



(12) 发明专利申请

(10) 申请公布号 CN 118541995 A

(43) 申请公布日 2024. 08. 23

(21) 申请号 202380016876.2

(22) 申请日 2023.01.09

(30) 优先权数据

22150868.2 2022.01.11 EP

(85) PCT国际申请进入国家阶段日

2024.07.11

(86) PCT国际申请的申请数据

PCT/EP2023/050310 2023.01.09

(87) PCT国际申请的公布数据

W02023/135075 EN 2023.07.20

(71) 申请人 皇家飞利浦有限公司

地址 荷兰艾恩德霍芬

(72) 发明人 J·G·H·科庞 S·M·杰夫斯

(74) 专利代理机构 永新专利商标代理有限公司

72002

专利代理师 刘兆君

(51) Int.Cl.

H04S 7/00 (2006.01)

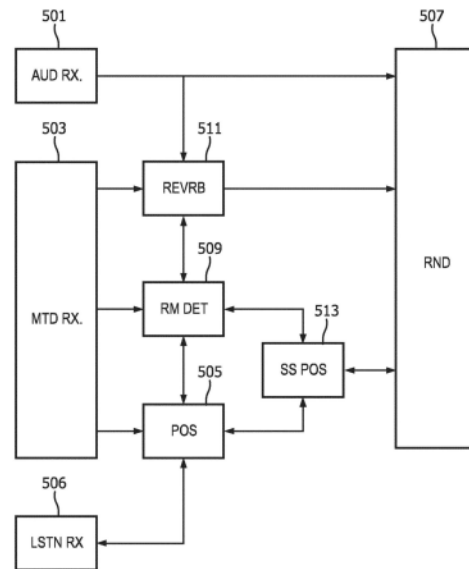
权利要求书2页 说明书23页 附图6页

(54) 发明名称

音频装置及其操作方法

(57) 摘要

一种音频装置包括第一接收器(501),所述第一接收器接收针对包括多个房间的场景中的音频源的音频数据。确定器(509)确定包括收听位置的房间和相邻房间。第二接收器(503)接收空间声学传输数据,所述空间声学传输数据描述针对所述收听房间的多个传输边界区域,所述多个传输边界区域具有超过阈值的从所述相邻房间到所述收听房间的声音的声学传输水平。第一混响器(511)确定针对所述相邻房间的混响音频信号。声源电路(513)确定针对所述传输边界区域的在所述相邻房间中的声源位置。呈现器(507)呈现针对所述收听位置的音频信号,所述音频信号包括通过从所述声源位置呈现所述第二房间混响音频信号而生成的音频分量。可以实现多房间场景的改进呈现。



1. 一种音频装置,包括:

第一接收器(501),其被布置为接收针对包括多个房间的场景中的音频源的音频数据;

位置电路(505),其被布置为确定所述场景中的收听位置;

确定器(509),其被布置为确定包括所述收听位置的第一房间和作为所述第一房间的相邻房间的第二房间;

第二接收器(503),其被布置为接收针对所述第一房间和所述第二房间的空间声学传输数据,所述空间声学传输数据描述针对所述第一房间的多个传输边界区域,每个传输边界区域的针对从所述第二房间到所述第一房间的声音的声学传输水平超过一阈值;

第一混响器(511),其被布置为根据所述第二房间中的至少一个音频源并且根据所述第二房间的几何性质和所述第二房间的声学性质中的至少一项来确定针对所述第二房间的第二房间混响音频信号;

声源电路(513),其被布置为:针对所述多个传输边界区域中的至少第一传输边界区域,确定针对音频源的在所述第二房间中的声源位置;

呈现器(507),其被布置为呈现所述收听位置的音频信号以包括通过从所述声源位置呈现所述第二房间混响音频信号而生成的第一音频分量。

2. 根据权利要求1所述的音频装置,其中,所述呈现取决于所述第一传输边界区域的几何性质和所述第一传输边界区域的声学性质中的至少一项。

3. 根据前述权利要求中的任一项所述的音频装置,其中,所述声源电路(513)被布置为受制于以下要求来确定所述声源位置:从所述声源位置到所述第一传输边界区域的距离不小于所述第一传输边界区域内的最大距离的十分之一。

4. 根据前述权利要求中的任一项所述的音频装置,其中,所述声源电路(513)被布置为受制于以下要求来确定所述声源位置:从所述声源位置到所述第一传输边界区域的距离不小于所述第二房间内的最大距离的十分之一。

5. 根据前述权利要求中的任一项所述的音频装置,其中,所述声源电路(513)被布置为受制于以下要求来确定所述声源位置:从所述声源位置到所述第一传输边界区域的距离不小于20cm。

6. 根据前述权利要求中的任一项所述的音频装置,其中,所述呈现器(507)被布置为通过沿着经过所述第一传输边界区域的声学路径从所述声源位置到所述收听位置呈现所述第二房间混响音频信号来生成所述第一音频分量。

7. 根据前述权利要求中的任一项所述的音频装置,其中,所述呈现包括通过将所述第二房间混响信号呈现为混响音频分量来生成第二音频分量。

8. 根据权利要求7所述的音频装置,其中,所述呈现包括响应于相对于所述第一传输边界区域的位置的所述收听位置而相对于所述第二音频分量的水平来调整所述第一音频分量的水平。

9. 根据权利要求8所述的音频装置,其中,所述呈现包括:针对从所述收听位置到所述第二房间的增加的距离,相对于所述第二音频分量的水平来增加所述第一音频分量的水平。

10. 根据权利要求7-9中的任一项所述的音频装置,其中,所述呈现包括响应于所述第一传输边界区域的尺寸而相对于所述第二音频分量的水平来调整所述第一音频分量的水

平。

11. 根据前述权利要求中的任一项所述的音频装置,其中,所述呈现器(507)被布置为将来自所述声源位置的所述第二房间混响音频信号呈现为空间扩展的声源。

12. 根据前述权利要求中的任一项所述的音频装置,其中,所述呈现器(506)包括:

路径呈现器(1001),其用于呈现针对声学路径的音频;

多个混响器(1003),其被布置为生成针对房间的混响信号,所述多个混响器包括被布置为生成所述第二房间混响音频信号的所述第一混响器;

耦合电路(1005),其用于将来自所述多个混响器(1003)的混响信号耦合到所述路径呈现器(1001);

组合电路(1005),其用于组合来自所述多个混响器(1003)的包括所述第二房间混响音频信号的混响信号与来自所述路径呈现器(1001)的输出信号以生成组合音频信号;以及

适配器(1007),其用于调整所述混响信号的水平以用于由所述耦合电路(1005)进行的耦合和用于由所述组合电路(1005)进行的组合。

13. 根据权利要求12所述的音频装置,其中,所述适配器(1007)被布置为响应于以下中的至少一项而调整所述混响信号的水平:

与所述音频源的所述音频数据一起接收的元数据;

所述第一传输边界区域的声学性质;

所述第一传输边界区域的几何性质;

收听位置;

从所述收听位置到所述声源位置的声学距离;以及

所述第一传输边界区域的尺寸。

14. 根据前述权利要求中的任一项所述的音频装置,其中,所述第一混响器(511)还被布置为响应于第一房间混响信号而生成所述第二房间混响信号。

15. 根据前述权利要求中的任一项所述的音频装置,其中,所述声源电路(513)被布置为基于所述第一传输边界区域的位置来生成所述声源位置。

16. 一种用于音频装置的操作的方法,所述方法包括:

接收针对包括多个房间的场景中的音频源的音频数据;

确定所述场景中的收听位置;

确定包括所述收听位置的第一房间和作为所述第一房间的相邻房间的第二房间;

接收针对所述第一房间和所述第二房间的空间声学传输数据,所述空间声学传输数据描述针对所述第一房间的多个传输边界区域,每个传输边界区域的针对从所述第二房间到所述第一房间的声音的声学传输水平超过一阈值;

根据所述第二房间中的至少一个音频源并且根据所述第二房间的几何性质和所述第二房间的声学性质中的至少一项来确定针对所述第二房间的第二房间混响音频信号;

针对所述多个传输边界区域中的至少第一传输边界区域,确定针对音频源的在所述第二房间中的声源位置;

呈现所述收听位置的音频信号以包括通过从所述声源位置呈现所述第二房间混响音频信号而生成的第一音频分量。

音频装置及其操作方法

技术领域

[0001] 本发明涉及用于生成音频信号,并且特别地但非排他地,用于针对多房间场景呈现音频来作为例如扩展现实体验的一部分的装置和方法。

背景技术

[0002] 基于视听内容的体验的种类和范围近年来已经随着利用和消耗持续开发和引入的这样的内容的新服务和方式而显著增加。特别地,许多空间和交互式服务、应用和体验正在被开发以给予用户更多参与和沉浸的体验。

[0003] 这样的应用的示例是扩展现实(XR),其是指迅速地变为主流的虚拟现实(VR)、增强现实(AR)和混合现实(MR)应用,其中,多个解决方案瞄准消费者市场。多个标准也在多个标准化主体的开发中。这样的标准化活动积极开发用于VR/AR/MR系统的各个方面的标准,包括例如流式传输、广播、呈现等。

[0004] VR应用往往提供对应于处于不同世界/环境/场景中的用户的用户体验,而AR(包括混合现实MR)应用往往提供对应于牌当前环境中的用户的用户体验,但是其中,添加了额外的信息或虚拟对象或信息。因此,VR应用往往提供完全沉浸式的合成生成的世界/场景,而AR应用往往提供叠加在用户物理存在于的真实场景的部分合成的世界/场景。然而,术语经常可交换地使用并且具有高重叠度。在下文中,术语扩展现实/XR将被用于指代虚拟现实和增强/混合现实两者。

[0005] 作为示例,日益流行的服务是以如下的方式提供图像和音频:使得用户能够主动并且动态地与系统交互以改变呈现的参数,使得这将适于用户的位置和取向的移动和改变。许多应用中的非常吸引人的特征是改变观看者的有效观看位置和观看方向的能力,例如像允许观看者在呈现的场景中移动和“四处查看”。

[0006] 这样的特征可以特别地允许虚拟现实体验被提供给用户。这可以允许用户在虚拟场景中(相对)自由地四处移动并且动态地改变他的位置以及他正在看何处。通常,这样的虚拟现实应用基于场景的三维模型,其中,模型被动态地评价以提供特定请求视图。该方法根据例如针对计算机和控制台的游戏应用(例如在第一人称射击的类别中)众所周知。

[0007] 还期望的是,特别是针对虚拟现实应用,呈现的图像是通常使用立体显示器呈现的三维图像。实际上,为了优化观看者的沉浸感,对于用户而言将呈现的场景体验为三维场景通常是优选的。实际上,虚拟现实体验应当优选地允许用户选择他/她自己的位置、视点、以及相对于虚拟世界的时刻。

[0008] 除了视觉呈现之外,大多数XR应用还提供对应的音频体验。在许多应用中,音频优选地提供空间音频体验,其中,音频源被感知为从与视觉场景中的对应对象的位置相对应的位置到达。因此,音频和视频场景优选地被感知为是一致的并且提供全空间体验。

[0009] 例如,许多沉浸式体验由通过使用双耳音频呈现技术的耳机重放生成的虚拟音频场景提供。在许多场景中,这种耳机重放可以基于头部跟踪,使得可以响应于用户的头部移动来进行呈现。这极大地增加了沉浸感。

[0010] 针对许多应用的重要特征是如何生成和/或分布可以提供音频场景的自然且逼真感知的音频。例如,当针对虚拟现实应用生成音频时,重要的是不仅生成期望的音频源,而且这些音频源还被生成为提供音频环境的真实感知,包括阻尼、反射、着色等。

[0011] 对于房间/环境声学,声波从墙壁、地板、天花板、对象等的反射引起声音源信号的延迟和衰减(通常是频率相关的)版本经由不同的路径到达收听者(即,XR系统的用户)。组合的效果可以通过脉冲响应来建模,所述脉冲响应可以被称为房间脉冲响应(RIR)。

[0012] 如图1中所图示,RIR通常包括取决于声音源到收听者的距离的直达声音,随后是表征房间的声学性质的反射部分。房间的尺寸和形状、声音源和收听者在房间中的位置以及房间表面的反射性质都在该混响部分的特性中起作用。

[0013] 反射部分可以被分解成通常交叠的两个时间区域。第一区域包含所谓的早期反射,其表示在到达收听者之前声音源在房间内部的墙壁或障碍物上的隔离反射。随着时间滞后/(传播)增加,在固定时间间隔中存在的反射的数量增加,并且路径可以包括二次或更高阶反射(例如,反射可以离开若干墙壁或墙壁和天花板两者等)。

[0014] 被称为混响部分的第二区域是这些反射的密度增加到它们不再被人脑隔离的程度的部分。该区域通常被称为漫射混响、后期混响或混响尾部,或简称为混响。

[0015] RIR包含给予听觉系统关于源的距离以及房间的大小和声学性质的信息的提示。混响部分的能量关于消声部分的能量在很大程度上决定了声音源的感知距离。最早反射的水平和延迟可以提供关于声音源与墙壁的接近程度的提示,并且通过人体测量的滤波可以加强对特定墙壁、地板或天花板的评估。

[0016] (早期)反射的密度有助于房间的感知尺寸。反射在能量水平上下降60dB所花费的时间(由混响时间 T_{60} 指示)是反射在房间中消散得有多快的常用度量。混响时间提供关于房间的声学性质的信息;例如具体地,墙壁是非常反射性的(例如,浴室)还是存在对声音的大量吸收(例如,具有家具、地毯和窗帘的卧室)。

[0017] 此外,当RIR是双耳房间脉冲响应(BRIR)的一部分时,RIR可能取决于用户的人体测量性质,这是由于RIR被头部、耳朵和肩部过滤;即头部相关的脉冲响应(HRIR)。

[0018] 由于后期混响中的反射不能被收听者区分和隔离,因此它们通常利用例如使用反馈延迟网络的参数混响器来参数化地模拟和表示,如在众所周知的Jot混响器中那样。

[0019] 对于早期反射,入射方向和距离相关延迟是人类提取关于房间和声音源的相对位置的信息的重要线索。因此,早期反射的模拟必须比后期混响更明确。因此,在有效的声学呈现算法中,早期反射与后期混响不同地且分开地被模拟。用于早期反射的众所周知的方法是对房间边界中的每一个中的声音源进行镜像,以生成表示反射的虚拟声音源。

[0020] 对于早期反射,用户和/或声音源相对于房间的边界(墙壁、天花板、地板)的位置是相关的,而对于后期混响,房间的声学响应是漫射的,并且因此在整个房间中趋于均匀。这允许后期混响的模拟通常比早期反射在计算上更有效。

[0021] 后期混响的两个主要性质是在给定阈值以上的时间内脉冲响应的斜率和幅度。在自然房间中,这些性质往往强烈地依赖于频率。通常使用表征这些性质的参数来描述混响。

[0022] 在图2中图示了表征混响的参数的示例。传统上,用于指示与漫射混响相对应的脉冲响应的斜率和幅度的参数的示例包括已知的 T_{60} 值和混响水平/能量。最近,已经提出了幅值水平的其他指示,诸如具体地指示漫射混响能量与总发射源能量之间的比率的参数。

[0023] 具体地,漫射对源比率(Diffuse to Source Ratio)DSR可以用于将用户接收的漫射混响能量或源的水平的量表示为该源的总发射能量的比率。DSR可以表示发射的源能量与漫射混响性质(诸如具体地漫射混响信号的能量或(初始)水平)之间的比率:

$$[0024] \quad DSR = \frac{\text{漫射混响能量}}{\text{总发射能量}}$$

[0025] 此后,这将被称为DSR(漫射对源比率)。

[0026] 这样的已知方法往往提供对房间中的音频传播的有效描述,并且往往导致对于收听者(虚拟地)存在于其中的房间而言被感知为自然的音频的呈现。

[0027] 然而,尽管用于在房间中表示和呈现声音的常规方法在许多实施例中可以提供合适的感知,但是它往往不适合于所有可能的场景。特别地,对于可以包括不同声学区域/房间的音频场景,使用所描述的混响方法生成的音频信号可能不会导致最佳体验或感知。它通常可能导致来自其他房间的音频没有由呈现的音频充分或准确地表示的情况,从而导致可能不完全反映声学场景和场景的感知。

[0028] 实际上,通常,考虑到房间的性质,为房间内的收听者建模混响。当收听者在房间外部或在不同的房间中时,混响器可以被关闭或者针对另一房间的性质被重新配置。即使当多个混响器可以并行运行时,混响器的输出通常是旨在在房间内呈现给收听者的漫射双耳(或多扬声器)信号。然而,这样的方法往往导致生成的音频通常不被感知为实际环境的准确表示。这可能例如导致场景的视觉感知与正在呈现的相关联的音频之间的感知到的脱钩或甚至冲突。

[0029] 因此,尽管用于呈现音频的典型方法在许多实施例中可以适合于呈现环境的音频,但是它们在一些场景中(包括特别地当为包括不同声学房间的场景呈现音频时)往往是次优的。

[0030] 因此,用于呈现场景的音频的改进方法将是有利的。特别地,允许改进的操作、增加的灵活性、降低的复杂度、方便的实施、改进的音频体验、改进的音频质量、减少的计算负担、改进的对变化位置的适合性、改进的虚拟/混合/增强现实应用的性能、增加的处理灵活性、改进的多个房间的音频和音频性质的表示和呈现、改进的多房间场景的音频呈现和/或改进的性能和/或操作的方法将是有利的。

发明内容

[0031] 因此,本发明试图优选单独地或以任何组合来减轻、缓解或消除上述缺点中的一个或多个。

[0032] 根据本发明的一个方面,提供了一种音频装置,包括:第一接收器,其被布置为接收针对包括多个房间的场景中的音频源的音频数据;位置电路,其被布置为确定所述场景中的收听位置;确定器,其被布置为确定包括所述收听位置的第一房间和作为所述第一房间的相邻房间的第二房间;第二接收器,其被布置为接收针对所述第一房间和所述第二房间的空间声学传输数据,所述空间声学传输数据描述针对所述第一房间的多个传输边界区域,每个传输边界区域的针对从所述第二房间到所述第一房间的声音的声学传输水平超过一阈值;第一混响器,其被布置为根据所述第二房间中的至少一个音频源并且根据所述第二房间的几何性质和所述第二房间的声学性质中的至少一项来确定针对所述第二房间的

第二房间混响音频信号;声源电路,其被布置为针对所述多个传输边界区域中的至少第一传输边界区域,确定所述音频源的在所述第二房间中的声源位置;呈现器,其被布置为呈现针对所述收听位置的音频信号以包括通过从所述声源位置呈现所述第二房间混响音频信号而生成的第一音频分量。

[0033] 该方法可以允许改进的用户体验,并且在许多场景中可以提供场景的音频的改进的呈现。该方法可以允许针对多房间场景的改进的音频呈现。在许多场景中可以实现场景的更自然和/或准确的音频感知。

[0034] 本发明可以提供包括混响分量的音频的改进和/或促进的呈现。音频信号的呈现通常可以以降低的复杂度和降低的计算资源要求来实现。

[0035] 该方法可以提供处理和/或呈现的音频的改进的、增加的和/或方便的灵活性和/或适应性。

[0036] 在许多实施例中,所述呈现器可以被布置为将第一音频分量呈现为局部音频源。局部音频源可以是(空间)范围局部音频源,或者可以例如是点源。

[0037] 所述第一混响器可以是漫射混响器。第一混响器可以包括(或是)参数混响器,诸如反馈延迟网络(FDN)混响器,并且具体地是Jot混响器。

[0038] 音频源可以是针对第一房间中的收听位置的在第二房间混响信号的音频源。

[0039] 声学传输水平可以是声学增益和/或透明度。

[0040] 根据本发明的可选特征,所述呈现取决于所述第一传输边界区域的几何性质和所述第一传输边界区域的声学性质中的至少一项。

[0041] 这可以在许多场景中提供改进的性能和/或方便的实施。当感知多房间场景的音频时,它可以帮助提供改进的用户体验。几何性质可以是空间性质,并且也可以被称为空间性质。

[0042] 根据本发明的可选特征,所述声源电路被布置为受制于以下要求来确定所述声源位置:从所述声源位置到所述第一传输边界区域的距离不小于所述第一传输边界区域内的最大距离的十分之一。

[0043] 这可以在许多场景中提供改进的性能和/或方便的实施。当感知多房间场景的音频时,它可以帮助提供改进的用户体验。

[0044] 根据本发明的可选特征,所述声源电路被布置为受制于以下要求来确定所述声源位置:从所述声源位置到所述第一传输边界区域的距离不小于所述第二房间内的最大距离的十分之一的。

[0045] 这可以在许多场景中提供改进的性能和/或方便的实施。当感知多房间场景的音频时,它可以帮助提供改进的用户体验。

[0046] 根据本发明的可选特征,所述声源电路被布置为受制于以下要求来确定所述声源位置:从所述声源位置到所述第一传输边界区域的距离不小于20cm。

[0047] 这可以在许多场景中提供改进的性能和/或方便的实施。当感知多房间场景的音频时,它可以帮助提供改进的用户体验。

[0048] 根据本发明的可选特征,所述呈现器被布置为通过沿着经过所述第一传输边界区域的声学路径从所述声源位置到所述收听位置呈现所述第二房间混响音频信号来生成所述第一音频分量。

[0049] 在许多实施例中,这可以允许改进的性能,并且可以允许改进的音频呈现和/或用户体验。它通常可以允许提供第二房间混响信号来自局部声源的感知的呈现。声学路径可以是直接声学路径。

[0050] 根据本发明的可选特征,所述呈现包括通过将所述第二房间混响信号呈现为混响音频分量来生成第二音频分量。

[0051] 在许多实施例中,这可以允许改进的性能,并且可以允许改进的音频呈现和/或用户体验。它可以允许提供更自然的声音音频场景的更灵活适配。混响音频分量可以是漫射音频分量。混响音频分量可以是不具有空间线索的分量。混响音频分量可以没有指示混响音频分量的空间源位置的空间线索。

[0052] 根据本发明的可选特征,所述呈现包括响应于相对于所述第一传输边界区域的位置的所述收听位置而相对于所述第二音频分量的水平调整所述第一音频分量的水平。

[0053] 在许多实施例中,这可以允许改进的性能,并且可以允许改进的音频呈现和/或用户体验。例如,它可以允许第一房间和第二房间的音频体验之间的更自然的声音灵活过渡。例如,可以提供当收听位置从在第一房间中改变为在第二房间中时的逐渐过渡。

[0054] 根据本发明的可选特征,所述呈现包括针对从所述收听位置到所述第二房间的增加了的距离,相对于所述第二音频分量的水平增加所述第一音频分量的水平。

[0055] 在许多实施例中,这可以允许改进的性能,并且可以允许改进的音频呈现和/或用户体验。

[0056] 根据本发明的可选特征,所述呈现包括响应于所述第一传输边界区域的尺寸而相对于所述第二音频分量的水平调整所述第一音频分量的水平。

[0057] 在许多实施例中,这可以允许改进的性能,并且可以允许改进的音频呈现和/或用户体验。例如,它可以允许第一房间和第二房间的音频体验之间的更自然的声音灵活过渡。例如,可以提供当收听位置从在第一房间中改变为在第二房间中时的逐渐过渡。

[0058] 在一些实施例中,所述呈现包括响应于所述第一传输边界区域的几何/空间性质而相对于所述第二音频分量的水平调整所述第一音频分量的水平。

[0059] 根据本发明的可选特征,所述呈现器被布置为将来自所述声源位置的所述第二房间混响音频信号呈现为空间扩展的声源。

[0060] 这可以允许在许多场景中呈现改进的音频场景。

[0061] 根据本发明的可选特征,所述呈现器包括:路径呈现器,其用于呈现针对声学路径的音频;多个混响器,其被布置为生成针对房间的混响信号,所述多个混响器包括被布置为生成所述第二房间混响音频信号的所述第一混响器;耦合电路,其用于将来自所述多个混响器的混响信号耦合到所述路径呈现器;组合电路,其用于组合来自所述多个混响器的包括所述第二房间混响音频信号的混响信号和来自所述路径呈现器的输出信号以生成组合音频信号;以及适配器,其用于调整所述混响信号的水平以用于由所述耦合电路进行的耦合和用于由所述组合电路进行的组合。

[0062] 这可以在许多场景中提供改进的性能和/或方便的实施。当感知多房间场景的音频时,它可以帮助提供改进的用户体验。在许多实施例中,该方法还可以允许非常有效且低复杂度的实施方式。

[0063] 根据本发明的可选特征,所述适配器被布置为响应于以下中的至少一项而调整所

述混响信号的水平；与针对所述音频源的所述音频数据一起接收的元数据；所述第一传输边界区域的声学性质；所述第一传输边界区域的几何性质；收听位置；从所述收听位置到所述声源位置的声学距离；以及所述第一传输边界区域的尺寸。

[0064] 这可以提供改进的性能，并且特别地在许多实施例中可以提供多房间音频场景的改进的和/或更灵活的和/或可适配的呈现。

[0065] 根据本发明的可选特征，所述第一混响器还被布置为响应于第一房间混响信号而生成所述第二房间混响信号。

[0066] 这可以允许在许多场景中呈现改进的音频场景。

[0067] 根据本发明的可选特征，根据前述权利要求中的任一项所述的音频装置，其中，所述声源电路被布置为基于所述第一传输边界区域的位置来生成所述声源位置。

[0068] 这可以允许在许多场景中呈现改进的音频场景。第一传输边界区域的位置可以是最近位置、中心位置和边缘位置、预定位置等。

[0069] 在一些实施例中，所述呈现包括响应于所述收听位置与所述第一传输边界区域的位置之间的距离而相对于所述第二音频分量的水平调整所述第一音频分量的水平。

[0070] 在一些实施例中，所述声音源电路可以被布置为根据所述第一传输边界区域的位置来生成所述声音源位置。

[0071] 这可以允许在许多场景中呈现改进的音频场景。第一传输边界区域的位置可以是最近位置、中心位置和边缘位置、预定位置等。

[0072] 根据本发明的一个方面，提供了一种用于音频装置的操作的方法，所述方法包括：接收针对包括多个房间的场景中的音频源的音频数据；确定所述场景中的收听位置；确定包括所述收听位置的第一房间和作为所述第一房间的相邻房间的第二房间；接收针对所述第一房间和所述第二房间的空间声学传输数据，所述空间声学传输数据描述针对所述第一房间的多个传输边界区域，每个传输边界区域的针对从所述第二房间到所述第一房间的声音的声学传输水平超过一阈值；根据所述第二房间中的至少一个音频源并且根据所述第二房间的几何性质和所述第二房间的声学性质中的至少一项来确定针对所述第二房间的第二房间混响音频信号；针对所述多个传输边界区域中的至少第一传输边界区域，确定针对音频源的在所述第二房间中的声源位置；呈现所述收听位置的音频信号以包括通过从所述声源位置呈现所述第二房间混响音频信号而生成的第一音频分量。

[0073] 本发明的这些和其他方面、特征和优点将参考下文描述的（一个或多个）实施例而显而易见并将参考下文描述的（一个或多个）实施例得到阐述。

附图说明

[0074] 现在将参考附图仅通过示例的方式描述本发明的实施例，其中，

[0075] 图1图示了房间脉冲响应的示例；

[0076] 图2图示了房间脉冲响应的示例；

[0077] 图3图示了虚拟现实系统的元件的示例；

[0078] 图4图示了具有三个房间的场景的示例；

[0079] 图5图示了根据本发明的一些实施例的用于生成音频输出的音频装置的示例；

[0080] 图6图示了呈现器的示例；

- [0081] 图7图示了根据本发明的实施例的两个房间之间的过渡区域的示例；
- [0082] 图8图示了根据本发明的实施例的两个房间之间的过渡区域的示例；
- [0083] 图9图示了根据本发明的实施例的两个房间之间的过渡区域的示例；并且
- [0084] 图10图示了根据本发明的一些实施例的用于生成音频输出的音频装置的示例。

具体实施方式

[0085] 以下描述将集中于用于扩展现实应用的音频处理和呈现,但是应当理解,所描述的原理和概念可以用于许多其他应用和实施例中。

[0086] 允许用户在虚拟世界中四处走动的虚拟体验正变得越来越流行,并且正在开发服务以满足这样的需求。

[0087] 在一些系统中,VR应用可以由例如未使用任何远程VR数据或处理或甚至不具有对任何远程VR数据或处理的任何访问的独立设备在本地提供给观看者。例如,设备(诸如游戏控制台)可以包括用于存储场景数据的存储、用于接收/生成观看者姿态的输入、以及用于根据场景数据来生成对应的图像的处理器。

[0088] 在其他系统中,VR应用可以远离观看者来实施并且执行。例如,在用户本地的设备可以检测/接收被传送到处理数据以生成观看者姿态的远程设备的移动/姿态数据。远程设备然后可以基于描述场景的场景数据来生成针对用户姿态的适合的视图图像和对应的音频信号。视图图像和对应的音频信号然后被传送到它们被呈现的在观看者本地的设备。例如,远程设备可以直接地生成由本地设备直接地呈现的视频流(通常是立体/3D视频流)和对应的音频流。因此,在这样的示例中,除传送移动数据和呈现接收到的视频数据之外,本地设备可以不执行任何VR处理。

[0089] 在许多系统中,功能可以跨本地设备和远程设备分布。例如,本地设备可以处理接收到的输入和传感器数据以生成连续地被传送到远程VR设备的用户姿态。远程VR设备可以然后生成对应的视图图像和对应的音频信号并且将这些传送到本地设备用于呈现。在其他系统中,远程VR设备可以不直接地生成视图图像和对应的音频信号,而是可以选择相关的场景数据并且将此传送到本地设备,其可以然后生成被呈现的视图图像和对应的音频信号。例如,远程VR设备可以识别最接近的捕获点并且提取对应的场景数据(例如,一组对象源及其位置元数据)并且将此传送到本地设备。本地设备可以然后处理接收到的场景数据以生成针对特定当前用户姿态的图像和音频信号。用户姿态通常对应于头部姿态,并且对用户姿态的提及通常可以被等效地认为对应于对头部姿态的提及。

[0090] 在许多应用中,尤其是对于广播服务,源可以以独立于用户姿态的场景的图像(包括视频)和音频表示的形式传送或流式传输场景数据。例如,与某个虚拟房间的范围内的音频源相对应的信号和元数据可以被传送或流式传输到多个客户端。然后,个体客户端可以在本地合成与当前用户姿态相对应的音频信号。类似地,源可以传送音频环境的一般描述,包括描述环境中的音频源和环境的声学特性。然后可以在本地生成音频表示并将其呈现给用户,例如使用双耳呈现和处理。

[0091] 图3图示了VR系统的这样的示例,其中,远程VR客户端设备301与VR服务器303例如经由网络305(例如,互联网)进行联络。服务器303可以被布置为同时支持潜在大量的客户端设备301。

[0092] VR服务器103可以例如通过传输包括图像数据形式的图像表示的图像信号来支持广播体验,客户端设备能够使用该图像信号来在本地合成与适当的用户姿态(姿态是指位置和/取向)相对应的视图图像。类似地,VR服务器303可以传送场景的音频表示,从而允许针对用户姿态在本地合成音频。具体地,当用户在虚拟环境中四处移动时,合成并呈现给用户的图像和音频被更新以反映用户在(虚拟)环境中的当前(虚拟)位置和取向。

[0093] 在许多应用(诸如图3的应用)中,因此可能期望对场景进行建模并生成可以有效地包括在数据信号中的有效图像和音频表示,该数据信号然后可以被传送或流式传输到可以针对与捕获姿态不同的姿态在本地合成视图和音频的各种设备。

[0094] 在一些实施例中,表示场景的模型可以例如在本地存储,并且可以在本地用于合成适当的图像和音频。例如,房间的音频模型可以包括可以在房间中听到的音频源的性质以及房间的声学性质的指示。然后可以使用模型数据来合成特定位置的适当音频。

[0095] 在许多场景中,场景可以包括具有不同声学性质并且具体地具有不同混响性质的若干不同声学环境或区域。具体地,场景可以包括或划分成各自具有均匀混响但混响在其间不同的不同声学环境/区域。对于声学环境/区域内的所有位置,在这些位置处接收的音频的混响分量可以是均匀的,并且具体地可以是基本上相同的(除了潜在地增益差异之外)。声学环境/区域可以是音频的混响分量均匀的一组位置。声学环境/区域可以是声学环境中的音频源的音频传播脉冲响应的混响分量是均匀的一组位置。具体地,声学环境/区域可以是除了可能的增益差异之外声学环境中的音频源的音频传播脉冲响应的混响分量具有相同的频率相关斜率和/或幅度性质的一组位置。具体地,声学环境/区域可以是除了可能的增益差异之外声学环境中的音频源的音频传播脉冲响应的混响分量是相同的一组位置。

[0096] 声学环境/区域通常可以是具有相同呈现混响参数的一组位置(通常是2D或3D区域)。用于呈现混响分量的混响参数对于声学环境/区域中的所有位置可以是相同的。特别地,相同的混响衰减参数(例如, T_{60})或漫射对源比率DSR可以应用于声学环境/区域内的所有位置。

[0097] 由于由引起混响的不同阶数的许多各种反射产生的“噪声”特性,脉冲响应在房间/声学环境/区域中的不同位置之间可能是不同的。然而,即使在这种情况下,频率相关的斜率和/或幅度性质也可以是相同的(除了可能的增益差异之外),尤其是当由例如混响时间(T_{60})或混响着色表示时。

[0098] 声学环境/区域也可被称为声学房间或简称为房间。房间可以被认为是如上所述的环境/区域。

[0099] 在许多实施例中,可以提供其中声学房间对应于用户可以(例如,虚拟地)在其间移动的不同虚拟或真实房间的场景。在图4中图示了具有三个房间A、B、C的场景的示例。在该示例中,用户可以通过门口和开口在三个房间之间移动,或在任何房间外部移动。

[0100] 对于具有基本混响性质的房间,它往往表示由具有完全或部分反射性质的几何表面充分界定的空间区域,使得该房间中的大部分反射保持反射回到该区域中,以在该区域中产生反射的漫射场,不具有显著的方向性质。几何表面不需要与任何视觉元素对齐。

[0101] 旨在向收听者提供自然且逼真效果的音频呈现通常包括声学场景的呈现。对于许多环境,这包括存在于环境中(诸如在收听者所在的房间中)的漫射混响的表示和呈现。已

经发现这种漫射混响的呈现和表示对环境的感知(诸如对音频是否被感知为表示自然且逼真的环境)具有显著影响。

[0102] 在场景包括多个房间的情况下,该方法通常仅为收听者所在的房间呈现音频和混响,并且忽略来自其他房间的任何音频。然而,这往往导致被认为不是最佳的音频体验,并且往往不提供最佳的自然体验,特别是当用户在房间之间转换时。尽管一些应用已经被实施为包括呈现来自相邻房间的音频,但是已经发现它们是次优的。

[0103] 在下文中,将描述用于呈现包含多个房间的音频场景的有利方法。为了清楚和简洁起见,将主要参考图4的示例性场景/场景来描述该方法,其中在场景中包括三个相邻房间A、B、C。

[0104] 图5图示了被布置为呈现音频场景的音频装置的示例。音频装置可以接收描述场景中的音频和音频源的音频数据。在特定示例中,音频装置可以接收图4的场景的音频数据。基于所接收的音频数据,音频装置可以呈现表示给定收听位置的场景的音频信号。呈现的音频可以包括来自在收听者所在的房间中生成的音频的贡献以及来自其他邻居(通常是相邻的)房间的贡献。

[0105] 音频装置被布置为生成表示场景中的音频的音频输出信号。具体地,音频装置可以生成表示由用户在具有多个音频源并且具有给定声学性质的场景中四处移动所感知的音频的音频。每个音频源由表示来自音频源的声音的音频信号以及可以描述音频源的特性(诸如提供音频信号的水平指示)的元数据来表示。另外,提供元数据来表征场景。

[0106] 在该示例中,呈现器是音频装置的一部分,其被布置为接收场景的音频数据和元数据,并且基于所接收的数据呈现表示环境的至少一部分的音频。

[0107] 图5的音频装置包括第一接收器501,第一接收器501被布置为接收场景中的音频源的音频数据。通常,多个例如点源可以被提供有反映要从那些音频点源呈现的声音的音频数据。在一些实施例中,还可以针对更漫射的音频源(例如背景或环境声源、或具有空间范围的声源)提供音频数据。

[0108] 音频装置包括第二接收器503,第二接收器503被布置为接收表征场景的元数据。元数据可以例如描述房间尺寸、房间的声学性质(例如, T_{60} 、DSR、材料性质)、房间之间的关系等。元数据还可以描述一些或所有音频源的位置和取向。

[0109] 元数据包括针对不同房间的空间声学传输数据。特别地,它包括描述针对场景的至少一个房间(通常是所有房间)的一个或多个传输边界区域的数据。具体地,传输边界区域可以是如下的区域,针对所述区域,从另一房间到提供传输边界区域的房间中的声音的声学传输水平超过阈值。具体地,传输边界区域可以定义两个房间之间的边界的区域(通常是面积),对于该区域,通过/跨边界的衰减小于给定阈值,而在该区域之外它可以更高。

[0110] 因此,传输边界区域可以限定两个房间之间的边界的区域,对于该区域,声传播/传输/透明度/耦合超过阈值。边界的未包括在传输边界区域中的部分可以具有高于阈值的声传播/传输/透明度/耦合。因此,传输边界区域可以限定两个房间之间的边界的区域,对于该区域,声学衰减低于阈值。边界的未被包括在传输边界区域中的部分可以具有高于阈值的声学衰减。

[0111] 因此,传输边界区域可以指示边界的区域,对于该边界的区域,声学透明度相对较高,而在区域外声学透明度可能较低。传输边界区域可以例如对应于边界中的开口。例如,

对于常规房间,传输边界区域可以例如对应于分隔两个房间的墙壁中的门口、打开的窗户或孔等。

[0112] 传输边界区域可以是三维或二维区域。在许多实施例中,房间之间的边界被表示为二维对象(例如,被认为没有厚度的墙壁),并且传输边界区域在这种情况下可以是具有低声学衰减的边界的二维形状或区域。

[0113] 声学透明度可以在标尺上表示。完全透明意味着不存在声抑制(例如,打开的门口)。部分透明度可能会对从一个房间过渡到另一个房间的能量引入衰减(例如,门口中的厚窗帘或单窗格窗)。在标尺的另一端是不允许房间之间的任何(显著的)声学泄漏(例如厚的混凝土墙)的房间分隔材料。

[0114] 因此,该方法可以(以传输边界区域的形式)提供描述两个房间如何在声学上链接的声学链接元数据。该数据可以在本地导出,或可以例如从接收的比特流获得。数据可以由内容作者手动提供,或间接地从房间的几何描述(例如,框、网格、体素化表示等)导出,包括声学性质,诸如指示多少音频能量通过材料传输或耦合到引起从一个房间到另一个房间的声学链路的材料的振动中的材料性质。

[0115] 在许多情况下,传输边界区域可以被认为指示房间泄漏,其中可以在两个房间之间交换声能。这可以是二进制指示(房间之间的边界中的开口)或可以是标量指示(反映能量的一部分被传输通过)。

[0116] 应当理解,在许多情况下,音频数据和元数据可以作为相同比特流的一部分被接收,并且第一接收器501和第二接收器503可以由相同的功能实施,并且有效地,相同的接收器功能可以实施第一接收器和第二接收器两者。图5的音频装置可以具体地对应于客户端设备或是客户端设备的一部分,并且可以在从服务器303传输的单个比特流中接收音频数据和元数据。

[0117] 该装置还包括被布置为确定场景中的收听位置的位置电路505。收听位置通常反映用户在场景中的(虚拟)位置。例如,第三接收器506可以耦合到用户跟踪装置,诸如VR头盔、眼睛跟踪装置、运动捕获相机等,并且可以由此接收用户移动(包含或可能限于头部移动和/或眼睛移动)数据。位置电路505可以从该数据连续地确定当前收听位置。

[0118] 该收听位置可以替代地由控制器输入表示或用控制器输入增强,用户可以利用控制器输入移动或瞬移场景中的收听位置。

[0119] 应当理解,许多方法和技术是已知的并且用于针对各种应用确定场景中的收听位置,并且可以使用任何合适的方法而不背离本发明。

[0120] 音频装置包括呈现器507,呈现器507被布置为生成表示收听位置处的场景的音频的音频输出信号。通常,可以生成音频信号以包括场景中的一系列不同音频源的音频分量。例如,可以将同一房间中的点音频源呈现为具有直接声学路径的点音频源,可以呈现或生成混响分量等。

[0121] 在下文中,将描述一种方法,其中,呈现的音频信号包括表示来自除了包括收听位置的房间之外的其他房间的音频的音频信号/分量。描述将集中于该音频分量的生成,但是应当理解,呈现给用户的呈现音频信号可以包括许多其他分量和音频源。这些可以根据任何合适的算法或方法来生成和处理,并且应当理解,技术人员将意识到大量这样的方法。

[0122] 以下描述将集中于生成/呈现反映除了当前包括收听位置的房间之外的一个或多

个其他房间中的音频的音频信号(分量)。

[0123] 音频装置具体地包括房间确定器509,房间确定器509被布置为确定包括收听位置的第一房间和作为相邻房间的第二房间,并且所述第二房间通常是第一房间的相邻房间。房间确定器509可以从位置电路505接收收听位置数据,并确定该收听位置的当前房间。然后,它可以继续选择当前房间的相邻房间,并且音频装置可以继续生成该相邻房间的收听位置的音频信号分量。作为具体示例,可以考虑如下场景,其中,收听位置当前在图4的房间B中并且位置电路505可以将房间A识别为相邻房间。然后,音频装置继续呈现从房间B中的收听位置听到的来自房间A的音频信号分量音频/声音。

[0124] 应当理解,对于房间C,可以遵循相同的过程,房间C也可以被识别为房间B的相邻房间。

[0125] 音频装置包括混响器511,混响器511被布置为针对所确定的相邻房间(即,在特定示例中针对房间A)来生成混响音频信号。

[0126] 房间确定器509向混响器511提供所确定的房间(即,房间A)的混响性质的信息。其可以直接地或间接地来进行此。例如,房间确定器509可以向混响器511指示所选择的相邻房间,并且然后这可以从所接收的元数据中提取所选择的房间(即,示例中的房间A)的混响参数。混响器511然后可以继续生成与相邻房间中存在的混响相对应的混响信号。

[0127] 混响器511因此继续基于相邻房间中的至少一个音频源和第二房间的至少一个性质(诸如几何性质(大小、边界/反射壁之间的距离等)或声学性质(衰减、频率响应等))来生成相邻房间的混响音频信号。

[0128] 例如,混响器511可以提取 T_{60} 或在元数据中针对邻居房间提供的DSR参数。然后,它可以继续选择相邻房间中的所有声源,并将这些声源的音频数据作为混响过程的输入提供。然后可以根据合适的混响算法生成混响信号。应当理解,许多算法和方法是已知的,并且可以使用任何合适的方法。例如,混响器511可以实施参数混响器,诸如Jot混响器。

[0129] 相邻房间混响信号与收听房间(即,包括收听位置的房间,即,在特定示例中的房间B)的音频源的音频源数据一起被馈送到呈现器507。呈现器507然后继续呈现收听位置的音频信号,除了收听房间中的音频源的分量之外,该音频信号还包括与相邻房间混响声音相对应的分量。

[0130] 相邻房间混响信号的呈现可以具体地包括将相邻房间混响信号呈现为局部音频源而不是漫射非空间背景源。因此,相邻房间混响器511不(总是)仅被呈现为漫射混响音频,而是被呈现为局部化或甚至点源。局部源可以是点源,或者可以具有范围但在空间上受约束。

[0131] 音频装置包括声源电路513,声源电路513被布置为确定相邻房间混响信号的声源位置。呈现器507然后被布置为将相邻房间混响信号呈现为来自该声源位置的局域化声源。相邻房间混响信号的呈现可以是来自声源位置的点源,或可以被呈现为响应于声源位置而被定位和/或包括声源位置的空间扩展的音频源(即,扩展音频源)。

[0132] 声源位置被具体地布置为基于传输边界区域来确定声源位置。特别地,它可以针对每个传输边界区域生成一个声源位置,并且可以针对每个声源位置/传输边界区域执行相邻房间混响信号的呈现。

[0133] 声源位置基于传输边界区域来确定,但是位于相邻房间内。因此,来自相邻房间的

混响音频也被呈现到收听房间中的收听位置,但是使得它可以被感知为源自相邻房间中的位置。因此,它可能不仅仅被感知为漫射的非空间声音,而是它可以提供空间分量。这可以在许多场景中提供对场景的更逼真的感知。

[0134] 在许多实施例中,声音源电路513可以针对传输边界区域被布置为基于传输边界区域的位置来生成声音源位置。在许多实施例中,声音源电路513可以针对传输边界区域被布置为根据传输边界区域的位置来生成声音源位置。传输边界区域的位置可以是最近位置、中心位置和边缘位置、预定位置等。

[0135] 在多房间场景中,每个房间通常具有其自己的混响特性。每个房间内的源将对它们所在的房间贡献更强,而对其他房间中的混响贡献弱得多。因此,所有房间中的所有源之间的平衡在房间之间也是不同的。

[0136] 此外,当模拟房间时,配置通常由 T_{60} 和DSR给出,但是房间尺寸也影响模式反射发生的速度和方式。

[0137] 当收听者从一个房间移动到另一个房间时,收听者所在的房间的混响可以被呈现为如本领域中已知的漫射混响。然而,根据所描述的方法,其他房间的混响可以被呈现为靠近房间之间的边界的位置上的可定位源,其中房间之间存在显著的声学透明度。这可能导致那些房间的混响仍然被感知,而不是被感知为漫射和非空间,它们被感知为在房间之间的边界的透明部分的方向上可定位。相邻房间的混响可以被感知为听到来自并通过房间之间的墙壁中的开口。此外,来自另一房间的混响的水平可以随着从收听者到混响源的距离的增加而衰减,类似于物理情况下的体验。

[0138] 声源位置被确定为在相邻房间内,即相邻房间混响信号不是从收听房间内或甚至在两个房间之间的边界处呈现的,而是从相邻房间内的位置呈现的。

[0139] 这样的方法是违反直觉的,因为房间中的音频的呈现被认为被执行以反映房间的几何和声学性质。然而,发明人已经可以认识到,可以通过确定声源位置在相邻房间内来感知改进的空间感知。特别地,已经发现,这导致在多房间场景中并且尤其是在收听位置在传输边界区域上或靠近传输边界区域的房间之间的过渡中感知到更自然的声音。

[0140] 在许多实施例中,声源位置可以被确定为在房间内给定的最小距离。最小距离可以是到最近的传输边界区域和/或到最近的边界点(即,边界上的点)的距离。最小距离可以是至少20cm,或在一些情况下,30cm、50cm或1米。最小距离可以是场景距离。在场景测量对应于现实生活距离的距离的意义上,场景通常可以对应于真实生活场景。可以参考这些来确定最小距离。

[0141] 在许多实施例中,最小距离可以是相对距离,并且具体地,最小距离可以取决于传输边界区域的尺寸。在许多实施例中,声源位置到传输边界区域的最小距离不小于传输边界区域的最大距离的十分之一。在一些实施例中,它可以不小于传输边界区域的五分之一、一半或最大距离。

[0142] 这种方法可以在许多场景中提供特别有利的操作,并且通常可以导致被感知为提供场景的自然印象的呈现。

[0143] 在一些实施例中,最小距离可以是相对于收听房间和/或相邻房间的相对距离。

[0144] 在许多实施例中,声源位置到传输边界区域的最小距离不小于收听房间和/或相邻房间的最大距离的十分之一。在一些实施例中,它可以不小于收听房间/相邻房间的五分

之一、一半或最大距离。

[0145] 这种方法可以在许多场景中提供特别有利的操作,并且通常可以导致被感知为提供场景的自然印象的呈现。

[0146] 在许多实施例中,声源位置可以被确定为在房间内给定的最大距离。最大距离可以是到最近的传输边界区域和/或到最近的边界点的距离。最大距离可以不超过1m,或在一些情况下为3m、5m或10米。最大距离可以是场景距离。在场景测量与现实生活距离相对应的距离的意义上,场景通常可以对应于现实生活场景。可以参考这些来确定最大距离。

[0147] 在许多实施例中,最大距离可以是相对距离,并且具体地,最大距离可以取决于传输边界区域的尺寸。在许多实施例中,声源位置到传输边界区域的最大距离不大于传输边界区域的最大距离的一半。在一些实施例中,它可以不小于传输边界区域的最大距离的一倍、两倍、三倍或五倍。

[0148] 这种方法可以在许多场景中提供特别有利的操作,并且通常可以导致被感知为提供场景的自然印象的呈现。

[0149] 在一些实施例中,最大距离可以是相对于收听房间和/或相邻房间的相对距离。

[0150] 在许多实施例中,针对声源位置到传输边界区域的最大距离不小于收听房间和/或相邻房间的最大距离的一半。在一些实施例中,它可以不小于收听房间/相邻房间的最大距离的五分之一或三分之一。

[0151] 这种方法可以在许多场景中提供特别有利的操作,并且通常可以导致被感知为提供场景的自然印象的呈现。

[0152] 在许多实施例中,可以基于考虑度量的组合来选择距离。例如源可以位于远离传输边界区域的最大传输边界区域的五分之一处,但是至少20cm并且至多相邻房间的最小房间尺寸的三分之一。

[0153] 表示相邻房间的混响信号的声源在相邻房间内并且具体地其靠近但不太靠近边界的定位可以提供高度有利的体验。

[0154] 声源在相邻房间中并且稍微远离边界的定位可以导致更真实的过渡,其中,定向接收的混响源自房间而不是边界。特别是在边界处,不与定向源重叠可能更现实。因此,例如,当用户从一个房间移动到相邻房间时,实现了改进的用户体验。

[0155] 特别地,所描述的方法通常可以允许在到达两个房间之间的传输边界之前从非定向的漫射混响到定向混响的基于用户位置的转换,其中,在边界处,混响基本上是定向的,源自房间。这与物理房间一致,其中,混响在没有来自传输区域的方向的混响贡献的这些边界处漫射小得多。

[0156] 如果源位置恰好在边界上并且对收听位置的音频信号有贡献,则其定位将不现实,因为它与收听位置交叠,并且甚至可能被感知为源自错误的房间。它可能对收听者跨边界移动并导致定位从收听者的一侧转到另一侧非常敏感。

[0157] 此外,通常使用非零参考距离来呈现可定位源,对于该非零参考距离,针对所述信号不需要距离衰减。在源在距边界的一定距离处的情况下,使得其距离衰减对于边界周围的收听位置和进入收听房间中更现实地操作。

[0158] 另外,在源被定位在相邻房间内的情况下,由直接声学路径呈现器(如下所述)进行的其呈现可以方便地对例如在传输边界区域中完全或部分关闭的门的遮挡和/或衍射进

行建模。在源在边界上的情况下,存在源未被门几何形状遮挡的风险。

[0159] 当从相邻房间内的位置呈现混响信号时,呈现器507被布置为呈现混响信号,使得它包括针对声源位置的一些空间线索。具体地,呈现包括呈现从声源位置到收听位置的声学路径,其中声学路径经过第一传输边界区域。声学路径可以是声源位置到收听位置的直接声学路径,或可以是反射声学路径。这种反射声路径通常可以包括不超过一次、两次、三次或五次反射。反射可以例如离开收听房间的墙壁或边界。

[0160] 图6图示了呈现器507的元件的示例。在该示例中,呈现器600包括用于每个音频源的路径呈现器601。每个路径呈现器601被布置为生成表示从音频源到收听者的直达路径的直达路径信号分量。直达路径信号分量是基于收听者和音频源的位置生成的,并且可以具体地通过根据距离和例如音频源在到用户的特定方向上的相对增益(例如,对于非全向源)针对音频源潜在地频率相关性地缩放音频信号来生成直达信号分量。

[0161] 在许多实施例中,呈现器601还可以基于在源位置与用户位置之间的遮挡或衍射(虚拟)元素来生成直达路径信号。

[0162] 在许多实施例中,路径呈现器601还可以针对个体路径生成另外的信号分量,其中,这些信号分量包括一个或多个反射。这可以例如通过评估墙壁、天花板等的反射来完成,如本领域技术人员将已知的。直达路径和反射路径分量可以被组合成针对每个路径呈现器的单个输出信号,并且因此可以针对每个音频源生成表示直达路径和早期/离散反射的单个信号。

[0163] 在一些实施例中,每个音频源的输出音频信号可以是双耳信号,并且因此每个输出信号可以包括左耳和右耳(子)信号两者,以包括其相对于用户取向的方向的定向呈现(例如,通过应用头部相关传递函数(HRTF)双耳房间脉冲响应(BRIR)或扬声器平移算法)。

[0164] 来自路径呈现器601的输出信号被提供给组合器603,该组合器603组合来自不同路径呈现器601的信号以生成单个组合信号。在许多实施例中,可以生成双耳输出信号,并且组合器可以执行来自路径呈现器601的个体信号的组合(诸如加权组合),即,可以将来自路径呈现器601的所有右耳信号加在一起以生成组合的右耳信号,并且可以将来自路径呈现器601的所有左耳信号加在一起以生成组合的左耳信号。

[0165] 路径呈现器601和组合器603可以以任何合适的方式实施,通常包括作为用于在合适的计算资源上(诸如微控制器、微处理器、数字信号处理器或包括诸如存储器等的支持电路的中央处理单元等)处理的可执行代码。应当理解,多个路径呈现器可以被实施为并行功能单元(诸如例如一组专用处理单元),或可以被实施为针对每个音频源的重复操作。通常,针对每个音频源/信号执行相同的算法/代码。

[0166] 除了个体路径音频分量之外,呈现器507还被布置为生成表示环境中的漫射混响的信号分量。在特定示例中,通过将源信号组合成下混信号并且然后将混响算法应用于下混信号以生成漫射混响信号来生成漫射混响信号。

[0167] 图6的音频装置包括下混器605,下混器605接收多个声音源(通常是混响器模拟漫射混响的声学环境内的所有源)的音频信号和用于将音频信号组合成下混的元数据(元数据可以例如由内容创建者提供作为视听数据流的一部分)。下混器将音频信号组合成相应地反映环境中生成的所有声音的下混。针对个体音频信号的系数/权重可以例如被设置为反映对应声源的(相对)水平,并且可选地与DSR组合以控制混响的水平。

[0168] 在许多实施例中,位于由混响器建模的房间外部的源也可以有助于混响。然而,这些源通常可以比房间内的源贡献少得多,因为这些外部源中的仅一部分通过任何传输边界区域到达房间。

[0169] 下混被馈送到混响呈现器/混响器607,混响呈现器/混响器607被布置为基于下混生成漫射混响信号。混响器607可以具体地是参数混响器,诸如Jot混响器。混响器607被耦合到组合器603,漫射混响信号被馈送到组合器603。然后,组合器603继续将漫射混响信号与表示个体路径的路径信号组合,以生成表示收听者感知的环境中的组合声音的组合音频信号。

[0170] 在该示例中,针对收听房间中的音频源的所有音频信号被馈送到路径呈现器,并且呈现器507继续生成包括来自所有这些的贡献的输出信号,包括对应于直接路径、反射路径和漫射混响的贡献。

[0171] 然而,另外,混响器511的输出(即相邻房间的混响信号)也可以被馈送到路径呈现器601。因此,用于在收听房间内呈现音频源也可以用于位于相邻房间中的相邻房间混响信号。

[0172] 在大多数情况下,混响信号也被馈送到混响器607,并且因此也从相邻房间的混响声音提供对收听房间中的漫射声音的贡献。

[0173] 类似地,在一些情况下,也可以基于收听房间的混响信号来生成相邻房间的混响信号。例如,混响器607可以被布置为生成混响信号,该混响信号不包括来自相邻房间混响信号(而是例如仅来自收听房间本身内的声源)的任何贡献。替代地,所生成的收听房间混响信号可以作为输入馈送到混响器511,并且可以有助于所生成的相邻房间混响信号。在许多场景中,这样的方法可以提供对场景的自然音频的改进且更准确的呈现。

[0174] 在许多实施例中,呈现器507被布置为将相邻房间混响信号呈现为来自声源位置的点源信号。然而,在其他实施例中,呈现器507可以被布置为将相邻房间混响信号呈现为空间扩展的音频源。因此,在一些实施例中,相邻房间混响信号可以被呈现为具有范围的音频源。

[0175] 例如,声源的空间扩展可以由声源电路513确定。作为示例,声音源的范围可以取决于传输边界区域的尺寸来确定。作为具体示例,声音源可以被确定为具有与传输边界区域的尺寸匹配的空间范围。

[0176] 然后,呈现器507可以继续呈现相邻房间混响信号,使得它被感知为具有与所确定的延伸匹配的空间延伸。

[0177] 应当理解,用于呈现具有空间范围的音频源的各种方法是已知的,并且可以使用任何合适的方法。

[0178] 在一些具体实施例中,呈现器507可以被布置为通过将音频源呈现为分布在音频源的范围内的多个点源来呈现范围音频源。例如,可以确定用于呈现相邻房间混响信号的范围以及分布在该范围内的相对大数量(例如10-50个)的点源。然后可以从每个点源呈现相邻房间混响信号,导致具有空间范围的单个音频源的整体感知。

[0179] 用与其他信号去相关的信号呈现范围中的每个点源通常有利于现实地生成感知范围。为此,可以使用去相关器。替代地,当使用反馈延迟网络(FDN)混响器时,可以用一组相互正交的提取矢量来从反馈回路中提取信号,以获得具有每个提取矢量的去相关混响信

号。例如,可以使用Gram-Schmidt过程导出一组正交向量。

[0180] 在许多实施例中,并且特别是对于大的传输边界区域,呈现具有一定范围的音频源可以提供改进的用户体验,并且特别是来自相邻房间的混响的更逼真和自然的发声贡献。

[0181] 在许多实施例中,音频装置可以被布置为适配混响信号的水平或增益。

[0182] 具体地,房间内的源将其全部发射的能量贡献给房间,并且因此贡献给混响。在许多实施例中,对于这样的源,源能量确定这些源在该房间的下混频器中被下混的相对水平。

[0183] 基于源性质,可以计算归一化源能量缩放因子,其指示将声源的信号转换成其对应的总发射能量的缩放因子。这些归一化的源能量缩放因子可以在呈现器507的下混频器605中使用,以获得表示源的总发射能量的下混信号。

[0184] 还应当认识到,许多实施例可以使用基于在距源的标称距离处的标称增益(方向性模式的平均值,并且包括诸如预增益和距离衰减增益的其他适用增益)的系数,其中混响能量数据(DSR)也以与在距源的该标称距离处的采样相对应的源能量而不是全部发射能量来表示。本领域技术人员将能够将基于全发射能量的示例和实施例转换为该替代源能量表示方案。

[0185] 对于所考虑的房间外部的源,并非所有能量都贡献于房间的混响。源在不同的房间或区域中发射其所有能量,并且该能量的一部分可能泄漏到所考虑的房间中。其能量泄漏到房间中的部分取决于若干因素,包括:

- [0186] • 源到房间泄漏的距离。
- [0187] • 另一房间中的障碍物对源的遮挡和衍射影响从源到房间泄漏的路径。
- [0188] • 房间泄漏的数量和大小。
- [0189] • 房间泄漏施加的衰减。

[0190] 在具有降低的复杂度的一些实施例中,能量分数可以基于已经针对收听者计算的(潜在频率相关的)增益。也就是说,收听者在房间内,并且其他房间中的源的直接路径呈现可能已经考虑了距离、遮挡和衍射,因此为从源到房间泄漏的路径提供了良好的近似。

[0191] 这种方法可能不是完全准确的,因为它还可能包括遮挡的衰减和/或在所考虑的房间内行进的距离。然而,它不需要附加计算。如果以算法知道每个房间施加哪些因素的这种方式确定收听者的增益,则仅从所考虑的房间之外的因素收集增益也可以是该过程的一部分,并且具有非常少的附加计算。

[0192] 在这些实施例中,该增益仅是能量缩放的一部分。它不考虑房间泄漏/传输边界区域的尺寸。它们通常确实包括由(至少一个)传输边界区域施加的衰减。例如,当传输边界区域是 2m^2 的门口时,进入房间的能量比当它是 0.25m^2 的小窗户时多得多。

[0193] 表示源和某个位置(即,收听者的头部或通过传输边界区域进入房间之后的右侧)之间的所有衰减的增益可以有利地被定义为对应于人耳的表面区域。这通常与使用HRTF对的连续呈现一致。因此,所获得的增益可以被平方并乘以表面比,以获得引入到房间中的能量的近似值,从而产生与所考虑的源相关联的信号*i*的下混系数:

$$[0194] \quad d_i = \sqrt{S_i \cdot g_i^2 \cdot \frac{A_{leak}}{A_{ear}}}$$

[0195] 可能存在源能量通过其到达所考虑的房间的多个房间泄漏,这些可以被聚合。例

如,利用以下等式:

$$[0196] \quad d_i = \sqrt{\sum_{j=0}^{J-1} S_i \cdot g_{i,j}^2 \cdot (t_j + c_j)^2 \cdot \frac{A_{leak,j}}{A_{gain}}}$$

[0197] 其中, d_i 表示信号*i*的下混系数, S_i 表示归一化的源能量比例因子, $g_{i,j}$ 表示施加在与信号*i*相关联的源到房间泄漏*j*的路径上的(潜在频率相关的)衰减增益, t_j 表示传输系数,并且 c_j 表示是房间泄漏*j*的耦合系数, $A_{leak,j}$ 表示是房间泄漏*j*的表面积,并且 A_{ear} 表示与增益相关联的表面积(例如人耳)。

[0198] 如本领域已知的,可以以不同的方式计算增益 $g_{i,j}$,模拟用于直接路径呈现的距离、遮挡和衍射。

[0199] 低复杂度方法的示例可以仅聚焦于从源到房间泄漏的直接路径距离衰减。

$$[0200] \quad g_{i,j} = \frac{d_{ref}}{d_{i,j}}$$

[0201] 其中, $d_{i,j}$ 是从与信号*i*相关联的源的位置到与房间泄漏*j*相关联的位置的距离,并且 d_{ref} 是信号/源的参考距离,其中,信号上的距离衰减等于1。

[0202] 如前所述,音频装置可以生成第一音频分量,该第一音频分量对应于来自声源位置(作为点源或作为范围源)的混响信号的局部化呈现。

[0203] 在一些实施例中,音频装置还可以被布置为通过将相邻房间混响信号呈现为收听房间的混响信号来生成第二音频分量。因此,在一些实施例中,相邻房间中的混响在收听房间中呈现为局部声音和漫射混响声音的组合。这样的方法可以例如在许多实施例中提供场景的更逼真的体验。

[0204] 在一些实施例中,混响信号因此可以被馈送到呈现器507的路径呈现器601,以导致相邻房间混响信号的空间局部化呈现。另外,相邻房间混响信号可以被馈送到组合器603并且与为收听房间本身生成的混响信号(以及与来自路径呈现器601的输出)组合。

[0205] 呈现器507可以包括用于呈现到收听位置的声学路径传播的路径呈现器,并且呈现器可以被布置为将相邻房间混响信号馈送到路径呈现器。呈现器507还可以被布置为将相邻房间混响信号与(一个或多个)路径呈现器的输出组合。

[0206] 在这种情况下,呈现器507可以被布置为适配两个音频分量的相对水平。

[0207] 在许多实施例中,呈现包括根据相对于第一传输边界区域的收听位置相对于第二音频分量(反映漫射和非局部混响)的水平调整第一音频分量(反映局部音频源)的水平。具体地,呈现器507可以被布置为针对从收听位置到传输边界区域/相邻房间的增加了的距离相对于第二音频分量的水平增加第一音频分量的水平。因此,收听者朝向传输边界区域和相邻房间移动得越近,相对于漫射声音贡献的局部声音的感知越强。

[0208] 在一些实施例中,呈现器可以被布置为根据传输边界区域的几何性质并且具体地根据传输边界区域的尺寸来相对于第二音频分量(反映漫射和非局部混响)的水平调整第一音频分量(反映局部音频源)的水平。

[0209] 在许多实施例中,呈现器507可以被布置为针对传输边界区的增加的尺寸而相对于第二音频分量的水平降低第一音频分量的水平。因此,传输边界区域越大,相对于漫射声音贡献的局部声音的感知越弱。

[0210] 水平适配可以例如用于在两个房间之间生成逐渐过渡。例如,当用户在房间之间移动时,通常可以实现音频从一个房间到另一个房间的更平滑和更自然的过渡。例如,可以为收听位置定义过渡或交叉衰落区域,其中局部和非局部(漫射)分量的加权根据区域内的收听位置动态地调整。

[0211] 图7-9图示了示例性传输边界区域的声源位置701和交叉衰落/过渡区域703的示例,其中收听房间由B表示,并且相邻房间由A表示。

[0212] 在图7中,声源是点源,并且过渡区域是边界开口周围的区域(由传输边界区域表示)。在图8的示例中,相邻房间混响信号的声源是广度声源,并且在图9的示例中,过渡区域仅形成在相邻房间中。

[0213] 在这样的示例中,两个分量的相对水平可以在过渡区域上逐渐变化,以提供平滑的交叉衰落过渡。

[0214] 如图10所示,在一些实施例中,音频装置可以包括路径呈现器1001,路径呈现器1001被布置为呈现音频源的声学路径。路径呈现器1001可以具体地实施图6的路径呈现器601。

[0215] 音频装置还可以包括被布置为生成房间的混响信号的多个混响器1003。混响器1003可以具体地包括混响器511以及下混器605和混响器607,并且因此可以分别生成用于相邻房间和收听房间的混响信号。混响器1003可以包括用于生成针对其他房间混响信号的混响器。

[0216] 音频装置还可以包括耦合电路1005,耦合电路1005被布置为选择性地将来自多个呈现器1003的输出的混响信号耦合到路径呈现器1001。因此,耦合电路1005能够将混响信号(诸如相邻房间混响信号)耦合到路径呈现器1001的输入,使得信号可以被呈现为局域化信号。

[0217] 音频装置还包括组合电路1005,组合电路1005被布置为选择性地将来自呈现器1001的混响信号彼此组合,并与来自路径呈现器1001的输出信号组合,并且直接与混响信号组合。结果是表示场景中的音频的音频信号。组合电路1005可以包括组合器603。

[0218] 在该示例中,耦合电路和组合电路1005由可以在路径呈现器1001的输入和组合器功能之间切换混响器1003的输出的开关实施。然而,应当理解,在许多实施例中,可以使用可以适配耦合到路径呈现器的输入之间的相对增益的单独增益。

[0219] 音频装置还包括适配器1007,适配器1007被布置为调整混响信号的水平以用于耦合和组合。例如,适配器1007可以控制图10的开关,或可以例如控制和调整从混响器1003到直接路径呈现器1001的输入和输出侧的路径的增益。

[0220] 该布置允许混响信号被适配并被呈现为局部化源和/或漫射混响信号。它提供了一种可以以低复杂度实施同时提供高性能和相当大的灵活性的非常有效的方法。

[0221] 适配器1007可以具体地根据以下中的一个或多个来适配分别针对直达路径呈现器输入和组合的混响信号的水平:

[0222] a) 与音频源的音频数据一起接收的元数据。例如,由内容创建者导出的重要性函数可以增加或减少超过(物理)模拟水平的不同房间的相对水平。

[0223] b) 所述第一传输边界区域的声学性质。例如,第一传输边界区域的反射系数,其中,更大的反射系数引起收听房间混响信号到组合器的输入的相对更高的增益。

[0224] c) 所述第一传输边界区域的几何性质。例如,第一传输边界区域的表面积,其中,更大的表面积引起相邻房间混响信号到组合器的输入的相对更高的增益。

[0225] d) 收听位置。例如,收听者到第一传输边界区域的距离,其中,更小的距离引起混响信号到组合器的输入的相对更小的增益。

[0226] e) 从收听位置到过渡区域边界的声学距离;

[0227] f) 从收听位置到声源位置的声学距离;以及

[0228] g) 相邻房间的尺寸。例如,相邻房间的垂直于第一传输边界区域的尺寸,其中,更大的尺寸引起相邻房间混响信号到组合器的输入的相对更小的增益。

[0229] 作为具体示例,在一些实施例中,该方法可以通过并行运行多个混响器来从具有显著不同特性的多个房间生成混响音频。通常,一个混响器可以用于需要呈现的每个房间/声学环境。

[0230] 为了优化处理,当所呈现的场景中的房间的数量增加到例如多于3或4个房间时,确定需要呈现哪些房间可能是重要的方面。这可以以许多不同的方式实现。为了改进质量,可以基于房间的感知相关性对房间进行排序。这可以通过根据房间在收听位置处的混响响度对房间进行排序来实现。显然,当收听者处于具有混响性质的环境中时,该房间可能是模拟的最重要的房间。

[0231] 房间中的源的数量及其响度可以起重要作用,理想地与房间的能量(DSR)组合并且进一步与房间泄漏/传输到收听房间的量组合。例如,房间k的房间相关性数可以被导出为:

$$[0232] \quad R_k = L_k * DSR_{100ms,k} * A_{leak,k \rightarrow k_c}$$

[0233] 其中,L表示房间中的源的组合平均响度,DSR_{100ms,k}表示从100ms开始测量的频率平均DSR,并且 $A_{leak,k \rightarrow k_c}$ 表示房间k和房间 k_c 之间的包括收听位置的有效泄漏表面(例如,对应于传输边界区域的有效大小)。

[0234] 有效泄漏表面可以基于其透明度来确定。例如

$$[0235] \quad A_{leak,k \rightarrow c} = \sum_{j=0}^{J-1} A_{leak,j} * (t_j + c_j)$$

[0236] 可以导出更高级的并且通常也更复杂的等式。例如,通过考虑房间k中的源与收听位置之间发生的距离和遮挡衰减。

[0237] 基本上,每个混响器可以表示单个房间。来自场景中的源的输入信号以适当的(相对)水平下混,以表示在混响器从其创建混响信号之前它们在房间中具有多少影响。这通常是用于在耳机上回放的双耳信号,但也可以是用于扬声器回放的多声道信号。

[0238] 当收听者在房间内时,这是呈现相关联的混响器的混响信号的适当方式。然而,来自其他房间的混响信号通常不应被呈现为从所有侧面到达收听者的完全漫射信号。替代地,它可以被呈现为邻近对应的传输边界区域的可定位源。

[0239] 混响信号的这种呈现可以与当呈现正常源时相同,距离衰减、遮挡、衍射和其他声学效应也可能对其起作用。通常,房间传输区域可以由具有与区域的尺寸匹配的空间范围的对象表示,使得声音看起来源自整个房间泄漏(通常是门或窗)。

[0240] 因此,相邻房间混响信号可以被馈送到已经存在的直接路径呈现器中,并且这可

以生成与相邻房间混响信号相关联的至少一个新源。

[0241] 在许多实施例中,路由可以不是硬开关,如图10中所示,而是可以由交叉衰落系数控制,其中,漫射表示以及混响源表示同时有效。这可以用于在收听者靠近房间泄漏时创建平滑过渡。在6DoF内容中,收听者通常具有从一个房间移动到另一个房间的自由度,并且因此受益于漫射表示平滑地过渡到基于源(并且反之亦然)的表示。

[0242] 例如,对于房间边界处的收听位置,房间A混响的交叉衰落系数 α_{xf} 可以是0.5,对于房间A中距边界至少1m距离的收听位置,可以是1,并且对于房间B中距边界至少1m距离的收听位置,可以是0。同时,来自房间B混响的交叉衰落系数可以具有反比关系。当收听者更远离房间泄漏时,收听者所在的房间的交叉衰落系数为1,并且对于所有其他房间为0,使得收听者所在的房间的混响是完全漫射的,并且所有其他房间的混响是完全定向的。

[0243] 另外,房间的混响信号可以被馈送到其他房间的早期反射处理和/或混响处理。在大多数实施例中,这些路由信号将不经受交叉衰落。

[0244] 实现从混响器的输出的输出到其他混响器的输入(或到早期反射输入)的适当映射的有利方式是使用映射矩阵。作为示例,映射矩阵可以将每个混响器输出信号映射到所有其他混响器的输入而不是映射到其自身。

$$[0245] \quad A_{\text{map}} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

[0246] 当存在多个传输边界区域时,可以针对多个混响源处理相同的混响输出信号。即,相同的信号可以用于呈现多个混响源。这可以通过生成参考相同信号的多个混响源来实现。

[0247] 当在混响信号的漫射表示和定向表示之间切换或交叉衰落时,可能期望对准这些不同的呈现,使得不发生伪迹。空间交叉衰落可以有助于这一点,因为硬切换通常难以掩蔽。用于在表示之间具有硬切换的实施例的最小伪像减少技术可以是滞后,其中在针对从房间A切换到房间B的阈值vs针对从房间B切换到房间A的阈值之间存在空间距离。

[0248] 此外,水平的对齐可能是有利的。在许多实施例中,确保两个表示的信号水平在整个交叉衰落区域中是稳定且相似的可能是有益的。这可以例如通过将混响源的参考距离(d_{ref})和最小收听者-源距离(d_{min})设置为等于交叉衰落边界到混响源的(公共、平均或最大)垂直距离来实现。

[0249] 通常不需要在整个交叉衰落区域中具有稳定的水平。一些实施例可以仅在某个子区域处对准水平。

[0250] 许多其他实施例可以以在收听者正在房间外移动时将混响水平显著衰减到更低响度为目标。

[0251] 当传输边界区域的尺寸增加时,它将在另一房间中引起更高的混响响度。即,大的门将引起比小的窗更多的混响能量通过。为了引入这种效应,可以根据房间泄漏的大小来缩放被呈现为可定位源的信号。没有额外增益的信号可以表示参考房间泄漏大小。例如, $A_{\text{ref}} = 4\text{m}^2$ 。具有不同大小的房间泄漏可以被分配与房间泄漏大小与该参考房间泄漏大小的比率成比例的增益。

$$[0252] \quad g_{rls} = \frac{A_{leak,k \rightarrow k_c}}{A_{ref}}$$

[0253] 替代地,源的范围呈现可以采用针对更大范围实现更高源响度的水平归一化模式。例如,不衰减作为跨越范围的点源呈现的信号以补偿点源的量,或确保由跨越范围的点源表示的组合信号功率根据来自上述等式的增益 g_{rls} 缩放。

[0254] 术语音频和声音可以被认为是等同的和可互换的,并且可以分别指代在上下文中诸如适当的物理声压和/或电信号表示。

[0255] 将认识到,为了清楚性,以上描述已经参考不同的功能电路、单元和处理器描述了本发明的实施例。然而,将显而易见的是,在不减损本发明的情况下,可以使用不同的功能电路、单元或处理器之间的功能性的任何适合的分布。例如,被图示为由分离的处理器或控制器执行的功能性可以由相同处理器或控制器执行。因此,对特定功能单元或电路的引用仅将被看作对用于提供所描述的功能性而不是指示严格逻辑或物理结构或组织的适合的装置的引用。

[0256] 本发明可以以任何适合的形式来实施,包括硬件、软件、固件或这些的任何组合。本发明可以可选地至少部分地被实施为在一个或多个数据处理器和/或数字信号处理器上运行的计算机软件。本发明的实施例的元件和部件可以以任何适合的方式物理地、功能地和逻辑地实施。实际上,功能可以被实施在单个单元中、在多个单元中或作为其他功能单元的一部分。因此,本发明可以被实施在单个单元中或可以物理地并且功能上被分布在不同的单元、电路和处理器之间。

[0257] 虽然本发明已经结合一些实施例描述,但是其不旨在限于本文阐述的特定形式。相反,本发明的范围仅由附图限制。此外,虽然特征可能看起来结合特定实施例描述,但是本领域的技术人员将意识到,所描述的实施例的各种特征可以根据本发明进行组合。在权利要求中,术语包括不排除其他元件或步骤的存在。

[0258] 此外,虽然单独地列出,但是多个装置、元件、电路或方法步骤可以通过例如单个电路、单元或处理器实施。此外,虽然单独特征可以被包括在不同的权利要求中,但是这些可以可能地被有利地组合,并且不同的权利要求中的包括不暗示特征的组合不是可行的和/或有利的。而且,一个类别的权利要求中的特征的包括不暗示对该类别的限制,而是相反指示特征同样地酌情适用于其他权利要求类别。此外,权利要求中的特征的次序不暗示特征必须工作的任何特定次序,并且特别地,方法权利要求中的单独步骤的次序不暗示步骤必须以该次序执行。相反,步骤可以以任何适合的次序执行。另外,单数引用不排除多个。因此,对“一”、“一个”、“第一”、“第二”等的引用不排除多个。权利要求中的附图标记仅被提供为澄清示例,而不应当被解释为以任何方式限制权利要求的范围。

[0259] 总体上,音频装置及其方法的示例由下面的实施例指示。

[0260] 实施例:

[0261] 1、一种音频装置:

[0262] 第一接收器(501),其被布置为接收包括针对多个房间的场景中的音频源的音频数据;

[0263] 位置电路(505),其被布置为确定所述场景中的收听位置;

[0264] 确定器(509),其被布置为确定包括所述收听位置的第一房间和作为所述第一房

间的相邻房间的第二房间；

[0265] 第二接收器(503),其被布置为接收针对所述第一房间和所述第二房间的空间声学传输数据,所述空间声学传输数据描述针对所述第一房间的多个传输边界区域,每个传输边界区域的针对从所述第二房间到所述第一房间的声音的声学传输水平超过一阈值；

[0266] 第一混响器(511),其被布置为根据所述第二房间中的至少一个音频源和所述第二房间的至少一个性质确定针对所述第二房间的第二房间混响音频信号；

[0267] 声源电路(513),其被布置为:针对所述多个传输边界区域中的至少第一传输边界区域,确定音频源在所述第二房间中的声源位置；

[0268] 呈现器(507),其被布置为呈现所述收听位置的音频信号,所述呈现包括通过从所述声源位置呈现所述第二房间混响音频信号来生成第一音频分量。

[0269] 2、根据权利要求1所述的音频装置,其中,所述呈现取决于所述第一传输边界区域的几何性质和声学性质中的至少一项。

[0270] 3、根据前述权利要求中的任一项所述的音频装置,其中,从所述声源位置到所述第一传输边界区域的距离不小于所述第一传输边界区域内的最大距离的十分之一。

[0271] 4、根据前述权利要求中的任一项所述的音频装置,其中,从所述声源位置到所述第一传输边界区域的距离不小于所述第二房间内的最大距离的十分之一。

[0272] 5、根据前述权利要求中的任一项所述的音频装置,其中,从所述声源位置到所述第一传输边界区域的距离不小于20cm。

[0273] 6、根据前述权利要求中的任一项所述的音频装置,其中,所述呈现包括对从所述声源位置通过所述第一传输边界区域到所述收听位置的声学路径进行呈现。

[0274] 7、根据前述权利要求中的任一项所述的音频装置,其中,所述呈现包括通过将所述第二房间混响信号呈现为混响音频分量来生成第二音频分量。

[0275] 8、根据权利要求7所述的音频装置,其中,所述呈现包括响应于相对于所述第一传输边界区域的所述收听位置而相对于所述第二音频分量的水平而调整所述第一音频分量的水平。

[0276] 9、根据权利要求8所述的音频装置,其中,所述呈现包括针对到所述第二房间的增加的距离,相对于所述第二音频分量的水平增加所述第一音频分量的水平。

[0277] 10、根据权利要求7-9中的任一项所述的音频装置,其中,所述呈现包括响应于所述第一传输边界区域的尺寸而相对于所述第二音频分量的水平调整所述第一音频分量的水平。

[0278] 11、根据前述权利要求中的任一项所述的音频装置,其中,所述呈现器(507)被布置为将来自所述声源位置的所述第二房间混响音频信号呈现为空间扩展的声源。

[0279] 12、根据前述权利要求中的任一项所述的音频装置,其中,所述呈现器(506)包括:

[0280] 路径呈现器(1001),其用于呈现声学路径的音频；

[0281] 多个混响器(1003),其被布置为生成房间的混响信号,所述多个混响器包括所述第一混响器；

[0282] 耦合电路(1005),其用于将来自所述多个呈现器的混响信号耦合到所述路径呈现器(1001)；

[0283] 组合电路(1005),其用于组合来自所述多个呈现器(1003)的混响信号和来自所述

路径呈现器(1001)的输出信号以生成组合音频信号;以及

[0284] 适配器(1007),其用于调整所述混响信号的水平以用于由所述耦合电路(1005)进行的耦合和用于由所述组合电路(1005)进行的组合。

[0285] 13、根据权利要求12所述的音频装置,其中,所述适配器(1007)被布置为响应于以下中的至少一项而调整所述混响信号的水平:

[0286] 与所述音频源的所述音频数据一起接收的元数据;

[0287] 所述第一传输边界区域的声学性质;

[0288] 所述第一传输边界区域的几何性质;

[0289] 收听位置;

[0290] 从所述收听位置到所述声源位置的声学距离;以及

[0291] 所述第一传输边界区域的尺寸。

[0292] 14、根据前述权利要求中的任一项所述的音频装置,其中,所述第一混响器(511)还被布置为响应于第一房间混响信号而生成所述第二房间混响信号。

[0293] 15、一种用于音频装置的操作的方法,所述方法包括:

[0294] 接收包括多个房间的场景中的音频源的音频数据;

[0295] 确定所述场景中的收听位置;

[0296] 确定包括所述收听位置的第一房间和为所述第一房间的相邻房间的第二房间;

[0297] 接收针对所述第一房间和所述第二房间的空间声学传输数据,所述空间声学传输数据描述针对所述第一房间的多个传输边界区域,每个传输边界区域的针对从所述第二房间到所述第一房间的声音的声学传输水平超过一阈值;

[0298] 从所述第二房间中的至少一个音频源和所述第二房间的至少一个性质确定针对所述第二房间的第二房间混响音频信号;

[0299] 针对所述多个传输边界区域中的至少第一传输边界区域,确定音频源在所述第二房间中的声源位置;

[0300] 呈现所述收听位置的音频信号,所述呈现包括通过从所述声源位置呈现所述第二房间混响音频信号来生成第一音频分量。

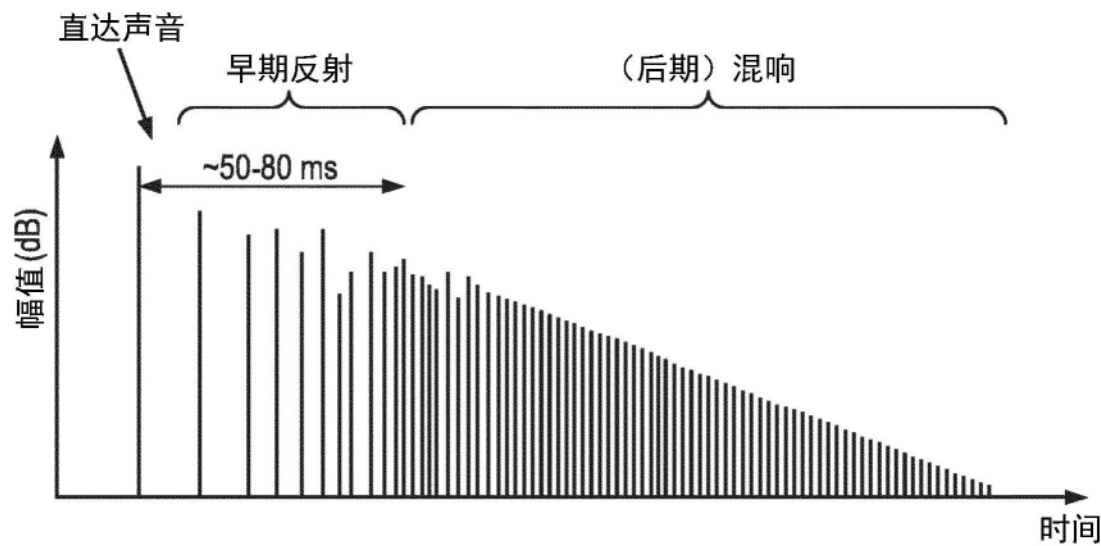


图1

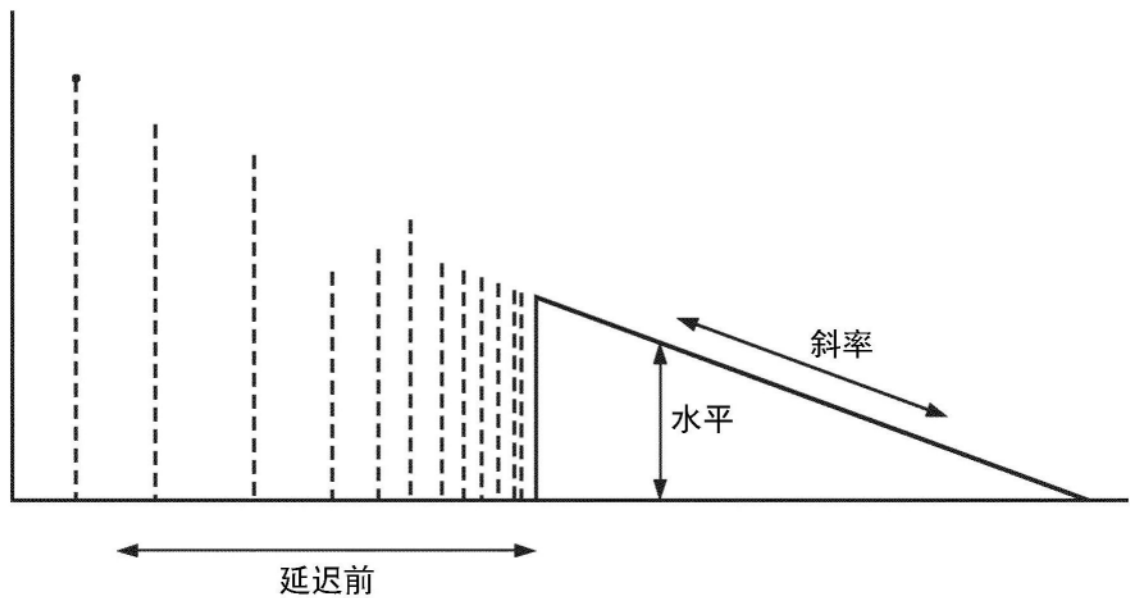


图2

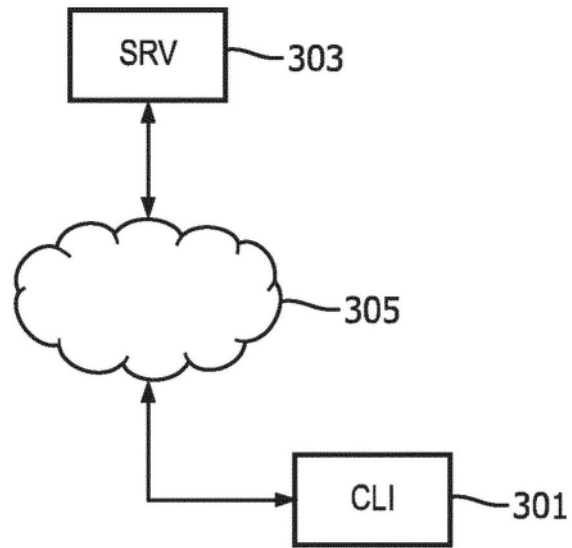


图3

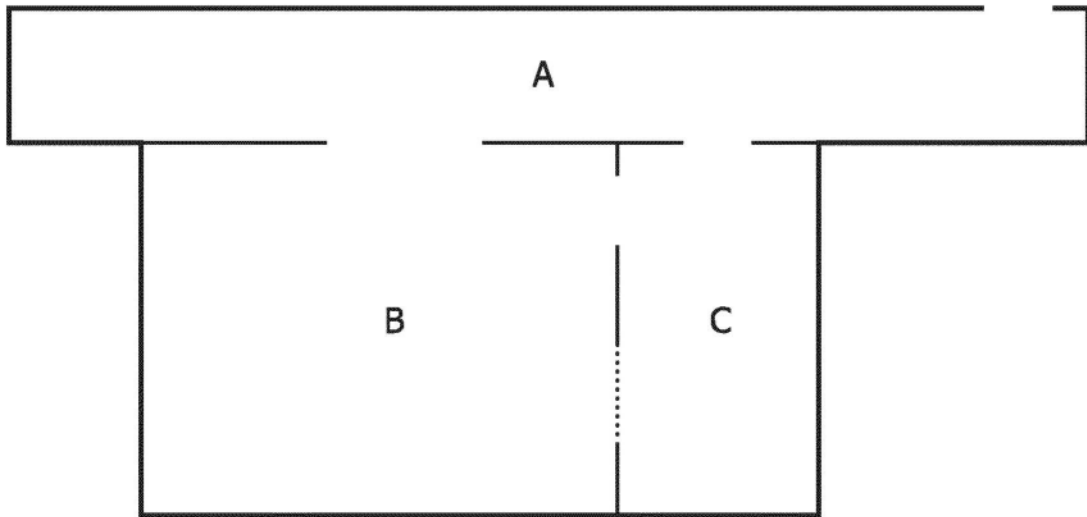


图4

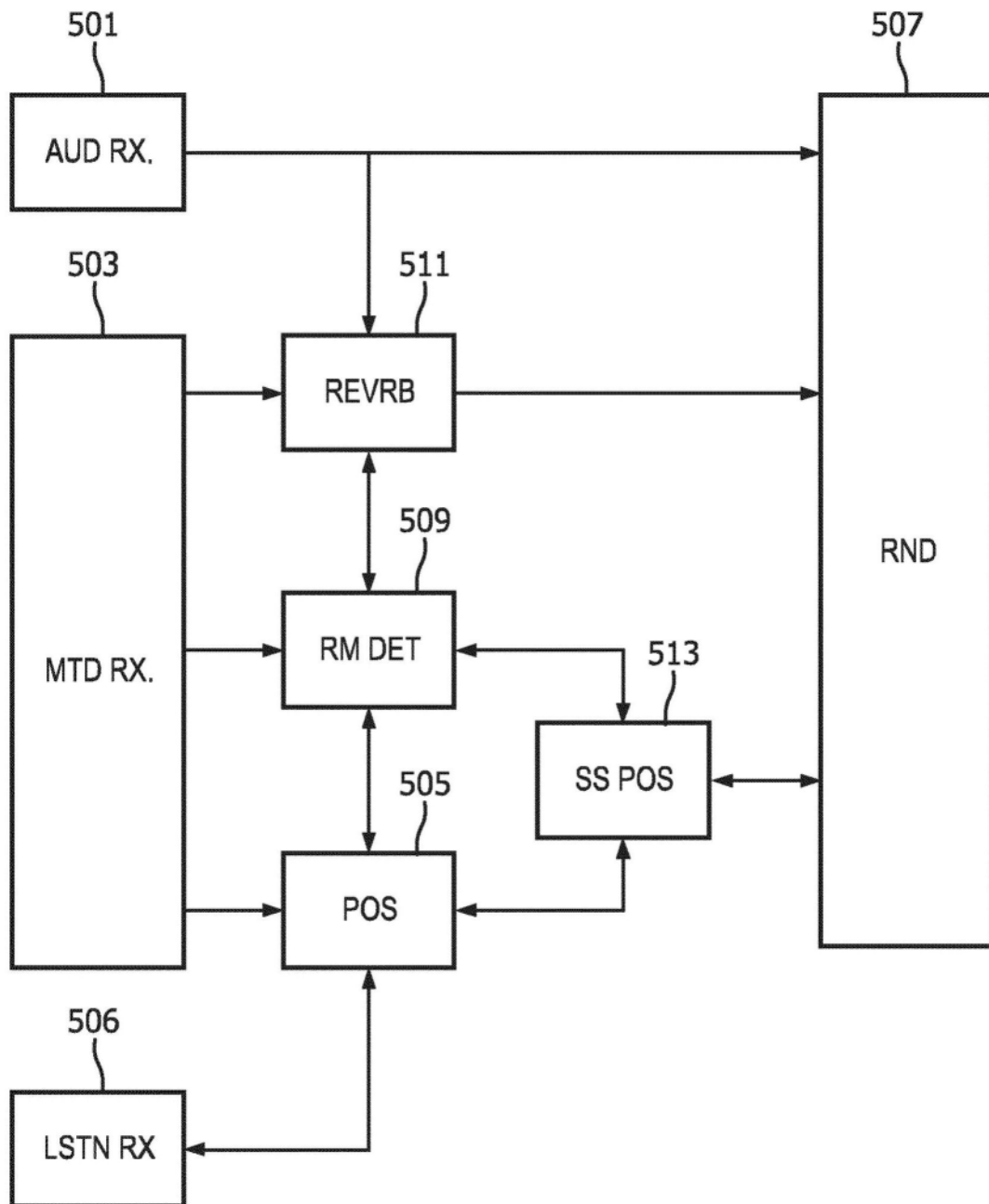


图5

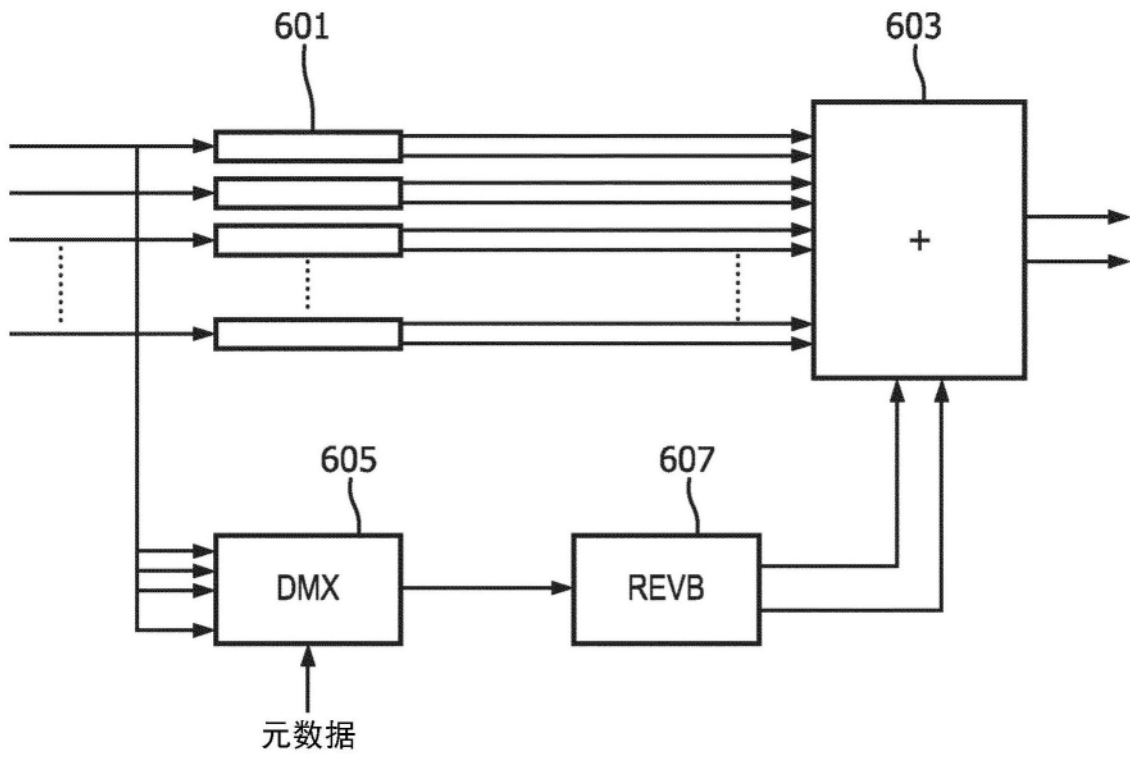


图6

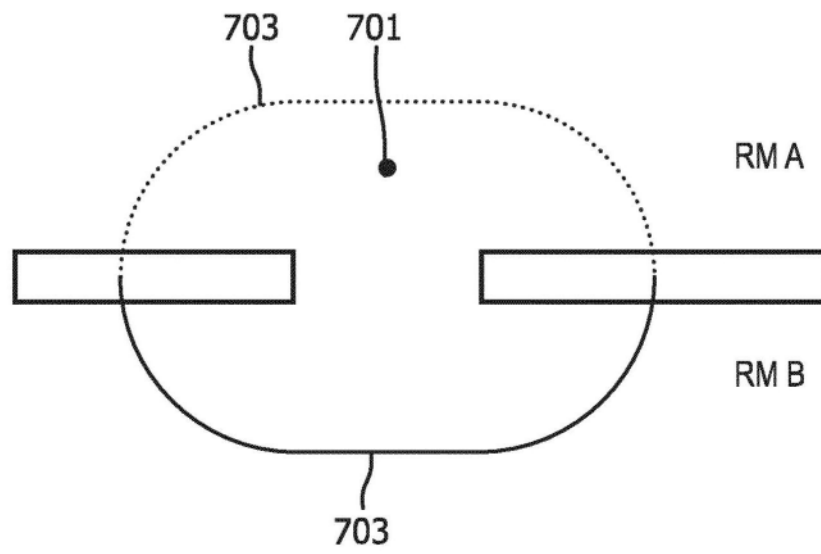


图7

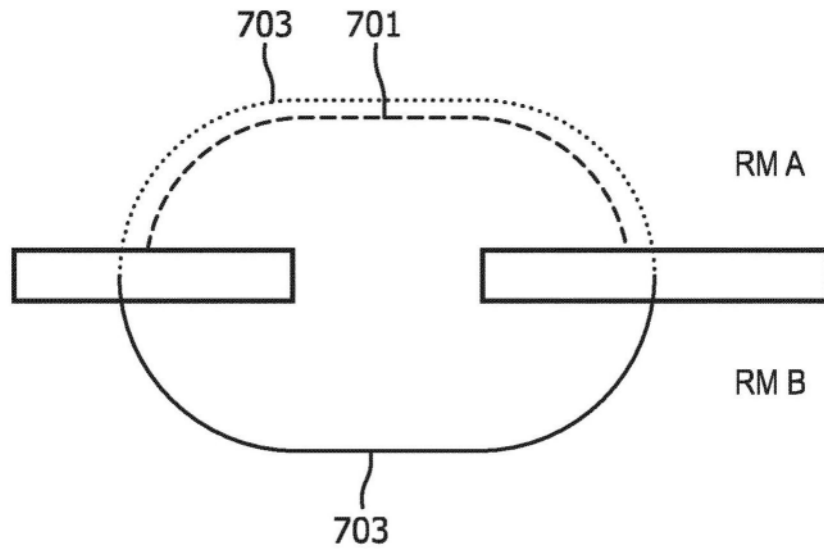


图8

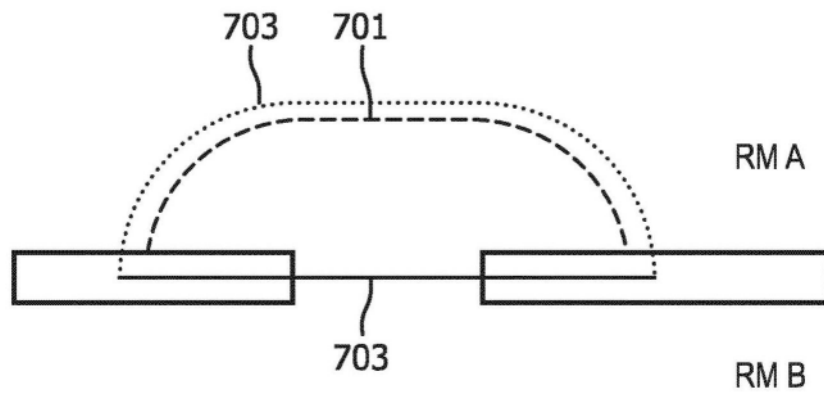


图9

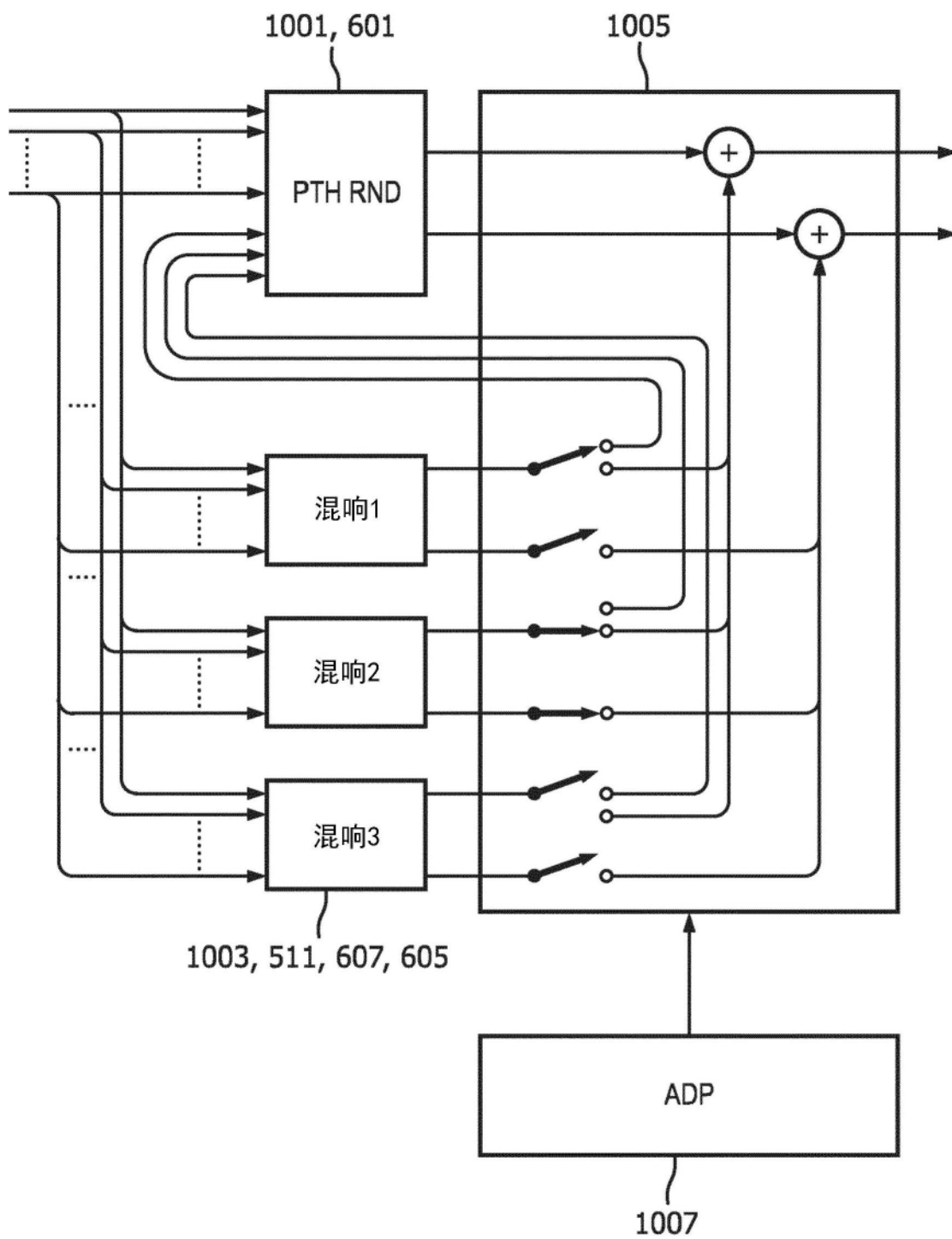


图10