

基于频率追踪的非线性器件虚拟低音增强算法

· 算法介绍 ·

朱睿 杨飞然 叶超 杨军

(中国科学院 声学研究所 通信声学实验室 北京 100190)

【摘要】目前经典的虚拟低音增强技术的时域算法存在比较明显的非线性失真。对基于非线性器件的虚拟低音增强算法进行了研究和改进,引入自适应格型陷波滤波器进行频率追踪,并改进了传统的谐波产生方法。Rnonlin 客观模型与隐藏参考和基准的多刺激法(multistimuli with the hidden reference and anchor, MUSHRA)主观听音实验的结果表明,改进后的时域虚拟低音增强算法音质明显得到改善,低音感知效果得到一定的提高。

【关键词】虚拟低音; 频率追踪; 非线性器件; 自适应陷波

【中图分类号】 TN912

【文献标志码】 A

DOI: 10.16311/j.audioe.2015.02.16

Frequency Tracking Nonlinear Devices Technique for Virtual Bass System

ZHU Rui, YANG Feiran, YE Chao, YANG Jun

(Communication Acoustics Laboratory Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China)

【Abstract】The classical time-based nonlinear devices technique(NLD) for the virtual bass system has obvious nonlinear distortions. To reduce distortions, a frequency tracking nonlinear devices technique is presented. This technique adopts adaptive notch filter(ANF) for frequency tracking and an improved method for harmonic generation. The Rnonlin distortion evaluation model test and the subjective listening test called multistimuli with the hidden reference and anchor(MUSHRA) are conducted. The results show that new technique achieves good bass impression with reduced distortion.

【Key words】 virtual bass; frequency tracking; nonlinear devices; adaptive notch filter

1 引言

随着便捷式计算机、超薄手机以及液晶电视等消费类电子产品的普及,市场对于体积小、重量轻、功耗低且声品质优良的小型扬声器的需求日益强烈。小型扬声器由于受到单元尺寸等物理条件的限制,存在低频还原能力差的问题^[1]。传统上采用均衡器或分频器的方法会带来系统耗能增加,效率降低,不利于携带等诸多问题^[2]。利用心理声学原理的虚拟低音增强技术是一种增强小型扬声器低音重放品质的有效方法。

虚拟音调指人耳在基频缺失的情况下可以利用基频信号的谐波组合重建基频信号。目前,虚拟低音增强技术按照谐波的产生方法可以分为时域和频域两大类。时域法包括利用乘法器回路产生谐波的MaxxBass技术^[2]和基于非线性器件的谐波生成技术(Nonlinear Devices, NLD)^[3-5]。时域算法的主要优点是算法简单,处理速度快,利于实时实现,但是会产生非线性畸变和互调失真。频域法原理是利用相位声码器进行频谱搬移来产生高次谐波,其优点是失真

小,响度控制精确,但是计算量和系统延迟较大,存在瞬态模糊和混响效应等缺陷^[6]。

本文在NLD算法的基础上引入自适应陷波技术(Adaptive Notch Filter, ANF),提出了基于频率追踪的非线性器件虚拟低音增强算法,降低了NLD的非线性失真,取得了更好的低音感知效果。

2 基于频率追踪的非线性器件虚拟低音增强技术

基于频率追踪(Frequency Tracking, FT)的虚拟低音增强是由Larsen和Aarts在其著作^[7]中提出的,书中的FT模块采用一种基于3个样点估计信号频率的快速算法。该方法的缺陷是只能追踪单一频率分量,在此基础上Lee提出了基于级联的二阶自适应格型陷波器的FT^[8],用来处理多个频率分量的情况。这种技术存在的问题是它只处理了稳态谐波分量,对于在低音听感中占很重要作用的瞬态冲击部分未做处理,造成最终感知的低音效果欠缺力度与饱满感。本文利用自适应陷波器频率追踪技术分析信号,将信号分解为稳态基频成分和瞬态残余分量,基频成分利

用谐波生成公式产生谐波,残余分量采用 NLD 产生谐波。该算法框图如图 1 所示。

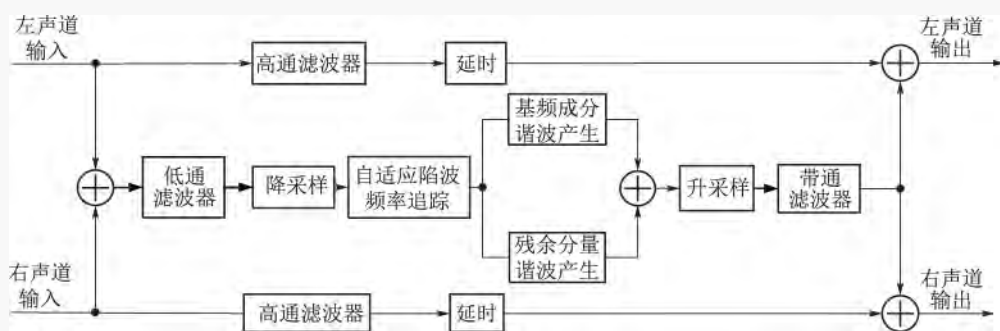


图 1 基于频率追踪的非线性器件虚拟低音增强整体框图

2.1 频率追踪

二阶格型陷波器可以写成

$$H(z) = \frac{N(z)}{D(z)} = \frac{1 + 2k_0 z^{-1} + z^{-2}}{1 + k_0(1 + \alpha)z^{-1} + \alpha z^{-2}} \quad (1)$$

式中: α 为零极点收敛因子,用来控制陷波器的带宽; k_0 是自适应系数。自适应滤波具体实现过程可参考文献 [9]。最终检测出来的基频频率表示为

$$f_0 = \frac{1}{2\pi} \arccsc(-k_0) \quad (2)$$

值得注意的是,式(2)中 k_0 取值范围被限定为 $[-1, 1]$ 因此需要对 k_0 做一定的约束。Lee 采用的是削波加平滑的方法^[8],这里可以借鉴 Romoli 在文献 [10] 中所采用的 Sigmoid 函数做约束来取得更精确的频率检测结果。本文采用两个陷波器级联,同时追踪信号中最强的两个频率分量。

2.2 谐波产生

在检测出信号基频后,下一步是分别生成基频成分和残余部分的谐波分量。谐波产生(Harmonic Generation HG)传统方法是采用下面的式子来产生谐波^[7-8]

$$x_h(t) = \sum_{k=M}^N A_k \sin(k\omega_0 t) \quad (3)$$

式中: ω_0 是估计出来的瞬时频率; k 是需要生成的谐波次数; A_k 是各次谐波的幅值。这样检测出主要频率分量的瞬时频率后,用式(3)可以避免 NLD 所带来的互调失真。然而对于非平稳信号,特别是频率随时间变化较快的信号时,该方法不能产生频率准确的谐波。产生这一问题的主要原因为:对于非平稳信号,当把通过 ANF 预测出来的时变的瞬时频率 $\omega_0(t)$ 直接带入式(3)时,得到信号

$$x_h(t) = \sum_{k=M}^N A_k \sin[k\omega_0(t)t] \quad (4)$$

其瞬时频率

$$\omega_h = \frac{d\Phi(t)}{dt} = k \cdot \left[\frac{d\omega_0(t)}{dt} \cdot t + \omega_0(t) \right] \neq k \cdot \omega_0(t) \quad (5)$$

从式(5)可以看出,式(4)不能严格产生 k 次谐波。因此需要对谐波生成公式做出改进,可以从高次谐波所需要的瞬时频率反向推出 HG 式(3)中所需要的参数频率。假设原始信号和所要产生的 k 次谐波信号的参数频率分别为 $f(t)$ 和 $f_k(t)$, 它们的瞬时频率需要满足式(6)中的关系

$$f'_k(t) \cdot t + f_k(t) = k \cdot [f'(t) \cdot t + f(t)] \quad (6)$$

问题转变成微分方程求解,即通过已知的频率函数 $f(t)$ 求解 k 次谐波频率函数 $f_k(t)$ 。将 $t = n \cdot t_s$ 代入式(6),并把微分改写为差分形式就可以得到方程的离散形式

$$[f_k(n) - f_k(n-1)] \cdot n + f_k(n) = k \cdot [f(n) - f(n-1)] \cdot n + k \cdot f(n) \quad (7)$$

通过式(7)就可以取得准确的谐波生成参数频率 $f_k(n)$ 。

对于提取出基频后的瞬态信号,采用 NLD 算法处理。笔者选取的是反正切开方非线性函数(AT-SR)^[5] 这个函数被验证在低音感知和非线性失真方面表现最好的一个函数

$$f(x) = 2.5 \tan^{-1}(0.9x) + 2.5 \sqrt{1 - (0.9x)^2} - 2.5 \quad (8)$$

3 仿真实验

对基于频率追踪的非线性器件虚拟低音增强算法进行了仿真实验研究。笔者选取的样本音频是钢琴的 C2 单音,采样率为 44.1 kHz。分别采用式(3)和式(7)生成 2~6 次谐波,比较结果如图 2 所示。其中图 2(a)是基于 1024 点 STFT 的钢琴 C2 单音低频分量时频分布图;图 2(b)是采用自适应格

型陷波器基频追踪的结果,同时追踪两个峰值;图2(c)是采用式(3)对于检测出来的峰值1产生谐波的时频分布图;图2(d)是采用式(7)对于检测出的峰值1产生的各次谐波的时频分布图,由于低频部分采用了降采样处理(c)和(d)选取512个点的STFT。从

仿真结果可以看出,基于陷波器的频率追踪算法可以较好地找到信号的基频。对比两种谐波产生方案发现,本文提出的谐波生成方案在产生各次谐波准确性以及产生信号的清晰度上要优于传统方法,特别是在频率变化的部分这种优势更加明显。

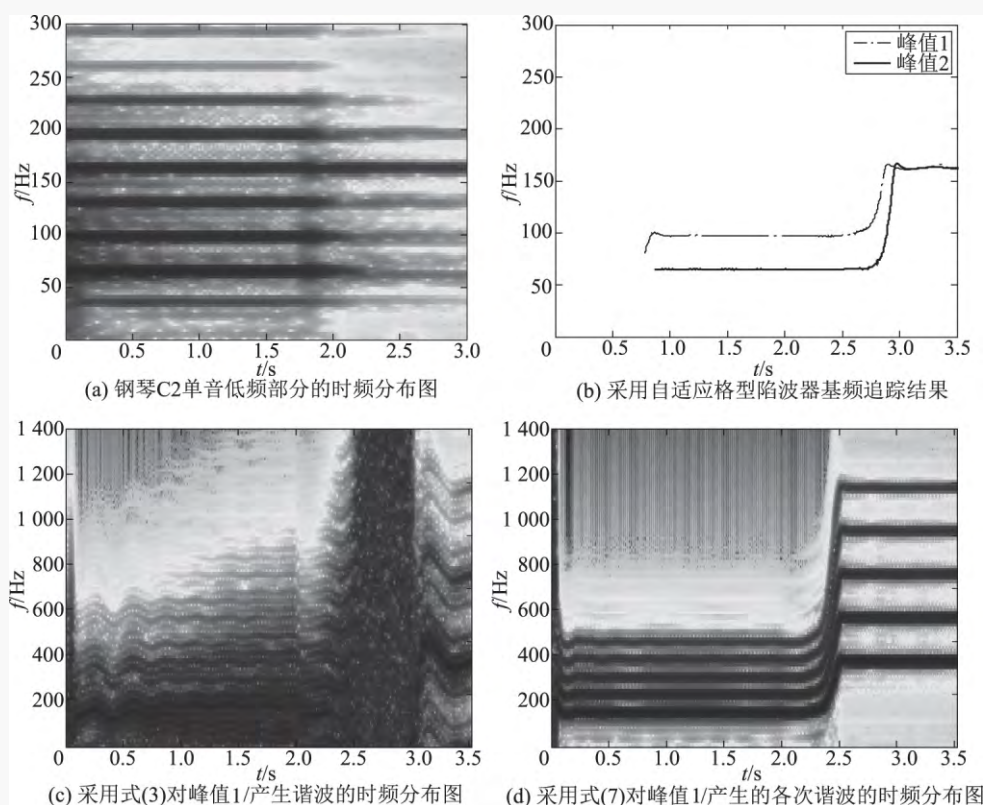


图2 谐波产生仿真比较结果

其次,对于本文提出算法在非线性谐波失真性能上进行了分析。这里采用 Tan 和 Moore 等人提出的 Rnonlin 模型来评价非线性失真的大小^[11]。Rnonlin 模型是充分考虑人耳听觉感知机理的一个非线性失真评价模型,被证明可以应用于 NLD 的非线性失真评估^[5]。测试选取的音乐包括:(1)钢琴 C2 单音片段;(2)低音大提琴片段;(3)舞曲鼓点片段;(4)流行音乐“渡口”片段。采用的处理方式有:(1)直接滤除 200 Hz 以下低频;(2)采用 NLD 法处理;(3)采用 ANF 法处理;(4)采用本文所提出的方法处理(NEW)。图3为几种处理方法的 Rnonlin 模型评估结果。

从图中可以看出,NLD 算法在 Rnonlin 模型中得分最低,这是由于其引入大量非线性谐波失真。ANF 与直接高通滤波的结果最接近,这是因为 ANF 对原始音频所带来的改变最小,只是增加了两个单频的谐波分量。而本文提出的算法降低了 NLD 中的非线性谐波失真。

4 主观评价实验

在这一部分笔者采用心理声学主观听音实验来检验本文提出算法的效果。这里采用 ITU-R BS.1534-1 标准提出的隐藏参考和基准的多刺激法(Multistimuli with the Hidden Reference and Anchor, MUSHRA)^[12]作为测试手段。主要评价音频两个方面的品质:低音感知效果和音质损伤程度。受试者选取声学专业的6名学生,其中男生3名,女生3名。

测试音乐的选取以及处理的方式与上面 Rnonlin 模型测试一样。只是在听音时加入未处理原始音频作为参考音源。测试采用的耳机为森海塞尔 HD380PRO。图4是低频感知效果的测试结果。

从图4可以看出,本文提出的方法在低音感知效果上优于 NLD 和 ANF 算法,但是离原始低音效果还有一定差距。图5是音质损伤程度的测试结果。

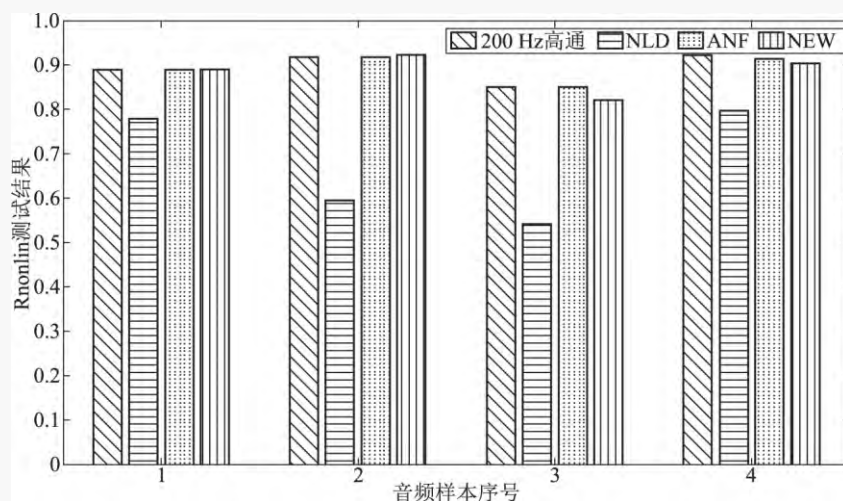


图3 几种算法的 R_{nonlin} 模型结果

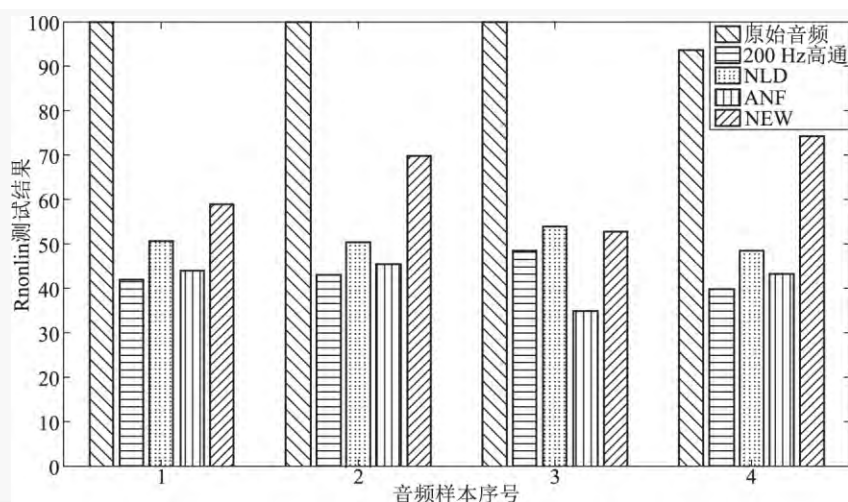


图4 低频感知效果的测试结果

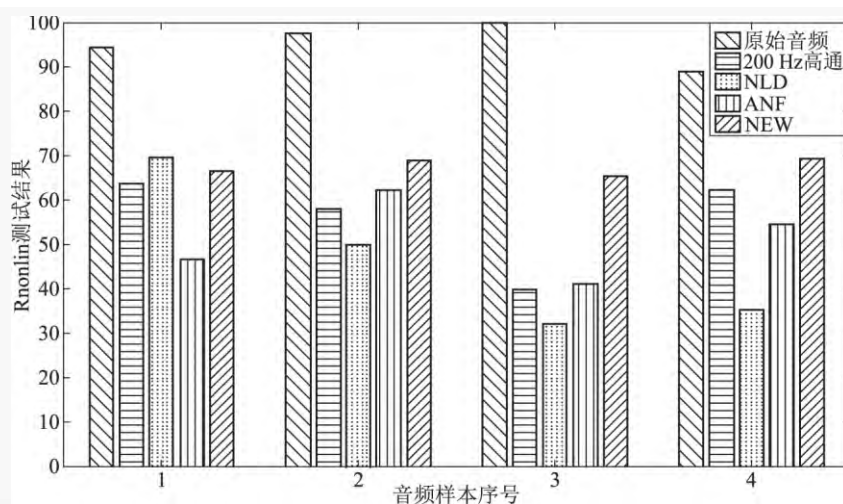


图5 音质损伤程度的测试结果

从图5可以看出本文提出算法对音质的损伤是几种算法里最小的,结果与Rnonlin模型仿真结果符合。

5 结论

本文针对NLD算法会带来非线性失真这一缺陷进行了研究,提出将ANF引入NLD算法,并改进了传统的谐波生成方法,经过综合对比Rnonlin客观仿真模型和Mushra主观听音实验结果可以看出,改进的NLD算法明显降低了原算法对音质的损害,且改进后的算法在低音感知效果上也有提高,但是与原始音乐的低频在听觉感知效果上还有一定差距。

参考文献

- [1] SMAL R H. Closed-box loudspeaker systems - part 1: analysis[J]. J. Audio Eng. Soc., 1972(20): 798-808.
- [2] BEN-TZUR D, COLLOMS M. The effect of the MaxxBass psychoacoustic bass enhancement system on loudspeaker design[J]. J. Audio Eng. Soc., 1999(47): 517.
- [3] OON GAN W S. Harmonic and intermodulation analysis of nonlinear devices used in virtual bass systems[C]//AES 124th Convention. Amsterdam. The Netherlands [s. n.], 2008.
- [4] LIM W T, OON GAN W S. Synthesis of polynomial-based nonlinear device and harmonic shifting technique for virtual bass system[C]//The IEEE International Symposium on Circuits and Systems (ISCAS). Taipei, Taiwan [s. n.], 2009.

(上接第71页)

便携式声测节点以及不间断实时声采集系统两个成功应用案例,可以看出,本文介绍的智能被动声探测节点在众多领域具有广泛的应用前景。

参考文献

- [1] 郭宝峰,林岳松,彭冬亮.声传感网中的语义增强型信息融合方法[J].杭州电子科技大学学报,2011,31(4):124-127.
- [2] 钟永信,黄建国,韩晶.基于空间唤醒的水声传感器网络节能路由协议[J].电子与信息学报,2011,33(6):1326-1331.
- [3] 林岳松,陈琳,郭宝峰.基于数据驱动的信息融合及其在车辆声辨识中的应用[J].电子与信息学报,2011,33(9):2158-2163.
- [4] 夏明,董亚波,鲁东明.无线传感网逐跳自适应FEC传输可靠性保证方法[J].浙江大学学报:工学版,2011,45

- [5] OON GAN W S, HAWKS FORD M O J. Perceptually-motivated objective grading of nonlinear processing in virtual-bass systems[J]. J. Audio Eng. Soc., 2011, 59: 804-824.
- [6] 郭杰,莫福源,蔡志博,等. EVB Phase Vocoder 算法与 MassBass 算法在虚拟低音增强应用中的比较[J].电声技术,2009,33(3):59-67.
- [7] LARSEN E R, AARTS R M. Audio bandwidth extension: application of psychoacoustics, signal processing and loudspeaker design[M]. [S. l.]: John Wiley & Sons, Ltd., 2004.
- [8] JUNHO L, EUNJUNG S, YOUNGCHEOL P, et al. Effective bass enhancement using second-order adaptive notch filter[J]. IEEE Transactions on Consumer Electronics, 2008(54): 663-668.
- [9] NAM-IK C, CHONG-HO C, UK L S. Adaptive line enhancement by using an IIR lattice notch filter[J]. IEEE Transactions on Acoustics, Speech and Signal Processing, 1989(37): 585-589.
- [10] ROMOLI L, CECCHI S, PERETTI P, et al. A mixed decorrelation approach for stereo acoustic echo cancellation based on the estimation of the fundamental frequency[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2012(20): 690-698.
- [11] TAN C T, MOORE B C J, ZACHAROV N, et al. Predicting the perceived quality of nonlinearly distorted music and speech signals[J]. J. Audio Eng. Soc., 2004(52): 699-711.
- [12] ITU-R BS. 1534-1 Method for the subjective assessment of intermediate sound quality (MUSHRA) [S]. Geneva, Switzerland: ITU, 2001.

责任编辑:闫雯雯

【收稿日期】2014-11-15

(2): 273-279.

- [5] 徐明. 水声传感器网中基于网络编码的多径路由协议[J]. 传感器与微系统, 2013, 32(1): 36-39.
- [6] COSTA A D, SAYEED A M. Data versus decision fusion in wireless sensor networks[C]//Proc. IEEE ICASSP, 2003. [S. l.]: IEEE Press, 2003: 832-835.
- [7] AGNOLI A, CHIUSO A, ERRICO P D, et al. Sensor fusion and estimation strategies for data traffic reduction in rooted wireless sensor networks[C]//Proc. IEEE ISCCSP, 2008. [S. l.]: IEEE Press, 2008: 677-682.
- [8] CHEN J C, YAO K, HUDSON R E. Source localization and beamforming[J]. IEEE Signal Processing Magazine, 2002, 19(2): 30-39.
- [9] 张福学. 传感器技术[J]. 电声技术, 1994(11): 47-48.

【责任编辑】丁雪

【收稿日期】2014-12-13