Displays 46 (2017) 52–60

Contents lists available at ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa

Hierarchical recovery for tampered images based on watermark self-embedding ☆

Fang Cao [a,b,*], Bowen An [a,*], Jinwei Wang [b], Dengpan Ye [c], Huili Wang [d]

[a] College of Information Engineering, Shanghai Maritime University, Shanghai 200135, China
[b] School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, Jiangsu, China
[c] School of Computer, Wuhan University, Wuhan 430072, Hubei, China
[d] School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

ARTICLE INFO

Article history:
Received 15 August 2016
Received in revised form 2 January 2017
Accepted 4 January 2017
Available online 5 January 2017

Keywords:
Self-embedding
Tampering detection
Hierarchical recovery
Reference sharing

ABSTRACT

In this paper, we propose a new self-embedding watermarking scheme with hierarchical recovery capability. The binary bits in the adopted MSB layers are scrambled and individually interleaved with different extension ratios according to their importance to image visual quality. The interleaved data, which are regarded as reference bits for tampering recovery, are segmented into a series of groups corresponding to the divided non-overlapping blocks, and then embedded into the LSB layers of blocks together with authentication bits of tampering detection. Because the extension ratios of MSB-layer bits are based on the hierarchical mechanism, the efficiency of reference bits is increased, and higher MSB layers of tampered regions have greater probabilities to be recovered than lower MSB layers, which can improve the visual quality recovered results, especially for larger tampering rates. Experimental results demonstrate the effectiveness and superiority of the proposed scheme compared with some of state-of-the-art schemes.

1. Introduction

In recent years, the rapid development of multimedia tools and Internet technology brings great convenience for transmitting and downloading multimedia data, which also leads to easier duplication and modification for digital contents than before in the meantime [1,2]. Therefore, how to protect the security and integrity of the multimedia data [3,4], especially the technique of image authentication [5–7], becomes an important research topic nowadays. The traditional image authentication scheme attaches digital signatures with the original image, and then compares the signatures of the received image with that of the original image to authenticate the integrity. However, it cannot locate and further recover the tampered region [8]. In order to solve these problems, fragile watermarking scheme for image authentication with the capability of tampering localization and content recovery has been proposed [9].

In the view of functions, fragile image watermarking schemes can be divided into two types. One type can just locate suspicious regions if the received image is tampered during transmission [10–17]. This type of fragile watermarking schemes usually regards the hash of principal contents retrieved from each image block as its watermark data for embedding, and on the receiver side, the re-calculated hash of the received image is compared with the extracted hash of the image to detect the tampered regions, because tampering operation destroys the matching relationship between the contents of original image and the corresponding watermark data [10,11]. In order to improve the accuracy of tampering detection, some researchers proposed the pixel-wise based fragile watermarking schemes. The watermark data derived from gray values of original pixels were embedded into the original pixels themselves, and then the tampered pixels can be located through the absence of watermark data [12–14]. In the scheme [15], a statistical mechanism was introduced into fragile image watermarking. The watermark data, including the tailor-made authentication data for each pixel and some additional test data, can be used to precisely locate the tampered pixels.

Another type of fragile watermarking schemes can not only locate the faked regions, but also can recover the located, tampered contents [18–34]. In many practical applications, only tampering detection cannot satisfy the requirement, and the reconstruction

☆ This paper was recommended for publication by Pen-Cheng Wang.

* Corresponding authors at: College of Information Engineering, Shanghai Maritime University, No. 1550 Haigang Ave, Shanghai 201306, China.
E-mail addresses: fangcao@shmtu.edu.cn (F. Cao), bwan@shmtu.edu.cn (B. An), wjwei_2004@163.com (J. Wang), yedp@whu.edu.cn (D. Ye), wanghuili2622@163.com (H. Wang).

http://dx.doi.org/10.1016/j.displa.2017.01.001
0141-9382/

for tampered regions is highly desirable. To achieve the self-recovery capability, Fridrich et al. conducted earlier attempt in this field, and they proposed fragile image watermarking scheme, which can realize content recovery after tampering detection [18]. They embedded the watermark of a block into the least significant bits (LSB) of other distant blocks, which was able to resist vector quantization (VQ) attack and collage attack. In order to accurately identify the faked image blocks, a digital watermarking method for image tampering detection and recovery was developed in [19], which was based on a 3-level hierarchical structure. This scheme can not only detect tampered areas accurately but also can deal with high tampering rate with acceptable recovered results. However, the above mentioned methods above cannot recover the tampered blocks whose watermarks embedded in other blocks were also destroyed, which was called as the tampering coincidence problem in Zhang et al.'s scheme [24]. Some watermarking schemes with self-correction capability were free of this problem. In [20], a fragile watermarking scheme with a hierarchical mechanism was presented, which can reconstruct original watermarked image without any error. The pixel-derived and block-derived watermark data were embedded into the LSBs of all pixels. This method had a limitation that the tampering rate must be no greater than 3.2% of the entire image to achieve the perfect restoration. In another work [21], an effective dual watermark scheme for image tampering detection and recovery was proposed by Lee and Lin. They applied two copies of watermark data for each block in the entire image, so it was able to provide the second chance for tampering recovery in case the first copy was damaged. However, the tampering coincidence problem still existed once both two copies of the embedded watermark data for the image block were destroyed. A self-embedding fragile watermarking scheme based on a reference sharing mechanism was proposed in [24], in which the watermark embedded into the three LSB layers of the whole image can be considered as the reference derived from the five most significant bits (MSB) layers of original image and shared by the whole image for further content restoration. As long as the content tampering was not too extensive, the five MSB layers of tampered regions can be perfectly recovered using the sufficient available data scattered in the intact blocks of image. Thus, it can effectively avoid the tampering coincidence problem. However, the way of reference data generation also caused the watermark wasting problem [26]. Huo et al. proposed an alterable-capacity fragile watermarking scheme in [28], which the watermark codes with the alterable-length consisted of three parts and were embedded into other three blocks. On the receiver side, two copies of significant-code were utilized to recover the tampered contents so that the recovery performance can be improved. However, this scheme was poor at dealing with the tampering form of random block missing.

In this work, in order to achieve better performance of visual quality for both watermarked image and recovered image, we propose a self-embedding watermarking scheme for tampering recovery based on hierarchical watermark embedding, which utilizes variable numbers of MSB layers to generate the shared reference data for content recovery and also variable extension ratios between the reference bits for each MSB-layer and the total reference bits for all adopted MSB layers. These parameters can be flexible according to different proportions of the tampered regions to achieve the satisfactory quality of recovered contents. During watermark embedding, the reference data are derived from each MSB layer, whose bits are interleaved and scrambled, and then are combined with the authentication data to form the watermark data to be embedded in the LSBs. Note that the proposed scheme is based upon the reference sharing mechanism and the extension ratio between the reference bits for each MSB-layer and total reference bits is variable. Thus, tampering coincidence problem is effectively avoided and the efficiency of watermark data can also be greatly improved.

The rest of this paper is organized as follows. Section 2 describes the procedure of watermark embedding, including watermark generation and data embedding. Section 3 presents the procedure of content recovery, including tampering detection and content recovery. Experimental results and comparison are given in Section 4. Section 5 concludes the paper.

2. Watermark embedding

The watermark embedding procedure of the proposed scheme consists of the following 3 stages: (1) Select the embedding parameters to generate reference data; (2) Generate authentication data using reference data with the embedding parameters; (3) Embed the watermark data, including reference data and authentication data, into original image to produce watermarked image.

2.1. Watermark generation

Denote the size of original image $\mathbf{I}_o$ as $H \times W$, and $N = H \times W$. In the design of the proposed scheme, the detection of tampered region is based on each non-overlapping image block sized $b \times b$. Thus, for simplicity, H and W are both assumed as the multiples of b. The number of MSB layers used for the generation of reference bits is denoted as m. The $(8 - m)$ LSB layers of original image are used to accommodate the watermark data.

For each non-overlapping block, we allocate h authentication bits for tampering detection and $(8 - m) \cdot b^2 - h$ reference bits for content recovery, respectively. How to generate reference bits and authentication bits is described detailedly in the following steps:

Step 1: Denote the gray value of each pixel in $\mathbf{I}_o$ as $p_i \in [0, 255]$, $i = 1, 2, \ldots, N$, and p_i can be represented by 8 binary bits, i.e., $q_{i,7}$, $q_{i,6}$, $\ldots$, $q_{i,0}$, see Eq. (1).

$$q_{i,k} = \lfloor p_i / 2^k \rfloor \bmod 2, \quad k = 0,\ 1,\ \ldots, 7. \tag{1}$$

Step 2: Collect the N bits of each MSB layer, and then randomly divide them into S subsets and each subset contains u bits, i.e., $u \cdot S = N$. Detailedly, for the x-th MSB layer $\mathbf{C}_x$, its N bits, i.e., $q_{i,8-x}$, $i = 1, 2, \ldots, N$, are collected and divided into S subsets randomly, i.e., $\mathbf{C}_x^{(1)}$, $\mathbf{C}_x^{(2)}$, $\ldots$, $\mathbf{C}_x^{(S)}$.

Step 3: Denote the u MSB bits in the j-th subset $\mathbf{C}_x^{(j)}$ as $c^{(j)}_{x,1}$, $c^{(j)}_{x,2}$, $\ldots$, $c^{(j)}_{x,u}$, and these u MSB bits are transformed into v_x reference bits $\mathbf{R}_x^{(j)}$, i.e., $r^{(j)}_{x,1}$, $r^{(j)}_{x,2}$, $\ldots$, $r^{(j)}_{x,vx}$, $j = 1, 2, \ldots, S$, through Eq. (2):

$$\begin{bmatrix} r^{(j)}_{x,1} \\ r^{(j)}_{x,2} \\ \vdots \\ r^{(j)}_{x,v_x} \end{bmatrix} = \mathbf{H}_x^{(j)} \cdot \begin{bmatrix} c^{(j)}_{x,1} \\ c^{(j)}_{x,2} \\ \vdots \\ c^{(j)}_{x,u} \end{bmatrix} \quad j = 1, 2, \ldots,\ S, \tag{2}$$

where $\mathbf{H}_x^{(j)}$ is the pseudo-random binary matrix sized $v_x \times u$ that is derived from a secret key. Denote the value t_x as the extension ratio between the generated reference bits for the x-th MSB-layer and the total reference bits for all m MSB layers, $x = 1, 2, \ldots, m$, see Eq. (3). Note that the two relationships in Eqs. (4) and (5) should be satisfied:

$$t_x = \frac{S \cdot v_x}{N \cdot [(8 - m) - h/b^2]}, x = 1,\ 2, \ldots, m. \tag{3}$$

$$t_1 \geqslant t_2 \geqslant \cdots \geqslant t_m, \tag{4}$$

$$\sum_{x=1}^{m} t_x = 1. \tag{5}$$