A new feature-based self-embedding image watermarking scheme

Yang Liu\(^1\), Zhikun Chen\(^{1,2,a}\), Haocui Liang\(^2\), Shaoyuan Fu\(^2\)

\(^1\)School of Computer Science, China University of Geosciences, Wuhan 430074, China;
\(^2\)Beibu Gulf Big Data Resources Utilisation Lab, Qinzhou University, Qinzhou 535000, China;
\(^3\)Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation, Qinzhou 535000, China.

Author: Yang Liu; email: 1491701161@qq.com; Author: Haocui Liang; email: lianghe8@163.com; Author: Shaoyuan Fu; email: hongqing_1196@qq.com.

\(^a\) Corresponding author: chzhikun@163.com

Abstract. The performance of digital watermarking algorithm is mainly manifested in invisibility and robustness. At present, many algorithms improve these two aspects by sacrificing embedding capacity for the reason that if watermark embedding amount is small, the watermark has little influence on the visuality of the cover image, and its robustness is also higher. Therefore, how to improve the anti-attack performance of the algorithm without reducing the amount of embedding has become a major research direction. This paper proposed a new robust watermarking scheme based on feature extraction techniques. Firstly, the SIFT (scale invariant feature transform) algorithm is used to extract the stable feature points from cover image to form embedding region. Then watermark is composed of the cover’s texture features obtained by LBP (local binary pattern) extracting method. Finally, embedding region is transformed by singular value decomposition (SVD) and the watermark is embedded into its singular value vector. The simulation results showed that the watermark embedding capacity of this algorithm is relatively high, and it has good resistance to common image processing and geometric transformation.

1. Introduction
Digital watermarking technology is mainly used for copyright protection. From the source of watermark, this technology can be roughly divided into carrier-independent algorithm and carrier-related algorithm. The main difference is whether the watermark information comes from the carrier. Researchers have found that using part of information of the carrier to generate watermark can improve the invisibility and robustness of the watermark algorithm [12]. So many papers related on that kind of algorithm appear successively.

He et al. [3] applied discrete wavelet transform (DWT) to generate watermark information from the host image, and then scrambled the watermark to embed it into lowest significant bits of the host. This algorithm chose low-frequency coefficients to generate low-frequency compressed matrix after the host image was transformed by one level DWT, and used the binary matrix to form the watermark. Chamlawi et al. [2] proposed a secure semi-fragile watermarking technique. The high-compression
version of the host image, used as watermark, was obtained by using JPEG coding technique. After host image was performed by integer wavelet transform, the watermark was embedded into the wavelet sub-band to achieve the dual purpose of image authentication and tamper recovery, which imposed little effect on the host image visually. Li et al [5] proposed a semi-fragile watermarking scheme, in which the host image was divided into many sub-blocks, and 5 bits of information was extracted from each sub-block to form the watermark information. Then the watermark was embedded into the intermediate frequency coefficients of the host. Liu et al. [6] proposed a color image authentication technology based on digital watermarking. Firstly, the bi-level moment-preserving technique [10] was used to generate the image feature information, and the double check strategy was used to obtain the authentication data from the cover image. Then the watermark was constructed by combining those feature information and authentication data. After carrier image was partitioned into non-overlapping sub-blocks, watermark was embedded in these blocks. Chunlei Li et al. [7] proposed a semi-fragile self-recovering watermarking algorithm based on packet quantization and double authentication. Firstly, the host image was divided into many non-overlapping blocks, and some bits were extracted from the sub-blocks by method of computing statistical moment to generate watermark information. After decomposing the sub blocks by two levels DWT, the watermark was embedded in the obtained sub-bands of those blocks. Chuan Qin et al. [9] used vector quantization and index sharing technology to restore image contents that may have been altered by attacks. First, a series of vector quantization index values were used to describe the important content of the original image concisely. Then binary numbers of those index values were expanded by a random binary matrix to generate reference bits, which were further processed by a hash function to form watermark. The watermark was embedded in the three least significant bit planes of the cover image.

These self-embedding watermarking schemes are not very robust. The reason is that embedding watermark into a certain kind of frequency domain coefficient of the carrier is only resistant to a part of attacks. Although watermark is generated from the carrier, the characteristics of the watermark produced by these schemes are not obvious, and the effect of improving the performance is not obvious, so they can only control the amount of watermark embedding to ensure the robustness of the algorithm. In view of this, this paper adopts a novel feature-based watermarking scheme to improve the robustness of the algorithm at the foundation of ensuring a relatively high embedding amount. The remaining organizational structure of this article is as follows: the second section introduces the feature extraction technology used in our scheme, and experimental processes are depicted in the third section. In the fourth section is the results and analysis of the experiment. The fifth section summarizes the whole article.

2. Preliminary

2.1 Local Binary Pattern LBP

Local binary pattern, or LBP, firstly developed by Ojala et al. [8], which is used to compare local features of images, is a visual operator used for classification in the field of computer vision. The flow of the algorithm is that, the image is first divided into small blocks, where each pixel is compared with its adjacent 8 pixels. The neighbor pixel value will be set to 0 when it is smaller, and it will be set to 1 when greater. Arrange the resulting 0 or 1 in either clockwise or counterclockwise directions into an 8-digit binary number (usually converted to a decimal for convenience). Then, the frequency histogram of each binary number is computed in that small block, and the histogram is selectively normalized. Finally, histogram of all small blocks is linked to constitute a feature vector of the whole image. The definition of generalized LBP can be formulated as:

\[ LBP(x_c, y_c) = \sum_{i=0}^{8} S(p_i - p_c)2^i \]

where \( p_c \) is the value in central pixel \((x_c, y_c)\) and \( p_i (i=0, 1, 2, ..., 7) \) refers to the value of corresponding neighbor pixel. \( S(x) \) is a sign function which is defined as:
Figure 1 shows the process of comparing a block pixel with its eight neighbor pixels. (a) represents a sub-block. The center pixel value 81 is viewed as a threshold. Pixel value 83 in the upper left corner of (a) is bigger than 81, so the binary number in corresponding position in (b) is 1.

![Figure 1. (a) 3×3 image block, (b) LBP.](image)

The obtained feature vector is closely linked to image’s texture characteristic, which are characterized by rotation invariance and grayscale invariance. Since its stability, LBP operator can be utilized to generate robust watermark from cover image.

2.2 SIFT

An import prerequisite for designing watermarking algorithm that can resist various attacks is to find an image description operator based on local region, which makes cover image still be able to use these robust region features for registration and correction, and find original embedded locations after impacted by image attacks. Scale invariant features transform (SIFT) [4] can be used to find a kind of scale invariant features descriptor in one image, maintaining invariance for rotation, scaling, brightness changes and visual angle change. To get the feature points, a signal’s scale space is defined first, and then the extreme points are detected as the feature points in the space.

The scale space of an image represents its feature set and contains a continuous scale parameters. Transformed on original image under different scales, and the multi-scale representation sequences of scale space are obtained. These sequences form the principal contour of that space. As a feature vector, the principal contour can detect edges, corners and extract features at different resolutions. Scale space kernel is described as

$$S(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$$

(2)

The convolution formula is as follows:

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) \times I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma)$$

(4)

As the formula above means, Gaussian difference images are obtained by subtracting the Gaussian image in each scale space, and then extreme points are detected in those difference images. Comparing the value of a pixel (x, y) in space D(x,y,\sigma) with pixel values of other 26 adjacent points, if the pixel (x,y) is the maximum or minimum value, the pixel is considered to be a grayscale extreme. Those 26 adjacent points include 8 adjacent points around the same layer difference image and 18 points in the same region 3 x 3 range in the upper and lower layers image, as shown in fig. 2. In this way, the extreme points detected are taken as the feature points, and information of location and scale of which is noted down.
3. Proposed scheme

This scheme involves three main steps: watermark generation, watermark embedding and watermark extraction. In the process of watermark generation, LBP technology is used to extract the texture features from the host image as watermark. The watermark embedding process is based on process of SIFT and Singular value decomposition (SVD). The stable feature points extracted by SIFT algorithm are used to locate the watermark again, so that the watermark is embedded in the robust position of the host, uniform and reliable, which eliminates the image instability when attacked. SVD [1] is a numerical analysis tool that implements diagonal form of matrix. From the point of view of linear algebra, a gray image is a discrete, non-negative array of matrices. After singular value decomposition, two orthogonal matrices and one diagonal matrix, singular value vector, are obtained. Singular vectors represent the structural information of an image. Compared with other watermark embedding methods, SVD can show better stability. When certain interference is applied to the image, the singular value of the diagonal line of the image fluctuates little. Although SIFT features can resist the attack of image geometric transformation, they are also sensitive to noise. SVD can help to make up that defect, because the singular value components obtained by SVD have strong anti-jamming ability to noise. In this way, the SIFT characteristic matrix is further decomposed by SVD and watermark is embedded into its singular value components, which further improves the robustness of the algorithm. The main flow of the scheme is that n points are obtained by LBP and form a feature matrix, viewed as watermark information. Then n SIFT feature points are extracted and their positions are marked. A SIFT matrix is composed of those n SIFT feature points with the same size as the watermark matrix. SVD is performed on SIFT matrix to embed the watermark into the obtained singular value vector. SIFT matrix in the cover image is replaced by the new SIFT matrix, and watermarked matrix is obtained.

3.1 Watermark embedding

Watermark information was generated by LBP operation and then was embedded into SIFT region of the cover image. The specific steps are as follow:

Step 1. The boundary of over image C was extend symmetrically so that the length of C can be divided by the length of LBP’s detection window with no remainder. A new image C_b was obtained after boundary extension.

Step 2. Then the C_b was partitioned into a lot of non-overlapping sub-blocks with size of 3×3, and the number of all blocks was denoted as n.

Step 3. The procedure of LBP was used in each sub-block to produce a string of 8-bit binary numbers, which was further converted into a decimal number by weighted-sum method to substitute the center pixel value of the cover. Watermark matrix W was composed of those newly formed pixel values. Let W_c be the element of W, and the range of c is [1, n]. W_c in the location of center pixel is obtained using weighted-sum method, which can be described as:

\[ w_c = x_0 + x_1 \cdot 2^1 + x_2 \cdot 2^2 + x_3 \cdot 2^3 + x_4 \cdot 2^4 + x_5 \cdot 2^5 + x_6 \cdot 2^6 + x_7 \cdot 2^7, \quad x_0, x_1, \ldots, x_7 \in \{0, 1\} \]  

(5)

Step 4. The scale and rotation invariant feature points were obtained by using SIFT algorithm. The feature points were extracted from the carrier image and a one-dimensional sequence was obtained, which was further constructed into a matrix C_SIFT of the same size as the watermark.

Step 5. SVD was performed on the matrix C_SIFT, and the orthogonal matrices U_{SIFT}, V_{SIFT} and a diagonal matrix S_{SIFT} were obtained respectively. The watermark W was embedded in S_{SIFT} to get S^*_{SIFT}: 

Figure 2. Extreme points in Gauss difference image
\[
U_{SIFT} \cdot S_{SIFT} \cdot V_{SIFT} = \text{SVD}(C_{SIFT}), \quad S'_{SIFT} = S_{SIFT} + \alpha \cdot W. \quad (6)
\]

Step 6. The inverse SVD transform was used to get the watermarked matrix \( C'_{SIFT} \).

\[
C'_{SIFT} = U_{SIFT} \cdot S'_{SIFT} \cdot V_{SIFT} \quad (7)
\]

Step 7. The watermarked matrix was performed by dimensionality reduction to replace the SIFT feature points of the original host image, and the watermarked image \( C_w \) was obtained.

3.2 Watermark extraction

In order to construct watermark information for copyright authentication, SIFT transform is first performed on the image to obtain the feature points which are invariant to rotation and scaling. Those feature points constitutes a feature matrix, which is further decomposed by SVD to get a diagonal matrix and watermark information is extracted from the diagonal matrix. The specific procedures are as follows:

Step 1. Applying the first three steps of watermark embedding process, and LBP characteristic matrix of the image \( W_{new} \) was obtained.

Step 2. According to the location information of the feature points, the corresponding SIFT feature points are obtained from the image, constructing a one-dimensional sequence.

Step 3. These points are further constructed into a matrix of the same size as the watermark, and the matrix was decomposed by SVD to generate the diagonal matrix \( S_{new} \).

Step 4. Watermark information \( W^*_{new} \) was extracted from \( S_{new} \):

\[
W^*_{new} = \frac{S_{new} - S_{SIFT}}{\alpha} \quad (8)
\]

Step 5. Comparing the \( W_{new} \) and \( W^*_{new} \), and the change of the watermark can be judged according to the comparison result, and the change of the cover can be further judged.

4. Experimental results

To quantify the watermarking property and to prove the advantage of less time consumption in the proposed scheme, an experiment was implemented on the platform of MATLAB 7.0 running on a PC with a CPU of Inter Core2 2.66 GHz and a memory chip of 4 GB. In this paper, the 512×512 Lena image was taken as the test object, as shown in figure Fig.5 (a). The watermark was generated from the Lena image and embedded in the SIFT feature region of the image. The watermark matrix was composed of 171×171 texture feature points from the carrier image by LBP method. The embedding intensity factor was set to 0.005. The watermark was embedded into the carrier by the embedding process described in 3.1. The obtained watermarked Lena was shown in Fig.5 (b).
The algorithm was tested through a series of attack experiments, and the results were shown in Table 1. It can be seen that the following six kinds of attacks have some influence on watermark images. After the image is attacked, there will be some distortion. Particularly in the case of histogram equalization and high noise intensity attacks, the PSNR value will be very low, below 30 db. When no attack occurs, the PSNR value exceeds 50dB and is in a relatively high state. This shows that the algorithm has good invisibility and cause little distortion to the carrier image.

In addition, the extracted watermark was compared with the original watermark to observe the similarity under various attacks. From the NCC values in the table 1, the similarity coefficients of the image under various attacks are relatively high, most of which are above 0.9, that is, the extracted watermark is almost identical with the original watermark. This shows that the algorithm is very robust under various attacks.

| Attack types        | Attack intensity | PSNR   | NCC   |
|---------------------|------------------|--------|-------|
| —                   | —                | 54.2701| 1     |
| Gauss noise         | 0.001            | 54.2701| 1     |
|                     | 0.005            | 46.5182| 0.9814|
|                     | 0.010            | 38.6094| 0.8961|
|                     | 0.050            | 26.8719| 0.8079|
| Salt and pepper noise| 0.005         | 32.1846| 0.9508|
|                     | 0.02             | 29.0891| 0.9037|
|                     | 0.1              | 26.6113| 0.7726|
| Median filtering    | 3×3              | 34.8901| 0.8762|
| Rotation            | 45°              | 24.7198| 0.9412|
| Scaling             | 1/2              | 31.0721| 0.9118|
| JPEG compression    | 80%              | 37.7205| 0.9471|
| Histogram equalization| —              | 23.5821| 0.9803|

Salt and pepper noise, also known as impulse noise, is a random white dot or black spot. This noise is very common and may occur in the process of image transmission, image cutting, image decoding and so on. Figure 6 presents results of salt and pepper noise attacks simulated on watermarking schemes proposed by Zhang, Wang and Zhou (2017) [13], Thakkar and Srivastava (2017) [11] and this paper. The noise intensity is respectively 0.002, 0.005, 0.02, 0.1. When the noise becomes stronger, the NCC value becomes smaller. Therefore, the noise has a certain effect on the embedded watermark.

In addition, the values on the red curve show that our scheme is better in resisting the noise, because most of the NCC values are larger than the other two schemes.
Figure 6. NCC after salt and pepper attack

Image compression is a common image processing method for saving the cost of network transmission, because it can reduce the volume of original image. Figure 7 shows the NCC values after the watermarked image was attacked with different degrees of JPEG compression. We tested five kinds of compression ratios, 30%, 50%, 60%, 70% and 80%, respectively. Take 80% for example, it means the tested image is compressed to ninety percent of its original one, and the compression degree of 70% is higher than that of 80%. Seeing from the fig. 7, the NCC values get smaller when the compression intensity increases, which signifies that the similarity between the original watermark and the extracted one gets smaller. Another point is that values in red curve are bigger than that in the other three curves mostly, illustrating that our algorithm is more robust to compression attacks.

Figure 7. NCC after JPEG compression

Changes of PSNR values in table 1 show that the watermarked image will be distorted to a certain extent after being attacked, especially histogram equalization and strong noise attack. However, in this case, the extracted watermark information is very close to the original watermark, indicating that the algorithm has good robustness against common image processing attacks. Watermarking schemes in Zhang, Wang and Zhou (2017) [13] and Thakkar and Srivastava (2017) [11] are typical domain algorithm and frequency algorithm, respectively. In the former scheme, feature information of the host image was obtained as watermark, and then was embedded into its least significant bit plane. This method is difficult to resist the influence of noise and image processing on watermark. In the latter, the host image was process by discrete wavelet transform, a commonly used frequency transform, and then the watermark was embedded into its low-frequency coefficients. The traditional frequency domain watermarking algorithm also has good robustness, but it sacrifices the watermark embedding capacity to a certain extent. In our scheme, using the texture features as watermark obtained from Lena image by LBP technology, the embedded data will hardly affect visual effect of the host image, implying good invisibility of our scheme. In addition, stable SIFT feature points are extracted as the embedding region, eliminating the image instability when attacked. Singular value vectors obtained by
SVD have strong anti-jamming ability, so watermark embedding in that vectors further strengthens the anti-noise ability of the algorithm.

5. Conclusion

This paper proposed a new robust watermarking scheme based on image features. In the scheme, the local binary pattern information is extracted from the host image as the watermark, and then some SIFT feature points of the host are obtained to form embedding region because those feature points have good invariance to geometric transformation attacks. Finally, watermark embedding is achieved through singular value decomposition and restoration process. The performance test results show that the algorithm can hide the copyright verification information well, and can extract a very complete watermark after the carrier is subjected to various common attacks, achieving the good effect of copyright protection. Our main contribution contains two points:

1) Robustness of watermark has been enhanced. SIFT has strong resistance to image geometric transformation, LBP has the property of rotation invariance and grayscale invariance, and SVD can resist noise interference. The combination of those advantages enables the watermarking algorithm to resist more attacks simultaneously, further improving its robustness.

2) The imperceptibility of the algorithm is improved without reducing the amount of embedding. The texture feature extracted from the carrier image not only contains a large amount of information, but also can not affect the visual effect of the carrier when used as watermark.

Acknowledgements

This research was financially supported by Qinzhou scientific research and technology development plan project (Grant No. 201714322).

References

[1] I. A. Ansari, M. Pant, C. W. Ahn, Robust and false positive free watermarking in IWT domain using SVD and ABC, Engineering Applications of Artificial Intelligence, 49, 114–125 (2016)
[2] R. Chamlawi, A. Khan, A. Idris, Wavelet Based Image Authentication and Recovery, Journal of Computer Science and Technology, 22(6), 795-804 (2007)
[3] H. He, J. Zhang, H. Tai, A Wavelet-Based Fragile Watermarking Scheme for Secure Image Authentication, International Workshop on Digital Watermarking, 422-432 (2006)
[4] D. G. Lowe, Distinctive image features from scale invariant key points, International Journal of Computer Vision, 1(5) (2004)
[5] C. Li, B. Ma, Y. Wang, D. Huang, Z. Zhang, A secure semi-fragile self-recoverable watermarking algorithm using group-based wavelet quantization, The 13rd pacific-rim conference on multimedia. Singapore,7674:327–336 (2012)
[6] K. Liu, Self-embedding watermarking scheme for colour images by bi-level moment-preserving technique, IET Image Processing, 8(6), 363–372 (2014)
[7] C. Li, A. Zhang, Z. Liu, L. Liao, D. Huang, Semi-fragile self-recoverable watermarking algorithm based on wavelet group quantization and double authentication, Multimed Tools Appl, 74:10581–10604 (2015)
[8] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no. 7, pp. 971-987 (2002)
[9] C. Qin, P. Ji, J. Wang, C. Chang, Fragile image watermarking scheme based on VQ index sharing and self-embedding, Multimedia Tools Applications, 76:2267–2287, doi 10.1007/s11042-015-3218-9 (2017)
[10]T. Wen-Hsiang, Moment-preserving thresholding: a new approach, Computer Vision Graph and Image Processing, 29, 377–393 (1985)
[11]F. N. Thakkar, V. K. Srivastava, A blind medical image watermarking: DWT-SVD based robust
and secure approach for telemedicine applications. Multimed Tools Appl, 76:3669–3697, doi: 10.1007/s11042-016-3928-7 (2017)

[12] L. Xie, G. R. Arce, A class of authentication digital watermarks for secure multimedia communication. IEEE Transactions on Image Processing, 10(11), 1754-1764 (2001)

[13] H. Zhang, C. Wang, X. Zhou, Fragile Watermarking for Image Authentication Using the Characteristic of SVD. Algorithms, 10, 27, doi:10.3390/a10010027 (2017)