Wavelet-Based Entropy for Digital Audio Watermarking

Shuo-Tsung Chen
Asian Institute of TeleSurgery,
Chang Bing Show Chwan Memorial Hospital, Lukang, Changhua County 505, Taiwan (ROC).
shough33@yahoo.com.tw

Huang-Nan Huang
Department of Mathematics,
Tunghai University,
Taichung 407, Taiwan (ROC).
nhuang@thu.edu.tw
(corresponding author)

Chih-Yu Hsu
Department of Information and Communication Engineering,
Chaoyang University of Technology,
Taichung County 413, Taiwan (ROC).
tccnchsu@gmail.com

Kuo-Kun Tseng
Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen Graduate School, China.
kuokun.tseng@googlemail.com

Chun-Hua Wu
Chang Bing Show Chwan Memorial Hospital, Lukang, Changhua County 505, Taiwan (ROC).
anis6699@yahoo.com.tw

Jeng-Shyang Pan
Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen Graduate School, China.
jengshyangpan@gmail.com

***Abstract*—Unlike traditional entropy in information theory, this work uses the normalized energy instead of probability to obtain a low-frequency amplitude transform (LAT) on coefficients of discrete wavelet transform (DWT). The watermark is embedded based on the properties and characteristics of this transform. Finally, performance of the proposed scheme is assessed by signal-to-watermark (SWR) and bit error rate (BER). Experimental results demonstrate that the embedded data are robust against most signal processing and attacks, such as re-sampling, low-pass filtering, and amplitude-scaling.**

Keywords-entropy; low-frequency amplitude transform; discrete wavelet transform

I. INTRODUCTION

In recent years, many watermarking techniques have been proposed [1-8]. For music copyright protection, audio watermarking has the following requirements: 1) The watermark should be imperceptible in embedded audio. 2) The embedding technique should offer more than 20 dB signal to watermark ratio (SWR). 3) The watermark should prevent common attacks, including filtering, re-sampling, and mp3 compression, etc.

Wu *et al.* [2] used quantization index modulation to embed information into the low-frequency sub-band coefficients of discrete wavelet transform (DWT). This technique has good watermarked audio quality and strong robustness against common signal processing and noise corruption. However, this method is vulnerable to amplitude and time scaling. Xiang *et al.* [3] proposed a DWT-based audio watermarking algorithm robust against the DA/AD conversions. The relative energy relation among different groups of the DWT coefficients in the low-frequency sub-band are utilized in embedding by adaptively controlling the embedding strength. However, the method has low capacity and SNR. Chen *et al.* [4] proposed an optimization-based watermarking scheme robustly against many attacks.

Unlike traditional entropy, this work uses normalized energy instead of probability to form a novel entropy. Based on this concept, this work presents a new technique that embeds information by using low-frequency amplitude transform (LAT). Some properties and characteristic curve of LAT are analyzed and proved to investigate the relationship between LAT and DWT coefficients. Finally, the performance of the proposed scheme is assessed by signal-to-watermark ratio (SWR) and bit error rate (BER). Experimental results demonstrate that the embedded data are robust against most signal processing and attacks.

The remainder of this paper is organized as follows. Section II introduces DWT and LAT. Section III derives the properties and characteristic curve of LAT to analyze the relationship between LAT and DWT coefficients. The proposed embedding and extraction processes are described in Section IV. Experiments are conducted to test the performance of our proposed method in Section V. Finally, conclusions are summarized in Section VI.

II. DWT AND LAT

Discrete wavelet transform (DWT) is first reviewed in this section. Based on the low-frequency DWT coefficients in level seven, which is also referred as the lowest-frequency DWT coefficients, traditional entropy is redefined as a novel low-frequency amplitude transform (LAT).

A. Discrete-time wavelet transform (DWT)

Since the conventional fast Fourier transform (FFT) efficiently decomposes a signal into uniform-resolution analysis, it is suitable to analyze the wide-sense-stationary condition but not in non-stationary signal. In this paper, the discrete wavelet transform (DWT) is adapted to decompose

978-0-7695-4517-2/11 $26.00 © 2011 IEEE
DOI 10.1109/IIHMSP.2011.39

the signal into the time-frequency domain. According to the multi-resolution property of DWT, it leads to low-frequency but high-temporal resolution in high frequency bands and low-temporal but high-frequency resolution in low frequency bands. Therefore, we let the low frequency bands enhance the periodic property by only decomposing low-frequency band in each level. In [9], a method to implement DWT by using filter bank decomposition is proposed.

B. LAT

Before to introduce the proposed watermarking technique, the LAT must be defined and discussed. If there are N non-negative random samples that are shown as $\tilde{X}_N = \{c_i \mid 0 \le i \le N-1\}$, the corresponding probabilities are $P(c_0) = p_0$, $P(c_1) = p_1$, …, $P(c_{N-1}) = p_{N-1}$. Based on information theory, the entropy of these samples is defined as follows:

$$H_r(\tilde{X}_N) = -\sum_{i=0}^{N-1} p_i \log_r p_i,\ 0 \le i \le N\text{-}1 \qquad (1)$$

where r is a base number of the logarithm function $\log$. This work adopts $r = 10$. Unlike the traditional way, this work uses the normalized energy instead of probability in wavelet domain as follows.

Definition 1. *Suppose that* $X_N = \{|c_i| \mid 0 \le i \le N-1\}$ *is a set of low-frequency coefficients in DWT, low-frequency amplitude transform (LAT) of* X_N *is then defined as*

$$LAT(X_N) = -\sum_{k=0}^{N-1}\left(|c_k| \Big/ \sum_{j=0}^{N-1}|c_j| \right) \log \left(|c_k| \Big/ \sum_{j=0}^{N-1}|c_j| \right) \qquad (2)$$

where $|c_k| \Big/ \sum_{j=0}^{N-1}|c_j|$ *is the normalized energy of coefficient* c_k.

From this definition, if the variation of low-frequency coefficients X_N is small, the corresponding $LAT(X_N)$ is big. For an example with $N = 20$ as shown in Figure 1, when X_{20} varies slowly, $LAT(X_{20})$ is 1.2998. However, when the variation of low-frequency coefficients X_N is large, the corresponding $LAT(X_N)$ is small. Figure 2 depicts that when X_{20} ($N = 20$) varies markedly, the corresponding $LAT(X_{20})$ is 1.1589. In this paper, we adopt $N = 2$ to have high embedding payload. Moreover, the standard deviations of this function for $N = 2$ before and after various attacks are approximately invariant. It is expected that the proposed low-frequency amplitude transform is robust against common attacks.

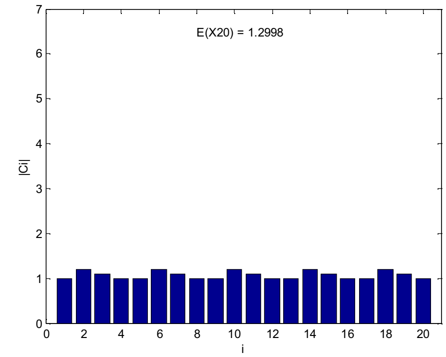

Fig. 1 When X_{20} varies slowly, $LAT(X_{20})$ is 1.2998.

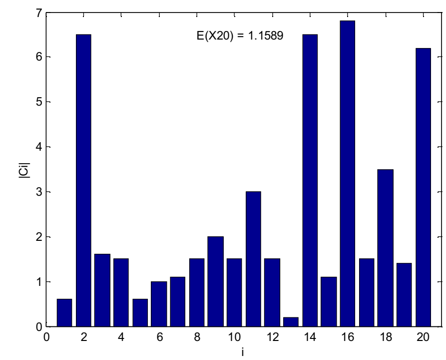

Fig. 2 When X_{20} varies markedly, $LAT(X_{20})$ is 1.1589.

Finally, the performance of our method is measured by signal-to-watermark ratio (SWR) and Bit-error-rate (BER) mathematically. They are defined as follows.

$$\text{SWR}=10\log_{10}\left\{ \frac{\sum_n (S(n))^2}{\sum_n (\overline{S}(n) - S(n))^2} \right\}, \qquad (3)$$

where $S(n)$ and $\overline{S}(n)$ denote the original and the modified audio, respectively;

$$\text{BER}=\left(B_{error}/B_{total}\right)\times 100\%, \qquad (4)$$

where B_{error} and B_{total} denote the number of error bits and the number of total bits, respectively.

III. Properties of the Characteristic Curve of Lat

This section discusses some properties of the characteristic curve of LAT (CCL). Based on these properties and the characteristic curve, this work presents a novel watermarking scheme.

A. Properties of CCL

In the proposed watermarking scheme, the host digital audio signal $S(n)$ is cut into segments. Then, every two low-frequency DWT coefficients in each segment are grouped and sorted according to their absolute value into a vector form $X_2 = [|c_0|, |c_1|]$, where $|c_0| < |c_1|$. Since the value of

$LAT(X_2)$ in (2) is a function of X_2, a weighting matrix W is used to control the variation of $LAT(X_2)$ as follows:

$$\hat{X}_2 = X_2 W \tag{5}$$

where

$$W = \begin{pmatrix} w_0 & 0 \\ 0 & 1 \end{pmatrix} = diag(w_0, 1) \tag{6}$$

In other words, only the smallest value $|c_0|$ will be modified. Hence, the corresponding $LAT(\hat{X}_2)$ is shown as follows.

$$LAT(\hat{X}_2) = -\left\{ \frac{w_0|c_0|}{w_0|c_0|+|c_1|}\log\frac{w_0|c_0|}{w_0|c_0|+|c_1|} + \frac{|c_1|}{w_0|c_0|+|c_1|}\log\frac{|c_1|}{w_0|c_0|+|c_1|} \right\}$$

with the following property:

***Lemma 1**. $LAT(\hat{X}_2)$ has an unique critical point (CP) $w_0 = |c_1|/|c_0|$.*

B. The Characteristic curve of CCL

Based on the previous discussion, the relation between $LAT(\hat{X}_2)$ and w_0 can be described as a CCL. For example, $|c_0| = 100$, $|c_1| = 370$. Their relation is shown in Fig. 3. Based on this CCL, $LAT(\hat{X}_2)$ has a CP at $w_0 = 3.7$ according to Lemma 1. In other words, the maximum of $LAT(\hat{X}_2)$ should occur at $w_0 = |c_1|/|c_0|$ with it value given by

$$LAT(\hat{X}_2)|_{w_0=|c_1|/|c_0|} = -\left\{ \frac{w_0|c_0|}{w_0|c_0|+|c_1|}\log\frac{w_0|c_0|}{w_0|c_0|+|c_1|} + \frac{|c_1|}{w_0|c_0|+|c_1|}\log\frac{|c_1|}{w_0|c_0|+|c_1|} \right\}\Bigg|_{w_0=|c_1|/|c_0|}$$

$$= -\log\frac{1}{2} \equiv LAT_{\max} \tag{7}$$

Since the minimum of $LAT(\hat{X}_2)$ should be attained when $w_0 \to 0$, i.e.,

$$LAT(\hat{X}_2)|_{w_0\to 0} = -\left\{ \frac{w_0|c_0|}{w_0|c_0|+|c_1|}\log\frac{w_0|c_0|}{w_0|c_0|+|c_1|} + \frac{|c_1|}{w_0|c_0|+|c_1|}\log\frac{|c_1|}{w_0|c_0|+|c_1|} \right\}\Bigg|_{w_0\to 0}$$

$$\equiv LAT_{\min} \to 0 \tag{8}$$

And we set $LAT_{\min} = 0.05$ to sufficiently approximate the smallest value of LAT for computational purpose. During the watermarking process, we also set LAT_{mid} to be

$$LAT_{\text{mid}} = \left(LAT_{\max} + LAT_{\min}\right)/2 \tag{9}$$

By Lemma 1, $LAT(\hat{X}_2)$ has two monotone subintervals which are called segment 1 and 2, referred to Fig. 3 as a typical example. In this work, we adopt the range $w_0 \in \left[0.05, |c_1|/|c_0|\right]$ of $LAT(\hat{X}_2)$ in segment 1 to embed data. The detail process will be introduced in the next section.

IV. The Proposed Watermarking Technique

In this section, the novel watermarking technique by using segment 1 in the characteristic curve of CCL is proposed. It contains embedding and extraction processes. These processed are introduced as follows.

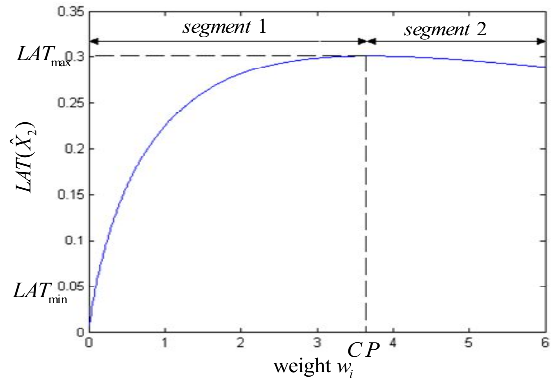

Fig. 3 The characteristic curve of LAT for $|c_0| = 100$, $|c_1| = 370$.

A. The embedding process

First of all, the synchronization codes and watermark are arranged into a binary pseudo-noise (PN) sequence B, for example, $B = \{0,1,10,1,...\}$. Secondly, as shown in Figure 4, the original audio $S(n)$ is split into proper segments, and DWT is applied to each segment. Then the synchronization codes and watermark are embedded into the lowest-frequency DWT coefficients. In this step, we group every two consecutive coefficients into $X_2 = \{|c_0|, |c_1|\}$ with $|c_0| < |c_1|$. The proposed embedding process is described as follows.

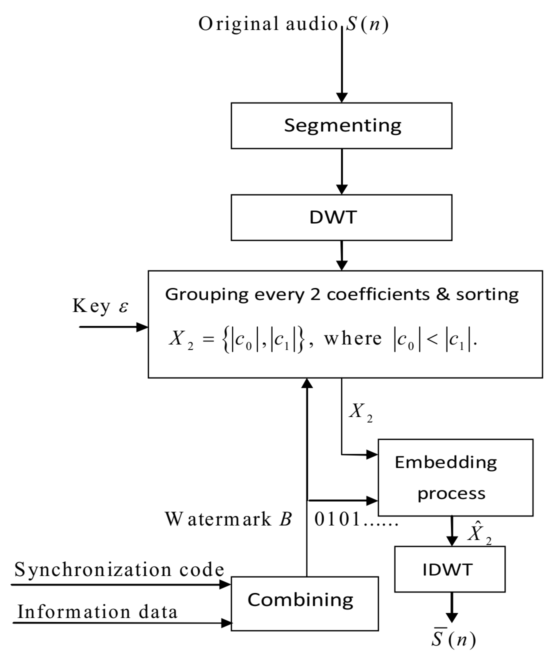

Fig. 4 The flowchart of watermark embedding process.

- If the binary bit "$1 \in \mathrm{B}$" is embedded, we choose w_0 such that

$$LAT(X_2W) = LAT(\hat{X}_2) = (LAT_{\max} + LAT_{\mathrm{mid}})/2 + \varepsilon \quad (10)$$

- If the binary bit "$0 \in \mathrm{B}$" is embedded, we choose w_0 such that

$$LAT(X_2W) = LAT(\hat{X}_2) = (LAT_{\min} + LAT_{\mathrm{mid}})/2 - \varepsilon \quad (11)$$

where $\varepsilon \in [0, 0.3]$ is a small positive number which can be used as a secret key.

B. *The extraction process*

The flowchart of watermark extraction is given in Figure 5. Every two consecutive lowest-frequency DWT coefficients is grouped into $X_2 = \{|c_0|, |c_1|\}$. To extract the watermark $\hat{B} = \{\hat{\beta}\}$, we apply (2) with $N = 2$ as follows.

- If $LAT(X_2) > LAT_{\mathrm{mid}}$, the extracted value $\hat{\beta} = 1$.
- If $LAT(X_2) < LAT_{\mathrm{mid}}$, the extracted value $\hat{\beta} = 0$.

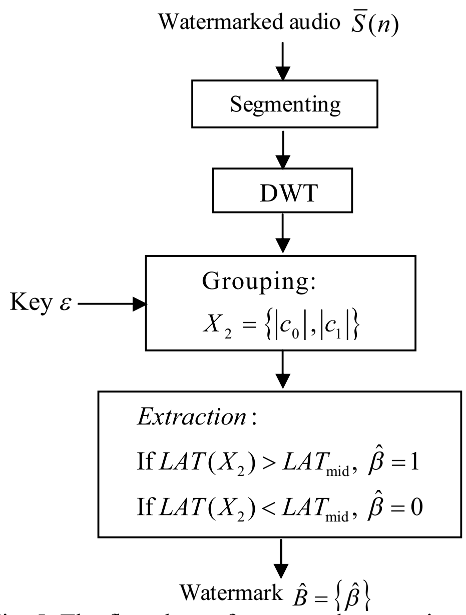

Fig. 5 The flowchart of watermark extraction.

V. EXPERIMENTAL RESULTS

The performance of the proposed audio watermarking technique is tested by using 16-bit mono audio signal sampled at 44.1 kHz. The length of each audio is about 11.6 seconds. We use two kinds of music which are symphony and popular. By setting the parameter ε to be 0.05, the synchronization code and watermark are embedded into the low-frequency DWT coefficients in level seven. Accordingly, the embedding capacity is 2000bits/11.6 secs. The SWR for the two audios are 20.8 dB (symphony) and 21.1 dB (popular). Moreover, we apply three types of attack to test the robustness: (1) re-sampling, (2) amplitude scaling, (3) low-pass filtering. The testing results are listed in TABLES I-III.

TABLE I. BER(%) after Re-Sampling

Rate (Hz)	22050	11025	8000
symphony	3.6	7.5	7.3
popular	9.2	12.9	14.7

TABLE II. BER(%) after Amplitude Scaling

Scaling factor	0.2	0.8	1.1	1.2
symphony	0.5	0.4	0.4	0.4
popular	0.5	0.5	0.4	0.4

TABLE III. BER(%) after Low-Pass Filter

Cut-off frequency(kHz)	3
symphony	24.1
popular	26.8

VI. CONCLUSIONS

A novel audio watermarking technique is proposed to embed the information by using LAT. When embedding the watermark, an analytical formula is provided to determine the weight on DWT coefficients. The experimental results show that the embedded data are robust against some attacks.

ACKNNOWLEDGEMENTS

The work was in part supported by the NSC, TAIWAN, under the grant: NSC 98-2115-M-029-006-MY2.

REFERENCES

[1] C.-Yu Yang, C.-H. Lin, and W.-C. Hu, "Embedding Limitations with Digital-audio Watermarking Method Based on Cochlear Delay Characteristics," *Journal of Information Hiding and Multimedia Signal Processing,* Vol. 2, No. 1, January 2011, pp. 1-23.

[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, March 2005, pp. 69-76.

[3] S. Xiang, and J. Huang, "Robust Audio Watermarking Against the D/A and A/D Conversions," CoRR abs/0707.0397, 2007.

[4] S.-T. Chen, G.-D. Wu, and H.-N. Huang, "Wavelet-Domain Audio Watermarking Scheme using Optimization-Based Quantization," *IET Proceedings on Signal Processing*, vol. 4, no. 6, 2010, pp. 720-727.

[5] H. G. Kaganami, S. K. Ali, and B. Zou, "Reversible Data Hiding By Adaptive IWT-coefficient Adjustment," *Journal of Information Hiding and Multimedia Signal Processing,* Vol. 2, No. 1, January 2011, pp. 24-32.

[6] R. M. Noriega, M. Nakano, B. Kurkoski, and K. Yamaguchi , "High Payload Audio Watermarking: toward Channel Characterization of MP3 Compression," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 2, April 2011, pp. 91-107.

[7] R. Martinez-Noriega, M. Nakano, B. Kurkoski, and K. Yamaguchi, "High Payload Audio Watermarking: Toward Channel Characterization of MP3 Compression," *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 2, No. 2, Apr. 2011, pp. 91-107.

[8] H. C. Huang and Y. H. Chen, "Genetic Fingerprinting for Copyright Protection of Multicast Media," *Soft Computing*, Vol. 13, No. 4, Feb. 2009, pp. 383-391.

[9] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Respresentation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, July 1989, pp.674-693.