

文章编号: 1000-5862(2014)01-0108-03

基于音频特征的自适应数字盲音频水印算法

杨得国, 王荣萍

(西北师范大学数学与信息科学学院, 甘肃 兰州 730070)

摘要: 提出了一种基于音频特征的自适应盲音频水印算法: 使用 Logistic 混沌二进制序列对水印进行加密, 通过对音频帧的过零率及短时能量的分析, 选取适当的阈值, 确定嵌入水印的帧, 并对选定音频帧整合分段后进行小波变换, 同时采用能量动态调整自适应确定嵌入强度的方法, 计算每段音频的近似分量的平均能量, 结合基础嵌入强度, 得音频帧中嵌入水印. 实验结果表明: 该算法具有较好的鲁棒性和不可感知性, 具有一定的实用性.

关键词: 音频水印; 短时能量; 过零率; 自适应; 鲁棒性

中图分类号: TP 391.41

文献标志码: A

0 引言

近年来, 随着计算机网络和多媒体技术发展的越来越快, 音频数字水印^[1]正逐渐成为数字水印中的一个重要发展方向^[2]. 目前, 嵌入音频水印的基本做法有: 回声隐藏、相位编码^[3]等, 其中基于小波变换^[4]的数字音频水印技术因其具有良好的多分辨率、时域局部分析等特性, 在计算机技术上得到广泛应用.

Logistic 映射^[5-6]混沌二进制序列对原始水印信息进行加密, 增强水印的安全性. 混沌序列对初值的敏感性高、适应性强, 具有一般伪随机序列所没有的低通特性, 以抵抗低通滤波等攻击. 本文对音频特征^[7]进行分析, 结合音频帧的过零率和短时能量特性, 构造出适合水印嵌入的音频帧. 为了更好地提高音频水印的鲁棒性和去同步攻击能力, 本文在所选定的音频帧中采用了能量动态调整嵌入强度自适应^[8]的方法来嵌入水印. 同时本算法为零水印算法, 实现了盲检测, 提高了算法的不可感知性和鲁棒性.

1 音频特征的分析及选取

1.1 音频帧的过零率与短时能量特征

设第 n 帧音频信号的短时能量用 E_n 表示, 则 $E_n = \sum_{j=1}^F x_n^2(j)$, 其中 F 为信号帧长, E_n 为一帧内所有采样

点幅值的平方和, 它反映了音频信号的强度. E_n 越强的音频帧具有较好的掩蔽特性.

设第 n 帧音频信号的过零率为 Z_n , 则 $Z_n = \frac{1}{2} \sum_{j=1}^F |\text{sgn}[x_n(j)] - \text{sgn}[x_n(j-1)]|$, 其中 sgn 为符号函数, 即

$$\text{sgn}[x] = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0, \end{cases}$$

Z_n 为在一帧内相邻的 2 个采样点改变符号的次数, 它直接表明了音频信号在一帧中幅值变化的强弱程度. 过零率较小的音频帧具有较好的稳定性. 因此采取一定阈值下的过零率, 且取到能量的较大值, 就能更好地体现时域掩蔽特性.

2 数字音频水印的嵌入

2.1 嵌入水印图像的预处理

设原始水印图像为 U , 则 $U = \{u(i, j) | 1 \leq i \leq M, 1 \leq j \leq N, u(i, j) \in \{0, 1\}\}$. 若将 2 维的图像嵌入到 1 维的音频中, 则首先将水印图像降维至 1 维序列 V 中, 则 $V = \{v(k) = u(i, j) | 1 \leq i \leq M, 1 \leq j \leq N, k = (i-1) \times N + j, 1 \leq k \leq M \times N\}$.

采用 Logistic 映射混沌二进制序列对水印进行混沌置乱加密. 首先对 Logistic 序列进行截断, 选取任意起始位置 a 开始截取, 设定阈值 λ , 初值 x_0 , 保留水印大小的数据对其进行处理, 得到 Logistic 序列 P 为: $P = \{p(k) | 1 \leq k \leq M \times N, p(k) \in \{0, 1\}\}$. 加

收稿日期: 2013-09-16

基金项目: 国家自然科学基金(61165002)资助项目.

作者简介: 杨得国(1971-), 男, 甘肃民勤人, 教授, 主要从事网络与多媒体的研究.

密后的水印序列 $W: W(k) = V(k) \oplus P(k)$, 其中 $V(k), P(k) \in \{0, 1\}, 1 \leq k \leq M \times N$. 如图 1 所示.



图 1 水印图像

2.2 原始音频信号的预处理

设音频信号 $A = \{a(k), 1 \leq k \leq L\}$, L 是音频信号长度. 将 A 进行分帧后表示为 $a_{kj} = a(F(k-1) + j), 1 \leq k \leq K, 1 \leq j \leq F$, 其中 a_{kj} 为第 k 帧音频的第 j 个信号值. K 为音频信息的总帧数, F 为每帧音频的样点数.

2.3 音频帧的选取

设置 2 个阈值 E_0, Z_0 , 选取满足条件: 过零率 $Z \leq Z_0$ 且短时能量 $E \geq E_0$ 的音频帧. 利用能量动态调整自适应方法实现水印的嵌入. 根据实验数据可知, 对于每一种类型的音频信号, 选取过零率 $Z_n \leq Z_0$ 并且短时能量 $E_n \geq E_0$ 的音频帧. 然后把音频帧连接到一起得到待嵌入水印的音频帧 Ae , 并将它分成 $M \times N$ 段. 那么每段含有 $Q = L/(M \times N)$ 个音频数据. 记为 $Ae = \{Ae(k), 1 \leq k \leq Q\}$.

2.4 数字音频水印的嵌入

对分段后的音频信号 $Ae(k)$ 分别进行 H 层小波变换^[9], 例如第 K 段音频信号进行 DWT 操作后可得到小波系数为: $Ae_K^H, De_K^H, De_K^{H-1}, \dots, De_K^1, Ae_K^H$ 是 H 层小波变换的低频系数. 本算法中的水印信息将被嵌入到低频系数中.

$$Ae_K^H = Ae_K^H [1 + \alpha km(i)], \quad (1)$$

其中 αk 为嵌入水印强度. 在现有音频水印嵌入方案中, 固定的嵌入强度比较常见. 但是这种固定性存在较大的缺陷性. 如果水印嵌入时强度越大, 那么受攻击时鲁棒性就越好, 对信号的影响就越大. 但是, 如果嵌入时强度越小, 这样不易被察觉, 但是抗干扰性就差. 本算法针对这一问题, 通过计算经过小波变化后每段音频的近似分量系数的平均能量, 再结合嵌入强度 β , 动态调整音频数据段嵌入强度这一方法, 为每个音频数据段确定一个嵌入强度.

$$\alpha k = \left(\sum_{k=1}^Q Ae(k)^2 \right) / Q + \beta, \quad (2)$$

按照 (2) 式得到修改后的小波系数. 再将其进行小波逆转换. 这样就可以得出嵌入水印的数字音频.

3 数字音频水印的获取

数字音频水印的获取按照以下步骤进行:

(i) 将嵌入水印后的音频信号进行 H 层的小波变换操作, 得到每段的低频系数 Ae_K^H ;

(ii) 计算每个音频数据段的嵌入强度 αk , 提取水印信息. 提取公式为

$$m(i) = |Ae_K^H - Ae_K^H| / \alpha k;$$

(iii) 按照初始位置、阈值, 初步得到 Logistic 序列. 根据阈值对它进行二值化处理, 得到 $P(k)$. 那么 1 维水印序列为 $V(k) = W(k) \oplus P(k)$. 然后将它提升为 2 维就可得到原始水印图像 $U = \{u(i, j), 1 \leq i \leq M, 1 \leq j \leq N, u(i, j) \in \{0, 1\}\}$.

4 仿真实验与结果分析

本文使用比特误码率 BER 和归一化相关系数 NC ^[10] 来度量算法的性能.

$$NC(W', W) = \frac{\sum_{i=1}^M \sum_{j=1}^N w(i, j) \times w'(i, j)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N w^2(i, j)} \sqrt{\sum_{i=1}^M \sum_{j=1}^N w'^2(i, j)}},$$

$$BER = \frac{\text{错误的比特数}}{\text{总比特数}} \times 100\%,$$

其中 BER 用来表示提取的水印与原始水印的误差程度. NC 用来表示提取的水印与原始水印的相似程度进行度量.

性能测试如图 2 所示, 分别表示为原始音频信号、嵌入水印后的音频信号、原始水印图像、从含水印的音频信号中提取出的水印图像.

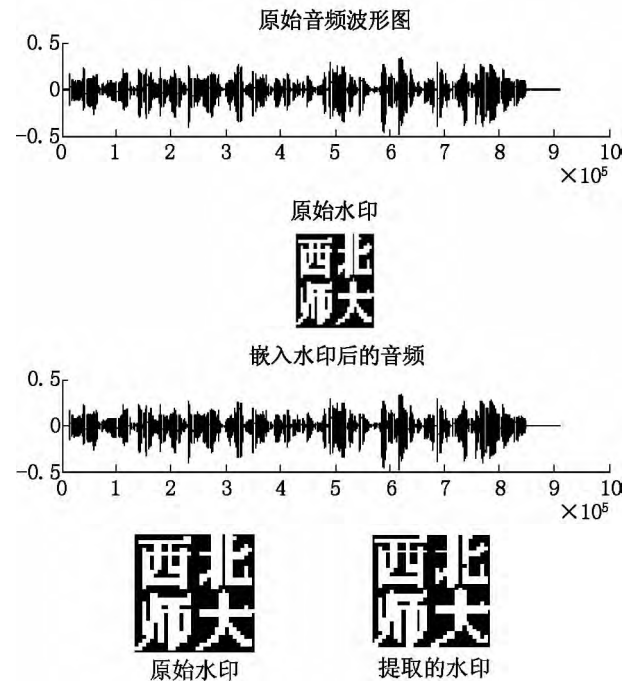




图 2 原始音频与嵌入水印音频及原始水印与提取的水印对比

表 1 为音频信号受各种攻击后提取的水印图像、相关系数 NC 和误码率 BER 的值,其中引入 50 ms 的回声干扰;采用截止频率为 8 kHz 低通滤波、最大衰减为 2 dB 的 Chebyshev I 型低通滤波器;重采样先将信号进行一次抽取和插值处理,抽取和插值的系数为 2;重量化是首先将 16 bit 的音频信号量化为 8 比特,然后再量化到 16 bit;添加噪声是在音频信号中加入均值为 0,方差为 0.01 的高斯白噪声;MP3 压缩是首先将音频信号压缩至 128 kB/s,然后再加压缩为 WAV 音频格式。

表 1 提取出的水印图像及其实验结果

	无攻击	低通滤波	重采样	重量化
NC	1	0.987 300	0.998 500	0.992 500
$BER/\%$	0	0	0	0
水印 图像				
	回声干扰	添加噪声	MP3 压缩	
NC	0.997 600	0.836 300	0.848 300	
$BER/\%$	0.191 460	0.974 254	8.540 690	
水印 图像				

5 结束语

该算法首先对水印序列进行混沌序列置乱,然后通过音频特征^[11]的过零率和短时能量选出需要

嵌入水印的音频部分,对新得到的音频分段后通过采用求每段的平均能量以动态调整嵌入强度从而对于固定嵌入强度所带来问题作了改进。另外,算法在提取水印时不需要原始音频信号。实验表明,该算法对一般常见攻击测试具有较好的鲁棒性和不可感知性。

6 参考文献

- [1] Petitcolas F A P, Anderson R J, Kuhn M G. Information hiding: A survey [J]. *Proc of the IEEE*, 1987(7): 1062-1078.
- [2] Lian S, Kanellopoulos D, Ruffo G. Recent advances in multimedia information system security [J]. *Informatica*, 2009, 33(1): 3-24.
- [3] 汪飞, 檀结庆. 基于 DWT 和均值量化的音频水印算法 [J]. *计算机应用*, 2009, 29(2): 444-446.
- [4] 赵红, 崔永瑞, 郑素娥. 小波域音频水印综述 [J]. *计算机安全*, 2007, 13(9): 1-3.
- [5] 张敏, 杨慧. 基于混沌加密的盲音频水印算法 [J]. *中国民航学院学报*, 2005, 23(3): 37-41.
- [6] 陈雪松, 张寒, 杨勇田. 基于混沌理论的音频数字水印算法研究 [J]. *计算机工程与应用*, 2007, 43(24): 59-60.
- [7] 彭宏, 王珣. 基于音频特征的多小波域水印算法 [J]. *计算机研究与发展*, 2010, 47(2): 216-222.
- [8] 康旭, 张秋余, 袁占亭. 基于能量自适应的同步小波音频水印算法 [J]. *科学技术与工程*, 2008, 8(23): 6240-6243.
- [9] 彭宏, 王珣等. 基于音频特征的多小波域水印算法 [J]. *计算机研究与发展*, 2010, 47(2): 216-222.
- [10] 李伟. 鲁棒性数字音频水印算法研究 [D]. 上海: 复旦大学, 2004.
- [11] 杨得国, 姜金娣等. 基于音频特征和逼近信号统计特征的数字零水印算法 [J]. *江西师范大学学报: 自然科学版*, 2011, 35(6): 633-636.

The Adaptive Blind Digital Audio Watermarking Algorithm Based on Audio Features

YANG De-guo, WANG Rong-ping

(Department of Mathematics and Information Science, Northwest Normal University, Lanzhou Gansu 730070, China)

Abstract: An adaptive blind audio watermarking algorithm based on audio features has been proposed. The algorithm uses the Logistic binary chaotic sequence of the watermark is encrypted, through the analysis and the short-time energy zero rate of the audio frame, select the appropriate threshold, is used to determine the embedding frame, and the wavelet transform of the sub audio frames selected after integration, method of simultaneous determination of embedding strength by dynamically adjusting the adaptive energy, average energy calculation of the approximate component of each audio, combining the embedding strength, embedding watermark in the audio frame. Experiments show that the algorithm has good robustness, has a certain practicality.

Key words: audio watermarking; zero watermark; short-time energy; zero-cross ratio; adaptively; robust

(责任编辑: 冉小晓)