Image Authentication and Restoration by Multiple Watermarking Techniques with Advance Encryption Standard in Digital Photography

Sidra Riaz*, Sang-Woong Lee**

*' **Department of Computer Engineering, Chosun University, Gwangju, South Korea

sidra.riaz426@gmail.com, swlee@chosun.ac.kr

**Corresponding Author

***Abstract*— In the past few years, semi-fragile watermarking techniques have become increasingly important to secure and verify the multimedia content and also to localize the tampered areas, while tolerating some non-malicious manipulations or attacks. In digital photography, watermarking schemes are very important for copyright protection purposes. In this paper, a multimedia authentication and restoration scheme is proposed with the security of AES-128 ciphered watermarking and correlated watermarking. An encrypted or ciphered image embedding is done by modified version of Closest Point Transform (CPT) in a digital photograph. We performed several security attacks e.g. noise attack, compression attack, and cropping attack on multiple watermarked photographs and evaluated the proposed watermarking technique to examine the system robustness. Image Authentication is done by locating the tempered areas and restoration is performed by correlated watermark on the tempered region of watermarked photograph. The PSNR values are checked to evaluate the proposed watermarking technique. The results of PSNR, MSE, and SSIM show that the imperceptibility of our scheme is high compared to existing methods.**

***Keywords*— Multimedia authentication, image restoration, multimedia security, digital photography, noise attacks, cropping attack, compression attack, Advance Encryption Standard**

I. INTRODUCTION

This is an era of advanced communication and modern technologies. With the evolution of digital and hand-held devices it is easy to make the digital contents. Multimedia content e.g. video, voice, and images are also saved in digital form. These contents are shared over online public community websites for various purposes [1] with the aspiration of copyright protection and authorization. To secure the multimedia content various techniques are used e.g. cryptography, steganography, and watermarking. Each technique is used for its purpose. Cryptographic techniques are used to change the meaning of the documents [2]. Steganographic techniques are used to conceal the existence of the important content [3]. Watermarking schemes are used for protection and or authentication of multimedia content [4]. There are three categories of watermarking techniques [5] developed for specific goals, i.e. 1) robust watermarking, 2) fragile watermarking, and 3) semi-fragile watermarking. Robust watermarks can stand up to malicious attacks and are difficult to remove from multimedia content, thus used for copy-right protection purposes. Fragile watermarks are easy to break and to remove from the multimedia contents, thus their existence or absence leads us to the conclusion of authentication. Semi-fragile watermarking techniques are used for both authentication and protection of data. Multimedia content authentication applications [4], [6] are in the domains but are not limited to defense, law, commerce, online banking, surveillance, emailing, and sharing multimedia data etc.

In this paper, we have used semi-fragile watermarking technique for digital photographs. Its real time application is that the photographers want to share their art and also want the ownership copyright protection. The photographs should be secure and also the embedded watermark should not degrade the quality of the photographs. The degradation of quality of photographs limits the amount of data being embedded. For security purposes, the watermark should be robust and perceptually invisible so it could not be removed from the photographs. Watermark must survive the malicious attacks too. We hereby deal with the mentioned criterions, problems and methods for watermarking the digital photographs. The authentication of the photographs can be achieved by locating the parts of the photographs that are changed. If some areas in the photograph are changed or removed to change the meaning of the picture, so restoration of these malicious parts is done to solve the problem. In short, the watermarking authentication and restoration process of digital photographs is a new and an important area of modern research.

Many researchers have put their efforts in the area of watermarking. Their contributions include content authentication [7], content verification [8], tampering localizations [9], content restoration [10] etc. However, robust watermarking techniques are insensitive to malicious attacks and thus it is difficult to detect the distortions or alterations performed in the watermarked content. Verification techniques are also proposed in literature which can detect that the content has undergone some distortions but they cannot

locate the tamper locations. Content tampers localization techniques [9] are available and they can localize the tampered areas but they fail to recover the tampered region. For authentication and additional ability of recovering the distorted image, multiple watermarking technique is required. Previous work in this area has been related to self-restoration techniques which are capable of restoring tampered regions once the regions localizations is done [10].

In this paper we use the security of AES-128 [11] and embed multiple watermarks: encrypted watermark and correlated watermark images. We encrypt the first watermark by AES, and correlate the second watermark with original photograph. The first watermark embedding is done in the original photographs by our modified CPT algorithm [12]. The second watermark is embedded in the wavelet sub-bands. First watermark is used for authentication purposes and second watermark is used for recovering the estimation of original photograph. By combining the restored version and the tampered one, we recover the photograph with a good quality.

Rest of the paper structure is organized as follows: Section 2 discusses our proposed approaches. In Section 3 we show the results and in Section 4 we conclude our paper.

II. PROPOSED APPROACHES

A. Encrypted Watermark with Advance Encryption Standard (AES-128)

AES is a symmetric block cipher [11]. It uses the block size of 128-bits and a key size of 128, 192 or 256 bits. Each full round of AES uses four functions: byte substitution, permutation, arithmetic operations, and XOR with the generated key. We use the security of AES-128 to encrypt the first watermark image as block by block approach by dividing the image into 4 blocks each of 128×128 bits. First 128 bits in the first row of block#1 are input to AES-128 module. The full block of 128×128 size is then converted to 4×4 square matrix of bytes. The cipher consists of N-rounds. The number of rounds depends on the key length. We used the key length of 128 bits and thus total rounds become 10. So we get the encrypted watermark, W_e in total 10 rounds. These rounds are taken to avoid the bruit force attack.

B. Correlated Watermark

The second watermark image is created through correlation of the original photograph *I(x,y)* by using below given Eq. 1

$$W_h = \frac{\sum_{x=1}^{X}\sum_{y=1}^{Y} I^{'}(x,y)\times I(x,y)}{\sum_{x=1}^{X}\sum_{y=1}^{Y} I^{'}(x,y)^2} \qquad \ldots(1)$$

where I'(x, y) is the inverse of the photograph. W_h is the correlated watermark image.

C. Embedding Watermark

The encrypted watermark (w_e) and correlated watermark (w_h) are embedded into the cover image respectively. First watermark is embedded through extended version of CPT algorithm. CPT algorithm was first introduced by Tseng et al., [12] for data hiding schemes in binary images. The extension of it was presented in [13] and [14]. We modified the CPT algorithm to name it as Featured-Closest Point Transform (F-CPT). Firstly, we define the steps of the CPT algorithm and then we explain the F-CPT methods and the motivation of its development.

1) CPT Algorithm:

C: is a cover image which is divided into m*n $\{C_1,.....C_y\}$ sized blocks.

W: is a weight matrix of size m*n, where $w_{i,j} = 1...2^r-1,1...2r-1,1...L, L \le 2^r-1$ $i=1,...m,\ j=1,...n$

R: embeddable bits in one block; $r \le \lfloor \log(mn+1) \rfloor$

BEGIN

Step 1: For collection of bits $\{b_1...b_r\}$ to be embedded in block C_i, do the following.

Step 2: Calculate $C_i \oplus K$, where $\oplus$ is XOR operation.

Step 3: Calculate $H = SUM((C_i \oplus K)\otimes W)$, where $\otimes$ is pair-wise multiplication of equal sized matrices.

Step 4: For each w, w=1,…, 2^r-1 Let $S_w = \{(9j,k);(W_{j,k}=w)$ and $[C_i \oplus K]_{j,k}=0$, OR $W_{j,k}=2^r-w$ and $[C_i \oplus K]$=1}

Step 5: let d= $b_1...(b_r-H) \bmod 2^r$

Step 6: IF d=0; there is no change in block C_i

ELSE

a) Randomly select hϵ{0,1,.., 2^r-1} such that $S_{hd} \neq \Theta$ and $S_{-(h-1)d} \neq \Theta$

b) Randomly select (j,k) ϵS_{hd} and complete the bit $[C_i]_{j,k}$

c) Randomly select (j,k) $\epsilon S_{-(h-1)d}$ and complement the bit $[C_i]_{j,k}$

IF (S_0) skip this step

ENDELSE

END

In this CPT algorithm, only 2 bits/block are changed in the first bit plane and remaining bits in C_i block are untouched. We can also use the remaining bits to increase the embeddable payload size in Kb's. We have performed the embedding in important locations per block by extracting the features from digital photographs. The features include saturation, intensity, and normalized contrast as defined in Eq. 2, 3 and 4 respectively [15]. Based on these features, we select the region of interest points per block C_i, and embed the encrypted watermark in the first bit plane, i.e. the decided threshold values of the features within each block give us the choice to use that point for embedding.

$$f1 = \frac{1}{X * Y} \sum_{x}^{X-1} \sum_{y}^{Y-1} I_s(x, y) \quad ...(2)$$

$$f2 = \frac{1}{X * Y} \sum_{x}^{X-1} \sum_{y}^{Y-1} I_v(x, y) \quad ...(3)$$

$$f3 = \sum_{s,y} \{ | x - y |^2 * I_{GLCM}(x, y) \} \quad ...(4)$$

The second watermark (W_h) is used for the recovery of the tampered areas. For embedding in the cover image, we used the wavelet coefficient of W_h and embedded in the second level wavelet sub-band {HH2} and in third level wavelet sub-band {VV3}. Detailed embedding of coefficients of W_h is explained in [16].

D. Retrieve and Decrypt Watermark

It is the inverse process of embedding and encrypting the watermark. First watermark is obtained by first bit plane of the feature points from every block of cover image. Inverse AES is applied by the authorized key (128 bits) for each 4 blocks. Second watermark is retrieved from the wavelet sub-bands {HH2} and {VV3}.

E. Image Authentication and Localization

A semi-fragile watermarking scheme should authenticate the watermarked image. The overall content authentication system has 3 parts: generation function (MD5 used in our work), embedding and extracting. First watermark is extracted and then decrypted by the inverse process of AES-128 [11]. Generation function (f_g) is added in the host image with the watermark and the AES key (128-bits). Watermark can be represented by generation function as given by Eq. 5.

$$W1 = f_g(i, K, I) \quad ...(5)$$

The watermark data W1 is added in the cover image by the embedding function as represented in Eq. 6. I'(x, y) is the watermarked data.

$$I' = E(I, W, K) \quad ...(6)$$

Extracting function which will work as an authenticator of watermark is given in Eq. 7.

$$W' = D(I', K) \quad ...(7)$$

Decrypted watermark and *W'* are correlated and the tampered locations are obtained i.e. the regions where they are different. The tampered locations are denoted by a cross on that pixel or by white pixels.

F. Image Recovery

For image recovery we used the correlated image itself which is called self-recovery process. We recover the input image through decomposed watermark. The recovery process gives good results for various tampering.

G. Malicious Attacks

Noise attack, cropping attack, and compression attacks are performed to check the imperceptibility of our system. The details of PSNR, MSE, and SSIM are shown in the results section.

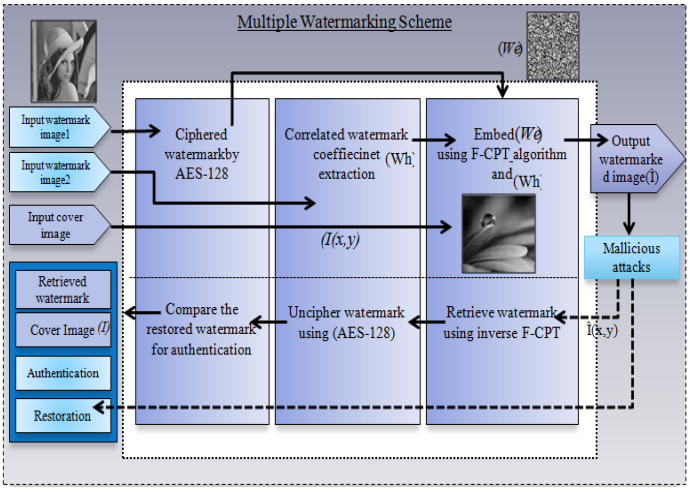
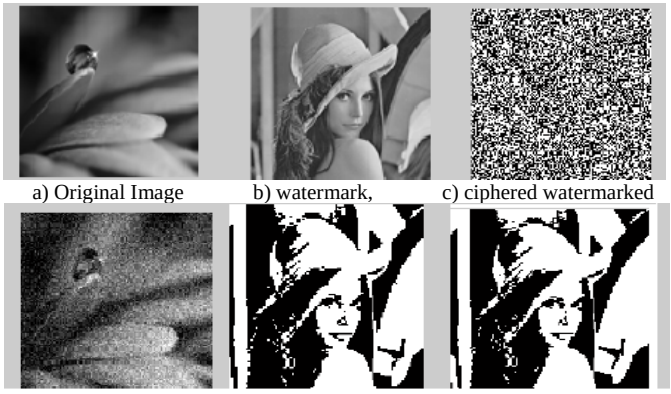


Figure 1. Overall system diagram with main modules



a) Original Image b) watermark, c) ciphered watermarked

d) noised attack, e) Original watermark, f) recovered watermark

Figure 2. The results of embedding the watermark with AES-128 ciphering and recovering the watermark after noise attack

a) Original Image b) tampered, c) recovered, d) tampered location

Figure 3. The results of compression and some other malicious attacks performed on Lena image and restoring the watermarked image

The overall system diagram of proposed approaches is shown in Fig. 1 with main modules.

III. RESULTS

The correlation of *W'* and decrypted watermark provides us with the tampered locations where malicious attacks are performed. The restoration is achieved by correlated watermark W_h.

The results of watermark recovery can be seen in Fig. 2 and Fig. 5 for different cases. The results of tamper localization and restoration can be seen in Fig. 3 and Fig. 4 for different cases. For imperceptibility we have calculated MSE (mean square error), PSNR (peak signal-to-noise ratio), SSIM (structural similarity index) etc. We have calculated the PSNR of multiple watermarked photographs to check if it is perceptually aesthetic or not. PSNR (Peak signal-to-noise ratio) is calculated by using Eq. 8. The MSE (mean square error) between original and watermarked photograph is calculated by using Eq. 9. Structural similarity index (SSIM) is a quality parameter and used to verify the content that is similar or changed by some attacks. It is given by Eq. 10.

$$PSNR = 10.\log_{10}\left(\frac{MAX_I^{\ 2}}{MSE}\right) \quad \ldots(8)$$

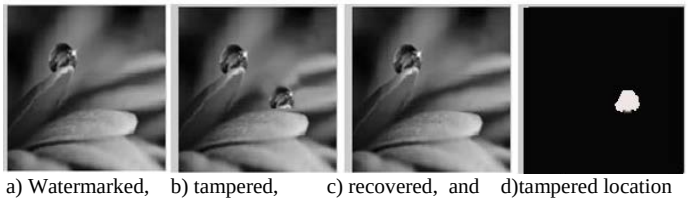
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(X_i - X_i^*)^2 \quad \ldots(9)$$

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^{\ 2} + \mu_y^{\ 2} + c_1)(\sigma_x^{\ 2} + \sigma_y^{\ 2} + c_2)} \quad \ldots(10)$$

where, μ, σ and σ_{xy} are mean, variance, and covariance of the image and c_1, c_2 are the stabilizing constants. The results of PSNR, MSE, and SSIM are shown in Table 1.

TABLE 1. RESULTS OF PSNR, MSE AND SSIM OF WATERMARKED IMAGES

Sr. No	Original Image	Watermark	MSE	PSNR	SSIM
1	Image#1	AES-128 Ciphered Lena Image and correlated watermark(W_h)	0.016	55.63dB	0.813
2	Image#2	AES-128 Ciphered Lena Image and correlated watermark(W_h)	0.0128	60.15dB	0.86
3	Image#3	AES-128 Ciphered Lena Image and correlated watermark(W_h)	0.0154	56.03dB	0.860



a) Watermarked, b) tampered, c) recovered, and d)tampered location

Figure 4. The results of some malicious attacks performed on a photograph

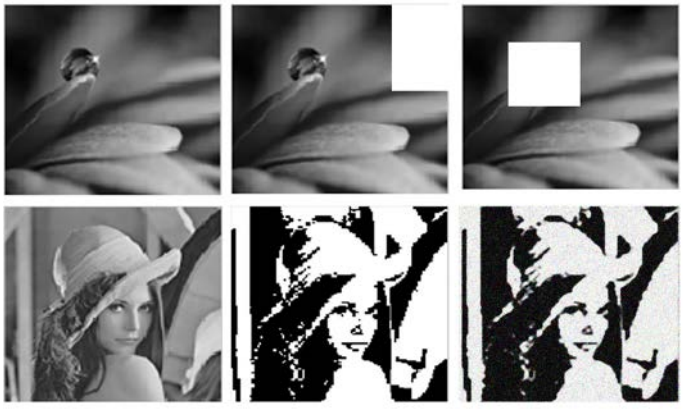


Figure 5. The results of cropping attacks and retrieved watermarks

IV. CONCLUSIONS

In this paper, a multiple watermarking scheme is proposed for digital photographs authentication and restoration. We used the security of AES-128 to make a ciphered watermark and embedded it in the cover image by F-CPT algorithm for content authentication. Image restoration is achieved by correlated watermark embedding in wavelet sub-bands. Several malicious attacks are performed i.e. noise attack, compression attack, and cropping attack etc. The results of PSNR, MSE and SSIM show that the imperceptibility of our system is high and the technique is highly robust.

ACKNOWLEDGMENT

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the Global IT Talents Program supervised by the NIPA (National IT Industry Promotion Agency)" (NIPA-2012-C1156-1001-0003).

REFERENCES

[1] G.-B. Shelly, M.-E. Vermaat, J.-J. Quasney, S.-L. Sebok, J. Jeffrey, "*Multimedia and content sharing,"* in Discovering Computers 2009, Boston, USA: Cengage Learning, 2008.

[2] C. Paar, J. Pelzl, "Introduction to cryptography and data security," in Understanding Cryptography: A Textbook for Students and Practitioners, reprint, Germarny: Springer, 2010.

[3] Z. Duric, M. Jacobs and S. Jajodia, "Information hiding: steganography and steganalysis," in Handbook of Statistics: Data Mining and Data Visualization, USA: Elsevier, 2005.

[4] B. Furht, D. Kirovski, "Protection of multimedia content in distribution networks," in Multimedia Watermarking Techniques and Applications, USA: CRC Press, Taylor and Francis Groups, 2006.

[5] E.-T. Lin, C.-I. Podilchuk, E.-J. Delp, "Detection of image alterations using semi-fragile watermarks," in *Proc.SPIE 3971*, 2000, p. 721-722.

[6] H.-T. Sencar, M. Ramkumar, A.-N. Akansu, "Communication with side information and data hiding," in Data Hiding Fundamentals and Applications: Content Security in Digital Media, London, UK: ELSEVIER Academic Press, 2004.

[7] W. Huang, "Watermarking-based content authentication of motion-JPEG sequences," *in Proc.* VIE, 2008, p. 813-818.

[8] R.-S. Alomari, A. Al-Jaber, "A Fragile watermarking algorithm for content authentication," *in IEEE Trans.* JCIS, vol. 2, No. 1, April, 2004.

[9] S. Dadkhah, A.-A. Manaf, S. Sadeghi, "Efficient digital image authentication and tamper localization techniques using 3Lsb watermarking," in IJCSI, Internation Journal of Computer Science, vol. 9, issue. 1, no. 2, January, 2012.

[10] X. Zhang, S. Wang, "Fragile watermarking with error-free restoration capability," in IEEE Trans. on Multimedia, vol. 10, no. 8, December, 2008.

[11] H. Dobbertin, V. Rijimen, A. Sowa Ed., "Advanced encryption standard-AES," ser. Leture Notes in Computer Science/Security and Cryptography, Bonn, Germany: Springer, 2004, vol. 3373.

[12] Y.-Y. Chen, H.-K. Pan, and Y.-C. Tseng, "A Secure Data Hiding Scheme for Two-Color Images," *in Proc.* of 5th IEEE Symposium on Computers and Communications 2000, 2000, pp. 750–755.

[13] Y.-C. Tseng and H.-K. Pan, "Secure and Invisible Data Hiding in 2-olor Images," *in Proc.* of INFOCOM 2001, 2001, pp. 887–896.

[14] P.-T. Huy, V.-P. Bac, N.-M. Thang, T.-D. Manh, V.-T. Duc, N.-T. Nam, "A new CPT extension for high data embedding ratio in binary images," *in Proc.* of KSE, International Conference on Knowledge and System Engineering, 2009, p. 61-66.

[15] S. Riaz, K.-H. Lee and S.-W. Lee, "Aesthetic score assessment based on generic features in digital photography," 5th AUN/SEED-Net Regional Conference on Information and Communication Technology, Manila, Philippine, October 2012, pp. 76-79.

[16] R. chamlawi, K. Asifullah, I. Usman, "Authentication and recovery of images using multiple watermarks," in Journal of Computers & Engineering, vol. 36, issue. 3, May 2010, p. 578-584.

Sidra Riaz received her B.S degree in Telecom Engineering from National University of Computer and Emerging Sciences (NUCES-FAST), Islamabad, Pakistan, in 2011. She is currently a master student and research assistant in Department of Computer Engineering, Chosun University, South Korea. Her research interests include multimedia and image processing, computational aesthetics, and pattern recognition. Ms. Sidra Riaz is the recipient of National ICT R&D Scholarship award from August 2007 to July 2011 and Global IT scholarship award in August 2011.

Sang-Woong Lee received his BS degree in Electronics and Computer Engineering from Korea University, Seoul, Korea, in 1996 and his MS and Ph.D. degrees in Computer Science and Engineering from Korea University, Seoul, Korea, in 2001 and 2006, respectively. From June 2006 to May 2007, he was a visiting scholar in Robotics Institute, Carnegie Mellon University. Currently, he is an assistant professor in Department of Computer Engineering at Chosun University, Gwangju, Korea. His present research interests include face recognition, computational aesthetics, and brain imaging analysis.