An SVD-based Fragile Watermarking Scheme With Grouped Blocks

Qingbo Kang*, Ke Li†, Hu Chen‡
*Chengdu Yufei Information Engineering Co.,Ltd., Chengdu, China
Email: qdsclove@gmail.com
†National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu, China
Email: likeneill@gmail.com
‡National Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu, China
Email: huchen@scu.edu.cn

***Abstract*—In this paper, a novel fragile watermarking scheme for digital image authentication is proposed based on Singular Value Decomposition(SVD) and grouped blocks. The watermark bits which include two types of bits are inserted into the least significant bit(LSB) plane of the host image using the adaptive chaotic map to determine the positions. The groped blocks break the block-wise independence and therefore can withstand the Vector Quantization attack(VQ attack). The inserting positions are related to the statistical information of image block data, in order to increase the security and provide an auxiliary way to authenticate the image data. The effectiveness of the proposed scheme is checked through a series of attacks, and the experimental results demonstrate it achieves superior tamper detection and localization accuracy.**

Keywords—Image Authentication, Tamper Detection, Fragile Watermarking, Singular Value Decomposition

I. Introduction

With the tremendous development of information technology, especially in network communication and multimedia, digital images had a paramount role in our daily life. However, the digital images can be easily manipulated and tampered with the help of powerful image processing software. In fact, lots of people can easily manipulate images in such a way that may lead to human casualty or financial loss [1]. So preserving the authenticity and integrity of digital images has become a considerable aspect of many organizations [2], [3]. The authentication schemes can be divided into two categories: cryptography based schemes [4], [5] and fragile watermark based schemes [6]–[8]. Image authentication schemes based on cryptography compute a message authentication code (MAC) from images using a hash function, they can detect if an image has been modified, but they don't have the ability to locate the modified regions [9].

In a fragile watermark based scheme, the watermark is embedded into the host image that need to be protected. The watermark is often generated using either image features extracted from host image or the random values induced by the selected random number seed, when the image need to be authenticated, the watermark is extracted from the watermarked image to detect the tampered areas [10]. Walton proposed the first fragile watermark-based authentication schemes [11]. It calculated the check-sums of the seven most significant bits(MSB) of gray-scales along pseudo-random order in the least significant bits(LSB) of pixels. It only provides very limited tamper detection [12]. Yeung and Mintzer [6] proposed a fragile watermarking scheme that uses a pseudo-random sequence and a modified error diffusion method to embed a binary watermark into an image. However, with this method the watermark to be embedded has a certain degree of certainty, the attacker could deduce the binary look-up table and make the counterfeit image [13]. Holliman and Memon proved that schemes which are block-wise independent are vulnerable to vector quantization attack [14]. The Vector Quantization attack(VQ attack) means the counterfeit image can be reconstructed using a vector quantization code-book generated from a set of watermarked images, since each block is authenticated by itself, the counterfeit images appears authentic to the watermarking scheme. To withstand the VQ attack, researches proposed a number of schemes. [12], [15] proposed the fragile watermarking schemes that use the chaotic pattern to generate the difference image and then mapping it into a binary image eventually insert into the LSB bit-plane of the host image. This way, some tampered pixels can be identified due to the absence of watermark information carried by them. Since some information derived from new pixel value may coincide with the watermark, modification to these pixels cannot be detected directly. In this case, localization of the tampered pixels is not complete, and detection of the tampering pattern is inaccurate [16].

Singular value decomposition(SVD) is a kind of effective method of algebraic feature extraction. It can not only capture the basic structure of the data in the matrix, but also reflecting the algebraic essence of the matrix. These excellent features make it have a wide application in signal processing, image compression, pattern recognition and other fields. SVD also have a widely used in robust watermarking field [17]. In the method proposed by Sun et.al [18], SVD is performed in the spatial domain, and watermark is embedded by quantizing the largest SV of an image block. However, this method is vulnerable to VQ attack. In this paper, a novel SVD-based watermarking scheme for image authentication is proposed. The blocks of the host image are disturbed with the help of Arnold scrambling method. Then, all scrambled image blocks are divided into grouped blocks. For each block, two types of watermark bits are embedded, one for the block itself, another for the grouped blocks. An adaptive chaotic image pattern is generated using the logistic map for each block to determining the embedded position of the watermark bits. The use of the watermark bits of the grouped blocks is to break the block-wise independence and withstand the VQ attack.The adaptive chaotic image pattern is not only in order to increase the security of the scheme but also provides an auxiliary way to authenticate the image data.

II. Singular Value Decomposition and Chaotic Maps

A. Singular Value Decomposition

In linear algebra, the SVD is a factorization of a real or complex matrix. Formally, any real or complex $m \times n$ matrix M of rank r can be decomposed as

$$M = USV^T \tag{1}$$

where $U_{m\times m}$ and $V_{n\times n}$ are unitary matrices, and $S_{m\times n}$ is an $m \times n$ rectangular diagonal matrix. Moreover,

$$S_{m\times n} = \begin{bmatrix} \triangle_{r\times r} & 0 \\ 0 & 0 \end{bmatrix}$$

$$\triangle_{r\times r} = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$$

$$\sigma_i = \sqrt{\lambda_i}(i = 1, 2, \ldots, r, \ldots, n)$$

The diagonal entries σ_i of $\triangle_{r\times r}$ are known as the singular value of the matrix M. $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r \geq 0, \lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_n = 0$ are the eigenvalues of both M^TM and MM^T. Under the limitation of $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r$, the vector $(\sigma_1, \sigma_2, \ldots, \sigma_r)$ is uniquely. It characterizes the distribution and retains the algebraic essence of the matrix data. From this viewpoint, we can observe that a discrete image is an array of non-negative scalar entries which may be regarded as a matrix [19].

B. Chaotic Maps

In recent years, chaotic system and permutation transform have been used for digital watermarking, in order to reinforce the security [15]. Use the logistic map and Arnold scrambling to increase the security and performance of our scheme.

1) Logistic Map: Logistic map is one of the simplest and most transparent systems exhibiting order to chaos transition. Mathematically it is defined as:

$$x_{n+1} = \mu x_n(1 - x_n), n \in Z, u \in [0, 4], x_n \in (0, 1) \tag{2}$$

The μ here is a positive constant sometimes known as the ”biotic potential”, when $3.5699456 < \mu < 4$ the map is in the region of fully developed chaos [20]. That is, at this point, the sequence $x_k; k = 0, 1, 2, 3, \ldots$ generated by (2) is non-periodic, non-convergent and sensitive to the initial value.

2) Arnold Transform: Arnold transform has been widely used in the field of image encryption. The classical Arnold transform is a two-dimensional invertible chaotic map described by

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & a \\ b & ab+1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} (mod\ N) \tag{3}$$

where a and b are positive integers, N represents the size of the image matrix, it indicates the start point (x_0, y_0) after n iterations, then through modular arithmetic to get the coordinates of the final result. The (3) is also chaotic and area preserving. The period T depends on the parameters a, b and the size N of the image matrix. Thus parameters a, b and the number of iterations k, all can be used as secret keys.

III. The Fragile Watermarking Scheme

In the proposed scheme, the watermark bits to be hidden are made up of two parts: block authentication bits for authenticate the image block itself, and group authentication bits for authenticate the grouped blocks. Our scheme is a block-based scheme. In order to withstand the VQ attack, the grouped blocks is used to break the independence between the blocks. For each one image block, these two type of authentication bits are hidden in the LSB of the block pixels. The positions used for insertion are determined by the chaotic sequence, which depends on the statistical information of the block pixels. If the statistical information is changed, all the watermark bits can’t be correctly extracted from the image block. This is an auxiliary way to authenticate the image data. So the use of chaotic sequence not only increases the security, but also increases the creditability of the proposed scheme.

Figure 1 shows the block diagram of the embedding procedure, the more details is described in below.

A. Watermark Embedding Procedure

1) Block Division: Before generation and insertion of watermark bits, first divide the original image into blocks. Denote the original grayscale image as I, and its rows and columns is M_1 and M_2, and the total number of pixels is $M(M = M_1 \times M_2)$. Assuming that both M_1 and M_2 are multiples of four, first divide the original image into $M/16$ non-overlapped blocks sized 4×4, and denote the pixel-blocks as $B_{m,n}(m \in [1, M_1/4], n \in [1, M_2/4])$ and the gray values of pixels in a block as $b_{m,n}(i, j)(1 \leq i, j \leq 4)$.

2) Arnold Scrambling: This step the Arnold transform described in (3) is used to scramble the original image blocks $B_{m,n}$, the unit of scrambling is image block. The times of transformation is k, where k is the security number. Let’s denote the scrambled image as $ScrI$ and the blocks in the scrambled image as $ScrB_{m,n}(m \in [1, M_1/4], n \in [1, M_2/4])$, and the gray values of pixels in a block as $Scrb_{m,n}(i, j)(1 \leq i, j \leq 4)$.

3) Set LSB to Zero: For all image blocks in $ScrB_{m,n}$, Set LSB bitplane of pixels as zero:

$$Scrb_{m,n}(i, i) = Scrb_{m,n}(i, j) - (Scrb_{m,n}(i, j)\ mod\ 2).$$
$$(m \in [1, M_1/4], n \in [1, M_2/4]), (1 \leq i, j \leq 4).$$

4) SVD on The Image Blocks: Treat the image pixels in one block as a matrix, and perform SVD on the image blocks. We obtain three matrixes U, S, V as in (1) for each one block. Then, the trace of the matrix S is calculated, eventually we map the traces to the range $[0, 1023]$ and we named the value as Block Authentication Number(BAN). Assuming the traces are $Trace_{m,n}(m \in [1, M_1/4], n \in [1, M_2/4])$, i.e.,

$$BAN_{m,n} = \lfloor(Trace_{m,n} mod\ 1024)\rfloor \tag{4}$$

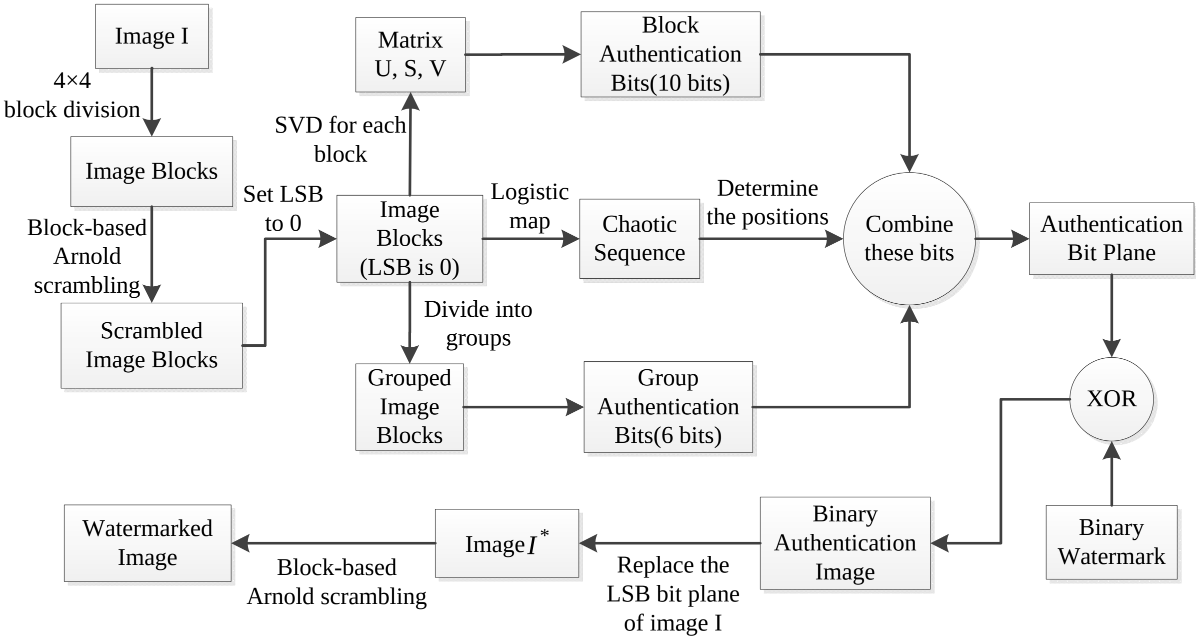


Fig. 1: Block diagram of embedding procedure

5) Group The Image Blocks: Grouping all the scrambled image blocks $ScrB_{m,n}$, each group have five blocks. The detailed procedure of Grouping is described as follows.

Firstly, the index of scrambled image block is converted from two-dimensional to one-dimensional, i.e.,

$$ScrB_k = Scrb_{m,n}$$

$$k = (m-1) \times M_1/4 + n$$

$$(m \in [1, M_1/4], n \in [1, M_2/4]), k \in [1, M]$$

Then, for each block in $ScrB_k$, we calculate the start position $StartIndex$ and end position $EndIndex$ of the group which the block is in:

$$StartIndex_k = \lfloor((k-1)/5)\rfloor \times 5 + 1$$

$$EndIndex_k = \lceil(k/5)\rceil \times 5,\ \ k \in [1, M]$$

Thirdly, for every block, we obtain the grouped scrambled image blocks GSB_k:

$$GSB_k = ScrB_{GroupIndexes}$$

$$GroupIndexes \in [StartIndex_k, EndIndex_k]$$

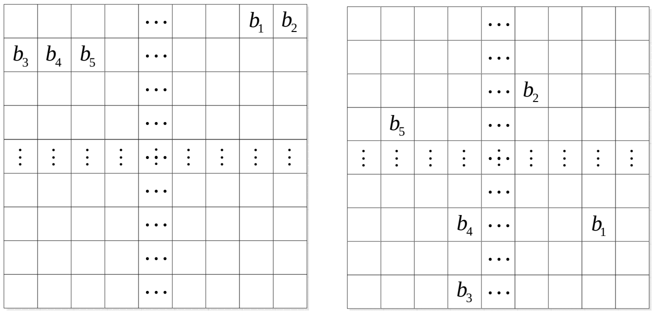
Finally, the index of grouped scrambled image blocks is converted from one-dimensional to two-dimensional, i.e.,

$$GSB_{p,q} = GSB_k$$

$$p = \lceil k/(M_2/4)\rceil, q = k\ mod\ (M_2/4)$$

$$p \in [1, M_1/4],\ \ q \in [1, M_2/4],\ \ k \in [1, M]$$

Figure 2 illustrates the basic concept of image blocks grouping method. In Figure 2, $\{b_1, b_2, b_3, b_4, b_5\}$ are grouped image blocks, they are neighbours in the scrambled imagebut they are spatially unrelated in the original image. Since the scrambling times is not known to the attacker, hence he can't obtain the grouped image blocks without the key. It improves the security of our scheme.



(a) scrambled image

(b) original image

Fig. 2: Grouped blocks in scrambled image and original image

6) Calculate Group Authentication Number: According to the Grouped image blocks $GSB_{p,q}$ and Block authentication number $BAN_{m,n}$ that we have got, we can calculate the Group Authentication Number(GAN) for each group, that is:

$$GAN_{p,q} = \sum_{(m,n)\ is\ in\ GSB_{p,q}} BAN_{m,n}/5$$

Eventually we map the GAN to the range $[0, 63]$:

$$GAN_{p,q} = \lfloor GAN_{p,q}\ mod\ 64\rfloor$$

$$p \in [1, M_1/4],\ \ q \in [1, M_2/4]$$

7) Generate The Adaptive Chaotic Sequence: For one block in the $ScrB_{m,n}$, computing the average value and the

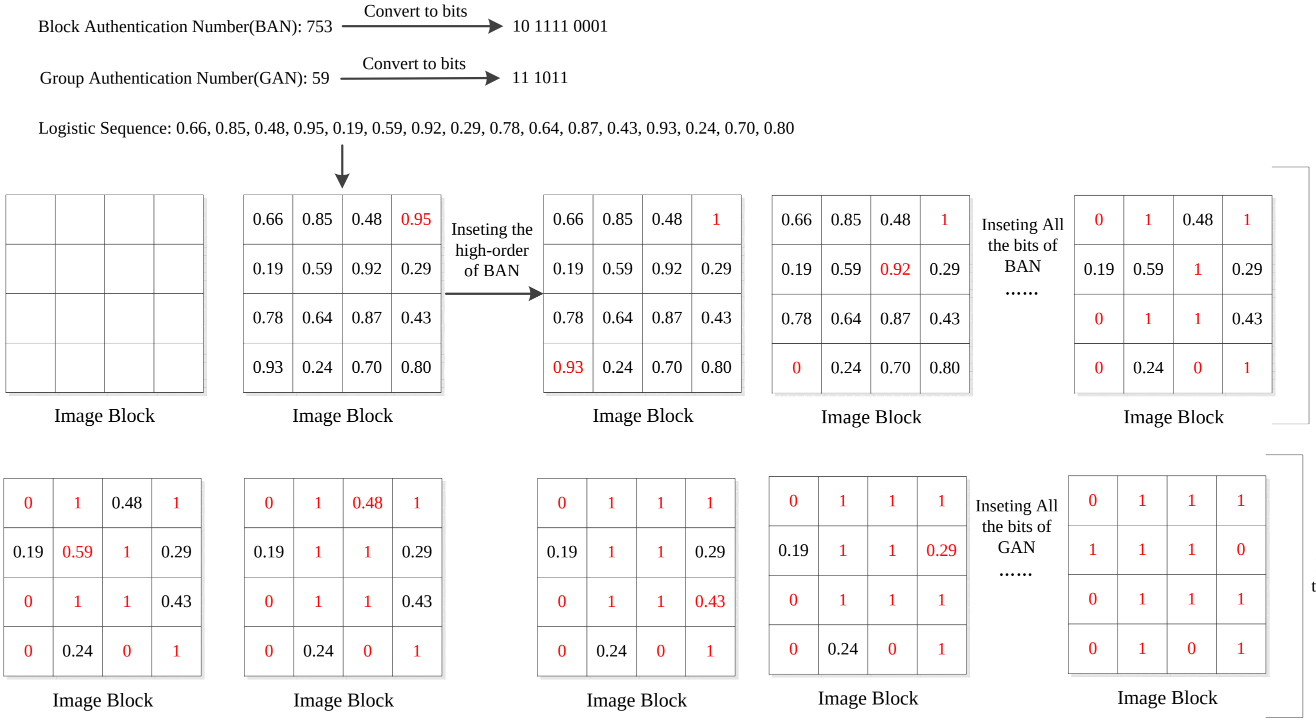


Fig. 3: An example of watermark bits inserting

standard deviation of the pixels. Denote them as $Average_{m,n}$ and $StDev_{m,n}$, respectively. Then:

$$Initial_{m,n} = (Average_{m,n} + 1)/257$$

$$Param_{m,n} = 3.5699456 + (StDev_{m,n} - \lfloor StDev_{m,n} \rfloor) \times 0.43$$

In this way we can obtain the initial value and μ of logistic sequence described in (2). At last we use these above values to generate the logistic sequences $LogSeq_{m,n}$, wherein the length of any one is 16:

$$LogSeq_{m,n} = \{x_1, x_2, \dots,\ x_{16}\}$$

This mechanism ensures that the different image blocks which have different average value or standard deviation obtain different logistic sequence.

8) Insert Watermark: First convert the BAN and GAN to binary bits. Since BAN are in the range [0, 1023] and GAN are in the range [0, 63], in that way the total length of these two binary bits is 16. The embedding positions depend on the logistic sequence, where the maximum value of the logistic sequence insert the high-order bit of BAN, and so forth, after inserting all the bits of BAN, the bits of GAN are then inserted one by one depend on the relationship in the logistic sequence, the low-order bit of GAN insert into the position where the smallest value of the logistic sequence is. Figure 3 give an example of inserting the watermark bits in one image block.

After all blocks are done watermark inserting, merge all 4×4 bit planes to obtain the Authentication Bit Plane (ABP), which has the same size to the original image.

9) XOR with the Binary Watermark Image: The watermark image is a visually meaningful binary image, denote it as W. The size of W is same with the original image I. Obtain the binary authentication image BAI using exclusive-or(XOR) operation between the binary image W and ABP as follows:

$$BAI = W \oplus ABP$$

10)Replace the LSB: Replace the LSB plane of $ScrI$ by BAI.

11)Arnold Scrambling: Owing to the Arnold transform is a periodic transformation. Apply it $(T - k)$ times on modified $ScrI$ to get the watermarked image, where T is the period of Arnold transform.

B. Watermark Extracting Procedure

Figure 4 schematically shows the watermark extract procedure, most of the steps are same or similar to the corresponding steps of the embedding procedure, therefore these steps will not repeat description here.

With the LSB plane of watermarked image, combining with the chaotic sequences generated by block to determining to positions, the original block authentication bits for each image block is obtained, denote it as $OrigBAN_{m,n}$. The block authentication bits and the group authentication bits calculated from the 7 MSB plane are $BAN_{m,n}$ and $GAN_{m,n}, (m \in [1, M_1/4], n \in [1, M_2/4])$, respectively. The procedure of matching to obtain the extracted watermark image and locate the modified regions of watermarked image is described as follows.

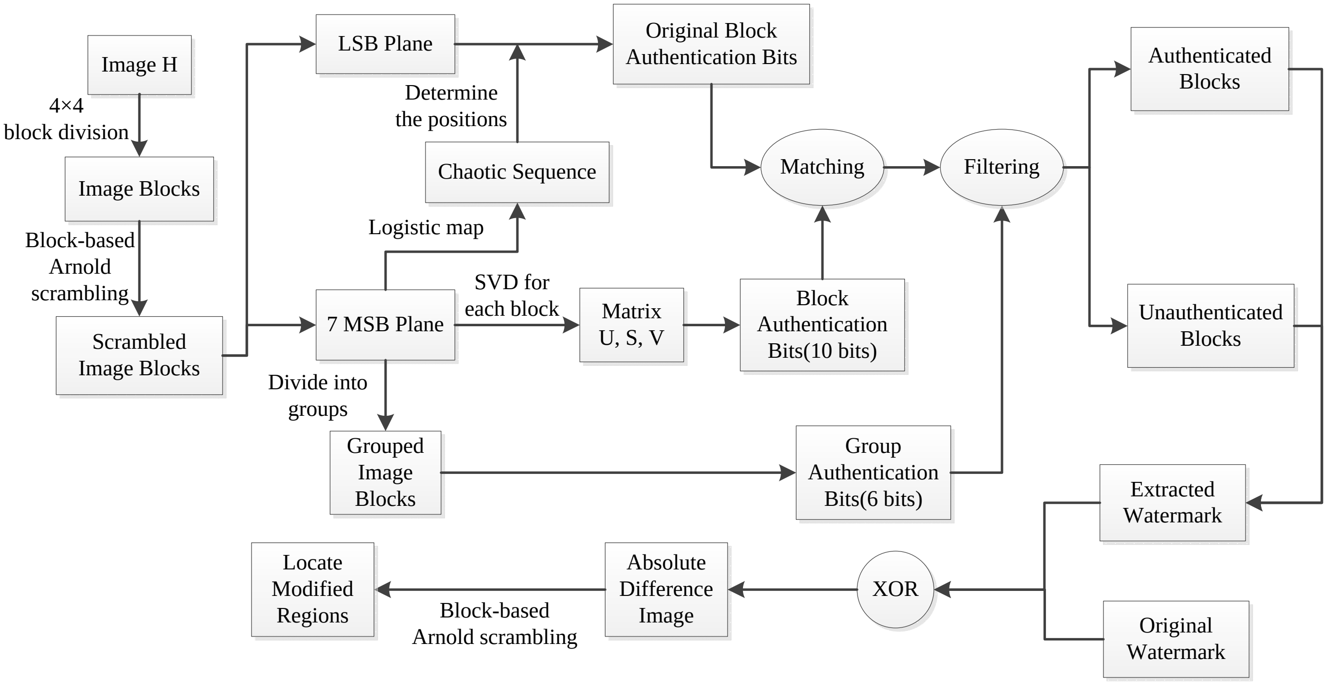


Fig. 4: Block diagram of extracting procedure

1) Block Authentication Bits Matching: This step is very simple and straightforward, if the original block authentication bits is not equally to the calculated block authentication bits, then the block is an unauthenticated block, meanwhile the watermark block generated by the block is in contrast with the original watermark at the same position. The extracted watermark $EW_{m,n}$ is a binary image and the original watermark $OW_{m,n}$ also is an binary image.

$$EW_{m,n} = \begin{cases} OW_{m,n} & if\ BAN_{m,n} = OrigBAN_{m,n} \\ \sim OW_{m,n} & otherwise \end{cases}$$

2) Group Authentication Bits Filtering: Firstly, divide all the blocks into groups use the method described above, find the most frequent element in one group, and denote as $FreqGAN_{m,n}$. Then,

$$EW_{m,n} = \begin{cases} OW_{m,n} & if\ GAN_{m,n} = FreqGAN_{m,n} \\ \sim OW_{m,n} & otherwise \end{cases}$$

3) Locating the Modified Regions: Merge the $EW_{m,n}$ to get the complete extracted watermark binary image W_{ext}, next take the absolute difference of W_{ext} and the original watermark image. Apply Block-based Arnold scrambling $(T - k)$ times to locate the tampered areas of the watermarked image.

IV. Simulation Experiments

Various experiments are carried out in this section, to assess the performance of the proposed scheme. A binary image of size 512×512 is used as watermark image in all the experiments. The parameters of Arnold transform used in our scheme are, $a = 1, b = 1$, and $k = 30$. PSNR(peak signal-to-noise ratio), is used in this paper to analyze the visual quality of the watermarked image in comparison with the original image.

A. *Performance under copy and paste attack*

In this experiment, 'Sailboat' image of size 512×512 is used. Figure 5 shows the host image, binary watermark and the corresponding watermarked image. The PSNR value of watermarked image is 51.1420 dB. Two kinds of copy and paste attacks are performed in our scheme. In first kind of copy and paste attack the watermarked image is modified by inserting two more boats in the image, where the boats are copied from the same watermarked image. The experiment result is shown in Figure 6.

In second kind of copy and paste attack the watermarked sailboat image is modified by inserting an U.S. Air Force jet in the image where the jet is copied from some other watermarked image. The experiment result is shown in Figure 7.

B. *Performance under text addition*

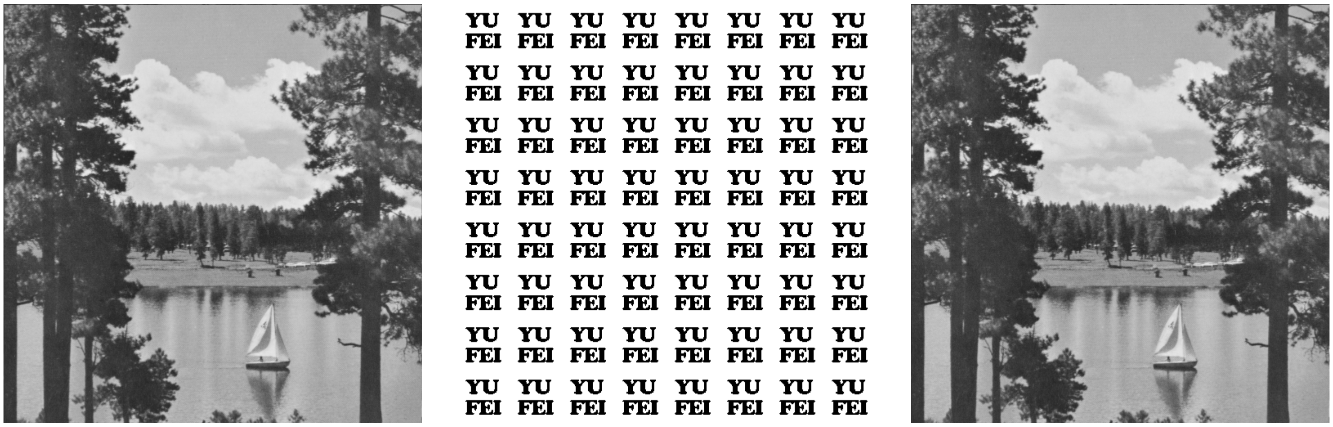
In this experiment, the watermarked couple image, shown in Figure 8(a) is modified by adding the text 'COUPLE' at the bottom of the image. Figure 8 shows the experiment result.

C. *Performance under content removal*

In this experiment, some content of the watermarked couple image is removed without degrading the image quality. The painting hanging on the wall had been removed from the watermarked image. The experiment result is shown in Figure 9.

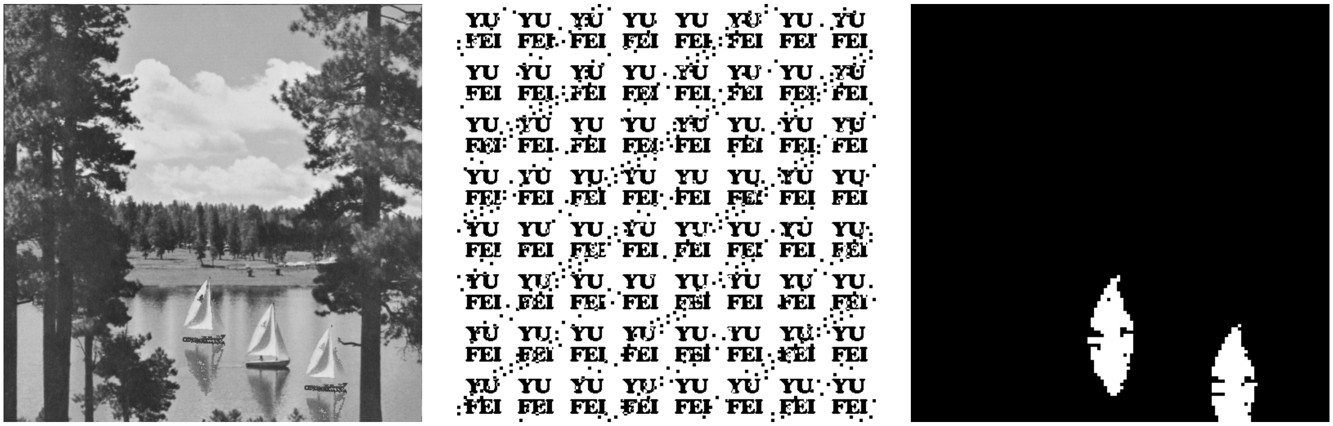
D. *Performance under VQ attack*

To evaluate the performance under VQ attack, a counterfeit image is formed by combining the portions of multiple watermarked images while preserving their relative spatial



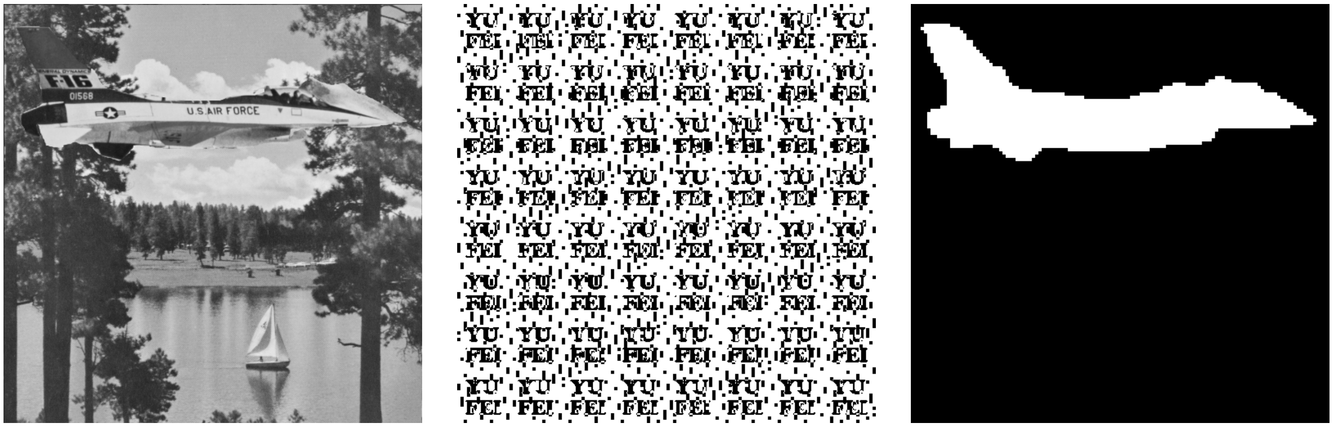
(a) Original sailboat image (b) Binary watermark image (c) Watermarked sailboat image

Fig. 5: Watermark experiment for sailboat image



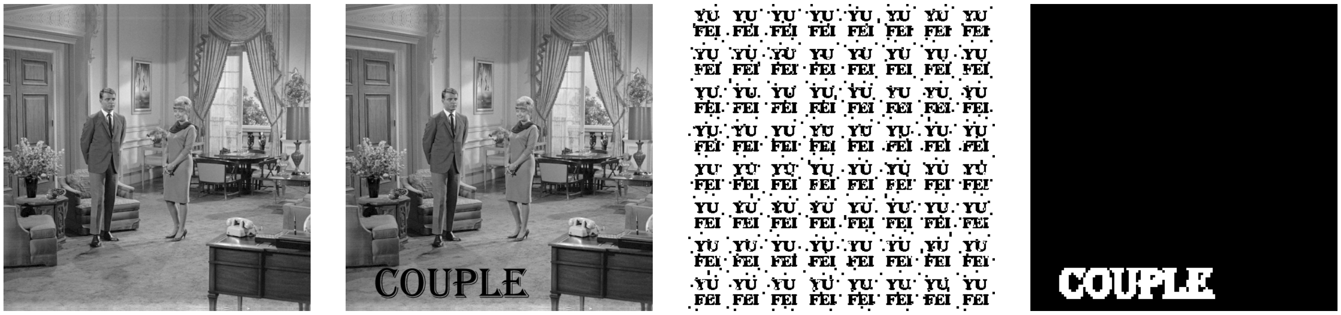
(a) Tampered sailboat image (b) Extracted watermark (c) Detected tampered region

Fig. 6: Experiment result of the first kind of copy and paste attack



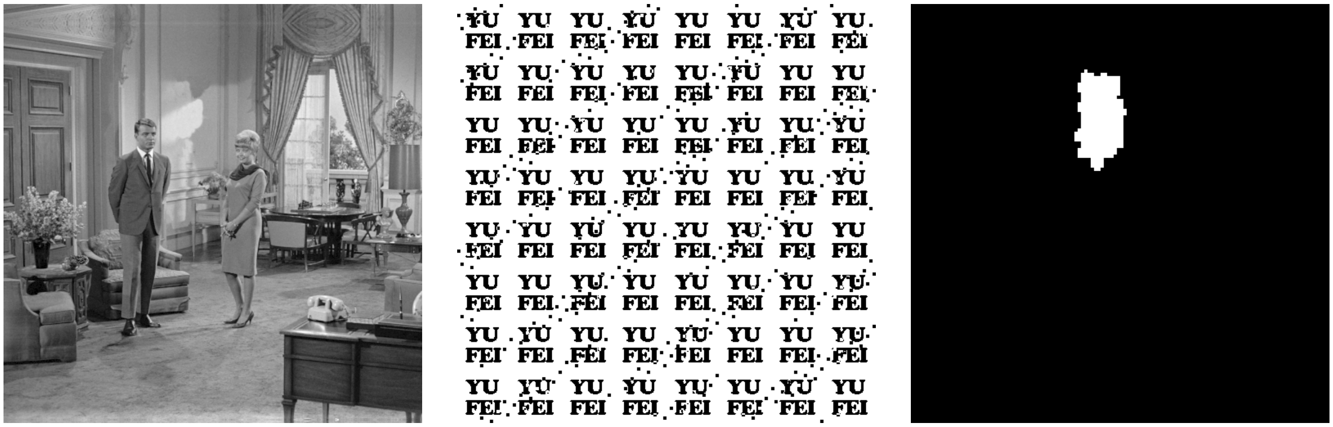
(a) Tampered sailboat image (b) Extracted watermark (c) Detected tampered region

Fig. 7: Experiment result of the second kind of copy and paste attack



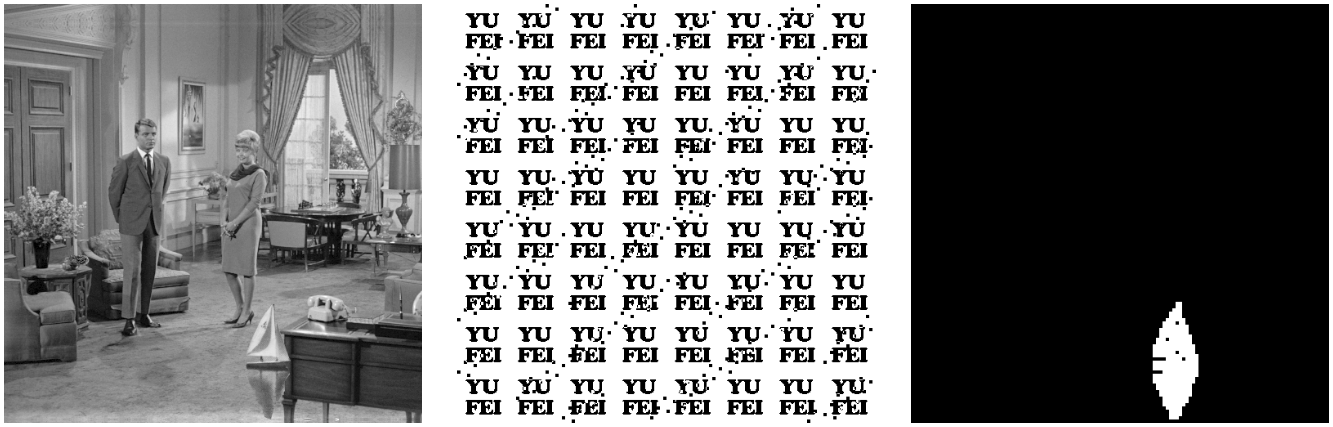
(a) Watermarked couple image (b) Tampered couple image (c) Extracted watermark (d) Detected tampered region

Fig. 8: Experiment result of the text addition attack



(a) Tampered couple image (b) Extracted watermark (c) Detected tampered region

Fig. 9: Experiment result of the content removal attack



(a) Tampered couple image with VQ attack (b) Extracted watermark (c) Detected tampered region

Fig. 10: Experiment result of the VQ attack

locations within the target image [15]. Figure 10 shows the experiment result. The counterfeit image, as shown in Figure 10(a) was constructed by copying the sailboat from Figure 5(c) and pasting it in Figure 8(a).

V. Conclusion

In this paper, A novel fragile watermarking scheme based on SVD and grouped blocks for image authentication is proposed. In order to withstand VQ attack, the watermark bits include two types of bits: the block authentication bits for authenticate the block image data and the group authentication bits for authenticate the grouped block image data. The chaotic sequence which is adaptive for each image block is generated in order to increase the security and provides an auxiliary way to authenticate the image block data. Experiments results have confirmed that this new scheme has high fidelity and excellent capability of localizing modified region in watermarked image. Therefore, it is a promising technique for fragile watermarking for high-quality and reliable still images.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No.61202160).

References

[1] Adil Haouzia and Rita Noumeir. Methods for image authentication: a survey. *Multimedia tools and applications*, 39(1):1–46, 2008.

[2] Weisi Lin, Dacheng Tao, Janusz Kacprzyk, Zhu Li, Ebroul Izquierdo, and Haohong Wang. *Multimedia Analysis, Processing and Communications*, volume 346. Springer, 2011.

[3] Gary L Friedman. The trustworthy digital camera: Restoring credibility to the photographic image. *Consumer Electronics, IEEE Transactions on*, 39(4):905–910, 1993.

[4] Toshihiko Matsuo and Kaoru Kurosawa. On parallel hash functions based on block-ciphers. *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, 87(1):67–74, 2004.

[5] Ninghui Li, Wenliang Du, and Dan Boneh. Oblivious signature-based envelope. *Distributed Computing*, 17(4):293–302, 2005.

[6] Minerva M Yeung and Fred Mintzer. An invisible watermarking technique for image verification. In *Image Processing, 1997. Proceedings., International Conference on*, volume 2, pages 680–683. IEEE, 1997.

[7] Ping Wah Wong and Nasir Memon. Secret and public key image watermarking schemes for image authentication and ownership verification. *Image Processing, IEEE Transactions on*, 10(10):1593–1601, 2001.

[8] Shan Suthaharan. Fragile image watermarking using a gradient image for improved localization and security. *Pattern Recognition Letters*, 25(16):1893–1903, 2004.

[9] Chun-Shien Lu and H-YM Liao. Structural digital signature for image authentication: an incidental distortion resistant scheme. *Multimedia, IEEE Transactions on*, 5(2):161–173, 2003.

[10] Yu-Chen Hu, Chun-Chi Lo, Chang-Ming Wu, Wu-Lin Chen, and Chia-Hsien Wen. Probability-based tamper detection scheme for btc-compressed images based on quantization levels modification. *International Journal of Security and Its Applications*, 7(3):11–32, 2013.

[11] S. Walton. Information authentication for a slippery new age. 1995.

[12] Shao-Hui Liu, Hong-Xun Yao, Wen Gao, and Yong-Liang Liu. An image fragile watermark scheme based on chaotic image pattern and pixel-pairs. *Applied Mathematics and Computation*, 185(2):869–882, 2007.

[13] N Memon, S Shende, and Ping Wah Wong. On the security of the yeung-mintzer authentication watermark. In *IS AND TS PICS CONFERENCE*, pages 301–306. SOCIETY FOR IMAGING SCIENCE & TECHNOLOGY, 1999.

[14] Matthew Holliman and Nasir Memon. Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes. *Image Processing, IEEE Transactions on*, 9(3):432–441, 2000.

[15] Sanjay Rawat and Balasubramanian Raman. A chaotic system based fragile watermarking scheme for image tamper detection. *Aeu-international Journal of Electronics and Communications*, 65:840–847, 2011.

[16] Xinpeng Zhang and Shuozhong Wang. Statistical Fragile Watermarking Capable of Locating Individual Tampered Pixels. *IEEE Signal Processing Letters*, 14:727–730, 2007.

[17] Chih-Chin Lai. A digital watermarking scheme based on singular value decomposition and tiny genetic algorithm. *Digital Signal Processing*, 21(4):522–527, 2011.

[18] Rui Sun, Hong Sun, and Tianren Yao. A svd-and quantization based semi-fragile watermarking technique for image authentication. In *Signal Processing, 2002 6th International Conference on*, volume 2, pages 1592–1595. IEEE, 2002.

[19] Ruizhen Liu and Tieniu Tan. An SVD-based watermarking scheme for protecting rightful ownership. *IEEE Transactions on Multimedia*, 4:121–128, 2002.

[20] Aidan Mooney, John G. Keating, and Daniel M. Heffernan. Performance analysis of chaotic and white watermarks in the presence of common watermark attacks. *Chaos Solitons Fractals*, 42:560–570, 2009.