NDFT-based Audio Watermarking Scheme with High Security

Ling Xie, Jia-shu Zhang*, Hong-jie He
Sichuan Key Lab of Signal and Information Processing Southwest Jiaotong University
Chengdu, 610031, China
jszhang@home.swjtu.edu.cn*

Abstract

Uniform discrete Fourier transform (DFT) has the drawback of public frequency points. On the other hand, Nonuniform discrete Fourier transform (NDFT) could set up random sampling points in frequency domain as desired rather than fixed frequency points in DFT. That is DFT is the special condition of NDFT. This paper utilizes NDFT to instead uniform DFT in the algorithm of audio watermarking scheme with the purpose of providing the probability of hidden embedding positions. Moreover, to further improve the systematic security, in our method, we use coupled-chaotic sequence to randomly select the NDFT-domain frequency points used for embedding watermark, which overcomes the flaws brought by single chaotic map which is finite word length effect and vulnerable to repeated group attack. Good experimental results have shown the proposed scheme possesses higher systematic security than DFT-based watermarking scheme.

1. Introduction

Due to fast growth of Internet applications, digitized data becomes more and more popular. Because of the ease of digital duplication and tampering, the intellectual property of multimedia becomes an important issue nowadays. Especially the large-scale proliferation of audio products over Internet represented by mp3 creates a demand for effective and reliable security of its copyright. Therefore, robust audio watermarking technique has shown itself as an emerging cutting edge technology.

Recently, robust audio watermarking algorithms in transform domain have attracted wide attention because of its higher robustness against attacks than that in temporal domain[1-5]. Discrete Fourier transform (DFT) belongs to transform domain. And many algorithms have been developed in DFT domain. One of the typical examples is quantizing DFT parameters (amplitude or phase) to embed watermark proposed by Wang and Sun [5], in which the high robustness against many common attacks, such as filtering, re-sampling, and compression. However, DFT transform has the vital disadvantage of public frequency points. That is, if the number of points implemented by DFT is certain, the corresponding frequency points are public. Accordingly, algorithms which embeds watermark on DFT parameters, are vulnerable to the known embedding position attack. The attacker could do anything with the embedded watermark, which leads to lower systematic security. On the other hand, a well designed watermarking system should fulfill the ‘Kerchoff’ principle of cryptography. According to his principle, the security of the system has to be based on the assumption that the adversary or attacker has full knowledge of the design and development details of the system. The only missing information for the attacker is the secret key, and without the secret key he has no way of extracting the hidden data from the carrier [6]. Obviously, DFT-based watermarking algorithm has no secret key. That is to say, the DFT-based watermarking schemes are susceptible and not secure.

To overcome the flaw of no secret key in DFT transform, in this paper, a NDFT-based audio watermarking scheme with high security is put forward. Nonuniform discrete Fourier transform (NDFT) could set up random sampling points in frequency domain. When NDFT chooses the uniform sampling points on the unit plane, NDFT has converted to DFT. That is DFT is the special condition of NDFT. In this paper, we utilize NDFT to instead uniform DFT transform to design secure audio watermarking system, where the random sampling points in NDFT domain are selected based on secret key. To further enhance the security, in the process of generating secret key, we use coupled-chaotic sequence to conquer the finite word length effect and resist repeated group attack. Experimental results have shown the superior security of the proposed scheme.

2. Proposed system model
As shown in Fig.1, first step of our proposed system is to divide original audio signal into non-overlap segments.

We utilize digital image as watermark and embed it into audio signal. Digital image is two dimension, it could be embedded into one-dimensional audio signal unless reducing its’ dimension to one dimension.

After reducing dimension, the watermark must be encrypted to protect from the so-called copy attack.

In our method, we use chaos to choose random sampling points in NDFT domain. The most attractive features of chaos in information hiding are its sensitivity to initial conditions and the outspreading of orbits over the entire space. In recent years, chaotic maps have been widely used for digital watermarking techniques to increase the security [3,4]. Here, we use coupled-chaotic sequence generated by secret keys to randomly select NDFT-domain frequency sampling points. The reason of using coupled-chaotic sequence is due to the single chaotic map exiting the following flaws: (1) It’s not a completely non-period sequence due to the finite word length effect; (2) It could be predicted because of the certain chaotic dynamics model; (3) It is susceptible to repeated group attack. Coupled-chaotic sequence is non-period, unpredictable, superior random, and resistant to brute-force attack because of large secret key space.

For each audio segment, one-bit watermark is embedded onto a NDFT-domain frequency point selected by coupled-chaos sequence. Then, implement inverse NDFT (INDFT) with the embedded NDFT-domain frequency points to form watermarked audio segment. At last, the whole watermarked audio signal is generated by combining every segment.

At the server end, the watermarked audio signal is received and watermark extraction takes place. A coupled-chaotic sequence is generated by the secret keys and applied on the NDFT-domain frequency points of watermarked audio signal to extract watermark, as shown in Fig.2. Then, decryption and increasing dimension are performed on the extracted watermark. The final result of this step is the watermark ready to identify the copyright of the audio signal.

3. Proposed scheme

Our proposed method, with the purpose of designing secure audio watermarking algorithm, consists of dividing audio signal into non-overlap segments, watermark processing, secret keys generation, embedding watermark into the audio signal, and extraction on the server side. In the following subsections, these processes are described in detail.

3.1. Dividing audio signal into non-overlap segments

We divides audio signal, noted as \( \mathbf{A} \) into \( n \) segments. Each segment includes 8 sampling points. Here, \( \frac{8}{L} \) represents the number of segments.

3.2. Watermark processing

In this paper, we utilize digital image as watermark. The process of dealing with watermark consists of the following steps.

1) Reduce dimension

The original digital image with size of \( L \times M \) is noted as \( \mathbf{W} = \{ w(i,j), 0 \leq i \leq L, 0 \leq j \leq M \} \), which is reduced to one dimension, noted as \( \mathbf{W} = \{ w(k), 0 \leq i \leq L, 0 \leq j \leq M, k = i \times M + j \} \).

2) Encryption

In our method, we adopt pseudo-sequence to scramble watermark.

3.3. Secret keys generation

In this paper, we add another chaotic map on the basis of single chaotic model to compose coupled-chaotic sequence. The process is elaborated as follows:

1) One map, which is logistic map, is used to generate 1-D sequence of real numbers that is used as sequence key. Eq.(1) shows the logistic chaotic map

\[
 y_{n+1} = g(y_n) = \mu y_n (1-y_n)
\]
Where, \( n=1,2,3,\ldots \) is the map iteration index and \( \mu \) is the system parameter. For \( 3.57 < \mu \leq 4.0 \), the sequence is non-convergent, and sensitive to the initial value.

2) Another chaotic map is Henon map, which is used to randomly select embedding points. The generalized Henon map is shown in

\[
x_{i+1} = [1+b(x_i-c) + 379x_i] \pmod{1} \quad (2)
\]

Where, \( b=0.3 \) and \( 1.07 \leq c \leq 1.09 \). Modulo operation is performed to restrict the chaotic sequence within limits and it also prevents the chaotic sequence from divergence.

3) Mapping: the chaotic sequence \( x(n) \) is interval within \((0,1)\), to utilize it to randomly select frequency points in NDFT domain, it should be mapped into the NDFT-domain suitable range of embedding watermark.

The NDFT-domain suitable range for embedding watermark is calculated by trying to embed watermark with frequency points \([0, \pi/4, f_1, \pi/4, f_2, \ldots, f_{15}/4, \pi/4] \) in NDFT domain. When the chosen step of \( f_i \) is \( \pi/2048 \), figure 3 shows the SNR curve varied with \( f_i \). To test the robustness, a certain Gaussian noise is added into watermarked signal, figure 4 shows the similarity curve between original watermark and extracted watermark from the noisy watermarked signal.

![Figure 3: SNR](image1)

![Figure 4: Similarity](image2)

The threshold is \( SNR = 32dB \) and \( \rho = 0.8 \), which is considered as having good insensitivity and successful extraction. From above experiments it can be easily made out that the middle frequency variable range \([817\pi/2048, 1329\pi/2048] \) is the selectable frequency range in NDFT domain.

Therefore, the chaotic sequence \( x(n) \) should be mapped into \([817\pi/2048, 1329\pi/2048] \), noted as \( key(n) \), which is used as the secret keys to select NDFT-domain frequency points.

### 3.4. Embedding watermark into audio signal

NDFT transform is carried out for each audio segment. In our method, we choose 8 sampling points to do NDFT, where the frequency points are chosen as: \([0, \pi/4, f_1, 3\pi/4, \pi, 5\pi/4, 2\pi-f_1, 7\pi/4] \). \( f_i = key(n) \).

1-D NDFT and its inverse transform (INDFT) are used as the secret keys to select NDFT-domain frequency points.

\[
F(k_m) = \sum_{n=0}^{M-1} f(n) e^{-jk_m n}, \quad m = 0, \ldots, M-1
\]

Where, \( k_m \) can be any real numbers within \([0, 2\pi] \), \( M \) expresses the number of frequency sampling points.

Quantize \( f_i \) to embed watermark \( \rho_i \). Then inverse NDFT is performed on the watermarked signal coefficient to reconstruct watermarked audio signal.

Now the watermarked audio signal can be transmitted to the server to perform identification of the copyright of audio signal.

### 3.5. Embedding Extraction on the server side

On the receiver side, we use the inverse process of embedding to extract watermark. NDFT transform is done for each segment of received audio signal. Then inverse quantization operation is performed to extract the watermark. At last, increase dimension of the decrypted extracted watermark to retrieve the original digital image.

### 4. Experimental results

#### 4.1 Performance of chaotic secret keys

The coupled-chaotic sequence, which is very sensitive to initial value, has good uniform distribution, randomness, large secret key space. Figure 4 (a) shows the distribution of coupled-chaotic sequence in time domain; (b) shows the chaotic attractor; (c) depicts appearance times of coupled-chaotic sequence within \((0,1)\); (d) shows appearance times of single logistic map. The total experimental number of both (c) and (d) is 30000. In (c), the maximum appearance times is 443, minimum=235, average=300; while in (d), maximum=1972, minimum=168, average=300. From (a) to (d), it illuminates coupled-chaotic sequence has better property of uniform distribution and randomness than single chaotic map. (e) shows the large secret key space up to \(10^{15}\); (f) means coupled-chaotic sequence is very sensitive to the initial values.
4.2 Performance of proposed scheme

1) Robustness against common attacks

The experimental results are shown in Table 1, where the mp3 compression is 128kbps; the frequency of re-sampling attack varies from 44.1khz to 96khz, and then changes back to 44.1khz; the quantization bits of re-quantization attack varies from 16-bit to 8-bit, and then changes back to 16-bit; the low-pass filtering is 6 order with 22.05khz cutoff frequency.

<table>
<thead>
<tr>
<th>Attack Type</th>
<th>DFT [5]</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>No attack</td>
<td>ρ: 35.4072</td>
<td>ρ: 34.5925</td>
</tr>
<tr>
<td>Mp3 compression</td>
<td>ρ: 0.9989</td>
<td>ρ: 35.1246</td>
</tr>
<tr>
<td>Re-sampling</td>
<td>SNR: 0.9998</td>
<td>SNR: 35.3736</td>
</tr>
<tr>
<td>Re-quantization</td>
<td>SNR: 0.9883</td>
<td>SNR: 33.7254</td>
</tr>
<tr>
<td>low-pass filtering</td>
<td>SNR: 35.1785</td>
<td>SNR: 34.4072</td>
</tr>
</tbody>
</table>

ρ: Similarity; SNR: Signal-to-noise ratio

From Table 1 we can easily see that the robustness against common signal processing operations is as high as DFT-based [5]. Even if the SNR is a little lower than [5], it only contributes slight influence in sense of hearing.

2) Robustness against malicious attack

In this paper, we adopt the same malicious attack method that is adding a slight quantity to the frequency points to attack DFT-based and the proposed scheme. The difference just lies in: the former is the fixed frequency points; the later is random sampling points. Fig.6 shows the experimental results. From the results, it is clear that no matter how much attack degree is, proposed watermarking algorithm always has higher similarity than DFT-based watermarking algorithm. Therefore, the proposed algorithm has stronger robustness against malicious attack due to the secret embedding position. That means the proposed scheme has higher security.

5. Conclusion

In this work, we have presented a NDFT-based audio watermarking scheme with high security, in which nonuniform discrete Fourier transform (NDFT) and coupled-chaotic sequence are combined to ensure the security of the audio watermarking scheme. The NDFT characteristic of setting up random sampling points provides the probability of security; the coupled-chaotic sequence generated by secret keys and applied in selecting frequency points in NDFT domain, brings the features of resistant to the finite word length effect brought by single chaotic map, very unpredictable, robust against attacks, and resistant to repeated group attack. Good experimental results have shown the proposed scheme has superior security.

References