mathematics

MDPI

Article

A Novel Image Tamper Detection and Self-Recovery Algorithm Based on Watermarking and Chaotic System

Yewen Li [1,2,3], Wei Song [1,2,*], Xiaobing Zhao [1,2], Juan Wang [1] and Lizhi Zhao [1]

[1] School of Information Engineering, Minzu University of China, Beijing 100081, China; liyewen@ncic.ac.cn (Y.L.); zxb_cn@163.com (X.Z.); 16049032@muc.edu.cn (J.W.); lizhi3285@126.com (L.Z.)
[2] National Language Resource Monitoring and Research Center of Minority Languages, Minzu University of China, Beijing 100081, China
[3] Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100081, China
* Correspondence: songwei@muc.edu.cn

Received: 3 September 2019; Accepted: 7 October 2019; Published: 12 October 2019

Abstract: With the development of image editing software techniques, the content integrity and authenticity of original digital images become more and more important in digital content security. A novel image tampering detection and recovery algorithm based on digital watermarking technology and a chaotic system is proposed, and it can effectively locate the tampering region and achieve the approximate recovery of the original image by using the hidden information. The pseudo-random cyclic chain is realized by the chaotic system to construct the mapping relationship between the image subblocks. It can effectively guarantee the randomness of the positional relationship between the hidden information and the original image block for the better ergodicity of the pseudo-random chain. The recovery value optimization algorithm can represent image information better. In addition to the traditional Level-1 recovery, a weight adaptive algorithm is designed to distinguish the original block from the primary recovery block, allowing 3×3 neighbor block recovery to achieve better results. The experimental results show that the hierarchical tamper detection algorithm makes tamper detection have higher precision. When facing collage attacks and large general tampering, it will have higher recovery image quality and better resistance performance.

Keywords: image authentication; image recovery; smooth function; chaotic system; hierarchical tamper detection

1. Introduction

With the development of digital portable devices, such as cell phone and digital camera, the images can be acquired more conveniently. The power of the image editing software become stronger. The authenticity and integrity are so important in digital content security that more and more researchers focus on this field. The watermarking technique, as one of the authentication methods, can detect the authenticity and localize the tampered area effectively, and it can also recover the modified or tampered image.

The algorithms of image self-recovery have three important parts: the authentication information, the recovery information and the mapping function to embed the authentication information and recovery information to the image. The authentication information has to detect the reality of the received image effectively, and it can localize the tampered area accurately. The mapping function can embed the information by modifying selected pixels, and it will influence the quality of the image. Therefore, a better mapping function is a key part of the proposed algorithm. The target of the modified algorithm is to improve the embedded image quality and the tampered area recovery accuracy.

The algorithms of image self-recovery based on watermark can be classified into two types from the embedding field: the spatial embedding and the transform embedding. The spatial embedding methods can modify the pixels directly in the spatial domain. They are simple and effective. For example, the author of [1] proposed the dual watermark to authenticate the tampered image and get a good recovery performance while the big tamper ratio appears. The image is divided into non-overlapping blocks of 2×2 pixels, and then the average values of each block are used to construct the recovery information, which is embedded into the two LSB planes. The scheme in [2] uses parity check and comparison between average intensities and the hierarchical structure is used to detect the tampered area. The authors obtained good recovery performance even in a high tamper ratio. To lower the risk of making an incorrect prediction, the method in [3] produces parity check bits from pixels whose bits have been rearranged. The parity check bits are produced from pixels whose bits have been rearranged. The Hamming code is used to construct the authentication information. To improve the security of Those algorithms, Arnold transform [4–6] is applied in the procedure to map the relationship of the blocks.

The algorithms in the transform domain first map the image into the other domain, such as discrete wavelet transform(DWT) [5,7,8], discrete cosine transform(DCT) [9–11], and lifting wavelet transform(LWT) [6]. Due to characteristics of the transform domain, the authentication and recovery information are generated by coefficients of the transform domain. The index value [12] of Vector Quantization (VQ) is used to recover information. This method can construct recovery information better. However, the index should be used in the watermarking extraction procedure, and this increases the extra information.

To improve the recovery performance and take advantage of spatial embedding, a novel image self-recovery algorithm based on watermarking is proposed. The contributions of the algorithm are as follows:

(1) The pseudo-random cyclic chain is used to construct the mapping relationship between the image subblocks. The pseudo-random cyclic chain is generated by the chaotic system, and the key space is large.
(2) The recovery information can better reflect the information of image blocks by constructing the recovery value optimization algorithm, and the hierarchical tamper detection algorithm makes tamper detection have higher precision.
(3) The modified smooth function is generated to improve the recovery performance, and it can resolve the overflow problem.

The paper is organized as follows. Section 2 presents the mapping mechanism named Double Pseudo-random Chain between the subblocks generated by the chaotic system. Section 3 describes the proposed method. Section 4 reports the performance of the proposed framework on test images. Finally, Section 5 concludes the paper and discusses some possible future work.

2. Double Pseudo-Random Chain

Logistic mapping is a classic model for studying behaviors of complex systems such as dynamic systems, chaos, fractals, etc. Because of its simplicity and randomness, it is widely used in watermark embedding technology [13]. It is defined as:

$$x_{n+1} = \lambda x_n (1 - x_n) \tag{1}$$

where n is the iteration times of chaotic sequences and λ is the control parameters of the logistic mapping system.

For any $x_0 \in (0, 1)$, the Logistic mapping has chaotic characteristics when $3.99 < \lambda < 4.0$ [14,15], it constitutes the pseudorandom chain we need. The double pseudo-random chain has two chaotic sequences which have different x_0, and it has the following characteristics [16]:

(1) The mapping relationship of the double pseudo-random chain is one-to-one and will not be repeated.
(2) The mapping relationship of the double pseudorandom chain is uniquely determined by the initial parameters, and different parameters will produce different sequences, thus ensuring that the mapping sequences are irrelevant.
(3) The neighbor elements of the double pseudo-random chain are far away from each other.

Figure 1 is to test the security of the pseudo-random cyclic chain, and the coefficients of scrambling map are $x_0 = 0.333$ and $\lambda = 3.991$. Figure 1a is a 256×256 8-bit grayscale Lena image, and Figure 1b is the image after reordering the pixels according to the logistic sequence. For instance, The original pixel vector is $[224, 210, 125, 78, 21, 45, 90, 255, ...]$, its index vector is $[0, 1, 2, 3, 4, 5, 6, 7, ...]$, and the mapping index vector is $[4, 1, 7, 0, 2, 6, 3, 5, ...]$ according to the pseudo-random cyclic chain. Then, we can get the reorder pixel vector as $[21, 210, 255, 224, 125, 90, 78, 45, ...]$. It can be seen that two images have little similarity to each other after scrambling, which can demonstrate the randomness of the pseudo-random chain.

Figure 1. The security of the pseudo-random cyclic chain: (**a**) The original Lena image; (**b**) The reordering Lena image.

3. Proposed Scheme

The self-recovery watermarking scheme proposed in this paper includes five parts. The flow chart of the scheme consisting of the watermark generation and embedding procedure is shown in Figure 2 and the tamper detection and recovery procedure is shown in Figure 3. A new mapping mechanism based on the double pseudo-random chain is proposed in the block mapping module, which can embed the current image block information into two corresponding different image blocks. The details of the mapping scheme are proposed in Section 3.1. In the watermark generation module, a new adaptive optimization algorithm is proposed to obtain the best recovery information that can replace the current block information, and generate authentication information according to the recovery information and current block index. The specific algorithm is shown in Section 3.2. In the watermark embedding module, the watermark data are embedded according to the mapping sequence, and the improved smooth function is used to enhance the quality of the embedded image. The algorithm is described with more details in Section 3.3. In the tampering module, combined with the work in [1,5], a four-level detection strategy is proposed, which can effectively resist various types of attacks. The algorithm is presented in Section 3.4. In the image recovery module, in addition to recovering the image with the extracted recovery information, an adaptive secondary recovery strategy is proposed. The algorithm is detailed in Section 3.5.

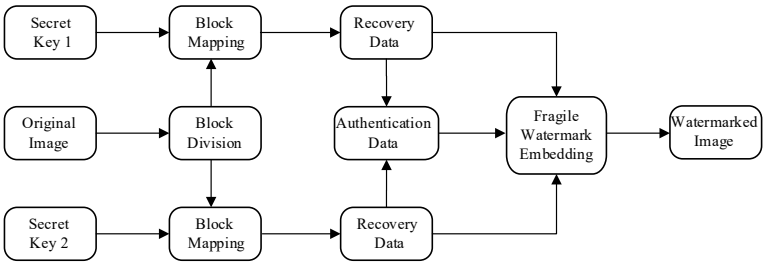

Figure 2. The watermark embedding procedure.

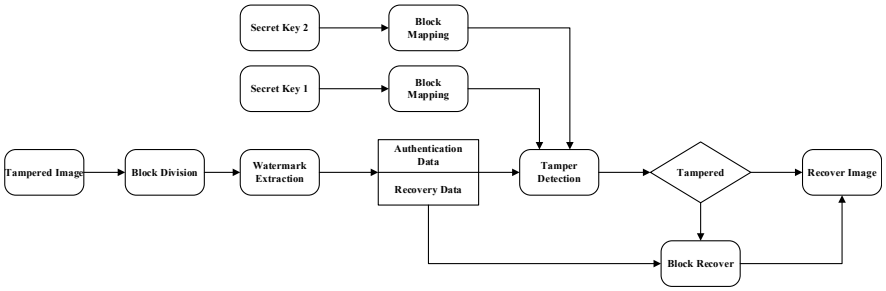

Figure 3. The tamper detection and recovery procedure.

3.1. Block Mapping

A mapping sequence can be represented as A $\rightarrow$ B $\longrightarrow$ C $\rightarrow$ D $\rightarrow \cdots \rightarrow$ A. That is, the recovery information of image block A is hidden in image block B, and the recovery information of image block B hidden in image blocks C, and so on. Thus, for the mapping sequence, the most significant thing is to ensure the existence of the recovery information under the condition of a large tampering ratio. In [1,9,12], the 1-D transform algorithm was used to build the blocks relationship mapping function. Although the sequence is random, the remainder of the elements in the same column are equal, and the performance in the face of column tampering is poor. Arnold transform [4–6] and Variant torus automorphism [3] are used to guarantee the randomness of the sequence, but it is not safe because both of them are periodic. Based on Arnold's Transform algorithm, Tai and Liao [5] remapped the watermark block mapped to the vicinity of the original block, which improves the security of the watermarked image. However, the mechanism of remapping is too ineffective. When the corner tampering rate exceeds 25%, 25% of the watermark will be completely lost, which seriously affects the quality of the image. Modulus operation is used as a mapping sequence by Hamid and Wang [10], but it also has the same problem as the methods in [1,12].

To solve the above problems, a double pseudo-random chain is used to form the block map sequence. Suppose that I_0 is a size of $M \times N$ 8-bit grayscale image, $M, N \in 2^R (R = 8)$. We split the image into nonoverlapping image blocks whose size are $m \times n$, as shown in Equation (2):

$$\boldsymbol{I}_o = \begin{bmatrix} \boldsymbol{I}_{(1,1)} & \cdots & \boldsymbol{I}_{(1,N/n)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{I}_{(M/m,1)} & \cdots & \boldsymbol{I}_{(M/m,N/n)} \end{bmatrix} \tag{2}$$

Next, two different pseudo-random block index sequences named L' and L'' are generated by Equation (1) with different initial values. We embed the data of the first chain in the upper 5 bits and embed the data of the second chain in the middle 5 bits.

As shown in Figure 4, suppose the value of the mapping sequences are

$$\begin{aligned} \boldsymbol{L} &= \left\{ l_0, l_1, l_2, \ldots, l_{\frac{M\times N}{m\times n}} \right\} \\ \boldsymbol{L'} &= \left\{ l'_0, l'_1, l'_2, \ldots, l'_{\frac{M\times N}{m\times n}} \right\} \\ \boldsymbol{L''} &= \left\{ l''_0, l''_1, l''_2, \ldots, l''_{\frac{M\times N}{m\times n}} \right\} \end{aligned} \tag{3}$$

For example, the recovery information of block l'_0 is embedded in the high 5 bits of block l_0, and the recovery information of block l''_0 is embedded in the middle 5 bits of block l_0. l_0, l'_0, and l''_0 are the elements of L, L', and L'', and they are the labels of the divided 2×2 image blocks.

Pearson Linear correlation coefficient(LCC) A is widely used in many fields; it is an effective measurement of the linear correlation between two variables [17]. The larger the correlation coefficient is, the greater the correlation and similarity between the original image and the transformed image is. We use this method to prove that the pseudo-random chain is more random and effective than other methods, such as Arnold transform. It is defined as:

$$\mathrm{LCC}\left(I_o, I'_o\right) = \frac{1}{M \times N - 1} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(\frac{\overline{I_{(i,j)} - \mu_{I_o}}}{\sigma_{I_o}} \right) \left(\frac{I'_{(i,j)} - \mu_{I'_o}}{\sigma_{I'_o}} \right) \tag{4}$$

In this Equation, I_o represents the original image, I'_0 represents the transformed image, μ_{I_o} and $\mu_{I'_o}$ are the average values of I_o and I'_o, and σ_{I_o} and $\sigma_{I'_o}$ are the standard deviations of $\overline{I_o}$ and $\overline{I'_o}$.

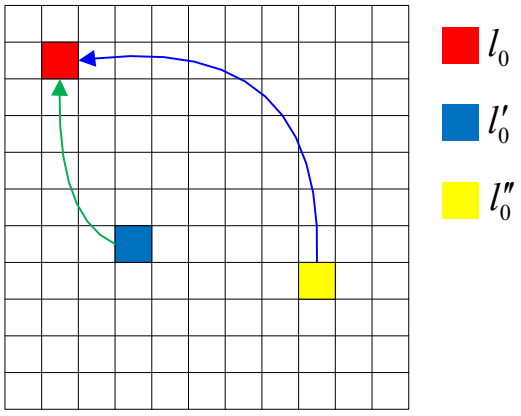

Figure 4. Example of block mapping.

Figure 5 is LCC of Arnold transform and the double pseudo-random chain. In Figure 5a, the two parameters of standard Arnold transform are $a = 1$ and $b = 1$ [18], the abscissa means the number of transform times, 1 means transform once, 2 means transform again based on 1, and so on. In Figure 5b, the parameter of the pseudo-random chain is $\lambda = 3.991$, the abscissa means the initial value range from 0 to 1. Although the abscissas of the two subfigures are not the same, they all represent the traversal of all the conditions under the given parameters and can be used for comparison.

As we can see, the mapping sequence generated by the Arnold transform has a strong correlation, and some of the correlation coefficients can even reach 0.3, while the correlation coefficient of pseudo-random chain is small, almost close to 0. This also proves that the pseudo-random chain has better randomness than Arnold transform and is more suitable as a mapping sequence.

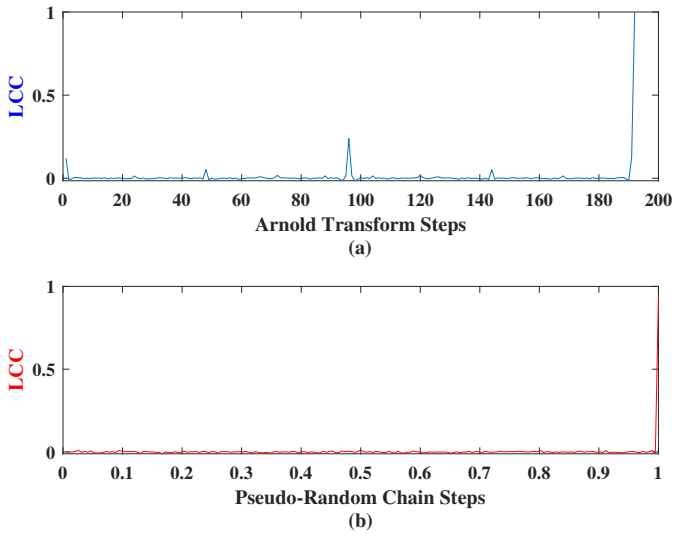

Figure 5. Linear correlation coefficient (LCC) comparison chart.

3.2. Watermark Generation

The watermarks of 8 × 8 image block and 4 × 4 image block are based on DWT and DCT [5,10]. Those methods got more embeddable information and higher embedding image quality. However, the block effect was obvious in this case, and the detection of a small tamper ratio had obvious error and the false alarm rate is high. To solve this problem, the image is divided into 2 × 2 non-overlapping blocks. However, the image block can be embedded with limited information bits, and the maximum capacity has only 12 bits. In the current paper, recovery information is based on the most significant 5 bits of the average of the 2 × 2 image block [1,4,19], which is better for the smooth image block tampering. However, this method is very poor for texture block recovery. A variable-length recovery information construction scheme is proposed by Chen [20], which uses 12 bits for texture blocks and 6 bits for smooth blocks. Although it can represent the texture information better, it affects the embedded information bit, which makes it difficult to tamper detection. Aiming at this problem, a pixel value adaptive optimization algorithm is designed in this paper, which makes the value express image block information better. Figure 6 shows the whole procedures.

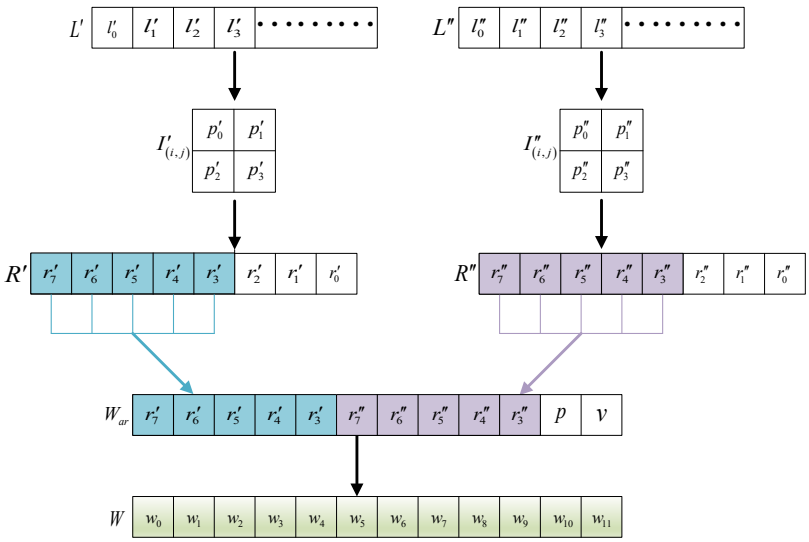

Figure 6. Watermark generation procedure.

Assume that the recovery information of block $I'_{(i,j)}$ and $I''_{(i,j)}$ will are embedded in $I_{(i,j)}$ according to mapping sequences L' and L''. The four pixels in the image block are represented by $I'_{(i,j)} = \{p'_0, p'_1, p'_2, p'_3\}$ and $I''_{(i,j)} = \{p''_0, p''_1, p''_2, p''_3\}$. Since the recovery information is formed by the upper 5-bits of the p'_i or p''_i, its maximum value $p'_{\max}$ and minimum value $p'_{\min}$ can be obtained by Equations (5) and (6).

$$p'_{\min} = \min\left\{\left\lfloor \frac{p'_i}{8} \right\rfloor \times 8, 0 \le i \le 3\right\} \tag{5}$$

$$p'_{\max} = \max\left\{\left\lfloor \frac{p'_i}{8} \right\rfloor \times 8, 0 \le i \le 3\right\} \tag{6}$$

Let R be the recovery information. Then, $R \in [p'_{\min}, p'_{\max}]$. To make the embedded value reflect the actual value of the image block better, we use the following Equation to get the best-embedded value R of the image block.

$$R' = \arg\min_{R} \sum_{i=0}^{3} \left(\left\lfloor \frac{R}{8} \right\rfloor \times 8 - p'_i \right)^2, R \in \left[p'_{\min}, p'_{\max}\right] \tag{7}$$

Then, the upper 5 bits recovery information r'_i can be calculated by:

$$r'_i = \bmod\left(\left\lfloor \frac{R'}{2^i} \right\rfloor, 2 \right), \text{ for } 3 \le i \le 7 \tag{8}$$

Similarly, the recovery information $I''_{(i,j)}$ can be got as:

$$r''_i = \bmod\left(\left\lfloor \frac{R''}{2^i} \right\rfloor, 2 \right), \text{ for } 3 \le i \le 7 \tag{9}$$

Through the above method, we generated 10-bits recovery information r'_i and r''_i. At the same time, we use the following Equation to generate 2-bits authentication information p and v. Here, $\oplus$ is exclusive OR operator, $\sim$ is reverse operator, and $||$ is bitwise stitching operator of binary characters.

$$p = r'_7 \oplus r'_6 \oplus r'_5 \oplus r'_4 \oplus r'_3 \oplus r''_7 \oplus r''_6 \oplus r''_5 \oplus r''_4 \oplus r''_3 \tag{10}$$

$$v = \sim p \tag{11}$$

Finally, the watermark value W_{ar} to be embedded in the image block $I_{(i,j)}$ can be generated by:

$$\boldsymbol{W}_{ar} = r'_7 \left\| r'_6 \right\| r'_5 \left\| r'_4 \right\| r'_3 \left\| r''_7 \right\| r''_6 \left\| r''_5 \right\| r''_4 \left\| r''_3 \right\| p \| v \tag{12}$$

To facilitate the following expression, W_{ar} is expressed as $W = \{w_0, w_1, ..., w_{11}\}$.

3.3. Watermark Embedding

The details of the watermark embedding procedure are shown in Figure 7.

Step 1: Divide the whole $M \times N$ image into $(M \times N)/(m \times n)$ non-overlapping image block $I_{(i,j)}$ whose size is $m \times n$ and let $I_{(i,j)} = \{p_0, p_1, p_2, p_3\}$.

Step 2: Generate two unrelated pseudo-random chains L' and L'' according to the method mentioned in Section 3.1.

Step 3: Calculate the recovery information and the authentication information that need to be embedded, and obtain the whole watermark information W for each image block according to Section 3.2.

Step 4: Embed W into the corresponding image block. However, if the recovery information is directly embedded into the lower 3 bits of the original image, the quality will be greatly affected. Lee [1]

and Yang [12] put forward the smooth function, which has a great effect on improving the quality of the embedded image. However, it also has some drawbacks, thus we make some modifications to it. The function is as follows:

$$v_i = 4 \times w_{3i} + 2 \times w_{3i+1} + w_{3i+2} - (p_i - \bmod (p_i, 8)), \quad \text{for } 0 \le i \le 3 \tag{13}$$

$$wp_i = \begin{cases} p_i + v_i + 8 & \text{if} \quad -7 \le v_i \le -5 \\ p_i + v_i + 0 & \text{if} \quad |v_i| < 5 \\ p_i + v_i - 8 & \text{if} \quad 7 \ge v_i \ge 5 \\ v_i & \text{if} \quad p_i = 0; v_i = 5, 6, 7 \\ v_i & \text{if} \quad p_i = 1; v_i = 6, 7 \\ v_i & \text{if} \quad p_i = 2; v_i = 7 \end{cases}, \text{ for } 0 \le i \le 3 \tag{14}$$

In this Equation, $\{w_{3i}, w_{3i+1}, w_{3i+2}\}$ are the values to be embedded, p_i is the value of original image, v_i is the difference between the lower 3-bit values of original image and $\{w_{3i}, w_{3i+1}, w_{3i+2}\}$, and wp_i are the pixel values of the embedded image generated by smooth function. The principle of this function is to add or subtract 8 from the embedded pixel value without affecting the lower 3-bit values, but it will reduce the difference between the embedded image value and the original image value. For instance, given $p_i = 232 = (11101000)_2$, $\{w_{3i}, w_{3i+1}, w_{3i+2}\} = \{1, 1, 1\}$, and $v_i = (111)_2 - (000)_2 = 7 - 0 = 7$. Using Equation (14), we obtain embedded value $wp_i = 232 + 7 - 8 = 231 = (11100111)_2$. The gap between original pixel value and embedded pixel value is $|wp_i - p_i| = 1$. We get better embedded pixel value without changing $\{w_{3i}, w_{3i+1}, w_{3i+2}\}$. Table 1 shows the effect of the smooth function on watermark embedding.

Table 1. The comparison of "without smoothing function" and "with smoothing function".

Images	Lena	Baboon	Barbara	Peppers	Cameraman
Without smoothing function (dB)	38.22	38.34	38.16	38.45	38.40
With smoothing function (dB)	40.7215	40.7084	40.6967	40.7587	40.7465

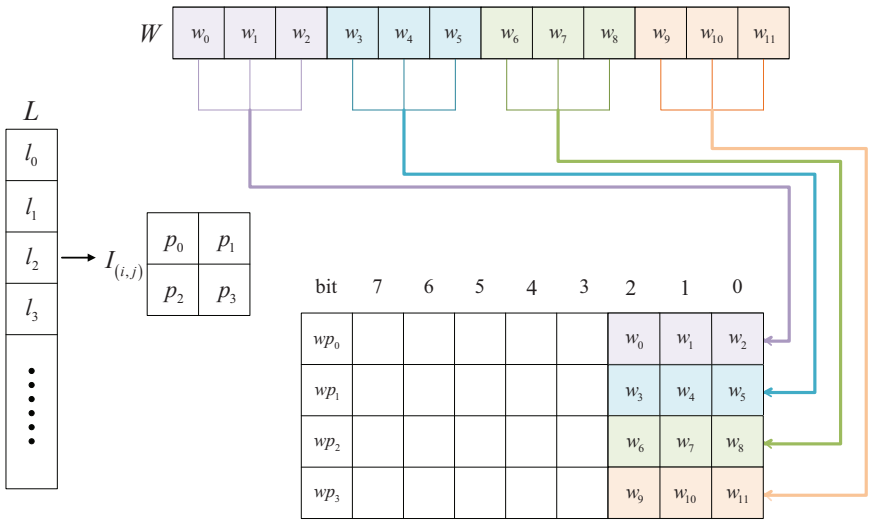

Figure 7. Watermark embedding procedure.

3.4. Hierarchical Tamper Detection

Firstly, the tampered image is partitioned into non-overlapping 2×2 image blocks. The proposed tamper detection algorithms are described below.

Step 1: Use the same x_0 in Section 3.3 to generate two pseudo-random chains L' and L''.

Step 2: Use Equation (15) to extract 12-bit watermark information of the current block $I_{(i,j)}$ from wp_0, wp_1, wp_2, and wp_3.

$$ew_i = \mathrm{mod}\left(\left\lfloor \frac{wp_{\lfloor i/3 \rfloor}}{2^{i \bmod 2}} \right\rfloor, 2\right), \text{ for } 0 \le i \le 11 \tag{15}$$

Let $W_e = \{ew_0, ew_1, ew_2, ..., ew_{11}\}$.

Step 3: Level-1 detection. According to Section 3.2, we can easily notice that W_e will satisfy Equations (10) and (11) if the image block has not been tampered. Thus, we can calculate and compare ew_0 with $ew_0 \oplus ew_1 \oplus ew_2 \oplus ew_3 \oplus ew_4 \oplus ew_5 \oplus ew_6 \oplus ew_7 \oplus ew_8 \oplus ew_9$, and then calculate and compare ew_{11} with $\sim ew_{10}$. If one of the comparison results is not equal, the current block is set to be invalid.

Step 4: Level-2 detection. If the current block is detected as valid in the Level-1 detection, use Equation (16) to decode W_e to obtain recovery information.

$$\begin{aligned} R'_e &= 128 \times ew_0 + 64 \times ew_1 + 32 \times ew_2 + 16 \times ew_3 + 8 \times ew_4 \\ R''_e &= 128 \times ew_s + 64 \times ew_6 + 32 \times ew_7 + 16 \times ew_s + 8 \times ew_0 \end{aligned} \tag{16}$$

According to Section 3.2, R'_e and R''_e will satisfy Equations (8) and (9). Thus, if $R'_e = \lfloor R'/8 \rfloor \times 8$ and $R''_e = \lfloor R''/8 \rfloor \times 8$, set the block to be a valid block, otherwise set it as an invalid block.

Step 5: Level-3 detection. If the current block is marked as valid in the Level-2 detection, the current block with 3×3 neighbor block is considered for detection. As shown in Figure 8, The neighbor block is divided into four groups of triples, which are (N, NE, E), (E, SE, S), (S, SW, W), (W, NW, N). If any triple is invalid, the current block is marked as invalid [1].

NW	*N*	*NE*
W	$I_{(i,j)}$	*E*
SW	*S*	*SE*

Figure 8. The four triples of current block.

Step 6: Level-4 detection. Consider the 3×3 neighbor block as shown in Figure 9. If there are 4 or more invalid blocks, the current block is marked as invalid.

Error	*Error*	
	$I_{(i,j)}$	*Error*
Error		*Error*

Figure 9. The 3×3 block-neighborhood of current block.

3.5. Image Recovery

When the tamper detection is completed, the blocks marked as invalid should be recovered. A scheme for secondary recovery is designed to get better recovered image quality. Suppose the block marked as invalid is $I_{(i,j)} = \{p_0, p_1, p_2, p_3\}$. The recovery information hide in $I'_{(i,j)}$ and $I''_{(i,j)}$ according to L' and L''.

The detailed procedure of Level-1 self-recovery is as follows:

Step 1: If $I'_{(i,j)}$ is a valid block, go to Step 2. If the $I'_{(i,j)}$ is an invalid block, go to Step 3.

Step 2: Extract the watermark information from the lower three bit planes of block $I'_{(i,j)}$ and use Equation (17) to get the recovery value $I^r_{(i,j)}$. Then, go to Step 4.

$$I^r_{(i,j)} = 128 \times rw'_0 + 64 \times rw'_1 + 32 \times rw'_2 + 16 \times rw'_3 + 8 \times rw'_4 \tag{17}$$

Step 3: If $I''_{(i,j)}$ is an invalid block, the $I_{(i,j)}$ is marked as the invalid block. If $I''_{(i,j)}$ is a valid block, then extract the watermark information from the lower 3-bits of $I''_{(i,j)}$ and use Equation (18) to get the recovery value $I'_{(i,j)}$, then go to Step 4.

$$I'_{(i,j)} = 128 \times rw''_5 + 64 \times rw''_6 + 32 \times rw''_7 + 16 \times rw''_8 + 8 \times rw''_9 \tag{18}$$

Step 4: Use $I^r_{(i,j)}$ to recover the current invalid block $I_{(i,j)}$.

If $I'_{(i,j)}$ and $I''_{(i,j)}$ are both invalid blocks, then the current tampering block cannot be recovered by extracting the recovery information. This situation is more likely to occur when the tampering ratio is high, so the design of the level-2 recovery strategy is important for image recovery.The mean of the 3×3 neighboring blocks is used to recover the image [1,5]. However, this method ignores the difference between the original block and the Level-1 recovery block. The pixel value of the original block is more accurate than the recovery block. In this paper, an adaptive weighted recovery algorithm is designed to resolve this problem.

The Level-2 recovery process is as follows:

Step 1: In the 3×3 neighboring block, let the number of original blocks is n_o, the pixel weight is μ_o and the pixel value is p_i $(0 \le i \le n_o)$. The number of recovery blocks is n_r, the corresponding pixel values and the pixel weight are p_i and μ_r $(0 \le j \le n_r)$. The number of unrecovered blocks is n_t, and the corresponding pixel weight is 0. k is the ratio of the original pixel block weight to the recovery pixel block weight and it is generally 1.5 or 2, then μ_o , μ_r can be calculated by Equations (19) and (20):

$$\begin{cases} \mu_o = k\mu_r (k \ge 1) \\ n_o\mu_o + n_r\mu_r = 1 \end{cases} \tag{19}$$

Solving Equation (19), μ_o and μ_r can be obtained:

$$\begin{cases} \mu_o = \dfrac{k}{n_o k + n_r} \\ \mu_r = \dfrac{1}{n_o k + n_r} \end{cases} \tag{20}$$

Then, the current invalid block pixel value can be calculated by:

$$\boldsymbol{I}^r_{(i,j)} = \mu_o \sum_{i=0}^{n_0} p_i + \mu_r \sum_{j=0}^{n_r} p_j \tag{21}$$

We can prove $I^r_{(i,j)} \in [0, 255]$ as follows.

Assume $p_{\max}$ is the maximum value of p_i and p_j, $p_{\min}$ is the minimum value of p_i and p_j, and ii is easy to get $p_{\max}, p_{\min} \in [0, 255]$. Then, Equation (22) can be obtained by Equation (21)

$$\begin{aligned} \boldsymbol{I}^r_{(i,j)} &= \mu_o \sum_{i=0}^{n_0} p_i + \mu_r \sum_{j=0}^{n_r} p_j \\ &\leq \mu_o \sum_{i=0}^{n_o} p_{\max} + \mu_r \sum_{j=0}^{n_r} p_{\max} \\ &= \mu_o n_o p_{\max} + \mu_r n_r p_{\max} \\ &= \left(\mu_o n_o + \mu_r n_r\right) p_{\max} \\ &= p_{\max} \leq 255 \end{aligned} \tag{22}$$

Similarly, we can prove

$$I_{(i,j)} \geq p_{\min} \geq 0 \tag{23}$$

Therefore, $I^r_{(i,j)} \in [0, 255]$, which can guarantee that the modified pixel value does not overflow.

The recovery procedure is shown in Figure 10. The brown block is marked as the recovered image block after tampering, and the purple block is marked as the original block. As can be seen in the figure, $n_r = 4$ and $n_o = 3$, and we set $k = 1.5$. Thus, we can calculate the weight of the original block and recovery block as $\mu_o = 0.18$ and $\mu_r = 0.12$. Finally, we can calculate $I^r_{(i,j)}$ according to Equation (21).

Step 2: Use $I^r_{(i,j)}$ to recover invalid block $I_{(i,j)}$.

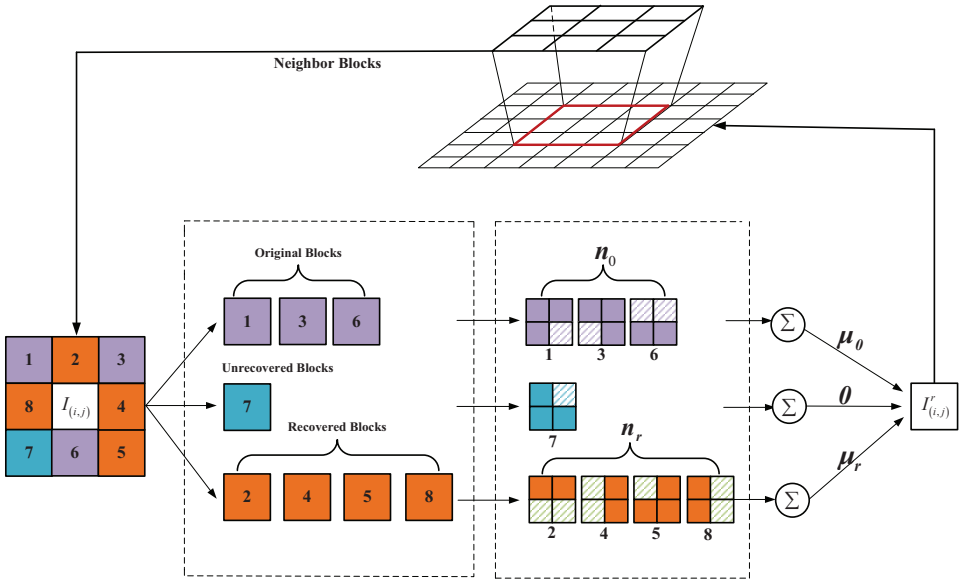

Figure 10. Level-2 recovery procedure.

4. Experimental Results and Performance Analysis

We performed a series of analyses and simulations of the performance of the proposed scheme in tamper detection. The types of attacks include collage attacks and large general tampering. We also compared the performance with the existing block-based approach [1,2,5,9]. All simulation environments were MATLAB R2018b.

4.1. PSNR

Peak signal-to-noise ratio ($PSNR$) is widely used in the field of image processing. It can measure the degree of deviation of the watermarked image or recovered image from the original image [21].

$$\begin{array}{c} MSE = \frac{1}{M\times N}\sum_{i=0}^{M}\sum_{j=0}^{N}\left[\boldsymbol{I}_{(i,j)}^{O}-\boldsymbol{I}_{(i,j)}^{R}\right]^{2} \\ PSNR = 10\times\log_{10}\frac{255^{2}}{MSE} \end{array} \tag{24}$$

where $I_{(i,j)}^{O}$ represents the pixel value of the original image and $I_{(i,j)}^{R}$ is the pixel value of the reconstructed image. Table 2 is the $PSNR$ of the watermarked image, which is about 2.82 dB higher than the theoretically value of 37.9 dB [21] due to the addition of the smooth function.

Table 2. The $PSNR$ (dB) of the watermarked images.

Images	Lena	Baboon	Barbara	Peppers	Cameraman
Wang [9]	39.82	- [1]	39.52	-	40.65
Lee [1]	40.68	40.73	40.72	40.73	-
Tai [5]	44.0119	44.0247	44.0736	44.0516	44.1004
Proposed	40.7215	40.7084	40.6967	40.7587	40.7465

[1] indicates that no data are provided.

4.2. Performance on Collage Attack

The first one is to collage the block of the current image to another block, and the second one is to collage the block of one image to the corresponding position of another image.

4.2.1. First Kind of Collage Attack

A 256 × 256 8-bit grayscale Lena image was used to simulate first kind collage attacks.

Figure 11 shows the details for the simulation. Figure 11a is the watermarked image in [1] with a $PSNR$ of 40.72 dB. Figure 11b shows the result of the collage attack in [1]; we paste the blocks of coordinates (47,41) to (210,124) into the image blocks of coordinates (47,133) to (210,216). The theoretical tampering rate is 21.02%. Figure 11c show the result of the tamper detection for paper [1]. As sown in this figure, the method in [1] could not detect the first kind collage attack. Figure 11d is the result of image recovery of Figure 11b. Since the image tampering cannot be detected, and the $PSNR$ is only 17.26 dB. Figure 11a1 is the watermarked image of Tai [5] with a $PSNR$ of 44.01 dB. Figure 11b1 is the collage image of Figure 11a1. In Figure 11c1, we can see that Tai's method can detect the first kind collage attack and Figure 11d1 shows the recovery image and the $PSNR$ is 29.21 dB. The block effect is obvious from the red part. Figure 11a2 is the watermarked image of our scheme, and the $PSNR$ is 40.71 dB. Figure 11b2 is the collage images, the ratio and location of tampering are the same as Figure 11b. Figure 11c2 shows the results of tampering detection. We can detect the first kind of collage attack and the detection tamper ratio is 21.02%, which is consistent with the theory. Figure 11d2 is recovery image with a $PSNR$ of 32.35 dB.

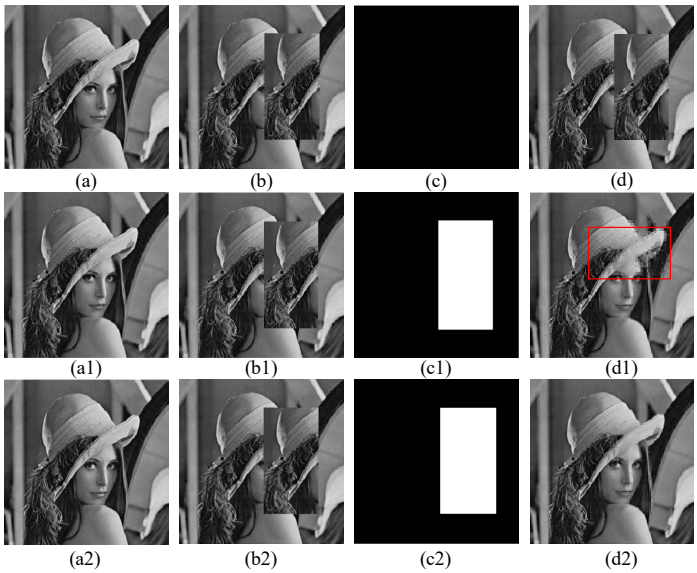

Figure 11. Collage attack simulation of first kind: (**a**) Lee [1] watermarked image ($PSNR$ = 40.72 dB); (**b**) Lee tampered image; (**c**) Lee tamper detective image; (**d**) Lee recovered image ($PSNR$ = 17.26 dB); (**a1**) Tai [5] watermarked image ($PSNR$ = 44.01 dB); (**b1**) Tai tampered image; (**c1**) Tai tamper detective image; (**d1**) Tai recovered image ($PSNR$ = 29.21 dB); (**a2**) Proposed watermarked image ($PSNR$ = 40.72 dB); (**b2**) proposed tampered image; (**c2**) proposed tamper detective image; and (**d2**) proposed recovered image ($PSNR$ = 32.35 dB).

4.2.2. Second Kind of Collage Attack

We also collaged 256 × 256 8-bit grayscale Lena to 256 × 256 8-bit grayscale Baboon to simulate the second kind collage attack.

Figure 12a is the watermarked image from [1] with a $PSNR$ of 40.74 dB. Figure 12b is the result of the second collage attack of the method in [1]; we paste Lena image blocks of coordinates (47,41) to (210,124) into Baboon image blocks with coordinates (47,133) to (210,216). The theoretical tampering ratio is 21.02%. Figure 12c is the result of the tamper detection for the method in [1], which cannot detect the collage attack. Figure 12d is the result of recovery image for Figure 12b with a $PSNR$ of 16.56 dB. Figure 12a1 is watermarked image of Tai [5] with a $PSNR$ of 44.02 dB. Figure 12b1 is the tampered image of Figure 12a1. As the first kind attack, the second collage attack can be detected by Tai's [5] scheme. Figure 12d1 shows the information of the recovery image, and the block effect can be observed. Figure 12a2 is a watermarked image of proposed scheme with a $PSNR$ of 40.71 dB. Figure 12b2 shows the collage images, and the ratio and location of tampering are the same as in Figure 12b. Figure 12c2 shows the results of the tampering detection. From this, we can see that we can detect the second kind collage attack and the detection tamper ratio is 21.02%, which is consistent with the theory. Figure 12d2 is a recovery image with a $PSNR$ of 32.35 dB.

In summary, compared with the the methods in [1,5], the proposed scheme is resistant to the collage attack.

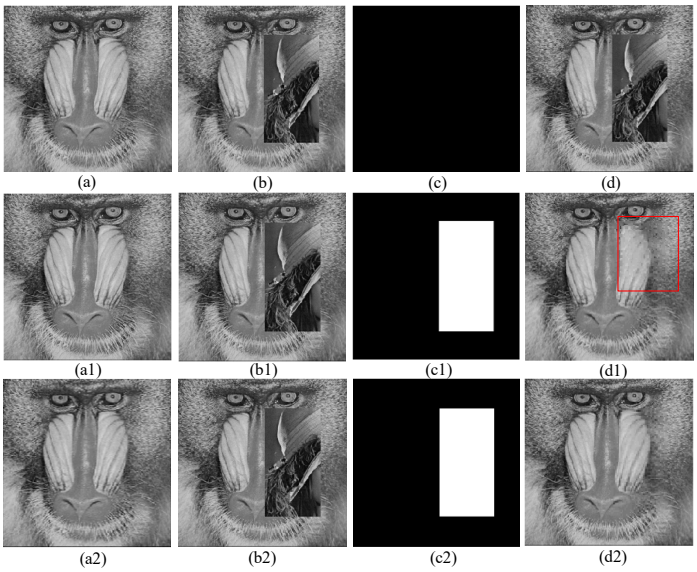

Figure 12. Collage attack simulation second kind: (**a**) Lee [1] watermarked image ($PSNR$ = 40.74 dB); (**b**) Lee tampered image; (**c**) Lee tamper detective image; (**d**) Lee recovered image ($PSNR$ = 16.55 dB); (**a1**) Tai [5] watermarked image ($PSNR$ = 44.02 dB); (**b1**) Tai tampered image; (**c1**) Tai tamper detective image; (**d1**) Tai recovered image ($PSNR$ = 31.57 dB); (**a2**) proposed watermarked image ($PSNR$ = 40.71 dB); (**b2**) proposed tampered image; (**c2**) proposed tamper detective image; and (**d2**) proposed recovered image ($PSNR$ = 32.35 dB).

4.3. Performance on Large General Tampering

To simulate the performance of the scheme on the large general tampering problem, we tampered with the watermarked image to the degree from 0% to 95%. The simulated image is a 256 × 256 8-bit Lena and the results are shown in Table 3.

Table 3. The $PSNR$ (dB) of the recovered image relative to the tampered size and location.

Tampered Size	Tampered	Location	Lin [2]	Lee [1]	Tai [5]	Proposed
40 × 40	2.4	Center	39.96	39.48	39.19	37.46
80 × 80	9.7	Center	36.24	35.17	33.85	34.26
256 × 64	25	Left	31.60	33.45	32.21	33.28
85 × 256	34	Top	27.37	33.01	35.03	33.49
164 × 164	40	Center	23.97	27.97	24.36	28.21
200 × 200	61	Center	19.47	25.20	21.83	25.99
206 × 206	65	Center	-	24.57	21.26	25.68
214 × 214	70	Center	-	24.16	20.45	24.99
222 × 222	75	Center	-	23.43	19.44	24.18
230 × 230	80	Center	-	22.55	18.57	22.99
236 × 236	85	Center	-	21.28	17.45	21.89
244 × 244	90	Center	-	19.86	16.18	19.97
250 × 250	95	Center	-	18.05	15.24	17.82

In Table 3, we can see that compared to the methods of Lin [2], Lee [1] and Tai [5], the proposed scheme has better performance when the tamper ratio is 33–90%. The mapping of double pseudo-random chain has better ergodicity and will have better effect on general tampering. In contrast, the mapping sequences of Lin and Lee are not random. For example, the remainder of each column of Lee's mapping sequence is equal, which led to better results in very low tampering rates and extremely high tampering rates. Although the mapping sequence of Tai [5] is a better random sequence, it is also periodic. The 2-bit embedding method makes the watermarked image of Tai have a higher $PSNR$. However, the recovery information is embedded just once; the secondary recovery scheme is not

effective. The performance is not as good as our method in the case of a high tampering ratio. Overall, the proposed approach is more general and practical.

The previous paragraph mentioned the limitations of Lin and Lee's method. To illustrate this problem, we specifically simulated a large number of column tampering. The contrast results of the column tampering Peppers are shown in Figure 13. The column tampering test results for the Lena, Baboon, Barbara, Peppers, and Cameraman are shown in Figure 14. In Figure 13, the proposed scheme has a better effect on tamper detection, and the *PSNR* of the recovery image is higher. Furthermore, from the red pane in Figure 13d, we can see that Lee's scheme has the Probability of False Rejection (PFR) for Peppers because the smooth function has not been improved. We can see in Figure 14 that, for a large number of column tampering, the proposed scheme has better resistance than Lin and Lee.

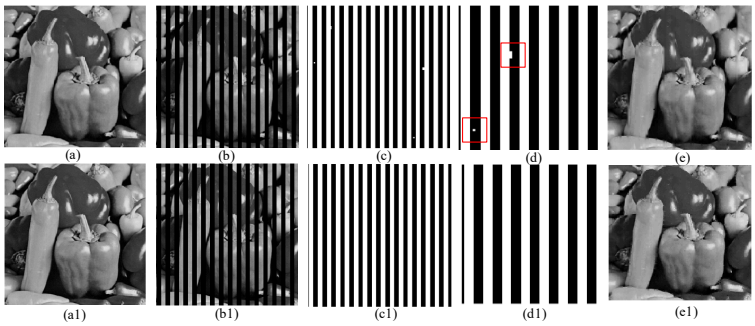

Figure 13. Longitudinal tampering distribution in Peppers (50% tampered): (**a**) Lee [1] watermarked image (*PSNR* = 40.73 dB); (**b**) Lee tampered image; (**c**,**d**) Lee tamper detective image; (**e**) Lee recovered image (*PSNR* = 27.42 dB); (**a1**) proposed watermarked image (*PSNR* = 40.76 dB); (**b1**) proposed tampered image; (**c1**,**d1**) proposed tamper detective image; and (**e1**) proposed recovered image (*PSNR* = 29.71 dB).

Figure 14. Longitudinal tampering distribution in Lena, Baboon, Barbara and Cameraman (50% tampered): (**a**–**d**) original images; (**a1**–**d1**) Lee recovered images (*PSNR* = 27.17 dB, 28.01 dB, 25.93 dB, and 24.67 dB); and (**a2**–**d2**) proposed recovered images (*PSNR* = 29.35 dB, 27.14 dB, 30.12 dB, and 29.80 dB).

Besides, to be more universal, five 8-bit grayscale images from the standard test set were taken for 0–95% central tampering simulation in this study. The images were: Lena, Baboon, Barbara, Peppers, and Cameraman. The results are shown in Figure 15. In Figure 15, for different types of images with

different characteristics, the $PSNR$ of the restored images is not much different, which proves the versatility of our scheme. Moreover, as the tampering ratio increases from 0% to 95%, the $PSNR$ of the recovered image decreases smoothly and the $PSNR$ range is always 40dB $\sim$ 18dB, which proves the efficiency of the algorithm.

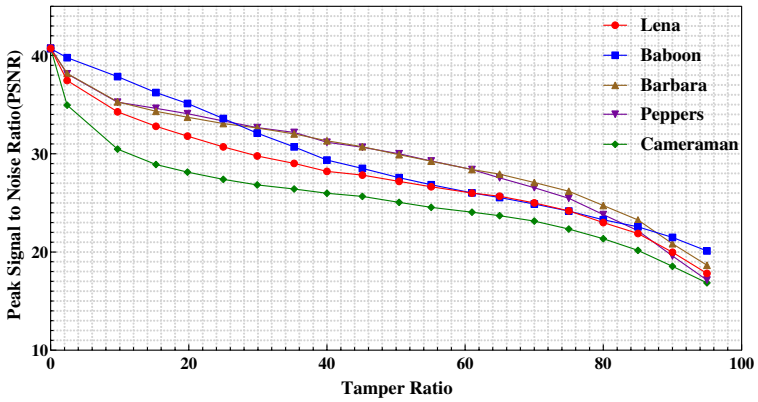

Figure 15. The $PSNR$ of the recovered image relative to the tampered ratio.

5. Conclusions

To authenticate the integrity of the digital image and locate the tamper area, a novel image self-recovery scheme based on watermarking technique is proposed. The mapping relationship between the subblocks is constructed by the chaotic system, thus the security of the algorithm is better. The authentication and the recovery information are generated by the image block content. The optimization algorithm is used to find the better recovery information, which makes the recovery performance better. A weight adaptive algorithm is proposed to assign different weight to the original block and the primary recovery block, and it is different from the traditional Level-2 recovery scheme, which makes the 3×3 neighbor block recovery achieve better results. Many experiments and analysis were done to show better performance of this method.

Author Contributions: The detailed contributions to this paper are as follows: conceptualization, W.S., Y.L. and X.Z.; methodology, W.S. and Y.L.; software, Y.L., J.W. and L.Z.; validation, W.S., J.W. and L.Z.; formal analysis, Y.L.; investigation, Y.L., J.W. and L.Z.; resources, W.S. and X.Z.; data curation, Y.L.; writing—original draft preparation, Y.L., W.S. and J.W.; writing—review and editing, W.S. and Y.L.; visualization, Y.L.; supervision, W.S., X.Z. and L.Z.; project administration, W.S., X.Z. and L.Z.; funding acquisition, W.S., X.Z. and L.Z.

Funding: This work was supported in part by National Science Foundation Project of P. R. China under Grant No.61701554 and State Language Commission Key Project (ZDl135-39), Promotion plan for young teachers' scientific research ability of Minzu University of China, MUC 111 Project, First-class University and First class Discipline of Minzu University of China ("intelligent computing and network security"), and the youth team leadership program

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, T.Y.; Lin, S.D. Dual watermark for image tamper detection and recovery. *Pattern Recognit.* **2008**, *41*, 3497–3506, doi:10.1016/j.patcog.2008.05.003. [CrossRef]
2. Lin, P.L.; Hsieh, C.K.; Huang, P.W. A hierarchical digital watermarking method for image tamper detection and recovery. *Pattern Recognit.* **2005**, *38*, 2519–2529, doi:10.1016/j.patcog.2005.02.007. [CrossRef]
3. Chan, C.S. An image authentication method by applying Hamming code on rearranged bits. *Pattern Recognit. Lett.* **2011**, *32*, 1679–1690, doi:10.1016/j.patrec.2011.07.023. [CrossRef]
4. Shehab, A.; Elhoseny, M.; Muhammad, K.; Sangaiah, A.K.; Yang, P.; Huang, H.; Hou, G. Secure and Robust Fragile Watermarking Scheme for Medical Images. *IEEE Access* **2018**, *6*, 10269–10278. [CrossRef]

5. Tai, W.L.; Liao, Z.J. Image self-recovery with watermark self-embedding. *Signal Process. Image Commun.* **2018**, *65*, 11–25, doi:10.1016/j.image.2018.03.011. [CrossRef]
6. Bolourian Haghighi, B.; Taherinia, A.H.; Harati, A. TRLH: Fragile and blind dual watermarking for image tamper detection and self-recovery based on lifting wavelet transform and halftoning technique. *J. Vis. Commun. Image Represent.* **2018**, *50*, 49–64, doi:10.1016/j.jvcir.2017.09.017. [CrossRef]
7. Wu, X.; Hu, J.; Gu, Z.; Huang, J. A Secure Semi-Fragile Watermarking for Image Authentication Based on Integer Wavelet Transform with Parameters. In *Proceeding ACSW Frontiers '05 Proceedings of the 2005 Australasian Workshop on Grid Computing and E-Research*; Australian Computer Society, Inc.: Darlinghurst, Australia, 2005; p. 6.
8. Bi, N.; Sun, Q.; Huang, D.; Yang, Z.; Huang, J. Robust Image Watermarking Based on Multiband Wavelets and Empirical Mode Decomposition. *IEEE Trans. Image Process.* **2007**, *16*, 1956–1966, doi:10.1109/TIP.2007.901206. [CrossRef] [PubMed]
9. Wang, C.; Zhang, H.; Zhou, X. A Self-Recovery Fragile Image Watermarking with Variable Watermark Capacity. *Appl. Sci.* **2018**, *8*, 548, doi:10.3390/app8040548. [CrossRef]
10. Hamid, M.; Wang, C. Adaptive Image Self-Recovery Based on Feature Extraction in the DCT Domain. *IEEE Access* **2018**, *6*, 67156–67165, doi:10.1109/ACCESS.2018.2879404. [CrossRef]
11. Huang, J.; Shi, Y.; Shi, Y. Embedding image watermarks in dc components. *IEEE Trans. Circuits Syst. Video Technol.* **2000**, *10*, 974–979, doi:10.1109/76.867936. [CrossRef]
12. Yang, C.W.; Shen, J.J. Recover the tampered image based on VQ indexing. *Signal Process.* **2010**, *90*, 331–343, doi:10.1016/j.sigpro.2009.07.007. [CrossRef]
13. Boeing, G. Visual Analysis of Nonlinear Dynamical Systems: Chaos, Fractals, Self-Similarity and the Limits of Prediction. *Systems* **2016**, *4*, 37, doi:10.3390/systems4040037. [CrossRef]
14. Lawnik, M. Generalized logistic map and its application in chaos based cryptography. *J. Phys. Conf. Ser.* **2017**, *936*, 012017, doi:10.1088/1742-6596/936/1/012017. [CrossRef]
15. Arroyo, D.; Alvarez, G.; Fernandez, V. On the inadequacy of the logistic map for cryptographic applications. *arXiv* **2008**, arXiv:0805.4355.
16. Song, W.; Hou, J.J.; Li, Z.H.; Huang, L. Chaotic system and QR factorization based robust digital image watermarking algorithm. *J. Cent. South Univ. Technol.* **2011**, *18*, 116–124, doi:10.1007/s11771-011-0668-8. [CrossRef]
17. Pearson Correlation Coefficient. 2019. Available online: https://en.wikipedia.org/w/index.php?title=Pearson_correlation_coefficient&oldid=917347908 (accessed on 25 September 2019).
18. Arnold's Cat Map. 2019. Available online: https://en.wikipedia.org/w/index.php?title=Arnold%27s_cat_map&oldid=908085373 (accessed on 25 September 2019).
19. Gull, S.; Loan, N.A.; Parah, S.A.; Sheikh, J.A.; Bhat, G.M. An efficient watermarking technique for tamper detection and localization of medical images. *J. Ambient. Intell. Hum. Comput.* **2018**. [CrossRef]
20. Chen, F. Variable capacity recovery watermarking algorithm for image authentication. *Chin. J. Comput.* **2012**, *35*, 154–162. (In Chinese) [CrossRef]
21. Sreenivas, K.; Kamkshi Prasad, V. Fragile watermarking schemes for image authentication: A survey. *Int. J. Mach. Learn. Cybern.* **2018**, *9*, 1193–1218, doi:10.1007/s13042-017-0641-4. [CrossRef]