

基于稳健特征点的平稳小波域数字水印算法

牛盼盼^{1,2}, 杨思宇¹, 王丽¹, 杨红颖¹, 李丽¹, 王向阳¹

(1. 辽宁师范大学计算机与信息技术学院, 辽宁 大连 116029; 2. 大连理工大学电子信息与电气工程学部, 辽宁 大连 116023)

摘要: 基于目前绝大多数数字音频水印方案仅仅能够对抗简单的常规信号处理, 尚无法有效抵抗破坏性较强的局部去同步攻击, 提出了一种基于稳健特征点的平稳小波域数字音频水印算法。首先利用高斯滤波技术计算出平稳小波域低频子带的一阶平滑梯度响应并确定基准阈值, 进而提取出分布均匀且性能稳定的音频特征点; 然后结合短时能量自适应确定适合于水印嵌入的局部特征音频段; 最后利用扩展抖动调制方法将水印信息嵌入局部特征音频段中。仿真实验结果证明, 所提算法具有较好的不可感知性, 并且对常规信号处理 (MP3 压缩等) 和去同步攻击 (音调伸缩等) 均具有较好的稳健性。利用描述能力强且性能稳定的平滑梯度刻画局部数字音频性质, 提出一种基于平滑梯度的平稳小波域音频特征点提取方法, 有效解决音频特征点稳定性差且分布极不均匀的缺点, 提高了数字音频水印对幅度伸缩、局部变调、随机剪切等去同步攻击的抵抗能力。

关键词: 数字音频水印; 去同步攻击; 特征点; 平滑梯度; 平稳小波变换

中图分类号: TP391

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2019228

Digital watermarking algorithm in SWT domain based on robust local feature

NIU Panpan^{1,2}, YANG Siyu¹, WANG Li¹, YANG Hongying¹, LI Li¹, WANG Xiangyang¹

1. School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China

2. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116023, China

Abstract: Aiming at the challenging work to design a robust digital audio watermarking algorithm against desynchronization attacks, a new second generation digital audio watermarking in stationary wavelet transform (SWT) domain based on robust local audio feature was proposed. First, the first-order smooth gradient response of the low-pass sub-band coefficient was calculated using Gaussian filter. Then, the short-term energy was utilized to adaptively determine local feature audio segments for embedding. Finally, the watermark information was embedded into local feature audio segments with spread transform dither modulation. The experimental results show that the proposed approach has not only good transparency, but also has strong robustness against common audio processing such as MP3 compression and good robustness against the desynchronization attacks such as pitch-scale modification et al. A SWT domain audio feature point extraction method based on smooth gradient is proposed, which effectively solves the drawbacks of poor stability and uneven distribution of audio feature points, and improves the resistance of digital audio watermarks to amplitude-scale modification, pitch-scale modification, random cropping, and jittering attacks.

Key words: digital audio watermarking, desynchronization attack, feature point, smooth gradient, stationary wavelet transform

收稿日期: 2019-07-29; 修回日期: 2019-10-21

基金项目: 国家自然科学基金资助项目 (No.61472171, No.61701212); 中国博士后科学基金资助项目 (No.2018T110220); 辽宁省教育厅科学研究经费基金资助项目项目 (重点攻关) (No.LZ2019001); 辽宁省自然科学基金资助项目 (No.2019-ZD-0468)

Foundation Items: The National Natural Science Foundation of China (No.61472171, No.61701212), China Postdoctoral Science Foundation Funded Project (No.2018T110220), Key Scientific Research Project of Liaoning Provincial Department (No.LZ2019001), The Natural Science Foundation of Liaoning Province (No.2019-ZD-0468)

1 引言

数字水印是一种可以在开放的网络环境下保护版权和认证来源及完整性的新技术, 目前已成为网络与信息安全领域的一个研究热点。不可感知性、稳健性及水印容量是数字水印系统必须满足的基本要求, 也是近年来数字水印领域的主要研究内容^[1]。用于版权保护的数字水印系统必须具有高度稳健性。现有大多数稳健音频水印算法能够抵抗简单的常规音频信号处理攻击(如叠加噪声、重新量化、重新采样、MP3 压缩等), 尚无法有效抵抗破坏性较强的去同步攻击(desynchronization attack)^[2-3]。去同步攻击能够改变水印信号嵌入位置, 即破坏水印信号分量的同步性, 使水印检测器找不到有效的水印信号^[4]。数字音频水印系统中的去同步攻击可以分为 2 类: 简单的全局去同步攻击和复杂的局部去同步攻击。其中, 全局去同步攻击是针对整段音频信号的攻击方式, 如时间缩放、幅度伸缩、音调伸缩、DA/AD 转换等, 即对音频整体内容进行了破坏; 局部去同步攻击是针对局部音频信号的攻击方式, 如随机剪切、局部变调、抖动攻击等, 即对音频信号的一部分内容进行了破坏^[5]。

到目前为止, 人们主要采纳 5 种策略解决数字音频水印领域的去同步攻击问题, 包括穷举搜索、构造同步不变特征、扩频水印扩频码相结合、同步码、利用音频重要特征等。

1) 基于穷举搜索的稳健音频水印

穷举搜索是数字音频信号遭受时域去同步攻击后恢复水印信息的最简单方法^[6]。该类水印算法存在 2 个致命问题, 一是对数字水印检测器的多次操作会显著增加虚警率; 二是计算复杂度太高, 其计算量将随着搜索空间的扩大而迅速膨胀。因此该类方法只适用于小数据量水印检测, 且无法有效抵抗诸如时间缩放、随机剪切、抖动等去同步攻击。

2) 基于同步不变特征的稳健音频水印

同步不变特征方案以其原理简单、性能稳定等优点而受到人们普遍重视。该类水印方案的基本思想是将水印信息隐藏在原始载体音频信号具有同步不变性的量中, 而目前所使用的同步不变量主要包括二进小波变换(DYWT, dyadic wavelet transform)、几何不变矩、直方图、奇异值、向量范数等^[7-9]。在含水印音频遭受各种攻击后, 同步不变量由于具有同步不变特性不会发生变化, 因此隐藏于

其中的数字水印信号得以保存。但目前该类水印方案仅仅能够抵抗轻微的全局音调伸缩和整体时间缩放, 尚不能有效抵抗诸如抖动、局部变调、随机剪切等较复杂的局部去同步攻击。

3) 基于扩频水印扩频码相结合的稳健音频水印

扩频数字水印技术往往先利用纠错编码技术, 对水印信号进行编码预处理; 然后利用正交频分复用、相移键控、移频键控等调制方法, 对数字水印信号进行频谱扩展; 最后在原始载体音频的变换系数中叠加扩频水印信号^[10-11]。扩频水印扩频码相结合算法的优势在于不需要额外同步信号, 且能够有效抵抗诸如幅度伸缩等轻微全局去同步攻击。但其仍然存在 2 个问题, 一是由于其缺乏必要的同步保护机制, 故普遍难以抵抗局部变调、随机剪切、抖动等破坏性较强的局部去同步攻击; 二是无法实现数字水印信号的盲检测, 不利于应用。

4) 基于同步码的稳健音频水印

基于同步码的音频水印方案需同时嵌入数字水印信号和同步码信息。进行水印检测时, 首先检测到同步码信息, 然后再根据同步码信息确定出水印信号的嵌入位置。显然, 这类数字音频水印方法的稳健性和安全性主要依赖于同步码信息。总体说来, 基于同步码的音频水印方案具有一定的抵抗诸如整体时间缩放等轻微去同步攻击能力, 但其无法有效抵御诸如局部变调、随机剪切等局部去同步攻击, 且数字水印容量受到一定限制。此外, 同步码音频水印算法还会产生 2 个严重问题, 一是安全性问题, 二是虚警问题。原因是当一个同步标记被用于一系列不同数字音频作品时, 其很容易被敌人发现和去除^[12-13]。

5) 基于音频重要特征的稳健音频水印

基于载体内容局部特征的数字水印方案, 其基本思想为, 利用相对稳定的载体特征点(也称显著点, salient point)来标识水印嵌入位置并确定相对应的特征区域用以嵌入水印信息。在检测含水印音频信号时, 同样利用载体特征点进行定位和提取水印信息, 从而有效抵御破坏性较强的去同步攻击。Wang 等^[14]结合人类听觉掩蔽特性, 提出了一种基于载体音频内容的新能量特征点提取方法。该算法具有一定的抵抗噪声干扰等常规攻击能力及轻微剪切攻击能力, 但无法有效抵抗时间缩放、音调伸缩、DA/AD 转换等去同步攻击能力。Li 等^[15]首先

提取出原始载体音频包络的峰值点, 然后以此作为音频特征点并从原始载体中选取 4 096 个局部音频片段用于水印信号嵌入, 最后利用 FFT 在每个局部音频片段内独立嵌入数字水印信号。该算法可抵抗轻微的时间缩放、随机剪切等去同步攻击, 但对高通滤波等常规音频处理比较敏感, 而且无法有效抵御联合攻击。Ma 等^[16]结合独立分量分析理论, 选取数字音频能量峰值点作为特征标识并依此确定数字水印嵌入位置, 而数字水印信号被嵌入子采样后的音频载体内, 该算法对噪声添加、重新采样等常规音频处理非常敏感。鲍德旺等^[17]首先利用非下采样小波分解所得的低频信息重构出新载体数字音频, 然后根据数字音频的局部能量特征, 从新载体中提取出稳定的音频特征点, 并以音频特征点为标识提取出长度固定的局部音频段, 最后利用量化调制方法将水印信息嵌入局部音频段内。该算法具有较好的抗噪声干扰能力, 但其抵抗时间缩放、音调伸缩、DA/AD 转换等攻击能力较弱。Pun 等^[18]提出了一种自同步小波域数字音频水印方案, 该方案利用音频载体信号的局部方差不变特性来定位水印嵌入位置, 具有较好的抵抗常规音频信号处理及轻微剪切攻击能力, 但对 DA/AD 转换、音调伸缩、时间缩放等比较敏感。Pun 等^[19]提出了一种基于稳健不变音频特征点的检测方法, 利用 SWT (stationary wavelet transform) 系数的平移不变性、不可感知性和稳健性嵌入水印信息。该算法在空域内提取特征点, 与频域相比抵抗攻击能力较差, 并且该算法是选取特征点周围的固定长度作为待嵌入水印特征区域, 而不是根据特征点周围的局部特征从而自适应确定待嵌入水印特征区域大小, 水印容量与其他算法相比也相对较小。Yuan 等^[20]首先在空间域内提取出载体音频的局部能量峰值点, 并将其作为稳定的数字音频特征点, 然后以音频特征点为标识, 确定固定尺寸的局部音频段, 最后对局部音频段进行小波变换, 并利用量化调制技术将水印信息嵌入低频子带内, 该方法计算简单、容易实现, 但其抵抗噪声干扰、时间缩放、音调伸缩等攻击能力均较差。Xiang 等^[21]提出了一种基于 DCT (discrete cosine transform) 域的大容量稳健水印检测方法, 利用 DCT 系数的不可感知性和稳健性嵌入正交 PN 序列作为水印信息, 该算法在提高算法稳健性的同时, 保证了较高的水印负载能力。

此外, 近年来已有学者将深度学习方法应用到数字水印技术中, 对基于深度学习的水印方法进行了深入的研究, 以寻求不可感知性及同时抵抗常规信号处理操作和去同步攻击的稳健性之间的平衡为优化目标, 利用粒子群优化等技术设计音频水印优化过程, 自动确定水印算法的最优水印参数和搜索最合适的嵌入位置等, 为所设计的水印系统提供了更大的灵活性^[22-23]。

综上所述, 穷举搜索、构造同步不变特征、扩频水印扩频码相结合、同步码等方案难以有效解决诸如抖动、局部变调、随机剪切等较复杂的局部去同步攻击问题, 而基于音频重要特征的数字音频水印方案为去同步攻击问题提供了一些可行而有效的解决方向, 但现有该类方案普遍存在诸如特征点稳定性差且分布极不均匀、计算量偏大、时间复杂度较高等一系列问题, 从而严重影响了数字水印的稳健性和不可感知性, 大大降低了整个数字音频水印系统的工作性能。鉴于此, 本文提出了一种基于稳健特征点的平稳小波域音频水印算法。该算法首先利用高斯滤波技术计算出平稳小波域低频子带的一阶平滑梯度响应, 并结合平滑梯度响应分布确定基准阈值, 进而提取出分布均匀且性能稳定的音频特征点; 然后以音频特征点为标识, 结合短时能量^[24]自适应确定适合于水印嵌入的局部特征音频段; 最后利用扩展抖动调制方法^[25]将水印信息嵌入局部特征音频段中。实验结果验证了该算法不仅不可感知性良好, 并且对常规音频信号处理和局部去同步攻击均具有较好的稳健性。

2 平稳小波变换简介

与 DWT 相比, 平稳小波变换 (SWT, stationary wavelet transform)^[26]具有平移不变性和冗余性等特性, 可以对连续小波变换做出更加近似的估计。在音频信号进行分解时, DWT 采用下采样抽样, 并采用差值补 0 的方法对音频信号进行重构; 而 SWT 在进行音频信号分解时采用非下采样抽样, 将信号分解成一系列与原信号长度相同的近似系数和细节系数, 同时在每 2 个滤波器系数之间插入零值以实现滤波器的延拓, 从而避免在音频信号进行重构时产生 Gibbs 震荡现象。

在一个正交小波变换中, 存在尺度函数 $\varphi_{j,u}(t)$ 与小波函数 $\psi_{j,u}(t)$, 尺度函数 $\varphi_{j,u}(t)$ 可以从多分辨

率函数 $L^2(R)$ 中获得。设 $f(t) \in L^2(R)$ 表示音频信号，SWT 的定义为

$$a_j(u) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{2^j}} \varphi \left[\frac{(t-u)}{2^j} \right] dt = \langle f(t), \varphi_{2^j}(t-u) \rangle$$

$$d_j(u) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{2^j}} \psi \left[\frac{(t-u)}{2^j} \right] dt = \langle f(t), \psi_{2^j}(t-u) \rangle$$

$$0 < j \leq J, u \in Z \quad (1)$$

其中， $a_j(u)$ 为近似系数或低频系数， $d_j(u)$ 为细节系数或高频系数。对于 $j \geq 0$ ，则有

$$a_{j+1}(u) = a_j * \bar{h}_j(u), \quad d_j(u) = a_j * \bar{g}_j(u),$$

$$a_j(u) = \frac{1}{2} (a_{j+1} * h_j(u) + d_{j+1} * g_j(u))$$

其中， $h(n)$ 、 $g(n)$ 为小波函数 $\psi(t)$ 对应的共轭镜像滤波器组， $\bar{h}_j(n)$ 为在 $h(n)$ 相邻两项之间插入 $2^j - 1$ 个 0 得到的扩展滤波器。不同于 DWT，SWT 并未将每个分级得到的近似系数和细节系数纳入到下采样范围中，故变换后的近似系数和细节系数能够保持与原始信号相同的长度。进行 SWT 操作可以得到每个分级的低通和高通滤波系数，此操作相当于扩频操作。随着分级层数的增加，SWT 系数的分辨率随之下降，利用滤波器，可以均匀地进行下采样操作并进一步分级低通系数（二次分级）。SWT 的结构如图 1 所示。

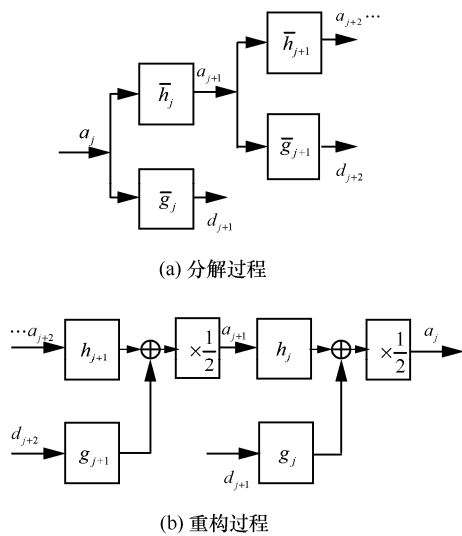


图 1 SWT 的结构

为证明 SWT 的有效性，图 2 给出了数字音频的 3 级 SWT 分解/重构示意。其中，图 2(a)为原始音频，图 2(b)为 3 级小波分解的低频子带，图 2(c)~

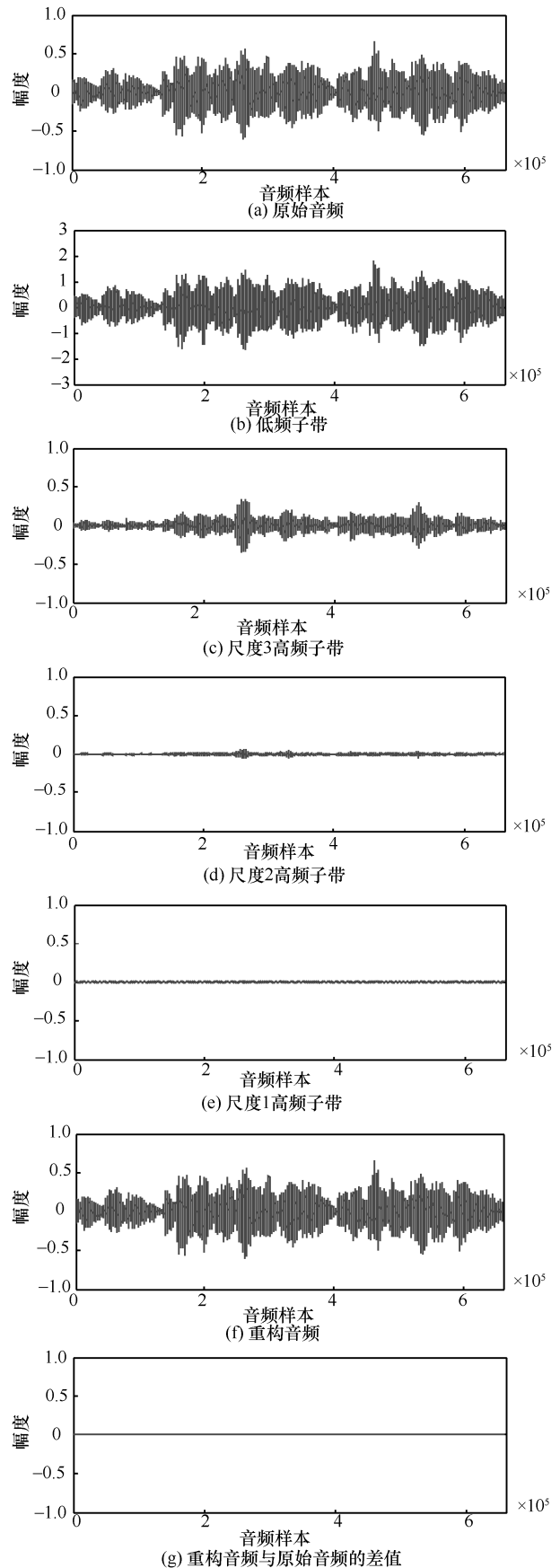


图 2 数字音频的 3 级非下采样小波分解/重构示意

图 2(e)分别为 3 级小波分解的尺度 3、尺度 2 和尺度 1 高频子带, 图 2(f)为小波重构音频, 图 2(g)为小波重构音频与原始音频的差值。

观察图 2 可以看出, 数字音频经过 3 级 SWT 分解后再重构, 重构波形与原始音频波形基本保持一致。

3 基于平滑梯度的音频特征点提取

特征点提取是基于载体内容局部特征的数字音频水印方案的关键环节之一。目前, 可应用于数字音频水印领域的音频特征点检测算子包括样本峰值法、包络峰值法^[14-22]。然而, 以上音频特征点检测算子均直接采用音频样本值刻画局部数字音频性质, 其不可避免存在 2 方面的问题: 一方面, 特征点检测算子对音程跳跃幅度较大的“高”音调区很敏感, 会在“高”音调区域检测出很多“噪声”特征点; 另一方面, 特征点过于集中在高频泛音充分的“尖”音色区。也就是说, 利用现有检测算子所提取出的音频特征点, 普遍存在稳定性差且分布极不均匀的问题, 严重影响数字音频水印对 DA/AD 转换、随机剪切、抖动等攻击的抵抗能力。尽管 Dessein 等^[27]结合最大似然参数估计理论, 提出了一种基于广义似然比极值的数字音频特征点检测算法, 但由于其所采纳的样本指数分布假设并不适合大多数实际音频, 故该算法的音频特征点检测性能较低。此外, 该特征点检测算法的时间复杂度较高, 不利于实际应用。

本文将利用描述能力强且性能稳定的平滑梯度刻画局部数字音频性质, 进而提出一种基于平滑梯度的平稳小波域音频特征点提取方法, 可以很好地满足音频水印的要求。

假设 $A = \{a(x), 1 \leq x \leq L\}$ 表示原始数字音频信号, 其中 L 为音频的样本数, $a(x) \in \{0, 1, 2, \dots, (2^p - 1)\}$ 表示每个样本的幅值, p 表示每个样本的量化位数, 则基于平滑梯度的音频特征点提取方法如下。

步骤 1 SW 变换。对原始载体音频 A 进行 3 级 SWT 分解, 以得到一个低频子带 F 和若干个高频子带。本文选取抗噪能力较强的低频子带提取音频特征点。

步骤 2 音频梯度计算。利用式(2), 计算出低频子带 $F(x) = \{f_1, f_2, \dots, f_x, \dots, f_L\}$ 的一阶梯度 $F'(x) = \{f'_1, f'_2, \dots, f'_x, \dots, f'_L\}$, 并令 $f'_1 = f'_L = 0$ 。其

中, 音频梯度值 $F'(x) (1 < x < L)$ 为

$$F'(x) = \frac{\delta F}{\delta x} = F(x+1) - F(x-1) \quad (2)$$

步骤 3 平滑处理。利用式(3), 对一阶音频梯度进行高斯滤波处理, 以得到性能稳定的一阶平滑梯度 $F'_s(x)$

$$F'_s(x) = G(x) * F'(x) \quad (3)$$

其中, $G(x) = e^{-\frac{x^2}{2\sigma^2}}$ 表示高斯核函数, $*$ 表示卷积运算。

步骤 4 响应计算。利用一阶平滑梯度幅值, 计算出原始音频样本 $a(x)$ 的响应值 $R(x)$, 并以此刻画局部数字音频性质。音频样本 $a(x)$ 的局部响应值 $R(x)$ 为

$$R(x) = b(F'_s(x))^2 \quad (4)$$

这里, b 为任意常量, 实验中 $b=0.5$ 。

步骤 5 特征点提取。对音频样本响应值 $R(x) = \{r_1, r_2, \dots, r_x, \dots, r_L\}$ 进行降序排列得到 $P_R = \{P_{r-L}, P_{r-L-1}, \dots, P_{r-1}\}$, 滤掉位于首尾两端的音频样本, 确保有足够的空间可以嵌入完整水印信息, 即不满足

$$\left(P_R - \frac{X}{2} \geq 1\right) \cap \left(P_R + \frac{X}{2} - 1 \leq L\right) \quad (5)$$

其中, X 代表相邻特征点之间的基本间距, 本实验中, $X = 1.5 \times 10^4$ 。

以最高响应值 P_{r-x} 为基准设置阈值, 滤掉与其相邻过近的特征点, 从而确保提取出的特征点分布均匀, 即不满足

$$|P_{r-y} - P_{r-x}| \geq X \quad (6)$$

其中, $P_{r-y} \in \{P_{r-x-1}, P_{r-x-2}, \dots, P_{r-1}\}$, 选取前 Y 个点 $P_r \in \{P_1, P_2, \dots, P_Y\}$ 作为本方法的音频特征点即同步标记, 本实验中, $Y=13$ 。

图 3 给出了本文特征点提取方法抵抗常规信号处理和去同步攻击的部分实验结果。其中, “○”表示正确提取, “●”表示错误提取(本实验中, 特征点位置偏移 5 个样本以上即视为错误点)。

4 数字水印嵌入与检测算法

4.1 数字水印的嵌入

本文采用冗余嵌入策略, 提出了一种基于稳健特征点的平稳小波域数字音频水印算法, 水印嵌入的关键步骤可描述如下。

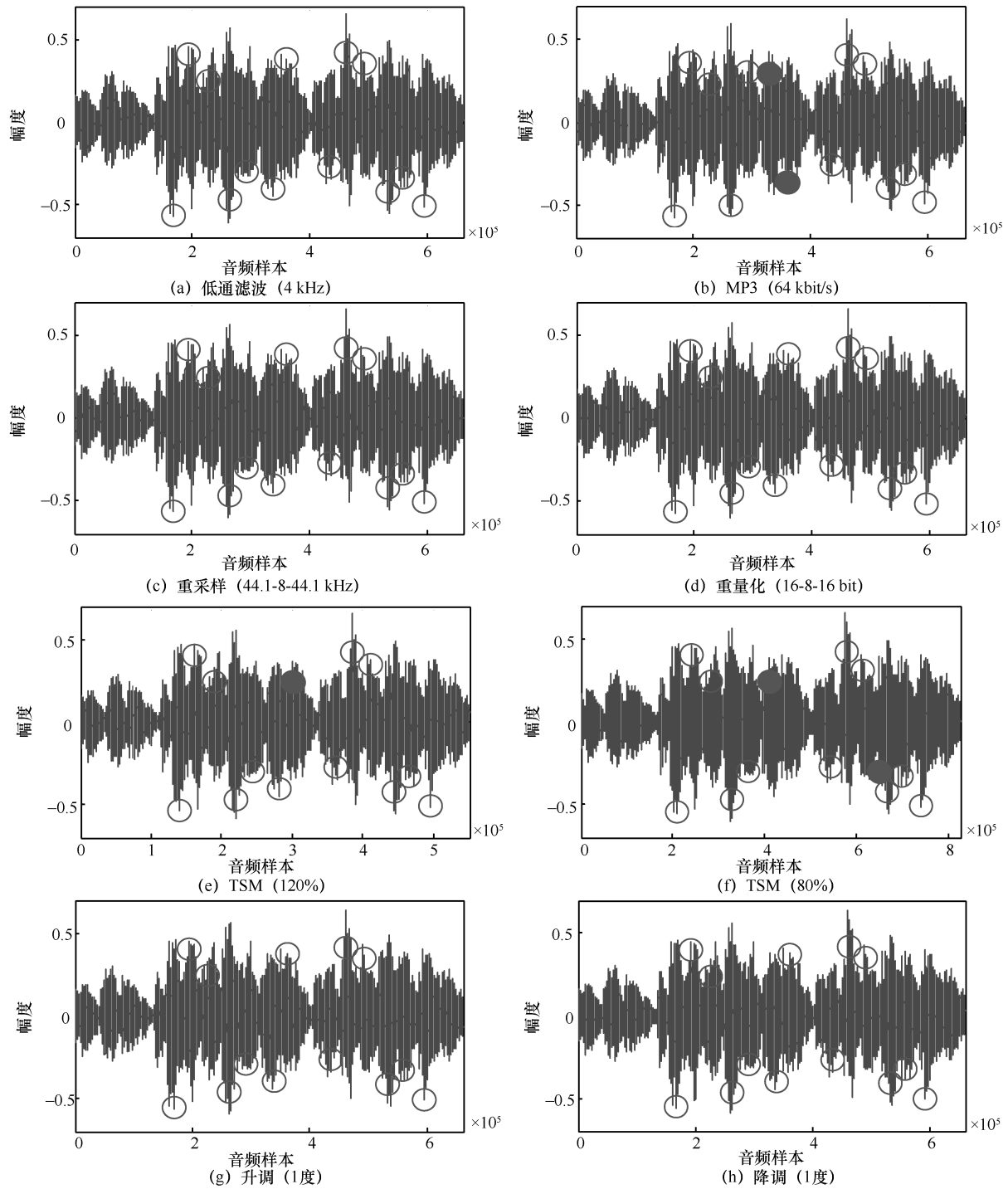


图 3 各种攻击下的音频特征点提取

步骤 1 数字水印产生。由密钥 K 产生一个伪随机序列 $C = \{c(i), i = 1, \dots, s\}$ 作为数字水印信息，其中， s 为水印大小， $c(i) \in \{0, 1\}$ 。

步骤 2 音频特征点提取。利用基于平滑梯度的平稳小波域音频特征点检测算子，从原始音频载体 A 中提取音频特征点，以得到音频特征点集

$$P_r \in \{P_1, P_2, \dots, P_Y\}。$$

步骤 3 局部特征音频段构造。所谓局部特征音频段，是指以数字音频特征点为标识，从原始载体音频中分割出的一部分子音频，以其作为水印信号的嵌入和检测区域。本文局部特征音频段的具体构造过程如下。

1) 将载体音频信号 A 进行降序排列, 得到载体音频数据最大值 A_{\max} 与最小值 A_{\min} , 并归一化到 $[-1,1]$ 的范围, 具体计算式为

$$A_{\text{norm}} = \frac{2(A - A_{\min})}{A_{\max} - A_{\min}} - 1 \quad (7)$$

其中, A_{norm} 是归一化后的音频信号。

2) 对 A_{norm} 进行 3 级 SWT 分解, 以得到一个低频子带 F_1 和若干个高频子带。

3) 在低频子带 F_1 下, 以特征点 P_r 为标识, 循环统计每个特征点 P_r 左右窗口长度为 M 范围内的能量和, 用以代表特征点周围音频短时能量。于是, 短时能量可表示为^[24]

$$\begin{aligned} E_n &= \sum_{m=0}^n [a_1(m)w(n-m)]^2 = \\ &= \sum_{m=0}^n a_1^2(m)h(n-m) = \\ &= a_1^2(n) \times h(n-m) \end{aligned} \quad (8)$$

其中, E_n 表示 n 时刻音频信号 $a_1(n)$ 的短时能量, $w(n-m)$ 表示窗函数, $h(n) = w^2(n)$, $n = 1 + 2M$ 。本文选取的窗函数为汉宁窗 $w(n)$, 即

$$w(n) = 0.5 \left(1 - \cos \left(2\pi \frac{n}{N+1} \right) \right), \quad n = 1, 2, \dots, N \quad (9)$$

其中, N 表示窗长, 特征点 P_r 周围的音频信号集可表示为

$$\{F_1(P_1 - M, P_1 + M - 1), F_1(P_2 - M, P_2 + M - 1), \dots, F_1(P_Y - M, P_Y + M - 1)\}$$

4) 利用特征点短时能量集 $E \in \{E_1, E_2, \dots, E_Y\}$ 再次计算

$$E_1 = \left\lfloor \frac{E}{100} \right\rfloor \quad (10)$$

得到 $E_1 \in \{E_{1-1}, E_{1-2}, \dots, E_{1-Y}\}$, $\lfloor \cdot \rfloor$ 为 round 运算。

5) 计算自适应局部特征音频段, 其中, 待嵌入水印音频段长度 $D_r = E_1 s K$, s 为水印大小, K 为常数, $D_r \in \{D_1, D_2, \dots, D_Y\}$ 。于是, 得到用于水印嵌入的局部特征音频段集为

$$\left\{ \left[P_1 - \frac{D_1}{2}, P_1 + \frac{D_1}{2} - 1 \right], \left[P_2 - \frac{D_2}{2}, P_2 + \frac{D_2}{2} - 1 \right], \dots, \left[P_Y - \frac{D_Y}{2}, P_Y + \frac{D_Y}{2} - 1 \right] \right\}$$

步骤 4 数字水印嵌入。本文采用量化调制方法, 数字水印的具体嵌入过程如下。

1) 对载体音频信号 A 进行 3 级 SWT 分解, 以得到一个低频子带 F 和若干个高频子带。本文选取性能稳定的 SWT 域低频子带嵌入水印信息。

2) 在低频子带 F 内, 依据特征点 P_r 和待嵌入水印音频段长度 D_r 找到对应的特征区域, 并根据水印位的长度 s 对待嵌入水印的局部特征音频段进行分段, 每段长度为 $t_1 = \left\lfloor \frac{D_r}{s} \right\rfloor$, 可得到音频子段

$$a_2 \in \{a_{2-1}, a_{2-2}, \dots, a_{2-t_1}\}。$$

3) 量化调制音频子段 a_2 内的 t_1 个低频系数, 每个音频子段嵌入一个水印位。首先根据水印信息的值, 使用 2 种不同的抖动量化^[25], 自适应确定水印嵌入强度。

① 当 $C = 0$ 时, 抖动量化 Q_0 为

$$Q_0(a_2, \Delta) = \left\lfloor \frac{(a_2 - d_0)}{\Delta} \right\rfloor \Delta + d_0 \quad (11)$$

② 当 $C = 1$ 时, 抖动量化 Q_1 为

$$Q_1(a_2, \Delta) = \left\lfloor \frac{(a_2 - d_1)}{\Delta} \right\rfloor \Delta + d_1 \quad (12)$$

其中, Δ 为量化步长, $d_0 = -\frac{\Delta}{4}$, $d_1 = \frac{\Delta}{4}$ 。

然后将水印信息利用 STDM (spread transform dither modulation) 量化方法嵌入音频子段, 量化调制方法为^[25]

$$a_2' = a_2 + \left(Q_C(a_2 \mathbf{p}^T, \Delta) - a_2 \mathbf{p}^T \right) \mathbf{p}, \quad C \in \{0, 1\} \quad (13)$$

$$a_2' = a_2 + \left(\left(\left\lfloor \frac{(a_2 \mathbf{p}^T - d_C)}{\Delta} \right\rfloor \Delta + d_C \right) - a_2 \mathbf{p}^T \right) \mathbf{p} \quad (14)$$

其中, \mathbf{p} 为投影向量, a_2' 为嵌入水印后的音频段。

4) 用含水印音频子段内的低频系数代替原低频系数, 并进行逆 SWT, 即可得到含水印局部特征音频段。

步骤 5 含水印音频获得。重复步骤 1~步骤 4，直到所有局部特征音频段都按照上述步骤处理完毕。最后，用含水印局部特征音频段替换原始的局部特征音频段，得到含水印数字音频 A_w 信号。

4.2 数字水印的检测

步骤 1 使用与嵌入过程相同的密钥 K 产生原始数字水印序列 $C = \{c(i), i = 1, \dots, s\}$ 。

步骤 2 利用基于平滑梯度的平稳小波域音频特征点检测算子，从含水印数字音频 A_w 中提取音频特征点，以得到音频特征点集 $P_{rw} \in \{P_1, P_2, \dots, P_{Y_w}\}$ 。

步骤 3 以所提取的音频特征点为标识，利用本文局部特征音频段构造方法，计算出短时能量 E_2 ，并根据 E_2 计算出含水印特征音频段长度 $D_{rw} = E_2 s K$ ， s 为水印大小， K 为常数， $D_{rw} \in \{D_1, D_2, \dots, D_{Y_w}\}$ ，从而得到用于水印提取的局部特征音频段集为

$$\left\{ \left[P_1 - \frac{D_1}{2}, P_1 + \frac{D_1}{2} - 1 \right], \left[P_2 - \frac{D_2}{2}, P_2 + \frac{D_2}{2} - 1 \right], \dots, \left[P_{Y_w} - \frac{D_{Y_w}}{2}, P_{Y_w} + \frac{D_{Y_w}}{2} - 1 \right] \right\}$$

步骤 4 对含水印数字音频 A_w 进行 3 级 SWT 分解，并按照如下过程提取水印信息。

1) 在低频子带 F_w 内，依据特征点 P_{rw} 和含水印特征音频段长度 D_{rw} 找到对应的特征区域，并根据水印位的长度 s 对待提取水印的局部特征音频段进行分段，每段长度为 $t_2 = \left\lfloor \frac{D_{rw}}{s} \right\rfloor$ ，得到音频子段 $a'_2 \in \{a'_{2_1}, a'_{2_2}, \dots, a'_{2_{t_2}}\}$ （每个音频子段提取一个水印位）。

2) 从每一音频子段 a'_2 内的 t_2 个低频系数中，采用最小距离检查方法提取水印位，计算式为^[25]

$$\text{ExtMessage} = \arg \min_{C \in \{0,1\}} \left| a'_2 \mathbf{p}^T - Q_C(a'_2 \mathbf{p}^T, \Delta) \right| \quad (15)$$

3) 重复上述步骤，即可检测出局部特征音频段内所隐藏的数字水印序列为

$$C^* = \left\{ \text{ExtMessage}^*(i) \right\} (i = 0, 1, \dots, s-1) \quad (16)$$

重复步骤 4，直到检测完所有局部特征音频段。只要有 2 个以上的局部特征音频段能够成功检测到水印信号，即可认为数字水印检测成功；否则，水印检测失败。

5 仿真实验与结论

本节分别给出了检测性能测试以及抗攻击能力测试的仿真实验结果。实验中，选用采样频率为 44.1 kHz、量化精度为 16 bit、长度为 15 s 的单声道数字音频信号作为原始载体，数字水印采用了 128 位的伪随机序列。另外，量化步长 $\Delta=73$ ，检测阈值 $T_s=24$ （即虚警率 $P_{F-A} \approx 5 \times 10^{-4}$ ）。

5.1 不可感知性测试

图 4 给出了本文算法的不可感知性能测试结果，包括原始音频、含水印音频、差值音频。从图 4(a) 与图 4(b) 中可以明显看出，原始音频与含水印音频的波形区别很小，而差值图 4(c) 也充分证明了这一观点。另外，本文算法的 PSNR 值为 39.831 2 dB，客观结果充分证明了本文算法的不可感知性良好。

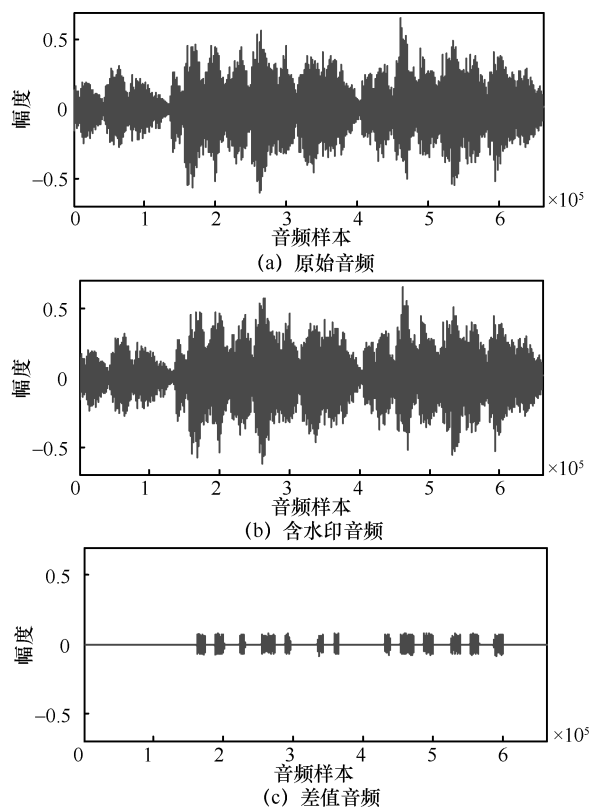


图 4 数字音频水印的不可感知性能测试结果

5.2 稳健性测试

为了验证本文数字音频水印算法的稳健性能，仿真实验分别对本文算法和文献[19-21]算法的含水印音频进行了一系列攻击。其中，文献[19-20]算法均是利用了平滑梯度刻画了局部数字音频性质，针对如何有效解决数字音频水印算法

的局部去同步攻击问题展开研究，因此本文首先与文献[19-20]算法进行了性能比较。文献[21]算法为较新的具有代表性的基于扩频水印扩频码相结合的数字音频水印方法。图 5 和图 6 分别给出了本文算法抵抗常规信号处理和去同步攻击的实验结果。其中，分子表示从攻击后含水印数字音频中成功检测到水印信号的局部特征音频段数目，分母表示原始载体音频中嵌有水印信号的局部特征音频段数目。这里，定义每个完整水印信息正确提取位数多于 $\frac{3}{4}$ ，即视为该自适应特征音频段

的水印被正确提取。本实验中，利用 RCES (the ratio of correctly extracted segment) 表示自适应特征音频段水印的正确提取率，正确提取的含水印局部特征音频段选用浅色填充“■”，错误提取的含水印局部特征音频段选用深色填充“■”。表 1 和表 2、图 7 和图 8 给出了本文算法和文献[19-21]算法的稳健性能对比。

表 1 和表 2 给出了各种攻击下本文算法与文献[19-20]的对比结果，图 7 和图 8 显示了本文算法与文献[21]算法在相同原始音频载体条件下，抵抗常

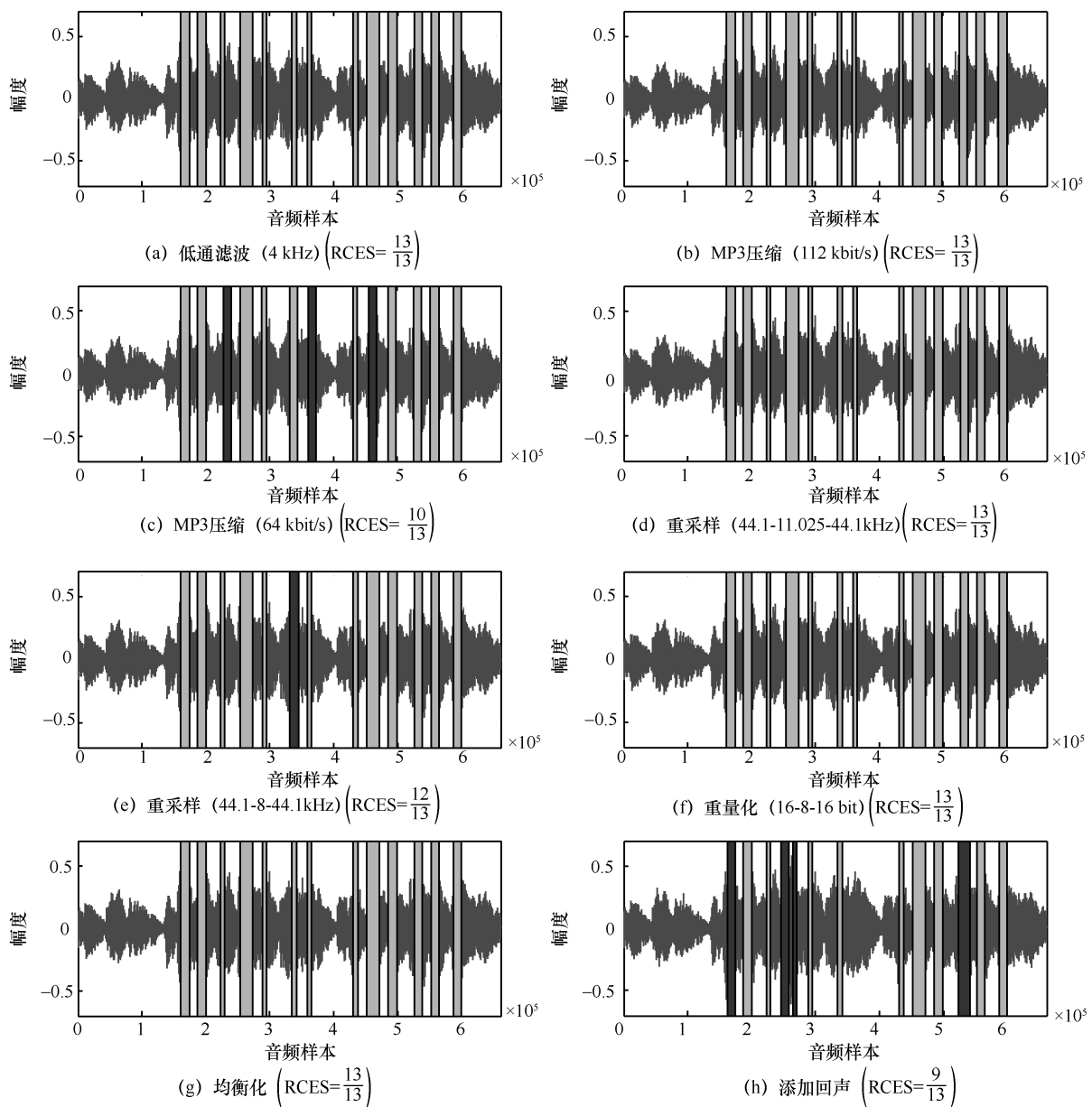


图 5 本文算法抵抗常规音频信号处理的实验结果

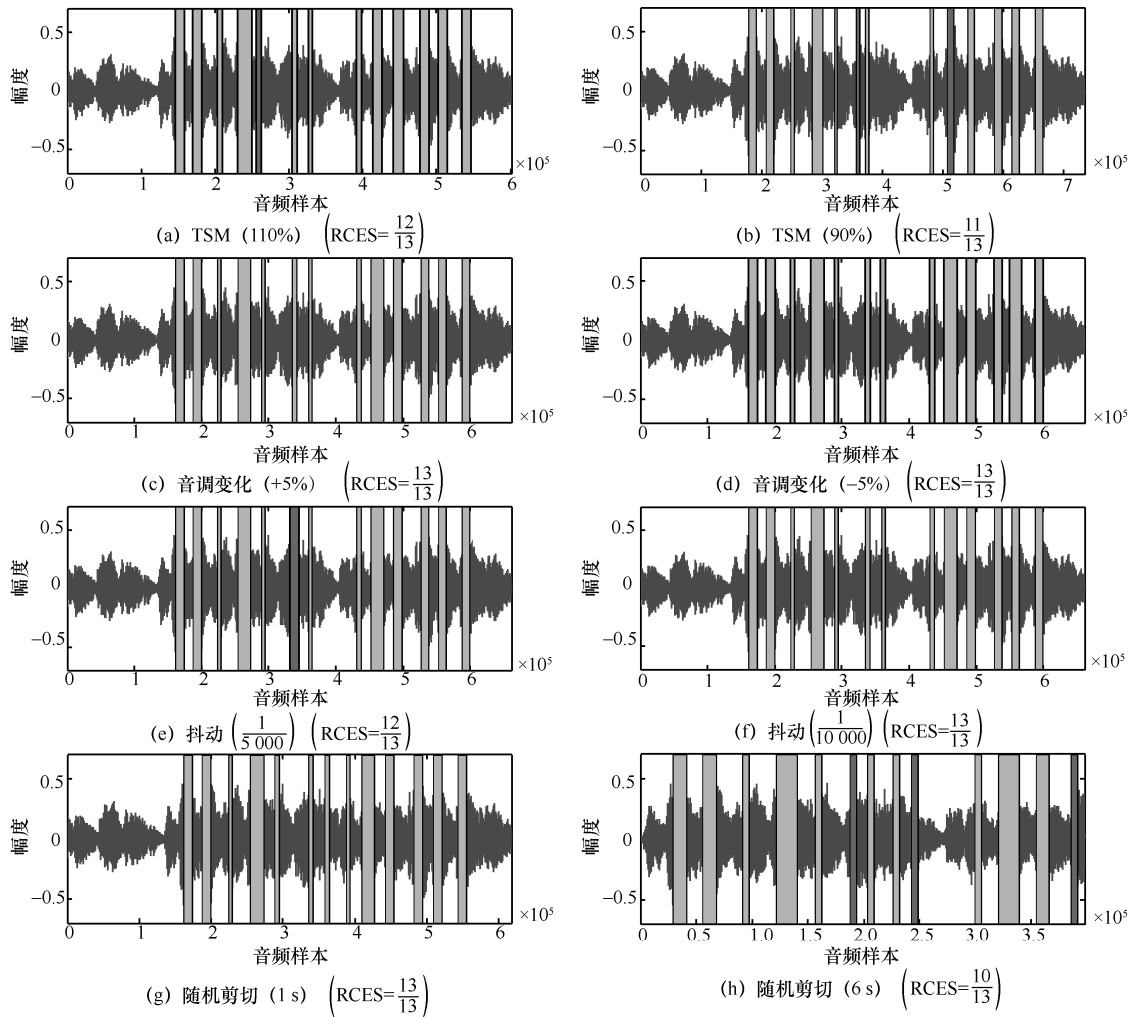


图 6 本文算法抵抗去同步攻击的实验结果

表 1 本文算法和文献[20]算法的稳健性对比结果

攻击方式	本文算法	文献[20]算法	攻击方式	本文算法	文献[20]算法
低通滤波 (8 kHz)	$\frac{13}{13}$ (100%)	$\frac{10}{15}$ (66.7%)	幅度伸缩 (+20%)	$\frac{9}{13}$ (69.2%)	$\frac{5}{15}$ (33.3%)
低通滤波 (4 kHz)	$\frac{13}{13}$ (100%)	$\frac{7}{13}$ (53.8%)	幅度伸缩 (+10%)	$\frac{7}{13}$ (53.8%)	$\frac{7}{15}$ (46.7%)
MP3 压缩 (64 kbit/s)	$\frac{10}{13}$ (76.9%)	$\frac{11}{15}$ (73.3%)	幅度伸缩 (-10%)	$\frac{9}{13}$ (69.2%)	$\frac{8}{15}$ (53.3%)
MP3 压缩 (32 kbit/s)	$\frac{4}{13}$ (30.8%)	$\frac{8}{13}$ (61.5%)	幅度伸缩 (-20%)	$\frac{12}{13}$ (92.3%)	$\frac{7}{15}$ (46.7%)
重采样 (44.1-11.025-44.1 kHz)	$\frac{13}{13}$ (100%)	$\frac{10}{13}$ (76.9%)	音调变化 (+5%)	$\frac{13}{13}$ (100%)	$\frac{12}{15}$ (80%)
重采样 (44.1-8-44.1 kHz)	$\frac{12}{13}$ (92.3%)	$\frac{10}{15}$ (66.7%)	音调变化 (-5%)	$\frac{13}{13}$ (100%)	$\frac{11}{15}$ (73.3%)
重量化 (16-8-16 bit)	$\frac{13}{13}$ (100%)	$\frac{11}{15}$ (73.3%)	随机剪切 (2 s)	$\frac{10}{13}$ (76.9%)	$\frac{6}{15}$ (40%)
高斯白噪声 (0.01)	$\frac{13}{13}$ (100%)	$\frac{8}{13}$ (61.5%)	抖动 ($\frac{1}{5000}$)	$\frac{12}{13}$ (92.3%)	$\frac{5}{15}$ (33.3%)
添加回声 (100 ms, 0.5)	$\frac{9}{13}$ (69.2%)	$\frac{5}{17}$ (29.4%)	TSM (110%)	$\frac{12}{13}$ (92.3%)	$\frac{4}{15}$ (26.7%)
均衡化 (baseboost)	$\frac{13}{13}$ (100%)	$\frac{5}{11}$ (45.5%)	TSM (90%)	$\frac{11}{13}$ (84.6%)	$\frac{6}{15}$ (40%)

表 2 本文算法和文献[19]算法的稳健性能对比结果

攻击方式	本文算法			文献[19]算法		
	Danube	Heart	Piano	Danube	Heart	Piano
MP3 压缩 (64 kbit/s)	$\frac{12}{13}$ (92.3%)	$\frac{9}{12}$ (75%)	$\frac{10}{13}$ (76.9%)	$\frac{14}{16}$ (87.5%)	$\frac{14}{16}$ (87.5%)	$\frac{14}{16}$ (87.5%)
MP3 压缩 (96 kbit/s)	$\frac{12}{13}$ (92.3%)	$\frac{11}{12}$ (91.7%)	$\frac{13}{13}$ (100%)	$\frac{15}{16}$ (93.8%)	$\frac{15}{16}$ (93.8%)	$\frac{16}{16}$ (100%)
低通滤波 (4 kHz)	$\frac{13}{13}$ (100%)	$\frac{11}{12}$ (91.7%)	$\frac{13}{13}$ (100%)	$\frac{14}{16}$ (87.5%)	$\frac{8}{16}$ (50%)	$\frac{14}{16}$ (87.5%)
低通滤波 (8 kHz)	$\frac{13}{13}$ (100%)	$\frac{12}{12}$ (100%)	$\frac{13}{13}$ (100%)	$\frac{15}{16}$ (93.8%)	$\frac{15}{16}$ (93.8%)	$\frac{16}{16}$ (100%)
均衡化	$\frac{13}{13}$ (100%)	$\frac{10}{12}$ (83.3%)	$\frac{13}{13}$ (100%)	$\frac{10}{16}$ (62.5%)	$\frac{12}{16}$ (75%)	$\frac{10}{16}$ (62.5%)
添加回声	$\frac{6}{13}$ (46.2%)	$\frac{6}{12}$ (50%)	$\frac{8}{13}$ (61.5%)	$\frac{14}{16}$ (87.5%)	$\frac{15}{16}$ (93.8%)	$\frac{15}{16}$ (93.8%)
消除噪声	$\frac{12}{13}$ (92.3%)	$\frac{10}{12}$ (83.3%)	$\frac{13}{13}$ (100%)	$\frac{15}{16}$ (93.8%)	$\frac{13}{16}$ (81.1%)	$\frac{16}{16}$ (100%)
重新采样 (44.1-16-44.1 kHz)	$\frac{13}{13}$ (100%)	$\frac{12}{12}$ (100%)	$\frac{13}{13}$ (100%)	$\frac{16}{16}$ (100%)	$\frac{14}{16}$ (87.5%)	$\frac{16}{16}$ (100%)
重新采样 (44.1-22.05-44.1 kHz)	$\frac{13}{13}$ (100%)	$\frac{12}{12}$ (100%)	$\frac{13}{13}$ (100%)	$\frac{16}{16}$ (100%)	$\frac{13}{16}$ (81.3%)	$\frac{16}{16}$ (100%)

规攻击和去同步攻击的对比结果。从对比结果可以看出，本文算法实验结果整体好于文献[19-21]，进一步证明了本文算法具有良好的稳健性。

本文数字音频水印算法不仅具有较好的不可感知性，而且对常规音频信号处理（包括重采样、重量化、MP3 压缩、均衡化等）和去同步攻击（包括随机剪切、幅度伸缩、局部变调、抖动等）均具有较好的稳健性。

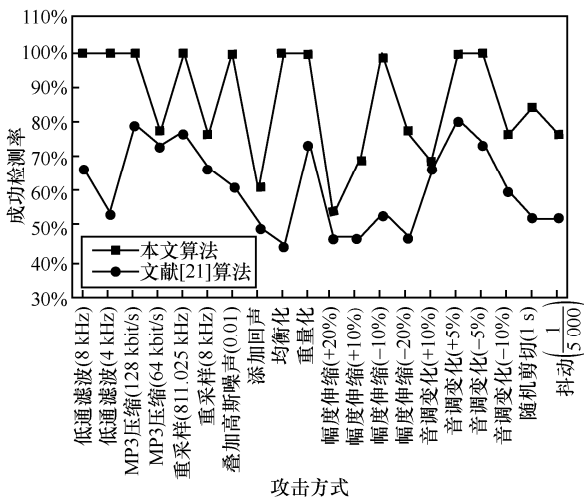


图 7 本文算法和文献[21]算法的稳健性能对比结果 ('Piano.wav')

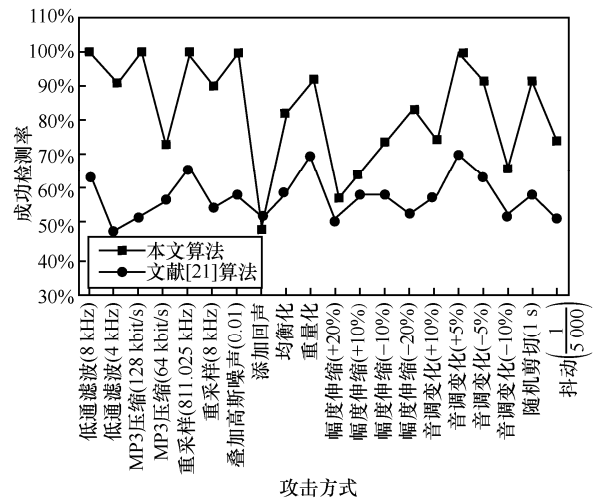


图 8 本文算法和文献[21]算法的稳健性能对比结果 ('Heart.wav')

6 结束语

数字水印是解决数字作品版权保护与内容认证的有效办法之一，稳健性是数字水印算法的重要性能指标之一。但现有绝大多数音频水印在经过去同步攻击后难以生存。因此，研究可有效抵抗去同步攻击的强稳健数字音频水印理论与方法仍是一项富有挑战性的工作。本文提出了一种基于稳健特征点的平稳小波域音频水印算法，实验结果表明，

参考文献:

- [1] QASIM A F, MEZIANE F, ASPIN R. Digital watermarking: applicability for developing trust in medical imaging workflows state of the art review[J]. Computer Science Review, 2018, 27: 45-60.
- [2] SU Z P, YUE F, JIANG J G, et al. SNR-constrained heuristics for optimizing the scaling parameter of robust audio watermarking[J]. IEEE Transactions on Multimedia, 2018, 20(10): 2631-2644.
- [3] KAUR A, DUTTA M K. An optimized high payload audio

- watermarking algorithm based on LU-factorization[J]. Multimedia Systems, 2018, 24(3): 341-353.
- [4] HUA G, HUANG J, SHI Y Q, et al. Twenty years of digital audio watermarking-a comprehensive review[J]. Signal Processing, 2016, 128: 222-242.
- [5] JIANG W Z, HUANG X H, QUAN Y J. Audio watermarking algorithm against synchronization attacks using global characteristics and adaptive frame division[J]. Signal Processing, 2019, 162: 153-160.
- [6] BASSIA P, PITAS I, NIKOLAIDIS N. Robust audio watermarking in the time domain[J]. IEEE Transactions on Multimedia, 2001, 3(2):232-241.
- [7] HU H T, HSU L Y. Robust, transparent and high-capacity audio watermarking in DCT domain[J]. Signal Processing, 2015, 109: 226-235.
- [8] HWANG M J, LEE J S, LEE M S, et al. SVD-based adaptive QIM watermarking on stereo audio signals[J]. IEEE Transactions on Multimedia, 2017, 20(1): 45-54.
- [9] TERCHI Y, BOUGUEZEL S. A blind audio watermarking technique based on a parametric quantization index modulation[J]. Multimedia Tools and Applications, 2018, 77(19): 25681-25708.
- [10] NADEAU A, SHARMA G. An audio watermark designed for efficient and robust resynchronization after analog playback[J]. IEEE Transactions on Information Forensics & Security, 2017, 12(6):1393-1405.
- [11] 唐鑫, 马兆丰, 钮心忻, 等. 基于变分贝叶斯学习的音频水印盲检测方法[J]. 通信学报, 2015, 36(1): 121-128.
TANG X, MA Z F, NIU X X, et al. Blind audio watermarking mechanism based on variational Bayesian learning[J]. Journal on Communications, 2015, 36(1): 121-128.
- [12] KHALDI K, BOUDRAA A O. Audio watermarking via EMD[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2009, 14(12): 2619-2622.
- [13] HU H T, CHANG J R, LIN S J. Synchronous blind audio watermarking via shape configuration of sorted LWT coefficient magnitudes[J]. Signal Processing, 2018, 147: 190-202.
- [14] WANG X Y, CUI Y R, YANG H Y, et al. A new content-based digital audio watermarking algorithm for copyright protection[C]// International Conference on Information Security(Infosecu'04). 2004: 62-69.
- [15] LI W, XUE X, LU P. Localized audio watermarking technique robust against time-scale modification[J]. IEEE Transactions on Multimedia, 2006, 8(1): 60-69.
- [16] MA X H, ZHANG B, DING X Y. Self-synchronization blind audio watermarking based on feature extraction and subsampling[C]// International Symposium on Neural Networks 2007, Lecture Notes in Computer Science. 2007: 40-46.
- [17] 鲍德旺, 杨红颖, 祁薇, 等. 基于音频特征的抗去同步攻击数字水印算法[J]. 中国图象图形学报, 2009, 14(12): 2619-2622.
BAO D W, YANG H Y, QI W, et al. A content based audio watermarking against desynchronization attacks[J]. Journal of Image and Graphics, 2009, 14(12): 2619-2622.
- [18] PUN C M, JIANG J J, LEE M C. Robust and self-synchronized audio watermarking by invariant feature points[J]. Information, 2012, 15(2): 815-829.
- [19] PUN C M, YUAN X C. Robust segments detector for de-synchronization resilient audio watermarking[J]. IEEE Transactions on Audio Speech and Language Processing, 2013, 21(11): 2412-2424.
- [20] YUAN X C, PUN C M, CHEN C L P. Robust mel-frequency cepstral coefficients feature detection and dual-tree complex wavelet transform for digital audio watermarking[J]. Information Sciences, 2015, 298: 159-179.
- [21] XIANG Y, NATGUNANATHAN I, PENG D, et al. Spread spectrum audio watermarking using multiple orthogonal PN sequences and variable embedding strengths and polarities[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2018, 26(3): 529-539.
- [22] 彭宏. 基于机器学习方法的数字音频水印技术研究[D]. 成都: 电子科技大学, 2010.
PENG H. Digital audio watermarking techniques based on machine learning method[D]. Chengdu: University of Electronic Science and Technology of China, 2010.
- [23] 张翼鹏, 陈亮, 郝欢. 采用量子神经网络的音频水印新算法[J]. 信号处理, 2013, 29(6): 684-690.
ZHANG Y P, CHEN L, HAO H. The new audio watermarking algorithm using quantum neural networks[J]. Signal Processing, 2013, 29(6): 684-690.
- [24] JLIL M, BUTT F A, MALIK A. Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals[C]//IEEE International Conference on Technological Advances. 2013: 208-212.
- [25] BITAR A W, DARAZI R, COUCHOT J. F. Blind digital watermarking in PDF documents using spread transform dither modulation[J]. Multimedia Tools and Applications, 2017, 76(1): 143-161.
- [26] ELLMAUTHALER A, PAGLIARI C L. Multiscale image fusion using the undecimated wavelet transform with spectral factorization and nonorthogonal filter banks[J]. IEEE Transactions on Image Processing, 2013, 22(3): 1005-1017.
- [27] DESSEIN A, CONT A. An information-geometric approach to real-time audio segmentation[J]. IEEE Signal Processing Letters, 2013, 20(4): 331-334.

[作者简介]



牛盼盼(1983-),女,辽宁大连人,博士,辽宁师范大学副教授,主要研究方向为网络信息安全、图像处理等。

杨思宇(1995-),女,辽宁铁岭人,辽宁师范大学硕士生,主要研究方向为信息隐藏与数字水印。

王丽(1996-),女,辽宁大连人,辽宁师范大学硕士生,主要研究方向为信息隐藏与数字水印。

杨红颖(1965-),女,辽宁大连人,辽宁师范大学教授,主要研究方向为多媒体信息安全、图像处理等。

李丽(1992-),女,辽宁大连人,辽宁师范大学硕士生,主要研究方向为图像处理与数字水印。

王向阳(1965-),男,辽宁大连人,辽宁师范大学教授,主要研究方向为网络信息安全、智能多媒体计算等。