Chamlawi R, Khan A, Idris A. Wavelet based image authentication and recovery. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 22(6): 795~804 Nov. 2007

Wavelet Based Image Authentication and Recovery

Rafiullah Chamlawi[1], Asifullah Khan[1,*], and Adnan Idris[2]

[1] *Pakistan Institute of Engineering and Applied Sciences (PIEAS), Nilore, Islamabad, Pakistan*

[2] *AJK University, Rawalakot, Azad Jammu and Kashmir (AJK), Pakistan*

E-mail: chamlawi@pieas.edu.pk; asif@pieas.edu.pk; adnanidris@gmail.com

Received March 14, 2007; revised September 19, 2007.

Abstract In this paper, we propose a secure semi-fragile watermarking technique based on integer wavelet transform with a choice of two watermarks to be embedded. A self-recovering algorithm is employed, that hides the image digest into some wavelet subbands for detecting possible illicit object manipulation undergone in the image. The semi-fragility makes the scheme tolerant against JPEG lossy compression with the quality factor as low as 70%, and locates the tampered area accurately. In addition, the system ensures more security because the embedded watermarks are protected with private keys. The computational complexity is reduced by using parameterized integer wavelet transform. Experimental results show that the proposed scheme guarantees safety of a watermark, recovery of image and localization of tampered area.

Keywords semi-fragile watermark, integer wavelet transform (IWT), discrete cosine transform (DCT), JPEG compression, authentication and self-recovery

1 Introduction

In modern society, which relies heavily on digitized information, the multimedia contents may easily be copied, manipulated and distributed, this gives rise to the challenging task of protecting digital content, especially for content owners and distributors and has gained much attention in recent times. The ease, by which digital multimedia data can be manipulated, has always raised many concerns about the reliability of their content[1]. Digital data authentication is thus one of the most important and investigated security applications in this regard. The fundamental role of watermarking is the reliable embedding and detection of information and therefore, it is generally considered as a form of communications. Consequently, the field of watermarking has great potential in both copyright and authentication based applications. However, the security aspect of watermarking is still an unsolved problem and many concepts in this context are borrowed from the field of Cryptography[2~5].

Authentication of image is the act of establishing or confirming that the image is credible[2]. Imperceptibility, fragility, security and efficient computation are the basic requirements for authentication. In this paper we strive for imperceptibility, efficient computation, security and also both the fragility and robustness, i.e., semi-fragility. The use of two different potential watermarks in our proposed scheme also enhances the security of our semi-fragile watermarking based authentication system.

Previous techniques[6] do not clarify that how and where the image is tampered but only identify that the image is tampered or not. Friedman *et al.*[7,8] proposed the digital camera in which the signature is embedded in each image and that signature is used to identify the camera that produced the image. Hua *et al.*[9], on the other hand, have proposed a new fragile watermarking technique, which is based on the Gaussian mixture model in different wavelet scales. In [10], the performance of the semi-fragile authentication watermarking is improved. The authors extend the JPEG DCT based watermarking technique to the integer wavelet transform (IWT) domain so that it could be compatible with JPEG2000 compression. Their objective is to improve the performance tradeoff between the alteration detection sensitivity and the false alarm rate and apply them to authenticating JPEG2000 images.

The presented work exploits the advantages of both the techniques[1,11] with some modifications enabling our proposed method of acquiring both authentication and recovery based attributes. Consequently, our approach is based on a comprehensive technique employing two watermarks[12], an image digest and a bi-

Regular Paper

*Corresponding Author

This work was supported by the Higher Education Commission of the Government of Pakistan under the Indigenous Ph.D. Scholarship Program (Grant No. 17- 5-I (Cu - 180) HEC/Sch/2004/4343).

nary image. The image digest is computed through a properly modified version of JPEG coding, operating at very high compression ratio on the original image[1]. Thus, image digest is a compressed version of the image itself that helps in obtaining an estimate of the original contents. The modification is introduced in the digest to make it insensitive to global, innocuous manipulations. The second watermark, binary signature is processed with a private key to ensure security[11]. Embedding binary image can help in accurately detecting manipulations made in image, but it cannot ensure recovery of an estimated image. Similarly embedding image digest can retrieve the estimated image but leaves the users to judge the authenticity by themselves. Thus, embedding image digest as well as binary image can lead to both authentication and recovery. In this regard, we make the following contributions.

- Embedding multiple watermarks, but with different intended applications, to achieve both authentication and recovery based attributes at the cost of only a small reduction in imperceptibility. Both these watermarks, as analyzed, strengthen each other in context of security.
- Utilization of IWT for image digest generation instead of using Discrete Wavelet Transform[1] to reduce computational cost effectively.

For the reason we use image digest as a compressed version of the original image, our technique can also be referred to as a self-recovery technique. The scheme is flexible enough with the choice of users, either to embed image digest or binary image, or both.

In the current communication, we discuss the watermarks generation, embedding and extraction in Section 2. Section 3 explains tamper detection. We report the experimental results in Section 4 and compare the performances in Section 5. Finally the paper is concluded in Section 6.

2 Watermark Generation

The scheme is based on embedding of two watermarks. We proceed for the watermark generation in the following subsection.

2.1 Binary Image Preprocessing

A binary signature (Binary Image) which is used for accurate authenticity of the cover image is preprocessed before being embedded. Let $\boldsymbol{W}$ be a binary signature of size $X \times Y$, then

$$\boldsymbol{W} = w(i,j) \quad (1 \leqslant i \leqslant X,\ 1 \leqslant j \leqslant Y) \tag{1}$$

where $w(i,j) \in \{0,1\}$ and $\boldsymbol{P}_{\mathrm{Rand}}$ be a pseudo-random matrix of the same size generated by a secret key.

$$\boldsymbol{P}_{\mathrm{Rand}} = R_n(i,j) \quad (1 \leqslant i \leqslant X,\ 1 \leqslant j \leqslant Y) \tag{2}$$

where $R_n(i,j) \in \{0,1\}$.

We adopt (3) to get the ultimate watermark $\widehat{\boldsymbol{W}}_1$:

$$\widehat{\boldsymbol{W}}_1 = \boldsymbol{W} \oplus \boldsymbol{P}_{\mathrm{Rand}} \tag{3}$$

where $\oplus$ denotes the exclusive OR.

2.2 Digest Generation

The image digest (second watermark) which is the highly compressed version of the original image, is generated using the following steps[1].

- One level Integer Wavelet Transform is applied on the original image of size $N \times N$. The levels are called approximation LL1, Horizontal HL1, Vertical LH1 and Diagonal HH1. We select LL1 to create the image digest after high compression.
- Full frame DCT on low pass version (LL1) is computed.
- DCT coefficients are quantized using JPEG quantization matrix[1], as shown in (4) to decrease their obtrusiveness.

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}. \tag{4}$$

- The scaled DCT values are ordered through a zigzag scan and the first M coefficients are selected and stored in vector $\boldsymbol{q}$:

$$\boldsymbol{q} = (q_1, q_2, q_3, \ldots, q_M) \tag{5}$$

where $M = N^2/32$. DC component is not included because of its high energy. If we include the DC component in digest generation, then the watermark will be perceptible because of its exceeding high energy.

- Vector $\boldsymbol{q}$ is further scaled, based on secret key (k_1):

$$\boldsymbol{q}_{\mathrm{scaled}}(i) = \boldsymbol{q}(i) \cdot \alpha \cdot \ln(i + 2 + r(i)) \tag{6}$$

where α is a strength factor and its value depend upon the image quality while r is the shift parameter ranging from -0.5 to 0.5.

- DCT coefficients are quadruplicated because we have $N^2/8$ available positions for embedding (Fig.1), which are four times in number to $M = N^2/32$ the number of DCT coefficients. Thus we obtain the new vector $\boldsymbol{V}$ as:

$$\boldsymbol{V} = (q_1, q_2, \ldots, q_M, q_1, q_2, \ldots, q_M, q_1, q_2, \ldots, q_M, q_1, q_2, \ldots, q_M). \quad (7)$$

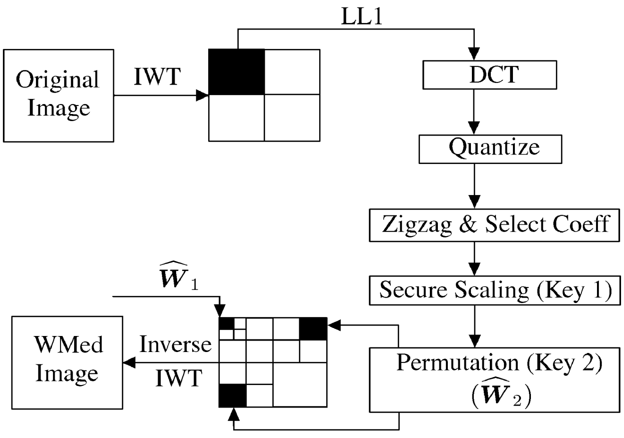


Fig.1. Embedding diagram.

- $\boldsymbol{V}_{\text{Permuted}}$ is obtained by scrambling the vector $\boldsymbol{V}$ with the help of a secret key (k_2) in order to make it more secure. Due to this permutation, the four copies of the DCT coefficients will occupy different locations in the two subbands, HL2 and LH2. Thus, $\boldsymbol{V}_{\text{Permuted}}$ is the resultant image digest to be embedded in the highlighted subbands, as shown in Fig.1, $\widehat{\boldsymbol{W}}_2$ is our second watermark ready for embedding.

$$\widehat{\boldsymbol{W}}_2 = \boldsymbol{V}_{\text{Permuted}}. \quad (8)$$

2.3 Watermark Embedding

Both the watermarks have been computed and are ready to be embedded into the original image with the following steps.

- Given an $N \times N$ image, after applying a 1-level Integer Wavelet Transform (IWT), the horizontal subband HL1 and vertical subband LH1 are further decomposed while the approximation subband LL1 is two times decomposed to get LL3. Embedding areas HL2, LH2 and LL3 are highlighted in Fig.1. Note that our scheme is more secure because we are using three secret keys, key1, key2 and key3. One key, i.e. key1, is used in generating Pseudo Random Number Matrix (PRNM), while pre-processing the first watermark (binary image) and other two keys, i.e. key2 ad key3, are used while generating the second watermark (image digest). Key2 is used in scaling the DCT coefficients and key3 is used in permuting (scrambling) the second watermark before embedding.
- We use the following formula to embed the watermark $\widehat{\boldsymbol{W}}_1$ in the LL3 subband coefficients[13]. Let $LFB(a)$ denote the five least significant bits of a, while $LFB(a,b)$ represent the substitution of b for the five least significant bits of a. The two choices "11000" and "01000" representing "1"and "0" respectively, are selected from the distance diagram[13], shown in Fig.2. We select these two choices according to the quality of the watermarked image.

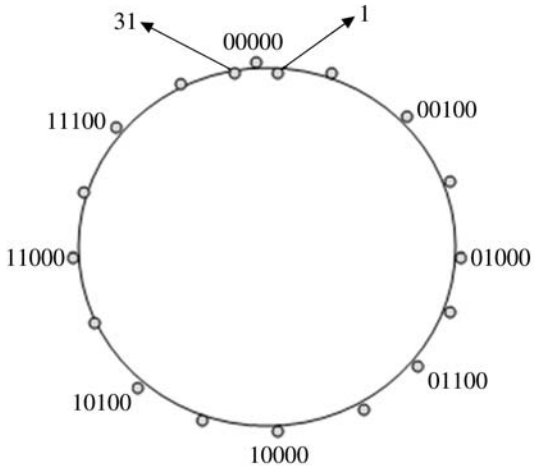


Fig.2. Distance diagram.

If we simply replace the least five significant bits of the coefficient by these choices, then the amplitude changes from -7 to 24 when "1" is embedded and from -23 to 8 when "0" is embedded. Keeping the performance of invisibility and robustness, the following embedding method is proposed[13].

When $\widehat{\boldsymbol{W}}_1(i,j) = 0$, (9) is adopted:

$$\boldsymbol{f}^*(i,j) = \begin{cases} LFB(f(i,j) - 01000, 11000), \\ \qquad\qquad \text{if } LFB(f(i,j)) \leqslant 01000, \\ LFB(f(i,j), 11000), \quad \text{otherwise.} \end{cases} \quad (9)$$

When $\widehat{\boldsymbol{W}}_1(i,j) = 1$, (10) is adopted:

$$\boldsymbol{f}^*(i,j) = \begin{cases} LFB(f(i,j) + 10000, 01000), \\ \qquad\qquad \text{if } LFB(f,(i,j)) \leqslant 11000, \\ LFB(f(i,j), 01000), \quad \text{otherwie,} \end{cases} \quad (10)$$

where $\boldsymbol{f}(i,j)$ is an IWT coefficient in LL3 subband before embedding, $\boldsymbol{f}^*(i,j)$ is the IWT coefficient after embedding. With such embedding, the amplitude of the coefficients changes from -15 to 16. The two choices "11000" and "01000" are used to represent the bits "1" and "0" respectively. On the authentication side, we will just examine the fifth least significant bit of these choices[13].

- The second watermark $\widehat{\boldsymbol{W}}_2$ is substituted in details, HL2 and LH2 subbands highlighted in Fig.1. Sizes of $\widehat{\boldsymbol{W}}_2$ and the two subbands HL2 and LH2 are the same.

- Performing the inverse IWT, we get the watermarked image.

In Fig.1, $\widehat{\boldsymbol{W}}_1$ is the binary image to be embedded in the LL3 subband[11] and $\widehat{\boldsymbol{W}}_2$ is the image digest to be embedded in HL2 and LH2 subbands[1].

Peak Signal to Noise Ratio (PSNR) is used to measure the induced distortion caused by the watermark[14]. PSNR in decibels (dB) is computed[15] using (11):

$$PSNR = 20\log_{10}\left[\frac{255^2}{\dfrac{1}{RS}\sum\limits_{i,j}(x(i,j)-y(i,j))^2}\right] \tag{11}$$

where $1 \leqslant i \leqslant R$ and $1 \leqslant j \leqslant S$.

Our proposed scheme use the parameterize integer wavelet transform (IWT) which is the fast approach of Discrete Wavelet Transform (DWT). Meerwald *et al.*[16] have proposed for the first time to use the parameterized integer wavelet transform. However, their scheme is based on conventional DWT. Lifting scheme is an effective method to improve the processing speed of DWT. On the other hand, Integer wavelet transform allows constructing lossless wavelet transforms and through lifting scheme, thus we construct such integer wavelet transform. Consequently, our approach is based on a novel idea of using integer wavelet transform with parameters for the development of secure semi-fragile watermarking for both image authentication and recovery. In our current work, we have used Daubechies wavelet.

2.4 Integrity Verification

In the integrity verification phase, the watermarked image undergoes a procedure, where the embedded watermarks ($\widehat{\boldsymbol{W}}_1$ and $\widehat{\boldsymbol{W}}_2$) are extracted. The binary watermark $\widehat{\boldsymbol{W}}_1$ is extracted from the LL3 subband, while the image digest $\widehat{\boldsymbol{W}}_2$ is extracted from the HL2 and LH2 subbands. The extraction procedure of $\widehat{\boldsymbol{W}}_1$, which is used for authentication, includes the following steps.

- Given an $N \times N$ watermarked image, after applying a 1-level IWT, the approximation subband is two times decomposed and LL3 is selected as shown in Fig.3(a).

- Let $\widehat{\boldsymbol{W}}_1^{*\prime}(i,j)$ denote the extracted watermark bit and $LFB_{5\text{th}}(a)$ denote the fifth least significant bit of a, then

$$\widehat{\boldsymbol{W}}_i^{*\prime}(i,j) = \begin{cases} 1, & LFB_{5\text{th}}(f^{*\prime}(i,j)) = 0,\\ 0, & LFB_{5\text{th}}(f^{*\prime}(i,j)) = 1,\end{cases} \tag{12}$$
$$(1 \leqslant i \leqslant X,\ 1 \leqslant j \leqslant Y).$$

- Now, as the watermark has been processed, therefore, at the verification phase, we again process it to obtain the ultimate watermark $\widehat{\boldsymbol{W}}_1^{\prime}$ (a binary image) using (13):

$$\widehat{\boldsymbol{W}}_1^{\prime}(i,j) = \widehat{\boldsymbol{W}}_1^{*\prime}(i,j) \oplus \boldsymbol{P}_{\text{Rand}}(i,j) \tag{13}$$
$$(1 \leqslant i \leqslant X,\ 1 \leqslant j \leqslant Y)$$

where $\boldsymbol{P}_{\text{Rand}}$ is the Pseudo Random Number Matrix (PRNM) and $\widehat{\boldsymbol{W}}_1^{*\prime}$ is the extracted binary signature.

- We express the difference mark as (14):

$$\boldsymbol{D}(i,j) = |\widehat{\boldsymbol{W}}_1(i,j) - \widehat{\boldsymbol{W}}_1^{\prime}(i,j)| \tag{14}$$
$$(1 \leqslant i \leqslant X,\ 1 \leqslant j \leqslant Y).$$

If $\boldsymbol{D}(i,j) = 1$ then, the pixel in the difference binary image is white, and represents mark extraction error. On the contrary, black pixel represents accurate mark extraction.

To obtain the estimated image, we move further to extract $\widehat{\boldsymbol{W}}_2$. Following steps are taken for the extraction of $\widehat{\boldsymbol{W}}_2$.

- Horizontal and vertical details are further decomposed and HL2 and LH2 are selected.

- Here the data is reversed into a vector $\boldsymbol{V}^{\prime}_{\text{Scrambled}}$, which is inversely scrambled by means of the same key, thus resulting in a sequence $\boldsymbol{V}^{\prime}$. An estimate of the hidden DCT coefficients is then obtained by averaging all four copies of each extracted coefficient. A unique set of authentication data $\boldsymbol{q}_{\text{extracted}}$ (i.e., M coefficients) is obtained.

- Inverse scaling operation is performed using (15):

$$\boldsymbol{q}_{\text{reconstructed}}(i) = \boldsymbol{q}_{\text{extracted}}(i) \cdot \frac{1}{\alpha} \cdot \frac{1}{\ln(i+2+r(i))}. \tag{15}$$

- The inverse scaled coefficients are then replaced in their correct positions, by means of an anti-zigzag scanning (missing elements are set to zero and a DC component with the value 128 is inserted)[1].

- These obtained values are weighed back with the JPEG quantization matrix and inverse DCT is applied to obtain an approximation of the original image with size $N/2 \times N/2$. The aim of the recovered image in our proposed scheme is that, if someone tampers (maliciously or incidentally) the watermarked image, we can recover the estimated image. Even if the image is tampered or not, in both the cases, the complete estimated image is recovered, i.e., we are not recovering only the tampered blocks[1].

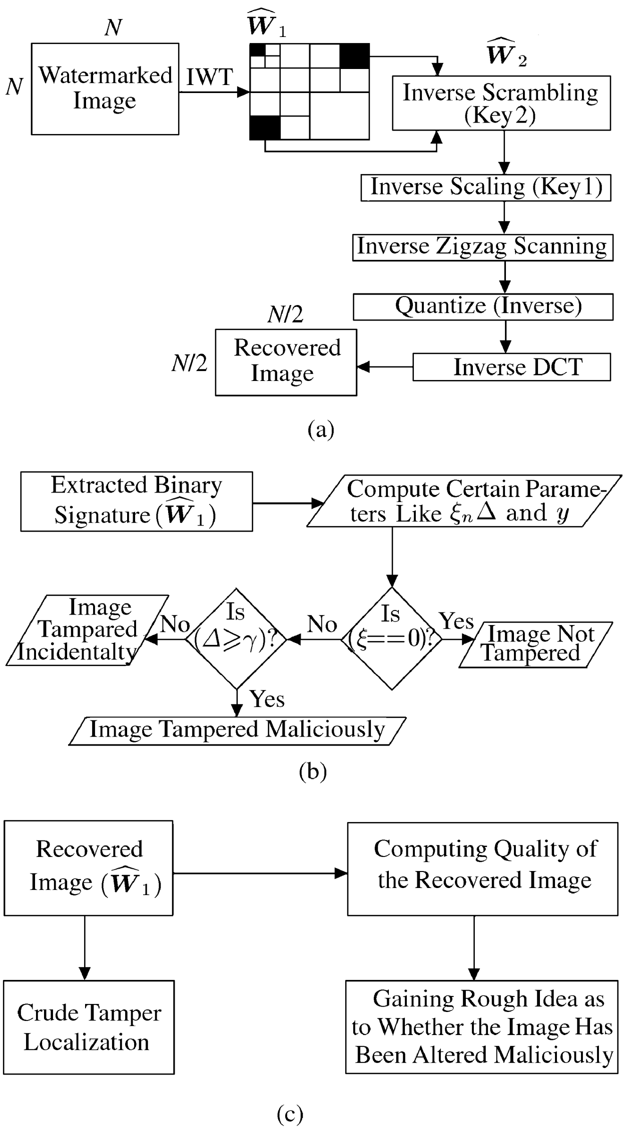

N
N
Watermarked Image
IWT
$\widehat{W}_1$
$\widehat{W}_2$
Inverse Scrambling (Key2)
Inverse Scaling (Key1)
Inverse Zigzag Scanning
Quantize (Inverse)
N/2
N/2
Recovered Image
Inverse DCT
(a)
Extracted Binary Signature ($\widehat{W}_1$)
Compute Certain Parameters Like $\xi_n\Delta$ and y
Image Tampered Incidentalty
No
Is ($\Delta\geqslant\gamma$)?
No
Is ($\xi==0$)?
Yes
Image Not Tampered
Yes
Image Tampered Maliciously
(b)
Recovered Image ($\widehat{W}_1$)
Computing Quality of the Recovered Image
Crude Tamper Localization
Gaining Rough Idea as to Whether the Image Has Been Altered Maliciously
(c)

Fig.3. (a) Extraction diagram. (b) Analysis of extracted signature. (c) Analysis of the recovered image $\widehat{W}_2$.

Both the watermarks $\widehat{W}_1$ and $\widehat{W}_2$ are extracted from the highlighted subbands, shown in Fig.3(a). The image digest $\widehat{W}_2$ is extracted from LH2 and HL2 subbands and the binary image $\widehat{W}_1$ is extracted from the LL3 subband.

Fig.3(b) shows the analysis of the extracted binary signature $\widehat{W}_1$, which is used for accurate authentication in our proposed technique. If somebody tampers the image, binary signature $\widehat{W}_1$ will detect it and in addition, is able to show whether the attack is malicious or incidental (see Section 4 for details).

Fig.3(c) shows the analysis of the image digest $\widehat{W}_2$ (recovered image), which can localize the tampering but not accurately. The image can be recovered after any type of attack; however, the quality of the recovered image degrades as the strength of the attack increases.

3 Tamper Detection

We express the difference between the original binary image and the extracted binary image watermark as:

$$Difference = |\widehat{W}_1(i,j) - \widehat{W}'_1(i,j)|. \qquad (16)$$

If *Difference* is "1" then it means that there exists a difference between the corresponding pixels of original and extracted binary watermarks. As we will see in the experimental results that "0", i.e., black pixel in the difference image corresponds to correctness while "1", i.e., white pixel in the difference image corresponds to error. Our proposed approach accurately locates the tampered area and distinguishes between malicious and incidental attacks. The details are given as follows.

Dense Pixel: for a mark error pixel in the difference image, it is a *dense pixel* if at least one of its eight neighbor pixels is a mark error pixel and a *sparse pixel* otherwise[11]. Thus, we have the following parameters.

Dense Area: the total number of dense pixels of LL subband.

Sparse Area: the total number of sparse pixels of LL subband.

Area: the total number of pixels of LL subband.

$$Total\ Area = Dense\ Area + Sparse\ Area,$$
$$\Delta = Dense\ Area\ /\ Sparse\ Area,$$
$$\xi = Total\ Area\ /\ Area,$$

- if $\xi = 0$, then the image is not tampered;
- if $\xi > 0$ and $\Delta < \gamma$, then tampering is incidental, where γ is set empirically between $0.5 \sim 1.0$;
- if $\Delta \geqslant \gamma$, then tampering is malicious.

Above parameters depict that if the difference image has sparse pixels, i.e., $\Delta < \gamma$, then the image is incidentally attacked like compression and file format change etc. Otherwise, in a case of dense pixels, the image is maliciously attacked i.e., tampered maliciously, as shown in Fig.4.

4 Experimental Results

We have tested our scheme on Lena and Cameraman images of different formats like bmp, tiff, etc. We apply two-and three-level IWT for embedding process. The PSNR of the watermarked images are 38~40db, which are quite reasonable. Fig.5 shows the original and watermarked images of Lena and a binary signature, which is embedded in LL3 subband of the Lena image.

The extracted watermark (binary signature) and recovered image without any attack are shown in Fig.6. Without attack, the recovered image is much better, very similar to the original image. The difference image, which is full black, shows that the image is not tampered. The number of dense pixels and sparse pixel in the difference image is zero.

The Lena image is tampered maliciously on the eye (cut/past). The image is recovered but the tampered areas are located in the extracted binary watermark and in the difference image as shown in Fig.6. The difference image is the difference in the original binary signature and extracted binary signature. In proposed scheme, we see that the quality of the recovered image degrades. If the watermarked image is highly tampered then it degrades the recovered image also. In addition, if we highly compress the watermarked image then the recovered image degrades accordingly.

We test the scheme also on the Cameraman.bmp image and obtain the results shown in Fig.7.

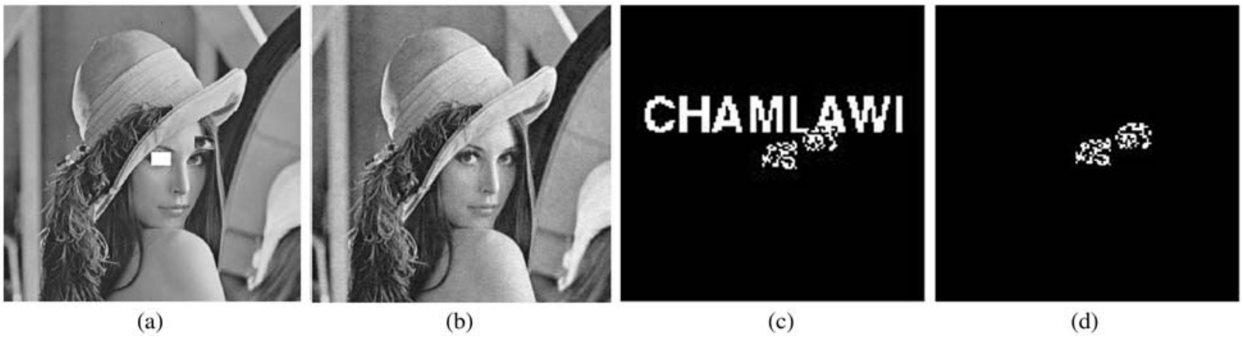


Fig.4. (a) Tampered image. (b) Recovered image. (c) Extracted watermark. (d) Difference between the original and extracted watermark.

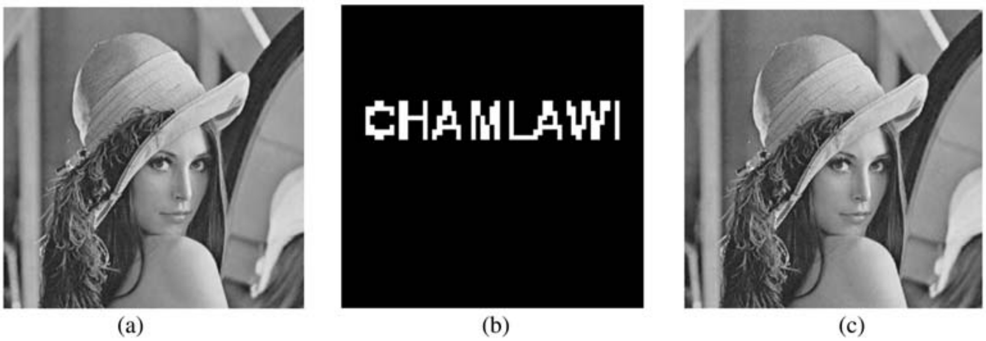


Fig.5. (a) Original Lena image. (b) Binary signature. (c) Watermarked image (PSNR 38dB).

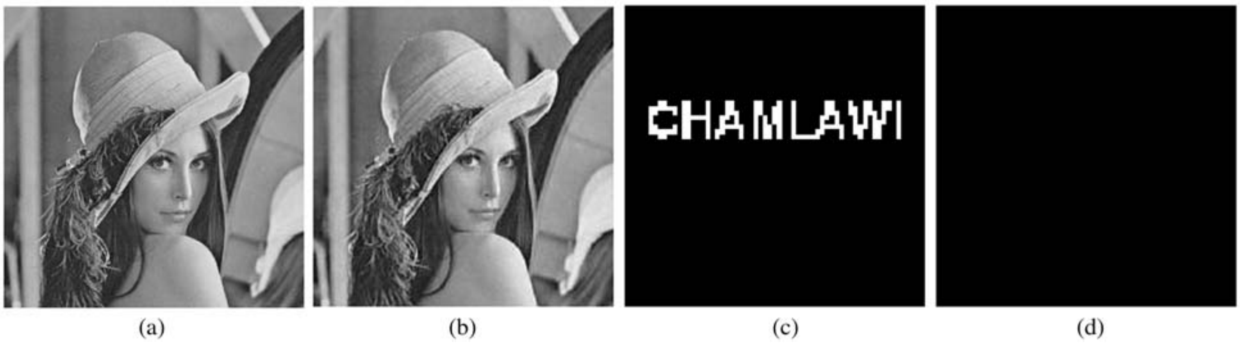


Fig.6. (a) Watermarked image. (b) Recovered image. (c) Extracted binary signature. (d) Difference in original and extracted binary signature.

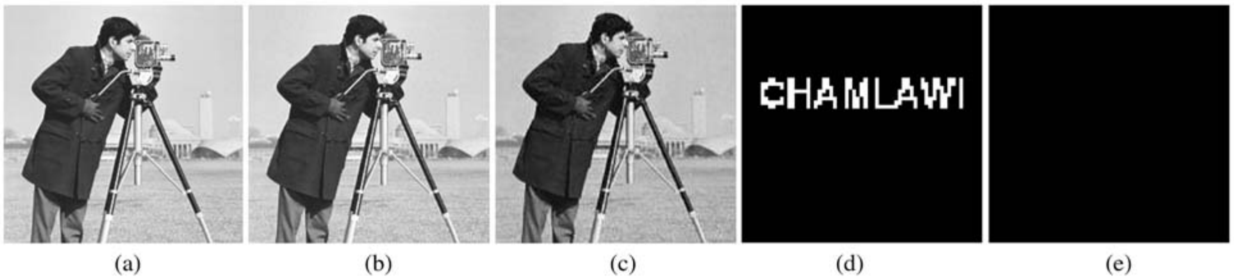


Fig.(7). (a) Original image of Cameraman. (b) Watermarked image, PSNR 38.2dB. (c) Recovered image. (d) Extracted binary image. (e) Difference in binary images.

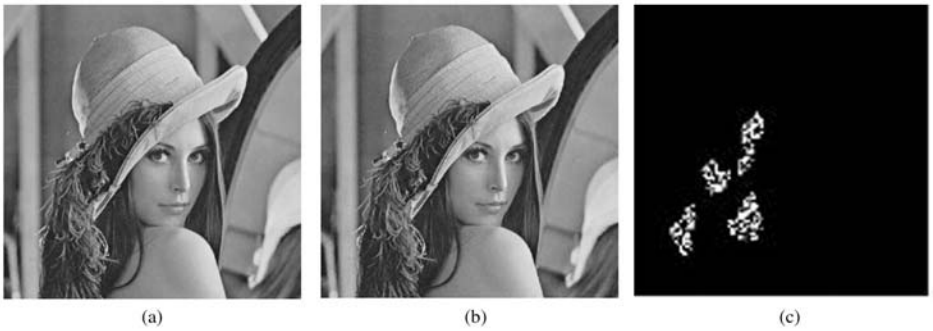


Fig.8. (a) Original image of Lena with bmp format. (b) Watermarked image. (c) Difference image which shows those areas which are tampered maliciously on hairs.

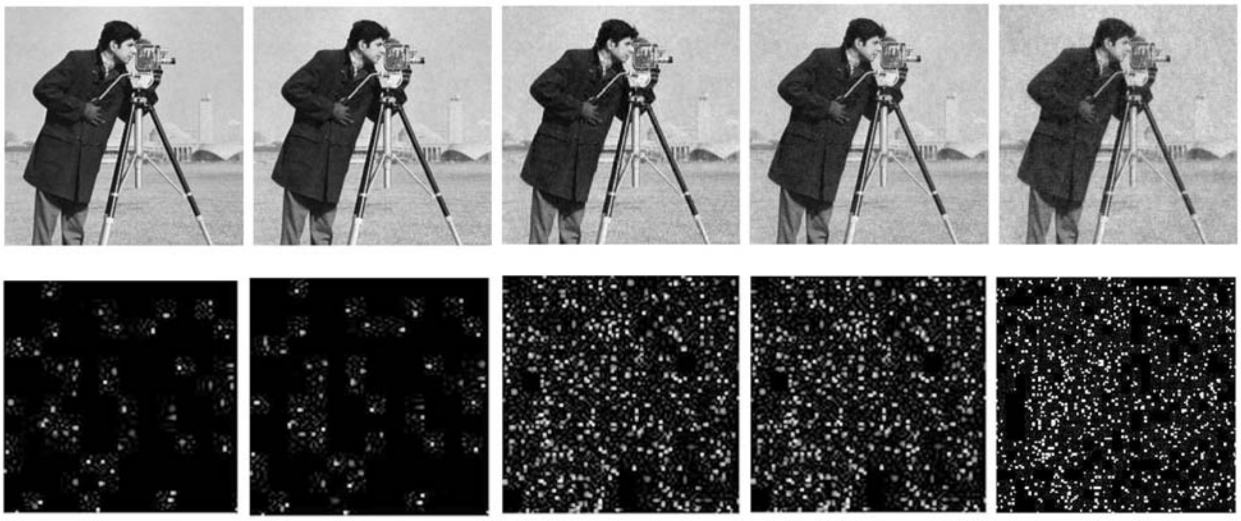
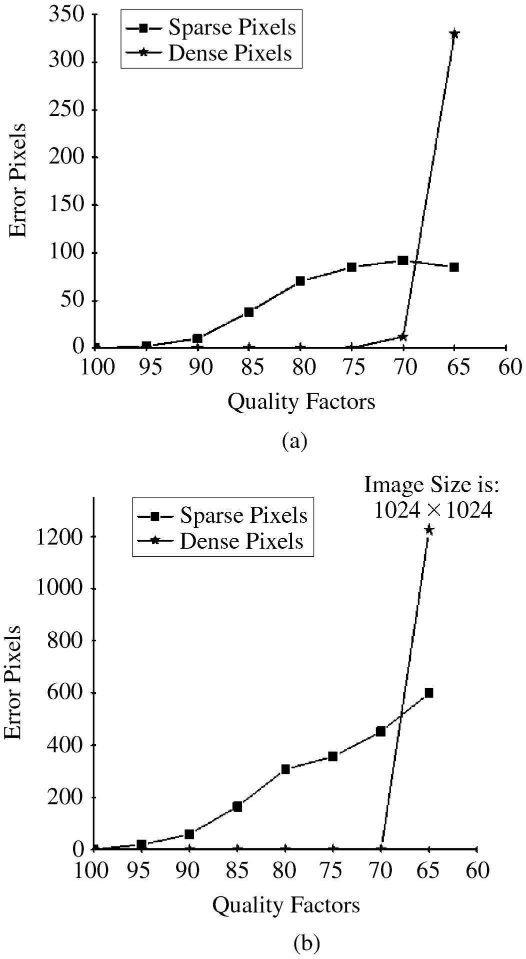


Fig.9. First row shows the recovered images and the second row shows the differences after compressing the watermarked image with quality factors 90, 80, 75, 70 and 60 respectively.

A result obtained after such tampering, which is not visible, is shown in Fig.8. The Lena image is tampered on hairs invisibly. In previous techniques[1], after such type of tampering, the image can be recovered but their scheme is unable to locate the tampered areas. Our proposed scheme recovers the approximated image and locates the tampered area accurately. In previous techniques[1], we observe that, if the image is invisibly tampered or highly compressed then users are left to judge the authenticity of the image by themselves. However, here as in Fig.8, we see that our proposed scheme is able to locate the tampered area accurately.

Fig.9 shows both the estimated image after recovery and extracted binary signatures from the watermarked image, compressed by JPEG at different quality factors (QF). We can see that the proposed scheme can resist as low as 70% JPEG compression while in this case, the quality less than 70% should be considered malicious manipulation and also the recovered image will be degraded. We compress the image of cameraman.tiff using quality factors 95, 90, 80, 75 and 70. When the quality factor is 70 or above then the difference image contains the error pixels but these pixels are not dense pixels but sparse pixels. It means that our scheme survives reasonable compression. If we use the quality factor less than 70 then the number of dense pixels increases in the difference image that shows the malicious manipulation.

In Fig.9, we have two rows. The first row shows the recovered images after compression of different quality factors, while the second row shows the difference in the two binary watermarks.



(a)

(b)

Fig.10. Error pixels versus strength of JPEG compression attack. (a) Image of size 512 × 512. (b) Image of size 1024 × 1024.

The dots on the difference images show the sparse pixels that the image is incidentally tampered, not maliciously. Below 70, dense pixels starts appearing in the difference image (last case of Fig.9).

Fig.10 illustrates the performance of our proposed scheme against JPEG compression attack. Images of size 512 × 512 and 1024 × 1024 are considered. In both cases, it is observed that up to quality factor 70, the number of sparse pixels increases gradually, while on the other hand, the numbers of dense pixels are almost negligible. This shows that the attack is incidental. However, when the quality factor decreases below 70, the number of dense pixels increases sharply, which shows that the attack is malicious (last case of Fig.9).

Table 1 shows the PSNR values of the image digest recovered after the watermarked image has been JPEG compressed with respect to the image digest extracted, when the watermarked image has not undergone any compression. For each PSNR value, the corresponding JPEG quality factor is shown. It should also be noted that when quality factor goes down to 70%, the PSNR is still satisfactory for applications in which data compaction is more important than image detail reconstruction. In our proposed approach, we are utilizing image quality measure, PSNR for two purposes as follows.

1) To analyze the strength of the watermarks in the watermarked image.

2) To check the degradation of the recovered image when the watermarked image is compressed using different quality factors, as shown in Table 1.

Table 1. PSNR of the Extracted Digest Image After JPEG Compression of the Watermarked Image, with Respect to the Digest Image Extracted from a Non-Compressed Watermarked Image

Quality Factor	PSNR
100	35.21
95	33.23
90	30.53
85	28.54
80	27.01
75	25.50
70	24.89
60	22.34

Note: PSNR values are given with the corresponding JPEG quality factor.

5 Performance Comparison with Previous Approaches

We analyze the performance of our proposed approach through experimental results and compare it with the previous approaches[1,11]. Table 2 shows the salient features of our proposed scheme.

Table 2. Salient Features of Our Proposed Scheme

Features	Piva *et al.*[1]	Xiaoyun *et al.*[11]	Proposed Scheme	Supporting Results and Discussions
Recovery	Recovers the estimated image	Cannot recover the image	Recovers the estimated image	Fig.5
Accurate Authentication	Cannot accurately authenticate the image in case of invisible tampering	Accurately authenticates after any type of tampering	Accurately authenticates after any type of tampering	Fig.8
Compression	Survives JPEG compression	Survives JPEG compression	Survives JPEG compression	Fig.9
Localization	If the image is tampered invisibly, then it cannot localize it	It localizes the invisible tampered area accurately	It localizes the invisible tampered area accurately	Fig.8
Security	More secure: uses two secret keys; for scaling, and for scrambling	Not substantial security, only one secret key is used during preprocessing of the binary watermark	More secure: uses three secret keys; for scaling, for permutation, and for preprocessing of the binary watermark	Subsections 2.1 and 2.2

In Table 2, we can observe that the method proposed in [1] can recover the estimated image and localize the tampered area. However, this approach cannot accurately localize the invisible attacks (intentional or unintentional), i.e., it has a crude way to tamper localization. This approach does not use any automatic method for localizing the tampered areas, but just computes the thresholded difference between the recovered image and the watermarked image. Therefore, if the tampering is visible, i.e., cut/copy and paste, then it will give the white pixels, otherwise black pixels on the difference image.

On the other hand, [11] cannot recover the image, but there is an automatic system to detect/localize the tampered areas. In this approach, any type of attack can be detected, either it is malicious or incidental.

In contrast, our proposed approach not only recovers the estimated image as in [1], but can also detect any type of tampering as in [11]. If the image is compressed, which is an incidental attack (Fig.9), or it is tampered maliciously; either visible (Fig.6) or invisible (Fig.8), our proposed approach is able to detect and localize all such attacks accurately.

We also analyzed practically, the effect of including the DC component in the image digest. It has been observed that by including the DC component in image digest; as opposed to [1] where it is approximated by the value of 128, the PSNR value of the recovered image degrades to 38.51 from 38.84.

It is to be noted that watermarking security is still an unsolved problem and therefore, is an active area of research. The security of our first watermark, the image signature, is still weak. However, the embedding of our 2nd watermark, the image digest, is secure through the use of scaling and permutation keys.

In the future, we plan to improve the security aspect of the image signature and analyze the interdependencies of the two watermarks as regards security, robustness and imperceptibility.

6 Conclusion

The proposed scheme is able to distinguish the malicious and incidental attacks and recovers a good estimate of original contents. In this approach, we sacrifice a little bit on the PSNR, which is approximately 38db to 40db for different images, but still the quality of the watermarked image is satisfactory. The technique is highly secure because of inclusion of three private keys at various stages of watermark generation. The proposed scheme also shows efficient authentication for a smallest scale transformation on an image. Embedding of two watermarks makes our proposed scheme more efficient for accurate detection of tampered area and recovery of estimated image. Invisible tamper detection another authentication attribute is achieved in our proposed semi-fragile secured watermarking scheme. The exploitation of machine learning approaches for localization of alterations and improving the quality of the recovered image seems to be perspective[17∼19].

References

[1] Alessandro Piva, Franco Bartolini, Roberto Caldelliy. Self-recovery authentication of images in the DWT domain. *International Journal of Image and Graphics*, 2005, 5(1): 149∼166.

[2] Fei C, Kundur D, Kwong R H. Analysis and design of secure watermark-based authentication systems. *IEEE Transactions on Information Forensics and Security,* 2006, 1(1): 43∼55.

[3] Ingemar J Cox, Gwenaë Doërr, Teddy Furon. Watermarking is not cryptography. *LNCS* 4283, Springer, 2006, pp.1∼15.

[4] Fei C, Kundur D, Kwong R H. Achieving computational and unconditional security in authentication watermarking:

Analysis, insights, and algorithm. In *Proc. SPIE: Security and Watermarking of Multimedia Contents VII*, Vol. 5681, San Jose, California, 2005, pp.697∼708.

[5] Luis Perez-Freire, Pedro Comesana, Juan Ramon Troncoso-Pastoriza, Fernando Perez-Gonzalez. Watermarking security: A survey. *Transactions on Data Hiding and Multimedia Security 1*, 4300, 2006, pp.41∼72.

[6] Hartung F, Kutter M. Multimedia watermarking techniques. In *Proc. IEEE*, USA, 1999, 87(7): 1079∼1107.

[7] Friedman G L. The trustworthy digital camera: Restoring credibility to the photographic image. *IEEE Transaction on Consumer Electronics*, Rosemont, IL, USA, 1993, 39(4): 905∼910.

[8] Xiang Zhou, Xiaohui Duan, Daoxian Wang. A semi-fragile watermark scheme for image authentication. In *Proc. IEEE, 10th International Multimedia Modeling Conference* (*MMM'04*), Brisbane, Australia, 2004, pp.374∼377.

[9] Hua Fiun, Xiuo-Ping Zhang. Fragile watermark based on the Gaussian mixture model in the wavelet domain for image authentication. *IEEE International Conference on Image Processing*, Barcelona, Spain, Vol. 1, 2003, pp.505∼508.

[10] Kurato Maeno, Qibin Sun, Shih-Fu Chang, Masayuki Suto. New semi-fragile image authentication watermarking techniques using random bias and non-uniform quantization. *IEEE Transactions on Multimedia*, 2006, 8(1): 32∼45.

[11] Xiaoyun Wu, Junquan Hu, Zhixiong Gu, Jiwu Huang. A secure semi-fragile watermarking for image authentication based on integer wavelet transform with parameters. In *Proc. the Australasian Workshop on Grid Computing and E-Research*, Newcastle, New South Wales, Australia, Vol. 44, 2005, pp.75∼80.

[12] Ching-Yang Lin, Shi Fu-Chang. Semi-fragile watermarking for authenticating JPEG visual content. *SPIE Security and Watermarking of Multimedia Contents* II, Vol. 3971, 2000, pp.140∼151.

[13] Liu H M, Liu J F, Huang J W, Huang D R, Shi Y Q. A robust DWT-based blind data hiding algorithm. In *Proc. IEEE on Circuits and Systems*, Phoenix-Scottsdale, AZ, USA, Vol. 2, 2002, pp.672∼675.

[14] Dima Pröfrock, Mathias Schlauweg, Erika Müller Richard Wagner. A new uncompressed-domain video watermarking approach robust to h.264/AVC compression. In *Proc. the 24th IASTED International Conference on Signal Processing, Pattern Recognition, and Applications,* Innsburk, Austria, 2006, pp.99∼104.

[15] http://www.math.cuhk.edu.hk/∼rchan/paper/impulse/definitions.html

[16] Meerwald, Uhl A. Watermark security via wavelet filter parameterization. In *Proc. IEEE International Conference on Image Processing*, Thessaloniki, Greece, Vol. 3, 2001, pp.1027∼1030.

[17] Khan A. A novel approach to decoding: Exploiting anticipated attack information using genetic programming. *International Journal of Knowledge-Based Intelligent Engineering Systems*, 2006, 10(5): 337∼347.

[18] A Khan, Anwar M Mirza, A Majid. Intelligent perceptual shaping of a digital watermark: Exploiting characteristics of human visual system. *International Journal of Knowledge-Based Intelligent Engineering Systems,* 2006, 10(3): 213∼223.

[19] A Khan, Anwar M Mirza. Genetic perceptual shaping: Utilizing cover image and conceivable attack information using genetic programming. *Information Fusion, Elsevier Science*, 2007, 8(4): 354∼365.

Rafiullah Chamlawi received his B.Sc and M.Sc. degrees in computer science from University of Peshawar, Pakistan in 1998 and 2000 respectively. He received his M.S. degree in computer system engineering from Ghulam Ishaq Khan (GIK) Institute of Engineering Sciences and Technology, Pakistan in 2006. He is currently pursuing his Ph.D. research in Department of Computer and Information Sciences (DCIS), Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan. His area of research is image processing, digital image watermarking, image authentication and machine learning techniques.

Asifullah Khan received his M.Sc. degree in physics from University of Peshawar, Pakistan in 1996 and his M.S. degree in nuclear engineering from Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan, in 1998. He received his M.S. and Ph.D. degrees in computer systems engineering from Ghulam Ishaq Khan Institute of Engineering Sciences and Technology (GIK Institute), Topi, Pakistan, in 2003 and 2006 respectively. He has more than 10 years of research experience and is working as assistant professor in Department of Computer and Information Sciences at PIEAS. Currently he is pursuing his Post-Doc Research at Signal and Image Processing Lab, Department of Mechatronics, Gwangju Institute of Science and Technology, South Korea. His research areas include digital watermarking, pattern recognition, image processing, genetic programming, data hiding, machine learning, and computational materials science.

Adnan Idris received his B.S. degree in computer software engineering from COMSATS Pakistan and his M.S. degree in computer system engineering from Ghulam Ishaq Khan (GIK) Institute of Engineering Sciences and Technology, Pakistan in 2006. He is currently working as a lecturer in computer science in AJK University, Azad Jammu and Kashmir. His research interest is digital image watermarking, image (color and gray scale) enhancement.