



(12) 发明专利申请

(10) 申请公布号 CN 112991445 A

(43) 申请公布日 2021.06.18

(21) 申请号 202110237771.2

(22) 申请日 2021.03.03

(71) 申请人 网易(杭州)网络有限公司  
地址 310052 浙江省杭州市滨江区长河街  
道网商路599号4幢7层

(72) 发明人 徐一凡 袁焱 黄子镭 范长杰  
胡志鹏

(74) 专利代理机构 北京超凡宏宇专利代理事务  
所(特殊普通合伙) 11463  
代理人 钟扬飞

(51) Int. Cl.  
G06T 7/73 (2017.01)  
G06K 9/62 (2006.01)

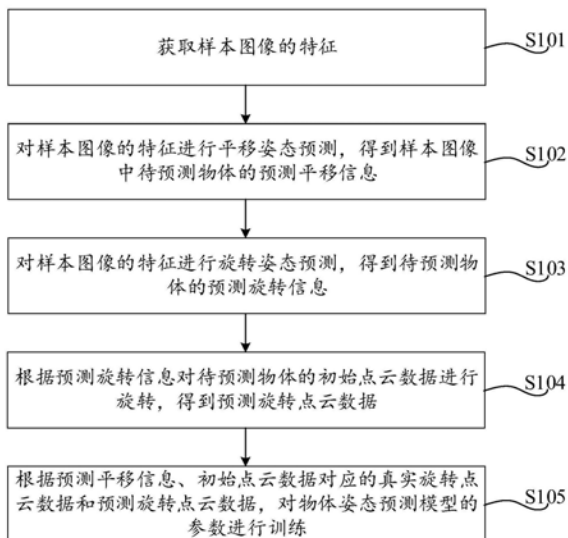
权利要求书2页 说明书10页 附图4页

(54) 发明名称

模型训练方法、姿态预测方法、装置、设备及  
存储介质

(57) 摘要

本申请提供一种模型训练方法、姿态预测方  
法、装置、设备及存储介质,涉及姿态预测领域。  
该方法包括:获取样本图像的特征;对所述样本  
图像的特征进行平移姿态预测,得到所述样本  
图像中待预测物体的预测平移信息;对所述样  
本图像的特征进行旋转姿态预测,得到所述待  
预测物体的预测旋转信息;根据所述预测旋转  
信息对所述待预测物体的初始点云数据进行旋  
转,得到预测旋转点云数据;根据所述预测平  
移信息、所述初始点云数据对应的真实旋转点  
云数据和预测旋转点云数据,对所述物体姿态  
预测模型的参数进行训练。相对于现有技术,避  
免了由于旋转和平移的预测存在耦合,导致模  
型训练的收敛速度慢的问题。



1. 一种物体姿态预测模型的训练方法,其特征在于,所述方法,包括:
  - 获取样本图像的特征;
  - 对所述样本图像的特征进行平移姿态预测,得到所述样本图像中待预测物体的预测平移信息;
  - 对所述样本图像的特征进行旋转姿态预测,得到所述待预测物体的预测旋转信息;
  - 根据所述预测旋转信息对所述待预测物体的初始点云数据进行旋转,得到预测旋转点云数据;
  - 根据所述预测平移信息、所述初始点云数据对应的真实旋转点云数据和预测旋转点云数据,对所述物体姿态预测模型的参数进行训练。
2. 如权利要求1所述的方法,其特征在于,所述根据所述初始点云数据对应的真实旋转点云数据和预测旋转点云数据,对所述物体姿态预测模型的参数进行训练,包括:
  - 根据所述真实旋转点云数据和所述预测旋转点云数据,计算第一损失函数值;
  - 根据所述第一损失函数值,对所述物体姿态预测模型的参数进行训练,直至达到预设迭代停止条件。
3. 如权利要求2所述的方法,其特征在于,所述预测旋转信息为预测旋转矩阵,所述预测平移信息为预测平移矩阵;所述根据所述预测旋转信息对所述待预测物体的初始点云数据进行旋转,得到预测旋转点云数据,包括:
  - 根据所述预测旋转矩阵对所述初始点云数据进行矩阵乘法,得到所述预测旋转点云数据;
  - 所述根据所述预测平移信息、所述真实旋转点云数据和所述预测旋转点云数据,计算第一损失函数值,包括:
    - 根据所述预测旋转点云数据、所述真实旋转点云数据、所述预测平移矩阵和所述待预测物体的平移真值,计算所述第一损失函数。
4. 如权利要求2所述的方法,其特征在于,所述根据所述预测平移信息、所述初始点云数据和旋转后的点云数据,计算第一损失函数值之前,所述方法还包括:
  - 根据所述待预测物体的旋转真值对所述初始点云数据进行矩阵乘法,得到所述真实旋转点云数据。
5. 如权利要求2所述的方法,其特征在于,所述根据所述第一损失函数值,对所述物体姿态预测模型的参数进行训练之前,所述方法还包括:
  - 对所述样本图像的特征进行尺寸预测,得到所述待预测物体的预测尺寸;
  - 根据所述预测尺寸和所述待预测物体的尺寸真值,计算第二损失函数值;
  - 所述根据所述第一损失函数值,对所述物体姿态预测模型的参数进行训练,包括:
    - 根据所述第一损失函数值和所述第二损失函数值,对所述物体姿态预测模型的参数进行训练。
6. 如权利要求5所述的方法,其特征在于,所述根据所述第一损失函数值和所述第二损失函数值,对所述物体姿态预测模型的参数进行训练,包括:
  - 根据所述第一损失函数值、所述第二损失函数值和预设尺寸权重,计算目标损失函数值;
  - 根据所述目标损失函数值,对所述物体姿态预测模型的参数进行训练。

7. 一种物体姿态预测方法,其特征在于,所述方法包括:

获取待预测图像的特征;其中,所述待预测图像中包括待预测物体;

根据预设物体姿态预测模型对所述待预测图像的特征进行姿态预测,得到所述待预测图像中待预测物体的预测姿态信息;所述预测姿态信息包括:预测平移信息和预测旋转信息;其中,所述物体姿态预测模型为采用上述权利要求1-6中任一所述的方法训练得到的模型。

8. 一种物体姿态预测模型的训练装置,其特征在于,所述装置包括:获取模块、预测模块、旋转模块和训练模块,其中:

所述获取模块,用于获取样本图像的特征;

所述预测模块,用于对所述样本图像的特征进行平移姿态预测,得到所述样本图像中待预测物体的预测平移信息;对所述样本图像的特征进行旋转姿态预测,得到所述待预测物体的预测旋转信息;

所述旋转模块,用于根据所述预测旋转信息对所述待预测物体的初始点云数据进行旋转,得到预测旋转点云数据;

所述训练模块,用于根据所述预测平移信息、所述初始点云数据对应的真实旋转点云数据和预测旋转点云数据,对所述物体姿态预测模型的参数进行训练。

9. 一种物体姿态预测装置,其特征在于,所述装置包括:获取模块和预测模块,其中:

所述获取模块,用于获取待预测图像的特征;其中,所述待预测图像中包括待预测物体;

所述预测模块,用于根据预设物体姿态预测模型对所述待预测图像的特征进行姿态预测,得到所述待预测图像中待预测物体的预测姿态信息;所述预测姿态信息包括:预测平移信息和预测旋转信息。

10. 一种电子设备,其特征在于,所述设备包括:处理器、存储介质和总线,所述存储介质存储有所述处理器可执行的机器可读指令,当所述电子设备运行时,所述处理器与所述存储介质之间通过总线通信,所述处理器执行所述机器可读指令,以执行上述权利要求1-6或7任一项所述的方法。

11. 一种存储介质,其特征在于,所述存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行上述权利要求1-6或7任一项所述的方法。

## 模型训练方法、姿态预测方法、装置、设备及存储介质

### 技术领域

[0001] 本申请涉及姿态预测技术领域，具体而言，涉及一种模型训练方法、姿态预测方法、装置、设备及存储介质。

### 背景技术

[0002] 物体在空间具有六个自由度，即沿x、y、z三个直角坐标轴方向的移动自由度和绕这三个坐标轴的转动自由度。要完全确定物体的位置，就必须清楚这六个自由度。因此如何基于单张图像对图像中的物体进行六自由度的估计变得尤为重要。

[0003] 现有技术中对于六自由度姿态估计一般可以通过卷积神经网络预测旋转矩阵和平移矩阵，随后利用预测的旋转平移矩阵对物体点云数据进行转置后得到的转置后的点云数据，与通过旋转平移矩阵真值进行转置后得到点云数据计算均方误差作为损失函数。

[0004] 但是这样的计算方式由于旋转矩阵和平移矩阵在计算过程中存在耦合，导致在平移矩阵的预测还不准确的时候，对旋转矩阵的预测就不够准确，反之在旋转矩阵的预测还不准确的时候，对平移矩阵的预测就不够准确，从而会明显影响姿态估计模型的收敛速度。

### 发明内容

[0005] 本申请的目的在于，针对上述现有技术中的不足，提供一种模型训练方法、姿态预测方法、装置、设备及存储介质，以解决现有技术中由于旋转和平移的预测存在耦合，导致模型训练的收敛速度慢的问题。

[0006] 为实现上述目的，本申请实施例采用的技术方案如下：

[0007] 第一方面，本申请一实施例提供了一种物体姿态预测模型的训练方法，所述方法，包括：

[0008] 获取样本图像的特征；

[0009] 对所述样本图像的特征进行平移姿态预测，得到所述样本图像中待预测物体的预测平移信息；

[0010] 对所述样本图像的特征进行旋转姿态预测，得到所述待预测物体的预测旋转信息；

[0011] 根据所述预测旋转信息对所述待预测物体的初始点云数据进行旋转，得到预测旋转点云数据；

[0012] 根据所述预测平移信息、所述初始点云数据对应的真实旋转点云数据和预测旋转点云数据，对所述物体姿态预测模型的参数进行训练。

[0013] 可选地，所述根据所述初始点云数据对应的真实旋转点云数据和预测旋转点云数据，对所述物体姿态预测模型的参数进行训练，包括：

[0014] 根据所述真实旋转点云数据和所述预测旋转点云数据，计算第一损失函数值；

[0015] 根据所述第一损失函数值，对所述物体姿态预测模型的参数进行训练，直至达到预设迭代停止条件。

[0016] 可选地,所述预测旋转信息为预测旋转矩阵,所述预测平移信息为预测平移矩阵;所述根据所述预测旋转信息对所述待预测物体的初始点云数据进行旋转,得到预测旋转点云数据,包括:

[0017] 根据所述预测旋转矩阵对所述初始点云数据进行矩阵乘法,得到所述预测旋转点云数据;

[0018] 所述根据所述预测平移信息、所述真实旋转点云数据和所述预测旋转点云数据,计算第一损失函数值,包括:

[0019] 根据所述预测旋转点云数据、所述真实旋转点云数据、所述预测平移矩阵和所述待预测物体的平移真值,计算所述第一损失函数。

[0020] 可选地,所述根据所述预测平移信息、所述初始点云数据和旋转后的点云数据,计算第一损失函数值之前,所述方法还包括:

[0021] 根据所述待预测物体的旋转真值对所述初始点云数据进行矩阵乘法,得到所述真实旋转点云数据。

[0022] 可选地,所述根据所述第一损失函数值,对所述物体姿态预测模型的参数进行训练之前,所述方法还包括:

[0023] 对所述样本图像的特征进行尺寸预测,得到所述待预测物体的预测尺寸;

[0024] 根据所述预测尺寸和所述待预测物体的尺寸真值,计算第二损失函数值;

[0025] 所述根据所述第一损失函数值,对所述物体姿态预测模型的参数进行训练,包括:

[0026] 根据所述第一损失函数值和所述第二损失函数值,对所述物体姿态预测模型的参数进行训练。

[0027] 可选地,所述根据所述第一损失函数值和所述第二损失函数值,对所述物体姿态预测模型参数进行训练,包括:

[0028] 根据所述第一损失函数值、所述第二损失函数值和预设尺寸权重,计算目标损失函数值;

[0029] 根据所述目标损失函数值,对所述物体姿态预测模型的参数进行训练。

[0030] 第二方面,本申请另一实施例还提供了一种物体姿态预测方法,所述方法包括:

[0031] 获取待预测图像的特征;其中,所述待预测图像中包括待预测物体;

[0032] 根据预设物体姿态预测模型对所述待预测图像的特征进行姿态预测,得到所述待预测图像中待预测物体的预测姿态信息;所述预测姿态信息包括:预测平移信息和预测旋转信息;其中,所述物体姿态预测模型为采用上述第一方面中任一所述的方法训练得到的模型。

[0033] 第三方面,本申请另一实施例提供了一种物体姿态预测模型的训练装置,所述装置包括:获取模块、预测模块、旋转模块和训练模块,其中:

[0034] 所述获取模块,用于获取样本图像的特征;

[0035] 所述预测模块,用于对所述样本图像的特征进行平移姿态预测,得到所述样本图像中待预测物体的预测平移信息;对所述样本图像的特征进行旋转姿态预测,得到所述待预测物体的预测旋转信息;

[0036] 所述旋转模块,用于根据所述预测旋转信息对所述待预测物体的初始点云数据进行旋转,得到预测旋转点云数据;

[0037] 所述训练模块,用于根据所述预测平移信息、所述初始点云数据对应的真实旋转点云数据和预测旋转点云数据,对所述物体姿态预测模型的参数进行训练。

[0038] 可选地,所述装置还包括:计算模块,用于根据所述真实旋转点云数据和所述预测旋转点云数据,计算第一损失函数值;

[0039] 所述训练模块,具体用于根据所述第一损失函数值,对所述物体姿态预测模型的参数进行训练,直至达到预设迭代停止条件。

[0040] 可选地,所述预测旋转信息为预测旋转矩阵,所述预测平移信息为预测平移矩阵;所述计算模块,具体用于根据所述预测旋转矩阵对所述初始点云数据进行矩阵乘法,得到所述预测旋转点云数据;根据所述预测旋转点云数据、所述真实旋转点云数据、所述预测平移矩阵和所述待预测物体的平移真值,计算所述第一损失函数。

[0041] 可选地,所述计算模块,具体用于根据所述待预测物体的旋转真值对所述初始点云数据进行矩阵乘法,得到所述真实旋转点云数据。

[0042] 可选地,所述预测模块,具体用于对所述样本图像的特征进行尺寸预测,得到所述待预测物体的预测尺寸;

[0043] 所述计算模块,具体用于根据所述预测尺寸和所述待预测物体的尺寸真值,计算第二损失函数值;

[0044] 所述训练模块,具体用于根据所述第一损失函数值和所述第二损失函数值,对所述物体姿态预测模型的参数进行训练。

[0045] 可选地,所述计算模块,具体用于根据所述第一损失函数值、所述第二损失函数值和预设尺寸权重,计算目标损失函数值;

[0046] 所述训练模块,具体用于根据所述目标损失函数值,对所述物体姿态预测模型的参数进行训练。

[0047] 第四方面,本申请另一实施例提供了一种物体姿态预测装置,所述装置包括:获取模块和预测模块,其中:

[0048] 所述获取模块,用于获取待预测图像的特征;其中,所述待预测图像中包括待预测物体;

[0049] 所述预测模块,用于根据预设物体姿态预测模型对所述待预测图像的特征进行姿态预测,得到所述待预测图像中待预测物体的预测姿态信息;所述预测姿态信息包括:预测平移信息和预测旋转信息;其中,所述物体姿态预测模型为采用上述第一方面中任一所述的方法训练得到的模型。

[0050] 第五方面,本申请另一实施例提供了一种电子设备,包括:处理器、存储介质和总线,所述存储介质存储有所述处理器可执行的机器可读指令,当所述电子设备运行时,所述处理器与所述存储介质之间通过总线通信,所述处理器执行所述机器可读指令,以执行如上述第一方面或第二方面任一所述方法的步骤。

[0051] 第六方面,本申请另一实施例提供了一种存储介质,所述存储介质上存储有计算机程序,所述计算机程序被处理器运行时执行如上述第一方面或第二方面任一所述方法的步骤。

[0052] 本申请的有益效果是:采用本申请提供的物体姿态预测模型的训练方法,在获取到样本图像的特征后,分别对预测平移信息和预测旋转信息进行预测,由于对预测平移信

息和预测旋转信息的计算是解耦的,是分别进行预测的,即在预测平移信息时假定旋转信息是正确的,并得到预测平移信息;预测旋转信息时假定平移信息是正确的,并得到预测旋转信息;因此提高了模型训练过程中参数的收敛速度,加速了模型的训练,提高了模型训练的效率。

### 附图说明

[0053] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,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。

[0054] 图1为本申请一实施例提供的物体姿态预测模型的训练方法的流程示意图;

[0055] 图2为本申请另一实施例提供的物体姿态预测模型的训练方法的流程示意图;

[0056] 图3为本申请另一实施例提供的物体姿态预测模型的训练方法的流程示意图;

[0057] 图4为本申请另一实施例提供的物体姿态预测模型的训练方法的流程示意图;

[0058] 图5为本申请一实施例提供的物体姿态预测方法的流程示意图;

[0059] 图6为本申请另一实施例提供的物体姿态预测方法的流程示意图;

[0060] 图7为本申请一实施例提供的物体姿态预测模型的训练装置的结构示意图;

[0061] 图8为本申请另一实施例提供的物体姿态预测模型的训练装置的结构示意图;

[0062] 图9为本申请一实施例提供的物体姿态预测装置的结构示意图;

[0063] 图10为本申请一实施例提供的电子设备的结构示意图。

### 具体实施方式

[0064] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本申请一部分实施例,而不是全部的实施例。

[0065] 通常在此处附图中描述和示出的本申请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申请的实施例的详细描述并非旨在限制要求保护的本申请的范围,而是仅仅表示本申请的选定实施例。基于本申请的实施例,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,都属于本申请保护的范畴。

[0066] 另外,本申请中使用的流程图示出了根据本申请的一些实施例实现的操作。应该理解,流程图的操作可以不按顺序实现,没有逻辑的上下文关系的步骤可以反转顺序或者同时实施。此外,本领域技术人员在本申请内容的指引下,可以向流程图添加一个或多个其他操作,也可以从流程图中移除一个或多个操作。

[0067] 本申请实施例所提供的物体姿态预测模型的训练方法,可应用在对物体姿态预测之前,可对物体姿态预测模型进行训练,以使得后续可以根据训练得到的物体姿态预测模型直接对图像中的待预测物体的姿态信息进行预测。应用场景例如可以为机器人抓取或自动驾驶感知等任何需要对图像中的待预测物体的姿态进行预测的场景中,应当理解本申请提供的方法的具体应用场景可以根据用户需要灵活调整,并不以上述实施例给出的为限。

[0068] 如下结合多个具体的应用示例,对本申请实施例所提供的一种物体姿态预测模型的训练方法进行解释说明。图1为本申请一实施例提供的一种物体姿态预测模型的训练方法的流程示意图,如图1所示,该方法包括:

[0069] S101:获取样本图像的特征。

[0070] 其中,样本图像例如可以为单张单视角的真彩(RGB)图像,每个样本图像中仅包括一个待预测物体,若存在包括多个待预测物体的图像,则需要对该图像进行切分,将切分得到多个只包括一个待预测物体的图像作为多个样本图像;样本图像的获取方式例如可以为从网络爬取的,或从数据库中获取的,或拍摄的,具体样本图像的获取方式并不以上述实施例给出的为限,可以根据用户需要灵活调整。

[0071] 示例地,在本申请的一个实施例中,物体姿态预测模型中例如可以包括:特征提取模块,则可以根据特征提取模块,对样本图像进行处理,得到样本图像的特征。

[0072] 在一些可能的实施例中,特征提取模块例如可以为卷积神经网络模块,在本申请的一个实施例中,卷积神经网络模块例如可以采用卷积神经网络 vgg16模块,将样本图像输入卷积神经网络后,通过卷积神经网络模块进行卷积和最大值池化,输出得到样本图像对应的1024维特征,并将输出的样本图像的特征命名为F<sub>cnn</sub>。

[0073] S102:对样本图像的特征进行平移姿态预测,得到样本图像中待预测物体的预测平移信息。

[0074] 在本申请的一个实施例中,物体姿态预测模型中例如可以包括:第一预测模块,第一预测模块例如可以为多层感知器MLP1,在本申请的实施例中,多层感知器即为全连接层,即可以采用第一预测模块对样本图像的特征进行平移姿态的预测,得到样本图像中待预测物体的预测平移信息;在本申请的一个实施例中,预测平移信息例如可以为预测平移距离信息,多层感知器MLP1的神经元数目例如可以分别为512,256,3,MLP1预测得到的预测平移信息例如可以由3维特征( $t_x, t_y, t_z$ )构成的,随后根据上述3维特征确定预测平移信息对应的平移矩阵,平移矩阵记为 $T_p$ 。

[0075] S103:对样本图像的特征进行旋转姿态预测,得到待预测物体的预测旋转信息。

[0076] 在本申请的一个实施例中,物体姿态预测模型中例如可以包括:第二预测模块,第二预测模块例如可以为多层感知器MLP2,在本申请的实施例中,多层感知器即为全连接层,即可以采用第二预测模块对样本图像的特征进行旋转姿态预测,得到样本图像中待预测物体的预测旋转信息;在本申请的一个实施例中,预测旋转信息例如可以为预测旋转矩阵;MLP2的神经元数目例如可以分别为512,256,6,MLP2预测得到的预测旋转信息为6维特征,分别用于预测待预测物体坐标系旋转后的x轴和y轴向量,然后根据预测得到的x轴向量和y轴向量计算出z轴的向量,最终得到包括x轴向量、y轴向量和z轴向量的预测旋转信息,进而根据预测旋转信息得到旋转矩阵 $R_p$ 。

[0077] S104:根据预测旋转信息对待预测物体的初始点云数据进行旋转,得到预测旋转点云数据。

[0078] 其中,初始点云数据为通过渲染得到的待检测物体对应的三维模型后,获取待检测物体的三维模型中在待检测物体对应的初始状态下,即未旋转状态下的点云数据。

[0079] S105:根据预测平移信息、初始点云数据对应的真实旋转点云数据和预测旋转点云数据,对物体姿态预测模型的参数进行训练。



[0080] 在一些可能的实施例中,真实旋转点云数据例如可以为通过渲染得到的待检测物体对应的三维模型后,获取到的待检测物体旋转后对应的真实旋转点云数据。

[0081] 采用本申请提供的物体姿态预测模型的训练方法,在获取到样本图像的特征后,分别对预测平移信息和预测旋转信息进行预测,由于对预测平移信息和预测旋转信息的计算是解耦的,是分别进行预测的,即在预测平移信息时假定旋转信息是正确的,并得到假定旋转信息正确的情况下的预测平移信息;预测旋转信息时假定平移信息是正确的,并得到假定平移信息是正确情况下的预测旋转信息;因此提高了模型训练过程中参数的收敛速度,加速了模型的训练,提高了模型训练的效率。

[0082] 可选地,在上述实施例的基础上,本申请实施例还可提供一种物体姿态预测模型的训练方法,如下结合附图对上述方法中对物体姿态预测模型的参数进行训练的实现过程进行示例说明。图2为本申请另一实施例提供的一种物体姿态预测模型的训练方法的流程示意图,如图2所示,S105可包括:

[0083] S106:根据真实旋转点云数据和预测旋转点云数据,计算第一损失函数值。

[0084] 示例地,在本申请的一个实施例中,预测旋转信息可以为预测旋转矩阵 $R_p$ ,预测平移信息可以为预测平移矩阵 $T_p$ ;则预测旋转点云数据的获取方式例如可以为:根据预测旋转矩阵 $R_p$ 对初始点云数据 $P$ 进行矩阵乘法,得到预测旋转点云数据 $P_{gt}$ ;对应的,第一损失函数值的确定方式例如可以为根据预测旋转点云数据 $P_{gt}$ 、真实旋转点云数据 $P_{pred}$ 、预测平移矩阵 $T_p$ 和待预测物体的平移真值 $T_{gt}$ ,计算第一损失函数。

[0085] 在一些可能的实施例中,真实旋转点云数据的获取方式例如可以为根据待预测物体的旋转真值对初始点云数据进行矩阵乘法,得到真实旋转点云数据,例如可以为通过 $R_{gt} * P + T_{gt}$ 计算得到的,其中 $R_{gt}$ 为旋转真值, $P$ 为初始点云数据, $T_{gt}$ 为平移真值。

[0086] 其中,旋转真值和翻译真值例如均可以为通过渲染得到的待检测物体对应的三维模型,并获取待检测物体对应的三维模型中的旋转真值和翻译真值,旋转真值例如可以记为 $R_{gt}$ ,翻译真值例如可以记为 $T_{gt}$ 。

[0087] 示例地,在本申请的一个实施例中,第一损失函数 $L_{RT}$ 中包括两部分:假定平移矩阵预测准确,即假定平移矩阵为翻译真值 $T_{gt}$ ,则确定通过翻译真值 $T_{gt}$ 和旋转矩阵对初始点云数据 $P$ 进行旋转后得到的点云数据,和真实旋转点云数据之间的均方误差作为第一损失函数的一部分;另一部分是假定旋转矩阵预测准确,即假定旋转矩阵为旋转真值 $R_{gt}$ ,则确定通过旋转真值 $R_{gt}$ 和平移矩阵对初始点云数据进行旋转后得到的点云数据,和真实旋转点云数据之间的均方误差作为第一损失函数的另一部分。因此第一损失函数 $L_{RT}$ 的计算方式例如可以为:

$$[0088] \quad L_{RT} = \text{MSE}(R_p * P + T_{gt}, R_{gt} * P + T_{gt}) + \text{MSE}(R_{gt} * P + T_p, R_{gt} * P + T_{gt})$$

$$[0089] \quad = \text{MSE}(R_p * P, R_{gt} * P) + \text{MSE}(T_p, T_{gt})$$

$$[0090] \quad = \text{MSE}(P_{pred}, P_{gt}) + \text{MSE}(T_p, T_{gt})$$

[0091] 其中,MSE为均方误差, $R_p$ 为旋转矩阵, $R_{gt}$ 为旋转真值, $P$ 为初始点云数据, $P_{gt}$ 为真实旋转点云数据, $P_{pred}$ 为预测旋转点云数据, $T_p$ 为平移矩阵, $T_{gt}$ 为翻译真值;即可以为通过计算预测旋转点云数据 $P_{pred}$ 与真实旋转点云数据 $P_{gt}$ 之间的均方误差,和预测平移矩阵 $T_p$ 和待预测物体的翻译真值 $T_{gt}$ 之间的均方误差,得到第一损失函数 $L_{RT}$ 。

[0092] S107:根据第一损失函数值,对物体姿态预测模型的参数进行训练,直至达到预设

迭代停止条件。

[0093] 可选地,在上述实施例的基础上,本申请实施例还可提供一种物体姿态预测模型的训练方法,如下结合附图对上述方法的实现过程进行示例说明。图3为本申请另一实施例提供的一种物体姿态预测模型的训练方法的流程示意图,如图3所示,S107之前,该方法还可包括:

[0094] S108:对样本图像的特征进行尺寸预测,得到待预测物体的预测尺寸。

[0095] 在本申请的一个实施例中,物体姿态预测模型中例如可以包括:第三预测模块,第三预测模块例如可以为多层感知器MLP3,在本申请的实施例中,多层感知器即为全连接层,即可以采用第三预测模块对待预测物体进行尺寸预测,得到待预测物体的预测尺寸;在本申请的一个实施例中,预测尺寸例如可以为预测长、预测宽和预测高;MLP3的神经元数目例如可以分别为512,256,3,通过MLP3输出的预测尺寸为3维特征分别为待预测物体的长、宽、高,预测尺寸记为 $LWH_{pred}$ 。

[0096] S109:根据预测尺寸和待预测物体的尺寸真值,计算第二损失函数值。

[0097] 在本申请的一个实施例中,待预测物体的尺寸真值可以记为 $LWH_{gt}$ ,也为通过渲染后的待检测物体对应的三维模型获取的,第二损失函数 $L_{LWH}$ 的计算例如可以为通过计算尺寸真值和预测尺寸的均方误差得到的,即通过公式 $L_{LWH} = MSE(LWH_{pred}, LWH_{gt})$ 计算得到的;其中,MSE为均方误差。

[0098] 对应的S107可包括:

[0099] S110:根据第一损失函数值和第二损失函数值,对物体姿态预测模型的参数进行训练。

[0100] 可选地,在本申请的一个实施例中,可以根据第一损失函数值、第二损失函数值和预设尺寸权重,计算目标损失函数值;根据目标损失函数值,对物体姿态预测模型的参数进行训练。

[0101] 由于第二损失函数值用于对第一损失函数值的监督,所以第二损失函数值的预设尺寸权重一般设置较小,在本申请的一个实施例中,预设尺寸权重例如可以设置为0.1,则目标损失函数值 $L_{loss}$ 的计算方式例如可以为: $L_{loss} = L_{RT} + 0.1 * L_{LWH}$ ,即第一损失函数值和与预设尺寸权重相乘后的第二损失函数值之间累加,得到的结果为目标损失函数;应当理解上述预设尺寸权重仅为示例性说明,具体预设尺寸权重的设置可以根据用户需要灵活调整,并不以上述实施例给出的为限制,只需不影响目标损失函数值的计算,仅起到监督作用即可。

[0102] 现有的技术方案在预测物体姿态时没有显性得考虑物体本身的尺寸。比如对于沙发来说,该物体的横向长度更大,对于立灯来说,该物体的纵向长度更大。如果不考虑物体本身的尺寸,直接将沙发和立灯均看作圆柱体,则两者的初始位置是相差90度的旋转的,可见这种不考虑物体本身尺寸的姿态估计就会对姿态估计带来歧义。

[0103] 因此采用本申请提供的物体姿态预测模型的训练方法,对于目标损失函数值的计算引入了根据预测尺寸和待预测物体的尺寸真值确定的第二损失函数值的监督,这种通过第一损失函数值和第二损失函数值共同确定目标损失函数值的方式,由于在引入了物体的尺寸的监督,因此在姿态预测时考虑了各物体的尺寸,从而得到的姿态预测结果更加准确。

[0104] 在本申请的一个实施例中,例如可以通过目标损失函数对物体姿态预测模型的参

数进行调整,直至目标损失函数收敛,则训练结束;对物体姿态预测模型的参数进行调整,即为分别对第一预测模块和第二预测模块的参数进行调整。

[0105] 图4为本申请另一实施例提供的一种物体姿态预测模型的训练方法的流程示意图,以物体姿态预测模型包括:卷积神经网络模块、第一预测模块、第二预测模块和第三预测模块为例进行说明,如图4所示,物体姿态预测模型的训练方法的完整流程例如可以为:

[0106] 输入样本图像和初始点云数据;卷积神经网络根据输入的图像,提取样本图像的特征;将样本图像的特征分别输入第一预测模块、第二预测模块和第三预测模块,并通过第一预测模块、第二预测模块和第三预测模块分别预测待预测物体的旋转矩阵、平移矩阵和预测尺寸;分别通过预测的旋转矩阵以及旋转矩阵真值对初始位置物体点云进行旋转,获取预测旋转点云数据;随后计算目标损失函数值,并通过梯度下降对物体姿态预测模型进行训练,优化物体姿态预测模型中的参数,随后重复上述步骤,直至目标损失函数值收敛。

[0107] 采用本申请提供的物体姿态预测模型的训练方法,不但解耦了旋转矩阵和平移矩阵的预测,并且引入了尺寸监督,在大幅度提高模型训练效率的情况下,提高了训练得到的物体姿态预测模型的预测精度。

[0108] 可选地,在上述实施例的基础上,本申请实施例还可提供一种物体姿态预测方法,如下结合附图对上述方法的实现过程进行示例说明。图5为本申请一实施例提供的一种物体姿态预测方法的流程示意图,如图5所示,该方法可包括:

[0109] S201:获取待预测图像的特征。

[0110] 其中,待预测图像中包括待预测物体。

[0111] S202:根据预设物体姿态预测模型对待预测图像的特征进行姿态预测,得到待预测图像中待预测物体的预测姿态信息。

[0112] 预测姿态信息例如可以包括:预测平移信息和预测旋转信息;例如可以为六自由度信息,即包括沿x、y、z三个直角坐标轴方向的移动自由度和绕x、y、z三个直角坐标轴的转动自由度。

[0113] 其中,物体姿态预测模型为采用上述图1-图4中任一所述的方法训练得到的模型,由于该物体姿态预测模型为根据尺寸信息监督训练得到的,因此在预测各待预测物体的预测姿态信息时,会考虑各待预测物体的尺寸信息。这种通过尺寸信息监督训练得到的预设物体姿态预测模型,由于在预测过程中考虑了各待预测物体的迟勋信息,所以得到的各待预测物体的姿态预测的准确性更强,从而得到的预测姿态信息更加准确。

[0114] 图6为本申请另一实施例提供的物体姿态预测方法的流程示意图,仍以物体姿态预测模型包括:卷积神经网络模块、第一预测模块、第二预测模块为例进行说明,如图6所示,该方法的完整流程可以为:获取待预测图像后,卷积神经网络根据获取的待预测图像,提取待预测图像的特征,并将提取的特征输入至第一预测模块和第二预测模块,分别预测待预测物体的预测平移信息和预测旋转信息,并确定预测平移信息和预测旋转信息为待预测物体的预测姿态信息。

[0115] 由于上述预设物体姿态训练模型是采用上述图1-图4提供的物体姿态预测模型的训练方法训练得到的,因此采用上述预设物体姿态训练模型进行物体姿态预测的方法带来的有益效果,与上述图1-图4带来的有益效果相同,本申请在此不再赘述。

[0116] 下述结合附图对本申请所提供的物体姿态预测模型的训练装置进行解释说明,该

物体姿态预测模型的训练装置可执行上述图1-图4任一物体姿态预测模型的训练方法,其具体实现以及有益效果参照上述,如下不再赘述。

[0117] 图7为本申请一实施例提供的物体姿态预测模型的训练装置的结构示意图,如图7所示,该装置包括:获取模块301、预测模块302、旋转模块 303和训练模块304,其中:

[0118] 获取模块301,用于获取样本图像的特征;

[0119] 预测模块302,用于对样本图像的特征进行平移姿态预测,得到样本图像中待预测物体的预测平移信息;对样本图像的特征进行旋转姿态预测,得到待预测物体的预测旋转信息;

[0120] 旋转模块303,用于根据预测旋转信息对待预测物体的初始点云数据进行旋转,得到预测旋转点云数据;

[0121] 训练模块304,用于根据预测平移信息、初始点云数据对应的真实旋转点云数据和预测旋转点云数据,对物体姿态预测模型的参数进行训练。

[0122] 图8为本申请另一实施例提供的物体姿态预测模型的训练装置的结构示意图,如图8所示,该装置还包括:计算模块305,用于根据真实旋转点云数据和预测旋转点云数据,计算第一损失函数值。

[0123] 训练模块306,具体用于根据第一损失函数值,对物体姿态预测模型的参数进行训练,直至达到预设迭代停止条件。

[0124] 可选地,预测旋转信息为预测旋转矩阵,预测平移信息为预测平移矩阵;计算模块305,具体用于根据预测旋转矩阵对初始点云数据进行矩阵乘法,得到预测旋转点云数据;根据预测旋转点云数据、真实旋转点云数据、预测平移矩阵和待预测物体的平移真值,计算第一损失函数。

[0125] 可选地,计算模块305,具体用于根据待预测物体的旋转真值对初始点云数据进行矩阵乘法,得到真实旋转点云数据。

[0126] 可选地,预测模块302,具体用于对样本图像的特征进行尺寸预测,得到待预测物体的预测尺寸;

[0127] 计算模块305,具体用于根据预测尺寸和待预测物体的尺寸真值,计算第二损失函数值;

[0128] 训练模块304,具体用于根据第一损失函数值和第二损失函数值,对物体姿态预测模型的参数进行训练。

[0129] 可选地,计算模块305,具体用于根据第一损失函数值、第二损失函数值和预设尺寸权重,计算目标损失函数值;

[0130] 训练模块304,具体用于根据目标损失函数值,对物体姿态预测模型的参数进行训练。

[0131] 下述结合附图对本申请所提供的物体姿态预测装置进行解释说明,该物体姿态预测装置可执行上述图5-图6提供的物体姿态预测方法,其具体实现以及有益效果参照上述,如下不再赘述。

[0132] 图9为本申请一实施例提供的物体姿态预测装置的结构示意图,如图9 所示,该装置包括:获取模块401和预测模块402,其中:

[0133] 获取模块401,用于获取待预测图像的特征;其中,待预测图像中包括待预测物体;

[0134] 预测模块402,用于根据预设物体姿态预测模型对待预测图像的特征进行姿态预测,得到待预测图像中待预测物体的预测姿态信息;预测姿态信息包括:预测平移信息和预测旋转信息;其中,物体姿态预测模型为采用上述图1-图4中任一所述的方法训练得到的模型。

[0135] 上述装置用于执行前述实施例提供的方法,其实现原理和技术效果类似,在此不再赘述。

[0136] 以上这些模块可以是配置成实施以上方法的一个或多个集成电路,例如:一个或多个特定集成电路(Application Specific Integrated Circuit,简称ASIC),或,一个或多个微处理器(digital signal processor,简称DSP),或,一个或者多个现场可编程门阵列(Field Programmable Gate Array,简称FPGA)等。再如,当以上某个模块通过处理元件调度程序代码的形式实现时,该处理元件可以是通用处理器,例如中央处理器(Central Processing Unit,简称CPU)或其它可以调用程序代码的处理器。再如,这些模块可以集成在一起,以片上系统(system-on-a-chip,简称SOC)的形式实现。

[0137] 图10为本申请一实施例提供的电子设备的结构示意图,该电子设备可以集成于终端设备或者终端设备的芯片。

[0138] 该电子设备包括:处理器501、存储介质502和总线503。

[0139] 处理器501用于存储程序,处理器501调用存储介质502存储的程序,以执行上述图1-图4或图5-图6对应的方法实施例。具体实现方式和技术效果类似,这里不再赘述。

[0140] 可选地,本申请还提供一种程序产品,例如存储介质,该存储介质上存储有计算机程序,包括程序,该程序在被处理器运行时执行上述方法对应的实施例。

[0141] 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。

[0142] 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。

[0143] 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。

[0144] 上述以软件功能单元的形式实现的集成的单元,可以存储在一个计算机可读存储介质中。上述软件功能单元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)或处理器(英文:processor)执行本申请各个实施例所述方法的部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(英文:Read-Only Memory,简称:ROM)、随机存取存储器(英文:Random Access Memory,简称:RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

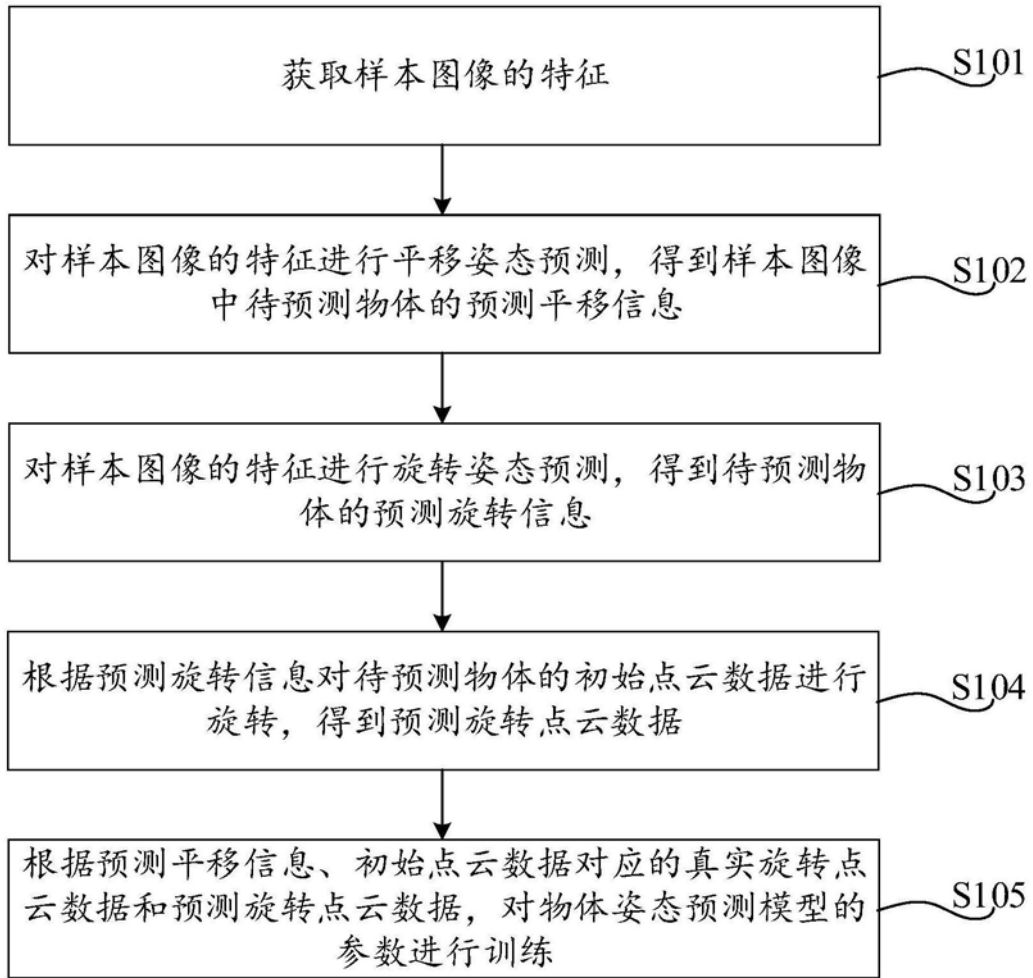


图1

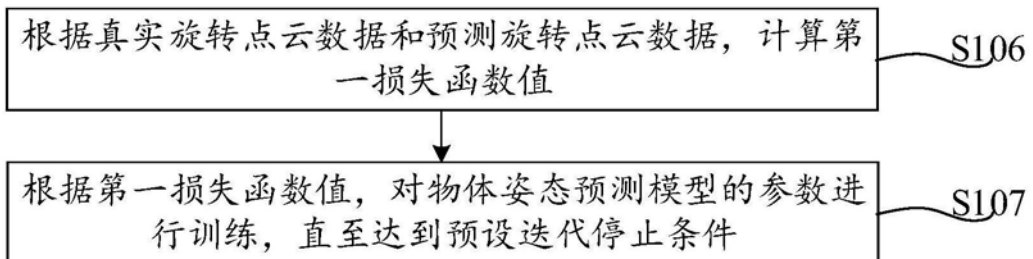


图2

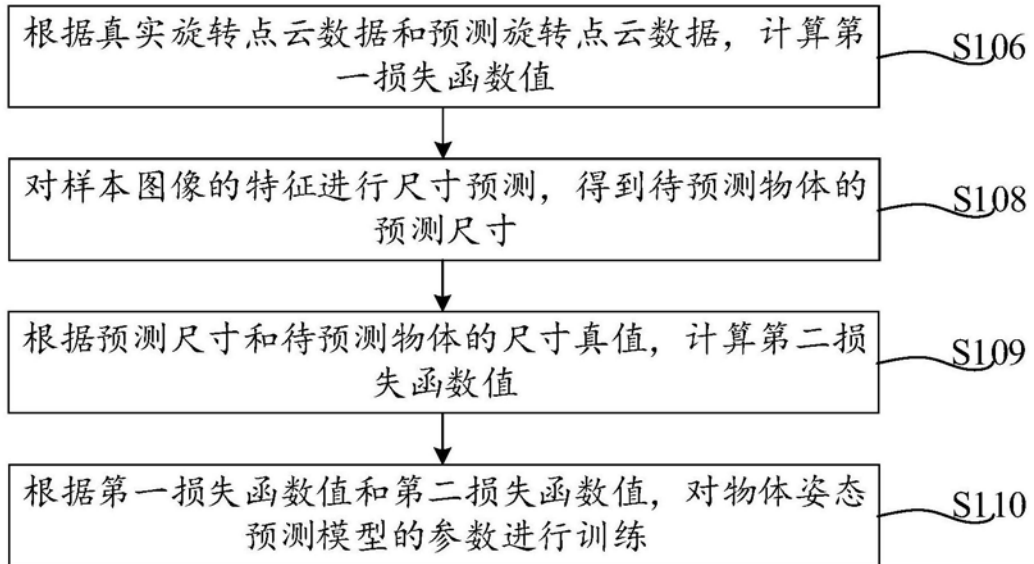


图3

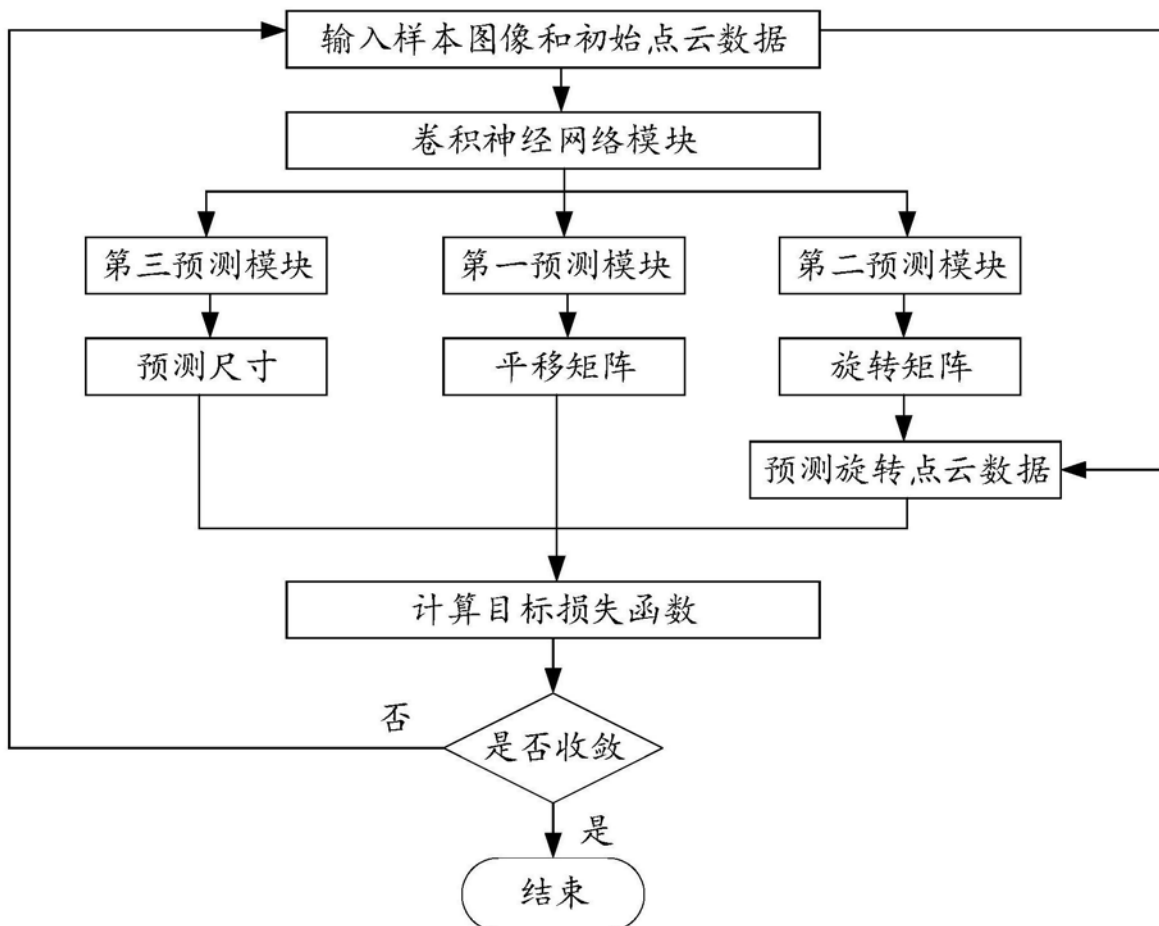


图4

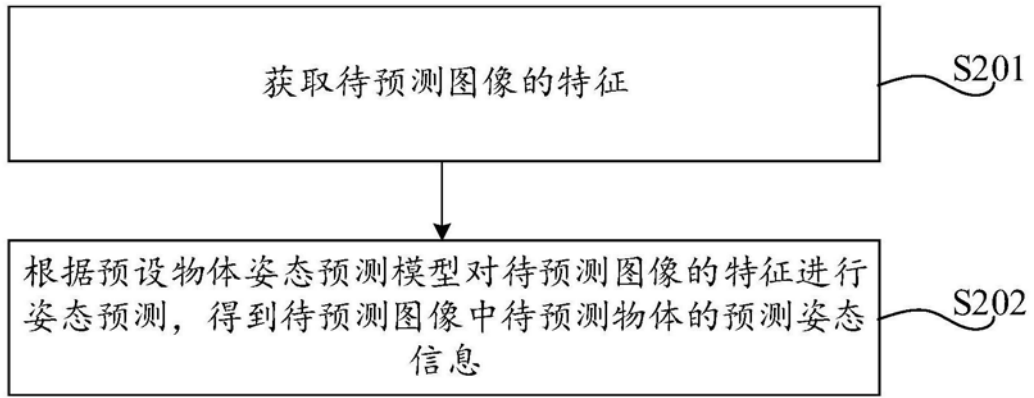


图5

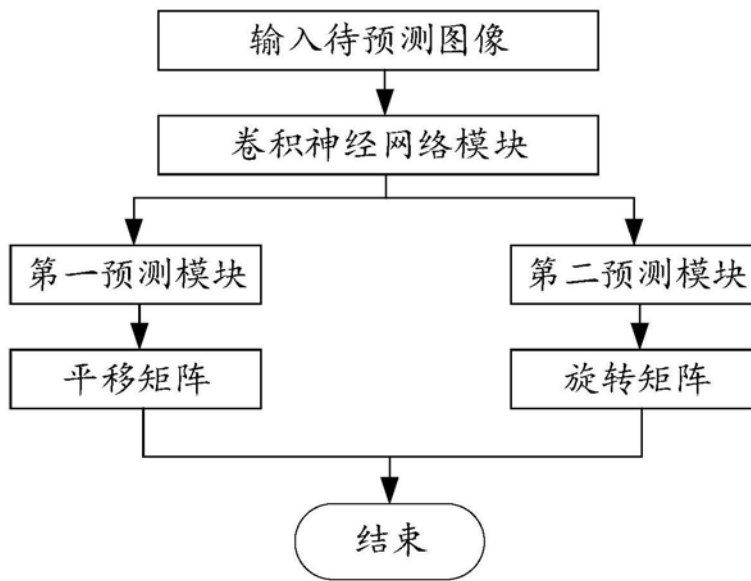


图6

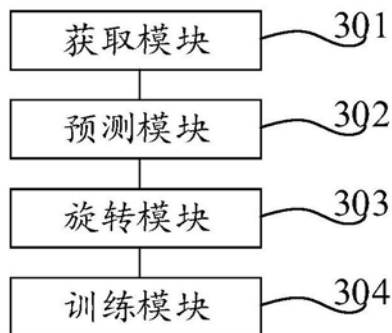


图7



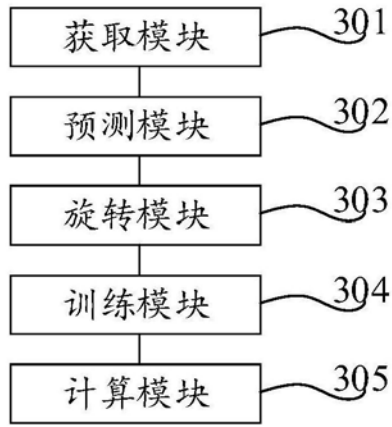


图8

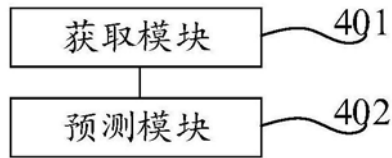


图9

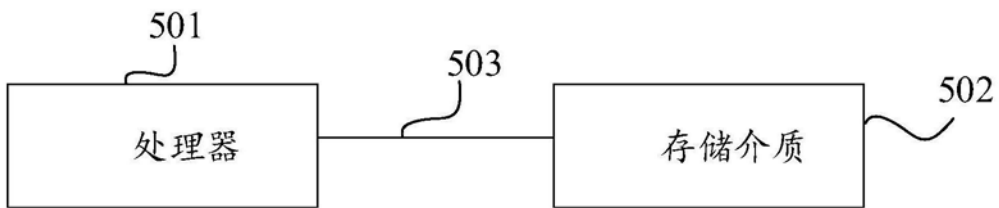


图10