2010 International Conference on Networking and Digital Society

An Audio Zero-Watermarking Algorithm Based on FFT

Liting Gao[1], Wei Zhao[2], Xiumei Wen[3] , Lixia Wang[4]

Computer Science and Technology Department, Hebei Institute of Architecture Civil Engineering
Zhang Jia Kou, China
gaoliting@163.com

***Abstract*—This paper proposes an audio zero-watermarking algorithm based on Fast Fourier Transform (FFT).This algorithm does not change any contents about the audio signals. It proposes a solution about the contradictions about imperceptible and robustness. The experimental results show this algorithm can effectively resist low-pass filtering, Gaussian noise and re-sampling attacks. This new audio zero-watermarking algorithm based on FFT can meet the requirements of watermarking security.**

***Keywords*-zero-watermarking; critical band; correlation value; probability of virtual detection; probability of undetected**

I. INTRODUCTION

With the rapid development of Internet and Electronic Commerce, intellectual property protection is the focus of the industrial and commercial fields. Digital watermarking technology adds some identity of ownership in the multimedia data (such as images, audio, video, etc.) in order to achieve copyright protection. At present, the audio digital watermarking can be mainly divided into two categories: the time-domain techniques and the transform domain techniques. While the two methods are different, their essence is the same. The time-domain changes the signal values and the transform domain changes the coefficients in the watermarking embedded in the audio signal. In digital watermarking technology, the data quantity and the robust characteristic becomes a pair of basic contradiction. The ideal watermarking algorithm should not only hide a great deal of data, but also resist various kinds of channel noise and signal deformation.This makes the security of digital watermarking has been limited.

The zero-watermarking algorithm is a new digital watermarking methods, it uses the important features of the image to construct the watermarking, without modifying the image itself. The algorithm proposed in this paper uses the human auditory system (HAS) characteristics, divides the audio signal into sub-frames, and then divides these short-time spectrum of the audio signal struck from the Fourier transform into 25 critical bands. This algorithm selects some frequency coefficients of the important critical bands to construct a series of information, and forms the final watermarking by pseudo-random scrambling on these information series, without changing anything of the audio signal itself. When extracting watermarking information, the algorithm uses the relevant detection methods, doesn't require the original audio signals.

The audio watermarking algorithm has the following three chateristics:

(1) imperceptibility

That is the audio signal into which has been embedded the watermarking signal is the same with the original audio signal to human auditory system(HAS).

(2) robustness

Robustness means that if the unauthorized individual or group delete or modify the embedded watermarking signal using some disposing methods the tone of the original audio signal will decrease significantly. But the embedded information should damage very slightly and can be detected or extracted to the familiar signal processing such as additive noise, filtering, resampling and so on.

(3) safety

The watermarking embedding and detecting algorithm are secret and can't be cracked easily by the unauthorized third party. But the legal owner or user can verify their legal behavior by the watermarking detecting process to achieve the purpose of copyright protection to achieve the purpose of copyright protection.

II. THE FUNDAMENTAL OF DIGITAL WATERMARKING

The digital watermarking system includes generating, embedding, extracting (distilling), checking and evaluating the watermarking etc. This can be expressed as a quadruple (G, E, D, C). Thereinto, G is the Generating algorithm, E is the embedding algorithm, D is the distilling algorithm and C is the checking algorithm.

The watermarking signal W is generated by key K and the digital media X which will be embedded the watermark. This is shown in formula (1).

$$G: X \times K \to W,\ W=G(X, K) \tag{1}$$

The watermarking embedding algorithm is shown in formula (2):

$$E: X \times W \times K \to Xw,\ Xw = E(X0,\ W,\ K) \tag{2}$$

Thereinto, X0 is the original digital media; Xw is the digital media with embedded watermarking.

The watermarking distilling algorithm is shown in formula (3):

$$D: W'=D(X, X0, K) \tag{3}$$

Thereinto, W' is the distilled watermarking signal, X is the digital media which will be detected whether it is embedded the watermark signal; function D uses the algorithm opposite

978-1-4244-5161-6/10/$26.00 ©2010 IEEE

with the embedding algorithm. If this algorithm needn't original digital media X_0, this algorithm is called blind watermarking. Its safety is higher than the watermarking distilling algorithm which needs the original digital media.

The watermarking checking algorithm is shown in formula (4) and (5):

$$C: W' \times W \rightarrow \{0, 1\}$$

$$(W'.W) = \begin{cases} 1, & \text{if } C >= \delta \quad \text{represent X has watermarking} \quad (4) \\ 0, & \text{otherwise} \quad \text{represent X hasn't watermarking} \quad (5) \end{cases}$$

(5)

Thereinto, δ is the checking threshold.

III. Watermark Generation AND Detection

A. Watermarking Generation Algorithm

(1) Perceptual characteristics of the human ear

The sound signal is a non-stationary process, while Fourier transform is apply to periodic, transient or stationary random signal, therefore the standard Fourier transform can not indicate the acoustic signals directly. The short-time Fourier transform should be used to analyse the spectrums of the sound signals. These corresponding spectrums are called the short-time spectrums. They are according to the actual frequency distribution, but the frequency distribution conforming to the human ear's auditory features are based on the critical band frequency distribution. So in this zero-watermarking algorithm the frequencies below 15 kHz will be divided into 24 critical bands, and the 25th critical bands occupy 15 ~ 20 kHz, as shown in Fig. 1.

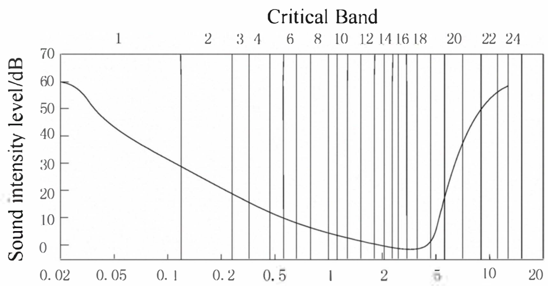

Fig. 1 Audio curve

Ordinary people can feel 20Hz~20kHz, -5~130dB sound signal. So the audio component outside of this range can't be heard and they can be ignored when the sound signals are disposed. From the hearing threshold curve we know that person's ear is most sensitive to the sounds between 1kHz and 5kHz. Person's ear can't keep up with so quick waveform change above 2kHz. Therefore, 1 kHz ~ 2 kHz is the important critical bands, it contains important information about the sound. The zero-watermarking algorithm combining the auditory masking effect, selects some frequency coefficients of the important critical bands as the important feature to form the final watermarking information.

(2) Construction watermarking

The algorithm selects the single-channel audio signals, which is the length of 4.2 s, the sampling frequency is 44.1 kHz, expressed it by A. The watermarking generation process is as follow:

Step 1: Divide the original audio signal A into the sub-frame, and each sub-frame is processed by short-time Fourier transform to get F.

Step 2: Get the first 10,11,12 and 13 critical bands in each sub-frame, which frequency between 1 and 2 kHz, where the lowest frequency is 1080 Hz, and the highest frequency is 2000Hz. Take the frequency coefficients, 0.2 times of the bandwidth, which distributed in the frequency band around the centre frequency, as an important feature coefficient P_k.

Step 3: Quantification of the P_k, $Q_k=[P_k/d+1/2]$, Q_k is the quantized coefficient, [ ] is a symbol to take an integer. d is the pre-set quantitative factors. Because the watermarking is not really embedded in the audio signal, so we needn't consider the imperceptibility of the watermarking. We can take d as a bigger integer, to enhance the robustness.

Step 4: Do mode 2 operation on the Q_k, $C_k=Q_k \bmod 2$, and C_k is the message sequence in binary form. Again do a pseudo-random scrambling on C_k, and form a new sequential D_k. Here we use a pseudo-random scrambling, the order of each element in the sequence randomly is disrupted, while the total number of elements isn't alteration. By this way, it can enhance the security of the algorithm.

Step 5: Do inverse short-time Fourier transform on the coefficients F, which already generated watermarking, and restore the original audio signal A.

Step 6: Map D_k from {0,1} to {-1,1} using the following function (6):

$$W_k = \begin{cases} 1, & (D_k = 1) \\ -1, & (D_k = 0) \end{cases} \tag{6}$$

W_k is just the watermarking.

The watermarking information generated by this algorithm varies with the audio signal and it does not change any eigenvalue of the audio signal. So this algorithm have some self-adaptive and a certain degree of security.

B. Watermarking Detection Algorithm

When the watermarking is being detecting, the algorithm calculates the audio signal watermarking information after a variety of signal attacking by the method mentioned below. Using the following function (7), we can detect the correlation between the two watermarking.

$$SIM(W_k W_k^*) = \sum_{k=1}^{M} W_k W_k^* / \sum_{k=1}^{M} W_k^* W_k^* \tag{7}$$

In the function (7), M is the length of the watermarking information, W* is the detected watermarking. When we are detecting the presence or absence of a watermarking, we should select a threshold T. If the calculated value SIM is greater than T, then there is the watermarking, otherwise there is not. When we are selecting the value of T, we should take into account the two kinds of error probability: the probability of virtual detection and the probability of undetected. The probability of virtual detection refers to the probability of detecting the watermarking, which doesn't have the correct watermarking. The probability of undetected refers to the probability of failure to detect the watermarking, which does have the correct watermarking in the audio signal. If the value of T is too small, it will reduce the probability of undetected and increase the probability of virtual detection. In contrast, if the value of T is too big, it will reduce the probability of virtual detection and increase the probability of undetected. The value of T should be the compromise between the two factors.

C. *Experimental Result*

Audio signal used in the experiment are shown in Fig.2. For the detection threshold, this experiment select 10000 values of (1,1) to form the random numbers sequence, among them the first 5000 is the original watermarking. The experiment detect the similarity of these sequences with the original watermarking respectively, the result is shown in Fig.3.

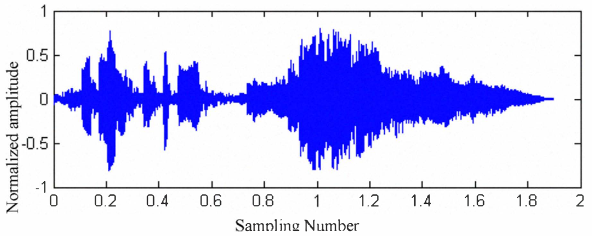

Fig. 2 Audio signal

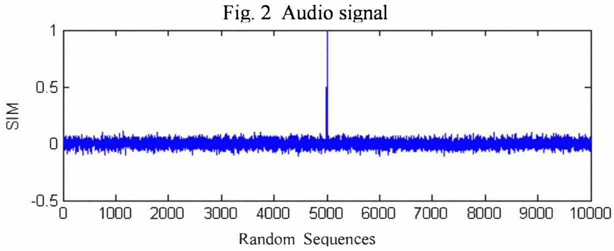

Fig. 3 Correlation detection

From the experimental results, we can see the maximum correlation value between the original watermarking and the watermarking generated randomly is 0.128. This shows the correlation value between the original watermarking and other random sequence is relatively lower. Therefore, the threshold is about 0.20. If the correlation tested is greater than 0.2, we can consider it is a correct watermarking. Fig.3 shows the output of the correct watermarking sequence is far greater than the output of the incorrect watermarking sequence. This indicates that the algorithm has a very low probability of virtual detection and a very low probability of undetected as the same time.

For different audio signals, the threshold may be different, so the threshold can also be re-amplification for the sake of accuracy, to reduce the occurrence of the probability of virtual detection.

In the experiments, we use a variety of signal processing operations on audio signals to detect the robustness of the algorithm. Fig.4 is an audio signal diagram after the low-pass filtering (cut-off frequency = 5 kHz). It shows audio signals after low-pass filtering are smoother a lot than the original. Fig.5 is a re-sampled audio signal diagram. It does the audio signal extraction and the interpolation operation respectively. The extraction and interpolation factor are both 2. After re-sampling, there isn't obvious change in the waveform, but the ratio of signal to noise is higher. SNR=37.5786 dB. Fig.6 is an audio signal diagram after Gaussian noise adding, the mean is 0, that M=0, and the mean square error is 0.001, that MSE=0.001, the diagram seem to become a little "thick" after Gaussian noise adding.

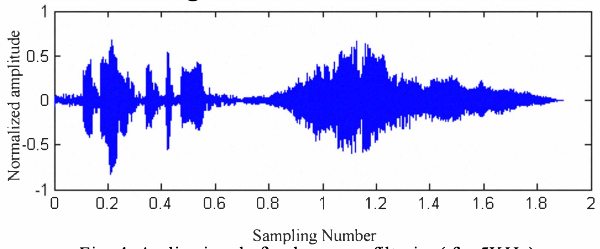

Fig. 4 Audio signal after low-pass filtering(f_H=5KHz)

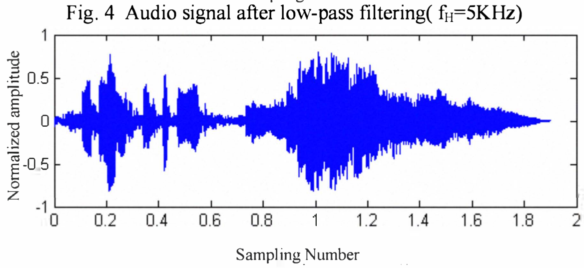

Fig.5 Audio signal after re-sampling (r=2)

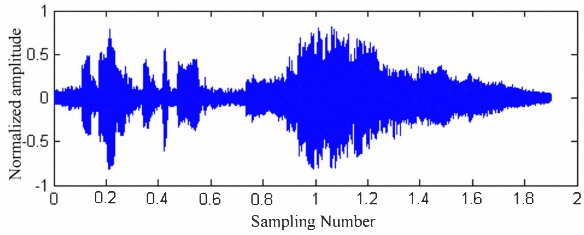

Fig.6 Audio signal after Gaussian noise adding(M=0;MSE=0.001)

Detect the watermarking on the audio signal after the signal processing operation above. The experimental result is shown in Table 1. The SNR in table 1 reflects the damage extent on the original audio signal after a sequence operating processor.

Table 1 Robustness testing

Signal Processing	Parameters	SNR: dB	SIM
low-pass filtering	Cut-off frequency F=5 kHz	18.6425	0.7548
	Cut-off frequency F=4 kHz	17.3868	0.6459
re-sampling	extraction and interpolation factor S=2	35.6875	0.8266
	extraction and interpolation factor S=3	26.6019	0.7891
Gaussian noise	M=0 MSE=0.001	16.2689	0.2864

M=0 MSE=0.003	10.7986	0.2547

Form the experimental result, we can see, even the audio signal has a certain degree of damage, generating a great deal of distortion, and even some post-processing SNR is 17 dB or so, the watermarking still can be detected. So we consider the robustness of this audio zero-watermarking algorithm based on FFT, including low-pass filter, re-sampled and Gaussian noise adding, is better than other watermarking algorithm.

IV. CONCLUSIONS

This paper presents a new zero-watermarking algorithm based on FFT. This algorithm takes full use of the hearing apperceive characteristic of person's ear which is not sensitive to phase. It does not modify the original audio signal, ensure the imperceptibility of the watermarking. The experimental result shows that this algorithm is simple and easy to implement. The extracted watermarking is clearly visible. This digital zero-watermarking algorithm has the stronger robustness, it can resist low-pass filtering, Gaussian noise and re-sampling attacks effectively.

REFERENCES

[1] Won-Gyum Kim, Jong Chan Lee, Won Don Lee. *An Audio Watermarking Scheme with Hidden Signatures*. International Conference on Signal Processing, 2000:21-24

[2] L.Qian, K.Nahrstedt. Watermarking Schemes and Protocols for Protecting Rightful, Ownership and Customer's Rights. Journal of Visual Communication and Image Representation.2005,29(4):194～210

[3] D. Swanson, B. Zhu, H. Tewfik etc. Robust Audio Watermarking Using Perceptual Masking, Signal Processing, 1998, 66(6):337-355

[4] Kirovski D , Malvar H . Robust spread — spectrum audio watermarking[C]. IEEE International Conference on A. coustics, Speech and Signal Processing— Proceedings,2001:(3):1345—1348

[5] Cox I J, Killian J, Leighton F T, etal. Secure spread spectrum watermarking for multimedia. In :IEEE Trans Image Processing [M], 1997,6(12):1673-1687.

[6] XU Chang-sheng, ZHU Yong-wei, FENG Da-gan. Digital audio watermarking based-on multiple-bit hopping and human auditory system[J]. International Multimedia Conference September, 2001, (3): 568-571.

[7] LIANG Juan, GENG Guo-hua. A audio watermark algorithm based on DWT and DCT. Science Journal of Northwest University. Jun. 2004,Vol.2,No. 6

[8] Wang Ye. A new watermarking method of digital audio content for copyright protection. Proceedings of ICSP'98, 1998,1:1420~1423

[9] C.C Chang, C.Y Chung. An Enhanced Buyer-Seller Watermarking Protocol. Communication Technology Processing,2006,15(2):1779～1783

[10] D. Swanson, B. Zhu, H. Tewfik etc. *Robust Audio Watermarking* Using *Perceptual Masking, Signal Processing*, 1998, 66(6):337-355

[11] Digimarc Watermarking Guide-Introduction to Digital Watermarking, http://www.assuredigit.com/tech_doc/more/,2008,3