Semi Fragile Watermarking for Content based Image Authentication and Recovery in the DWT-DCT Domains

Jayashree Pillai[1] and Padma Theagarajan[2]
[1]Department of Computer Science, Acharya Institute of Management and Sciences, India
[2]Department of Computer Applications, Sona College of Technology, India

Abstract: *Content authentication requires that the image watermarks must highlight malicious attacks while tolerating incidental modifications that do not alter the image contents beyond a certain tolerance limit. This paper proposed an authentication scheme that uses content invariant features of the image as a self authenticating watermark and a quantized down sampled approximation of the original image as a recovery watermark, both embedded securely using a pseudorandom sequence into multiple sub bands in the Discrete Wavelet Transform (DWT) domain. The scheme is blind as it does not require the original image during the authentication stage and highly tolerant to Jpeg2000 compression. The scheme also ensures highly imperceptible watermarked images and is suitable in application with low tolerance to image quality degradation after watermarking. Both Discrete Cosine Transform (DCT) and DWT transform domain are used in the watermark generation and embedding process.*

Keywords: *Content authentication, self authentication, recovery watermark, DWT, PQ sequence.*

Received May 29, 2015; accepted February 22, 2016

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 [5, 8] 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) [4] and Discrete Cosine Transform (DCT) [7, 11, 15] exploit the frequency properties of the image to ensure robustness of the watermark, but they lack spatial information. Discrete Wavelet Transform (DWT) [1, 9, 14, 19] exploits the spatial-frequency properties of the image to imperceptibly embed the watermark.

Semi fragile watermarks are desired for image authentication to ensure that the watermark is tolerant to incidental operation that arise during the regular storage and transmission of media like compression and noise that affects the entire image but at the same time should be intolerant to malicious manipulation that alter the content or meaning of the image. A number of semi fragile techniques [6, 7, 9, 10, 12, 13, 15, 18] have been researched and addressed in literature.

This paper proposes a blind, semi fragile, self authenticating watermarking scheme in dual domains of DCT and DWT, which is highly robust to Jpeg2000 compression and mildly robust to Jpeg compression. The scheme is also tolerant to most incidental noises that happen during storage and transmission of the media. The watermark is the content based feature vector extracted from the image and secured using a pseudorandom sequence generated using novel number theoretic concepts of irrational numbers and continued fractions [16]. The scheme is practical 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. Once identified as authentic, the watermarked image can be partially restored to its original form.

2. Related Work

Dual watermarking schemes use two watermarks [10, 11, 19], usually embedded in mutually exclusive domains to achieve image authentication. Chamlawi *et al.* [2] addressed a secure semi-fragile watermarking 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 *LL3* sub-band and a compressed version of original image is generated as the image digest and is embedded in the HL_2 and LH_2 sub-bands and offers high degree of robustness against JPEG compression up to 70%. Qi *et al.* [17], propose a content-based image features from the approximation sub-band in the wavelet domain are

extracted to generate two complementary watermarks, one is to detect any changes after manipulations and the other is used to localize tampered regions. Both watermarks are embedded into the high-frequency wavelet domain to ensure the watermark invisibility.

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.

3. The Proposed Scheme

In this scheme, two watermarks are used-the Authentication Watermark W_A and Recovery Watermark W_R. W_A is used to detect incidental or malicious tampering and is embedded in the HL_1 sub band, where as W_R is the approximated version of the original image and is used for visual authentication.

3.1. Generation and Embedding of Authentication Watermark W_A

3.1.1. Generation of W_A

The DCT of LL_1 sub band, as shown in Figure 1 is considered to generate the feature vector which is scrambled using the PQ sequence, generated in [16], to get the Authentication Watermark W_A Multiple copies of W_A are embedded in the horizontal and vertical sub bands obtained after further DWT decomposition of the HL_1 sub band. The watermark generation procedure can be summarized as follows:

1. Apply first level DWT to host image 1 resulting in the approximate sub band LL_1 and detail sub bands –HL_1, LH_1 and HH_1.
2. The LL_1 subband is divided into non-overlapping m^*m blocks and DCT applied to each block.
3. Blocks are placed in two disjoint groups A and B based on a secret key and pairs of blocks (p, q), are formed based on secret function, using one from group A and the other from group B
4. For each pair of blocks (p, q),

a) 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 $FV_{pq}(v)$ is computed as:

$$FV_{pq}(v) = \begin{cases} 1 & if\ DCT_p(v) \geq DCT_q(v) \\ 0 & if\ DCT_p(v) \prec DCT_q(v) \end{cases} \quad (1)$$

Where v=1..n and v is the corresponding pair of coefficients, one from block p and the other from block q.

b) Scramble $FV_{pq}(v)$ by exclusive OR-ing with the corresponding Key Vector KV generated from the PQ sequence [416] to get the scrambled feature vector $SFV_{pq}(v)$.

Repeat Step 4 for all the pairs of blocks to obtain the SFV for each pair of blocks.

extract the Majority bit M_b for each pair of blocks from the SFV and concatenate to generate the content based authentication watermark W_A.

3.1.2. Embedding of W_A

The generated watermark is embedded into the horizontal and vertical detail sub bands as follows:

1. Apply DWT to the $HL1$ sub band to obtain the HHL_2 and HLH_2 sub bands where W_A will be embedded.
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, evaluate the ratio of the coefficients

$$R(i,j) = (sgn)\ HHL_2(i,j)/HLH_2(i,j) \quad (2)$$

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

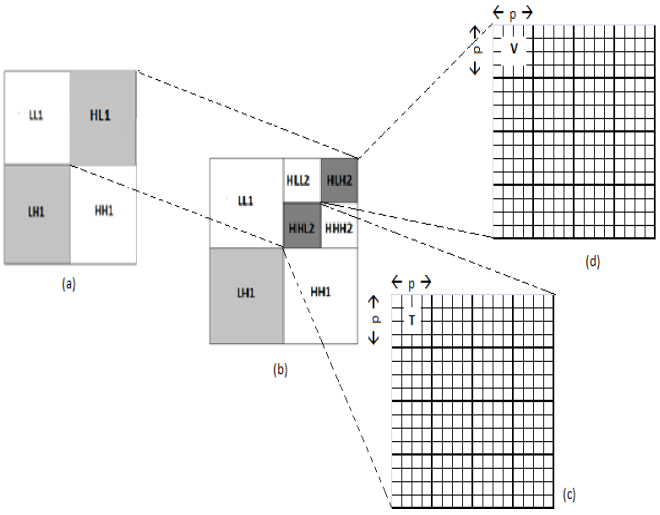

Figure 1. Watermark embedding locations (a) 1st level 2D DWT of image (b) 2nd level DWT of HL_1 sub band indicating embedding locations (c) and (d) are sub blocks of $HHL2$ and HLH2 sub band of size p used to embed the Authentication Watermark W_A.

3. Modify the amplitude of the corresponding coefficients of the HHL_2 and HLH_2 sub bands to embed the watermark:

$$if\ M_b = 1, \begin{cases} HHL_2 = HHL_2 * \alpha \\ and \\ HLH_2 = \frac{HLH_2}{\alpha} \end{cases} \quad (3)$$

$$if\ M_b = 0, \begin{cases} HHL_2 = HHL_2/\alpha \\ and \\ HLH_2 = HLH_2 * \alpha \end{cases} \quad (4)$$

Where α is the watermark strength factor and can be experimentally determined. A value of a=0.9 gives good imperceptibility in the experiments conducted. The ratio $R'(i, j)$ after embedding M_b will increase if M_b=1and decrease if M_b=0

4. Apply inverse DWT to get the watermarked image W_m.

3.2. Generation and Embedding of Recovery Watermark W_R

3.2.1. Generation of W_R

The recovery watermark is generated by second level DWT decomposition of the LL_1 sub band to obtain a coarse representation $LL3$. The coefficients of $LL3$ sub band are then suitably quantized using Quantized Index Modulation [3] to decrease the obtrusiveness of the coefficients and represent it uniformly by dividing the entire range of coefficients from the smallest $LL3_{\min}$ to the largest $LL3_{\max}$ into various bins as shown in Table 1 and Δ is the bin size.

Table 1. Quantization table.

Bin no	Low value	High value
b_1	$LL3_{\min}$-Δ	$LL3_{\min}$
B_2	$LL3_{\min}$	$LL3_{\min}$+Δ
B_3	$LL3_{\min}$+Δ	$LL3_{\min}$+2Δ
.	.	.
.	.	.
$b_{n-1}b_{n-1}$	$LL3_{\max}$-Δ	$LL3_{\max}$
b_n	$LL3_{\max}$	$LL3_{\max}$+Δ

The value of Δ is appropriately selected to ensure a good approximated image and at the same time a good quality of the watermarked image. A large Δ gives better quality of the watermarked image but a heavily quantized recovery watermark where as a small Δ gives a better approximation of the recovery watermark at the cost of quality of the watermarked image and is given by

$$\Delta = (LL3_{\max} - LL3_{\min})/2^{k} \quad (5)$$

Where k is the no of bits to represent each block. On identifying the bin number bl, the coefficient $LL3(i, j)$ is quantized to a value given by of the corresponding bin number

$$(highvalue_{bl} - lowvalue_{bl})/2 \quad (6)$$

The quantized coefficients form the Recovery Watermark W_R and the overall scheme for generating and embedding the dual watermarks is illustrated in Figure 2

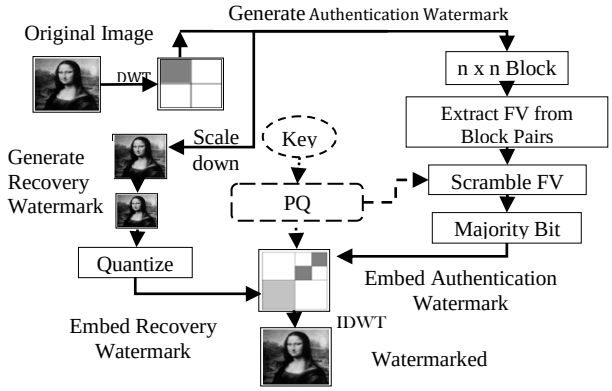

Figure 2. Overall scheme for generation of authentication and recovery watermark.

3.2.2. Embedding of W_R

W_R is embedded in the selected coefficients of LH_1 sub band by replacing five LSBs of the selected coefficients with the quantized binary equivalent of W_R.

3.3. Extraction of the Authentication Watermark and Verification of Integrity

The watermark extraction procedure is similar to the watermark generation and insertion procedure. The vector R needs to be secretly shard with the authenticator along with the seed used to generate the PQ sequence.

3.3.1. Generation of Authentication Watermark W_A^*

The content based authentication watermark W_A^* is generated for the received image by following the procedure in section 3.1.1.

3.3.2. Evaluate Ratio of Corresponding Coefficients

The ratio of coefficients $R^{\sim}$ for the received image is evaluated using the procedure in steps 1 and 2 of section 3.1.2.

3.3.3. Extraction of Authentication Watermark $W_A^{\sim}$

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 $W_A^{\sim}$.

3.3.4. Verification of Integrity

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 3.1.2, step 3.

3.4. Extraction of the Recovery Watermark W_R

To extract the estimated image, the reverse procedure of the Recovery Watermark generation and embedding is performed. The corresponding LH_1 sub band is selected and five LSBs from the selected coefficients is extracted. The extracted bits are used to reconstruct the quantized coefficient values using Table 1. The vector representing the quantized coefficients represents the extracted Recovery Watermark $W_R^{\sim}$.

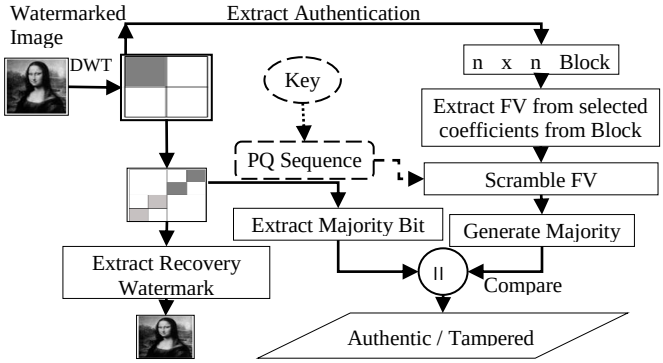

Figure 3. Overall scheme for verification and recovery.

4. Performance Evaluation

The authentication scheme described in this paper is implemented in Matlab 7.10.0.5 (R2011a) environment. Images of type tiff, bmp, png and color images of various sizes and complexities were considered for the study.

4.1. Watermark Embedding Results

4.1.1. Quality of Watermarked Image

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.

The quality of the watermarked images after embedding only W_A, only W_R and both W_A and W_R for Δ=3 are shown in Figure 4. Tables 3 and 4 summarizes the experimental results. The Peak Signal to Noise Ratio (PSNR) of the images watermarked with only W_A are in the range 63-71 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.

a) only W_A. b) onlyW_{VA}. c)after embedding both W_A*and* W_{VA} for embedding Strength a=0.9.

Figure 4. Watermarked images after watermark embedding.

4.1.2. PEG Compression

Any image authentication system should be robust to compression attacks. The robustness of the scheme is evaluated by compressing the watermarked image with different quality indices and then trying to extract the image.

Table 2. Quality metrics after watermark embedding-only W_A, only W_{VA} and after embedding both W_A *and* W_{VA} for embedding strength a=0.9.

Image	Only W_A			Only W_{VA}			Both W_A *and* W_{VA}		
	PSNR	SSIM	PCC	PSNR	SSIM	PCC	PSNR	SSIM	PCC
Lena	67.4	1	1	52.1	0.94	0.9	51.3	0.94	0.9
Pirate	66.8	1	1	50.1	0.96	0.9	50.0	0.96	0.9
Camera man	70.4	1	1	55.8	0.9	0.9	55.1	0.89	0.9
Living Room	63.0	1	1	46.5	0.93	0.9	46.2	0.93	0.9

a) dual watermarksW_A and W_{VA} and after JPEG2000 compression With. b)QF=10. c)QF=15. d) QF=20.

Figure 5. Watermarked pentagon image with.

a) dual watermarked image and dual watermarked image after JPEG2000 compression with. b)QF=10. c)QF=15. d) QF=20.

Figure 6. Visual authentication watermark extracted from.

The results in Tables 3 and 4, and Figures 5 and 6 indicate the scheme is robust to JPEG2000 compression with quality factor up to 20 which is the highest permissible value in JPEG2000 compression.

Even though JPEG2000 is the current standard for image compression, it will take some time before it is universally implemented. Till such time, JPEG compression will be prevalent. The results in Figures 7-8 and Table 4 indicate the scheme is not very robust to JPEG compression beyond QF= 90%. The quality of the recovery watermark extracted from the JPEG compressed image deteriorates beyond QF=90%.

4.1.3. Addition of Noise

The proposed authentication scheme is evaluated for robustness against common signal processing operations that occur during storage and transmission of images. Noise distorts and degrades the image which in turn distorts the embedded watermarks.

a) QF=95. b) QF=90.

Figure 7. Watermarked Pentagon image with dual watermarks W_A and W_{VA} and after JPEG compression with.

a) QF=95. b) QF=90.

Figure 8. Visual authentication watermark extracted from dual watermarked Pentagon image after JPEG compression with.

The output of the authenticator is in terms of percent of tampered blocks φ, which is a measure of strength of manipulation between the generated and extracted watermarks W_A^* and $W_A^{\sim}$ and is given by $\varphi = \frac{|R-R^{\sim}|}{max(R,R^{\sim})} * 100$, where R and $R^{\sim}$ are the received and computed ratios as explained in section 3.3.3. The results indicate that the scheme is robust to compression, salt and pepper noise and mild smoothening. The security of the scheme is evident in the authentication result when the wrong key is used to de scramble the watermark and 78% of the blocks are detected as tampered.

Table 3. Performance of the authentication scheme under various attacks for watermarked image of Lena. Correlation and PSNR are calculated for the recovered watermark after the attack.

Attacks	Authentication Results	Recovery Results	
	Percentage of blocks detected as tampered	Correlation value	PSNR value
Original watermarked image	Nil	1	65
Wrong key	78%	-	-
JPEG= QF95%	10%	0.97	42.92
JPEG QF=90%	28%	0.93	38
JPEG2000 QF=5	2%	1	55.3
JPEG2000 QF=20	10%	1	50
Median Filter [3 3]	40%	0.3	13
Median Filter [4 4]	36%	0.4	17
Gaussian blur M=0; V=0.001	8%	0.87	31
Salt and pepper Noise density =0.01	12%	0.82	30

5. Conclusions

This paper describes an authentication scheme that uses content invariant features of the image as a self authenticating watermark and a quantized down sampled approximation of the original image as a recovery watermark, both embedded securely using a pseudorandom sequence into multiple sub bands in the DWT domain. The second watermark is used to reinforce the authentication decision and the combination of DWT and DCT domains makes the image mildly robust to JPEG compression and highly robust to JPEG2000 compression. The scheme also ensures good quality of the watermarked images as the spatio-frequency properties of DWT are utilized to embed the dual watermarks.

References

[1] Abbasi A., Seng W., and Ahmad I., "Multi Block based Image Watermarking in Wavelet Domain using Genetic Programming," *The International Arab Journal of Information Technology*, vol. 11, no. 6, pp. 582-589, 2014.

[2] 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. 2, no. 11, pp. 3866-3870, 2006.

[3] Chen B. and Gregory W., "Digital Watermarking and Information Embedding Using Dither Modulation," *in Proceedings of the 2nd IEEE workshop on Multimedia Signal Processing*, Redondo Beach, pp. 273-278, 1998.

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

[5] Hu M., Der-Chyuan D., and Chang M., "Dual-Wrapped Digital Watermarking Scheme for Image Copyright Protection," *Computers and Security*, vol. 26, no. 4, pp.319-330, 2007.

[6] Hu Y., Lo C., Chen W., and Wen C., "Joint Image Coding and Image Authentication Based on AMBTC," *Journal of Electronic Imaging*, vol. 22, no. 1, pp. 013012 -1-013012-11, 2013.

[7] Hu Y., Chuang J., Lo C., and Chen W., "Grayscale Image Tamper Detection and Recovery based on Vector Quantization," *International journal of Security and its Applications*, vol. 7, no. 6, pp. 209-228, 2013.

[8] Hussein J., "Spatial Domain Watermarking Scheme for Colored Images Based on Log-Average Luminance," *Journal of Computing*, vol. 2, no. 1, 2010.

[9] Kundur D. and Dimitrios H., "Digital Watermarking for Telltale Tamper Proofing and Authentication," *in Proceedings of the IEEE*, vol. 87, no. 7, pp. 1167-1180, 1999.

[10] Lie W., Lin G., and Cheng S., "Dual Protection of JPEG Images based on Informed Embedding and Two-Stage Watermark Extraction Techniques," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 3, pp. 330-341, 2006.

[11] Lin C. and Chang S., "SARI: Self-authentication-and-Recovery Image Watermarking System,"

in Proceedings of the 9th ACM International Conference on Multimedia, Ontario, pp. 628-629, 2001.

[12] Lin C. and Chang S., "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 2, pp. 153-168, 2001.

[13] Lu C. and Liao H., "Multipurpose Watermarking for Image Authentication and Protection," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1579-1592, 2001.

[14] Parthasarathy A. and Kak S., "An Improved Method of Content Based Image Watermarking," *IEEE Transactions on Broadcasting*, vol. 53, no. 2, pp. 468-479, 2007.

[15] Patra J., Phua J., and Bornand C., "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.

[16] Pillai J. and Padma T., "The Analysis of PQ Sequences Generated from Continued Fraction for use as Pseudorandom Sequences in Cryptographic Applications," *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, New Delhi, pp. 633-644, 2016.

[17] Qi X., Xin X., and Chang R., "Image Authentication and Tamper Detection using Two Complementary Watermarks," *in Proceedings of 16th IEEE International Conference on Image Processing*, Cairo, 2009.

[18] Qi X. and Xin X., "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.

[19] Seng, W., Jiang Du J., and Pham B., "Semi Fragile Watermark with Self Authentication and Self Recovery," *Malaysian Journal of Computer Science*, vol. 22, no. 1, pp. 64-84, 2009.

Jayashree Pillai earned her BE degree in Computer Science from Tamilnadu College of Engineering, Coimbtore, M.Tech degree in Computer Science from M. S Ramaiah Institute of Technology, Bangalore and is presently pursuing her research from Mother Teresa Women's University, Kodaikanal, in the area of image authentication. She is presently working as Associate Professor, Department of MCA, AIMS Institute of Higher Education, Bangalore. Her interests are in the field of Computation, networking and security. She has a number of national and international publications in the area of wireless sensor networks and image authentication.

Padma Theagarajan has a Masters degree in Computer Applications from Alamelu Angappan College for Women in 1988 and M. Tech from AAU. She did her Ph. D in C. Sc from Mother Teresa Women's University, Kodaikanal. At present, she is the Professor and Head of MIS, Dept. of MCA, Sona college of Technology, Salem. She has a number of National and International publications to her credit and serves as editor and reviewer of a number of well known national and international She also serves as a subject expert in the BOS of various institutions. She is a Life Member of ISTE and CSI and her areas of specialization include Artificial intelligence, Data mining, DSS and Knowledge based systems.