

一种基于回声隐藏的数字音频水印改进算法^①

雷 鸣, 王新房, 刘 侨, 孙 鑫

(西安理工大学 自动化与信息工程学院, 西安 710048)

摘 要: 在前后向双核回声隐藏思想的基础上, 提出了一种基于回声隐藏的改进算法. 根据设定的水印的嵌入率来确定音频段的能量阈值, 选择短时能量高于此阈值的音频段嵌入水印信息, 并且引入功率倒谱检测法来检测水印信息. 仿真结果表明改进后的算法具有不可感知性更高, 隐藏效果好并且显著降低了水印信息检测的误码率.

关键词: 回声隐藏; 短时能量; 能量阈值; 功率倒谱

Digital Audio Watermarking Adaptive Algorithm Based on Echo Hiding

LEI Ming, WANG Xin-Fang, LIU Qiao, SUN Xin

(Faculty of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: Based on the method of dual backward and forward delay kernels, an improved algorithm of echo hiding in which watermark codes embedded the audio signal is proposed in this paper. The energy of each segment into which the original audio is first divided was computed. The energy threshold of audio segments is set by the given rate of embedding watermark codes. Then the audio segments whose energy value above the energy threshold are embedded with the watermark codes and the secret information is detected with a power cepstrum computing method. The simulation results show that the improved algorithm has a significant improvement in imperceptibility and can remarkably decrease the bit error rate of detected watermark codes.

Key words: echo hiding; short-time energy; energy threshold; power cepstrum

数字水印技术是通过在数字多媒体中嵌入水印信息(保密信息)以实现版权保护、保密通信等功能的技术. 按照数字多媒体不同, 它可以分为: 图像水印技术、音频水印技术、文本技术等.

回声隐藏作为音频水印技术的一种, 是利用人耳的听觉遮掩效应, 在原始声音中引入人耳不可感知的回声以达到信息隐藏的目的, 是一种有效的音频信息隐藏方法. 它具有隐藏算法简单, 算法不产生噪声, 隐藏效果好等优点, 但同时又存在着缺陷: 当回声幅度较小时, 提取水印的误码率就会较高而当回声幅度较大时, 又达不到很好的隐藏效果^[1]. 因此, 在保证音频具有良好的感知透明性的情况下如何降低水印信息检测的误码率成为回声隐藏算法的难点.

对此, 国内外专家、学者针对回声隐藏的特点展

开了大量的深入研究, 取得了许多建设性的成果, 总结起来, 主要集中在两方面: 首先是对回声内核的改造及衰减系数的自适应调节. 比如 C.Xu 等人提出的多回声核隐藏算法^[2]、H.O.Oh 等人提出的基于双极性的回声核隐藏算法^[3,4]、Kim 等人的前后向回声核方法^[5]等; 然后是对传统的提取水印方法进行改进. 比如赵朝阳等人提出的指数序列提取方法^[6]、殷凯等人提出的空余隐藏量改进方法^[7]、比如段柳云等人的倒谱优化算法^[8]和 BinYan 等人提出的功率倒谱检测法^[9]等.

本文在文献[5]提到的前后向回声核算法的基础上提出了一种改进的回声隐藏算法, 计算音频信号不同音频段的短时能量, 按能量值从大到小排序后, 根据水印信息的嵌入率确定能量阈值, 选择短时能量高于此阈值的音频段嵌入水印信息, 以求在保证水印隐藏

^① 收稿时间:2012-12-09;收到修改稿时间:2013-01-21

效果的同时提高检测水印信息的准确率, 并且引入了文献[9]提到的功率倒谱检测法, 进一步降低了水印信息检测的误码率.

1 前后向核算法思想

回声隐藏的思想是利用人耳的听觉遮掩效应, 也就是一个高音可以遮盖一个极短时间内相邻的低音, 使得叠加在原始语音上的回声信号几乎不被感知, 以达到良好的透明性和不可察觉性. 1996 年 bender 首次提出了回声隐藏算法^[10], 嵌入水印的数学模型为:

$$\begin{aligned}
 y(n) &= x(n) * h(n) \\
 h(n) &= \delta(n) + \alpha \cdot \delta(n-d)
 \end{aligned}
 \tag{1}$$

其中, $y(n)$ 为含回声信息的音频信号; $x(n)$ 为原始音频信号; $h(n)$ 为回声内核; $\delta(n)$ 为脉冲信号; α 为回声衰减系数; d 为延迟时间. 回声隐藏就是通过音频信号中引入不同的回声延迟来嵌入水印比特值的.

Kim 提出了前后向回声核^[5]的思想: 该回声内核由两个延迟时间相同, 但是方向相反的回声引入因子构成, 可表示为

$$h(n) = \delta(n) + \alpha\delta(n-d) + \alpha\delta(n+d)
 \tag{2}$$

其中 $\delta(n-d)$ 称为后向回声引入因子, $\delta(n+d)$ 称为前向回声引入因子.

对含有上述回声核的音频信号 $y[n]$ 进行倒谱分析:

$$\begin{aligned}
 C_{y'}[n] &= C_x[n] + C_h[n] \\
 C_h[n] &= \alpha\{\delta[n-d] + \delta[n+d]\} - \frac{\alpha^2}{2}\{\delta[n-2d] + 2\delta[n] \\
 &+ \delta[n+2d]\} + \frac{\alpha^3}{3}\{\delta[n-3d] + 3\delta[n-d] + 3\delta[n+d] \\
 &+ \delta[n+3d]\} - \dots \quad \alpha \in (0,1)
 \end{aligned}
 \tag{3}$$

上式在 $n=d$ 处的幅值近似为

$$\alpha + \alpha^3 + \alpha^5 + \alpha^7 + \dots \approx \frac{\alpha}{1-\alpha^2} \quad \alpha \in (0,1)
 \tag{4}$$

显然它是大于传统算法在 $n=d$ 处的幅度 α 的, 这样水印信息的检测率得到提高. 所以理论上前后向回声核算法在提高水印检测率上是优于传统算法的.

前后向回声核算法正是利用前后向内核上述特性来提高水印信息的不可感知性, 并且在检测过程中获得较低的误码率的. 但是音频信号是随时间而变化的, 各个音频段的短时能量是不同的. 能量较小的音频段, 水印信息的隐藏效果较差, 而降低回声幅度后又难以检测出隐藏的水印信息, 误码率较高. 所以若能降低

能量较小音频段的影响, 音频信号将会获得较高的信息隐藏能力和较低检测误码率.

2 算法的改进

本文提出的改进算法主要是在音频信号的分段方式、嵌入方法上对文献[5]提出的前后向回声核算法进行改进, 并且在水印信息的检测过程中引入文献[9]的功率倒谱检测方法. 其框图如图 1 所示.

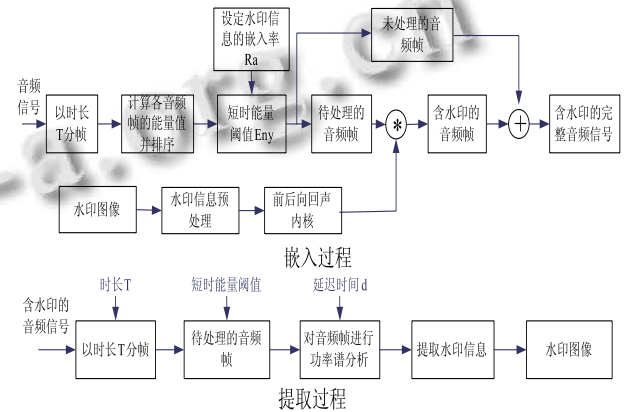


图 1 整体算法流程图

图 1 中, *表示卷积运算, +表示音频段合成. 水印嵌入过程是指水印信息的隐藏到音频信号中的过程, 水印提取过程是指从含回声的音频信号中检测出水印信息的过程.

采用时长 T 代替以固定采样数 N 为段长对音频信号分段: 1) 音频信号的长短通常是以时间为单位衡量的. 由于采样频率 f_s 的不同(可分为 11KHz、22.05KHz 和 44.1KHz 等), 同一时长的音频信号含有的采样点数差别较大, 若以固定采样数 N 为段长分段, 可能导致采用的算法不具有普遍性, 比如: 以采样点数 N 分段, 在 11KHz 下可以嵌入 3 个完整水印图像, 在 44.1KHz 下却无法嵌入一个完整水印图像, 则嵌入算法将失效或者无法达到最佳的效果, 而以时长 T 分段则可以避免采用频率 f_s 的影响; 2) 回声延时是有范围的, 在文献[3]中提到以 2ms 为界, 延迟时间大于 2ms 称为 echo, 小于 2ms 则成为 coloration. 前者由于引入的回声延时过大, 影响听觉质量; 后者听觉上可理解为对原始声音的润色, 所以回声延时一般要小于 2ms 的. 以时长 T 分段, 可以人为控制它们的比例, 避免由于采样频率 f_s 的不同, 不同音频的延迟点相对靠近音频段的边缘而受到音频段间频谱噪声对延时处倒谱值的干扰.

2.1 嵌入方法的改进

对原始音频信号以时长 T 进行均匀分段, 共分为 M 个音频段(不计入不足一个时长的部分), 则每个音频段含有 $N(N=\text{fix}(T*fs))$ 个采样点. 计算出第 m 个音频段的短时能量^[10].

$$E_m = \sum_{n=(m-1)*N+1}^{m*N} x^2(n) \quad (m=1,2,3,\dots,M) \quad (5)$$

然后根据能量值从大到小进行排序, 在这一能量序列 $\{E_p\}(p=1,2,3,\dots,M)$ 中寻找能量阈值 E_n , 使得 $E_m > E_n$ 的音频段个数 n 与 M 的比值最接近给定的水印嵌入率 R_a (音频信号实际嵌入的水印比特与理论上能够嵌入的最大比特数之比). 将短时能量大于 E_n 的音频段称为待处理音频段, 向其中嵌入音频信息, 其他音频段及不足音频段的部分则不做处理, 最后按照原来的位置进行段合并得到含有回声的音频信号. 由于隐藏信息的过程是在短时能量较高的音频段进行的, 水印信息的隐藏能力得到提升, 隐藏效果较好; 音频信号的振动幅度较大, 又降低了隐藏信息检测的误码率.

实验发现, 存在少部分音频段隐藏信息后, 其短时能量 $E_k \leq E_n$, 显然若以能量阈值检测隐藏信息, 这样这些音频段将无法检测出水印信息, 并且造成检测到的水印信息产生错位, 大大增加了水印信息的误码率.

为了解决这一问题, 本文采用比例系数法: 人为抬升短时能量 $E_k \leq E_n$ 的音频段的能量, 为其能量

$$E_k' = E_n \times K \quad (6)$$

其中 $K = E_n / \min(E_k) \times (1 + \alpha)$

式中 $K > 1$, α 一般取所用软件的数值精度, 当考虑信道噪声影响时, 也可相应的增大, 力求处理后所得音频段的短时能量值 $E_k' > E_n$ 的情况下, 所取的 K 值最小. 如图 2 所示, 由于只是为这一少部分的音频段乘了同一个常数, 所以音频信号的不可感知性几乎不受影响.

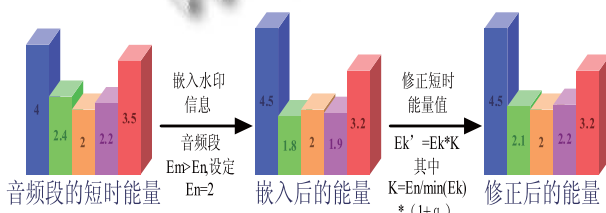


图 2 嵌入水印信息后能量处理示意图

2.2 功率倒谱检测法

功率倒谱检测法^[9]是 BinYan 等人提出的一种回

声隐藏信息的提取方法, 它利用嵌入隐藏信息的音频数据段的功率倒谱值来测定回声延时的位置, 提取出隐藏信息. 功率倒谱的计算方法^[10]如下:

$$C_{Power}[n] = (IFFT(\ln|FFT(y[n])|^2))^2 \quad (7)$$

由于在功率倒谱中, 对应于回声延时处的峰值要大大高于同等条件下倒谱法得到的峰值, 而相对于自相关复倒谱法, 其峰值更加突出, 远远高于其周围的值, 尤其是在语音信号的静音段, 效果更加明显. 所以, 采用功率倒谱的检测法代替自相关倒谱法能显著降低提取水印信息的误码率.

3 实验结果及分析

试验在 MATLAB 开发环境下进行仿真: 采用 3 类, 每类含 25 首时长为 3min, 格式为 wav 的不同音频信号来测试改进算法的性能. 将这 3 类音频信号分别标记为 A, B, C, 其中 A 为萨克斯器乐, B 为轻音乐, C 为摇滚乐. 它们同为单声道音频, 采样频率均选为 44.1KHz, 量化精度均为 16bit. 水印图像选取名为“版权所有”, 像素为 56×56 的二值图像, 当嵌入水印比特为“0”时 $d = d_0$, 嵌入水印比特为“1”时 $d = d_1$.

实验参数: 采用音频段时长 $T=0.01s$, $fs=44.1KHZ$, 长度 $N=\text{fix}(T*fs)=441$ (采样点数) 的矩形窗对上述音频分别进行分段. 回声衰减系数 $\alpha=0.2$, 回声延时 $d_0=75$ (约为 1.7ms), $d_1=50$ (约为 1.13ms). 水印的嵌入率为 0.7 (刚好嵌入 4 幅完整图像).

3.1 提取水印图像效果比较

本文采用归一化相关系数(NC)和误码率(BER)来评价提取的水印图像的效果.

1) 归一化相关系数(NC)是对提取的水印和原始水印的相似度进行定量的评价.

$$NC(w, w') = \frac{\sum_{i=1}^N w(i)w'(i)}{\sqrt{\sum_{i=1}^N w(i)^2} \sqrt{\sum_{i=1}^N w'(i)^2}} \quad (NC \in [0,1]) \quad (8)$$

式中, $w(i)$ 为原始二值水印图像的一维化序列, $w'(i)$ 为提取出的水印序列, NC 的值越接近 1, 说明原始水印图像和提取出的水印图像越相似, 相关性越好.

2) 误码率(BER)是指提取出的水印占原始水印比特值的百分率.

$$BER = \frac{\text{错误的比特数}}{\text{总比特数}} \times 100\% \quad (9)$$

BER 的值越接近 0, 说明原始水印图像和提取出的水印图像越相似, 相关性越好.

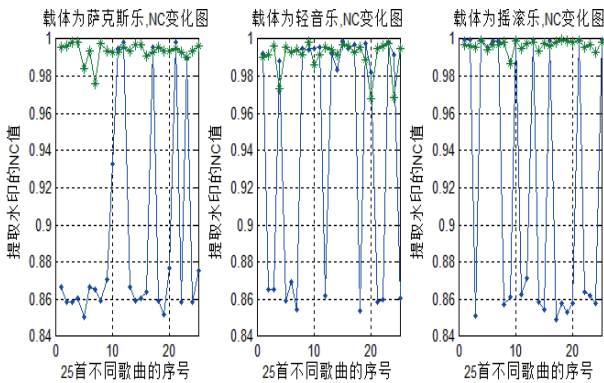


图 3 三类音频的 NC 变化图

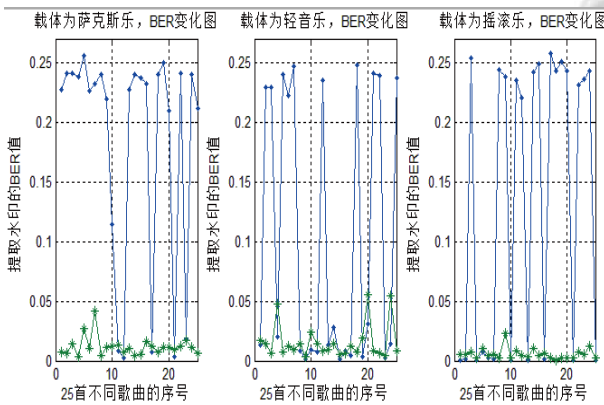


图 4 三类音频的 BER 变化图

图 3 和图 4 中图线振幅波动较大的点划线均指前后向回声核算法得到的变化曲线, 图线振幅波动较小的星划线为改进算法得到的变化曲线, 从图中可以看出, 无论是采用 NC 评价还是 BER 评价, 改进算法的得到的绝大多数结果都能显示出提高 NC 值或者降低 BER 值的效果, 并且对于三类不同类型的音频信息都能取得非常明显的效果, 显然本文提出的改进算法在检测水印信息的准确率上是优于前后向回声核算法的. 而从表 1 中也可以得出改进算法不但可以获得较高的准确率, 并且稳定性也是大大高于前后向回声核算法的.

3.2 音频感知透明性比较

通常采用音频信噪比(SNR)来衡量音频信号的感知透明性的, 它能客观评价隐藏水印信息(嵌入水印信息)后原始音频信号受影响的程度, 一般来说, SNR 值越大, 隐藏信息(水印信息)越难以被人耳感知, 音频信号的感知透明性越高.

表 1 NC 和 BER 的平均值和方差

音乐类型	算法名称	平均值		方差	
		NC	BER	NC	BER
萨克斯乐	前后向核算法	0.8919	0.1842	0.0030	0.0088
	改进算法	0.9935	0.0117	$2.2 \cdot 10^{-5}$	$6.5 \cdot 10^{-5}$
轻音乐	前后向核算法	0.9402	0.1014	0.0044	0.0128
	改进算法	0.9910	0.0156	$7.2 \cdot 10^{-5}$	$2.2 \cdot 10^{-4}$
摇滚乐	前后向核算法	0.9203	0.1361	0.0051	0.0149
	改进算法	0.9967	0.0057	$7.3 \cdot 10^{-6}$	$2.2 \cdot 10^{-5}$

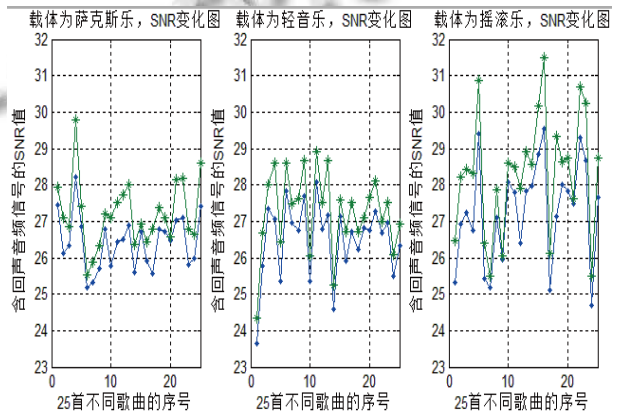


图 5 三类音频的 SNR 变化图

图 5 为分别采用上述两种算法得到的含有回声的音频信号的音频信噪比(SNR)值. 其中图线振幅波动较大的点划线表示前后向回声核算法得到的 SNR 变化曲线, 图线振幅波动较小的星划线表示改进算法得到的 SNR 变化曲线. 从图中可以看出采用改进算法得到的音频信号的 SNR 值绝大部分是高于前后向回声核算法的, 也就是说采用改进算法得到的音频的感知透明性更好, 音频效果更高.

4 结语

本文针对音频信号存在的振动幅度较小即音频段的短时能量值较小而造成的水印信息的隐藏效果较差, 检测误码率较高的情况, 在充分研究文献[2]前后向回声核算法和文献[9]功率倒谱检测法的基础上, 提出了基于短时能量阈值的过滤效应的回声隐藏算法. 在嵌入水印信息前, 先计算各音频段的短时能量, 根据设定的水印嵌入率确定的能量阈值过滤掉短时能量较低的音频段而向能量较高的音频段嵌入信息, 则得到的音频信号回声隐藏效果好, 不可感知性强, 并且隐藏

(下转第 94 页)

参考文献

- 1 Harman M, Mcminn P, Wegener J. The impact of input domain reduction on search based test data generation. Antonia Bertolino. Proc. of the the 6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on The Foundations of Software Engineering. New York: ACM Press, 2007: 155-156.
- 2 Zheng L, Harman M, Hierons RM. Search algorithm for regression test case prioritization. IEEE Trans. on Software Engineering, 2007,33(4):225-228.
- 3 Berndt D, Fisher J, Johnson, et al. Breeding software testcases with genetic algorithm. Proc. of the 36th Hawaii International Conference on System Sciences. IEEE Press, 2002:1-4.
- 4 Bryce RC, Colbourn CJ. Constructing interaction test suites with greedy algorithm. The Proc. of ASE'05. California, ACM Press, 2005. 124-136.
- 5 莱伟,谢军凯,奚红宇,等.遗传算法在软件测试数据生成中的应用.北京航空航天大学学报,1998,24(4):434-437.
- 6 Girgis MR. Automatic test data generation for data flow testing using a genetic algorithm. Journal of Universal Computer Science, 2005,11(6):898-915.
- 7 Rapps S, Weyuker EJ. Selecting software test data using data flow information. IEEE Trans. on Software Engineering, 1985,11(4):367-375.
- 8 Frankl PG, Weyuker EJ. An applicable family of data flow testing criteria. IEEE Trans. on Software Engineering, 1988: 1483-1498.
- 9 王小平,曹立明.遗传算法-理论、应用与软件实现.西安:西安交通大学出版社,2002:19-67.
- 10 AHmed SG, Harrold MJ. Using genetic algorithms to aid test-data generation for data-flow coverage. Proc. of the 14th Asia-Pacific Software Engineering Conference. Nagoya: IEEE Press, 2007: 41-48.

(上接第84页)

信息检测的误码率更低.

参考文献

- 1 姚钟涵,王慧琴,毛力.一种基于回声隐藏的数字音频水印算法.微计算机信息,2008,24:1-1.
- 2 Xu C, Wu J, Sun Q, Xin K.Applications of watermarking technology in audio signals. Journal Audio Engineering Society, October,1999,47(10):1995-2007.
- 3 Oh HO, Kim HW, Seok JW. Transparent and robust audio watermarking with a new echo embedding technique. Multimedia and Expo, 2001:317-320.
- 4 Huang DY, Yeo TY. Robust and inaudible multi-echo audio watermarking. IEEE Pacific Rim Conference on Multimedia, 2002. 615-622.
- 5 Kim HJ, Choi YH. A novel echo-hiding scheme with backward and forward kernels. Circuits and Systems for Video Technology, 2003,13:885-889.
- 6 赵朝阳,刘振华,王挺.数字音频信号的回声隐藏技术.计算机应用研究,2000,17:42-44.
- 7 殷凯,周辉.音频数字水印中回声隐藏技术的时域提取方法.计算机与数字工程,2006,34(5).
- 8 段柳云,宗节保,等.一种基于回声隐藏的倒谱提取优化算法.电子设计工程,2010,18(7).
- 9 Yan B, Sun SH, Lu ZM. Improved echo hiding using power cepstrum and simulated annealing based synchronization technique. The 2nd Int. Conf. on Machine learning and Cybernetics. Xi'an, China, 2003,2(5):2142-2147.
- 10 Bender W, Gruhl D, Morimoto N, et al. Techniques for data hiding. IBM Systems Journal, 1996,35(3/4):313-316.
- 11 唐升,侯榆青,克兢.一种基于短时能量自适应的回声隐藏算法.计算机应用,2008,28(11).