Content Authentication Based on JPEG-to-JPEG Watermarking

Hong-Xia Wang, Jie Hou, and Ke Ding

Abstract—A content authentication technique based on JPEG-to-JPEG watermarking is proposed in this paper. In this technique, each 8×8 block in a JPEG compressed image is first processed by entropy decoding, and then the quantized discrete cosine transform (DCT) is applied to generate DCT coefficients: one DC coefficient and 63 AC coefficients in frequency coefficients. The DCT AC coefficients are used to form zero planes in which the watermark is embedded by a chaotic map. In this way, the watermark information is embedded into JPEG compressed domain, and the output watermarked image is still a JPEG format. The proposed method is especially applicable to content authentication of JPEG image since the quantized coefficients are modified for embedding the watermark and the chaotic system possesses an important property with the high sensitivity on initial values. Experimental results show that the tamper regions are localized accurately when the watermarked JPEG image is maliciously tampered.

Index Terms—Chaotic map, compressed domain, content authentication, digital watermarking.

1. Introduction

In recent years, one of the most popular image formats used is JPEG (joint photographic experts group) on the Internet and in storage media such as CD-ROM and DVD since the JPEG format can achieve high compression while retaining high image quality. The JPEG-to-JPEG (J2J) image watermarking is a watermark technique that if the input is a JPEG image file, the output is also a JPEG file after watermark embedding. Hence J2J watermarking is suitable for embedding watermarks in JPEG compressed images. However, many existing fragile watermarking schemes which embed the watermark in the uncompressed images may not be suitable for authentication of the JPEG compressed images (.jpg files). There are a few existing schemes for J2J watermarking. One problem of embedding watermarks in JPEG compressed images is that the watermarking scheme needs to be JPEG compatible. This implies that all DCT coefficients need to be re-quantized with the same quantization factor after the watermark insertion. If the output images are not JPEG compatible, the existence of the watermark may be detectable using steganalysis techniques. While the output images are JPEG compatible which is the J2J framework, obviously the efficiency of watermark processing will be greatly improved due to the watermark is added directly into the JPEG compressed domain rather than going through decoding, watermarking, and re-encoding process.

Many watermarking algorithms were proposed to embed digital watermarks in the uncompressed images\cite{1} to \cite{3}, but those algorithms may not be suitable for JPEG compressed images (.jpg files). Especially for the content authentication technique, the watermark is often embedded into bitmap BMP uncompressed image\cite{4} to \cite{5}. Because the quantization and re-quantization of discrete cosine transform (DCT) coefficients carry some unavoidable quantization error, there will be great influence on the content-based watermark and the tamper detection if the watermark is processed directly in the spatial domain of JPEG lossy compressed image. Consequently, the content authentication for JPEG image can not be implemented. Besides, it is well known that a BMP image has larger storage capacity than JPEG image, which is not suitable for the bandwidth requirement of the watermarking application. Hence compressed domain watermarking schemes are more attentive to resolve these problems\cite{6} to \cite{8}.

In the present paper, we develop a content authentication technique based on J2J image watermarking. The watermark is added directly to the quantized DCT coefficients. Thus the watermark can be embedded directly into a JPEG file, and the output is also a JPEG file. The simulation results show our scheme can directly localize the tamper regions on the watermarked JPEG images by using the high sensitivity on initial values in the chaotic system.

Manuscript received July 7, 2008; revised September 25, 2008. This work was supported by the National Natural Science Foundation of China under Grant No. 60702025, the Research Fund for the Doctoral Program of Higher Education under Grant No. 20070613024, and Sichuan Youth Science & Technology Foundation under Grant No. 07ZQ026-004.

H.-X. Wang is with School of Information Science & Technology, Southwest Jiaotong University, Chengdu, 610031, China (e-mail: hxwang@home.swjtu.edu.cn).

J. Hou is with School of Computer Science & Engineering, University of Electronic Science and Technology, Chengdu, 610065, China.

K. Ding is with Marvell Semiconductor Inc., Shanghai, 200030, China (e-mail: kding1@marvell.com).
2. Watermark Embedding

For color JPEG images, the watermark is embedded only in the luminance (Y) component of the image. The JPEG compression process consists of conversion of RGB (red/green/blue) to YCbCr color image, composition of two-dimensional discrete cosine transform (2-D DCT), quantization of DCT coefficients, run-length coding, and Huffman coding. In this paper, watermark embedding and extracting will be performed on quantized DCT coefficients. As shown in Fig. 1, for each 8×8 block of quantized DCT coefficients, the AC components are selected to embed the watermark. The least significant bit (LSB) plane of the quantized AC coefficients is set to zero, and $f_{AC}$ is obtained. Each $f_{AC}$ is mapped into the defining field of initial values $x_0$ of a chaotic system by using the mapping $C$ as follows:

$$ C: Y \rightarrow X, \quad X, Y \in \mathbb{R}, \quad x_0 = C(\mu, f_{AC}), \quad \mu \in \mathbb{R}, \quad f_{AC} \in Y, \quad x_0 \in X $$

(1)

where $\mu$ is the parameter to control $f_{AC}$ into the defining field of $x_0$. Obviously, the mapped initial value $x_0$ is closely connected with each frequency coefficient of an image. That is, the frequency coefficient of an image is modified, corresponding $x_0$ will be different from the original one.

It is well known that one of the important properties of chaos is its sensitivity of initial value, i.e. a very small change in the initial state of a chaotic system can cause a large effect on its later state. Based on this property of chaos, we can use $x_0$ as an initial value which is associated with the quantized DCT coefficients to generate the fragile watermark for authenticate image purpose. In our work, the initial value $x_0$ mapped by $f_{AC}$ is performed $n$ times iteration by a chaotic system, and a real chaotic sequence $\{x_n \mid x_n \in X, \quad n = 0, 1, 2, \cdots \}$ is obtained. The number of iteration $n$ is random, which is decided by the decimal random number generated by a chaotic system whose initial value $K$ is used as the key of watermarking system. After binary quantization to the real chaotic sequence, the $n$th element is used as the watermark bit $w(i, j) \in \{0, 1\}$. Namely,

$$ w(i, j) = Q_{\text{binary}}(x_n \mid x_0 \in X), \quad 1 \leq i, j \leq 8 $$

(2)

The watermark bit is embedded into the LSB plane of the quantized AC coefficients in each 8×8 block. Thus the watermarked quantized DCT coefficients $f_{AC}'$ are obtained. When all of 8×8 blocks are embedded by the watermark information, the watermark embedding is implemented. After embedding procedure by Huffman coding and run-length coding, the watermarked JPEG image is obtained.

3. Watermark Extraction and Authentication

The watermark extraction will be performed in the compressed domain. This means that our embedding process can skip inverse quantization, inverse DCT, and YCbCr-to-RGB conversion in JPEG decoding procedure. After decoding Huffman code and run-length code, the 8×8 blocks with watermarked quantized DCT coefficients are obtained and the watermark bits are extracted from the LSB plane of the AC components to form a watermark version. On the other hand, like watermark embedding process, the quantized DCT coefficients whose LSB plane is set to zero are applied to generate another watermark version by chaotic iteration. Finally, a comparison can be made between these two versions of the watermark.

If the content of JPEG image is tampered, $f_{AC}'$ will be correspondingly changed to $f_{AC}''$. Assume $f_{AC}'$ is mapped to an initial value $x_0$, then $f_{AC}''$ is mapped to another initial value $x_0'$. Obviously, an error will exist between $x_0$ and $x_0'$:

$$ |x_0 - x_0'| = \delta $$

(3)

Let $V$ be a set, and $F: V \rightarrow V$ is chaos on $V$. According to the properties of sensitive dependence on the initial conditions in the chaotic system, $\exists \delta > 0$ for any $x_0 \in X$, and a neighborhood $\Omega$ of $x_0$, and $\exists x_0' \in \Omega$ for $n \geq 0$, such that

$$ |F^n(x_0) - F^n(x_0')| > \delta $$

(4)

where $F^n$ denotes chaotic iteration $n$ times. Now we define

Fig. 1. The process of watermark embedding.
a tampered matrix $\mathbf{T}(i, j)$ by

$$
\mathbf{T}(i, j) = \begin{cases} 
0 & \text{image is not tampered} \\
1 & \text{image is tampered}
\end{cases}
$$

(5)

and

$$
\mathbf{T}(i, j) = \left\| W(x_i^0) - \text{LSB}(f_{AC}(w_x^0)) \right\|, \quad \mathbf{T}(i, j) \in \mathbb{T}_{8 \times 8}, \quad 1 \leq i, j \leq 8
$$

(6)

where $W(x_i^0)$ represents the watermark generated by tampered image and $\text{LSB}(f_{AC}(w_x^0))$ represents the watermark extracted by LSB plane of watermarked AC coefficients $f_{AC}(w_x^0)$. If $W(x_i^0) = \text{LSB}(f_{AC}(w_x^0))$, we have

$$
\mathbf{T}(i, j) = 0
$$

(7)

It means that the image is not tampered, else, the non-zero $\mathbf{T}(i, j)$ is the tampered region. After all of $8 \times 8$ blocks in an image are detected, the tampered region is obtained, and the image authentication is completed. From the process of authentication, it can be seen that the watermarking system is sensitive to the tamper on the JPEG image.

Let $\mathbf{D}_{8 \times 8}$ denote an $8 \times 8$ image block, and assume some pixels $p(i, j) \in \mathbf{D}_{8 \times 8}$, $(1 \leq i, j \leq 8)$ are maliciously tampered, then the whole corresponding $8 \times 8$ DCT coefficient block will be different from the original one. So the watermark bits generated by the $8 \times 8$ block are usually different from those watermark bits extracted from the LSB plane. Therefore, the tampered localization precision is based on $8 \times 8$ block in our scheme. That is, even if only one pixel in an $8 \times 8$ block is tampered, the whole $8 \times 8$ image block is deemed to be tampered.

### 4. Simulation Results

In order to test the performance of the proposed scheme, the standard color JPEG image of size $512 \times 512$ ‘Pepper’ and ‘Airplane’ is used as the test images, as shown in Fig. 2 (a) and (b), respectively. The chaotic system that used in the test is a logistic map described by

$$
x_n = 1 - 2x_n^2, \quad x_n \in [-1, 1], n = 1, 2, \ldots
$$

(8)

Fig. 3 (a) and (b) show the JPEG images watermarked by the proposed scheme, which is implemented by using the simulate tool VC++ 6.0. Obviously, the embedded watermark does not visually degrade the quality of the image. The peak signal-to-noise ratio (PSNR) for the Pepper and Airplane is 47.73 dB and 46.98 dB, respectively.

Fig. 4 (a) and (b) show the images tampered through addition and modification to Fig. 3 (a) and (b) under Photoshop 6.0. Two papers are tampered into the original watermarked ‘Pepper’ image, but the watermarked ‘Airplane’ image is modified from ‘F16’ to ‘F18’ on the type of plane. For such tampering, our scheme can authenticate the tampered regions, as shown in Fig. 5 (a) and (b). It is clear that the localizations of tampered regions are very accurate in the authenticated JPEG images. By using a computer with Intel P4 2.66GHz CPU and 512M RAM, the runtime of the proposed scheme for authenticating ‘Pepper’ and ‘Airplane’ images of size $512 \times 512$ pixels are 1.09 s and 1.02 s, respectively. But in the case that the watermark is embedded in spatial domain, the runtime will be 1.98 s and 1.91 s. Obviously, the efficiency is greatly improved to operate directly watermark in compressed domain.

The performance comparison between the proposed scheme and related works are shown in Table 1. The invisibility is evidently enhanced by the proposed scheme because the watermark in the scheme is embedded into the LSB plane of AC component of the quantized DCT coefficients.

![Fig. 2. Original JPEG images: (a) original ‘Pepper’ and (b) original ‘Airplane’.](image)

![Fig. 3. Watermarked JPEG images: (a) Pepper, (b) Airplane, and (c) Airplane by magnification.](image)

![Fig. 4. Tampered JPEG images: (a) Pepper, (b) Airplane, and (c) tampered Airplane.](image)
Fig. 5. Tampered regions obtained by content authentication: (a) Pepper and (b) Airplane.

The performance comparison between the proposed scheme and related works are shown in Table 1. The invisibility is evidently enhanced by the proposed scheme because the watermark in the scheme is embedded into the LSB plane of AC component of the quantized DCT coefficients.

Table 1: Performance comparison between the proposed scheme and related works

<table>
<thead>
<tr>
<th></th>
<th>PSNR(dB) Pepper</th>
<th>PSNR(dB) Airplane</th>
<th>JPEG decoding steps for authentication</th>
<th>Localization precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>47.73</td>
<td>46.98</td>
<td>Huffman decoding</td>
<td>8×8 block</td>
</tr>
<tr>
<td>scheme</td>
<td></td>
<td></td>
<td>Huffman decoding+inverse quantization+inverse DCT</td>
<td></td>
</tr>
<tr>
<td>Ref. [1]</td>
<td>≈37.9</td>
<td>≈37.9</td>
<td>Huffman decoding+inverse quantization</td>
<td>Pixel point</td>
</tr>
<tr>
<td>Ref. [8]</td>
<td>36.97</td>
<td>35.68</td>
<td>Huffman decoding+inverse quantization</td>
<td>8×8 block</td>
</tr>
</tbody>
</table>

5. Conclusions

The proposed J2J watermarking scheme for content authentication avoids the quantization error from JPEG quantization and re-quantization process. Experimental results demonstrate that the JPEG format image can be accurately located where the tampering has occurred. The potential applications of the scheme are the authentications of JPEG images on Internet network and trust-worthy digital cameras. The future research could emphasize methods of more precise localization technique for smaller block or point unit.

Acknowledgment

The authors would like to thank Southwest Jiaotong University Development Foundation for the financial support under Grant No. 2006A04.

References


Hong-Xia Wang was born Hebei Province, China, in 1973. She received the B.S. degree from Hebei Normal University, Shijiazhuang, in 1996, and the M.S. and Ph.D. degrees from University of Electronic Science and Technology of China (UESTC), Chengdu, in 1999 and 2002, respectively. She engaged in postdoctoral research work in Shanghai Jiaotong University from 2002 to 2004. Currently she is a professor with School of Information Science and Technology, Southwest Jiaotong University, Chengdu. Her research interests include multimedia information security, intelligent information processing, and the coding technique. She has published 30 peer research papers and won 6 authorized patents.

Jie Hou was born in Jilin Province, China, in 1974. She received the B.S. degree from Sichuan University, Chengdu, in 1997. She is currently pursuing the M.S. degree in software engineering with UESTC, Chengdu, China. Her research interests include computational intelligence and encryption theory with application.

Ke Ding was born in Zhejiang Province, China, in 1980. He received the B.S. degree from the Zhejiang University, Hangzhou, in 2002 and the M.S. degree from Shanghai Jiaotong University, Shanghai in 2005. Currently, he is an engineer in Marvell Semiconductor Inc., Shanghai. His research interests include digital image and audio signal processing.