(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105934227 B
(45)授权公告日 2019.12.03

(21)申请号 201580005723.3
(22)申请日 2015.01.22
(65)同一申请的已公布的文献号
申请公布号 CN 105934227 A
(43)申请公布日 2016.09.07
(30)优先权数据
14/163,943 2014.01.24 US
(85)PCT国际申请进入国家阶段日
2016.07.25
(86)PCT国际申请的申请数据
PCT/US2015/012329 2015.01.22
(87)PCT国际申请的公布数据
WO2015/112651 EN 2015.07.30
(73)专利权人 微软技术许可有限责任公司
地址 美国华盛顿州
(72)发明人 R·杰拉尔德
(74)专利代理机构 上海专利商标事务所有限公司 31100
代理人 杨洁
(51)Int.Cl.
G01C 21/20(2006.01)
A61H 3/06(2006.01)
(56)对比文件
CN 103458838 A,2013.12.18,
CN 1712900 A,2005.12.28,
US 2007016425 A1,2007.01.18,
US 2013216093 A1,2013.08.22,
CN 202409427 U,2012.09.05,
审查员 贾仁杰

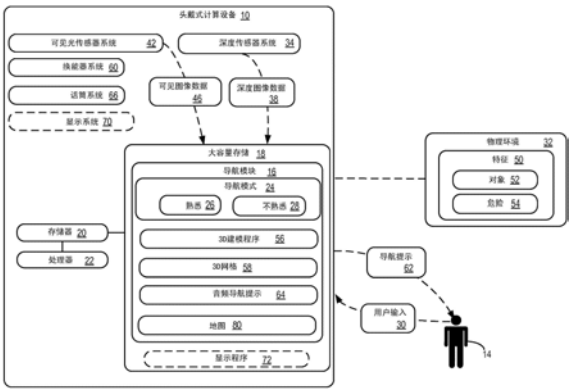
权利要求书2页 说明书10页 附图9页

(54)发明名称

音频导航辅助

(57)摘要

揭示了有关用于通过音频输出提供导航辅助的头戴式计算设备的各实施例。例如,在一个所揭示的实施例中,深度图像数据和可见图像数据被用于生成环境的至少一个部分的三维网格。使用该网格,环境中的至少一个特征被检测。当操作在熟悉导航模式时并基于检测到特征,第一音频导航提示被输出给用户。当操作在不熟悉导航模式时并基于检测到特征,第二音频导航提示被输出给用户,其中第二音频导航提示不同于第一音频导航提示。



CN 105934227 B

1.一种用于通过音频输出在环境中提供导航辅助的头戴式计算设备,所述头戴式计算设备包括:

用于从所述环境中生成深度图像数据的一个或多个深度传感器;

用于从所述环境中生成可见图像数据的一个或多个可见光传感器;

一个或多个换能器;以及

由所述头戴式计算设备的处理器执行的导航模块,所述导航模块包括熟悉导航模式和不熟悉导航模式,所述导航模块被配置以:

使用所述深度图像数据和可见图像数据来生成所述环境的至少一个部分的三维网格;

接收标识所述环境的用户输入;

根据接收到的用户输入给所述环境打上标签,并生成所述环境的至少一个特征;

检测所述环境是否被打上标签使用所述三维网格来检测所述环境中的至少一个特征;

响应于检测到所述环境是被打上标签的,操作在所述熟悉导航模式中;

当操作在所述熟悉导航模式时并基于检测到所述特征,通过所述一个或多个换能器将所述环境的第一音频导航提示输出给所述用户;以及

当操作在不熟悉导航模式时并基于检测到所述特征,通过所述一个或多个换能器将所述环境的第二音频导航提示输出给所述用户,其中所述第二音频导航提示不同于所述第一音频导航提示。

2.如权利要求1所述的头戴式计算设备,其特征在于,所述导航模块被进一步配置来以致使所述用户察觉所述提示好像源自所述特征的三维位置处的方式来向所述用户输出所述第一音频导航提示和所述第二音频导航提示的一个或多个。

3.如权利要求1所述的头戴式计算设备,其特征在于,所述特征包括斜坡,并且所述第一音频导航提示和所述第二音频导航提示包括指示上坡的多个升序定调的声音或指示下坡的多个降序定调的声音。

4.如权利要求3所述的头戴式计算设备,其特征在于,所述第一音频导航提示和所述第二音频导航提示进一步包括指示连续斜坡的多个持续定调的声音或指示多个台阶的多个离散定调的声音。

5.如权利要求1所述的头戴式计算设备,其特征在于,所述导航模块被进一步配置来以随着所述特征到所述用户的邻近度减少而增大的音量向所述用户输出所述第一音频导航提示和所述第二音频导航提示的一个或多个。

6.一种用于在头戴式计算设备中通过音频输出在环境中提供导航辅助的方法,所述方法包括:

从所述环境接收深度图像数据;

从所述环境接收可见图像数据;

使用所述深度图像数据和可见图像数据,生成所述环境的至少一个部分的三维网格;

接收标识所述环境的用户输入;

根据接收到的用户输入给所述环境打上标签,并生成所述环境的至少一个特征;

检测所述环境是否被打上标签;

使用所述三维网格,检测所述环境中的至少一个特征;

响应于检测到所述环境是被打上标签的,所述头戴式计算设备在熟悉导航模式中操

作；

当所述头戴式计算设备在所述熟悉导航模式中操作时并基于检测到所述特征，通过一个或多个换能器将所述环境的第一音频导航提示输出给所述用户；以及

当所述头戴式计算设备操作在不熟悉导航模式时并基于检测到所述特征，通过所述一个或多个换能器将所述环境的第二音频导航提示输出给所述用户，其中所述第二音频导航提示不同于所述第一音频导航提示。

7.如权利要求6所述的方法，其特征在于，进一步包括：

使用所述深度图像数据和所述可见图像数据的一个或多个来识别所述环境中的脸；

将所述脸与人的身份相关联；以及

通知所述用户此人的所述身份。

8.如权利要求7所述的方法，其特征在于，进一步包括：

向此人的所述身份分派身份耳标；以及

通过所述一个或多个换能器向所述用户输出所述身份耳标。

9.如权利要求6所述的方法，其特征在于，进一步包括通过所述一个或多个换能器向所述用户输出触觉导航提示。

10.如权利要求6所述的方法，其特征在于，进一步包括，基于检测到所述特征，经由所述头戴式计算设备的显示系统向所述用户显示视觉导航提示，其中所述视觉导航提示至少基于所述环境中的特征的三维位置。

音频导航辅助

[0001] 背景

[0002] 具有生理视觉缺陷的人可利用一个或多个移动性帮助来导航他们周围。此类移动性帮助可包括，例如，拐杖、辅助动物、视觉增强设备等。类似地，遇到诸如黑暗或充满烟的房间的视觉缺陷的情况的看得见的人也可得益于对他们的当前环境提供导航辅助的移动性帮助。

[0003] 某些移动性帮助提供指示用户周围的各方面的触觉反馈，诸如接触街道上路缘的拐杖。然而，使用这样的帮助可能占用用户的大量注意力和用户的手。此外，此类帮助一般提供用户近周围的相同量和类型的反馈，而不考虑那些周围的上下文。因此，这些帮助一般不能基于用户当前环境的上下文来修改或增强它们的反馈的用处。

[0004] 概述

[0005] 在此公开了涉及在环境中通过音频输出提供导航辅助的各个实施例。例如，一个所揭示的实施例在头戴式计算设备中提供一种包括从环境中接收深度图像数据和可见图像数据的方法。使用该深度图像数据和可见图像数据，环境的至少一个部分的三维网格被生成。使用该三维网格，环境中的至少一个特征被检测到。当头戴式计算设备在熟悉导航模式中操作时并基于检测到特征，环境的第一音频导航提示通过一个或多个换能器被输出给用户。当头戴式计算设备在不熟悉导航模式中操作时并基于检测到特征，环境的第二音频导航提示通过一个或多个换能器被输出给用户，其中第二音频导航提示不同于第一音频导航提示。

[0006] 提供本概述以便以简化的形式介绍以下在详细描述中进一步描述的一些概念。本发明内容并不旨在标识所要求保护主题的关键特征或必要特征，也不旨在用于限制所要求保护主题的范围。而且，所要求保护的主题不限于解决该公开的任一部分中所注的任何或全部缺点的实现方式。

[0007] 附图简述

[0008] 图1是根据本公开的一实施例的用于提供导航辅助的头戴式计算设备的示意图。

[0009] 图2示出根据本公开的一实施例的示例头戴式计算设备。

[0010] 图3是家中佩戴图2的头戴式计算设备的用户的示意图。

[0011] 图4是宾馆中佩戴头戴式计算设备的图3的用户的示意图。

[0012] 图5是接近包括楼梯的斜坡的佩戴图2的头戴式计算设备的用户的示意图。

[0013] 图6是接近持续下坡的图6的用户的示意图。

[0014] 图7是根据本公开的一实施例的通过头戴式计算设备观看的室外环境的示意图。

[0015] 图8A和8B是根据本公开的一实施例的用于在环境中通过音频输出向用户提供导航辅助的方法的流程图。

[0016] 图9是计算设备的一实施例的简化示意解说。

[0017] 详细描述

[0018] 图1示出了用于向用户14提供导航辅助的头戴式计算设备10的一个实施例示意图。头戴式计算设备10包括可被存储在头戴式计算设备10的大容量存储18中的导航模块

16。导航模块16可被加载到存储器20中并由头戴式计算设备10的处理器22执行以执行以下更为详细地描述的方法和过程中的一个或多个。

[0019] 导航模块16包括包含熟悉导航模式26和不熟悉导航模式28的多个导航模式24。有利地并且如以下更详细地描述的，导航模式24的一个或多个可被选择性地使用来以上下文合适的方式提供针对用户特定需求定制的导航辅助。例如，当用户14导航该用户熟悉的环境时，熟悉导航模式26可以程序性的或通过用户输入30来被使用。在熟悉导航模式26中，导航辅助可被定制以说明用户与她周围的熟悉度。

[0020] 以类似方式，当用户14导航该用户不熟悉的环境时，不熟悉导航模式28可以程序性的或通过用户输入30来被使用。在不熟悉导航模式28中，导航辅助可被定制以说明用户与她周围的不熟悉度。将理解，导航模块16也可包括与具有特定特性或方面的环境或周围相关的一个或多个附加导航模式24诸如，例如，室内导航模式、室外导航模式、城市导航模式、乡村导航模式等。在某些示例中，两个或更多个导航模式可被同时使用，诸如例如，熟悉导航模式和室内导航模式。

[0021] 头戴式计算设备10可包括从物理环境32接收物理环境数据的各种传感器和相关的系统。现在还参考图2，提供了一副可戴的眼镜形式的头戴式计算设备200的一个示例。在某些示例中，头戴式计算设备200可以包括被支撑在查看者的一只或两只眼睛前方的透明、半透明或不透明显示器204。在其它示例中，头戴式计算设备200可以采取可包括或可不包括被支撑在用户的一只或两只眼睛前方的显示器的其它形式。

[0022] 例如，在一些实施例中，头戴式计算系统10可以包括不包含显示器的一副眼镜框架。因此，具有各种形状因子的许多其它类型和配置的头戴式计算设备100也可被使用并且在本公开的范围内。还将理解，图1中所示的头戴式计算设备10可以采取图2中所示的头戴式计算设备200(如以下更详细地描述的)或任何其他合适的头戴式计算设备的形式。

[0023] 参考图1和图2，在此示例中，头戴式计算设备10包括深度传感器系统34，该深度传感器系统包括生成深度图像数据38的一个或多个深度相机208。深度传感器系统34还可以检测其视野内的运动，如用户视野内的用户14或人或物理对象所执行的基于姿势的输入或其他运动。在一个示例中，每一深度相机208可包括立体视觉系统的左和右相机。来自这些深度相机中的一个或多个的时间分辨的图像可被彼此配准和/或与来自另一光学传感器(如可见光谱相机)的图像配准，且可被组合以产生深度分辨的视频。

[0024] 在其他示例中，结构化光深度相机可被配置成投影结构化红外照明并对从照明被投影到其之上的场景中反射的该照明进行成像。基于所成像的场景的各个区域内邻近特征之间的间隔，可构造该场景的深度图。在其他示例中，深度相机可以采取飞行时间深度相机的形式，其被配置成将脉冲的红外照明投影到该场景上以及检测从该场景反射的照明。可以理解，在本发明的范围内可使用任意其他合适的深度相机。

[0025] 该头戴式计算设备10还可包括利用至少一个面向外的传感器212(如RGB相机或光学传感器)的可见光传感器系统42。该可见光传感器系统42可生成被提供给导航模块16的可见图像数据46。面向外的传感器212还可从物理环境32和该环境内的特征50捕捉二维图像信息。如以下更详细描述的，特征50可包括环境32中的物理对象52和危险54。在一些示例中，物理对象52还可包括危险54。面向外的传感器212可以捕捉用户14位于其中的物理环境32的图像。

[0026] 头戴式计算设备10还可包括使用深度图像数据38和可见图像数据46来生成三维网络58的3D建模程序56，该建模程序对围绕用户14的物理环境的至少一部分建模。在某些示例中，3D建模程序56可利用多边形建模技术，其中3D空间的顶点由线段连接以生成纹理化的3D多边形网格。还应当理解，用于3D建模的其它技术可被利用，包括，但不限于，曲线建模和数字雕刻技术。

[0027] 头戴式计算设备10还可包括换能器系统60，换能器系统包括将来自导航模块16的电信号转换成另一形式能量的一个或多个致动器。如以下更详细描述的，导航模块16可利用该换能器系统60来向用户14输出环境32的导航提示62。再次参考图2，并且在某些示例中，换能器系统60可包括用于生成音频导航提示64的一个或多个扬声器216。在其它示例中，一个或多个扬声器216可采取用户14佩戴的头戴式耳机或耳内耳塞的形式。

[0028] 在某些示例中，导航模块16和/或换能器系统60可处理音频导航提示64以使得用户14能够察觉发生在物理环境32的3D空间中特定位置的提示。例如，一个或多个串音抵消机制可被利用并被配置成使得第一音频信号(例如，左通道)被递送至第一耳朵(例如，左耳)而第二音频信号(例如，右通道)被递送至第二耳朵(例如，右耳)，同时显著衰减第一音频信号到第二耳朵的递送以及第二音频信号到第一耳朵的递送。

[0029] 在其它示例中，提供三维音频可基于与头部有关的传递函数“HRTF”和/或与头部有关的脉冲响应“HRIR”来创建声音源自3D声学空间中的特定位置的错觉。HRTF描述了在声音抵达耳膜和内耳之前给定声波输入如何被头部和耳廓的衍射和反射属性过滤。换言之，HRTF可以基于自由空间中的声音与到达耳膜时的声音之间的差异来定义。

[0030] 在其它示例中，换能器系统60可包括一个或多个触觉换能器220，用于向用户14生成触觉导航提示，诸如振动。

[0031] 头戴式计算设备10还可包括话筒系统66和用于接收来自物理环境32的音频输入的一个或多个话筒224。在某些示例中，包括置于头戴式计算设备200上不同位置的多个话筒224的话筒阵列可被提供。话筒224可包括被配置用于接收语音和来自物理环境32的其它音频输入的全向话筒和/或单向话筒。

[0032] 在某些示例中，话筒系统66和/或导航模块16可利用一个或多个声源定位技术来在物理环境32的3D空间中定位音频源，包括位于用户14身后的源。例如，话筒系统66和/或导航模块16可将一个或多个波束成形技术应用于来自话筒阵列的音频输入的至少一个部分。例如，可以以任何合适的方式来确定单个定向自适应声音信号。可以基于不随时间改变的波束成形技术、自适应波束成形技术或不随时间改变和自适应波束成形技术的组合来确定定向自适应声音信号。所得到的已组合信号可以具有窄方向性模式，该模式在音频源方向上前进。还将理解，可以使用任何合适的声源定位技术来标识音频源的位置。

[0033] 如上面注意的，在某些示例中，头戴式计算设备200可包括显示系统70和使视觉导航提示能够被递送到用户14的眼睛的显示器204。在一个示例中，显示器204可包括透明显示器，该透明显示器被配置来向通过该透明显示器查看环境的用户14视觉地增强物理环境32的外观。例如，物理环境32的外观可以由经由透明显示器呈现的图形内容(例如，一个或多个像素，每一像素具有相应色彩和亮度)来增强以创建混合现实环境。

[0034] 在这个示例中，透明显示器204还可被配置成使用户能够透过显示虚拟对象表示的一个或多个部分透明的像素来查看物理环境32中的物理现实世界对象52。在一个示例

中，透明显示器204可包括位于透镜210内的图像生成元件(诸如例如透视有机发光二极管(OLED)显示器)。作为另一示例，透明显示器204可包括在透镜210边缘上的光调制器。在这一示例中，透镜210可以担当光导以供将光从光调制器递送到用户的眼睛。这样的光导可以使用户能够察觉位于用户正在查看的物理环境32内的2D图像或3D全息图像，同时还允许用户查看物理环境中的物理对象52。

[0035] 再次参考图1的头戴式计算设备10，将理解处理器22可包括具逻辑子系统而存储器20可包括存储子系统，如以下参考图9更详细地讨论的，所述逻辑子系统和存储子系统与上述各种传感器和对应系统进行通信。在一个示例中，存储子系统可包括可由逻辑子系统执行的指令，用以接收来自传感器的信号输入并将此类输入(以未经处理或经处理的形式)转发到导航模块16。

[0036] 此外，图1中示出的示例示出将大容量存储18、存储器20和处理器22合并到头戴式计算设备10中。将理解，在其它示例中大容量存储18、存储器20和处理器22的一个或多个可被置于头戴式计算设备10通信地耦合到的一个或多个其它计算设备中。例如，头戴式计算设备10可使用有线连接，或可采用经由WiFi、蓝牙或任何其他合适的无线通信协议的无线连接来与另一计算设备在操作上连接。关于头戴式计算设备10和其它计算设备的组件和计算方面的附加细节以下参考图9更详细地描述。

[0037] 将理解，头戴式计算设备10和相关的传感器以及上面描述的并在图1和2中示出的其他组件是作为示例来提供的。这些示例不旨在以任何方式进行限制，因为任何其他合适的传感器、组件，和/或传感器和组件的组合可被使用。因此，应当理解，头戴式计算机设备10可以包括未偏离本公开文本范畴的附加和/或替代的传感器、相机、话筒、输入设备、输出设备等。此外，头戴式计算设备10的物理配置及其各种传感器和子组件可以采取不偏离本公开文本范畴的各种不同形式。

[0038] 现在参考图3-7，现在将提供对头戴式计算系统10的示例用例和实施例的描述。在图3中，戴着图2的头戴式计算设备200的用户14可能正在他家300中走动。用户14可能具有视觉缺陷并且可将头戴式计算设备200用于导航辅助。在一个示例中，用户14可通过在他家300中从房间到房间走动并且，对于每个房间，提供标识该房间的房间标识符的用户输入30，来初始化头戴式计算设备200。

[0039] 例如，当用户14在厨房304中四处走动时，用户说，“这是我的厨房”。导航模块16可接收这个用户输入30，使用语音识别程序来处理它，并用“厨房”标识符给厨房304打上标签。导航模块16可使用厨房304的可见图像数据46和深度图像数据38来生成厨房的3D网格及其各种特征50，诸如设施、家具、墙、轮廓、表面等。导航模块16还可生成家300的对应的地图80，该地图包括被用“厨房”标识符打了标签的厨房304。以类似的方式，导航模块16可使用起居室308的可见图像数据46和深度图像数据38来生成该房间的3D网格及其各种特征50。导航模块可基于来自用户14的用户输入30用“起居室”标识符给该起居室打上标签，并可将用“起居室”标识符打上标签的该起居室308加入家300的地图80。当用户14围绕家300从房间走到房间时，导航模块16在导航房间时还可识别并存储用户采用的常用路径。

[0040] 随后，当用户14进入这些房间之一时，导航模块16可通过地图80标识该房间并将该房间检测为用对应的房间标识符打上标签的。因此并至少基于将该房间检测为用对应的房间标识符打上标签的，导航模块16可使用熟悉导航模式26。在另一示例中，房间或物理环

境32的其它部分可能不与标识符相关联，但可能已经先前被用户14访问过。在这些示例中，导航模块16可通过利用例如机器学习技术从用户的一个或多个先前访问来识别房间或环境32的其它部分。如果用户已经访问过该房间或环境32的其它部分达至少预定次数，诸如1、3、5或任何其它合适的次数，则导航模块16可使用熟悉导航模式26。在又一个其它示例中，用户14可向使用熟悉导航模式26的导航模块16提供用户输入30。

[0041] 继续参考图3并如以下更详细描述的，尽管处在熟悉导航模式26中，导航模块16仍可向用户14提供与导航模块16操作在不熟悉导航模式28中时所提供的导航提示所不同的导航提示62。在一个示例中，用户14可能向着位于平常位置314处的坐凳312的通常方向走。在熟悉导航模式26中，在用户从前次到访该起居室308中记住了坐凳312的平常位置314的假设下，导航模块16可能不提供对应于坐凳312的导航提示62。在另一个示例中，导航模块16可输出包含对坐凳312的温和的提醒(诸如通过话筒216用柔和的声音发出“坐凳”)的音频导航提示64。

[0042] 在另一示例中，用户14可以是家300的首次到访客人。因此，导航模块16可使用不熟悉导航模式28。在此模式中，随着用户14朝着坐凳312方向步行，导航模块16可向用户提供提醒用户坐凳位置的音频导航提示64。例如，导航模块16可使用一个或多个声源定位技术来输出简短的，特殊声音(或耳标)，示为熟悉耳标320，来使得用户14感觉到该耳标好像发自平常对应于坐凳312的3D位置314处。在另一个示例中，导航模块16可使用声源定位技术从平常位置314向用户14发出单词“坐凳”的声音。

[0043] 在另一个示例中且尽管在熟悉导航模式26中，导航模块16仍可能识别出坐凳312处在不同于其平常位置314的非平常位置316。进一步并且仍基于用户的移动方向，导航模块16可确定在非平常位置316坐凳312在用户14的直接路径上。导航模块16可而确定在此情况中，坐凳312对用户14造成了绊倒危险。

[0044] 因此，导航模块16可将坐凳312分类成危险54并利用声源定位技术以警告耳标322的形式输出音频导航提示64，该提示警告用户坐凳312的非平常位置316。在这个示例中，坐凳312可被导航模块16分类成对象52和危险54两者。此外并且为了向用户14传达紧急的感觉，警告耳标322可用相比于熟悉耳标320更大的音量来输出，如图3中用具有比熟悉耳标320更大的警告耳标322描绘的。在其它示例中，警告耳标322可包括相比于熟悉耳标320而言不同的特殊声音。

[0045] 在某些实施例中，随着环境32中的特征50到用户14的邻近度减少，导航模块16可针对该特征以渐增的音量输出音频导航提示64。例如并现在参考图4，用户14可能第一次在宾馆400的休息室404中。因此，导航模块16可能操作在不熟悉导航模式28中。用户14可处在距雕塑414第一距离的初始位置410，并且可能正在朝着休息室404中央的雕塑走去。当用户14处在初始位置410时，导航模块16可以以第一音量输出雕塑耳标418，这可致使用户14察觉该雕塑耳标好像发自对应于该雕塑的3D位置420。随着用户14持续朝着雕塑414走并到达后续位置422，导航模块16可以以高于第一音量的第二音量输出雕塑耳标，如雕塑耳标418’所指示的，以警告用户他更接近雕塑414了。

[0046] 在不熟悉导航模式28的某些示例中，导航模块16可输出对应于用户14所处于的物理环境32中的每个特征50的音频导航提示64。例如，在图4中，导航模块16可输出对应于用户14的视野内的每个特征的音频导航提示64。当用户14处在初始位置410时，导航模块16可

输出对应于雕塑414、桌子430、沙发434、人438和墙壁422的音频导航提示64。

[0047] 在另一示例中，导航模块16可使用深度图像数据38和可见图像数据46的一个或多个来识别坐在沙发434上的人438的脸。导航模块可接着使用例如存储在远程服务器中的人脸识别数据库来将人的脸与人的身份相关联。导航模块可接着通知用户14人438的身份。有利地，用户14可接着预见该用户邻近的人。

[0048] 进一步且在某些示例中，身份耳标可被分派到人438的身份，身份耳标通过用户输入30提供给导航模块16或被导航模块16程序性地生成。在这些示例中，当导航模块16将脸与人438的身份相关联时，导航模块16可向用户输出身份耳标以通知用户此人的身份。

[0049] 现在参考图5，在另一个示例中，音频导航提示64可被定制以标识物理环境32中的斜坡。在一个示例中，用户14可接近上坡502。导航模块16可识别上坡502并可对应地向用户14输出包括指示上坡的多个升序定调的声音506的音频导航提示64。

[0050] 在另一个示例中并且现在参考图6，用户14可接近下坡602。导航模块16可识别下坡602并可对应地向用户14输出包括指示下坡的多个降序定调的声音606的音频导航提示64。

[0051] 再次参考图5，在某些示例中，导航模块16可将上坡502识别为一个或多个台阶510，诸如在一段楼梯中。作为响应，导航模块16可对应地向用户14输出包括多个离散升序定调的声音506的音频导航提示64，如各八分音符514所指示的。以此方式，导航模块16可将前方有向上台阶传递给用户15。以类似的方式，当导航模块16将下坡识别为一个或多个台阶时，导航模块16可对应地向用户14输出包括多个离散降序定调的声音的音频导航提示64。以此方式，导航模块16可将前方有向下台阶传递给用户15。

[0052] 在某些示例中，导航模块可标识上坡或下坡中的台阶数量，并可通过对应于台阶数量的离散定调声音的数量将台阶数量传递给用户14。例如，并且如在图5中所示的，导航模块16可识别上坡502中的三个台阶510。因此，导航模块16可以以三个八分音符514的形式向用户14输出三个离散定调的声音。有利地，以此方式，用户14可方便地被通知前方有三个向上台阶。

[0053] 再次参考图6，在某些示例中，导航模块16可将下坡602识别为包括具有降级形式的连续坡度，诸如，向下倾斜的人行道、街道、走廊等，其如图6所示可能是平坦的，或波状的或有点起伏的。作为响应，导航模块16可对应地向用户14输出包括多个连续降序定调的声音606的音频导航提示64，如图6中的连接四个八分音符的连奏线610所指示的。

[0054] 在某些示例中，向用户14输出以标识斜坡的音频导航提示64在熟悉导航模式26中与在不熟悉导航模式28中可不同。例如并再次参考图5，在导航模块16操作在熟悉导航模式26中时(诸如，当用户14在他家中时)指示台阶510的离散升序定调的声音506可以以第一较低音量输出。然而，当导航模块16操作在不熟悉导航模式28中时(诸如，例如，当用户14第一次在某一建筑物中时)，指示台阶510的离散升序定调的声音506可以以第二较大音量输出以向用户提供对楼梯的更多的注意。

[0055] 在某些示例中并且如上提到的，导航模块16可以被配置以通过一个或多个触觉换能器220来向用户14输出触觉导航提示。例如并还参考图4，用户14佩戴的头戴式计算设备200可在设备的右镜腿240上包括触觉换能器220，其中触觉换能器被配置来生成用户14感知的振动。当用户14在初始位置410时，并警告用户桌子430在用户的右边很近，导航模块16

可致使在头戴式计算设备200的右镜腿240上的触觉换能器220以预定方式振动以指示对象在附近并且在用户14的右边。

[0056] 参考图7，在某些示例中，导航模块16还可被配置来通过头戴式计算设备10的显示系统70向用户14显示视觉导航提示。在一个示例中，用户14可能具有包括限制用户看见远处对象的能力的高度近视的轻度缺陷。用户14可能沿着城市街道步行，具有通过图7所示的头戴式计算设备200的透明显示器204的视野。厢式货车704可能正朝着用户14开来。

[0057] 导航模块16可使用深度图像数据38来随着厢式货车704朝着用户14移动来确定其三维位置，并可确定该厢式货车在可能撞到用户14的路径上。因此，导航模块16可在透明显示器204上显示多个闪烁的警示图标708以警告用户接近的危险。在某示例中，导航模块16还可输出警告用户接近的危险的音频导航提示64。例如，音频导航提示64可包括一个或多个发出声音的警示和相关的指令712，诸如“警示!汽车快速接近!向你的右方移动!”。

[0058] 如上面所注意的，在一些示例中，头戴式计算设备10可包括显示系统70。其中头戴式计算设备10包括显示系统70，在某些示例中，头戴式计算设备10可输出如上所述的音频导航提示64而不通过显示系统70向用户显示图像或其它视觉内容。

[0059] 图8A和8B示出根据本公开的一实施例的用于在环境中通过音频输出向用户提供导航辅助的方法800的流程图。参考以上描述并在图1-7中示出的头戴式计算设备10的软件和硬件组件来提供方法800的以下描述。可以理解，方法800还可在使用其他合适的硬件和软件组件的其他上下文中来执行。

[0060] 参考图8A，在804，方法800可包括接收来自环境的深度图像数据。在808，方法800可包括接收来自环境的可见图像数据。在812，方法800可包括，使用该深度图像数据和可见图像数据，生成环境的至少一个部分的三维网格。在816，方法800可包括，使用该三维网格，检测环境中的至少一个特征。在820，方法800可包括，当头戴式计算设备在熟悉导航模式中操作时并基于检测到特征，通过一个或多个换能器将环境的第一音频导航提示输出给用户。

[0061] 在824，方法800可包括，当头戴式计算设备在不熟悉导航模式中操作时并基于检测到特征，通过一个或多个换能器将环境的第二音频导航提示输出给用户，其中第二音频导航提示不同于第一音频导航提示。在828，方法800可包括，以致使用户察觉提示好像源自特征的三维位置处的方式输出第一音频导航提示和第二音频导航提示的一个或多个。在832，特征可包括斜坡，并且第一音频导航提示和第二音频导航提示包括指示上坡的多个升序定调的声音或指示下坡的多个降序定调的声音。

[0062] 在836，第一音频导航提示和第二音频导航提示可包括指示连续斜坡的多个持续定调的声音或指示多个台阶的多个离散定调的声音。现在参考图8B，在840，方法800可包括以随着检测到的对象到用户的邻近度减少而增大的音量向用户输出第一音频导航提示和第二音频导航提示。在844，方法800可包括使用深度图像数据和可见图像数据的一个或多个来识别环境中的脸。在848，方法800可包括将脸与人的身份相关联。在852，方法800可包括通知用户此人的身份。

[0063] 在856，方法800可包括向此人的身份分派身份耳标。在860，方法800可包括通过一个或多个换能器向用户输出该身份耳标。在864，方法800可包括通过一个或多个换能器向用户输出触觉导航提示。在868，方法800可包括，基于检测到特征，通过头戴式计算设备的

显示系统向用户显示视觉导航提示，其中视觉导航提示是至少基于环境中的特征的三维位置的。

[0064] 能够理解，方法800是以举例方式提供的，并且不旨在为限制性的。因此，可以理解，方法800可包括相比于图8A和8B中示出的那些步骤而言附加的和/或替换的步骤。并且，可以理解，方法800可用任何适当的次序执行。而且，可以理解，一个或多个步骤可从方法800中省略，而不背离本发明的范围。

[0065] 图9示意性示出了可以执行上述方法和过程之中的一个或更多个的计算系统900的非限制性实施例。头戴式计算设备10可采取计算系统900的形式或包括计算系统900的一个或多个方面。以简化形式示出了计算系统900。应当理解，可使用基本上任何计算机架构而不背离本公开的范围。在不同的实施例中，计算系统900可以采取大型计算机、服务器计算机、台式计算机、膝上型计算机、平板计算机、家庭娱乐计算机、网络计算设备、移动计算设备、移动通信设备、游戏设备等等的形式。

[0066] 如图9所示，计算系统900包括逻辑子系统904、存储子系统908、以及传感器子系统912。计算系统900可任选地包括显示子系统916、通信子系统920、输入子系统922和/或在图9中未示出的其他子系统和组件。计算系统900还可包括计算机可读介质，其中该计算机可读介质包括计算机可读存储介质和计算机可读通信介质。计算系统900还可以任选地包括其他用户输入设备，诸如例如键盘、鼠标、游戏控制器，和/或触摸屏等等。此外，在某些实施例中，此处所述的方法和过程可被实现为计算机应用、计算机服务、计算机API、计算机库，和/或包括一个或多个计算机的计算系统中的其他计算机程序产品。

[0067] 逻辑子系统904可包括被配置为执行一个或多个指令的一个或多个物理设备。例如，逻辑子系统904可被配置为执行一个或多个指令，该一个或多个指令是一个或多个应用、服务、程序、例程、库、对象、组件、数据结构、或其他逻辑构造的一部分。可实现这样的指令以执行任务、实现数据类型、变换一个或多个设备的状态、或以其他方式得到所希望的结果。

[0068] 逻辑子系统904可包括被配置成执行软件指令的一个或多个处理器。附加地或可替代地，逻辑子系统可以包括被配置为执行硬件或固件指令的一个或多个硬件或固件逻辑机器。逻辑子系统的处理器可以是单核或多核，且在其上执行的程序可被配置为并行或分布式处理。逻辑子系统可以任选地包括遍布两个或更多设备分布的独立组件，所述设备可远程放置和/或被配置为进行协同处理。该逻辑子系统的一个或多个方面可被虚拟化并由以云计算配置进行配置的可远程访问的联网计算设备执行。

[0069] 存储子系统908可包括被配置为保持可由逻辑子系统904执行以实现此处所述的方法和过程的数据和/或指令的一个或多个物理持久设备。在实现此类方法和过程时，存储子系统908的状态可以被变换(例如，以保持不同的数据)。

[0070] 存储子系统908可以包括可移动介质和/或内置设备。存储子系统908可包括光学存储设备(例如，CD、DVD、HD-DVD、蓝光盘等)、半导体存储器设备(例如，RAM、EPROM、EEPROM等)和/或磁性存储设备(例如，硬盘驱动器、软盘驱动器、磁带驱动器、MRAM等)等等。存储子系统908可包括具有以下特性中的一个或多个特性的设备：易失性、非易失性、动态、静态、读/写、只读、随机存取、顺序存取、位置可寻址、文件可寻址，以及内容可寻址。

[0071] 在一些实施例中，可以将逻辑子系统904和存储子系统908的各方面集成在一个或

多个共同设备中，通过该一个或多个共同设备，可以至少部分地实施本文所述的功能。这样的硬件逻辑组件可包括：例如，现场可编程门阵列(FPGA)、程序和应用专用集成电路(PASIC/ASIC)、程序和应用专用标准产品(PSSP/ASSP)、片上系统(SOC)系统以及复杂可编程逻辑设备(CPLD)。

[0072] 图9还示出以可移动计算机可读存储介质924形式的存储子系统908的一方面，该介质可以用于存储可执行以实现此处所述的方法和过程的数据和/或指令。可移动计算机可读存储介质924尤其是可以采取CD、DVD、HD-DVD、蓝光盘、EEPROM和/或软盘的形式。

[0073] 将明白，存储子系统908包括一个或多个物理持久设备。相反，在一些实施例中，本文描述的指令的各方面可以按暂态方式通过不由物理设备在至少有限持续时间期间保持的纯信号(例如电磁信号、光信号等)传播。此外，与本公开有关的数据和/或其他形式的信息可以经由计算机可读通信介质通过纯信号来传播。

[0074] 传感器子系统912可包括被配置成感测不同的物理现象(例如，可见光、红外光、声音、加速度、取向、位置等)的一个或多个传感器，如上所述。传感器子系统912例如可以被配置为向逻辑子系统904提供传感器数据。如上所述，此类数据可包括图像信息、环境光信息、深度信息、音频信息、位置信息、运动信息，用户位置信息，和/或可被用来执行上述方法和过程的任何其他合适的传感器数据。

[0075] 在被包括时，显示子系统916可用于呈现由存储子系统908保存的数据的视觉表示。由于以上所描述的方法和过程改变了由存储子系统908保持的数据，并由此变换了存储子系统的状态，因此同样可以转变显示子系统916的状态以在视觉上表示底层数据的改变。显示子系统916可包括利用几乎任何类型的技术的一个或多个显示设备。可以将此类显示设备与逻辑子系统904和/或存储子系统908一起组合在共享封装中，或者此类显示设备可以是外围触摸显示设备。显示子系统916可包括，例如，头戴式计算设备10的可选显示系统70和显示器程序72。

[0076] 在被包括时，通信子系统920可以被配置成将计算系统900与一个或多个网络和/或一个或多个其他计算设备可通信地耦合。通信子系统920可以包括与一个或多个不同通信协议兼容的有线和/或无线通信设备。作为非限制性示例，通信子系统920可被配置为经由无线电话网、无线局域网、有线局域网、无线广域网、有线广域网等进行通信。在一些实施例中，通信子系统可允许计算系统900经由诸如因特网之类的网络发送消息至其他设备和/或从其他设备接收消息。

[0077] 在被包括时，输入子系统922可包括一个或多个传感器或用户输入设备(诸如游戏控制器、姿势输入检测设备、语音识别器、惯性测量单元、键盘、鼠标、或触摸屏)或与它们对接。在某些实施例中，输入子系统922可以包括所选的自然用户输入(NUI)部件或与其结合。这种元件部分可以是集成的或外围的，输入动作的转导和/或处理可以在板上或板外被处理。NUI部件的示例可包括用于语言和/或语音识别的话筒；用于机器视觉和/或姿势识别的红外、色彩、立体显示和/或深度相机；用于运动检测和/或意图识别的头部跟踪器、眼睛跟踪器、加速计和/或陀螺仪；以及用于评估脑部活动的电场感测部件。

[0078] 术语“模块”和“程序”可被用于描述被实现以执行一个或更多个具体功能的头戴式计算设备10的一方面。在某些情况下，可以通过执行由存储子系统908所保持的指令的逻辑子系统904来实例化此类模块或程序。可以理解，可以从同一应用、服务、代码块、对象、

库、例程、API、函数等来实例化不同的模块和程序。同样，可以由不同的应用程序、服务、代码块、对象、例程、API、函数等实例化同一模块或程序。术语“模块”和“程序”意在涵盖单个或成组的可执行文件、数据文件、库、驱动程序、脚本、数据库记录等。

[0079] 应该理解，此处所述的配置和/或方法在本质上是示例性的，并且这些具体实施例或示例不应被认为是局限性的，因为多个变体是可能的。此处描述的具体例程或方法可以表示任何数量的处理策略中的一个或多个。由此，所示出的各个动作可以按所示次序执行、按其他次序执行、并行地执行，或者在某些情况下被省略。同样，上述过程的次序可以改变。

[0080] 本公开的主题包括各种过程、系统和配置以及此处公开的其它特征、功能、动作和/或属性、以及它们的任一和全部等价物的所有新颖且非显而易见的组合和子组合。

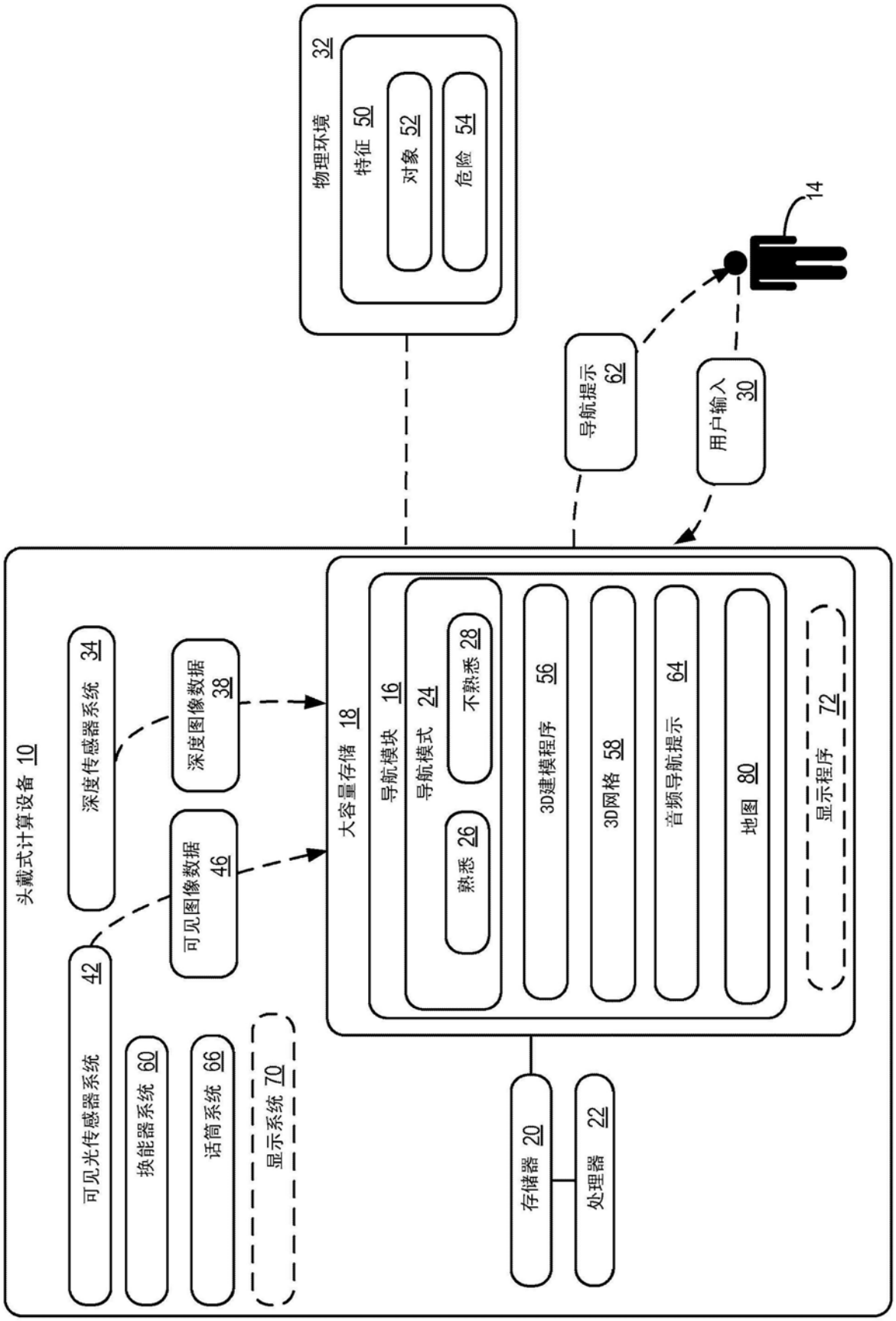

物理环境 32
特征 50
对象 52
危险 54
14
导航提示 62
用户输入 30
头戴式计算设备 10
深度传感器系统 34
深度图像数据 38
可见图像数据 46
可见光传感器系统 42
换能器系统 60
话筒系统 66
显示系统 70
大容量存储 18
导航模块 16
导航模式 24
熟悉 26
不熟悉 28
3D建模程序 56
3D网格 58
音频导航提示 64
地图 80
显示程序 72
存储器 20
处理器 22

图1

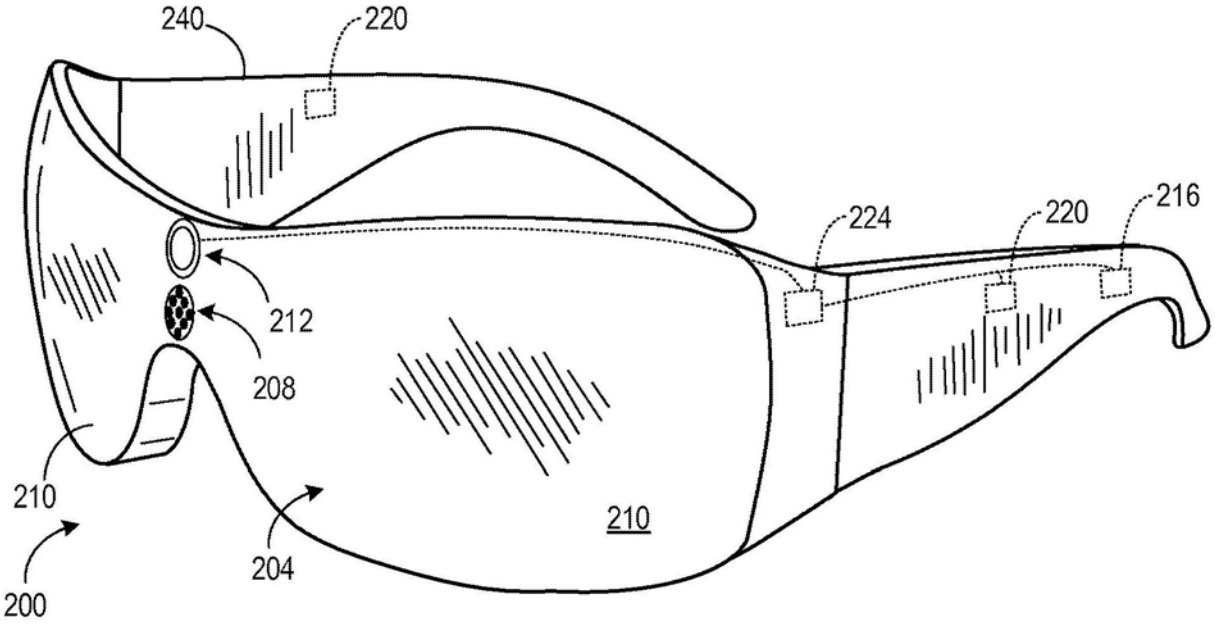


图2

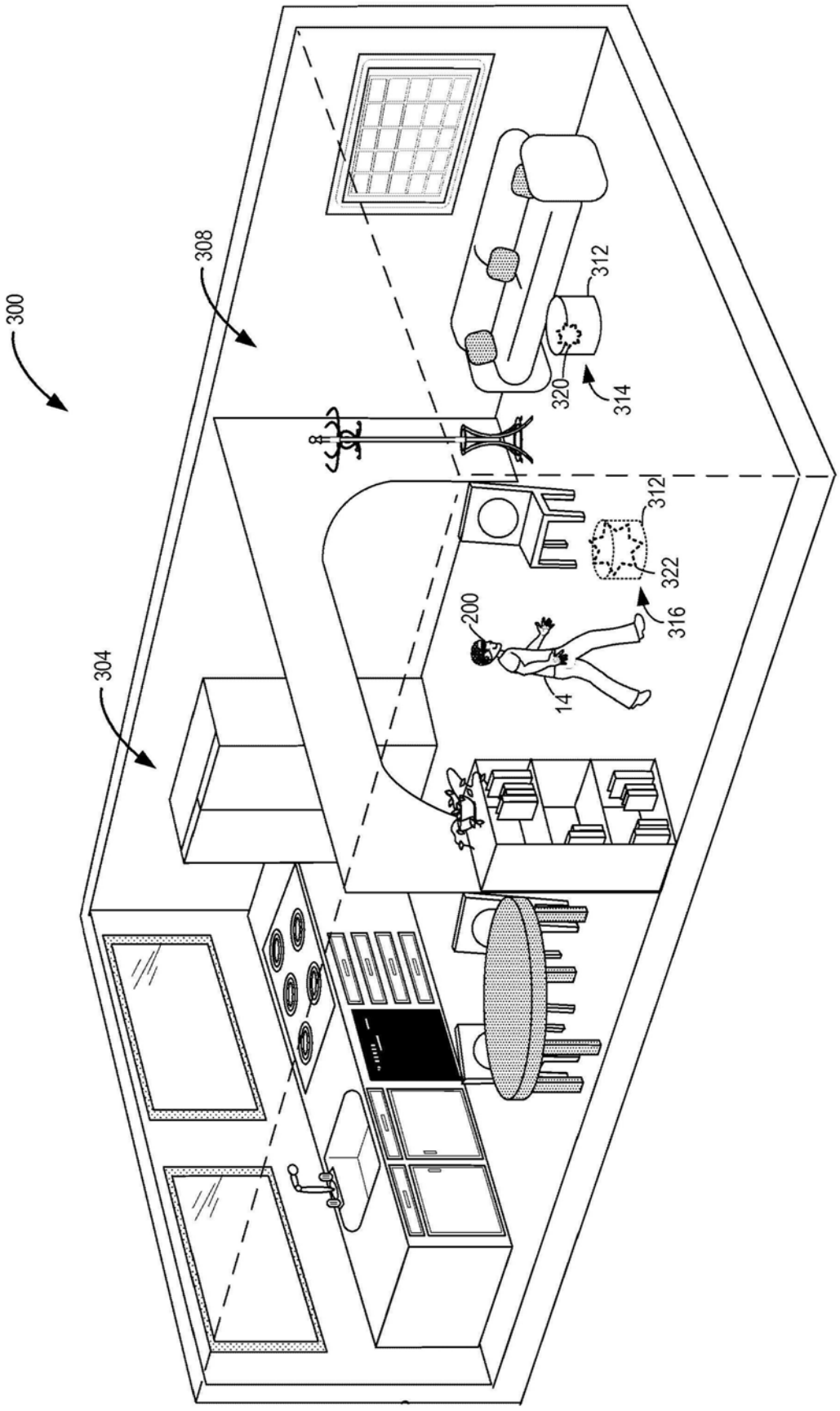


图3

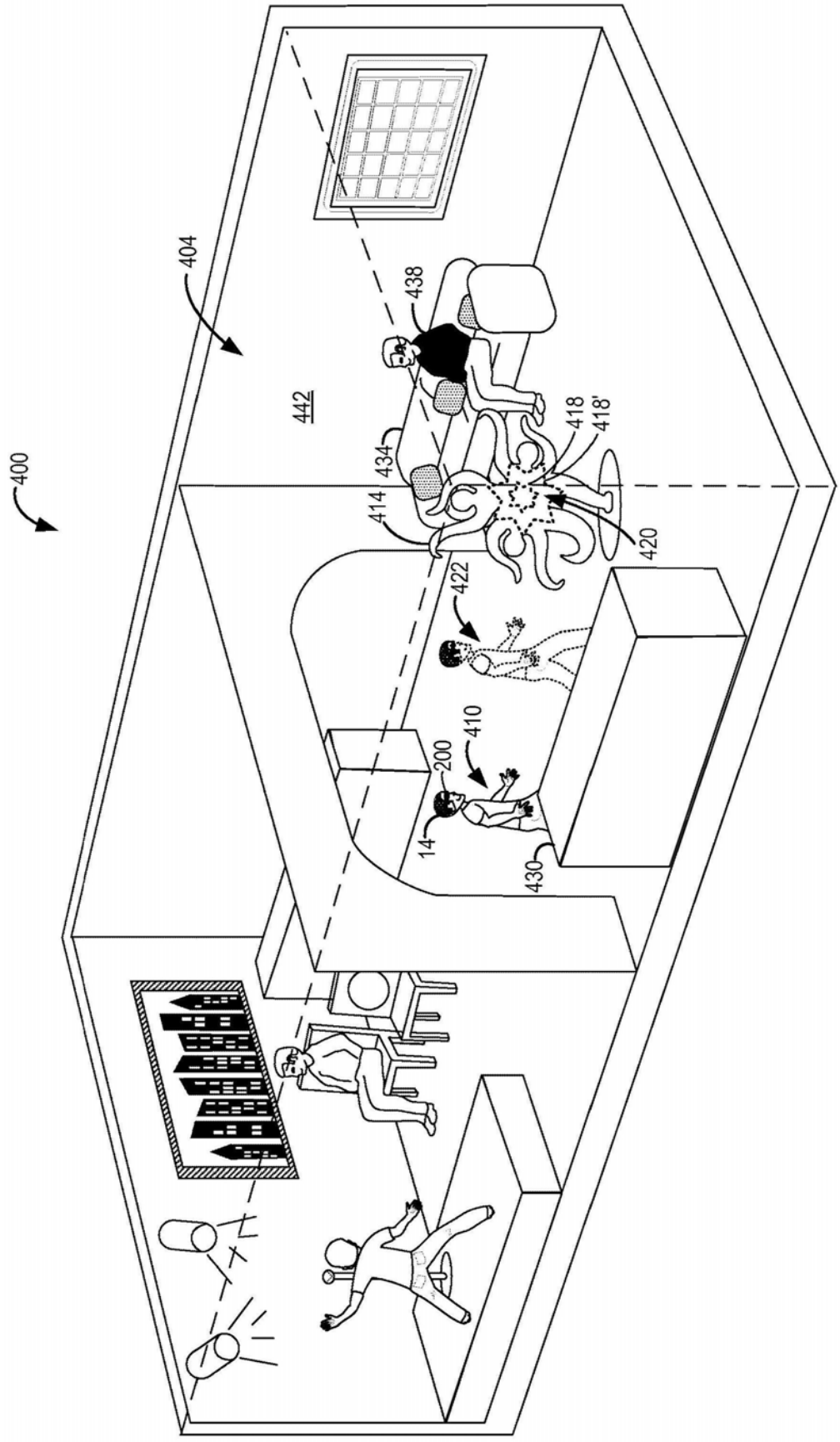


图4

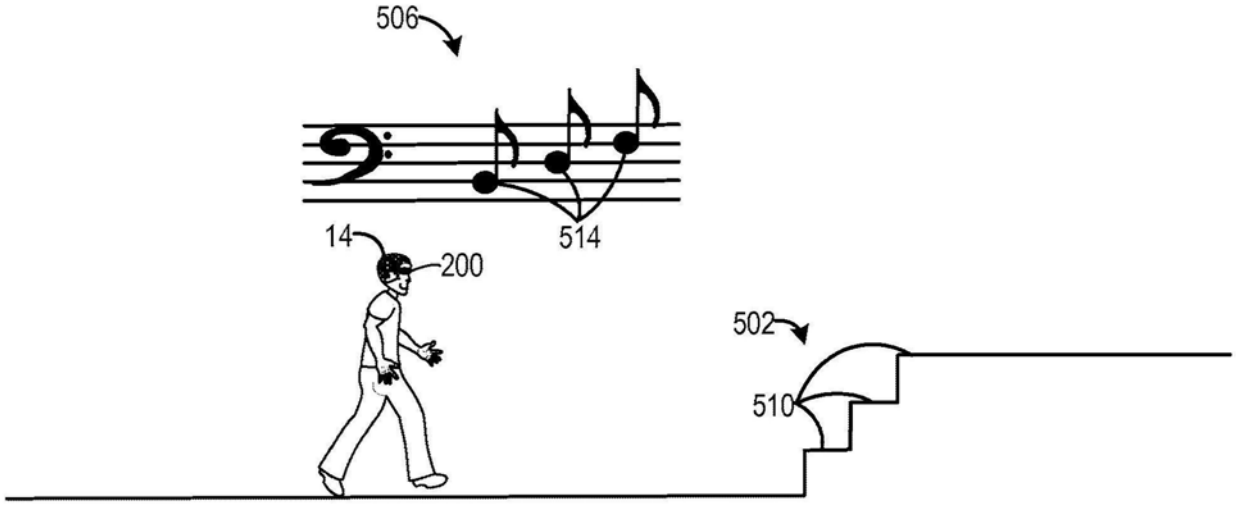


图5

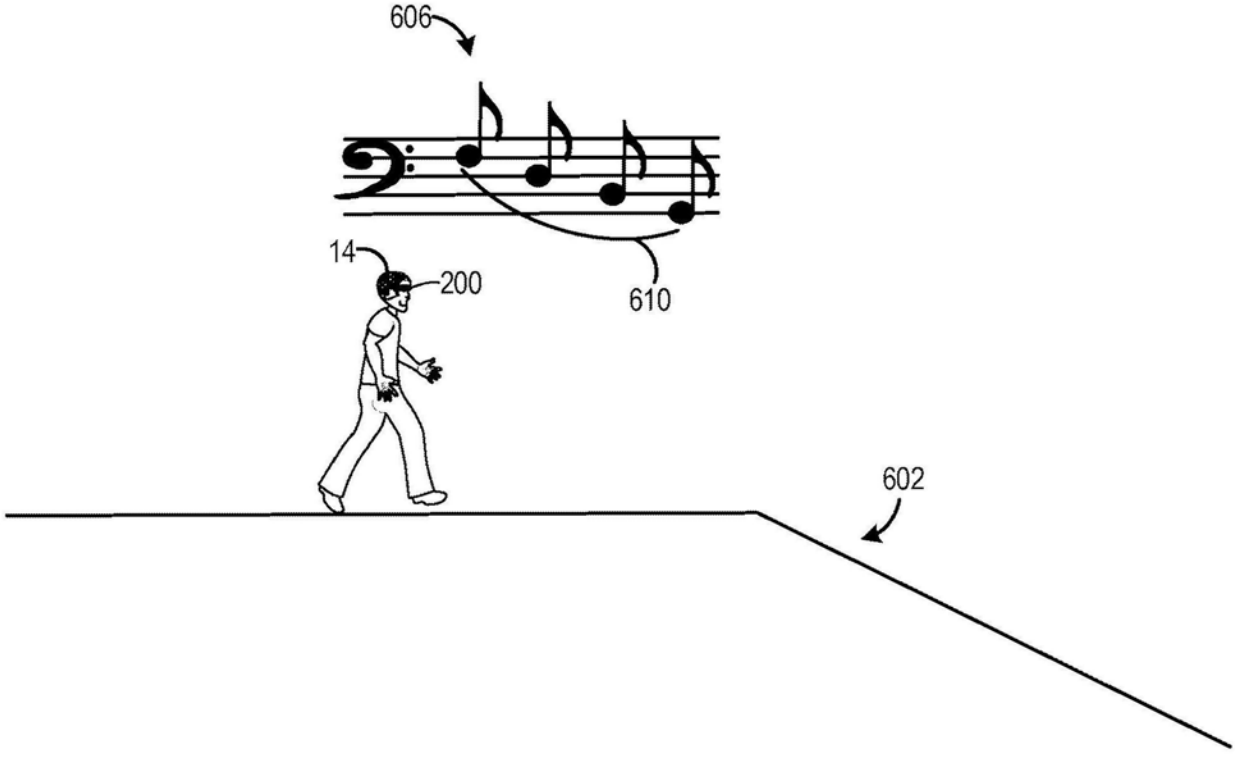


图6

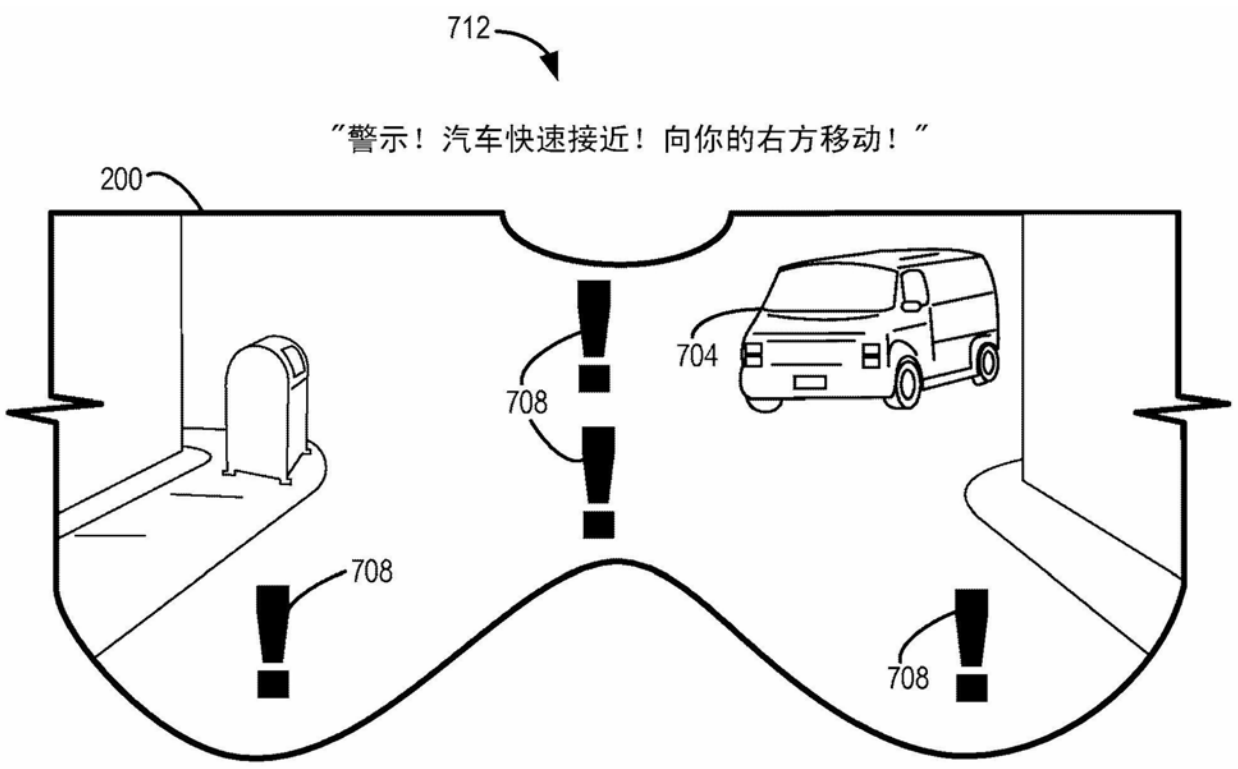


图7

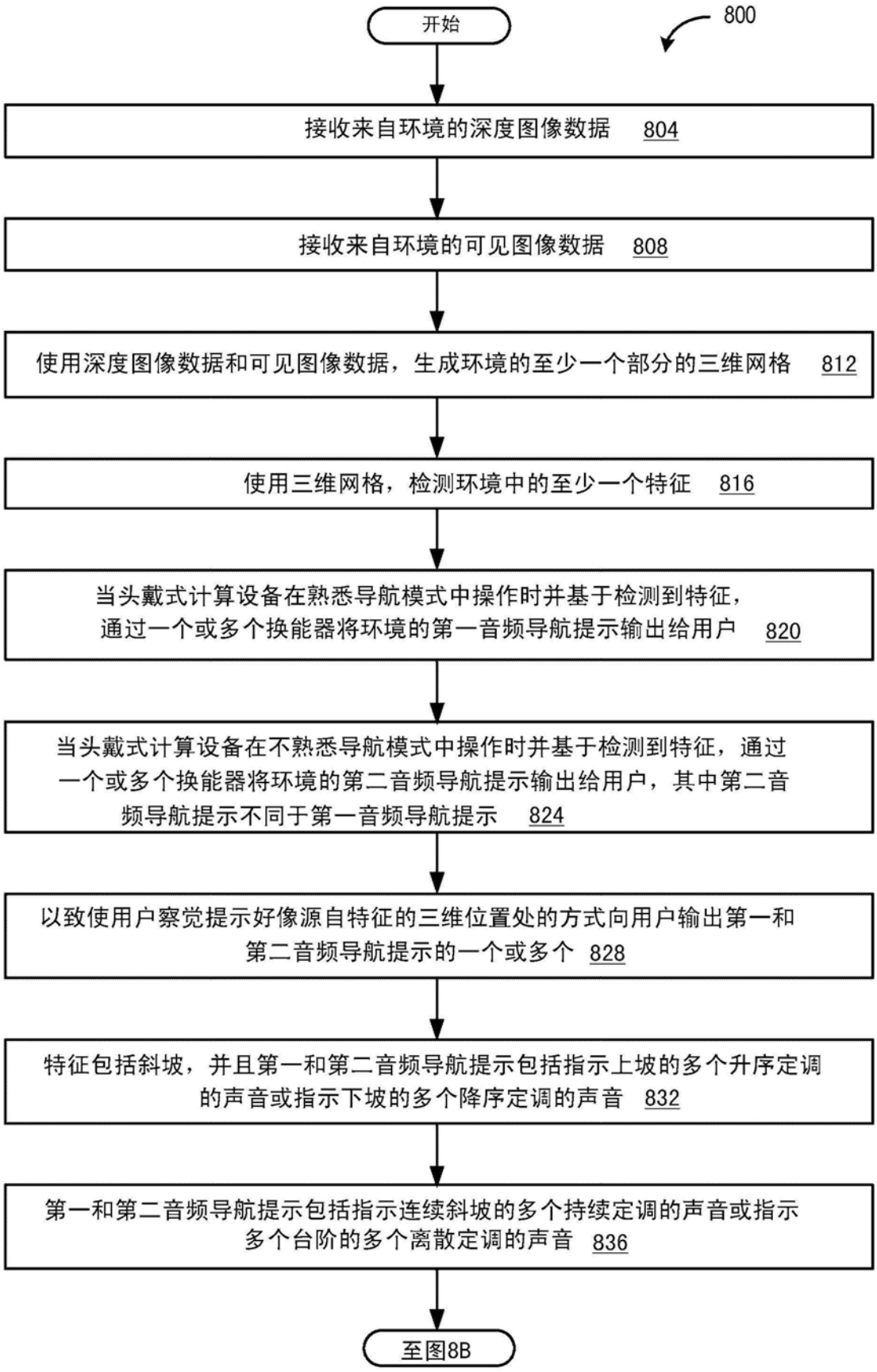

开始
800
接收来自环境的深度图像数据 804
接收来自环境的可见图像数据 808
使用深度图像数据和可见图像数据，生成环境的至少一个部分的三维网格 812
使用三维网格，检测环境中的至少一个特征 816
当头戴式计算设备在熟悉导航模式中操作时并基于检测到特征，通过一个或多个换能器将环境的第一音频导航提示输出给用户 820
当头戴式计算设备在不熟悉导航模式中操作时并基于检测到特征，通过一个或多个换能器将环境的第二音频导航提示输出给用户，其中第二音频导航提示不同于第一音频导航提示 824
以致使用户察觉提示好像源自特征的三维位置处的方式向用户输出第一和第二音频导航提示的一个或多个 828
特征包括斜坡，并且第一和第二音频导航提示包括指示上坡的多个升序定调的声音或指示下坡的多个降序定调的声音 832
第一和第二音频导航提示包括指示连续斜坡的多个持续定调的声音或指示多个台阶的多个离散定调的声音 836
至图8B

图8A

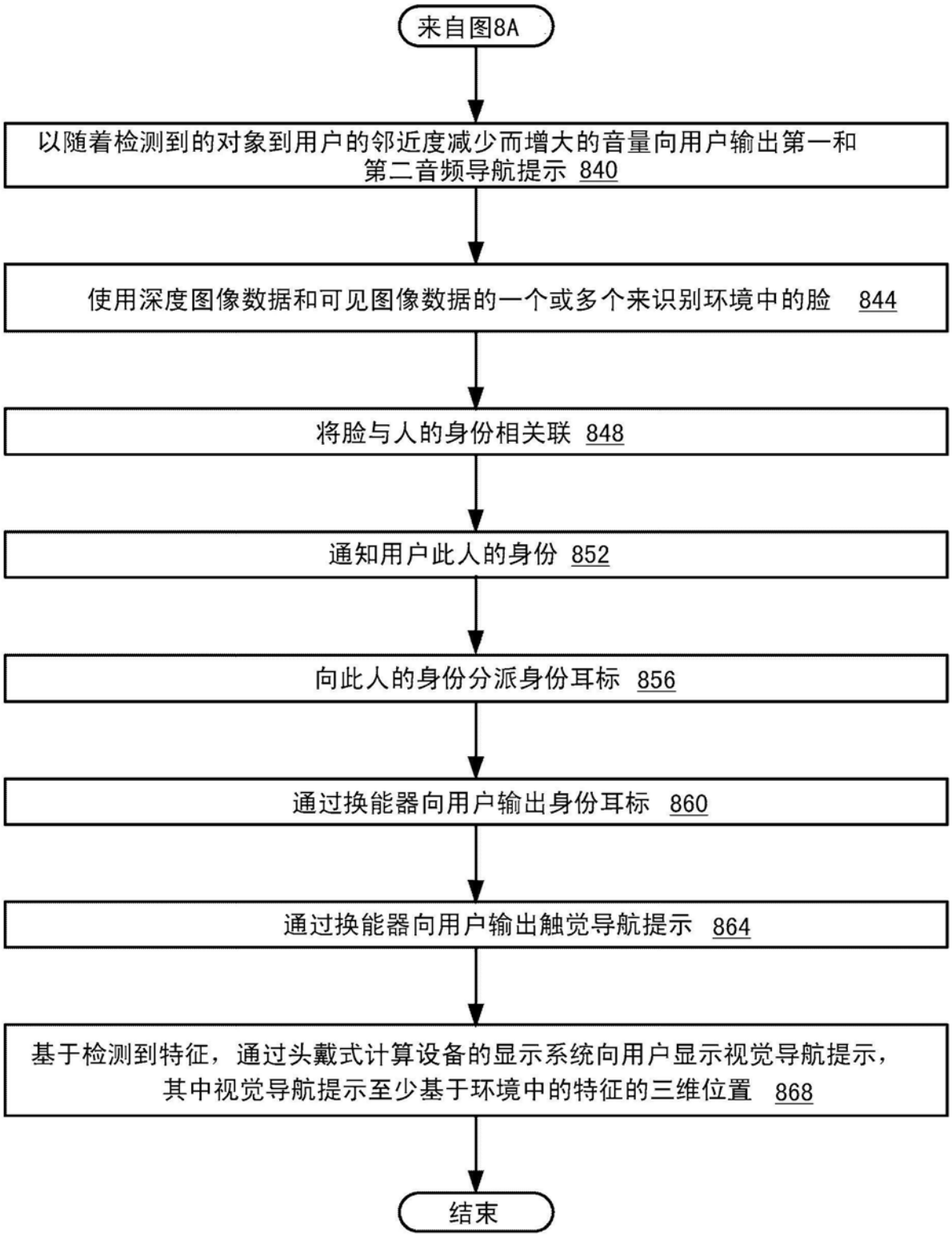

来自图8A
以随着检测到的对象到用户的邻近度减少而增大的音量向用户输出第一和第二音频导航提示 840
使用深度图像数据和可见图像数据的一个或多个来识别环境中的脸 844
将脸与人的身份相关联 848
通知用户此人的身份 852
向此人的身份分派身份耳标 856
通过换能器向用户输出身份耳标 860
通过换能器向用户输出触觉导航提示 864
基于检测到特征，通过头戴式计算设备的显示系统向用户显示视觉导航提示，其中视觉导航提示至少基于环境中的特征的三维位置 868
结束

图8B

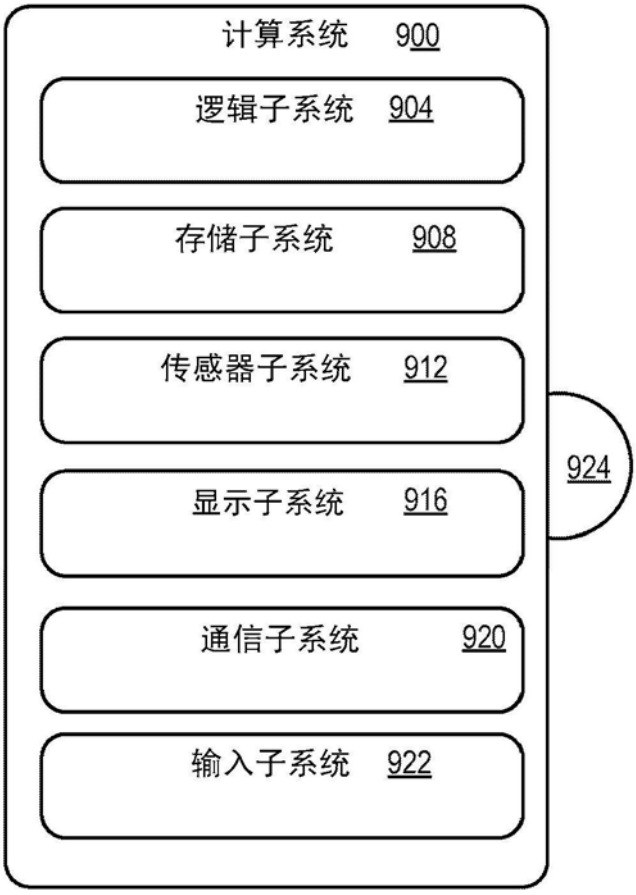

计算系统 900
逻辑子系统 904
存储子系统 908
传感器子系统 912
显示子系统 916
924
通信子系统 920
输入子系统 922

图9