



(12) 发明专利申请

(10) 申请公布号 CN 112106020 A

(43) 申请公布日 2020.12.18

(21) 申请号 201980029718.4

(22) 申请日 2019.04.23

(30) 优先权数据

62/666,201 2018.05.03 US

(85) PCT国际申请进入国家阶段日

2020.11.02

(86) PCT国际申请的申请数据

PCT/US2019/028631 2019.04.23

(87) PCT国际申请的公布数据

WO2019/212794 EN 2019.11.07

(71) 申请人 苹果公司

地址 美国加利福尼亚州

(72) 发明人 A·巴-兹夫

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 李兴斌

(51) Int.Cl.

G06F 3/16 (2006.01)

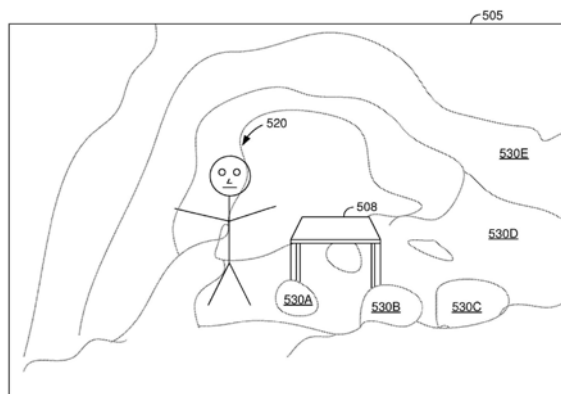
权利要求书2页 说明书11页 附图7页

(54) 发明名称

用于合成现实布景的声音处理的方法和设  
备

(57) 摘要

在一种具体实施中,一种将声音转换成用于合成现实(SR)布景的虚拟声音的方法由头戴式设备(HMD)执行,该头戴式设备包括一个或多个处理器、非暂态存储器、麦克风、扬声器和显示器。该方法包括:在显示器上显示合成现实(SR)布景的图像表示,该合成现实(SR)布景包括与SR布景的声学混响特性相关联的多个表面。该方法包括:经由麦克风记录在物理布景中产生的真实声音。该方法进一步包括:使用一个或多个处理器通过基于SR布景的声学混响特性转换真实声音来生成虚拟声音。该方法进一步包括:经由扬声器播放虚拟声音。



1. 一种方法,所述方法包括:

在位于物理布景中并且包括一个或多个处理器、非暂态存储器、麦克风、扬声器和显示器的头戴式设备处:

在所述显示器上显示合成现实(SR)布景的图像表示,所述合成现实(SR)布景包括与所述SR布景的声学混响特性相关联的多个表面;

经由所述麦克风记录在所述物理布景中产生的真实声音;

使用所述一个或多个处理器通过基于所述SR布景的所述声学混响特性转换所述真实声音来生成虚拟声音;以及

经由所述扬声器播放所述虚拟声音。

2. 根据权利要求1所述的方法,其中,所述多个表面分别位于所述SR布景中的多个位置处,并且其中,基于所述声学混响特性转换所述真实声音包括基于所述多个位置改变所述真实声音的混响量。

3. 根据权利要求2所述的方法,其中,改变所述混响量包括:当所述多个位置远离用户的虚拟位置时,将所述混响量改变为第一量,或者当所述多个位置靠近所述用户的所述虚拟位置时,将所述混响量改变为小于所述第一量的第二量。

4. 根据权利要求1至3中任一项所述的方法,其中,所述多个表面分别与多个声学阻尼特点相关联,并且基于所述声学混响特性转换所述真实声音包括基于所述多个声学阻尼特点改变所述真实声音的混响量。

5. 根据权利要求4所述的方法,其中,改变所述混响量包括:当所述多个声学阻尼特点低时,将所述混响量改变为第一量,或者当所述多个声学阻尼特点高时,将所述混响量改变为小于所述第一量的第二量。

6. 根据权利要求1至5中任一项所述的方法,进一步包括基于所述多个表面将所述SR布景分类为多个离散类中的一类,并且基于所述分类改变所述混响量。

7. 根据权利要求1至6中任一项所述的方法,其中,转换所述真实声音包括基于所述多个表面确定传递函数并且将所述传递函数应用于所述真实声音。

8. 根据权利要求1至7中任一项所述的方法,其中,播放所述虚拟声音进一步包括对所述真实声音进行噪声消除。

9. 根据权利要求1至8中任一项所述的方法,其中,生成所述虚拟声音进一步包括基于所述SR布景的一个或多个其他声学特性转换所述真实声音。

10. 根据权利要求1至9中任一项所述的方法,进一步包括在播放所述虚拟声音之后:

在所述显示器上显示第二SR布景的图像表示,所述第二SR布景具有与所述第二SR布景的第二声学混响特性相关联的多个表面,所述第二SR布景的所述第二声学混响特性不同于所述SR布景的所述声学混响特性;

经由所述麦克风记录在所述物理布景中生成的第二真实声音;

使用所述一个或多个处理器通过基于所述第二SR布景的所述第二声学混响特性转换所述第二真实声音来生成第二虚拟声音;以及

经由所述扬声器播放所述第二虚拟声音。

11. 一种设备,所述设备包括:

一个或多个处理器;

非暂态存储器；

一个或多个麦克风；

一个或多个SR显示器；以及

一个或多个程序，所述一个或多个程序存储在所述非暂态存储器中，当由所述一个或多个处理器执行时，所述一个或多个程序使得所述设备执行根据权利要求1至10所述的方法中的任一种方法。

12. 一种存储一个或多个程序的非暂态存储器，当由具有一个或多个麦克风和一个或多个SR显示器的设备的一个或多个处理器执行时，所述一个或多个程序使得所述设备执行根据权利要求1至10所述的方法中的任一种方法。

13. 一种设备，所述设备包括：

一个或多个处理器；

非暂态存储器；

一个或多个麦克风；

一个或多个SR显示器；以及

用于使得所述设备执行根据权利要求1至10所述的方法中的任一种方法的装置。

14. 一种设备，所述设备包括：

麦克风；

显示器；以及

一个或多个处理器，所述一个或多个处理器用于：

在所述显示器上显示合成现实(SR)布景的图像表示，所述合成现实(SR)布景包括与所述SR布景的声学混响特性相关联的多个表面；

经由所述麦克风记录在物理布景中产生的真实声音；

通过基于所述SR布景的所述声学混响特性转换所述真实声音来生成虚拟声音；以及

经由所述扬声器播放所述虚拟声音。

## 用于合成现实布景的声音处理的方法和设备

### 技术领域

[0001] 本公开整体涉及合成现实布景的声音处理,并且特别地涉及用于基于合成现实布景的声学混响特性将真实声音转换为虚拟声音的系统、方法和设备。

### 背景技术

[0002] 物理布景是指无需电子系统帮助个体就能够感知和/或个体能够交互的世界。物理布景(例如,物理森林)包括物理元素(例如,物理树、物理结构和物理动物)。个体可直接与物理布景相互作用和/或感知物理布景,诸如通过触摸、视觉、嗅觉、听觉和味觉。

[0003] 相比之下,合成现实(SR)布景是指经由电子系统个体能够感知并且/或者个体能够与其交互的完全或部分由计算机创建的布景。在SR中,监测个体移动的子集,并且响应于此,以符合一个或多个物理定律的方式使SR布景中的一个或多个虚拟对象的一个或多个属性变化。例如,SR系统可检测到个体向前走几步,并且响应于此,以类似于此类情景和声音在物理环境中会如何变化的方式来调节呈现给个体的图形和音频。也可响应于移动(例如,音频指令)而进行对SR布景中一个或多个虚拟对象的一个或多个属性的修改。

[0004] 个体可使用他的任何感觉(包括触觉、嗅觉、视觉、味觉和声音)与SR对象进行交互和/或感知SR对象。例如,个体可与创建多维(例如,三维)或空间听觉布景和/或实现听觉透明性的听觉对象进行交互和/或感知听觉对象。多维或空间听觉布景为个体提供了在多维空间中对离散听觉源的感知。在具有或不具有计算机创建的音频的情况下,听觉透明性选择性地结合来自物理布景的声音。在一些SR布景中,个体可仅与听觉对象进行交互和/或仅感知听觉对象。

[0005] SR的一个示例是虚拟现实(VR)。VR布景是指被设计为仅包括针对至少一种感觉的计算机创建的感官输入的模拟布景。VR布景包括个体可与之交互和/或对其进行感知的多个虚拟对象。个体可通过在计算机创建的布景内模拟个体动作的子集和/或通过对个体或其计算机创建的布景内的存在的模拟,来与VR布景中的虚拟对象进行交互和/或感知VR布景中的虚拟对象。

[0006] SR的另一个示例是混合现实(MR)。MR布景是指被设计为将计算机创建的感官输入(例如,虚拟对象)与来自物理布景的感官输入或其表示集成的模拟布景。在现实谱系上,混合现实布景介于一端的VR布景和另一端的完全物理布景之间并且不包括这些布景。

[0007] 在一些MR布景中,计算机创建的感官输入可以适应于来自物理布景的感官输入的变化。另外,用于呈现MR布景的一些电子系统可以监测相对于物理布景的取向和/或位置,以使虚拟对象能够与真实对象(即来自物理布景的物理元素或其表示)交互。例如,系统可监测运动,使得虚拟植物相对于物理建筑物看起来是静止的。

[0008] 混合现实的一个示例是增强现实(AR)。AR布景是指至少一个虚拟对象叠加在物理布景或其表示之上的模拟布景。例如,电子系统可具有不透明显示器和至少一个成像传感器,成像传感器用于捕获物理布景的图像或视频,这些图像或视频是物理布景的表示。系统将图像或视频与虚拟对象组合,并在不透明显示器上显示该组合。个体使用系统经由物理

布景的图像或视频间接地查看物理布景,并且观察叠加在物理布景之上的虚拟对象。当系统使用一个或多个图像传感器捕获物理布景的图像,并且使用那些图像在不透明显示器上呈现AR布景时,所显示的图像被称为视频透传。另选地,用于显示AR布景的电子系统可具有透明或半透明显示器,个体可通过该显示器直接查看物理布景。该系统可在透明或半透明显示器上显示虚拟对象,使得个体使用该系统观察叠加在物理布景之上的虚拟对象。又如,系统可包括将虚拟对象投影到物理布景中的投影系统。虚拟对象可例如在物理表面上或作为全息图被投影,使得个体使用该系统观察叠加在物理布景之上的虚拟对象。

[0009] 增强现实布景也可指其中物理布景的表示被计算机创建的感官信息改变的模拟布景。例如,物理布景的表示的一部分可被以图形方式改变(例如,放大),使得所改变的部分仍可代表一个或多个初始捕获的图像但不是忠实再现的版本。又如,在提供视频透传时,系统可改变传感器图像中的至少一者以施加不同于一个或多个图像传感器捕获的视点的特定视点。再如,物理布景的表示可通过以图形方式将其部分进行模糊处理或消除其部分而被改变。

[0010] 混合现实的另一个示例是增强虚拟(AV)。AV布景是指计算机创建布景或虚拟布景并入来自物理布景的至少一个感官输入的模拟布景。来自物理布景的一个或多个感官输入可为物理布景的至少一个特征的代表。例如,虚拟对象可呈现由一个或多个成像传感器捕获的物理元素的颜色。又如,虚拟对象可呈现出与物理布景中的实际天气条件相一致的特征,如经由天气相关的成像传感器和/或在线天气数据所识别的。在另一个示例中,增强现实森林可具有虚拟树木和结构,但动物可具有从对物理动物拍摄的图像精确再现的特征。

[0011] 许多电子系统使得个体能够与各种SR布景进行交互和/或感知各种SR布景。一个示例包括头戴式系统。头戴式系统可具有不透明显示器和一个或多个扬声器。另选地,头戴式系统可以被设计为接收外部显示器(例如,智能电话)。头戴式系统可具有分别用于拍摄物理布景的图像/视频和/或捕获物理布景的音频的一个或多个成像传感器和/或麦克风。头戴式系统也可具有透明或半透明显示器。透明或半透明显示器可结合基板,表示图像的光通过该基板被引导到个体的眼睛。显示器可结合LED、OLED、数字光投影仪、激光扫描光源、硅上液晶,或这些技术的任意组合。透射光的基板可以是光波导、光组合器、光反射器、全息基板或这些基板的任意组合。在一个实施方案中,透明或半透明显示器可在不透明状态与透明或半透明状态之间选择性地转换。又如,电子系统可以是基于投影的系统。基于投影的系统可使用视网膜投影将图像投影到个体的视网膜上。另选地,投影系统还可将虚拟对象投影到物理布景中(例如,投影到物理表面上或作为全息图)。SR系统的其他示例包括平视显示器、能够显示图形的汽车挡风玻璃、能够显示图形的窗口、能够显示图形的镜片、耳机或耳塞、扬声器布置、输入机构(例如,具有或不具有触觉反馈的控制器)、平板电脑、智能电话,以及台式计算机或膝上型计算机。

[0012] 当佩戴SR头戴式受话器的用户发出声音(例如,说话、吹口哨、拍手等)时,SR头戴式受话器可忽略该声音(允许用户自然听到声音)或记录该声音并且通过集成扬声器播放该声音。在任一种情况下,声音的特性可能都与SR头戴式受话器所呈现的SR布景中预期的那些特性不一致。因此,可能令人期望的是使SR头戴式受话器处理声音,使得其听起来好像是在所呈现的SR布景中产生的那样。

## 附图说明

[0013] 因此,本公开可被本领域的普通技术人员理解,更详细的描述可参考一些例示性具体实施的方面,其中一些具体实施在附图中示出。

[0014] 图1是根据一些具体实施的例示性操作环境的框图。

[0015] 图2是根据一些具体实施的例示性控制器的框图。

[0016] 图3是根据一些具体实施的例示性头戴式设备(HMD)的框图。

[0017] 图4示出了根据一些具体实施的例示性物理布景。

[0018] 图5示出了根据一些具体实施的第一SR布景的例示性图像表示。

[0019] 图6示出了根据一些具体实施的第二SR布景的例示性图像表示。

[0020] 图7是根据一些具体实施的基于SR布景来转换声音的方法的流程图表示。

[0021] 根据通常的做法,附图中示出的各种特征部可能未按比例绘制。因此,为了清楚起见,可以任意地扩展或减小各种特征部的尺寸。另外,一些附图可能未描绘给定的系统、方法或设备的所有部件。最后,在整个说明书和附图中,类似的附图标号可用于表示类似的特征部。

## 发明内容

[0022] 本文所公开的各种具体实施包括用于SR布景的声音处理的设备、系统和方法。在各种具体实施中,该方法在头戴式设备上执行,该头戴式设备位于物理布景中并且包括一个或多个处理器、非暂态存储器、麦克风、扬声器和显示器。该方法包括:在显示器上显示SR布景的图像表示,该SR布景包括与SR布景的声学混响特性相关联的多个表面。该方法包括:经由麦克风记录在物理布景中产生的真实声音。该方法进一步包括:使用一个或多个处理器通过基于SR布景的声学混响特性转换真实声音来生成虚拟声音。该方法进一步包括:经由扬声器播放虚拟声音。

[0023] 根据一些具体实施,一种设备包括一个或多个处理器、非暂态存储器以及一个或多个程序;一个或多个程序被存储于非暂态存储器中并且被配置为由一个或多个处理器执行。一个或多个程序包括用于执行或促使执行本文描述的任何方法的指令。根据一些具体实施,一种非暂态计算机可读存储介质中存储有指令,这些指令在由设备的一个或多个处理器执行时使该设备执行或使执行本文所述方法中的任一种。根据一些具体实施,一种设备包括:一个或多个处理器、非暂态存储器以及用于执行或使执行本文所述方法中的任一种的装置。

## 具体实施方式

[0024] 描述了许多细节以便提供对附图中所示的例示具体实施的透彻理解。然而,附图仅示出了本公开的一些例示方面,因此不应被视为限制。本领域的普通技术人员将理解,其他有效方面和/或变体不包括本文所述的所有具体细节。此外,没有详尽地描述众所周知的系统、方法、部件、设备和电路,以免模糊本文所述的例示性具体实施的更多相关方面。

[0025] 如上所述,当佩戴SR头戴式受话器的用户发出声音(例如,说话、吹口哨、拍手、跺脚等)或在用户的物理布景中产生另一声音时,SR头戴式受话器可忽略该声音(允许用户自然听到声音)或记录该声音并且通过集成扬声器播放该声音。然而,因为已发现音频是完全

沉浸式体验的重要方面,所以可能令人期望的是使SR头戴式受话器处理声音,使得其听起来好像是在所呈现的ST布景中产生的那样。作为示例,用户可处于其客厅的物理布景和音乐厅的SR布景中。记录在用户的客厅中产生的声音(例如,由用户或另一个人发出的言语)并且用回声回放该声音以使其听起来好像用户在(合成)音乐厅中说话那样。

[0026] 图1是根据一些具体实施的示例性操作环境100的框图。尽管示出了相关特征,但本领域的普通技术人员将从本公开中认识到,为简洁起见并且为了不模糊本文所公开的示例性具体实施的更多相关方面,未示出各种其他特征。为此,作为非限制性示例,操作环境100包括控制器110和头戴式设备(HMD)120。

[0027] 在一些具体实施中,控制器110被配置为管理和协调用户的合成现实(SR)体验。在一些具体实施中,控制器110包括软件、固件和/或硬件的合适组合。下文参考图2更详细地描述控制器110。在一些具体实施中,控制器110是相对于场景105处于本地或远程位置的计算设备。例如,控制器110是位于场景105内的本地服务器。又如,控制器110是位于场景105之外的远程服务器(例如,云服务器、中央服务器等)。在一些具体实施中,控制器110经由一个或多个有线或无线通信信道144(例如,蓝牙、IEEE 802.11x、IEEE 802.16x、IEEE 802.3x等)与HMD 120通信耦接。

[0028] 在一些具体实施中,HMD 120被配置为向用户呈现SR体验。在一些具体实施中,HMD 120包括软件、固件和/或硬件的合适组合。下文参考图3更详细地描述HMD 120。在一些具体实施中,控制器110的功能由HMD120提供和/或与HMD 120结合。

[0029] 根据一些具体实施,当用户虚拟地和/或物理地存在于场景105内时,HMD 120向用户呈现合成现实(SR)体验。在一些具体实施中,在呈现增强现实(AR)体验时,HMD 120被配置为呈现AR内容。在一些具体实施中,在呈现虚拟现实(VR)体验时,HMD 120被配置为呈现VR内容。

[0030] 在一些具体实施中,用户将HMD 120戴在头上。因此,HMD 120包括被提供用于显示SR内容的一个或多个SR显示器。例如,在各种具体实施中,HMD 120包围用户的视野。在一些具体实施中,使用被配置为呈现SR内容的手持设备(诸如智能电话或平板电脑)代替HMD 120,其中用户不佩戴HMD 120而是手持该设备,同时使显示器朝向用户视野,并且使相机朝向场景105。在一些具体实施中,采用被配置为呈现SR内容的SR舱、外壳或室代替HMD 120,用户在其中不再佩戴或手持HMD 120。

[0031] 图2是根据一些具体实施的控制器110的示例的框图。尽管示出了一些具体特征,但本领域的技术人员将从本公开中认识到,为简洁起见并且为了不模糊本文所公开的具体实施的更多相关方面,未示出各种其他特征。为此,作为非限制性示例,在一些具体实施中,控制器110包括一个或多个处理单元202(例如,微处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、图形处理单元(GPU)、中央处理单元(CPU)、处理核心等)、一个或多个输入/输出(I/O)设备206、一个或多个通信接口208(例如,通用串行总线(USB)、FIREWIRE、THUNDERBOLT、IEEE 802.3x、IEEE 802.11x、IEEE 802.16x、全球移动通信系统(GSM)、码分多址(CDMA)、时分多址(TDMA)、全球定位系统(GPS)、红外(IR)、蓝牙、ZIGBEE和/或相似类型接口)、一个或多个编程(例如,I/O)接口210、存储器220以及用于互连这些部件和各种其他部件的一条或多条通信总线204。

[0032] 在一些具体实施中,所述一条或多条通信总线204包括互连系统部件和控制系统

部件之间的通信的电路。在一些具体实施中,一个或多个I/O设备206包括键盘、鼠标、触控板、操纵杆、一个或多个麦克风、一个或多个扬声器、一个或多个图像传感器、一个或多个显示器等中的至少一种。

[0033] 存储器220包括高速随机存取存储器,诸如动态随机存取存储器(DRAM)、静态随机存取存储器(SRAM)、双倍数据速率随机存取存储器(DDR RAM)或者其他随机存取固态存储器设备。在一些具体实施中,存储器220包括非易失性存储器,诸如一个或多个磁盘存储设备、光盘存储设备、闪存存储器设备或其他非易失性固态存储设备。存储器220任选地包括远离一个或多个处理单元202定位的一个或多个存储设备。存储器220包括非暂态计算机可读存储介质。在一些具体实施中,存储器220或存储器220的非暂态计算机可读存储介质存储下述程序、模块和数据结构或它们的子集,其中包括任选的操作系统230和合成现实(SR)体验模块240。

[0034] 操作系统230包括用于处理各种基础系统服务和用于执行硬件相关任务的过程。在一些具体实施中,SR体验模块240被配置为管理和协调一个或多个用户的单重或多重SR体验(例如,一个或多个用户的单重SR体验,或一个或多个用户的相应群组的多重SR体验)。为此,在各种具体实施中,SR体验模块240包括数据获取单元242、跟踪单元244、协调单元246和数据传输单元248。

[0035] 在一些具体实施中,数据获取单元242被配置为至少从HMD 120获取数据(例如,呈现数据、交互数据、传感器数据、位置数据等)。为此,在各种具体实施中,数据获取单元242包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0036] 在一些具体实施中,跟踪单元244被配置为映射场景105并且至少跟踪HMD 120相对于场景105的位置/定位。为此,在各种具体实施中,跟踪单元244包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0037] 在一些具体实施中,协调单元246被配置为管理和协调HMD 120向用户呈现的SR体验。为此,在各种具体实施中,协调单元246包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0038] 在一些具体实施中,数据传输单元248被配置为至少向HMD 120传输数据(例如,呈现数据、位置数据等)。为此,在各种具体实施中,数据传输单元248包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0039] 尽管数据获取单元242、跟踪单元244、协调单元246和数据传输单元248被示为驻留在单个设备(例如,控制器110)上,但应当理解,在其他具体实施中,数据获取单元242、跟踪单元244、协调单元246和数据传输单元248的任何组合可位于单独计算设备中。

[0040] 此外,图2更多地用作存在于特定具体实施中的各种特征部的功能描述,与本文所述的具体实施的结构示意图不同。如本领域的普通技术人员将认识到的,单独显示的项目可以组合,并且一些项目可以分开。例如,图2中单独示出的一些功能模块可以在单个模块中实现,并且单个功能块的各种功能可在各种具体实施中通过一个或多个功能块来实现。模块的实际数量和特定功能的划分以及如何在其中分配特征将根据具体实施而变化,并且在一些具体实施中,部分地取决于为特定实施方案选择的硬件、软件和/或固件的特定组合。

[0041] 图3是根据一些具体实施的头戴式设备(HMD) 120的示例的框图。尽管示出了一些



具体特征,但本领域的技术人员将从本公开中认识到,为简洁起见并且为了不模糊本文所公开的具体实施的更多相关方面,未示出各种其他特征。为此,作为非限制性示例,在一些具体实施中,HMD 120包括一个或多个处理单元302(例如,微处理器、ASIC、FPGA、GPU、CPU、处理核心等)、一个或多个输入/输出(I/O)设备及传感器306、一个或多个通信接口308(例如,USB、FIREWIRE、THUNDERBOLT、IEEE802.3x、IEEE 802.11x、IEEE 802.16x、GSM、CDMA、TDMA、GPS、IR、BLUETOOTH、ZIGBEE和/或类似类型的接口)、一个或多个编程(例如,I/O)接口310、一个或多个SR显示器312、一个或多个任选的面向内部和/或面向外部的图像传感器314、存储器320以及用于互连这些部件和各种其他部件的一条或多条通信总线304。

[0042] 在一些具体实施中,一条或多条通信总线304包括互连和控制系统部件之间的通信的电路。在一些具体实施中,一个或多个I/O设备和传感器306包括惯性测量单元(IMU)、加速度计、陀螺仪、温度计、一个或多个生理传感器(例如,血压监测仪、心率监测仪、血氧传感器、血糖传感器等)、一个或多个麦克风307A、一个或多个扬声器307B、触觉引擎以及/或者一个或多个深度传感器(例如,结构光、飞行时间等)等中的至少一者。

[0043] 在一些具体实施中,一个或多个SR显示器312被配置为向用户呈现SR体验。在一些具体实施中,一个或多个SR显示器312对应于全息、数字光处理(DLP)、液晶显示器(LCD)、硅上液晶(LCoS)、有机发光场效应晶体管(OLET)、有机发光二极管(OLED)、表面传导电子发射显示器(SED)、场发射显示器(FED)、量子点发光二极管(QD-LED)、微机电系统(MEMS)和/或类似的显示器类型。在一些具体实施中,一个或多个SR显示器312对应于衍射、反射、偏振、全息等波导显示器。例如,HMD 120包括单个SR显示器。在另一示例中,HMD 120包括针对用户的每只眼睛的SR显示器。在一些具体实施中,一个或多个SR显示器312能够呈现AR和VR内容。在一些具体实施中,一个或多个SR显示器312能够呈现AR或VR内容。

[0044] 在一些具体实施中,一个或多个图像传感器314被配置为获取与用户面部的包括用户眼睛的至少一部分对应的图像数据。例如,一个或多个任选的图像传感器314对应于一个或多个RGB相机(例如,具有互补金属氧化物半导体(CMOS)图像传感器或电荷耦合器件(CCD)图像传感器)、红外(IR)相机、基于事件的相机等。

[0045] 存储器320包括高速随机存取存储器,诸如DRAM、SRAM、DDR RAM或其他随机存取固态存储器设备。在一些具体实施中,存储器320包括非易失性存储器,诸如一个或多个磁盘存储设备、光盘存储设备、闪存存储设备或其他非易失性固态存储设备。存储器320任选地包括远离一个或多个处理单元302定位的一个或多个存储设备。存储器320包括非暂态计算机可读存储介质。在一些具体实施中,存储器320或者存储器320的非暂态计算机可读存储介质存储下述程序、模块和数据结构或者它们的子集,其中包括任选的操作系统330和SR呈现模块340。

[0046] 操作系统330包括用于处理各种基础系统服务和用于执行硬件相关任务的过程。在一些具体实施中,SR呈现模块340被配置为经由一个或多个SR显示器312向用户呈现SR内容。为此,在各种具体实施中,SR呈现模块340包括数据获取单元342、SR呈现单元344、声音处理单元346和数据传输单元348。

[0047] 在一些具体实施中,数据获取单元342被配置为至少从控制器110获取数据(例如,呈现数据、交互数据、传感器数据、位置数据等)。为此,在各种具体实施中,数据获取单元342包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0048] 在一些具体实施中,SR呈现单元344被配置为经由一个或多个SR显示器312呈现SR内容。为此,在各种具体实施中,SR呈现单元344包括指令和/或用于指令的逻辑以及启发法和用于启发法的元数据。

[0049] 在一些具体实施中,声音处理单元346被配置为通过基于SR布景的声学混响特性转换真实声音来生成虚拟声音。为此,在各种具体实施中,声音处理单元346包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0050] 在一些具体实施中,数据传输单元348被配置为至少向控制器110传输数据(例如,呈现数据、位置数据等)。为此,在各种具体实施中,数据传输单元348包括指令和/或用于指令的逻辑部件以及启发法和用于启发法的元数据。

[0051] 尽管数据获取单元342、SR呈现单元344、声音处理单元346和数据传输单元348被示为驻留在单个设备(例如,HMD 120)上,但应当理解,在其他具体实施中,数据获取单元342、SR呈现单元344、声音处理单元346和数据传输单元348的任何组合可位于单独计算设备中。

[0052] 此外,图3更多地用作存在于特定具体实施中的各种特征部的功能描述,与本文所述的具体实施的结构示意图不同。如本领域的普通技术人员将认识到的,单独显示的项目可以组合,并且一些项目可以分开。例如,图3中单独示出的一些功能模块可以在单个模块中实现,并且单个功能块的各种功能可在各种具体实施中通过一个或多个功能块来实现。模块的实际数量和特定功能的划分以及如何在其中分配特征将根据具体实施而变化,并且在一些具体实施中,部分地取决于为特定实施方案选择的硬件、软件和/或固件的特定组合。

[0053] 图4示出了根据一些具体实施的示例性物理布景405。物理布景405包括第一用户(未示出,但从其角度来看,示出了物理布景)、第二用户420和桌子408。物理布景405包括多个表面430A-430D,该多个表面430A-430D包括地板表面430A、两个墙壁表面430B-430C和天花板表面430D。

[0054] 当在物理布景405中产生声音时,声音沿着多个路径行进到第一用户的耳朵,潜在地从表面430A-430D反射出来。因此,物理布景405向所产生的声音引入一定量的混响。

[0055] 图5示出了根据一些具体实施的由物理地位于图4的物理布景405中的用户体验的第一SR布景505(例如,洞穴环境)的示例性图像表示。第一SR布景505的图像表示包括第二用户520的图像表示、桌子508的图像表示和各种虚拟对象(例如,岩石和墙壁)的图像表示。第一SR布景505包括分别位于第一SR布景505中的第一多个位置处的第一多个表面。因此,第一SR布景505的图像表示包括在第一多个位置处的第一多个表面530A-530E的图像表示,该第一多个表面530A-530E包括三个岩石表面530A-530C、洞穴墙壁表面530D和洞穴天花板表面530E。

[0056] 当在物理布景405中产生声音时,物理布景405引入一定量的混响。然而,该混响量可能与将由第一SR布景505的真实版本(例如,洞穴)引入的混响量不一致。特别地,可以预期的是,洞穴将引入较大混响量,特别是与小室内房间中的混响量相比。因此,在各种具体实施中,为了产生更沉浸式的体验,HMD记录在物理布景中产生的真实声音,通过转换真实声音以改变(例如,增加)与真实声音相关联的混响量来生成虚拟声音,并且将虚拟声音播放给第一用户。换言之,当第一用户说话时,其HMD拾取其语音,引入与SR布景可能预期的混

响一致的混响,并且利用该HMD(以及参与同一SR布景的其他用户的HMD)的扬声器中的附加混响回放其语音,从而增加模拟现实体验的真实性水平。

[0057] 图6示出了根据一些具体实施的由物理地位于图4的物理布景405中的用户体验的第二SR布景605(例如,海滩环境)的示例性图像表示。第二SR布景605的图像表示包括第二用户520的图像表示、桌子508的图像表示和各种虚拟对象(例如,沙子、水和叶子)的图像表示。第二SR布景605包括分别位于第二SR布景605中的多个第二位置处的第二多个表面。因此,第二SR布景605的图像表示包括在第二多个位置处的第二多个表面630A-630D的图像表示,该第二多个表面630A-630D包括沙床表面630A、海水表面630B以及两个叶子表面630C和630D。

[0058] 如上所述,当在物理布景405中产生声音时,物理布景405引入一定量的混响。然而,该混响量可能与将由第二SR布景605的真实版本引入的混响量不一致。特别地,可以预期的是,海滩将引入很少的混响甚至不引入混响,特别是与没有声学阻尼的大房间(例如,具有木地板的大型客厅、洞穴等)中的混响量相比。因此,在各种具体实施中,为了产生更沉浸式的体验,HMD记录在物理布景405中产生的真实声音,通过转换真实声音以改变(例如,减少)与真实声音相关联的混响量来生成虚拟声音,并且将虚拟声音播放给第一用户。

[0059] 特别地,应当理解,预期第一SR布景505和第二SR布景605的真实版本将引入不同的混响量。在确定预期将由SR布景的真实版本引入的混响量时,可使用多个表面的特性,诸如,它们的位置(例如,相对于彼此或相对于用户的位置)、声学阻尼特点或其他特性。

[0060] 图7是根据一些具体实施的基于SR布景来转换声音的方法700的流程图表示。在各种具体实施中,方法700由位于物理布景中的具有一个或多个处理器、非暂态存储器、麦克风、扬声器和一个或多个SR显示器的设备(例如,图3中的HMD 120)执行。在一些具体实施中,方法700由处理逻辑部件(包括硬件、固件、软件或其组合)执行。在一些具体实施中,方法700由执行存储在非暂态计算机可读介质(例如,存储器)中的指令(例如,代码)的处理器执行。简而言之,在一些情况下,方法700包括:记录在物理布景中产生的真实声音;通过基于合成现实布景的声学混响特性转换真实声音来生成虚拟声音;以及播放虚拟声音。

[0061] 在框700中,方法700始于设备在显示器上显示SR布景的图像表示,该SR布景具有与SR布景的声学混响特性相关联的多个表面。例如,图5示出了第一SR布景505的图像表示,并且图6示出了第二SR布景605的图像表示。

[0062] 在各种具体实施中,SR布景的声学混响特性指示预期将由SR布景的真实版本引入的混响量(例如,音量或持续时间)。在各种具体实施中,声学混响特性指示SR布景的脉冲响应函数(或传递函数)。

[0063] 在各种具体实施中,多个表面分别位于SR布景中的多个位置处。因此,SR布景的图像表示包括多个位置处的表面的图像表示。在各种具体实施中,多个位置限定SR布景的声学混响特性。例如,当多个表面远离用户的虚拟位置时,声学混响特性指示较大混响量,并且当多个表面靠近用户的虚拟位置时,声学混响特性指示较小的混响量。作为另一示例,当多个表面被定向为面向用户的虚拟位置时,声学混响特性指示较大混响量,并且当多个表面未被定向为面向用户的虚拟位置时,声学混响特性指示较小的混响量。

[0064] 在各种具体实施中,多个表面分别与多个声学阻尼特点相关联。在各种具体实施中,多个声学阻尼特点限定SR布景的声学混响特性。例如,当多个声学阻尼特点高(例如,指

示较高程度的声学阻尼)时,声学混响特性指示较低的混响量,并且当多个声学阻尼特点低(例如,指示较低程度的声学阻尼)时,声学混响特性指示较高的混响量。

[0065] 作为示例,第一SR布景(在图5中示出了其图像表示505)包括多个表面530A-530E,该多个表面530A-530E远离第一用户的虚拟位置定位(因为洞穴较大)、被定向为面向第一用户(因为墙壁和天花板是弯曲的并且面向多个方向)并且与低声学阻尼特点相关联(因为岩石能反射较大比例的声音)。因此,第一SR布景的声学混响特性指示较大混响量。

[0066] 作为另一示例,第二SR布景(在图6中示出了其图像表示605)包括多个表面630A-630D,该多个表面630A-630D进一步远离第一用户的虚拟位置定位(因为海洋浩瀚无边)、未被定向为面向第一用户(因为沙子和水是平坦的,通常面朝上)并且与高声学阻尼特点相关联(因为沙子和叶子是吸声的)。因此,第二SR布景的声学混响特性指示较小混响量。

[0067] 在框720处,方法700继续,其中设备经由麦克风记录在物理布景中产生的真实声音。在各种具体实施中,真实声音由第一用户产生(例如,通过说话或拍手)。在各种具体实施中,真实声音由第二用户产生(例如,通过说话或跺脚)。在各种具体实施中,真实声音由物理布景中的其他噪声源(例如,立体声系统)产生。在各种具体实施中,设备取决于声音的源头以不同方式转换(如下文相对于框730所描述)。

[0068] 在框730处,方法700继续,其中设备通过基于SR布景的声学混响特性转换真实声音来生成虚拟声音。特别地,在各种具体实施中,设备转换真实声音,使得其听起来好像是在SR布景的真实版本中产生的那样。

[0069] 在各种具体实施中,设备通过改变(例如,增加或减少)真实声音的混响量基于声学混响特性来转换真实声音。在各种具体实施中,这包括:确定真实声音的混响量;确定虚拟声音的混响量;以及改变真实声音的混响量以匹配虚拟声音的所确定的混响量。

[0070] 在各种具体实施中,多个表面分别位于SR布景中的多个位置处,并且基于声学混响特性转换真实声音包括:基于多个位置改变真实声音的混响量。例如,在各种具体实施中,改变混响量包括:当多个位置远离用户的虚拟位置时,将混响量改变为第一量,或者当多个位置靠近用户的虚拟位置时,将混响量改变为小于第一量的第二量。

[0071] 在各种具体实施中,多个表面分别与多个声学阻尼特点相关联,并且基于声学混响特性转换真实声音包括:基于多个声学阻尼特点改变真实声音的混响量。例如,在各种具体实施中,改变混响量包括:当多个声学阻尼特点低时,将混响量改变为第一量,或者当多个声学阻尼特点高时,将混响量改变为小于第一量的第二量。

[0072] 在各种具体实施中,设备基于多个表面(例如,其位置、声学阻尼特点和/或其他特性)将SR布景分类为多个离散类中的一类。例如,多个离散类可包括大房间(与较高混响量相关联)、小房间(与较小混响量相关联)或户外位置(与较小混响量相关联)。作为另一示例,多个离散类可包括具有较小体积的封闭空间(例如,天花板较低的客厅)和具有较大体积的封闭空间(例如,带拱形天花板的同一客厅)。

[0073] 在各种具体实施中,设备基于多个表面(例如,其位置和/或声学阻尼特点)确定传递函数。例如,设备可应用光线跟踪算法来确定传递函数。因此,在各种具体实施中,设备通过将传递函数应用于真实声音来转换真实声音。在各种具体实施中,设备进一步均衡真实声音以补偿物理布景的传递函数(例如,使用MMSE(最小均方误差)或迫零)。

[0074] 在各种具体实施中,除了基于SR布景的声学混响特性来转换真实声音之外,设备

还通过基于SR布景的一个或多个其他声学特性进一步转换真实声音来生成虚拟声音。在各种具体实施中,设备过滤真实声音。例如,由于SR布景在水下,因此设备低通过滤真实声音,或者由于SR布景在空间中,因此设备高通过滤真实声音。在各种具体实施中,设备调节真实声音的频率。例如,由于SR布景是充满氦气的房间,因此设备增加真实声音(例如,第二用户的语音)的频率。如上所述,在各种具体实施中,设备取决于声音的源头以不同方式转换真实声音。因此,设备可调节第二用户的语音的频率,但不调节环境声音的频率。在各种具体实施中,设备调节真实声音的音量。例如,由于SR布景是摇滚音乐会,因此设备增加声音的音量,或者由于SR布景是图书馆,因此设备减小声音的音量。

[0075] 在框740处,方法700继续,其中设备经由扬声器播放虚拟声音。在各种具体实施中,播放虚拟声音进一步包括:对真实声音进行噪声消除。因此,用户仅听到处理过的声音,而不是听到处理过的声音和真实声音两者。

[0076] 在各种具体实施中,方法700在不同SR布景中重复,其中基于不同SR布景的不同声学混响特性进行不同的转换。因此,在各种具体实施中,方法700进一步包括:在显示器上显示(如相对于框710所描述)第二SR布景的图像表示,该第二SR布景具有与第二SR布景的第二声学混响特性相关联的多个表面,该第二SR布景的第二声学混响特性不同于SR布景的声学混响特性;经由麦克风记录(如相对于框720所描述)在物理布景中生成的第二真实声音;通过基于第二SR布景的第二声学混响特性转换第二真实声音来生成(如相对于框730所描述)第二虚拟声音;以及经由扬声器播放(如相对于框740所描述)第二虚拟声音。

[0077] 虽然上文描述了在所附权利要求书范围内的具体实施的各个方面,但是应当显而易见的是,上述具体实施的各种特征可通过各种各样的形式体现,并且上述任何特定结构和/或功能仅是例示性的。基于本公开,本领域的技术人员应当理解,本文所述的方面可以独立于任何其他方面来实现,并且这些方面中的两个或更多个可以采用各种方式组合。例如,可以使用本文阐述的任何数量的方面来实现装置和/或可以实践方法。另外,除了本文阐述的一个或多个方面之外或者不同于本文阐述的一个或多个方面,可以使用其他结构和/或功能来实现这样的装置和/或可以实践这样的方法。

[0078] 还将理解的是,虽然术语“第一”、“第二”等可能在本文中用于描述各种元素,但是这些元素不应当被这些术语限定。这些术语只是用于将一个元件与另一元件区分开。例如,第一节点可以被称为第二节点,并且类似地,第二节点可以被称为第一节点,其改变描述的含义,只要所有出现的“第一节点”被一致地重命名并且所有出现的“第二节点”被一致地重命名。第一节点和第二节点都是节点,但它们不是同一个节点。

[0079] 本文中所使用的术语仅仅是为了描述特定具体实施并非旨在对权利要求进行限制。如在本具体实施的描述和所附权利要求中所使用的那样,单数形式的“一个”和“该”旨在也涵盖复数形式,除非上下文清楚地另有指示。还将理解的是,本文中所使用的术语“和/或”是指并且涵盖相关联的所列出的项目中的一个或多个项目的任何和全部可能的组合。还将理解的是,术语“包括”在本说明书中使用是指定存在所陈述的特征、整数、步骤、操作、元件和/或部件,但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、部件,和/或其分组。

[0080] 如本文所使用的,术语“如果”可以被解释为表示“当所述先决条件为真时”或“在所述先决条件为真时”或“响应于确定”或“根据确定”或“响应于检测到”所述先决条件为

真,具体取决于上下文。类似地,短语“如果确定[所述先决条件为真]”或“如果[所述先决条件为真]”或“当[所述先决条件为真]时”被解释为表示“在确定所述先决条件为真时”或“响应于确定”或“根据确定”所述先决条件为真或“当检测到所述先决条件为真时”或“响应于检测到”所述先决条件为真,具体取决于上下文。

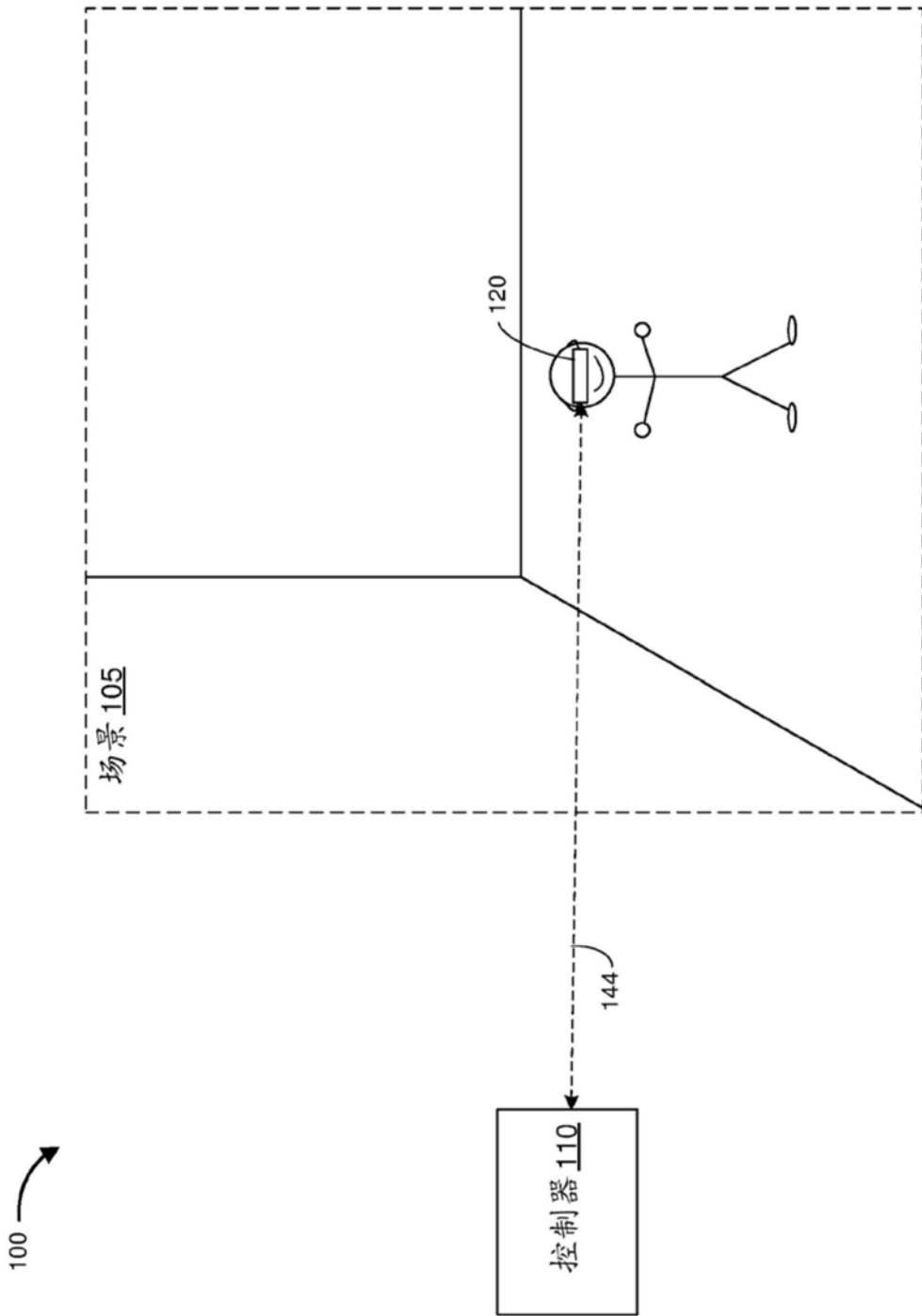


图1

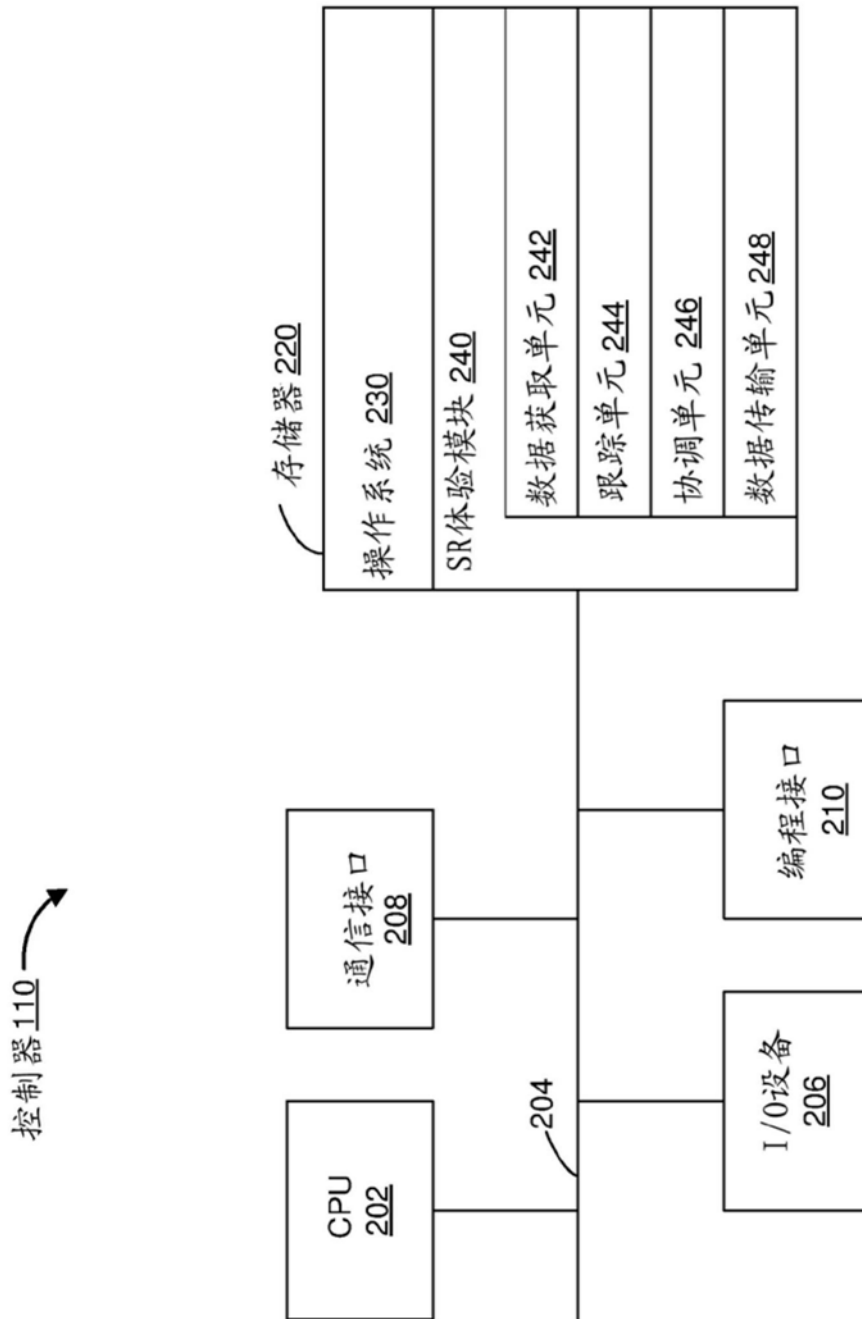


图2



头戴式设备 (HMD) 120

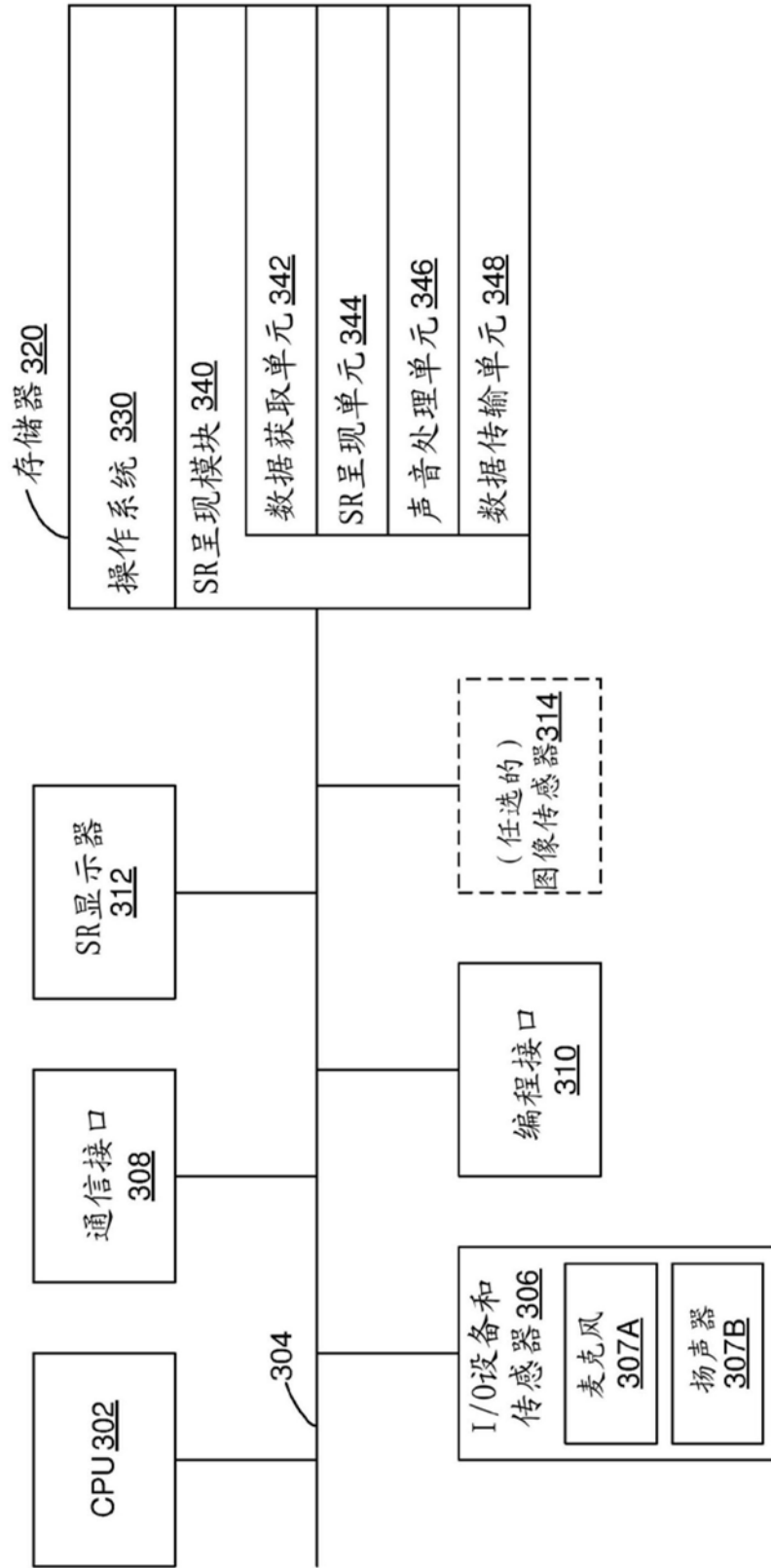


图3

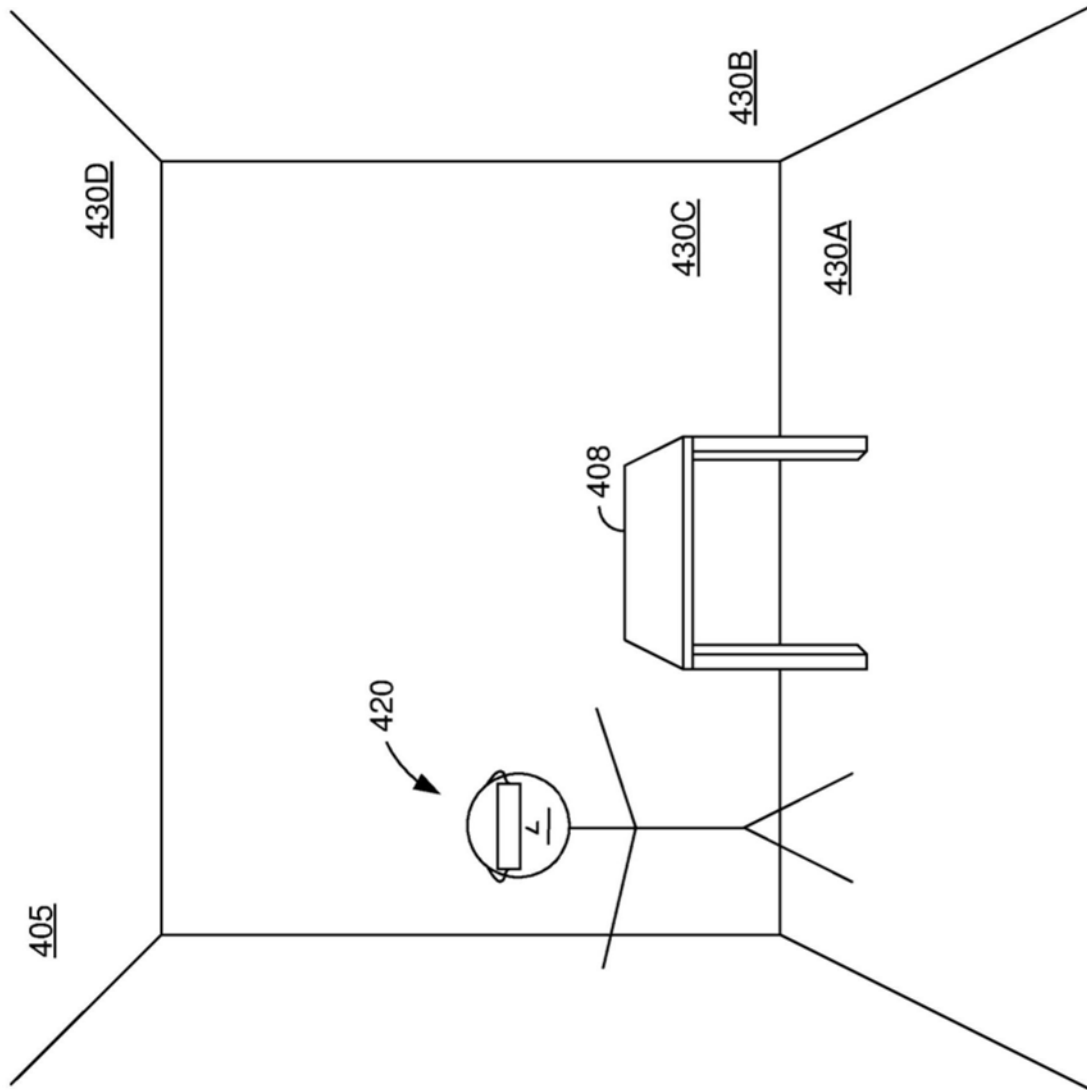


图4

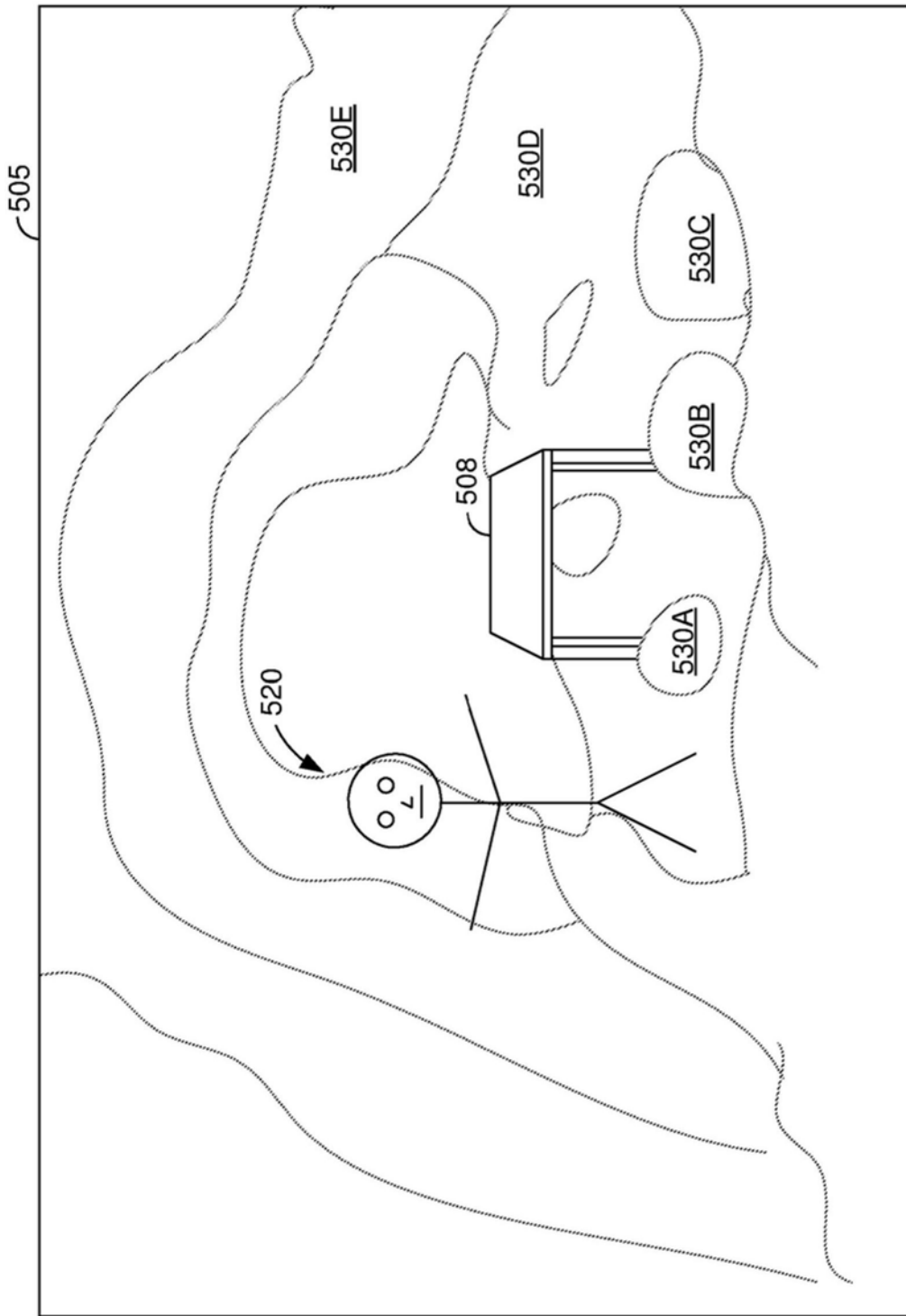


图5

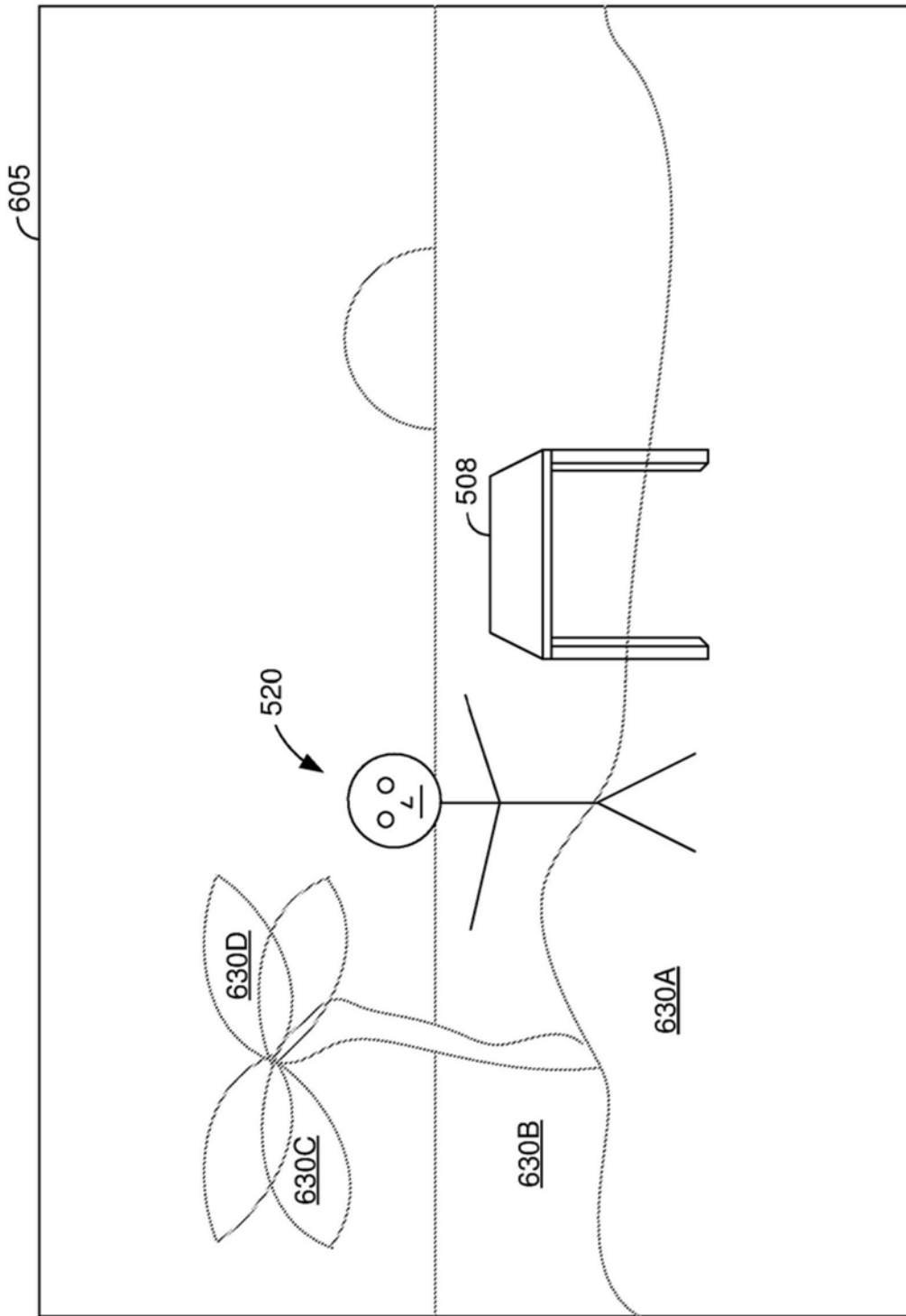


图6

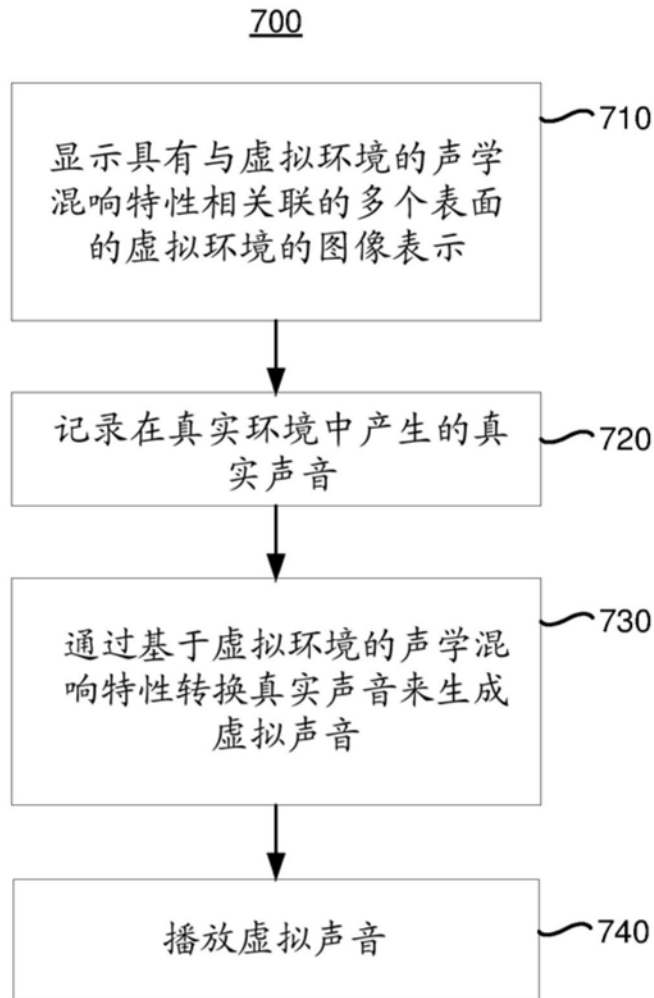


图7