A Chaos-based Fragile Watermarking Method in Spatial Domain for Image Authentication

Rinaldi Munir
School of Electrical Engineering and Informatics
Institut Teknologi Bandung (ITB)
Bandung, Indonesia
E-mail: rinaldi-m@stei.itb.ac.id

Abstract— **In this paper, a fragile watermarking method based on chaos map for image authentication is proposed. The watermark is a binary logo of the same size with the host image. Before embedding, the watermark is encrypted by XOR–ing it with a chaotic image. The chaotic image is generated by using Logistic Map. Next, the least significant bits (LSB) of pixels is used for embedding where the LSB's of the host image are replaced by the encrypted watermark. Authentication is done by extracting the watermark from the watermarked image and then compare it with the original watermark. If found the extracted watermark fragile then it indicates that the image has been altered. Various typical attacks to the watermarked images are carried out to assess the performance of the proposed method. Experiment results show that the proposed method can detect image integrity due to the attacks. For increasing security, the chaos system give sensitivity aspect so that the method is secure from the exhaustive attack.**

Keywords— ***fragile watermarking; image authentication, chaos; least significant bit***

I. INTRODUCTION

The rapid development of information technology make multimedia data such as the images can be easily transferred via public channels, such as Internet, wireless networks, etc. Once an image have been transferred, it can easily copied and manipulated or tampered by using image processing tools such as Photoshop. For example, somoene can change contrast/ brightness of the image, or replace an object in the image with another object, or delete some parts of the image, add a new object into the image, or combine the image with another image, etc This ease may cause problems referring to the protection of intellectual rights and integrity of information [1]. Once an image manipulated or tampered, the image is not authentic anymore. Solution to this problem is by using fragile watermarking. Fragile watermarking provides a convenient technique for authentication, tamper detection, and verification of image integrity [2]. Fragile watermarking is defined as a technique for embedding the mark (usually called waternark) into an image so that whenever the image is modified then the watermark also changed (fragile). Even though the change is slightly and no impact on the visual quality of the image, it will result a tampered or fragile watermark. The fragile watermark is the indication that the image is not original (authentic) anymore. Thus fragile watermark are commonly used for tamper detection or integrity proof of the images. The watermark embedded into the image usually represent the information that refer to image's owner which could be binary logo or some other binary data.

The fragile watermarking schemes should fulfill the some following requirements [5]: (1) Tamper detection; (2) Perceptual transparancy, (3) Detection should not require the original image; (4) Detector should be able to locate and characterize alterations made to a marked image, (5) The watermarking key spaces should be large so that brute force attack be impossible; (6) The watermarking key(s) should be difficult to deduce from the detection side information; (7) The insertion of a watermarks by unauthorized parties should be difficult.

According to working domain, the fragile watermarking techniques can be divided into two categories: spatial domain or frequency (transform) domain. In spatial domain, the watermark bits is embedded into pixel values, whereas in frequency (transform) domain the watermark bits is embedded into the transform coefficients of the host image.

In this paper, we focus on spatial domain only. In the spatial domain, the watermark bits are generally embedded into the least significant bits (LSB) of the pixels in order to maintain the image quality. The first original fragile watermarking techniques was proposed by Walton in 1995 [3]. Walton used the concept of a checksums to detect the changes in the image. The watermark bits are checksums generated from the seven most significant bits (MSB) of pixels and replace the LSB bits of the pixels by the checksums. For increasing security, Walton suggested to hide the checksums along pseudo-random order in the LSBs of pixels. The Walton scheme has a drawback, because the method only can detect the change of odd number of bits so that the attacker is easy to forge a valid watermarking after manipulation [4]. In addition, the scheme is blockwise technique that only can not detect pixel-level tampering [6].

In this paper, a fragile watermarking algorithm based on chaotic map in spatial domain is proposed. Unlike the Walton's scheme, the watermark is a binary logo of the same size with the host image. Why using chaos? Because a chaos system is sensitive to small change in the initial conditions. A sensitivity to initial conditions is required on security because it is relevant to principle of diffusion from Shannon.

In this proposed method, a chaos system is used for increasing security. Before embedding, the watermark is encrypted by XOR–ing it with a chaotic image. The chaotic image is obtained by using logistic map. The least significant bits (LSB) of pixels is used for watermark embedding. Because of watermark embedding on every pixels of the image, then the method can detect pixel-level tampering.

II. CHAOS MAP

In recent years, chaos have been used for digital watermarking to increase security [7]. The most characteristic of chaotic systems is a sensitivity to initial conditions. This sensitivity means that the small changes to the initial conditions, after the function is iterated a number of times, will produce the function values that differ significantly. As a result of this sensitivity, the behavior of chaotic systems appears to be random.

One of the simplest chaotic maps is a Logistic Map, described by

$$x_{k+1} = \mu x_k (1 - x_k) \quad (1)$$

where $0 < \mu \leq 4$. The bifurcation diagram of a logistic map is visualized in Fig. 1. The map is in chaotic state when $3.57 < \mu \leq 4$ [8]. The chaotic state is displayed by shaded area. In this state, the resulting values appear random, even though the system deterministic [9].

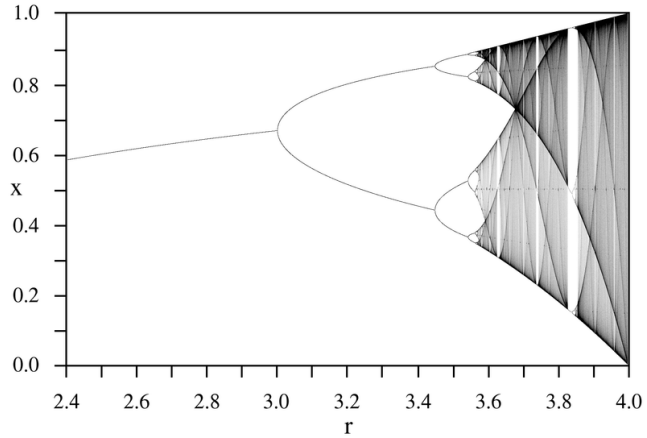


Fig. 1. Bifurcation diagram of a logistic map (image source: Wikipedia)

Because of its random behavior, a chaotic system can be used as a pseudo-random generator. Hence, initial value of Logistic Map, x_0, and constant μ serve as secret keys. When we iterate equation (1) from an initial value (x_0), we get a random sequences between 0 and 1. The random values generated from Logistic Map are sensitive to small changes in the initial values. By changing x_0 slightly becomes $x_0 + \Delta$ (for example $\Delta = 10^{-9}$), the random values generated after iterated several times are significantly different from the previous chaotic values with initial value x_0.

III. THE PROPOSED METHOD

This section explains the proposed fragile watermarking method. The proposed method embed and extract watermark in spatial domain. Withous loss of generalization for color images, let H is the host grayscale image of size $M \times N$ and W is the watermark which is the binary image (logo) of size $m \times n$. In general, $m \leq M$ and $n \leq N$, so that we must arrange periodically the watermark W in order to result binary watermark W has the same size with host image H. The watermark embedding and extraction algorithm is described in each sub-section below.

A. Watermark Embedding Algorithm

The watermark embedding algorithm is follows (The block diagram of watermark embedding is shown in Fig. 2): Generate a chaotic image C of size M × N using Logistic Map with initial value x_0 and constant μ (serve as watermarking keys).

1. Obtain encrypted waternark W_e using XOR operation between W and C as follows:

$$W_e = W \oplus C \quad (2)$$

2. Extract the least significant bit (LSB) of pixels of host image H into a plane image. Let the LSB plane is H_{LSB}.
3. Replace H_{LSB} by W_e as follows:

$$H_{\text{LSB}}' \leftarrow W_e$$

4. Place H_{LSB}' into the host image to get watermarked image H_w.

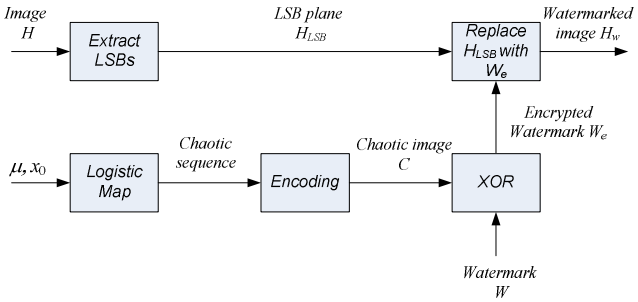


Fig. 2. Block diagram of watermark embedding

B. Watermark Extraction Algorithm

The watermark extraction algorithm is follows (the block diagram of watermark extraction is shown Fig.3):

1. Generate a chaotic image C of size $M \times N$ using Logistic Map with initial value x_0 and constant μ.
2. Extract the least significant (*LSB*) bit of pixels of watermarked image H_w into a plane image. The LSB plane is H_{wLSB}.
3. Apply XOR operation between H_{wLSB} and C to get extracted watermark W_{ext} as follows:

$$W_{ext} = H_{\text{wLSB}} \oplus C \tag{3}$$

4. Compare original watermark W with W_{ext} to decide authentication of the image. If W = W_{ext} then the image is authentic, otherwise the image is not authentic anymore.

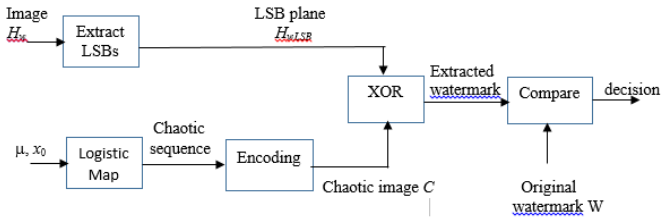


Fig.3. Block diagram of watermark extraction.

C. *Chaotic Image Encoding*

Because of Logistic Map yields real numbers between 0 and 1, we need to convert them to integers and extract binary bits to get a chaotic image. In this proposed method, conversion from real to integer is based on a truncation function [10]. Suppose x_i is a real number between 0 and 1 that generated by a logistic map. Conversion x_i to integer is obtained as follows: x_i multiplied by 10 repeatedly until it reach a desired long number (size), and then truncate to take the integer part. Mathematically, this process is described by function T as follows:

$$T(x, size) = \left\| x * 10^{count} \right\|, x \neq 0 \tag{4}$$

where *count* is begined from 1 until $x * 10^{count} > 10^{size-1}$ and symbol || || represents truncation. The least significant bit of binary representation of the integer is then extracted to get a chaotic image C.

IV. Experiment Results

The proposed method is programmed using software MATLAB and some experiments to embed watermarks into the host images are performed. After embedding, the watermarks are extracted from the watermarked images, and the extracted watermarks are compared with the original watermarks to make conclusion if the watermarked image is aletred. The host image are two grayscale images of size 512 × 512 ('ship' and 'bird'). The watermarks are two binary logo ('ganeca' and 'simplemark') which after arranged periodically in order to result the same size with the host images. The secret keys are parameters of Logistic Map which are chosen as $x_0 = 0.675$ and $\mu = 3.9762$.

Fig. 4 shows respectively the original images, the binary watermarks, and the watermarked images for the specified parameters. Visually no difference between the host image and the watermarked image. The watermarked images have PSNR's respectively are 51.1451 dB and 51.1446 dB. Experiments to extract watermarks from the watermarked images result the original watermarks exactly.

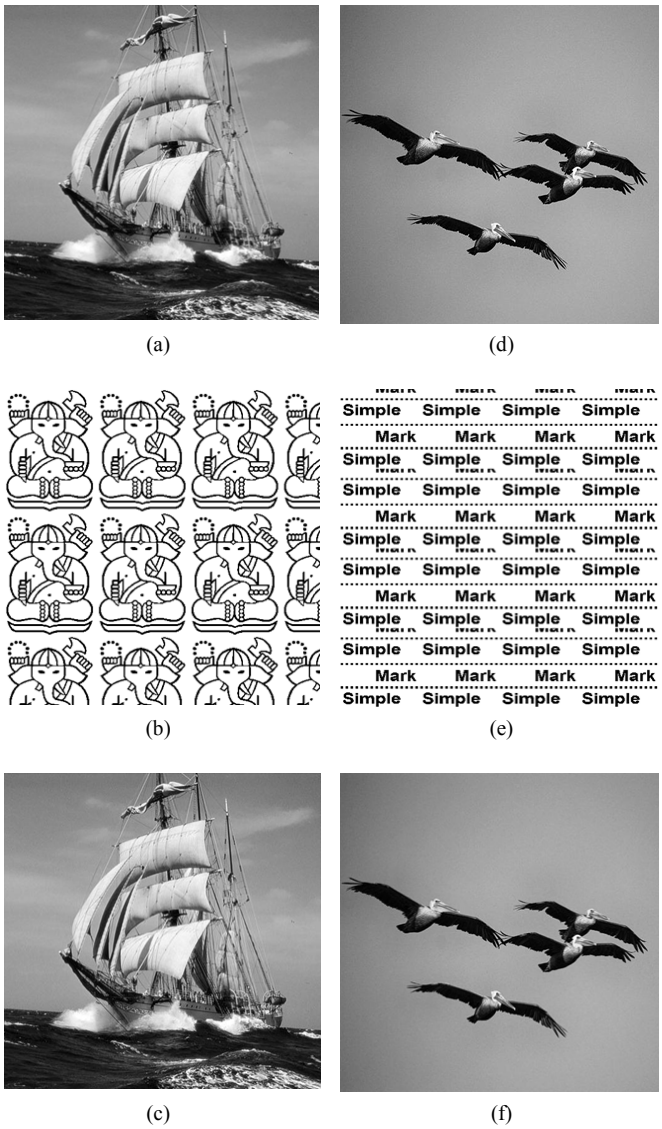


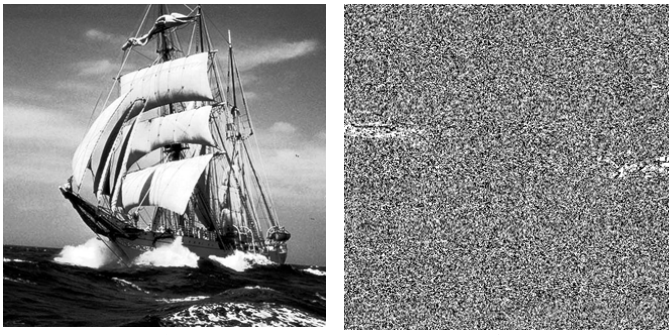
Fig. 4. (a) and (b) are original images; (b) and (e) are binary watermarks; (c) and (f) are watermarked images.

In section below, various attacks to the watermarked images are carried out to assess the performance of the proposed method. The attacks is performed using an image processing tool such as Photoshop. The typical attacks are

histogram equalization, text addition, image flipping, copy-paste attack in the same image, and copy-paste attack into another watermarked image.

A. Performance under histogram equalization

In this experiment, we manipulated the watermarked image by adjusting intensity values by performing histogram equalization operation. Fig 5. shows respectively the watermarked image, the watermarked image after histogram equalization, and the extracted watermark. Majority of intensity values of pixels has changed after the operation. As a result, the extracted watermark has tampered and looks like as a random image, so that we conclude that the watermarked image has changed (not authentic anymore).

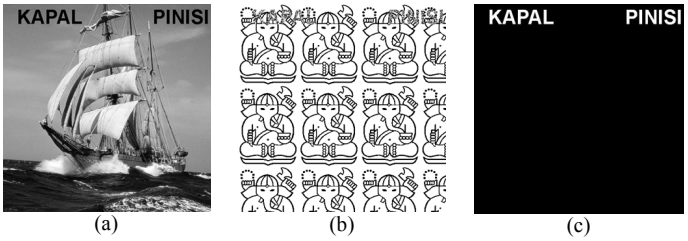


(a) (b)

Fig. 5. (a) watermarked image after histogram equalization, (b) extracted watermark.

B. Performance under text addition

In this experiment, the watermarked image is modified by adding a text 'KAPAL PINISI' at te top of the image (Fig. 5(a)). Only the region with the text changed, and as a result the extracted watermark contains image of the text (Fig. 6(b)). Because of the extracted watermark is not same with the original watermark, we conclude that the watermarked image has been altered. When we substract the extracted watermark with the original watermark, and then remove the parts that no contain objects in the substraction result, we obtain tampered region (Fig. 6(c)).

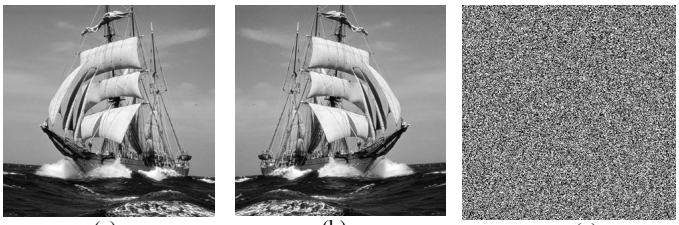


(a) (b) (c)

Fig. 6. (a) Watermarked imaged after text addition, (b) extracted watermarked; (c) detected tampered region.

C. Performance under image flipping

In this experiment, the watermarked image is flipped horizontally and then the watermark is extracted. As a result, reference of embedded watermark is lost, and the extracted watermark looks like as a random image. Based on the experiment result, we conclude that the watermarked image has been altered. Fig. 7 shows the experiment results of image flipping to 'ship' image.

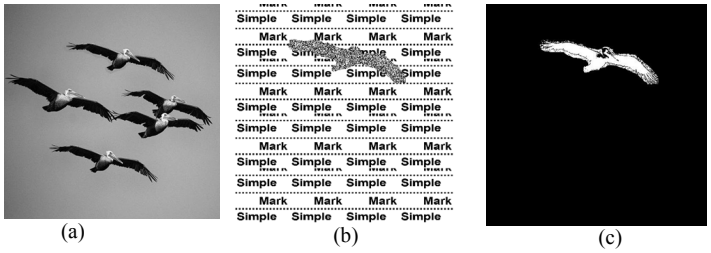


(a) (b) (c)

Fig. 7. (a) Watermarked imaged, (b) watermarked image after horizontal flipping; (c) extracted waternarked.

D. Performance under copy-paste attack

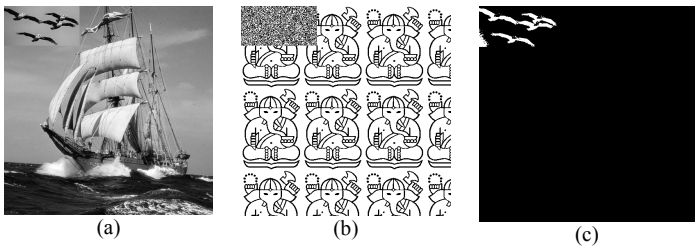
In this experiment, a part of object in the watermarked image is copied and pasted into the image itself. Fig. 8 shows a lowest bird is copied and pasted on top of image. Visually, the extracted watermark contains the new object, and we conlcude that the watermarked image is not authentic anymore (Fig 6(b)).



(a) (b) (c)

Fig. 8. (a) Watermarked image after adding one bird, (b) extracted watermark, (c) detected tampered region.

Meanwhile, in Fig. 9 we copy all of birds and paste them into another watermarked image (in this case 'ship' image) on left top of image. The extracted watermark contains the pasted image, so that we conclude that the 'ship' image is not authentic.

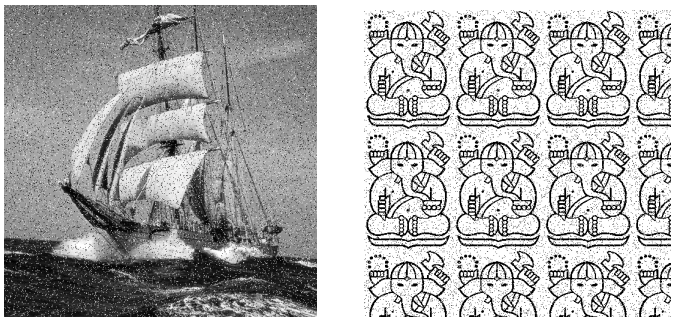


(a) (b) (c)

Fig. 9. (a) Watermarked image after joining an image into another image, (b) extracted watermark, (c) detected tampered region.

E. Performance under image noising

In this experiments, salt and pepper noise is added to the watermarked image with density 0.1. After that, the watermark is extracted from the image. Fig. 10 shows the noisy image and the extracted watermark. The watermark is fragile so we conclude the image has been tampered.



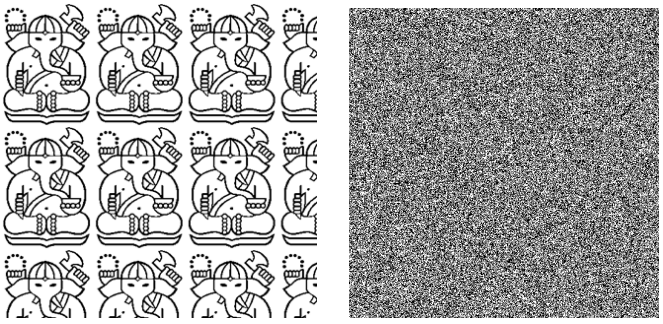
(a) (b)

Fig. 10. (a) Watermarked image after noising, (b) extracted watermark

F. Performance under Sensitivity attack

One of important characteristics of chaos is the sensitivity to small changes in initial values. Small change (Δ) to initial value (x_0) yields random values that significantly different after iterating the Logistic Map a number of times. As a result, the chaotic map C is significantly different.

Let $\Delta = 10^{-10}$ so that the initial value of a logistic map $x_0 = 0.67500000001$. Use this x_0 as a key in the watermark extraction. Figure 11 shows the extracted watermark. The extracted watermark is a random image. This experiment shows that sensitivity characteristics of chaos provide good security from exhaustive attack.



(a) (b)

Fig. 11. (a) original watermark; (b) extracted watermark due to small change x_0

V. CONCLUSION

In this paper, the fragile watermarking method based on chaos has been presented. The method can detect pixel-level tampering, because embedding of watermark is performed on every pixels of the image. Experiment results show that the method can detect image authentication due to the various typical attacks.

REFERENCES

[1] Caragata, D., Radu, A. L., Assad, S., Fragile Watermarking using Chaotic Sequences, *International Journal for Information Security Research* (IIJSR), Vol. 1, Issues ½, Marc/June 2011.

[2] Fridrich, J., Goljan, M., Baldoza, A., New Fragile Watermarking for Images.

[3] Walton, S., "Image Authentication for a Slippery New Age," *Dr. Dobb's Journal*, vol. 20, no. 4, pp. 18–26, 1995.

[4] Liu, S., Yao, H., Gao, W., Liu, Y., An image fragile watermark scheme based on chaotic image pattern and pixel-pairs, *Applied Mathematics and Computation 185* (2007) 869 – 882.

[5] Jain, P., Rajawat, A., Fragile Watermarking for Image Authentication Survey, International Journal of Electronics and Computer Science Engineering (JSECSE).

[6] Suthaharan, S., Logistic Map-Based Fragile Watermarking for Pixel Level Tamper Detection and Resistance, EUROSIP Journal on Information Security, Vol. 2010.

[7] Dawei, Z., Guanrong, C., Wenbo, L., A Chaos-Based Robust Wavelet-Dmain Watermarking Algorithm, *Chaos Solitons and Fractals 22* (2004) 47-54.

[8] Bose, R., Banerjee, A., Implementing Symmetric Cryptography Using Chaos Function, Indian Institute of Technology.

[9] Robinson, R.C., An Introduction to Dynamical Systems, Continuous and Discrete, Pearson Prentice Hall, 2004.

[10] Lampton, J., Chaos Cryptography: Protecting Data Using Chaos, Mississippi School for Mathematics and Science