Qing-qing WEI*, Ru-chun CUI, Zi-chao CHEN, Ya LI

A DCT-domain-based Research and Application of the Algorithm of Digital Audio Watermark

Abstract: Digital watermarking technology has important application value in the aspect of authenticity identification, covert communication, hidden identification and electronic identity authentication. Digital watermarking is the important means to solve the problem of copyright protection of audio media. Through the analysis of the basic characteristics and basic methods of audio digital watermark embedding and extraction technology, the paper further discusses the domain audio digital watermarking algorithm that is based on the discrete cosine transform (DCT) and with the aid of MATLAB tool, A specific implementation method of the algorithm based on the two valued gray image as a watermark is proposed. This approach implements the copyright identification of specific application Settings for audio signal.

Keywords: audio digital watermarking; DCT domain algorithm; watermarking embedded; watermarking extraction

1 Introduction

Discrete Cosine Transform (DCT) is the best way for image conversion. And it has lots of advantages. First, DCT is a change of orthogonal, and it can convert 8*8 spatial expression to a frequency domain, which only needs a small number of data points to present the image; Second, DCT coefficient can easily be quantified, so it can get a good block compression; Third, DCT algorithms performs very well, and it has a fast algorithm, such as fast Fourier Transform for efficient operation, so it is easy to implement, whether in hardware or software; Fourth, DCT algorithm is symmetrical, so inverse DCT algorithm can be used to unzip image [1].

Digital watermarking system consists of digital watermark generation, digital watermark embedding, watermark detection, digital watermark extraction and the attack on the part of the digital watermarking system [2]. The paper mainly focuses on the watermark generation, embedding, detection and extraction process.

***Corresponding author: Qing-qing WEI,** Department of Computer Science, Foshan University, Foshan, Guangdong, China, E-mail: fswqq@126.com
Ru-chun CUI, Ya LI, Department of Computer Science, Foshan University, Foshan, Guangdong, China
Zi-chao CHEN, Department of Educational Information Technology, Foshan University, Foshan, Guangdong, China

2 The Construction of Watermark

2.1 the Generation of Watermark

Suppose that watermark is a binary image of m1*m2, it can be expressed as W={w(i,j),0≤ i<M_1, 0≤ j< M_2}, in which w(i,j)∈{0,1}. In order to embed the two dimensional binary image into the one dimensional audio signal, we take it into dimension reduction processing, then two dimensional will be reduced to one dimensional:

$$V=\{v(k)=w(i,j),0\leq i<M_1,0\leq j< M_2,k=i\,M_2+j\} \quad (1)$$

In order to eliminate the sequence V in adjacent element correlations (and) improve the watermark embedding stability, Arnold scrambling transform is used to generate pseudo random sequence of V in all the elements of the pseudo-random sequence:

$$V_p=\text{Arnold}(V)=\{Vp(k)=V(k'),0\leq k,k' < M_1M_2\} \quad (2)$$

Through the pseudo random sorting operation, the v' elements in V sequence is moved to the position of the k-th in VP sequence. In order to further improve the resistance of the watermark damage, the watermark sequences spreads spectrum modulation, with Key for random seed producing m series {m (k)}. Use {m (k)} and watermark sequences {Vp (L)} to spread spectrum modulation and get a sequence {s (k)}, spread spectrum factor for m, S(k)= Vp([k/m]) Λ m(k),Vp(k),m(k)∈{0,1}(0≤ k<m M_1M_2-1),type in "Λ" for XOR operator, [x] for not more than x nearest integer. To spread the spectrum modulation is to use the bandwidth for signal-to-noise ratio and, to use the increase of watermark channel capacity for the improvement of the stabilization algorithm [3].

2.2 the Processing of Watermark Embedded

Watermark embedded process can be divided into three steps: DCT transformation, embedded watermark weight and DCT inverse transform. If A is set as the original digital audio signal, and the number of data as K, it can be expressed as:

$$A=\{a(k),0\leq k\leq K\rightarrow 1\} \quad (3)$$

In this formula, a (k) ∈ {1, 2,......, (a 1)} is the range of data value of No. K, and P, the number of bits used for each hat data. For the sake of argument, the original digital audio signal is put into two parts:

$$A = Ae+Ar \quad (4)$$

In the expression, Ae is relevant to watermark while Ar not.

Because watermark is embedded into the original digital audio signal as noise, such embedding should not interfere with the use of digital audio signals as the prerequisite. Generally, to improve the robustness of the embedded watermark, one embedding needs at least N audio data [4].

In consideration of the spread spectrum of modulation factor of spread spectrum m, if all the M1*M2 pixels of the watermark is embedded, the size of Ae relevant to the embedded watermark is N.

2.2.1 DCT Transformation

The processing of the audio data related to the embedded watermark and divided into segments of $M_1*M_2$and the DCT Transformation of each segment results in Dk(U) expressed as follows:

$$Dk(u)=DCT(Ar(k))=\{dk(x),0\leq x\leq N\rightarrow 1,0\leq k\leq M_1\times M_2\} \quad (5)$$

In the expression, dk(x) is the factor of No. x in the discrete cosine transformation of the audio data segment of No. k [5].

2.2.2 Watermark Weight Embedding

First of all, confirm the area for digital audio embedding in the discrete cosine transformation. That is, choose DC component and (m-1) low frequency AC coefficient as the embedding point in Dk(u) to embed m values of modulation sequence {S (k)}, namely, a pixel.

And then modify the selected wavelet coefficients dk(x) and embed in a sequence of modulation element S (k) :

$$dk'(x)= dk(x)+\beta S(k) \quad (6)$$

β is predetermined quantitative coefficient, which is used to regulate the embedding depth and whose value should be given according to the specific conditions of the embedded watermark because the watermark robustness will be poorer if the value is too small or otherwise will reduce excessively the value of original digital audio signal [6].

2.2.3 DCT Inverse Transformation

The digital audio signal containing watermark information is achieved when the inverse transformation of the embedded watermark to audio data section d (u) in the discrete cosine is finished. The expression is as follows:

$$A'e=IDCT(D'k(u))=\{d'\ k(u),0\leq k\leq M_1\times M_2\} \quad (7)$$

The digital audio signals containing watermark is obtained when Ae is substituted by A e' and in (7). The expression is As=A'e+Ar.

2.3 The Extraction and Detection of Watermark

Only when the expected watermarking information can be tested and extracted in the watermark embedded into audio signal can the protection of copyright and the integrity be protected.

Watermark extraction can be expressed as W = F (I ‘, I), in which F is the function of watermarking extraction. Original audio signal is an option for watermark extraction and detection. It will be difficult for watermark technology to achieve its product release and network communication as the use of original audio signal in watermark detection is a defect, which explains why current watermark detection algorithm is carried out without the participation of original audio signal.

Still, the original digital audio signal and the watermarked signal need to be preserved for testing use. The detection process and the watermark embedded process are on the opposite.

If Aw is the digital audio signal to be tested, the inspection process of watermark extraction can be described as follows:

Segment the digital audio signals Aw, namely, Aw = Awe + Awr, then transform the discrete cosine containing watermark segments:

$$Dwk(u)=DCT(Awe(k))=\{dwk(x),0\leq x\leq N\rightarrow 1,0\leq k\leq M_1\times M_2\} \tag{8}$$

As for the DC component and (m-1) AC frequency components, use the original audio signal to find the hidden m binary image sequences of spread spectrum positions, which results in) (k) = [k (x)-dk (x)] /β

According to the m sequence produced by the individual key, the solution to and the expansion of {(k)} results in another sequence of {p(k)}, which is done in way of taking exclusively or operation, namely:

$$r'p(k)= S'(k)\ \Lambda\ M_1 m(k) \tag{9}$$

Transform the dimension of {p (k)}, that is, change the sequence of 1d into binary images of 2d:

$$Ws=\{ws(i,j)= v'\ p(k),0\leq i<M_1,0\leq j<M_2,k=i\ M_2\ +j\} \tag{10}$$

Compare the detected image of watermark Ws with the original one W to discern the false from the genuine watermark through the following formula:

$$sim(W,Ws)=(W^*Ws)/SQRT(W^*Ws) \tag{11}$$

3 DCT Algorithm

Discrete Cosine Transform (DCT) is one of mathematical operations that is closely related to Fourier transform. In Fourier series expansion, if an unfolded function is an even one, it only contains cosine term in the Fourier series, and is called discrete cosine transform when discretized [7].

3.1 DCT algorithm analysis

The specific process of the extraction and detection of the embedded watermark through DCT algorithm is shown in Figure 1 and Figure 2.

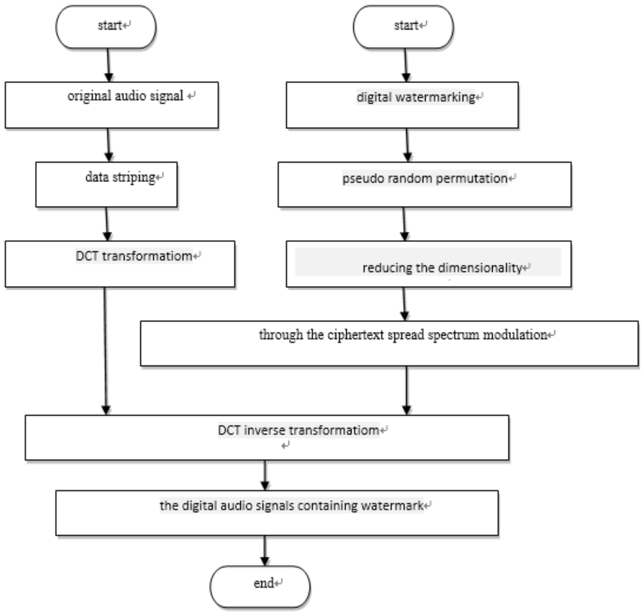

Figure 1. DCT watermark algorithm

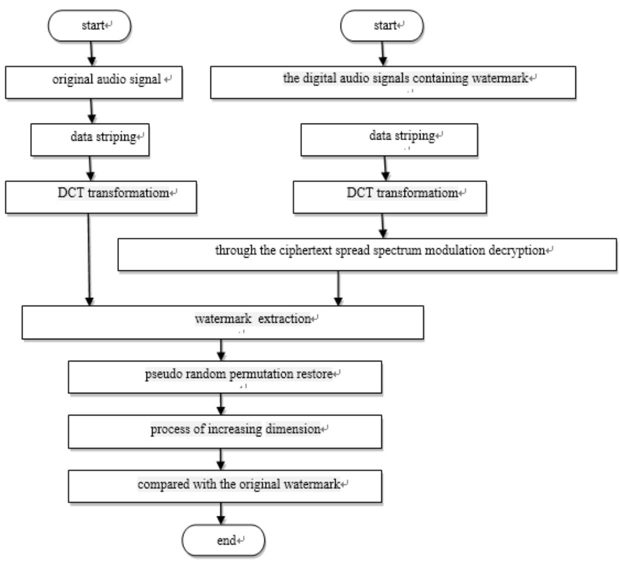

Figure 2. DCT Algorithm for watermark extraction and detection

3.2 Simulation and Testing

In the study, the binary gray image used as a watermark (as shown in Figure 3) is a 278*278 pixel grayscale one. The procedures to improve the robustness and the anti-interpolation of the embedded watermark are as follows:

1. Valuing the quantitative coefficient of β in function 1-6 1.5;
2. Using proper Arnold transform.

Theatermarkimage become illegible when the pictures

Undergo six times of Arnold transform (as shown in Figure 4 and Figure 5), but they restitute when undergone adequate iteration cycles. Because the watermark image size is 278*278, the restitution will be achieved after 24 times of iteration [8].

Figure 3. Binary gray watermark

Figure 4. The watermark undergone one Arnold scrambling

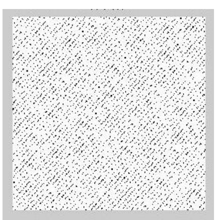

Figure 5. The watermark undergone six times of Arnold scrambling

3. In addition, we can use the digital audio signal of 1411 kbps in the wav format as the original one. First, the original digital audio signal (A = Ae + Ar) is divided into two parts: Ae and Ar, in which Ae is the embedded watermark, and Ar is the redundancy part. The embedding of watermark has no effect on Ar. Second, divide Ae into M1 * M2 N- bytes of audio data segment (M1 is the watermark image pixels, and M2, the width watermark image pixels. In this paper, value the watermark image 8 and size of 278 * 278). The audio signals containing digital watermark can be achieved when the watermark undergoes six times of Arnold scrambling and receives a dimensional decrease of one dimension binary sequences and all audio data segments each embedded with a watermark pixel undergo both discrete cosine and reverse discrete cosine transform [9].

The results of the experiments are shown in Figure 6. It is clear that the audio signals embedded with watermark become a bit stronger than the original ones. There is little significant distortion of sound quality when such signals are normally played, which shows the anti-effect and the concealment of watermark.

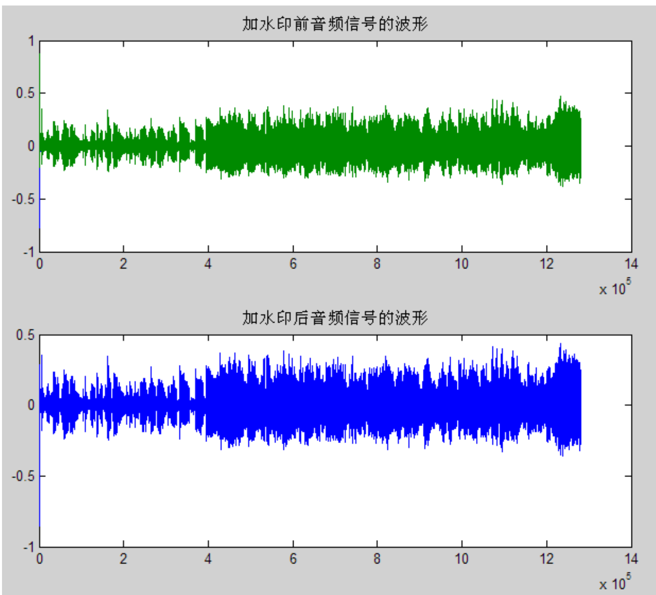

Figure 6. Figure of waveform in experiment

In the process of watermark extraction and inspection, MATLAB is used to extract a 1d binary watermarking signal which becomes a 2d one that has previously undergone six times of Arnold scrambling transformation and the original watermark is obtained when the 2d binary watermark signal undergoes 18 times of replacement. The experiment results shown in Figure 7 show that the watermark has been successfully embedded in the audio and can be correctly extracted [10].

a) The original image b) the watermark extraction

Figure 7. Contrast of extracted experiment watermark

3.3 Audio Attack

Compact the audio file embedded with watermark into wma form before transform it into wav one.

The signal waveform obviously change if the audio file is compacted, as shown in Figure 8. The watermark extracted from the compacted audio file is quite different from the original one as shown in Figure 9.

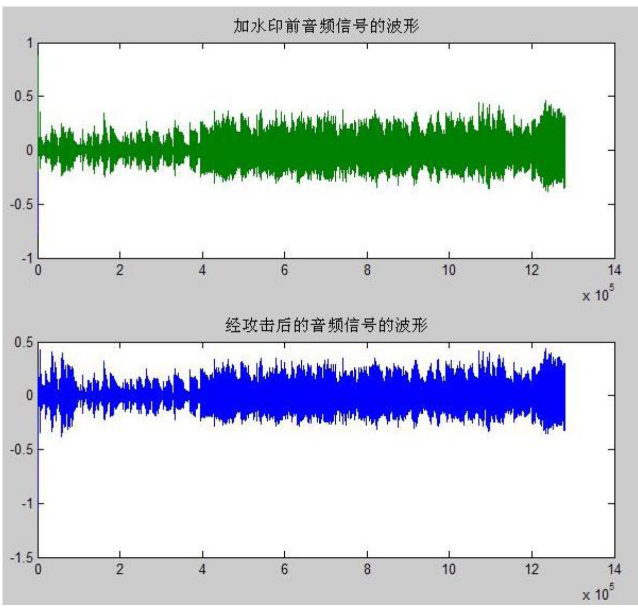

Figure 8. The different waveform of audio signal before and after compact attack

a) the original watermark image b) the watermark imageafter being attacked

Figure 9. The different watermark before and after compact attack

4 The Experimental Results and Analysis

In light of the characteristics of embedding watermark into audio signal and the outstanding advantages of the discrete cosine transform used in audio data processing, a solution to the embedding of binary gray 2d image watermark by using discrete cosine transform (DCT). The advantage of image watermark is to offer users an intuitive and friendly analytical result while that of discrete cosine transform (DCT) is to provide users with software implementation that reflects the superior performance of digital and audio signals [11].

1. Because the watermark is, in the form of noise, embedded into audio signal, the embedding of watermark into audio signal should be based on a number of audio signals or the audio quality will be damaged because of the effect of watermark noise on the audio signals. It has been proved in this paper that the proportion between the watermark and the audio signals should not be higher than 1:8 – that is, the watermark image resolution is 278*278, and the shortest value of audio signal is L = 278*278*8 = 618272 – or the shortest length of audio signals will be limited if such proportion is too low and the length of watermark signals are limited. That's why an additional analysis of the length of audio signals before the embedding of watermark.
2. The spreading spectrum and encryption of digital watermark is to improve the robustness of digital watermarking, but in practice, the robustness of digital watermarking provided by the software suggested in the paper does not perform well, this is basically because of the contradictory choice of abandoning the higher signal-to-noise ratio to reserve the integrity and the origin of the audio signals and prevent the audio signals from excessive distortion when they are embedded with watermark.

5 Conclusion

Through the analysis of the algorithm and simulation tests, it can be seen that using the audio digital watermark algorithm based on discrete cosine transform (DCT), with the aid assistant of MATLAB tool, it is effective to embed audio signals with watermark which is based on binary gray image, whose corresponding algorithm and application are illustrated in the paper for setting copyright marks by using audio signals.

References

[1] Jia Yong-Hong. Computer image processing and analysis [M]. Wuhan: Wuhan University Press.. 2001.9
[2] Xu Wen-Li, Wang Ming-Yu. Digital watermarking technology and its application[M]. Beijing: Publishing House of Electronics Industry. 2013.1
[3] Sun Sheng-He, Liu Zhe-Ming, Niu Xia-Mu etc. Digital watermarking technology and application [M]. Beijing: Science Press. 2004.
[4] Zhong Hua, Zhang Xiao-Hua, Xiong Li-Cheng. Digital watermarking and image authentication algorithm and its application [M]. Xian: Xian University of Electronic Science and Technology Press, 2006.
[5] Yang Yu. Information hiding and digital watermarking experiment tutorial [M]. Beijing: National Defense Industry Press. 2010.1
[6] Wang Yu-Ming, Li Hui, Liang Chuan-Jia. Information theory and coding theory [M]. Beijing: Higher Education Press. 2005
[7] Yang Yi-Xian, Niu Xin-Xin. Theory and applications of Digital watermarking [M]. Beijing: Higher Education Press.2006.3
[8] Zhang Xian-Ming. MATLAB language and its application cases [M]. Nanjing: Southeast University Press. 2010.9
[9] Zhang Liang, Guo Shi-Jian Wang Bao-Shun etc. MATLAB7.X system modeling and simulation [M]. Beijing: Posts & Telecom Press. 2006.
[10] Meng Yi-Zheng. MATLAB and technique [M]. Beijing: Science Press. 1999.11
[11] Qiu Zhi-Cong. Encryption and decrypt skill practice [M], Beijing: China Railway Publishing House). 2005.1