



(12) 发明专利

(10) 授权公告号 CN 114026606 B

(45) 授权公告日 2025.06.17

(21) 申请号 202080046371.7

(22) 申请日 2020.06.25

(65) 同一申请的已公布的文献号
申请公布号 CN 114026606 A

(43) 申请公布日 2022.02.08

(30) 优先权数据
62/868,640 2019.06.28 US

(85) PCT国际申请进入国家阶段日
2021.12.24

(86) PCT国际申请的申请数据
PCT/US2020/039601 2020.06.25

(87) PCT国际申请的公布数据
W02020/264149 EN 2020.12.30

(73) 专利权人 奇跃公司

地址 美国佛罗里达州

(72) 发明人 D·G·莫利纳

(74) 专利代理机构 北京市中咨律师事务所
11247

专利代理师 于静 牛南辉

(51) Int.Cl.
G06F 17/00 (2019.01)

(56) 对比文件
CN 103472909 A, 2013.12.25
CN 109471521 A, 2019.03.15

审查员 张琳

权利要求书2页 说明书17页 附图12页

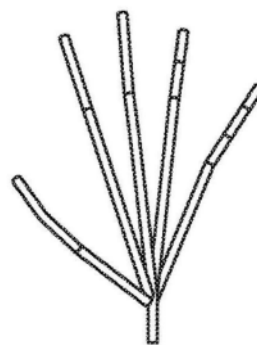
(54) 发明名称

用于动态遮挡的快速手部网格划分

(57) 摘要

一种提供沉浸式用户体验的增强现实/混合现实系统。以更高的更新速度、较少计算资源用于表示用户手部的遮挡数据来提供这种体验,遮挡数据可以用于遮挡虚拟对象。手部遮挡数据是通过根据实况深度传感器数据计算手部网格来生成的。通过使用其中手部简单地由从关键点识别出的多个分段表示的模型对实况深度数据进行掩蔽来处理少量数据。手部遮挡数据可以通过预测捕获深度数据的时间和手部网格将用于遮挡处理的时间之间的手部姿势的变化来生成。手部网格可以变形以表示预测姿势中的手部。

22个关键点



1. 一种操作计算系统以重建用于动态遮挡虚拟对象的手部的方法,所述方法包括:
从在场景中渲染虚拟对象的应用接收针对与所述场景中的手部相关的数据的查询;
从用户佩戴的设备获取所述场景的信息,所述设备包括一个或多个传感器,所述场景的所述信息包括指示所述用户佩戴的所述设备与所述场景中的物理对象之间的距离的深度信息;

检测所述场景中的所述物理对象是否包括手部;

当检测到所述手部时,至少部分地基于所述场景的所述信息来计算所述手部的模型;

用所述手部的所述模型来掩蔽指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息;

基于被掩蔽到所述手部的所述模型的所述深度信息,计算手部网格,所述计算包括:随着所述设备与所述手部之间的相对位置改变,实时更新所述手部网格;以及

将所述手部网格提供给所述应用,以使得所述应用渲染所述虚拟对象的未被所述手部网格遮挡的部分。

2. 根据权利要求1所述的方法,其中:

所述手部的所述模型包括指示所述手部的分段上的点的所述手部的多个关键点。

3. 根据权利要求2所述的方法,其中:

所述手部的所述多个关键点的至少一部分对应于所述手部的关节和所述手部的指尖。

4. 根据权利要求2所述的方法,其中:

所述方法还包括:基于所述多个关键点来确定所述手部的轮廓;以及

掩蔽所述深度信息包括:

滤除在所述手部的所述模型的所述轮廓之外的所述深度信息;以及

至少部分地基于经过滤的深度信息,生成所述手部的深度图像,所述深度图像包括多个像素,每个像素指示到所述手部的点的距离。

5. 根据权利要求4所述的方法,其中,滤除在所述手部的所述模型的所述轮廓之外的所述深度信息包括:

移除与所述场景中的所述物理对象相关联的深度信息。

6. 根据权利要求2所述的方法,其中:

用所述手部的所述模型来掩蔽指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息包括:将深度图像的部分关联到手部分段;以及

实时更新所述手部网格包括:选择性地更新表示所述手部分段的适当子集的所述手部网格的部分。

7. 根据权利要求6所述的方法,还包括:

在计算所述手部网格之前,填充所述深度图像中的空洞。

8. 根据权利要求7所述的方法,其中,填充所述深度图像中的空洞包括:

从所述设备的立体相机生成立体深度信息,所述立体深度信息对应于所述深度图像中的所述空洞的区域。

9. 根据权利要求7所述的方法,其中,填充所述深度图像中的空洞包括:

从手部的3D模型访问表面信息,所述表面信息对应于所述深度图像中的所述空洞的区域。

10. 根据权利要求1所述的方法,其中,基于被掩蔽到所述手部的所述模型的所述深度信息来计算所述手部网格包括:

根据在时间 t 从在所述场景中渲染所述虚拟对象的所述应用接收的针对与所述场景中的手部相关的所述数据的查询,预测延时 n ;

预测在查询时间 t 加上所述延时 n 的时间的手部姿势;以及

用所预测的在所述查询时间 t 加上所述延时 n 的所述时间的姿势来变形所述手部网格。

11. 根据权利要求1所述的方法,其中,指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息包括以至少每秒30帧的帧速率的深度图像序列。

12. 一种能够由用户携带的电子系统,包括:

由所述用户佩戴的设备,其中,所述设备包括被配置为渲染虚拟对象的显示器并且包括一个或多个传感器,所述一个或多个传感器被配置为捕获佩戴所述设备的所述用户的头部姿势和包括一个或多个物理对象的场景的信息,所述场景的信息包括指示所述设备与所述一个或多个物理对象之间的距离的深度信息;

手部网格划分部件,其被配置为执行计算机可执行指令以检测所述场景中的手部,并计算所检测到的手部的所述手部网格,并随着所述头部姿势改变和/或所述手部移动来实时地更新所述手部网格;以及

应用,其被配置为执行计算机可执行指令以在所述场景中渲染所述虚拟对象,其中,所述应用从所述手部网格划分部件接收所述手部网格和所述虚拟对象的被所述手部遮挡的部分。

13. 根据权利要求12所述的电子系统,其中:

所述手部网格划分部件被配置为通过以下方式来计算手部网格:

识别所述手部上的关键点;

计算所述关键点之间的分段;

基于与一个或多个所计算的分段的接近度,从所述深度信息中选择信息;以及

基于所选择的深度信息,计算表示所述手部网格的至少一部分的网格。

14. 根据权利要求13所述的电子系统,其中:

所述深度信息包括多个像素,所述多个像素中的每个像素表示到所述场景中的对象的距离;以及

计算所述网格包括:对表示距离差小于阈值的相邻像素进行分组。

用于动态遮挡的快速手部网格划分

技术领域

[0001] 本申请一般涉及使用3D世界重建来渲染场景的交叉现实系统。

背景技术

[0002] 计算机可以控制人类用户界面以创建X Reality (X现实, XR或交叉现实) 环境, 在该环境中, 由计算机生成用户所感知的一些或全部XR环境。这些XR环境可以是虚拟现实 (VR)、增强现实 (AR) 和混合现实 (MR) 环境, 其中的一些或所有XR环境可以由计算机部分使用描述环境的数据来生成。例如, 该数据可以描述虚拟对象, 该虚拟对象可以以用户的感知或感知为物理世界的一部分的方式渲染, 并且可以与虚拟对象进行交互。由于数据是通过用户接口设备 (诸如例如头戴式显示设备) 渲染和呈现的, 因此用户可以体验这些虚拟对象。数据可以显示给用户看, 或者可以控制被播放给用户听的音频, 或者可以控制触觉 (或触知) 界面, 从而使用户能够体验到用户感觉或感知为感觉到虚拟对象的触摸感觉。

[0003] XR系统可以被用于跨科学可视化、医学培训、工程设计和原型、远程操纵和远程呈现以及个人娱乐领域的许多应用。与VR相比, AR和MR包括与物理世界的真实对象相关的一个或多个虚拟对象。虚拟对象与真实对象交互的体验显著地增强了用户使用XR系统的乐趣, 并且也为呈现有关如何改变物理世界的现实且易于理解的信息的各种应用打开了大门。

[0004] XR系统可以将系统的用户周围的物理世界表示为“网格”。网格可以由多个相互连接的三角形表示。每个三角形具有连接物理世界内对象的表面上的点的边, 使得每个三角形表示该表面的一部分。关于表面的该部分的信息, 例如颜色、纹理或其他属性可以相关联地存储在三角形内。在操作中, XR系统可以处理图像信息以检测点和面, 从而创建或更新网格。

发明内容

[0005] 本申请的各方面涉及用于动态遮挡的快速手部网格划分 (meshing) 的方法和设备。在此描述的技术可以一起使用、单独使用或以任何合适的组合使用。

[0006] 一些实施例涉及一种操作计算系统以重建用于动态遮挡虚拟对象的手部的方法。所述方法包括: 从在场景中渲染虚拟对象的应用接收针对与所述场景中的手部相关的数据的查询; 从用户佩戴的设备获取所述场景的信息, 所述设备包括一个或多个传感器, 所述场景的所述信息包括指示所述用户佩戴的所述设备与所述场景中的物理对象之间的距离的深度信息; 检测所述场景中的所述物理对象是否包括手部; 当检测到所述手部时, 至少部分地基于所述场景的所述信息来计算所述手部的模型; 用所述手部的所述模型来掩蔽指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息; 基于被掩蔽到所述手部的所述模型的所述深度信息, 计算手部网格, 所述计算包括: 随着所述设备与所述手部之间的相对位置改变, 实时更新所述手部网格; 以及将所述手部网格提供给所述应用, 以使得所述应用渲染所述虚拟对象的未被所述手部网格遮挡的部分。

[0007] 在一些实施例中,所述手部的所述模型包括指示所述手部的分段上的点的所述手部的多个关键点。

[0008] 在一些实施例中,所述手部的所述多个关键点的至少一部分对应于所述手部的关节和所述手部的指尖。

[0009] 在一些实施例中,所述方法还包括:基于所述多个关键点来确定所述手部的轮廓;以及用所述手部的模型来掩蔽指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息。掩蔽所述深度信息包括:滤除在所述手部的所述模型的所述轮廓之外的所述深度信息;以及至少部分地基于经过滤的深度信息,生成所述手部的深度图像,所述深度图像包括多个像素,每个像素指示到所述手部的点的距离。

[0010] 在一些实施例中,滤除在所述手部的所述模型的所述轮廓之外的所述深度信息包括:移除与所述场景中的所述物理对象相关联的深度信息。

[0011] 在一些实施例中,用所述手部的所述模型来掩蔽指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息包括:将所述深度图像的部分关联到手部分段;以及实时更新所述手部网格包括:选择性地更新表示所述手部分段的适当子集的所述手部网格的部分。

[0012] 在一些实施例中,所述方法还包括:在计算所述手部网格之前,填充所述深度图像中的空洞。

[0013] 在一些实施例中,填充所述深度图像中的空洞包括:从所述设备的立体相机生成立体深度信息,所述立体深度信息对应于所述深度图像中的所述空洞的区域。

[0014] 在一些实施例中,填充所述深度图像中的空洞包括:从手部的3D模型访问表面信息,所述表面信息对应于所述深度图像中的所述空洞的区域。

[0015] 在一些实施例中,基于被掩蔽到所述手部的所述模型的所述深度信息来计算所述手部网格包括:根据在时间 t 从在所述场景中渲染所述虚拟对象的所述应用接收的针对与所述场景中的手部相关的所述数据的查询,预测延时 n ;预测在查询时间 t 加上所述延时 n 的时间的手部姿势;以及用所预测的在所述查询时间 t 加上所述延时 n 的所述时间的姿势来变形所述手部网格。

[0016] 在一些实施例中,指示所述用户佩戴的所述设备与所述场景中的所述物理对象之间的所述距离的所述深度信息包括以至少每秒30帧的帧速率的深度图像序列。

[0017] 一些实施例涉及一种能够由用户携带的电子系统。所述电子系统包括由所述用户佩戴的设备。所述设备包括被配置为渲染虚拟对象的显示器并且包括一个或多个传感器,所述一个或多个传感器被配置为捕获佩戴所述设备的所述用户的头部姿势和包括一个或多个物理对象的场景的信息,所述场景的信息包括指示所述设备与所述一个或多个物理对象之间的距离的深度信息。所述电子系统包括:手部网格划分部件,其被配置为执行计算机可执行指令以检测所述场景中的手部,并计算所检测到的手部的所述手部网格,并随着所述头部姿势改变和/或所述手部移动来实时地更新所述手部网格;以及应用,其被配置为执行计算机可执行指令以在所述场景中渲染所述虚拟对象,其中,所述应用从所述手部网格划分部件接收所述手部网格和所述虚拟对象的被所述手部遮挡的部分。

[0018] 在一些实施例中,所述手部网格划分部件被配置为通过以下方式来计算手部网格:识别所述手部上的关键点;计算所述关键点之间的分段;基于与一个或多个所计算的分

段的接近度,从所述深度信息中选择信息;以及基于所选择的深度信息,计算表示所述手部网格的至少一部分的网格。

[0019] 在一些实施例中,所述深度信息包括多个像素,所述多个像素中的每个像素表示到所述场景中的对象的距离。计算所述网格包括:对表示距离差小于阈值的相邻像素进行分组。

[0020] 一些实施例涉及一种操作AR系统以在包括物理对象的场景中渲染虚拟对象的方法。所述AR系统包括至少一个传感器和至少一个处理器。所述方法包括:用所述至少一个传感器捕获场景的信息,所述场景的所述信息包括指示到所述场景中的物理对象的距离的深度信息;用至少一个处理器:处理所捕获的信息,以检测所述场景中的手部并计算所述手部上的点;基于与所述手部上的所计算的点的接近度,选择所述深度信息的子集;以及基于所选择的深度信息来计算所述手部的表示,其中,所述手部的所述表示指示所述手部的表面。

[0021] 在一些实施例中,所述方法还包括:存储所计算的所述手部的表示;以及连续处理所捕获的信息以更新所存储的所述手部的表示。

[0022] 在一些实施例中,计算所述手部的所述表示包括:基于所捕获的信息,计算所述手部的一个或多个运动参数;基于所述一个或多个运动参数,推算所述手部在未来时间的位置,所述未来时间是基于与使用所计算的所述手部的表示来渲染虚拟对象相关联的延时而确定的;以及使所计算的所述手部的表示变形以表示推算位置的所述手部。

[0023] 在一些实施例中,所述方法还包括:基于所述手部的所述表示,渲染所述虚拟对象的所选部分,其中,所选部分表示所述虚拟对象的未被所述手部遮挡的部分。

[0024] 在一些实施例中,所述深度信息包括深度图,所述深度图包括多个像素,每个像素表示距离。基于所选的深度信息来计算所述手部的所述表示包括:识别表示表面分段的像素组。

[0025] 在一些实施例中,计算所述手部的所述表示包括:基于所识别的像素组,定义表示所述手部的网格。

[0026] 在一些实施例中,定义所述网格包括:识别与所识别的表面分段对应的三角形区域。

[0027] 前述概述是通过说明的方式提供的,并非旨在进行限制。

附图说明

[0028] 附图并非旨在按比例绘制。在附图中,在各个附图中示出的每个相同或几乎相同的部件由相似的数字表示。为了清楚起见,并非每个部件都可以在每个附图中标记。在附图中:

[0029] 图1是示出根据一些实施例的简化的增强现实 (AR) 场景的示例的示意图。

[0030] 图2是根据一些实施例的示例性简化AR场景的示意图,示出了示例性世界重建用例,包括视觉遮挡、基于物理的交互和环境推理。

[0031] 图3是示出根据一些实施例的被配置为提供与物理世界交互的AR内容的体验的AR系统中的数据流的示意图。

[0032] 图4是示出根据一些实施例的AR显示系统的示例的示意图。

[0033] 图5A是示出根据一些实施例的当用户佩戴AR显示系统时,该AR显示系统在用户移

动穿过物理世界环境时渲染AR内容的示意图。

[0034] 图5B是示出根据一些实施例的观看光学组件和附属部件的示意图。

[0035] 图6是示出根据一些实施例的使用世界重建系统的AR系统的示意图。

[0036] 图7是示出根据一些实施例的被配置为实时生成用于动态遮挡的手部网格的AR系统的示意图。

[0037] 图8是示出根据一些实施例的实时生成用于动态遮挡的手部网格的方法的流程图。

[0038] 图9A是根据一些实施例的由对应于一只眼睛的一个传感器捕获的示例性图像。

[0039] 图9B是根据一些实施例的由对应于左眼和右眼的两个传感器捕获的两个示例性图像。

[0040] 图9C是根据一些实施例的示例性深度图像,其可以至少部分地从图9A的图像或图9B的图像获得。

[0041] 图9D是示出根据一些实施例的图8的手部的模型的轮廓的示例性图像。

[0042] 图9E是示出根据一些实施例的手部的示例性八个关键点模型的示意图。

[0043] 图9F是示出根据一些实施例的手部的示例性二十二个关键点模型的示意图。

[0044] 图9G是示出根据一些实施例的密集手部网格的示意图。

[0045] 图10是示出根据一些实施例的利用图8的手部模型来掩蔽(masking)深度信息的细节的流程图。

[0046] 图11是示出根据一些实施例的基于掩蔽到图8的手部分段的深度信息来实时计算手部网格的细节的流程图。

具体实施方式

[0047] 本文描述了在X现实(XR)系统中用于动态遮挡的快速手部网格划分的方法和设备。XR系统可以创建和使用三维(3D)世界重建。为了向用户提供逼真的XR体验,XR系统必须知道用户的物理环境,以便正确相对于真实对象来关联虚拟对象的位置。世界重建可以从关于那些物理环境的图像和深度信息来构建,图像和深度信息是利用作为XR系统一部分的传感器收集的。世界重建然后可以被这种系统的多个部件中的任何一个使用。例如,世界重建可由执行视觉遮挡处理、计算基于物理的交互、或执行环境推理的部件使用。

[0048] 遮挡处理识别不应为用户渲染和/或显示给用户的虚拟对象的部分,因为物理世界中存在对象阻挡了用户对虚拟对象将被用户感知的位置的观看。计算基于物理的交互以确定虚拟对象在何处或如何出现在用户面前。例如,虚拟对象可以被渲染为看起来停留在物理对象上、移动穿过空白空间、或与物理对象的表面碰撞。世界重建提供了一个模型,从中可以获得关于物理世界中对象的信息以进行此类计算。

[0049] 提供这样的系统存在重大挑战。可能需要大量处理来计算世界重建和遮挡信息。此外,XR系统必须正确地知道如何相对于用户的头部、身体等定位虚拟对象。随着用户相对于物理环境的位置发生变化,物理世界的相关部分也会发生变化,这可能需要进一步处理。此外,随着对象在物理世界中的移动(例如,杯子在桌子上移动),通常需要更新3D重建数据。必须在不使用生成XR环境的计算机的大量计算资源的情况下,快速执行对于表示用户正在体验的环境的数据的更新,因为在执行世界重建时无法执行其他功能。此外,“消耗”数

据的部件对于重建数据的处理会加剧对计算机资源的需求。

[0050] 动态遮挡处理识别不应为用户渲染和/或向用户显示的虚拟对象的部分,因为存在物理对象阻挡了用户对虚拟对象将被用户感知的位置的观看,并且物理对象和虚拟对象之间的相对位置随时间变化。考虑到用户手部的遮挡处理对于提供理想的XR体验尤其重要。然而,发明人已经认识到并意识到,专门针对手部的改进的遮挡处理可以为用户提供更逼真的XR体验。例如,XR系统可能会基于图形图像来生成用于在遮挡处理中使用的对象的网格,这些图形图像以每秒5帧(fps)左右的帧速率拍摄。然而,由于手部运动和/或头部运动,例如高于15fps、高于30fps或高于45fps,该速率可能无法满足手部和手部后面的虚拟对象之间的位置变化速度。

[0051] XR设备的用户可以通过用手部做手势来与设备交互。手部是延时的关键,因为它们直接用于用户交互。在与设备交互期间,用户的手部可以快速移动,例如,比用户移动以扫描物理环境来重建世界的速度快。进一步地,用户的手部离用户佩戴的XR设备更近。因此,用户手部和手部后面的虚拟对象之间的相对位置也对头部运动敏感。如果用于遮挡处理的手部的表示没有足够快地更新以跟上这些相对运动源,则遮挡处理将不会基于手部的的位置,并且遮挡处理将不准确。如果在手部运动和/或头部运动期间,手部后面的虚拟对象没有被正确渲染为被手部遮挡,则XR场景对用户来说将显得不真实。虚拟对象可能出现在手部的顶部,就好像手部是透明的一样。否则虚拟对象可能不会出现在其预期位置。手部可能看起来具有虚拟对象的颜色图案,或者可能会出现其他伪影。因此,手部的运动破坏了用户对XR体验的沉浸感。

[0052] 发明人已经认识到并意识到,当遮挡虚拟对象的对象是用户的手部时,可能需要特别高的计算需求。然而,通过以低计算资源、高速率生成手部遮挡数据的技术可以减轻这种计算负担。手部遮挡数据可以通过根据实况(live)深度传感器数据计算手部网格来生成,实况深度传感器数据以比图形图像更高的频率来获取。在一些实施例中,可以以至少30fps的帧速率获取实况深度传感器数据。为了能够快速处理该数据,可以处理少量数据以通过使用其中手部简单地由从关键点识别的多个分段表示的模型对于实况深度数据进行掩蔽来制作在遮挡处理中使用的手部模型。此外,为了增加遮挡处理的准确性,可以通过预测捕获深度数据的时间和手部网格将用于遮挡处理的时间之间的手部姿势的变化来生成手部遮挡数据。手部网格可能会变形(distort)以表示预测姿势中的手部。

[0053] 在此描述的技术可以与许多类型的设备以及对于许多类型的场景一起或单独使用,包括具有有限的计算资源的提供交叉现实场景的可穿戴或便携式设备。在一些实施例中,可以通过形成XR系统的一部分的服务来实现该技术。

[0054] 图1和图2示出了此类场景。为了说明的目的,AR系统被用作XR系统的示例。图3-6示出示例性AR系统,其包括可以根据在此描述的技术操作的一个或多个处理器、存储器、传感器和用户界面。

[0055] 参考图1,描绘了室外AR场景4,其中AR技术的用户看到了物理世界的类似公园的设置6,其特征是人、树、背景中的建筑物以及混凝土平台8。除了这些项目之外,AR技术的用户还感知到他们“看到”了站在物理世界混凝土平台8上的机器人雕像10,以及似乎是大黄蜂的头像的飞行着的卡通般的头像角色2,即使这些元素(例如,头像角色2和机器人雕像10)在物理世界中不存在。由于人类视觉感知和神经系统的极端复杂性,产生一种对于在其

它虚拟或物理世界图像元素中促进舒适、自然感觉、丰富的虚拟图像元素呈现的AR技术具有挑战性。

[0056] 这样的AR场景可以利用包括世界重建部件的系统来实现,该系统可以构建和更新用户周围的物理世界表面的表示。该表示可以用于在基于物理的交互中遮挡渲染、放置虚拟对象,以及用于虚拟角色路径规划和导航,或用于使用有关物理世界的信息的其他操作。图2描绘了根据一些实施例的室内AR场景200的另一个示例,示出了示例性世界重建用例,包括视觉遮挡202、基于物理的交互204、和环境推理206。

[0057] 示例性场景200是具有墙壁、在墙壁的一侧上的书架、在房间的拐角的落地灯、地板、沙发和地板上的咖啡桌的客厅。除了这些物理物品外,AR技术的用户还可以感知虚拟对象,诸如沙发后面的墙壁上的图像,飞过门的鸟儿,从书架上窥视的鹿,以及以放置在咖啡桌上的风车形式的装饰品。对于墙壁上的图像,AR技术不仅需要有关墙壁表面的信息,还需要有关房间内对象和表面的信息(诸如灯的形状),该信息会遮挡图像以正确渲染虚拟对象。对于飞翔的飞翔鸟类,AR技术需要有关房间周围所有对象和表面的信息,以便以逼真的物理效果渲染鸟类,以避免对象和表面或避免在鸟类碰撞时反弹。对于鹿,AR技术需要有关表面(诸如地板或咖啡桌)的信息,以计算鹿的放置位置。对于风车,系统可以识别出是与桌子分离的对象,并且可以推断其是可移动的,而架子的拐角或墙壁的拐角可以被推断为是静止的。这种区别可用于推断在各种操作中的每个操作中使用或更新场景的哪些部分。

[0058] 可以经由包括多个部件的系统向用户呈现场景,该多个部件包括可以刺激一种或多种用户感觉的用户界面,包括视觉声音和/或触摸。另外,该系统可以包括一个或多个传感器,该传感器可以测量场景的物理部分的参数,包括用户在场景的物理部分内的位置和/或运动。此外,该系统可以包括一个或多个计算设备,以及相关联的计算机硬件,诸如存储器。这些部件可以集成到单个设备中,或更多地跨多个互连设备分布。在一些实施例中,这些部件中的一些或全部可以集成到可穿戴设备中。

[0059] 图3描绘了根据一些实施例的AR系统302,其被配置为提供与物理世界306交互的AR内容的体验。AR系统302可以包括显示器308。在所示的实施例中,显示器308可以由用户作为头戴式耳机的一部分来佩戴,使得用户可以像一副护目镜或眼镜一样将显示器戴在他们的眼睛上。显示器的至少一部分可以是透明的,使得用户可以观察到透视现实310。透视现实310可以对应于处于AR系统302的当前视点内的物理世界306的部分,在用户戴上结合了AR系统的显示器和传感器的头戴式耳机以获取有关物理世界的信息的情况下,这可以与用户的视点相对应。

[0060] AR内容也可以呈现在显示器308上,覆盖在透视现实310上。为了在显示器308上提供AR内容和透视现实310之间的准确交互,AR系统302可以包括被配置为捕获有关物理世界306的信息的传感器322。

[0061] 传感器322可包括输出深度图312的一个或多个深度传感器。每个深度图312可具有多个像素,每个像素可表示距在特定方向中相对于深度传感器的物理世界306中的表面的距离。原始深度数据可来自深度传感器以创建深度图。该深度图可以与深度传感器可以形成新图像的速度一样快地被更新,其可以是每秒成百上千次。然而,该数据可能是嘈杂的和不完全的,并且在所示的深度图上具有示为黑色像素的孔。

[0062] 该系统可以包括其它传感器,诸如图像传感器。图像传感器可以获取信息,该信息

可以被处理为以其它方式表示物理世界。例如,可以在世界重建部件316中对图像进行处理以创建网格,该网格表示物理世界中对象的连接部分。有关这种对象的元数据(包括例如颜色和表面纹理)可以类似地采用传感器获取,并作为世界重建的一部分进行存储。

[0063] 系统还可以获取关于用户相对于物理世界的头部姿势的信息。在一些实施例中,传感器可包括可用于计算和/或确定头部姿势314的惯性测量单元(IMU)。用于深度图的头部姿势314可指例如以六个自由度(6DoF)捕获深度图的传感器的当前视点,但是头部姿势314可用于其它目的,诸如将图像信息与物理世界的特定部分相联系或将佩戴在用户头部上的显示器的位置与物理世界相联系。在一些实施例中,头部姿势信息可以以不同于IMU(诸如分析图像中的对象)的其它方式来导出。

[0064] 世界重建部件316可以从传感器接收深度图312和头部姿势314以及任何其它数据,并将该数据集成到重建318中,该重建可以至少看起来是单个组合重建。重建318可以比传感器数据更完整并且噪声更少。世界重建部件316可以使用随时间推移的来自多个视点的传感器数据的空间和时间平均来更新重建318。

[0065] 重建318可以包括一种或多种数据格式(包括例如体素、网格、平面等)的物理世界的表示。不同格式可以表示物理世界的相同部分的替代表示或可以表示物理世界的不同部分。在所示的示例中,在重建318的左侧,物理世界的部分被呈现为全局表面;在重建318的右侧,物理世界的部分被呈现为网格。

[0066] 重建318可以用于AR功能,诸如产生用于遮挡处理或基于物理的处理的物理世界的表面表示。该表面表示可能随着用户移动或真实世界中的对象改变而改变。重建318的各方面可以例如由在世界坐标中产生改变的全局表面表示的部件320使用,其可以由其它部件使用。

[0067] 可以基于该信息,诸如通过AR应用304来生成AR内容。AR应用304可以是例如游戏程序,其基于关于物理世界的信息来执行一个或多个功能,诸如视觉遮挡、基于物理的交互和环境推理。它可以通过查询来自世界重建部件316所产生的重建318的不同格式的数据来执行这些功能。在一些实施例中,部件320可以被配置为当物理世界的关注区域中的表示改变时输出更新。例如,该关注区域可以被设置为近似于系统用户附近的物理世界的一部分,诸如用户视野内的部分,或者被投影(预测/确定)为进入用户的视野内。

[0068] AR应用304可以使用该信息来生成和更新AR内容。AR内容的虚拟部分可以结合透视现实310呈现在显示器308上,从而创建真实的用户体验。

[0069] 在一些实施例中,可以通过可穿戴显示系统向用户提供AR体验。图4示出了可穿戴显示系统80(以下称为“系统80”)的示例。系统80包括头戴式显示设备62(以下称为“显示设备62”),以及支持显示设备62的功能的各种机械和电子模块和系统。显示设备62可以耦合至框架64,该框架64可由显示系统用户或观看者60(以下称为“用户60”)佩戴,并被配置为将显示设备62定位在用户60的眼前。根据各种实施例,显示设备62可以顺序显示。显示设备62可以是单眼的或双眼的。在一些实施例中,显示设备62可以是图3中的显示器308的示例。

[0070] 在一些实施例中,扬声器66耦合到框架64并定位在用户60的耳道附近。在一些实施例中,未示出的另一扬声器定位在用户60的另一耳道附近,以提供立体声/可塑声音控制。显示设备62诸如通过有线导线或无线连接68可操作地耦合到本地数据处理模块70,该本地数据处理模块70可以以各种配置安装,诸如固定地附接到框架64,固定地附接到用户

60所戴的头盔或帽子,嵌入耳机中或以其它方式可移除地附接到用户60(例如,以背包式配置,以腰带耦合式配置)。

[0071] 本地数据处理模块70可以包括处理器以及诸如非易失性存储器(例如,闪存)的数字存储器,二者都可以用于协助数据的处理、缓存和存储。数据包括:a)从传感器(例如,可以可操作地耦合到框架64)捕获或以其它方式附接到用户60的数据,诸如图像捕获设备(诸如相机)、麦克风、惯性测量单元、加速度计、指南针、GPS单元、无线电设备和/或陀螺仪;和/或b)使用远程处理模块72和/或远程数据存储库74获取和/或处理,可能在该处理或取得之后传递给显示设备62的数据。本地数据处理模块70可以通过通信链路76、78,例如经由有线或无线通信链路,分别可操作地耦合到远程处理模块72和远程数据存储库74,使得这些远程模块72可操作地彼此耦合并且可用作本地处理和数据模块的资源。在一些实施例中,图3中的世界重建部件316可以至少部分地在本地数据处理模块70中实现。例如,本地数据处理模块70可以被配置为执行计算机可执行指令以至少部分地基于数据的至少一部分来生成物理世界表示。

[0072] 在一些实施例中,本地数据处理模块70可以包括被配置为分析和处理数据和/或图像信息的一个或多个处理器(例如,图形处理单元(GPU))。在一些实施例中,本地数据处理模块70可以包括单个处理器(例如,单核或多核ARM处理器),这将限制模块的计算预算,但是实现更小型的设备。在一些实施例中,世界重建部件316可以使用小于单个ARM核心的计算预算来在非预定义空间上实时生成物理世界表示,使得单个ARM核心的剩余计算预算可以访问以用于其它用途,诸如例如提取网格。

[0073] 在一些实施例中,远程数据存储库74可以包括数字数据存储设施,该数字数据存储设施可以通过互联网或“云”资源配置中的其它联网配置而可用。在一些实施例中,所有数据被存储在并且所有计算在本地数据处理模块70中执行,从而允许来自远程模块的完全自主的使用。例如,世界重建可以全部或部分地存储在该存储库中。

[0074] 在一些实施例中,本地数据处理模块70可操作地耦合到电池82。在一些实施例中,电池82是可移动电源,诸如在计数器电池上方。在其它实施例中,电池82是锂离子电池。在一些实施例中,电池82既包括在系统80的非操作时间期间可由用户60充电的内部锂离子电池,又包括可移除的电池,使得用户60可以在更长的时间段内操作系统80,而不必连接到电源以为锂离子电池充电,或者不必关闭系统80以更换电池。

[0075] 图5A示出了当用户30在移动穿过物理世界环境32(以下称为“环境32”)时,佩戴有渲染AR内容的AR显示系统的用户30。用户30将AR显示系统定位在位置34,并且AR显示系统记录可通行世界相对于位置34的环境信息(例如,物理世界中真实对象的数字表示,其可以随着对真实对象在物理世界的改变被存储和更新),例如与映射特征或定向音频输入的姿势关系。位置34被聚集到数据输入36,并且至少由可通行世界模块38处理,该可通行世界模块38可以例如通过在图4的远程处理模块72上的处理来实现。在一些实施例中,可通行世界模块38可以包括世界重建部件316。

[0076] 可通行世界模块38确定如从数据输入36确定的AR内容40可以被放置在物理世界中的位置和方式。AR内容通过经由用户界面呈现物理世界呈现和AR内容二者来“放置”在物理世界中,AR内容渲染为好像与物理世界中的对象进行交互,并且物理世界中的对象呈现为好像AR内容在当时掩盖了用户对这些对象的观看。在一些实施例中,可以通过从重建

(例如重建318)中适当地选择固定元素42(例如桌子)的部分来确定AR内容40的形状和位置来放置AR内容。作为示例,固定元素可以是桌子,并且虚拟内容可以定位成使得其看起来好像在该桌子上。在一些实施例中,可以将AR内容放置在视野44中的结构内,该视野可以是当前视野或估计的未来视野。在一些实施例中,AR内容可以相对于物理世界的映射网格模型(mapped mesh model)46放置。

[0077] 如所描绘的,固定元素42用作可存储在可通行世界模块38中的物理世界内的任何固定元素的代理,使得用户30可以感知固定元素42上的内容而无需每次用户30看到固定元素42时系统都必须地图构建到固定元素42。因此,固定元素42可以是来自先前建模会话的映射网格模型,或者可以由单独的用户确定但是仍然存储在可通行世界模块38上,以供多个用户将来参考。因此,可通行世界模块38可以从先前地图构建的环境中识别环境32并显示AR内容,而无需用户30的设备首先对环境32进行地图构建(mapping),从而节省了计算过程和周期并避免了任何渲染的AR内容的时延。

[0078] 可以通过AR显示系统来创建物理世界的映射网格模型46,并且用于交互和显示AR内容40的适当的表面和度量可以被映射和存储在可通行世界模块38中,以供用户30或其他用户将来取得,而无需重新映射或模型。在一些实施例中,数据输入36是诸如地理位置、用户标识和当前活动的输入,以向可通行世界模块38指示一个或多个固定元素中的哪个固定元素42可用,哪个AR内容40最后被放置在固定元素42上,以及是否显示该相同的内容(无论用户如何观看特定的可通行世界模型,这种AR内容都是“持久”内容)。

[0079] 即使在其中对象被认为是固定的实施例中,可通行世界模块38也可以不时被更新,以考虑到物理世界中变化的可能性。固定对象的模型可能以非常低的频率更新。物理世界中的其它对象可能正在移动或以其它方式不被认为是固定的。为了渲染具有真实感的AR场景,AR系统可以以比用于更新固定对象的频率高得多的频率来更新这些非固定对象的位置。为了能够准确跟踪物理世界中的所有对象,AR系统可以从多个传感器(包括一个或多个图像传感器)中获取信息。

[0080] 图5B是观看光学组件48和附属部件的示意图。在一些实施例中,指向用户眼睛49的两个眼睛跟踪相机50检测用户眼睛49的度量,诸如用户眼睛49上的眼睛形状、眼睑闭塞、瞳孔方向和闪烁。在一些实施例中,传感器之一可以是深度传感器51,诸如飞行时间传感器,其向世界发射信号并检测来自附近对象的那些信号的反射以确定与给定对象的距离。深度传感器例如可以由于那些对象的运动或用户姿势的变化而快速确定对象是否已经进入用户的视野。然而,关于对象在用户视野中的位置的信息可以可替代地或另外地由其它传感器收集。深度信息例如可以从立体视觉图像传感器或全光传感器获得。

[0081] 在一些实施例中,世界相机52记录大于外围的视图以对环境32进行映射环境32,并检测可影响AR内容的输入。在一些实施例中,世界相机52和/或相机53可以是灰度和/或彩色图像传感器,该传感器可以以固定的时间间隔输出灰度和/或彩色图像帧。相机53可以在特定时间进一步捕获用户视野内的物理世界图像。即使基于帧的图像传感器的像素的值不变,也可以重复对其像素进行采样。世界相机52、相机53和深度传感器51中的每一个具有相应的视野54、55和56,以从诸如图5A中描绘的物理世界环境32的物理世界场景中收集数据并记录该物理世界场景。

[0082] 惯性测量单元57可以确定观看光学组件48的运动和取向。在一些实施例中,每个

部件可操作地耦合到至少一个其它部件。例如,深度传感器51可操作地耦合到眼睛跟踪相机50,以相对于用户眼睛49正在注视的实际距离来确认所测量的适应。

[0083] 应当理解,观看光学组件48可以包括图5B中所示的一些部件,并且可以包括代替所示部件或除了所示部件之外的部件。例如,在一些实施例中,观看光学组件48可以包括两个世界相机52而不是四个。可替代地或另外地,相机52和53不需要捕获其全部视野的可见光图像。观看光学组件48可以包括其它类型的部件。在一些实施例中,观看光学组件48可以包括一个或多个动态视觉传感器(DVS),其像素可以异步地响应超过阈值的光强度的相对变化。

[0084] 在一些实施例中,基于飞行时间信息,观看光学组件48可以不包括深度传感器51。例如,在一些实施例中,观看光学组件48可以包括一个或多个全光相机,其像素可以捕获光强度和入射光的角度,由此可以确定深度信息。例如,全光相机可以包括覆盖有透射衍射掩模(TDM)的图像传感器。可替代地或另外地,全光相机可以包括图像传感器,该图像传感器包含角度敏感像素和/或相位检测自动聚焦像素(PDAF)和/或微透镜阵列(MLA)。代替深度传感器51或除了深度传感器51之外,这种传感器可以用作深度信息源。

[0085] 还应当理解,图5B中的部件的配置被示出为示例。观看光学组件48可以包括具有任何适当配置的部件,该适当配置可以设定为向用户提供对于一组特定部件实际可行的最大视野。例如,如果观看光学组件48具有一个世界相机52,则该世界相机可以放置在观看光学组件的中心区域而不是侧面。

[0086] 来自观看光学组件48中的传感器的信息可以耦合到系统中的一个或多个处理器。处理器可以生成可以被渲染以便使用户感知与物理世界中的对象交互的虚拟内容的数据。可以以任何合适的方式来实现该渲染,包括生成描绘物理和虚拟对象二者的图像数据。在其它实施例中,可以通过调制用户在物理世界中浏览的显示设备的不透明度来在一个场景中描绘物理和虚拟内容。可以控制不透明度以便创建虚拟对象的外观,并且还阻止用户看到物理世界中被虚拟对象遮挡的对象。在一些实施例中,当通过用户界面观看时,图像数据可以仅包括虚拟内容,该虚拟内容可以被修改以使得虚拟内容被用户感知为与物理世界现实地交互(例如,剪辑内容以考虑遮挡)。无论如何向用户呈现内容,都需要物理世界的模型,使得可以正确计算可受到物理对象影响的虚拟对象的特性,包括虚拟对象的形状、位置、运动和可见性。在一些实施例中,模型可以包括物理世界的重建,例如重建318。

[0087] 可以根据从用户的可穿戴设备上的传感器收集的数据来创建该模型。但是,在一些实施例中,可以从多个用户收集的数据中创建模型,该数据可以在远离所有用户的计算设备中汇总(并且该数据可以在“云中”)。

[0088] 模型可以至少部分地由世界重建系统创建,例如,图6中更详细地描绘的图3的世界重建部件316。世界重建部件316可以包括感知模块160,该感知模块160可以生成、更新和存储物理世界的一部分的表示。在一些实施例中,感知模块160可以将传感器的重建范围内的物理世界的部分表示为多个体素。每个体素可以对应于物理世界中预定体积的3D立方体,并且包括表面信息,该表面信息指示在体素表示的体积中是否存在表面。可以为体素分配值,该值指示是否已经确定了它们的对应体积以包括物理对象的表面,确定为空的还是尚未用传感器测量,并且因此其值是未知的。应当理解,不需要明确地存储指示被确定为空或未知的体素的值,因为可以以任何合适的方式将体素的值存储在计算机存储器中,包括

不存储被确定为空或未知的体素的信息。

[0089] 除了生成用于持久世界表示的信息之外,感知模块160还可以识别并输出AR系统的用户周围的区域中的变化的指示。这种变化的指示可以触发对作为持久世界的一部分存储的体积数据的更新,或者触发其它功能,诸如触发生成AR内容以更新AR内容的触发部件304。

[0090] 在一些实施例中,感知模块160可以基于符号距离函数(SDF)模型来识别变化。感知模块160可以被配置为接收诸如例如深度图160a和头部姿势160b的传感器数据,并且然后将传感器数据融合到SDF模型160c中。深度图160a可以直接提供SDF信息,并且图像可以被处理以获得SDF信息。SDF信息表示距用于捕获该信息的传感器的距离。由于那些传感器可以是可穿戴单元的一部分,因此SDF信息可以从可穿戴单元的角度以及因此从用户的角度表示物理世界。头部姿势160b可以使SDF信息能够与物理世界中的体素相关。

[0091] 在一些实施例中,感知模块160可以生成、更新和存储在感知范围内的物理世界的部分的表示。感知范围可以至少部分地基于传感器的重建范围来确定,该传感器的重建范围可以至少部分地基于传感器的观察范围的限制来确定。作为特定示例,使用主动IR脉冲操作的主动深度传感器可以在一定距离范围内可靠地操作,从而创建传感器的观察范围,该观察范围可以从几厘米或几十厘米到几米。

[0092] 世界重建部件316可以包括可以与感知模块160交互的附加模块。在一些实施例中,持久世界模块162可以基于由感知模块160获取的数据来接收对物理世界的表示。持久世界模块162还可以包括物理世界的各种格式的表示。例如,可以存储诸如体素的体积元数据162b以及网格162c和平面162d。在一些实施例中,可以保存其它信息,诸如深度图。

[0093] 在一些实施例中,感知模块160可以包括以各种格式生成物理世界的表示的模块,这些格式包括例如网格160d、平面和语义160e。这些模块可以基于在生成表示时一个或多个传感器的感知范围内的数据以及在先前时间捕获的数据和持久世界162中的信息来生成表示。在一些实施例中,这些部件可以关于采用深度传感器捕获的深度信息来操作。然而,AR系统可以包括视觉传感器,并且可以通过分析单眼或双眼视觉信息来生成这种表示。

[0094] 在一些实施例中,这些模块可以在物理世界的区域上操作。当感知模块160检测到物理世界的子区域中的物理世界的变化时,可以触发那些模块以更新物理世界的子区域。例如,可以通过在SDF模型160c中检测新的表面或其它标准(例如改变表示子区域的足够数量的体素的值)来检测这种变化。

[0095] 世界重建部件316可以包括可以从感知模块160接收物理世界的表示的部件164。关于物理世界的信息可以由这些部件根据例如来自应用的使用请求来提取。在一些实施例中,可以诸如经由对预识别区域中的变化或感知范围内的物理世界表示的变化的指示,将信息推送到使用部件。部件164可以包括例如游戏程序和执行用于视觉遮挡、基于物理的交互以及环境推理的处理的其它部件。

[0096] 响应于来自部件164的查询,感知模块160可以一种或多种格式发送用于物理世界的表示。例如,当部件164指示该使用是用于视觉遮挡或基于物理的交互时,感知模块160可以发送表面的表示。当部件164指示该使用是用于环境推理时,感知模块160可以发送物理世界的网格、平面和语义。

[0097] 在一些实施例中,感知模块160可以包括格式化信息以提供给部件164的部件。这

种部件的示例可以是光线投射部件160f。使用部件(例如,部件164)例如可以从特定的视点查询关于物理世界的信息。光线投射部件160f可以从该视点从视野内的物理世界数据的一种或多种表示中选择。

[0098] 关于物理世界的信息也可用于遮挡处理。该信息可以由视觉遮挡部件164a使用,该视觉遮挡部件164a可以是世界重建部件316的一部分。例如,视觉遮挡部件164a可以向应用提供指示视觉对象的哪些部分被物理对象遮挡的信息。替代地或附加地,视觉遮挡部件164a可以向应用提供关于物理对象的信息,应用可以使用该信息进行遮挡处理。如上所述,关于手部位置的准确信息对于遮挡处理很重要。在如本文所述的示例中,视觉遮挡部件164a可以响应于应用的请求,维护手部模型并且在被请求时将该模型提供给应用。图7示出了这种处理的示例,其可以跨越图6中图示的一个或多个部件来执行,或者,在一些实施例中,通过不同的或附加的部件来执行。

[0099] 图7描绘了根据一些实施例的被配置为实时生成用于动态遮挡处理的手部网格的AR系统700。AR系统700可以在AR设备上实现。AR系统700可以包括数据收集部702,被配置为使用AR设备上的传感器来捕获佩戴AR设备(例如,显示设备62)的用户的姿势(例如,头部姿势、手部姿势等)和场景信息。场景信息可以包括深度信息,该深度信息指示AR设备与场景中的物理对象之间的距离。

[0100] 数据收集部702包括手部跟踪部件。手部跟踪部件可以处理传感器数据,诸如深度信息和图像信息,以检测场景中一只或多只手。可以处理其他传感器数据以检测场景中一只或多只手。当检测到时,一只或多只手可以稀疏方式表示,例如由一组关键点表示。例如,关键点可以表示关节、指尖或手部分段(segment)的其他边界。由数据收集部702收集或生成的信息可以传递到手部网格划分部704,以用于基于稀疏表示来生成一只或多只手的更丰富的模型,例如网格。

[0101] 手部网格划分部704被配置为计算检测到一只或多只手的手部网格并且随着姿势改变和/或手部移动实时更新手部网格。

[0102] AR系统700可以包括应用706,该应用706被配置为从手部网格划分部704接收手部网格并且渲染场景中的一个或多个虚拟对象。在一些实施例中,应用706可以从手部网格划分部704接收遮挡数据。在一些实施例中,遮挡数据可以指示虚拟对象被一只或多只手遮挡的部分。在一些实施例中,例如在所示实施例中,遮挡数据可以是一只或多只手的模型,应用706可以根据该模型计算遮挡数据。作为特定示例,遮挡数据可以是来自手部网格划分部704接收的一只或多只手的手部网格。

[0103] 图8是示出根据一些实施例的实时生成用于动态遮挡的手部网格的方法800的流程图。在一些实施例中,方法800可由AR系统700内的一个或多个处理器执行。方法800可以在AR系统700的手部网格划分部704从AR系统700的应用706接收(动作802)针对与场景中一只或多只手相关的数据的查询时开始。方法800可以包括基于由AR系统700的数据收集部702捕获的场景的信息来检测(动作804)场景中一只或多只手。

[0104] 当检测到一只或多只手时,方法800可以包括基于场景的信息计算(动作806)一只或多只手的一个或多个模型。一只或多只手的一个或多个模型可能是稀疏的,指示手上关键点的位置,而不是表面。那些关键点可以表示手部分段的关节或末端部分。可以从关于一只或多只手的传感器数据识别关键点,包括例如一只或多只手的立体图像。深度信息以及

在一些情况下一只或多只手的单眼图像可以替代地或附加地用于识别关键点。

[0105] 题为“手部姿势估计 (Hand Pose Estimation)”的美国临时专利申请No.62/850,542描述了获得关于手的位置和姿势的信息以及基于所获得的信息对手进行建模的示例性方法和装置。美国申请No.62/850,542的已提交版本的副本作为附录附在此处,并且出于所有目的通过引用将其全部内容并入本文。在该申请中描述的技术可用于构建手部的稀疏模型。

[0106] 在一些实施例中,可以基于由AR设备的传感器捕获的场景信息来计算一只或多只手的一个或多个模型。场景信息的示例包括由与单只眼睛对应的一个传感器捕获的图9A的示例图像、由对应于左眼和右眼的两个传感器所捕获的图9B的两个示例图像、以及图9C的示例性深度图像,该示例性深度图像可以至少部分地从图9A的图像或图9B的图像获得。

[0107] 在一些实施例中,一只或多只手的一个或多个模型可以包括手部的多个关键点,这些关键点可以指示手部的分段上的点。一些关键点可以对应于手的关节和手的指尖。图9E和图9F描绘了分别示出手的示例性八个关键点模型和手的示例性二十二个关键点模型的示意图。

[0108] 在一些实施例中,手部模型的关键点可以用于确定手部的轮廓。图9D描绘了手的示例性轮廓,其可以基于手部模型的关键点来确定。例如,相邻的关键点可以通过线连接,如图9E和图9F示意性所示,并且手的轮廓可以表示为与线的距离。与线的距离可以根据手部的图像、关于人体解剖结构的信息和/或其他信息来确定。应当理解,一旦手部的关键点被识别,手部的模型可以在稍后使用先前获取的关于手的信息来更新。例如,连接关键点的线的长度可能不会改变。

[0109] 稀疏手部模型可以用于选择有限数量的数据,从中可以构建更丰富的手部模型,例如,包括表面信息。在一些实施例中,可以通过使用手的轮廓掩蔽附加数据(例如深度数据)来进行选择。因此,方法800可以包括利用一只或多只手的一个或多个模型掩蔽(动作808)指示AR设备与场景中的物理对象之间的距离的深度信息。

[0110] 图10描绘了示出根据一些实施例的用一只或多只手的一个或多个模型掩蔽(动作808)深度信息的细节的流程图。动作808可以包括过滤掉(动作1002)在一只或多只手的一个或多个模型的轮廓之外的深度信息。该动作导致移除与场景中的手以外的物理对象相关联的深度信息。动作808可以包括基于过滤的深度信息生成(动作1004)一只或多只手的深度图像。

[0111] 深度图像可以包括像素,每个像素可以指示到一只或多只手的点的距离。在一些实施例中,可以通过不针对一只或多只手的所有表面捕获深度信息的方式来捕获深度信息。例如,深度信息可以利用有源IR传感器捕获。例如,如果用户佩戴带有深色宝石的戒指,IR可能不会被深色戒指反射,从而在收集到的不充分信息中出现空洞(hole)。动作808可以包括填充(动作1006)深度图像中的空洞。在一些实施例中,可以通过识别深度图像中的空洞并从增强现实设备的立体相机生成与识别的区域对应的立体深度信息来填充空洞。在一些实施例中,可以通过识别深度图像中的空洞并从一只或多只手的一个或多个3D模型访问与识别的区域对应的表面信息来填充这些空洞。

[0112] 可选地,一个或多个手部网格可以被计算为多个子网格,每个子网格表示一个或多个手的分段。分段可以对应于由关键点界定的分段。由于这些分段中的许多分段由关节

界定,因此这些分段对应于可以独立于手部的至少一些其他分段进行移动的手部的部分,从而可以通过更新与自从最后的手部网格计算后发生移动的分段相关联的子网格来快速更新手部网格。在这样的实施例中,动作808可以包括识别(动作1008)与一只或多只手的一个或多个模型的关键点对应的深度图像中的关键点。可以计算由关键点分开的手部分段。动作808可以包括将深度图像的部分关联(动作1010)到由深度图像中识别出的关键点分开的手部分段。

[0113] 方法800可以包括基于掩蔽到一只或多只手的一个或多个模型的深度信息来计算(动作810)一个或多个手部网格。一个或多个手部网格可以是一只或多只手部的表示,指示一只或多只手的表面。图9G描绘了根据一些实施例的计算出的密集手部网格。然而,本申请不限于计算密集的手部网格。在一些实施例中,稀疏手部网格可能足以用于动态遮挡。

[0114] 在一些实施例中,可以根据深度信息计算一个或多个手部网格。例如,网格可以是区域的集合,通常表示为三角形,表示表面的一部分。可以通过对于深度图像中距离差小于阈值的相邻像素进行分组来识别这样的区域,指示这些像素很可能在同一表面上。可识别界定此类像素区域的一个或多个三角形并将其添加至网格。然而,可以使用用于根据深度信息形成网格的其他技术。

[0115] 计算一个或多个手部网格可以包括随着AR设备与一只或多只手之间的相对位置改变而实时更新一个或多个手部网格。在用于计算一个或多个手部网格的场景信息被捕获的时间与一个或多个计算出的手部网格被例如应用使用(例如,以渲染内容)的时间之间可能存在延时。在一些实施例中,可以跟踪一只或多只手的分段的运动,从而可以推算(project)/预测一只或多只手的那些分段的未来位置。在使用一个或多个手部网格处理的虚拟对象将被渲染时,一个或多个手部网格可以被变形以符合分段的推测位置。

[0116] 图11描绘了示出根据一些实施例的基于掩蔽到手部分段的深度信息来实时计算(动作810)一个或多个手部网格的细节的流程图。动作810可以包括确定(动作1102)手部网格划分部704从应用706接收对于与一只或多只手有关的数据的查询的时间 t 。

[0117] 动作810可以包括根据在时间 t 从应用706接收到的查询来预测(动作1104)延时 n 。动作810可以包括预测(动作1106)在查询时间 t 加上延时 n 的时间的姿势(例如,手部姿势)。动作810可以包括使用在查询时间 t 加上延时 n 的时间的预测姿势使一个或多个手部网格变形(动作1108)。在一些实施例中,预测手部姿势可以包括预测在查询时间 t 加上延时 n 的时间的预测深度图像中一只或多只手的关键点的运动。可以基于随时间跟踪关键点的位置来做出这样的预测。这种跟踪使得能够确定诸如速度或加速度之类的运动参数。假设所确定的运动参数保持相同,可以基于从先前位置到未来位置的外推来进行位置的推算。替代地或附加地,可以使用卡尔曼滤波器或类似的推算技术来确定推算。

[0118] 在一些实施例中,利用预测姿势使一个或多个手部网格变形可以包括使与手部分段的子集对应的一个或多个手部网格的部分变形,这些部分位于预测在查询时间 t 加上延时 n 的时间发生变化的关键点之间。

[0119] 在使先前计算的网格变形以表示时间 t 加上延时 n 时的一只或多只手时,可能会考虑多种因素。例如, n 的值可以反映使一个或多个网格变形的处理以及应用在渲染对象中使用一个或多个网格所需的时间。该值可以从软件操作中的结构或测试中估计或测量。作为另一示例,该值可以基于测量使用中的延时或基于做出对于一个或多个网格的请求时的处

理负载调整先前建立的延时来动态确定。

[0120] 在使一个或多个网格变形时,变形量可以基于捕获用于形成一个或多个手部网格的数据的时间以及直到将使用一个或多个网格的延时。在一些实施例中,除了响应于来自应用的请求之外,还可以创建一个或多个手部网格。例如,一旦应用指示(例如通过API进行调用)它被配置用于遮挡处理,AR系统就可以定期计算一个或多个更新的手部网格。替代地,手部跟踪处理可以使用系统的一定量的计算资源连续运行。在任何情况下,可以相对频繁地更新一个或多个手部网格,例如每秒至少30次。尽管如此,在捕获数据以制作网格与收到网格请求之间可能存在延迟,并且在使手部网格变形中该延迟也可以被考虑在内。

[0121] 方法800可以包括向应用706提供(动作812)一个或多个手部网格,使得应用渲染虚拟对象的未被一个或多个手部网格遮挡的部分。

[0122] 至此已经描述了一些实施例的几个方面,应当理解,本领域技术人员将容易想到各种改变、修改和改进。

[0123] 作为一个示例,结合增强(AR)环境来描述实施例。应当理解,在此描述的一些或全部技术可以在MR环境中或更普遍地在其它XR环境和VR环境中应用。

[0124] 作为另一个示例,结合诸如可穿戴设备的设备来描述实施例。应当理解,可以经由网络(诸如云)、分立应用和/或设备、网络和分立应用的任何合适的组合来实现在此描述的一些或全部技术。

[0125] 这种改变、修改和改进旨在是本公开的一部分,并且旨在处于本公开的精神和范围内。此外,尽管指示了本公开的优点,但是应当理解,并非本公开的每个实施例将包括每个描述的优点。在一些情况下,一些实施例可能不实现在此描述为有利的任何特征。因此,前面的描述和附图仅作为示例。

[0126] 本公开的上述实施例可以以多种方式中的任何一种方式来实现。例如,可以使用硬件、软件或其组合来实现实施例。当以软件实现时,软件代码可以在任何合适的处理器或处理器集合上执行,无论是在单个计算机中提供还是在多个计算机中分布。这种处理器可以被实现为集成电路,在集成电路部件中具有一个或多个处理器,包括本领域已知的可商购的集成电路部件,其名称诸如CPU芯片、GPU芯片、微处理器、微控制器或协处理器。在一些实施例中,可以在定制电路(诸如ASIC)或者在通过配置可编程逻辑器件而产生的半定制电路中实现处理器。作为另一替代方案,处理器可以是较大的电路或半导体设备的一部分,无论是可商购的,半定制的还是定制的。作为特定示例,一些可商购的微处理器具有多个核心,使得这些核心中的一个或子集可以构成处理器。但是,可以使用任何适当格式的电路来实现处理器。

[0127] 此外,应当理解,计算机可以以多种形式中的任何一种形式来体现,诸如机架式计算机、台式计算机、膝上型计算机或平板计算机。另外,计算机可以被嵌入通常不被认为是计算机但具有合适的处理能力的设备中,该设备包括个人数字助理(PDA)、智能电话或任何其它合适的便携式或固定电子设备。

[0128] 此外,计算机可以具有一个或多个输入和输出设备。这些设备尤其可以用于呈现用户界面。可以用于提供用户界面的输出设备的示例包括用于视觉呈现输出的打印机或显示屏,以及用于听觉呈现输出的扬声器或其它声音生成设备。可以用于用户界面的输入设备的示例包括键盘和指向设备,诸如鼠标、触摸板和数字化平板计算机。作为另一个示例,

计算机可以通过语音识别或其它可听格式接收输入信息。在所示的实施例中,输入/输出设备被示为与计算设备物理上分开。然而,在一些实施例中,输入和/或输出设备可以物理地集成到与处理器或计算设备的其它元件相同的单元中。例如,键盘可能实现为触摸屏上的软键盘。在一些实施例中,输入/输出设备可以与计算设备完全断开,并且通过无线连接在功能上集成。

[0129] 这种计算机可以通过任何适当形式的一个或多个网络互连,包括作为局域网或诸如企业网络或互联网的广域网。这种网络可以基于任何适当的技术,并且可以根据任何适当的协议进行操作,并且可以包括无线网络、有线网络或光纤网络。

[0130] 此外,在此概述的各种方法或过程可以被编码为可在采用多种操作系统或平台中的任何一种操作系统或平台的一个或多个处理器上执行的软件。另外,可以使用多种合适的编程语言和/或编程或脚本工具中的任何一种来编写此类软件,并且还可以将其编译为可执行的机器语言代码或在框架或虚拟机上执行的中间代码。

[0131] 在这方面,本公开可以体现为采用一个或多个程序编码的计算机可读存储介质(或多个计算机可读介质)(例如,计算机存储器、一个或多个软盘、压缩光盘(CD)、光盘、数字视频盘(DVD)、磁带、闪存、现场可编程门阵列或其它半导体设备或其它有形计算机存储介质中的电路装置),该程序在一个或多个计算机或其它处理器上执行时将执行实现以上讨论的本公开的各种实施例的方法。从前述示例显而易见,计算机可读存储介质可以将信息保留足够的时间,以提供非暂态形式的计算机可执行指令。这种一个或多个计算机可读存储介质可以是可移动的,使得可以将存储在其上的一个或多个程序加载到一个或多个不同的计算机或其它处理器上,以实现如上所述的本公开的各个方面。如在此所使用的,术语“计算机可读存储介质”仅涵盖可以被认为是制品(即,制造品)或机器的计算机可读介质。在一些实施例中,本公开可以体现为除计算机可读存储介质之外的计算机可读介质,诸如传播信号。

[0132] 在一般意义上在此使用术语“程序”或“软件”来指代可用于对计算机或其它处理器进行编程以实现如上所述的本公开的各个方面的计算机代码或计算机可执行指令集。另外,应当理解,根据该实施例的一方面,当被执行时执行本公开的方法的一个或多个计算机程序不需要驻留在单个计算机或处理器上,而是可以以模块化的方式在多个不同的计算机或处理器之间分布,以实现本公开的各个方面。

[0133] 计算机可执行指令可以具有许多形式,诸如由一个或多个计算机或其它设备执行的程序模块。通常,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、部件、数据结构等。通常,在各种实施例中,程序模块的功能可以根据需要进行组合或分布。

[0134] 此外,数据结构可以以任何合适的形式存储在计算机可读介质中。为了简化说明,可以示出数据结构具有通过数据结构中的位置而相关的字段。同样地,可以通过传达字段之间的关系在计算机可读介质中的位置为字段分配存储来实现此类关系。然而,可以使用任何合适的机制来建立数据结构的字段中的信息之间的关系,包括通过使用指针、标签或在数据元素之间建立关系的其它机制。

[0135] 本公开的各个方面可以单独、组合或在前述实施例中未具体讨论的各种布置中使用,并且因此,在其应用中不限于在前面的描述中阐述或在附图中示出的部件的细节和布

置。例如,一个实施例中描述的方面可以以任何方式与其它实施例中描述的方面组合。

[0136] 此外,本公开可以体现为一种方法,已经提供了方法的一个示例。作为该方法的一部分执行的动作可以以任何合适的方式排序。因此,可以构造实施例,其中以与所示出的顺序不同的顺序执行动作,即使在说明性实施例中被示为顺序动作,该动作也可以包括同时执行一些动作。

[0137] 在权利要求中使用诸如“第一”、“第二”、“第三”等序数术语来修改权利要求元素本身并不表示一个权利要求元素相对于执行方法动作的另一个顺序或临时顺序的任何优先级、优先权或顺序,但仅用作标记,以区分具有某些名称的一个权利要求元素与具有相同名称的另一个元素(但用于序数词)来区分权利要求元素。

[0138] 另外,在此使用的措词和术语是出于描述的目的,并且不应被视为限制。在此“包括”、“包含”或“具有”、“包含了”、“涉及”及其变体的使用旨在涵盖其后列出的项目及其等同物以及其它项目。

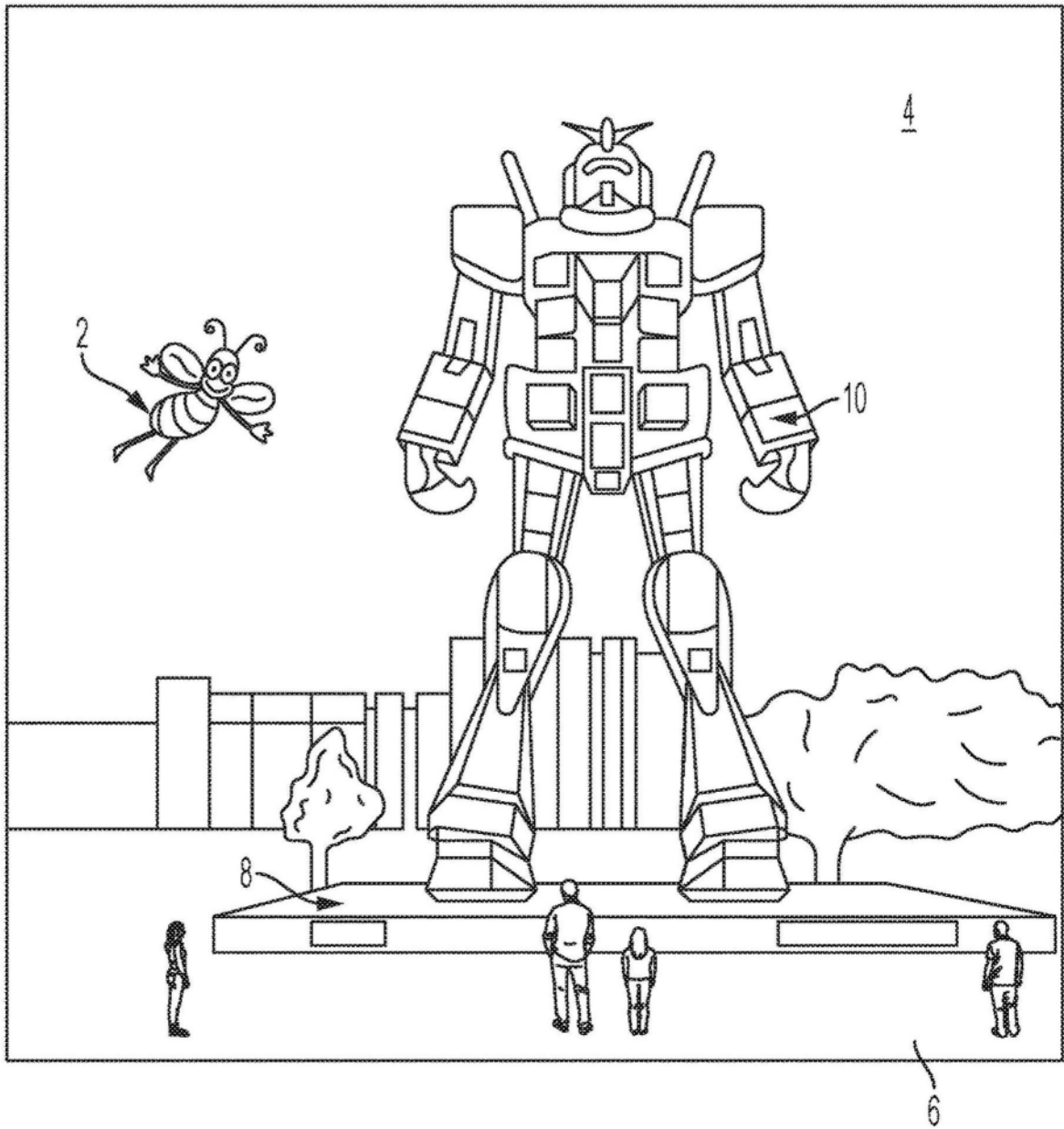


图1

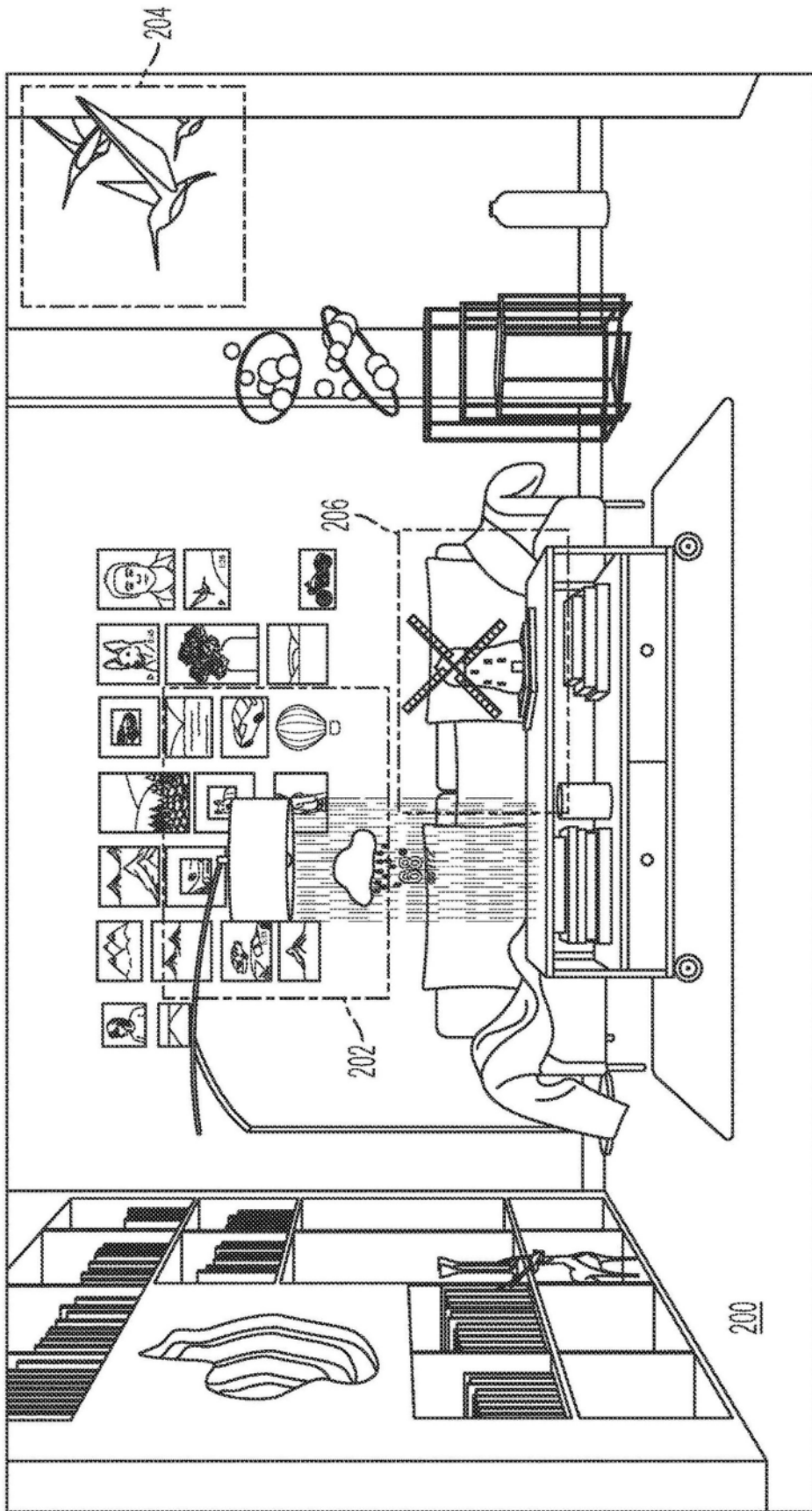


图2

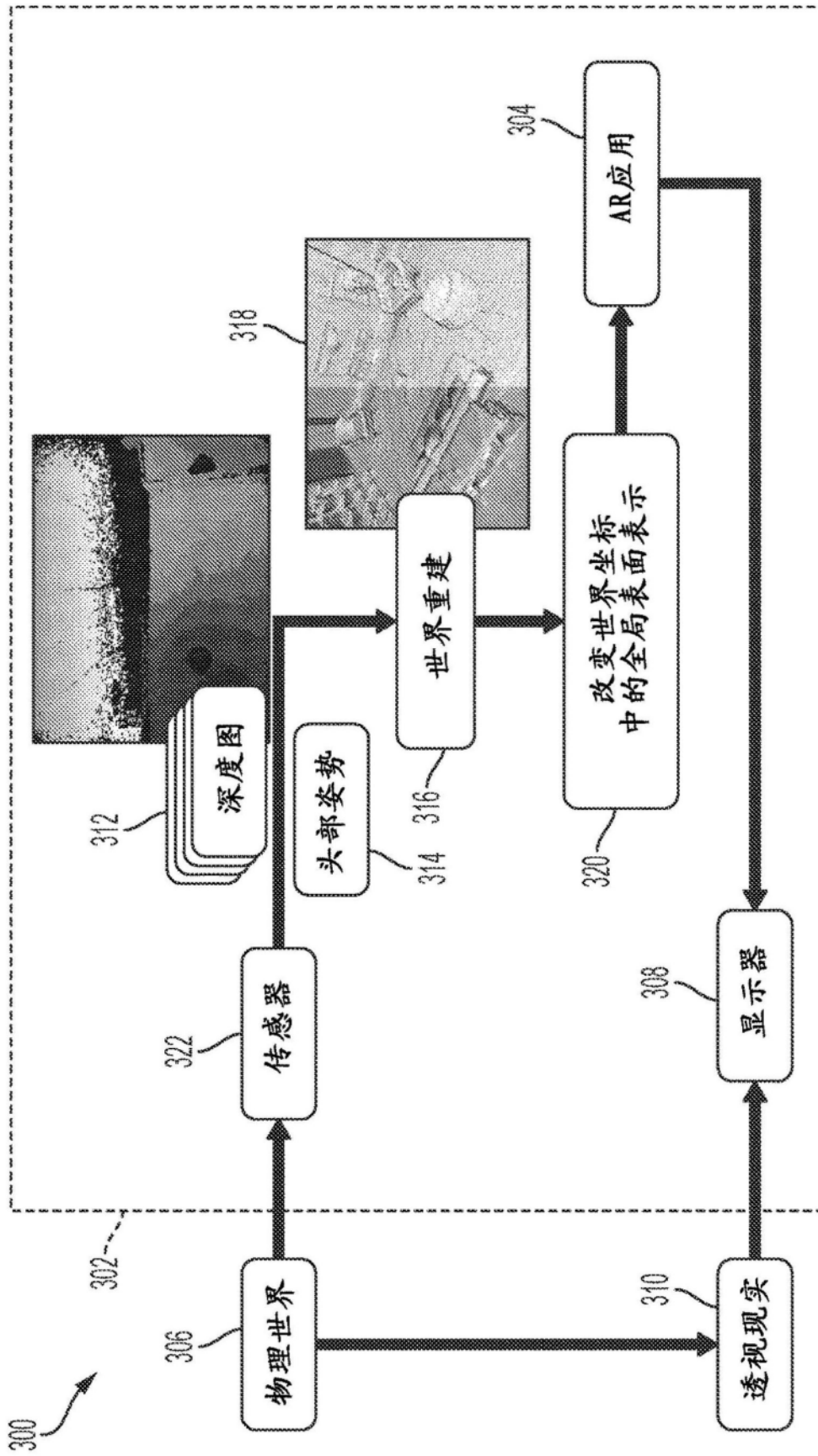


图3

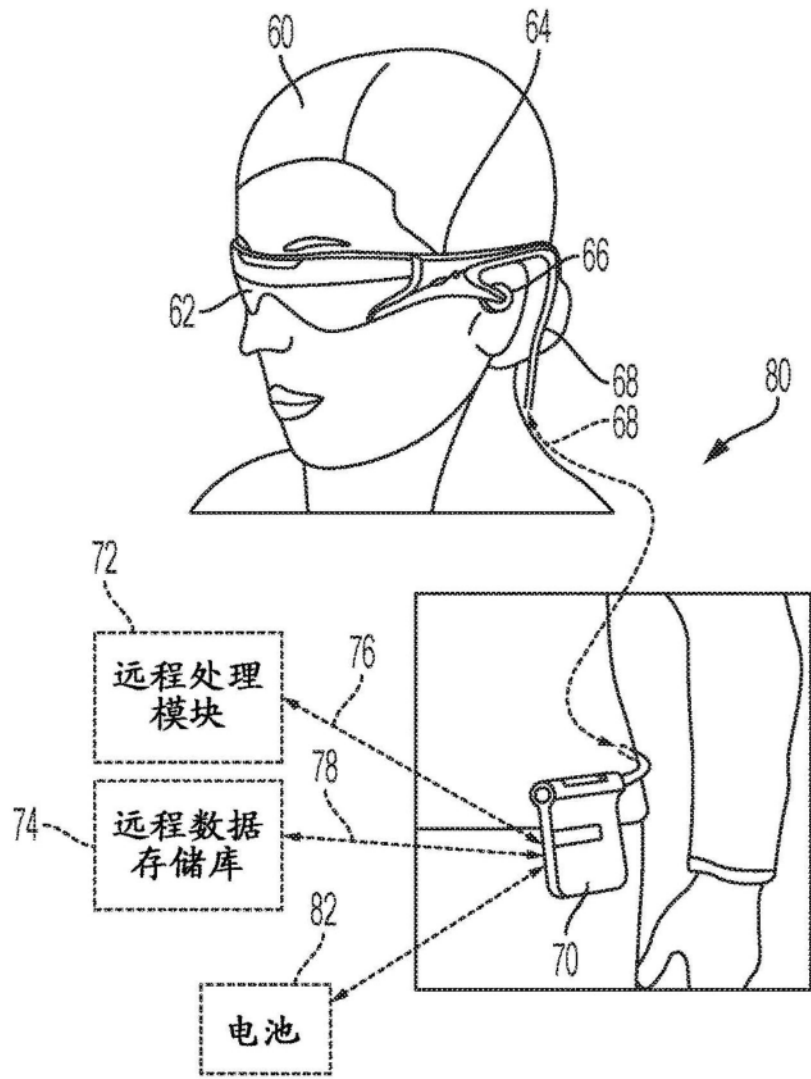


图4

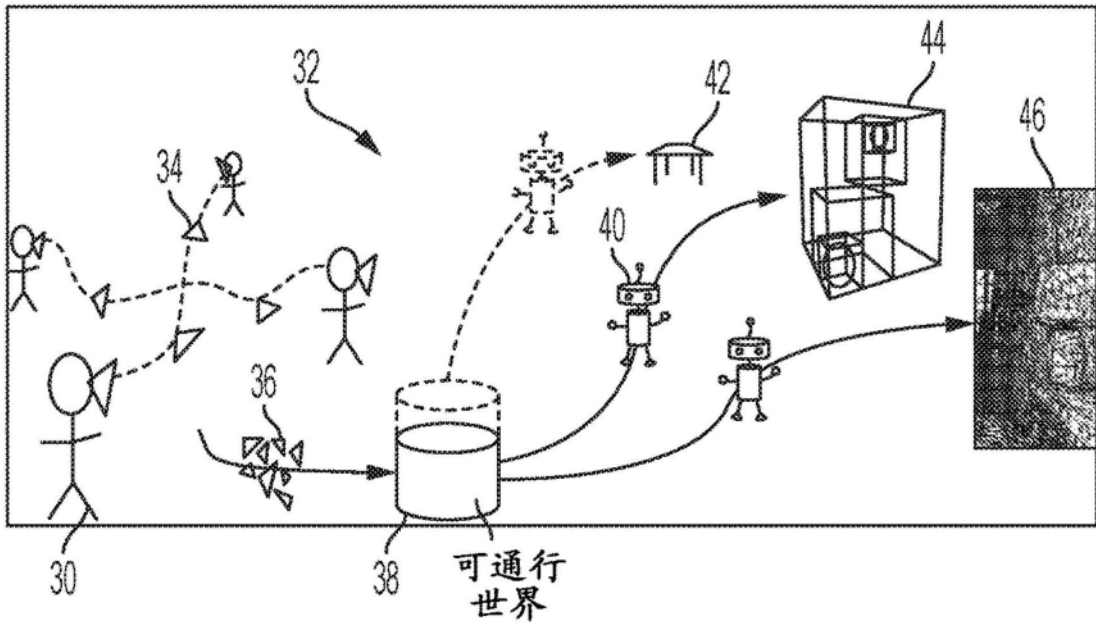


图5A

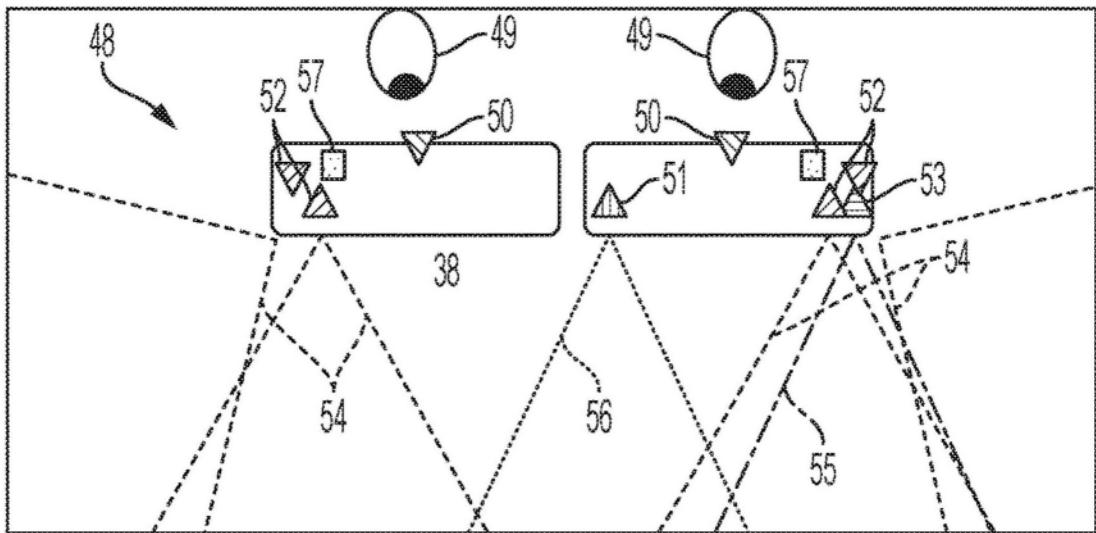


图5B

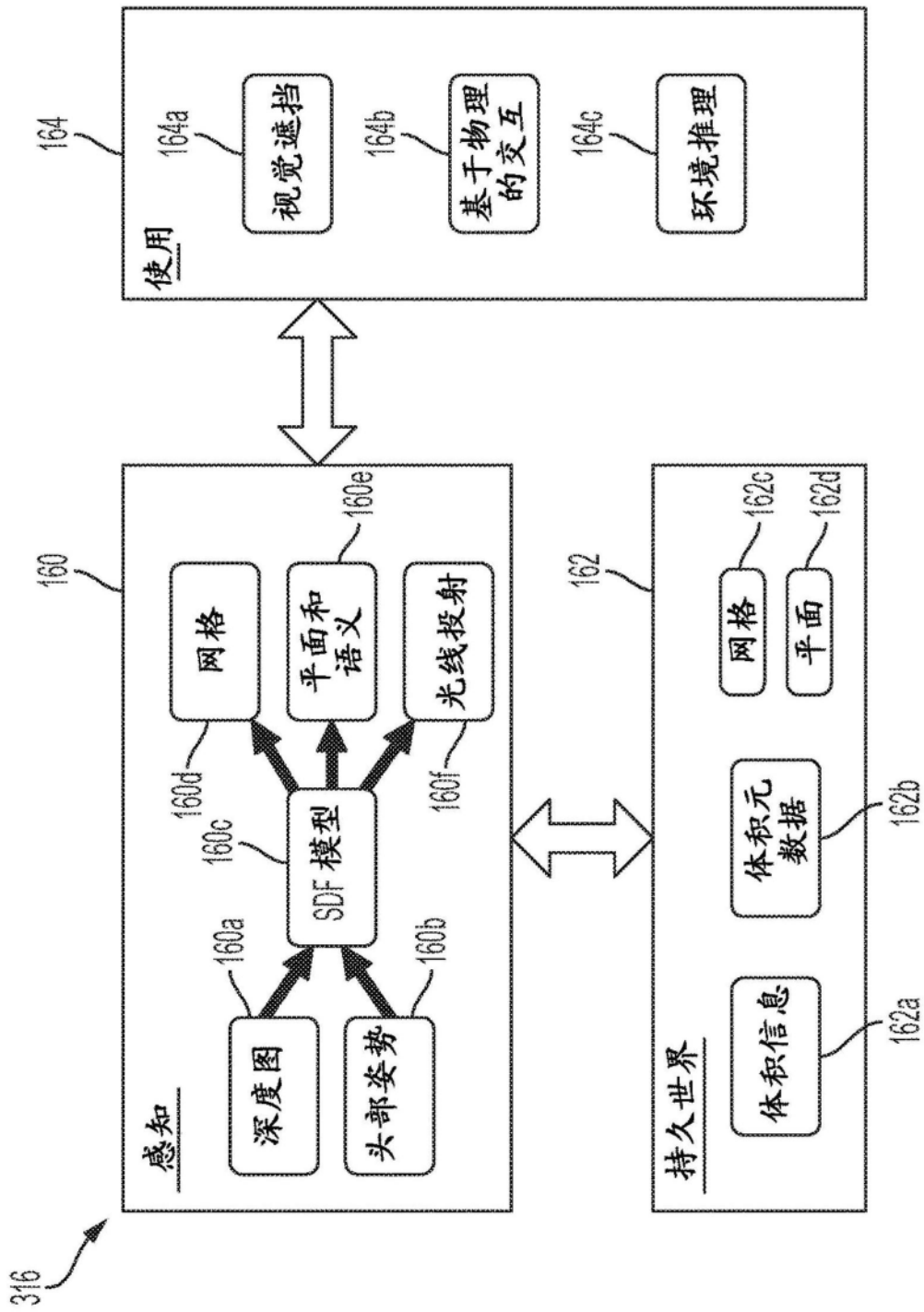


图6

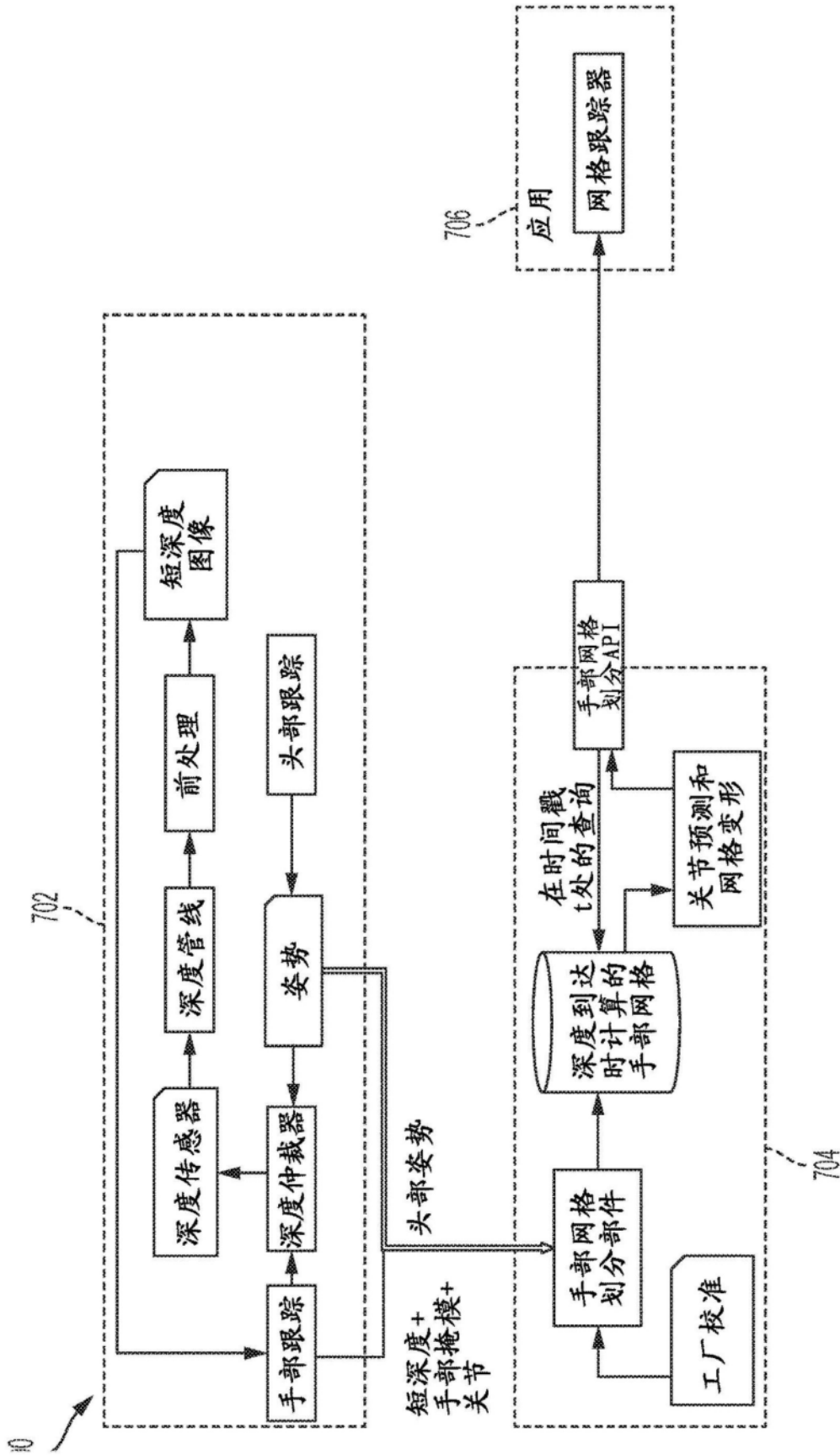


图7

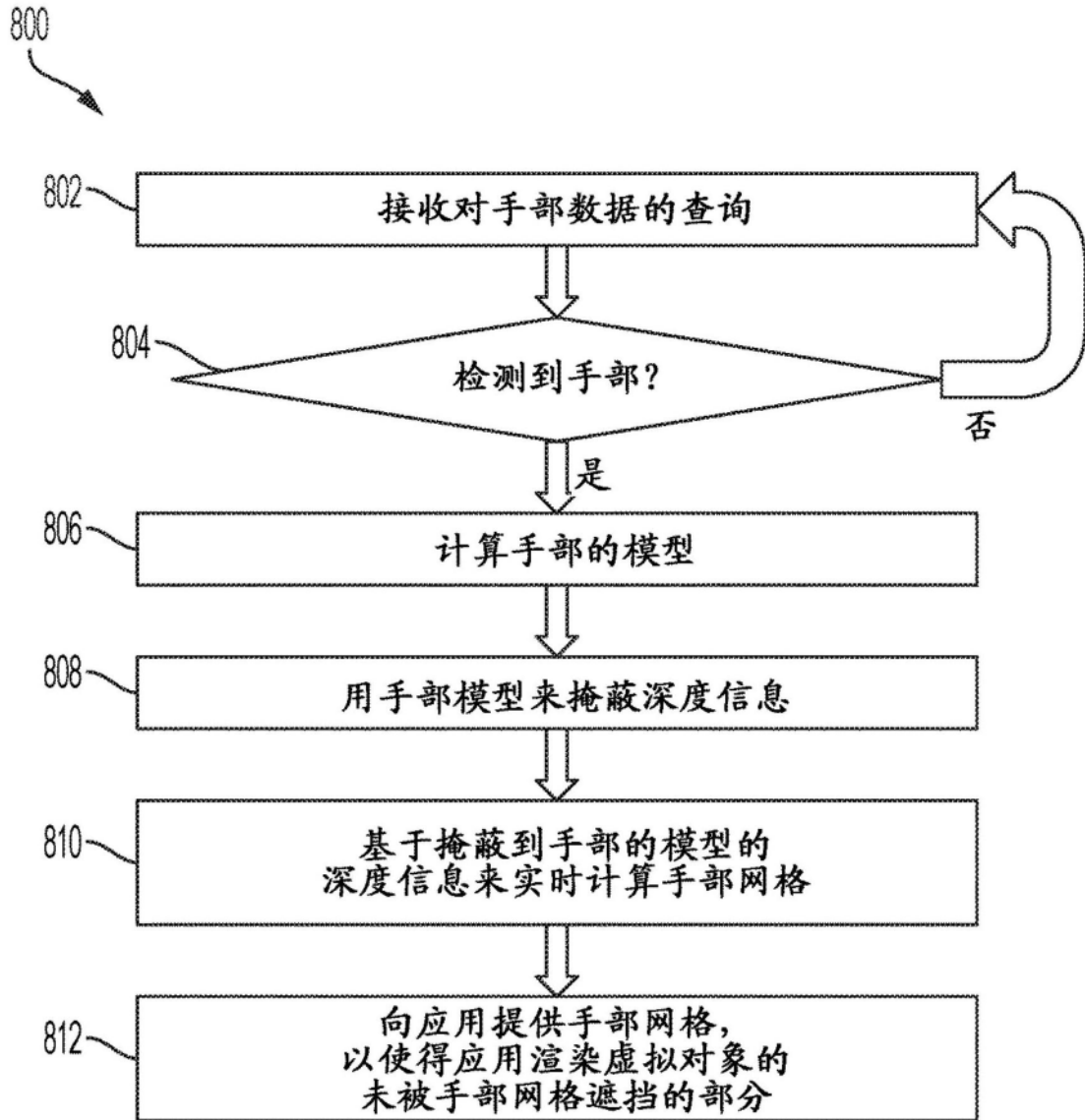


图8

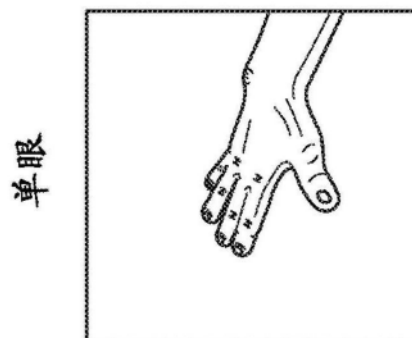


图9A

立体

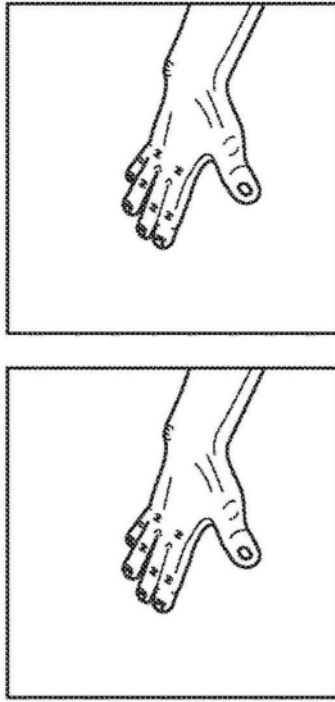


图9B

深度



图9C

分段

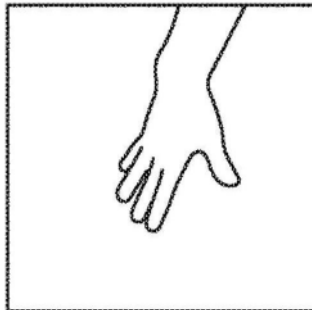


图9D

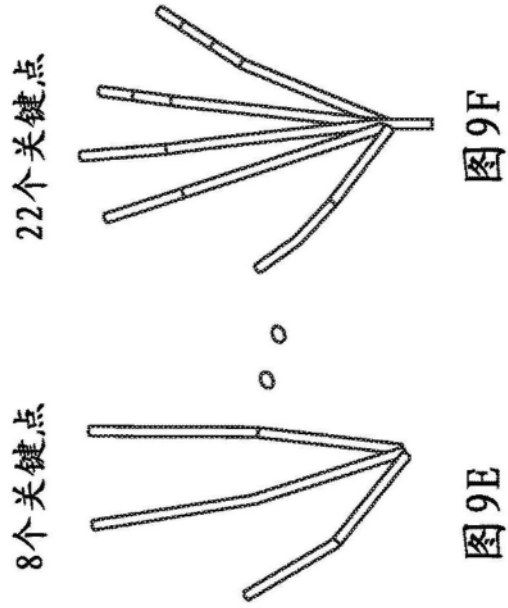


图9G

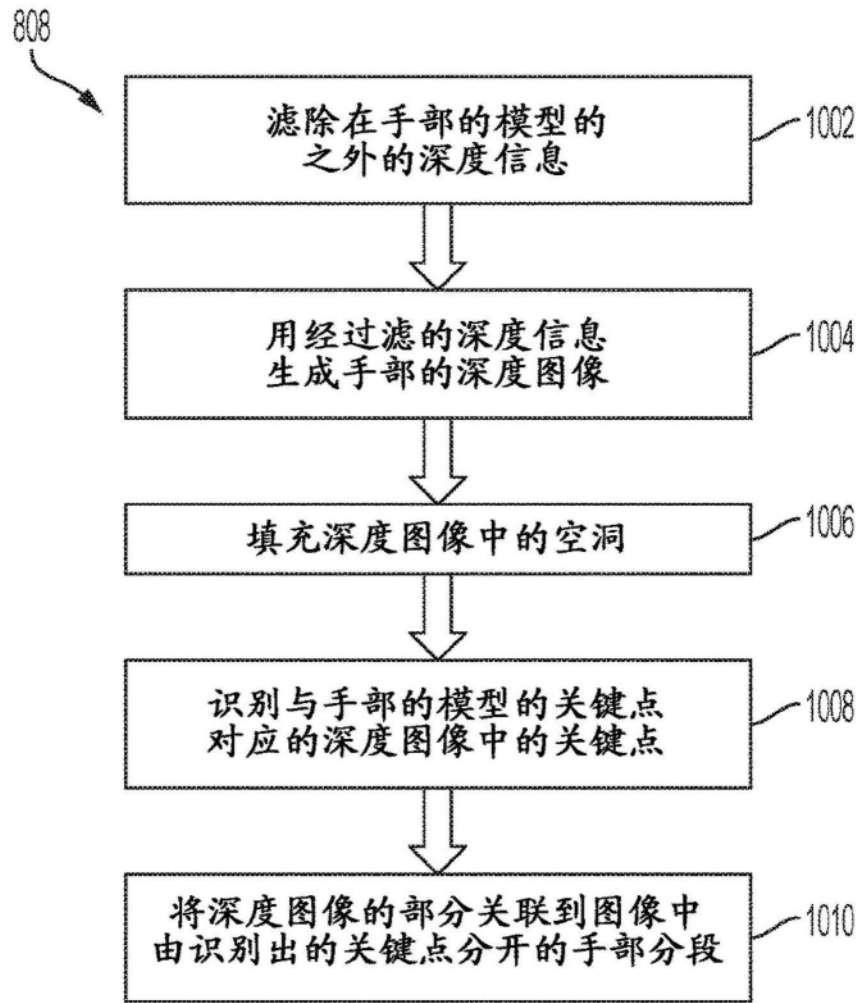


图10

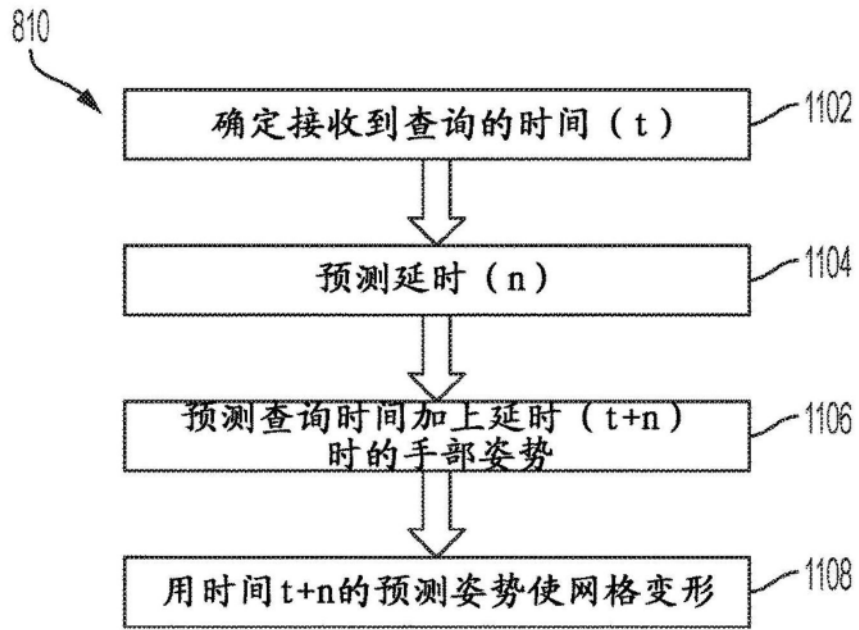


图11