A Dual Audio Watermark Embedding Algorithm

Chao Yin [1], ShuJuan Yuan [2]

1 Department of Physics and Electronics
Langfang Teachers University
Langfang, Hebei,China
2. College of Science
North China University of Science and Technology
TangShan, Hebei,China

Abstract - This paper proposes a dual audio watermark embedding algorithm. We embed the watermark into the DWT domain of the audio. First we choose a binary image as the watermark, and encrypt it with a chaotic sequence. The audio signal is decomposed with an appropriate wavelet basis. The low-frequency coefficients are chosen to be embedded with robust watermark which is used to protect the work holder's copyright, and the high-frequency coefficients are chosen to be embedded with fragile watermark which is used for content authentication. Experimental results show that these two kinds of watermarks can function respectively.

Keywords-*Dual-watermarking; robust watermarking; fragile watermarking; Discrete Wavelet Transform; Mean-Quantization; Single-Coefficient Quantization*

I. INTRODUCTION

The network and multimedia technology develop rapidly, and they provide facilities for our life but also bring some negative influence. For example, work tort and falsify problems emerge in an endless stream. Specially audio works deliberately editor and illegally copy happen frequently in recent years. These problems hurt the interest of the copyright holders. So audio watermark technology appears and becomes more and more important [1].By embedding meaningful watermark into the audio, we can protect the audio' copyright, communicate secretly and so on.

Similar to image file, the audio file can also be modified to be embedded information. These modifications can't be removed if the attackers do not destroy the original audio. The main application areas are as follows[2]:

(1) Copyright protection: Audio watermarking is an effective way to protect the audio owner's copyright. Through embedding watermarking into audio signal, we can extract watermark to prove who is the owner when a copyright dispute happens

(2) Use control: It can control the user's authority by embedding watermarking into audio signal. Only someone with a permit can use or copy these products.

(3) Content authentication: By detecting the integrity of the watermark extracted, we can judge whether the audio is a complete one. For example, judge the integrity of an evidence recording in court, or judge the integrity and accuracy of the photo-taking. So our research on it is tremendously significant.

(4) Digital fingerprint: With the rapid development of information technology, digital products can be copied and transmitted easily through simple point-and-click operations. Those operations hurt legal owners greatly. By embedding watermark (digital fingerprint) into every legal copy, we can extract watermark to follow its source when the digital products are copied without permission,

(5) Secret communication: We can embed secret information into the audio carrier. Thus the secret information can't be found readily, and it has a better invisibility and security.

(6) Broadcast monitoring: After business users buying advertisements, they want to know whether the advertisement is broadcast on time. It is very expensive to employ someone to monitor it. So we can embed watermark into the audio part of the advertisement, and thus they can monitor the advertisement broadcasting automatically real-timely with a monitor.

(7) Information sign: We can embed some assistant information into audio signal, such as the time of sound recording and so on.

The characteristics of audio watermark are as follows[3-4]:

(1) Invisibility, which means we can't feel the change of the audio after being embedded with watermark.

(2) Fidelity, which means the watermark can't destroy the normal use of the original audio carrier.

(3) Safety, which means the watermark can't be deleted easily, and can provide copyright evidence.

(4) Robustness, which means the watermark can resist attacks to the audio signals. That is to say, we can still extract watermark even when the audio is attacked.

(5) Watermark capacity, which means the maximum data quantity of the audio system.

According to different purpose, the digital watermark generally falls into two categories: fragile watermark and robust watermark. The fragile watermark is used for protecting the authenticity and integrity of the multimedia contents. This watermark emphasize on the susceptibility to the attacks. It means that the watermark is damaged easily when the digital audio is modified even if the modification is very small, so it's very fragile; The robust watermark is used for protecting the digital products' copyright, such as DVD copyright protection. This robust watermark embed meaningful watermark into the audio and is necessary to resist some kinds of attacks [5].

Most audio watermark embedding algorithms have only one of the functions, copyright protection or content authentication. But in some fields, we want to embed several watermarks into the audio for several purposes: For example, we embed robust watermark for copyright protection; we embed fragile watermark for content authentication; We embed annotation watermark for marking the multimedia. That is to say we can embed several watermarks into a single audio for several purposes [6].

When embedding several watermarks, we should consider the capacity of the carrier audio, as well as the interaction between the different kinds of watermarks, because it's possible that the robust watermark damage the fragile watermark and it's also possible that the fragile watermark damage the robust watermark [7]. If you want to embed robust watermark and fragile watermark in a single audio, there are three schemes: (1) Embed the robust watermark fist, and than embed the fragile watermark; (2) Embed the fragile watermark fist, and than embed the robust watermark; (3) Embed the robust watermark and the fragile watermark at the same time. The two papers [8-9] propose two audio fragile embedding algorithms, and the paper[10] proposes an audio multiple watermarks embedding algorithm.

This paper proposes a dual audio watermarking embedding algorithm. The audio signal is decomposed by an appropriate wavelet basis. The low-frequency coefficients are chosen to be embedded with robust watermark which is used to protect the work holder' copyright, and the high-frequency coefficients are chosen to be embedded with fragile watermark which is used for content authentication. The experiments show that these two kinds of watermarks can work well separately.

II. PROPOSED SCHEME

This paper presents a blind audio dual-watermark embedding algorithms in discrete wavelet transform domain of the audio signal. Why we embed watermark in the DWT domain of the audio?

Because discrete wavelet transform has its own characteristic features:

(1) Multiscale, which can be used to analog characteristics of human auditory system;

(2) Wavelet transform is adept at analyzing audio signals with its time-frequency localization ability;

(3) Various basic functions provide us with flexible and diversified schemes;

(4) It has fast algorithms.

Based on the characteristics above, we choose discrete wavelet transform domain to embed watermark. So it is a very effective tool for us to research on embedding algorithms.

(1) Choose a planar image as watermark, and convert it into a one-dimensional sequence in order to embed watermark bits into audio signal in mono.

$$V = \{v(k) = w(i,j), 0 \le i < M1, 0 \le j < M2, k = i \times M2 + j\} \quad (1)$$

Where the one-dimensional sequence is represented by V; $w(i,j)$ is a watermark bit; M_1 is the number of lines of the planar image; M_2 is the number of rows of the planar image.

(2) In order to ensure the safety of the watermark, we need to disorder the watermark with a chaotic sequence. Chaos system[11-12] is sensitive to initial value and it has many good characteristics similar with white noise. So it is usually used to encrypt image. We gets the binary chaotic sequence of $\{h(k)\}$ using Logistic chaos mapping, and assume that $\{v_p(k)\}$ represents the disordered watermark sequence.

$$v_p(k) = v(k) \oplus h(k) \quad (2)$$

(3) Perform discrete wavelet transform process on the audio signal. The audio signal is decomposed with an appropriate wavelet basis. A_4 is a set composed of low-frequency coefficients after the audio signal is decomposed for 4 levels using Daubechies-3 wavelet basis, and D_1 is a set composed of high-frequency coefficients after the audio signal is decomposed for 1 level using Daubechies-3 wavelet basis.

(4) Embed robust watermark using mean-quantization method[13-14]. The elements of A_4 are chosen to be embedded watermark in order to guarantee the robustness of the watermark.

Here A_4 is a set whose elements are chosen to be embedded or modified. We change the set of A_4 into a matrix, whose number of lines is represented by K and the number of rows is represented by L. And then each row's mean value which is represented by $\overline{Y}(l), (l = 0,1,\cdots L-1)$ is calculated. Calculate $z(l)$ using the following equation.

$$z(l) = \lfloor \overline{Y}(l) / Q_1 + 1/2 \rfloor, \ l = 0,\cdots L-1 \quad (3)$$

Here the symbol of $\lfloor \ \rfloor$ represents the operation of rounding down, and Q_1 is a parameter predetermined, called quantization step.

We embed watermark like this:

If

$$z(l)\%2 = w(l) \quad (4)$$

Where the symbol of % represents the operation of modular arithmetic, and $w(l)$ is the remainder.

thus

$$\overline{Y^*}(l) = Z(l) \times Q_1 \quad (5)$$

Where $\overline{Y^*}(l)$ stands for the new mean value of each row.

else if

$$z(l)\%2 \neq w(l), \tag{6}$$

$$z(l) = \lfloor \overline{Y}(l)/Q_1 \rfloor \tag{7}$$

$$\overline{Y^*}(l) = (z(l)+1)\times Q_1 \tag{8}$$

else if

$$z(l)\%2 \neq w(l) \tag{9}$$

$$z(l) \neq \lfloor \overline{Y}(l)/Q_1 \rfloor \tag{10}$$

$$\overline{Y^*}(l) = (z(l)-1)\times Q_1 \tag{11}$$

(5) Embed fragile watermark using single coefficient-quantization method. The elements of D_1 are chosen to be embedded with fragile watermark.

The Single-coefficient quantization principle[15-16] diagram is shown in Fig.1.

Here f is a coefficient to be embedded or modified, and f' is a coefficient having been modified.

We separate the coefficients into two kinds of intervals: A and B. k is an integer and ρ is a quantization step. w is a watermark bit. If w is 1, f is quantized to the nearest middle of A. Otherwise, f is quantized to the nearest middle of B.

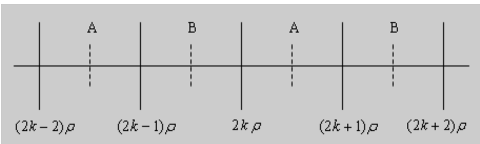

Figure 1. The quantization principle diagram

Modular arithmetic is expressed by following equations.

$$m = f \bmod \rho \tag{12}$$

$$r = f - m\ \rho \tag{13}$$

Where m is the quotient after modular arithmetic, and r is the remainder.

The quantization process is as follows:

If $f \geq 0$ and $w=1$, thus

$$f' = \begin{cases} (2k+\frac{1}{2})\rho & m = 2k \\ (2k+\frac{1}{2})\rho & m = 2k+1 and |r| \leq \frac{1}{2}\rho \\ (2k+2\frac{1}{2})\rho & m = 2k+1 and |r| > \frac{1}{2}\rho \end{cases} \tag{14}$$

If $f \geq 0$, and $w=0$, thus

$$f' = \begin{cases} (2k+1+\frac{1}{2})\rho & m = 2k+1 \\ (2k-\frac{1}{2})\rho & m = 2k and |r| \leq \frac{1}{2}\rho \\ (2k+2\frac{1}{2})\rho & m = 2k and |r| > \frac{1}{2}\rho \end{cases} \tag{15}$$

If $f < 0$, and $w=1$, thus

$$f' = \begin{cases} -(2k+1+\frac{1}{2})\rho & m = -(2k+1) \\ -(2k-\frac{1}{2})\rho & m = -2k and |r| \leq \frac{1}{2}\rho \\ -(2k+2\frac{1}{2})\rho & m = -2k and |r| > \frac{1}{2}\rho \end{cases} \tag{16}$$

If $f < 0$, and $w=0$, thus

$$f' = \begin{cases} -(2k+\frac{1}{2})\rho & m = -2k \\ -(2k+\frac{1}{2})\rho & m = -(2k+1) and |r| \leq \frac{1}{2}\rho \\ -(2k+2\frac{1}{2})\rho & m = -(2k+1) and |r| > \frac{1}{2}\rho \end{cases} \tag{17}$$

By above knowable, the maximum error of coefficient after quantization is ρ, and we call it quantified coefficient. The bigger the quantified coefficient is, the stronger the watermark's robustness is and the worse of the invisibility of the watermark. In contrast, the smaller the quantified coefficient is, the better the invisibility of the watermark is and the weaker of the robustness is. So we must choose a appropriate one.

(6) Perform discrete wavelet inverse-transform.

Maybe, the audio signal is attacked, so we need to extract the watermark when necessary.

The extracting process is as follows:

The block diagram of extracting process is shown in Figure.2.

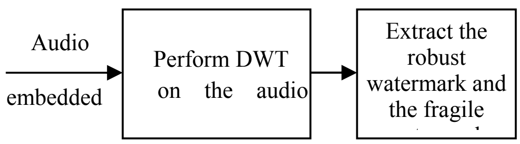

Figure 2. The block diagram of extracting process

(1) Perform discrete wavelet transform on the audio signal. The audio signal being embedded is also decomposed with the same wavelet basis as the one when embedding.

(2) Extract the robust watermark from low-frequency coefficients which is represented by a set of A_4, and change it

into a matrix; Calculate the mean of each column whose mean value is represented by $\overline{Y'(l)}$.

The extracted watermark bit is represented by $w'(j)$. Calculate $w'(j)$ using equation (18).

$$w'(j) = \lfloor \overline{Y'(l)} / Q1 + 1/2 \rfloor \%2, \quad l = 0, \cdots L-1 \qquad (18)$$

At last, we change $w'(j)$ to a two-dimension image. Thus we get the robust watermark.

(3) Extract the fragile watermark from high-frequency coefficients which is represented by a set of D_1.

If the value of the high-frequency coefficient is in the interval of A, then we get the watermark bit of 1; Otherwise, we get the watermark bit of 0. Then we change the watermark bits to a matrix which is the extracted fragile watermark.

III. SIMULATION RESULTS

In this part, we give the experimental results based on MATLAB.

We choose a 12 seconds length' mono pop music in these simulation experiments. Its sample frequency is 44.1 kHz and its resolution is 16bit. Considering the complexity of this algorithm, the audio signal is decomposed for 4 levels using Daubechies-3 wavelet basis. The watermark is a 64×64 binary image. The value of mean-quantization step is 0.02, The value of single-quantization step is 0.001.

SNR is used to evaluate the invisibility of the watermarks [17].

$$SNR = 10 * \log_{10}\left(\frac{\sum_{k=1}^{N} I^2(k)}{\sum_{k=1}^{N} \left[I(k) - I'(k) \right]^2} \right) \qquad (19)$$

Where $I(k)$ stands for an original audio sampling value; $I'(k)$ stands for an audio sampling value having been modified after embedding; N stands for the number of sampling value.

In these experiments, the SNR of the audio signal having embedded is 37.9730 dB, and we can't distinguish the differences between the original audio and the audio having been embedded.

To test the robustness of the watermark, the music is attacked as follows [18]:

Add white Gaussian noise whose mean value is zero and variance is 0.01 to the audio signal embedded.

Filter the audio with a low-pass filter, whose cut-off frequency is 11.025 kHz;

Requantization: change the audio's resolution from 16 bits to 8 bits, and then change its resolution from 8 bits to 16 bits again;

Resample the audio. Change the audio' sampling frequency to 22.05 kHz, and than change the audio' sampling frequency to 44.1 kHz again.

Denoise the audio with one-dimensional discrete stationary wavelet transform method ;

MP3 compression: Compress the audio's bits rate into 128 Kbit/s.

The robust watermarks extracted from the audio attacked are shown in Figure 3.

The fragile watermarks extracted from the audio attacked are shown in Figure 4.

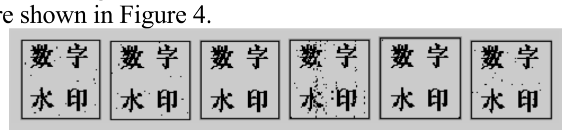

Figure 3. The robust watermarks

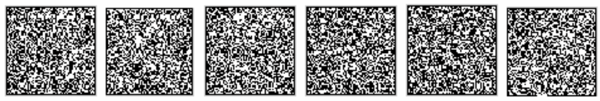

Figure 4. The fragile watermarks

There are three methods to evaluate the quality of the watermark.

(1) Subjective listening testing: To test the quality of the watermark, we use Subjective Difference Grade (SDG) for subjective listening testing [19], which is shown in table I.

In these experiments, we let ten persons listen to the extracted music. At last, these people gave the grade of 0.0. The nearer the given mean of grade approaches zero, the better the invisibility is.

TABLE I SUBJECTIVE DIFFERENCE GRADE (SDG) FOR SUBJECTIVE LISTENING TESTING

SDG	Description
0.0	Imperceptible
-1.0	Perceptible, but not annoying
-2.0	Slightly annoying
-3.0	Annoying
-4.0	Very annoying

(2) Bits error ratio (BER)

We also can use BER to evaluate the robustness of the watermark. BER is the ratio of the wrong bits number of the watermark to the total number bits of the watermark.

(3) Normalized cross-correlation coefficient(NC)

NC is used for evaluating the similarity between the original watermark and the extracted watermark.

$$NC(W, W') = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w(i,j) w'(i,j)}{\sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w(i,j)^2} \sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} w'(i,j)^2}} \qquad (20)$$

Where $w(i,j)$ stands for an original watermark bit; $w'(i,j)$ represents a extracted watermark bit. M_1 is the number of lines of the watermark ; M_2 is the number of rows of the watermark.

IV. CONCLUTIONS

This paper proposes a dual audio watermarking embedding algorithm. Choose a two-dimensional binary image as the watermark, and encrypt it with a chaotic sequence to ensure the security of watermark. Embed the watermark into the discrete wavelet transform domain (DWT) of the audio signal: The audio signal is decomposed by an appropriate wavelet basis. The low-frequency coefficients are chosen to be embedded with robust watermark which is used to protect the work holder' copyright, and the high-frequency coefficients are chosen to be embedded with fragile watermark which is used to authenticate the content. Experiments show that these two different kinds of watermarks can function respectively.

ACKNOWLEDGMENT

This research is supported by Self-financed Program of Science & Technology Research and Develop of Lang Fang City (No.2014011064). I would like to express my gratitude to my department leaders and colleagues. I was thankful for their help and support.

REFERENCES

[1] Chen Haiman., "Research on digital audio dual-watermarking algorithms", Master paper, Northeast Petroleum University, pp. 49-69, 2013.

[2] Chen Pan. "Digital audio watermarking algorithm". Master paper, Shanghai Jiao Tong University, pp.13-14, 2009.

[3] Wang Hongxia, "Study on adaptive audio watermark algorithm",.Southwest Jiaotong University, pp.7-8, 2011.

[4] Arnold M. "Audio watermarking:features，applications and algorithms". IEEEI nternational Conference on Multimedia and Expo (ICME), pp. 1013-1016, 2000.

[5] Wang Yan, "Research on robust watermarking algorithms". Master paper, Fudan University, pp. 5-6, 2006.

[6] Wang Rangding, Xu Dawen. "Multiple audio watermarks based on lifting wavelet transform". Journal of Electronics & Information Technology. Vol 28 ,No10, pp.1820-1826, 2006.

[7] Hua Xiansheng, Shi qingyun, "To embed multiple watermarks simultaneously". Acta Scientiarum Naturalium Universitatis Pekinensis.,vol 37, No3, pp.305-314, 2001.

[8] Fu Xingbin. "Digital audio fragile watermarking algorithm based on mean-quantization". Applied Science and Technology. Vol 32,No8, pp.17-19. 2005.

[9] Wang Xiangyang,Yang Hongying. "A novel fragile digital audio watermarking based on DCT and HAS". Journal of Dalian University of Technology, pp. 23-26. 2003.

[10] Yang Yang, Chen Xiaoping "A digital audio multi-watermarking algorithm in transform-domain".Microelectronics Computer. vol21, No6, pp.63-66, 2004.

[11] Liu Sunan, Fang Zhihong, Li Shuzhi, "Audio watermarking algorithm in complex cepstrum domain based on chaotic encrypting and repeatencode". Computer Engineering and Applications,vol 46, No4, pp.79-82, 2010.

[12] Fan Mingquan. "Study on techniques of digital image and audio watermarking", Master paper, Southwest Jiaotong University, pp.19-19, 2010.

[13] Chen Jiner, Wang Rangding, Xu Ji, "Mean Quantization Audio Watermarking Algorithm Based on Wavelet Transform", Application Research of Computers, pp. 107-111, 2006.

[14] Fu Xingbin. "Digital audio fragile watermarking algorithm based on mean-quantization", Applied Science and Technology.vol.32 No8, pp17-19, 2005.

[15] Wang hong, Huang Junbo. "A digital Audio Watermarking Algorithm Based on Quantization in DCT Domain", Computer and communication. Vol.22,No1. pp.26-28, 2004.

[16] Zhao Chunhui, Li fuchang. "A Blind Audio Watermark Algorithm Based on DWT and DCT". Journal of Electronics and Information Technology. Vol 25,No 11, pp.1570-1574, 2003.

[17] Yan Qingqing. "Study of Digital Audio Watermarking Technology Based on Transform", Master paper, Hunan Normal University, pp.14-14, 2011.

[18] Zhang Min. "Research on Technology of Digital Audio Watermarking", Master paper, Harbin Engineering University, pp. 61-67, 2004.

[19] Xiang Shijun. "Robust Audio Watermarking Algorithms", doctoral paper, pp.7-7, 2006.