



(12) 发明专利

(10) 授权公告号 CN 113674421 B

(45) 授权公告日 2023. 10. 13

(21) 申请号 202110980060.4

US 2019206071 A1, 2019.07.04

(22) 申请日 2021.08.25

CN 111739005 A, 2020.10.02

(65) 同一申请的已公布的文献号

US 2020167941 A1, 2020.05.28

申请公布号 CN 113674421 A

CN 108509918 A, 2018.09.07

CN 110060331 A, 2019.07.26

(43) 申请公布日 2021.11.19

CN 110264468 A, 2019.09.20

(73) 专利权人 北京百度网讯科技有限公司

CN 110689008 A, 2020.01.14

地址 100085 北京市海淀区上地十街10号

CN 111223135 A, 2020.06.02

百度大厦2层

CN 111291714 A, 2020.06.16

CN 111723721 A, 2020.09.29

(72) 发明人 叶晓青 孙昊

CN 112132829 A, 2020.12.25

(74) 专利代理机构 北京银龙知识产权代理有限公司

US 2020118331 A1, 2020.04.16

公司 11243

US 2021004974 A1, 2021.01.07

专利代理师 黄灿 刘念

US 2021065391 A1, 2021.03.04

CN 110766170 A, 2020.02.07

(51) Int. Cl.

US 2021073997 A1, 2021.03.11

G06T 17/20 (2006.01)

李宇杰;李焯鹏;张为公.基于视觉的三维目标检测算法研究综述.计算机工程与应用.2019,(第01期),全文.

G06V 10/774 (2022.01)

G06V 10/82 (2022.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

审查员 李丽萍

(56) 对比文件

CN 112862006 A, 2021.05.28

权利要求书3页 说明书12页 附图4页

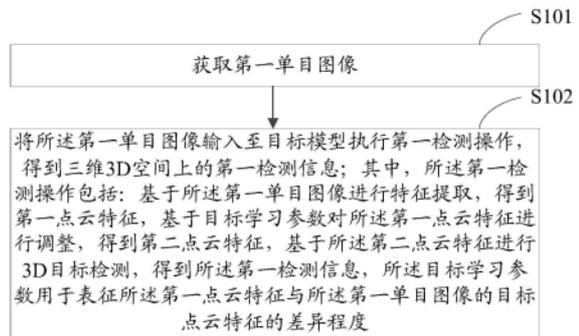
(54) 发明名称

3D目标检测方法、模型训练方法、相关装置及电子设备

的差异程度。

(57) 摘要

本公开提供了一种3D目标检测方法、模型训练方法、相关装置及电子设备,涉及计算机视觉、深度学习等人工智能技术领域。具体实现方案为:获取第一单目图像;将所述第一单目图像输入至目标模型执行第一检测操作,得到三维3D空间上的第一检测信息;其中,所述第一检测操作包括:基于所述第一单目图像进行特征提取,得到第一点云特征,基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,基于所述第二点云特征进行3D目标检测,得到所述第一检测信息,所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征的差异程度



CN 113674421 B

1. 一种3D目标检测方法,包括:

获取第一单目图像;

将所述第一单目图像输入至目标模型执行第一检测操作,得到三维3D空间上的第一检测信息;

其中,所述第一检测操作包括:基于所述第一单目图像进行特征提取,得到第一点云特征,基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,基于所述第二点云特征进行3D目标检测,得到所述第一检测信息,所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征的差异程度;

所述目标学习参数用于表征所述第一点云特征与所述目标点云特征的分布差异程度,所述基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,包括:

对所述第一点云特征进行归一化;

基于所述目标学习参数,对归一化的第一点云特征进行调整,得到所述第二点云特征。

2. 根据权利要求1所述的方法,其中,所述基于所述第一单目图像进行特征提取,得到第一点云特征,包括:

对所述第一单目图像进行深度预测,得到所述第一单目图像的深度信息;

基于所述深度信息和所述第一单目图像对应的相机内参,将所述第一单目图像中的像素点转换为第一3D点云数据;

对所述第一3D点云数据进行特征提取,得到所述第一点云特征。

3. 一种模型训练方法,包括:

获取训练样本数据,所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签;

将所述第二单目图像输入至目标模型执行第二检测操作,得到3D空间上的第二检测信息,所述第二检测操作包括:基于所述第二单目图像进行特征提取,得到第三点云特征,基于所述点云特征标签,对所述第三点云特征进行调整,得到第四点云特征和目标学习参数,所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数,基于所述第四点云特征进行3D目标检测,得到所述第二检测信息;

确定所述目标模型的损失值,所述损失值包括:所述点云特征标签与所述第四点云特征差异值的差异值,以及所述检测标签与所述第二检测信息的差异值;

基于所述损失值更新所述目标模型的网络参数。

4. 根据权利要求3所述的方法,其中,所述基于所述点云特征标签,对所述第三点云特征进行调整,得到第四点云特征和目标学习参数,包括:

分别对所述第三点云特征和所述点云特征标签进行归一化;

基于学习参数,对归一化后的第三点云特征进行调整,得到第五点云特征;

确定所述第五点云特征和归一化的点云特征标签的差异值;

基于所述第五点云特征和归一化的点云特征标签的差异值更新所述学习参数,以得到所述目标学习参数和所述第四点云特征。

5. 一种3D目标检测装置,包括:

第一获取模块,用于获取第一单目图像;

第一执行模块,用于将所述第一单目图像输入至目标模型执行第一检测操作,得到三

维3D空间上的第一检测信息；

其中，所述第一检测操作包括：基于所述第一单目图像进行特征提取，得到第一点云特征，基于目标学习参数对所述第一点云特征进行调整，得到第二点云特征，基于所述第二点云特征进行3D目标检测，得到所述第一检测信息，所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征的差异程度；

所述目标学习参数用于表征所述第一点云特征与所述目标点云特征的分布差异程度，所述第一执行模块包括：

第一归一化单元，用于对所述第一点云特征进行归一化；

第一调整单元，用于基于所述目标学习参数，对归一化的第一点云特征进行调整，得到所述第二点云特征。

6. 根据权利要求5所述的装置，其中，所述第一执行模块包括：

深度预测单元，用于对所述第一单目图像进行深度预测，得到所述第一单目图像的深度信息；

转换单元，用于基于所述深度信息和所述第一单目图像对应的相机内参，将所述第一单目图像中的像素点转换为第一3D点云数据；

第一特征提取单元，用于对所述第一3D点云数据进行特征提取，得到所述第一点云特征。

7. 一种模型训练装置，包括：

第二获取模块，用于获取训练样本数据，所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签；

第二执行模块，用于将所述第二单目图像输入至目标模型执行第二检测操作，得到3D空间上的第二检测信息，所述第二检测操作包括：基于所述第二单目图像进行特征提取，得到第三点云特征，基于所述点云特征标签，对所述第三点云特征进行调整，得到第四点云特征和目标学习参数，所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数，基于所述第四点云特征进行3D目标检测，得到所述第二检测信息；

模型损失确定模块，用于确定所述目标模型的损失值，所述损失值包括：所述点云特征标签与所述第四点云特征差异值，以及所述检测标签与所述第二检测信息的差异值；

网络参数更新模块，用于基于所述损失值更新所述目标模型的网络参数。

8. 根据权利要求7所述的装置，所述第二执行模块包括：

第二归一化单元，用于分别对所述第三点云特征和所述点云特征标签进行归一化；

第二调整单元，用于基于学习参数，对归一化后的第三点云特征进行调整，得到第五点云特征；

特征差异确定单元，用于确定所述第五点云特征和归一化的点云特征标签的差异值；

学习参数更新单元，用于基于所述第五点云特征和归一化的点云特征标签的差异值更新所述学习参数，以得到所述目标学习参数和所述第四点云特征。

9. 一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-2中任一项所述的方法,或者执行权利要求3-4中任一项所述的方法。

10.一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行权利要求1-2中任一项所述的方法,或者执行权利要求3-4中任一项所述的方法。

## 3D目标检测方法、模型训练方法、相关装置及电子设备

### 技术领域

[0001] 本公开涉及人工智能技术领域,尤其涉及计算机视觉、深度学习技术领域,具体涉及一种3D目标检测方法、模型训练方法、相关装置及电子设备。

### 背景技术

[0002] 随着图像处理技术的高速发展,3D目标检测得到了广泛的应用,而单目图像的3D目标检测指的是基于单目图像进行3D目标检测,得到3D空间上的检测信息。

[0003] 目前,单目图像的3D目标检测通常基于彩色RGB图像,结合几何约束或语义知识的方式来进行3D目标检测,或者对单目图像进行深度估计,结合深度信息与图像特征进行3D目标检测。

### 发明内容

[0004] 本公开提供了一种3D目标检测方法、模型训练方法、相关装置及电子设备。

[0005] 根据本公开的第一方面,提供了一种3D目标检测方法,包括:

[0006] 获取第一单目图像;

[0007] 将所述第一单目图像输入至目标模型执行第一检测操作,得到三维3D空间上的第一检测信息;

[0008] 其中,所述第一检测操作包括:基于所述第一单目图像进行特征提取,得到第一点云特征,基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,基于所述第二点云特征进行3D目标检测,得到所述第一检测信息,所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征的差异程度。

[0009] 根据本公开的第二方面,提供了一种模型训练方法,包括:

[0010] 获取训练样本数据,所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签;

[0011] 将所述第二单目图像输入至目标模型执行第二检测操作,得到3D空间上的第二检测信息,所述第二检测操作包括:基于所述第二单目图像进行特征提取,得到第三点云特征,基于所述点云特征标签,对所述第三点云特征进行特征蒸馏,得到第四点云特征和目标学习参数,所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数,基于所述第四点云特征进行3D目标检测,得到所述第二检测信息;

[0012] 确定所述目标模型的损失值,所述损失值包括:所述点云特征标签与所述第四点云特征的特征差异值,以及所述检测标签与所述第二检测信息的差异值;

[0013] 基于所述损失值更新所述目标模型的网络参数。

[0014] 根据本公开的第三方面,提供了一种3D目标检测装置,包括:

[0015] 第一获取模块,用于获取第一单目图像;

[0016] 第一执行模块,用于将所述第一单目图像输入至目标模型执行第一检测操作,得到三维3D空间上的第一检测信息;

[0017] 其中,所述第一检测操作包括:基于所述第一单目图像进行特征提取,得到第一点云特征,基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,基于所述第二点云特征进行3D目标检测,得到所述第一检测信息,所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征的差异程度。

[0018] 根据本公开的第四方面,提供了一种模型训练装置,包括:

[0019] 第二获取模块,用于获取训练样本数据,所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签;

[0020] 第二执行模块,用于将所述第二单目图像输入至目标模型执行第二检测操作,得到3D空间上的第二检测信息,所述第二检测操作包括:基于所述第二单目图像进行特征提取,得到第三点云特征,基于所述点云特征标签,对所述第三点云特征进行特征蒸馏,得到第四点云特征和目标学习参数,所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数,基于所述第四点云特征进行3D目标检测,得到所述第二检测信息;

[0021] 模型损失确定模块,用于确定所述目标模型的损失值,所述损失值包括:所述点云特征标签与所述第四点云特征差异值,以及所述检测标签与所述第二检测信息的差异值;

[0022] 网络参数更新模块,用于基于所述损失值更新所述目标模型的网络参数。

[0023] 根据本公开的第五方面,提供了一种电子设备,包括:

[0024] 至少一个处理器;以及

[0025] 与至少一个处理器通信连接的存储器;其中,

[0026] 存储器存储有可被至少一个处理器执行的指令,该指令被至少一个处理器执行,以使至少一个处理器能够执行第一方面中的任一项方法,或者执行第二方面中的任一项方法。

[0027] 根据本公开的第六方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,该计算机指令用于使计算机执行第一方面中的任一项方法,或者执行第二方面中的任一项方法。

[0028] 根据本公开的第七方面,提供了一种计算机程序产品,包括计算机程序,该计算机程序在被处理器执行时实现第一方面中的任一项方法,或者执行时实现第二方面中的任一项方法。

[0029] 根据本公开的技术解决了3D目标检测的精度比较低的问题,提高了3D目标检测的精度。

[0030] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

## 附图说明

[0031] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0032] 图1是根据本公开第一实施例的3D目标检测方法的流程示意图;

[0033] 图2是目标模型执行第一检测操作的整体框架示意图;

[0034] 图3是根据本公开第二实施例的模型训练方法的流程示意图;

- [0035] 图4是目标模型训练的整体框架示意图；
- [0036] 图5是根据本公开第三实施例的3D目标检测装置的结构示意图；
- [0037] 图6是根据本公开第四实施例的模型训练装置的结构示意图；
- [0038] 图7是用来实施本公开的实施例的示例电子设备的示意性框图。

### 具体实施方式

[0039] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开实施例的各种细节以助于理解,应当将它们认为仅仅是示范性的。因此,本领域普通技术人员应当认识到,可以对这里描述的实施例做出各种改变和修改,而不会背离本公开的范围和精神。同样,为了清楚和简明,以下的描述中省略了对公知功能和结构的描述。

#### [0040] 第一实施例

[0041] 如图1所示,本公开提供一种3D目标检测方法,包括如下步骤:

[0042] 步骤S101:获取第一单目图像。

[0043] 本实施例中,3D目标检测方法涉及人工智能技术领域,尤其涉及计算机视觉、深度学习技术领域,其可以广泛应用于单目3D目标检测场景中,即对单目图像进行3D目标检测。本公开实施例的3D目标检测方法,可以由本公开实施例的3D目标检测装置执行。本公开实施例的3D目标检测装置可以配置在任意电子设备中,以执行本公开实施例的3D目标检测方法。该电子设备可以为服务器,也可以为终端,这里不做具体限定。

[0044] 该步骤中,单目图像是相对于双目图像和多目图像来说的,双目图像指的是在同一场景拍摄的左目图像和右目图像,多目图像指的是在同一场景拍摄的多个图像,而单目图像则指的是在同一场景拍摄的单个图像。

[0045] 本实施例的目的即是对单目图像进行3D目标检测,以获取单目图像中3D空间上的检测信息,该检测信息包括单目图像中物体的3D检测框。在一可选场景中,当单目图像中包括车辆图像数据时,可以对单目图像进行3D目标检测,得到物体类别和车辆的3D检测框,以确定单目图像中的物体类别以及车辆所在位置。

[0046] 第一单目图像可以为彩色RGB图像,也可以为灰度图像,这里不进行具体限定。

[0047] 第一单目图像的获取方式可以包括多种,比如,可以采用单目相机拍摄一张图像作为第一单目图像,也可以获取预先存储的第一单目图像,还可以接收其他电子设备发送的第一单目图像,亦或是从网络上下载一张图像。

[0048] 步骤S102:将所述第一单目图像输入至目标模型执行第一检测操作,得到三维3D空间上的第一检测信息;其中,所述第一检测操作包括:基于所述第一单目图像进行特征提取,得到第一点云特征,基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,基于所述第二点云特征进行3D目标检测,得到所述第一检测信息,所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征的差异程度。

[0049] 该步骤中,目标模型可以为神经网络模型,如卷积神经网络或残差神经网络ResNet等。该目标模型可以用于针对单目图像进行3D目标检测。其输入为任意一张图像,输出为该图像中3D空间上的检测信息,该检测信息可以包括物体类别和物体的3D检测框。

[0050] 可以将第一单目图像输入至目标模型执行第一检测操作,以通过目标模型对第一单目图像进行3D目标检测,得到3D空间上的第一检测信息。其中,第一检测信息包括第一单

目图像中的物体类别和物体的3D检测框,物体类别指的是第一单目图像中物体的分类属性,如物体类别为车辆、猫或人物等,3D检测框指的是物体在第一单目图像中的具体位置,其包括围成物体的长、宽和高的框体,以及通过框体的朝向角来表征物体在第一单目图像中的朝向位置。

[0051] 具体的,第一检测操作可以包括三个部分,第一部分为点云特征提取,第二部分可以为点云特征蒸馏,第三部分为基于点云特征进行3D目标检测。

[0052] 其中,点云特征提取指的是基于第一单目图像进行点云特征提取,得到第一点云特征,第一点云特征可以为与第一单目图像对应的点云三维图像相关的特征,也就是说,其为3D空间上的特征,相对于与二维图像相关的特征来说,其具备图像深度特征。由于点云三维图像可以通过鸟瞰图来表征,因此,第一点云特征也可以称之为鸟瞰图特征,指的是与第一单目图像对应的鸟瞰图相关的特征。

[0053] 点云特征提取的方式可以包括多种,在一可选实施方式中,可以对第一单目图像进行深度估计,得到深度信息,基于预测得到的深度信息确定第一单目图像的点云数据,之后结合该点云数据将2D图像特征转换为体素数据,基于该体素数据进行点云特征提取,得到体素的图像特征即第一点云特征。

[0054] 在另一可选实施方式中,可以对第一单目图像进行深度估计,得到深度信息,基于预测得到的深度信息确定第一单目图像的点云数据,之后将点云数据转换为鸟瞰图,并对鸟瞰图进行点云特征提取,得到第一点云特征。

[0055] 点云特征蒸馏指的是从第一点云特征中蒸馏出可表达第一单目图像的目标点云特征的特征,即蒸馏出与目标点云特征相似的特征。其中,目标点云特征指的是基于第一单目图像的点云数据标签所提取出来的点云特征,其可以称之为点云特征标签,点云数据标签可以为通过激光雷达,针对第一单目图像的同场景采集到的该场景下比较精确的点云数据。

[0056] 可以通过目标学习参数对第一点云特征进行特征蒸馏,得到第二点云特征,第二点云特征可以与目标点云特征相似,具体可以基于目标学习参数对第一点云特征进行调整,得到第二点云特征。

[0057] 其中,目标学习参数可以表征第一点云特征与目标点云特征的差异程度,其是通过目标模型进行训练得到。在一可选实施方式中,目标学习参数可以包括第一点云特征与目标点云特征之间像素点的特征差异值。相应的,基于该特征差异值,对第一点云特征中像素点的特征值进行调整,可以得到与目标点云特征相似的第二点云特征。

[0058] 在另一可选实施方式中,目标学习参数具体可以表征第一点云特征与目标点云特征的分布差异程度,该目标学习参数可以包括第一点云特征与目标点云特征的分布之间的均值差异值和方差差异值。

[0059] 在该实施方式中,设第一点云特征为 $BEV_{img}$ ,目标学习参数为 $(\Delta \mu_{img}, \Delta \sigma_{img})$ ,基于该目标学习参数对第一点云特征进行调整的步骤具体可以为:计算 $BEV_{img}$ 的均值和方差,记为 $(\mu_{img}, \sigma_{img})$ ,基于该均值和方差,对 $BEV_{img}$ 进行归一化,得到归一化的第一点云特征,用

$\overline{BEV}_{img}$ 表示,  $\overline{BEV}_{img} = \frac{BEV_{img} - \mu_{img}}{\sigma_{img}}$ ,基于该目标学习参数,采用下式(1)对归一化的

第一点云特征进行调整,得到第二点云特征。

$$[0060] \quad \overline{BEV}_{img} = \overline{BEV}_{img} * \Delta\sigma_{img} + \Delta\mu_{img} \quad (1)$$

[0061] 其中,  $\overline{BEV}_{img}$  为第二点云特征。

[0062] 之后,可以采用现有的或新的检测方式,基于第二点云特征进行3D目标检测,得到第一检测信息,这里的检测方式不进行具体限定。

[0063] 需要说明的是,目标模型在使用之前,需要对其进行训练,以学习目标模型参数,包括目标学习参数,其训练过程将在以下实施例中进行详细说明。

[0064] 本实施例中,通过目标模型对第一单目图像进行点云特征提取,得到第一点云特征,并基于目标学习参数对第一点云特征进行点云特征蒸馏,得到与目标点云特征相似的第二点云特征,之后基于第二点云特征进行3D目标检测,得到第一检测信息。如此,可以通过目标模型对单目图像进行点云特征提取和点云特征蒸馏,使得单目图像学到的特征与目标点云特征相似,从而可以提高单目3D目标检测的精度。

[0065] 可选的,所述基于所述第一单目图像进行特征提取,得到第一点云特征,包括:

[0066] 对所述第一单目图像进行深度预测,得到所述第一单目图像的深度信息;

[0067] 基于所述深度信息和所述第一单目图像对应的相机内参,将所述第一单目图像中的像素点转换为第一3D点云数据;

[0068] 对所述第一3D点云数据进行特征提取,得到所述第一点云特征。

[0069] 本实施方式中,目标模型执行第一检测操作的整体框架示意图如图2所示,目标模型可以包括2D编码器和用于对单目图像进行深度预测的网络分支,2D编码器用于提取第一单目图像的2D图像特征,且用于对单目图像进行深度预测的网络分支串接在2D图像编辑器之后。

[0070] 可以对第一单目图像进行深度估计,得到深度信息,基于预测得到的深度信息确定第一单目图像的点云数据,之后结合该点云数据将2D图像特征转换为体素数据,基于该体素数据进行点云特征提取,得到体素的图像特征即第一点云特征。

[0071] 具体的,给定W x H大小的RGB图像作为目标模型的输入,基于该网络分支,采用现有的或新的深度预测方法对该RGB图像进行深度预测,得到该RGB图像的深度信息。

[0072] 基于该深度信息确定第一单目图像的点云数据,在一可选实施方式中,可以结合深度信息和第一单目图像对应的相机内参将第一单目图像中每个像素点转换为三维点云。

具体为,即相机内参为  $K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ , 预测的深度图为D(u, v), 针对第一单目图像中

每一像素点,记为I(u, v), 可以采用下式(2), 基于相机内参和深度图转换为三维点云。

$$[0073] \quad D \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = KP_c = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (2)$$

[0074] 其中,  $P_c$  为三维点云, 对上式(2)进行转换, 采用下式(3)表示  $P_c$ 。

$$[0075] \quad P_c = DK^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3)$$

[0076] 针对每个三维点,可以基于该三维点将2D图像特征转换为体素,得到体素数据,之后目标模型中可以串接一个现有的或者新的用于提取点云特征的网络对该体素数据进行点云特征提取,得到体素的图像特征即第一点云特征。

[0077] 本实施方式中,通过对所述第一单目图像进行深度预测,得到所述第一单目图像的深度信息;基于所述深度信息和所述第一单目图像对应的相机内参,将所述第一单目图像中的像素点转换为第一3D点云数据;对所述第一3D点云数据进行特征提取,得到所述第一点云特征。如此,可以实现基于第一单目图像进行点云特征提取,得到第一点云特征,且点云特征提取方式简单易实现。

[0078] 可选的,所述目标学习参数用于表征所述第一点云特征与所述目标点云特征的分布差异程度,所述基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,包括:

[0079] 对所述第一点云特征进行归一化;

[0080] 基于所述目标学习参数,对归一化的第一点云特征进行调整,得到所述第二点云特征。

[0081] 本实施方式中,目标学习参数具体可以表征第一点云特征与目标点云特征的分布差异程度,该目标学习参数可以包括第一点云特征与目标点云特征的分布之间的均值差异值和方差差异值。

[0082] 设第一点云特征为 $BEV_{img}$ ,目标学习参数为 $(\Delta \mu_{img}, \Delta \sigma_{img})$ ,其中 $\Delta \mu_{img}$ 表征第一点云特征与目标点云特征的分布之间的均值差异值, $\Delta \sigma_{img}$ 表征第一点云特征与目标点云特征的分布之间的方差差异值。

[0083] 基于该目标学习参数对第一点云特征进行调整的步骤具体可以为:计算 $BEV_{img}$ 的均值和方差,记为 $(\mu_{img}, \sigma_{img})$ ,基于该均值和方差,对 $BEV_{img}$ 进行归一化,得到归一化的第一点云特征 $\overline{BEV}_{img}$ ,基于该目标学习参数,采用上式(1)对归一化的第一点云特征进行调整,得到第二点云特征 $\overline{BEV}_{img}$ 。

[0084] 本实施方式中,在目标学习参数用于表征所述第一点云特征与所述目标点云特征的分布差异程度的情况下,通过对所述第一点云特征进行归一化;基于所述目标学习参数,对归一化的第一点云特征进行调整,得到所述第二点云特征。如此,可以实现对第一点云特征进行点云特征蒸馏,得到第二点云特征,且点云特征蒸馏方式简单易实现。

[0085] 第二实施例

[0086] 如图3所示,本公开提供一种模型训练方法,包括如下步骤:

[0087] 步骤S301:获取训练样本数据,所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签。

[0088] 步骤S302:将所述第二单目图像输入至目标模型执行第二检测操作,得到3D空间上的第二检测信息,所述第二检测操作包括:基于所述第二单目图像进行特征提取,得到第三点云特征,基于所述点云特征标签,对所述第三点云特征进行特征蒸馏,得到第四点云特

征和目标学习参数,所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数,基于所述第四点云特征进行3D目标检测,得到所述第二检测信息;

[0089] 步骤S303:确定所述目标模型的损失值,所述损失值包括:所述点云特征标签与所述第四点云特征差异值,以及所述检测标签与所述第二检测信息的差异值;

[0090] 步骤S304:基于所述损失值更新所述目标模型的网络参数。

[0091] 本实施例描述的是目标模型的训练过程。

[0092] 在步骤S301中,训练样本数据中可以包括很多个第二单目图像以及每个第二单目图像对应的点云特征标签和3D空间上的检测标签。

[0093] 第二单目图像的获取方式可以包括多种,可以采用一种或多种方式获取训练样本数据中的第二单目图像。比如,可以通过单目相机直接拍摄一张单目图像,并将该单目图像作为第二单目图像,也可以获取预先存储的该第二单目图像,还可以接收其他电子设备发送的该第二单目图像,亦或是从网络上下载单目图像作为该第二单目图像。

[0094] 第二单目图像对应的点云特征标签可以指的是基于第二单目图像的点云数据标签所提取出来的点云特征,该点云特征标签可以比较准确地表达第二单目图像的特征。第二单目图像的点云数据标签可以为通过激光雷达,针对第二单目图像的同场景采集到的该场景下比较精确的点云数据。

[0095] 第二单目图像对应的点云特征标签的获取方式可以包括多种,比如,在准确获知第二单目图像的点云数据标签的情况下,可以对该点云数据标签进行点云特征提取,得到点云特征标签,也可以获取预先存储的该第二单目图像对应的点云特征标签,还可以接收其他电子设备发送的该第二单目图像对应的点云特征标签。

[0096] 第二单目图像对应的3D空间上的检测标签可以包括表征第二单目图像中物体类别的标签,以及表征第二单目图像中物体位置的3D检测框的标签,其获取方式可以包括多种,比如,可以对点云特征标签进行3D目标检测,得到检测标签,也可以获取预先存储的该第二单目图像对应的检测标签,还可以接收其他电子设备发送的该第二单目图像对应的检测标签。

[0097] 在一可选实施方式中,可以基于点云预训练网络模型来获取,该模型的参数已经固定,比如,常见的点云三维检测框架Second或PointPillars等。可以将第二单目图像对应的真实雷达点云输入至点云预训练网络模型进行3D目标检测,中间特征图为点云特征标签,输出为第二单目图像对应的检测标签。

[0098] 目标模型训练的整体框架示意图如图4所示,可以输入真实雷达点云至点云预训练网络模型,点云预训练网络模型对该真实雷达点云进行体素化,得到体素数据,并基于3D编码器进行特征提取,得到点云特征标签 $BEV_{cloud}$ ,归一化点云特征标签后,得到归一化的点云特征标签,用 $\overline{BEV}_{cloud}$ 表示。

[0099] 在步骤S302中,可以将第二单目图像输入至目标模型执行第二检测操作,得到第二检测信息。其中,第二检测操作也可以包括点云特征提取、点云特征蒸馏和基于点云特征进行3D目标检测。

[0100] 第二检测操作中的点云特征提取与第一检测操作中的点云特征提取的方式类似,以及第二检测操作中的基于点云特征进行3D目标检测与第一检测操作中的基于点云特征

进行3D目标检测的方式类似,这里不进行赘述。

[0101] 第二检测操作中的点云特征蒸馏方式也可以包括多种,在一可选实施方式中,可以设置初始学习参数,该初始学习参数可以包括两个点云特征之间像素点的特征差异值,基于该初始学习参数对第三点云特征中像素点的特征值进行调整,得到另一点云特征,确定调整后的点云特征与点云特征标签之间像素点的特征差异值,基于该特征差异值,采用梯度下降等方法调整初始学习参数,最终得到目标学习参数。

[0102] 其中,目标学习参数可以包括第三点云特征与目标点云特征之间像素点的特征差异值,基于该特征差异值,对第三点云特征中像素点的特征值进行调整,可以得到与点云特征标签相似的第四点云特征。

[0103] 在另一可选实施方式中,可以设置初始学习参数,该初始学习参数可以表征两个点云特征之间的分布差异值,基于该初始学习参数对第三点云特征的分布进行调整,得到另一点云特征,确定调整后的点云特征与点云特征标签之间的分布差异值,基于该分布差异值,采用梯度下降等方法调整初始学习参数,最终得到目标学习参数。

[0104] 其中,目标学习参数具体可以表征第三点云特征与点云特征标签的分布差异程度,该目标学习参数可以包括第三点云特征与点云特征标签的分布之间的均值差异值和方差差异值。基于该均值差异值和方差差异值,对第三点云特征的分布进行调整,可以得到与点云特征标签分布相似的第四点云特征。

[0105] 另外,第二检测信息中包括的内容与第一检测信息中包括的内容类似,这里不进行赘述。

[0106] 在步骤S303中,可以确定目标模型的损失值,该损失值可以包括点云特征标签与第四点云特征的差异值,以及检测标签与第二检测信息的差异值,具体可以下式(4)来计算目标模型的损失值。

$$L=L_{\text{distill}}+L_{\text{class}}+L_{\text{box3d}} \quad (4)$$

[0108] 其中, $L$ 表示目标模型的损失值, $L_{\text{distill}}$ 表示点云特征标签与第四点云特征的差异值, $L_{\text{distill}} = \|\overline{BEV}_{\text{img}} - \overline{BEV}_{\text{cloud}}\|_{L_2}$ , $L_{\text{class}}$ 为检测标签中物体类别的标签与第二检测信息中物体类别的差异值, $L_{\text{box3d}}$ 为检测标签中3D检测框与第二检测信息中3D检测框的差异值,包括两个3D检测框之间的长、宽、高以及朝向角的差异值。

[0109] 在步骤S304中,基于该损失值,采用梯度下降等方法更新目标模型的网络参数,直至目标模型的损失值小于某一阈值,且达到收敛,此时目标模型训练完成。

[0110] 本实施例中,通过获取训练样本数据,所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签;将所述第二单目图像输入至目标模型执行第二检测操作,得到3D空间上的第二检测信息,所述第二检测操作包括:基于所述第二单目图像进行特征提取,得到第三点云特征,基于所述点云特征标签,对所述第三点云特征进行特征蒸馏,得到第四点云特征和目标学习参数,所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数,基于所述第四点云特征进行3D目标检测,得到所述第二检测信息;确定所述目标模型的损失值,所述损失值包括:所述点云特征标签与所述第四点云特征差异值,以及所述检测标签与第二检测信息的差异值;基于所述损失值更新所述目标模型的网络参数。如此,可以实现目标模型的

训练,采用该目标模型对单目图像进行3D目标检测,可以提高单目3D目标检测的精度。

[0111] 可选的,所述基于所述点云特征标签,对所述第三点云特征进行特征蒸馏,得到第四点云特征和目标学习参数,包括:

[0112] 分别对所述第三点云特征和所述点云特征标签进行归一化;

[0113] 基于学习参数,对归一化后的第三点云特征进行调整,得到第五点云特征;

[0114] 确定所述第五点云特征和归一化的点云特征标签的差异值;

[0115] 基于所述第五点云特征和归一化的点云特征标签的差异值更新所述学习参数,以得到所述目标学习参数和所述第四点云特征。

[0116] 本实施方式中,第三点云特征和点云特征标签的归一化方式与第一点云特征的归一化方式类似,这里不进行赘述。

[0117] 可以设置初始学习参数,该初始学习参数可以表征两个点云特征之间的分布差异值,基于该初始学习参数对第三点云特征的分布(即归一化的第三点云特征)进行调整,得到另一点云特征即第五点云特征,确定第五点云特征与点云特征标签之间的分布差异值,即确定第五点云特征与归一化的点云特征之间的差异值。基于该分布差异值,采用梯度下降等方法调整初始学习参数,最终得到目标学习参数。

[0118] 其中,目标学习参数具体可以表征第三点云特征与点云特征标签的分布差异程度,该目标学习参数可以包括第三点云特征与点云特征标签的分布之间的均值差异值和方差差异值。基于该均值差异值和方差差异值,对第三点云特征的分布进行调整,可以得到与点云特征标签分布相似的第四点云特征。

[0119] 在训练过程中,可以首先确定目标学习参数,在该目标学习参数下,确定目标模型的损失值,以更新目标模型的网络参数,之后在更新的目标模型的网络参数下,由于第三点云特征进行了更新,相应的,再次更新目标学习参数,直至目标模型的损失值小于某一阈值,且达到收敛,此时,将最后一次更新的网络参数和目标学习参数用于实际的单目3D目标检测。

[0120] 本实施方式中,通过分别对所述第三点云特征和所述点云特征标签进行归一化;基于学习参数,对归一化后的第三点云特征进行调整,得到第五点云特征;确定所述第五点云特征和归一化的点云特征标签的差异值;基于所述第五点云特征和归一化的点云特征标签的差异值更新所述学习参数,以得到所述目标学习参数和所述第四点云特征。如此,可以在目标模型的训练过程中实现对第三点云特征的特征蒸馏,得到与点云特征标签相似的第四点云特征,且点云特征蒸馏方式简单易实现。

[0121] 第三实施例

[0122] 如图5所示,本公开提供一种3D目标检测装置500,包括:

[0123] 第一获取模块501,用于获取第一单目图像;

[0124] 第一执行模块502,用于将所述第一单目图像输入至目标模型执行第一检测操作,得到三维3D空间上的第一检测信息;

[0125] 其中,所述第一检测操作包括:基于所述第一单目图像进行特征提取,得到第一点云特征,基于目标学习参数对所述第一点云特征进行调整,得到第二点云特征,基于所述第二点云特征进行3D目标检测,得到所述第一检测信息,所述目标学习参数用于表征所述第一点云特征与所述第一单目图像的目标点云特征之间的差异程度。

- [0126] 可选的,所述第一执行模块502包括:
- [0127] 深度预测单元,用于对所述第一单目图像进行深度预测,得到所述第一单目图像的深度信息;
- [0128] 转换单元,用于基于所述深度信息和所述第一单目图像对应的相机内参,将所述第一单目图像中的像素点转换为第一3D点云数据;
- [0129] 第一特征提取单元,用于对所述第一3D点云数据进行特征提取,得到所述第一点云特征。
- [0130] 可选的,所述目标学习参数用于表征所述第一点云特征与所述目标点云特征的分布差异程度,所述第一执行模块502包括:
- [0131] 第一归一化单元,用于对所述第一点云特征进行归一化;
- [0132] 第一调整单元,用于基于所述目标学习参数,对归一化的第一点云特征进行调整,得到所述第二点云特征。
- [0133] 本公开提供的3D目标检测装置500能够实现3D目标检测方法实施例实现的各个过程,且能够达到相同的有益效果,为避免重复,这里不再赘述。
- [0134] 第四实施例
- [0135] 如图6所示,本公开提供一种模型训练装置600,包括:
- [0136] 第二获取模块601,用于获取训练样本数据,所述训练样本数据包括第二单目图像、所述第二单目图像对应的点云特征标签和3D空间上的检测标签;
- [0137] 第二执行模块602,用于将所述第二单目图像输入至目标模型执行第二检测操作,得到3D空间上的第二检测信息,所述第二检测操作包括:基于所述第二单目图像进行特征提取,得到第三点云特征,基于所述点云特征标签,对所述第三点云特征进行特征蒸馏,得到第四点云特征和目标学习参数,所述目标学习参数为使所述第四点云特征和所述点云特征标签的差异值小于预设阈值的学习参数,基于所述第四点云特征进行3D目标检测,得到所述第二检测信息;
- [0138] 模型损失确定模块603,用于确定所述目标模型的损失值,所述损失值包括:所述点云特征标签与所述第四点云特征差异值,以及所述检测标签与所述第二检测信息的差异值;
- [0139] 网络参数更新模块604,用于基于所述损失值更新所述目标模型的网络参数。
- [0140] 可选的,所述第二执行模块602包括:
- [0141] 第二归一化单元,用于分别对所述第三点云特征和所述点云特征标签进行归一化;
- [0142] 第二调整单元,用于基于学习参数,对归一化后的第三点云特征进行调整,得到第五点云特征;
- [0143] 特征差异确定单元,用于确定所述第五点云特征和归一化的点云特征标签的差异值;
- [0144] 学习参数更新单元,用于基于所述第五点云特征和归一化的点云特征标签的差异值更新所述学习参数,以得到所述目标学习参数和所述第四点云特征。
- [0145] 本公开提供的模型训练装置600能够实现模型训练方法实施例实现的各个过程,且能够达到相同的有益效果,为避免重复,这里不再赘述。

[0146] 本公开的技术方案中,所涉及的用户个人信息的收集、存储、使用、加工、传输、提供和公开等处理,均符合相关法律法规的规定,且不违背公序良俗。

[0147] 根据本公开的实施例,本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。

[0148] 图7示出了可以用来实施本公开的实施例的示例电子设备700的示意性框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0149] 如图7所示,设备700包括计算单元701,其可以根据存储在只读存储器(ROM)702中的计算机程序或者从存储单元708加载到随机访问存储器(RAM)703中的计算机程序,来执行各种适当的动作和处理。在RAM 703中,还可存储设备700操作所需的各种程序和数据。计算单元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。

[0150] 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。

[0151] 计算单元701可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元701的一些示例包括但不限于中央处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元701执行上文所描述的各个方法和处理,例如3D目标检测方法或模型训练方法。例如,在一些实施例中,3D目标检测方法或模型训练方法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并由计算单元701执行时,可以执行上文描述的3D目标检测方法或模型训练方法的一个或多个步骤。备选地,在其他实施例中,计算单元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3D目标检测方法或模型训练方法。

[0152] 本文中以上描述的系统和技术各种实施方式可以在数字电子电路系统、集成电路系统、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)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0153] 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或控制器,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器上执行。

[0154] 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。

[0155] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0156] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0157] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,也可以为分布式系统的服务器,或者是结合了区块链的服务器。

[0158] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0159] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

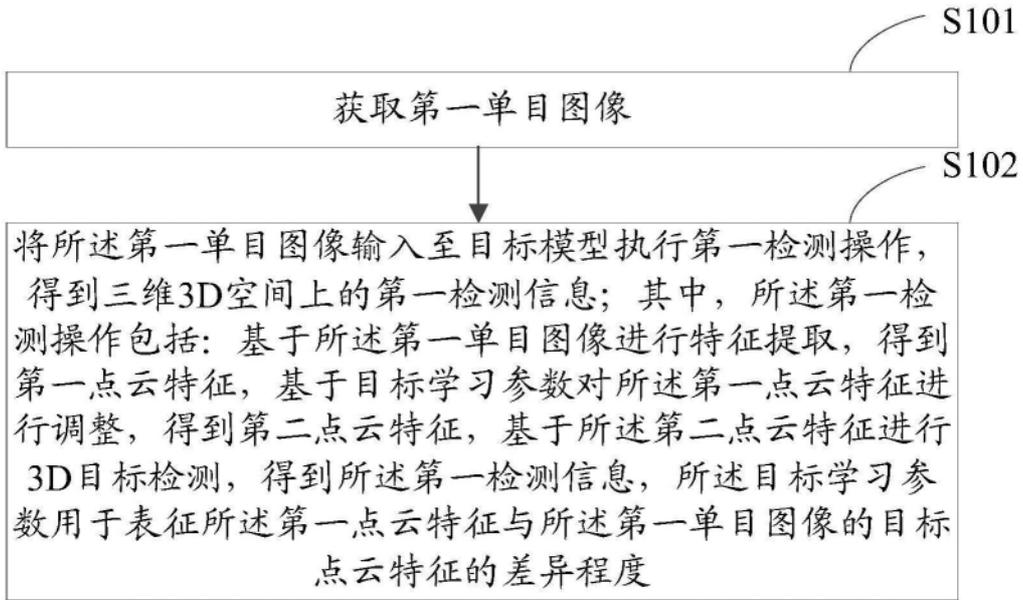


图1

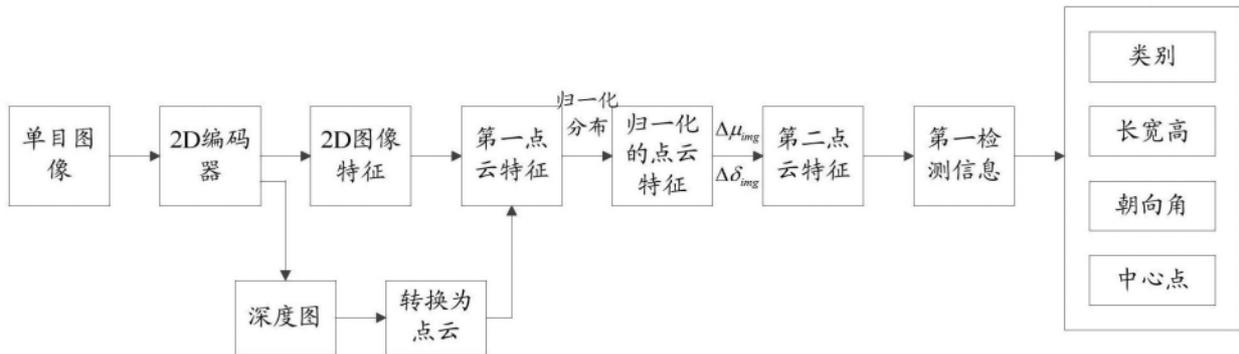


图2

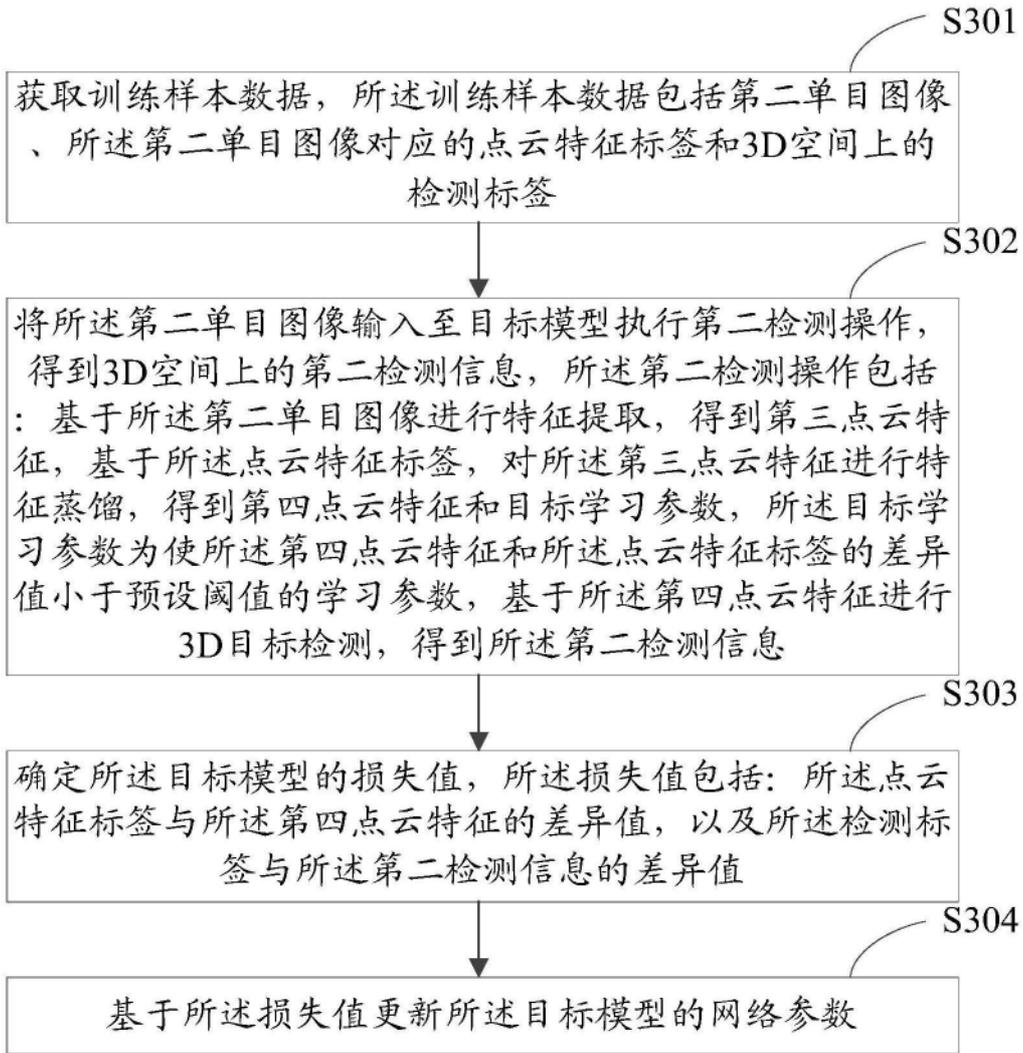


图3

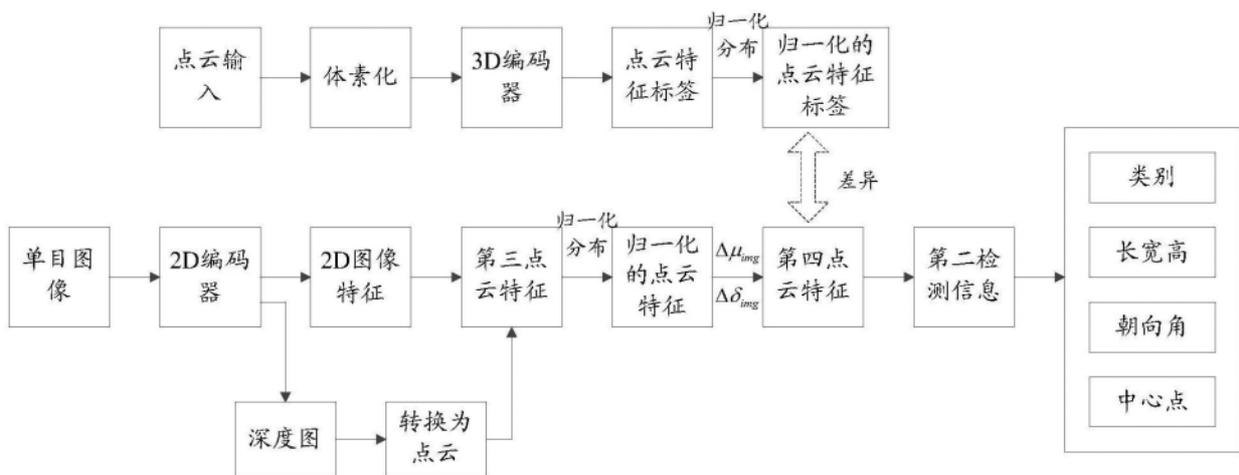


图4

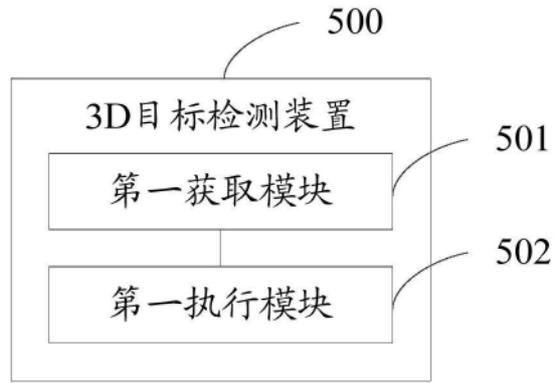


图5

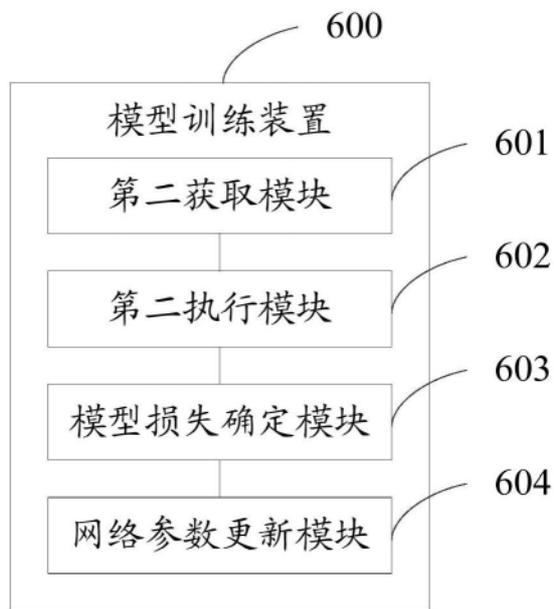


图6

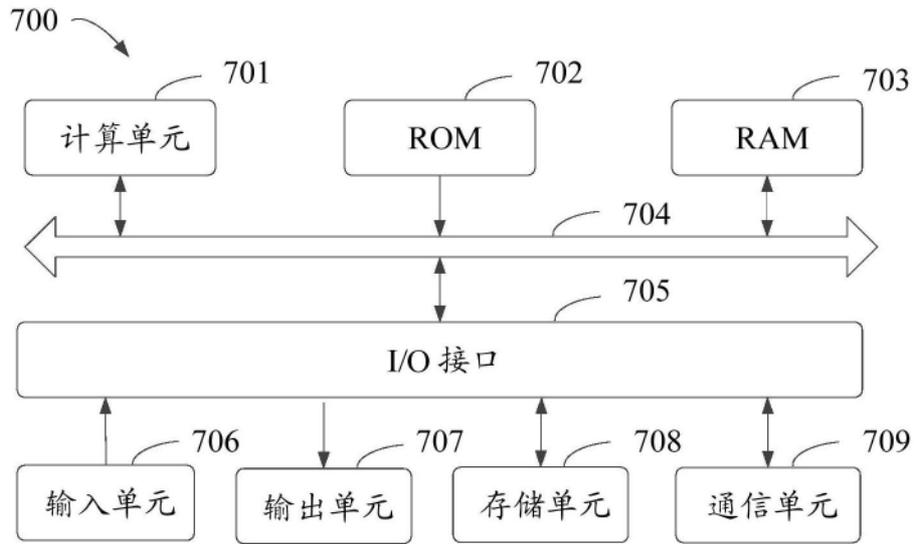


图7