

基于神经网络的扩展回声隐藏算法

王慧琴¹, 修可山², 姚钟涵¹

(1. 西安建筑科技大学信息与工程学院, 西安 710055; 2. 西安交通大学电信学院, 西安 710049)

摘 要: 提出了一种神经网络自适应扩展回声隐藏算法。利用 PN 序列将音频信号的单回声内核进行扩展后作为水印信号, 提高了水印算法的安全性。该算法利用了神经网络的非线性映射能力确定扩展回声内核的幅值, 从而避免了复杂的心理声学模型的计算过程, 实现了水印嵌入的强度的自适应。仿真实验证明了该算法的有效性和可靠性。

关键词: 音频; 数字水印; 神经网络; 自适应; 心理声学模型

Spread Echo Hiding Algorithm Based on Neural Network

WANG Huiqin¹, XIU Keshan², YAO Zhonghan¹

(1. School of Information and Control Engineering, Xi'an University of Architecture & Technology, Xi'an 710055;

2. School of Electronic & Information Engineering, Xi'an Jiaotong University, Xi'an 710049)

【Abstract】 This paper proposes a neural network self-adaptive digital audio watermarking algorithm based on time-spread echo hiding algorithm. The single echo kernel of audio signal is spread by pseudo noise (PN) sequence and taken as watermark signal. Therefore, the algorithm achieves higher robustness and secrecy merits. By exploiting the abilities of neural networks and considering the characteristics of human audio system (HAS), a just noticeable differences (JDN) threshold controller is designed to ensure the strength of the embedded data adapting to the host audio itself entirely. The simulation experiment results show that the algorithm is robust to common digital audio processing methods and the quality of the audio is guaranteed.

【Key words】 Audio; Digital watermarking; Neural network; Self-adaptive; Human audio system(HAS)

1 概述

随着网络通信的普及,许多传统媒体内容都向数字化转变,这使得信息的存取和共享更加快捷和方便。越来越多的数字媒体得以广泛传播,信息表达的效率和准确度也得到了显著的提高,在电子商务中占据巨大市场份额。然而,在数字媒体给社会带来进步、改善人们日常生活和工作的同时,也暴露出越来越严重的安全问题。例如多媒体作品的版权侵犯、软件或文档的非法拷贝,电子商务中的非法篡改、复制和盗版现象非常普遍,严重阻碍了媒体数字化进程和数字媒体的正常合法使用^[1,2]。

数字水印是一种全新的信息安全领域技术,它利用信息隐藏技术实现对多媒体数据安全的保护。随着数字化音像制品和音乐制品的大量制作和发行,音频数据的版权保护也显得越来越重要,通过在音频载体中嵌入水印信息,可以实现拷贝限制、使用跟踪、盗版确认等功能。

基于 HAS 的音频水印技术是当前该领域研究的一个重要方面。该技术利用音频信号的主要特性,结合 HAS 的特征解决音频数字水印的鲁棒性和不可知觉性之间的矛盾。在保证不可知觉性前提下,在合适的位置嵌入尽可能高强度的水印信号,使水印自适应于宿主音频信号。

水印编码可看作在强背景下叠加一个弱信号。只要叠加的信号低于某一个阈值,听觉系统就无法感觉到信号的存在。

2 神经网络自适应扩展回声隐藏算法

算法主要包括水印生成、水印嵌入、水印提取和验证等 3 个方面。

2.1 扩展回声音频水印的生成

回声是由声源产生的声音,经反射后得到的。回声可以通过将宿主音频信号和冲激响应进行卷积计算得到^[3]。回声隐藏通过在时域中引入回声将数据嵌入到音频信号中,基于回声隐藏的数字水印算法利用了 HAS 时域掩蔽特性的向后掩蔽特性实现水印的嵌入。向后掩蔽是指弱信号在强信号消失之后变得无法听见。它可以在强信号消失之后 50~200ms 作用而不被人耳觉察。回声普遍存在于周围环境中,对于延迟时间很短的回声,例如 2ms,人类听觉系统不能够将它与直接到达的声音分辨出来,只是感觉声音的音色发生了变化,音色变得更丰富。回声隐藏正是利用了人类听觉系统的这种特性达到隐藏目的。因此,回声隐藏与其它方法不同,它不是将水印数据当作随机噪声嵌入到载体数据中,而是作为载体数据大环境条件,对一些有损压缩的算法具有一定的稳健性^[2,3]。

在时域扩展回声隐藏算法中将一个伪随机序列作为回声内核的一部分,如果用 $p(n)$ 表示 PN 序列, a 表示 PN 序列的幅值,则时域扩展回声内核可以表示成如下形式^[4,5]:

$$k(n) = \delta(n) + ap(n-m), \quad 0 < a \leq 1 \quad (1)$$

基金项目: 陕西省自然科学研究计划基金资助项目(2005F50); 陕西省教育厅产业化培育基金资助项目(02JC40)

作者简介: 王慧琴(1970—),女,副教授、博士后,主研方向:网络安全,数字水印,智能信息处理;修可山,工程师、博士生;姚钟涵,硕士生

收稿日期: 2006-01-17 **E-mail:** whq463@263.net

其中, m 用来模拟回声信号由于反射导致的与直接到达声音之间的延迟, 在回声编码中通过修改 m 嵌入水印信息。 g 用来模拟由于反射导致的声音能量的衰减。

2.2 基于神经网络的扩展回声水印的嵌入

利用计算宿主音频信号和时域扩展回声内核的线性卷积的形式, 实现水印的嵌入。嵌入公式可由下式表示:

$$y(n) = s(n) * k(n) \quad (2)$$

其中, $s(n)$ 是原始宿主音频信号, $k(n)$ 是时域扩展回声内核, $*$ 是线性卷积。

数字水印的两个最重要的特性是不可感知性和鲁棒性, 二者之间存在着竞争行为。如何折中二者之间的矛盾, 使水印具有良好的自适应性是数字水印技术的关键问题。在时域扩展回声隐藏算法中, 如果 PN 序列的幅值取得不合适, 对整个算法的影响很大。幅值取得太大, 鲁棒性好, 并且水印易于检测, 但水印的嵌入将严重影响到原信号的质量。因此, 如何控制好水印嵌入的强度, 是算法的关键所在。本文提出了一种基于神经网络的自适应扩展回声隐藏算法, 利用了 HAS 的频域掩蔽特性, 在满足水印信号的不可知觉性的前提下, 能够保证良好的鲁棒性。算法框图如图 1 所示。

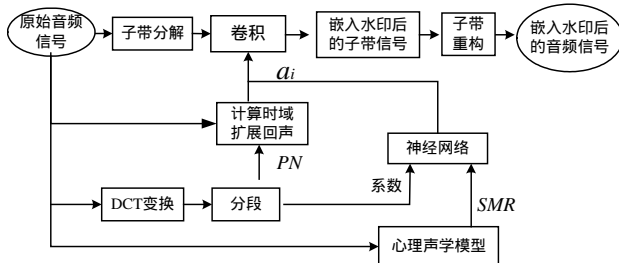


图 1 基于神经网络的扩展回声隐藏算法框图

算法采用神经网络方法估计水印的嵌入强度。如果将所有音频信号段的频谱组成一个集合, 相应的掩蔽阈值组成另一个集合, 集合间的映射关系是复杂的非线性映射关系。利用神经网络来逼近这种非线性映射关系, 从而可以根据输入信号来估计水印信号的缩放因子。

利用心理声学模型, 获得一个全局掩蔽阈值 T_g 。当量化噪声的声压级等于这个掩蔽阈值时, 听觉系统刚好能够感知到这个噪声, 所以这个掩蔽阈值又称为临界感知失真 (Just Noticeable Distortion, JND)。音频水印使用临界感知失真来限制嵌入的水印的最大能量, 当水印的能量被限制在临界感知失真阈值以下时, 就能够保证水印的不可感知性。

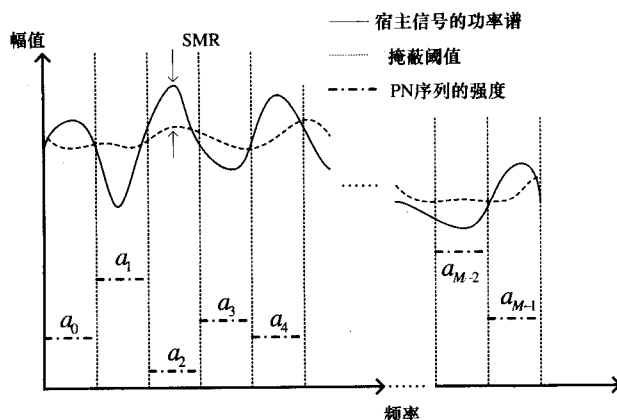


图 2 基于子带分解频率掩蔽阈值的确定示意图
算法采用三层 BP 神经网络, 分别为输入层、隐层、输

出层, 各层之间采用全连接, 各神经元的激活函数均为 Sigmoid 函数。设扩频后的回声内核大小为 N , 载体音频信号的大小为 S , 首先将载体音频信号分成 S/N 段, 若不能整除则舍余取整。对每一段分别作 DCT 变换, 根据 N 的大小, 将 DCT 变换后的系数划分成 M 个频段, 将每个频段变换后的频率系数和 SMR 作为神经网络的输入。如图 2 所示。其输出值为对应该频率段扩展回声核的幅值。

为更好地模拟 HAS 特性, 对应不同的 DCT 域变换系数采用不同的神经网络计算该系数所能嵌入的最大水印强度, 隐层的神经元个数也根据不同的 DCT 域变换系数而不同。

设 $s(n)$ 是原始宿主音频信号, $y(n)$ 是嵌入水印后的音频信号, 计算误差信号:

$$e(n) = y(n) - s(n) \quad (3)$$

使用此误差, 采用后向传播算法调整神经网络的权值。当均方误差小于预先确定的阈值时, 停止训练, 否则重复以上训练过程。训练好的神经网络的输出为扩展回声幅值 α_i 。

2.3 水印的检测

利用复频谱分析可以实现水印的提取^[4]。嵌入水印的信号复频谱等于原始宿主信号的复频谱和时域扩展回声内核的复频谱的和, $y(n)$ 的复频谱表示为

$$\hat{y}(n) = \hat{s}(n) + \hat{k}(n) \quad (4)$$

其中:

$$\hat{s}(n) = \text{IDFT}[\log(S(w))] \quad (5)$$

$$\hat{k}(n) = \text{IDFT}[\log(K(w))] \quad (6)$$

$$S(w) = \text{DFT}[s(n)] \quad (7)$$

$$K(w) = \text{F}[k(n)] \quad (8)$$

对等式(8)两边取傅立叶变换可以得到 $k(n)$ 的复频谱:

$$K(w) = 1 + ae^{-j\omega m} P(w) \quad (9)$$

对 $K(w)$ 取对数后记为 K_1 , 则:

$$k_1(w) = \log(K(w)) = \log(1 + ae^{-j\omega m} p(w)) \quad (10)$$

因为:

$$0 < a << \frac{1}{\max|P(w)|} \quad (11)$$

所以有:

$$|ae^{-j\omega m} P(w)| < 1 \quad (12)$$

因此式(10)可以展开成如下泰勒级数的形式:

$$k_1(w) = ae^{-j\omega m} p(w) - \frac{a^2}{2} (ae^{-j\omega m} p(w))^2 + \frac{a^3}{3} (ae^{-j\omega m} p(w))^3 + \dots \approx ae^{-j\omega m} p(w) \quad (13)$$

对式(13)进行反变换得到 $k(n)$ 的复频谱为

$$\hat{k}(n) = ap(n-m) \quad (14)$$

将式(14)代入到(4), 得

$$\hat{y}(n) = \hat{s}(n) + ap(n-m) \quad (15)$$

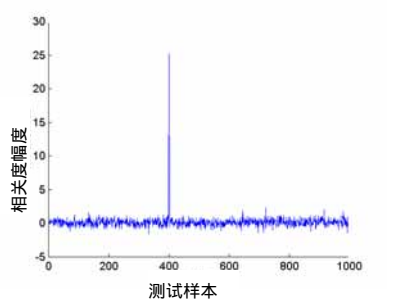
利用复频谱对水印信号进行检测的过程可用下式描述:

$$d_c(n) = \hat{y}(n) \otimes p(n) = n_s(n) + ap(n-m) \otimes p(n) \quad (16)$$

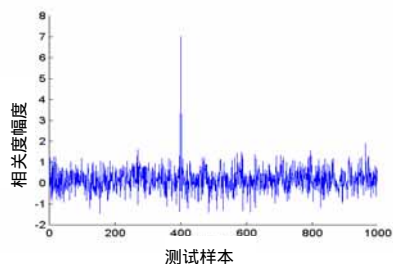
其中 \otimes 表示互相关运算, $n_s(n)$ 是考虑到在水印提取时存在的误差的附加项。

2.4 仿真实验

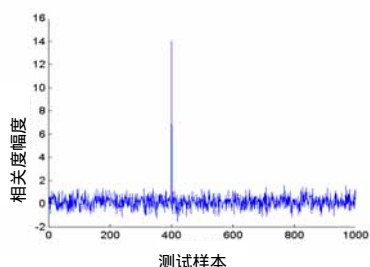
为了更客观地估计待检测信号, 选取 1 000 个随机序列作为校验数据, 分别与 $\hat{y}(n)$ 作互相关运算, 其中, 第 400 个为实际嵌入的水印序列。由相应曲线可以看出, 在各种情况下, 实际嵌入的水印信号的相关度响应有明显的峰值。实验结果如图 3 所示。



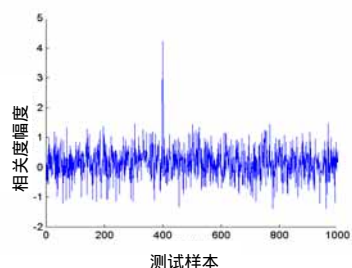
(a) 检测器音频信号的响应图



(b) 检测器对音频信号低通滤波后的响应



(c) 检测器对音频信号噪声干扰后的响应



(d) 检测器对音频信号 1/4 重采样后的响应

图 3 仿真实验结果

图 3(a)是不加任何干扰情况下的响应,响应值很明显。图 3(b)采用长度为 5 阶、截止频率为 250Hz 比雪夫 I 型低通滤波器滤波后检测出水印的情况,效果良好。图 3(c)为噪声干扰实验。在音频信号中加入信噪比为 18dB 的高斯白噪声进行检验,可检测出水印信息。图 3(d)为重采样实验。以低采样率对含水印的信号进行下采样,然后再以原采样频率进行上采样。由采样定理可知,这样会丢失音频的高频成分。试验表明,重采样频率为原来的 1/3,仍可以检测出水印信息。

3 结论

本文提出一种基于模糊集合理论的自适应时阈扩展回声隐藏算法。算法利用 PN 序列将音频信号的单回声内核进行扩展作为水印信号,提高了水印算法的安全性。算法利用了神经网络的非线性映射能力实现逼近音频信号的频谱和掩蔽阈值非线性映射,从而避免了复杂的心理声学模型的计算过程,并且可以实现算法的自适应。仿真实验证明了算法的有效性和可靠性。

参考文献

- 1 孙圣和. 数字水印技术及应用[M]. 北京: 科学出版社, 2004.
- 2 Hyen O O, Jong W S, Jin W H, et al. New Echo Embedding Technique for Robust and Imperceptible Audio Watermarking[C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, Utah, USA, 2001: 121-126.
- 3 Say W F, Theng H Y. An Adaptive Audio Watermarking System[C]. Proceedings of the 10th IEEE Region International Conference on Electrical and Electronic Technology, Singapore, 2001: 509-513.
- 4 Byeong-Seob K, Ryouichi N, Yoiti S. Time-Spread Echo Method for Digital Audio Watermarking Using PN Sequences[C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, Florida, USA, 2002: 2001-2004.
- 5 Byeong-Seob K, Ryouichi N, Yoiti S. Robust Watermarking Based on Time-spread Echo Method with Subband Decomposition[J]. IEICE Trans. Fundamentals, 2004, 87-A(7): 1647-1650.

(上接第 24 页)

失败连接的流量分布变化更加明显,因此本方法对早期的蠕虫攻击更加敏感;和基于 TCP SYN、TCP RST 和 ICMP 主机不可达消息的统计方法相比,本方法充分利用了蠕虫攻击所引起的失败连接的流量分布变化,因此具有更低的误报率。

参考文献

- 1 Zou C C, Gong W, Towsley D, et al. The Monitoring and Early Detection of Internet Worms[J]. IEEE/ACM Transactions on Networking, 2005,13(5): 961-974.
- 2 Chen Shigang, Ranka S. Detecting Internet Worms at Early Stage[J]. IEEE Journal on Selected Areas in Communications, 2005, 23(10): 2003-2012.
- 3 Berk V H, Gray R S, Bakos G. Using Sensor Networks and Data Fusion for Early Detection of Active Worms [C]. SPIE AeroSense Symp., Orlando, FL, 2003.

- 4 He Hui, Zhang Hongli, Zhang Weizhe, et al. Early Warning of Active Worms Based on Multi-similarity[C]. Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 2005-08: 3876-3883.
- 5 胡昌华, 李国华, 刘 涛等. 基于 Matlab 6.x 的系统分析与设计——小波分析[M]. 西安: 西安电子科技大学出版, 2004.
- 6 Mallat S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1989,11(7): 674-693.
- 7 Jump@ Network Information Audit System (JAudit-N100) [EB/OL]. http://www.jump.net.cn, 2005-12.
- 8 Chen Shigang, Tang Yong. Slowing Down Internet Worms[C]. Proc. of the 24th International Conference on Distributed Computing Systems, Tokyo, Japan, 2004-03.