



(12) 发明专利申请

(10) 申请公布号 CN 118657932 A

(43) 申请公布日 2024.09.17

(21) 申请号 202411118951.9

G01S 17/86 (2020.01)

(22) 申请日 2024.08.15

(71) 申请人 新石器慧通(北京)科技有限公司  
地址 100176 北京市大兴区北京经济技术  
开发区宏达北路12号A幢二区6层613  
室

(72) 发明人 张雄 郭和炆 苗乾坤

(74) 专利代理机构 北京嘉科知识产权代理事务  
所(特殊普通合伙) 11687  
专利代理师 彭杰

(51) Int. Cl.

G06V 10/25 (2022.01)

G06V 10/82 (2022.01)

G06V 20/56 (2022.01)

G01S 17/931 (2020.01)

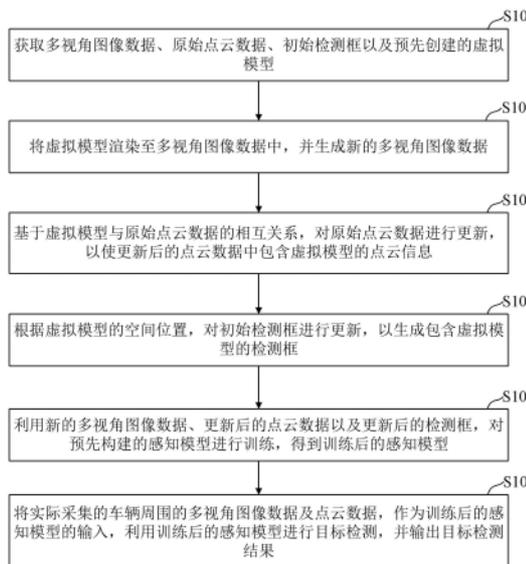
权利要求书2页 说明书12页 附图2页

(54) 发明名称

基于数据合成的目标检测方法、装置、电子设备  
及介质

(57) 摘要

本申请提供一种基于数据合成的目标检测方法、装置、电子设备及介质。该方法包括：获取多视角图像数据、原始点云数据、初始检测框及预先创建的虚拟模型；将虚拟模型渲染至多视角图像数据中，并生成新的多视角图像数据；基于虚拟模型与原始点云数据的相互关系，对原始点云数据进行更新，以使更新后的点云数据中包含虚拟模型的点云信息；根据虚拟模型的空间位置，对初始检测框进行更新，以生成包含虚拟模型的检测框；利用新的多视角图像数据、更新后的点云数据以及更新后的检测框，对预先构建的感知模型进行训练，得到训练后的感知模型；将实际采集的车辆周围的多视角图像数据及点云数据，作为训练后的感知模型的输入，利用训练后的感知模型进行目标检测，输出目标检测结果。本申请能够降低感知数据标注成本，提高模型检测精度和准确性。



1. 一种基于数据合成的目标检测方法,其特征在于,包括:
  - 获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;
  - 将所述虚拟模型渲染至所述多视角图像数据中,并生成新的多视角图像数据;
  - 基于所述虚拟模型与所述原始点云数据的相互关系,对所述原始点云数据进行更新,以使更新后的点云数据中包含所述虚拟模型的点云信息;
  - 根据所述虚拟模型的空间位置,对所述初始检测框进行更新,以生成包含所述虚拟模型的检测框;
  - 利用所述新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;
  - 将实际采集的车辆周围的多视角图像数据及点云数据,作为所述训练后的感知模型的输入,利用所述训练后的感知模型进行目标检测,并输出目标检测结果。
2. 根据权利要求1所述的方法,其特征在于,所述将所述虚拟模型渲染至所述多视角图像数据中,并生成新的多视角图像数据,包括:
  - 利用计算机图形学方法将所述虚拟模型渲染到每个摄像头对应的视角空间中;
  - 将所述虚拟模型与原始的多视角图像数据进行叠加,生成包含虚拟模型的多视角图像数据;
  - 将所有视角对应的包含虚拟模型的多视角图像数据进行校正和融合,得到所述新的多视角图像数据。
3. 根据权利要求1所述的方法,其特征在于,所述基于所述虚拟模型与所述原始点云数据的相互关系,对所述原始点云数据进行更新,包括:
  - 在场景中插入虚拟模型后,确定所述原始点云数据中被所述虚拟模型遮挡的点,并将所述原始点云数据中被所述虚拟模型遮挡的点删除;
  - 计算每条激光雷达线束与所述虚拟模型的交点,将最近的交点作为新增的点云数据,将所述新增的点云数据与原始点云数据融合,生成包含虚拟模型的点云数据。
4. 根据权利要求3所述的方法,其特征在于,所述确定所述原始点云数据中被所述虚拟模型遮挡的点,包括:
  - 确定激光雷达设备的原点与所述原始点云数据中每个点的射线,根据所述射线与所述虚拟模型的每个面片之间的相交性,判断是否为所述原始点云数据中被所述虚拟模型遮挡的点;
  - 其中,当所述射线与所述虚拟模型中的面片相交时,判断所述原始点云数据中的点属于被所述虚拟模型遮挡的点,否则,判断所述原始点云数据中的点不属于被所述虚拟模型遮挡的点。
5. 根据权利要求1所述的方法,其特征在于,所述根据所述虚拟模型的空间位置,对所述初始检测框进行更新,包括:
  - 计算所述虚拟模型的外接长方体,将计算得到的外接长方体转换到统一的坐标空间中,生成所述虚拟模型的检测框;
  - 将所述虚拟模型的检测框添加到所述初始检测框中,以便对所述初始检测框进行更新,生成包含虚拟模型的检测框。
6. 根据权利要求1所述的方法,其特征在于,所述利用所述新的多视角图像数据、更新

后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,包括:

确定预先构建的感知模型的模型结构及模型参数,并利用所述新的多视角图像数据、更新后的点云数据及更新后的检测框构建训练数据集;

将所述训练数据集输入到所述感知模型中,利用梯度下降算法对所述模型参数进行优化,通过最小化损失函数调整所述模型参数;

在模型训练过程中,反复利用梯度下降算法执行模型参数的优化操作,直至损失函数收敛,得到训练后的感知模型。

7.根据权利要求1所述的方法,其特征在于,所述将实际采集的车辆周围的多视角图像数据及点云数据,作为所述训练后的感知模型的输入,利用所述训练后的感知模型进行目标检测,并输出目标检测结果,包括:

将实际采集的车辆周围的多视角图像数据及点云数据输入到所述训练后的感知模型中;

在所述训练后的感知模型中,利用特征提取模块对数据进行特征提取;

将提取到的特征转换到统一的鸟瞰视角空间中,并对所述鸟瞰视角空间中的特征进行时序融合,得到时序相关的多模态特征;

将所述多模态特征输入到与预设任务相关的解码器中,利用所述解码器对所述多视角图像数据中的目标对象进行识别,并输出目标检测结果。

8.一种基于数据合成的目标检测装置,其特征在于,包括:

获取模块,被配置为获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;

渲染模块,被配置为将所述虚拟模型渲染至所述多视角图像数据中,并生成新的多视角图像数据;

更新模块,被配置为基于所述虚拟模型与所述原始点云数据的相互关系,对所述原始点云数据进行更新,以使更新后的点云数据中包含所述虚拟模型的点云信息;

生成模块,被配置为根据所述虚拟模型的空间位置,对所述初始检测框进行更新,以生成包含所述虚拟模型检测框;

训练模块,被配置为利用所述新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;

检测模块,被配置为将实际采集的车辆周围的多视角图像数据及点云数据,作为所述训练后的感知模型的输入,利用所述训练后的感知模型进行目标检测,并输出目标检测结果。

9.一种电子设备,包括存储器,处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现如权利要求1至7中任一项所述方法的步骤。

10.一种计算机可读存储介质,所述计算机可读存储介质存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求1至7中任一项所述方法的步骤。

## 基于数据合成的目标检测方法、装置、电子设备及介质

### 技术领域

[0001] 本申请涉及自动驾驶技术领域,尤其涉及一种基于数据合成的目标检测方法、装置、电子设备及介质。

### 背景技术

[0002] 感知作为自动驾驶系统中的关键部分,通过传感器数据对周边环境进行建模,包括识别障碍物的位置、大小、朝向和速度,以及车道线的位置、类型和长度等。感知模块的主要算法依赖于由人工智能(AI)主导的神经网络,这些神经网络的核心要素包括数据、算法和算力。在这三要素中,数据的质量直接决定了感知模块的性能。然而,在工业界生产环境中,数据标注费用极高,通常需要数百万至数十亿的投入。

[0003] 即使投入大量标注经费,感知系统仍难以完全覆盖长尾问题(即在自动驾驶过程中极少遇到但需准确识别的特殊情况)。解决这些长尾问题的成本往往与解决前99%问题的费用相当,因此,降低感知长尾问题的成本成为一个亟待解决的技术需求。

[0004] 目前,自动驾驶行业广泛采用的BEV(Bird's Eye View,鸟瞰视角)架构自Tesla AI-Day 2021发布以来,已经成为最主流的感知架构。与传统的后融合架构相比,BEV架构具备更强的伸缩性与可扩展性,完全依赖数据驱动。然而,作为一个数据驱动的感知架构,BEV架构需要大量高质量数据进行训练,这使得数据长尾问题变得更加严重,并进一步提高了数据标注的成本。

[0005] 因此,综上所述,现有技术自动驾驶感知系统中存在以下问题:数据标注成本高昂,难以覆盖所有场景;长尾问题处理费用巨大,解决效率低下;现有数据驱动架构(如BEV)对高质量数据需求量大,进一步加剧了数据标注的负担,导致模型的检测精度降低,准确性变差。

### 发明内容

[0006] 有鉴于此,本申请实施例提供了一种基于数据合成的目标检测方法、装置、电子设备及介质,以解决现有技术存在的感知数据标注成本高,数据长尾问题严重,导致模型检测精度和准确性降低的问题。

[0007] 本申请实施例的第一方面,提供了一种基于数据合成的目标检测方法,包括:获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;将虚拟模型渲染至多视角图像数据中,并生成新的多视角图像数据;基于虚拟模型与原始点云数据的相互关系,对原始点云数据进行更新,以使更新后的点云数据中包含虚拟模型的点云信息;根据虚拟模型的空间位置,对初始检测框进行更新,以生成包含虚拟模型的检测框;利用新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;将实际采集的车辆周围的多视角图像数据及点云数据,作为训练后的感知模型的输入,利用训练后的感知模型进行目标检测,并输出目标检测结果。

[0008] 本申请实施例的第二方面,提供了一种基于数据合成的目标检测装置,包括:获取

模块,被配置为获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;渲染模块,被配置为将虚拟模型渲染至多视角图像数据中,并生成新的多视角图像数据;更新模块,被配置为基于虚拟模型与原始点云数据的相互关系,对原始点云数据进行更新,以使更新后的点云数据中包含虚拟模型的点云信息;生成模块,被配置为根据虚拟模型的空间位置,对初始检测框进行更新,以生成包含虚拟模型的检测框;训练模块,被配置为利用新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;检测模块,被配置为将实际采集的车辆周围的多视角图像数据及点云数据,作为训练后的感知模型的输入,利用训练后的感知模型进行目标检测,并输出目标检测结果。

[0009] 本申请实施例的第三方面,提供了一种电子设备,包括存储器,处理器及存储在存储器上并可在处理器上运行的计算机程序,处理器执行计算机程序时实现上述方法的步骤。

[0010] 本申请实施例的第四方面,提供了一种计算机可读存储介质,该计算机可读存储介质存储有计算机程序,该计算机程序被处理器执行时实现上述方法的步骤。

[0011] 本申请实施例采用的上述至少一个技术方案能够达到以下有益效果:

通过获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;将虚拟模型渲染至多视角图像数据中,并生成新的多视角图像数据;基于虚拟模型与原始点云数据的相互关系,对原始点云数据进行更新,以使更新后的点云数据中包含虚拟模型的点云信息;根据虚拟模型的空间位置,对初始检测框进行更新,以生成包含虚拟模型的检测框;利用新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;将实际采集的车辆周围的多视角图像数据及点云数据,作为训练后的感知模型的输入,利用训练后的感知模型进行目标检测,并输出目标检测结果。本申请能够降低感知数据标注成本,提高模型检测精度和准确性。

## 附图说明

[0012] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其它的附图。

[0013] 图1是本申请实施例提供的基于数据合成的目标检测方法的流程示意图;

图2是本申请实施例提供的基于数据合成的目标检测装置的结构示意图;

图3是本申请实施例提供的电子设备的结构示意图。

## 具体实施方式

[0014] 以下描述中,为了说明而不是为了限定,提出了诸如特定系统结构、技术之类的具体细节,以便透彻理解本申请实施例。然而,本领域的技术人员应当清楚,在没有这些具体细节的其它实施例中也可以实现本申请。在其它情况中,省略对众所周知的系统、装置、电路以及方法的详细说明,以免不必要的细节妨碍本申请的描述。

[0015] 感知作为自动驾驶系统中重要的一部分,旨在基于传感器数据完成周边的环境建

模,如:给出障碍物的位置/大小/朝向/速度,给出车道线的位置/类型/长度等。感知模块中大部分算法是由AI主导的神经网络实现的,对于神经网络而言最核心的三要素为:数据、算法、以及算力。数据占据感知模块研发中最核心的位置,其质量直接决定感知模块的实际性能,在工业界的生产环境中,数据通常要花费数百万至数十亿的标注经费。即使投入海量的标注经费也很难覆盖感知的长尾问题,解决最后剩下1%的长尾问题可能需要跟其它99%的问题一样的费用,因而如何降低感知长尾问题的花费是一个亟待解决的问题。

[0016] 当前主流的感知架构是BEV(Bird's Eye View)架构,自Tesla AI-Day 2021以后, BEV成为自动驾驶行业最主流的感知架构,相对于过往的后融合架构, BEV 架构有着更强的伸缩性与可扩展性,是一个真正的数据驱动的架构。作为一个完全的数据驱动感知架构, BEV架构需要更多的数据,因而该架构面临更严重的数据长尾问题。

[0017] BEV架构通过多个视角的摄像头图像、激光雷达点云数据以及毫米波雷达数据进行环境建模。这种数据驱动的架构需要大量数据进行训练,而这些数据的标注成本较高且难以覆盖所有自动驾驶场景,导致模型在某些场景下表现不佳。

[0018] 因此,现有感知系统面临的主要问题是感知数据标注成本高,尤其是处理长尾问题时,费用更为巨大。本申请旨在通过计算机图形学合成感知3D检测模型的长尾训练样本,以降低数据标注的成本。为了更详细的讲述本申请技术方案,以下实施例将通过合成一个包含大巴车和皮卡车的训练样本来详细展示本申请提供的数据合成思路。

[0019] 下面结合附图以及具体实施例对本申请技术方案的内容进行详细描述。

[0020] 图1是本申请实施例提供的基于数据合成的目标检测方法的流程示意图。图1的基于数据合成的目标检测方法可以由自动驾驶系统来执行。如图1所示,该基于数据合成的目标检测方法具体可以包括:

S101,获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;

S102,将虚拟模型渲染至多视角图像数据中,并生成新的多视角图像数据;

S103,基于虚拟模型与原始点云数据的相互关系,对原始点云数据进行更新,以使更新后的点云数据中包含虚拟模型的点云信息;

S104,根据虚拟模型的空间位置,对初始检测框进行更新,以生成包含虚拟模型的检测框;

S105,利用新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;

S106,将实际采集的车辆周围的多视角图像数据及点云数据,作为训练后的感知模型的输入,利用训练后的感知模型进行目标检测,并输出目标检测结果。

[0021] 首先,本申请实施例给定环视 $n$ 视角的图像 $\{I_1, I_2, \dots, I_n\}$ 、Lidar点云 $P$ (即原始点云数据)、以及对应的检测框 $B$ (即初始检测框),对于美术建模的模型 $M$ (即虚拟模型),本申请使用渲染的方式将模型 $M$ 放置到场景中并更新相应的图像 $\{I_1, I_2, \dots, I_n\}$ 、点云 $P$ 、以及3D检测框 $B$ ,使得更新后的环视图像 $\{I_1, I_2, \dots, I_n\}$ 中能看到加入的模型素材 $M$ 、Lidar点云 $P$ 中有模型 $M$ 的成像点、以及检测框 $B$ 中包含模型 $M$ 的检测框。

[0022] 在一些实施例中,将虚拟模型渲染至多视角图像数据中,并生成新的多视角图像

数据,包括:

利用计算机图形学方法将虚拟模型渲染到每个摄像头对应的视角空间中;

将虚拟模型与原始的多视角图像数据进行叠加,生成包含虚拟模型的多视角图像数据;

将所有视角对应的包含虚拟模型的多视角图像数据进行校正和融合,得到新的多视角图像数据。

[0023] 具体地,利用计算机图形学方法,将虚拟模型渲染到每个摄像头对应的视角空间中。这一过程包括为每个摄像头视角创建一个虚拟场景,并在该场景中插入预先创建的虚拟模型。例如,使用3D建模软件创建一个虚拟汽车模型,并将该模型放置在摄像头视角的适当位置。

[0024] 进一步地,将虚拟模型与原始的多视角图像数据进行叠加,生成包含虚拟模型的多视角图像数据。具体地,基于计算机图形学技术,将模型M(虚拟模型)渲染到每一个摄像头相关的空间中,并将该虚拟模型的图像与对应摄像头的原始图像进行叠加处理。这样可以在原始图像上准确地展示虚拟模型的存在。

[0025] 例如,在一个示例中,将一个虚拟汽车模型渲染到前视相机中,使其出现在前视图图像数据中,还可以将同一虚拟汽车模型在所有6个环视相机中成像,以确保虚拟模型在所有视角下的一致性和连续性。

[0026] 进一步地,对所有视角对应的包含虚拟模型的多视角图像数据进行校正和融合,得到新的多视角图像数据。校正步骤包括调整各个视角下虚拟模型的几何形状和位置,以确保虚拟模型在不同视角下的显示效果一致。融合步骤包括将校正后的多视角图像数据进行整合,生成统一的、多视角的图像数据集。这一过程可以利用图像处理技术,如多视角立体匹配和图像拼接技术,以确保最终生成的多视角图像数据在视觉上无缝衔接,且符合真实世界的物理规律。

[0027] 例如,在一个示例中,可以将一个虚拟汽车模型渲染到前视相机中,生成前视图图像数据,并将该虚拟汽车模型在所有6个环视相机中成像,生成对应的环视图图像数据。

[0028] 通过上述本实施例的方法,可以得到包含虚拟模型的多视角图像数据,基于本申请中提出的方式合成的3D汽车模型,具有高度的真实性和一致性,能够有效模拟真实世界中的车辆场景,提高感知模型的训练效果。

[0029] 在一些实施例中,基于虚拟模型与原始点云数据的相互关系,对原始点云数据进行更新,包括:

在场景中插入虚拟模型后,确定原始点云数据中被虚拟模型遮挡的点,并将原始点云数据中被虚拟模型遮挡的点删除;

计算每条激光雷达线束与虚拟模型的交点,将最近的交点作为新增的点云数据,将新增的点云数据与原始点云数据融合,生成包含虚拟模型点云数据。

[0030] 具体地,在场景中插入预先创建的3D虚拟模型。例如,将一个虚拟的小汽车模型和一个大巴车模型插入到原始点云数据的场景中。这些虚拟模型应与原始点云数据中的实际环境保持一致,以确保更新后的点云数据的真实性和准确性。

[0031] 进一步地,在插入虚拟模型后,确定原始点云数据中被虚拟模型遮挡的点。具体地,判断激光雷达设备的原点与原始点云中每个点的射线与虚拟模型每个面片的相交性。

如果射线与虚拟模型的某个面片相交,则认为该点被遮挡,需从原始点云数据中删除。通过这一过程,可以确保更新后的点云数据不包含被虚拟模型遮挡的无效点。

[0032] 进一步地,计算每条激光雷达线束与虚拟模型的交点,将最近的交点作为新增的点云数据。例如,对每条激光雷达线束,求解其与虚拟模型的相交点,并将这些相交点加入到点云数据中。新增的点云数据应与虚拟模型的表面特征匹配,以确保更新后的点云数据能够准确反映虚拟模型的形状和位置。

[0033] 进一步地,将新增的点云数据与原始点云数据融合,生成包含虚拟模型的更新点云数据。融合过程中需要确保新增点云数据与原始点云数据在空间上的一致性和连续性,从而得到完整、准确的点云数据。例如,在原始点云数据中新增小汽车和大巴车后的新点云数据。

[0034] 例如,在一个示例中,当场景中插入3D虚拟模型后,原有的点云中与被遮挡的点应该删除,同时Lidar打到新增加物体上的点应该加入到点云中。为了删除原始点云中的点,可以判断Lidar设备的原点与原始点云中每个点的射线与虚拟模型每个面片的相交性。如果射线与虚拟模型的某个面片相交,则该点被遮挡,需删除;否则,保留该点。同时,为了使数据更真实,还需要增加由该虚拟物体所引入的新的点云数据。为实现该目的,可以求解每条Lidar线束与虚拟模型的相交点,将最近的交点更新到点云数据中。

[0035] 通过上述本实施例的方法,可以有效地将虚拟模型融入到原始点云数据中,生成包含虚拟模型的更新点云数据。这种方法能够提高感知系统对虚拟目标的识别精度,增强感知模型在复杂场景下的适应性。

[0036] 在一些实施例中,确定原始点云数据中被虚拟模型遮挡的点,包括:

确定激光雷达设备的原点与原始点云数据中每个点的射线,根据射线与虚拟模型的每个面片之间的相交性,判断是否为原始点云数据中被虚拟模型遮挡的点;

其中,当射线与虚拟模型中的面片相交时,判断原始点云数据中的点属于被虚拟模型遮挡的点,否则,判断原始点云数据中的点不属于被虚拟模型遮挡的点。

[0037] 具体地,首先,从激光雷达设备的原点发出射线,并确定这些射线与原始点云数据中每个点的路径。假设激光雷达设备的原点为 $O$ ,原始点云数据中的每个点为 $P_i$ ,则每条射线可以表示为从 $O$ 到 $P_i$ 的直线。

[0038] 进一步地,计算射线与虚拟模型的每个面片的相交性:对于每条从激光雷达设备原点 $O$ 到点 $P_i$ 的射线,判断其是否与虚拟模型 $M$ 的面片相交。虚拟模型 $M$ 通常由多个面片(如三角形或四边形)组成,可以通过计算射线与这些面片的相交性来判断点 $P_i$ 是否被虚拟模型遮挡。

[0039] 例如,可以采用几何算法来求解射线与多边形面片的相交问题。例如,利用射线-三角形相交测试算法来判断射线是否与虚拟模型 $M$ 的三角形面片相交。

[0040] 进一步地,当射线与虚拟模型 $M$ 中的面片相交时,判断原始点云数据中的点 $P_i$ 属于被虚拟模型遮挡的点,需从原始点云数据中删除该点。否则,判断点 $P_i$ 不属于被虚拟模型遮挡的点,保留该点。

[0041] 通过上述判断,可以更新原始点云数据,删除被虚拟模型遮挡的点。具体地,为了删除原始点云中的点,可以执行以下步骤:

从激光雷达设备的原点发出射线,确定这些射线与原始点云数据中每个点的路

径。

[0042] 计算每条射线与虚拟模型M的每个面片的相交性。若射线与M的某个面片相交,则认为点云数据中的该点被遮挡,需删除该点;否则,保留该点。

[0043] 例如,在一个具体示例中,假设激光雷达设备的原点为O,原始点云数据中的点为 $P_1, P_2, \dots, P_n$ ,虚拟模型为M,其由多个三角形面片组成。对于每个点 $P_i$ 计算从O到 $P_i$ 的射线路径,判断该射线是否与M的任意一个三角形面片相交。如果相交,则删除点 $P_i$ ;否则,保留点 $P_i$ 。

[0044] 通过上述本实施例的方法,可以有效地确定原始点云数据中被虚拟模型遮挡的点,并将其删除,从而更新点云数据,使其更加真实和准确。

[0045] 在一些实施例中,根据虚拟模型的空间位置,对初始检测框进行更新,包括:

计算虚拟模型的外接长方体,将计算得到的外接长方体转换到统一的坐标空间中,生成虚拟模型检测框;

将虚拟模型检测框添加到初始检测框中,以便对初始检测框进行更新,生成包含虚拟模型检测框。

[0046] 具体地,给定待插入的3D虚拟模型M,首先计算该模型的外接长方体。外接长方体是完全包围虚拟模型M的最小长方体,其每个面与虚拟模型的边界相切。为了计算外接长方体,可以利用虚拟模型M的所有顶点坐标,确定其在各个坐标轴上的最小和最大值,从而定义外接长方体的六个面。

[0047] 进一步地,将计算得到的外接长方体转换到统一的坐标空间中。统一的坐标空间是整个感知系统所使用的参考坐标系,可以确保不同数据源(如摄像头、激光雷达)的数据在同一个坐标系下进行处理和融合。具体地,可以通过坐标变换矩阵将虚拟模型M的外接长方体从其本地坐标系转换到全局坐标系。

[0048] 进一步地,在统一坐标空间中,生成虚拟模型检测框。检测框是一个矩形框,用于标识虚拟模型在空间中的位置和尺寸。此步骤确保虚拟模型M在全局坐标系下有一个明确的空间表示,可以被后续的感知处理模块识别和处理。

[0049] 进一步地,将生成的虚拟模型检测框添加到初始检测框B中。初始检测框B包含了原始点云数据和多视角图像数据中所有已知物体的检测框。通过将虚拟模型检测框加入到初始检测框B中,生成包含虚拟模型检测框集合。这一过程确保所有的检测框在同一个坐标系下进行更新和处理。

[0050] 例如,在一个示例中,假设给定的3D虚拟模型M,初始检测框的更新包括以下步骤:

计算虚拟模型M的外接长方体,例如,通过获取模型M在x、y、z三个坐标轴上的最小和最大坐标值,从而定义外接长方体的六个面。

[0051] 利用坐标变换矩阵,将外接长方体从虚拟模型M的本地坐标系转换到感知系统的全局坐标系中。此步骤可能涉及平移、旋转和缩放等变换操作,以确保虚拟模型检测框在全局坐标系下的正确位置和方向。

[0052] 在全局坐标系中生成虚拟模型M检测框,该检测框用于标识虚拟模型在空间中的位置和尺寸。

[0053] 将生成的虚拟模型检测框添加到初始检测框B中,更新初始检测框B,得到包含虚拟模型检测框集合。

[0054] 通过上述本实施例的方法,可以有效地将虚拟模型的检测框与原始检测框进行融合,生成包含虚拟模型的更新检测框。这种方法可以提高感知系统对虚拟目标的检测能力,增强系统在复杂场景下的适应性。

[0055] 在一些实施例中,利用新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,包括:

确定预先构建的感知模型的模型结构及模型参数,并利用新的多视角图像数据、更新后的点云数据及更新后的检测框构建训练数据集;

将训练数据集输入到感知模型中,利用梯度下降算法对模型参数进行优化,通过最小化损失函数调整模型参数;

在模型训练过程中,反复利用梯度下降算法执行模型参数的优化操作,直至损失函数收敛,得到训练后的感知模型。

[0056] 具体地,首先确定预先构建的感知模型的模型结构 $G$ 及初始模型参数 $\theta$ 。模型结构 $G$ 包括感知模型的神经网络架构,如卷积层、池化层、全连接层等。初始模型参数 $\theta$ 包括这些层中的权重和偏置。

[0057] 进一步地,利用新的多视角图像数据、更新后的点云数据及更新后的检测框,构建训练数据集。具体地,新的多视角图像数据作为模型的输入数据,更新后的点云数据和检测框作为模型的标签数据。训练数据集由多个样本对 $(x_i, y_i)$ 组成,其中 $x_i$ 是模型的输入数据(如多视角图像和点云数据), $y_i$ 是对应的标签数据(如检测框)。

[0058] 进一步地,将训练数据集输入到感知模型中,利用梯度下降算法对模型参数进行优化。具体地,使用随机梯度下降算法(SGD)来最小化损失函数,从而调整模型参数 $\theta$ 。损失函数用于衡量模型预测结果与实际标签数据之间的差异,常用的损失函数包括交叉熵损失和均方误差等。

[0059] 进一步地,在模型训练过程中,反复利用梯度下降算法执行模型参数的优化操作。具体步骤如下:

计算模型的预测输出,并与标签数据比较,得到损失值。

[0060] 计算损失函数相对于模型参数的梯度,利用反向传播算法更新模型参数。

[0061] 重复上述步骤,迭代更新模型参数,直至损失函数收敛或达到预设的训练轮数。

[0062] 当损失函数收敛后,得到训练后的感知模型。此时,模型参数 $\theta$ 已经经过优化,能够在新的多视角图像数据和点云数据中进行准确的目标检测。

[0063] 具体地,给定模型结构 $G$ 及模型参数 $\theta$ ,模型训练的主要任务是优化参数 $\theta$ 的值。为了实现该目的,通常需要一些带标签的训练数据 $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ ,并使用梯度下降算法SGD(Stochastic Gradient Descent)优化模型权重 $\theta$ ,其中 $x_i$ 是模型的输入, $y_i$ 是人工标注的标签。训练过程包括以下步骤:

输入数据:将训练数据集 $\{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ 输入到感知模型中。

[0064] 计算损失:模型预测输出与实际标签数据 $y_i$ 比较,计算损失函数值。

[0065] 计算梯度:通过反向传播算法,计算损失函数相对于模型参数 $\theta$ 的梯度。

[0066] 更新参数:利用梯度下降算法SGD,根据梯度更新模型参数 $\theta$ 。

[0067] 迭代训练:重复上述步骤,直至损失函数收敛,得到优化后的模型参数 $\theta$ 。

[0068] 通过上述本实施例的方法,可以利用新的多视角图像数据、更新后的点云数据及更新后的检测框,对预先构建的感知模型进行有效训练,提升模型的目标检测准确性和鲁棒性。

[0069] 在一些实施例中,将实际采集的车辆周围的多视角图像数据及点云数据,作为训练后的感知模型的输入,利用训练后的感知模型进行目标检测,并输出目标检测结果,包括:

将实际采集的车辆周围的多视角图像数据及点云数据输入到训练后的感知模型中;

在训练后的感知模型中,利用特征提取模块对数据进行特征提取;

将提取到的特征转换到统一的鸟瞰视角空间中,并对鸟瞰视角空间中的特征进行时序融合,得到时序相关的多模态特征;

将多模态特征输入到与预设任务相关的解码器中,利用解码器对多视角图像数据中的目标对象进行识别,并输出目标检测结果。

[0070] 具体地,将实际采集到的车辆周围的多视角图像数据及点云数据输入到训练后的感知模型中。多视角图像数据通常包括多个摄像头的环视图像,点云数据则来自多个激光雷达(Lidar)传感器。

[0071] 进一步地,在训练后的感知模型中,利用特征提取模块对输入数据进行特征提取。特征提取模块包括多个卷积神经网络(CNN)层,用于从多视角图像和点云数据中提取高维特征表示。这些特征表示包含了输入数据的空间和语义信息。

[0072] 进一步地,将提取到的特征转换到统一的鸟瞰视角空间(BEV)。BEV模型架构通过几何变换将来自不同传感器的数据映射到统一的二维平面上,从而实现数据的统一表示。具体地,可以利用摄像头的内外参数进行图像特征的投影,并利用点云数据的空间坐标进行点云特征的映射。

[0073] 进一步地,对鸟瞰视角空间中的特征进行时序融合,得到时序相关的多模态特征。时序融合模块通过多帧数据的联合处理,捕捉动态变化信息,提升特征表示的时效性和稳定性。这一过程可以通过循环神经网络(RNN)或长短时记忆网络(LSTM)来实现。

[0074] 进一步地,将多模态特征输入到与预设任务相关的解码器中,利用解码器对多视角图像数据中的目标对象进行识别。解码器根据具体感知任务的需求,采用不同的神经网络结构,如卷积层、全连接层等,输出目标检测结果。常见的感知任务包括3D目标检测、车道线识别和占用网格识别等。

[0075] 在一个示例中,本申请实施例的感知模型采用BEV模型架构,其输入一般包括环视的多视角图像、多个激光雷达点云数据及毫米波/超声波雷达数据。模型通过各个特征提取模块提取相应传感器数据的特征,将这些特征转换到统一的BEV空间(即鸟瞰视角空间)。例如,通过几何变换和投影,将多视角图像数据和点云数据映射到BEV空间中。

[0076] 在BEV空间中,感知模型对特征进行时序融合,利用多帧数据增强特征的时效性。随后,时序相关的BEV多模态特征被输入到任务相关的解码器中。对于3D目标检测任务,解码器识别并输出目标对象的三维位置、大小和类型信息;对于车道线识别任务,解码器输出车道线的位置和类型信息;对于占用网格识别任务,解码器输出车辆周围的占用情况。

[0077] 通过上述本实施例的方法,训练后的感知模型能够利用多视角图像数据和点云数

据进行准确的目标检测,并输出高精度的检测结果。该方法提高了感知系统在复杂驾驶环境中的性能和可靠性。

[0078] 根据本申请实施例提供的技术方案,本申请技术方案至少具有以下优点:

本申请所提出的方法可以有效地合成大量的长尾训练样本,与传统的2D层面的复制-粘贴合成方式相比,本申请的方法更符合真实的物理规律。具体而言,本申请的合成过程完全遵循客观的物理逻辑,通过将虚拟模型渲染到多视角图像数据和点云数据中,实现了与真实场景的高度一致性。

[0079] 首先,通过计算机图形学方法将虚拟模型精确地渲染到多视角图像数据中,并根据虚拟模型的位置和形状对原始点云数据进行更新,生成包含虚拟模型的多模态数据。这种方法确保了虚拟模型在不同视角下的一致性和连续性,使得合成数据更具真实性。

[0080] 其次,与完全虚拟渲染合成的方式相比,本申请的合成训练样本除了新增加的目标是渲染的之外,整个场景完全基于真实传感器采集的数据。由于整个场景中的背景和环境数据是真实的传感器数据,本申请合成的训练样本与真实数据的特性更加接近,从而提高了感知模型的训练效果和实用性。

[0081] 另外,本申请提出的方法能够有效解决自动驾驶训练数据中的长尾样本问题,显著降低自动驾驶训练样本的标注成本。通过合成高质量的长尾训练样本,可以减少人工标注的工作量和费用,提升数据覆盖范围和模型的泛化能力。这种高效的数据合成方法不仅提高了感知系统的性能,还加速了自动驾驶技术的实际应用和落地进程。

[0082] 综上所述,本申请的方法通过合成符合物理规律的高质量训练样本,不仅增强了自动驾驶感知系统的能力,还降低了训练数据的标注成本,为自动驾驶技术的推广和应用提供了重要支持。

[0083] 窗体底端下述为本申请装置实施例,可以用于执行本申请方法实施例。对于本申请装置实施例中未披露的细节,请参照本申请方法实施例。图2是本申请实施例提供的基于数据合成的目标检测装置的结构示意图。如图2所示,该基于数据合成的目标检测装置包括:

获取模块201,被配置为获取多视角图像数据、原始点云数据、初始检测框以及预先创建的虚拟模型;

渲染模块202,被配置为将虚拟模型渲染至多视角图像数据中,并生成新的多视角图像数据;

更新模块203,被配置为基于虚拟模型与原始点云数据的相互关系,对原始点云数据进行更新,以使更新后的点云数据中包含虚拟模型的点云信息;

生成模块204,被配置为根据虚拟模型的空间位置,对初始检测框进行更新,以生成包含虚拟模型的检测框;

训练模块205,被配置为利用新的多视角图像数据、更新后的点云数据以及更新后的检测框,对预先构建的感知模型进行训练,得到训练后的感知模型;

检测模块206,被配置为将实际采集的车辆周围的多视角图像数据及点云数据,作为训练后的感知模型的输入,利用训练后的感知模型进行目标检测,并输出目标检测结果。

[0084] 在一些实施例中,图2的渲染模块202利用计算机图形学方法将虚拟模型渲染到每个摄像头对应的视角空间中;将虚拟模型与原始的多视角图像数据进行叠加,生成包含虚

拟模型的多视角图像数据;将所有视角对应的包含虚拟模型的多视角图像数据进行校正和融合,得到新的多视角图像数据。

[0085] 在一些实施例中,图2的更新模块203在场景中插入虚拟模型后,确定原始点云数据中被虚拟模型遮挡的点,并将原始点云数据中被虚拟模型遮挡的点删除;计算每条激光雷达线束与虚拟模型的交点,将最近的交点作为新增的点云数据,将新增的点云数据与原始点云数据融合,生成包含虚拟模型的点云数据。

[0086] 在一些实施例中,图2的更新模块203确定激光雷达设备的原点与原始点云数据中每个点的射线,根据射线与虚拟模型的每个面片之间的相交性,判断是否为原始点云数据中被虚拟模型遮挡的点;其中,当射线与虚拟模型中的面片相交时,判断原始点云数据中的点属于被虚拟模型遮挡的点,否则,判断原始点云数据中的点不属于被虚拟模型遮挡的点。

[0087] 在一些实施例中,图2的生成模块204计算虚拟模型的外接长方体,将计算得到的外接长方体转换到统一的坐标空间中,生成虚拟模型的检测框;将虚拟模型的检测框添加到初始检测框中,以便对初始检测框进行更新,生成包含虚拟模型的检测框。

[0088] 在一些实施例中,图2的训练模块205确定预先构建的感知模型的模型结构及模型参数,并利用新的多视角图像数据、更新后的点云数据及更新后的检测框构建训练数据集;将训练数据集输入到感知模型中,利用梯度下降算法对模型参数进行优化,通过最小化损失函数调整模型参数;在模型训练过程中,反复利用梯度下降算法执行模型参数的优化操作,直至损失函数收敛,得到训练后的感知模型。

[0089] 在一些实施例中,图2的检测模块206将实际采集的车辆周围的多视角图像数据及点云数据输入到训练后的感知模型中;在训练后的感知模型中,利用特征提取模块对数据进行特征提取;将提取到的特征转换到统一的鸟瞰视角空间中,并对鸟瞰视角空间中的特征进行时序融合,得到时序相关的多模态特征;将多模态特征输入到与预设任务相关的解码器中,利用解码器对多视角图像数据中的目标对象进行识别,并输出目标检测结果。

[0090] 应理解,上述实施例中各步骤的序号的大小并不意味着执行顺序的先后,各过程的执行顺序应以其功能和内在逻辑确定,而不应对本申请实施例的实施过程构成任何限定。

[0091] 图3是本申请实施例提供的电子设备3的结构示意图。如图3所示,该实施例的电子设备3包括:处理器301、存储器302以及存储在该存储器302中并且可以在处理器301上运行的计算机程序303。处理器301执行计算机程序303时实现上述各个方法实施例中的步骤。或者,处理器301执行计算机程序303时实现上述各装置实施例中各模块/单元的功能。

[0092] 示例性地,计算机程序303可以被分割成一个或多个模块/单元,一个或多个模块/单元被存储在存储器302中,并由处理器301执行,以完成本申请。一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述计算机程序303在电子设备3中的执行过程。

[0093] 电子设备3可以是桌上型计算机、笔记本、掌上电脑及云端服务器等电子设备。电子设备3可以包括但不限于处理器301和存储器302。本领域技术人员可以理解,图3仅仅是电子设备3的示例,并不构成对电子设备3的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件,例如,电子设备还可以包括输入输出设备、网络接入设备、总线等。

[0094] 处理器301可以是中央处理单元(Central Processing Unit,CPU),也可以是其它通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)或者其它可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0095] 存储器302可以是电子设备3的内部存储单元,例如,电子设备3的硬盘或内存。存储器302也可以是电子设备3的外部存储设备,例如,电子设备3上配备的插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)等。进一步地,存储器302还可以既包括电子设备3的内部存储单元也包括外部存储设备。存储器302用于存储计算机程序以及电子设备所需的其它程序和数据。存储器302还可以用于暂时地存储已经输出或者将要输出的数据。

[0096] 所属领域的技术人员可以清楚地了解到,为了描述的方便和简洁,仅以上述各功能单元、模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能单元、模块完成,即将装置的内部结构划分成不同的功能单元或模块,以完成以上描述的全部或者部分功能。实施例中的各功能单元、模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中,上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。另外,各功能单元、模块的具体名称也只是为了便于相互区分,并不用于限制本申请的保护范围。上述系统中单元、模块的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0097] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述或记载的部分,可以参见其它实施例的相关描述。

[0098] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0099] 在本申请所提供的实施例中,应该理解到,所揭露的装置/计算机设备和方法,可以通过其它的方式实现。例如,以上所描述的装置/计算机设备实施例仅仅是示意性的,例如,模块或单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通讯连接可以通过一些接口,装置或单元的间接耦合或通讯连接,可以是电性,机械或其它的形式。

[0100] 作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0101] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0102] 集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时,可以存储在一个计算机可读存储介质中。基于这样的理解,本申请实现上述实施例方法中的全部或部分流程,也可以通过计算机程序来指令相关的硬件来完成,计算机程序可以存储在计算机可读存储介质中,该计算机程序在被处理器执行时,可以实现上述各个方法实施例的步骤。计算机程序可以包括计算机程序代码,计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。计算机可读介质可以包括:能够携带计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory,RAM)、电载波信号、电信信号以及软件分发介质等。

[0103] 以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围,均应包含在本申请的保护范围之内。

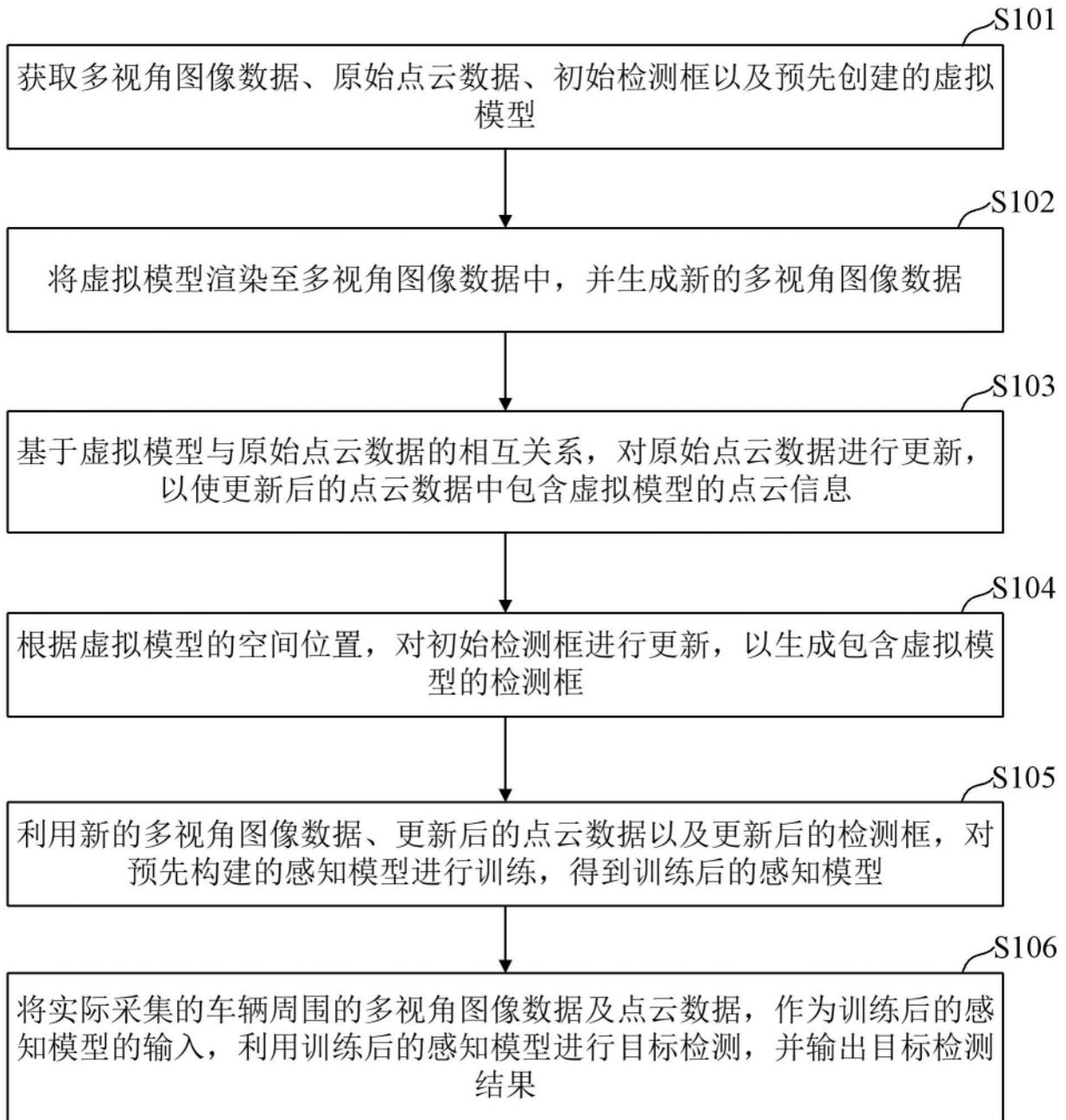


图1

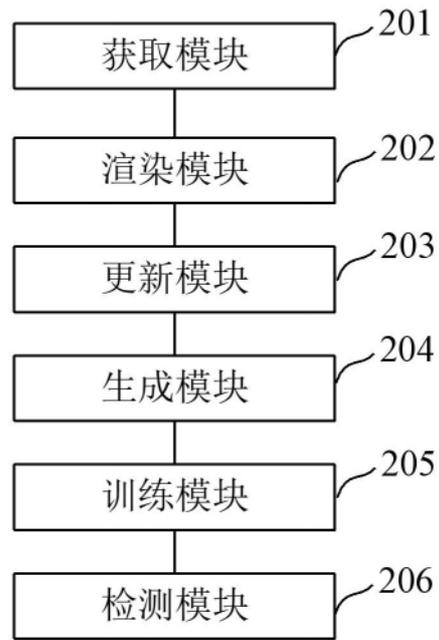


图2

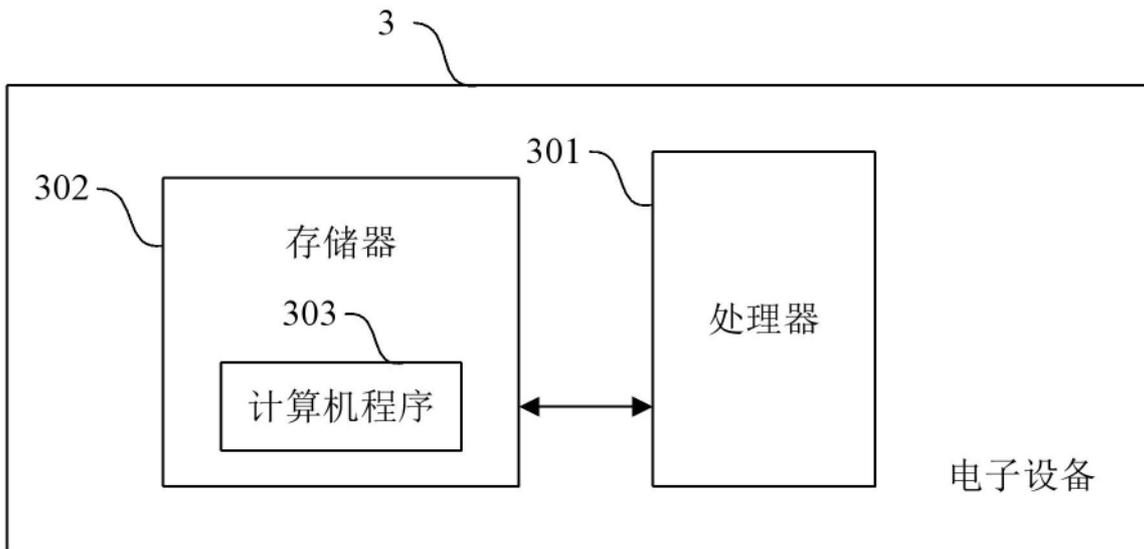


图3