Improved Detection for Robust Image Watermarking

Corina Nafornita[1]
[1] "Politehnica" University of Timisoara, Communications Dept., Timisoara, Romania, e-mail corina@etc.utt.ro

***Abstract*— In a previous paper, the author proposed a robust watermarking method for still images that embeds the binary watermark into the detail subbands of the image wavelet transform. The perceptually significant coefficients are selected for each subband using a different threshold. For greater invisibility of the mark, the approximation subband is left unmodified. The watermark is embedded several times in each subband to achieve robustness. Here, we propose a new type of detection and we test the performance against different types of attacks (lossy compression, AWGN, scaling, cropping, intensity adjustment, filtering and collusion attack).**

I. INTRODUCTION

Protection of multimedia transmitted over the Internet can be made through encryption and watermarking. Encryption makes the multimedia data unintelligible, therefore protecting it in transmission over insecure channels, while watermarking embeds into the host data in some invisible way a signal called watermark that is supposed to identify the owner [1]. Important properties of an image watermarking system are [2-3]: *perceptual transparency* (the watermarking process should not degrade the image significantly), *robustness* (resistance of the mark against intentional or unintentional attacks like AWGN, filtering, lossy compression, scaling, cropping), and *data hiding capacity* (the amount of information that can be embedded into the original cover work without causing serious distortions).

Most of existing watermarking systems proposed in the literature can be classified depending on the watermarking domain, where the embedding takes place: spatial domain techniques [4], where the pixels are directly altered, or transform domain techniques. Popular transforms are the Discrete Cosine Transform (DCT) [5], the Discrete Wavelet Transform (DWT) [6-11,13], and the Discrete Fourier Transform (DFT) [12].

In this paper, we present a watermarking technique in the DWT domain, for copyright protection purposes. A new type of detection is proposed. The detection is non-blind, thus increasing the probability of detecting the watermark.

II. PREVIOUS WORK

Several papers that deal with copyright protection for images argue that the mark should be embedded in some transform domain selecting only perceptually significant coefficients (PSCs), because those are the most likely to survive compression [5-8]. Cox et al. [5] embed a continuous watermark in the largest 1000 DCT coefficients of the original image, except the DC coefficient, thus spreading its energy on several bins of frequency. Detection is made using the similarity between the two watermarks.

Xia et al. [6] insert several watermarks in the DWT domain in each detail image, except the approximation subband, suggesting that the detection could be done hierarchically, computing crosscorrelations of the watermark and the difference between the two images for each resolution level.

Other authors [7, 8] embed the watermark into perceptually significant coefficients for each subband of the DWT using statistical properties of the human visual system (HVS) and of the original image.

Nafornita [11] proposes a technique that embeds the watermark into the wavelet domain, into perceptually significant coefficient using subband adaptive thresholding. The watermark is embedded repeatedly into the detail subbands, thus increasing the robustness of the method. An average of the extracted watermarks is computed at the detector.

III. BRIEF DESCRIPTION OF THE PROPOSED METHOD

In two-dimensional separable dyadic DWT, each level of decomposition produces four bands of data, one corresponding to the low pass band (LL), and three other corresponding to horizontal (HL), vertical (LH), and diagonal (HH) high pass bands. The decomposed image shows a coarse approximation image in the lowest resolution low pass band, and three detail images in higher bands. The low pass band can further be decomposed to obtain another level of decomposition. This process is continued until the desired number of levels determined by the application is reached. Taking into account the fact that the HVS is not sensitive to small changes in high frequencies of the image, but rather sensitive to changes affecting the smooth parts of the image (the coarsest resolution level of the image), the embedding of the same watermark is made several times into the HH, HL and LH detail images, leaving the LL band unaffected.

A. Embedding the watermark

Consider X the original gray-level image and the watermark W a pseudo random binary sequence, of length N_w with $w(i) \in \{-1, 1\}$. The image is decomposed into L resolution levels using the DWT, thus obtaining for each resolution level "l", three detail subbands HH_l, HL_l, LH_l and one approximation subband (last level) LL_L. The watermark is repeatedly embedded of M>>1 times in the transform image. Each repetition is denoted by W_r, with $r = 1,2,...,M$. This can be viewed as a form of transmitting the watermark in different subchannels. It has been shown by Kundur et al. in [13] that diversity techniques can give very good results in detecting the watermark, considering the fact that many watermark attacks are more appropriately modeled as fading like.

This work was financed by a grant from the National Council of Scientific Research and Education, Romania, CNCSIS code 47 TD.

Roughly speaking, the current watermark bit $w_r(i)$ is embedded at the location (m, n) of subband s, level l if the wavelet coefficient $d_{s,l}(m, n)$ is higher than a subband dependent threshold $T_{s,l}$. The watermarked coefficient is given by

$$d^w_{s,l}(m,n) = \begin{cases} d_{s,l}(m,n)[1+\alpha \cdot w_r(i)], & \text{if } d_{s,l}(m,n) > T_{s,l} \\ d_{s,l}(m,n), \text{otherwise} \end{cases}, \quad (1)$$

where α is the embedding strength, $r = 1,2,\ldots,M$ and

$$T_{s,l} = q_l \max_{m,n} \{ d_{s,l}(m,n) \}. \quad (2)$$

The watermarked image X^w is computed with the IDWT from the new coefficients. It is obvious that the higher the strength of the mark α, and the lower the parameters q_l are, the more robust yet visible the watermark will be.

B. Detecting the watermark

The detection requires the original watermark and the original image, or some significant vector extracted from its wavelet transform, specifically in this case, the detail coefficients with a value above the computed threshold for each subband. The watermark bit $\hat{w}_r(i)$ is obtained from the wavelet coefficient $\hat{d}_{s,l}(m,n)$ of the possibly distorted image $\hat{X}^w$, and the original wavelet coefficient $d_{s,l}$(m, n):

$$\hat{w}_r(i) = \text{sgn}\left(\frac{\hat{d}_{s,l}(m,n) - d_{s,l}(m,n)}{d_{s,l}(m,n)} \right). \quad (3)$$

A random guess is made for the watermark bit in the location (m, n) if the two coefficients are equal or if $d_{s,l}(m, n)$=0. In [11] extraction of the watermark is made using the majority rule: the most common bit value is assigned for the recovered watermark bit. This is done from all levels (detector type I) or from level 3 since the lowest frequencies are not so affected by compression (detector type II). The correlation coefficient compares the original and the extracted mark:

$$c(W,\hat{W}) = \frac{\sum_{i=1}^{N_w} w(i)\hat{w}(i)}{\sqrt{\sum_{i=1}^{N_w} w^2(i)} \cdot \sqrt{\sum_{i=1}^{N_w} \hat{w}^2(i)}} \quad (4)$$

where $\hat{w}(i) = \text{sgn}\left(\sum_r \hat{w}_r(i)\right)$ and $w_r(i)=w(i)$. If the correlation coefficient is above a specified threshold, the watermark is positively detected in the image. We consider that if the watermark length is large enough, setting the threshold at 0.5 will not result in large probability of false negative.

The third detector extracts every estimate $\hat{W}_r$ of the original watermark, and computes the correlation coefficient of $\hat{W}_r$ and W_r (where $W_r=W$). The highest correlation value will result in the most likely estimate $\hat{W}_r$ of the embedded watermark. The embedding and extraction procedure are shown in Fig. 1 and 2.

IV. Simulation Results

We performed simulations using the test image Peppers, size 256 x 256, and a 256-bits watermark. The Daubechies 10pt wavelet was used to produce the wavelet coefficients. In all tests we used the following parameters: the number of resolution levels L = 3, the level-dependent parameters q_1 = 0.06, q_2 = 0.04, q_3 = 0.02, and the strength of the watermark α = 0.2. Specifically, we affected 8448 coefficients from a total of 65536 (including the LL subband). The repetition number of the original watermark was for this image M=33. Human observers cannot make a distinction between the original and the watermarked image. The distortion introduced by the watermark can be measured with the peak signal-to-noise ratio PSNR, in this case 40.28 dB.

To prove the robustness of the new type of detection, we investigate the effect of common signal distortions on the correlation coefficient between the original and the recovered mark and compare the new performances with the results obtained using the method previously proposed in [11].

In Table I we have the detector response for the three types of detectors, when the watermarked Peppers image is attacked by lossy compression (JPEG, compression rate 15 and JPEG2000, compression rate 10 and 15); AWGN with the SNR = 11.4 dB, rescaling to half of the image, median filtering, filter size 3, intensity adjustment, cropping. We can clearly see that detector II yields in higher performances in the case of lossy compression, median filtering and scaling, whereas detector I has better results in the case of AWGN attack.

In the cropping attack, the two types of detector have the same results. In the case of intensity adjustment the watermark is not detected.

The 3rd detector is improved compared to the first one; the 2nd has similar or better results than the 3rd except in the cases where the image is cropped or in the case of intensity adjustment.

In Table II we have the detector response for the collusion attack: when four watermarked images are averaged. It is obvious that the 3rd detector works better than the first two, because its output is dependent of the original watermark. In other words, the third detector searches the most resembling watermark to the original.

In Fig. 3-10 we give for the 3rd detector the correlation values as a function of 1000 randomly generated watermarks. Only the 500th watermark should be positively detected, except in the collusion attack where watermarks 200, 400, 600 and 800 should be detected. We also give the values of the PSNR between the distorted images and the watermarked image.

Conclusions

We proposed a new type of detection for a robust wavelet-based watermarking method that embeds the mark in a transparent manner. The embedding system transmits the watermark over many subchannels, in the hope that at least one of it will survive the attacks. Employing diversity can yield in

better results when the distortions are unpredictable (cropping, filtering etc.). The new detector is more resistant against cropping, intensity adjustment and collusion attacks.

TABLE I. COMPARISON BETWEEN THE THREE TYPES OF DETECTION FOR VARIOUS ATTACKS

Attack vs. detector response	**Detection type**		
	I	***II***	***III***
JPEG compression, CR = 14.85	0.21	0.78	0.69
Median filtering, filter size 3	0.13	0.82	0.81
Resizing, 256->128->256	0.03	0.45	0.31
AWGN attack, SNR = 11.4 dB	0.82	0.57	0.49
Cropping ½	0.42	0.44	0.64
Intensity adjustment	0	0.22	0.31
JPEG 2000, CR=10	0.56	0.93	0.85
JPEG 2000, CR=15	0.24	0.78	0.64

TABLE II. COMPARISON BETWEEN THE THREE TYPES OF DETECTION FOR THE COLLUSION ATTACK (AVERAGING FOUR WATERMARKED IMAGES)

Detection type vs. detector response	**Original watermark**			
	W_1	W_2	W_3	W_4
Type I	0.35	0.25	0.41	0.49
Type II	0.36	0.30	0.39	0.44
Type III, W = W_1	**0.47**	0.35	0.30	0.38
Type III, W = W_2	0.37	**0.40**	0.29	0.42
Type III, W = W_3	0.39	0.30	**0.42**	0.38
Type III, W = W_4	0.28	0.35	0.36	**0.49**

REFERENCES

[1] G. Voyatzis, I. Pitas, "Problems and Challenges in Multimedia Networking and Content Protection," *TICSP Series* No. 3, Editor Iaakko Astola, March 1999.

[2] I. Cox, M. Miller, J. Bloom, *Digital Watermarking*, Morgan Kaufmann Publishers, 2002.

[3] M. Borda, I. Nafornita, "Digital Watermarking – Principles and Applications," *Proc. Of Int. Conf. Communications* Bucharest, 2004, pp. 41-54.

[4] N. Nikolaidis, I. Pitas, "Robust Image Watermarking in the Spatial Domain," *IEEE Trans. Signal Processing*, Vol. 66, No. 3, pp. 385-403, 1998.

[5] I. Cox, J. Killian, T. Leighton, T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Trans. Image Processing*, Vol. 6, No. 12, pp. 1673-1687, 1997.

[6] X. Xia, C. G. Boncelet, and G. R. Arce, "Wavelet transform based watermark for digital images," *Optics Express*, Vol. 3, No. 12, 1998, pp. 497-505.

[7] J.R. Kim and Y.S. Moon, "A Robust Wavelet-Based Digital Watermarking Using Level-Adaptive Thresholding," *Proc. of IEEE ICIP*, Vol. 2, Kobe, Japan, Oct. 1999, pp. 226-230.

[8] B. S. Kim, K. K. Kwon, S. G. Kwon, K. N. Park, K. N. Park, K. I. Song, K. I. Lee, "A robust wavelet-based digital watermarking using statistical characteristic of image and human visual system," *Proc. of ITC-CSCC* 2002, vol. 2, pp. 1019-1022.

[9] C. Nafornita, A. Isar, "Digital Watermarking of Still Images using the Discrete Wavelet Transform," *Scientific Bulletin of Politehnica University of Timisoara*, Electronics and Telecommunications, tom 48(62), fascicola 1, 2003, pp. 73-78.

[10] C. Nafornita, M. Borda, A. Kane, "A Wavelet-Based Digital Watermarking using Subband Adaptive Thresholding for Still Images," *microCAD 2004*, Miskolc, Hungary, 18 – 19 March 2004, pp. 87 - 92.

[11] C. Nafornita, "A Wavelet-Based Watermarking for Still Images," *Scientific Bulletin of Politehnica University of Timisoara*, Electronics and Telecommunications, tom 49(63), fascicola 2, 2004, *Symposium of Electronics and Telecommunications ETc 2004*, 22 - 23 October 2004, Timisoara, pp. 126-131.

[12] M.Ramkumar, A.N. Akansu, A.A.Alatan, "A Robust Data Hiding Schemes for Images Using DFT," *IEEE International Conference on Image Processing*, II, pp. 211-215, October 1999.

[13] D. Kundur, D. Hatzinakos, "Diversity and Attack Characterization for Improved Robust Watermarking," *IEEE Trans. Signal Processing*, Vol. 49, No. 10, 2003, pp. 2383-2396.

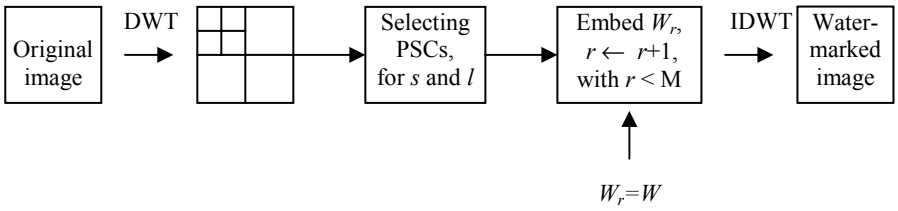

Figure 1. Embedding part

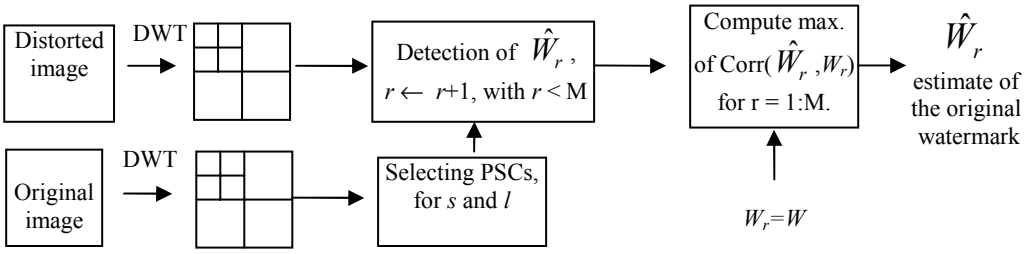

Figure 2. Detection part, type 3

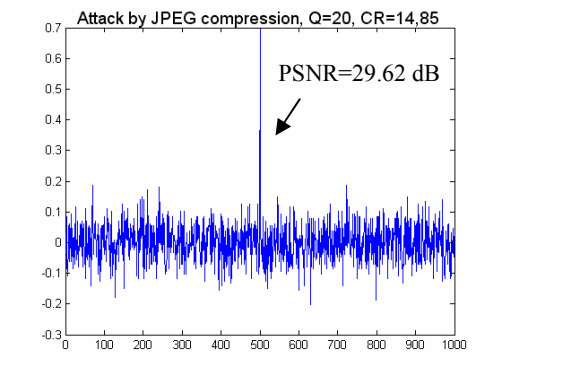

Figure 3. Detector response to 1000 randomly generated watermarks.

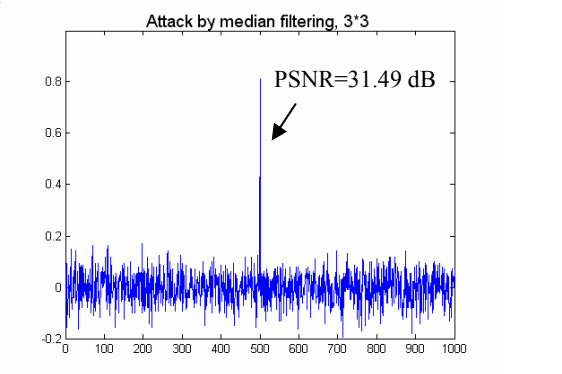

Figure 4. Detector response to 1000 randomly generated watermarks.

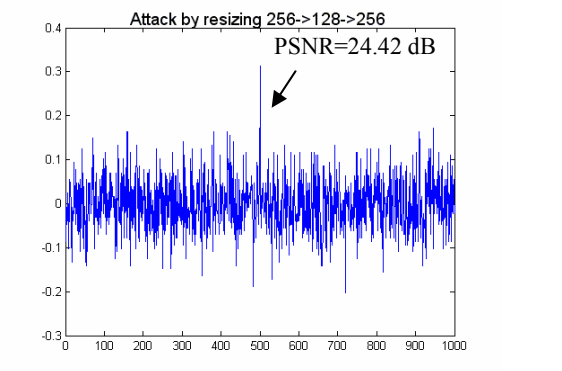

Figure 5. Detector response to 1000 randomly generated watermarks.

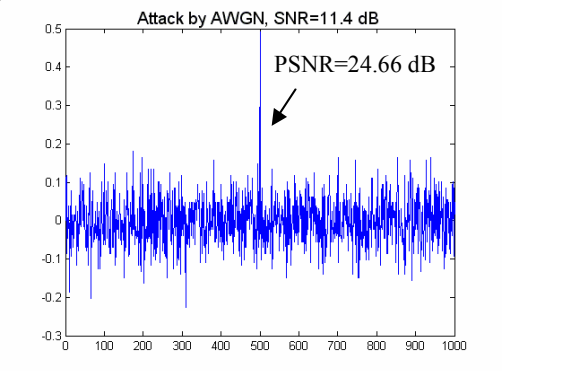

Figure 6. Detector response to 1000 randomly generated watermarks.

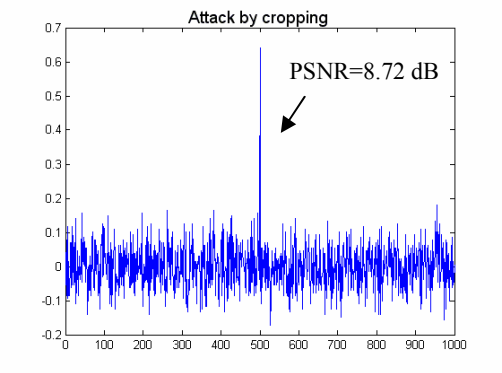

Figure 7. Detector response to 1000 randomly generated watermarks.

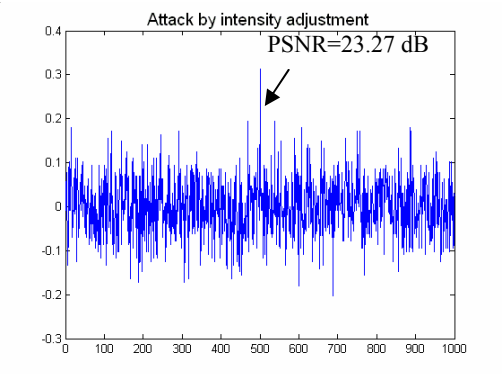

Figure 8. Detector response to 1000 randomly generated watermarks.

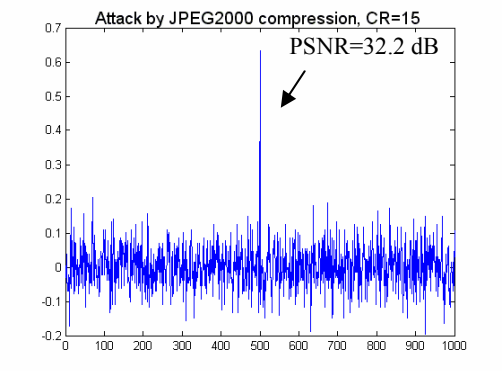

Figure 9. Detector response to 1000 randomly generated watermarks.

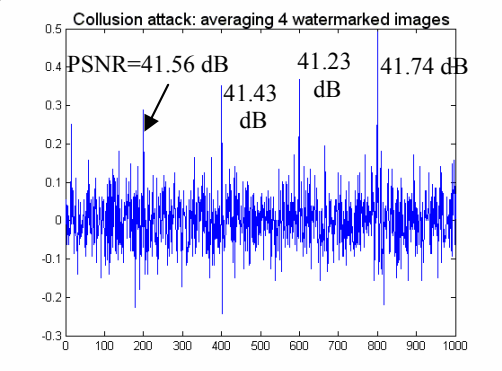

Figure 10. Detector response to 1000 randomly generated watermarks.