A Feature-Based Fragile Watermarking of Color Image for Secure E-Government Restoration

Lusia Rakhmawati[1,2], Wirawan[1], Suwadi[1],Titiek Suryani[1], Endroyono[1]
1Department of Electrical Engineering, Faculty of Electrical Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
2Department of Electrical Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Surabaya, Indonesia
1lusiarakhmawati17071@mhs.its.ac.id, 2lusiarakhmawati@unesa.ac.id, 1{wirawan, suwadi, titiks, endroyono}@ee.its.ac.id

Abstract— this research developed a method using fragile watermarking technique for color images to achieve secure e-government tamper detection with recovery capability. Before performing the watermark insertion process, the RGB image is converted first into YCbCr image. The watermark component is selected from the image feature that approximates the original image, in which the chrominance value features as a watermark component. For a better detection process, 3-tuple watermark, check bits, parity bits, and recovery bits are selected. The average block in each 2 x 2 pixels is selected as 8 restoration bits of each component, the embedding process work on the pixels by modifying the pixels value of three Least Significant Bit (LSB) . The secret key for secure tamper detection and recovery, transmitted along with the watermarked image, and the algorithm mixture is used to extract information at the receiving end. The results show remarkably effective to restore tampered image.

Keywords— fragile watermarking, tamper detection, tamper recovery

I. INTRODUCTION

In recent years the implementation of Electronic Government (e-government) has been widely used, but the implementation of e-government is still not considered a very important security factor to maintain the integrity, confidentiality, and availability of such digital documents [1]. The integrity aspect relates to the integrity of the data. This aspect ensures that data cannot be altered (tampered, altered, modified) without the permission of the entitled. Some threats to aspects of integrity can be done through access breakthroughs, spoofing, viruses that alter or erase data, and man in the middle attack is attack by inserting themselves in the middle of data transmission [2].

Protection against the attack can be done either by using watermarking technique [1-3], which aims to insert a digital sign (referred to as a watermark) into the origin medium (called host). Watermark is a data string that can be seen (perceptible) or invisible (imperceptible). The invisible watermark can be another image, part of the original image, as well as the original image itself in different sizes [1, 4]. Then the recipient side of the image can be extracted for authenticity of ownership and content integrity.

Current developments, watermarking techniques are also proposed for recovery of tampered digital documents [5-7; 10-12]. In this method not only able to detect the tampered image, but also can localize the areas and recover it. This method uses the image feature information itself as a watermark. The original image is modified by a particular method, mostly reduced to the corresponding size while still considering the adequacy of information that can represent the original image [5], then inserted to obtain the watermarked image. When there is irresponsible tamper or modification, the watermark can be extracted to restore it.

In addition, several studies have used watermarking techniques for the security of e-government documents. Zhang [6] using encryption and descriptions schemes for the security of e-government documents. However, this scheme is vulnerable to use in distributed networks. Therefore, it is proposed a combination with watermarking technique as an additional layer for document security protection. In Al-Haj et al. method [8] used transform domain as a process for e-government protection. This technique used transform domain, their drawback is on their low capacity. Their weaknesses can be overcome by increasing the storage capacity space by inserting more than one watermark while keeping the image quality of the watermark.

Talking about adding insertion capacity, the use of digital watermarking applications as part of the detection of damage and recovery of digital documents can refer to other pre-existing methods, so that it can be applied to the case of e-government documents. Lin et al. [2] proposes that each block can be validated by using additional authentication data that has been inserted into each block itself, while data recovery is embedded in different blocks. The quality of the recovery image depends on the tampered image, if the damage is reduced, so the quality of the recovery results decreases. Lin et al [10] proposed a hieratical method by storing watermark in two different region. While He et al. [4] presents a fragile watermarking scheme for self-recovery using neighbouring block methods to recover it.

A representative watermarking technique for tamper detection with restoration capability is Lin's scheme [11], where the insertion process uses a simple technique, LSB (Least Significant Bit). In this method, detection of block damage is done in 3 levels of hierarchy. If not detected at level 1, it will be detected at the next level with probability close to 1. The advantage of this method is effective in terms of memory because it only needs storage for secret keys and simple algorithms to restore the error. In addition to these advantages, there is a weakness in Lin's algorithm,

XXX-X-XXXX-XXXX-X/XX/$XX.00©20XX IEEE
978-1-5386-8402-3/18/$31.00 ©2018 IEEE

which is too sensitive to pixel damage. For example, even if only one pixel in a 4x4 pixel block is damaged, the entire block will be detected as a tampered block, so recovery will be made to all pixels in the block, thus adding degradation to the image recovery end result.

The method proposed by He [4], selects important information from an image by performing an average value of 3 x 3 neighbouring blocks, and uses 2 LSB for its insertion process, and works on blocks of smaller size, 2 x 2 pixels, so it can make image resolution better after image recovery. This method works well for collage attacks because it uses the 3x3 neighbourhood as part of the authentication bit. Another method Hassan [13] proposed colour fragile watermarking for image authentication. This method used average intensity in each 2x2 non-overlapping blocks as a watermark for tamper detection and recovery. Using a particular image object as an important part of the authentication process, this method is not quite successful if the modified not certain part, but can damage all parts of the image.

This research takes the idea of the last two methods, [4] and [13], which proposes the development of watermark component selection, the watermark insertion process, so it can be applied to improve the integrity and ability of restoration of damaged colour digital documents. The proposed fragile watermarking technique is discussed in Section II. Section III presents the experimental results and comparative analysis. The conclusion is summarized in Section IV.

II. Proposed Self-Embedding Fragile Watermarking Scheme

In this section, the proposed self-embedding fragile watermarking scheme is described and the block diagram of the system shown in Fig.1. This scheme with watermarking techniques adopted, some important feature information is extracted from the original image as a watermark and reinvested into the image itself into a watermarked image. The watermark image is encoded and transmitted over a communication channel. At the receiving end, modified data is reconstructed using a watermark along with other post-reconstruction methods if the resulting results have not shown good quality.

In general, as shown in Fig. 1, there are four main processes: pre-process the original image, watermark generation and embedding process, tamper detection, and tamper recovery. The detail of proposed technique is described in detail as follows.

A. *Pre-process the original image*

RGB color space is not recommended when dealing with computer-based analyzers, RGB is best suited for image display. This is due to the high correlation between components R, G, and B. Numerically the cross-correlation value between component R and G is 0.98, G and B components of 0.94, whereas between B and R has a value of 0, 78, so it is not appropriate when used for signal processing especially the use of watermarking techniques. Under the scheme [14], the use of color space that has lace correlation can use the YCbCr channel.

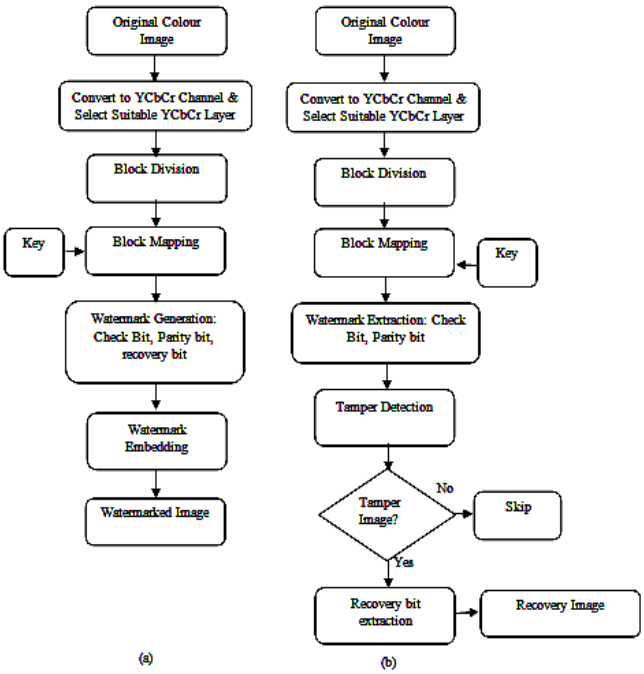

Fig. 1. Block diagram a feature-based fragile watermarking scheme of the proposed method. (a)encoder side; (b) decoder side.

Thus, the proper selection of channels for watermark insertion is required to reduce the degradation of the watermarked imagery. The first thing to do is transform the channel from RGB to YCbCr using equation (1). After that the best channel selection for the insertion process is done. In this research used Cr channel as the place of insertion. Then, by utilizing the block-based watermark method, the channel is divided into 2x2 pixels non-overlapped blocks, which can form a series $I_i\ (i = 1,2,3,\dots,N)$. N is the total number of blocks in the image.

$$Y = (77/256)R + (150/256)G + (29/256)B$$
$$Cb = -(44/256)R - (87/256)G + (131/256)B + 128$$
$$Cr = (131/256)R - (110/256)G - (21/256)B + 128 \quad (1)$$

Initial process before watermark insertion is done by using 1-D linear transformation [2] as seen in equation (2). This transformation resulted in one-to-one mapping, where the watermark for the tamper recovery process will be inserted.

$$I_i^r = [f(I_i) = (k \times I_i) mod\ N] + 1, \quad (2)$$

Where $\{(I_i, I_i^r) | i \in [1, N]\}$ is the block number in each 2x2 pixel, k is a secret key which must be a prime number.

B. *Watermark Generation and Embedding Process*

In the proposed scheme, the selected YCbCr channel of the original image I is partitioned into non-overlapping N block $I_i\ (i = 1,2,3,\dots,N)$ of size 2x2 pixels. Each block of I_i is contained four pixels, each pixel is converted to binary bit, 8 bits.

We adopted He's method [4] and Hassan's method [13] for watermark generation by utilizing the original image features, which was used in Rakhmawati [11] for

gray images. As seen in Fig. 3, we generate two watermarks in each 2x2 blocks. Four detection bits (c_1, c_2, p_1, p_2) and an eight-bit recovery $(r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7)$. The four detection bits replace the 2x2 block in LSB 3, and the recovery bits replace the corresponding block in LSB 2 and LSB 1 by the 1-D transformation. The following algorithm describes watermark generation and embedding.

First, to insert the watermark component, we must create the intensity value of three LSBs in each 8-bits pixel in the 2x2 block to zero, then calculate the average intensity of the five MSBs remaining in each pixel (AI_i) as illustrated in Fig. 2. Then through the equation (3), generate the value of c_1, c_2 of each 2x2 block.

$$(c_1, c_2) = ((AI_i)_2 \; x \; K_i) \; mod \; 2 \tag{3}$$

Where K_i is key generated randomly of size 20 x 2, it is from 5 MSB which have four pixels and need two check bit, thus the size is 20x2.

Second, to get the value of parity bits can be seen through equation (4) and (5), where $(r_7 r_6 r_5 r_4 r_3 r_2 r_1 r_0)_{(2)}$ is the binary form of the average intensity AI_i and also as a recovery bits.

$$p_1 = r_7 \oplus r_5 \oplus r_3 \oplus r_1 \oplus r_0 \tag{4}$$

$$p_2 = \begin{cases} 1, if \;\; p_1 = 0 \\ 0, if \;\; p_1 = 1 \end{cases} \tag{5}$$

After four detection bits and an eight-bit recovery with totally 12-bit watermark are obtained, replace the LSB 1, LSB 2, and LSB 3 of each 2x2 block respectively as shown in Fig. 3.

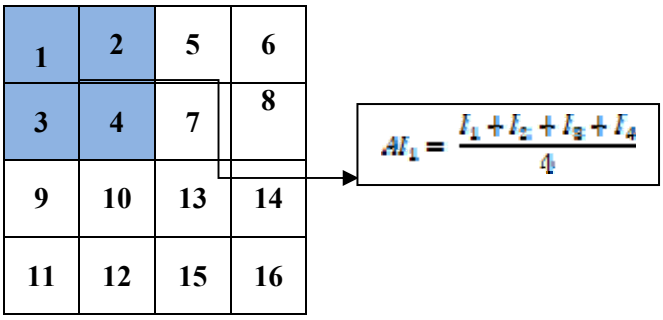

Fig. 2. Illustration watermark generation procedure for 2x2 blocks of 4x4 pixels.

Bit	8	7	6	5	4	3	2	1
Pixel 1						c_1	r_0	r_1
Pixel 2						c_2	r_2	r_3
Pixel 3						p_1	r_4	r_5
Pixel 4						p_2	r_6	r_7

Fig. 3. 12-bit *watermark* (c, p, r) embedded in pixels 1,2,3,4 of each 2x2 blocks non-overlapping

C. *Tamper Detection*

The detection process is done on the receiver side, which basically has a process flow approximately the same as the embedding process on the sender side. The detection process bilaman damage can be explained as follows. 1) the received image is first transform to YCbCr, then the same selected channel that is used in encoder is divided into 2x2 non-overlapping blocks; 2) extract c_1', c_2', p_1', p_2' in LSB 3 from each block; 3) compare c_1', c_2', p_1', p_2' with c_1, c_2, p_1,p_2 from original image block. If the four detection bits in each block are not the same, then mark the block as a modified block; otherwise, this block is marked as authentic.

D. *Tamper Recovery*

After detection process, the authentic and tampered block can be indicated. In this proposed method only restore the damaged blocks and keep the original block unchanged. For destructed blocks, use the appropriate blocks of one-to-one mapping results as described earlier to obtain the recovery information bit. The next step covers the damaged part by extracting information from the coresponding block, which can be obtained in the last 2 bits in every block.

III. RESULT AND ANALYSIS

In this section some experiments was conducted to test the extent of success of the proposed method. It was evaluated using the Ministry of Research Technology and Higher Education of the Republic of Indonesia logo as a representative of e-government host image, which has 200 x 200 pixels size. The scheme is evaluated based on the watermark embedding and recovery process.

A. *Watermark Embedding*

As mentioned earlier, the selection of watermark components is important to maintain the quality of the watermarked imagery. To evaluate the imperceptibility of the watermarked image can be done by calculating the value of Peak Signal to Noise Ratio (PSNR) which can be seen in equation (6).

$$PSNR(I, I_w) = 10 \; x \; log_{10} \frac{MAX_I^2}{MSE} \tag{6}$$

Where I is original image, I_w is the watermarked image MAX_I is the maximum possible pixel value of the original image I and Mean Square Error (MSE) refer to equation (7) . In which M and N shows the image dimensions.

$$MSE = \frac{1}{MN} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (I(i,j) - I_w(i,j))^2 \tag{7}$$

(a) (b)

Fig. 4. Test Image (a) original image; (b)watermarked image

In paper [15], it is mentioned that if the larger PSNR values show similarity between the original image and the watermarked image, which has a minimum value of 30 dB as a generally accepted perceptual value.

The PSNR value of the watermarked image will decrease when there is an increase in the number of watermark components. This can be demonstrated from the acquisition of PSNR in the proposed scheme was 36.91 dB for the watermark payload 3 bit per pixel. When compared with the method [4] which was used 2 bit per pixel, the PSNR value larger, that was 42.32 dB. The visual effect of adding the watermark component is shown in Fig. 4.

B. Tamper Detection and Recovery

At this stage, the calculation of the performance of the detection process is done by adding degradation as a form of tampered image. In this simulation, we used two type of attacks: the cropped image which are included in the general tamper types [4] and the collage attack. Fig. 5(a) illustrates an example of modification of the watermarked image by removing some of the name information on the logo, tamper detection results in Fig. 5(b), while Fig. 5(c) shows the recovery image. To know objectively, we use the PSNR calculation between the recovery image and the watermark image. The proposed method shows better results, i.e. 37.62 dB, while method [4] produces a PSNR value of 32.40 dB. It shows clearly that in our scheme the tampers can be detected and localized more accurate.

For collage attacks can be done by replacing certain part of the watermarked image with another image, in our experiment we used Chrome's logo of the same size as shown in Fig. 6(a), an example to illustrate the form of collage modification shown in Fig. 6(b), tampered detected area shown in Fig. 5(c), and the recovered image with PSNR 38.87 dB in Fig. 6(d). The results obtained can be seen in Fig. 7, where the PSNR value for cropped modification has a higher PSNR value than the collage modification. In addition, there are differences in graphs that tend to fluctuate for the type of collage attack compared with the type of cropped image. This is due to the high colour degradation difference between the Chrome's logo and the Dikti's logo, so the colour image recognition is more varied.

Furthermore, we compared the proposed scheme with the He's scheme [4] which had modified to RGB image for both types of degradation to show the performance. Fig. 8 and Fig.9 shows the results of cropping and collage attacks with a tamper region ranging from 5% to 60%. The results show that the proposed scheme gives better results in recovered image quality which experienced an average increase sharply. In addition, from the PSNR of the recovered image have a greater visual quality, with under 10% cropped attacks have more than 35 dB.

(a) (b) (c)

Fig. 5. Test cropped image of proposed method (a) tampered image; (b) tamper detection; (c) recovered image

(a) (b)

(c) (d)

Fig. 6. Test collage attackof proposed method (a) chrome'logo ; (b)tampered image (c) tamper detection; (d) recovered image

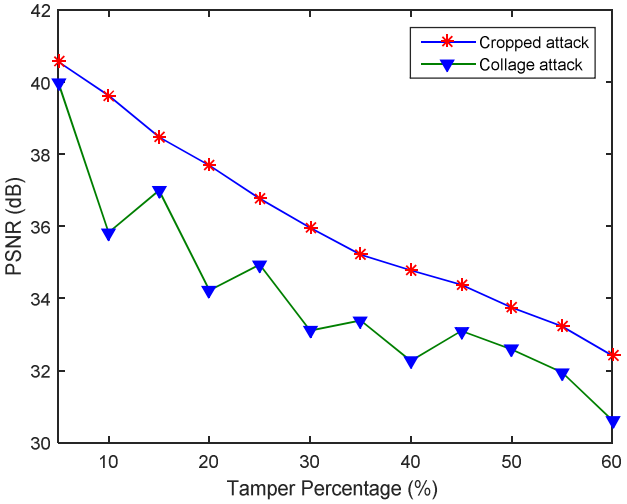

Fig. 7. PSNR of recovered image performance comparison under the collage attack and cropped attack of proposed method

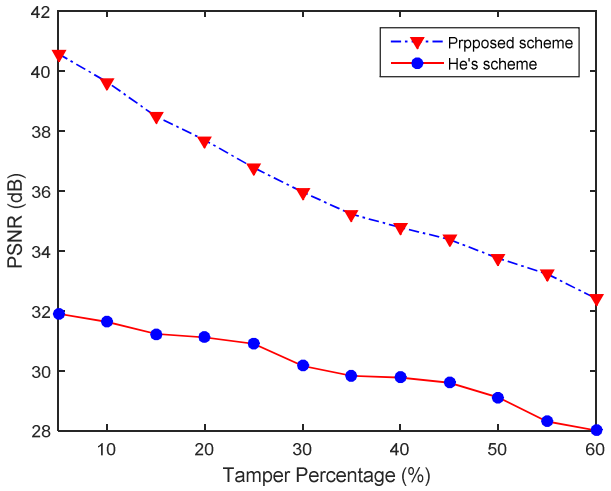

Fig. 8. PSNR of recovered image performance comparison under the cropped attack

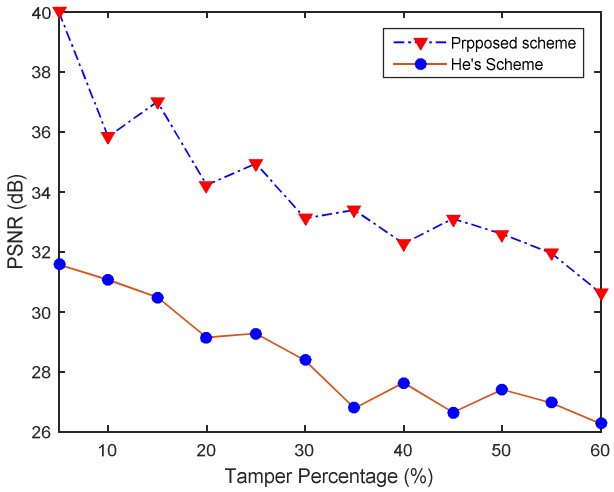

Fig. 9. PSNR of recovered image performance comparison under the collage attacks

IV. CONCLUSION

In this paper, we proposed fragile watermarking of color image for secure E-Government restoration in the spatial domain using LSB technique. We generate two watermarks that serve to detect and restore the tampered image. The detection bit is inserted in the 2x2 pixel in current image blocks, while the recovery bit is inserted in the corresponding block as a result of block mapping used 1-D transformation. In addition, the proposed scheme recover deliberate error created to remove image information by extracting the inserted watermark before the image sent.. When compared with the methods proposed previously, using 3 bit per pixel, the PSNR of watermarked image normally have around 37 dB. The results show the average value of PSNR increased significantly for both attacks: cropped and collage attacks. In addition, the results indicate that less than 10% of the tampered blocks have been recovered well.

ACKNOWLEDGMENT

This research was supported by Institut Teknologi Sepuluh Nopember (ITS) through Laboratory Research Grant (Penelitian Laboratorium) Scheme 2018.

REFERENCES

[1] Bravo-Solorio, S. & A.K. Nandi, "Secure fragile watermarking method for image authentication with improved tampering localisation and self-recovery capabilities", *Signal Processing*, Vol. 91, pp. 728–739, 2011.

[2] Phen-Lan Lin, Po-Whei Huang, An-Wei Peng. "A fragile watermarking scheme for image authentication with localization and recovery." IEEE Sixth International Symposium on multimedia software engineering Washington DC. pp. 146-153, 2004.

[3] Wu C-M, Shih Y-S. "A Simple Image Tamper Detection and Recovery Based on Fragile Watermark with One Parity Section and Two Restoration Sections". Optics and Photonics., vol. 3, no.2, pp. 103–107, 2013.

[4] He H, Chen F, Tai H, Member S, Kalker T, Zhang J. Performance Analysis of a Block-Neighborhood- Based Self-Recovery Fragile Watermarking Scheme. IEEE Transaction on Information Forensics and Security.vol 7, no.1, pp.185–196, 2012

[5] Qin, C., P. Ji, X. Zhang, J. Dong, and J. Wang, "Fragile Image Watermarking With Pixel-wise Recovery Based on Overlapping Embedding Strategy'', *Signal Processing*, Vol. 138, pp. 280–293, 2017.

[6] Chen,T-Y, M. Hwang, & J. Jan, "A secure image authentication scheme for tamper detection and recovery", *The Imaging Science Journal*, vol. 60, pp. 219–233, 2012

[7] X. Zhang, G. Han, K. Zou, W. Li, and B. Li., "An Effective Mechanism Based on Watermark for E-government Information", In Proceedings of the 2007 IEEE International Conference on Convergence Information Technology, pp. 580-585, 2007.

[8] Ali Al-Haj , Hussam Barouqa. "Copyright Protection of E-Government Document Images Using Digital Watermarking", Proceedings of IEEE 3rd International Conference on Information Management, pp. 441-446, 2017.

[9] Chen T, Lu H., "Robust Spatial LSB Watermarking of Color Images Against JPEG Compression", IEEE 5th International Conference on Advance Computational Intelligence ICACI.pp.872–875, 2012

[10] Lin PL, Hsieh CK, Huang PW. "A Hierarchical Digital Watermarking Method for Image Tamper Detection and Recovery," Pattern Recognition., Vol. 38, No.12, pp.2519–2529, 2005.

[11] Rakhmawati, Lusia, Wirawan, Suwadi, "Image Fragile Watermarking with Two Authentication Components for Tamper Detection and Recovery", In Proceedings of 2018 International Conference on Intelligent Autonomous Systems (ICoIAS'2018),pp.35-38,2018

[12] L Rakhmawati and N Rochmawati, " Review of Some Existing Work for Self-Recovery Fragile Watermarking Algorithms", IOP Conf. Ser.: Mater. Sci. Eng. 288 012093, 2018

[13] M. Hamad Hassan, and S.A.M. Gilani. "A Fragile Watermarking Scheme for Color Image Authentication" Proceeding of World Academy of Science, Engineering, and technology, Vol. 13, pp.312-316, May 2006.

[14] L. R. Roldan, M. C. Hernández, J. Chao, M. N. Miyatake and H. P. Meana."Watermarking-based Color Image Authentication With Detection And Recovery Capability", IEEE Latin America Transactions, Vol. 14, No. 2, Feb. 2016

[15] Mousavi, S.M. and A. Naghsh "Watermarking Techniques used in Medical Images : a Survey", *J Digit Imaging*, 2014.