An Audio Mask Algorithm Based on Cepstrum Transform and Wavelet Transform

Jian-yong LI¹, Di-long WU², Yan-qiu YANG³, Rui Li⁴ and Yu LEI⁴

¹Chongqing Lilong Technology Industry Group Co. Ltd, Chongqing China, 401123
²Laser fusion research center of China Academy of Engineering Physics, 401147
³Chongqing Public Security Bureau, Chongqing, 401147
⁴Communication Sergeant School of the Army Engineering University, Chongqing, 400035
*Corresponding author

Keywords: Audio mask, Cepstrum coefficient, Robustness, Discrete wavelet transform, The synchronization attack.

Abstract. In order to better improve audio watermark robustness and perceptual, the paper puts forward a kind of based on cepstrum transform and the wavelet transform of digital audio watermarking algorithm. The algorithm will be audio signal wavelet decomposition, selection of low frequency coefficient of framing and cepstrum transform. And according to the statistical mean thought, using the cepstrum domain characteristics of the energy concentration near zero, a pseudo random sequence watermark embedded in cepstrum domain. Experimental results show that the algorithm not only has good perceptual, but also has better robustness and can withstand attacks of common signal such as low-pass filter resampling quantitative lossy compression and sync to attack.

Introduction

Audio information hiding technology as an important branch of multimedia information hiding, more and more get people's attention and use of, and put forward many representative [1,2,3]. Literature [4] algorithm of wavelet coefficient of the secondary cepstrum literature [5] this paper presents a mathematical transformation based on wavelet decomposition and complex cepstrum transform audio digital watermark algorithm. In this chapter, combining the wavelet and cepstrum transform, this paper proposes a hybrid domain audio watermarking algorithm, and has strong robustness.

Watermark Embedding Algorithm

If carrier audio is \( A = \{a(c)|1 \leq c \leq \text{Length}\} \), in \( \text{Length} \) is signal sampling points of carrier audio, \( a(c) \) is the \( c \) sample point amplitude values. Embedded steps are as follows:

Step1: Discrete wavelet transform. Firstly discrete wavelet transform (DWT) to \( A \) using “db4”:
\[
[X, H] = DWT(A), \quad \text{then } X
\]
is the low frequency component, for \( X = (x_1, x_2, \cdots, x_{L-1}, x_L) \). Using a binary \( m \) sequence. Using “seed” To generate a binary pseudo random sequence as a watermark, for \( W = \{m(i)|i=1, 2, \cdots, M\} \), then \( m(i) \in \{0,1\} \) is the sequence of the \( i \)th a value.

Step2: In carrier audio after wavelet transform of the low frequency part of the former \( L_1 \) length, using literature [6] of the spread spectrum synchronization scheme embedded \( L_1 \) length of \( m \) sequence as synchronization code.

Step3: Select the appropriate frame length, then frame to the rest of the low frequency component \( Ca \), each frame length is \( M \), get \( Ca_1, Ca_2, \cdots, Ca_{\lfloor (L-L_1)/M \rfloor} \);
\[
Ca = \left(\frac{ca(1), ca(2), \ldots, ca(M), ca(M + 1), \ldots, ca(2 \times M), \ldots, \text{ca}\left(\left\lfloor \frac{(L - L_t)}{M} \right\rfloor - 1 \times M + 1, \ldots, ca\left(\left\lfloor \frac{(L - L_t)}{M} \right\rfloor \times M \right)\right)}{ca_{1}, ca_{2}}\right)
\]

them \(\left\lfloor \frac{(L - L_t)}{M} \right\rfloor\) is the total number of framing, \(len = \left\lfloor \frac{(L - L_t)}{M} \right\rfloor\).

Step4: Coefficient of cepstrum transform and choice. Cepstrum transform to each frame data, choose the middle part smooth length is N to embed the watermark, the collection of cepstrum coefficient:
\[
C = \{c_k(j) | k = 1, 2, \ldots, len, j = 1, 2, \ldots, N\}
\]

(2)

Step5: Watermark embedding. Watermark embedding method is as follows: for each frame cepstrum coefficients \(c_k(j), (1 \leq j \leq N)\), according to the formula (2) to embed the watermark. Getting:
\[
c'_k(j) = c_k(j) + (\beta w(k) - \lambda c_k(j))u(j)
\]

(3)

Step6: Inverse cepstrum transform. To embed watermark information embedded cepstrum coefficient and the coefficient of consolidation, and inverse cepstrum transform low frequency wavelet coefficient.

Step7: The discrete wavelet transform. Embedded synchronization code and the watermark information of the low frequency coefficient and high frequency coefficients of wavelet reconstruction of the original, you can get audio watermark carrier.

Watermark Extraction Algorithm

Watermark extraction algorithm process is shown in figure 1. In the watermark extraction, want to know as the seeds of a key “seed”, without the original audio information carrier, it is blind extraction process.

Specific steps are as follows:

Step1: Discrete wavelet transform. To contain the watermark audio category five wavelet decomposition.

Step2: Using pseudo random \(m\) sequence, using the algorithm principle of literature [2], using the correlation detection method, find out the starting position of the embedded watermark information.

Step3: According to the coefficient of each frame M from its starting position to the low frequency coefficient for framing and cepstrum transform, framing for total \(len\). Choose between smooth has hidden watermark cepstrum coefficient, the collection is \(C' = \{c'_k(j) | k = 1, 2, \ldots, len, j = 1, 2, \ldots, N\}\), to \(c'_k(j), (1 \leq j \leq N)\) and \(u(j)(1 \leq j \leq N)\) relevant calculation, According to the polarity of the relevant value judgment to extract the watermark sequence.
Simulation and Experimental Results

In the experiments, we use a length of 12 seconds, sample rate 44.1 kHz, 16 quantitative mono pop music as a carrier signal, and embed watermarking information in the carrier signal. Validation process, frame selection long $M = 200$, hidden watermark for smooth cepstrum coefficient $N = 100$, the watermark information hiding capacity of 7bps.

**Best Control Parameters $\beta, \lambda$ Selection.** Analysis by Matlab simulation experiment, when $\beta = 0.05, \lambda = 0.5$ in (4), signal-to-noise ratio and achieve a better balance, robustness under various conventional signal processing attacks, the error rate is under 5%, considered in this algorithm the optimum parameters of the control factors.

$$\begin{align*}
\text{Pe} &= \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\sigma_r} \exp\left(-\frac{(x - \mu_r)^2}{2\sigma^2_r}\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\beta\sigma_u \sqrt{N}}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt
\end{align*}$$

(4)

Concealment Test

In concealment test, then a long length of 48s, the sampling rate is 8KHZ, a speed of 128KBPS, 16 quantization of speech signal. Fig2 is the original audio and contain audio watermark figure the time domain waveform.

Robustness Testing

In order to evaluate the robustness of the watermarking algorithm, the experiment used to add white Gaussian noise of the mean $u = 0$, variance $\sigma = 0.1$, audio files with 22.05kHz and 16kHz resampling, again with 44.1kHz and 8kHz sampling again back to the original audio. MP3 compression attacks test will be a speed of 705kbps and 128kbps audio files compressed into 80kbps respectively, and to the synchronous attack and attack DA/AD conversion, a series of experiments:

(1) Sync to attack. In contains audio watermark randomly selected from a number of sampling points, and then cut off. Cut out the sample number is 200.

(2) DA/AD conversion. Use of Winamp music player to play the original audio, the use of Windows's own recorder through the line output way.
The proposed algorithm in this chapter in resistance to low pass filtering, adding noise, resampling, quantitative, MP3 compression, has good effect to synchronization from table 1. Compared with the literature [5, 6], the robustness had significantly improve.

Table 1. The robust watermark detection results under attack and algorithm (random cutting 200 sample points).

<table>
<thead>
<tr>
<th>Attack types</th>
<th>Ber（%）</th>
</tr>
</thead>
<tbody>
<tr>
<td>low pass filter</td>
<td>1.56(5kHz)</td>
</tr>
<tr>
<td>white Gaussian noise</td>
<td>0</td>
</tr>
<tr>
<td>resampling</td>
<td>0.1</td>
</tr>
<tr>
<td>Re-quantitative</td>
<td>0</td>
</tr>
<tr>
<td>MP3 compression</td>
<td>4.10 (64kbps)</td>
</tr>
<tr>
<td>Sample cutting</td>
<td>12.532</td>
</tr>
<tr>
<td>DA/AD</td>
<td>—</td>
</tr>
</tbody>
</table>

By the experimental results it can be seen that the algorithm is able to withstand attack most Stirmark, strong robustness of the proposed algorithm.

**Conclusion**

This chapter is given based on wavelet and cepstrum domain digital audio watermarking algorithm, with the human ear auditory model and frequency mask, we effect, select the low frequency is not sensitive to the human ear hearing part and cepstrum transform, using the cepstrum domain coefficient has the characteristics of stability, in the middle of a pseudo random sequence using the proposed algorithm is embedded in the cepstrum coefficient of relatively stable, thus realizes the watermark embedding. Using the correlation detection method, to extract the watermark according to the relative value of polarity. Experiments also show that the algorithm not only for a variety of music signal, also for speech signal also has very strong robustness and good perceptual.
References


