

三维目标检测方法及装置

本公开要求在 2020 年 4 月 3 日提交中国专利局、申请号为 202010258375.3、申请名称为“三维目标检测方法及装置”的中国专利申请的优先权，其全部内容通过引用结合在本公开中。

技术领域

本公开涉及图像检测技术领域，具体而言，涉及一种三维目标检测方法、行驶控制方法、装置、电子设备及存储介质。

背景技术

近年来，三维 3D 物体检测是计算机视觉中一个非常重要的任务，比如，3D 物体检测可以应用在无人驾驶等场景中。其中，对 3D 物体检测时，可以通过激光雷达获取视野范围内的点云数据，进而可以将获取的点云数据输入神经网络得到视野范围内所有对象的三维位姿数据。

发明内容

有鉴于此，本公开至少提供一种三维目标检测方法、行驶控制方法、装置、电子设备及存储介质。

第一方面，本公开提供了一种三维目标检测方法，包括：

获取双目图像；所述双目图像包括左目图像和右目图像；分别对所述左目图像和右目图像进行目标检测，得到待检测对象在所述左目图像中的第一区域图像和该待检测对象在所述右目图像中的第二区域图像；基于所述待检测对象对应的所述第一区域图像和所述第二区域图像，生成所述待检测对象的目标视差图，并基于所述待检测对象的目标视差图、以及所述第一区域图像和所述第二区域图像，计算得到所述待检测对象的点云数据；基于所述点云数据，确定所述待检测对象的三维位姿数据。

采用上述方法，通过从左目图像中得到待检测对象的第一区域图像以及从右目图像中得到待检测对象的第二区域图像，并基于第一区域图像以及第二区域图像，确定每个待检测对象的目标视差图，进而基于该目标视差图，计算得到每个待检测对象的点云数据，采用这种基于视差图的方式生成点云数据，不需要依赖激光雷达获取点云数据，可以生成较为稠密的点云数据，进而可以基于视差图生成的点云数据确定待检测对象的三维位姿数据。

一种可能的实施方式中，在生成所述待检测图像的目标视差图之前，所述方法还包括：

基于所述第一区域图像与所述第二区域图像，确定横坐标调整值；

基于所述横坐标调整值，调整所述第一区域图像或所述第二区域图像中像素点的横

坐标值,使得所述第一区域图像与所述第二区域图像中对应边界点的横坐标值重合。

上述实施方式中,通过确定横坐标调整值,该横坐标调整值可以表征第一区域图像与第二区域图像之间的粗略的视差值,进而可以基于该横坐标调整值粗略的对第一区域图像和第二区域图像之间的视差值进行修正,使得后续在基于调整后的第一区域图像和
5 所述第二区域图像,生成每个待检测对象的目标视差图时,可以减少区域图像搜索范围,以及减少目标视差图确定过程的运算量。

一种可能的实施方式中,所述基于所述第一区域图像和所述第二区域图像,生成所述目标视差图,包括:

10 确定所述待检测对象在所述第一区域图像或所述第二区域图像中的轮廓图像,并基于所述第一区域图像和所述第二区域图像,生成所述待检测对象对应的区域视差图;

基于所述轮廓图像,从所述区域视差图中截取所述目标视差图。

上述实施方式下,通过待检测对象的轮廓图像,可以从区域视差图中截取待检测对象的目标视差图,进一步保留跟目标相关的有用的视差图,可以使得基于目标视差图确定的待检测对象的点云数据可以较好的表征待检测对象,进而可以提高检测的精准度。

15 一种可能的实施方式中,基于所述待检测对象的所述第一区域图像和所述第二区域图像,生成所述待检测对象对应的区域视差图,包括:

将所述第一区域图像和第二区域图像输入至训练的视差估计网络中进行视差估计,得到所述区域视差图。

20 上述实施方式中,通过训练的视差估计网络,确定待检测对象的区域视差图,在保证得到的区域视差图的准确度的同时,可以较快速的确定待检测对象的区域视差图。

一种可能的实施方式中,通过下述步骤对视差估计网络进行训练:

获取同一类型下不同待检测对象的形状数据、多组标注好待检测对象的三维位姿数据的双目图像样本,以及所述双目图像样本对应的点云数据;所述双目图像样本包括左目图像样本和右目图像样本;

25 基于所述同一类型下不同待检测对象的形状数据,确定该类型下的待检测对象的平均形状数据;

针对所述双目图像样本,基于所述待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据,得到该组双目图像样本对应的目标视差图;

30 将所述双目图像样本以及所述双目图像样本对应的目标视差图作为训练样本,对所述视差估计网络进行训练,得到训练后的视差估计网络。

这里,可以基于平均形状数据、每组双目图像样本中包括的待检测对象的三维位姿数据、以及每组双目图像样本对应的点云数据,确定每组双目图像样本对应的目标视差图,基于上述实施方式得到的目标视差图可以更准确地表征待检测对象的视差信息;进

一步的,通过双目图像样本以及每组双目图像样本对应的目标视差图构成的训练样本对视差估计网络训练时,得到的视差估计网络的准确度更高。

- 一种可能的实施方式中,针对一组双目图像样本,基于所述待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据,得到该组双目图像样本对应的目标视差图,包括:

针对一组双目图像样本,生成变换后的双目图像样本,其中,所述变换后的双目图像样本与所述平均形状数据以及所述三维位姿数据相匹配;

基于所述双目图像样本对应的点云数据、以及所述变换后的双目图像样本,生成所述双目图像样本对应的目标视差图。

- 上述实施方式中,在不改变双目图像样本中待检测对象的三维位姿数据的前提下,生成变换后的双目图像样本,使得变换后双目图像样本中的形状数据与平均形状数据匹配,由于平均形状数据可以表征待检测对象的形状特征,且物体形状表面具有连续性,从而基于双目图像样本对应的点云数据、以及变换后的双目图像样本,得到的双目图像样本对应的目标视差图更准确和稠密。

- 一种可能的实施方式中,所述基于所述双目图像样本对应的点云数据、以及所述变换后的双目图像样本,生成所述双目图像样本对应的目标视差图,包括:

基于所述双目图像样本对应的点云数据,对所述变换后的双目图像样本的形状数据进行调整,得到调整后的双目图像样本;其中,调整后的双目图像样本的形状数据与所述点云数据所对应的形状相匹配;

- 基于所述调整后的双目图像样本,生成所述双目图像样本对应的深度图;

基于所述深度图,生成该组双目图像样本对应的目标视差图。

- 在上述实施方式中,通过基于双目图像样本对应的点云数据,对变换后的双目图像样本的形状数据进行调整,使得调整后的形状数据对应的形状与双目图像样本变换前包含的形状数据对应的形状相符,即使得调整后的形状数据与双目图像样本中待检测对象的真实形状相符,进而基于调整后的双目图像样本,生成的双目图像样本对应的深度图的准确度和稠密度较高,进而可以提高生成的目标视差图的准确度。

一种可能的实施方式中,所述三维位姿数据包括以下信息中的至少一种:

所述待检测对象的三维检测框的长度、高度、宽度、所述三维检测框的中心点的坐标信息、以及所述三维检测框的基准面与设置的基准线的夹角。

- 第二方面,本公开提供了一种行驶控制方法,包括:

获取行驶装置在行驶过程中采集的双目道路图像;

利用第一方面所述的三维目标检测方法对所述双目道路图像进行三维目标检测,得到所述双目道路图像中包括的目标对象的三维位姿数据;

基于所述双目道路图像中包括的目标对象的三维位姿数据,控制所述行驶装置。

以下装置、电子设备等的效果描述参见上述方法的说明，这里不再赘述。

第三方面，本公开提供了一种三维目标检测装置，包括：

双目图像获取模块，用于获取双目图像；所述双目图像包括左目图像和右目图像；

5 检测模块，用于分别对所述左目图像和右目图像进行目标检测，得到待检测对象在所述左目图像中的第一区域图像和该待检测对象在所述右目图像中的第二区域图像；

点云数据确定模块，用于基于所述第一区域图像和所述第二区域图像，生成所述待检测对象的目标视差图，并基于所述待检测对象的目标视差图、以及所述第一区域图像和所述第二区域图像，计算得到所述待检测对象的点云数据；

10 三维位姿数据确定模块，用于基于所述点云数据，确定所述待检测对象的三维位姿数据。

一种可能的实施方式中，所述装置还包括：

横坐标调整值确定模块，用于基于所述第一区域图像与所述第二区域图像，确定横坐标调整值；

15 调整模块，用于基于所述横坐标调整值，调整所述第一区域图像或所述第二区域图像中像素点的横坐标值，使得所述第一区域图像与所述第二区域图像中对应边界点的横坐标值重合。

一种可能的实施方式中，所述点云数据确定模块，在基于所述第一区域图像和所述第二区域图像，生成所述目标视差图的情况下，用于：

20 确定所述待检测对象在所述第一区域图像或所述第二区域图像中的轮廓图像，并基于所述第一区域图像和所述第二区域图像，生成所述待检测对象对应的区域视差图；

基于所述轮廓图像，从所述区域视差图中截取所述目标视差图。

一种可能的实施方式中，所述点云数据确定模块，在基于所述第一区域图像和所述第二区域图像，生成所述待检测对象对应的区域视差图的情况下，用于：

25 将所述第一区域图像和第二区域图像输入至视差估计网络中进行视差估计，得到所述区域视差图。

一种可能的实施方式中，所述装置还包括：网络训练模块；

所述网络训练模块，用于通过下述步骤对视差估计网络进行训练：

30 获取同一类型下不同待检测对象的形状数据、多组标注好待检测对象的三维位姿数据的双目图像样本，以及所述双目图像样本对应的点云数据；所述双目图像样本包括左目图像样本和右目图像样本；

基于所述同一类型下不同待检测对象的形状数据，确定该类型下的待检测对象的平均形状数据；

针对一组双目图像样本，基于所述待检测对象的平均形状数据、该组双目图像样本

中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据，得到该组双目图像样本对应的目标视差图；

将所述双目图像样本以及所述双目图像样本对应的目标视差图作为训练样本，对所述视差估计网络进行训练，得到训练后的视差估计网络。

- 5 一种可能的实施方式中，所述网络训练模块，在针对一组双目图像样本，基于所述待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据，得到该组双目图像样本对应的目标视差图的情况下，用于：

- 10 针对一组双目图像样本，生成变换后的双目图像样本，其中，所述变换后的双目图像样本与所述平均形状数据以及所述三维位姿数据相匹配；

基于所述双目图像样本对应的点云数据、以及所述变换后的双目图像样本，生成所述双目图像样本对应的目标视差图。

- 15 一种可能的实施方式中，所述网络训练模块，在基于所述双目图像样本对应的点云数据、以及所述变换后的双目图像样本，生成所述双目图像样本对应的目标视差图的情况下，用于：

基于所述双目图像样本对应的点云数据，对所述变换后的双目图像样本的形状数据进行调整，得到调整后的双目图像样本；其中，调整后的双目图像样本的形状数据与所述点云数据所对应的形状相匹配；

基于所述调整后的双目图像样本，生成所述双目图像样本对应的深度图；

- 20 基于所述深度图，生成该组双目图像样本对应的目标视差图。

一种可能的实施方式中，所述三维位姿数据包括以下信息中的至少一种：

所述待检测对象的三维检测框的长度、高度、宽度、所述三维检测框的中心点的坐标信息、以及所述三维检测框的基准面与设置的基准线的夹角。

第四方面，本公开提供了一种行驶控制装置，包括：

- 25 道路图像获取模块，用于获取行驶装置在行驶过程中采集的双目道路图像；

三维目标检测模块，用于利用第一方面所述的三维目标检测方法对所述双目道路图像进行三维目标检测，得到所述双目道路图像中包括的目标对象的三维位姿数据；

控制模块，用于基于所述双目道路图像中包括的目标对象的三维位姿数据，控制所述行驶装置。

- 30 第五方面，本公开提供一种电子设备，包括：处理器、存储器和总线，所述存储器存储有所述处理器可执行的机器可读指令，当电子设备运行时，所述处理器与所述存储器之间通过总线通信，所述机器可读指令被所述处理器执行时执行如上述第一方面或任一实施方式所述的三维目标检测方法的步骤，或执行如上述第二方面或任一实施方式所述的行驶控制方法的步骤。

第六方面，本公开提供一种计算机可读存储介质，该计算机可读存储介质上存储有计算机程序，该计算机程序被处理器运行时执行如上述第一方面或任一实施方式所述的三维目标检测方法的步骤，或执行如上述第二方面或任一实施方式所述的行驶控制方法的步骤。

- 5 第七方面，本公开提供一种计算机程序，其中，所述计算机程序包括计算机可读代码，当所述计算机可读代码在电子设备中运行时，所述电子设备中的处理器执行上述第一方面或任一实施方式所述的三维目标检测方法的步骤，或执行如上述第二方面或任一实施方式所述的行驶控制方法的步骤。

10 为使本公开的上述目的、特征和优点能更明显易懂，下文特举较佳实施例，并配合所附附图，作详细说明如下。

附图说明

15 为了更清楚地说明本公开实施例的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，此处的附图被并入说明书中并构成本说明书中的一部分，这些附图示出了符合本公开的实施例，并与说明书一起用于说明本公开的技术方案。应当理解，以下附图仅示出了本公开的某些实施例，因此不应被看作是对范围的限定，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他相关的附图。

图 1 示出了本公开实施例所提供的一种三维目标检测方法的流程示意图；

20 图 2 示出了本公开实施例所提供的一种三维目标检测方法中，基于第一区域图像和第二区域图像，生成目标视差图的具体方法的流程示意图；

图 3 示出了本公开实施例所提供的一种三维目标检测方法中，对视差估计网络进行训练的具体方法的流程示意图；

25 图 4 示出了本公开实施例所提供的一种三维目标检测方法中，针对一组双目图像样本，基于待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据，得到该组双目图像样本对应的目标视差图的具体方法的流程示意图；

图 5 示出了本公开实施例所提供的一种三维目标检测方法中，基于双目图像样本对应的点云数据、以及变换后的双目图像样本，生成双目图像样本对应的目标视差图的具体方法的流程示意图；

30 图 6 示出了本公开实施例所提供的一种行驶控制方法的流程示意图；

图 7 示出了本公开实施例所提供的一种三维目标检测装置的架构示意图；

图 8 示出了本公开实施例所提供的一种行驶控制装置的架构示意图；

图 9 示出了本公开实施例所提供的一种电子设备 900 的结构示意图；

图 10 示出了本公开实施例所提供的另一种电子设备 1000 的结构示意图。

具体实施方式

为使本公开实施例的目的、技术方案和优点更加清楚，下面将结合本公开实施例中的附图，对本公开实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本公开一部分实施例，而不是全部的实施例。通常在此处附图中描述和示出的本公开实施例的组件可以以各种不同的配置来布置和设计。因此，以下对在附图中提供的本公开的实施例的详细描述并非旨在限制要求保护的本公开的范围，而是仅仅表示本公开的选定实施例。基于本公开的实施例，本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例，都属于本公开保护的范围。

目前可以通过激光雷达获取场景内的点云数据，再基于获取的点云数据，确定场景内包括的每个对象的三维位姿数据，即确定每个对象的三维检测框。本公开实施例提供了一种三维目标检测方法，可减少依赖激光雷达获取对象的点云数据，减少获取点云数据的成本。

本公开可以应用于图像或视频的目标检测、目标定位等场景，例如，可以在自动驾驶、辅助驾驶等的交通场景中，通过行驶过程中实时采集到的双目图像，确定出现在道路上的车辆、行人等目标对象的三维位姿数据，从而辅助驾驶员驾驶。

本公开实施例提供的三维目标检测方法可应用于服务器，或者应用于具有中央处理器的终端设备。服务器可以是本地服务器也可以是云端服务器等，终端设备可以是智能手机、平板电脑、智能电视、车载电脑等，本公开对此并不限定。

为便于对本公开实施例进行理解，首先对本公开实施例所公开的一种三维目标检测方法进行详细介绍。

参见图 1 所示，为本公开实施例所提供的三维目标检测方法的流程示意图，该方法包括 S101-S104，其中：

S101，获取双目图像；双目图像包括左目图像和右目图像。

S102，分别对左目图像和右目图像进行目标检测，得到待检测对象在左目图像中的第一区域图像和该待检测对象在右目图像中的第二区域图像。

S103，基于第一区域图像和第二区域图像，生成待检测对象的目标视差图，并基于待检测对象的目标视差图、以及第一区域图像和第二区域图像，计算得到待检测对象的点云数据。

S104，基于点云数据，确定待检测对象的三维位姿数据。

上述方法中，通过从左目图像中得到待检测对象的第一区域图像以及从右目图像中得到待检测对象的第二区域图像，并基于第一区域图像以及第二区域图像，确定待检测对象的目标视差图，进而基于该目标视差图，计算得到待检测对象的点云数据，采用这种基于视差图的方式生成点云数据，可减少依赖激光雷达获取点云数据，可以生成较为稠密的点云数据，进而可以基于视差图生成的点云数据确定待检测对象的三维位姿数据。

以下对 S101-S104 进行说明。

针对 S101:

这里, 可以从存储有双目图像的存储装置中获取该双目图像, 也可以实时的从摄像设备 (比如双目相机) 上获取双目图像。一般的, 双目图像可以包括左目图像和右目图像。

5 针对 S102:

本公开实施例中, 可以通过对象检测神经网络分别对左目图像和右目图像进行目标检测, 得到至少一个待检测对象中每个待检测对象在左目图像中的第一检测框以及在右目图像中的第二检测框, 并根据第一检测框, 从左目图像中截取得到每个待检测对象的第一区域图像, 以及根据第二检测框, 从右目图像中截取得到每个待检测对象的第二区域图像。

10 示例性的, 在第一检测框的尺寸与第二检测框的尺寸不同时, 则可以选择较大的尺寸作为截取区域图像时对应的检测框的尺寸。比如, 若第一检测框的尺寸为第一尺寸, 第二检测框的尺寸为第二尺寸, 第一尺寸大于第二尺寸, 则将第二检测框的尺寸调整为第一尺寸, 并基于第一检测框和尺寸调整后的第二检测框, 得到第一区域图像以及第二区域图像, 其中, 第一区域图像与第二区域图像的尺寸相同。

15 这里, 可以通过多组检测训练样本对对象检测神经网络进行训练, 其中, 每组检测训练样本中包括标注对象的左目图像样本和右目图像样本。示例性的, 将多组检测训练样本输入至对象检测神经网络中, 对对象检测神经网络进行训练, 直至对象检测神经网络满足设置的条件为止。

20 针对 S103:

本公开实施例中, 视差图中像素点的像素值可以为该像素点对应在第一区域图像中的横坐标值与该像素点对应在第二区域图像中的横坐标值之差。比如, 若视差图中的像素点A对应在第一区域图像中的横坐标值为 x_1 , 该像素点A对应在第二区域图像中的横坐标值为 x_2 , 则视差图中该像素点A的像素值为 $x_1 - x_2$ 。其中, x_1 、 x_2 可以为像素点A 在图像坐标系下的横坐标值。

25 这里, 可以通过反投影公式, 基于每个待检测对象的目标视差图、相机内参、第一区域图像和第二区域图像, 计算得到每个待检测对象的点云数据。

一种可选实施方式中, 在生成所述待检测图像的目标视差图之前, 所述方法还包括:

基于第一区域图像与第二区域图像, 确定横坐标调整值;

30 基于横坐标调整值, 调整第一区域图像或第二区域图像中像素点的横坐标值, 使得第一区域图像与第二区域图像中对应边界点的横坐标值重合。

可以基于第一区域图像与第二区域图像, 确定坐标调整值。这里, 可以基于位置点在第一区域图像上的横坐标值以及在第二区域图像上的横坐标值, 确定横坐标调整值, 其中, 该位置点可以为区域图像上的任一位置点, 比如, 该位置点可以为区域图像的顶点、中心点、边界线上的中点等。示例性的, 可以确定第一区域图像的左上角的顶点位

置的第一横坐标值、以及确定第二区域图像的左上角的顶点位置的第二横坐标值，求解第一横坐标值与第二横坐标值的差值，将该差值确定为横坐标调整值。这里，第一区域图像上的像素点的坐标值与第二区域图像上的像素点的坐标值位于同一图像坐标系中。

5 示例性的，还可以通过确定第一区域图像的左边界线的第一横坐标值与第二区域图像的左边界线的第二横坐标值之差，确定横坐标调整值。或者，可以通过确定第一区域图像的右边界线的第一横坐标值与第二区域图像的右边界线的第二横坐标值之差，确定横坐标调整值。其中，同一区域图像中的左边界线和右边界线上任一位置点的横坐标值相同。

进一步的，可以基于横坐标调整值，调整第一区域图像或第二区域图像中像素点的横坐标值。示例性说明，若第一区域图像中左边界线的第一横坐标值为 60，第二区域图像中左边界线的第二横坐标值为 30，则确定横坐标调整值为 30，并可以基于横坐标调整值 30，调整第一区域图像中像素点的横坐标值，即将第一区域图像中多个像素点的横坐标值均减去 30，使得调整后的第一区域图像与第二区域图像中左边界线上边界点的横坐标值重合；或者，可以基于横坐标调整值 30，调整第二区域图像中像素点的横坐标值，
15 即将第二区域图像中多个像素点的横坐标值均加上 30，使得调整后的第二区域图像与第一区域图像中左边界线上边界点的横坐标值重合。

这里，可以基于横坐标值调整后的第一区域图像和第二区域图像，得到区域视差图。比如，可以将对应边界点的横坐标值重合的第一区域图像和第二区域图像输入至视差估计网络中进行视差估计，得到该待检测对象的区域视差图。

20 上述实施方式中，通过确定横坐标调整值，该横坐标调整值可以表征第一区域图像与第二区域图像之间的粗略的视差值，进而可以基于该横坐标调整值粗略的对第一区域图像和第二区域图像之间的视差值进行修正，使得在基于调整后的第一区域图像和所述第二区域图像，生成任意一个待检测对象的目标视差图时，可以减少区域图像搜索范围，以及减少目标视差图确定过程的运算量。

25 一种可选实施例，参见图 2 所示，基于第一区域图像和第二区域图像，生成目标视差图，可以包括：

S201，确定待检测对象在第一区域图像或第二区域图像中的轮廓图像，并基于待检测对象的第一区域图像和第二区域图像，生成待检测对象对应的区域视差图。

S202，基于轮廓图像，从区域视差图中截取目标视差图。

30 示例性的，可以通过对象检测神经网络确定每个待检测对象在第一区域图像或第二区域图像中的轮廓图像。比如，对象检测神经网络对左目图像进行目标检测，得到至少一个待检测对象中每个待检测对象在左目图像中的第一检测框以及该待检测对象在第一检测框内的轮廓图像，即得到待检测对象在第一区域图像中的轮廓图像。

这里，在基于每个待检测对象的第一区域图像和第二区域图像，得到该待检测对象对应的区域视差图之后，可以基于该待检测对象的轮廓图像，从对应的区域视差图中截
35 取该待检测对象的目标视差图。或者，在基于横坐标值重合的第一区域图像和第二区域

图像,得到该待检测对象对应的区域视差图之后,可以基于该待检测对象的轮廓图像,从对应的区域视差图中截取该待检测对象的目标视差图。

上述实施方式下,通过待检测对象的轮廓图像,可以从区域视差图中截取待检测对象的目标视差图,进一步保留与目标相关的有用的视差图,可以使得基于目标视差图确定的待检测对象的点云数据较好的表征待检测对象,进而可以提高检测的精准度。

一种可选实施例中,基于第一区域图像和第二区域图像,生成待检测对象对应的区域视差图,包括:将第一区域图像和第二区域图像输入至视差估计网络中进行视差估计,得到区域视差图。

示例性的,可以将待检测对象的第一区域图像和第二区域图像输入至视差估计网络中,视差估计网络基于构建的每个像素点的像素值的最小损失量,确定该像素点在第一区域图像中的位置以及在第二区域图像中对应的位置,进而基于每个像素点在第一区域图像中的位置的横坐标值以及对应的像素点在第二区域图像中对应的位置的横坐标值,得到待检测对象的区域视差图。或者,可以将待检测对象的横坐标值重合后的第一区域图像和第二区域图像输入至视差估计网络中,得到待检测对象的区域视差图。

上述实施方式中,通过视差估计网络,确定待检测对象的区域视差图,在保障得到的区域视差图的准确度的同时,可以较快速的确定待检测对象的区域视差图。

一种可选实施方式中,参见图3所示,可以通过下述步骤对视差估计网络进行训练:

S301,获取同一类型下不同待检测对象的形状数据、多组标注好待检测对象的三维位姿数据的双目图像样本,以及每组双目图像样本对应的点云数据;每组双目图像样本包括左目图像样本和右目图像样本。

S302,基于同一类型下不同待检测对象的形状数据,确定该类型下的待检测对象的平均形状数据。

S303,针对每组双目图像样本,基于待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据,得到该组双目图像样本对应的目标视差图。

S304,将双目图像样本以及每组双目图像样本对应的目标视差图作为训练样本,对视差估计网络进行训练,得到训练后的视差估计网络。

对 S301 进行说明,以待检测对象为车辆为例进行说明,这里,车辆类型可以包括微型车辆类型、轻型车辆类型、中型车辆类型、大型车辆类型等,或者,车辆类型还可以包括轿车类型、跑车类型、越野车类型等。具体的,待检测对象的类型的划分可以根据实际情况进行确定,此处仅为示例性说明。这里,同一类型下不同待检测对象可以为轻型车辆类型下,不同车辆(不同车辆可以为不同品牌和/或不同型号的车辆的)的形状数据。

这里,可以用截断符号距离函数表征待检测对象的形状数据。示例性的,可以将待检测对象划分为 n 个小正方体,每个小正方体对应一个数值,该数值的大小可以为该正

方体距离待检测对象的各个表面的距离值中的最小值,可以得到待检测对象对应的 n 个数值,该 n 个数值构成了待检测对象的截断符号距离函数值, n 为正整数。其中,同一类型下不同待检测对象中每个待检测对象的形状数据中均包括 n 个数值, n 的值可以根据需要进行设置。

- 5 本公开实施例中,可以获取同一类型下不同待检测对象的形状数据、多组双目图像样本,以及每组双目图像样本对应的点云数据;其中,每组双目图像样本包括左目图像样本和右目图像样本,每组双目图像样本中的左目图像样本和右目图像样本中均标注有每个待检测对象的三维位姿数据。这里,每组双目图像样本对象的点云数据为稀疏的点云数据,比如,每组双目图像样本对应的点云数据可以为通过卷积神经网络模型估计得到的数据,其中,该卷积神经网络可以为基于训练数据集训练后的神经网络。或者,也可以为通过激光雷达获取到的点云数据。

对 S302 进行说明,可以将同一类型下不同待检测对象的形状数据求平均,得到该类型下的待检测对象的平均形状数据。这里,该类型下的待检测对象的平均形状数据对应一平均形状。

- 15 可以将每个待检测对象对应的截断距离符号函数值求平均,得到该类型下的待检测对象的平均形状数据。示例性的,若同一类型下的待检测对象包括待检测对象A、待检测对象B、以及待检测对象C,待检测对象A的形状数据为 $\{a_1, a_2, \dots, a_n\}$ 、待检测对象B的形状数据为 $\{b_1; b_2; \dots; b_n\}$ 、以及待检测对象C的形状数据为 $\{c_1; c_2; \dots; c_n\}$,则确定的该类型下的待检测对象的平均形状数据可以为 $\{m_1; m_2; \dots; m_n\}$ 。其中, $m_1;$
- 20 $m_2; \dots; m_n$ 的计算过程可以为:

$$m_1 = \frac{(a_1+b_1+c_1)}{3}; m_2 = \frac{(a_2+b_2+c_2)}{3}; m_n = \frac{(a_n+b_n+c_n)}{3}。$$

- 对 S303 进行说明,针对每组双目图像样本,可以确定该双目图像样本中包括的待检测对象的类型,进而可以确定该类型下的待检测对象对应的平均形状数据;再可以基于该类型下的待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据,得到该组双目图像样本对应的目标视差图。
- 25

- 上述实施方式中,可以基于平均形状数据、每组双目图像样本中包括的待检测对象的三维位姿数据、以及每组双目图像样本对应的点云数据,确定每组双目图像样本对应的目标视差图,基于上述实施方式得到的目标视差图可以更准确的表征待检测对象的视差信息;进一步的,通过双目图像样本以及每组双目图像样本对应的目标视差图构成的训练样本对视差估计网络训练时,得到的视差估计网络的准确度更高。
- 30

一种可选实施例中,参见图 4 所示,针对任意一组双目图像样本,基于待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据,得到该组双目图像样本对应的目标视差图,可以包括:

- 35 S401,针对一组双目图像样本,生成变换后的双目图像样本,其中,所述变换后的双目图像样本与所述平均形状数据以及所述三维位姿数据相匹配。

本公开实施例中，针对每组双目图像样本，确定该组双目图像样本中待检测对象对应的平均形状数据，再根据该双目图像样本中待检测对象对应的三维位姿数据，调整该平均形状数据对应的平均形状的尺寸以及方向，并基于调整后的平均形状以及该组双目图像样本，得到变换后的双目图像样本。其中，变换后的双目图像样本中待检测对象的形状为平均形状，且该待检测对象的三维位姿数据与变换前的三维位姿数据相同。

S402，基于双目图像样本对应的点云数据、以及变换后的双目图像样本，生成双目图像样本对应的目标视差图。

上述实施方式中，在不改变双目图像样本中待检测对象的三维位姿数据的前提下，生成变换后的双目图像样本，使得变换后双目图像样本中的形状数据与平均形状数据匹配，由于平均形状数据可以表征待检测对象的形状特征，且物体形状表面具有连续性，从而基于双目图像样本对应的点云数据、以及变换后的双目图像样本，得到的双目图像样本对应的目标视差图更准确和稠密。

示例性的，如图5所示，基于双目图像样本对应的点云数据、以及变换后的双目图像样本，生成双目图像样本对应的目标视差图，可以包括：

S501，基于双目图像样本对应的点云数据，对变换后的双目图像样本的形状数据进行调整，得到调整后的双目图像样本；其中，调整后的双目图像样本的形状数据与所述点云数据所对应的形状相匹配。

这里，可以基于双目图像样本对应的点云数据，对变换后的双目图像样本的形状数据进行调整，得到调整后的双目图像样本。具体的，待检测对象的表面位置处对应的截断符号距离函数的数值为零，即点云数据对应的点云点的截断符号距离函数的目标数值可以为零，因此，可以根据点云数据对应的点云点的截断符号距离函数的真实数值以及目标数值，对变换后的双目图像样本的形状数据进行调整，得到调整后的双目图像样本，使得调整后的双目图像样本的形状数据与点云数据所对应的形状相匹配。

S502，基于调整后的双目图像样本，生成双目图像样本对应的深度图。

本公开实施例中，可以基于调整后的双目图像样本以及渲染器，生成双目图像样本对应的深度图。

S503，基于深度图，生成该组双目图像样本对应的目标视差图。

示例性的，可以通过深度图，生成该组双目图像样本对应的目标视差图。比如，可以将深度图中每一像素点通过几何关系转换的方式，生成该组双目图像样本对应的目标视差图。

这里，可以基于深度图，生成该组双目图像样本对应的区域视差图，再基于区域视差图以及待检测对象的轮廓图像，生成该组双目图像样本对应的目标视差图。

上述实施方式中，基于双目图像样本对应的点云数据，对变换后的双目图像样本的形状数据进行调整，使得调整后的形状数据对应的形状与双目图像样本变换前包含的形状数据对应的形状相符，即使得调整后的形状数据与双目图像样本中待检测对象的真实

形状相符，进而基于调整后的双目图像样本，生成的双目图像样本对应的深度图的准确度和稠密度较高，进而可以提高生成的目标视差图的准确度。

承接 S303 的说明，继续对 S304 进行说明，在得到每组双目图像样本对应的目标视差图之后，将双目图像样本以及每组双目图像样本对应的目标视差图作为训练样本，对视差估计网络进行训练，直至训练完成的视差估计网络满足条件为止，得到训练后的视差估计网络。其中，条件可以为视差估计网络的准确度大于设置的准确度阈值，或者，视差估计网络的损失量小于设置的损失量阈值等。

针对 S104：

示例性的，可以将点云数据输入至三维检测神经网络中，确定每个待检测对象的三维位姿数据。其中，三维检测神经网络的结构可以根据实际需要进行设置，此处不进行具体限定。

示例性的，三维位姿数据包括以下信息中的至少一种：待检测对象的三维检测框的长度、高度、宽度、三维检测框的中心点的坐标信息、以及三维检测框的基准面与设置的基准线的夹角。其中，基准线和三维检测框的基准面可以根据实际情况进行设置；比如，基准面可以为三维检测框的前表面、后表面、左侧表面、右侧表面等；基准线可以为图像坐标系的横坐标轴。

基于相同的构思，本公开实施例还提供了一种行驶控制方法，参见图 7 所示，为本公开实施例所提供的行驶控制方法的流程示意图，该方法包括 S601-S603，其中：

S601，获取行驶装置在行驶过程中采集的双目道路图像；

S602，利用本公开实施例提供的三维目标检测方法对双目道路图像进行三维目标检测，得到双目道路图像中包括的目标对象的三维位姿数据；

S603，基于双目道路图像中包括的目标对象的三维位姿数据，控制行驶装置。

示例性的，行驶装置可以为自动驾驶车辆、装有高级驾驶辅助系统（Advanced Driving Assistance System，ADAS）的车辆、或者机器人等。双目道路图像可以为行驶装置在行驶过程中实时采集到的双目图像。目标对象可以为道路中可以出现的任一物体和/或、任一对象。比如，目标对象可以为出现在道路上的动物、行人等，也可以为道路上的其他车辆等。

其中，在控制行驶装置时，可以控制行驶装置加速、减速、转向、制动等，或者可以播放语音提示信息，以提示驾驶员控制行驶装置加速、减速、转向、制动等。

本领域技术人员可以理解，在具体实施方式的上述方法中，各步骤的撰写顺序并不意味着严格的执行顺序而对实施过程构成任何限定，各步骤的具体执行顺序应当以其功能和可能的内在逻辑确定。在不违背逻辑的情况下，本公开不同实施方式之间可以相互结合，不同实施方式描述有所侧重，不同实施方式所侧重描述的部分可参见相应实现方式的记载。

基于相同的构思，本公开实施例还提供了一种三维目标检测装置，参见图 7 所示，

为本公开实施例提供的三维目标检测装置的架构示意图，包括双目图像获取模块 701、检测模块 702、点云数据确定模块 703、三维位姿数据确定模块 704、横坐标调整值确定模块 705、调整模块 706、网络训练模块 707，具体的：

5 双目图像获取模块 701，用于获取双目图像；所述双目图像包括左目图像和右目图像；

检测模块 702，用于分别对所述左目图像和右目图像进行目标检测，得到待检测对象在所述左目图像中的第一区域图像和该待检测对象在所述右目图像中的第二区域图像；

10 点云数据确定模块 703，用于基于所述第一区域图像和所述第二区域图像，生成所述待检测对象的目标视差图，并基于所述待检测对象的目标视差图、以及所述第一区域图像和所述第二区域图像，计算得到所述待检测对象的点云数据；

三维位姿数据确定模块 704，用于基于所述点云数据，确定所述待检测对象的三维位姿数据。

一种可能的实施方式中，所述装置还包括：

15 横坐标调整值确定模块 705，用于基于所述第一区域图像与所述第二区域图像，确定横坐标调整值；

调整模块 706，用于基于所述横坐标调整值，调整所述第一区域图像或所述第二区域图像中像素点的横坐标值，使得所述第一区域图像与所述第二区域图像中对应边界点的横坐标值重合。

20 一种可能的实施方式中，所述点云数据确定模块 703，在基于所述第一区域图像和所述第二区域图像，生成所述目标视差图的情况下，用于：

确定所述待检测对象在所述第一区域图像或所述第二区域图像中的轮廓图像，并基于所述第一区域图像和所述第二区域图像，生成所述待检测对象对应的区域视差图；

基于所述轮廓图像，从所述区域视差图中截取所述目标视差图。

25 一种可能的实施方式中，所述点云数据确定模块 703，在基于所述第一区域图像和所述第二区域图像，生成所述待检测对象对应的区域视差图的情况下，用于：将所述第一区域图像和第二区域图像输入至视差估计网络中进行视差估计，得到所述区域视差图。

一种可能的实施方式中，所述装置还包括：网络训练模块 707；

所述网络训练模块 707，用于通过下述步骤对视差估计网络进行训练：

30 获取同一类型下不同待检测对象的形状数据、多组标注好待检测对象的三维位姿数据的双目图像样本，以及每组双目图像样本对应的点云数据；每组双目图像样本包括左目图像样本和右目图像样本；

基于所述同一类型下不同待检测对象的形状数据，确定该类型下的待检测对象的平均形状数据；

针对每组双目图像样本，基于所述待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据，得到该组双目图像样本对应的目标视差图；

5 将所述双目图像样本以及每组双目图像样本对应的目标视差图作为训练样本，对所述视差估计网络进行训练，得到训练后的视差估计网络。

一种可能的实施方式中，所述网络训练模块 707，在针对每组双目图像样本，基于所述待检测对象的平均形状数据、该组双目图像样本中包括的待检测对象的三维位姿数据、以及该组双目图像样本对应的点云数据，得到该组双目图像样本对应的目标视差图的情况下，用于：

10 针对每组双目图像样本，生成变换后的双目图像样本，其中，所述变换后的双目图像样本与所述平均形状数据以及所述三维位姿数据相匹配；

基于所述双目图像样本对应的点云数据、以及所述变换后的双目图像样本，生成所述双目图像样本对应的目标视差图。

15 一种可能的实施方式中，所述网络训练模块 707，在基于所述双目图像样本对应的点云数据、以及所述变换后的双目图像样本，生成所述双目图像样本对应的目标视差图的情况下，用于：

基于所述双目图像样本对应的点云数据，对所述变换后的双目图像样本的形状数据进行调整，得到调整后的双目图像样本；其中，调整后的双目图像样本的形状数据与所述点云数据所对应的形状相匹配；

20 基于所述调整后的双目图像样本，生成所述双目图像样本对应的深度图；

基于所述深度图，生成该组双目图像样本对应的目标视差图。

一种可能的实施方式中，所述三维位姿数据包括以下信息中的至少一种：

所述待检测对象的三维检测框的长度、高度、宽度、所述三维检测框的中心点的坐标信息、以及所述三维检测框的基准面与设置的基准线的夹角。

25 本公开实施例还提供了一种行驶控制装置，参见图 8 所示，为本公开实施例提供了一种行驶控制装置的架构示意图，包括道路图像获取模块 801、三维目标检测模块 802、控制模块 803，具体的：

道路图像获取模块 801，用于获取行驶装置在行驶过程中采集的双目道路图像；

30 三维目标检测模块 802，用于利用上述实施例所述的三维目标检测方法对所述双目道路图像进行三维目标检测，得到所述双目道路图像中包括的目标对象的三维位姿数据；

控制模块 803，用于基于所述双目道路图像中包括的目标对象的三维位姿数据，控制所述行驶装置。

在一些实施例中，本公开实施例提供的装置具有的功能或包含的模板可以用于执行上文方法实施例描述的方法，其具体实现可以参照上文方法实施例的描述，为了简洁，

这里不再赘述。

基于同一技术构思，本公开实施例还提供了一种电子设备。参照图 9 所示，为本公开实施例提供的电子设备 900 的结构示意图，包括处理器 901、存储器 902、和总线 903。其中，存储器 902 用于存储执行指令，包括内存 9021 和外部存储器 9022；这里的内存 9021 也称内存储器，用于暂时存放处理器 901 中的运算数据，以及与硬盘等外部存储器 9022 交换的数据，处理器 901 通过内存 9021 与外部存储器 9022 进行数据交换，当电子设备 900 运行时，处理器 901 与存储器 902 之间通过总线 903 通信，使得处理器 901 在执行以下指令：

获取双目图像；所述双目图像包括左目图像和右目图像；

- 10 分别对所述左目图像和右目图像进行目标检测，得到待检测对象在所述左目图像中的第一区域图像和该待检测对象在所述右目图像中的第二区域图像；

基于所述第一区域图像和所述第二区域图像，生成所述待检测对象的目标视差图，并基于所述待检测对象的目标视差图、以及所述第一区域图像和所述第二区域图像，计算得到所述待检测对象的点云数据；

- 15 基于所述点云数据，确定所述待检测对象的三维位姿数据。

- 基于同一技术构思，本公开实施例还提供了另一种电子设备。参照图 10 所示，为本公开实施例提供的电子设备 1000 的结构示意图，包括处理器 1001、存储器 1002、和总线 1003。其中，存储器 1002 用于存储执行指令，包括内存 10021 和外部存储器 10022；这里的内存 10021 也称内存储器，用于暂时存放处理器 1001 中的运算数据，以及与硬盘等外部存储器 10022 交换的数据，处理器 1001 通过内存 10021 与外部存储器 10022 进行数据交换，当电子设备 1000 运行时，处理器 1001 与存储器 1002 之间通过总线 1003 通信，使得处理器 1001 在执行以下指令：

获取行驶装置在行驶过程中采集的双目道路图像；

- 25 利用上述实施例提供的三维目标检测方法对所述双目道路图像进行三维目标检测，得到所述双目道路图像中包括的目标对象的三维位姿数据；

基于所述双目道路图像中包括的目标对象的三维位姿数据，控制所述行驶装置。

此外，本公开实施例还提供一种计算机可读存储介质，该计算机可读存储介质上存储有计算机程序，该计算机程序被处理器运行时执行上述方法实施例中所述的三维目标检测方法的步骤。

- 30 本公开实施例所提供的三维目标检测方法的计算机程序产品，包括存储了程序代码的计算机可读存储介质，所述程序代码包括的指令可用于执行上述方法实施例中所述的三维目标检测方法的步骤，具体可参见上述方法实施例，在此不再赘述。

- 此外，本公开实施例还提供另一种计算机可读存储介质，该计算机可读存储介质上存储有计算机程序，该计算机程序被处理器运行时执行上述方法实施例中所述的行驶控制方法的步骤。

本公开实施例所提供的行驶控制方法的计算机程序产品，包括存储了程序代码的计算机可读存储介质，所述程序代码包括的指令可用于执行上述方法实施例中所述的行驶控制方法的步骤，具体可参见上述方法实施例，在此不再赘述。

5 本公开实施例还提供了一种计算机程序，所述计算机程序包括计算机可读代码，当所述计算机可读代码在电子设备中运行时，所述电子设备中的处理器执行所述的三维目标检测方法，或实现所述的行驶控制方法。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统 and 装置的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。在本公开所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，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，又例如，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信连接可以是通过一些通信接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

15 所述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。

20 另外，在本公开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。

所述功能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，可以存储在一个处理器可执行的非易失的计算机可读取存储介质中。基于这样的理解，本公开的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案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一个存储介质中，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，服务器，或者网络设备等）执行本公开各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（Read-Only Memory, ROM）、随机存取存储器（Random Access Memory, RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

30 以上仅为本公开的具体实施方式，但本公开的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本公开揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本公开的保护范围之内。因此，本公开的保护范围应以权利要求的保护范围为准。