PERFORMANCE ANALYSIS OF A NOVEL IMAGE CONTENT AUTHENTICATION SCHEME USING MULTIPLE WATERMARKS IN DUAL DOMAINS.

***Jayashree Nair**
Associate Professor, AIMS Institutes, Bangalore

****Padma T**
Professor, Sona College of Technology, Salem

ABSTRACT

This paper presents a multiple watermark based semi-fragile watermarking scheme that tries to balance the requirements of an effective authentication scheme ensuring imperceptibility, tamper detection and localization of tampered areas. The proposed scheme uses multiple watermarks to enable authentication at multiple levels in the wavelet domain. The scheme is flexible and permits the user to decide on the number of watermarks to be embedded based on the nature of the application. It can also accurately locate the tampered locations.

Keywords : *Content authentication, authentication watermark , recovery watermark, DWT, tamper localization, robustness, feature vector.*

1. INTRODUCTION

Image authentication is the process of verifying and validating the integrity of watermarked data. It is also the act of confirming if the image is credible or not. Semi-fragile watermarks for image authentication have been proposed and designed in the spatial and transform domains. Spatial domain techniques [1][12] exploit the statistical properties of the pixels of the image to embed the watermark but are normally fragile. Transform domain techniques like Discrete Fourier Transform (DFT) [13] and Discrete Cosine Transform (DCT) [2] [5][10] exploit the frequency properties of the image to ensure robustness of the watermark, but they lack spatial information. Discrete Wavelet Transform (DWT) [3][6][14] exploits the spatial-frequency properties of the image to imperceptibly embed the watermark.

II. RELATED WORK

Dual watermarking schemes uses two watermarks [2][3][9], usually embedded in mutually exclusive domains to achieve image authentication. In the schemes proposed by Yuping et. al., in [4], [7], two watermarks were generated from the low-frequency bands and embedded into the high-frequency bands, one for detecting the intentional content modification and indicating the modified location and another for recovering the image. Chamlawi et al. [14] addressed a secure semi-fragile watermarking scheme for image authentication and recovery based on integer wavelet transform based on embedding two watermarks namely a binary signature and an image digest. The binary signature is embedded in the LL_3 sub-band and a compressed version of original image is generated as the image digest which is embedded in the HL_2 and LH_2 sub-bands and offers high degree of robustness against JPEG compression. Qi et.al., in [16], proposed a scheme where content-based image features from the approximation sub-band in the wavelet domain are extracted to generate two complementary watermarks, one to detect manipulations and the other to localize tampered regions. Both watermarks are embedded in the high-frequency wavelet domain to ensure the watermark invisibility. In [11], the semi-fragile watermark is designed from low-frequency band of wavelet-transformed image and is embedded into the high-frequency band by exploiting the Human Visual System (HVS). The robustness to mild modification such as JPEG compression, channel additive white Gaussian Noise (AWGN) and fragility to malicious attack are analyzed. In [8], a tamper detection and retrieval scheme is proposed. Special characteristic values of the low-frequency sub-band are embedded in the middle frequency sub-bands. The embedded data with a digital signature and a public key are used to prove the authenticity of the image. Recovery with visually acceptable quality has also been achieved. Woo et al. [3] use a down scaled version of the host image as the content based watermark where a quantization function maps the identified DWT coefficient to its binary equivalent. A random key is used to identify four MSB's each in the horizontal and vertical sub bands of the 2nd level 2D DWT of the LL_1 sub band. During authentication, the embedded watermark is extracted and compared with a down sampled version of the original host image.

Dual watermarks offer a kind of backup in case of situations of false alarm that is triggered when the authentication results fail to appropriately diagnose manipulations.

This paper presents a blind self authenticating watermarking scheme in dual domains of DCT and DWT. Multiple watermarks are generated from the image and embedded at different levels of the image and thus no external watermarks are used. The scheme is practical and blind as it does not require the original or watermarked image as a reference for authentication but retrieves a quantized and down sampled approximation of the original image for visual verification.

III. PROPOSED SCHEME

Multiple watermarks are derived from features of the image. Authentication Watermark W_A is generated for every pair of sub blocks and embedded in the HL_1 sub band obtained after 1st level DWT of the image. Recovery Watermark W_R is an approximated and quantized version of the original image embedded in the LH_1 sub bands. Multiple copies of W_R are generated and is used for visual authentication. The steps are elaborated in the forthcoming sections.

Generation and embedding of Authentication Watermark W_A

Image I is divided into non-overlapping m* m blocks and DCT applied to each block. Pairs of blocks are formed according a predetermined randomizing function and for each pair of blocks, n low frequency DCT coefficients, including the DC coefficient and n-1 low frequency AC coefficients, from each block p and q are considered to generate the feature vector for the block pair. The feature vector FV for a pair of blocks is computed as per Lin's Model [8] which is based on the relationship between corresponding pair of coefficients that preserves the content invariant features in the presence of compression and mild noise. The Majority bit M_b for each pair of blocks is extracted from the FV and concatenated to generate the content based authentication watermark W_A. The generated watermark is embedded into the horizontal and vertical detail sub bands as follows:

First level DWT is applied to the host image I to decompose it to the approximate sub band *LL* and detail sub bands – *HL, LH* and *HH* as in Figure 1. 2nd level DWT is applied to the HL sub band to obtain the HHL_2 and HLH_2 sub bands where W_A will be embedded as shown in Figure 2.

For the corresponding positions *(i,j)* of the selected pair of blocks, both HHL_2 and HLH_2 sub bands, as determined by the PQ sequence [17], ratio of the coefficients are evaluated as:

$$R(i,j)=(sgn)\ HHL_2\,(i,j)/HLH_2\,(i,j) \qquad (1)$$

This vector will be the side information to be shared with the authenticator in a secure manner for authentication of the image.

The watermark is embedded by modifying the coefficients as:

$$if\ \ M_b = 1, \begin{cases} HHL_2 = HHL_2 * \alpha \\ \quad and \\ HLH_2 = HLH_2/\alpha \end{cases}$$

$$if\ \ M_b = 0, \begin{cases} HHL_2 = HHL_2/\alpha \\ \quad and \\ HLH_2 = HLH_2 * \alpha \end{cases}$$

where α is the watermark strength factor and can be experimentally determined. A value of α=1.2 gives good imperceptibility in the experiments conducted.

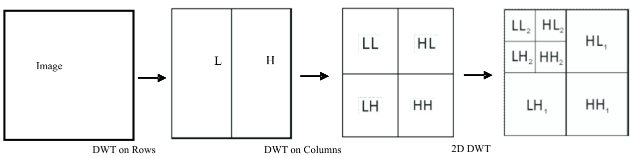


Figure 1: DWT decomposition of Image

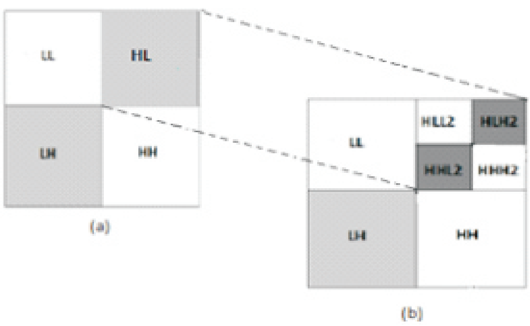


Figure 2: (a) 1st level 2D DWT of image (b) 2[nd] level 2D DWT of HL_1 sub band indicating embedding locations HHL_2 and HLH_2

B. Generation and embedding of Recovery Watermark W_R

Recovery Watermarks, W_R are generated from second level DWT decomposition of the *LL* sub band to obtain a coarse representation LL_2 of the image. The coefficients of LL_2 sub band are then suitably quantized using Dither Modulation [18] to decrease the obtrusiveness of the coefficients. The quantized coefficients form the Recovery Watermark W_R and are embedded in the selected coefficients of LH sub band by replacing five LSBs of the selected coefficients with the quantized binary equivalent of W_R . Figure 3 depicts the quantized approximation of the image and the sub band after embedding the quantized approximation.

Inverse DWT is applied to get the watermarked image W_M.

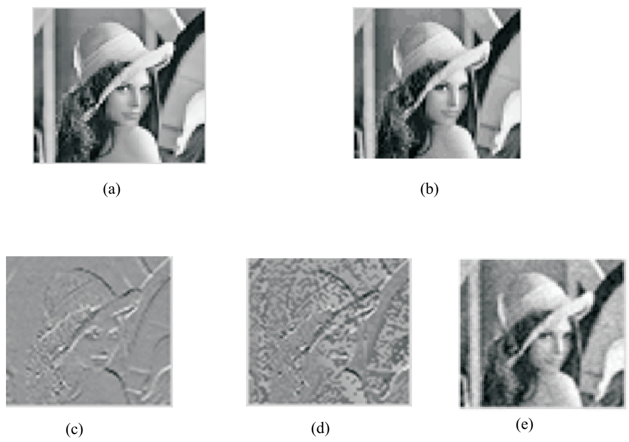


Figure 3: a) Coarse approximation of Recovery Watermark after first level DWT of image b) Quantized approximation of the Recovery Watermark, W_{VA} c) Identified DWT sub band of Lena image to embed the Recovery Watermark W_{VA} d) Sub band after embedding W_{VA} e) Extracted Recovery watermark at the time of Authentication

Table 1 Quality metrics of Watermarked image after embedding – only W_A, only W_R and after embedding both W_A and W_R for embedding strength α = 1.2

Image	Only W_A			Only W_{VA}			Both W_A and W_R		
	PSNR	SSIM	PCC	PSNR	SSIM	PCC	PSNR	SSIM	PCC
Lena	67.48	1	1	52.1	0.94	0.998	53.39	0.94	0.998

From the results in Table 1, it is observed that embedding the Visual Authentication watermark causes a fair degradation in the quality of the image and is comparable to the quality after embedding both the watermarks *WA* and W_{VA}. Depending the nature of the application, if the emphasis on quality is non-negotiable, then the watermarking scheme need embed only W_A

C. Visual Authentication using W_R

Visual observation is a basic step in any authentication process. The scheme is blind and hence the original image need not be available at the authentication end. The approximated watermark W_R is used as a representation of the original image to compare with the received image. To extract the estimated image, the reverse procedure of the Recovery Watermark generation and embedding is performed. The corresponding *LH* sub band is selected and five LSBs from the selected coefficients are extracted. The extracted bits are used to reconstruct the quantized coefficient values using the same quantization table used in the watermark generation phase.

D. Authentication of watermarked image

Authentication of the query image is carried out on a block by block basis by comparing the generated and extracted watermarks for each block. The Authentication watermark W_A^* is extracted by a procedure similar to the watermark generation and insertion procedure in section III (A). The ratio of coefficients R~ for the received image is also evaluated using Equation 1. The Majority bit $M_b^{\sim}$ embedded in the received watermarked image for each block pair is extracted using the relationship

$$M_b^{\sim} = \begin{cases} 1 & if\ R^{\sim}/R > 0 \\ 0 & otherwise \end{cases}$$

The string of majority bits of each block pair will give the extracted Authentication Watermark $W_A^{\sim}$.

The received image is authenticated by correlating the generated watermark W_A^* and the extracted watermark $W_A^{\sim}$. If the integrity is verified, then the watermarked image can be reversed back to a better approximation of the original image using the reverse of the procedure in section III (B).

IV. Experimental Results

The authentication scheme described in this paper is implemented in Matlab 7.10.0.5 (R2011a) environment.

A. Imperceptibility of Watermarked images

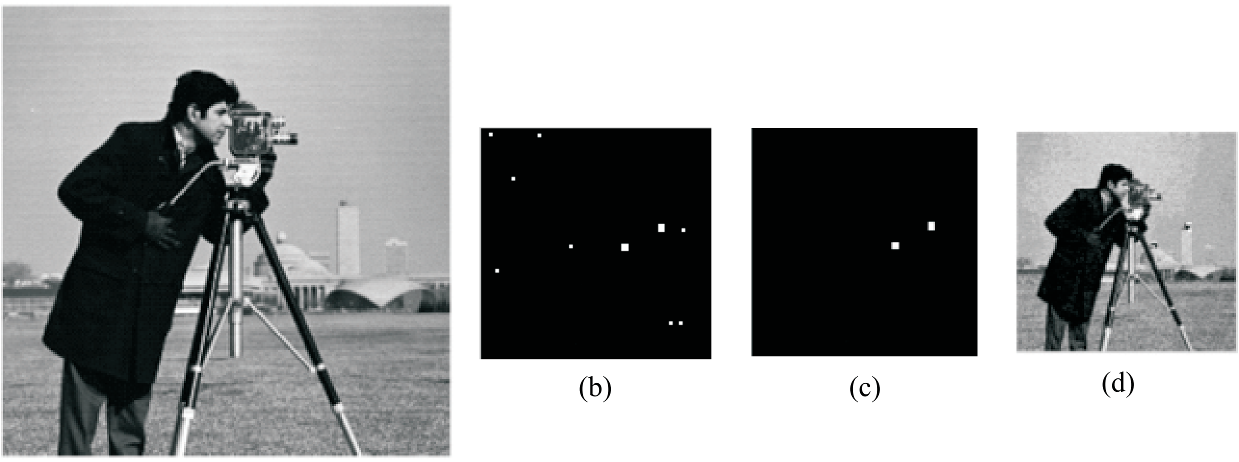
The choice of embedding the authentication Watermark W_A or Recovery watermark W_R or both can be decided based on the requirement of the application. The embedding of the Recovery Watermark W_R slightly reduces the quality of the watermark but is still above acceptable limits.

Figure 4: Watermarked Images after watermark embedding – a) only W_A b) only W_{VA} and c) after embedding both W_A and W_{VA} for embedding strength α = 1.2.

Quality of the watermarked images after embedding only W_A, only W_R and both W_A and W_R are shown in Figure 4. The Peak Signal to Noise Ratio (PSNR) of the images watermarked with only W_A are in the range 63- 70 and after embedding both W_A and W_R are in the range 46 – 55dB. A PSNR of 30dB and above indicates good quality of the watermarked image.

B. Tamper Detection

As a part of the watermarking scheme, features are extracted from each unique pair of sub blocks selected randomly from the LL sub band after first level DWT of the image. During authentication, the generated and extracted feature vectors are correlated to determine the integrity of the image. In case of mismatch, both the sub blocks involved in feature vector extraction are identified as manipulated. Hence the additional locations marked as tampered in Figure (c). By calculating the percentage of difference of the identified pair of blocks, the manipulated block is identified and the other pair is designated as authentic. Figure (d) represents the image after reversing the identified pair as authentic.



(b) (c) (d)

(a)

Figure 5: a) Watermarked image of Cameraman that is altered c) Detection of possible tampered locations d) Localization of the tampered locations e) Approximation of Recovery Watermark extracted from the tampered image performance analysis

V. PERFORMANCE ANALYSIS

The features and performance of the proposed scheme are compared with dual watermarked based peer schemes proposed in [3] and [13] and is illustrated in Table 2.

Table 2 Performance analysis of proposed scheme with peer schemes

Feature	Woo et al.'s scheme [3]	Chamlawi et al.'s scheme [14]	Proposed scheme[79], [80]
Quality metrics - PSNR	≈41 dB	≈39 dB	≈60 dB
Tamper detection	Non Blind - By difference with test image	Non Blind - By difference with test image	Blind – By Correlation
Localization	Yes	Pixel level	4 × 4 Block level
Watermark type for recovery	Down scaled version of original image	JPEG compressed DCT of image	Quantized down scaled version of original image
Randomizing watermark embedding locations	Using secret key	Using secret key	PQ Sequences [17]

VI. CONCLUSION

This paper describes an authentication scheme that uses multiple watermarks, the Authentication watermark and Visual Authentication watermark, one to authenticate the watermarked images and the other to reinforce the authentication in the DCT-DWT domains. The scheme ensures good quality of the watermarked images as the frequency – temporal properties of DWT are utilized to embed the multiple watermarks. The scheme also offers good tamper localization capabilities. The choice of whether to embed one or both the watermarks is open and be decided based on the application.

VII. REFERENCES

[1] Hu M., Lou D. and Chang M., "Dual-Wrapped Digital Watermarking Scheme for Image Copyright Protection," Computers & Security, Vol.26, pp.319-330, 2007.

[2] Lin, C.Y. and S.F. Chang, "SARI: Self-Authentication-and-Recovery Image Watermarking System," ACM Multimedia, Ottawa, Canada: ACM Press. pp.628-629, 2001

[3] Woo Chaw Seng1 et.al., "Semi-Fragile Watermark with Self Authentication and Self Recovery, " Malaysian Journal of Computer Science, Vol. 22(1), pp 64-84, 2009.

[4] Yuping, H., Guangjun, G.: Watermarking-based authentication with recovery mechanism. In: 2nd International Workshop on Computer Science and Engineering. doi:10.1109/WCSE. 2009.856 (2009)

[5] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper proofing and authentication," Proceedings of the IEEE, vol. 87, no. 7, pp. 1167–1180,1999.

[6] C. S. Lu and H. Y. M. Laio, "Multipurpose watermarking for image authentication and protection," IEEE Transactions on Image Processing, vol. 10, no. 10, pp. 435-439, 2001

[7] Hui, L., Yuping, H.: A wavelet-based watermarking scheme with authentication and recovery mechanism. In: International Conference on Electrical and Control Engineering (ICECE). doi:10.1109/iCECE. 2010.86 (2010)

[8] Chen, T.S., Chen, J., Chen, J.G.: Tamper detection and retrieval technique based on JPEG2000 with LL subband. In Proceedings of IEEE International Conference OD Networking, Sensing & Control, Taipei, Taiwan (2004)

[9] J. C. Patra, J. E. Phua and C. Bornand, "A novel DCT domain CRT-based watermarking scheme for image authentication surviving JPEG compression," Digital Signal Processing, vol. 20, no. 6, pp. 1597–1611, 2010.

[10] X. Qi and X. Xin, "A quantization-based semi-fragile watermarking scheme for image content authentication," Journal of Visual Communication and Image Representation, vol. 22, no. 2, pp. 187–200, 2011.

[11] Tsai, M.J., Chien, C.C.: A wavelet-based semi-fragile watermarking with recovery mechanism. In: IEEE International Symposium on Circuits and Systems, ISCAS 2008. doi:10.1109/ISCAS.2008. 4542097 (2008)

[12] De Rosa A., Barni M., Bartolini F., Cappellini V. and Piva A., "Optimum Decoding of Non-additive Full Frame DFT Watermarks", Proceedings of the 3rd Workshop of Information Hiding, pp.159-171,1999.

[13] Parthasarathy A. K. and Subhash Kak, "An Improved Method of Content Based Image Watermarking." IEEE Transactions on Broadcasting, Vol.53, No.2, pp.468-479, 2007.

[14] Chamlawi R., Khan A., Idris A. and Munir Z., "A Secure Semi-Fragile Watermarking Scheme for Authentication and Recovery of Images based on Wavelet Transform," World Academy of Science, Engineering and Technology, Vol.23, pp.49-53, 2006.

[15] Qi, X., Xin, X., Chang, R. "Image authentication and tamper detection using two complementary watermarks," 16th IEEE International Conference on Image Processing (ICIP). doi:10.1109/ICIP.2009. 5413681, 2009.

[16] A Chen, B., Wornell, G.W., "Digital watermarking and information embedding using dither modulation," Proceedings of the IEEE workshop on Multimedia Signal Processing, Redondo Beach, CA, December, 1998.Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 198.

[17] J. S. Pillai, T. Padma. "The analysis of PQ sequences generated from continued fractions for use as pseudorandom sequences in Cryptographic Applications." In Proc. Advances in Intelligent Systems and Computing, Springer, Vol. 394, Dec. 2015, pp. 633-644.