



Research on Image Content Authentication Algorithm of Semi-fragile Watermarking for Adaptive Embedding Based on Characteristic of Image

Shan Yang, Liang Chen, and De Li^(✉)

Department of Computer Science, Yanbian University, Yanji, China
379606352@qq.com, 371482788@qq.com,
leader1223@ybu.edu.cn

Abstract. In this paper, we propose a novel quantization-based semi-fragile watermarking algorithm for image authentication and tamper localization based on discrete wavelet transform (DWT). In this algorithm, the watermarking is generated by extracting image feature from the approximation subband in the wavelet domain. The linear congruential pseudo random number generator is used to control the watermark embedding location and realize adaptive embedding. The generation and embedding of watermark are carried out in the original image. The authentication are not require original image and any additional information, which improves the security and confidentiality of watermark. The scheme can resist the mild modification of digital image and able to detect the malicious modifications precisely. Experimental results show that the proposed algorithm can distinguish high quality JPEG compression, small amount of noise from shear attack, classification of intentional and incidental tampering, and has good localization ability.

Keywords: Semi-fragile watermarking · Content authentication
Discrete wavelet transform (DWT) · Adaptive embedding · Tamper location

1 Introduction

In the era of rapid development of information technology, multimedia data is convenient for storage, transmission and sharing, and enriches people's means of obtaining information. But at the same time, people can easily replicate and modify digital image, and multimedia data has a large amount of data, high redundancy, and the difficulty of tracking changes sources also make digital media security issues into a major problem in the informationization process, this will seriously hindered the development of information industry. In some applications, such as military documents, forensic evidence, news images or video, history literature, medical images, ticket anti-counterfeiting etc., the illegal tampering of digital content will have a serious impact on people or society [1]. So how to verify the authenticity, integrity and credibility of digital multimedia content in the complex network environment becomes an urgent problem.

Digital watermarking technology, as an effective means of protecting intellectual property rights of digital media works, has received extensive attention from the society. Among them, semi-fragile watermarking technology can not only tolerate to some extent some common signal processing operations (such as adding noise, smooth filtering, lossy compression etc.), which contained on watermarking digital media, but also can make alarm reaction to malicious tampering with and have the ability to locate tamper with the region, so it is more important application value in the Internet age [2]. In order to improve the safety and reduce the computational complexity of the algorithm, Hu [3] proposed a fragile watermarking algorithm based on wavelet transform, the main innovation point of the algorithm lies in ascension format parametric integer wavelet coefficients. but this plan in resistance to JPEG compression cannot achieve a satisfactory effect. Zhao et al. [4] proposes a semi semi-fragile zero-watermark algorithm based on block compression perception, that the image is divided into several blocks. Then each image block is observed according to the compressed sensing theory, and the observed eigenvalues are preserved as semi-fragile zero-watermark information. Liu [5] proposed a semi-fragile watermarking algorithm based on Zernike moment. First, the original image was transformed by discrete wavelet transform, then the Zernike moment was taken as the image feature. Finally Zernike moment of low-frequency wavelet coefficient was taken as the watermark, which embedded into the DWT coefficient of the image.

This paper puts forward a semi-fragile watermarking algorithm based on wavelet domain, which can pass the authentication of common signal processing operations, and refuse to pass the authentication of malicious tampering and has strong localization ability. In the algorithm implementation process, first of all multilayer discrete wavelet transform is applied to the carrier image, and then the watermark is generated according to the image feature information of the carrier image. Embedded position determined randomly by random linear congruence generator, then inserts and updates the wavelet coefficients. When the watermark is extracted, the carrier image embedded with watermark information is firstly transformed by discrete wavelet transform. Then generate watermark information, and locate tamper areas. So the dependence on original information reduces and improves the security of watermark information. The experimental results show that this algorithm can realize the image content authentication function, and the location of malicious tamper. At the same time, it can withstand general non-malicious attacks such as JPEG compression and a small amount of noise.

2 Research on Related Technology

2.1 Image Authentication Technology Based on Digital Watermarking

Digital watermarks can be divided into fragile watermarks and semi-fragile watermarks, both of which are used to achieve image integrity authentication. Among them, the fragile watermark technology is very sensitive to any slight change of the detection object, while the semi-fragile watermarking technology allows general processing operations such as JPEG compression of images under the guarantee of the authenticity of image content. Both of these watermarking techniques can detect and locate tamper areas well.

2.2 Linear Congruent Generator

In order to ensure a good visual effect of carrier images and the invisibility of watermark information. For the selection of embedded locations, random embedding with the same probability is carried out in the intermediate frequency components of wavelet coefficients. In the random position determination, it is determined by the linear congruent pseudo random number generator.

Linear congruent pseudo-random number generator [6] is a uniformly distributed pseudo-random number generator. Its mathematical formula is:

$$x_{i+1} = (ax_i + c)(\text{mod}m) \quad (1)$$

There are three parameter multiplier a , increment c , and module m . Where m is a prime number, a is a prime root of m , $c \in [0, m)$. Suppose there is a minimum positive integer T , So that for any I there is a $x_{i+T} = x_i$. T is the period of the pseudo-random number. The values of the non-negative integers of T modules m are different in a period, So maximum period $T_{\max} = m$. In theory, the bigger T is, the better the pseudo-random number is. In this paper, the value of m is selected as 2^{31-1} . When $c = 0$, the most general application of linear congruent pseudo random number can be obtained. It's also called the multiplication congruence generator.

The mathematical formula is as follows:

$$x_{i+1} = ax_i(\text{mod}m) \quad (2)$$

In this chapter, a pseudo-random number subject to uniform distribution is generated by the generator. Then the binary pseudo-random number is obtained, so that the embedding domain of the watermark is selected.

3 Semi-fragile Watermarking Based on Feature Recognition

Digital watermarking is an important way to ensure multimedia information security technology. It is mainly to verify the authenticity and integrity of digital multimedia works. Precision certification is sensitive to any minor modification and is suitable for using in situations where extreme precision is required [7]. However, in practical application, multimedia data may be changed by various attacks in network transmission, but small changes in data bits cause little visual change and make any significant difference. In the era of information sharing, network transmission may inevitably suffer from signal attacks such as JPEG compression. This attack will change a lot of data in the image, but will not affect the visual perception of the human eye. However, such processing operations will be rejected for accurate certification. In order to protect the integrity of information content and the inevitable routine signal processing operations, so the semi-fragile watermark has the two functions of fragile watermark and robust watermark.

3.1 Watermark Information Generation Algorithm

The advantage of wavelet transform is to provide rich feature information of image, and it has a great advantage in watermark information extraction. The coefficient component of the third layer wavelet decomposition $LL3(i, j)$, contains most of the energy of the image and can effectively characterize the image features. It has certain robustness before and after JPEG compression and has strong anti-interference ability. Therefore, it is an ideal choice to use it to extract image feature watermark.

The flow chart of watermark certification is shown in Fig. 1:

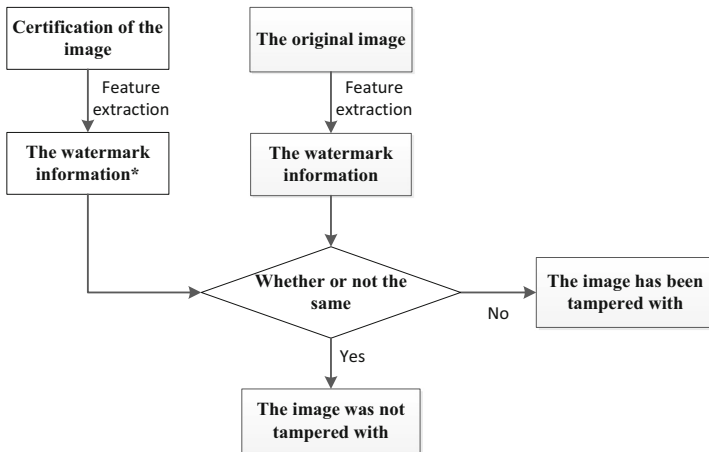


Fig. 1. Watermark certified flow chart

The watermark information generation algorithm is as follows:

Step 1: The carrier image is transformed by 3 - layer discrete wavelet transform.

Hypothesis: the carrier image I is a grayscale image of $M \times N$ size, and the 3 - layer wavelet transform is performed on I . The size of the wavelet coefficient Z is $M/8 \times N/8$. Where, $LL3(i, j)$ is the low frequency coefficient of three-layer wavelet decomposition.

Step 2: The calculation of watermark information.

Extract the binary highest bit of the coefficient component $LL3(i, j)$ as image binary watermark information, and generate W for embedded watermark of $M \times N$ size.

3.2 Watermark Information Embedding Algorithm

The process of watermark extraction is the inverse process of constructing zero-watermark. The algorithm is described as follows:

Step 1: Block processing of wavelet embedded domain.

The vertical components $HL2$ and horizontal components $LH2$ of the image are obtained by two-layer wavelet transform of the carrier image.

The image should be embedded with the watermark W with the size of $M/8 \times N/8$, and then segmented by HL2 and LH2. The size of each chunk is 2×2 , so the two components of the wavelet coefficients are divided into $M/8 \times N/8$ pieces. To embed the size of $M/8 \times N/8$ for the watermark information W on and block processing, so the two components of the wavelet coefficients are divided into $M/8 \times N/8$ pieces. The position of each piece in the wavelet coefficient can be denoted as $B(i, j)$, among them: ($1 \leq i \leq 8M, 1 \leq j \leq 8N$). We call it HL2 for $B_0(i, j)$, and LH2 for $B_1(i, j)$.

Step 2: Selection of wavelet embedded domain.

The linear congruence generator described in 1.2 is used to generate pseudo-random numbers that obey uniform distribution. After the initial value is determined, the 0,1 matrix choice(i, j) of the binary value is generated.

Thus, the embedding domain determination principle is: at the time when choice(i, j) = 0, the corresponding watermark information is embedded into block $B_0(i, j)$; at the time when choice(i, j) = 1, The corresponding watermark information is embedded into block $B_1(i, j)$. In this way, the distribution of watermark information is more uniform and the invisibility of watermark is improved. When the initial value is uncertain, it is difficult to crack the watermark information even if you know the rules of watermark embedding. The security of watermark is guaranteed.

Step 3: Determine the quantification degree difference of watermark information embedding.

Each block $B_0(i, j)$ or $B_1(i, j)$ is embedded with a bit of watermark information. In order to ensure the invisibility of watermark embedding, the size of embedded watermark information can be adaptive embedding according to the size of wavelet coefficient. The definition parameter weight(i, j) represents the maximum degree of quantization that each block can withstand weight(i, j) is defined as:

$$weight(i, j) = mean(i, j) + 3 \times std(i, j) + 10 \times entropy(i, j) \quad (3)$$

Where mean(i, j) represents the mean of $B(i, j)$, std(i, j) represents the standard deviation of $B(i, j)$, and entropy(i, j) represents the entropy of $B(i, j)$.

Step 4: Determine the quantitative step size of watermark information embedding.

Do the same thing for all the chunks, Find the matrix weight₀(i, j) and weight₁(i, j) that all the blocks of HL2 and LH2 can bear the maximum quantization degree. Determine the quantized step size of block $B(i, j)$, The definition formula is:

$$\Delta(i, j) = c + (d - c) \times \frac{weight(i, j) - w_{min}}{w_{max} - w_{min}} \quad (4)$$

Among them, w_{min} represents the minimum value of weight(i, j) matrix and w_{max} represents the maximum value of weight(i, j) matrix. C and d are the range of the selected quantitative step length, and c and d are selected according to the experiment.

Step 5: Watermark information embedding.

For block $B(i, j)$ to be embedded in the watermark, the adaptive quantization step is $\Delta(i, j)$, and x_i represents the wavelet coefficient within the block. $W(i, j)$ represents the

watermark information to be embedded, and the watermark embedding process is the process of quantifying the absolute mean of the wavelet coefficient block. The absolute mean of the blocks $B(i, j)$ is defined as:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n |x_i| \tag{5}$$

Define the formula: $(\bar{x} + \Delta(i, j)/2)/\Delta(i, j) = Q + r$. Among them, Q is the quotient and r is the remainder. Embed the watermark by modifying absolute mean \bar{x} , and set the modified absolute mean to \bar{x}' .

Watermark embedding rules are:

$$\begin{cases} \bar{x}' = \Delta(i, j) \times Q + \Delta(i, j)/2, Q \text{ and } W \text{ have the same parity} \\ \bar{x}' = \Delta(i, j) \times Q - \Delta(i, j)/2, Q \text{ and } W \text{ have different parity} \end{cases} \tag{6}$$

The embedding of watermark information $W(i, j)$ can be completed by updating the wavelet coefficients with the new absolute mean, so define the amount of changing in absolute mean as $g = \bar{x}' - \bar{x}$. The updating method of wavelet coefficient is:

$$\bar{x}' = x_i \times \left(1 + \frac{g}{\bar{x}}\right). \tag{7}$$

\bar{x}' represents the updated coefficient, and the image with watermark is obtained through wavelet inverse transform.

3.3 Watermark Information Extraction Algorithm

When the watermark is extracted, two-layer wavelet transform is carried out for the watermark image I_w . The wavelet components HL2 and LH2 of the I_w can be obtained. Using the same algorithm as 3.2, the vertical detail sub-band and horizontal detail sub-band are segmented into blocks with a size of 2×2 . The maximum watermark embedding degree weight $\Delta'(i, j)$ and the quantization step length $\Delta'(i, j)$ of each block are obtained. Random selection matrix choice (i, j) is generated by using the same initial value conditions as the embedded algorithm. Define:

$$\bar{x}_w / \Delta'(i, j) = \tilde{Q} + r' \tag{8}$$

Among them, \tilde{Q} is the quotient and r' is the remainder. When \tilde{Q} is even, $W_m^*(i, j) = 0$, When \tilde{Q} is odd, $W_m^*(i, j) = 1$. So get the extracted watermark W_m^* .

3.4 Image Authentication and Tamper Location

The watermark information W_m^* is generated by using image I_w containing watermark. After obtaining the two extracted watermark information W_m^* and W_0^* , because

watermark is the authentication information of image, the tamper part of image is judged by comparing the difference between W_m^* and W_0^* . Tamper positioning matrix is defined as follows:

$$T = |W_0^* - W_m^*| \quad (9)$$

If $T = 0$, the image is not tampered with; if $T \neq 0$, then the point in the T matrix where is not 1, is considered to be the tamper area.

4 The Experimental Results and Analysis

4.1 Experimental Analysis of Semi-fragile Watermark Based on Image Authentication

The algorithm proposed in this paper, which extracts the features of the original image and inserts it into itself as watermark information to generate watermark images. The watermark information generated by the image under test is compared with the extracted watermark information.

- (1) The algorithm uses PSNR (Peak Signal Noise Ratio) to accurately describe the change of image quality of the carrier before and after watermark embedding. Its definition is as follows:

$$PSNR = 10 \lg \frac{255^2}{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [f(m, n) - g(m, n)]^2} \quad (10)$$

Normalize the NC value of the correlation function: NC value is often used to evaluate the robustness of hidden algorithms. Its definition is as follows:

$$NC(f, g) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (f(m, n) \times g(m, n))}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(m, n)^2} \quad (11)$$

As can be seen from the above formula, the closer the NC value is to 1, the stronger the anti-attack capability of the algorithm is.

Among them: $f(m, n)$ is the pixel value of the original image; $g(m, n)$ is the modified pixel value. In general, if the value of PSNR is larger, the difference between watermark image and carrier image is smaller, so the algorithm has high transparency of watermark.

The 256 grayscale Lena images with the original carrier image size of 256×256 were tested. Without any attack, Fig. 2(a) is the original Lena image and Fig. 2(b) is the generated watermark image.

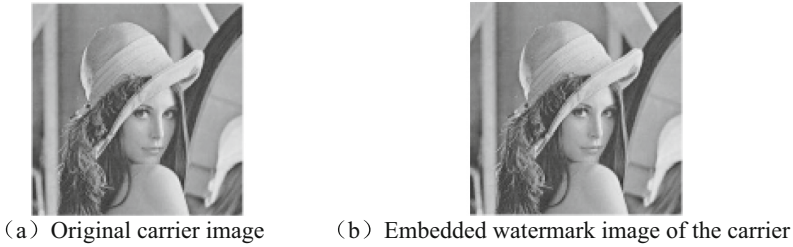


Fig. 2. Host image and watermarked image

It can be concluded from the experiment that the two images are hard to be distinguished by the eye, indicating that the watermark information in the algorithm is highly invisible. By calculation, the PSNR of Lena graph is 50.55 dB. The comparison results of PSNR with other algorithms are shown in Table 1.

Table 1. Comparison of watermarked image’s PSNR values of different algorithms

Algorithm	Image’s PSNR values (dB)
This algorithm	50.55
The literature [8]	40.28
The literature [9]	41.04

- (2) Judging whether the matrix is tampered according to the number of elements not equal to 1 in tamper matrix T [10]. Define BER as bit error rate
- Compression attacks: JPEG compression attacks with different compression factors for authentication images are treated separately.

Specific results are shown in Table 2.

Table 2. The change rate of image under different JPEG compression factors

JPEG compression factor	BER	PSNR	NC
30	23.8%	43.48	0.9992
40	21.7%	44.84	0.9959
50	4.2%	52.08	0.9980
60	2.6%	53.88	0.9984
70	2.3%	54.08	0.9985
80	2.3%	54.08	0.9985

- (3) Noise attack: add different types of noise to the image, and the results are shown in the figure below (Table 3):


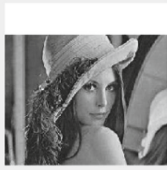
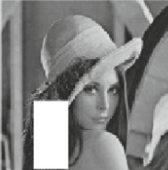





Table 3. The change rate of image under different JPEG noise attacks

Shear type	STREBGT	BER	PSNR	NC
Salt and pepper noise	0.005	13.5%	38.8018	0.9987
Salt and pepper noise	0.01	19.8%	36.3102	0.9976
Gaussian noise	0.005	15.4%	29.4708	0.9997
Gaussian noise	0.01	23.6%	28.7037	0.9984

It can be concluded from the above experiments (2) (3) that JPEG compressed and slightly noisy images can be authenticated, $BER \leq 25\%$, so this algorithm allows images to be processed in a conventional way.

- (4) Shear attack: it is a very common malicious tamper method, which randomly conducts shear experiments and tamper localization on Lena.

Table 4. Cropping experiment of Lena

Shear mode	1	2	3	4
The cut image				
Locate the image				
BER	11.36%	19.33%	8.66%	38.54%

As can be seen from Table 4, this algorithm can well indicate the tamper position of the image.

4.2 Algorithm Comparison and Experimental Results Analysis

We compared the experimental data of the algorithm in this paper with the experimental data in literature [11]. The comparison results are shown in Table 5:

Table 5. Watermark robustness test results contrast

The experimental data	The literature [11]		This algorithm	
	PSNR	NC	PSNR	NC
Attack types				
The evaluation index				
JPEG compression 30%	34.10	0.9379	43.48	0.9992
Salt and pepper noise 0.02	21.77	0.9824	29.34	0.9976
Gaussian noise 0.01	19.69	0.799	28.70	0.9984

As can be seen from the comparison results in Table 5, the algorithm in this paper is characterized by high robustness and high invisibility of the extracted watermark, strong resistance against various kinds of attacks and good image quality. The overall performance of the algorithm is superior to that of literature [11].

The experiment is carried out in the wavelet domain, which is combined with the compression standard of image, and has strong practicability. The image is transformed from the spatial domain to the transform domain, which can enhance the transparency of the watermark. This has good robustness to noise, filtering and JPEG compression.

The tamper location can be well realized in the experiment, but there are some errors in the experimental results. The extraction and generation of the printing information are provided by the image to be authenticated. It does not need to provide the original image and any additional information, which improves the security and practicability of the watermark.

5 Conclusions and Further Work

This paper presents a content authentication algorithm based on semi-fragile watermark. Make full use of the feature of multilayer multi-resolution of wavelet and combine image features to generate watermark information. The watermark information was embedded into the wavelet domain coefficients of the original carrier image. In the process of image verification, there is no need to provide the original carrier image. This allows for true blind extraction. The experimental results show that the algorithm is vulnerable to malicious attacks.

There are still many problems in this algorithm, that is, the image is not attacked and the extracted watermark and embedded watermark are not guaranteed to be the same. So on that basis, the further work is: firstly improve the algorithm of extracting and embedding watermarks to make them exactly the same; Secondly, it can reduce the error detection rate and improve the robustness.

Acknowledgments. 1. This research project was supported by the Education Department of Jilin Province (“Thirteen Five” Scientific Planning Project, Grant No. 249 [2016]).

2. This research project was supported by the National Natural Science Foundation of China (Grant No. 61262090).

References

1. Huang, S.: Image authentication algorithm based on semi fragile watermarking. Hunan University (2008)
2. Wang, X.Y., Yang, H.Y., Chen, L.K., Zhao, H.: Semi-fragile watermark embedding algorithm for image content authentication. *Microcomput. Syst.* **11**, 147–150 (2005)
3. Zhang, J.S., Yang, M., Zhou, L.X.: Semi-fragile watermark authentication algorithm for image content based on DWT. *Sci. Technol. Bull.* **33**(6), 192–195 (2017)
4. Zhao, C.H., Liu, W.: Image semi-fragile zero-watermark algorithm based on block compression perception. *J. Autom.* **38**(4), 609–617 (2012)
5. Liu, H., Lin, J., Huang, J.: Image authentication using content based watermark. In: IEEE International Symposium on Circuits and Systems, ISCAS 2005 (2005)
6. Wang, L.N., Guo, C., Li, P.: Experimental Course on Information Hiding Technology, pp. 22–23. Wuhan University Press, Hubei (2004)
7. Wang, B.B.: Research on digital watermarking algorithm of document image for content authentication. Shandong Normal University (2010)
8. Ma, X.M., Li, Y.S., Xie, J.L.: Experiment and analysis of point cloud de-noising using bilateral filtering method. Survey report, 89–115 (2017)
9. Qi, X.J., Xin, X.: A quantization-based semi-fragile watermarking scheme for image content authentication. *J. Vis. Commun. Image* **22**(2), 187–200 (2011)
10. Zhang, B.: Research on content authentication technology based on perception hash and digital watermark image. Beijing University of Posts and Telecommunications (2011)
11. Li, Z.W.: Research on dual digital watermarking algorithm based on wavelet transform domain. Anhui University of Technology (2016)