



(12) 发明专利

(10) 授权公告号 CN 109643373 B

(45) 授权公告日 2023. 06. 27

(21) 申请号 201780053000.X

(22) 申请日 2017.05.17

(65) 同一申请的已公布的文献号
申请公布号 CN 109643373 A

(43) 申请公布日 2019.04.16

(30) 优先权数据
62/357,285 2016.06.30 US

(85) PCT国际申请进入国家阶段日
2019.02.27

(86) PCT国际申请的申请数据
PCT/US2017/033139 2017.05.17

(87) PCT国际申请的公布数据
W02018/004863 EN 2018.01.04

(73) 专利权人 奇跃公司
地址 美国佛罗里达州

(72) 发明人 A·克勒 G·布拉德斯基

(74) 专利代理机构 北京市中咨律师事务所
11247

专利代理师 贺月娇 杨晓光

(51) Int.Cl.
G06V 10/10 (2022.01)
H04N 13/398 (2018.01)
G06T 7/11 (2017.01)
H04N 13/344 (2018.01)
G01S 3/00 (2006.01)

(56) 对比文件
WO 2015161307 A1, 2015.10.22
US 2016026253 A1, 2016.01.28
US 2015234477 A1, 2015.08.20

审查员 赵立莎

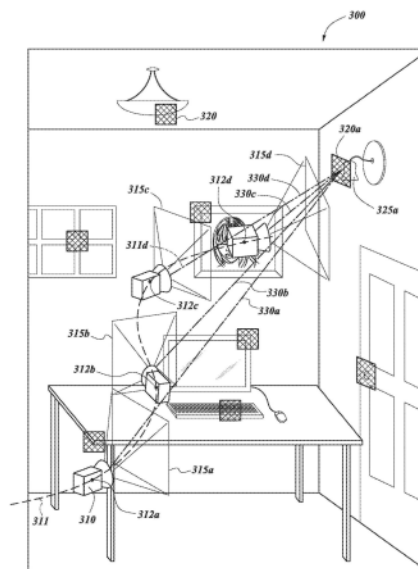
权利要求书3页 说明书22页 附图10页

(54) 发明名称

估计3D空间中的姿态

(57) 摘要

描述了用于估计设备在3D环境内的方位的方法和设备。这些方法的实施例包括顺序地接收形成图像的多个图像片段,所述图像表示包括环境的一部分的视场(FOV)。该图像包括多个稀疏点,所述多个稀疏点部分地基于所述多个图像片段中的对应图像片段子集而可辨识。该方法还包括在接收到与所述多个稀疏点中的一个或多个稀疏点对应的每个图像片段子集时顺序地辨识所述一个或多个稀疏点,并且基于所辨识的所述一个或多个稀疏点估计设备在环境内的方位。



1. 一种成像系统,包括:

图像捕获设备,其包括透镜和图像传感器,所述透镜被配置为将来自所述图像捕获设备周围的环境的光引导到所述图像传感器,所述图像传感器被配置为:

基于来自所述环境的光顺序地捕获图像的第一多个图像片段,所述图像表示所述图像捕获设备的视场FOV,所述FOV包括所述环境的一部分并且包括多个稀疏点,以及

顺序地捕获第二多个图像片段,所述第二多个图像片段是在所述第一多个图像片段之后被捕获的并形成所述图像的至少另一部分;

非暂时性数据存储装置,其被配置为从所述图像传感器顺序地接收所述第一多个图像片段和所述第二多个图像片段,并且存储用于估计所述图像捕获设备在所述环境内的方位和取向中的至少一者的指令;以及

至少一个硬件处理器,其可操作地耦接到所述非暂时性数据存储装置并且由所述指令配置为:

部分地基于所述第一多个图像片段的对应子集辨识第一组稀疏点,当在所述非暂时性数据存储装置处接收到所述第一多个图像片段时所述第一组稀疏点被辨识,

基于所述第一组稀疏点确定所述成像设备在所述环境内的方位和取向中的至少一者,

部分地基于所述第二多个图像片段的对应子集辨识第二组稀疏点,当在所述非暂时性数据存储装置处接收到所述第二多个图像片段时所述第二组稀疏点被辨识,以及

基于所述第一组稀疏点和所述第二组稀疏点更新所述成像设备在所述环境内的方位和取向中的所述至少一者。

2. 根据权利要求1所述的成像系统,其中所述图像传感器是滚动快门图像传感器。

3. 根据权利要求1所述的成像系统,其中所述非暂时性数据存储装置包括非暂时性缓冲存储装置,所述非暂时性缓冲存储装置被配置为在所述图像片段被所述图像传感器捕获时顺序地接收所述第一多个图像片段和所述第二多个图像片段,所述非暂时性缓冲存储装置具有至少部分地基于在每个图像片段子集中包括的图像片段的数量的存储容量。

4. 根据权利要求1所述的成像系统,其中所述第一组稀疏点或所述第二组稀疏点包括在10个稀疏点与20个稀疏点之间的稀疏点数量。

5. 根据权利要求1所述的成像系统,其中所述硬件处理器被配置为基于最近辨识的稀疏点的数量更新所述图像捕获设备的方位和取向中的所述至少一者,其中所述最近辨识的稀疏点包括所述第一组稀疏点中的至少一个、所述第二组稀疏点、以及所述第一组稀疏点和所述第二组稀疏点中的一个或多个。

6. 根据权利要求5所述的成像系统,其中所述最近辨识的稀疏点的数量等于所述第一组稀疏点中的稀疏点的数量。

7. 根据权利要求1所述的成像系统,其中所述硬件处理器被配置为执行视觉同步定位与地图构建(V-SLAM)算法。

8. 根据权利要求1所述的成像系统,其中所述多个稀疏点基于真实世界对象、虚拟图像元素和被投射到所述环境中的不可见指示符中的至少一者而被辨识。

9. 一种头戴式显示器HMD,其被配置为穿戴在用户的头部上,所述HMD包括:

框架;

显示器,其由所述框架支撑并且被设置在所述用户的眼睛前面;

面向外的图像捕获设备,其被设置在所述框架上并且包括透镜和图像传感器,所述透镜被配置为将来自所述HMD周围的环境的光引导到所述图像传感器,并且所述图像传感器被配置为基于来自所述环境的光顺序地捕获图像的多个图像片段,所述图像表示所述面向外的图像捕获设备的视场FOV,所述FOV包括环境的一部分并且包括多个稀疏点,其中每个稀疏点部分地基于所述多个图像片段的对应子集而可辨识;

非暂时性数据存储装置,其被配置为从所述图像传感器顺序地接收所述多个图像片段,并且存储用于估计所述HMD在所述环境内的方位和取向中的至少一者的指令;以及

至少一个硬件处理器,其可操作地耦接到所述非暂时性数据存储装置并且由所述指令配置为:

当在所述非暂时性数据存储装置处接收到与所述多个稀疏点中的一个或多个稀疏点对应的每个图像片段子集时,顺序地辨识所述一个或多个稀疏点,以及

基于所辨识的所述一个或多个稀疏点估计所述HMD在所述环境内的方位和取向中的所述至少一者。

10. 根据权利要求9所述的HMD,其中所述非暂时性数据存储装置包括循环缓冲器或滚动缓冲器。

11. 根据权利要求9所述的HMD,其中所述多个图像片段至少包括第一多个图像片段和第二多个图像片段,并且其中所述图像传感器被配置为将所述第一多个图像片段和所述第二多个图像片段顺序地发送到所述非暂时性数据存储装置。

12. 根据权利要求11所述的HMD,其中所述硬件处理器被配置为:

在接收到与第一组的一个或多个稀疏点对应的所述第一多个图像片段时,顺序地辨识所述第一组的一个或多个稀疏点;以及

在接收到与第二组的一个或多个稀疏点对应的所述第二多个图像片段时,顺序地辨识所述第二组的一个或多个稀疏点,

其中所述第二多个图像片段在所述第一多个图像片段之后被接收。

13. 根据权利要求12所述的HMD,其中所述硬件处理器被配置为基于所辨识的第一组的一个或多个稀疏点估计所述HMD的方位和取向中的所述至少一者。

14. 根据权利要求13所述的HMD,其中所述第一组的稀疏点或所述第二组的稀疏点包括2个与20个之间的稀疏点数量。

15. 根据权利要求13所述的HMD,其中所述第一组的稀疏点或所述第二组的稀疏点包括10个与20个之间的稀疏点数量。

16. 根据权利要求13所述的HMD,其中所述硬件处理器被进一步配置为基于所辨识的第二组的一个或多个稀疏点更新所述HMD的方位和取向中的所述至少一者。

17. 根据权利要求12所述的HMD,其中,所述硬件处理器被进一步配置为在一数量的所述顺序地辨识的一个或多个稀疏点被辨识时,更新所述HMD的方位和取向中的所述至少一者。

18. 根据权利要求17所述的HMD,其中所述数量的顺序地辨识的一个或多个稀疏点包括所述第一组的一个或多个稀疏点中的至少一个稀疏点。

19. 根据权利要求9所述的HMD,其中基于真实世界对象、虚拟图像元素和被投射到所述环境中的不可见指示符中的至少一者辨识所述多个稀疏点。

20. 根据权利要求9所述的HMD,其中所述硬件处理器被进一步配置为:
从所述多个图像片段的对应子集中提取所述顺序地辨识的一个或多个稀疏点;以及
对所述顺序地辨识的一个或多个稀疏点执行视觉同步定位与地图构建 (VSLAM) 算法,
以估计所述图像捕获设备的方位和取向中的至少一者。

估计3D空间中的姿态

[0001] 相关申请的交叉引用

[0002] 本申请根据根据35U.S.C.§119(e)要求2016年6月30日提交的名称为“ESTIMATING POSE IN 3D SPACE (估计3D空间中的姿态)”的美国临时专利申请No.62/357,285的优先权,该申请的整个公开内容通过引用并入本文中。

技术领域

[0003] 本公开涉及虚拟现实和增强现实成像及可视化系统,更特别地涉及三维(3D)空间中的稀疏(sparse)姿态估计。

背景技术

[0004] 现代计算和显示技术促进了用于所谓的“虚拟现实”或“增强现实”体验的系统开发,其中数字再现的图像或其部分以看起来是或可被感知为真实的方式呈现给用户。虚拟现实或“VR”场景通常涉及数字或虚拟图像信息的呈现而对其它实际的真实世界视觉输入不透明;增强现实或“AR”场景通常涉及将数字或虚拟图像信息呈现为对用户周围的实际世界的可视化的增强。例如,参考图1,描绘了增强现实场景1000,其中AR技术的用户看到以人、树木、背景中的建筑物和混凝土平台1120为特征的真实世界公园状设置1000。除了这些项目之外,AR技术的用户还感知他“看到”站在真实世界平台1120上的机器人雕像1110,以及看起来是大黄蜂的化身的正在飞行的卡通式的化身角色1130,尽管这些元素在真实世界中不存在。事实证明,人类视觉感知系统非常复杂,而且,产生促进除其他虚拟或真实世界图像元素之外的虚拟图像元素的舒适的、感觉自然的、丰富的呈现的VR或AR技术是具有挑战性的。本文公开的系统和方法解决了与VR和AR技术相关的各种挑战。

发明内容

[0005] 本公开的一方面提供了通过图像捕获设备在图像帧中捕获稀疏点时执行的稀疏姿态估计。因此,可以在捕获整个图像帧之前执行稀疏姿态估计。在一些实施例中,可以在捕获图像帧时细化或更新稀疏姿态估计。

[0006] 在一些实施例中,公开了用于估计图像捕获设备在环境内的位置的系统、设备和方法。在一些实施方式中,该方法可以包括顺序地接收第一组多个图像片段。所述第一组多个图像片段可以形成表示来自图像捕获设备前面的视场(FOV)的图像的至少一部分,该视场可以包括图像捕获设备周围的环境的一部分和多个稀疏点。每个稀疏点可以对应于图像片段子集。该方法还可以包括辨识(identify)第一组稀疏点,所述第一组稀疏点包括在接收所述第一组多个图像片段时辨识的一个或多个稀疏点。该方法然后可以包括通过位置估计系统基于所述第一组稀疏点确定所述图像捕获设备在环境内的位置。该方法还可以包括顺序地接收第二组多个图像片段,所述第二组多个图像片段可以在所述第一组多个图像片段之后被接收并且形成所述图像的至少另一部分。该方法然后可以包括辨识第二组稀疏点,所述第二组稀疏点可以包括在接收所述第二组多个图像片段时辨识的一个或多个稀疏

点。然后,该方法可以通过所述位置估计系统基于所述第一组稀疏点和所述第二组稀疏点更新所述图像捕获设备在所述环境内的位置。

[0007] 在一些实施例中,公开了用于估计图像捕获设备在环境内的位置的系统、设备和方法。在一些实施方式中,一种方法可包括顺序地接收多个图像片段,所述多个图像片段可以形成表示来自所述图像捕获设备前面的视场(FOV)的图像。所述FOV可以包括所述图像捕获设备周围的环境的一部分并且包括多个稀疏点。每个稀疏点可以是部分地基于所述多个图像片段的对应图像片段子集而可辨识的。该方法还可以包括:在接收与所述多个稀疏点中的一个或多个稀疏点对应的每个图像片段子集时,顺序地辨识一个或多个稀疏点。该方法然后可以包括基于所辨识的一个或多个稀疏点估计所述图像捕获设备在所述环境内的位置。

[0008] 在一些实施例中,公开了用于估计图像捕获设备在环境内的位置的系统、设备和方法。在一些实施方式中,所述图像捕获设备可以包括被配置为捕获图像的图像传感器。可以通过顺序地捕获表示所述图像捕获设备的视场(FOV)的多个图像片段来捕获图像。所述FOV可以包括所述图像捕获设备周围的环境的一部分和多个稀疏点。每个稀疏点可以部分地基于所述多个图像片段的对应子集而可辨识。所述图像捕获设备可以还包括:存储器电路,其被配置为存储与一个或多个稀疏点对应的图像片段子集;以及计算机处理器,其可操作地被耦接到所述存储器电路。所述计算机处理器可以被配置为在所述图像捕获设备接收与所述多个稀疏点中的一个或多个稀疏点对应的每个图像片段子集时,顺序地辨识所述一个或多个稀疏点。该计算机处理器还可以被配置为提取顺序地辨识的一个或多个稀疏点,以基于所辨识的一个或多个稀疏点估计所述图像捕获设备在所述环境内的位置。

[0009] 在一些实施例中,公开了用于估计图像捕获设备在环境内的位置的系统、设备和方法。在一些实施方式中,公开了一种增强现实系统。该增强现实系统可以包括面向外的成像设备、计算机硬件、以及可操作地耦接到所述计算机硬件和所述面向外的成像设备的处理器。该处理器可以被配置为执行指令以执行本文公开的方法的至少一部分。

[0010] 在一些实施例中,公开了用于估计图像捕获设备在环境内的位置的系统、设备和方法。在一些实施方式中,公开了一种自主实体(autonomous entity)。该自主实体可以包括面向外的成像设备、计算机硬件、以及可操作地耦接到所述计算机硬件和所述面向外的成像设备的处理器。该处理器可以被配置为执行指令以执行本文公开的方法的至少一部分。

[0011] 在一些实施例中,公开了用于估计图像捕获设备在环境内的位置的系统、设备和方法。在一些实施方式中,公开了一种机器人系统。该机器人系统可包括面向外的成像设备、计算机硬件、以及可操作地耦接到所述计算机硬件和所述面向外的成像设备的处理器。该处理器可以被配置为执行指令以执行本文公开的方法的至少一部分。

[0012] 所附权利要求的范围内的方法和装置的各种实施方式各自具有若干个方面,其中没有一个方面单独负责本文所述的所需属性。在不限制所附权利要求的范围的情况下,本文描述了一些突出特征。

[0013] 在附图和以下描述中阐述了本说明书中描述的主题的一个或多个实施方式的细节。通过说明书、附图和权利要求,其它特征、方面和优点将变得显而易见。本发明内容和以下具体实施方式都并非旨在限定或限制本发明主题的范围。

附图说明

[0014] 图1绘制了具有由人观看到的某些虚拟现实对象和某些实际的现实对象的增强现实场景的图示。

[0015] 图2示意性地示出了可穿戴显示系统的示例。

[0016] 图3示意性地示出了当成像设备在3D空间(在该示例中为房间)中移动时该成像设备的多个位置的示例。

[0017] 图4A和4B示意性地示出了针对图像帧的剪切效果的示例。

[0018] 图5A和5B示意性地示出了针对多个稀疏点的图4A和4B的剪切效果的示例。

[0019] 图6是示例性AR架构的框图。

[0020] 图7是用于姿态的示例性坐标系。

[0021] 图8是确定成像设备在3D空间内的姿态的方法的示例的过程流程图。

[0022] 图9A和9B示意性地示出了基于接收多个图像片段而从图像帧中提取一个或多个稀疏点的示例。

[0023] 图10是确定成像设备在3D空间内的姿态的方法的另一示例的过程流程图。

[0024] 在全部附图中,参考标号可以被重复使用以指示所提及的要素之间的对应关系。所提供的附图并非按比例绘制,是为了说明本文描述的示例性实施例而提供的,并非旨在限制本公开的范围。

具体实施方式

[0025] 概述

[0026] 当使用AR设备或在三维(3D)空间内移动的其它设备时,设备可能需要跟踪其通过3D空间的移动并对3D空间进行地图构建(map)。例如,AR设备可能因为用户的移动或者独立于用户(例如,机器人或其它自主实体)而在3D空间内四处移动,并且为了从其它虚拟图像元素或真实世界图像元素中突出显示虚拟图像元素,对3D空间进行地图构建并确定设备在3D空间内的位置(location)、方位(position)或取向(orientation)中的一者或多者以用于后续处理是有益的。例如,为了准确地呈现虚拟和真实世界的图像元素,设备可能需要知道它在真实世界中所处的位置和取向,并且在位于真实世界空间内的特定取向时准确地特定位置处呈现虚拟图像。在另一实施例中,可能需要再现到设备通过3D空间的轨迹。因此,当设备在3D空间内四处移动时,可能需要实时地确定设备在3D空间内的方位、位置或取向(下文统称为“姿态”)。在一些实施方式中,3D空间内的稀疏姿态估计可以根据来自成像设备的图像帧的连续流来确定,该成像设备作为例如AR设备的一部分而被包括。连续流中的每个图像帧可以被存储以用于处理,并且还用于由此估计设备的姿态以包括在稀疏姿态估计中。然而,由于将每个帧整体传送到存储器以执行后续处理,因此,这些技术可能导致估计姿态的延迟。

[0027] 本公开提供了被配置为估计设备(例如,AR设备或诸如机器人之类的自主设备)在3D空间内的姿态的示例性设备和方法。作为一个示例,当设备移动通过3D空间时,设备可以基于接收到多个图像帧以及根据每个图像帧估计设备的姿态来执行稀疏姿态估计。每个图像帧可以表示设备前面的3D空间的一部分,该部分指示设备在3D空间内的方位。在一些实施例中,每个图像帧可以包括能够由稀疏点、关键点、点云或其它类型的数学表示来表示的

一个或多个特征或对象。对于每个图像帧,可以通过顺序地接收多个图像片段来捕获图像帧,所述多个图像片段在被组合时构成整个图像帧。由此,该设备可以被配置为在接收到包括每个稀疏点的图像片段时辨识图像帧内的稀疏点。该设备可以提取包括一个或多个稀疏点的第一组稀疏点。第一组稀疏点可以是稀疏姿态估计过程的至少一个输入。随后,该设备可以辨识并提取第二组稀疏点,并基于第二组稀疏点更新稀疏姿态估计。在一种示例性实施方式中,在辨识后续稀疏点(例如,第二组稀疏点)之前,可以利用第一组稀疏点估计设备的姿态。后续稀疏点在被辨识时可用于更新稀疏姿态估计。

[0028] 虽然在此参考AR设备描述了方法、设备和系统的实施例,但是这并不旨在限制本公开的范围。在此描述的方法和设备不限于AR设备或头戴式设备;其它设备是可能的(例如,移动机器人、数字相机、自主实体等)。适用的设备包括但不限于这样的能够独立地或在用户介入下移动通过3D空间的设备。例如,在此描述的方法可以应用于在3D空间内四处移动的对象,该对象由对于该对象为远程的相机跟踪。在一些实施例中,还可以对对象远程地执行远处理。

[0029] 在3D空间内移动的示例性AR设备

[0030] 为了使3D显示器便于舒适、自然、丰富地从其它虚拟或真实世界的图像元素当中突出呈现虚拟图像元素,需要对显示器周围的真实世界进行地图构建并再现显示器通过3D空间的轨迹。例如,可以执行稀疏姿态估计处理以确定3D空间的地图。如果不以最小延迟实时地执行稀疏姿态估计,用户可能体验到不稳定的成像、有害的眼疲劳、头痛以及通常不愉快的VR和AR观看体验。因此,在此描述的各种实施例被配置为确定或估计AR设备的方位、位置或取向中的一者或多者。

[0031] 图2示出了可穿戴显示系统100的示例。显示系统100包括显示器62,以及支持该显示器62的功能的各种机械和电子模块和系统。显示器62可以被耦接到框架64,该框架可由显示系统用户、穿戴者或观看者60穿戴,并且可以被配置为将显示器62定位在用户60的眼睛前面。显示系统100可以包括戴在穿戴者的头上的头戴式显示器(HMD)。增强现实显示器(ARD)可以包括可穿戴显示系统100。在一些实施例中,扬声器66被耦接到框架64并且被定位在用户的耳道附近(在一些实施例中,另一扬声器(未示出)可以被定位在用户的另一耳道附近以提供立体声/可塑形的声音控制)。显示系统100可以包括一个或多个面向外的成像系统110,该成像系统观察穿戴者周围的环境(例如,3D空间)中的世界。显示器62可以通过通信链路68,例如通过有线引线或无线连接,可操作地耦接到本地处理和数据模块70,本地处理和数据模块70可以以各种配置安装,例如固定地附接到框架64,固定地附接到用户穿戴的头盔或帽子,嵌入在耳机中,或以其它方式可移除地附接到用户60(例如,采取背包式配置,采取腰带耦接式配置)。

[0032] 显示系统100可以包括设置在框架64上的一个或多个面向外的成像系统110a或110b(在下文中单独或统称为“110”)。在一些实施例中,面向外的成像系统110a可以设置在框架64的位于用户眼睛之间的大约中心部分处。在另一实施例中,替代地或组合地,面向外的成像系统110b可以设置在框架的与用户的一只或两只眼睛相邻的一侧或多侧上。例如,面向外的成像系统110b可以位于用户的左侧和右侧,与两只眼睛相邻。虽然上面提供了面向外的相机110的示例性布置,但是其它配置也是可能的。例如,面向外的成像系统110可以被定位在相对于显示系统100的任何取向或方位上。

[0033] 在一些实施例中,面向外的成像系统110捕获显示系统100前面的世界的一部分的图像。可供观看者观看或成像的整个区域可以被称为能视域(field of regard, FOR)。在一些实施方式中, FOR可以包括显示系统100周围的基本上所有立体角,因为显示器可以在环境内四处移动以对显示器周围(在穿戴者的前面、后面、上面、下面或侧面)的对象进行成像。显示系统前面的FOR的部分可以被称为视场(FOV),并且面向外的成像系统110有时被称为FOV相机。从面向外的成像系统110获得的图像可用于辨识环境的稀疏点并估计用于稀疏姿态估计处理的姿态等等。

[0034] 在一些实施方式中,面向外的成像系统110可以被配置为包括光学透镜系统和图像传感器的数字相机。例如,来自显示器62前面的世界(例如,来自FOV)的光可以被面向外的成像系统110的透镜聚焦到图像传感器上。在一些实施例中,面向外的成像系统100可以被配置为在红外(IR)光谱、可见光光谱或任何其它合适的波长范围或电磁辐射波长范围中工作。在一些实施例中,成像传感器可以被配置为CMOS(互补金属氧化物半导体)或CCD(电荷耦合器件)传感器。在一些实施例中,图像传感器可以被配置为检测IR光谱、可见光谱或任何其它合适的波长范围或电磁辐射波长范围中的光。在一些实施例中,数字相机的帧速率可涉及图像数据可以从数字相机发送到存储器或存储单元(例如,本地处理和数据模块70)的速率。例如,如果数字相机的帧速率是30赫兹,则由图像传感器的像素捕获的数据可以每30毫秒被读入存储器中(例如,时钟控制(clock off))。因此,数字相机的帧速率可能给图像数据的存储和后续处理带来延迟。

[0035] 在一些实施例中,在面向外的成像系统110是数字相机的情况下,面向外的成像系统110可以被配置为全局快门相机或滚动快门(rolling shutter)(例如,也被称为逐行扫描相机)。例如,在面向外的成像系统110是全局快门相机的情况下,图像传感器可以是CCD传感器,其被配置为通过单个操作捕获代表显示器62前面的FOV的整个图像帧。然后将整个图像帧读入本地处理和数据模块70以执行处理,例如执行如本文所述的稀疏姿态估计。因此,在一些实施例中,例如如上所述,由于帧速率和存储图像中的延迟,利用整个图像帧可在姿态估计中产生延迟。例如,具有30赫兹帧速率的全局快门数字相机可以在可以执行任何姿态估计之前产生30毫秒的延迟。

[0036] 在其它实施例中,在面向外的成像系统110被配置为滚动快门相机的情况下,图像传感器可以是CMOS传感器,其被配置为顺序地捕获多个图像片段并扫描场景以发送所捕获的图像片段的图像数据。当以捕获的顺序组合时,图像片段构成面向外的成像系统110的FOV的图像帧。在一些实施例中,扫描方向可以是水平的,例如,面向外的成像系统110可以捕获在向左或向右的方向上水平相邻的多个垂直图像片段。在另一实施例中,扫描方向可以是竖直的,例如,面向外的成像系统110可以捕获在向上或向下方向上竖直相邻的多个水平图像片段。当在图像传感器处捕获相应的图像片段时,可以将每个图像片段顺序地读入本地处理和数据模块70中。因此,在一些实施例中,如上所述,由数字相机的帧速率导致的延迟可以通过在数字相机捕获图像片段时顺序地发送图像片段而被减少或最小化。

[0037] 本地处理和数据模块70可以包括一个或多个硬件处理器,以及数字存储器,诸如非易失性存储器(例如,闪存),这两者都可用于辅助数据的处理、缓冲、缓存和存储。数据可以包括a)从传感器(其可以例如可操作地耦接到框架64或以其它方式附接到用户60)捕获的数据,传感器诸如是图像捕获设备(面向外的成像系统110)、麦克风、惯性测量单元

(IMU)、加速度计、指南针、全球定位系统(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/或远程数据储存库74获取和/或处理的数据,可能用于在这样的处理或检索之后向显示器62传送。本地处理和数据模块70可以通过通信链路76和/或78(诸如经由有线或无线通信链路)可操作地耦接到远程处理模块72和/或远程数据储存库74,使得这些远程模块可用作本地处理和数据模块71的资源。此外,远程处理模块72和远程数据储存库74可以操作地彼此耦接。在一些实施例中,本地处理和数据模块70可以可操作地耦接到图像捕获设备、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/或陀螺仪中的一者或多者。在一些其它实施例中,这些传感器中的一者或多者可以被附接到框架64,或者可以是通过有线或无线通信路径与本地处理和数据模块70通信的独立结构。

[0038] 在一些实施例中,本地处理和数据模块70的数字存储器或其一部分可以被配置为持续一临时时间段存储数据的一个或多个元素(例如,作为非暂时性缓冲存储装置)。例如,数字存储器可以被配置为接收部分或所有数据并短期存储部分或所有数据,同时数据在本地处理和数据模块70的进程之间移动。在一些实施方式中,数字存储器的一部分可以被配置为缓冲器,该缓冲器从面向外的成像系统110顺序地接收一个或多个图像片段。因此,该缓冲器可以是非暂时性数据缓冲器,其被配置为在设定数量的图像片段被发送到本地处理和数据模块70(或远程数据储存库74)以进行永久存储或后续处理之前,存储这些图像片段(如下面参考图9A和9B所述)。

[0039] 在一些实施例中,远程处理模块72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硬件处理器。在一些实施例中,远程数据储存库74可以包括数字数据存储设施,该设施可以通过因特网或“云”资源配置中的其它网络配置获得。在一些实施例中,远程数据储存库74可以包括一个或多个远程服务器,远程服务器将信息(例如,用于产生增强现实内容的信息)提供给本地处理和数据模块70和/或远程处理模块72。在一些实施例中,所有数据都被存储,所有计算都在本地处理和数据模块70中执行,允许从远程模块完全自主地使用。

[0040] 虽然在此描述了示例性AR设备,但是应当理解,在此公开的方法和设备不限于AR设备或头戴式设备。其它配置是可能的,例如,移动机器人、数字相机、自主实体等。适用的设备包括但不限于这样的能够独立地或通过使用介入而移动通过3D空间的设备。

[0041] 通过3D空间的AR设备的示例性轨迹

[0042] 图3示意性地示出了移动通过3D空间300的成像设备310。例如,图3示出了当成像设备310沿着示意性地表示轨迹311的虚线移动时,位于环境300内的多个方位312(例如,312a、312b、312c和312d)和取向处的成像设备310。在每个方位312处,成像设备310可以被配置为捕获特定位置和取向的环境300的图像帧,其可以用作连续的图像帧流,例如用于执行稀疏姿态估计。轨迹311可以是移动通过环境300的任何轨迹或路径。尽管图3示出了四个方位312,但方位数量可以不同。例如,方位312的数量可以少到两个方位,也可以是可接受的确定性水平执行稀疏姿态估计所需的那么多的方位(例如,5、6、7个等)。在一些实施例中,成像设备312可以被配置为捕获一系列图像帧,例如,如在视频中,其中视频的每个图像帧可以被用于通过如本文所述的计算机视觉技术执行稀疏姿态估计。

[0043] 在一些实施例中,成像设备310可以被配置为图1中的显示系统100,包括面向外的成像系统110、内含成像系统的移动机器人,或者被配置为独立的成像设备。成像设备310可

以被配置为在其移动通过环境300时,在每个方位312处捕获图像帧,这些图像帧描绘环境300的来自成像设备310前面的部分。如上所述,由成像设备在每个方位312和取向处捕获的环境300的部分可以是来自成像设备310前面的FOV。例如,方位312a的FOV被示意性地示出为FOV 315a。成像设备310的每个后续方位和取向(例如,312b、312c和312d)包括对应的FOV 315(例如,FOV 315b、315c和315d)。可以对从成像设备310获得的每个图像帧执行计算机视觉技术,以估计每个方位312处的成像设备310的姿态。姿态估计可以是例如用于确定或产生环境300的地图(或其部分)并跟踪成像设备310通过环境300的移动的稀疏点估计过程的输入。

[0044] 环境300可以是任何3D空间,例如,办公室(如图3所示)、起居室、室外空间等。环境300可以包括在整个环境300内设置的多个对象325(例如,家具、个人物品、周围结构、纹理、可检测图案等)。对象325可以是与该环境中的其它特征(例如,每堵墙可以不是可唯一辨识的)相比能够唯一辨识的个体对象。此外,对象325可以是在两个或更多个图像帧中捕获的共同特征。例如,图3示出了在沿着相应视线330a-d(为了说明的目的,示出为虚线)在每个方位312处,位于成像设备310的每个FOV315中的对象325a(在该示例中为灯)。因此,对于每个方位312(例如,312a),表示每个FOV 315(例如,315a)的图像帧包括沿着视线330(例如,330a)成像的对象325a。

[0045] 成像设备310可以被配置为从表示FOV 315的每个图像帧中检测并提取多个稀疏点320,每个稀疏点320(或多个稀疏点)与对象325或对象325的部分、纹理或图案对应。例如,成像设备310可以提取与对象325a对应的稀疏点320a。在一些实施例中,对象325a可以与一个或多个稀疏点320相关联,其中每个稀疏点320可以与对象325的不同部分(例如,灯的角部、顶部、底部、侧面等)相关联。因此,每个稀疏点320可以在图像帧内被唯一地辨识。计算机视觉技术可用于从与每个稀疏点320对应的图像帧或图像片段中提取和辨识每个稀疏点320(例如,如下面结合图9A和9B所述)。

[0046] 在一些实施例中,稀疏点320可用于估计成像设备310在环境300内的方位、位置或取向。例如,成像设备310可以被配置为提取多个稀疏点320作为稀疏姿态估计处理的输入。用于稀疏姿态估计的示例性计算机视觉技术可以是同步定位与地图构建(SLAM或V-SLAM,是指其中输入只是图像/视觉的配置)处理或算法。这样的示例性计算机视觉技术可用于输出成像设备310周围世界的稀疏点表示,如下面更详细描述。在使用方位312的多个图像帧的传统稀疏姿态估计系统中,可以从每个图像帧收集稀疏点320,计算连续图像帧(例如,方位312a至312b)之间的对应关系,并且基于所发现的对应关系估计姿态变化。因此,在一些实施例中,可以确定成像设备310的方位和/或取向。在一些实施例中,稀疏点位置的3D地图是估计处理所需的,或者可以是辨识一图像帧或多个图像帧中的稀疏点的副产物。在一些实施例中,稀疏点320可以与一个或多个描述符相关联,所述描述符可以被配置为稀疏点320的数字表示。在一些实施例中,描述符可以被配置为便于计算连续图像帧之间的对应关系。在一些实施例中,姿态确定可以由位于成像设备上的处理器(例如,本地处理和数据模块70),或者位于成像设备远程的处理器(例如,远程处理模块72)执行。

[0047] 在一些实施例中,可以以可操作地与成像设备310进行通信的方式(例如作为本地处理和数据模块70或远程处理模块和数据储存库72、74的一部分)包括计算机视觉模块。示例性计算机视觉模块可以实现一种或多种计算机视觉技术,并且可以用于分析由面向外的

成像相机获得的图像片段,例如以辨识稀疏点,确定姿态等等,例如,如参考图8和10的方法800、100所描述的。计算机视觉模块可以辨识成像设备310周围环境中的对象,例如结合图3描述的那些。当成像设备在环境中移动时,计算机视觉模块可以从图像帧中提取稀疏点并使用所提取的稀疏点跟踪和辨识通过各种图像帧的对象。例如,可以将第一图像帧的稀疏点与第二图像帧的稀疏点进行比较,以跟踪成像设备的移动。在一些实施例中,第二图像帧的一个或多个稀疏点可以包括第一图像帧的一个或多个稀疏点,例如,作为用于在第一和第二图像帧之间进行跟踪的参考点。可以类似地使用第三、第四、第五等图像帧,并将其与先前和后续的图像帧进行比较。计算机视觉模块可以处理稀疏点以基于所辨识的稀疏点估计成像设备在环境内的方位或取向。计算机视觉技术的非限制性示例包括:尺度不变特征变换(SIFT)、加速鲁棒特征(SURF)、定向FAST和旋转BRIEF(ORB)、二进制鲁棒不变可扩展关键点(BRISK)、快速视网膜关键点(FREAK)、Viola-Jones算法、特征脸方法、Lucas-Kanade算法、Horn-Schunk算法、均值漂移(Mean-shift)算法、视觉同步定位与地图构建(v-SLAM)技术、序列贝叶斯估计器(例如,卡尔曼滤波器、扩展卡尔曼滤波器等)、光束法平差(bundle adjustment)、自适应阈值分割(和其它阈值分割技术)、迭代最近点(ICP)、半全局匹配(SGM)、半全局块匹配(SGBM)、特征点直方图、各种机器学习算法(例如,支持向量机、k最近邻算法、朴素贝叶斯、神经网络(包括卷积或深度神经网络)或其它有监督/无监督模型等)等等。

[0048] 如上所述,当前姿态估计处理可能包括在估计成像设备姿态时的延迟。例如,部分地由于将整个图像帧从成像设备传送到存储器,成像设备的帧速率可能导致延迟。不受任何特定科学理论的束缚,稀疏姿态估计可能被延迟,这是因为直到将整个图像帧从成像设备读取到存储器为止,不从图像帧中提取稀疏点。因此,部分地基于成像设备的帧速率能力传送整个图像帧可能成为稀疏姿态估计所经历的延迟的一个因素。在此描述的系统和设备中的某些的一个非限制性优点是,当部分图像帧被读入图像传感器或存储器中时,可以在飞行中(on the fly)执行用于估计姿态的稀疏点的提取或辨识,因此,当使用整个图像帧时,可以在早于其它可能的时间点的时间点处估计姿态。此外,由于针对关键点仅分析帧的一部分,因此可以提高处理速度和效率。

[0049] 虽然前面的描述在环境300中的物理对象的上下文中描述了稀疏点320,但是这并不旨在为限制性的,其它实施方式是可能的。在一些实施例中,对象325可以指环境的任何特征(例如,真实世界对象、虚拟对象、不可见对象或特征等)。例如,投影设备可以被配置为在整个环境中投射多个指示符(indicator)、纹理、标识符(identifier)等,它们可以是可见的或不可见的(例如,在IR光谱、近红外光谱,紫外光谱,或者在任何其它合适的波长范围或电磁辐射波长范围内投射的)。指示符、纹理、标识符等可以是能够被成像设备310检测到的独特特征或形状。成像设备310可以被配置为检测这些指示符并根据多个指示符提取稀疏点320。例如,指示符可以在电磁辐射的IR光谱中被投射在环境的壁上,并且成像设备310可以被配置为在IR光谱中操作以辨识指示符并根据其提取稀疏点。在另一实施例中,替代地或组合地,成像设备310可以被包括在被配置为显示虚拟图像元素(例如,在显示器62上)的AR设备中。成像设备或AR设备可以被配置为辨识虚拟图像元素并从中提取稀疏点320。AR设备可以被配置为使用这些稀疏点320来确定AR设备相对于虚拟图像元素的姿态。

[0050] 被赋予到示例性图像帧和稀疏点的剪切效果的示例

[0051] 如上所述,面向外的成像系统110可以被实现为滚动快门相机。滚动快门相机的一个非限制性优点是能够发送所捕获的场景的部分(例如,图像片段)并同时捕获场景的其它部分(例如,并非图像帧的所有部分都在完全相同的时间被捕获)。然而,这可能导致在捕获图像帧的同时相对于相机移动的对象失真,因为在捕获图像的整个时间内,成像设备可能相对于对象不位于同一位置。

[0052] 例如,图4A和4B是应用于场景图像的滚动快门效果(例如,有时在此被称为“剪切”、“移位”或“失真”)的示意图。图4A示意性地示出了包括对象425a(例如,在该示例中为正方形)的场景400a。该场景可以是图像捕获设备(例如,图2中的面向外的成像系统110)的FOV。在图4A所示的实施例中,场景可以在方向430上相对于图像捕获设备移动。图4B示出了所捕获的场景400a的产生的图像400b,其可以被存储在存储器或存储单元(例如,本地处理和数据模块70)中。如图4B所示,由于对象425a的相对移动,产生的图像400b是失真对象425b(例如,显示为剪切的正方形或菱形),其中失真对象的虚线未被捕获在产生的图像400b中。不受任何特定科学理论的束缚,这可能是由成像设备逐行向下扫描方向造成的,因此对象的顶部首先被捕获,其失真度小于对象的底部。

[0053] 图5A和5B是被赋予到由成像设备捕获的FOV(例如,图3中的FOV315a、315b、315c或315d)中包括的多个稀疏点上的滚动快门效果的示意图。例如,当AR设备在3D空间内四处移动时,各种稀疏点相对于AR设备移动并且以类似于上面结合图4B描述的方式,如图5B中示意性地示出的那样失真。图5A示出了包括多个稀疏点320(例如,320a、320b和320c)的场景(例如,类似于图3中的场景300)。图4B示意性地示出了产生的捕获图像帧包括失真稀疏点525(例如,525a、525b和525c)。例如,每个失真稀疏点525与说明性的对应箭头522相关联。仅出于说明性目的,箭头522的大小与赋予稀疏点525的失真量成比例。因此,类似于上面结合图4B所述,箭头522a小于箭头522e,这可以指示与稀疏点525e相比,与箭头522a相关联的稀疏点525a的失真度较不严重。

[0054] 示例性AR架构

[0055] 图6是AR架构600的示例的框图。AR架构600被配置为从一个或多个成像系统接收输入(例如,来自面向外的成像系统110的视觉输入、来自房间相机的输入等)。成像设备不仅提供来自FOV相机的图像,还可以配备有各种传感器(例如,加速度计、陀螺仪、温度传感器、运动传感器、深度传感器、GPS传感器等)以确定位置和用户环境的各种其它属性。该信息可以进一步由可以从不同的视点提供图像和/或各种线索的来自房间中的固定相机的信息进行增补。

[0056] AR架构600可以包括多个相机610。例如,AR架构600可以包括图1中的面向外的成像系统110,该面向外的成像系统110被配置为输入来自可穿戴显示系统100前面的FOV的多个已捕获图像。在一些实施例中,相机610可以包括相对宽的视场或设置在用户脸部侧面的无源相机对,以及被定位在用户前面以处理立体成像过程的不同相机对。然而,其它成像系统、相机和布置也是可能的。

[0057] AR架构600还可以包括地图数据库630,其包括世界的地图数据。在一个实施例中,地图数据库630可以部分地驻留在用户可穿戴系统(例如,本地处理和数据模块70)上,或者可以部分地驻留在可通过有线或无线网络访问的网络存储位置(例如,远程数据储存库74)处。在一些实施例中,地图数据库630可以包括真实世界地图数据或虚拟地图数据(例如,包

括定义虚拟地图或覆盖在现实世界环境上的虚拟图像元素)。在一些实施例中,计算机视觉技术可用于生成地图数据。在一些实施例中,地图数据库630可以是预先存在的环境地图。在其它实施例中,可以基于所辨识的稀疏点来填充地图数据库630,这些稀疏点被读入存储器中并被存储以相对于随后辨识的稀疏点进行比较和处理。在另一实施例中,单独地或组合地,地图数据库630可以是预先存在的地图,其基于来自一个或多个图像帧(或滚动快门相机系统的帧的部分)的所辨识的稀疏点而被动态地更新。例如,一个或多个稀疏点可用于辨识环境中的对象(例如,图3中的对象325)并且用于用环境的辨识特征来填充该地图。

[0058] AR架构600可以还包括缓冲器620,其被配置为从相机610接收输入。缓冲器620可以是非暂时性数据缓冲器,例如与非暂时性数据存储装置(例如,图2中的本地处理和数据模块70)分离或者是非暂时性数据存储装置的一部分,并且被配置为临时存储图像数据。然后,缓冲器620可以临时存储一些或所有接收的输入。在一些实施例中,缓冲器620可以被配置为在例如执行进一步处理并将数据移到AR架构600的另一组件之前存储所接收的数据的一个或多个部分或片段(例如,如下面结合图9A和9B所述)。在一些实施例中,当用户体验在环境中操作的可穿戴显示系统100时,可以将由相机610收集的图像数据读入缓冲器620中。这样的图像数据可以包括由相机610捕获的图像或图像片段。然后,表示图像或图像片段的图像数据可以首先在由本地处理和数据模块处理之前被发送到并存储在缓冲器620中,并且被发送到显示器62以进行可视化并向可穿戴显示系统100的用户表示。替代地或组合地,图像数据也可以被存储在地图数据库630中。或者,在被存储到缓冲器620中之后,数据可以从存储器(例如,本地处理和数据模块70或远程数据储存库74)中被移除。在一个实施例中,缓冲器620可以部分地驻留在用户可穿戴系统(例如,本地处理和数据模块70)上,或者可以部分地驻留在可通过有线或无线网络访问的网络存储位置(例如,远程数据储存库74)处。

[0059] AR架构600可以还包括一个或多个对象识别器650。对象识别器可以被配置为爬过(crawl through)所接收的数据并辨识和/或标记对象,然后借助于地图数据库630例如经由计算机视觉技术将信息附到对象上。例如,对象识别器可以扫描或爬过存储在缓冲器620中的图像数据或图像片段,并辨识在图像数据中捕获的对象(例如,图3中的对象325)。在缓冲器中辨识的对象可以参考地图数据库而被标记或描述信息被附加到其上。地图数据库630可以包括在所捕获的图像数据与其对应的对象之间随时间辨识的各种对象(例如,在第一图像帧中辨识的对象与在后续图像帧中辨识的对象的比较)以生成地图数据库630或用于生成环境地图。在一些实施例中,地图数据库630可以用预先存在的环境地图填充。在一些实施例中,地图数据库630存储在AR设备(例如,本地处理和数据模块70)上。在其它实施例中,AR设备和地图数据库可以通过网络(例如,LAN、WAN等)彼此连接以访问云存储(例如,远程数据储存库74)。

[0060] 在一些实施例中,AR架构600包括姿态估计系统640,其被配置为执行指令以部分地基于存储在缓冲器620和地图数据库630中的数据来执行姿态估计处理,以确定可穿戴计算硬件或设备的位置和取向。例如,当用户正在体验可穿戴设备并在世界中操作时,可以在由相机610收集的数据被读入缓冲器620时根据该数据计算方位、位置或取向数据。例如,基于从数据辨识并被存储在缓冲器620中的信息和对象集合,对象识别器610可识别(recognize)对象325并将这些对象作为稀疏点320提取到处理器(例如,本地处理和数据模块70)。在一些实施例中,当给定图像帧的一个或多个图像片段被读入缓冲器620中并被用

于估计AR设备在相关联的图像帧中的姿态时,可以提取稀疏点320。当该图像帧的另外的图像片段被读入缓冲器620中并被用于辨识另外的稀疏点时,可以更新姿态的估计。可选地,在一些实施例中,姿态估计系统640可以访问地图数据库630并检索在先前捕获的图像片段或图像帧中辨识的稀疏点320,并在AR设备移动通过3D空间时比较先前图像帧和后续图像帧之间的对应稀疏点320,从而跟踪AR设备在3D空间中的移动、方位或取向。例如,参考图3,对象识别器650可以在多个图像帧中的每一者中识别作为灯325a的稀疏点320a。AR设备可以附加一些描述符信息以将一个图像帧中的稀疏点320a与其它图像帧中的对应稀疏点320a进行关联,并将该信息存储在地图数据库650中。对象识别器650可以被配置为针对任何数量的稀疏点320(例如,1、2、3、4个等等稀疏点)识别对象。

[0061] 一旦对象被识别,姿态估计系统640就可以使用该信息来确定AR设备的姿态。在一个实施例中,对象识别器650可以在图像片段被接收时辨识与图像片段对应的稀疏点,并且随后可以在同一图像帧的后续图像片段被接收时辨识另外的稀疏点。姿态估计系统640可以执行指令以基于第一辨识的稀疏点来估计姿态,并且通过将随后辨识的稀疏点整合到估计处理中来更新估计。在另一实施例中,单独地或组合地,对象识别器650可以识别第一帧中的两个对象(例如,图3所示的对象325a和另一对象)的两个稀疏点320a、320b,然后辨识第二帧和后续帧中相同的两个稀疏点(例如,可以考虑最多为任何数量的后续帧)。基于两个或更多个帧中的稀疏点之间的比较,3D空间内的姿态(例如,取向和位置)也可以被估计或被跟踪通过该3D空间。

[0062] 在一些实施例中,姿态估计的精度或姿态估计结果的降噪可以基于由对象识别器640识别的稀疏点的数量。例如,在3D空间中,成像设备的方位、位置或取向可以基于环境内的平移和旋转坐标。这样的坐标可以包括例如X、Y和Z平移坐标或偏航、翻滚、俯仰旋转坐标,如下面结合图7所述。在一些实施例中,从图像帧提取的一个稀疏点可能无法传达成像设备的完整姿态。然而,单个稀疏点可以是指对姿态估计的至少一个约束,例如,通过提供一个或多个坐标有关的信息。随着稀疏点数量的增加,姿态估计的精度可以提高,或者姿态估计中的噪声或误差可以降低。例如,两个稀疏点可以基于由稀疏点表示的对象而指示成像设备在3D空间中的X、Y位置。然而,成像设备可能无法确定其相对于对象的Z位置(例如,在对象的前面还是后面)或其翻滚坐标。因此,在一些实施例中,可以使用三个稀疏点来确定姿态,然而,可以使用任何数量的稀疏点(例如,1、2、4、5、6、7、10个或更多个等等)。

[0063] 在一些实施例中,姿态确定可以由AR设备上的处理器(例如,本地处理和数据模块70)执行。所提取的稀疏点可以被输入到被配置为执行计算机视觉技术的姿态估计系统640中。在一些实施例中,姿态估计系统可以包括由姿态估计系统640执行的SLAM或V-SLAM(例如,指其中输入只是图像/视觉的配置),然后姿态估计系统640可以输出AR设备周围的世界的稀疏点表示670。在一些实施例中,姿态估计系统640可以被配置为执行连续更新的递归贝叶斯估计器(例如,卡尔曼滤波器)。然而,贝叶斯估计器旨在作为用于由姿态估计系统640执行姿态估计的至少一种方法的说明性示例,在本公开的范围内可以设想其它方法和处理。该系统可以被配置为不仅找出各种组件在世界中的位置,而且还找出世界的构成。姿态估计可以是实现许多目标的构建块,其包括填充地图数据库630并使用来自地图数据库630的数据。在其它实施例中,AR设备可以被连接到处理器,该处理器被配置为通过网络(例如,LAN、WAN等)访问云存储(例如,远程数据储存库74)来执行姿态估计。

[0064] 在一些实施例中,一个或多个远程AR设备可以被配置为基于包括AR架构600的单个AR设备的姿态确定来确定每个AR设备的姿态。例如,一个或多个AR设备可以与包括AR架构600的第一AR设备进行有线或无线通信。第一AR设备可以基于从环境中提取的稀疏点来执行姿态确定,如本文所述。第一AR设备还可以被配置为发送可以由一个或多个远程AR设备(例如,第二AR设备)接收的辨识信号(例如,IR信号或其它合适的介质)。在一些实施例中,第二AR设备可以尝试显示与第一AR设备类似的内容,并从第一AR设备接收辨识信号。根据辨识信号,第二AR设备能够确定(例如,解释或处理辨识信号)其相对于第一AR设备的姿态,而无需提取稀疏点并对第二AR设备执行姿态估计。该布置的一个非限制性优点是可以通过链接两个AR设备,避免在第一和第二AR设备上显示的虚拟内容的差异。该布置的另一非限制性优点是第二AR系统能够基于从第一AR设备接收的辨识信号更新其估计的方位。

[0065] 成像设备姿态和坐标系的示例

[0066] 图7是用于成像设备姿态的坐标系的示例。设备700可具有多个自由度。当设备700朝着不同方向移动时,设备700的方位、位置或取向将相对于起始位置720而改变。图7中的坐标系示出了移动的三个平移方向(例如,X、Y和Z方向),其可用于测量相对于设备的起始位置720的设备移动,以确定3D空间内的位置。图7中的坐标系还示出了可用于测量相对于设备的起始方向720的设备取向的三个角度自由度(例如,偏航、俯仰和翻滚)。如图7所示,设备700也可以水平移动(例如,X方向或Z方向)或竖直移动(例如,Y方向)。设备700还可以前后倾斜(例如,俯仰),左右转动(例如,偏航),以及从一侧向另一侧倾斜(例如,翻滚)。在其它实施方式中,可以使用用于测量头部姿态的其它技术或角度表示,例如,任何其它类型的欧拉角系统。

[0067] 图7示出了设备700,其可以被实现为例如可穿戴显示系统100、AR设备、成像设备或本文描述的任何其它设备。如贯穿本公开所描述的,设备700可以用于确定姿态。例如,在设备700是包括图6中的AR架构600的AR设备的情况下,姿态估计系统640可以使用图像片段输入来提取稀疏点以用于姿态估计处理,如上所述,以跟踪设备在X、Y或Z方向上的移动或跟踪偏航、俯仰或翻滚的角度移动。

[0068] 用于估计3D空间中的姿态的示例性例程

[0069] 图8是用于确定成像设备(例如,图2中的面向外的成像系统110)在3D空间(例如,图3)内的姿态的说明性例程的过程流程图,成像设备在该3D空间内移动。例程800描述了可以如何从表示FOV(例如,FOV315a、315b、315c或315d)的图像帧中提取多个稀疏点以确定成像设备在3D空间中的方位、位置或取向中的一者。

[0070] 在框810处,成像设备可捕获有关AR设备的周围环境的输入图像。例如,成像设备可以基于从周围环境接收的光顺序地捕获输入图像的多个图像片段。这可以通过各种输入设备(例如,AR设备上的数字相机或远离AR设备的数字相机)来实现。输入可以是表示FOV(例如,FOV 315a、315b、315c或315d)的图像并且包括多个稀疏点(例如,稀疏点320)。当成像设备捕获图像片段时,FOV相机、传感器、GPS等可以将包括顺序地捕获的图像片段的图像数据的信息传送到系统(框810)。

[0071] 在框820处,AR设备可以接收输入图像。在一些实施例中,AR设备可以顺序地接收形成在框810处捕获的图像的一部分的多个图像片段。例如,如上所述,面向外的成像系统110可以是滚动快门相机,其被配置为顺序地扫描场景,从而在数据被捕获时顺序地捕获多

个图像片段并顺序地将图像数据读出到存储单元。该信息可以被存储在用户可穿戴系统(例如,本地处理和数据模块70)上,或者可以部分地驻留在可通过有线或无线网络访问的网络存储位置(例如,远程数据储存库74)处。在一些实施例中,可以将信息临时存储在被包括在存储单元内的缓冲器中。

[0072] 在框830处,AR设备可以基于所接收的图像片段辨识一个或多个稀疏点。例如,对象识别器可以爬过与所接收的图像片段对应的图像数据,并辨识一个或多个对象(例如,对象325)。在一些实施例中,辨识一个或多个稀疏点可以基于接收与一个或多个稀疏点对应的图像片段,如下面参考图9A和9B所述。然后,对象识别器可以提取稀疏点,这些稀疏点可以被用于确定姿态数据(例如,3D空间内的成像设备姿态)的输入。然后可以将该信息传送到姿态估计处理(框840),并且AR设备可以相应地利用姿态估计系统来对通过3D空间的AR设备进行地图构建(框850)。

[0073] 在各种实施例中,例程800可以由硬件处理器(例如,图2中的本地处理和数据模块70)执行,该硬件处理器被配置为执行存储在存储器或存储单元中的指令。在其它实施例中,具有计算机可执行指令的远程计算设备(与显示装置进行网络通信)可以使显示装置执行例程800的各方面。

[0074] 如上所述,当前姿态估计处理可以包括由于将数据(例如,所提取的稀疏点)从图像捕获设备传送到姿态估计系统而导致的估计AR设备姿态时的延迟。例如,当前实施方式可能需要将整个图像帧从图像捕获设备传送到姿态估计器(例如,SLAM、VSLAM或类似物)。一旦传送了整个图像帧,就允许对象识别器辨识稀疏点并将其提取到姿态估计器。传送整个图像帧可以是估计姿态的延迟的一个促成因素。

[0075] 从图像帧中提取稀疏点的示例

[0076] 图9A和9B示意性地示出了基于接收多个图像片段,从图像帧中提取一个或多个稀疏点的示例。在一些实施方式中,图9A和9B还可以示意性地示出最小化在估计通过3D空间的成像设备(例如,图2中的面向外的成像设备110)的姿态时的延迟的示例性方法。在一些实施例中,图9A和9B还示意性地示出了辨识图像帧900的一个或多个稀疏点的示例。在一些实施方式中,图9A和9B示出了如上所述,在滚动快门相机将图像帧从成像设备读入存储单元时的图像帧。图像帧900可以由被配置为逐行扫描成像设备的面向外的成像系统110捕获。图像帧可包括多个图像片段(有时被称为扫描线)905a至905n,当成像设备捕获图像片段时,所述多个图像片段从成像设备被读入存储单元(例如,本地处理和数据模块70)中。图像片段可以水平排列(如图9A所示)或竖直排列(未示出)。虽然示出了15个图像片段,但是图像片段的数量不限于此,可以根据给定应用的需要或基于成像系统的能力而为任何数量的图像片段905a至905n。在一些实施方式中,图像片段可以是光栅扫描模式中的线(例如,行或列),例如,图像片段可以是由面向外的成像设备110捕获的图像的光栅扫描模式中的像素的行或列。光栅扫描模式可以由滚动快门相机执行或实施,如贯穿本公开所述。

[0077] 再次参考图9A,图像帧900可以包括被顺序地捕获并被读入存储单元中的多个图像片段905。图像片段905可以被组合以表示由成像设备捕获的视场(FOV)。图像帧900还可以包括多个稀疏点320,例如,如上面参考图3所述。在一些实施方式中,如图9A所示,每个稀疏点320可以由一个或多个图像片段905生成。例如,稀疏点320a可以由图像片段905的子集910生成,并因而与其相关联。因此,当在存储单元处接收到图像片段时,可以在接收到与每

个给定稀疏点对应的图像片段905的子集时辨识每个稀疏点。例如，一旦在AR设备的存储单元处接收到图像片段906a至906n，就可以由对象识别器（例如，对象识别器650）辨识稀疏点320a。图像片段906a至906n可以对应于表示稀疏点320a的图像片段905的子集910。因此，一旦从图像捕获设备（例如，逐行扫描相机）接收到对应的图像片段，AR设备就能够确定各个稀疏点。图像片段905的子集910可包括图像片段906a至906n。在一些实施方式中，图像片段906的数量可基于沿垂直方向解析（resolve）或捕获整个稀疏点所需的、在垂直方向上顺序地接收的图像片段的数量。尽管图9B示出了与稀疏点320a相关联的7个图像片段，但情况不一定如此，任何数量的图像片段（例如，2、3、4、5、6、8、9、10、11等个图像片段）可以与稀疏点320a相关联以辨识与稀疏点320a对应的对象325a。

[0078] 在示例性实施方式中，可以通过实现循环缓冲器或滚动缓冲器来辨识稀疏点320。例如，该缓冲器可以类似于图6中的缓冲器620。该缓冲器可以被构造为存储在AR设备上的存储器或存储单元的一部分（例如，本地处理和数据模块70），或者可以远离AR设备（例如，远程数据储存库74）。该缓冲器可以被配置为从图像捕获设备（例如，图2中的面向外的成像系统110）接收图像信息。例如，当图像传感器捕获每个顺序图像片段时，缓冲器可以顺序地从图像传感器接收表示图像片段的图像数据。该缓冲器还可以被配置为存储图像数据的一部分，以进行图像内容的后续处理和辨识。在一些实施例中，该缓冲器可以被配置为存储一个或多个图像片段，其中图像片段的数量可以小于总图像帧900。在一些实施例中，存储在该缓冲器中的图像片段的数量可以是预定数量，例如子集910中的数量。在一些实施例中，替代地或组合地，该缓冲器可以被配置为存储与稀疏点对应的图像片段的子集910。例如，参考图9B，稀疏点320a可能需要 7×7 像素窗口（例如，呈现图像片段906的7行像素，其中每个图像片段包括7个像素）。在该实施例中，该缓冲器可以被配置为足够大以存储图像片段906的子集910，例如所示的7个图像片段。

[0079] 如上所述，该缓冲器可以被配置为临时存储图像数据。因此，当从成像捕获设备接收新图像片段时，从该缓冲器中移除较旧的图像片段。例如，可以接收第一图像片段906a，并且在该缓冲器处接收与稀疏点320a对应的后续图像片段。一旦接收到所有图像片段906a至906n，便可辨识稀疏点320a。随后，接收新的图像片段（例如，906n+1），从而从该缓冲器中移除图像片段906a。在一些实施例中，片段906a从该缓冲器移到数字存储器（例如，本地处理和数据模块70）中的存储装置以执行进一步处理。

[0080] 用于估计3D空间中姿态的示例性例程

[0081] 图10是用于确定成像设备（例如，图2中的面向外的成像系统110）在3D空间（例如，图3）内的姿态的说明性例程的过程流程图，成像设备在该3D空间内移动。例程1000描述了在接收与第一组稀疏点中的稀疏点对应的图像片段时，如何从图像帧中提取第一组稀疏点的示例。在各种实施例中，可在捕获表示成像设备的FOV的整个图像帧之前捕获对应的图像片段。例程1000还描述了后续稀疏点或第二组稀疏点可以如何被提取和整合以更新姿态确定。如上所述，例程1000可以由可操作地耦接到面向外的成像系统（例如，面向外的成像系统110）和数字存储器或缓冲器的硬件处理器（例如，图2中的本地处理和数据模块70）来执行。面向外的成像系统110可包括滚动快门相机。

[0082] 在框1010处，成像设备可捕获有关AR设备的周围环境的输入图像。例如，成像设备可以基于从周围环境接收的光而顺序地捕获输入图像的多个图像片段。这可以通过各种输

入设备(例如,AR设备上的数字相机或远离AR设备的数字相机)来实现。输入可以是表示FOV(例如,FOV 315a、315b、315c或315d)的图像帧并且包括多个稀疏点(例如,稀疏点320)。当成像设备捕获图像片段时,FOV相机、传感器、GPS等可以将包括顺序地捕获的图像片段的图像数据的信息传送到系统(框1010)。

[0083] 在框1020处,AR设备可以接收输入图像。在一些实施例中,AR设备可以顺序地接收形成在框1010处捕获的图像的一部分的第一多个图像片段。例如,成像设备可以被配置为顺序地扫描场景,从而顺序地捕获第一多个图像片段,如上面参考图9A和9B所述。当数据被捕获时,图像传感器还可以顺序地将图像数据读出到存储单元。该信息可以被存储在用户可穿戴系统(例如,本地处理和数据模块70)上,或者可以部分地驻留在可通过有线或无线网络访问的网络存储位置(例如,远程数据储存库74)处。在一些实施例中,可以将信息临时存储在被包括在存储单元内的缓冲器中。

[0084] 在框1030处,AR设备可以基于接收与每个稀疏点对应的第一多个图像片段(有时被称为“预列表(pre-list)”)来辨识第一组稀疏点。例如,参考图9A和9B,AR设备可以基于接收与一个或多个稀疏点320对应的图像片段905的子集910(例如,第一多个图像片段)来辨识一个或多个稀疏点320,如上面参考图9A和9B所述。一旦在存储单元(例如,本地处理和数据模块70)处接收到与稀疏点320对应的图像片段905的子集910(例如,图像片段906),便可辨识稀疏点320。

[0085] 在一些实施方式中,第一组稀疏点包括任意数量的稀疏点(N_1)。数量(N_1)可以是被选择用于估计AR设备在环境内的姿态的任何数量的稀疏点。在一些实施例中,数量(N_1)可以不小于三个稀疏点。在其它实施例中,数量(N_1)在10至20个稀疏点之间。更大数量(N_1)的一个非限制性优点是可以拒绝异常值(outlier)数据点,这样,由于正常值(inlier)数据点,可以提供对噪声具有一定鲁棒性的姿态确定。例如,由于被赋予物理成像设备的事件,成像设备可能抖动(jilt)或摇晃,或者被记录的场景可能被临时改变(例如,在前景中移动的人)。该事件可能仅影响一个或多个图像帧中的一小组稀疏点。使用更大数量(N_1)的稀疏点或根据本说明书更新姿态估计,可以至少部分地降低由于这些异常值或单个实例事件导致的姿态估计中的噪声。

[0086] 在一种实施方式中,可以从图像帧中提取第一组稀疏点(例如,通过对象识别器650)并将其传送到被配置为执行姿态确定(例如,如上所述的SLAM、VSLAM或类似物)的姿态估计系统(例如,图6中的姿态估计系统640)(框1040)。在各种实施例中,在基于接收对应的第一多个图像片段而辨识稀疏点的数量(N_1)后,将第一组稀疏点传送到姿态估计系统。因此,可以在仅接收到图像帧的一部分时传送第一组稀疏点,这是因为成像设备尚未接收到整个图像帧;后续图像片段(例如,在第一多个图像片段之后获得的第二多个图像片段)仍有待接收。在一个实施例中,一旦基于扫描对应的图像片段子集而辨识每个稀疏点,便可提取第一组稀疏点(例如,从AR设备的存储单元或其一部分,例如缓冲器)。在另一实施例中,一旦辨识出数量(N_1)的稀疏点,便可提取第一组稀疏点(例如,从AR设备的存储单元或缓冲器),并且通过单个过程发送稀疏点。

[0087] 在框1045处,AR设备可以接收第二多个图像片段(有时被称为“跟随列表(follow-list)”)。在一些实施例中,在框1020处接收第一多个图像片段之后,AR设备还可以顺序地获得第二多个图像片段。例如,成像设备可被配置为顺序地扫描场景,从而顺序地捕获第一

多个图像片段(例如,框1020),随后在框1030之后或框1030期间顺序地扫描场景以获得第二多个图像片段,如上面参考图9A和9B所述。在另一实施例中,第二多个图像片段或其一部分可以从成像设备捕获的第二图像中获得,第二图像是在第一图像之后捕获的。信息可以被存储在AR设备(例如,本地处理和数据模块70)上,或者可以部分地驻留在可通过有线或无线网络访问的网络存储位置(例如,远程数据储存库74)处。在一些实施例中,可以将信息临时存储在被包括在存储单元内的缓冲器中。

[0088] 再次参考图10,在框1050处,AR设备可以基于第二多个图像片段辨识第二组稀疏点。例如,在一个实施例中,在框1040处确定姿态之前尚未接收到整个图像帧,并且在框1045处,可以从成像设备接收第二多个图像片段。因此,AR设备可以基于接收与一个或多个新稀疏点(例如,第二组稀疏点)对应的第二多个图像片段而辨识一个或多个新稀疏点,如上面参考图9A和9B所述。在另一实施例中,在框1010处捕获第一图像之后,成像设备可以捕获第二图像,并且可以从第二图像中获得第二多个图像片段。因此,AR设备可以基于接收来自第二图像的第二多个图像片段辨识一个或多个新稀疏点,该第二多个图像片段可以与第二组稀疏点对应。在一些实施例中,第二组稀疏点可以包括任何数量的新稀疏点(例如,1、2、3个等等)。在一个实施方式中,第二组稀疏点可以被提取并整合到姿态确定中,例如,通过将第二组稀疏点传送到姿态估计系统。下面是将第二组稀疏点与第一组稀疏点整合到图10的地图构建例程中的示例性方法。例如,在此描述的示例性整合方法可以被重新整合、滑动缩放整合或块整合。但是,这些示例性整合方法并非旨在是穷举的。可以使误差最小化并减少姿态确定时的延迟的其它方法也是可能的。

[0089] 在框1060处,姿态估计系统可以被配置为基于框1040处的姿态确定以及框1050处的第二组稀疏点的接收而更新姿态确定。

[0090] 上述例程1000的一个非限制性优点可以是在姿态估计过程之前减少从图像帧中提取稀疏点所导致的延迟。例如,通过在缓冲器620处接收与各个稀疏点对应的图像片段时计算和辨识这些稀疏点,可以将单独的稀疏点或选定的一组稀疏点提取到姿态估计系统并且由姿态估计系统进行处理,无需等待捕获整个图像帧。因此,可以在将整个图像传送到存储器之前以及在能够从整个图像中提取所有稀疏点之前很好地执行姿态估计。然而,一旦提取了特定图像帧的第一组和所有后续组的稀疏点,整个图像帧便可用于姿态估计。

[0091] 在各种实施方式中,第二组稀疏点可以包括在框1040处确定姿态之后辨识的设定数量的稀疏点。在一些实施例中,该设定数量可以是一个稀疏点。例如,每当后续稀疏点被辨识时,可以将该稀疏点传送到姿态估计系统,并且在框1060处执行新的姿态估计处理以更新AR设备的方位、位置或取向中的一者或多者。该方法有时可被称为重新整合方法。因此,每个随后辨识的稀疏点可以表示后续稀疏点组(例如,第二、第三、第四等组稀疏点)。在另一实施例中,该设定数量可以是任何数量的后续辨识的稀疏点(例如,2、3、4个等等)。例如,在设定数量为3的情况下,每次辨识3个新的稀疏点(例如,后续稀疏点组),在框1050处将该组传送到姿态估计系统,并且在框1060处执行新的姿态估计处理。因此,姿态估计处理可以利用整个图像帧中所包括的所有稀疏点。

[0092] 在其它实施方式中,整合方法可以被配置为解释上面参考图4A至5B描述的滚动快门效果。例如,可以针对固定数量(N_2)的稀疏点执行姿态估计处理。该方法有时可被称为滑动缩放整合方法。在该实施例中,第二组稀疏点可以包括在框1040处确定姿态之后辨识的

选定数量(k_2)的稀疏点。每当一数量(k_2)的稀疏点可以被辨识时,可以更新姿态确定。然而,只有最近期的 N_2 个稀疏点可用于在框1060处更新姿态。在一些实施例中,该方法利用最近期的 N_2 个稀疏点,而不管它们对应于哪个组。例如,如果 N_1 被设定为10, N_2 被设定为15,并且 k_2 被设定为5,则第一组稀疏点包括在框1030处辨识的第一10个稀疏点。因此,基于第一10个稀疏点在框1040处确定姿态。随后,新稀疏点被辨识出,但不更新姿态。一旦辨识出包括第二组稀疏点的5个新稀疏点,便可基于第一(N_1)和第二(k_2)组稀疏点更新姿态。如果辨识出第三组稀疏点(例如,在第二组之后的5个稀疏点),则在框1060处再次更新姿态,然而,该更新可以基于第一组中的部分稀疏点(例如,稀疏点6至10)、第二组稀疏点(例如,稀疏点11至15)和第三组稀疏点(例如,稀疏点16至21)。因此,可以将该整合视为稀疏点的滑动窗口或滑动列表,由此仅使用设定数量的稀疏点来估计姿态,并且所使用的稀疏点从第一组至第二组和第三组滑动。该方法的一个非限制性优点可以是,当从较早接收的图像片段中辨识的稀疏点变得老旧或陈旧时,其可以从框1060处的姿态确定中被删除。在一些情况下,如果AR设备相对于稀疏点移动,则可以通过移除旧稀疏点并捕获所辨识的新稀疏点之间的姿态变化来缩减滚动快门效果。

[0093] 在一些实施例中,可以在图像帧之间,例如,在面向外的成像系统110在捕获图3的FOV 315a的图像帧和捕获图3的FOV 315b的图像帧之间移动时,利用前述整合方法。例如,可以从与第一方位312a(例如,FOV315b)相关联的图像帧接收第一组稀疏点,并且可以从与第二方位312b(例如,FOV 315b)相关联的图像帧接收第二组稀疏点。可以实现滑动列表方法以缩减这些图像帧之间的滚动快门效果。然而,在一些实施例中,不必保留比来自第一帧的最近期(N_2-1)个稀疏点更多的稀疏点。

[0094] 在另一实施方式中,可以针对固定数量的稀疏点或稀疏点块执行在框1060处的姿态确定。该方法有时可被称为块整合方法。在一些实施例中,多组稀疏点中的每一组可以包括等于该块的一数量的稀疏点。例如,如果块被设定为10,则第一组的固定数量(N_1)是10,并且在辨识和提取该第一组后在框1040处确定姿态。随后,可以辨识包括接下来的10个稀疏点的第二组稀疏点,并且使用该第二组稀疏点在框1060处更新姿态。在一些实施例中,该过程可以针对多个组(例如,第三、第四、第五组等等)继续。在一些实施例中,当图像片段被存储在缓冲器(例如,图6中的缓冲器620)中时,缓冲器的大小可以被选择并配置为至少存储可以包括在块中的数量的稀疏点(例如,在上述示例中,缓冲器可以被选择为具有被配置为存储至少10个稀疏点的大小)。在一些实施例中,缓冲器可以具有被限制为仅存储块中所包括的稀疏点数量的大小。

[0095] 虽然贯穿本公开参考头戴式显示设备或AR设备描述了方法、设备和系统的各种实施例,但是这并不旨在限制本申请的范围,并且仅用作用于说明目的的示例。在此描述的方法和设备可以应用于可以实现在此描述的方法和设备以对设备所在的3D空间进行地图构建、并且跟踪设备通过3D环境的移动的其它设备,例如机器人、数字相机和其它自主实体。

[0096] 其它方面

[0097] 在第1方面,公开了一种用于估计图像捕获设备在环境内的方位的方法。所述方法包括:顺序地接收第一多个图像片段,所述第一多个图像片段形成表示所述图像捕获设备的视场(FOV)的图像的至少一部分,所述FOV包括围绕所述图像捕获设备的环境的一部分并包括多个稀疏点,其中每个稀疏点对应于图像片段的子集;辨识第一组稀疏点,所述第一组

稀疏点包括在接收所述第一多个图像片段时辨识的一个或多个稀疏点;通过位置估计系统基于所述第一组稀疏点确定所述图像捕获设备在所述环境内的方位;顺序地接收第二多个图像片段,所述第二多个图像片段是在所述第一多个图像片段之后接收的并形成所述图像的至少另一部分;辨识第二组稀疏点,所述第二组稀疏点包括在接收所述第二多个图像片段时辨识的一个或多个稀疏点;以及通过所述位置估计系统基于所述第一组稀疏点和所述第二组稀疏点更新所述图像捕获设备在所述环境内的方位。

[0098] 在第2方面,根据第1方面所述的方法,进一步包括在所述图像捕获设备的图像传感器处顺序地捕获所述多个图像片段。

[0099] 在第3方面,根据第1方面或第2方面所述的方法,其中所述图像传感器是滚动快门图像传感器。

[0100] 在第4方面,根据第1方面至第3方面中任一项所述的方法,进一步包括:在顺序地接收所述图像片段时,将所述第一多个图像片段和所述第二多个图像片段存储在缓冲器中,所述缓冲器具有与所述图像片段的子集中的图像片段的数量对应的大小。

[0101] 在第5方面,根据第1方面至第4方面中任一项所述的方法,进一步包括将所述第一组稀疏点和所述第二组稀疏点提取到所述位置估计系统。

[0102] 在第6方面,根据第1方面至第5方面中任一项所述的方法,其中所述第一组稀疏点包括一数量的稀疏点。

[0103] 在第7方面,根据第6方面所述的方法,其中所述稀疏点的数量在10个与20个稀疏点之间。

[0104] 在第8方面,根据第1方面至第7方面中任一项所述的方法,其中所述第二组稀疏点包括第二数量的稀疏点。

[0105] 在第9方面,根据第1方面至第8方面中任一项所述的方法,其中所述更新所述图像捕获设备的方位是基于最近辨识的稀疏点的数量,其中所述最近辨识的稀疏点是所述第一组、所述第二组中的至少一者,或所述第一组和所述第二组中的一者或多者。

[0106] 在第10方面,根据第9方面所述的方法,其中所述最近辨识的稀疏点的数量等于所述第一组稀疏点中的稀疏点的数量。

[0107] 在第11方面,根据第1方面至第10方面中任一项所述的方法,其中所述位置估计系统被配置为执行视觉同步定位与地图构建(V-SLAM)。

[0108] 在第12方面,根据第1方面至第11方面中任一项所述的方法,其中所述多个稀疏点基于真实世界对象、虚拟图像元素和被投射到所述环境中的不可见指示符中的至少一者而被提取。

[0109] 在第13方面,公开了一种用于估计图像捕获设备在环境内的方位的方法。该方法包括:顺序地接收多个图像片段,所述多个图像片段形成表示所述图像捕获设备的视场(FOV)的图像,所述FOV包括所述图像捕获设备周围的环境的一部分并包括多个稀疏点,其中每个稀疏点部分地基于所述多个图像片段的对应图像片子集而可辨识;在接收到与所述一个或多个稀疏点对应的每个图像片子集时,顺序地辨识所述多个稀疏点中的一个或多个稀疏点;基于所辨识的一个或多个稀疏点估计所述图像捕获设备在所述环境内的方位。

[0110] 在第14方面,根据第13方面所述的方法,其中顺序地接收所述多个图像片段进一

步包括接收一数量的图像片段并将该数量的图像片段存储在缓冲器中。

[0111] 在第15方面,根据第13方面或第14方面所述的方法,其中顺序地接收所述多个图像片段包括接收至少第一图像片段和第二图像片段,其中所述第一图像片段被存储在所述缓冲器中。

[0112] 在第16方面,根据第13方面至第15方面中任一项所述的方法,进一步包括:在接收到第二图像片段时更新所述缓冲器;将所述第二图像片段存储在所述缓冲器中;以及,在接收到所述第二图像片段时,移除所述第一图像片段。

[0113] 在第17方面,根据第16方面所述的方法,其中顺序地辨识一个或多个稀疏点进一步包括在更新所述缓冲器时扫描被存储在所述缓冲器中的所述图像片段。

[0114] 在第18方面,根据第13方面至第17方面中任一项所述的方法,其中在接收到与所述一个或多个稀疏点对应的每个图像片子集时顺序地辨识所述多个稀疏点中的一个或多个稀疏点进一步包括:在接收到与第一组的一个或多个稀疏点对应的第一多个图像片段时顺序地辨识所述第一组的所述一个或多个稀疏点;以及在接收到与第二组的一个或多个稀疏点对应的第二多个图像片段时顺序地辨识所述第二组的所述一个或多个稀疏点,其中所述第二多个图像片段是在所述第一多个图像片段之后被接收的。

[0115] 在第19方面,根据第13方面至第18方面中任一项所述的方法,其中估计所述图像捕获设备的方位是基于辨识所述第一组的一个或多个稀疏点,其中所述第一组包括一数量的稀疏点。

[0116] 在第20方面,根据第19方面所述的方法,其中所述稀疏点的数量在2和20之间。

[0117] 在第21方面,根据第19方面所述的方法,其中所述稀疏点的数量在10和20之间。

[0118] 在第22方面,根据第13方面至第21方面中任一项所述的方法,进一步包括基于辨识第二组的一个或多个稀疏点而更新所述图像捕获设备的方位。

[0119] 在第23方面,根据第13方面至第22方面中任一项所述的方法,其中所述第二组的一个或多个稀疏点包括第二数量的稀疏点。

[0120] 在第24方面,根据第13方面至第23方面中任一项所述的方法,进一步包括基于辨识一数量的顺序地辨识的稀疏点而更新所述图像捕获设备的方位。

[0121] 在第25方面,根据第24方面所述的方法,其中所述顺序地辨识的稀疏点的数量等于所述稀疏点的数量。

[0122] 在第26方面,根据第24方面所述的方法,其中所述顺序地辨识的稀疏点的数量包括所述第一组稀疏点中的至少一个稀疏点。

[0123] 在第27方面,根据第13方面至第26方面中任一项所述的方法,其中所述多个稀疏点基于真实世界对象、虚拟图像元素和被投射到所述环境中的不可见指示符中的至少一者而被提取。

[0124] 在第28方面,根据第13方面至第27方面中任一项所述的方法,进一步包括:从所述缓冲器中提取所述顺序地辨识的稀疏点;将所述顺序地辨识的稀疏点发送到视觉同步定位与地图构建(VSLAM)系统,其中所述VSLAM系统基于所述顺序地辨识的一个或多个稀疏点而估计所述图像捕获设备的方位。

[0125] 在第29方面,公开了一种增强现实(AR)系统。所述AR系统包括面向外的成像设备、计算机硬件、以及可操作地耦接到所述计算机硬件和所述面向外的成像设备并且被配置为

执行指令以执行根据第1方面至第28方面中任一项所述的方法的处理器。

[0126] 在第30方面,根据第29方面所述的AR系统,其中所述面向外的成像设备被配置为检测不可见光谱中的光。

[0127] 在第31方面,根据第29方面或第30方面所述的AR系统,其中所述AR系统被配置为显示一个或多个虚拟图像元素。

[0128] 在第32方面,根据第29方面至第31方面中任一项所述的AR系统,进一步包括:收发器,其被配置为将指示所述AR系统的估计的方位的辨识信号发送到远程AR系统,其中所述远程AR系统被配置为基于所接收的辨识信号而更新其估计的方位。

[0129] 在第33方面,公开了一种自主实体。所述自主实体包括面向外的成像设备、计算机硬件、以及可操作地耦接到所述计算机硬件和所述面向外的成像设备并且被配置为执行指令以执行根据第1方面至第28方面中任一项所述的方法的处理器。

[0130] 在第34方面,根据第33方面所述的自主实体,其中所述面向外的成像设备被配置为检测不可见光谱中的光。

[0131] 在第35方面,公开了一种机器人系统。所述机器人系统包括面向外的成像设备、计算机硬件、以及可操作地耦接到所述计算机硬件和所述面向外的成像设备并且被配置为执行指令以执行根据第1方面至第28方面中任一项所述的方法的处理器。

[0132] 在第36方面,公开了一种用于估计图像捕获设备在环境内的方位的图像捕获设备。所述图像捕获设备包括:图像传感器,其被配置为通过顺序地捕获多个图像片段来捕获图像,所述图像表示所述图像捕获设备的视场(FOV),所述FOV包括所述图像捕获设备周围的环境的一部分,其中包括多个稀疏点,其中每个稀疏点部分地基于所述多个图像片段的对应子集而可辨识;存储器电路,其被配置为存储与一个或多个稀疏点对应的图像片段子集;计算机处理器,其可操作地耦接到所述存储器电路并且被配置为:在接收到与所述一个或多个稀疏点对应的每个图像片段子集时,顺序地辨识所述多个稀疏点中的一个或多个稀疏点;以及提取顺序地辨识的一个或多个稀疏点,以基于所辨识的所述一个或多个稀疏点估计所述图像捕获设备在所述环境内的方位。

[0133] 在第37方面,根据第36方面所述的图像捕获设备,进一步包括方位估计系统,其被配置为:接收顺序地辨识的一个或多个稀疏点;以及基于所辨识的一个或多个稀疏点估计所述图像捕获设备在所述环境内的方位。

[0134] 在第38方面,根据第36方面或第37方面所述的图像捕获设备,其中所述方位估计系统是视觉同步定位与地图构建(VSLAM)系统。

[0135] 在第39方面,根据第36方面至第38方面中任一项所述的图像捕获设备,其中所述图像传感器被配置为检测不可见光谱中的光。

[0136] 在第40方面,根据第36方面至第39方面中任一项所述的图像捕获设备,进一步包括:收发器,其被配置为将指示其估计的方位的辨识信号发送到远程图像捕获设备,其中所述远程图像捕获设备被配置为基于所接收的辨识信号更新其估计的方位。

[0137] 其它考虑事项

[0138] 在此描述的和/或在附图中描绘的过程、方法和算法中的每一个可以体现在由一个或多个物理计算系统、硬件计算机处理器、专用电路和/或被配置为执行具体和特定计算机指令的电子硬件所执行的代码模块中,并且由以上完全或部分自动化。例如,计算系统可

以包括编程有特定计算机指令的通用计算机(例如,服务器)或专用计算机、专用电路等等。代码模块可以被编译并链接到可执行程序中,安装在动态链接库中,或者可以用解释的编程语言写入。在一些实施方式中,特定操作和方法可以由给定功能特定的电路来执行。

[0139] 此外,本公开的功能的某些实施方式在数学上、计算上或技术上是足够复杂的,以致于可能需要专用硬件或一个或多个物理计算设备(利用适当的专用可执行指令)或专用图形处理单元来执行功能,例如由于所涉及的计算的量或复杂性或为了基本实时地提供结果,例如,姿态估计输入。例如,视频可以包括许多帧,每帧具有数百万个像素,并且特别编程计算机硬件需要处理视频数据,从而在商业上合理的时间量内提供期望的图像处理任务或应用。

[0140] 代码模块或任何类型的数据可以存储在任何类型的非瞬态计算机可读介质上,诸如物理计算机存储器,其包括硬盘驱动器、固态存储器、随机存取存储器(RAM)、只读存储器(ROM)、光盘、易失性或非易失性存储器、相同的组合和/或类似物。方法和模块(或数据)也可以在各种计算机可读传输介质上作为生成的数据信号(例如,作为载波或其它模拟或数字传播信号的一部分)传输,所述传输介质包括基于无线的和基于有线/线缆的介质,并且可以采取多种形式(例如,作为单个或多路复用模拟信号的一部分,或者作为多个离散数字分组或帧)。所公开的过程或过程步骤的结果可以持久地或以其它方式存储在任何类型的非瞬态、有形计算机存储器中,或者可以经由计算机可读传输介质来传送。

[0141] 在此所描述和/或附图中描绘的流程图中的任何过程、框、状态、步骤或功能应当被理解为潜在地表示代码模块、代码段或代码部分,这些代码模块、代码段或代码部分包括用于实现特定功能(例如,逻辑或算术)或步骤的一个或多个可执行指令。过程、框、状态、步骤或功能可以与在此提供的说明性示例相组合、重新排列、添加、删除、修改或以其它方式改变。在一些实施例中,附加的或不同的计算系统或代码模块可以执行在此描述的功能中的一些或全部。在此描述的方法和过程也不限于任何特定的顺序,并且与其相关的块、步骤或状态可以以适当的其它顺序来执行,例如串行、并行或以某种其它方式。任务或事件可以添加到所公开的示例实施例或者从中去除。此外,在此描述的实施方式中的各种系统部件的分离是出于说明的目的,并且不应该被理解为在所有实施方式中都需要这种分离。应该理解,所描述的程序组件、方法和系统通常可以一起被集成在单个计算机产品中或者被封装到多个计算机产品中。许多实施方式变化是可能的。

[0142] 过程、方法和系统可以实现在网络(或分布式)计算环境中。网络环境包括企业范围的计算机网络、内联网、局域网(LAN)、广域网(WAN)、个人局域网(PAN)、云计算网络、众包(crowd-sourced)计算网络、互联网和万维网。网络可以是有线或无线网络或任何其它类型的通信网络。

[0143] 本公开的系统和方法各自具有若干创新性方面,其中没有单独一个对在此公开的期望属性完全负责或需要。上述各种特征和过程可以彼此独立地使用,或者可以以各种方式组合。所有可能的组合和子组合均旨在落入本公开的范围。对于本公开中所描述的实施方式的各种修改对于本领域技术人员来说可以是容易清楚的,并且在不脱离本公开的精神或范围的情况下,可将在此定义的一般原理应用于其它实施方式。因此,权利要求不旨在限于在此示出的实施方式,而是应被给予与本公开、在此公开的原理和新颖特征一致的最宽范围。

[0144] 本说明书中在分开的实现方式中描述的某些特征也可以在单个实现方式中组合地实现。相反地,在单个实施方式的上下文中描述的各种特征也可以分别在多个实施方式中分开地或以任何合适的子组合方式来实现。此外,尽管上文可以将特征描述为以某些组合起作用并且甚至最初如此要求,但是来自所要求的组合的一个或多个特征可以在一些情况下从组合中删除,并且所要求的组合可以针对子组合或子组合的变体。没有单个特征或特征组是每个实施例必需或不可缺少。

[0145] 除非另有特别说明,或者在所使用的上下文中进行理解,在此使用的条件语言,诸如“能”、“能够”、“可能”、“可以”、“例如”等等,一般意在表达某些实施例包括而其它实施例不包括某些特征、要素和/或步骤。因此,这样的条件语言通常不旨在暗示特征、要素和/或步骤以任何方式对于一个或多个实施例是必需的,或者一个或多个实施例必然包括用于在有或者没有作者输入或提示的情况下决定这些特征、要素和/或步骤是否包括在或执行在任何特定实施例中。术语“包含”、“包括”、“具有”等是同义词,并且以开放式的方式包含性地使用,并且不排除附加的要素、特征、动作、操作等等。此外,术语“或”以其包含性含义(而不是其排他含义)使用,因此当用于例如连接要素列表时,术语“或”表示列表中的一个、一些或全部要素。另外,除非另有说明,否则本申请和所附权利要求书中使用的冠词“一”、“一个”和“所述”应被解释为表示“一个或多个”或“至少一个”。

[0146] 如在此所使用的,提及项目列表中的“至少一者”的短语是指这些项目的任何组合,包括单个成员。作为例子,“A,B或C中的至少一者”旨在涵盖:A、B、C、A和B、A和C、B和C,以及A、B和C。除非另有特别说明,否则诸如短语“X、Y和Z中的至少一者”的连接语言如所使用的利用上下文进行理解,通常用于表达项目、术语等可以是X、Y或Z中的至少一者。因此,这样的连接语言通常不旨在暗示某些实施例需要X中的至少一个、Y中的至少一个和Z中的至少一个都存在。

[0147] 类似地,虽然可以在附图中以特定顺序示出操作,但是应该认为,不需要以所示出的特定顺序或按顺次顺序执行这样的操作,或者不需要执行所有示出的操作以实现期望的结果。此外,附图可以以流程图的形式示意性地描绘一个或多个示例过程。然而,未示出的其它操作可以并入示意性说明的示例性方法和过程中。例如,一个或多个附加操作可以在任何所示操作之前、之后、同时或之间执行。另外,在其它实施方式中,操作可以重新安排或重新排序。在特定情况下,多任务和并行处理可能是有利的。而且,上述实施方式中的各种系统组件的分离不应当被理解为在所有实施方式中都需要这种分离,并且应该理解的是,所描述的程序组件和系统通常可以一起集成在单个软件产品中或者封装到多个软件产品。另外,其它的实施方式在以下权利要求的范围中。在一些情况下,权利要求中列举的动作能够以不同的顺序执行并且仍然实现期望的结果。

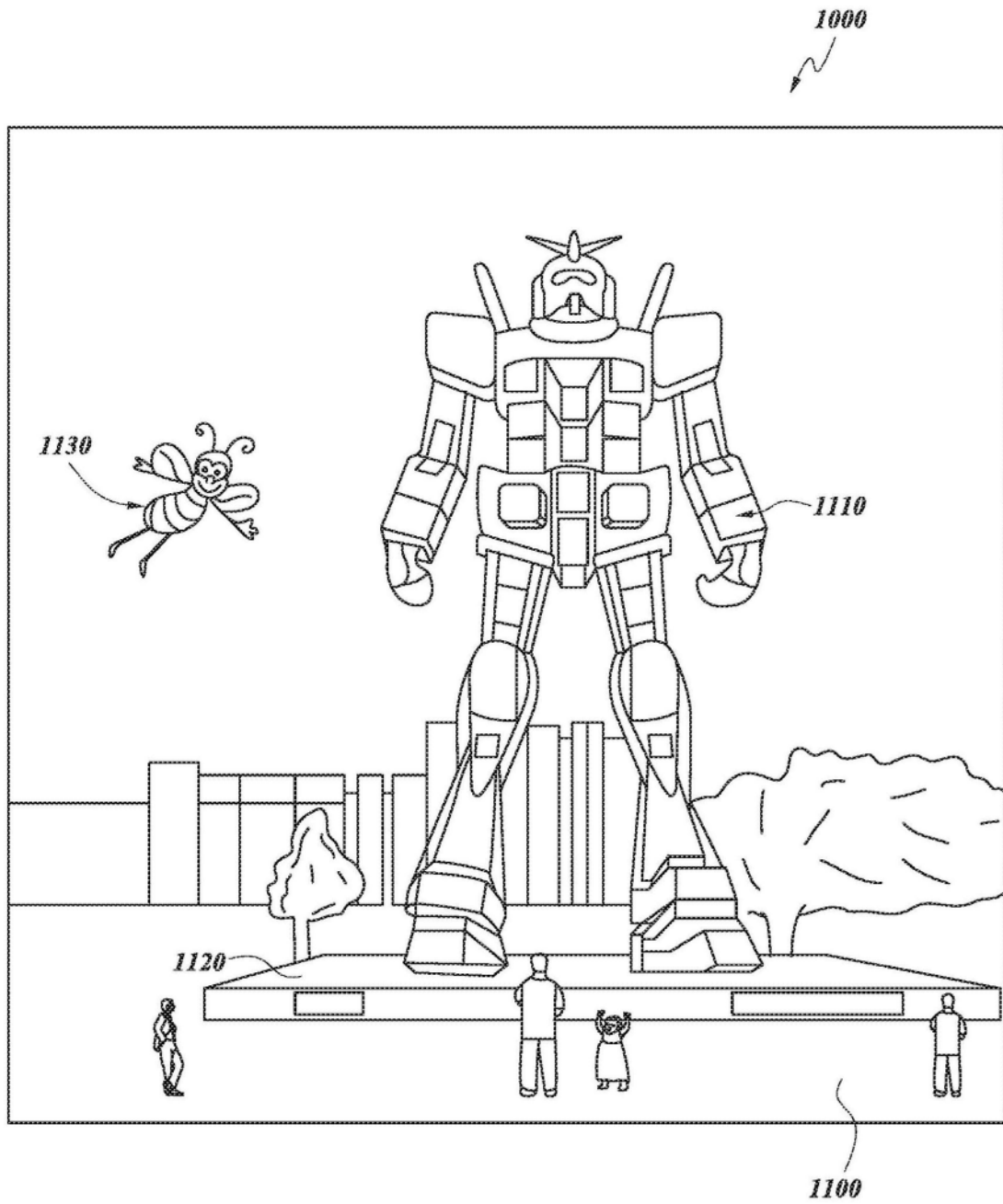


图1

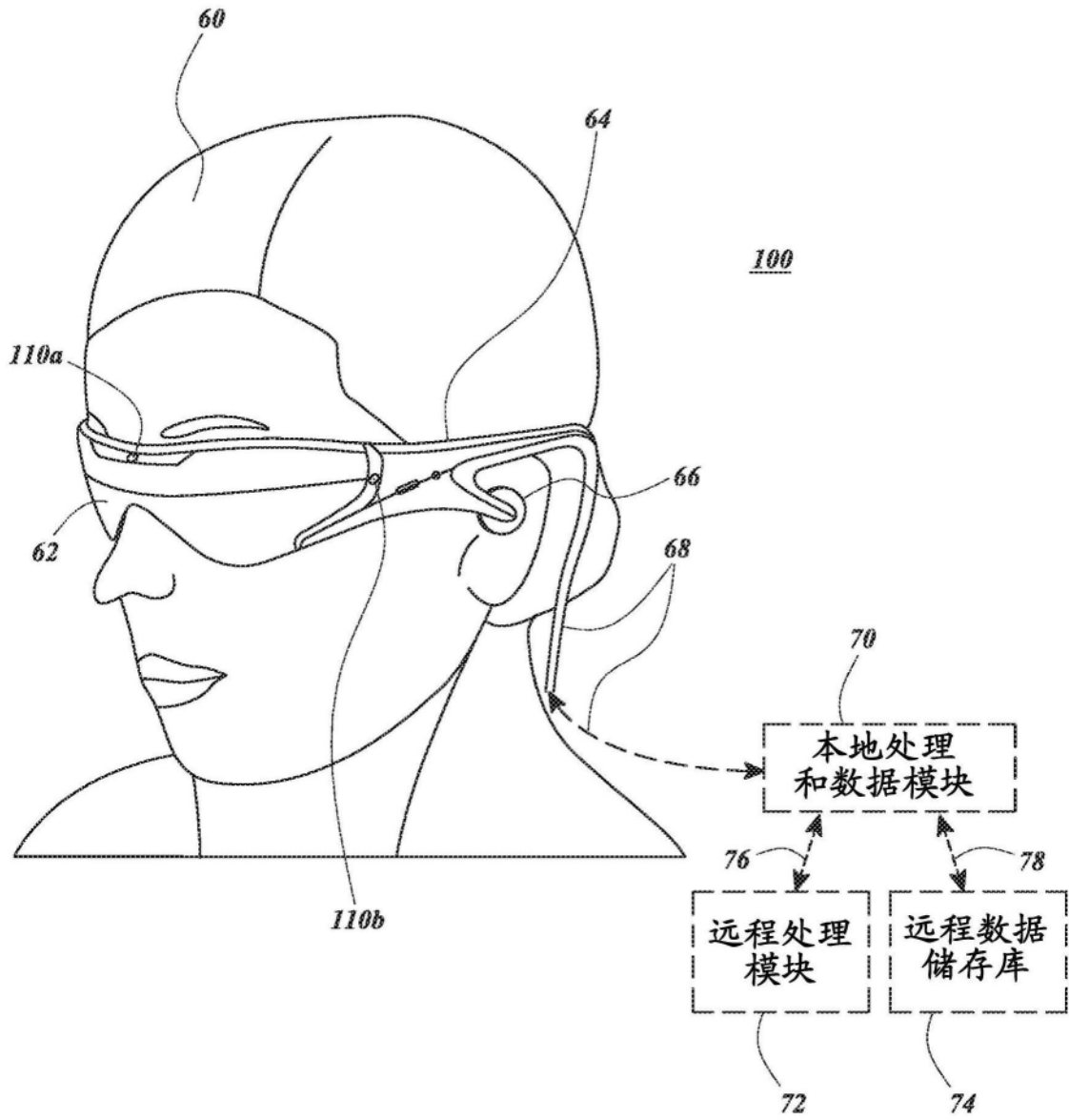


图2

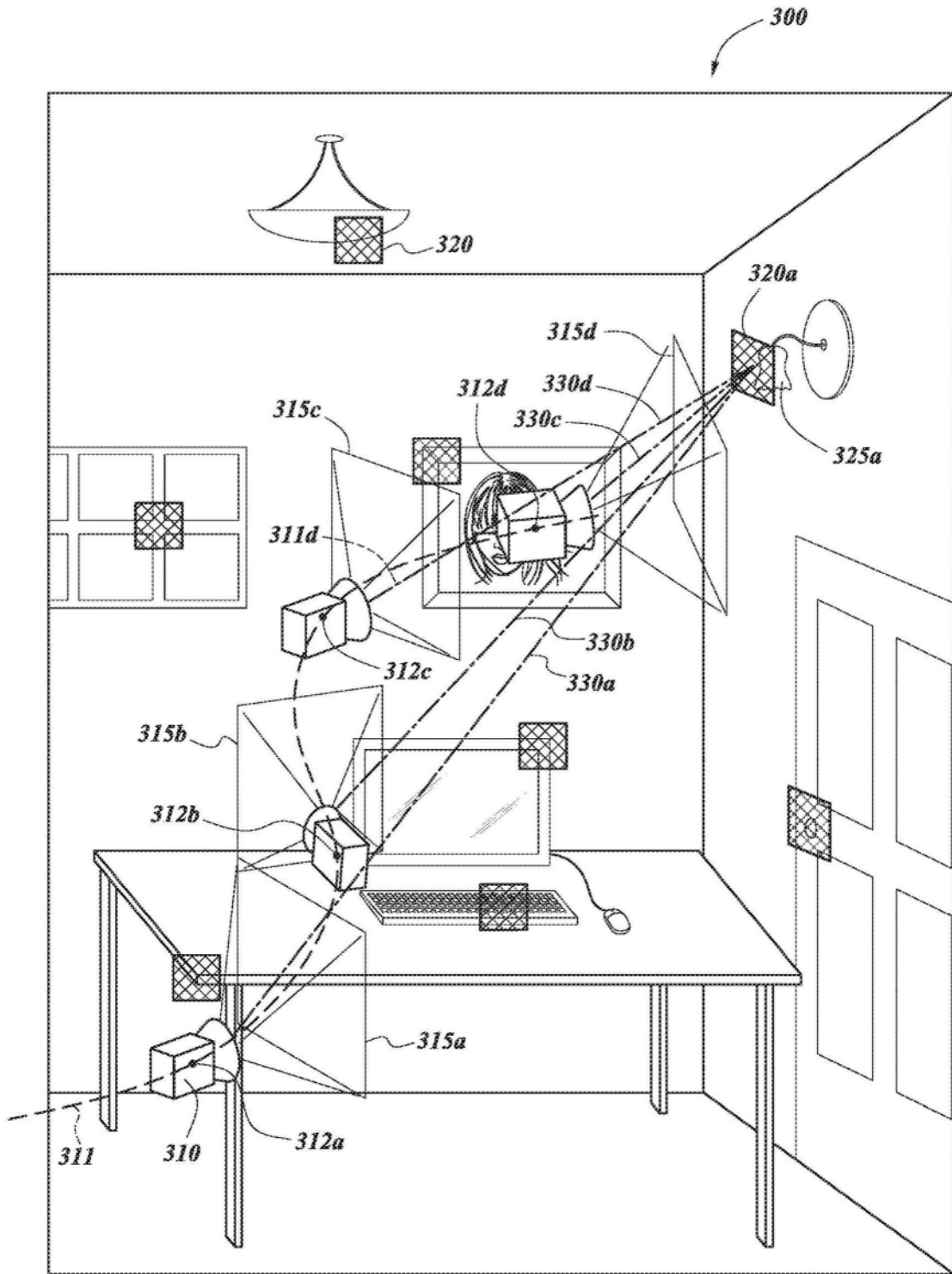


图3

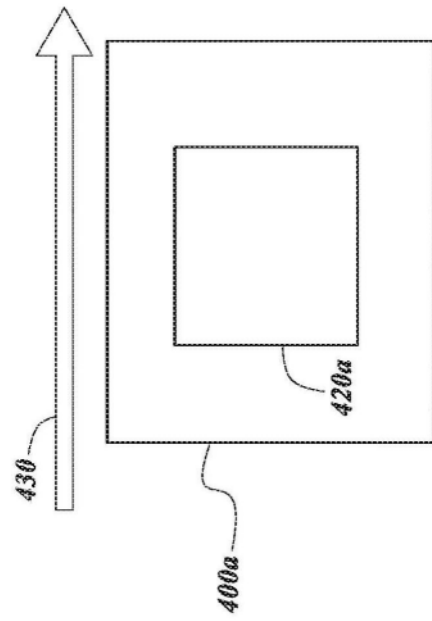


图4A

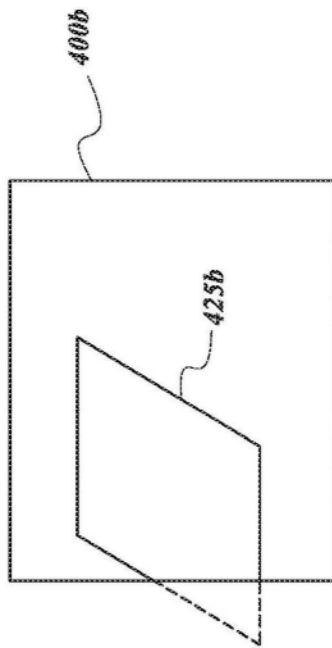


图4B

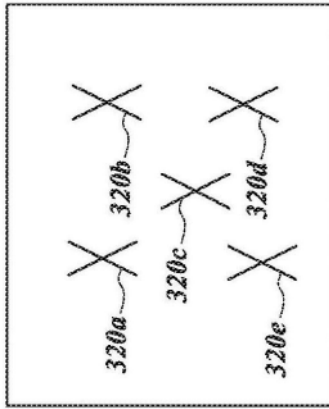


图5A

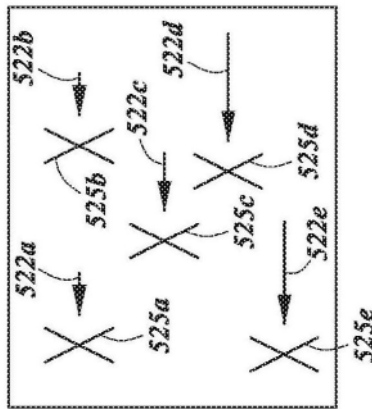


图5B

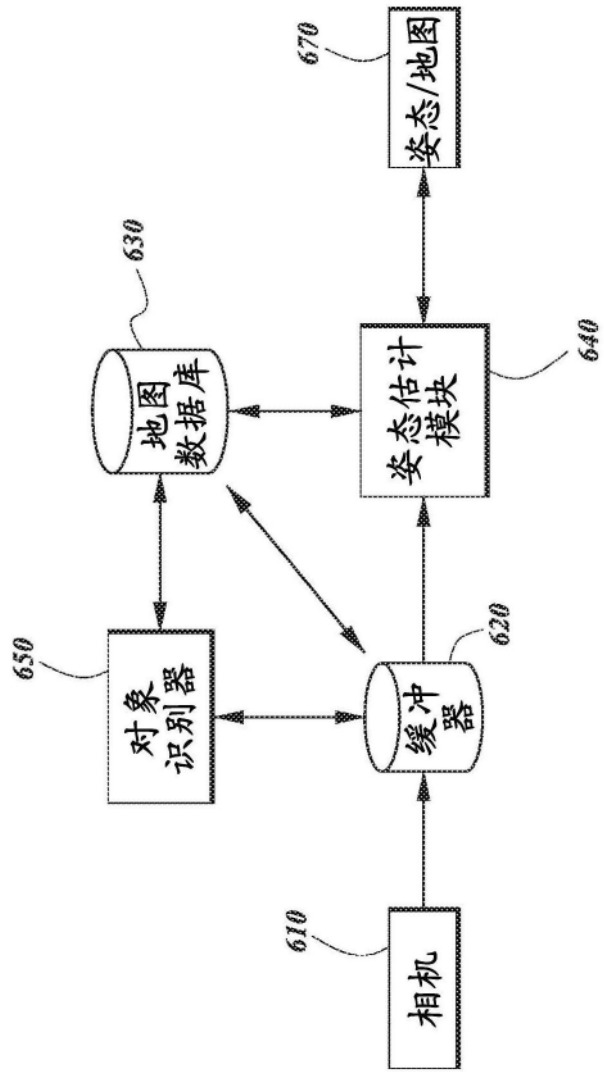


图6

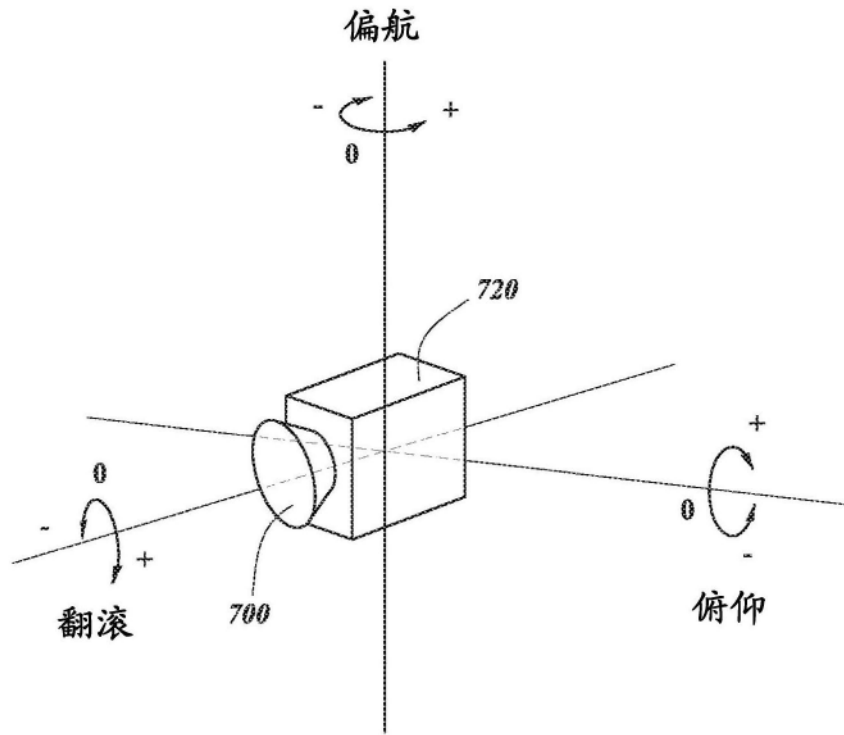


图7

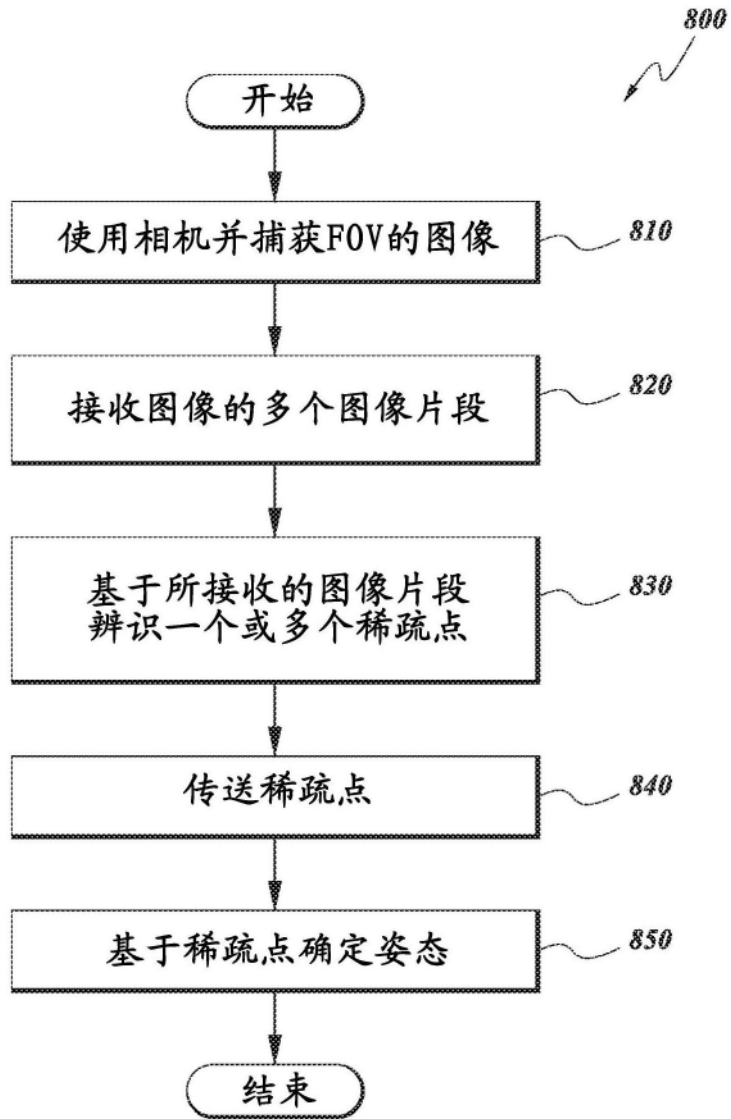


图8

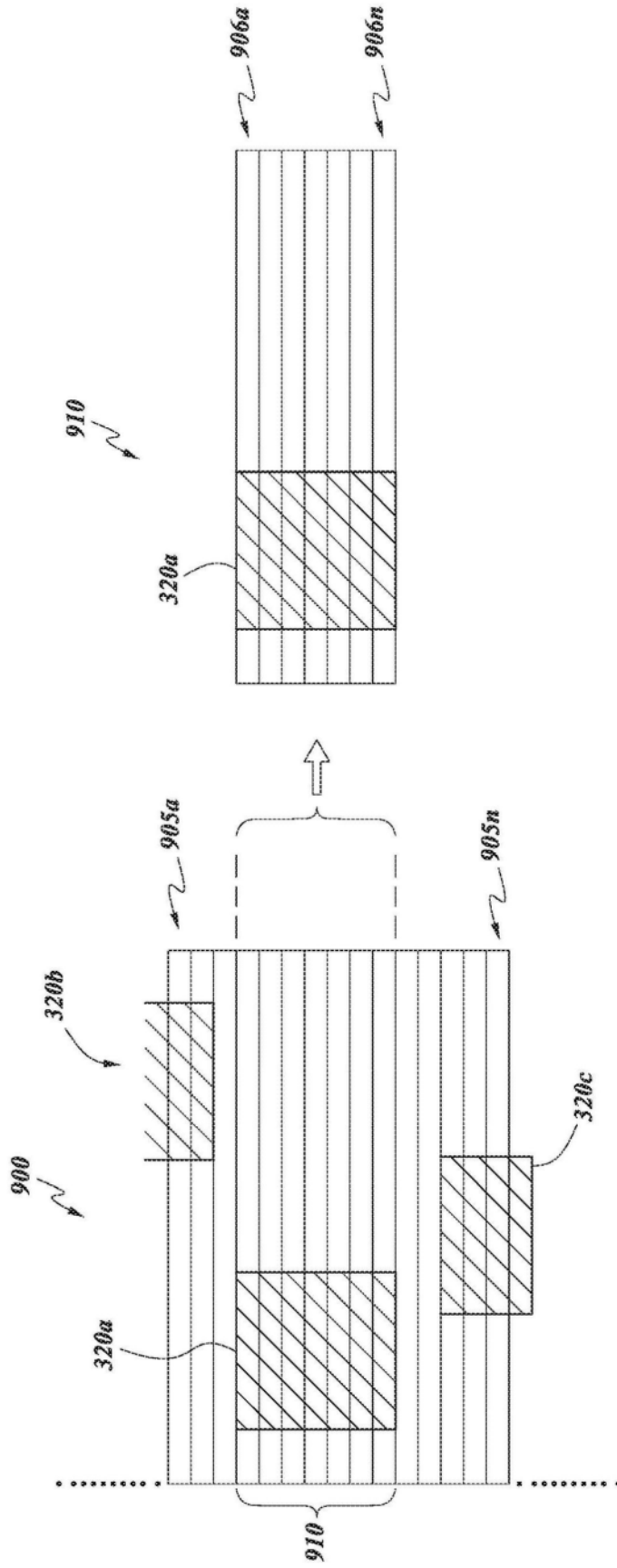


图9B

图9A

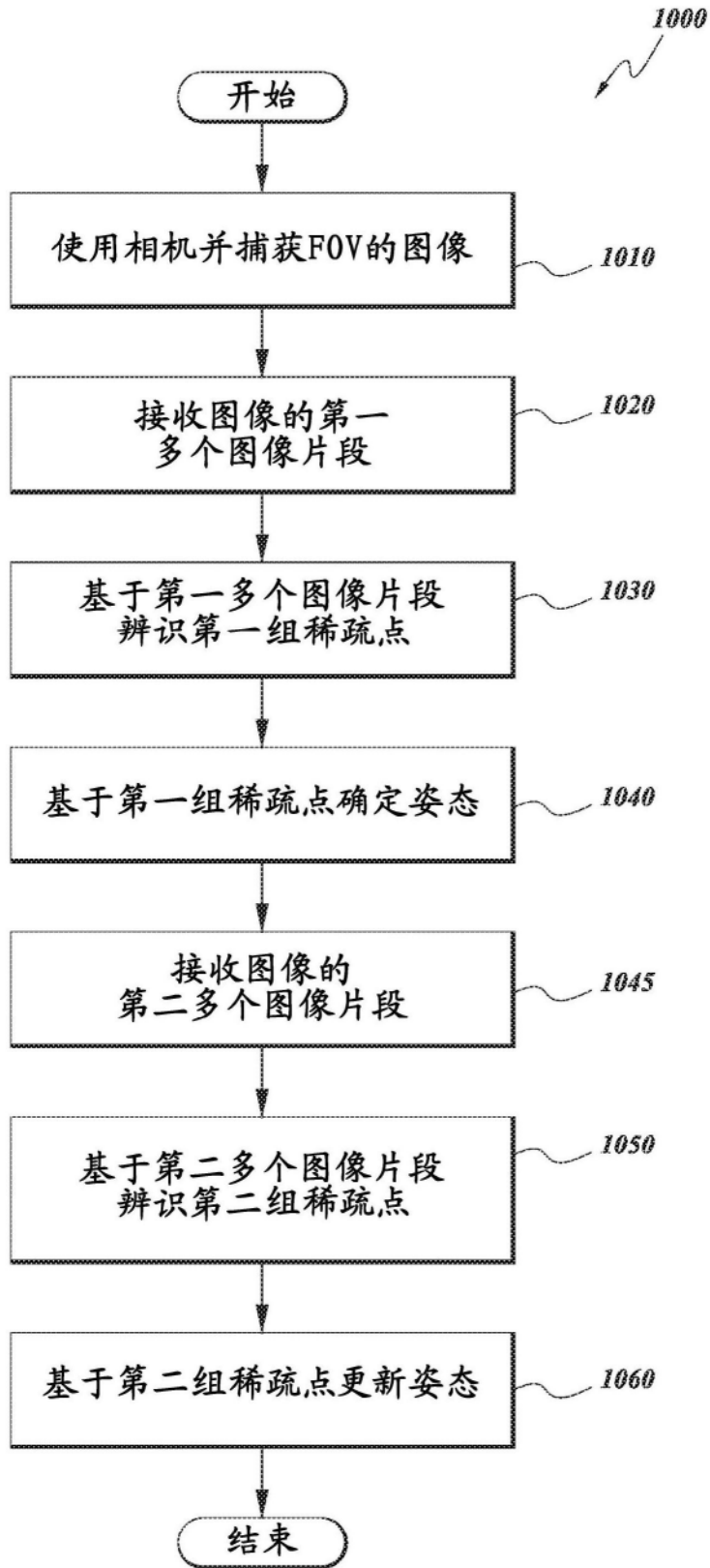


图10