

DOI:10.13232/j.cnki.jnju.2019.05.012

适用于虚拟低音音质的客观评价方法研究

王 旻,林志斌*,卢 晶

(近代声学教育部重点实验室,南京大学声学研究所,南京,210093)

摘 要:由于受平板电视扬声器物理尺寸的限制,嵌入其中的扬声器低频重放能力有限,难以回放令人满意的音效.虚拟低音算法通过诱使人脑从高次谐波中感知基频,使扬声器低频性能的主观感受得到提升.对虚拟低音音质的主观评价耗时耗力,基于 ITU-R BS. 1387 建议书(Perceptual Evaluation of Audio Quality)的模型输出参量(Model Output Variables)、客观等级差异(Objective Difference Grade, ODG)、失真指数(Distortion Index)以及 MPEG-7 标准中的底层描述符音频频谱重心(Audio Spectrum Centroid, ASC),利用多元线性回归分析方法,提出一种适用于虚拟低音音质的客观评价方法.与单独使用 ODG 或 ASC 的客观评价方法相比,多参数优选混合模型的客观评价方法显著提升了主客观相关系数.论文通过实验验证了模型的有效性.

关键词:虚拟低音,音频质量评价,多元线性回归,多参数优选混合模型

中图分类号:O429

文献标识码:A

Research on objective evaluation method for sound quality of virtual bass system

Wang Min, Lin Zhibin*, Lu Jing

(Key Laboratory of Modern Acoustics, Ministry of Education, Institute of Acoustics of Nanjing University, Nanjing, 210093, China)

Abstract: Due to the physical size constraints of flat panel television loudspeakers, low frequency reproduction from these loudspeakers is generally limited and unsatisfactory. The virtual bass algorithm enhances the low frequency performance of the loudspeakers by inducing the human brain to perceive the fundamental frequency from its higher harmonics. The normal subjective listening tests are often time-consuming. In this paper, we introduce an objective evaluation method for virtual bass algorithms based on the Model Output Variables (MOVs), the Objective Difference Grade (ODG), and the Distortion Index (DI) from the ITU Recommendation ITU-R BS.1387, as well as the Audio Spectrum Centroid (ASC), one of the low-level descriptors in MPEG-7 standard. The multiple linear regression analysis method is utilized to exploit these metrics. Compared with the objective evaluation method merely by ODG or ASC, the correlation coefficient between subjective and objective results has been significantly improved by this multi-parameter optimization hybrid model. The efficacy of the proposed method is validated by experiments.

Key words: virtual bass, audio quality evaluation, multiple linear regression, multi-parameter optimization hybrid model

基金项目:国家自然科学基金(11874219)

收稿日期:2019-02-27

* 通讯联系人, E-mail: zblin@nju.edu.cn

随着多媒体音频技术的发展,音频音质的评价问题获得了广泛的关注。目前,音频音质的评价主要有主观评价和客观评价两种方式。主观评价方法通过实验统计分析总结主观感受活动背后的心理潜在规律,将模糊的主观感受进行数值上的量化,用具体的数字表述这些主观感觉或主观感觉之间的相对关系。目前,经典的主观评价实验方法有对偶比较法、排序法、系列范畴法、语义细分法以及量值估计法等^[1],然而主观评测方法耗时耗力,同时实验结果与被试人员当天的状态、情绪以及之前受到的相关听音训练程度等有关。近年来,很多学者陆续提出针对音频信号的客观评价方法。1979年 Schroeder et al^[2]提出噪声响度模型(Noise Loudness, NL);1979年 Schroeder^[3]提出基于滤波器组的听觉频谱差异模型(Auditory Spectral Difference, ASD);1987年 Brandenburg^[4]提出噪声掩蔽比模型(Noise to Mask Ratio, NMR)。在此之后的近十年间,又先后有 Beerends and Stemerding^[5], Paillard et al^[6], Colomes et al^[7], Sporer^[8], Thiede et al^[9]等学者陆续提出关于音频信号的感知模型或者算法。1998年,ITU-R (International Telecommunication Union - Radio-communication, 国际电信联盟无线电通信组)完善和综合了 Disturbance Index(DIX), Noise-to-Mask Ratio(NMR), Perceptual Audio Quality Measure(PAQM), PERCEVAL, Perceptual Objective Measure(POM)和 The Toolbox Approach等六种方案,形成了 ITU-R BS. 1387 (Perceptual Evaluation of Audio Quality, PEAQ)建议^[10]。

音频低音音质评价作为音频音质评价的重要组成部分,近年来也获得了较多的关注^[11-15],其中也囊括了虚拟低音的音质评价。虚拟低音的获取主要是通过诱使人脑从高次谐波中感知基频,主要用于解决小型扬声器低音表现力差的问题。目前虚拟低音算法可以分为两大类:时域算法-非线性器件算法(Nonlinear Devices, NLD)^[16-17]与频域算法-相位声码器算

法(Phase Vocoder, PV)^[18]。虚拟低音算法的音质评价目前以主观评价为主,客观评价方法的研究很少。曾有学者使用总谐波丰满度(Total Harmonic Richness, THR), ASC (Audio Spectrum Centroid) 和 ODG (Objective Difference Grade)等客观参量评价虚拟低音算法的主观音质,但发现两者相关性不高^[13-15]。虚拟低音算法是用谐波的组合来代替基频的听感,谐波成分的多少与虚拟低音听感并非完全正相关关系^[19-20],同时音频材料中缺少真实低频成分,仅用总谐波丰满度或频谱重心评价虚拟低音音质效果不佳。此外,总谐波丰满度和频谱重心均未引入人耳听觉系统模型。PEAQ引入了人耳听觉模型,是用来评价音频质量的相对有效的客观评测工具,本文采用 PEAQ 的 11 个模型输出参量 MOVs(Model Output Variables)以及两个输出值客观差异等级 ODG 和失真指数 DI (Distortion Index),结合频谱重心 ASC,利用多元线性回归分析方法,提出了一种多参数优选混合模型的虚拟低音评价方法。与单独使用 ODG 或 ASC 的客观评价方法相比,多参数优选混合模型的客观评价方法的主客观相关性系数得到了明显提高,并通过实验验证了模型的有效性。

1 主观评价方法与实验

1.1 赋值的语义细分法 语义细分法是评价者运用意义相反的形容词对测试音频进行等级描述,如安静-吵闹、平滑-粗糙、强有力的-微弱的等,并将这些形容词安置在等级的两端,中间使用一些量度性的副词,评价者根据听音感受对测试音频做出评判。

本文研究选用两个与低音相关度最高的主观评价参量丰满度和力度^[21],其中丰满度选取的形容词组为“单薄-丰满”,力度选取的形容词组为“弱-强”。并采用经典的 7 级评价等级,并相应赋值 1~7 分,如表 1 和表 2 所示。

量化打分精确到小数点后一位。同时,为了解决实验过程中存在的“上限效应”(即若当

表 1 丰满度评分表

Table 1 Rating scale of richness

丰满度	极度	非常	有些	都不是	有些	非常	极度
单薄	1	2	3	4	5	6	7
丰满							

表 2 力度评分表

Table 2 Rating scale of strength

力度	极度	非常	有些	都不是	有些	非常	极度
弱	1	2	3	4	5	6	7
强							

前样本量化评分使用极端值,会因为下一个样本量化评分超过当前样本量化评分导致无法对下一个样本评分赋值),量化打分允许突破上限 7 分或低于下限 1 分,但需要保证量化打分的线性化差异^[1].

1.2 实验环境 主观试听实验在南京大学声学所的试听室中进行,其建筑特性如下:体积 123.75 m³,高 3.0 m,长 7.5 m,宽 5.5 m,平均混响时间约为 0.3 s. 实验采用耳机 AKG K702 进行重放.

1.3 试听人员 试听人员共 15 人,其中男 12 人,女 3 人,年龄 22~28 岁,所有试听人员均为南京大学声学所研究生,听力正常,且都有听音经历.

1.4 音频材料 音频材料均选取低频成分丰富的流行音乐,包含人声(男声、女声)、弦乐和打击乐等多种无损音频材料(表 3). 主观试听实验的音频材料包含音频原音与经过虚拟低音算法处理后的音频材料.

1.5 实验步骤 试听实验评价两个主观参量:丰满度和力度. 试听实验含训练实验和正式实验,首先,在训练实验中,依次连续播放存在明显差异的音频材料,帮助试听者熟悉评价参量的主观感受并构建线性化心理尺度,以便于正式实验的量化打分;其次,在正式实验中,试听者量化评分每段音频材料的丰满度与力度,试听过程中可多次重放任何一个音频材料或比较任何两个音频材料. 每次试听实验只有一个试听者参与,单一试听者试听时长不超过 30 min.

表 3 音频材料

Table 3 Audio stimuli

名称	时长(s)
Piano Beat	23
渡口	31
渡口前奏	13
光辉岁月	19
被遗忘的时光	33
炎黄第一鼓片段一	24
炎黄第一鼓片段二	23

1.6 结果处理分析 由于不同试听者的量化评分结果所遵循的比例尺度不同,统计分析之前需对评价结果进行标准化处理. 本文采用的标准化处理方法是百分制法,规定每一个试听者评分的最大值为 100,最小值为 0,中间分值按照比例进行转换.

主观试听实验的最终结果为主观差异等级 (Subjective Difference Grade, SDG):

$$SDG = Score_{TS} - Score_{RS} \quad (1)$$

其中, $Score_{TS}$ 为虚拟低音算法处理后的音频材料的主观评分, $Score_{RS}$ 为原音音频材料的主观评分. 试听数据为 15 个试听者的九组音频材料的量化评分.

数据处理中剔除个体评分与全体平均得分相差大于三倍标准差的相关数据并重新计算该音频材料平均得分^[22],表 4 给出了力度和丰满度的主观评价平均分数值,图 1 给出了 95% 置信区间的力度和丰满度的主观评价平均分数线.

2 客观评价系统的建立

为了获得精确的客观评价系统,本文研究选取了 PEAQ 的 11 个模型输出参量 MOVs、客观差异等级 ODG、失真指数 DI 和 MPEG-7 标准中的底层描述符音频频谱重心 ASC 作为客观评价系统的客观参量,利用多元线性回归分析方法,拟合得出了虚拟低音主观评价参量中的丰满度与力度的客观评价表达式.

表4 力度、丰满度主观试听实验评价均分

Table 4 Mean subjective scores of strength and richness

编号	名称	力度	丰满度
P1	Piano Beat 7	35.80	37.60
P2	Piano Beat 10	61.96	63.74
P3	渡口 7	33.33	35.20
P4	渡口 10	59.83	59.20
P5	渡口前奏 10	62.12	54.63
P6	光辉岁月 10	42.35	36.67
P7	被遗忘的时光 10	29.46	42.83
P8	炎黄第一鼓片段一 10	40.90	44.42
P9	炎黄第一鼓片段二 10	44.01	48.01

2.1 PEAQ ITU-R BS. 1387 标准 (PEAQ)

是目前业界广泛应用的音频质量客观评价方法. 本文采用了PEAQ中的11个模型输出参量 MOVs, 参数含义见表5.

PEAQ中采用人工神经网络 ANN (Artificial Neural Network) 整合模型输出参量 MOVs, 得到输出参数 DI 和 ODG, 作为最终的音质客观评价结果^[10]. 其中 ANN 的激活函数 (activation function) 为:

$$\text{sig}(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

该 ANN 包括 I 个输入, 并且 ANN 的隐层 (hidden layer) 中有 J 个节点. 在映射过程定义

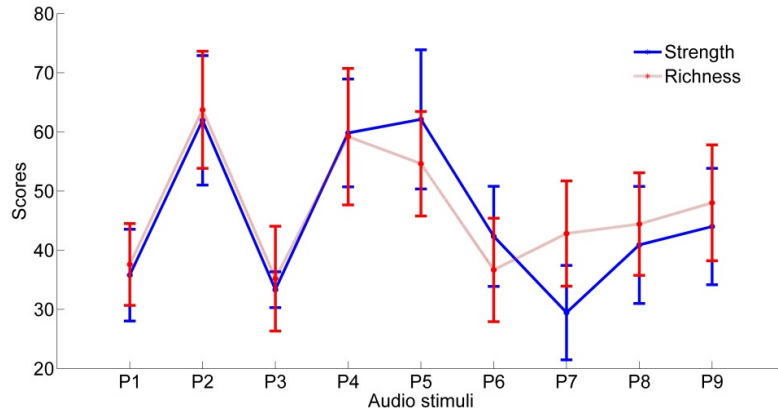


图1 力度、丰满度主观试听实验评价均分(95%置信区间)

Fig. 1 Mean subjective scores of strength and richness with 95% confidence intervals

集合 $a_{\min}[i]$, $a_{\max}[i]$ 作为输入比例系数, $w_x[i, j]$ 作为输入加权, $w_y[j]$ 作为输出加权以及一对输出比例系数 b_{\min} 和 b_{\max} . 其中, 失真指数 DI:

$$\text{DI} = w_y[j] + \sum_{j=0}^{J-1} \left(w_y[j] \text{sig} \left(w_x[I, j] + \sum_{i=0}^{I-1} w_x[i, j] \frac{x[i] - a_{\min}[i]}{a_{\max}[i] - a_{\min}[i]} \right) \right) \quad (3)$$

DI 可以直接与 ODG 相联系:

$$\text{ODG} = b_{\min} + (b_{\max} - b_{\min}) \text{sig}(\text{DI}) \quad (4)$$

2.2 ASC 音频频谱重心 ASC 是 MPEG-7 标准^[23]中的底层描述符之一, 是音频材料对数频率功率谱的重心. ASC 是根据信号的功率谱计算的, 同时为防止直流分量的存在以及极低频

分量的比重过低, 将 62.5 Hz 以下的成分作为一个整体, 为:

$$K_{\text{low}} = \text{floor} \left(62.5 \frac{N_{\text{FFT}}}{F_s} \right) \quad (5)$$

$$P'(k) = \begin{cases} \sum_{k'=0}^{K_{\text{low}}} P(k') (k'=0) \\ P(k' + K_{\text{low}}) \left(1 \leq k' \leq \frac{N_{\text{FFT}}}{2} - K_{\text{low}} \right) \end{cases} \quad (6)$$

$$f'(k) = \begin{cases} 31.25 (k'=0) \\ f(k' + K_{\text{low}}) \left(1 \leq k' \leq \frac{N_{\text{FFT}}}{2} - K_{\text{low}} \right) \end{cases} \quad (7)$$

表 5 模型输出参量 MOVs

Table 5 Model output variables

MOVs	各参量含义
BandwidthRef	参考信号带宽
BandwidthTest	测试信号带宽
TotalNMR	总信掩比
WinModDiff1	窗选平均调制差异
ADB	总失真与总严重失真比的对数值
EHS	误差谐波结构
AvgModDiff1	平均调制差异 1
AvgModDiff2	平均调制差异 2
RmsNoiseLoud	平均噪声响度的均方根值
MFPD	检测到失真存在的最大滤波概率
RelDistFrames	受干扰帧的相对分数

其中, $\text{floor}(x)$ 表示对 x 取整, F_s 为采样频率, N_{FFT} 为 FFT 的长度, $P'(k')$ 为调整后的信号功率谱, $f'(k')$ 为调整后的信号离散化频谱.

ASC 可由功率谱 $P'(k')$ 和频谱 $f'(k')$ 计算而得:

$$ASC = \frac{\sum_{k=0}^{\frac{N_{FFT}}{2} - K_{low}} \log_2 \left(\frac{f'(k')}{1000} \right) P'(k')}{\sum_{k=0}^{\frac{N_{FFT}}{2} - K_{low}} P'(k')} \quad (8)$$

2.3 多参数优选混合模型的客观评价系统

首先将所选取的音频材料经过虚拟低音算法进行不同程度的虚拟低音增强处理,并分别计算获得每组原音和虚拟低音增强后的音频材料 PEAQ 的 11 个模型输出参量 MOVs 以及 PEAQ 的两个映射输出值 ODG, DI^[24] 和音频频谱重心 ASC 共 14 个客观测量值.

主客观评价参数的相关性分析表明,力度评分在 0.05 双尾检测显著性水平^[25]与 ODG, DI 相关性显著;而丰满度评分在 0.01 双尾检测显著性水平与 ODG, DI 相关性显著,在 0.05 双尾检测显著性水平与 ADB 相关性显著.

为进一步描述虚拟低音主客观评价参量的相关性,本文通过多元线性回归分析,采用多参

数优选混合模型,获得了丰满度和力度的拟合模型表达式.

在多元线性回归分析过程中,为了避免多重共线性,需要对自变量进行筛选,选择自变量的原则是对统计量进行显著性检验,检验的根据是:将一个或一个以上的自变量引入回归模型中时,是否使残差平方和显著减小.如果增加一个自变量使残差平方和显著减小,则说明有必要将这个变量引入回归模型中.自变量的选择方式分为三种:向前选择,向后剔除和逐步回归.向前选择方法的步骤为:先将各个自变量分别与因变量进行一元线性回归,找到 F 统计量的值最大的模型及其对应的自变量并将其引入模型,然后再加入另一个自变量,挑选出 F 值最大的含有两个自变量的模型,依次循环,直到增加自变量不能导致残差平方和显著减小为止.向后剔除方法的步骤为:先对所有的自变量进行多元线性回归建立模型,然后考察去掉一个自变量的模型,使模型的残差平方和减小最少的自变量被挑选出来从模型中剔除,依次循环,直到剔除自变量不会使残差平方和显著减小为止.逐步回归方法是向前选择和向后剔除这两种方法的结合:在增加了一个自变量后,会对模型中所有自变量进行考察,如果增加一个新的自变量后,前面增加的某个自变量对模型的贡献变得不显著,前面这个自变量就会被剔除.按照该方法不停地增加自变量并考虑剔除以前增加的自变量的可能性,直到增加自变量不能导致残差平方和显著减小为止.在前面步骤中增加的自变量在后面的步骤中有可能被剔除,而在前面步骤中剔除的自变量在后面的步骤中也有可能重新进入到模型中^[26].本文采用了逐步回归的方法对自变量进行了筛选,得到了丰满度和力度的多元线性回归拟合模型.

其中,丰满度模型为:

$$C_r = 29.350 - 29.125ODG - 0.001BandwidthRef \quad (9)$$

表 6 为丰满度模型摘要,表中的 R 表示模型的拟合优度,用来衡量估计的模型对观测值的拟合程度, R 值越接近 1 表明模型拟合程度

越高. 表7为丰满度模型方差分析(Analysis of Variance, ANOVA)结果, 其中显著性数值小于0.05则表明拟合模型是有效的.

表6 丰满度模型摘要

Table 6 Summary of richness model

R	R 方	调整后 R 方	标准估算的 误差
0.923	0.852	0.802	4.571

表7 丰满度模型方差分析

Table 7 Variance analysis of richness model

模型	平方和	自由度	均方	F	显著性
回归	721.008	2	360.504	17.251	0.003
残差	125.382	6	20.897		
总计	846.390	8			

力度模型为:

$$C_s = 39.708 - 45.129\text{ODG} - 114.626\text{RmsNoiseLoud} - 23.646\text{EHS} + 40.304\text{ASC} + 17.560\text{RelDistFrames} \quad (10)$$

表8为力度模型摘要, 表9为力度模型方差分析结果. 从表中数据可以看出, 丰满度模型和力度模型的拟合优度高, 分别为0.923和0.999. 模型的显著性均低于0.05, 验证了模型的有效性.

为进一步验证模型的准确性, 本文将多参数优选混合模型的丰满度和力度的客观评价模型与单独使用ODG或ASC客观评价方法进行了比较, 主客观皮尔逊相关性系数和双尾显著性见表10. 从表中可以看出, 多参数优选混合模型的相关性系数明显高于ODG或ASC的客

观评价方法, 多参数优选混合模型得到的客观评分与主观感受更加吻合.

表8 力度模型摘要

Table 8 Summary of strength model

R	R 方	调整后 R 方	标准估算的误差
0.999	0.998	0.995	0.918

表9 力度模型方差分析

Table 9 Variance analysis of strength model

模型	平方和	自由度	均方	F	显著性
回归	1282.414	5	256.483	304.567	0.000
残差	2.526	3	0.842		
总计	1284.941	8			

为了进一步验证模型的有效性, 本文选取了三组不同于训练集的实验数据作为验证集. 并引入了曲线趋势拟合优度 R^2 这一评价指标来衡量验证集的实验数据与客观模型之间的匹配程度, 其参数定义如式(11):

$$R^2 = 1 - \frac{\sum_i [(y_i - Y_i) - (\bar{y} - \bar{Y})]^2}{\sum_i y_i^2} \quad (11)$$

其中 Y 为通过式(9)和式(10)计算获得的客观模型的结果, y 为通过主观试听实验得到的结果, \bar{Y} 和 \bar{y} 分别表示均值.

表11给出了验证集中3个实验数据样本与客观拟合模型的丰满度模型拟合优度 R_r^2 和力度模型拟合优度 R_s^2 . 从表中数据可以看出, 验证集实验数据与客观模型的拟合优度均高于0.85, 拟合程度高, 验证了拟合模型的有效性.

表10 客观评价方法的主客观皮尔逊相关性系数和双尾显著性(括号内的数)比较

Table 10 Pearson's linear correlation coefficients between mean subjective and predicted scores of different objective evaluation methods (two-tailed significance in brackets)

	ODG	ASC	C_r	C_s
丰满度	-0.837(0.005)	0.596(0.090)	0.923(0.003)	0
力度	-0.688(0.041)	0.357(0.346)	0	0.999(0.000)

表 11 验证集与模型之间的拟合优度

Table 11 Goodness of fit between validation set and model

	1	2	3
R_r^2	0.932	0.975	0.851
R_s^2	0.962	0.952	0.872

3 结 论

本文在 PEAQ 的 11 个模型输出参量 MOVs 以及 PEAQ 的两个输出值 ODG, DI 和音频频谱重心 ASC 的基础上, 利用多元线性回归分析方法, 提出了一种多参数优选混合模型用于虚拟低音评价的客观评价方法, 分别得到了关于丰满度和力度的客观评分拟合表达式. 与单独使用 ODG 或 ASC 的客观评价方法相比, 多参数优选混合模型的客观评价方法的主客观相关性系数得到了明显提高, 客观评分与主观感受更加贴近, 并通过实验验证了模型的有效性.

致 谢 对所有参加主观试听实验的南京大学声学所的同学表示诚挚的谢意.

参考文献

- [1] 孟子厚. 音质主观评价的实验心理学方法. 北京: 国防工业出版社, 2008, 227.
- [2] Schroeder M R, Atal B S, Hall J L. Optimizing digital speech coders by exploiting masking properties of the human ear. The Journal of the Acoustical Society of America, 1979, 66(6): 1647—1652.
- [3] Karjalainen M. A new auditory model for the evaluation of sound quality of audio systems//1985 IEEE International Conference on Acoustics, Speech, and Signal Processing. Tampa, FL, USA: IEEE, 1985, 10: 608—611.
- [4] Brandenburg K. Evaluation of quality for audio encoding at low bit rates//Audio Engineering Society Convention 82. London, UK: Audio Engineering Society Press, 1987.
- [5] Beerends J G, Stemerdink J A. A perceptual audio quality measure based on a psychoacoustic sound representation. Journal of the Audio Engineering Society, 1992, 40(12): 963—974.
- [6] Paillard B, Mabilieu P, Morissette S, et al. Perceval: Perceptual evaluation of the quality audio signals. Journal of the Audio Engineering Society, 1992, 40(1—2): 21—31.
- [7] Colomes C, Lever M, Rault J B, et al. A perceptual model applied to audio bit - rate reduction. Journal of the Audio Engineering Society, 1995, 43(4): 233—240.
- [8] Sporer T. Objective audio signal evaluation - applied psychoacoustics for modeling the perceived quality of digital audio//Audio Engineering Society Convention 103. New York, NY, USA: Audio Engineering Society Press, 1997.
- [9] Thiede T, Kabot E. A new perceptual quality measure for bit rate reduced audio//Audio Engineering Society Convention 100. Copenhagen, Denmark: Audio Engineering Society Press, 1996.
- [10] ITU - R Recommendation BS. 1387—1. Method for objective measurements of perceived audio quality. 2001.
- [11] Creusere C D, Kallakuri K D, Vanam R. An objective metric of human subjective audio quality optimized for a wide range of audio fidelities. IEEE Transactions on Audio, Speech, and Language Processing, 2008, 16(1): 129—136.
- [12] Creusere C D, Hardin J C. Assessing the quality of audio containing temporally varying distortions. IEEE Transactions on Audio, Speech, and Language Processing, 2011, 19(4): 711—720.
- [13] Oo N, Gan W S. Harmonic analysis of nonlinear devices for virtual bass system. In: 2008 IEEE International Conference on Audio, Language and Image Processing. Shanghai, China: IEEE, 2008: 279—284.
- [14] Mu H, Gan W S, Tan E L. A timbre matching approach to enhance audio quality of psychoacoustic bass enhancement system//2013

- IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, Canada:IEEE,2013:36—40.
- [15] Mu H, Gan W S, Tan E L. An objective analysis method for perceptual quality of a virtual bass system. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 2015, 23(5): 840—850.
- [16] Lim W T, Oo N, Gan W S. Synthesis of polynomial-based nonlinear device and harmonic shifting technique for virtual bass system//2009 IEEE International Symposium on Circuits and Systems. Taipei, China:IEEE,2009.
- [17] Arora M, Jang S, Moon H. Low complexity virtual bass enhancement algorithm for portable multimedia device//Audio Engineering Society Conference: 29th International Conference: Audio for Mobile and Handheld Devices. Seoul, Korea: Audio Engineering Society Press,2006.
- [18] Bai M R, Lin C. Synthesis and implementation of virtual bass system with a phase - vocoder approach. *Journal of the Audio Engineering Society*,2006,54(11):1077—1091.
- [19] Oo N, Gan W S. Harmonic and intermodulation analysis of nonlinear devices used in virtual bass systems//Audio Engineering Society Convention 124. Amsterdam, Netherlands: Audio Engineering Society Press,2008.
- [20] Mu H, Gan W S. Perceptual quality improvement and assessment for virtual bass systems. *Journal of the Audio Engineering Society*, 2015, 63(11): 900—913.
- [21] 秦伶俐,包紫薇. 影响音质主观感觉的因素. *电声技术*,1990(5):1—11. (Qin L J, Bao Z W. Factors of affecting the subjective perception of sound quality. *Audio Engineering*,1990(5):1—11.)
- [22] 国家技术监督局. 声学·语言清晰度测试方法. GB/T 15508 - 1995. (State Bureau of Technical Supervision. Acoustics-Speech articulation testing method. GB/T 15508-1995.)
- [23] Kim H G, Moreau N, Sikora T. MPEG-7 audio and beyond: audio content indexing and retrieval. New York:John Wiley & Sons Press,2006.
- [24] Kabal P. An examination and interpretation of ITU-R BS. 1387: Perceptual evaluation of audio quality. Technical Report. TSP Lab, Electrical & Computer Engineering Department, McGill University, Montreal, Canada,2002.
- [25] 盛骤,解式千,潘承毅. 概率论与数理统计. 第4版. 高等教育出版社,2008,428.
- [26] Pallant J. SPSS survival manual: a step by step guide to data analysis using SPSS for windows (version 15). Maidenhead: Open University Press, 2007,334.

(责任编辑 章 强)