WATERMARKING SCHEME FOR IMAGE CONTENT AUTHENTICATION

[1]K.BALASAMY M.E., [2]ISHWARYA G, [3]MAHALAKSHMI M

Department of Information Technology, Dr. Mahalingam College of Engineering and Technology, Pollachi, Coimbatore, India
Email: ishumcet@gmail.com

Abstract— Our project proposes a semi-fragile watermarking scheme for image content authentication with tampering localization. Here we use a non-traditional quantization method to modify one of the chosen approximation coefficient of each non-overlapping block to ensure its robustness against incidental attacks and fragileness. The image content authentication starts with watermark extracting process using the parity of quantization results from the probe image, where the round operation is used to ensure the semi-fragile property. It then constructs a binary error map and computes two authentication measures with M1 measuring the overall similarity between extracted and embedded watermarks and M2 measuring the overall clustering level of tampered error pixels. These two measures are further integrated to confirm the image content and localize the possible tampered areas.

Keywords—Quantization, Error map, Authentication measures.

I. INTRODUCTION

Digital multimedia plays an important role in applications such as news reporting, intelligence information gathering, criminal investigation, security surveillance, and health care. However, this trustworthiness could no longer be granted since users can easily manipulate, modify, or forge digital content without causing noticeable traces using low-cost and easy-to-use digital multimedia editing software. Therefore, digital multimedia authentication has become an important issue. Recently, digital watermarking techniques have been considered as one of the promising techniques for multimedia authentication. Among these, semi-fragile watermarking techniques have been proposed to protect copyright and prove tampering of the digital content. These techniques allow acceptable content preserving manipulations (i.e., changing the quality of the image without modifying the image content) such as common image processing (e.g., image blurring, Gaussian low-pass filtering, median filtering, and salt and peppers noise attacks) and JPEG/JPEG2000 compression, while detecting content-altering malicious manipulations such as removal, addition, and modification of objects. We briefly review several representative semi-fragile watermarking schemes in the domain of discrete cosine transform (DCT) or discrete wavelet transform (DWT). In general, these schemes use the chosen transform domain as the media to embed and extract watermarks. They then use the extracted watermarks to authenticate the digital content and localize the tampered areas if possible.

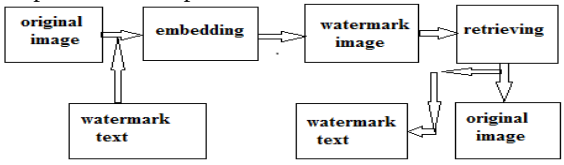


Fig 1: Block diagram for watermarking process

II. EXISTING SYSTEM AND NEED FOR PROPOSED SCHEME

In the existing system they use two ways to recover the image: confined recovery and self-recovery. Many techniques are proposed to protect the image and recover it where watermarked image has been tampered. These techniques embed two semi-fragile watermarks to authenticate the image and recover it, if tampered. One watermark is used for authentication purpose and other one is used for the recovery. Both the techniques have the ability of authentication and recovery of the image but at the cost of imperceptibility. However, the issue with these techniques is the use of two watermarks, which affect the imperceptibility of the watermark. Similarly, they use the semi-fragile watermark for authenticating and recovering the images but with cost concise authentication. This approach is unable to authenticate the content concisely and this technique is up to some extend only, it will not work on jpeg images. The high visual quality of stego-images, the data embedding capacity, and the robustness of the proposed lossless data hiding scheme against compression are acceptable for many applications, including semi-fragile image authentication. Specifically, it has been successfully applied to authenticate losslessly compressed JPEG2000 images, followed by possible transcoding.

The system not only achieve dual protection of the image content, but also maintain higher visual quality (an average of 6.69 dB better than a comparable approach) for a specified level of watermark robustness. In addition, the overall computing load is low enough to be practical in real-time applications. The security of medical images and reviewed some work done regarding them. A fragile watermarking scheme was then proposed that could detect tamper and subsequently recover the image. It required a

secret key and a public chaotic mixing algorithm to embed and recover a tampered image.

III. DISCRETE COSINE TRANSFORM

This scheme can identify altered regions within a watermarked image with 75% accuracy under moderate compression and near 90% accuracy under light compression. It protects the authenticity of the compressed watermarked image when the JPEG quality is higher than their predefined lowest authenticable quality. Here we proposes two methods, the first method adds a random bias factor to the fixed decision boundary to catch the malicious manipulation and keep the false alarm rate low. The second method uses the non-uniform quantization scheme to improve the accuracy in encoding the relationships between paired transform coefficients and increase the alteration detection sensitivity.

IV. DISCRETE WAVELET TRANSFORM

We propose a novel semi-fragile watermarking scheme by embedding a private-key-based random watermark bit sequence in the wavelet domain using the quantization method. The proposed watermarking scheme further utilizes two authentication measures derived from a binary error map to authenticate the image content and localize the tampered areas. Our proposed scheme also possesses the desired properties for an effective authentication watermarking scheme including invisibility, tamper detection, security, identification of manipulated areas, oblivion with no transmission of any secret information, and discrimination of incidental distortion and malicious tampering.

A. Watermark Embedding

Here we divide the original image into non-overlapping 4 × 4 blocks and embed the private-key-based random watermark bit sequence W in the wavelet domain of each unique randomly chosen 4 × 4 block. We choose the low-frequency components in each 4 × 4 block to embed a watermark bit since high frequency components are affected by most image processing techniques and small blocks lead to high capability in localizing the possible tampered areas. Specifically, we utilize the parity of the quantized value of the approximation coefficient to embed the watermark. The parity of a value is 0 when the value is divisible by 2 and the parity of a value is 1 when the value is not divisible by 2. To ensure the watermark invisibility and increase the robustness against common image processing attacks, we choose to use one of four values of the approximation subband as the media for the embedding process.

The following figure shows the embedding procedure where the original image is embedded with the cover image and using the recovery technique the original image is retrieved. While embedding the size of the original image should be less than the cover image.

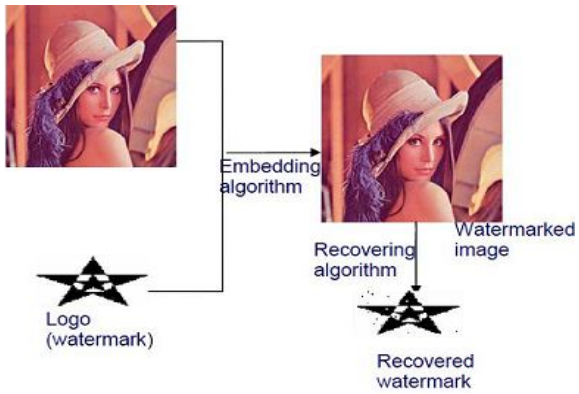


Fig 2: Block diagram for watermark embedding

B. Watermark Extraction

The watermark extraction process uses the same one way hash function together with the two secret keys K2 and K3 to choose the order of non-overlapping 4 × 4 blocks for extracting watermark. It then uses the parity of the quantized value X0 of the approximation sub band of each block to extract the watermark bit.

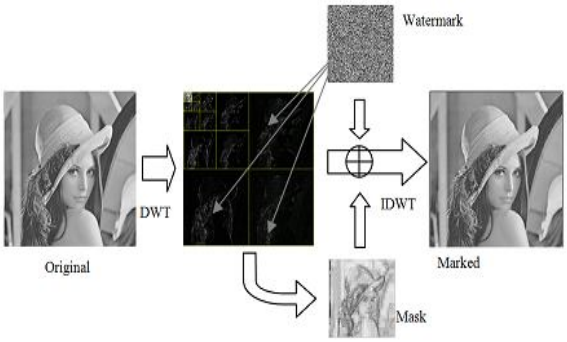


Fig 3: Block diagram for watermark extraction

C. Watermark Authentication

Here we define two authentication measures, M1 and M2, to protect copyright and prove tampering, where M1 measures the overall similarity between extracted and embedded watermarks and M2 measures the overall clustering level of the tampered error pixels. We compute M1 as the percentage of error pixels in ErrorMap. We compute M2 as the ratio of the number of strongly tampered error pixels to the number of tampered error pixels in ErrorMap. The value of M2 is set to 0's when the number of tampered error pixels is zero. Finally, we design a quantitative method to decide the authenticity of the probe image based on our defined two authentication measures.

Compute M1 using Errormap, if M1 value is lesser than or equal to median Threshold then update as its 3×3 median filtering. Compute using errormap if M2 value is less than the malicious threshold then the probe image is authenticated otherwise it is maliciously attacked. If the M1 value is greater than

the threshold half error bit and less than the error bit and also M2 is less than the threshold malicious the probe image will be incidentally attacked else it will be maliciously attacked. In the other case, if M1 value exceeds the threshold error bit the image will not be embedded with watermark.

V. QUALITY OF THE WATERMARKED IMAGE

The quality of the watermarked image can be determined by the following factor

1. Mean Squared Error(MSE)
2. Peak Signal to Noise Ratio(PSNR)

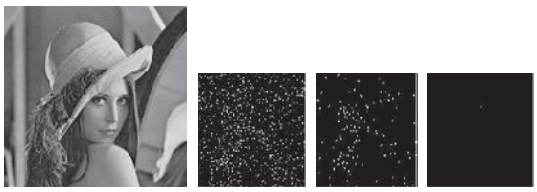
A. Mean Squared Error

The mean squared error (MSE) is derived in the embedding process. Since there is roughly an equal distribution of all values in the approximation subband. we assume that the original wavelet coefficients in the approximation subband are uniformly distributed over the range of [kq, (k + 1)q] for k ϵ Z. When the parity of the quantization result of the original wavelet coefficient LLi(x, y) matches the embedded watermark bit Wi, Li (x, y) is modified to the lower-bound kq, and the MSE cause by this quantization is:

$$MSE = \frac{1}{q}\int_0^q \tau^2 d\tau = \frac{q2}{3} \quad \text{---------- (1)}$$

(a) $M_1 = 0.0325$ and $M_2 = 0.8218$

Fig 4: Maliciously attacked watermarked image



(b) $M_1 = 0.0542$ and $M_2 = 0$

Fig 5: 70% JPEG compressed watermarked image

MSE of embedding p watermark bits in the block-based wavelet domain is:

$$MSE = \frac{p \times q^2}{3 \times W \times H} \quad \text{---------- (2)}$$

B. Peak Signal to Noise Ratio

Peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. PSNR is most commonly used to measure the quality of reconstruction of lossy compression codecs (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression. PSNR is an approximation to human perception of reconstruction quality. Although a higher PSNR generally indicates that the reconstruction is of higher quality, in some cases it may not.

Therefore the PSNR value of watermarked image is:

$$PSNR = 20 \log\left(\frac{255 \times bs}{q}\sqrt{3}\right) \quad \text{---------- (3)}$$

VI. TAMPERING DETECTION SENSITIVITY

The tampering detection sensitivity of our project is determined by the quantizer. The error map captures the changes in the quantization results and makes the tampering detectable for k 2 Z in the following two cases:

(1) The wavelet coefficient LLi'(x, y) of the watermarked image is 2kq and the manipulation causes a shift of LLi' (x, y) in the range of [(0.5 + 2k)q, (1.5 + 2k)q).

(2) The wavelet coefficient LLi'(x, y) of the watermarked image is 2kq + q and the manipulation causes a shift of LLi'(x, y) in the range of [(1.5 + 2k)q, (2.5 + 2k)q).

VII. THE EFFECT OF QUANTIZER

The following figure clearly demonstrates the average values of M1's and M2's of different watermarked images after no attack and after ten levels of JPEG compression with a quality factor of 100% down to 10% with a step size of 10%. We also show the two threshold lines of 0.4837 and 0.2418 for M1 and the threshold line of 0.6085 for M2 to facilitate comparison.

The figure clearly demonstrates the following: (1) The M1 and M2 values are 0's under no attack and after the JPEG compression of 100% quality factor. (2) For each chosen quantizer, the M1 and M2 values increase as the JPEG quality factor decreases. (3) For each JPEG compression quality factor, the M1 and M2 values increase as the quantizer decreases. (4) For all quantizers, the M1 value is below the threshold of 0.2418 for quality factors higher than 60% and the M2 value is below the threshold of 0.6085 for quality factors higher than 10%. To ensure our scheme achieves excellent watermark invisibility with the PSNR value around 41 db and is robust to JPEG compression of quality factors higher than 50%, we choose q as 15.

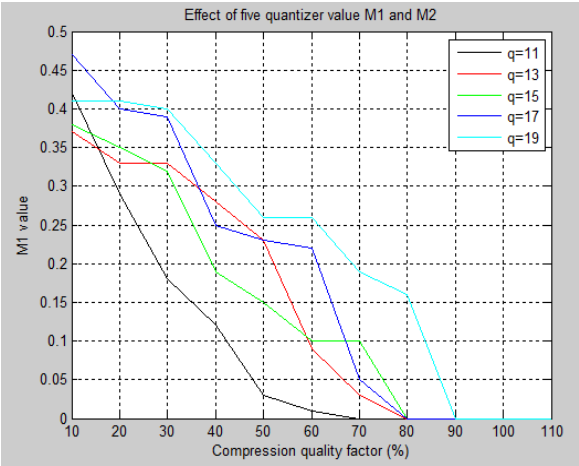


Fig 6: illustration of the effects of five quantizers on the M1 value

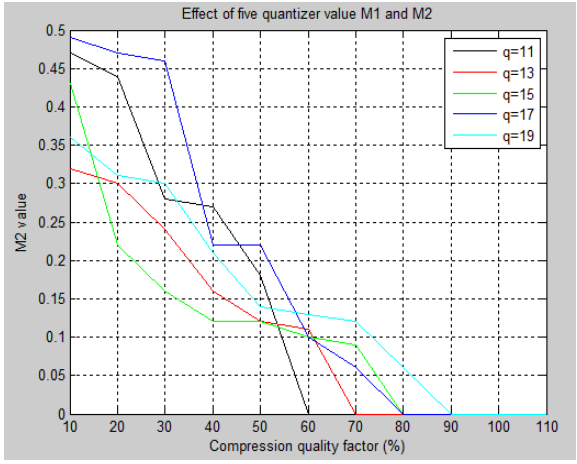


Fig 7: illustration of the effects of five quantizers on the M2 value

VIII. LOCALIZATION CAPABILITY

In the proposed scheme, the image content is monitored by the embedded and extracted watermark. Specifically, changing a value in the 2 * 2 corner of each 4 * 4 block may result in a mismatch in ErrorMap. To compensate the misclassification, we employ a window size of 3 * 3 to categorize each non-isolated error pixel as either strongly tampered or mildly tampered.

IX. WATERMARK INVISIBILITY

Table 1 summarizes the PSNR values after embedding watermarks in four representative images and the average PSNR values after embedding watermarks in some test images using our scheme and four peer schemes, respectively. This table clearly shows that our PSNR values for four representative images.

IMAGE	PSNR	MSE
SUNFLOWER	14.6914	2.2077
LENA	15.1782	1.9736
PUPPY	19.1303	794.4161
CAMERAMAN	21.7376	435.8329

Table 1: PSNR and MSE values for four images

X. ROBUSTNESS TO COMMON IMAGE PROCESSING ATTACKS

We performed four kinds of representative image processing attacks on some watermarked images. These attacks are ten levels of image blurring attacks using circular averaging filters, ten levels of Gaussian low pass filtering attacks using rotationally symmetric Gaussian low pass filter, five levels of salt and peppers noise attacks using noise density.

CONCLUSION

We present a novel semi-fragile watermarking scheme for image content authentication with tampering localization. The contributions of the proposed scheme are:

- Applying the quantization method to embed the private key-based watermark in the wavelet domain so that a majority of image distortions, which cause the intensity shift by a value larger than a half of the quantizer q, can be detected in the authentication process. Unlike traditional quantization based approaches, our quantization method modifies only one chosen wavelet coefficient in the approximation subband of the Haar wavelet transform of each block to ensure its robustness against moderate incidental attacks and fragileness against malicious attacks. In addition, our quantization method extracts the watermark bit using the round operation instead of the traditional floor operation to further ensure the semi-fragile property.

- Defining two kinds of tampered error pixels and two authentication measures to detect the authenticity of the probe image and prove tampering. Specifically, we consider an error pixel as strongly tampered if at least four of its eight neighbors are error pixels and an error pixel as mildly tampered if less than four of its eight neighbors are error pixels. We compute the percentage of error pixels in the error map as the first authentication measure, M1, which quantitatively evaluates the overall similarity between extracted and embedded watermarks. We compute the ratio of the number of our defined strongly tampered error pixels to the number of our defined tampered error pixels in the error map as the second authentication measure, M2, which evaluates the overall clustering level of the tampered error pixels.

- Using a binary error map together with the two authentication measures in the authentication process to compensate the possible misclassification in the error map, capture all possible distortions, and localize all possible tampered areas.
- Applying randomness strategies to increase the security of the proposed system. To this end, we first apply the Mersenne Twister algorithm to generate a watermark bit sequence using a private key. We then apply the one-way hash function to choose the order of the blocks for embedding watermark using two secret keys. We finally apply the Mersenne Twister algorithm to generate the embedding positions for each block using three private keys.

REFERENCES

[1] Robust Lossless Image Data Hiding Designed for Semi-Fragile Image Authentication,Zhicheng Ni, Yun Q. Shi, Fellow, IEEE, Nirwan Ansari, Senior Member, IEEE, Wei Su, Senior Member, IEEE, Qibin Sun, and Xiao Lin, Senior Member, IEEE,2012

[2] Dual Protection of JPEG Images Based on Informed Embedding and Two-Stage Watermark Extraction Techniques ,Wen-Nung Lie, Member, IEEE, Guo-Shiang Lin, and Sheng-Lung Cheng,2011.

[3] Medical Image Watermarking with Tamper Detection and Recovery, Jasni M Zain and Abdul R M Fauzi,IEEE,2011

[4] A Semi-Fragile Lossless Digital Watermarking Scheme Based on Integer Wavelet Transform Dekun Zou, Member, IEEE, Yun Q. Shi, Fellow, IEEE, Zhicheng Ni, Member, IEEE, and Wei Su, Senior Member, IEEE,2009

[5] Digital Watermarking Algorithm Based On DCT and DWT Mei Jiansheng1, Li Sukang1 and Tan Xiaomei2 , Nanchang Power Supply Company,Nanchang, China,2011

[6] Analysis and Design of Secure Watermark-Based Authentication Systems Chuhong Fei, Student Member, IEEE, Deepa Kundur, Senior Member, IEEE, and Raymond H. Kwong, Member, IEEE,2012

[7] A Grayscale Image Authentication Method with a Pixel-level Self-Recovering Capability against Image Tampering. Che Wei Lee, Wen Hsiang Tsai,2011.

[8] Efficient Digital Image Authentication and Tamper Localization Technique Using 3Lsb Watermarking , Sajjad Dadkhah , Azizah Abd Manaf and Somayeh Sadeghi.2012. MultipurposeWatermarking for Image Authenticationand Protection Chun-Shien Lu, Member, IEEE, and Hong-Yuan Mark Liao, Member, IEEE,2010

[9] " http:// WWW.mathworks.com "

[10] " http://WWW.pudn.com"

[11] " http:// WWW.matlabtutorials.com "

★★★