中正嶺學報 第三十二卷 第二期 民國 93.5.
JOURNAL OF C.C.I.T., VOL.32, NO.2, MAY, 2004

Robust Watermarking Scheme for Digital Images Using Significant Coefficients and the De-correlating Principle

Der-Chyuan Lou* and Chia-Hung Sung**

**Department of Electrical Engineering, Chung Cheng Institute of Technology, National Defense University Tahsi, Taoyuan 33509, Taiwan, E-mail: dclou@ccit.edu.tw*
***Graduate School of National Defense Science, Chung Cheng Institute of Technology, National Defense University*

ABSTRACT

A robust watermarking scheme is proposed to protect the copyright of digital images. The imperceptibility and robustness of a digital watermark are crucial. However, these requirements are frequently conflicting. The proposed scheme lessens this conflict by employing the spread-spectrum technique and the de-correlating process, which introduces randomness and significance into transformed coefficients, and provides a complementary embedding to yield better robustness. Besides, the proposed scheme is a blind method, meaning that the watermark can be detected from the watermarked image without the original image. From simulation experiments, results indicate that our scheme is remarkably effective in resisting various attacks.

Keywords: blind watermarking, robustness, imperceptibility, copyrights protection, de-correlating principle, spread-spectrum communication

使用重要係數與除相關原則之強韌型數位影像浮水印技術

婁德權*　宋嘉宏**

*****國防大學中正理工學院電機工程學系
******國防大學中正理工學院國防科學研究所

摘　　要

本文提出一個強韌型浮水印技術來保護數位影像之智慧財產權。對浮水印技術而言，浮水印的透明度及強韌性是很重要的需求，但是這些需求常常是衝突的。本文所提出的方法是要藉著應用展頻通訊技術及在除相關的過程中導入混亂及重要性到轉換領域的係數上來降低這樣的衝突。同時本文也提出一個互補式的嵌入策略來將浮水印嵌入轉換領域的係數以得到較佳的強韌性。此外，本文所提出的浮水印技術是一個盲目的方法，意味著浮水印的偵測是不需要使用原始的影像。由模擬實驗的結果顯示我們所提出的方法在抵抗各種攻擊上有很好的效果。

關鍵詞：盲目浮水印，強韌性，透通性，智慧財產權保護，除相關原則，展頻通訊

文稿收件日期 92.6.15; 文稿修正後接受日期 92.11.19.
Manuscript received June 15, 2003; revised Nov. 19, 2003.

I. INTRODUCTION

The rapid growth of the Internet, broadband networks, large archives, and the development of efficient compression algorithms have provided producers of multimedia a range of methods for distributing their products. These advances have also facilitated the creation and distribution of unauthorized copies of multimedia data by copyright violators, depriving the producers of their financial, legal, and intellectual rights. A digital watermark is an imperceptible signal embedded in the media content. The embedding and detecting of digital watermarks can help identify the source or ownership of multimedia data and prevent multimedia data from being illegally copied.

To be effective, a watermarking method for digital images must meet some basic requirements. It should be statistically and perceptually undetectable (imperceptibility). It must tolerate signal manipulations and be difficult for an attacker to destroy the watermark (robustness). It must have the ability to resist intentional tampering (security). However, imperceptibility and robustness of a digital watermark are conflicting. For example, a robust watermark not only survives signal processing but also produces distortions, which may be visible in a few distinct edges of images.

From the viewpoint of embedding a watermark, most schemes embed a watermark to images in the spatial domain [1,2] or in the transform domain [3-14]. The former schemes embed the watermark by directly adjusting the magnitude of the pixels' values, and the latter schemes modulate the transformed coefficients of the original image. Generally, spatial domain methods yield good perceptual results but they suffer from poor security and weak robustness. Although transform domain methods require more operations, they are typically much more robust than the spatial domain schemes.

From the viewpoint of detecting a watermark, watermarking methods can be classified into two classes according to whether the detection of the watermark depends on the original image. For example, a "cover escrow" method uses the original image to detect a watermark [4,9], and a blind or oblivious method does not require the original image [1,15]. The former is very robust to various image processing operations but it has the shortcoming that it cannot be used in some applications, such as the automatic search engines that search for watermarks automatically. The latter includes some limitations on the insertion of the watermark, because the watermark must be detectable without the original image. Such a scheme is practical in a network environment because it avoids the transmission of a duplicate copy when bandwidth is limited. Furthermore, sending the original image is insecure because an attacker could intercept it.

Transform-domain watermarking schemes are further divided by embedding methods using the discrete cosine transform (DCT) [3-9] and the discrete wavelet transform (DWT) [10-14]. In 1996, Swanson *et al.* [3] presented a watermarking method that applies a human perceptual model in the DCT domain. In 1997, Cox *et al.* [4] proposed a method for adding a watermark to the DCT coefficients of images; the method was robust and exclusively used the DC

coefficient. In 1998, Barni *et al*. [5] proposed a blind watermarking scheme using many insignificant DCT coefficients to trade off between robustness and transparency. This method requires a substantial number of computations to compute the global DCT, and is unsuitable to a large image.

In 1998, Kutter [6] developed an adaptive scheme based on the luminance sensitivity function of the human visual system (HVS) [16]. Unfortunately, the masking function based on the estimate of the image luminance for watermark embedding is inefficient to wavelet compression or denoising attacks. In 1998, Delaigle *et al.* [7] presented a perceptual modulation function to overcome the problem of the visibility of a watermark around edges. In 1999, Voloshynovskiy *et al*. [8] presented an effective stochastic model for content-adaptive watermarking. With knowledge of stochastic models of the watermark and the host image, the problem of watermark detection can be formulated using a posteriori probability. In 1999, Hsu and Wu [9] presented a transform domain technique that uses DCT transformation and involves fixed blocks. The method embeds data by modifying middle-band coefficients according to a residual mark to reverse the polarity.

In 1997, Kundur *et al*. [10] decomposed a binary image through DWT and repeatedly added it to the subbands of the DWT decomposition. Before it is added, the watermark is scaled by a salience factor, computed block by block, and related to the sensitivity of the local image noise. In 1998, Inoue *et al*. [11] took inspiration from the wavelet-based compression algorithms to modify significant DWT coefficients to carry the watermark. Some side information is required to recover the watermark. In 1998, Xia *et al*. [12] added a pseudo-random sequence to the largest coefficients in the detail bands. Perceptual considerations are addressed by setting the amount of modification proportional to the strength of each coefficient. The watermark is detected by comparison with the original un-watermarked image. This scheme fails under the JPEG attack, since JPEG quantization sets most coefficients of higher frequency components in each block to zero, so the watermark embedded in these coefficients is lost. In 1999, Wang and Wederhold [13] proposed the hiding of a watermark in less significant coefficients of DWT subband. The coefficients to be modified are selected according to the results of a perceptual analysis step, performed on each block, by estimating the local variance. In 2000, Lumini and Maio [14] presented to embed a watermark into the low-resolution part of the image in the DWT domain.

In this paper, we propose a robust watermarking scheme to embed two pseudo-random watermarks into significant coefficients of an image to protect copyright using the spread-spectrum technique and the de-correlating principle. The de-correlating process introduces randomness into the un-protected image to increase in number of the significant coefficients and to enhance the orthogonality between the image and the watermark. Two watermarks play complementary roles to yield better robustness of watermark detection. The simulation result shows that the

proposed method can withstand more attacks than other methods can in terms of the same quality.

The rest of this paper is organized as follows. Section II reviews the spread-spectrum technique. Section III describes the proposed watermarking scheme, including the de-correlating principle, selecting significant coefficients, embedding and detection processes. Section IV presents the experimental results that show the effectiveness of the proposed scheme over that of other schemes. Finally, Section V draws conclusions.

II. THE SPREAD SPECTRUM TECHNIQUE

Spread spectrum (SS) technologies [17] were first used to secure military communication, because of their robustness against jamming, interference and multi-path distortion in fading channels. In spread spectrum communication, a narrow-band signal is modulated into a wide-band signal with a noise-like and pseudo-random fashion. The power of the signal to be transmitted can be large; the signal-to-noise ratio in every frequency band is small. The embedded signal is cryptologically secure and hard to detect. Even if parts of the signal are removed from some frequency bands, sufficient information should remain in other bands to allow the signal to be recovered. The strength of the SS technique is that the signal detection is self-synchronized. The embedded signal can be detected without knowledge of the original cover signal.

Let x be the original signal used to hide a signal to be transmitted. The stego-signal y is given by $y = x + w$, where w is the random sequence of elements w_i with two equiprobable values ; that is, $w_i \in \{-\Delta, +\Delta\}$ generated independently with respect to x. The magnitude Δ is set based on the sensitivity of HVS to amplitude changes. A detector detects the embedded signal by correlation tests C between the stego-signal y and the signal w.

$$\begin{aligned} C &= y \cdot w = (x + w) \cdot w \\ &= x \cdot w + w \cdot w = x \cdot w + N\Delta^2, \end{aligned} \quad (1)$$

where N is the cardinality of involved signals. Since the original signal x can be modeled as a Gaussian random sequence, $x = N(\mu_x, \sigma_x), \sigma_x >> \Delta$, the normalized value of the correlation test can be represented as:

$$Q = \frac{C}{N\Delta^2} = \rho + \frac{1}{\Delta} N(0, \sigma_x / \sqrt{N}), \quad (2)$$

where $\rho = 1$ if the signal w is present and $\rho = 0$ otherwise. The optimal detection rule is to declare the presence of signal w if $Q > T_C$. The choice of the threshold T_C controls the trade-off between false alarm and detection probabilities. According to the central limit theorem, the probability of $Q > T_C$ is equal to,

$$\lim_{N \to \infty} \Pr[Q > T_C] = \frac{1}{2} \mathbf{erfc}\left(\frac{T_C \sqrt{N}}{\sigma_x \sqrt{2}} \right). \quad (3)$$

III. ROBUST WATERMARKING SCHEME

In this section, we propose a complementary watermarking scheme in order to obtain higher detecting responses based on the de-correlating principle. The scheme embeds two watermarks in an image and makes them hard to remove simultaneously. Then, at least one watermark survives various attacks. Other parts of the proposed scheme, including selecting significant coefficients, embedding watermarks and detecting watermarks, are also presented in details.

A. De-correlating Principle

This subsection describes the reasons for applying the de-correlating principle to an un-protected image. From Eq. (3), the size of the watermark must be large to obtain a sufficiently small error probability. However, a long watermark increases the complexity and delay of detection. To improve the problem, the watermark is reduced to save the computational load. However, a short watermark may decline orthogonality between the watermark and the original image. The declined orthogonality produces a bias interfering with the correct correlation obviously to raise error probability. If the bias suppresses the embedded watermark, correct detection is impossible. To cope with this case, the watermark is increased both in the length and strength, and the de-correlating principle is used to enlarge space and strength of the watermark.

The de-correlating process introduces randomness to provide more significant coefficients. Most image-watermarking techniques use DCT transform to compact the energy of an image into relatively few low-frequency coefficients, which may provide insufficient space to embed a watermark. Introducing some significant coefficients can increase the space available [11]. The de-correlating process transforms the un-protected image into meaningless data, and pixels' amplitudes are modeled as un-correlated and identically distributed random variables. The spectrum of the un-correlated variables contains significant middle-frequency and high-frequency coefficients, and the energy is spread uniformly. From the perspective of communication theory [18], uniformly spreading spectrum energy notably increases the bandwidth for data transmission. The de-correlating process spreads energy to introduce more significant coefficients and space for embedding a long and strong watermark, and that enhances orthogonality between the watermark and the cover image.

The de-correlating process utilizes four functions: an initial permutation (**IP**); a transposition function titled $\mathbf{F_K}$, which involves both permutation and substitution operations and depends on a key input; a permutation function titled **SW** that simply switches the two halves of the location data, and a permutation function that is the inverse of the initial permutation ($\mathbf{IP^{-1}}$). The use of multiple stages of permutation and substitution results in a confusion operation, which increases the difficulty of steganalysis. We can concisely express the de-correlating operator $\mathbf{D_K}$ as a composition of functions shown as Eq. (4):

$$\mathbf{D_K} = \mathbf{IP^{-1}} \cdot \mathbf{F_{K_2}} \cdot \mathbf{SW} \cdot \mathbf{F_{K_1}} \cdot \mathbf{IP}, \qquad (4)$$

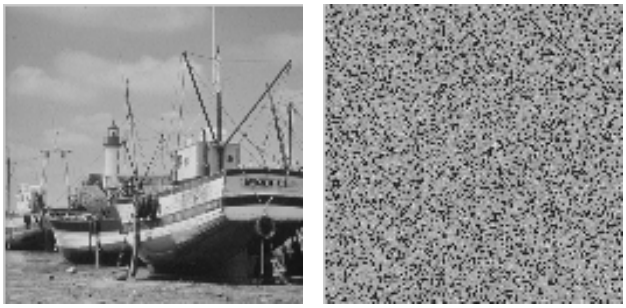
where $\mathbf{K_1}$ is the output of the shift operation performed on the input key and $\mathbf{K_2}$ is the output of the shift operation base on $\mathbf{K_1}$.

The output of this de-correlating operation $\mathbf{D_K}$ can be written as Eq.(5).

$$\begin{aligned}\mathbf{I}^*(n_1,n_2) &= \mathbf{D_K}(\mathbf{I}(n_1,n_2)) \\ &= \left(\mathbf{IP}^{-1}\left(\mathbf{F_{K_2}}\left(\mathbf{SW}\left(\mathbf{F_{K_1}}\left(\mathbf{IP}(\mathbf{I}(n_1,n_2))\right)\right)\right)\right)\right),\end{aligned} \qquad (5)$$

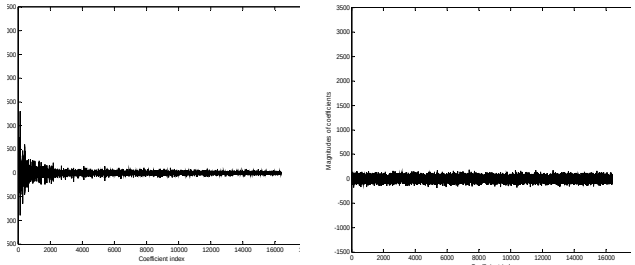
where $\mathbf{I}(n_1,n_2)$ is two-dimensional signals and $\mathbf{I}^*(n_1,n_2)$ is the results of the de-correlating process. For example, Fig. 1(a) displays the cover image "Boat", with a size of 128×128. Fig. 1(b) shows the de-correlated image. Fig. 2(a) shows the spectral distribution of DCT coefficients

before the image is scrambled. Fig. 2(b) shows the spectral distribution of DCT coefficients after the image is scrambled. Notably, the magnitudes of most coefficients after scrambling become large to allow a robust watermark to be embedded, because a de-correlating operation distributes an image uniformly over the entire spectrum frequencies.



(a) (b)

Fig. 1. (a) An example image "Boat," (b) the de-correlated result of Boat image.



(a) (b)

Fig. 2. (a) The spectral distribution of DCT coefficients before Boat image was scrambling, (b) the spectral distribution of DCT coefficients after Boat image was scrambling.

Based on such characteristics, extra space and strength from the coefficients that are significant after the image is de-correlated can be used to embed a robust watermark. A watermark detector can produce a high correlation response when the watermark is enhanced in strength and length.

B. Selecting Significant Coefficients

The proposed watermarking scheme involves wavelet transform, which is a hierarchical system and described as follows. An image is first decomposed into four parts； that is, LL1, HL1, LH1 and HH1 subbands by subsampling horizontal and vertical channels using subband filters. The subbands HL1, LH1 and HH1 are specified by the finest scaled wavelet coefficients. The subband LL1 is further decomposed and subsampled to find the next coarser scaled wavelet coefficients. This process is repeated several times, which is determined by the application. Fig. 3 shows an image decomposed into ten subbands by three levels decomposition. Each level includes various subbands, including the low-low, low-high, high-low and high-high frequency subbands. Next, selecting significant coefficients using the hierarchical structure is described.

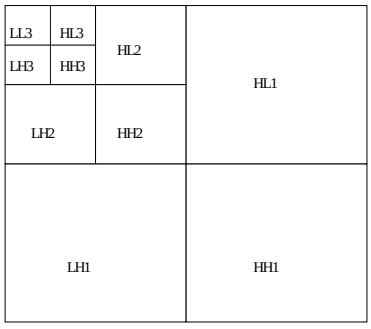


Fig. 3. Hierarchical representation of three-level wavelet decomposition structure.

The significance and the number of coefficients are important factors to compromise between robustness and transparency of a watermark. Signal processing and compression attacks generally do not greatly change significant coefficients, which have a large magnitude and

make the embedded watermark robust against attacks. If the coefficients do change substantially, the reconstructed image will perceivably differ from the original image and the value of the protection of the intellectual property that corresponds to such a severely degraded image becomes low. Besides, the watermark cannot be long because modifying many significant coefficients will severely degrade the visual quality of the image. Usually, a coarse subband corresponds to fewer significant coefficients than a fine one in the hierarchical structure of DWT. Embedding a watermark into a coarse subband only increases the strength of the watermark but decreases the orthogonality due to few coefficients. On the contrary, embedding the watermark into a finest subband has the disadvantages of high computational complexity and vulnerability to attacks. A reasonable tradeoff is achieved by embedding the watermark into middle components of an image.

Suppose the middle-frequency coefficients of each subband are assumed to follow the Gaussian distribution after de-correlation. For a Gaussian distribution with zero mean and variance σ^2, the simplified rate-distortion model is as follows.

$$D = \begin{cases} T^2/12, & \sigma^2 > T^2/12, \\ \sigma^2, & \sigma^2 \le T^2/12. \end{cases} \qquad (6)$$

As specified in Eq. (6), *D* is proportional to the square of the threshold T if the corresponding variance is large. A large *D* implies that this subband has more energy and is more significant than other subbands. Thus, significant subbands can be searched based on a comparison between a threshold and its variance. According to Eq. (3), the detection probability by correlation must be maintained by a square number of coefficients in accordance with the variance. Thus, define a significant subband, in which the wavelet coefficients on a coarse subband satisfy $\lambda = (\sigma_s / N_s^2) > T$ and $N_s > N_\delta^{1/2}$, where N_δ is a low-bound number of coefficients involved in correlation computation. A significant subband with few coefficients may be selected when N_δ is small. Embedding a short watermark into fewer coefficients may result in a worse correlation even the coefficients are significant. In the proposed scheme, the targets for embedding a watermark are middle components, which are located in the same level as the significant subband. Those middle components include low-high and high-low subbands, which have the same number and similar significance as the coarse subband. Modifying them to embed a watermark may not only degrade the image less than modifying the coarse subband but also own the robustness attribute in resisting attacks. Besides, T is the threshold given by $T = C_{\max,s} \times \delta$, where $C_{\max,s}$ represents the maximum absolute value of coefficients in coarse subband *S*, and δ is the threshold ratio that increases the flexibility of the proposed watermarking scheme.

C. Embedding the Watermark

The proposed embedding process includes four parts: wavelet decomposition, coefficients de-correlation, DCT transformation, and watermark casting. Fig. 4 shows a block diagram of embedding watermarks. The image *X* is first

decomposed into hierarchical subbands, and the significant subband is selected as described in Section III-B. Next, a complementary embedding strategy is proposed.

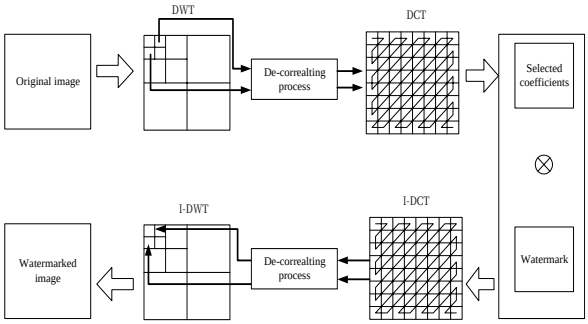


Fig. 4. The block diagram of the proposed watermark embedding process.

The proposed scheme embeds two watermarks, which play complementary roles in resisting various kinds of attacks. The values of the two watermarks are drawn from the same watermark sequence. The difference is that they are embedded in two different subbands. If one of them is destroyed by attacks, the other has the chance to maintain the detecting response to identify the ownership except that the watermarked image is degraded severely. Another reason for casting watermarks to two subbands is to resist a removal attack. Consider the following cases of malicious manipulation. First, if someone tries to remove the watermark, the constraint of perceptual distortion can banish his attempt. Complete removal of two embedded watermarks by an attacker will degrade the image markedly, since the watermarks are embedded into two subbands, which include many significant coefficients. Second, suppose someone wants to maliciously manipulate the watermarked coefficients such as by replacing them with other un-marked coefficients. He needs the original image to obtain the un-marked coefficients. Otherwise, the marked coefficients would be replaced with uncertain ones, and the modification would be obvious because of the localization characteristics of DWT. The proposed scheme prevents from such an attack by using a unique image that is destroyed after embedding the watermark. An attacker will be unable to intercept the original image to replace the watermarked coefficients, since the proposed method is a blind approach.

The image to be watermarked is first decomposed through DWT decomposition. Wavelet transformation provides both frequency and spatial localization of the energy in an image. Coefficients with local information in the subband are unsuitable to be target coefficients, because embedding a watermark directly into such coefficients raises suspicion. To solve this case, we use the de-correlating process introducing randomness and significance into coefficients, which are dispersed over the entire sub-band spatially. The watermark is embedded not by changing individual coefficients but by changing the statistical characteristics of a certain local area of a subband, to minimize the perceptual degradation and guarantee the robustness of the proposed scheme. A watermark embedded in these spread coefficients can easily survive typical image processing and such spreading also avoids revealing the embedded watermark.

DCT has good frequency resolution to provide a good shelter for the watermark signal from human perception. Suppose subbands LH_n and HL_n be selected for embedding the two watermarks. The choice of subbands is described

in Section III-B. The DCT is applied to the two subbands and the resultant coefficients are reordered into two sets in a zig-zag order, $C_{LH_n} = \{c_{i,LH_n}\}$ and $C_{HL_n} = \{c_{i,HL_n}\}$, where $i = 1 \cdots m$, and m is the number of coefficients in the subbands. Consequently, the SS technique is used to cast the watermark into selected coefficients, since a broadband watermark cannot be easily removed. The most important idea is casting the watermark in selected coefficients to generate protected coefficients by weighting absolute values of these coefficients. The watermark is localized in DCT coefficients, where the embedded watermark is less visible than elsewhere. Two watermarks with the same mean-zero, unit-variance and size m pseudo-random sequence, $W = \{w_i\}$, are embedded into these two sets of selected coefficients to yield two new sets of coefficients, $C'_{LH_n} = \{c'_{i,LH_n}\}$ and $C'_{HL_n} = \{c'_{i,HL_n}\}$, where $i = 1 \cdots m$. The watermark is cast into selected coefficients according to the following way,

$$c'_i = c_i + \alpha \log_r |c_i| w_i , \qquad (7)$$

where $i = 1 \cdots m$ and r is the radix of the logarithmic function.

The weighting powers of the watermark are determined by the weighting factor α and the introduced distortions, which are the absolute values of coefficients. A large α can increase robustness against signal processing and generate a distortion, which may be visible in images with a few distinct edges. To compromise this case, we use a logarithmic function to scale down the absolute values of the coefficients to increase α values inversely with the same weighting powers. The increased α enhances the robustness and the visual quality is not degraded by the introduced distortions obviously. Therefore, the loss of fidelity due to casting a watermark can be set a bound, and the proposed embedding method handless the tradeoff between visual distortion and robustness.

The watermarked coefficients are then inserted into their original positions in a zig-zag order, and the inverse DCT is performed to yield new subbands LH'_n and HL'_n. Then, the de-correlating process and the inverse DWT are used to obtain the watermarked image. The distortions due to watermark embedding are spread over the entire image. The difference between X and X' is inappreciable.

D. Detecting the Watermark

In the literature, a number of authors [4, 9, 19] have proposed extracting a watermark resorting to the original image under strong attacks. The above-mentioned methods used a similarity measure between the embedded watermark and the actual watermark for watermark detection. Fig. 5 shows a block diagram of proposed watermark detecting method, which makes watermark detection possible without the original image.

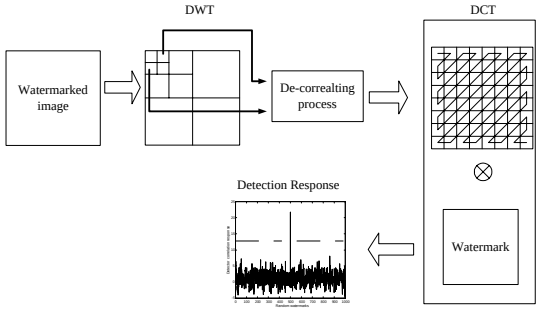


Fig. 5. The block diagram of the proposed watermark detecting process.

First, the wavelet transformation

decomposes a possibly watermarked image X^*. Subbands LH_n^* and HL_n^* are selected for watermark detection as described in Section III-B. Consequently, the de-correlating process disperses the selected subbands into meaningless data and then the DCT is applied to these data. Calculating the correlation between the watermark and these two sets of marked coefficients, $C^*_{LH_n} = \{c^*_{i,LH_n}\}$ and $C^*_{HL_n} = \{c^*_{i,HL_n}\}$, is for watermark detection, respectively. The correlation Z is defined as follows.

$$\begin{aligned} Z &= \max(Z_{LH_n}, Z_{HL_n}) \\ &= \max\left(\frac{1}{m}\sum_{i=1}^{m} C^*_{i,LH_n}\cdot w_i,\ \ \frac{1}{m}\sum_{i=1}^{m} C^*_{i,HL_n}\cdot w_i\right) \end{aligned} \tag{8}$$

Both c_i and w_i are zero-mean, mutually independent, and equally distributed random variables. The expected mean of Z is,

$$\mu_z = \begin{cases} \alpha\mu_{\log_r|x|}, & \textbf{if } w_e = w_t, \\ 0, & \textbf{if } w_e \neq w_t, \end{cases} \tag{9}$$

where w_e is the embedded watermark and w_t is the test watermark. From Eq. (8), the correct watermark should yield a higher detecting response than other watermarks. Recall that two watermarks are embedded into an image to yield a better correlation. The maximum value between these two correlation results concludes the final response of detecting watermark according to Eq. (7). Even if one of these two watermarks is destroyed, the other still survives attacks except that the image is degraded severely.

The value of the correlation is compared with a threshold to determine the presence of the watermark. The detector must consider the probability of error to determine the threshold. The probability of error P_E in detecting the watermark can be expressed as,

$$P_E = p(0|1)p(0) + p(1|0)p(0), \tag{10}$$

where $p(0|1)$ is the probability of missing the presence of the watermark and $p(1|0)$ is the probability of revealing the presence of the watermark when no watermark is present. $p(0)$ and $p(1)$ are priori probabilities of un-watermarking the image and watermarking the image, respectively. A threshold must be chosen to minimize the error probability P_E. According to statistical analysis, the threshold should be set to the middle value between zero and μ_Z. However, an intentional or unintentional attack can corrupt an image, and then this setting is no longer valid since the marked coefficients may have been changed by the attacks. The corrupted coefficients decline toward zero and their variances increase. Accordingly, the threshold TH should be set closer to zero corresponding to possible attacks, as follows.

$$TH = \frac{\alpha}{3m}\sum_{i=1}^{m}\log_r|c_i|. \tag{11}$$

IV. EXPERIMENTAL RESULTS AND PERFORMANCE EVALUATION

The following experiments used the image labeled “Lena” size 512×512, as shown in Fig. 6(a). The low-bound length of watermark was experimentally set to 2000. The threshold was set to 0.1. The λ_{LL1}, λ_{LL2} and λ_{LL3} were 0.0015, 0.0131, and 0.1088, respectively. The LL3 over the threshold was selected as the significant subband, and the middle components including LH3 and HL3 were selected to embed the watermarks with zero mean, unity variance and a

length of 4096. The peak signal-noise ratio (PSNR) of the watermarked image was 42.36 dB. Fig. 6(b) shows the watermarked image retaining a reasonable visual quality. The PSNR value of the watermarked image without applying the de-correlating process was 40.38 dB. The difference in image quality is around 2 dB. The PSNR is defined as,

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}, \qquad (12)$$

where *MSE* is the mean square-error between a watermarked image and its original image.

(a)

(b)

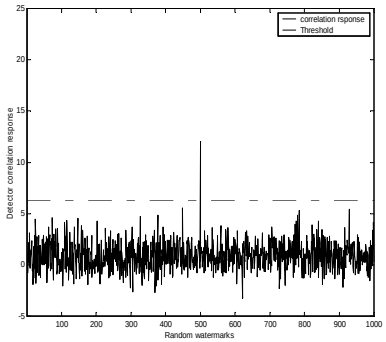
Fig. 6. (a) The original image, (b) the watermarked image with PSNR=42.36 dB.

Several image manipulations, including lossy JPEG compression, SPIHT compression, additive noise, median filter, and resizing, were performed on the watermarked image to evaluate the robustness of the proposed scheme. The correlation tests between the inserted watermark and 1000 randomly generated watermarks are measured for the response of detecting the watermark.

A. JPEG Attack

JPEG is a compression standard for still images. It was applied to the watermarked image to simulate an attack. Fig. 7(a) shows the image corrupted by JPEG compression with a quality factor (QF) of 10% and PSNR of 31.01 dB. The compression ratio is 40. Fig. 7(b) shows that the response of detecting the 500-th watermark exceeds the threshold to visually identify the ownership rights. This result implies that the proposed scheme can withstand a JPEG attack with a low quality factor.

(a)



(b)

Fig. 7. (a) The compressed image with a quality factor of 10% and a PSNR of 31.01 dB, (b) the detected response of the watermark.

B. SPIHT Attack

Set partitioning in hierarchical trees (SPIHT) compression [20] was used to elucidate the effectiveness of the proposed scheme. The SPIHT attack results in the quality of the watermarked image with a PSNR of 34.30 dB when the compression ratio is 32. Table 1 lists the response values of detecting watermark at different compression ratio (CR). The response of the embedded watermark can be easily identified. The fact that these detected responses exceed the thresholds implies that the proposed scheme can survive SPIHT attack.

Table 1. The correlation values using the proposed method under SPIHT attack.

Lena (512×512)					
CR	4:1	8:1	16:1	32:1	64:1
Bits per pixel	2	1	0.5	0.25	0.125
PSNR (dB)	40.9138	39.2588	37.0574	34.2969	31.2947
Threshold	6.2914	6.3111	6.3076	6.2644	6.0982
Responses	17.4325	16.0252	11.9835	8.4011	5.5822

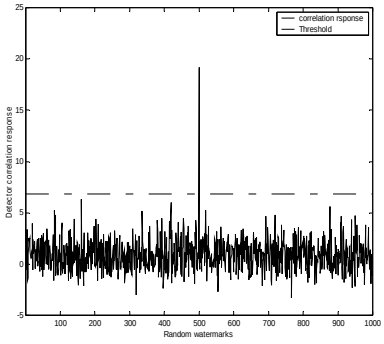
C. Additive Noise Attack

Next, adding Gaussian noise corrupted the watermarked image. Fig. 8(a) shows the noisy image obtained by adding noise with zero mean and a variance of 400 with a PSNR of 22.13 dB. The image is degraded so obviously that it is unacceptable in practical applications. Fig. 8(b) shows the response of detecting the watermark in the noisy image. One peak clearly shows that the 500-th watermark is the correct watermark, meaning that the proposed scheme is remarkably robust to noise attack.

D. Media Filtering Attack

The watermarked image was filtered through a media filter. Fig. 9(a) shows the resultant image by median filtering with a window size of 5×5 and a PSNR of 32.58 dB. The response of detecting the watermark exceeds the threshold, as shown in Fig. 9(b). The test shows that the proposed scheme survives the filter attack. The window size 5×5 of filtering does not influence the frequencies that carry most of the watermark information, because we randomize the transformed coefficients before embedding the watermark.

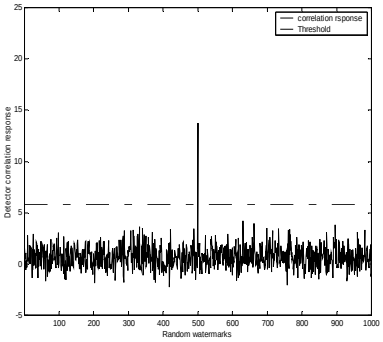
(a)



(b)

Fig. 8. (a) The noisy image with additive noise variance 400 and a PSNR of 22.13 dB, (b) the detected response of the watermark.

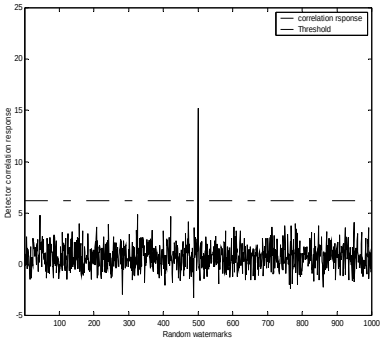
(a)



(b)

Fig. 9. (a) The median filter with a window size 5×5 and a PSNR of 32.58 dB, (b) the detected response of the watermark.

(a)



(b)

Fig. 10. (a) The corrupted image by resizing attack with a PSNR of 25.79 dB, (b) the detected response of the watermark.

E. Resizing Attack

Resizing occurs frequently in the editing of digital images. The watermarking technique must be robust against the resizing attack. A watermarked image is shrunk to a quarter of its original size and then rescaled to its original dimensions. The image quality is degraded obviously with a PSNR of 25.79 dB, shown in Fig. 10(a). However, the watermark is still clearly detectable in the corrupted version of the watermarked image, as shown in Fig. 10(b).

F. Comparisons with Previous Methods

Barni *et al*. [5] method is a blind approach, in which embedding watermark is weighted with the absolute values of coefficients. Fig. 11 shows responses of detecting the embedded watermark between the proposed method and the method of Barni *et al*. [5]. Four images “Pepper”, “Airplane”, “Lena” and “Boat” are used to examine the relationship between detection responses and compression QF. These QFs range from 50% to 5%. From Fig. 11, the detection responses of the proposed scheme appear to be obviously better than those of the method of Barni *et al*. [5]. In the propose scheme, the Boat image has better

detection response than other images. This image has more randomness patches than other images, and that enhances the robustness of the watermark after JPEG attack.

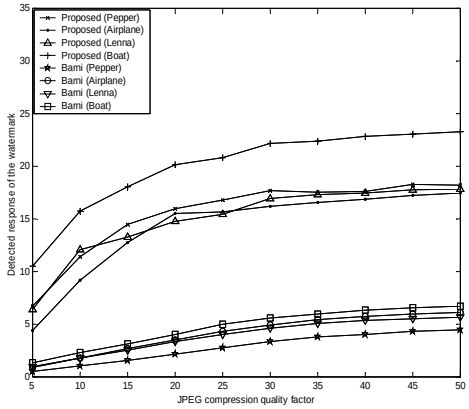


Fig. 11. The correlation responses of detecting the watermark compared the proposed method with Barni *et al*'s method [5] under various quality factors of JPEG lossy compression.

Table 2 shows the relationship between the weighting factor of the watermark and the degradation of the image based on PSNR. In this table, the proposed scheme always has better quality than other methods [5,12] under fixed weighting factor of the watermark. According to the results from a diagonal axis, the proposed scheme has the higher weighting factor than other methods for a given quality. Recall that the proposed scheme reduces the introduced distortions to coefficients and increases the watermark strength inversely. According to Eq. (9), the increasing strength can improve the response of detecting the watermark under an expected correlation value.

Table 2. The PSNR comparisons of proposed method, the method in reference 5, and the method in reference 12.

Lena (512×512)			
Weighting factor	5	0.4	0.06
Proposed method	42.3584	46.2348	62.7129
Barni *et al.* method [5]	18.1531	39.2588	56.5695
Xia *et al.* method [12]	2.3562	24.2944	40.7726

Table 3 summarizes the comparisons among the proposed scheme, the method of Barni *et al*. [5] and the method of Xia *et al*. [12] under various image manipulations. The results reveal that high responses of detecting the watermark using the proposed scheme are due to the strong watermark. It is shown that the proposed scheme has more chances than other methods to survive these attacks.

Table 3. The correlation comparisons of proposed method, the method in reference 5, and the method in reference 12 against various attacks.

Attacks	Proposed method	The method in [5]	The method in [12]
Weighting factor	5	0.4	0.06
JPEG (QF=5%)	6.4445	0.9390	0.0369
Scaling (1/4)	15.1939	2.1078	0.0006
Gaussian noise	19.1972	6.2949	0.7110
Median filtering	13.7661	1.6899	0.0332
Blurring	17.8928	4.3578	0.0575
Sharpening	23.0896	8.2593	1.3874

G. Security Considerations

The security of the proposed scheme is based on the de-correlating key. The key scrambles the coefficients of wavelet transform, and DCT transform packs energy of the un-correlated coefficients in a zig-zag order. The de-correlating process sandwiched between two transformations provides confusion and diffusion to prevent from deriving the embedding process. Without the correct key, attackers cannot reveal the watermarks for a removal or a remarking attack.

An authorized person can detect the watermark using the de-correlating key even if the image is attacked, because the orthogonality between the watermark and the image is enhanced.

V. CONCLUSIONS

This work presents a robust watermarking scheme that embeds two watermarks to an image to protect copyright based on the de-correlating principle. The de-correlating process introduces significance and randomness to improve the robustness of the watermark without degrading the image. This process wedged between two transformations also provides the security to resist a remarking attack. The two watermarks play complementary roles to resist a removal attack. The proposed scheme can withstand various signal processing, including lossy compression, sharpening, blurring, filtering, resizing, and adding noise. The experimental results show that the proposed method outperforms other schemes in terms of robustness.

In summary, the proposed method offers two advantages. (1) The scheme does not depend on the original image to detect the watermark. Thus, it may be applied easily to networks such as the Internet. (2) The wavelet's hierarchical structure reduces computational overhead for watermark detection. Thus, the scheme is suitable to large images. Future work will focus on a public watermarking system to improve the security of the scheme without revealing the contents of the watermark.

REFERENCES

[1] Pitas, I., "A method for signature casting of digital images," Proceedings of the 1996 IEEE International Conference on Image Processing, 1996, pp. 215-218.

[2] Wu, D.-C., and Tsai, W.-H., "Spatial-domain image hiding using an image differencing," IEE Proceedings-Vision, Image and Signal Processing, vol. 147, no. 1, pp. 29-37, 2000.

[3] Swanson, M. D., Zhu, B., and Tewfik, A. H., "Transparent robust image watermarking," Proceedings of the 1996 International Conference on Image Processing, 1996, pp. 211-214.

[4] Cox, I. J., Kilian, J., Leighton, F. T., and Shamoon, T., "Secure spread spectrum watermarking for multimedia," IEEE Transactions on Image Processing, vol. 6, no. 12, pp. 1673-1687, Dec. 1997.

[5] Barni, M., Bartolini, F., Cappellini, V., and Piva, A., "A DCT-domain system for robust image watermarking," Signal Processing, vol. 66, pp. 357-372, May 28, 1998.

[6] Kutter, M., "Watermarking resistance to translation, rotation, and scaling," Multimedia Systems and Applications, vol. 3528, SPIE, 1999, pp. 423-431.

[7] Delaigle, J. F., Vleeschouwer, D. C., and Macq, B., "Watermarking algorithm based on a human visual model," Signal Processing, vol. 66, no. 3, pp. 319-335, May 28, 1998.

[8] Voloshynovskiy, S., Herrigel, A., Baumgaertner, N., and Pun, T., "A stochastic approach to content adaptive digital image watermarking," Proceedings of Third International Workshop on Information Hiding, 1999, pp. 211-236.

[9] Hsu, C.-T. and Wu, J.-L., “Hidden digital watermarks in images,” IEEE Transactions on Image Processing, vol. 8, no. 1, pp. 58-68, Jan. 1999.

[10] Kundur, D. and Hatzinakos, D., “A robust digital image watermarking method using wavelet-based fusion,” Proceedings of the 1997 International Conference on Image Processing, 1997, pp. 544-0547.

[11] Inoue, H., Miyazaki, A., Yamamoto, A., and Katsura, T., “A digital watermark based on the wavelet transform and its robustness on image compression,” Proceedings of the 1998 IEEE International Conference on Image Processing, 1998, pp. 391-395.

[12] Xia, X.-G., Boncelet, C. G., and Arce, G. R., “Wavelet transform based watermark for digital images,” Optics Express, vol. 3, no. 12, pp. 497-511, Dec. 1998.

[13] Wang, J. Z. and Wiederhold, G., “WaveMark: Digital image watermarking using daubechies’ wavelets and error correcting codes,” Multimedia Systems and Applications, vol. 3528, SPIE, 1999, pp. 432-439.

[14] Lumini, A. and Maio, D., “A wavelet-based image watermarking scheme,” Proceedings of the IEEE International Conference on Information Technology: Coding and Computing, 2000, pp. 122-127.

[15] Su, P.-C., Kuo, C.-C., and Wang, H.-J., “Blind digital watermarking for cartoon and map images,” Security and Watermarking of Multimedia Contents, vol. 3657, SPIE, 1999, pp. 296-306.

[16] Jayant, N., Johnston, J., and Safranek, R., “Signal compression based on models of human perception,” Proceedings of the IEEE, vol. 81, no. 10, pp. 1385-1422, Oct. 1993.

[17] Pickholtz, R. L., Schilling, D. L., and Millstein, L. B., “Theory of spread spectrum communications-A tutorial,” IEEE Transactions on Communication, vol. COMM-30, pp. 855-884, 1982.

[18] Shannon, C. E., “A mathematical theory of communication,” Bell System Technology Journal, vol. 27, pp. 379-423 and pp. 623-656, 1948.

[19] Lu, C.-S., Huang, S.-K., Sze, C.-J., and Liao, H.-Y., “Cocktail watermarking for digital image protection,” IEEE Transactions on Multimedia, vol. 2, no. 4, pp. 209-224, Dec., 2000.

[20] Said, A. and Pearlman, W. A., “A new fast and efficient image codec based on set partitioning in hierarchical trees,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 3, pp. 243-250, June 1996.