Effective Fragile Watermarking for Image Authentication with High-quality Recovery Capability

Chuan Qin[1], Chin-Chen Chang[2, 3] and Tai-Jung Hsu[4]
[1] School of Optical-Electrical and Computer Engineering,
University of Shanghai for Science and Technology, Shanghai 200093, China
[e-mail: qin@usst.edu.cn]

[2] Department of Information Engineering and Computer Science,
Feng Chia University, Taichung 40724, Taiwan
[e-mail: alan3c@gmail.com]

[3] Department of Computer Science and Information Engineering,
Asia University, Taichung 41354, Taiwan
[e-mail: alan3c@gmail.com]

[4] Department of Computer Science and Information Engineering,
National Chung Cheng University, Chiayi 62102, Taiwan
[e-mail: andyblack77@gmail.com]
*Corresponding author: Chin-Chen Chang

Received June 23, 2013; revised September 3, 2013; revised October 2, 2013; accepted November 9, 2013; published November 29, 2013

Abstract

In this paper, we propose an effective fragile image watermarking scheme for tampering detection and content recovery. Cover image is divided into a series of non-overlapping blocks and a block mapping relationship is constructed by the secret key. Several DCT coefficients with direct current and lower frequencies of the MSBs for each block are used to generate the reference bits, and different coefficients are assigned with different bit numbers for representation according to their importance. To enhance recovery performance, authentication bits are generated by the MSBs and the reference bits, respectively. After LSB substitution hiding, the embedded watermark bits in each block consist of the information of itself and its mapping blocks. On the receiver side, all blocks with tampered MSBs can be detected and recovered using the valid extracted reference bits. Experimental results demonstrate the effectiveness of the proposed scheme.

Keywords: Fragile watermarking, image authentication, tampering detection, content recovery

This research was supported by the Natural Science Foundation of China (61303203), the Natural Science Foundation of Shanghai, China (13ZR1428400), and the Innovation Program of Shanghai Municipal Education Commission (14YZ087).

http://dx.doi.org/10.3837/tiis.2013.11.023

1. Introduction

Digital communication and signal processing technologies have been developed rapidly in recent years. As a result, various operations, such as editing, copying, and distribution of digital contents, have become more and more convenient. But, if the digital contents are distributed illegally or modified without authorization, copyright infringements and harmful social effects might occur. Therefore, how to protect digital contents with the capabilities of ownership identification and integrity authentication has aroused the significant research interest among the academia [1-7].

Many techniques can be utilized to realize the identification of the trustworthiness and the integrity for digital content. Multimedia hashing technique can produce a fixed-length string that is a compact representation of the principle features of the multimedia data [8-10]. Multimedia data that are perceptually similar have similar hash strings, whereas perceptually different data have very distinct hash strings. Thus, multimedia hashing can be applied in authentication. However, the hash string must be appended with the original multimedia data and transmitted to the receiver side together. Digital forensic technique can be utilized to decide whether received multimedia data have undergone certain malicious operations without having any knowledge about the original data [11-13]. Intrinsic traces and inconsistencies, such as the color filter array (CFA) of image capturing devices and lighting directions, are analyzed to produce the authenticity judgment. However, such forensic schemes have relatively low accuracy and involve considerable computational complexity. Fragile watermarking technique can realize multimedia authentication through embedding the auxiliary information imperceptibly, i.e., watermark, into the multimedia cover data [14-27]. The integrity of the received data can be judged easily using the extracted watermark and the re-calculated watermark. If the embedded watermark is generated from the multimedia cover data, such schemes are called self-embedding fragile watermarking schemes. In this work, we mainly focus on the fragile watermarking for digital images.

The purpose of earlier research on the fragile image watermarking was to realize the localization of tampered image regions [14-18]. There are two categories of fragile watermarking, i.e., block-wise schemes and pixel-wise schemes, which differ in their accuracy of locating the tampered regions. Block-wise schemes often divide the cover image into non-overlapping blocks and embed the watermark into each block [14-16], and the embedded watermark can be a hash of the principal content of each cover image block. If the watermarked image is tampered by an attacker, the extracted watermark and the image content corresponding to the tampered blocks are mismatched so that the localization of the tampered blocks can be achieved. Pixel-wise schemes often generate the watermarked image by embedding the watermark information derived from the cover image pixels [17-18]. Zhang and Wang proposed a pixel-wise fragile watermarking based on a statistical mechanism [18]. In this method, a set of tailor-made authentication data for each cover pixel and some additional test data were hidden into the cover image. On the authentication side, two different distributions of tampered and original pixels can be utilized to locate the tampered pixels. Although the pixel-wise scheme has more precise localization capability than the block-wise scheme, it allows only a relatively smaller tampered area.

Recent research works on fragile image watermarking methods have focused on content recovery in addition to the tampering localization [19-28]. These reported methods often embed the compressed code of image content into the cover image. Once the watermarked

image is tampered, the extracted watermark can be decoded for content recovery.

Fridrich *et al*. encoded the DCT coefficients of each cover image block into 64 or 128 bits and used them to replace the least significant bits (LSB) of another block [19]. After the tampering identification, the quantized DCT coefficients were extracted from the intact regions and decoded to recover the content of the tampered areas. Zhang *et al*. integrated the block-wise scheme and the pixel-wise scheme by using a hierarchical mechanism [20]. After identifying the tampered blocks, the watermark bits embedded in the intact blocks were exploited to locate the tampered pixels. In [21], by using a reversible data hiding method, reference-bits and check-bits were hidden into the cover image as the watermark. When the tampered region is not too large, this scheme can recover the cover image with no errors, but the visual quality of the watermarked image is unsatisfactory. A fragile watermarking scheme with the capability of content restoration based on an adaptive bit allocation mechanism was proposed in [22]. In this scheme, the restoration-bits for tampering recovery were generated according to the priority of each block by using the nonsubsampled contourlet transform (NSCT) coefficients. Due to the low embedding capacity, the visual quality of the watermarked image is high. Two self-embedding schemes based on a reference sharing mechanism were proposed in [24]. In these two schemes, the watermark to be embedded was a reference calculated using the original principle contents from the different regions and shared by these regions, which can achieve good recovery performance for higher tampering rate. However, the reference sharing based watermark generation and extraction processes of these two schemes were relatively complex and time-consuming. Qian *et al*. proposed an image self-embedding scheme, in which the cover image was compressed into a number of bits by multi-level encoding [25] and each block was encoded into 64 bits on average. On the receiver side, after de-quantization, inverse DCT, and rounding operation, the reference-bits were decompressed, and the tampered blocks can be recovered. Instead of embedding gray-level information or frequency coefficients, Yang *et al*. created an index table of cover image via vector quantization (VQ) and used a secret key to determine where to embed the VQ indices of all the blocks securely [26]. After tampering detection, the VQ index table can be reconstructed and the tampered areas can be recovered by VQ codewords. However, if all embedded copies of the VQ index of the block were destroyed, the quality of the recovered image by this method was not high enough. In order to decrease the amount of embedding data while maintaining good recovery quality, the scheme in [27] generated reference bits by encoding different types of blocks into different number of bits and by integrating the inpainting technique. But, if the flag bits used to indicate the block type were damaged, the recovery procedure can not be conducted.

In this work, we propose a fragile watermarking scheme for image authentication, which can detect tampered regions correctly and recover image contents with high-quality. In order to produce the reference bits of each block, several direct-current and low-frequency DCT coefficients of the most significant bits (MSBs) are represented using different numbers of binary bits and are embedded into other different blocks according to the constructed mapping relationship. In the proposed scheme, the authentication bits of each block for tampering detection are generated by its MSBs and its reference bits, respectively, which can be utilized to differentiate the tampering manipulations on the MSBs and LSBs effectively and can make full use of the valid reference bits for content recovery.

The rest of the paper is organized as follows. Section 2 describes the watermark embedding procedure of the proposed scheme. Section 3 presents the procedures of tampering detection and content recovery. Experimental results and analysis are given in Section 4, and Section 5 concludes the paper.

2. Watermark Embedding Procedure

In the proposed scheme, the λ LSB planes of each cover image block are used to carry the watermark bits. The embedded bits in each block are generated from the block itself and the other k different blocks, which can be utilized for tampering localization and image recovery on the receiver side. In other words, the watermark bits include the authentication bits for tampering judgment and the reference bits for block content recovery, and the authentication bits can be considered as the guide for the reference bits in block recovery. The flowchart of the watermark embedding procedure for each cover image block is illustrated in **Fig. 1**. In the following, we describe the watermark generation and embedding detailedly.

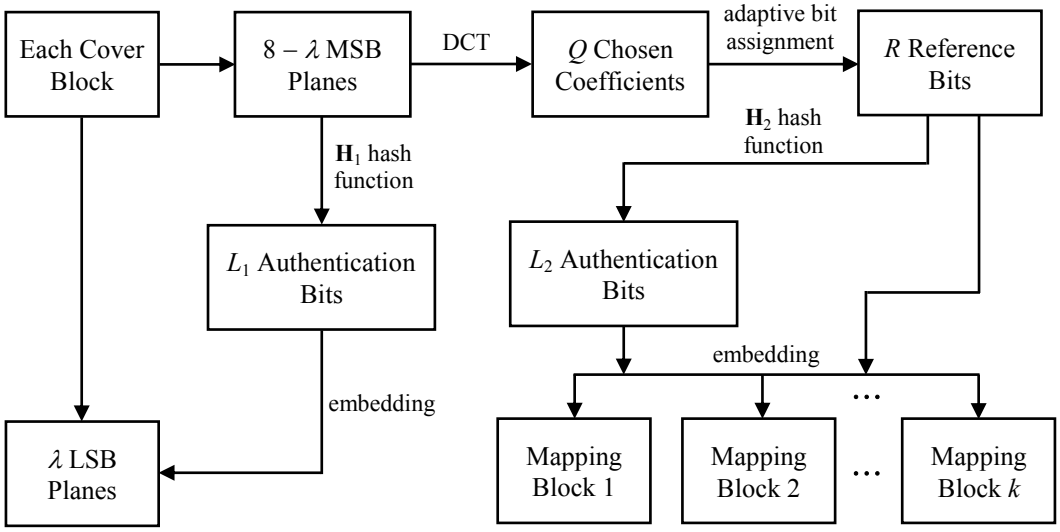


Fig. 1. Flowchart of watermark embedding procedure

2.1 Block Mapping

As mentioned above, the λ LSBs of each block are embedded with the bits generated from itself and the other k blocks. Thus, we first construct the block mapping relationship for each block and its k corresponding blocks. Suppose that the cover image $\mathbf{I}$ is divided into non-overlapped $s \times s$ blocks $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n$, where n is the number of all divided blocks. The k kinds of random arrangements, i.e., $\mathbf{P}^{(i)} = \{p_1^{(i)}, p_2^{(i)}, \ldots, p_n^{(i)}\}$ for $\{1, 2, \ldots, n\}$ are produced using secret keys ($i = 1, 2, \ldots, k$). Note that each $\mathbf{P}^{(i)}$ is a permuted version for $\{1, 2, \ldots, n\}$, and the $k + 1$ numbers, i.e., $j, p_j^{(1)}, p_j^{(2)}, \ldots, p_j^{(k)}$, should be different with each other ($j = 1, 2, \ldots, n$). Therefore, for each cover block $\mathbf{B}_j$, its corresponding mapping blocks can be easily exploited, see Eq. (1).

$$\mathbf{B}_{p_j^{(1)}}, \mathbf{B}_{p_j^{(2)}}, \ldots, \mathbf{B}_{p_j^{(k)}} \Rightarrow \mathbf{B}_j, \tag{1}$$

where the symbol $\Rightarrow$ implies that the blocks in its left part map into the block in its right part, and $j = 1, 2, \ldots, n$. According to the established relationship of block mapping, k blocks can map into one block $\mathbf{B}_j$ and the block $\mathbf{B}_j$ can also map into the other k different blocks.

2.2 Watermark-Bits Generation

To guarantee the effective tampering recovery on the receiver side, reference bits that represent the principle content of each cover block should be produced and embedded. These reference bits can also be seen as the compressed code of the cover block [28].

For a given block $\mathbf{B}_j$ ($j = 1, 2,\ldots, n$), we first divide all $s \times s$ pixel values in $\mathbf{B}_j$ by 2^{λ} to shorten the range of gray levels from [0, 255] to [0, $2^{8-\lambda} - 1$], see Eq. (2).

$$\mathbf{B}_j^{'}(x,y) = \left\lfloor \frac{\mathbf{B}_j(x,y)}{2^{\lambda}} \right\rfloor, \tag{2}$$

where $\lfloor \cdot \rfloor$ is the function to obtain the nearest integer in the direction of negative infinity, $\mathbf{B}_j(x, y)$ denotes the original pixel value in $\mathbf{B}_j$ at the coordinate (x, y), $x, y \in \{1, 2,\ldots, s\}$, and $\mathbf{B}_j^{'}(x, y)$ denotes the pixel value after processing. Then, DCT is conducted for each processed block $\mathbf{B}_j^{'}$. Denote the DCT coefficient matrix for $\mathbf{B}_j^{'}$ as $\mathbf{C}_j$. As we know, for a block or an image, the direct-current and low-frequency coefficients in the upper left corner of its DCT coefficient matrix represent the main content and structural information. Thus, in order to generate the compressed representation of $\mathbf{B}_j^{'}$, i.e., the reference bits for $\mathbf{B}_j$, only Q DCT coefficients in the upper left corner of $\mathbf{C}_j$ are chosen ($0 < Q < s^2$). Denote the Q chosen coefficients in zigzag scanning order as $c_1, c_2, \ldots, c_Q$. c_1 is the direct current coefficient that is always non-negative. Different numbers of binary bits are assigned to represent these Q coefficients due to their different importance. Denote the assigned bits for the representation of $c_1, c_2, \ldots, c_Q$ as $r_1, r_2, \ldots, r_Q$. Note that r_i is always not smaller than r_j ($1 \le i < j \le Q$). According to Eqs. (3)-(4), we can regularize the value of c_i ($i = 1, 2,\ldots, Q$) into the integer $c_i^{'}$ within the range [0, $2^{r_i} - 1$]. Then, each $c_i^{'}$ can be easily transformed into r_i binary bits.

$$c_1^{'} = \begin{cases} \text{round}(c_1), & \text{if } c_1 < 2^{r_1} - 1, \\ 2^{r_1} - 1, & \text{if } c_1 \ge 2^{r_1} - 1. \end{cases} \tag{3}$$

$$c_i^{'} = \begin{cases} 0, & \text{if } c_i \le -2^{r_i - 1} + 1, \\ \text{round}(c_i + 2^{r_i - 1} - 1), & \text{if } -2^{r_i - 1} + 1 < c_i < 2^{r_i - 1}, \\ 2^{r_i} - 1, & \text{if } c_i \ge 2^{r_i - 1}. \end{cases} \quad i = 2, 3, \ldots \tag{4}$$

Therefore, the total number of reference bits for each $\mathbf{B}_j$ is:

$$R = \sum_{i=1}^{Q} r_i . \tag{5}$$

Besides the reference bits for content recovery, the authentication bits for tampering judgment and localization should also be generated for each block $\mathbf{B}_j$. Since the λ LSB planes of each block are used for watermark embedding, only the image contents in the $8 - \lambda$ MSB planes require to be protected, and the block needs to be recovered only when its $8 - \lambda$ MSB planes are tampered. On the other hand, the reference bits for each block should also be authenticated because reference bits can be used for tampering recovery only if they are intact. Therefore, for each block $\mathbf{B}_j$, two groups of authentication bits should be generated based on

its $8-\lambda$ MSB planes and its reference bits, respectively.

The $8-\lambda$ MSB planes of each block sized $s \times s$ contain $(8-\lambda)s^2$ bits. Two one-way hash functions, i.e., $\mathbf{H}_1$ and $\mathbf{H}_2$, with the security in the sense of cryptography are adopted. Detailedly, the function $\mathbf{H}_1$ is utilized to generate the authentication bits with the length of L_1 for the $(8-\lambda)s^2$ bits in $8-\lambda$ MSB of $\mathbf{B}_j$, and the function $\mathbf{H}_2$ is utilized to generate the authentication bits with the length of L_2 for the R reference bits of $\mathbf{B}_j$. In fact, we can realize the hash functions of $\mathbf{H}_1$ and $\mathbf{H}_2$ through the multiplication of random binary matrix. For example, n random binary matrices $\mathbf{M}_j$ ($j = 1, 2,\ldots, n$) are derived through n different secret keys, and each $\mathbf{M}_j$ is with the equal size of $L_1 \times (8-\lambda)s^2$. By multiplying $\mathbf{M}_j$ with the $(8-\lambda)s^2$ bits in the $8-\lambda$ MSB of $\mathbf{B}_j$ in modulo-2 arithmetic, we can easily obtain the L_1 authentication bits for the $8-\lambda$ MSB of $\mathbf{B}_j$. Similarly, L_2 authentication bits for the R reference bits of $\mathbf{B}_j$ can also be acquired in this way.

Because reference bits are critical for block recovery, we extend R reference bits for each block into k copies, i.e., $k \times R$ bits, to reduce error possibility. Correspondingly, there are totally $k \times L_2$ authentication bits for the k copies of reference bits. Thus, the final produced watermark bits for each $\mathbf{B}_j$ consist of k copies of reference bits, k copies of the authentication bits for the reference bits, and one copy of the authentication bits for the $8-\lambda$ MSB. The total number of watermark bits for each $\mathbf{B}_j$ is $k(R+L_2)+L_1$, and the following relationship in Eq. (6) should be satisfied.

$$k(R+L_2)+L_1 \equiv \lambda s^2 . \tag{6}$$

The procedure of watermark embedding is described detailedly in the next subsection.

2.3 Watermark-Bits Embedding

All watermark bits including authentication bits and reference bits are embedded into the λ LSB planes of n cover image blocks to produce the watermarked image. Obviously, in order to ensure the accuracy of tampering judgment and the effectiveness of content recovery, for each block, its L_1 authentication bits for the $8-\lambda$ MSB should be embedded into the λ LSB of the block itself, and k copies of the R reference bits and the corresponding L_2 authentication bits should be embedded into the λ LSB of the other k different blocks according to the mapping relationship described in Subsection 2.1. Thus, in the λ LSB of each $\mathbf{B}_j$, besides the L_1 authentication bits for the $8-\lambda$ MSB of $\mathbf{B}_j$ itself, $k(R+L_2)$ bits from k mapping blocks of $\mathbf{B}_j$ in Eq. (1) are embedded. Note that, before conducting LSB substitution hiding, the $k(R+L_2)+L_1$ embedding bits for each $\mathbf{B}_j$ are scrambled for security. After all n blocks finish watermark embedding in the λ LSB planes, the final watermarked image $\mathbf{I}_W$ can be acquired.

3. Tampering Detection and Content Recovery Procedures

The watermarked image $\mathbf{I}_W$ is transmitted to the receiver side through the public channels. During the transmission, the contents of the watermarked image may be tampered by the adversaries. Therefore, the receiver should first authenticate the integrity of the received image and locate the tampered blocks. After that, the detected tampered blocks should be recovered. The flowchart of the tampering detection and content recovery procedures for each watermarked block is illustrated in **Fig. 2**.

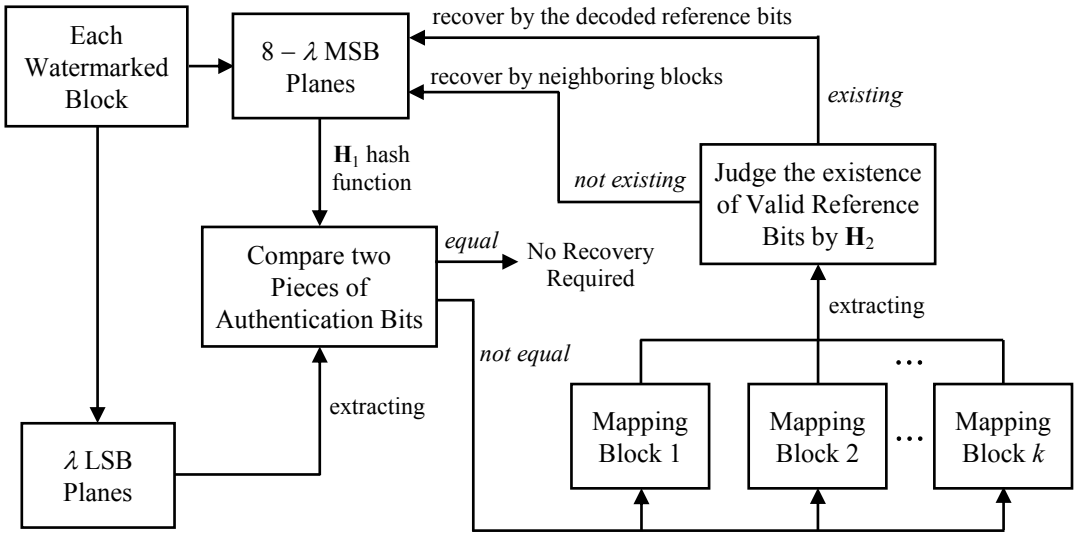


Fig. 2. Flowchart of tampering detection and content recovery procedures

3.1 Tampered-Block Detection

Conventional methods often regard any modification imposed on the block as the tampering operation. Actually, since λ LSB planes has lower importance and are just used to embed the watermark, thus, we think that only those blocks with the modification on the $8 - \lambda$ MSB planes need to be recovered. In the proposed scheme, we divide all blocks into two categories, i.e., $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$, which denote the blocks with the damaged $8 - \lambda$ MSB and the blocks with the intact $8 - \lambda$ MSB, respectively.

Denote the non-overlapped $s \times s$ blocks in the received watermarked image $\mathbf{I}_w$ as $\mathbf{W}_1$, $\mathbf{W}_2$, …, $\mathbf{W}_n$. For each $\mathbf{W}_j$ ($j = 1, 2, \ldots, n$), we collect the $(8 - \lambda)s^2$ bits in its $8 - \lambda$ MSB and feed them into the hash function $\mathbf{H}_1$ to generate L_1 bits. Then, these generated L_1 bits are compared with the extracted L_1 authentication bits from its λ LSB planes. If these two pieces of L_1 bits are equal, the block $\mathbf{W}_j$ belongs to $\mathbf{\Omega}_2$ and doesn't need to be recovered. Otherwise, the block $\mathbf{W}_j$ belongs to $\mathbf{\Omega}_1$ and needs further content recovery. The probability for a block with damaged $8 - \lambda$ MSB but being falsely categorized into $\mathbf{\Omega}_2$ is only 2^{-L_1}. After all n blocks, i.e., $\mathbf{W}_1$, $\mathbf{W}_2$,…, $\mathbf{W}_n$, are checked using the above way, the tampered blocks that belong to $\mathbf{\Omega}_1$ and need further recovery can be detected.

3.2 Content Recovery

By the shared secret keys, the receiver can also construct the same block mapping relationship of the sender side. Suppose that the block $\mathbf{W}_j$ belongs to $\mathbf{\Omega}_1$ and $j \in \{1, 2, \ldots, n\}$. The k copies of the R reference bits and the corresponding L_2 authentication bits for $\mathbf{W}_j$ are embedded into the λ LSB of its k mapping blocks. Note that the k mapping blocks of $\mathbf{W}_j$ might belong to $\mathbf{\Omega}_1$ or $\mathbf{\Omega}_2$. In order to conduct content recovery for $\mathbf{W}_j$, we extract the R reference bits and the corresponding L_2 authentication bits of $\mathbf{W}_j$ from the λ LSB of each mapping block for $\mathbf{W}_j$. Thus, k groups of the R reference bits and the corresponding L_2 authentication bits for $\mathbf{W}_j$ can be acquired. The R reference bits of each group are then fed into the function $\mathbf{H}_2$ to generate L_2 bits. The L_2 generated bits are compared with the corresponding, extracted L_2 authentication bits. If these two pieces of L_2 bits are equal, it means the extracted R reference bits of $\mathbf{W}_j$ in this

group are valid. Otherwise, the extracted R reference bits of $\mathbf{W}_j$ in this group are seen as invalid. The valid R reference bits can be decompressed and then be used to recover $\mathbf{W}_j$. If no valid reference bits of $\mathbf{W}_j$ exist among these k groups, the block $\mathbf{W}_j$ should be repaired after all the tampered blocks with valid reference bits are recovered.

If $\mathbf{W}_j$ for content recovery has valid reference bits, the valid R reference bits are divided into Q sections and the lengths of these Q divided sections are $r_1, r_2, \ldots, r_Q$. After transforming these Q binary sections into Q decimal numbers, the Q regularized integers, i.e., $c'_1, c'_2, \ldots, c'_Q$, can be obtained and the Q approximated DCT coefficients of the $8-\lambda$ MSB for $\mathbf{W}_j$ are:

$$\hat{c}_i = \begin{cases} c'_1, & \text{if } i = 1, \\ c'_i - 2^{r_i - 1} + 1, & \text{if } i = 2,3,\ldots,Q. \end{cases} \tag{7}$$

Besides these Q DCT coefficients with lower frequencies, the other $s^2 - Q$ higher frequency coefficients for $\mathbf{W}_j$ are set to zero. Thus, by using the inverse DCT, the reconstructed result of the $8-\lambda$ MSB for $\mathbf{W}_j$ can be obtained and used to replace the tampered $8-\lambda$ MSB of $\mathbf{W}_j$ for recovery. After all blocks that belong to $\mathbf{\Omega}_1$ and also have valid reference bits finish the above content recovery procedure, the remaining blocks that belong to $\mathbf{\Omega}_1$ but have no valid reference bits can be recovered by the neighborhood interpolation. The candidate blocks used for the neighborhood interpolation consist of the blocks belonging to $\mathbf{\Omega}_2$ and the recovered blocks by valid reference bits. Finally, the $8-\lambda$ MSB of all the tampered blocks belonging to $\mathbf{\Omega}_1$ can be recovered, and the whole recovered image $\mathbf{I}_O$ can be acquired successfully.

4. Experimental Results and Analysis

Experiments were conducted on a group of gray-level images to verify the effectiveness of the proposed scheme. In the experiment, the sizes of the divided non-overlapping image blocks were 8×8, i.e., $s = 8$. The parameter k for block mapping was set to 2, which means two blocks map into one block and each block can also map into the other two different blocks. Obviously, the larger the parameter value of λ was set, the better performances of tampering detection and content recovery become, because larger embedding capacity can provide more valid reference bits and corresponding detailed information to recover the tampered blocks and more authentication bits can reduce the probabilities of false judgment for the tampered block and the reference bits. But, on the other hand, the larger the value of λ was set, more hiding payload were caused, which led to the degradation of visual quality for the watermarked image. During the watermark embedding procedure of our scheme, we set the parameter λ to 3 for sufficient embedding capacity, which means that three LSBs of each cover block were replaced with watermark bits and the five MSBs were preserved. Assume that the distribution of the three LSBs of cover image is uniform. Thus, the average energy of distortion caused by the watermark embedding for each pixel is:

$$D = \frac{1}{64} \cdot \sum_{\alpha=0}^{7} \sum_{\beta=0}^{7} (\alpha - \beta)^2 = 10.5. \tag{8}$$

The theoretical peak signal-to-noise ratio (PSNR) of the watermarked image can be calculated approximately using Eq. (9), which demonstrates the visual quality of watermarked image

after embedding with $\lambda = 3$ is also satisfactory:

$$\text{PSNR} \approx 10 \cdot \log_{10} \frac{255^2}{D} = 37.92 \text{ dB}. \tag{9}$$

In order to meet the relationship in Eq. (6), 15 DCT coefficients of direct-current and lower frequencies representing the principle content in each block were used to produce reference bits, i.e., $Q = 15$. The assigned bits, i.e., $r_1, r_2, \ldots, r_{15}$, for the representation of the 15 DCT coefficients, i.e., $c_1, c_2, \ldots, c_{15}$, in each block were: 8, 6, 6, 4, 4, 4, 3, 3, 3, 3, 2, 2, 2, 2, 2, individually. Thus, the length R of the reference bits for each block was 54. The lengths L_1 and L_2 of the authentication bits for five MSBs and 54-bits reference bits were 40 and 22, respectively. **Fig. 3** shows the standard cover image *Lena* sized 512 × 512 and its watermarked version. The PSNR value of the watermarked image is 37.98 dB. It can be observed from **Fig. 3** that the visual distortion due to watermark embedding is imperceptible.

Fig. 3. Cover image Lena and its watermarked version

For convenience of description, we divide the 512 bits of each 8 × 8 received watermarked block into three sets: 1) $\mathbf{\Phi}_1$: 320 bits of the five MSBs, 2) $\mathbf{\Phi}_2$: 40 bits of the authentication bits for the five MSBs, 3) $\mathbf{\Phi}_3$: 152 bits of the reference bits and corresponding authentication bits for the two mapping blocks. The 192 bits in $\mathbf{\Phi}_2$ and $\mathbf{\Phi}_3$ come from the three LSBs of each block. We first conducted the testing of meaningful content tampering for the proposed scheme, in which the bits of $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$, and $\mathbf{\Phi}_3$ for each modified block may all be changed randomly. **Figs. 4-6** show the results of tampering detection and content recovery for our scheme. In **Figs. 4-6**, subfigures (a) show the original images *Lena*, *Elaine*, and *Lake*, and all of them are sized 512 × 512; subfigures (b) show the tampered, watermarked versions of (a) correspondingly; subfigures (c) are the tampered detection results, in which the black blocks indicate the tampered regions; subfigures (d) show the content recovery results for (b) correspondingly. The tampering percentages of subfigures (b) in **Figs. 4-6** are 2.44%, 7.40%, and 9.67%, respectively, and the PSNR values of the recovered images in subfigures (d) of **Figs. 4-6** are 37.39 dB, 36.11 dB, and 32.66 dB, respectively. Because the watermark

embedding capacity of the scheme in [19] was two LSB planes at most, the decompressed visual quality of the embedded reference bits was not high, which was approximately equivalent to a 50% quality JPEG compressed original image. Additionally, in the scheme of [19], the reference bits of each block was hidden into only one mapping block, thus, it had only one chance for each tampered block to extract its corresponding reference bits. However, when the tampered region was relatively extensive, such as **Fig. 6**(b), it had relatively high probability that the two mapping blocks were both damaged, which may lead to the failure of the recovery operation. The PSNR values of the recovered images for the tampered images in subfigures (b) of **Figs. 4-6** by the scheme in [19] are 34.46 dB, 32.38 dB, and 28.54 dB, respectively, which were lower than those of the proposed scheme. Note that, due to the low embedding payload, the average PSNR value of the watermarked images for the scheme in [19] was 50.17 dB, which was superior to that of our scheme, i.e., 37.92 dB. However, in fact, human eyes can not distinguish the significant visual difference between them.

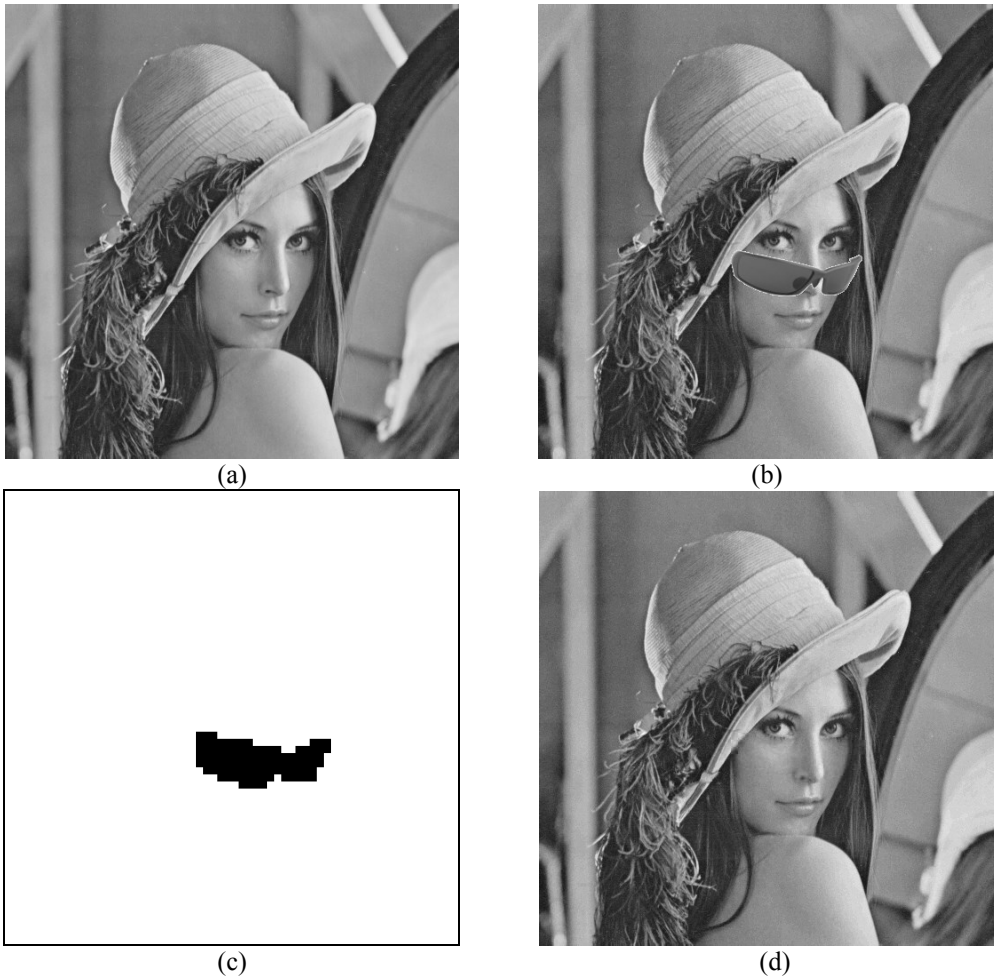
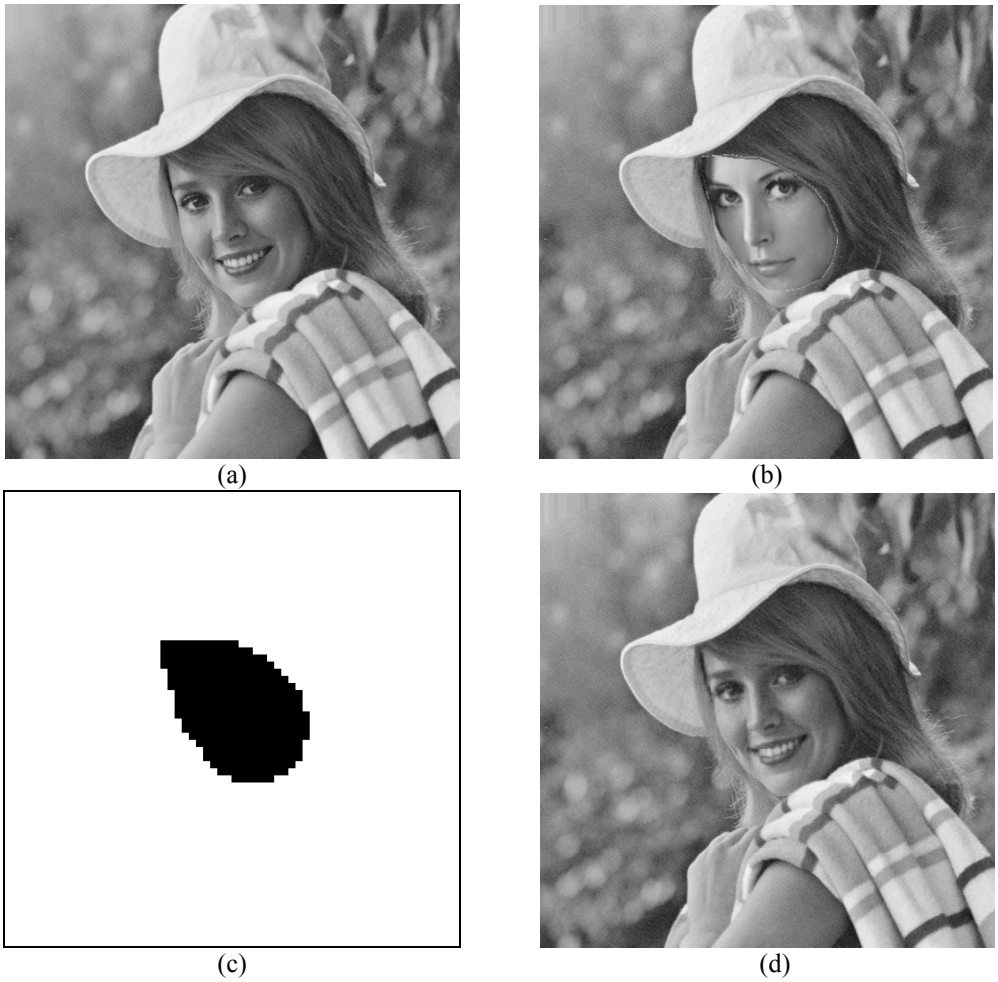
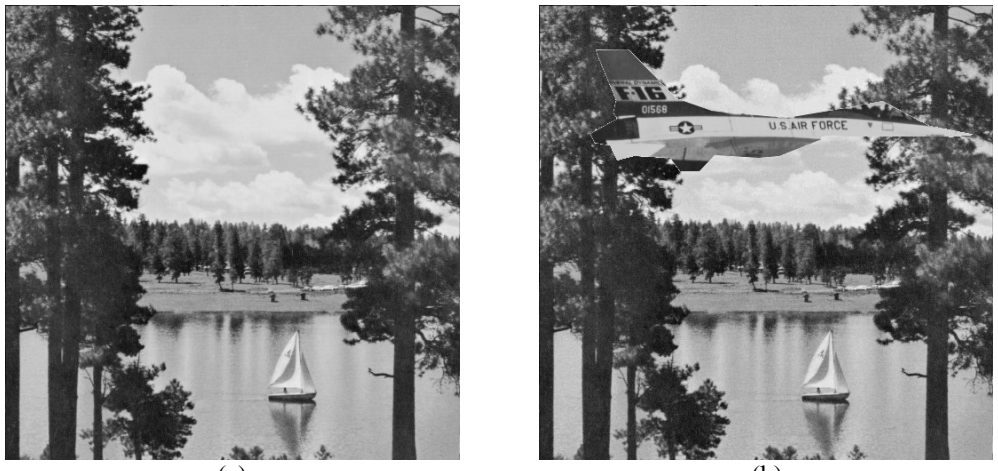


Fig. 4. Recovery result for the tampered image *Lena*. (a) Original image *Lena*; (b) Tampered, watermarked image *Lena* (tampering percentage: 2.44%); (c) Tampering detection result; (d) Content recovery result (PSNR = 37.39 dB).

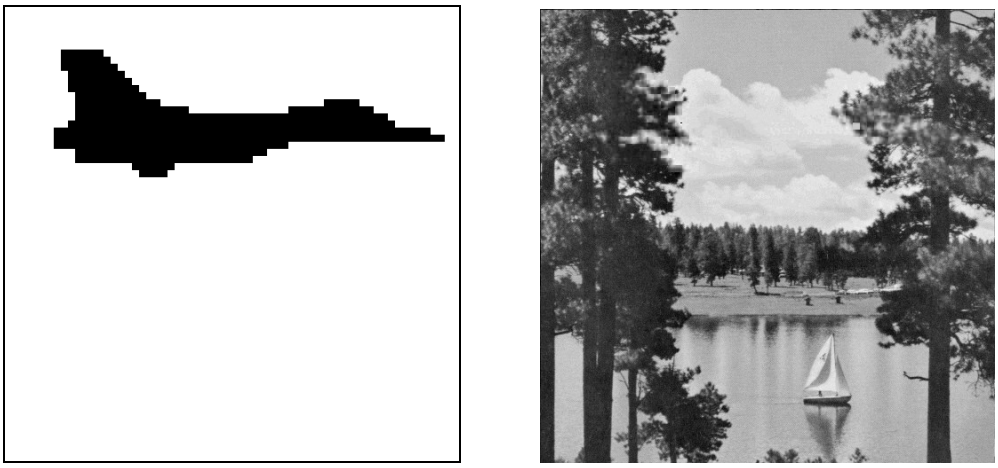


(a) (b) (c) (d)

Fig. 5. Recovery result for the tampered image *Elaine*. (a) Original image *Elaine*; (b) Tampered, watermarked image *Elaine* (tampering percentage: 7.40%); (c) Tampering detection result; (d) Content recovery result (PSNR = 36.11 dB).



(a) (b)



(c) (d)

Fig. 6. Recovery result for the tampered image *Lake*. (a) Original image *Lake*; (b) Tampered, watermarked image *Lake* (tampering percentage: 9.67%); (c) Tampering detection result; (d) Content recovery result (PSNR = 32.66 dB).

Besides the random modification on all three sets $\mathbf{\Phi}_1$, $\mathbf{\Phi}_2$ and $\mathbf{\Phi}_3$ of the watermarked blocks, we also conducted two kinds of specific modifications to show the superiority of our scheme: 1) the bits in $\mathbf{\Phi}_1$ or $\mathbf{\Phi}_2$ are modified and the bits in $\mathbf{\Phi}_3$ are intact; 2) the bits in $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ are intact and only the bits in $\mathbf{\Phi}_3$ are modified. For the both two scenarios, the conventional methods, such as [20, 23, 28], always try to recover the modified blocks. However, our scheme doesn't implement the content recovery for the second scenario because only the bits in LSBs, i.e., $\mathbf{\Phi}_3$, are modified. In fact, the recovered result of the conventional methods for the second scenario is generally worse than the unprocessed version due to the lossy representation of reference bits. For the first scenario, both our scheme and the conventional methods conduct content recovery for the modified blocks because the MSBs of the blocks are not authenticated. For our scheme, because the bits in $\mathbf{\Phi}_3$ that are embedded in the LSBs of the modified block are intact, the reference bits belonging to $\mathbf{\Phi}_3$ can be authenticated using $\mathbf{H}_2$ and can still be utilized to recover other tampered blocks. But, for the conventional methods, because the modifications on MSBs and LSBs can not be differentiated, all bits in this block are considered as invalid. Thus, the reference bits embedded in the LSBs of this block can not be used to recover other tampered blocks although these reference bits are intact. For example, if block A and block B are a pair of mapping blocks, i.e., one copy of reference bits of block A is embedded in the LSBs of block B, and if the MSBs of block A and block B are both modified, for the conventional methods, the reference bits embedded in block B can not be used to recover block A, which leads to the loss of one chance for recovery. During the experiments, we conducted the tampering only on the MSBs of watermarked blocks, which belongs to the first scenario, and the tampered blocks are mapped in pairs. **Table 1** shows the comparison results of tampering recovery for Scenario 1. The first column is the numbers of MSB-tampered mapping block pairs that are distributed randomly in the watermarked image *Lena*. The second and third columns are the percentages of the recoverable blocks for [23] and the proposed scheme, respectively. The last column is the PSNR values of the recovered results by the proposed scheme.

In the experiments, we also conducted the tampering only on the bits in $\mathbf{\Phi}_3$ of watermarked blocks, which belongs to the second scenario, and the tampered blocks are also mapped in

pairs. **Table 2** shows the comparison results of tampering recovery for Scenario 2. The first column is the numbers of tampered mapping block pairs that are distributed randomly in the watermarked image *Lena*. The second column is the percentages of the recoverable blocks for [23]. The third column is the PSNR values of the recovered results for [23]. Since our scheme doesn't conduct any recovery operations for Scenario 2, the last column is the PSNR values of the unprocessed versions in the proposed scheme. It can be found from **Tables 1-2** that the proposed scheme has better recovery performance than the conventional method [23].

Table 1. Results of tampering recovery for Scenario 1

Number of tampered block pairs	Recovered percentage of [23]	Recovered percentage of proposed scheme	PSNR of recovered results by proposed scheme (dB)
50	48.00%	100%	37.14
100	48.50%	100%	36.53
250	46.69%	100%	35.08
500	40.90%	100%	33.92
750	36.44%	100%	32.92
1000	31.80%	100%	32.22
2000	16.78%	100%	30.42

Table 2. Results of tampering recovery for Scenario 2

Number of tampered block pairs	Recovered percentage of [23]	PSNR of the scheme [23] (dB)	PSNR of the proposed scheme (dB)
50	48.00%	37.66	37.74
100	48.50%	37.48	37.70
250	46.69%	36.70	37.61
500	40.90%	36.15	37.45
750	36.44%	35.59	37.32
1000	31.80%	35.49	37.20
2000	16.78%	35.65	36.83

5. Conclusion

In this paper, a fragile image watermarking scheme with high-quality recovery capability was proposed. The LSB planes of each cover block are embedded with the watermark bits

generated from itself and other different blocks according to the constructed mapping relationship. The reference bits in the watermark bits of each block can be seen as the compressed codes of the MSBs, which are represented by a group of DCT coefficients using different bit numbers. The authentication bits among the watermark bits of each block are generated from its MSBs and reference bits, respectively, in order that the tampering manipulations on the MSBs and the LSBs can be differentiated on the receiver side. Only the blocks with the tampered MSBs are needed to repair and only the intact reference bits in LSBs can be used for content recovery. After detecting the blocks with tampered MSBs by authentication bits, all valid reference bits embedded in the LSB planes can be fully exploited and then be decompressed to recover the contents of MSBs. Experimental results show that the proposed scheme has satisfactory performances of tampering detection and content recovery.

References

[1] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE,* vol. 87, no. 7, pp. 1079-1107, 1999. Article (CrossRef Link).

[2] C. S. Chan and C. C. Chang, "An efficient image authentication method based on Hamming code," *Pattern Recognition*, vol. 40, no. 2, pp. 681-690, 2007. Article (CrossRef Link).

[3] C. C. Wu, S. J. Kao and M. S. Hwang, "A high quality image sharing with steganography and adaptive authentication scheme," *Journal of Systems and Software*, vol. 84, no. 12, pp. 2196-2207, 2011. Article (CrossRef Link).

[4] C. Qin, C. C. Chang and K. N. Chen, "Adaptive self-recovery for tampered images based on VQ indexing and inpainting," *Signal Processing*, vol. 93, no. 4, pp. 933-946, 2013. Article (CrossRef Link).

[5] P. D. Sheba Kezia Malarchelvi, "A semi-fragile image content authentication technique based on secure hash in frequency domain," *International Journal of Network Security*, vol. 15, no. 5, pp. 365-372, 2013. Article (CrossRef Link).

[6] P. Y. Lin, J. S. Lee and C. C. Chang, "Dual digital watermarking for internet media based on hybrid strategies," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 8, pp. 1169-1171, 2009. Article (CrossRef Link).

[7] S. S. Sujatha and M. M. Sathik, "A novel DWT based blind watermarking for image authentication," *International Journal of Network Security*, vol. 14, no. 4, pp. 223-228, 2012. Article (CrossRef Link).

[8] A. Swaminathan, Y. N. Mao and M. Wu, "Robust and secure image hashing," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 215-230, 2006. Article (CrossRef Link).

[9] D. Wu, X. B. Zhou and X. M. Niu, "A novel image hash algorithm resistant to print-scan," *Signal Processing*, vol. 89, no. 12, pp. 2415-2424, 2009. Article (CrossRef Link).

[10] V. Monga and M. K. Mhcak, "Robust and secure image hashing via non-negative matrix factorizations," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 376-390, 2007. Article (CrossRef Link).

[11] H. T. Sencar and N. Memon, "Overview of state-of-the-art in digital image forensics," *World Scientific Press*, 2007. Article (CrossRef Link).

[12] A. Swaminathan, M. Wu and K. J. R. Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 101-117, 2008. Article (CrossRef Link).

[13] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 154-160, 2009. Article (CrossRef Link).

[14] P. W. Wong and N. Memon, "Secret and public key image watermarking schemes for image authentication and ownership verification," *IEEE Transactions on Image Processing*, vol. 10, no.

10, pp. 1593-1601, 2001. Article (CrossRef Link).

[15] S. Suthaharan, "Fragile image watermarking using a gradient image for improved localization and security," *Pattern Recognition Letters*, vol. 25, no. 16, pp. 1893-1903, 2004. Article (CrossRef Link).

[16] H. Yang and A. C. Kot, "Binary image authentication with tampering localization by embedding cryptographic signature and block identifier," *IEEE Signal Processing Letters*, vol. 13, no. 12, pp. 741-744, 2006. Article (CrossRef Link).

[17] H. Lu, R. Shen and F. L. Chung, "Fragile watermarking scheme for image authentication," *Electronics Letters*, vol. 39, no. 12, pp. 898-900, 2003. Article (CrossRef Link).

[18] X. P. Zhang and S. Z. Wang, "Statistical fragile watermarking capable of locating individual tampered pixels," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 727-730, 2007. Article (CrossRef Link).

[19] J. Fridrich and M. Goljan, "Images with self-correcting capabilities," in *Proc. of IEEE International Conference on Image Processing*, pp. 792-796, 1999. Article (CrossRef Link).

[20] X. P. Zhang and S. Z. Wang, "Fragile watermarking scheme using a hierarchical mechanism," *Signal Processing*, vol. 89, no. 4, pp. 675-679, 2009. Article (CrossRef Link).

[21] X. P. Zhang and S. Z. Wang, "Fragile watermarking with error-free restoration capability," *IEEE Transactions on Multimedia*, vol. 10, no. 8, pp. 1490-1499, 2008. Article (CrossRef Link).

[22] C. Qin, C. C. Chang and P. Y. Chen, "Self-embedding fragile watermarking with restoration capability based on adaptive bit allocation mechanism," *Signal Processing*, vol. 92, no. 4, pp. 1137-1150, 2012. Article (CrossRef Link).

[23] H. J. He, J. S. Zhang and F. Chen, "Adjacent-block based statistical detection method for self-embedding watermarking techniques," *Signal Processing*, vol. 89, no. 8, pp. 1557-1566, 2009. Article (CrossRef Link).

[24] X. P. Zhang, S. Z. Wang, Z. X. Qian and G. R. Feng, "Reference sharing mechanism for watermark self-embedding," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 485-495, 2011. Article (CrossRef Link).

[25] Z. X. Qian, G. R. Feng, X. P. Zhang and S. Z. Wang, "Image self-embedding with high-quality restoration capability," *Digital Signal Processing*, vol. 21, no. 2, pp. 278-286, 2011. Article (CrossRef Link).

[26] C. W. Yang and J. J. Shen, "Recover the tampered image based on VQ indexing," *Signal Processing*, vol. 90, no. 1, pp. 331-343, 2010. Article (CrossRef Link).

[27] Z. X. Qian and G. R. Feng, "Inpainting assisted self recovery with decreased embedding data," *IEEE Signal Processing Letters*, vol. 17, no. 11, pp. 929-932, 2010. Article (CrossRef Link).

[28] X. P. Zhang, S. Z. Wang and G. R. Feng, "Fragile watermarking scheme with extensive content restoration capability," in *Proc. of the 8th International Workshop on Digital Watermark (IWDW 2009)*, pp. 268-278, 2009. Article (CrossRef Link).

Chuan Qin received the B.S. and M.S. degrees in electronic engineering from Hefei University of Technology, Anhui, China, in 2002 and 2005, respectively, and the Ph.D. degree in signal and information processing from Shanghai University, Shanghai, China, in 2008. Since 2008, he has been with the faculty of the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, where he is currently a Lecturer. He also has been with Feng Chia University at Taiwan as a Postdoctoral Researcher from July 2010 to June 2012. His research interests include image processing and multimedia security.

Chin-Chen Chang received his B.S. degree in applied mathematics in 1977 and the M.S. degree in computer and decision sciences in 1979, both from the National Tsing Hua University, Hsinchu, Taiwan. He received his Ph.D in computer engineering in 1982 from the National Chiao Tung University, Hsinchu, Taiwan. During the academic years of 1980-1983, he was on the faculty of the Department of Computer Engineering at the National Chiao Tung University. From 1983-1989, he was on the faculty of the Institute of Applied Mathematics, National Chung Hsing University, Taichung, Taiwan. From August 1989 to July 1992, he was the head of, and a professor in, the Institute of Computer Science and Information Engineering at the National Chung Cheng University, Chiayi, Taiwan. From August 1992 to July 1995, he was the dean of the college of Engineering at the same university. From August 1995 to October 1997, he was the provost at the National Chung Cheng University. From September 1996 to October 1997, Dr. Chang was the Acting President at the National Chung Cheng University. From July 1998 to June 2000, he was the director of Advisory Office of the Ministry of Education of Taiwan. From 2002 to 2005, he was a Chair Professor of National Chung Cheng University. Since February 2005, he has been a Chair Professor of Feng Chia University. In addition, he has served as a consultant to several research institutes and government departments. His current research interests include database design, data structures, computer cryptography and image processing. He is a fellow of the IEEE.

Tai-Jung Hsu received the B.S. degree in computer science and engineering from National University of Tainan, Taiwan, in 2011. He is currently pursuing the M.S. degree in computer science and engineering from National Chung Cheng University, Taiwan. His research interests include data hiding and watermarking.