(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 109565633 B
(45)授权公告日 2022.02.11

(21)申请号 201780024940.6
(22)申请日 2017.04.20
(65)同一申请的已公布的文献号
申请公布号 CN 109565633 A
(43)申请公布日 2019.04.02
(30)优先权数据
20165348 2016.04.20 FI
(85)PCT国际申请进入国家阶段日
2018.10.22
(86)PCT国际申请的申请数据
PCT/FI2017/050300 2017.04.20
(87)PCT国际申请的公布数据
WO2017/182716 EN 2017.10.26

(73)专利权人 珍尼雷克公司
地址 芬兰伊沙尔米
(72)发明人 哈维尔·戈麦斯-博拉诺斯
阿基·梅基维塔 维莱·普尔基
(74)专利代理机构 北京聿宏知识产权代理有限公司 11372
代理人 吴大建 张杰

(51)Int.Cl.
*H04R 5/033*(2006.01)
*H04S 3/00*(2006.01)
*H04R 1/10*(2006.01)
审查员 陈世元

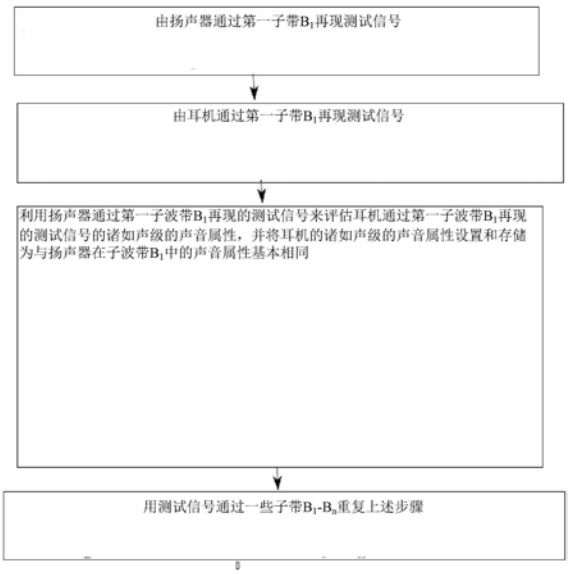
权利要求书2页 说明书26页 附图14页

(54)发明名称

有源监听耳机及其双声道方法

(57)摘要

根据本发明的示例方法,提供了一种用于形成立体声耳机的双耳滤波以保持耳机的声音质量的方法,由此从扬声器到每只耳朵的直接路径和串扰路径的总和具有平坦的幅度响应。




CN 109565633 B

1.一种形成用于立体声耳机的双耳滤波器以保持所述耳机的音质的方法，其特征在于，形成从扬声器到每只耳朵的直接路径和串扰路径的加和，以使得振幅大致上不作为频率的函数而变化，这也称为全通系统，

其中，双耳滤波器形成为使得通过有利地为340ms的预定混响时间针对试听室内的立体声扬声器装置来测量仿真头部的双耳时间响应$h_{ij}(t)$，并且使用所测量的响应，通过对所述响应的42ms的第一预设时间进行窗口化来设计一组双耳滤波器$H_{bin}$，

$$H_{\mathrm{bin}} = \mathscr{F}\left\{h_{\mathrm{ij}}(t)w(t)\right\}, i \in \{L,R\}, j \in \{l,r\}, \tag{15}$$

其中，$\mathscr{F}\{\cdot\}$表示傅里叶变换，L和R表示左右扬声器，l和r分别表示左耳和右耳，以及w(t)是预定义的42ms时长的时间窗口，并且，在进行非正式听力测试之后，采用这种滤波器长度作为外在化能力和由于室内混响引起的音色效果之间的最佳折中。

2.根据权利要求1所述的方法，其中，对耳机应用而言，与恒定振幅值的偏差小于+/-3dB，或者小于+/-0.1dB。

3.根据权利要求1或2所述的方法，其中，仅进行相位均衡。

4.根据权利要求1所述的方法，其中，使用均衡后的$H_{bin}$作为双耳滤波器。

5.根据权利要求1所述的方法，其中，使用均衡后的$H_{ph}$作为双耳滤波器，其中获得$H_{ph}$的方法包括：

-通过将根据直接路径和串扰路径的平滑响应的幅度比得到的结果幅度求平均，来估计直接路径和串扰路径之间的水平差异$H_{LD}$，

-直接滤波器和串扰滤波器的幅度分别设计成$Hd_{ph}$和$Hx_{ph}$

$$|H_{\mathrm{d_{ph}}}| = \frac{1}{H_{\mathrm{LD}}+1} \quad , \quad |H_{\mathrm{x_{ph}}}| = \frac{H_{\mathrm{LD}}}{H_{\mathrm{LD}}+1}$$

-将双耳全通滤波器与其对应的$Hd_{ph}$滤波器和$Hx_{ph}$滤波器进行卷积以生成双耳滤波器$H_{ph}$。

6.一种用于校准根据上述权利要求中的任一项所述的立体声耳机(1)的方法，所述立体声耳机(1)包括具有存储器和信号处理特性的放大器(2)，所述方法包括步骤：比对设定的参考耳罩或驱动器来校准所述立体声耳机(1)的驱动器或耳罩，并且将校准设置存储在所述放大器(2)的存储器中。

7.根据权利要求6所述的方法，其中，通过设置所述放大器(2)中的信号处理参数来确定所述立体声耳机(1)的期望的声音属性，以便通过测量或基于从所述立体声耳机(1)的用户处接收的输入信息来获得所述期望的声音属性。

8.根据权利要求6所述的方法，其中，所述方法包括如下工厂校准步骤：校准至少幅度响应。

9.根据权利要求7所述的方法，其中，所述声音属性包括以下特性中的至少一个：“频率响应”、“时间响应”、“相位响应”或“灵敏度”。

10.根据权利要求9所述的方法，其中，所述期望的声音属性基于特定房间的耳机系统的校准参数来确定。

11.一种存储有计算机程序的计算机可读介质，所述计算机程序配置为促使根据权利

要求1-10中任一项所述的方法被执行。

# 有源监听耳机及其双声道方法

## 技术领域

[0001] 本发明涉及有源监听耳机和涉及这些耳机的方法。

## 背景技术

[0002] 大多数耳机是无源的，因此性能取决于所使用的外部放大器。因此，各个单元之间以及各个设计之间的性能差异很大。有一些有源耳机，其具有置于耳机耳罩内的电子器件。电子器件占用空间并(经常)降低声学性能。电学功能只是放大器，或放大器和ANC(有源噪声消除器)。获得计算机/数字音频/模拟音频信号所必需的接口造价昂贵。有两种类型的耳机：开放式耳机和封闭式耳机。虽然开放式耳机有其自身的优点，但它们对环境噪声的衰减很差，这妨碍听到音频材料中的细节(并且环境声学甚至可能影响到耳机的音频)，但据说开放式耳机设计可避免时常与封闭式耳机设计相关的“盒子”声音(音频着色)和有限的低频扩展。另外在封闭式耳机中，用户的收听仅限于耳罩区域，因此用户之间的沟通可能受阻。

[0003] 当使用耳机来补充和继续播放也能使用扬声器播放的任务时，需要设计耳机和相关的信号处理，使得耳机的校准具有与房间内基于监听系统的扬声器的声音相同的声音特性，以便当从一个系统切换到另一个系统时音质能保持一致。

## 发明内容

[0004] 本发明涉及有源监听耳机(AMH)及其校准方法。

[0005] 本发明由独立权利要求的特征限定。一些特定实施例在从属权利要求中限定。

[0006] 根据本发明的第一方面，提供了一种用于自动校准有源监听耳机的方法，该有源监听耳机包括具有存储器和信号处理特性的放大器，该方法包括如下步骤：确定耳机(1)的期望声音属性；在放大器(2)中设置信号处理参数和校准算法，以便通过测量或基于从耳机用户所接收的输入信息获得期望的声音属性。

[0007] 根据本发明的第二方面，提供了一种方法，其中声音属性包括以下特征中的至少一个：“频率响应”、“时间响应”，“相位响应”或“声级”。

[0008] 根据本发明的第三方面，提供了一种方法，其中基于用于特定房间的扬声器系统的校准参数并根据房间中的声学测量，来确定诸如频率响应的期望声音属性。

[0009] 根据本发明的第四方面，提供了一种方法，其中测试信号通过软件或硬件接口启动、由放大器或接口设备产生并由扬声器通过第一子波带($B_1$)再现。该测试信号由耳机(1)通过第一子波带($B_1$)再现，利用扬声器通过第一子波带($B_1$)再现的测试信号来评估耳机(1)通过第一子波带($B_1$)再现的测试信号的诸如声级的声音属性，并将耳机的诸如声级的声音属性设置和存储为与扬声器在子波带$B_1$中的声音属性基本相同，通过几个子波带$B_1$-$B_n$用测试信号重复上述过程。

[0010] 根据本发明的第五方面，提供了一种方法，其中测试信号是粉红噪声。

[0011] 根据本发明的第六方面，提出，其中测试信号是包括具有宽频谱含量的音频信号

的类音乐音频文件。

[0012] 根据本发明的第七方面，提供了一种方法，其中测试信号的持续时间为1-10秒。

[0013] 根据本发明的第八方面，提出，其中测试信号是连续地重复。

[0014] 根据本发明的第九方面，提供了一种有源监听耳机系统，包括耳机和通过电缆连接到耳机的放大器，该系统包括罩耳式耳罩、用于在放大器(2)中进行信号处理的机构、用于在放大器(2)中存储至少两个预定义的均衡设置的机构，以及用于消除200Hz以下的频率中噪声的机构。

[0015] 根据本发明的第十方面，提供了一种有源耳机系统，其中耳机和耳机放大器是通过线缆彼此连接的分开的独立单元。

[0016] 根据本发明的第十一方面，提供了一种有源耳机系统，其中耳机的每个驱动器或耳罩参照设定的参考耳罩或驱动器进行工厂校准，并存储在放大器的存储器中，由此工厂校准使得耳机系统中的所有耳罩在声学上基本相同，例如基于设定的参考耳罩或驱动器而响应相同、响度相同。

[0017] 根据本发明的第十一方面，提供了一种有源耳机系统，其中耳机放大器和耳机是基于工厂校准的唯一性配对。

[0018] 所要求保护的发明涉及如何通过最小化耳朵附近的物理声音再现的变化来使换能器(驱动器)的声音从第一收听环境(扬声器)向第二收听环境(耳机)均衡的技术效果。

[0019] 换句话说，本发明创造了一种在听众耳朵处变化最小的情况下如何将为扬声器创建的声音信息均衡到耳机驱动器的技术方案。

## 附图说明

[0020] 图1示出了根据本发明的至少一些实施例的一个有源耳机。

[0021] 图2示出了根据本发明如何将音频信号划分成子波带的图形。

[0022] 图3以框图示出了根据本发明的一种校准方法的一个实施例。

[0023] 图4以框图示出了根据本发明的电子器件的一个实施例。

[0024] 图5以框图示出了根据本发明的软件的一个实施例。

[0025] 图6示出了根据本发明的系统的第一布局。

[0026] 图7示出了根据本发明的系统的第二布局。

[0027] 图8示出了重新定位对耳机均衡的影响。使用公式1的耳机响应的反演滤波器用于补偿重新定位耳机后测量的两个响应。对于低于2kHz的频率没有明显的差异。

[0028] 图9示出了使用直接反演(DI)、具有β＝0.01的正则化反演(RI)以及维纳反卷积(WI)的耳机响应的反演。

[0029] 图10示出了使用公式6(实线)和公式7(虚线)定义的α(ω)的正则化参数β(ω)的值，并且$\hat{H}(\omega)$是半倍频程平滑版的耳机响应。

[0030] 图11示出了使用直接反演(虚线)和所提出的sigma反演方法(实线)的耳机响应的反演。

[0031] 图12a示出了放置在开放式耳道内的微型扩音器的示意图。

[0032] 图12b示出了扩音器引线的图片，其围绕耳廓弯曲并且在两个位置处用胶带固定以避免在放置耳机时扩音器位移。

[0033] 图13示出了表示用于公式9的参数的表格，以使用维纳反卷积(WI)、常规正则化反演(RI)、复杂平滑法(SM)和提出的sigma反演(SI)等方法来获得耳机响应的反演。

[0034] 图14示出了测量了四次并在测量之间重新定位耳机的耳机的归一化幅度响应。在每次测量之前，受试者自己移除并重新应用耳机。第一次测量用于反演(实线)。其他三个响应用虚线、点划线和阴影线表示。在低于2kHz的频率没有明显的差异。

[0035] 图15示出了使用利用Wiener维纳反卷积(WI)法、传统正则化反演法(RI)、复杂平滑法(SM)和提出的sigma反演法(SI)获得的反演滤波器来补偿单个耳机响应的效果。对于低于2kHz的频率没有明显的差异。

[0036] 图16示出了当三次重新定位耳机时使用利用维纳反卷积(WI-顶部的框)、正则化反演方法(RI-从顶部开始的第二个框)、复杂平滑方法(SM-从顶部开始的第三个框)和提出的方法(SI-底部的框)获得的反演滤波器所补偿的响应的稳定性。对应于第一、第二和第三测量的所补偿的响应分别表示为实线、点线和虚线。对于低于2kHz的频率没有明显的差异。

[0037] 图17示出了显示针对无耳机均衡(NF)、传统正则化反演(RI)、平滑方法(SM)和提出的方法(SI)中每种反演方法而应用于10个受试者所获得的平均评分μ和标准偏差(SD)的表格。

[0038] 图18示出了显示使用Games-Howell过程的多重比对测试的p值的表格。识别出的方法为:无耳机均衡(NF)、传统正则化反演(RI)、平滑方法(SM)和提出的方法(SI)。

[0039] 图19示出了应用于10个受试者所计算的反演方法的平均值及这些平均值的95％置信区间。方法为无耳机均衡(NF)、传统正则化反演方法(RI)、平滑方法(SM)和提出的方法(SI)。

[0040] 图20示出了扬声器立体声设置的双耳道渲染的示意图。

[0041] 图21示出了通过放置在中心的幻象源的耳机的双耳立体声再现的示意图。

[0042] 图22示出了通过放置在中心的幻像源的立体声信号的耳机的直接再现的示意图。仅显示一只耳朵。

[0043] 图23示出了通过幻象源完全向左平移的耳机的双耳立体声再现的示意图。

[0044] 图24示出了通过放置在中心的幻象源的响应得到均衡下的耳机的双耳立体声再现的示意图。

[0045] 图25示出了由滤波器$H_{d_{ph}}$(实线)和$H_{x_{ph}}$(虚线)引入的增益。

[0046] 图26示出了基于Kirkeby,O.发表于2002年音频工程学会会议第22届国际会议“虚拟、综合及娱乐音频”中的“A Balanced Stereo Widening Network for Headphones”的滤波器$H_{d_k}$(实线)和$H_{x_k}$(虚线)引入的增益。

[0047] 图27示出了在对左耳处的直接路径和串扰路径求和之后的均衡滤波器的一个倍频程平滑幅度响应。$H_{binEQ}$、$H_{phEQ}$以及$H_{roomEQ}$的响应分别以实线、阴影线和虚线示出。

[0048] 图28示出了表示空间质量测试(测试1)的事后测试结果的表格。从分析中移除了低锚节点。小于$2\times10^{-3}$的p值被舍入为零并且大于$\alpha=0.05$的p值以粗体表示。

[0049] 图29示出了空间质量测试结果。示出了测试1中每个案例获得的四分位数和中位数。方框中的缺口表示中位数的95％置信区间。$H_{bin_}$用作参考(评分＝100)。

[0050] 图30示出了显示音色/声音平衡质量测试(Test2)的事后测试的结果的表格。从分

析中移除了低锚节点。小于$2\times10^{-3}$的p值被舍入为零并且大于α＝0.05的p值以粗体表示。

[0051] 图31示出了音色/声音平衡质量测试结果。示出了在测试2中为每个案例得到评分的四分位数和中位数表征。方框中的缺口表示中位数的95％置信区间。通过耳机的立体声信号直接再现被用作参考(评分＝100)}。

[0052] 图32示出了显示用于总体质量测试(测试3)的事后测试的结果的表格。从分析中移除了低锚节点。小于$2\times10^{-3}$的p值被舍入为零并且大于α＝0.05的p值以粗体表示。

[0053] 图33示出了整体质量测试结果。示出了测试3中每个案例获得的评分的四分位数和中位数表征。方框中的缺口表示中位数的95％置信区间。

## 具体实施方式

[0054] 定义

[0055] 在本上下文中,术语“音频范围”是20Hz至20kHz的频率范围。

[0056] 在本上下文中,术语“子波带”$B_n$表示在比音频频率范围窄的音频范围内的通带。

[0057] 在本上下文中,“评估声音特性”的定义是指通过使用扩音器进行测量或者通过人进行主观判定。

[0058] 在本上下文中,“声音属性”的定义包括“频率响应”、“时间响应”、“相位响应”、“音级”和“子波带内的频率提升”定义。

[0059] 当使用耳机来补充和继续播放也能使用扬声器播放的监听任务时,需要设计耳机和相关的信号处理,使得耳机的校准具有与房间内基于监听系统的扬声器的声音相同的声音特性。这对于确保从一个监控系统切换到另一个监控系统时监听质量尽可能保持一致是必要的。

[0060] 图1示出了根据本发明的至少一些实施例的一个有源监听耳机,其中,具有用于双耳的驱动器的有源监听立体声耳机1借助于连接电缆3连接到耳机放大器2。方框60描述了该实施例的特征,即工厂校准,其中耳机1的每个驱动器相对于所述参考进行电子均衡,以使每个耳朵的驱动系统各自具有与参考相同的响应,根据本发明的至少一些实施例,这去除了每个耳朵的驱动器系统之间的任何差异以及进行动态控制,保护用户免受过高的声级。

[0061] 在一优选的实施例中,耳机是这样的:它包括两个耳罩,每个耳罩从所有侧面(罩耳式)围绕耳朵,使得所使用的耳罩的类型是在音频范围内闭合的,从而对环境声音或噪声提供声学衰减。根据本发明的耳机线缆的连接器是四(或更多)引脚连接器,允许电子信号分别接入耳机内的每个驱动器。然后,如果在耳机的每个耳罩内使用多个驱动器,则耳机放大器可以单独应用校准,并且还可以进行交叉滤波。

[0062] 增强的有源LF(低频)隔离(EAI)使用通过耳机线缆中的额外导体附接到耳罩外部或内部的扩音器,允许耳机放大器接入扩音器信号。耳机放大器以频率选择性增益来反演并放大扩音器信号,并将该反演的信号添加到馈送到耳机驱动器的信号中,使得泄漏到耳罩内部的噪声被衰减或完全消除。增益的频率选择性使得该衰减主要作用在低频上,更具体地说,作用在低于500Hz的频率上。通过这样方式,朝着低频方面来加强封闭式耳机设计的典型无源衰减减小,产生出能与耳机放大器结合地使低频显著衰减的耳机。

[0063] 通常,耳机的机械低频声音隔离不好。本发明的一些实施例可以使用电子增强来

改善LF隔离。目的是能够更加详细地听取LF处的音频细节。通常，此增强功能工作在200Hz(波长1.7米)以下。在实际实施中，至少一个耳罩包括扩音器。扩音器带宽有限，以消除中间范围内的噪声增加。扩音器信号通过耳机线缆发送回耳机放大器。负反馈应用于放大器的模拟部分，以降低耳机内部能听见的低频音级。针对低频的耳机隔离看起来有所增加。因此，根据本发明的耳机的明显声音隔离看起来比现有技术更好。

[0064] 工厂校准

[0065] 在一个优选的实施例中，工厂校准用于耳机的每个驱动器。工厂校准使耳机中的所有耳罩完全相同，基于设置的参考驱动器或耳罩的响应相同、响度相同。这也将每个耳罩的灵敏度设置为完全相同。工厂校准对于每个耳机和耳机的耳罩都是独一无二的，因此耳机放大器和耳机是唯一的一对，就如同放大器和附件可用于有源监听扬声器。因此，不能将任意耳机放大器与任意其他有源耳机混合使用。这些工厂校准的耳机与特定耳机放大器单元形成系统，并且它们不能与设备中的第三方放大器或普通耳机输出一起使用。

[0066] 室内校准，版本1

[0067] 这是一种可以免于测量的耳机声音特性的室内校准方法。用户可以在听音室中迭代地设置该校准。参考图5的设置和图2和图3的方法，室内校准对有源监听耳机放大器2中的滤波器进行设置。连接到有源耳机放大器2的软件提供测试信号并显示校准期间测量过程的进度。这是通过连接到耳机放大器2的诸如PC或MAC 51的计算机中提供的用户界面完成的。测试信号被馈送到有源耳机放大器2，并且图形用户界面指导该过程。用户通过用户界面调整软件中的滤波器设置，实现有源监听耳机放大器2的设置，使得诸如测试信号的音量之类的声音属性与扬声器系统相同。将监听扬声器系统校准测试测量和均衡设置用作调整有源监听耳机声音属性的参考。参考测试信号可以包括基于存储或实时测量的一组不同设置。用户可以于任何时间在监视扬声器系统和耳机1之间切换，直到软件用户界面检测到变化太小或是随机的为止，这意味着没有发生系统改进，并且这终止了该进程。根据图2和图3，设置操作逐步通过音频带宽的不同子波带B1-Bn，实现整个音频通带的均衡。此过程将有源监听耳机放大器2的声音属性设置为频率响应与使用扬声器系统的监听室声音着色相类似。

[0068] 换句话说，耳机1的用户通过跨越不同频率范围的测试信号交替地收听扬声器和有源监听耳机。这意味着用带通滤波器对测试信号进行了滤波，使得音频频率范围被分成根据图2的几个子波带B1-Bn。用户通过几个子波带B1-Bn收听测试信号，调整每个子波带B1-Bn的耳机声音属性(例如声级)与具有相同频带的扬声器系统相同。该评估还可以通过使用包括扩音器的仿真头部的测量来进行，使得耳机1被戴在仿真头部上并从仿真头部取下，并且来自仿真头部中的扩音器的输出端是监视器。持续进行该过程，直到监视扬声器系统和有源耳机之间没有本质区别，然后软件将调整产生的设置作为其中一组预定设置存储到耳机放大器中。通常，子波带B1-Bn的带宽Δf是一个倍频程。由于声音属性也可以在子波带B1-Bn内使用频率调整，以致于要在子波带B1-Bn内对低频或高频进行加强。

[0069] 有利地，测试信号是包括如下这样信号的wav文件：

[0070] a.粉红噪声，换句话说，该信号的功率谱密度(每Hz的能量或功率)与该信号的频率成反比。在粉红噪声中，每个倍频程(频率减半/加倍)携带等量的噪声功率。

[0071] b.备选地，测试信号可以是类似音乐的信号的伪序列，该伪序列基本上包括频谱

遍布在宽频率区域上、通常基本上覆盖各子频带的频率范围的频率含量。

[0072] c.伪序列可以重复,这创建了用于调整的样本参考,并且重复之前的持续时间通常为1到10秒。

[0073] 关于用户界面,可以通过以下方式描述该校准过程:

[0074] • 免测量校准允许用户将声音校准为与其扬声器系统的声音在着色(相同的声音属性)上相同

[0075] • 该过程例如基于软件所生成的声音

[0076] • 校准过程按以下方式进行:

[0077] -计算机为每个子波带播放声音样本(可以是WAV文件)

[0078] -在软件控制下,此样本可在监听器或有源耳机中播放

[0079] -软件呈现了图形用户界面,用户可以通过监听系统输出将耳机中的级别调整为相似

[0080] -这是针对左、右(或环绕)系统共同完成的

[0081] -软件从一个子波带前进到下一个子波带,直到所有子波带都被覆盖

[0082] -用户评估输出结果并将校准保存到有源耳机放大器2存储器

[0083] 室内校准,版本2

[0084] 备选地,可以通过测量进行校准。这是一种基于测量的室内校准耳机声音特性的方法。可以在软件校准借助监听扬声器系统和扩音器测量过听音室之后,再设置这种类型的室内校准。这里,使用扩音器测量来确定听音室的脉冲响应。脉冲响应能够计算出室内频率响应。室内校准测量用于设置有源监听耳机放大器2中的滤波器。此方法将有源监听耳机放大器的输出信号属性设置为与测量的室内频率响应相匹配。该方法对室内频率响应的主要特征进行建模。用户可以选择建模精度的精确性。室内模型是:前30ms的FIR,和五个子波带的IIR(无限脉冲响应)混响模型的室内衰退剩余量。FIR(有限脉冲响应)适用于室内IR。子波带IIR适合于检测到的子波带中的衰减特性和速度。通常应用外在化滤波器。无需用户交互。关于外在化,以下过程是与本发明相关的一种选择:外在化滤波器实施为双耳滤波器,以使其成为全通滤波器。换句话说,此滤波器具有恒定幅度响应的滤波器(幅度/振幅不随频率函数而变化),但仅实现双耳滤波器的相位响应。在本申请中,恒定幅度/振幅是指:对于耳机应用而言与恒定振幅值的偏差优选地小于+/-3dB,或者优选地小于+/-0.1dB。

[0085] 这种类型的滤波器可以有利地实现为FIR滤波器,但理论上可以获得与IIR滤波器相同的结果。由于滤波器的高程度,IIR实施并不总是能实现的。通过这种方法,可以获得一些优点:如果用普通的双耳滤波器对幅度的反演进行建模,则可以容易地产生清晰的听觉着色。根据本发明,这可以通过全通实施方案来避免。此外,全通方案永远不会带来大的增益,因此对动态方面的要求极低。全通实施方案创建了经历过已测量过的空间的外在化配置。此外,全通实施方案对于HRTF滤波器的形式不像普通双耳滤波器那样敏感,由此也可以使用由第三人的头部进行的测量。因此,可以向用户提供与最近使用的收听空间相对应的默认外在化过滤器。

[0086] 可以例如以下列方式对扬声器执行该室内校准:

[0087] 经工厂校准的声学测量扩音器用于调整每个扬声器的声级和补偿距离差。合适的软件可以为每个扬声器提供针对测量响应、滤波器补偿和所得到系统响应的精确图形显

示,并完全手动控制声学设置。单点或多点扩音器位置可用于一人环境、两人环境或三人混合环境。

[0088] 从软件的角度来看,可以以下列方式呈现该校准:

[0089] • 校准将有源耳机1的声音设置为类似于用户先前测量的扬声器监听系统的声音

[0090] -校准过程如下:

[0091] -用户将有源耳机放大器2连接到运行合适软件(例如GLM)的计算机51

[0092] -用户选择现有的系统校准

[0093] -软件选择左右监听响应

[0094] -软件计算滤波器设置,以使有源耳机中的声音与监听扬声器中的声音类似

[0095] -包括早期反射、子波带衰减、声音着色和外在化滤波器设置

[0096] -用户可以聆听均衡结果,并将这些设置永久地保存在有源耳机放大器存储器中

[0097] 图4示出了能够支持本发明的至少一些实施例的示例装置。根据图4,耳机放大器2包括用于接收模拟音频信号的模拟输入端35。该信号由模数转换器36转换成数字形式并馈送到数字信号处理块37,在此之后数字信号被转换回模拟形式以馈送到功率放大器39和功率放大器40,功率放大器39和功率放大器40将放大的信号馈送到耳机1的驱动器。耳机放大器2还包括局部简化的用户界面34,用户界面34可以是具有彩色信号灯的开关或旋钮或小显示器。此外,耳机放大器2包括能够将电力输入到电源和电池管理系统32的USB连接器33,电源和电池管理系统32将电力进一步馈送到充电子系统31并从那里馈送到电池30,电池30用作耳机放大器2的电子器件的主电源。USB连接器33还用作数字信号处理块37的数字输入端。

[0098] 图5示出了能够支持本发明的至少一些实施例的示例软件系统。根据图5,该软件包括:用于AutoCal室内均衡器41的软件模块,用于处理室内校准;用于EarCal用户均衡器42的软件模块,用于为耳机1创建定制化均衡。工厂均衡模块43代表存储在耳机放大器2的存储器中的工厂均衡,其中,耳机的每个驱动器都相对于参考进行过工厂校准,使得离开工厂的每一耳机1-耳机放大器2组对都能产生具有基本相似的声音属性的音频信号。此外,软件包包括用于USB接口功能47、软件接口(GLM)功能48、存储器管理功能49以及电源和电池管理功能50的软件功能度。

[0099] 风格化耳机使用

[0100] 根据图6和图7,有源监听耳机1通过线缆3连接到耳机放大器2。放大器2通过线缆52连接到程序源51,56的线路输出端或监控输出端。程序源可以是包括计算机平台51的专业型或普通型便携式设备56。用户打开有源监听耳机放大器2并调整信号属性。

[0101] 根据本发明的一些实施例,例如图6,需要将耳机放大器2连接到计算机USB连接器并安装合适的(例如GLM)软件。用户在用户界面中导航到“耳机”页面。可用选项可以例如是:

[0102] • 音量控制,具有所有相关尺寸、预设项等。

[0103] • 个人平衡控制(以将声像图像设置在中间)

[0104] • 声音特性曲线调整

[0105] • 启动音量设定功能

[0106] • ISS控制功能(多久后休眠)

[0107] • 最大SPL限制功能(保护听力)开/关、限制调整

[0108] • EAI(增强型LF隔离)开/关功能以及隔离级别(反馈)的低/中/高控制

[0109] • 用于将这些设置永久存储到有源耳机放大器中的功能

[0110] 校准之间的切换

[0111] 当用户在有源耳机放大器中存储校准时,可以参考图6和7选择均衡。利用诸如音量控制之类的开关,可以按以下方式选择其中一个校准:向下推动音量控制器54(单击),然后转动音量控制器来选择均衡(未设置均衡或设置享乐均衡,均衡方法1,均衡方法2),然后释放音量控制器来选择均衡。

[0112] 本发明的一些实施例在以下基本系统质量中的益处为:包括了专用和单独均衡的耳机放大器2。工厂均衡消除了单元与单元间的音质差异。耳罩之间没有(随机变化的)单元与单元间的差异,始终保持平衡。与大多数其他耳机不同,音频再现始终是中性的。此外,声音隔离非常优异(在中/高频率由闭环式耳罩进行无源隔离,在低音频率通过校准提高隔离度)。室内均衡(方法1和方法2)允许模拟现有监听系统的声音特性;例如不在工作室时,通过耳机进行准确可靠的工作。电池容量和电子设计允许全天操作,而无需将放大器连接到电源。

[0113] 利用所描述的实施例,可以获得若干益处。在使用与耳机分开的放大器模块中的电子器件实现(手动)音量控制的方案中,不对电池(电源处理)或电子器件进行空间限制。在该解决方案中,可以使用所有需要的输入类型和连接。同样,不对所包含的信号处理进行限制。

[0114] 该解决方案可以从USB连接器供电。单独的放大和布线避免了驱动器之间的任何相互作用,例如,当在耳机线缆中共享导体时可能发生这种相互作用。在有源耳机中,信号处理可以是极度线性的。耳机中的每个耳/驱动器可以单独地被工厂均衡到参考值,因此每个驱动器可以呈现完全平坦和中性的响应。对于每只耳朵具有多路驱动器的情况,可以进行多路系统的交叉,以具有理想的性能。可以进行风格化校准。可以进行享乐校准(例如,优选的声音、响应曲线),以及将耳机校准为听起来与参考系统(例如,听音室)相同;这种校准可以自动化。

[0115] 耳机传递函数反演的自动正则化参数

[0116] 提出了一种用于自动调整用于耳机均衡的耳机传递函数的反演的方法。该方法通过比较半倍频程平滑化之前和之后的测量响应来估量正则化的量。因此,正则化完全取决于耳机响应。该方法将反演测量响应时的传统正则化反演法的精度与在陷波频率处使用平滑法的反演的感知鲁棒性相结合。进行主观评估以确认所提出的方法的效果,该方法用于获得主观可接受的、用于均衡双耳再现应用耳机的自动正则化。结果表明,与使用固定正则化因子的正则化反演方法或与半倍频程平滑窗一起使用的复杂平滑法相比,所提出的方法可以产生感知上更好的均衡。

[0117] 双耳合成使得耳机的音频呈现能够渲染出与听者在原始声场中感知到的听觉印象相同的听觉印象。为了将通过耳机呈现的虚拟源放置在特定方向上,源声音的消声记录与表示从目标源位置到听者耳朵的声学路径的滤波器进行卷积。已知这些过滤器为双耳响应。在消声呈现的情况下,已知这些响应为头部相关脉冲响应(HRIR)。在混响呈现的情况下,已知这些响应被称为双耳室内响应(BRIR)。双耳响应可以通过在听众的听觉通道、双耳

扩音器(仿真头部)的听觉通道进行测量来获得,或通过计算机模拟来获得。为了保持双耳响应的频谱特征,必须在通过耳机呈现音频时补偿耳机传递函数(HpTF)。这是通过将双耳响应与在相同位置处测量的耳机响应的反演进行卷积来完成的。当针对每个听众单独测量响应时,可以实现更好的结果。

[0118] 耳机传递函数通常包含峰值和陷波,这是由于在由耳机和收听者的耳朵限定的音量内产生的共振和散射。

[0119] 耳机的复频率响应的直接反演

[0120] $$H^{-1}(\omega)=\frac{1}{H(\omega)} \qquad (1)$$

[0121] 在测量响应具有陷波的频率处包含大峰值。在耳机传递函数测量中看到的峰值和陷波在个体之间会变化,并且当耳机被取下然后再针对相同的受试者再次佩戴时也可能改变。尽管如果受试者自己放置耳机会使得由于耳机的重新定位导致的耳机传递函数的可变性降低,但是使用耳机传递函数的直接反演来均衡耳机的过程可能导致声音的着色。此外,当陷波频率由于耳机的重新定位而移位并且均衡器的升高不再与实际响应中的陷波的频率和增益匹配时,通过应用深陷波的精确反演而产生的大峰值可被感知为共振振铃伪像。该效果在图8中示出,其中在重新定位之后测量的耳机的两个幅度响应已经使用在重新定位之前测量的响应的直接反演来补偿。在图8中所示的响应中看到的窄带共振是用于反演的响应和与重新定位耳机之后测量的响应中的陷波频率之间的失配的结果。通过对测量响应中的陷波进行反演来限制峰值增益,可以最小化这种不匹配的可听性。

[0122] 为了最小化陷波反演的听觉效果,通常采用感知激励式修改,以直接对测量响应进行反演。由于人类对相同幅度和Q因子下的峰值的感知要比陷波更好,因此进行反演时应当使得在对测量响应中的峰值进行反演的同时忽略陷波,或者在反演之前降低陷波幅度。用于在反演之前减小陷波幅度的方法包括平滑所测量的响应,对通过重新定位耳机所取得的若干响应进行平均,或者使用统计方法来近似整体响应。然而,这些方法可能会影响该响应的其余部分的反演精度。

[0123] 反演的正则化是一种在减少陷波反演的同时确保响应的精确反演的方法。正则化参数限定了尽量在特定频率处反演,这限制了响应中的陷波反演和噪声。必须选择正则化参数,使得其引起声音的最小主观劣化。然而,正则化参数的合适值取决于待反演的响应,因此必须使用收听测试为每个反演选择该值。

[0124] 在该工作中,提出了一种方法,用于在对用于双耳合成应用的耳机响应进行反演时自动获得频率相关的正则化参数。关于除了大陷波以外的响应反演精度和对抗耳机重新定位的均衡的稳定性,将所提出的正则化的性能与传统的正则化反演、维纳(Wiener)反卷积和复杂平滑法进行比较。使用个性化双耳室内响应来执行主观评估,以确认所提出的正则化的主观表现。

[0125] 应用于耳机均衡的正则化反演

[0126] 可以在反演过程中引入频率相关的正则化因子,以限制在陷波的反演中的作用。正则化因子由滤波器B(ω)组成,滤波器B(ω)由比例因子β缩放。然后将响应H(ω)的正则化反演$H_{RI}^{-1}(\omega)$表示为:

[0127] $$H_{RI}^{-1}(\omega)=\frac{H^{*}(\omega)}{|H(\omega)|^{2}+\beta|B(\omega)|^{2}}D(\omega)\quad,\qquad(2)$$

[0128] 其中，*表示复共轭，|·|是绝对值运算符，以及D(ω)是引入的延迟滤波器，用于产生因果反演$H_{RI}^{-1}(\omega)$。

[0129] 当$|H(\omega)|^{2}>>\beta|B(\omega)|^{2}$时，反演是精确的，而当$\beta|B(\omega)|^{2}\geq|H(\omega)|^{2}$时，反演的作用是有限的。正则化的效果可以在图9中看到，其中除了直接反演(虚线)中呈现的大共振以外，β＝0.01和B(ω)＝1的正则化反演(实线)产生了耳机响应的精确反演。进一步地，由于该方法避免了在幅度小于正则化因子的频率处的反演，因此耳机的有用带宽之外的频率不进行反演，例如所示的低于30Hz的频率那样。

[0130] 通常选择参数β和B(ω)以获得最小的声音质量劣化，并同时精确地反演除狭窄的陷波之外的响应。通常，B(ω)是基于使得反演所需的带宽能被评估成主观可接受质量来定义的，例如响应的第三倍频程平滑版本的反演，或者使用高通滤波器。然后，使用收听测试来调整β，以便缩放B(ω)以使音质的降底最小。在S. G. Norcross, G. A. Soulodre和 M. C. Lavoie在音频工程学会2004年期刊，第52卷，第10期，第1003-1028页发表的“Subjective investigations of inverse filtering”中，使用三个不同的B(ω)滤波器(平坦响应、具有截止频率为80 Hz和18 kHz的带阻滤波器)评估扬声器响应的正则化反演，并且对第三倍频程平滑响应进行反演。然后，针对每个B(ω)测试不同的β值。S. G. Norcross, G. A.Soulodre和 M. C. Lavoie在音频工程学会2004年期刊，第52卷，第10期，第1003-1028页发表的“Subjective investigations of inverse filtering”的结果表明β的校正值取决于待反演的响应和选择用于正则化的滤波B(ω)。进一步地，针对用于双耳再现的耳机响应的反演的不同方法的性能研究表明，专家听众对β的调整也会依赖于B(ω)而产生不同的结果。在他们的实验中，B(ω)被限定为耳机响应的倍频程平滑响应的反演，或者限定为具有截止频率为8kHz的高通滤波。然而，使用由专家听众调整的正则化的正则化反演而获得的耳机均衡在感知上要比使用复杂平滑法所得到的反演来获得的耳机均衡更加可接受。因此，尽管可以先验地选择B(ω)，但是应该根据待反演的响应、H(ω)以及正则化滤波B(ω)来调整β。

[0131] 关于维纳反卷积

[0132] 如果噪声功率谱$|N(\omega)|^{2}$是已知的，则在公式(2)中的项$\beta|B(\omega)|^{2}$可以估算为信噪比(SNR)的反演，

[0133] $$\mathrm{SNR}(\omega)=\frac{|H(\omega)|^{2}}{|N(\omega)|^{2}}\quad.\qquad(3)$$

[0134] 这产生了维纳反卷积，其提供关于SNR的最佳反演带宽。维纳反卷积滤波器$H_{WI}^{-1}(\omega)$可以得出自：

[0135] $$H_{WI}^{-1}(\omega)=\frac{H^{*}(\omega)}{|H(\omega)|^{2}+\frac{|N(\omega)|^{2}}{|H(\omega)|^{2}}}D(\omega)\quad.\qquad(4)$$

[0136] 对于大的SNR，维纳反卷积等效于直接反演，但具有最佳的反演带宽，因为只有具有大SNR的带宽才能被准确地反演。这在图9中示出，其中显示了使用维纳反卷积(阴影线)计算的反演耳机响应。尽管该方法提供了最佳的反演带宽，但是陷波被精确地反演，这是以与直接反演(虚线)类似的方式产生了大的共振，从而产生了振铃伪像。为了避免反演响应中的大共振，可以应用比例因子，使得维纳反卷积等效于正则化反演方法(参见公式2)。

[0137] 推荐的正则化

[0138] $\beta|B(\omega)|^2$项可以定义为频率相关参数$\hat{\beta}(\omega)$，以使得可以精确地对响应进行反演，但是对于窄的陷波和在耳机的再现带宽之外的频率，不需要反演效果。参数$\hat{\beta}(\omega)$可以结合耳机再现带宽的估计值$\alpha(\omega)$和该带宽内所需的正则化的估计值$\sigma(\omega)$来确定。

[0139] 然后，将参数$\hat{\beta}(\omega)$定义为：

[0140] $$\hat{\beta}(\omega) = \alpha(\omega) + \sigma^2(\omega) \quad . \tag{5}$$

[0141] 参数$\alpha(\omega)$确定反演的带宽，反演的带宽被定义为使得$\alpha(\omega)$接近或等于零的频率范围。新的正则化因子$\sigma(\omega)$控制由$\alpha(\omega)$定义的带宽内的反演效果。

[0142] 如果耳机带宽已知，则可以使用整体增益滤波器$W(\omega)$来将$\alpha(\omega)$定义为

[0143] $$\alpha(\omega) = \left(\frac{1}{|W(\omega)|^2} - 1\right) \quad , \tag{6}$$

[0144] $W(\omega)$的平坦通带对应于耳机的再现带宽，对于高质量耳机通常为20Hz至20kHz。

[0145] 以类似的方式，如果可获得噪声功率谱估计值，则$\alpha(\omega)$可以定义为

[0146] $$\alpha(\omega) = \frac{1}{\mathrm{SNR}(\omega)} = \frac{|N(\omega)|^2}{|H(\omega)|^2} \quad . \tag{7}$$

[0147] 为了避免响应中相邻频率区间之间的强烈变化，应使用噪声包络$N(\omega)$的估计值，例如平滑光谱。

[0148] 新的正则化因子$\sigma(\omega)$被定义为测量响应$H(\omega)$与经减小陷波幅度的响应$\hat{H}(\omega)$的负偏差。例如，$\hat{H}(\omega)$可以使用平滑版本的耳机响应来定义。基于此，$\sigma(\omega)$可以定义为

[0149] $$\sigma(\omega) = \begin{cases} |H(\omega)| - |\hat{H}(\omega)| \text{ , if } |\hat{H}(\omega)| \geq |H(\omega)| \\ 0 \qquad\qquad \text{ , if } |\hat{H}(\omega)| < |H(\omega)| \end{cases} \quad . \tag{8}$$

[0150] 由于对于$|\hat{H}(\omega)| > |H(\omega)|$，$\sigma^2(\omega) > 0$，参数$\hat{\beta}(\omega)$在比平滑窗窄的陷波频率处包含大的正则化值。作为示例，在图9中使用耳机响应所得到的$\hat{\beta}(\omega)$在图10中示出。为了得到$\hat{\beta}(\omega)$，使用公式6确定参数$\alpha(\omega)$，其中，$W(\omega)$选择为使得带宽限制在20Hz和20kHz之间(实线)。此外，$\alpha(\omega)$也使用公式7(虚线)来确定，其中，根据测量的耳机脉冲响应的尾部估算$N(\omega)$。在这两种情况下，$\hat{H}(\omega)$是半倍频程平滑版本的耳机响应。最高正则化值与图9中所示的直接反演中的共振频率一致。正则化参数$\hat{\beta}(\omega)$对于响应的其余部分保持接近或等于零，

这保证了准确的反演。由α(ω)引起的带宽限制可以在低于20Hz和高于20kHz的频率处看出，其中，$\hat{\beta}(\omega)$包括较大值。当α(ω)使用公式7(虚线)进行限定时，反演带宽更偏向于略微向低频扩展，并且在高频时不受限制，其中使用公式6时，如前所述，反演带宽限制在20Hz和20kHz之间。对于20Hz和20kHz之间的频率，$\hat{\beta}(\omega)$对于这两种方法都类似，确认使用任一种方法来确定α(ω)都会产生类似的结果。

[0151] 将公式5应用到公式2产生了所提出的对常规正则化反演方程的修改，即sigma反演$H_{SI}^{-1}(\omega)$

[0152] $$H_{SI}^{-1}(\omega)=\frac{H^*(\omega)}{|H(\omega)|^2+\hat{\beta}(\omega)}D(\omega)=\frac{H^*(\omega)}{|H(\omega)|^2+[\alpha(\omega)+\sigma^2(\omega)]}D(\omega) \quad . \tag{9}$$

[0153] 将在图11中所提出的sigma反演法与图9中使用的耳机响应的直接反演进行比较。用于渲染$H_{SI}^{-1}(\omega)$的参数$\hat{\beta}(\omega)$为图10中如实线所呈现出的。由耳机响应中的陷波的精确反演产生的共振不存在于由所提出的方法(实线)产生的反演中。此外，不对所定义带宽之外的频率进行补偿，并且对响应的其他部分被准确地反演。

[0154] 装置和方法

[0155] 本部分描述了在评估所提出方法的性能时测量装置和所执行的信号处理。还解释了评估测量和收听测试的设计。

[0156] 测量装置

[0157] 测量装置包括两个微型扩音器(FG-23329，∅=2.59mm，Knowles牌)，该微型扩音器设置在人类受试者的开放式听道内并连接到音频接口(UltraLite Hybrid3，MOTU牌)。响应以48kHz采样率数字化。将扩音器设置在开放式听道内以避免耳机负载对双耳滤波器的影响。微型扩音器引入耳道内而不到达鼓膜但足够深，以便在将引线绕着耳朵弯曲时微型扩音器仍保持在原位(如图12a所示)。考虑到要确保当将耳机放在耳朵上时扩音器不移动，通过在如图12b所示的两个位置处用胶带固定引线。

[0158] 归一化

[0159] 使用比例因子g将测量的耳机响应H(ω)归一化为单位能量先验反演，使得

[0160] $$\frac{1}{2\pi}\int_{-\pi}^{\pi}|gH(\omega)|^2d\omega=1 \quad . \tag{10}$$

[0161] 如图9和图11所示，这允许使反演居中在0dB的等级处，当待反演的响应的幅度非常小时，这避免了在反演带宽之外的频率处的反演响应的不连续性。在反演之后，可以针对该比例因子补偿响应，以恢复原始信号增益。进一步地，该归一化能够将正则化定义为动态限值，例如，如果在反演带宽内B(w)＝1，则β＝0.01＝-20dB。因此，如图2所示，归一化响应的反演不会产生大于如图9所示的|β|-6dB的放大率，其中β＝0.01＝-20dB的常规正则化反演不会超过14dB的放大率。

[0162] 反演滤波器

[0163] 使用公式9通过改变α(ω)和$\sigma^2$((ω)的值来获得针对不同方法的反演滤波器。在图13中示出了通过维纳反卷积法、常规正则化反演法、复杂平滑法以及所提出的sigma反演正则化法来获得反演响应的参数值。为了确保本工作中使用的所有方法具有相同的带宽，

使用公式6限定α(ω),其中W(ω)在20Hz和20kHz之间具有恒定单位增益。维纳反卷积使用公式7,但所产生的带宽与其他方法没有很大差别。正则化比例因子β通过使用收听测试进行的调整来进行选择。半倍频程平滑法与复杂平滑法和sigma反演方法一起使用,以在方法之间呈现直接的对比。该平滑窗基于非正式的收听测试进行选择。与倍频程、三分之一倍频程以及ERB平滑窗相比,半倍频程平滑产生了最小的声音劣化。

[0164] 平滑响应HSM(ω)实施在使用起始于ω1处结束于ω2处的半倍频程方窗$W_{SM,_}$的频域内,以分别地平滑幅度

[0165] $$|H_{SM}(\omega)| = \frac{1}{\omega_2 - \omega_1}\int_{\omega_1}^{\omega_2} W_{SM}|H(\omega)|d\omega \quad , \tag{11}$$

[0166] 和解缠相位

[0167] $$\angle H_{SM}(\omega) = \frac{1}{\omega_2 - \omega_1}\int_{\omega_1}^{\omega_2} W_{SM}\angle H(\omega)d\omega \quad . \tag{12}$$

[0168] 平滑后响应获得为

[0169] $$H_{SM}(\omega) = |H_{SM}(\omega)|e^{j\angle H_{SM}(\omega)} \quad , \tag{13}$$

[0170] 并且,然后通过公式9计算反演$H_{SM}^{-1}(\omega)$

[0171] 性能评估测量

[0172] 由单个受试者佩戴的耳机(HD600,Sennheiser牌,德国)被测量四次,在每次测量之后重新定位耳机。为了重新定位耳机,受试者在测量之间移除然后重新应用耳机,以减少测量响应的可变性。将测量响应归一化在幅度为0dB等级周围。得到的响应如图14所示,以允许在响应之间进行比较。第一耳机响应(实线)用于反演,并且其也用于获得图9和图11所示的反演响应。选择一个特定的受试者,该选择是根据他的个人均衡滤波器在反演时产生振铃伪影的早期非正式测量结果得出的。假设9.5kHz处的陷波的精确反演是造成伪像的原因。基于由该受试者进行的调整测试,选择β=-20dB的值用于常规的正则化反演法。在图13中给出了每种方法的参数。

[0173] 针对主观评价的收听测试设计

[0174] 进行一组测量以主观评估所提出的方法。针对每个测试参与者,测量在ITU-R BS.1116兼容室内的耳机响应(SR-307,Stax牌,日本)和立体声扬声器装置(8260A,Genelec牌,芬兰)的个体双耳室内响应。所测量的耳机响应在反演之前进行了归一化,并且在反演之后对增益因子进行了补偿。这允许耳机上的再现等级与扬声器上的再现声级相匹配。

[0175] 设计收听测试以在感知上评估所提出的方法的性能。该测试的范例是相对于立体声扬声器装置的耳机来评估双耳合成演示的保真度。目的在于当重新定位耳机时与扬声器演示相比较地来评估整体音质。受试者的任务是取下耳机,然后收听扬声器,最后再戴上耳机以收听双耳再现。这导致在测试期间重新定位的效果。工作假设是:所提出的方法在统计上表现得与常规正则化反演和平滑法的最佳情况同样好或更优。这证实了所提出方法的适用性。

[0176] 所使用的测试信号是在2kHz处具有的截止频率的高通粉红噪声、宽带粉红噪声以及两种不同的音乐样本。测试信号具有宽带频含量。因此,可以检测到高频伪影和着色。噪

声信号由两个不相关的粉红噪轨道组成，每个扬声器对应一个噪声轨道。音乐信号是可以循环无缝再现的摇滚和疯克音乐的短立体声轨道。为了获得测试样本，将测试信号与使用正则化反演法、平滑法和所提出的sigma反演法获得的双耳滤波器进行卷积。通过非正式测试(其中三个听众对用不同正则化β值获得的声音质量进行分级)选择出常规正则化反演的比例因子β＝-18dB。没有耳机均衡的双耳滤波器用作低锚点。由于听道内的扩音器的响应和耳机响应相互不均衡，因此预期这些未经补偿的滤波器会扭曲声音的音色和空间特性。

[0177] 十名受试者参加了测试。他们在类似的需要区分音色和空间失真的测试中有经验。要求受试者使用0到100的标度对音频样本的耳机呈现的保真度进行分级。扬声器上的再现用作参考。要求受试者仅在他们没有感觉到任何差异并因此不能区分声音是来自扬声器还是耳机时给出最大评分。如果耳机再现不能再现扬声器演示的任何特征，则给出最低分。将待评估的这些特征向受试者描述为音色、空间特性和伪像的存在。然而，受试者可以自由地以不同方式对每个特征进行加权，例如，不同于音色，空间再现的微小差异可以更显着地被分级。测试样本以连续循环方式再现，受试者可以自由选择是收听扬声器还是耳机再现。图形界面允许受试者在四个双耳滤波器和扬声器再现之间进行选择。对于每个测试信号，双耳滤波器是随机排序的，并且允许在滤波器之间的对比。

[0178] 结果

[0179] 性能的评估

[0180] 通过与维纳反卷积法、常规正则化反演法和复杂平滑法的比较，来评估所提出的正则化的适用性。比较的标准是除了那些由于重新定位可能产生伪影的陷波处的响应之外的响应的反演精度。选择维纳反卷积法和常规正则化反演法进行比较，是因为它们具有与所提方法类似的公式，仅在所使用的正则化参数方面不同(参考如上的“应用于耳机均衡的正则化反演”)。维纳反卷积也表示具有最佳带宽限制的直接反演。选择平滑法进行比较，是因为在所提出的方法中也使用幅度平滑来评估正则化参数$\sigma^2(\omega)$(参见公式8)。

[0181] 在图14中作为实线示出的耳机响被用于使用前述方法来获得反演滤波器。将原始响应与不同的反演滤波器进行卷积的结果如图15所示。曲线显示了能显现出差异的2kHz到20kHz之间的数据。维纳反卷积法(虚线)产生平坦的响应，精确地对陷波进行反演。平滑法(阴影线)在陷波频率之间产生5dB的共振，其中预期反演是准确的。常规的正则化反演法(点划线)产生比平滑法更平坦的响应，同时在陷波频率处保持类似的衰减。所提出的方法(实线)产生补偿响应，在陷波频率处具有最大衰减，但在陷波之间仍然提供平坦响应。陷波频率处的强衰减表明，当将该反演滤波器应用于在重新定位耳机之后测量的耳机响应时，陷波频率的小偏移可能不会导致共振。在图16中可以看到这种效果的示例，给出了将先前获得的反演滤波器与重新定位后测量的三个响应进行卷积的结果。这些耳机重新定位后的响应在图14中以虚线、点划线以及阴影线示出。对于所有的方法，在16kHz以上，通过第三次测量获得的响应的均衡相对于原始耳机响应相差高达10dB。然而，如果再现宽带声音，预计不会对判断产生很大影响。因此，评估针对低于16kHz的频率进行。尽管图14中的耳机响应没有很大差异，但是图16中使用维纳反卷积(顶部的框)的均衡耳机响应包含了可以被视为振铃伪像的共振。这些共振在其他方法中没有经历，但是常规的正则化反演法(从顶部开始的第二个框)、平滑法(从顶部开始的第三个框)和提出的方法(底部框)之间在这些频率处存在一些差异。所提出的方法对于所有响应而言都在陷波频率(9.5kHz和15kHz)下产生稳

定的大衰减。其他方法则不是这种情况。它们的衰减随重新定位而变化。此外，所提出的方法仍然保持类似于常规正则化反演法的平坦总体响应。这些结果表明所提出的方法在保持最小的声音劣化的同时可以增加一定的对抗重新定位效果的鲁棒性。但是，这应该通过听力测试来评估。

[0182] 主观评估

[0183] 在参与测试的10个受试者中所评估的样品平均值(μ)和标准偏差(SD)在图17中示出。为了评估给予每种方法的评分平均值之间的差异的统计学显着性，进行单向ANOVA检验。使用Levene's检验($F(3,156)=14.05$，$p<0.001$)测试方差齐性，这导致违反了齐性假设。因此，使用具有alpha＝0.05的Welch's测试来代替常规的单向ANOVA检验。Welch的测试报告了针对给予不同方法的评分的至少一个平均值的统计学显着性差异($F(3,79.48)=145.48$，$p<0.001$)。给定评分与反演方法之间关联强度的度量($\omega^2=0.73$)表明评分中73％的方差可归因于反演方法。由于违反了方差齐性，因此使用Games-Howell的事后检验来确定哪些方法的平均评分在统计上有所不同。检验的结果在图18中示出。除了由常规正则化反演法(μ＝79.8，SD＝14.33)和平滑法(μ＝69.92，SD＝25.7，其中零假设不能被排除(p＝0.139))形成的配对之外，所有方法都显示评分平均值之间存在统计学上的显着差异。

[0184] 在图19中描绘了平均值和它们的95％置信区间。常规正则化反演法的评分平均值和置信区间优于平滑法的评分均值和置信区间，尽管平均值的差异在统计上不显着但这表明在感知上优越的性能。这与Z.Schärer and A.Lindau在2009年5月的音频工程学会会议126中发表的“Evaluation of equalization methods for binaural signals”(其中β由专家听众选择)中的结果一致。基于此，当前测试中使用的β值可能被认为与专家获得的值一致，因此可用于评估所提出的方法的性能。所提出的方法呈现出最大的质量评分平均值，表明所提出的方法比其他方法引起更小的声音劣化。进一步地，所提出的方法的平均值的置信区间很窄，表明受试者对该方法的评分一致。这些结果证实了这一假设：即所提出的方法在统计学上比在该测试中使用的其他方法更好。

[0185] 讨论和结束语

[0186] 由于对原始测量的耳机响应的陷波的反演，最佳正则化因子产生耳机响应的主观上可接受且精确的反演，同时仍然使声音质量的主观劣化最小化。

[0187] 因为预期到一定的频率依赖性，所以为了最佳主观接受性而单独调整正则化因子的做法是繁琐且耗时的。用于定义反演耳机响应的正则化因子的方法是基于缩放预定义的正则化滤波器。首先设计正则化滤波器以限制反演的带宽，然后将固定比例因子调整到可接受的值。由于正则化因子取决于待反演的响应，因此固定的比例因子可能导致某些陷波过度正则化而其它陷波正则化不充分，这会降低声音质量。

[0188] 所提出的方法通过使用耳机响应本身对频率相关的正则化因子进行估计，来自动生成该频率相关的正则化因子。测量的耳机响应与其平滑版本之间的比较提供了对每个频率所需的正则化的估计。这种正则化在陷波频率处很大，并且当原始响应和平滑后响应相似时接近于零。可以使用SNR的估计或再现带宽的先验知识，根据测量响应来定义反演的带宽。因此，可以单独且自动地获得正则化因子。

[0189] 用于估计正则化量的平滑窗口应该导致对音质的最小劣化。窄平滑窗口可以产生更准确的耳机响应反演，因为平滑后的响应更类似于原始数据。然而，由于在原始测量中在

陷波周围的频率处的反演引入的过度放大，使得这可能导致刺耳的声音质量。发现耳机响应的半个倍频程平滑可以充分估计所需的正则化量，但是使用不同方法获得的其他平滑后响应也可能是合适的，就例如B.Masiero和J.Fels在2011年5月的音频工程学会会议130中提到的“Perceptually robust headphone equalization for binaural reproduction”。此外，对于不同于该工作中分析目的的某些目的而言，不同的平滑窗口可能是更优的。

[0190] 对所提出的方法的评估表明，它提供了一种反演滤波器，其能够保持用于反演测量响应的传统正则化反演法的精度，同时以保守的、主观上可接受的方式限制陷波的反演。与常规正则化反演中使用的固定正则化相比，该正则化更强并且在原始响应的陷波周围跨越更宽的频率范围。这导致有效的正则化，尽管重新定位耳机时通常出现陷波频率小位移，但造成的主观效果却更小，因此表明对耳机重新定位具有更好的鲁棒性。基于主观测试，由所提出的方法引起的较大正则化似乎不会降低感知的声音质量。

[0191] 传统正则化反演方法的正则化因子的调整基于仅由三个受试者进行的主观测试。对所有十个受试者应用这种单一正则化对于其中一些人可能不是最佳的。然而，正则化反演法获得了良好的分数(μ＝79.8,SD＝14.33)并且通常比复杂平滑法(μ＝69.9,SD＝25.7)的等级更好，这与先前的研究一致。这表明为常规正则化反演法选择的正则化因子可以在主观实验中用作验证所提出的方法的效果的参考。

[0192] 受试者的数量足以观察所提出的方法相对于传统的正则化反演方法的性能。关联度量($\omega^2$＝0.73)表明主观得分主要受反演方法的影响，并且事后检验表明，该方法与传统的正则化反演法存在显著差异(p＝0.002)。因此，通过所提出的方法获得的分数不是偶然的。通过所提出的方法获得的平均分数(μ＝89.62,SD＝8.04)证实了该实验中的研究假设。该假设是：所提出的耳机响应反演的正则化在感知上优于使用固定值正则化参数，并且结果对于耳机重新定位具有主观上的鲁棒性。

[0193] 较小的标准偏差以及评估得分的较窄置信区间表明受试者认同所提出的方法产生的感知声音质量。在测试期间重新定位耳机的效果似乎对所提出方法的得分影响更少于参考方法的得分。

[0194] 所提出的方法代表了对传统正则化反演的改进。所提出的方法的一个重要好处的是正则化是频率特定的，它导致最小的声音质量劣化，并且它完全基于测量的耳机响应数据来自动设置。

[0195] 所提出的方法避免了单独调整每个受试者的正则化因子所需的时间，从而允许更快和更准确的耳机均衡。该方法在主观测试中呈现的保真度表明该方法可以作为进一步研究耳机双耳合成的参考方法，或者，如听力测试设计所表明的那样，可以通过耳机来模拟扬声器装置，并同时保持原始扬声器室内系统的音色特征。

[0196] 使用均衡后的双耳响应的耳机立体声增强功能来保持耳机音质

[0197] 为了保持耳机的声音质量，描述并评估用于均衡双耳立体声渲染网络的输出的标准。目的在于对双耳滤波器进行均衡，以使得从扬声器到每只耳朵的直接路径和串扰路径的总和具有平坦的幅度响应。使用收听测试来评估该均衡标准，其中使用了几个双耳滤波器设计。结果表明，保持双耳滤波器的直接路径和串扰路径之间的差异对于保持双耳渲染的空间质量是必要的，并且双耳滤波器的后均衡可以保持耳机的原始声音质量。此外，发现测量的双耳响应的后均衡更好地满足测试参与者对从扬声器进行立体声再现的虚拟呈现

的期望。

[0198] 介绍

[0199] 由于便携性和对周围环境的隔离性，耳机通常用于便携式设备的立体声收听。耳机的音质主要受其频率响应的影响，并且一些研究已经提出了用于设计高音质耳机的不同目标功能。由此带来的耳机设计可以在立体声声音再现中提供出色的音质。然而，已知通过耳机再现立体声信号会在耳朵之间产生听觉图像(侧向化)并产生疲劳。这是由耳机产生的双耳提示与通过扬声器的立体声再现产生的双耳提示的差异引起的。用于耳机再现的立体声增强方法可以通过滤波的方式人工地引入类似于扬声器产生的双耳提示。立体声扬声器装置的双耳渲染在图20中示出。从扬声器到耳朵的双耳响应由滤波器$H_{ij}(\omega)$表示(大写下标“L”和“R”表示左右扬声器，小写“l”和“r”分别表示左耳和右耳)。在将立体声音频信号与这些滤波器进行卷积之后，在通过耳机收听时会再现类似于由一对扬声器产生的听觉图像。

[0200] 由于耳间时间和水平差异(分别为ITD和ILD)是水平面中定位的主要提示，因此可以使用模拟立体声扬声器系统的ITD和ILD的滤波器来减小侧向化效果。此外，通过使用头部相关传递函数、HRTF或双耳室内响应、BRIR(它们更准确地近似于听者的真实ITD，ILD和单声道响应)来改善耳机上立体声再现的空间特性。

[0201] 虽然双耳渲染已经广泛用于听觉定位研究，但是声音质量评估测试表明，听众更喜欢通过耳机再现立体声信号，而无需增强方法。这可能是由于非个性化双耳滤波器在声音中引起的光谱着色。为了使用双耳滤波器产生更“自然”的声音，已经提出了HRTF的均衡。还研究了使用专家听众来设计双耳滤波器的后均衡以便将双耳声音质量与扬声器声音质量相匹配。但是，很少有关于在使用双耳渲染时保留原始耳机音质的研究。

[0202] 在增强听觉图像的空间特性的同时保持耳机的原始音质激发了本发明的。在本发明中，双耳滤波器被设计成使得双耳室内响应的相位信息被保留的同时，以不同的方式均衡幅度信息。这些双耳滤波器的设计目的是增强空间立体声图像，同时最小化耳机声音质量的劣化。如Kirkeby,O.发表于2002年音频工程学会会议第22届国际会议“虚拟、综合及娱乐音频”中的“A Balanced Stereo Widening Network for Headphones”所说，保持双耳立体声网络输出的平坦幅度响应以便在两个声道中获得相等的信号幅度被采用作为保持耳机声音质量的标准。通过听力测试评估滤波器，其中针对空间质量、音色/声音平衡质量和整体立体声呈现质量分别进行测试。

[0203] 首先，给出了在双耳立体渲染中保持耳机音质的标准。其次，描述了测量方法、过滤方法和用于评估的听力测试的设计。随后，提出并讨论了听力测试的结果。接下来，将阐述结束语。

[0204] 立体声双耳渲染中保持耳机音质的标准

[0205] 在立体声混合中，通过在两个声道之间均等地分配信号，将幻象单声道声源放置在听觉图像的中心。当应用双耳渲染来用耳机模拟扬声器立体声再现时，每个立体声声道总是由一对滤波器处理，这对滤波器代表从头部同一侧的扬声器到耳朵的直接路径$H_d$，以及来自头部相对侧的扬声器的串扰路径$H_x$。滤波器$H_d$等效于$H_{LI}$和$H_{Rr}$，而$H_{x_}$等效于图20中的$H_{Lr}$和$H_{Rl_}$。放置在中心的幻象源在耳机上的双耳立体声再现如图21所示，其中s是音频信号，s’是双耳滤波过程后产生的信号，$H_{HP_}$是耳机的传递函数，$s'_{HP}$是传输到耳朵的声学信号。图

22示出了在没有双耳处理的情况下通过耳机对相同信号s的再现，其中$s_{HP}$是所得到的传输到耳朵的声学信号。我们假设从每个扬声器到耳朵的路径之间存在对称性，因此图21中呈现的网络对于两只耳朵是相似的。

[0206] 图23中示出完全向左平移的幻象源的双耳立体声再现。在这种情况下，音频信号包含在立体声信号的左声道$s_L$中，而右声道不包含任何信号。由于假定了对称性，反向布置则将幻象源完全平移到右侧。

[0207] 与图21中的网络相反，在大脑内完成信号的加和。这被称为双耳总和。术语“双耳总和”应理解为信号(仅呈现给一只耳朵的信号)的单声道再现与信号(呈现给双耳的信号)的双色再现之间的感知响度的感知增量。已经发现响度的增量取决于再现水平。然而，我们在此假设，针对单耳呈现而言，双耳呈现产生6dB的增益，因为双耳呈现近似于中等水平的感知增益。这相当于两个相等相关信号的总和。由于假设两个耳朵的滤波器$H_{x_}$是相同的，所以图23中的网络变得等同于图21。这证明使用图21中的系统来获得保持耳机原始声音质量的均衡是合理的。

[0208] 为了保持耳机的音质，当耳机由中心幻象源的立体声信号直接驱动时(参见图21)双耳网络的输出s’应当接近耳机的输入。然而，导致s’＝s的过滤器$H_{EQ_}$将删除针对空间化所进行的所有双耳处理。如果声音质量是根据幅度响应来定义的，则可以将滤波器$H_{EQ_}$定义为能产生幅度响应近似于s的幅度响应的信号s”。这意味着$H_{EQ_}$应该平坦化双耳网络输出的幅度。该滤波器可以设计为具有如下计算的幅度响应的线性滤波器

[0209] $$|H_{\mathrm{EQ}}| = \frac{1}{|H_{\mathrm{d}} + H_{\mathrm{x}}|} \approx \frac{1}{|H_{\mathrm{SM}}|} \quad . \tag{14}$$

[0210] 由于$H_{d_}$和$H_{x_}$可能包含室内效果，因此可能需要平滑版本的反演$|H_{d_}+H_x|$和$|H_{SM}|$。本发明使用了一个倍频程宽的平滑窗口。图24示出了用于保持耳机音质的双耳立体声再现网络。

[0211] 方法

[0212] 为了评估用于保持耳机音质的双耳立体声网络，设计了三个双耳滤波器并进行了听力测试。双耳室内响应用于添加反射，以改善过滤器所产生的外在化效果。

[0213] 测量和过滤器设计

[0214] 对于试听室内的具有340ms混响时间的立体声扬声器装置(Genelec 8260A)，测量了仿真头(Cortex Mk II)的双耳时间响应$h_{ij}(t)$。使用测量过的响应，将一组双耳过滤器$H_{bin}$是通过对如下响应的前42ms(2048样本，48kHz的采样率)进行窗口化来设计的，

[0215] $$H_{\mathrm{bin}} = \mathscr{F}\left\{h_{\mathrm{ij}}(t)\,w(t)\right\}, i \in \{L,R\}, j \in \{l,r\}, \tag{15}$$

[0216] 其中$\mathscr{F}\{\cdot\}$表示傅里叶变换，w(t)是42ms时长的时间窗口。在进行非正式听力测试后，采用这种滤波器长度作为外在化能力和由于室内混响引起的音色效果之间的最佳折中。

[0217] 然后，应用上述过程以获得一组均衡后的双耳滤波器$H_{binEQ}$。首先，使用双耳的双耳网络来获得平均滤波器$H_{SM_}$

[0218] $$|H_{SM}| = \frac{|\widehat{H_{RI} + H_{LI}}| + |\widehat{H_{Rr} + H_{Lr}}|}{2} \quad , \qquad (16)$$

[0219] 其中^表示在直接滤波器和串扰滤波器加和之后的一个倍频程平滑处理。滤波器$H_{EQ_}$的幅度作为50Hz和20KHz之间的$|H_{SM}|_$的反演。然后，双耳滤波器$H_{bin}$与$H_{EQ_}$进行卷积，以获得均衡后的双耳滤波器$H_{binEQ}$

[0220] $H_{binEQ}=H_{bin}H_{EQ}$ • (17)

[0221] 还进行了对双耳滤波器的进一步修改以去除单声道提示。仅仅通过保留双耳滤波器的相位信息，来生成全通版本的$H_{bin_}$。这会保留过滤器中的时态信息，但会删除ILD和单声道提示。然后，通过将根据直接路径和串扰路径的平滑响应的幅度比得到的结果幅度求平均，来估计直接路径和串扰路径之间的水平差异$H_{LD}$，

[0222] $$H_{LD} = \frac{\left(\frac{\hat{H}_{RI}}{\hat{H}_{LI}} + \frac{\hat{H}_{Lr}}{\hat{H}_{Rr}}\right)}{2} \quad , \qquad (18)$$

[0223] 其中^表示一个倍频程平滑的滤波器幅度响应。在此之后，直接滤波器和串扰滤波器的幅度分别设计成$H_{d_{ph}}$和$H_{x_{ph}}$

[0224] $$|H_{d_{ph}}| = \frac{1}{H_{LD}+1} \quad , \quad |H_{x_{ph}}| = \frac{H_{LD}}{H_{LD}+1} \qquad (19)$$

[0225] 由$H_{d_{ph}}$(实线)和$H_{x_{ph}}$(虚线)引入的频率相关增益如图25所示。将双耳全通滤波器与其对应的$H_{d_{ph}}$和$H_{x_{ph}}$滤波器进行卷积以生成双耳滤波器$H_{ph}$

[0226] $$H_{ph} = \begin{cases} \arg\{H_{LI}\} \times H_{d_{ph}} \\ \arg\{H_{RI}\} \times H_{x_{ph}} \\ \arg\{H_{Lr}\} \times H_{x_{ph}} \\ \arg\{H_{Rr}\} \times H_{d_{ph}} \end{cases} , \qquad (20)$$

[0227] 其中arg{•}代表过滤器的参数(相位)。在此之后，使用公式16和公式14设计均衡滤波器，然后，将得到的滤波器与$H_{EQ_}$进行卷积，以得到均衡后的双耳滤波器$H_{phEQ}$。

[0228] 此外，还使用位于聆听位置左侧和右侧9cm处的全向扩音器(G.R.A.S.型号40DP)在试视听室中测量立体扬声器装置。来自一个扬声器的直接声音到达每个扩音器位置的时间差异近似于用虚拟头部获得的ITD。这些响应被窗口化为42ms并以与$H_{phEQ}$类似的方式处理，但ILD是由Kirkeby,O发表于2002年音频工程学会会议第22届国际会议“虚拟、综合及娱乐音频”中的“A Balanced Stereo Widening Network for Headphones”中提出的直接滤波器和串扰滤波器引入的。这些滤波器表示为$H_{d_{ph}}$和$H_{x_{ph}}$，它们的频率响应如图26所示。得到的均衡双耳滤波器表示为$H_{oomEQ}$。

[0229] 对于左耳机信道，在图27中示出了在直接滤波器和串扰滤波器的加和(图24中的s”)之后的滤波器$H_{binEQ}$，$H_{phEQ}$，和$H_{roomEQ_}$的响应。与平坦响应的偏差是由于为了近似对称滤波器而在耳朵之间进行平均化和在该过程中选择的平滑窗口导致的。

[0230] 听力测试设计

[0231] 设计了由三个独立部分组成的听力测试，以便分别评估空间立体声质量、音色/音质和整体音质。听力测试使用前一节中测量的房间内专用的耳机(Stax SR-307)进行。要评估的案例是通过耳机直接再现立体声信号，以及使用在滤波器设计章节(即$H_{bin}$，$H_{binEQ}$，$H_{phEQ}$，和$H_{roomEQ}$)中描述的处理之后获得的双耳滤波器的双耳立体声再现。在测试中引入低通滤波(3.5kHz截止频率)单声道信号作为低锚点。

[0232] 选择了四个立体声音轨道进行测试。第一作者通过向各个方向平移的不同乐器循环将两个立体声轨道混合。另外两个立体声音轨是几首简短商业音乐的混音(乡村和摇滚)。这些立体声轨道与每个双耳滤波器进行卷积，并且使用由测试参与者控制的图形用户界面以无缝连续循环的方式再现所得到的信号。图形用户界面允许参与者随意愿多次选择测试用例和参考，然后使用滑块使用0到100的数字等级对每个测试用例进行评分。质量描述符(坏、差、尚可、良好、优秀)在滑块的右侧可见。参与者被指示将最差用例评为0，最佳用例评为100。然后，应根据感知差异对剩余用例进行评分。这对所有测试都有效。

[0233] 第一个测试(表示为测试1)比对这参考项产生的空间立体声质量来评估不同用例的空间立体声质量。参考项是$H_{bin}$，因此它在测试1中用作隐藏参考。要参与测试，参与者需要在收听参考项时应该能感知到外在化。否则，参与者的数据就不被包括在分析中。在测试1中，指示参与者通过关注于听觉图像中的幻象源的定位、宽度和分布来避免音色的变化可能对空间特征的感知造成的任何影响。

[0234] 在测试2中，将每种用例产生的声音质量与参考项进行比较。参考项是通过耳机直接再现立体声信号。因此，测试包括隐藏的参考。参与者被指示在评分时忽略空间化的影响，同时关注于不同幻象源、声音平衡和声音伪像的响度/音色差异。

[0235] 测试3在再现立体声时基于整体声音质量来评估不同用例。此测试中没有参考项，但参与者被指示假设虚拟参考。这个虚拟参考是参与者对通过扬声器播放音乐的立体声再现的个人期望。对于该测试，参与者应根据他的个人期望来考虑空间和音色。

[0236] 共有14名年龄在23岁至45岁之间的受试者参加了测试。其中一位参与者在测试1中没有感知到外在化。因此，在所有测试中他的数据被排除在分析之外，并且针对剩余的13名参与者进行结果分析。

[0237] 结论

[0238] 使用$x^2$拟合优度程序测试数据的正态性。通过如下滤波器获得的分数违反了正态性假设

[0239] 测试1中的$H_{binEQ}$($x^2(4,52)=13.22$，$p=0.01$)；

[0240] 测试2中的($x^2(4,52)=10.75$，$p=0.0294$)；以及通过

[0241] $H_{binEQ}$($x^2(2,52)=6.98$，$p=0.0304$)以及

[0242] 测试2中的$H_{roomEQ}$($x^2(4,52)=12.11$，$p=0.0165$)

[0243] 发现：三个听力测试的数据也违反了方差齐性的假设(针对测试1、测试2和测试3分别为$p=0.00206$，$p=2.87\times10^{-5}$，和$p=1.327\times10^{-11}$)。因此，对于从每个听力测试获得的数据，执行弗里德曼氏非参数统计分析和具有Bonferroni校正的双尾Wilcoxon符号铁事后检验。

[0244] 测试1：空间质量

[0245] 测试1的数据的非参数分析

[0246] ($x^2(3)=107.06, p=4.69\times10^{-23}$)表明由不同过滤器获得的分数不共享相同的分布。事后检验证实了所有用例都不同(见图28)。图29示出汇总数据的中位数和四分位数。通过耳机直接再现立体声信号表示为“Direct”,参考项是$H_{bin}$。参考和低锚点未在图中示出,因为它们总是分别为100和0。方框中的缺口表示中位数的95%置信区间,异常值标记为十字形。每个过滤器的中位数按照与$H_{bin}$中包含的双耳信息的劣化一致的趋势进行排序。发现:包含了与$H_{bin}$相同的耳间差异的滤波器$H_{binEQ}$要比$H_{phEQ}$(仅包含与$H_{bin}$相同的相位)和$H_{roomEQ}$能更好地再现参考项的空间特征,并具有人工引入的双耳信息。发现:通过耳机直接再现立体声信号很难再现参考项的空间特性。

[0247] 测试2:音色/声音平衡质量

[0248] 非参数分析(($x^2(3)=104.38, p=1.77\times10^{-22}$))发现不同用例的得分的分布存在显著差异。事后检验的结果如图30所示。事后检验证实,除了$H_{binEQ_}$和$H_{phEQ_}$($Z=0.915, p=0.845$)之外,用例之间的数据分布差别很大。这也可以在图31中看到,其中$H_{binEQ_}$和$H_{phEQ_}$显示出相似的分布和相似的中位数置信区间。在该测试中,通过耳机直接再现立体声信号被用作参考项。不同用例的分数按滤波器引入的幅度失真量进行排序。$H_{roomEQ_}$中使用的直接滤波器和串扰滤波器是平滑的并且设计用于产生平坦响应,因此引入了较小的幅度失真。$H_{binEQ_}$包含$H_{bin}$的耳间差异,但它与被人工引入耳间水平差异的$H_{binEQ_}$得分相同。此外,$H_{bin}$明显优于此测试中的其他过滤器,但$H_{binEQ_}$和$H_{phEQ_}$相对接近$H_{roomEQ_}$的分数。与图27中的响应相比,这些结果表明,与通过耳机的直接再现相比,平滑滤波器响应可以改善音质。然而,如在$H_{phEQ}$中那样去除单声道和ILD提示以产生更平滑的滤波器,并没有改善$H_{phEQ}$的音色质量,$H_{phEQ}$包含与$H_{bin}$相同的双耳信息。

[0249] 测试3:整体质量

[0250] 测试3中的数据分布之间存在显著差异(($x^2(4)=114.21, p=9.17\times10^{-24}$))。事后检验结果证实,除了通过耳机的直接再现与$H_{bin_}$($Z=0.77, p=0.43$)形成的一对以及由$H_{binEQ_}$和$H_{phEQ_}$($Z=0.87, p=0.38$)形成的一对之外,每种用例的得分都不同。事后检验的结果如图32所示。

[0251] 虽然事后检验发现$H_{binEQ_}$和$H_{phEQ_}$之间没有差异,但图33中的箱线图显示$H_{binEQ_}$的得分略高。具有后均衡的双耳滤波器(用下标EQ表示)优于通过耳机的直接再现和$H_{bin}$获得的分数。直接立体声再现和$H_{bin}$的相似分布表明,参与者相似地针对缺乏空间印象和音色失真进行了罚分。这些结果不同于Lorho,G.,Isherwood,D.,Zacharov,N.,和Huopaniemi,J发表于2002年音频工程学会会议第22届国际会议“虚拟、综合及娱乐音频”中的“Round Robin Subjective Evaluation of Stereo Enhancement System for Headphones”中得出的结果,其中该文献涉及选择虚拟参考(扬声器装置),而不涉及音质的抽象定义。

[0252] 结束语

[0253] 本研究的重点是使用双耳滤波器来重现扬声器立体声组对的空间印象,同时保留原始耳机音质。定义和评估了用于在扬声器立体声再现的双耳渲染中保持耳机的原始声音质量的标准。设计后均衡滤波器,使其平坦化从扬声器到每个耳朵的直接路径和串扰路径的总和的输出。这不同于其他那些将同侧和对侧HRTF修改为期望方向的均衡方法。所提出的均衡方法与Kirkeby,O发表于2002年音频工程学会会议第22届国际会议“虚拟、综合及娱

乐音频”中的“A Balanced Stereo Widening Network for Headphones”中提出的概念相同，但这里概括为使用双耳室内响应。测量的双耳室内响应(42ms)用于设计双耳滤波器，这允许很少的早期反射，同时避免由于混响导致的过度音质效应。设计改进的双耳滤波器以使得一些原始双耳属性被平滑或由人工双耳信息代替。上述标准用于设计后均衡滤波器，其用于平坦化不同双耳滤波器的直接滤波器和串扰滤波器的总和。进行听力测试以评估双耳滤波器在空间质量、音质/声音平衡质量和整体质量方面的性能。结果表明，保持原始双耳滤波器的直接路径和串扰路径之间的差异是必要的，以便保持双耳渲染的空间质量，并且使得这种双耳滤波器的后均衡仍然保持耳机的声音质量。当听众被问及他们对于立体声音乐再现应该是怎样的个人期望时，所设计的滤波器优于典型的双耳渲染和耳机上的典型立体声再现。这证实了所提出的标准对于保持耳机的音质的同时增强声音的空间特性方面的适用性。

[0254] 应当理解的是所公开的本发明的实施方案不限于本文公开的具体结构，工艺步骤或材料，而是扩展到相关领域的普通技术人员将认识到的其等同物。还应该理解的是本文采用的术语仅用于描述特定实施方案的目的，而不是限制性的。

[0255] 贯穿本说明书对一个实施例的引用意味着包括在本发明的至少一个实施例中结合该实施例描述特定的特征，结构或特性。因此，贯穿本说明书中在各个地方出现的短语“在一个实施例中”或“在实施例中”不一定都指代相同的实施例。在使用术语(例如，大约或基本上)参考数值的情况下，还公开了精确的数值。

[0256] 如本文所使用的，为方便起见，可以在公共列表中呈现多个项目，结构元素，组成元素和/或材料。但是，这些列表应该被解释为列表中的每个成员都被单独标识为一个独立且独特的成员。因此，在没有相反指示的情况下，不应仅仅根据个体成员在公共团体中的表象就将此类列表中的这些个体成员解释为同一列表中其他成员的事实等同物。另外，本文中的各种实施方案和实施例可以与其各种组件的替代物一起提及。应当理解，这些实施例，示例和替代方案不应被解释为彼此的事实上的等同物，而是应被视为本发明的单独和自主的表示。

[0257] 此外，所描述的特征，结构或特性可以在一个或多个实施例中以任何合适的方式组合。在以下描述中，提供了许多具体细节，例如长度，宽度，形状等示例，已提供对本发明实施例的透彻理解。然而，相关领域的技术人员将认识到，可以在没有一个或多个具体细节的情况下或者利用其他方法，组件，材料等实践本发明。在其他情况下，未详细示出或描述公知的结构，材料或操作以避免模糊本发明的各方面。

[0258] 虽然前述示例在一个或多个特定应用中说明了本发明的原理，对于本领域普通技术人员来说显而易见的是，可以在不运用创造性能力的情况下对形式，用法和实施细节进行多种修改，并且不脱离本发明的原理和概念。因此，除了下面提出的权利要求之外，并不意图限制本发明。

[0259] 动词“包含”和“包括”在本文中用作开放式限制，既不排除也不要求存在未记载的特征。除非另有明确说明，否则从属权利要求中所述的特诊可相互自由组合。此外，应该理解的是在整个本文件中使用“一个”，即单数形式，并不排除多个。

[0260] 工业实用性

[0261] 本发明的至少一些实施例在声音再现装置及系统中得到了工业应用。

[0262] 还可以以下列方式来考虑本发明:耳机具有两个声道,但其不会像成对的立体声扬声器那样重现相同的听觉印象。本发明涉及通过技术手段最小化这两种方案(扬声器<->耳机)的差异。

[0263] 在以下段落中描述了本发明的一些方面。

[0264] 段落1、一种形成用于立体声耳机的双耳滤波器以保持所述耳机的音质的方法,其特征在于,从扬声器到每只耳朵的直接路径和串扰路径的加和具有平坦的幅度响应。

[0265] 段落2、根据段落1所述的方法,其中,仅进行相位均衡。

[0266] 段落3、根据上述段落中的任一项所述的方法,其中,双耳滤波器形成为使得通过有利地为340ms的预定混响时间针对试听室内的立体声扬声器装置来测量仿真头部的双耳时间响应$h_{ij}(t)$,并且使用所测量的响应,通过对所述响应的例如42ms的第一预设时间进行窗口化来设计一组双耳滤波器$H_{bin}$,

[0267] $$H_{\text{bin}} = \mathscr{F}\left\{h_{\text{ij}}(t)\,w(t)\right\},\ i \in \{L,R\},\ j \in \{l,r\}, \tag{15}$$

[0268] 其中,$\mathscr{F}\{\cdot\}$表示傅里叶变换,以及w(t)是预定义的例如42ms时长的时间窗口,并且,在进行非正式听力测试之后,采用这种滤波器长度作为外在化能力和由于室内混响引起的音色效果之间的最佳折中。

[0269] 段落4、根据上述段落中的任一项所述的方法,其中,使用$H_{binEQ}$作为双耳滤波。

[0270] 段落5、根据上述段落中的任一项所述的方法,其中,使用$H_{phEQ}$作为双耳滤波。

[0271] 段落6、一种用于校准根据上述段落中的任一项所述的立体声耳机(1)的方法,所述立体声耳机(1)包括具有存储器和信号处理特性的放大器(2),所述方法包括步骤:比对设定的参考耳罩或驱动器来校准所述耳机(1)的驱动器或耳罩,并且将校准设置存储在所述放大器(2)的存储器中。

[0272] 段落7、根据段落1所述的方法,其中,通过设置所述放大器(2)中的信号处理参数来确定所述耳机(1)的期望的声音属性,以便通过测量或基于从所述耳机(1)的用户处接收的输入信息来获得所述期望的声音属性。

[0273] 段落8、根据上述段落中的任一项所述的方法,其中,所述方法包括如下工厂校准步骤:校准至少幅度响应,典型地为包括了相位响应的频率响应。

[0274] 段落9、根据上述段落中的任一项或其结合所述的方法,其中,所述声音属性包括以下特性中的至少一个:“频率响应”、“时间响应”、“相位响应”或“灵敏度”。

[0275] 段落10、根据上述段落中的任一项或其结合所述的方法,其中,诸如频率响应的所述期望的声音属性基于特定房间的耳机系统的校准参数来确定。

[0276] 段落11、根据前述段落中任一项所述的方法,其中,为信号处理参数执行外在化函数,以为耳机的用户创建室内效果。

[0277] 段落12、根据段落11所述的方法,其中,通过对双耳滤波器执行外在化函数,使得其为全通滤波器。

[0278] 段落13、根据段落11所述的方法,其中,所述双耳滤波器具有恒定幅度响应(幅频/振幅不作为频率的函数而变化),但只实施双耳滤波器的相位响应。

[0279] 段落14、根据段落11所述的方法,其中,双耳滤波器为FIR滤波器。

[0280] 段落15、根据上述段落中的任一项所述的方法,其中,

[0281] a.测试信号由扬声器通过第一子波带($B_1$)再现，

[0282] a.所述测试信号由耳机(1)通过所述第一子波带($B_1$)再现，

[0283] b.利用由所述扬声器通过所述第一子波带($B_1$)再现的测试信号来评估由所述耳机(1)通过所述第一子波带($B_1$)再现的测试信号的诸如声级的声音属性，并将所述耳机的诸如声级的声音属性设置并存储为与所述扬声器在子波带($B_1$)中的声音属性基本相同，

[0284] c.通过若干子波带$B_1$-$B_n$使用所述测试信号重复上述步骤。

[0285] 段落16、根据段落15所述的方法，其中，所述测试信号为粉红噪声。

[0286] 段落17、根据段落15或16所述的方法，其中，所述测试信号是包括具有宽频谱含量的音频信号的类音乐音频文件。

[0287] 段落18、根据段落15-17中任一项所述的方法，其中，所述测试信号的持续时间为1-10秒。

[0288] 段落19、根据段落15-18中任一项所述的方法，其中，所述测试信号连续地重复。

[0289] 段落20、一种有源立体声/双声道耳机系统，其包括具有用于每个耳罩的至少一个驱动器的耳机(1)和通过线缆(3)连接到所述耳机(1)放大器(2)，所述系统(1,2,3)包括：

[0290] b.耳罩，

[0291] c.在所述放大器(2)中用于信号处理的机构，

[0292] d.比对例如耳罩或驱动器的设定参考，对所述耳机(1)的所述驱动器或所述耳罩中的每一个进行工厂校准，并且储存在所述放大器(2)的存储器中，

[0293] e.用于将至少两个预定的均衡设置存储在放大器(2)中的机构，以及

[0294] f.用于消除低于200Hz的频率的噪音的机构。

[0295] 段落12、根据段落11所述的系统，其中，所述耳罩例如以罩耳的方式完全地覆盖耳朵。

[0296] 段落21、根据段落20所述的系统，其中，所述耳罩例如以罩耳的方式完全地覆盖耳朵。

[0297] 段落22、根据段落20或21所述的系统，其中，所述参考是通过测试获得的或来自于参考驱动器或耳罩的预定频率响应。

[0298] 段落23、根据上述段落中的任一项所述的有源耳机系统，其中，所述耳机(1)和所述耳机放大器(2)是通过线缆(3)彼此连接的分开的独立单元。

[0299] 段落24、根据上述段落中的任一项所述的有源耳机系统，其中，比对设定的参考耳罩或驱动器，来对所述耳机(1)的每个驱动器或耳罩进行工厂校准，并且储存在所述放大器(2)的存储器中，由此所述工厂校准使所述耳机系统中的所有耳罩在声学上大致相同，例如基于设定的参考耳罩或驱动器而响应相同、响度相同。

[0300] 段落25、根据上述段落中的任一项所述的有源耳机系统，其中，所述耳机放大器和所述耳机基于工厂校准之后构成唯一性的配对。

[0301] 段落26、根据前述段落中的任一项所述的有源耳机系统，其中，其中，所述有源耳机系统包括使用信号处理参数对音频进行外在化的机构，以为耳机的用户创建室内效果。

[0302] 段落27、根据上述段落中的任一项所述的有源耳机系统，其中，通过对双耳滤波器执行外在化函数。

[0303] 段落28、根据上述段落中的任一项所述的有源耳机系统，其中，所述双耳滤波器是

[0304] g.全通滤波器,或者

[0305] h.具有相位响应和幅度响应的滤波器。

[0306] 段落29、根据上述段落中的任一项所述的有源耳机系统,其中,将所述扬声器的传递函数引入至所述耳机系统。

[0307] 段落30、根据上述段落中的任一项所述的有源耳机系统,其中,将所述耳机系统的传递函数输出到扬声器系统。

[0308] 段落31、根据上述段落中的任一项所述的有源耳机系统,其中,对于所述扬声器和所述耳机而言,音量控制是相同的。

[0309] 段落32、一种计算机程序,配置为促使根据前述方法段落中的至少一项被执行。

[0310] 段落33.一种用于形成双耳滤波器的方法,所述方法通过耳机模拟室内的扬声器立体声再现的听觉印象,或增强耳机再现中的立体声空间特性,同时保持耳机的音质,其特征在于,从扬声器到每只耳朵的直接路径和串扰路径形成为使得直接路径和串扰路径的加和基本上不作为频率的函数而变化。

[0311] 缩略语列表

[0312] IIR 无限脉冲响应

[0313] FIR 有限脉冲响应

[0314] IR 脉冲响应

[0315] ARM 自适应多速率音频数据压缩方案

[0316] GLM 真力扬声器管理

[0317] SPL 声压级

[0318] ISS 休眠控制

[0319] EAI 增强式低频隔离

[0320] 引用列表

[0321] 非专利文献

[0322] Kirkeby,O.,“ABalanced Stereo Widening Network for Headphones,”in Audio Engineering Society Conference:22nd International Conference:Virtual, Synthetic,and Entertainment Audio,2002.

[0323] Lorho,G.,Isherwood,D.,Zacharov,N.,and Huopaniemi,J.,“Round Robin Subjective Evaluation of Stereo Enhancement System for Headphones,”in Audio Engineering Society Conference:22nd International Conference:Virtual, Synthetic,and Entertainment Audio,2002.

[0324] B.Masiero and J.Fels,“Perceptually robust headphone equalization for binaural reproduction,”in Audio Engineering Society Convention 130,May 2011S.G.Norcross,G.A.Soulodre,and M.C.Lavoie,“Subjective investigations of inverse filtering,”J.Audio Eng.Soc,vol.52,no.10,pp.1003-1028,2004

[0325] Z. Schärer and A.Lindau,“Evaluation of equalization methods for binaural signals,”in Audio Engineering Society Convention 126,May 2009

[0326] 附图标记列表

[0327] 1 包括双耳驱动器的立体声耳机

[0328] 2 耳机放大器

[0329] 3 耳机线缆

[0330] 30 电池

[0331] 31 充电子系统

[0332] 32 SMPS电源和电池管理

[0333] 33 USB输入

[0334] 34 本地用户界面

[0335] 35 模拟输入

[0336] 36 模数转换器(ADC)

[0337] 37 自适应多速率(AMR)和数字信号处理(DSP)

[0338] 38 数模转换器(DAC)

[0339] 39 功率放大器

[0340] 40 功率放大器

[0341] 41 自动校准模块

[0342] 42 耳部校准模块

[0343] 43 工厂均衡器/校准

[0344] 45 音量控制器

[0345] 46 动态处理器

[0346] 47 USB接口功能

[0347] 48 软件接口

[0348] 49 存储器管理

[0349] 50 电源和电池管理

[0350] 51 计算机运行软件

[0351] 52 用于用户界面的连接器线缆

[0352] 54 耳机放大器的控制钮

[0353] 55 电源线缆

[0354] 56 便携式终端

[0355] 60 耳机改进元件

[0356] 61 监听改进元件

[0357] $B_1$-$B_n$音频的子波带

[0358] Δf 子波带的带宽,通常是一个倍频程。

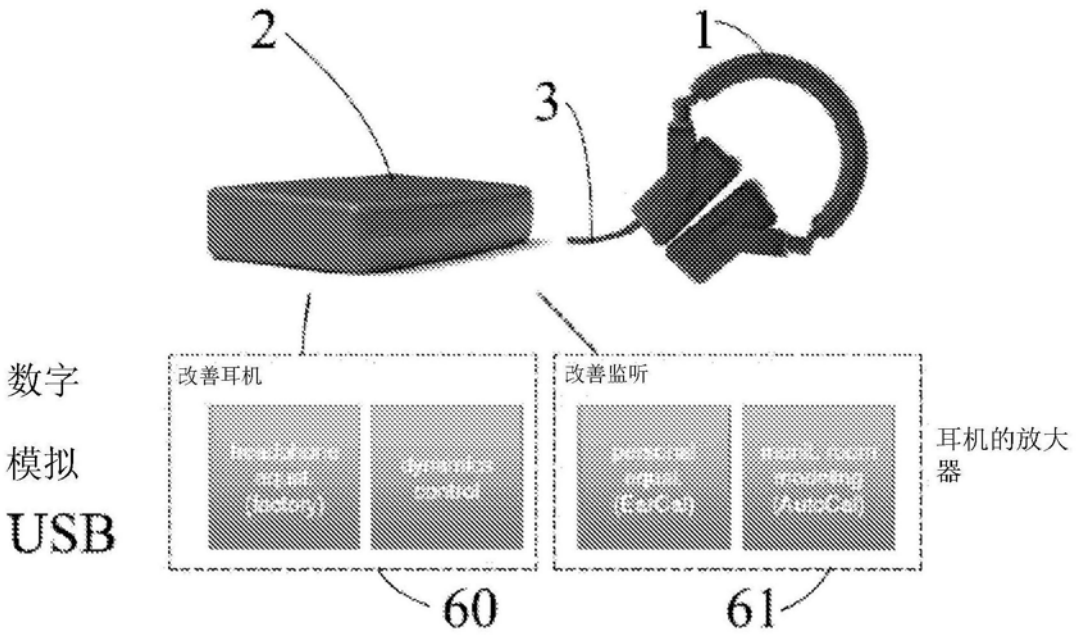



图1

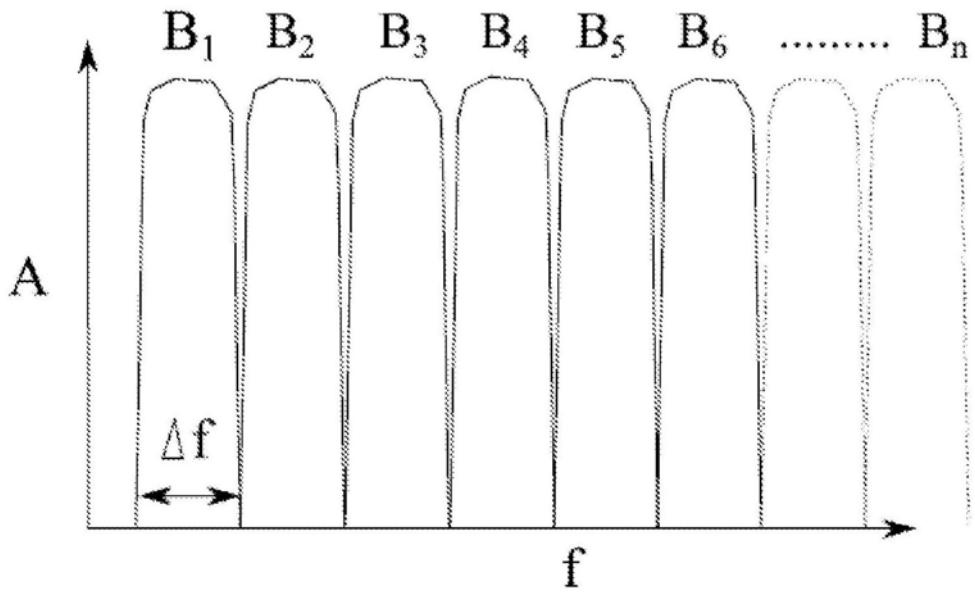



图2

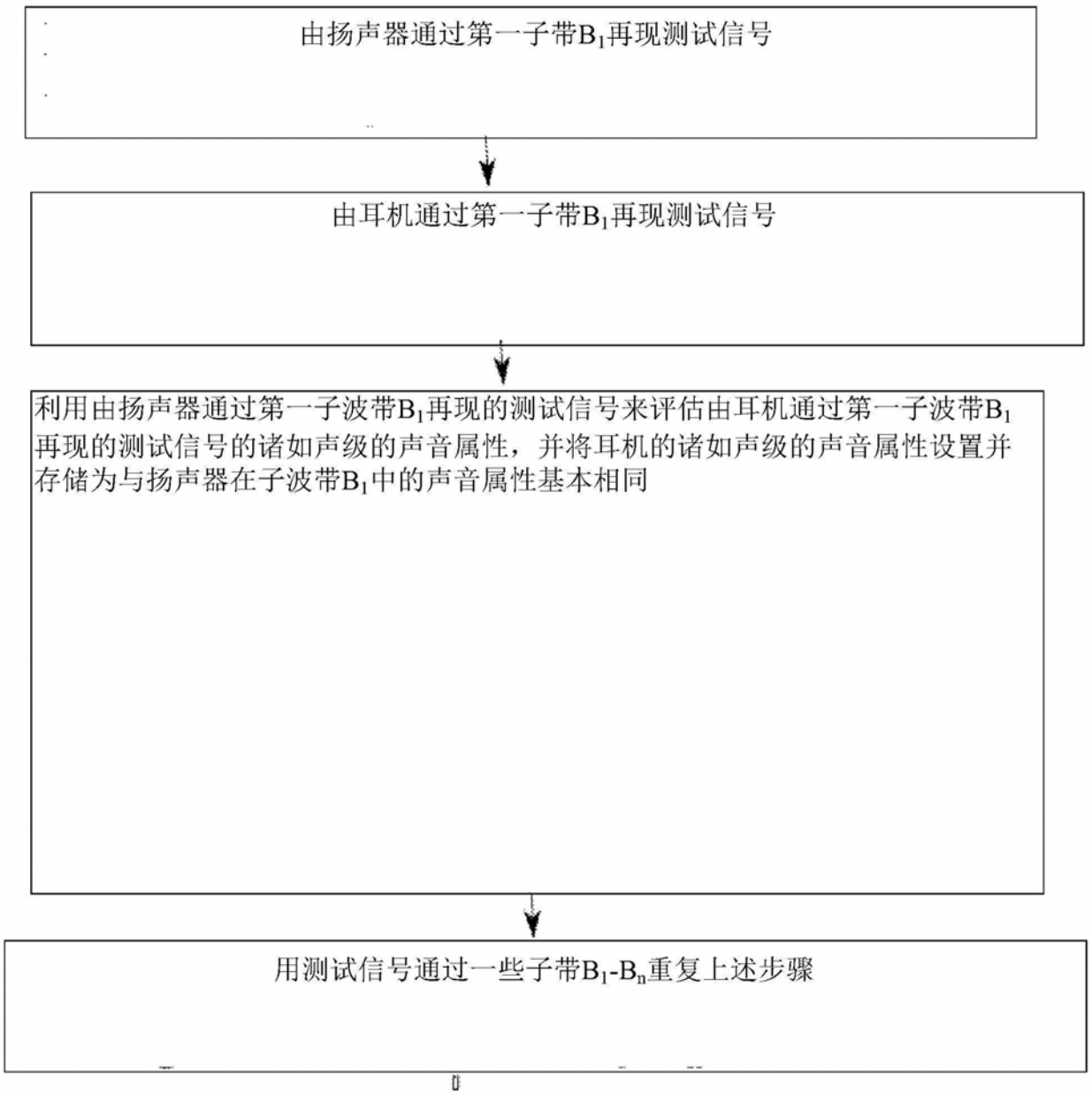

由扬声器通过第一子带$B_1$再现测试信号
由耳机通过第一子带$B_1$再现测试信号
利用由扬声器通过第一子波带$B_1$再现的测试信号来评估由耳机通过第一子波带$B_1$再现的测试信号的诸如声级的声音属性，并将耳机的诸如声级的声音属性设置并存储为与扬声器在子波带$B_1$中的声音属性基本相同
用测试信号通过一些子带$B_1$-$B_n$重复上述步骤

图3

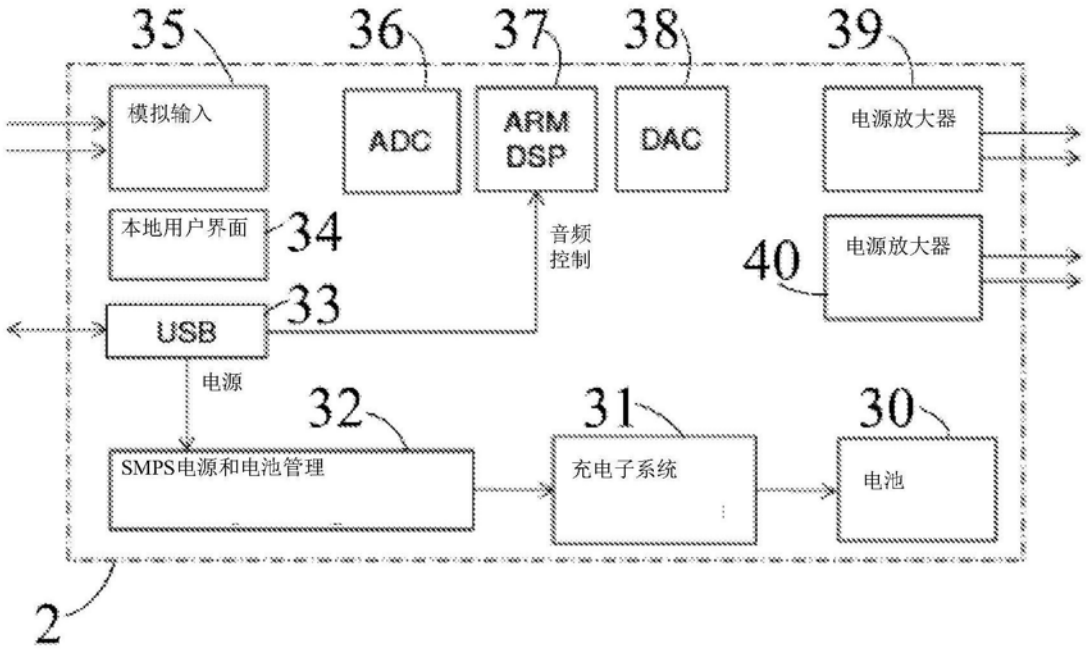

35
36
37
38
39
模拟输入
ADC
ARM
DSP
DAC
电源放大器
本地用户界面
34
音频
控制
40
电源放大器
USB
33
电源
32
31
30
SMPS电源和电池管理
充电子系统
电池
2

图4

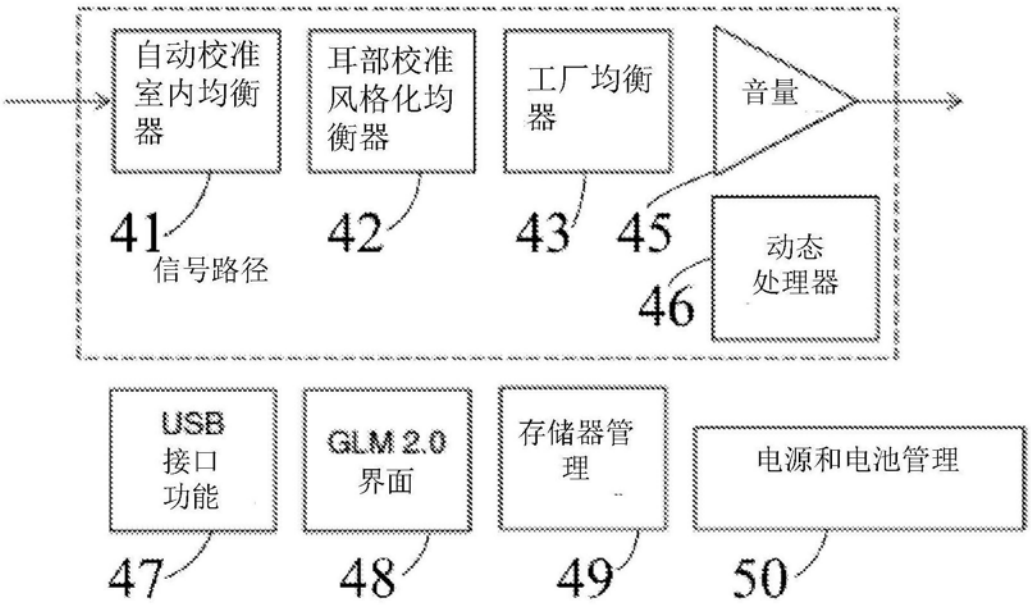

自动校准
室内均衡
器
耳部校准
风格化均
衡器
工厂均衡
器
音量
41
信号路径
42
43
45
动态
处理器
46
USB
接口
功能
GLM 2.0
界面
存储器管
理
电源和电池管理
47
48
49
50

图5

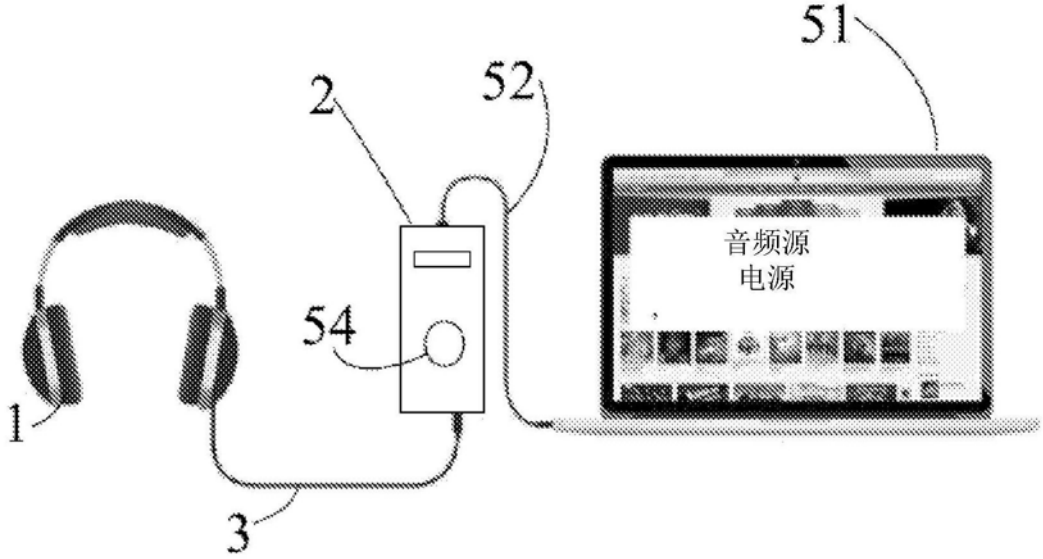



图6

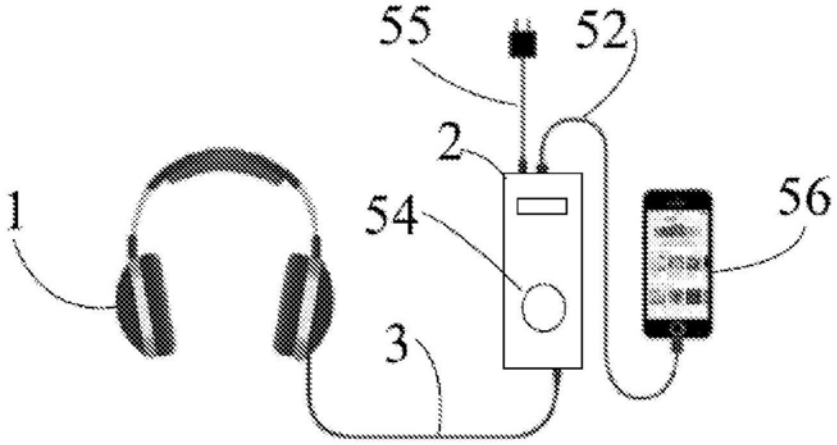



图7

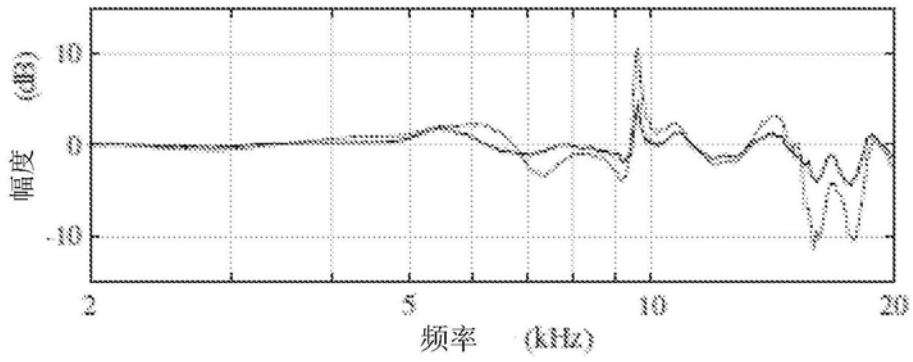



图8

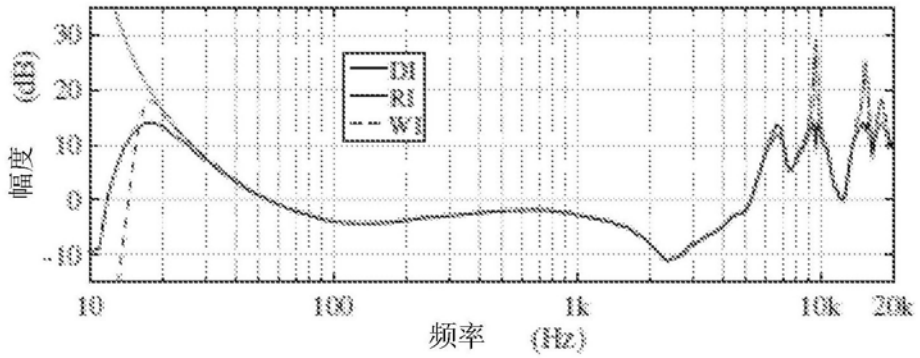

幅度 (dB)
30
20
10
0
-10
D1
R1
W1
10
100
1k
10k
20k
频率 (Hz)

图9

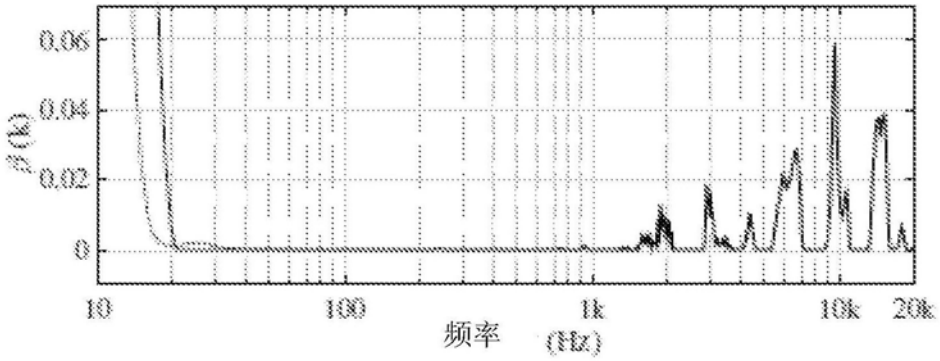

β(k)
0.06
0.04
0.02
0
10
100
1k
10k
20k
频率 (Hz)

图10

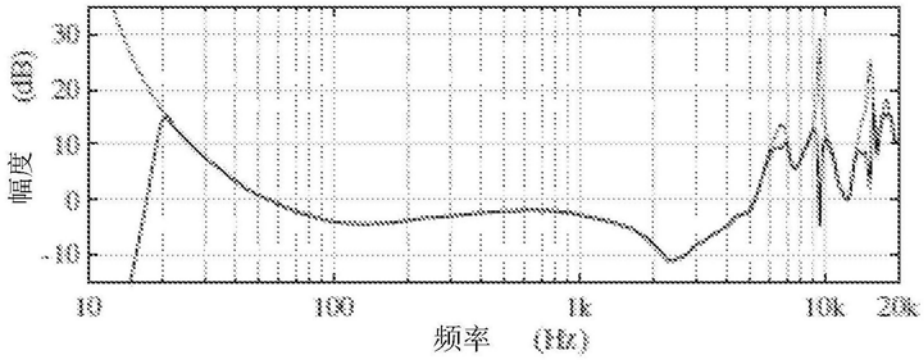

幅度 (dB)
30
20
10
0
-10
10
100
1k
10k
20k
频率 (Hz)

图11

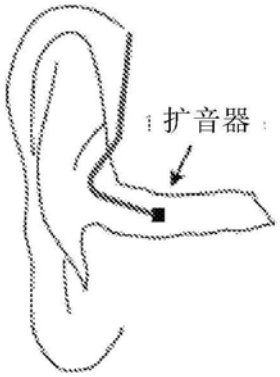



图12a



图12b

| 方法 | $\alpha$ | $\sigma^2$ | 平滑因子 |
|---|---|---|---|
| WI | 公式7 | 0 | … |
| RI | 公式6 $\beta$ | $\beta$ | … |
| SM | 公式6 | 0 | 半倍频程 |
| SI | 公式6/公式7 | 公式8 | 半倍频程 |

图13

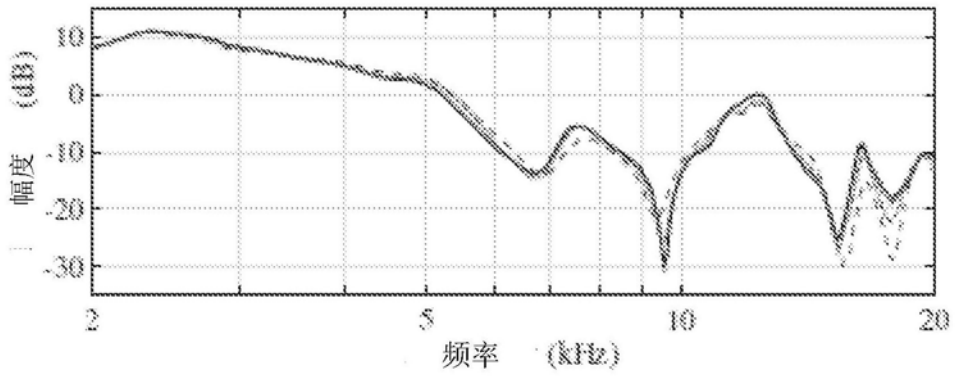

10
0
-10
-20
-30
幅度 (dB)
2
5
10
20
频率 (kHz)

图14

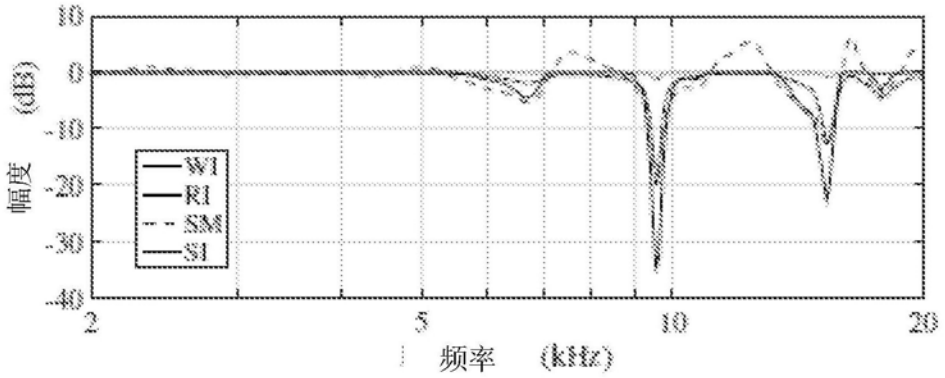

10
0
-10
-20
-30
-40
幅度 (dB)
W1
R1
SM
S1
2
5
10
20
频率 (kHz)

图15

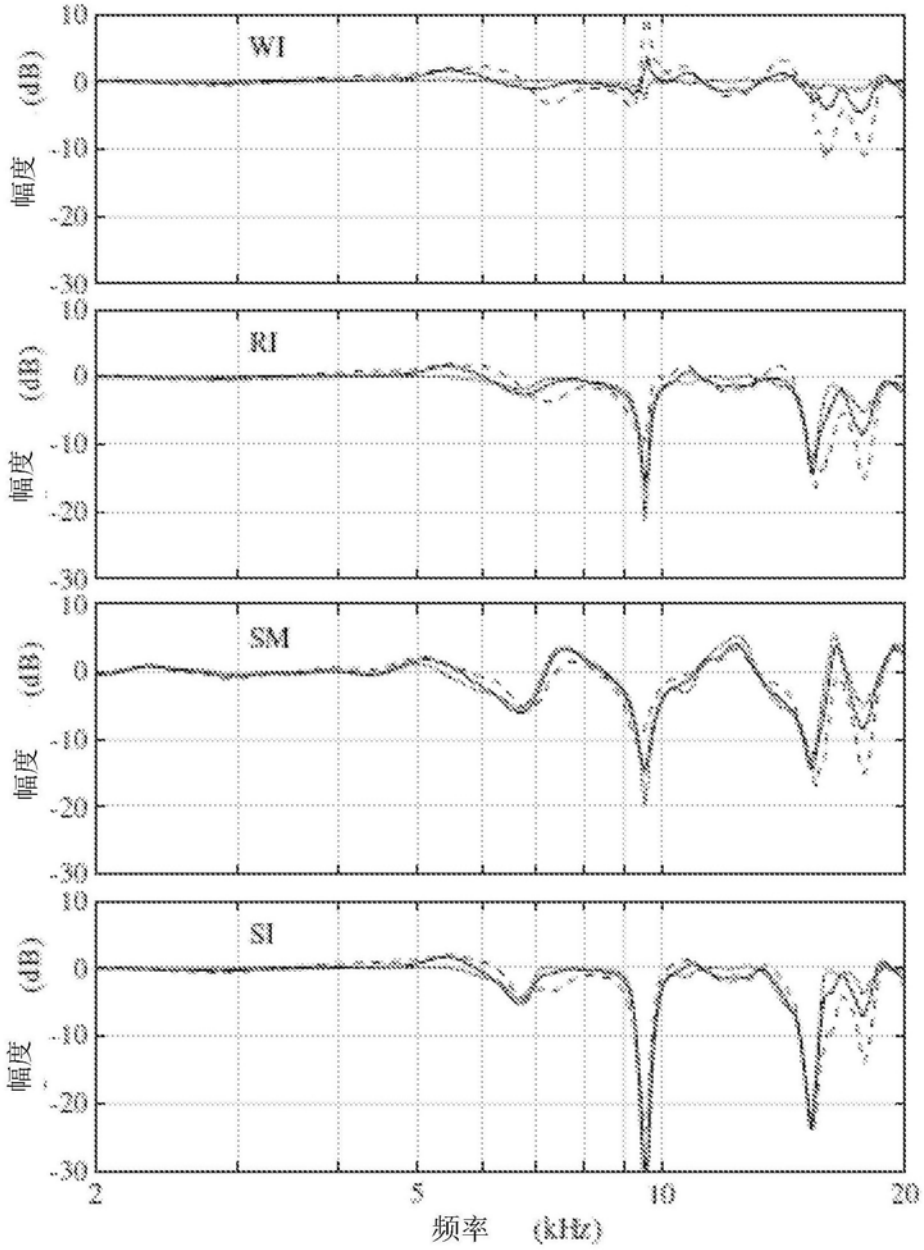



图16

| 方法 | μ | SD |
|---|---|---|
| NF | 19.95 | 19.45 |
| RI | 79.8 | 14.33 |
| SM | 69.62 | 25.7 |
| SI | 89.62 | 8.04 |

图17

| 方法 | RI | NF | SM |
|---|---|---|---|
| SI | .002 | < .001 | < .001 |
| SM | .139 | < .001 | -- |
| NF | < .001 | -- | -- |

图18

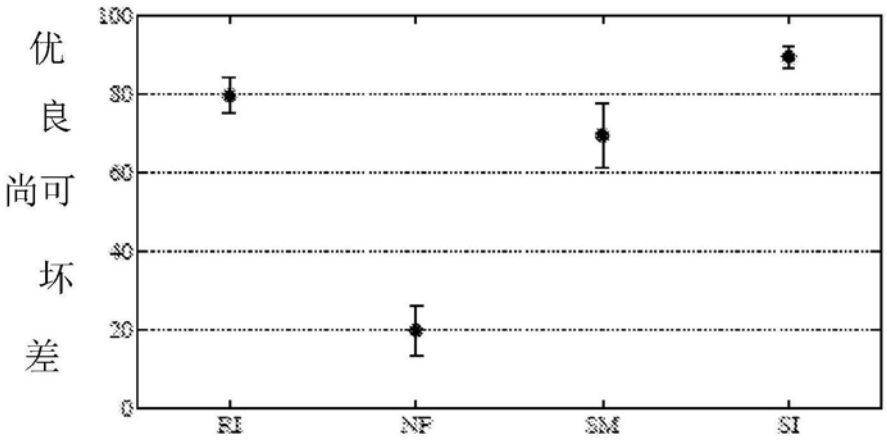



图19

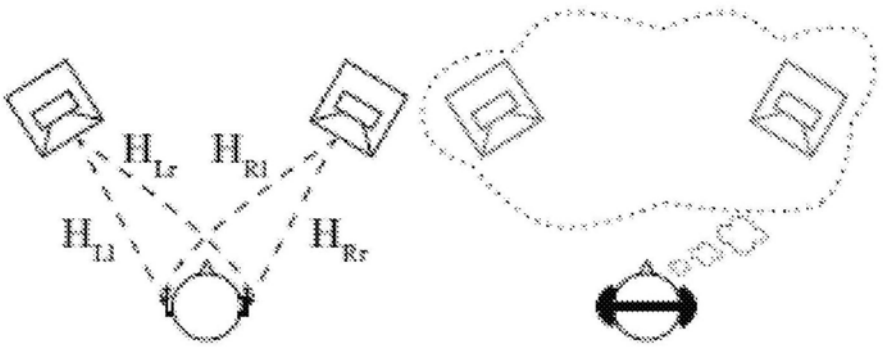



图20

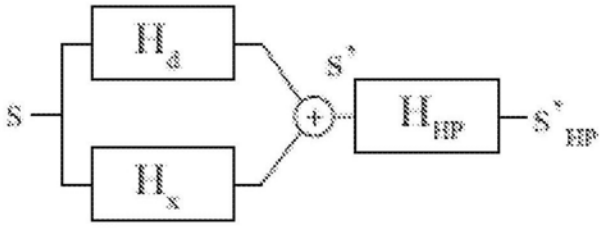



图21

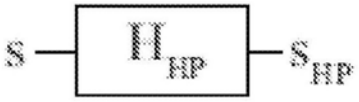



图22

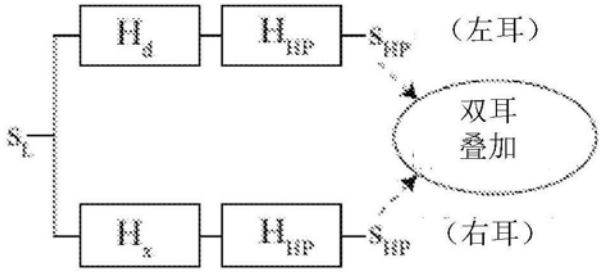



图23

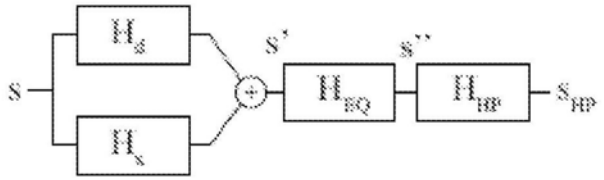



图24

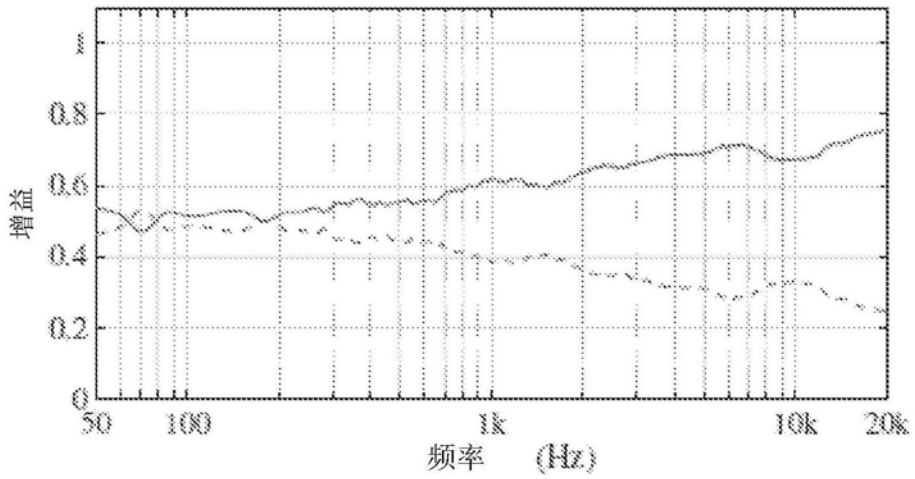

增益
1
0.8
0.6
0.4
0.2
0
50
100
1k
10k
20k
频率 (Hz)

图25

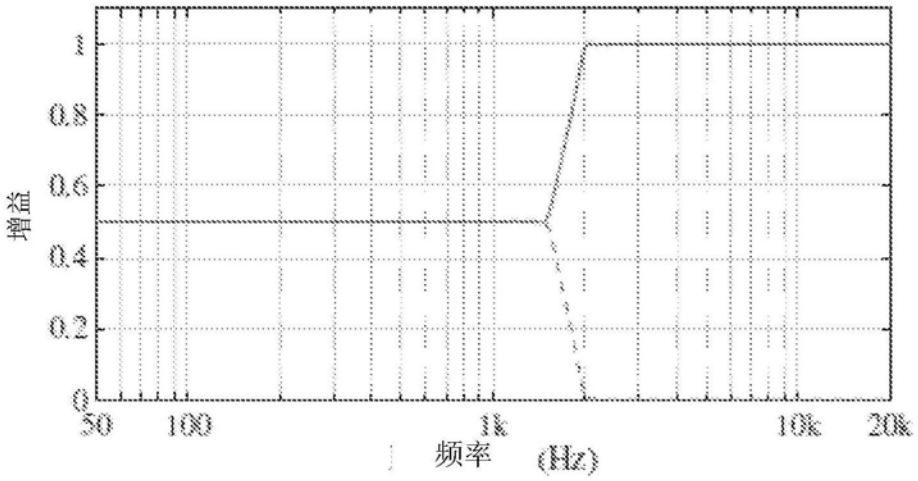

增益
1
0.8
0.6
0.4
0.2
0
50
100
1k
10k
20k
频率 (Hz)

图26

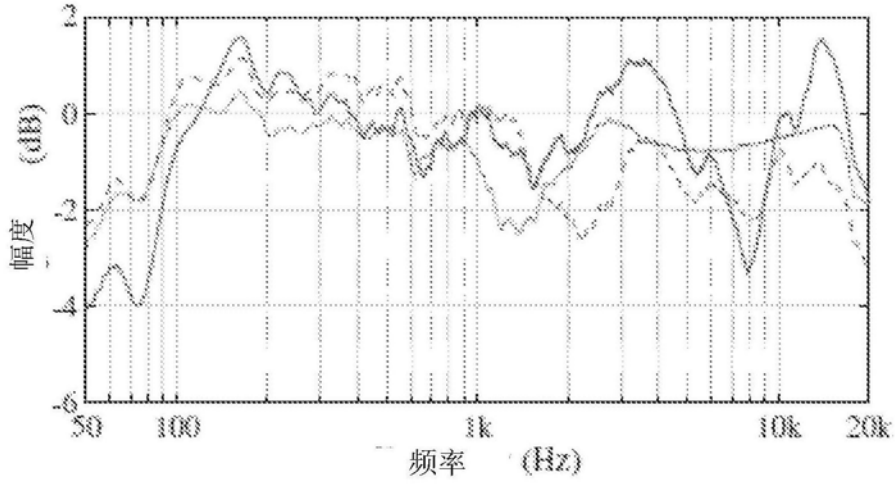



图27

| $(Z/p)$ | Direct | $H_{bin}$ | $H_{binEQ}$ | $H_{phEQ}$ |
|---|---|---|---|---|
| $H_{bin}$ | 9.25/0 | – | – | – |
| $H_{binEQ}$ | 7.9/0 | 8.55/0 | – | – |
| $H_{phEQ}$ | 7.12/0 | 8.97/0 | 4.03/0 | – |
| $H_{roomEQ}$ | 5.69/0 | 9.25/0 | 6.36/0 | 3.23/0 |

图28

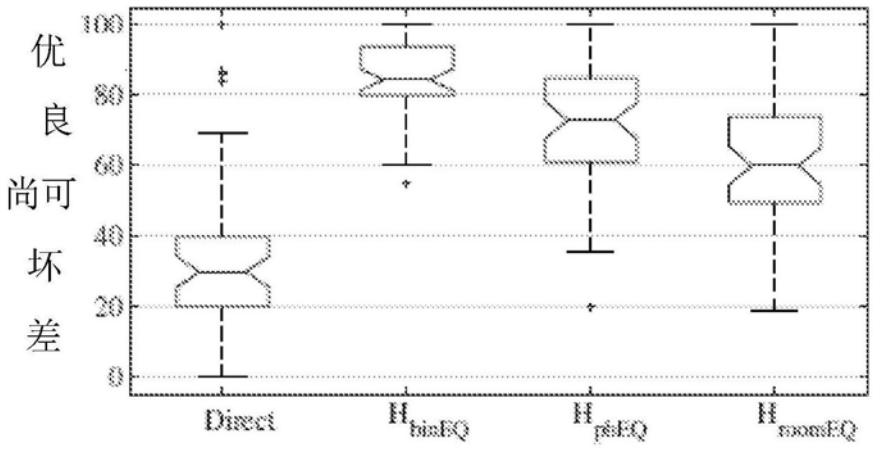



图29

| $(Z/p)$ | Direct | $H_{bin}$ | $H_{binEQ}$ | $H_{phEQ}$ |
|---|---|---|---|---|
| $H_{bin}$ | 9.393/0 | – | – | – |
| $H_{binEQ}$ | 9.396/0 | 6.98/0 | – | – |
| $H_{phEQ}$ | 9.25/0 | 7.28/0 | **0.915/0.845** | – |
| $H_{roomEQ}$ | 8.83/0 | 8.5/0 | 4.05/0 | 4.33/0 |

图30

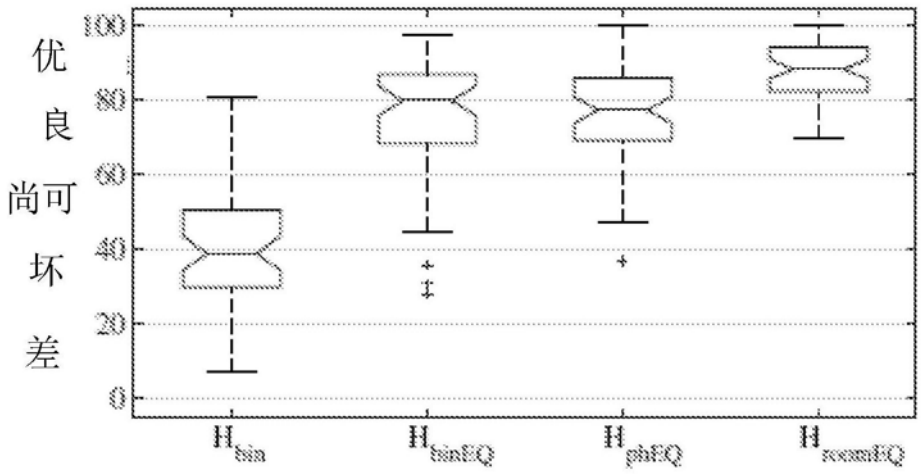



图31

| $(Z/p)$ | Direct | $H_{bin}$ | $H_{binEQ}$ | $H_{phEQ}$ |
|---|---|---|---|---|
| $H_{bin}$ | **0.77/0.43** | – | – | – |
| $H_{binEQ}$ | 6.88/0 | 7.49/0 | – | – |
| $H_{phEQ}$ | 6.68/0 | 7.44/0 | **0.87/0.38** | – |
| $H_{roomEQ}$ | 3.78/0 | 4.34/0 | 5.58/0 | 5.22/0 |

图32

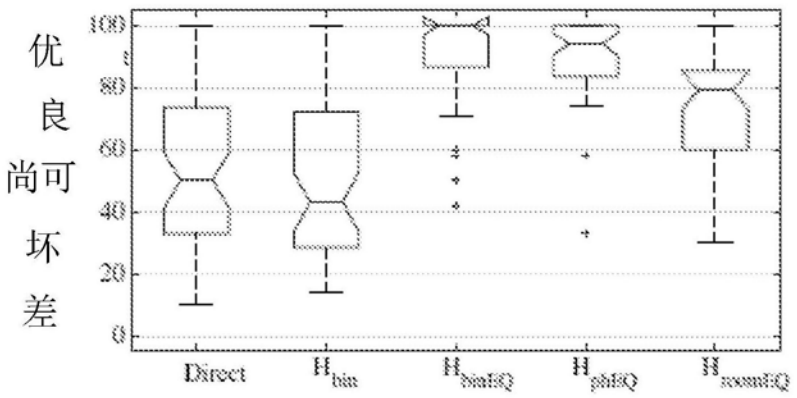



图33