SPATIALLY ADAPTIVE WAVELET THRESHOLDING FOR IMAGE WATERMARKING

Mahmood Al-khassaweneh and Selin Aviyente

Department of Electrical and Computer Engineering,
Michigan State University, East Lansing, MI
e-mail:{alkhassa, aviyente}@egr.msu.edu

ABSTRACT

In this paper, we introduce a new robust image watermarking technique based on the Discrete Wavelet Transform (DWT). The proposed method extends the concept of image denoising to watermarking. A spatially adaptive wavelet thresholding method is used to select the coefficients to be watermarked. A multi-bit watermark is embedded into the discrete wavelet coefficients of the host image. A semi-blind watermark extraction algorithm is presented and the threshold for a given probability of false alarm is derived. The simulation results show that the proposed method outperforms a well-known DWT based watermarking method under most attacks including JPEG compression.

1. INTRODUCTION

With the rapid increase in accessing digital images, many watermarking techniques have been proposed for copyright protection and ownership authentication. The two most common approaches for image watermarking are the spatial and spectral domain techniques. In the spatial domain, the watermark is embedded in selected regions chosen based on the texture of the given image [1, 2]. In the spectral domain, the watermark is embedded in the transform domain using methods such as DCT and DWT, in the mid-frequency range to ensure transparency and robustness of the watermark simultaneously [3].

Despite the different approaches introduced for image watermarking, the DWT approach remains one of the most effective and easy to implement techniques for image watermarking [4]. The biggest issue in DWT-based image watermarking is how to choose the coefficients to be used for watermark embedding. The most common approaches include modifying the largest DWT coefficients in all decomposition levels or quantizing certain DWT coefficients in different levels and scales. Other approaches mark the host image by setting modulo 2 difference between the largest and the smallest coefficients according to the watermark bit value. Some DWT watermarking methods use a Human Visual System (HVS) model to determine the coefficients to be watermarked [5, 6, 7].

The effectiveness of DWT-based image denoising in separating the wavelet coefficients that represent the noise from the signal motivates us to apply it to watermarking. In this paper, a spatially adaptive wavelet thresholding method is used to determine the coefficients to be watermarked [8].

The paper is organized as follows: Section 2 gives a brief background on the use of DWT in watermarking and denoising. Sections 3 and 4 describe the embedding and extraction algorithms respectively. In Section 5, the performance of the algorithm under different attacks is illustrated. A summary of the paper and conclusions are given in Section 6.

2. BACKGROUND

The multiresolution wavelet transform of an image decomposes the image into bands of approximately equal bandwidth on a logarithmic scale. Similarly, the retina of the human eye splits the image into several components, each having a bandwidth of approximately one octave. Therefore, it is believed that the use of DWT for watermarking produces an imperceptible watermark [5, 10]. The use of the DWT domain for image watermarking and denoising has been studied in detail. The DWT splits the signal into high and low frequency parts. The high frequency part contains information about the edge components, while the low frequency part is split again into high and low frequency parts as shown in Fig 1. The high frequency components are usually used for watermarking since the human eye is less sensitive to changes in edges.

In watermarking, the main concern besides invisibility of the watermark is how to choose the coefficients to be watermarked such that they will survive the possible attacks that the transmitted image may undergo. For denoising, the major concern is to get rid of the coefficients that do not carry important information, e.g noise.

The proposed method uses the concept of wavelet thresholding with context modeling to determine which coefficients to embed the watermark into [8]. A spatially adaptive wavelet threshold is found for every wavelet coefficient and is compared with the coefficient itself. In image denoising, the coefficient that has a value less than the corresponding threshold is considered "insignificant" or noise, while in watermarking, the coefficient is considered "significant" or watermarkable if it has a value greater than the threshold. In this paper, we leave the "insignificant" coefficients unchanged and choose a subset of the "significant" coefficients for watermarking as discussed in the next section. The embedding algorithm ensures that no watermark is embedded into the largest coefficients. This approach ensures simultaneously that the "noise" or the "insignificant" coefficients and the large coefficients which correspond to the "visible" part of the image are not altered. This, in turn, ensures imperceptibility and robustness of the proposed watermarking algorithm.

3. WATERMARK EMBEDDING

In this paper, we assume that the original image I is of size $N \times N$ and the watermark, w, is a binary sequence of length R, which takes values from $\{1, -1\}$. We use the spatially
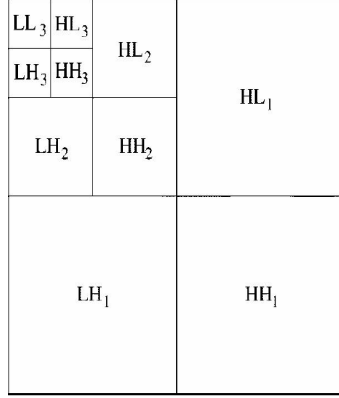
Figure 1: Three levels discrete wavelet decompositions.

adaptive wavelet thresholding method in [8] to find the coefficients to be watermarked. This method can be summarized as follows [1].

Each coefficient, $Y[i,j]$, in a particular subband with M^2 coefficients is treated as a random variable. The variance of this random variable is estimated by considering the neighborhood of $Y[i,j]$. The absolute value of the eight nearest neighbors of $Y[i,j]$ plus its parent coefficient are placed in a vector u_{ij} of size $p \times 1$. The context of the current pixel is calculated as the weighted average of the absolute value of its neighbors,

$$Z[i,j] = d^t u_{ij}, \qquad (1)$$

where d is the weight found by minimizing the mean square error,

$$d_{LS} = argmin \sum_{i,j} \left(|Y[i,j]| - d^t u_{ij} \right)^2 = \left(U^t U\right)^{-1} U^t |Y|, \qquad (2)$$

where U is a $M^2 \times p$ matrix and each row is u_{ij}^t, and Y is a vector of size $M^2 \times 1$ that contains all coefficients in the given subband.

The variance of $Y[i,j]$ is estimated by considering the other coefficients whose context is close in value to $Z[i,j]$. For a given coefficient $Y[i_0, j_0]$, a window is placed around $Z[i_0, j_0]$, and the points in $Y[i,j]$ whose context falls within this window are used to estimate the variance. The L points above $Z[i_o, j_o]$ and the L points below $Z[i_o, j_o]$ in value are chosen, which yields a total of $2L+1$ points. To ensure that enough points are used for estimating the variance, L is set to $L = \max(50, 0.02M^2)$. The variance is estimated as,

$$\hat{\sigma_X^2}[i_0, j_0] = \max\left(\frac{1}{2L+1} \sum_{[k,l] \in B_{i_o j_o}} Y[k,l]^2 - \sigma_n^2, 0 \right), \quad (3)$$

where $B_{i_o j_o}$ contains the points $\{Y[i,j]\}$ whose context falls in the window. The threshold for $Y[i_0, j_0]$ is estimated using,

$$T_B[i_0, j_0] = \frac{\sigma_n^2}{\hat{\sigma}_X[i_0, j_0]}, \qquad (4)$$

where σ_n^2 is estimated using the robust median estimator in the highest subband in the wavelet transform,

$$\hat{\sigma}_n = \frac{\text{Median}(|Y[i,j]|)}{0.6745}, \qquad (5)$$

[1]Readers are encouraged to refer to [8] for more details

where $Y[i,j] \in$ subband HH_1. The same procedure can be applied for every location $[i,j]$ to obtain a spatially adaptive threshold for every coefficient $Y[i,j]$. This algorithm makes the threshold values adaptive to the spatially changing statistics of the image, which in turn improves the wavelet thresholding performance because it incorporates additional local information about the image such as the identification of the smooth regions and the edges.

The watermark embedding algorithm can be summarized as follows:

1. Compute the Lth level DWT of the original image to obtain the detail (horizontal HL_l, vertical LH_l and diagonal HH_l) images at each level $l = 1, 2, ...L$, plus the approximation at the Lth level. In this paper, L is set to 5.
2. For every coefficient at each orientation in level l, $s_l \in \{HL_l, LH_l, HH_l\}$, find the spatially adaptive threshold, $T_B[i,j]$, for every coefficient $Y_{s_l}[i,j]$.
3. Find all coefficients that satisfy: $Y_{s_l}[i,j] > T_B[i,j]$.
4. Arrange the coefficients found in step 3 in descending order with respect to their absolute values.
5. Take the smallest R coefficients found in step 4 at each scale and modify them as follows:

$$Y_{s_l}(i,j) = w_r \cdot \alpha |Y_{s_l}(i,j)|, \\ r = 1, 2, ...R, \qquad (6)$$

   where α is a positive constant that controls the strength of the watermark embedding and w_r is the rth watermark bit.
6. Save the locations of the modified coefficients as a key K. The key has value one if the coefficient is modified and zero if not.
7. Compute the inverse DWT to obtain the watermarked image.

This method modifies a subset of the coefficients with values greater than $T_B[i,j]$ by scaling them with a constant α. Since the high coefficients that correspond to the "visible" part of the image are not altered, the imperceptibility of the watermark is ensured even if the watermark is added at all levels. Moreover, the watermark is not embedded into the "noise" part to ensure the robustness of the watermarking algorithm.

4. WATERMARK EXTRACTION/DETECTION

In this section, we provide a semi-blind algorithm for extracting the watermark using the key K. The extraction of the watermark from an attacked image, $\hat{I}$, can be summarized as follows:

1. Find the Lth level DWT of the received image $\hat{I}$.
2. Find the modified coefficients according to the given key K.
3. For each orientation s_l at every level l, estimate the value of the watermark bit according to the sign of the corresponding wavelet coefficient $\hat{Y}_{s_l}(i,j)$:

$$\begin{aligned} &\text{if} \quad \hat{Y}_{s_l}(i,j) > 0, \qquad \hat{w}_r = 1 \\ &\text{else} \quad \hat{w}_r = -1, \quad r = 1, 2, ...R. \end{aligned} \qquad (7)$$

4. Using the results from step (3), a majority vote decision rule is used to determine the value of the watermark bit embedded.

The following correlation based detector is applied to quantify the performance of the extraction algorithm,

$$\langle w(n), \hat{w}(n) \rangle \quad \overset{watermark}{\underset{no\ watermark}{\overset{>}{\underset{<}{}}}} \quad \eta. \tag{8}$$

The expected value of the maximum of this correlation is,

$$\eta_{max} = R. \tag{9}$$

Let,

$$z = \sum_n w(n)\hat{w}(n). \tag{10}$$

The threshold η is derived by applying the Neyman-Pearson criterion to this detection statistic, which corresponds to solving,

$$\overline{P_{FA}} = \int_{\eta}^{+\infty} f_z(z|H_0)\ dz, \tag{11}$$

where $f_z(z|H_0)$ is the pdf of z for which the watermark is not present. The mean and the variance of z, since $w(n)\hat{w}(n)$ takes only two values -1 and 1 with equal probability, are given as,

$$\mu_z = 0, \tag{12}$$

$$\sigma_z^2 = R. \tag{13}$$

By applying the central limit theorem [9], the pdf of z can be assumed to be a normal distribution as $R \rightarrow \infty$. Therefore, $\overline{P_{FA}}$ is given by,

$$\overline{P_{FA}} = Q\left(\frac{\eta}{\sqrt{R}}\right), \tag{14}$$

where $Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^{+\infty} \exp\left(-\left(\frac{t^2}{2}\right)\right) dt$. For a given $\overline{P_{FA}}$, the threshold, normalized by its maximum value, can be written as,

$$\eta = \frac{Q^{-1}(\overline{P_{FA}})}{\sqrt{R}}. \tag{15}$$

The result in (15) shows that the choice of η should be dependent on the length of the watermark R. In particular, for $R = 256$ and $\overline{P_{FA}} = 0.01$, $\eta = 0.144$.

5. RESULTS

The watermark embedding algorithm proposed in this paper was applied to the well-known Lena image of size 512×512. The watermark is a multi-bit sequence of $\{1, -1\}$ with length 256. The watermark is embedded into every resolution level using five level wavelet decomposition with haar filter and $\alpha = 6$ unless otherwise mentioned. The watermarked image is similar to the original one with no visible differences and PSNR=45.1dB. The algorithm was tested under different attacks. Table 1 shows the effect of the choice of α on the robustness of the proposed algorithm under additive white gaussian noise "AWGN", "JPEG" compression, median filtering "MF" and rotation. It is clear that increasing α improves the robustness of the algorithm. The watermark is detectable in all attacks as the correlation values in Table 1

are greater than the threshold $\eta = 0.144$. The constant α was chosen such that the imperceptibility of the watermark is maintained, so the choice of α is image-dependent. In case of the Lena image, it is found that $\alpha < 8$ produces an invisible watermark and high PSNR. The PSNR value ranges from 57.3dB for $\alpha = 1$ to 43.2dB for $\alpha = 7$.

Table 1: The correlation between the extracted and original watermarks under different types of attack with different α values.

α	1	4	7
AWGN (PSNR=45db)	0.77	1	1
AWGN (PSNR=40db)	0.65	0.97	0.99
AWGN (PSNR=20db)	0.13	0.88	0.94
JPEG (Q=70%)	0.36	0.79	0.81
JPEG (Q=80%)	0.54	0.89	0.89
JPEG (Q=90%)	0.66	0.96	0.97
JPEG (Q=100%)	1	1	1
MF (3×3)	0.66	0.91	0.92
MF (5×5)	0.33	0.59	0.68
MF (7×7)	0.25	0.44	0.46
Rotation ($1°$)	0.89	1	1
Rotation ($3°$)	0.88	0.98	0.98
Rotation ($5°$)	0.86	0.97	0.97
Rotation ($7°$)	0.80	0.97	0.97

We compared our algorithm ($\alpha = 7$) with another well-known DWT algorithm [10]. Authors in [10] use five level wavelet decomposition is used for embedding the watermark. They transform the binary watermark into a real-numbered image and embed it in one of the bands in the highest level. They save the filters bank for watermark detection. The same watermark was embedded in both methods. The two methods were tested under attacks, AWGN, JPEG compression, rotation and median filtering. Table 2 shows the correlation between the extracted watermark and the original one under different attacks for both methods. The variance of the noise in AWGN case is 100, the size of filter is 3×3, the rotation degree is $3°$ and the quality factor in the JPEG case is 80. Both methods perform equally well under AWGN noise, while the proposed method performs better under median filtering, rotation and JPEG compression. The proposed method always chooses coefficients greater than T_B and embeds the watermark in a multiplicative way at all levels, which makes it more robust against attacks.

6. CONCLUSIONS

In this paper, we have presented a simple robust watermarking algorithm based on DWT. The algorithm uses the idea of spatially adaptive wavelet thresholding for image de-noising to choose the coefficients to be watermarked. The performance of the proposed semi- blind detection algorithm is quantified by deriving the optimum threshold for a given false alarm rate. The proposed algorithm is shown to be transparent and highly robust under attacks. The algorithm performs better than a well-known DWT based method.

REFERENCES

Table 2: A comparison between the proposed methods and method in [10] under different attacks.

Image	Attack	Proposed	Method in [10]
Lena	AWGN	0.96	0.97
Lena	MF	0.92	0.39
Lena	Rotation	0.97	0.75
Lena	JPEG	0.90	0.51
Peppers	AWGN	0.98	0.98
Peppers	MF	0.90	0.42
Peppers	Rotation	0.96	0.77
Peppers	JPEG	0.88	0.47
Barbara	AWGN	0.98	0.98
Barbara	MF	0.89	0.52
Barbara	Rotation	0.96	0.74
Barbara	JPEG	0.91	0.53

[1] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Processing*, vol. 66, pp. 385–403, 1998.

[2] D. P. Mukherjee, S. Maitra, and S. T. Acton, "Spatial domain digital watermarking of multimedia objects for buyer authentication," *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 1–15, Feb. 2004.

[3] J. R. Hernández, M. Amado, and F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Transaction on Image Processing*, vol. 9, pp. 55–68, Jan. 2000.

[4] P. Meerwald and A. Uhl, "Survey of wavelet-domain watermarking algorithms," in *Proc. SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents III*, Jan. 2001, vol. 4314.

[5] D. Kundur and D. Hatzinakos, "Toward robust logo watermarking using multiresolution image fusion principles," *IEEE Transactions on Multimedia*, vol. 6, pp. 185–198, 2004.

[6] A. A. Reddy and B. N. Chatterji, "A new wavelet based logo-watermarking scheme," *Pattern Recognition Letters.*, vol. 26, no. 7, pp. 1019–1027, May 2005.

[7] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking," *IEEE Transactions on Image Processing*, vol. 10, pp. 783–791, May 2001.

[8] G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modelling for image denoising," *IEEE Transactions on Image Processing*, vol. 9, pp. 1522–1531, Sept. 2000.

[9] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1991.

[10] Y. Wang, J. Doherty, and R. Van Dyck, "A Wavelet-Based Watermarking Algorithm for Ownership Verification of Digital Images," *IEEE Transactions on Image Processing*, vol. 11, pp. 77–88, Feb. 2002.