

INFORMED AUDIO WATERMARKING SCHEME USING DIGITAL CHAOTIC SIGNALS

G. C. M. Silvestre[†], N. J. Hurley[†], G. S. Hanau[†] and W. J. Dowling[‡]

[†]National University of Ireland, Dublin, Ireland

[‡]University of Dublin, Ireland

ABSTRACT

In this paper, an informed audio watermarking scheme using chaotic waveforms is proposed. The watermark consists of an attenuated signal embedded in the original audio signal. A modulation scheme based on vector quantization is used to achieve blind detection of the watermark signal. A method for self-synchronization of the watermark detector is proposed using chaotic orthogonal waveforms. A perceptual model is employed to ensure inaudibility of the embedded signal that is shaped to the spectrum of the host signal. High robustness to external noise is achieved by spreading the watermark data over a period of time. The robustness of the scheme is assessed by computer simulations.

1. INTRODUCTION

With recent advances in perceptual audio coding, CD-quality digital audio can be distributed quickly and economically by means of broadband networks such as the Internet. However such electronic distribution has been impeded by the fear of copyright infringements and illicit dissemination. In this context, digital watermarking has emerged as a promising technique to combat such piracy.

2. INFORMED WATERMARKING

In previous work [1, 2], it was shown that high robustness can be achieved if the watermark signal is embedded in the perceptually most significant components of the host signal. In second generation watermarking systems [3–5], information about the original signal is exploited explicitly by the encoder. The watermarking process is understood as a channel coding problem, which requires an accurate characterization of the channel. Second generation watermarking schemes do not consider the host signal as noise to the watermark signal.

2.1. Watermarking Channel

In watermarking applications, the audio signal constitutes the channel for transmission of the watermark data. Defining the channel capacity is essential and, in the case of digital music, is a difficult problem that involves subjective considerations such as the maximum distortion that can be introduced to the music without producing noticeable artefacts. Psychoacoustic models can be used to determine non-perceptible components and the watermark energy is concentrated in un-exploited redundancy of the audio signal [7]. Watermarking algorithms are therefore in direct competition with lossy audio encoders such as MPEG-1 layers [8]. In addition, large interference and deformation of the channel may result from coding, signal processing operations or malicious attacks, and may seriously narrow the capacity. Providing that the bit-rate of the embedded data does not exceed the channel capacity, Shannon has shown that it is possible to achieve reliable transmission of the watermark [9]. In this work, we propose the use of Differential Chaos Shift Keying (DSCK) spread-spectrum techniques [10, 11] to embed watermarking data with good robustness against noisy environments.

2.2. Blind Watermarking

In blind watermarking schemes, decoding is achieved without recourse to the original signal. Therefore, if $w(t) = s(t) - s_w(t)$ is the watermark introduced in the original audio signal, $s(t)$, to produce the watermarked audio signal, $s_w(t)$, blind recovery of the embedded information requires $w(t)$ to be uncorrelated with $s(t)$. In [4–6], a dither modulation scheme was employed to modulate a digital watermark signal into images. In this approach, the noise introduced by a quantizer is dither-modulated using a digital pseudo-random watermarking map. If the size of the codebook is sufficiently large, the noise signal introduced by the quantizer is approximately white and uncorrelated with the audio signal, which guarantees blind recovery. A high degree

of robustness can be achieved by amplification of the modulated quantization noise [4]. However, invisibility of the mark cannot be guaranteed. In addition, high performance perceptual coders or malicious attackers will tend to filter out uncorrelated noise and destroy the performance of such systems. In [5], elements of perceptual models for the human visual system are introduced to select the DFT components most relevant to the human eye, which are subsequently used to carry the mark. This results in the colouring of the quantization noise according to the signal spectrum, which defines a tradeoff between robustness and perception of the watermark.

3. PROPOSED SCHEME

3.1. Model Formulation

In this work, we investigate modulation schemes based on vector quantization [12], which exploit side information available at the watermarking encoder. In addition, we employ psychoacoustic models to determine the perceptual energy of the signal suitable to carry our watermark.

Our informed watermarking framework can be described using a similar description to that found in [3]. We extract a vector \vec{v} from the discrete time audio signal $s(n)$ so that $\vec{v} = f(s)$. The dimension of \vec{v} is that corresponding to a data symbol spread over a period of T samples. A modulation function, Q , is then applied to \vec{v} to embed T samples of a digital chaotic signal, $w(n)$, resulting in a watermarked vector $\vec{v}_w = Q(\vec{v}, w)$. The watermarked audio signal is given by $s_w(n) = g(\vec{v}_w, s)$, where g is the embedding process. The function Q is chosen such that the noise introduced in $s_w(n)$ is perceptually insignificant.

In this implementation, the extraction process f is applied to a number of audio frames, N . From each frame vector, we extract 25 vectors corresponding to each critical band so that $T = 25 \times N$. The modulation takes place in the Fourier domain using a dither vector quantization scheme and the digital chaotic map. Psychoacoustic models [13] are introduced in the quantization process to ensure good perceptual characteristics.

3.2. Chaotic Map Construction and DSCK

Digital chaotic maps are a class of deterministic dynamical system admitting non-periodic signals characterized by a continuous ‘noiselike’ broad spectrum, which appears random in the time domain. The use of chaotic signals in digital communications has received much attention recently [10]. Conventional spread spectrum techniques use synchronizable pseudo-random se-

quences to spread the transmitted signal. These techniques are limited by the periodic nature of the pseudo-random sequence and by the fact that all sequences must be synchronized at the receiver. Researchers in chaotic communications have examined the use of chaotic spreading sequences as a means of achieving asynchronous CDMA. An alternative approach is to make the transmission ‘noiselike’ without spreading by representing the transmitted symbols as chaotic basis functions.

We exploit this latter approach in the development of a watermarking technique based on chaotic maps. Each watermark symbol bit, b_w , is represented by a chaotic map using a binary DSCK technique given by

$$w(n) = \begin{cases} c(n) & \text{for } 0 \leq n < T/2 \\ -c(n - T/2) & \text{for } T/2 \leq n < T \end{cases}$$

for a bit value 0, and by

$$w(n) = \begin{cases} c(n) & \text{for } 0 \leq n < T/2 \\ c(n - T/2) & \text{for } T/2 \leq n < T \end{cases}$$

for a bit value 1, where $c(n)$ is the digital chaotic map. The first half of the basis function is considered as the reference chip while the second part is the information bearing chip [10].

3.3. Watermark Detection

In conventional communications systems, each symbol to be transmitted is represented as a linear combination of periodic functions, and the symbol duration T is an integer multiple of the period of the basis functions. In a chaotic communications system, the sample function which represents the symbol is a chaotic basis function. The decision as to which symbol is received is made by estimating some property (for example, the energy) of the sample function at the receiver. Because the chaotic waveform is non-periodic, each sample function will be different over each period T . Hence, the energy over the symbol duration T is not constant, but can be considered as a random variable, centered at some mean value and taking values distributed around this mean. It can be shown that the standard deviation of the energy of a chaotic waveform scales approximately as $1/(T \times BW)$, where BW is the statistical bandwidth of the chaotic signal [10]. Hence, by choosing the symbol period large enough, we can ensure that energy is almost constant over each period. Furthermore, the auto-correlation

$$R(\tau) = \sum_{n=1}^T c(n)c(n - \tau)$$

falls rapidly to zero as τ is increased from zero.

The properties of the chaotic map allow watermark symbols to be retrieved at the decoder without knowledge of the chaotic waveform used by the transmitter. Binary DCSK can be demodulated using a differential coherent receiver. The received signal is delayed by half the bit duration and the correlation between the received signal and the delayed copy of itself is determined. Given the correlation properties of the chaotic waveform, a 0 is determined to have been received if the correlation is negative and, in absolute value, approximately equal to the energy over the period $T/2$. A 1 is determined if the correlation is positive (see Fig.1). The strong correlation property of chaotic maps combined with DSCK is used to self-synchronize the signal at the receiver. For secure applications, a pseudo-random delay can be added to the reference and information bearing chips. This pseudo-random sequence becomes a private user-key.

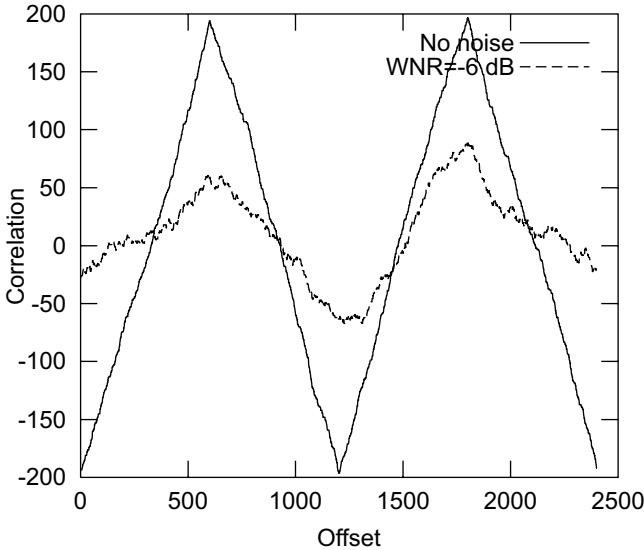


Fig. 1. The correlation of the chaotic signal retrieved at the decoder, with $T = 600$ and with $\text{WNR} = 10\log_{10}(\sigma_w^2/\sigma_n^2) = -6$ dB. T samples, starting from offset k are chosen, and the first half of the signal is correlated with the second half. The correlation peaks at values of $k = m \times T$ for integer values of m . A positive peak implies that a symbol 1 was received, while a negative peak implies a symbol 0 was received.

4. RESULTS

Since the watermark is placed in the perceptual domain, the watermark-to-signal ratio (WSR) is considered as an insufficient measurement of the noise introduced in the signal. Subjective listening tests were therefore conducted to determine the threshold of per-

ceptibility for the proposed algorithm. WSR values below -23 dB were not perceptible by untrained listeners. Computer simulations were carried out to test the robustness of the proposed watermarking scheme against perceptual filtering resulting from MPEG-1 coding (Fig. 2) and Additive White Gaussian Noise (AWGN) attacks (Fig. 3).

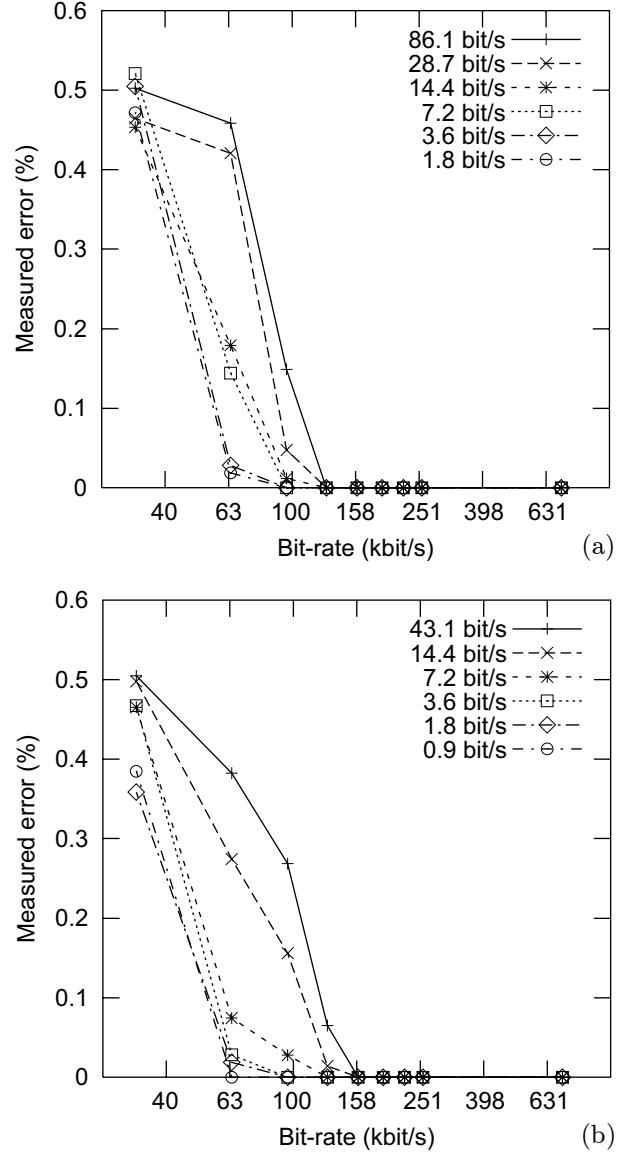


Fig. 2. Robustness of a DCSK watermarking scheme as a function of perceptual coding attacks using a MPEG-1 algorithm for different values of the watermark bit-rate. A dither-modulated vector quantization scheme was employed with a $\text{WSR} = 10\log_{10}(\sigma_w^2/\sigma_s^2)$ of -23 dB and a frame size of (a) 512 samples and (b) 1024 samples.

The watermarking algorithm was tested against compression rates ranging from 32 kbit/s up to 256 kbit/s

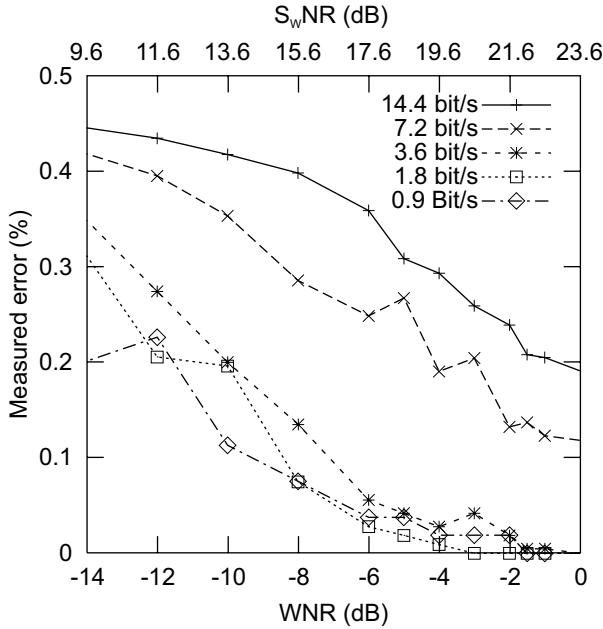


Fig. 3. Robustness of a DCSK watermarking scheme as a function of AWGN attacks for different values of the watermark bit-rate. A dither-modulated vector quantization scheme was employed with a frame size of 1024 samples and a WSR=-23 dB. $S_wNR = 10\log_{10}(\sigma_{s_w}^2 / \sigma_n^2)$.

using a 30 s mono audio signal sampled at 44.1 kHz, 16 bits/sample. It was found possible to embed a watermark signal of 7.2 bit/s which survived, without detection error, perceptual filtering down to 64 kbit/s. At very low rates, the error probability was found to be large. However, audio signals at such compression rates are severely decimated and coarsely quantized.

5. CONCLUSION

A robust informed watermarking scheme for embedding data in digital audio signals has been presented. Watermark symbols are spread over a large period of time using a digital chaotic map and DCSK. This ensures self-synchronization of the watermarking decoder and provides strong robustness to channel deformations frequently encountered in watermarking attacks. Elements of psychoacoustic models are used to devise a vector quantizer for modulating the digital chaotic map with suitable perceptual properties. Computer simulations and subjective listening tests were used to determine a tradeoff between robustness and perceptibility of the watermark. Also, it was shown that the proposed technique is robust to the perceptual filtering of audio encoding algorithms.

6. REFERENCES

- [1] G. C. M. Silvestre and W. J. Dowling, "Image watermarking using digital communication techniques," in Proc. of 6th IEE Int. Conf. on Image Proc. and its Applications, IPA'97, vol. I, pp. 443-448, 1997.
- [2] G. C. M. Silvestre and W. J. Dowling, "A data-embedding technique for digital images," in IEE Colloquium on Secure Images and Image Authentication, Savoy Place, London, pp. 81-86, April 10, 2000.
- [3] M. Miller, I. Cox, J. Bloom, "Informed embedding: exploiting image and detector information during watermark insertion," in Proc of 2000 IEEE Int. Conf. on Image Proc., Vancouver, Canada, pp. 1-4, v. III, September 10-13, 2000.
- [4] J. J. Eggers, J. K. Su, and B. Girod, "A blind watermarking scheme based on structured codebooks," in IEE Colloquium on Secure Images and Image Authentication, Savoy Place, London, pp. 41-46, April 10, 2000.
- [5] G. C. M. Silvestre and W. J. Dowling, "Embedding data in digital images using CDMA techniques," in Proc of 2000 IEEE Int. Conf. on Image Proc., Vancouver, Canada, pp. 589-592, v. I, September 10-13, 2000.
- [6] B. Chen and G. W. Wornell, "Preprocessed and post-processed quantization index modulation methods for digital watermarking," in Proc. of SPIE Security and Watermarking of Multimedia Contents II, San Jose, CA, USA, pp. 4859, v. 3971, January 2000.
- [7] L. Boney, A. H. Tewfik and K. N. Hamdy, "Digital watermarks for audio signals," in Proc. of IEEE Int. Conf. on Multimedia Computing and Systems., Hiroshima, Japan, pp. 473-480, v. I, June 17-23, 1996.
- [8] "Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s," Tech. Rep., ISO/CEI 11172-3,1993.
- [9] C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J., v. 27, pp. 379-423, 623-656, 1948.
- [10] M. P. Kennedy, R. Rovatti and G. Setti, Chaotic Electronics in Telecommunications, CRC Press LLC, 2000.
- [11] J. G. Proakis, Digital Communications, McGraw-Hill, 1983.
- [12] R. M. Gray and D. L. Neuho, "Quantization," IEEE Trans. on Information Theory, v. 44, no. 6, pp. 1-63, 1998.
- [13] E. Zwicker and H. Faslt, Psychoacoustics, Springer-Verlag, Berlin, 1999.