Audio Watermarking Algorithm for Tracing the Re-recorded Audio Source

Zhenghui Liu, Xuelin Zhao, and Yang Jin

***Abstract*—The popularization of recording equipment and the increase of audio signals propagated through air channels, re-recording becomes a common way to remove watermarks. Audio watermarking scheme robust against re-recording attack is challenging work. This paper proposed a robust audio watermarking algorithm used for source tracking. The logarithmic mean of coefficient (LMC) feature is defined, and the stability of this feature in re-recording attacks is analyzed. We given the method of embedding information by quantizing DCT intermediate frequency coefficients to quantize LMC features. For embedding, frame number and the watermark generated by copyright information are embedded in the carrier signal audio frame. The extraction end extracts the frame number to synchronize the watermarked audio frame, and then extracts the watermark to prove the audio owner and trace the source. Compared with common audio watermarks used for copyright protection, this algorithm is robust against signal processing and de-synchronization attacks, but also robust against re-recording attacks.**

***Index Terms*—digital watermarking, copyright protection, re-recording attacks, resource tracing**

I. Introduction

WITH the explosive increase of audio devices and the widespread use of Internet, people around the world keep creating and spreading much audio data each day. General users can re-record the copyrighted audio data and unauthorized redistribute for their own purposes. Traditional watermarking approach has been used for copyright protection, content authentication, ownership verification, and extensively researched [1], [2]. While, it is a challenging works for the existing literature to address the re-recorded audio piracy and illegal distribution problem. For such applications, watermark should be extracted exactly from the re-recorded data, to track and protect the piracy of the copyrighted audio. Nowadays, how to protect digital audio and trace the source of illegal audio dissemination has become a technical problem, which is public media and judicial authentication agencies urgently need to solve.

Digital audio watermarking utilizes the redundancy in audio and some insensitive characteristics of the human auditory system, and embeds special-purpose information into the carrier signal through a certain algorithm without reducing the auditory quality. The embedded information can protect the copyright of audio works, prove its authenticity and integrity, and track illegal acts such as piracy. There are a large number of published studies that describe digital watermark technology [1], [2], [3], [4]. Large parts of the digital watermarking literature focuses particularly on how to resistance to common signal processing operations (resampling, re-quantization, MP3 compression, etc.) and de-synchronization attacks (pitch-scaling, time-scaling attacks, jittering attack, etc.)[5], [6], [7], [8], [9]. Authors in [6] proposed a new audio watermarking algorithm that against de-synchronization attack. Firstly, the watermark bit is embedded into the digital audio (obtained watermarked signal), and then the synchronization code is embedded into the watermarking signal. The extracter determines whether the watermark signal is scaled by the position changes of the synchronization code and determines scale factor. According to the scale factor, with processing watermark signal and extracting watermark content, extracter can reduce the impact of attacks and improve the accuracy of watermark extraction. In [10], authors proposed a data hiding scheme by using audio files to hide the medical record and to secure it. However, previous studies of digital watermarking have not dealt with re-recorded audio. And it is hard to extract the accurate watermark from the re-recorded signal, obtained by recording watermarked signal spread in air channel, due to the interference of air channels.

The main work of this paper includes two aspects: (1) we proposed a robust audio feature, defined logarithmic mean of coefficient (LMC), which provides a relatively stable embedding domain for the algorithm. (2) We give the embedding method by using the LMC feature, and propose the audio watermarking algorithm robust against re-recording attacks based on the LMC feature. In this paper, frame number and copyright information are embedded in the host audio. The extraction end extracts the frame number to synchronize the watermarked audio frame, and then extracts the copyright information to prove the audio owner and trace the source. While maintaining robustness against common signal processing operations and de-synchronization attacks, the algorithm has a certain ability to resist re-recording attacks, and has accumulated certain technical experience for the protection of audio content transmitted through air channels in the current environment.

Manuscript received December 7, 2020; revised July 12, 2021. This work is supported by National Natural Science Foundation of China (No. 61902085), Nanhu Scholars Program for Young Scholars of XYNU.

Zhenghui Liu is an associate professor in the school of Computer and Information Technology, Xinyang Normal University, Henan Xinyang, China, 464000. Email: zhenghui.liu@163.com

Xuelin Zhao is a lecturer in the school of Computer and Information Technology, Xinyang Normal University, Henan Xinyang, China, 464000. Email: xlz_cit@xynu.edu.cn

Yang Jin is a lecturer in the School of Medical Science, Xinyang Vocational and Technical College, Henan Xinyang, China, 464000. Email: 823294869@qq.com

II. LMC FEATURE

A. The definition of LMC feature

For audio signal $A = \{a_n | 1 \leq n \leq N\}$, we perform DCT to A and obtain the DCT coefficients, denoted by $C = \{c_n | 1 \leq n \leq N\}$, where c_n can be calculated based on the

Eq. (1).

$$c_n = \omega(k) \sum_{n=1}^{N} a_n cos\left(\frac{\pi(2n-1)(k-1)}{2N}\right) \quad (1)$$

In Eq. (1), when $k = 1$, $\omega(k) = 1/\sqrt{N}$, when $2 \leq k \leq N$, $\omega(k) = \sqrt{2/N}$.

We select N DCT intermediate frequency coefficient denoted by $C_{med} = \{c_{n_1} | 1 \leq n_1 \leq N_{med}\}$, $1 \leq N_med \leq N$. We define the LMC feature as F, calculated by using the Eq. (2).

$$F = \left| \sum_{n_1=1}^{N_{med}} log_2\left(\frac{c_{n_1}}{\alpha}\right) \right| \Big/ N_{med} \quad (2)$$

where α is a value greater than zero, satisfying $|c_{n_1}|/\alpha < 1$. Assuming the DCT intermediate frequency coefficient amplitude(α is often not greater than 10, we set $\alpha = 10$ in this paper). According to logarithmic nature, when $|c_{n_1}| > 1$, $log_2|c_{n_1}| > 0$; $0 < |c_{n_1}| < 1$, $log_2|c_{n_1}| < 0$; If we sum $log_2|c_{n_1}|$ $(1 \leq n_1 \leq N_{med})$, then the terms greater than 0 and terms less than 0 will cancel each other out, which can weaken the represent the feature of audio signal. Therefore, in this paper, let each participating item be less than 0, and then the amplitude is calculated. This method enhances the correlation between each coefficient and feature involved in the calculation, and enhances the characterization of the feature.

Since there is a case where the DCT coefficient is 0. If $c_{n_1} = 0$, we take c_{n_1} to a non-zero smaller value and participate in the calculation of the above formula. In this paper, we set $c_{n_1} = 0.001$ when $c_{n_1} = 0$.

B. Robustness Analysis

The audio signal is converted into an analog signal by D/A, then amplified and played by the speaker, and spread in the air channel in the form of sound waves. Audio signal energy is mainly concentrated in the low frequency part. It can also be said that the difference between different signals lies in the difference in the low frequency part. Therefore, the low frequency part of the audio signal changes greatly after re-recording.

In order to show the experimental results more clearly, an audio signal with 12,000 sample points was randomly selected and divided into 4 frames on average, as shown in Fig.1. We perform a re-recording attack on this audio, and the re-recorded signal as shown in Fig.2. Then we apply DCT to the original signal and the re-recorded signal (the corresponding DCT coefficients are depicted in Fig.3 and Fig.4). Table I shows the energy of the DCT low-frequency coefficient and the intermediate frequency coefficient of each frame, and the energy change rate R (original energy/re-recording energy) of different frequencies. From the results shown in Table I, it can be seen that for the re-recording signal, the low-frequency part has a greater energy loss compared with the original signal. Although there is a certain degree of energy loss in the intermediate frequency part, the energy loss is relatively small and closer to the original signal when compared with the low frequency. Therefore, we select the intermediate frequency coefficient to define the LMC feature. Next, we experimentally test the ability of LMC features to resist re-recording attacks.

Generally, the frequency of audio signal is 44.1kHz. In this section, we randomly select an audio signal of length N, with 44.1kHz sample rate, denoted as A, as illustrated in Fig.5.

Step 1: Let us divide A into P frames, we denote the i-th frame as A_i, and the length of each frame is L.

Step 2: We perform DCT to $A^{'}$. Then we select the intermediate frequency coefficient. According to the Eq. (2), we can calculate the LMC feature of $A^{'}$.

We perform a re-recording attack on the selected audio signal as shown in Fig.6. The re-recorded signal is denoted as $A^{'}$. We repeat the steps 1 and 2 above, divide $A^{'}$ into P frames, and calculate the LMC feature $F_i^{'}$ of each frame. Fig.7 shows the comparison of the LMC feature of the audio signal A and the signal $A^{'}$ after the re-recording attack, where the length of each frame is set to 12000.

It can be seen from Fig.7 that the LMC feature of each audio frame after re-recording changes small compared with the original signal, which illustrates the robustness of this feature against re-recording attacks. Next, we test the applicability of LMC features to other audio signals, we selected 240 different audio signals, and re-recorded the selected audio signals (SONY recorder, model PCM-D100). According to the above method, we divide each audio signal into 200 frames. Calculate the difference between the LMC feature of each frame of 240 audio signals before and after the attack. Then we calculate the difference between the LMC feature of each frame of 240 audio signals before and after the attack. Fig.8 shows the statistical mean of the LMC feature difference of each frame before and after the attack. It can be concluded that the maximum change of LMC feature before and after the attack is about 0.5. Compared with LMC feature, the magnitude of change is relatively small (the range of change is about 1/20 of the LMC feature value in Fig.7).

III. Proposed algorithm

Let us denote the watermark embedded into i-th frame as $W_i = \{w_1, w_2, \cdots, w_M\}$, $w_m \in \{0, 1\}(m = 1, 2, \cdots, M)$. We divide W_i into three segments, denoted by $W1_i = \{w_m | 1 \leq m \leq M_1\}$, $W2_i = \{w_m | M_1 + 1 \leq m \leq 2M_1\}$, $W3_i = \{w_m | 2M_1 + 1 \leq m \leq M\}$. In this paper, $W1_i$ and $W2_i$ are the synchronization code generated from the frame number and $W1_i = W2_i$. $W3_i$ is the complete (or partial) watermark information generated by copyright. At the watermark detection, we start from the watermarked signal, we divide the watermarked signal of the first frame into 3 segments. We extract synchronization code from the first and second segments, if the synchronization code extracted from the two segments are the same, we think the frame is complete. Then we extract the watermark information from the third segment.

A. Watermark Embedding

Let us denote the audio signal as $A = \{a_l | 1 \leq l \leq L\}$, where a_l is the l-th sample point, and L is the length of the signal A. The steps of watermarking embedding are summarized as follows.

Step 1: We split A into P equal length frames, the i-th frame is recorded as A_i. The length of each frame is N, $N = L/P$. Then we split A_i into M equal length segments,

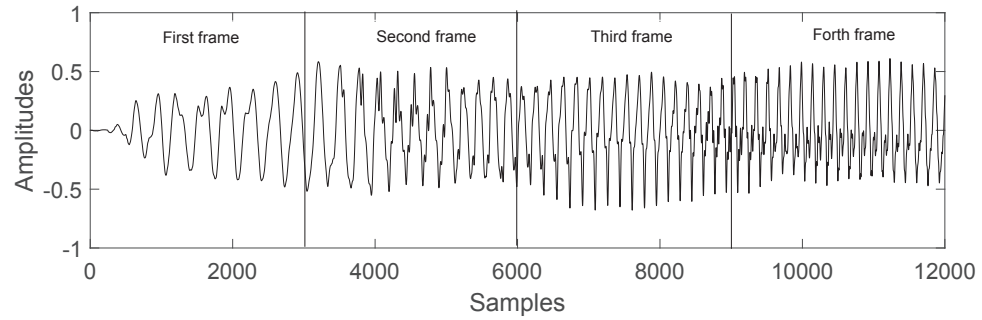

Fig. 1. The audio selected randomly.

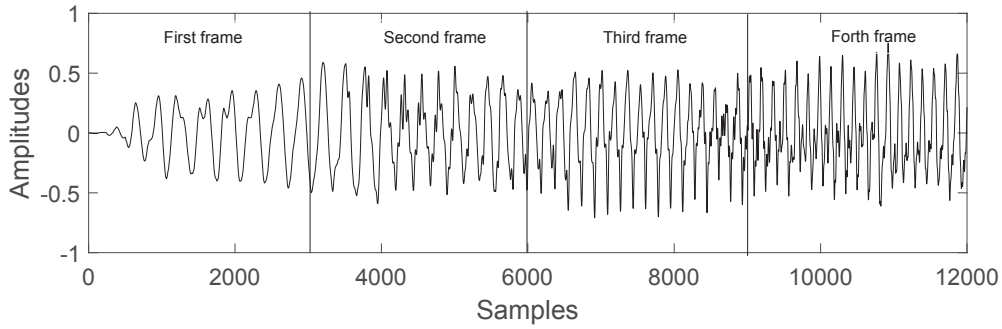

Fig. 2. The re-recording attacked signal of that shown in Fig 1.

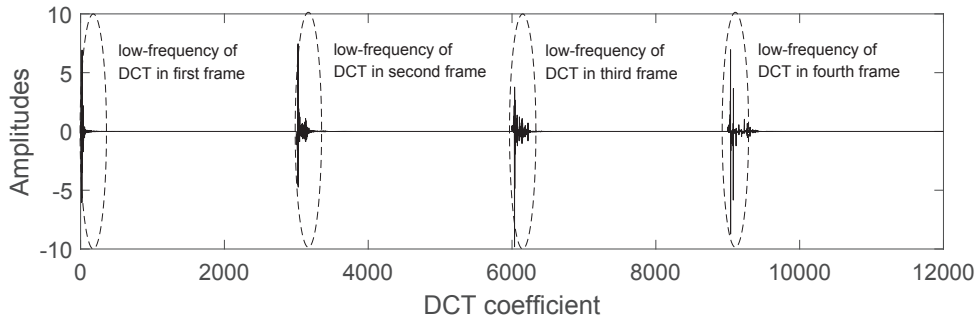

Fig. 3. DCT coefficients of 4 frames of the original signal.

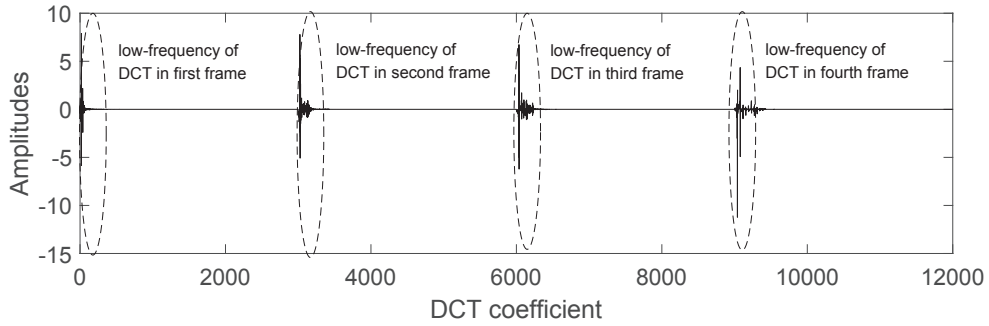

Fig. 4. DCT coefficients of 4 frames of the re-recorded signal.

TABLE I
COMPARISON OF LOW FREQUENCY AND INTERMEDIATE FREQUENCY ENERGY OF DCT BETWEEN ORIGINAL SIGNAL AND RE-RECORDED SIGNAL.

		Original / re-recorded			
		First frame	Second frame	Third frame	Fourth frame
Low frequency		45/11	26/11	24/17	32/6
	R	4.1	2.4	1.4	5.3
intermediate frequency		3/3	7/6	17/12	24/21
	R	1	1.2	1.4	1.1

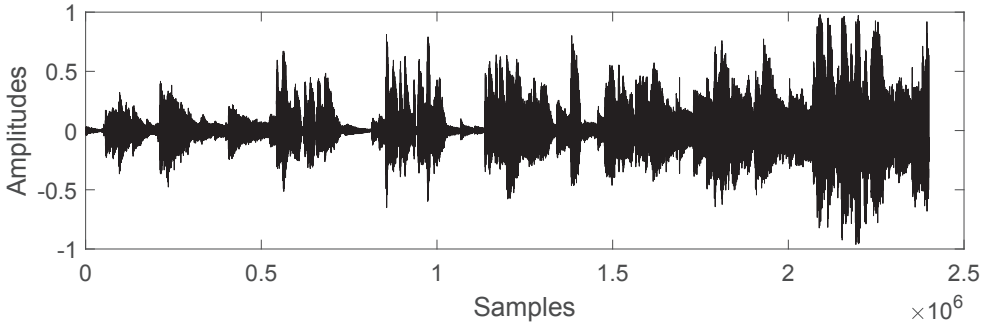

Fig. 5. Original audio signal.

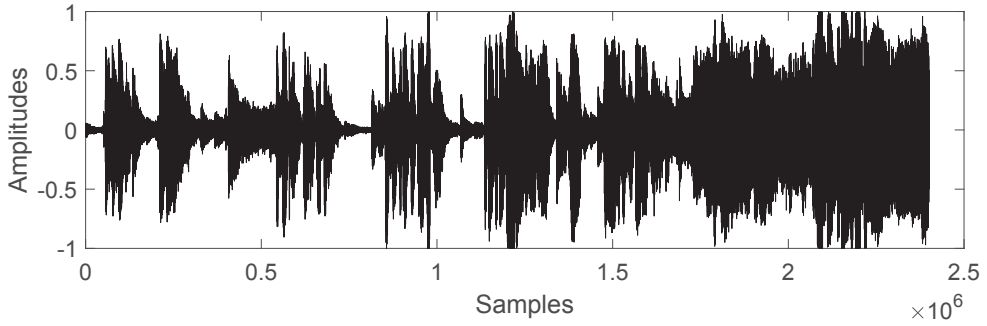

Fig. 6. The re-recorded audio signal.

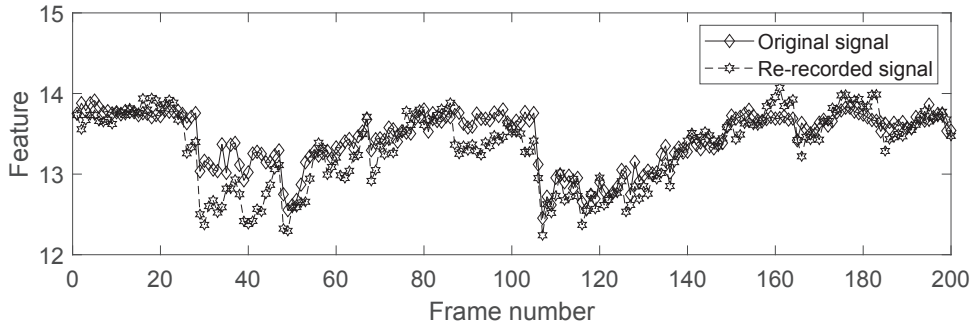

Fig. 7. The LMC feature of original and re-recorded audio.

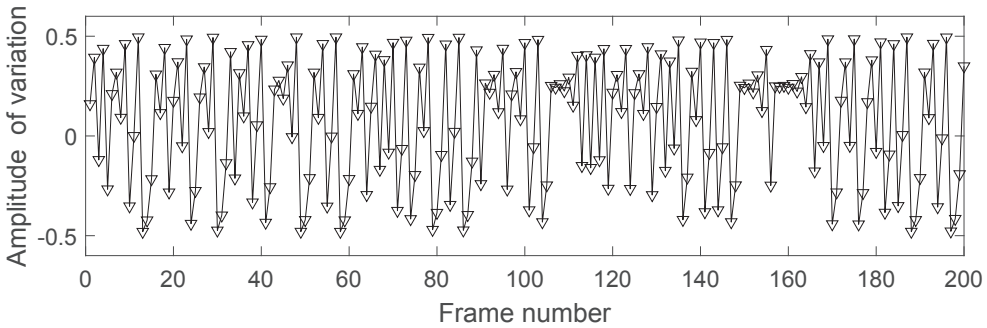

Fig. 8. The amplitude of variation of LMC feature after being re-recorded.

denoted as $A_{i1}, \cdots, A_{iM}$, the length of each segment is N_1 .

Step 2: We use chaotic mapping method to scramble each signal, denoted by $S_{i1}, \cdots, S_{iM}$. The method is described as follows.

1) We use logistic mapping to generate pseudo-random sequence. In this paper, the logistic mapping given in [11], and the generated sequence is recorded as $X = \{x_l | l = 1, 2, \cdots, N_1\}$. In addition, we can get x_l based on the Eq. (3).

$$x_{l+1} = \mu x_l(1 - x_l), 3.5699 \leq \mu \leq 4 \tag{3}$$

where $x_0 = k$ is the initial value of logistic mapping, which is used as the key of the watermarking system.

2) Assume $A_{i,1} = \{a_l | 1 \leq l \leq N_1\}$. Arrange the elements in the pseudo-random sequence in ascending order, the address index sequence in ascending order can be obtained by using the Eq. (4).

$$x_{h(l)} = ascend(x_l), l = 1, 2, \cdots, N_1 \tag{4}$$

The signal after scrambling is recorded as $S_{i,1}$, $S_{i,1} = \{a_{h(l)} | 1 \leq l \leq N_1\}$. Then we use the same method to scramble the audio signals of other segments.

Step 3: We apply DCT to the scrambled signal $S_{i,1}$ and obtain the DCT coefficients $C_{i,1} = \{c_1, c_2, \cdots, c_{N_1}\}$. Then we select the intermediate frequency coefficient and calculate the LMC feature, denoted by $F_{i,1}$.

Step 4: We then quantify $F_{i,1}$ according to the watermark w_1.

$$QF_{i,1} = \begin{cases} \eta \times \triangle + \rho, & \eta \mod 2 = w_1 \\ \eta \times \triangle - \rho, & else \end{cases} \tag{5}$$

where $QF_{i,1}$ is the quantized LMC feature, $\lfloor \cdot \rfloor$ is rounded down and $\rho = \triangle /2, \eta = \lfloor QF_{i,1} / \triangle \rfloor$.

Step 5: After feature $F_{i,1}$ being quantified, we can obtain the DCT intermediate frequency coefficient corresponding to the quantized feature $QF_{i,1}$ according to the Eq. (6).

$$c_n^* = sign(c_n) \cdot |c_n|^{\frac{QF_{i,1}}{F_{i,1}}} \cdot \alpha^{1 - \frac{QF_{i,1}}{F_{i,1}}}, 1 \leq n \leq N_{med} \tag{6}$$

where c_n is intermediate frequency coefficient before quantization, c_n^* is intermediate frequency coefficient after quantization; $sign(c_n)$ is sign function, when $c_n > 0$, $sign(c_n) = 1$, else $sign(c_n) = -1$.

Step 6: Finally, we combine the quantized DCT intermediate frequency coefficients and other low and high frequency coefficients, then we apply the inverse DCT and anti-scrambling. After that, we can embed w_1 in the frame.

Repeat the steps above, we can embed other watermark information $w_2, w_3, \cdots, w_M$ into other segments $A_{i1}, \cdots, A_{iM}$ of A_i.

We show the watermark embedding of the method in Fig.9.

B. Watermark Extraction

The watermark extraction process can be regarded as the inverse process of the watermark embedding process. Let us denote the watermarked signal as A', and length of A' is L'. Then we divide A' into frames, and denote A' as the i-th frame. Finally, we divide A' into M equal length frames, denoted as $A_{i,1}^{'}, A_{i,2}^{'}, \cdots, A_{i,M}^{'}$. The steps of watermarking extraction are summarized as follows.

Step 1: Based on Eq. (4), we can scramble the sample points of the $A_{i,1}^{'}$, and then we apply DCT to the scrambled signal.

Step 2: We select the DCT intermediate frequency coefficient. Based on the intermediate frequency coefficients, we can calculate the LMC feature and denoted as $F_{i,1}^{'}$, by using the Eq. (2).

Step 3: According to the Eq. (7), we extract the information embedded in $A_{i,1}^{'}$, denoted as $F_{i,1}^{'}$.

$$w^{'} = \lfloor F_{i,1}^{'} / \triangle \rfloor \mod 2 \tag{7}$$

Step 4: Repeat steps 1 to 3 above to extract the embedded information in $A_i^{'}$. We record the extracted information as $W_i^{'} = \{w_1, w_2, \cdots, w_M\}$ and divide $W_i^{'}$ into three segments, denoted as $W_{1i}^{'} = \{w_m^{'} | 1 \leq m \leq M_1\}$, $W_{2i}^{'} = \{w_m^{'} | M_1 + 1 \leq m \leq 2M_1\}$, $W_{3i} = \{w_m^{'} | 2M_1 + 1 \leq m \leq M\}$, respectively.

Step 5: According to the three segments of information by step 4, we can determine whether the frame is a synchronized audio frame. If $\sum_{i=1}^{M_1} (w_m^{'} \oplus w_{m+M_1}^{'}) = 0$, according to the embedding algorithm, we can accurately extract synchronization information from this frame; if $\sum_{i=1}^{M_1} (w_m^{'} \oplus w_{m+M_1}^{'}) \neq 0$, we move the frame with a sliding window and extract the synchronization information until we find a frame that satisfies $\sum_{i=1}^{M_1} (w_m^{'} \oplus w_{m+M_1}^{'}) = 0$, and the frame is a new synchronized audio frame. After detecting the synchronized audio frame, the watermark information is extracted from the third segment.

We show the watermark extraction of the method in Fig.10.

IV. Experimental results

he performance of the algorithm is verified by simulation with MATLAB. We selected database containing 200 speech signals and they are 16-bit quantified mono signals with 44.1 kHz sampling rate. We use different recording equipment to re-record the watermarked signal (SONY PCM-D100, HUAWE P20 and iPhone 6s). The parameters used are set as $L = 600000$, $N = 12000$, $M = 10$, $M_1 = 3$, $\triangle = 0.8$, $\alpha = 12$.

A. Quality of Watermarked Signal

In this paper, in order to measure the quality of watermarked signals, we take subjective difference grade (SDG) and signal noise ratio (SNR) to evaluate the quality of watermarked signal. The meaning of each grade of SDG is shown in Table II [12]. The score of SDG is between -4 and 0, and the perceptual quality improves with the increase of the SDG value. The calculation formula of SNR is shown in Eq. (8) [9].

$$SNR = 10 \lg \left(\frac{\sum_{l=1}^{N} a(l)^2}{\sum_{l=1}^{N} (a(l) - a^{'}(l))^2} \right) \tag{8}$$

where $a(l)$ is original audio signal, $a^{'}(l)$ is the audio signal with watermark, N is the length of audio signal. The SNR value should be above 20dB without affecting the hearing.

In this paper, we calculate the average SNR and SDG of the selected 200 test audio signals (SDG value is calculated by 14 listeners scoring). Then, we listed these values in Table

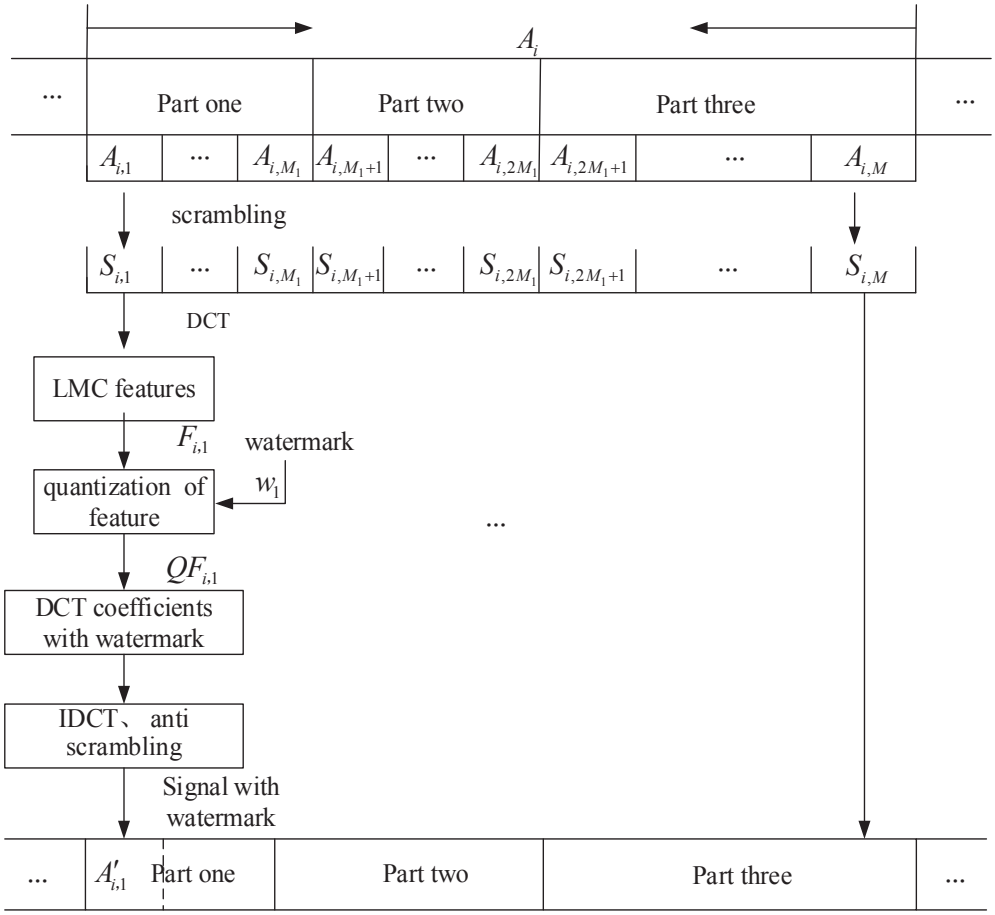

Fig. 9. The process of embedding.

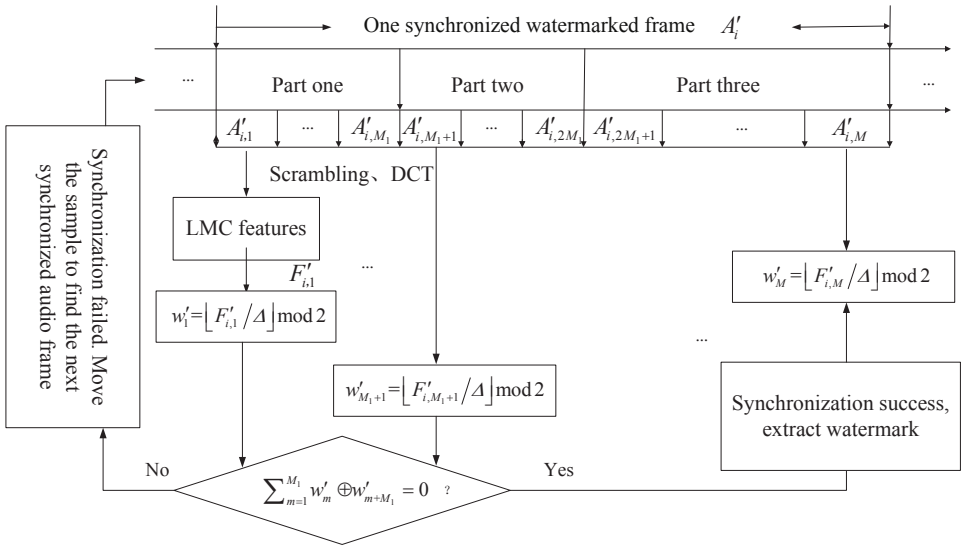

Fig. 10. The process of watermark extraction and synchronization.

TABLE II
SDG EVALUATION CRITERIA

SDG	Describe
0.0	imperceptible
-1.0	felt but not harsh,
-2.0	slightly harsh
-3.0	harsh
-4.0	very harsh

TABLE III
SNR, AND SDG MEAN VALUES OF WATERMARKED SPEECH

		Watermarking Audio
SNR (dB)	Maximum	33.1
	Mean	29.38
	Minimum	25.6
SDG	Maximum	-0.4
	Mean	-0.61
	Minimum	-0.77

III, and the result is that the mean SNR is 29.38 and the mean SDG is -0.61, which can achieve the performance index of imperceptibility (SNR value is greater than 20, SDG value is greater than -1). Therefore, the embedding of the watermark in the algorithm does not affect the auditory quality of the original signal.

B. Robustness

We hope to solve the problem for source tracking of audio watermarking when the audio signal is transmitted through the air channel and re-recorded by the audience in this paper. The performance and application effect of the algorithm depend on the ability to accurately extract the watermark information from the re-recorded audio signal. In this section, we evaluate the robustness against various attacks by using the bit error rates (BER), which is defined based on Eq. (9).

$$BER = \frac{1}{M}\sum_{m=1}^{M}\Big(w(m) \oplus w^{'}(m)\Big) \tag{9}$$

where M is length of embedded watermark, $w(m)$ and $w^{'}(m)$ are the original watermark information and the extracted watermark information, respectively. $\oplus$ is the exclusive (XOR) operator. BBesides, the small BER value shows that the lower the error rate of extracting the watermark, the better the robustness of the algorithm, and the stronger the anti-attack ability.

In this part, we evaluate the robustness of the proposed method against signal processing, de-synchronization attacks.

1) Robustness against signal processing and de-synchronization attacks: The de-synchronization attacks can stretch or shorten the length of the watermarked audio signal. Therefore, in the watermark extraction stage, in order to improve the extraction rate of watermark information from the watermarked audio signal undergoing de-synchronization attack, we try a variety of different lengths for framing. For example, the length of each frame can be $90\% \times L/P, L/P, 110\% \times L/P, L^{'}/P$, respectively. Then we take the frame with the highest watermark extraction accuracy as the effective frame length.

The BER values of watermarked signals for common signal processing operations and de-synchronization attacks are shown in Table . The common signal processing operations include MP3 compression (with compression bit rates of 64kbps and 128kbps), resampling (44.1kHz22.05kHz44.1kHz). The de-synchronization attacks contain jittering attacks. The selected parameters are 1/100 and 1/1000, which means that one sample is randomly deleted from every 100 and 1000 samples from the watermarked signals.

TABLE IV
PERFORMANCE COMPARISON (BER (%)) OF THE PROPOSED METHOD AND THE METHODS IN [13],[14]

Attacks		[13]	[14]	Prosed
MP3	64bps	5	6	1
	128bps	3	3	0
Resampling		8	8	0
Jittering	1/100	7	8	0
	1/1000	2	3	0
Time-scaling	90%	13	15	7
	110%	12	13	5

TABLE V
PERFORMANCE COMPARISON (BER (%)) OF THE PROPOSED METHOD AND THE METHODS IN [13], [14] FOR THE RE-RECORDING ATTACK

Attacks			[13]	[14]	Prosed
No post processing			12	18	4
Recapturing and	MP3	64bps	15	19	6
		128bps	12	18	4
	Resampling		13	18	4
	Jittering	1/100	13	21	4
		1/1000	12	19	4
	Time-scaling	90%	18	23	11
		110%	15	22	9

The results show that our proposed method achieves lower BER values than the methods in [13], [14] for all the attacks in Table . Under common signal processing operations and de-synchronization attacks, our method better than the methods in [13], [14]. It shows that our method has greater anti-attack ability in signal processing operations and de-synchronization attacks compared with the methods in [13], [14].

2) Robustness against re-recorded attacks: In this part, we test the ability of our method to resist re-recording attacks. Firstly, we recaptured the watermarked audio signal, then we perform common signal processing operations and de-synchronization attacks on the re-recorded signal.

When comparing our method with the methods in [13], [14], the BER value of our method and the methods in [13], [14] have all increased to varying degrees in Table , but in the case of our method, the increase is slight. As to re-recorded attacks, the BER value of our method is 4, and the methods in [13] and [14] are 12 and 18, respectively. It indicates that the BER value of our method reduced by up to 14 percentage points. For other attacks, our method has the lowest BER value than the methods in [13], [14]. It demonstrates that the proposed scheme has greater robustness against to the above attacks.

V. CONCLUSION

In this paper, a robust audio watermarking algorithm is proposed. It can solve the problem which can trace to the source after air channel propagation and re-recording. Firstly, the LMC feature of audio signal is defined, besides, the robustness of LMC features in re-recording attack is analyzed. The method of quantifying LMC features is used to embed the frame number and watermark into the host

carry signal. When the watermark signal is attacked, the frame number is used to synchronize the audio frame with watermark, and then the watermark information is extracted to trace the source. Compared with the existing robust audio watermarking algorithms for copyright protection, the algorithm in this paper not only improves the ability to resist de-synchronization attacks, but also has a certain degree of robustness against re-recording attacks.

REFERENCES

[1] Hua, Guang, Huang, Jiwu, Shi, Yun, Q., Thing, Vrizlynn, and L., "Twenty years of digital audio watermarking-a comprehensive review," *Signal Processing*, vol. 128, no. 11, pp. 222–242, 2016.

[2] N. Akira, "Audio watermarking based on subband amplitude modulation," *Acoustical Science and Technology*, vol. 31, no. 5, pp. 328–336, 2010.

[3] Z. Liu and H. Wang, "A novel speech content authentication algorithm based on besselcfourier moments," *Digital Signal Processing*, vol. 31, no. 5, pp. 328–336, 2010.

[4] L. I. Chen, K. Wang, and L. Tian, "Audio watermarking algorithm in mp3 compressed domain based on low frequency energy ratio of channels," *Journal of Computer Applications*, vol. 38, no. 8, pp. 2301–2305, 2018.

[5] H. Peng, D. Peng, Y. Zhang, and X. Yong, "Robust time-spread echo watermarking using characteristics of host signals," *Electronics Letters*, vol. 52, no. 1, pp. 5–6, 2016.

[6] X. Yong, I. Natgunanathan, G. Song, W. Zhou, and S. Nahavandi, "Patchwork-based audio watermarking method robust to de-synchronization attacks," *IEEE/ACM Transactions on Audio Speech & Language Processing*, vol. 22, no. 9, pp. 1413–1423, 2014.

[7] M. Marwan, F. Alshahwan, F. Sifou, A. Kartit, and H. Ouahmane, "Improving the security of cloud-based medical image storage," *Engineering Letters*, vol. 27, no. 1, pp. 175–193, 2019.

[8] X. Wang, T. Ma, and P. P. Niu, "A pseudo-zernike moment based audio watermarking scheme robust against desynchronization attacks," *Computers & Electrical Engineering*, vol. 37, no. 4, pp. 425–443, 2011.

[9] B. Soni, P. K. Das, and D. M. Thounaojam, "Dual system for copy-move forgery detection using block-based lbp-hf and fwht features," *Engineering Letters*, vol. 26, no. 1, pp. 171–180, 2018.

[10] M. B. Andra, T. Ahmad, and T. Usagawa, "Medical record protection with improved grde data hiding method on audio files," *Engineering Letters*, vol. 25, no. 2, pp. 112–124, 2017.

[11] J. Wang, Z. Liu, Q. I. Chuanda, and H. Wang, "A content authentication algorithm for digital speech signal robust against feature-analyzed substitution attack," *Journal of the China Railway Society*, vol. 28, no. 6, pp. 73–78, 2016.

[12] Z. Liu, Z. Fan, W. Jing, H. Wang, and J. Huang, "Authentication and recovery algorithm for speech signal based on digital watermarking," *Signal Processing*, vol. 123, no. 1, pp. 157–166, 2016.

[13] A. Nadeau and G. Sharma, "An audio watermark designed for efficient and robust resynchronization after analog playback," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 6, pp. 1393–1405, 2017.

[14] J. F. Li, H. X. Wang, T. Wu, X. M. Sun, and Q. Qian, "Norm ratio-based audio watermarking scheme in dwt domain," *Multimedia Tools & Applications*, vol. 77, no. 12, pp. 14 481–14 497, 2017.