TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7516~7524
e-ISSN: 2087-278X

A Novel HVS-based Watermarking Scheme in Contourlet Transform Domain

Hongbo Bi*[1,2], Xueming Li[1], Yubo Zhang[2]
[1]Beijing Key Laboratory of Network System and Network Culture, Beijing, China
[2]School of Electrical Information Engineering, NorthEast Petroleum University, Daqing, China
*Corresponding author, e-mail: bhbdq@126.com

Abstract

In this paper, a novel watermarking technique in contourlet transform (CT) domain is presented. The proposed algorithm takes advantage of a multiscale framework and multi-directionality to extract the significant frequency, luminance and texture component in an image. Unlike the conventional methods in the contourlet domain, mask function is accomplished pixel by pixel by taking into account the frequency, the luminance and the texture content of all the image subbands including the low-pass subband and directional subbands. The adaptive nature of the novel method allows the scheme to be adaptive in terms of the imperceptibility and robustness. The watermark is detected by computing the correlation. Finally, the experimental results demonstrate the imperceptibility and the robustness against standard watermarking attacks.

Keywords: *digital watermarking, human visual system, mask function, blind detection*

1. Introduction

The rapid development of multimedia technique and computer networks brought about the issue of copyright protection. Digital watermarking can be considered as a solution to the problem of intellectual property rights (IPR) of data contents, which embedded the secret identity information into the host data in such a way that it usually exhibits imperceptibility and robustness against the intentional or unintentional operations such as compression, linear and nonlinear filtering, noise and geometric transformations [1-3]. Digital watermarking has been widely applied in many fields such as content authentication, fingerprinting, and service tracing etc [4, 5].

Imperceptibility and robustness are two most important requirements of digital watermarking. For meet these two requirements simultaneously, people have proposed many schemes in either spatial or transform domain. Space domain methods hide information by changing space domain characteristic of host data, while transform domain methods by changing some coefficients in transform domain of host data. The familiar watermark algorithms in transform domain include discrete fourier transform (DFT), discrete cosine transform (DCT), discrete wavelet transform (DWT), singular value decomposition (SVD), discrete contourlet transform (CT), and so on [6-8].

As an efficient geometric representation of natural images, CT has attracted many researchers' eyes. CT is a true representation of the digital image, which provides a flexible multiresolution representation for two-dimensional signal [9]. Compared with DWT, CT possesses the characteristics of the directionality and anisotropy. Since its coefficients are sparse, and represent the intrinsic property, CT has been widely used in a variety of image processing such as compression, denoising, enhancement, etc. Meanwhile, researchers began to apply CT to the image watermarking [10-13].

This paper proposes an adaptive watermarking method in contourlet domain. The mask of frequency, luminance and texture by human eyes has been examined in the framework of contourlet transform technique. According to the results, the embedding strength of watermark components is determined adaptively. Experimental results demonstrate the novel embedding strategy and demonstrate that the proposed watermarking algorithm are invisible and very robust against noise and common image processing techniques such as lossy compression, filtering, cropping and noise addition.

Received June 27, 2013; Revised August 11, 2013; Accepted September 2, 2013

The outline of the rest of the paper is organized as follows. In section 2, the contourlet transform algorithm is mentioned, the Laplacian pyramid and directional filter banks are introduced. In section 3, we describe the just noticeable distortion profile of human vision, which considers the influence of frequency, luminance and texture. In section 4, the proposed watermarking scheme is presented, we give the embedding and the blind detection of the watermarking. In section 5, some experimental results are mentioned. Finally, conclusions are given.

2. The Contourlet Transform

DWT provides the characteristics of multi-scale decomposition and time-frequency localization which facilities the representation of the image. However it can only offer the information in few directions and low-pass components. The contourlet transform is a novel image decomposition scheme, which can capture the intrinsic geometrical structure in visual information. It exploits the Laplacian Pyramid (LP) for the multiresolution decomposition of the image. Afterwards, a directional decomposition is performed on every bandpass image using directional filter banks. It can obtain a sparse expansion for natural image by specifying the number of directional bands at any level. As a result, the image is represented as a set of directional subbands at multiple scales.

CT is illustrated in Figure 1.

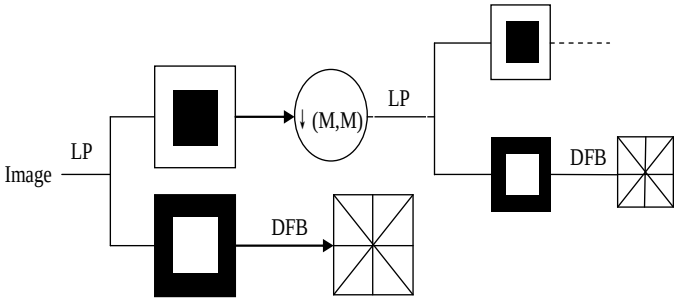

Figure 1. Block Diagram of CT

Composed of low-pass filtering and downsampling, LP is indeed the high frequency components of the gaussian pyramid (GP) at the same scale, that is, the detailed parts of the image. LP image can be obtained by subtracting the two neighbouring images in GP. Generally, we need to expand the finer scale image to the coarser scale, namely, perform the interpolation to the rows and columns of the image, afterwards, the interpolated results through a low-pass filter will subtract the image at the coarser scale. The reconstruction of LP is the inverse of the decomposition.

Suppose the image $IMG \in L_2(\Re^2)$, LP in CT uses orthogonal filters and downsampling by 2 in each dimension, the l-level of LP decomposes IMG into a coarser image V_J and a sequence of detail image I_j, where V_J is the approximation subspace at the scale l ($l = 2^J$). I can be denoted as:

$$L_2(\Re^2) = V_J \oplus (\bigoplus_{j=J}^{1} I_j) \tag{1}$$

The directional filter bank (DFB) is generally implemented via an l-level binary tree decomposition that leads to 2^l subbands with wedge-shaped frequency partition as shown in Figure 2. In CT, the simplified DFB is intuitively constructed from two building blocks. The first building block is a two-channel quincunx filter bank with fan filters that divides a 2-D spectrum

into two directions: horizontal and vertical. The second building block of the DFB is a shearing operator, which amounts to just reordering of image samples.

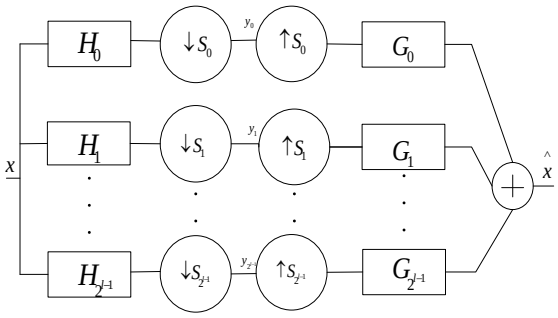

Figure 2. l-level Binary Tree Decomposition

In CT, the DFB partition further I_j from LP. The results are detail subbands in multiple directions.

$$I_j = \bigoplus_{k=0}^{2^{l_j}-1} I_{j,k}^{l_j}, (j,k) \in Z^2 \quad (2)$$

Figure 3 shows an example of CT on the “Lena” image.

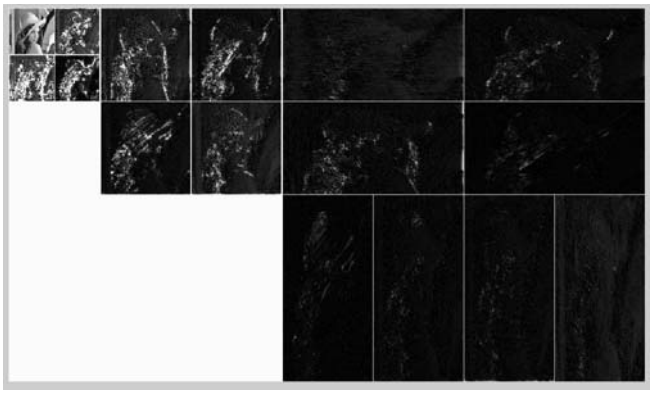

Figure 3. Decomposition of Lena by CT

The image is decomposed into a low-pass subband and a set of directional subbands. We notice that CT effectively represents the true image where edges are localized in both location and direction.

3. Just Noticeable Distortion (JND) Profile Analysis in CT

The contourlet transform provides a multiscale and multidirectional representation of an image. It is easily adjustable for detecting fine texture detail in any orientation at various scale level. In order to conform to the multiresolution nature of human visual system and enhance the performance of the watermark system, we calculate the weight parameters according to the frequency, luminance and texture complication of detail subband.

In order to embed into the host the maximum, but still imperceptible, HVS has to be considered. Lewis and Knowles tackle the problem of DWT coefficients quantization for compression purposes. Details of the improved HVS-based watermarking method are presented by Barni et al. in 2001 and some modifications of the model is proposed in order to better fit the model to the watermarking system, which has been widely cited by many researcher. However, DWT can only provide information in 3 directions, that is, horizontal, vertical and diagonal, while CT can decompose the host along arbitrary directions at arbitrary scale, which provides more details information than DWT and can be considered as the potential solution to improve the performance of the watermarking system. Details of the HVS characteristics in CT are mentioned here.

The Just Noticeable Distortion (JND) in the CT domain can be typically expressed as the product of three terms. As a result, the HVS mask function is adaptive to the CT coefficients.

Considering the sensitivity of the human eye, Xiao et al. proposed the weight coefficient calculation as shown below:

$$M(J,x,y) = M_f(k,J) \cdot M_l(J,x,y) \cdot M_t(J,x,y)^{0.25} \tag{3}$$

Where $M_f(k,J)$ denotes frequency sensitivity as shown in Equation (4). $M_l(J,x,y)$ denotes local luminance for gray levels in with reference to Equation (5) and (6). $M_t(J,x,y)$ denotes the influence of the texture as indicated in the Equation (7), (8) and (9).

$$M_f(k,J) = \begin{Bmatrix} \frac{\sqrt{2}}{2}, k = 4n \\ 1, k = 4n+1 \\ \sqrt{2}, k = 4n+2 \\ 2, k = 4n+3 \end{Bmatrix} \times \begin{Bmatrix} 1, J = 1 \\ 0.32, J = 2 \\ 0.16, J = 3 \\ 0.1, J = 4 \end{Bmatrix} \tag{4}$$

$$M_l(J,x,y) = \begin{cases} \overline{L}(J) + \left(\overline{L}(J) - L(J,x,y)\right) & \overline{L}(J) > L(J,x,y) \\ L(J,x,y) & \overline{L}(J) \le L(J,x,y) \end{cases} \tag{5}$$

$$L(J,x,y) = \frac{1}{256} c_{k+1}\left(\left\lfloor \frac{x}{2^{4-J}} \right\rfloor, \left\lfloor \frac{y}{2^{4-J}} \right\rfloor \right) \tag{6}$$

$$M_t(J,x,y) = E(J,x,y) + D(J,x,y) \tag{7}$$

$$E(J,x,y) = \sum_{0}^{4-J} \sum_{i=0}^{1} \sum_{j=0}^{1} \left[c_{k+1}\left(j + \frac{x}{2^{4-J}}, i + \frac{y}{2^{4-J}} \right) \right]^2 \tag{8}$$

$$D(J,x,y) = \delta \left\{ c_{k+1}\left(1 + j + \frac{x}{2^{4-J}}, 1 + i + \frac{y}{2^{4-J}} \right) \right\}_{\substack{i=0,1 \\ j=0,1}} \tag{9}$$

Where $c_k(x,y)$ denotes the decomposition coefficient at (x,y) in the subband k. $E(J,x,y)$ denotes the local sum of squares, $D(J,x,y)$ denotes the local variance.

Assuming changes smaller than one half of the calculated mask function are visually imperceptible, $M(J,x,y)$ gives maximum embedding threshold in the quantization of CT coefficients using:

$$T(J,i,j) = M(J,x,y)/2 \quad (10)$$

According to the mention above, it is apparent that the computed mask function $M(J,x,y)$ at each pixel enables HVS-based watermarking to obtain high level of imperceptibility and robustness without consideration of the computations in equations.

4. Proposed Watermarking System

4.1. Watermarking Embedding

CT coefficients contain an approximate low-pass subband and several detail directional subbands at each level. The low-pass subband represents basic information of the image, which is the most important part for image reconstruction, embedding the watermark in these coefficients may achieve robustness against intentional or unintentional attacks, but may degrade the visual quality, since the human visual system (HVS) is less sensitive to high frequencies, embedding the watermark in the high frequency subbands improves the perceptibility of the watermarked image, but it is hardly robust.

In our scheme, as shown in Figure 4, the watermark is embedded into both the low-pass subband and the directional subbands by different HVS characteristics. Consequently, the proposed watermarking scheme is robust to the widely spectral attacks resulting from both the low and high frequency image processing.

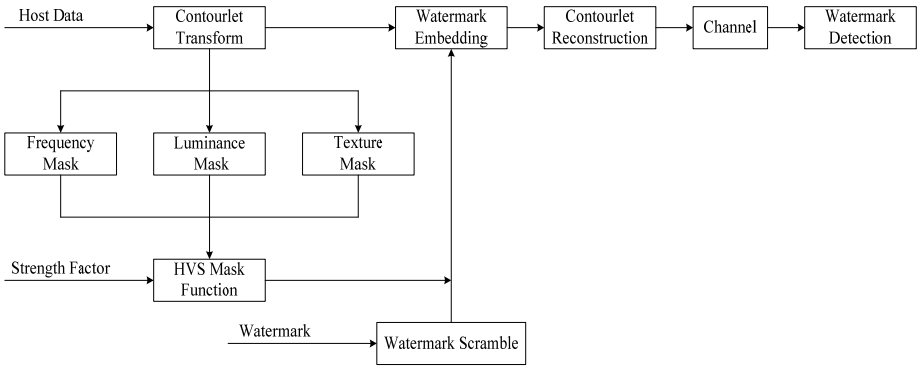

Figure 4. Watermarking Scheme

The general embedding steps for the proposed watermarking scheme are described as follows.

Step 1. Watermarking preprocessing

The watermark information need be preprocessed in order to weaken the correlation of watermark image pixels and enhance system robustness. In our scheme, we treat the watermark image using Arnold scramble as shown in Equation (11).

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \pmod{N} \quad (11)$$

Where (x,y) is the pixel of the watermarking image, (x',y') is the pixel of the watermarking image after scramble.

Since the Arnold transform is periodic, the number of scrambling can be also considered as the key to enhance the security.

Step 2. CT of the whole image

We propose to embed the watermark in the low-pass subband and one of the directional subbands of the highest level, that is, J^{th} level, which actually are mid-frequency

subbands in terms of the whole frequency band of the image. This is an appropriate tradeoff between the visual perceptibility and the robustness.

Step 3. Determination of the embedding position

Since HVS is less sensitive to the texture, the watermarking embedding directional subband can be chosen combined with the texture in the image. The energy distribution corresponds to the texture property, so we embed the watermarking in the most textured directional subband.

The energy is computed as Equation (12).

$$E_{J,d} = \frac{1}{M_{J,d} M_{J,d}} \sum_{x=1}^{M_{J,d}} \sum_{y=1}^{M_{J,d}} C_{J,d}^2(x,y) \tag{12}$$

Where $M_{J,d}$ represents the width and height of the d[th] directional subband in the J[th] level, $C_{J,d}$ represents the corresponding coefficients, (x, y) is the coordinate in the subband. The bigger $E_{J,d}$ is, the more texture is, which means the subband contributes more to the image.

Step 4. Embedding the watermark

According to the frequency, luminance and texture complication value of the subband, a binary watermark is embedded into the host image, by modifying the value of the corresponding coefficients. Thus, the embedding can be described as:

$$IM_w(i,j) = IM(i,j) + \alpha \cdot T(J,x,y) \cdot W(i,j) \tag{13}$$

Where $IM_w(i,j)$ represents the watermarked subband coefficients, $IM(i,j)$ represents the original subband coefficients, α denotes the global embedding factor that determines the embedding strength, $T(J,x,y)$ is the HVS mask function as formulated in (10), $W(i,j)$ denotes the binary pseudorandom distributed watermark, which is arranged in the form of two dimensions.

Step 5. Image Reconstruction

Finally the watermarked image is obtained by the contourlet reconstruction of subband coefficients combined with the remained coefficients followed by Arnold scramble.

4.2. Watermarking Detection

Maximum-likelihood detection is used to extract each embedded bit from the watermarked signal coefficients. The watermark is recovered by means of contourlet reconstruction and calculations in reverse order of embed process.

Step 1. Perform contourlet transform for the watermarked image to get contourlet coefficient of all the subbands and determine the subband with maximum energy.

Step 2. Detect the correlation. The correlation is calculated as:

$$\rho = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} c'_{l,k}(x,y)\, w(x,y) \tag{14}$$

Due to there are some distortions of detected watermark in some degree, τ is set to be threshold, if $\rho > \tau$, the watermark is exist, if $\tau < P$, the watermark is not exist. The threshold τ is related to the false alarm probability and false dismissal probability.

We make the assumption that $c_{l,k}(x,y)$ are zero mean, independent variables. By exploiting the central limit theorem, we can also consider that ρ is normally distributed [14-15]. Under these hypotheses, it can be easily deduced that the mean values of ρ in cases A (not watermarked), B (watermarked with another wrong watermark) and C (watermarked with the correct watermark) are:

$$A: \mu_{\rho A} = 0 \tag{15}$$

$$B: \mu_{\rho B} = 0 \tag{16}$$

$$C: \mu_{\rho C} = \frac{\alpha}{M_{l,k} N_{l,k}} \sum_{x=1}^{M} \sum_{y=1}^{N} E\left[H_{l,k}(x,y) \right] \tag{17}$$

Where $E[\cdot]$ denotes the expectation.

The false alarm probability $P_f = \Pr ob\left(\rho > \tau \left| AorB \right. \right)$, for case A:

$$\sigma_{\rho A}^2 = \frac{\sigma_W^2}{\left(M_{l,k} N_{l,k} \right)^2} \sum_{x=1}^{M} \sum_{y=1}^{N} E\left[c_{l,k}(x,y)^2 \right] \tag{18}$$

For case B:

$$\sigma_{\rho B}^2 = \frac{\sigma_W^2}{\left(M_{l,k} N_{l,k} \right)^2} \sum_{x=1}^{M} \sum_{y=1}^{N} E\left[c_{l,k}(x,y)^2 \right] + \alpha^2 \sigma_W^2 E\left[H_{l,k}(x,y)^2 \right] \tag{19}$$

Where σ^2 denotes the variance, and:

$$E\left[c_{l,k}^{'}(x,y)^2 \right] = E\left[c_{l,k}(x,y)^2 \right] + \alpha^2 E\left[H_{l,k}(x,y)^2 W(x,y)^2 \right]$$
$$+2\alpha E\left[c_{l,k}(x,y) H_{l,k}(x,y) W(x,y) \right] \tag{20}$$

According to the assumption $\sigma_W^2 = 1$, $E\left[c_{l,k}(x,y) \right] = E\left[W(x,y) \right] = 0$, and $W(x,y)$ is irrelated with $c_{l,k}(x,y)$ with each other, so:

$$\sigma_{\rho B}^2 = \frac{1}{\left(M_{l,k} N_{l,k} \right)^2} \sum_{x=1}^{M_{l,k}} \sum_{y=1}^{N_{l,k}} E\left[c_{l,k}^{'}(x,y)^2 \right] \tag{21}$$

Adopting the unbiased estimation of $\sigma_{\rho B}^2$:

$$\sigma_{\rho B}^2 \approx \frac{1}{\left(M \times N \right)^2} \sum_{x=1}^{M} \sum_{y=1}^{N} c_{l,k}^{'}(x,y)^2 \tag{22}$$

From which it is readily seen that case B is the worse case, since the higher the variance the higher the error probability. So, we get:

$$P_f \le \frac{1}{2} erfc\left(\frac{\tau}{\sqrt{2\sigma_{\rho B}^2}} \right) \tag{23}$$

When $P_f \le 10^{-8}$, we get:

$$\tau = 3.97\sqrt{2\sigma_{\rho B}^2} \tag{24}$$

5. Experimental Results

The algorithm has been extensively tested on various standard images and different kinds of attacks. Some of the most significant results will be shown. For the experiments presented in the following, the host image is "Lena" with the size of 512*512, the false alarm is 10^{-8}, CT exploits "9-7" pyramid filter and "pkva" directional filter to obtain a 2-level decomposition. First, watermark invisibility is valuated. In Figure 5, the original "Lena" image is presented on the left, while the watermarked copy is shown on the right. We can see that the images are evidently undistinguishable, proving the effectiveness of CT watermarking and the masking procedure.

Figure 5. Host Image (left) and the Watermarked Image (right) under No Attacks

Other results after attacks are listed in Table 1. It is apparent that the proposed scheme is resistant to many signal operation.

Table 1. Robustness Experimental Results

Attacks		parameters	PSNR	Detection Results
noise	Peper & Salt	0.001	35.54	Yes
		0.010	25.43	Yes
		0.020	22.57	Yes
	Gauss	0.001	30.00	Yes
		0.003	25.22	Yes
JPEG Compression		70%	37.29	Yes
		50%	35.78	Yes
		30%	34.26	Yes
Filtering	Average	3*3	33.75	Yes
		6*6	27.91	Yes
	Median	3*3	35.41	Yes
		6*6	27.77	Yes
Geometric	Scaling	25%	28.58	Yes
		75%	38.64	Yes
		150%	44.77	Yes
	Cropping	15%	21.79	Yes
		25%	18.80	Yes

6. Conclusion

In this paper, a novel image watermarking scheme has been presented. The algorithm embeds the watermark code by modifying the CT coefficients of the image, and exploits a model adapting the watermark strength to the characteristics of the HVS. The method uses frequency, luminance and texture analysis to model the HVS characteristics, increasing the watermark strength without great perceptible distortion. The experimental results show that the proposed method is robust against many signal processing attacks and the behavior of the watermark detector was good.

Acknowledgement

This work is supported by the Science & Technology Project of Heilongjiang province under Grant No. 12521056.

References

[1] Ingemar J Cox, Joe Kilian, F Thomson Leighto, et al. Secure Spread Spectrum Watermarking for Multimedia. *IEEE Transactions On Image Processing*. 1997; 6(12): 1673-1687.
[2] Scott Craver, Nasir Memon, Boon-Lock Yeo, et al. Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications. *IEEE Journal On Selected Areas In Communications*. 1998; 16(4): 573–586.
[3] Qiming Liu. An Adaptive Blind Watermarking Algorithm for Color Image. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(1): 302-309.
[4] Deepa Kundur, Dimitrios Hatzinakos. Diversity and attack characterization for improved robust watermarking. *IEEE Transactions On Signal Processing.* 2001; 49(10): 2383-2396.
[5] Huming Gao, Liyuan Jia, Meiling Liu. A Digital Watermarking Algorithm for Color Image Based on DWT. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(6): 3271-3278.
[6] Hongbo BI, Yubo ZHANG, Xueming LI. Video watermarking robust against spatio-temporal attacks. *Journal of Networks*. 2011; 6(6): 932-936.
[7] Braci S, Boyer R, Delpha C. *Analysis of the resistance of the Spread Transform against Temporal Frame Averaging attack*. International Conference on Image Processing. HongKong. 2010; 1: 213-216.
[8] Yubo ZHANG, Hongbo BI. Transparent Video Watermarking Exploiting Spatio-Temporal Masking in 3D-DCT Domain. *Journal of Computational and Information Systems*. 2011; 7(5): 1706-1713.
[9] MN Do, Martin Vetterli. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Transactions On Image Processing*. 2005; 14(12): 2091-2106.
[10] Shiva Zaboli, M Shahram Moin. *CEW: A Non-Blind Adaptive Image Watermarking Approach Based On Entropy in Contourlet Domain*. IEEE International Symposium on Industrial Electronics. Vigo. 2007; 1: 1687-1692.
[11] Mohan BC, Kumar SS. Robust digital watermarking scheme using contourlet transform. *International Journal of Computer Science and Network Security*. 2008; 8(2): 43-51.
[12] Haifeng Li, Weiwei Song, Shuxun Wang. Robust image watermarking algorithm based on contourlet transform. *Journal on Communications*. 2006; 27(4): 87-94.
[13] Shangqin Xiao, Hefei Ling, Fuhao Zou, et al. *Adaptive Image Watermarking Algorithm in Contourlet Domain*. The Japan-China Joint Workshop on Frontier of Computer Science and Technology. WuHan. 2007; 1: 125-130.
[14] Barni M, Bartolini F, Pive A. Improved wavelet-based watermarking through pixel-wise masking. *IEEE Transactions on Image Processing*. 2001; 5(10): 783–91.
[15] Hazem A. Al-Otum, Allam O. Al-Taba'a. Adaptive color image watermarking based on a modified improved pixel-wise masking technique. *Computers and Electrical Engineering*. 2009; 35(5): 673-695.