



(12) 发明专利

(10) 授权公告号 CN 112424832 B

(45) 授权公告日 2025. 05. 06

(21) 申请号 201980047687.5

(22) 申请日 2019.07.17

(65) 同一申请的已公布的文献号
申请公布号 CN 112424832 A

(43) 申请公布日 2021.02.26

(30) 优先权数据
16/037,899 2018.07.17 US

(85) PCT国际申请进入国家阶段日
2021.01.15

(86) PCT国际申请的申请数据
PCT/KR2019/008843 2019.07.17

(87) PCT国际申请的公布数据
W02020/017890 EN 2020.01.23

(73) 专利权人 三星电子株式会社
地址 韩国京畿道水原市灵通区三星路129号

(72) 发明人 玉真秀夫 李硕竣
玛尼什·戈埃尔

(74) 专利代理机构 北京英赛嘉华知识产权代理有限公司 11204
专利代理师 王达佐 杨莘

(51) Int.Cl.
G06T 7/70 (2006.01)
G06T 7/11 (2006.01)
G06T 7/13 (2006.01)
G06T 15/10 (2006.01)

(56) 对比文件
US 2017243352 A1, 2017.08.24

审查员 封泉州

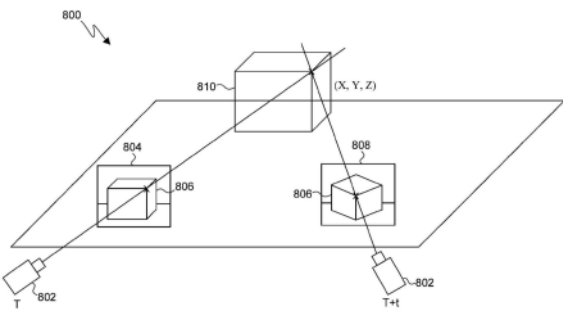
权利要求书3页 说明书14页 附图9页

(54) 发明名称

用于检测对象的3D关联的系统与方法

(57) 摘要

提供了用于检测对象的3D关联的电子装置、方法和计算机可读介质。电子装置包括至少一个图像传感器,惯性测量传感器,存储器,以及联接到至少一个图像传感器、惯性测量传感器和存储器的至少一个处理器。至少一个处理器配置为使用至少一个图像传感器来拍摄环境的图像;在所拍摄的图像中检测对象;限定图像中围绕所检测的对象的有界区域;从惯性测量传感器接收头部姿态数据,以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。



1. 电子装置, 包括:
至少一个图像传感器;
惯性测量传感器;
存储器; 以及
至少一个处理器, 联接到所述至少一个图像传感器、所述惯性测量传感器和所述存储器,
其中, 所述至少一个处理器配置为:
使用所述至少一个图像传感器拍摄环境的图像;
在所拍摄的图像中检测对象;
限定所述图像中围绕所检测的对象的有界区域;
从所述惯性测量传感器接收头部姿态数据;
确定可用系统资源;
基于所确定的可用系统资源切换到多个三维3D关联方法之中的指定3D关联方法; 以及
根据所述指定3D关联方法, 使用所述头部姿态数据和所拍摄的图像中的所述有界区域来确定所检测的对象在3D空间中的位置,
其中, 所述至少一个处理器还配置为:
基于确定出所述可用系统资源不足以允许在所拍摄的图像中生成2D特征点, 切换到所述多个3D关联方法之中的像素匹配三角测量方法,
其中, 在所述像素匹配三角测量方法中, 为了确定所检测的对象的位置, 所述至少一个处理器还配置为:
在位于所述有界区域内的像素之间实施像素匹配,
将匹配的像素设定为将在三角测量中使用的特征点, 以及
使用所设定的特征点进行三角测量以确定所检测的对象的位置。
2. 根据权利要求1所述的电子装置, 其中, 所述至少一个处理器还配置为: 基于确定出所述可用系统资源足以提供3D特征点, 切换到所述多个3D关联方法之中的重投影方法,
其中, 在所述重投影方法中, 为了确定所检测的对象的位置, 所述至少一个处理器还配置为:
使用所述至少一个图像传感器和所述惯性测量传感器生成所述环境的至少一部分的3D点云;
将所述3D点云的特征点投影到所拍摄的图像中, 其中, 所拍摄的图像包括所述有界区域;
确定所述特征点中位于所述对象的表面上的特征点; 以及
移除不位于所述对象的所述表面上的特征点。
3. 根据权利要求2所述的电子装置, 其中, 所述至少一个处理器还配置为:
确定出稀疏图被用于生成所述3D点云;
生成所述有界区域的密集图3D点云; 以及
将所述密集图3D点云的特征点投影到所拍摄的图像的所述有界区域中。
4. 根据权利要求1所述的电子装置, 其中, 所述至少一个处理器还配置为:
基于确定出可用系统资源足以允许在所拍摄的图像中生成2D特征点但不足以提供3D

特征点,切换到所述多个3D关联方法之中的2D特征点三角测量方法,

其中,在所述2D特征点三角测量方法中,为了确定所检测的对象的位置,所述至少一个处理器还配置为:

使用所述至少一个图像传感器拍摄多个图像;

在所述多个图像中的每个中检测所述对象;

限定所述多个图像中的有界区域,其中,所述有界区域中的每个有界区域均限定在所述多个图像中的一个图像中的所检测的对象周围;以及

利用所述多个图像中的至少另一个图像中的所述对象的位置对所述多个图像中的一个图像中的所述对象的位置进行三角测量。

5.根据权利要求4所述的电子装置,其中,为了对所述对象的位置进行三角测量,所述至少一个处理器还配置为:

在所述多个图像中的每个中生成2D特征点;

确定位于所述多个图像中的每个所述有界区域内的所述2D特征点;以及

利用位于所述多个图像中的另一个图像的所述有界区域中的另一个有界区域内的所述2D特征点对位于所述多个图像中的一个图像的所述有界区域中的一个有界区域内的所述2D特征点实施三角测量。

6.根据权利要求4所述的电子装置,其中,为了对所述对象的位置进行三角测量,所述至少一个处理器还配置为:

在所述多个图像中的至少两个图像的每个所述有界区域之间实施像素匹配;以及

在所述多个图像中的所述至少两个图像的匹配像素之间实施三角测量。

7.根据权利要求1所述的电子装置,其中,所述至少一个处理器还配置为将虚拟对象与所检测的对象相邻地放置在所述3D空间中,其中,所检测的对象和所述虚拟对象与对象类关联。

8.用于检测对象的3D关联的方法,所述方法包括:

使用至少一个图像传感器拍摄环境的图像;

在所拍摄的图像中检测对象;

限定所述图像中围绕所检测的对象的有界区域;

从惯性测量传感器接收头部姿态数据;

确定可用系统资源;

基于所确定的可用系统资源切换到多个三维3D关联方法之中的指定3D关联方法;以及

根据所述指定3D关联方法,使用所述头部姿态数据和所拍摄的图像中的所述有界区域、使用处理器来确定所检测的对象在3D空间中的位置,

其中,切换到所述指定3D关联方法包括:基于确定出所述可用系统资源不足以允许在所拍摄的图像中生成2D特征点,切换到所述多个3D关联方法之中的像素匹配三角测量方法,

其中,在所述像素匹配三角测量方法中,确定所检测的对象的位置包括:

在位于所述有界区域内的像素之间实施像素匹配,

将匹配的像素设定为将在三角测量中使用的特征点,以及

使用所设定的特征点进行三角测量以确定所检测的对象在3D空间中的位置。

9. 根据权利要求8所述的方法, 其中, 切换到所述指定3D关联方法包括: 基于确定出所述可用系统资源足以提供3D特征点, 切换到所述多个3D关联方法之中的重投影方法,

其中, 在所述重投影方法中, 确定所检测的对象的位置包括:

使用所述至少一个图像传感器和所述惯性测量传感器生成所述环境的至少一部分的3D点云;

将所述3D点云的特征点投影到所拍摄的图像中, 其中, 所拍摄的图像包括所述有界区域;

确定所述特征点中位于所述对象的表面上的特征点; 以及

移除不位于所述对象的所述表面上的特征点。

10. 根据权利要求9所述的方法, 确定所检测的对象的所述位置还包括:

确定出稀疏图被用于生成所述3D点云;

生成所述有界区域的密集图3D点云; 以及

将所述密集图3D点云的特征点投影到所拍摄的图像的所述有界区域中。

11. 根据权利要求8所述的方法, 其中, 切换到所述指定3D关联方法包括: 基于确定出可用系统资源足以允许在所拍摄的图像中生成2D特征点但不足以提供3D特征点, 切换到所述多个3D关联方法之中的2D特征点三角测量方法,

其中, 在所述2D特征点三角测量方法中, 确定所检测的对象的位置包括:

使用所述至少一个图像传感器拍摄多个图像;

在所述多个图像中的每个图像中检测所述对象;

在所述多个图像中限定有界区域, 其中, 所述有界区域中的每个有界区域均限定在所述多个图像中的一个图像中的所检测的对象周围; 以及

利用所述多个图像中的至少另一个图像中的所述对象的位置对所述多个图像中的一个图像中的所述对象的位置进行三角测量。

12. 根据权利要求11所述的方法, 其中, 对所述对象的位置进行三角测量包括:

在所述多个图像中的每个图像中生成2D特征点;

确定位于所述多个图像中的每个所述有界区域内的所述2D特征点; 以及

利用位于所述多个图像中的另一个图像的所述有界区域中的另一个有界区域内的所述2D特征点对位于所述多个图像中的一个图像的所述有界区域中的一个有界区域内的所述2D特征点进行三角测量。

13. 根据权利要求11所述的方法, 其中, 对所述对象的位置进行三角测量包括:

在所述多个图像中的至少两个图像的每个所述有界区域之间匹配像素; 以及

对所述多个图像中的所述至少两个图像的匹配像素进行三角测量。

14. 根据权利要求8所述的方法, 还包括将虚拟对象与所检测的对象相邻地放置在所述3D空间中, 其中, 所检测的对象和所述虚拟对象与对象类关联。

用于检测对象的3D关联的系统与方法

技术领域

[0001] 本公开大体上涉及增强现实系统与方法。更具体地,本公开涉及用于检测对象的3D关联的系统与方法。

背景技术

[0002] 3D空间中的对象检测是增强现实应用的重要方面。然而,增强现实应用提出了关于确定3D空间中检测对象的比例和深度的挑战。

发明内容

[0003] 问题的方案

[0004] 本公开提供了用于检测对象的3D关联的系统与方法。

[0005] 在一个实施例中,提供了用于检测对象的3D关联的电子装置。电子装置包括至少一个图像传感器、惯性测量传感器、存储器,以及联接到至少一个图像传感器、惯性测量传感器和存储器的至少一个处理器。至少一个处理器配置为使用至少一个图像传感器拍摄环境的图像;在所拍摄的图像中检测对象;限定图像中围绕所检测的对象的有界区域;从惯性测量传感器或从配置为使用惯性测量传感器和图像传感器数据来计算头部姿态的另一处理器接收头部姿态数据;以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。

[0006] 在另一个实施例中,提供了用于检测对象的3D关联的方法。该方法包括:使用至少一个图像传感器拍摄环境的图像;在拍摄的图像中检测对象;限定图像中围绕所检测的对象的有界区域;从惯性测量传感器或配置为使用惯性测量传感器和图像传感器数据计算头部姿态的另一处理器接收头部姿态数据;以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。

[0007] 在又一实施例中,提供了用于操作电子装置的体现计算机程序的非暂时性介质,其中,电子装置用于检测对象的3D关联。电子装置包括至少一个图像传感器、惯性测量传感器、存储器和至少一个处理器。当由至少一个处理器执行程序代码时,程序代码使得电子装置使用至少一个图像传感器拍摄环境的图像;在所拍摄的图像中检测对象;限定图像中围绕所检测的对象的有界区域;从惯性测量传感器或从配置为使用惯性测量传感器和图像传感器数据计算头部姿态的另一处理器接收头部姿态数据;以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。

[0008] 根据以下附图、说明和权利要求,其他技术特征对于本领域技术人员来说可能是显而易见的。

[0009] 发明的有益效果

[0010] 本公开提供了三角测量过程,该三角测量过程允许增强现实标记应用即使在较少的系统资源可用时也能继续有效地运行。这对于其中用户利用摄像机来提供场景或环境的连续视图的增强现实应用是重要的,因为用户期望该应用的用户装置上显示关于环境中的

对象的准确且连续更新的信息。

附图说明

[0011] 为了更完整地理解本公开及其优点,现参考结合附图的以下描述,其中,相同的附图标记代表相同的部件:

[0012] 图1示出了根据本公开的实施例的示例网络配置;

[0013] 图2示出了根据本公开的实施例的用于与增强现实标记应用一起使用的示例拍摄图像;

[0014] 图3示出了根据本公开的实施例的增强现实对象检测和3D关联架构的示例;

[0015] 图4示出了根据本公开的实施例的示例对象检测和3D关联过程的流程图;

[0016] 图5示出了根据本公开的实施例的用于检测对象的3D关联的3D特征点重投影过程的图解视图;

[0017] 图6示出了根据本公开的实施例的3D特征点重投影过程的另一个说明性示例的图;

[0018] 图7示出了根据本公开的实施例的示例3D特征点重投影过程的流程图;

[0019] 图8示出了根据本公开的实施例的示例特征点三角测量过程的图;以及

[0020] 图9示出了根据本公开的实施例的示例检测对象三角测量过程的流程图。

具体实施方式

[0021] 在一个实施例中,提供了用于检测对象的3D关联的电子装置。电子装置包括至少一个图像传感器,惯性测量传感器,存储器,以及联接到至少一个图像传感器、惯性测量传感器和存储器的至少一个处理器。至少一个处理器配置为使用至少一个图像传感器来拍摄环境的图像;在所拍摄的图像中检测对象;限定图像中围绕所检测的对象的有界区域;从惯性测量传感器或从配置为使用惯性测量传感器和图像传感器数据来计算头部姿态的另一处理器接收头部姿态数据;以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。

[0022] 在实施例中,为了确定所检测的对象的位置,至少一个处理器还配置为使用至少一个图像传感器和惯性测量传感器生成环境的至少一部分的3D点云;将3D点云的特征点投影到所拍摄的图像中,其中,所拍摄的图像包括有界区域;确定特征点中的哪些特征点位于对象的表面上,并移除不位于对象的表面上的特征点。

[0023] 在实施例中,至少一个处理器还配置为确定稀疏图被用于生成3D点云,生成有界区域的密集图3D点云,以及将密集图3D点云的特征点投影到所拍摄的图像的有界区域中。

[0024] 在实施例中,为了确定所检测的对象的位置,至少一个处理器还配置为使用至少一个图像传感器拍摄多个图像;在多个图像中的每个图像中检测对象,限定多个图像中的有界区域,其中,有界区域中的每个有界区域均被限定为在多个图像中的一个中围绕所检测的对象,并且利用对象在多个图像中的至少另一个图像中的位置来对对象在多个图像中的一个图像中的位置进行三角测量。

[0025] 在实施例中,为了对对象的位置进行三角测量,至少一个处理器还配置为在多个图像中的每个图像中生成2D特征点,确定位于多个图像中的有界区域中的每个有界区域内

的2D特征点,以及利用位于多个图像中的另一个图像的另一个有界区域内的2D特征点对位于多个图像中的一个图像的有界区域中的一个有界区域内的2D特征点实施三角测量。

[0026] 在实施例中,为了对特征点进行三角测量,至少一个处理器还配置为在多个图像中的至少两个图像的每个有界区域之间实施像素匹配,以及在多个图像中的至少两个图像的匹配像素之间实施三角测量。

[0027] 在实施例中,至少一个处理器还配置为将虚拟对象与所检测的对象相邻地放置在3D空间中,其中,所检测的对象和虚拟对象与对象类有关。

[0028] 在另一个实施例中,提供了用于检测对象的3D关联的方法。该方法包括:使用至少一个图像传感器拍摄环境的图像;在所拍摄的图像中检测对象;限定图像中围绕所检测的对象的有界区域;从惯性测量传感器或从配置为使用惯性测量传感器和图像传感器数据计算头部姿态的另一处理器接收头部姿态数据;以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。

[0029] 在又一实施例中,提供了用于操作电子装置的、体现计算机程序的非暂时性介质,其中,电子装置用于检测对象的3D关联。电子装置包括至少一个图像传感器、惯性测量传感器、存储器和至少一个处理器。当由至少一个处理器执行程序代码时,程序代码使得电子装置使用至少一个图像传感器拍摄环境的图像,在所拍摄的图像中检测对象,限定图像中围绕所检测的对象的有界区域,从惯性测量传感器或从配置为使用惯性测量传感器和图像传感器数据计算头部姿态的另一处理器接收头部姿态数据,以及使用头部姿态数据和所拍摄的图像中的有界区域来确定所检测的对象在3D空间中的位置。

[0030] 发明模式

[0031] 参考附图描述了下面讨论的图1至图9以及本公开的各种实施例。然而,应了解,本公开不限于实施例,并且其所有改变和/或等效物或替代也属于本公开的范围。在整个说明书和附图中,相同或相似的附图标记可用于指代相同或相似的元件。

[0032] 在进行以下详细说明之前,阐述贯穿本专利文件所使用的某些词语和短语的定义可能是有利的。术语“传输”、“接收”和“通信”及其派生词包括直接和间接通信。术语“包括”和“包含”及其派生词意指非限制性地包括。术语“或”是包括性的,意指和/或。短语“与…有关”及其派生词意指包括、被包括在…内、与…互连、包含、被包含在…内、连接到或与…连接、联接到或与…联接、可与…通信、与…协作、交织、并列、接近、绑定到或与…绑定、具有、具有…的特性、与…有关系,等等。

[0033] 此外,下面所描述的各种功能可由一个或多个计算机程序来实现或支持,每个计算机程序由计算机可读程序代码形成,并体现在计算机可读介质中。术语“应用”和“程序”是指一个或多个计算机程序、软件组件、指令集、过程、函数、对象、类、实例、相关数据,或适于在适当的计算机可读程序代码中实现的其部分。短语“计算机可读程序代码”包括任何类型的计算机代码,该计算机代码包括源代码、对象代码和可执行代码。短语“计算机可读介质”包括能够由计算机访问的任何类型的介质,诸如,只读存储器(ROM)、随机存取存储器(RAM)、硬盘驱动、光盘(CD)、数字视频光盘(DVD)或任何其它类型的存储器。“非暂时性”计算机可读介质排除了有线、无线、光或其它通信链路,这些链路运输暂时性的电信号或其它信号。非暂时性计算机可读介质包括可永久存储数据的介质以及可存储数据并随后重写数据的介质,诸如,可重写光盘或可擦除存储装置。

[0034] 如本文所用,术语“具有”、“可具有”、“包括”或“可包括”特征(例如,数字、函数、操作或诸如部件的组件)指示该特征的存在且不排除其它特征的存在。

[0035] 如本文所用,术语“A或B”、“A和/或B中的至少一个”或“A和/或B中的一个或多个”可包括A和B的所有可能的组合。例如,“A或B”、“A和B中的至少一个”、“A或B中的至少一个”可指示(1)包括至少一个A、(2)包括至少一个B,或(3)包括至少一个A和至少一个B中的所有。

[0036] 如本文所用,不管重要性如何,术语“第一”和“第二”可修改各种组件,且不限制这些组件。这些术语仅用于将一个组件与另一个组件区分开。例如,不管装置的顺序或重要性如何,第一用户装置和第二用户装置可指示彼此不同的用户装置。例如,第一组件可被表示为第二组件,且在不背离本公开的范围的情况下,反之亦然。

[0037] 应当理解,当元件(例如,第一元件)被称为(可操作地或通信地)“与另一元件(例如,第二元件)联接或联接至另一元件(例如,第二元件)”或“与另一元件(例如,第二元件)连接或连接至另一元件(例如,第二元件)”时,其可直接地或经由第三元件与另一元件联接或联接至另一元件或与另一元件连接或连接至另一元件。相反,应当理解,当元件(例如,第一元件)被称为“直接与另一个元件(例如,第二元件)联接/联接至另一个元件(例如,第二元件)”或“直接与另一个元件(例如,第二元件)连接/连接至(例如,第二元件)”时,没有其它元件(例如,第三元件)介于该元件与其它元件之间。

[0038] 如本文所用,术语“配置(或设定)为”可与术语“适于”、“具有…的能力”、“设计为”、“适应于”、“使…做”或“能够”依据状况而互换使用。术语“配置(或设定)为”基本上不意指“专门设计为硬件以”。相反,术语“配置为”可意指装置可与另一个装置或部件一起实施操作。

[0039] 例如,术语“配置(或设定)为实施A、B和C的处理器”可意指可通过实施存储在存储器装置中的一个或多个软件程序来执行操作的通用处理器(例如,CPU或应用处理器),或者用于实施操作的专用处理器(例如,嵌入式处理器)。

[0040] 本文所用的术语仅提供以描述其一些实施例,而不是限制本公开的其它实施例的范围。应当理解,单数形式“一个”、“一种”和“该”包括复数引用,除非上下文另有明确规定。本文所用的包括技术和科学术语的所有术语具有与本公开的实施例所属领域的普通技术人员通常理解的相同的含义。还应当理解,诸如在常用词典中定义的那些术语应当被解释为具有与它们在相关技术的上下文中的含义一致的含义,且将不会被解释为理想化的或过分正式的意义,除非在本文明确地如此定义。在一些情况下,本文定义的术语可被解释为排除本公开的实施例。

[0041] 例如,根据本公开的实施例的电子装置的示例可包括智能电话、平板个人计算机(PC)、移动电话、视频电话、电子书阅读器、台式PC、膝上型计算机、网络书计算机、工作站、PDA(个人数字助理)、便携式多媒体播放器(PMP)、MP3播放器、移动医疗装置、摄像机或可穿戴装置(例如,智能眼镜、头戴式装置(HMD)、电子衣服、电子手镯、电子项链、电子附件、电子纹身、智能镜或智能手表)中的至少一个。

[0042] 在整个本专利文件中提供了对其它某些单词和短语的定义。本领域技术人员应了解,在许多(如果不是大多数)实例中,此类定义适用于此类经定义的词语和短语的先前及将来的使用。

[0043] 本申请中的任何被理解为暗示任何特定元素、步骤或功能的描述都不是必须包括在权利要求范围内的必要元素。专利可授权主题的范围仅由权利要求限定。此外,权利要求中的任何一个都不旨在调用35U.S.C.112(f),除非确切的词语“用于……的装置”后面跟有分词。在权利要求内使用的任何其它术语,包括但不限于“机构”、“模块”、“装置”、“单元”、“组件”、“元件”、“构件”、“设备”、“机器”、“系统”、“处理器”或“控制器”,被申请人理解是指相关领域的技术人员已知的结构,并不旨在调用35U.S.C.112(f)。

[0044] 尽管已经利用示例性实施例描述了本公开,但是可向本领域技术人员建议各种改变和修改。本公开旨在包括落入所附权利要求的范围内的此类改变和修改。

[0045] 3D空间中的对象检测是增强现实应用的重要方面。增强现实应用提出了关于确定3D空间中检测对象的比例和深度的挑战。例如,检测对象的比例被用于确定所检测的对象的性质。例如,如果在图像中检测到汽车,则汽车的比例可确定该汽车是玩具还是真实的汽车。尽管3D对象检测(诸如,通过使用卷积神经网络和几何结构)是可能的,但是此类方法在计算上可能是昂贵的。因此,可在2D域中实施用于增强现实的对象检测,这提供了更少的计算,但是对象仍需要与3D空间关联。虽然可使用深度或立体摄像机来确定3D空间理解,但是也可使用同时定位和映射(SLAM)来实现3D空间理解。

[0046] 图1示出了根据本公开的各种实施例的示例网络配置100。图1所示的网络配置100的实施例仅用于说明。在不背离本公开的范围的情况下,可使用网络配置100的其它实施例。

[0047] 根据本公开的实施例,电子装置101被包括在网络配置100中。电子装置101可包括总线110、处理器120、存储器130、输入/输出(I/O)接口150、显示器160、通信接口170或传感器180中的至少一个。在一些实施例中,电子装置101可排除至少一个组件或者可添加另一个组件。

[0048] 总线110包括用于将组件120至组件170彼此连接并在组件之间传送通信(例如,控制消息和/或数据)的电路。

[0049] 处理器120包括中央处理单元(CPU)、应用处理器(AP)或通信处理器(CP)中的一个或多个。处理器120能够对电子装置101的其它组件中的至少一个实施控制,和/或实施涉及通信的操作或数据处理。在一些实施例中,处理器可以是图形处理器单元(GPU)。

[0050] 例如,处理器120可接收在拍摄事件期间由摄像机拍摄的多个帧。处理器120可检测位于所拍摄的多个帧中的一个或多个帧内的对象。处理器120可限定围绕检测对象的有界区域。处理器120还可从惯性测量单元接收头部姿态信息。处理器120可使用头部姿态信息和所拍摄的帧中的有界区域来确定所检测的对象在3D空间中的位置。处理器120可将虚拟对象在所检测的对象的确定的位置附近放置在3D空间中。处理器120可操作显示器以显示包括所检测的对象和虚拟对象的3D空间。

[0051] 存储器130可包括易失性存储器和/或非易失性存储器。例如,存储器130可存储涉及电子装置101的至少一个其它组件的命令或数据。根据本公开的实施例,存储器130可存储软件和/或程序140。程序140包括,例如:内核141、中间件143、应用编程接口(API) 145,和/或应用程序(或“应用”) 147。内核141、中间件143或API 145的至少一部分可被表示为操作系统(OS)。

[0052] 例如,内核141可控制或管理用以实施在其它程序(例如,中间件143、API 145或应

用程序147)中实现的操作或功能的系统资源(例如,总线110、处理器120或存储器130)。内核141提供允许中间件143、API 145或应用147访问电子装置101的各个组件以控制或管理系统资源的接口。应用147包括一个或多个应用,用于图像拍摄、对象检测、头部或摄像机姿态、同时定位和映射(SLAM)处理、以及对象分类和标记。这些功能可由单个应用或多个应用来实施,其中,多个应用中的每个应用实行这些功能中的一个或多个功能。

[0053] 例如,中间件143可用作允许API 145或应用147与内核141通信数据的中继。可设置多个应用147。中间件143能够例如通过将使用电子装置101的系统资源(例如,总线110、处理器120或存储器130)的优先级分配给多个应用147中的至少一个,来控制从应用147接收的工作请求。

[0054] API 145是允许应用147控制从内核141或中间件143提供的功能的接口。例如,API 145包括用于档案控制、窗口控制、图像处理或文本控制的至少一个接口或功能(例如,命令)。

[0055] IO接口150充当可以例如将从用户或其它外部装置输入的命令或数据传送到电子装置101的其它组件的接口。此外,IO接口150可将从电子装置101的其它组件接收的命令或数据输出到用户或其它外部装置。

[0056] 显示器160包括,例如:液晶显示器(LCD)、发光二极管(LED)显示器、有机发光二极管(OLED)显示器,或微机电系统(MEMS)显示器,或电子纸显示器。显示器160还可以是深度感知显示器,诸如,多焦点显示器。显示器160能够向用户显示例如各种内容(例如,文本、图像、视频、图标或符号)。显示器160可包括触摸屏,并可使用电子笔或用户的身体部分接收例如触摸、手势、接近或悬停输入。

[0057] 例如,通信接口170能够在电子装置101与外部电子装置(例如,第一电子装置102、第二电子装置104或服务器106)之间建立通信。例如,通信接口170可通过无线通信或有线通信与网络162或网络164连接,以与外部电子装置通信。通信接口170可以是有线收发器或无线收发器或用于传输和接收信号(诸如,图像、头部或摄像机姿态数据或其它信息)的任何其它组件。

[0058] 电子装置101还包括一个或多个传感器180,一个或多个传感器180可计量物理量或检测电子装置101的激活状态,并将所计量的或所检测的信息转换为电信号。例如,传感器180可包括用于触摸输入的一个或多个按钮、一个或多个摄像机、手势传感器、陀螺仪或陀螺仪传感器、气压传感器、磁传感器或磁力计、加速度传感器或加速度计、握持传感器、接近传感器、颜色传感器(例如,红绿蓝(RGB)传感器)、生物物理传感器、温度传感器、湿度传感器、照度传感器、紫外(UV)传感器、肌电图(EMG)传感器、脑电图(EEG)传感器、心电图(ECG)传感器、IR传感器、超声传感器、虹膜传感器、指纹传感器等。传感器180还可包括惯性测量单元,该惯性测量单元可包括一个或多个加速度计、陀螺仪和其它组件。传感器180还可包括用于控制其中所包括的至少一个传感器的控制电路。这些传感器180中的任何传感器都可位于电子装置101内。一个或多个摄像机可拍摄用于对象检测和3D环境映射的图像。惯性测量单元可跟踪电子装置101的头部姿态和/或其它运动。

[0059] 第一外部电子装置102或第二外部电子装置104可以是可穿戴装置或可安装电子装置101的可穿戴装置(例如,头戴式显示器(HMD))。当电子装置101安装在电子装置102(例如,HMD)中时,电子装置101可通过通信接口170与电子装置102进行通信。电子装置101可直

接地与电子装置102连接,以在不涉及单独的的网络的情况下与电子装置102进行通信。电子装置101还可以是增强现实可穿戴装置,诸如,眼镜,该眼镜包括用于检测对象的3D关联的SLAM和对象检测的一个或多个摄像机。

[0060] 无线通信能够使用例如长期演进 (LTE)、长期高级演进 (LTE-A)、第五代无线系统 (5G)、毫米波或60GHz无线通信、无线USB、码分多址 (CDMA)、宽带码分多址 (WCDMA)、通用移动通信系统 (UMTS)、无线宽带 (WiBro) 或全球移动通信系统 (GSM) 中的至少一种作为蜂窝通信协议。有线连接可包括通用串行总线 (USB)、高清晰度多媒体接口 (HDMI)、推荐的标准232 (RS-232) 或普通老式电话服务 (POTS) 中的至少一种。

[0061] 网络162包括通信网络中的至少一种,例如,计算机网络(例如,局域网 (LAN) 或广域网 (WAN))、因特网或电话网络。

[0062] 第一外部电子装置102和第二外部电子装置104以及服务器106均可以是与电子装置101相同类型或不同类型的装置。根据本公开的某些实施例,服务器106包括一组一个或多个服务器。根据本公开的某些实施例,在电子装置101上实施的所有操作或一些操作可在另一个或多个其他电子装置(例如,电子装置102和电子装置104或服务器106)上实施。根据本公开的某些实施例,当电子装置101应自动地或根据请求执行一些功能或服务时,电子装置101可请求另一装置(例如,电子装置102和电子装置104或服务器106)执行与其关联的至少一些功能,而不是独立地或附加地执行该功能或服务。其它电子装置(例如,电子装置102和电子装置104或服务器106)能够执行所请求的功能或附加功能,并将执行的结果传送到电子装置101。电子装置101可通过按照原样或附加地处理所接收的结果来提供请求的功能或服务。为此,例如,可使用云计算、分布式计算或客户端-服务器计算技术。

[0063] 尽管图1显示了电子装置101包括通信接口170以经由网络162与外部电子装置104或服务器106进行通信,但是根据本公开的实施例,电子装置101可以在没有单独的通信功能的情况下独立地运行。

[0064] 服务器106可支持通过实施在电子装置101上实现的操作(或功能)中的至少一个来驱动电子装置101。例如,服务器106可包括可支持在电子装置101中实现的处理器120的处理模块或处理器。

[0065] 尽管图1示出了网络配置100的一个示例,但是可对图1进行各种改变。例如,网络配置100可在任何适当的排列中包括任何数量的每种组件。通常,计算和通信系统具有多种配置,并且图1不会将本公开的范围限制于任何特定配置。虽然图1示出了其中可使用本专利文件中公开的各种特征的一个操作环境,但是这些特征可用在任何其它合适的系统中。

[0066] 图2示出了根据本公开实施例的用于与由处理器(诸如电子装置101的处理器120)执行的增强现实标记应用一起使用的示例拍摄图像200。如图2的示例拍摄图像200所示,所拍摄的图像200包括位于所拍摄的图像200的帧内的各种可检测对象。当处理器检测所拍摄的图像200内的对象的位置时,围绕每个所检测的对象创建并显示有界区域202。处理器通过关联的检测对象在所拍摄的图像200内的定位或位置来限定有界区域202。例如,处理器可确定检测对象在帧内的中心(x,y)坐标,并定义有界区域的宽度(w)和高度(h),以使得整个检测对象在有界区域内。当有界区域是正方形或矩形区域时,且因为所检测的对象可以是各种形状,所以可存在有界区域202的不包括所检测的对象的表面的部分,诸如,包括有界区域202中的周围环境的部分。

[0067] 由处理器执行的应用还对检测对象进行分类,并用虚拟对象对所检测的对象进行标记。可实现图像识别以在预定对象类中对每个对象进行分类。例如,如图2所示,在所拍摄的图像200中的桌子上检测到的杯子被分类为“杯子”类,检测到的时钟被分类为“时钟”类,检测到的香蕉被分类为“香蕉”类,以及检测到的蜂窝电话被分类为“蜂窝电话”类。然后,处理器可用图形204对每个所检测的项目进行标记,在所检测的对象周围的有界区域202附近显示图形204。例如,如图2所示,邻近每个有界区域202的所显示的图形204是具有在球附近文本地显示的所检测的对象的对象类的球。在一些实施例中,所显示的图形也可基于类来着色。例如,图2中所检测的杯子、蜂窝电话、时钟和香蕉的图形204可分别显示为红球、蓝球、绿球和黄球。

[0068] 所检测的对象也可用视觉上类似于所关联的对象类的虚拟对象进行标记。例如,如图2所示,虚拟杯子对象206可放置在所检测/杯子附近。虚拟杯子对象206可具有与所检测的杯子不同的尺寸。为了虚拟对象的放置对用户具有吸引力,虚拟对象相对于所拍摄的图像200中的虚拟对象的关联的检测对象将是成比例的且被良好地放置在3D空间中。因此,应用确定了所检测的对象的比例(诸如,确定是真实的汽车还是玩具汽车),并在所检测的对象附近以正确的比例呈现虚拟对象。例如,如果在桌子上检测到时钟,诸如在图2中,则可将类似但具有不同尺寸的虚拟时钟放置在所检测的时钟附近以进行比较。处理器还确定所检测的对象的深度,以便虚拟对象可被放置在所拍摄的图像中的正确深度处。这对于诸如多焦点显示器的深度感知显示器尤其重要。为了准确地显示虚拟对象,依据本公开的各种实施例,使用SLAM对在其中检测到对象的环境的3D空间进行映射,并且2D有界区域202与3D空间关联。

[0069] 图3示出了根据本公开的实施例的增强现实对象检测和3D关联架构300的示例。架构300包括3D关联块302。3D关联块302限定了由处理器(诸如处理器120)执行的各种功能。处理器从两个有区别的数据管道接收数据以实施检测对象的3D关联。一个管道提供对象检测信息,并且另一个管道提供3D头部或摄像机姿态和3D环境信息(SLAM数据)。对于对象检测管道,对象检测块304通过关联的图像信号处理器(ISP)308从RGB摄像机306接收数据。在对象检测块304处接收的数据包括RGB帧或图像。每个RGB帧或图像可包括时间戳。

[0070] RGB摄像机306与ISP 308之间到对象检测块304的连接可以是运行中(OTF)连接,以使得该连接独立于且平行于相对于架构300描述的其它OTF连接运行。由处理器执行的对象检测块304检测由RGB摄像机306拍摄的RGB图像中的对象。对象检测块304在RGB图像中限定了围绕每个检测对象的有界区域,并向所检测的对象和/或有界区域分配对象ID、对象类和时间戳。对象检测块304还依据本公开的各种实施例来限定围绕所检测的对象的有界区域的位置和尺寸,诸如,通过确定检测对象的(x,y)位置并确定在所拍摄的图像中围绕所检测的对象的有界区域的尺寸(h,w)。所检测的对象和有界区域信息(诸如,对象ID、对象类、时间戳、位置和尺寸)存储在存储器310中。存储器310可以是执行对象检测块304和/或3D关联块302的处理器本地的临时存储器位置或永久存储位置,可以是远程存储位置(诸如服务器)或其它存储位置。

[0071] 对于3D环境管道,例如,3D头部姿态块312通过关联的ISP 316从单色摄像机314接收数据。在3D头部姿态块312处接收的数据包括单色帧或图像,并且每个帧或图像可包括时间戳。在单色摄像机314与ISP 316之间到3D头部姿态块312的连接可以是OTF连接,以使得

该连接独立于并平行于相对于架构300描述的其它OTF连接运行。3D头部姿态块312还通过接口 (IF) 320从惯性测量单元 (IMU) 318接收包括时间戳 (诸如, 每个头部姿态或摄像机姿态的时间戳) 的运动和定向数据。IMU 318与IF 320之间到3D头部姿态块312的连接可以是OTF连接, 以使得该连接独立于并平行于相对于架构300描述的其它OTF连接运行。应当理解, 在一些实施例中, 诸如RGB摄像机306的单个摄像机可用于两个数据管道。例如, 单个摄像机可拍摄用于对象检测的图像, 并用于与IMU数据结合使用以进行SLAM处理。在一些实施例中, RGB摄像机306可以是配置为拍摄对象检测的图像的单色摄像机, 或者可以是另一种类型的图像感测装置。在一些实施例中, 单色摄像机314可以是第二RGB摄像机以向3D头部姿态块312提供数据, 或者可以是另一种类型的图像感测装置。在其它实施例中, 可使用诸如激光扫描仪或声纳装置的其它装置来提供图像、用于生成3D点云的3D空间环境数据、以及其它环境和SLAM数据。

[0072] 由处理器执行的3D头部姿态块312使用由单色摄像机314拍摄的图像以及从IMU 318接收的诸如头部或摄像机姿态的运动和定向数据来生成环境的3D点云。来自3D头部姿态块312的数据 (诸如, 头部姿态, 3D点云, 具有2D特征点的图像, 或其它信息) 可存储在存储器322中。存储器322可以是执行3D头部姿态块312、对象检测块304和/或3D关联块302的处理器本地的临时存储器位置或永久存储位置, 可以是诸如服务器的远程存储位置, 或其它存储位置。存储器322也可以是与存储器310相同的存储器, 或单独的存储器。

[0073] 由处理器执行的3D关联块302从两个数据管道接收数据, 并确定检测对象在3D空间中的位置。依据本文所述的各种实施例, 可通过至少使用与所检测的对象关联的有界区域和从数据管道接收的头部姿态数据来确定检测对象的位置。一旦确定了所检测的对象的3D位置, 3D位置连同诸如对象ID和时间戳的其他对象信息就一起被存储在存储器324中。存储器324可以是执行3D关联块302、对象检测块304和/或3D头部姿态块312的处理器本地的临时存储器位置或永久存储位置, 可以是诸如服务器的远程存储位置, 或其它存储位置。存储器324还可以是与存储器310或存储器322相同的存储器, 或者是单独的存储器。

[0074] 存储在存储器324中的对象信息被由处理器执行的呈现块326使用, 以将图形和虚拟对象在3D空间中附着于所检测的对象。呈现块326从存储器324检索对象信息、从存储器328检索3D模型、并从3D头部姿态块312检索头部姿态数据。可根据所检测的对象的类别来检索3D模型。呈现块326基于所检测的对象的确定比例缩放3D模型, 并基于所检测的对象的确定3D位置为每个3D模型确定包括深度的3D布局。存储器328可以是执行呈现块326、对象检测块304、3D头部姿态块312和/或3D关联块302的处理器本地的临时存储器位置或永久存储位置, 可以是诸如服务器的远程存储位置, 或其它存储位置。存储器328还可以是与存储器310、存储器322或存储器324相同的存储器, 或单独的存储器。在一些实施例中, 3D关联块302可具有到对象检测管道和SLAM管道的快速通信/数据信道, 诸如通过访问两个管道之间的共享存储器。呈现块326向显示处理单元 (DPU) 或显示处理器330提供呈现信息, 以在显示器332上显示图像, 其中, 该图像包括附着于检测对象的图形和虚拟对象。

[0075] 尽管图3示出了增强现实对象检测和3D关联架构300的一个示例, 但是可对图3进行各种改变。例如, 架构300可包括任何适当排列中的任何数量的每种组件。应当理解, 由架构300的各个块实施的功能可由单个处理器实施, 或者分布在两个或更多个处理器之间, 或者在同一电子装置内或者布置在单独的电子装置内。还应当理解, RGB摄像机306、单色摄像

机314和IMU 318可以是相对于图1描述的传感器180中的一个。所述功能可在至少一个处理器、图形处理单元 (GPU) 或其它组件中实现。通常,计算架构具有多种配置,并且图3不会将本公开的范围限制于任何特定配置。

[0076] 图4示出了根据本公开的实施例的示例对象检测和3D关联过程400的流程图。虽然该流程图描绘了一系列的连续的步骤,但是除非明确说明,否则不应该从关于实施的特定顺序的该序列中得出推断,步骤或部分步骤连续地而不是同时地或以重叠的方式的实施,或者在不发生介于中间或中间步骤的情况下专门描绘的步骤的实施。图4中描绘的过程可由图1中的电子装置101实施,并可由处理器120执行。

[0077] 在框402处,处理器控制一个或多个摄像机来拍摄环境的图像。在框404处,处理器在所拍摄的图像中检测对象。在框406处,处理器限定了围绕所检测的对象的有界区域。该有界区域包括计算以包围所检测的对象的 (x, y) 位置和尺寸 (w, h) 。在框408处,处理器生成头部姿态数据。头部姿态数据可以由IMU 318提供给处理器的头部或摄像机姿态,并可包括诸如3D点云的其它信息、具有2D特征点的图像、时间戳或其它信息。在框410处,依据本公开的各种实施例,处理器使用有界区域和头部姿态数据来确定所检测的对象在与环境关联的3D空间中的位置。

[0078] 尽管图4示出了示例过程,但是可对图4进行各种改变。例如,虽然被示为一系列步骤,但是在每个图中的各个步骤可重叠、并行发生、以不同的顺序发生,或多次发生。

[0079] 图5示出了根据本公开的实施例的用于检测对象的3D关联的3D特征点重投影过程的图解示图。该过程可由电子装置101实施,并可由处理器120执行。重投影过程包括通过诸如RGB摄像机306的传感器从特定的头部姿态 (P_{trgb}) 拍摄RGB帧502。处理器诸如经由对象检测块304检测RGB帧502内的对象504,并限定围绕所检测的对象504的有界区域506。有界区域506在RGB帧502内具有 (x, y) 坐标,并且处理器根据确定的宽度 (w) 和高度 (h) 确定有界区域506的大小以包围所检测的对象504。所检测的对象和有界区域信息(诸如,对象ID、对象类、时间戳、位置和大小)可存储在存储器中(诸如,在存储器310中),以在3D关联期间使用。使用SLAM,处理器基于一个或多个参数(诸如,从诸如IMU 318的IMU接收的头部姿态数据)并通过根据所生成的图像数据(例如来自单色摄像机314的数据)来确定环境的多个3D特征点,来生成3D点云508(诸如经由3D头部姿态块312)。SLAM数据(诸如头部姿态、3D点云、具有2D特征点的图像、或其它信息)可存储在存储器中(诸如,在存储器322中),以便在3D关联期间使用。

[0080] 应当理解,可使用各种SLAM算法来生成环境的3D点云,并且本公开不限于任何特定的SLAM算法。依据本公开的实施例,可使用拍摄RGB帧502的相同摄像机或者使用单独的摄像机,结合来自IMU的运动和定向数据来创建3D点云508。3D点云508的多个特征点与环境中的检测特征(对象、墙等)相对应,且提供环境内的特征的映射。

[0081] 重投影过程还包括:诸如经由3D关联块302,处理器通过所拍摄的RGB帧502和有界区域506对3D点云进行映射或重投影,以使得3D特征点位于姿态 P_{trgb} 的RGB摄像机的视场(FOV) 510内。用于RGB帧502的RGB摄像机参数512可用于提供FOV 510。处理器将包括所检测的对象504的有界区域506映射到FOV 510。一旦有界区域506在FOV 510内被映射,且3D点云508被重投影到FOV 510中,则有界区域506将包括来自通过有界区域506投影的3D点云508的特征点的子集,如图5所示。在一些实施例中,处理器可具体地生成特征点并将特征点重

投影到有界区域506中。例如,处理器可使用稀疏图而不是密集图以生成3D点云508。稀疏图可减少用以生成3D点云508的计算,但是可能丢失有界区域506中的所检测的对象504上的特征点。如果使用稀疏图来生成3D点云508,则由处理器执行的3D关联块302可请求也可由同一处理器执行的SLAM块312,以在所检测的有界区域506内具体地生成3D特征点,并将有界区域506的3D特征点保持在3D点云中。

[0082] 一旦3D特征点被重投影到FOV 510中,处理器就确定有界区域506内的哪些3D特征点位于所检测的对象504的表面上。处理器从有界区域506移除不位于所检测的对象504的表面上3D特征点。处理器根据所检测的对象504的表面上剩余的3D特征点确定所检测的对象504在3D空间514中的包括(x,y,z)坐标的3D位置。3D位置与诸如对象ID和时间戳的其它对象信息一起存储在存储器中(诸如在存储器324中),以由增强现实标记应用使用。在增强现实标记应用的操作期间,处理器可使用所检测的对象504的确定3D位置以将所检测的对象504放置在呈现块326处的3D空间514中,并用于在显示器332上显示。

[0083] 图6示出了根据本发明的实施例的3D特征点重投影过程的另一说明性示例的图600。示出的FOV 510包括有界区域506。处理器将3D特征点602从3D点云508重投影到有界区域506中。处理器确定有界区域中的哪些3D特征点602没有被布置在所检测的对象504的表面上,并移除那些特征点。然后,处理器可根据布置在所检测的对象504的一个或多个表面上的3D特征点602来确定3D空间514内的所检测的对象504的位置。

[0084] 图7示出了根据本公开的实施例的示例3D特征点重投影过程700的流程图。虽然该流程图描绘了一系列连续的步骤,但是除非明确说明,否则不应该从关于实施的特定顺序的该序列中得出推断,步骤或部分步骤连续地而不是同时地或以重叠的方式的实施,,或者在不发生介于中间或中间步骤的情况下专门描绘的步骤的实施。图7中描绘的过程可由图1中的电子装置101实施,并可由处理器120执行。

[0085] 在框702处,处理器使用一个或多个摄像机拍摄图像,并限定图像中围绕检测对象的有界区域。该过程还可将头部姿态数据与在拍摄图像时的头部姿态相对应的图像关联。在框704处,处理器生成3D点云。依据本公开的实施例,可根据由处理器从惯性测量单元接收的一个或多个拍摄的图像和数据生成3D点云。在框706处,处理器通过在框702中拍摄的图像的视场来对3D点的3D特征点进行重投影。在决策框708处,处理器确定在框704处是否使用稀疏图生成3D点云。如果不是,则过程700移至决策框710。

[0086] 在决策框710,处理器确定是否存在投影到有界区域中、不在所检测的对象的表面上的特征点。如果是,则过程700移至框712。在框712处,处理器移除有界区域内不位于所检测的对象的表面上的3D特征点。然后,该过程移至框714。如果在决策框710处处理器确定在有界区域中没有不位于所检测的对象的表面上的特征点,则过程从决策框710移至框714。在框714处,处理器根据布置在所检测的对象上的3D特征点存储对象在3D空间中的位置。根据框714,可发生电子装置101的其它操作,诸如,呈现和显示与所检测的对象的3D位置相邻的虚拟对象。

[0087] 如果在决策框708处处理器确定在框704处使用稀疏图生成3D点云,则过程从决策框708移至决策框716。在决策框716处,处理器确定SLAM是否可用。在某些情境下,诸如,如果由于在电子装置101上运行的其它进程而存在有限的可用系统资源,则SLAM可离线。在一些情境下,SLAM可以是在线的,但是或者在对象检测管道过远之前或者之后太远运行。如果

在决策框716处处理器确定SLAM管道不可用,则过程移至框718。在框718处,处理器使用针对所检测的对象的分配的对象类的预先配置的设定,以将对象放置在3D空间中。例如,处理器可在存储器中存储对象类的默认比例,并使用该默认比例来确定将虚拟对象放置在所拍摄的图像中的距离或深度。处理器还可使用对象类的预先配置的距离。在这种情况下,处理器根据所检测的对象的2D(x,y)坐标和预先配置的距离将所检测的对象放置在3D空间中。

[0088] 如果在决策框716处处理器确定SLAM是可用的,则过程700移至框720。在框720处,处理器在有界区域内生成3D特征点的密集图,并将用于有界区域的该密集图保持3D点云中。创建有界区域的单独的密集图提供了在不使用系统资源创建整个环境的密集图的情况下,在有界区域中的对象的表面上生成足够的特征点。过程700从框720移至框710,在框710,处理器确定来自在框720处生成的密集图的任何特征点是否不位于所检测的对象的表面上,在框712处移除这些特征点,并在框714处存储所检测的对象的3D位置。

[0089] 由SLAM生成的3D特征点提供关于3D环境的数据,包括对象的深度和比例以及环境的特征。通过将3D特征点重投影到所检测的对象的表面上,可确定所检测的对象在环境中的3D位置,包括所检测的对象的深度和比例。这允许确定所检测的对象的性质。例如,通过确定所检测的对象的深度和比例,可将玩具车与实际的车区分开,然后可通过相对于图2描述的增强现实标记应用对所检测的对象进行正确地分类。对象的深度也可用于向用户提供信息(诸如从用户到对象的距离)。此外,图像中所检测的对象的深度和比例可用于将虚拟对象准确地附着于图像中所检测的对象,以使得应用将虚拟对象放置在图像中所检测的对象附近,并如应用所限定地缩放该虚拟对象(该虚拟对象可有意地具有与所检测的对象不同的比例),并使得即使当用户在不同的姿态之间移动摄像机时,虚拟对象也保持在图像中所检测的对象附近。

[0090] 使用用于3D空间理解的SLAM的同时使用2D图像实施对象检测,比尝试完全3D对象检测需要更少的计算和系统资源,这提供了更快的对象检测和对象标记。这对于增强现实应用是重要的,在增强现实应用中,用户利用摄像机来提供场景或环境的连续视图,因为用户期望应用在用户装置上显示关于环境中的对象的准确且连续更新的信息。

[0091] 尽管图7示出了示例过程,但是可对图7进行各种改变。例如,虽然被示为一系列步骤,但是在每个图中的各个步骤可重叠、并行发生、以不同的顺序发生、或多次发生。

[0092] 在一些场景中,电子装置101的系统资源,或实施本公开中所描述的一个或多个操作的其它装置,可能被由电子装置101正在执行的其它过程限制。在此类场景中,可用的系统资源可能防止SLAM生成3D点云。在这种情况下,可由SLAM管道处理的所拍摄的图像的有界区域,以在所拍摄的帧上提供2D特征点。如果可用的系统资源不允许生成2D特征点,则对象检测管道可实施在单独拍摄的图像的有界区域之间的像素匹配,以经由像素匹配提供特征点。一旦提供了特征点(诸如通过生成2D特征点或者通过实施像素匹配),电子装置101就可在匹配特征点之间实施三角测量,以将所检测的对象与3D空间关联。因此,电子装置101可监控当前可用的系统资源,并还可监控3D点云的质量,以及根据当前可用的系统资源实时地切换3D关联方法(重投影、2D特征点三角测量、像素匹配三角测量)。

[0093] 图8示出了根据本公开的实施例的示例特征点三角测量过程的图800。该过程可由电子装置101实施,并可由处理器120执行。摄像机802在时间T和第一姿态处拍摄包括在环境中的对象806的第一帧804。摄像机802或另一个摄像机在时间T+t和第二姿态处拍摄对象

806的第二帧808。处理器处理第一帧804和第二帧808以及第一姿态和第二姿态的姿态数据,以检测每个图像中的对象806,并在对象806的表面上生成特征点。如果可用的系统资源允许SLAM管道在图像中提供2D特征点,则特征点可以是2D特征点。如果SLAM管道不提供2D特征点,并由于可用的系统资源仅提供姿态,则对象检测管道可在第一帧804与第二帧808的有界区域之间实施像素匹配,以确定相同的像素以及从而相同的对象特征位于第一帧804和第二帧808内的位置。例如,处理器可确定包括相同RGB值和可能的其它像素值的单独的帧中的像素是匹配的像素。

[0094] 处理器使用由SLAM管道提供的姿态信息对第一帧804和第二帧808内的特征点进行三角测量。如图8所示,当分别拍摄第一帧804和第二帧808时,可使用姿态数据对立方体形对象的角处的特征点进行三角测量,以确定摄像机802在时间T和T+t的定向。处理器对来自每个帧804和帧808的对象806的特征点进行三角测量,以确定相同特征点在3D空间内的位置。可对多个特征点(2D特征点或匹配像素)实施三角测量,以将包括(x,y,z)坐标的三角测量的对象810放置在3D空间中。处理器可对多个特征点进行三角测量,以确定对象的3D位置和比例。例如,处理器可对在图8所示的立方体形对象的每个角处的特征点进行三角测量,这可提供足够的信息以在3D空间中准确地呈现整个形状。应当理解,虽然图8示出了两个姿态和两个拍摄的帧,但是可使用多于两个的姿态和帧来提供多个三角测量点。

[0095] 图9示出了根据本公开的实施例的示例检测对象三角测量过程900的流程图。虽然该流程图描绘了一系列连续的步骤,但是除非明确说明,否则不应该从关于实施的特定顺序的该序列中得出推断,步骤或部分步骤连续地而不是同时地或以重叠的方式的实施,或者在不发生介于中间或中间步骤的情况下专门描绘的步骤的实施。图9中所描绘的过程可由图1中的电子装置101实施,并可由处理器120执行。

[0096] 在框902处,处理器确定由于可用的系统资源而未提供3D特征点。在框904处,由电子装置拍摄多个帧,其中每个帧与姿态关联。处理器对多个帧中的每个帧实施对象检测,并在多个帧中的每个帧中限定围绕至少一个检测对象的有界区域。在决策框908处,处理器确定可用的系统资源是否允许在所拍摄的图像中生成2D特征点。如果是,则过程900移至框910。

[0097] 在框910处,处理器在多个帧中的每个帧的至少有界区域内生成2D特征点。在框912处,处理器将有界区域内的一个或多个2D特征点设定为三角测量参数。在框914处,处理器使用设定的三角测量参数在多个帧中的两个或更多个帧中对所检测的对象的3D位置进行三角测量。在处理器在框912中将三角测量参数设定为所生成的2D特征点的情况下,处理器使用两个或更多个图像的头部姿态数据对一个或多个特征点进行三角测量。

[0098] 如果在决策框908处处理器确定可用的系统资源不允许产生2D特征点,则过程900移至框916。在框916处,处理器在位于多个帧中的两个或更多个帧中的特定对象的有界区域内的像素之间实施像素匹配。在框918处,处理器将一个或多个匹配的像素设定为将在三角测量中使用的三角测量参数或特征点。过程900从框918移至框914,其中,处理器使用设定的三角测量参数(在这种情况下是匹配的像素)对所检测的对象的位置进行三角测量,以确定所检测的对象的3D位置和比例。

[0099] 如本文所述,使用2D特征点和像素匹配的三角测量提供了所检测的对象的深度和比例。这允许确定所检测的对象的性质。例如,通过确定所检测的对象的比例,可将玩具车

与实际的车区分开,然后可通过相对于图2描述的增强现实标记应用对所检测的对象进行正确的分类。对象的深度也可用于向用户提供信息(诸如从用户到对象的距离)。此外,图像中所检测的对象的深度和比例可用于将虚拟对象准确地附着于图像中所检测的对象,以使得应用将虚拟对象放置在图像中所检测的对象附近,并如应用所限定地缩放虚拟对象(虚拟对象可有意地具有与所检测的对象不同的比例),并使得即使当用户在不同的姿态之间移动摄像机时,虚拟对象也保持在图像中所检测的对象附近。

[0100] 本文所述的三角测量过程允许增强现实标记应用即使在较少的系统资源可用时也继续有效地运行。这对于增强现实应用(在增强现实应用中,用户利用摄像机提供场景或环境的连续视图)是重要的,因为用户期望应用为用户装置上显示关于环境中的对象的准确且连续更新的信息。

[0101] 尽管图9示出了示例过程,但是可对图9进行各种改变。例如,虽然被示为一系列步骤,但是在每个图中的各个步骤可重叠、并行发生、以不同的顺序发生、或多次发生。

[0102] 本申请中的任何描述都不应被理解为暗示任何特定元素、步骤或功能都是必须包括在权利要求范围内的必要元素。专利可授权主题的范围仅由权利要求限定。此外,权利要求中的任何一个权利要求都不旨在调用35U.S.C.112(f),除非确切的词语“用于…的装置”后面跟有分词。

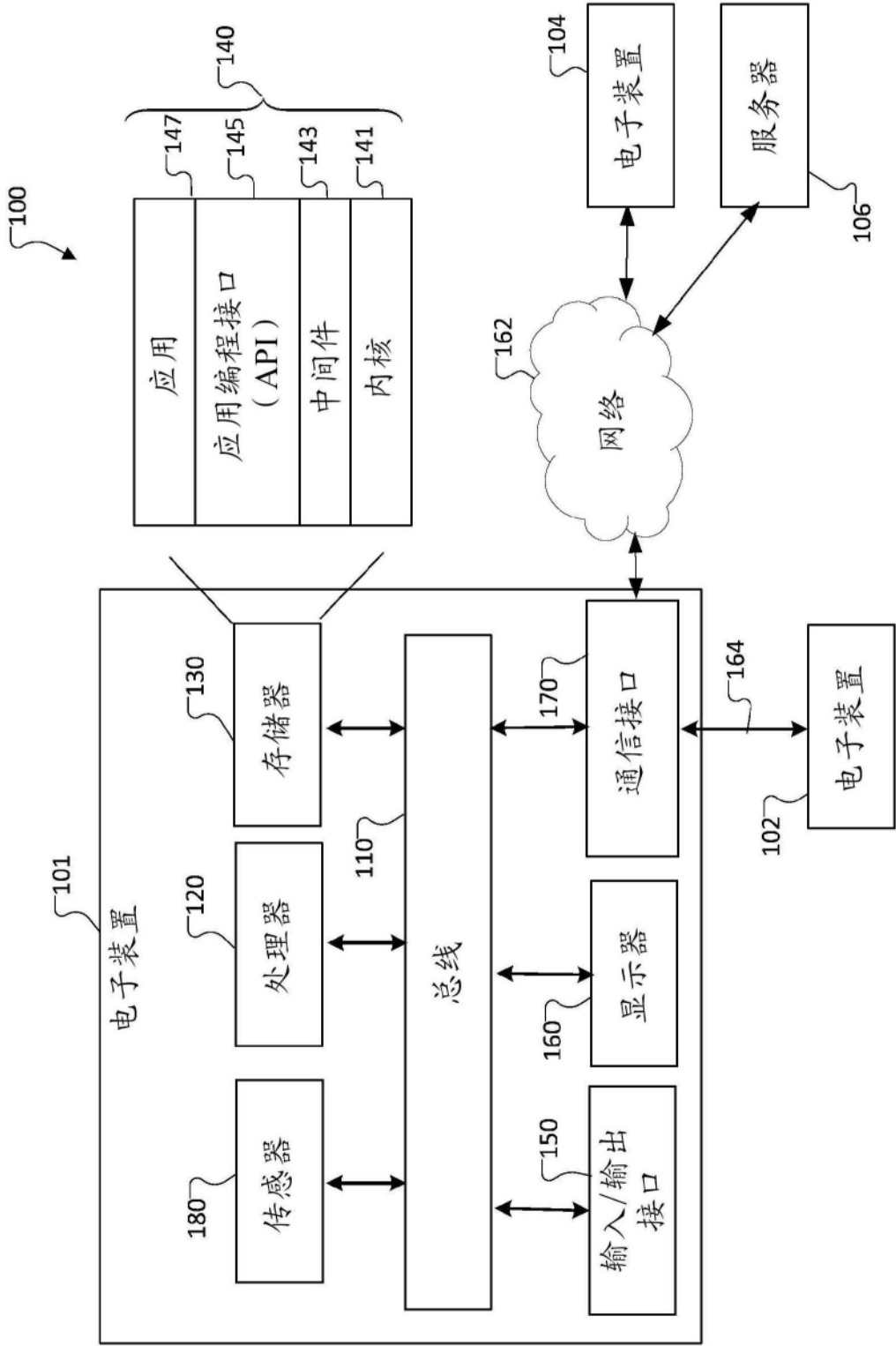


图1

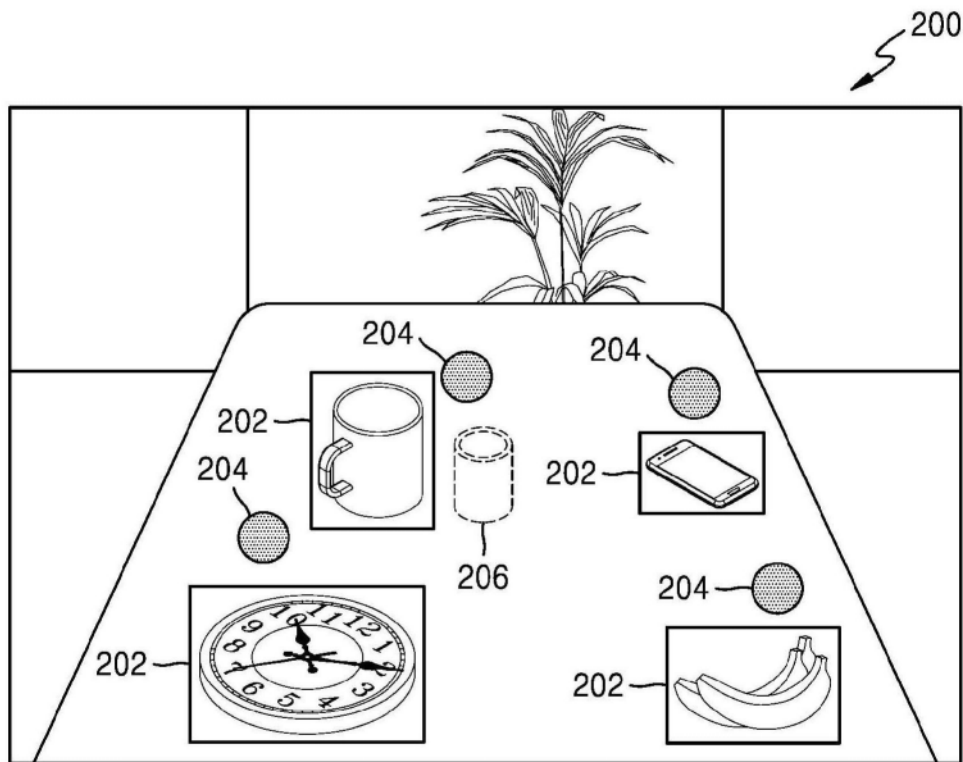


图2

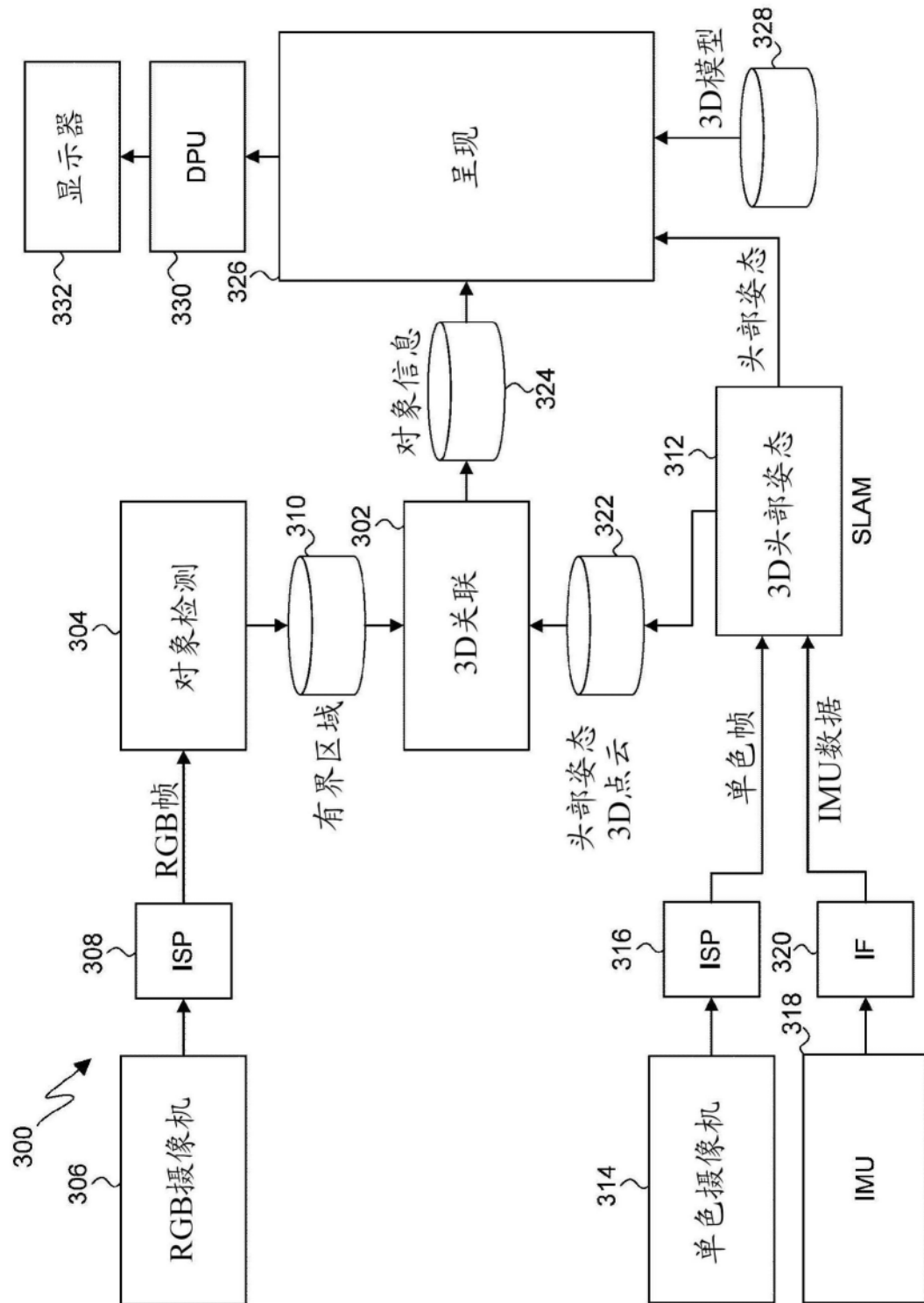


图3

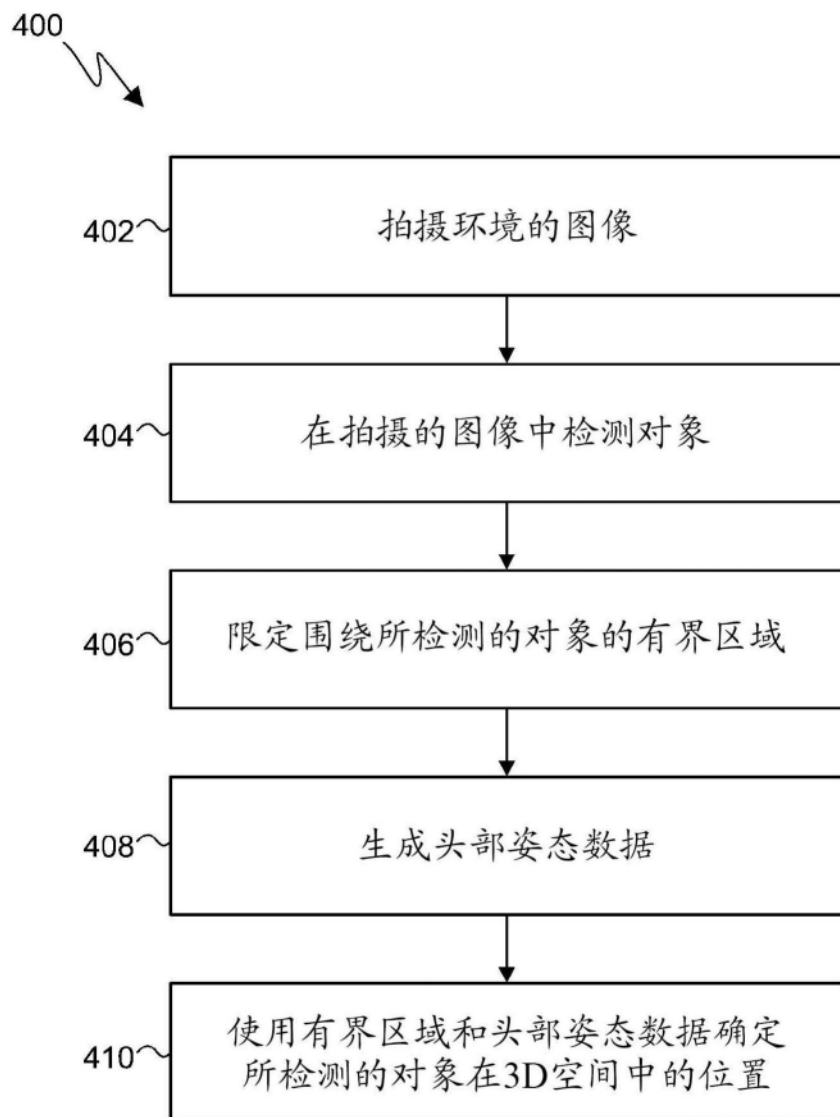


图4

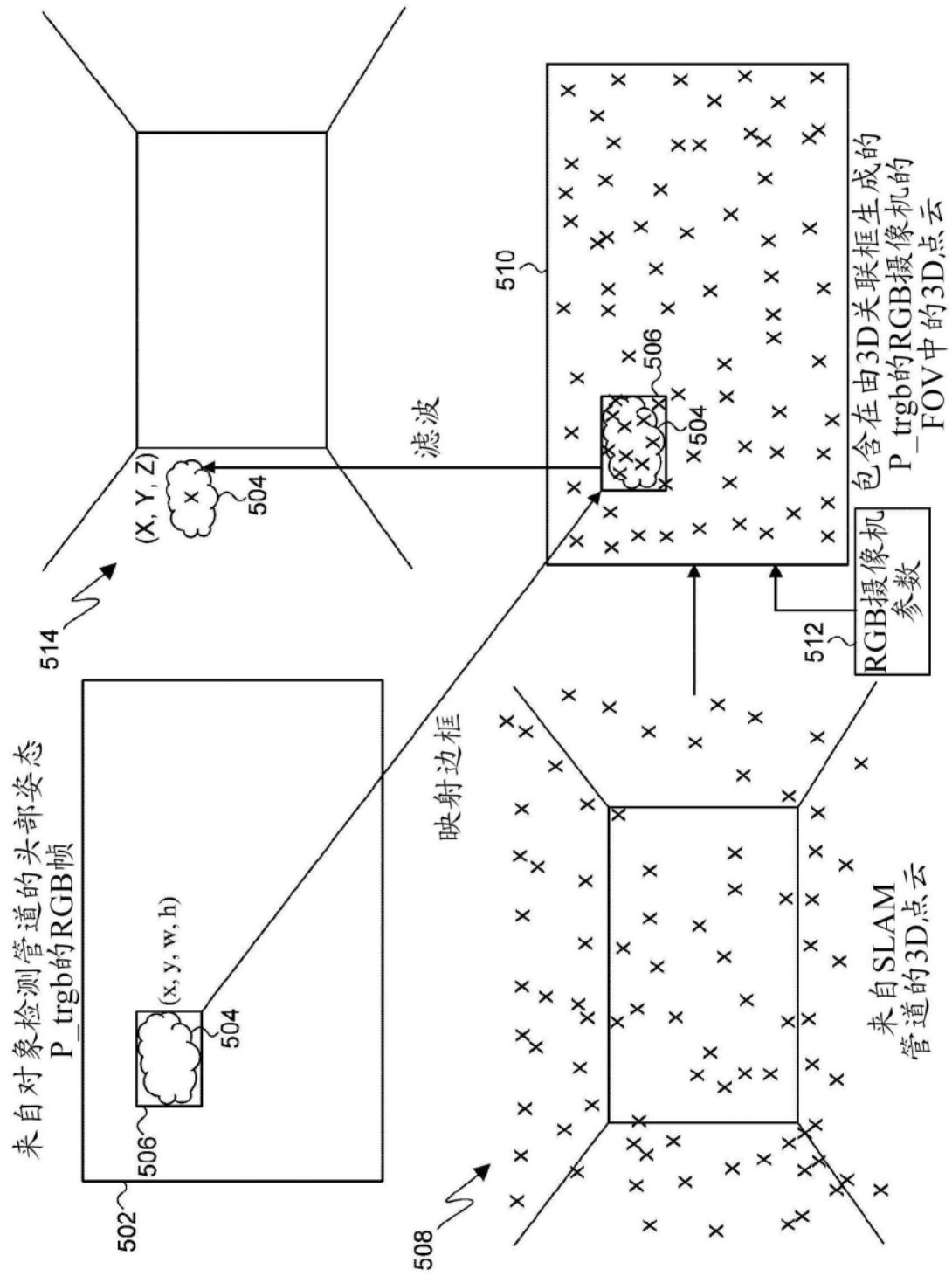


图5

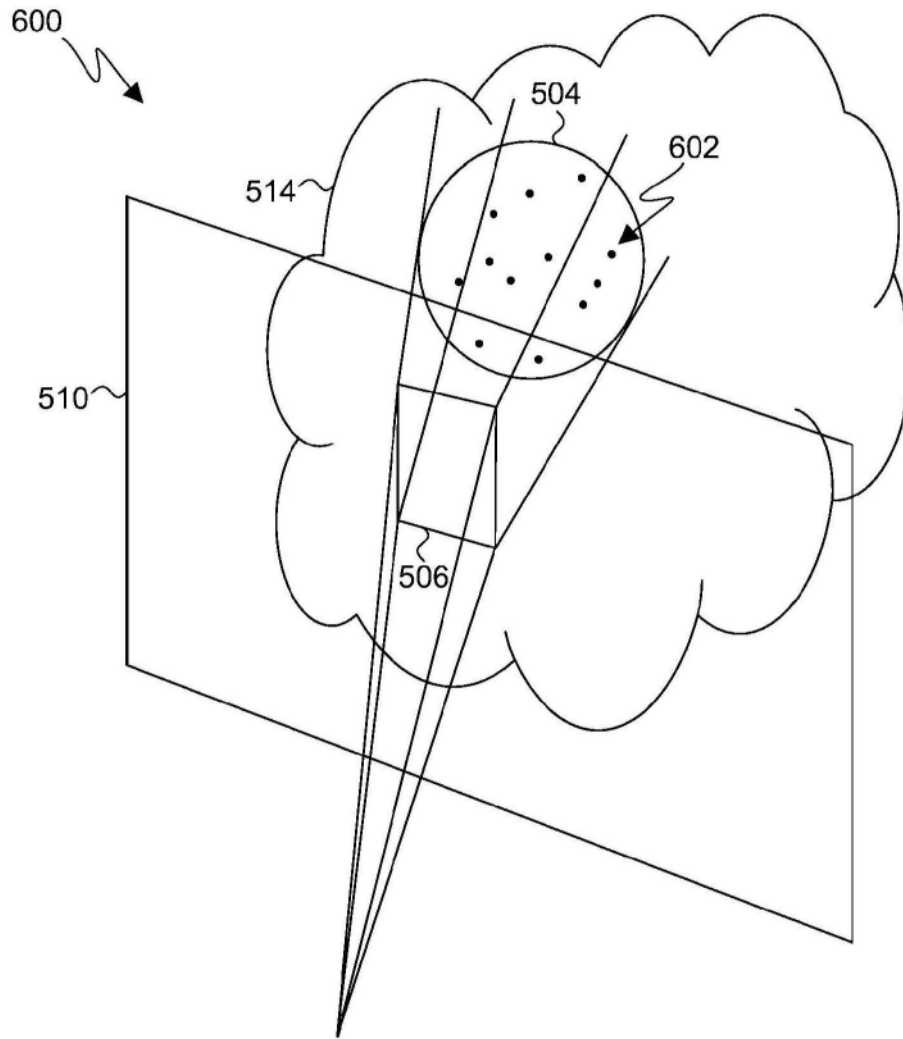


图6

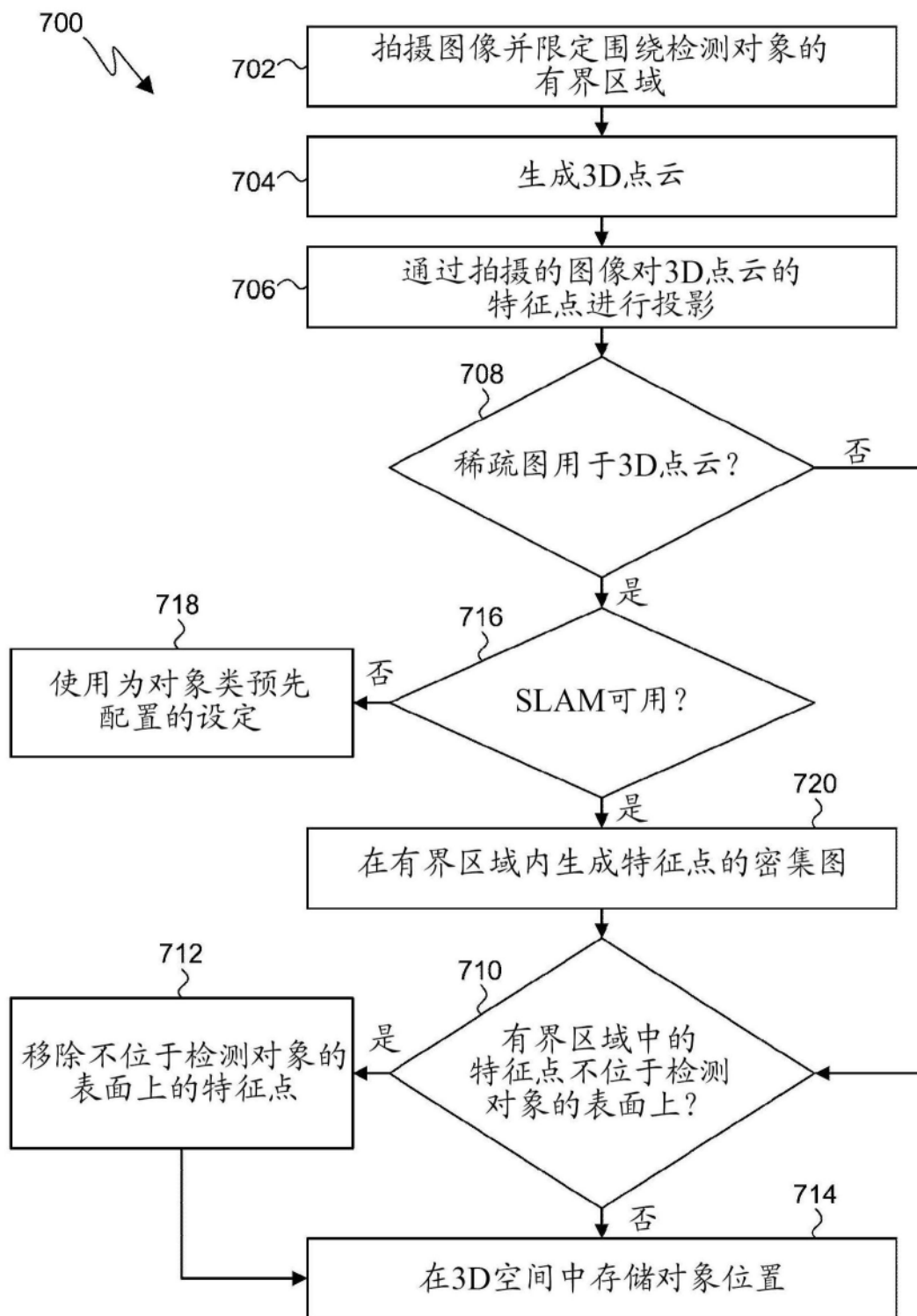


图7

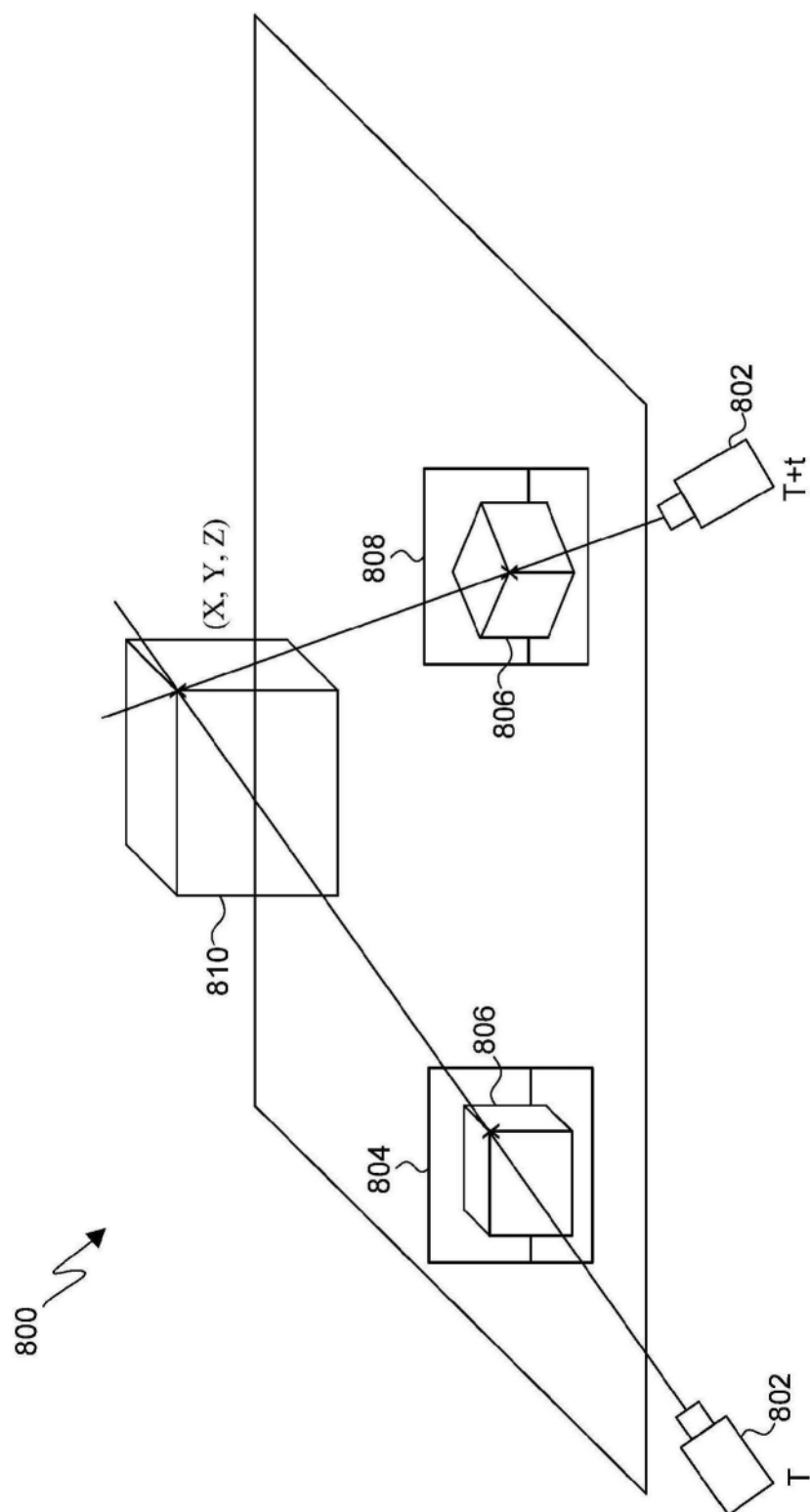


图8

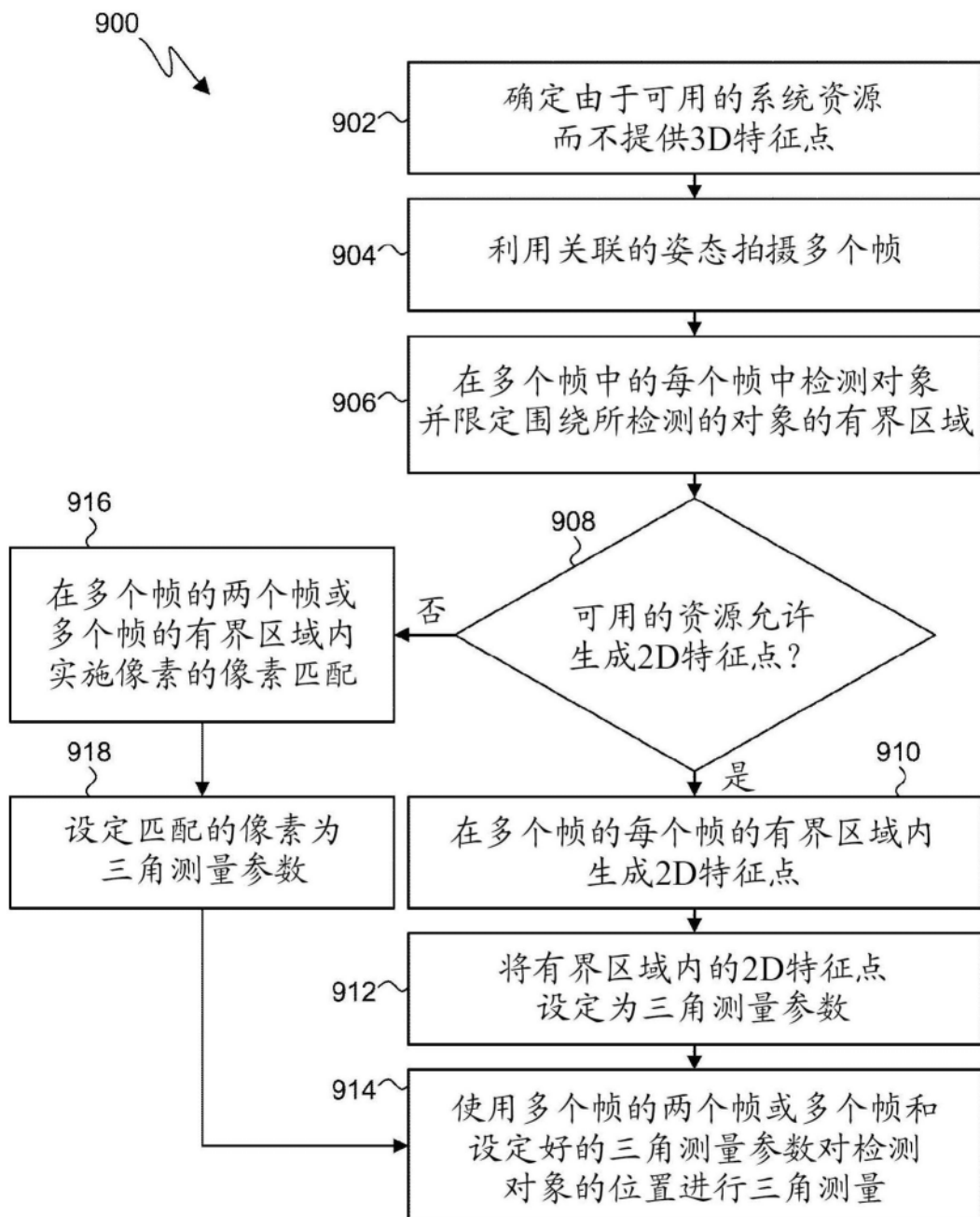


图9