An Audio Copyright Protection Schemes Based on SMM in Cepstrum Domain*

Shenghong Li[1], Lili Cui[1], Jonguk Choi [2], and Xuenan Cui[2]

[1] School of Information Security Engineering,
Shanghai Jiaotong University
Shanghai, 200030, China
shli@sjtu.edu.cn,
cuilili@sjtu.edu.cn
[2] Copyright Protection Research Institute,
Sangmyung University, 5096, Korea
juchoi@sangmyung.ac.kr,
cuixuenan00@163.com

Abstract. In this paper, we present an audio scheme protective of copyright protection using information hiding. We propose visually recognizable binary image and text information as watermark (copyright) information embedded in audio signal. Cepstrum representation of audio can be shown to be very robust to a wide range of attacks. We apply SMM(statistical mean manipulation) theory in embedding image watermarking, and address attacks against lossy audio compression like MP3, white Gaussian noise and so on. A blind detection watermarking can be realized with the proposed scheme.

Keywords: copyright protection, information hiding, watermark, statistical mean manipulation, cepstrum domain.

1 Introduction

The digital watermark technique is a technique to solve the copyright problem. The media owner can use this technique to insert some information into the media. There has been a fair amount of research on diverse applied techniques of audio watermark, i.e. Spread Spectrum method [1-4], echo hiding [5-7], a method Replica Signal [9] etc.

However in most audio watermarking methods, the embedding algorithms embed a chaos sequence or pseudo-random array to be watermarking in the content, insert mean information is very peculiar. In this paper we will insert a still binary image being audio watermarking into the ceptrum domain. Extensive experimental results prove that the embedded watermark is inaudible and robust.[1]

* The work is fully supported by the international co-operation project of the ministry Science and Technology of Korea: Co-Development of Broadcasting Sync. Equipment and DRM Watermark Chipset for Digital Broadcasting Content based on Original Watermarking Technology of Korea. (Project No: M60401000150-05A0100-15010).

D.-Y. Yeung et al. (Eds.): SSPR&SPR 2006, LNCS 4109, pp. 923 – 927, 2006.

2 Details of the Proposed Algorithm

The cepstrum domain analysis is used commonly in speech application, such as recognition area. In speech recognition, the cepstral coefficients are regarded as the main features of a voice. The cepstral coefficients vary less after general signal processing then samples in time domain .Due to the advantage of cepstral coefficients, Li and Yu [10] proposed a robust audio data hiding technique in cepstrum domain.

Cepstral analysis utilizes a form of homomorphic system which converts the convolution operation an addition operation. It consists of three consecutive steps: Fourier transform, take logarithm and inverse Fourier transform. It is easy to see that those three operations are all linear It should be noted that the logarithm we take at the second step is complex logarithm and $X(n)$ is formally called "*complex cepstrum*". But in practice, people often define the real part of complex cepstrum to be the "real" cepstrum for convenience.

$$X(n) = IFFT\left(\log\left(REAL\left(FFT\left(x(n)\right)\right)\right)\right) \tag{1}$$

And we can exactly recover the original signal in time domain from its cepstrum domain representation by taking correspondent inverse operations

$$x(n) = IFFT\left(\exp\left(REAL\left(FFT\left(X(n)\right)\right)\right)\right) \tag{2}$$

Cepstrum coefficients are around zero except the last, therefore we shall modify small cepstrum coefficients except the last. Experimental studies have shown that most common signal processing could change individual cepstrum coefficients dramatically, but their statistical mean often experiences much less disturbance, offering an appropriate candidate for information carrying.

3 Scheme on Binary Image Watermark Embedding

In embedding process, we adopt the concept of the cepstrum, and embed the data based on statistical mean theory which is much more robust, especially for attacks destroying synchronization structure of audio signal. We shall focus on the statistical mean of cepstrum coefficients to be a real number for embedding '1',and another number for embedding '0', then we can detect the watermarking by adjudging the threshold derived from the two number. The detail watermarking embedding works as following:

1. Transform time domain signal to cepstrum domain.
2. Divide audio cepstrum into frames, which is depend on the size of binary image.
3. Calculate the mean of each frame of cepstral coefficients. Modify the mean of cepstral coefficients to zero. Then the embedding algorithm is following:

To embedding '1':

$$X(n)' = X(n) + \alpha * W_m(n) \tag{3}$$

To embedding '0':

$$X(n)' = X(n) \tag{4}$$

Where α is the factor controlling the allowable distortion for individual cepstrum component $X(n)$. $W_m(n)$ is watermarking information, m denotes the number of frame.

4. Create the final watermarked audio.

4 Binary Image Watermark Detecting

The watermark should be extractable even if common signal processing (including data compression and some kinds of noise attacks) operations are applied to the host audio. In detection procession, don't need original audio signal, is total blind detection process. The detection method is base on statistical mean manipulation, calculate the sum of every frame cepstrum coefficients, and set the threshold Td to identify the watermark information.

5 Experiment Results

In the experiment, Matlab6.1 is used as emulation software, the music used as the watermarked media is 102.06 seconds music, 11025Hz of sampling rate and 16 bit recorded for each sampling. The embedding capacity is 62kbps.The watermark is 64x64 binary image , given in Fig1 (a). A blind listening test was used to confirm the transparency of the watermarked signal and most listeners couldn't distinguish the difference of the watermarked signals.

The following are the test results, where Fig1 (a) is original watermark (binary image), and Fig1 (b) is picked up without attacked by our detection method.

(a) (Sim=1) (b) (Sim=1) (c) (Sim=0.805) (d) (Sim=0.624)

Fig. 1. (a) Original watermark, (b) Detected watermark, (c) MP3 compression at 64kbps, (d) MP3 compression at 32kbps

To test the robust of our scheme, we evaluate the performance of the watermark against lossy attacks by diving the test results into four subtest, the performance can refer Fig1 and Fig4:

Subtest1 (MP3 Attack): We compare the effect marked the audio and the decoded audio given by MP3 compression at different bit rate. Fig1 (c) is under attack of MP3 compression at the rate of 64kbps, and provides transparent audio quality. Fig1 (d) is at 32kbps. Each similarity value corresponding to the compression rate is shown in Fig2. In Fig2 with the rate of MP3 increase, accordingly the similarity value increase, here we list four kinds of conditions, the lowest rate is 32kbps, and under 32kbps, the image can't be extracted. So we can see that embedded image can be extracted for MP3 compression at the rate of above 32kbps.

Subtest2 (White Gaussian Noise Attack): Fig3. (a), (b) are under attack of white Gaussian noise with mean zero, covariance 1 and 0.1. We can see our proposed scheme

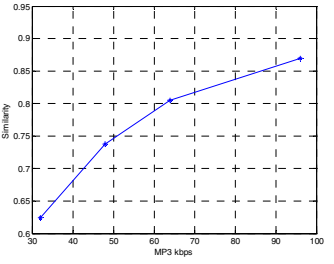

Fig. 2. Similarity comparison with various Mp3 compression rate

(a) (Sim=0.89) (b) (Sim=0.81) (c) (Sim=0.88)

Fig. 3. (a)White Gaussian noise (0,1); (b) White Gaussian noise (0,0.1); (c) Median filter

demonstrates good survivability with (0,1) white Gaussian noise. But if mean is nonzero, or covariance is above 0.1, we can't detect the watermark.

Subtest3 (Filter Attack): Fig3 (c) shows the detection performance after median filter, which is nonlinear filter. It can be seen from Fig3 that our scheme demonstrates good robustness.

(a) (Sim=0.95) (b)(Sim=0.86) (c)(Sim=0.98) (d)(Sim=0.58)

Fig. 4. (a) 22050Hz (b) 44100Hz (c) 8000Hz (d)8 bit

Subtest4 (Repeat Sampling and Repeat Quantification): for repeat sampling test we subsampling watermarked signal at 22050Hz and 44100Hz and 8000Hz ,then revert original sampling frequency, Fig4 (a-c) show the performance under these condition.for repeat quantification test we quantify watermarked signal from 16 bit at first to 8 bit, then restore signal, the performance just as (d) show.

$$W_m' = \begin{cases} 1 & \sum_{i=1}^{N} x_m(i) > T_d \\ 0 & \sum_{i=1}^{N} x_m(i) < T_d \end{cases} \tag{5}$$

where the $x_m(i)$ denote the mth frame of the audio signal, W_m' is detected watermarking information, the value is '1' or '0'. We embed 1 bit each frame. To show the performance of our test, we compare the extracted watermark with original watermark. In this comparison, we use the similarity measure given in (6).

$$Sim(W,W') = W \cdot W' / \sqrt{W \cdot W} \tag{6}$$

6 Conclusion

In this paper, an audio scheme protective of copyright protection using information hiding is proposed. The binary image watermark scheme based on SMM(statistical mean manipulation) theory, and divide frames to embedding watermark. The audio scheme is robustness against the data compression and some kinds of attacks such as MP3, Audio Stirmark, white Gaussian noise and repeat sampling and repeat quantification.

The following is our future work:

(1) Study on the performance of SMM further.
(2) Research for the robust performance of the other audio attack.
(3) Research for the robust performance of the other embedding domain.
(4) Multi-watermark embedding.
(5) Study on the performance of Text as watermarks.

References

1. P.Bassia,I.Pitas, and N. Nikolaidis: Robust audio watermarking in the time domain. IEEE Transactions on Multimedia, vol. 3, June (2001), pp. 232-241.
2. D.Kirovski and H.Malvar: Robust spread–spectrum audio watermarking. IEEE International Conference on Acoustics, Speech, and Signal processing, vol. 3, (2001). pp. 1345-1348
3. L.Boney, A.H.Tewfik, and K.N. Hamdy: Digital watermark for audio signals. In International Conference on Multimedia Computing and Systems, IEEE, Hiroshima, Japan, June, (1996), pp.473-480.
4. H.Malik, S.Khokhar, and A.Rashid: Robust audio watermarking using frequency selective spread spectrum theory, In International Conference on Accoustic, Speech and Signal Processing, IEEE, Montreal, Canada, May, (2004). pp. 385-388.
5. D.Gruhl, A.Lu, W.Bender: Echo hiding. in Proc. Information Hiding Workshop , University of Cambridge, U.K., (1996), pp. 295-315.
6. S. W. Foo, T. H. Yeo, and D. Y. Huang: An Adaptive Audio Watermarking System. Electrical and Electronic Technology, TENCON. Proceedings of IEEE Region 10 International Conference on, Vol2, (2001), pp.509-513.
7. H. O. Oh, J. W. Seok, J.W. Huang and D. H. Youn: New echo embedding technique for robust and imperceptible audio watermarking. Acoustics, Speech, and Signal Processing, Proceedings. 2001 IEEE International Conference on, Vol3, (2001), pp.1341-1344.
8. S. Shin, J. W. Kim, J. Choi: Audio watermarking using Digital Filter. Korea Information Security, Conference, vol. 11. No.1 (2001). pp.464-468.
9. R. Petrovic: Audio signal watermarking based on replica modulation. Telecomfmunications in Modern Satellite, Cable and Broadcasting Service, TELSIKS 2001. 5th International Conference on vol 1.(2001). pp.227-234.
10. X. Li, H. H. Yu: Transparent and Robust Audio Data hiding in cepstrum Domain. ICME2000, vol. 1, (2000). pp.397-400.