

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022 年 1 月 27 日 (27.01.2022)



(10) 国际公布号
WO 2022/016368 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2020/103263

(22) 国际申请日: 2020 年 7 月 21 日 (21.07.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 深圳市大疆创新科技有限公司 (SZ DJI TECHNOLOGY CO., LTD.) [CN/CN]; 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。

(72) 发明人: 陈创荣 (CHEN, Chuangrong); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。 徐斌 (XU, Bin); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。 陈晓

智 (CHEN, Xiaozhi); 中国广东省深圳市南山区高新区南区粤兴一道9号香港科大深圳产学研大楼6楼, Guangdong 518057 (CN)。

(74) 代理人: 北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM); 中国北京市海淀区西直门北大街32号枫蓝国际A座8F-6, Beijing 100082 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: 3D FRAME LABELING METHOD AND DEVICE AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 3D框标注方法、设备及计算机可读存储介质

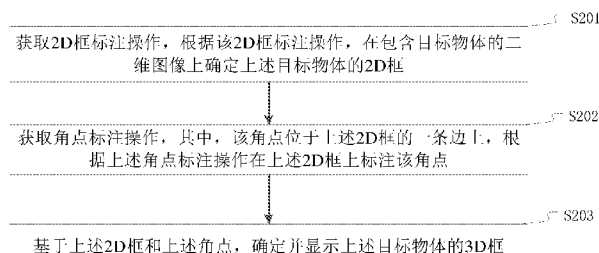


图 2

- S201 Acquire a 2D frame labeling operation, and according to the 2D frame labeling operation, determine in a two-dimensional image containing a target object a 2D frame of the target object
- S202 Acquire a corner point labeling operation, wherein the corner point is located on one side of the 2D frame, and label the corner point on the 2D frame according to the corner point labeling operation
- S203 On the basis of the 2D frame and the corner point, determine and display a 3D frame of the target object

(57) Abstract: A 3D frame labeling method and device and a computer-readable storage medium. The method directly labels a 3D frame at an image level without relying on an extra depth sensor, thereby reducing cost. The method can acquire a 3D frame of a target object only by labeling, on a 2D image containing the target object, a 2D frame of the target object and a corner point on the 2D frame, thus being simple in processing and reducing workload. In addition, the difference between the 3D frame obtained by means of the embodiment of the present application and the actual size of an object is small, solving the problem that the difference between a 3D frame obtained on the basis of point cloud data labeling and the actual size of the object is large, thereby guaranteeing the subsequent accurate training of a neural network and improving the recognition effect of the neural network.

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种3D框标注方法、设备及计算机可读存储介质, 该方法在图像层面直接对3D框进行标注, 无需依赖额外的深度传感器, 降低成本, 而且该方法仅通过在包含目标物体的二维图像上标注目标物体的2D框和该2D框上的一角点, 就能获得目标物体的3D框, 处理过程简单, 降低工作量, 另外, 本申请实施例获取的3D框与物体实际尺寸相差较小, 解决现有基于点云数据标注获得的3D框与物体的实际尺寸相差较大的问题, 从而, 保证后续对神经网络的准确训练, 提高神经网络的识别效果。

3D 框标注方法、设备及计算机可读存储介质

技术领域

5 本申请实施例涉及图像处理技术，尤其涉及一种 3D 框标注方法、设备及计算机可读存储介质。

背景技术

随着神经网络至深度学习等人工智能（AI）领域技术的快速发展，人们已经能够使用这些 AI 技术来实现对周围环境的感知功能。例如，在自动驾驶中，可以利用神经网络对搭载于车辆上的摄像头所采集的图像进行识别，从而得到周围目标物体（如周围的车辆、行人、树木等）的 2D 或 3D 信息。然而，想要获得高准确率的识别结果，需要首先对所使用的神经网络进行训练。例如，如需要使用神经网络对图像进行识别得到目标物体及其 3D 信息，则需要首先利用训练图像和图像中已知目标物体和其 3D 信息对该神经网络进行训练。

获得图像中已知的目标物体和其 3D 信息（如伪 3D 框，即二维图像平面上的 3D 框投影，后续简述为 3D 框），通常需要人工进行标注。现有的标注方法一个是依赖于外部激光雷达传感器或深度相机等主动式深度传感器的标注方法。该方法依赖外部深度传感器获取深度信息，生成点云数据，并直接在 3D 空间标注物体的实际 3D 框，再根据传感器间坐标转换关系将其 3D 框投影到图像上，从而得到 3D 框。

然而，这类标注方法过程复杂，成本较高，而且基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大，影响后续对神经网络的训练，从而
25 影响神经网络的识别效果。因此，有必要提供一种更为高效准备的标注方法。

发明内容

本申请实施例提供一种 3D 框标注方法、设备及计算机可读存储介质，
30 以克服上述至少一个问题。

第一方面，本申请实施例提供一种 3D 框标注方法，包括：

获取 2D 框标注操作，根据所述 2D 框标注操作，在包含目标物体的二维图像上确定所述目标物体的 2D 框；

5 获取角点标注操作，其中，所述角点位于所述 2D 框的一条边上，根据所述角点标注操作在所述 2D 框上标注所述角点；

基于所述 2D 框和所述角点，确定并显示所述目标物体的 3D 框。

第二方面，本申请实施例提供另一种 3D 框标注方法，包括：

在包含目标物体的二维图像上确定所述目标物体的 2D 框；

获取标注的角点，其中，所述角点位于所述 2D 框的一条边上；

10 基于所述 2D 框和所述角点，确定所述目标物体的 3D 框。

第三方面，本申请实施例提供一种 3D 框标注设备，包括存储器、处理器、交互单元，以及存储在所述存储器中并可在所述处理器上运行的计算机指令，所述处理器执行所述计算机指令时实现如下步骤：

15 通过所述交互单元获取 2D 框标注操作，根据所述 2D 框标注操作，在包含目标物体的二维图像上确定所述目标物体的 2D 框；

通过所述交互单元获取角点标注操作，其中，所述角点位于所述 2D 框的一条边上，根据所述角点标注操作在所述 2D 框上标注所述角点；

基于所述 2D 框和所述角点，确定所述目标物体的 3D 框，并通过所述交互单元显示。

20 第四方面，本申请实施例提供另一种 3D 框标注设备，包括存储器、处理器，以及存储在所述存储器中并可在所述处理器上运行的计算机指令，所述处理器执行所述计算机指令时实现如下步骤：

在包含目标物体的二维图像上确定所述目标物体的 2D 框；

获取标注的角点，其中，所述角点位于所述 2D 框的一条边上；

25 基于所述 2D 框和所述角点，确定所述目标物体的 3D 框。

第五方面，本申请实施例提供一种神经网络训练方法，包括：

利用如上第一方面以及第一方面各种可能的设计所述的 3D 框标注方法确定的目标物体的 3D 框，以及包含所述目标物体的二维图像，进行神经网络的训练。

30 第六方面，本申请实施例提供另一种神经网络训练方法，包括：

利用如上第二方面以及第二方面各种可能的设计所述的 3D 框标注方法确定的目标物体的 3D 框，以及包含所述目标物体的二维图像，进行神经网络的训练。

5 第七方面，本申请实施例提供一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机指令，当处理器执行所述计算机指令时，实现如上第一方面以及第一方面各种可能的设计所述的 3D 框标注方法。

第八方面，本申请实施例提供另一种计算机可读存储介质，所述计算机可读存储介质中存储有计算机指令，当处理器执行所述计算机指令时，实现如上第二方面以及第二方面各种可能的设计所述的 3D 框标注方法。

10 本申请实施例提供的 3D 框标注方法、设备及计算机可读存储介质，该方法在图像层面直接对 3D 框进行标注，无需依赖额外的深度传感器，降低成本，而且该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，就能获得目标物体的 3D 框，处理过程简单，降低工作量，另外，本申请实施例获取的 3D 框与物体实际尺寸相差
15 较小，解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题，从而，保证后续对神经网络的准确训练，提高神经网络的识别效果。

附图说明

20 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本申请的实施例，并与说明书一起用于解释本申请的原理。

图 1 为本申请实施例提供的 3D 框标注系统架构示意图；

图 2 为本申请实施例提供的一种 3D 框标注方法的流程示意图；

图 3 为本申请实施例提供的一种 2D 框示意图；

25 图 4 为本申请实施例提供的一种 2D 框上的角点示意图；

图 5 为本申请实施例提供的一种 3D 框示意图；

图 6 为本申请实施例提供的另一种 3D 框标注方法的流程示意图；

图 7 为本申请实施例提供的再一种 3D 框标注方法的流程示意图；

图 8 为本申请实施例提供的又一种 3D 框标注方法的流程示意图；

30 图 9 为本申请实施例提供的又一种 3D 框标注方法的流程示意图；

图 10 为本申请实施例提供的又一种 3D 框标注方法的流程示意图；
图 11 为本申请实施例提供的一种 2D 框和 3D 框的对应关系示意图；
图 12 为本申请实施例提供的又一种 3D 框标注方法的流程示意图；
图 13 为本申请实施例提供的一种 3D 框标注装置的结构示意图；
5 图 14 为本申请实施例提供的另一种 3D 框标注装置的结构示意图；
图 15 为本申请实施例提供的一种 3D 框标注设备的基本硬件架构；
图 16 为本申请实施例提供的另一种 3D 框标注设备的基本硬件架构。

通过上述附图，已示出本申请明确的实施例，后文中将有更详细的描述。这些附图和文字描述并不是为了通过任何方式限制本申请构思的范围，
10 而是通过参考特定实施例为本领域技术人员说明本申请的概念。

具体实施方式

这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相
15 似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”及“第四”等（如果存在）是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，
20 以便这里描述的本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，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、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地
25 列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

本申请实施例提供的 3D 框标注方法可应用在神经网络训练的前期数据标注中，其中，该神经网络可以用于获得目标物体（如车辆、房屋等）的 2D 或 3D 信息等，本申请实施例对此不做特别限制。需要说明的是，本申请
30 实施例中标注的 3D 框指的是在二维图像上的物体的三维框的投影，而非实际三维空间中的三维框。

可选地，本申请实施例提供的 3D 框标注方法可以应用于如图 1 所示的应用场景中。图 1 只是以示例的方式描述了本申请实施例提供的 3D 框标注方法的一种可能的应用场景，本申请实施例提供的 3D 框标注方法的应用场景不限于图 1 所示的应用场景。

图 1 为 3D 框标注系统架构示意图。在图 1 中，以获得周围的车辆的 3D 信息为例。上述架构包括处理装置 11 和多个摄像头，这里，该多个摄像头以第一摄像头 12、第二摄像头 13 和第三摄像头 14 为例。

可以理解的是，本申请实施例示意的结构并不构成对 3D 框标注架构的具体限定。在本申请另一些可行的实施方式中，上述架构可以包括比图示更多或更少的部件，或者组合某些部件，或者拆分某些部件，或者不同的部件布置，具体可根据实际应用场景确定，在此不做限制。图 1 所示的部件可以以硬件，软件，或软件与硬件的组合实现。

在具体实现过程中，本申请实施例中第一摄像头 12、第二摄像头 13 和第三摄像头 14 可以分别采集周围的车辆的图像。在上述应用场景中，第一摄像头 12、第二摄像头 13 和第三摄像头 14 在采集图像后，可以把采集的图像发送至处理装置 11。处理装置 11 将接收的上述图像作为样本数据，该样本数据经过标注后，可用于训练神经网络。处理装置 11 在图像层面获取到用户的基本操作后，可以直接生成标注完成的 3D 框，从而，利用上述图像中已知车辆和其 3D 信息对神经网络进行训练。

本申请实施例所涉及的 3D 框标注是指在图像层面直接对目标物体的 3D 框进行标注，指的是目标物体的三维空间中的实际 3D 框在二维图像上投影的 3D 框，即通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，获得目标物体的 3D 框，无需依赖额外的深度传感器，降低成本，处理过程简单，降低工作量。

另外，本申请实施例描述的系统架构以及业务场景是为了更加清楚的说明本申请实施例的技术方案，并不构成对于本申请实施例提供的技术方案的限定，本领域普通技术人员可知，随着系统架构的演变和新业务场景的出现，本申请实施例提供的技术方案对于类似的技术问题，同样适用。

下面结合附图详细介绍本申请实施例提供的 3D 框标注方法。该方法的执行主体可以为图 1 中的处理装置 11。处理装置 11 的工作流程主要包括 2D

框阶段和 3D 框阶段。在 2D 框阶段，处理装置 11 获取在包含目标物体的二维图像上标注的目标物体的 2D 框和该 2D 框上的一角点。在 3D 框阶段，处理装置 11 根据上述 2D 框和 2D 框上的一角点，生成目标物体的 3D 框，无需依赖额外的深度传感器，降低成本，处理过程简单，降低工作量。

5 下面以几个实施例为例对本申请的技术方案进行描述，对于相同或相似的概念或过程可能在某些实施例不再赘述。

图 2 为本申请实施例提供的一种 3D 框标注方法的流程示意图，本实施例的执行主体可以为图 1 中的处理装置 11，具体执行主体可以根据实际应用场景确定。如图 2 所示，在图 1 所示应用场景的基础上，本申请实施例提供的 3D 框标注方法包括如下步骤：

S201：获取 2D 框标注操作，根据该 2D 框标注操作，在包含目标物体的二维图像上确定上述目标物体的 2D 框。

这里，上述目标物体可以根据实际情况确定，例如车辆、房屋等，本申请实施例对此不做特别限制。

15 本申请实施例中，以目标物体为车辆为例，处理装置 11 在获取 2D 框标注操作后，在包含上述车辆的二维图像上确定上述车辆的 2D 框。示例性的，2D 框标注操作可由标注人员操作完成，如图 3 所示，上述车辆的 2D 框在包含上述车辆的二维图像上完全框住上述车辆，具体上述 2D 框大小可以根据实际需要确定，例如上述 2D 框的大小与上述二维图像上的上述车辆的大小越近似越好，本申请实施例对此不做特别限制。

另外，处理装置 11 在确定上述目标物体的 2D 框后，还可以获取 2D 框调整操作，根据该操作，调整上述 2D 框，例如调整上述 2D 框的大小、位置等。

25 S202：获取角点标注操作，其中，该角点位于上述 2D 框的一条边上，根据上述角点标注操作在上述 2D 框上标注该角点。

本申请实施例中，角点指的是被标注物体实际三维框的某个角投影在二维图像上的点。例如，当物体为车辆时，通常其实际三维框为长方体，则标注的角点为该长方体的六个角在二维图像上的投影的点。上述角点具体在上述 2D 框的哪条边上，可以根据实际情况确定，本申请实施例对此

30 不做特别限制。

示例性的，还以目标物体为车辆为例，如图 4 所示，上述角点位于上述 2D 框的底边上。

S203: 基于上述 2D 框和上述角点，确定并显示上述目标物体的 3D 框。

这里，还以目标物体为车辆为例，如图 5 所示，在该车辆的上述 2D 框和上述角点基础上，获得该车辆的 3D 框。

本申请实施例，在图像层面直接对 3D 框进行标注，无需依赖额外的深度传感器，降低成本，而且该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，就能获得目标物体的 3D 框，处理过程简单，降低工作量，另外，本申请实施例获取的 3D 框与物体实际尺寸相差较小，解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题，从而，保证后续对神经网络的准确训练，提高神经网络的识别效果。

另外，本申请实施例在确定并显示上述目标物体的 3D 框之前，还获取上述角点的角点编号。图 6 为本申请实施例提出的另一种 3D 框标注方法的流程示意图。如图 6 所示，该方法包括：

S601: 获取 2D 框标注操作，根据该 2D 框标注操作，在包含目标物体的二维图像上确定上述目标物体的 2D 框。

S602: 获取角点标注操作，其中，该角点位于上述 2D 框的一条边上，根据上述角点标注操作在上述 2D 框上标注该角点。

其中，步骤 S601- S602 与上述步骤 S201- S202 的实现方式相同，此处不再赘述。

S603: 获取上述角点的角点编号，该角点编号用于指示上述角点相对于上述目标物体的位置。

这里，角点编号的顺序可以不限，只要角点与其角点编号之间有映射关系即可。

其中，上述映射关系可以理解为角点相对于物体的位置与其角点编号之间有对应关系，例如位于物体的 2D 框的底边上四个角点，从后面的右边位置开始，顺时针，确定角点编号可以为 p0、p1、p2 和 p3。如果某一角点位于物体的 2D 框的底边的后面左边位置，则其角点编号为 p1。

上述处理装置 11 可以预存上述映射关系，从而，基于该映射关系获取

上述角点的角点编号，该角点编号用于指示上述角点相对于目标物体的位置。

另外，上述角点编号可以是用户输入的，也可以是预先配置的，本申请实施例对此不做特别限制。

- 5 S604：基于上述角点编号、2D 框和角点，确定并显示上述目标物体的 3D 框。

本申请实施例，在确定并显示上述目标物体的 3D 框之前，还获取上述角点的角点编号，进而，基于上述角点编号、2D 框和角点，准确确定并显示上述目标物体的 3D 框，满足应用需要。示例性的，当所述目标物体为车辆时，通常由于车辆位于前方且朝向前方行驶，因此可以默认设定为车辆朝向为正前方；在确定了车辆的 3D 框的一个角点编号后，由于 3D 框在图像上的各个角点都落在已标注的 2D 框的各个边上，因此，可以根据 2D 视图下三维物体的消失点及已确定的角点编号及映射关系，直接生成车辆的 3D 框。本申请实施例在图像层面直接对 3D 框进行标注，无需
10 依赖额外的深度传感器，降低成本，该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，就能获得目标物体的 3D 框，处理过程简单，降低工作量，并且相对于人工标注目标物体 3D 框，能更加满足实际三维框在二维图像上投影的几何关系约束，避免人工标注造成的误差或错误，从而提高标注数据精确度及可用性。本申请实施例获取的 3D 框与物体实际尺寸相差较小，解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题，从而，保证后续对神经网络的准确训练，提高神经网络的识别效果。
15 20

另外，在另一些实施例中，本申请在确定并显示上述目标物体的 3D 框之前，还获取上述目标物体在 3D 空间的朝向角。图 7 为本申请实施例提出的再一种 3D 框标注方法的流程示意图。如图 7 所示，该方法包括：
25

S701：获取 2D 框标注操作，根据该 2D 框标注操作，在包含目标物体的二维图像上确定上述目标物体的 2D 框。

S702：获取角点标注操作，其中，该角点位于上述 2D 框的一条边上，根据上述角点标注操作在上述 2D 框上标注该角点。

30 其中，步骤 S701- S702 与上述步骤 S201- S202 的实现方式相同，此处不

再赘述。

S703: 获取上述目标物体在 3D 空间的朝向角, 该朝向角用于指示上述目标物体的朝向。

其中, 上述朝向角可以是用户输入的, 也可以是预设设置的, 本申请
5 实施例对此不做特别限制, 例如, 以目标物体为车辆为例, 用户可以不输入朝向角, 默认目标物体朝向正前方且朝向角为 0 度。当然用户也可以输入其他朝向角的角度。例如, 当目标物体为车辆时, 当车辆朝向正前方行驶时, 则可以选择默认的朝向角; 而当出现某些斜向行驶或逆向行驶的车辆时, 则可以根据实际情况输入实际的朝向角。

10 **S704:** 基于上述朝向角、2D 框和角点, 确定并显示上述目标物体的 3D 框。

本申请实施例, 在确定并显示上述目标物体的 3D 框之前, 还获取上述目标物体在 3D 空间的朝向角, 进而, 基于上述朝向角、2D 框和角点, 确定并显示上述目标物体的 3D 框, 使得获得的 3D 框与实际更相符。具体的,
15 的, 当朝向角取默认设置, 例如为 0 度时, 表示目标物体朝向为正前方, 此时目标物体对应的消失点位于图像中心, 从而可以根据消失点确定出目标物体其他角点的位置完成 3D 框的生成; 而当朝向角为用户输入的其他数值时, 表示目标物体朝向其他方向, 此时消失点位于其他位置, 但同样地可以根据消失点完成 3D 框的生成。本申请实施例在图像层面直接对 3D
20 框进行标注, 无需依赖额外的深度传感器, 降低成本, 该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点, 就能获得目标物体的 3D 框, 处理过程简单, 降低工作量, 另外, 本申请实施例获取的 3D 框与物体实际尺寸相差较小, 解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题, 从而, 保证后续对神经网络
25 的准确训练, 提高神经网络的识别效果。

另外, 本申请实施例在确定并显示上述目标物体的 3D 框之后, 还可以调整上述 2D 框。图 8 为本申请实施例提出的又一种 3D 框标注方法的流程示意图。如图 8 所示, 该方法包括:

S801: 获取 2D 框标注操作, 根据该 2D 框标注操作, 在包含目标物体
30 的二维图像上确定上述目标物体的 2D 框。

S802: 获取角点标注操作, 其中, 该角点位于上述 2D 框的一条边上, 根据上述角点标注操作在上述 2D 框上标注该角点。

S803: 基于上述 2D 框和上述角点, 确定并显示上述目标物体的 3D 框。

其中, 步骤 S801- S803 与上述步骤 S201- S203 的实现方式相同, 此处不再赘述。

S804: 上述 2D 框标注操作包括框选操作、移动操作和旋转操作中至少一个, 根据上述 2D 框标注操作调整上述 2D 框。

其中, 上述 2D 框标注操作除上述外, 还可以包括其它操作, 本申请实施例对此不做特别限制。

本申请实施例, 在确定并显示上述目标物体的 3D 框之后, 还可以调整上述 2D 框, 从而, 可以基于调整后的 2D 框生成新的 3D 框, 满足多种应用需要。而且, 本申请实施例在图像层面直接对 3D 框进行标注, 无需依赖额外的深度传感器, 降低成本, 该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点, 就能获得目标物体的 3D 框, 处理过程简单, 降低工作量, 另外, 本申请实施例获取的 3D 框与物体实际尺寸相差较小, 解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题, 从而, 保证后续对神经网络的准确训练, 提高神经网络的识别效果。

图 9 为本申请实施例提供的又一种 3D 框标注方法的流程示意图, 本实施例的执行主体可以为图 1 中的处理装置 11, 具体执行主体可以根据实际应用场景确定。如图 9 所示, 在图 1 所示应用场景的基础上, 本申请实施例提供的 3D 框标注方法包括如下步骤:

S901: 在包含目标物体的二维图像上确定上述目标物体的 2D 框。

这里, 处理装置 11 在确定上述目标物体的 2D 框后, 还可以获取 2D 框调整操作, 根据该操作, 调整上述 2D 框, 例如调整上述 2D 框的大小、位置等。

S902: 获取标注的角点, 其中, 该角点位于上述 2D 框的一条边上。

在一种可能的实现方式中, 上述角点位于上述 2D 框的底边上。

S903: 基于上述 2D 框和角点, 确定上述目标物体的 3D 框。

本申请实施例, 在图像层面直接对 3D 框进行标注, 无需依赖额外的深

度传感器，降低成本，而且该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，就能获得目标物体的 3D 框，处理过程简单，降低工作量，另外，本申请实施例获取的 3D 框与物体实际尺寸相差较小，解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题，从而，保证后续对神经网络的准确训练，提高神经网络的识别效果。

另外，本申请实施例在确定上述目标物体的 3D 框之前，还获取上述角点的角点编号。图 10 为本申请实施例提出的又一种 3D 框标注方法的流程示意图。如图 10 所示，该方法包括：

- 10 S1001：在包含目标物体的二维图像上确定上述目标物体的 2D 框。
- S1002：获取标注的角点，其中，该角点位于上述 2D 框的一条边上。
- 其中，步骤 S1001- S1002 与上述步骤 S901- S902 的实现方式相同，此处不再赘述。

- S1003：获取上述角点的角点编号，该角点编号用于指示上述角点相对
- 15 于上述目标物体的位置。

 其中，上述角点编号可以是用户输入的，也可以是预先配置的。

 S1004：基于上述角点编号、2D 框和角点，确定上述目标物体的 3D 框。

- 在一种可能的实现方式中，上述处理装置 11 可以基于上述角点编号和
- 20 上述角点，确定上述 2D 框和 3D 框的对应关系，进而，根据该对应关系和上述 2D 框，确定上述 3D 框。

 示例性的，还以目标物体为车辆为例，上述 2D 框和 3D 框的对应关系，如图 11 所示，图中 front 表示前方，rear 表示后方。

- 在本申请实施例中，上述确定上述 2D 框和 3D 框的对应关系，可以
- 25 包括：

 获取预存的物体 2D 框与其 3D 框的角点的对应规则；

 根据上述对应规则，以及上述角点编号和上述角点，确定上述对应关系。

- 其中，上述对应规则可以根据实际情况设置，本申请实施例对此不做
- 30 特别限制。

本申请实施例，在确定上述目标物体的 3D 框之前，还获取上述角点的角点编号，进而，基于上述角点编号、2D 框和角点，准确确定上述目标物体的 3D 框，满足应用需要。而且，本申请实施例在图像层面直接对 3D 框进行标注，无需依赖额外的深度传感器，降低成本，该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，就能获得目标物体的 3D 框，处理过程简单，降低工作量，另外，本申请实施例获取的 3D 框与物体实际尺寸相差较小，解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差较大的问题，从而，保证后续对神经网络的准确训练，提高神经网络的识别效果。

另外，本申请实施例在确定上述目标物体的 3D 框之前，还获取上述目标物体在 3D 空间的朝向角。图 12 为本申请实施例提出的又一种 3D 框标注方法的流程示意图。如图 12 所示，该方法包括：

S1201：在包含目标物体的二维图像上确定上述目标物体的 2D 框。

S1202：获取标注的角点，其中，该角点位于上述 2D 框的一条边上。

其中，步骤 S1201- S1202 与上述步骤 S901- S902 的实现方式相同，此处不再赘述。

S1203：获取上述目标物体在 3D 空间的朝向角，该朝向角用于指示上述目标物体的朝向。

其中，上述朝向角可以是用户输入的，也可以是预先配置的。

S1204：基于上述朝向角、2D 框和角点，确定上述目标物体的 3D 框。

在一种可能的实现方式中，上述处理装置 11 可以根据上述朝向角，确定上述 3D 框对应的三个消失点，进而，根据该三个消失点、上述 2D 框和角点，确定上述 3D 框。示例性的，结合长方体的边的平行关系，根据上述三个消失点、上述 2D 框和角点，确定上述 3D 框。

其中，上述根据上述朝向角，确定上述 3D 框对应的三个消失点，可以包括：

获取上述二维图像对应的图像获取装置的投影矩阵；

根据上述投影矩阵和上述朝向角的投影矩阵，确定上述三个消失点。

示例性的，将上述投影矩阵和上述朝向角的投影矩阵相乘，根据矩阵相乘结果确定上述三个消失点，例如，如图 11 所示，三个消失点 $vp0$ 、 $vp1$

和 vp2。

另外，在上述根据上述三个消失点、2D 框和角点，确定上述 3D 框之前，还包括：

5 获取上述角点的角点编号，该角点编号用于指示上述角点相对于所述目标物体的位置。

相应的，上述根据上述三个消失点、2D 框和角点，确定上述 3D 框，包括：

根据上述角点编号、三个消失点、2D 框和角点，确定上述 3D 框。

10 示例性的，上述处理装置 11 可以基于上述角点编号和角点，确定上述 2D 框和上述 3D 框的对应关系，进而，结合长方体的边的平行关系，根据该对应关系、上述三个消失点和 2D 框，求解 3D 框投影的 8 个角点位置，例如如图 11 中角点 p0、p1、p2、p3、p4、p5、p6 和 p7，从而，确定上述 3D 框。

15 这里，本申请实施例将长方体的内在几何关系与上述图像获取装置的投影模型相结合，从而可以保证标注得到的伪 3D 框满足长方体的内在几何关系，标注精度和标注一致性更高。

本申请实施例，在确定并显示上述目标物体的 3D 框之前，还获取上述目标物体在 3D 空间的朝向角，进而，基于上述朝向角、2D 框和角点，确定并显示上述目标物体的 3D 框，使得获得的 3D 框与实际更相符。而且，
20 本申请实施例在图像层面直接对 3D 框进行标注，无需依赖额外的深度传感器，降低成本，该方法仅通过在包含目标物体的二维图像上标注目标物体的 2D 框和该 2D 框上的一角点，就能获得目标物体的 3D 框，处理过程简单，降低工作量，另外，本申请实施例获取的 3D 框与物体实际尺寸相差较小，解决现有基于点云数据标注获得的 3D 框与物体的实际尺寸相差
25 较大的问题，从而，保证后续对神经网络的准确训练，提高神经网络的识别效果。

对应于上文实施例的 3D 框标注方法，图 13 为本申请实施例提供的一种 3D 框标注装置的结构示意图。为了便于说明，仅示出了与本申请实施例相关的部分。图 13 为本申请实施例提供的一种 3D 框标注装置的结构示
30 意图，该 3D 框标注装置 1300 包括：第一获取模块 1301、第二获取模块

1302 以及显示模块 1303。这里的 3D 框标注装置可以是上述处理装置 11 本身，或者是实现处理装置 11 的功能的芯片或者集成电路。这里需要说明的是，第一获取模块、第二获取模块以及显示模块的划分只是一种逻辑功能的划分，物理上两者可以是集成的，也可以是独立的。

5 其中，第一获取模块 1301，用于获取 2D 框标注操作，根据所述 2D 框标注操作，在包含目标物体的二维图像上确定所述目标物体的 2D 框。

第二获取模块 1302，用于获取角点标注操作，其中，所述角点位于所述 2D 框的一条边上，根据所述角点标注操作在所述 2D 框上标注所述角点；

10 显示模块 1303，用于基于所述 2D 框和所述角点，确定并显示所述目标物体的 3D 框。

在一种可能的实现方式中，在所述显示模块 1303 确定并显示所述目标物体的 3D 框之前，还用于：

15 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

在一种可能的实现方式中，所述显示模块 1303，具体用于：

基于所述角点编号、所述 2D 框和所述角点，确定并显示所述 3D 框。

在一种可能的实现方式中，在所述显示模块 1303 确定并显示所述目标物体的 3D 框之前，还用于：

20 获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

在一种可能的实现方式中，所述显示模块 1303，具体用于：

基于所述朝向角、所述 2D 框和所述角点，确定并显示所述 3D 框。

在一种可能的实现方式中，所述角点位于所述 2D 框的底边上。

25 在一种可能的实现方式中，所述 2D 框标注操作包括框选操作、移动操作和旋转操作中至少一个。

在一种可能的实现方式中，在所述显示模块 1303 确定并显示所述目标物体的 3D 框之后，还用于：

根据所述 2D 框标注操作调整所述 2D 框。

30 在一种可能的实现方式中，所述角点编号是用户输入或预先配置的。

在一种可能的实现方式中，所述朝向角是用户输入或预先配置的。

本申请实施例提供的装置，可用于执行上述方法实施例的技术方案，其实现原理和技术效果类似，本申请实施例此处不再赘述。

图 14 为本申请实施例提供的另一种 3D 框标注装置的结构示意图，该 3D 框标注装置 1400 包括：第一确定模块 1401、第三获取模块 1402 以及第二确定模块 1403。这里的 3D 框标注装置可以是上述处理装置 11 本身，或者是实现处理装置 11 的功能的芯片或者集成电路。这里需要说明的是，第一确定模块、第三获取模块以及第二确定模块的划分只是一种逻辑功能的划分，物理上两者可以是集成的，也可以是独立的。

其中，第一确定模块 1401，用于在包含目标物体的二维图像上确定所述目标物体的 2D 框。

第三获取模块 1402，用于获取标注的角点，其中，所述角点位于所述 2D 框的一条边上；

第二确定模块 1403，用于基于所述 2D 框和所述角点，确定所述目标物体的 3D 框。

在一种可能的实现方式中，在所述第二确定模块 1403 确定所述目标物体的 3D 框之前，还用于：

获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

在一种可能的实现方式中，所述第二确定模块 1403，具体用于：基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框。

在一种可能的实现方式中，所述第二确定模块 1403 基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框，包括：

基于所述角点编号和所述角点，确定所述 2D 框和所述 3D 框的对应关系；

根据所述对应关系和所述 2D 框，确定所述 3D 框。

在一种可能的实现方式中，所述第二确定模块 1403 确定所述 2D 框和所述 3D 框的对应关系，包括：

获取预存的物体 2D 框与其 3D 框的角点的对应规则；

根据所述对应规则，以及所述角点编号和所述角点，确定所述对应关

系。

在一种可能的实现方式中，在所述第二确定模块 1403 确定所述目标物体的 3D 框之前，还用于：

获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

在一种可能的实现方式中，所述第二确定模块 1403，具体用于：

基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框。

在一种可能的实现方式中，所述第二确定模块 1403 基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框，包括：

根据所述朝向角，确定所述 3D 框对应的三个消失点；

根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

在一种可能的实现方式中，所述第二确定模块 1403 根据所述朝向角，确定所述 3D 框对应的三个消失点，包括：

获取所述二维图像对应的图像获取装置的投影矩阵；

根据所述投影矩阵和所述朝向角的投影矩阵，确定所述三个消失点。

在一种可能的实现方式中，在所述第二确定模块 1403 根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框之前，还用于：

获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置；

所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框，包括：

根据所述角点编号、所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

在一种可能的实现方式中，所述角点位于所述 2D 框的底边上。

在一种可能的实现方式中，所述角点编号是用户输入或预先配置的。

在一种可能的实现方式中，所述朝向角是用户输入或预先配置的。

本申请实施例提供的装置，可用于执行上述方法实施例的技术方案，其实现原理和技术效果类似，本申请实施例此处不再赘述。

可选地，图 15 示意性地提供本申请所述 3D 框标注设备的一种可能的基本硬件架构。

参见图 15，3D 框标注设备 1500 包括至少一个处理器 1501 以及存储器 1502。进一步可选的，还可以包括通信接口 1503 和总线 1504。

其中，3D 框标注设备 1500 可以是计算机或服务器，本申请对此不作特别限制。3D 框标注设备 1500 中，处理器 1501 的数量可以是一个或多个，图 15 仅示意了其中一个处理器 1501。可选地，处理器 1501，可以是中央处理器（central processing unit, CPU）、图形处理器（graphics processing unit, GPU）或者数字信号处理器（digital signal processor, DSP）。如果 3D 框标注设备 1500 具有多个处理器 1501，多个处理器 1501 的类型可以不同，或者可以相同。可选地，3D 框标注设备 1500 的多个处理器 1501 还可以集成为多核处理器。

存储器 1502 存储计算机指令和数据；存储器 1502 可以存储实现本申请提供的上述 3D 框标注方法所需的计算机指令和数据，例如，存储器 1502 存储用于实现上述 3D 框标注方法的步骤的指令。存储器 1502 可以是以下存储介质的任一种或任一种组合：非易失性存储器（例如只读存储器（ROM）、固态硬盘（SSD）、硬盘（HDD）、光盘），易失性存储器。

通信接口 1503 可以为所述至少一个处理器提供信息输入/输出。也可以包括以下器件的任一种或任一种组合：网络接口（例如以太网接口）、无线网卡等具有网络接入功能的器件。

可选的，通信接口 1503 还可以用于 3D 框标注设备 1500 与其它计算设备或者终端进行数据通信。

进一步可选的，图 15 用一条粗线表示总线 1504。总线 1504 可以将处理器 1501 与存储器 1502 和通信接口 1503 连接。这样，通过总线 1504，处理器 1501 可以访问存储器 1502，还可以利用通信接口 1503 与其它计算设备或者终端进行数据交互。

在本申请中，处理器 1501 执行存储器 1502 中的计算机指令，实现如下步骤：

获取 2D 框标注操作，根据所述 2D 框标注操作，在包含目标物体的二维图像上确定所述目标物体的 2D 框；

获取角点标注操作，其中，所述角点位于所述 2D 框的一条边上，根据所述角点标注操作在所述 2D 框上标注所述角点；

基于所述 2D 框和所述角点，确定并显示所述目标物体的 3D 框。

在一种可能的实现方式中，在所述确定并显示所述目标物体的 3D 框之前，所述处理器 1501 执行所述计算机指令时还实现如下步骤：

5 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

在一种可能的实现方式中，所述确定并显示所述目标物体的 3D 框，包括：

基于所述角点编号、所述 2D 框和所述角点，确定并显示所述 3D 框。

10 在一种可能的实现方式中，在所述确定并显示所述目标物体的 3D 框之前，所述处理器 1501 执行所述计算机指令时还实现如下步骤：

获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

在一种可能的实现方式中，所述确定并显示所述目标物体的 3D 框，包括：

15 基于所述朝向角、所述 2D 框和所述角点，确定并显示所述 3D 框。

在一种可能的实现方式中，所述角点位于所述 2D 框的底边上。

在一种可能的实现方式中，所述 2D 框标注操作包括框选操作、移动操作和旋转操作中至少一个。

20 在一种可能的实现方式中，在所述确定并显示所述目标物体的 3D 框之后，所述处理器 1501 执行所述计算机指令时还实现如下步骤：

根据所述 2D 框标注操作调整所述 2D 框。

在一种可能的实现方式中，所述角点编号是用户输入或预先配置的。

在一种可能的实现方式中，所述朝向角是用户输入或预先配置的。

25 另外，从逻辑功能划分来看，示例性的，如图 15 所示，存储器 1502 中可以包括第一获取模块 1301、第二获取模块 1302 以及显示模块 1303。这里的包括仅仅涉及存储器中所存储的指令被执行时可以分别实现第一获取模块 1301、第二获取模块 1302 以及显示模块 1303 的功能，而不限定是物理上的结构。

30 上述的 3D 框标注设备除了可以像上述图 15 通过软件实现外，也可以作为硬件模块，或者作为电路单元，通过硬件实现。

可选地，图 16 示意性地提供本申请所述 3D 框标注设备的另一种可能的基本硬件架构。

参见图 16，3D 框标注设备 1600 包括至少一个处理器 1601 以及存储器 1602。进一步可选的，还可以包括通信接口 1603 和总线 1604。

5 其中，3D 框标注设备 1600 可以是计算机或服务器，本申请对此不作特别限制。3D 框标注设备 1600 中，处理器 1601 的数量可以是一个或多个，图 16 仅示意了其中一个处理器 1601。可选地，处理器 1601，可以是中央处理器(central processing unit, CPU)、图形处理器(graphics processing unit, GPU)或者数字信号处理器(digital signal processor, DSP)。如果
10 3D 框标注设备 1600 具有多个处理器 1601，多个处理器 1601 的类型可以不同，或者可以相同。可选地，3D 框标注设备 1600 的多个处理器 1601 还可以集成为多核处理器。

存储器 1602 存储计算机指令和数据；存储器 1602 可以存储实现本申请提供的上述并行执行单元的管理方法所需的计算机指令和数据，例如，
15 存储器 1602 存储用于实现上述并行执行单元的管理方法的步骤的指令。存储器 1602 可以是以下存储介质的任一种或任一种组合：非易失性存储器(例如只读存储器(ROM)、固态硬盘(SSD)、硬盘(HDD)、光盘)，易失性存储器。

通信接口 1603 可以为所述至少一个处理器提供信息输入/输出。也可以
20 包括以下器件的任一种或任一种组合：网络接口(例如以太网接口)、无线网卡等具有网络接入功能的器件。

可选的，通信接口 1603 还可以用于 3D 框标注设备 1600 与其它计算设备或者终端进行数据通信。

进一步可选的，图 16 用一条粗线表示总线 1604。总线 1604 可以将处
25 理器 1601 与存储器 1602 和通信接口 1603 连接。这样，通过总线 1604，处理器 1601 可以访问存储器 1602，还可以利用通信接口 1603 与其它计算设备或者终端进行数据交互。

在本申请中，处理器 1601 执行存储器 1602 中的计算机指令，实现如下步骤：

30 在包含目标物体的二维图像上确定所述目标物体的 2D 框；

获取标注的角点，其中，所述角点位于所述 2D 框的一条边上；

基于所述 2D 框和所述角点，确定所述目标物体的 3D 框。

在一种可能的实现方式中，在所述确定所述目标物体的 3D 框之前，所述处理器 1601 执行所述计算机指令时还实现如下步骤：

- 5 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

在一种可能的实现方式中，所述确定所述目标物体的 3D 框，包括：

基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框。

- 10 在一种可能的实现方式中，所述基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框，包括：

基于所述角点编号和所述角点，确定所述 2D 框和所述 3D 框的对应关系；

根据所述对应关系和所述 2D 框，确定所述 3D 框。

- 15 在一种可能的实现方式中，所述确定所述 2D 框和所述 3D 框的对应关系，包括：

获取预存的物体 2D 框与其 3D 框的角点的对应规则；

根据所述对应规则，以及所述角点编号和所述角点，确定所述对应关系。

- 20 在一种可能的实现方式中，在所述确定所述目标物体的 3D 框之前，所述处理器 1601 执行所述计算机指令时还实现如下步骤：

获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

在一种可能的实现方式中，所述确定所述目标物体的 3D 框，包括：

基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框。

- 25 在一种可能的实现方式中，所述基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框，包括：

根据所述朝向角，确定所述 3D 框对应的三个消失点；

根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

- 30 在一种可能的实现方式中，所述根据所述朝向角，确定所述 3D 框对应的三个消失点，包括：

获取所述二维图像对应的图像获取装置的投影矩阵；

根据所述投影矩阵和所述朝向角的投影矩阵，确定所述三个消失点。

在一种可能的实现方式中，在所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框之前，所述处理器 1601 执行所述计算机指令时还实现如下步骤：

获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置；

所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框，包括：

10 根据所述角点编号、所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

在一种可能的实现方式中，所述角点位于所述 2D 框的底边上。

在一种可能的实现方式中，所述角点编号是用户输入或预先配置的。

在一种可能的实现方式中，所述朝向角是用户输入或预先配置的。

15 另外，从逻辑功能划分来看，示例性的，如图 16 所示，存储器 1602 中可以包括第一确定模块 1401、第三获取模块 1402 以及第二确定模块 1403。这里的包括仅仅涉及存储器中所存储的指令被执行时可以分别实现第一确定模块 1401、第三获取模块 1402 以及第二确定模块 1403 的功能，而不限定是物理上的结构。

20 上述的 3D 框标注设备除了可以像上述图 16 通过软件实现外，也可以作为硬件模块，或者作为电路单元，通过硬件实现。

另外，本申请实施例提供一种神经网络训练方法，包括：利用如上所述的 3D 框标注方法确定的目标物体的 3D 框，以及包含上述目标物体的二维图像，进行神经网络的训练。

25 本申请提供一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得所述计算机执行上述 3D 框标注方法。

本申请提供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指令，当其在计算机上运行时，使得所述计算机执行上述 3D 框标注方法。

30 本申请提供一种可移动平台，所述可移动平台可以为智能设备或者运

输工具，例如无人机、无人车或者机器人等，其上包含上述 3D 框标注设备。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，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，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。

10 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。

15 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，也可以采用硬件加软件功能单元的形式实现。

权利要求书

1、一种 3D 框标注方法，其特征在于，包括：

获取 2D 框标注操作，根据所述 2D 框标注操作，在包含目标物体的二维图像上确定所述目标物体的 2D 框；

5 获取角点标注操作，其中，所述角点位于所述 2D 框的一条边上，根据所述角点标注操作在所述 2D 框上标注所述角点；

基于所述 2D 框和所述角点，确定并显示所述目标物体的 3D 框。

2、根据权利要求 1 所述的方法，其特征在于，在所述确定并显示所述目标物体的 3D 框之前，还包括：

10 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

3、根据权利要求 2 所述的方法，其特征在于，所述确定并显示所述目标物体的 3D 框，包括：

基于所述角点编号、所述 2D 框和所述角点，确定并显示所述 3D 框。

15 4、根据权利要求 1 所述的方法，其特征在于，在所述确定并显示所述目标物体的 3D 框之前，还包括：

获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

5、根据权利要求 4 所述的方法，其特征在于，所述确定并显示所述目标物体的 3D 框，包括：

基于所述朝向角、所述 2D 框和所述角点，确定并显示所述 3D 框。

6、根据权利要求 1 至 5 任一项所述的方法，其特征在于，所述角点位于所述 2D 框的底边上。

7、根据权利要求 1 至 6 中任一项所述的方法，其特征在于，所述 2D 框标注操作包括框选操作、移动操作和旋转操作中至少一个。

8、根据权利要求 7 所述的方法，其特征在于，在所述确定并显示所述目标物体的 3D 框之后，还包括：

根据所述 2D 框标注操作调整所述 2D 框。

9、根据权利要求 2 或 3 所述的方法，其特征在于，所述角点编号是用户输入或预先配置的。

10、根据权利要求 4 或 5 所述的方法，其特征在于，所述朝向角是用户输入或预先配置的。

11、一种 3D 框标注方法，其特征在于，包括：

在包含目标物体的二维图像上确定所述目标物体的 2D 框；

5 获取标注的角点，其中，所述角点位于所述 2D 框的一条边上；

基于所述 2D 框和所述角点，确定所述目标物体的 3D 框。

12、根据权利要求 11 所述的方法，其特征在于，在所述确定所述目标物体的 3D 框之前，还包括：

10 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

13、根据权利要求 12 所述的方法，其特征在于，所述确定所述目标物体的 3D 框，包括：

基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框。

15 14、根据权利要求 13 所述的方法，其特征在于，所述基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框，包括：

基于所述角点编号和所述角点，确定所述 2D 框和所述 3D 框的对应关系；

根据所述对应关系和所述 2D 框，确定所述 3D 框。

20 15、根据权利要求 14 所述的方法，其特征在于，所述确定所述 2D 框和所述 3D 框的对应关系，包括：

获取预存的物体 2D 框与其 3D 框的角点的对应规则；

根据所述对应规则，以及所述角点编号和所述角点，确定所述对应关系。

25 16、根据权利要求 11 所述的方法，其特征在于，在所述确定所述目标物体的 3D 框之前，还包括：

获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

17、根据权利要求 16 所述的方法，其特征在于，所述确定所述目标物体的 3D 框，包括：

30 基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框。

18、根据权利要求 17 所述的方法，其特征在于，所述基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框，包括：

根据所述朝向角，确定所述 3D 框对应的三个消失点；

根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

5 19、根据权利要求 18 所述的方法，其特征在于，所述根据所述朝向角，确定所述 3D 框对应的三个消失点，包括：

获取所述二维图像对应的图像获取装置的投影矩阵；

根据所述投影矩阵和所述朝向角的投影矩阵，确定所述三个消失点。

10 20、根据权利要求 18 或 19 所述的方法，其特征在于，在所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框之前，还包括：

获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置；

所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框，包括：

15 根据所述角点编号、所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

21、根据权利要求 11 至 20 中任一项所述的方法，其特征在于，所述角点位于所述 2D 框的底边上。

20 22、根据权利要求 12 至 15 中任一项所述的方法，其特征在于，所述角点编号是用户输入或预先配置的。

23、根据权利要求 16 至 20 中任一项所述的方法，其特征在于，所述朝向角是用户输入或预先配置的。

25 24、一种 3D 框标注设备，其特征在于，包括存储器、处理器，以及存储在该存储器中并可在该处理器上运行的计算机指令，该处理器执行该计算机指令时实现如下步骤：

获取 2D 框标注操作，根据所述 2D 框标注操作，在包含目标物体的二维图像上确定所述目标物体的 2D 框；

获取角点标注操作，其中，所述角点位于所述 2D 框的一条边上，根据所述角点标注操作在所述 2D 框上标注所述角点；

30 基于所述 2D 框和所述角点，确定并显示所述目标物体的 3D 框。

25、根据权利要求 24 所述的设备，其特征在于，在所述确定并显示所述目标物体的 3D 框之前，所述处理器执行所述计算机指令时还实现如下步骤：

5 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

26、根据权利要求 25 所述的设备，其特征在于，所述确定并显示所述目标物体的 3D 框，包括：

基于所述角点编号、所述 2D 框和所述角点，确定并显示所述 3D 框。

10 27、根据权利要求 24 所述的设备，其特征在于，在所述确定并显示所述目标物体的 3D 框之前，所述处理器执行所述计算机指令时还实现如下步骤：

获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

15 28、根据权利要求 27 所述的设备，其特征在于，所述确定并显示所述目标物体的 3D 框，包括：

基于所述朝向角、所述 2D 框和所述角点，确定并显示所述 3D 框。

29、根据权利要求 24 至 28 中任一项所述的设备，其特征在于，所述角点位于所述 2D 框的底边上。

20 30、根据权利要求 24 至 29 中任一项所述的设备，其特征在于，所述 2D 框标注操作包括框选操作、移动操作和旋转操作中至少一个。

31、根据权利要求 30 所述的设备，其特征在于，在所述确定并显示所述目标物体的 3D 框之后，所述处理器执行所述计算机指令时还实现如下步骤：

根据所述 2D 框标注操作调整所述 2D 框。

25 32、根据权利要求 25 或 26 所述的设备，其特征在于，所述角点编号是用户输入或预先配置的。

33、根据权利要求 27 或 28 所述的设备，其特征在于，所述朝向角是用户输入或预先配置的。

30 34、一种 3D 框标注设备，其特征在于，包括存储器、处理器，以及存储在该存储器中并可在所述处理器上运行的计算机指令，所述处理器

执行所述计算机指令时实现如下步骤：

在包含目标物体的二维图像上确定所述目标物体的 2D 框；
获取标注的角点，其中，所述角点位于所述 2D 框的一条边上；
基于所述 2D 框和所述角点，确定所述目标物体的 3D 框。

5 35、根据权利要求 34 所述的设备，其特征在于，在所述确定所述目标物体的 3D 框之前，所述处理器执行所述计算机指令时还实现如下步骤：

获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置。

10 36、根据权利要求 35 所述的设备，其特征在于，所述确定所述目标物体的 3D 框，包括：

基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框。

37、根据权利要求 36 所述的设备，其特征在于，所述基于所述角点编号、所述 2D 框和所述角点，确定所述 3D 框，包括：

15 基于所述角点编号和所述角点，确定所述 2D 框和所述 3D 框的对应关系；

根据所述对应关系和所述 2D 框，确定所述 3D 框。

38、根据权利要求 37 所述的设备，其特征在于，所述确定所述 2D 框和所述 3D 框的对应关系，包括：

获取预存的物体 2D 框与其 3D 框的角点的对应规则；

20 根据所述对应规则，以及所述角点编号和所述角点，确定所述对应关系。

39、根据权利要求 34 所述的设备，其特征在于，在所述确定所述目标物体的 3D 框之前，所述处理器执行所述计算机指令时还实现如下步骤：

25 获取所述目标物体在 3D 空间的朝向角，所述朝向角用于指示所述目标物体的朝向。

40、根据权利要求 39 所述的设备，其特征在于，所述确定所述目标物体的 3D 框，包括：

基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框。

30 41、根据权利要求 40 所述的设备，其特征在于，所述基于所述朝向角、所述 2D 框和所述角点，确定所述 3D 框，包括：

根据所述朝向角，确定所述 3D 框对应的三个消失点；

根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

42、根据权利要求 41 所述的设备，其特征在于，所述根据所述朝向角，确定所述 3D 框对应的三个消失点，包括：

5 获取所述二维图像对应的图像获取装置的投影矩阵；

根据所述投影矩阵和所述朝向角的投影矩阵，确定所述三个消失点。

43、根据权利要求 41 或 42 所述的设备，其特征在于，在所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框之前，所述处理器执行所述计算机指令时还实现如下步骤：

10 获取所述角点的角点编号，所述角点编号用于指示所述角点相对于所述目标物体的位置；

所述根据所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框，包括：

15 根据所述角点编号、所述三个消失点、所述 2D 框和所述角点，确定所述 3D 框。

44、根据权利要求 34 至 43 中任一项所述的设备，其特征在于，所述角点位于所述 2D 框的底边上。

45、根据权利要求 35 至 38 中任一项所述的设备，其特征在于，所述角点编号是用户输入或预先配置的。

20 46、根据权利要求 39 至 43 中任一项所述的设备，其特征在于，所述朝向角是用户输入或预先配置的。

47、一种神经网络训练方法，其特征在于，包括：

利用权利要求 1 至 10 任一项所述的 3D 框标注方法确定的目标物体的 3D 框，以及包含所述目标物体的二维图像，进行神经网络的训练。

25 48、一种神经网络训练方法，其特征在于，包括：

利用权利要求 11 至 23 任一项所述的 3D 框标注方法确定的目标物体的 3D 框，以及包含所述目标物体的二维图像，进行神经网络的训练。

49、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中存储有计算机指令，当处理器执行所述计算机指令时，实现如权利要求 1 至 10 任一项所述的 3D 框标注方法。

30

50、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中存储有计算机指令，当处理器执行所述计算机指令时，实现如权利要求 11 至 23 任一项所述的 3D 框标注方法。

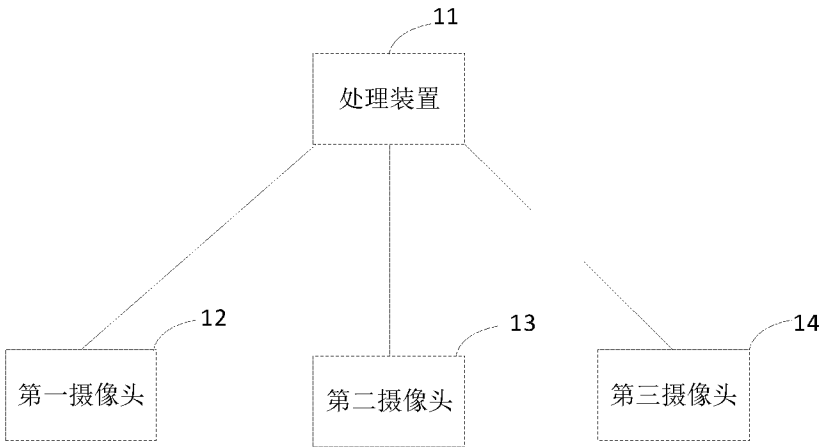


图 1

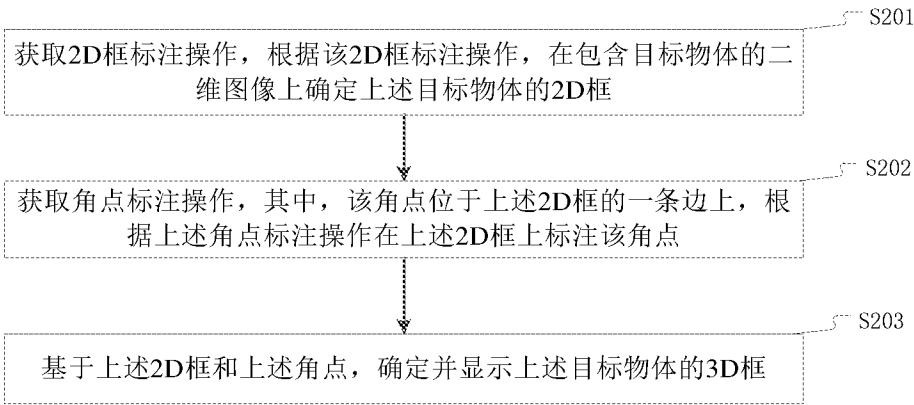


图 2

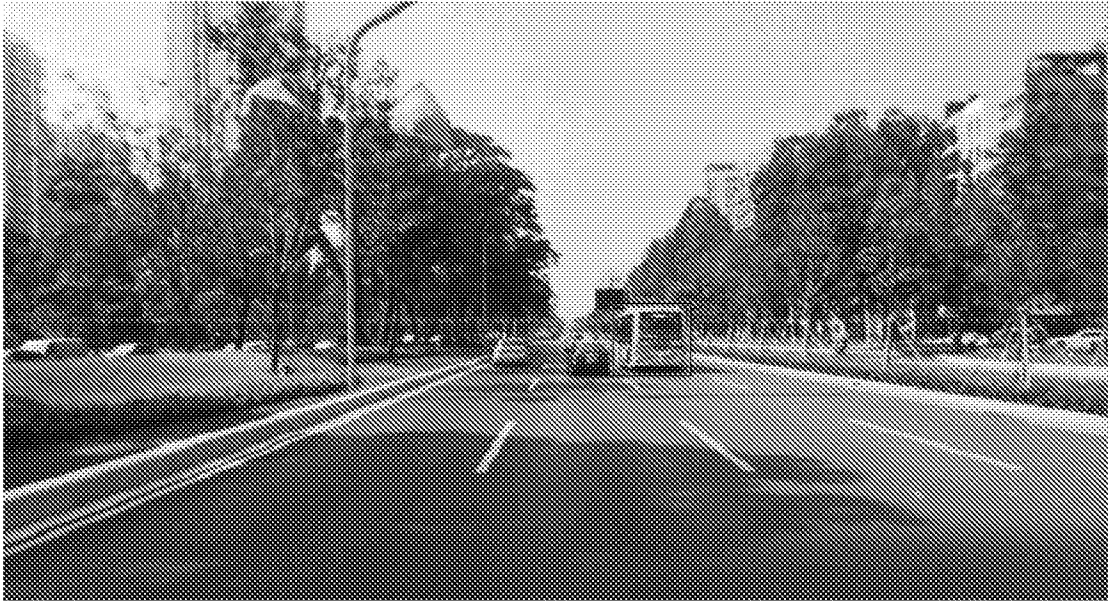


图 3



图 4



图 5

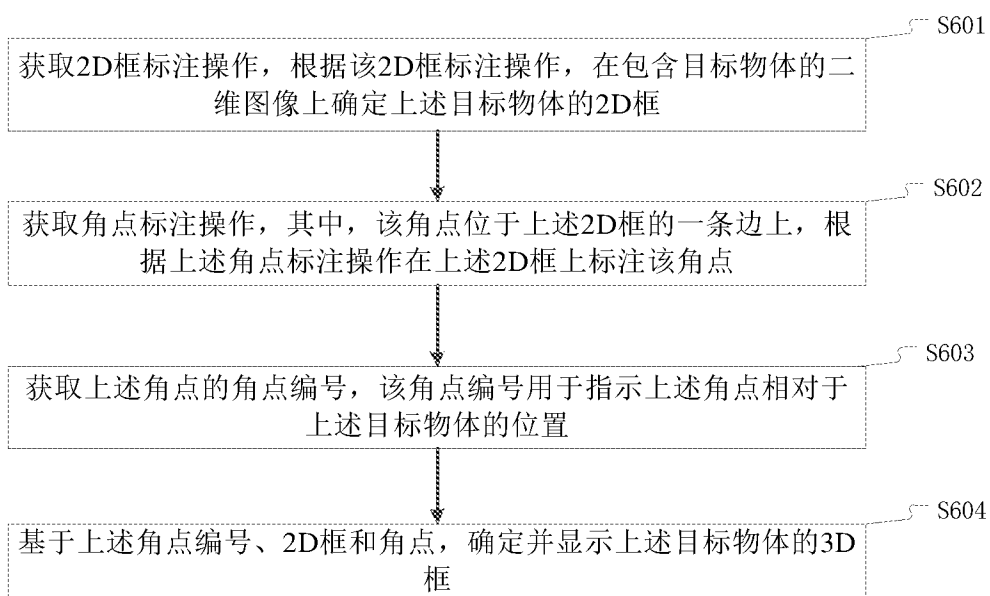


图 6

4/8

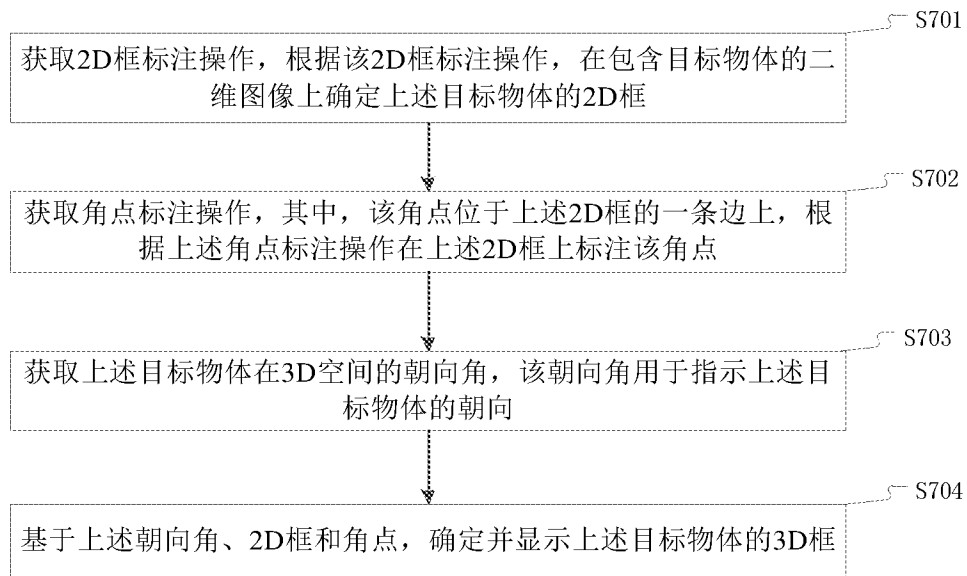


图 7

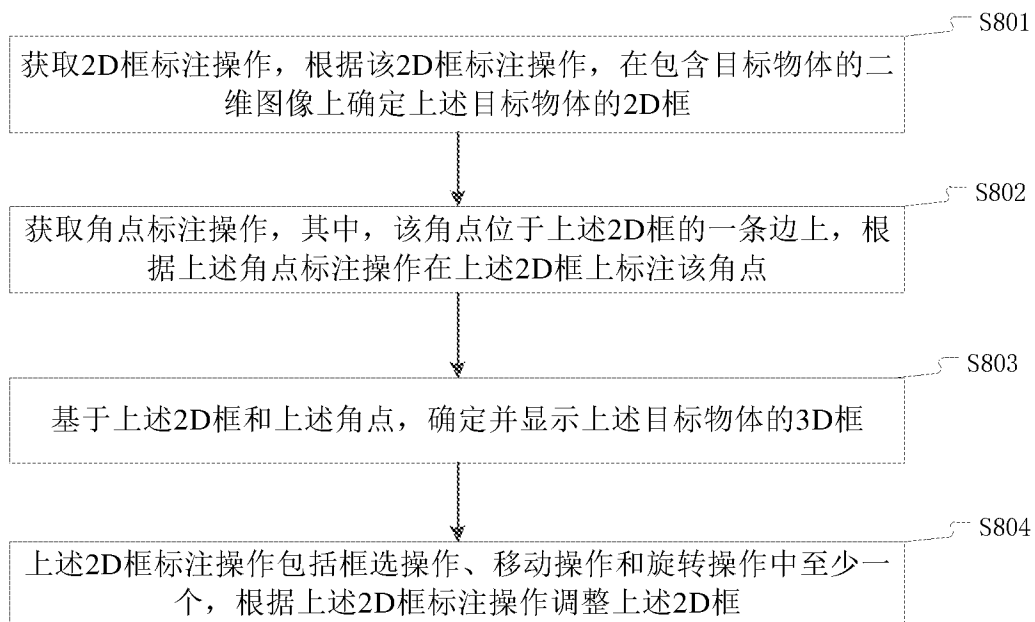


图 8

5/8

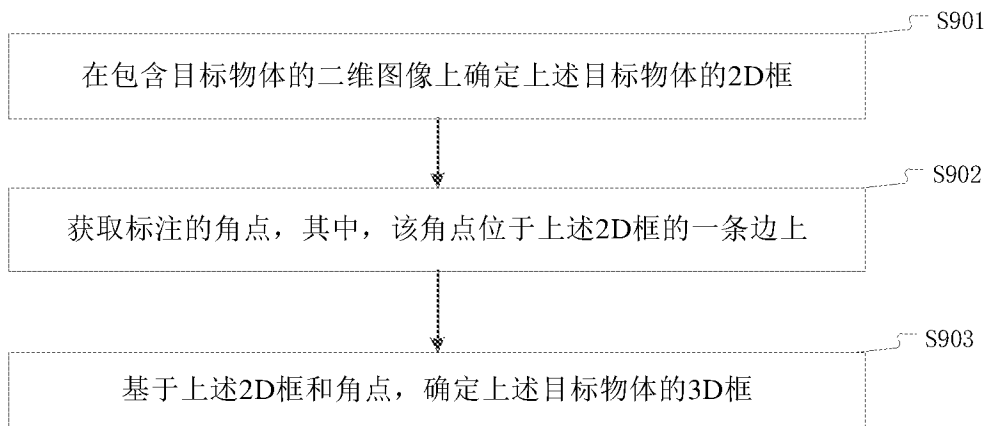


图 9

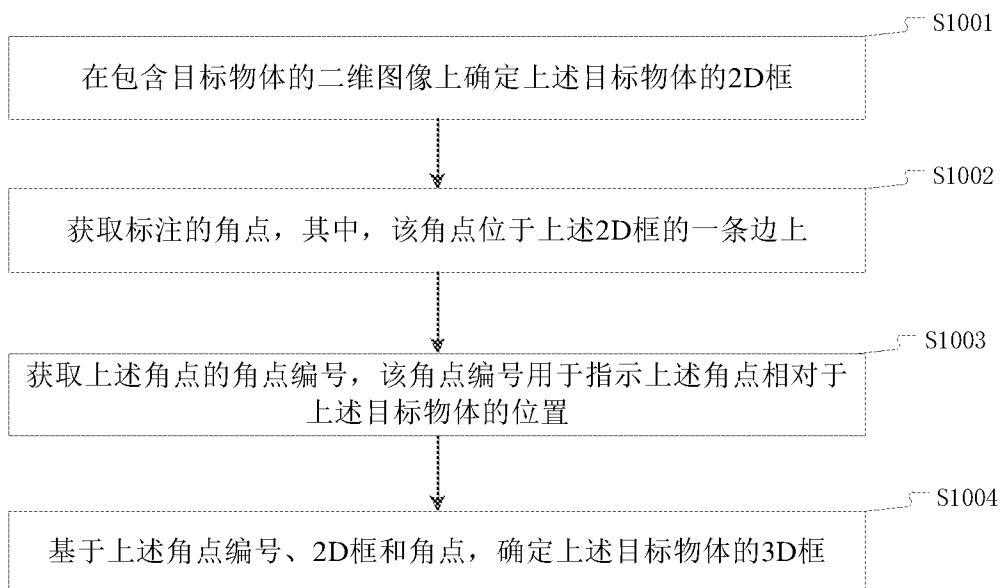


图 10

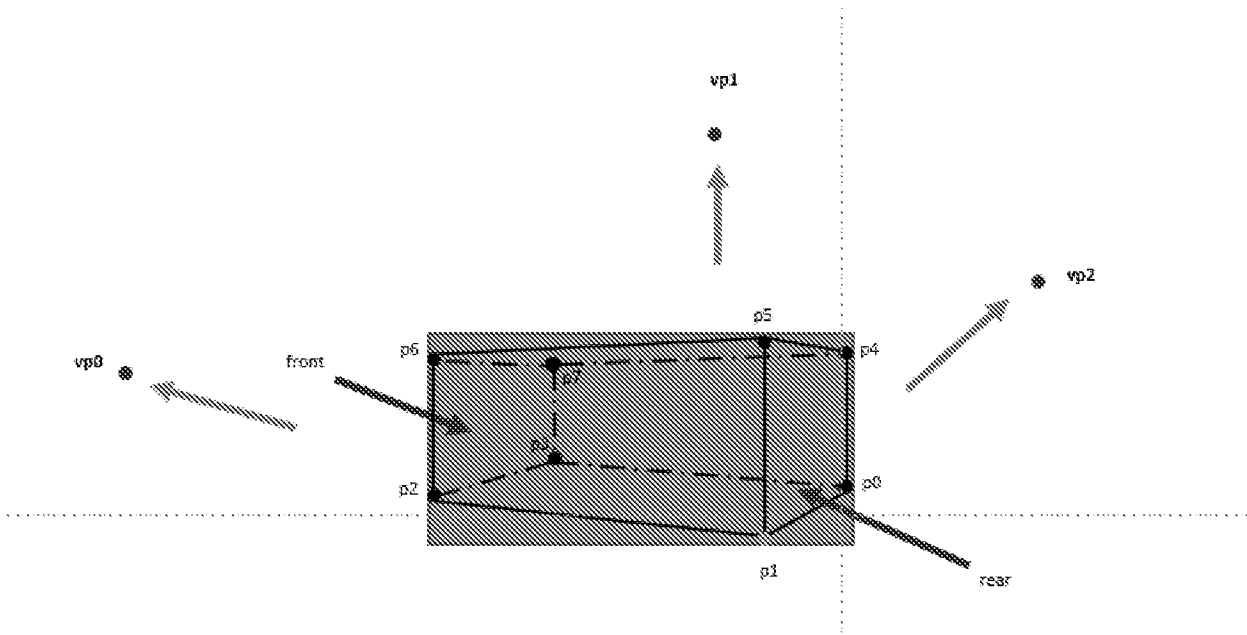


图 11

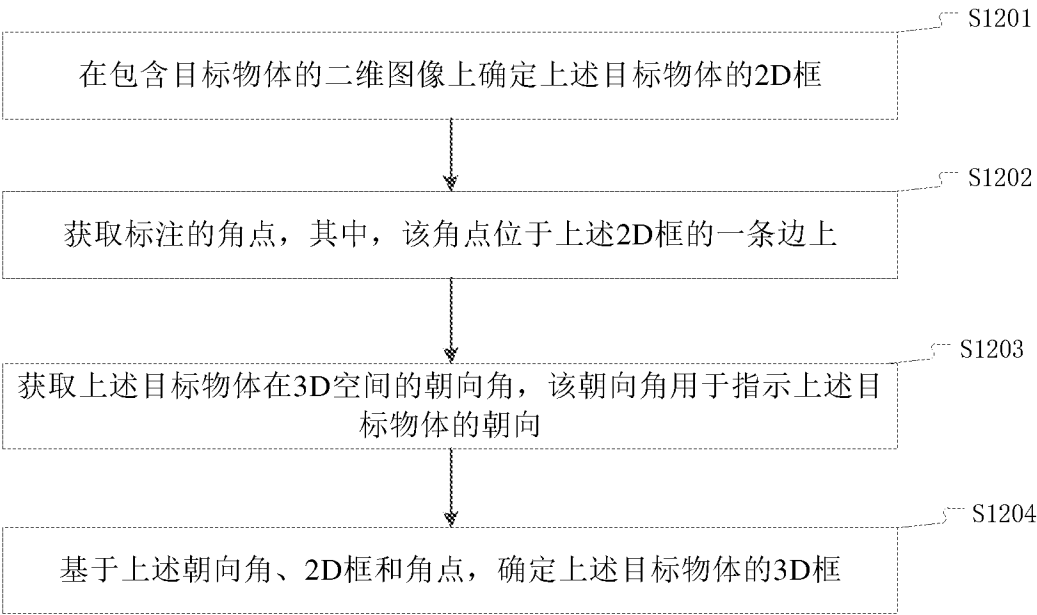


图 12

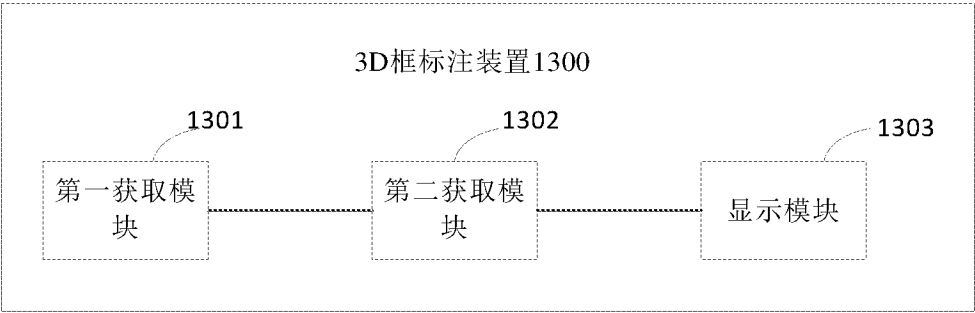


图 13

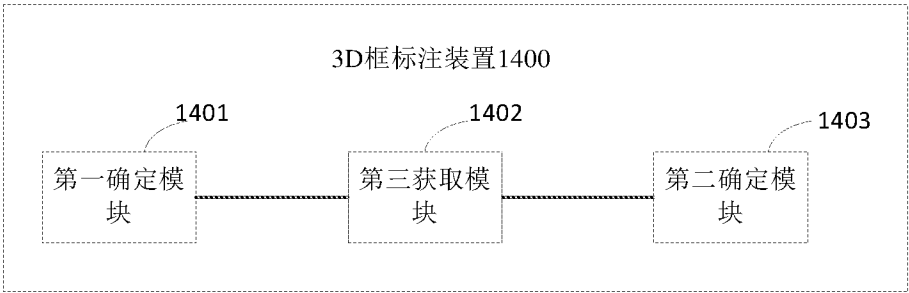


图 14

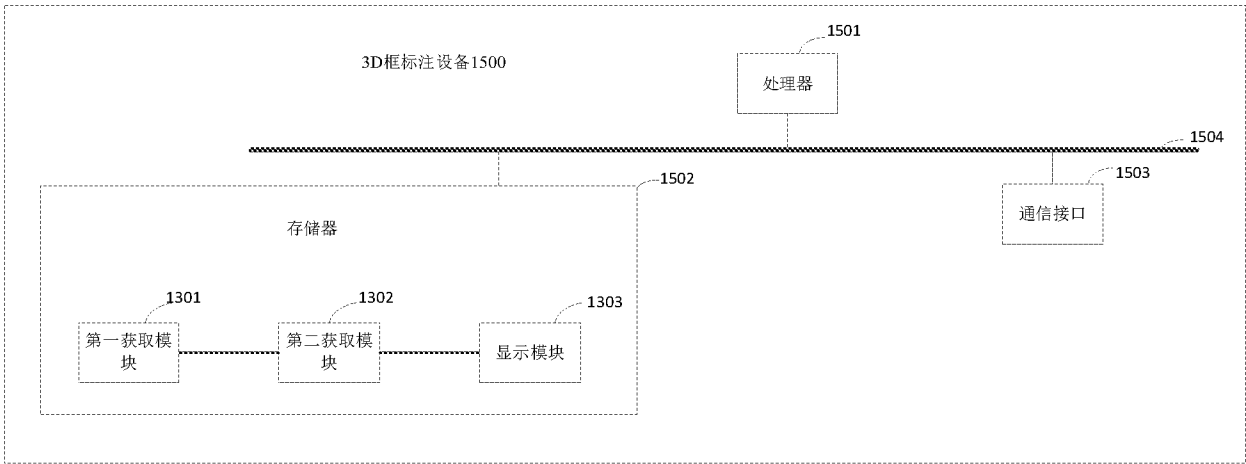


图 15

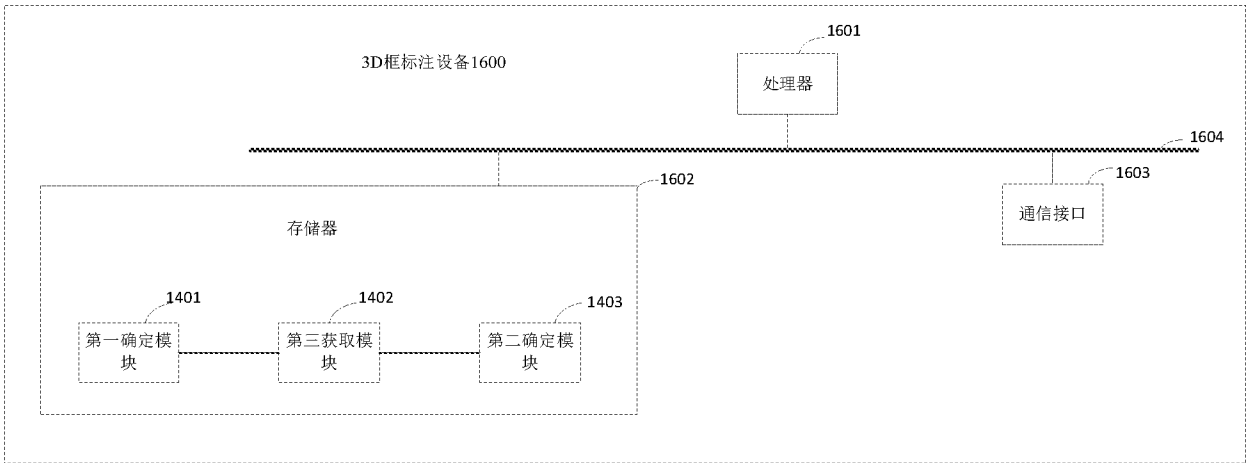


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/103263

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; DWPI; VEN; USTXT; EPTXT; WOTXT; CNKI: 框, 标注, 标记, 目标, 角点, 顶点, 二维, 三维, 朝向, frame, label, mark, target, corner point, vertex, two dimension, 2D, three dimension, 3D, orientation

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 110390258 A (SOUTHEAST UNIVERSITY) 29 October 2019 (2019-10-29) description paragraphs 4-22, 29-40	1-46, 49-50
X	CN 111126161 A (BEIJING UNION UNIVERSITY) 08 May 2020 (2020-05-08) description, paragraphs 4-31	1-46, 49-50
Y	CN 110390258 A (SOUTHEAST UNIVERSITY) 29 October 2019 (2019-10-29) description paragraphs 4-22, 29-40	47, 48
Y	CN 111126161 A (BEIJING UNION UNIVERSITY) 08 May 2020 (2020-05-08) description, paragraphs 4-31	47, 48
Y	CN 111079619 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 28 April 2020 (2020-04-28) description, paragraphs 3-25	47, 48
A	CN 111310667 A (BEIJING XIAOMA HUIXING TECHNOLOGY CO., LTD.) 19 June 2020 (2020-06-19) entire document	1-50
A	CN 109829447 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 31 May 2019 (2019-05-31) entire document	1-50



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

14 April 2021

Date of mailing of the international search report

08 May 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/103263

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2019088027 A1 (SHOPPAR INC.) 21 March 2019 (2019-03-21) entire document	1-50

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/103263

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110390258	A	29 October 2019	None			
CN	111126161	A	08 May 2020	None			
CN	111079619	A	28 April 2020	None			
CN	111310667	A	19 June 2020	None			
CN	109829447	A	31 May 2019	None			
US	2019088027	A1	21 March 2019	US	10762713	B2	01 September 2020

国际检索报告

国际申请号

PCT/CN2020/103263

A. 主题的分类

G06K 9/00 (2006.01) i

按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06K

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用))

CNABS; CNTXT; DWPI; VEN; USTXT; EPTXT; WOTXT; CNKI:框, 标注, 标记, 目标, 角点, 顶点, 二维, 三维, 朝向, frame, label, mark, target, corner point, vertex, two dimension, 2D, three dimension, 3D, orientation

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 110390258 A (东南大学) 2019年 10月 29日 (2019 - 10 - 29) 说明书第4-22、29-40段	1-46、49-50
X	CN 111126161 A (北京联合大学) 2020年 5月 8日 (2020 - 05 - 08) 说明书第4-31段	1-46、49-50
Y	CN 110390258 A (东南大学) 2019年 10月 29日 (2019 - 10 - 29) 说明书第4-22、29-40段	47-48
Y	CN 111126161 A (北京联合大学) 2020年 5月 8日 (2020 - 05 - 08) 说明书第4-31段	47-48
Y	CN 111079619 A (北京百度网讯科技有限公司) 2020年 4月 28日 (2020 - 04 - 28) 说明书第3-25段	47-48
A	CN 111310667 A (北京小马慧行科技有限公司) 2020年 6月 19日 (2020 - 06 - 19) 全文	1-50
A	CN 109829447 A (百度在线网络技术北京有限公司) 2019年 5月 31日 (2019 - 05 - 31) 全文	1-50
A	US 2019088027 A1 (SHOPPAR INC) 2019年 3月 21日 (2019 - 03 - 21) 全文	1-50

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2021年 4月 14日

国际检索报告邮寄日期

2021年 5月 8日

ISA/CN的名称和邮寄地址

中国国家知识产权局 (ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

房琦

电话号码 86-(20)-28950714

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/103263

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	110390258	A	2019年 10月 29日	无	
CN	111126161	A	2020年 5月 8日	无	
CN	111079619	A	2020年 4月 28日	无	
CN	111310667	A	2020年 6月 19日	无	
CN	109829447	A	2019年 5月 31日	无	
US	2019088027	A1	2019年 3月 21日	US 10762713 B2	2020年 9月 1日