entropy

MDPI

Article

A Robust Image Tampering Detection Method Based on Maximum Entropy Criteria

Bo Zhao [1], Guihe Qin [1,2] and Pingping Liu [1,2,*]

Received: 12 August 2015 ; Accepted: 18 November 2015 ; Published: 1 December 2015
Academic Editors: Umberto Lucia and Giuseppe Grazzini

1 College of Computer Science and Technology, Jilin University, Changchun 130012, China; wolfers509@126.com (B.Z.); qingh@jlu.edu.cn (G.Q.)
2 Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China
* Correspondence: liupp@jlu.edu.cn

Abstract: This paper proposes a novel image watermarking method based on local energy and maximum entropy aiming to improve the robustness. First, the image feature distribution is extracted by employing the local energy model and then it is transformed as a digital watermark by employing a Discrete Cosine Transform (DCT). An offset image is thus obtained according to the difference between the extracted digital watermarking and the feature distribution of the watermarked image. The entropy of the pixel value distribution is computed first. The Lorenz curve is used to measure the polarization degree of the pixel value distribution. In the pixel location distribution flow, the maximum entropy criteria is applied in segmenting the offset image into potentially tampered regions and unchanged regions. All-connected graph and 2-D Gaussian probability are utilized to obtain the probability distribution of the pixel location. Finally, the factitious tampering probability value of a pending detected image is computed through combining the weighting factors of pixel value and pixel location distribution. Experimental results show that the proposed method is more robust against the commonly used image processing operations, such as Gaussian noise, impulse noise, *etc.* Simultaneously, the proposed method achieves high sensitivity against factitious tampering.

Keywords: image tampering detection; watermarking; maximum entropy; offset image

1. Introduction

1.1. Relevant Background

Image watermarking is currently a research hot spot as it is an effective secure information transformation approach. In a real network transmission environment, a digital image is vulnerable to human tampering attacks, such as cropping, replacement, *etc*. Meanwhile, during the data transmission process, especially in a wireless environment, an image might experience noise interference generated by many factors, such as the distance, terrain, the spatial environment, and lack of power in both the transmitter and/or receiver. These types of interference cause further different degrees of change in the image content. The image data changes caused by natural noise are usually considered to be acceptable, whereas those caused by factitious tampering are unacceptable. Image tampering detection that must distinguish noise interference from factitious tampering often requires high robustness. When the image information is almost similar, a certain degree of distortion is allowed. However, malicious tampering, which could obviously change image information, is not allowed. Image watermarking algorithms that can detect the changes in image content caused by factitious tampering with a certain degree of tolerance to natural noise have been developed and

widely used [1–7]. Cox [8] embedded watermarks into the perceptually most significant spectral components of original images. This method used a watermark comparison method so it could find the image changes, but this method didn't use a content-based image authentication technique so it could not identify whether the image content was tampered with or not. Kang [9] used the difference of the frequency coefficients derived from a discrete cosine transform or a discrete wavelet transform. The BCH code as the sync signal has been used to error-correct the watermark bits to withstand JPEG compression and cropping attacks, but this method can't distinguish JPEG compression and cropping attacks. Walton [10] used the quantized wavelet coefficient to embed watermarking and then quantized the image by employing the Haar wavelet coefficient. The robustness of the watermarking is controlled by the size of the quantization step. To identify whether an image has suffered from malicious tampering, an attack estimate function is employed. The disadvantage of this algorithm is that determining the quantization step length is easy to perform, so attackers can change the image content while keeping the watermarking information. Egger [11] applied the scalar encoder to encode the watermarking information and quantized the selected 8×8 block coefficient through the corresponding scalar quantization function. He then used binary pseudo-random sequence images to embed the watermark. At the authentication step, a mixed operation is applied on the corresponding quantitative function and the pending detection images to obtain a verified value. However, the disadvantage of this algorithm is its high sensitivity to the histogram equalization and sharpening operations. Lu [12] separated the masking threshold wavelet coefficients into masking threshold units (MTUs) and chose wavelet coefficients in the frequency domain that has similar dimensions and orientations, and the absolute value is larger than the Just Noticeable Distortion (JND) threshold. Afterward, a cocktail watermarking scheme is used to adjust the wavelet coefficients to complete the watermark embedding. The original quantitative information is used as a secret key to restore the original image. Two types of watermark detection were operated to complete the authentication. The method took the features of two types of watermarking into account, but when used in image tampering detection, the original watermarking has to be stored, which could affect the robustness and lead to equilibrium fragility. Yu [13] proposed embedding the watermarking through quantifying a weighted average of the wavelet coefficients. By assuming that the change of wavelet coefficients follow a normal distribution, factitious tampering can cause a larger variance of the local change. However, the variance of the image distortion caused by noise is always smaller. Under this assumption, image tampering detection that differentiates factitious tampering from malicious tampering can be realized. Compared with the direct quantitative wavelet coefficient, it has better robustness. However, when the tampering happens in a small area, the accuracy is usually poor. After studying the above mentioned algorithms, a new image watermarking method with the aim of achieving a high detection accuracy for image operations, such as cutting, replacement, and conventional factitious tampering is investigated in this paper. Meanwhile, satisfactory robustness against noise attacks should be achieved. The algorithm should also ensure excellent accuracy, even in a communication channel with a low signal-to-noise ratio.

1.2. Motivations and Contributions

This paper proposes a robust image watermarking algorithm combining the information of pixel value and pixel location distributions based on maximum entropy criteria. First, the Itti model [14] is applied to get the original image feature, which is referred to as the saliency map. Then the saliency map is used as the watermarking, and then the watermarking is embedded into the original image using the Discrete Cosine Transform-Singular Value Decomposition (DCT-SVD) technique [15]. In the watermarking extraction step, the inverse transformation of DCT-SVD is used to extract the watermarking from the pending detection image, which is the saliency map of the original image. After that the local energy model [16] is also applied to get the saliency map of the watermarked image. Then the two saliency maps are substracted to get the offset image. The information in the offset image will be used to carry out the image tampering detection. Two main types of

information are necessary in this study: pixel value and pixel location distributions. Therefore, the image tampering detection algorithm is divided into two workflows: pixel value and pixel location. In the pixel value flow, the entropy of the pixel distribution in the offset image is calculated. The difference algorithm based on the Lorenz curve is also utilized to obtain the divergence indicator of the pixel value distribution. The pixel value entropy and divergence indicator are then combined to produce a joint weighting factor. In the pixel location flow, the maximum entropy model is adopted to segment the offset image into potential tampered and untampered regions. Next, an all-connected graph of pixel difference matrix is applied to obtain the average distance between the potential tampering pixels. The deviation function of high-difference pixels is obtained based on the distribution probability of the coordinate. The next step is to compute the weighting factor of the pixel location distribution. Finally, the weighting coefficients of the pixel value and pixel location are combined. The results of image tampering detection are obtained by comparing the coefficients with the pre-set threshold. The whole image tampering detection workflow includes the following steps:

1. Watermark generation: The original image's saliency map is adopted as the watermark.
2. Watermark embedding: Embedding the watermark into the original image by DCT-SVD.
3. Image transmission: Sender sends the image to receiver (through a lossy channel).
4. Watermark extraction: Extracting the watermark from the pending detection image.
5. Saliency map extraction: Extracting the saliency map from the pending detection image.
6. Tamper detection: Determining whether the image has been factitiously tampered with.

The simple process based on the proposed method is shown in Figure 1.

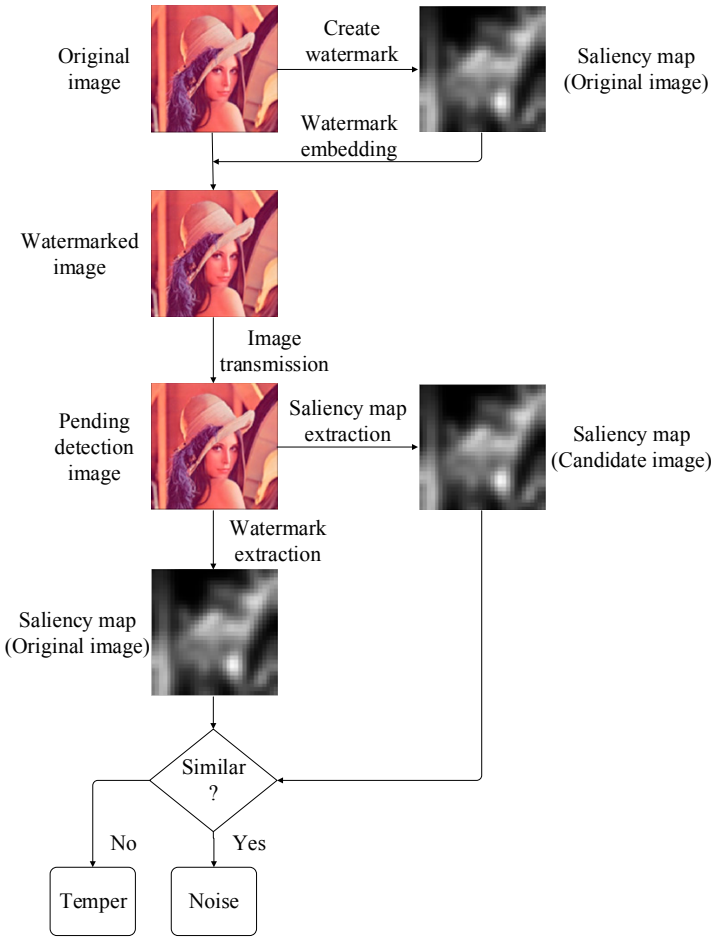

Figure 1. The simple process based on the proposed method.

The main contributions of this study are as follows:

1. A new watermarking model based on the local energy and maximum entropy model is proposed. The watermarking itself contains relevant information to the original image. Unlike previous digital watermarking algorithms, this watermarking is embedded with high redundancy, which ensures that the watermarking remains high robustness.
2. To increase the accuracy of image tampering detection, the difference algorithm based on the Lorentz curve, combined with information entropy, is adopted to verify the uniformity of the pixel distribution.
3. An all-connected graph pixel average distance algorithm based on the maximum entropy model is proposed. The distribution uniformity of the pixel coordinates is evaluated through the probability distribution function to obtain the degree of deviation between the average distance and the maximum likelihood estimate of the pixels, which can further improve the effect of image tampering detection.

The detailed flowchart based on the proposed method is shown in Figure 2.

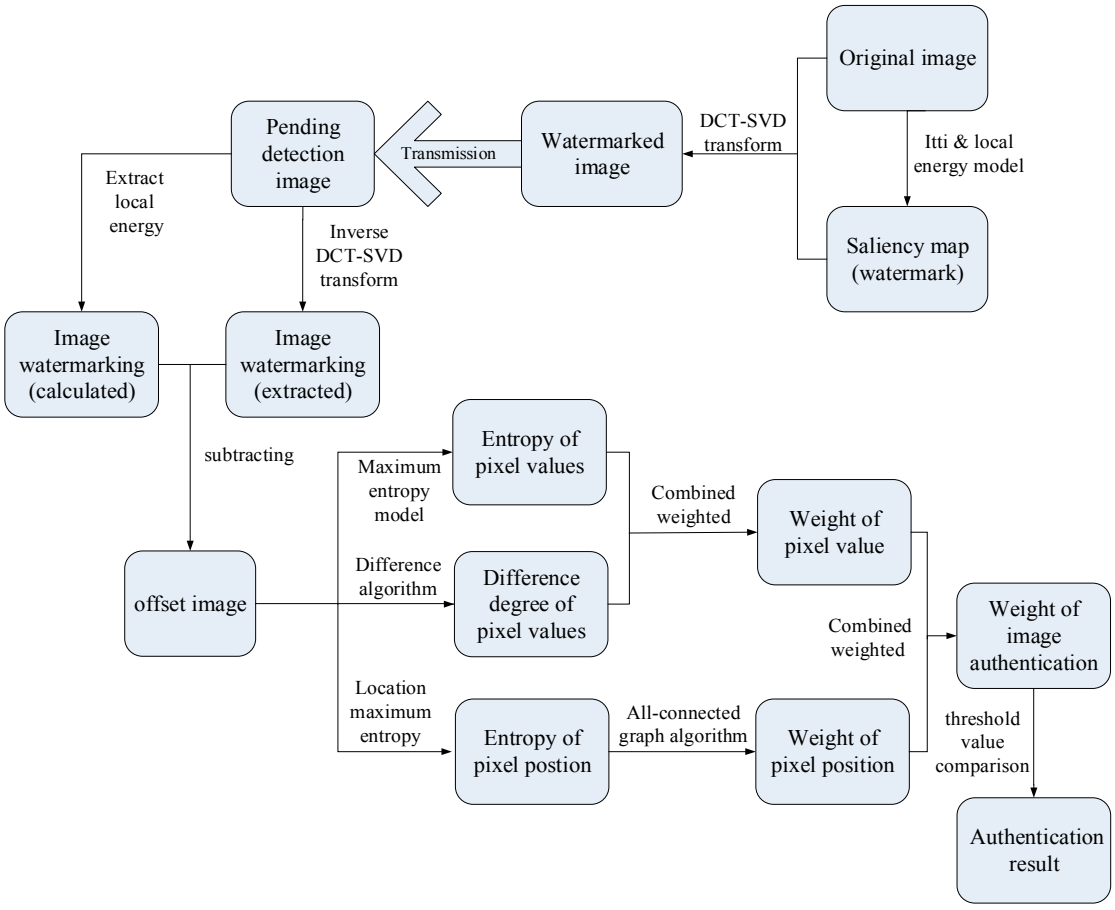

Figure 2. The flowchart of the proposed method based on the maximum entropy model.

2. Materials and Methods

2.1. Saliency Map Based on Local Energy Channel

In different areas of the image, the distribution of the information varies. Thus, the proposed method could extract the saliency map based on the distribution of the information of image distribution. A similar known method is the Itti model, which uses feature integration theory.

This method utilizes center-around mechanism, low-level features and normalization method, to build the saliency map.

This study uses the Itti model and local-energy model to extract the image distribution features. Henriksson [17] found the phase congruency pixels have a tendency to cluster. Compared with phase congruency, the local energy can obviously reflect the whole image features. It could also compensate for the limitations of phase congruency, which has relatively sparse information. The local energy model is structured by the odd and even of the Gabor function. Venkatesh and Owens [18] provided the plural forms of local energy:

$$E(x) = \sqrt{I^2(x) + H^2(x)} \tag{1}$$

where $I^2(x)$ indicates the component of the real part and $H^2(x)$ represents the component of the imaginary part. There is a kind of relationship between phase congruency and local energy:

$$E(x) = PC(x)\sum_n A_n \tag{2}$$

where $E(x)$ represents of local energy, $PC(x)$ represents the phase congruency and A_n represents local range. The phase is at every position of the image. A measurement method of the phase similarity in each frequency component, the phase congruency $PC(x)$ can be represented as follows:

$$PC(x) = max\frac{\sum_n A_n \cos(\theta_n(x) - \bar{\theta}(x))}{\sum_n A_n} \quad (\bar{\theta}(x) \in [0, 2\pi] \tag{3}$$

Here, $\theta_n(x)$ represents the phase of the n frequency component and $\bar{\theta}(x)$ represents the superimposed phase of all frequency components. The original image's feature information is obtained by the color, direction, intensity and the local energy channel. Then different feature information is combined to produce saliency map. The process of generating the saliency map is shown in Figure 3.

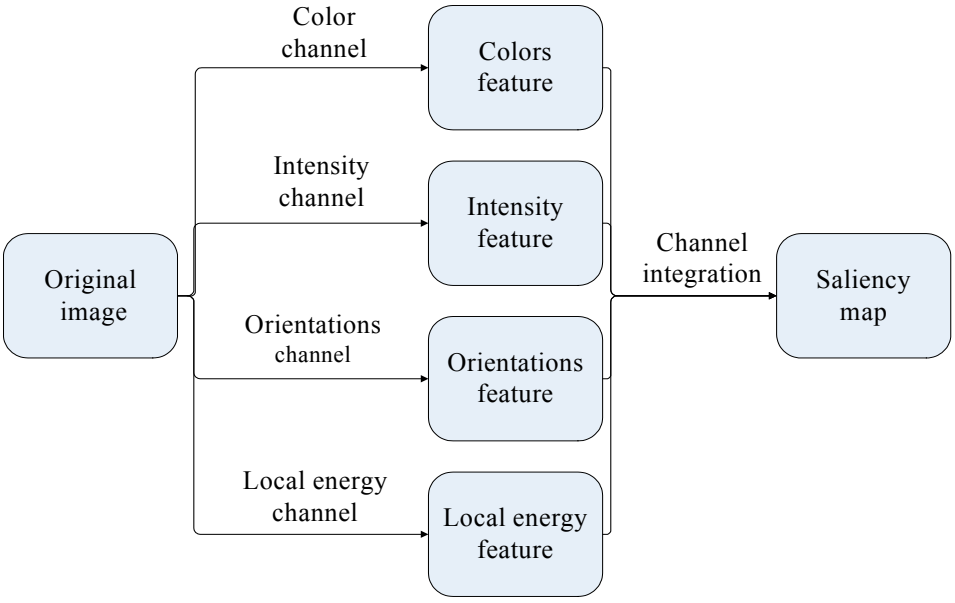

Figure 3. The process of generating saliency map.

Compared with the Itti model, the local energy model combines the local energy features to calculate the saliency degree of each pixel and express it as a grayscale image which reflects the image

information distribution. The image watermark obtained by the variable-size block model is used to contract the grayscale image. In our proposed method we generate 8×8 blocks to get watermarking.

2.2. Watermark Embedding

At the sender's terminal, the DCT-SVD of the original image is computed and the watermark is added to the low frequency component of the image. Here we use hamming error-correcting code to generate the watermarking checking information, hamming check bit k and data bit m has the following expression:

$$2^k \geqslant m + k + 1 \tag{4}$$

Then the robustness of the watermark has observably increased.

2.3. Image Preprocessing

At the receiver's terminal, the inverse DCT-SVD transformation is applied to extract the watermarking in the watermarked image. After the embedded watermarking is obtained, the saliency map of the received image can also be extracted by local energy model algorithm. The offset image is obtained by subtracting the original image saliency map from the received image.

2.4. Maximum Entropy Value Based on Image Pixel Value Distribution

Given the strong direction and purpose, the factitious tampered image parts will inevitably cause significant changes to the image information, whereas the rest of the area of the image information will remain unchanged. This can result in higher pixel values of the tampered area than those in the untampered area (nearly zero). If a pixel value distribution histogram is built, obvious differences will appear in the tampered area, but for the noise-interfered area, no obvious regularity will be observed. Therefore, the pixel value distribution entropy of the offset image is as follows:

$$H_{PV} = -\sum_{i=0}^{n} p(i) \log p(i) \tag{5}$$

where n stands for the grayscale number, $p(i)$ is the probability of the ith grayscale, and H_{PV} is the pixel value distribution entropy of the whole image. Compared with the factitious tampered image, the noise-interfered image presents more chaos. Thus, its entropy is higher. In other words, compared with the results of the noise interference, H_{PV} of factitious tampered image area will be significantly lower, the H_{PV} value as shown in Figure 4.

Figure 4. Maximum entropy results between image noise and factitious tampering (the ratio of the changed region area is 1%).

The exact changing degree of factitious tampering and noise interference is uncertain. If a slight change or noise happens in an image, the offset image is generally near 0. Thus, even if the distribution of the non-zero pixel in the offset image is uniform, the entropy of the whole offset image is not high. However, the offset image has numerous non-zero pixels if an image is considerably

changed, because of its substantial differences from the original, which will increase its whole information entropy. Therefore, the pixel values based on entropy distribution exhibit a certain degree of sensitivity according to how much the image is changed. To compensate, a difference algorithm is utilized based on the Lorenz curve to measure the difference degree of the pixel value distribution.

The main function of the Lorenz curve, which is widely used in economics, is to calculate the Gini coefficient that could measure the gap of national population between the rich and the poor. In this study, the Lorenz curve is utilized to calculate the difference degree between pixel values. First, the offset image is sorted by pixel values in ascending order. Second, the number of each pixel value V and the sum of all pixel values N are obtained. For the number of the ith pixel, $y(i)$ means number of pixels no less than i, and $x(i)$ means the sum of all pixel values less than i. If $y(i)/V$ is set as the ordinate and $x(i)/N$ as the abscissa, the Lorenz curve is obtained, as shown in Figure 5.

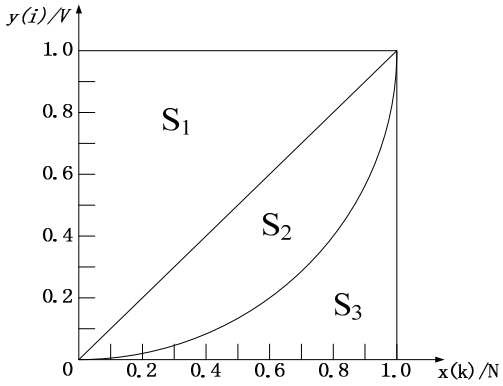

Figure 5. Lorenz curve diagram.

Figure 5 shows that if a great difference exists between the image pixel values, the area of the S3 is smaller. If not, the area of the S3 is larger. A difference parameter φ is used to describe the difference degree as follows:

$$\varphi = S_2/(S_3 + S_2) \tag{6}$$

If the value φ is higher, the difference degree is greater, thus, the image is more likely to be tampered by an attacker, as shown in Figure 6.

Compared with the entropy model, the sensitivity of difference parameter significantly decreases, as shown in Figure 7.

Figure 6. Comparison of the parameter φ between noise interference and factitious tampering.

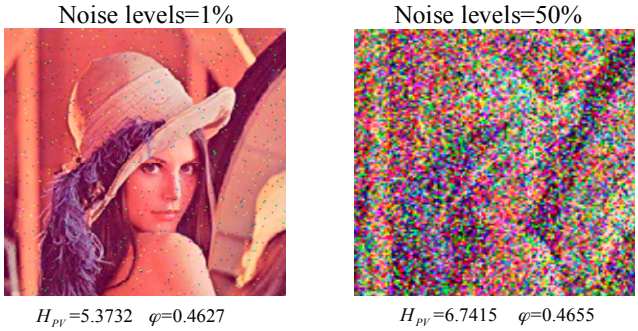

Figure 7. H_{PV} and φ in different levels of random noise interference.

By utilizing H_{PV} and ϕ, a combined weighted factor can be described as follows:

$$\phi = \frac{\varphi}{H_{PV}} \times \gamma \tag{7}$$

where γ is the compensation factor, and in this study, γ = 3.

If the entropy of image pixel value distribution and pixel value difference are combined through the Lorenz curve, ϕ can effectively distinguish the change of image information between irregular noise interference and regular factitious tampering.

2.5. Image Pixel Position Weight Based on Maximum Entropy

The offset images after factitious tampering and noise interference not only show great differences in pixel values, but also exhibit otherness in pixel location distribution, because in the factitious tampered image, changes often happen in one or a few concentrated areas, and other areas do not show any changes. Thus, the offset image would show the highest value pixels in one area. Meanwhile, the changes for the image caused by noise interference will be distributed in various parts. Therefore, noise interference will not cause regional aggregation, and on the contrary, its distribution will show randomness. Based on these reasons, an algorithm of distance difference between the pixel distributions and the maximum entropy model is proposed. An all-connected graph to describe the aggregation degree in the offset image is investigated in this study.

The offset image entropy is calculated based on the maximum entropy criterion. The received image is assumed to be composed of changed (potential tampered) and unchanged regions. The entropy means the degree of pixel value similarity in each region. If the entropy of a region is large, it indicates the differences of pixel values distribution are not significant. On the contrary, when the entropy of a region is small, it indicates that there are significant differences in the pixel values distribution. When the entropies of the changed and unchanged regions are both large, the entropy of the whole image would be the maximum. The gray level that maximizes entropy is the threshold to segment the image into changed and unchanged regions. Given that the image whose size is $M \times N$, the image gray level collection is $G = \{0, 1, \cdots, i, \cdots, L-1\}$, and n_i is the occurrence time of the ith gray value. If p_i is the probability of the i-th gray value occurrence, then p_i is as follows:

$$p_i = \frac{n_i}{M * N} \tag{8}$$

The probability of changed and unchanged regions conforms to the following conditions, here t is the threshold to segment the image into changed region and unchanged region:

$$\begin{aligned} p_{Changed}(t) &= \sum_{i=0}^{t} p_i \\ p_{Unchanged}(t) &= \sum_{i=t+1}^{L-1} p_i \end{aligned} \tag{9}$$

Thus, the entropy $H(t)$ of the two regions could be expressed as follows:

$$\begin{aligned} H_{Changed}(t) &= -\sum_{i=0}^{t} \frac{p_i}{p_{Changed}(t)} \log \frac{p_i}{p_{Changed}(t)} \\ H_{Unchanged}(t) &= -\sum_{i=t+1}^{L-1} \frac{p_i}{p_{Unchanged}(t)} \log \frac{p_i}{p_{Unchanged}(t)} \end{aligned} \tag{10}$$

The image of general entropy is as follows:

$$H_{total}(t) = H_{Changed}(t) + H_{Unchanged}(t) \tag{11}$$

When the entropy $H_{total}(t)$ is maximum, it means the differences in each internal region are the smallest, but the differences between changed region and unchanged region is the biggest. Then t is the optimal threshold to segment the whole image into changed (potentially tampered) and unchanged regions. If the number of the pixels in potential tampering region is assumed as m, the 2-D coordinate distribution of these pixels can be obtained. The triangular adjacency matrix A of the all-connected graph is established based on the pixel coordinates:

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1m} \\ 0 & \ddots & a_{23} & \ddots & a_{2m} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & 0 & \cdots & \ddots & a_{m-1,m} \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} \tag{12}$$

In this matrix, a_{ij} $(i < j)$ means the Manhattan distance d between pixels i and j in potential tampering region. Assuming that $\Delta\, x$ means the abscissa difference between the coordinates of pixel i and pixel j, Δy means the ordinate difference, and Manhattan distance d is defined as follows:

$$d = ||\,\Delta x + \Delta y\,||_1 \tag{13}$$

When all the elements in the triangular adjacency matrix are summed, the sum value of Manhattan distance S is obtained in potential tampering pixels. The average distance value $\overline{d}$ is computed by S and m:

$$\overline{d} = \frac{S}{C_m^2} \tag{14}$$

Here $C_m^2 = m(m-1)/2$, it means the number of all connection lines for every two pixels in potential tampering region. Given the strong pertinence of a factitious tampered image, its potential tampered points are clustered, and the average distance value $\overline{d}$ is small. Thus, if the value of $\overline{d}$ is small, the probability of the image being changed by artificial tampering is high. By contrast, if the value of $\overline{d}$ in potential tampering region is large, this indicates the distribution of potential tampering points is scattered. Thus the image is more likely changed by noise interference.

If the pixels in potential tampering region completely follow a random distribution, the pixel coordinates follow the 2-D uniform distribution:

$$f(x,y) = \begin{cases} \dfrac{1}{D}, & (x,y) \in D \\ 0, & (x,y) \notin D \end{cases} \tag{15}$$

For the regular rectangular image, the abscissa and the ordinate are linearly independent. Thus, the 2-D uniform distribution can be expressed by multiplying two 1-D uniform distributions:

$$P(XY) = P(X)P(Y) \tag{16}$$

The Manhattan distance d of two random pixels is as follows:

$$d = d_1 + d_2 \tag{17}$$

where d_1 is the absolute value of abscissa difference between two pixels, and d_2 is the absolute value of ordinate difference between two pixels. Thus, for the two pixels randomly distributed in the region, the random probability density function could be expressed as follows:

$$f(x) = \frac{2(\alpha - x)}{\alpha^2} \quad f(y) = \frac{2(\beta - y)}{\beta^2} \tag{18}$$

where α and β are the length and width of images, respectively. Then based on probability theory, the following equation is obtained:

$$\begin{aligned} E(X) &= \frac{\alpha}{3} & E(Y) &= \frac{\beta}{3} \\ D(X) &= \frac{\alpha^2}{18} & D(Y) &= \frac{\beta^2}{18} \\ E(X+Y) &= E(X) + E(Y) = \frac{\alpha + \beta}{3} \end{aligned} \tag{19}$$

Given the linear independence between $P(X)$ and $P(Y)$, the following equation could also be obtained:

$$D(X+Y) = D(X) + D(Y) + 2cov(XY) = D(X) + D(Y) = \frac{\alpha^2 + \beta^2}{18} \tag{20}$$

The expectation of the Manhattan distance between two random pixels in the offset image is $(\alpha + \beta)/3$, and the variance is $(\alpha^2 + \beta^2)/18$. When the number of random pixels m is large enough ($m \geqslant 10$), the number of lines in the all-connected graph C_m^2 is also very large. The length of any lines in the all-connected graph d_i can hardly affect the average Manhattan distance $\overline{d}$, and the length between any two lines almost have no correlation. Therefore, based on central limit theorem, the average Manhattan distance $\overline{d}$ nearly follows a Gaussian distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(x-\mu)^2}{2\sigma^2}] \tag{21}$$

In addition, the expectation of Gaussian distribution $\mu = (\alpha + \beta)/3$, and the variance $\sigma^2 = C_m^2 \times (\alpha^2 + \beta^2)/18$, based on the maximum likelihood probability. When the average Manhattan distance $P(x \notin \mu \pm 3\sigma) \approx 0.003$ of potential tampering pixels is $(\alpha + \beta)/3$, the distribution of image pixel is extremely random, simultaneously we can obtain the maximum entropy.

Compared with the potential tampering pixels caused by noise interference, the potential tampered pixels caused by factitious tampering have strong aggregation. The average Manhattan distance $\overline{d}_n$ of the all-connected graph caused by factitious tampering is much lower than that caused by noise interference.

Given that Gaussian distribution outside of $x = \mu \pm 3\sigma$ region has a low probability $P(x \notin \mu \pm 3\sigma) \approx 0.003$), the deviation degree parameter div is introduced:

$$div = \frac{\left|\mu - \overline{d}\right|}{\sigma} \tag{22}$$

As a deviation degree parameter based on maximum entropy of image pixel position distribution, *div* not only reflects the information of pixel position distribution difference, but also shows satisfactory discrimination between factitious tampering and noise interference, as shown in Figure 8.

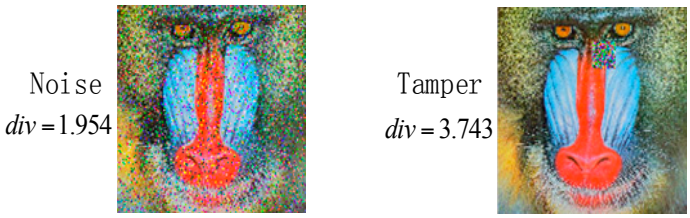

Figure 8. Compared deviation degree parameter $\varepsilon = \sqrt{\phi^2 + div^2}$ with factitious tampering and noise interference.

2.6. Image Tampering Detection Based on Combined Weighted Threshold

ϕ and *div* are applied as a joint weighting factor. An exception should be considered, that is, if only one weight is overly large, the possibility of factitious tampering is much larger than that in two weights that are both moderately large. Therefore, this study adopted two-norm as the final weight function:

$$\varepsilon = \sqrt{\phi^2 \cdot div^2} \tag{23}$$

Comparing ε with the pre-set threshold helps verifying whether the image has been factitious tampered under high noisy environment or not.

3. Experimental Verification

3.1. Watermark Robustness Testing

To test the validity of the method proposed in this study, the algorithm is simulated in MATLAB. In the watermarking embedding stage, the local energy image that contains information of the image itself as watermarking is embedded into the image spatial domain by employing the DCT-SVD transformation. The peak signal-to-noise ratio (PSNR) is introduced to determine the image distortion degree. PSNR is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Given that the range of signal fluctuation is usually wide, PSNR is normally expressed in decibels (dB). Its formula is as follows:

$$PSNR = 10\log_{10}(\frac{255^2}{MSE}) \tag{24}$$

The MSE is the mean square error between the two images, such as two RGB images I and K whose sizes are both $m \times n$. If the noise in the two images is approximately similar, their mean square error is defined as follows:

$$MSE = \frac{1}{3mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{3} || I(i,j,k) - K(i,j,k) ||_2 \tag{25}$$

The abovementioned formula shows that the larger the image changes caused by noise interference, the bigger the MSE value, and the smaller the PSNR value. Based on this principle, the change of embedded watermarking in the original image is measured, as shown in Figure 9 below.

The watermark computed by DCT-SVD and hamming error-correcting is robust to noise interference and factitious tampering. The watermark can be almost fully tested under different sorts

of image changes. The image transmission is simulated in a noisy channel using many sorts of noise interference and factitious tampering. The results are shown in Figure 10.

Figure 9. peak signal-to-noise ratio (PSNR) for original image and watermarking image.

Figure 10. Watermarking extraction effect of noise interference and factitious tampering.

Figure 10 shows that even in high degree of noise environments, the extracted watermarking has excellent robustness. This r can guarantee that the watermarking is tamper-proof. Compared with the traditional robust watermarking algorithm, watermarking generated by DCT-SVD and hamming error-correction has outstanding performance. Here we use the parameter normalized correlation *NC* (normalized correlation) to measure the similarity between two images. For two images y_1 and y_2, their *NC* values are defined as follows:

$$NC(y_1, y_2) = \frac{y_1^T y_2}{||y_1|| \cdot ||y_2||} \tag{26}$$

When the value of *NC* is closer to 1, the two images is more similar. Table 1 shows that compared with other robust watermarking algorithms based on DCT [19,20], our method has better robustness.

Table 1. The NC value of different robust watermarking algorithm.

Various Noise	NC Value (Our Method)	NC Value (Method [19])	NC Value (Method [20])
Contrast adjustment	0.9912	0.9558	0.9338
Average filter (size = [3, 3])	0.9937	0.9914	0.9909
Poisson noise	0.9884	0.9811	0.9754
Salt and pepper (noise density = 0.02)	0.9937	0.9295	0.9289
Salt and pepper (noise density = 0.05)	0.9818	0.8612	0.8420
Gaussian noise (noise density = 0.02)	0.9688	0.8937	0.8116
Gaussian noise (noise density = 0.05)	0.9493	0.8923	0.6814

Table 1 shows that in normal noise environments, the watermark obtained by our method is more robust than state-of-the-art alternatives. This result can further ensure the veracity of the extracted offset information image.

3.2. The Image Tampering Detection Performance Testing

In the image tampering detection stage, image distortion caused by noise interference can be distinguished from that caused by factitious tampering through setting a tolerance threshold in advance. The final weight parameter ε has a satisfactory capability of distinguishing the two different causes, as shown in Figure 11.

Figure 11 shows that during the noise growth process, the value of the weighted function displays no obvious change. Instead, it remains steadily around 50. However, for the bottom of two factitious tampered pictures, the value of the weighted function obviously increased. Therefore, image distortion caused by noise interference can be distinguished from that caused by factitious tampering by setting a reasonable threshold. The results show a satisfactory distinctiveness, and the image tampering detection algorithm is robust to noise interference. As the most three classic test images in image watermarking(Lena, Baboon and Plane), the value of ε are shown in Table 2.

As shown in Table 2, compared with noise interference, the value of ε by factitious tampering like clipping or replacing is much higher, so it is intuitional to distinguish noise from factitious tampering by choosing a reasonable threshold (in this experiments, the threshold can be set as 50).

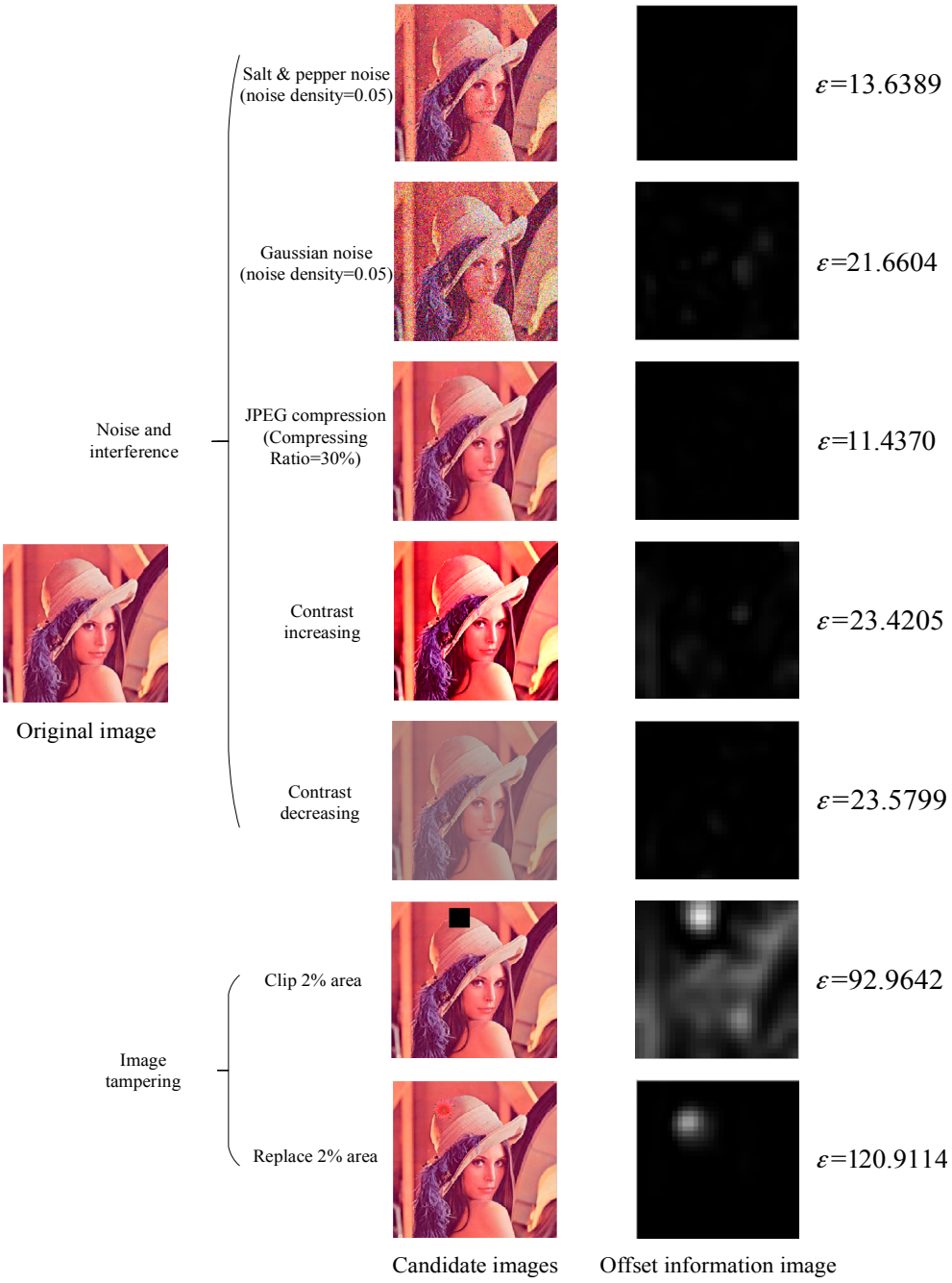

Figure 11. Combined weights ε of the images after different kinds of attack.

Table 2. The value ε of test image in different sort of image changing.

Image Changing	ε Value (Lena)	ε Value (Baboon)	ε Value (Plane)
Contrast increasing	23.4205	21.4458	19.5439
Contrast decreasing	23.5799	22.0004	20.0933
Average filter	11.7935	10.9523	11.4423
Image compression	11.4370	11.9632	12.0045
Salt and pepper noise	13.6389	14.6534	14.6238
Gaussian noise	21.6604	22.0954	20.9625
Clipping	92.9642	89.0045	101.4322
Replacing	120.9114	115.4493	109.4830

We also tested the performance of our proposed algorithm and some other robust watermarking algorithms [21–23] to compare their capabilities of distinguishing between noise (first four lines) and factitious tampering (last two lines). The results are shown in Table 3, where Y means the image identified as being changed by factitious tampering and N means the image identified as just being changed by noise.

Table 3. The result of noise and factitious tampering.

Image Changing	Result of Image Tampering Detection			
	Our Method	**Method [21]**	**Method [22]**	**Method [23]**
JPEG compress (30)	N	N	N	N
Salt and pepper noise (10%)	N	N	N	Y
Gaussian noise (10%)	N	Y	Y	Y
Average filter (3×3)	N	N	Y	Y
Clipping (2%)	Y	Y	Y	Y
Replacing (2%)	Y	Y	Y	Y

As Table 3 shows, under high intensity noise environment conditions, some other robust watermarking algorithms make wrong judgments, but our algorithm can accurately distinguish noise from factitious tampering.

We define M as the pixel number ratio of the factitious tampered pixels to all changed pixels in pending detection image as follows:

$$M = \frac{M_{\text{tamper}}}{M_{image}} \tag{27}$$

Here M_{image} is identified as the total image area and M_{tamper} is identified as the size of the tampering area in pending detection image, the definition M_{tamper} is as follows:

$$M_{\text{tamper}} \in \left(\frac{W_{offset}}{\min(W_{original}, W_{extracted})} \geqslant \eta\right) \tag{28}$$

Here the W_{offset} is represented as the pixel value of the offset saliency map, the W_{original} is represented as the pixel value of the original image's saliency map, and the $W_{\text{extracted}}$ is represented as the pixel value of the pending detection image's saliency map. η is the threshold, and in this study, $\eta = 5$. The value of M is set to 0's when the number of factitious tampered pixels is 0. Here, others used the methods similar to our M's which is also to measure the proportion of the image that is a malicious tampering area. This method keeps the clustered changed pixels intact and the isolated changed pixels disappear. As a result, the malicious attack in concentrated area leads to a larger M value due to the removal of mildly distorted changed pixels. Compared with Qi's method [24], Xiao's method [25] and Yang's method [26], extensive experiments show that the M value fits well on all 200 test images in five kinds of noise. The result is shown in Figure 12.

Figure 12 clearly indicates that all the average values of our M's are significantly lower than Qi's and Yang's for five kinds of noise, indicating that our method is more robust in classifying a watermarked image under any of these image changes such as noise interference or malicious attack. This is due to the fact our method introduces the offset saliency map algorithm, that is the offset image between the two saliency map of the original image and pending detection image. Because the saliency map based on image information has high robustness against operations which do not change image content such as image processing or noise interference, so the offset saliency map has good invariance.

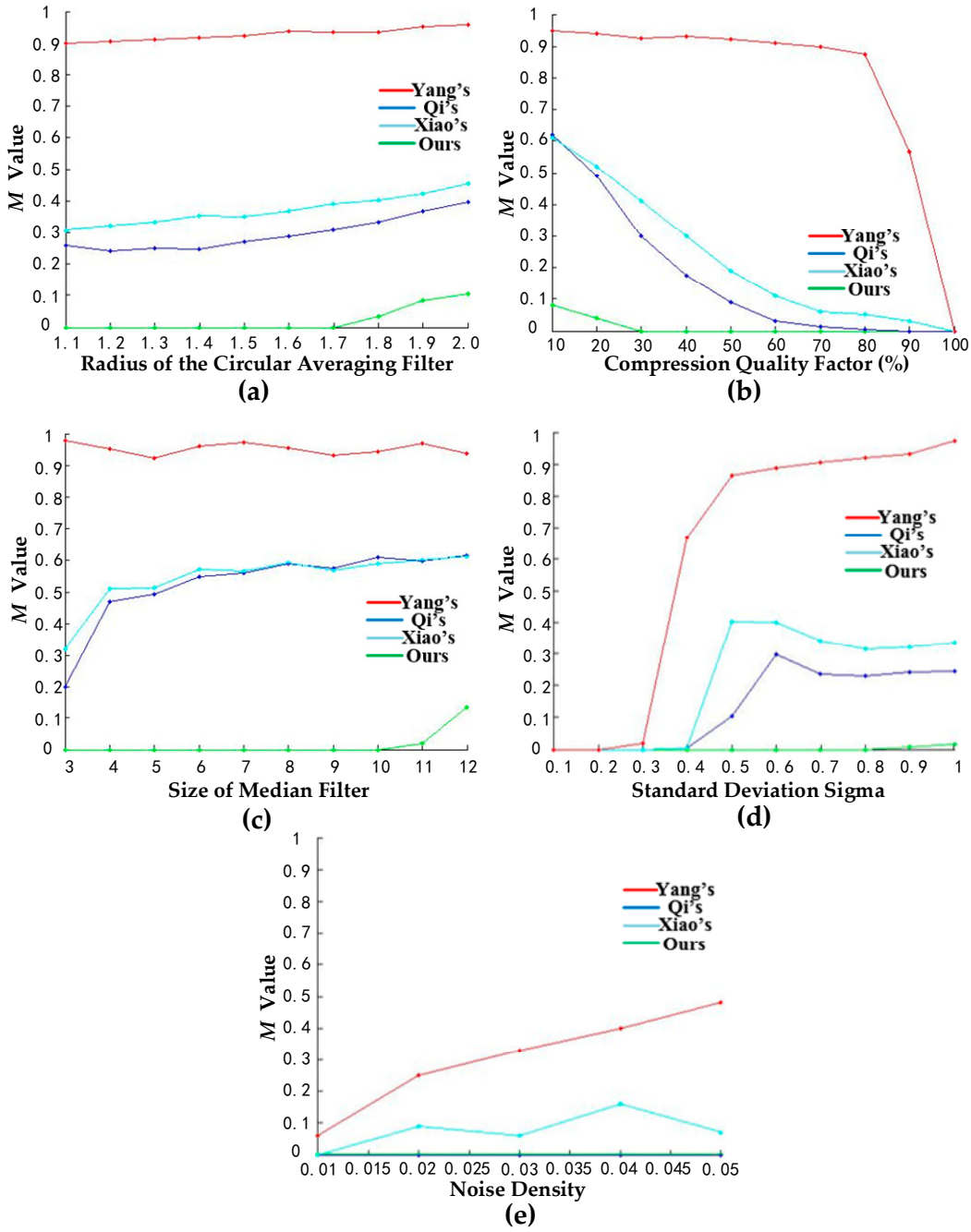

Figure 12. Comparison of various common image processing attacks on M value. (**a**) Image blurring attacks.(**b**) JEPG compression attacks. (**c**) Median filtering attacks. (**d**) Gaussian low-pass filtering attacks. (**e**) Salt and peppers noise attacks.

In order to test our method in more detail we performed four kinds of representative image processing attacks on 200 watermarked images of 512 × 512 size. These attacks included Gaussian low-pass filtering attacks using rotationally symmetric Gaussian low-pass filters of size 3 × 3 and standard deviation ranging from 0.1 to 1; median filtering attacks using filters of radii of 3–12; salt and peppers noise attacks using noise intensity ranging from 0.01 to 0.05; JPEG compression attacks using quality factors of 100% down to 10%. We also performed each of the three different pixel size replacements (e.g., 24 × 24, 36 × 36, and 48 × 48 pixels replacements), and we then we compared the performance of our scheme and five peer schemes: namely, Qi *et al.*'s method,

Maeno *et al.*'s method [27], Yang and Sun's method, Che *et al.*'s method [28] and Cruz *et al.*'s method [29]. The result are summarized in Table 4.

Table 4. Detection results under each simulated attack for 200 watermarked images.

Method	Actual Noise Interference				Actual Artificial Tampering		
	Probability of False Alarm (%)				Probability of Miss (%)		
	Gaussian	Median	S & P	JPEG	24 × 24	36 × 36	48 × 48
Ours	0	17.8	0.8	0.4	32.4	27.2	13.2
Qi	0	39.6	0	9.8	59.2	31.6	13.8
Yang	100	76.4	94.7	8.4	71.2	66.8	59.6
Che	100	64.3	94.5	66.6	100	40.6	40.4
Maeno	100	100	100	22.4	100	100	100
Cruz	8.5	88.6	40.3	24.3	100	100	100

It can be learned from the above table that our method has the lowest probability of false alarm in actual noise interference and has the lowest miss probability in actual artificial tampering among the six kinds of method. That means our proposed method obtains the most excellent comprehensive performance both in the probability of false alarm for four kinds of actual noise interference and the probability of missing actual artificial tampering. This is due to the fact our saliency map based on image content is composed of four channel features in different scales. Conventional image noise can only change a single image feature, such as contrast, color, texture, *etc.*, so the effect of the saliency map is limited, but malicious tampering operations, such as cutting, replacing, *etc.*, will make comprehensive information changes in a local area. Therefore, the saliency map will show the strong changes in this local area.

4. Conclusions

This study proposes a robust image watermarking algorithm based on maximum entropy. The algorithm uses local energy information as watermarking to embed and extract watermarking. In the image tampering detection stage, both pixel value and pixel location weighted algorithms based on the maximum entropy model are utilized. The image tampering detection scheme is improved by utilizing the difference parameters based on the Lorenz curve and the divergence degree of the average Manhattan distance between pixels of an all-connected graph. In the experiments, the proposed algorithm shows excellent performance under high noise interference conditions. In future study, we will combine a stronger encryption algorithm, such as the asymmetric encryption algorithm or chaos encryption algorithm, to increase the difficulty of cracking our image watermarking.

Supplementary Materials: The matlab code of our proposed are available online at http://www.mdpi.com/1099-4300/17/12/7854/s1.

Acknowledgments: This work was supported in part by the grants of the National Natural Science Foundation of China (No. 6110115), the Natural Science Foundation of Jilin Province (No. 20150520063JH), Postdoctoral Science Foundation of China (No. 2015M571363). Anonymous reviewers for their insightful comments which have helped to improve the quality of this paper.

Author Contributions: Guihe Qin conceived the research subject of this paper. Bo Zhao carried out the method of the robust image watermarking and validated results, drafted the paper and final approved the version to be published. Pingping Liu revised the paper and directed this study. All authors have read and approved the final manuscript.

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

References

1. Celik, M.U.; Sharma, G.; Tekalp, A.M. Lossless watermarking for image authentication: A new framework and an implementation. *IEEE Trans. Image Process.* **2006**, *15*, 1042–1049. [CrossRef] [PubMed]

2. Akhaee, M.A.; Sahraeian, S.M.E.; Sankur, B.; Marvasti, F. Robust scaling-based image watermarking using maximum-likelihood decoder with optimum strength factor. *IEEE Trans. Multimed.* **2009**, *11*, 822–833. [CrossRef]
3. Agoyi, M.; Çelebi, E.; Anbarjafari, G. A watermarking algorithm based on chirp z-transform, discrete wavelet transform, and singular value decomposition. *Signal Image Video Process.* **2015**, *9*, 735–745. [CrossRef]
4. Yang, S.Y.; Lu, Z.D.; Zou, F.H. A Novel Semi-Fragile Watermarking Technique for Image Authentication. In Proceedings of the 2004 7th International Conference on Signal Processing (ICSP'04), Beijing, China, 31 August–4 September 2004.
5. Ho, C.K.; Li, C.-T. Semi-fragile Watermarking Scheme for Authentication of JPEG Images. In Proceedings of the International Conference on Information Technology: Coding and Computing, Las Vegas, NV, USA, 5–7 April 2004.
6. Lin, C.-Y.; Chang, S.-F. A robust image authentication method distinguishing JPEG compression from malicious manipulation. *IEEE Trans. Circuits Syst. Video Technol.* **2001**, *11*, 153–168.
7. Maeno, K.; Sun, Q.; Chang, S.F.; Suto, M. New semifragile image authentication watermarking techniques using random bias and nonuniform quantization. Available online: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=877919 (accessed on 27 November 2015).
8. Cox, I.J.; Kilian, J.; Leighton, F.T.; Shamoon, T. Secure spread spectrum watermarking for multimedia. *IEEE Trans. Image Process.* **1997**, *6*, 1673–1687. [CrossRef] [PubMed]
9. Kang, H.; Iwamura, K. Information hiding method using best DCT and wavelet coefficients and its watermark competition. *Entropy* **2015**, *17*, 1218–1235. [CrossRef]
10. Walton, S. Image authentication for a slippery new age. *Dr. Dobb's J.* **1995**, *20*, 18–27.
11. Eggers, J.J.; Girod, B. Blind watermarking applied to image authentication. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'01), Salt Lake City, UT, USA, 7–11 May 2001.
12. Lu, C.-S.; Liao, H.-Y.M. Multipurpose watermarking for image authentication and protection. *IEEE Trans. Image Process.* **2001**, *10*, 1579–1592. [PubMed]
13. Yu, G.-J.; Lu, C.-S.; Liao, H.-Y.M.; Sheu, J.-P. Mean quantization blind watermarking for image authentication. In Proceedings of the 2000 International Conference on Image Processing, Vancouver, BC, Canada, 10–13 September 2000.
14. Itti, L.; Koch, C.; Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 1254–1259. [CrossRef]
15. Liu, R.-Z.; Tan, T.-N. SVD based digital watermarking method. *Acta Electron. Sin.* **2001**, *29*, 168–171.
16. Zhao, H.-W.; Chen, X.; Liu, P.-P.; Geng, Q.-T. Adaptive segmentation for visual salient object. *Opt. Precis. Eng.* **2013**, *21*, 531–538. [CrossRef]
17. Henriksson, L.; Hyvärinen, A.; Vanni, S. Representation of cross-frequency spatial phase relationships in human visual cortex. *J. Neurosci.* **2009**, *29*, 14342–14351. [CrossRef] [PubMed]
18. Venkatesh, S.; Owens, R. An energy feature detection scheme. In Proceedings of the IEEE International Conference on Image Processing (ICIP'89), Singapore, Singapore, 5–8 September 1989.
19. Singh, A.K.; Dave, M.; Mohan, A. Hybrid technique for robust and imperceptible image watermarking in DWT–DCT–SVD domain. *Natl. Acad. Sci. Lett.* **2014**, *37*, 351–358. [CrossRef]
20. Singh, A.; Tayal, A. Choice of wavelet from wavelet families for DWT-DCT-SVD image watermarking. *Int. J. Comput. Appl.* **2012**, *48*, 9–14. [CrossRef]
21. Lei, X. Image spatial semi-fragile watermarking algorithm with high classification capability of attack types. *Comput. Sci.* **2010**, *2*, 075.
22. Zhao, Y.; Sun, X. A Semi-fragile Watermarking Algorithm Based on HVS Model and DWT. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Wuhan, China, 12–14 December 2008.
23. Seng, W.C.; Du, J.; Pham, B. Semi fragile watermark with self authentication and self recovery. *Malays. J. Comput. Sci.* **2009**, *22*, 64–84.
24. Qi, X.; Xin, X. A quantization-based semi-fragile watermarking scheme for image content authentication. *J. Vis. Commun. Image Represent.* **2011**, *22*, 187–200. [CrossRef]

25. Qi, X.; Xin, X. A singular-value-based semi-fragile watermarking scheme for image content authentication with tamper localization. *J. Vis. Commun. Image Represent.* **2015**, *30*, 312–327. [CrossRef]
26. Yang, H.; Sun, X. Semi-fragile watermarking for image authentication and tamper detection using HVS model. In Proceedings of the International Conference on Multimedia and Ubiquitous Engineering (MUE'07), Seoul, Korea, 26–28 April 2007.
27. Maeno, K.; Sun, Q.; Chang, S.-F.; Suto, M. New semi-fragile image authentication watermarking techniques using random bias and nonuniform quantization. *IEEE Trans. Multimed.* **2006**, *8*, 32–45. [CrossRef]
28. Che, S.-B.; Ma, B.; Che, Z.-G. Semi-fragile image watermarking algorithm based on visual features. In Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR'07), Beijing, China, 2–4 November 2007.
29. Cruz, C.; Reyes, R.; Nakano, M.; Perez, H. Image content authentication system based on semi-fragile watermarking. In Proceedings of the 51st Midwest Symposium on Circuits and Systems (MWSCAS 2008), Knoxville, TN, USA, 10–13 August 2008.