A Digital Audio Watermarking Scheme Using Majority Energy Distribution Technique in Cepstrum Domain

Wen-Yuan Chen, Jhen-Syun Liao, Chin-Ho Chung
Department of Electronic Engineering,
National Chin-Yi University of Technology, Taichung, Taiwan, R.O.C.
Department of Electronic Engineering,
Ta Hwa Institute of Technology, Hsin-Chu, Taiwan, R.O.C.

E-mail:cwy@ncut.edu.tw, lifrscamus@yahoo.com.tw, chc@thit.edu.tw

Abstract

In this paper, a novel audio-watermarking scheme in cepstrum domain using majority energy distribution techniques is proposed. Cepstrum domain transform is firstly hired to transfer the audio signal from time domain to cepstrum domain. In our approach, three cepstrum domain coefficients of a frame are treated as a group of a signal, and majority energy distribution of the groups is adopted for the secret bit embedding. On the other side, a logo image "NCUT" in chinese is used as a watermark. Meanwhile, a chaotic mechanism hashes the watermark image for enhancing the security.

In order to demonstrate the effectiveness of the proposed scheme, simulation under various conditions were conducted. Several kinds of music like piano music, classical music, and pop music are used as the test audio signals. The experimental results show our scheme can sustain all kind of attacks like added noise, MP3 compression, and re-sampling.

Keyword: Majority Energy Distribution, Cepstrum Domain Transform (CDT), MP3 audio compression, Audio watermarking scheme, Fast Fourier Transform (FFT).

1. Introduction

In application, digital watermarking scheme can be classified into image watermarking scheme, video watermarking scheme and audio watermarking scheme. Several audio watermarking techniques have been reported. However, the most audio watermarking methods have some common problems as follows [1-6]: (a). The amount of hidden information is small; (b). Some of them detect watermark with the help of the original digital audio signals; (c).They all embeds the watermark by choosing the individual transform coefficient result in instability.

Wang and Cui [7] introduce a new digital audio watermarking algorithm that is based on modifying the whole low-band wavelet coefficients under integer wavelet transform. There are features of his approach as follows: (a). Embedded watermark as meaningful gray image; (b). Blind detection without resorting to the original digital audio signals. (c). Modified the wavelet coefficients as a whole to increase the robustness. (d) Integer lifting wavelet transforms to solve rounder-off error problem.

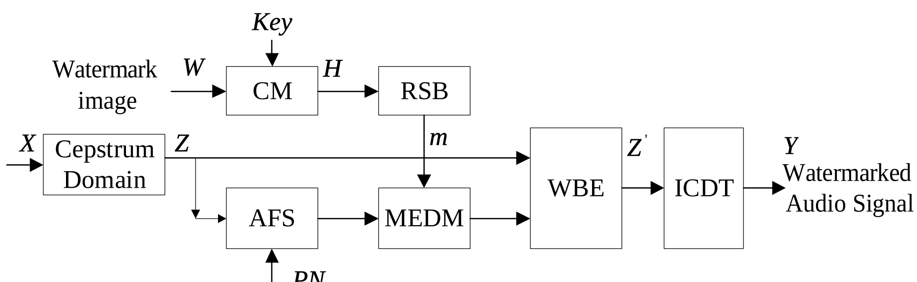
According the watermarks ability against attacks, the algorithms used in the transform domain is better than that used in the time domain. The complex cepstral analysis is a homomorphic mapping and is the most effective way to characteristic extraction in audio identification. S.K.Lee [8] perform the audio watermarks embedding according to the distribution of cepstral coefficients and the frequency masking features of Human Audio System (HAS). Meanwhile, the C.T.Hsith [9] discussed the detection of audio blind watermarking in cepstrum domain using the conception of energy characteristic radix in frequency domain.

Tang et. al [10] process the audio watermark embedding based on the human auditory characters using the wavelet and cepstrum transform. In his paper, the watermark embedding realized to embed binary image watermarks into original audio signal. And the experimental results show that this algorithm had a better robustness against common signal processing such as noise, filtering, resampling and lossy compression.

This paper is aim at audio watermarking scheme. We use cepstrum transform and majority energy distribution technique to achieve the goal. The remainder of this paper is organized as follows. In section 2, the proposed concealing algorithm is presented. The watermark extracting process is presented in section 3. Empirical results are presented in section 4. Finally, we conclude this paper in Section 5.

2. Watermark embedding Algorithm

In this paper, the watermark signal is a significant binary image. Chaotic mechanism and dimension reduction are performed on the watermark signal. The overall concealing process of our proposed scheme is shown in Fig. 1. The watermark W is first transferred to H by chaotic mechanism (CM) using a fast pseudorandom number traversing method with a private password (key) to enhance the security. Following, a dimension reduction step is used to convert hashed watermark into a binary sequence $\{m\}$ by rearrange to series bits (RSB) process. On the other hand, the audio signal X is transformed to Z by cepstrum domain transform (CDT). Successively, an audio frame is selected from Z by using the PN sequence for increasing security. Three CDT coefficients are selected serve as a group and a group embedding a secret bit by the majority energy distribution modulation (MEDM) technique. In the MEDM modulation, the majority energy distribution of the groups signal E is modulated by m into E'. After MEDM, the mean of group E' contains the secret information is then embedded into Z' at its corresponding locations by watermark bits embedding (WBE). The above embedding process is repeated for each secret bit in $\{m\}$. The resultant signal Z' after embedded is then inversely transformed by inverse cepstrum domain transform to obtain the watermarked audio signal Y. The details of each block of our embedding algorithm are described in the following.



AFS:Audio Frame Selection
CM:Chaotic Mechanism
RSB:Rearrange to Series Bits Stream
ICDT:Inverse Cepstrum Domain Transform
WBE:Watermark Bit Embedding
MEDM : Majority Energy Distribution Modulation of Group Signal

Fig. 1 The flow chart of audio watermark embedding

2.1 Chaotic Mechanism and Rearrange to Series Bits

In data hidden system, we want the data to be chaotic because a chaotic data is not easy to be stolen or attacked. Therefore, in our scheme needs a chaotic mechanism to hash the watermark stream. In this paper, a fast pseudorandom number traversing method is used as the chaotic mechanism to permute the watermark stream. Fig. 2 shows the result of watermark hashed by chaotic mechanism. Owing to the watermark is a two dimension image, it is needed to rearrange into a bits stream for secret data embedding. Fig. 3 shows the concept of the watermark image transfer to bits stream.

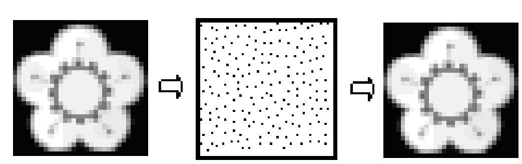


Fig. 2 The flow chart of watermark hashed by chaotic mechanism

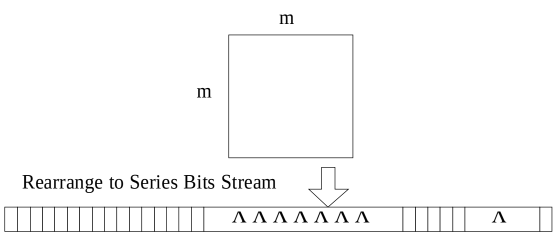
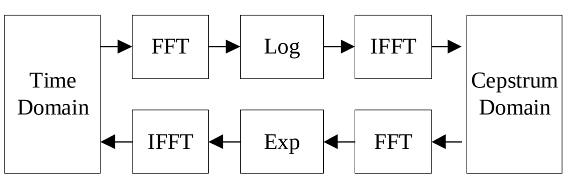


Fig. 3 The concept of the watermark image transfer to bits stream.

2.2 Cepstrum Domain Transform

An audio watermarking embedding process is used to embed secret data into audio signal. However, the audio signal must be divided into several non-overlapped frames for secret data embedding. But, what domain is superior to the watermark concealing? cepstrum coefficient has much less disturbance when it encounters the statistical-mean manipulation. Thus the cepstrum domain is widely used for voice identifier. Because the cepstrum domain with attack invariant feature, it is suitable for the audio watermarking. Fig. 4 shows the flow chart of a signal from time domain to cepstrum domain. In cepstrum domain transform, a signal pass to the fast Fourier transform firstly, then a logarithm operation and an inverse Fourier transform used to achieve the goal. Inversely, a fast Fourier transform, an exponential operation, and Inverse fast Fourier transform can easily back the cepstrum domain to time domain.



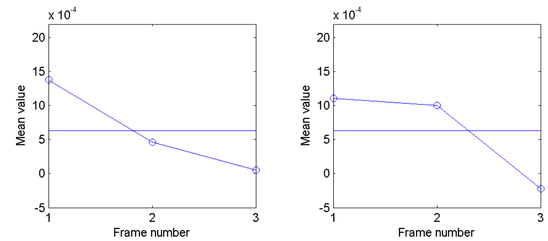
FFT: Fast Fourier Transform

IFFT: Inverse Fast Fourier Transform

Fig. 4 The relationship between time domain and cepstrum domain

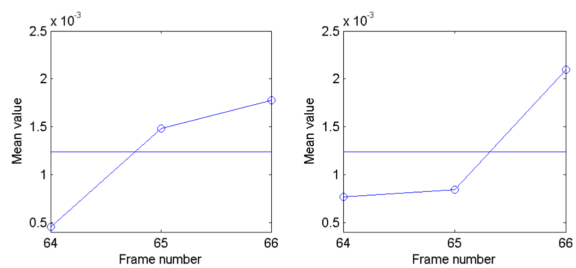
2.3TheMajority Energy Distribution Modulation

In our approach, we take 1024 coefficients of cepstrum domain serve as a frame and three frames used as a group. Meanwhile, a majority energy distribution modulation (MEDM) technique is used for secret bit embedding. In MEDM, we compute energy of each frame in the group, then adjust its mean to embed secret bit. Fig.5 and Fig. 6 show the details of the MEDM. Fig.5 shows the concept of embed secret bit '1'; Fig. 5(a) shows the original situation and Fig. 5(b) shows the result after secret bit '1' embedded, it is the mean of frame 2 adjusted. On the other hand, Fig.6 shows the concept of embed secret bit '0'; Fig. 6(a) shows the original situation and Fig. 6(b) shows the result after secret bit '0' embedded.



(a) (b)

Fig.5 the concept of embed secret bit '1'; (a) the original situation. (b) after secret bit '1' embedded.

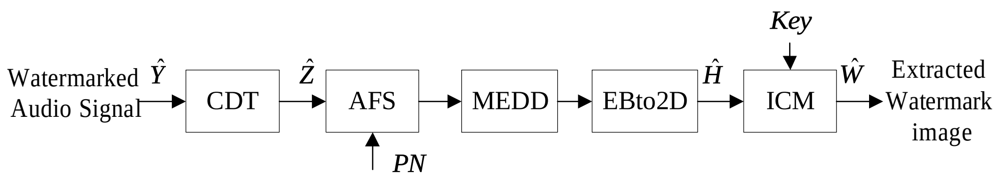


(a) (b)

Fig.6 the concept of embed secret bit '0'; (a) the original situation. (b) after secret bit '0' embedded.

3. Watermark Extracting Algorithm

Audio watermark extracting process is inverse the action of the audio watermark embedding. The block diagram of the Audio watermark extracting is shown in Fig. 6. At the beginning, the watermarked audio signal $\hat{Y}$ is transformed to $\hat{Z}$ by cepstrum domain transform. Then the same PN sequence used in the concealing process is hired to select the embedded frame from $\hat{Z}$ for the majority energy distribution demodulation (MEDD). After MEDD, the secret bit $\hat{m}$ is extracted. When all the secret bits are extracted by the MEDD, they are contracted by inverse rearrange to series bits and the two-dimensional image $\hat{H}$ is obtained by extracted bits to 2D (EBto2D). Finally, by passing $\hat{H}$ through the inverse chaotic mechanism (ICM), the recovered audio watermark $\hat{W}$ is gotten.



CDT:Cepstrum Doamin Transform
ICM:Inverse Chaotic Mechanism
EBto2D:Extracted Bits to 2 Dimension
AFS:Audio Frame Selection
MEDD : Majority Energy Distribution DeModulation of Group Signal

Fig. 7 The flow chart of audio watermark extracting

3.1 The Majority Energy Distribution De-Modulation

The majority energy distribution de-modulation (MEDD) process is inverse the action of the majority energy distribution modulation. During the MEDM does, the energy of each group is calculated, and then stored the values to the variables $\hat{E}_1 \Lambda \ \hat{E}_3$. Successively, we compute the average of group's energy. And then the secret bit is extracted easily by using the majority check rule.

3.2 Extracted Bits to 2D image

After watermark bits extracted, the watermark stream was obtained. However, the original watermark is a two-dimensional image. Thus, a rearrange step is needed to map the watermark into 2D style. Fig. 8 shows the schematic of the extracted watermark bits stream mapping to two-dimensional image.

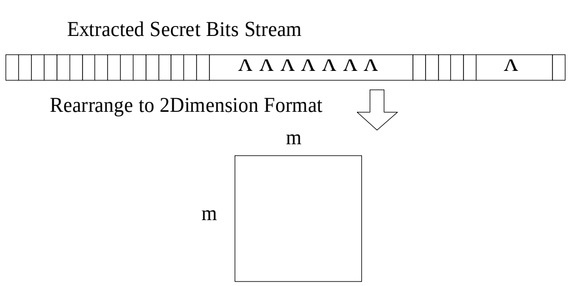


Fig. 8 The schematic of the extracted watermark bits transfer to two-dimensional image

3.3 Inverse Chaotic Mechanism

In the secret data embedding process, a fast pseudorandom number traversing method is used as the chaotic mechanism to permute the output streams. Likely, in the secret data extracting, the inverse fast pseudorandom number traversing method is used as the inverse chaotic mechanism to re-permute the output stream.

4. Experimental Results

The watermark similarity measurement is dependent on factors such as the knowledge of the experts, the experimental conditions, etc. Therefore a quantitative measurement is necessary to provide fair judgment of the extracted fidelity.

$$NC = \frac{\sum_i \sum_j w(i,j) w'(i,j)}{\sum_i \sum_j [w(i,j)^2]} \tag{1}$$

In order to illustrate the robustness of our watermarking scheme, various experiments were performed. The proposed algorithm applied one channel of digital audio pieces with 16 bits/sample, 44.1kHz sample rates, and 71.3 seconds long. Meanwhile, we use two 64×64 gray-scale logo image "ncut" in chinese as our watermark. In addition, the NC value is adopted to evaluate the similarity between the origin watermark and exacted watermark signal. PSNR is applied to evaluate the difference between the origin audio and watermarked audio signal. The audio pieces used in the experiment include piano, Classical, and Pop music as the table 1. Table 2 is the comparison of specification among our method and others. Beside, the NC value under attacks like, the MP3 compression, pitch shifting, re-sampling, low pass, and noise reduction are shown in table 3. Finally, The NC value, Error bits, and extracted watermark under different music and different attacks are shown in table 4.

Table 1 the test audio signal types and its information

Music / Infomation	Pop	Single Instrument (piano)	Single Instrument (cello)	Classic	Newage
Artist	L'arc~en~Ciel	Kevin Kern	Yo-Yo Ma	Mozart	Bandari
Title	Driver's High	Fantasia's Lullaby	Vistas	Symphonies No.41'Jupiter'	New Morning

Table 2 the comparison of specification among our method and others

References / parameter	W.N. Lie and L. C. Chang [11]	C. T. Hsieh and P. Y. Tsou [12]	W. Li et. al 13]	C. M. Huang [14]	P. Y. Tsou[15]	S. C. Liu [16]	Proposed technique
Sample rate	44.1 KHz	44.1 KHz	44.1 KHz	44.1 KHz	44.1 KHz	44.1 KHz	44.1 KHz
Resolution	16 Bit	16 Bit	16 Bit	16 Bit	16 Bit	8 Bit	16 Bit
Channel	-	-	-	Mono	Mono	Mono	Mono
Frame size	2048	1152	2048	1024	1024	256	1024
Watermark	Bitstream	-	24x24 Binary image	32x32 Binary image	Bitstream	64x64 Binary image	32x32 Binary Image
Music	-	5 min	45 sec	25 sec	10 sec	50 sec	71.3 sec

Table 3 the NC value under attacks among our method and others

Attack & Parameters / methods		[11]	[12]	[13]	[14]	[15]	[16]	Our method
Mp3 Compression	128 kbps		1				1	0.99
	96 kbps		1				1	0.99
	64 kbps		1	0.94	0.99		1	0.97
Pitch Shifting	-3 db				1	0.82	1	0.99
	+3 db						1	0.99
	+10%		0.8212				1	0.99
	+20%		0.8212				1	0.99
Resampling	22.05 KHz			0.8942	1		1	0.99
	96 KHz						1	0.99
Lowpass	4 KHz	0.9440		0.9041	0.9956		0.9785	0.99
	8 KHz						1	0.99
Noise Reduction	"Hiss Removal" of Goldwave			0.9027			1	0.97

Table 4 the NC value and error bits under attacks for different type's music

Music \ Attack & Parameters		MP3 128kb	MP3 96kb	MP3 64kb	+3db	-3db	Pitch +10%	Pitch +20%	Down Sample 22.05KHz	Up Sample 96KHz	Low Pass 4KHz	Low Pass 8KHz	Noise Reduction - Hiss Removal
Pop	NC	0.96	0.93	0.85	0.96	0.99	0.99	0.98	0.99	0.99	0.99	0.99	0.97
	Error Bits	52	108	189	51	4	4	13	19	4	8	4	37
	Result												
Single Instrument (piano)	NC	0.98	0.98	0.96	0.99	0.99	0.99	0.99	0.97	0.99	0.99	0.99	0.63
	Error Bits	24	22	52	15	17	13	13	32	12	14	12	442
	Result												
Single Instrument (cello)	NC	0.99	0.99	0.97	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.96
	Error Bits	8	12	48	3	4	4	3	10	3	5	4	52
	Result												
Classic	NC	0.99	0.98	0.95	0.99	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.96
	Error Bits	12	23	63	7	7	5	5	15	7	6	6	44
	Result												
Newage	NC	0.97	0.96	0.89	0.98	0.98	0.98	0.98	0.96	0.98	0.98	0.98	0.82
	Error Bits	42	53	143	22	22	21	20	47	22	26	20	227
	Result												

5. Conclusions

Cepstrum coefficient encounters the statistical-mean manipulation has much less disturbance. Thus the cepstrum domain is widely used for voice identifier. Owing to the cepstrum domain with attack invariant feature, it is suitable for the audio watermarking. In this paper, a novel audio-watermarking scheme in cepstrum domain using majority energy distribution techniques is proposed. Cepstrum domain transform is firstly hired to transfer the audio signal from time domain to cepstrum domain. In our approach, three cepstrum domain coefficients of a frame are treated as a group of a signal, and majority energy distribution of the groups is used for the secret bit embedding. On the other side, a logo image "NCUT" in Chinese is serve as a watermark. Meanwhile, a chaotic mechanism is used to hash the watermark image for enhancing the security.

In order to demonstrate the effectiveness of the proposed scheme, simulations under various conditions were conducted. Several kinds of music like piano music, classical music, and pop music are used as the test audio signals. The experimental results show our scheme can sustain all kind of attacks like added noise, MP3 compression, and re-sampling.

References

[1] Bassia P., Pitas I., and Nikolaidis N. " Robust

audio watermarkingin the time domain “, **IEEE Transactions on Multimedia**, Vol.3, No.2, pp.232-241, 2001.

[2] Swanson D., Zhu B.,and Tewfik H.”Robust audio watermarking using perceptual masking”, **Signal Processing**, Vol.66, No.6, pp.337-355, 1998.

[3] Xiangyang Wang and Hongying Yang, “ Digital audio watermarking algorithm based on adaptive quantization”, **Technical Acoustics**, Vol.23, No.2, pp.117-121, 2004.

[4] Xiangyang Wang and Hongying Yang, “Adaptive digital audio watermarking based on discrete cosine transform “, **Mini- Micro Systems**, Vol.25, No.10, pp.1825-1828, 2004.

[5] H. J. Kim, “ Audio Watermarking Techniques”, **Pacific Rim Workshop on Digital Steganography**, Kyushu Institute of Technology, Kitakyushu, Japan, July 3-4, 2003.

[6] Huang J. W., Wang Y., and Shi Y. Q. ,”A blind audio watermarking algorithm with self-synchronization”, In: **Proceedings of IEEE International Symposium on Circuits and System**, Arizona, USA, Vol.3, pp.627-630, 2002.

[7] Xiangyangt Wang and Yongrui Zhao Cui, “ Hong 2-Dimention Digital Audio Watermarking Based on Integer Wavelet Transform”, **Proceedings of ISCIT2005**, pp.877-880,2005.

[8] Sang-Kwang Lee and Yo-Sung Ho, “Digital Audio Watermarking in the Cepstrum Domain,” **IEEE Transactions on Consumer Electronic**s, Vol.46 No.3, pp.744-750, 2000.

[9] C.T. Hsieh and PY. Tsou, “Blind cepstrum Domain Audio Waternnaking based on Time Energy Features”, 14th of, **International conference on Digital signal Processing es AProceeding**, Greece: Santorini, pp.705- 708, 2002.

[10] Tang Xianghong, Niu Yamei, and Li Qiliang, “ A Digital Audio Watermark Embedding Algorithm with WT and CCT”, **2005 IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications Proceedings**, pp.970-973, 2005.

[11] Wen-Nung Lie, and Li-Chun Chang, “Robust and high-quality time-domain audio watermarking subject to psychoacoustic masking”, **Circuits and Systems, The 2001 IEEE International Symposium** on, Volume: 2, 6-9 May, 2001, pp.45-48.

[12] Ching-Tang Hsieh, and Pei-Ying Tsou, “Blind cepstrum domain audio watermarking based on time energy features”, **Digital Signal Processing, 2002 14th International Conference** on, Volume: 2, 1-3 July, 2002, pp. 705-708.

[13] Wei Li, Xiangyang Xue, Xiaoqiang Li, Peizhong Lu, “A novel feature-based robust audio watermarking for copyright protection”, Information Technology: **Coding and Computing [Computers and Communications], 2003. Proceedings. ITCC 2003. International Conference** on, 28-30 April 2003, pp.554-558.

[14] Ching-Ming Huang, “Key-based Audio Watermarking System using Wavelet Packet Decomposition”, **國立中央大學電機工程研究所碩士論文, June 2002.**

[15] Pei-Ying Tsou, “A study of the synchronization problem of audio watermark”, **淡江大學電機工程學系通訊系統組碩士論文,** February 2003.

[16] Shi-Cheng Liu, “A BCH Code-Based Robust Audio Watermarking in Cepstrum Domain”, **國立東華大學資訊工程學系碩士論文**, July 2004.