



(12)发明专利申请

(10)申请公布号 CN 108734120 A

(43)申请公布日 2018.11.02

(21)申请号 201810464372.8

(22)申请日 2018.05.15

(71)申请人 百度在线网络技术(北京)有限公司

地址 100080 北京市海淀区上地十街10号

(72)发明人 程新景 杨睿刚

(74)专利代理机构 北京市金杜律师事务所

11256

代理人 李辉 杜波

(51)Int.Cl.

G06K 9/00(2006.01)

G06T 7/10(2017.01)

G06T 7/50(2017.01)

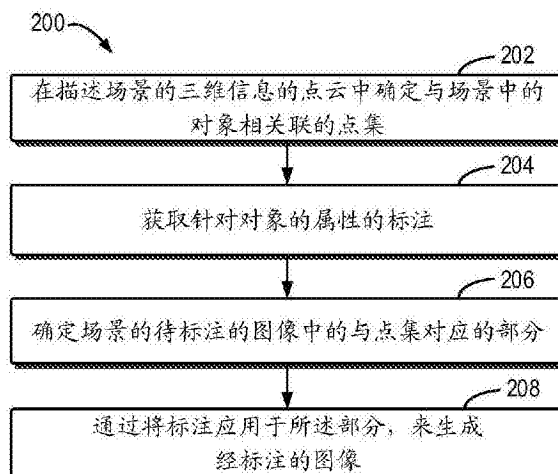
权利要求书3页 说明书9页 附图4页

(54)发明名称

标注图像的方法、装置、设备和计算机可读存储介质

(57)摘要

本公开的实施例提供了用于标注图像的方法、装置、设备和计算机可读存储介质。用于标注图像的方法包括在描述场景的三维信息的点云中确定与场景中的对象相关联的点集。该方法还包括获取针对对象的属性的标注。该方法还包括确定场景的待标注的图像中的与点集对应的部分。该方法还包括通过将标注应用于该部分，来生成经标注的图像。以此方式，可以提高图像的标注效率，从而获得大规模的经标注的图像。



1. 一种用于标注图像的方法,包括:
在描述场景的三维信息的点云中确定与所述场景中的对象相关联的点集;
获取针对所述对象的属性的标注;
确定所述场景的待标注的图像中的与所述点集对应的部分;以及
通过将所述标注应用于所述部分,来生成经标注的图像。
2. 根据权利要求1所述的方法,其中在所述点云中确定与所述场景中的对象相关联的点集包括:
从所述点云中去除与预定对象相关联的预定对象点集;
在去除了所述预定对象点集的所述点云中分割出与所述场景中的多个对象相关联的多个点集;以及
从所述多个点集中选择与所述对象相关联的所述点集。
3. 根据权利要求1所述的方法,其中在所述点云中确定与所述场景中的对象相关联的点集包括:
基于所述点云中的点之间的相似性,将所述点云划分成多个集合;以及
基于所述多个集合之间的凹凸关系,对所述多个集合进行聚类,以获取与所述场景中的多个对象相关联的多个点集;以及
从所述多个点集中选择与所述对象相关联的所述点集。
4. 根据权利要求1所述的方法,其中获取针对所述对象的属性的标注包括:
向用户呈现与所述对象相关联的所述点集;以及
接收所述用户输入的针对所述对象的属性的标注。
5. 根据权利要求4所述的方法,还包括:
接收所述用户对所述点集的调整;以及
基于所述调整,更新所述点集。
6. 根据权利要求1所述的方法,其中获取针对所述对象的属性的标注包括:
利用属性预测模型来预测所述对象的属性,所述属性预测模型是基于对象已被标注的点云来训练的;以及
将预测的所述属性确定为针对所述对象的属性的标注。
7. 根据权利要求1所述的方法,其中获取针对所述对象的属性的标注包括:
利用属性预测模型来预测所述对象的属性,所述属性预测模型是基于对象已被标注的点云来训练的;
向用户提供预测的所述属性;
响应于接收到用户对预测的所述属性的确认,将预测的所述属性确定为针对所述对象的属性的标注;以及
响应于接收到所述用户对预测的所述属性的修改,将经修改的属性确定为针对所述对象的属性的标注。
8. 根据权利要求6或7所述的方法,还包括:
利用所述点云和针对所述对象的属性的标注来更新所述属性预测模型。
9. 根据权利要求1所述的方法,其中确定所述场景的待标注的图像中的与所述点集对应的部分包括:

获取相机在拍摄所述图像时的参数;以及
基于所述参数确定与所述点集对应的所述部分。

10. 根据权利要求1所述的方法,其中所述对象的属性包括以下至少一项:所述对象的类别、名称和深度信息。

11. 根据权利要求1所述的方法,其中获取针对所述对象的属性的标注包括:

确定所述对象的深度信息,所述深度信息指示所述对象与拍摄所述图像时相机所处的空间位置之间的距离;以及

将所述深度信息确定为针对所述对象的属性的标注。

12. 根据权利要求11所述的方法,还包括:

通过将所述距离与所述部分相关联,生成与所述图像对应的深度图像。

13. 一种用于标注图像的装置,包括:

点集确定模块,被配置为在描述场景的三维信息的点云中确定与所述场景中的对象相关联的点集;

标注获取模块,被配置为获取针对所述对象的属性的标注;

对应部分确定模块,被配置为确定所述场景的待标注的图像中的与所述点集对应的部分;以及

标注图像生成模块,被配置为通过将所述标注应用于所述部分,来生成经标注的图像。

14. 根据权利要求13所述的装置,其中所述点集确定模块包括:

去除模块,被配置为从所述点云中去除与预定对象相关联的预定对象点集;

分割模块,被配置为在去除了所述预定对象点集的所述点云中分割出与所述场景中的多个对象相关联的多个点集;以及

点集选择模块,被配置为从所述多个点集中选择与所述对象相关联的所述点集。

15. 根据权利要求13所述的装置,其中所述点集确定模块包括:

划分模块,被配置为基于所述点云中的点之间的相似性,将所述点云划分成多个集合;

聚类模块,被配置为基于所述多个集合之间的凹凸关系,对所述多个集合进行聚类,以获取与所述场景中的多个对象相关联的多个点集;以及

点集选择模块,被配置为从所述多个点集中选择与所述对象相关联的所述点集。

16. 根据权利要求13所述的装置,其中所述标注获取模块包括:

呈现模块,被配置为向用户呈现与所述对象相关联的所述点集;以及

标注接收模块,被配置为接收所述用户输入的针对所述对象的属性的标注。

17. 根据权利要求16所述的装置,还包括:

调整接收模块,被配置为接收所述用户对所述点集的调整;以及

更新模块,被配置为基于所述调整,更新所述点集。

18. 根据权利要求13所述的装置,其中所述标注获取模块包括:

预测模块,被配置为利用属性预测模型来预测所述对象的属性,所述属性预测模型是基于对象已被标注的点云来训练的;以及

标注确定模块,被配置为将预测的所述属性确定为针对所述对象的属性的标注。

19. 根据权利要求13所述的装置,其中所述标注获取模块包括:

预测模块,被配置为利用属性预测模型来预测所述对象的属性,所述属性预测模型是

基于对象已被标注的点云来训练的；

提供模块，被配置为向用户提供预测的所述属性；以及

标注确定模块，被配置为：

响应于接收到用户对预测的所述属性的确认，将预测的所述属性确定为针对所述对象的属性的标注；以及

响应于接收到所述用户对预测的所述属性的修改，将经修改的属性确定为针对所述对象的属性的标注。

20. 根据权利要求18或19所述的装置，还包括：

预测模型更新模块，被配置为利用所述点云和针对所述对象的属性的标注来更新所述属性预测模型。

21. 根据权利要求13所述的装置，其中所述对应部分确定模块包括：

相机参数获取模块，被配置为获取相机在拍摄所述图像时的参数；以及

确定模块，被配置为基于所述参数确定与所述点集对应的所述部分。

22. 根据权利要求13所述的装置，其中所述对象的属性包括以下至少一项：所述对象的类别、名称和深度信息。

23. 根据权利要求13所述的装置，其中所述标注获取模块包括：

深度信息确定模块，被配置为确定所述对象的深度信息，所述深度信息指示所述对象与拍摄所述图像时相机所处的空间位置之间的距离；以及

标注确定模块，被配置为将所述深度信息确定为针对所述对象的属性的标注。

24. 根据权利要求23所述的装置，还包括：

深度图像生成模块，被配置为通过将所述距离与所述部分相关联，生成与所述图像对应的深度图像。

25. 一种电子设备，所述电子设备包括：

一个或多个处理器；以及

存储器，用于存储一个或多个程序，当所述一个或多个程序被所述一个或多个处理器执行时，使得所述电子设备实现根据权利要求1-12中任一项所述的方法。

26. 一种计算机可读存储介质，其上存储有计算机程序，所述程序被处理器执行时实现根据权利要求1-12中任一项所述的方法。

标注图像的方法、装置、设备和计算机可读存储介质

技术领域

[0001] 本公开的实施例一般涉及图像处理领域,并且更具体地,涉及用于标注图像的方法和装置、设备和计算机可读存储介质。

背景技术

[0002] 近些年,随着人工智能的迅速发展,大规模的训练数据集对于人工智能算法的精度起着越来越重要的作用。如何高效地获取高质量的经标注的训练数据是推进人工智能算法向前发展的先决条件。然而,在无人驾驶领域中,由于目前缺少对图像进行高效标注的方案,所以还不存在大规模的经标注的室外图像数据集,从而无法为无人驾驶及其相应算法提供良好的数据支持。

发明内容

[0003] 根据本公开的实施例,提供了一种基于点云来标注图像的方案。

[0004] 在本公开的第一方面,提供了一种用于标注图像的方法。该方法包括:在描述场景的三维信息的点云中确定与场景中的对象相关联的点集;获取针对对象的属性的标注;确定场景的待标注的图像中的与点集对应的部分;以及通过将标注应用于所述部分,来生成经标注的图像。

[0005] 在本公开的第二方面,提供了一种用于标注图像的装置。该装置包括:点集确定模块,被配置为在描述场景的三维信息的点云中确定与场景中的对象相关联的点集;标注获取模块,被配置为获取针对对象的属性的标注;对应部分确定模块,被配置为确定场景的待标注的图像中的与点集对应的部分;以及标注图像获取模块,被配置为通过将标注应用于所述部分,来生成经标注的图像。

[0006] 在本公开的第三方面,提供了一种电子设备。该电子设备包括:一个或多个处理器;以及存储器,用于存储一个或多个程序,当所述一个或多个程序被所述一个或多个处理器执行,使得电子设备实现根据本公开的第一方面的方法。

[0007] 在本公开的第四方面中,提供了一种计算机可读介质,其上存储有计算机程序,该程序被处理器执行时实现根据本公开的第一方面的方法。

[0008] 应当理解,发明内容部分中所描述的内容并非旨在限定本公开的实施例的关键或重要特征,亦非用于限制本公开的范围。本公开的其它特征将通过以下的描述变得容易理解。

附图说明

[0009] 结合附图并参考以下详细说明,本公开各实施例的上述和其他特征、优点及方面将变得更加明显。在附图中,相同或相似的附图标记表示相同或相似的元素,其中:

[0010] 图1示出了能够在其中实现本公开的实施例的示例性环境的示意图;

[0011] 图2示出了根据本公开的实施例的用于标注图像的方法的流程图;

[0012] 图3示出了根据本公开的实施例的用于获取针对对象的属性的标注的方法的流程图；

[0013] 图4示出了根据本公开的实施例的用于获取针对对象的属性的标注的方法的流程图；

[0014] 图5示出了根据本公开的实施例的用于生成深度图像的方法的流程图；

[0015] 图6示出了根据本公开的实施例的用于标注图像的装置的方框图；以及

[0016] 图7示出了能够实施本公开的实施例的电子设备的方框图。

具体实施方式

[0017] 下面将参照附图更详细地描述本公开的实施例。虽然附图中显示了本公开的某些实施例，然而应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为限于这里阐述的实施例，相反提供这些实施例是为了更加透彻和完整地理解本公开。应当理解的是，本公开的附图及实施例仅用于示例性作用，并非用于限制本公开的保护范围。

[0018] 如上文所提及的，在无人驾驶领域中，需要大规模的已标注的室外图像集来开发并优化相关算法。传统上，采用以下流程来对在室外采集的RGB图像进行语义标注：在RGB图像上进行物体分割，获取物体在RGB图像中的大致边界以及边界控制点；人工在RGB图像上用多边形框选出包括标注物体的边界，然后利用算法进行细化；通过人工拖动边界控制点来进一步修正物体边界，从而得到像素级别边界精度的物体，并将其标注成相应类别。

[0019] 然而，上述流程存在着诸多不足。例如，在RGB图像上进行物体分割，是基于相邻物体边界像素在RGB颜色空间中具有较大的差异并且同一物体像素在RGB空间差异较小的假设。但是，在实际情况中，这种假设几乎很难满足。例如，在一个穿着白衣服的人站在一堵白色的墙前面的场景中，利用现有的方法，很难将人与墙分割开，从而后期需要投入大量的人力进行修改。

[0020] 另一方面，通过人工选择物体边界，不仅增加了人工成本，同时一些室外形状和空间分布比较奇特的物体（例如树木的树叶）也给人工选择带来了很大的挑战。通过人工拖动边界控制点来修正物体边界，将进一步消耗大量人力成本。

[0021] 以上这些诸多不足，导致对一张室外RGB图像进行标注以获得像素级别的语义标注通常需要耗费很长的时间，从而限制了经标注的室外RGB图像集的规模。这也是在无人驾驶领域中为何缺少大规模的经标注的室外图像集的原因所在。

[0022] 针对上述问题，本公开的实施例提供了一种用于标注图像（例如，室外RGB图像）的方案。该方案在三维点云上进行物体（下文也称为对象）分割，并在点云中对物体的属性进行标注。利用点云中的点（下文也称为体素）与待标注的图像上的像素之间的对应关系，将经标注的点云投影到图像上，从而得到带标注信息的图像。

[0023] 由于与图像相比，对象在点云中相对比较孤立，所以在点云中分割对象能够实现更高的精度，从而后续仅需要花费很少的工作量即可完成对对象的边界的调整。同时，利用点云与同一场景的图像之间的对应关系，可以将标注好的点云投影到多张图像上，从而得到大规模的经标注的图像集。与现有技术相比，本公开的实施例的方案在标注成本和标注效率方面都具有显著的优势，从而能够为室外场景理解算法提供大规模的语义数据集。

[0024] 下面将结合图1到图7来具体描述本公开的实施例。

[0025] 图1示出了能够在其中实现本公开的实施例的示例性环境100的示意图。在环境100中,安装有激光雷达和相机(图1中未示出)的采集实体112(例如车辆112)运行在路面114上,以对与采集实体112所处的场景110相关的数据进行采集。

[0026] 在本公开的上下文中,术语“采集实体”是能够采集点云、数据和/或其他适当数据的实体,例如车辆、人或其他能够运动的设备。激光雷达可以是单线激光雷达、多线激光雷达、3D激光雷达等。相机可以是高精相机、全景相机、单目相机等。应当理解,上述示例仅仅是为了说明的目的,而无意限制本公开的实施例的范围。

[0027] 在采集实体112运动过程中,激光雷达采集描述采集实体112所处的场景110的三维信息的点云。如图1所示,场景110包括采集实体112当前所处的路面114、采集实体112两侧的树木116以及采集实体112前方的建筑物118。所采集的点云数据可以描述路面114、树木116以及建筑物118上各点的三维信息,例如空间坐标。

[0028] 同时,与激光雷达关联设置的相机可以采集同一场景110的图像。在本公开的实施例中,术语“图像”是指相机所拍摄的二维图像,例如RGB图像。相机通过调整成像参数(例如焦距、位置、角度等)可以采集到同一场景110的大量图像。根据相机成像原理,相机的成像参数可以用来确定点云中的点与图像中的哪个像素相对应,即可以用来确定点云中的点与图像中的像素之间的对应关系。

[0029] 采集实体112将运动过程中采集的点云、图像以及采集图像时相机的成像参数关联地上传到云存储设备120中。本领域技术人员将理解的是,也可以将点云、图像和成像参数存储在其他存储设备上,而限于存储到云存储设备120。例如,点云、图像和成像参数也可以存储在诸如硬盘之类的常规存储器中。

[0030] 计算设备130从云存储设备120下载描述场景110的三维信息的点云、场景110的多张图像以及相机在拍摄这些图像时的成像参数。计算设备130在点云中对描述场景110中的对象的点进行划分。例如,计算设备130可以将点云中的点划分成与建筑物118相关联的点集和与各树木116分别相关联的点集。由于在点云中,表示建筑物118的点与表示树木116的点具有一定的孤立性,所以基于点云完成的对象分割能够具有较高的精度。

[0031] 计算设备130将点云中与对象相关联的点集呈现给用户140,例如将与建筑物118相关联的点集呈现给用户140。用户140可以对所呈现的点集进行标注,即对与点集相关联的对象的属性进行标注,例如将所呈现的点集标注为“建筑物”。用户140也可以对所呈现的点集进行调整,而计算设备130可以基于该调整来更新与对象相关联的点集,以使得对象的分割更加精确。

[0032] 计算设备130接收用户140对对象的属性的标注以及对点集的调整。计算设备130可以基于用户140对点集的调整,来更新与对象相关联的点集。对于该场景110的多张图像中的每一张图像,计算设备130基于相机在拍摄该图像时的成像参数,在该图像中确定对应于点云中的与对象相关联的点集的部分,并将从用户140获取的对对象的属性的标注应用于这些部分。以这种方式,完成对该场景的多张图像的标注。

[0033] 在本公开的实施例中,在点云中实现高精度的对象分割,并从用户140获取对分割的对象的属性(例如,对象的类别、对象的名称)的标注,然后利用点云与图像的对应关系,将所获取的标注投影到图像中的对应部分。以这种方式,对一个场景中的点云进行标注之后,能够获得该场景的大量图像的标注,从而能够获得大规模的经标注的图像。

[0034] 计算设备130可以将经标注的图像上传到云存储设备120中进行存储,以形成经标注的图像数据集。执行机器学习算法(例如与无人驾驶相关的机器学习算法)的计算设备150可以从云存储设备120下载这种经标注图像数据集,以实现机器学习算法的优化。应当理解的是,虽然在图1中将计算设备130和计算设备150示出为是分离的,但是在其他实施例中,计算设备130和计算设备150可以被集成在一起。

[0035] 计算设备130还可以基于已经标注过的点云,来训练预测模型,来预测点云中的对象的属性并呈现给用户140,以方便用户140的标注。以这种方式,本公开的实施例能够充分利用已经标注过的先验知识,来进一步加快图像的标注过程。

[0036] 应当理解,图1所示的各个部件的数目、结构、连接关系和布局都是示例性的,而非限制性的,并且其中一些部件是可选的。本领域技术人员在本公开的范围内可以在数目、结构、连接关系和布局等方面进行调整。

[0037] 图2示出了根据本公开的实施例的用于标注图像的方法200的流程图。方法200可以由图1中所示的计算设备130来执行。如之前所描述的,为了建立大规模的经标注的室外图像数据集,计算设备130事先从云存储设备120下载描述场景110的三维信息的点云、同一场景110的大量图像(例如室外RGB图像)以及相机在拍摄这些图像时的成像参数。

[0038] 在框202,计算设备130在描述场景的三维信息的点云中确定与该场景中的对象相关联的点集。即,计算设备130在点云中实现场景中的对象的分割。场景的三维信息可以包括场景中的对象上的各个点的空间坐标。点云中的与对象相关联的点集可以是点云中描述该对象的点(以下也称为“体素”)的集合。在点云中,由于对象本身具有一定的孤立性,因此与图像相比,在点云中能够实现精度更高的对象分割。

[0039] 虽然在点云中,对象之间具有孤立性,但是对象几乎全部与某个预定对象有关联。如在图1所示的例子中,建筑物118、树木116都与地面114相关联。为此,在本公开的一些实施例中,为了在点云中实现更好的对象分割,计算设备130可以从点云中去除与预定对象相关联的预定对象点集。例如计算设备130可以对点云进行腐蚀-膨胀处理,来去除点云中与预定对象(例如地面114)相关联的预定对象点集,使得点云中的其他对象(例如建筑物118和树木116)彼此之间更加孤立,以有助于这些对象的分割。

[0040] 计算设备130可以在去除了预定对象点集之后的点云中分割出与场景中的多个对象相关联的多个点集。例如,以图1为例,计算设备130可以从去除了路面点集的点云中,分割出与建筑物118相关联的点集、与两颗树木116分别关联的点集。

[0041] 在本公开的实施例中,为了更精确地实现对象分割,计算设备130可以基于点云中的点之间的相似性,将点云划分成多个集合。例如,计算设备130可以计算点云中的点之间的欧式距离,将欧式距离小于预定阈值的点划分到一个集合中。之后,计算设备130可以利用多个集合之间的凹凸关系,对所述多个集合进行聚类,由此获取与场景中的多个对象相关联的多个点集。具体地,如果两个集合的并集中的任意两点之间的连线上的所有点都属于该并集,则表明这两个集合的并集为凸集,可以将这两个集合聚类到一起。在一些实施例中,计算设备130可以将点云聚集成超体素,并且可以通过分析相邻超体素之间的凹凸性,来实现点云的聚类。

[0042] 在本公开的实施例中,计算设备130可以从所分割的多个点集中选择一个点集,以便呈现给用户140以供用户140对与所选择的点集相关联的对象的属性进行标注。

[0043] 在框204,计算设备130获取针对对象的属性的标注。在本公开的实施例中,对象的属性可以包括对象的类别、对象的名称以及对象的深度信息中的至少一个。

[0044] 在本公开的实施例中,计算设备130可以从用户140获取针对对象的属性的标注。在一些实施例中,为了进一步减轻人力消耗,计算设备130也可以利用属性预测模型来预测对象的属性,并将预测的属性确定为针对对象的属性标注。在一些实施例中,也可以由其他设备利用属性预测模型来预测对象的属性,计算设备130可以从其他设备获取利用属性预测模型预测的属性,并将预测的属性确定为针对对象的属性标注。稍后将会在图3和图4中描述用于获取针对对象的属性的标注的具体过程。

[0045] 在框206,计算设备130确定点云所描述的场景的图像中的与该点集对应的部分。图像是待标注的图像,例如室外RGB图像。点集可以包括点云中的一个或多个体素。图像中与该点集对应的部分可以包括该图像的一个或多个像素。计算设备130可以从云存储设备120获取(例如下载)相机在拍摄该图像时的成像参数信息,例如相机在拍摄该图像时所处的空间位置、角度以及焦距等信息。利用相机成像原理,计算设备130可以基于所获取的成像参数信息,确定该图像中的与该点集对应的部分。在本公开的实施例中,计算设备130可以基于相机的成像参数,确定该点集中的每个体素所表示的空间点经相机成像后对应于该图像中的哪个像素,从而确定出图像中的对应部分。

[0046] 在框208,计算设备130通过将所获取的针对对象的属性的标注应用于在框206所确定的对应部分,来生成经标注的图像。计算设备130已经在框204中获取了针对对象的属性的标注,即已经获取了点云中与对象相关联的点集中的每个体素的标注信息,例如该体素是表示建筑物118上的点还是表示树木116上的点。计算设备130可以将点云中与对象相关联的点集中的每个体素的标注信息应用于图像中与该体素对应的像素,从而获得该像素的标注信息。

[0047] 在一些实施例中,计算设备130可以针对每个所分割的对象,分别执行方法200的过程,从而为图像中的对应像素赋予标注信息。以此方式,能够实现针对图像的像素级别的标注。

[0048] 对于同一场景的大量图像,方法200通过在点云中实现对象分割和属性标注,并利用相机在拍摄这些图像时的成像参数,将基于点云的标注应用于这些图像。与现有的图像标注方法相比,方法200在点云上实现对象的分割,而不是直接在图像上进行对象分割,从而方法200的分割精度更高。同时方法200利用点云与同一场景下的大量图像的对应关系,将基于点云获取的属性标注映射到这些图像上,而无须针对每个图像进行人工标注,这提高了标注的效率。实验表明,根据本公开的实施例的方法200能够比现有的直接对图像进行标注的方法快30倍左右。对于室外图像而言,通过方法200能够获得大规模的经标注的数据集。

[0049] 图3示出了根据本公开的实施例的用于获取针对对象的属性的标注的方法300的流程图。方法300可以由图1中所示的计算设备130来执行。

[0050] 在框302,计算设备130可以向用户140呈现与对象相关联的点集。为了便于用户140对对象的属性进行标注,计算设备130可以将点云渲染在RGB图像上以呈现给用户140,从而便于用户140对点集中的零散点进行标注。用户140可以对所呈现的点集进行标注,例如用户140可以标注与该点集相关联的对象的类别、名称等。

[0051] 在计算设备130向用户140呈现与对象相关联的点集时,用户140可能会发现该对象的三维边界不够理想。为了获得更精确对象分割,从而获得更精确的标注图像,计算设备130可以允许用户140对所呈现的点集进行调整,例如对所分割的对象的边界进行调整。为此,在框304,计算设备130可以确定用户140是否通过调整点集来调整对象的边界。如果用户140对所呈现的点集进行了调整,则方法300前进到框306。如果用户140没有对所点集进行调整,则方法300前进到框308。

[0052] 在框306,计算设备130可以响应于用户140调整了点集,而对与对象相关联的该点集进行更新,以使得更新后的点集与用户140的调整相适应。对点集进行更新会使方法200中框206所确定的对应部分发生相应的变化,从而能够获得更精确的经标注的图像。

[0053] 在框308,计算设备130可以接收用户140输入的针对该对象的属性的标注,并将该标注与该点集关联起来。以图1为例,用户140可以基于所呈现的与建筑物118相关联的点集,将建筑物118标记为“建筑物”这一类别。计算设备130可以将点云中与建筑物118相关联的所有点都标注为属于“建筑物”这一类别。

[0054] 方法300描述了用于从用户140获取针对对象的属性的标注的过程。为了进一步减轻人力消耗,本公开的实施例也可以基于机器学习的方法来获取针对对象的属性的标注。下面结合图4来描述这一过程。图4示出了根据本公开的实施例的用于获取针对对象的属性的标注的方法400的流程图。方法400可以由图1中所示的计算设备130来执行。

[0055] 在按照方法300通过用户140手工在点云中对对象进行标注,从而获得大量的对象被标注的点云之后,计算设备130可以基于这些已经标注的点云,利用机器学习的方法来训练一个属性预测模型,用于对后续的点云中的对象的属性进行预测。在框402,计算设备130可以利用属性预测模型来预测对象的属性。属性预测模型是利用对象已经被标注的点云来训练的。例如,计算设备130可以基于3D-CNN来训练属性预测模型,从而获得该模型的参数。

[0056] 在框404,计算设备130可以向用户140呈现与对象相关联的点集和该对象的预测的属性。由于机器学习的方法不可避免地会存在一些预测错误,所以为了避免这种预测错误影响图像的标注精度,在本公开的实施例中,计算设备130可以向用户140提供预测的属性,以由用户140判断所预测的属性是否正确。

[0057] 在框406,计算设备130可以确定用户140是否对预测的属性进行了修改。如果用户140未对该属性进行修改,则在框408,计算设备130可以将预测的属性作为针对该对象的属性的标注。作为替代方式,用户140也可以直接确认所预测的属性是正确的。在此情况下,计算设备130同样可以将预测的属性作为针对该对象的属性的标注。

[0058] 另一方面,如果用户140判断所预测的属性不正确,用户140可以对该属性进行修改。如果用户140对该属性进行修改,则在框410,计算设备130可以将经修改的属性确定为针对对象的属性的标注。通过这种方式,在利用先验知识提高标注效率的同时,保证了标注的精度。

[0059] 在框412,计算设备130可以利用新标注的点云来更新属性预测模型。随着已经标注的点云的增加,计算设备130可以利用更多的已标注点云来重新训练属性预测模型,从而实现属性预测模型的更新,以不断提高属性预测模型的预测精度。这样,方法400充分地利用了先验知识,减轻了用户的负担,从而进一步提高了标注效率。

[0060] 根据本公开的实施例,不仅可以基于点云来对图像进行语义标注(例如,对图像中

的对象的类别、名称进行标注),还可以对对象的深度信息进行标注,深度信息指示对象与拍摄图像时相机所处的空间位置之间的距离,更具体地深度信息指示对象上的点的空间位置与相机所处的空间位置之间的距离。计算设备130可以将所述距离与图像中的与该对象对应的部分关联起来,从而生成与图像对应的深度图像。

[0061] 在本公开的实施例中,也可以在不进行对象分割的情况下,确定与图像对应的深度图像。图5示出了根据本公开的实施例的用于生成深度图像的方法500的流程图。方法500可以由图1中所示的计算设备130来执行。

[0062] 在框502,计算设备130获取描述场景的三维信息的点云和同一场景的图像。如之前所描述的,场景的三维信息可以包括场景中对象上的点的空间坐标。例如,计算设备130可以从云存储设备120下载点云和图像。

[0063] 在框504,计算设备130获取点云中的点(即体素)与图像中的像素之间的对应关系。如之前参考图1所描述的,采集实体112在采集点云、图像的同时,还获取相机在拍摄这些图像时的成像参数,例如相机的空间位置、相机的角度、相机的焦距等,并将点云、图像和成像参数关联地上传到云存储设备120中。为此,计算设备130可以从云存储设备120下载成像参数,并基于成像参数来确定点云中的每个体素所表示的空间点经相机成像后对应于该图像中的哪个像素,从而确定出点云中的体素与图像中的像素之间的对应关系。

[0064] 在框506,计算设备130确定点云中的体素所表示的空间点与相机的空间位置之间的距离。例如,计算设备130可以基于体素所表示的三维坐标与相机拍摄图像时所处位置的三维坐标来计算两者之间的欧式距离、曼哈顿距离、切比雪夫距离等。应当理解,以上对于各种距离的举例仅仅是为了说明的目的,而无意限制本公开的范围。在根据本公开的其他实施例中,可以采用任何适当的距离来表示深度信息。

[0065] 在框508,计算设备130可以将所确定的距离与图像中的相应像素关联起来,从而生成与该图像对应的深度图像。深度图像中的像素的值表示图像中的相应像素所表示的空间点与拍摄图像时相机的位置之间的距离。计算设备130可以基于框504中确定的对应关系,将所确定的点云中的体素所表示的空间点与相机的空间位置之间的距离,应用于图像中对应于所述体素的像素,从而获得深度图像中的相应像素的值。

[0066] 利用方法500,能够生成稠密的室外深度图像,并使得经标注的图像数据集具有多样性,填充了目前尚无室外深度图像数据集的空白。

[0067] 图6示出了根据本公开的实施例的用于标注图像的装置600的方框图。装置600可以被包括在图1的计算设备130中或者被实现为计算设备130。如图6所示,装置600包括点集确定模块610,被配置为在描述场景的三维信息的点云中确定与场景中的对象相关联的点集。装置600还包括标注获取模块620,被配置为获取针对对象的属性的标注。装置600还包括对应部分确定模块630,被配置为确定场景的待标注的图像中的与点集对应的部分。此外,装置600还包括标注图像获取模块640,被配置为通过将标注应用于所述部分,来生成经标注的图像。

[0068] 在一些实施例中,点集确定模块610可以包括:去除模块,被配置为从点云中去除与预定对象相关联的预定对象点集;分割模块,被配置为在去除了预定对象点集的点云中分割出与场景中的多个对象相关联的多个点集;以及点集选择模块,被配置为从多个点集中选择与对象相关联的点集。

[0069] 在一些实施例中,点集确定模块610可以包括:划分模块,被配置为基于点云中的点之间的相似性,将点云划分成多个集合;以及聚类模块,被配置为基于多个集合之间的凹凸关系,对多个集合进行聚类,以获取与场景中的多个对象相关联的多个点集;以及点集选择模块,被配置为从多个点集中选择与对象相关联的点集。

[0070] 在一些实施例中,标注获取模块620可以包括:呈现模块,被配置为向用户呈现与对象相关联的点集;以及标注接收模块,被配置为接收用户输入的针对对象的属性的标注。

[0071] 在一些实施例中,装置600还可以包括:调整接收模块,被配置为接收用户对点集的调整;以及更新模块,被配置为基于调整,更新点集。

[0072] 在一些实施例中,标注获取模块620可以包括:预测模块,被配置为利用属性预测模型来预测对象的属性,属性预测模型是基于对象已被标注的点云来训练的;标注确定模块,被配置为将预测的属性确定为针对对象的属性的标注。

[0073] 在一些实施例中,标注获取模块620可以包括:预测模块,被配置为利用属性预测模型来预测对象的属性,属性预测模型是基于对象已被标注的点云来训练的;提供模块,被配置为向用户提供预测的属性;以及标注确定模块,被配置为:响应于接收到用户对预测的属性的确认,将预测的属性确定为针对对象的属性的标注;以及响应于接收到用户对预测的属性的修改,将经修改的属性确定为针对对象的属性的标注。

[0074] 在一些实施例中,装置600还可以包括:预测模型更新模块,被配置为利用点云和针对对象的属性的标注来更新属性预测模型。

[0075] 在一些实施例中,对应部分确定模块630可以包括:相机参数获取模块,被配置为获取相机在拍摄图像时的参数;以及确定模块,被配置为基于参数确定与点集对应的部分。

[0076] 在一些实施例中,对象的属性包括以下至少一项:对象的类别、名称和深度信息。

[0077] 在一些实施例中,标注获取模块620可以包括:深度信息确定模块,被配置为确定对象的深度信息,深度信息指示对象与拍摄图像时相机所处的空间位置之间的距离;标注确定模块,被配置为将深度信息确定为针对对象的属性的标注。

[0078] 在一些实施例中,装置600可以包括:深度图像生成模块,被配置为通过将距离与对应部分相关联,生成与图像对应的深度图像。

[0079] 图7示出了可以用来实施本公开的实施例的电子设备700的示意性框图。设备700可以用于实现图1的计算设备130。如图所示,设备700包括中央处理单元(CPU)701,其可以根据存储在只读存储器(ROM)702中的计算机程序指令或者从存储单元708加载到随机访问存储器(RAM)703中的计算机程序指令,来执行各种适当的动作和处理。在RAM 703中,还可以存储设备700操作所需的各种程序和数据。CPU 701、ROM 702以及RAM 703通过总线704彼此相连。输入/输出(I/O)接口705也连接至总线704。

[0080] 设备700中的多个部件连接至I/O接口705,包括:输入单元706,例如键盘、鼠标等;输出单元707,例如各种类型的显示器、扬声器等;存储单元708,例如磁盘、光盘等;以及通信单元709,例如网卡、调制解调器、无线通信收发机等。通信单元709允许设备700通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0081] 处理单元701执行上文所描述的各个方法和处理,例如方法200、300、400、500。例如,在一些实施例中,方法200、300、400、500可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元708。在一些实施例中,计算机程序的部分或者全部可以经

由ROM 702和/或通信单元709而被载入和/或安装到设备700上。当计算机程序加载到RAM 703并由CPU 701执行时,可以执行上文描述的方法200、300、400、500的一个或多个步骤。备选地,在其他实施例中,CPU 701可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行方法200、300、400、500。

[0082] 本文中以上描述的功能可以至少部分地由一个或多个硬件逻辑部件来执行。例如,非限制性地,可以使用的示范类型的硬件逻辑部件包括:场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统的系统(SOC)、负载可编程逻辑设备(CPLD)等等。

[0083] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器的控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务上执行。

[0084] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0085] 此外,虽然采用特定次序描绘了各操作,但是这应当理解为要求这样操作以所示出的特定次序或以顺序次序执行,或者要求所有图示的操作应被执行以取得期望的结果。在一定环境下,多任务和并行处理可能是有利的。同样地,虽然在上面论述中包含了若干具体实现细节,但是这些不应当被解释为对本公开的范围的限制。在单独的实施例的上下文中描述的某些特征还可以组合地实现在单个实现中。相反地,在单个实现的上下文中描述的各种特征也可以单独地或以任何合适的子组合的方式实现在多个实现中。

[0086] 尽管已经采用特定于结构特征和/或方法逻辑动作的语言描述了本主题,但是应当理解所附权利要求书中所限定的主题未必局限于上面描述的特定特征或动作。相反,上面所描述的特定特征和动作仅仅是实现权利要求书的示例形式。

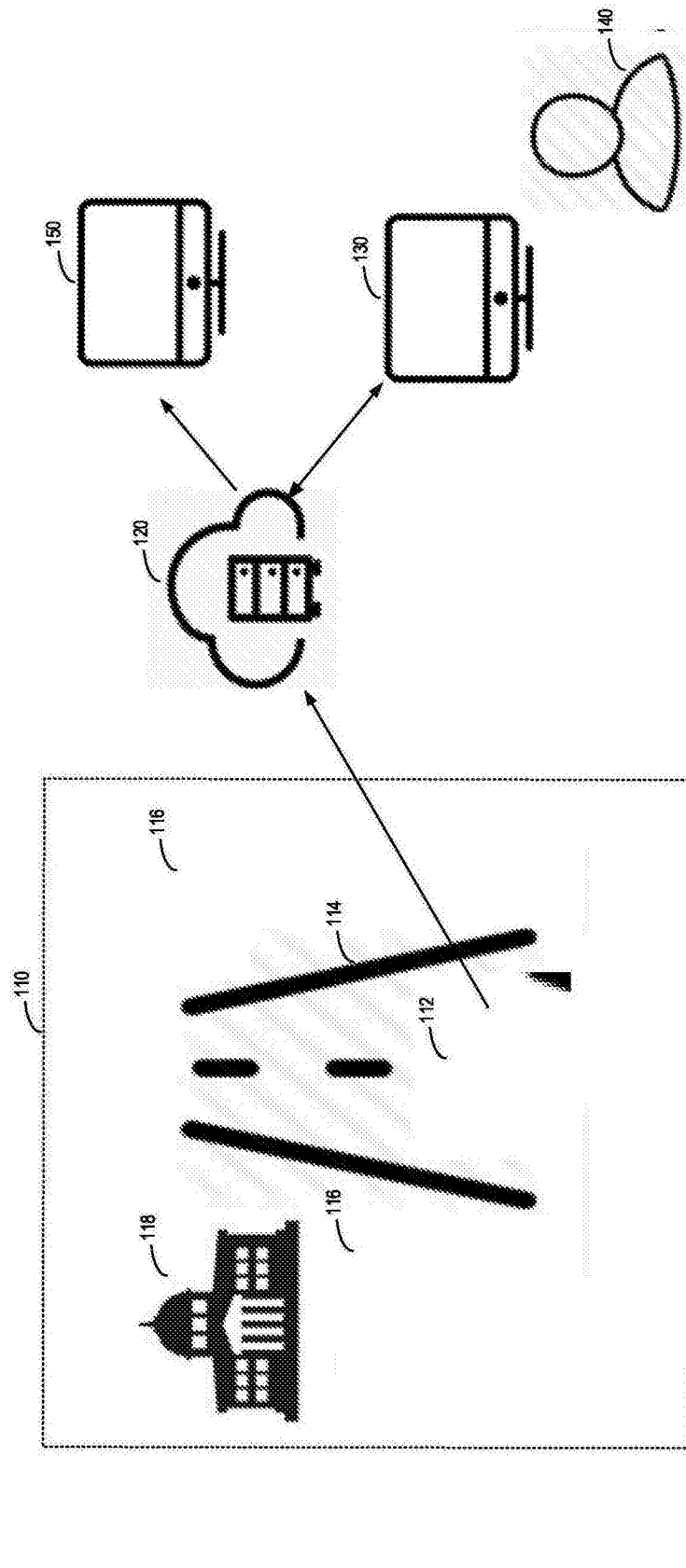


图1

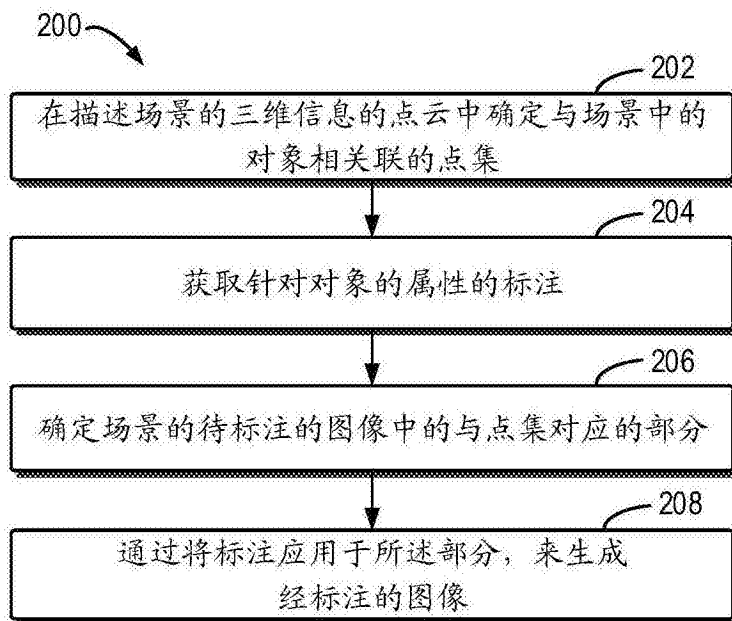


图2

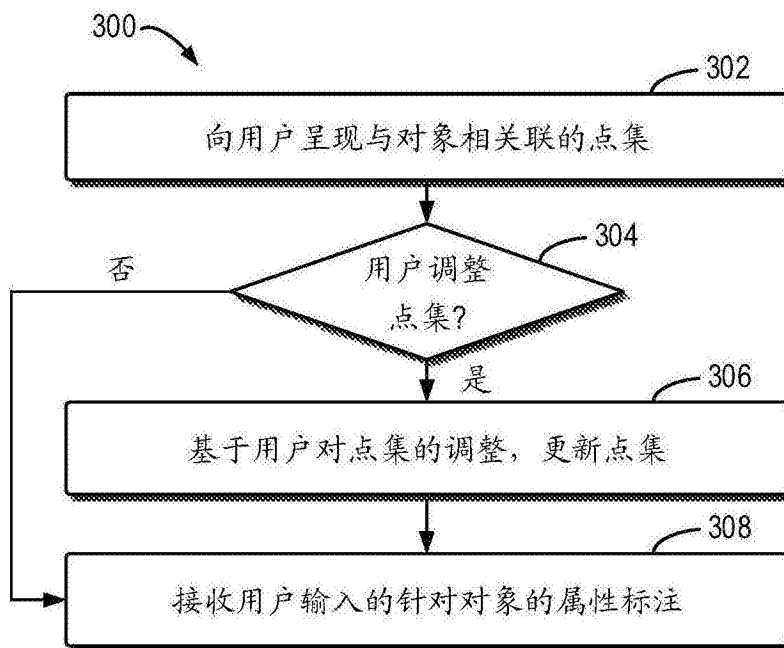


图3

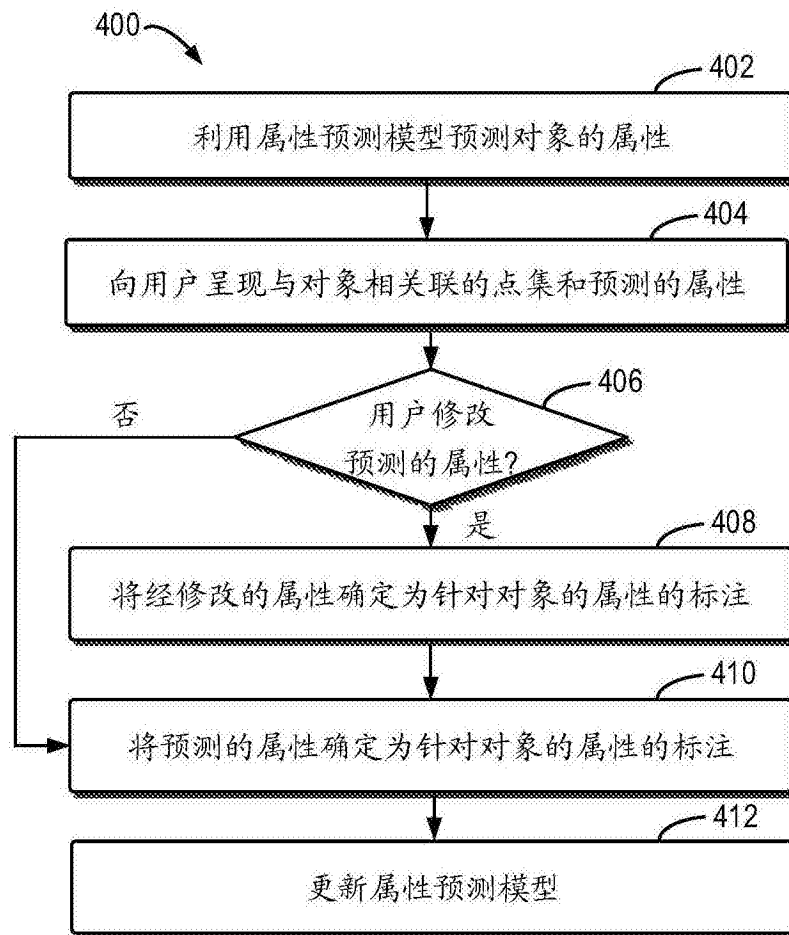


图4

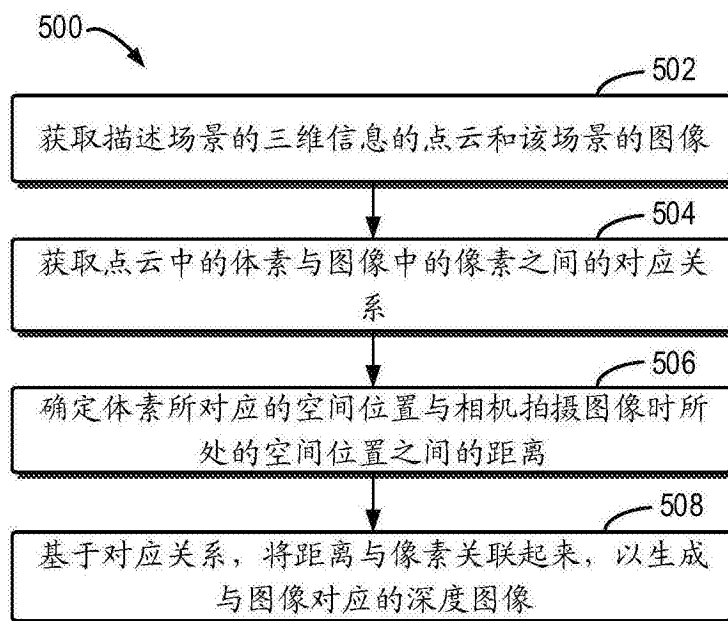


图5

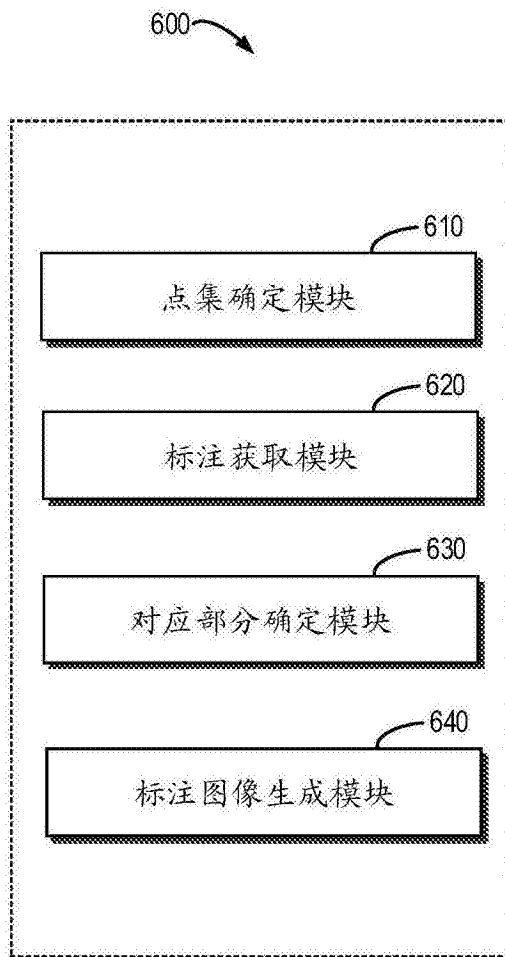


图6

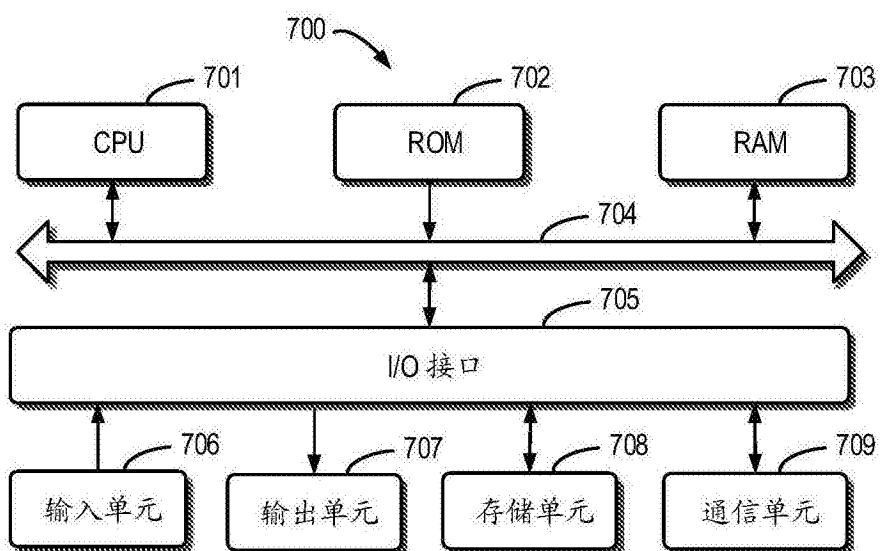


图7