IEICE TRANS. FUNDAMENTALS, VOL.E91–A, NO.6 JUNE 2008

PAPER Special Section on Image Media Quality

Wavelet Based Multipurpose Color Image Watermarking by Using Dual Watermarks with Human Vision System Models

Min-Jen TSAI[†a)], *Member and* **Chih-Wen LIN**[†], *Nonmember*

SUMMARY In this study, we propose a complete architecture based on digital watermarking techniques to solve the issue of copyright protection and authentication for digital contents. We apply visible and semi-fragile watermarks as dual watermarks where visible watermarking is used to establish the copyright protection and semi-fragile watermarking authenticates and verifies the integrity of the watermarked image. In order to get the best tradeoff between the embedding energy of watermark and the perceptual translucence for visible watermark, the composite coefficients using global and local characteristics of the host and watermark images in the discrete wavelet transform (DWT) domain is considered with Human Vision System (HVS) models. To achieve the optimum noise reduction of the visibility thresholds for HVS in DWT domain, the contrast-sensitive function (CSF) and noise visible function (NVF) of perceptual model is applied which characterizes the global and local image properties and identifies texture and edge regions to determine the optimal watermark locations and strength at the watermark embedding stage. In addition, the perceptual weights according to the basis function amplitudes of DWT coefficients is fine tuned for the best quality of perceptual translucence in the design of the proposed watermarking algorithm. Furthermore, the semi-fragile watermark can detect and localize malicious attack effectively yet tolerate mild modifications such as JPEG compression and channel additive white Gaussian noise (AWGN). From the experimental results, our proposed technique not only improves the PSNR values and visual quality than other algorithms but also preserves the visibility of the watermark visible under various signal processing and advanced image recovery attacks.

key words: *HVS, semi-fragile watermarking, tamper detection, visible watermarking*

1. Introduction

We are now in an era of knowledge-based economy. At the core of such an economy, intellectual property becomes the critical issue we concerned. Intellectual property acts like real property and surrounds us in nearly everything we do. Books, music, digital multimedia, and any kind of arts actually belong to the authors who made it, and the authors have the rights to restrict access to intellectual property [1].

Because of the advantages of digital media and rapid development of digital signal processing, a variety of multimedia contents have been digitalized and easily distributed or duplicated without any reduction in quality through both authorized and unauthorized distribution channels. Digital watermarking [2] has been extensively studied and regarded as a potentially effective means for protecting copyright of digital media in recent years, Visible watermarking schemes protect copyrights in a more active method. They not only prevent pirates but also recognize copyright of multimedia data. Digital contents embedded with visible watermarks will overlay recognizable but unobtrusive copyrights patterns identifying its ownership. Therefore, a useful visible watermarking technique should remain details of contents and ensure embedded patterns difficult or even impossible to be removed, and no one could use watermarked data illegally. Watermark removal, at a minimum, should be more costly and labor intensive than purchasing rights to use the digital data.

For content authentication and integrity verification, fragile (or semi-fragile) watermarks are used because they are fragile to certain alterations and modifications of the authenticated multimedia. Semi-fragile watermarks are more practical than fragile watermarks, since they are robust to some mild modifications such as JPEG compression and channel AWGN (additive white Gaussian noise) causing by exchange and storage but fragile to malicious attacks like image cropping which crops objects from a source and pastes them onto a target.

The goal of this paper is to propose a novel scheme for copyright protection and authentication of color images by using visible watermark and semi-fragile watermark with HVS models. For copyright protection, we present a differential visible watermarking algorithm based on noise reduction and HVS model to get the best tradeoff between the embedding energy of watermark and the perceptual translucence for visible watermark. The collaboration of CSF and NVF for HVS models is leveraged with the noise reduction of the visibility thresholds for HVS in DWT domain. The perceptual weights is fine tuned for watermark embedding which results significant improvement over the watermarked images by CSF only algorithms regarding the image quality, translucence and robustness of the watermarking. For authentication and verifying the integrity of the watermarked images, we applied a semi-fragile watermark algorithm which can detect and localize malicious attack and the order of embedding is visible watermark first and semi-fragile watermark next.

The rest of this paper is organized as follows. Related works about visible watermarking and image authentication will be introduced briefly in Sect. 2. The details of the algorithm will be explained in Sect. 3. Section 4 will show the experiments results with discussion and conclusion is in Sect. 5, respectively.

Manuscript received June 4, 2007.
Manuscript revised October 16, 2007.
[†]The authors are with Institute of Information Management, National Chiao Tung University, Taiwan, R.O.C.
a) E-mail: mjtsai@cc.nctu.edu.tw
DOI: 10.1093/ietfec/e91–a.6.1426

2. Related Works

2.1 Visible Watermarking

Visible watermarking techniques are used to protect copyright of digital multimedia (audio, image or video) that have to be delivered for certain purpose, such as digital multimedia used in exhibition, digital library, advertisement or distant learning web, while illegal duplicate is forbidden. From the literature survey, the visible watermarking has captured significant attention [3] since there are not only different visible watermarking approaches either in spatial or transform domain but also various visible watermark removal schemes.

Braudaway et al. [4] proposed one of the early approaches for visible watermarking by formulating the nonlinear equation to accomplish the luminance alteration in spatial domain. Meng and Chang [5] applied the stochastic approximation for Braudaway's method in the discrete cosine transform (DCT) domain by adding visible watermarks in video sequences. Mohanty et al. [6] proposed a watermarking technique called dual watermarking by combining of a visible watermark and an invisible watermark in the spatial domain. The visible watermark adopted to establish the owner's right to the image and invisible watermark to check the intentional and unintentional tampering of the image. Chen [7] has proposed a visible watermarking mechanism to embed a gray level watermark into the host image based on the statistic approach by the standard deviation of blocks.

Hu and Kwong [8], [9] implemented an adaptive visible watermarking in the wavelet domain by using the truncated Gaussian function to approximate the effect of luminance masking for the image fusion. Based on image features, they first classify the host and watermark image pixels into different perceptual classes. Then, they use the classification information to guide pixel-wise watermark embedding. In high-pass subbands, they focus on image features, while in the low-pass subband, they use truncated Gaussian function to approximate the effect of luminance masking. Yong et al. [10] also proposed a translucent digital watermark in the DWT domain and use error-correct code to improve the ability to anti-attack.

Each of above schemes wasn't devoted to better feature-based classification and the use of sophisticated visual masking models. Recently, Huang and Tang [3] presented a contrast sensitive visible watermarking scheme with the assistance of HVS. They first compute the CSF mask of the discrete wavelet transform domain. They later use square function to determine the mask weights for each subband. At last, they adjust the scaling and embedding factors based on the block classification with the texture sensitivity of the HVS for watermark embedding. However, their scheme doesn't consider the following issues:

1. The basis function of the wavelet transform plays an important role during the application of CSF for the HVS in the wavelet transform domain.
2. The embedding factors emphasize more weights in the low frequency domain instead of the medium-to-high frequency domain.
3. The interrelationship of block classification and the characteristics of the embedding location.

For the first issues, the direct application of CSF for the HVS in the wavelet transform domain needs to be further studied [11]–[13] while the basis function of the wavelet transform is a critical factor to affect the visibility of the noise in the DWT domain. For the second issue, the watermark embedding in the low frequency components results high degradation of the image fidelity. In addition, the high frequency components of the watermarked image easily suffer common image signal processing attacks with low robustness. For the third issue, the plane, edge and texture block classification in [3] is a genuine approach should the local and global characteristics of wavelet coefficients be further considered.

2.2 Image Authentication and Temper Detection

Semi-fragile watermarking schemes have been proposed to verify the integrity of digital contents and tolerate some degree of mild modifications such as JPEG compression and channel AWGN. Interested readers could refer [14], [15] for latest development in this topic. Regarding the watermark embedding approach, the fragile (semi-fragile) watermarks can be embedded in the spatial domain or the transformed domain while the schemes embedding watermarks in the transformed domain offer a higher degree of robustness [2]. Recently, many semi-fragile methods are based on wavelet transform domain since it can resist a certain degree of attacks and have the spatial and frequency localization of digital data by the nature of multiresolution discrete wavelet decomposition.

Kundur and Hatzinakos [16] proposed one of the first approaches to semi-fragile watermarking called telltale tamper proofing. They embed a watermark in the discrete wavelet domain of the image by quantizing the corresponding coefficients. They also use a statistics-based tamper assessment function as measurement for tamper proofing and authentication. H.P. Alexandre et al. [17] proposed a novel technique for content authentication of digital images by quantizing wavelet packet coefficients and adopting characteristics of the human visual system to maximize the embedding weights for improving good imperceptibility of watermarked image. Hua Yuan and Xiao-Ping Zhang [18] proposed a semi-fragile watermarking method based on image modeling using the Gaussian mixture model (GMM) in the wavelet domain. They modify selected wavelet coefficients according to the GMM parameters obtained through an EM algorithm. Ding et al. [19] proposed a wavelet-based chaotic semi-fragile watermarking scheme based on chaotic map and odd-even quantization. Their scheme can detect and

localize malicious attacks with high peak signal-to-noise ratio (PSNR), while allowing more JPEG compression and channel additive white Gaussian noise (AWGN) tolerance. Since [19] is superior in resisting JPEG and AWGN attacks among other semi-fragile approaches, we further modify the scheme and integrate it into the proposed dual watermark approach which will be explained in the next section.

3. The Proposed Approach

The most important requirements in the visible watermarking scheme are the robustness and translucence, but unfortunately these are in confliction with each other. If we increase the energy of watermark to improve its robustness, the problem we get is perceptual translucence decreasing with less image fidelity and vice versa. Therefore, we have to decrease the energy of the watermark to get good perceptual translucence and the embedded watermark will still be robust to intentional or unintentional signal processing attacks. HVS (Human Visual System) is the key factor we have found in providing the good translucence of the watermarked image and a better robustness [3]. A lot of works have been devoted to understanding HVS and offering mathematical models of how humans see the world [12], [20]. Psychovisual studies have shown that human vision has different sensitivity from various spatial frequencies (frequency subbands). Common HVS models are composed of image dependent or independent Just Noticeable Difference (JND) thresholds, so the HVS by using the contrast sensitive function (CSF) and noise visibility function (NVF) is integrated in this study and will be explained in brief as following.

3.1 CSF (Contrast Sensitive Function)

For watermarked images, there has been a need for good metrics for image quality that incorporates properties of the HVS. The visibility thresholds of visual signals are studied by psychovisual measurements to determine the thresholds. These measurements were performed on sinusoidal gratings with various spatial frequencies and orientations by given viewing conditions. The purpose of such study was to determine the contrast thresholds of gratings by the given frequency and orientation. Contrast as a measure of relative variation of luminance for periodic pattern such as a sinusoidal grating is given by the equation

$$C = (L_{\max} - L_{\min})/(L_{\max} + L_{\min}) \quad (1)$$

where L_{max} and L_{min} are maximal and minimal luminance of a grating. Reciprocal values of contrast thresholds express the contrast sensitivity (CS), and Mannos and Sakrison [20] originally presented a model of the contrast sensitive function (CSF) for luminance (or grayscale) images is given as follows:

$$H(f) = 2.6 * (0.0192 + 0.114 * f) * e^{-(0.114*f)^{1.1}} \quad (2)$$

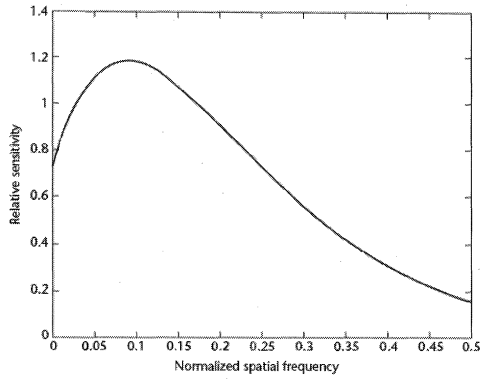


Fig. 1 Luminance sensitivity curve of CSF.

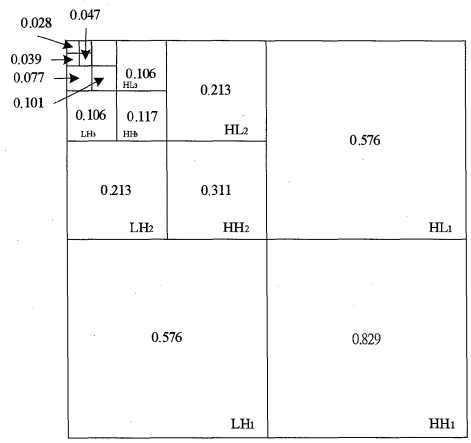


Fig. 2 $\beta_{\lambda,\theta}$ in different DWT level and orientation.

where $f = \sqrt{f_x^2 + f_y^2}$ is the spatial frequency in cycles/degree of visual angle (f_x and f_y are the spatial frequencies in the horizontal and vertical directions, respectively). Figure 1 depicts the CSF curve which characterizes luminance sensitivity of the HVS as a function of normalized spatial frequency. According to the CSF curve, we can see that the HVS is most sensitive to normalized spatial frequencies between 0.025 and 0.125 and less sensitive to low and high frequencies [3]. Therefore, this knowledge from CSF can be used to develop a simple image independent HVS model.

CSF masking [11], [12] is one way to apply the CSF in the discrete wavelet domain. CSF masking refers to the method of weighting the wavelet coefficients according to their perceptual importance. Some well-designed CSF masks which transforms the CSF curve in Fig. 2 into perceptual importance weight are presented in [11]. Huang and Tang [3] use the same method led to 11-weight DWT CSF mask in the five-level wavelet transform. Figure 3 illustrates the 11-weight DWT CSF mask with the weights shown for each subband.

Psychovisual studies have shown that the HVS has a general band-pass characteristic [3]. For a five-level pyramidal DWT decomposition, the HVS is most sensitive to

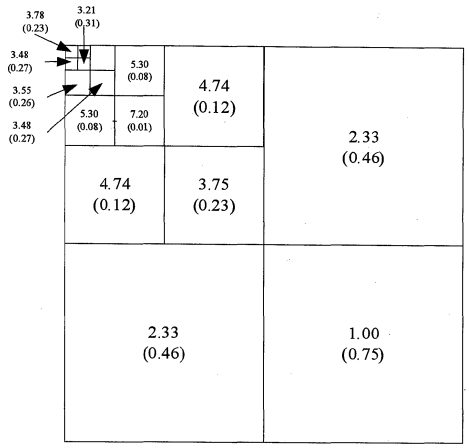


Fig. 3 A five-level DWT structure. $r^{\lambda}(\beta^{\lambda})$ values for each level λ are indicated at the center of each band.

the distortion in mid-frequency regions and sensitivity falls off as the frequency value drifts on both sides. The square function (3) in [3] is applied to approximate the effect of CSF masking. The adequate modulation rate β^{λ} for each subband is determined by:

$$\beta^{\lambda} = 0.01 + \frac{(7.20 - r^{\lambda})^2}{7.20^2} \tag{3}$$

where λ denotes the decomposed level and β^{λ}s for each subband are shown in Fig. 2 where the level 3 has the smallest rate for modulation. In addition,r^{λ} represents the wavelet coefficient CSF of the perceptual importance weight as shown in Fig. 3. It is apparent that r^{λ}s have the highest values at level 3 and the values falls off for level 1, 2, 4 and 5 in Fig. 3.

3.2 NVF (Noise Visibility Function)

Many schemes embedded the watermark as random noise in the whole host image with the same strength regardless of the local properties of the host image, so the visible artifacts are easy taken placed at flat regions. S. Voloshynovskiy et al. [11] presented a stochastic approach based on the computation of a NVF (Noise Visibility Function) that characterizes the local image properties and identifies texture and edge regions. Accordingly, when the local variance is small, the image is flat, and a large enough variance indicates the presence of edges or highly texture areas. Because human eyes are sensitive to changes in flat than edges regions of the image, ones can increase the energy of watermark in the edges and high textured areas of the image, and reducing it in smooth regions in similar peak-signal noise rate (PSNR). This allows us to determine the optimal watermark locations and strength for the watermark embedding stage. Therefore, this concept from NVF can be used to develop a simple image dependent HVS model.

They developed three such NVF Functions and we adopt the NVF Function with Stationary GG (Generalized Gaussian) Model in this study. The formula is as following:

Table 1 Basis function amplitudes for a five-level 9/7 DWT.

Orientation	Level				
	1	2	3	4	5
LL	0.62171	0.34537	0.18004	0.09140	0.045943
HL	0.67234	0.41317	0.22726	0.11792	0.059758
LH	0.67234	0.41317	0.22726	0.11792	0.059758
HH	0.72709	0.49428	0.28688	0.15214	0.077727

$$NVF(i,j) = \frac{w(i,j)}{w(i,j) + \sigma_x^2} \tag{4}$$

where $w(i,j) = \gamma[\eta(\gamma)]^{\gamma} \frac{1}{\|r(i,j)\|^{2-\gamma}}$ and σ_x^2 is the global variance of host image. $\eta(\gamma) = \sqrt{\Gamma(3/\gamma)/\Gamma(1/\gamma)}$, $\Gamma(t) = \int_0^{\infty} e^{-u}u^{t-1}du$ (gamma function) and $r(i,j) = \frac{x(i,j)-\bar{x}(i,j)}{\sigma_x}$. γ is the shape parameter and $r(i,j)$ is determined by the local mean and the local variance. For most of real images, the shape parameter is in the range $0.3 \le \gamma \le 1$. In our scheme, the estimated shape parameter for $\gamma = 0.65$ and width of window is 1.

3.3 Detection Thresholds for DTW Coefficients

In order to further improve the HVS model for better image quality, the knowledge of detection thresholds for DWT coefficients should be also studied. A.B. Watson, et al. [13] proposed a mathematical model for DWT noise detection thresholds which is a function of level, orientation, and display visual resolution. The model is given by

$$\log Y_{\lambda,\theta} = \log a + k(\log f_{\lambda} - \log g_{\theta} f_0)^2 \tag{5}$$

where a is the minimum threshold occurs at spatial frequency $g_{\theta}f_0$, f_{λ} is the spatial frequency of decomposition level λ, and g_{θ} shifts the minimum by an amount that is a function of orientation. Table 1 shows the basis function amplitudes for a 5-level DWT. In this study, we use $A_{\lambda,\theta}$ indicating the basis function amplitudes, λ as DWT level, and θ as orientation.

3.4 Visible Watermarking Embedding Algorithm

The complete design of the visible watermarking algorithm is summarized as following and the flow chart is shown in Fig. 4:

(1) The host color image is converted in the color space domain from RGB to YCrCb.
(2) By using Bi9/7 filter from [13], compute the 5-level 2-D wavelet coefficients of Y component from host color image and grayscale watermark image. If the width of watermark is not the same as the one of the host image, it should be proportionally scaled to the host image.
(3) Modify the DWT coefficients of the host image by using the following equation

$$Y_{i,j} = \alpha_{\lambda,\theta} \times X_{i,j} + (1 - NVF_{i,j}) \times S_{i,j} \times \beta_{\lambda,\theta} \tag{6}$$

Note: (i,j) indicates the spatial location. X and S are the decomposed wavelet coefficients of the host image

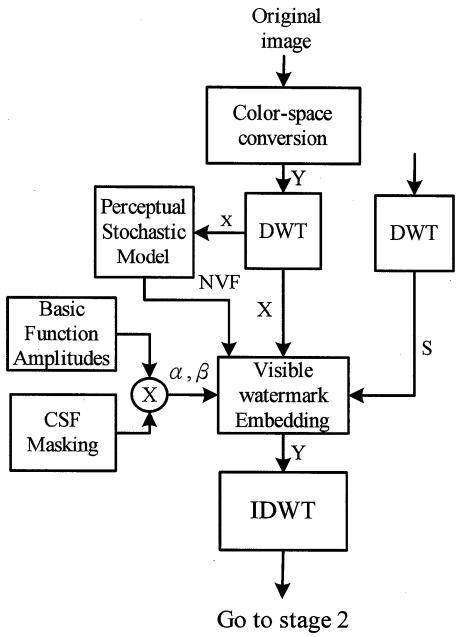


Fig. 4 The flow chart of the proposed visible watermarking approach.

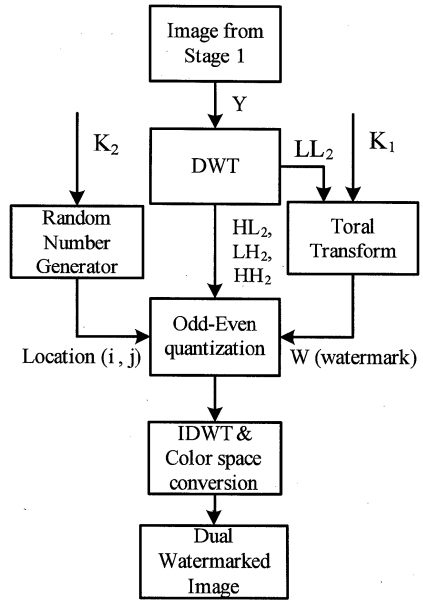


Fig. 5 The flow chart of the proposed visible watermarking approach.

and the watermark image. *NVF* is defined in formula (4) and the relationship of $\alpha_{\lambda,\theta}$ and $\beta_{\lambda,\theta}$ is defined in (3) where $\alpha_{\lambda,\theta} = 1 - \beta_{\lambda,\theta}$.

(4) Inverse transform the DWT coefficients of the host image to obtain a watermarked image (Y component).

3.5 Semi-Fragile Watermark Generation and Embedding Algorithm

It is difficult to develop a visible watermarking algorithm that can avoid any attack by expensive human labors, especially while the texture content of the host image is not complicated. In order to detect such kind of tampering and verify the integrity of the visible watermarked images, we modified the image authentication (semi-fragile watermark) algorithm from [19] into the proposed visible watermarked image as a dual watermarking scheme for our complete architecture.

The semi-fragile watermark embedding procedures are as following and the flow chart of semi-fragile watermarking is shown in Fig. 5.

(1) Select parameters: K_1 and K_2 are the private keys of the scheme. q_1 and q_2 are the quantization parameter.
(2) Select the Y (Luminance) component from visible watermarked image and compute the 2-level 2-D wavelet coefficients of it, $r \times c$ is the size of LL_2.
(3) In order to get high security watermark (pseudo random number), we refer to [22]'s chaotic system called toral automorphisms as chaotic map. Map $Q_{num} = \lfloor LL_2/q_1 \rfloor$ and K_1 as controlling parameter. Using Eqs. (7), (8), we obtain the binary watermark $W(i, j) \in \{0, 1\}$, $1 \le i \le r$, $1 \le j \le c$.

$$A_r(k_1) : \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ Q_{num} + k_1 & Q_{num} + k_1 + 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} (\text{mod } r) \quad (7)$$

$$W(i, j) = (x' + y') \bmod 2 \quad (8)$$

(4) We use K_2 as random seed to create two-dimensional pseudo-random array $location(i, j) \in \{1, 2, 3\}$, $1 \le i \le r$, $1 \le j \le c$ to determine the watermark embedding location corresponding to $\{LH_2, HL_2, HH_2\}$.
(5) The binary watermark is embedded into the visible watermarked image by using simple odd-even quantization. We define odd-even quantization function in formula (9), (10), (11), (12), (13) as the function f in Step (6). The formula performs quantization on $X(i, j)$ into odd-even region according the binary watermark W. q_2 is the quantization parameter. (i, j) indicates the spatial location. X is the decomposed wavelet coefficients of the visible watermarked image.

$$y(i, j) = f(x(i, j), W, q_2) \quad (9)$$

$$x \in R \quad W \in \{0, 1\} \quad q_2 \in Z^+$$

$$I = \begin{cases} 0 & \lfloor x(i, j)/q_2 \rfloor \text{ is even} \\ 1 & \lfloor x(i, j)/q_2 \rfloor \text{ i odd} \end{cases} \quad (10)$$

Note: $\lfloor \cdot \rfloor$denotes the floor function.

$$y(i, j) = \begin{cases} \lfloor x(i, j)/q_2 \rfloor \times q_2 + q_2/2 + x_r & \text{if } I = b \\ y' + x_r & \text{if } I \ne b \end{cases} \quad (11)$$

$$y' = \begin{cases} \lfloor x(i, j)/q_2 - 1 \rfloor \times q_2 + q_2/2 \text{ if} \\ \quad x \in [\lfloor x(i, j)/q_2 \rfloor \times q_2, \lfloor x(i, j)/q_2 \rfloor \\ \quad \times q_2 + q_2/2] \\ \lfloor x(i, j)/q_2 + 1 \rfloor \times q_2 + q_2/2 \text{ if} \\ \quad x \in [\lfloor x(i, j)/q_2 \rfloor \times q_2 + q_2/2, \\ \quad \lfloor x(i, j)/q_2 \rfloor \times q_2 + q_2] \end{cases} \quad (12)$$

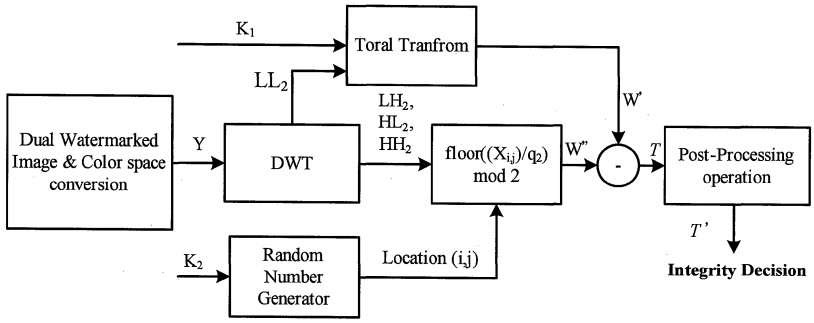


Fig. 6 The flow chart of authentication and tamper detection algorithm approach.

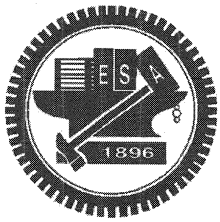


Fig. 7 NCTU logo.

$$x_r = \text{sgn}(x(i, j))(|x(i, j)| \bmod 2) \tag{13}$$

(6) Perform quantization on wavelet coefficients as follows
FOR i =1 to r
FOR j =1 to c
SWITCH location(i, j)
CASE 1: $HL_2(i, j) = f(HL_2(i, j), W(i, j), q_2)$
CASE 2: $LH_2(i, j) = f(LH_2(i, j), W(i, j), q_2)$
CASE 3: $HH_2(i, j) = f(HH_2(i, j), W(i, j), q_2)$

(7) Inverse transform the DWT coefficients of the Y component. The Y component with visible and semi-fragile watermark is converted in the color space domain from YCrCB to RGB.

3.6 Semi-Fragile Watermark Authentication and Tamper Detection Algorithm

Figure 6 shows the flow chart of watermark detection scheme, which is similar to the part of semi-fragile watermark embedding. The tamper detection procedure as follows:

(1) Select parameters: K_1 and K_2 are the private keys of the scheme. q_1 and q_2 are the quantization parameter. The value of K_1, K_2, q_1 and q_2 are the same in embedding and extraction processes.

(2) The obtained visible watermarked image is converted in the color space domain from RGB to YCrCb.

(3) Select the Y (Luminance) component and compute the 2-level 2-D wavelet coefficients of it, $r \times c$ is the size of LL_2.

(4) Use K_1 and K_2 to create two-dimensional pseudo-random arrays; $W'(i, j) \in \{0, 1\}$, $1 \le i \le r$, $1 \le j \le c$ and $location(i, j) \in \{1, 2, 3\}$, $1 \le i \le r$, $1 \le j \le c$.

(5) According to the $location(i, j)$, we find the sub-band and the quantized coefficient, defined as $u(i, j)$. The extract watermark may be obtained by the following formula (14):

$$W''(i, j) = (\lfloor (u(i, j)/q_2) \rfloor) \bmod 2 \tag{14}$$

(6) Having obtained two watermarks W' and W'', we define the tamper detection matrix as formula (15), If $W' = W''$, then T=0. It means the visible watermarked image was not tampered. Otherwise, the '1' element in the tamper detection matrix indicates the pixels that were tampered.

$$T = |W' - W''| \tag{15}$$

(7) Since the algorithm is designed to be semi-fragile watermarking scheme which would want to be robust to mild modifications in all cases, it is inevitable that we can not detect all malicious attack in pixel-wise. However, for practical cases such as removal visible watermark using neighbor pixels and image cropping which crops objects from a source and pastes them onto a target, the malicious attacks always be applied in a certain region in the watermarked image. That is to say, we assume tamper pixels are always continues. Therefore, for a certain tamper detection matrix element $T(i, j)$, if the number of tampered neighboring element for $T(i, j)$ is greater than a given threshold, we regard $T(i, j)$ as a tampered one. The summary of such post-processing operation of tamper detection matrix is shown as following formula (16):

$$T' = \begin{cases} 1, & \sum_{k=-L}^{L} \sum_{l=-L}^{L} T(i+k, j+l) > \beta \\ 0, & \sum_{k=-L}^{L} \sum_{l=-L}^{L} T(i+k, j+l) \le \beta \end{cases} \tag{16}$$

Note: L as width of window, β as threshold.

(8) According to the DWT decomposition of the watermarked image, the size of tamper detection matrix is $r \times c$, which is about 1/16 of the watermarked image. Thus one element in the matrix indicates a corresponding 4×4 block in the watermarking image. Finally, we rescale the tamper detection matrix to have the same size of the watermarked image and obtain the tamper detection image.

4. Experiments and Discussion

The proposed multipurpose watermarking algorithm using dual watermarks has been implemented and intensively tested by using the commonly available color images from USC image database [23]. Because the evaluation standards

Table 2 PSNR summary of watermarked color images.

Image	Method of [3] 【T1=1, T2=350】	Proposed Approach	
		Visible Watermark only	Dual watermarks
Lena	27.0 dB	31.5 dB	31.2 dB
lake	26.2 dB	30.7 dB	30.5 dB
Peppers	26.9 dB	31.4 dB	31.1 dB

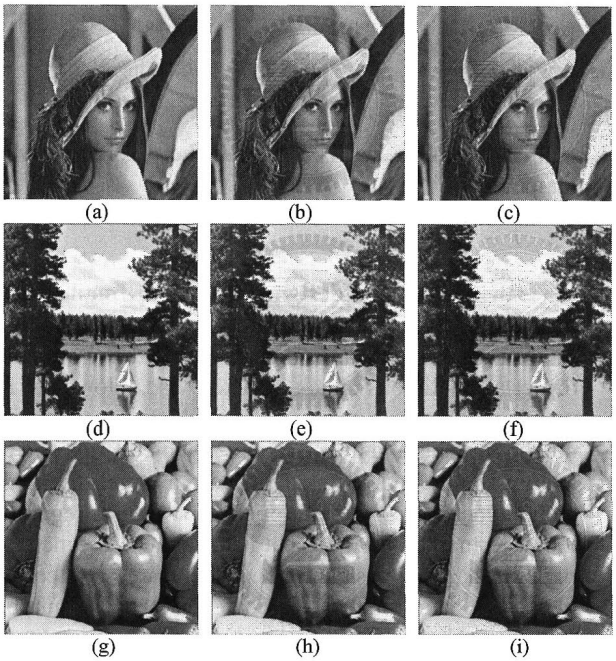


Fig. 8 Original images of (a) Lena (d) Lake (g) Peppers. (b), (e), (h) are watermarked images by the method of [3]. (c), (f), (i) are watermarked images by the proposed algorithm.

for visible watermarking system are absent, we would compare our algorithm with previously proposed ones. To make the fair comparison with other visible watermarking considering HVS, the simulation of [3] is highly addressed here.

Since the CSF based visible watermark technique from [3] has shown better performance than the methods from [11] and AiS Watermark Pictures Protector [24], we compared the results by [3] with the proposed approach and the performance of 512×512 colors images. In the Huang and Tang's method [3], they didn't describe the value of two thresholds used to classify the blocks of each subband, so we choose T1 = 1 and T2 = 350 to match the PSNR values in [3] during the comparison.

One grayscale watermark of logo image is embedded for illustration in Fig. 7 NCTU logo (school logo). The performance of 512×512 experimental images is tabulated in Table 2 for comparison purpose and Fig. 8 shows the original host images of "Lena," "Lake" and "Peppers" respectively. Figures 8(b), (e), (h) are the watermarked images by [3] and Figs. 8(c), (f), (i) are watermarked images by the proposed approach. The performance analysis can be categorized as following.

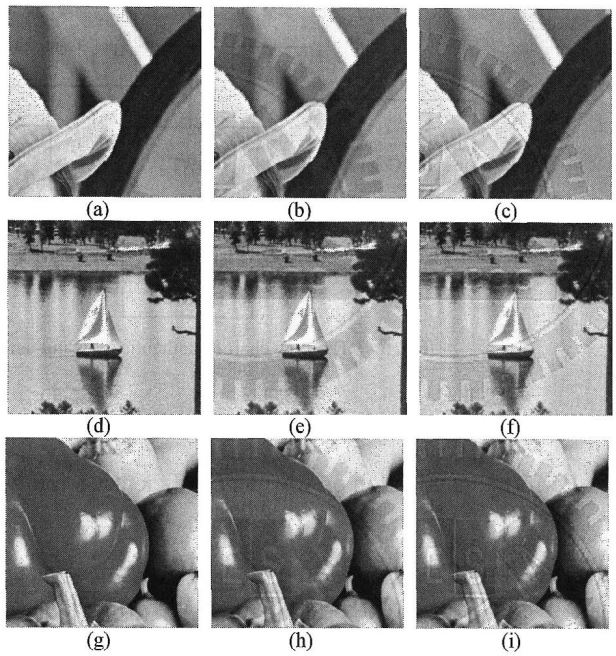


Fig. 9 The visual comparison of close-ups for images from Fig. 8. (a), (d), (e) are close-ups of the original images. (b), (e), (h) are close-ups of the watermarked images by the method of [3]. (c), (f), (i) are close-ups of the watermarked images by the proposed algorithm.

4.1 Visual Quality

From Figs. 8(b) (c), (e) (f) and (h) (i) image pairs, the proposed method has the closest luminance and chrominance maintenance compared with the original ones which are shown clearly from the photos even the difference is sometimes identified subjectively. The watermarked images by using [3] have more bright effect in the unmarked areas. On the other hand, translucence effect is one of requirements for an effective visible watermarking algorithm. The results from our proposed method have better translucence effect than the method of [3] to make photos look more natural since the watermarked images by using [3] affect the details of the host (original) image more, especially in Figs. 8(b), (e), (h).

To further compare the details from the watermarked images, Fig. 9 demonstrates some of close-ups for comparison. Figures 9(a), (d), (g) are the close-ups from original images in Fig. 8. Figures 9(b), (e), (h) are the close-ups from the watermarked images by using [3]'s method. Figures 9(c), (f), (i) are the close-ups from the watermarked images by using the proposed technique. It is very clear that the watermark's edges and thin lines are blurred in those images by using the method of [3]. In addition, the watermark patterns in our proposed method still have sharp edge and the logo watermark is evidently embedded. For the text pattern, the text of character *A* in our results is with sharper edge than the ones in Figs. 9(b), (h). Besides, the outlines in our results are clearer than those from method of [3].

Table 3 PSNR summary of watermarked color images before and after JPEG 2000 compression.

Image	PSNR value (dB)					
	Method of [3]			Proposed method		
	Before	After	After(wn)	Before	After	After(wn)
Lena	27.0	26.0	34.5	31.2	29.2	33.8
lake	26.2	24.2	30.2	30.5	26.9	30.1
Peppers	26.9	22.7	27.4	31.1	24.8	26.8

4.2 PSNR (Peak Signal-to-Noise Ratios)

Since we need to make the complete comparison with the data in [3] where the watermarked images are generally with PSNR(peak signal-to-noise ratio) values below 30 dB for many common used 512 × 512 colour images, PSNR values comparison for the proposed approach are also listed for comparison purpose. Even PSNR values don't truly represent the subjective visual quality for visible watermark images, the values do have the positive correlation with the image fidelity. Higher PSNRs do have better image similarity with the host images. The definition of PSNR is as following:

$$PNSR\,(\text{dB}) = 10\log_{10}(255^2/MSE) \tag{17}$$

where MSE is the mean square error of the watermarked image and the original image per pixel for the Y component of color image.

To make a fair comparison with the method from [3], it is better to embed the same watermark for the same cover image. However, the watermark used in [3] is not available. We then embed a logo watermark from Fig. 7 to make the best effort for performance comparison. The tabulated results from Table 2 disclose that our watermarking scheme can achieve higher PSNR values than the method in [3]. This denotes the fidelity of images from our method is better than those by method of [3]. In addition, the PSNR values of dual watermarked images are only 0.2–0.3 dB less than those of visible watermark only images. This means that our proposed multipurpose design could achieve as good as high image quality of visible watermarking but also with extra function of invisible watermarks.

4.3 JPEG 2000 Compression

The robustness of the proposed visible watermark technique should be tested for comparison. For JPEG 2000 compression, software from [25] is adopted as the compression tool. The PSNR values before and after the jpeg 2000 compression are tabulated in Table 3. The compression ratio is 100:3 between the uncompressed image and compressed image. There are two columns of PSNR values for both methods labeled "after." The pure "after" column means those PSNR values are compared between the compressed watermarked image and the original image. The after(wn) column means those PSNR values are compared between the compressed watermarked image and the watermarked image. From Table 3, we can find that the PSNR values are almost the same for both methods while the compressed watermarked images are compared with the watermarked images (after(wn) column). However, the PSNR values are higher while the compressed watermarked images are compared with the original images by the proposed approach than by the method of [3] (after column). Therefore, this statistic indicates that the image quality of watermarked image before and after compressed is higher by the proposed approach than the method of [3]. To further investigate the effect of compression, the visual difference can be illustrated by the close-up comparison. Figure 10 shows the close-ups of original images. Figures 10(b), (e), (h) are the close-ups of watermarked images by the method of [3]. Figures 10(c), (f), (i) are the close-ups of watermarked images by our proposed method. By comparing Figs. 10(b), (c), (e), (f), (h), (i), the compressed images maintain the details of the logo pattern but the characters E, S, A of watermarked images by our proposed method are more apparent than those of watermarked images by the method of [3]. This observation is consistent with the claim of our discussion in Sect. 2 that the embedding factors in [3] emphasize more weights in the low frequency domain instead of the medium-to-high frequency domain while the high frequency components of the watermarked image easily suffer common image signal processing attacks like compression. Therefore, we can conclude that our proposed method is more robust than Huang and Tang's method by JPEG 2000 compression attack from above observation where the visibility of watermark is surely higher by the proposed approach.

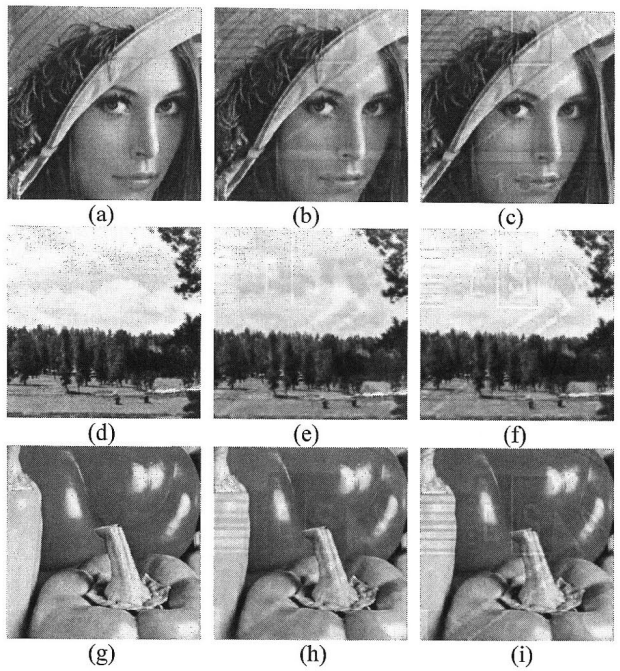


Fig. 10 The visual quality comparison of close-ups of watermarked image after jpeg 2000 compression ratio of 100:3. (a), (d), (e) are close-ups of the original images. (b), (e), (h) are close-ups of the watermarked images by the method of [3]. (c), (f), (i) are close-ups of the watermarked images by the proposed algorithm.

Table 4 PSNR summary of watermarked color images before and after Median Filter.

Image	PSNR value (dB)					
	Method of [3]			Proposed method		
	Before	After	After(wn)	Before	After	After(wn)
Lena	27.0	21.2	24.7	31.2	23.1	24.4
lake	26.2	19.3	21.8	30.5	20.7	21.9
Peppers	26.9	18.4	20.8	31.1	19.8	20.6

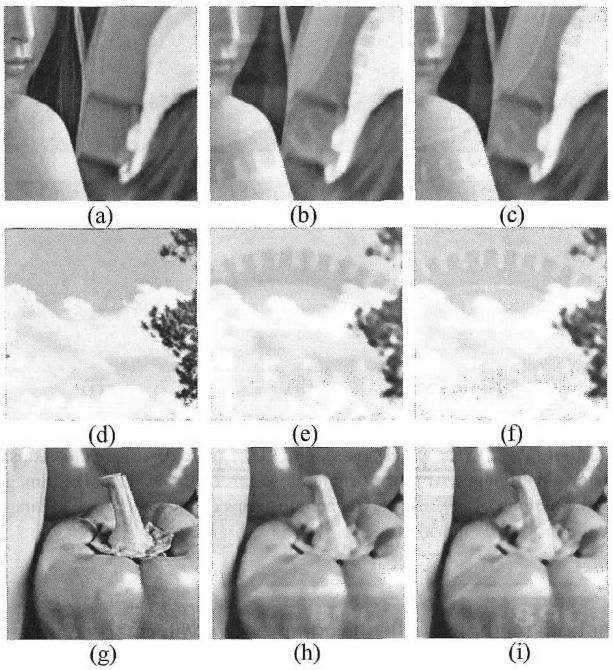


Fig. 11 The visual quality comparison of close-ups of 7×7 median filtering of watermarked image (a) original image (b) watermarked images the Huang and Tang's method (c) watermarked image by the proposed algorithm.

4.4 Median Filtering

The robustness of Median filtering attack is also tested here and StirMark [26] software is adopted here for this attack. Since the results of 3×3 and 5×5 median filtering are similar to the illustration as shown in Fig. 8, a stronger attack as 7×7 median filtering is applied here for the comparison. The PSNR values before and after the median filtering are tabulated in Table 4. There are two columns of PSNR values for both methods labeled "after" and their meaning is the same as mentioned in the section of JPEG 2000 compression. From Table 4, we can find that the PSNR values are almost the same for both methods while the filtered watermarked images are compared with the watermarked images (after(wn) column). However, the PSNR values are higher while the filtered watermarked images are compared with the original images by the proposed approach than by the method of [3] (after column). Therefore, this statistic indicates that the image quality of watermarked image before and after filtering is higher by the proposed approach than the method of [3]. To further investigate the effect of median filtering, the visual difference can be illustrated by the close-up comparison. Figures 11(a), (d), (g) are close-ups of original images. Figures 11(b), (e), (h) are close-ups of 7×7 median filtering of watermarked image by the the method of [3]. Figures 11(c), (f), (i) are close-ups of 7×7 median filtering of watermarked image by the proposed method. By comparing Figs. 11(b), (c), (e), (f), (h), (i), the median filtered images became blurry but Figs. 11(c), (f), (i) have sharper contour than Figs. 11(b), (e), (h). It is apparent that the logo pattern (i.e. the characters of E, S, A, or the characters of 1896) is still evidently existed in Fig. 11(c) but is blurred and hard to be recognized in Fig. 11(b). Therefore, the proposed technique outperforms [3] by the median filtering attack from above observation where the visibility of watermark is surely higher by the proposed approach.

Table 5 PSNR summary of watermarked color images before and after Gaussian Filter.

Image	PSNR value (dB)					
	Method of [3]			Proposed method		
	Before	After	After(wn)	Before	After	After(wn)
Lena	27.0	26.6	33.5	31.2	29.9	33.6
lake	26.2	24.8	30.1	30.5	27.2	30.6
Peppers	26.9	26.3	32.6	31.1	29.4	32.9

Table 6 PSNR summary of watermarked color images before and after Average Filter.

Image	PSNR value (dB)					
	Method of [3]			Proposed method		
	Before	After	After(wn)	Before	After	After(wn)
Lena	27.0	25.8	32.7	31.2	28.8	32.5
lake	26.2	24.1	30.0	30.5	26.6	30.2
Peppers	26.9	22.7	27.4	31.1	24.9	26.9

Other attacks from [26] like Gaussian filtering and average filtering are also preformed and the experimental results shown in Tables 5 and 6 are consistent with the above findings which indicate our visible watermarking scheme has better visual effect and higher PSNR values than other schemes like [3].

4.5 ICA (Independent Component Analysis) Image Recovery Attack

Regarding the removal technique, the image recovery method [27] can remove visible watermarking patterns consisting of thin lines and a few human interventions of image-inpainting approach of [28] can deal with patterns of thick lines. However, the iterative process of image-inpainting is costly and time-consuming. The image inpainting attack for the proposed visible watermarking is performed for Lena image only and Fig. 12 shows the watermarked image in Fig. 12(a), the mask used during the image inpainting attack in Fig. 12(b) and the recovered image in Fig. 12(c). From Fig. 12, we have found that the image inpainting based watermark removal approach could not remove the watermark completely since the watermark outline are still existing and the detailed image content can not be fully reconstructed.

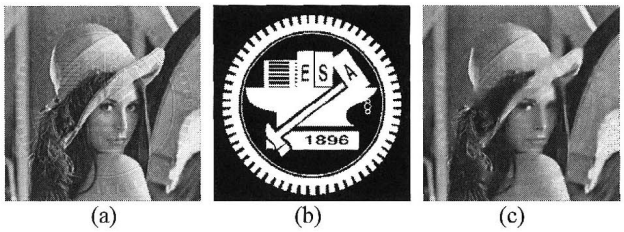


Fig. 12 Recovering the watermarked images by image inpainting approach (a) is the watermarked images with NCTU logo by the proposed approach. (b) is the mask used during the watermark removal (c) is the recovered image after the image inpainting attack.

Therefore, a better image recovery attack is needed for further comparison.

Pei and Zeng [29] proposed another image recovery algorithm for removing visible watermarks which is simple, fast with less human intervention. The method mainly utilized independent component analysis (ICA) to separate host images from watermarked and reference images. The algorithm included three phases: watermarked area segmentation, reference image generation, and image recovery. In their experiments, five different visible watermarking methods [3]–[7] and three public domain images are tested. The experimental results showed that their algorithm can successfully removed the visible watermarks, and the algorithm itself is independent of both the adopted ICA approach and the visible watermarking method. Interested readers can refer [29] for detailed information.

In this study, we have implemented the method of [29] and tested several public images used in [29] for comparison. By applying the method of [29] to our proposed visible watermarking approach, Fig. 13 illustrates the results of the watermark removal attack where the logo patterns slightly disappear but still exist and the contours are recognizable in Figs. 13(b), (d), (f). Besides, the watermark removal scheme in [29] can remove the watermark by the method in [3]–[7] but the proposed approach can resist such attack. We can conclude that the proposed visible scheme certainly outperforms the methods in [3]–[7].

The robustness of a visible watermarking scheme mainly depends on how difficult or impossible it is to remove the embedded visible watermark patterns without exhaustive and costly user interventions. Even the results in Figs. 12 and 13 show that the proposed visible watermarking approach can resist the watermark removal attacks, we fully understand subsequent signal processing operations like equalization, sharpening or background blurring could completely remove the watermark eventually. Even the visible watermark will be removed completely, our dual watermark can still disclose the information about where the attacks occur and the tamper detection simulations are performed next.

4.6 Tamper Detection

To evaluate the validity of the proposed image authentication algorithm and make up tampered images, we use Adobe Photoshop CS2 for implementation of image processing operations. In our experiments, we set parameters $q_1 = 30$, $q_2 = 10$, $K_1 = 1234$, $K_2 = 1234$, $L = 1$, $\beta = 3$. Figure 14 and Fig. 15 demonstrate the result images (visible and semi-fragile watermarks embedded), tampered image, and tampering detection image respectively. In Fig. 14(b), an object (a company logo of blue character A with copyright symbol ®) is inserted into the right bottom of the dual watermarked Lena image. In the shoulder part of the watermarked Lena image, we use neighboring pixels to remove the visible watermark. In Fig. 15(b), three objects (the same company logo of Fig. 14(b) and two peppers) are inserted into the watermarked Peppers image. From the detection results of Figs. 14(c) and 15(c), the marked points indicate the tampered parts of watermarked image in Figs. 14(b) and 15(b) where these parts are located correctly.

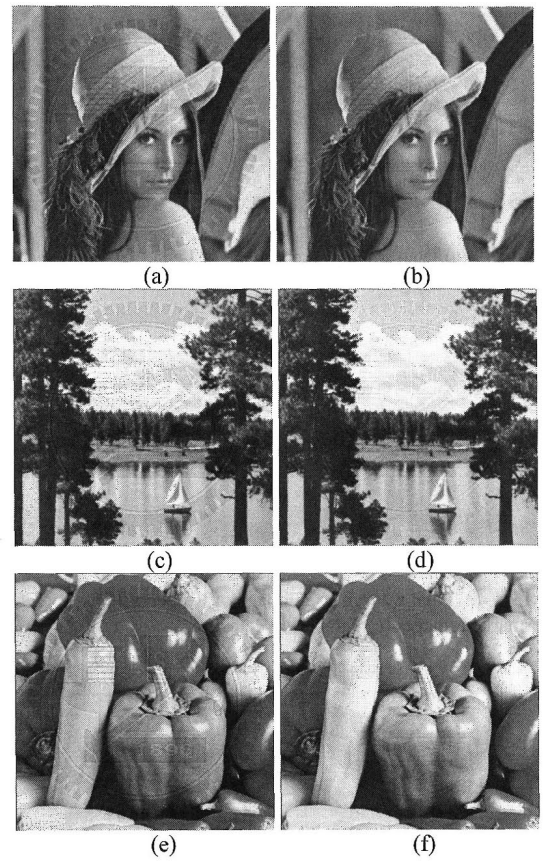


Fig. 13 Recovering the watermarked images from the proposed method (a), (c), (e) are the watermarked images with NCTU logo. (b), (d), (f) are the recovered image of (a), (c), (e) respectively under image recovery attack.

For the combination of tampering operation and mild modification, Fig. 16 and Fig. 17 show the tamper detection after AWGN with different σ^2 and JPEG compression with different quality factor (QF) setting. Figure 16(a) shows the watermarked Lake image and Fig. 16(b) shows the tampered Lake image: one object (a boat) is inserted into the watermarked Lake image. From Figs. 16(c), (d), (e), we can see

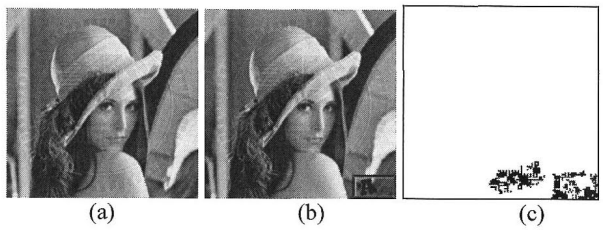


Fig. 14 (a) Dual watermarked image of Lena. (b) Tampered dual watermarked image. (c) Tampering detection.

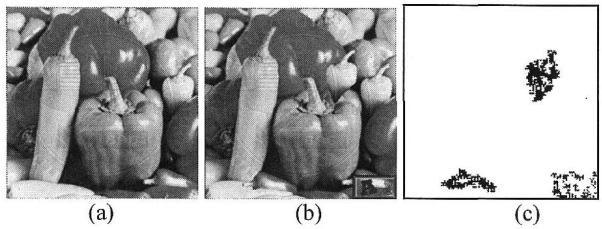


Fig. 15 (a) Dual watermarked image of Peppers. (b) Tampered dual watermarked image. (c) Tampering detection.

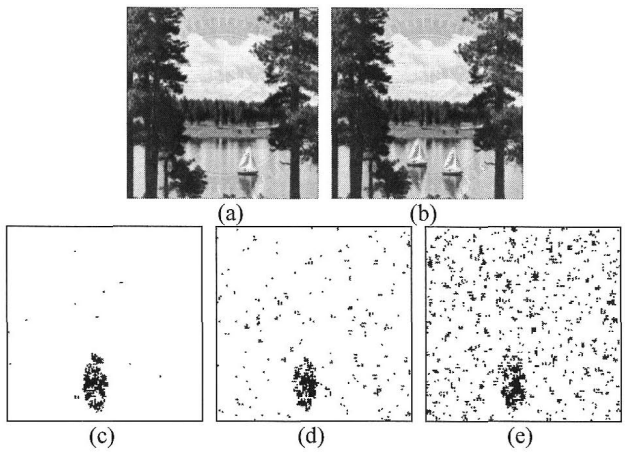


Fig. 16 Tamper detection for the combined tampering operations and AWGN with different σ^2. (a) Watermarked image of Lake (b) tampered image of Lake (c) $\sigma^2 = 6$ (d) $\sigma^2 = 12$ (e) $\sigma^2 = 18$.

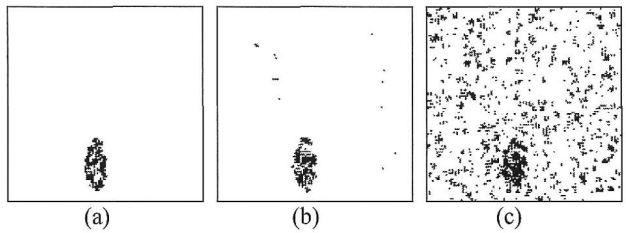


Fig. 17 Tamper detection for the combined tampering operations and JPEG compression. (a) QF = 100 (b) QF = 80 (c) QF = 60.

the detection result of tampered Lake image is located correctly after AWGN with different σ^2. From Figs. 17(a), (b), (c), we can see the detection result of tampered Lake image is located correctly after JPEG compression with different quality factor (QF) settings.

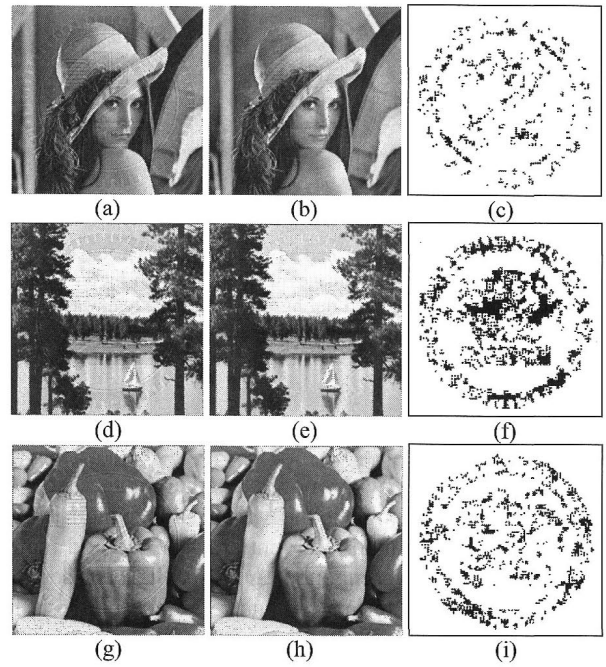


Fig. 18 Dual watermarked images of (a) Lena (d) Lake (g) Peppers respectively. (b), (e), (h) are tampered dual watermarked images with watermark removal attack. (c), (f), (i) are images after tampering detection.

For more serious attacks like watermark removal, we are also interested in the detection capability by the proposed approach. Figure 18 demonstrates the temper detection result. We can clearly see the tampered areas are labeled in Figs. 18(c), (f), (i) and reflected the evidence of tempering.

After the intensive performance comparison, the results of different attacks, visual quality analyses and temper detection demonstrate that the proposed multipurpose color image watermarking by using dual watermarks with HVS method is more robust with better image quality. In summary, we are convinced that the proposed complete architecture is a superior scheme among the referred published techniques.

5. Conclusion

A novel watermarking-based technique for copyright protection and authentication has been presented in this study. In copyright protection, we propose a new visible watermarking technique where the intensity of the watermark in different regions of the image depends on the underlying content of the image and human sensitivity to spatial frequencies. The collaboration of CSF and NVF for HVS models is leveraged with the noise reduction of the visibility thresholds for HVS in DWT domain. The perceptual weights is fine tuned for watermark embedding which results significant improvement over the watermarked images by CSF only based algorithms regarding the image quality, translucence and robustness of the watermarking. For authentication and verification of the integrity for the

dual watermarked images, we applied a semi-fragile watermark algorithm which can detect and localize malicious attack effectively yet tolerate mild modifications such as JPEG compression and channel additive white Gaussian noise (AWGN). In addition, the experimental results demonstrate the proposed visible watermarking scheme has achieved high PSNR values with better visual fidelity and robustness to attacks than other schemes and the semi-fragile watermarking scheme has the capability to verify the integrity of the images.

Acknowledgments

This work was supported by the National Science Council in Taiwan, Republic of China, under NSC95-2416-H009-027 and NSC96-2416-H009-015.

References

[1] World Intellectual Property Organization (WIPO), http://www.wipo.int/

[2] I.J. Cox, J. Kilian, J., F.T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Process., vol.6, no.12, pp.1673–1687, Dec. 1997.

[3] B.B. Huang and S.X. Tang, "A contrast-sensitive visible watermarking scheme," IEEE Multimed., vol.13, no.2, pp.60–66, April-June 2006.

[4] G.W. Braudaway, K.A. Magerlein, and F.C. Mintzer, "Protecting publicly-available images with a visible image watermark," Proc. Conf. Optical Security and Counterfeit Deterrence Techniques, SPIE, pp.126–132, 1996.

[5] J. Meng and S.-F. Chang, "Embedding visible video watermarks in the compressed domain," Proc. ICIP, vol.1, pp.474–477, Oct. 1998.

[6] S.P. Mohanty, K.R. Ramakrishnan, and M.S. Kankanhalli, "A dual watermarking technique for image," Proc. 7th ACM Int. Multimedia Conf., vol.2, pp.49–51, Oct./Nov. 1999.

[7] P.-M. Chen, "A visible watermarking mechanism using a statistic approach," Proc. 5th Int. Conf. Signal Processing, vol.2, pp.910–913, 2000.

[8] Y. Hu and S. Kwong, "Wavelet domain adaptive visiblewatermarking," Electron. Lett., vol.37, no.20, pp.1219–1220, Sept. 2001.

[9] Y. Hu and S. Kwong, "An image fusion-based visible watermarking algorithm," Proc. 2003 Int'l Symp. Circuits and Systems, pp.25–28, IEEE Press, 2003.

[10] L. Yong, L.Z. Cheng, Y. Wu, and Z.H. Xu, "Translucent digital watermark based on wavelets and error-correct code," Chinese J. of Computers, vol.27, no.11, pp.1533–1539, Nov. 2004.

[11] A.P. Beegan, L.R. Iyer, and A.E. Bell, "Design and evaluation of perceptual masks for wavelet image compression," Proc. 10th IEEE Digital Signal Processing Workshop, pp.88–93, IEEE CS Press, 2002.

[12] D. Levický and P. Foriš, "Human visual system models in digital image watermarking," Radioengineering, vol.13, no.4, pp.38–43, 2004.

[13] A.B. Watson, G.Y. Yang, J.A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," IEEE Trans. Image Process., vol.6, no.8, pp.1164–1175, 1997.

[14] Ö. Ekici, B. Sankur, B. Coşkun, U. Nazi, and M. Akcay, "Comparative evaluation of semifragile watermarking algorithms," J. Electronic Imaging, vol.13, no.1, pp.209–216, 2004.

[15] C. Fei, D. Kundur, and R. Kwong, "Analysis and design of secure watermark-based authentication systems," IEEE Trans. Information Forensics and Security, vol.1, no.1, pp.43–55, March 2006.

[16] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper proofing and authentication," Proc. IEEE, vol.87, no.7, pp.1167–1180, July 1999.

[17] H.P. Alexandre and K.W. Rabab, "Wavelet-based digital watermarking for image," IEEE Canadian Conference on Electrical and Computer Engineering, vol.I, pp.879–884, 2002.

[18] H. Yuan and X.P. Zhang, "A multiscale fragile watermarking based on the Gaussian mixture model in the wavelet domain," Proc. 2004 Int. Conf. on Acoustics, Speech and Signal Processing, vol.3, pp.413–416, Montreal, QC, Canada, May 2004.

[19] K. Ding, C. He, L.G. Jiang, and H.X. Wang, "Wavelet-based semi-fragile watermarking with tamper detection," IEICE Trans. Fundamentals, vol.E88-A, no.3, pp.787–790, March 2005.

[20] J.L. Mannos and D.J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," IEEE Trans. Inf. Theory, vol.20, no.4, pp.525–536, July 1974.

[21] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," Proc. 3rd Int. Workshop Information Hiding, pp.211–236, Dresden, Germany, Sept. 1999.

[22] M.J. Tsai, K.Y. Yu, and Y.Z. Chen, "Joint wavelet and spatial transformation for digital watermarking," IEEE Trans. Consum. Electron., vol.46, no.1, pp.241–245, Feb. 2000.

[23] USC SIPI–The USC-SIPI Image Database [Online]:http://sipi.usc.edu/services/database/Database.html

[24] AiS Watermark Pictures Protector: http://www.watermarker.com

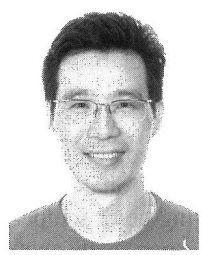
[25] JPEG 2000 compression, [Online]: http://www.ece.uvic.ca/mdadams/hasper/~

[26] StirMark, [Online]: http://www.petitcolas.net/fabien/software/StirMarkBench mark_4_0_129.zip

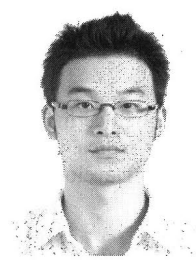
[27] M. Bertalmio, V. Caselles, and C. Ballester, "Image inpainting," SIGGRAPH 2000, pp.417–424, Aug. 2000.

[28] C.-H. Huang and J.-L. Wu, "Attacking visible watermarking," IEEE Trans. Multimed., vol.6, no.1, pp.16–30, Feb. 2004.

[29] S.C. Pei and Y.C. Zeng, "A novel image recovery algorithm for visible watermarked image," IEEE Trans. Information Forensics and Security, vol.1, no.4, pp.543–550, Dec. 2006.



Min-Jen Tsai received the B.S. degree in Electrical Engineering from National Taiwan University in 1987, the M.S. degree in Industrial Engineering and Operations Research from University of California at Berkeley in 1991, the Engineer and Ph.D. degrees in Electrical Engineering from University of California at Los Angeles in 1993 and 1996, respectively. From 1996 to 1997, he was a senior researcher at America Onlinc Inc. In 1997, he joined the Institute of Information Management at the National Chiao Tung University in Taiwan and is currently an associate professor. His research interests include multimedia system and applications, digital forensic, digital watermarking and authentication, web services, enterprise computing for electronic commerce. Dr. Tsai is a member of IEEE, ACM, and Eta Kappa Nu.



Chih-Wen Lin has received B.S. degree in Industrial Education from National Taiwan Normal University in 2004, the M.S. degree in Institute of Information Management at the National Chiao Tung University in the year 2007. He currently serves in the Telecommunication Laboratory of Chunghwa Telecom Co. Ltd.