文章编号: 1004-4353(2019)03-0228-06

基于 PN 序列和完全互补码的鲁棒音频水印算法

陈亮, 李德*
(延边大学 工学院, 吉林 延吉 133002)

摘要: 为了解决音频作品容易引发版权纠纷的问题,提出了一种基于 PN(Pseudo-Noise)序列和完全互补码的音频扩频水印算法. 首先,在时域内引入回声,利用 PN 序列改变前后向回声内核的极性,并对回声内核进行加密,以此提高水印的安全性和透明性;然后,利用完全互补码技术对水印序列进行扩频,以此提高水印算法的鲁棒性和嵌入容量;再次,利用 PN 序列的短时能量和随机特性,将扩频调制后的水印信息自适应地嵌入到音频信号的时域波形上;最后,利用 PN 序列和完全互补码扩频得到水印. 实验结果显示,在保证水印透明性的基础上,且在回声嵌入强度小于 0.01 的情况下,本文提出算法的水印信息检测率超过 98%,对高斯噪声、滤波和重采样等攻击的水印信息检测率超过 90%,这表明本文算法鲁棒性较强,具有良好的应用价值.

关键词: 音频水印; 伪噪声序列; 回声隐藏; 完全互补码
中图分类号: TP391.41 **文献标识码**: A

Robust watermarking based on PN sequence and complete complementary code

CHEN Liang, LI De*
(*College of Engineering*, *Yanbian University*, *Yanji* 133002, *China*)

Abstract: In order to solve the copyright disputes of audio works, an audio spread spectrum watermarking algorithm based on PN (Pseudo-Noise) sequence and complete complementary code is proposed. Firstly, echo is introduced in time domain, and the polarity of forward and backward echo cores is changed by PN sequence, and the echo cores are encrypted to improve the security and transparency of watermarking. Secondly, the watermarking sequence is spread spectrum using complete complementary code technology to improve the robustness and embedding capacity of the watermarking algorithm. Thirdly, the short-term energy and random characteristics of PN sequence are used to expand the watermarking algorithm. After frequency modulation, the watermarking information can be adaptively embedded into the time domain waveform of audio signal. Finally, the watermarking is obtained by PN sequence and complete complementary code spread spectrum. The experimental results show that, on the basis of ensuring the transparency of the watermarking, the detection rate of the watermarking information exceeds 98% when the echo hiding strength is less than 0.01, and the detection rate of the assault watermarking information such as adding noise, filtering and resampling exceeds 90%. This shows that the proposed algorithm has strong robustness and good application value.

Keywords: audio watermarking; PN sequence; echo hiding; complete complementary code

0 引言

随着计算机网络技术的蓬勃发展,数字音频作品的存储、复制和传播变得越来越容易,而随之而来的版权纠纷问题时有发生[1]. 近年来,一些学者利用数字音频水印技术对数字产品的版权问题进行了研

收稿日期: 2019-06-10 * **通信作者**: 李德(1972—),男,博士,教授,研究方向为信息安全与版权保护.
基金项目: 吉林省教育厅"十三五"科学技术研究项目(吉教科合字[2016]第 249 号)

究，并取得了较好的研究成果. 例如：高雪丽等[2]利用奇异值分解技术提出了一种轮廓波变换的水印算法，该算法在抵抗高斯噪声、裁剪攻击和压缩攻击方面具有很强的鲁棒性，但当压缩比逐渐提高后，鲁棒性会逐渐降低. 杨振仁等[3]利用图像归一化技术提出了一种采用 Arnold 和扩频技术将音频进行离散余弦变换(discrete cosine transform，DCT)的水印算法，该算法对仿射变换、高斯噪声等攻击具有很好的鲁棒性，但在抵抗剪切攻击方面相对较弱. Hu 等[4]在变换域的基础上应用量化索引调制技术，并根据人耳的听觉掩蔽效应提出了一种 DCT 水印算法，该算法对高斯噪声和滤波攻击具有较好的鲁棒性，但对幅值修改、剪减等几何攻击的鲁棒性较差. Bhat 等[5]通过对音频信号块进行奇异值分解，提出了一种离散小波变换(discrete wavelet transformation，DWT)算法，该算法在水印检测方面的误检率和漏检率较低. Lei 等[6]提出了一种基于音频向量关系的算法，经计算机模拟验证表明该算法对水印有良好的鲁棒性，但对音频的随机剪切等几何攻击的鲁棒性较差. Fallahpour 等[7]提出了一种基于奇异值分解(singular value decomposition，SVD)的 DCT 域水印算法，该算法在水印的鲁棒性和可嵌容量方面较为平衡，但因缺乏同步机制容易造成误检测. Kumsawat 等[8]提出了一种遗传算法(genetic algorithm，GA)，该算法在理论上可以平衡水印的不可感知性和鲁棒性，但实验结果显示鲁棒性并未有明显提升. 基于上述研究，本文针对数字音频作品的版权保护，将回声隐藏技术和扩频技术相结合，提出一种对多种攻击透明性好、鲁棒性强的音频水印算法.

1　相关技术介绍

1.1　PN 序列回声隐藏

最长线性反馈移位寄存器序列是回声隐藏技术中最为常见的伪随机序列(M 序列)，其通常由反馈移位寄存器生成. 由于 M 序列和随机噪声的特性都具有随机特性和周期性，因此 M 序列又被称为伪噪声序列(PN 序列)[7]. 在信息隐藏过程中，首先需对音频信号进行分割，并将回声引入每个音频段的前后核，然后通过计算音频信号中每个音频段的短时能量值，以此自适应地调整前后回声核的衰减系数 a_n. 自适应调整规则如下：

$$a_n=\begin{cases}a_1,\ E_n\in\left(0,\dfrac{1}{2}\bar{E}\right);\\ a_2,\ E_n\in\left[\dfrac{1}{2}\bar{E},\bar{E}\right);\\ a_3,\ E_n\in\left[\bar{E},\dfrac{3}{2}\bar{E}\right);\\ a_4,\ E_n\in\left[\dfrac{3}{2}\bar{E},\infty\right).\end{cases}\tag{1}$$

其中 E_n 为第 n 个音频段的短时能量，$\bar{E}$ 为所有音频段的短时平均能量. E_n 值所属的区间分别为 $\left(0,\frac{1}{2}\bar{E}\right)$、$\left[\frac{1}{2}\bar{E},\bar{E}\right)$、$\left[\bar{E},\frac{3}{2}\bar{E}\right)$、$\left[\frac{3}{2}\bar{E},\infty\right)$，每个区间依次为 a_1、a_2、a_3、a_4，即由每个音频段的短期能量值的大小来确定音频段选择的衰减系数的大小. 回声隐藏方法的核心是通过引入 *PN* 序列来修正前向和后向的回声核，然后利用 *PN* 序列在时域中均匀地引入回声，并根据短时能量值自适应地调整衰减系数.

1.2　完全互补码

完全互补码(complete complementary code，CCC)通常由互相关函数来衡量. 互相关函数可描述为：对于任意长度为 L 的有限复序列 $c_a=\{c_{a,0},c_{a,1},\cdots,c_{a,L-1}\}$，定义其有限长序列 $C_a=\{C_a(t)\}$ 为：$C_a(t)=\left\{\sum_{i=0}^{L-1}C_{a,i}\delta_{i,t}\right\}$，其中 $\delta_{i,t}$ 代表克罗内克 δ 函数. 当 $t\leqslant-1$，$L\leqslant t$ 时，$C_a(t)$ 为 0. 由此可定义 $C_a(t)$ 为 C_a，则对任意 2 个有限长序列有 $c_1=\{c_{1,0},c_{1,1},\cdots,c_{1,L-1}\}$，$c_2=\{c_{2,0},c_{2,1},\cdots,c_{2,L-1}\}$.

由以上描述可将互相关函数表示为 $R_{c_1c_2}(\tau)$，且当 $c_1=c_2$ 时 $R_{c_1c_2}(\tau)$ 为自相关函数. 由此可知，可将完全互补码定义为一组 M 序列的自动互补码，由其构成的矩阵为：

$$\boldsymbol{CCC}=\begin{bmatrix} C_0^{(0)} & C_1^{(0)} & \cdots & C_{N-1}^{(0)} \\ C_0^{(1)} & C_1^{(1)} & \cdots & C_{N-1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ C_0^{(M-1)} & C_1^{(M-1)} & \cdots & C_{N-1}^{(M-1)} \end{bmatrix}_{M\times N}. \tag{2}$$

其中 M 和 N 分别表示完全互补码构造矩阵的行数和列数，$\boldsymbol{C}$ 表示一段完全互补序列. 公式(2)中每行构成一组完整的完全互补序列，即每行在行内形成自相关，每两行序列码之间形成互相关.

2 音频水印算法的设计

2.1 音频水印的嵌入算法

本文算法的音频水印嵌入流程如图 1 所示，构造步骤见步骤 1— 步骤 3.

步骤 1 水印序列的编码. 采用 12 位的二进制信息作为水印信息，利用 Hadamard 矩阵生成 (m,n,l) -完全互补码 (N,N,N^2)，其生成方法如下：① 构造一个 N 阶 Hadamard 矩阵，$\boldsymbol{A}=[A_1A_2\cdots A_N]^{\mathrm{T}}$. ② 取 $\boldsymbol{B}=[B_1B_2\cdots B_N]^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{A}\times\mathrm{diag}(A_1)\\ \boldsymbol{A}\times\mathrm{diag}(A_2)\\ \vdots\\ \boldsymbol{A}\times\mathrm{diag}(A_N)\end{bmatrix}$. ③ 令 $\boldsymbol{C}=\begin{bmatrix}B_1 & B_2\\ B_1 & -B_2\\ B_3 & B_4\\ B_3 & -B_4\\ \cdots & \cdots\end{bmatrix}$. ④ 定义 $\overset{\sim(i)}{S}=\begin{bmatrix}B_i & B_{i+1}\\ B_i & -B_{i+1}\end{bmatrix}$ 是一组完全互补码. ⑤ 选取合适的密钥 (P,Q) 为生成密码. 根据前面定义的完全互补码可知 (P,Q) 为密钥对，其中 P 代表由 $\overset{\sim(i)}{S}$ 定义的所有选基序列的序列号 i，Q 代表其循环移位位数. 例如：$(2,100)$ 表示基序列 $\overset{\sim(i)}{S}$ 的自互补码是 $\{c_0^{(2)},c_1^{(2)},\cdots,c_{n-1}^{(2)}\}$，同时 $S^{(i)}$ 需要右移 100 位.

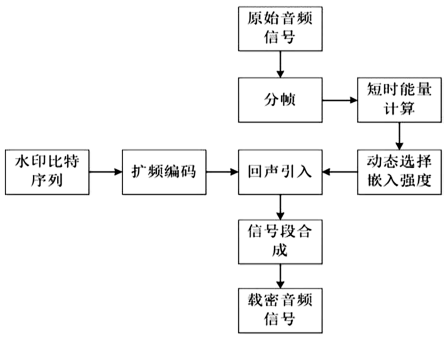

图 1 音频水印信息的嵌入流程

步骤 2 载体信息处理. 本文在载体信息处理中使用的 PN 序列为 10 阶的 M 序列. 通过计算水印序列 $w(n)$ 和 PN 序列 $c(n)$ 即可生成扩频信号 $s(n)$，$s(n)$ 的计算公式为 $s(n)=w(n)\oplus c(n)$. 由于 PN 序列具有伪随机特性，因此扩频信号 $s(n)$ 的频谱宽度远大于原始水印的频谱宽度. 设水印信息 $d(n)$ 的比特率为 Kbps，如果扩频序列每个码片的传送时间为 T_c，则扩频信号的带宽 $W_{ss}=1/T_c$. 在载体信息处理阶段，基带信号也需要处理，其处理过程为：首先对二进制数字基带信号进行差分编码，将绝对码表示的二进制信息转换成相对码表示的二进制信息，然后进行绝对相位调制.

步骤 3 水印信息的嵌入. 水印隐藏信息的整个嵌入过程具体如下：

1) 使用矩形窗口均匀分割音频信号时，窗口长度以 T 为单位，即每个窗长包含 $N(N=T\times f_s$，f_s 为音频信号的采样频率) 个采样点数，且将音频信号分为 M 个音频段(不足一个时长 T 的部分不计).

2) 计算每个音频帧的短期能量和平均短期能量. 平均短期能量的计算公式[8] 为：$\bar{E}=\frac{1}{M}\sum_{m}^{M}E_m$.

3) 对前后向回声核进行改造得到新的回声内核，改造公式为：$k_n=\delta(n)+(-1)^{p+1}(\alpha\delta(n-d)+\alpha\delta(n+d))$，$0<\alpha<1$. 将 d_0 和 d_1 定义为延时，然后按照公式(3)完成水印嵌入.

$$y'(n)=\begin{cases}x(n)\times\delta(n)+(-1)^{p+1}(\alpha\delta(n-d_1)+\alpha\delta(n+d_1)),\\ x(n)\times\delta(n)+(-1)^{p+1}(\alpha\delta(n-d_0)+\alpha\delta(n+d_0)).\end{cases} \tag{3}$$

4) 将经过扩频调制的序列 $x(n)$ 嵌入至原始音频序列中，嵌入的规则为：$sig^{\mathrm{w}}(n)=sig(n)+\alpha x(n)$. 其中 $sig(n)$ 为原始音频信号，$sig^{\mathrm{w}}(n)$ 为含水印的音频信号，α 为水印嵌入强度系数. 由该规则可以看出，水印的透明性和鲁棒性由 α 值决定. 因在扩频处理中均方根值的误差小于平均值的误差，因此本文

算法利用音频信号的均方根值来调整参数 α. 假设水印信号嵌入到一段长度为 N 的音频信号 $x(n)$ 中，则该段音频信号的均方根值 $scont$ 可表示为 $scont=\sqrt{\frac{1}{N}\sum_{i=1}^{n}x_i^2}$. 根据该公式，水印嵌入强度系数 α 的计算公式可表示为 $\alpha=\begin{cases}0.005, & scont<0.05;\\ 0.01, & scont>0.1;\\ 0.1, & \text{otherwise.}\end{cases}$

5) 将所有音频段按照它们被分段的顺序重新组合成完整的音频信号.

2.2　音频水印的提取算法

音频水印的提取过程与水印的嵌入流程相反. 图 2 为水印信息的提取流程图，具体提取步骤如下：

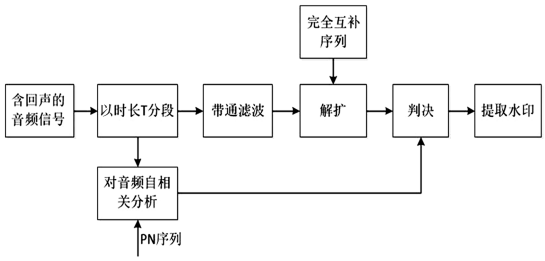

图 2　音频水印信息提取流程图

步骤 1　采用窗函数设计带通滤波器(FIR 滤波器). 过滤器的顺序是 127，函数选择汉明窗. 本段带通滤波器的主要参数为：中心频率为 4 900 Hz，起始频率为 3 400 Hz，截止频率为 6 400 Hz，滤波器带宽为3 000 Hz(大于发送端带通信号的带宽).

步骤 2　对上一步得到的基带信号进行去噪，提取水印序列.

步骤 3　将步骤 2 中得到的相关值与阈值进行比较，得到水印序列.

步骤 4　在时域内提取水印信息. 在嵌入隐藏信息的过程中，由于水印信息隐藏的音频信号具有与 PN 序列相同的分布特征，且回声核衰减系数的符号由 PN 序列控制，因此本文使用 PN 序列对每个音频段执行自相关检测.

3　实验结果与分析

3.1　实验环境

实验采用 3 种波形音频信号(分别标记为 1(轻音乐)、2(钢琴乐)和 3(摇滚乐))，每种波形音频信号时长为 3 min. 音频采样频率为 44.1 kHz，量化精度为 16 位. 使用 12 个长度的随机二进制序列作为水印序列，水印强度系数 α 为动态嵌入强度.

3.2　透明性验证

音频信噪比(SNR)是评价噪声信号在原始音频信号中嵌入强度的一种有效方法. 通常，信噪比越大，人耳感知噪声的能力越困难，即表明水印信息的隐藏效果越好，水印信息的透明性越高. 图 3 为原始音频和含水印信息音频图. 由图 3 可以看出，二者之间虽然存在差异，但并不明显，这表明本文水印算法具有较高的透明性.

表 1 为不同嵌入强度的信噪比. 由表 1 可以看出，本文算法的信噪比 SNR 随着嵌入强度的增大而逐渐减小，由此可知人耳对音频载体嵌入噪声的感知随嵌入强度的增大逐渐减小，这表明本文算法得到的音频水印具有较好的不可感知性，音频效果良好.

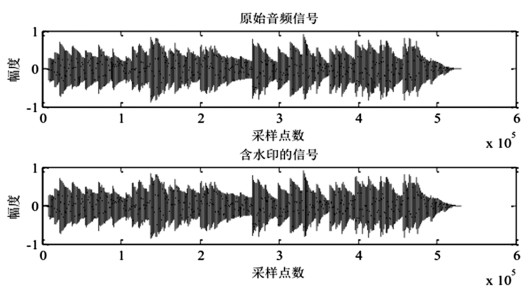

图 3　原始音频图和含水印音频图

表 1　不同嵌入强度的信噪比

嵌入强度	SNR/dB
0.001	51.49
0.003	46.65
0.005	41.98
0.007	37.46
0.009	33.97
0.010	32.12
0.020	25.01

3.3 鲁棒性验证

为了测试本文算法的鲁棒性,对音频载体分别进行滤波、剪切、噪声、重采样和 MP3 压缩攻击实验,具体的实验结果如下.

1)滤波攻击. 采用截止频率为 1 kHz 的巴特沃斯低通滤波器对音频信号进行滤波. 滤波阶数分别为 1、2 和 4 时,(滤波攻击的结果(相似度(NC)与比特出错概率(BEP))如表 2—表 4 所示. 由表 2—表 4 可以看出,本文算法能够有效地抵抗低通滤波攻击,说明本文算法对低通滤波攻击具有很强的鲁棒性.

表 2 滤波攻击轻音乐(1 号)的结果

	低通滤波阶数		
	1	2	4
NC	0.994	0.987	0.968
BER/%	0.06	0.14	0.35

表 3 滤波攻击钢琴乐(2 号)的结果

	低通滤波阶数		
	1	2	4
NC	0.996	0.984	0.961
BER/%	0.07	0.15	0.38

表 4 滤波攻击摇滚乐(3 号)的结果

	低通滤波阶数		
	1	2	4
NC	0.997	0.982	0.962
BER/%	0.09	0.16	0.41

2)剪切攻击. 用回声隐藏信息剪切 30、60 s 和 90 s 的音频信号,实验结果如表 5—表 7 所示. 由表 5—表 7 可以看出,音频经过 3 个时长的剪切攻击后,仍能提取出较为准确的水印信息,这说明本文算法对剪切攻击具有较强的鲁棒性.

表 5 剪切攻击轻音乐(1 号)的结果

	剪切时长/ms		
	30	60	90
NC	0.891	0.918	0.995
BER/%	0.41	0.25	0.05

表 6 剪切攻击钢琴乐(2 号)的结果

	剪切时长/ms		
	30	60	90
NC	0.855	0.899	0.914
BER/%	0.51	0.43	0.14

表 7 剪切攻击摇滚乐(3 号)的结果

	剪切时长/ms		
	30	60	90
NC	0.821	0.875	0.905
BER/%	0.65	0.48	0.21

3)噪声攻击. 在音频信号中引入 100、300 dB 和 900 dB 的白噪声,实验结果如表 8—表 10 所示. 由表 8—表 10 可以看出,本文算法能够有效地抵抗噪声攻击,说明本文算法对噪声攻击具有很强的鲁棒性.

表 8 噪声攻击轻音乐(1 号)的结果

	噪声/dB		
	100	300	900
NC	0.987	0.956	0.935
BER/%	0.06	0.08	0.11

表 9 噪声攻击钢琴乐(2 号)的结果

	噪声/dB		
	100	300	900
NC	0.995	0.974	0.924
BER/%	0.03	0.05	0.09

表 10 噪声攻击摇滚乐(3 号)的结果

	噪声/dB		
	100	300	900
NC	0.993	0.954	0.917
BER/%	0.05	0.09	0.12

4)重采样攻击. 对音频信号分别进行 22.05、16 kHz 和 11.025 kHz 的采样攻击,结果如表 11—表 13 所示. 由表 11—表 13 可以看出,本文算法能够有效地抵抗重采样攻击,说明本文算法对重采样攻击具有很强的鲁棒性.

表 11 重采样攻击轻音乐(1 号)的结果

	采样率/kHz		
	22.05	16	11.025
NC	0.985	0.974	0.925
BER/%	0.06	0.15	0.38

表 12 重采样攻击钢琴乐(2 号)的结果

	采样率/kHz		
	22.05	16	11.025
NC	0.954	0.921	0.956
BER/%	0.08	0.18	0.41

表 13 重采样攻击摇滚乐(3 号)的结果

	采样率/kHz		
	22.05	16	11.025
NC	0.956	0.947	0.943
BER/%	0.05	0.24	0.4

5)MP3 压缩攻击. 对音频信号进行压缩编码(压缩比为 12.5 ∶ 1),解码后提取隐藏信息,结果如表 14 所示. 由表 14 可以看出,本文算法可较好地抵抗 MP3 压缩攻击,说明本文算法对 MP3 压缩攻击也具有较好的鲁棒性.

表 14 MP3 压缩 3 种类型音乐的结果

	音乐类型		
	轻音乐	钢琴乐	摇滚乐
NC	0.854	0.881	0.886
BER/%	0.18	0.09	0.11

3.4　算法性能比较

为验证本文算法在抵抗攻击方面的性能，将本文算法分别与文献[9]（语音内容认证算法）、文献[10]（线性预测倒谱滤波算法）和文献[11]（基于音频特征的鲁棒水印算法）的算法进行对比，检测结果如表 15 所示. 从表 15 可以看出：本文算法在抵抗高斯噪声、滤波、重采样攻击时其性能均优于文献[9-11]的算法；在抵抗剪切和 MP3 压缩攻击时，其性能均优于文献[9]和文献[11]的算法，但略低于文献[10]的算法. 综合来看，本文算法不仅明显优于文献[9]和文献[11]的算法，而且略优于文献[10]的算法.

表 15　不同算法的性能

算法	不同攻击类型的 NC 值					
	未攻击	高斯噪声	滤波	剪切	MP3 压缩	重采样
文献[9]算法	0.911	0.875	0.857	0.844	0.646	0.873
文献[10]算法	0.975	0.958	0.977	0.945	0.913	0.922
文献[11]算法	0.955	0.968	0.977	0.887	0.844	0.955
本文算法	0.996	0.975	0.988	0.917	0.882	0.984

4　结论

研究表明，本文提出的基于 PN 序列和完全互补码的数字音频水印算法对水印信息的检测率超过 98%，对高斯噪声、滤波和重采样等攻击的水印信息检测率超过 90%，且在抵抗高斯噪声、滤波、重采样、剪切和 MP3 压缩攻击时其性能均优于文献[9]和文献[11]的算法，在抵抗高斯噪、滤波、重采样攻击时，其性能优于文献[10]的算法，说明本文算法的鲁棒性较强，具有很好的应用价值. 本文算法在抵抗剪切和 MP3 压缩攻击时的效果略低于文献[10]的算法，因此在今后的研究中，我们将探讨利用 QR 分解和小波变换的方法来进一步提高本文算法对抗剪切攻击和 MP3 压缩攻击的鲁棒性.

参考文献：

[1] SYHYNDEL P. Watermarking schemes evaluation[J]. IEEE Signal Processing Leters, 2011,17(5):58-64.

[2] 高雪丽，李德. 基于边缘特征的动画零水印算法[J]. 延边大学学报(自然科学版)，2017,43(4):353-358.

[3] 杨振仁，李德. 基于图像归一化与扩频技术的 DTC 零水印算法[J]. 延边大学学报(自然科学版)，2017,43(2):179-183.

[4] HU K, KUNDUR D, HATZINAKOS D. Statistical invisibility for collusion resistant digital video watermarking [J]. IEEE Trans on Multimedia, 2015,17(1):43-51.

[5] BHAT H, THENG Y. Robust and inaudible multi-echo audio watermarking[J]. IEEE Pacific Rim Conference on Multimedia, 2002,15(3):713-720.

[6] LEI M, CHENG M, LIU B. An audio zero-watermark scheme based on energy comparing[J]. China Communications, 2011,11(7):110-116.

[7] FALLAHPOUR M, MEGÍAS D. Secure logarithmic audio watermarking scheme based on the human auditory system[J]. Multimedia System, 2014,20(2):155-164.

[8] KHALDI K, BOUDRAA A O. Audio watermarking via EMD[J]. IEEE Transactions on Audio Speech and Language Processing, 2013,21(3):675-680.

[9] 钱清. 基于数字水印的语音内容认证算法研究[D]. 成都：西南交通大学，2018:23-26.

[10] 吕冰. 基于线性预测倒谱滤波的音频水印检测技术研究[D]. 武汉：华中师范大学，2017:19-22.

[11] 马志伟. 音频数字水印算法研究与优化[D]. 哈尔滨：哈尔滨理工大学，2017:22-26.