# A SEMI-BLIND ROBUST WATERMARKING FOR DIGITAL IMAGES

*He Quan, Su Guangchuan*

505 Lab, Dept. of Electronic Engineering, Beijing Institute of Technology, 100081, Beijing, China
E-mail: wjhequan@sina.com

**ABSTRACT**

A new robust watermarking scheme based on triplet of wavelet coefficients is proposed. Triplets of wavelet coefficients are defined then selected and classified into two classes to embed and extract the watermark according to the relationship of their three wavelet coefficients. The shuffle algorithm is also adopted to randomize the watermark sequence for the purpose of security. Experimental results show that the degradation caused by the proposed algorithm is very slight, and this watermarking scheme is very robust to common image processing and malicious attack.

## 1. INTRODUCTION

With the fast development of multimedia technology and rapid growth of network distributions of images and video, there is an urgent need for copyright protection against pirating. Digital watermarking has been proved to be an effective way to protect the ownership of multimedia data. Digital watermarking can be described as the process of embedding, by means of a secret key, an imperceptible digital signal(the watermark) into multimedia content. To be effective a good watermarking scheme should satisfy following requirements: transparency, robustness, universal and unambiguous [1]. There exist two basic classes of electronic watermarks: fragile and robust. One major difference between watermarking techniques is whether or not the watermark detection or extraction step requires the original image. Watermarking techniques that do not require the original image during the extraction process are called *oblivious*(or blind) watermarking techniques [2].

The general method of traditional frequency-based watermarking algorithm is to add the watermark on perceptual significant DCT or DWT coefficients [3-4]. Watermarking energy is limited by the requirement of HVS(Human Visual System) to achieve the balance between transparency and robustness. But the embedding of watermark will inevitably cause the degradation of the host image. Although Cox *et.al.* point out that perceptual significant coefficients are less affected during the process of image processing and attack, traditional watermark detector will be seriously degraded under severe image processing or attack, especially geometrical distortions such as cropping, rotation, and affine transform, etc. The computation of JND(Just Noticeable Difference) is also a burden for watermarking process. Our proposed watermarking scheme casts watermarks on DWT domain; the embedding and extraction of watermark is based on the relationship of three wavelet coefficients in the selected triplets of wavelet coefficients. The modification of wavelet coefficients is very slight so that good transparency is achieved. And the original host image is not needed during the watermark extraction process. This algorithm is also relatively easy to implement. Experimental results show that this watermarking scheme is very robust to common image processing and attacks.

## 2. PROPOSED METHOD

Triplet of wavelet coefficients is defined as the three wavelet coefficients in three detail orientation(HL, LH, HH) of one certain wavelet decomposition level, as illustrated in Figure 1. The basic idea of proposed algorithm is to map 1 bit of a watermark to a selected triplet of wavelet coefficient, and the value of wavelet coefficients is modified according to the bit that will be embedded. Embedding and extraction of watermark is based on the relationship of three coefficients in the selected triplet of wavelet coefficients. The watermark can be a pseudo-random sequence or a binary image.

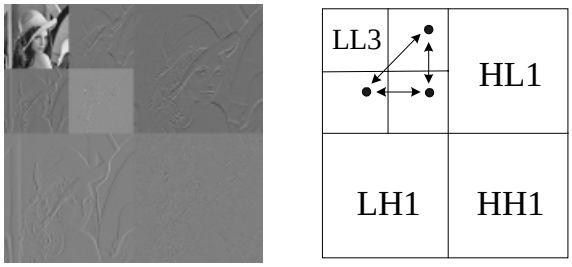



Figure 1. Illustration of triplet of wavelet coefficients

Triplets of wavelet coefficients are selected and classified into following two classes according to the watermark bit that will be embedded:

Class I：

1. $|c(i,j)| > T_1$

where $T_1 = \alpha \max\{|c_1(i,j)|, |c_2(i,j)|, |c_3(i,j)|\}$.

2. $std(c_1(i,j), c_2(i,j), c_3(i,j)) > T_2$

here $c(i,j)$ denotes the value of wavelet coefficient, $c_1(i,j), c_2(i,j), c_3(i,j)$ denotes the three coefficients in a selected triplet of wavelet coefficients, and $std(\cdot)$ denotes the computation of standard deviation.

Class II：

1. $c(i,j) \notin I \qquad (I \cap II = \phi)$
2. $std(c_1(i,j), c_2(i,j), c_3(i,j)) < T_3$
3. $T_3 < T_2$

here $T_1$, $T_2$, $T_3$ is three thresholds related to the host image. The two value of a watermark (1 and –1, or 1 and 0) will be marked by different class of triplet of wavelet coefficients(here 1 is marked by class I). The value of each wavelet coefficient is modified by the following rule [5]:

· embed 1:

$$c_1(i,j) = c_1(i,j) + \lambda_1 \times c_1(i,j) \quad (1)$$

$$c_2(i,j) = c_2(i,j) + \lambda_2 \times c_2(i,j) \quad (2)$$

$$c_3(i,j) = c_3(i,j) + \lambda_3 \times c_3(i,j) \quad (3)$$

where $|c_1(i,j)| \le |c_2(i,j)| \le |c_3(i,j)|$; $0 \le \lambda_1 \le \lambda_2 \le \lambda_3$;

· embed –1:

$$c_1(i,j) = c_2(i,j) = c_3(i,j) = mean(c(i,j)) \quad (4)$$

where $mean(c(i,j))$ denotes the mean value of three wavelet coefficients in a triplet of wavelet coefficients. Choosing proper $\lambda$ will decrease the degradation caused by the watermarking process. Triplets of wavelet coefficients will be selected from these two classes to embed each bit of the watermark, and a position table will be generated at the same time for future extraction. The original host image is not required during the extraction process, so this watermarking scheme is semi-blind. To improve the security of proposed system, shuffle algorithm is adopted to randomize the watermark sequence [6]. Without the correct shuffle table the embedded watermark cannot be correctly extracted.

The extraction of watermark is based on the position table referred above. When $std(c_1(i,j), c_2(i,j), c_3(i,j)) > T_4$ ($T_3 < T_4 < T_2$), 1 is extracted, and –1 vice versa. If the embedded watermark is a binary image, the similarity value will be computed for evaluation; if a pseudo-random sequence is embedded, the correlation analysis will be performed:

$$\rho = \frac{\sum w(i,j)w'(i,j)}{\sum w^2(i,j)} \quad (5)$$

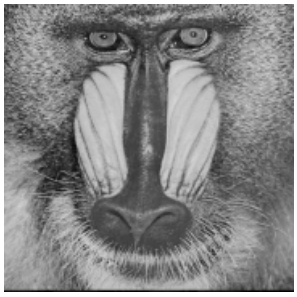
were $w(i,j)$ is the original watermark, and $w'(i,j)$ is the extracted watermark. In the case of pseudo-random sequence as a watermark: when $\rho$ is greater than a threshold $T$, the watermark is considered extracted correctly. Given the positive false probability $P_{fp}$, the threshold $T$ is determined by the following equation [7]:

$$P_{fp} = \sum_{m=[N_w(T+1)/2]}^{N_w} \binom{N_w}{m} 0.5^{N_w} \quad (6)$$

## 3. EXPERIMENATL RESULTS

### 3.1. Binary image as a watermark

The image we used in this experiment is the standard test image Baboon (512×512).

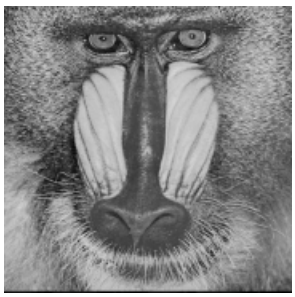


(a) Host image, Baboon (512×512)

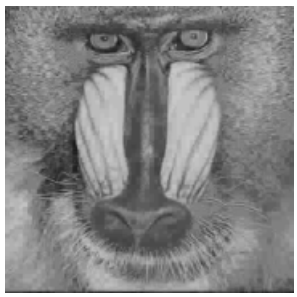

(b) Signature image(64×64)

(c) Shuffled signature image

Figure 2. Host image(a), signature image(b), and shuffled signature image(c)



(a) Watermarked image, PSNR =52.0949

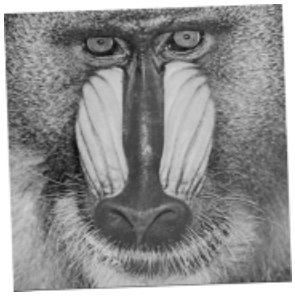


(b) JPEG lossy compression, QF=10%, CR=16.5:1

(a*) $\rho$ = 1

(b*) $\rho$= 0.8909

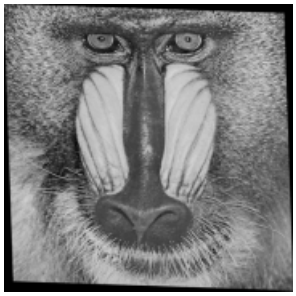
(c) JPEG2000 lossy compression, CR=150:1



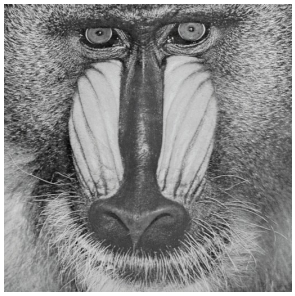
(d) Rotate (Widdershins, 2 ° )

(c*) $\rho$= 0.7383

(d*) $\rho$= 0.7595



(e) Affine transform T = [3 0.1; 0.1 3; 0 0];



(f) Stirmark with random geometric distortions

(e*) $\rho$= 0.7854

(f*) $\rho$ = 0.8188

Figure 3. Watermarked image (a), watermarked image after various image processing and attacks (b-f), and corresponding extracted signature image (a*-f*) with similarity measurement

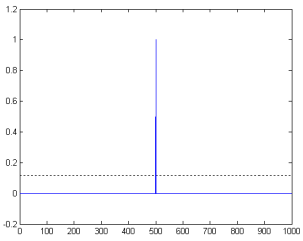
Figure 3. shows the extracted signature from watermarked image Baboon after various modifications, where QF = Quality Factor, and CR = Compression Rate.

### 3.2. Pseudo-random sequence as a watermark

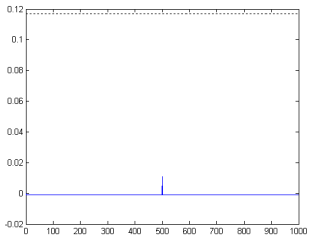
The length of watermark is $N_w = 1023$ . Given the positive false probability $P_{fP} = 0.008\%$ , the threshold is $T = 0.117$ . Because the high frequency component of an image is easy to remove during the process of image processing and attack, and the low-pass band is very important for the quality of reconstructed host image, the middle frequency component is exploited in our experiment.

(a) Host image, Lena (512×512)

(b) Watermarked image, PSNR =73.7610



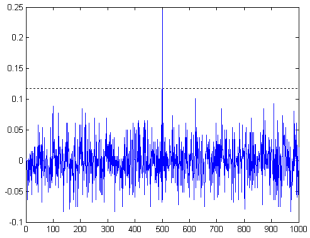
(a*) Watermarked image (Unmodified)



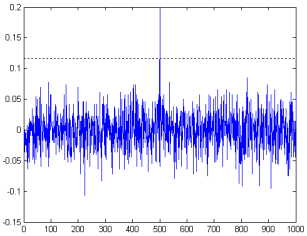
(b*) Without reverse shuffle process

(c) Image cropping (6.25% left), and median filtering(3×3)

(d) Rotate(Clockwise,5 ° ) , and Gaussian blur(2×2)



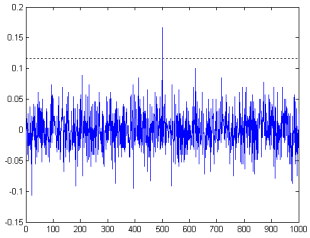
(c*) Detector response



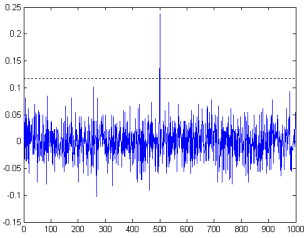
(d*) Detector response

(e) Affine transform, T=[1, 0.1; 0.1, 1; 0, 0], and JPEG lossy compression, QF = 10%, CR = 32.3:1

(f) JPEG2000 lossy compression (120:1), and Gaussian noise~N(0,0.01)



(e*) Detector response



(f*) Detector response

Figure 4. Host image(a), watermarked image after various image processing and attacks (b-f), and corresponding detector responses (a*-f*)

Figure 4. shows the watermark detector responses to the extracted watermark from the watermarked Lena after various modifications and other 999 randomly generated watermarks. Only one peak(the 500$^{th}$) exceeds the threshold *T*, which corresponds to the original watermark. To extract the watermark correctly, all watermarked host images that are attacked by the geometric transforms are rescaled to their original size(512 × 512). From the simulation results we can see that the extracted watermark that are not reverse shuffled is almost not correlative to the original watermark. The experimental results also show that the proposed algorithm presents good robustness to common image processing and malicious attacks: the watermark detector still has good response under severe image processing operations, miscellaneous attacks and geometric distortions.

## 4. CONCLUSION

A watermarking scheme based on the triplets of wavelet coefficients is proposed. Experimental results show that this watermarking system has good transparency property, and is very robust to the image processing operations and attacks such as median filtering, JPEG lossy compression, JPEG2000 lossy compression, Gaussian blur, and geometric distortions such as cropping, rotate and affine transform, etc.

## 5. REFERENCES


[1] I. Cox, J. Kilian, F. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," IEEE Trans. on Image Proceeding, vol.6, no.12, pp.1673-1687, 1997.

[2] R. Chandramouli and N. Memon, "How many pixels to watermark?," Proc. of IEEE International Conference on Information Technology: Coding and Computing, pp. 11-15, March 2000.

[3] Christine I. Podilchuk, and Wenjun Zeng, "Image-Adaptive Watermarking Using Visual Models", IEEE Jounal on Selected Areas in Communications, vol.16, no.4, pp.525-539, May 1998.

[4] X. G. Xia, C. G. Boncelet, G. R. Arce, "A Multiresolution Watermark for Digital Images," ICIP'97, pp.548-551, 1997.

[5] Davoine. F., "Comparison of Two Wavelet Based Image Watermarking Schemes," Proc. of ICIP, pp.682-685, Vancouver, Canada, Sept. 2000.

[6] M. Wu, B. Liu, "Digital Watermarking Using Shuffing", Proc. of IEEE International Conference on Image Processing (ICIP ' 99), vol.1, pp.291-295, Kobe, Japan, Oct. 1999.

[7] D. Kundur and D. Hatzinakos, "Digital Watermarking using Multiresolution Wavelet Decomposition," Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Seattle, Washington, vol. 5, pp. 2969-2972, May 1998.