

抗重录音的溯源追踪音频水印算法

杨艳聪¹ 王锐¹ 秦兴红² 刘正辉^{1,2*}

¹(信阳师范学院计算机与信息技术学院 河南 信阳 464000)

²(深圳大学深圳市媒体信息内容安全重点实验室 广东 深圳 518060)

摘要 录音设备的普及,使重录音成为一种去除音频信号中嵌入水印的普遍攻击方式。为抵抗重录音攻击提出一种新的用于溯源追踪的鲁棒音频水印算法,给出系数对数均值(Logarithmic Mean of Coefficient, LMC)的定义,分析系数对数均值的性质。提出基于DCT的中频域水印嵌入方法。同步信息由帧号生成,和由版权生成的水印信息一起嵌入到载体信号中。仿真实验结果表明,提出的算法具有较好的不可听性、鲁棒性,同时能够有效抵抗一定程度的重录音攻击。

关键词 数字水印 版权保护 重录音攻击 溯源追踪

中图分类号 P237 TP3 文献标志码 A DOI:10.3969/j.issn.1000-386x.2024.08.053

AUDIO WATERMARKING ALGORITHM FOR TRACING THE RE-RECORDED AUDIO SOURCE

Yang Yancong¹ Wang Rui¹ Qin Xinghong² Liu Zhenghui^{1,2*}

¹(College of Computer and Information Technology, Xinyang Normal University, Xinyang 464000, Henan, China)

²(Shenzhen Key Laboratory of Media Security, Shenzhen University, Shenzhen 518060, Guangdong, China)

Abstract With the popularity of recording equipment, re-recording becomes a popular attack method to remove watermarks embedded in audio signals. In order to resist re-recording attacks, a new robust audio watermarking algorithm for source tracking is proposed. The definition of logarithmic mean of coefficient (LMC) was given, and the properties of the logarithm mean of coefficients were analyzed. A DCT-based middle frequency domain watermark embedding method was proposed. The synchronization information was generated after the frame number was binarized, and the frame number and the watermark were embedded in the carrier signal audio frame. The simulation results show that the proposed algorithm has good imperceptibility and robustness, and can effectively resist a certain degree of re-recording attacks.

Keywords Digital watermarking Copyright protection Re-recording attacks Source tracing

0 引言

数字信号处理技术的进一步发展和5G时代网络传输速率的提高,为高清音频的传输和共享提供技术支持。在此情况下,在线音频行业一线品牌喜马拉雅的付费内容已从单一的知识品类发展到了有声书、广播剧在内的知识加娱乐领域。诸如此类贴合人们生活

的音频的内容、产品将更加丰富。然而,随着大量音频作品的涌现、传输带宽的增加,音频播放、录音设备的普及,对受版权保护音频的录制、篡改以及肆意传播等问题将更加凸显。如未经版权所有者的允许,从某网络平台非法复制、经空气信道重录音来获取未授权的音频内容;对音频内容进行篡改、去除原音频作品中的版权信息,再编辑伪造成自己所有的音频作品并重新发布、销售。当下,如何对越来越多、内容越来越丰富

的音频作品进行保护,对音频非法传播行为的追踪溯源成为公共媒体和司法鉴定机构急需解决的技术难题。

数字音频水印是利用音频中存在的冗余和人类听觉系统的一些不敏感特性,在不降低听觉质量的前提下,通过一定的算法将特殊用途的信息嵌入到载体信号中。嵌入的信息可以保护音频作品的版权,证明其真实性 and 完整性,追踪盗版等非法行为。数字水印技术研究至今,已有相当丰硕的研究成果^[1-4]。用于版权保护的音频水印算法已经做到能够抵抗常见信号处理操作(重采样、重量化、MP3 压缩等),也能够抵抗一定程度的去同步攻击(变速、变调、抖动攻击等)^[5-7]。文献[6]提出了一种抗去同步攻击的音频水印算法,嵌入端首先进行水印信息嵌入,然后将同步码嵌入到含水印的信号中。提取端通过同步码的位置变化来确定攻击是否对含水印信号进行缩放,并确定缩放系数。依据缩放系数对含水印信号再处理,然后提取水印内容,来减小攻击带来的影响,提高水印提取正确率。然而,对于由扬声器播放,经空气信道传播,被录音笔等录音设备重录音的信号,由于空气信道各因素的干扰,使同步码提取的准确率存在大幅度的下降,进一步地影响了水印的准确提取。

考虑到抗重录音水印技术的实际需求,以及当前鲁棒音频水印技术存在重录音攻击可以去除水印信息的风险,本文以数字音频为研究对象,提出一种用于溯源追踪的鲁棒音频水印技术。本文主要工作包括两个方面:(1) 提出了一种对重录音攻击具有一定鲁棒性的音频特征,系数对数均值(Logarithmic Mean of Coefficient, LMC),为音频水印算法提供了较为稳定的嵌入域;(2) 给出了基于 LMC 特征的信息嵌入方法,构造了基于 LMC 特征的抗重录音攻击的音频水印算法。本文将帧号和版权信息一起嵌入到载体信号中,提取端由帧号同步含水印的音频帧,然后提取版权信息来证实音频内容的所有者,并进行溯源追踪。和现有的部分鲁棒音频水印算法相比,本文算法在对常见信号处理操作和去同步攻击保持鲁棒性的同时,具有一定的抗重录音攻击能力,为当前环境下经空气信道传播的音频内容的保护积累了一定的技术经验。

1 LMC 特征

1.1 特征定义

对于含有 N 个样本的音频信号 $A = \{a_n \mid 1 \leq n \leq N\}$,由式(1)计算其 DCT 系数,记为 $C = \{c_n \mid 1 \leq n \leq N\}$ 。

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

式中: $k=1$ 时, $w(k) = 1/\sqrt{N}$; $2 \leq k \leq N$ 时, $w(k) = \sqrt{2/N}$ 。

从 DCT 系数 C 中选中频系数 $C_{\text{med}} = \{c_{n_1} \mid 1 < n_1 < N_{\text{med}}\}$,其中 $1 < N_{\text{med}} < N$ 。由式(2)计算音频信号 A 的 LMC 特征,记为 F 。

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

式(2)中,且保证 $|c_{n_1}|/\alpha < 1$ 。假设音频信号 DCT 中频系数幅值 $|c_{n_1}| < \alpha$ (常不小于 10,本文中取 $\alpha = 12$)。依据对数性质,当 $1 < |c_{n_1}| < \alpha$ 时, $\log_2 |c_{n_1}| > 0$; 当 $0 < |c_{n_1}| < 1$ 时, $\log_2 |c_{n_1}| < 0$ 。如果对 $\log_2 |c_{n_1}|$ ($1 < n_1 < N_{\text{med}}$) 进行求和,如 $\sum_{n_1=1}^{N_{\text{med}}} \log_2 |c_{n_1}|$,则其中的大于 0 的项和小于 0 的项会相互抵消,降低了特征对音频信号特性的表征能力。于是,本文中让参与求和每一项都小于 0,然后求幅值,以此方法增强参与计算的每个系数和特征的关联,增强特征的表征力。

由于 DCT 系数存在取值为 0 的情况,为了使式(2)有意义,若 $c_{n_1} = 0$,先将 c_{n_1} 赋值为非零的较小的值,然后再参与式(2)的计算。如本文中当 $c_{n_1} = 0$ 时,将 c_{n_1} 赋值为 0.001。由式(2)可得,中频系数幅值越大,特征 F 则越小;反之,特征 F 则越大。

1.2 鲁棒性分析

音频信号经 D/A 转换为模拟信号,放大后经扬声器播放,以声波的方式在空气信道中发散传播。由于空气信道的噪声干扰、吸收等影响,使重录音信号和原始信号相比存在较大的能量损失。音频信号能量主要集中在低频部分,也可以说,不同信号的区别主要在于低频部分的区别。于是,音频信号重录音后低频部分改变较大。

为了更清晰地展示实验结果,随机选取一段 12 000 个样本点的音频信号,平均分为 4 帧,如图 1 所示。对此音频进行重录音攻击,重录音的信号如图 2 所示。然后对原始信号和重录音信号各帧进行 DCT(对应信号 DCT 系数如图 3、图 4 所示)。表 1 给出了各帧信号 DCT 低频系数和中频系数的能量,以及不同频率能量变化率 R (原始能量/重录音能量)。从表 1 所示结果可得,对重录音信号而言,低频部分和原始信号相比能量存在较大的损失。虽然中频部分能量也有一定程度的损失,然而和低频相比,损失的能量相对较小,更接近原始信号。于是,本文选取中频系数来定义 LMC 特

征。下面我们实验测试 LMC 特征的抗重录音攻击能力。

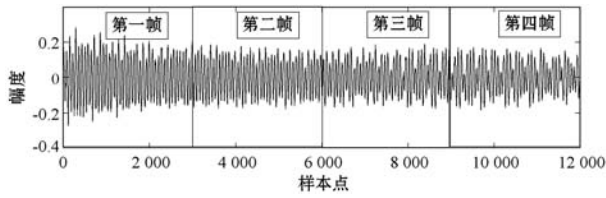


图1 随机选取 12 000 个样本点的音频信号

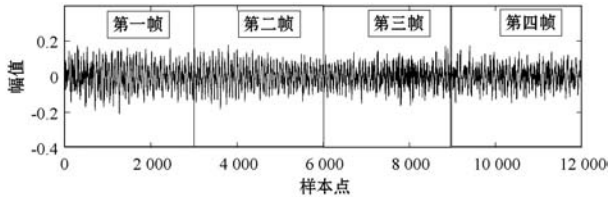


图2 对图1所示信号重录音攻击后的信号

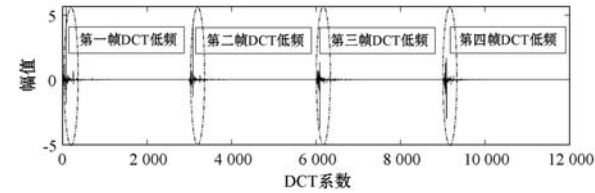


图3 原始信号不同帧的 DCT 系数

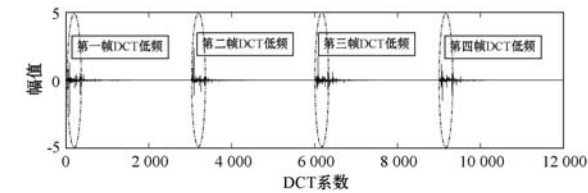


图4 重录音信号不同帧的 DCT 系数

表1 原始信号和重录音信号 DCT 低频、中频能量对比

音频信号		原始/重录音			
		第一帧	第二帧	第三帧	第四帧
低频		45/11	26/11	24/17	32/6
	R	4.1	2.4	1.4	5.3
中频		3/3	7/6	17/12	24/21
	R	1	1.2	1.4	1.1

音频信号常见的采样频率为 44.1 kHz, 这里我们随机选取一段长为 L , 采样频率为 44.1 kHz 的音频信号, 记为 A , 如图 5 所示。

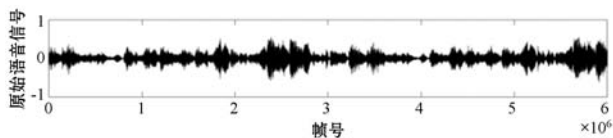


图5 原始音频信号

步骤 1 将 A 分为 P 帧, 第 i 帧记为 A_i , 每帧长为 N 。

步骤 2 对 A_i 进行 DCT, 取中频系数, 由式(2)计算 A_i 的 LMC 特征 F_i 。

对选取的音频信号进行重录音攻击, 重录音后的信号记为 A' , 如图 6 所示。重复上述步骤 1 和步骤 2, 将 A' 分帧, 并计算各帧的 LMC 特征 F'_i 。图 7 给出了音频信号 A 和重录音攻击后信号 A' 各帧的 LMC 特征的对比情况, 这里每帧长设置为 12 000。

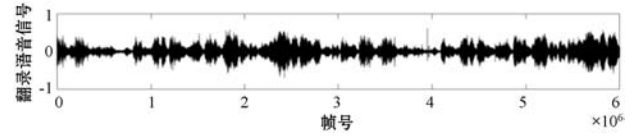


图6 重录音信号

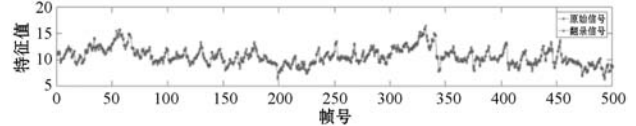


图7 原始信号和重录音信号的 LMC 特征

可以看出, 重录音后各音频帧的 LMC 特征和原始信号相比变化较小, 一定程度上说明了该特征对重录音攻击的鲁棒性。为了测试 LMC 特征对其他音频信号的适用性, 我们选取 240 段不同的音频信号, 并对选取的音频信号进行重录音攻击 (SONY 录音笔, 型号 PCM-D100)。按照上述方法, 将各音频信号分为 500 帧。计算 240 段音频信号各帧攻击前后 LMC 特征的差值。图 8 给出了攻击前后各帧 LMC 特征差值的统计均值。易得, 攻击前后 LMC 特征变化的最大幅度大约为 0.5。和 LMC 特征相比, 变化幅度相对较小 (变化幅度约为图 7 中 LMC 特征值的 1/20)。将水印嵌入在 LMC 特征中, 该特征的鲁棒性为从重录音信号中提取准确的水印信息提供有力的支撑。

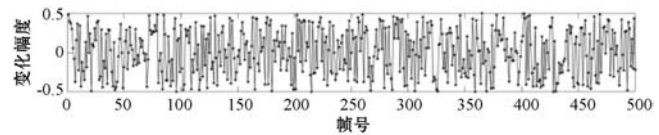


图8 原始信号重录音后 LMC 特征的变化幅度

2 水印方案

记 $W_i = \{w_1, w_2, \dots, w_M\}$ 为第 i 帧要嵌入的信息, $w_m \in \{0, 1\}$, $1 \leq m \leq M$ 。 W_i 由三部分构成, $W_i^{(1)} = \{w_m \mid 1 \leq m \leq M_1\}$, $W_i^{(2)} = \{w_m \mid M_1 + 1 \leq m \leq 2M_1\}$, $W_i^{(3)} = \{w_m \mid 2M_1 + 1 \leq m \leq M\}$ 。 $W_i^{(1)}$ 和 $W_i^{(2)}$ 为由帧号二值化后生成的同步信息, $W_i^{(1)} = W_i^{(2)}$ 。 $W_i^{(3)}$ 为由版权生成的水印 (或部分) 信息。水印检测端, 从含水印信号开始位置, 将第一帧含水印的信号分为三部分。从第一、第二部分中提取同步信息, 若从两段信号中提取的同步信息相同, 则认为该帧是完整的, 然后从第三段中提取水印信息。

2.1 水印嵌入

记音频信号为 $A = \{a_l \mid 1 \leq l \leq L\}$, 其中 a_l 表示第 l 个样本点, L 表示信号 A 的长度。

步骤 1 将音频信号 A 等分为 P 帧, 第 i 帧记为 A_i 。每帧长为 $N, N = L/P$ 。将 A_i 等分为 M 段, 分别为 $A_{i,1}, A_{i,2}, \dots, A_{i,M}$, 每段长为 N_1 。

步骤 2 采用混沌映射的方法将每段信号置乱, 置乱后的信号记为 $S_{i,1}, S_{i,2}, \dots, S_{i,M}$ 。方法如下:

(1) 选择 Logistic 映射来生成伪随机序列, 这里采用文献[8]中给出的 Logistic 映射, 生成的序列记为 $X = \{x_l \mid l = 1, 2, \dots, N_1\}$, 其中 x_l 可由式(3)来得到。

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

式中: $x_0 = k$ 是 Logistic 映射的初值, 作为水印系统的密钥。

(2) 假设 $A_{i,1} = \{a_l \mid 1 \leq l \leq N_1\}$ 。将伪随机序列中的元素进行升序排列, 升序排列后的地址索引序列可由式(4)得到, 置乱后的信号记为 $S_{i,1}$, 则 $S_{i,1} = \{a_{h(l)} \mid 1 \leq l \leq N_1\}$ 。

$$x_{h(l)} = \text{ascend}(x_l) \quad l = 1, 2, \dots, N_1 \quad (4)$$

采用同样的方法, 将其他段的音频信号置乱。

步骤 3 对置乱后的信号 $S_{i,1}$ 进行 DCT, 得到的系数为 $C_{i,1} = \{c_1, c_2, \dots, c_{N_1}\}$ 。选择中频系数并由式(2)计算 $S_{i,1}$ 的 LMC 特征, 记为 $F_{i,1}$ 。

步骤 4 采用式(5)的方法, 根据第一个水印信息 w_1 量化 $F_{i,1}$ 。

$$QF_{i,1} = Q_{\Delta}^w(F_{i,1}) = \text{round}[(F_{i,1} + d(w))/\Delta] \times \Delta - d(w) \quad (5)$$

式中: $QF_{i,1}$ 表示量化后的 LMC 特征; “ $\text{round}(\cdot)$ ” 表示向下取整; Δ 表示量化步长; w 为水印比特信息, 其值为 0 或 1; $d(w)$ 是与水印信息 w 对应的抖动量。

步骤 5 特征 $F_{i,1}$ 量化之后, 由式(6)来得到和量化后的特征 $QF_{i,1}$ 相对应的 DCT 中频系数(即量化前后中频系数的对应关系)。

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

式中: c_n 为量化前的中频系数; c_n^* 为量化后的中频系数; $\text{sign}(c_n)$ 为符号函数, 当 $c_n \geq 0$ 时, $\text{sign}(c_n) = 1$; 当 $c_n < 0$ 时, $\text{sign}(c_n) = -1$ 。

步骤 6 合并量化后的 DCT 中频系数和其他低频、高频系数进行逆 DCT, 并反置乱, 即可得到嵌入 w_1 的音频信号。

重复上述步骤, 将其他水印信息 w_2, w_3, \dots, w_M 嵌入到 A_i 的其他各段 $A_{i,2}, A_{i,3}, \dots, A_{i,M}$ 中。嵌入方法如图 9 所示。

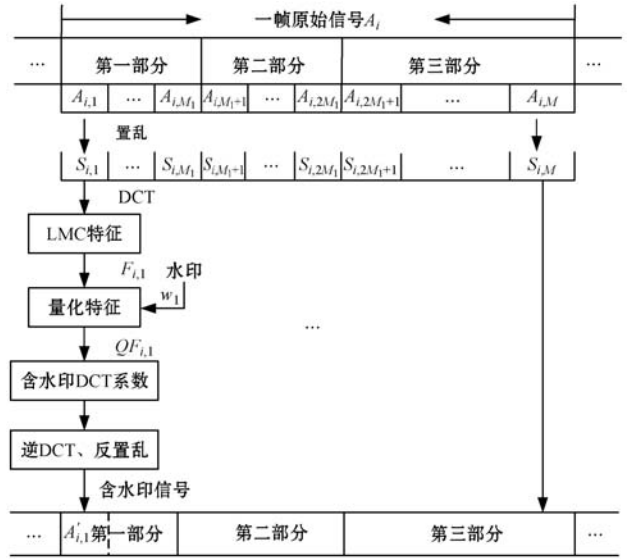


图 9 信息嵌入过程框

2.2 水印提取

水印提取过程可视为水印嵌入过程的逆过程。这里假设含水印信号为 A' , 长度为 L' 。对含水印信号 A' 进行分帧, 记第 i 帧信号为 A'_i 。将 A'_i 等分为 M 段, $A'_{i,1}, A'_{i,2}, \dots, A'_{i,M}$ 。下面以 $A'_{i,1}$ 中水印提取为例, 来介绍水印提取方法。

步骤 1 采用式(4)的方法, 对 $A'_{i,1}$ 中的样本点置乱, 并对置乱后的信号进行 DCT。

步骤 2 选择 DCT 中频系数, 由式(2)计算特征, 记为 $F'_{i,1}$ 。

步骤 3 由式(7)提取 $A'_{i,1}$ 中嵌入的信息, 记为 w'_1 。

$$w'_1 = \arg \min_{l \in \{0,1\}} (F'_{i,1} - F'_{i,1}[l]) \quad (7)$$

式中: $F'_{i,1}$ 表示接收到的含水印信号, $F'_{i,1}[0]$ 和 $F'_{i,1}[1]$ 分别表示采用 $Q_{\Delta}^0(\cdot)$ 和 $Q_{\Delta}^1(\cdot)$ 对接收到的数据 $F'_{i,1}$ 进行抖动调制所得到的量化值。

重复上述步骤 1 到步骤 3, 提取 A'_i 中各段嵌入的信息 w'_1, w'_2, \dots, w'_M 。记 $W'_i = \{w'_1, w'_2, \dots, w'_M\}$, 将 W'_i 分为三段, 分别记为 $W'_{1,i} = \{w'_m \mid 1 \leq m \leq M_1\}$, $W'_{2,i} = \{w'_m \mid M_1 + 1 \leq m \leq 2M_1\}$, $W'_{3,i} = \{w'_m \mid 2M_1 + 1 \leq m \leq M\}$, 即 $W'_i = [W'_{1,i} \parallel W'_{2,i} \parallel W'_{3,i}]$ 。

根据划分的三段信息, 首先判断该帧是否为同步

的音频帧。若 $\sum_{m=1}^{M_1} w'_m \oplus w'_{m+M_1} = 0$, 根据本文嵌入算

法, 说明该帧能够准确提取同步信息; 若 $\sum_{m=1}^{M_1} w'_m \oplus w'_{m+M_1} \neq 0$, 移动样本, 将紧接着的连续的 N 个样本作为新的一帧, 并根据上述方法确定是否能从新的一帧中提取同步信息, 直到满足 $\sum_{m=1}^{M_1} w'_m \oplus w'_{m+M_1} = 0$ 为止, 则该帧即为确定的新的同步音频帧。在检测到同步的

音频帧之后,从第三段中提取水印信息。提取方法如图 10 所示。

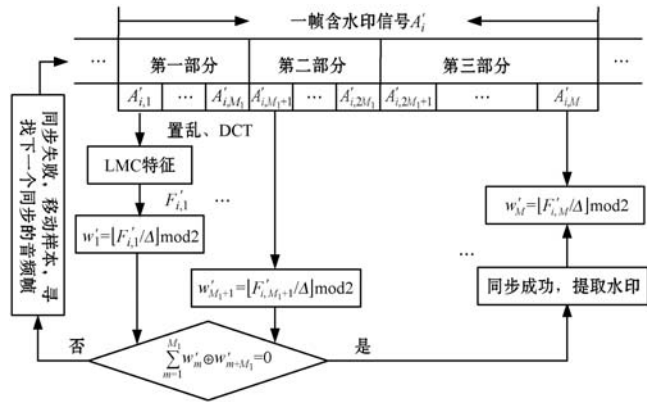


图 10 提取和同步水印的过程

3 性能测试

为了测试本文算法的性能,下面采用 MATLAB 进行实验仿真。从样本库随机选取 200 段 16 位量化、采样频率为 44.1 kHz 的不同类型的音频信号。首先对选取的音频信号进行水印嵌入,然后采用不同录音设备对含水印信号进行重录音攻击(三种不同录音设备:SONY 录音笔 PCM-D100,华为 P20 和 iPhone 6s)。算法涉及到的参数取值分别为:帧长 $L = 600\ 000$, $N = 12\ 000$, $P = 50$, $M = 10$, $M_1 = 3$, $\Delta = 0.8$, $\alpha = 12$ 。

3.1 不可听性

水印系统常用不可听性来衡量水印的嵌入对原信号听觉质量的影响程度。常用的评价方法有主观和客观两种。主观方法是将原音频信号及含水印的信号提供给一组听众,由听众根据主观感觉来区分两个信号之间的差别,并按照主观区分度(Subjective Difference Grades, SDG)来打分,打分的平均值作为主观听觉质量测试的结果。SDG 各等级分值的含义如表 2 所示^[9],易得,SDG 值越大,水印的不可感知性越好。

表 2 SDG 评价标准

SDG	描述
0.0	不可感觉
-1.0	可感觉但不刺耳
-2.0	轻微刺耳
-3.0	刺耳
-4.0	非常刺耳

客观评价常见的方法有信噪比(Signal to Noise Ratio, SNR),是一种把嵌入的水印信息作为添加到原始音频中的噪声。计算信噪比,用信噪比来衡量水印

信息的嵌入对原信号的改变程度。SNR 的计算式如式(8)所示^[9]。

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

式中: $a(l)$ 表示原始信号; $a'(l)$ 表示含水印的信号; N 为信号的长度,在不影响听觉的情况下,SNR 值应在 20 dB 以上。

本文对选取的 200 段测试音频信号,计算其 SNR 均值和 SDG 均值(由 14 位听众现场打分统计所得)。所得结果为 SNR 均值为 30.08,SDG 均值 -0.59,满足水印不可听性的性能指标(SNR 值大于 20,SDG 值大于 -1)。于是,本文算法水印的嵌入不影响原信号的听觉质量。

3.2 鲁棒性

本文算法旨在解决音频信号经空气信道传播,听众重录音后非法传播的溯源追踪问题。从重录音信号中能否准确地提取水印信息,关系到本文算法的性能和适用效果。下面我们采用误码率(Bit Error Rates, BER)测试含水印信号被攻击后水印提取的准确率。BER 的定义如式(9)所示。

$$B_{ER} = \frac{1}{M} \sum_{m=1}^M w(m) \oplus w'(m) \quad (9)$$

式中: M 表示嵌入水印的长度; $w(m)$ 和 $w'(m)$ 分别表示原始水印信息和提取的水印信息; \oplus 表示异或操作。易得,BER 值越小,表明提取水印的错误率越低,算法的鲁棒性越好,抗攻击能力越强。

此部分鲁棒性的测试主要包括含水印信号经过信号处理操作、部分去同步攻击,以及在重录音攻击基础上,再进行信号处理和去同步攻击水印提取的 BER 值。

(1) 信号处理和去同步攻击的鲁棒性。通常,去同步攻击会一定程度上拉伸或缩短含水印信号的长度。为了从去同步攻击的含水印信号中尽可能地准确提取水印信息,水印提取端,本文尝试多种不同的长度,如分别采用 $90\% \times L/P$, L/P , $110\% \times L/P$, L'/P (L 和 L' 的含义如第 2 部分所示)进行分帧,以水印提取准确率最高的分帧作为有效的分帧长度。

含水印信号进行信号处理和去同步攻击的 BER 值如表 3 所示。信号处理操作包括 MP3 压缩(64 kbit/s 和 128 kbit/s)、重采样(44.1 kHz \rightarrow 22.05 kHz \rightarrow 44.1 kHz)。去同步攻击包括抖动(Jittering)攻击,所选参数分别为 1/100 和 1/1 000。这里 1/100 表示每隔 100 个样本删除一个样本点;1/1 000 表示每隔 1 000 个样本删除一个样本点。

表 3 本文算法和文献[10-11]在不同信号处理和去同步攻击后水印提取的 BER 值(%)

攻击类型		文献[10]	文献[11]	本文
MP3 压缩	64 kbit/s	5	6	1
	128 kbit/s	3	3	0
重采样	44.1→22.05→ 44.1 kHz	8	8	0
抖动 Jittering	1/100	7	8	0
	1/1 000	2	3	0
变速 Time-scaling	90%	13	15	7
	110%	12	13	5

从表 3 所示结果可得,本文算法生成的含水印信号在 128 kbit/s 的 MP3 压缩、重采样,以及抖动攻击后,水印能够完全准确地被提取。对于其他类型的信号处理操作以及去同步攻击,虽水印不能被完全准确地被提取,存在一定的错误提取的情况,但提取错误的信息数量小于文献[10-11]所给的算法。这表明在信号处理操作和去同步攻击中,与文献[10-11]相比本文算法有更好的抗攻击能力。

(2) 重录音攻击、信号处理和去同步攻击的鲁棒性。下面测试本文算法的抗重录音攻击的能力。首先对含水印音频进行重录音攻击,在此基础上,对重录音的信号进行如表 3 所示的信号处理和去同步攻击。攻击后信号的水印提取 BER 值如表 4 所示,其中,信号处理和去同步攻击的参数同表 3。

表 4 本文算法和文献[10-11]算法在重录音攻击和不同信号处理,以及去同步攻击后水印提取的 BER 值(%)

攻击类型		文献[10]	文献[11]	本文	
重 录 音	重录音	12	18	4	
	MP3 压缩	64 kbit/s	15	19	6
		128 kbit/s	12	18	4
	重采样	44.1→22.05→ 44.1 kHz	13	18	4
	抖动 Jittering	1/100	13	21	5
		1/1 000	12	19	5
	变速 Time-scaling	90%	18	23	11
		110%	15	22	9

由表 4 可得,对于重录音攻击,以及在重录音攻击的基础上进行的信号处理和去同步攻击,本文算法和文献[10-11]给出的算法的 BER 值均有不同程度的增加。相比而言,本文算法水印提取 BER 值的增加程度远小于文献[10-11],表明对于重录音攻击的情况,本文水印提取误码率更低,能够有更好的鲁棒性抵抗重录音攻击。

4 结 语

本文提出了一种鲁棒的抗重录音的音频水印算法,旨在解决音频信号经空气信道传播、重录音后的溯源追踪问题。首先定义了音频信号 LMC 特征,分析得到 LMC 特征对重录音攻击的鲁棒性。采用量化 LMC 特征的方法将帧号和水印一起嵌入到载体信号中。含水印信号被攻击后,帧号用来同步含水印的音频帧,然后提取水印信息来进行溯源追踪。和现有的用于版权保护的鲁棒音频水印算法相比,本文算法不仅提高了算法的抗去同步攻击能力,同时对于重录音攻击也具有一定的鲁棒性。

参 考 文 献

- [1] Hua G, Huang J W, Shi Y Q, et al. Twenty years of digital audio watermarking—A comprehensive review [J]. Signal Processing, 2016, 128(11): 222-242.
- [2] Nishimura A. Audio watermarking based on Subband amplitude modulation [J]. Acoustical Science and Technology, 2010, 31(5): 328-336.
- [3] Liu Z H, Wang H X. A novel speech content authentication algorithm based on Bessel-Fourier moments [J]. Digital Signal Processing, 2014, 24(1): 197-208.
- [4] 李晨,王可鑫,田丽华.基于声道低频能量比的 MP3 压缩域音频水印算法[J]. 计算机应用, 2018, 38(8): 2301-2305.
- [5] Hu P, Peng D Z, Yi Z, et al. Robust time-spread echo watermarking using characteristics of host signals [J]. Electronics Letters, 2016, 52(1): 5-6.
- [6] Xiang Y, Natgunanathan I, Guo S, et al. Patchwork-based audio watermarking method robust to de-synchronization attacks [J]. IEEE/ACM Transactions on Audio Speech & Language Processing, 2014, 22(9): 1413-1423.
- [7] Wang X Y, Ma T X, Niu P. A pseudo-Zernike moments based audio watermarking scheme robust against desynchronization attacks [J]. Computers and Electrical Engineering, 2011, 37(4): 425-443.
- [8] 王静,刘正辉,祁传达,等.一种抗特征分析替换攻击的数字语音取证算法[J]. 铁道学报, 2016, 38(6): 73-78.
- [9] Liu Z H, Zhang F, Wang J, et al. Authentication and recovery algorithm for speech signal based on digital watermarking [J]. Signal Processing, 2016, 123(1): 157-166.
- [10] Nadeau A, Sharm G. An audio watermark designed for efficient and robust resynchronization after analog playback [J]. IEEE Transactions on Information Forensics and Security, 2017, 12(6): 1393-1405.
- [11] Li J F, Wang H X, Wu T, et al. Norm ratio-based audio watermarking scheme in DWT domain [J]. Multimedia Tools and Applications, 2018, 77(12): 14481-14497.