Fragile Watermarking to detect change of small range on image

Hye-Joo Lee*, Yun-Hee Oh**, Ji-Hwan Park***, Kwangjo Kim*

*Cryptology and Information Security Lab., Information and Communications Univ.

**Dept. of Computer & Information Science, Pukyong Nat'l Univ.

*** Dept. of Computer & Multimedia Engineering, Pukyong Nat'l Univ.

Abstract

Fragile watermarking is a technique for authentication/integrity of digital data. Unlike robust watermarking, it has to design to be vulnerable against some slight processing to verify the modification of digital data. Feature of fragile watermarking is to identify the modifications of data and to locate some places modification occurred at the same time, so it has to identify slight changes of small range if possible. In this paper, fragile watermarking is proposed that the changes of small range on image can be identified using the watermark sequence with period and the values of low bit planes in an image.

1. Introduction

The techniques to protect multimedia data have been researched with increasing use of multimedia data. In particular, digital watermarking, which embeds a watermark as special information into multimedia data, has been proposed for protecting multimedia since 1990[1]. Digital watermarking techniques are classified into two categories, one is a technique for copyright protection of multimedia and the other is for authentication/integrity. The former means robust watermarking that the watermark must be preserved in multimedia data after processing such as compression, filtering, etc[2]. The latter is fragile watermarking that the watermark is embedded into multimedia data when creating multimedia, and it allows creator to identify some changes of data by destruction of watermark. Fragile watermarking embeds a verification data such as cryptographic hash value of image or random bits into image, and then verifies the modification of multimedia data by extracting the verification data.

In this paper, fragile watermarking is proposed that it can verify some changes in small range of size 8×8 at least for digital image. In section 2, we introduce the concepts about fragile watermarking and the existing schemes. A proposed method is described that the periodic watermark consisting of –1 and 1 is embedded into low bit planes using DCT(discrete cosine transformation) in section 3. In section 4, we simulate our proposed method to evaluate its efficiency, and future work is described in section 5.

2. Related Works

The requirements for fragile watermarking are as follows:

1) It must to detect slight forgery/modification on the image.
2) It must to locate the places in which the modification occurred.
3) It must to extract the watermark W from the watermarked image I' without the original image I.

4) The watermark has to be imperceptible.

Specially, as features of fragile watermarking, two requirements of 1) and 2) are distinct from robust watermarking

A method proposed by Wolfgang et al. embedded the watermark W of M-sequence of pseudo-random binary sequence into block b of size 8×8 as

$$Y(b)=X(b)+W(b), \qquad (1)$$

and a watermarked image Y is obtained[3]. To verify the modification by detecting the watermark,

$$\delta(b)=Y(b)\cdot W(b)-Z(b)\cdot W(b) \qquad (2)$$

are computed for a possibly modified watermarked image Z. Given a threshold value T, if $\delta < T$, then the image Z is not modified. This method is not required the original image, but it has disadvantage that the original watermarked image Y is needed as specified in Equation (2).

A method proposed by Kundur et al. embedded the watermark using L-level discrete wavelet decomposition of image[4]. Some coefficients $f_{k,l}$ to embed a binary watermark of length N, $w(i)$, $i=1,\cdots,N$, are randomly selected using $ckey(i)$ in the same level and then the watermark is embedded using function Q as following:

If $Q(f_{k,l}(m,n))=w(i)$, then coefficients are not modified, where function Q is defined as

$$Q(f)=\begin{cases}0, \text{ if } r\Delta\le f<(r+1)\Delta \text{ for } r=0,\pm2,\pm4,\cdots\\ 1, \text{ if } r\Delta\le f<(r+1)\Delta \text{ for } r=\pm1,\pm3,\pm5,\cdots\end{cases}. \qquad (3)$$

Otherwise, coefficients are modified by

$$f_{k,l}(m,n)=\begin{cases}f_{k,l}(m,n)+\Delta \text{ if } f_{k,l}(m,n)\le 0\\ f_{k,l}(m,n)-\Delta \text{ if } f_{k,l}(m,n)>0\end{cases}. \qquad (4)$$

Inverse wavelet transform is applied to this result, the watermarked image is obtained. For detecting the watermark, $\widetilde{w}(i)$ is calculated from the value of wavelet coefficients using function Q. The determination of modification applies tamper assessment function,

$$\mathrm{TAF}(w,\widetilde{w})=\frac{1}{N_w}\sum_{i=1}^{N_w} w(i)\oplus\widetilde{w}(i) \qquad (5)$$

to the original watermark and the extracted one. If $\mathrm{TAF}(w,\widetilde{w})\ge T$, $0\le T\le 1$, then we determine that the image is modified. In Kundur's method, there is a case that if an attacker modified the image and then the coefficients are modified using $\Delta' << \Delta$ while satisfying the function Q, it is impossible to determine the modification of image.

3. Fragile Watermarking Using DCT of low bit planes

When an image is separated to each bit plane, a meaning about the image is clear as being high bit plane. Namely when considering a bit plane consists of low bit, it appears almost random sequences. Embedding watermark using high bit planes does not satisfy the imperceptibility of watermark because it has an influence on information about the image. Embedding watermark in low bit planes does not the robustness of watermark but fragile watermarking allows it because the robustness is not required. Thus, a method is described that the watermark is embedded into low bit plane using DCT.

The proposed method is performed as follows. An image I of size $N\times M$ is partitioned into block B_{ij} of size $W\times H$, $1\le i\le\lfloor N/W\rfloor, 1\le j\le\lfloor M/H\rfloor$. Using block pixel values, $b=(b_1,b_2,\cdots,b_K)$, $K=W\times H$ arranged in form of one dimension for block B_{ij}, the watermark are embedded as illustrated in Figure 1.

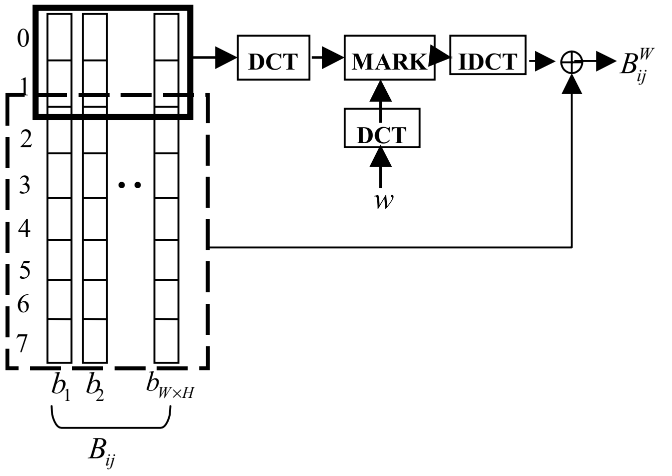


Figure 1. Watermarking Process

Let values consist of two low bits extracted from pixel values b_k of block B_{ij} as Figure 1 denote $b' = (b'_1, b'_2, \cdots, b'_K)$, and let DCT coefficients calculated by applying DCT to these values denote $\hat{b} = (\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_K)$.

Embedding process of watermark performs that DCT is applied to the watermark w and then these coefficients are added to the DCT coefficients of pixel values, $\hat{b}$. The watermark $w = (w_1, w_2, \cdots, w_L)$ consists of two values, –1 and 1 as the periodic sequence, and length L of w sets to be shorter than block size as

$$L = \left\lfloor \frac{W \times H}{2} \right\rfloor, \tag{6}$$

for satisfying the imperceptibility of watermark.
The coefficients of watermark, $\hat{w} = (\hat{w}_1, \hat{w}_2, \cdots, \hat{w}_L)$ are added to $\hat{b}$ as

$$\tilde{b}_{l+p} = \hat{b}_{l+p} + \hat{w}_l \text{, where } 1 \le l \le L \tag{7}$$

from $p(p > 1)$-th coefficient among $\hat{b}$. By adding the watermark w as Equation (7) and performing IDCT(inverse discrete cosine transform), $b'' = (b''_1, b''_2, \cdots, b''_K)$ is constructed. The watermarked block B^W_{ij} is obtained by combining this value b'' with high bit planes of the original block B_{ij}.

The extraction of watermark is as follows. First, by extracting low bit planes and applying DCT to it, L DCT coefficients are obtained. The obtained coefficients are different from the original coefficients because 'Round' operation in IDCT are performed during embedding process, so let the obtained coefficients be $\bar{b} = (\bar{b}_1, \bar{b}_2, \cdots, \bar{b}_L)$.
For the embedded coefficients, $\hat{w}$, similarity s with the extracted coefficients $\bar{b}$ is computed as

$$s = \frac{(\bar{b} \cdot \hat{w})}{\sqrt{(\hat{w} \cdot \hat{w})}}. \tag{8}$$

Given threshold value T, if $s > T$, then the watermark w is present in block. Because the watermark is affected by simple processing on image, it is possible to locate the distorted places by searching block of $s \le T$.

4. Simulation and Results

To verify an efficiency of the proposed method, we simulated the proposed method with image 'Lena(size 256×256, 8bits/pixel)' of Figure 2.

Figure 2. Original image(Lena)

Figure 3 shows each bit plane of two high bit planes and two low ones for Figure 2. As shown in Figure 3, high bit planes provide more information about image than low bit plane.

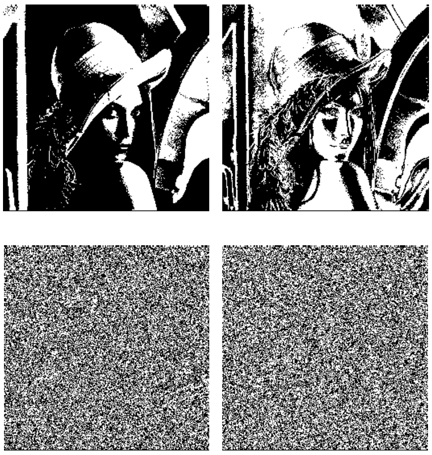
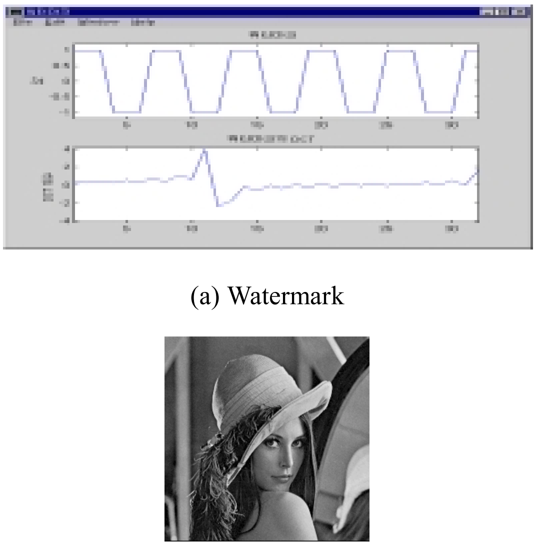


Figure 3. Bit planes for an original image

Therefore, a method for fragile watermarking is desirable to use low bit plane for satisfying the imperceptibility of watermarking inserting and for easily destructing the watermark by various processing at the same time.

In simulation, the size of block is set as 8×8. Using the watermark w of Figure 4(a), DCT is applied to it and coefficients of w are added to the coefficients of block from $p = 2$, and the watermarked image is shown in Figure 4(b). The SNR(signal-to-noise) for the watermarked image is measured at 41.29[dB].



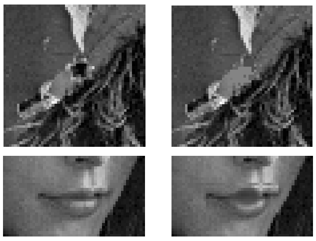
(a) Watermark

(b) A watermarked image

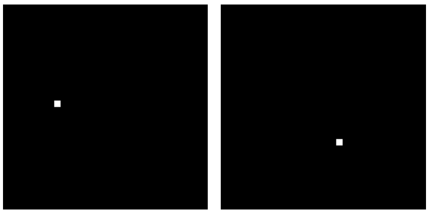
Figure 4. A watermark and watermarked image

In this paper, a reason to use periodic sequence rather than random sequence as the watermark is as follows. When the random sequence is used for the watermark, shape of DCT coefficients appears almost random. This increases the secrecy for the watermark, but clearness of watermark during detection is decreased. On the other hand, the periodic watermark of Figure 4(a) is clear about a shape of DCT coefficients for the watermark. Thus, the detection result of watermark can be improved if distorted by any processing.

The proposed method is applied to two cases making alterations on the watermarked image. Figure 5(a) and Figure 5(b) are zoomed images distorted in a small range on image, one removes ornament of hat and the other makes alteration on a lip, respectively.



(a) Modified image 1 (b) Modified Image 2



(c) The location of modification

Figure 5. Modification of the watermarked Image

To determine the modification of image, the threshold value sets as $T = 2.0$. As a result, it is possible to locate the distorted place in spite of difficulty not to identify the modification of small range by eyes as shown in Figure 5(c).

5. Conclusion

In this paper, we proposed a method that the watermark is embedded into low bit planes for satisfying the imperceptibility and fragileness of the watermark. It is not replacement of low bit planes with the watermark but by applying DCT to the watermark and low bit planes, respectively, the watermark is spread out low bit planes. Also using periodic sequence as watermark not random sequence is allowed to identify a modification of very small range. As future work, research about secrecy of the proposed method is needed for attack that modifies the watermarked image so as to detect the watermark after modification.

[References]

1. Security and Watermarking of Multimedia Contents, P. W. Wong & E. J. Delp Editors, Proc. of SPIE, Volume 3657, 1999
2. J. Cox, J. Killian, T. Leighton, and T. Shammon, “Secure spread spectrum watermarking for multimedia,” NEC Research, Inst., Tech. Report. 95-10, 1995
3. R. B. Wolfgang, E. J. Delp, “Fragile watermarking using the VW2D watermark,” Security and Watermarking of Multimedia Contents, Proc. of SPIE, Vol. 3657, pp.204-213, 1999
4. D. Kundur, D. Hatzinakos, “Digital Watermarking for Telltale Tamper Proofing and Authentication,” Proc. of the IEEE, Vol.87, No.7, pp.1167-1180, 1999