JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 22, 535-543 (2006)

BCH Code-Based Robust Audio Watermarking in the Cepstrum Domain*

SHI-CHENG LIU AND SHINFENG D. LIN+
Department of Computer Science and Information Engineering
National Dong Hwa University
Hualien, 974 Taiwan
E-mail: ko.hsien@msa.hinet.net
+E-mail: david@mail.ndhu.edu.tw

In this article, a BCH code-based audio watermarking approach performed in the cepstrum domain is proposed. The technique takes advantage of the attack-invariant feature of the cepstrum domain and the error-correction capability of BCH code to increase the robustness of audio watermarking. In addition, the watermarked audio has very high perceptual quality. A blind watermark detection technique is developed to identify the embedded watermark under various types of attacks. Experiment results demonstrate that the proposed technique outperforms the existing audio watermarking techniques against most of the asynchronous attacks.

Keywords: audio watermarking, cepstrum domain, permutation, error correction coding, BCH code

1. INTRODUCTION

The rapid evolution of the digital world has greatly facilitated the manipulation and transmission of digital media, such as text, images, audio, and video. Easy access and replication, however, have led to serious problems with copyright protection for media. Watermarking is a new technology that draws from many different fields, such as cryptography, communication theory, information theory, etc. Therefore, media owners can use this technique to insert some information into their media for the purpose of copyright recognition. Three capabilities are required in digital watermarking: transparency, robustness, and capacity. It must be noted that these capabilities are often mutually contradictory, so compromises must be made among them. Most of the existing watermarking algorithms are applicable to images or video. However, the literature on audio watermarking techniques is comparatively limited.

The widespread use of the Internet and the proliferation of digital audio distribution in MP3 form have made the copyright protection of digital audio works more and more necessary. Some research works have been published on audio watermarking. These approaches work in the time domain [1], temporal domain [2], DCT domain [3], DWT domain [4], cepstrum domain [5, 6], or subband domain [7, 8]. The methods for encoding audio signal adopt delay [9, 10], a pseudo-random sequence [11, 12], or linear

Received January 1, 2005; accepted June 30, 2005.
Communicated by C. C. Chiang, H. Y. Mark Liao and K. C. Fan.
* This work was supported in part by the National Science Council of Taiwan, R.O.C., under grant No. NSC 92-2213-E-259-018.

block code [13, 14]. In previous articles, much attention has been paid to audio watermarking techniques performed in the cepstrum domain [5, 6] and to the use of linear block code [13, 14].

The cepstrum domain has been widely adopted for phonetic analysis and recognition, which include a series of operations: (1) Fourier transform, (2) take logarithm, and (3) inverse Fourier transform. It is obvious that these three operations are linear and that the original signal in the time domain can be exactly recovered from its cepstrum domain representation. After a general attack, the statistical mean of the cepstrum coefficients for an audio signal experience much less variance. Due to the attack-invariant feature, the watermark information can be preserved. BCH code belongs to linear block code. In order to increase security for watermark hiding, BCH encoding is incorporated. However, BCH code has only a finite error-correction capability.

This article is organized as follows. Section 2 introduces cepstrum analysis and BCH code. The proposed watermarking scheme is presented in section 3. Experimental results are presented and discussed in section 4. Finally, conclusions are drawn in section 5.

2. CEPSTRUM DOMAIN AND BCH CODE

The steps in cepstrum analysis are shown in Fig. 1. It should be noted that the logarithm in the second step is a complex logarithm. In practice, researchers often refer to the real part of the complex cepstrum as the real cepstrum for short. In our research, however, the complex cepstrum is still required for converting back to the initial signal sequence. To more clearly explain cepstrum analysis, we will present an original signal and an attacked signal in the following figures. Among these figures, Fig. 4 demonstrates the attack-invariant feature in the cepstrum domain.

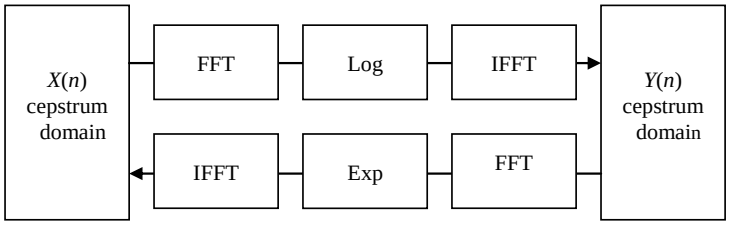

Fig. 1. Cepstrum analysis.

BCH codes are block-based error correcting codes with a wide range of applications in digital communications and storage. Defined by the polynomial of the code generator, the length of BCH code is exactly n, and $n = 2^m - 1$, $m \geq 3$. The numeric available for correction is t, and $t < (2^m - 1)/2$, $n - k \leq mt$ [15]. Thus, not all of the values of n and k can satisfy the previously mentioned inequality equation. Table 1 shows the available parameters for BCH code.

By using the attack-invariant feature of cepstrum analysis and the error-correction capability of BCH code, we can increase the robustness of audio watermarking.

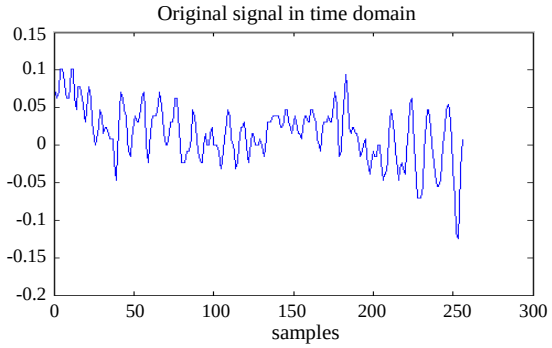

Fig. 2. Original signal in the time domain.

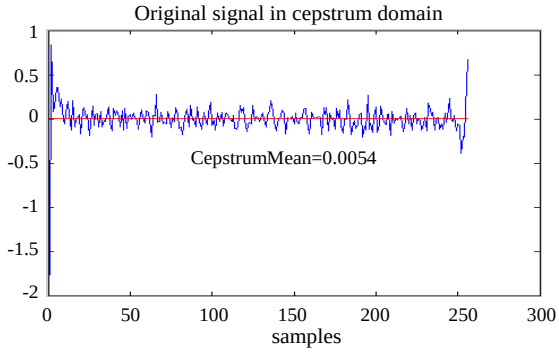

Fig. 3. Original signal in the cepstrum domain.

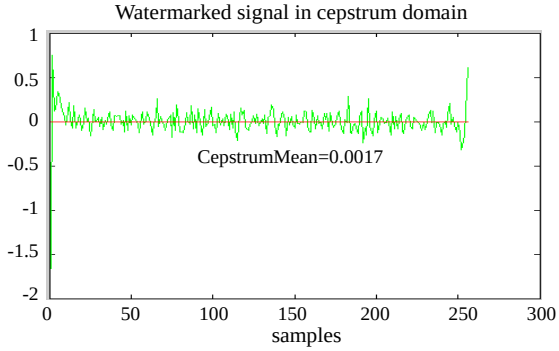

Fig. 4. An echo-attacked watermark signal in the cepstrum domain.

3. THE PROPOSED AUDIO WATERMARKING SCHEME

3.1 Embedding Algorithm

A general outline of the basic steps involved in the proposed scheme is shown in Fig. 5.

Table 1. Available parameters for BCH code.

N	*k*	*t*	*n*	*k*	*t*
7	4	1	**127**	99	4
15	11	1		92	5
	7	2		85	6
	5	3		78	7
31	26	1		71	9
	21	2		**64**	**10**
	16	3		57	11

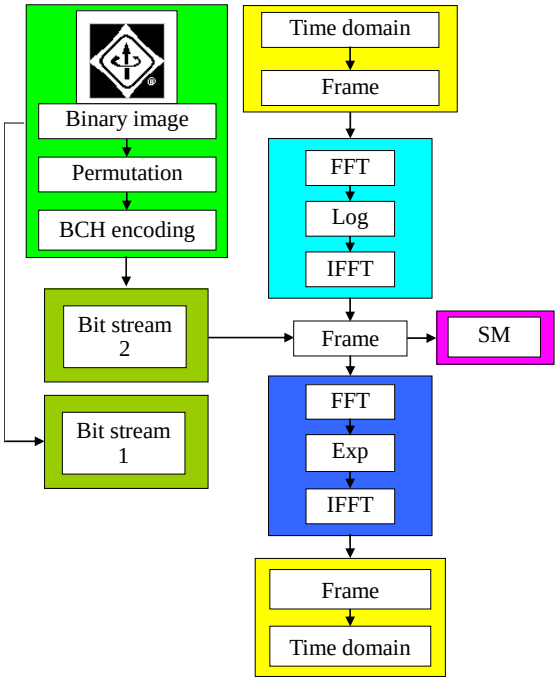

Fig. 5. Flowchart of watermark embedding.

Step 1: The input audio signal, sampled at 44100 Hz, is first segmented into non-overlapping frames, with a frame size of 1024 samples. Each music song has a duration of 300 seconds and is recorded in mono at a sampling rate of 16 bits.

Step 2: The frame is transformed into another type of logarithm. It is easy to see that these three operations are all linear.

Step 3: The mean of a frame in the cepstrum domain is the SM (statistical mean). We remove the bias of SM and then accumulate these coefficients as a threshold. The sequence of frames is denoted as *A*. The sequence of thresholds is denoted as *B*.

Step 4: The matrix of the binary image can be permuted by the specified vector order. This order can be considered as a secure key(*C*).

Step 5: The logo image can be encoded by using the BCH technique and the generator polynomial:

Code = encode(n, k, *msg*)

where n is the codeword length, k is the logo length, and *msg* is a binary matrix with k columns. Each row of *msg* represents a codeword. The code is in the form of a binary matrix with n columns. The sequence of code is denoted as D.

Step 6: It is easy to perceive the mappings between A and D. Every single bit of D is embedded into every single frame of A:

To embed "1":
$x'(n) = x(n)$ + parameter, if $x(n) < 0$;
To embed "0":
$x'(n) = x(n)$ – parameter, if $x(n) > 0$.

Step 7: The frame is transformed into an original type of exponential. This method is similar to the inverse of step 2. In this way, we can then transform the embedded signal into the time domain.

3.2 Extraction Algorithm

A general outline of the basic steps involved in the proposed scheme is shown in Fig. 6.

Step 1: The input watermarked audio signal sampled at 44100 Hz is first segmented into non-overlapping frames, with a frame size of 1024 samples.

Step 2: The frame is transformed into the cepstrum domain.

Step 3: We calculate the mean value of a frame's coefficients as the CM (coefficient mean). The sequence of the threshold is denoted by F.

Step 4: It is easy to perceive the contrast between B and F. Suppose the sequence of F is F_i (i = 1, …, n) and that the sequence of B is B_i (i = 1, …, n). If F_i is greater than B_i, a bit "1" is extracted in the corresponding frame. Else if F_i is less than B_i, a bit "0" is extracted in the corresponding frame. We collect these "1s" or "0s" into a sequence (G).

Step 5: The logo image can be decoded by using the BCH technique:

H = bchdeco(G, k, t)

where t is the error-correction capability and H is the logo image.

Step 6: The logo image can be inpermuted by using the secure key(C):

I = inpermuted(H, C)

where I is the extracted watermark. Finally, the normalized correlation is calculated.

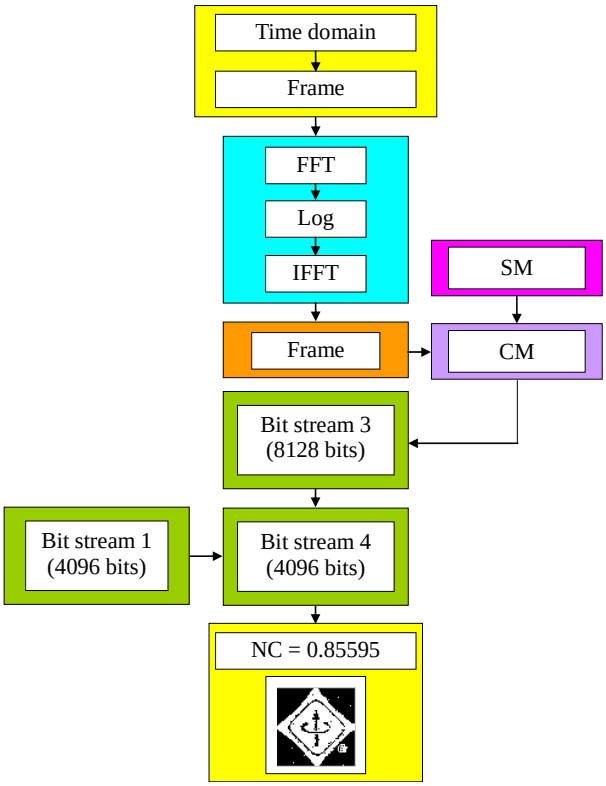

Fig. 6. Flowchart of watermark extraction.

4. EXPERIMENTAL RESULTS

The sampling rate, f_s, was used for playback. The value typically supported by sound cards is 44100 Hz. Each frame had 1024 samples. Each song had a duration of 300 seconds and was recorded in mono at a sampling rate of 16 bits.

The audio editing and attacking tools adopted in this experiment were CoolEdit Pro 2.0 and GoldWave 4.26. In Tables 2 and 3, comparisons of experimental data obtained in previous studies for asynchronous and synchronous attacks, respectively, are shown. It can be seen that our technique outperformed the existing audio watermarking techniques against most of the asynchronous attacks. However, the experiment results obtained against the synchronous attacks leave room for improvement.

5. CONCLUSION

Judging from the experimental data, cepstrum domain analysis is applicable to audio watermarking. Due to its attack-invariant feature, even after numerous asynchronous attacks, it still remains very robustness. From our research, however, it seems no excellent algorithms from the scientific research documents for resisting synchronous attacks. This will be our research topic in the near future.

Table 2. Comparison with previously reported experimental data for asynchronous attacks.

attack	References / parameter	[16]	[17]	[14]	[8]	[18]	Proposed technique
amplitude	-3db						1
	+3db						1
	+5%		0.9942				1
	+10%		0.8212				1
	+15%		0.8370				1
	+20%		0.8212				1
resample	44100→22050→44100			0.8942	1		1
	44100→96000→44100						1
lowpass	Frequency : 4 KHZ	0.9440		0.9041	0.9956		0.9785
	Frequency : 8 KHZ						1
Mp3	32Kbps		1		0.9	1	1
	48Kbps				0.99		1
	64Kbps		1	0.9411	0.99		1
	80Kbps	0.9410					1
	96Kbps		1		1		1
	128 Kbps	0.9916	1		1		1
	160 Kbps	1					1
	256 Kbps				1		1
	100ms		0.9826	0.9103			1
	200ms		0.9942				1
	300ms		1				1
	400ms		1				1
noise	'Hiss removal' of GoldWave			0.9027			1
time stretch	95%						1
	97%			0.8274			1
	103%			0.8829			1
	105%						1

Table 3. Comparison with previously reported experimental data for synchronous attacks.

attack	references / parameter	[16]	[17]	[14]	[8]	[18]	Proposed technique
Cut-samples	…	1			0.0677	0.8000	
	100/1		0.8170	0.8646			1
	200/1						1
	500/1			0.8832			1

REFERENCES

1. P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Transactions on Multimedia*, Vol. 3, 2001, pp. 232-241.
2. A. N. Lemma, J. Aprea, W. Oomen, and L. van de Kerkhof, "A temporal domain

audio watermarking technique," *IEEE Transactions on Signal Processing*, Vol. 51, 2003, pp. 1088-1097.
3. I. K. Yeo and H. J. Kim, "Modified patchwork algorithm: a novel audio watermarking scheme," *IEEE Transactions on Speech and Audio Processing*, Vol. 11, 2003, pp. 381-386.
4. L. Cui, S. X. Wang, and T. F. Sun, "The application of wavelet analysis and audio compression technology in digital audio watermarking," in *Proceedings of the IEEE International Conference on Neural Networks and Signal Processing*, Vol. 2, 2003, pp. 1533-1537.
5. S. K. Lee and Y. S. Ho, "Digital audio watermarking in the cepstrum domain," *IEEE Transactions on Consumer Electronics*, Vol. 46, 2000, pp. 744-750.
6. X. Li and H. H. Yu, "Transparent and robust audio data hiding in cepstrum domain," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, Vol. 1, 2000, pp. 397-400.
7. X. Li and H. H. Yu, "Transparent and robust audio data hiding in subband domain," in *Proceedings of the IEEE International Conference on Information Technology: Coding and Computing*, 2000, pp. 74-79.
8. J. M. Huang, "Key-based audio watermarking system using wavelet packet decomposition," Master's thesis, Dept. of Electrical Engineering, National Central University, Taiwan, 2002.
9. P. Eneroth, "Joint filterbanks for echo cancellation and audio coding," *IEEE Transactions on Speech and Audio Processing*, Vol. 11, 2003, pp. 342-354.
10. H. R. Wu, "Design of digital audio watermarking techniques based on echo data hiding," Master's thesis, Dept. of Computer Science, National Chiao Tung University, Taiwan, 2000.
11. M. Ikeda, K. Takeda, and F. Itakura, "Audio data hiding by use of band-limited random sequences," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 4, 1999, pp. 2315-2318.
12. B. S. Ko, R. Nishimura, and Y. Suzuki, "Time-spread echo method for digital audio watermarking using PN sequences," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 2, 2002, pp. 2001-2004.
13. G. D. Forney, Jr., "On decoding BCH codes," *IEEE Transactions on Information Theory*, Vol. 11, 1965, pp. 549-557.
14. W. Li, X. Xue, X. Li, and P. Lu, "A novel feature-based robust audio watermarking for copyright protection," in *Proceedings of the IEEE International Conference on Information Technology: Coding and Computing Computers and Communications*, 2003, pp. 554-558.
15. 翁芳標, "通訊工程概論," 全華科技圖書公司, 1998, pp. 5-22.
16. W. N. Lie and L. C. Chang, "Robust and high-quality time-domain audio watermarking subject to psychoacoustic masking," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 2, 2001, pp. 45-48.
17. C. T. Hsieh and P. Y. Sou, "Blind cepstrum domain audio watermarking based on time energy features," in *Proceedings of the IEEE International Conference on Digital Signal Processing*, Vol. 2, 2002, pp. 705-708.
18. P. Y. Zou, "A study of the synchronization problem of audio watermark," Master's thesis, Dept. of Electrical Engineering, Private Tamkang University, Taiwan, 2003.

Shi-Cheng Liu (劉適程) received his B.S. degree from the Department of Elementary Education of National Hualien Teacher's College, Hualien, Taiwan, in 1992, and his M.S. degree from the Department of Computer Science and Information Engineering of National Dong Hwa University, Taiwan, R.O.C., in 2004. He is currently the Dean of Studies at Chung Yuan Primary School, Hualien, Taiwan, R.O.C. His research interests include image processing and audio watermarking.

Shinfeng D. Lin (林信鋒) received his B.S. degree in Automatic Control Engineering from Feng Chia University, Taichung, Taiwan, in 1980, and his M.E. and Ph.D. degrees in Electrical Engineering from Mississippi State University in 1985 and 1991, respectively. He is currently a Professor in the Department of Computer Science and Information Engineering at National Dong Hwa University, Taiwan, R.O.C. Dr. Lin was the Director of the Bureau of Education, Hualien County, Taiwan, from Jan. 2002 to Sep. 2003. His research interests include signal/image processing, image/video compression, and information security. Dr. Lin is a member of Tau Beta Pi and Eta Kappa Nu.