Semi-Fragile Watermark for Visual Content Authentication

CHAMIDU ATUPELAGE, KOICHI HARADA.
Department of Information Engineering
Hiroshima University
1-7-1 Kagamiyama, Higashi-Hiroshima 739-8521
JAPAN

Abstract: - Digital watermarking was initiated to copyright protection and ownership verification of multimedia data. The evolution of the watermark has focused on different aspects of security issues in multimedia data such as integrity and authenticity. Fragile and semi-fragile watermarking schemes were introduced to accomplish these security requirements. In this paper, we propose semi-fragile watermarking scheme to authenticate visual content of the image. The proposed system provides high security and it facilitates to locate the tampered region in the image. Since the system inherits PKC (public key cryptography), the proposed scheme is considered as public watermarking approach.

Key-Words: - semi-fragile watermarking, public key cryptography, discrete cosine transformation, image authentication, imperceptibility, elliptic curve digital signature algorithm, JPEG.

1 Introduction

As the immense growth of the digital multimedia technologies, multimedia data creation and distribution have become very simple for the content owners. However illegal usage of multimedia data also has been increased such as copying and editing, facilitate unauthorized use, misappropriation and misrepresentation. Thus, the significance of protecting the integrity of the digital media elements and the intellectual property rights of its owners has became a great exertion among owners and content distributors. Digital watermarking is a promising approach for the content owners to defeat these substances. Initially digital watermarking proposed as a solution for intellectual property right management and copyright management and it was considered as robust approach. However evolution defined other types of watermark called fragile and semi-fragile watermarks that facilitate integrity and authenticity of multimedia data. Fragile watermarks are highly sensitive for both malicious and non-malicious manipulations. In practice non-malicious manipulations are accepted, thus fragile watermarks are poor solution for real-time applications. This constrain has been defeated in semi-fragile watermarks, which robust against non-malicious manipulations and fragile for malicious manipulations.

In this paper we propose new semi-fragile watermarking scheme which provides integrity and authenticity for images. The system is capable of localizing the tampered area. Inheriting the existing public key cryptography, the proposed scheme provides strong security. Most of the semi-fragile watermarking proposals, the watermark is another image or random bit sequence. If the attacker can safely alter the content without harming to the watermark the watermarking scheme will certainly compromised. In our approach the watermark is a digital signature of the visual content of the image. Therefore this watermarking scheme sensitive to visual content alternations such as cropping, replacing,etc .

In section 2 we will discuss the essence of semi-fragile watermarks and its features. Section 3 briefs existing multimedia authentication watermarking approaches and their limitations. Then we look into the system overview and design issues in section 4. Consequently we will discuss system definition and implementation issues in section 5 and 6. Simulation results will be presented in section & and in section 8 we will present the analysis and discussion of the experimented results. Finally we conclude everything with future direction of the scheme.

2 Semi-Fragile Watermarks for Image Authentication

Image and video authentication is applicable for some e-commerce applications such as law, defense, journalism, and video conferencing etc, which are intended to show that no tampering has occurred during the situations where the credibility of an image or video may be questioned.

In a fragile marking system, a signal (watermark) is embedded within an image such that subsequent alternations to the watermarked image can be detected with higher probability. Therefore fragile watermarking provides a promising approach for image and video authentication [1]. In history, fragile watermark authentication schemes intended to provide some set of common features. Our literature review of semi-fragile watermarking and information security concludes set of features which should incorporated in an effective semi-fragile watermarking scheme.

1. *Tamper detection*: watermarking systems should detect any tampering marked on the image.
2. *Perceptible content authentication*: watermarking scheme should be able to authenticate the visual content.
3. *Localizing alternations*: the system should be able to localize the alternation.
4. *Perceptual transparency*: The watermark should not be visible under normal circumstances.
5. *Large marking space*: Security should be increased, by providing large space for watermark.
6. *Robust to non-malicious attacks*: The watermark withstands to the predefined image manipulation technique.
7. *Incorporate to PKI*: watermarks should provide high level of security.
8. *Independent recovering*: recovery credentials should be independent from the watermark generation credentials.

3 Related Work and Motivation

Min Wu and Bede Liu have proposed a watermark scheme for image authentication with localization. However inconsistent values of look-up-table may cause noise in the watermarked image [2]. Visible Watermarking and Verifiable Digital Seal Image [3] is a fragile watermarking approach which has integrated to PKI based digital signature algorithm. This method does not localize the tampering and the seal image distorts the visual information. [4] proposed a blind watermarking approach, which tamper the alternation with localizing, but both parties have to share the same secrete. Multimedia authenticating method proposed in [5], which uses original watermark at the authenticator, thus it does not suit for real-time applications. [6] is vulnerable to JPEG and provides erroneous results for images having more edges and texts. Jessica proposed the new fragile watermarking approach in [7], though it is too fragile for a simple image processing technique.

The literature review is evident that there is not any method proposed been able to fulfill the entire semi-fragile watermarking requirements. Therefore in this paper, we propose new content based effective semi-fragile watermarking scheme for image authentication which reaches all defined requirements together with strong security authentic secure infrastructure.

4 System Overview & Design Issues

Lossy compression algorithms discard the information in digital assets which are not perceived by HVS. It is apparent; authenticating all information in an image will increase the computational cost and complexity. Thus in our approach we only authenticate the perceptual information in an image. Also the proposal is robust against the JPEG compression, and sensitive for the threats like content removal, visualization damages, cropping, etc.

DCT domain is desired to be used to retrieve visual content and embed the watermark, thus the verification is also carried out in the same domain. Watermark (digital signature) is generated from the visual content and it will be embedded into the image by choosing the invariant properties in JPEG system. Moreover it gains large space for the watermark and the length of the watermark is proportional to the security strength.

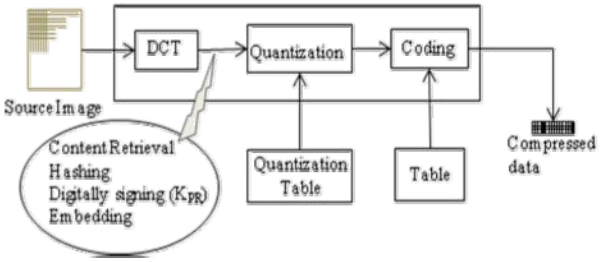

Fig. 1 Watermark insertion in DCT domain

JPEG lossy compression algorithm consists of three major steps which are DCT computation, Quantization and finally variable length code assignment [8]. Fig.1 depicts how proposed system involved in JPEG. Proposed method is engaged after DCT transformation and authentication is carried out after de-quantization.

4. 1 Visible Content Retrieval

Recipient of digital media always anticipate the assurance of the content integrity than the

degradation of quality (Ex: JPEG and MPEG). If some attack does not alter the important visual content recipient might accept it because of the persistance of the content. (Ex: identifying text or number or face in the image despite the high quality).

JPEG discards the information which cannot be perceived under ordinary conditions. Analysis of DCT coefficients concludes higher frequency coefficients are carrying more than half of the visual information in an image. Li Weng and Bart Preneel [9] has precisely benchmarked that 60% of the visual information conserved in the DC coefficient and 70% of visual information is drawn to DC and first two AC coefficients (according to zigzag scanning in 8×8 pixels block). Fig. 2 shows how much information is conserved only in DC coefficients.

Therefore authenticating only DC coefficient or DC and the first two AC coefficients, we can assure the proposed scheme provides integrity for 60% or 70% of the visual content respectively.

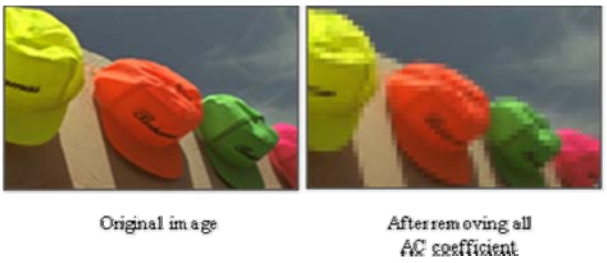

Fig. 2: Visual information preserve in DC coefficient

4.2 Security Infrastructure

According to the definitions in PKI, the key length is an important factor than the secrecy of the algorithm. Large watermarks need large space for embedding and intuitively it increases the noise. Elliptic curve digital signature algorithm (ECDSA) is a promising approach to overcome this limitation which provides maximum security even for small key lengths. The remarks and the benchmarks of the security of the elliptic curve cryptosystem are available in [10]. Therefore the security infrastructure of the proposed method entirely follows the PKI principles.

5 System Definition

In this section we will explain two main intermediate procedures in the proposed system. Firstly we will express the technique; watermark (signature) generation and embedding in DCT domain, then secondly retrieving and verification of the watermark.

Following notations will be using in each section for easy understanding of the routines.

- I_s : Source image.
- I_w : Watermarked image.
- $F_{RET}()$: Retrieving the visual content.
- $F_{INS}()$: Embedding the signature.
- $F_{EXT}()$: Extracting the signature.
- $H_{MD5}()$: Retrieving the visual content.
- K_{PR} : Signer's private key.
- K_{PU} : Signer's public key.
- s : Signature.
- md : Message digest.
- vc : Visual content.
- $ECDSA()$: ECDSA function

5.1 Signature Generation Embedding

At first the system should retrieve the desired visual content and preprocess them. (Arranging predefined manner). Hashing function generates the message digest from the visual content and we call to the signature generating algorithm (ECDSA) together with signer's private key. The signature is embedded into the image at the same frequency domain. The whole procedure is represented as algorithmic format as follow.

1. $vc \leftarrow F_{RET}(I_s)$
2. $md \leftarrow H_{MD5}(vc)$
3. $s \leftarrow ECDSA(K_{PR}, md)$
4. $I_w \leftarrow F_{INS}(I_s, s)$

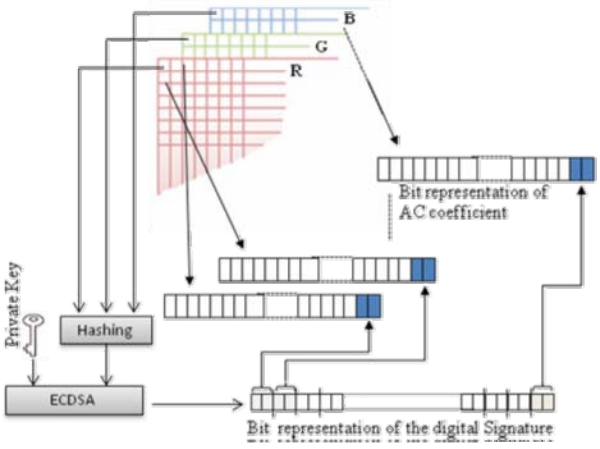

Fig.3: Signature generation and embedding (bit representation)

Fig.3 graphically represents the above procedure. For clear representation we assume the system authenticates only DC coefficients. We have given bit representation of digital signature and AC coefficients. In DCT domain three color components (RGB) considered independently and transformation is carried for blocks of 8×8 pixels. First we take three DC coefficients from each color component and send them to hash function and consequently to ECDSA with the private key. Generated signature is divided into two bit pieces and each piece replaces the last two significant bits of AC coefficients.

5.2 Signature Extraction and Verification

Visual content is retrieved from DC coefficients and will be sent to the hashing function to get the message digest. Signature is also recovered from the AC coefficients. Signature verification expects three parameters; message-digest, original signature and public key, then it will return the validity of the signature. The whole process can be illustrated as algorithmic format in four steps.

1. $vc \leftarrow F_{RET}(I_w)$
2. $s \leftarrow F_{EXT}(I_w)$
3. $md \leftarrow H_{MD5}(vc)$
4. $ECDSA(K_{PU}, md, s) \rightarrow Acceptance$

The graphical representation of this technique is almost similar to the Fig3. F_{EXT} is opposite to F_{INS} and ECDSA takes three parameters for signature verification.

6 Implementation Issues & Conquerors

In this section we will discuss the practical issue occurred in the implementation of this scheme and how we have defeated them.

6.1 Altering Quantization Table

In general DCT coefficients in low frequency area of each block capture the "essence" of the image while the values in the high frequency area capture the fine details and noise. The proposed system should have guaranty of preserving the low frequency coefficients and simply we achieve it by altering the quantization table. In fact this alternation directly effect to the compression ratio. However our experimental results are evident that we can ignore this consequence. To keep up recommended security standards, here we mark first 15 values as 1 in quantization table (refer Fig.5 (a)) in order to zigzag scanning.

6.2 Minimum Security Block Size

In most cryptographic functions, the key length is an important security parameter and some accepted organizations (such as ECRYPT, NIST) release the specifications in different security protocols. Both ECRYPT and NIST have recommended in [11] as 160 bits length key for elliptic curve cryptography. Then the corresponding signature length becomes 336 bits long. Conversely our scheme only allows AC coefficients to carry the signature. One 8×8 pixels block gives 14×2 bits space and including three colors it becomes 14×2×3 = 84 bits. Considering four blocks together (Fig.4) we can increase the space up to 84×4 = 336 bits. It is an adequate space for watermark. In this behavior the minimum authentication block size will become 16×16 pixels instead 8×8 pixels block. However 16×16 pixel block is still small to HVS and it can be considered as sufficient block size to authenticate.

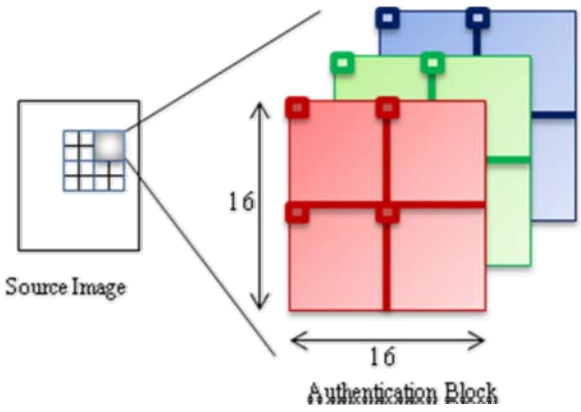

Fig. 4 Authentication block represent in source image

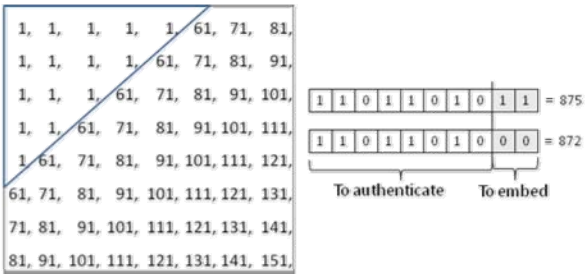

Fig. 5 (a) Quantization Matrix (b) Bit representation of AC coefficient.

6.3 Integrity of Higher Perceptual Information

If the system is expected to provide higher integrity (more than 60%), AC coefficients are required to be fed into signature generation. Since embedding scheme crashes two LSBs in AC coefficients, we separate AC bit sequence into two parts (refer Fig.5(b)); one portion is for authentication and the other is available for carrying the signature. Fig. 5(b) expresses that left portion holds considerable amount of information compared to the entire bit sequence. Therefore we recommend discounting 2 LSBs when AC coefficients are considered to be authenticated.

7 Simulations and Results

In this section we are proving the merit of the proposed scheme against to the compression ratio, visible quality degradation.

In JPEG, quantization contributes considerable endeavor to the compression ratio. Conversely the proposed scheme alternates the quantization table and this promise significance degradation of the compression ratio of the resulted image.

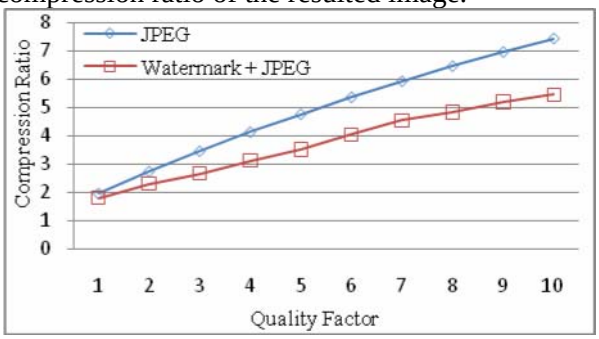

Fig. 6 Comparison of compression ratios of JPEG and watermarked JPEG images

Graph shown in Fig.6 compares the compression ratios of the general JPEG scheme and the proposed scheme (with altered quantization table). Graph is plotted for quality factors which lead from 1 to 10 against to average compression ratios of 11 images. Figure shows proposed technique having low compression ratio compared to the original JPEG system, however this degradation is acceptable, because numerically the difference of compression ratios is nearly 2 when quality factor is at 10. The degradation becomes nearly 1 when quality factor is 5. We can conclude that the difference is being minimized for low quality factors (Low compression rations).

Embedding some extra signal might make quality degradation of the original source. Our second experiment is carried out to account this quality degradation. Image quality is highly subjective measurement, though here we used objective quality matrix measure called PSNR (peak signal-to-noise ratio). Fig.7 expresses the PSNR values of R color component of 11 images in different quality factors which lead from 1 to 10. Since other color components show the similar pattern as in Fig.7 only one color component is presented. PSNRs of the watermarked and non-watermarked images are very close at small quality factors (low compression), though PSNRs of the watermarked image exceed the non-watermarked images for large quality factors. More precisely quality of the watermarked images is better than non-watermarked image for large quality factors (large compressions) because watermark image shows low compression.

The experiment results are apparent that proposed watermarking scheme does not make any extra visual distortion.

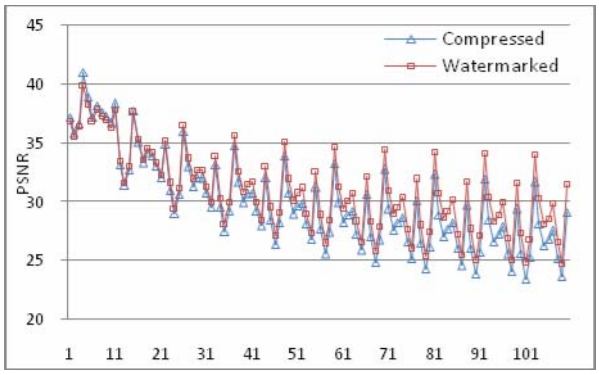

Fig. 7 Quality evaluation of watermarked and non-watermarked images

8 Analysis & Discussion

Though the most secure systems do not need to be perfect, the contrary high enough degree of security should be reached. In other word watermark breaking does not need to be impossible, but only difficult enough.

The objective of this research is to accomplish above defined fragile watermarking requirements. Table1 summarizes the achievement of the defined requirements.

Table 1: Features of effective semi-fragile watermarking scheme

Requirement	Status
Tamper detection	Satisfied
Perceptible content authentication	Satisfied
Localizing alternation	Satisfied
Perceptual transparency	Satisfied
Large marking space	Satisfied
Robust to non-malicious attacks	Satisfied
Incorporate to the PKI	Satisfied
Independent recovering	Satisfied

The proposed fragile watermarking scheme detects any tampering, which is being effective on DC coefficient or DC with first two AC coefficients. More precisely proposed scheme can be use to authenticate 60% or 70% percent of the visual quality of the image, conversely imperceptible alternations are smoothly avoided. The system can identify the compromised block mutually; locating the tampered block of size 16 ×16 pixels. We have evidently proved that the proposed scheme is transparent under normal visual conditions, and by choosing the frequency domain we have offered large space for embedding our signature. Watermark embedding is proceeding before quantization in JPEG compression and it promises withstand the watermark to JPEG compression. Information security point of view our approach provides high level security contrast to the current security requirements by incorporating ECDAS. Thus, the system is strong against to the common security threats such as predicting the secret credentials. According to the security architecture of the system, public key of the owner can be distributed in order to the cryptographic definition (publishing at secure certificate server with the signed certificate of authorized certificate authority) and receiver uses it for integrity verification.

9 Conclusion and Future Directions

Our literature review defined set of features, which should be achieved by an effective semi fragile watermarking scheme. In this paper we proposed an effective fragile watermarking scheme to accomplish all defined requirements. Paper discussed designed aspects, issues and conquerors, implementation in detail. Experimented results prove the essence of the proposed scheme.

Our future directions include: 1) applying same directions for MPEG and JPEG2000 definitions, 2) More detail evaluation of the performance of the scheme, 3) More precise analysis of security attacks and survivability of the watermark.

References:

[1] Lin E T, Delp E J. A review of fragile image watermarks. *Multimedia and Security Workshop (ACM Multimedia'99)*, Orlando, 1999, pp. 25-29.

[2] Min Wu, Bede Liu, Watermarking for Image Authentication. *International Conference on Image Processing (ICIP)*, 1998, pp. 437-441.

[3] Hyuncheol Park, Kwangjo Kim,Visible Watermarking using Verifiable Digital Seal Image, *Symposium on Cryptography and Information Security*, Japan,2001, pp. 103-108.

[4] J.J. Eggers, and B.Girod, Blind watermarking applied to image authentication, *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 3, 2002, pp. 1977 - 1980.

[5] Ming-Shing Hsieh, Din-Chang Tseng, Perceptual Digital Watermarking for Image Authentication in Electronic Commerce, *Kluwer Academic Publishers Norwell*, MA, USA, 2004, Vol. 4, pp. 157 - 170.

[6] Eugene T. Lin, Christine I. Podilchuk, Edward J. Delp, Detection of image alterations using semi-fragile watermarks, *SPIE proceedings series*, San Jose , 2000., Vol. 3971, pp. 152-163.

[7] Jessica, Fridrich, Security of fragile authentication watermarks with localization, *SPIE proceedings series*, Bellingham, 2002, Vol. 4675, pp. 691-700.

[8] Wallace, G.K., The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 1992, Vol. 38, pp. xviii - xxxiv.

[9] L. Weng, B. Preneel, On encryption and authentication of the DC DCT coefficient, *International Conference on Signal Processing and Multimedia Applications*, 2007.

[10] Certicom, The Elliptic Curve Cryptosystem, http://www.comms.scitech.susx.ac.uk/fft/crypto/EccWhite3.pdf, 2000.

[11] Keylength.com, http://www.keylength.com/en/compare/