Kingdom of Saudi Arabia
Majmaah University
Ministry of Higher
Education
College of Science Al Zulfi

المملكة العربية السعودية
وزارة التعليم العالي
جامعة المجمعة
كلية العلوم بالزلفي
قسم علوم الحاسب والمعلومات

Student Affairs System
For College of science Al Zulfi
Department of Computer Science and Information

Authentication of Digital Images

Graduation Project

Submitted in partial fulfilment of the requirements for the award of
Bachelor Degree of the Majmaah University

Submitted by:

Abdullatif Ali Alsaif

351101042

Under the supervision of:

D. Rafi Ullah Habib

TABLE OF CONTENTS

ABSTRACT

In this project, a secure semi-fragile watermarking technique based on integer wavelet transform with a choice of two watermarks has been proposed. A random watermark is generated by using a secret key and then embedded in the least four significant bits. The semi-fragility makes the scheme tolerant against JPEG lossy compression with the quality factor as low as 70% and locates the tempered area accurately. The computational complexity is reduced by using a parameterized integer wavelet transform. Experimental results show that the proposed scheme guarantees the safety of a watermark, recovery of image and localization of tampered area.

Chapter 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. Instead of some other tools like cryptography, watermarking based authentication is more advantageous 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. Digital data authentication is thus one of the most important and investigated security applications in this regard. Authentication of the image is the act of establishing or confirming that the image is credible. Imperceptibility, fragility and efficient computation are the basic requirements for authentication. In this project, we strive for imperceptibility, efficient computation and also both the fragility and robustness i.e. semi-fragility. Cryptographic tools like MD4, MD5 and SHA are also available to authenticate the images but it cannot survive any legitimate transformation like compression and format change etc and also is unable to localize the tampered area. Due to these limitations, the interest in the alternative of the cryptography, which is authentication, has rapidly increased [2].

Previous techniques [8] do not clarify how and where the image is tampered but only identify that the image is tampered or not. Friedman et al [9-10] propose 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 [13], on the other hand, have proposed a new fragile watermarking technique, which is based on the Gaussian mixture model in the different wavelet scales. In [14], 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 present work exploits the advantages of both the techniques [2, 3] 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 [1], an image digest and a binary image. The image digest is computed through a properly modified version of JPEG coding, operating at very high compression ratio on the original image [2]. 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, the binary signature is processed with a private key to ensure security [3]. Embedding binary image can help in accurately detecting manipulations made in the 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.

Chapter 2

LITERATURE SURREY

Significant numbers of robust watermarking algorithms are proposed, which are aimed to protect copyrights [19,31]. In [17,23–27,32,33,34,37,40], the authors use DCT based perceptually tuned robust watermarking to protect digital images. However, it is difficult for these approaches to detect a variety of distortions including malicious attacks, rather the main aim of these approaches is to resist the attacks. Therefore fragile watermarking algorithms have been proposed, which are sensitive to distortions. In this regard, the tamper assessment function is used to check the integrity of images [29]. DCT, vector quantization, quantization index modulation, and image structure based watermarking approaches are also proposed to prevent tampering and make the content secure [30,35,36,38,41,42]. These approaches are able to detect the distortions but they may fail against incidental manipulations like JPEG compression. Semi-fragile watermarking techniques [18,20,28] have thus been proposed which are tolerant towards incidental manipulation, but sensitive to malicious attacks. Recently besides authentication, some researchers have proposed watermarking techniques that are also able to perform image recovery. For example, the altered areas of the image are recovered in [21,22] using vector quantization technique. The discrete wavelet transform (DWT) technique has been utilized in [31] to extract image information from the low-frequency coefficients and is hidden in the mid-frequency components for tamper detection and recovery. In [39], fractal image compression and watermarking schemes are combined to accomplish image authentication as well as recovery. In [16], a hierarchical watermarking based approach is used whereby, the image can be recovered elegantly but the authenticity of the system is weak as it authenticates/localizes the content blockwise.

Similarly, in [15], a multiple-watermarking approach is used to accurately authenticate the image and recover it by using the self-recovery technique, but at the cost of imperceptibility. The embedding of the first watermark, which is generated with a secret key, is weak. The attacker can easily modify the significant bits leaving five LSBs intact. The wavelet subbands LL1 (approximation subband at level 1), HH2 and VV2 (detailed subbands at level 2) are secured by using the second watermark. The second watermark is being secured by using keys k1 and k2. But subbands other than LL1, HH2 and VV2 are not protected. The attacker is free to modify these areas to alter the watermarked image without being detected. Therefore, as regards the security aspects of the scheme proposed in [15], it is hard for the attacker to successfully substitute the concerned subbands with another image. However, manipulations to degrade the quality of the watermarked image are possible. Very recently, security consideration is designated as one of the most important aspects of the watermarking system.

PROBLEM STATEMENT

The motivation for taking up “Image Watermarking” as the topic for this project was to make images on web prone to illegal duplication of images without the consent of the owner. Watermarking was proposed to resolve this so that a watermark is hidden in the image as a token of ownership. But the intelligent attacker performed various attacks on the image to destroy the watermark without harming the image much. This motivated us to think on the lines to propose a robust algorithm which resists various attacks like JPEG Compression, Addition of Noise and Cropping attack. On further researching, we decided to work on DCT-based watermarking schemes. We found that there was a scope of improvement in these schemes. However, before that, we needed to bring certain facts about the image’s response to various attacks after doing minor changes in the image. These will be discussed in detail in later chapters.

The goal of this project is to analyze the performance of colour channel for DCT-based watermarking scheme and to propose a robust algorithm, which resists various attacks like JPEG Compression, Addition of Noise and cropping attack.

Chapter 3

PROPOSED METHOD

WATERMARK GENERATION

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

BINARY IMAGE PREPROCESSING:

A binary signature (Binary Image) which is used for accurate authenticity of the cover image is preprocessed before being embedded. Let W be a binary signature of size M×N, then

$$W = w(i, j) \qquad (1 \le i \le M, 1 \le j \le N) \qquad (1)$$

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

and Rand is a pseudo-random matrix of the same size generated by a secret key.

$$Rand = R_n(i, j) \qquad (1 \le i \le M, 1 \le j \le N) \qquad (2)$$

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

We adopt formula (3) to get the ultimate watermark W :

$$W_1 = W \oplus Rand \qquad (3)$$

where $\oplus$ denotes the exclusive OR.

4.1. EMBEDDING

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

1. Given an N×N image, after applying a 1-level 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 Figure 1.
2. We use the following formula [6] to embed the watermark in the LL3 subband coefficients. Let LFB (a) denote the five least significant bits of a, while LFB(a,b) represents 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 as shown in figure 1 [6]. The selection is based on the quality of the watermarked image

When $W_1(i, j) = 0$ formula (9) is adopted

$$f^*(i, j) = \begin{cases} LFB(f(i, j) - 01000, 11000) & if \quad LFB(f(i, j)) \le 01000 \\ LFB(f(i, j), 11000) & if \quad otherwise \end{cases} \tag{9}$$

When $W_1(i, j) = 1$ formula (10) is adopted

$$f^*(i, j) = \begin{cases} LFB(f(i, j) + 10000, 01000) & if \quad LFB(f(i, j)) \le 11000 \\ LFB(f(i, j), 01000) & if \quad otherwise \end{cases} \tag{10}$$

where $f(i, j)$ is an IWT coefficient in LL3 subband before embedding, $f^*(i, j)$ is the IWT coefficient after embedding. 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 [6].

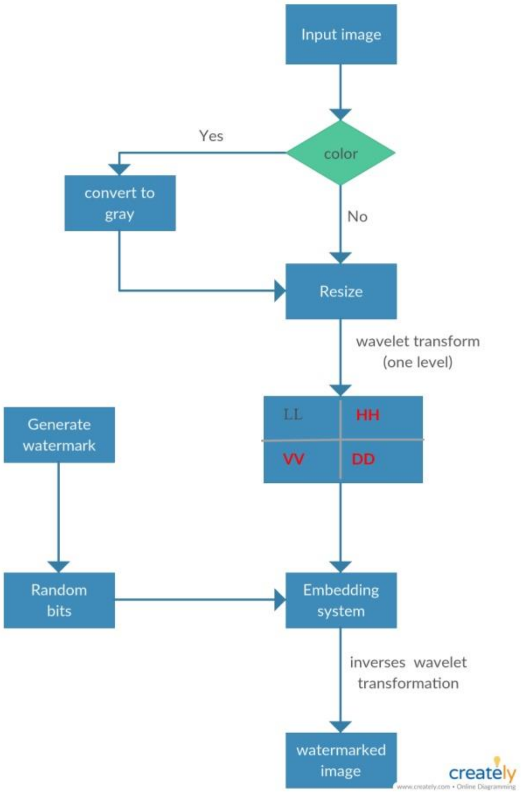

Figure 1. Watermark embedding process

Peak Signal to Noise Ratio (PSNR) is used to measure the induced distortion caused by the watermark [11]. PSNR in decibels (dB) is computed [12] using formula (11)

$$PSNR = 20\log_{10}\left[\frac{255^2}{\frac{1}{MN}\sum_{i,j}(x(i,j)-y(i,j))^2}\right] \qquad (11)$$

where $1 \le i \le M$ and $1 \le j \le N$

Our proposed scheme uses the parameterize integer wavelet transform (IWT) which is the fast approach of Discrete Wavelet Transform. Based on the idea, Meerward *et al* [4] proposed for the first time to use the parameterized wavelet transform. However, their scheme is based on conventional DWT.

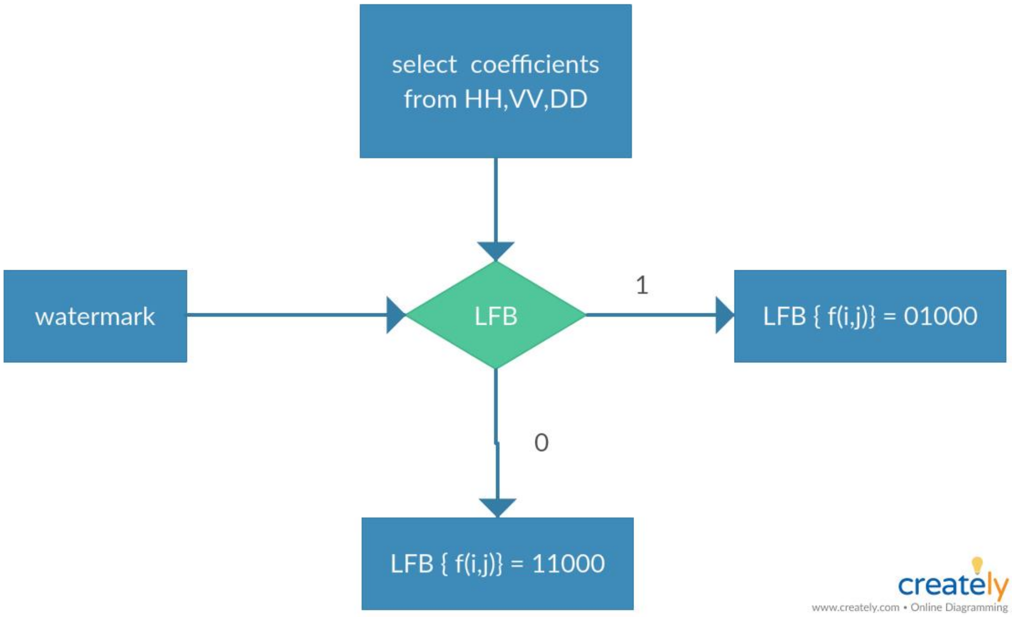
select coefficients from HH,VV,DD
watermark
LFB
1
LFB { f(i,j)} = 01000
0
LFB { f(i,j)} = 11000
creately
www.creately.com • Online Diagramming

Figure 2. Embedding System

4.2. EXTRACTION

In the integrity verification phase, the watermarked image undergoes a procedure, where the embedded watermarks (W_1 and W_2) are extracted. The binary watermark W_1 is extracted from the LL3 subband while the image digest W_2 is extracted from the HL2 and LH2 subbands. The extraction procedure of W_1 , which is used for authentication, includes the following steps:

- Given an N×N watermarked image, after applying a 1-level IWT, the approximation subband is two times decomposed and LL3 is selected as shown in Figure 2.
- Let $W_1^{*'}(i,j)$ denote the extracted watermark bit and $LFB(a)$ denote the fifth least significant bit of a, then

$$W_1^{*'}(i,j) = \begin{cases} 1 & LFB(f^{*'}(i,j)) = 0 \\ 0 & LFB(f^{*'}(i,j)) = 1 \end{cases} \qquad (1 \le i \le M, 1 \le j \le N) \qquad (12)$$

- Now as the watermark has been processed, therefore, at the verification phase, we again process it to obtain the ultimate watermark. $W_1^{'}$ (a binary image) using formulas (13):

$$W_1^{'}(i,j) = W_1^{*'}(i,j) \oplus Rand_n(i,j) \qquad (1 \le i \le M, 1 \le j \le N) \qquad (13)$$

where Rand is the pseudo-random number matrix and $W_1^{*'}$ is the extracted binary signature.

- We express the difference mark as (14):

$$D(i,j) = \left| W_1(i,j) - W_1^{'}(i,j) \right| \quad (1 \le i \le M, 1 \le j \le N) \qquad (14)$$

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

4.3. TAMPER DETECTION

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

$$Difference = \left| W_1(i,j) - W_1^{'}(i,j) \right| \qquad (16)$$

If the 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 the error. Our proposed approach accurately locates the tampered area and distinguishes between malicious and incidental attacks. The details are given as follow:

Dense pixel: For a mark error pixel in the difference image, it is a dense pixel if at least one of its eight neighbour pixels is a mark error pixel and a sparse pixel otherwise [3]. 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}

$Total\ Area\ =\ Dense\ Area + Sparse\ Area$

$\xi = Total\ area / Area$

$\Delta = Dense\ Area / Sparse\ Area$

- $if\ \xi = 0$ Then the image does not tamper)
- $if\ \xi > 0$ and $\Delta < \gamma$ then tampering is incidental, where γ is set empirically between 0.5 - 1.0.

$if\ \Delta \geq \gamma$ 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 figure 5.

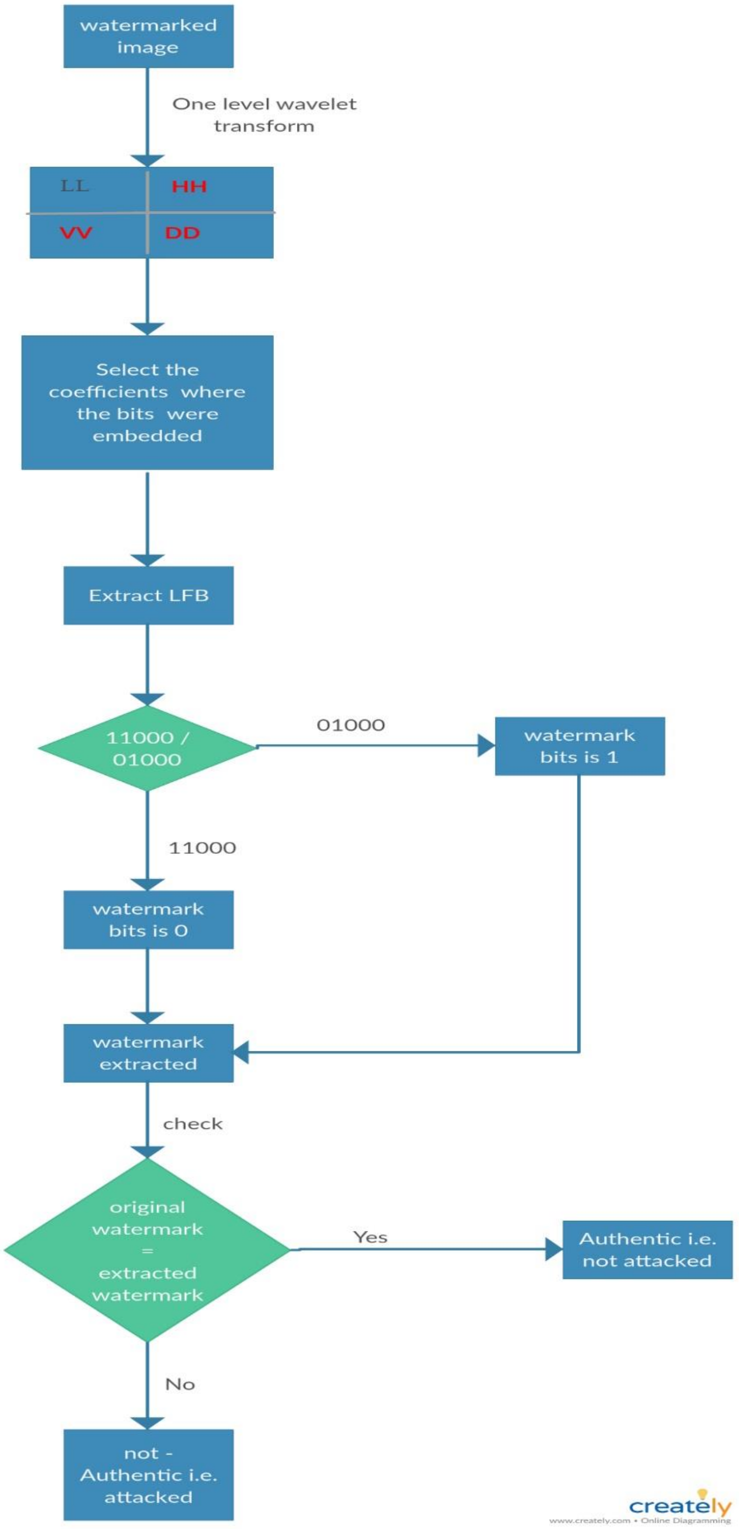
watermarked image
One level wavelet transform
LL
HH
VV
DD
Select the coefficients where the bits were embedded
Extract LFB
11000 / 01000
01000
watermark bits is 1
11000
watermark bits is 0
watermark extracted
check
original watermark = extracted watermark
Yes
Authentic i.e. not attacked
No
not - Authentic i.e. attacked
creately
www.creately.com • Online Diagramming

Figure 3. Watermark extraction

Chapter 4

EXPERIMENTAL RESULTS

To be provided in the Project 2 phase

Chapter 5

CONCLUSIONS

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 40 dB for different images, but still, the quality of the watermarked image is satisfactory. The technique is highly secure because of the inclusion of three private keys at various stages of the watermark generation. The proposed scheme also shows efficient authentication for the 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 is another authentication attribute achieved in our proposed semi-fragile secured watermarking scheme. We may use classification techniques like Genetic Programming (GP) and Support Vector Machine (SVM) to differentiate between the no alterations, intentional and unintentional alterations using γ, ζ, Δ, Dense Area and Sparse Area as functions/features.

REFERENCES

[1] Ching-Yang Lin and Shi Fu-Chang, Semi-Fragile Watermarking for authentication of JPEG visual contents, SPIE security and watermarking of Multimedia contents II, EI-00 San Jose, CA, Jan 2000.

[2] Alessandro Piva, Franco Bartolini and Roberto Caldelliy, Self recovery authentication of images in the DWT domain, International Journal of Image and Graphics Vol. 5, No. 1 149-165 (2005)

[3] Xiaoyun Wu, Junquan Hu, Zhixiong Gu, Jiwu Huang (contacting author), A Secure Semi-Fragile Watermarking for Image Authentication Based on Integer Wavelet Transform with Parameters, Copyright © 2005 Australian Computer Society, Inc.

[4] Meerwald, P.and Uhl,A. watermark security via wavelet filter parameterization. Proc. IEEE Int. Conf. on image processing, (3): 1027-1030 (2001)

[5] Ingemar J Cox, Methiw L Miller and Jeffery A Bloom, Digital Watermarking. (2002).

[6] Liu, H.M., Liu, J.F, Huang, J.W, Huang, D.R. and Shi, Y.Q. (2002): A robust DWT-based blind data hiding algorithm. Proc. of IEEE on Circuits and Systems, (2):672 - II-675.

[7] 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, Vol 8, No 1, (2006).

[8] F.Hartung and M.Kutter "Multimedia Watermarking Techniques" Proc. of IEEE, Vol. l87, PP 1079-1107, July 1999

[9] GL.Friedman "The trustworthy digital camera restoring credibility to the photographic image" IEEE Trans, Consumer Electron, Vol. 39, PP.905-910, Nov. 1993

[10] Xiang Zhou, Xiaohui Duan, Daoxian Wang "A Semi-Fragile Watermark Scheme for Image Authentication Proceedings of the 10th International Multimedia Modeling Conference (MMM'04) 0-7695-2084-7/04 $ 20.00 © 2004 IEEE

[11] Dima Pröfrock, Mathias Schlauweg, Erika Müller Richard Wagner "A new uncompressed-domain video watermarking approach robust to h.264/avc compression" From Proceeding (520) Signal Processing, Pattern Recognition, and Applications - 2006

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

[13] Hua fiun, Xiuo-Ping Zhang "Fragile watermark based on the Gaussian mixture Model in the wavelet domain for image Authentication" IEEE 2003

[14] Kurato Maeno, Qibin Sun, Shih-Fu Chang, Fellow, IEEE, and Masayuki Suto “New Semi-Fragile Image Authentication Watermarking Techniques Using Random Bias and Non uniform Quantization” IEEE Transactions on multimedia, vol. 8, no. 1, February 2006

[15] R. Chamlawi, A. Khan, A. Idris, Wavelet based image authentication and recovery, Journal of Computer Science and Technology 22 (6) (2007) 795–

804.

[16] C.C. Chang, Y.H. Fan, W.L. Tai, Four-scanning attack on hierarchical digital watermarking method for image tamper detection and recovery, Pattern

Recognition, Elsevier Sciences 41 (2008) 654–661.

[17] C-C. Chang, P-Y. Lin, J-S. Yeh, Preserving robustness and removability for digital watermarks using subsampling and difference correlation, Information

Sciences, Elsevier Science 179 (13) (2009) 2283–2293.

[18] K.H. Chi and L. C-T Li, Semi-fragile watermarking scheme for authentication of JPEG images,in: Proceedings of the International Conference on

Information Technology: Coding and Computing (ITCC’04), l.1, 2004, pp. 7–11.

[19] M.S. Hsieh, D.C. Tseng, Hiding digital watermarks using multiresolution wavelet transform, IEEE Transactions on Industrial Electronics 48 (7) (2001)

875–882.

[20] C-H. Huang, J-L. Wu, Fidelity-guaranteed robustness enhancement of blind-detection watermarking schemes, Information Sciences, Elsevier Science

179 (6) (2009) 791–808.

[21] K.L. Hung, C.C. Chang, A Robust and recoverable tamper proofing technique for image authentication, Springer-Verlag, Germany, 2003. 44-53.

[22] K.L. Hung, C.C. Chang, T.S. Chen, Secure discrete cosine transform based technique for recoverable tamper proofing, Optical Engineering 40 (2001)

1950–1958.

[23] A. Khan, Intelligent perceptual shaping of a digital watermark, PhD thesis, Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute

of Engineering Sciences and Technology, 2006.

[24] A. Khan, A Novel approach to decoding: exploiting anticipated attack information using genetic programming, International Journal of Knowledge-

Based Intelligent Engineering System 10 (5) (2006) 337–347.

[25] A. Khan, Anwar M. Mirza, Genetic perceptual shaping: Utilizing cover image and conceivable attack information during watermark embedding,

Information Fusion, Elsevier Science 8 (4) (2007) 354–365.

[26] A. Khan, A.M. Mirza, A. Majid, Optimizing perceptual shaping of a digital watermark using genetic programming, Iranian Journal of Electrical and

Computer Engineering 3 (2) (2004) 144–150.

[27] A. Khan, S.F. Tahir, A. Majid, Tae-Sun Choi, Machine learning based adaptive watermark decoding in view of an anticipated attack, Pattern Recognition,

Elsevier Science 41 (2008) 2594–2610.

[28] C.C. Ko and C.H. Huang, A novel semi-fragile watermarking technique for image authentication, Proceeding of the 6th IASTED International Conference

on Signal and Image Processing (SIP'04), Honolulu, Hawaii, 2004, pp. 24–29.

[29] D. Kundur and D. Hatzinakos, Digital watermarking for telltale tamper proofing and authentication, in: Proceeding of IEEE 87, vol. 7, 1999, pp. 1167–

1180.

[30] C.-T. Li, Digital Fragile Watermarking Scheme for Authentication of JPEG Images, IEE Proceedings-Vision, Image, and Signal Processing 151 (6) (2004)

60–466.

[31] K.F. Li, T.S. Chen, and S.C. Wu, Image tamper detection and recovery system based on discrete wavelet transformation, in: Proceedings of the

International Conference on Communications, Computers, and Signal Processing, vol. 1, 2001.

[32] T-C. Lin, M-C. Lin, Wavelet-based copyright-protection scheme for digital images based on local features, Information Sciences, Elsevier Science 179

(19) (2009) 3349–3358.

[33] C-S. Shieh, H-C. Huang, F-H. Wang, J-S. Pan, Genetic watermarking based on transform-domain techniques, Pattern Recognition, Elsevier Science 37

(2004) 555–565.

[34] H. Si, C.-T. Li, Copyright Protection in Virtual Communities through Digital Watermarking, in: S. Dasgupta (Ed.), Encyclopaedia of Virtual Communities

and Technologies, Part C, Idea Group Publishing, 2005, pp. 61–65.

[35] S. Suthaharan, Fragile image watermarking using a gradient image for improved localization and security, Pattern Recognition Letters, Elsevier Science

25 (16) (2004) 1893–1903.

[36] S. Suthaharan, S.W. Kim, H.K. Lee, S. Sathananthan, Perceptually tuned robust watermarking scheme for digital images, Pattern Recognition Letters,

Elsevier Science 21 (2) (2000) 139–145.

[37] S. Tonegawa, N. Morimoto, and K. Kamijoh, Alteration detection apparatus and method thereof, US Patent (6963655 B1) 2005.

[38] K. Toyokawa, N. Morimoto, S. Tonegawa, K. Kamijo, and A. Koide, Secure digital photograph handling with watermarking technique in insurance claim

process, in: Proceedings of the SPIE, 3971 2000, pp. 438–445.

[39] C.T. Wang, T.S. Chen, S.H. He, Detecting and restoring the tampered images based on iteration-free fractal compression, Journal of Systems and

Software 67 (2003) 131–140.

[40] F-H. Wang, K.K. Yen, L.C. Jain, J-S. Pan, Multiuser-based shadow watermark extraction system, Information Sciences, Elsevier Science 177 (12) (2007)

2522–2532.

[41] A.B. Watson, Visual optimization of DCT quantization matrices for individual images, in: Proceedings of the AIAA Computing in Aerospace 9, San Diego,

CA, 1993, pp. 286–291.

[42] D. Yu, F. Sattar, B. Barkat, Multiresolution fragile watermarking using complex chirp signals for content authentication, Pattern Recognition, Elsevier

Science 39 (5) (2006) 935–952.