2010 Fourth International Conference on Genetic and Evolutionary Computing

Energy-Proportion Audio Watermarking Scheme in the Wavelet Domain

Shuo-Tsung Chen
Department of Mathematics
Tunghai University
Taichung, Taiwan 407
shough33@yahoo.com.tw

Huang-Nan Huang
Department of Mathematics
Tunghai University
Taichung, Taiwan 407
nhuang@thu.edu.tw

***Abstract*—This work presents a novel audio watermarking by using energy proportion instead of probability of entropy in information theory. The energy-proportion function is obtained by mapping low-frequency coefficients of discrete wavelet transform (DWT) into the domain of this function. Based on the characteristic curve and the properties of energy-proportion function, this study proposes a novel audio watermarking scheme. In addition, the proposed energy-proportion scheme can extract the watermark without original audio signal. Experimental results demonstrate that the embedded data are robust against re-sampling, low-pass filtering, and amplitude scaling.**

Keywords-Audio watermarking; energy-proportion function; DWT;

I. INTRODUCTION

With the development of internet and wireless transmission, digital watermarking has become more and more important. Many watermarking techniques [1-6] are proposed in recent years. In the time domain, Lie *et al.* [1] proposed a robust watermarking algorithm using the amplitude modification. Unfortunately, it has very low capacity. In the frequency domain, Wu *et al.* [2] use quantization index modulation to embed watermarks into the low-frequency coefficients of discrete wavelet transform (DWT). It has strong robustness against re-sampling, MP3 compression, and filtering. However, a small amount of amplitude or time scaling will be able to fail the extraction of the hidden watermarks. Kim *et al.* [3] proposed a new algorithm on the piecewise constant DWT to improve the traditional patchwork algorithm.

Entropy is a measure of the amount of information or uncertainty [7]. It quantifies the information contained in a message, usually in bits. Since a watermark is usually represented by binary bits, we apply this entropy to audio watermarking. This work rewrites entropy as an energy-proportion function and maps the low-frequency coefficients of DWT into the domain of this function. Then, the characteristic curve of energy-proportion function (CCEP) is obtained. Based on the CCEP and the properties of energy-proportion function, this study proposes an energy-proportion based scheme. In addition, the proposed scheme can extract the watermark without original audio signal.

This paper is organized as follows. Section II illustrates CCEP. For the design of watermarking, some properties of energy-proportion function are also derived in this section. Based on CCEP and the properties, the energy-proportion based scheme for embedding and extraction are proposed in section III. Experiments are also conducted to test the performance of proposed scheme in section IV. Finally, some conclusions are summarized in section V.

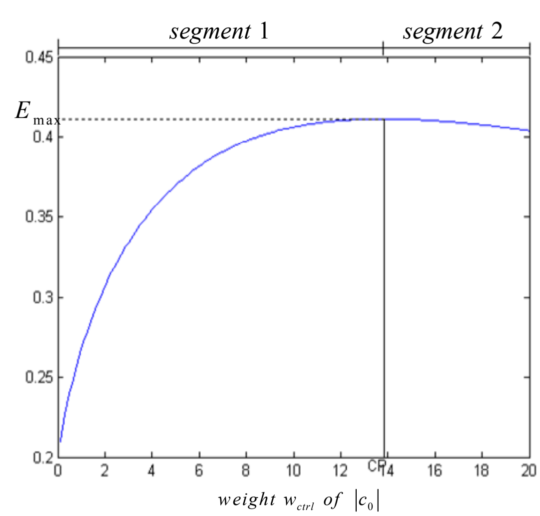
II. THE CHARACTERISTIC OF ENERGY-PROPORTION FUNCTION

In this section, we use energy proportion instead of probability to form an energy-proportion function. Furthermore, we analyze its characteristic and properties.

A. CCEP

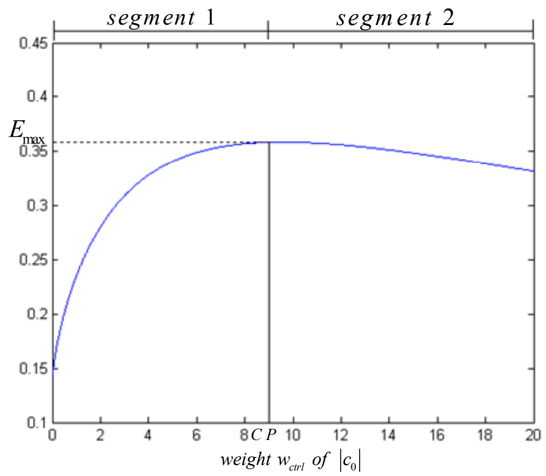
Let $X_N = \{|c_i| \mid i = 0,1,\cdots,N-1\}$ be a subset of DWT lowest-frequency coefficient. Then energy-proportion function is defined as

$$E(X_N) = -\sum_{k=0}^{N-1} z_k \log z_k$$



(a)

978-0-7695-4281-2/10 $26.00 © 2010 IEEE
DOI 10.1109/ICGEC.2010.173



(b)

Fig. 1. Characteristic curve of energy- proportion function (CCEP) for (a) $|c_0|=100$, $|c_1|=370$, $|c_2|=1810$, and (b) $|c_0|=122$, $|c_1|=130$, $|c_2|=1491$.

B. Some properties of energy-proportion function

To briefly describe the relationship in general case, X_N is rewritten as a vector form $X_N = [|c_0| \;\; |c_1| \;\; \cdots \;\; |c_{N-1}|]$ and $|c_0|$ is assumed to be the smallest value. Then,

$$\hat{X}_N = X_N W_N$$

(2)

where

$$W_N = diag(w_{ctrl}, 1, \cdots, 1)$$

(3)

is a weighting matrix and w_{ctrl} is the weight for controlling the smallest value $|c_0|$. Based on (2) and (3), we defined energy-proportion function as

$$E(\hat{X}_N) = -\left\{ \frac{w_{ctrl}|c_0|}{M} \log\left(\frac{w_{ctrl}|c_0|}{M}\right) + \frac{|c_1|}{M} \log\left(\frac{|c_1|}{M}\right) + \cdots\cdots + \frac{|c_{N-1}|}{M} \log\left(\frac{|c_{N-1}|}{M}\right) \right\}$$

(4)

where $M = w_{ctrl}|c_0| + \sum_{i=1}^{N-1} |c_i|$.

Based on (4), the relationship between the energy-proportion function and the smallest value $|c_0|$ of X_N is derived in the following theorem.

Theorem 1. *Suppose that* $X_N = [|c_0|, |c_1|, \cdots, |c_{N-1}|]$ *is given and* $|c_0|$ *is the smallest value, then* $E(\hat{X}_N)$ *has an absolute maximum at critical point*

$$w_{ctrl} = \left(|c_1|^{|c_1|} \times |c_2|^{|c_2|} \times \cdots \times |c_{N-1}|^{|c_{N-1}|}\right)^{\frac{1}{\sum_{i=1}^{N-1}|c_i|}} \Big/ |c_0|$$

(5)

proof: Since

$$\frac{dE(\hat{X}_N)}{dw_{ctrl}} = -\frac{1}{M^2}\left\{ |c_0||c_1|\log\left(\frac{w_{ctrl}|c_0|}{|c_1|}\right) + |c_0||c_2|\log\left(\frac{w_{ctrl}|c_0|}{|c_2|}\right) + \cdots + |c_0||c_{N-1}|\log\left(\frac{w_{ctrl}|c_0|}{|c_{N-1}|}\right) \right\}$$

(6)

Setting (6) to equal zero gives

$$\frac{-1}{M^2}\left\{ |c_0||c_1|\log\left(\frac{w_{ctrl}|c_0|}{|c_1|}\right) + |c_0||c_2|\log\left(\frac{w_{ctrl}|c_0|}{|c_2|}\right) + |c_0||c_3|\log\left(\frac{w_{ctrl}|c_0|}{|c_3|}\right) + \cdots + |c_0||c_{N-1}|\log\left(\frac{w_{ctrl}|c_0|}{|c_{N-1}|}\right) \right\} = 0$$

(7)

and then we have

$$\left\{ |c_0||c_1|\log\left(\frac{w_{ctrl}|c_0|}{|c_1|}\right) + |c_0||c_2|\log\left(\frac{w_{ctrl}|c_0|}{|c_2|}\right) + |c_0||c_3|\log\left(\frac{w_{ctrl}|c_0|}{|c_3|}\right) + \cdots\cdots + |c_0||c_{N-1}|\log\left(\frac{w_{ctrl}|c_0|}{|c_{N-1}|}\right) \right\} = 0$$

(8)

which implies

$$w_{ctrl} = \left(|c_1|^{|c_1|} \times |c_2|^{|c_2|} \times \cdots \times |c_{N-1}|^{|c_{N-1}|}\right)^{\frac{1}{\sum_{i=1}^{N-1}|c_i|}} \Big/ |c_0|$$

(9)

III. THE PROPOSED WATERMARKING SCHEME

In this section, the CCEP and the properties of energy-proportion function presented in the previous section are used in the design of embedding and extraction processes.

A. Watermarking embedding

In order to have high embedding capacity, this work considers the proposed scheme in case $N = 3$. The original audio $S(n)$ is first segmented into proper sections. Secondly, DWT is applied to each section. Thirdly, the synchronization code and watermark are arranged into a binary pseudo-noise (PN) sequence $B = \{0,1,0,1,1,\cdots\}$. The watermarks are adopted as the first secret key k_1. To achieve high robustness, they are embedded into the lowest-frequency coefficients. In this step, every three consecutive coefficients is grouped into $X_3 = [|c_0| \;\; |c_1| \;\; |c_2|]$ with $|c_0| < |c_1| < |c_2|$. Then the corresponding weighting matrix and $E(\hat{X}_3)$ are

$$W_3 = \begin{pmatrix} w_{ctrl} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(10)

and

$$E(\hat{X}_3) = E(X_3 W_3)$$

(11)

According to Theorem 1, $E(\hat{X}_3)$ has an absolute maximum at critical point (CP)

$$w_{ctrl}{}^{*}=\frac{(|c_1|^{|c_1|}\times|c_2|^{|c_2|})^{\frac{1}{|c_1|+|c_2|}}}{|c_0|}$$

(12)

Replace (12) into (5), $E(\hat{X}_3)$ has an absolute maximum

$$E_{\max}=-\left\{\frac{w_{ctrl}{}^{*}|c_0|}{w_{ctrl}{}^{*}|c_0|+|c_1|+|c_2|}\log(\frac{w_{ctrl}{}^{*}|c_0|}{w_{ctrl}{}^{*}|c_0|+|c_1|+|c_2|})+\frac{|c_1|}{w_{ctrl}{}^{*}|c_0|+|c_1|+|c_2|}\log(\frac{|c_1|}{w_{ctrl}{}^{*}|c_0|+|c_1|+|c_2|})+\frac{|c_2|}{w_{ctrl}{}^{*}|c_0|+|c_1|+|c_2|}\log(\frac{|c_2|}{w_{ctrl}{}^{*}|c_0|+|c_1|+|c_2|})\right\}$$

(13)

The corresponding minimum is

$$\begin{aligned}E_{\min}&=\lim_{w_{ctrl}\to 0}E(\hat{X}_3)\\&=-\lim_{w_{ctrl}\to 0}\left\{\frac{w_{ctrl}|c_0|}{w_{ctrl}|c_0|+|c_1|+|c_2|}\log(\frac{w_{ctrl}|c_0|}{w_{ctrl}|c_0|+|c_1|+|c_2|})+\frac{|c_1|}{w_{ctrl}|c_0|+|c_1|+|c_2|}\log(\frac{|c_1|}{w_{ctrl}|c_0|+|c_1|+|c_2|})+\frac{|c_2|}{w_{ctrl}|c_0|+|c_1|+|c_2|}\log(\frac{|c_2|}{w_{ctrl}|c_0|+|c_1|+|c_2|})\right\}\\&=-\left\{\frac{|c_1|}{|c_1|+|c_2|}\log(\frac{|c_1|}{|c_1|+|c_2|})+\frac{|c_2|}{|c_1|+|c_2|}\log(\frac{|c_2|}{|c_1|+|c_2|})\right\}\end{aligned}$$

(14)

which reduces to the case *N*=2. In this step, $|c_1|$ becomes the minimal coefficient and is scaled as follows.

$$E(\hat{X}_2)=E(X_2W_2)$$

(15)

where

$$W_2=\begin{pmatrix}w_{ctrl}&0\\0&1\end{pmatrix}=diag(w_{ctrl},1)$$

Besides, the least upper bound of range of $E(X_3)$ is derived as

$$\begin{aligned}E_{\sup}&=E(\hat{X}_3)|_{c_0=c_1=c_2}\\&=-\left\{\frac{|c_0|}{|c_0|+|c_0|+|c_0|}\log(\frac{|c_0|}{|c_0|+|c_0|+|c_0|})+\frac{|c_0|}{|c_0|+|c_0|+|c_0|}\log(\frac{|c_0|}{|c_0|+|c_0|+|c_0|})+\frac{|c_0|}{|c_0|+|c_0|+|c_0|}\log(\frac{|c_0|}{|c_0|+|c_0|+|c_0|})\right\}\\&=-\log\frac{1}{3}\end{aligned}$$

(16)

and the greatest lower bound $E_{\inf}$ reduces to the greatest lower bound of case N=2. i.e. ,

$$\begin{aligned}\lim_{w_{ctrl}\to 0}E(\hat{X}_2)&=-\lim_{w_{ctrl}\to 0}\left\{\frac{w_{ctrl}|c_0|}{w_{ctrl}|c_0|+|c_1|}\log\frac{w_{ctrl}|c_0|}{w_{ctrl}|c_0|+|c_1|}+\frac{|c_1|}{w_{ctrl}|c_0|+|c_1|}\log\frac{|c_1|}{w_{ctrl}|c_0|+|c_1|}\right\}\\&=\lim_{w_{ctrl}\to 0}\left\{\frac{w_{ctrl}|c_0|}{w_{ctrl}|c_0|+|c_1|}\log\frac{w_{ctrl}|c_0|}{w_{ctrl}|c_0|+|c_1|}\right\}+\lim_{w_{ctrl}\to 0}\left\{\frac{|c_1|}{w_{ctrl}|c_0|+|c_1|}\log\frac{|c_1|}{w_{ctrl}|c_0|+|c_1|}\right\}\\&=\lim_{x\to 0^+}\left\{\frac{x}{x+|c_1|}\log\frac{x}{x+|c_1|}\right\}+1\log 1\ (Let\ x=w_{ctrl}|c_0|)\\&=\lim_{y\to 0^+}\left\{\frac{\log y}{\frac{1}{y}}\right\}+0\ (Let\ y=\frac{x}{x+|c_1|})\\&=0\ (\equiv E_{\inf})\qquad (by\ L'Hospital)\end{aligned}$$

(17)

Based on the previous discussions, we use following rules to embed the sequence B.

- If the bit "$1\in B$" is embedded into $X_3=[|c_0|\ \ |c_1|\ \ |c_2|]$, then $E(X_3)$ is modified to

$$E_{\text{mid}}+\varepsilon\le E(\hat{X}_3)\le E_{\text{mid}}+2\varepsilon$$

(18)

- If the bit "$0\in B$" is embedded into $X_3=[|c_0|\ \ |c_1|\ \ |c_2|]$, then $E(X_3)$ is modified to

$$E_{\text{mid}}-2\varepsilon\le E(\hat{X}_3)\le E_{\text{mid}}-\varepsilon$$

(19)

where $E_{\text{mid}}=(E_{\sup}+E_{\inf})/2$ and ε is a small positive number which can be adopted as the second secret Key k_2.

B. Watermarks extraction

Without sorting, every three consecutive coefficients is grouped into $X_3=[|c_0|\ \ |c_1|\ \ |c_2|]$ from DWT lowest-frequency. The watermarks are extracted with the key k_1 and key k_2 by the following rules.

- If $E_{\text{mid}}+\varepsilon\le E(\hat{X}_3)\le E_{\text{mid}}+2\varepsilon$, the extracted bit is "1".
- If $E_{\text{mid}}-2\varepsilon\le E(\hat{X}_3)\le E_{\text{mid}}-\varepsilon$, the extracted bit is "0".

IV. EXPERIMENTAL RESULTS

This work embeds the watermarks into the lowest-frequency sub-band by applying seven-level DWT. The parameter ε is set to 0.05. The quality of the watermarked audio is measured by signal-to-noise ratio (SNR) which is defined as

$$\mathrm{SNR} = 10\log_{10}\left\{\sum_n (S(n))^2 \Big/ \sum_n (\bar{S}(n) - S(n))^2\right\}$$

where $S(n)$ and $\hat{S}(n)$ denote the original and the watermarked audio signal. The SNRs are 20.7dB and 17.2dB for the two audios, popular and symphony, which are 16-bit mono-type with sampling rate 44.1 kHz and length 11.609seconds. Besides, some attacks are applied to test the robustness of hidden data. Bits-error-rate (BER) defined as

$$\mathrm{BER} = (\text{Number of error bits} / \text{Total bits}) \times 100\%$$

is used to measure the robustness. Under the resample with 22050 Hz and 11025 Hz, the BERs are 3.1 and 4.5 for popular and 4.2 and 4.8 for symphony. The BERs after adopting a low-pass filter with the cutoff frequency 3 kHz are 25.1% and 28.2%. The BERs after amplitude scaling with scaling factor 0.8 and 1.2 are all 0.4 for popular and symphony.

V. CONCLUSION

The characteristic curve of energy-proportion function (CCEP) is first illustrated and proven in this work. Then the CCEP and the properties are applied to propose an energy-proportion based scheme. In addition, the proposed scheme can extract the watermark without original audio signal. The experimental results show that the proposed scheme has strong robustness.

REFERENCES

[1] W. N. Lie and L. C. Chang, "Robust and High-Quality Time-Domain Audio Watermarking Based on Low-Frequency Amplitude Modification" *IEEE. Transactions on Multimedia,* Vol.8, No.1, pp. 46-59, February 2006.

[2] S. Wu, J. Huang, D. Huang, and Y. Q. Shi, "Efficiently self- synchronized audio watermarking for assure audio data transmission." *IEEE. Transactions on Broadcasting,* Vol.51, No.1, pp. 69-76, March 2005.

[3] H. O. Kim, B. K. Lee, and N. Y. Lee, "Wavelet-based audio watermarking techniques: Robustness and fast synchronization," September 2007, http://citeseer.ist.psu.edu/644572.html.

[4] X. Huang, Y. Abe, and I. Echizen, "Capacity Adaptive Synchronized Acoustic Steganography Scheme", Journal of Information Hiding and Multimedia Signal Processing, Vol. 1, No. 2, pp. 72-90, April 2010.

[5] H. Al-Qaheri, A. Mustafi, and S. Banerjee, "Digital Watermarking using Ant Colony Optimization in Fractional Fourier Domain," Journal of Information Hiding and Multimedia Signal Processing, Vol. 1, No. 3, pp. 179-189, July 2010.

[6] K. Yamamoto and M. Iwakiri, "Real-Time Audio Watermarking Based on Characteristics of PCM in Digital Instrument", Journal of Information Hiding and Multimedia Signal Processing, Vol. 1, No. 2, pp. 59-71, April 2010.

[7] R. M. Gray, Entropy and information theory. New York: Springer-Verlag, c1990.