(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117897687 A
(43)申请公布日 2024.04.16

(21)申请号 202280059219.1

(22)申请日 2022.06.28

(30)优先权数据
17/467,883 2021.09.07 US

(85)PCT国际申请进入国家阶段日
2024.02.29

(86)PCT国际申请的申请数据
PCT/US2022/073199 2022.06.28

(87)PCT国际申请的公布数据
WO2023/039308 EN 2023.03.16

(71)申请人 高通股份有限公司
地址 美国加利福尼亚州

(72)发明人 滕迪炎 J·韩 R·库尔卡尼

(74)专利代理机构 北京市柳沈律师事务所 11105
专利代理师 安之斐

(51)Int.Cl.
G06F 3/16(2006.01)

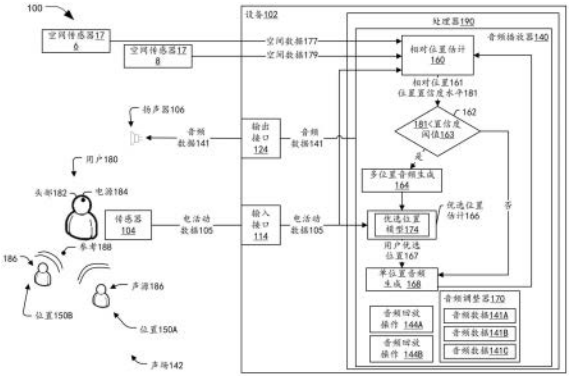
权利要求书3页 说明书21页 附图10页

(54)发明名称

基于用户电信号的音频调整

(57)摘要

一种设备包括存储器和一个或多个处理器。该存储器被配置为存储指令。该一个或多个处理器被配置为执行该指令以获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据。该一个或多个处理器还被配置为执行该指令以基于该电活动数据呈现音频数据,以在该音频数据的回放期间调整声源在声场中的位置。



CN 117897687 A

1.一种设备,包括:

存储器,所述存储器被配置为存储指令;和

一个或多个处理器,所述一个或多个处理器被配置为执行所述指令以:

获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据;以及

基于所述电活动数据呈现音频数据,以在所述音频数据的回放期间调整声源在声场中的位置。

2.根据权利要求1所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以经由一个或多个扬声器输出所述音频数据。

3.根据权利要求1所述的设备,其中所述电活动数据包括眼电图(EOG)数据、脑电图(EEG)数据或两者。

4.根据权利要求1所述的设备,还包括被配置为从一个或多个传感器接收所述电活动数据的接口。

5.根据权利要求4所述的设备,其中所述接口包括以太网接口、通用串行总线(USB)接口、Wi-Fi接口、蓝牙接口、串行端口接口或并行端口接口中的至少一者。

6.根据权利要求4所述的设备,其中所述一个或多个传感器包括入耳式传感器。

7.根据权利要求4所述的设备,其中所述一个或多个传感器包括电极帽、神经植入物、导电屏、非可穿戴传感器或它们的组合。

8.根据权利要求1所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以:

经由一个或多个扬声器启动所述音频数据的第一回放操作,所述音频数据在所述第一回放操作期间被呈现为包括所述声源在所述声场中的多个位置,其中所述电活动数据基于在所述音频数据的所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及

基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述位置。

9.根据权利要求8所述的设备,其中所述一个或多个处理器被配置为执行所述指令以基于确定位置置信度水平是否小于置信度阈值而确定是否将所述音频数据呈现为包括所述多个位置,其中所述位置置信度水平与所述用户相对于参考的估计位置相关联。

10.根据权利要求9所述的设备,其中所述用户相对于所述参考的所述估计位置包括所述用户相对于所述参考的估计取向的估计取向、所述用户相对于所述参考的估计距离或两者。

11.根据权利要求9所述的设备,其中所述参考包括所述设备、显示设备、物理参考或虚拟参考中的一者或多者。

12.根据权利要求9所述的设备,其中所述一个或多个处理器被配置为执行所述指令以将所述估计位置初始化为对应于取向在所述用户前方的所述参考。

13.根据权利要求9所述的设备,其中所述一个或多个处理器被配置为将所述位置置信度水平初始化为小于位置阈值。

14.根据权利要求9所述的设备,其中所述一个或多个处理器被配置为执行所述指令以更新所述用户相对于所述参考的估计位置,所述估计位置是基于所述用户的头戴式耳机的惯性测量单元(IMU)数据、所述参考的空间数据、用户注视估计或它们的组合来更新的。

15.根据权利要求1所述的设备,其中所述一个或多个处理器被配置为执行所述指令以:

基于所述电活动数据来确定用户注视估计;以及

至少部分地基于所述用户注视估计来更新所述用户相对于参考的估计位置。

16.根据权利要求1所述的设备,其中所述一个或多个处理器被配置为执行所述指令以:

使用机器学习模型来处理所述电活动数据以确定用户注视估计;以及

基于所述用户注视估计来更新所述用户相对于参考的估计位置。

17.根据权利要求1所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以使用机器学习模型来处理所述电活动数据以确定所述声源的用户优选位置。

18.根据权利要求1所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以:

将所述音频数据呈现为在所述声场中包括语音源的第一位置和非语音源的第二位置;

经由一个或多个扬声器启动所述音频数据的第一回放操作,其中所述电活动数据基于在所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及

基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述声源的所述位置。

19.根据权利要求18所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以响应于确定所述电活动数据指示单个声源被跟踪而确定所述用户优选位置对应于所述语音源的所述第一位置。

20.根据权利要求18所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以响应于确定所述电活动数据指示所述语音源和所述非语音源被跟踪而确定所述用户优选位置对应于所述非语音源的所述第二位置。

21.一种方法,包括:

在设备处获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据;以及

基于所述电活动数据呈现音频数据,以在所述音频数据的回放期间调整声源在声场中的位置。

22.根据权利要求21所述的方法,还包括:

经由一个或多个扬声器启动所述音频数据的第一回放操作,所述音频数据在所述第一回放操作期间被呈现为包括所述声源在所述声场中的多个位置,其中所述电活动数据基于在所述音频数据的所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及

基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述位置。

23.根据权利要求22所述的方法,还包括基于确定位置置信度水平是否小于置信度阈值而确定是否将所述音频数据呈现为包括所述多个位置,其中所述位置置信度水平与所述用户相对于参考的估计位置相关联。

24.根据权利要求23所述的方法,其中所述用户相对于所述参考的所述估计位置包括所述用户相对于所述参考的估计取向的估计取向、所述用户相对于所述参考的估计距离或

两者。

25.根据权利要求23所述的方法,其中所述参考包括所述设备、显示设备、物理参考或虚拟参考中的一者或多者。

26.一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所述一个或多个处理器:

获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据;以及

基于所述电活动数据呈现音频数据,以在所述音频数据的回放期间调整声源在声场中的位置。

27.根据权利要求26所述的非暂态计算机可读介质,其中所述指令在由所述一个或多个处理器执行时使所述一个或多个处理器更新所述用户相对于参考的估计位置,所述估计位置是基于所述用户的头戴式耳机的惯性测量单元(IMU)数据、所述参考的空间数据、用户注视估计或它们的组合来更新的。

28.根据权利要求26所述的非暂态计算机可读介质,其中所述指令在由所述一个或多个处理器执行时使所述一个或多个处理器:

基于所述电活动数据来确定用户注视估计;以及

基于所述用户注视估计来更新所述用户相对于参考的估计位置。

29.一种装置,包括:

用于获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据的部件;和

用于基于所述电活动数据呈现音频数据以在所述音频数据的回放期间调整声源在声场中的位置的部件。

30.根据权利要求29所述的装置,其中用于获得的所述部件或用于呈现的所述部件中的至少一者集成在头戴式耳机、增强现实头戴式耳机、混合现实头戴式耳机、虚拟现实头戴式耳机、交通工具、通信设备、显示设备、电视、游戏控制台、音乐播放器、收音机、数字视频播放器、相机、导航设备、飞行器、家庭自动化系统、语音激活设备、无线扬声器和语音激活设备、便携式电子设备、汽车、计算设备、物联网(IoT)设备、移动设备或它们的任何组合中。

基于用户电信号的音频调整

[0001] 相关申请的交叉引用

[0002] 本申请要求来自2021年9月7日提交的共同拥有的美国非临时专利申请17/467,883号的优先权的权益,该美国非临时专利申请的全部内容以引用方式明确地并入本文。

技术领域

[0003] 本公开整体涉及基于用户电信号调整音频。

[0004] 相关技术

[0005] 技术进步已导致更小且更强大的计算设备。例如,当前存在各种各样的便携式个人计算设备,包括小型、轻量级且易于用户携带的无线电话(诸如移动和智能电话、平板设备和膝上型计算机)。这些设备可以通过无线网络传送语音和数据分组。此外,许多此类设备并入有附加功能性,诸如数字静态相机、数字摄像机、数字记录器和音频文件播放器。此外,此类设备可以处理可执行指令,包括软件应用,诸如网络浏览器应用,其可用于访问互联网。照此,这些设备可以包括显著的计算能力。

[0006] 此类计算设备常常并入有用以回放具有可被感知为来自音频源的方向的声音的空间音频的功能性。音频源的方向通常被映射到回放设备。作为示例,音频可以表示人物的被感知为来自查看回放设备的用户前方的语音。然而,如果用户将回放设备放在桌子上,则语音被感知为来自桌子而不是用户前方。

发明内容

[0007] 根据本公开的一个具体实施,一种设备包括存储器和一个或多个处理器。该存储器被配置为存储指令。该一个或多个处理器被配置为执行该指令以获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据。该一个或多个处理器还被配置为执行该指令以基于该电活动数据呈现音频数据,以在该音频数据的回放期间调整声源在声场中的位置。

[0008] 根据本公开的另一具体实施,一种方法包括在设备处获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据。该方法还包括基于该电活动数据呈现音频数据,以在该音频数据的回放期间调整声源在声场中的位置。

[0009] 根据本公开的另一具体实施,一种非暂态计算机可读介质包括指令,该指令在由一个或多个处理器执行时使该一个或多个处理器获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据。该指令在由该一个或多个处理器执行时使该一个或多个处理器基于该电活动数据呈现音频数据,以在该音频数据的回放期间调整声源在声场中的位置。

[0010] 根据本公开的另一具体实施,一种装置包括用于获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据的部件。该装置还包括用于基于该电活动数据呈现音频数据以在该音频数据的回放期间调整声源在声场中的位置的部件。

[0011] 本公开的其他方面、优点和特征将在审阅包括以下部分的整个申请后变得显而易

见:附图说明、具体实施方式和权利要求书。

附图说明

[0012] 图1是根据本公开的一些示例的可操作以基于用户电信号调整音频的系统的特定例示性方面的框图。

[0013] 图2A和图2B是根据本公开的一些示例的与可由图1的系统执行的相对位置估计相关联的操作的例示性方面的示图。

[0014] 图3是根据本公开的一些示例的图1的系统的组件的操作的例示性方面的示图。

[0015] 图4示出了根据本公开的一些示例的可操作以基于用户电信号调整音频的集成电路的示例。

[0016] 图5是根据本公开的一些示例的可操作以基于用户电信号调整音频的移动设备的示图。

[0017] 图6是根据本公开的一些示例的可操作以基于用户电信号调整音频的头戴式耳机的示图。

[0018] 图7是根据本公开的一些示例的可操作以基于用户电信号调整音频的可穿戴电子设备的示图。

[0019] 图8是根据本公开的一些示例的可操作以基于用户电信号调整音频的声控扬声器系统的示图。

[0020] 图9是根据本公开的一些示例的可操作以基于用户电信号调整音频的头戴式耳机(诸如虚拟现实、混合现实或增强现实头戴式耳机)的示图。

[0021] 图10是根据本公开的一些示例的可操作以基于用户电信号调整音频的交通工具的示图。

[0022] 图11是根据本公开的一些示例的可操作以基于用户电信号调整音频的耳塞的示图。

[0023] 图12是根据本公开的一些示例的可由图1的设备执行的基于用户电信号调整音频的方法的特定具体实施的示图。

[0024] 图13是根据本公开的一些示例的可操作以基于用户电信号调整音频的设备的特定例示性示例的框图。

具体实施方式

[0025] 在空间音频的回放期间,声音可被感知为来自音频源的被映射到回放设备的方向。作为示例,音频可以表示人物的被感知为来自查看回放设备的用户前方的语音。然而,如果用户或回放设备改变定位使得回放设备不再位于用户前方,则语音被感知为来自不在用户前方的回放设备。

[0026] 公开了基于用户电信号调整音频的系统和方法。音频播放器可以将音频数据呈现为在第一回放操作期间包括声源在声场中的多个位置。例如,在第一回放操作期间收听对呈现的音频数据的回放的用户将同时感知来自同一声源的声音,该声音来自声场中的多个位置中的每个位置。为了例示,用户将感知该声音,如同同一声源在多个位置中的每个位置处被复制一样。音频播放器在第一回放操作期间获得与从用户头部内的电源(例如,脑细

胞)生成的电信号相对应的电活动数据。作为示例,电活动数据包括从入耳式传感器接收的脑电图(EEG)数据。音频播放器基于电活动数据将多个位置中的一个位置标识为声源的用户优选位置。音频播放器基于该用户优选位置呈现音频数据,以在第二回放操作期间调整声源的位置。例如,在第二回放操作期间收听对呈现的音频数据的回放的用户将感知来自声源的声音,该声音来自声场中的用户优选位置。因此,音频播放器使得声源的方向能够基于用户偏好进行调整,而不是被映射到回放设备。

[0027] 下面参考附图描述本公开的特定方面。在本说明书中,共用的特征由共用的参考标号来指定。如本文所使用的,各种术语仅用于描述特定具体实施的目的,而并不旨在对具体实施进行限制。例如,单数形式“一”、“一个”和“该”旨在也包括复数形式,除非上下文另外明确指示。此外,本文所描述的一些特征在一些具体实施中是单数,而在其他具体实施中是复数。为了例示,图1描绘了包括一个或多个处理器(图1的“处理器”190)的设备102,这指示在一些具体实施中,设备102包括单个处理器190,并且在其他具体实施中,设备102包括多个处理器190。

[0028] 在一些附图中,使用了特定类型的特征的多个实例。尽管这些特征在物理上和/或逻辑上不同,但对每个特征使用相同的参考标号,并且这些不同实例通过向参考标号添加字母来区分。当作为一组或一种类型的特征在本文中被引用时(例如,当没有引用这些特征中的特定特征时),使用了参考标号,而不使用区分字母。然而,当同一类型的多个特征中的一个特定特征在本文中被引用时,参考标号与区分字母一起使用。例如,参考图1,多个位置被示出并且与参考标号150A和150B相关联。当参考这些位置中的特定位置(诸如位置150A)时,使用了区分字母“A”。然而,当参考这些位置中的任何的任意位置或将这些位置作为组参考时,使用了参考标号150,而没有区分字母。

[0029] 如本文所使用的,术语“包括”可以与“包含”互换使用。此外,术语“其中(wherein)”可以与“其中(where)”互换使用。如本文所使用的,“示例性的”指示示例、具体实施和/或方面,而不应被解释为限制性或指示偏好或优选的具体实施。如本文所使用的,用于修饰元素(诸如,结构、组件、操作等)的序数术语(例如,“第一”、“第二”、“第三”等)本身并不指示该元素相对于另一元素的任何优先级或顺序,而只是将该元素与具有相同名称(但使用序数术语)的另一元素相区分。如本文所使用的,术语“集合”是指一个或多个特定元素,而术语“多个”是指多个(例如,两个或更多个)特定元素。

[0030] 如本文所使用的,“耦合”可以包括“通信地耦合”、“电耦合”或“物理地耦合”,以及还可以(或另选地)包括它们的任何组合。两个设备(或组件)可经由一个或多个其他设备、组件、导线、总线、网络(例如,有线网络、无线网络或它们的组合)等直接或间接耦合(例如,通信地耦合、电耦合或物理地耦合)。作为例示性、非限制性示例,电耦合的两个设备(或组件)可包括在相同的设备中,也可包括在不同的设备中,并且可经由电子器件、一个或多个连接器或电感耦合来连接。在一些具体实施中,通信地耦合(诸如电连通)的两个设备(或组件)可经由一个或多个导线、总线、网络等直接或间接地发送和接收信号(例如,数字信号或模拟信号)。如本文所使用的,“直接耦合”可包括在没有中间组件的情况下耦合(例如,通信地耦合、电耦合或物理地耦合)的两个设备。

[0031] 在本公开中,诸如“确定”、“计算”、“估计”、“移位”、“调整”等术语可以用于描述如何执行一个或多个操作。应当注意,此类术语不应被解读为限制性的,并且可以利用其他技

术来执行类似的操作。此外,如本文所提及的,“生成”、“计算”、“估计”、“使用”、“选择”、“访问”和“确定”可以互换使用。例如,“生成”、“计算”、“估计”或“确定”参数(或信号)可以是指主动生成、估计、计算或确定该参数(或信号),或者可以是指使用、选择或访问(诸如,由另一组件或设备)已经生成的参数(或信号)。

[0032] 参考图1,公开了被配置为基于用户电信号调整音频的系统的特定例示性方面,并且总体上将其指定为100。系统100包括被配置为经由输出接口124耦合到一个或多个扬声器106的设备102。

[0033] 设备102被配置为经由输入接口114耦合到一个或多个传感器104。在特定方面,一个或多个传感器104包括入耳式传感器、电极帽、神经植入物、导电屏、非可穿戴传感器或它们的组合。设备102被配置为耦合到一个或多个空间传感器176,该一个或多个空间传感器被配置为生成指示用户180的空间信息(例如,移动、定位或取向中的至少一者)的空间数据177。在特定方面,一个或多个空间传感器176包括惯性测量单元(IMU)、相机、全球定位系统(GPS)传感器或它们的组合。

[0034] 在一些方面,空间数据177(例如,IMU数据、图像数据或两者)指示用户180的位置的改变、用户180的取向的改变或这两种情况皆有。例如,集成在由用户180穿戴的头戴式耳机中的一个或多个空间传感器176的IMU生成IMU数据,该IMU数据指示与用户180的头部182的移动相对应的头戴式耳机的移动。空间数据177包括IMU数据。

[0035] 在一些方面,空间数据177(例如,GPS数据、图像数据或两者)指示用户180的位置、取向或两者。例如,一个或多个空间传感器176的相机在第一时间捕获用户180的第一图像并且在第二时间捕获用户180的第二图像。第一图像指示头部182在第一时间的第一取向,并且第二图像指示头部182在第二时间的第二取向。空间数据177包括指示在第一时间的第一取向和在第二时间的第二取向并且还指示取向从第一取向到第二取向的改变的第一图像和第二图像。

[0036] 设备102被配置为耦合到一个或多个空间传感器178,该一个或多个空间传感器被配置为生成指示参考188的空间信息(例如,移动、定位或取向中的至少一者)的空间数据179。在特定方面,一个或多个空间传感器178包括IMU、相机、GPS传感器或它们的组合。在一些方面,空间数据179(例如,IMU数据、图像数据或两者)指示参考188的位置的改变、参考188的取向的改变或这两种情况皆有。在一些方面,空间数据179(例如,GPS数据、图像数据或两者)指示参考188的位置、取向或两者。在一些方面,参考188具有固定位置、固定取向或两者。在这些方面,空间数据179可以指示固定位置、固定取向或两者。例如,与由一个或多个空间传感器178生成相比,空间数据179可以基于指示固定位置、固定取向或两者的配置设置、默认数据、用户输入或它们的组合。设备102被配置为使用音频播放器140基于用户电信号调整音频。

[0037] 在一些具体实施中,系统100的一个或多个组件被包括在设备102中,并且系统100的一个或多个组件被包括在被配置为耦合到设备102的第二设备中。在例示性的非限制性示例中,音频播放器140被包括在设备102(例如,电话、平板设备、游戏控制台、计算设备等)中,并且一个或多个空间传感器176、一个或多个扬声器106、一个或多个传感器104或它们的组合被包括在第二设备(例如,用户头戴式设备,诸如用户180的头戴式耳机)中。

[0038] 一个或多个传感器104被配置为生成电活动数据105,该电活动数据对应于来自用

户180的头部182内的一个或多个电源184(例如,脑细胞)的电信号(例如,脑波)。在特定方面,电活动数据105包括眼电图(EOG)数据、EEG数据或两者。输入接口114被配置为从一个或多个传感器104接收电活动数据105。在特定方面,输入接口114包括以太网接口、通用串行总线(USB)接口、Wi-Fi接口、蓝牙®(华盛顿的蓝牙SIG公司的注册商标)接口、串行端口接口、并行端口接口或其他类型的数据接口中的至少一者。

[0039] 设备102包括一个或多个处理器190。在特定方面,输入接口114、输出接口124或两者耦合到一个或多个处理器190。一个或多个处理器190包括音频播放器140。在特定方面,音频播放器140包括被配置为调整音频数据的音频调整器170。在特定方面,音频数据141A对应于由一个或多个麦克风捕获的声音。在特定方面,音频数据141A对应于由游戏引擎、音频应用等生成的音频。在特定方面,音频数据141A对应于所捕获声音与虚拟声音的组合。音频数据141A表示声场142(例如,三维(3D)声场)。在音频数据141A的回放期间,声场142(例如,3D声场)可被重建,使得收听者能够区分收听者与3D声场的一个或多个声源之间的定位和/或距离。

[0040] 在例示性的非限制性示例中,音频数据141A基于或被转换为如下格式中的一种格式:(i)传统的基于声道的音频,其意在通过预先指定的定位处的扩音器播放;(ii)基于对象的音频,其涉及单个音频对象的离散脉冲编码调制(PCM)数据,其中相关联的元数据包含其位置坐标(以及其他信息);或(iii)基于场景的音频,其涉及使用球谐基函数的系数(也称为"球谐系数"或SHC、"高阶高保真度立体声响复制"或HOA、以及"HOA系数")来表示声场。

[0041] 音频播放器140被配置为执行多位置音频生成164。例如,音频调整器170被配置为通过将音频数据141A呈现为具有声源186的与声场142的多个位置150相对应的声音而生成音频数据141B。音频播放器140被配置为在音频回放操作144A期间将音频数据141B输出到一个或多个扬声器106。一个或多个传感器104被配置为在音频回放操作144A期间生成电活动数据105。音频播放器140被配置为通过基于电活动数据105执行优选位置估计(pref.loc.estimation)166来确定用户优选位置167。音频播放器140被配置为基于用户优选位置167执行单位置音频生成(single location audio gen.)168。例如,音频调整器170被配置为通过基于用户优选位置167呈现音频数据141A来生成音频数据141C,以调整声源186在声场142中的位置。

[0042] 在一些具体实施中,设备102对应于或被包括在各种类型的设备中的一种类型的设备中。在例示性示例中,一个或多个处理器190集成在头戴式耳机设备中,该头戴式耳机设备包括一个或多个扬声器106,并且包括或耦合到一个或多个传感器104,诸如参考图6进一步描述的。在其他示例中,一个或多个处理器190集成在以下各项中的至少一项中:(如参考图5所描述的)移动电话或平板计算机设备、(如参考图7所描述的)可穿戴电子设备、(如参考图8所描述的)声控扬声器系统、(如参考图9所描述的)虚拟现实、混合现实或增强现实头戴式耳机、(如参考图11所描述的)一个或多个耳塞。在另一例示性示例中,一个或多个处理器190被集成到耦合到一个或多个扬声器106和一个或多个传感器104的交通工具中,诸如参考图10进一步描述的。

[0043] 在操作期间,用户180激活或启动音频播放器140回放音频数据141A的操作。音频数据141A对应于至少表示声源186的来自声场142中的位置150A的声音(例如,人的语音、来自鸟的声音、来自乐器的音乐等)的空间音频数据。例如,在音频数据141A的回放期间,声场

142可被重建使得来自声源186的声音将被收听者感知为来自声场142(例如,3D空间)中的位置150A。为了例示,来自演员的语音(例如,用于电影音频或有声读物)可被感知为来自收听者前方,并且来自经过的汽车的声音可被感知为在收听者后方从右到左传播。音频数据141A可以表示来自声场142中的一个或多个附加声源的附加声音。

[0044] 音频播放器140被配置为基于用户优选位置167或基于相对位置161(例如,用户180相对于参考188的估计位置)调整声源186在声场142中的位置。例如,相对位置161对应于对用户优选位置167的估计(或替代)。确定用户优选位置167包括向用户180回放具有可从声场142的多个位置150感知的声源186的声音的音频。另一方面,可以在不知晓用户180的情况下在后台估计相对位置161。在一些具体实施中,当相对位置161的位置置信度水平181大于或等于置信度阈值(conf.threshold)163时,音频播放器140基于相对位置161调整声源186的位置。然而,当位置置信度水平181小于置信度阈值163时,音频播放器140通过回放具有声源186的来自多个位置150的声音的音频来确定用户优选位置167,并且基于用户优选位置167调整声源186的位置。因此,当位置置信度水平181未能满足置信度阈值163时,从多个位置150选择性地回放声源186的声音以确定用户优选位置167。

[0045] 在特定方面,音频播放器140基于空间数据177、空间数据179、电活动数据105或它们的组合执行相对位置估计160,以确定相对位置161和位置置信度水平181,如参考图2A至图2B进一步描述的。例如,相对位置161对应于用户180相对于参考188的估计位置,并且位置置信度水平181指示与相对位置161相关联的估计置信度。在一些具体实施中,相对位置161对应于用户180相对于参考188的估计定位(例如,位置、取向或两者)的估计定位(例如,位置、取向或两者)。

[0046] 在特定方面,参考188包括设备102、显示设备、回放设备、一个或多个扬声器106、物理参考、虚拟参考、固定位置参考或移动参考中的一者或多者。例如,参考188可包括在虚拟场景中具有固定位置的虚拟参考(例如,建筑物)。作为另一示例,参考188可包括能够在虚拟场景中改变位置的虚拟参考(例如,移动虚拟人物)。在一些示例中,参考188可包括在物理空间中具有固定位置(例如,附着到墙壁)的物理参考(例如,广告显示器)。在其他示例中,参考188可包括能够在物理空间中改变位置的物理参考(例如,移动设备)。

[0047] 作为例示性示例,参考188被示为与设备102分离。在其他示例中,参考188可被集成在设备102中。在一些具体实施中,参考188是指参考点(例如,特定位置)。在其他具体实施中,参考188可以具有多维(例如,二维或三维)形状,诸如正方形、立方体、矩形、平面、棱柱、三角形、棱锥、圆形、球形、椭圆形、卵形等。

[0048] 在特定具体实施中,音频播放器140将相对位置161初始化为对应于在距用户180预定距离(例如,12英寸)处取向在用户180前方(例如,面向用户)的参考188(例如,移动电话屏幕)。在特定方面,预定距离基于配置设置、默认值、用户输入或它们的组合。在特定方面,音频播放器140将位置置信度水平181初始化为小于置信度阈值163。在特定方面,置信度阈值163基于配置设置、默认值、用户输入或它们的组合。在一些方面,音频播放器140基于参考188的移动、用户180的移动或两者来更新相对位置161和位置置信度水平181,如参考图2A至图2B进一步描述的。

[0049] 音频播放器140执行比较162以确定是将相对位置161用于单位置音频生成168还是确定用户优选位置167。例如,音频播放器140将位置置信度水平181与置信度阈值163进

行比较。响应于确定位置置信度水平181大于或等于置信度阈值163,音频播放器140基于相对位置161进行单位置音频生成168,如参考图2A至图2B进一步描述的。例如,音频调整器170基于音频数据141A和相对位置161生成音频数据141C。为了例示,生成音频数据141C包括基于相对位置161呈现音频数据141A。另选地,响应于确定位置置信度水平181小于置信度阈值163,音频播放器140执行多位置音频生成164以确定用户优选位置167。

[0050] 在一些方面,比较162包括相对位置161与相对位置161的先前确定的比较。例如,响应于确定相对位置161与相对位置161的先前确定之间的差小于位置改变阈值并且位置置信度水平181大于第二置信度阈值,音频播放器140返回到相对位置估计160而不调整声源186的位置。

[0051] 在多位置音频生成164期间,音频调整器170基于音频数据141A生成音频数据141B。例如,生成音频数据141B包括将音频数据141A呈现为具有声源186的多个位置。为了例示,音频数据141B表示来自声场142的位置150A的声源186的声音、来自声场142的位置150B的声源186的声音、来自声场142的一个或多个附加位置的声源186的声音或它们的组合。

[0052] 在特定方面,音频数据141A表示来自位置150A的声源186的声音,并且生成音频数据141B包括添加来自位置150B、一个或多个附加位置或它们的组合的声源186的声音。在替代方面,音频数据141A不包括声源186的任何声音,并且生成音频数据141B包括添加来自多个位置150中的每一者的声源186的声音。

[0053] 音频播放器140基于音频数据141B执行优选位置估计166。例如,音频播放器140经由一个或多个扬声器106启动音频数据141B的音频回放操作144A。例如,音频播放器140经由输出接口124将音频数据141B提供给一个或多个扬声器106。在特定方面,输出接口124包括以太网接口、通用串行总线(USB)接口、Wi-Fi接口、蓝牙®(华盛顿的蓝牙SIG公司的注册商标)接口、串行端口接口、并行端口接口或其他类型的数据接口中的至少一者。

[0054] 音频数据141B包括在音频回放操作144A期间声源186在声场142中的多个位置150。在音频回放操作144A期间,音频播放器140经由输入接口114从一个或多个传感器104获得电活动数据105。电活动数据105对应于在音频回放操作144A期间来自用户180的头部182内的一个或多个电源184的电信号。例如,电信号由一个或多个电源184(例如,脑细胞)生成,同时音频数据141B被播放给用户180。在特定方面,音频播放器140在音频回放操作144A期间输出指示正在执行音频配置的警示(例如,视觉警示)。

[0055] 音频播放器140基于电活动数据105确定声源186的用户优选位置167。例如,音频播放器140使用优选位置模型(pref.location model)174来处理电活动数据105,并且优选位置模型174的输出(例如,人工神经网络、机器学习模型或两者)指示位置150B对应于声源186的用户优选位置167。

[0056] 在一些具体实施中,音频播放器140基于执行优选源估计来确定用户优选位置167,如相对于图3进一步描述的。例如,音频调整器170通过在音频回放操作144A期间将音频数据141A呈现为具有语音源的第一位置和非语音源(例如,汽车)的第二位置来生成音频数据141B。音频调整器170基于在音频回放操作144A期间获得的电活动数据105来确定用户优选位置167。例如,响应于确定电活动数据105指示单个源正被用户180跟踪,音频播放器140确定语音源的第一位置对应于用户优选位置167。另选地,响应于确定电活动数据105指

示多个源正被用户180跟踪,音频播放器140确定非语音源的第二位置对应于用户优选位置167。例如,即使当用户180正在收听非语音声音(例如,汽车驾驶)时,人脑也跟踪语音,因此指示多个源正被用户180跟踪的电活动数据105对应于用户180正在收听非语音声音(例如,汽车)。

[0057] 音频播放器140基于用户优选位置167执行单位置音频生成168。例如,音频调整器170基于用户优选位置167和音频数据141A生成音频数据141C。为了例示,生成音频数据141C包括将音频数据141A呈现为具有声源186在声场142中的用户优选位置167(例如,位置150B)。作为示例,音频播放器140将声源186的位置从声场142中的位置150A(如音频数据141A中所表示)调整到声场142中的位置150B(如音频数据141C中所表示)。

[0058] 音频播放器140经由一个或多个扬声器106启动音频数据141C的音频回放操作144B。例如,音频播放器140经由输出接口124将音频数据141C提供给一个或多个扬声器106。音频数据141C包括在音频回放操作144B期间声源186在声场142中的用户优选位置167(例如,位置150B)。例如,在音频回放操作144B期间,将声源186的位置调整到用户优选位置167(例如,位置150B)。为了例示,在音频回放操作144B期间,声源186被感知为来自声场142的单个位置。在特定方面,在音频回放操作144B期间,将声源186的单个位置固定在用户优选位置167(例如,位置150B)处。在替代方面,声源186的单个位置在用户优选位置167(例如,位置150B)处被初始化并且在音频回放操作144B期间改变。例如,声源186对应于飞鸟,并且来自鸟的声音的位置在声场142中移动。

[0059] 因此,系统100使得能够呈现音频,该音频的声音可被感知为来自声源186的被映射到用户优选位置167或用户优选位置167的估计(例如,相对位置161)的方向。作为示例,如果用户180将回放设备(例如,参考188)放在桌子上,则声源186的位置可被调整为继续被感知为来自用户180前方(例如,用户优选位置167)。

[0060] 参考图2A,示出了与相对位置估计160相关联的操作的例示性方面的示图200。相对位置估计160可由图1的音频播放器140执行。

[0061] 音频播放器140执行相对位置估计160以确定相对位置161。相对位置161包括用户180相对于参考188的距离212、相对取向263或两者。例如,距离212指示用户位置220(例如,用户180的估计位置)与参考位置230(例如,参考188的估计位置)之间的距离。相对取向263指示相对于参考取向232(例如,参考188的估计取向)的用户取向222(例如,用户180的估计取向)。

[0062] 音频播放器140基于位置数据270执行相对位置估计160。例如,音频播放器140将相对位置161初始化为对应于在距用户180预定距离(例如,12英寸)处取向在用户180前方(例如,面向用户)的参考188(例如,移动电话屏幕)。音频播放器140基于对位置数据270的更新来更新相对位置161。

[0063] 示例202A-202C示出了三维空间中的水平平面的俯视图。在特定方面,水平平面由三维空间中的X轴和Y轴限定,并且竖直平面由三维空间中的X轴和Z轴限定。示例202A对应于音频播放器140将相对取向263和距离212分别初始化为相对取向263A和距离212A。示例202B-202C对应于音频播放器140基于对位置数据270的更新来更新相对取向263和距离212。

[0064] 在示例202A中,音频播放器140将参考188的参考位置230初始化为参考位置230A,

将参考188的参考取向232初始化为参考取向232A,将用户180的用户位置220初始化为用户位置220A,并且将用户180的用户取向222初始化为用户取向222A。

[0065] 在一些具体实施中,空间数据179(例如,GPS数据、配置数据、图像数据等)指示在三维空间中在具有参考取向232A的参考位置230A处检测到参考188,并且音频播放器140将参考位置230和参考取向232分别初始化为参考位置230A和参考取向232A。在替代具体实施中,音频播放器140将参考位置230初始化为与三维空间的原点(例如,沿X轴为0英寸、沿Y轴为0英寸并且沿Z轴为0英寸)相对应的参考位置230A,并且参考取向232A对应于参考188面向三维空间中的预定方向(例如,在水平平面中为0度并且在竖直平面中为0度)。

[0066] 在一些具体实施中,空间数据177(例如,GPS数据、图像数据等)指示在三维空间中在具有用户取向222A的用户位置220A处检测到用户180,并且音频播放器140将用户位置220和用户取向222分别初始化为用户位置220A和用户取向222A。在替代具体实施中,音频播放器140将用户位置220初始化为与三维空间中距参考位置230A的预定点(例如,在预定距离和预定方向处)相对应的用户位置220A。例如,用户位置220A对应于在三维空间中距原点在相对方向265A(例如,在水平平面中为0度并且在竖直平面中为0度)上的距离212A(例如,预定距离,诸如12英寸)的点(例如,沿X轴为12英寸,沿Y轴为0英寸并且沿Z轴为0英寸)。音频播放器140将用户取向222A初始化为三维空间中的预定方向(例如,在水平平面(例如,XY平面)中为180度并且在竖直平面(例如,XZ平面)中为0度),以对应于面向参考188的用户180。因此,音频播放器140在时间T0处将距离212初始化为距离212A并且将相对取向263初始化为相对取向263A。相对取向263A(例如,在水平平面中为0度并且在竖直平面中为0度)基于参考取向232A(例如,在水平平面中为0度并且在竖直平面中为0度)、用户取向222A(例如,在水平平面中为180度并且在竖直平面中为0度)以及用户位置220A到参考位置230A的相对方向265A(例如,在水平平面中为0度并且在竖直平面中为0度)。

[0067] 音频播放器140基于位置数据270更新相对位置161。例如,位置数据270指示用户180的位置的改变、用户180的取向的改变、参考188的位置的改变、参考188的取向的改变或它们的组合,并且音频播放器140基于由位置数据270指示的改变来更新相对位置161。在特定方面,位置数据270包括用户180的空间数据177、参考188的空间数据179、用户注视估计275或它们的组合。

[0068] 电活动数据105可以指示用户注视的方向(例如,用户注视估计275)。在特定方面,用户注视的方向相对于头部182的取向。例如,如果用户180正在查看特定注视目标并且继续查看同一注视目标,同时将头部182的取向改变特定量(例如,10度),则用户注视的方向改变特定量(例如,10度)。在特定方面,音频播放器140基于用户注视估计275来估计用户180的取向的改变。例如,音频播放器140基于空间数据177来确定用户180的头部182的取向的改变。在一些具体实施中,头部182的取向的改变对应于对用户取向222的改变的宽泛估计,并且音频播放器140基于用户注视估计275来细化对用户取向222的改变的估计。例如,如果用户180移动头部182但保持注视朝向同一位置,则用户取向222可能不会改变。

[0069] 在特定方面,音频播放器140使用注视估计模型274(例如,人工神经网络、机器学习模型或两者)处理电活动数据105(例如,EOG数据)以确定用户注视估计275。基于空间数据177确定用户取向222(例如,头部取向)并且基于用户注视估计275(例如,用户注视方向)更新用户取向222作为例示性的非限制性示例而提供。在一些示例中,音频播放器140可以

处理空间数据177和电活动数据105(指示用户注视的方向)以确定用户取向222。

[0070] 根据一些研究,EOG数据(例如,电活动数据105)被建模以检测扫视(例如,眼睛在凝视状态之间的快速移动),并且指示扫视的EOG信号的变化反映注视移位的方向。例如,EOG信号的变化的增加指示向右的注视移位,并且EOG信号的变化的减小指示向左的注视移位。EOG信号的振幅指示注视移位的角度。例如,振幅的较高绝对值指示较大的注视移位。在特定方面,响应于检测到凝视而确定用户注视估计275。例如,基于确定用户注视的方向在至少阈值持续时间内未改变而检测到凝视。

[0071] 在一些具体实施中,空间数据177(例如,GPS数据、图像数据或两者)可以直接指示用户位置220、用户取向222或两者,而不是音频播放器140基于改变来估计用户位置220、用户取向222或两者。在一些具体实施中,空间数据179(例如,GPS数据、图像数据或两者)可以直接指示参考位置230、参考取向232或两者,而不是音频播放器140基于改变来估计参考位置230、参考取向232或两者。

[0072] 在一些具体实施中,参考188具有固定位置(例如,参考位置230A)。在一些方面,空间数据179指示参考取向232、固定位置(例如,参考位置230A)或两者的改变(如果有的话)。在替代方面,参考188具有固定取向(例如,参考取向232A)。在一些示例中,空间数据179指示固定位置(例如,参考位置230A)、固定取向(例如,参考取向232A)或两者。在一些示例中,空间数据179指示对参考位置230没有改变、对参考取向232没有改变或这两种情况皆有。在又其他示例中,位置数据270可以不包括空间数据179。例如,音频播放器140基于空间数据177、用户注视估计275或两者来估计(例如,更新)用户位置220和用户取向222,并且基于参考188的固定位置和固定取向以及用户180的估计位置和估计取向来执行相对位置估计160。

[0073] 示例202B和202C例示与不同用户取向222、不同参考取向232和不同参考位置230相对应的相同相对取向263的示例。在一些示例中,相同相对取向263可对应于不同用户取向222、不同参考取向232、不同用户位置220、不同参考位置230或它们的组合。在示例202B中,音频播放器140在时间T0之后的时间T1处获得位置数据270。音频播放器140基于空间数据177确定用户180具有用户取向222B(例如,在水平平面中为135度并且在竖直平面中为0度)并且处于用户位置220B处。音频播放器140基于空间数据179确定参考188处于参考位置230B处并且具有参考取向232A(例如,在水平平面中为0度并且在竖直平面中为0度)。音频播放器140基于用户位置220B与参考位置230B之间的差来确定距离212B。

[0074] 音频播放器140确定用户位置220B(例如,3D空间中的第一点)具有与参考位置230B(例如,3D空间中的第二点)的相对方向265A(例如,在水平平面中为0度并且在竖直平面中为0度)。例如,相对方向265在示例202B中与在示例202A中相同。

[0075] 相对方向265(例如,3D空间中第一点相对于第二点的方向)基于用户位置220(例如,第一点)相对于参考位置230(例如,第二点)的方向,并且与用户取向222和参考取向232无关。相比而言,除了用户位置220到参考位置230的相对方向265之外,相对取向263还基于用户取向222和参考取向232。例如,相对取向263指示与用户180相对应的至少第一平面(例如,包括第一点)相对于与参考188相对应的至少第二平面(例如,包括第二点)的取向。在例示性的非限制性示例中,第一平面对应于用户180的头部182的竖直横截面并且第二平面对应于参考188(例如,移动设备)的显示屏。

[0076] 音频播放器140基于用户取向222B(例如,在水平平面中为135度并且在竖直平面中为0度)、参考取向232A(例如,在水平平面中为0度并且在竖直平面中为0度)以及用户位置220B到参考位置230B的相对方向265B(例如,在水平平面中为0度并且在竖直平面中为0度)来确定相对取向263B(例如,在水平平面中为45度并且在竖直平面中为0度)。在一些示例中,对于具有相同用户取向222和相同参考取向232的不同相对方向265,相对取向263可以是不同的,如参考图2B进一步描述的。

[0077] 在示例202C中,音频播放器140在时间T0之后的时间T2处获得位置数据270。音频播放器140基于空间数据177确定用户180具有用户取向222A(例如,在水平平面中为180度并且在竖直平面中为0度)并且处于用户位置220B处。音频播放器140基于空间数据179确定参考188处于参考位置230C处并且具有参考取向232C(例如,在水平平面中为45度并且在竖直平面中为0度)。音频播放器140基于用户位置220B与参考位置230C之间的差来确定距离212B。

[0078] 音频播放器140基于用户位置220B(例如,3D空间中的第一点)与参考位置230C(例如,3D空间中的第三点)的比较来确定相对方向265C。例如,用户位置220B具有与参考位置230C的相对方向265C(例如,在水平平面中为45度并且在竖直平面中为0度)。音频播放器140基于用户取向222A(例如,在水平平面中为180度并且在竖直平面中为0度)、参考取向232C(例如,在水平平面中为45度并且在竖直平面中为0度)以及用户位置220B到参考位置230C的相对方向265C(例如,在水平平面中为45度并且在竖直平面中为0度)来确定相对取向263B(例如,在水平平面中为45度并且在竖直平面中为0度)。对于不同参考位置230、不同参考取向232、不同用户取向222、不同相对方向265、相同距离212和相同用户位置220,相对取向263在示例202C中与在示例202B中相同(例如,相对取向263B)。例如,与示例202B相比,在示例202C中,与用户180相对应的至少第一平面(例如,头部182的竖直横截面)相对于与参考188相对应的至少第二平面(例如,显示屏)具有相同取向。

[0079] 参考图2B,示出了与相对位置估计160相关联的操作的例示性方面的示图250。相对位置估计160可由图1的音频播放器140执行。示例202D示出了由于用户位置220到参考点位置230的不同相对方向265而对应于不同相对取向263的相同用户取向222和相同参考取向232的示例。

[0080] 示例202D对应于音频播放器140基于对位置数据270的更新来更新相对取向263和距离212。在示例202D中,音频播放器140在时间T0之后的时间T3处获得位置数据270。音频播放器140基于空间数据177确定用户180具有用户取向222A(例如,在水平平面中为180度并且在竖直平面中为0度)并且处于用户位置220D处。音频播放器140基于空间数据179确定参考188具有参考取向232A(例如,在水平平面中为0度并且在竖直平面中为0度)并且处于参考位置230D处。音频播放器140基于用户位置220D与参考位置230D之间的差来确定距离212A。

[0081] 音频播放器140基于用户位置220D(例如,3D空间中的第三点)与参考位置230D(例如,3D空间中的第四点)的比较来确定相对方向265D。例如,用户位置220D具有与参考位置230D的相对方向265D(例如,在水平平面中为39度并且在竖直平面中为0度)。音频播放器140基于用户取向222A(例如,在水平平面中为180度并且在竖直平面中为0度)、参考取向232A(例如,在水平平面中为0度并且在竖直平面中为0度)以及用户位置220D到参考点位置

230D的相对方向265D(例如,在水平平面中为39度并且在竖直平面中为0度)来确定相对取向263D(例如,在水平平面中为39度并且在竖直平面中为0度)。

[0082] 对于相同用户取向222(例如,用户取向222A)、相同参考取向232(例如,参考取向232A)、不同相对方向265、不同用户位置220和不同参考位置230,示例202D的相对取向263D不同于示例202A的相对取向263A。例如,与示例202A相比,在示例202D中,与用户180相对应的第一平面(例如,头部182的竖直横截面)具有相同用户取向222,并且与参考188相对应的第二平面(例如,显示屏)具有相同参考取向232。与示例202A相比,在示例202D中,第一平面具有与第二平面不同的相对取向263,因为第三点相对于第四点的相对方向265D不同于第一点相对于第二点的相对方向265A。

[0083] 参考图3,示出了可操作以基于用户电信号调整音频的系统的示图,并且总体上将其指定为300。在特定方面,图1的系统100包括系统300的一个或多个组件。

[0084] 在图3中,示出了多位置音频生成164和优选位置估计166的具体实施的示例。例如,多位置音频生成164包括多源音频生成364。为了例示,音频播放器140通过将音频数据141A呈现为在声场142中具有语音源386的位置350和非语音源388的位置352来生成音频数据141B。

[0085] 音频播放器140经由一个或多个扬声器106启动音频数据141B的音频回放操作144A。一个或多个传感器104在音频回放操作144A期间生成电活动数据105。例如,电活动数据105基于在音频回放操作144A期间来自一个或多个电源184的电信号。

[0086] 在一些示例中,优选位置估计166包括优选源估计(pref.source estimation)366。例如,音频播放器140基于电活动数据105将语音源386或非语音源388中的一者标识为用户优选源367。音频播放器140使用优选源模型374(例如,人工神经网络、机器学习模型或两者)处理电活动数据105以生成所跟踪声源的计数。

[0087] “所跟踪声源”对应于用户180的听觉系统在音频回放操作144A期间随着声源在声场142中移动而聚焦(例如,关注)的声源。根据一些研究,线性映射(例如,时间响应函数(TRF))可以在EEG数据与关注以及未关注的声源轨迹之间导出。即使在存在分散注意力的刺激的情况下,也可以从EEG的Δ相位和α功率可靠地重建关注的声源的轨迹(例如,路径)。对未关注的非语音声源(例如,噪声)的跟踪低于检测水平,并且未关注的语音(例如,通过EEG的Δ相位)被弱跟踪。

[0088] 如果用户180在音频回放操作144A期间关注语音源386,则电活动数据105(例如,EEG数据的Δ相位和α功率)跟踪语音源386,而对非语音源388的跟踪低于检测水平。如果用户180在音频回放操作144A期间关注非语音源388,则电活动数据105跟踪非语音源388并且弱跟踪语音源386。例如,EEG的Δ相位和α功率跟踪非语音源388,并且EEG的Δ相位弱跟踪语音源386。

[0089] 在一些具体实施中,优选源模型374被训练以确定由电活动数据指示的所跟踪声源的计数。例如,训练电活动数据通过以下方式生成:使用一个或多个扬声器106回放与在声场中移动(例如,改变位置)的声源(例如,一个或多个语音源、一个或多个非语音源或它们的组合)相对应的音频、请求用户180在回放期间聚焦(例如,关注)特定声源、从一个或多个传感器104收集训练电活动数据、以及用关注的声源的计数来标记训练电活动数据。优选源模型374用于处理训练电活动数据以生成所跟踪声源的估计计数,损失度量基于估计计

数与标记计数的比较,并且优选源模型374的配置设置(例如,权重、偏差或它们的组合)基于损失度量进行调整。

[0090] 响应于确定所跟踪声源的计数具有指示单个声源被跟踪的第一值(例如,1),音频播放器140确定语音源386对应于用户优选源367并且位置350对应于用户优选位置167。另选地,响应于确定所跟踪声源的计数具有指示多个声源(例如,语音源386和非语音源388)被用户180跟踪的第二值(例如,大于1),音频播放器140确定非语音源388对应于用户优选源367并且位置352对应于用户优选位置167。为了例示,人脑即使在收听者(例如,用户180)正关注非语音音频时也在某种程度上跟踪语音,并且在收听者正关注语音音频时不跟踪非语音。音频播放器140基于用户优选位置167(例如,位置350或位置352中的一者)执行单位置音频生成168,如参考图1所描述的。

[0091] 图4描绘了设备102作为包括一个或多个处理器190的集成电路402的具体实施400。集成电路402还包括输入接口114(诸如一个或多个总线接口),以使得能够接收电活动数据105以进行处理。集成电路402还包括输出接口124(诸如总线接口),以使得能够发送音频数据141。集成电路402使得能够基于用户电信号调整音频。在一些示例中,集成电路402对应于系统中的耦合到一个或多个传感器104、一个或多个扬声器106或它们的组合的组件,诸如如图5中所描绘的移动电话或平板设备、如图6中所描绘的头戴式耳机、如图7中所描绘的可穿戴电子设备、如图8中所描绘的声控扬声器系统、如图9中所描绘的虚拟现实、混合现实或增强现实头戴式耳机、如图10中所描绘的交通工具或如图11中所描绘的一个或多个耳塞。

[0092] 图5描绘了具体实施500,其中设备102包括移动设备502,诸如电话或平板设备,作为例示性的非限制性示例。移动设备502包括一个或多个扬声器106、显示屏504或它们的组合。一个或多个处理器190的组件(包括音频播放器140)被集成在移动设备502中并使用虚线示出以指示通常对移动设备502的用户不可见的内部组件。

[0093] 移动设备502耦合到一个或多个传感器104。在一些具体实施中,移动设备502对应于参考188并且包括一个或多个运动传感器178。使用虚线示出一个或多个运动传感器178以指示通常对移动设备502的用户不可见的内部组件。在一些具体实施中,一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176或它们的组合被集成在用户头戴式设备(例如,头戴式耳机或耳塞)中,并且音频播放器140被集成在移动设备502中。在一些具体实施中,一个或多个空间传感器176(例如,相机)被集成在移动设备502中。

[0094] 在特定示例中,音频播放器140操作以基于用户电信号调整音频,该用户电信号也可被处理以在移动设备502处执行一个或多个操作,诸如启动图形用户界面或以其他方式(例如,经由集成的“智能助理”应用)在显示屏504处显示与调整音频相关联的信息或与音频中所检测到的语音相关联的信息。

[0095] 图6描绘了具体实施600,其中设备102包括头戴式耳机设备602。头戴式耳机设备602包括或耦合到一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176、一个或多个空间传感器178或它们的组合。一个或多个处理器190的组件(包括音频播放器140)被集成在头戴式耳机设备602中。在特定示例中,音频播放器140操作以基于用户电信号调整音频,该用户电信号可使得头戴式耳机设备602在头戴式耳机设备602处执行一个或多个操作、将调整的音频数据发射到第二设备(未示出)以进行进一步处理、或它们的

组合。

[0096] 在一些具体实施中,头戴式耳机设备602包括一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176或它们的组合,并且耦合到包括音频播放器140的第二设备。一个或多个空间传感器178可包括在头戴式耳机设备602、第二设备或两者中。在一些方面,第二设备包括交通工具、移动设备、电话、游戏控制台、通信设备、可穿戴电子设备、声控扬声器系统、无人驾驶交通工具或它们的组合。

[0097] 图7描绘了具体实施700,其中设备102包括被示为"智能手表"的可穿戴电子设备702。音频播放器140和一个或多个扬声器106被集成到可穿戴电子设备702中或耦合到该可穿戴电子设备。

[0098] 可穿戴电子设备702耦合到一个或多个传感器104。在一些具体实施中,可穿戴电子设备702对应于参考188并且包括一个或多个运动传感器178。使用虚线示出一个或多个运动传感器178以指示通常对可穿戴电子设备702的用户不可见的内部组件。在一些具体实施中,一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176或它们的组合被集成在用户头戴式设备(例如,头戴式耳机或耳塞)中,并且音频播放器140被集成在可穿戴电子设备702中。在一些具体实施中,一个或多个空间传感器176(例如,相机)被集成在可穿戴电子设备702中。

[0099] 在特定示例中,音频播放器140操作以基于用户电信号调整音频,该用户电信号随后被处理以在可穿戴电子设备702处执行一个或多个操作,诸如启动图形用户界面或以其他方式在可穿戴电子设备702的显示屏704处显示与调整音频相关联的信息或与音频中所检测到的语音相关联的信息。为了例示,可穿戴电子设备702可包括显示屏,该显示屏被配置为在由可穿戴电子设备702进行的音频回放操作144A期间显示通知。在特定示例中,可穿戴电子设备702包括触觉设备,该触觉设备在音频回放操作144A期间提供触觉通知(例如,振动)。例如,触觉通知可使用户查看可穿戴电子设备702以看见指示正在进行的音频配置的所显示的通知。因此,可穿戴电子设备702可以向具有听力障碍的用户或佩戴头戴式耳机的用户警示正在执行音频配置。

[0100] 图8是具体实施800,其中设备102包括无线扬声器和语音激活设备802。无线扬声器和语音激活设备802可具有无线网络连接性,并且被配置为执行辅助操作。包括音频播放器140、一个或多个扬声器106或它们的组合的一个或多个处理器190被包括在无线扬声器和语音激活设备802中。

[0101] 无线扬声器和语音激活设备802耦合到一个或多个传感器104。在一些具体实施中,无线扬声器和语音激活设备802对应于参考188并且包括一个或多个运动传感器178。使用虚线示出一个或多个运动传感器178以指示通常对无线扬声器和语音激活设备802的用户不可见的内部组件。在一些具体实施中,一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176或它们的组合被集成在用户头戴式设备(例如,头戴式耳机或耳塞)中,并且音频播放器140被集成在无线扬声器和语音激活设备802中。在一些具体实施中,一个或多个空间传感器176(例如,相机)被集成在无线扬声器和语音激活设备802中。

[0102] 在操作期间,响应于经由音频播放器140的操作接收到被标识为用户语音的口头命令,无线扬声器和语音激活设备802可以执行辅助操作(例如,集成辅助应用)。该辅助操作可包括调节温度、播放音乐、打开灯等。例如,该辅助操作响应于在关键字或关键短语(例

如,“你好,助手”)后接收命令来执行。

[0103] 图9描绘了具体实施900,其中设备102包括与虚拟现实、混合现实或增强现实头戴式耳机902相对应的便携式电子设备。音频播放器140、一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176、一个或多个空间传感器178或它们的组合被集成到头戴式耳机902中。

[0104] 可执行基于用户电信号的音频调整,并且可经由头戴式耳机902的一个或多个扬声器106输出调整的音频信号。将视觉接口设备定位在用户的眼睛前方,以使得能够在穿戴头戴式耳机902时向用户显示增强现实、混合现实或虚拟现实图像或场景。在一些具体实施中,参考188对应于可由视觉接口设备显示的虚拟参考。在特定示例中,视觉接口设备被配置为显示指示正在进行音频配置的通知。

[0105] 图10描绘了具体实施1000,其中设备102对应于交通工具1002(被示为有人驾驶或无人驾驶空中设备(例如,包裹递送无人机))或集成在该交通工具内。音频播放器140、一个或多个扬声器106或它们的组合被集成到交通工具1002中。

[0106] 交通工具1002耦合到一个或多个传感器104。在一些具体实施中,交通工具1002包括一个或多个运动传感器178。在一些具体实施中,交通工具1002对应于参考188。使用虚线示出一个或多个运动传感器178以指示通常对交通工具1002的用户不可见的内部组件。在一些具体实施中,一个或多个传感器104、一个或多个扬声器106、一个或多个空间传感器176或它们的组合被集成在用户头戴式设备(例如,头戴式耳机或耳塞)中,并且音频播放器140被集成在交通工具1002中。在一些具体实施中,一个或多个空间传感器176(例如,相机)被集成在交通工具1002中。可执行基于用户电信号的音频调整,并且可经由交通工具1002的一个或多个扬声器106输出调整的音频信号。

[0107] 图11是可操作以基于用户电信号执行音频调整的耳塞1100(例如,图1的设备102的另一特定示例)的示图。在图11中,第一耳塞1102包括一个或多个空间传感器176中的至少一个空间传感器,并且第二耳塞1104包括一个或多个空间传感器176中的至少一个空间传感器。第一耳塞1102和第二耳塞1004中的每一者还包括一个或多个扬声器106中的至少一个扬声器。耳塞1200中的一者或两者还可包括音频播放器140、一个或多个空间传感器178或它们的组合。

[0108] 参考图12,示出了基于用户电信号调整音频的方法1200的特定具体实施。在特定方面,方法1200的一个或多个操作由音频播放器140、音频调整器170、一个或多个处理器190、设备102、图1的系统100或它们的组合中的至少一者执行。

[0109] 在1202处,方法1200包括获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据。例如,图1的音频播放器140获得与来自用户180的头部182内的一个或多个电源184的电信号相对应的电活动数据105。

[0110] 在1204处,方法1200还包括基于电活动数据呈现音频数据,以在该音频数据的回放期间调整声源在声场中的位置。例如,音频播放器140基于电活动数据105呈现音频数据141,以在音频数据141的回放期间调整声源186在声场142中的位置。为了例示,音频播放器140通过呈现音频数据141A来生成音频数据141C,以将声源186的位置从声场142中的位置150A调整到位置150B。

[0111] 因此,方法1200使得能够呈现音频,该音频的声音可被感知为来自与用户优选位

置167或用户优选位置167的估计(例如,相对位置161)相对应的声源186(例如,位置150B)的方向。作为示例,如果用户180将回放设备(例如,参考188)放在桌子上,则声源186的位置可被调整为继续被感知为来自用户180前方(例如,用户优选位置167)。

[0112] 图12的方法1200可由现场可编程门阵列(FPGA)设备、专用集成电路(ASIC)、处理单元(诸如中央处理单元(CPU))、数字信号处理器(DSP)、控制器、另一硬件设备、固件设备、或它们的任何组合实现。作为示例,图12的方法1200可由执行指令的处理器执行,诸如参考图13所描述的。

[0113] 参考图13,描绘了设备的特定例示性具体实施的框图,并且总体上将其指定为1300。在各种具体实施中,设备1300可以具有比图13所示的更多或更少的组件。在例示性具体实施中,设备1300可以对应于设备102。在例示性具体实施中,设备1300可以执行参考图1至图12所描述的一个或多个操作。

[0114] 在特定具体实施中,设备1300包括处理器1306(例如,CPU)。设备1300可包括一个或多个附加处理器1310(例如,一个或多个DSP)。在特定方面,图1的一个或多个处理器190对应于处理器1306、处理器1310或它们的组合。处理器1310可包括语音和音乐编码器-解码器(编解码器)1308,其包括语音译码器(“声码器”)编码器1336、声码器解码器1338、音频播放器140或它们的组合。

[0115] 设备1300可包括存储器1386和编解码器1334。存储器1386可包括指令1356,该指令可由一个或多个附加处理器1310(或处理器1306)执行以实现参考音频播放器140所描述的功能性。设备1300可包括经由收发器1350耦合到天线1352的调制解调器1348。设备1300可包括或耦合到一个或多个空间传感器176、一个或多个空间传感器178、一个或多个传感器104或它们的组合。

[0116] 设备1300可包括耦合到显示控制器1326的显示器1328。一个或多个扬声器106和一个或多个麦克风1390可耦合到编解码器1334。编解码器1334可包括数模转换器(DAC)1302、模数转换器(ADC)1304或两者。在特定具体实施中,编解码器1334可以从一个或多个麦克风1390接收模拟信号,使用模数转换器1304将模拟信号转换为数字信号(例如,音频数据141A),并且将数字信号提供给语音和音乐编解码器1308。语音和音乐编解码器1308可以处理数字信号,并且数字信号可以进一步由音频播放器140进行处理。在特定具体实施中,语音和音乐编解码器1308可将数字信号(例如,音频数据141C)提供给编解码器1334。编解码器1334可以使用数模转换器1302将数字信号转换为模拟信号,并且可将模拟信号提供给一个或多个扬声器106。例如,响应于确定用户180正在远离用户180或在用户后方的位置150A处跟踪声源186,音频播放器140可以生成音频数据141C以将声源186的位置调整到更靠近用户180或在用户前方的位置150B。对应于音频数据141C的模拟信号可经由集成在用户180的头戴式耳机或耳塞中的一个或多个扬声器106播放。

[0117] 在特定具体实施中,设备1300可被包括在系统级封装件或片上系统设备1322中。在特定具体实施中,存储器1386、处理器1306、处理器1310、显示控制器1326、编解码器1334和调制解调器1348被包括在系统级封装件或片上系统设备1322中。在特定具体实施中,输入设备1330和电源1344耦合到系统级封装件或片上系统设备1322。此外,在特定具体实施中,如图13所示,显示器1328、输入设备1330、一个或多个扬声器106、一个或多个麦克风1390、天线1352和电源1344在系统级封装件或片上系统设备1322的外部。在特定具体实施

中,显示器1328、输入设备1330、一个或多个扬声器106、一个或多个麦克风1390、天线1352和电源1344中的每一者可耦合到系统级封装件或片上系统设备1322的组件,诸如接口(例如,输入接口114或输出接口124)或控制器。

[0118] 设备1300可包括智能扬声器、扬声器条、移动通信设备、智能电话、蜂窝电话、膝上型计算机、计算机、平板设备、个人数字助理、显示设备、电视、游戏控制台、音乐播放器、收音机、数字视频播放器、数字视频光盘(DVD)播放器、调谐器、相机、导航设备、交通工具、头戴式耳机,增强现实头戴式耳机、混合现实头戴式耳机、虚拟现实头戴式耳机、飞行器、家庭自动化系统、语音激活设备、无线扬声器和语音激活设备、便携式电子设备、汽车、计算设备、通信设备、物联网(IoT)设备、虚拟现实(VR)设备、基站、移动设备或它们的任何组合。

[0119] 结合所描述的具体实施,一种装置包括用于获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据的部件。例如,用于获得的部件可对应于一个或多个传感器104、输入接口114、音频播放器140、一个或多个处理器190、设备102、图1的系统100、图3的系统300、处理器1306、处理器1310、设备1300、被配置为获得电活动数据的一个或多个其他电路或组件、或它们的任何组合。

[0120] 该装置还包括用于基于该电活动数据呈现音频数据以在该音频数据的回放期间调整声源在声场中的位置的部件。例如,用于呈现的部件可对应于音频播放器140、一个或多个处理器190、设备102、图1的系统100、图3的系统300、处理器1306、处理器1310、设备1300、被配置为获得电活动数据的一个或多个其他电路或组件、或它们的任何组合。

[0121] 在一些具体实施中,一种非暂态计算机可读介质(例如,计算机可读存储设备,诸如存储器1386)包括指令(例如,指令1356),该指令在由一个或多个处理器(例如,一个或多个处理器190、一个或多个处理器1310或处理器1306)执行时使该一个或多个处理器获得与来自用户头部(例如,头部182)内的一个或多个电源(例如,一个或多个电源184)的电信号相对应的电活动数据(例如,电活动数据105)。该指令在由一个或多个处理器执行时使该一个或多个处理器基于电活动数据呈现音频数据(例如,音频数据141A),以在音频数据(例如,音频数据141C)的回放期间调整声源(例如,声源186)在声场(例如,声场142)中的位置(例如,位置150A)。

[0122] 下面在各组相关条款中描述了本公开的特定方面:

[0123] 根据条款1,一种设备包括:存储器,所述存储器被配置为存储指令;和一个或多个处理器,所述一个或多个处理器被配置为执行所述指令以:获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据;以及基于所述电活动数据呈现音频数据,以在所述音频数据的回放期间调整声源在声场中的位置。

[0124] 条款2包括根据条款1所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以经由一个或多个扬声器输出所述音频数据。

[0125] 条款3包括根据条款1或条款2所述的设备,其中所述电活动数据包括眼电图(EOG)数据、脑电图(EEG)数据或两者。

[0126] 条款4包括根据条款1至条款3中任一项所述的设备,还包括被配置为从一个或多个传感器接收所述电活动数据的接口。

[0127] 条款5包括根据条款4所述的设备,其中所述接口包括以太网接口、通用串行总线(USB)接口、Wi-Fi接口、蓝牙接口、串行端口接口或并行端口接口中的至少一者。

[0128] 条款6包括根据条款4或条款5所述的设备,其中所述一个或多个传感器包括入耳式传感器。

[0129] 条款7包括根据条款4或条款5所述的设备,其中所述一个或多个传感器包括电极帽、神经植入物、导电屏、非可穿戴传感器或它们的组合。

[0130] 条款8包括根据条款1至条款7中任一项所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以:经由一个或多个扬声器启动所述音频数据的第一回放操作,所述音频数据在所述第一回放操作期间被呈现为包括所述声源在所述声场中的多个位置,其中所述电活动数据基于在所述音频数据的所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述位置。

[0131] 条款9包括根据条款8所述的设备,其中所述一个或多个处理器被配置为执行所述指令以基于确定位置置信度水平是否小于置信度阈值而确定是否将所述音频数据呈现为包括所述多个位置,其中所述位置置信度水平与所述用户相对于参考的估计位置相关联。

[0132] 条款10包括根据条款9所述的设备,其中所述用户相对于所述参考的所述估计位置包括所述用户相对于所述参考的估计取向的估计取向、所述用户相对于所述参考的估计距离或两者。

[0133] 条款11包括根据条款9或条款10所述的设备,其中所述参考包括所述设备、显示设备、物理参考或虚拟参考中的一者或多者。

[0134] 条款12包括根据条款9至条款11中任一项所述的设备,其中所述一个或多个处理器被配置为执行所述指令以将所述估计位置初始化为对应于取向在所述用户前方的所述参考。

[0135] 条款13包括根据条款9至条款12中任一项所述的设备,其中所述一个或多个处理器被配置为将所述位置置信度水平初始化为小于位置阈值。

[0136] 条款14包括根据条款9至条款13中任一项所述的设备,其中所述一个或多个处理器被配置为执行所述指令以更新所述用户相对于所述参考的估计位置,所述估计位置是基于所述用户的头戴式耳机的惯性测量单元(IMU)数据、所述参考的空间数据、用户注视估计或它们的组合来更新的。

[0137] 条款15包括根据条款1至条款14中任一项所述的设备,其中所述一个或多个处理器被配置为执行所述指令以:基于所述电活动数据来确定用户注视估计;以及至少部分地基于所述用户注视估计来更新所述用户相对于参考的估计位置。

[0138] 条款16包括根据条款1至条款15中任一项所述的设备,其中所述一个或多个处理器被配置为执行所述指令以:使用机器学习模型来处理所述电活动数据以确定用户注视估计;以及基于所述用户注视估计来更新所述用户相对于参考的估计位置。

[0139] 条款17包括根据条款1至条款16中任一项所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以使用机器学习模型来处理所述电活动数据以确定所述声源的用户优选位置。

[0140] 条款18包括根据条款1至条款17中任一项所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以:将所述音频数据呈现为在所述声场中包括语音源的第

一位置和非语音源的第二位置;经由一个或多个扬声器启动所述音频数据的第一回放操作,其中所述电活动数据基于在所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述声源的所述位置。

[0141] 条款19包括根据条款18所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以响应于确定所述电活动数据指示单个声源被跟踪而确定所述用户优选位置对应于所述语音源的所述第一位置。

[0142] 条款20包括根据条款18或条款19所述的设备,其中所述一个或多个处理器被进一步配置为执行所述指令以响应于确定所述电活动数据指示所述语音源和所述非语音源被跟踪而确定所述用户优选位置对应于所述非语音源的所述第二位置。

[0143] 根据条款21,一种方法包括:在设备处获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据;以及基于所述电活动数据呈现音频数据,以在所述音频数据的回放期间调整声源在声场中的位置。

[0144] 条款22包括根据条款21所述的方法,还包括:经由一个或多个扬声器输出所述音频数据。

[0145] 条款23包括根据条款21或条款22所述的方法,其中所述电活动数据包括眼电图(EOG)数据、脑电图(EEG)数据或两者。

[0146] 条款24包括根据条款21至条款23中任一项所述的方法,还包括经由接口从一个或多个传感器接收所述电活动数据。

[0147] 条款25包括根据条款24所述的方法,其中所述接口包括以太网接口、通用串行总线(USB)接口、Wi-Fi接口、蓝牙接口、串行端口接口或并行端口接口中的至少一者。

[0148] 条款26包括根据条款24或条款25所述的方法,其中所述一个或多个传感器包括入耳式传感器。

[0149] 条款27包括根据条款24或条款25所述的方法,其中所述一个或多个传感器包括电极帽、神经植入物、导电屏、非可穿戴传感器或它们的组合。

[0150] 条款28包括根据条款21至条款27中任一项所述的方法,还包括:经由一个或多个扬声器启动所述音频数据的第一回放操作,所述音频数据在所述第一回放操作期间被呈现为包括所述声源在所述声场中的多个位置,其中所述电活动数据基于在所述音频数据的所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述位置。

[0151] 条款29包括根据条款28所述的方法,还包括基于确定位置置信度水平是否小于置信度阈值而确定是否将所述音频数据呈现为包括所述多个位置,其中所述位置置信度水平与所述用户相对于参考的估计位置相关联。

[0152] 条款30包括根据条款29所述的方法,其中所述用户相对于所述参考的所述估计位置包括所述用户相对于所述参考的估计取向的估计取向、所述用户相对于所述参考的估计距离或两者。

[0153] 条款31包括根据条款29或条款30所述的方法,其中所述参考包括所述设备、显示设备、物理参考或虚拟参考中的一者或多者。

[0154] 条款32包括根据条款29至条款31中任一项所述的方法,还包括将所述估计位置初始化为对应于取向在所述用户前方的所述参考。

[0155] 条款33包括根据条款29至条款32中任一项所述的方法,还包括将所述位置置信度水平初始化为小于位置阈值。

[0156] 条款34包括根据条款29至条款33中任一项所述的方法,还包括更新所述用户相对于所述参考的估计位置,所述估计位置是基于所述用户的头戴式耳机的惯性测量单元(IMU)数据、所述参考的空间数据、用户注视估计或它们的组合来更新的。

[0157] 条款35包括根据条款21至条款34中任一项所述的方法,还包括:基于所述电活动数据来确定用户注视估计;以及基于所述用户注视估计来更新所述用户相对于参考的估计位置。

[0158] 条款36包括根据条款21至条款35中任一项所述的方法,还包括:使用机器学习模型来处理所述电活动数据以确定用户注视估计;以及基于所述用户注视估计来更新所述用户相对于参考的估计位置。

[0159] 条款37包括根据条款21至条款36中任一项所述的方法,还包括使用机器学习模型来处理所述电活动数据以确定所述声源的用户优选位置。

[0160] 条款38包括根据条款21至条款37中任一项所述的方法,还包括:将所述音频数据呈现为在所述声场中包括语音源的第一位置和非语音源的第二位置;经由一个或多个扬声器启动所述音频数据的第一回放操作,其中所述电活动数据基于在所述第一回放操作期间来自所述一个或多个电源的所述电信号;以及基于所述电活动数据确定所述声源的用户优选位置,其中基于所述用户优选位置呈现所述音频数据以在所述音频数据的第二回放操作期间调整所述声源的所述位置。

[0161] 条款39包括根据条款38所述的方法,还包括响应于确定所述电活动数据指示单个声源被跟踪而确定所述用户优选位置对应于所述语音源的所述第一位置。

[0162] 条款40包括根据条款38或条款39所述的方法,还包括响应于确定所述电活动数据指示所述语音源和所述非语音源被跟踪而确定所述用户优选位置对应于所述非语音源的所述第二位置。

[0163] 根据条款41,一种设备包括:存储器,所述存储器被配置为存储指令;和处理器,所述处理器被配置为执行所述指令以执行根据条款21至40中任一项所述的方法。

[0164] 根据条款42,一种非暂态计算机可读介质存储指令,所述指令在由处理器执行时使所述处理器执行根据条款21至条款40中任一项所述的方法。

[0165] 根据条款43,一种装置包括用于执行根据条款21至条款40中任一项所述的方法的部件。

[0166] 根据条款44,一种非暂态计算机可读介质存储指令,所述指令在由一个或多个处理器执行时使所述一个或多个处理器:获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据;以及基于所述电活动数据呈现音频数据,以在所述音频数据的回放期间调整声源在声场中的位置。

[0167] 条款45包括根据条款44所述的非暂态计算机可读介质,其中所述指令在由所述一个或多个处理器执行时使所述一个或多个处理器更新所述用户相对于参考的估计位置,所述估计位置是基于所述用户的头戴式耳机的惯性测量单元(IMU)数据、所述参考的空间数

据、用户注视估计或它们的组合来更新的。

[0168] 条款46包括根据条款44或条款45所述的非暂态计算机可读介质,其中所述指令在由所述一个或多个处理器执行时使所述一个或多个处理器:基于所述电活动数据来确定用户注视估计;以及基于所述用户注视估计来更新所述用户相对于参考的估计位置。

[0169] 条款47包括一种装置,所述装置包括:用于获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据的部件;和用于基于所述电活动数据呈现音频数据以在所述音频数据的回放期间调整声源在声场中的位置的部件。

[0170] 条款48包括根据条款47所述的装置,其中用于获得的所述部件或用于呈现的所述部件中的至少一者集成在头戴式耳机、增强现实头戴式耳机、混合现实头戴式耳机、虚拟现实头戴式耳机、交通工具、通信设备、显示设备、电视、游戏控制台、音乐播放器、收音机、数字视频播放器、相机、导航设备、飞行器、家庭自动化系统、语音激活设备、无线扬声器和语音激活设备、便携式电子设备、汽车、计算设备、物联网(IoT)设备、移动设备或它们的任何组合中。

[0171] 技术人员还将明白的是,结合本文所公开的具体实施来描述的各个例示性的逻辑块、配置、模块、电路和算法步骤可以被实现为电子硬件、由处理器执行的计算机软件、或这两者的组合。各种例示性的组件、块、配置、模块、电路和步骤已在上文根据其功能性进行了一般性描述。此类功能性是被实现为硬件还是处理器可执行指令取决于特定应用和施加于整体系统的设计约束。技术人员可针对每种特定应用以不同方式来实现所描述的功能性,此类具体实施决策将不被解释为致使脱离本公开的范围。

[0172] 结合本文中所公开的具体实施所描述的方法或算法的步骤可直接在硬件中、在由处理器执行的软件模块中、或在这两者的组合中实施。软件模块可以驻留在随机存取存储器(RAM)、闪存存储器、只读存储器(ROM)、可编程只读存储器(PROM)、可擦除可编程只读存储器(EPROM)、电可擦除可编程只读存储器(EEPROM)、寄存器、硬盘、可移动盘、压缩光盘只读存储器(CD-ROM)、或本领域中已知的任何其他形式的非暂态存储介质。示例性存储介质耦合到处理器,使得处理器可以从该存储介质读取信息以及向该存储介质写入信息。在替代方案中,存储介质可与处理器成一整体。处理器和存储介质可以驻留在专用集成电路(ASIC)中。该ASIC可以驻留在计算设备或者用户终端中。在替换方案中,处理器和存储介质可作为分立组件驻留在计算设备或用户终端中。

[0173] 提供对所公开各方面的先前描述是为使本领域技术人员能够制作或使用所公开各方面。对这些方面的各种修改对于本领域技术人员而言将是显而易见的,并且本文中定义的原理可被应用于其他方面而不会脱离本公开的范围。由此,本公开并非旨在限定于本文中示出的各方面,而是应被授予可能与如由以下权利要求书所定义的原理和新颖性特征一致的最广义的范围。

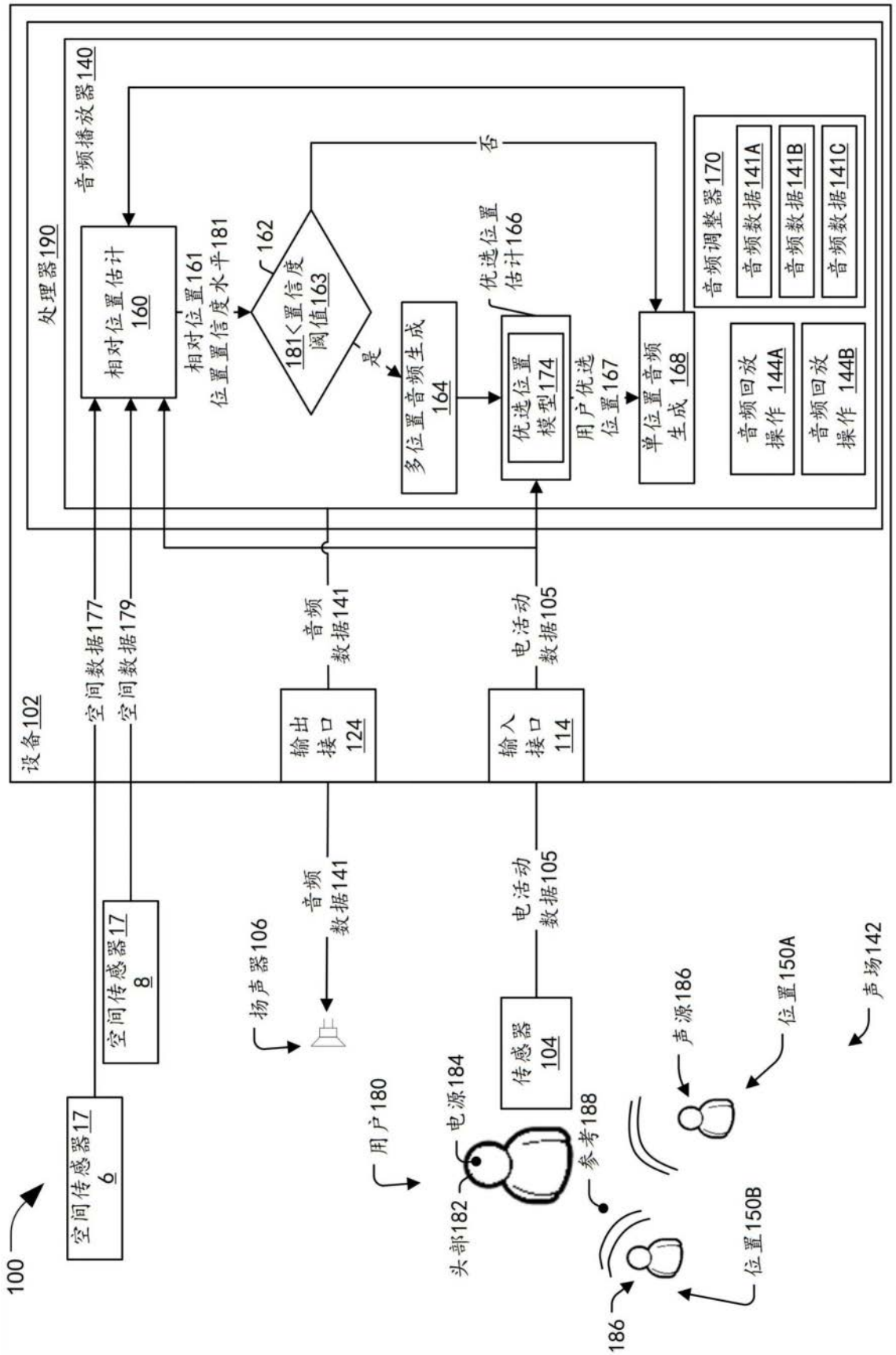

100
设备102
处理器190
音频播放器140
相对位置估计160
相对位置161
位置置信度水平181
162
181<置信度阈值163
是
否
多位置音频生成164
优选位置估计166
优选位置模型174
用户优选位置167
单位置音频生成168
音频调整器170
音频数据141A
音频数据141B
音频数据141C
音频回放操作144A
音频回放操作144B
空间数据177
空间数据179
音频数据141
电活动数据105
输出接口124
输入接口114
空间传感器176
空间传感器178
扬声器106
音频数据141
电活动数据105
传感器104
用户180
电源184
头部182
参考188
声源186
位置150A
声场142
186
位置150B

图1

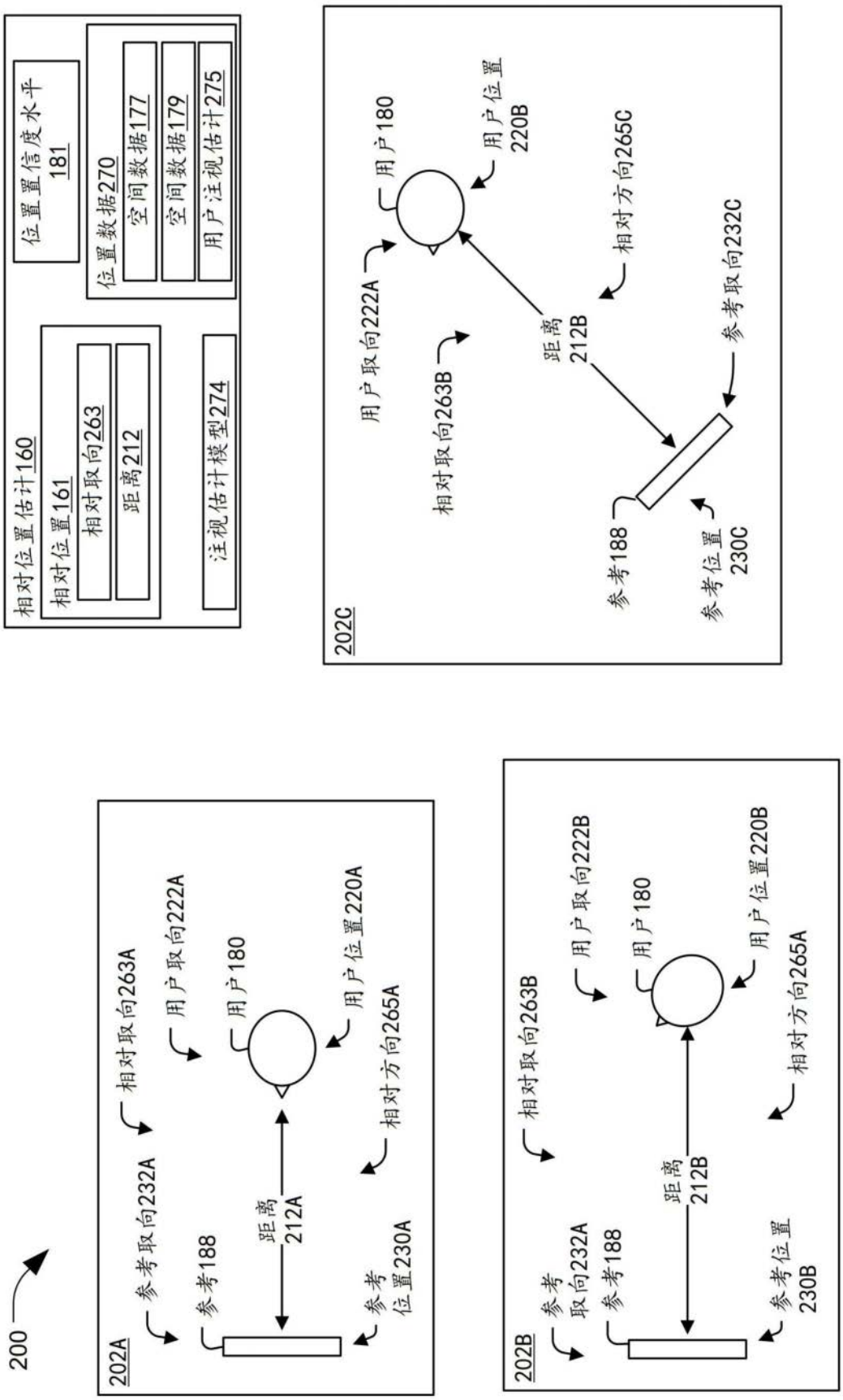

200
202A
参考取向232A
相对取向263A
用户取向222A
用户180
用户位置220A
参考188
距离212A
相对方向265A
参考位置230A
202B
参考取向232A
相对取向263B
用户取向222B
用户180
用户位置220B
参考188
距离212B
相对方向265A
参考位置230B
202C
用户取向222A
用户180
用户位置220B
相对取向263B
距离212B
相对方向265C
参考取向232C
参考188
参考位置230C
相对位置估计160
相对位置161
相对取向263
距离212
注视估计模型274
位置数据270
空间数据177
空间数据179
用户注视估计275
位置置信度水平181

图2A

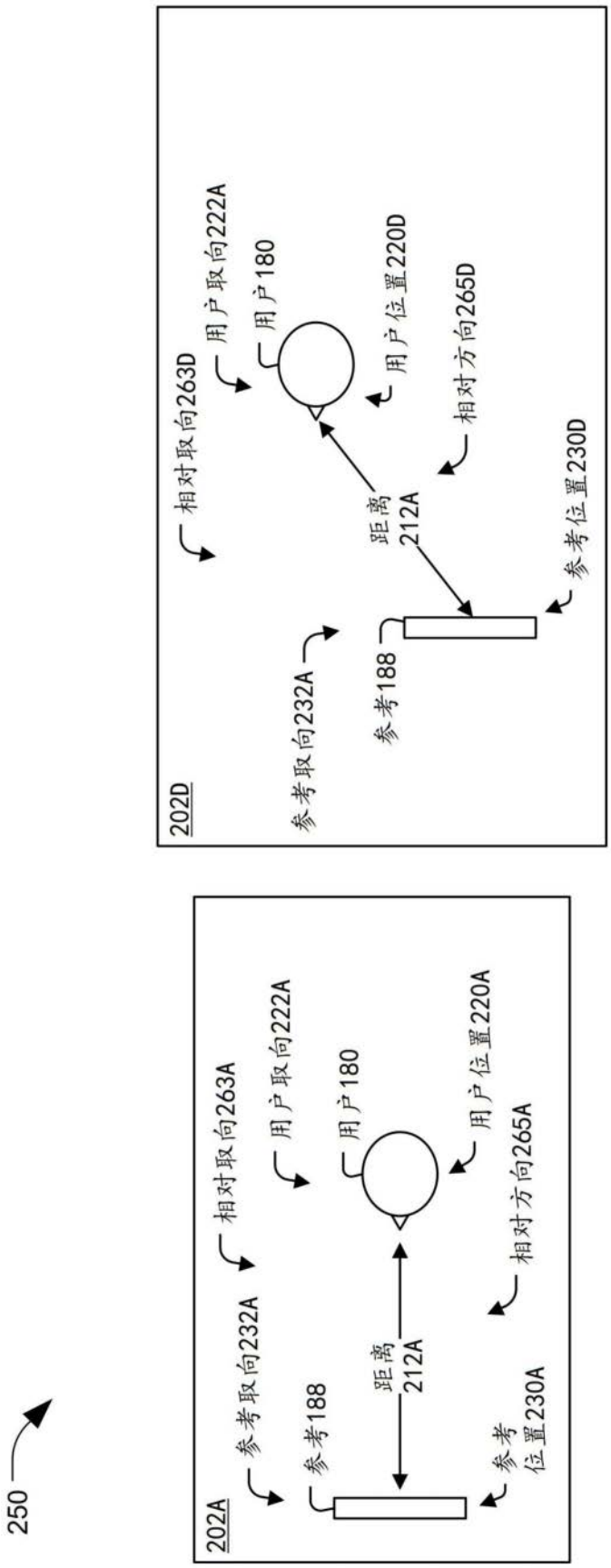

250
202A
参考取向232A
参考188
距离
212A
参考
位置230A
相对取向263A
用户取向222A
用户180
用户位置220A
相对方向265A
202D
参考取向232A
参考188
距离
212A
参考位置230D
相对取向263D
用户取向222A
用户180
用户位置220D
相对方向265D

图2B

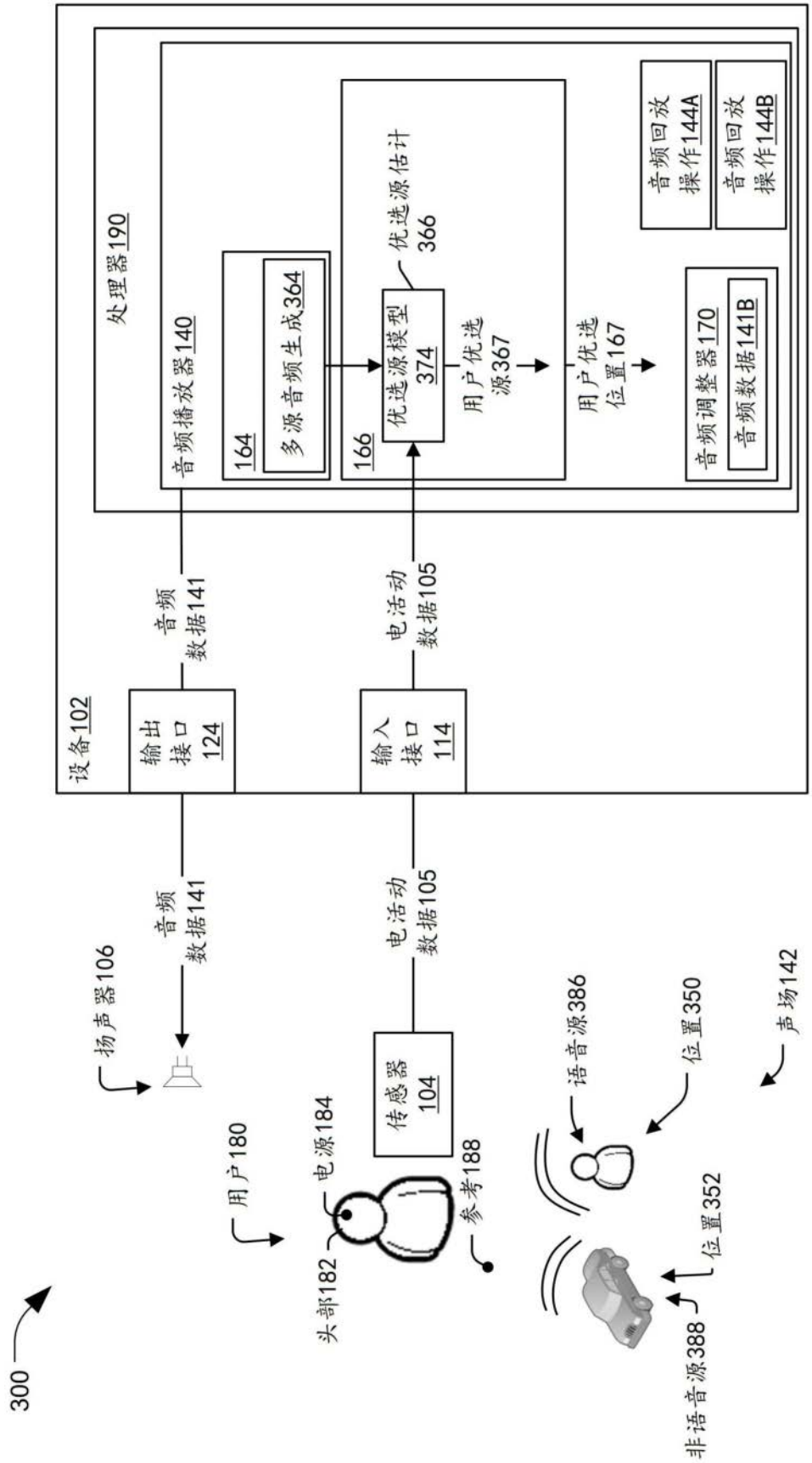

300
设备102
处理器190
音频播放器140
164
多源音频生成364
166
优选源估计366
优选源模型374
用户优选源367
用户优选位置167
音频回放操作144A
音频回放操作144B
音频调整器170
音频数据141B
输出接口124
输入接口114
音频数据141
电活动数据105
音频数据141
电活动数据105
扬声器106
传感器104
用户180
电源184
头部182
参考188
语音源386
位置350
位置352
非语音源388
声场142

图3

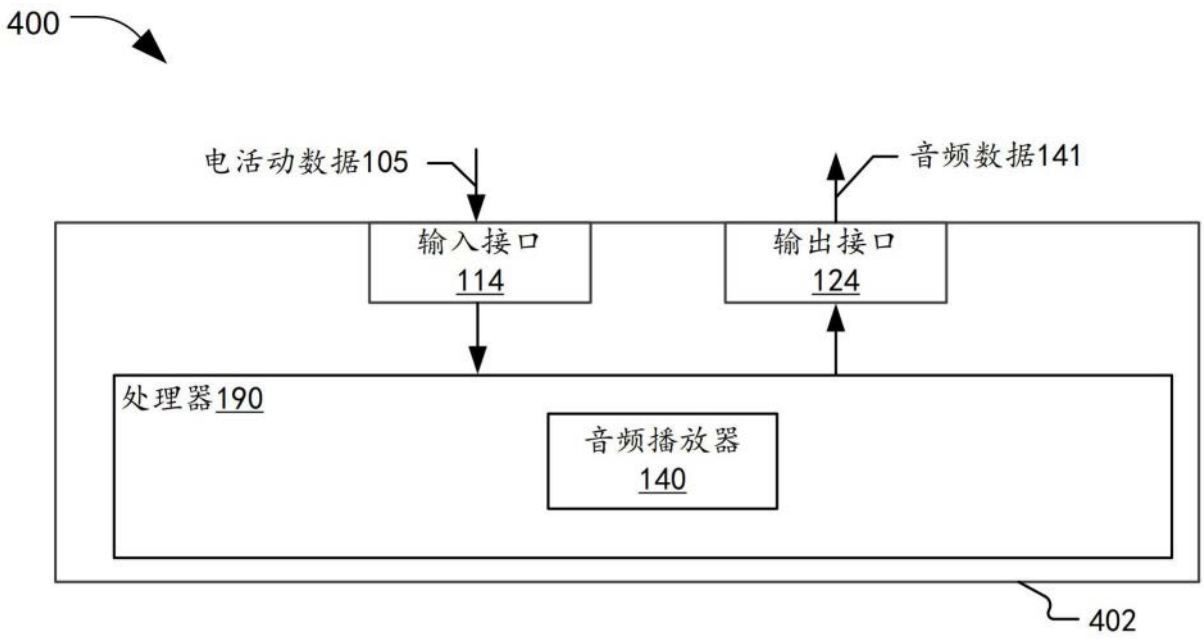

400
电活动数据105
音频数据141
输入接口
114
输出接口
124
处理器190
音频播放器
140
402

图4

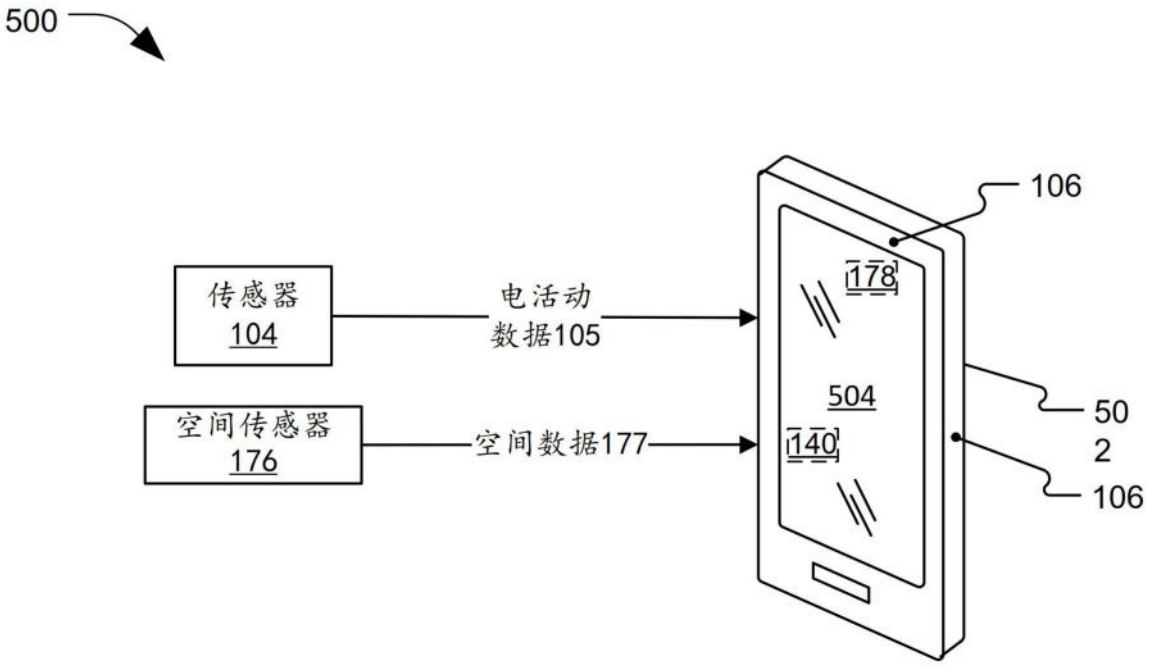

500
106
传感器
104
电活动
数据105
178
504
空间传感器
176
空间数据177
140
50
2
106

图5

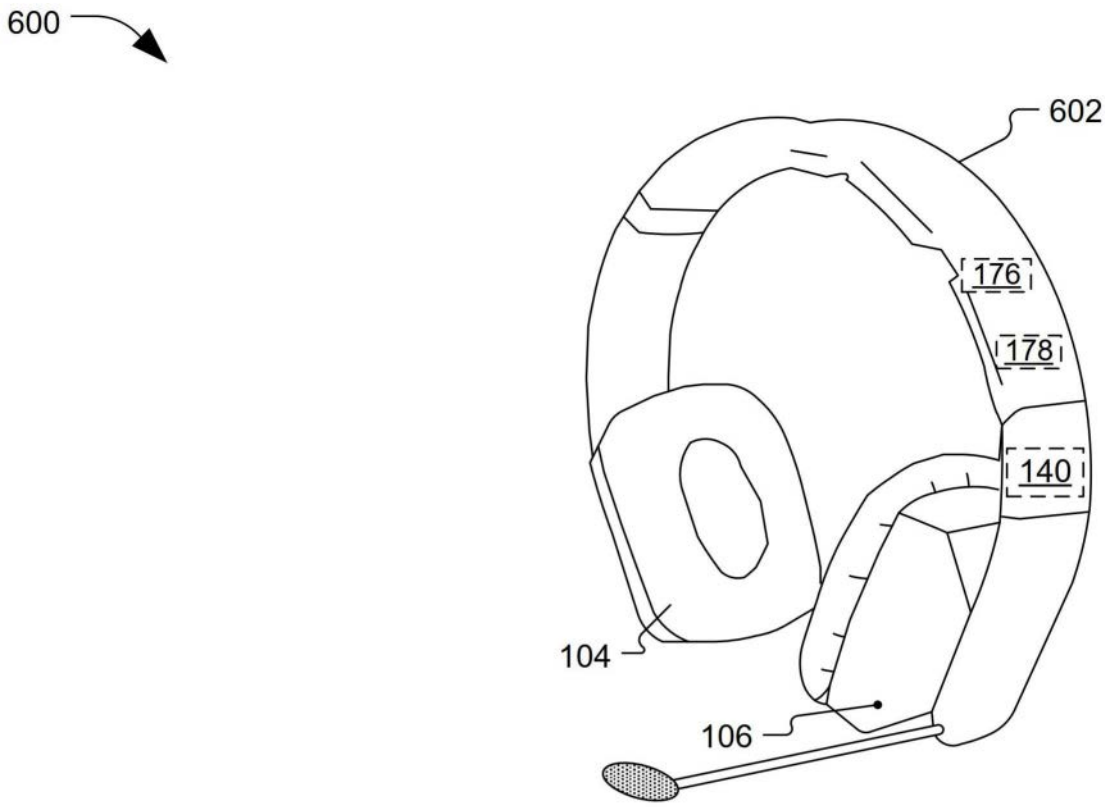

600
602
176
178
140
104
106

图6

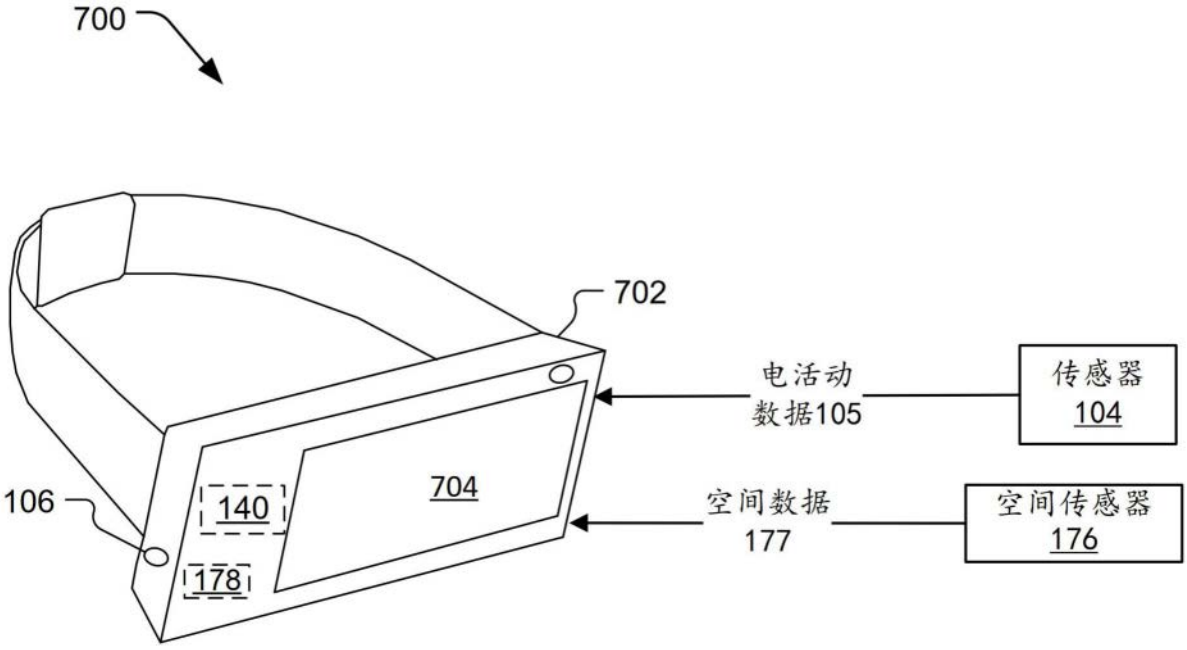

700
702
704
140
178
106
电活动
数据105
传感器
104
空间数据
177
空间传感器
176

图7

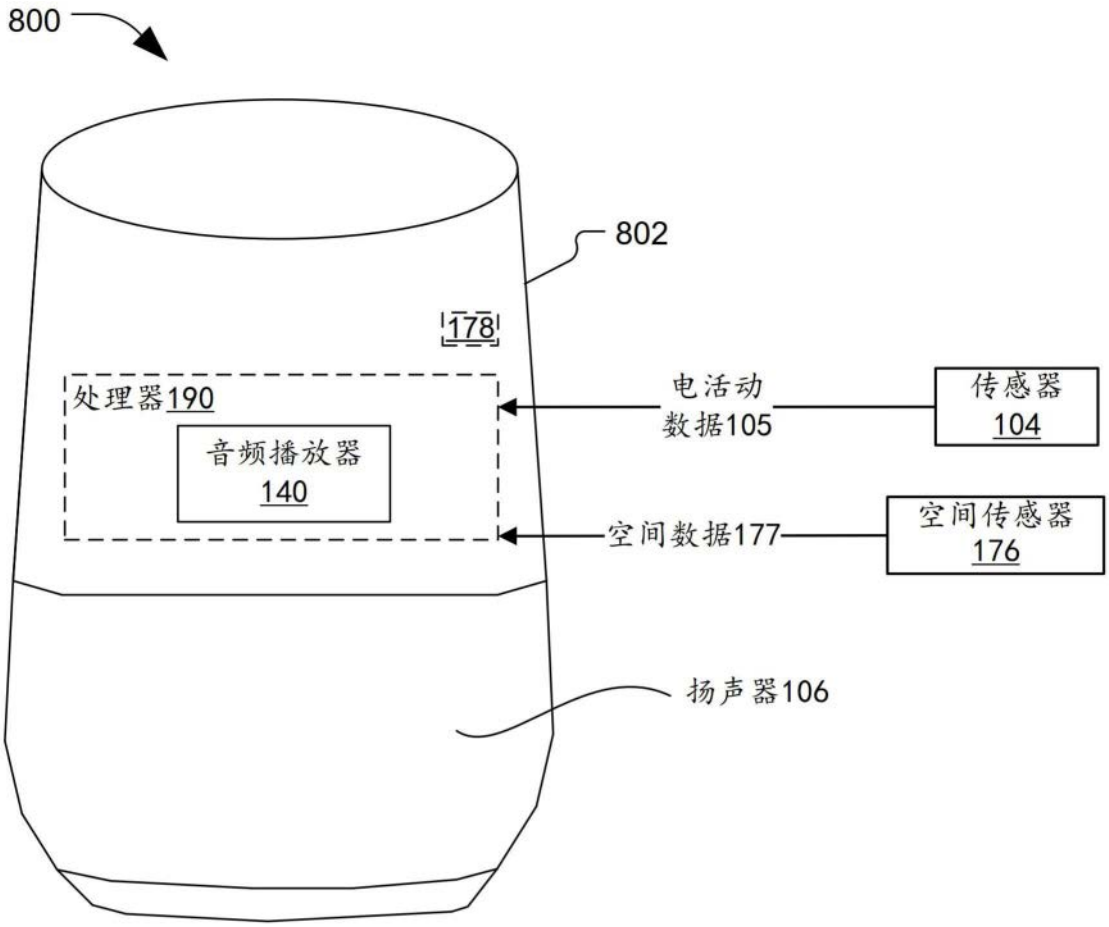

800
802
178
处理器190
音频播放器
140
电活动
数据105
传感器
104
空间数据177
空间传感器
176
扬声器106

图8

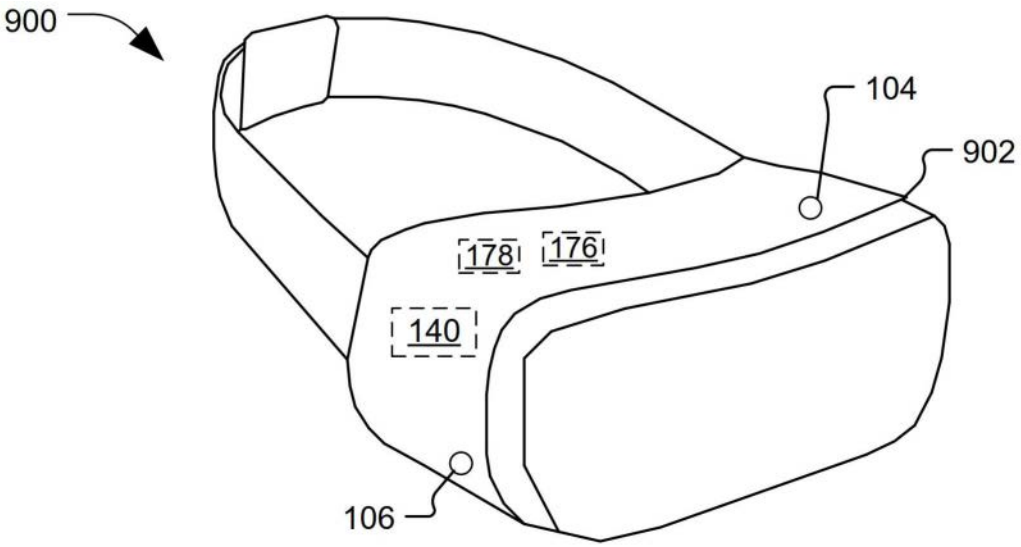

900
104
902
178
176
140
106

图9

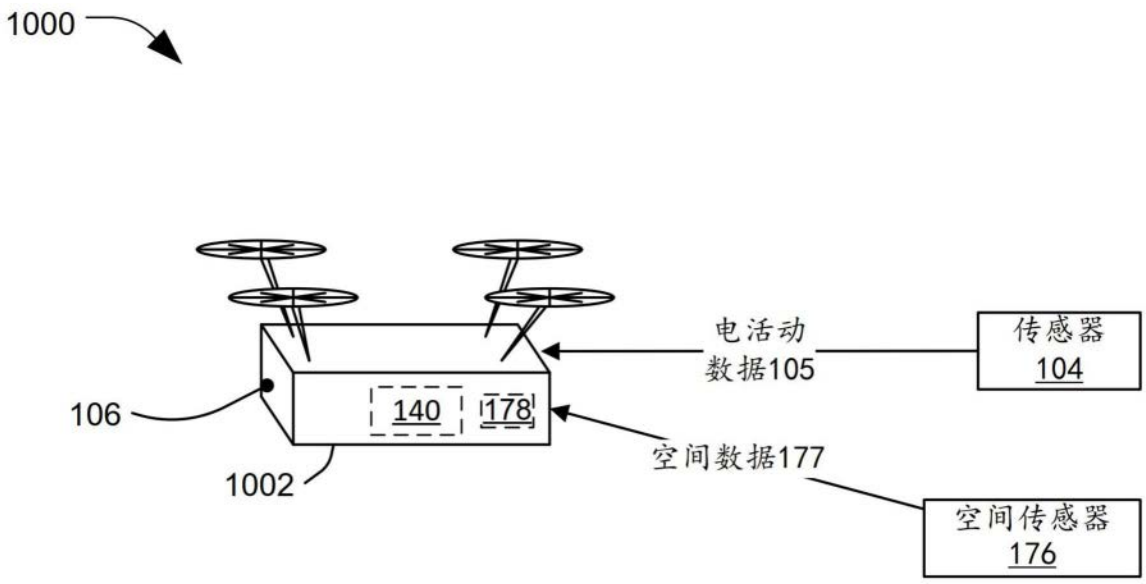

1000
106
140
178
1002
电活动
数据105
传感器
104
空间数据177
空间传感器
176

图10

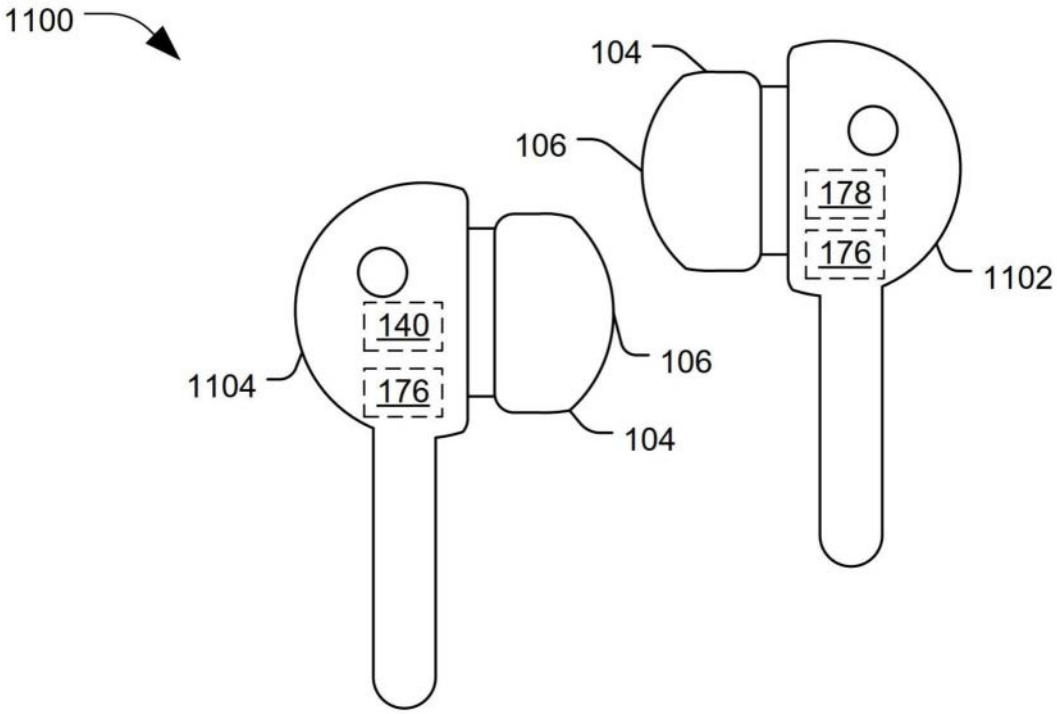

1100
104
106
178
176
1102
140
176
1104
106
104

图11

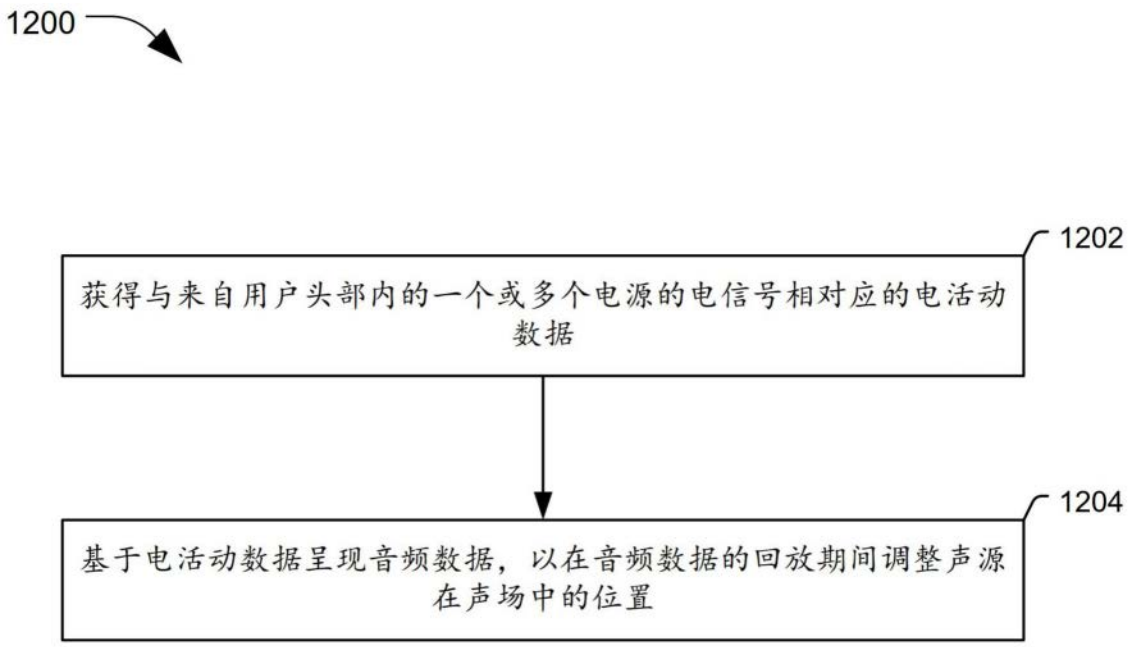

1200
1202
获得与来自用户头部内的一个或多个电源的电信号相对应的电活动数据
1204
基于电活动数据呈现音频数据，以在音频数据的回放期间调整声源在声场中的位置

图12

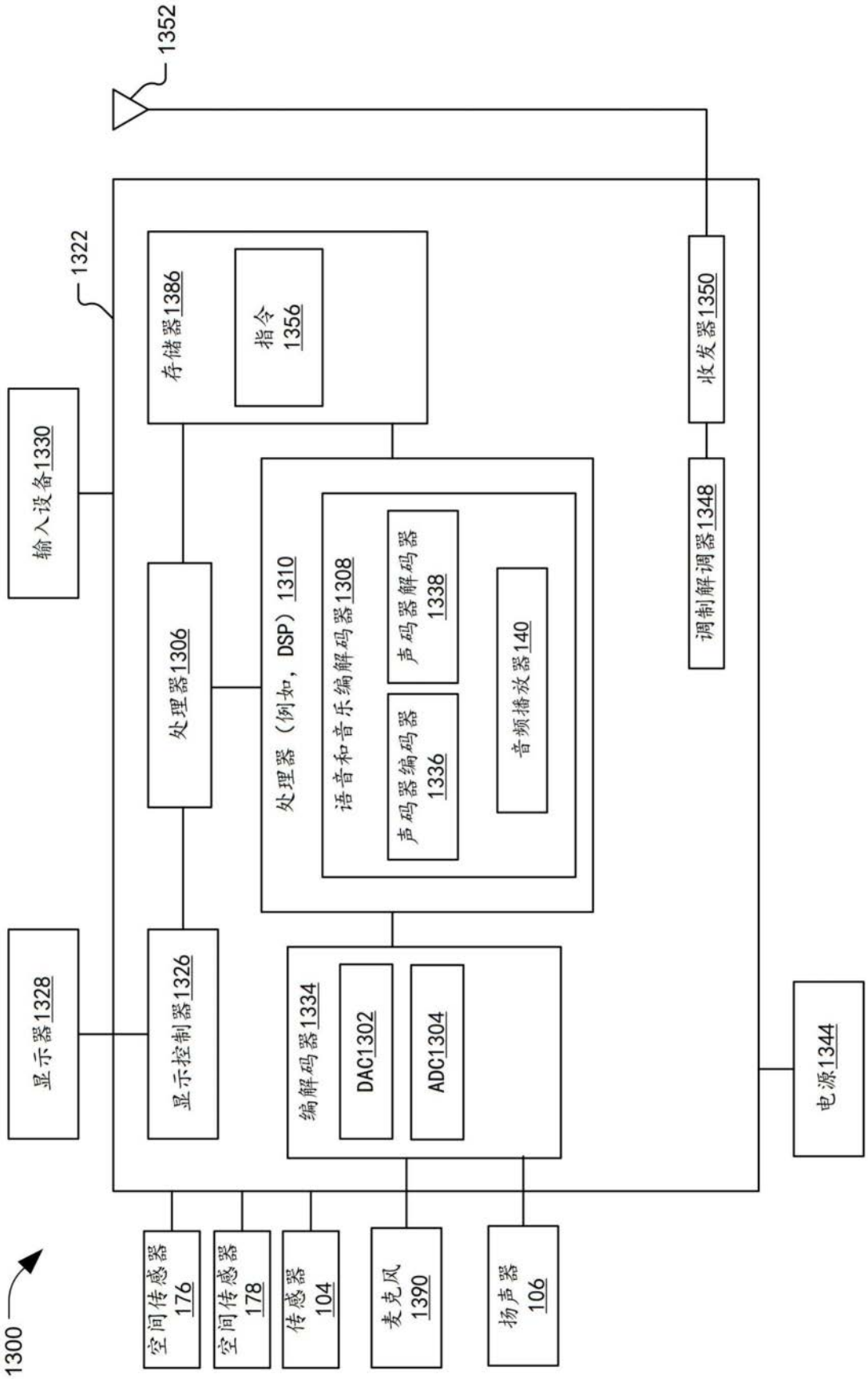

1352
1322
存储器1386
指令
1356
输入设备1330
处理器1306
处理器（例如，DSP）1310
语音和音乐编解码器1308
声码器解码器
1338
声码器编码器
1336
音频播放器140
收发器1350
调制解调器1348
显示器1328
显示控制器1326
编解码器1334
DAC1302
ADC1304
电源1344
空间传感器
176
空间传感器
178
传感器
104
麦克风
1390
扬声器
106
1300

图13