



一种基于离散小波变换的抗重录音鲁棒音频水印算法

李艳丽, 祁传达

引用本文:

李艳丽, 祁传达. 一种基于离散小波变换的抗重录音鲁棒音频水印算法[J]. 信阳师范学院学报自然科学版, 2024, 37(4): 449–453. doi: 10.3969/j.issn.1003–0972.2024.04.005

LI Yanli, QI Chuanda. A Discrete Wavelet Transform–based Anti–replay Audio Watermarking Algorithm[J]. *Journal of Xinyang Normal University (Natural Science Edition)*, 2024, 37(4): 449–453. doi: 10.3969/j.issn.1003–0972.2024.04.005

在线阅读 View online: <https://doi.org/10.3969/j.issn.1003–0972.2024.04.005>

您可能感兴趣的其他文章

Articles you may be interested in

一种安全的抗去同步攻击数字语音取证算法

A Security Algorithm for Digital Speech Forensics Robust Against Desynchronization Attacks

信阳师范学院学报自然科学版, 2017, 30(1): 134–139. <https://doi.org/10.3969/j.issn.1003–0972.2017.01.030>

一种基于Arnold技术的彩色图像数字水印算法研究

Research of Digital Watermarking Algorithm of Color Image Watermark Embed Based on Arnold Transformation

信阳师范学院学报自然科学版, 2018, 31(3): 484–488. <https://doi.org/10.3969/j.issn.1003–0972.2018.03.028>

基于ICLBP的多特征自适应加权融合目标跟踪算法

A Target Tracking Algorithm of Adaptive Multiple Feature Weighted Fusion Based on ICLBP

信阳师范学院学报自然科学版, 2021, 34(3): 478–483. <https://doi.org/10.3969/j.issn.1003–0972.2021.03.023>

一种NG–DSL系统的低复杂度时频同步算法

Synchronization Algorithm about Time and Frequency with Low Complexity Based on NG–DSL System

信阳师范学院学报自然科学版, 2020, 33(2): 292–297. <https://doi.org/10.3969/j.issn.1003–0972.2020.02.019>

低维显式语义空间下的语义关联度计算方法

Semantic Relatedness Computation Method Under Low Dimensional Explicit Semantic Space

信阳师范学院学报自然科学版, 2019, 32(4): 675–682. <https://doi.org/10.3969/j.issn.1003–0972.2019.04.029>

DOI:10.3969/j.issn.1003-0972.2024.04.005

文章编号:1003-0972(2024)04-0449-05

一种基于离散小波变换的抗重录音鲁棒音频水印算法

李艳丽*, 祁传达

(信阳师范大学 计算机与信息技术学院, 河南 信阳 464000)

摘要:为了解决音频内容被重录音后非法传播和使用的问题,提出了一种抗重录音攻击的鲁棒音频水印算法。在对音频信号进行离散小波变换基础上,定义了数字音频特征 DWT-CLM (Discrete Wavelet Transform-Coefficient Logarithmic Mean),分析得出了该特征和音频信号 DWT 近似分量之间的关系,在此基础上提出了基于音频信号 DWT-CLM 特征的水印嵌入方法。实验分析结果表明,本算法具有较好的不可听性、抗信号处理的能力,尤其在对信号处理满足一定鲁棒性的同时,有一定的容忍重录音攻击的能力。

关键词:离散小波变换; 数字水印; 版权保护; 特征提取; 重录音攻击

中图分类号:TP309.7

文献标识码:A

开放科学(资源服务)标识码(OSID):



A Discrete Wavelet Transform-based Anti-replay Audio Watermarking Algorithm

LI Yanli*, QI Chuanda

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

Abstract:In order to solve the problem of illegal transmission and use of audio content after rerecording, a robust audio watermarking algorithm was proposed. Based on discrete wavelet transform, Discrete Wavelet Transform-Coefficient Logarithmic Mean (DWT-CLM) was defined, the relationship between DWT-CLM and the approximate components of DWT was analyzed, and a watermark embedding method was proposed based on DWT-CLM. Experimental results showed that the proposed algorithm had the good inaudibility and anti-signal processing ability, especially in the signal processing to meet a certain robustness, and a certain tolerance for re-recording attacks.

Key words:discrete wavelet transform(DWT); digital watermark; copyright protection; feature extraction; re-recording attack

0 引言

随着 5G 数字时代的到来,大量音频作品的涌现、传输带宽的增加、音频播放和录音设备的普及,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^[1-2]。例如,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。虽然数字水印技术^[3]

为音频内容版权保护以及内容取证提供了一定的技术支持,然而新时代下,如何既可以充分发挥 5G 网络的优势,又能有效地保护音频原创者合法权益,确保音频信息不被盗用,追踪用户的非法行为,依然是当下数字多媒体领域值得研究的问题^[4]。

数字音频水印技术是通过利用音频的冗余性和人类听觉系统的某些不敏感属性,在不损害听觉体验的情况下,采用特定算法将特定用途的信息嵌入到载体信号之中。在目前音频水印算法的研究领域,大部分研究都集中在增强对信号处理和去同

收稿日期:2022-12-01;修回日期:2023-05-14;*. 通信联系人,E-mail:wenyueli@126.com

基金项目:国家自然科学基金项目(61902085);河南省本科高校教学改革研究与实践项目(2022SYJXLX061)

作者简介:李艳丽(1982—),女,河南信阳人,讲师,硕士,主要从事密码学、深度学习等方向的研究。

引用格式:李艳丽,祁传达.一种基于离散小波变换的抗重录音鲁棒音频水印算法[J].信阳师范学院学报(自然科学版),2024,37(4):449-453.

LI Yanli, QI Chuanda. A Discrete Wavelet Transform-based Anti-replay Audio Watermarking Algorithm[J]. Journal of Xinyang Normal University (Natural Science Edition), 2024,37(4):449-453.

步攻击的抵抗力上,而对重录攻击的相关研究甚少。

鉴于现有鲁棒音频水印技术在重录攻击中去除水印信息的潜在风险,本文定义了一个数字音频特征 DWT-CLM (Discrete Wavelet Transform-Coefficient Logarithmic Mean),通过对常见信号的处理和对重录音攻击进行鲁棒性实验,分析得出了该特征和音频信号 DWT 近似分量之间的关系,在此基础上提出了基于音频信号 DWT-CLM 特征的水印嵌入方法。

1 理论基础

1.1 离散小波变换

离散小波变换 (Discrete Wavelet Transform, DWT)^[5-7] 是一种对信号进行时间-频率分析的方法。它在时频域内都具有体现信号局部特征的能力,是一种时间窗和频率窗都能够改变的时频域局部化分析方法,具有多分辨率的特点。离散小波变换具有良好的时频局部化特征,在数字信号处理中应用广泛。

记为 $A = \{a_l | 1 \leq l \leq L\}$ 为长为 L 的音频信号, a_l 表示第 l 个样本点。对 A 进行 D 级 DWT, 可得小波系数 $F_A^D, G_A^D, G_A^{D-1}, \dots, G_{pq_c}^1$ 。 F_A^D 即为小波变换的 D 级近似分量,集中了信号的主要能量; $G_{pq_c}^D, G_{pq_c}^{D-1}, \dots, G_{pq_c}^1$ 为小波变换的 $1 \sim D$ 级细节分量。为了更清晰地展示音频信号 A 的 DWT 结构,图 1 给出了 3 级 DWT 的结构图。

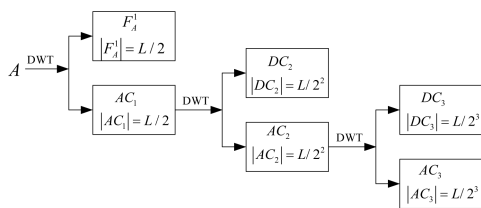


图 1 音频信号 A 的 3 级 DWT 结构图

Fig. 1 Level 3 DWT structure of audio signal A

1.2 特征定义

记 $A = \{a_l | 1 \leq l \leq L\}$ 为长为 L 的音频信号, 对 A 进行 D 级 DWT, 得到小波系数的近似分量 F_A^D , 以及细节分量 $G_A^D, G_A^{D-1}, \dots, G_{pq_c}^1$ 。由于信号 DWT 后能量主要集中在近似分量中, 以近似分量来构造本文的特征 DWT-CLM, 方法如式 (1) 所示, 这里记 $F_A^D = \{f_1, f_2, \dots, f_N\}$, 长度为 N , $N = L/2^D$ 。

$$D = \left| \sum_{n=1}^N \log_2(|f_n|/\lambda) \right| / N. \quad (1)$$

由式 (1) 定义所得, 近似分量 F_A^D 中, 包含幅值较大系数的个数越多, 得到的特征 F_A^D 越大; 反之, 得到的特征越小。

1.3 DWT-CLM 特征的性质

1.3.1 DWT-CLM 特征的鲁棒性

本文提取音频信号的 DWT-CLM 特征, 并将水印嵌入其中生成含水印信号。经过信号处理以及重录音攻击, 含水印信号中 DWT-CLM 特征应该保持稳定, 确保嵌入其中的水印能够准确地提取。若含水印音频被非法传播, 提取被传播音频的水印信息, 为证明版权所有者提供依据。下面通过实验验证音频信号 DWT-CLM 特征的鲁棒性。

从样本库中任意选取一段采样频率为 44.1 kHz 的音频信号 (含 L 个样本), 如图 2 所示。

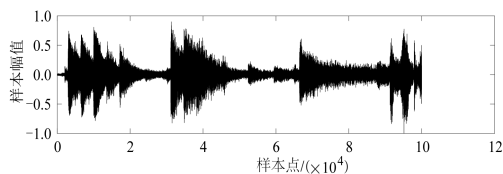


图 2 随机选取的音频信号

Fig. 2 Randomly selected audio signals

下面实验分析 DWT-CLM 在信号处理以及重录音攻击后的改变程度, 以此来测试该特征的鲁棒性。

第一步: 将 A 分为 P 帧 (P 取 100), 第 i 帧记为 A_i , 每帧长为 N 。

第二步: 对 A_i 进行 DWT (这里选择 3 级), 计算其 DWT-CLM 特征。

第三步: 对音频信号 A 进行常见的信号处理操作, 包括 64 kb/s 的 MP3 压缩和截止频率为 8 kHz 的低通滤波。图 3 给出了原始信号 DWT-CLM 特征以及信号处理后的特征。

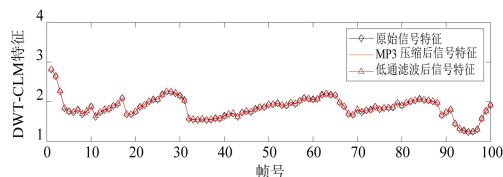


图 3 原始信号和信号处理后信号的 DWT-CLM 特征

Fig. 3 Characteristics of DWT-CLM in raw and processed signals

第四步: 对音频信号 A 进行重录音攻击, 攻击后的信号如图 4 所示。根据 DWT-CLM 特征的计算方法得到原始信号和重录音信号的 DWT-CLM 特征, 如图 5 所示。

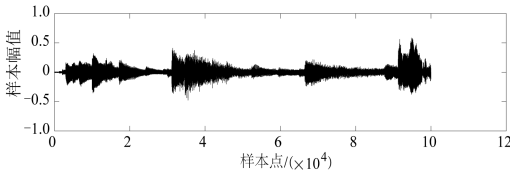


图4 重录音攻击后的音频信号

Fig. 4 Recording audio signals after the attack

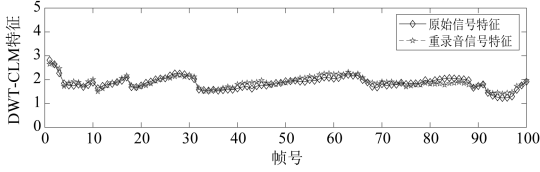


图5 原始信号和重录音攻击信号的DWT-CLM特征

Fig. 5 DWT-CLM features of raw and rerecorded attack signals

由测试对比结果可得,不同类型语音信号 FDCR 特征信号处理前后几乎保持不变,具有较好的鲁棒性。若采用量化 DWT-CLM 特征的方法来嵌入水印,可以保证含水印信号在信号处理后,以较高的概率正确地提取嵌入的信息。

1.3.2 DWT-CLM 特征和 DWT 近似分量的关系

为了更直观地得到含水印信息,下面分析量化后的特征和对应 DWT 近似分量之间的关系。

根据要嵌入的水印信息,假设量化后的特征为 D' ,由式(1)可得

$$D' = \left| \sum_{n=1}^N \log_2(f'_n/\lambda) \right| / N, \quad (2)$$

式中: f'_n 表示量化特征对应的 DWT 近似分量。假设 D, D' 和 f_n 为已知量,结合式(1)和式(2)可得

$$f'_n/\lambda = (f_n/\lambda)^{D'/D}, \quad (3)$$

式中: f'_n 为量化后的特征对应的 DWT 近似分量。由式(3)可得

$$f'_n = (f_n/\lambda)^{D'/D} \times \lambda. \quad (4)$$

式(4)给出了含水印信号 DWT 近似分量和原始信号 DWT 近似分量之间的关系。可以根据(4)式量化 DWT-CLM 特征的方法,得到含水印信号 DWT 近似分量,进一步逆 DWT 生成含水印信息。

基于此,本文给出了一种通过量化音频信号 DWT-CLM 特征的鲁棒音频水印算法。

2 本文算法

假设 $A = \{a_l \mid 1 \leq l \leq L\}$ 表示载体音频信号,符号 a_l 代表音频信号 A 的第 l 个样本点, L 代表音频信号 A 含有的样本点个数。

2.1 水印嵌入

(1)分帧、分段。将载体信号 A 按照等长的方法分帧,第 i 帧记为 A_i ,长为 N 。将 A_i 等分为前

后两段,分别记为 $A_{i,1}$ 和 $A_{i,2}$,长度均为 $N/2$ 。

(2)由式(1)计算第 i 帧前后两段信号 $A_{i,1}$ 和 $A_{i,2}$ 的 DWT-CLM 特征,分别记为 $D_{i,1}$ 和 $D_{i,2}$ 。

(3)假设 w 为要嵌入的水印信息,这里通过量化 $D_{i,1}$ 和 $D_{i,2}$ 来将 w 嵌入到 A_i 中。

①取 $QD_{i,1} = \lfloor D_{i,1}/\Delta \rfloor \times \Delta + \Delta/2$,同时记 $QD_{i,2} = \lfloor D_{i,2}/\Delta \rfloor \times \Delta + \Delta/2$,式中 Δ 为量化步长。

② 计算 $R_i = |QD_{i,1} - QD_{i,2}|$ (这里假设 $QD_{i,1} > QD_{i,2}$,对于 $QD_{i,1} \leq QD_{i,2}$ 的情况类似)。若 $(R_i/\Delta) \bmod 2 = w$,将第 i 帧前一段信号 $A_{i,1}$ 的 DWT-CLM 特征 $D_{i,1}$ 量化为 $Q_{i,1}$, $Q_{i,1} = QD_{i,1} + \Delta/2$;否则,将 $D_{i,1}$ 量化为 $Q_{i,1}$, $Q_{i,1} = QD_{i,1} - \Delta/2$ 。同时,将第 i 帧后一段信号 $A_{i,2}$ 的 DWT-CLM 特征 $D_{i,2}$ 量化为 $Q_{i,2}$, $Q_{i,2} = QD_{i,2}$ 。

③基于 1.3.2 部分的分析,由式(4)可以得到和量化后特征 $Q_{i,1}$ 相对应的 DWT 近似分量;类似地,可以得到量化后特征 $Q_{i,2}$ 相对应的 DWT 近似分量。

④将细节分量和量化后的近似分量进行逆 DWT,完成 w 的嵌入,得到第 i 帧的含水印信号 A'_i 。

(4)采用以上方法和步骤,将水印信号嵌入其他音频帧中,即可得到完整的含水印的音频信号。水印嵌入大致过程如图 6 所示。

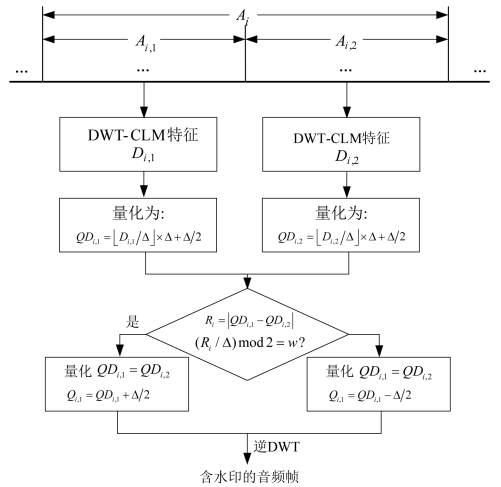


图6 水印嵌入过程框图

Fig. 6 Block diagram of watermark embedding process

2.2 水印提取

假设 $A' = \{a'_l \mid 1 \leq l \leq L'\}$ 表示验证端收到的含水印的音频信号,这里 a'_l 代表含水印音频信号 A' 的第 l 个样本点, L' 代表音频信号 A' 含有的样本点个数。和水印嵌入过程相似,将 A' 分帧、分段。将第一个音频帧记为 A'_1 。以从第一帧中提取水印信息为例,来介绍水印提取步骤(如图 7)。

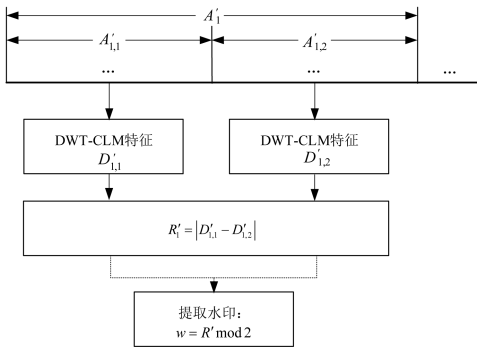


图 7 水印提取过程框图

Fig. 7 Watermark extraction process block diagram

(1) 将 A'_1 分为前后两段, 分别记为 $A'_{1,1}$ 和 $A'_{1,2}$ 。

(2) 由式(1)计算 $A'_{1,1}$ 和 $A'_{1,2}$ 的 DWT-CLM 特征, 记为 $D'_{1,1}$ 和 $D'_{1,2}$ 。并计算 $D'_{1,1}$ 和 $D'_{1,2}$ 的残差, $R'_1 = |D'_{1,1} - D'_{1,2}|$ 。

(3) 由 R'_1 提取水印 $w, w = R' \bmod 2$ 。

依照上述步骤完成其它音频帧中的水印提取, 即可得到完整的水印信息。

3 性能分析

用户为了不同的目的, 可能会对含水印的信号进行信号处理(如压缩、滤波等), 甚至是有针对性的攻击, 导致需要验证的含水印信号和原始信号相比, 样本点会有增多或者减少, 也可能导致含水印信号的水印位置和原信号不同步。为了解决此类问题, 一种可行的方法是, 将载体信号帧分为三段, 由帧号生成的水印嵌入分别嵌入到第一、二段中, 由版权信息生成的水印信息嵌入到第三段中。验证端可搜索能够正确提取帧号的音频帧, 作为含水印的音频帧, 以此来同步含水印的内容。在此基础上提取水印用来进行版权保护。

下面选取长为 600 000 的 300 段音频信号作为测试样本, 测试样本均为 16 位量化、采样频率为 44.1 kHz 的音频信号。水印分为三组, 将各帧号映射为 6 比特的二进制序列, 分别作为第一、二组的水印信息; 版权信息映射为 18 比特的水印信息, 作为第三组的水印信息。将水印信息嵌入后, 用三种不同的设备(SONY PCM-D100 录音笔、苹果 12 和三星 S6 手机)进行重录音攻击, 得到攻击后的信号。

3.1 不可听性

不可听性刻画了水印的嵌入对原始信号的改变程度, 算法要求水印的嵌入不改变语音信号的听觉质量。下面采用主观和客观两种方法对本文算

法不可听性进行测试。主观的评价方法是将原始语音信号及含水印的语音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 SDG (Subjective Difference Grades) 来打分, 打分标准如表 1 所示。

表 1 SDG 值的评分标准
Tab. 1 Score criteria for SDG values

SDG	描述	质量等级
0.0	不可感觉	优
-1.0	可感觉但不刺耳	良
-2.0	轻微刺耳	中
-3.0	刺耳	差
-4.0	非常刺耳	极差

将这一组听众最后打分的平均值, 作为原始语音信号和含水印语音信号主观听觉质量测试的结果。客观评价采用信噪比(SNR)来进行测试, SNR 的定义如式(5)所示。

$$SNR = 10 \lg \left(\sum_{l=1}^L a_l^2 / \sum_{l=1}^L (a_l - a'_l)^2 \right), \quad (5)$$

式中: a_l 表示原始音频信号的第 l 个样本点, a'_l 表示含水印音频信号的第 l 个样本点, L 表示音频信号长度。为满足不可听性要求, SNR 值应大于 20 dB。

表 2 列出了含水印音频的 SDG、SNR 的最大、最小和均值。其中 SDG 值由 12 位听众打分所得。从测试结果可以看到, SDG 最小值和均值都大于 -1, 同样地 SNR 最小值和均值大于 20, 表明本文所给算法满足水印不可听性的要求。

表 2 含水印信号的 SDG 值和 SNR 值
Tab. 2 SDG and SNR values for water-printed signals

类型	SDG	SNR /dB
最大值	-0.85	30.3
最小值	-0.72	24.9
均值	-0.67	26.8

3.2 鲁棒性

假定攻击后的含水印信号各帧和原信号保持同步, 采用误码率 BER(Bit Error Rate)来测试本文算法水印提取的可靠性, 即水印的抗信号处理和重录音攻击能力。BER 的定义如式(6)^[5]。

$$BER = \frac{B_E}{T} \times 100\%, \quad (6)$$

式中: B_E 表示错误提取水印的个数, T 表示嵌入水印的总个数。易得, BER 值越小, 提取水印错误

的个数就越少,算法抗信号处理的能力也就越强。

表 3 列出了 300 段含水印的语音信号在经过一些常见的信号处理(压缩率为 32、64 kb/s 的 MP3 压缩、截止频率为 11 kHz 和 16 kHz 的重采样、截止频率为 16 kHz 的低通滤波)以及重录音攻击后,水印提取 BER 的统计均值,并和文献[8]进行了对比。

由测试结果可得,和文献[8]相比,本文算法的误码率较低,表明本文算法具有一定的容忍信号处理的能力以及抗重录音攻击的能力。

4 结论

基于离散小波变换,提出了音频信号 DWT-CLM 特征。实验验证了该特征对信号处理和重录音攻击的鲁棒性,理论分析了该特征和音频信号

表 3 信号处理和重录音攻击后含水印信号的 BER 值
Tab. 3 BER values of water-imprinted signals after signal processing and rerecording attacks

部分信号处理和重录音攻击	BER	
	文献[8]	本文
Mp3 压缩(32 kb/s)	6.28	1.54
Mp3 压缩(64 kb/s)	2.49	0.00
重采样(44.1→11.0→44.1 kHz)	1.97	0.00
重采样(44.1→16.0→44.1 kHz)	0.61	0.00
低通滤波(16 kHz)	0.00	0.00
重录音	21.62	4.32

DWT 近似分量之间的关系。在此基础上,给出了基于 DWT-CLM 特征的量化方法和水印嵌入方法,所给的算法具有较好的不可听性和抗信号处理的能力以及一定的抗重录音攻击能力,进一步提高了水印系统的实用性。

参考文献:

[1] 侯翔, 闵连权, 唐立文. 定位篡改实体组的矢量地图脆弱水印算法[J]. 武汉大学学报(信息科学版), 2020, 45(2): 309-316.
HOU Xiang, MIN Lianquan, TANG Liwen. Fragile watermarking algorithm for locating tampered entity groups in vector map data[J]. Geomatics and Information Science of Wuhan University, 2020, 45(2): 309-316.

[2] WU Shiqiang, GUAN Hu, HUANG Ying, et al. STAW: An audio watermarking in short time DCT for copyright protection[C]//2020 International Conference on Culture-oriented Science & Technology (ICCST). Beijing: IEEE, 2020: 28-33.

[3] LIU Zhenghui, YANG Yancong, LUO Da, et al. Speech watermarking robust against recapturing and de-synchronization attacks[J]. Multimedia Tools and Applications, 2020, 79(9): 6009-6024.

[4] 刘正辉, 张钰, 秦兴红. 抗翻录攻击的鲁棒语音水印算法[J]. 武汉大学学报(信息科学版), 2021, 46(2): 303-308.
LIU Zhenghui, ZHANG Yu, QIN Xinghong. Robust speech watermarking algorithm against recapturing attacks[J]. Geomatics and Information Science of Wuhan University, 2021, 46(2): 303-308.

[5] LIU Z H, LUO D, HUANG J W, et al. Tamper recovery algorithm for digital speech signal based on DWT and DCT [J]. Multimedia Tools and Applications, 2017, 76(10): 12481-12504.

[6] 赵岩松, 钱清, 熊晶晶. 一种基于离散余弦变换的数字语音篡改恢复算法[J]. 信阳师范学院学报(自然科学版), 2022, 35(4): 645-650.
ZHAO Yansong, QIAN Qing, XIONG Jingjing. A tamper recovery scheme for digital speech forensics based on discrete cosine transform[J]. Journal of Xinyang Normal University (Natural Science Edition), 2022, 35(4): 645-650.

[7] 师春灵, 钱清. 鲁棒的数字语音取证算法[J]. 计算机工程与设计, 2021, 42(9): 2455-2461.
SHI Chunling, QIAN Qing. Robust digital speech forensics algorithm[J]. Computer Engineering and Design, 2021, 42(9): 2455-2461.

[8] 王静, 刘正辉, 周新建, 等. 一种精确篡改定位的数字语音取证算法[J]. 信阳师范学院学报(自然科学版), 2016, 29(2): 289-293.
WANG Jing, LIU Zhenghui, ZHOU Xinjian, et al. A precise tamper location algorithm used for content authentication of digital speech signal[J]. Journal of Xinyang Normal University (Natural Science Edition), 2016, 29(2): 289-293.

责任编辑: 郭红建