

## 虚拟低音算法的设计与实现\*

· 算法设计 ·

王红梅<sup>1</sup>, 刘华平<sup>2</sup>

(1. 浙江天格信息技术有限公司, 浙江 金华 321000; 2. 上海大学 通信与信息工程学院, 上海 200072)

【摘要】先介绍虚拟低音算法中最有代表性的两种算法, MaxxBass 算法和 EVB Phase Vocoder 算法。并给出了基于 VB Phase Vocoder 方法设计的虚拟低音增强系统, 并对虚拟低音增强结果进行对比分析。实验结果表明, 该虚拟低音增强系统能产生较为浑厚饱满的低音增强效果。

【关键词】虚拟低音; 谐波; 增强

【中图分类号】TN912.35

【文献标志码】A

## Design and Implement of the Virtual Bass Enhancement Algorithm

WANG Hongmei<sup>1</sup>, LIU Huaping<sup>2</sup>

(1. Zhejiang Tiange Information Co., Ltd., Jinhua Zhejiang 321000, China;

2. School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

【Abstract】The applications of MaxxBass algorithm and EVB Phase Vocoder algorithm in virtual bass enhancement are introduced. And an virtual bass enhancement system based on VB Phase Vocoder algorithm is implemented, and then, the result of the experiment is analyzed. The result shows that this virtual bass enhancement system can provide heavy bass effect, create the perfect environment different from the sound field, gain the effect is very obvious.

【Key words】virtual bass; harmonic; enhancement

## 1 引言

随着移动互联网时代的到来, 手持智能终端的普及率越来越高, 各大手机厂商对系统易用性及多媒体表现方面也原来越重视。在音频方面, 手机厂商将好音质等作为推广重点也可以看出, 消费者对于多媒体影音方面的要求越来越高。但同时, 智能终端如手机、平板计算机都趋向于纤薄和细小, 受到体积尺寸的影响, 在这些设备中无法防止尺寸过大的扬声器。所以, 从传统意义上做硬件改进比如增大扬声器尺寸改进材料以及腔体结构已经没有太多提升的空间。较小的扬声器口径通常其低频谐振频率较高, 低频放音能力差, 也就是低音表现不足。消费者对高品质音乐的追求和小扬声器低音表现力差的矛盾成为了急需改进的问题。

对于该问题, 目前多采用均衡器的方法来解决。但简单地通过均衡器提高低频能量容易产生信号畸变, 并降低扬声器的使用寿命。而利用虚拟低音增强技术实现的低音增强, 则避免了上述问题。

其中, 虚拟低音增强方案最有代表性的两个方案分别为 MaxxBass<sup>[1]</sup> 算法和 VB Phase Vocoder 算法。后续章节将分别介绍这两种方法, 并给出基于 VB Phase Vocoder<sup>[2]</sup> 方法设计的虚拟低音增强系统, 并对虚拟低音增强结果进行对比分析。

## 2 虚拟低音原理

虚拟低音来源于一种被称为“虚拟音调<sup>[3]</sup>”的现象。对于一段包含谐波的音频, 其基频部分决定了该音频信号的音调。而音色是人们对个谐波成分比例的主观感受。例如, 当存在  $f$ ,  $2f$ ,  $3f$  和  $4f$  四个频率的音频信号, 其音调会被认为是  $f$ 。将  $f$  去掉后的音频信号, 人们会感觉到音色有些变化, 但其音调仍会被认为是  $f$ 。这种现象被称为“基音缺失”。虚拟低音增强系统正是利用了这个现象来增强低音的表现。

虚拟低音增强系统的基本步骤大致相似, 下面以单声道音频为例, 其步骤如图 1 所示。

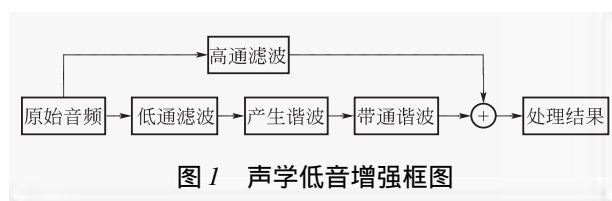


图 1 声学低音增强框图

\* [基金项目] 金华市科技计划重点项目(2013-1-007)

首先,原始音频输入到系统,由于只有 40 ~ 120 Hz 的低音部分产生谐波,所以通过低通滤波保存该范围基频部分,同时对该原始音频通过高通滤波保留其高频部分信息。低频信息通过谐波生成系统产生谐波,用于虚拟低音的作用。同时,对于谐波生成系统生成的过高的次谐波等噪声通过带通滤波进行滤除。最后,将保留的高频信息和处理过的谐波组合成最后输出结果。低音增强中最重要的系统就是谐波产生系统。

基于非线性电路实现的 MaxxBass 算法是最早投入市场的虚拟低音增强技术。对于频率为  $f_0$  的信号,可表示为

$$X = Ae^{-j2\pi f_0 t} \quad (1)$$

则通过乘法器输出的信号为两个信号的乘积,即

$$X_{out} = A^2 e^{-j2\pi 2f_0 t} \quad (2)$$

为频率为  $2f_0$  的信号,同理,可以通过乘法器产生  $3f_0$  等信号。

虚拟低音李路值保证率谐波叠加所得的虚拟低音与原基频信号音调保持一致,但要得到理想的虚拟低音效果,还需要保持新信号的音色与原信号一致。为此,MaxxBass 算法定义了响度变化量与声压级变化量的比值  $R$ ,该值随频率的变化改变。 $R$  称为声压级-响度扩展比率 (SPL-to-Phon Expansion Ratio)。当响度为 20 ~ 80 方,频率为 20 ~ 700 Hz,  $R$  可近似表示为

$$R(f) = \frac{1.0}{\ln(f) \cdot 0.241 - 0.579} \quad (3)$$

保持谐波信号响度对声压级的动态特性与基频处相一致,对于保持虚拟低音与原信号的音色一直具有重要作用。

设

$$RR(f, n) = \frac{R(f)}{R(n \cdot f)} = 1 + \ln(f) \cdot 0.241 \cdot R(f) \quad (4)$$

式中:  $n$  为谐波阶次。谐波能量与基频信号能量应满足关系

$$E_{nf} = RR(f, n) \cdot E_f + k \quad (5)$$

式中:  $E_{nf}$ ,  $E_f$  分别为倍频点和极品点的信号能量,以分贝为单位;  $k$  为增益常量。其对应的线性坐标公式为

$$X_{nf}^2 = 10 \lg \frac{k}{10} \cdot (X_f^2)^{RR(f, n)} = g \cdot (X_f^2)^{RR(f, n)} \quad (6)$$

式中:  $X_{nf}$ ,  $X_f$  分别为倍频点和基频点的信号幅度。

MaxxBass 算法简单、高效,但容易产生非线性畸变,影响输出的音乐音质。与 MaxxBass 方法不同,Phase Vocoder 通过视频分析方法实现信号的音调变换。通过 Phase Vocoder 方法,可以精确控制谐波的数量及能量,一定程度上弥补了 MaxxBass 算法的不足。

Phase Vocoder 算法采用时频分析方法中的窗口傅里叶变换方法,可实现音调变换<sup>[4]</sup>。该方法利用短时窗函数  $h(n)$  与原始信号  $x(n)$  相乘,截取小段信号,然后进行傅里叶变换得到其频谱信息。得到其信号的幅度和频率信息后,利用这些参数生成一组正弦信号的以完成新信号的合成。信号的短时加窗傅里叶变换可以写为

$$X(n, k) = \sum_{m=-\infty}^{+\infty} x(m) h(n-m) e^{-j2\pi mk/N} = |X(n, k)| e^{j\phi(n, k)} \quad (7)$$

在此基础上,信号可以写成一组幅度和相位随时间变化的正弦信号的叠加:

$$\begin{aligned} X(n) &= \frac{X(n, 0) + X(n, N/2)}{N} + \\ &\sum_{k=1}^{N/2-1} \frac{2}{N} |X(n, k)| \cdot \cos \left[ \phi(n, k) + \frac{2\pi kn}{N} \right] = \\ &\frac{X(n, 0) + X(n, N/2)}{N} + \sum_{k=1}^{N/2-1} \frac{2}{N} |X(n, k)| \cdot \\ &\cos [\phi(n, k) + \Omega_k n] = \\ &A_0(n) + \sum_{k=1}^{N/2-1} A(n, k) \cdot \cos [\theta(n, k)] \quad (8) \end{aligned}$$

改变各频率通道的幅度和相位参数即可实现对信号的相应变换。对英语音调变换,设帧移位为  $R_a$ ,由相邻 2 帧的相位值渴求的频率通道  $k$  的顺势频率  $d\theta(k)$  为

$$d\theta(k) = \Delta\theta(k) / R_a \quad (9)$$

其中  $\Delta\theta(k)$  为解卷积后 ide 相位差,再将瞬时频率诚意音调变换系数  $\alpha$  累加,便可得到音调变换后的信号相位

$$\psi(n+1, k) = \psi(n, k) + \alpha \cdot d\theta(k) \quad (10)$$

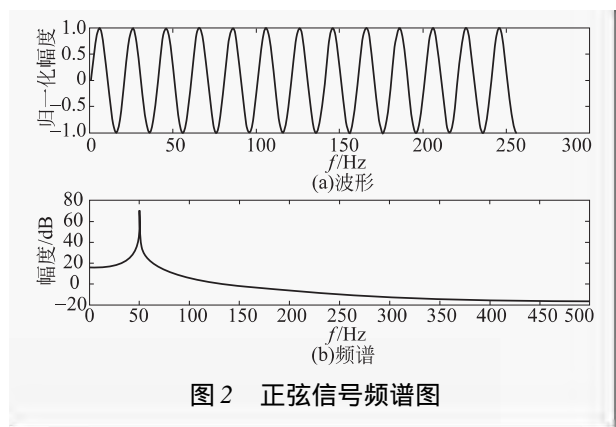
将各正弦分量叠加即可得到音调变换后的信号

$$y(n) = A_0(n) + \sum_{k=1}^{N/2-1} A(n, k) \cdot \cos [\psi(n, k)] \quad (11)$$

### 3 虚拟低音系统的设计

虚拟低音系统最重要的部分就是谐波生成部分,

例如对于正弦这种纯音信号,其波形和对应的频率如图2所示。

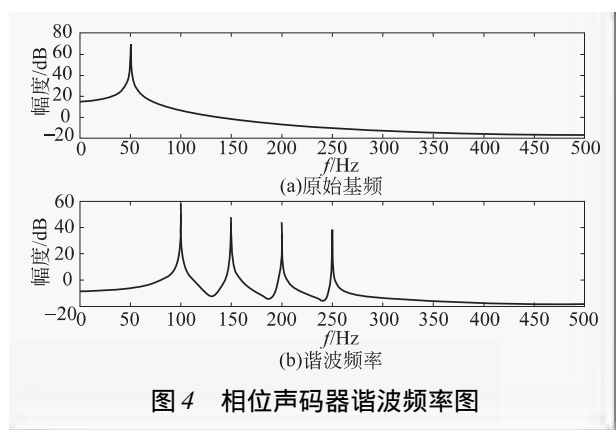
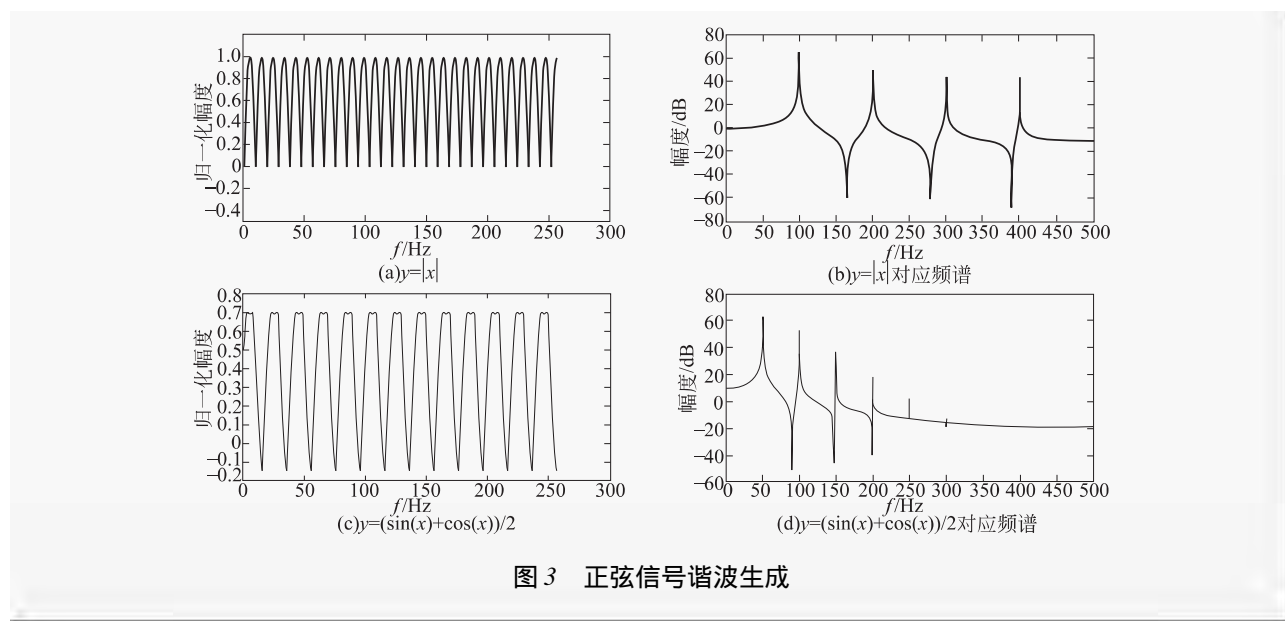


简单通过一些特殊函数即可构造其谐波,如图3所示。

从图3可以看出,虽然由特殊函数产生谐波算法简单,运行效率高。但不同函数其产生的谐波质量不同,波形的振幅及相位都是不可控的,也无法控制产生的谐波对音乐质量的影响。

利用相位声码器产生的波形机器频率如图4所示。

从图4可以看出,通过相位声码器,可以对产生的谐波次数、幅度以及相位进行精确控制。因此,本系统选用 VB Phase Vocoder 算法来生成谐波。



如前所述,不同阶次谐波的各种组合,虽然可以将音调与原始基频信号保持一致,但其音色会发生变化。文献[6]中提到,对于音色的鉴别有影响的只涉及到最初的7阶谐波,更高次的谐波对音色的贡献并不显著。令根据基音缺失理论,至少需要三个谐波组

合成的复音才能虚拟出对应的音频音调,且2、3、4次谐波对虚拟低音的贡献最重要,缺一不可。所以,本文选择了基频信号对应的2~5次谐波。这在保证可以虚拟出基频音调的同时,减少计算量,提高谐波生成模块的运行效率。

谐波的相对强度与虚拟低音音效有较强的关系。文献[6]中提到,在心里听力实验中,6种虚拟低音信号的谐波成分进行2-3-4-5组合,强度如表1所示。其中TP1、TP2、TP3分别按照0.1、0.2、0.25的系数线性递减,TP4按照 $0.5^{n-2}$ 衰减,TP5按照黄金比例衰减,TP6按照 $(n-1)^{-1}$ 衰减, $n$ 为谐波次数。最终对各个谐波组合的偏爱程度如表1所示。

从表中可以看出,TP4的偏爱程度最高,TP1的偏爱程度最小。高次谐波的相对强度越小,低次谐波的相对强度越大,虚拟信号的音色越接近基频信号。

表 1 谐波与虚拟低音音效强弱关系

	TP1	TP2	TP3	TP4	TP5	TP6
2Th	1.00	1.00	1.00	1.00	1.00	1.00
3Th	0.90	0.80	0.75	0.50	0.62	0.50
4Th	0.80	0.60	0.50	0.25	0.50	0.33
5Th	0.70	0.40	0.25	0.13	0.31	0.25

表 2 谐波组全偏爱程度

	TP1	TP2	TP3	TP4	TP5	TP6
偏爱程度	-0.92	-0.35	0.25	0.83	-0.10	0.63

所以 本算法按照 TP4 的比例来生成 2~5 次谐波。另外 处理后的谐波组合需要乘以一定的衰减系数再与高频分量叠加。

## 4 实验结果与结论

本文采用对比本算法与开方法产生的谐波生成的低音音质 来检验算法的有效性。本算法分别选取了《Hotel California》、《渡口》和《The Mass》。

其中 分别对三首歌曲利用谐波生成模块生成其 2~4 次谐波 并赋予不同的衰减系数。如表 3~5 所示。

表 3 2 次谐波增强

测试文件	衰减系数	低音增强
Hotel California	0.3	明显
渡口	0.3	轻微破音
The Mass	0.3	较严重破音

表 4 3 次谐波增强

测试文件	衰减系数	低音增强
Hotel California	0.15	不明显
渡口	0.15	明显
The Mass	0.15	轻微破音

表 5 4 次谐波增强

测试文件	衰减系数	低音增强
Hotel California	0.08	不明显
渡口	0.08	不明显
The Mass	0.08	明显

由上述三个表对比可以看出 不同音乐所需要的谐波整体分量是不同的。《Hotel California》本身低频不多 所以可以为其加入更多分量的谐波来提升其低音音质 但同样的谐波分量加入到《渡口》或是《The Mass》中 会引起其破音 这是因为《渡口》本身既是富含低频的音乐 在加入分量过多的谐波 会导致整体音乐表现的失常 引发破音。而《The Mass》更是在重低音歌曲的代表 所以其对谐波所需的分量更低。

综上所述 本文所述的虚拟低音增强方法能为音乐产生较为浑厚且听感明显的谐波成份 且由于快速傅里叶方法的应用 易于多平台的移植 并且效率较高。另外 如上述提到 不同音乐所适合的谐波分量不尽相同 甚至不同乐器 演唱者的歌唱习惯 男声女声等因素都对低音质感有较为明显的影响 后续的研究还需要将这些音色考虑在内 使得算法能有更好的鲁棒性和自适应性。

## 参考文献

- [1] DANIEL B T, MARTIN C. The effect of maxxbass psychoacoustic bass enhancement system on loudspeaker design[C]// AES Convention 106. [S. l.]: AES Press, 1999: 4892.
- [2] MINGSIAN R B, LIN Wan-chi. Synthesis and implementation of virtual bass system with a phase-vocoder approach [J]. Audio Eng. Soc. 2006, 54(11): 1077-1091.
- [3] ZWICKER E, FASTL H. Psychoacoustics facts and models [M]. 2nd ed. Berlin: Springer verlag, 1999.
- [4] ZHU Yongwei, KANKANHALLI M S. Robust and efficient pitch tracking for query-by-humming [C]// Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information Communications and Signal Processing 2003 and the Fourth Pacific Rim Conference on Multimedia. [S. l.]: IEEE Press, 2003: 1586-1590.
- [5] 孟子厚. 音质主观评价的实验心理学方法 [M]. 北京: 国防工业出版社, 2008.
- [6] 周加成. 虚拟低音音色与成分的关系 [J]. 声学技术, 2010, 29(6): 336-337.

[责任编辑] 丁雪

[收稿日期] 2014-08-25

(上接第 48 页)

- [3] 潘杨, 孟子厚. 视听注意属性对听觉掩蔽效应的影响 [J]. 电声技术, 2011, 35(2): 42-44.
- [4] BEERENDS J G, DE CALUWE F E. The influence of video quality on perceived audio quality and vice versa [J]. J. Audio Eng. Soc., 1999, 47(5): 355-362.

- [5] 马庆贺. 认知训练和视觉环境对汉语听感清晰度评测的影响 [D]. 北京: 中国传媒大学, 2013.
- [6] 国家技术监督局. GB/T 15508-1995 声学、语言清晰度测试方法 [S]. 北京: 中国标准出版社, 1995.

[责任编辑] 李薇

[收稿日期] 2014-07-26