MCIS 2006
Conference in technical cooperation with the IEEE and IEEE Computer Society

Increasing Robustness of Audio Watermarking DM using ATHC Codes.

R. Martínez-Noriega[2,3], H. Kang[1], B. Kurkoski[2], K. Yamaguchi[2], K. Kobayashi[2], M. Nakano[3]
[1]Graduate School of Information Systems, [2]Dept. of Inf. and Communications Eng.
University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-Shi, Tokyo 182-8585, Japan
{ kang,kurkoski,yama,kingo}@ice.uec.ac.jp
[3]Secction of graduate students and research.
SEPI-ESIME-Culhuacan, Av.Santa Ana No. 1000, 04430, Mexico City, Mexico
raul_mtz@prodigy.net.mx

Abstract— **This paper presents a robust technique for embedding digital watermarks into digital audio signals. Watermarking is a technique used to label digital media by hiding copyright or other information into the underlying data. In the case of digital audio the main constraint is the trade-offs between imperceptibility and robustness, due to hearing is one of the most sensitive senses. In our method, the original watermark is modified by a system of error control coding (ECC) for create a robust watermark. We use an asymmetric turbo-Hadamard code (ATHC) in the ECC process, it has good performance compared to original turbo-Hadamard code (THC) in waterfall region. The embedding process is by mean of dither modulation (DM) in the time domain, this one is an implementation of quantization index modulation (QIM). As a result we obtain a robust technique against MPEG-1 layer III (mp3) and better than others algorithms in time domain as LBM, DC-Shifting.**

I. INTRODUCTION

With the explosive growth of the computers and Internet, multimedia data has become more popular, due to digital advantages like the ease of transmission in Internet, copying without loss of quality and so on [1]. This is a problem because exists many copies that infringe the copyright. Many techniques have been researched, one of them is cryptographic systems but it has the disadvantage that once you break the security you can have free access at digital data. Watermarking arose as a complement to security systems, this technique attempts to embed additional information into digital data without introducing perceptual changes and at the same time the watermark must be robust to common degradations.

Nowadays there exist many methods like spread-spectrum, which embed additional information by lineal combination of the host signal with a small pseudo-noise signal that is modulated by the embedded signal, but this kind of method has some problems when the original host signal is unknown in the decoding process. QIM is a method which has very good rate-distortion-robustness trade-offs and it is based in the idea that each quantizer can represent one different symbol of the watermark [2].

Many times the robustness of the watermarking systems is not enough, for this reason we need improvement by applying ECC to the watermark. Actually there exist many kinds of ECC, one of these are the turbo codes whose astonishing performance as given rise to a large interest in the coding community were first proposed by Berrou in 1993 [3]. The encoder of turbo codes is formed by two or more constituent systematic encoders joined through the interleaver. The binary information data sequence is fed into the first encoder, and it is also fed into the second encoder after permutation by the interleaver. The coded sequences of the parallel concatenated code consist of the information sequence to first encoder and the parity check bits of both encoders [4].

We propose a robust watermarking method against mp3 based in dither modulation with the addition of ECC in the input, this ECC is called asymmetric turbo-Hadamard code (ATHC) [5] [6]. An important feature of ATHC and THC is that they are low rate codes, in this proposed method there are a large number of places to embed data and small amount user information. The objective of use an ATHC is improvement of watermark robustness before apply watermarking process.

The rest of this paper is organized as follows. In section II, introduce the background of QIM method with DM application. In section III, we show and describe the characteristics of ATHC using in this method. In section IV, the proposed algorithm is presented: embedding phase and watermark detection. Experimental results are given in the section V, and conclusions are addressed in section VI.

II. QIM

QIM is the method of embedding information by first modulating an index or sequence of indices with the information and then quantizing the host signal with the associated quantizer or sequence of quantizers [7]. We can see the embedding function (1), as an ensemble of samples p quantized by the quantizer indexed for m_i symbol.

$$p' = q(p\ , m_i) \qquad i = 1,2,\ldots,n \tag{1}$$

where:

n = number of different symbols in the watermark.

p = audio sample.

m_i = i-symbol of the watermark.

In other words, if we want to embed a symbol m_i of the watermark into the sample p of host signal, we will quantize the sample p with the quantizer indexed by m_i. In this way if the minimum distance between two different quantizers is higher than the distortion (1) in the host signal, due to attacks or some commons manipulations, we hope that in the decoding process the sample quantized p' was closest to its quantizer that any other quantizers.

Fig. 1 illustrates the example of quantizers for binary symbols. The reconstruction points for $1's$ are represented with x and for $0's$ with o. If we embed 1, the host signal sample p is quantized with the x-reconstruction point closest to p, in other way for $0's$ we use the closest o-reconstruction point. In QIM the quantizer shape is very important, for example in Fig. 2 we have the x-shape and all the points inside this Vornoi-region have closest distance with the reconstruction point in the center of x-shape. When host signal is distorted due to an attack, the quantized point p' changes but if this distortion remains inside x-shape, the point p' always will be inside and the bit recovered correctly.

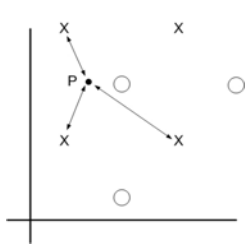

Figure 1. Quantizer for binary symbols.

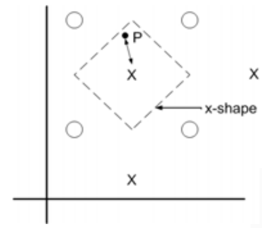

Figure 2. Quantizer Region.

A. Dither Modulation.

Chen et. al. proposed a way to add security to the watermark by dithered quantizers, this technique are called DM. This technique has the property that the quantization points of any given quantizer in the ensemble are shifted versions of the quantization points of any other quantizer in the ensemble [7]. The shifts typically correspond to adding pseudorandom vectors called dither vectors to host signal. In this way in the embedding function (2) we only use one base quantizer and n different dither vectors $d(i)$.

$$p' = q(p + d(i), m_i) - d(i) \tag{2}$$

The first dither vector is generated by a private key and the next dither vectors $d_j(k)$, where j is the j-element of dither vector k, are generated by (3). The length of the vectors is based on the dimensions or the sample number that we modulate for each embedded bit.

$$d_j(k) = \begin{cases} d_j(k-1) + \Delta/2 & d_j(k-1) < 0 \quad j = 1,\ldots,L \\ d_j(k-1) - \Delta/2 & d_j(k-1) > 0 \quad k = 2,\ldots,n \end{cases} \tag{3}$$

where

L is the number of dimensions.

Q is total number of quantizers.

III. ATHC

In information hiding applications, low-rate codes can be used to reliably embed information into digitized audio or video because we have many places where embedded the information. Recently a powerful low-rate was proposed by L. Ping et. al. in 2003, turbo-Hadamard codes (THC) [8]. THC can achieve performance of bit error rate = 10^{-5} at Eb/N0 ≈ -1.2 dB with an interleaver size of 65534. THC outperform superorthogonal turbo codes (SOTC) known as good low-rate codes in several interleaver sizes and with low complexity [9]. One kind of THC is ATHC (Fig. 3), this one has a good performance compared to original THC in waterfall region and faster convergence speed.

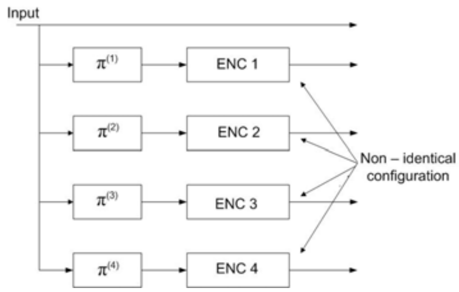

Figure 3. ATHC – Four parallel concatenation.

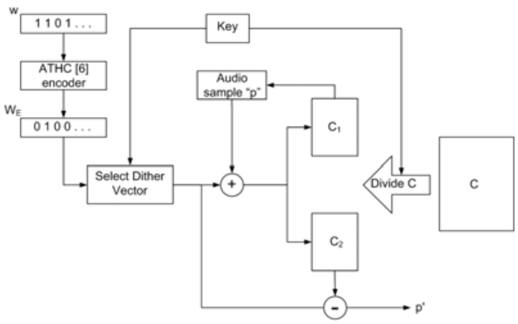

Figure 4. Coding process.

In ATHC the N information symbols are split into N/r blocks of r bits, w_k, $k = 1,\ldots,N/r$. A single parity bit is q'_k, is computed for each block. The bits q'_k are input to an S-state convolutional code $\hat{C}$ with a specified generator polynomial, to produce q_k. An order-r produce the systematic codeword $\{w_k, q_k, p_k\}$ given the codeword $\{w_k, q_k\}$, where p_k is the parity.

In this method we use an ATHC based on feed-forward encoder with four $(1,1/3)_8$ component convolutional codes, and two asymmetric designs, code A and code B which use mixtures of $(1,3)_8$ and $(1,1/3)_8$ component convolutional codes.

IV. Proposed Method

A. Coding.

First we obtain a robust watermark W_E from ATHC encoder. With a private key generated the corresponding dither vector (section 3-A) and divide the main codebook C in c_1 and c_2. From C we select the codeword which has the lowest quantization error and put in c_1 and the other ones in c_2.

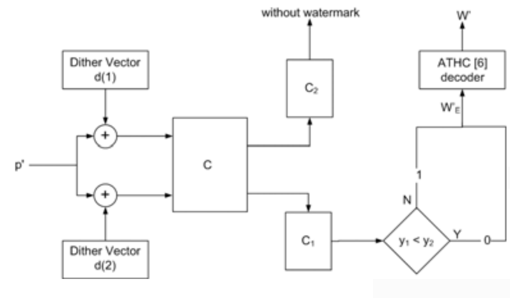

Figure 5. Decoding process.

According to watermark W_E, we select the corresponding dither vector $d(k)$ and add the audio sample $p(t)$. If $p(t)$ is closer to c_1 than c_2, then it is watermarked according to (2), otherwise the sample is left unchanged. In this way we spread the watermark into all audio file and at the same time add security in the decoding process.

B. Decoding.

In the decoding process (Fig. 5) we need the private key and the code-book C. With the private key generate the dither vectors and divide the codebook C. Take the first watermark sample $p'(t)$ and add, one by one, all dither vectors. As a result we will obtain two new samples y_1 and y_2. Search the closest codeword in C for y_1 and y_2, if the closest codeword is into c_2 the sample $p'(t)$ don't have embedded bit, but if is into c_1 we have a bit inside the sample $p'(t)$. Then, if the distance between the c_1-codeword and y_2 is higher that y_1 and the same c_1-codeword we have a 0 embedded in this audio sample, the other hand we have 1.

The distance between the codeword and watermarked audio samples is measure by squared Euclidean distance (4).

$$d = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{4}$$

Repeat this process for all audio samples until we have all watermark bits W'_E. Once we have all bits, we divide the bit stream in n blocks of r length and feed the ATHC decoder to correct error introduced by distortion. As a result we will obtain the recovered watermark w'.

V. Experimental Results

We evaluate the algorithm in base to watermark's imperceptibility and watermark's clarity in the extraction.

The watermark is a binary logotype of 64x64 pixels, the host signal is a digital audio file sampling at 44.1 kHz. In the embedding process each bit was embedded in four audio samples. The quantization was made in third least significant bit of the audio file.

Watermark's imperceptibility was measure by the relationship signal to noise (5).

$$SNR(dB) = 10\log_{10}\left(\frac{P_x}{P_{noise}}\right) \quad (5)$$

Where P_x is average power of audio signal and P_{noise} is the difference between original audio signal and watermarked audio signal.

In the case of watermark's clarity, we use the normal correlation (NC), given by (6). The NC is a similarity measure between two binary images, if the NC is equal to 1, this indicates that both images are identical. And if NC equal to 0, they are uncorrellated.

$$NC = \frac{\sum_{i,j} W_{i,j} W'_{i,j}}{\sum_{i,j} W_{i,j}^2} \quad (6)$$

Table 1 summarize the results, this table show the attacks made to watermarked audio file, the distortion suffered and the watermark's clarity.

TABLE I. RESULTS – SNR AND NC.

Attack.	***SNR***	***NC***
Without attack.	29.62505 dB	1
Compression mp3 128 kbps	22.46018 dB	.9977444
Add noise	17.04922 dB	.7759398
Normalize	27.04391 dB	.9856456
Zero cross attack	25.52342 dB	.7578947
Increasing stereo part	28.46380 dB	.9856186
Statistical evaluation	21.41455 dB	.7616541
Set all least significant bits to "0"	28.46170 dB	1
Inver all samples	28.46380 dB	.7984962

VI. CONCLUSION

In this paper we described a technique to embed a binary logotype into digital audio. First we create a robust watermark using ATHC codes, then we embedded the robust watermark by means of DM technique in the time domain.

The DM technique is a good method against attacks such as compression (mp3), additive noise and some frequency attacks but it results inefficient over attacks that change the wave-form in the time domain like a modulation.

This deficiency is results of embedded in time domain, but DM technique is better that other algorithms in time domain, such as low bit modulation, DC-Shifting. The advantage of embedded in time domain is a higher speed in the coding-decoding process.

REFERENCES.

[1] T. Furon and P. Duhamel. "Copy protection of distributed contents: An application of watermarking technique". *Workshop COST 254: Friendly Exchange through the net*, Bordeaux, France, March 2000.

[2] B. Chen, G.W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding", *IEEE Transaction on Information Theory,* vol. 47, No. 4, May 2001.

[3] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes (1)," *Proceedings of IEEE ICC '93*, (Geneva, Switzerland), pp. 1064–1070, May 1993.

[4] K. Harada, "Separate coding and iterative demodulation-decoding in digital communications", Ph.D. dissertation, Osaka City University, Osaka, Japan, 2003.

[5] N. Shimanuki, "On the construction of very low rate concatenated codes and their decoding", M.S. thesis, University of Electro-Communications, Tokyo, Japan, 2006.

[6] N. Shimanuki, K. Yamaguchi, K. Kobayashi, B. M. Kurkoski, "Improvements and extensions of low-Rate turbo-Hadamard codes", *IEEE International Symposium on Information Theory and its Applications,* Seoul, Korea, October 2006.

[7] B. Chen, G.W. Wornell, "Quantization index modulation methods for digital watermarking and Information Embedding of Multimedia, *Journal of VLSI Signal Processing*, vol. 27, pp. 7-33, 2001.

[8] L. Ping, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," *IEEE Transactions Information Theory*, vol.49, pp.3213-3224, Dec. 2003.

[9] P. Komulainen and K. Pehkonen, "Performance evaluation of superorthogonal turbo codes in AWGN and flat Rayleigh fading channels," *IEEE Journal on Selected Areas Communications*, vol.16, pp.196-205, Feb. 1998.