1. INTRODUCTION

Digital watermarking has been proposed as an effective tool for protecting and authenticating the copyright of multimedia such as music, picture and video [1]. The watermarking is a process of inserting data into the multimedia content. The watermarking schemes can be classified into two classes. The first is non-blind scheme [2, 3], which an original image is provided. Using the original image, a synchronization problem can be resolved effectively by the difference. The other is blind scheme [4-8], which does not use the original image in watermark detection. This scheme is obviously more interesting and more widely used. The prominent problem of watermarking is that it must remain after unintended distortion or hostile attacks. A distortion such as rotation, scaling, shearing or changing aspect ratio, can defeat most of existing watermarking schemes [9]. Such distortions can destroy the synchronization, which is required for watermark detection.

In this paper, we propose a blind watermarking scheme, which based on the tree structure of the discrete wavelet transform. The tree structure is similar to the embedded zerotree wavelet (EZW) compression. We adapt an imperceptibility of watermark by using a perceptual model of Just-Noticeable distortion (JND). In addition to the watermark, a desired template is inserted into the discrete Fourier transform (DFT) domain. The cross-ratio of four peaks is kept as key, which is required in detection.

2. EMBEDDING METHOD

2.1 Watermark Embedding

According to the EZW concept, the wavelet coefficients across different scales are correlated. In Fig. 1, the concept can be described as parent-children relationship [5], where a coefficient of a coarser scale is the parent of the four coefficients of the next finer scale at the same spatial location and orientation, and so on. It can be assumed with a high probability that if a parent is smaller than a certain threshold then all its descendants are smaller than the threshold.

Fig. 1 The parent-children relationship of wavelet coefficients
quadruplet \( s_k = (b_k, x_k, y_k, v_k) \), where \( b_k \) is the sub-band number (sub-band 2...10 in Fig.1), \( x_k \) and \( y_k \) are the position in the sub-band and \( v_k \) is the value of the coefficient.

\[ v_k = v_k + \left( \alpha \cdot \text{jnd}_k \cdot \rho_k \right) \]  

where \( v_k \) is the value of \( s_k \) after embedding, \( \alpha \) is the JND scaling factor, \( \text{jnd}_k \) is the JND value of each spatial position, and \( \rho_k \) is a value produced by the pseudo-random number generator. The JND and JND scaling factor will be discussed in Sec. 2.2.

(c) Let \( C_k = \{c_{k,1}, c_{k,2}, c_{k,3}, c_{k,4}\} \) be the children of \( s_k \) in the wavelet decomposition.

d) The children are watermarked:

\[ v_{k,i} = v_{k,i} + \left( \alpha \cdot \text{jnd}_{k,i} \cdot \rho_{k,i} \right) \]  

where \( v_{k,i} \) is the value of \( c_{k,i} \) after embedding, \( \alpha \) is the JND scaling factor, \( \text{jnd}_{k,i} \) is the JND value of each spatial position, and \( \rho_{k,i} \) is a value produced by the pseudo-random number generator.

e) If the magnitude of any \( c_{k,i} \) is larger than \( T / 2 \), then the watermark is embedded in the children of \( c_{k,i} \). The 3(d) and 3(e) step are repeated for all the children of \( c_{k,i} \) and all their descendants as necessary.

(f) The threshold is halved for each sub-band level.

The inverse DWT is performed to obtain the watermarked image.

### 2.2 Perceptual model

According to the fact that JND can be quite accurately estimated from the watermarked image [2]. The JND value obtained from the perceptual model are used to determine the maximum strength of the watermark that can be embedded in every portion of image without affecting the image quality. The model used here can be described in terms of three properties of the human visual system that have been studied in the context of image coding: frequency sensitivity, luminance sensitivity and contrast masking [10, 11]. Our JND profile is shown in Fig.3, where the bright pixels indicate that the strong watermarks can be embedded.

#### 2.3 Template Embedding

Since the DWT coefficients are not invariant under geometric transformation. To achieve the robustness to affine transform, we produced a template in DFT domain [5, 6]. The template embedding is as follows:

1. In order to have a high resolution, the image \( f(x,y) \) is padded with zeros to be extended to the size of 1024x1024. Then the fast Fourier transform (FFT) is performed.

2. Two sets of four points randomly distributed along two lines in the upper half (Fig.4) are selected by the key \( Ke \). Each line lies at different angles. All of the points are between \( h_1 \) and \( h_2 \) shown in Fig.4. The strength of template points is equal to the local average magnitude plus 3 times of the standard deviation.

3. The corresponding sets of collinear points are also embedded in the lower half plane to fulfill the symmetry constraint.

4. Perform the inverse FFT to produce the complete watermarked image.

As defined template, the cross-ratio of length of four collinear points (CR) is kept for detection. The CR of four points located at \((r_{i,j}, \theta_i)\) in polar form is defined as follows:

\[ CR = \frac{(r_{i,j} - r_{i}) (r_{j} - r_{i,j})}{(r_{j} - r_{i}) (r_{i,j} - r_{i})} \]  

![Fig. 4 The template in DFT domain.](image-url)
3. DETECTION METHOD

Since the watermarked image is possibly distorted with the affine transform. The watermark detection process has two phases, template detection and watermark detection, as shown in Fig.5.

3.1 Template Detection

Since a linear transform applied in spatial domain results in a corresponding linear transform in DFT domain [6, 8]. To avoid high computational payload, we use the mean square error (MSE) of the affine transformation matrix. The procedures are performed as follows:

1. The image is padded with zero to the size of 1024x1024. Then the Fast Fourier Transform (FFT) is applied.
2. Extract and record the position of all local peaks in DFT magnitude. According to the angle, divide the peaks into equally spaced bins. This procedure reduces time consumption of the exhaustive search.
3. Search for sets of four collinear peaks \( \theta = \{ \theta_1, \theta_2, \ldots, \theta_4 \} \), \( \theta_i = (r_{ij}, r_{ik}, r_{il}, r_{i}) \) which are arranged according to their radius.
4. The cross-ratio of each \( \theta_i \in \Gamma \) is computed by Eq.(4). Then each matched template is used for estimating the affine matrix \( A_i \), and finding the MSE, as follows:

\[
A_i = (P_i^T P_i)^{-1} (P_i^T P_d) \\
MSE_i = \frac{1}{4} || P_d - A_i P_i ||^2
\]

where \( P_i \) is the set of original points, \( P_d \) is the set of detected point.
5. An \( A_i \), which produces the smallest \( MSE_i \) is chosen for reversing the transformation.

3.2 Watermark Detection

The watermark detector is correlation-based, similar to 1-bit detection [5, 7]. The result of the detection is only a presence or absent of the watermark. The watermark detection process is similar to the embedding process in Sec.2.1. We obtain the watermark from an input image with detection key \( Kd = K e \).

The significance of wavelet coefficients is determined by a threshold \( T_d \) which \( T_d \geq T_e \). That is, \( T_d \) is larger than \( T_e \) to avoid correlating coefficients that are not watermarked. Once the watermark is detected, it is correlated with the significant coefficients of the test image:

\[
Z = \frac{1}{M} \sum_{i=1}^{M} \hat{V}_i \rho_i
\]

where \( \hat{V}_i \) is the \( i-th \) wavelet coefficient of the test image, \( M \) is the total number of the watermark, \( \rho_i \) is PN which is corresponding to the \( i-th \) wavelet coefficient. The detection threshold is similar to Dugad method [7].

\[
S = \frac{1}{2M} \sum_{i=1}^{M} |jnd_i|
\]

If \( Z \geq S \), it means the watermark is present in the test image.

4. EXPERIMENTAL RESULTS

We have tested the proposed algorithm. In our tests, we choose \( \sigma = 3.0 \), \( T_e = 30 \), \( T_d = 45 \), \( h_1 = 0.25 \) normalized frequency and \( h_2 = 0.35 \) normalized frequency. We use C language as programming tool on 664 MHz Pentium PC.

Fig. 6(a) shows a watermarked image that has been undergone rotation by 30 degree, while Fig. 6(b) shows the distorted image by general linear transform of the Stirmark 3.1; linear_1.01_0.013_0.009_1.011. Fig 6(c) and Fig. 6(d) show the images after recovered with our template detection, respectively. Fig. 7 shows a result of robustness against JPEG compression that can be detected under high compression. Table 1 shows results of our algorithm by using the StirMark 3.1. In Table 1, “1” represents the embedded watermark can be detected successfully while “0” means the watermark can not be detected successfully. However, robustness against translation is not included in this paper yet.

We observed that the PSNR of the watermarked images, as in Fig. 8, are higher than 40 dB. From the advantage of JND, the watermarks are perceptually invisible. The embedding process takes about 5 seconds, while the detection process take about 10-20 seconds.
5. CONCLUSIONS

This paper described a new watermarking technique based on the DWT-DFT composite scheme. The technique uses a perceptual model and collinear cross-ratio for synchronization template. Since the advantage of perceptual model, the watermarked image is imperceptible. The results show that the proposed scheme is robust against various signal processing attacks and JPEG compression. However, the proposed scheme does not provide a method against the translation attack. It needs to be improved.

REFERENCES