Journal of Information Security and Applications 40 (2018) 236–246

Contents lists available at ScienceDirect

Journal of Information Security and Applications

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

Image authentication scheme based on reversible fragile watermarking with two images

Yinyin Peng[a], Xuejing Niu[b], Lei Fu[a], Zhaoxia Yin[a,c,*]

[a] Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230601, PR China
[b] School of Computer Science, Fudan University, Shanghai 201203, PR China
[c] Department of Computer Science, Purdue University, West Lafayette 47906, United States

ARTICLE INFO

Article history:

Keywords:
Image authentication
Reversible data hiding
Fragile watermarking

ABSTRACT

Image authentication receives growing awareness because digital images can be easily modified and the modification is hard to detect. A reversible image authentication method that can improve the accuracy of tamper detection and the quality of watermarking image is presented in this paper. In the proposed method, reversible data hiding is implemented with two identical host images, where the secret information is embedded in one host image while the distortion information is embedded in the other image. In the authentication process, we compared the information extracted from the image with the original authentication information to judge whether the image was tampered with. Experimental results show that the algorithm can improve the accuracy of tamper detection and guarantee the image quality while a large number of authentication information is embedded.

1. Introduction

Digital media data is widely used in various fields and images account for about 35% of the multimedia data on the Internet. As the advantages of having rich information, being intuitive and being easy to understand, images have become the most widely used medium in our daily life. Image integrity authentication has aroused much concern, because digital images can be easily modified and the modification is difficult to be detected. Currently, more and more methods are explored to deal with the problems of image modification.

There are two criteria to evaluate the image integrity authentication algorithms, which are detection accuracy and image quality. The digital signature-based algorithms [1–3] and fragile watermarking-based authentication algorithms [4–9] are two existing types of image integrity authentication algorithms. In the digital signature-based algorithm, digital signature is saved by the third party. When authenticating, the digital signature is extracted from the image and then compared with the digital signature preserved by the third party to detect the integrity of image. The authentication methods based on fragile watermarking can be divided into semi-fragile watermarking and complete fragile watermarking method. The semi-fragile watermarking method is robust to some attacks and can distinguish the common signal processing and malicious tampering. However, the semi-fragile watermarking authentication method is insensitive to tampering. On the contrary, the complete fragile watermarking method is sensitive to tampering and any images modification would be detected, so the complete fragile watermarking method can be used for the accurate integrity authentication. The general data hiding algorithms in image spatial domain can't resist any modifications and belong to complete fragile watermarking algorithms. The information will be embedded into the host image to generate watermarking image by using the general spatial domain data hiding method. After the watermarking image is modified by some image processing software, the pixel values are changed and the information can't be extracted correctly. In this paper, the image integrity authentication method takes characteristic of its low robustness to embed the authentication information into the host image, and then the integrity of the image can be determined exactly by the integrity of the authentication information.

The data hiding algorithms that are used in embedding authentication information can be divided into two categories: the irreversible data hiding algorithms [10–17] and the reversible data hiding algorithms [18–28]. Irreversible data hiding algorithm modified pixel values of a host image to embed data. The embedding capacity is relatively large and the image quality and detection accuracy are high. However, irreversible hiding methods bring damage to the host image in itself and the host image can't be recovered after data extraction. Therefore, many reversible data hiding

* Corresponding author at: Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230601, PR China.
E-mail address: yinzhaoxia@ahu.edu.cn (Z. Yin).

https://doi.org/10.1016/j.jisa.2018.04.007
2214-2126/

algorithms have been proposed. LSB (least significant bit) [12] replacement method uses binary secret information to replace the least significant bit of host pixels directly. Then scholars put forward the LSBM(LSB matching) [13] method, commonly known as ± 1 method, which modifies the least significant bit randomly by $+1$ or -1 to improve the security. In 2006, LSB-MR (LSB matching revisited) [14] improved the efficient of LSB data hiding by using the element equation to modify pixel pairs. In the same year, EMD (exploiting modification direction) method [15] was proposed to enhance the data hiding capacity by using different directions to indicate different secret data. Later, in order to meet the demand of data hiding capacity and quality, some algorithms with large data hiding capacity and good quality were put forward. In 2009, Chao et al. proposed an adaptive data hiding algorithm, named DE (diamond encoding) [16]. This method can embed a B-nary number into a host pixel according to different payload requirements, but the image distortion is serious with the increase of capacity. In 2014, Yin et al. [17] proposed another adaptive algorithm called second-order steganographic (SOS). By using a particular matrix MB, it can embed two B-nary numbers into a pair of pixels, which not only improves the efficiency of data hiding, but also produces better data hiding images under the same payload.

A reversible data hiding algorithm based on histogram shifting was first proposed by Ni et al. [18]. The main idea is to modify the pixel values between peak and zero point of histogram produced by image to embed data into peak point. In 2013, Li et al. [20] proposed pixel-value-ordering (PVO) method. The host image is divided into non-overlapping and equal blocks. Differences between the largest and the second largest value, the smallest and the second smallest value constitute histograms. The largest value and the smallest value of each block are modified to embed data. This method achieves high embedding capacity with high quality, but it does not make full use of the image redundancy space. Peng et al. [21] proposed an improved PVO algorithm by utilizing more peak points. It achieves higher embedding capacity and better image quality. Wang et al. [22] dynamically adjusted the block size on the basis of [21], which improved the experimental results further. In addition, Chang et al. [26] used two host images to implement reversible data hiding. In this method, two 5-nary numbers are embedded into the same pixel pairs of two identical host images with a special matrix. In order to hide more information, Lyu et al. [27] proposed a high capacity data hiding method based on two host images and the method also use a special matrix so that the two 9-nary numbers and a 4-nary number can be hidden in two pixel pairs, while maintaining low distortion rate. Although the data hiding capacity of these two methods is large and quality is good, their adaptability is not strong.

In 2017, Zhang et al. [29] proposed a robust forgery detection algorithm which achieves very good detection accuracy. In order to further improve the accuracy of tampering detection as well as the marked image quality, in this paper, a reversible image tampering authentication method based on high-capacity reversible data hiding algorithm is proposed, which uses two watermark images. In the proposed method, reversible data hiding is implemented with two identical host images, where the secret information is embedded in one host image while the distortion information is embedded in the other image. In the authentication process, we compare the information extracted from the image with the original authentication information to judge whether the image has been tampered with. Experimental results demonstrate that the proposed algorithm can reduce the image distortion and improve the accuracy of tamper authentication while embedding a large number of authentication information. Image redundancy can be exploited more fully and more flexible by using adaptive high-capacity reversible data hiding with two marked images.

The rest of this paper is organized as follows. In Section 2, there is a brief introduction of the SOS method and two compared methods. In Section 3, a detailed description of the proposed method is given. Experimental results are shown in Section 4. Finally, Section 5 concludes the paper.

2. Related works

In this section, the SOS method is introduced in Section 2.1 and then the two reversible data hiding algorithms based on two images are described in Sections 2.2 and 2.3.

2.1. SOS algorithm

The SOS algorithm [17] is an irreversible data hiding algorithm. The matrix MB used by the SOS algorithm varies according to the secret digital B, but these matrices have some basic commonalities: they are $B \times B$ in size and contain an unrepeatable combination of all possible two B-nary numbers. When $B = 3$, as shown in Fig. 1(a), the left and right of the matrix are connected and roll into a barrel. From the view of the row, there are all combinations of 00, 01, 02... 21, 22. The matrix MB of $B = 4$ and $B = 5$ is also given in Fig. 1(b) and (c).

For a pair of host pixel pairs (p_1, p_2) and two B-nary secret numbers $(d_1, d_2)_B$, the SOS method modifies the pixel pair (p_1, p_2) to (p'_1, p'_2) using maximum amplitude $\lfloor B/2 \rfloor$, enabling $MB(p'_1\%B, p'_2\%B) = d_1$ and $MB(p'_1\%B, (p'_2 + 1)\%B) = d_2$ to be embedded. The $(d_1, d_2)_B$ can be obtained by calculating $MB(p'_1\%B, p'_2\%B)$ and $MB(p'_1\%B, (p'_2 + 1)\%B)$ directly during the extracting process.

In short, with the help of a specific $B \times B$ matrix MB, the SOS algorithm can embed two B-nary digits into a pair of pixels, which not only improves the efficiency of data hiding, but also produces better data hiding images under the same payload. Therefore, on the basis of SOS algorithm, the rational using of matrix MB is a breakthrough of adaptive high capacity reversible data hiding.

2.2. Chang et al.'s method

Chang et al. achieved reversible data hiding with two host images. Chang et al. [26] used a special matrix to embed two 5-nary numbers into two identical pixel pairs of the same host image, each of which hides a 5-nary number. At this point, the payload reaches $log_2 5/2$ bpp. Chang et al.'s algorithm uses a matrix which is similar to the one produced under the condition of $n = 2$ in EMD. The matrix $M1$, as shown in Fig. 2, can be generated by Eq. (1), where p_1 and p_2 are two-pixel values and % represents modulo operations (the same below).

$$y = (p_1 + 2p_2)\%5, p_1, p_2 \in [0, 255] \quad (1)$$

It is observed that the diagonal of each 5×5 block in the matrix of Fig. 2 is a non-repeating 5-nary number and each diagonal can hide a 5-nary number. Assume that there are two same pixel pairs (p_1, p_2), (q_1, q_2) and two 5-nary numbers $(d_1, d_2)_5$ with two identical host images. In matrix $M1$, take (p_1, p_2), (q_1, q_2) as the center of the block, modify (p_1, p_2) to (p'_1, p'_2) along the main diagonal, make $(p'_1 + 2p'_2)\%5 = d_1$ to hide d_1, modify (q_1, q_2) to (q'_1, q'_2) along the sub-diagonal, make $(q'_1 + 2q'_2)\%5 = d_2$ to hide d_2. Using Eq. (1), according to the modified pixel (p'_1, p'_2) and (q'_1, q'_2), $(d_1, d_2)_5$ can be extracted directly. For the (p'_1, p'_2), (q'_1, q'_2) that are main diagonal and secondary diagonal in the matrix $M1$, the coordinate intersection in $M1$ is in accordance with the pixel values of the original image, and then the original pixel can be restored from the host.

The algorithm of Chang et al. mainly embeds the fixed B-nary number. Compared with other methods, it is easier to understand,