Information Sciences 373 (2016) 233–250

Contents lists available at ScienceDirect

Information Sciences

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

# Self-embedding fragile watermarking based on reference-data interleaving and adaptive selection of embedding mode

Chuan Qin[a,*], Huili Wang[a], Xinpeng Zhang[b], Xingming Sun[c]

[a] Shanghai Key Lab of Modern Optical System, and Engineering Research Center of Optical Instrument and System, Ministry of Education, University of Shanghai for Science and Technology, Shanghai 200093, China
[b] School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China
[c] School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, Jiangsu, China

ARTICLE INFO

*Article history:*
Received 13 April 2016
Revised 28 August 2016
Accepted 1 September 2016
Available online 3 September 2016

*Keywords:*
Fragile watermarking
Self-embedding
Embedding mode
Tampering detection
Content recovery

ABSTRACT

In this paper, we propose a novel self-embedding fragile image watermarking scheme for tampering recovery based on reference-data interleaving mechanism and adaptive selection of embedding mode. During watermark embedding, reference bits are derived from the interleaved, scrambled MSB bits of original image, and then are combined with authentication bits to form the watermark bits for LSB embedding. Different with the reported schemes with the fixed embedding mode, the proposed scheme not only has two types of embedding modes, i.e., overlapping-free embedding and overlapping embedding, but also utilizes the adaptively flexible numbers of MSB and LSB layers to achieve satisfactory performances for different tampering rates. Also, detailed analyses are given to provide the theoretical values of watermarked-image quality, perfect recovery probability, and recovered-image quality, which are used to conclude the optimal choice of embedding modes. Experimental results show the effectiveness and superiority of the proposed scheme compared with some state-of-the-arts schemes.

## 1. Introduction

In the past two decades, with the development of the Internet and multimedia processing tools, the transmission, duplication, and modification of digital contents have become much easier than before [34,5]. As a result, protecting the intellectual property contained in the multimedia data has become an important challenge [8]. In addition to properly identifying ownership [14], determining the authenticity of multimedia data and protecting their integrity also are important current issues that must be solved [21,38,16]. When the traditional cryptographic technique is used to authenticate multimedia data by attaching digital signatures, it cannot locate suspicious regions if the multimedia data that are received have been tampered during transmission [18]. Also, the computation burden associated with digital signatures for multimedia data is extremely heavy, and additional storage space is required to attach the signatures. Thus, in order to deal with these problems effectively, in recent years, researchers have proposed the technique of fragile watermarking for the authentication of multimedia data [31]. In this paper, we mainly focus on the authentication of digital images.

[*] Corresponding author at: School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Shanghai 200093, China. Fax: 86 21 55272982.
*E-mail addresses:* qin@usst.edu.cn (C. Qin), wanghuili2622@163.com (H. Wang), xzhang@shu.edu.cn (X. Zhang), sunnudt@163.com (X. Sun).

http://dx.doi.org/10.1016/j.ins.2016.09.001
0020-0255/© 2016 Elsevier Inc. All rights reserved.

One category of reported fragile watermarking schemes focused on integrity authentication and tampering detection for digital images, which are sensitive to any modifications of the images and can effectively identify locations where tampering has occurred [33,29,22,35,2,41,10,23,3,20]. The embedded watermark data for this category of fragile watermarking schemes usually are the hash of principal contents retrieved from each image pixel or block [33,29,22,35]. Since tampering manipulations destroy the matching relationship between the contents of the original image and the corresponding watermark data, the tampered regions can be detected. Chang et al. proposed a fragile watermarking scheme for image ownership and tampering authentication [2]. The aim of their scheme was to protect the rightful ownership and detect malicious manipulations of embedded images using the authentication data inserted in adaptive least significant bits (LSB) of the original pixels. Zhang and Wang proposed a statistical scheme of fragile watermarking to locate tampered regions with pixel-wise accuracy [41]. The watermark data of this scheme consisted of tailor-made authentication data for each pixel and some additional test data that can be used to reveal the exact pattern of the tampered contents. However, in many real applications, just detecting tampering is not enough, and it is highly desirable to recover the original content from the tampered regions. Therefore, many researchers have investigated ways of recovering the original content after tampering has been detected [4,17,24,15,12,6,26,25,19,40,37,7,28,13,42,36,43,39,11].

Fridrich and Goljan developed a fragile watermarking scheme with self-recovery capability, which encoded the DCT coefficients of each block into 64 or 128 bits and embedded them into the LSBs of other distant blocks [4]. This embedding strategy made the scheme can resist the collage attack. When the tampered blocks were detected, the quantized DCT coefficients were extracted from the intact regions and decoded to recover the original contents of the tampered image. In [17], Lin et al. presented a fragile watermarking method for detecting image tampering and recovering the original content, which was based on a 3-level hierarchical structure to ensure the accuracy of tampering localization. The scheme can deal with high tampering rates and obtain acceptable recovery results. However, these schemes had the problem of not being able to recover original content of the tampered regions if the hidden reference information used for the recovery also was destroyed. This situation is referred to as a tampering coincidence problem [42]. To solve this problem, Lee and Lin proposed an effective dual watermark scheme for image tampering detection and recovery in [19]. Their scheme provided two copies of watermark data for each non-overlapping block in the entire image, thereby providing a second chance for tampering recovery in case the first copy of the watermark was damaged. In [40], a tailor-made watermark consisting of reference-bits and check-bits was hidden in the original image using a reversible data hiding method [9]. On the receiver side, the extracted and calculated check-bits can be used to locate tampered image blocks. The original image can be reconstructed exactly by the reliable reference-bits extracted from other blocks. If the tampering rate is less than 3.2%, this scheme can restore the information about the original image with no errors, but, in this scheme, the visual quality of the watermarked image is relatively low. The scheme [37] reduced the encoding length of block features by establishing an index table for the original image via vector quantization (VQ). In this scheme, several copies of the VQ indices for all image blocks were embedded into the original image as watermark data according to a pseudo-random sequence. The tampered image can be recovered by the decoded VQ codewords. However, if all of the copies of the embedded watermark for the image block were damaged, the visual quality of recovered result was not very good. A self-embedding fragile watermarking scheme based on a reference sharing mechanism was proposed in [42]. In this scheme, the shared reference bits derived from the five most significant bit (MSB) layers of original image were scrambled and then embedded into the three LSB layers of the entire image. As long as the area where tampering occurred was not too extensive, sufficient available data scattered in the intact regions of image can be retrieved to recover the five MSB layers of tampered regions, effectively avoiding the tampering coincidence problem. However, the fixed embedding capacity of this scheme made the usage efficiency of watermark bits lower when dealing with variable tampering rates, and the way of reference-bits generation also caused the watermark wasting problem [39].

In this work, in order to achieve better visual quality of watermarked images and recovered images, we propose a novel, self-embedding, fragile watermarking scheme, which integratedly considers visual quality of watermarked image and recovered image from different tampering rates based on the reference-data interleaving mechanism. Different from earlier schemes, our scheme utilizes flexible numbers of the MSB layers to generate the interleaved reference bits for content recovery, and it also uses variable numbers of LSB layers to accommodate watermark bits. The embedding modes of the proposed scheme can be categorized as overlapping-free embedding and overlapping embedding. Detailed analysis and calculation of the theoretical PSNR values of the watermarked/recovered images, as well as the probability of the perfect recovery of tampered images, are given so that the optimal choice of embedding modes can be made for different tampering rates. To the best of our knowledge, our work is the first to present the relationship for the overall performance of self-embedding scheme, the embedding modes that are used, and the ranges of tampering rates. The experimental results show the effectiveness and superiority of our scheme compared with some state-of-the-art schemes.

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

## 2. Proposed scheme

In the proposed self-embedding scheme, there are two main procedures: (1) watermark embedding procedure, which embeds watermark bits (including authentication bits and reference bits) derived from MSBs of the original image into LSBs;