Blind dual watermarking for color images' authentication and copyright protection

Xiao-Long Liu, Chia-Chen Lin*, *Member, IEEE,* and Shyan-Ming Yuan

***Abstract*—This paper presents a blind dual watermarking mechanism for digital color images in which invisible robust watermarks are embedded for copyright protection and fragile watermarks are embedded for image authentication. For the purpose of copyright protection, the first watermark is embedded by using the discrete wavelet transform (DWT) in YCbCr color space, and it can be extracted blindly without access to the host image. However, fragile watermarking is based on an improved least significant bits (LSB) replacement approach in RGB components for image authentication. The authenticity and integrity of a suspicious image can be verified blindly without the host image and the original watermark. The combination of robust and fragile watermarking makes the proposed mechanism suitable for protecting valuable original images. The experimental results indicated that the proposed watermarking mechanism can withstand various processing attacks and accurately locate the tampered area of an image.**

***Index Terms*—Authentication, copyright protection, discrete wavelet transform (DWT), LSB, watermarking**

I. Introduction

WITH the rapid progress of an information-oriented society, increasingly large quantities of digitalized material are being transmitted over the Internet. Concerns pertaining to the enhancement of security and protection against violations of digital images have become critical over the past decade. Digital watermarking [1-6] is now a relatively focused technique aimed at providing a reliable way to authenticate images or protect copyrights protection; in this technique a watermark usually is embedded invisibly in the digital image to avoid attracting the attention of malicious attackers. In accordance with the desired robustness of the embedded watermark, digital watermarking techniques are divided into robust watermarking [4-6] and fragile watermarking [1-3]. The main purpose of robust watermarking technique often is to protect the ownership of host images, while the fragile watermarking technique is used to authenticate the integrity of images.

Robust watermarking is typically used for copyright protection, thus it is designed to resist attacks that attempt to remove or destroy the watermark without significantly degrading the visual quality of the watermarked image. In robust watermarking, verifiable watermarks of users, such as logos or copyright information, are embedded into the host images. Later, the verifiers can extract the watermarks and confirm ownership through the watermarked images. Robustness is one of the major points of concern, which means the extracted watermark must be robust enough for the ownership of the host image to be verified even after the watermarked image has been subjected to signal processing attacks. However, for the sake of increasing the robustness of a watermarked image, previous robust watermarking techniques often alter significant areas of the host image, which can cause serious distortion of the quality of the watermarked image. This distortion may help malicious attackers identify which data are valuable, allowing them to perform cryptanalysis and acquire the confidential data. Therefore, it is still a salient issue in the watermarking field to develop a robust watermarking scheme that can deliver outstanding robustness while maintaining good visual quality of watermarked images. On the other hand, fragile watermarking was developed particularly for image authentication, in which the embedded watermark should be fragile so that any modifications of the images will be apparent. The authenticity of the image should be verified definitively if the watermarked image has been manipulated in any way, such as JPEG compression, collage, or cropping. Since the less significant areas of the host image are altered in fragile watermarking, the visual quality of a fragile watermarked image is usually better than that of the robust watermarked image. The accuracy of the authentication is the major concern in fragile watermarking, and techniques that were developed before 2000 focused mainly on detecting whether an image had been tampered with or not. However, they did not specify clearly where the image had been modified. Over the last 15 years, several image authentication schemes have been developed for the purpose of locating the tampered areas, but capability of doing so is barely satisfactory, and not every modified pixel is guaranteed to be detected correctly. Furthermore, most of the fragile watermarking schemes are non-blind, and original watermark information is required during the verification procedure.

Manuscript received May 26, 2015; revised April 29, 2016; accepted June 28, 2016. This work was supported in part by the Ministry of Science and Technology under Grant 104-2410-H-126 -005.

X.-L. Liu is with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China, and also with the Department of Computer Science, National Chiao Tung University, Hsinchu 300, Taiwan, ROC (e-mail: shallen548@gmail.com).

C.-C. Lin is with the Department of Computer Science and Information Management, Providence University, Taichung 43301, Taiwan, ROC (corresponding author, e-mail: mhlin3@pu.edu.tw).

S.-M. Yuan is with the Department of Computer Science, National Chiao Tung University, Hsinchu 300, Taiwan, ROC (e-mail: smyuan@cs.nctu.edu.tw).

Concerning previous watermarking literature, most publications addressed grey-scale images rather than color images. However, the processing and transmission of color images now has a central role in our information-oriented society, and these activities amount to more than just an extension of the activities associated with grey-scale images. Recently, several watermarking studies [9] have focused on color images rather than grey-scale images. In particular, RGB and YCbCr are two major color models concerned in designing watermarking schemes. RGB, which is composed of three highly-correlated channels, i.e., a red channel, a green channel, and a blue channel, is a natural space for representing real-world color. YCbCr, however, breaks the visual information into one luminance channel, i.e., Y, and two chrominance channels, i.e., Cb and Cr. Previous research [9] indicated that watermarks that are hidden in the YCbCr color space are relatively robust and recovered better after JPEG compression and Gaussian noise attacks than watermarks that are hidden in the RGB color space. Thus, it works well to use the YCbCr color space for copyright protection and to use the RGB color space for image authentication.

To satisfy the essentials of robust and fragile watermarking schemes explained above, in this paper, we present a blind dual watermarking mechanism for color images. An invisible and robust watermark is embedded for copyright protection and an invisible and fragile watermark is embedded for image authentication at the same time in our scheme. The discrete wavelet transform (DWT) in the Y channel of the YCbCr color space is used for robust watermarking. After one level of DWT decomposition, the low-low (LL) sub-band of Y is quantized by the luminance quantization table. The robust watermark is embedded in the high-high (HH) sub-band by expertly replacing it with the result of the LL quantization. In addition, the fragile watermark for image authentication is embedded independently on each RGB color channel according to an improved LSB replacement approach [10]. The copyright and authenticity of watermarked image can be verified blindly without the host image. The simulation results showed that the proposed dual watermarking mechanism can withstand various processing attacks and locate the tampered area of the image accurately. Furthermore, the proposed mechanism provides watermarked image with superb visual quality, which makes it suitable for protecting valuable original images.

The remainder of this paper is organized as follows. We briefly describe the previous relevant work in Section II. Section II contains a detailed exposition of the proposed dual watermarking mechanism. Experimental results and analyses of the robustness and accuracy of our work are provided in Section IV. Also in Section IV, we compare the performance of the proposed dual watermarking mechanism with other existing dual watermarking schemes. Our conclusion are presented in Section V.

II. Related Studies

In this section, we briefly describe the previous literature relevant to fragile watermarking, robust watermarking, and dual watermarking. All of the literature references that are mentioned have been either used extensively in several applications or cited by other research.

Fragile watermarking is used mainly to authenticate the integrity of images; such authentications must be very sensitive to changes in the signal of host image, thus most of them are based on the spatial domain. The state of fragile watermarking allows us to determine whether the image has been tampered with. One of the original fragile watermarking schemes for image authentication was proposed by Yeung and Mintzer [1] in 1997; in their scheme, a secret key was used to generate a binary valued function. This binary function was used to ensure that the watermark extracted from the unaltered watermarked image was equal to the watermark W. Thus, if the watermarked image had been altered or damaged, the extracted watermark would be different with watermark W. The visual quality of the watermarked images provided by Yeung and Mintzer's scheme are satisfactory. However, they focused mainly on detecting whether an image had been tampered with or not, instead of specifing clearly how and where the image was tampered, which means they do not have the capability of identifying the location at which tampering occurred during the image verification procedure. In 2008, Chang et al. [2] applied the concept of the Chinese remainder theorem (CRT) and embedding four authentication bits obtained from CRT into each block instead of one authentication bit in a host image to improve the tamper localization performance. Recently, some scholars also studied the removable fragile watermarking to provide the capability of reconstructing an authentic image [3]. The removable fragile watermarking used in image authentication allows the user to restore the watermarked image to a lossless original image once a watermarked image has been authenticated. Nevertheless, the tampered areas still cannot be restored once a watermarked image is under attack by any kind of manipulation. This characteristic would thwart the hypothetical reconstruction capability when the watermarked image has been tampered with.

The robust watermarking schemes given in the literature can be classified into two categories, i.e., spatial domain schemes and frequency domain schemes, based on the domain into which the watermark is inserted. In spatial domain scheme [4], the watermark is straightforwardly inserted into the host image by altering the pixel values. It has the advantages of low complexity and easy implementation, but, generally, it is not very robust to resistant affine transformations, and to some image processing attacks. However, frequency domain schemes [5] typically make images more difficult to perceive and can provide more robustness against many common attacks. In frequency domain watermarking schemes, there is a tradeoff between robustness and imperceptibility of the host image. Recently, the performance of robust watermarking was improved further by some computational intelligence-based techniques, such as genetic algorithms (GAs) and differential evolution (DE). Maity et al. [6] proposed an optimized spread spectrum (SS) image watermarking scheme using GA and multiband wavelets. Although the computational intelligence-based techniques can handle the automatic balance between imperceptibility and robustness in watermarking, their disadvantage of slow computing speed for intelligence-based watermarking is not acceptable for most watermarking applications. Therefore, an adaptive, robust, frequency-domain watermarking scheme is presented in this paper to balance the robustness and imperceptibility of the watermarked image adaptively, while maintaining the advantage of satisfactory computing speed.

The dual watermarking technique is proposed by embedding different watermarks for multi-purpose image protection. In 2001, Lu and Liao [7] proposed a blind watermarking scheme by embedding robust and fragile watermarks in a host image for simultaneous copyright protection and image authentication. However, the embedded robust watermark is generated from the host image itself instead of a defined logo image. Also, the tamper localization capability is not satisfactory in Lu and Liao's scheme and not all of the tampered area can be located accurately. In 2009, Lin et al. [8] concentrated on presenting a dual watermarking scheme for intensive copyright protection. In their scheme, a visible watermark image is directly appended on the spatial domain of the host image, and an invisible watermark image is embedded in the frequency domain by utilizing the just noticeable distortion [JND] technique. In order to increase the robustness of a watermarking system for copyright protection, Lusson et al. [9] proposed a dual robust watermarking scheme by embedding one image watermark into the RGB color space and the other binary watermark into the YCbCr color space of a host color image. Two watermarks were embedded into different spaces of a host image increase the probability of the image's surviving multiple attacks. However, the original host image is required to extract the watermark in RGB color space, which is a non-blind watermarking scheme. Moreover, both Lin et al.'s scheme and Lusson et al.'s scheme concentrate only on dual watermarking for intensive copyright protection, and authentication of the image is not a concern. It is not secure enough for protecting images once its integrity is infringed, but such infringement cannot be detected.

III. Proposed Scheme

This section explains the proposed watermarking mechanism in detail. In order to satisfy image authentication and copyright protection requirements, a blind invisible dual watermarking mechanism for color images is presented. The details of dual watermark embedding and extraction phases of our proposed mechanism are described in subsections III-A and III-B, respectively.

A. Dual Watermark Embedding

Fig. 1 shows our watermark embedding procedure. The process of embedding the watermark is divided into two phases, i.e. 1) embedding the invisible robust watermark and 2) embedding the invisible fragile watermark.

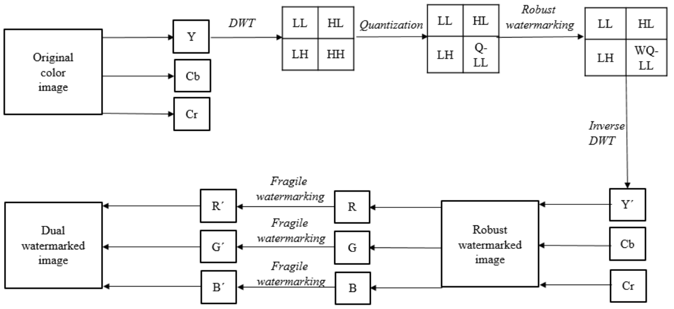


Fig. 1. Our watermark embedding procedure

Phase 1: Embedding the invisible robust watermark

Initially, the original RGB color image is converted into YCbCr color space. The basic equations used to convert RGB into YCbCr are:

$$\begin{aligned} \mathrm{Y} &= 0.299\mathrm{R} + 0.587\mathrm{G} + 0.114\mathrm{B} \\ \mathrm{Cb} &= -0.172\mathrm{R} - 0.339\mathrm{G} + 0.511\mathrm{B} + 128\ . \\ \mathrm{Cr} &= 0.511\mathrm{R} - 0.428\mathrm{G} - 0.083\mathrm{B} + 128 \end{aligned} \quad (1)$$

After YCbCr conversion, the one level DWT decomposition of Y is performed to generate the low-low (LL), low-high (LH), high-low (HL), and high-high (HH) sub-bands, where LL consists of the approximation part of the original Y channel, and the remaining three resolution sub-bands consist of the detailed parts, which are very difficult for the human eye to discern. Thus, we used this characteristic for our robust watermarking scheme.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Fig. 2. Luminance quantization table [11]

First, the LL sub-band is divided into several 8×8-sized blocks, and, then, each pixels of a block is quantized by directly dividing it by the corresponding value of the luminance quantization table which is defined by U. It [11], as shown in Fig. 2. After all of the blocks have been quantized, a quantized LL (Q-LL) sub-band is generated, and the HH sub-band is replaced directly by the resulting Q-LL sub-band. This is because the quantized LL sub-band would hardly be changed while suffering malicious attacks. This characteristic allows the blind and robust decoding during watermark extraction procedure. After that, the robust watermark is embedded in the Q-LL sub-band to produce a watermarked quantized LL (WQ-LL) sub-band with the following equation:

$$\mathrm{WQLL_n} = \mathrm{QLL_n} + W_{Rn} * \mathrm{k}\ , \quad (2)$$

where $\mathrm{QLL_n}$ is the pixel value in the Q-LL sub-band, $\mathrm{WQLL_n}$ is the resulting pixel value in the WQ-LL sub-band, $\mathrm{WR_n}$ is the embedded robust watermark bit, and k is a constant parameter that corresponds to the strength of the watermark. A higher k can increase the strength of the embedded watermark, but it makes the watermarked image easier to perceive.

After embedding the watermark, the inverse DWTs of the four resulting sub-bands, i.e., LL, HL, LH, and WQ-LL, are performed to generate the watermarked Y' channel. Then, the watermarked Y', Cb, and Cr channels are converted back into RGB and saved as the robust watermarked image.

Phase 2: Embedding the invisible fragile watermark

The robust watermarked color image is generated after robust watermarking. In this phase, the same fragile watermark (WF) for image authentication is embedded independently on each RGB color channel of the robust watermarked image.

For each channel, i.e., R or G or B, first, we transform the fragile watermark bitstream $\mathrm{W_F}$ into a sequence of digits $\mathrm{W_F}'$ by using a 3^n-base notational system, where *n* is a parameter. Later, each digit in sequence $\mathrm{W_F}'$ is treated as one hidden digit *s*, which can be embedded into an *n*-pixel unit U= $(p_1, p_2, \ldots, p_n)$ of each channel. The basic idea of the embedding procedure is to perform the LSB replacement for unit U in 3^n-base notational system. The detailed embedding process of each hidden digit *s* is described in detail as follows:

1) Step 1: A digit E is extracted from each n-pixel unit U by using (3). Here, function $\Im()$ is defined as an extracting function that also is used for extracting the fragile watermark in subsection III-B. This equation is equivalent to extract the LSB of unit U in 3^n-base notational system, and p_i can be regarded as a digit in unit U.

$$\mathrm{E} = \Im(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} 3^{i-1} p_i \mod 3^n . \quad (3)$$

2) Step 2: To embed hidden digit s, the digits in Unit U need to be adjusted and make sure the resulted E equals to s. Therefore, a temporary value t is generated for adjusting the original unit U by using (4).

$$\mathrm{t} = (s - \mathrm{E} + \left\lfloor \frac{3^n - 1}{2} \right\rfloor) \bmod 3^n . \quad (4)$$

3) Step 3: To match each digit p_i in Unit U, the temporary value t is transformed into a sequence t' by using a 3-base notational system, where t' = $b_1 b_2 \ldots b_n$, b_i is a digit in t' and $1 \leq i \leq n$.

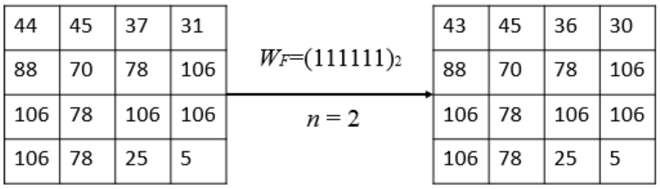
4) Step 4: Based on the ternary modulation property, each digit in sequence t' is then reduced by 1 to generate a subtracted sequence $\mathrm{t}'' = d_1 d_2 \ldots d_n$, where $d_i = b_i - 1$.

5) Step 5: Each pixel of the original n-pixel unit U is added to a corresponding digit of subtracted sequence t'' by using (5) to generate the watermarked pixel unit $U' = (p'_1, p'_2, \ldots, p'_n)$. Therefore, the extracted E from unit *U'* would be identical to hidden digit *s*.

$$p_i' = p_i + d_j , \quad \text{where } 1 \leq \mathrm{i} \leq \mathrm{n} \text{ and } \mathrm{j} = \mathrm{n} - \mathrm{i} + 1 . \quad (5)$$

These steps are repeated until all of the digits in sequence WF' are embedded in each RGB color channel of the robust watermarked image, resulting in a dual watermarked image. To make it easier to understand, an example of the fragile watermarking procedure is illustrated as follows:

Assume that the left table in Fig. 3 is a 4×4-pixels block of the R channel, watermark $\mathrm{W_F}$ is $(111111)_2$, and *n* is set as 2. First, $\mathrm{W_F} = (111111)_2$ is transformed into $\mathrm{W_F}' = (70)_9$ by using a 3^2-base notational system. Therefore, there are two hidden digits in $\mathrm{W_F}' = (70)_9$, one of which is 7, and the other is 0. We choose pixels 44 and 45 as the *2*-pixel unit U = (44, 45) to embed hidden digit *s* = 7. By using Equation (3), digit *E* = 8 is extracted from U. Later, by using Equation (4), the temporary value *t = 3* is calculated, and it is transformed into a sequence $t' = (10)_3$ by using a 3-base notational system. Then, each digit in sequence t' is reduced by 1 to generate a reduced sequence t'' = (0,-1). The watermarked pixel unit U' = (43, 45) can be obtained by using Equation (5), i.e., $p_1' = p_1 + d_2 = 44 - 1 = 43, p_2' = p_2 + d_1 = 45 + 0 = 45$. In a similar way, unit (37, 31) can be generated as a watermarked unit (36, 30) while embedding the other hidden digit 0. Therefore, after embedding watermark $\mathrm{W_F}$= $(111111)_2$, a new 4×4-pixels block of the R channel is generated, as shown in the right table in Fig. 3.



44	45	37	31
88	70	78	106
106	78	106	106
106	78	25	5

43	45	36	30
88	70	78	106
106	78	106	106
106	78	25	5

Fig. 3. Example of embedding a fragile watermark

As a special case, underflow or overflow would occur when the watermarked pixel p_i' is less than 0 or greater than 255, respectively. To solve with either an underflow or overflow situation, the original pixel p_i , which would result in underflowed or overflowed p_i', is increased by 1 or reduced by 1, respectively. After that, the proposed embedding algorithm is repeated to generate a new pixel unit U'. This procedure is repeated until all pixels in the new pixel unit U' range between 0 and 255.

B. *Extraction of the Dual Watermark*

In the proposed extraction procedure, the robust watermark and the fragile watermark can be extracted separately for the purpose of copyright detection and image authentication, respectively. The two different usages are described in detail below.

In terms of copyright protection, the process of extracting the robust watermark begins with converting the watermarked RGB image into YCbCr color space. After the YCbCr conversion, the one-level DWT decomposition of Y is performed to generate the LL, LH, HL, and HH sub-bands. Then, the LL sub-band is divided into several 8×8-sized blocks, and each pixel of the block of host image is quantized by directly dividing it by the corresponding value of the luminance quantization table, as shown in Fig. 2. After all of the blocks are quantized, a quantized LL (Q-LL) sub-band is generated. Then, the robust watermark is extracted by the following equation:

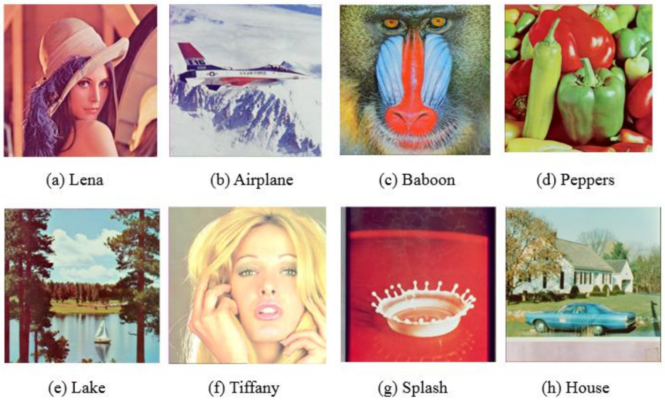
$$W_{Rn} = (\mathrm{HH_n} - \mathrm{QLL_n}) / \mathrm{k}, \quad (6)$$

where QLL_n is the pixel value in Q-LL sub-band, HH_n is the corresponding pixel value in HH sub-band, W_{Rn} is the extract robust watermark bit, and *k* is a constant parameter that is the same as that in the watermark embedding procedure. After all of the robust watermark bits have been extracted from the watermarked image, the owner of this image can be identified. Note that the extraction of the robust watermark is blind in that it does not require the original image or any information regarding the watermark.

Concerning image authentication, the extraction of the fragile watermark is proceeded by extracting three completed watermark bitstreams in the R, G, and B channel of the watermarked image separately. For each n-pixel unit in each channel, a hidden digit can be extracted by using the extracting function, as shown in (3). After extracting all of the hidden digits, the hidden sequence is transformed by using a 2-base notational system to provide an extracted watermark bitstream of each channel. The three extracted watermark bitstreams are compared with the original fragile watermark, and, if the watermark bit of any one of the extracted watermark bitstreams is not the same as the original watermark bit, the corresponding pixel is regarded as being modified, and we mark it with a dark color. Therefore, the integrity of the image can be detected, and the tampered area can be located accurately.

IV. Experimental Results and Discussion

Performance results of imperceptibility, fragility, and robustness of the proposed dual watermarking mechanism are presented in Subsections IV-A, IV-B, and IV-C, respectively. In Subsection IV-D, functionality comparisons with related studies are presented to demonstrate the superiority of the proposed scheme. The experiments were performed with eight commonly used color images, i.e., “Lena,” “Airplane,” “Baboon,” “Peppers,” “Lake,” “Tiffany,” “Splash,” and “House.” All of the images were the same size, i.e., 512 × 512, as shown in Fig. 4. Although all eight images were tested during the experiment, most of results presented in this section are based on “Lena” for succinct presentation. The robust watermark embedded in the test images was a 64 × 64-sized grayscale logo image. The embedded fragile watermark was a random binary bitstream and the parameter n was chosen as 2 in our experiments.


(a) Lena (b) Airplane (c) Baboon (d) Peppers

(e) Lake (f) Tiffany (g) Splash (h) House

Fig. 4. Eight original test images

A. Imperceptibility Results

In the experiments, two image quality assessment (IQA) metrics [12] (i.e. peak signal-to-noise ratio (PSNR) and structural similarity (SSIM)) were used to measure the imperceptibility performance of the dual watermarking mechanism. The PSNR is a conventional IQA metric which operate directly on the pixel-based stage of the images. The well-known SSIM brings IQA from conventional pixel-based stage to structure-based stage, which is based on the hypothesis that human visual system (HVS) is highly adapted to extract the structural information from the visual scene. Note that the values of SSIM are always in the range of 0 to 1. All of the three IQA metrics were demonstrated suitable and widely used to measure the image quality, and a higher IQA value implies higher visual quality of the watermarked image for a given test image.

The parameter *k* is a crucial parameter of the proposed dual watermarking mechanism. A higher *k* can increase the strength of the embedded watermark, but it makes the watermarked image easier to perceive. Therefore, we first tested the imperceptibility performance of the dual watermarking mechanism with various *k* from value 0.1 to 0.4. The results of the three IQA metrics of the proposed mechanism with various *k* are shown in Fig. 5 and Fig. 6, respectively. The results of PSNR values in Fig.5 show an inverse relationship with *k*, but all of the PSNR values are greater than 30 dB regardless of the *k* value. The average PSNR value of the eight dual watermarked images is highest (nearly 40dB) when *k* is 0.1, and it decreases to approximate 31 dB when *k* is 0.4. Moreover, there is very little difference between the PSNR values for the eight test images with the same *k* value, although each test image has different characteristics. The features can also find in the results of SSIM values in Fig. 6. The average SSIM values see a fall from nearly 0.98 to about 0.96 when *k* is increased from 0.1 to 0.4. Generally, a PSNR value greater than 30 dB or a SSIM value greater than 0.9 means the two compared images are not visibly different. The results of three IQA metrics in Fig. 5 and Fig. 6 indicate that the proposed dual watermarking mechanism achieves very impressive and outstanding imperceptibility performance.

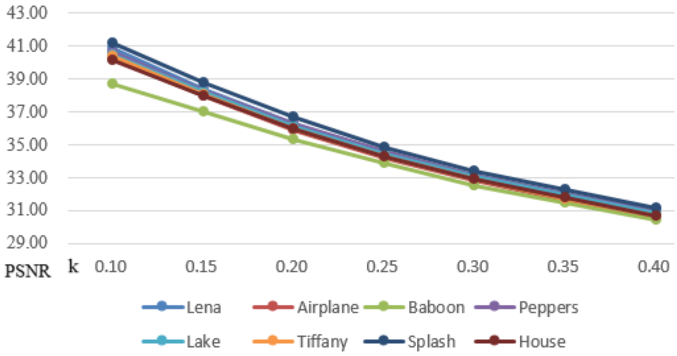


Fig. 5. PSNR values of the dual watermarked images with various k

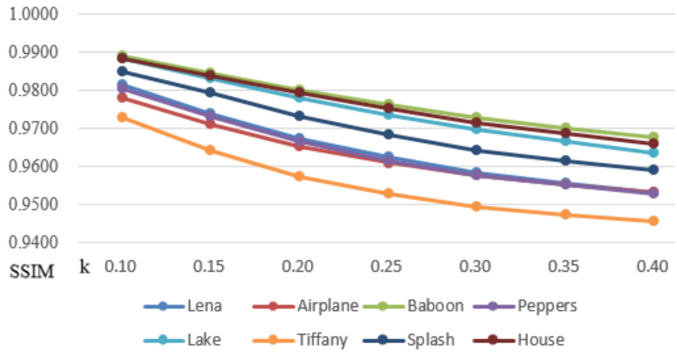


Fig. 6. SSIM values of the dual watermarked images with various k

Table I summarizes the three average IQA values of the eight test images after the watermark was embedded when k is 0.1, where column “Robust watermarked” presents the result of the image after robust watermarking only, column “Fragile watermarked” is the IQA value of the image after fragile watermarking only, and column “Dual watermarked” means the result of both robust and fragile watermarking. The average PSNR values of robust watermarked, fragile watermarked, and dual watermarked images are 40.85, 49.75, and 40.32 dB, respectively. Fragile watermarking only produces a higher PSNR value than robust watermarking only because the improved LSB embedding in the fragile watermarking procedure has less visual impact than the sub-band replacement operation in the robust watermarking procedure. In addition, dual watermarking just results in a small decrease in the PSNR value (about 0.5 dB) compared with robust watermarking only. In terms of SSIM, the average values are approximate to 0.99 which are quite satisfactory. In addition, the variance of SSIM for different watermarked images is very slight.

TABLE I
AVERAGE IQA VALUES OF THE WATERMARKED IMAGES

IQA	Robust watermarked	Fragile watermarked	Dual watermarked
PSNR	40.85	49.75	40.32
SSIM	0.9858	0.9968	0.9830

B. *Results of Fragile Watermarking*

We developed our proposed fragile watermarking particularly for authenticating images and locating tampered areas. Fig. 7 shows the attacked images and their corresponding authentication results, where N_{MP} represents the total number of modified pixels in the tampered image, N_P is the number of detected pixels, and accurate rate (AR) is the percentage of detected pixels in modified pixels. As Fig. 7(a) shows, some malicious attacks were performed on the test image “Lena,” including adding an extra object, adding a fake watermark, and blurring Lena’s eyes. Fig. 7(b) shows the results of image authentication and tampered area localization. Fig. 7(c) shows the difference between an attacked image and an authenticated image. The results displayed in Figs. 7(b) and (c) show that the tampered area of the attacked image can be located accurately (nearly 100%), irrespective of the attack that is performed.

In the proposed scheme, the same fragile watermark was embedded three times in the RGB channel of the watermarked image, thus any pixel modification of R or G or B would be detected. In addition, *2* pixels were chosen as a pixel unit U in the embedding strategy, thus any change in the pixels in unit U would be reflected in the extracted watermark. The multiple reflecting function in our proposed scheme can enhance the accuracy of tamper detection significantly during the image authentication procedure. The experimental results also implied that our scheme is indeed fragile enough to authenticate the attacked image and locate the tampered area accurately.

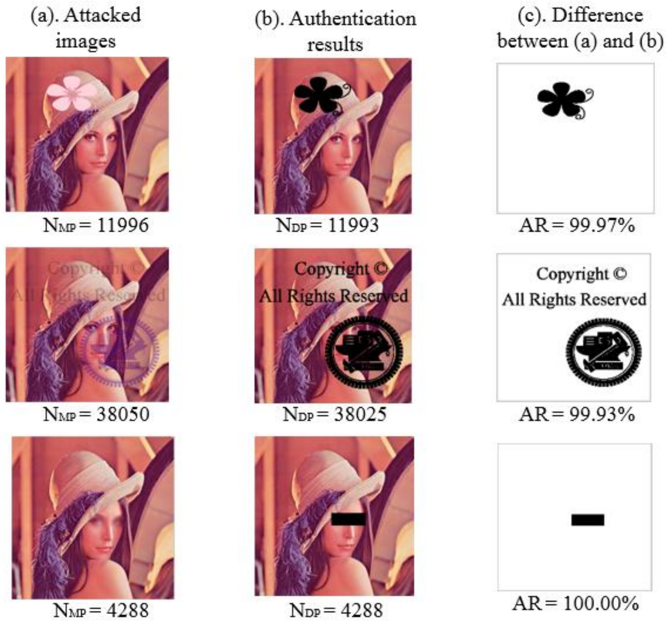


Fig. 7. Attacked images and their corresponding authentication results

To further demonstrate the fragile watermarking performance, we compared the proposed scheme with traditional LSB substitution [13]. The comparison results of LSB substitution and the proposed scheme are shown in Table II. Although the same fragile watermark is also embedded three times in the RGB channel, the average tamper detection accurate rate of LSB substitution is 88.58%. In Table II, the average detection accurate rate of the proposed scheme is 99.97%, which is much larger than that of LSB substitution scheme. No matter how many pixels are modified in the watermarked image, the tampered area can be located accurately in the proposed scheme. It is obvious that the proposed scheme achieves the optimal performance and superiority.

TABLE II
COMPARISON RESULTS OF LSB SCHEME [13] AND THE PROPOSED SCHEME

Attacked Images	LSB scheme [13]			Proposed scheme		
	N_{MP}	N_{DP}	AR	N_{MP}	N_{DP}	AR
Adding flower	11996	10458	87.18%	11996	11993	99.97%
Fake watermark	38050	33541	88.15%	38050	38025	99.93%
Blurring eyes	4288	3876	90.40%	4288	4288	100.00%
Average	18111	15959	88.58%	18111	18102	99.97%

C. *Results of Robust Watermarking*

Robustness is a significant concern for copyright protection mechanisms. In this subsection, we report the results of several signal processing attacks that were performed on the test image “Lena” to demonstrate the robustness of our watermarking scheme. The signal processing attacks that were performed in our experiments were adding an object, JPEG compression, salt and pepper, Gaussian noise, brighten, darken, resizing, cropping, blurring, contrast, tone mapping, and twisting. After extracting the watermark, the well-known metric normalized correlation coefficient (NC) was computed using the original watermark and the extracted watermark to measure the robustness of the watermarking scheme. The calculated NC is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSVT.2016.2633878, IEEE Transactions on Circuits and Systems for Video Technology

in the range of 0 to 1, if the value is closer to 1, the extracted watermark is getting more similar to the embedded one.

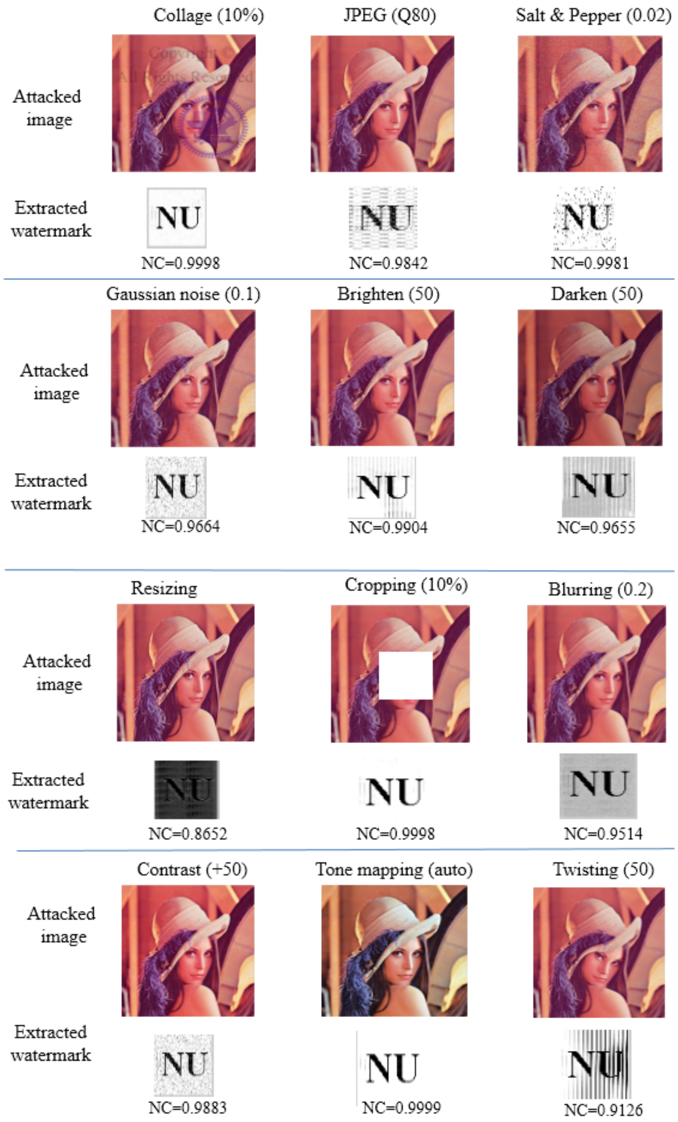


Fig. 8. Attacked images and extracted robust watermarks after various attacks

Fig. 8 illustrates the images subjected to the above-mentioned attacks with given parameter and the visual inspection of the extracted watermark images. Fig. 10 shows that even though the quality of the attacked image has been seriously distorted, the extracted watermark images are recognizable and the NC values are quite close to 1. An exception was that the extracted watermark image was not very satisfactory after we performed the resizing attack on the test image (NC is 0.8652). This is because the resizing attack smoothed the detailed part of the watermarked image, which significantly distorted the HH sub-band of DWT decomposition. Moreover, if we want to extract the watermark after a resizing attack, the attacked image should be scaled back to the original size of the watermarked image. This operation would result in further distortion of the HH sub-band. However, the proposed robust watermarking scheme was highly correlated with the HH sub-band of watermarked image's Y channel, and distorting it excessively would decrease the robustness of the watermark. Hence, we defined a higher value of parameter k in our robust watermarking scheme in order to resist resizing attacks. In general, the proposed scheme is robust against most malicious attacks, and the extracted watermark images are still recognizable by the human eye. This experimental result demonstrated that the proposed scheme can protect against the most common attacks and is beneficial in protecting the copyright of valuable images.

To highlight the robustness of the proposed scheme, variant parameters of each attack were also tested and a quantitative comparison with the other similar schemes was presented. Table III lists the IQA and NC results compared with the similar robust watermarking schemes under some common attack types. All of the compared robust watermarking schemes are based on frequency domain except Lusson et al.'s scheme [9], which is based on spatial domain. Therefore, Lusson et al.'s scheme [9] is not robust to most of the common attacks. Compared with Su et al.'s scheme [5], both of our schemes can resist against the common attacks with outstanding NC results. However, the imperceptibility performance of Su et al.'s scheme [5] is satisfactory, especially for SSIM result. They embed the robust watermark by roughly modified the U matrix of singular value decomposition (SVD) in original image, which would generate lots of obvious horizontal stripe in the watermarked image. Therefore, after considering the robustness and imperceptibility performance of the watermarked image adaptively. As can be seen from the above comparison, the overall imperceptibility and robustness performance of the proposed method outperform the other similar schemes.

TABLE III
IQA and NC comparison results of related robust watermarking schemes

Attack types	Parameters	Lusson et al. [9]	Su et al. [5]	proposed scheme

		PSNR = 38.97
		SSIM = 0.9793
			PSNR = 36.30
			SSIM = 0.9050
				PSNR = 40.85
				SSIM = 0.9814

Salt & pepper noise	0.01	0.9986	0.9952	0.9990
	0.02	0.9946	0.9889	0.9981
	0.04	0.9913	0.9847	0.9959
	0.08	0.9898	0.9600	0.9893
JPEG	90	0.9118	0.9988	0.9967
	80	0.8976	0.9950	0.9842
	70	0.8873	0.9907	0.9711
	60	0.8844	0.9860	0.8687
Blurring	0.1	0.9998	1	0.9997
	0.2	0.4436	0.9877	0.9514
	0.3	0.3279	0.9393	0.8735
Gaussian noise	0.1	0.9664	0.9767	0.9664
	0.3	0.9334	0.9003	0.9171
	0.5	0.9064	0.8697	0.8735
Brighten	50	0	0.9208	0.9904
	80	0	0.8040	0.9806
Darken	50	0	0.9999	0.9655
	80	0	0.9921	0.8813

Twist	50	0.1681	0.8952	0.9126
	-50	0.1404	0.8926	0.9123
Resizing	200%	0.9998	0.9997	0.8652
	50%	0.5188	0.0031	0.5712
Contrast	+50	0.0010	1	0.9883
	-50	0.0736	1	0.9652
Collage	10%	0.9021	0.9999	0.9998
	30%	0.8281	0.9931	0.9954
Cropping	10%	0.9132	0.9998	0.9998
	30%	0.8325	0.9820	0.9989
Tone mapping	auto	0.0414	0.9999	0.9999

D. Comparisons among Dual Watermarking Mechanisms

In this subsection, the superior performance of the proposed dual watermarking mechanism is demonstrated by comparing its functionality with that of the related well-known dual watermarking mechanisms [7, 8, 9]. Table IV illustrates the different functionalities of the dual watermarking mechanisms.

The major difference between the four schemes is that both of scheme [7] and our proposed scheme present invisible hybrid watermarking for copyright protection and image authentication, whereas scheme [8] and scheme [9] concentrate on intensive robust watermarking for copyright protection with no concern about image authentication. Therefore, after the integrity of the protected image has been compromised, scheme [8] and scheme [9] cannot detect the infringement. However, the infringement can be detected accurately by our proposed scheme and scheme [7] by using the image authentication mechanism. With respect to copyright protection, all four of the dual watermarking schemes can resist multiple common attacks. However, in Lussion et al.'s scheme [9], the original host image is required to extract the second robust watermark in RGB color space, which is a non-blind watermarking scheme. In terms of the PSNR value of the watermarked image, the proposed mechanism achieved the optimal result (nearly 40 dB) among the four mechanisms that were evaluated. This means that our mechanism is suitable for protecting valuable images without attracting the attention of malicious attackers.

In addition, we compared the performances of our proposed scheme and Lu and Liao's scheme [7] in terms of image authentication and copyright protection. We chose scheme [7] for this purpose because their scheme and ours had the same objective, i.e., multi-purpose watermarking. In comparing their performances in authenticating images, the YCbCr color space was chosen in scheme [7] for image watermarking, and the watermark was embedded in the Y channel, which is the luminance channel. However, their scheme cannot detect chrominance changes if the color is modified and the intensity is left unchanged. However, our proposed scheme can detect the modified areas accurately irrespective of which channel is changed. Moreover, considering with the accuracy of locating normal tampered areas, scheme [7] cannot detected all the altered regions, and the detection rate is lower that 80%, which is not satisfactory. However, based on the discussion in Section IV-B, it is evident that all the altered regions can be detected successfully by the proposed scheme. These results show that the proposed scheme is outstanding among related watermarking schemes during image authentication. However, concerning the performance of copyright protection, the watermark extraction of scheme [7] is non-blind, and the embedded robust watermark is generated from the host image itself instead of a pre-defined logo image. This means that, during copyright verification, the ownership of image can be verified only by fuzzy detection instead of by extracting a recognizable watermarked image. In fact, such recognizable watermarks are very common in many practical applications. Thus, after considering the global functionally of various mechanisms, the proposed dual watermarking mechanism is demonstrably superior.

TABLE IV
COMPARISONS OF THE FUNCTIONALITIES OF OUR DUAL WATERMARKING METHOD AND RELATED DUAL WATERMARKING MECHANISMS

Functionality	Lu and Liao [7]	Lin et al. [8]	Lusson et al. [9]	Our Scheme
Dual watermarks	Fragile + Robust	Robust+ Robust	Robust + Robust	Fragile + Robust
Embedding domain	DWT + DWT	Spatial + DCT	Spatial + Spatial	Spatial + DWT
Visibility	Invisible + Invisible	Visible + Invisible	Invisible + Invisible	Invisible + Invisible
Blind extraction	Yes + No	Yes + Yes	Yes + No	Yes + Yes
Target image	Color	Grayscale	Color	Color
PSNR	~40 dB	~30 dB	~39 dB	~40 dB
Copyright protection	Yes	Yes	Yes	Yes
Image authentication	Yes	No	No	Yes

E. Security analysis

Concerning the security of embedded watermarks, the advantages of the fragile and robust watermarking in this paper is illustrated in this subsection. Different with the traditional LSB substitution, the LSB bit plane is not mapped directly to a watermark bit in our fragile watermark embedding scheme. The embedded fragile watermark is imperceptible, and it cannot be extracted by the attackers using LSB attack tools to generate a counterfeit image that can pass the image authentication procedure. Therefore, the proposed fragile watermarking scheme is reliable for image authentication.

In terms of robust watermarking, the experiment results mentioned above have demonstrated the proposed scheme is robust against the most common attacks and beneficial in protecting the copyright of valuable images. However, Craver et al. [14] indicated that for most of the robust watermarking schemes, counterfeit attacks can easily allow anyone to claim ownership of any images he or she has access to. Inevitably, the same problem exists in the proposed robust watermarking scheme, the attackers may claim multiple ownerships by manipulating the HH sub-band or even using entirely different watermarking schemes.

Nonetheless, the counterfeit attacks problem can be successfully solved by the proposed dual watermarking mechanism. Any image suffers counterfeit attacks would be detected as a manipulated one through image authentication procedure, which means the ownership information extracted from this image is trustless . Only if the user who provides the original image where both fragile and robust watermarks are extracted can claim the ownership. Therefore, with the combination of fragile and robust watermarking, the proposed dual watermarking mechanism provides a security way for both copyright protection and image authentication.

V. Conclusions

In this paper, we presented a blind dual watermarking mechanism for color the authentication of images and copyright protection. The invisible, fragile, and robust watermarks are embedded into the spatial domain of the RGB color space and into the frequency domain of the YCbCr color space. The major contribution of this work is that our mechanism can achieve copyright protection and image authentication simultaneously, and the extraction of watermarks from the protected image can be processed blindly without the original host image and watermarks. According to the experimental results, the proposed watermarking mechanism can withstand various processing attacks and locate the tampered area of the image accurately. Moreover, the dual watermarked image is imperceptible, which makes the proposed mechanism suitable for protecting valuable original images. Comparisons of the functionality of our proposed mechanism with the functionalities of other, well-known dual watermarking mechanisms cleared demonstrated that the superiority of the proposed mechanism.

References

[1] M. Yeung, F. Mintzer, "An invisible watermarking technique for image verification," *Proceedings of the International Conference on Image Processing*, vol. 2, pp. 680-683,1997.

[2] C. C. Chang, Y. P. Hsieh, C. H. Lin, "Sharing secrets in stego images with authentication," *Pattern Recognition*, vol. 41, no. 10, pp. 3130-3137, 2008.

[3] A. Khan, S. A. Malik, "A high capacity reversible watermarking approach for authenticating images: Exploiting down-sampling, histogram processing, and block selection," *Information Sciences*, vol. 256, pp. 162-183, 2014.

[4] W. J. Chen, C. C. Chang, T. H. Ngan Le, "High payload steganography mechanism using hybrid edge detector," *Expert Systems with Applications*, vol. 37, pp.3292-3301, 2010.

[5] 5Q. Su, Y. Niu, H. Zou, X. X. Liu, "A blind dual color images watermarking based on singular value decomposition," *Applied Mathematics and Computation*, vol. 219, no. 16, pp. 8455-8466, 2013.

[6] S. P. Maity, S. Maity, J. Sil, C. Delpha, "Collusion resilient spread spectrum watermarking in M-band wavelets using GA-fuzzy hybridization," *Journal of Systems and Software*, vol. 86, no. 1, pp. 47-59, 2013.

[7] C. S. Lu, H. Y. M. Liao, "Multipurpose watermarking for image authentication and protection," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1579-1592, 2001.

[8] P. Y. Lin, J. S. Lee, C. C. Chang, "Dual digital watermarking for internet media based on hybrid strategies," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 8, pp. 1169-1177, 2009.

[9] F. Lusson, K. Bailey, M. Leeney, K. Curran, "A novel approach to digital watermarking, exploiting colour spaces," *Signal Processing*, vol. 93, no. 5, pp. 1268-1294, 2013.

[10] C. C. Lin, Y. H. Chen, C. C. Chang, "LSB-based high-capacity data embedding scheme for digital images," *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 11, pp. 4283-4289, 2009.

[11] Int. Telecommunication Union, "Information technology-digital compression and coding of continuous-tone still images-requirements and guidelines," *CCITT Recommendation T.81*, 1992.

[12] L. Zhang, L. Zhang, X. Q. Mou, D Zhang, "FSIM: A feature similarity index for image quality assessment", *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, 2011.

[13] R. Z. Wang, C. F. Lin, J. C. Lin, "Image hiding by optimal LSB substitution and genetic algorithm," *Pattern recognition*, vol. 34, no. 3, pp. 671-683, 2001.

[14] Q. Su, Y. Niu, X. Liu, Y. Zhu, "A blind dual color images watermarking based on IWT and state coding," *Optics Communications*, vol. 285, no. 7, pp. 1717-1724, 2012.

Xiao-Long Liu received his B.S. degree from Xiamen University, China in 2011, his M.S. degree in Computer Science and Information Management from Providence University, Taiwan in 2013, his Ph.D. degree in Institute of Computer Science and Engineering, National Chiao Tung University, Taiwan in 2016. His current research interests include image processing, multimedia security and cloud computing.

Chia-Chen Lin (also known as Min-Hui Lin) received her Ph.D. degree in information management in 1998 from the National Chiao Tung University, Hsinchu, Taiwan. Dr. Lin is currently a professor of the Department of Computer Science and Information Management, Providence University, Sha-Lu, Taiwan. Her research interests include image and signal processing, information hiding, mobile agent, and electronic commerce.

Shyan-Ming Yuan received his Ph.D. degree in Computer Science from the University of Maryland, College Park in 1989. Since September 1990, he has been an Associate Professor at the Department of Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan. He became a professor in June 1995. His current research interests include distance learning, internet technologies and multimedia security.