

文章编号: 1673-3851 (2011) 03-0394-04

一种零水印广告监播算法

张 雷, 张华熊

(浙江理工大学信息电子学院, 杭州 310018)

摘 要: 提出一种基于音频特征值的零水印广告监播算法。在一定帧长下,经过 DA/AD 转换、MP3 压缩等攻击后的音频与原始音频波形的相似度很高,而经过小波变换后的音频最低频部分又是音频中能量最集中、最稳定的部分。基于以上特点,算法首先对原始音频进行三级小波变换,并提取出最低频部分。对最低频部分进行相应的转换,然后按一定帧长分段后保存为音频特征值。在实际的监播中,通过计算播出后音频的特征值与原始音频特征值的相似度值进行自适应判断来完成音频的同步和监播。本算法不需要在音频中嵌入任何水印,避免了复杂的水印嵌入和提取过程,对广告音频本身的品质没有任何影响。实验结果表明:该算法可以在 DA/AD 转换、MP3 压缩等攻击下有效地完成广告监播,具有很好的鲁棒性。

关键词: 音频特征值; 零水印; 广告监播; 小波变换; 相似度值; 鲁棒性

中图分类号: TP319.42 **文献标识码:** A

0 引 言

广告是人们最常见、也是运用最广泛的媒介形式。广播电视中会出现广告少播、广告不能按约定的数量播放等各种问题,所以对广告播出过程进行实时的不间断的监控,是保证广告准确播出的必要手段。而这些问题单纯靠人工来监控不但费时、费力,效果也不理想。

当前基于音频的广告监播算法可以分为两类:时域算法和频域算法。时域的典型算法有文献[1]提出的 LSB 算法和文献[2]提出的回声隐藏算法,以及文献[3]提出的时域能量算法等。这些算法的优点是简单易行,工作量小,缺点是鲁棒性差,无法满足正常广告监播的需要。频域的典型算法有文献[4]提出的离散傅立叶变换 DFT 和文献[5-6]提出的离散余弦变换 DCT 以及文献[7]提出的离散小波变换等。这些算法都是通过改变其频域的系数来嵌入秘密信息,具有较好的鲁棒性。鲁棒的音频信息隐藏算法是广告监播技术中的一个研究热点和难点,它要求在诸如 DA/AD 转换、噪声、低通滤波、重采样、MP3 压缩等各种攻击下,能够有效地完成广告监播工作。现在的鲁棒算法一般只能抵抗其中的几种攻击,特别是在抵抗 DA/AD 转换和 MP3 压缩方面,发表的论文很少,能同时抵抗这两种攻击的算法就更少了。

本文经过大量实验分析,发现经过 DA/AD 转换、MP3 压缩、低通滤波等处理以后的音频频域波形与原始音频频域波形也具有很高的相似性,在总结上述算法的优缺点的同时并结合现实中广告监播的需求,提出了一种全新的零水印广告监播算法。该算法的主要优点是:摒弃了以往复杂的水印嵌入和提取过程,实现了零水印监播;在提取特征值时对音频本身没有任何影响,工作量小,能够满足长时间的、高强度的广告监播要求;使用音频中稳定的最低频部分作为特征值进行广告监播,提高了算法的鲁棒性;在进行音频的同步和监播时,通过计算播出后音频特征值与原始音频特征值的相似度值进行自适应判断,大大提高了广告监播的准确度;可以在 DA/AD 转换、MP3 压缩等攻击下有效的完成广告监播。

1 零水印广告监播算法

零水印是指不修改原始数据的信息,通过利用原始数据的重要特征来构造一个水印。零水印方法是一种新的数字水印方法,很好地解决了不可见水印的不可感知和鲁棒性之间的矛盾。本文所提出的一种零水印广告监播算法是将某套节目中出现的某广告的音频特征信息提取出来,制作为该广告的音频特征值(音频特征水印),以该特征值为依据对后续广告音频进行连续自动识别,并将识别出的广告记录到数据库中,从而实现广告自动化监播过程。以往的大部分算法之所以很难有效抵抗 DA/AD 转换、MP3 压缩,因为它们研究的重点是水印的有效嵌入与有效提取,但嵌入的水印在经 DA/AD 转换或 MP3 压缩后很难保证其完整性,因此也就无法有效而又完整地提取出水印信息。本文算法是以经过 DA/AD 转换、MP3 压缩等攻击后的音频与原始音频波形的相似度很高为基础提出的,并使用音频中稳定的最低频部分做为特征值来完成广告监播,所以可以有效抵抗 DA/AD 转换、MP3 压缩。监播算法主要步骤如下。

a)特征水印获取

对原始广告音频进行三级小波变换,然后提取出最低频部分。假设最低频部分为 $X(l), l=0, 1, \dots, L-1$, 特征水印为 $SY(l), l=0, 1, \dots, L-1$ 。通过式(1)将最低频部分转换为特征水印。

$$SY(l) = \begin{cases} 1, & \text{if } X(l) > 0 \\ 0, & \text{if } X(l) \leq 0 \end{cases} \quad l=0, 1, \dots, L-1 \quad (1)$$

b)广告音频的同步及监播(如图 1 所示)

由文献[8]中的结论得知在时域中,音频在经过 DA/AD 转换后,帧长大于 1 000 时相同音频的相似度值始终保持在 0.8 附近。在经过 MP3 压缩后,帧长大于 1 000 时相同音频的相似度值始终保持在 0.99 以上。而在小波域中,通过大量的实验发现音频在经过 DA/AD 转换后,帧长大于 1 500 时相同音频的相似度值始终保持在 0.9 以上(经 MP3 压缩后的相似度值更高),具有很好的鲁棒性,而不同音频的相似度值始终保持在 0.5 左右,如图 2 和图 3 所示。同时,经小波变换后音频的样本点是服从相同分布的独立随机变量,由概率论的中心极限定理知道当最低频部分序列长度足够大时,对应的相似值近似符合高斯分布。

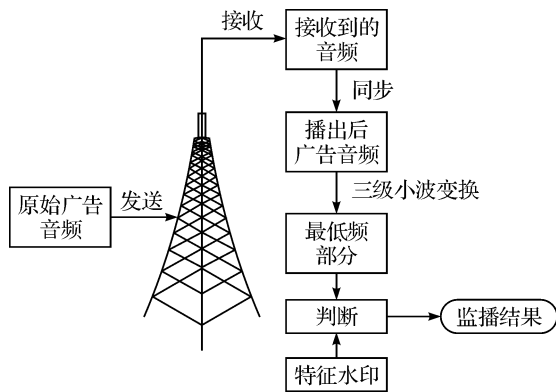


图 1 零水印广告监播流程

为了有效判断两音频是否是同一音频,本文设定一阈值 N 。由上述可以得知,相同音频在经过 DA/AD 转换或 MP3 压缩后频域波形的相似度值始终保持在 0.9 以上,而不同音频在经过 DA/AD 转换或 MP3 压缩后频域波形的相似度值始终保持在 0.5 左右。在实际应用中,一般用期望加 5 倍的标准差作为阈值 N 的值,这样就可以根据不同的音频样本进行自适应的判断,如果两段音频的相似度值大于 N ,则认为二者为同一音频,否则二者不是同一音频。

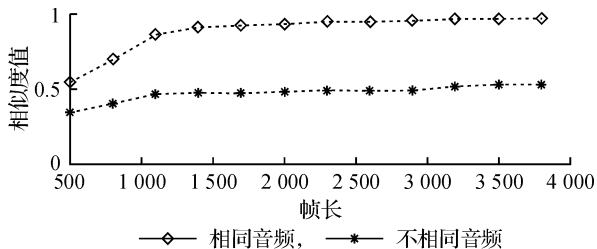


图 2 DA/AD 转换下不同帧长音频相似度值分布

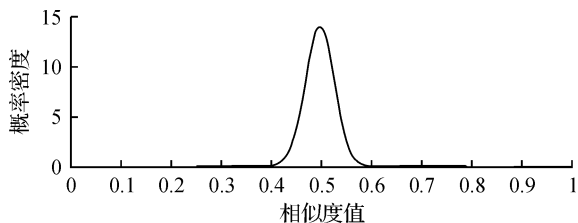


图 3 DA/AD 转换下音频相似度值概率分布

因此在同步时,先获取原始广告音频和接收到广告音频的最低频部分,然后通过下式求两者最低频部分开始一段的相似度值 M 。

$$M = \frac{\sum_{n_1=1}^{L_1} a(n_1)b(n_1)}{\sqrt{\sum_{n_1=1}^{L_1} a^2(n_1)} \times \sqrt{\sum_{n_1=1}^{L_1} b^2(n_1)}} \quad (2)$$

式(2)中, M 为两音频的相似度值, $a(n_1)$ 和 $b(n_1)$ 分别代表原始广告音频和接收到广告音频的最低频部分。因为广告监播不要求实时性, 所以在实际的监播中我们截取接收到的音频中前 3 s 的音频, 并对其进行三级小波变换得到其最低频部分, 直接从该最低频部分的开始位置向后逐点地求所有的相似度值 M , 然后找出其中最大的相似度值 M , 当最大的相似度值 M 大于阈值 N 时, 则该最大值 M 所对应时域中的点就是播出后广告音频的起始位置, 进而完成音频的同步工作。

同步完成后, 便可获取到播出后广告音频的最低频部分。经式(1)转换后, 对其和特征水印进行分段, 然后逐个求其与特征水印中对应段的相似度值。如果所有相似度值都大于阈值 N , 则证明广告音频完整播出了。如果相似度有小于阈值 N 的, 则证明广告音频没有完整播出。

2 实验与分析

实验中采用时长为 16 s、抽样速率为 44.1 kHz、16 bit 量化的单声道 WAV 格式音频。为了模拟 DA/AD 转换攻击, 在真实的环境中, 把原始广告音频和其他若干音频循环播放, 在抗 MP3 压缩攻击实验时, 先把 WAV 格式音频转换为 MP3 格式。然后通过无线调频发射机发送出去, 在另一端用音频接收器接收并得到转换后的音频。为了确保接收到的音频中能完整包含播出后的广告音频, 应提前 2 s 接收原始广告音频, 然后在播出结束时再延时 2 s 停止接收。

2.1 抗 DA/AD 转换攻击实验

先通过步骤 a) 获取特征水印。然后通过步骤 b) 完成音频的同步和监播。实验中采用期望加 5 倍的标准差来作为阈值 N 的值, 在进行音频同步时, 通过计算所有 M 值的期望和标准差得到 N 的值为 0.9127。找出 M 的最大值为 0.9834, 所以该 M 值所对应时域中的点就是播出后音频的起始点。在同步后分别进行了完整播出实验和不完整播出实验。如果原始广告音频完整的播出了, 则如图 4(a) 所示。如果没有完全播出, 本实验中少播了最后的 1.5 s, 则如图 4(b) 所示。

由步骤 b) 知, 在判断广告音频是否完整播出时, 要求出播出后广告音频的最低频部分和特征水印中对应每小段的相似度值, 试验中把特征水印和广告音频的最低频部分对应都分成 175 个小段。如果所有的 175 个相似度值都大于阈值 N , 则证明广告音频完整播出了, 否则证明广告音频没有完整播出。试验中, 在进行音频同步和完整性判断时理论上准确度达 99.84%, 对应的漏警率为 0.16%。

2.2 抗 MP3 压缩攻击实验

先通过步骤 a) 获取特征水印。然后通过步骤 b) 完成音频的同步和监播。实验中依然采用期望加 5 倍的标准差来作为阈值 N 的值, 在进行音频同步时, 通过计算所有 M 值的期望和标准差得到 N 的值为 0.9293。找出 M 的最大值为 0.9916, 所以该 M 值所对应时域中的点就是播出后音频的起始点。在同步后分别进行了完整播出实验和不完整播出实验。如果原始广告音频完整的播出了, 则如图 5(a) 所示。如果没有完全播出, 本实验中少播了最后的 1.5 s, 则如图 5(b) 所示。

由步骤 b) 知, 在判断广告音频是否完整播出时, 要求出播出后广告音频的最低频部分和特征水印中对应每小段的相似度值, 试验中把广告音频的最低频部分和特征水印对应都分成 175 个小段。如果所有的 175 个相似度值都大于阈值 N , 则证明广告音频完整播出了, 否则证明广告音频没有完整播出。试验中, 在

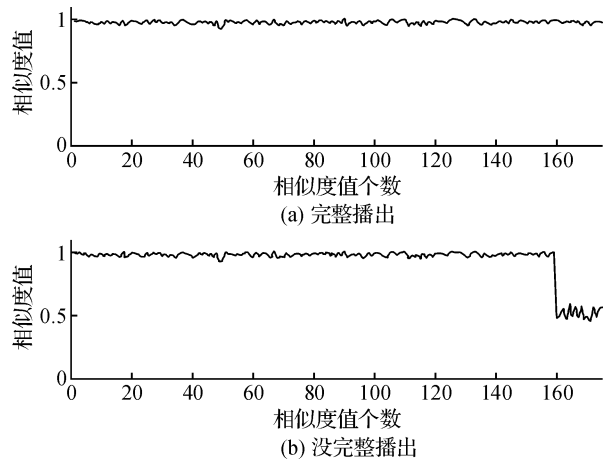


图4 抗 DA/AD 转换攻击效果

进行音频同步和完整性判断时理论上准确度达 99.97%, 对应的漏警率为 0.03%。

为了检查在不同强度噪音环境下的监播效果, 分别计算了信噪比为 50、45、40、35、30、25、20、15、10 dB 情况下的广告识别准确度, 如表 1 所示。

表 1 不同强度噪音环境下的监播效果表

信噪比/dB	DA/AD 转换		MP3 压缩	
	准确度/%	漏警率/%	准确度/%	漏警率/%
50	99.84	0.16	99.97	0.03
45	99.84	0.16	99.97	0.03
40	99.84	0.16	99.97	0.03
35	99.84	0.16	99.97	0.03
30	99.82	0.18	99.97	0.03
25	99.82	0.18	99.97	0.03
20	99.81	0.19	99.95	0.05
15	99.81	0.19	99.95	0.05
10	99.80	0.20	99.93	0.07

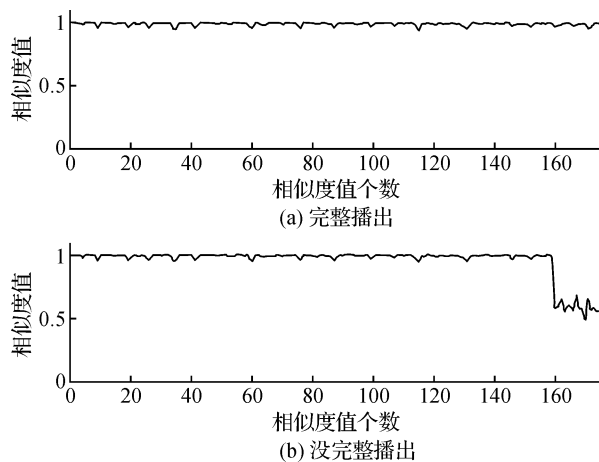


图 5 抗 MP3 压缩攻击效果

结果无论是抗 DA/AD 转换攻击, 还是抗 MP3 压缩攻击, 理论上准确度都达到了 99.8% 以上。可见基于零水印的广告监播算法具有很好的鲁棒性。

3 结 论

本文以音频波形经过各种转换和攻击后仍保持很高的相似度为基础, 设计一种以音频相似度值为判断标准的零水印广告监播方法。实验结果表明: 该方法不仅可以有效而精确地完成音频的同步工作, 也可以有效地判断广告是否完整的播出。本文的零水印广告监播方法不需要将水印直接嵌入到音频中, 避免了复杂的水印嵌入和提取过程, 对广告音频本身的品质没有任何影响, 且能抵抗住 MP3 解压缩和 DA/AD 转换, 具有很好的鲁棒性。

参考文献:

- [1] Cooperman M, Moskowitz S. Steganographic method and device: USA, 5687236[P]. 1997-11-11.
- [2] Kim H J, Choi Y H. A novel echo hiding scheme with backward and forward kernels[J]. IEEE Trans Circuits Syst, Video Technol, 2003, 13(7): 885-889.
- [3] Lie Wennung, Chang Lichun. Robust and highqualitytimedomain audio watermarking based on lowfrequency amplitude modification[J]. IEEE Transactions on Multimedia, 2006, 8(1): 46-59.
- [4] Scok J W, Hong J W. A novel audio watermarking algorithm for copyright protection of digital audio data[J]. ETRI Journal, 2002, 24(3): 181-189.
- [5] 马翼平, 韩纪庆. DCT 域音频水印: 嵌入对策和算法[J]. 电子学报, 2006, 34(7): 1260-1264.
- [6] 马翼平, 韩纪庆. 基于能量特性分块的 DCT 域自适应音频水印算法[J]. 信号处理, 2006, 22(4): 519-522.
- [7] 项世军. 稳健音频水印研究: 音频水印抗 D/A 和 A/D 变换的研究[D]. 广州: 中山大学, 2006.
- [8] 由守杰, 柏 森, 曹巍巍, 等. 一种抗 DA/AD 转换的音频信息隐藏算法[J]. 计算机工程与应用, 2008, 44(7): 113-116.

(下转第 420 页)