from the providers.
5) Decompress encrypted JPEG image $\hat{I}_{ec}$.
6) Decrypt manipulated image $\hat{I}$.
7) Compute the PSNR value between original image $I$ and decrypted image $\hat{I}$.

To compare PSNR values in step 7, original image $I$ was compressed without any encryption and then uploaded and downloaded. The downloaded images were decompressed, and then, the average PSNR values for 1338 images per $Q_{fu}$ were calculated.

We focused on Facebook and Twitter, which always recompress uploaded images when the images meet the conditions for $Q_{fu}$ as shown in Table II.

2) Experimental Results: Table III shows the experimental result of Facebook. As described in Sec.III-E.1, the images decrypted with the conventional scheme and 4:2:0 sub-sampling ratio included block distortion like in Fig.8(a). Therefore, the PSNR values of the decrypted images were much lower than the others. In comparison, images decrypted by using the proposed scheme always avoided the interpolation on Facebook, so the PSNR values of decrypted images were higher than the others.

Table IV shows the result for Twitter. As indicated in the Table, a similar tendency to Facebook was shown. While Facebook recompresses all uploaded images by users, Twitter recompresses uploaded images under some conditions such as $Q_{fu} \geq 85$ for 4:2:0 sub-sampling ratio. For this reason, we uploaded images compressed with $Q_{fu} = 90, 95, 100$. Twitter recompresses uploaded images with 4:2:0 sub-sampling in the DCT domain, thereby generating no distortion [25]. However, interpolation is carried out by users when users download and decompress the images. Therefore, the PSNR values of the decrypted images with 4:2:0 sub-sampling ratio were lower than with the proposed scheme.

As shown in Tables III and IV, most PSNR values of the images decrypted with the proposed scheme using the luminance table were higher than with the chrominance one. It has been known that most SNS providers use the luminance table for the uploaded grayscale images in recompression. That is to say, uploaded images compressed with the luminance table by users are recompressed with the luminance table again by SNS providers. As a result, using the luminance table would decrease the quantization error in recompression.

C. Robustness Against Jigsaw Puzzle Solver Attacks

Next, the security of grayscale-based block scrambling image encryption against the jigsaw puzzle solver was evaluated.

1) Experimental Conditions: We assembled encrypted images by using the extended jigsaw puzzle solver [23], [24]. The following three measures [29], [36] were used to evaluate the results.

a) Direct comparison ($D_c$): Represents the ratio of the number of pieces which are in the correct position. $Dc$ for image $I_d$, namely, $Dc(I_d)$ is calculated as

$$D_c(I_d) = \frac{1}{n} \sum_{i=1}^{n} d_c(i),$$

where $d_c(i)$ represents the position of a piece $i$ in image $I_d$

b) Neighbor Comparison ($N_c$): Is the ratio of the number of correctly joined blocks. $Nc$ for image $I_d$, namely, $Nc(I_d)$ is
correctly assembled area, and \( \text{lc} \) calculated as Fig. 12. Images assembled from encrypted ones with the proposed scheme. (a) Assembled image (Proposed, \( B_x = B_y = 16, D_x = 0, N_c = 0.042, L_c = 0.036 \)).

(c) Assembled image (Grayscale, \( B_x = B_y = 16, D_x = 0, N_c = 0.042, L_c = 0.036 \)).

![Image 104x474 to 236x574]

![Image 104x634 to 236x734]

![Figure 13](image)

(c) Assembled image (Proposed, \( B_x = B_y = 32, D_x = 0.001, N_c = 0.000, L_c = 0.021 \)). (b) Assembled image (Proposed, \( B_x = B_y = 16, D_x = 0.001, N_c = 0.001, L_c = 0.006 \)). (c) Assembled image (Proposed, \( B_x = B_y = 8, D_x = 0.001, N_c = 0.001, L_c = 0.002 \)).

![Image 240x474 to 372x574]

![Image 240x634 to 372x734]

![Figure 12](image)

Fig. 11. Images assembled from encrypted ones with the conventional scheme. The image (c) is assembled from grayscale-based encrypted image with one channel. (a) Original image \( (X \times Y = 256 \times 192) \). (b) Assembled image (Conventional scheme, \( B_x = B_y = 16, D_x = 0, N_c = 0.111, L_c = 0.078 \)). (c) Assembled image (Grayscale, \( B_x = B_y = 16, D_x = 0, N_c = 0.042, L_c = 0.036 \)).

![Figure 13](image)

![Figure 12](image)

\[ Nc(I_d) = \frac{1}{B} \sum_{k=1}^{B} n_c(k), \]

\[ n_c(k) = \begin{cases} 1, & \text{if } b_k \text{ is joined correctly} \\ 0, & \text{otherwise} \end{cases} \quad (12) \]

where \( B \) is the number of pieces in \( I_d \), and \( b_k \) is the \( k \)th boundary among pieces in \( I_d \). For an image with \( u \times v \) blocks, there are \( B = 2uv - u - v \) boundaries in the image.

c) Largest component (\( L_c \)): Is the ratio of the number of the largest joined blocks that have correct adjacencies to the number of blocks in an image. \( Lc \) for image \( I_d \), namely, \( Lc(I_d) \) is calculated as

\[ Lc(I_d) = \frac{1}{n} \max \{ l_c(I_d, j), j = 1, 2, \ldots , m \} \quad (13) \]

where \( l_c(I_d, j) \) is the number of blocks in the \( j \)th partial correctly assembled area, and \( m \) is the number of partial correctly assembled areas.

In the measures, \( D_c, N_c, L_c \in [0, 1] \), a larger value means a higher compatibility.

We used 20 images randomly chosen from the UCID dataset, and each image was resized to \( 256 \times 192 \). Forty different encrypted images were generated from one ordinary image by using different keys. We assembled the encrypted images by using the jigsaw puzzle solver and chose the image that had the highest sum of \( D_c, N_c, \) and \( L_c \). We performed this procedure for each encrypted image independently and calculated the average \( D_c, N_c, \) and \( L_c \) for the 20 images.

2) Experimental Results: Table V shows robustness against an extended jigsaw puzzle solver attack [23], [24]. To evaluate the effect of reducing the number of color channels on puzzle assembly, images with only one channel, namely grayscale-based images, were encrypted in accordance with Step 5 in Sec. III-A and then assembled. Examples of images assembled from encrypted ones with the proposed and conventional scheme are shown in Figs. 11(b) and (c); Fig. 11(a) shows the original one. As shown in Fig. 11(b), it is possible to partially assemble encrypted images with the conventional scheme if the number of blocks is small. Since the scores for the encrypted images with one color channel were lower than those of the ones assembled with the conventional scheme, reducing the number of color channels makes assembling puzzles more difficult.

Figures 12(a), (b) and (c) show the examples of images assembled from encrypted ones with the proposed scheme, where Fig. 11(b) is the original image. As shown in Table V and Figs. 12(a), (b) and (c), the scores for the proposed scheme were much lower than the other scores. Even when large block sizes (\( B_x = B_y = 16 \)) were used for the proposed encryption, the scores of assembled images were far lower than with the conventional scheme with \( L_c = 0.021 \). This is because images encrypted with the proposed scheme have a large number of encrypted blocks and less color information in the blocks. The proposed scheme thus has higher security against jigsaw puzzle solver attacks than the conventional scheme.

3) Running Time to Assemble Jigsaw Puzzles: Figure 13 shows the running time to assemble encrypted images by using the jigsaw puzzle solver [23], [24], where the average time of 15 images from resized UCID dataset were plotted.
much more difficult. In comparison, decrypted images with the proposed scheme include a larger number can be assumed to occur as cipher-text only attacks, images ciphertext-only attacks. Although jigsaw puzzle solver attacks of assembled results.

The running time to assemble images encrypted in terms of both computational complexity and the accuracy scheme can enhance security against ciphertext-only attacks by the conventional scheme. The reason is that the proposed scheme can offer a smaller block size, the larger number of assembled results.

As shown in Fig. 13, although the images encrypted by using the proposed method \((B_x = B_y = 8)\) were solved in 166.11 minutes, the scores of assembled images were very low as \(L_c = 0.002\) (See Table V). It obviously takes more time to assemble encrypted images than that of images encrypted by the conventional scheme. The reason is that the proposed scheme can offer a smaller block size, the larger number of blocks and less color information. As a result, the proposed scheme can enhance security against cipher-text-only attacks in terms of both computational complexity and the accuracy of assembled results.

V. CONCLUSION

We proposed a novel block-scrambling image encryption scheme that enhances the security of EtC systems for JPEG images. Although \(B_x = B_y = 16\) is used as the smallest block size in the conventional scheme to avoid the effect of color sub-sampling, the proposed scheme enables us to use \(B_x = B_y = 8\) as a block size, which enhances robustness against ciphertext-only attacks. Although jigsaw puzzle solver attacks can be assumed to occur as cipher-text only attacks, images encrypted with the proposed scheme include a larger number of small blocks, which makes assembling encrypted images much more difficult. In comparison, decrypted images with the conventional scheme sometimes include some block distortion due to the interpolation on social media. The proposed scheme makes it possible to avoid the effect of the interpolation on social media due to the use of grayscale-based images. As a result, the proposed scheme has a better performance than the conventional one in terms of the image quality. Experimental results showed the EtC systems with the proposed scheme are applicable to Twitter and Facebook. The proposed scheme is also applicable to other SNS providers and cloud photo services like Tumblr, iCloud and Google Photos. In addition, the robustness of the proposed scheme against jigsaw puzzle solver attacks was confirmed in the experiment.

Fig. 13. Running time of assembling encrypted images from resized UCID dataset (256 × 192).

We compared the running time to assemble images encrypted with the conventional scheme \((B_x = B_y = 16)\) and the proposed one \((B_x = B_y = 16\) and \(B_x = B_y = 8\)). The jigsaw puzzle solver was implemented in MATLAB2015a on a PC with a 3.3GHz processor and a main memory 16Gbytes (Processor:Intel Core i7-5820K 3.3GHz, OS:Ubuntu 16.04 LTS).

As shown in Fig. 13, although the images encrypted by using the proposed method \((B_x = B_y = 8)\) were solved in 166.11 minutes, the scores of assembled images were very low as \(L_c = 0.002\) (See Table V). It obviously takes more time to assemble encrypted images than that of images encrypted by the conventional scheme. The reason is that the proposed scheme can offer a smaller block size, the larger number of blocks and less color information. As a result, the proposed scheme can enhance security against cipher-text-only attacks in terms of both computational complexity and the accuracy of assembled results.

V. CONCLUSION

We proposed a novel block-scrambling image encryption scheme that enhances the security of EtC systems for JPEG images. Although \(B_x = B_y = 16\) is used as the smallest block size in the conventional scheme to avoid the effect of color sub-sampling, the proposed scheme enables us to use \(B_x = B_y = 8\) as a block size, which enhances robustness against ciphertext-only attacks. Although jigsaw puzzle solver attacks can be assumed to occur as cipher-text only attacks, images encrypted with the proposed scheme include a larger number of small blocks, which makes assembling encrypted images much more difficult. In comparison, decrypted images with the conventional scheme sometimes include some block distortion due to the interpolation on social media. The proposed scheme makes it possible to avoid the effect of the interpolation on social media due to the use of grayscale-based images. As a result, the proposed scheme has a better performance than the conventional one in terms of the image quality. Experimental results showed the EtC systems with the proposed scheme are applicable to Twitter and Facebook. The proposed scheme is also applicable to other SNS providers and cloud photo services like Tumblr, iCloud and Google Photos. In addition, the robustness of the proposed scheme against jigsaw puzzle solver attacks was confirmed in the experiment.

REFERENCES

[1] C. T. Huang et al., “Survey on securing data storage in the cloud,” IPSIPA Trans. Signal Inf. Process., vol. 3, p. e7, Jan. 2014.
[2] R. L. Lagendijk, Z. Erkin, and M. Barni, “Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation,” IEEE Signal Process. Mag., vol. 30, no. 1, pp. 82–105, Jan. 2013.
[3] 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.
[4] M.-R. Ra, R. Govindan, and A. Ortega, “P3: Toward privacy-preserving photo sharing,” in Proc. 10th USENIX Conf. Netw. Syst. Design Implement., vol. 13, Apr. 2013, pp. 515–528.
[5] 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.
[6] I. Ito and H. Kiya, “A new class of image registration for guaranteeing secure data management,” in Proc. 15th IEEE Int. Conf. Image Process. (ICIP), 2008, pp. 269–272.
[7] K. Ito and H. Kiya, “Image matching between scrambled images for secure data management,” in Proc. 16th Eur. Signal Process. Conf. (EUSIPCO), Aug. 2008, pp. 1–5.
[8] 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, Dec. 2009, Art. no. 841045.
[9] Z. Tang, X. Zhang, and W. Lan, “Efficient image encryption with block shuffling and chaotic map,” Multimedia Tools Appl., vol. 74, no. 15, pp. 5429–5448, Aug. 2015.
[10] C. Li, D. Lin and J. Li, “Cryptanalyzing an image-scrambling encryption algorithm of pixel bits,” IEEE Trans. Multimedia, vol. 24, no. 3, pp. 64–71, Aug. 2017.
[11] Z. Erkin et al., “Protection and retrieval of encrypted multimedia content: When cryptography meets signal processing,” EURASIP J. Inf. Security, vol. 2007, no. 78943, Jan. 2007, Art. no. 17.
[12] K. G. Nimbokar and M. V. Sarode, “A survey based on designing an efficient image encryption-then-compression system,” in Proc. IJCA Nat. Level Tech. Conf. (X-PLORE), 2014, pp. 6–8.
[13] 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, Oct. 2004.
[14] 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, Apr. 2010.
[15] X. Zhang, “Lossy compression and iterative reconstruction for encrypted image,” IEEE Trans. Inf. Forensics Security, vol. 6, no. 1, pp. 53–58, Mar. 2011.
[16] R. Hu, X. Li, and B. Yang, “A new lossy compression scheme for encrypted gray-scale images,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2014, pp. 7387–7390.
[17] T. Y. Liu, K. J. Lin, and H. C. Wu, “ECG data encryption then compression using singular value decomposition,” IEEE J. Biomed. Health Inform., vol. 22, no. 3, pp. 707–713, May 2018.
[18] O. Watanabe, A. Uchida, T. Fukuhara, and H. Kiya, “An encryption-then-compression system for JPEG 2000 standard,” in Proc. IEEE Inf. Conf. Acoust., Speech Signal Process. (ICASSP), Apr. 2015, pp. 1226–1230.
[19] K. Kurihara, S. Shiota, and H. Kiya, “An encryption-then-compression system for JPEG standard,” in Proc. Picture Coding Symp. (PCS), Jun. 2015, pp. 119–123.
[20] K. Kurihara, M. Kikuchi, S. Imaizumi, S. Shiota, and H. Kiya, “An encryption-then-compression system for JPEG/Motion JPEG standard,” IEICE Trans. Fundam. Electron., Commun. Comput. Sci., vol. 98, no. 11, pp. 2238–2245, Nov. 2015.

[21] K. Kurihara, O. Watanabe, and H. Kiya, “An encryption-then-compression system for JPEG XR standard,” in Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB), Jun. 2016, pp. 1–5.

[22] K. Kurihara, S. Imaizumi, S. Shiota, and H. Kiya, “An encryption-then-compression system for lossless image compression standards,” IEICE Trans. Inf. Syst., vol. E100-D, no. 1, pp. 52–56, Jan. 2017.

[23] T. Chuman, K. Kurihara, and H. Kiya, “On the security of block scrambling-based ETC systems against jigsaw puzzle solver attacks,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2017, pp. 2157–2161.

[24] T. Chuman, K. Kurihara, and H. Kiya, “Image manipulation on social media document Recommendation T.871, International Telecommunication Union, 2012.

[25] T. Chuman, K. Iida, and H. Kiya, “Image manipulation on social media for encryption-then-compression systems,” in Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC), Dec. 2017, pp. 858–863.

[26] T. Chuman, K. Kurihara, and H. Kiya, “On the security of block scrambling-based ETC systems against extended jigsaw puzzle solver attacks,” IEICE Trans. Inf. Syst., vol. E101-D, no. 1, pp. 37–44, Jan. 2018.

[27] T. Chuman and H. Kiya, “On the security of block scrambling-based image encryption including JPEG distortion against jigsaw puzzle solver attacks,” in Proc. IEEE Int. Workshop Signal Design Appl. Commun. (IWSDA), Sep. 2017, pp. 64–68.

[28] Information Technology—Digital Compression and Coding of Continuous-Tone Still Images: JPEG File Interchange Format (JFIF), document Recommendation T.871, International Telecommunication Union, 2012.

[29] A. Gallagher, “Jigsaw puzzles with pieces of unknown orientation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 382–389.

[30] K. Son, D. Moreno, J. Hays, and D. B. Cooper, “Solving small-piece jigsaw puzzles by growing consensus,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 1193–1201.

[31] G. Paikin and A. Tal, “Solving multiple square jigsaw puzzles with missing pieces,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 4832–4839.

[32] F. A. Andaló, G. Taubin, and S. Goldenstein, “PSQP: Puzzle solving by quadratic programming,” IEEE Trans. Pattern Anal. Machine Intell., vol. 39, no. 2, pp. 385–396, Feb. 2017.

[33] Independent JPEG Group. Accessed: Feb. 2, 2018. [Online] Available: http://www.iijg.org/

[34] M. T. Gaata and F. F. Hantoosh, “An efficient image encryption technique using chaotic logistic map and RC 4 stream cipher,” Int. J. Modern Trends Eng. Res., vol. 3, no. 9, pp. 213–218, 2016.

[35] Y. Wu, J. P. Noonan, G. Yang, and H. Jin, “Image encryption using the two-dimensional logistic chaotic map,” J. Electron. Imag., vol. 21, no. 1, pp. 013014, Mar. 2012.

[36] T. S. Cho, S. Avidan, and W. T. Freeman, “A probabilistic image jigsaw puzzle solver,” in Proc. IEEE Conf. Comput.Vis. Pattern Recognit. (CVPR), Jun. 2010, pp. 183–190.

Tatsuya Chuman received the B.Eng. and M.Eng. degrees from Toyo University, Japan, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with Tokyo Metropolitan University. His research interests are in the areas of image processing.

Warit Sirichotedumrong received the B.Eng. and M.Eng. degrees from King Mongkut’s University of Technology Thonburi, Thailand, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with Tokyo Metropolitan University. His research interests are in the area of image processing.

Hitoshi Kiya (F’16) received the B.E. and M.E. degrees from the Nagaoka University of Technology, Japan, in 1980 and 1982, respectively, and the Dr. Eng. degree from Tokyo Metropolitan University in 1987. In 1982, he joined Tokyo Metropolitan University, Japan, where he became a Full Professor in 2000. From 1995 to 1996, he attended the University of Sydney, Australia, as a Visiting Fellow. He is currently a Professor with Tokyo Metropolitan University. His research interests are in the areas of signal processing theory, image/video processing, and information security and forensics. In these areas, he has published over 150 refereed journal papers, over 300 international conference papers, and over 10 books. He is a fellow of the IEICE and ITE. He was the Chair of two technical committees and was a member of seven technical committees, including the IEEE Information Forensics and Security TC. He is/was Regional Director-at-Large for Region 10 of the IEEE Signal Processing Society, the Japan Chapter Chair of the IEEE Signal Processing Society, President of the IEICE Engineering Sciences Society, and President-Elect of APSIPA. He was also an Associate Editor of eight journals, including IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.