Semi-Fragile Watermarking Scheme based on Feature in DWT Domain

Chaitanya Kommini
Dept of IT,
SVEC, Tirupati
J.N.T.Univeristy, Anantapur
A.P – 517 102, India

Kamalesh Ellanti
Dept of IT,
SVEC, Tirupati
J.N.T.Univeristy, Anantapur
A.P – 517 102, India

E. Harshavardhan Chowdary
Dept of CSE,
SVEC, Tirupati
J.N.T.Univeristy, Anantapur
A.P – 517 102, India

ABSTRACT:
In DWT domain, A novel semi-fragile watermarking is the scheme for the purpose of authenticating the image and also tamper localization. In this scheme, to generate the watermark, we extract the content-based image features in the wavelet domain in particular from the approximation subband. Later, we embed the watermark into the middle subband. Then, the main goals of this scheme i.e., image authentication and tamper localization are achieved by comparing the matrices of extracted content based image with the extracted watermark. It is proven that this proposed scheme is very robust for common content preserving image processing like noise-adding and lossy JPEG compression through experimental results. The impact of this scheme to malicious tamper in images is fragile. The proposed scheme is very accurate in detecting the tamper location.

Keywords:
Semi-fragile Watermarking, Watermark, Tamper localization, Image authentication.

1. INTRODUCTION:
In present days, it is very easy to replicate and modify the digital images as the image processing software is widely known. So, we need to preserve the integrity of the image content. For this, digital watermarking is very useful which is proposed to protect the copyright and integrity of the digital images. This digital watermarking scheme provides authenticity of the original data by hiding the special mark into the multimedia data [1-2]. To verify the authenticity of the original data, a fragile or semi-fragile watermark [3-4] is enough but for protecting the copyright we need a robust watermark. The algorithm used for detecting the tamper must possess certain properties. First one is transparency i.e., the process of embedding the watermark must preserve the quality of the original image and it must be invisible perceptually. The second property is sensitivity i.e., this semi-fragile watermarking must be fragile for content-modifications whereas it must be robust for content-preserving manipulations. The third property is security i.e., the process of embedding the watermark should be secure. In [5], image authentication research papers are given. In [6], semi-fragile watermarking by adaptive gradient partitioning is proposed by Xiang. In [7], semi-fragile watermarking using the image contour as a watermark and embedding it in middle frequency DWT coefficients is proposed by Zhang. The properties of the proposed scheme are: watermarks are made invisible; tampering done to the watermarked image can be detected easily; the altered regions can be located easily.

2. PROPOSED SCHEME:
The proposed scheme for detecting the tampering has four parts: Watermark generation is the initial stage, second section is watermark embedding, third section is watermark extraction and fourth section is image authentication.

2.1 Watermark generation:
From the two-level wavelet decomposition of the original image, we generate the watermark.

The steps are:

Step 1: From the P x Q original image I, we take the two-level DWT.

Step 2: Based on the three wavelet subbands: HL2, LH2, and HH2, we calculate the image feature matrix M.

$$M(i,j) = |HL2(i,j)| + |LH2(i,j)| + |HH2(i,j)|$$

Step 3: Using the feature matrix M and adaptive threshold T, the watermark W is obtained:

$$W(i,j) = \begin{cases} 1, if\ M(i,j) \geq T \\ 0, if\ M(i,j) < T \end{cases}$$

Also, the Adaptive threshold T is calculated from the feature matrix M by distributing the values in M.

$$T = E(M) + \sigma(M)$$

Where $E(M)$ is the calculated mean of the feature matrix M and $\sigma(M)$ is the calculated standard deviation.

Step 4: Watermark Encryption: As we are extracting the watermark publicly, we need encryption for the watermark. For this purpose, first we generate a pseudorandom sequence using logistic map.

$$X_{K+1} = \mu X_K (1 - X_K), \text{where } 1 \leq \mu \leq 4$$

X_0 and μ are used as the secret key K1. The watermark sequence is then changed to the binary mask sequence F. Later, we perform an XOR operation on Watermark W and the obtained mask sequence F. The result is the encrypted watermark $\hat{W}$:$\hat{W} = W \oplus F$

The figure 1 shows the watermark generation section.

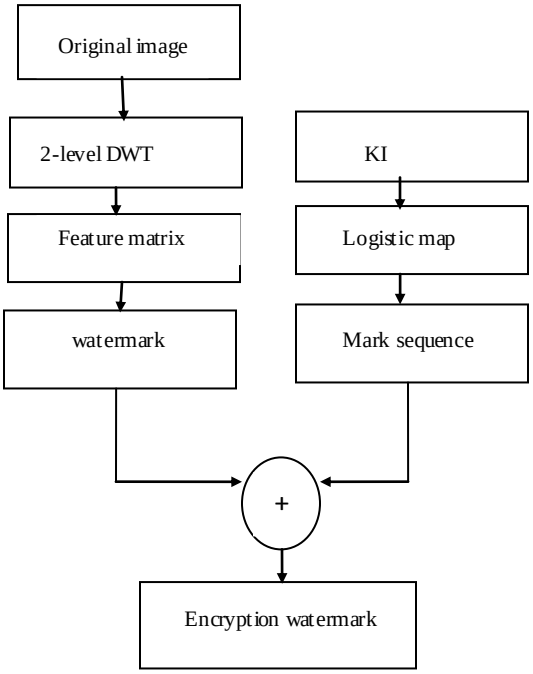


Fig 1: Watermark generation

2.2 Watermark embedding:

The reason for embedding the watermark in the middle frequency subbands: HL2 and LH2 are to have the perceptual invisibility and robust watermark. The steps are:

Step 1: In the initial step, the middle frequency subbands HL1 and LH1 are divided into 2*2 blocks which were non-overlapping. As security is the desired property, to enhance it we select the positions of the embedding blocks of middle frequency subbands in pseudo random way using a secret key k2.

Step 2: Let c11, c12, c21 and c22 be the four coefficients of the block then to embed one bit of watermark, we adjust the c12 and c21 as follows:

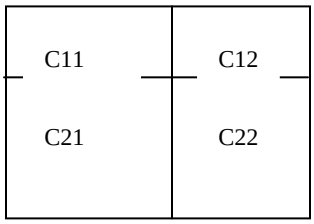


Fig 2: HL1 and LH1 bock to embedd watermark

Two parameters are defined as follows:

$$v1 = max\left\{c12, c21, \frac{(c12 + c21)}{2} + \delta\right\}$$

$$v2 = min\left\{c12, c21, \frac{(c12 + c21)}{2} - \delta\right\}$$

Now, the watermark is embedded is as follows:

$$\begin{cases} c12 = v1, c21 = v2, & if\ wi = 1 \\ c12 = v2, c21 = v1, & if\ wi = 0 \end{cases}$$

The robustness and visual quality of this scheme is guaranteed by the value δ.

Step 3: IDWT is performed which results in watermarked image.

2.3 Watermark extraction:

The step by step procedure for extracting the watermark is as follows:

Step 1: first, we take the 2-level DWT on the watermark image.

Step 2: Now, we need to find the 2*2 blocks in HL1 and LH1 subbands for the purpose of watermark embedding. This is done using secret key K2. The watermark bit wi from each block is obtained as:

$$Wi = \begin{cases} 1, if\ c12 \geq c21 \\ 0, if\ c12 < c21 \end{cases}$$

Step 3: Now the watermark W1 is obtained after all the watermark bits are obtained from the blocks.

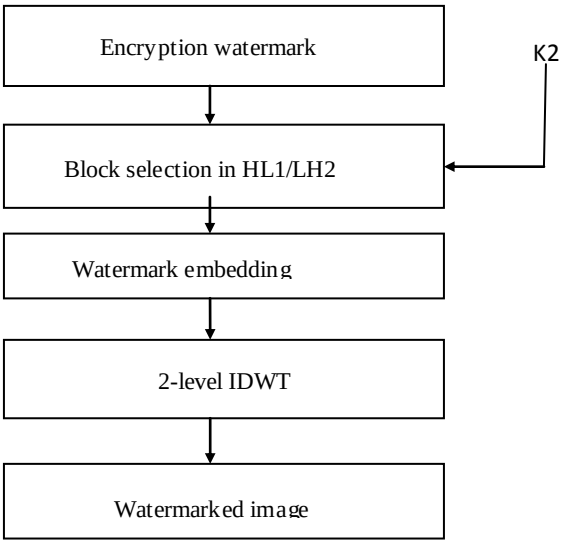


Fig 3: watermark embedding

Step 4: As in the previous section, the mask matrix F is obtained from logistic map using the secret key K1. Later we calculate the watermark Wb as: $W_b = W1 \oplus F$

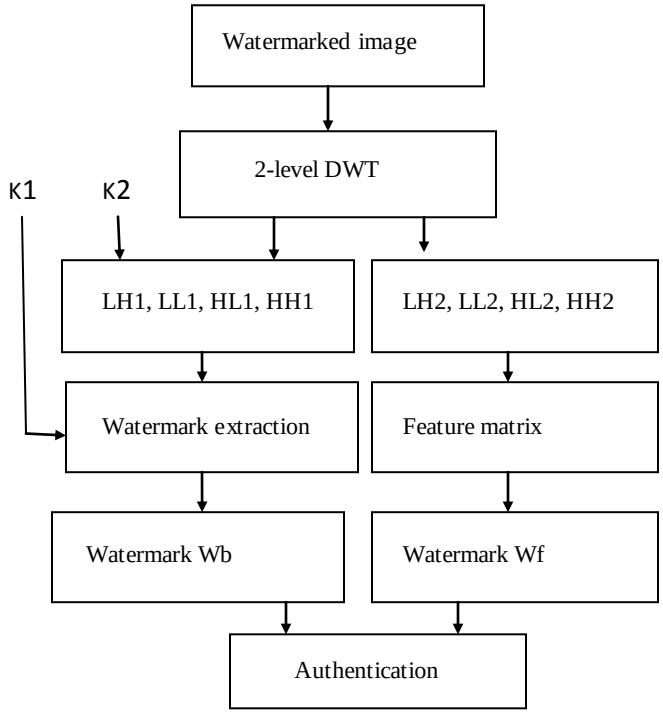


Fig 4: watermark extraction flowchart

Step 5: Now using the subbands HL2, LH2 and HH2, we obtain the extracted watermark Wf. This extraction of Wf is similar to that of Wb.

Step 6: These watermarks Wb and Wf are used for the later section i.e., Image authentication.

2.4 Image Authentication:

We perform the image authentication using two watermarks obtained in the previous sections: Wb and Wf. The tamper matrix Wt is defined as the XOR of Wb and Wf.

$$W_t = W_b \oplus W_f$$

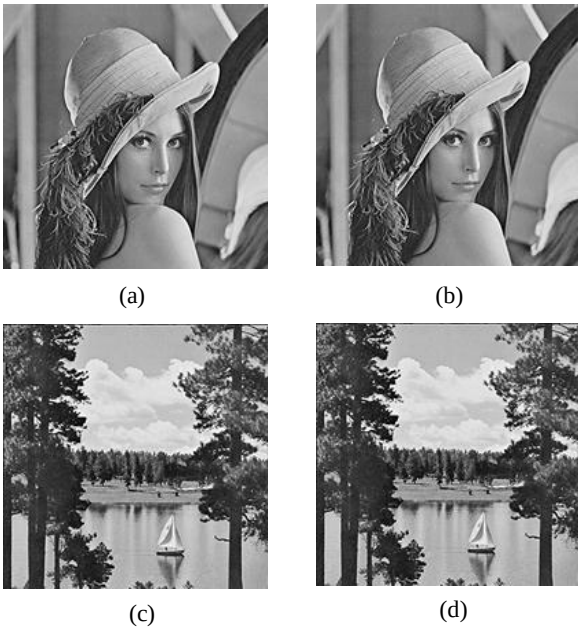
If any malicious attacks are made on the watermarked image, then we can observe that many of the watermark error pixels which have a value of 1 in Wt are clustered in the attacked regions. If any content preserving attacks are made on the watermarked image, then we can observe that the watermark error pixels are spread across Wt. Hence, before authentication, we have to delete all the isolated error pixels in the tamper matrix Wt. We can say an error point as an isolated one if there are no error points in any of its neighbour eight points. Based on the pattern of the filtered matrix i.e., obtained after deleting all the isolated error points, we can clearly distinguish the content preserving attacks from malicious attacks.

3 EXPERIMENTAL RESULTS:

Numerous Experiments are conducted on different standard 8-bit grayscale images and also different kinds of attacks for the purpose of evaluating the performance of the proposed scheme. Let the watermark extracted from blocks in HL1/LH1 be Wb and the watermark extracted from the feature matrix be Wf.

3.1 Quality of watermarked image:

We take the PSNR as the quality metric here. In figure 5, 5(a) and 5(c) shows the original images and 5(b) and 5(d) shows the watermarked images. The PSNR values taken are 34.16dB and 33.16dB.



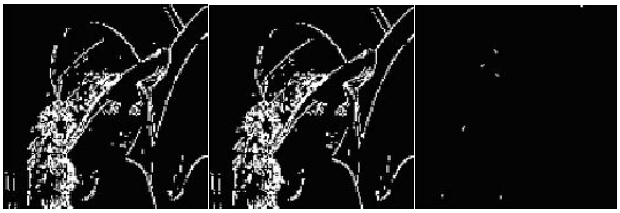
(a) (b)

(c) (d)

Fig 5: Original and Watermarked Images

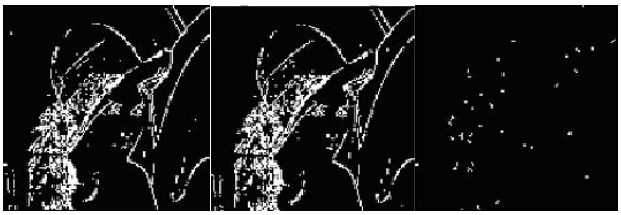
3.2 JPEG COMPRESSION:

JPEG compression is the widely used image operation hence the authentication algorithm should also be robust enough to this compression operation so that it is distinguishable from malicious tamper. Figure 6, 7 and 8 shows the results of JPEG compression with qualities taken as 90, 80, 70 respectively. In the figure, (a) represents the Wf matrix, (b) represents the Wb matrix and (c) represents the authentication result. The authentication algorithm applied here shows that the watermarked image is not tampered.



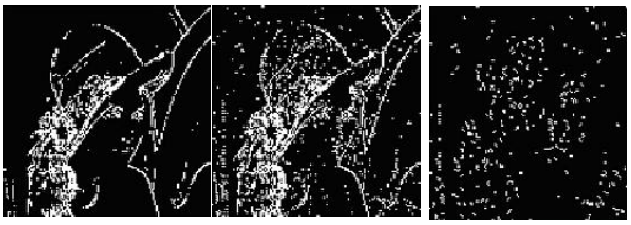
(a) (b) (c)

Fig 6: JPEG Compression, quality=90



(a) (b) (c)

Fig 7: JPEG Compression, quality=80

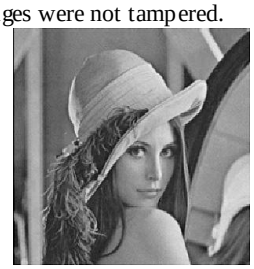


(a) (b) (c)

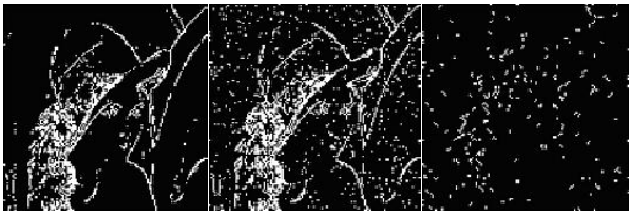
Fig 8: JPEG Compression, quality=70

3.3 ADDING NOISE:

Consider that we added Gaussian noise to the watermarked image with the standard deviation of the noise as σ=0.255 (figure 9). In figure 10, salt-and-pepper noise is added to the watermarked image with the rate of the noise as 0.005. The authentication experiments conducted here showed that the watermarked images were not tampered.



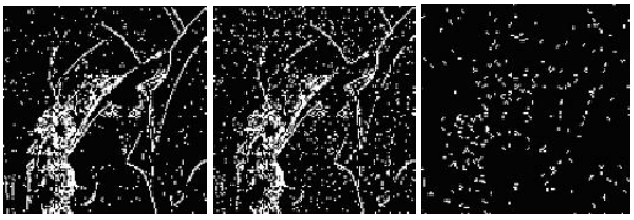
(a) Noise added watermarked image



(b) *Wf* (c) *Wb* (d) result

Fig 9: Gaussian Noise, rate=0.255

(a) Noise added watermarked image

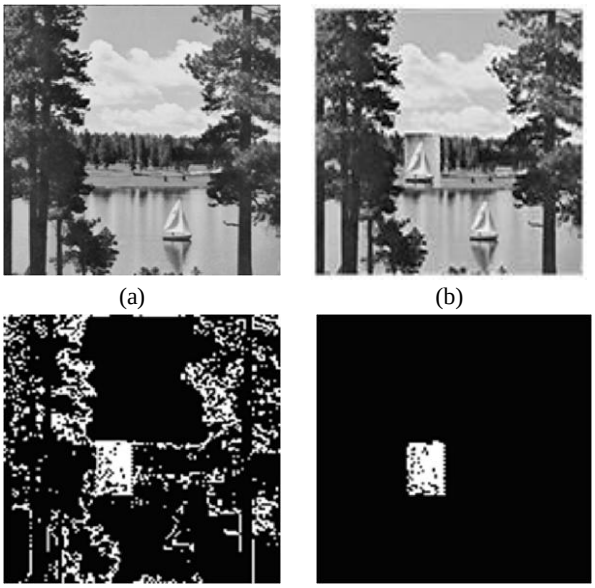


(b) *Wf* (c) *Wb* (d) result

Fig 10: Salt-and-pepper noise, rate=0.005

3.4 TAMPERING ATTACKS:

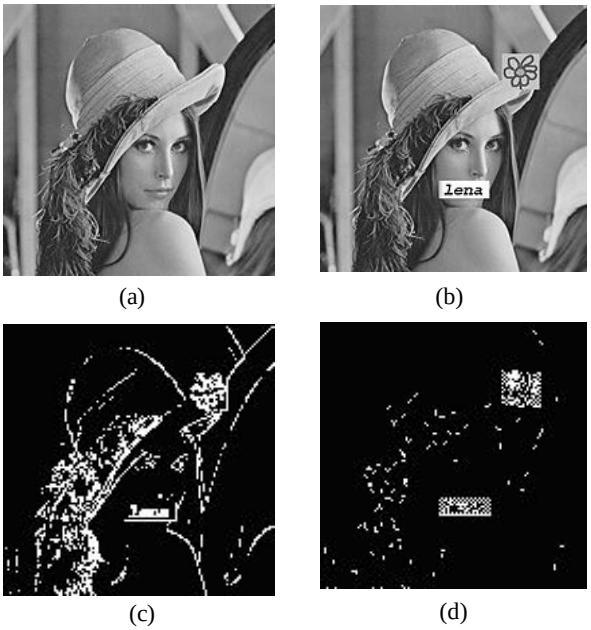
Figures 11, 12 and 13 shows that tamper can be detected and localized accurately. (a) Represents the actual watermarked image, (b) represents the tampered image, (c) represents the Wb matrix which is extracted from tampered image and (d) represents the result.



(a) (b)

(c) (d)

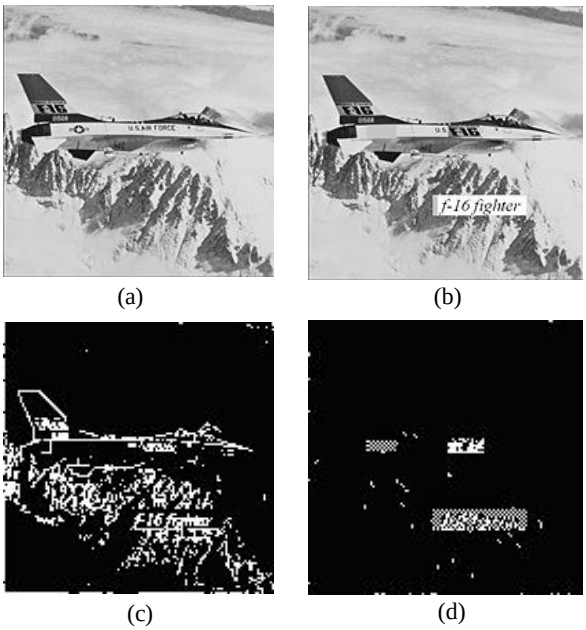
Fig 11: Tampering test of boat



(a) (b)

(c) (d)

Fig 12: Tampering test of Lena



(a) (b)

(c) (d)

Fig 13: Tampering test of fighter

4. CONCLUSIONS:

In this paper, a semi-fragile watermarking scheme is proposed for the purpose of authenticating the image and also localizing the tamper in DWT domain. In the first step, we extract the feature matrix from the approximation subband of DWT. Then we extract the watermark from the feature matrix. Later the generated watermark is embedded in the middle frequency subband. The proposed scheme can distinguish the malicious attack from the content preserving attacks by comparing the extracted image feature matrix from the extracted watermark. Experimental results show that this scheme can effectively localize the malicious tampered regions.

5. REFERENCES

[1] J. Cox, M. L. Miller, and J. A. Bloom, Digital Watermarking, New York: Academic Press, 2002.

[2] Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," IEEE Transactions on Image Processing, vol. 6, pp. 1673-1687, December 1997.

[3] S. Walton, "Authentication for a slippery new age," Dr. Dobb's Journal, vol. 20, no. 4, pp. 18-26, April 1995.

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

[5] J. Wu and F. Lin, "Image authentication based on digital watermarking," Chinese Journal of Computers, vol. 27, no. 9, pp. 1153-1161, September 2004.

[6] X. Wang, "A novel adaptive semi-fragile watermarking scheme based on image content," ACTA AUTOMATICA SINICA, vol. 33, no. 4, pp. 361-366, 2007.

[7] D. Zhang and Z. Pan, "A contour-based semi fragile image watermarking algorithm in DWT domain," Proc. ETCS, vol. 3, pp. 228- 231, 2010.