

基于音频特征的抗去同步攻击数字水印算法

鲍德旺¹⁾ 杨红颖¹⁾ 祁薇¹⁾ 王向阳^{1),2)}

¹⁾(辽宁师范大学计算机与信息技术学院,大连 116029) ²⁾(中国科学院软件研究所信息安全国家重点实验室,北京 100039)

摘要 抗去同步攻击的强鲁棒数字音频水印方法研究是一项富有挑战性的工作. 本文结合数字音频自身特征,提出了一种基于音频内容的抗去同步攻击数字水印算法. 该算法首先根据数字音频的局部能量特征,从原始载体中提取出稳定的特征点;然后以音频特征点为标识,确定用于水印嵌入的候选音频段;最后采纳量化调制策略,将数字水印嵌入到音频载体内. 进行数字水印检测时,系统通过分析音频内容提取特征点,再以特征点为标识提取水印信息,水印检测无需原始音频信号参与. 仿真实验结果表明,本文算法不仅具有较好的不可感知性,而且对常规信号处理(MP3压缩、低通滤波、添加噪声等)和去同步攻击(随机剪切、幅度缩放、时间延展、抖动等)均具有较好的鲁棒性.

关键词 音频水印 去同步攻击 特征点 局部能量

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2009)12-2619-04

A Content Based Audio Watermarking Against Desynchronization Attacks

BAO De-wang¹⁾, YANG Hong-ying¹⁾, QI Wei¹⁾, WANG Xiang-yang^{1),2)}

¹⁾(School of Computer and Information Technology, Liaoning Normal University, Dalian 116029)

²⁾(State Key Laboratory of Information Security, Institute of Software of Chinese Academy of Sciences, Beijing 100039)

Abstract It is a challenging work to design a robust audio watermarking scheme against desynchronization attacks. In this paper, a feature-based digital audio watermarking scheme robust against desynchronization attacks is proposed. Firstly, the steady feature points are extracted from the host audio according to the local energy distribution. Then, the candidate audio segments for embedding watermark are defined by using the feature points. Finally, the digital watermark is embedded into a host audio by modulating the statistics average value of audio samples. In digital watermark detection, the feature points are selected by the same technique as the embedding, and the digital watermark is extracted from the watermarked audio after the feature points. Experimental results show that the proposed scheme is inaudible and robust against common signals processing such as MP3 compression, low-pass filtering, noise addition, and equalization etc, and is robust against desynchronization attacks such as random cropping, amplitude variation, time-scale modification, and jittering etc.

Keywords audio watermarking, desynchronization attack, feature point, local energy

1 引言

数字音频水印技术发展到今天,已有大量不同算法,它们广泛提出了“鲁棒性”声明.但遗憾的是,现有绝大多数数字音频水印算法仅仅能够对抗常规

的信号处理^[1-2]. Wang等人以回归型支持向量机(SVR)理论为基础,结合同步码重定位技术,提出了一种可同时抵抗常规信号处理和部分轻微去同步攻击的数字音频水印算法^[3],但其抵抗联合攻击(如叠加噪声+时间拉伸)能力较差. Wei等人结合人类心理模型理论,提出了一种基于音频内容分析的

基金项目:国家自然科学基金项目(60773031,60873222);中国科学院软件研究所信息安全国家重点实验室开放基金项目(03-06);北京大学视觉与听觉信息处理国家重点实验室开放基金项目(0503);南京大学计算机软件新技术国家重点实验室开放基金项目(A200702)

收稿日期:2009-06-11; **改回日期:**2009-09-18

第一作者简介:鲍德旺(1984~),男.辽宁师范大学硕士研究生.主要研究方向为信号处理、音频水印. E-mail:baodwang@126.com

DCT 域数字水印算法^[4],该算法选取音频载体的能量峰值点作为音频特征,并以此确定水印嵌入位置,而水印嵌入是在 DCT 域内完成的。Ma 等人以独立分量分析(ICA)理论为基础,提出了一种基于特征点提取与子采样的盲音频水印方法^[5],该方法同样选取音频能量峰值点作为特征标识并依此确定水印嵌入位置,而水印信息被嵌入到子采样后的载体音频内。该算法具有一定的抵抗去同步攻击能力,但对叠加噪声等常规信号处理却很敏感。Mansour 等人首先提取出原始音频的包络信息,进而对音频包络实施小波变换并利用极值方法提取出音频特征点,最后通过改变音频特征点间距离等措施嵌入数字水印^[6]。该方法具有一定的抗时间拉伸、音高位移等去同步攻击能力,但无法抵抗随机剪切、联合攻击等。Li 等人首先计算出原始音频包络,然后选取音频包络的所有峰值点作为音频特征点,继而以特征点为标识从原始载体音频中选取 4 096 个局部特征区域用于水印嵌入,最后利用 FFT 变换在每个局部特征区域内嵌入数字水印^[7]。该方法可以抵抗轻微的时间拉伸等去同步攻击,但对重新量化、均衡化等常规信号处理较敏感。

本文结合数字音频自身特征,提出了一种基于音频内容的抗去同步攻击数字水印算法。同时,仿真实验结果也证明了本文算法的有效性。

2 音频特征点提取

对数字音频水印技术而言,总希望将数字水印嵌入到音频信号能量较强处,因为:(1)由于载体信号能量较强,故嵌入水印能量可相应增加,从而提高水印检测可靠性;(2)由于高能量处是音频信号的主要部分,对它的恶意攻击如果强度过大,则会破坏信号的品质,因此可一定程度限制对水印的攻击,进而增强水印的稳健性。这里音频特征点指的是音频信号中能量值急速爬升到峰值的点。由于能量值的急速爬升反映了信号短时能量的大小,因此该点也是信号局部能量较强部分的起始点。将这样的点作为水印嵌入的起始点符合将水印嵌入信号能量较强部分的要求。音频特征点的具体检测过程如下:

(1)对采样点数为 N 的数字音频信号 $x(n)$ ($n=0,1,\dots,N-1$),由式(1)分别计算各个采样点向前 L 个点及向后 L 个点的短时能量总和。

$$\begin{aligned} M_b(n) &= \sum_{i=1}^L x^2(n-i) \\ M_a(n) &= \sum_{i=0}^{L-1} x^2(n+i) \end{aligned} \quad (1)$$

式中, $n=L+1,L+2,\dots,N-L+1$, M_b 和 M_a 分别代表前向短时能量总和与后向短时能量总和。

(2)利用式(2)计算采样点 n 前后短时能量总和改变的比值 $r(n)$ 。

$$r(n) = \frac{M_a(n)}{M_b(n)} \quad (2)$$

(3)计算 $r(n)$ 的局部极大值。若 $r(n)$ 同时满足以下 4 个条件

$$\begin{cases} r(n) > r(n-1) \\ r(n) > r(n+1) \\ R(n) > T_1 \\ M_a(n) > T_2 \end{cases} \quad (3)$$

则称 $r(n)$ 为局部极大值。其中,第 4 个条件中的 T_2 是短时能量总和的门限值,设置该门限是为了防止 M_b 接近于 0 而将某采样点标记为符合条件的点。本文将 $r(n)$ 取局部最大值的采样点作为音频特征点。

3 数字水印的嵌入

假设原始数字音频信号为 $\mathbf{A} = \{a(i), 0 \leq i < Length\}$ 。其中, $Length$ 为音频数据的个数, $a(i) \in \{0,1,2,\dots,(2^p-1)\}$ 是第 i 个音频数据的幅度值, p 表示每个数据所使用的比特数。

假设二值水印图像为 $\mathbf{W} = \{w(i,j), 0 \leq i < P, 0 \leq j < Q\}$ 。其中, $w(i,j) \in \{0,1\}$ 代表二值水印图像的第 i 行、第 j 列像素值。则该音频水印嵌入方案可描述如下。

3.1 数字水印的预处理

为了消除二值水印图像的像素空间相关性,提高整个数字水印系统的安全性能,确保数字音频某一部分受到破坏后仍能全部或部分地恢复水印,宜首先对二值水印图像进行置乱加密。为此,首先引入 Arnold 变换,将二值水印图像 \mathbf{W} 置乱加密为安全水印矩阵 \mathbf{W}_1 。

接下来,将安全水印矩阵 \mathbf{W}_1 降维成 1 维二进制数字水印序列 \mathbf{W}_2 。

3.2 非下采样小波变换及低频信息重构

对原始载体音频 A 实施非下采样小波变换,并利用所得到的低频信息重构出新的载体音频 \tilde{A} 。

3.3 新载体音频的特征点提取

将新载体音频 \tilde{A} 均分成 K 个音频数据段,并从每个音频数据段中提取出音频特征点(参见第 2 节)。于是可获得分布较为均匀的新载体音频 \tilde{A} 特征点 $p_k (k=0,1,\dots,K-1)$ 。

3.4 水印信号的嵌入

以新载体音频 \tilde{A} 特征点为标识(起始点标识),选取长度为 $P \times Q \times s$ 的音频数据段 PA 用于水印嵌入。这里, s 为常数(本文选取为 5)。数字水印嵌入步骤为

(1) 将所选取的音频数据段 PA 按数字水印大小 $P \times Q$ 划分成 $P \times Q$ 段,每一段 $PA(m)$ 含有 s 个音频样本,即

$$PA(m) = \{pa(m)(i), 0 \leq i < s, 0 \leq m < P \times Q\} \quad (4)$$

(2) 计算 $PA(m)$ 的平均值,即

$$\overline{PA(m)} = \frac{1}{s} \sum_{i=0}^{s-1} pa(m)(i) \quad (5)$$

$$(0 \leq m < P \times Q)$$

(3) 采用量化方法嵌入数字水印,即对每一段 $PA(m)$, 修改其均值 $\overline{PA(m)}$, 以嵌入一位水印信息。修改策略为

$$pa'(m)(i) = pa(m)(i) + (\overline{PA'(m)} - \overline{PA(m)}) \quad (6)$$

式中, $PA(m) = \{pa(m)(i), 0 \leq i < s\}$ 为修改前音频样本值, $PA' = \{pa'(m)(i), 0 \leq i < s\}$ 为修改后的音频样本值,且有

$$\overline{PA'(m)} = \begin{cases} IQ(\overline{PA(m)}) \times S_1 + S_1/2, & \text{如果 } Q(\overline{PA(m)}) = w_2(m) \\ IQ(\overline{PA(m)}) \times S_1 - S_1/2, & \text{如果 } Q(\overline{PA(m)}) \neq w_2(m) \\ IQ(\overline{PA(m)}) = [\overline{PA(m)}/S_1], & \\ Q(\overline{PA(m)}) = \text{mod}(IQ(\overline{PA(m)}), 2) \end{cases}$$

这里, S_1 为量化步长。

3.5 逆非下采样小波变换

结合原始音频小波域高频信息进行逆非下采样小波变换,即可得到含有水印的音频信号。

4 数字水印的检测

抗去同步攻击数字音频水印算法属于盲水印算法,即检测数字水印信号时不需要原始音频载体的参与。数字水印的检测过程如下:

(1) 对待检测音频 A^* 实施非下采样小波变换,并利用所得到的低频信息重构出新的载体音频 \tilde{A}^* 。

(2) 将新载体音频 \tilde{A}^* 均分成 K 个音频数据段,并从每个音频数据段中提取出音频特征点(参见第 3 节)。于是可获得分布较为均匀的新载体音频 \tilde{A}^* 特征点 $p_k^* (k=0,1,\dots,K-1)$ 。

(3) 以新载体音频 \tilde{A}^* 特征点为标识(起始点标识),选取长度为 $P \times Q \times s$ 的音频数据段 PA^* 用于水印提取。

(4) 将所选取的音频数据段 PA^* 按数字水印大小 $P \times Q$ 划分成 $P \times Q$ 段,每一段 $PA^*(m)$ 含有 s 个音频样本。

(5) 计算 $PA^*(m)$ 的平均值 $\overline{PA^*(m)}$, 并按如下方法提取水印信息:

$$w_2^*(m) = \begin{cases} 1 & \overline{PA^*(m)} > 0 \\ 0 & \overline{PA^*(m)} \leq 0 \end{cases} \quad (7)$$

式中, $m = 0, 1, \dots, P \times Q - 1$ 。

(6) 对二进制序列 W_2^* 进行升维处理,并进行逆置乱解密,即得到所提取的二值图像水印 W^* 。

此外,为了消除观测者的经验、身体条件、实验条件等主客观因素的影响,需采用归一化相关系数(normalized cross-correlation, NC),对提取出的水印和原始水印的相似性进行定量评价。如果该归一化相关系数 NC 超过某一检测阈值,我们就判定待检测音频信号中存在数字水印,否则不含数字水印。

最后,选取归一化相关系数 NC 取值最大的数字水印作为最终水印信号检测结果。

5 仿真实验结果

为了验证本文音频水印算法的高效性,以下给出了抗攻击能力测试的实验结果。实验中,所选用的原始音频载体是采样频率为 44.1 kHz、分辨率为 16 比特的单声道数字音频信号,数字水印采用了 32×32 的二值图像。小波变换采用了常见的 Daubechies-1

小波基,非下采样小波变换级数选取为 $H = 3$, 数字水印检测阈值设置为 0.6, 参数 $K = 10$ 。

另外,还采用归一化相关系数(NC)定量分析了提取水印与原始水印的相似度,采用峰值信噪比

(PSNR)评价了原始数字音频信号与含水印数字音频信号之间的差别。表 1 给出了本文数字水印嵌入方案的抗攻击能力实验结果(包括数字水印的归一化相关系数 NC、数字音频的峰值信噪比 PSNR)。

表 1 数字水印对各种攻击的抵抗能力

Tab. 1 The watermark detection results against various attacks

	未攻击	重新量化	重新采样 22 050 Hz	重新采样 11 025 Hz	高斯噪声	低通滤波 4 kHz	MP3-256k
NC	1.00	0.88	0.99	0.92	0.99	0.96	1.00
PSNR	41.95	30.83	31.04	35.20	40.51	39.50	41.98
	MP3-112k	MP3-64k	回声	均衡化	TSM (+1%)	TSM (+2%)	TSM (-1%)
NC	1.00	0.96	1.00	0.99	1.00	1.00	1.00
PSNR	42.06	41.83	38.93	41.08	27.26	26.82	27.59
	剪切前面 1s	剪切中间 1s	前面加 1s	幅度 放大 150%	幅度 缩小 50%	抖动攻击 (1/5 000)	抖动攻击 (1/10 000)
NC	1.00	1.00	1.00	0.96	0.96	0.98	1.00
PSNR	27.25	29.68	29.83	40.75	38.15	27.76	28.40

6 结 论

结合数字音频自身特征,提出了一种基于音频内容的抗去同步攻击数字水印算法。该算法首先根据数字音频的局部能量特征,从原始载体中提取出稳定的特征点;然后以音频特征点为标识,确定用于水印嵌入的候选音频段;最后采纳量化调制策略,将数字水印嵌入到音频载体内。仿真实验表明,该算法不仅具有较好的不可感知性,而且对常规信号处理和去同步攻击均具有较好的鲁棒性。

此外,提出的数字音频水印嵌入算法还具有计算简单、容易实现、抽取水印时无需原始载体音频等特点,这大大增强了其用于数字音频作品版权保护的实用性。

参考文献 (References)

1 Lian S, Kanellopoulos D, Ruffo G. Recent advances in multimedia

information system security[J]. Informatica, 2009, 33(1): 3-24.

2 Cvejic N, Seppänen T. Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks[M]. Hershey, PA, USA: Information Science Reference, 2007.

3 Wang X Y, Qi W, Niu P P. 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.

4 Foo Say Wei, Xue Feng, Li Mengyuan. A blind audio watermarking scheme using peak point extraction[A]. In: Proceedings of IEEE International Symposium on Circuits and Systems[C], Kobe, Japan, 2005, 5: 4409-4412.

5 Ma Xiao-hong, Zhang Bo, Ding Xiao-yan. Self-synchronization blind audio watermarking based on feature extraction and subsampling[A]. In: Proceedings of the International Symposium on Neural Networks 2007[C], Nanjing, China, 2007: 40-46.

6 Mansour M F, Tewfik A H. Data embedding in audio using time-scale modification[J]. IEEE Transactions on Speech and Audio Processing, 2005, 13(3): 432-440.

7 Li W, Xue X Y, Lu P Z. Localized audio watermarking technique robust against time-scale modification[J]. IEEE Transactions on Multimedia, 2006, 8(1): 60-69.