



(12) 发明专利申请

(10) 申请公布号 CN 115398936 A

(43) 申请公布日 2022. 11. 25

(21) 申请号 202180027939.5

M · Z · 兰德

(22) 申请日 2021.02.11

(74) 专利代理机构 北京市中咨律师事务所  
11247

(30) 优先权数据

62/976,569 2020.02.14 US

专利代理师 尹丽丹 于静

(85) PCT国际申请进入国家阶段日

2022.10.11

(51) Int. Cl.

H04R 5/00 (2006.01)

G06F 3/01 (2006.01)

(86) PCT国际申请的申请数据

PCT/US2021/017728 2021.02.11

G06F 3/16 (2006.01)

(87) PCT国际申请的公布数据

W02021/163382 EN 2021.08.19

(71) 申请人 奇跃公司

地址 美国佛罗里达州

(72) 发明人 R · S · 奥德弗雷

M · B · 赫腾施泰因尔

S · C · 迪克尔 B · I · 伍德

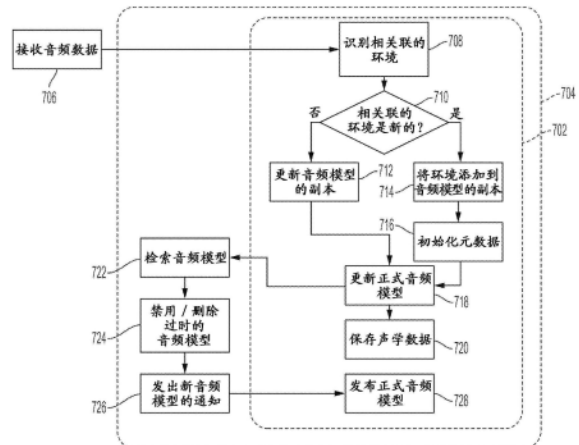
权利要求书2页 说明书25页 附图14页

(54) 发明名称

多应用音频渲染

(57) 摘要

本文公开了用于有效地渲染音频的系统和方法。一种方法可包括：接收呈现第一音轨的请求，其中，第一音轨基于包括共享模型分量和第一模型分量的第一音频模型；接收呈现第二音轨的请求，其中，第二音轨基于包括共享模型分量和第二模型分量的第二音频模型；基于第一音轨、第二音轨、共享模型分量、第一模型分量和第二模型分量，渲染声音；以及经由一个或多个扬声器呈现包括所渲染的声音的音频信号。



1. 一种系统,包括:
  - 一个或多个扬声器;以及
  - 一个或多个处理器,其被配置为执行一种方法,所述方法包括:
    - 接收呈现第一音轨的请求,其中,所述第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;
    - 接收呈现第二音轨的请求,其中,所述第二音轨基于包括所述共享模型分量和第二模型分量的第二音频模型;
    - 基于所述第一音轨、所述第二音轨、所述共享模型分量、所述第一模型分量和所述第二模型分量,渲染声音;以及
    - 经由所述一个或多个扬声器呈现包括所渲染的声音的音频信号。
2. 根据权利要求1所述的系统,其中,渲染所述声音包括:基于所述共享模型分量,渲染所述第一音轨和所述第二音轨。
3. 根据权利要求1所述的系统,其中,渲染所述声音包括:基于所述第一音频分量,渲染所述第一音轨。
4. 根据权利要求1所述的系统,其中,渲染所述声音包括:基于所述第二音频分量渲染所述第二音轨。
5. 根据权利要求1所述的系统,其中,所述一个或多个扬声器是头戴式设备的扬声器。
6. 根据权利要求5所述的系统,其中,所述第一音频模型基于所述头戴式设备的物理环境。
7. 根据权利要求1所述的系统,其中,所述共享模型分量对应于直接路径声学行为。
8. 根据权利要求1所述的系统,其中,所述共享模型分量对应于反射声学行为。
9. 根据权利要求1所述的系统,其中,所述共享模型分量对应于混响声学行为。
10. 一种方法,包括:
  - 接收呈现第一音轨的请求,其中,所述第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;
  - 接收呈现第二音轨的请求,其中,所述第二音轨基于包括所述共享模型分量和第二模型分量的第二音频模型;
  - 基于所述第一音轨、所述第二音轨、所述共享模型分量、所述第一模型分量和所述第二模型分量,渲染声音;以及
  - 经由一个或多个扬声器呈现包括所渲染的声音的音频信号。
11. 根据权利要求10所述的方法,其中,渲染所述声音包括:基于所述共享模型分量,渲染所述第一音轨和所述第二音轨。
12. 根据权利要求10所述的方法,其中,渲染所述声音包括:基于所述第一音频分量,渲染所述第一音轨。
13. 根据权利要求10所述的方法,其中,渲染所述声音包括:基于所述第二音频分量,渲染所述第二音轨。
14. 根据权利要求10所述的方法,其中,呈现所述音频信号包括:经由头戴式设备的一个或多个扬声器呈现所述音频信号。
15. 根据权利要求14所述的方法,其中,所述第一音频模型基于所述头戴式设备的物理

环境。

16. 根据权利要求10所述的方法,其中,所述共享模型分量对应于直接路径声学行为。

17. 根据权利要求10所述的方法,其中,所述共享模型分量对应于反射声学行为。

18. 根据权利要求10所述的方法,其中,所述共享模型分量对应于混响声学行为。

19. 一种存储指令的非暂态计算机可读介质,所述指令在由一个或多个处理器执行时使得所述一个或多个处理器执行一种方法,所述方法包括以下各项:

接收呈现第一音轨的请求,其中,所述第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;

接收呈现第二音轨的请求,其中,所述第二音轨基于包括所述共享模型分量和第二模型分量的第二音频模型;

基于所述第一音轨、所述第二音轨、所述共享模型分量、所述第一模型分量和所述第二模型分量,渲染声音;以及

经由一个或多个扬声器呈现包括所渲染的声音的音频信号。

20. 根据权利要求19所述的非暂态计算机可读介质,其中,渲染所述声音包括:基于所述共享模型分量,渲染所述第一音轨和所述第二音轨。

21. 根据权利要求19所述的非暂态计算机可读介质,其中,渲染所述声音包括:基于所述第一音频分量,渲染所述第一音轨。

22. 根据权利要求19所述的非暂态计算机可读介质,其中,渲染所述声音包括:基于所述第二音频分量,渲染所述第二音轨。

23. 根据权利要求19所述的非暂态计算机可读介质,其中,呈现所述音频信号包括:经由头戴式设备的一个或多个扬声器呈现所述音频信号。

24. 根据权利要求23所述的非暂态计算机可读介质,其中,所述第一音频模型基于所述头戴式设备的物理环境。

25. 根据权利要求19所述的非暂态计算机可读介质,其中,所述共享模型分量对应于直接路径声学行为。

26. 根据权利要求19所述的非暂态计算机可读介质,其中,所述共享模型分量对应于反射声学行为。

27. 根据权利要求19所述的非暂态计算机可读介质,其中,所述共享模型分量对应于混响声学行为。

## 多应用音频渲染

[0001] 相关申请的交叉引用

[0002] 本申请要求于2020年2月14日提交的美国临时申请NO.62/976,569的权益,该申请的全部内容通过引用并入本文。

### 技术领域

[0003] 本公开大体涉及用于管理和存储音频数据的系统和方法,并且特别地,涉及用于在混合现实环境中管理和存储音频数据的系统和方法。

### 背景技术

[0004] 虚拟环境在计算环境中是普遍存在的,虚拟环境应用于视频游戏(其中,虚拟环境可表示游戏世界);地图(其中,虚拟环境可表示待导航的地形);模拟(其中,虚拟环境可模拟真实环境);数字叙事(其中,虚拟角色可在虚拟环境中彼此交互);和许多其他应用。现代计算机用户通常舒适感知虚拟环境并且与虚拟环境交互。然而,关于虚拟环境的用户的体验可能由用于呈现虚拟环境的技术限制。例如,常规显示器(例如,2D显示屏)和音频系统(例如,固定扬声器)可能不能够以产生令人信服、逼真并且沉浸式体验的方式实现虚拟环境。

[0005] 虚拟现实(“VR”)、增强现实(“AR”)、混合现实(“MR”)、和相关技术(统称为“XR”)共享向XR系统的用户呈现与由计算机系统中的数据表示的虚拟环境相对应的感觉信息。这样的系统可通过将虚拟视觉和音频线索与真实视线和声音组合来提供唯一升高的沉浸感和逼真感。因此,可能希望以这样的方式向XR系统的用户呈现数字声音:声音好像自然地且与用户在用户的真实环境中对声音的期望一致地发生。一般而言,用户期望虚拟声音将呈现听到他们的真实环境的声学特性。例如,大型音乐厅中的XR系统的用户将期望XR系统的虚拟声音具有大的、海绵状的音质;相反,小公寓中的用户将期望声音更衰减、接近和即时。除了将虚拟声音与真实和/或虚拟环境的声学特性相匹配之外,通过空间化虚拟声音进一步增强了真实感。例如,虚拟对象可能从后面在视觉上飞过用户,并且用户可能期望对应的虚拟声音类似地反映虚拟对象相对于用户的空间运动。

[0006] 现有技术往往达不到这些期望,诸如通过呈现不考虑用户周围环境或不与虚拟对象的空间运动相对应的虚拟音频,导致可能损害用户体验的不真实的感觉。对XR系统用户的观察表明,虽然用户可能对虚拟内容和真实环境之间的视觉失配(例如,照明不一致)相对宽容;用户可能对听觉失配更敏感。我们自己的听觉体验,在我们的一生中不断完善,可以使我们敏锐地意识到我们的物理环境如何影响我们听到的声音;我们可以高度感知与那些期望不一致的声音。利用XR系统,这样的不一致可能是刺耳的,并且可能将身临其境且引人入胜的体验变为花哨的模仿体验。在极端示例中,听觉不一致可能导致晕动病和其他不良影响,因为内耳无法将听觉刺激与其对应的视觉线索相协调。

[0007] 需要系统架构来组织和管理用于生成虚拟音频的系统。生成虚拟音频可能涉及管理和存储关于用户环境的信息,以使得该信息可用于产生逼真的虚拟声音。因此,音频系统

架构可能需要与其他系统接口以接收和利用与音频引擎相关的信息。还可能希望一种音频系统架构,该架构可以在使用期间无中断地呈现逼真的声音。可以在不中断用户的情况下更新音频引擎的音频系统架构可以产生身临其境的用户体验,其中听觉信号持续反映用户的环境。

[0008] 通过考虑用户物理环境的特性,本文所描述的系统和方法可以模拟如果虚拟声音是在该环境中自然生成的真实声音,用户将听到什么。通过以忠实于声音在现实世界中表现方式的方式呈现虚拟声音,用户可以体验与混合现实环境的高度关联感。类似地,通过呈现响应于用户运动和环境的位置感知虚拟内容,内容变得更主观、互动和真实——例如,用户在A点处的体验可能与他或她在B点处的体验完全不同。这种增强的逼真感和交互性可以为混合现实的新应用提供基础,诸如使用空间感知音频来使能新颖形式的游戏玩法、社交特征或交互行为的应用。

### 发明内容

[0009] 本公开的示例描述了用于有效地渲染音频的系统和方法。根据本公开的示例,一种方法可以包括:接收呈现第一音轨的请求,其中,所述第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;接收呈现第二音轨的请求,其中,所述第二音轨基于包括所述共享模型分量和第二模型分量的第二音频模型;基于所述第一音轨、所述第二音轨、所述共享模型分量、所述第一模型分量和所述第二模型分量,渲染声音;以及经由一个或多个扬声器呈现包括所渲染的声音的音频信号。

### 附图说明

[0010] 图1A至图1C示出了根据一些实施例的示例混合现实环境。

[0011] 图2A至图2D示出了根据一些实施例的可用于生成并且与混合现实环境相互作用的示例混合现实环境的部件。

[0012] 图3A示出了根据一些实施例的可用于向混合现实环境提供输入的示例混合现实手持式控制器。

[0013] 图3B示出了根据一些实施例的可与示例混合现实系统一起使用的示例辅助单元。

[0014] 图4示出了根据一些实施例的用于示例混合现实系统的示例功能框图。

[0015] 图5示出了根据一些实施例的虚拟音频系统的示例。

[0016] 图6示出了根据一些实施例的用于更新音频模型的示例过程。

[0017] 图7示出了根据一些实施例的用于更新音频模型的示例过程。

[0018] 图8示出了根据一些实施例的示例性模型管理架构。

[0019] 图9A至图9B示出了根据一些实施例的示例性渲染架构。

### 具体实施方式

[0020] 在示例的以下描述中,对形成其一部分并且在其中通过图示的方式示出可以实践的示例的附图进行参考。应理解,在不脱离所公开的示例的范围的情况下,可以使用其他示例并且可以做出结构改变。

[0021] 混合现实环境

[0022] 像所有人一样,混合现实系统的用户存在于真实环境中——即,“真实世界”的三维部分和其内容的全部可由用户感知。例如,用户使用个人的普通人类感觉感知真实世界——视线、声音、触碰、味道、气味——并且通过在真实环境中移动自己的身体与真实环境交互。真实环境中的位置可被描述为坐标空间中的坐标;例如,坐标可包括纬度、经度和相对于海平面的海拔;在三个正交维度上距参考点的距离;或其他适合的值。同样地,矢量可描述具有坐标空间中的方向和数量的数量。

[0023] 计算设备可例如在与设备相关联的存储器中维护虚拟环境的表示。如本文所使用的,虚拟环境是三维空间的计算表示。虚拟环境可包括与该空间相关联的任何对象、动作、信号、参数、坐标、矢量、或其他特性的表示。在一些示例中,计算设备的电路(例如,处理器)可以维护和更新虚拟环境的状态;即,处理器可以在第一时间 $t_0$ 处基于与虚拟环境相关联的数据和/或由用户提供的输入来确定虚拟环境在第二时间 $t_1$ 处的状态。例如,如果虚拟环境中的对象在时间 $t_0$ 处位于第一坐标处,并且具有某个编程的物理参数(例如,质量、摩擦系数);以及从用户接收的输入指示力应当以方向矢量施加到该对象;处理器可应用运动学定律来使用基础力学确定对象在时间 $t_1$ 处的位置。处理器可以使用关于虚拟环境已知的任何适合的信息和/或任何适合的输入来确定虚拟环境在时间 $t_1$ 处的状态。在维护和更新虚拟环境的状态时,处理器可执行任何适合的软件,包括与虚拟环境中的虚拟对象的创建和删除有关的软件;用于定义虚拟环境中的虚拟对象或角色的行为的软件(例如,脚本);用于定义虚拟环境中的信号(例如,音频信号)的行为的软件;用于创建和更新与虚拟环境相关联的参数的软件;用于生成虚拟环境中的音频信号的软件;用于处理输入和输出的软件;用于实现网络操作的软件;用于应用资产数据(例如,随时间移动虚拟对象的动画数据)的软件;或许多其他可能性。

[0024] 输出设备(诸如显示器或者扬声器)可以向用户呈现虚拟环境的任何或所有方面。例如,虚拟环境可包括可以向用户呈现虚拟对象(其可包括无生命对象;人;动物;光;等的表示)。处理器可以确定虚拟环境的视图(例如,与具有坐标原点、视轴和视锥的“相机”相对应);以及向显示器渲染与该视图相对应的虚拟环境的可视场景。出于该目的,可以使用任何适合的渲染技术。在一些示例中,可视场景可以仅包括虚拟环境中的一些虚拟对象,并且不包括某些其他虚拟对象。类似地,虚拟环境可包括可向用户呈现为一个或多个音频信号的音频方面。例如,虚拟环境中的虚拟对象可生成起源于对象的位置坐标的声音(例如,虚拟角色可以说话或引起声音效果);或者虚拟环境可以与可以或可以不与特定位置相关联的音乐线索或环境声音相关联。处理器可确定与“听众”坐标相对应的音频信号——例如,与虚拟环境中的声音的复合相对应的音频信号,并且被混合并处理以模拟将由听众坐标处的听众听到的音频信号——并且经由一个或多个扬声器向用户呈现音频信号。

[0025] 由于虚拟环境仅存在作为计算结构,所以用户不能使用个人的普通感觉直接感知虚拟环境。相反,用户可以仅间接地感知如呈现给用户的虚拟环境,例如通过显示器、扬声器、触觉输出设备等。类似地,用户不能直接触碰、操控、或以其他方式与虚拟环境交互;但是可以经由输入设备或传感器将输入数据提供给可使用设备或传感器数据来更新虚拟环境的处理器。例如,相机传感器可提供指示用户正试图移动虚拟环境中的对象的光学数据,并且处理器可使用该数据使得对象因此在虚拟环境中作出反应。

[0026] 混合现实系统可以向用户呈现组合真实环境和虚拟环境的各方面的混合现实环

境(“MRE”),例如使用透射式显示器和/或一个或多个扬声器(其可以例如被并入可穿戴头部设备中)。在一些实施例中,一个或多个扬声器可以在头戴式可穿戴单元的外部。如本文所使用的,MRE是真实环境和对应的虚拟环境的同时表示。在一些示例中,对应的真实环境和虚拟环境共享单个坐标空间;在一些示例中,真实坐标空间和对应的虚拟坐标空间通过变换矩阵(或其他适合的表示)彼此相关。因此,单个坐标(在一些示例中,连同变换矩阵一起)可以定义真实环境中的第一位置,以及虚拟环境中对应的第二位置;反之亦然。

[0027] 在MRE中,虚拟对象(例如,在与MRE相关联的虚拟环境中)可以对应于真实对象(例如,在与MRE相关联的真实环境中)。例如,如果MRE的真实环境包括位置坐标处的真实灯杆(真实对象),则MRE的虚拟环境可包括对应的位置坐标处的虚拟灯杆(虚拟对象)。如本文所使用的,真实对象与其对应的虚拟对象一起组合构成“混合现实对象”。不需要虚拟对象与对应的真实对象完美匹配或者对准。在一些示例中,虚拟对象可以是对应的真实对象的简化版本。例如,如果真实环境包括真实灯杆,则对应的虚拟对象可以包括具有与真实灯杆粗略地相同高度和半径的圆柱体(可以在形状方面反映该灯杆是粗略地圆柱形)。以这种方式简化虚拟对象可以允许计算效率,并且可以简化将在这样的虚拟对象上执行的计算。进一步地,在MRE的一些示例中,并非所有真实环境中的真实对象可以与对应的虚拟对象相关联。同样地,在MRE的一些示例中,并非所有虚拟环境中的虚拟对象可以与对应的真实对象相关联。即,一些虚拟对象可以仅在MRE的虚拟环境中而没有任何现实世界配对物。

[0028] 在一些示例中,虚拟对象可以具有与对应的真实对象的特性不同(有时急剧地)的特性。例如,虽然MRE中的真实环境可以包括绿色双臂仙人掌——多刺无生命对象——MRE中的对应的虚拟对象可以包括具有人类面部特征和粗暴行为的绿色双臂虚拟角色的特性。在该示例中,虚拟对象在某些特性(颜色、臂数)方面类似其对应的真实对象;但是在其他特性(面部特征、个性)方面与真实对象不同。以这种方式,虚拟对象具有以创造性、抽象、夸大、或想象的方式表示真实对象的潜在性;或者向其他无生命真实对象给予行为(例如,人类个性)。在一些示例中,虚拟对象可以是纯想象创造而没有现实世界配对物(例如,虚拟环境中的虚拟怪物,也许在与真实环境中的空白空间相对应的位置处)。

[0029] 与VR系统相比较,该VR系统向用户呈现虚拟环境同时模糊真实环境,呈现MRE的混合现实系统提供在虚拟环境被呈现时真实环境保持可感知的优点。因此,混合现实系统的用户能够使用与真实环境相关联的视觉和音频线索来体验对应的虚拟环境并且与该对应的虚拟环境交互。作为示例,当VR系统的用户可以努力感知或与虚拟环境中显示的虚拟对象交互时——因为如上所述,用户不能直接感知或与虚拟环境交互——MR系统的用户可以发现通过看到、听到和触碰他或她自己的真实环境中的对应的真实对象来与虚拟对象交互是直观并且自然的。该交互性水平可以提高用户与虚拟环境的沉浸感、连接感和接合感。类似地,通过同时呈现真实环境和虚拟环境,混合现实系统可以减少与VR系统相关联的负面心理感觉(例如,认知失调)和负面身体感觉(例如,运动病)。混合现实系统进一步为可以增加或更改我们的真实世界的体验的应用提供许多可能性。

[0030] 图1A示出了用户110使用混合现实系统112的示例真实环境100。混合现实系统112可以包括显示器(例如,透射式显示器)及一个或多个扬声器、以及一个或多个传感器(例如,相机),例如如下文所描述的。示出的真实环境100包括用户110正站在的矩形房间104A;以及真实对象122A(灯)、124A(桌子)、126A(沙发)和128A(油画)。房间104A还包括位置坐标

106,其可以被称为是真实环境100的原点。如图1A所示,具有在点106(世界坐标)处的原点的环境/世界坐标系108(包括x轴108X、y轴108Y和z轴108Z)可以定义用于真实环境100的坐标空间。在一些实施例中,环境/世界坐标系108的原点106可以与混合现实环境112被供电的位置相对应。在一些实施例中,环境/世界坐标系108的原点106可以在操作期间被重置。在一些示例中,用户110可以被认为是真实环境100中的真实对象;类似地,用户110的身体部分(例如,手部、脚部)可以被认为是真实环境100中的真实对象。在一些示例中,具有在点115(例如,用户/听众/头部坐标)处的原点的用户/听众/头部坐标系114(包括x轴114X、y轴114Y和z轴114Z)可以定义针对混合现实系统112所在的用户/听众/头部的坐标空间。用户/听众/头部坐标系114的原点115可以相对于混合现实系统112的一个或多个部件来定义。例如,用户/听众/头部坐标系114的原点115可以相对于混合现实系统112的显示器来定义,诸如在混合现实系统112的初始校准期间。矩阵(其可以包括平移矩阵和四元数矩阵或其他旋转矩阵)或其他适合的表示可以表征用户/听众/头部坐标系114空间与环境/世界坐标系108空间之间的变换。在一些实施例中,左耳坐标116和右耳坐标117可以相对于用户/听众/头部坐标系114的原点115来定义。矩阵(其可以包括平移矩阵和四元数矩阵或其他旋转矩阵)或其他合适的表示可以表征左耳坐标116和右耳坐标117与用户/听众/头部坐标系114空间之间的变换。用户/听众/头部坐标系114可以简化相对于用户的头部或头戴式设备(例如,相对于环境/世界坐标系108)的位置的表示。使用同时定位和地图创建(SLAM)、视觉里程计或其他技术,用户坐标系114与环境坐标系108之间的变换可以被实时确定和更新。

[0031] 图1B示出了与真实环境100相对应的示例虚拟环境130。示出的虚拟环境130包括与真实矩形房间104A相对应的虚拟矩形房间104B;与真实对象122A相对应的虚拟对象122B;对应于真实对象124A的虚拟对象124B;以及与真实对象126A相对应的虚拟对象126B。与虚拟对象122B、124B、126B相关联的元数据可以包括从对应的真实对象122A、124A、126A导出的信息。虚拟环境130附加地包括虚拟怪物132,该虚拟怪物132不与真实环境100中的任何真实对象相对应。真实环境100中的真实对象128A不与虚拟环境130中的任何虚拟对象相对应。具有点134(持久坐标)处的原点的持久坐标系133(包括x轴133X、y轴133Y和z轴133Z)可以定义用于虚拟内容的坐标空间。持久坐标系133的原点134可以相对于/关于一个或多个真实对象,诸如真实对象126A,来定义。矩阵(其可以包括平移矩阵和四元数矩阵或其他旋转矩阵)或其他适合的表示可以表征持久坐标系133空间与环境/世界坐标系108空间之间的变换。在一些实施例中,虚拟对象122B、124B、126B和132中的每个虚拟对象可以具有其自己相对于持久坐标系133的原点134的持久坐标点。在一些实施例中,可以存在多个持久坐标系,并且虚拟对象122B、124B、126B和132中的每个虚拟对象可以具有其自己相对于一个或多个持久坐标系的持久坐标点。

[0032] 相对于图1A和图1B,环境/世界坐标系108定义用于真实环境100和虚拟环境130的共享坐标空间。在示出的示例中,坐标空间具有点106处的原点。进一步地,坐标空间由相同的三个正交轴(108X、108Y、108Z)来定义。因此,真实环境100中的第一位置和对应的虚拟环境130中的第二位置可以相对于相同坐标系来描述。这简化了标识和显示真实环境和虚拟环境中的对应的位置,因为相同坐标可以用于标识这两个位置。然而,在一些示例中,对应的真实环境和虚拟环境不需要使用共享坐标空间。例如,在一些示例中(未示出),矩阵(其可以包括平移矩阵和四元数矩阵或其他旋转矩阵)或其他适合的表示可以表征真实环境坐

标空间与虚拟环境坐标空间之间的变换。

[0033] 图1C示出了经由混合现实系统112同时向用户呈现真实环境100和虚拟环境130的各方面的示例MRE 150。在示出的示例中，MRE 150同时向用户110呈现来自真实环境100的真实对象122A、124A、126A和128A（例如，经由混合现实系统112的显示器的透射部分）；以及来自虚拟环境130的虚拟对象122B、124B、126B和132（例如，经由混合现实系统112的显示器的有源显示部分）。如上文，原点106充当用于与MRE 150相对应的坐标空间的原点，并且坐标系108定义用于坐标空间的x轴、y轴和z轴。

[0034] 在示出的示例中，混合现实对象包括占用坐标空间108中的对应位置的对应的真实对象和虚拟对象对（即，122A/122B、124A/124B、126A/126B）。在一些示例中，真实对象和虚拟对象二者可以对用户110同时可见。这可以在例如虚拟对象呈现被设计为增加对应的真实对象的视图的信息的实例中（诸如在虚拟对象呈现缺失的古代损坏雕塑件的博物馆应用中）是期望的。在一些示例中，虚拟对象（122B、124B和/或126B）可以被显示（例如，经由使用像素化遮挡快门的有源像素化遮挡）以便遮挡对应的真实对象（122A、124A和/或126A）。这可以在例如虚拟对象充当用于对应的真实对象的视觉替换的实例中（诸如在无生命真实对象变为“活的”角色的交互式叙事应用中）是期望的。

[0035] 在一些示例中，真实对象（例如，122A、124A、126A）可以与虚拟内容或可以不必构成虚拟对象的辅助数据（helper data）相关联。虚拟内容或辅助数据可以促进混合现实环境中的虚拟对象的处理或处置。例如，这样的虚拟内容可以包括以下各项的二维表示：对应的真实对象；与对应的真实对象相关联的自定义资产类型；或与对应的真实对象相关联的统计数据。该信息可以使能或者促进涉及真实对象的计算而不招致不必要的计算开销。

[0036] 在一些示例中，上文所描述的呈现还可以包含音频方面。例如，在MRE 150中，虚拟怪物132可以与一个或多个音频信号相关联，诸如当怪物在MRE 150到处走时生成的脚步声效果。如下文进一步描述的，混合现实系统112的处理器可以计算对应于MRE 150中的所有此类声音的混合和处理复合的音频信号，并且经由被包括在混合现实系统112中的一个或多个扬声器和/或一个或多个外部扬声器向用户110呈现音频信号。

[0037] 示例混合现实系统

[0038] 示例混合现实系统112可以包括可穿戴头部设备（例如，可穿戴增强现实或混合现实头部设备），其包括：显示器（其可以包括左透射式显示器和右透射式显示器，其可以是近眼显示器，以及用于将来自显示器光耦合到用户的眼睛的相关联的部件）；左扬声器和右（例如，其分别邻近用户的左和右耳被定位）；惯性测量单元（IMU）（例如，其被安装到头部设备的支撑臂）；正交线圈电磁接收器（例如，其被安装到左支撑件）；左相机和右相机（例如，深度（飞行时间）相机），其被远离用户取向；以及左眼相机和右眼相机，其朝向用户取向（例如，用于检测用户的眼运动）。然而，混合现实系统112可以包含任何适合的显示技术，以及任何适合的传感器（例如，光学、红外、声学、LIDAR、EOG、GPS、磁性）。另外，混合现实系统112可以包含与其他设备和系统（包括其他混合现实系统）通信的网络特征（例如，Wi-Fi能力）。混合现实系统112还可以包括电池（其可以被安装在辅助单元中，诸如被设计为穿戴在用户的腰部周围的腰带包）、处理器和存储器。混合现实系统112的可穿戴头部设备可以包括跟踪部件，诸如IMU或其他适合的传感器，其被配置为输出可穿戴头部设备相对于用户的环境的一组坐标。在一些示例中，跟踪部件可以向执行同时定位和地图创建（SLAM）和/或视觉里

程计算法的处理器提供输入。在一些示例中,混合现实系统112还可以包括手持式控制器300和/或辅助单元320,其可以是可穿戴腰带包,如下文进一步描述的。

[0039] 图2A-2D示出了可以用于向用户呈现MRE(其可以对应于MRE 150)或其他虚拟环境的示例混合现实系统200(其可以对应于混合现实系统112)的部件。图2A示出了在示例混合现实系统200中包括的可穿戴头部设备2102的透视图。图2B示出了在用户的头部2202上穿戴的可穿戴头部设备2102的俯视图。图2C示出了可穿戴头部设备2102的前视图。图2D示出了可穿戴头部设备2102的示例目镜2110的侧视图。如图2A-2C所示,示例可穿戴头部设备2102包括示例左目镜(例如,左透明波导集目镜)2108和示例右目镜(例如,右透明波导集目镜)2110。每个目镜2108和2110可以包括:透射元件,通过该透射元件,真实环境可以是可见的;以及显示元件,其用于呈现重叠真实环境的显示(例如,经由成像调制光)。在一些示例中,这样的显示元件可以包括用于控制成像调制光流的表面衍射光学元件。例如,左目镜2108可以包括左耦入光栅集2112、左正交光瞳扩展(OPE)光栅集2120和左出射(输出)光瞳扩展(EPE)光栅集2122。类似地,右目镜2110可以包括右耦入光栅集2118、右OPE光栅集2114和右EPE光栅集2116。成像调制光可以经由耦入光栅2112和2118、OPE 2114和2120、和EPE 2116和2122传递到用户的眼睛。每个耦入光栅集2112、2118可以被配置为朝向其对应的OPE光栅集2120、2114偏转光。每个OPE光栅集2120、2114可以被设计为朝向其相关联的EPE 2122、2116向下递增地偏转光,从而水平延伸形成的出射光瞳。每个EPE 2122、2116可以被配置为将从其对应的OPE光栅集2120、2114接收的光的至少一部分向外递增地重引导到定义在目镜2108、2110后面的用户眼动范围(eyebow)位置(未示出),垂直延伸在眼动范围处形成的出射光瞳。可替代地,代替耦入光栅集2112和2118、OPE光栅集2114和2120、和EPE光栅集2116和2122,目镜2108和2110可以包括光栅和/或用于控制将成像调制光耦合到用户的眼睛的折射和反射特征的其他布置。

[0040] 在一些示例中,可穿戴头部设备2102可以包括左支撑臂2130和右支撑臂2132,其中,左支撑臂2130包括左扬声器2134并且右支撑臂2132包括右扬声器2136。正交线圈电磁接收器2138可以被定位在左支撑件中,或者在可穿戴头部单元2102中的另一适合的位置中。惯性测量单元(IMU)2140可以被定位在右边撑臂2132中,或者在可穿戴头部设备2102中的另一适合的位置中。可穿戴头部设备2102还可以包括左深度(例如,飞行时间)相机2142和右深度相机2144。深度相机2142、2144可以在不同的方向上适合地取向以便一起覆盖更宽的视场。

[0041] 在图2A-2D中示出的示例中,左成像调制光源2124可以通过左耦入光栅集2112被光学耦合到左目镜2108中,并且右成像调制光源2126可以通过右耦入光栅集2118被光学耦合到右目镜2110中。成像调制光源2124、2126可以包括例如光纤扫描器;投影仪,包括电子光调制器,诸如数字光处理(DLP)芯片或硅上液晶(LCoS)调制器;或发射显示器,诸如微发光二极管( $\mu$ LED)或微有机发光二极管( $\mu$ OLED)面板,其每侧使用一个或多个透镜被耦合到耦入光栅集2112、2118中。输入耦合光栅集2112、2118可以将来自成像调制光源2124、2126的光偏转到大于针对目镜2108、2110的全内反射(TIR)的临界角的角度。OPE光栅集2114、2120朝向EPE光栅集2116、2122递增地偏转通过TIR传播的光。EPE光栅集2116、2122将光朝向用户的面部,包括用户的眼睛的瞳孔,递增地耦合。

[0042] 在一些示例中,如图2D所示,左目镜2108和右目镜2110中的每一个包括多个波导

2402。例如,每个目镜2108、2110可以包括多个单独波导,每个波导专用于相应的颜色通道(例如,红色、蓝色和绿色)。在一些示例中,每个目镜2108、2110可以包括这样的波导的多个集,其中,每个集被配置为向发射光给予不同的波前曲率。波前曲率可以相对于用户的眼睛是凸的,例如以呈现定位在用户的前面一定距离的虚拟对象(例如,通过对应于波前曲率的倒数的距离)。在一些示例中,EPE光栅集2116、2122可以包括通过变更跨每个EPE出射光的坡印廷矢量来实现凸波前曲率的弯曲光栅凹陷。

[0043] 在一些示例中,为了创建所显示的内容是三维的感知,可以通过成像光调制器2124、2126和目镜2108、2110向用户呈现立体调节的左眼影像和右眼影像。三维虚拟对象的呈现的感知的逼真感可以通过选择波导(以及因此对应的波前曲率)增强,使得虚拟对象在近似由立体的左右图像指示的距离的距离处被显示。该技术还可以减少一些用户经历的运动病,其可以由立体的左右眼影像提供的深度感知线索与人眼的自动调节(例如,对象距离-相关焦点)之间的差引起。

[0044] 图2D示出了从示例可穿戴头部设备2102的右目镜2110的顶部的面向边缘的视图。如图2D所示,多个波导2402可以包括三个波导2404的第一子集和三个波导2406的第二子集。波导2404、2406的两个子集可以通过向出射光给予不同波前曲率的不同光栅线曲率为特征的不同EPE光栅来区分。在波导2404、2406的子集中的每一个内,每个波导可以用于将不同光谱信道(例如,红色、绿色和蓝色光谱信道之一)耦合到用户的右眼2206。(虽然未在图2D中示出,但是左目镜2108的结构类似于右目镜2110的结构。)

[0045] 图3A示出了混合现实系统200的示例手持式控制器部件300。在一些示例中,手持式控制器300包括柄部346和沿着顶部表面348设置的一个或多个按钮350。在一些示例中,按钮350可以被配置用于用作光学跟踪目标,例如,用于结合相机或其他光学传感器(其可以被安装在混合现实系统200的头部单元(例如,可穿戴头部设备2102)中)跟踪手持式控制器300的六自由度(6DOF)运动。在一些示例中,手持式控制器300包括用于检测位置或取向(诸如相对于可穿戴头部设备2102的位置或取向)的跟踪部件(例如,IMU或其他适合的传感器)。在一些示例中,这样的跟踪部件可以被定位在手持式控制器300的手柄中,和/或可以被机械耦合到手持式控制器。手持式控制器300可以被配置为提供与按钮的按压状态中的一个或多个相对应的一个或多个输出信号;或手持式控制器300的位置、取向和/或运动(例如,经由IMU)。这样的输出信号可以用作混合现实系统200的处理器输入。这样的输入可以与手持式控制器的位置、取向和/或运动(例如,通过扩展,握持控制器的用户的手部的位置、取向和/或运动)。这样的输入还可以对应于用户按钮350。

[0046] 图3B示出了混合现实系统200的示例辅助单元320。辅助单元320可以包括提供操作系统200的能量的电池,并且可以包括用于执行操作系统200的程序的处理器。如图所示,示例辅助单元320包括芯片2128,诸如用于将辅助单元320附接到用户的腰带。其他形状因子适合于辅助单元320并且将是明显的,包括不涉及将单元安装到用户的腰带的形状因子。在一些示例中,辅助单元320通过多导管电缆耦合到可穿戴头部设备2102,该多导管电缆可以包括例如电线和光纤。还可以使用辅助单元320与可穿戴头部设备2102之间的无线连接。

[0047] 在一些示例中,混合现实系统200可以包括检测声音并且将对应的信号提供给混合现实系统的一个或多个麦克风。在一些示例中,麦克风可以被附接到可穿戴头部设备2102或与其集成,并且被配置为检测用户的语音。在一些示例中,麦克风可以被附接到手持

式控制器300和/或辅助单元320或与其集成。这样的麦克风可以被配置为检测环境声音、环境噪声、用户或第三方的语音或其他声音。

[0048] 图4示出了可以对应于示例混合现实系统,诸如上文所描述的混合现实系统200(其可以对应于相对于图1的混合现实系统112),的示例功能框图。如图4所示,示例手持式控制器400B(其可以对应于手持式控制器300(“图腾”))包括图腾到可穿戴头部设备六自由度(6DOF)图腾子系统404A,并且示例可穿戴头部设备400A(其可以对应于可穿戴头部设备2102)包括图腾到可穿戴头部设备6DOF子系统404B。在示例中,6DOF图腾子系统404A和6DOF子系统404B合作确定手持式控制器400B相对于可穿戴头部设备400A的六个坐标(例如,在三个平移方向上的偏移和沿着三个轴的旋转)。六个自由度可以相对于可穿戴头部设备400A的坐标系表示。三个平移偏移可以表示为这样的坐标系中的X、Y和Z偏移、平移矩阵、或某种其他表示。旋转自由度可以表示为偏航、俯仰和滚动旋转的序列、旋转矩阵、四元数或某种其他表示。在一些示例中,可穿戴头部设备400A;包括在可穿戴头部设备400A中的一个或多个深度相机444(和/或一个或多个非深度相机);和/或一个或多个光学目标(例如,如上文所描述的手持式控制器400B的按钮450,或包括在手持式控制器400B中的专用光学目标)可以用于6DOF跟踪。在一些示例中,手持式控制器400B可以包括如上文所描述的相机;并且可穿戴头部设备400A可以包括用于结合相机光学跟踪的光学目标。在一些示例中,可穿戴头部设备400A和手持式控制器400B各自包括一组三个正交取向的螺线管,其用于无线地发送和接收三个可区分的信号。通过测量用于接收的线圈中的每一个中接收的三个可区分信号的相对大小,可以确定可穿戴头部设备400A相对于手持式控制器400B的6DOF。此外,6DOF图腾子系统404A可以包括惯性测量单元(IMU),该惯性测量单元(IMU)可用于提供关于手持式控制器400B的快速运动的经改进的准确度和/或更及时的信息。

[0049] 在一些示例中,可能变得需要将来自局部坐标空间(例如,相对于可穿戴头部设备400A固定的坐标空间)的坐标变换到惯性坐标空间(例如,相对于真实环境固定的坐标空间),例如以便补偿可穿戴头部设备400A相对于坐标系108的运动。例如,这样的变换可能为可穿戴头部设备400A的显示器进行以下内容所必需:将虚拟对象呈现在相对于真实环境的期望位置和取向处(例如,坐在真实椅子中、面向前的虚拟人,而不管可穿戴头部设备的位置和取向),而不是在显示器上的固定位置和取向处(例如,在显示器的右下角的相同位置处),以保持虚拟对象存在于真实环境中的错觉(并且例如当可穿戴头部设备400A移动和旋转时不显得不自然地定位在真实环境中)。在一些示例中,坐标空间之间的补偿变换可以通过使用SLAM和/或视觉里程计程序处理来自深度相机444的影像确定以便确定可穿戴头部设备400A相对于坐标系108的变换。在图4所示的示例中,深度相机444被耦合到SLAM/视觉里程计块406并且可以向块406提供影像。SLAM/视觉里程计块406实施方式可以包括处理器,该处理器被配置为处理该影像并且确定用户的头部的位置和取向,其然后可以用于标识头部坐标空间与另一坐标空间(例如,惯性坐标空间)之间的变换。类似地,在一些示例中,从IMU 409获得关于用户的头部姿势和位置的信息的附加源。来自IMU 409的信息可以与来自SLAM/视觉里程计块406的信息集成以提供关于用户的头部姿势和位置的快速调节的经改进的准确度和/或更及时的信息。

[0050] 在一些示例中,深度相机444可以将3D影像供应到手势跟踪器411,该手势跟踪器411可以被实现在可穿戴头部设备400A的处理器中。手势跟踪器411可以标识用户的手势,

例如通过将深度相机444接收的3D影像与表示手势的存储的图案匹配。标识用户的手势的其他适合的技术将是明显的。

[0051] 在一些示例中,一个或多个处理器416可以被配置为从可穿戴头部设备的6DOF头盔子系统404B、IMU 409、SLAM/视觉里程计块406、深度相机444和/或手势跟踪器411接收数据。处理器416还可以发送和接收来自6DOF图腾系统404A的控制信号。处理器416可以被无线耦合到6DOF图腾系统404A,诸如在不限于手持式控制器400B的示例中。处理器416还可以与附加部件通信,诸如音频-视觉内容存储器418、图形处理单元(GPU) 420、和/或数字信号处理器(DSP) 音频空间化器。DSP音频空间化器422可以被耦合到头部相关传递函数(HRTF)存储器425。GPU 420可以包括被耦合到左成像调制光源424的左信道输出和被耦合到右成像调制光源426的右信道输出。GPU 420可以将立体图像数据输出到成像调制光源424、426,例如如上文相对于图2A-2D所描述的。DSP音频空间化器422可以向左扬声器412和/或右扬声器414输出音频。DSP音频空间化器422可以从处理器419接收指示从用户到虚拟声源(其可以由用户移动,例如,经由手持式控制器320)的方向矢量的输入。基于该方向矢量,DSP音频空间化器422可以确定对应的HRTF(例如,通过访问HRTF、或通过内插多个HRTF)。DSP音频空间化器422然后将所确定的HRTF应用到音频信号,诸如与由虚拟对象生成的虚拟声音相对应的音频信号。这可以通过内插用户相对于混合现实环境中的虚拟声音的相对位置和取向——即,通过呈现匹配该虚拟声音听起来就像真实环境中的真实声音的用户的期望的虚拟声音——来提高虚拟声音的可信度和逼真感。

[0052] 在一些示例中,诸如图4所示,处理器416、GPU 420、DSP音频空间化器422、HRTF存储器425和音频/视觉内容存储器418中的一个或多个可以被包括在辅助单元400C中(其可以对应于上文所描述的辅助单元320)。辅助单元400C可以包括对其部件供电和/或向可穿戴头部设备400A或手持式控制器400B供电的电池427。将这样的部件包括在可被安装到用户的腰部的辅助单元中可以限制可穿戴头部设备400A的大小和重量,其进而可以减少用户的头部和颈部的疲劳。

[0053] 虽然图4呈现了与示例混合现实系统的各种部件相对应的元件,但是这些部件的各种其他适合的布置将对于本领域技术人员来说变得明显。例如,在图4中呈现为与辅助单元400C相关联的元件可以相反地与可穿戴头部设备400A或手持式控制器400B相关联。此外,一些混合现实系统可以完全放弃手持式控制器400B或辅助单元400C。这样的改变和修改将被理解为被包括在所公开的示例的范围内。

[0054] 环境声学持久性

[0055] 如上文所描述的,MRE(诸如经由混合现实系统,例如混合现实系统112,体验的,其可包括诸如上文所描述的可穿戴头部单元200、手持式控制器300或辅助单元320的部件)可以向MRE的用户呈现看起来在具有MRE的原点坐标的声源处起源的音频信号。即,用户可以感知这些音频信号,就像它们是源自声源的原点坐标的真实音频信号。

[0056] 在一些情况下,音频信号可以被认为是虚拟的,因为它们对应于虚拟环境中的计算信号。虚拟音频信号可以作为可由人耳检测到的真实音频信号向用户呈现,例如如经由图2中的可穿戴头部单元200的扬声器2134和2136生成的。

[0057] 声源可以对应于真实对象和/或虚拟对象。例如,虚拟对象(例如,图1C的虚拟怪物132)可以发射MRE中的音频信号,该音频信号在MRE中被表示为虚拟音频信号,并且作为真

实音频信号向用户呈现。例如,图1C的虚拟怪物132可以发射与怪物的语音(例如,对话)或声音效果相对应的虚拟声音。类似地,可以使真实对象(例如,图1C的真实对象122A)看起来发射MRE中的虚拟音频信号,该虚拟音频信号在MRE中被表示为虚拟音频信号,并且作为真实音频信号向用户呈现。例如,真实灯122A可以发射与灯被接通或关断的声音效果相对应的虚拟声音——即使灯在真实环境中未被接通或关断。虚拟声音可以与声源(无论真实的还是虚拟的)的位置和取向相对应。例如,如果作为真实音频信号(例如,经由扬声器2134和2136)向用户虚拟声音呈现,则用户可以将虚拟声音感知为源自声源的位置。声源在本文中被称为“虚拟声源”,即使明显地发射声音的底层对象可以自己对应于真实或虚拟对象,诸如上文所描述的。

[0058] 一些虚拟或混合现实环境存在环境不感觉真实或可信的感知的问题。针对该感知的一个原因在于,音频和视觉线索在这样的环境中不总是彼此匹配。例如,如果用户被定位在MRE中的大砖墙后面,则用户可能期望来自砖墙后面的声音比起源于用户旁边的声音更安静并且更听不清。该期望基于现实世界中用户的听觉体验,其中,声音在其通过大而密集对象时变得安静并且听不清。当用户被呈现有据称源自砖墙后面但是强劲并以全音量呈现的音频信号时,声音源自砖墙后面的错觉被损害。整个虚拟体验可能感觉假并且不真实,部分是因为它不基于现实世界交互适合用户的期望。进一步地,在一些情况下,“恐怖谷”问题出现,其中,虚拟体验与真实体验之间的更细微的差异可能引起升高的不舒适的感觉。通过在MRE中呈现看起来逼真地与用户的环境中的对象交互——甚至以细微的方式——的音频信号改善用户的体验是期望的。这样的音频信号与用户的期望越一致,基于真实世界经验,用户的MRE体验可以越沉浸并且啮合。

[0059] 用户感知和理解他们周围的环境的一种方式是通过音频线索。在真实世界中,用户听到的真实音频信号受那些音频信号源自何处并且音频信号与什么对象相互作用影响。例如,在所有其他因素相同的情况下,在源自距用户长距离的声音(例如,在远处的狗叫)将显得比起源自短距离的相同声音(例如,与用户相同房间中的狗叫)更安静。因此,用户可以部分基于狗叫的感知音量,识别真实环境中的狗的位置。同样地,在所有其他因素相同的情况下,远离用户行进的声音(例如,远离用户的人的语音)将比朝向用户行进的相同声音(例如,面向用户的人的语音)显得更不清楚并更听不清(即,低通滤波的)。因此,用户可以基于该人的语音的感知频率特性,识别真实环境中的人的取向。

[0060] 真实音频信号的用户感知还可能受音频信号与之相互作用的环境中的对象的存在影响。即,用户可以不仅感知由声源生成的音频信号,而且感知该音频信号在附近对象的反射以及由周围声学空间赋予的混响特征。例如,如果用户在具有封闭墙的小房间中说话,则那些墙可能引起短的自然混响信号产生,因为人的语音从墙反射。用户可以根据那些混响推断他们处于具有封闭墙的小房间中。同样地,大音乐厅或教堂可能引起更长的混响,根据该混响,用户可以推断他们处于大的宽敞的房间中。类似地,音频信号的混响可以基于那些信号反射的表面的位置或取向、或那些表面的材料,呈现各种声特性。例如,针对拼接墙的混响将听起来与针对砖、地毯、干式墙或其他材料的混响不同。这些混响特性可以由用户用于理解——声学上——他们居住的空间的大小、形状和材料组成。

[0061] 以上示例示出了音频线索可以如何通知用户的感知他们周围的环境。这些线索可以结合视觉线索起作用:例如,如果我们看到远处的狗,则我们期望狗叫的声音与该距离一

致(并且如果其不是,则我们可能感觉困惑或迷惑,如在一些虚拟环境中)。在一些示例中,诸如在低光环境中,或者相对于视觉障碍用户,视觉线索可能是有限或不可用的;在这样的情况下,音频线索可以呈现特定重要性,并且可以用作用户理解他们的环境的主要手段。

[0062] 系统架构可能有利于组织、存储、调用和/或管理呈现逼真的虚拟音频所需的信息。例如,MR系统(例如,MR系统112、200)可以管理环境信息,例如用户可能处于什么真实环境中、真实环境可能具有什么声学特性、和/或用户可能位于该真实环境中的什么位置。MR系统可以进一步管理关于真实和/或虚拟环境中的对象(例如,可能影响真实环境的一般声学特性的对象和/或可能影响与对象交互的虚拟声源的声学特性的对象)的信息。MR系统还可以管理关于虚拟声源的信息。例如,虚拟声源位于何处可能与渲染逼真的虚拟音频相关。

[0063] 除了管理虚拟音频系统之外,可能还需要同时管理其他系统以呈现完整的MR体验。例如,完整的MR体验可能需要虚拟视觉系统,该系统可以管理用于渲染虚拟对象的信息。完整的MR体验可能需要同时定位和建图系统(“SLAM”),它可以构建、更新和/或维护用户环境的三维模型。除了虚拟音频系统之外,MR系统(例如,MR系统112、200)可以管理这些系统和更多系统,以呈现完整的MR体验。虚拟音频系统架构可能有助于管理这些系统之间的交互以促进数据传送、管理、存储和/或保全。

[0064] 在一些实施例中,系统(例如,虚拟音频系统)可以与其他更高级别的系统交互。在一些实施例中,较低级别的系统(例如,虚拟音频系统)可以与硬件级别的输入和/或输出更紧密地交互,而较高级别的系统(例如,应用)可以与较低级别的系统交互。较高级别的系统可以利用较低级别的系统来执行其功能(例如,游戏应用可能依赖于较低级别的虚拟音频系统来渲染逼真的虚拟音频)。虚拟音频系统可以受益于被设计为管理与较高级别系统的交互同时维护虚拟音频系统的完整性的系统架构。例如,多个较高级别的系统(例如,多个第三方应用)可以同时或基本上同时与虚拟音频系统相互交流。在一些实施例中,维护可以渲染虚拟音频的单个虚拟音频系统可能比让每个较高级别的系统维护单独的音频系统在计算上更有效。例如,在一些实施例中,单个数字混响器可用于处理来自多个较高级别系统的声音对象,当那些对象旨在位于相同的虚拟或真实声学空间(例如,用户所在的房间)中时。设计良好的系统架构还可以保护可能在其他应用中使用的信息的完整性(例如,防止数据损坏和/或篡改)。

[0065] 在一些实施例中,设计系统架构以使得可以实时做出改变而不中断对其他系统(例如,较高级别的系统)的服务可能是有利的。例如,虚拟音频系统可以存储、维护或以其他方式管理考虑真实环境(例如房间)的声学特性的音频模型。如果用户改变真实环境(例如,移动到不同的房间),则音频模型可能被更新以考虑真实环境中的变化。如果MR系统当前正在使用中(例如,MR系统正在向用户呈现虚拟视觉和/或虚拟音频),则当不同系统(例如,较高级别的系统)仍在使用音频模型来渲染虚拟音频时,可能需要更新音频模型。

[0066] 在一些实施例中,设计系统架构以在其他系统中传播改变可能是有利的。例如,一些系统可以维护音频模型的单独副本,或者一些系统可以存储使用由虚拟音频系统维护的音频模型渲染的特定重复声音效果。因此,将在虚拟音频系统中做出的改变传播到其他系统可能是有利的。例如,如果用户改变环境(例如,移动房间)并且新的音频模型可能更准确,则虚拟音频系统可以修改其音频模型并通知任何客户端(例如,使用和/或依赖虚拟音频系统的系统)改变。在一些实施例中,客户端然后可以查询虚拟音频系统并且相应地更新

内部数据。

[0067] 图5示出了根据一些实施例的示例性虚拟音频系统。虚拟音频系统500可包括持久性模块502。模块(例如,持久性模块502)可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。在一些实施例中,模块(例如,持久性模块502)可以被配置为执行由音频服务522管理的进程、子进程、线程和/或服务(例如,由持久性模块502执行的指令可以在音频服务522内运行),音频服务522可以在一个或多个计算机系统上运行。在一些实施例中,音频服务522可以是可以在运行时环境中运行的进程,并且由模块(例如,持久性模块502)执行的指令可以是音频服务522的部件(例如,由持久性模块502执行的指令可以是音频服务522的子进程)。在一些实施例中,音频服务522可以是父进程(parent process)的子进程。由模块(例如,持久性模块502)执行的指令可包括一个或多个部件(例如,由定位状态子模块506、声学数据子模块508和/或音频模型子模块510执行的进程、子进程、线程和/或服务)。在一些实施例中,由模块(例如,持久性模块502)执行的指令可以作为音频服务522的子进程和/或作为与音频服务522的其他部件不同的位置中的单独进程运行。例如,由模块(例如,持久性模块502)执行的指令可以在通用处理器中运行,并且音频服务522的一个或多个其他部件可以在音频专用处理器(例如,DSP)中运行。在一些实施例中,由模块(例如,持久性模块502)执行的指令可以在与音频服务522的其他部件不同的进程地址空间和/或存储器空间中运行。在一些实施例中,由模块(例如,持久性模块502)执行的指令可以作为音频服务522内的一个或多个线程运行。在一些实施例中,由模块(例如,持久性模块502)执行的指令可以在音频服务522内实例化。在一些实施例中,由模块(例如,持久性模块502)执行的指令可以与音频服务522的其他部件共享进程地址和/或存储器空间。

[0068] 在一些实施例中,持久性模块502可包括定位状态子模块506。定位状态子模块506可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,由定位状态子模块506执行的指令可以是持久性模块502的子进程。在一些实施例中,定位状态子模块506可以指示是否已经实现定位(例如,定位状态子模块506可以指示MR系统是否已经识别了真实环境和/或将其自身定位在真实环境中)。在一些实施例中,定位状态子模块506可以与定位系统相互交流(例如,经由API)。定位系统可以确定用于MR系统(和/或使用MR系统的用户)的位置。在一些实施例中,定位系统可以利用类似SLAM的技术来创建真实环境的三维模型并估计系统(和/或用户)在环境内的位置。在一些实施例中,定位系统可以依赖于可联结的世界系统(下文进一步详细描述)和一个或多个传感器(例如,MR系统112、200的一个或多个传感器)来估计MR系统(和/或用户)在环境内的位置。在一些实施例中,定位状态子模块506可以查询定位系统以确定当前是否实现了用于MR系统的定位。类似地,定位系统可以通知定位状态子模块506成功定位。(例如,MR系统112、200的)定位状态可用于确定是否应该更新音频模型(例如,因为用户的真实环境已经改变)。

[0069] 在一些实施例中,持久性模块502可包括声学数据子模块508。声学数据子模块508可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,由声学数据子模块508执行的指令可以是持久性模块502的子进程。在一些实施例中,声学数据子模块508可以存储一个或多个数据结构,该数据结构表示可用于创建音频模型的声学数据。在一些实施例中,声学数据子模块508可以与可联结的世界系统相互交流(例如,经由API)。可联结的世界系统可以包括关于已知真实环境(例如,房间、建筑物和/或外部空

间)和相关联的真实和/或虚拟对象的信息。在一些实施例中,可联结的世界系统可以包括持久坐标系和/或锚点。持久坐标系和/或锚点可以是MR系统可能已知的空间中固定的点(例如,通过唯一标识符)。虚拟对象可以相对于一个或多个持久坐标系和/或锚点定位以使用能对象持久性(例如,虚拟对象可以看起来保持在真实环境中的相同位置而不管谁在查看虚拟对象并不管用户的任何运动)。当具有单独MR系统的两个或两个以上用户使用不同的世界坐标系(例如,每个用户的位置被指定为用于其相应世界坐标系的原点)时,持久坐标系和/或锚点可能尤其有利。跨用户的对象持久性可以通过在单独世界坐标系和通用持久性坐标系之间进行转换以及相对于持久性坐标系放置/引用虚拟对象来实现。在一些实施例中,可联结的世界系统可以通过例如对新区域建图和创建新的持久坐标系来管理和维护持久坐标系;通过重新对已知区域建图并将新的持久坐标系与先前确定的持久坐标系相协调;和/或将持久坐标系与可识别信息(例如,位置和/或附近的对象)相关联。在一些实施例中,声学数据子模块508可以向单独的系统(例如,可联结的世界系统)查询一个或多个持久坐标系。在一些实施例中,声学数据子模块508可以检索相关的持久坐标系(例如,用户位置的阈值半径内的持久坐标系)以促进访问和管理,包括相关联的声学数据的创建、修改和/或删除。

[0070] 在一些实施例中,存储在声学数据子模块508中的声学数据可以被组织成物理相关的模块化单元(例如,房间可以由模块化单元表示,并且房间内的椅子可以由另一个模块化单元表示)。例如,模块化单元可以包括物理环境(例如,房间)的物理和/或感知相关属性。物理和/或感知相关属性可包括可能影响房间声学特性的属性(例如,房间的尺寸和/或形状)。在一些实施例中,物理和/或感知相关属性可以包括功能和/或行为属性,这些属性可以由渲染引擎解释(例如,房间外部的源是否应该被遮挡)。在一些实施例中,物理和/或感知相关属性可以包括已知和/或识别对象的属性。例如,固定(例如,地板、墙壁、家具等)和/或可移动(例如,杯子)对象的几何形状可以被存储为物理和/或感知相关属性,并且可以与特定环境相关联。在一些实施例中,物理和/或感知相关属性可包括传输损耗、散射系数和/或吸收系数。在一些实施例中,模块化单元可以包括物理和/或感知相关的链接(例如,在其他模块化单元之间或在模块化单元内)。例如,物理和/或感知相关的链接可以将两个或两个以上房间链接在一起,并描述房间可以如何彼此交互(例如,模拟房间的数字混响器之间的交叉耦合增益水平和/或两个空间之间的视线路径)。

[0071] 在一些实施例中,物理和/或感知相关属性可以包括声学属性,例如混响时间、混响延迟和/或混响增益。混响时间可包括声音衰减一定量(例如,60分贝)所需的时间长度。声音衰减可能是声音从真实环境中的表面(例如,墙壁、地板、家具等)反射的结果,同时由于例如房间边界(例如,墙壁、地板、天花板等)、房间内的对象(例如,椅子、家具、人等)以及房间内的空气引起的声音吸收而损失能量。混响时间可能受环境因素影响。例如,吸收性表面(例如,软垫)除了几何扩展之外还可以吸收声音,并且因此可以减少混响时间。在一些实施例中,可能不需要具有关于起源的信息来估计环境的混响时间。混响增益可以包括声音的直接/源/原始能量与声音的混响能量(例如,由直接/源/原始声音产生的混响的能量)的比率,其中,听众和源基本上共同定位(例如,用户可以拍手,产生可以被认为与安装在头戴式MR系统上的一个或多个麦克风基本上共同定位的源声音)。例如,脉冲(例如,拍手)可以具有与该脉冲相关联的能量,并且来自该脉冲的混响声音可以具有与该脉冲的混响相关联

的能量。原始/源能量与混响能量的比率可以是混响增益。真实环境的混响增益可能受到例如可以吸收声音并从而降低混响能量的吸收性表面的影响。

[0072] 在一些实施例中,声学数据可包括元数据(例如,物理和/或感知相关属性的元数据)。例如,关于何时和/或何处收集声学数据的信息可以被包括在声学数据中。在一些实施例中,与声学数据相关联的置信度数据(例如,估计的测量准确度和/或重复测量的计数)可以被包括作为元数据。在一些实施例中,一种类型的模块化单元(例如,用于房间的模块化单元或模块化单元之间的链接)和/或数据版本可以被包括作为元数据。在一些实施例中,与声学数据、持久坐标系和/或锚点相关联的唯一标识符可以被包括作为元数据。在一些实施例中,来自持久坐标系和/或锚点和相关联的虚拟对象的相对变换可以被包括作为元数据。在一些实施例中,元数据可以作为单个束与声学数据一起存储。

[0073] 在一些实施例中,声学数据可以由持久坐标系和/或锚点来组织,并且持久坐标系和/或锚点可以被组织成地图。在一些实施例中,音频模型可以考虑由持久坐标系和/或锚点组织的声学数据,其可以对应于环境内的位置。在一些实施例中,声学数据可以在成功的定位事件(这可以由定位状态子模块506指示)时被加载到声学数据子模块508中。在一些实施例中,所有可用的声学数据可以被加载到声学数据子模块508中。在一些实施例中,仅相关的声学数据可以被加载到声学数据子模块508中(例如,用于持久坐标系和/或在MR系统的位置的特定距离内的锚点的声学数据)。

[0074] 在一些实施例中,声学数据可以包括可以根据真实环境的变化而变化的不同状态。例如,可以表示房间的给定模块化单元的声学数据可以包括处于空状态的房间的声学数据和处于占用状态的房间的声学数据。在一些实施例中,房间家具布置的变化可以通过与房间相关联的声学数据的状态变化来反映。在一些实施例中,模块化单元(例如,表示房间)可以包括用于门打开或关闭时的状态的不同声学数据。状态可以表示为二进制值(例如,0或1)或连续值(例如,门的打开程度、房间的占用情况等)。

[0075] 在一些实施例中,持久性模块502可包括音频模型子模块510。音频模型子模块510可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,音频模型子模块510可包括表示用于真实和/或虚拟环境的音频模型的一个或多个数据结构。在一些实施例中,音频模型可以至少部分地由存储在声学数据子模块508中的声学数据生成。音频模型可以表示声音在特定环境中如何行为。例如,可以由音频模型子模块510中的音频模型修改由MR系统(例如,MR系统112、200)生成的虚拟声音,以反映环境的声学特性。呈现给坐在大音乐厅中的用户的虚拟音乐会可能具有与呈现在相同音乐厅中的真实音乐会相似的声学特性。MR系统可以将自己定位到已标识的音乐厅,加载相关的声学数据,并生成音频模型以对音乐厅的声学属性进行建模。

[0076] 在一些实施例中,音频模型可用于对环境中的声音传播进行建模。例如,传播效应可包括遮挡、阻塞、早期反射、衍射、飞行时间延迟、多普勒效应和其他效应。在一些实施例中,音频模型可以考虑与频率相关的吸收和/或传输损耗(例如,基于加载到声学数据子模块508中的声学数据)。在一些实施例中,存储在音频模型子模块510中的音频模型可以通知音频引擎的其他方面。例如,音频模型可以使用声学数据在程序上合成音频(例如,虚拟和/或真实对象之间的碰撞)。

[0077] 在一些实施例中,音频渲染服务522可包括渲染轨道模块514。渲染轨道模块514可

包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,渲染轨道模块514可包括稍后可以呈现给用户的音频信息。在一些实施例中,MR系统可以呈现包括混合在一起的几个声源(例如,两把剑碰撞的声源和人大喊的声源)的虚拟声音。渲染轨道模块514可存储可以与其他轨道混合以呈现给用户的一个或多个轨道。在一些实施例中,渲染轨道模块514可包括关于空间源的信息。例如,渲染轨道模块514可以包括关于声源位于何处的信息,这可以在音频模型和/或渲染算法中加以考虑。在一些实施例中,渲染轨道模块514可包括关于模块化单元和/或声源之间的关系的的信息。例如,一个或多个渲染轨道和/或音频模型可以作为单个组关联在一起。

[0078] 在一些实施例中,音频渲染服务522可包括位置管理器模块516。位置管理器模块516可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,位置管理器模块516可以管理与音频引擎相关的位置信息(例如,真实环境中的MR系统的当前位置)。在一些实施例中,位置管理器模块516可包括感知包装器子模块。感知包装器子模块可以是感知数据周围的包装器(例如,MR系统已经检测到或正在检测什么)。在一些实施例中,感知包装器可以在感知数据与位置管理器模块516之间进行交流和/或转换。在一些实施例中,位置管理器模块516可包括头部姿势子模块,该头部姿势子模块可包括头部姿势数据。头部姿势数据可包括真实环境中的MR系统(或对应的用户)的位置和/或取向。在一些实施例中,可以基于感知数据确定头部姿势。

[0079] 在一些实施例中,音频渲染服务522可包括音频模型模块518。音频模型模块518可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,音频模型模块518可包括音频模型,该音频模型可以是包括在音频模型子模块510中的相同音频模型。在一些实施例中,模块510和518可以维护相同音频模型的复制副本。例如,在更新音频模型但应向用户呈现声音时,维护不止一个音频模型的副本可能是有利的。当模型当前正在使用时更新模型的副本,并且然后在过时的模型变得可用(例如,不再使用)时更新过时的模型可能是有利的。在一些实施例中,可以通过序列化在模块510和518之间传送音频模型。模块510中的音频模型可以被序列化和反序列化,以促进数据传送到模块518。序列化可以促进处理器(例如,通用处理器和音频专用处理器)之间的数据传送,以使得不需要共享类型化的存储器。

[0080] 在一些实施例中,音频渲染服务522可包括渲染算法模块520。渲染算法模块520可包括被配置为执行指令和/或存储一个或多个数据结构的一个或多个计算机系统。例如,渲染算法模块520可包括渲染虚拟声音的算法,以使得它们可以被呈现给用户(例如,经由MR系统的一个或多个扬声器)。渲染算法模块520可以考虑特定环境的音频模型(例如,模块510和/或518中的音频模型)。

[0081] 在一些实施例中,音频服务522可以是在一个或多个计算机系统(例如,在MR系统112、200中)上运行的进程、子进程、线程和/或服务。在一些实施例中,单独的系统(例如,第三方应用)可以请求呈现音频信号(例如,经由MR系统112、200的一个或多个扬声器)。这样的请求可以采取任何适合的形式。在一些实施例中,呈现音频信号的请求可包括呈现音频信号的软件指令;在一些实施例中,这样的请求可以是硬件驱动的。可以在有或没有用户参与的情况下发出请求。进一步地,可以经由本地硬件(例如,从MR系统本身)、经由外部硬件(例如,与MR系统通信的单独计算机系统)、经由互联网(例如,经由云服务器)、或经由任何

其他适合的源或源的组合来接收这样的请求。在一些实施例中，音频服务522可以接收请求，渲染所请求的音频信号（例如，通过渲染算法520，其可以考虑来自框510和/或518的音频模型），并且将请求的音频信号呈现给用户。在一些实施例中，音频服务522可以是在MR系统的操作系统正在运行时连续运行（例如，在后台）的进程。在一些实施例中，音频服务522可以是父后台服务的实例，其可以用作一个或多个后台进程和/或子进程的主机进程。在一些实施例中，音频服务522可以是MR系统的操作系统的一部分。在一些实施例中，可以在MR系统上运行的应用可以访问音频服务522。在一些实施例中，MR系统的用户可能不直接向音频服务522提供输入。例如，用户可以向在MR系统上运行的应用（例如，角色扮演游戏）提供输入（例如，移动命令）。应用可以向音频服务522提供输入（例如，以渲染脚步声），并且音频服务522可以向用户（例如，经由扬声器）和/或其他过程和/或服务提供输出（例如，所渲染的脚步声）。

[0082] 图6示出了根据一些实施例的用于更新音频模型的示例性过程。在步骤606处，可以确定定位（例如，MR系统可以成功地识别其在环境内的位置）。在可发生在持久性模块602（其可对应于持久性模块502）内的步骤607处，可以发出成功定位的通知。在一些实施例中，成功定位的通知可以触发更新音频模型的过程（例如，因为先前的音频模型可能不再适用于当前位置）。

[0083] 在步骤608处，可以确定是否启动声学数据的调用。可能希望设置一个或多个用于启动调用的条件，使得音频模型不太频繁地更新。例如，如果使用MR系统的用户仅在房间内轻微移动，则可能不希望更新音频模型（例如，因为更新的模型在感知上可能无法与现有模型区分，和/或持续更新模型可能在计算上很昂贵）。在一些实施例中，步骤608处的阈值条件可以基于时间。例如，仅当在前5秒内尚未启动调用时，才可以启动调用。在一些实施例中，步骤608处的阈值条件可以基于定位。例如，仅当用户已经将位置改变了阈值距离量时，才可以启动调用。应当注意，也可以使用其它阈值条件。在一些实施例中，步骤607和/或608可以发生在定位状态子模块（例如，定位状态子模块506）内。

[0084] 如果确定应该启动调用，则可以在步骤610处检索持久坐标系。在一些实施例中，可以在步骤610处仅检索可用持久坐标系的子集。例如，仅可以检索定位附近的持久坐标系。

[0085] 在步骤612处，可以检索声学数据。在一些实施例中，在步骤612处所检索的声学数据可以与存储在声学数据子模块508中的声学数据相对应。在一些实施例中，在步骤612处可以仅检索可用声学数据的子集。例如，可以检索与一个或多个持久坐标系和/或锚点相关联的声学数据。在一些实施例中，步骤612和/或614可以发生在声学数据子模块（例如，声学数据子模块508）内。

[0086] 在步骤614处，可以构建和/或修改音频模型。在一些实施例中，音频模型可以考虑在步骤612处所检索的声学数据，并且音频模型可以对特定环境的声学特性进行建模。在一些实施例中，步骤614可以发生在音频模型子模块（例如，音频模型子模块510）内。

[0087] 在步骤616处，可以确定是否应该更新音频模型的副本。可能希望设置一个或多个条件用于更新音频模型的副本以避免中断服务（例如，向用户呈现音频）。例如，一个条件可以评估音频模型的副本是否存在（例如，在音频渲染服务604内但在持久性模块602之外）。如果不存在音频模型的副本，则音频模型可以由音频渲染服务604（其可以对应于音频渲染

服务522)检索。在一些实施例中,条件可以评估音频模型的副本当前是否正在使用(例如,音频模型是否正被用于渲染音频以呈现给用户)。如果音频模型的副本未在使用,则更新的音频模型(例如,在步骤614处生成的音频模型)可以被传播到该副本。

[0088] 在步骤618处,音频渲染服务604可以检索音频模型(例如,在步骤614所生成的音频模型)的副本。可以使用数据序列化和/或反序列化来传送音频模型。

[0089] 在步骤620处,可以可选地删除和/或禁用过时的音频模型。例如,音频渲染服务604可以具有它先前已经使用的第一现有音频模型。音频渲染服务604可以检索第二个更新的音频模型(例如,从持久性模块602)并且删除和/或禁用第一现有音频模型。

[0090] 在步骤622处,可以发出新音频模型的通知。在一些实施例中,通知可以包括对可以订阅以在音频模型改变时听到的客户端(例如,第三方应用)的回调函数。

[0091] 图7示出了根据一些实施例的用于更新音频模型的示例性过程。在步骤706处,可以接收音频数据(例如,经由MR系统112、200的一个或多个传感器)。在一些实施例中,可以手动输入音频数据(例如,用户和/或开发者可以手动输入混响时间、混响延迟、混响增益等)。在可发生在持久性模块702(其可与持久性模块502相对应)中的步骤708处,可以识别相关联的环境。相关联的环境可以由可以伴随音频数据的元数据来识别(例如,元数据可以携带关于MR系统可能已知的一个或多个持久坐标系和/或锚点的信息)。

[0092] 在步骤710处,可以确定相关联的环境是否是新的。例如,如果相关联的环境可能没有被识别和/或与未知的标识符相关联,则可以确定音频数据与新环境相关联。如果确定相关联的环境不是新的,则可以更新正式(official)音频模型的副本(例如,利用可以从音频数据导出的房间属性)。如果确定相关联的环境是新的,则可以将新环境添加到正式音频模型的副本,并且可以相应地更新副本音频模型。在一些实施例中,新环境可以由新的模块化单元来表示。

[0093] 在步骤716处,可以初始化与新环境相关联的元数据(例如,与新模块化单元相关联的元数据)。例如,可以创建与测量计数、置信度或其他信息相关联的元数据并将其与新的模块化单元捆绑在一起。

[0094] 在步骤718处,可以更新持久性模块702内的正式音频模型。例如,可以从更新的副本音频模型复制正式音频模型。在一些实施例中,正式音频模型可以在步骤718处被锁定以防止对正式音频模型做出进一步的改变。在一些实施例中,当正式音频模型被锁定时,仍然可以对正式音频模型的副本(可能仍然存在于持久性模块702内)做出改变。

[0095] 在步骤720处,可以保存与新音频数据相关联的声学数据。例如,与新房间相关联的新模块化单元可以被保存和/或传递给可联结的世界系统(这可以使其在未来根据需要被MR系统访问)。在一些实施例中,步骤720可以发生在声学数据子模块508内。在一些实施例中,步骤720可以在步骤718之后顺序发生。在一些实施例中,步骤720可以在如由其他部件确定的独立时间处发生,例如,基于可以保存声学数据的可联结的世界系统的可用性。

[0096] 在步骤722处,音频渲染服务704(其可以对应于音频渲染服务522)可以检索音频模型的副本。在一些实施例中,音频模型的副本可以是在步骤718处更新的相同音频模型。在一些实施例中,数据传送可以通过序列化和反序列化发生。在一些实施例中,可以在执行序列化过程时锁定音频模型,这可以防止音频模型在创建音频模型的快照(snapshot)时改变。

[0097] 在步骤724处,可以删除和/或禁用音频模型的过时的副本。

[0098] 在步骤726处,可以发出关于新音频模型的通知。通知可以是订阅的客户端的回调函数,以在模型更新时听到。

[0099] 在步骤728处,可以发布正式音频模型,这可以指示可以删除对应于音频模型的序列化包。在一些实施例中,可能希望在将正式音频模型的副本正传送到音频渲染服务704时将正式音频模型锁定在持久性模块702内。一旦音频渲染服务704已经完成检索正式音频模型的副本,可能希望发布正式音频模型的锁定,以使得正式音频模型可以继续更新。

[0100] 在一些实施例中,音频渲染服务(例如,音频渲染服务522)可能不管理持久性模块(例如,持久性模块502)与渲染算法(例如,渲染算法模块520)之间的交互。例如,渲染算法520可以直接与持久性模块502通信以检索更新的音频模型。在一些实施例中,渲染算法520可以包括它自己的音频模型的副本。在一些实施例中,渲染算法520可以访问持久性模块502内的音频模型。

[0101] 多应用音频渲染

[0102] MR系统(例如,MR系统112、200)可以利用各种板载传感器为用户的MRE开发定制的音频模型。MR系统可以解释用户MRE中的真实特征(例如,地板、墙、人)以及用户MRE(例如,虚拟沙发)中的虚拟对象,以生成可适当反映用户MRE的真实和/或虚拟物理现象的定制音频模型。例如,虚拟声源可能位于用户的MRE中。音频模型可以识别从虚拟声源到用户的直接路径,以确定虚拟声音是否应该被直接路径中的真实或虚拟对象遮挡。在一些实施例中,音频模型可确定来自虚拟声源的声音如何从用户的MRE中的真实和/或虚拟对象(例如,硬表面可反射更多声音,某些表面可反射更高频声音,粗糙表面可在多个方向上散射声音等)反射。在一些实施例中,音频模型可以确定来自虚拟声源的声音如何在用户的MRE中混响。在一些实施例中,可能希望允许对模型进行定制,该模型对于用户的MRE的真实和/或虚拟物理现象可能是“准确的”。例如,MR应用可以从用户在其当前物理空间中开始,并慢慢地将用户的环境转换为异域环境。当MR应用变换用户的环境时(例如,通过添加虚拟树、虚拟叶子等),MR应用可以适配“正式”音频模型(例如,从正式音频模型中移除物理墙和天花板),以适合期望的用户体验。

[0103] 然而,当多于一个MR应用在MR系统上同时运行时,可能出现的问题。在一些实施例中,每个应用可能希望使用不同的音频模型来向用户呈现声音。例如,MR系统可以同时运行网络浏览应用和播客应用。在一些实施例中,网络浏览器可以播放信息视频,并且网络浏览器可以使用由MR系统生成的正式音频模型。因为网络浏览器可以使用正式音频模型,所以信息视频可以听起来就像扬声器位于用户的MRE中(例如,扬声器的声音可以基于用户MRE中的真实和/或虚拟对象被适当地遮挡、反射和混响)。在一些实施例中,播客应用可以播放位于洞穴中的播客。因此,可能希望播客应用将声音呈现为高度反射和/或混响。

[0104] 然而,如果网络浏览应用和播客应用同时尝试使用不同的音频模型呈现音频,则可能出现冲突。在一些实施例中,来自每个应用的音轨可以使用单独的音频模型单独渲染,并且稍后混合在一起以向用户呈现声音(例如,包括一个或多个渲染声音的音频信号)。然而,完全单独的渲染管线可能在计算上很昂贵,并且在一些情况下可能使渲染来自单个应用的音轨所需的计算资源加倍。随着更多应用同时运行,计算资源可能变得更加紧张,并且渲染来自例如使用单独的定制音频模型的十个不同同时运行的应用的音轨可能在计算上

不可行。因此,可能希望设计系统和方法以适应可以利用不同音频模型的同时运行的MR应用。

[0105] 图8示出了根据一些实施例的模型管理架构。在一些实施例中,MR系统803(其可以对应于MR系统112、200)可以包括被配置为执行指令的一个或多个计算机系统。MR系统803可以包括音频服务模块802(其可以对应于音频服务522)。音频服务模块802可以管理音频渲染,并且可以包括被配置为执行一个或多个进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中,音频服务模块802可以被配置为执行服务(例如,后台服务)作为一个或多个计算机系统的操作系统的一部分。

[0106] 在一些实施例中,单独的系统(例如,第三方应用)可以请求呈现音频信号(例如,经由MR系统112、200的一个或多个扬声器)。这样的请求可以采取任何适合的形式。在一些实施例中,呈现音频信号的请求可包括呈现音频信号的一个或多个软件指令;在一些实施例中,这样的请求可以是硬件驱动的。可以在有或没有用户参与的情况下发出请求。进一步地,可以经由本地硬件(例如,MR系统本身)、经由外部硬件(例如,与MR系统通信的单独计算机系统)、经由互联网(例如,经由云服务器)、或经由任何其他适合的源或源的组合来接收这样的请求。在一些实施例中,音频服务802可以接收请求,渲染所请求的音频信号(例如,通过一个或多个渲染算法,其可以利用由模型层804管理的音频模型),并向用户呈现所请求的音频信号。在一些实施例中,音频服务802可以被配置为在MR系统的操作系统正在运行时执行连续运行(例如,在后台)的进程。在一些实施例中,音频服务802可以被配置为执行父后台服务的实例,其可以用作一个或多个后台进程和/或子进程的主机进程。在一些实施例中,音频服务802可以被配置为执行作为MR系统的操作系统的一部分的指令。在一些实施例中,可以在MR系统上运行的应用可以访问音频服务802。在一些实施例中,MR系统的用户可能不直接向音频服务802提供输入。例如,用户可以向在MR系统上运行的应用(例如,角色扮演游戏)提供输入(例如,移动命令)。应用可以向音频服务802提供输入(例如,以渲染脚步声),并且音频服务802可以向用户(例如,经由扬声器)和/或其他过程和/或服务提供输出(例如,包括所渲染的脚步声的音频信号)。

[0107] 在一些实施例中,音频服务802可以包括模型层804,该模型层804可以被配置为管理一个或多个音频模型。在一些实施例中,模型层804可以包括被配置为执行一个或多个进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中,模型层804可以包括被配置为存储信息的一个或多个计算机系统。例如,模型层804可以存储一个或多个音频模型806a、806b和/或806c。在一些实施例中,音频模型可以包括用于渲染音频的数据的一个或多个抽象表示。在一些实施例中,音频模型可以包括被配置为存储信息的一个或多个数据结构。在一些实施例中,音频模型(例如,音频模型806a)可以包括一个或多个音频模型分量(例如,分量808a和808b)。在一些实施例中,音频模型分量可以包括被配置为存储信息的一个或多个数据结构。

[0108] 在一些实施例中,音频模型分量可以存储用于渲染音频的数据的一个或多个抽象表示。例如,音频模型分量可以包括关于一个或多个真实和/或虚拟对象的数据。在一些实施例中,音频模型分量可包括关于一个或多个真实和/或虚拟对象的位置数据。在一些实施例中,音频模型分量可包括关于一个或多个真实和/或虚拟对象的材料数据(例如,音频反射率、透射率、吸收、散射、衍射等)。在一些实施例中,一个音频模型分量可以表示一个真实

和/或虚拟对象及其相关联的特性。

[0109] 在一些实施例中,音频模型可包括关于一个或多个虚拟声源的数据。例如,音频模型可以包括关于一个或多个虚拟声源的位置数据。在一些实施例中,一个音频模型分量可以表示一个虚拟声源及其相关联的特性。

[0110] 在一些实施例中,音频模型可包括关于一个或多个虚拟声源的音频参数数据(例如,声辐射特性、音量等)。在一些实施例中,音频模型可包括环境的一个或多个音频参数。例如,音频模型可包括环境的尺寸(例如,用户正在占用的房间)。在一些实施例中,音频模型可包括环境的混响时间和/或混响增益。在一些实施例中,一个音频模型分量可以表示环境的一个音频参数。

[0111] 在一些实施例中,模型层804可以被配置为管理不同的音频模型。例如,模型806a可以被认为是正式模型(例如,模型806a可以被配置为表示虚拟声音与用户环境中的真实和/或虚拟对象之间的逼真物理交互)。在一些实施例中,模型806a可包括表示虚拟墙(其可以驻留在用户的环境中)的模型分量808a。在一些实施例中,模型806a可包括表示混响时间(例如,用户环境的)的模型分量808b。在一些实施例中,模型806a可以由应用814a使用,应用814a可以被配置为在MR系统803上运行。在一些实施例中,应用814a可以是已经请求将虚拟音频呈现给用户的MR应用。在一些实施例中,应用814a与模型806a之间的关联可以存储在模型层804中(例如,经由与模型806a相关联的元数据)。在一些实施例中,应用814a可以存储数据(例如,模型标识符),并且应用814a可指定使用该数据来使用特定音频模型。

[0112] 在一些实施例中,模型层804可以被配置为管理模型806b。在一些实施例中,模型806b可以与模型806a不同。例如,模型806b可以包括模型分量808c,该模型分量808c可以与模型分量808a相对应和/或与模型分量808a相同(例如,模型分量808a和808c都可以表示相同的虚拟墙)。然而,模型806b可以包括模型分量808d,该模型分量808d可以不与模型分量808b相对应和/或与模型分量808b不同。例如,模型分量806d可以表示第二混响时间,该第二混响时间可以是比与模型分量806b相对应的混响时间更长的混响时间。

[0113] 在一些实施例中,模型806b可以是全音频模型。例如,模型806b可用于渲染音频而不依赖于其他模型(例如,模型806a)。在一些实施例中,模型806b可包括对其他模型的一个或多个依赖性。例如,模型806b可包括对模型806a中的模型分量808a的一个或多个指针(例如,因为模型分量808c与模型分量808a相对应)。在一些实施例中,与存储完全独立的音频模型相比,存储具有依赖性的音频模型对存储器和/或存储设备的要求可能较低。

[0114] 在一些实施例中,模型806c可以不与模型806a和/或806b共享分量。例如,模型分量808e可以表示虚拟沙发,模型分量808f可以表示虚拟白板,以及模型分量808g可以表示第三混响时间,其可以短于分别与模型分量808b和808d相关联的第二混响时间和第一混响时间。在一些实施例中,模型806c可以与应用814b相关联,应用814b可以被配置为在MR系统803上运行。

[0115] 尽管图8将模型层804描绘为管理作为音频服务802的一部分的一个或多个音频模型,但是也可以构想其他实施例。在一些实施例中,每个应用可以管理其自己的一个或多个音频模型。例如,应用814a可以将模型806a本地地存储在应用814a内。在一些实施例中,模型806a可以是正式模型,并且应用814a可以将对模型806a的引用存储在模型层804中。在一些实施例中,应用814b可以本地地存储模型806c,并且模型806b可以不存储在模型层804

中。在一些实施例中，模型层804可以生成音频模型。例如，MR系统的用户可以改变物理环境，并且新的音频模型可以更准确地表示新环境的声学特性。在一些实施例中，应用（例如，应用814a）可以请求生成新的音频模型，因为该应用可以利用一个或多个定制和/或因为该应用可以利用完全定制的音频模型。

[0116] 在一些实施例中，模型层804可以存储与当前在MR系统803上运行的应用相关联的音频模型。例如，MR系统803当前可能仅运行应用814a和814b，并且模型806b可能不与应用814a或814b相关联。在一些实施例中，可以从模型层804移除模型806b（例如，以节省存储器）。在一些实施例中，正式模型可以被连续地存储在模型层804中。在一些实施例中，如果音频模型没有被使用了阈值时间量，则模型层804可以从存储器中移除音频模型。在一些实施例中，应用可以请求模型层804在存储器中保留音频模型。

[0117] 在一些实施例中，音频服务802可以包括渲染层812，该渲染层812可以被配置为管理一个或多个音频模型。在一些实施例中，渲染层812（其可以对应于渲染算法模块520）可以包括被配置为执行一个或多个进程、子进程、线程和/或服务的一个或多个计算机系统。渲染层812可以渲染来自一个或多个输入的音频（例如，模型分量、声源信息等）。在一些实施例中，渲染层812可以渲染音频信息流，并且可以实时渲染音频流。在一些实施例中，渲染层812可以渲染来自固定量的输入数据的音频。渲染的音频可以作为包括所渲染的音频的一个或多个音频信号呈现给用户。音频信号可以经由设备的一个或多个扬声器（例如，图4所示的扬声器412和414）被呈现。

[0118] 尽管图8描绘了具有不同数量的模型分量的三个模型，但是应构想任何数量的模型可以由音频服务802管理。还构想任何模型可以包括任何数量的模型分量。

[0119] 图9A示出了根据一些实施例的渲染架构。在一些实施例中，渲染层902（其可以对应于渲染层812）可以包括被配置为执行一个或多个进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中，渲染层902可以被配置为渲染来自一个或多个输入的音频（例如，模型分量、声源信息等）。在一些实施例中，渲染层902可以渲染音频信息流，并且可以实时渲染音频流。在一些实施例中，渲染层902可以从固定量的输入数据渲染音频。

[0120] 渲染层902可以包括直接层904。在一些实施例中，直接层904可以包括被配置为执行进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中，直接层904可以被配置为渲染声源与收听节点（例如，用户）之间的直接路径。直接层904可以确定声音是否应当被遮挡（例如，因为真实和/或虚拟对象阻碍了从声源到收听节点的直接路径）。在一些实施例中，直接层904可以确定声音是否应当被衰减（例如，由于声源与收听节点之间的距离）。

[0121] 渲染层902可以包括反射层906。在一些实施例中，反射层906可以包括被配置为执行进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中，反射层906可以被配置为渲染声源与收听节点（例如，用户）之间的一个或多个反射音频路径。例如，来自声源的声音可以在到达收听节点之前从墙反射，并且反射层906可以渲染反射的一种或多种效果（例如，与直接声音相比较，所反射的声音可能被延迟，并且与直接声音相比较，所反射的声音可能被衰减）。在一些实施例中，反射层906可以被配置为将一个或多个滤波器应用于音频以近似一般环境中反射的声音的行为。在一些实施例中，反射层906可以被配置用于音频射线跟踪，其可以计算用于一个或多个音频射线的反射路径。

[0122] 渲染层902可以包括混响层908。在一些实施例中，混响层908可以包括被配置为执行进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中，混响层908可以被配置为渲染音频的混响行为(例如，后期混响行为)。在一些实施例中，后期混响行为可能受环境特性(诸如环境的混响时间和/或混响增益)影响和/或确定。

[0123] 渲染层902可以包括虚拟化器910。在一些实施例中，虚拟化器910可以包括被配置为执行进程、子进程、线程和/或服务的一个或多个计算机系统。在一些实施例中，虚拟化器910可以被配置为将音频呈现给一个或多个虚拟扬声器。例如，MR系统可以向用户呈现源自用户头部周围的扬声器阵列中布置的六个虚拟扬声器之一的音频。在一些实施例中，虚拟化器910可以被配置为确定应该在虚拟扬声器阵列中的哪个虚拟扬声器或扬声器组合处播放什么声音或声音组合。

[0124] 在图9A所示的示例实施例中，音频流912和音频流914可以由渲染层902渲染。在一些实施例中，音频流912可以与源自第一应用的音频流相对应，并且音频流914可以与源自第二应用的音频流相对应。在一些实施例中，第一应用可以使用与第二应用相同的音频模型，并且第一应用和第二应用可以在MR系统上被同时运行。在一些实施例中，一起渲染音频流912和音频流914可能比单独渲染音频流更有效(例如，因为两个音频流可能依赖于相同的音频模型)。例如，对应于音频流912和音频流914的直接路径可以被渲染并且可以被传递到虚拟化器910。在一些实施例中，可以使用相同的直接路径计算来渲染音频流912和音频流914(例如，因为相同的真实和/或虚拟对象可能阻碍一个或多个声源与一个或多个收听节点之间的直接路径)。

[0125] 在一些实施例中，音频流912和音频流914可以通过反射层906一起渲染。在一些实施例中，反射层906可以从直接层904接收一个或多个输入音频流。在一些实施例中，反射层906可以直接从应用接收一个或多个输入音频流(例如，不首先通过直接层904)。在一些实施例中，一起渲染音频流912和音频流914可能比单独渲染音频流更有效。例如，同一组滤波器可以被应用于两个音频流(和/或与两个音频流的混合相对应的单个音频流)以近似反射行为。在一些实施例中，渲染的反射可以被传递到虚拟化器910。

[0126] 在一些实施例中，音频流912和音频流914可以通过混响层908一起渲染。在一些实施例中，混响层908可以从反射层906和/或直接层904接收一个或多个输入音频流。在一些实施例中，混响层908可以直接从应用接收一个或多个输入音频流(例如，不通过反射层906和/或直接层904)。在一些实施例中，一起渲染音频流912和音频流914可能比单独渲染音频流更有效。例如，后期混响行为可以根据共享音频模型来确定，并且后期混响行为可以被应用于两个音频流。在一些实施例中，渲染的混响行为可以被传递到虚拟化器910。

[0127] 图9B示出了根据一些实施例的渲染架构。在一些实施例中，依赖于不同音频模型的多个音频流可以被一起渲染。例如，音频流912可以源自使用第一音频模型的第一应用，并且音频流914可以源自使用第二音频模型的第二应用。在一些实施例中，第一音频模型可以包括与可能不存在于第二音频模型中的虚拟对象相对应的模型分量(例如，因为第一应用已经引入了第二应用可能不利用的虚拟对象)。在一些实施例中，由第一应用引入的虚拟对象可能影响用于音频流912的直接渲染路径(例如，直接声音可能由于障碍而被遮挡)，但可能不影响用于音频流914的直接渲染路径。在一些实施例中，音频流912和音频流914可以在直接层904的单独实例中渲染。在一些实施例中，音频流912和音频流914可以各自在直接

层904处被独立渲染。

[0128] 在一些实施例中,可以在音频流912与音频流914之间共享其他模型分量。例如,虚拟对象可以被配置为不干扰音频反射,并且音频流912和音频流914可以在反射层906处使用相同的反射计算来渲染。在一些实施例中,虚拟对象可能不影响MRE的后期混响,并且音频流912和音频流914可以在混响层908处使用相同的混响计算来渲染。在一些实施例中,音频流912和音频流914两者可以在虚拟化器910处混合在一起。

[0129] 在一些实施例中,当音频流912和音频流914依赖于共享模型分量时,可以利用效率(即使模型分量中也可能存在一个或多个差异)。例如,反射、混响和虚拟化器计算可以跨音频流共享,即使在直接路径可能需要跨两个音频流的独立计算时。

[0130] 尽管图9B示出了具有不同直接路径实例的渲染架构,但是也构想其他实施例。例如,音频流912可以与将虚拟对象材料组成改变为更能反映音频的第一应用(与音频流914相比较,音频流914可以对应于没有修改虚拟对象材料组成的第二应用)。在一些实施例中,音频流912和音频流914可以在直接层904处被一起渲染,但可以在反射层906处被单独(例如,在单独的实例中)渲染。在一些实施例中,音频流912和音频流914可以在混响层908再次被一起渲染(例如,因为对应于后期混响的模型分量在由音频流912和音频流914依赖的音频模型之间共享)。

[0131] 尽管图9A至图9B示出了包括直接层904、反射层906、混响层908和虚拟化器910的渲染架构,但是也可以使用其他架构。在一些实施例中,一个或多个层可能不包括在渲染架构中(例如,由于计算限制)。在一些实施例中,可以添加一个或多个层以更逼真地模拟声学行为。

[0132] 公开了示例系统、方法和计算机可读介质。根据一些示例,一种系统包括:一个或多个扬声器;以及一个或多个处理器,其被配置为执行一种方法,所述方法包括:接收呈现第一音轨的请求,其中,第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;接收呈现第二音轨的请求,其中,第二音轨基于包括共享模型分量和第二模型分量的第二音频模型;基于第一音轨、第二音轨、共享模型分量、第一模型分量和第二模型分量,渲染声音;以及经由一个或多个扬声器呈现包括所渲染的声音的音频信号。在一些示例中,渲染声音包括:基于共享模型分量,渲染第一音轨和第二音轨。在一些示例中,渲染声音包括:基于第一音频分量,渲染第一音轨。在一些示例中,渲染声音包括:基于第二音频分量,渲染第二音轨。在一些示例中,一个或多个扬声器是头戴式设备的扬声器。在一些示例中,第一音频模型基于头戴式设备的物理环境。在一些示例中,共享模型分量对应于直接路径声学行为。在一些示例中,共享模型分量对应于反射声学行为。在一些示例中,共享模型分量对应于混响声学行为。

[0133] 根据一些示例,一种方法:接收呈现第一音轨的请求,其中,第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;接收呈现第二音轨的请求,其中,第二音轨基于包括共享模型分量和第二模型分量的第二音频模型;基于第一音轨、第二音轨、共享模型分量、第一模型分量和第二模型分量,渲染声音;以及经由一个或多个扬声器呈现包括所渲染的声音的音频信号。在一些示例中,渲染声音包括:基于共享模型分量,渲染第一音轨和第二音轨。在一些示例中,渲染声音包括:基于第一音频分量,渲染第一音轨。在一些示例中,渲染声音包括:基于第二音频分量,渲染第二音轨。在一些示例中,呈现音频信号包括:

经由头戴式设备的一个或多个扬声器呈现音频信号。在一些示例中,第一音频模型基于头戴式设备的物理环境。在一些示例中,共享模型分量对应于直接路径声学行为。在一些示例中,共享模型分量对应于反射声学行为。在一些示例中,共享模型分量对应于混响声学行为。

[0134] 根据一些示例,一种存储指令的非暂态计算机可读介质,该指令在由一个或多个处理器执行时使得一个或多个处理器执行一种方法,该方法包括:接收呈现第一音轨的请求,其中,第一音轨基于包括共享模型分量和第一模型分量的第一音频模型;接收呈现第二音轨的请求,其中,第二音轨基于包括共享模型分量和第二模型分量的第二音频模型;基于第一音轨、第二音轨、共享模型分量、第一模型分量和第二模型分量,渲染声音;以及经由一个或多个扬声器呈现包括所渲染的声音的音频信号。在一些示例中,渲染声音包括:基于共享模型分量,渲染第一音轨和第二音轨。在一些示例中,渲染声音包括:基于第一音频分量,渲染第一音轨。在一些示例中,渲染声音包括:基于第二音频分量,渲染第二音轨。在一些示例中,呈现音频信号包括:经由头戴式设备的一个或多个扬声器呈现音频信号。在一些示例中,第一音频模型基于头戴式设备的物理环境。在一些示例中,共享模型分量对应于直接路径声学行为。在一些示例中,共享模型分量对应于反射声学行为。在一些示例中,共享模型分量对应于混响声学行为。

[0135] 虽然所公开的示例已经参考附图充分描述,但是,应注意到,各种改变和修改对于本领域技术人员来说将变得明显。例如,一个或多个实施方式的元素可以组合、删除、修改、或补充以形成进一步的实施方式。这样的改变和修改将被理解为包括在如由附加的权利要求限定的所公开的示例的范围内。

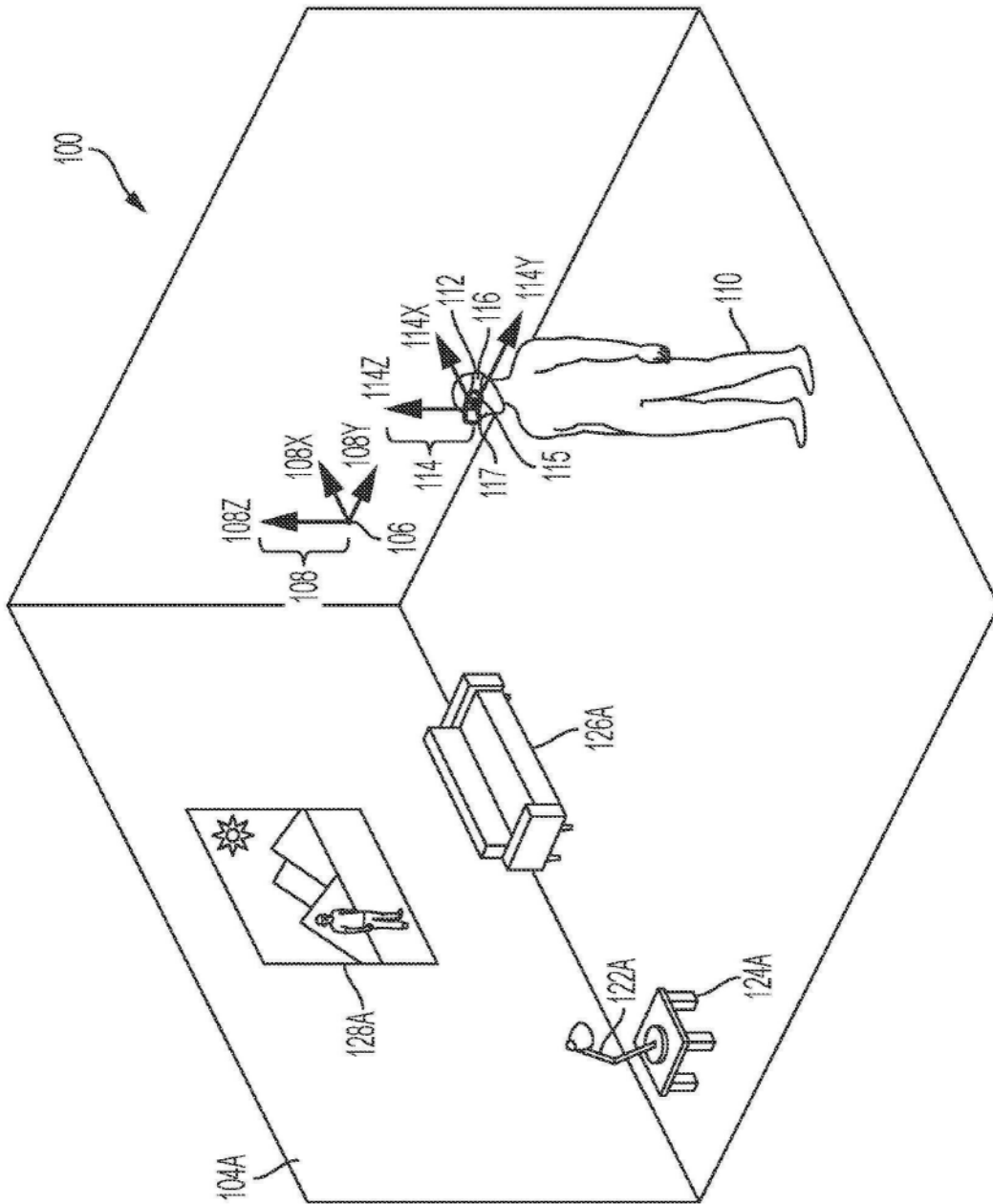


图1A

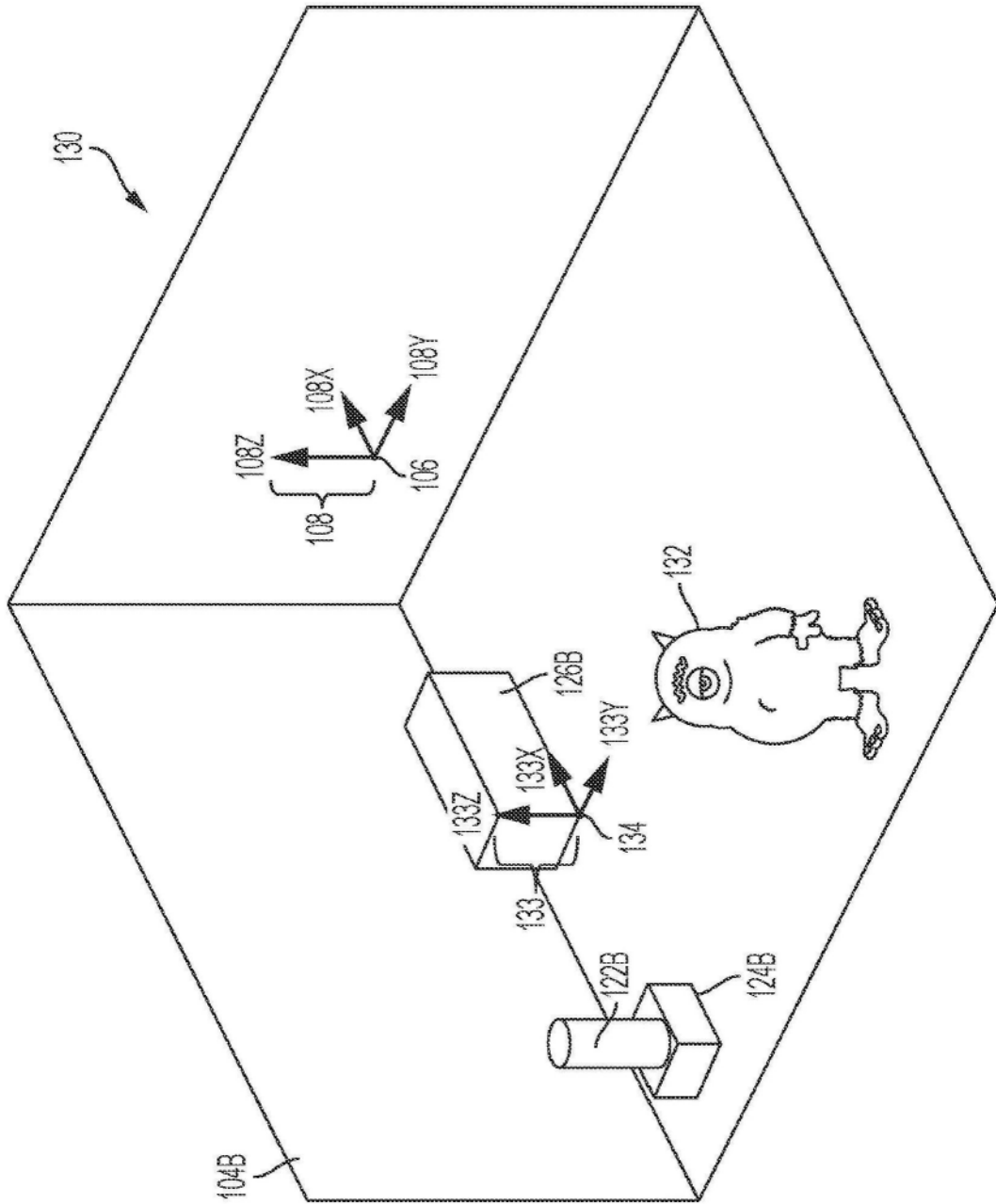


图1B

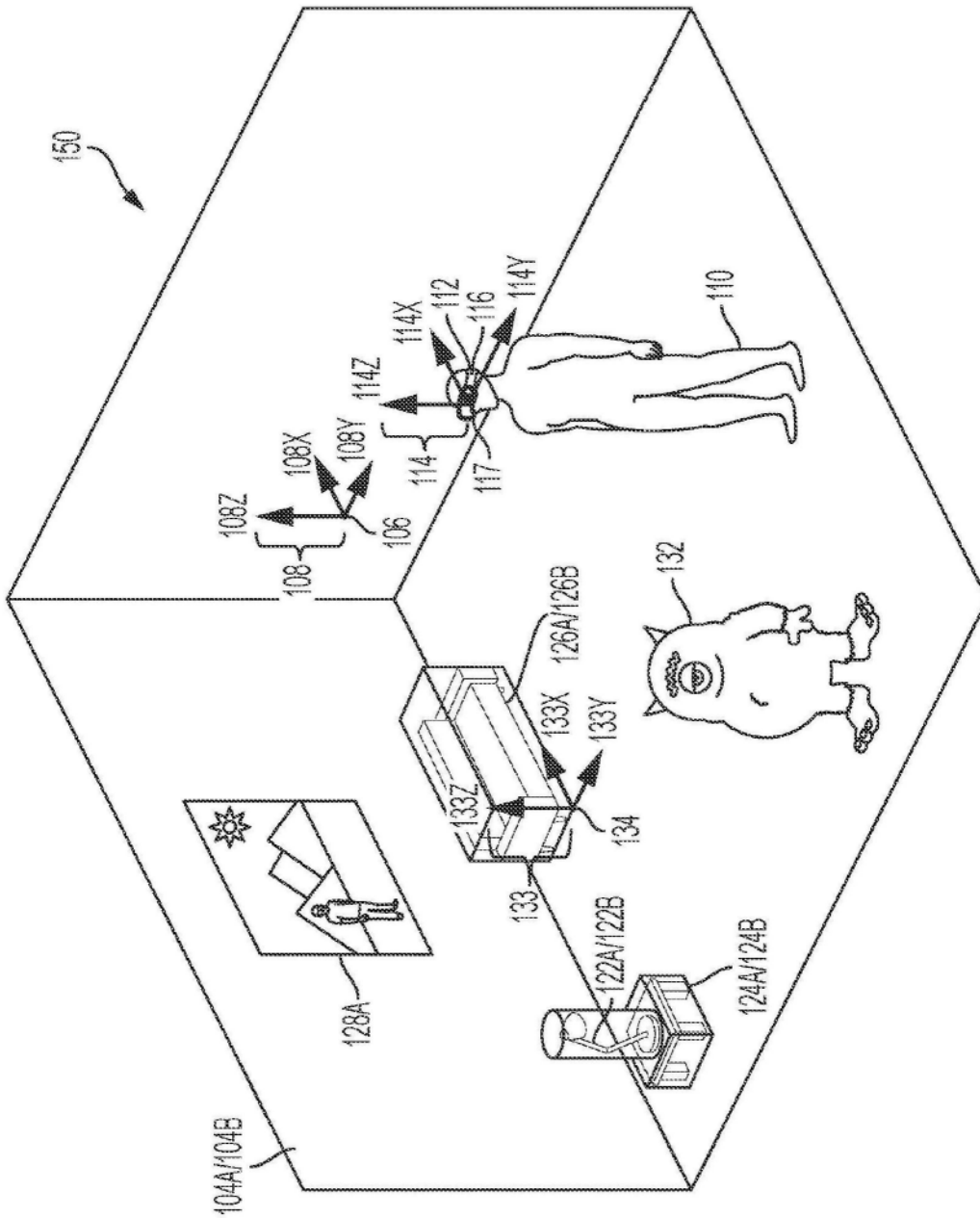


图1C

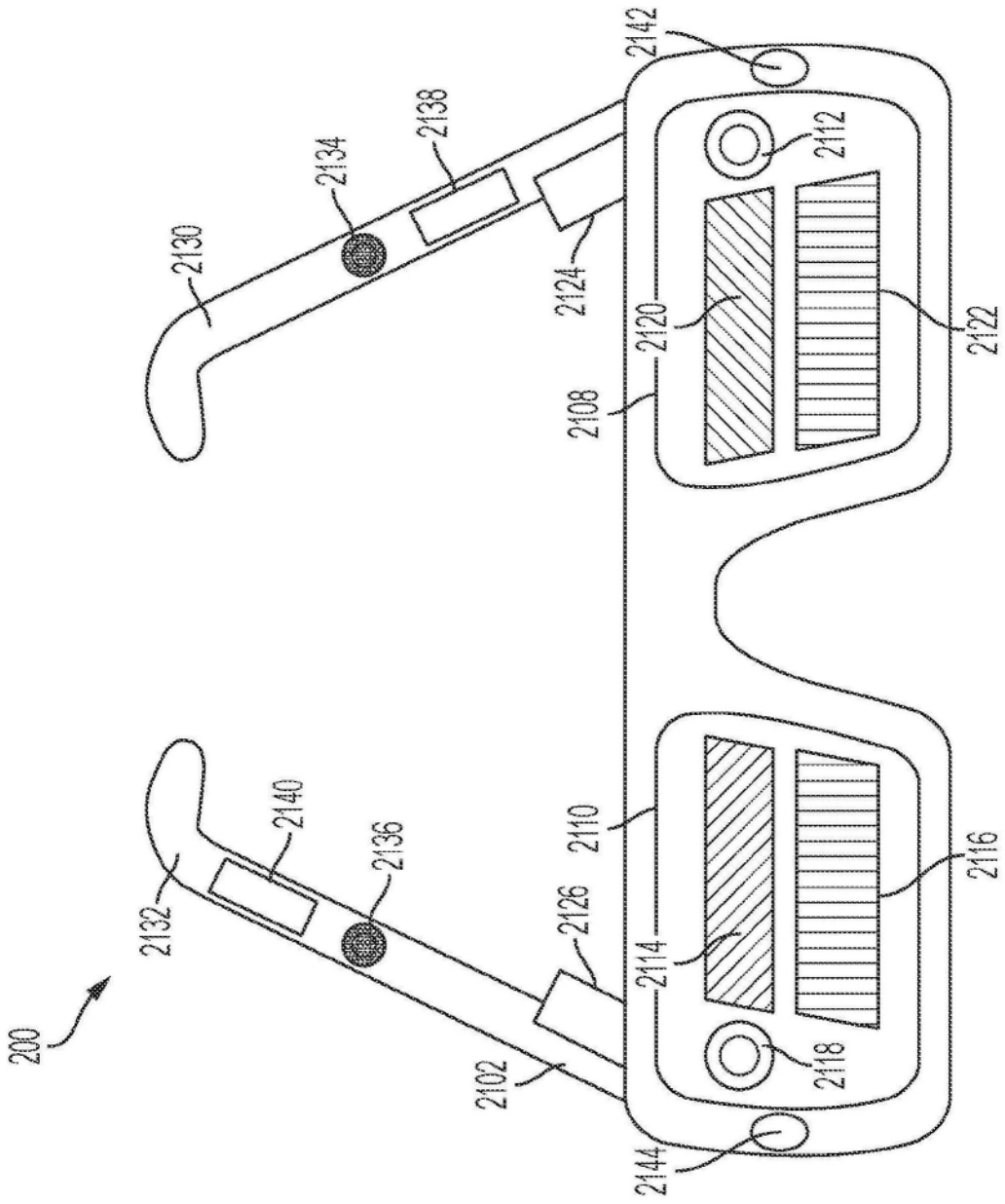


图2A

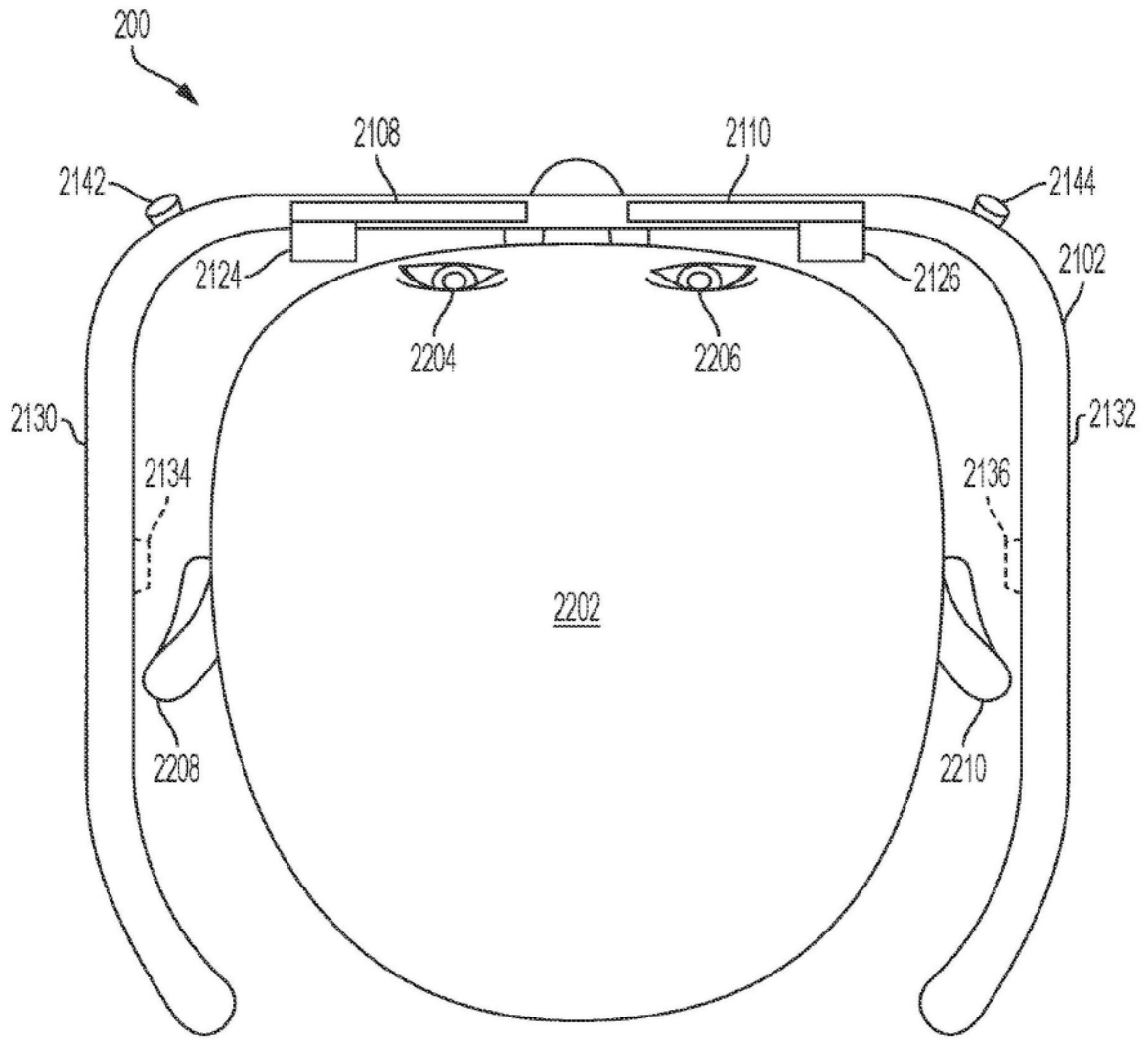


图2B

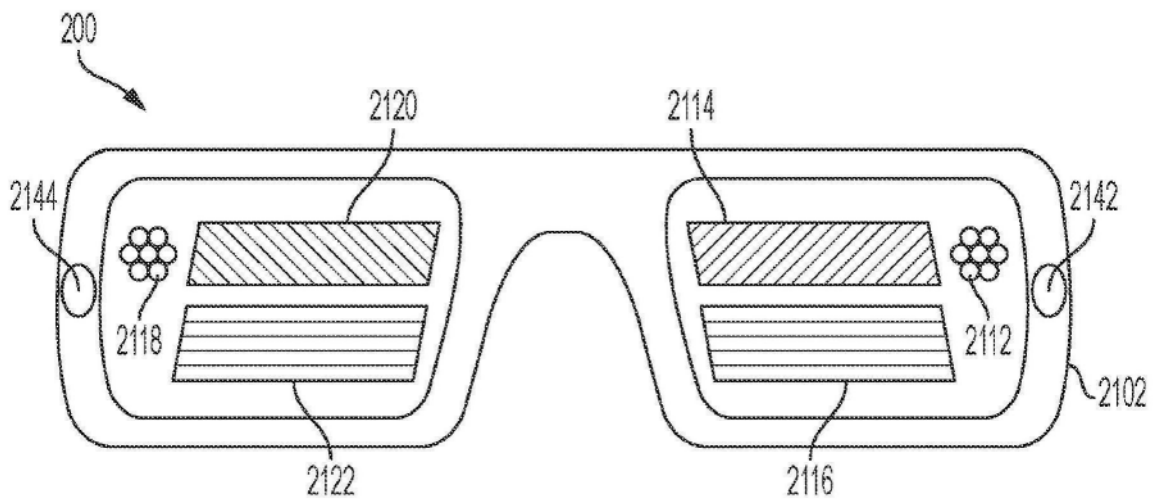


图2C

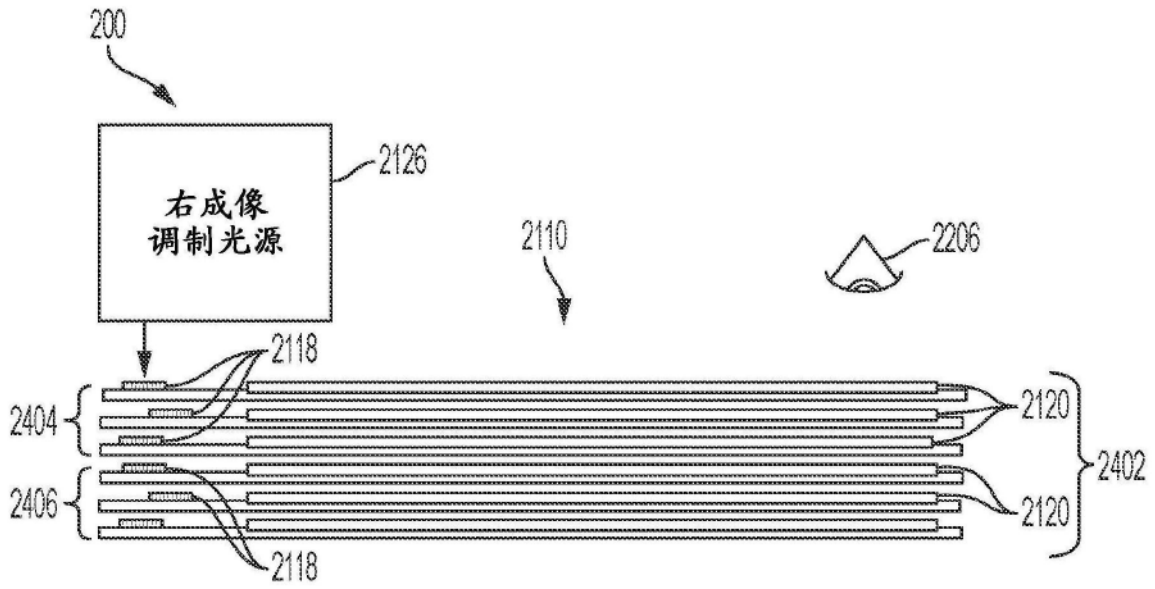


图2D

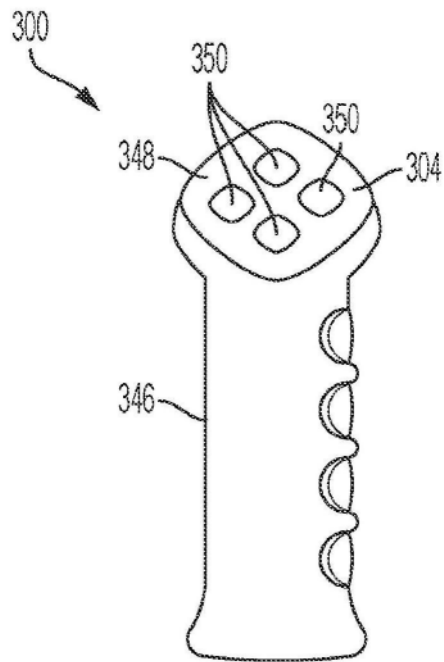


图3A

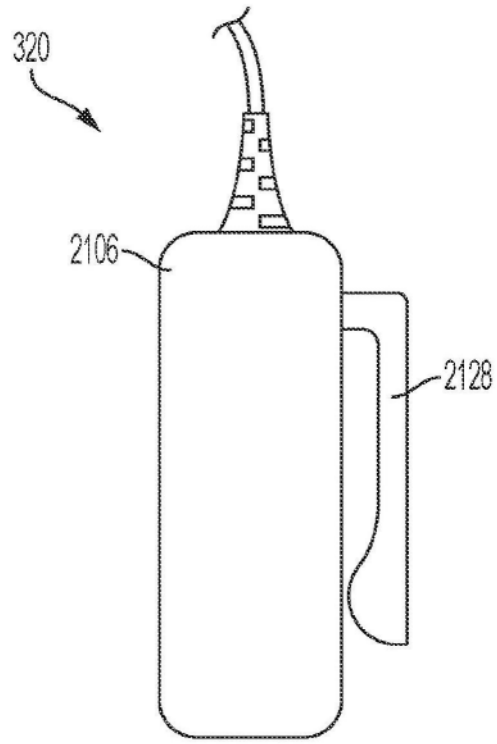


图3B

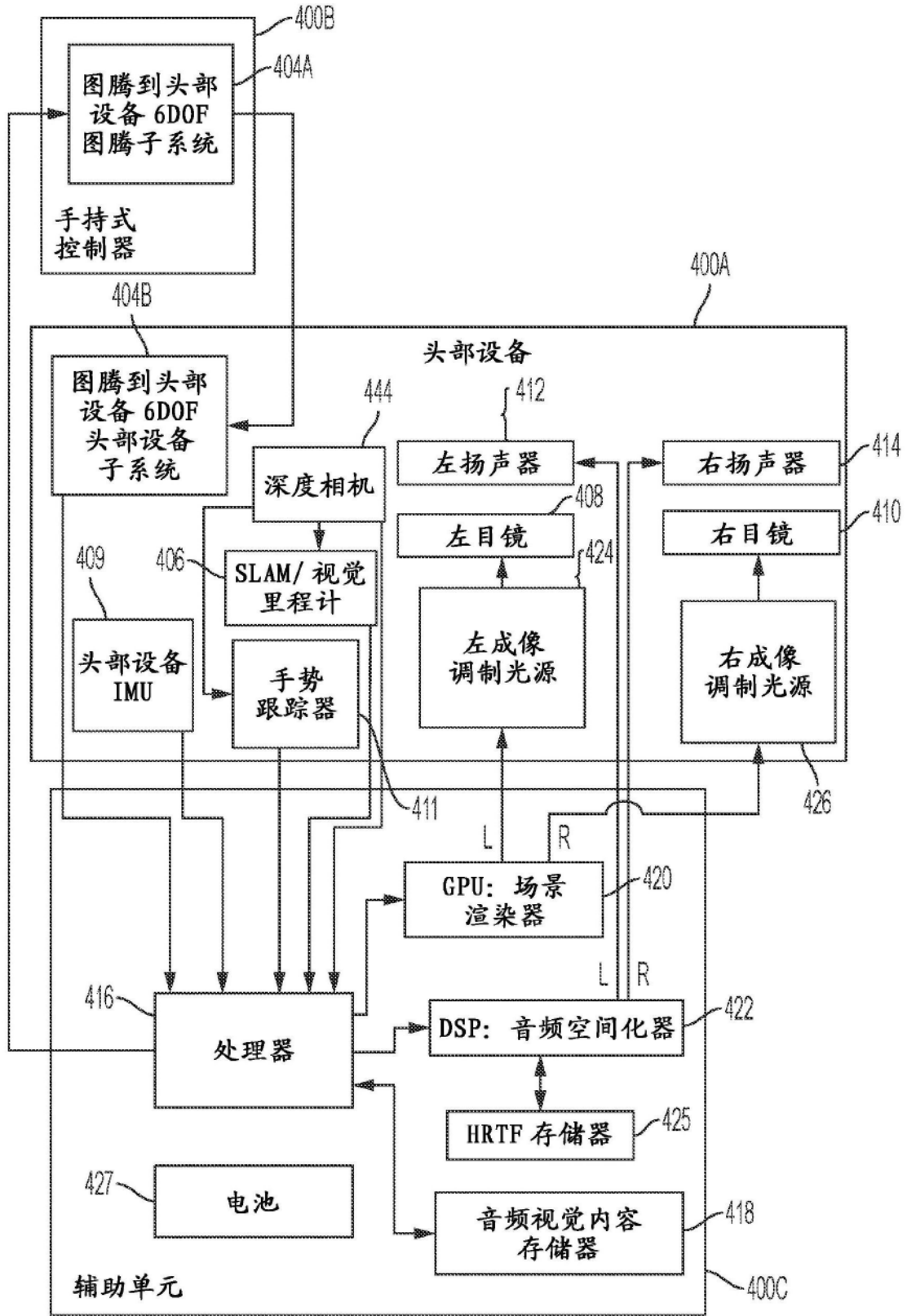


图4

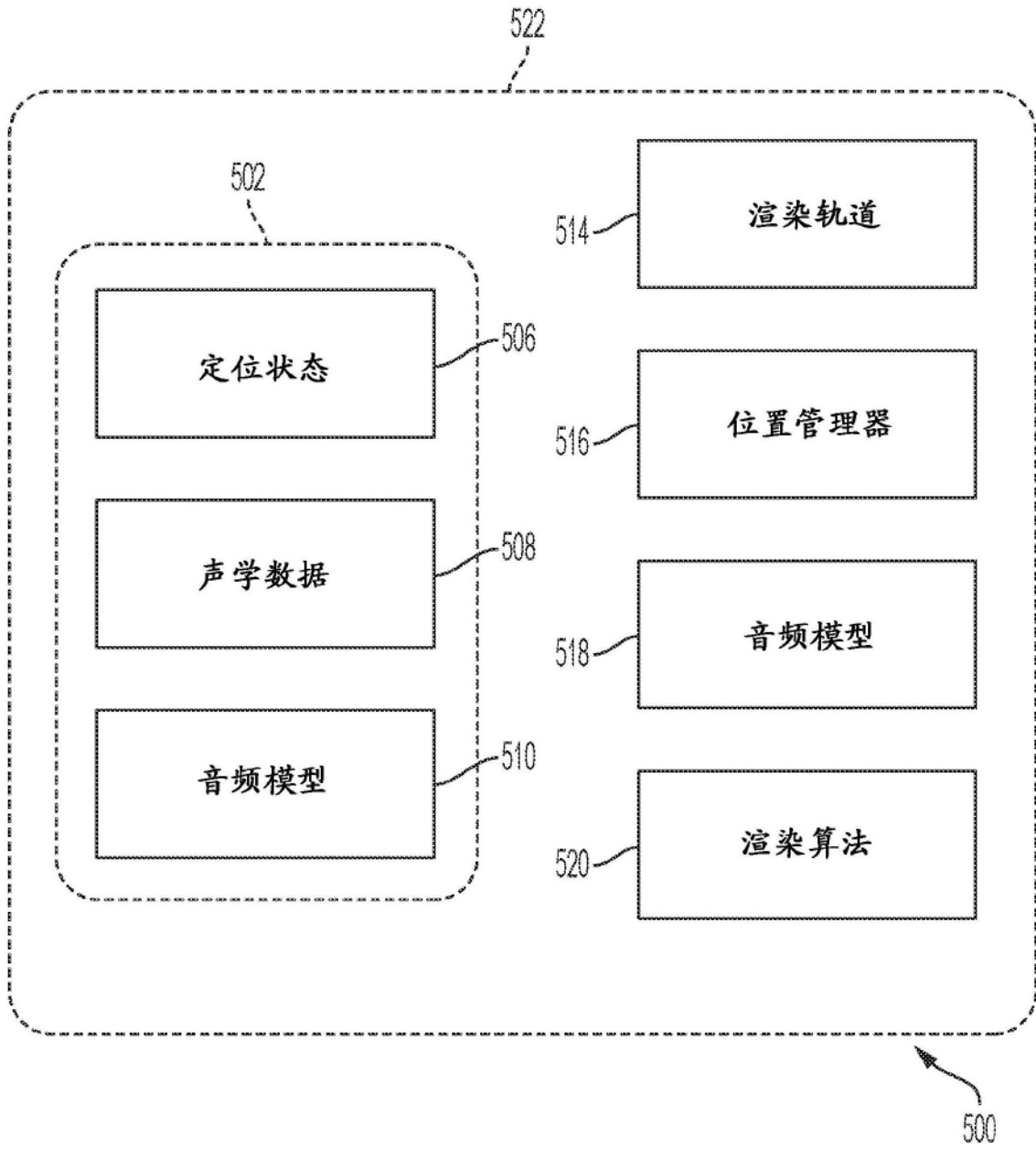


图5

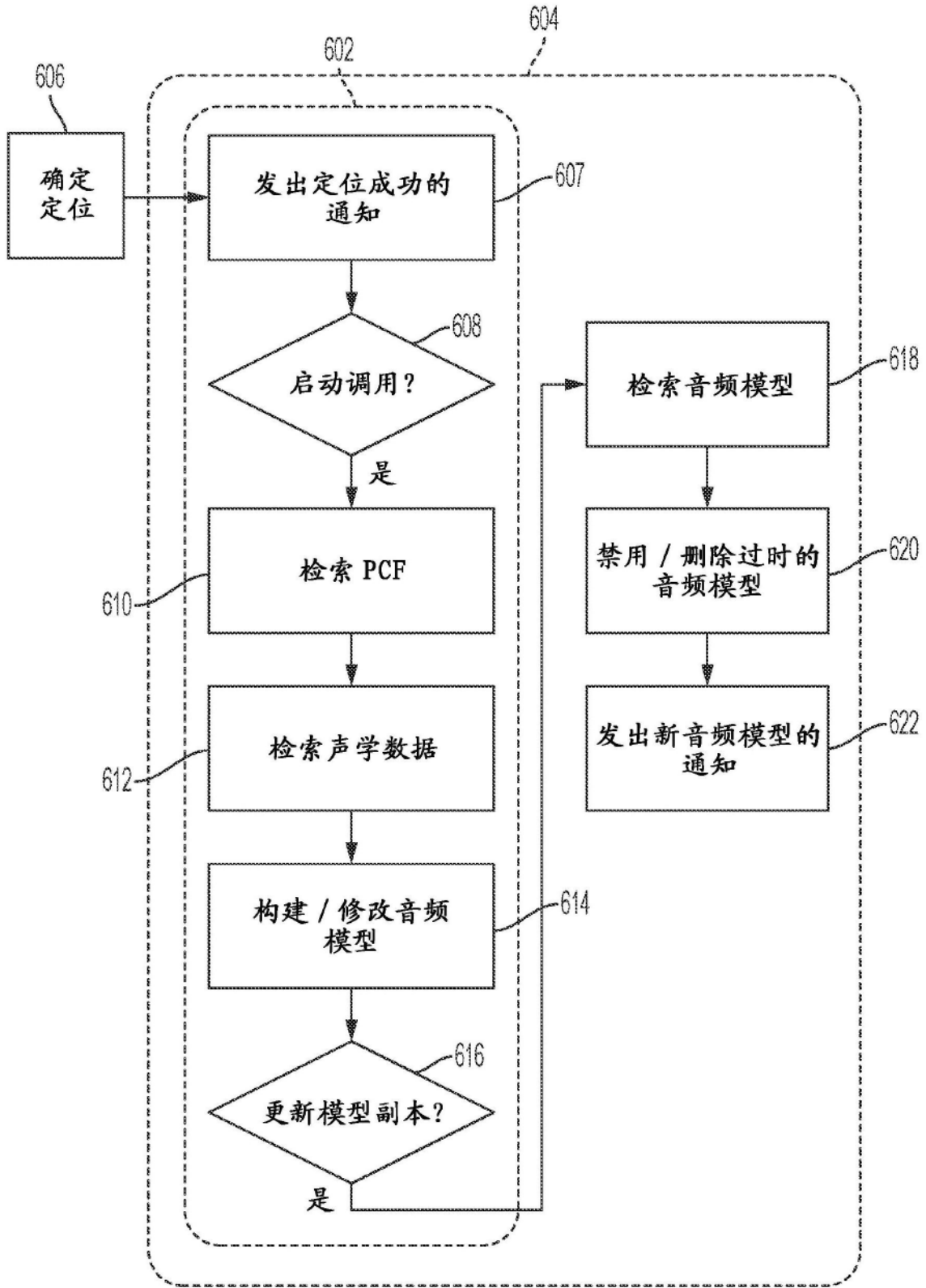


图6

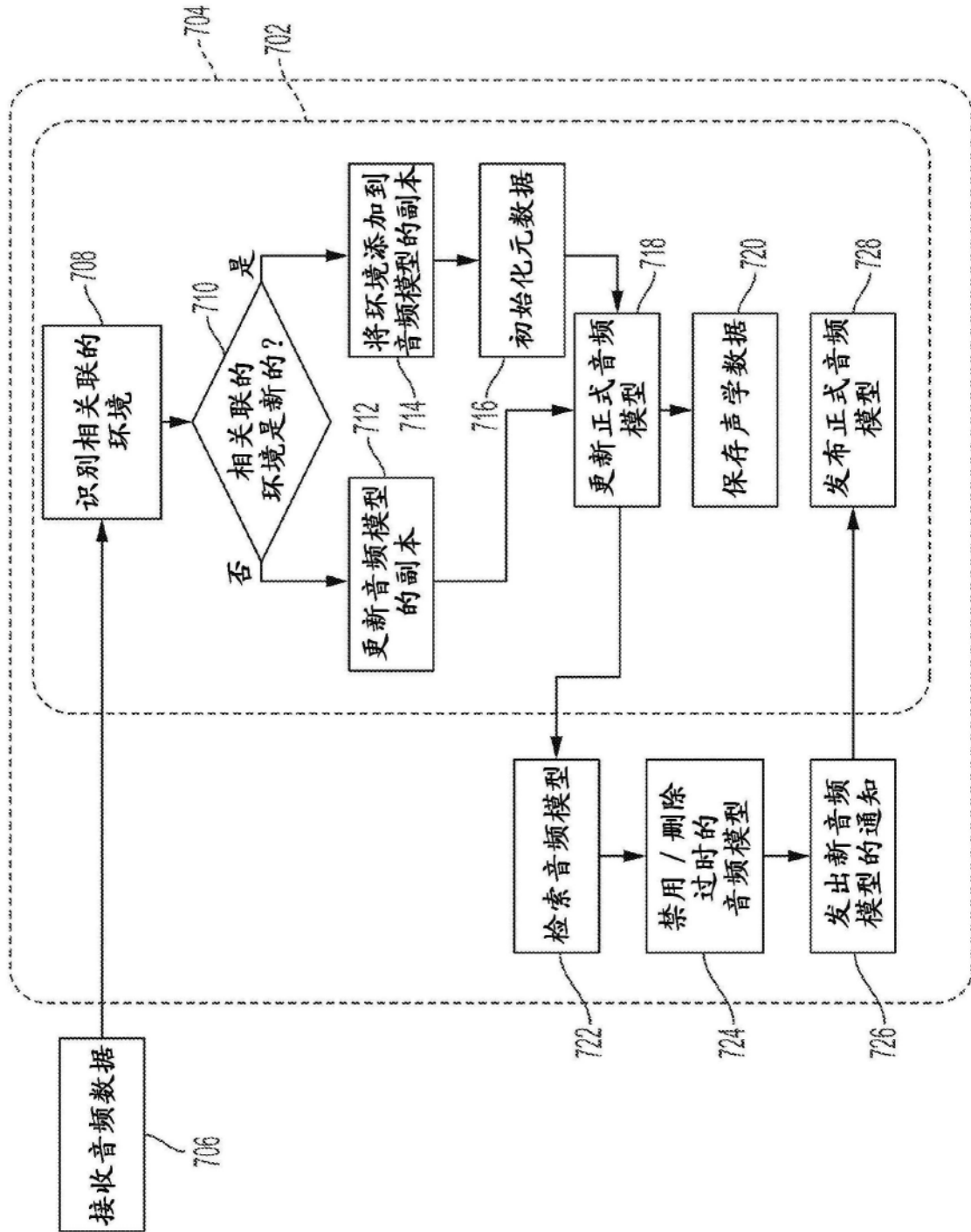


图7

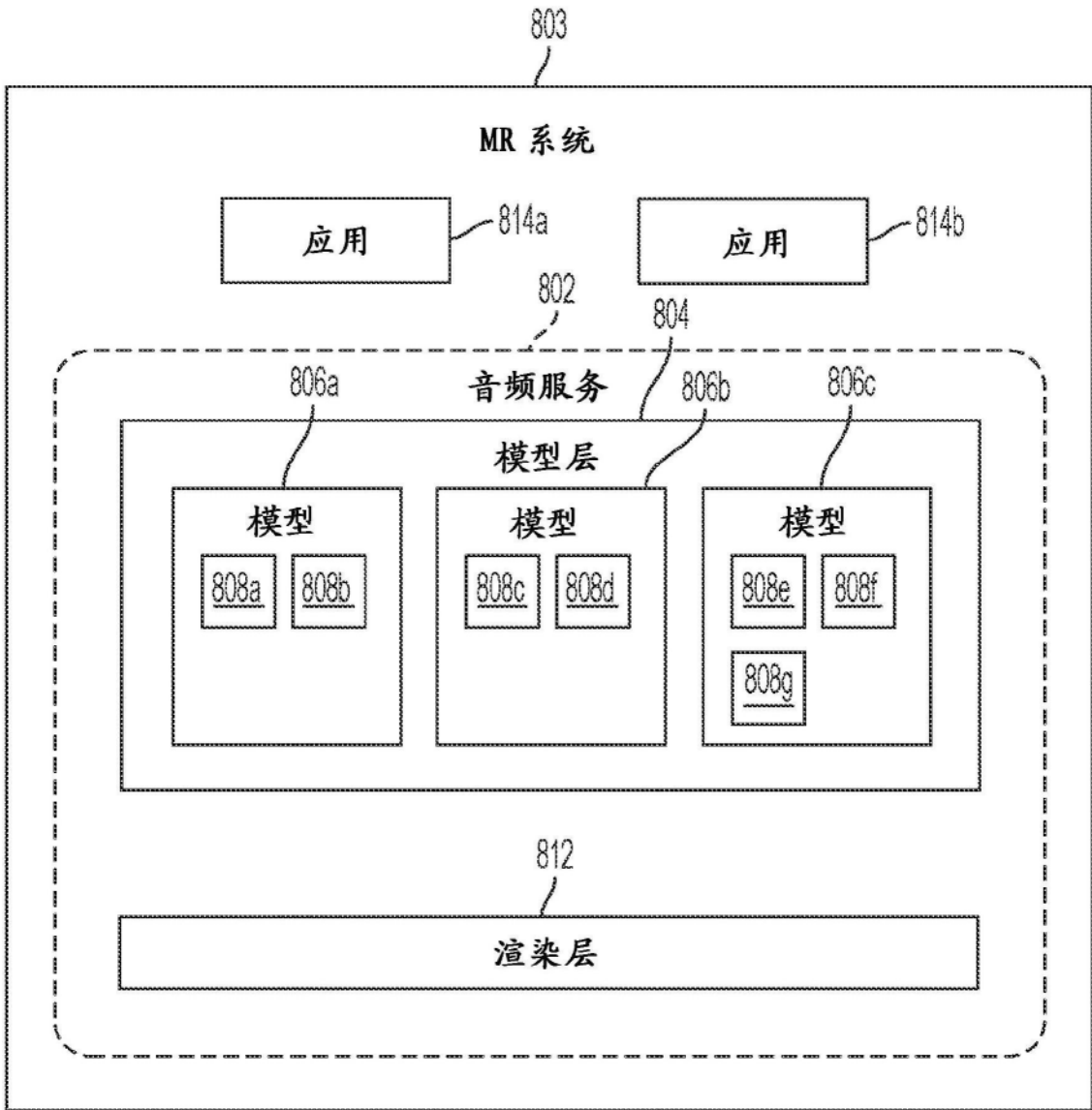


图8

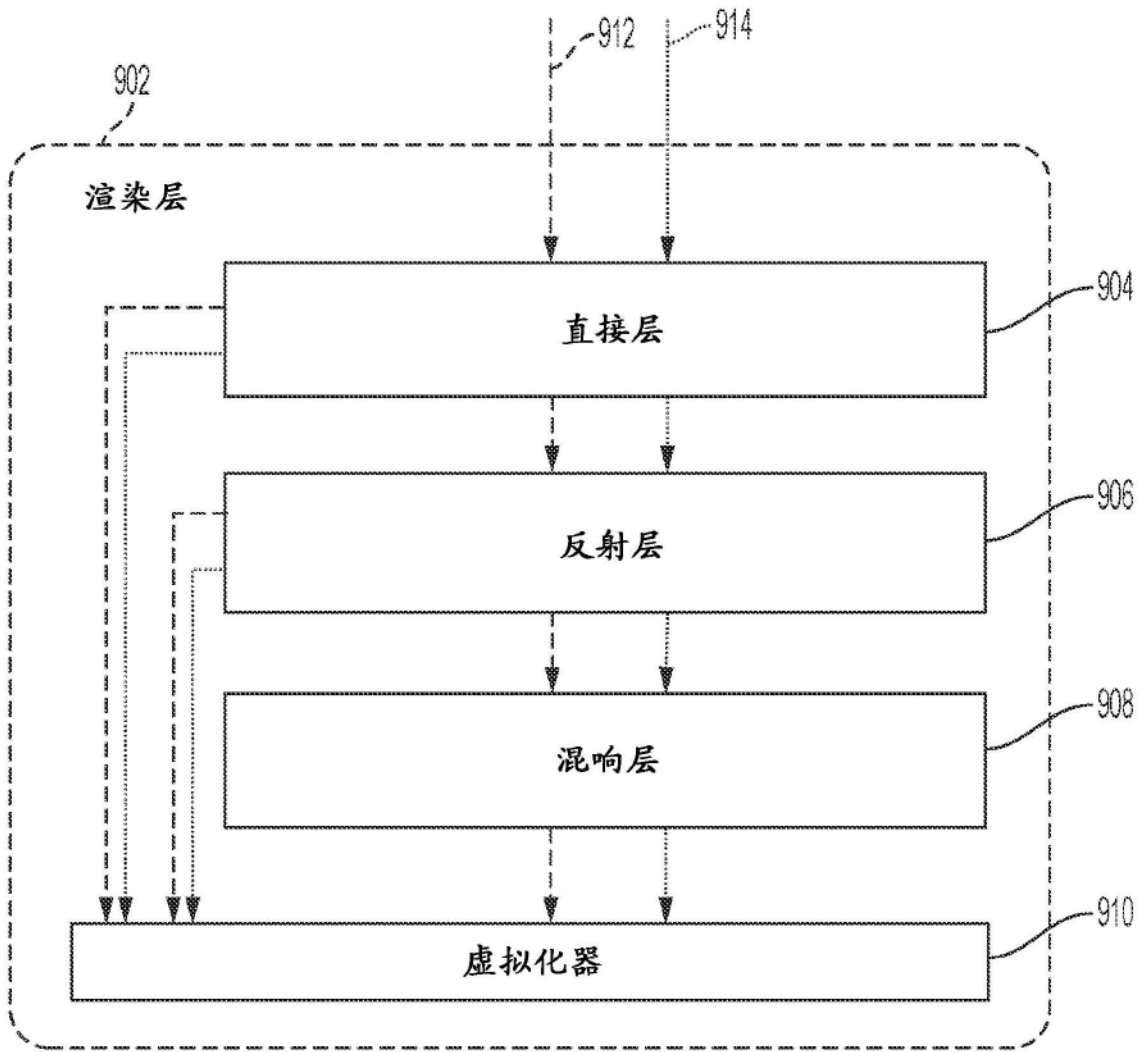


图9A

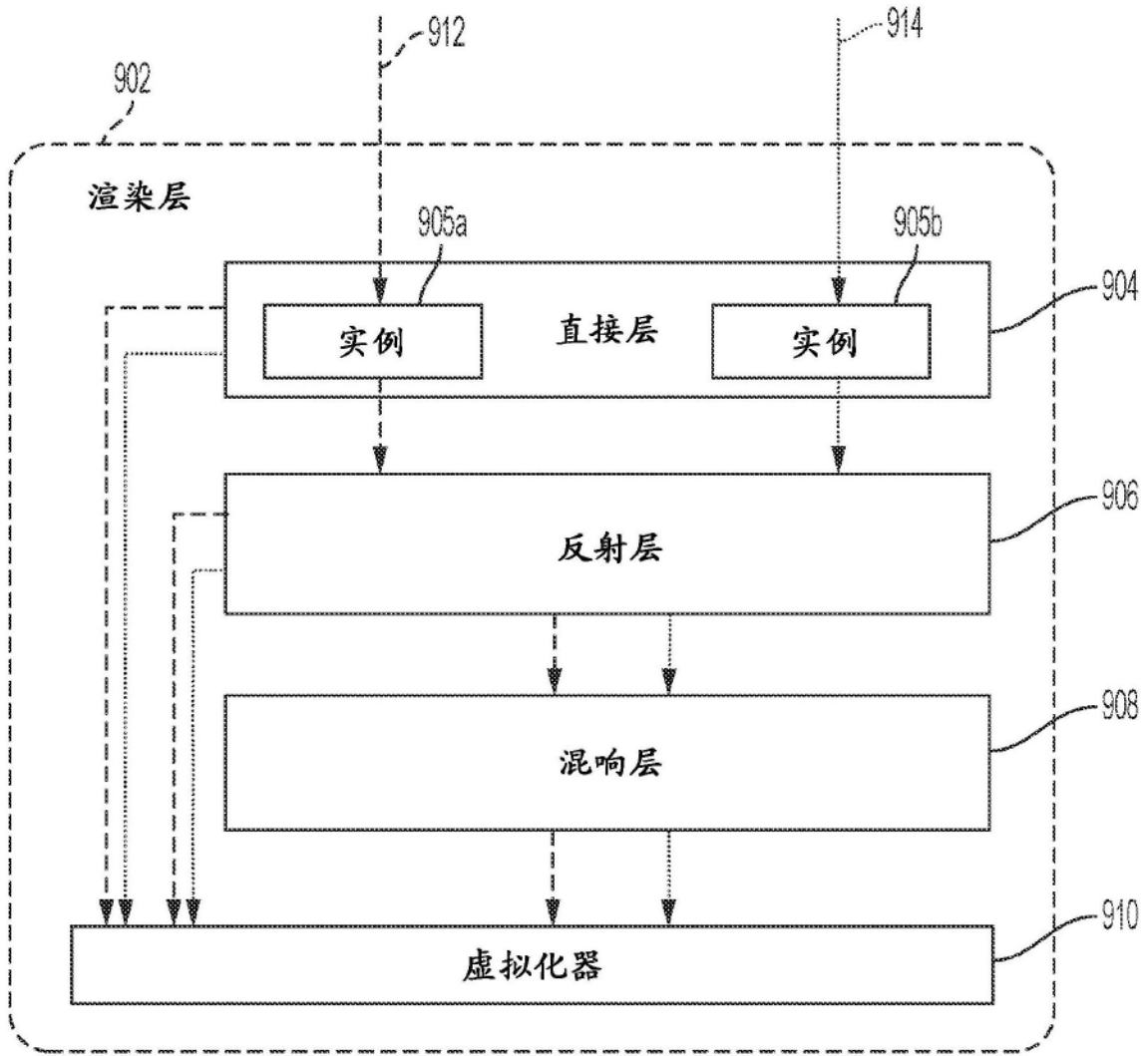


图9B