A Semi-Fragile Watermarking for ROI Image Authentication

Delong Cui
College of Computer and Electronic Information
Maoming University
Maoming China
delongcui@163.com

Abstract—In this paper, a watermarking algorithm for ROI (Regions of Interesting) authentication and tamper detection is proposed. The watermarking is embedded in the DWT (Discrete Wavelet Transform) coefficients of the ROB (Regions of Background). The algorithm uses a hashing to detect whether the ROI of image is tampered, and then uses the watermarking to reconstruct the data of ROI. The experiment results show that the embedded watermark is perceptually invisible and robust to various image processing operations. Most importantly, the tampered place can be identified accurately.

Keywords- digital watermark; ROI; wavelet transform; DPCM; ECC; SHA1

I. INTRODUCTION

With the rapid development of computers and internet, copyright protection of digital media as image, video and audio becomes more and more important. Digital watermarking is a new resolution for copyright protection and media security, and it is becoming a very active research area for information security. Digital watermarking is the technology which embeds copyright information including product ID, text content or image logo into multimedia for copyright protection, secret communication and discrimination of the data files. Perceptual invisibility, robustness and safeness are the digital watermarking system’s central characteristics.

There have been a large number of authentications watermarking schemes proposed for tamper detection and content integrity verification of multimedia content [1]. In watermarking, the authentication data are embedded in the digital media with a secure and efficient way. The watermarking techniques for content authentication are usually referred to fragile/semi-fragile watermarking. The existing fragile/semi-fragile watermarking techniques differ mainly in the way to embed the authentication data into the host data [2]. The embedding techniques proposed include spatial-domain embedding and transform-domain (DCT, wavelet, etc.) embedding. Many fragile/semi-fragile watermarking schemes benefit from the transform domain embedding by preserving moderate robustness to compression attacks. Wu and Liu [3] proposed fragile watermarking scheme based on modifying the quantized DCT coefficients according to a special LUT. Xie and Arce’s technique [4] firstly implements the signature in DWT (Discrete Wavelet Transform) domain and then embeds the signature using SPIHT compression algorithm. Kundur and Hatzinakos [5,6] embedded the pseudo-random sequences by modifying the quantization process of Haar wavelet coefficients.

In this paper, a watermarking algorithm for ROI (Regions of Interesting) authentication and tamper detection is proposed. Specially, we make use of the data of ROI as watermarks. The authentication does not require any a priori information about the original watermark as well as the original image. The remaining of this paper is organized as follows. The concept of image DWT (Discrete Wavelet Transform) is briefly introduced and torus automorphism is provided in section II. The framework and the details of the proposed scheme are described in section III. In section IV, some experimental results are presented. A conclusion is drawn in section V.

II. PRELIMINARIES

A. DWT

DWT is identical to a hierarchical sub bands system, where the sub bands are logarithmically spaced in frequency and represent the octave-band decomposition. The image is first divided into four sub bands LL1, LH1, HL1, HH1 by cascading horizontal and vertical two channels critically sub sampled filter banks. The sub bands labeled LH1, HL1, and HH1 represent the finest scale wavelet coefficients. To obtain the next coarse scale of wavelet coefficients, the sub band LL1 is further decomposed and critically sub sampled. The process continues until some final scale is reached.

B. Torus automorphism

Chaos technique has been applied in secure field, which owns both confusion and diffusion properties. There is a well-known technique called torus automorphism. The equation is depicted as follows:

\[ A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \mod N \]  \tag{1}

where \((x,y)\) and \((x',y')\) represent the pixel location before and after the transformation, \(a, b, c,\) and \(d\) represent the secrete
key. If the $A$ has period $T$, the pixel location $(x, y)$ can return to the original location after $T$ times of the $A$ operation. The period $T$ is thus termed the orbit. If the pixel point $(x, y)$ has been performed $k$ times of the $A$ operation, it needs $T-k$ times of the $A$ operation to make $(x, y)$ return to the original location.

III. PROPOSED SEMI-FRAGILE WATERMARK EMBEDDING SCHEME

A. Watermark embedding algorithm

To describe the proposed watermark embedding algorithm, the ROB (Regions of Background) of Lena image is used as the host image; and ROI pixels data are used as the watermark, as shown in Fig. 1. Watermark embedding scheme, which includes the following steps:

Step 1: Generate the watermark: coded the ROI into binary sequences, using the DPCM (Differential Pulse Code Modulation) and ECC (Error Correct Code) technique. The torus automorphism technique is also used to enhance the security of the watermarking system.

Step 2: Discrete wavelet transform of the image: The original image is decomposed to $L$-level wavelet coefficients. Select the sub bands of ROB to embed watermark using a key. Sub bands coefficient and its root-father have the position relation as shown in Fig. 2.

Step 3: Key generation: a random sequencing order is generated in order to disperse the embedding locations of the watermark samples. The watermark bits are to be embedded based on the randomly selected locations.

Step 4: Watermark embedding: the detailed watermark embedding technique is presented in [7].

Embedding value 0: Suppose 0 is embedding into the $5^{th}$ bit of the coefficient, the modify value and original coefficient might have the difference of $2^7=32$. As shown in Fig. 3, Value 1 is equal to $(282)_{10}$. In order to lower image loss, we chose a value 0 to replace the $5^{th}$ bit. Since there are too many values to choose from, we will only consider 3 conditions. The first one is Value 1. The second one is to change the $6^{th}$ bit into 1 in Value 2, while the $4^{th}$, $3^{rd}$ and $2^{nd}$ are changed to 0 with the $1^{st}$ and $0^{th}$ bits flipped form 0 to 1 or vice versa depending on its original value. Value 2 is equal to $(321)_{10}$. The third one is to change the $4^{th}$, $3^{rd}$ and $2^{nd}$ to 1 with the $1^{st}$ and $0^{th}$ bits flipped as in Value 2. Here, Value 3 is equal to $(286)_{10}$. Finally, we will choose a value closest to the original coefficient from those 3 condition to replace and to get the smallest image lossy possible.

| value 1: | 1 0 0 0 1 1 0 1 0 |
| value 2: | 1 0 1 0 0 0 0 0 1 |
| value 3: | 1 0 0 0 1 1 1 1 0 |

MSB  LSB

Figure 3. The 3 candidate conditions for embedding 0

Embedding value 1: The rules for modifications here are similar to embedding value 0. As shown in Fig 4, for value 1 is embedded into the 5th bit giving $(314)_{10}$. It is coincidental that the embedded value happens to be the original value. Our test will perform well for other test data, too. The $6^{th}$ bit in Value 2 is modified to 0 and the $5^{th}$ to $2^{nd}$ bits are modified to 1 with the $1^{st}$ and $0^{th}$ bits flipped between 0 and 1 depending on its original values. This gives the value $(317)_{10}$. Whereas in Value 3, the $5^{th}$ bit is modified to 1 and the $4^{th}$ to $2^{nd}$ bits are modified to 0 while the $1^{st}$ and $0^{th}$ bits are flipped giving $(318)_{10}$. The one with the closest value to the original coefficient will be chosen.

| value 1: | 1 0 0 1 1 1 0 1 0 |
| value 2: | 1 0 0 1 1 1 0 1 0 |
| value 3: | 1 0 0 1 1 1 1 0 0 |

MSB  LSB

Figure 4. The 3 candidate conditions for embedding 1

Step 5: Inverse wavelet transform: the final watermarked image is obtained by an inverse DWT using Haar bases.

Step 6: SHA1 [8] (Secure Hash Algorithm 1) is employed to enhance the proposed algorithm. The hashing value of ROI is appended to the watermarked image for authentication. Fig. 5 shows the flowchart of the watermark embedded procedure.
B. watermark extraction and verification scheme

This section presents the proposed blind watermark extraction scheme. The nice feature of our proposed semi-fragile watermarking scheme is that the original watermark is not necessarily required during authentication.

Extraction of the embedded watermark bits: based on the watermark embedding locations provided by the key. The extracted watermark bits then reconstruct the ROI of the received image.

Content authentication: in the case when the original hashing is available at the authentication stage, the authentication is performed by simple comparisons between the extracted hashing and the original hashing. Compared with the two hashing, if they are equal, that means the ROI of the image is through the authentication; else the extracted watermark bits are then converted back to the watermark sample.

IV. PERFORMANCE EVALUATION: ROBUSTNESS VERSUS TAMPER SENSITIVITY

The watermark extraction performance of the proposed watermarking scheme against common signal processing is evaluated. The sensitivity or the robustness of the proposed scheme is investigated under various attacks in the following scenarios: JPEG compression, histogram equalization (uniform distortion), sharpening (low-pass filtering processed), blurring (high-pass filtering processed), additive Gaussian noise, salt-and-pepper noise (value set to 255 or 0, respectively), as show in Fig. 6.

The watermark extraction of the watermarked Lena images is tested under the abovementioned manipulations. The performance is measured in terms of the BER (Bit Error Rate) of the extracted watermark bits, which is the ratio of bits in errors, $Ne$, relative to the total number of the hidden watermark bits, $Nw$, given by

$$BER=\frac{Ne}{Nw}$$  \hfill (2)

The BER of the extracted watermark of all the wavelet decomposition L-level are analyzed. Table I shows the BERs. Table II shows the PSNR (peak signal to noise ratio) of the watermarked image under various signal processing attacks.

The PSNR is defined as follows:

$$PSNR = 10 \log_{10} \left( \frac{2^8 - 1}{MSE} \right) (dB)$$  \hfill (3)

$$MSE = \frac{1}{m \times n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (X_{ij} - X'_{ij})^2$$  \hfill (4)

where $X_{ij}$ and $X'_{ij}$ represents the pixel location before and after the attacked image.

<table>
<thead>
<tr>
<th>Attacks</th>
<th>BERs(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No attacks</td>
<td>0</td>
</tr>
<tr>
<td>JPEG comp.</td>
<td>2.75</td>
</tr>
<tr>
<td>Gaussian noise</td>
<td>7.46</td>
</tr>
<tr>
<td>Salt-and-pepper noise</td>
<td>2.88</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attacks</th>
<th>PSNR(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No attacks</td>
<td>36.61</td>
</tr>
<tr>
<td>JPEG comp.</td>
<td>28.94</td>
</tr>
<tr>
<td>Histogram equalization</td>
<td>27.89</td>
</tr>
<tr>
<td>Sharpening</td>
<td>27.78</td>
</tr>
<tr>
<td>Blurring</td>
<td>28.97</td>
</tr>
<tr>
<td>Gaussian noise</td>
<td>29.47</td>
</tr>
<tr>
<td>Salt-and-pepper noise</td>
<td>35.93</td>
</tr>
</tbody>
</table>

V. CONCLUSION

This paper presents a wavelet domain semi-fragile watermarking scheme. The hashing of ROI is employed as authentication watermark. More significant bits are embedded in the sub bands of the wavelet decomposition of the ROB. In particular, a secure embedding zone is exploited such that all the embedded wavelet coefficients are adjusted to fall into this secure zone. It has shown that the proposed scheme successfully reconstructs the data of ROI which are through malicious modulations. The sensitivity of the watermark under various attacks is also demonstrated.

The nice feature by using semi-fragile watermark in authentication is that no original watermark is required. Thus, arbitrary watermark can be used for different images. Based on the embedding locations provided by the key, the embedded watermark can be extracted by the image process of the wavelet coefficients for the received content. In the absence of the original watermark, the magnitude of the extracted watermarking is tested to detect and locate the possible alterations in the content. Two potential applications of our proposed semi-fragile watermarking scheme are presented. The first one is for content integrity verification digital medicine image, satellite remote sensing image, which have high quality of ROI. The second one is the blind quality assessment for the received watermarked image. The semi-fragile authentication watermarking that can sustain moderate degree of compressions needs to pay attention for further research.
Figure 6. Various attacks and watermark extraction

REFERENCE:


