MULTIRESOLUTION FRAGILE WATERMARKING USING COMPLEX CHIRP SIGNAL FOR CONTENT INTEGRITY VERIFICATION

Dan Yu, Student Member, IEEE, Farook Sattar, Member, IEEE, and Braham Barkat, Member, IEEE
School of Electrical and Electronic Engineering
Nanyang Technological University, Nanyang Avenue, Singapore 639798

ABSTRACT

This paper proposes a wavelet-domain multiresolution fragile watermarking scheme using an improved Quantization-Index-Modulation (QIM) embedding technique. A secure embedding zone is exploited in our proposed scheme to reduce the false detection rate of Kundur's scheme. The frequency modulated (FM) complex chirp signal is employed as watermark. Both the real and the imaginary parts of the chirp signal are embedded simultaneously in a hierarchical manner. Unlike the conventional schemes, the proposed scheme does not require the original watermark for content integrity verification. The *blind* authentication process allows embedding of arbitrary FM chirp watermarks.

1. INTRODUCTION

As a great volume of multimedia data is stored in a digital form, it is very easy to modify and forge its content by widely available software editing tools. The content authentication has become very important, to which can guarantee that the content has not been altered, or at least that the semantic characteristic of the image is still preserved. Recently, there have been a number of fragile authentication watermarking schemes proposed for tamper detection and content integrity verification of multimedia content [1]. One classification of content-authentication watermarking techniques is as follows [2]:

1) Visually-authenticated technique, where a visual pattern is hidden in the image (e.g., replacing the least significant bit (LSB) plane), and tamper detection is based on the visual assessment;

2) Statistically-authenticated technique, where an estimate of tampering likelihood is obtained based on correlation coefficient or measured mismatch between the original and the recovered authentication sequences;

3) Self-embedding techniques, where the features extracted from the content are embedded as authentication data for proof of authenticity and image protection.

The authentication based on the LSB visual pattern is rather insecure, for example, the modification of the content while maintaining the LSB plane unchanged. The self-embedding authentication would fail if large portions of the image are corrupted, since the authentication data may probably be lost in such a case. The statistically-authenticated watermarking may require the original watermark or a matched filter to estimate the correlation coefficient for authentication. The objective of this paper is to propose a new multiresolution fragile watermarking scheme by using complex chirp signal to achieve *blind* content authentication. An improved QIM technique is also proposed for watermark embedding to reduce the false detection rate of Kundur's scheme [3].

2. MULTIRESOLUTION REPRESENTATIONS

The proposed fragile watermarking scheme is performed in the wavelet domain of the original image. A grayscale Lena image of 256×256 pixels, as shown in Fig. 1(a), is used to illustrate the proposed scheme. In particular, the wavelet decomposition level, l, is set equal to three.

Chirp signals are sinusoidal signals with time-varying frequencies (FM-frequency modulation) and/or amplitudes (AM-amplitude modulation) [4]. A generalized form of a complex chirp signal, $x(t)$, can be written as $x(t) = A(t)\ e^{\vartheta(t)}$, where the magnitude $A(t)$ and the phase $\vartheta(t)$ are arbitrary functions of time. If the magnitude is constant, it becomes a complex FM chirp signal as

$$x(t) = A[\cos(\vartheta(t)) + j\ \sin(\vartheta(t))] \quad (1)$$

where A is an arbitrary real number, denoting the constant magnitude of the chirp signal.

For instance, a quadratic FM signal is used as the watermark, and in the discrete domain it is given by [5]

$$\mathbf{s}_i = \cos[2\pi(\rho_0 i + \rho_1 i^2 + \rho_2 i^3)] + j\sin[2\pi(\rho_0 i + \rho_1 i^2 + \rho_2 i^3)] \quad (2)$$

where i=0, 1, ..., N–1, and N is the signal length, and ρ_0, ρ_1 and ρ_2 are the chirp parameters. Note that the constant magnitude of the complex signal in (2) is chosen as one in this paper. The real and the imaginary parts for a quadratic FM signal of 256 samples are shown in Figs. 1 (c) and (d), respectively. By applying Wigner Distribution (WD) [6], the time-frequency representation (TFR) of this signal is obtained as shown in Fig. 1 (b), where a quadratic waveform is clearly observed within frequency range of [0.1, 0.4].

For embedding, the signal, $\mathbf{s}_i$(=$a_i + jb_i$), is quantized by Pulse-Code Modulation (PCM) technique [7], converting to digital codes using 7 bits. To simplify the coding process, we, first, map the values of the watermark samples to the non-negative integers ranging from 0 to 127 (=$2^7 - 1$) using

$$w_i = \text{round}\{(\mathbf{s}_i + 1) \times 63.5\} \quad (3)$$

where $w_i = [w_{ia} + jw_{ib}]$ denotes the i-th watermark sample after mapping. Then w_{ia} and w_{ib} are represented in binary

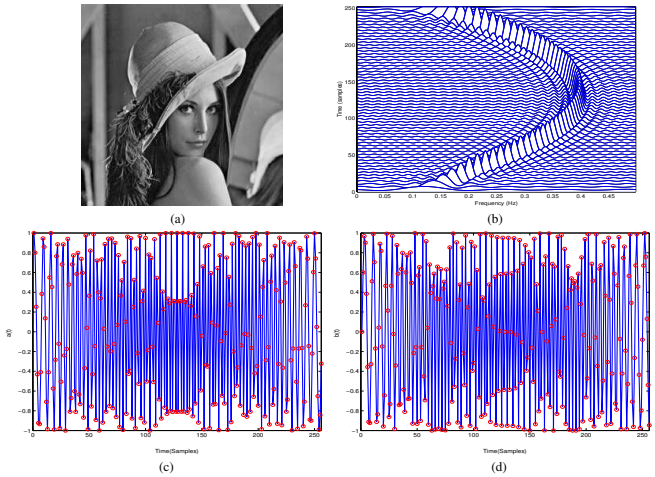

Fig. 1. (a) The original Lena image (256×256 pixels), (b) the TFR display of a quadratic chirp watermark signal $\mathbf{s}$, and the time plots of (c) the real part and (d) the imaginary part of $\mathbf{s}$ ('○' denotes the samples of the watermark).

form, as a_{in} and b_{in} ($i = 1, 2, ..., 256$ and $n = 1, 2, ..., 7$), respectively, which satisfy

$$w_{ia} = \sum_{n=1}^{7} a_{in} \cdot 2^{7-n}; \quad w_{ib} = \sum_{n=1}^{7} b_{in} \cdot 2^{7-n}. \tag{4}$$

In this way, watermark samples are successively sliced into seven bit-planes of different significance. The 1st bit-plane (n=1) is of the most significance, while the 7th bit-plane (n=7) is the least significant one. To correspond with the 3-level wavelet decomposition of the host image, these seven bit-planes are divided into 3 groups having different resolutions: (1) Low resolution group containing the most significant bits (MSB) a_{i1} and b_{i1}; (2) Intermediate resolution group consisting of a_{i2}, a_{i3}, b_{i2} and b_{i3}; and (3) High resolution group of a_{in} and b_{in} for $n = 4, 5, 6, 7$.

3. PROPOSED MULTIRESOLUTION WATERMARK EMBEDDING SCHEME

The wavelet coefficients are firstly obtained using 3-level discrete wavelet transform (DWT) with Harr bases. The rules for embedding are presented in Section 3.1, and the improved QIM embedding technique is elaborated in Section 3.2. The final watermarked image is obtained by an inverse DWT of the modified wavelet coefficients.

3.1. Watermark Embedding Rules

A wavelet coefficient in the 3rd decomposition level has successive correspondence with (2×2) components in the 2nd level and (4×4) components in the 1st decomposition level, forming a *cluster* of coefficients which is originated from the initial non-overlapping 8×8 image block.

The proposed watermark embedding scheme contains the following rules (see Fig. 2):

Rule 1: The sub-bands containing horizontal details, i.e., HL1, HL2 and HL3, are selected for embedding the real parts of the complex watermark sample bits of a_{in}; and the sub-bands containing vertical details, i.e., LH1, LH2 and LH3, are selected to embed the corresponding imaginary parts b_{in}.

Rule 2: One 7-bit watermark sample would be embedded into one cluster of wavelet coefficients in a multiresolution manner, by grouping them into three different resolution levels as described in Section 2. That is to embed the most significant bit of the watermark sample into the low frequency sub-bands of the host image, i.e., sub-band HL3 or LH3. Consequently, 2nd and 3rd bits are embedded into the HL2 or LH2 sub-band, while the remaining bits (from 4th bit to 7th bit) are embedded into HL1 or LH1 sub-band.

Rule 3: For a given decomposition level, only one of the randomly chosen horizontal or vertical wavelet coefficient is marked for each position. The embedding of the imaginary part of a watermark sample is coupled with the embedding of its real part, so that if the embedding position for the 1st bit of the real part is (p, q) in sub-band HL3, the embedding position for the 1st bit of the imaginary part will be then at (p+1, q) in the sub-band LH3.

Rule 4: To enhance the security of the watermarking scheme, a random permutation is performed to scramble the frequency modulation relationships between the consecutive samples of the watermark signal. This random permutation is the *key* of the proposed watermarking scheme. The random permutation can also be interpreted as randomly selected locations for watermark embedding.

To illustrate our watermark embedding approach, let us consider a host image of size 256×256, where each of the sub-band HL3 and LH3 is of size 32×32=1,024 pixels. The complex chirp watermark of 256 samples is to be embedded twice for better localization as elaborated in Section 4. So the total number of watermark samples (including both the real and imaginary parts of each complex watermark sample) to be embedded is then 256×2×2=1,024. Note that the number of watermark samples is the same as the number of pixels in a 3rd- level sub-band. In this context, every position (either from the horizontal or the vertical sub-band) of the 3rd decomposition level can be embedded with one authentication watermark bit. To authenticate an image, one would like to embed at least one authentication bit into ev-

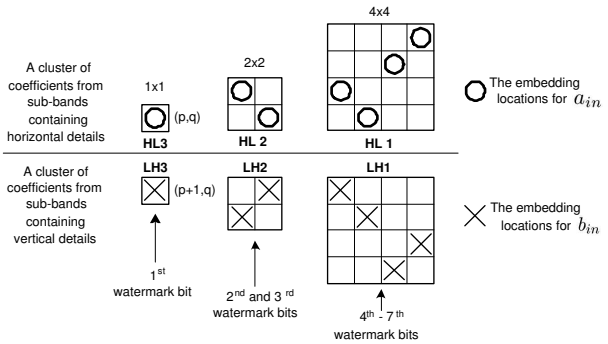

Fig. 2. A pair of clusters of the wavelet coefficients to embed a pair of i-th watermark sample bits of a_{in} and b_{in}, $n = 1, 2, .., 7$.

ery non-overlapping partition of the image. In this example, the minimum authentication block has a size of 8×8, since each position in the 3rd decomposition level is embedded with one watermark bit according to **Rule 3**. Fig. 2 illustrates one possible choice to select the embedding locations for a pair of watermark samples a_{in} and b_{in} considering the above rules.

3.2. Improved QIM Embedding Technique

Let the i-th wavelet coefficient of the original image be C_i. The real-valued coefficients are quantized into two symbols - {0,1}, as shown in Fig. 3. The quantization function $Q(\cdot)$, which maps the wavelet coefficient C_i to 0 or 1, is given by

$$Q(C_i) = \begin{cases} 0, & \text{if } z\triangle \le C_i < (z+1)\triangle \text{ for } z = 0, \pm 2, ... \\ 1, & \text{if } z\triangle \le C_i < (z+1)\triangle \text{ for } z = \pm 1, \pm 3, ... \end{cases} \quad (5)$$

where $\triangle$, the quantization step, is a positive real number.

Let $m \in (a_{in}, b_{in})$ denotes any watermark bit to be embedded into a randomly selected wavelet coefficient C_i based on the location provided by the key. A secure embedding zone, $\mathbb{Z}$, of width r ($r \le \triangle$), is defined in our scheme to improve the extraction performance for the Kundur's QIM method [3]. As shown in Fig. 4, the shaded region represents the zone $\mathbb{Z}$, and the mid point of the zone coincides with the mid point of the quantization step $\triangle$. The rules for embedding a watermark sample bit m are the followings.

Rule 1: If $Q(C_i) = m$ and the coefficient C_i falls within the zone $\mathbb{Z}$ (as illustrated in Fig. 4(a)), then no change in this coefficient C_i is necessary such that the i-th wavelet coefficient of the watermarked image, $\widetilde{C}_i$, is

$$\widetilde{C}_i = C_i. \quad (6)$$

Rule 2: If $Q(C_i) = m$ but the coefficient C_i is not within the zone $\mathbb{Z}$, the following adjustment is done to shift the coefficient into the half portion of the zone $\mathbb{Z}$ that is closer to this coefficient, as shown in Fig. 4(b):

$$\widetilde{C}_i = \begin{cases} \triangle \text{ floor}(\frac{C_i}{\triangle}) + \frac{\triangle}{2} - \xi, & \text{if floor}(\frac{C_i}{\triangle}) = \text{round}(\frac{C_i}{\triangle}) \\ \triangle \text{ floor}(\frac{C_i}{\triangle}) + \frac{\triangle}{2} + \xi, & \text{if floor}(\frac{C_i}{\triangle}) \ne \text{round}(\frac{C_i}{\triangle}) \end{cases} \quad (7)$$

where ξ is a non-negative random real number such that $\xi \le \frac{r}{2}$, and 'floor(·)' is to round the element to the nearest integer towards minus infinity.

Rule 3: If $Q(C_i) \ne m$, the coefficient C_i is then shifted to the secure zone of its nearest neighboring quantization step, as illustrated in Fig. 4(c). The watermarked coefficient $\widetilde{C}_i$ in this scenario is given by

$$\widetilde{C}_i = \begin{cases} \triangle \text{ floor}(\frac{C_i}{\triangle}) - \frac{\triangle}{2} + \xi, & \text{if floor}(\frac{C_i}{\triangle}) = \text{round}(\frac{C_i}{\triangle}) \\ \triangle \text{ round}(\frac{C_i}{\triangle}) + \frac{\triangle}{2} - \xi, & \text{if floor}(\frac{C_i}{\triangle}) \ne \text{round}(\frac{C_i}{\triangle}) \end{cases} \quad (8)$$

where ξ ($\le \frac{r}{2}$) is a non-negative random real number.

The *random* number ξ is used to adjust the coefficient C_i to a *randomly* selected position within the half-portion of the secure embedding zone that is closer to C_i. As a result,

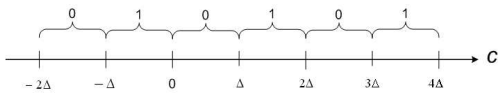

Fig. 3. The quantization function for the wavelet coefficients, C.

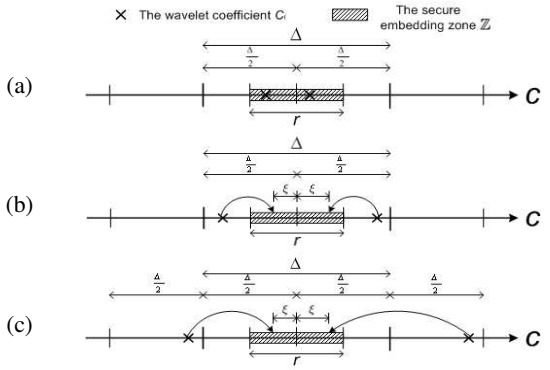

Fig. 4. The rules for embedding of watermark sample bits.

the use of ξ enhances the security level of the embedding scheme, in the sense that it would be hard for a pirate to detect the embedding of watermark bits. Furthermore, the embedding is performed by quantization of C_i to its nearer half-portion of zone $\mathbb{Z}$. The change of a coefficient's value is always smaller than a quantization parameter $\triangle$, which is the amount of change required for embedding in Kundur's scheme [3]. The use of a secure zone $\mathbb{Z}$ can also reduce the bit error rates efficiently at the extraction stage. The narrower the width of the zone $\mathbb{Z}$, the higher possibility to decode the embedded watermark bit correctly.

4. BLIND CONTENT INTEGRITY VERIFICATION

This section presents the *blind* content verification scheme. Neither the original image nor the original watermark is required during extraction and verification. The key includes the random permutation of watermark embedding locations, the wavelet type, the quantization step $\triangle$ and the quantization function $Q(\cdot)$.

The 3-level DWT of the received image is performed using Harr wavelets, where the i-th wavelet coefficient is denoted as $\widetilde{C}_i^{'}$. Based on the embedding locations provided by the key, the selected coefficients are quantized into symbol 0 or 1 using the same quantization function $Q(\cdot)$ for embedding as shown in (5). The watermark bits $\{m^{'} \in (a_{in}^{'}, b_{in}^{'})\}$ are then retrieved from quantization of the selected coefficients by

$$m^{'} = Q(\widetilde{C}_i^{'}). \quad (9)$$

In the absence of the original watermark, the extracted watermark bits are converted back to the watermark sample, $\mathbf{s}_i^{'}$ ($= a_i^{'} + jb_i^{'}$), with the values within the range of [-1, 1], in the following way:

$$\begin{aligned} a_i^{'} &= \frac{w_{ia}^{'}}{63.5} - 1, \text{ where } w_{ia}^{'} = \sum_{n=1}^{7} a_{in}^{'} \cdot 2^{7-n}; \\ b_i^{'} &= \frac{w_{ib}^{'}}{63.5} - 1, \text{ where } w_{ib}^{'} = \sum_{n=1}^{7} b_{in}^{'} \cdot 2^{7-n}. \end{aligned} \quad (10)$$

The magnitude of the extracted watermark samples $\mathbf{s}^{'}$ is then calculated as the quantitative measurement for content authentication, given by

$$mag_i = \sqrt{a_i'^2 + b_i'^2}. \quad (11)$$

The parameter mag_i should be a constant, equal to 1, in the absence of any attack on the protected content. The authentication block of size 16×8 corresponds to the region which is embedded with a pair of real and imaginary watermark samples of the complex FM watermark. Therefore, an image of 256×256 pixels is partitioned into 512 such 16×8 blocks with index labelled column-wise from 1 to 512. Since the same watermark sample is embedded twice as described in Section 3.1, it is possible to further locate the error pixels within a smaller block of size 8×8, using the other extracted copy for the same watermark sample as reference. If the real parts of the two extracted samples differ, the alternation occurs in the upper 8×8 block. On the other hand, the alteration occurs within the lower 8×8 block if the imaginary parts are not matched. Hence, size of the minimum block, which can be authenticated by the proposed scheme is 8×8, however, it depends on whether the repeating watermark sample can be decoded correctly from the corresponding block.

A watermarked Lena image is shown in Fig. 5(a), by setting $\triangle$=4,8,16 for l=1,2,3, respectively, and the width r of the embedding zone $\mathbb{Z}$ as $\frac{\triangle}{2}$, i.e., 2,4,8 for l=1,2,3, respectively. The experiments show that when $r \leq \frac{\triangle}{2}$, an error-free watermark decoding process can be achieved by our modified QIM method. On the contrary, there exists a false detection rate of 0.1% in the original Kundur's QIM method even if there is no attack [2, 3].

An illustrative example is presented next to demonstrate the detection and the localization of alterations without using the original watermark. As shown in Fig. 5(b), the mouth region of the watermarked Lena image is replaced deliberately by a different mouth image. From the image itself without comparing with its original image, it would be difficult to tell the alterations. The detector detects errors in ten 16×8 blocks shown in Fig. 5(d), and those ten blocks form a larger rectangular block corresponding to the marked Lena image with four corner locations: (161, 129), (161, 168), (192, 129) and (192, 168), as shown in Fig. 5(e). The tampered region can be further located in a smaller region by comparing the error samples with the correct repeated samples embedded in the other distant locations. The alterations are further located within the shaded areas. This shaded region matches the deliberately tampered mouth region of the watermarked Lena image as shown in Fig. 5(c).

5. CONCLUSION

This paper presents a wavelet-domain multiresolution fragile watermarking scheme. The FM complex chirp signal is employed as watermark. Both the real part and the imaginary part of the chirp signal are embedded simultaneously into the HL and LH sub-bands, respectively, in a hierarchical manner. The embedding algorithm improves the Kundur's QIM technique. A secure embedding zone is exploited such that all the embedded wavelet coefficients are adjusted

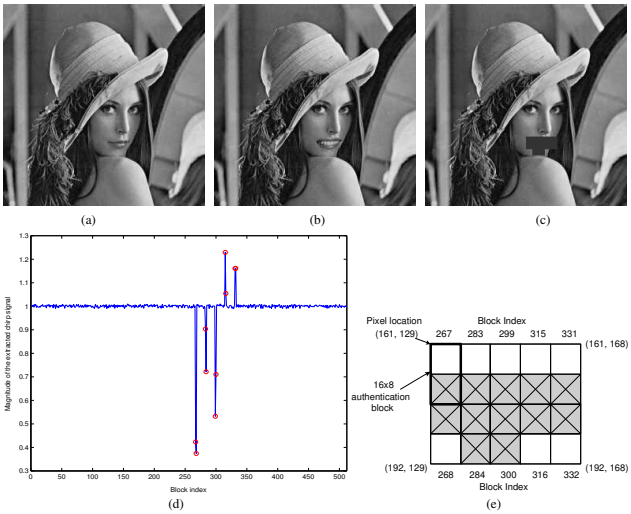

Fig. 5. (a) A watermarked Lena image (PSNR=45.97dB), (b) a tampered watermarked Lena image, (c) the tampered region of (b), (d) the detector response for image (b) ('○' denotes samples of non-one magnitude), (e) localization of the tampered region.

to fall into this zone. It has shown that the proposed scheme successfully reduces the bit error rates for the extracted watermark. The nice feature of using complex chirp signal for authentication is that no information on the original watermark is required. Therefore, an arbitrary FM chirp watermark can be used for authentication of different images. In the absence of the original watermark, the magnitude of extracted watermark is verified to detect and locate the alterations in the content. The application of the proposed scheme for *blind* content integrity verification is demonstrated. The amplitude modulated (AM) complex chirp signal could also be applied. In that case, the feature used for authentication would be its constant frequency rather than the constant magnitude used for FM signal. The principle of using AM complex chirp signal would be the same as using FM complex chirp signal as presented in this paper.

6. REFERENCES

[1] E. T. Lin and E. J. Delp, "A review of fragile image watermarks," *Proc. of ACM Multimedia'99 Multimedia Contents*, pp. 25-29, Oct. 1999.

[2] Ö. Ekici, B. Sankur and M. Akcay, "Comparative assessment of semifragile watermarking methods," *Journal of Electronic Imaging*, 2002.

[3] D. Kundur and D. Hatzinakos, "Towards a telltale watermarking technique for tamper-proofing," *Proc. ICIP*, vol. 2, pp. 409-413, Oct. 1998.

[4] A. S. Kayhan, "Representation and analysis of complex chirp signals," *Signal Processing*, vol. 66, no. 1998, pp. 111-116, 1999.

[5] B. Barkat and F. Sattar, "A new time-frequency based private fragile watermarking scheme for image authentication," *Proc. of the IEEE ISSPA*, 2003.

[6] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21-67, April 1992.

[7] R. E. Ziemer and R. L. Peterson, *Introduction to digital communication*, New Jersey: Prentice Hall, 2001.