

文章编号: 1000-5013(2011)02-0153-03

一种基于支持向量机的数字音频认证方法

林晓丹

(华侨大学 信息科学与工程学院, 福建 泉州 362021)

摘要: 提出一种基于数字水印技术的音频内容认证方法. 选取稳定的梅尔倒谱特征构造特征向量, 利用支持向量机自适应地选取合适的帧进行水印的嵌入和提取. 结果表明, 算法在保证嵌入水印的不可感知性的同时, 能够有效地区分恶意的内容篡改和非恶意的常规信号处理操作, 准确地定位篡改发生的位置.

关键词: 数字音频; 水印技术; 认证; 梅尔倒谱系数; 支持向量机

中图分类号: TN 911.72; TN 912.3

文献标志码: A

数字音频认证包括精确认证和选择性认证, 前者要求需要认证的音频和原始音频完全一致. 但人们更关心的是音频所传达的信息是否真实, 是否遭受了篡改, 并不要求精确到每个数据位完全相同. 文献[1]用量化小波系数的方法嵌入水印, 检测时不需原始音频, 通过比较归一化相关系数判断音频是否遭受篡改. 文献[2]采用奇偶调制的方法对指数刻度下的傅里叶变换系数进行量化, 对某些信号处理操作具有鲁棒性, 对恶意篡改具有脆弱性. 然而, 大多认证算法未能利用多媒体信号自身特征以自适应地嵌入水印, 需要较多的人工干预, 导致算法性能过分依赖于参数的选择. 支持向量机(SVM)由于其良好的学习和泛化能力, 近年来开始被引入数字水印领域. 文献[3]利用支持向量机, 把水印的检测问题转化为一个支持向量机的二分类问题, 提高水印检测的性能. 文献[4]采用自适应量化的水印嵌入策略并将回归型支持向量机用于水印的检测, 实现水印鲁棒性和不可感知性的良好平衡. 本文利用支持向量机选择合适的帧在音频离散余弦变换(DCT)域中嵌入水印, 并将训练好的支持向量机用于水印的检测.

1 水印信息的构造

选用二值图像作为水印信息, 图像大小为 $M \times N$. 为了去除图像像素点之间的相关及增强认证系统的安全性, 对二值图像进行 Arnold 置乱. 将置乱后的图像 G 经过降维处理, 可以得到一维序列 $W_1 = \{g(l) = G(i, j), 1 \leq i \leq M, 1 \leq j \leq N, l \leq i \times M + j\}$. 用两个不相关的长度均为 c_r 的伪随机序列 $PN(0)$ 和 $PN(1)$ 对序列 W_1 扩频^[5], 若嵌入 0, 则映射成 $PN(0)$; 否则, 映射成 $PN(1)$. 即

$$W = w_j \mid w_j = \begin{cases} PN(0, k), & g(l) = 0, \\ PN(1, k), & g(l) = 1, \end{cases} \quad k = j \bmod c_r, 0 \leq j \leq c_r \times M \times N - 1.$$

式中, $PN(0, k)$ 代表 $PN(0)$ 的第 k 位, \bmod 表示取模运算, W 为最终要嵌入音频载体的水印信号.

2 构造特征向量

先对音频分帧, 计算各音频帧的 12 阶 Mel 系数. 然后, 取出第 1 阶和第 12 阶频率倒谱系数(MF-CC)^[6], 构造出特征向量 $p = \{M_{r_1}, M_{r_{12}}\}$. 其中: r 是帧号. 由音频掩蔽特性可知, 能量大的区域嵌入水印透明性好. 因此, 定义目标向量 d_r : 当 M_{r_1} 大, $M_{r_{12}}$ 大时, $d_r = 1$; 而当 M_{r_1} 小, $M_{r_{12}}$ 小时, $d_r = -1$. 将

收稿日期: 2009-12-19

通信作者: 林晓丹(1983-), 女, 助教, 主要从事多媒体信号处理及安全认证技术的研究. E-mail: echo.linxd@gmail.com.

基金项目: 华侨大学科研基金资助项目(09HZR12)

两类特征向量(样本数分别为 S_1 和 S_2)与目标向量 d_r ($1 \leq r \leq S_1 + S_2$)一起构成 $S_1 + S_2$ 组训练样本以训练支持向量机.

3 水印的嵌入与提取

对原始音频信号 A 分帧,每帧包含 N 个采样点,共 L 帧.计算各帧音频的 12 阶 MFCC,构造特征向量 $\delta_r = \{M_{r_1}, M_{r_{12}}\}$ ($1 \leq r \leq L$).将 δ_r 输入训练好的 SVM,得到预测输出 d_r .若 d_r 值为 1,则此帧被选中用于水印的嵌入;否则,不用于水印的嵌入.对被选中的音频帧作离散余弦变换(DCT)变换,选择 c_r 个中频系数作为水印的嵌入点.然后,根据 $c'_i = c_i + (2\omega_i - 1) \cdot \alpha$ 修改 DCT 系数.其中: α 为水印嵌入强度, ω_i 是扩频后需要嵌入的水印位.最后,对修改后的帧施行 DCT 逆变换,得到含水印的音频 A' .

水印提取不需要原始音频,能够实现盲检测.首先,将待检测音频按照嵌入时的长度分帧,计算各帧特征,得到各音频帧的特征向量.由于输入之前已训练好 SVM,若 SVM 预测输出的结果为 1,则表明该帧包含水印,需要在此帧中进行水印的提取;否则,跳过该帧.然后,对包含水印信息的帧进行 DCT 变换取出 c_r 个中频系数 $f_j(l)$ ($i=1, 2, \dots, c_r$), l 为帧号,计算水印嵌入分量 $f(l)$ 与伪随机序列 $PN(0)$, $PN(1)$ 的相关值 cor_0 和 cor_1 . 即

$$cor_0 = cor[f(l), PN(0)] = \frac{\sum_{j=1}^{c_r} f_j(l) \cdot PN(0, j)}{\sigma_{f(l)} \cdot \sigma_{PN(0)}},$$

$$cor_1 = cor[f(l), PN(1)] = \frac{\sum_{j=1}^{c_r} f_j(l) \cdot PN(1, j)}{\sigma_{f(l)} \cdot \sigma_{PN(1)}}.$$

提取水印时,不需设置检测阈值,只比较哪个相关值大即可,使算法的检测性能不受阈值设置的影响. $cor_0 > cor_1$, 提取出水印位“0”, 否则提取水印位“1”. 对选中的音频帧进行相同操作,提取各水印比特,再进行升维、反置乱处理,便可恢复嵌入时所使用的二值图像 G^* . 计算篡改评估函数 TAF , 有

$$TAF = \frac{1}{M \cdot N} \sum_{i=1, \dots, M} \sum_{j=1, \dots, N} G^*(i, j) \oplus G(i, j).$$

其中: G 为原始二值图像; \oplus 表示异或操作. 若 TAF 大于某个阈值则认证失败. $G^*(i, j) \oplus G(i, j) = 1$, 说明该比特发生错误, 对应于该比特的帧即为篡改发生区域.

4 仿真结果分析

实验选用长度 10.38 s 的音频信号,采样率 44.1 kHz, 16 位量化,单声道.选择 512 作为音频帧长,水印嵌入强度 $\alpha=0.1$, SVM 选用 RBF 核,参数 σ 取 0.01, TAF 检测阈值设为 0.3. 内容篡改后提取的水印图像,如图 1 所示.图 1(a)为嵌入的原始二值图像,大小为 24 px \times 24 px;图 1(b)为含水印的音频未受任何攻击时所提取的二值图像.为了验证算法对音频内容篡改的检测性能,剪切中间第 5~6 s 的音频,图 1(c)是提取出的二值图像;替换第 7~8 s 间的一段音频后,提取的二值图像如图 1(d)所示;在 8 s 时插入 10 帧其他音频后提取的二值图像,如图 1(e)所示.



(a) 原始 (b) 未受攻击 (c) 提取 (d) 替换 (e) 插入

图 1 内容篡改后提取的水印图像

Fig. 1 Extracted watermarks after audio tampering

当音频遭受恶意替换时,算法对篡改定位的结果如图 2 所示.图 2(a)中:嵌入了水印的音频峰值信噪比为 39.816 dB; φ 为归一化幅值.

此外,算法对常规信号处理操作具有一定的鲁棒性. 含水印的音频信号在不同常规信号处理操作后,提取的水印图像的误比特率(R_{BE})不同. $64 \text{ kbit} \cdot \text{s}^{-1}$ 的 MP3 压缩, R_{BE} 为 0; 截止频率为 11.025 kHz 的低通滤波, R_{BE} 为 1.56% ; 高斯噪声(均值为 0, 方差为 0.01), R_{BE} 为 0.35% ; 重采样(下采样到 22.05 kHz , 再还原成 44.1 kHz), R_{BE} 为 1.04% ; 重量化(从 16 bit 量化成 8 bit , 再量化成 16 bit), R_{BE} 为 0.

5 结束语

提出一种选择性音频认证方法, 利用支持向量机自适应选取合适的帧进行水印的嵌入和提取. 实验结果表明, 算法在保证嵌入水印的不可感知性的同时, 能够有效地区分恶意的内容篡改和非恶意的常规信号处理操作. 此外, 认证方法能够实现盲检测, 在需要对音频语义内容进行认证的应用中具有一定的实用价值.

参考文献:

- [1] 孙圣和, 王秋生. 数字音频信号的脆弱水印嵌入算法[J]. 计算机学报, 2002, 25(5): 520-525.
- [2] WU Hung-ping, KUO C C J. Fragile speech watermarking based on exponential scale quantization for tamper detection[C] // Proceedings of IEEE International Conference on Acoustic, Speech, and Signal Processing. Florida: IEEE, 2002: 3305-3308.
- [3] 王剑, 林福宗. 基于支持向量机(SVM)的数字音频水印[J]. 计算机研究与发展, 2005, 42(9): 1605-1611.
- [4] WANG Xiang-yang, QI Wei, NIU Pan-pan. A new adaptive digital audio watermarking based on support vector regression[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2007, 15(8): 2270-2277.
- [5] LANGELAAR G C, SETYAWAN I, LAGENDIJK R L. Watermarking digital image and video data[J]. IEEE Signal Processing Magazine, 2000, 17(5): 20-46.
- [6] PFEIFFER S, FISCHER S, EFFELBERG W. Automatic audio content analysis[C] // Proceedings of the 4th ACM Multimedia Conference. New York: ACM, 1997: 21-30.

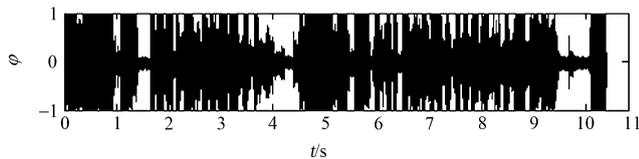
An SVM-Based Digital Audio Authentication Method

LIN Xiao-dan

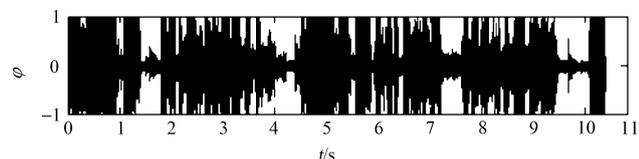
(College of Information Science and Engineering, Huaqiao University, Quanzhou 362021, China)

Abstract: A watermarking-based method for audio content authentication is proposed. Mel frequency cepstral coefficient (MFCC) is adopted to construct training vectors for support vector machine (SVM). Adaptive watermark embedding and extraction is achieved by the well-trained SVM. Experimental results demonstrate that this approach not only can distinguish malicious tampering from content-preserving operations, but also accurately locate regions that have undergone malicious manipulations while keeping the inaudibility of the embedded watermark.

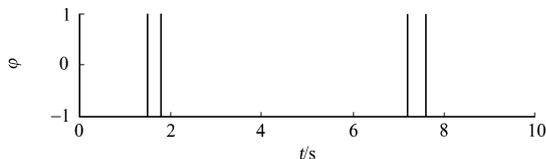
Keywords: digital audio; watermarking; authentication; Mel frequent cepstral coefficient; support vector machines



(a) 含水印的音频



(b) 篡改后的音频



(c) 篡改区域的定位

图 2 算法对篡改定位的结果

Fig. 2 Results of tamper detection