Analysis of QIM-Based Audio Watermarking Using LDPC Codes

⋆R. Martínez-Noriega, †M. Nakano, ⋆⋆B. Kurkoski, ⋆⋆K. Yamaguchi and ‡K. Kobayashi
⋆*Graduate School of Electro-Comms.*, ⋆⋆*Dept. of Inf. and Comms. Eng.*,
University of Electro-Communications, 1-5-1 Chofugaoka, 182-8585 Tokyo, Japan.
Email: {raul,kurkoski,yama}@ice.uec.ac.jp
†*Sec. of Research and Graduate Students, National Polytechnic Institute,*
Av. Santa Ana 1000, Mexico City 04430, Mexico.
‡*National Institute of Information and Communications Technology,*
4-2-1 Nukui-Kitamachi, Tokyo 184-8795, Japan.

Abstract

In this paper, we analyze audio watermarking methods based on quantization index modulation and low-density parity-check (LDPC) codes. We found that dither modulation (DM) can achieve better performance using half-rate Margulis LDPC code even better than some low-rate codes. Then, we propose a scheme based on LDPC codes and DM with distortion-compensation (DC) property which has a robustness benefit of 6 dB versus uncoded case, 2 dB versus algebraic codes, 1 dB versus DM with LDPC code. In DM with DC property, we show that it is possible to achieve .5 dB better robustness using a scale parameter α lower than the theoretically optimal and LDPC codes. Finally our proposal was evaluated against more practical attacks. These results show that our scheme could be a good option for robust watermarks.

1. Introduction

In 2001 Chen *et al.* introduced a new watermarking scheme, called quantization index modulation (QIM) [1], based on the framework of communications with side information [2]. Due to its ease of implementation and amenability to theoretical analysis, several variations of QIM has been proposed. For example, dither modulation (DM) is the lowest complexity implementation of QIM and distortion-compensation (DC) is a post-quantization processing which improves QIM's robustness.

Error control coding (ECC) and watermarking have been used before, however important results have been obtained only on images. Baudry *et al.* [3] introduced ECC strategies in image watermarking, they analyzed the performance of BCH codes, repetition codes and their concatenations. In [4], Gu *et al.* claimed that for common signal processing including compression and noise corruption in images, ECC can not improve the robustness of watermarking.

For audio watermarking, turbo codes have been used in [5] with spread spectrum but not with QIM-based techniques. Hernandez *et al.* [6] asserted that repetition codes seem to be the optimal choice for rational dither modulation. In [7] QIM and low-density parity-check (LDPC) codes [8] are used, however they do not show any results related with watermark robustness besides synchronization. Also our research group has published an article on LDPC codes for digital watermarking [9].

It has been shown in [4] and [6] that BCH, Golay and Hamming codes are weak codes in watermarking channels. On the other hand, LDPC codes work under the general principle that the longer codeword length, the closer they are to the theoretical channel capacity and its decoding complexity increases linearly to its codeword length.

In this paper, we explicit show the advantage of LDPC codes in watermarking channels and we propose an scheme based on DM with DC property (DC-DM) and LDPC codes. We divide the analysis in two parts: results against additive white Gaussian noise (AWGN) and results against common signal processing.

In the analysis against AWGN, we found that DC-DM with half-rate Margulis LDPC code obtains the best performance. We show that better performance can be achieved if DC-DM uses a scale factor α lower than the theoretically optimal. Our proposal has benefit of 2 dB against BCH codes and 5 dB against repetition codes (uncoded schemes). Moreover, results against common signal processing show that the watermark resists a compression of 96 kbps and low-pass filtering of 800 Hz.

2. Quantization index modulation

In QIM the host signal x is quantized according to the watermark symbol to be embedded. Each quantizer Q_j is

978-1-4244-4480-9/09/$25.00 ©2009 IEEE

associated with a different watermark symbol $j \in \mathbb{W}$, where $\mathbb{W}$ is the set of watermark symbols. For a binary watermark message m only two quantizers, Q_0 and Q_1, are defined. $s = Q_j(x, \Delta)$ describes the embedding function, where Δ is the step quantization size and s is the watermarked signal.

The watermark extraction is done by computing the distance between s and its closest quantization points. The minimum distance decides which quantizer was the most likely quantizer used in the embedding phase. Therefore, if the closest distance belongs to Q_j, the watermark symbol j is recovered. Extraction function is $m' = \arg\min_j |s - Q_j(s, \Delta)|$, where m' is the recovered watermark.

2.1. Dither modulation

DM is a QIM's variation which uses scalar, uniform and lattice-based quantizers. DM only uses one base quantizer Q and $|\mathbb{W}|$ dithered vectors. The embedding function is $s = Q(x + d_i^j, \Delta) - d_i^j$ where d^j is a dither vector which modulates the watermark symbol j. If the watermark is binary, two dither vectors are computed. The first vector is pseudo-random generated using an uniform distribution between $[-\Delta/2, \Delta/2]$ and the second with

$$d_i^1 = \begin{cases} d_i^0 + \frac{\Delta}{2}, & d_i^0 < 0 \\ d_i^0 - \frac{\Delta}{2}, & d_i^0 \geq 0 \end{cases},$$

for $i = 1, ..., l$ where l is the number of host samples per embedded symbol. Watermark extraction is similar to QIM's decoding.

2.2. Distortion compensation

DC is a process that improves distortion-robustness tradeoff of QIM. Given a quantizer ensemble, DC scales all the quantizers by $0 < \alpha \leq 1$. Then the square minimum distance between reconstruction points will be increased, therefore the robustness too. However the distortion is also increased, adding back a fraction $1 - \alpha$ of the quantization error to quantization value compensates this additional distortion [1]. The embedding function of QIM with DC property is $s = Q_j(x, \Delta/\alpha) + (1 - \alpha)\big[x - Q_j(x, \Delta/\alpha)\big]$.

The optimal choice for the scale parameter α in ideal case depends on the watermark-to-noise power ratio, WNR$= 10\log_{10}(\sigma_w^2/\sigma_v^2)$, and it is given by

$$\alpha_{opt} = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_v^2} = \frac{1}{1 + 10^{-\text{WNR[dB]}/10}}, \quad (1)$$

where σ_v^2 is the noise power and σ_w^2 is the watermark signal power, $w = s - x$. The ideal case assumes that the quantizers are based in a huge random codebook.

3. Proposal scheme

In our scheme the watermark is encoded with LDPC encoder and then it is embedded in the audio file using DM. At decoding phase, soft-information is obtained by applying again DM to the watermarked audio. Then, the soft-information is forwarded to LDPC decoder and the watermark is recovered.

3.1. Encoder

The audio file is divided in blocks of 512 samples, for each block *haar* wavelet transform in five levels is computed. After wavelet transformation, 32 frequency samples are obtained. The first 16 samples are viewed to be approximation coefficients and the last 16 samples are called detail coefficients. Since small changes in approximation coefficients produce big distortion in the audio file, the watermark is embedded in the detail coefficients x'.

The watermark is encoded using LDPC encoder and each bit from encoded watermark is repetitive embedded l times using DM with DC property (DC-DM) which is represented by $s' = \{Q\big(x' + d_i^j, \frac{\Delta}{\alpha}\big) + (1-\alpha)[x' - Q(x' + d_i^j, \frac{\Delta}{\alpha})]\} - d_i^j$.

Detail coefficients x' are replaced by watermarked detail coefficients s'. Reconstruction of the watermarked audio is done by inverse wavelet transform.

3.2. Decoder

An audio signal $y = s + v$, where v is the noise produced by attacks, is expected at the decoder. Attack is defined as any process that makes changes to audio signal s. The attacked audio y is decomposed with wavelet transform.

There are several ways to obtain soft-information from DM, the traditional fashion is computing the distance between the watermarked coefficient y' and its closest quantizer. We propose more reliable soft-information $r = D_1 - D_0$, where D_j means Euclidean distance between y' and $Q(y' + d_i^j, \Delta) - d_i^j$, which gives better performance than the traditional one. According to several experiments, we have noticed that the soft-information r behaves similar to Gaussian distribution. Therefore LLR$= 2r\mu/\sigma^2$ was computed using a Gaussian model, where μ is the mean of $|r|$ and σ^2 is the variance of v. LLR information is forwarded to LDPC decoder and using sum-product algorithm [10], the watermark is recovered.

4. Results

Audio files in WAVE format sampled at 44.1 kHz were used. The quantization step size was fixed to $\Delta = .02$. This Section is divided in two parts, in the first part we will focus on the analysis against AWGN and in the second part we will show results against audio signal processing.

Let R_t and R_f be the watermark rates computed in time and frequency domain respectively. $R_f = r_w \times r_{ldpc}$ is measured in bit per samples, where r_w is the embedding rate and r_{ldpc} refers to LDPC code rate. On the other hand, R_t is measured in bit per second (bps).

4.1. Analysis against AWGN attack

Fig. 1 shows a simulation using DM with LDPC codes and uncoded DM, all of them with $R_f = 1/4$ bit per sample. Uncoded schemes use repetition codes and no other ECC. First, we observe that using LDPC codes better performance can be achieved in comparison with repetition codes (uncoded scheme). Second, the performance of DM with LDPC codes is variable depending on the code, the best result is obtained with a Margulis LDPC code $r_{ldpc} = 1/2$ and DM with $r_w = 1/2$. Since the watermark rate is the same for all schemes, the watermarked audio quality is the same in the uncoded schemes as well as coded schemes. Those LDPC codes were chosen because they obtained the best performance among 30 tested codes.

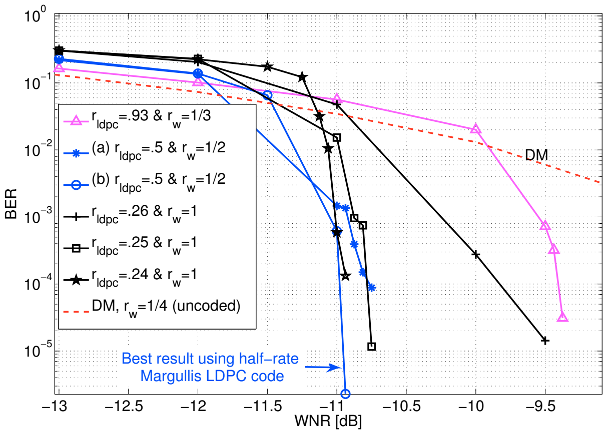

Figure 1. Performance of DM with LDPC codes and uncoded DM.

DC increases DM's robustness by scaling the step quantization size with a scale parameter α. In the ideal case, the optimal α_{opt} is given by (1), however ideal case is not practical because it involves huge random codes. For suboptimal codebook, e.g. the scalar and uniform codebook used in DM, the optimum value of α can be different. An approximation of optimal α for practical cases, [11], can be computed with

$$\alpha^*_{opt} = \frac{\sigma_w \sqrt{12}}{\Delta}. \tag{2}$$

Using (2) for DM with $R_f = 1/4$, we obtain $\alpha^*_{opt} = .9$. In Fig. 2 empirical results about different values of α reaffirm that DC-DM and $\alpha^*_{opt} = .9$ is the best result. DM has better performance than DC-DM with $\alpha < \alpha^*_{opt}$, nevertheless there is an $\alpha = .8$ which has good robustness in low WNR and acceptable in high WNR.

In the next simulations we will show results only with half-rate Margulis LDPC code which obtained the best performance in Fig. 1. Simulation of uncoded DM and DC-DM with LDPC codes is shown in Fig. 3. This result shows

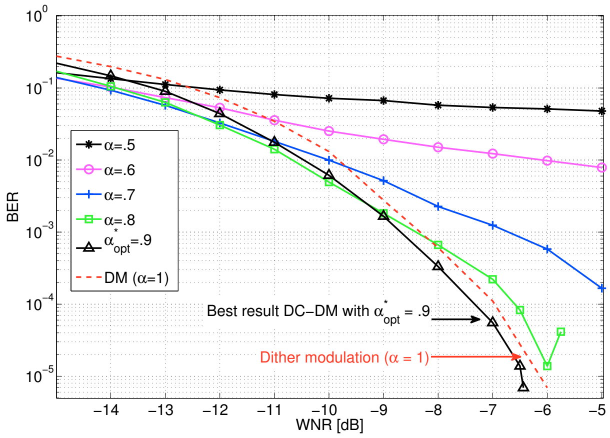

Figure 2. Performance of uncoded DM and uncoded DC-DM schemes with different α.

that it is possible to achieve better performance using DC-DM and LDPC codes with α lower than the theoretically optimal. DC-DM with LDPC codes and $\alpha = .8$ is the best result among all simulations against AWGN.

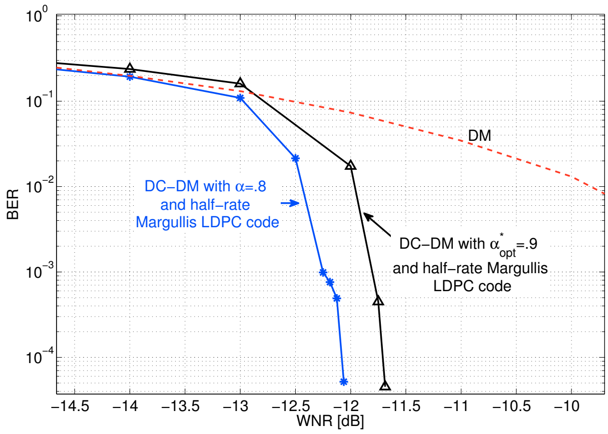

Figure 3. Performance of DM and DC-DM with α lower than theoretically optimal.

For the next simulation, we fix $\alpha = .8$ and we only use half-rate Margullis LDPC code. Fig. 4 is a summary of the watermarking schemes described in this paper and it also includes a comparison with half-rate BCH codes. Our proposal, DC-DM with $\alpha < \alpha^*_{opt}$ and half-rate Margulis LDPC code, has a benefit of 5 dB against DC-DM using $\alpha^*_{opt} = .9$, 2 dB against DC-DM with BCH codes and 1 dB against DM with LDPC codes.

From this Section we conclude that LDPC codes achieve better performance than repetition codes against AWGN and we found that the best robustness for DM with LDPC codes is using a half-rate Margulis code.

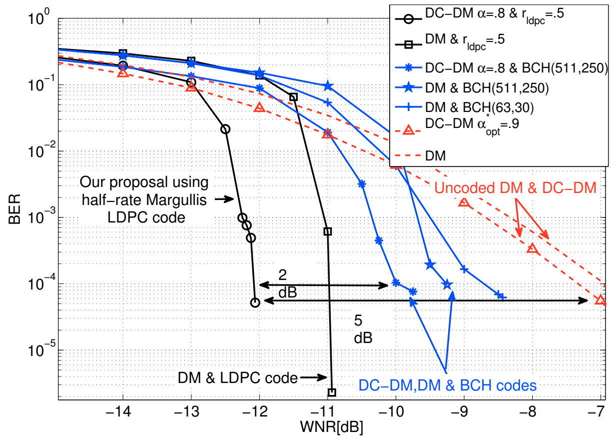

Figure 4. Comparison of proposal schemes against schemes with BCH codes.

4.2. Results against audio signal processing

We took the best result from Sec. 4.1, DC-DM with $\alpha = .8$ and half-rate Margulis LDPC code (DC-DM LDPC), and it is tested against more practical attacks. In these simulations watermarked audio quality, SNR= 33 dB, is good which means that the watermark does not introduce audible distortion in the audio files. Our definition of SNR is the ratio between the power audio signal x and the power of watermarked audio signal s. Fig. 5 shows MP3 compression and low pass filtering results with three different rates. The watermark is recovered from a compression of 96 kbps and also it resists a low-pass filtering with a cutoff frequency of 800 Hz.

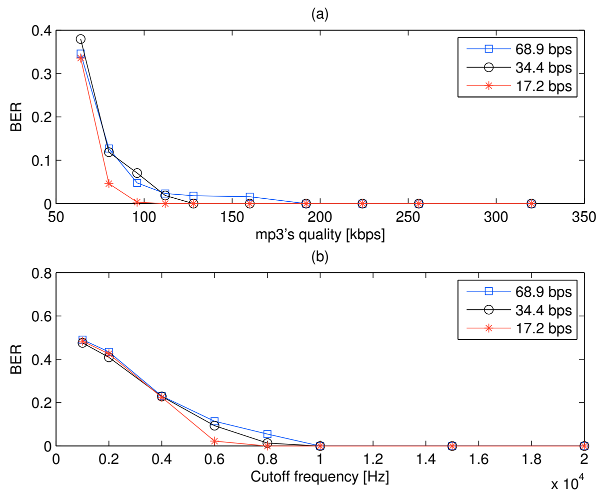

Figure 5. Watermark robustness against: (a) MP3 and (b) low-pass filtering.

We also tested our scheme with StirMark Audio benchmark with good results. Only attacks like *copysample* and *zerocross* destroy the watermark because they applied desynchronization at the decoder. The aim of this paper is not synchronization, therefore we consider that the decoder is in perfect synchronization with the watermark.

5. Conclusion

QIM-based watermarking schemes with LDPC codes were analyzed. We show that DM with LDPC codes achieve better performance than repetition codes. The best combination is using half embedding rate and half-rate LDPC codes. Specially, we found that half-rate Margulis code obtains the best performance, even better than some low-rate codes.

We proposed a watermarking scheme which uses LDPC codes and DC-DM with α lower than the optimal, it obtained a benefit of .5 dB in comparison with the traditional scheme using α^*_{opt}. Our scheme was also compared with algebraic codes like half-rate BCH codes and it obtains a benefit of 2 dB.

Finally, our scheme was tested against common signal processing and it shows that our proposal is a good candidate for robust audio watermarking schemes.

References

[1] B. Chen and G. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Inf. Theory*, vol. 47, no. 4, 2001.

[2] M. Costa, "Writing on dirty paper," *IEEE Trans. on Inf. Theory*, vol. IT-29, pp. 731–739, 1983.

[3] S. Baudry, J. F. Delaigle, B. Sankur, B. Macq, and H. Matre, "Analyses of error correction strategies for typical communication channels in watermarking," *Signal Processing*, vol. 81, no. 6, pp. 1239–1250, 2001.

[4] L. Gu, Y. Fang, and J. Huang, "Revaluation of error correcting coding in watermarking channel*," *Lecture Notes in Computer Science*, vol. 3810, pp. 274–287, 2005.

[5] N. Cvejic, D. Tujkovic, and T. Seppanen, "Increasing robustness of an audio watermark using turbo codes," in *Int. Conf. on Mult. and Expo*, 2003, pp. 217–220.

[6] J. Hernandez, M. Nakano, and H. Meana, "Rational dither modulation in audio signals," in *50th Midwest Symp. Circuits and Systems*, 2007, pp. 109–112.

[7] D. J. Coumou and G. Sharma, "Watermark synchronization for feature-based embedding: application to speech," in *IEEE Int. Conf. on Mult. and Expo*, 2006, pp. 849–852.

[8] R. Gallager, "Low-density parity-check codes," *IRE Trans. on Inf. Theory*, vol. 8, no. 10, pp. 21–28, 1962.

[9] R. Martinez-Noriega, M. Nakano, B. Kurkoski, K. Yamaguchi, and K. Kobayashi, "Robustness analysis of audio QIM watermarking methods with LDPC codes," in *Proc. of Triangle Symp. on Advanced ICT*, 2008, pp. 231–236.

[10] D. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inf. Theory*, vol. 45, pp. 399–431, 1999.

[11] J. J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar costa scheme for information embedding," *IEEE Trans. on Signal Proc.*, vol. 51, no. 4, pp. 1003–1019, 2003.