AC Prediction Error Propagation-based Encryption for Texture Protection of JPEG Compressed Images

Kosuke Shimizu, Qifan Wang, and Taizo Suzuki
University of Tsukuba, Ibaraki, Japan
Email: {kshimizu, wangqf}@wmp.cs.tsukuba.ac.jp, taizo@cs.tsukuba.ac.jp

***Abstract*—A recent attention of JPEG format-compliant encryptions has been concentrated to randomize the specific syntaxes of JPEG file interchange format (JFIF) or to flip the signs of quantized discrete cosine transform (QDCT) coefficients. Although the random sign flip (RSF) achieves low-complexity encryption while preserving the JPEG coding efficiency, the texture information represented by nonzero AC coefficients are still easy to reveal for an attacker using replacement attacks. To reinforce the perceptual degradation and the texture protection at the same complexity as the RSF, we propose an AC prediction error propagation-based encryption (AC-WPE) in the JPEG QDCT domain. Since the AC-WPE encrypts the coefficient signs in the residual QDCT domain obtained by a provisional AC predictor such as differential pulse-code modulation (DPCM) and propagates the encrypted prediction errors among the several blocks by the AC reconstructor, the attack resilience is obviously reinforced. Moreover, to suppress the decrease in coding efficiency, we introduce the randomized prediction intervals into the AC predictor/reconstructor. Experiments with JPEG encryption show that the AC-WPE reinforces the perceptual degradation and the resilience to the replacement attacks, while reasonably expensing the bitrate overheads of crypto-compressed images.**

***Index Terms*—AC predictor, JPEG, prediction error propagation, QDCT domain, sign flip, texture protection.**

I. Introduction

Many JPEG format-compliant encryptions (FEs) [1]–[7] have been proposed to display image textures and/or colors that are in imperceptibly distorted states. Since the JPEG is the de-facto image coding standard implemented in many multimedia apparatuses, e.g., cameras and personal computers, the JPEG FEs are required for application fields of social networking services (SNSs), e.g., Twitter and Facebook, and subscription-aware broadcasting services (SBSs), e.g., Video on Demand and Stock Photo. The popular ways of JPEG FEs, which lead to the great visual distortions while preserving the JPEG coding efficiency, are to transcode-likely XOR or permute the carefully-extracted bits [1], [5], [7] in JPEG file interchange format (JFIF) syntax elements. However, encrypting only the specific portions does not fully confuse the stochastic properties of image textures and provides the hints to recover or leak the image-related information to attackers. Said has stated that if the transcoder does not encrypt the side-information (e.g., the thumbnail image) of JPEG file, most of the encrypted information are recovered with only low-complexity attacks using the information [8].

This work was supported by Grant-in-Aid for JSPS Fellows, Grant Number 20J14599.

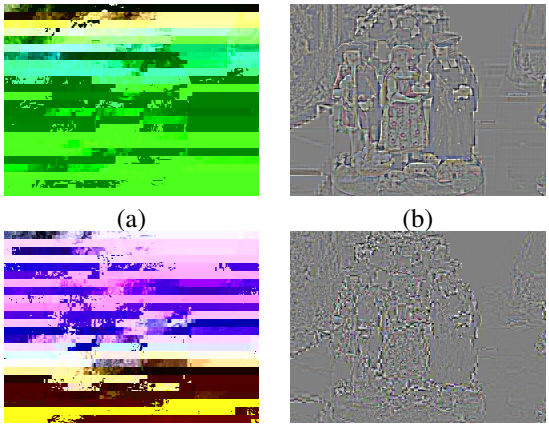


Fig. 1. Texture information leakages by the replacement attack against sign-flipped images: (a) only DC residuals are sign-flipped, (b) DC coefficient deletion to (a), (c) both DC residuals and AC coefficients are sign-flipped, and (d) DC coefficient deletion to (c).

This work focuses on the quantized discrete cosine transform (QDCT) domain JPEG FEs (QJEs). Since the thumbnail image is generated by the JPEG file packetizer, if the original image is encrypted in QDCT domain, the thumbnail image is also encrypted. Here, let the *density* properties for QJEs be as follows:

density-1: Positions of original nonzero coefficients (block-internally and externally).
density-2: Values/magnitudes of original nonzero coefficients.
density-3: Ratio of zero and nonzero coefficients.

Full inter-block shuffle (FIBS) [2] confuses *density*-1 (block-externally) by randomly permuting the QDCT coefficients in the same band among the QDCT coefficient blocks and provides a certain level of the attack resilience at the expense of the coding efficiency. Run-level pair scrambling (RLP) [3] also confuses *density*-1 (block-internally) and preserves the coding efficiency. The QJEs that confuse only *density*-1 are quite high resilient against the replacement attacks, which overwrite the encrypted DC coefficients as zeros (DC coefficient deletion) and the encrypted AC coefficient signs as positive signs '+' (AC sign affirmation) in low-complexity. However, they remain potentially vulnerable to pairwise matching attacks such as jigsaw-puzzle assembling [9] after the JPEG decoding, because the "values" of all coefficients remain unencrypted.

978-1-6654-2545-2/21/$31.00 ©2021 IEEE

TABLE I
density PROPERTIES ACHIEVED BY QJES.

QJE	*density*-1	*density*-2	*density*-3
FIBS [2]	✓		
RLP [3]	✓		
RSF [4], [6]		✓	
FIBS/RLP and RSF	✓	✓	
Proposed AC-WPE	✓	✓	✓

Otherwise, random sign flip (RSF) [4], [6] confuses *density*-2 by flipping the signs of QDCT coefficients and provides a certain level of visual distortions, while preserving the coding efficiency and the low-complexity. A joint approach with the FIBS/RLP and the RSF can confuse both of *density*-1 and 2. However, the QJE that confuses only *density*-2 (signs) is low resilient against the replacement attacks. The encrypted (sign-flipped) DC residuals (prediction errors) after the differential pulse-code modulation (DPCM) are propagated among the QDCT coefficient blocks and the propagation superficially conceals the texture information represented by nonzero AC coefficients. On the other hand, since the AC ones themselves are directly sign-flipped without the DPCM, the replacement attacks reveal the outline of the texture information (Fig. 1). Furthermore, there is no existing QJE, which confuses *density*-3.

To reinforce the perceptual degradation and the texture protection against the replacement attacks at the same complexity as the RSF, we propose an AC prediction error propagation-based encryption (AC-WPE) in the QDCT domain. We build a provisional AC predictor and its reconstructor in the QDCT domain of both the encoder and decoder. In addition, we place an encryptor (a decryptor in the case of the decoder), which randomly flips the signs of the AC residuals (prediction errors), between the AC predictor and reconstructor. As a result, the AC reconstructor propagates the errors among several 8×8 QDCT coefficient blocks, i.e., the AC-WPE confuses all of *density*-1, 2, and 3 as summarized in Table I. Moreover, to suppress the decrease in coding efficiency, we introduce the randomized prediction intervals into the AC predictor/reconstructor. Experiments with JPEG encryption show that the AC-WPE reinforces the perceptual degradation and the resilience to the replacement attacks, while reasonably expensing the bitrate overheads of crypto-compressed images.

II. PRELIMINARIES

A. JPEG

JPEG is a de-facto image coding standard consisted of the following procedure:

1) Transform RGB signals of a full-color image to YCbCr signals and then spatially downsample the Cb and Cr ones.
2) Transform the YCbCr signals to the frequency signals with discrete cosine transform (DCT) with respect to each 8×8 transform coefficient block and then quantize the block with the quantization matrix derived from the quality factor $Q \in \mathbb{Z}_{[1\ 100]}$. Note that each 8×8 block is composed of 1 DC coefficient and 63 AC coefficients, and this quantized domain is called QDCT domain in this work.
3) Transform fixed-length binary strings of QDCT coefficients into variable-length ones to compactly packetize them with Huffman coding after predicting inter-block DC coefficients with the DPCM and bundling zig-zag scanned AC run-level pairs with run-length encoding (RLE) in each block. The packetized bitstreams are stored into the entropy-coded segments (ECS) of JFIF syntax structure.

On the other hand, JPEG decoder reconstructs the decompressed images in the inverse fashion of the above procedure.

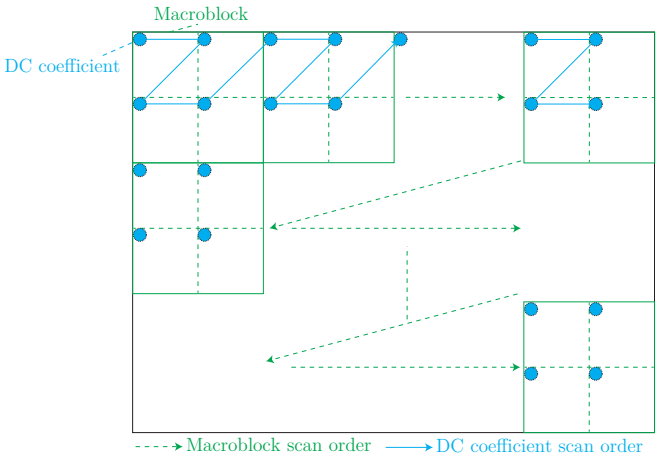


Fig. 2. Forward/inverse DPCM scan order for DC coefficients in the case of luma component at [4:2:0] option.

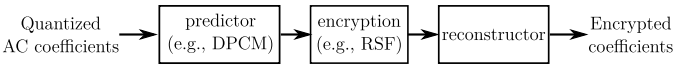


Fig. 3. Flow of AC-WPE encryptor.

B. Differential Pulse-code Modulation

Since the DC coefficients are larger than any AC coefficients in each block but the DC ones correlate between the adjacent blocks, JPEG employs the DPCM to efficiently decorrelate them in QDCT domain. The DC coefficients are predicted among the same component from 8×8 blocks to $8h \times 8v$ $(h, v \in \{1, 2\})$ macroblocks as shown in Fig. 2[1]. When the predictor reaches the rightest macroblock of the same row, it turns back to the leftest macroblock of the next row.

Let $\{DC_i\}_{i=1}^{n}$ $(n \in \mathbb{N})$ and $\{\widetilde{DC}_i\}_{i=1}^{n}$ be DC coefficients of n 8×8 blocks in any QDCT domain of Y, Cb, and Cr signals

[1]By default, the JPEG employs the default option [4:2:0] ($h, v = 2$ for the luma component and $h, v = 1$ for the chroma components), and it can be changed to [4:4:4] ($h, v = 2$ for all compoments) or [4:2:2] ($h, v = 2$ for the luma component and $h = 1, v = 2$ for the chroma components).

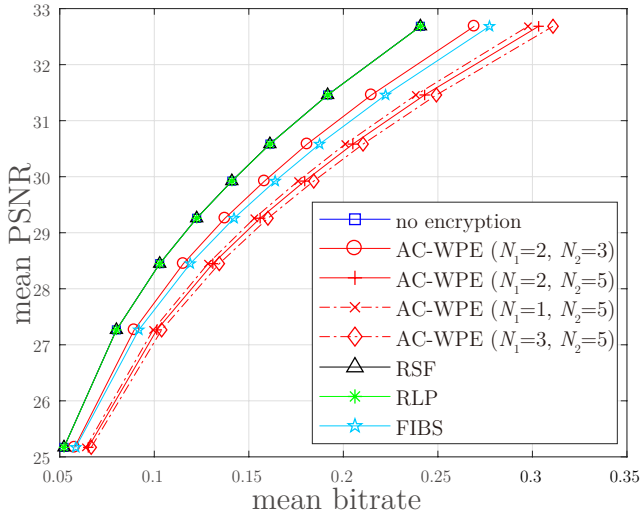


Fig. 4. Rate-distortion performances of AC-WPE, FIBS, RLP, and RSF.

and its residuals with the forward DPCM at the encoder as

$$\widetilde{DC}_{i+1} = DC_{i+1} - DC_i \quad \text{if} \quad i > 1, \\ \widetilde{DC}_1 = DC_1 \quad \text{otherwise}, \tag{1}$$

respectively. Also, the residuals are reconstructed to the original ones by the inverse DPCM at the decoder as

$$DC_{i+1} = \widetilde{DC}_{i+1} + DC_i \quad \text{if} \quad i > 1, \\ DC_1 = \widetilde{DC}_1 \quad \text{otherwise}. \tag{2}$$

C. Random Sign Flip

The RSF is a low-complexity encryption technique, which randomly flips only the signs of nonzero QDCT coefficients. Let $\{QC\}_{i=1}^{m}$ be the original m QDCT coefficients composed of $\lfloor m/64 \rfloor$ DC coefficients $\{DC\}_{j=1}^{\lfloor m/64 \rfloor}$ and $m - \lfloor m/64 \rfloor$ AC coefficients $\{AC\}_{j=1}^{m \quad \lfloor m/64 \rfloor}$. They are randomly sign-flipped as

$$QC'_i = \begin{cases} -QC_i & \text{if} \quad \text{rand}(i) = 1, \\ QC_i & \text{otherwise}, \end{cases} \tag{3}$$

where $\text{rand}(i)$ is a random binary number generated at ith QDCT coefficient from pseudo random binary sequence (PRBS) seeded by the SHA-256 [10] digest-like long encryption key.

If $\{DC\}_{j=1}^{\lfloor m/64 \rfloor}$ has already been predicted, i.e., only DPCM is conducted within 3. in the procedure of Section II-A, the DC residuals are sign-flipped as

$$\widetilde{DC}'_j = \begin{cases} -\widetilde{DC}_j & \text{if} \quad \text{rand}(j) = 1, \\ \widetilde{DC}_j & \text{otherwise}, \end{cases} \tag{4}$$

which are stored as binary strings into the back of Huffman codewords denoting their bit-lengths in the JFIF ECS. Since JPEG decoder reconstructs the sign-flipped DC residuals with the inverse DPCM, the encrypted residuals are propagated among 8×8 blocks (in the order of Fig. 2) as

$$DC'_{j+1} = \widetilde{DC}'_{j+1} + DC'_j \quad \text{if} \quad j > 1, \\ DC'_1 = \widetilde{DC}'_1 \quad \text{otherwise}. \tag{5}$$

Since the most of the energy in QDCT domain is concentrated in the DC coefficients, the encrypted DC residuals are propagated to conceal the most of texture information represented by nonzero AC coefficients (Fig. 1(a)). However, the texture information are revealed by the replacement attacks to the DC residuals (Fig. 1(b)). Even if the AC coefficients are simultaneously sign-flipped (Fig. 1(c)), the most of texture information are still not confused (Fig. 1(d)).

III. AC PREDICTION ERROR PROPAGATION-BASED ENCRYPTION

A. Encryption and Decryption Algorithms within JPEG Coder

A provisional AC predictor and its reconstructor are set in QDCT coefficient processing part of the JPEG coder to encrypt/decrypt arbitrary QDCT residuals as shown in Fig. 3. By reconstructing the encrypted residuals, the errors are propagated among the several QDCT coefficient blocks. In this work, we employ the DPCM, which is one of the most simplest predictors, as the provisional AC predictor to keep the complexity low. The AC predictor changes the b ($\in \mathbb{Z}_{[1\ 63]}$)th band of n AC coefficients $\{AC_i^{(b)}\}_{i=1}^{n}$ into the residuals $\{\widetilde{AC}_i^{(b)}\}_{i=1}^{n}$ as

$$\widetilde{AC}_{i+1}^{(b)} = AC_{i+1}^{(b)} - AC_i^{(b)} \quad \text{if} \quad i > 1, \\ \widetilde{AC}_1^{(b)} = AC_1^{(b)} \quad \text{otherwise}, \tag{6}$$

in the same scan order as Fig. 2, and then the residuals $\{\widetilde{AC}_i^{(b)}\}_{i=1}^{n}$ are randomly sign-flipped by using (3) as

$$\widetilde{AC}_i^{(b)\prime} = \begin{cases} -\widetilde{AC}_i^{(b)} & \text{if} \quad \text{rand}(i) = 1, \\ \widetilde{AC}_i^{(b)} & \text{otherwise}. \end{cases} \tag{7}$$

Note that the difference of (7) from (3) for the AC coefficients is that the sign-flipped signals are residuals like (4) for the DC ones. To propagate the encrypted prediction error, the residuals $\{\widetilde{AC}_i^{(b)\prime}\}_{i=1}^{n}$ are reconstructed by the AC reconstructor as

$$AC_{i+1}^{(b)\prime} = \widetilde{AC}_{i+1}^{(b)\prime} + AC_i^{(b)\prime} \quad \text{if} \quad i > 1, \\ AC_1^{(b)\prime} = \widetilde{AC}_1^{(b)\prime} \quad \text{otherwise}. \tag{8}$$

Also, the encrypted bth band coefficients $\{AC_i^{(b)\prime}\}_{i=1}^{n}$ are completely decrypted with the same encryption key and the same fashion of (6) to (8) in the JPEG decoder.

B. Randomized Prediction Intervals

Since the proposed encryption further increases the magnitudes of original AC coefficients, whose values are originally not smooth, they tend to significantly increase the bitrate overheads. To suppress the overheads, we should mitigate

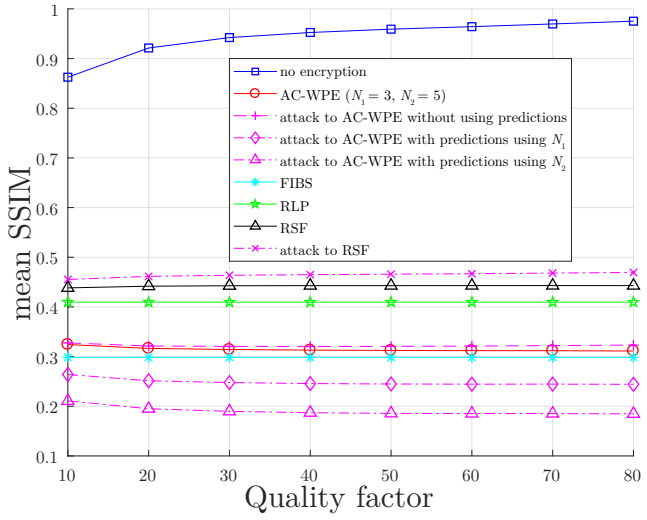


Fig. 5. Resilience against replacement attacks of AC-WPE, FIBS, RLP, and RSF.

the increase in the encrypted AC coefficient values. In this work, we set the prediction intervals to block the increasing prediction errors as a special case. However, if the encryptor uses only an explicit fixed interval N ($\in \mathbb{Z}_{[1\ n]}$), an attacker may know and abuse the first AC coefficient positions for each interval. To confuse the first coefficient positions for each interval, we randomize the prediction interval in a pre-defined section (N_1, N_2) ($N_1, N_2 \in \mathbb{Z}_{[1\ n]};\ N_1 < N_2$). Although a fixed N equally splits the entire prediction interval $\mathcal{I}$, the section (N_1, N_2) splits the $\mathcal{I}$ into

$$\{I_i | I_i = \mathrm{rand}(i)_{[N_1\ N_2]}\}, \tag{9}$$

where $\mathrm{rand}(i)_{[N_1\ N_2]} (\in \mathbb{Z}_{[N_1\ N_2]})$ is a random number from N_1 to N_2 generated at the ith interval by the pseudo random number generator (PRNG) [11] seeded by the SHA-256 [10] digest-like long encryption key. For each I_i, the AC coefficients $\{AC_k^{(b)}\}_{k=S+1}^{S+I_i}$ ($S = \sum_{j=1}^{i} I_j$) are encrypted by using (6) to (8). Since I_i is randomly determined among (N_1, N_2), an attacker cannot easily know the position of the first AC coefficient for each interval.

C. Complexity and Property

First, we show the complexity as the computational order. Since the RSF encrypts all n QDCT coefficients for each $\mathcal{O}(1)$ order, the total order is $\mathcal{O}(n)$. For each band, the AC-WPE predicts all $m = n/64$ coefficients in $\mathcal{O}(m)$ order, sign-flips them for each in $\mathcal{O}(1)$ order, and finally reconstructs them in $\mathcal{O}(m)$ order: every band of all m coefficients are encrypted in $\max(\mathcal{O}(m), \mathcal{O}(m), \mathcal{O}(m)) = \mathcal{O}(m)$ order. Because $\mathcal{O}(64m) = \mathcal{O}(n)$, the AC-WPE encrypts all coefficients in $\mathcal{O}(n)$. Therefore, the AC-WPE also requires the same order as the RSF.

Next, we show the property about the density. Although the original AC coefficient positions are not changed, *density*-1 and *density*-3 are confused because even the zero (nonzero) AC coefficients become nonzero (zero) ones after the AC-WPE. Also, since the encrypted prediction errors change the original coefficient values, *density*-2 is naturally confused. Therefore, the AC-WPE reinforces the resilience against the jigsaw-puzzle assembling and the replacement attacks. Although the self-evident combination of the AC-WPE and the other methods such as the FIBS further reinforces the resilience, the actual demonstration is omitted here for simplicity.

IV. Experiments

A. Conditions and Procedure

We exploited 100 test images from UCID dataset [12], whose resolutions are of 512×384 or 384×512, used *libjpeg-turbo* [13] as the JPEG software, and chose the FIBS [2], RLP [3], and RSF [4], [6] as the conventional QJEs to be compared. Also, we did not process the DC coefficients at the objective comparisons on coding efficiency and attack resilience to compare the effectiveness of AC coefficient encryptions. On the other hand, we involved the sign-flip of DC coefficients after the DPCM at the subjective comparisons of encryption/attack results to observe the actual texture leakages. The experimental procedure was as follows:

i: JPEG-encode a test image with a quality factor $Q = 10, 20, \cdots, 80$, while encrypting the QDCT coefficients by the FIBS, RLP, RSF, and AC-WPE.
ii: Measure the bitrates of nonencrypted- and encrypted-encoded bitstreams.
iii: JPEG-decode them while decrypting the encrypted QDCT coefficients (and attacking the encrypted QDCT coefficients with the replacement attacks as needed).
iv: Measure PSNRs and SSIMs [14] of the nonencrypted-, encrypted-, and decrypted-decoded images from the original one.
v: Iterate from 1. to 4. for the test images to obtain the mean results.

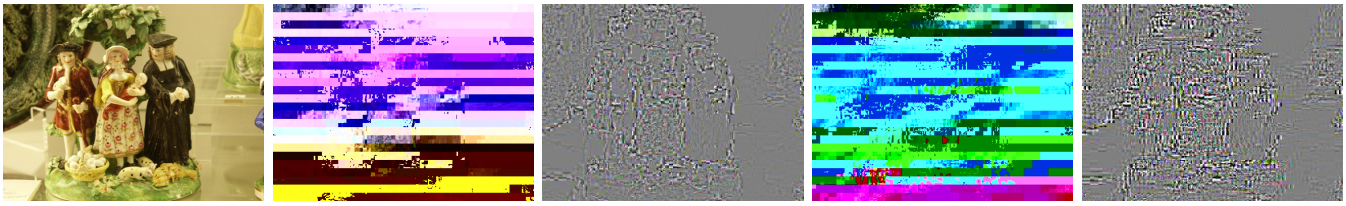
B. Results and Discussions

We experimented two kinds of trials: coding efficiency and resilience against the replacement attacks such as the DC coefficient deletion and the AC sign affirmation, to show the characteristics of the AC-WPE.

First, we compared the coding efficiency of encrypted QDCT coefficients by using the AC-WPE, FIBS, RLP, and RSF. One can see that although the AC-WPE imposed the larger amount of bitrate overheads than the RSF and RLP at any selection of (N_1, N_2), it partly suppressed the larger amounts of ones compared with the FIBS, as shown in Fig. 4.

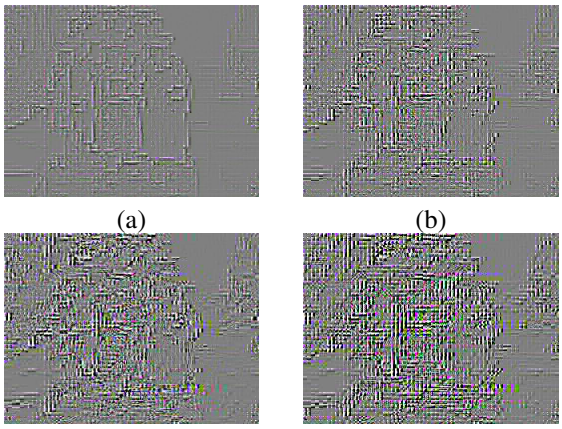
Next, we compared the the perceptual degradation amounts and the resilience against the replacement attacks. The AC-WPE provided the larger amounts of perceptual degradation than any of FIBS, RLP, and RSF, as shown in Fig. 5. Also, the AC-WPE provided the higher resilience against the replacement attacks than the RSF[2]. Even if an attacker replaced

[2]It is obvious that the FIBS and RLP are resilient against the replacement attacks because they confuse *density*-1.



(a) (b) (c) (d) (e)

Fig. 6. Comparison of perceptual degradation encrypted with RSF and AC-WPE and attacked with the replacement attacks (JPEG $Q = 70$): (a) original (*ucid00055*), (b) RSF, (c) DC coefficient deletion to (b), (d) AC-WPE ($N_1 = 3, N_2 = 5$), and (e) DC coefficient deletion to (d) without predictions.



(a) (b) (c) (d)

Fig. 7. Reinforcement of texture leakage prevention: (a) AC sign affirmation to Fig. 6(c), (b) AC sign affirmation to Fig. 6(e), (c) AC sign affirmation to Fig. 6(e) with predictions using $N_1 = 3$ and (d) AC sign affirmation to Fig. 6(e) with predictions using $N_2 = 5$.

the AC coefficient signs without any predictions (i.e., without considering any provisional AC predictor/reconstructor), the AC-WPE achieved the higher resilience against the attacks than the RSF, as shown in the pink dotted lines symbolized with '$+$' and '$\times$' of Fig. 5. Even subjectively, the AC-WPE induced more significant amount of perceptual degradation that cannot be revealed with the attacks than the RSF, as shown in Fig. 6; note that although the DC coefficient deletion easily leaked the outline superficially concealed with the error-propagated DC coefficients (Fig. 6(b,c)), the proposed AC-WPE prevented it (Fig. 6(d,e)). Moreover, even if an attacker tries to replace the signs with the predictions, the resilience was kept high as shown in the pink dotted lines symbolized with '$\Diamond$' and '$\triangle$' of Fig. 5, because the attacker is anticipated to predict the AC coefficients with respect to the equally-split interval by using N_1 or N_2. On the contrary, we observed the error-propagated AC coefficients, whose signs were affirmed in the residual state by using mistaken intervals, further reinforced the prevention of the texture leakages (Fig. 7).

V. Conclusion

To reinforce the resilience of the conventional random sign flip (RSF), we proposed an AC prediction error propagation-based encryption (AC-WPE) in JPEG quantized discrete cosine transform (QDCT) domain. We built a provisional AC predictor and its reconstructor in the QDCT domain to design the AC-WPE. Since the encrypted prediction errors are propagated among several 8×8 QDCT coefficient blocks, the AC-WPE confused the positions, values, and ratio of zero and nonzero AC coefficients. Moreover, we suppressed the decrease in coding efficiency by introducing the randomized prediction intervals into the AC predictor/reconstructor. Experiments with JPEG encryption showed that the AC-WPE reinforced the perceptual degradation and the resilience against the replacement attacks, while reasonably expensing the bitrate overheads of crypto-compressed images.

References

[1] C. Shi and B. Bhargava, "A fast MPEG video encryption algorithm," in *Proc. ACMMM*, Bristol, UK, Sep. 1998, pp. 81–88.

[2] W. Li and Y. Yuan, "A leak and its remedy in JPEG image encryption," *Int. J. Comput. Math*, vol. 84, no. 9, pp. 1367–1378, Sep. 2007.

[3] A. Unterweger and A. Uhl, "Length-preserving bit-stream-based JPEG encryption," in *Proc. MM&Sec*, Coventry, UK, Sept. 2012, pp. 85–90.

[4] B. Zeng, A. Yeung, S. Kei, S. Zhu, and M. Gabbouj, "Perceptual encryption of H.264 videos: Embedding sign-flips into the integer-based transforms," *IEEE Trans. Inf. Forensics Secur.*, vol. 9, no. 2, pp. 309–320, Feb. 2014.

[5] A. Unterweger, K. V. Ryckegem, D. Engel, and A. Uhl, "Building a post-compression region-of-interest encryption framework for existing video surveillance systems," *Multimedia Syst.*, vol. 22, no. 5, pp. 617–639, Oct. 2016.

[6] P. Li and K.-T. Lo, "Joint image compression and encryption based on order-8 alternating transforms," *J. Vis. Commun. Image*, vol. R, no. 44, pp. 61–71, Apr. 2017.

[7] V. Itier, P. Puteaux, and W. Puech, "Recompression of JPEG crypto-compressed images without a key," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 3, pp. 646–660, Mar. 2020.

[8] A. Said, "Measuring the strength of partial encryption schemes," in *Proc. ICIP*, Genova, Italy, Sep. 2005, pp. 1126–1129.

[9] G. Paikin and A. Tal, "Solving multiple square jigsaw puzzles with missing pieces," in *Proc. CVPR*, Boston, MA, June 2015, pp. 4832–4839.

[10] National Institute of Standards and Technology, *FIPS PUB 180-4: Secure Hash Standard (SHS)*, pub-NIST, Aug. 2015.

[11] M. Matsumoto and T. Nishimura, "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Trans. Model. Comput. Simul.*, vol. 8, no. 1, pp. 3–30, Jan. 1998.

[12] G. Schaefer and M. Stich, "UCID: An uncompressed color image database," in *Proc. SPIE*, San Jose, CA, Jan. 2004, pp. 472–480.

[13] "JPEG software," https://jpeg.org/jpeg/software.html.

[14] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.