



(12) 发明专利

(10) 授权公告号 CN 112700552 B

(45) 授权公告日 2025.06.27

(21) 申请号 202011641585.7

G06T 7/507 (2017.01)

(22) 申请日 2020.12.31

G06T 7/73 (2017.01)

G06T 7/80 (2017.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 112700552 A

(43) 申请公布日 2021.04.23

(73) 专利权人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(56) 对比文件

CN 106407947 A, 2017.02.15

CN 110286387 A, 2019.09.27

CN 112102409 A, 2020.12.18

审查员 吕梅利

(72) 发明人 王凯 薛景涛 贺亚农 陈辰

(74) 专利代理机构 北京中博世达专利商标代理

有限公司 11274

专利代理师 张影

(51) Int. Cl.

G06T 19/00 (2011.01)

G06T 19/20 (2011.01)

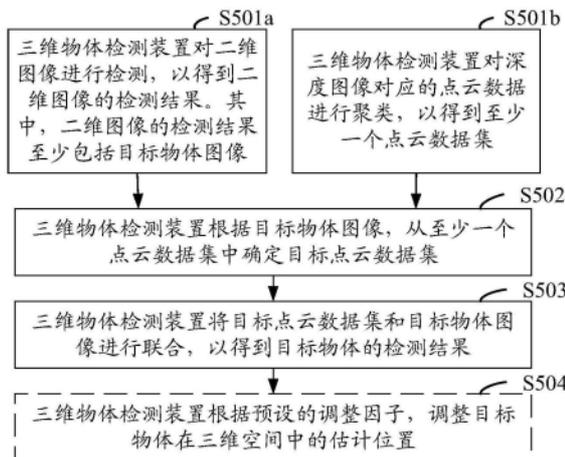
权利要求书3页 说明书22页 附图8页

(54) 发明名称

三维物体检测方法、装置、电子设备及介质

(57) 摘要

本申请提供了一种三维物体检测方法、装置、电子设备及介质,涉及人工智能领域,能够提高三维物体检测精度。该方法包括:三维物体检测装置获取二维图像和至少一个点云数据集。其中,二维图像包括至少一个物体的图像,点云数据集包括多个点云数据,点云数据用于描述至少一个物体在三维空间中的候选区域。之后,三维物体检测装置根据二维图像中目标物体图像,从至少一个点云数据集中确定目标点云数据集。其中,目标点云数据集中的点云数据用于描述目标物体在三维空间中的候选区域。三维物体检测装置将目标点云数据集和目标物体图像进行关联,以得到检测结果。其中,检测结果指示目标物体在三维空间中的估计位置。



1. 一种三维物体检测方法,其特征在于,包括:

获取二维图像和至少一个点云数据集,所述二维图像包括至少一个物体的图像,所述点云数据集包括多个点云数据,所述点云数据用于描述所述至少一个物体在三维空间中的候选区域,所述二维图像是图像传感器采集的信息,所述点云数据是深度传感器采集的信息;

根据所述二维图像中目标物体图像,从所述至少一个点云数据集中确定目标点云数据集,其中,所述目标物体图像包括所述至少一个物体中目标物体的图像,所述目标点云数据集中的点云数据用于描述所述目标物体在所述三维空间中的候选区域,所述目标点云数据集所表示的特征点在所述二维图像中的目标投影区域与目标图像区域之间的相似度大于相似度阈值,所述目标图像区域是所述目标物体图像在所述二维图像中的区域,所述目标投影区域与所述目标图像区域之间的相似度等于所述目标投影区域与所述目标图像区域之间的交并比与所述目标投影区域的投影点间距相乘后,与所述目标投影区域的参考点与所述目标图像区域的参考点之间的距离的比值,所述投影点间距是目标特征点在所述二维图像中的投影点之间的距离;

根据所述目标点云数据集中点云的深度范围,将所述目标物体图像中的部分像素点逆映射至所述三维空间,以得到目标逆映射点;

将所述目标逆映射点对应的三维标注框所标注的区域,作为检测结果,其中,所述检测结果指示所述目标物体在所述三维空间中的估计位置。

2. 根据权利要求1所述的方法,其特征在于,所述根据所述二维图像中目标物体图像,从所述至少一个点云数据集中确定目标点云数据集,包括:

确定第一点云数据集在所述二维图像中的第一投影区域,其中,所述第一点云数据集是所述至少一个点云数据集中的集合;

根据所述第一投影区域和所述目标图像区域,确定所述第一点云数据集为所述目标点云数据集。

3. 根据权利要求2所述的方法,其特征在于,所述确定第一点云数据集在所述二维图像中的第一投影区域,包括:

根据所述第一点云数据集中点云的深度范围,从所述第一点云数据集所表示的特征点中确定第一特征点;

根据所述点云数据与所述二维图像之间的转换参数,确定所述第一特征点在所述二维图像中的第一投影点;

将所述第一投影点对应的二维标注框所标注的区域,作为所述第一投影区域。

4. 根据权利要求2或3所述的方法,其特征在于,所述根据所述第一投影区域和所述目标图像区域,确定所述第一点云数据集为所述目标点云数据集,包括:

根据所述第一投影区域和所述目标图像区域之间的重合程度,以及所述第一投影区域的大小,确定所述第一点云数据集为所述目标点云数据集。

5. 根据权利要求4所述的方法,其特征在于,所述目标点云数据集所表示的特征点在所述二维图像中的目标投影区域满足:

$$\begin{cases} S = IOU * Lj/Dij \\ IOU = S_n/S_U \\ S > T \end{cases}$$

其中,S表示所述目标投影区域与所述目标图像区域之间的相似度,IOU表示所述目标投影区域与所述目标图像区域之间的交并比, S_n 表示所述目标投影区域与所述目标图像区域之间的重叠面积, S_U 表示所述重叠面积与不重叠面积之和,所述不重叠面积为所述目标投影区域与所述目标图像区域之间未重叠的面积, L_j 表示所述目标投影区域的投影点间距,所述投影点间距是目标特征点在所述二维图像中的投影点之间的距离,所述目标特征点属于所述目标点云数据集所表示的特征点,且指示所述目标点云数据集中点云的深度范围, D_{ij} 表示所述目标投影区域的参考点与所述目标图像区域的参考点之间的距离,T表示相似度阈值。

6. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

根据预设的调整因子,调整所述检测结果指示的估计位置,其中,所述调整因子指示所述目标物体在所述三维空间中的真实位置与估计位置之间的差异。

7. 根据权利要求1所述的方法,其特征在于,所述点云数据集所表示的特征点的数量小于数量阈值。

8. 一种三维物体检测装置,其特征在于,包括:

获取单元,用于获取二维图像和至少一个点云数据集,所述二维图像包括至少一个物体的图像,所述点云数据集包括多个点云数据,所述点云数据用于描述所述至少一个物体在三维空间中的候选区域,所述二维图像是图像传感器采集的信息,所述点云数据是深度传感器采集的信息;

处理单元,用于根据所述二维图像中目标物体图像,从所述至少一个点云数据集中确定目标点云数据集,其中,所述目标物体图像包括所述至少一个物体中目标物体的图像,所述目标点云数据集中的点云数据用于描述所述目标物体在所述三维空间中的候选区域,所述目标点云数据集所表示的特征点在所述二维图像中的目标投影区域与目标图像区域之间的相似度大于相似度阈值,所述目标图像区域是所述目标物体图像在所述二维图像中的区域,所述目标投影区域与所述目标图像区域之间的相似度等于所述目标投影区域与所述目标图像区域之间的交并比与所述目标投影区域的投影点间距相乘后,与所述目标投影区域的参考点与所述目标图像区域的参考点之间的距离的比值,所述投影点间距是目标特征点在所述二维图像中的投影点之间的距离;

所述处理单元,还用于根据所述目标点云数据集中点云的深度范围,将所述目标物体图像中的部分像素点逆映射至所述三维空间,以得到目标逆映射点;

所述处理单元,还用于将所述目标逆映射点对应的三维标注框所标注的区域,作为检测结果,其中,所述检测结果指示所述目标物体在所述三维空间中的估计位置。

9. 根据权利要求8所述的装置,其特征在于,所述处理单元,用于根据所述二维图像中目标物体图像,从所述至少一个点云数据集中确定目标点云数据集,具体包括:

确定第一点云数据集在所述二维图像中的第一投影区域,其中,所述第一点云数据集是所述至少一个点云数据集中的集合;

根据所述第一投影区域和所述目标图像区域,确定所述第一点云数据集为所述目标点

云数据集。

10. 根据权利要求9所述的装置,其特征在于,所述处理单元,用于确定第一点云数据集在所述二维图像中的第一投影区域,具体包括:

根据所述第一点云数据集中点云的深度范围,从所述第一点云数据集所表示的特征点中确定第一特征点;

根据所述点云数据与所述二维图像之间的转换参数,确定所述第一特征点在所述二维图像中的第一投影点;

将所述第一投影点对应的二维标注框所标注的区域,作为所述第一投影区域。

11. 根据权利要求9或10所述的装置,其特征在于,所述处理单元,用于根据所述第一投影区域和所述目标图像区域,确定所述第一点云数据集为所述目标点云数据集,具体包括:

根据所述第一投影区域和所述目标图像区域之间的重合程度,以及所述第一投影区域的大小,确定所述第一点云数据集为所述目标点云数据集。

12. 根据权利要求11所述的装置,其特征在于,所述目标点云数据集所表示的特征点在所述二维图像中的目标投影区域满足:

$$\begin{cases} S = IOU * L_j / D_{ij} \\ IOU = S_{\cap} / S_{\cup} \\ S > T \end{cases}$$

其中,S表示所述目标投影区域与所述目标图像区域之间的相似度,IOU表示所述目标投影区域与所述目标图像区域之间的交并比, S_{\cap} 表示所述目标投影区域与所述目标图像区域之间的重叠面积, S_{\cup} 表示所述重叠面积与不重叠面积之和,所述不重叠面积为所述目标投影区域与所述目标图像区域之间未重叠的面积, L_j 表示所述目标投影区域的投影点间距,所述投影点间距是目标特征点在所述二维图像中的投影点之间的距离,所述目标特征点属于所述目标点云数据集所表示的特征点,且指示所述目标点云数据集中点云的深度范围, D_{ij} 表示所述目标投影区域的参考点与所述目标图像区域的参考点之间的距离,T表示相似度阈值。

13. 根据权利要求8所述的装置,其特征在于,所述处理单元,还用于:

根据预设的调整因子,调整所述检测结果指示的估计位置,其中,所述调整因子指示所述目标物体在所述三维空间中的真实位置与估计位置之间的差异。

14. 根据权利要求8所述的装置,其特征在于,所述点云数据集所表示的特征点的数量小于数量阈值。

15. 一种电子设备,其特征在于,包括:处理器和存储器,所述处理器和所述存储器耦合,所述存储器存储有程序指令,当所述存储器存储的程序指令被所述处理器执行时,如权利要求1至7中任一项所述的三维物体检测方法被执行。

16. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储程序,所述程序被处理器调用时,权利要求1至7中任一项所述的三维物体检测方法被执行。

17. 一种计算机程序产品,其特征在于,所述计算机程序产品包括计算机指令,当所述计算机指令在计算机上运行时,使得计算机执行如权利要求1至7中任一项所述的三维物体检测方法。

三维物体检测方法、装置、电子设备及介质

技术领域

[0001] 本申请涉及人工智能(artificial intelligence, AI)技术领域,尤其涉及一种三维物体检测方法、装置、电子设备及介质。

背景技术

[0002] 机器人对环境中的物体具备辨识能力,从而实现路径规划、避障等功能。其中,物体的三维(three-dimensional, 3D)空间尺寸对机器人理解环境尤为重要。示例性的,设备获取场景的点云之后,基于场景的点云确定候选物体区域,再从点云中选择位于候选物体区域中的目标点,利用目标点的位置信息,调整候选物体区域,从而定位出物体的三维空间位置。

[0003] 然而,在某些观测视角下,设备无法获取足够的点云,导致物体无法辨识,使得三维物体检测精度低、假阳率高。

发明内容

[0004] 本申请实施例提供一种三维物体检测方法、装置、电子设备及介质,能够提高三维物体检测精度。

[0005] 为达到上述目的,本申请实施例采用如下技术方案:

[0006] 第一方面,本申请实施例提供一种三维物体检测方法,该方法的执行主体可以是三维物体检测装置。该方法包括:获取二维图像和至少一个点云数据集,所述二维图像包括至少一个物体的图像,所述点云数据集包括多个点云数据,所述点云数据用于描述所述至少一个物体在三维空间中的候选区域,所述二维图像是图像传感器采集的信息,所述点云数据是深度传感器采集的信息;根据所述二维图像中目标物体图像,从所述至少一个点云数据集中确定目标点云数据集,其中,所述目标物体图像包括所述至少一个物体中目标物体的图像,所述目标点云数据集中的点云数据用于描述所述目标物体在所述三维空间中的候选区域;将所述目标点云数据集和所述目标物体图像进行关联,以得到检测结果,其中,所述检测结果指示所述目标物体在所述三维空间中的估计位置。

[0007] 本方法中通过三维物体检测装置获取二维图像和至少一个点云数据集。其中,二维图像包括至少一个物体的图像。点云数据集包括多个点云数据,点云数据用于描述至少一个物体在三维空间中的候选区域。二维图像是图像传感器采集的信息,点云数据是深度传感器采集的信息。之后,三维物体检测装置根据二维图像中目标物体图像,从至少一个点云数据集中确定目标点云数据集。其中,目标物体图像包括至少一个物体中目标物体的图像,目标点云数据集中的点云数据用于描述目标物体在三维空间中的候选区域。然后,三维物体检测装置将目标点云数据集和目标物体图像进行关联,以得到检测结果。其中,检测结果指示目标物体在三维空间中的估计位置。

[0008] 在本申请实施例提供的三维物体检测方法中,由于二维图像的处理精度高,目标物体图像能够准确地呈现目标物体在二维图像中的区域,利用目标物体图像来筛选目标点

云数据集,实现对点云数据集几何分割聚类,无需获取大量的三维训练数据。即使物体被遮挡,也能够获取到目标点云数据集,在一定程度上提高了目标物体对应的目标点云数据集的精准度。并且,三维物体检测装置将目标点云数据集和目标物体图像进行关联,以得到检测结果。由于二维图像的处理精度高,即使目标物体的点云数据不足,也能够精准地确定目标物体在三维空间中的估计位置,避免假阳性高的问题。本申请实施例三维物体检测方法无需获取三维训练数据,避免了“基于三维训练数据训练模型”所导致的“泛化性差”的问题。

[0009] 在一种可能的设计中,所述根据所述二维图像中目标物体图像,从所述至少一个点云数据集中确定目标点云数据集,包括:确定第一点云数据集在所述二维图像中的第一投影区域,其中,所述第一点云数据集是所述至少一个点云数据集中的集合;根据所述第一投影区域和目标图像区域,确定所述第一点云数据集为所述目标点云数据集,其中,所述目标图像区域是所述目标物体图像在所述二维图像中的区域。

[0010] 在本方法中,三维物体检测装置根据二维图像中目标物体图像,从至少一个点云数据集中确定目标点云数据集,包括:三维物体检测装置确定第一点云数据集在二维图像中的第一投影区域。其中,第一点云数据集是至少一个点云数据集中的集合。然后,三维物体检测装置根据第一投影区域和目标图像区域,确定第一点云数据集为目标点云数据集。其中,目标图像区域是目标物体图像在二维图像中的区域。

[0011] 也就是说,三维物体检测装置是基于两个区域(即目标图像区域与一个点云数据集在二维图像上的投影区域),来确定一个点云数据集是否为目标点云数据集。由于目标物体图像属于二维图像,三维物体检测装置对二维图像的检测与识别精度高,联合目标物体图像识别出目标点云数据集,也就能够相应提高目标点云数据集的识别精度。

[0012] 在一种可能的设计中,所述确定第一点云数据集在所述二维图像中的第一投影区域,包括:根据所述第一点云数据集中点云的深度范围,从所述第一点云数据集所表示的特征点中确定第一特征点;根据所述点云数据与所述二维图像之间的转换参数,确定所述第一特征点在所述二维图像中的第一投影点;将所述第一投影点对应的二维标注框所标注的区域,作为所述第一投影区域。

[0013] 在本方法中,三维物体检测装置确定第一点云数据集在二维图像中的第一投影区域,包括:三维物体检测装置根据第一点云数据集中点云的深度范围,从第一点云数据集所表示的特征点中确定第一特征点,如最远点、最近点。然后,三维物体检测装置根据点云数据与二维图像之间的转换参数,如深度传感器的内参、旋转矩阵、或平移矩阵,确定第一特征点在二维图像中的第一投影点。三维物体检测装置将第一投影点对应的二维标注框所标注的区域,作为第一投影区域。示例性的,二维标注框是以第一投影点为对角点的框。

[0014] 也就是说,在三维物体检测装置确定了一个点云数据集中的第一特征点的情况下,首先确定第一特征点在二维图像上的投影点,即第一投影点。由于第一投影点是第一点云数据集中最远点、最近点在二维图像上的投影。所以,第一投影区域是第一投影点之间的区域,即第一投影点对应的二维标注框所标注的区域,从而实现了第一点云数据集在二维图像上的准确投影。

[0015] 在一种可能的设计中,所述根据所述第一投影区域和目标图像区域,确定所述第一点云数据集为所述目标点云数据集,包括:根据所述第一投影区域和所述目标图像区域

之间的重合程度,以及所述第一投影区域的大小,确定所述第一点云数据集为所述目标点云数据集。

[0016] 在本方法中,三维物体检测装置根据第一投影区域和目标图像区域,确定第一点云数据集为目标点云数据集,包括:三维物体检测装置根据第一投影区域和目标图像区域之间的重合程度,以及第一投影区域的大小,确定第一点云数据集为目标点云数据集。

[0017] 也就是说,即使第一投影区域与目标图像区域重合,若“第一投影区域”面积偏小,则第一点云数据集中点云数据所表示的特征点可能是目标物体的一部分。由于目标物体的一部分无法准确地表征目标物体整体在三维空间中的估计位置,所以,此类的点云数据集不作为目标点云数据集。如此,三维物体检测装置在确定目标点云数据集的过程中,需要结合“第一投影区域和目标图像区域之间的重合程度”和“第一投影区域的大小”两个因素,以更准确地确定目标点云数据集。

[0018] 在一种可能的设计中,所述目标点云数据集所表示的特征点在所述二维图像中的目标投影区域满足:

$$[0019] \quad \begin{cases} S = IOU * L_j / D_{ij} \\ IOU = S_{\cap} / S_{\cup} \\ S > T \end{cases}$$

[0020] 其中,S表示所述目标投影区域与所述目标图像区域之间的相似度,IOU表示所述目标投影区域与所述目标图像区域之间的交并比, S_{\cap} 表示所述目标投影区域与所述目标图像区域之间的重叠面积, S_{\cup} 表示所述重叠面积与不重叠面积之和,所述不重叠面积为所述目标投影区域与所述目标图像区域之间未重叠的面积, L_j 表示所述目标投影区域的投影点间距,所述投影点间距是目标特征点在所述二维图像中的投影点之间的距离,所述目标特征点属于所述目标点云数据集所表示的特征点,且指示所述目标点云数据集中点云的深度范围, D_{ij} 表示所述目标投影区域的参考点与所述目标图像区域的参考点之间的距离,T表示相似度阈值。

[0021] 在一种可能的设计中,所述将所述目标点云数据集和所述目标物体图像进行关联,以得到检测结果,包括:根据所述目标点云数据集中点云的深度范围,将所述目标物体图像中的部分像素点逆映射至所述三维空间,以得到目标逆映射点;将所述目标逆映射点对应的三维标注框所标注的区域,作为所述检测结果。

[0022] 在本方法中,三维物体检测装置将目标点云数据集和目标物体图像进行关联,以得到检测结果,包括:三维物体检测装置根据目标点云数据集中点云的深度范围,将目标物体图像中的部分像素点逆映射至三维空间,以得到目标逆映射点。例如,若目标物体图像在二维图像中是矩形区域的情况下,将位于对角点上的像素点逆映射至三维空间,以得到目标逆映射点。然后,三维物体检测装置将目标逆映射点对应的三维标注框所标注的区域,作为检测结果。

[0023] 也就是说,三维物体检测装置是利用了目标点云数据集和目标物体图像来确定检测结果的,以避免目标物体被遮挡、视角不全所导致的“假阳率高”的问题。

[0024] 在一种可能的设计中,本申请实施例三维物体检测方法还包括:根据预设的调整因子,调整所述检测结果指示的估计位置,其中,所述调整因子指示所述目标物体在所述三维空间中的真实位置与估计位置之间的差异。

[0025] 本方法还包括：三维物体检测装置根据预设的调整因子，调整检测结果指示的估计位置。其中，调整因子指示目标物体在三维空间中的真实位置与估计位置之间的差异，以使三维物体检测装置所确定的估计位置更贴合实际物体尺寸，提高了物体检测的精准度。

[0026] 在一种可能的设计中，所述点云数据集所表示的特征点的数量小于数量阈值，以剔除描述背景物体的点云数据集，有助于降低三维物体检测装置的运算量。

[0027] 第二方面，本申请实施例提供一种三维物体检测装置，该三维物体检测装置可以为上述第一方面或第一方面任一种可能的设计中的装置，或者实现上述功能的芯片；所述三维物体检测装置包括实现上述方法相应的模块、单元、或手段 (means)，该模块、单元、或 means 可以通过硬件实现，软件实现，或者通过硬件执行相应的软件实现。该硬件或软件包括一个或多个与上述功能相对应的模块或单元。

[0028] 该三维物体检测装置包括获取单元和处理单元。其中，所述获取单元，用于获取二维图像和至少一个点云数据集，所述二维图像包括至少一个物体的图像，所述点云数据集包括多个点云数据，所述点云数据用于描述所述至少一个物体在三维空间中的候选区域，所述二维图像是图像传感器采集的信息，所述点云数据是深度传感器采集的信息；

[0029] 所述处理单元，用于根据所述二维图像中目标物体图像，从所述至少一个点云数据集中确定目标点云数据集，其中，所述目标物体图像包括所述至少一个物体中目标物体的图像，所述目标点云数据集中的点云数据用于描述所述目标物体在所述三维空间中的候选区域；

[0030] 所述处理单元，还用于将所述目标点云数据集和所述目标物体图像进行关联，以得到检测结果，其中，所述检测结果指示所述目标物体在所述三维空间中的估计位置。

[0031] 在一种可能的设计中，所述处理单元，用于根据所述二维图像中目标物体图像，从所述至少一个点云数据集中确定目标点云数据集，具体包括：

[0032] 确定第一点云数据集在所述二维图像中的第一投影区域，其中，所述第一点云数据集是所述至少一个点云数据集中的集合；

[0033] 根据所述第一投影区域和目标图像区域，确定所述第一点云数据集为所述目标点云数据集，其中，所述目标图像区域是所述目标物体图像在所述二维图像中的区域。

[0034] 在一种可能的设计中，所述处理单元，用于确定第一点云数据集在所述二维图像中的第一投影区域，具体包括：

[0035] 根据所述第一点云数据集中点云的深度范围，从所述第一点云数据集所表示的特征点中确定第一特征点；

[0036] 根据所述点云数据与所述二维图像之间的转换参数，确定所述第一特征点在所述二维图像中的第一投影点；

[0037] 将所述第一投影点对应的二维标注框所标注的区域，作为所述第一投影区域。

[0038] 在一种可能的设计中，所述处理单元，用于根据所述第一投影区域和目标图像区域，确定所述第一点云数据集为所述目标点云数据集，具体包括：

[0039] 根据所述第一投影区域和所述目标图像区域之间的重合程度，以及所述第一投影区域的大小，确定所述第一点云数据集为所述目标点云数据集。

[0040] 在一种可能的设计中，所述目标点云数据集所表示的特征点在所述二维图像中的目标投影区域满足：

$$[0041] \quad \begin{cases} S = IOU * L_j / D_{ij} \\ IOU = S_n / S_u \\ S > T \end{cases}$$

[0042] 其中, S表示所述目标投影区域与所述目标图像区域之间的相似度, IOU表示所述目标投影区域与所述目标图像区域之间的交并比, S_n 表示所述目标投影区域与所述目标图像区域之间的重叠面积, S_u 表示所述重叠面积与不重叠面积之和, 所述不重叠面积为所述目标投影区域与所述目标图像区域之间未重叠的面积, L_j 表示所述目标投影区域的投影点间距, 所述投影点间距是目标特征点在所述二维图像中的投影点之间的距离, 所述目标特征点属于所述目标点云数据集所表示的特征点, 且指示所述目标点云数据集中点云的深度范围, D_{ij} 表示所述目标投影区域的参考点与所述目标图像区域的参考点之间的距离, T表示相似度阈值。

[0043] 在一种可能的设计中, 所述处理单元, 用于将所述目标点云数据集和所述目标物体图像进行关联, 以得到检测结果, 具体包括:

[0044] 根据所述目标点云数据集中点云的深度范围, 将所述目标物体图像中的部分像素点逆映射至所述三维空间, 以得到目标逆映射点;

[0045] 将所述目标逆映射点对应的三维标注框所标注的区域, 作为所述检测结果。

[0046] 在一种可能的设计中, 所述处理单元, 还用于:

[0047] 根据预设的调整因子, 调整所述检测结果指示的估计位置, 其中, 所述调整因子指示所述目标物体在所述三维空间中的真实位置与估计位置之间的差异。

[0048] 在一种可能的设计中, 所述点云数据集所表示的特征点的数量小于数量阈值。

[0049] 第三方面, 本申请实施例提供一种电子设备, 该电子设备包括处理器和存储器。该处理器和该存储器进行相互的通信。该处理器用于执行该存储器中存储的指令, 以使得电子设备执行如第一方面或第一方面的任一种设计中的三维物体检测方法。

[0050] 第四方面, 本申请实施例提供一种计算机可读存储介质, 该计算机可读存储介质中存储有指令, 该指令用于指示设备执行上述第一方面或第一方面的任一种设计中的三维物体检测方法。

[0051] 第五方面, 本申请提供了一种包含指令的计算机程序产品, 当其在设备上运行时, 使得设备执行上述第一方面或第一方面的任一种设计中的三维物体检测方法。

[0052] 第六方面, 本申请实施例提供一种芯片, 包括逻辑电路和输入输出接口。其中, 输入输出接口用于与芯片之外的模块通信, 例如, 该芯片可以为实现上述第一方面或第一方面任一种可能的设计中的三维物体检测装置功能的芯片。输入输出接口输入二维图像和至少一个点云数据集, 输入输出接口输出检测结果。逻辑电路用于运行计算机程序或指令, 以实现上述第一方面或第一方面任一种可能的设计中的三维物体检测方法。

[0053] 第七方面, 本申请实施例提供一种机器人, 包括: 图像传感器、深度传感器、处理器、用于存储处理器可执行指令的存储器。其中, 图像传感器用于采集二维图像, 深度传感器用于采集至少一个点云数据集, 处理器被配置可执行指令, 以实现如上述第一方面或第一方面任一种可能的设计中的三维物体检测方法。

[0054] 第八方面, 本申请实施例提供一种服务器, 包括: 处理器、用于存储处理器可执行指令的存储器。其中, 处理器被配置可执行指令, 以实现如上述第一方面或第一方面任一种

可能的设计中的三维物体检测方法。

[0055] 其中,第二方面至第八方面中任一种设计所带来的技术效果可参考上文所提供的对应的方法中的有益效果,此处不再赘述。

附图说明

- [0056] 图1为本申请实施例提供的一种系统架构示意图;
- [0057] 图2为本申请实施例提供的又一种系统架构示意图;
- [0058] 图3为本申请实施例提供的一种三维物体检测方法的流程示意图;
- [0059] 图4a为本申请实施例提供的一种棋盘格图像;
- [0060] 图4b为本申请实施例提供的再一种三维物体检测方法的流程示意图;
- [0061] 图5为本申请实施例提供的又一种三维物体检测方法的流程示意图;
- [0062] 图6a为本申请实施例提供的一种模型训练阶段的流程示意图;
- [0063] 图6b为本申请实施例提供的一种模型应用阶段的流程示意图;
- [0064] 图6c为本申请实施例提供的一种2D检测框的示意图;
- [0065] 图7a为本申请实施例提供的又一种三维物体检测方法的流程示意图;
- [0066] 图7b为本申请实施例提供的一种法线估计的示意图;
- [0067] 图8为本申请实施例提供的又一种三维物体检测方法的流程示意图;
- [0068] 图9a为本申请实施例提供的一种最远点和最近点的位置示意图;
- [0069] 图9b为本申请实施例提供的一种投影区域的位置示意图;
- [0070] 图9c为本申请实施例提供的一种目标投影区域和目标图像区域的位置示意图;
- [0071] 图10为本申请实施例提供的又一种三维物体检测方法的流程示意图;
- [0072] 图11为本申请实施例提供的再一种设备的结构示意图。

具体实施方式

[0073] 本申请的说明书以及附图中的术语“第一”和“第二”等是用于区别不同的对象,或者用于区别对同一对象的不同处理,而不是用于描述对象的特定顺序。此外,本申请的描述中所提到的术语“包括”和“具有”以及它们的任何变形,意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元,而是可选地还包括其他没有列出的步骤或单元,或可选地还包括对于这些过程、方法、产品或设备固有的其它步骤或单元。需要说明的是,本申请实施例中,“示例性地”或者“例如”等词用于表示作例子、例证或说明。本申请实施例中被描述为“示例性地”或者“例如”的任何实施例或设计方案不应被解释为比其它实施例或设计方案更优选或更具优势。确切而言,使用“示例性地”或者“例如”等词旨在以具体方式呈现相关概念。

[0074] 为了使得本申请更加的清楚,首先对本申请提到的部分概念和处理流程作简单介绍。

[0075] 1、鲁棒(robust)性

[0076] 鲁棒性,是指在异常和危险情况下系统生存的能力,或控制系统在一定(结构,大小)的参数摄动下,维持其它某些性能的特性。

[0077] 2、假阳率

[0078] 假阳率,是指“通过深度学习模型获取的结果为错误的正类”的概率,即深度学习模型将非目标样本确定为正确的目标样本的概率。

[0079] 3、二维(two dimensions,2D)图像

[0080] 二维图像,是指不包含深度信息的平面图像。二维图像可以包括红绿蓝(red green blue,RGB)图像、灰度图像等。

[0081] 4、深度图像(depth image)

[0082] 深度图像,也被称为距离影像(range image),是指将从深度传感器到空间中各点的距离(或称为深度)作为像素值的图像。深度图像直接反映了空间中物体可见表面的几何形状。

[0083] 5、点云数据

[0084] 点云,是指在一定的空间参考系下表达目标物体的空间分布和表面特性的点的集合。在本申请实施例中,点云数据用于表征该点云中每个点在空间参考坐标系下的三维坐标值。空间参考坐标系可以是深度传感器对应的坐标系。

[0085] 6、点云簇

[0086] 点云簇,是指对点云数据进行一系列计算(如几何分割、聚类处理等)之后,所得到的满足预设划分规则的一部分点云数据所表征的点。其中,计算方法可以是基于点云数据密度的聚类方法、基于kdtree的最近邻方法、k-means方法,以及深度学习等方法。

[0087] 在本申请实施例中,将一个点云簇对应的点云数据描述为“点云数据集”。

[0088] 7、三维(three dimensions,3D)物体检测

[0089] 三维物体检测,能够提供物体地图,以使机器人更好的定位。由于物体是机器人理解环境的基础,所以,物体可以作为一种语义,以提高机器人导航智能性。三维物体检测能够将物体从图像平面扩展到现实世界,更好的实现人机交互。下面,给出了基于深度学习的三维物体检测方法的实现过程:

[0090] 设备获取场景的点云之后,基于场景的点云确定候选物体区域,再从点云中选择位于候选物体区域中的目标点,利用目标点的位置信息,调整候选物体区域,从而定位出物体的三维空间位置。然而,在某些观测视角下,设备无法获取足够的点云,导致物体无法辨识,从而使得三维物体检测精度低、假阳率高。

[0091] 有鉴于此,本申请实施例提供了一种三维物体检测方法。本申请实施例提供的三维物体检测方法可以应用于如图1所示的设备中。该设备包括第一设备101和第二设备102。第一设备01为图像采集设备,该图像采集设备包括图像传感器和深度传感器。其中,图像传感器用于采集二维图像,如RGB图像、灰度图像等。图像传感器可以例如但不限于如下介绍:RBG相机、数字单反(digital single-lens reflex,DSLR)相机、傻瓜相机、摄像机、可穿戴设备、增强现实技术(augmented reality,AR)设备、虚拟现实(virtual reality,VR)设备、车载设备、智慧屏等。深度传感器用于采集深度图像。深度传感器可以例如但不限于如下介绍:深度相机、飞行时间(time of flight,TOF)相机、或激光雷达、照相式扫描仪、或激光雷达探测与测量(light detection and ranging,LiDAR)传感器。第二设备102为处理设备,该处理设备具有中央处理器(central processing unit,CPU)和/或图形处理器(graphics processing unit,GPU),用于对图像采集设备采集的图像进行处理,从而实现三维物体检测。

[0092] 需要说明的是,第一设备101和第二设备102可以设置在机器人本体上,如图1所示。例如,第一设备101和第二设备102可以设置于机器人的头部(图1未示出),也可以设置于机器人的躯体部位,如图1所示。当然,第一设备101和第二设备102还可以设置在机器人本体的其他部位,本申请实施例对此不作限定。

[0093] 另外,第一设备101和第二设备102可以是独立的设备,也可以合设。例如,第一设备101是第二设备102的一部分。此种情况下,第一设备101与第二设备102之间通过总线连接。示例性的,总线可以实现为双向同步串行总线,双向同步串行总线包括一根串行数据线(serial data line,SDA)和一根串行时钟线(derail clock line,SCL)。此种情况下,第一设备101和第二设备102包括集成电路(inter-integrated circuit,I2C)接口。第一设备101和第二设备102之间通过I2C接口连接的双向同步串行总线进行通信。或者,第一设备101和第二设备102包括移动产业处理器(mobile industry processor interface,MIPI)接口。第一设备101和第二设备102之间通过MIPI接口连接的双向同步串行总线进行通信。或者,第一设备101和第二设备102包括通用输入输出(general-purpose input/output,GPIO)接口。第一设备101和第二设备102之间通过GPIO接口连接的双向同步串行总线进行通信。

[0094] 在本申请实施例中,以“第一设备101和第二设备102是独立的设备”为例,进行描述。在“第一设备101和第二设备102是独立的设备”的情况下,第一设备101和第二设备102可以设置在不同的位置上。例如,第一设备101设置于机器人的躯体部位,第二设备102设置于机器人本体的外部,如图2所示。此种情况下,第二设备102可以是物理设备或物理设备集群,例如终端、服务器、或服务器集群。第二设备102也可以是虚拟化的云设备,例如云计算集群中的至少一个云计算设备。第一设备101和第二设备102均可以包括支持无线通信技术的装置或芯片。其中,无线通信技术可以例如但不限于如下介绍:近距离无线通信(near field communication,NFC)技术、红外(infrared,IR)技术、全球移动通讯系统(global system for mobile communications,GSM)、通用分组无线服务(general packet radio service,GPRS)、码分多址接入(code division multiple access,CDMA)、宽带码分多址(wideband code division multiple access,WCDMA)、时分码分多址(time-division code division multiple access,TD-SCDMA)、长期演进(long term evolution,LTE)、蓝牙(bluetooth,BT)、全球导航卫星系统(global navigation satellite system,GNSS)、或调频(frequency modulation,FM)等。GNSS可以包括全球卫星定位系统(global positioning system,GPS)、全球导航卫星系统(global navigation satellite system,GLONASS)、北斗卫星导航系统(beidou navigation satellite system,BDS)、准天顶卫星系统(quasi-zenith satellite system,QZSS)和/或星基增强系统(satellite based augmentation systems,SBAS)。

[0095] 其中,图1和图2中的机器人可以是服务型机器人,如居家环境中的扫地机器人、送货上门的机器人、儿童教育机器人等。图1和图2中的机器人也可以是机械型机器人,如工厂中运输货物的机器人。另外,在图1和图2中,仅以机器人为例进行描述,机器人可以替换为智能家电设备,如智能音箱、智能电视等,以定位人体在三维空间中的估计位置,从而切换自身的工作状态。例如,当智能音箱定位人体在三维空间中的估计位置大于一定阈值的情况下,停止播放音频。反之,当智能音箱定位人体在三维空间中的估计位置小于一定阈值的

情况下,开始播放音频。图1和图2中的机器人还可以替换为无人机,如送货上门的无人机、监控森林火险的无人机、用于喷洒农药、化肥的无人机等。

[0096] 为了使得本申请的技术方案更加清楚、易于理解,下面,通过两个阶段对本申请实施例提供的第一种三维物体检测方法进行介绍:

[0097] 本申请实施例还提供了第二种三维物体检测方法,该方法包括两个阶段,具体说明如下:

[0098] 第一阶段,采集阶段。在此阶段中,三维物体检测装置获取二维图像和深度图像对应的点云数据。参见图3,该阶段的步骤介绍如下:

[0099] S301a、第一设备采集二维图像。

[0100] 其中,二维图像包括第一场景中至少一个物体的平面图像。第一场景是第一设备扫描范围内的场景。例如,在第一设备处于客厅的情况下,第一场景可以是第一设备扫描到的客厅范围内的场景,第一场景中的物体可以例如但不限于人、电视机、桌子、椅子、或沙发等。在第一设备处于卧室的情况下,第一场景可以是第一设备扫描到的卧室范围内的场景,第一场景中的物体可以例如但不限于床、或衣柜等。在第一设备处于厨房的情况下,第一场景可以是第一设备扫描到的厨房范围内的场景,第一场景中的物体可以例如但不限于冰箱、酒杯、或盘子等。在第一设备处于运输车道的情况下,第一场景可以是第一设备扫描到的车道场景,第一场景中的物体可以例如但不限于车辆、或轨道等。在第一设备处于监控森林火险的情况下,第一场景可以是第一设备扫描到的森林场景,第一场景中的物体可以例如但不限于树木、或障碍物等。

[0101] 示例性的,第一设备可以包括图像传感器,图像传感器可以例如但不限于图1中的示例。由图像传感器采集二维图像。

[0102] S302a、第一设备向三维物体检测装置发送二维图像。相应的,三维物体检测装置接收来自第一设备的二维图像。

[0103] 示例性的,在第一设备与第二设备中的三维物体检测装置通过有线连接方式连接的情况下,第一设备通过总线向第二设备中的三维物体检测装置发送二维图像。相应的,第二设备中的三维物体检测装置通过总线接收来自第一设备的二维图像。其中,“总线”的介绍可以参见图1中的相关说明,此处不再赘述。在第一设备与第二设备中的三维物体检测装置通过无线通信技术通信的情况下,第一设备通过互联网向第二设备中的三维物体检测装置发送二维图像。相应的,第二设备中的三维物体检测装置通过互联网接收来自第一设备的二维图像。其中,“无线通信技术”的介绍可以参见图1中的相关说明,此处不再赘述。

[0104] S301b、第一设备采集深度图像。

[0105] 其中,深度图像包括第一场景中至少一个物体的深度值构成的图像。“第一场景”和“物体”的介绍可以参见S301a的相关说明,此处不再赘述。

[0106] 示例性的,第一设备可以包括深度传感器,深度传感器可以例如但不限于图1中的示例。由深度传感器采集深度图像。

[0107] S302b、第一设备向三维物体检测装置发送深度图像。相应的,三维物体检测装置接收来自第一设备的深度图像。

[0108] 示例性的,在第一设备与第二设备中的三维物体检测装置通过有线连接方式连接的情况下,第一设备通过总线向第二设备中的三维物体检测装置发送深度图像。相应的,第

二设备中的三维物体检测装置通过总线接收来自第一设备的深度图像。其中，“总线”的介绍可以参见图1中的相关说明，此处不再赘述在第一设备与第二设备中的三维物体检测装置通过无线通信技术通信的情况下，第一设备通过互联网向第二设备中的三维物体检测装置发送深度图像。相应的，第二设备中的三维物体检测装置通过互联网接收来自第一设备的深度图像。其中，“无线通信技术”的介绍可以参见图1中的相关说明，此处不再赘述。

[0109] S303b、三维物体检测装置将深度图像中的像素点反投影到深度传感器的坐标系，以得到3D空间中的点云数据。

[0110] 示例性的，三维物体检测装置利用深度传感器的内参，将深度图像的像素点坐标 (u', v') 反投影到深度传感器坐标系，以获得3D空间中的点云数据。其中，3D空间中点云数据与该深度图像的像素点坐标之间满足如下公式：

$$[0111] \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = K_1^{-1} * \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} * z \quad \text{公式 (1)}$$

[0112] 其中， u' 表示像素点在深度图像中的横坐标， v' 表示像素点在深度图像中的纵坐标。 x 表示像素点在深度传感器坐标系下x轴的坐标(或3D空间中点云数据在x轴的坐标)， y 表示像素点在深度传感器坐标系下y轴的坐标(或3D空间中点云数据在y轴的坐标)， z 表示像素点在深度传感器坐标系下z轴的坐标(或3D空间中点云数据在z轴的坐标)。 K_1^{-1} 表示深度传感器的内参的逆矩阵。

[0113] 需要说明的是，深度传感器的内参 K_1 和图像传感器的内参 K_2 是预先标定的参数。标定过程可以例如但不限于如下介绍：

[0114] 首先，三维物体检测装置获取多组不同角度的棋盘格图像。

[0115] 其中，上述“多组不同角度的棋盘格图像”中每组棋盘格图像可以包括二维图像和深度图像，且是图像传感器和深度传感器在同一时刻采集的图像。棋盘格是黑白相间的A4纸大小的棋盘格，方格分布可以是10行8列，如图4a所示。在图4a中，斜线填充的方格表示黑方格，无斜线填充的方格表示白方格。

[0116] 然后，三维物体检测装置通过高斯-牛顿法对棋盘格图像中棋盘格对角点的坐标进行计算，以得到相机内参，即深度传感器的内参 K_1 和图像传感器的内参 K_2 。

[0117] 另外，三维物体检测装置还能够根据二维图像和3D空间中的点云数据确定相机外参。

[0118] 其中，相机外参包括旋转矩阵和平移矩阵。示例性的，参见图4b，对“三维物体检测装置确定相机外参”的过程进行介绍：

[0119] S3041、三维物体检测装置将3D空间中的点云数据变换到图像传感器的坐标系，以得到第一坐标。

[0120] 其中，第一坐标是指点云数据在图像传感器的坐标系下的坐标。

[0121] 示例性的，第一坐标与3D空间中的点云数据之间满足如下公式：

$$[0122] \quad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = r * \begin{bmatrix} x \\ y \\ z \end{bmatrix} + t \quad \text{公式 (2)}$$

[0123] 其中， x 表示3D空间中点云数据在x轴的坐标， y 表示3D空间中点云数据在y轴的坐标， z 表示3D空间中点云数据在z轴的坐标。 x' 表示3D空间中点云数据所表示的点在图像传

传感器的坐标系中x轴的坐标, y' 表示3D空间中点云数据所表示的点在图像传感器的坐标系中y轴的坐标, z' 表示3D空间中点云数据所表示的点在图像传感器的坐标系中z轴的坐标。 r 表示3*3的旋转矩阵, t 表示3*1的平移矩阵。

[0124] S3042、三维物体检测装置将第一坐标变换到二维图像坐标系,以得到第二坐标。

[0125] 其中,第二坐标是3D空间中点云数据(即S303b所确定的点云数据)在二维图像坐标系下的坐标。

[0126] 示例性的,第一坐标与第二坐标满足如下公式:

$$[0127] \quad \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_2 * \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad \text{公式(3)}$$

[0128] 其中, x' 表示3D空间中点云数据所表示的点在图像传感器的坐标系中x轴的坐标, y' 表示3D空间中点云数据所表示的点在图像传感器的坐标系中y轴的坐标, z' 表示3D空间中点云数据所表示的点在图像传感器的坐标系中z轴的坐标。 u 表示3D空间中点云数据所表示的点在二维图像坐标系中的横坐标, v 表示3D空间中点云数据所表示的点在二维图像坐标系中的纵坐标。 K_2 表示图像传感器的内参。

[0129] S3043、三维物体检测装置根据深度图像中的像素点坐标和第二坐标,确定相机外参。

[0130] 示例性的,三维物体检测装置确定深度图像中像素点坐标 (u, v) 和第二坐标 (u', v') 之间的误差 $(u-u', v-v')$,基于误差,来调整旋转矩阵 r 和平移矩阵 t 。三维物体检测装置通过重复执行上述S3041至S3043,以确定最小误差对应的旋转矩阵 R 和平移矩阵 T 。

[0131] 第二阶段,检测阶段。参见图5,在此阶段中,三维物体检测装置对二维图像和深度图像对应的点云数据进行检测,以确定目标物体在三维空间中的估计位置。其中,目标物体是至少一个物体中的一个物体。第二阶段的具体步骤介绍如下:

[0132] 首先,对“二维图像”的处理过程进行说明:

[0133] S501a、三维物体检测装置对二维图像进行检测,以得到二维图像的检测结果。

[0134] 其中,二维图像的检测结果至少包括目标物体图像。目标物体图像是至少一个物体中目标物体的图像。

[0135] 示例性的,S501a的实现过程如下:三维物体检测装置将二维图像输入2D物体检测模型,采用2D物体检测模型对二维图像进行检测,以获取二维图像的检测结果。其中,2D物体检测模型可以例如但不限于如下介绍:SSD(single shot multibox detector)模型、DSSD(deconvolution single shot multibox detector)模型、YoloV4或其他自研模型等。示例性的,2D物体检测模型可以是预先训练的模型。参见图6a,模型训练阶段的步骤说明如下:

[0136] 步骤a1、图像数据标注。在此步骤中,对预先获取的样本集中的二维图像进行标注。

[0137] 步骤a2、数据增强。在此步骤中,对标注后的二维图像进行数据增强处理,如亮度变换,以得到数据增强后的图像。

[0138] 步骤a3、输入神经网络。在此步骤中,将数据增强后的图像输入神经网络,如卷积神经网络。

[0139] 步骤a4、计算损失函数。在此步骤中,采用卷积神经网络计算输入的数据增强后的图像与标注的信息之间的特征向量,此过程称为“计算损失函数”。

[0140] 步骤a5、保存训练权重。在此步骤中,经过上述训练过程,三维物体检测装置保存卷积神经网络计算得到的权重。

[0141] 如此,经过步骤a1至步骤a5,三维物体检测装置即可得到2D物体检测模型。

[0142] 在模型应用阶段,即S502a的实现过程。参见图6b,模型应用阶段的步骤说明如下:

[0143] 步骤a6、确定二维图像。在此步骤中,三维物体检测装置确定待处理的二维图像,即S501a

[0144] 中获取的二维图像。

[0145] 步骤a7、加载训练权重和网络模型。在此步骤中,三维物体检测装置加载训练权重和网络模型,以构建2D物体检测模型,并将二维图像输入2D物体检测模型。

[0146] 步骤a8、前向传播。在此步骤中,采用2D物体检测模型对输入的二维图像进行计算,此过程即可描述为“前向传播”。

[0147] 步骤a9、预测2D检测框。在此步骤中,三维物体检测装置采用2D物体检测模型对二维图像进行检测,以得到目标物体图像。示例性的,三维物体检测装置采用2D检测框标识出目标物体图像。其中,2D检测框可以是一个矩形框,包括左上角的像素点坐标(x,y)、宽度参数和高度参数。示例性的,目标物体的数量记为N, $N \geq 1$ 。第i个目标物体的二维图像的检测结果记为 $DR = \{0i\}$ 。其中,DR表示二维图像的检测结果,0i表示第i个目标物体的2D检测框参数。 $1 \leq i \leq N$ 。示例性的,参加图6c,图6c中示出了两种目标物体,2D检测框分别标识出了人的图像、椅子的图像,如图6c中的粗实线框所示。

[0148] 可选的,二维图像的检测结果还包括以下至少一项:

[0149] 第一、目标物体的类别。其中,目标物体的类别可以是例如但不限于人、桌子、椅子等。

[0150] 第二、置信度。其中,置信度指示二维图像的检测结果的可信程度。置信度的取值不大于1。置信度的取值越高,表征二维图像的检测结果的可信程度越高。示例性的,在置信度大于置信度阈值的情况下,三维物体检测装置执行S502。反之,在置信度小于或等于置信度阈值的情况下,三维物体检测装置重新执行步骤a8和步骤a9,直至置信度超过置信度阈值,或对步骤a6的二维图像重复执行次数达到第一预设值。由于三维物体检测结果是基于目标物体图像确定的,且目标物体图像是满足置信度要求的图像,所以,本申请实施例三维物体检测方法能够准确地筛选出目标点云数据集,有助于提升三维物体检测结果的准确性。

[0151] 然后,再对“深度图像对应的点云数据”的处理过程进行说明:

[0152] S501b、三维物体检测装置对深度图像对应的点云数据进行聚类,以得到至少一个点云数据集。

[0153] 其中,一个“点云数据集中的点云数据”是上述“S303b得到的点云数据”中的一部分。点云数据集中的点云数据用于描述第一场中物体的候选区域。其中,一个“点云数据集”所表示的点也可以描述为一个“点云簇”。

[0154] 示例性的,如图7a所示,S502b的实现过程可以例如如下介绍:

[0155] 步骤b1、滤波。在此步骤中,三维物体检测装置对“S303b得到的点云数据”进行降

采样处理,以提高计算效率。

[0156] 步骤b2、法线估计。在此步骤中,三维物体检测装置对“步骤b1降采样处理后的点云数据”进行法线估计,以确定表面法线。

[0157] 示例性的,参见图7b,以一个采样点 P_i 为例,从“步骤b1降采样处理后的点云数据”所表示的点中,确定满足第一预设条件的点。例如,第一预设条件可以实现为:半径为3厘米的圆形区域内的点。以“K个满足第一预设条件的点”为例,根据K个点的坐标创建协方差矩阵C。然后,分解出协方差矩阵C的特征值和特征向量。其中,协方差矩阵C满足如下公式:

$$[0158] \quad C = \frac{1}{K} \sum_{i=1}^K (P_i - \bar{P}) * (P_i - \bar{P})^T, C * \vec{v}_j = \lambda_i * \vec{v}_j, j \in \{0,1,2\} \quad \text{公式(4)}$$

[0159] 其中,C表示协方差矩阵,K表示满足第一预设条件的点的数量, P_i 表示K个点中的第i个采样点, \bar{P} 表示K个点的坐标的平均值, λ_i 是协方差矩阵C的第i个特征值, \vec{v}_j 是第j个特征向量。将特征值最小且与深度传感器感知方向相同的特征向量,作为法线。

[0160] 步骤b3、平面检测。首先,基于法线方向进行聚类,即将满足欧式距离约束的法线进行聚类,找出法线方向接近的点构成的点云数据集S。然后,再基于空间位置进行聚类,即对点云数据集S中的点进行聚类,找出满足欧式距离的点。最后,将满足欧式距离的点带入平面方程中计算形如 $AX=B$ 的最小二乘解。其中,平面方程 $AX=B$ 满足如下公式:

$$[0161] \quad \left\{ \begin{array}{l} \begin{bmatrix} \frac{x_1}{z_1} & \frac{y_1}{z_1} & 1 \\ \vdots & \vdots & \vdots \\ \frac{x_m}{z_m} & \frac{y_m}{z_m} & 1 \end{bmatrix} * \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1/z_1 \\ \vdots \\ 1/z_m \end{bmatrix} \\ A = \begin{bmatrix} \frac{x_1}{z_1} & \frac{y_1}{z_1} & 1 \\ \vdots & \vdots & \vdots \\ \frac{x_m}{z_m} & \frac{y_m}{z_m} & 1 \end{bmatrix} \\ X = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ B = \begin{bmatrix} 1/z_1 \\ \vdots \\ 1/z_m \end{bmatrix} \end{array} \right. \quad \text{公式(5)}$$

[0162] 其中, x_1 表示“满足欧式距离的点”中第一个点在深度传感器坐标系中x轴的坐标, y_1 表示“满足欧式距离的点”中第一个点在深度传感器坐标系中y轴的坐标, z_1 表示“满足欧式距离的点”中第一个点在深度传感器坐标系中z轴的坐标。 x_m 表示“满足欧式距离的点”中第m个点在深度传感器坐标系中x轴的坐标, y_m 表示“满足欧式距离的点”中第m个点在深度传感器坐标系中y轴的坐标, z_m 表示“满足欧式距离的点”中第m个点在深度传感器坐标系中z轴的坐标。解析解为 $X = (AA^T)^{-1}A^TB$,即为所求法向量。 a 表示法向量中x轴的坐标, b 表示法向量中y轴的坐标, c 表示法向量中z轴的坐标。如此,三维物体检测装置即求解出 a 、 b 、 c 取值,从而得到拟合平面。

[0163] 步骤b4、欧式聚类。

[0164] 首先,确定步骤b3拟合平面中点的数量,剔除点的数量大于数量阈值的拟合平面。由于深度图像中包括大量的背景图像,如地面的图像,所以,深度图像中包括大量的背景物体的像素点。若某一拟合平面中点的数量大于数量阈值,则该拟合平面属于背景物体的图

像区域的概率较大,相应的,该拟合平面属于目标物体的图像区域的概率较小,需要剔除,以提高计算效率。

[0165] 然后,对剩余的拟合平面中的点进行聚类,将满足欧式距离条件的点的坐标构成点云数据集,作为深度图像的点云数据集。其中,欧式距离条件可以例如但不限于如下介绍:拟合平面中两点之间的欧式距离小于距离阈值。距离阈值可以是2CM或者其他值,可以根据调试经验、或者实验测试确定。

[0166] 示例性的,深度图像包括N个物体的图像,深度图像对应的点云数据集记为 $S = \{C_i\}$ 。其中, C_i 表示第i个物体的点云数据集。

[0167] 如此,通过上述步骤b1至步骤b4,三维物体检测装置即可获取到深度图像的至少一个点云数据集。

[0168] 需要说明的是,三维物体检测装置可以先执行二维图像的处理步骤(即S501a),再执行点云数据的处理步骤(即S501b),也可以先执行点云数据的处理步骤,再执行二维图像的处理步骤,还可以同时执行二维图像的处理步骤和点云数据的处理步骤,本申请实施例对此不作限定。

[0169] 最后,再对目标物体图像和点云数据集的处理过程进行说明:

[0170] S502、三维物体检测装置根据目标物体图像,从至少一个点云数据集中确定目标点云数据集。

[0171] 其中,目标点云数据集中的点云数据用于描述第一场景中存在目标物体的估计区域。在本申请实施例中,“目标点云数据集”所表示的点也可以描述为“目标点云簇”。

[0172] 示例性的,参见图8,将至少一个点云数据集中的一个点云数据集描述为“第一点云数据集”。其中,“第一点云数据集”所表示的点也可以描述为“第一点云簇”。以第一点云数据集为例,在“第一点云数据集投影到二维图像”的情况下,对“目标点云数据集的确定过程”进行介绍:

[0173] S5021、三维物体检测装置确定第一点云数据集在二维图像中的第一投影区域。

[0174] 其中,二维图像是S501a获取的图像。示例性的,S5021的实现过程如下:

[0175] 步骤1、三维物体检测装置根据第一点云数据集中点云的深度范围,从第一点云数据集所表示的特征点中确定第一特征点。

[0176] 示例性的,第一特征点可以是以下至少一项:第一点云数据集所表示的特征点中的最远点、第一点云数据集所表示的特征点中的最近点。

[0177] 示例性的,第一点云数据集记为点云数据集 C_i 。三维物体检测装置查找点云数据集 C_i 中的最远点 P_{max} 和最近点 P_{min} ,以作为第一特征点。

[0178] 步骤2、三维物体检测装置根据点云数据与二维图像之间的转换参数,确定第一特征点在二维图像中的第一投影点。

[0179] 示例性的,点云数据与二维图像之间的转换参数可以是以下至少一项:深度传感器的内参 K_1 、旋转矩阵 R 、平移矩阵 T 。

[0180] 示例性的,以最远点 P_{max} 为例,先利用公式(6)确定最远点 P_{max} 在图像传感器的坐标系下的坐标。

$$[0181] \quad \begin{bmatrix} x'_{max} \\ y'_{max} \\ z'_{max} \end{bmatrix} = R * \begin{bmatrix} x_{max} \\ y_{max} \\ z_{max} \end{bmatrix} + T \quad \text{公式 (6)}$$

[0182] 其中, x_{max} 表示3D空间中最远点Pmax在x轴的坐标, y_{max} 表示3D空间中最远点Pmax在y轴的坐标, z_{max} 表示3D空间中最远点Pmax在z轴的坐标。 x'_{max} 表示最远点Pmax在图像传感器的坐标系中x轴的坐标, y'_{max} 表示最远点Pmax在图像传感器的坐标系中y轴的坐标, z'_{max} 表示最远点Pmax在图像传感器的坐标系中z轴的坐标。R表示3*3的旋转矩阵, T表示3*1的平移矩阵。

[0183] 再利用公式(7)确定最远点Pmax在二维图像的坐标系下的坐标。

$$[0184] \quad \begin{bmatrix} u_{max} \\ v_{max} \\ 1 \end{bmatrix} = K_2 * \begin{bmatrix} x'_{max} \\ y'_{max} \\ z'_{max} \end{bmatrix} \quad \text{公式 (7)}$$

[0185] 其中, x'_{max} 表示最远点Pmax在图像传感器的坐标系中x轴的坐标, y'_{max} 表示最远点Pmax在图像传感器的坐标系中y轴的坐标, z'_{max} 表示最远点Pmax在图像传感器的坐标系中z轴的坐标。 u_{max} 表示最远点Pmax在二维图像坐标系中的横坐标, v_{max} 表示最远点Pmax在二维图像坐标系中的纵坐标。 K_2 表示图像传感器的内参。

[0186] 以最近点Pmin为例,先利用公式(8)确定最近点Pmin在图像传感器的坐标系下的坐标。

$$[0187] \quad \begin{bmatrix} x'_{min} \\ y'_{min} \\ z'_{min} \end{bmatrix} = R * \begin{bmatrix} x_{min} \\ y_{min} \\ z_{min} \end{bmatrix} + T \quad \text{公式 (8)}$$

[0188] 其中, x_{min} 表示3D空间中最近点Pmin在x轴的坐标, y_{min} 表示3D空间中最近点Pmin在y轴的坐标, z_{min} 表示3D空间中最近点Pmin在z轴的坐标。 x'_{min} 表示最近点Pmin在图像传感器的坐标系中x轴的坐标, y'_{min} 表示最近点Pmin在图像传感器的坐标系中y轴的坐标, z'_{min} 表示最近点Pmin在图像传感器的坐标系中z轴的坐标。R表示3*3的旋转矩阵, T表示3*1的平移矩阵。

[0189] 再利用公式(9)确定最近点Pmin在二维图像的坐标系下的坐标。

$$[0190] \quad \begin{bmatrix} u_{min} \\ v_{min} \\ 1 \end{bmatrix} = K_2 * \begin{bmatrix} x'_{min} \\ y'_{min} \\ z'_{min} \end{bmatrix} \quad \text{公式 (9)}$$

[0191] 其中, x'_{min} 表示最近点Pmin在图像传感器的坐标系中x轴的坐标, y'_{min} 表示最近点Pmin在图像传感器的坐标系中y轴的坐标, z'_{min} 表示最近点Pmin在图像传感器的坐标系中z轴的坐标。 u_{min} 表示最近点Pmin在二维图像坐标系中的横坐标, v_{min} 表示最近点Pmin在二维图像坐标系中的纵坐标。 K_2 表示图像传感器的内参。

[0192] 步骤3、三维物体检测装置将第一投影点对应的二维标注框所标注的区域,作为第一投影区域。

[0193] 也就是说,二维标注框在二维图像上标注的区域,即为第一投影区域。

[0194] 示例性的,二维标注框可以是矩形框,如图9b所示。二维标注框可以是以第一投影点为对角点的标注框。

[0195] 如此,三维物体检测装置即可确定第一点云数据集在二维图像中的第一投影区域,进而判断该第一点云数据集是否为目标点云数据集。

[0196] S5022、三维物体检测装置确定目标物体图像在二维图像中的目标图像区域。

[0197] 示例性的,目标图像区域可以是S501a中2D检测框参数指示的区域,具体参见S501a中的介绍,此处不再赘述。

[0198] S5023、三维物体检测装置根据第一投影区域和目标图像区域,确定第一点云数据集为目标点云数据集。

[0199] 其中,S5023的实现方式有多种,可以例如但不限于如下介绍:

[0200] 三维物体检测装置根据第一投影区域和目标图像区域之间的重合程度,以及第一投影区域的大小,确定第一点云数据集为目标点云数据集。

[0201] 也就是说,在确定“第一点云数据集是否为目标点云数据集”的情况下,三维物体检测装置除了考量“第一投影区域和目标图像区域之间的重合程度”之外,还参考了“第一投影区域的大小”这一指标。若“第一投影区域”面积偏小,则第一点云数据集中点云数据所表示的特征点可能是目标物体的一部分。例如,在目标物体是“椅子”的情况下,第一点云数据集中点云数据所表示的特征点可能属于“椅背”部分,也可能属于“扶手”部分。此种情况下,第一投影区域与目标图像区域之间仍重合,但目标物体的一部分无法准确地表征目标物体整体在三维空间中的估计位置,所以,此类的点云数据集不作为目标点云数据集。在考量了上述两个指标的情况下,有助于提升目标点云数据集筛选的精准度。

[0202] 示例性的,通过两个示例对S5023的实现过程进行说明:

[0203] 示例一、目标点云数据集所表示的特征点在二维图像中的目标投影区域满足:

$$[0204] \begin{cases} S_s = IOU_s * L_{j1} / D_{ij1} \\ IOU_s = S_{\cap} / S_{\cup} \\ S_s > T_s \end{cases} \quad \text{公式 (10)}$$

[0205] 其中, S_s 表示目标投影区域与目标图像区域之间的相似度。 IOU_s 表示目标投影区域与目标图像区域之间的交并比。 S_{\cap} 表示目标投影区域与目标图像区域的交集的面积(或重叠面积), S_{\cup} 表示目标投影区域与目标图像区域的并集的面积(或重叠面积与不重叠面积之和)。 L_{j1} 表示目标投影区域的投影点间距,投影点间距是目标特征点在二维图像中的投影点之间的距离。目标特征点属于目标点云数据集所表示的特征点,且指示目标点云数据集所表示的特征点的深度范围。 D_{ij1} 表示目标投影区域的参考点与目标图像区域的参考点之间的距离。其中,参考点可以是中心点、左上角的顶点、侧边的中心点等。例如,目标投影区域的参考点可以是目标投影区域的中心点、左上角的顶点、左侧边的中心点等。类似的,目标图像区域的参考点也可以是目标图像区域的中心点、左上角的顶点、左侧边的中心点等。 T_s 表示相似度阈值。

[0206] 以图9c为例,目标投影区域记为 R_i ,目标图像区域记为 O_i 。两者之间的重叠区域如图9c中斜线填充的区域所示,两者之间的未重叠区域如图9c中无斜线填充的区域所示。 S_{\cap} 表示上述重叠面积, S_{\cup} 表示上述重叠面积与不重叠面积之和。 L_{j1} 表示目标投影区域的投影点间距,如图9c中 R_i 的对角线所示。 D_{ij1} 表示目标投影区域的中心点与目标图像区域的中心点之间的距离,如图9c中的粗实线所示。如此,三维物体检测装置确定第一点云数据集是否满足上述公式(10),若满足,则将第一点云数据集作为目标点云数据集,反之,若不满足,

则第一点云数据集不是目标点云数据集。

[0207] 示例二、三维物体检测装置在确定公式(10)中的 IOU_s 大于第二预设值(如0.5)的情况下,三维物体检测装置再结合公式(10),判断第一点云数据集是否为目标点云数据集,具体详见“S5023的示例一”的相关描述,此处不再赘述。

[0208] S503、三维物体检测装置对目标点云数据集和目标物体图像进行联合,以得到目标物体的检测结果。

[0209] 其中,目标物体的检测结果指示目标物体在三维空间中的估计位置。

[0210] 示例性的,S503的实现步骤如下步骤1和步骤2:

[0211] 步骤1、三维物体检测装置根据目标点云数据集中点云的深度范围,将目标物体图像中的部分像素点逆映射至三维空间,以得到目标逆映射点。

[0212] 其中,目标物体图像中的部分像素点可以是目标物体图像的对角点。以图9c为例,将2D检测框 O_i 的对角点 $P_{i1}(u_1, v_1)$,反投影到3D空间,以得到 $PP_{i1}(x_1, y_1, z_1)$ 。其中, PP_{i1} 和 P_{i1} 之间的坐标满足如下公式:

$$[0213] \quad \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = K_2^{-1} * \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} * z_{min_i} \quad \text{公式(11)}$$

[0214] 其中, z_{min_i} 表示目标点云数据集的深度范围的最小值, u_1 表示对角点 P_{i1} 的横坐标, v_1 表示对角点 P_{i1} 的纵坐标, K_2^{-1} 表示图像传感器的内参的逆矩阵, x_1 表示 PP_{i1} 在x轴的坐标, y_1 表示 PP_{i1} 在y轴的坐标, z_1 表示 PP_{i1} 在z轴的坐标。

[0215] 将2D检测框 O_i 的对角点 $P_{i2}(u_2, v_2)$,反投影到3D空间,以得到 $PP_{i2}(x_2, y_2, z_2)$ 。其中, PP_{i2} 和 P_{i2} 之间的坐标满足如下公式:

$$[0216] \quad \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = K_2^{-1} * \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} * z_{max_i} \quad \text{公式(12)}$$

[0217] 其中, z_{max_i} 表示目标点云数据集的深度范围的最大值, u_2 表示对角点 P_{i2} 的横坐标, v_2 表示对角点 P_{i2} 的纵坐标, K_2^{-1} 表示图像传感器的内参的逆矩阵, x_2 表示 PP_{i2} 在x轴的坐标, y_2 表示 PP_{i2} 在y轴的坐标, z_2 表示 PP_{i2} 在z轴的坐标。在图9c所示的场景下,目标逆映射点为 PP_{i1} 和 PP_{i2} 。

[0218] 步骤2、三维物体检测装置将目标逆映射点对应的三维标注框所标注的区域,作为检测结果,即目标物体的检测结果,以指示目标物体在三维空间中的估计位置。

[0219] 示例性的,以图9a为例,三维标注框的对角点为 PP_{i1} 和 PP_{i2} ,三维标注框所标注的区域即为目标物体在三维空间中的估计位置。

[0220] 可选的,在一些实施例中,三维物体检测装置还执行S504:

[0221] S504、三维物体检测装置根据预设的调整因子,调整目标物体在三维空间中的估计位置。

[0222] 其中,调整因子指示目标物体在第一场景中的真实位置与估计位置之间的差异。示例性的,基于大量的数据统计,“物体在三维空间中的真实位置”通常小于“物体在三维空间中的估计位置”,调整因子可以是小于1的一个系数。标识三维物体检测结果的3D标注框的各个顶点坐标均乘以调整因子,以得到调整后的三维空间估计位置,以贴合目标物体的实际位置。

[0223] 参见图10,本申请实施例提供的第二种三维物体检测方法包括如下步骤:

[0224] S1001、三维物体检测装置获取二维图像和至少一个点云数据集。

[0225] 其中,二维图像是图像传感器采集的信息。二维图像包括至少一个物体的图像。关于“二维图像”的介绍可以详见S301a和S302a的相关描述。

[0226] 其中,点云数据是深度传感器采集的信息。点云数据集包括多个点云数据,点云数据用于描述至少一个物体在三维空间中的候选区域。关于“点云数据集”的介绍可以详见S501b、S301b、S302b和S303b的相关描述。

[0227] S1002、三维物体检测装置根据二维图像中目标物体图像,从至少一个点云数据集中确定目标点云数据集。

[0228] 其中,目标物体图像包括至少一个物体中目标物体的图像,具体可以参见S501a的相关描述。目标点云数据集中的点云数据用于描述目标物体在三维空间中的候选区域。

[0229] 示例性的,将“至少一个点云数据集”中的一个集合描述为“第一点云数据集”,以第一点云数据集为例,S1002的实现过程可以参见S5021至S5023的相关描述。其中,目标点云数据集所表示的特征点在二维图像中的目标投影区域满足:

$$[0230] \quad \begin{cases} S = IOU * L_j / D_{ij} \\ IOU = S_{\cap} / S_{\cup} \\ S > T \end{cases} \quad \text{公式 (13)}$$

[0231] 其中,S表示目标投影区域与目标图像区域之间的相似度。IOU表示目标投影区域与目标图像区域之间的交并比(intersection-over-union, IOU)。S_∩表示目标投影区域与目标图像区域之间的重叠面积。S_∪表示重叠面积与不重叠面积之和,不重叠面积为目标投影区域与目标图像区域之间未重叠的面积。L_j表示目标投影区域的投影点间距,投影点间距是目标特征点在二维图像中的投影点之间的距离,目标特征点属于目标点云数据集所表示的特征点,且基于目标点云数据集的深度范围的端值确定。D_{ij}表示目标投影区域的参考点与目标图像区域的参考点之间的距离,T表示相似度阈值。在三维物体检测装置执行S5023的情况下,上述公式(13)可以实现为公式(10)。

[0232] S1003、三维物体检测装置将目标点云数据集和目标物体图像进行关联,以得到检测结果。

[0233] 其中,检测结果指示目标物体在三维空间中的估计位置。示例性的,在S1003具体实现为S503时,S1003中的“检测结果”实现为S503中的“目标物体的检测结果”,详见S503的相关描述。

[0234] 由于二维图像的处理精度高,目标物体图像能够准确地呈现目标物体在二维图像中的区域,利用目标物体图像来筛选目标点云数据集,实现对点云数据集几何分割聚类,无需获取大量的三维训练数据。即使物体被遮挡,也能够获取到目标点云数据集,在一定程度上提高了目标物体对应的目标点云数据集的精准度。并且,三维物体检测装置将目标点云数据集和目标物体图像进行关联,以得到检测结果。由于二维图像的处理精度高,即使目标物体的点云数据不足,也能够精准地确定目标物体在三维空间中的估计位置,避免假阳率高的问题。本申请实施例三维物体检测方法无需获取三维训练数据,避免了“基于三维训练数据训练模型”所导致的“泛化性差”的问题。

[0235] 在一些实施例中,三维物体检测装置还执行S1004:

[0236] S1004、三维物体检测装置根据预设的调整因子,调整检测结果指示的估计位置,以使三维物体检测装置所确定的估计位置更贴合实际物体尺寸,从而提高物体检测的精准度。

[0237] 其中,调整因子指示目标物体在三维空间中的真实位置与估计位置之间的差异,具体参见S504的相关说明。示例性的,在S1004具体实现为S504时,S1004中的“检测结果”实现为S504中的“目标物体在三维空间中的估计位置”,详见S504的的相关描述。

[0238] 上述主要从方法的角度对本申请实施例提供的方案进行了介绍。下面将结合附图,描述根据本申请所提供的三维物体检测装置1020和第二设备102。

[0239] 参见图1所示的系统架构图中三维物体检测装置1020的结构示意图,如图1所示,该三维物体检测装置1020包括:获取单元1121和处理单元1122。

[0240] 获取单元1121,用于获取二维图像和至少一个点云数据集。其中,二维图像是图像传感器采集的信息,二维图像包括至少一个物体的图像。点云数据是深度传感器采集的信息,点云数据集包括多个点云数据,点云数据用于描述至少一个物体在三维空间中的候选区域。

[0241] 处理单元1122,用于根据二维图像中目标物体图像,从至少一个点云数据集中确定目标点云数据集。其中,目标物体图像包括至少一个物体中目标物体的图像。目标点云数据集中的点云数据用于描述目标物体在三维空间中的候选区域。

[0242] 处理单元1122,还用于将目标点云数据集和目标物体图像进行关联,以得到检测结果。其中,检测结果指示目标物体在三维空间中的估计位置。

[0243] 其中,获取单元1121的具体实现可以参见图3所示实施例中S302a、S302b和S303b相关内容描述,处理单元1122的具体实现可以参见图5所示实施例中S501a、S501b、S502、S503相关内容描述,在此不再赘述。

[0244] 在一种可能的设计中,处理单元1122,用于根据二维图像中目标物体图像,从至少一个点云数据集中确定目标点云数据集时,具体包括:处理单元1122用于确定第一点云数据集在二维图像中的第一投影区域。其中,第一点云数据集是至少一个点云数据集中的集合。处理单元1122用于根据第一投影区域和目标图像区域,确定第一点云数据集为目标点云数据集。其中,所示目标图像区域是目标物体图像在二维图像中的区域。

[0245] 其中,处理单元1122的具体实现可以参见图8所示实施例中S5021、S5022和S5023相关内容描述,在此不再赘述。

[0246] 在一种可能的设计中,处理单元1122,用于确定第一点云数据集在二维图像中的第一投影区域,具体包括:处理单元1122用于根据第一点云数据集中点云的深度范围,从第一点云数据集所表示的特征点中确定第一特征点。处理单元1122用于根据点云数据与二维图像之间的转换参数,确定第一特征点在二维图像中的第一投影点。处理单元1122用于将第一投影点对应的二维标注框所标注的区域,作为第一投影区域。

[0247] 其中,处理单元1122的具体实现可以参见S5021中的步骤1、步骤2和步骤3相关内容描述,在此不再赘述。

[0248] 在一种可能的设计中,处理单元1122,用于根据第一投影区域和目标图像区域,确定第一点云数据集为目标点云数据集,具体包括:处理单元1122用于根据第一投影区域和目标图像区域之间的重合程度,以及第一投影区域的大小,确定第一点云数据集为目标点

云数据集。

[0249] 其中,处理单元1122的具体实现可以参见S5023中的相关内容描述,在此不再赘述。

[0250] 在一种可能的设计中,处理单元1122,用于将目标点云数据集和目标物体图像进行关联,以得到检测结果,具体包括:处理单元1122用于根据目标点云数据集中点云的深度范围,将目标物体图像中的部分像素点逆映射至三维空间,以得到目标逆映射点。处理单元1122用于将目标逆映射点对应的三维标注框所标注的区域,作为检测结果。

[0251] 其中,处理单元1122的具体实现可以参见S503中步骤1和步骤2的相关内容描述,在此不再赘述。

[0252] 在一种可能的设计中,处理单元1122,还用于根据预设的调整因子,调整检测结果指示的估计位置。其中,调整因子指示目标物体在三维空间中的真实位置与估计位置之间的差异。

[0253] 其中,处理单元1122的具体实现可以参见图8中S504相关内容描述,在此不再赘述。

[0254] 根据本申请实施例的三维物体检测装置1020可对应于执行本申请实施例中描述的方法,并且,三维物体检测装置1020中的各个模块的上述和其它操作和/或功能分别为了实现图2、图3、图4b、图5、图6a、图6b、图7a、图8中的各个方法的相应流程,为了简洁,在此不再赘述。

[0255] 另外需说明的是,以上所描述的实施例仅仅是示意性的,其中所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块显示的部件可以是或者也可以不是物理模块,即可以位于一个地方,或者也可以分布到多个网络模块上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外,本申请提供的装置实施例附图中,模块之间的连接关系表示它们之间具有通信连接,具体可以实现为一条或多条通信总线或信号线。

[0256] 本申请实施例还提供了一种第二设备102,用于实现上述图1所示的系统架构图中三维物体检测装置1020的功能。其中,第二设备102可以是物理设备或物理设备集群,也可以是虚拟化的云设备,如云计算集群中的至少一个云计算设备。为了便于理解,本申请以第二设备102为独立的物理设备对该第二设备102的结构进行示例说明。

[0257] 图11提供了一种第二设备102的结构示意图,如图11所示,第二设备102包括总线1101、处理器1102、通信接口1103和存储器1104。处理器1102、存储器1104和通信接口1103之间通过总线1101通信。总线1101可以是外设部件互连标准(peripheral component interconnect, PCI)总线或扩展工业标准结构(extended industry standard architecture, EISA)总线等。总线可以分为地址总线、数据总线、控制总线等。为便于表示,图11中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。通信接口1103用于与外部通信。例如,获取二维图像和深度图像等。

[0258] 其中,处理器1102可以为中央处理器(central processing unit, CPU)。存储器1104可以包括易失性存储器(volatile memory),例如随机存取存储器(random access memory, RAM)。存储器1104还可以包括非易失性存储器(non-volatile memory),例如只读存储器(read-only memory, ROM),快闪存储器,硬盘驱动器(hard disk drive, HDD)或固态

盘(solid-state disk,SSD)。

[0259] 存储器1104中存储有可执行代码,处理器1102执行该可执行代码以执行前述三维物体检测方法。

[0260] 具体地,在实现图1所示实施例的情况下,且图1实施例中所描述的三维物体检测装置1020的各模块为通过软件实现的情况下,执行图1中的获取单元1121、处理单元1122功能所需的软件或程序代码存储在存储器1104中。处理器1102执行存储器1104中存储的各模块对应的程序代码,如获取单元1121和处理单元1122对应的程序代码,以提取出目标物体图像和目标点云数据集,进而得到目标物体的检测结果。如此,通过关联目标物体图像和目标点云数据集,从而实现三维物体检测。

[0261] 本申请实施例还提供了一种电子设备,该电子设备包括处理器和存储器。该处理器和该存储器进行相互的通信。该处理器用于执行该存储器中存储的指令,以使得电子设备执行上述三维物体检测方法。

[0262] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质包括指令,所述指令指示第二设备102执行上述应用于三维物体检测装置1020的三维物体检测方法。

[0263] 本申请实施例还提供了一种计算机程序产品,所述计算机程序产品被计算机执行时,所述计算机执行前述三维物体检测方法的任一方法。该计算机程序产品可以为一个软件安装包,在需要使用前述三维物体检测方法的任一方法的情况下,可以下载该计算机程序产品并在计算机上执行该计算机程序产品。

[0264] 本申请实施例还提供了一种芯片,包括逻辑电路和输入输出接口。其中,输入输出接口用于与芯片之外的模块通信,例如,该芯片可以为实现上述三维物体检测装置功能的芯片。输入输出接口输入二维图像和至少一个点云数据集,输入输出接口输出检测结果。逻辑电路用于运行计算机程序或指令,以实现上述三维物体检测方法。

[0265] 本申请实施例还提供了一种机器人,包括:图像传感器、深度传感器、处理器、用于存储处理器可执行指令的存储器。其中,图像传感器用于采集二维图像,深度传感器用于采集至少一个点云数据集,处理器被配置可执行指令,以实现如上述三维物体检测方法。

[0266] 本申请实施例还提供了一种服务器,包括:处理器、用于存储处理器可执行指令的存储器。其中,处理器被配置可执行指令,以实现如上述三维物体检测方法。

[0267] 通过以上的实施方式的描述,所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现,当然也可以通过专用硬件包括专用集成电路、专用CPU、专用存储器、专用元器件等来实现。一般情况下,凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现,而且,用来实现同一功能的具体硬件结构也可以是多种多样的,例如模拟电路、数字电路或专用电路等。但是,对本申请而言更多情况下软件程序实现是更佳实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在可读的存储介质中,如计算机的软盘、U盘、移动硬盘、ROM、RAM、磁碟或者光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,训练设备,或者网络设备等)执行本申请各个实施例所述的方法。

[0268] 在上述实施例中,可以全部或部分地通过软件、硬件、固件或者其任意组合来实

现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。

[0269] 所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一计算机可读存储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、训练设备或数据中心通过有线(例如同轴电缆、光纤、数字用户线(DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、训练设备或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的训练设备、数据中心等数据存储设备。所述可用介质可以是磁性介质,(例如,软盘、硬盘、磁带)、光介质(例如,DVD)、或者半导体介质等。

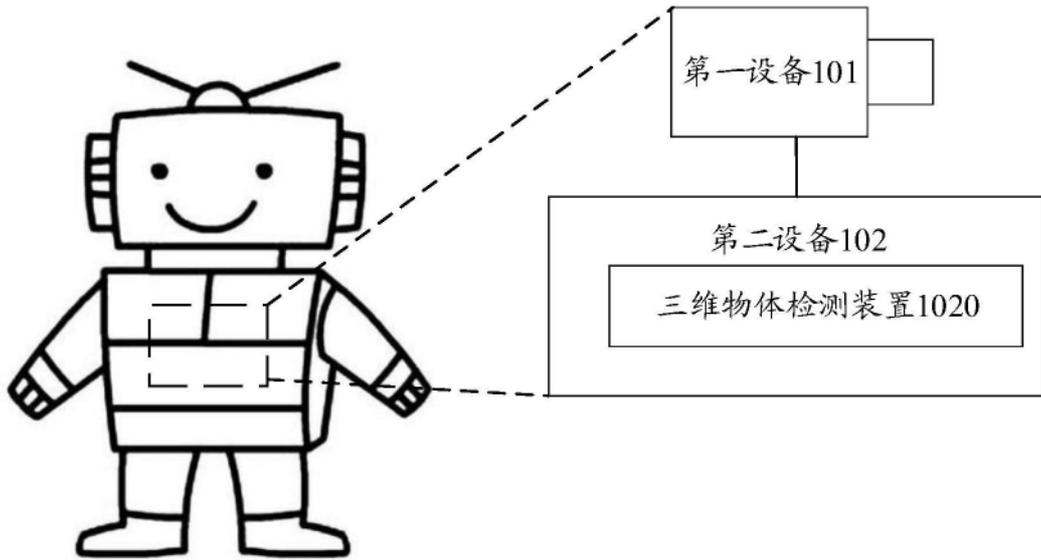


图1

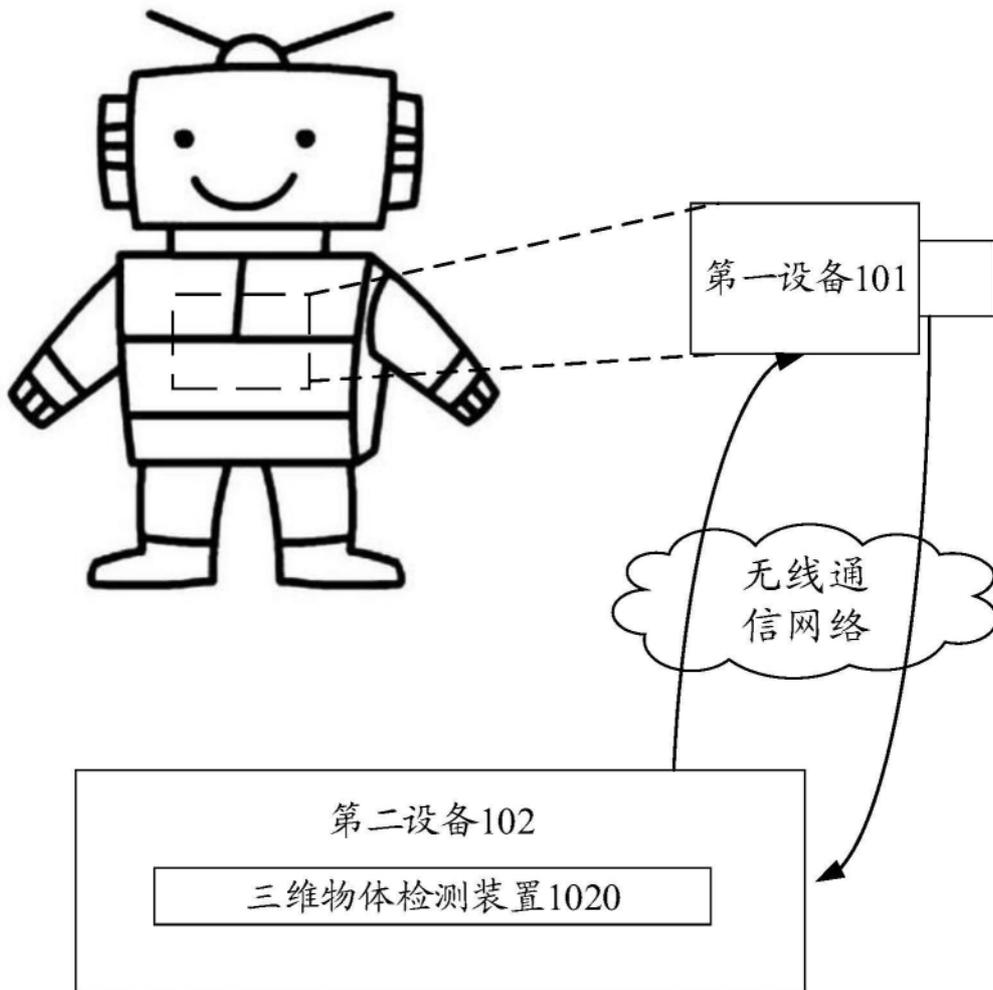


图2

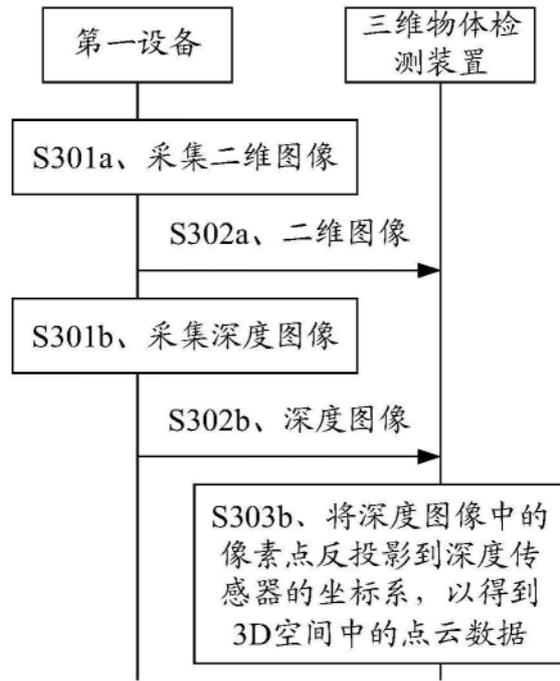


图3

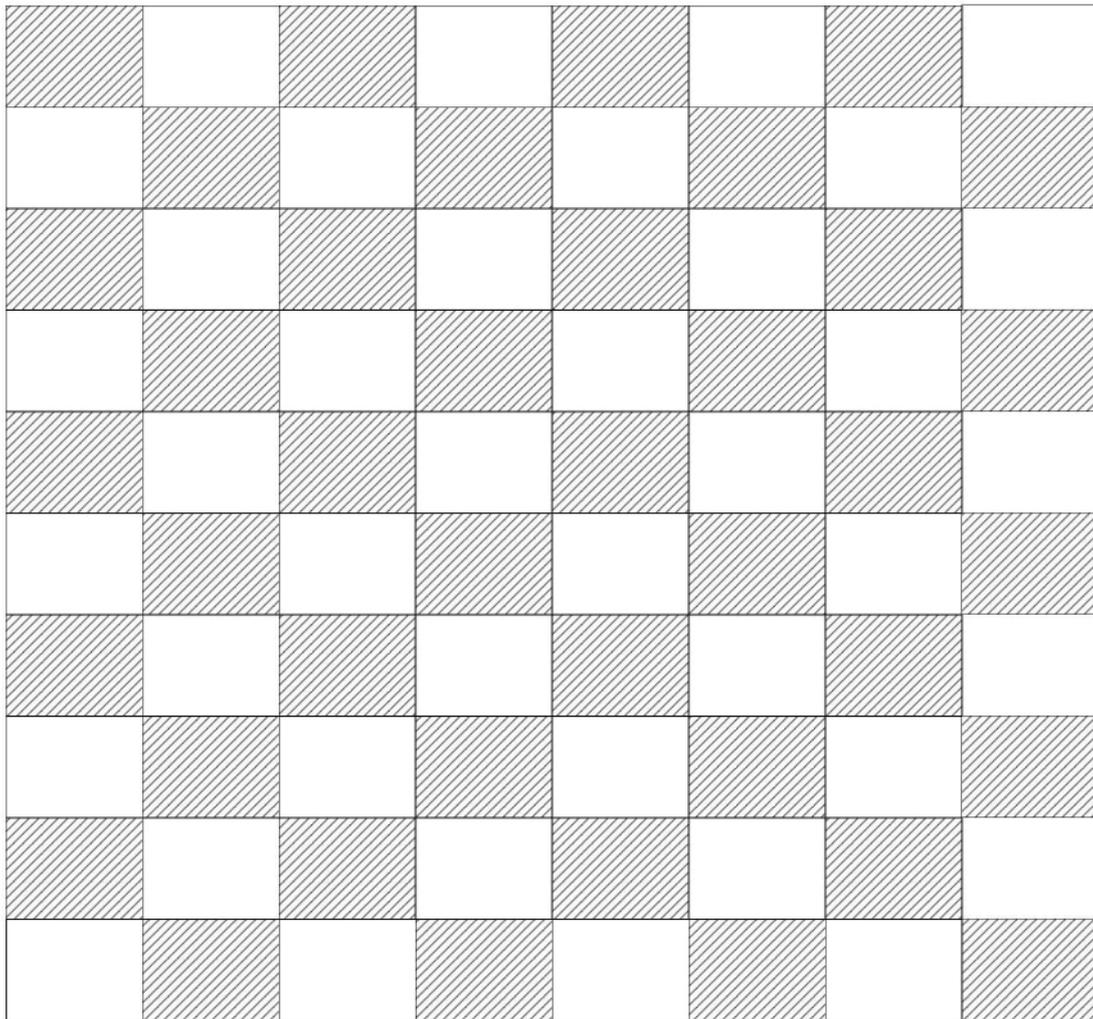


图4a

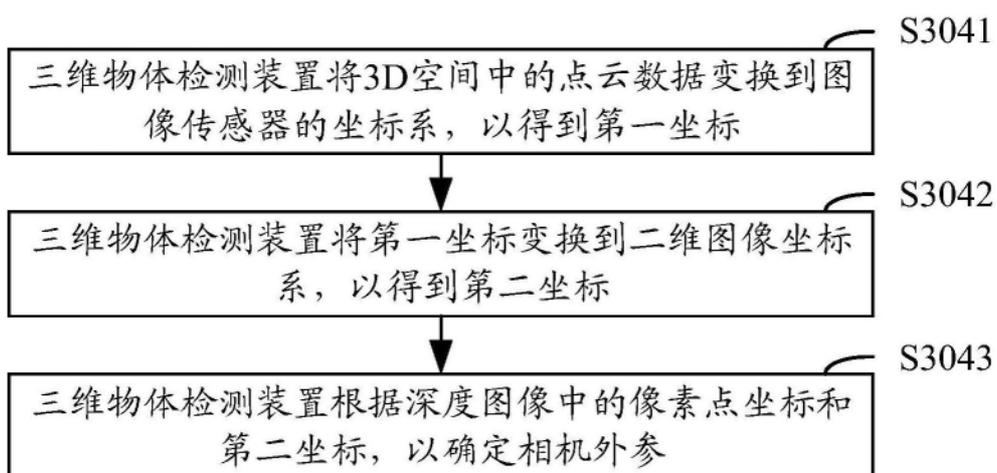


图4b

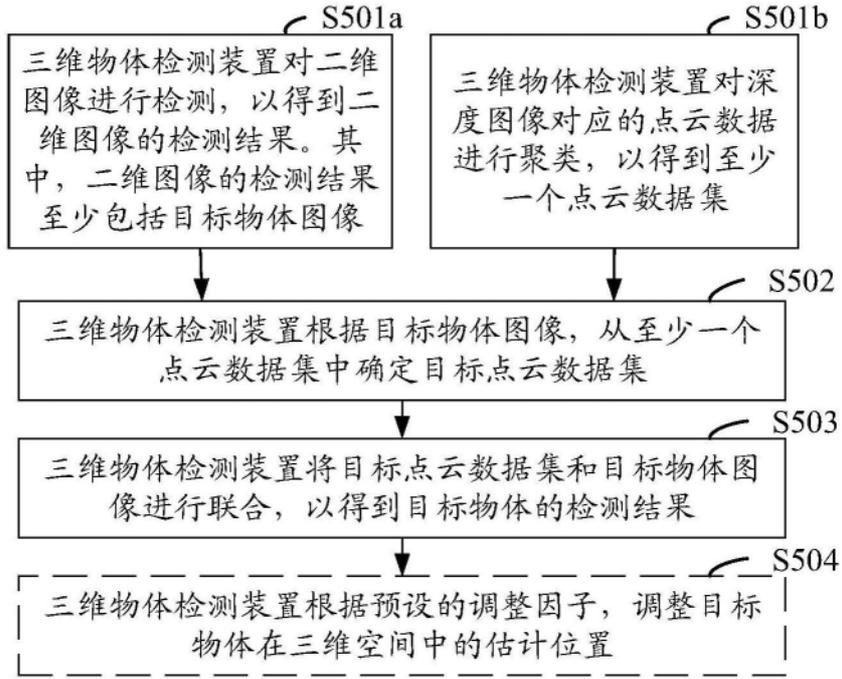


图5

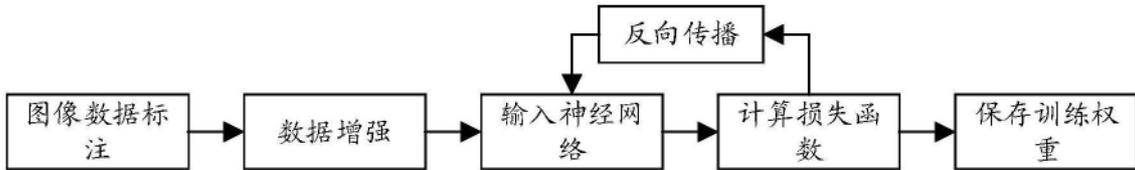


图6a

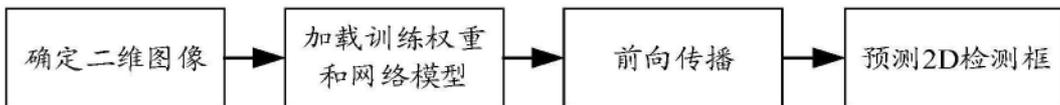


图6b

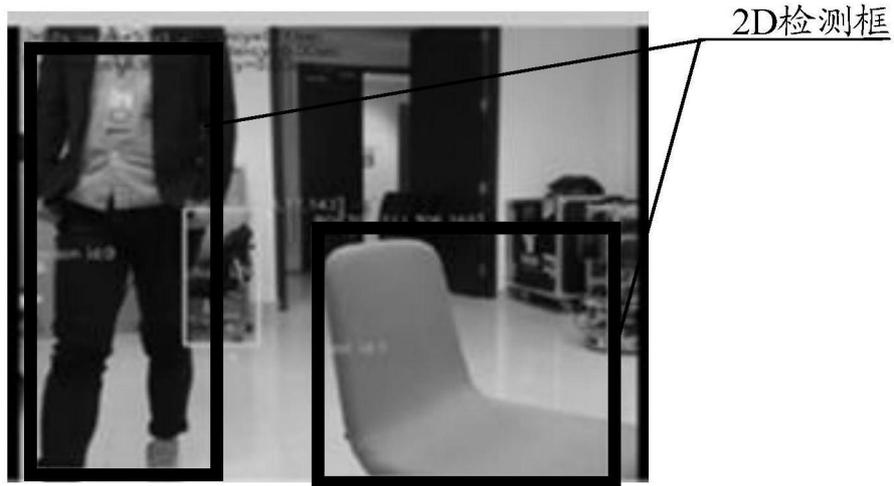


图6c

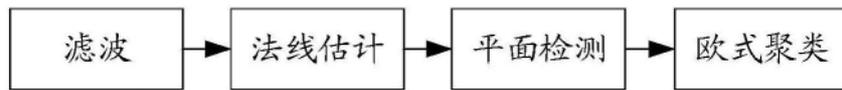


图7a

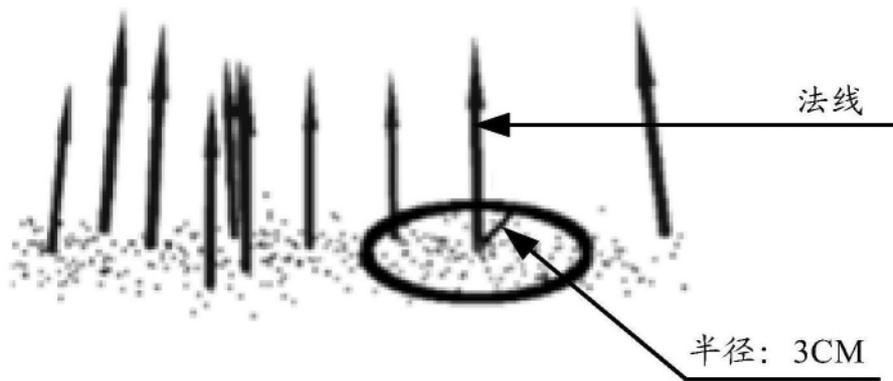


图7b

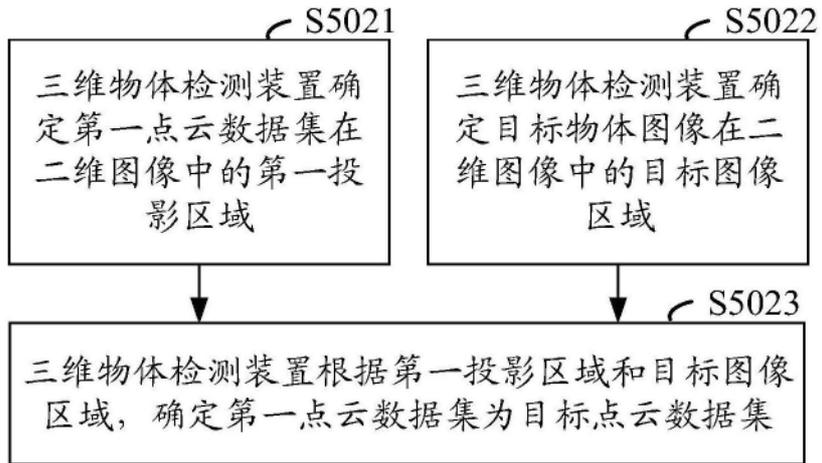


图8

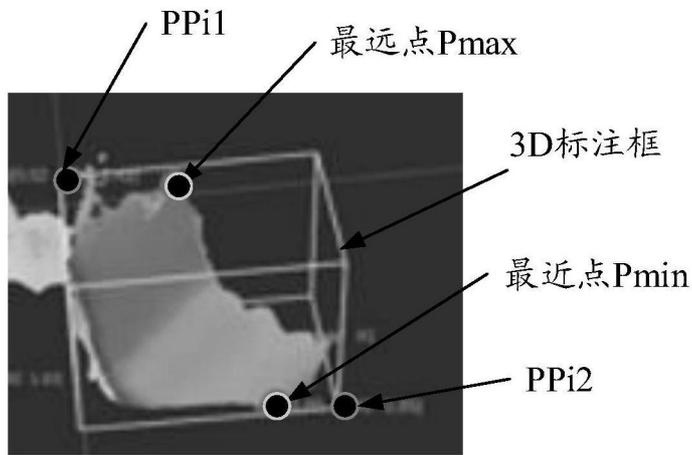


图9a

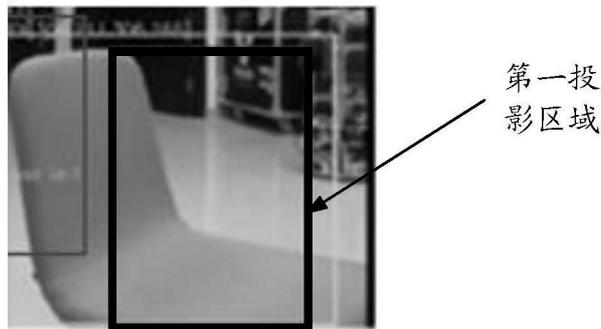


图9b

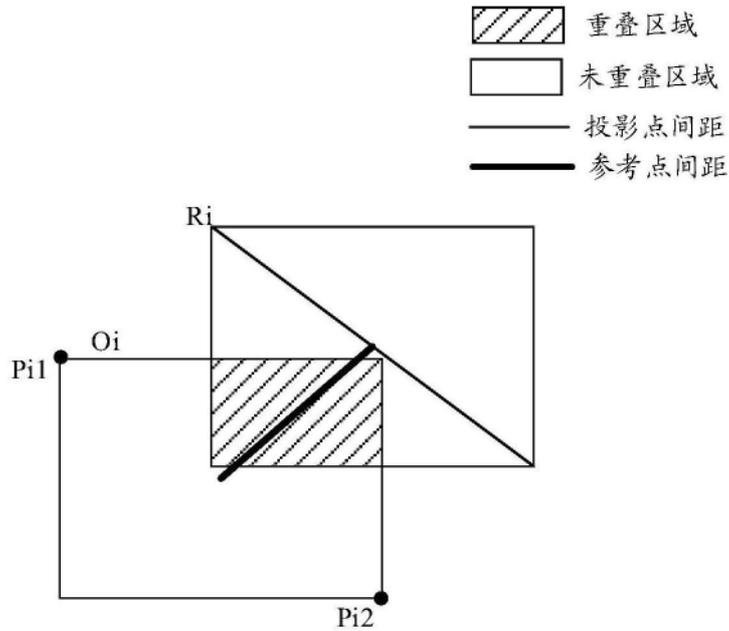


图9c

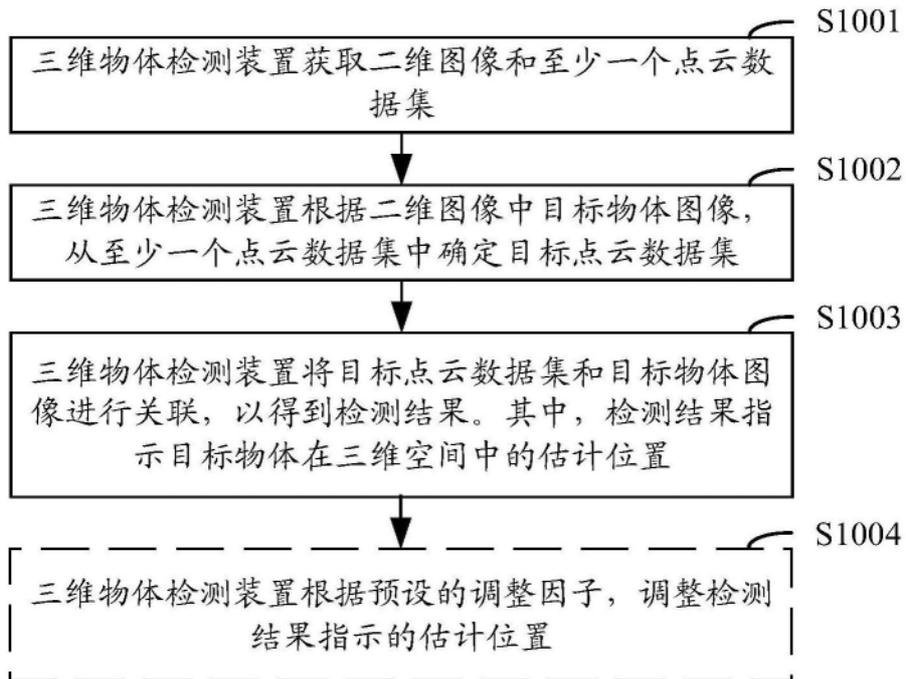


图10

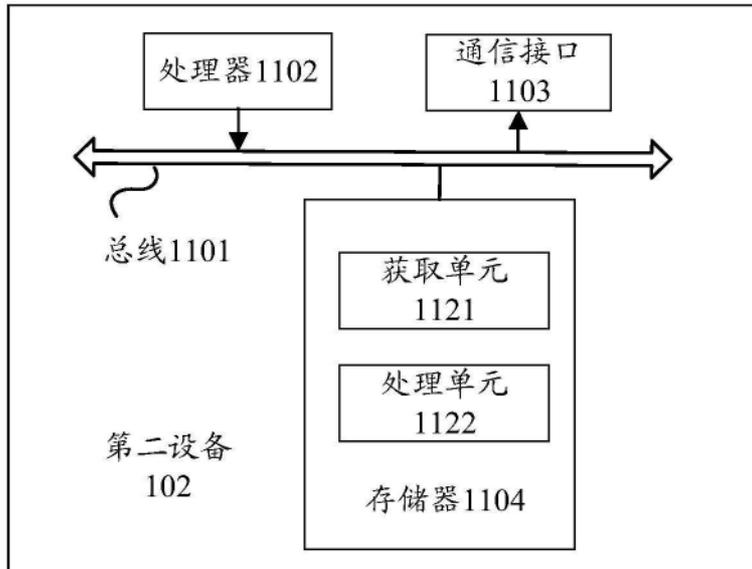


图11