



(12)发明专利

(10)授权公告号 CN 109923877 B

(45)授权公告日 2020.08.25

(21)申请号 201680090777.9

(22)申请日 2016.11.11

(65)同一申请的已公布的文献号
申请公布号 CN 109923877 A

(43)申请公布日 2019.06.21

(85)PCT国际申请进入国家阶段日
2019.05.10

(86)PCT国际申请的申请数据
PCT/EP2016/077376 2016.11.11

(87)PCT国际申请的公布数据
W02018/086701 EN 2018.05.17

(73)专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 金文字 彼得·格罗舍

(74)专利代理机构 广州三环专利商标代理有限公司 44202

代理人 熊永强 李稷芳

(51)Int.Cl.
H04S 1/00(2006.01)
H04S 7/00(2006.01)

(56)对比文件
US 5305386 A,1994.04.19
EP 1696702 A1,2006.08.30
US 2010290643 A1,2010.11.18

审查员 文蕊佳

权利要求书2页 说明书10页 附图6页

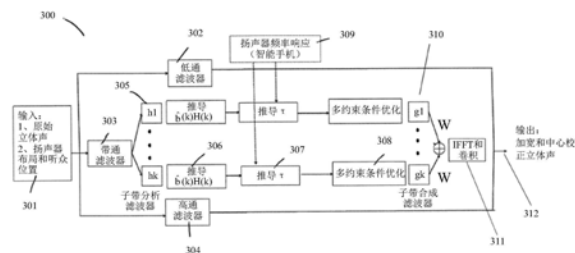
(54)发明名称

对立体声音频信号进行加权的装置和方法

(57)摘要

一种信号发生器具具有一个滤波器组,用于:接收至少两个音频信号,向所述音频信号施加权重,并且将所述音频信号的加权版提供给至少两个扬声器。所述信号发生器中的所述滤波器组用于施加权重,所述权重通过识别第一约束条件来推导出,所述第一约束条件限制可以向将要提供给第一扬声器的音频信号施加的权重。还确定第二扬声器的特征,所述特征影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号。基于所述确定的特征和所述第一约束条件来确定第二约束条件。然后确定所述权重,以便使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化。向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定。向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。所述信号发

生器可以同时实现最佳点校正和音场加宽。还通过施加权重来实现平衡音场,所述权重基于影响现实生活中的扬声器的约束条件来确定。还通过考虑各扬声器的约束条件如何影响用户对所述扬声器输出的音频信号的感知来进一步增强所述平衡音场,尤其是在这些扬声器具有某种形式的不对称排列的情况下。



1. 一种信号发生器,其特征在于,包括:

滤波器组,用于:接收至少两个音频信号,向所述音频信号施加权重,并且将所述音频信号的加权版提供给至少两个扬声器;

其中,由所述滤波器组向所述音频信号施加的所述权重通过以下方式推导出:

识别第一约束条件,所述第一约束条件限制可以向提供给第一扬声器的音频信号施加的权重;

确定第二扬声器的特征,所述特征将影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号;

基于所述确定的特征和所述第一约束条件来确定第二约束条件;以及

确定所述权重,以便使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化,其中,向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定,向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。

2. 根据权利要求1所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过以下方式推导出:

根据所述第二扬声器的所述特征来确定立体声平衡的衰减因子;以及

根据所述衰减因子来确定所述第一约束条件。

3. 根据权利要求1所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过以下方式推导出:当所述第一和第二扬声器与用户的距离不同时,确定所述特征为所述第二扬声器距离所述用户相比于所述第一扬声器距离所述用户的相对距离。

4. 根据权利要求3所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过将所述相对距离确定为 $\tau(k) = \frac{d_1^2}{d_2^2}$ 来推导出,其中, d_1 为所述第二扬声器与所述用户之间的距离, d_2 为所述第一扬声器与所述用户之间的距离, k 为频率指数。

5. 根据权利要求1至4任意一项所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过以下方式推导出:当所述第一和第二扬声器具有不同频率响应时,将所述特征确定为所述第二扬声器相比于所述第一扬声器的相对频率响应。

6. 根据权利要求5所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过将所述相对频率响应确定为 $\tau(k) = \frac{|t_1(k)|^2}{|t_2(k)|^2}$ 来推导出,其中, $t_1(k)$ 为所述第二扬声器的脉冲响应, $t_2(k)$ 为所述第一扬声器的脉冲响应, k 为频率指数。

7. 根据权利要求1至4任意一项所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过确定所述第一约束条件为与两个或更多扬声器相关联的最大增益来推导出。

8. 根据权利要求7所述的信号发生器,其特征在于,对于所述信号发生器用于将所述音频信号提供给汽车内至少两个扬声器的情况,所述第一约束条件为与距离所述用户较远的扬声器相关联的最大增益。

9. 根据权利要求1至4任意一项所述的信号发生器,其特征在于,由所述滤波器组施加的所述权重通过以下方式推导出:确定所述权重,使得向将要提供给其中一个所述扬声器

的所述音频信号施加的所述权重的平方和不超过所述扬声器的约束条件。

10. 根据权利要求1至4任意一项所述的信号发生器,其特征在於,由所述滤波器组施加的所述权重通过以下方式推导出:根据所述两个或更多扬声器相对于用户的物理排列来确定所述目标平衡。

11. 根据权利要求1至4任意一项所述的信号发生器,其特征在於,由所述滤波器组施加的所述权重通过以下方式推导出:确定所述目标平衡,以便模拟相对于所述用户对称排列的扬声器。

12. 根据权利要求1至4任意一项所述的信号发生器,其特征在於,由所述滤波器组施加的所述权重通过以下方式推导出:确定所述目标平衡,以便模拟比所述两个或更多扬声器远的扬声器。

13. 一种音频信号处理方法,其特征在於,包括:

接收至少两个音频信号,向所述音频信号施加权重,并且将所述音频信号的加权版提供给至少两个扬声器;

其中,向所述音频信号施加的所述权重通过以下方式推导出:

识别第一约束条件,所述第一约束条件限制可以向将要提供给第一扬声器的音频信号施加的权重;

确定第二扬声器的特征,所述特征将影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号;

基于所述确定的特征和所述第一约束条件来确定第二约束条件;以及

确定所述权重,以便使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化,其中,向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定,向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。

14. 一种非瞬时性机器可读存储介质,其特征在於,其上存储有处理器可执行指令,用于控制计算机执行以下操作:

接收至少两个音频信号,向所述音频信号施加权重,并且将所述音频信号的加权版提供给至少两个扬声器;

其中,向所述音频信号施加的所述权重通过以下方式推导出:

识别第一约束条件,所述第一约束条件限制可以向将要提供给第一扬声器的音频信号施加的权重;

确定第二扬声器的特征,所述特征将影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号;

基于所述确定的特征和所述第一约束条件来确定第二约束条件;以及

确定所述权重,以便使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化,其中,向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定,向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。

对立体声音频信号进行加权的装置和方法

技术领域

[0001] 本发明涉及一种对音频信号进行加权的装置和方法,从而在用户听到这些音频信号时获得期望的音频效果。

背景技术

[0002] 立体声播放常用于娱乐系统中,通过两个或更多独立的音频通道来再现声音,从而产生从不同方向听到声音的感觉,如同自然听觉。立体声优选通过相对于用户对称放置的一对立体声扬声器进行播放。然而,现实中不可避免地会遇到不对称或不平衡的立体声扬声器。例如,汽车内相对于驾驶员位置的立体声结构和小型移动设备上的不平衡扬声器设置。不对称扬声器设置不会产生良好的空间效果。这是因为,如果听众不在最佳点上,那么立体声图像就会被破坏。作为响应,许多声音图像都集中在最近的扬声器位置。这导致声场分布狭窄且空间效果不佳。

[0003] 不对称扬声器排列的一个常见示例出现在智能手机等移动设备中。在移动设备中配备立体声扬声器越来越受欢迎。然而,由于硬件存在约束条件(例如尺寸、电池),很难嵌入一双对称扬声器,特别是对智能手机而言。一种方案是使用嵌入式耳机接收器作为扬声器单元。然而,接收器和扬声器的频率响应必然会有所不同(例如,由于挡板大小不同),这导致立体声效果差并且立体声图像不平衡。接收器/扬声器响应的均衡化可以解决立体声图像不平衡的问题,但不能实现音场加宽。

[0004] 一种创建加宽音场的选择是实现支持串音消除的虚拟声源渲染。先前的研究探究了通过‘非常规’扬声器排列进行虚拟声源渲染的可能性(例如,参见Glasgel在第123次AES大会上发表的“通过4.x RACE处理进行的360度定位(360localisation via 4.x RACE processing)”以及Kahana等人在关于空间声音再现的第16次国际会议上发表的“关于汽车内饰中虚拟声源合成的实验(Experiments on the synthesis of virtual acoustic sources in automotive interiors)”)。此研究仅限于单一虚拟声源的渲染,不考虑平衡立体音场的优化。此外,这两种方法都只考虑几何不对称的情况,无法减少由于其它不对称而造成的差异,例如两个扬声器的自然频率响应差异。因此,这些方法无法优化智能手机上的不对称扬声器设置,而且播放质量差(包括滤波器设计中的重要预回声),并且声场加宽效果的鲁棒性受到限制,特别是在汽车环境恶劣时。本发明的目标是提供用于通过不平衡扬声器设置改进音频信号播放的概念。

发明内容

[0005] 前述和其它目标通过独立权利要求的特征来实现。其它实施形式从从属权利要求、描述内容和附图中显而易见。

[0006] 根据第一方面,提供了一种信号发生器。所述信号发生器包括滤波器组,用于:接收至少两个音频信号,向所述音频信号施加权重,并且将所述音频信号的加权版提供给至少两个扬声器。所述滤波器组可以对所述信号进行加权,使得当所述扬声器输出所述加权

信号时,模拟扬声器之间的距离与它们的实际距离不同的效果。所述信号发生器中的所述滤波器组用于施加通过识别第一约束条件来推导出的权重,所述第一约束条件限制可以向将要提供给第一扬声器的音频信号施加的权重。还确定第二扬声器的特征,所述特征影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号。基于所述确定的特征和所述第一约束条件来确定第二约束条件。然后确定所述权重,以便使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化。向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定。向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。所述信号发生器可以同时实现最佳点校正和音场加宽。还通过施加权重来实现平衡音场,所述权重基于影响现实生活中的扬声器的约束条件来确定。还通过考虑各扬声器的约束条件如何影响用户对所述扬声器输出的音频信号的感知来进一步增强所述平衡音场,尤其是在这些扬声器具有某种形式的不对称排列的情况下。这种不对称性可能是由于扬声器的物理排列(例如,一个扬声器可能比另一个扬声器距离用户远,例如在车内)或者由于扬声器具有不同的脉冲响应(移动设备通常是这种情况)。

[0007] 在所述第一方面的第一实施形式中,由所述滤波器组施加的所述权重可以通过以下方式推导出:根据所述第二扬声器的所述特征来确定立体声平衡的衰减因子;以及根据所述衰减因子来确定所述第一约束条件。所述衰减因子获得了不对称扬声器排列对用户如何感知这些相应扬声器的约束条件产生的影响。因此,根据所述衰减因子来推导所述滤波器权重改进了所得音场的平衡。

[0008] 在所述第一方面的第二实施形式中,上述实施形式中的由所述滤波器组施加的所述权重可以通过以下方式推导出:当所述第一和第二扬声器与用户的距离不同时,确定所述特征为所述第二扬声器距离所述用户相比于所述第一扬声器距离所述用户的相对距离。这解决了立体声扬声器排列中常见的不对称性之一:扬声器相对于用户的物理排列中的不对称,意味着来自一个扬声器的音频信号必须比来自另一扬声器的音频信号传播得更远以到达用户。

[0009] 在所述第一方面的第三实施形式中,由所述滤波器组施加的所述第二实施形式中的所述权重可以通过将所述相对距离确定为 $(k) = \frac{d_1^2}{d_2^2}$ 来推导出,其中, d_1 为所述第二扬声器与所述用户之间的距离, d_2 为所述第一扬声器与所述用户之间的距离, k 为频率指数。这获得了使扬声器与用户的距离不同能够对收听音频信号的用户如何感知约束条件产生的影响,从而能够补偿该影响。

[0010] 在所述第一方面的第四实施形式中,由所述滤波器组施加的任意上述实施形式中的所述权重可以通过以下方式推导出:当所述第一和第二扬声器具有不同频率响应时,将所述特征确定为所述第二扬声器相比于所述第一扬声器的相对频率响应。这解决了立体声扬声器排列中的另一种常见不对称性:扬声器频率响应中的不对称性意味着音频信号的特定频带可以由每个扬声器有区别地放大。

[0011] 在所述第一方面的第五实施形式中,由所述滤波器组施加的所述第四实施形式中的所述权重可以通过将所述相对频率响应确定为 $\tau(k) = \frac{|t_1(k)|^2}{|t_2(k)|^2}$ 来推导出,其中, $t_1(k)$ 为所述

第二扬声器的脉冲响应, $t_2(k)$ 为所述第一扬声器的脉冲响应, k 为频率指数。这获得了使扬声器具有不同频率响应能够对收听音频信号的用户如何感知约束条件产生的影响, 从而能够补偿该影响。

[0012] 在所述第一方面的第六实施形式中, 由所述滤波器组施加的任意上述实施形式中的所述权重通过确定所述第一约束条件为与两个或更多扬声器相关联的最大增益来推导出。这限制了所述权重, 从而实际上可以实现通过所述扬声器播放所得音频信号。

[0013] 在所述第一方面的第七实施形式中, 对于所述信号发生器用于将所述音频信号提供给汽车内至少两个扬声器的情况, 所述第六实施形式中的所述第一约束条件可以为与距离所述用户较远的扬声器相关联的最大增益。这说明了以下事实: 来自较远扬声器的音频信号必须传播得更远才能到达用户, 因此, 如果用户需要感知这些音频信号与来自另一扬声器的音频信号具有相同的音量, 通常必须在播放时放大更多。

[0014] 在所述第一方面的第八实施形式中, 由所述滤波器组施加的任意上述实施形式中的所述权重可以通过以下方式推导出: 确定所述权重, 使得向将要提供给其中一个所述扬声器的所述音频信号施加的所述权重的平方和不超过所述扬声器的约束条件。这有助于确保所推导出的权重不超过实际上可以在现实扬声器排列中实现的权重。

[0015] 在所述第一方面的第九实施形式中, 由所述滤波器组施加的任意上述实施形式中的所述权重可以通过以下方式推导出: 根据所述两个或更多扬声器相对于用户的物理排列来确定所述目标平衡。这使得所述滤波器权重能够补偿所述扬声器的物理排列中的不对称性。

[0016] 在所述第一方面的第十实施形式中, 由所述滤波器组施加的任意上述实施形式中的所述权重可以通过以下方式推导出: 确定所述目标平衡, 以便模拟相对于所述用户对称排列的扬声器。用户可以由用户头部模型表示, 目标平衡可以旨在再现围绕所述头部模型的对称虚拟扬声器排列。这使得所述权重能够在用户处产生平衡音场的效果。

[0017] 在所述第一方面的第十一实施形式中, 由所述滤波器组施加的任意上述实施形式中的所述权重可以通过以下方式推导出: 确定所述目标平衡, 以便模拟比所述两个或更多扬声器更远的扬声器。这具有加宽音场的效果。

[0018] 根据第二方面, 提供了一种方法, 包括: 接收至少两个音频信号, 向所述音频信号施加权重, 以及将所述音频信号的加权版提供给至少两个扬声器。向所述音频信号施加的所述权重通过以下方式推导出: 识别第一约束条件, 所述第一约束条件限制可以向将要提供给第一扬声器的音频信号施加的权重。还确定第二扬声器的特征, 所述特征影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号。基于所述确定的特征和所述第一约束条件来确定第二约束条件。然后确定所述权重, 以便使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化。向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定。向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。

[0019] 根据第三方面, 提供了一种非瞬时性机器可读存储介质, 其上存储有处理器可执行指令, 用于控制计算机来实现一种方法, 所述方法包括: 接收至少两个音频信号, 向所述音频信号施加权重, 以及将所述音频信号的加权版提供给至少两个扬声器。向所述音频信

号施加的所述权重通过以下方式推导出：识别第一约束条件，所述第一约束条件限制可以向将要提供给第一扬声器的音频信号施加的权重。还确定第二扬声器的特征，所述特征影响用户相对于所述第一扬声器输出的音频信号如何感知所述第二扬声器输出的音频信号。基于所述确定的特征和所述第一约束条件来确定第二约束条件。然后确定所述权重，以使当所述扬声器输出所述加权信号时预期用户会听到的每个信号的实际平衡与目标平衡之间的差异最小化。向将要提供给所述第一扬声器的音频信号施加的所述权重进一步根据所述第一约束条件来确定。向将要提供给所述第二扬声器的音频信号施加的所述权重进一步根据所述第二约束条件来确定。

附图说明

- [0020] 现将参考附图通过示例对本发明进行描述。在附图中：
- [0021] 图1示出了根据本发明一个实施例的信号发生器；
- [0022] 图2为汽车内传统立体声结构与音场扩展之间的比较；
- [0023] 图3示出了一种用于推导权重以向音频信号施加的信号结构；
- [0024] 图4示出了听众和不对称扬声器排列的示例；
- [0025] 图5示出了听众和实现平衡扬声器设置的虚拟加宽扬声器排列的示例；
- [0026] 图6示出了一种用于推导权重以向音频信号施加的方法的示例；以及
- [0027] 图7示出了将使用根据传统串音算法推导出的权重的滤波器与使用通过多约束条件优化推导出的权重的滤波器进行比较的模拟结果。

具体实施方式

[0028] 图1示出了信号发生器的示例。信号发生器100包括用于接收两个或更多音频信号的输入端101。这些音频信号表示立体声系统的不同通道，因此可用于不同扬声器。信号发生器包括可选变换单元102，用于通过对每个音频信号进行傅里叶变换来将该信号分解成其相应频率分量。在其它实施方式中，滤波器组103可以对所需的音频信号全部进行分割。该滤波器组包括多个独立滤波器104。每个独立滤波器可以用于对音频信号的特定频带进行滤波。这些滤波器可以是带通滤波器。每个滤波器可以用于向音频信号施加权重。这些权重通常是预先计算好的，对每个频带分别施加一个权重。预先计算的权重优选通过在下文更详细描述的多约束条件优化技术来推导出。该技术适于推导能够实现不对称扬声器排列的音场平衡的权重。由于一个扬声器与一个扬声器的距离比与另一个扬声器的距离远（例如汽车内），所以扬声器排列可以是不对称的。由于一个扬声器与另一扬声器具有不同的脉冲响应（例如智能手机场景中），所以扬声器排列可以是不对称的。信号发生器100用于同时实现音场加宽和最佳点校正。

[0029] 在一些实施例中，信号发生器可以包括用于存储多组不同滤波器权重的数据存储单元105。每个滤波器组可以适用于不同场景。滤波器组可以用于根据用户输入和/或内部或外部生成的建议特定场景适用的观察结果来使用一组滤波器权重。例如，当信号发生器正在将音频信号提供给汽车内的立体声系统时，用户通常可能想要为驾驶员优化音场，但也可以为其中一名乘客优化音场。这可能是用户可以通过与汽车立体声系统相关联的用户界面进行选择的一个选项。在另一个示例中，实现音场优化的合适权重可能取决于如何使用

智能手机等移动设备。例如,如果设备的传感器指示其水平放置在平面上或者如果传感器输出指示设备垂直放置并且可能靠近用户的脸部,则不同权重可能都是合适的。

[0030] 在许多实施方式中,信号发生器很可能是较大设备的一部分。例如,该设备可以是手机、智能手机、平板电脑、笔记本电脑、立体声系统或任何通用用户设备,尤其是具有音频播放功能的用户设备。

[0031] 图1所示的结构(以及本文包含的所有装置框图)意在对应多个功能块。这仅用于说明目的。图1并非意在定义芯片上硬件的不同部分之间或者软件中不同程序、过程或功能之间的严格划分。在一些实施例中,本文描述的部分或所有信号处理技术很有可能全部或部分在硬件中执行。这尤其适用于包含重复操作的技术,例如傅里叶变换和滤波。在一些实施方式中,至少一些功能块很可能全部或部分由在软件控制下运行的处理器实施。任何此类软件都可以存储在非瞬时性机器可读存储介质上。处理器可以是例如手机、智能手机、立体声系统或任何具有音频播放功能的通用用户设备中的DSP。

[0032] 不对称扬声器排列的一个常见示例出现在汽车内。这是音场加宽尤其有益的场景。图2 示出了汽车内传统立体声结构与音场扩展之间的比较。对于传统立体声设置(201),对于所有乘客,特别是对于驾驶员,所生成的声场分布很狭窄并且不理想,这是因为收听位置偏离中心。扬声器放置受约束,导致固定设置不灵活。一种选择是采用基于延迟和增益调整的最佳点校正方法(202)。这为各个收听位置(例如驾驶员的收听位置)重定义了立体音场。这时,系统具有非常窄的音场,不能产生令人满意的空间效果。一个优选选择是通过创建与另一个扬声器的距离远大于与实际扬声器的距离的“虚拟扬声器”来加宽音场(203)。在图2 中示为位于汽车外部的虚拟扬声器,表示听众感受到的声音加宽效果。

[0033] 图3示出了用于确定滤波器权重的系统结构的示例,这些滤波器权重可以用于处理图2 所示的不平衡扬声器排列的类型。该系统结构包括旨在模拟当扬声器输出立体声音频信号时发生在这些立体声音频信号上的情况的功能块,还包括用于计算滤波器权重的功能块,这些滤波器权重可以为不对称扬声器排列重新平衡立体音场。下面结合图6所示的过滤器权重生成过程对这些功能块进行更详细的描述。在大多数实际实施方式中,预期预先计算滤波器权重并将这些滤波器权重存储在信号发生器100的滤波器组103中。

[0034] 系统结构原始左右立体声信号具有作为其输入301。这些信号是扬声器输出的音频信号。系统结构描述如下,具体参考涉及两个音频信号的示例:一个用于左手扬声器,一个用于右手扬声器,但是下面描述的技术可以很容易地扩展到两个以上的音频通道。

[0035] 功能块302至305主要用于模拟当输入音频信号301由扬声器输出并在空气中传播以使听众听到时所发生的情况。预期会绕过非常低和非常高的频率,这在图3的系统结构中通过低通滤波器302和高通滤波器304表示。这一假设是合适的,这是由于大多数场景(例如汽车场景和智能手机场景)中的设备尺寸有限以及预计大多数实施中只需要两个扬声器这一事实。合适的低截止频率和高截止频率约分别为300Hz和7kHz。带通滤波器303将音频信号分割成子带,并执行快速傅里叶变换。这将为下一阶段的合成过程准备音频信号,在合成过程中,由于传递函数存在频率依赖性,所以音频信号的不同频带在空气中传播时会有效地受制于这些不同传递函数。子带分析滤波器305表示当音频信号从扬声器传播到听众耳朵时应用到这些音频信号的传递函数。这在图4中示出。

[0036] 从扬声器到听众耳朵的声音传播的频率相关传递函数 $h_{m1}(k)$ 根据扬声器的位置和

听众耳朵的位置来确定。这在图4中示出,图4示出了听众401相对于左右扬声器402、403不对称放置。标记m标识听众的耳朵(例如,m=1表示左耳,m=2表示右耳),而标记l标识扬声器(例如l=1表示左扬声器,l=2表示右扬声器)。传递函数 $h_{ml}(k)$ (其中, $m,l \in \{1;2\}$)可以按 2×2 矩阵 $H(k)$ 排列。矩阵 $H(k)$ 也称为工厂矩阵。

$$[0037] \quad \mathbf{H}(k) = \begin{bmatrix} \mathbf{h}_{11}(k) & \mathbf{h}_{12}(k) \\ \mathbf{h}_{21}(k) & \mathbf{h}_{22}(k) \end{bmatrix} \quad (1)$$

[0038] $h_{11}(k)$ 、 $h_{12}(k)$ 、 $h_{21}(k)$ 、 $h_{22}(k)$ 可以通过球形头部模型基于相应扬声器和听众位置来确定。

[0039] 在图3的系统中,子带分析滤波器之后是系数推导单元306、约束条件推导单元307和多约束条件优化单元308。这些功能单元一起工作以确定合适的滤波器权重来解决不对称扬声器设置。加强滤波器权重确定的理论概述如下。

[0040] 对于每个频率点k,有可能利用两个(也可能两个以上)约束条件来确切地表达优化。这种确切表达开始于表示 2×2 维度的扬声器权重矩阵:

$$[0041] \quad \mathbf{W}(k) = \begin{bmatrix} w_{11}(k) & w_{12}(k) \\ w_{21}(k) & w_{22}(k) \end{bmatrix} \quad (2)$$

[0042] $W(k)$ 的对角元素表示左立体声通道和右立体声通道的同侧滤波器增益。非对角元素表示两个通道的对侧滤波器增益。这些增益特定于频率点,所以该矩阵位于频域中。

[0043] 立体声信号的短时间傅里叶变换(short-time Fourier transform,STFT)系数可表示为 $s_n(k)$ ($n \in \{1,2\}$),其中n为通道索引。通过将音频信号划分为长度相等的短片段,然后分别对每个短片段计算FFT,可以计算出STFT系数。因此,STFT系数具有振幅和时间延长。左通道 $n=1$,而右通道 $n=2$ 。因此,驱动第1个扬声器的播放信号可以写成:

$$[0044] \quad x_l(k) = \sum_{n=1}^2 w_{ln}(k) s_n(k) \quad (3)$$

[0045] 其中, $l \in \{1,2\}$ 。这表示被带通滤波到各个频率点的音频信号,每个频率点在播放之前分别进行加权。

[0046] 参考图4所示的两个扬声器相对于用户的物理排列,可以看出,对于频率点k,到达耳朵m的音频信号由下式给出:

$$[0047] \quad y_m(k) = \sum_{l=1}^2 h_{ml}(k) \sum_{n=1}^2 w_{ln}(k) s_n(k) \quad (4)$$

[0048] 其中 $m \in \{1;2\}$ 。

[0049] 因此,向扬声器输出的音频信号施加的权重与通过球形头部模型确定的传递函数相结合,从而形成响应系数 $b_{mn}(k)$:

$$[0050] \quad b_{mn}(k) = \sum_{l=1}^2 h_{ml}(k) \sum_{n=1}^2 w_{ln}(k) \quad (5)$$

[0051] 响应系数将左右通道信号 $s_1(k)$ 和 $s_2(k)$ 变换成听众感知到的信号 $y_m(k)$ ($m \in \{1;2\}$)。原则上,权重 $w_{ln}(k)$ 可以自由选择。传递函数 $h_{ml}(k)$ 由系统中的几何结构来固定。

[0052] 目的是为实际设置选择权重 $w_{ln}(k)$,以使所得响应系数 $b_{mn}(k)$ 与期望虚拟设置的

响应系数相同或至少相近。

$$[0053] \quad \hat{b}_{mn}(k) = \sum_{l=1}^2 \hat{h}_{ml}(k) \sum_{n=1}^2 \hat{w}_{ln}(k) \quad (6)$$

[0054] 与虚拟设置相关联的 (2×2) 矩阵 $\hat{\mathbf{b}}(k) = [\hat{b}_{mn}(k)]$ 表示在听众耳朵处观察到的期望频率响应。优选地选择目标矩阵 $\hat{\mathbf{b}}(k)$ ，以使所得滤波器显示最小的预回声，从而使得播放质量良好、声音加宽感知更佳。

[0055] 期望虚拟设置是一种假想的设置，在这种设置中，在音场加宽和良好播放质量方面，两个扬声器的放置比实际设置中更有利。图5示出了期望虚拟设置的示例。该图示出了汽车场景，其中两个真实扬声器501、502相对于用户不对称排列。在期望设置中，两个虚拟扬声器503、504相对于用户（该示例中为汽车驾驶员）对称排列。在图5的示例中，两个虚拟扬声器中的一个与实际系统中的距离远的扬声器重合（这是实际设置中的右侧扬声器（1=2））。

[0056] 对于汽车场景，其中两个扬声器通常相对于驾驶员不对称放置，通常希望物理加宽其中至少一个扬声器。参考图4所示的两个扬声器相对于用户的物理排列，图5的汽车场景中的 $\hat{\mathbf{b}}(k)$ 矩阵的第一列表示期望的左手虚拟扬声器的频率响应。这个期望扬声器与右手物理扬声器对称。右手扬声器距离驾驶员相对较远，因此足够宽。图5的汽车场景中的 $\hat{\mathbf{b}}(k)$ 矩阵的第二列表示期望的右手虚拟扬声器的频率响应。右手虚拟扬声器可以放置在右手物理扬声器附近，优选地位于完全相同的位置。理想的排列是模拟扬声器情况如下的扬声器排列：(i) 相对于用户对称排列；以及 (ii) 提供宽音场。

[0057] 对于智能手机场景，两个扬声器通常相对于用户对称放置。在这个场景下， $\hat{\mathbf{b}}(k)$ 矩阵的第一和第二列可以表示一对对称的左右虚拟扬声器的频率响应，其中那些虚拟声源的空间间隔比物理扬声器宽。智能手机场景下的不对称性与扬声器的频率响应有关，而不是与它们的物理排列有关。这两个物理扬声器很可能具有不同的频率响应。

[0058] 返回图3的系统结构，确定一组合适的滤波器权重的第一阶段是：系数推导单元306确定物理扬声器排列的工厂矩阵 $\mathbf{H}(k)$ 和一组理想响应系数 $\hat{\mathbf{b}}(k)$ 。这还可以通过图6中的步骤 S601和S602来表示。

[0059] 一种选择是：系统一旦确定好工厂矩阵和一组理想响应系数（例如通过等式(6)），系统就直接确定滤波器权重。然而，这不是最优的，因为没有考虑物理扬声器排列中固有的一个或多个约束条件，而且会影响用户如何感知不同扬声器输出的音频信号。具体而言，可能存在物理约束条件，限制在将音频信号提供给物理扬声器之前可以向这些音频信号施加的权重。一种这样的约束条件与特定扬声器的增益上限相关联。这种约束条件可以表示为 N 。

[0060] 在图3的系统结构中，约束条件推导单元307用于确定限制可以向用于由特定扬声器播放的音频信号施加的权重的约束条件（步骤S603）。对于含两个扬声器的排列，这些约束条件可以表示为第一约束条件 N_1 和第二约束条件 N_2 ，它们可以如下定义：

$$[0061] \quad \|\mathbf{w}(1, :) (k)\|^2 \leq N_1 \text{ 即, } \sum_{n=1}^2 |w_{1,n}(k)|^2 \leq N_1 \text{ 以及}$$

$$[0062] \quad \|w(2, :) (k)\|^2 \leq N_2, \text{ 即, } \sum_{n=1}^2 |w_{2,n}(k)|^2 \leq N_2 \quad (7)$$

[0063] 因此,每个扬声器的权重的平方和不应该超过该扬声器的约束条件。

[0064] 约束条件推导单元可以确定其中一个约束条件通过与两个扬声器都相关联的最大增益来设置。这样就为任一扬声器设置了滤波器增益的上限。例如,如果两个扬声器具有不同的增益限制,则扬声器对的上限可能是这些增益限制中的较小者。上限也可能受扬声器相对于用户的相应位置和/或其相应频率响应的影响。例如,如果两个扬声器相对于用户不对称放置,则上限可以由两个扬声器中距离较远的扬声器来确定。预期这尤其适用于向汽车内扬声器提供音频信号的情况。对于移动设备,通常情况是任一扬声器都可以提供增益上限。这在下文结合图4所示的场景更详细地描述,其中扬声器相对于用户不对称排列。

[0065] 约束条件推导单元307可以用于使用预设增益上限并且将其分配给认为上限更合适的扬声器,6dB可能是预设增益上限的一个合适示例。例如,在图4中,右手扬声器(在本示例中表示为扬声器2)距离用户较远,因此其输出的音频信号必须比左手扬声器(在本示例中表示为扬声器1)输出的音频信号响,以使用户感知音量相同的两个音频信号。因此,右手扬声器可与预设上限相关联,意味着 N_2 设为6dB。如果忽略这种约束条件,则滤波器组可能向音频信号施加不会反映在输出音频信号中的权重,因为这些权重超出了扬声器的播放能力。

[0066] 通常,同一约束条件并不会适用于所有扬声器。这可能是由于各扬声器自身之间存在固有差异和/或因为这些扬声器相对于用户进行物理排列的方式存在差异。约束条件推导单元(307)优选地用于通过确定一个扬声器的特征来解决这个问题(步骤S604),该特征影响用户如何相对于另一扬声器输出的音频信号感知该扬声器输出的音频信号。目的是创建平衡音场,其中用户感知到立体声信号如同由虚拟扬声器平等地输出。

[0067] 在一个示例中,约束条件推导单元307用于通过确定立体声平衡的衰减因子来量化另一扬声器的该特征。衰减因子表示为 $\tau(k)$,所述另一扬声器的约束条件可以确定为:

$$[0068] \quad N_1 = \tau(k) N_2 \quad (8)$$

[0069] 对于典型的汽车场景,约束条件推导单元307可以假设扬声器基本上相同,所以它们具有相同的频率响应和相同的增益限制,意味着确定用户如何感知音频信号的特征取决于每个相应扬声器与用户之间的相对距离。在这种场景下, $\tau(k)$ 可以通过基于距离的振幅平移(distance-based amplitude panning, DBAP)来推导出:

$$[0070] \quad \tau(k) = \frac{d_1^2}{d_2^2} \quad (9)$$

[0071] 在图4中, d_1 和 d_2 分别表示从左手扬声器到听众头部中心的距离以及从右手扬声器到用户头部中心的距离。

[0072] 对于典型的智能手机场景,约束条件推导单元307可以假设扬声器与用户的距离相同但是具有不同的频率响应。在这种场景下, $\tau(k)$ 可以根据左右扬声器/接收器的测量脉冲响应来推导出:

$$[0073] \quad \tau(k) = \frac{|t_1(k)|^2}{|t_2(k)|^2} \quad (10)$$

[0074] 其中, $t_1(k)$ 和 $t_2(k)$ 分别是左手和右手扬声器在频率 k 处的频率响应。

[0075] 约束条件推导单元可以具有合适的频率响应309。例如,可以基于来自加州大学戴维斯分校的CIPIC HRTF数据库来确定虚拟声源的频率响应。

[0076] 确定第二扬声器的影响用户相比于第一扬声器输出的音频信号如何感知第二扬声器输出的音频信号的特征之后,约束条件确定单元能够根据第一扬声器的约束条件和所确定的特征,通过应用等式8等来确定第二扬声器的约束条件(步骤S605)。

[0077] 在图3的系统结构中,约束条件推导单元(307)用于将约束条件输出到优化单元(308)。优化单元可以用于实施多约束条件优化,旨在当扬声器输出音频信号时预期用户会听到的每个音频信号的实际平衡与目标平衡之间的差异最小化。这可以表示为:

$$[0078] \quad \min_{\mathbf{W}(k)} \|\mathbf{H}(k)\mathbf{W}(k) - \hat{\mathbf{b}}(k)\|^2$$

[0079] 受限于:

$$[0080] \quad \|\mathbf{w}(1, :)(k)\|^2 \leq N_1, \text{ 即, } \sum_{n=1}^2 |w_{1,n}(k)|^2 \leq N_1, \text{ 以及}$$

$$[0081] \quad \|\mathbf{w}(2, :)(k)\|^2 \leq N_2, \text{ 即, } \sum_{n=1}^2 |w_{2,n}(k)|^2 \leq N_2$$

[0082] 其中, $\mathbf{H}(k)\mathbf{W}(k)$ 表示预期用户会听到的每个音频信号的实际平衡, $\hat{\mathbf{b}}(k)$ 表示目标平衡, N_1 和 N_2 限制复杂维度中的权重增益。

[0083] 如上所述,目标平衡可能旨在模拟对称扬声器排列,即物理扬声器排列,其中扬声器相对于用户对称排列(通过使用用户头部模型表示用户来实现,模拟扬声器围绕头部模型对称排列),和/或两个扬声器都显示相同频率响应的扬声器排列。目标平衡还可以旨在模拟比现实中的扬声器更远的扬声器。

[0084] 因此,优化单元308能够生成准确渲染预期虚拟声源的权重,同时还满足左通道扬声器相比于右通道扬声器的衰减约束条件。如果优化单元使用等式8,则其将在相比于复杂频域中的预期虚拟声源响应使再现误差最小化的最小均方差(minimum mean square error, MMSE)意义上求解到全局最优解,同时还通过指定的滤波器增益衰减进行有效约束。

[0085] 图3所示的系统结构还用于通过施加优化单元(308)已经确定的权重来合成将由信号发生器输出的信号。通过施加由优化单元308(如通过滤波器组310表示)生成的权重来对音频信号进行滤波。音频信号的每个频带通过该频带的合适权重 $w(k)$ 进行加权。通过变换单元311执行FFT和叠加操作来生成所得信号,从而推导出加宽和平衡立体声信号(312)。实际上,滤波器组310和变换单元311模拟同样包含在信号发生器100中的功能块,这些功能块最终施加推导出的滤波器权重,以形成音频信号通过两个或更多扬声器进行播放。

[0086] 图3所示的结构(以及本文包含的所有装置框图)意在对应于多个功能块。这仅用于说明目的。图3并非意在定义芯片上硬件的不同部分之间或者软件中不同程序、过程或功能之间的严格划分。在一些实施例中,图3的系统结构执行的部分或所有信号处理技术很有可能全部或部分在硬件中执行。这尤其适用于包含重复操作的技术,例如傅里叶变换、滤波和优化。在一些实施方式中,至少一些功能块很可能全部或部分由在软件控制下运行的处理器实施。任何此类软件都可以存储在非瞬时性机器可读存储介质上。该处理器可以是DSP等。

[0087] 图7将用于根据传统串音算法(701)对信号进行加权的滤波器的响应与用于使用根据本文描述的具有多个约束条件的优化虚拟声源渲染技术(702)推导出的权重对信号进行加权的滤波器的响应进行比较。这两种技术都用于为同一组不对称扬声器创建一对加宽虚拟声源。利用所提议的方法可以清楚地看到左通道滤波器增益的约束能量衰减(703),从而得到一个平衡立体声最佳点。此外,所提议的方法大大减少了滤波器的预回声,使得播放质量更佳、伪像更少。利用人类听众进行主观收听测试,并且利用本文所述的具有多个约束条件的优化虚拟声源渲染技术来验证虚拟声音加宽和立体声最佳点平衡的有效性。

[0088] 申请方在此单独公开本文描述的每一个体特征及两个或两个以上此类特征的任意组合。以本领域技术人员的普通知识,能够基于本说明书将此类特征或组合作为整体实现,而不考虑此类特征或特征的组合是否能解决本文所公开的任何问题;且不对权利要求书的范围造成。本申请表明本发明的各方面可由任何这类单独特征或特征的组合构成。鉴于前文描述可在本发明的范围内进行各种修改对本领域技术人员来说是显而易见的。

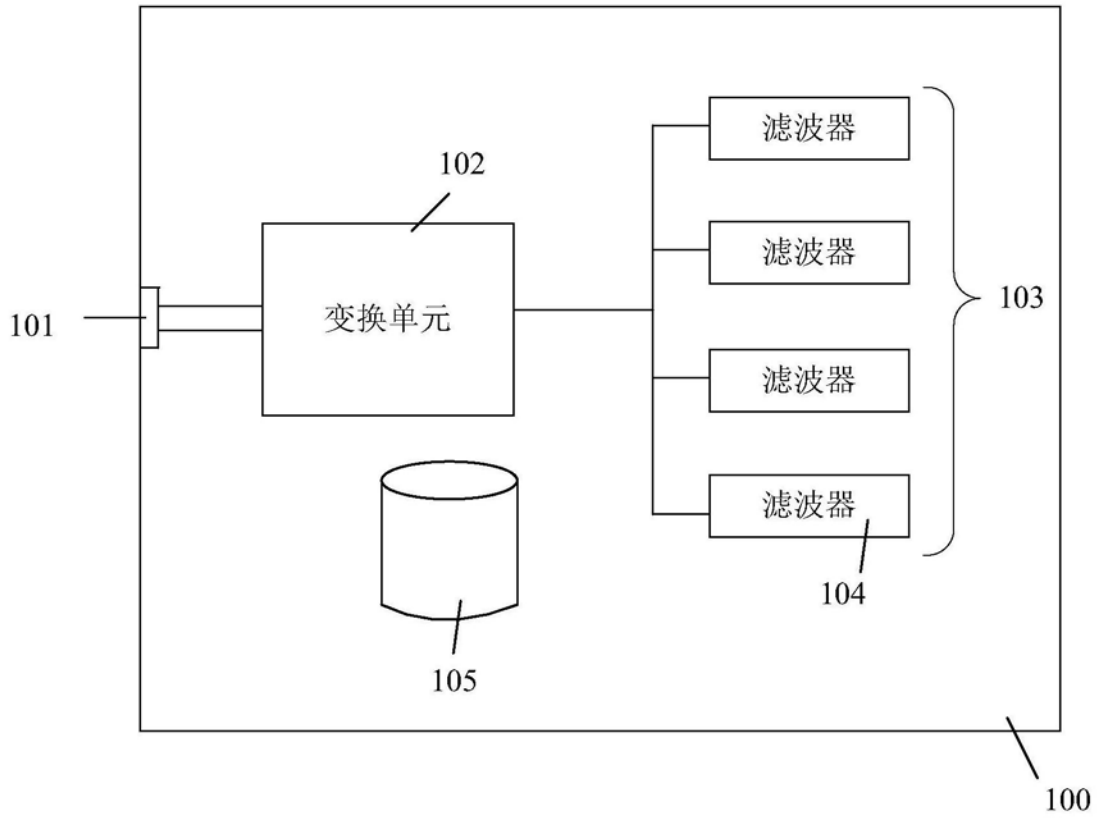


图1

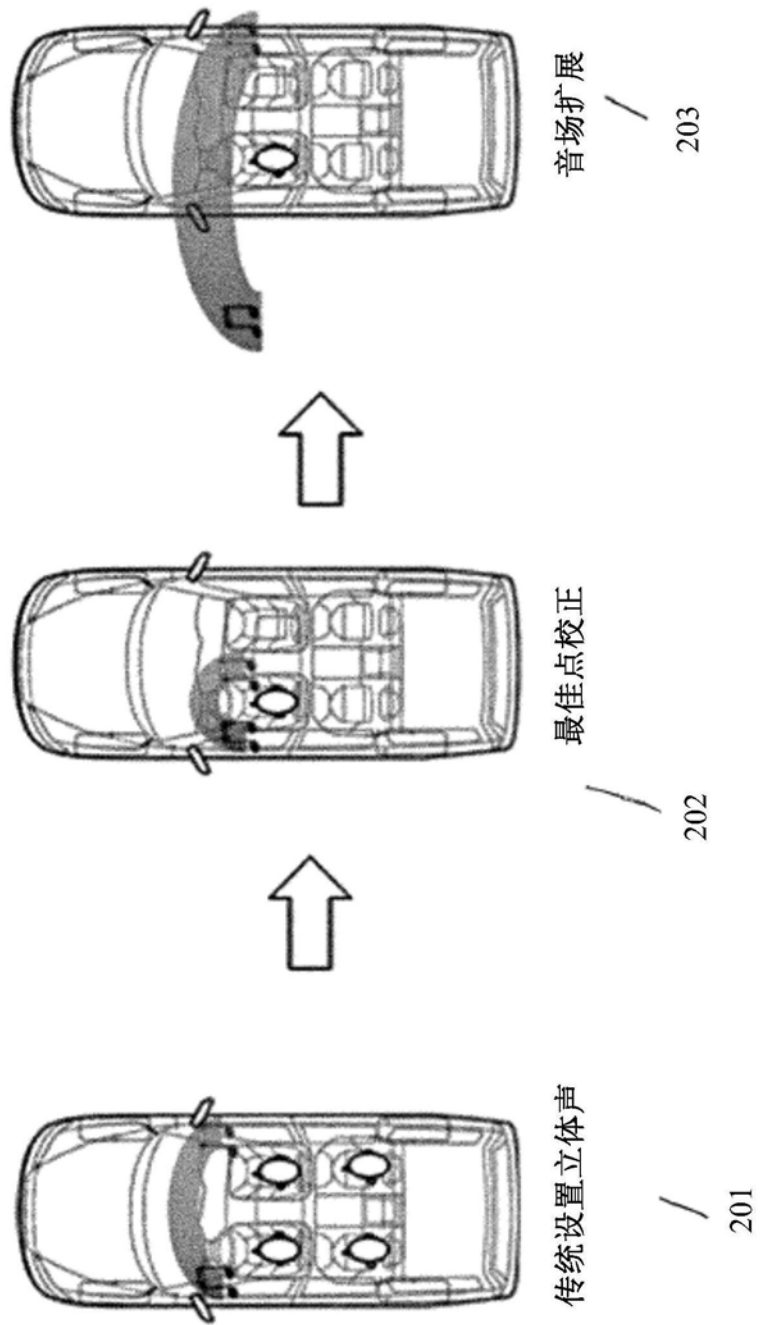


图2

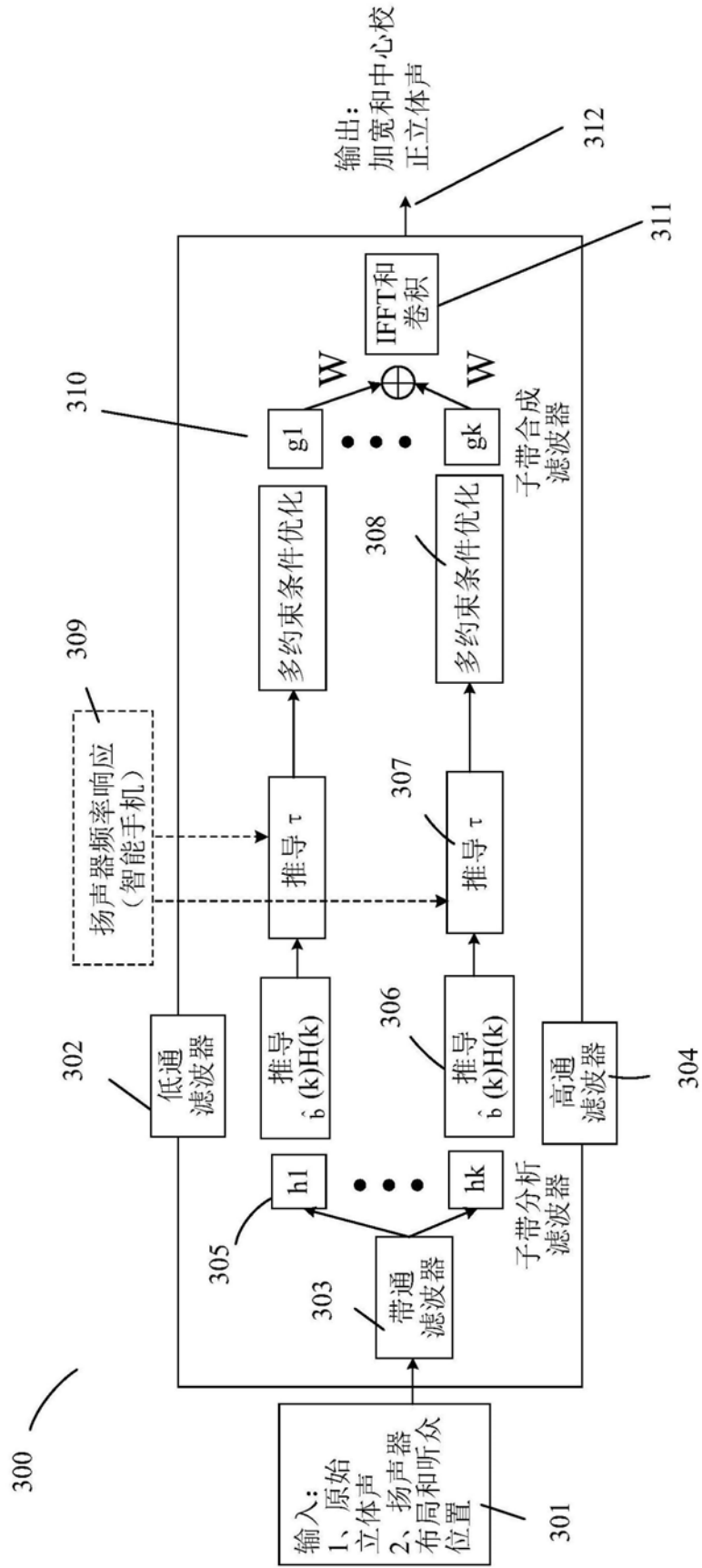


图3

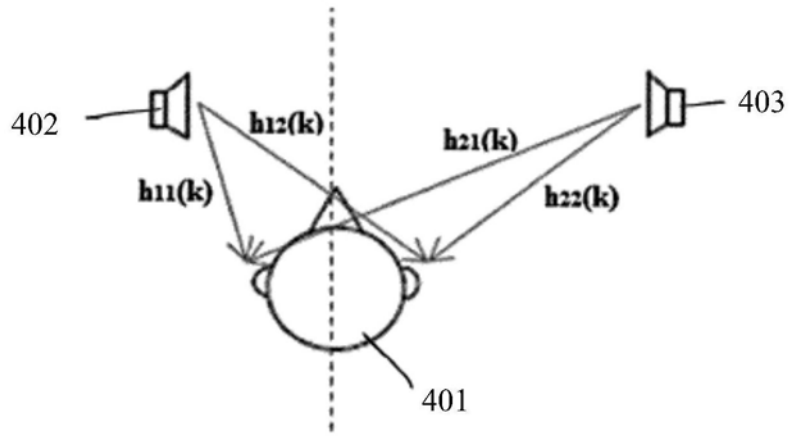


图4

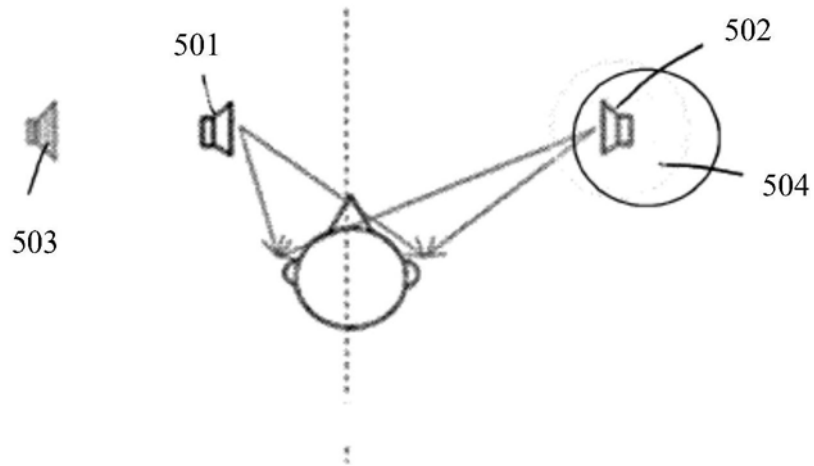


图5

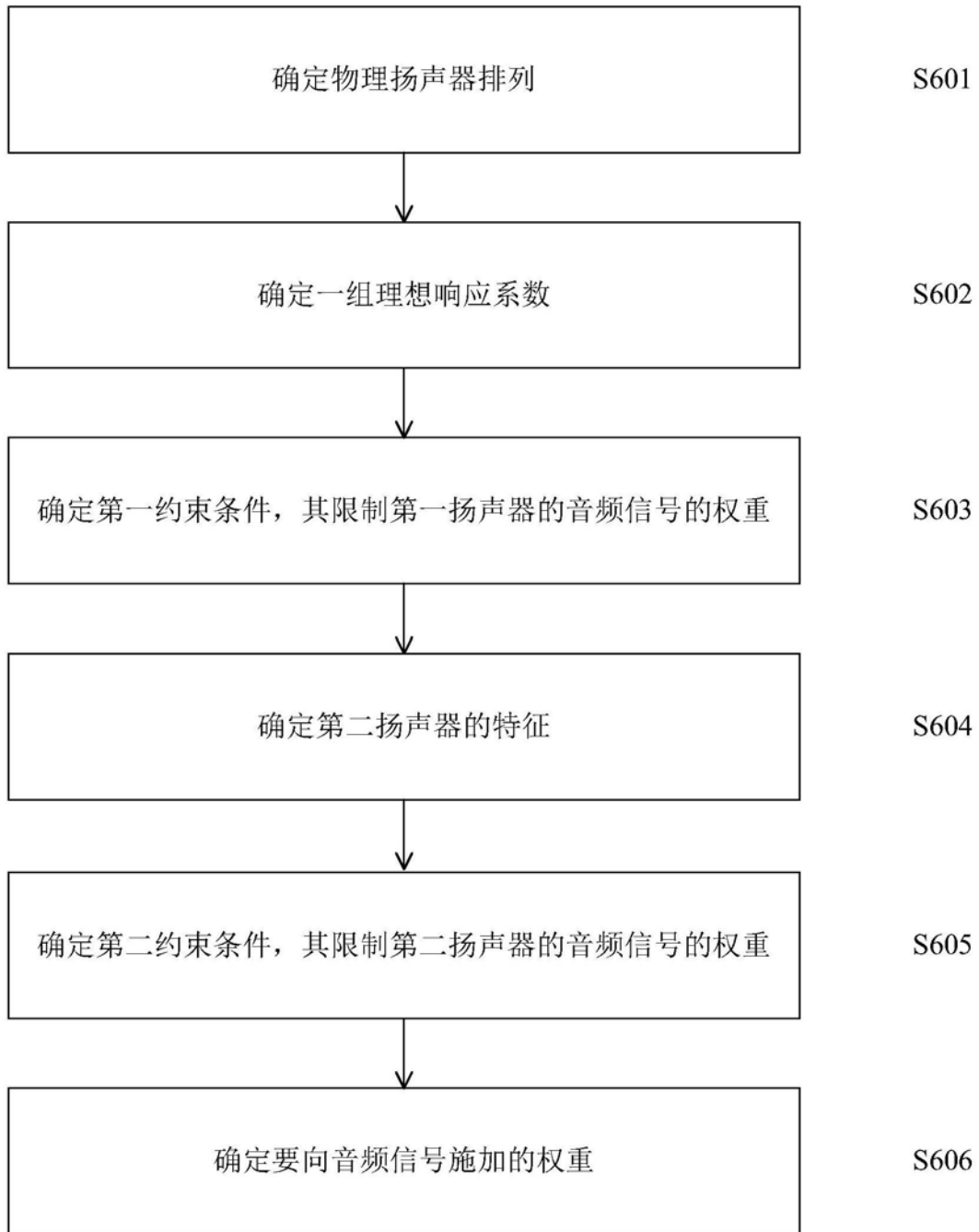


图6

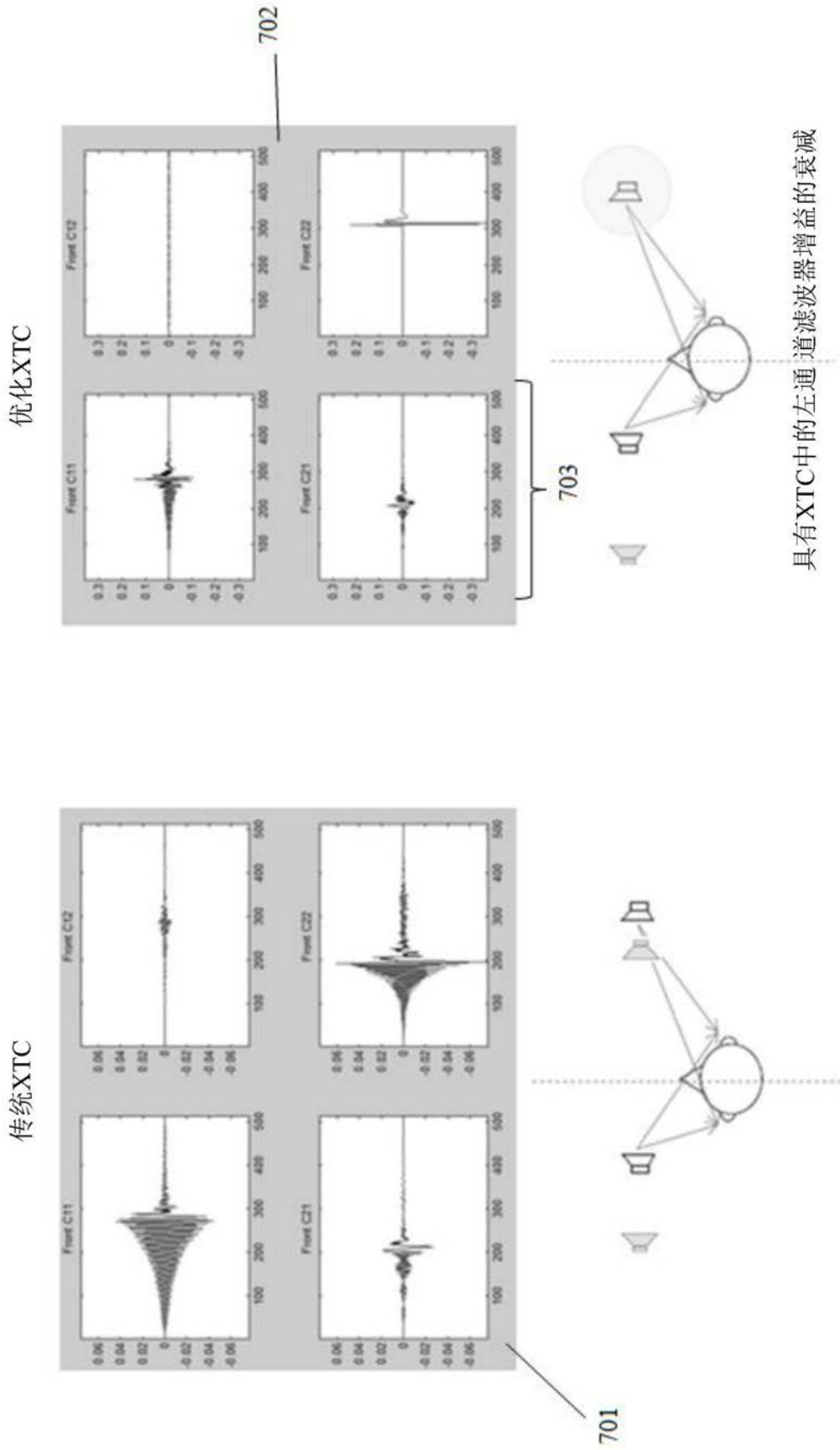


图7