IEICE Electronics Express, Vol.5, No.7, 217–222

Data hiding in audio signal using Rational Dither Modulation

Jose Juan Garcia-Hernandez[a], **Mariko Nakano-Miyatake**[b], **and Hector Perez-Meana**[c]

Mechanical and Electrical Engineering School, National Polytechnic Institute of Mexico, Mexico City, Mexico

a) *jjgarcia@calmecac.esimecu.ipn.mx*

b) *mariko@calmecac.esimecu.ipn.mx*

c) *hmpm@calmecac.esimecu.ipn.mx*

Abstract: A data hiding system in audio signals based on a Rational Dither Modulation is proposed, which hides information in the Modulated Complex Lapped Transform (MCLT) domain. The proposed system is able to hide around of 689 bits per second, while keeping a CD-quality audio signal. Objective and subjective evaluations show robustness to classical attacks and transparency to the Human Auditory System (HAS), respectively. Comparison results, in similar conditions, with some other algorithms reported in the literature are also provided.

Keywords: data hiding, QIM, RDM, MCLT domain, HAS

Classification: Science and engineering for electronics

References

[1] B. Chen and G. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1423–1443, May 2001.

[2] J. Eggers, R. Bäuml, and B. Girod, "Estimation of amplitude modifications before scs watermark detection," in *Security and Watermarking of Multimedia Contents IV, Proc. SPIE*, vol. 4675, pp. 387–398, 2002.

[3] K. Lee, D. Kim, and K. A. Moon, "Em estimation of scale factor for quantization-based audio watermarking," in *Proceedings of Second International Workshop on Digital Watermarking*, Springer Verlag, pp. 316–327, 2003.

[4] F. P. Gonzalez, C. Mosquera, M. Barni, and A. Abrardo, "Rational dither modulation: a high rate data-hiding method invariant to gain attacks," *IEEE Trans. Signal Process.*, vol. 53, pp. 3960–3975, Oct. 2005.

[5] M. Steinebach, J. Dittmann, C. Seibel, L. C. Ferri, F. A. Petitcolas, N. Fates, C. Fontaine, and F. Raynal, "Stirmark benchmark: Audio watermarking attacks," in *International Conference on Information Technology: Coding and Computing (ITCC '01)*, 2001.

[6] W. Yang, M. Dixo, and R. Yantorno, "A modified bark spectral distortion measure which uses noise masking threshold," in *IEEE Speech Coding Workshop*, pp. 55–56, 1997.

[7] W. Yang, M. Benbouchta, and R. Yantorno, "Performance of the modified bark spectral distortion as an objective speech quality measure," in

DOI: 10.1587/elex.5.217
Received October 01, 2007
Accepted November 20, 2007
Published April 10, 2008

IEICE Electronics Express, Vol.5, No.7, 217–222

ICASSP, vol. 1, pp. 541–544, 1998.
[8] D. Kirovski and H. Malvar, “Spread spectrum watermarking of audio signals,” *IEEE Trans. Signal Process.*, vol. 51, pp. 1020–1033, April 2003.
[9] I.-K. Yeo and H. J. Kim, “Modified patchwork algorithm: A novel audio watermarking scheme,” *IEEE Trans. Speech Audio Process.*, vol. 11, no. 4, 2003.
[10] H. Tsai, J. Cheng, and P. Yu, “Audio watermarking based in has and neural networks in dct domain,” *EURASIP Journal on Applied Signal Processing*, pp. 252–263, 2003.
[11] H. S. Malvar, Signal Processing with Lapped Transforms, Artech House, Inc., 1992.

1 Introduction

The Quantization Index Modulation (QIM) [1] appears to be a practical solution to the digital information hiding problem. The main task in QIM based method, such as Dither Modulation, is the design of suitable quantizers used to embed the data. However such simple method, as several other QIM based methods, presents lack of robustness against the gain attack. In order to reduce that vulnerability, several schemes have been proposed [2, 3]. Rational Dither Modulation (RDM) was introduced as a solution to gain attack in high-rate data hiding schemes [4], however, its performance was studied only for the Gaussian channels case without considering other kinds of signal distortions. On the other hand, Modulated Complex Lapped Transform (MCLT) is considered as an excellent domain to realize imperceptible data hiding. In this paper, a RDM scheme based on MCLT is proposed in which robustness against typical attacks to audio data hiding system; including gain attack and higher embedding payload are got, while keeping the imperceptibility of the hidden data.

2 Proposed Data Hiding System

The proposed data hiding system consists of embedding and decoding process, respectively. The embedding process is carried out as follows: firstly host audio signal is divided in frames of 128 samples per frame. Next each frame is transformed using the MCLT. Subsequently magnitude and phase of MCLT are computed. The RDM is then applied to the MCLT magnitudes, while keeping phase without change. Finally inverse MCLT is applied to processed magnitude and original phase to get the audio signal with hidden data. In the receiver side, the decoding process is carried out as follows: firstly audio signal with hidden data is divided in frames with 128 samples per frame. Next each frame is transformed using MCLT whose magnitude is then computed, finally RDM decoding algorithm is applied to the MCLT magnitude to extract hiding message. These processes are shown in Fig. 1. The coding and decoding rules of RDM are given as follows [4],

$$y_k = g(\mathbf{y_{k-1}})Q_{b_k}\left(\frac{x_k}{g(\mathbf{y_{k-1}})}\right) \tag{1}$$

DOI: 10.1587/elex.5.217
Received October 01, 2007
Accepted November 20, 2007
Published April 10, 2008

IEICE Electronics Express, Vol.5, No.7, 217–222

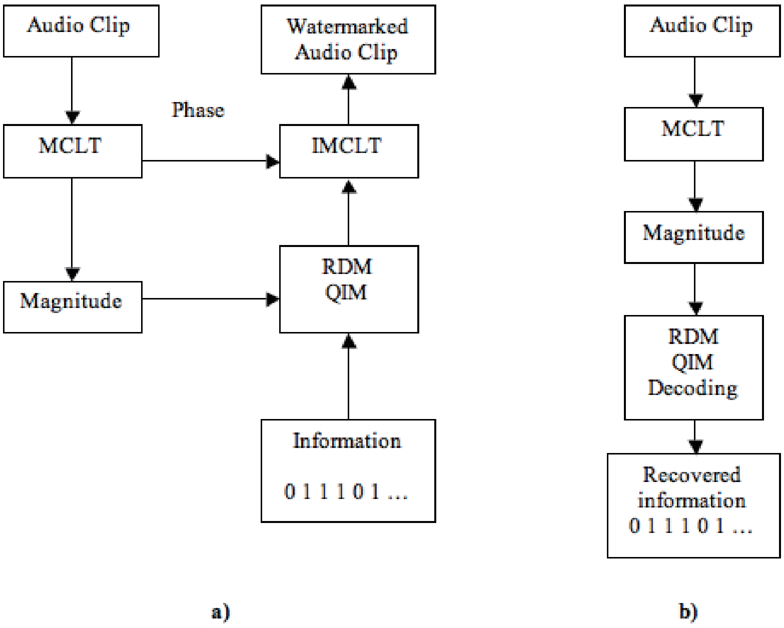

Fig. 1. Proposed data hiding system based on MCLT and RDM. a) Data embedding stage, b) data recovery stage.

where y_k is the RDM sample, $\mathbf{y_{k-1}}$ is a vector of $k-1$ past RDM samples, x_k is the host sample, Q_{b_k} is a message dependent quantizer and g is a function satisfying properity:

$$g(p\mathbf{y}) = pg(\mathbf{y}) \tag{2}$$

where p is the gain factor. On the other hand, the decoding is given by

$$b_k = arg\ min \left| \frac{z_k}{g(\mathbf{z_{k-1}})} - Q_{b_k}\left(\frac{z_k}{g(\mathbf{z_{k-1}})}\right)\right| \tag{3}$$

where b_k is the decoded bit, z_k is the received signal and $\mathbf{z_{k-1}}$ is a vector of $k-1$ past received signals. The problem of choosing a particular g function is very involved due to the intrinsic nonlinearity of the quantization process, authors of [4] suggests one subset based on Holder or l_p vector-norms:

$$g(\mathbf{y_{k-1}}) = \left(\frac{1}{L}\sum_{m=k-L}^{k-1} y_m^p\right)^{\frac{1}{p}}, p \geq 1 \tag{4}$$

where L is the number of past RDM samples utilized in the data hiding process. In this paper the average function ($p = 1$ in Eq. (4)) is considered as function g.

RDM in audio signals was carried out using frames with 128 MCLT coefficients instead of single samples in order to get a robust scheme not only to gain attacks but also to typical attacks in data hiding systems, such as the Audio Stirmark benchmark test [5].

3 System Evaluation

The system evaluation is carried out from imperceptibility and robustness point of view.

DOI: 10.1587/elex.5.217
Received October 01, 2007
Accepted November 20, 2007
Published April 10, 2008

IEICE Electronics Express, Vol.5, No.7, 217–222

3.1 Hidden data imperceptivity

In order to evaluate the audio quality after the embedding process, subjective and objective evaluations were carried out.

3.1.1 MOS (Mean Opinion Score)

The MOS regarding to the hidden data imperceptibility was carried out with a population of 100 persons using 5 different kinds of music: Classic music, Rock music, Pop music, Instrumental music and Latin music. The evaluation criterion used here are 5- imperceptibile, 4- Perceptible but not annoying, 3- Slightly annoying, 2- Annoying, 1- Very annoying. The MOS evaluation results are approximately 4.7 (4.8 for Rock, Pop and Latin music and 4.7 for Classic music and 4.5 for Instrument music). MOS evaluation results show that embedding data is imperceptible to Human Auditory System (HAS) and it is difcult to discriminate between original signal and the embedded one.

3.1.2 MBSD (Modified Bark Spectral Distortion)

The second evaluation to measure imperceptibility of hidden data was the MBSD [6, 7]. The MBSD measure estimates speech distortion in loudness domain taking in account the noise-masking threshold in order to include only audible distortions in the calculation of the distortion measure. This evaluation was carried out using 5 different kinds of music with the following results: for classic music −63.2 dB, for rock music −65.3, for pop music −62.8 dB, for instrumental music −61.7 and finally for latin music −62.5 dB.

3.2 Robustness

Table I shows the bit error rate (BER) of extracted hidden data after the Audio Stirmark attacks [5] are applied. Generally, the embedded data are robust to various types of attacks, except copysample, cutsample, echo, ffttest, flippsample and voiceremove attacks. However these attacks also considerably distorts to audio signal.

3.3 Comparison with conventional methods

Several audio data hiding methods have been proposed in the literature in the last several years. In this paper three algorithms representative of the state of the art [8, 9, 10] and a DCT based RDM scheme are considered to compare with performance of the proposed system. These representative algorithms are: a watermarking algorithm based in the spread spectrum theory (SS) [8], a modified patchwork algorithm (MP) [9] and neural networks based information hiding scheme (NN) [10]. Data hiding system based on RDM realized in DCT domain (DCT-RDM) is also provided for comparison. The comparison is performed in terms of payload (bit per second) and imperceptibility measured using MBSD. Table II shows that summary. Here the transparency in terms of MBSD is similar in all of the five compared systems; however, the proposed systems payload is the highest. The robustness to Audio Stirmark attacks of all systems is quite similar, as shown in to Table I. However if

DOI: 10.1587/elex.5.217
Received October 01, 2007
Accepted November 20, 2007
Published April 10, 2008

Table I. Bit Error Rate Results on Audio Signals Attacked With the Stirmark Audio Benchmark.

Attack	BER	Attack	BER	Attack	BER
Original	0.0	Addbrum 100	0.0	Addbrum 1100	0.0
Addbrum 3100	0.0	Addbrum 4100	0.0	Addbrum 5100	0.0
Addbrum 7100	0.0	Addbrum 8100	0.0	Addbrum 9100	0.0
addfftnoise	0.11	Addnoise 100	0.0	Addnoise 300	0.04
Addnoise 700	0.13	Addnoise 900	0.2	addsinus	0.0
compressor	0.2	copysample	0.8	cutsamples	0.8
echo	0.63	exchange	0.23	Extraestereo 30	0.18
Extraestereo 70	0.33	Fft hlpass	0.05	Fft invert	0.02
Fft stat	0.1	Fft test	0.88	flippsample	0.78
lsbzero	0.0	normalize	0.0	nothing	0.0
lowpass	0.0	resampling	0.0	smooth	0.11
stat1	0.08	stat2	0.07	voiceremove	0.75
Addbrum 2100	0.0	Addbrum 6100	0.0	Addbrum 10100	0.0
Addnoise 500	0.1	amplify	0.0	dynnoise	0.2
Extraestereo 50	0.22	Fft real reverse	0.01	invert	0.0
highpass	0.0	smoth2	0.12	zerocross	0.23

Table II. Metrics of several audio watermarking schemes.

	SS [8]	MP [9]	NN [10]	DCT-RDM N=256	MCLT-RDM (proposed) N=128
Payload (bps)	0.5-1	10	34	344	689
MBSD (dB)	−64.1	−63.6	−61.7	−64	−63.2

DOI: 10.1587/elex.5.217
Received October 01, 2007
Accepted November 20, 2007
Published April 10, 2008

these systems are modified in order to achieve similar payload to that of the proposed system, their robustness falls dramatically.

3.4 Performance of MCLT-RDM compared with DCT-RDM

The Table II shows that, with very similar MBSD, the MCLT-RDM provides higher payload than the DCT-RDM. This is because the distortion introduced by RDM, as in transform coding, increases as the block length N become shorter. This effect is not present in MCLT because this transform does not present blocking artifact [11], providing in such way lower MBSD values. To perform this comparison the MCLT-RDM was taken as reference, choosing the block length N in DCT-RDM in order to obtain a similar MBSD value; which was obtained with $N = 256$. On the other hand, if $N = 128$, as in the MCLT-RDM, then the MBSD value in the DCT-RDM system becomes around -50.5 dB.

4 Conclusions

In this paper a data hiding system in audio signals based on RDM in MCLT domain was proposed which achieve high payload, hidden data imperceptibility and robustness. Subjective and objective evaluations were carried out, in order to evaluate the hidden data imperceptibility. Evaluation results show that in comparison with conventional methods included RDM conventional scheme (DCT-RDM), the proposed system provides a higher payload, while keeping hidden data imperceptibility and robustness. The proposed system was implemented in DSP system such that its real-time behavior is guaranteed.

Acknowledgments

The authors would like to thanks The National Science and Technology Council (CONACyT) for financial support.

DOI: 10.1587/elex.5.217
Received October 01, 2007
Accepted November 20, 2007
Published April 10, 2008