



# (12) 发明专利申请

(10) 申请公布号 CN 114782510 A

(43) 申请公布日 2022. 07. 22

(21) 申请号 202210467512.3

(22) 申请日 2022.04.29

(71) 申请人 北京地平线信息技术有限公司  
地址 100094 北京市海淀区丰豪东路9号院  
2号楼3层1单元301

(72) 发明人 刘方原

(74) 专利代理机构 北京思源智汇知识产权代理  
有限公司 11657  
专利代理师 王晓多

(51) Int. Cl .  
G06T 7/50 (2017.01)  
G06V 10/40 (2022.01)

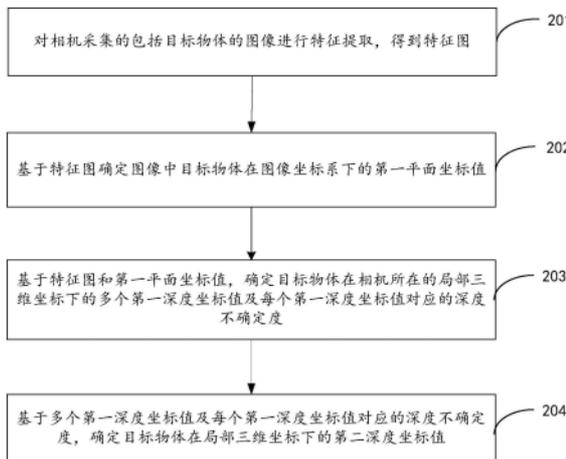
权利要求书2页 说明书11页 附图5页

## (54) 发明名称

目标物体的深度估计方法和装置、存储介质、电子设备

## (57) 摘要

本公开实施例公开了一种目标物体的深度估计方法和装置、存储介质、电子设备,其中,方法包括:对相机采集的包括目标物体的图像进行特征提取,得到特征图;基于所述特征图确定所述图像中目标物体在图像坐标系下的第一平面坐标值;基于所述特征图和所述第一平面坐标值,确定所述目标物体在所述相机所在的局部三维坐标下的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度;基于所述多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度,确定所述目标物体在所述局部三维坐标下的第二深度坐标值。



1. 一种目标物体的深度估计方法,包括:

对相机采集的包括目标物体的图像进行特征提取,得到特征图;

基于所述特征图确定所述图像中目标物体在图像坐标系下的第一平面坐标值;

基于所述特征图和所述第一平面坐标值,确定所述目标物体在所述相机所在的局部三维坐标下的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度;

基于所述多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度,确定所述目标物体在所述局部三维坐标下的第二深度坐标值。

2. 根据权利要求1所述的方法,其中,所述基于所述特征图和所述第一平面坐标值,确定所述目标物体在所述相机所在的局部三维坐标下的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度,包括:

基于所述特征图和所述第一平面坐标值,确定所述目标物体对应的多个预测值及每个所述预测值对应的预测不确定度、所述目标物体的实际高度预测值;

基于所述多个预测值和所述实际高度预测值,确定所述局部三维坐标下的多个所述第一深度坐标值;

基于所述多个预测不确定度,确定每个所述第一深度坐标值对应的所述深度不确定度。

3. 根据权利要求2所述的方法,其中,所述基于所述特征图和所述第一平面坐标值,确定所述目标物体对应的多个预测值及每个所述预测值对应的预测不确定度、所述目标物体的实际高度预测值,包括:

基于多个预测分支模型对所述特征图进行预测,确定所述图像中至少一个物体中每个物体对应的候选预测结果;

基于所述第一平面坐标值从所述至少一个候选预测结果中,确定所述目标物体对应的候选预测结果;

基于所述目标物体对应的候选预测结果,得到所述多个预测值及每个所述预测值对应的预测不确定度、所述目标物体的实际高度预测值。

4. 根据权利要求3所述的方法,其中,所述多个预测值包括以下至少之二:所述局部三维坐标下的直接深度预测值、所述局部三维坐标下的多个关键点预测值、所述局部三维坐标下的投影高度预测值;

所述基于所述多个预测值和所述实际高度预测值,确定所述局部三维坐标下的多个所述第一深度坐标值,包括以下至少之二:

基于所述直接深度预测值作为一个所述第一深度坐标值;

基于所述多个关键点预测值确定所述目标物体在图像中的投影高度值,基于所述投影高度值、所述实际高度预测值和所述相机的内参矩阵确定一个所述第一深度坐标值;

基于所述投影高度预测值、所述实际高度预测值和所述相机的内参矩阵确定一个所述第一深度坐标值。

5. 根据权利要求1-4任一所述的方法,其中,所述基于所述多个第一深度坐标值及每个所述第一深度坐标值对应的不确定度,确定所述目标物体在所述局部三维坐标下的第二深度坐标值,包括:

基于多个所述不确定度,确定每个所述第一深度坐标值的权重值;

基于多个所述权重值对所述多个第一深度坐标值加权,得到所述目标物体在所述相机对应的相机坐标下的第二深度坐标值。

6. 根据权利要求1-5任一所述的方法,还包括:

基于所述第二深度坐标值和所述相机的内参矩阵,对所述第一平面坐标值进行从所述图像坐标系到所述局部三维坐标系的逆投影变换;

确定所述目标物体在所述局部三维坐标系下的第二平面坐标值。

7. 根据权利要求6所述的方法,还包括:

基于所述特征图,确定所述目标物体在所述图像坐标系下的中心偏移值;

基于所述中心偏移值对所述第二平面坐标值进行修正,得到修正后的第三平面坐标值。

8. 一种目标物体的深度估计装置,包括:

特征提取模块,用于对相机采集的包括目标物体的图像进行特征提取,得到特征图;

平面坐标预测模块,用于基于所述特征提取模块得到的特征图确定所述图像中目标物体在图像坐标系下的第一平面坐标值;

物体坐标估计模块,用于基于所述特征提取模块得到的特征图和所述平面坐标预测模块确定的第一平面坐标值,确定所述目标物体在所述相机所在的局部三维坐标下的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度;

目标物体估计模块,用于基于所述物体坐标估计模块确定的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度,确定所述目标物体在所述局部三维坐标下的第二深度坐标值。

9. 一种计算机可读存储介质,所述存储介质存储有计算机程序,所述计算机程序用于执行上述权利要求1-7任一所述的目标物体的深度估计方法。

10. 一种电子设备,所述电子设备包括:

处理器;

用于存储所述处理器可执行指令的存储器;

所述处理器,用于从所述存储器中读取所述可执行指令,并执行所述指令以实现上述权利要求1-7任一所述的目标物体的深度估计方法。

## 目标物体的深度估计方法和装置、存储介质、电子设备

### 技术领域

[0001] 本公开涉及计算机视觉技术,尤其是一种目标物体的深度估计方法和装置、存储介质、电子设备。

### 背景技术

[0002] 深度估计,就是获取图像中场景里的每个空间中的点到相机上摄像头的距离信息,这种距离信息组成的图我们称之为深度图;其中,单目深度估计是使用单个相机从其采集的RGB图像中,估计图像中目标相对于相机的距离;基于单目相机估计深度信息被定义为单目相机估计(MDE,Monocular Depth Estimation)问题;在汽车辅助驾驶/自动驾驶领域中,车辆的深度估计是非常重要的任务需求,是场景重建和理解任务的关键步骤。

### 发明内容

[0003] 本公开的实施例提供了一种目标物体的深度估计方法和装置、存储介质、电子设备。

[0004] 根据本公开实施例的一个方面,提供了一种目标物体的深度估计方法,包括:

[0005] 对相机采集的包括目标物体的图像进行特征提取,得到特征图;

[0006] 基于所述特征图确定所述图像中目标物体在图像坐标系下的第一平面坐标值;

[0007] 基于所述特征图和所述第一平面坐标值,确定所述目标物体在所述相机所在的局部三维坐标下的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度;

[0008] 基于所述多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度,确定所述目标物体在所述局部三维坐标下的第二深度坐标值。

[0009] 根据本公开实施例的另一方面,提供了一种目标物体的深度估计装置,包括:

[0010] 特征提取模块,用于对相机采集的包括目标物体的图像进行特征提取,得到特征图;

[0011] 平面坐标预测模块,用于基于所述特征提取模块得到的特征图确定所述图像中目标物体在图像坐标系下的第一平面坐标值;

[0012] 物体坐标估计模块,用于基于所述特征提取模块得到的特征图和所述平面坐标预测模块确定的第一平面坐标值,确定所述目标物体在所述相机所在的局部三维坐标下的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度;

[0013] 目标物体估计模块,用于基于所述物体坐标估计模块确定的多个第一深度坐标值及每个所述第一深度坐标值对应的深度不确定度,确定所述目标物体在所述局部三维坐标下的第二深度坐标值。

[0014] 根据本公开实施例的又一方面,提供了一种计算机可读存储介质,所述存储介质存储有计算机程序,所述计算机程序用于执行上述任一实施例所述的目标物体的深度估计方法。

[0015] 根据本公开实施例的还一方面,提供了一种电子设备,所述电子设备包括:

- [0016] 处理器；
- [0017] 用于存储所述处理器可执行指令的存储器；
- [0018] 所述处理器，用于从所述存储器中读取所述可执行指令，并执行所述指令以实现上述任一实施例所述的目标物体的深度估计方法。
- [0019] 基于本公开上述实施例提供的一种目标物体的深度估计方法和装置、存储介质、电子设备，通过先预测多个第一深度坐标值，每个第一深度坐标值对应不同的深度估计算法，基于多个第一深度坐标值及其不确定度确定第二深度坐标值，实现了结合多种深度估计算法的深度估计，克服了单一深度估计算法导致的对该深度估计算法的依赖问题，鲁棒性好，抗干扰能力强。
- [0020] 下面通过附图和实施例，对本公开的技术方案做进一步的详细描述。

### 附图说明

- [0021] 通过结合附图对本公开实施例进行更详细的描述，本公开的上述以及其他目的、特征和优势将变得更加明显。附图用来提供对本公开实施例的进一步理解，并且构成说明书的一部分，与本公开实施例一起用于解释本公开，并不构成对本公开的限制。在附图中，相同的参考标号通常代表相同部件或步骤。
- [0022] 图1a是本公开一示例性实施例涉及的深度估计网络模型的网络结构示意图。
- [0023] 图1b是图1a提供的深度估计网络模型中预测分支模型的一种可选示例结构图。
- [0024] 图2是本公开一示例性实施例提供的目标物体的深度估计方法的流程示意图。
- [0025] 图3是本公开图2所示的实施例中步骤203的一流程示意图。
- [0026] 图4是本公开图3所示的实施例中步骤2031的一流程示意图。
- [0027] 图5是本公开图2所示的实施例中步骤204的一流程示意图。
- [0028] 图6是本公开另一示例性实施例提供的目标物体的深度估计方法的流程示意图。
- [0029] 图7是本公开一示例性实施例提供的目标物体的深度估计装置的结构示意图。
- [0030] 图8是本公开另一示例性实施例提供的目标物体的深度估计装置的结构示意图。
- [0031] 图9是本公开一示例性实施例提供的电子设备的结构图。

### 具体实施方式

- [0032] 下面，将参考附图详细地描述根据本公开的示例实施例。显然，所描述的实施例仅仅是本公开的一部分实施例，而不是本公开的全部实施例，应理解，本公开不受这里描述的示例实施例的限制。
- [0033] 应注意到：除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值不限制本公开的范围。
- [0034] 本领域技术人员可以理解，本公开实施例中的“第一”、“第二”等术语仅用于区别不同步骤、设备或模块等，既不代表任何特定技术含义，也不表示它们之间的必然逻辑顺序。
- [0035] 还应理解，在本公开实施例中，“多个”可以指两个或两个以上，“至少一个”可以指一个、两个或两个以上。
- [0036] 还应理解，对于本公开实施例中提及的任一部件、数据或结构，在没有明确限定或

者在前后文给出相反启示的情况下,一般可以理解为一个或多个。

[0037] 另外,本公开中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本公开中字符“/”,一般表示前后关联对象是一种“或”的关系。

[0038] 还应理解,本公开对各个实施例的描述着重强调各个实施例之间的不同之处,其相同或相似之处可以相互参考,为了简洁,不再一一赘述。

[0039] 同时,应当明白,为了便于描述,附图中所示出的各个部分的尺寸并不是按照实际的比例关系绘制的。

[0040] 以下对至少一个示例性实施例的描述实际上仅仅是说明性的,决不作为对本公开及其应用或使用的任何限制。

[0041] 对于相关领域普通技术人员已知的技术、方法和设备可能不作详细讨论,但在适当情况下,所述技术、方法和设备应当被视为说明书的一部分。

[0042] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步讨论。

[0043] 本公开实施例可以应用于终端设备、计算机系统、服务器等电子设备,其可与众多其它通用或专用计算系统环境或配置一起操作。适于与终端设备、计算机系统、服务器等电子设备一起使用的众所周知的终端设备、计算系统、环境和/或配置的例子包括但不限于:个人计算机系统、服务器计算机系统、瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统、大型计算机系统和包括上述任何系统的分布式云计算技术环境,等等。

[0044] 终端设备、计算机系统、服务器等电子设备可以在由计算机系统执行的计算机系统可执行指令(诸如程序模块)的一般语境下描述。通常,程序模块可以包括例程、程序、目标程序、组件、逻辑、数据结构等等,它们执行特定的任务或者实现特定的抽象数据类型。计算机系统/服务器可以在分布式云计算环境中实施,分布式云计算环境中,任务是由通过通信网络链接的远程处理设备执行的。在分布式云计算环境中,程序模块可以位于包括存储设备的本地或远程计算系统存储介质上。

[0045] 申请概述

[0046] 在实现本公开的过程中,发明人发现,现有的深度估计方法,通常为通过单一的深度估计算法进行估计,但现有技术至少存在以下问题:由于基于单一的深度深度估计算法进行估计,易受单一深度估计算法的准确性的影响,对噪声以及误判不够鲁棒。

[0047] 示例性网络结构

[0048] 图1a是本公开一示例性实施例涉及的深度估计网络模型的网络结构示意图。如图1所示,深度估计网络模型包括特征提取分支模型101、预测分支模型102和特征融合分支模型103。

[0049] 可选地,图1b是图1a提供的深度估计网络模型中预测分支模型的一种可选示例结构图。如图1b所示,预测分支模型102可以包括但不限于:平面预测分支模型1021、深度预测分支模型1022、关键点分支模型1023、投影高度分支模型1024和实际高度预测分支模型1025。

[0050] 利用图1a及图1b提供的深度估计网络模型进行深度估计的过程可包括:将相机采

集的单张图像(例如,RGB图像)输入深度估计网络模型的特征提取分支模型101,通过特征提取分支模型101对图像进行特征提取,得到特征图;将该特征图分别输入多个预测分支模型102;通过平面预测分支模型1021对特征图进行处理,得到目标物体(例如,车辆等)在图像对应的图像坐标系下的第一平面坐标(u,v)和中心偏移值,其中,第一平面坐标表示的是目标物体的中心点坐标值,通过该第一平面坐标可从关键点分支模型1023、投影高度分支模型1024和实际高度预测分支模型1025输出的对应图像中至少一个物体的结果中确定目标物体对应的结果;通过深度预测分支模型1022对特征图进行处理,得到目标物体对应的第一深度估计值和第一不确定度,其中,深度预测分支模型1022可输出图像中至少一个物体对应的至少一个候选估计值,通过第一平面坐标确定与第一平面坐标距离最近的候选估计值作为第一深度估计值;通过关键点分支模型1023对特征图进行处理,得到目标物体的多个关键点预测值和第二不确定度;其中,目标物体在图像中的关键点可以是目标物体在图像中的最小包含立方体对应的8个关键点的坐标值,具体确定过程可包括:关键点分支模型1023可预测得到图像中包括的至少一个物体对应的至少一组关键点,通过第一平面坐标对多个关键点预测值进行筛选(距离最近),确定与第一平面坐标相近的8个关键点坐标作为关键点分支模型1023的输出结果;例如,目标物体为车辆时,关键点为将车辆包含在内的立方体对应的8个顶点坐标值;通过投影高度分支模型1024对特征图进行处理,得到目标物体的投影高度值和第三不确定度,该投影高度表示目标物体在图像中的高度,可选地,投影高度分支模型1024可预测出多个预测投影高度值(例如,图像中包括多个物体等),通过第一平面坐标对多个预测投影高度值进行筛选,确定目标物体的投影高度值;通过实际高度预测分支模型1025对特征图进行处理,得到目标物体的实际高度预测值,例如,目标物体为车辆时,该实际高度预测值为预测的车辆的实际高度。

[0051] 基于上述多个预测分支模型102的输出可确定三个深度估计值,可选地,包括第一深度估计值 $Z_1$ ,基于第一平面坐标、多个关键点预测值和相机的内参矩阵确定第二深度估计值 $Z_2$ ,基于第一平面坐标、投影高度值和相机的内参矩阵确定第三深度估计值 $Z_3$ ;

[0052] 可选地,第二深度估计值可基于以下公式(1)确定:

$$[0053] \quad Z_2 = \frac{f \cdot H}{(k_2 - k_1)} \quad \text{公式 (1)}$$

[0054] 其中, $k_1$ 表示目标物体在图像中的下关键点坐标值(例如,通过8个关键点中属于底面的4个关键点坐标值的均值确定), $k_2$ 表示目标物体在图像中的上关键点坐标值(例如,通过8个关键点中属于顶面的4个关键点坐标值的均值确定),具体地,本实施例中的8个关键点是基于第一平面坐标筛选得到; $f$ 表示相机内参中的焦距; $H$ 表示实际高度预测分支模型预测的实际高度预测值。

[0055] 第三深度估计值可基于以下公式(2)确定:

$$[0056] \quad Z_3 = \frac{f \cdot H}{h} \quad \text{公式 (2)}$$

[0057] 其中, $h$ 表示投影高度分支模型预测的投影高度值,具体地,本实施例中的投影高度值是基于第一平面坐标筛选确定的; $f$ 表示相机内参中的焦距; $H$ 表示实际高度预测分支模型预测的实际高度预测值。

[0058] 特征融合分支模型103根据第一深度估计值、第二深度估计值和第三深度估计值,

以及第一不确定度、第二不确定度和第三不确定度,确定目标物体在相机所在的局部三维坐标下的目标深度估计值。可选地,可基于以下公式(3)确定目标深度估计值:

$$[0059] \quad Z = \left( \sum_{i=1}^3 \frac{z_i}{\sigma_i} \right) / \left( \sum_{i=1}^3 \frac{1}{\sigma_i} \right) \quad \text{公式 (3)}$$

[0060] 其中,Z表示目标深度估计值;i为预测分支模型102输出的深度估计值的数量,在本实施例中取值为3; $\sigma_i$ 表示不确定度, $\sigma_1$ 、 $\sigma_2$ 和 $\sigma_3$ 分别表示第一不确定度、第二不确定度和第三不确定度,不确定度的倒数可表示对应的深度估计值的权重值,即公式(3)实现加权求平均。

[0061] 以上,即可实现估计目标物体局部三维坐标下的目标深度估计值,并且,为了更好的使用目标深度估计值,结合内参矩阵K,对第一平面坐标(u,v)进行逆投影变换,得到目标物体的中心点在局部三维坐标下的X轴和Y轴的坐标值,即,本实施例实现了确定目标物体所在局部三维坐标下的三维坐标值。

[0062] 本实施例中通过三种深度估计算法对目标物体的深度值进行估计,相较于基于现有技术中的基于单一深度估计算法的深度估计方法,本实施例对单个的深度估计算法的依赖较小,对噪声误判不敏感,从而实现更加鲁棒抗干扰;并且,相较直接深度估计算法,从图像上预测关键点和投影高度更加简单,使深度估计更容易,并且结合三种深度估计算法及其不确定度提升了深度估计的精度。

[0063] 通常情况下,网络模型在应用之前需要进行训练,本公开实施例上述图1a及图1b提供的深度估计网络模型在应用之前,同样需要进行网络训练,训练过程包括,将已知深度真值的样本图像输入深度估计网络模型,输出深度估计值;基于深度估计值和不确定度确定网络损失,利用网络损失对深度估计网络模型进行训练,可选地,网络损失可以基于分支损失 $L_i$ 与深度损失 $L_{\text{depth}}$ 的和确定;其中,分支损失可基于以下公式(4)确定,i的取值为1、2、3:

$$[0064] \quad L_i = \frac{|z_i - z^*|}{\sigma_i} + \log \sigma_i \quad \text{公式 (4)}$$

[0065] 其中, $\sigma_i$ 表示不确定度(i取不同值时,分别对应深度预测分支模型1022、关键点分支模型1023和投影高度分支模型1024输出的不确定度), $z_i$ 表示估计的深度值(i取不同值时,分别对应深度预测分支模型1022、关键点分支模型1023和投影高度分支模型1024输出的深度度), $z^*$ 表示样本图像对应的深度真值。

[0066] 深度损失 $L_{\text{depth}}$ 可基于以下公式(5)确定:

$$[0067] \quad L_{\text{depth}} = |z - z^*| \quad \text{公式 (5)}$$

[0068] 其中,z表示基于深度估计网络模型预测得到的目标物体的深度估计值; $z^*$ 表示样本图像对应的深度真值。

[0069] 示例性方法

[0070] 图2是本公开一示例性实施例提供的目标物体的深度估计方法的流程示意图。本实施例可应用在电子设备上,如图2所示,包括如下步骤:

[0071] 步骤201,对相机采集的包括目标物体的图像进行特征提取,得到特征图。

[0072] 本实施例中,可通过上述图1a所示的深度估计网络模型中的特征提取分支模型101进行特征提取,得到特征图;本实施例可应用于任意场景,例如,利用车载相机采集车辆

周围图像,此时,目标物体可以为其他车辆、行人、建筑物等障碍物,此时,通过本实施例的深度估计方法,可为车辆的辅助驾驶或自动驾驶提供更多信息,提高驾驶安全性。

[0073] 步骤202,基于特征图确定图像中目标物体在图像坐标系下的第一平面坐标值。

[0074] 可选地,本实施例可基于上述图1b中提供的平面预测分支模型1021实现对第一平面坐标值的预测。

[0075] 步骤203,基于特征图和第一平面坐标值,确定目标物体在相机所在的局部三维坐标下的多个第一深度坐标值及每个第一深度坐标值对应的深度不确定度。

[0076] 可选地,本实施例可通过图1b提供多个预测分支模型102中的深度预测分支模型1022、关键点分支模型1023和投影高度分支模型1024和实际高度预测分支模型1025中的多个分支模型对特征图进行处理,并结合第一平面坐标值,输出目标物体在相机所在的局部三维坐标下的多个第一深度坐标值及每个第一深度坐标值对应的深度不确定度。

[0077] 步骤204,基于多个第一深度坐标值及每个第一深度坐标值对应的深度不确定度,确定目标物体在局部三维坐标下的第二深度坐标值。

[0078] 可选地,可通过每个第一深度坐标值对应的深度不确定度,确定每个第一深度坐标值的权重值,结合权重值进行加权求平均,即可确定第二深度坐标值。

[0079] 本公开上述实施例提供的一种目标物体的深度估计方法,通过多个预测分支模型预测多个第一深度坐标值,每个第一深度坐标值多个预测分支模型对应不同的深度估计算法,基于多个第一深度坐标值及其不确定度确定第二深度坐标值,实现了结合多种深度估计算法的深度估计,克服了单一深度估计算法导致的对该深度估计算法的依赖问题,鲁棒性好,抗干扰能力强。

[0080] 如图3所示,在上述图2所示实施例的基础上,步骤203可包括如下步骤:

[0081] 步骤2031,基于特征图和第一平面坐标值,确定目标物体对应的多个预测值及每个预测值对应的预测不确定度、目标物体的实际高度预测值。

[0082] 本实施例中的预测值可以包括但不限于:深度预测值、关键点预测值、投影高度预测值;可选地,实际高度预测值可通过图1b提供的投影高度分支模型1024预测获得。

[0083] 步骤2032,基于多个预测值和实际高度预测值,确定局部三维坐标下的多个第一深度坐标值。

[0084] 本实施例中,当预测值为关键点预测值或投影高度预测值时,可通过图1b提供的实施例中的公式(1)或公式(2)进行计算,确定对应的第一深度坐标值。

[0085] 步骤2033,基于多个预测不确定度,确定每个第一深度坐标值对应的深度不确定度。

[0086] 由于每个预测值对应一个预测不确定度,因此,以该预测值确定的对应的第一深度坐标值对应的深度不确定度为该预测值对应的预测不确定度。

[0087] 本实施例中提出了包括深度、关键点、投影高度等信息对应的多个预测值,基于多个预测值分别通过不同的深度估计算法确定的多个第一深度坐标值,增加了深度估计算法,克服了仅依赖一种深度估计算法而导致过于依赖该深度估计算法的问题,避免了由于单一深度估计算法不准确而导致的深度估计结果不准确的问题,提升了本实施例提供的深度估计方法的鲁棒性。

[0088] 如图4所示,在上述图3所示实施例的基础上,步骤2031可包括如下步骤:

[0089] 步骤401,基于多个预测分支模型对特征图进行预测,确定图像中至少一个物体中每个物体对应的候选预测结果。

[0090] 可选地,每个候选结果中包括候选预测值和候选预测值对应的不确定度;可通过上述图1b所示实施例中提供的深度预测分支模型1022、关键点分支模型1023、投影高度分支模型1024中的多个进行预测,得到每个物体分别对应的多个候选预测结果。

[0091] 步骤402,基于第一平面坐标值从至少一个候选预测结果中,确定目标物体对应的候选预测结果。

[0092] 本实施例中,第一平面坐标值包括目标物体的中心点的水平方向坐标值和垂直方向坐标值,即表示目标物体的中心点在图像中的中心点坐标,以该中心点坐标为基准,可从对应多个物体的候选预测结果中,筛选确定对应目标物体的候选预测结果,可选地,可通过距离该中心点坐标的距离进行筛选等。

[0093] 步骤403,基于目标物体对应的候选预测结果,得到多个预测值及每个预测值对应的预测不确定度、目标物体的实际高度预测值。

[0094] 本实施例中,利用平面坐标值从多个候选预测结果中确定目标物体对应的候选预测结果,由于图像中可能包括多个物体(例如,路面图像中可能包括多个车辆、行人、建筑物等),通过目标物体的平面坐标值确定对应目标物体的一组预测值;通过第一平面坐标值筛选,使后续操作仅基于目标物体对应的候选预测结果,无需对每个物体对应的候选预测结果都执行后续操作,解决了由于对所有候选预测结果都进行深度估计导致的结果不是针对目标物体而不准确的问题,提升了目标物体对应的至少一个预测值的准确性。

[0095] 可选地,在上述实施例的基础上,步骤2032可包括以下至少之二:

[0096] a1,基于直接深度预测值作为一个第一深度坐标值。

[0097] 本实施例中,可基于图1b提供的深度预测分支模型1022确定第一深度坐标值,可选地,本实施例中的深度预测分支模型可以采用任一现有技术中进行深度预测的网络模型,本实施例对深度预测的网络结构不做具体限制。

[0098] a2,基于多个关键点预测值确定目标物体在图像中的投影高度值,基于投影高度值、实际高度预测值和相机的内参矩阵确定一个第一深度坐标值。

[0099] 可选地,基于上述图1b提供的实施例中的公式(1)实现一个第一深度坐标值的预测,其中,关键点包括包围目标物体的立方体的8个关键点,以上下平面的关键点的差值确定投影高度值。

[0100] a3,基于投影高度预测值、实际高度预测值和相机的内参矩阵确定一个第一深度坐标值。

[0101] 可选地,基于上述图1b提供的实施例中的公式(2)实现一个第一深度坐标值的预测。

[0102] 本实施例中,多个预测值包括以下至少之二:局部三维坐标下的直接深度预测值、局部三维坐标下的多个关键点预测值、局部三维坐标下的投影高度预测值,分别通过不同的分支网络模型确定不同预测值,其中,每个预测值通过不同的深度估计算法确定一个第一深度坐标值,以多个第一深度坐标值及其不确定度确定第二深度坐标值,实现了结合多种深度估计算法的深度估计,克服了单一深度估计算法导致的对该深度估计算法的依赖问题,提升了深度估计方法的鲁棒性;并且,相对于直接估计深度值,对目标图像进行关键点

预测和投影高度预测更加简单,容易实现,使本实施例的深度估计更易于实现,且结合多种预测值时得到的深度估计值更准确。

[0103] 如图5所示,在上述图2所示实施例的基础上,步骤204可包括如下步骤:

[0104] 步骤2041,基于多个不确定度,确定每个第一深度坐标值的权重值。

[0105] 可选地,基于每个不确定度的倒数作为每个第一深度坐标值的权重值。

[0106] 步骤2042,基于多个权重值对多个第一深度坐标值加权,得到目标物体在相机对应的相机坐标下的第二深度坐标值。

[0107] 本实施例中,在确定每个第一深度坐标值的权重值之后,可基于上述图1b提供的实施例中的公式(3)确定第二深度坐标值,本实施例中通过加权求平均确定第二深度坐标值,结合了多种深度估计算法,克服了现有深度估计方法对单一深度估计算法的依赖问题,提供了更加鲁棒的深度估计方法。

[0108] 图6是本公开另一示例性实施例提供的目标物体的深度估计方法的流程示意图。如6所示,本实施例在图2所示的实施例的基础上,还可以包括如下步骤:

[0109] 步骤601,基于第二深度坐标值和相机的内参矩阵,对第一平面坐标值进行从图像坐标系到局部三维坐标系的逆投影变换。

[0110] 可选地,相机的内参矩阵在确定相机时即可确定为已知数据;本实施例结合第二深度坐标值和相机内参矩阵,将在图像坐标系下的平面坐标通过逆投影变换映射到相机所在局部三维坐标下,进而得到在局部三维坐标系下的X轴和Y轴的坐标值。

[0111] 步骤602,确定目标物体在局部三维坐标系下的第二平面坐标值。

[0112] 本实施例中,为了更好的使用目标深度估计值,结合内参矩阵K,对第一平面坐标进行逆投影变换,得到目标物体的中心点在局部三维坐标下的X轴和Y轴的坐标值,即,本实施例实现了确定目标物体在局部三维坐标下的三维坐标值,实现更好的对目标物体进行定位,为后续其他操作提供了基础。

[0113] 可选地,在上述实施例的基础上,还可以包括:

[0114] b1,基于特征图,确定目标物体在图像坐标系下的中心偏移值。

[0115] 可选地,如图1b提供的实施例所示,通过平面预测分支模型1021不仅可以输出第一平面坐标,还可以输出中心偏移值。

[0116] b2,基于中心偏移值对第二平面坐标值进行修正,得到修正后的第三平面坐标值。

[0117] 本实施例中,通过输出的中心偏移值,对第二平面坐标值进行修正,得到修正后的目标物体的中心点在局部三维坐标系下的X轴和Y轴的坐标值;通过修正,提高了目标物体中心点坐标的预测准确性。

[0118] 本公开实施例提供的任一种目标物体的深度估计方法可以由任意适当的具有数据处理能力的设备执行,包括但不限于:终端设备和服务器等。或者,本公开实施例提供的任一种目标物体的深度估计方法可以由处理器执行,如处理器通过调用存储器存储的相应指令来执行本公开实施例提及的任一种目标物体的深度估计方法。下文不再赘述。

[0119] 示例性装置

[0120] 图7是本公开一示例性实施例提供的目标物体的深度估计装置的结构示意图。如图7所示,本实施例提供的装置包括:

[0121] 特征提取模块71,用于对相机采集的包括目标物体的图像进行特征提取,得到特

征图。

[0122] 平面坐标预测模块72,用于基于特征提取模块71得到的特征图确定图像中目标物体在图像坐标系下的第一平面坐标值。

[0123] 物体坐标估计模块73,用于基于特征提取模块71得到的特征图和平面坐标预测模块72确定的第一平面坐标值,确定目标物体在相机所在的局部三维坐标下的多个第一深度坐标值及每个第一深度坐标值对应的深度不确定度。

[0124] 目标物体估计模块74,用于基于物体坐标估计模块73确定的多个第一深度坐标值及每个第一深度坐标值对应的深度不确定度,确定目标物体在局部三维坐标下的第二深度坐标值。

[0125] 本公开上述实施例提供的一种目标物体的深度估计装置,过先预测多个第一深度坐标值,每个第一深度坐标值对应不同的深度估计算法,基于多个第一深度坐标值及其不确定度确定第二深度坐标值,实现了结合多种深度估计算法的深度估计,克服了单一深度估计算法导致的对该深度估计算法的依赖问题,鲁棒性好,抗干扰能力强。

[0126] 图8是本公开另一示例性实施例提供的目标物体的深度估计装置的结构示意图。如图8所示,本实施例提供的装置中,物体坐标估计模块73,包括:

[0127] 物体预测单元731,用于基于特征图和第一平面坐标值,确定目标物体对应的多个预测值及每个预测值对应的预测不确定度、目标物体的实际高度预测值;

[0128] 坐标值确定单元732,用于基于多个预测值和实际高度预测值,确定局部三维坐标下的多个第一深度坐标值;

[0129] 深度确定单元733,用于基于多个预测不确定度,确定每个第一深度坐标值对应的深度不确定度。

[0130] 可选地,物体预测单元731,具体用于基于多个预测分支模型对特征图进行预测,确定图像中至少一个物体中每个物体对应的候选预测结果;基于第一平面坐标值从至少一个候选预测结果中,确定目标物体对应的候选预测结果;基于目标物体对应的候选预测结果,得到多个预测值及每个预测值对应的预测不确定度、目标物体的实际高度预测值。

[0131] 可选地,多个预测值包括以下至少之二:局部三维坐标下的直接深度预测值、局部三维坐标下的多个关键点预测值、局部三维坐标下的投影高度预测值;

[0132] 坐标值确定单元732,具体用于基于直接深度预测值作为一个第一深度坐标值;和/或,基于多个关键点预测值确定目标物体在图像中的投影高度值,基于投影高度值、实际高度预测值和相机的内参矩阵确定一个第一深度坐标值;和/或,基于投影高度预测值、实际高度预测值和相机的内参矩阵确定一个第一深度坐标值。

[0133] 在一些可选的实施例中,目标物体估计模块74,包括:

[0134] 权重值单元741,用于基于多个不确定度,确定每个第一深度坐标值的权重值;

[0135] 坐标值估计单元742,用于基于多个权重值对多个第一深度坐标值加权,得到目标物体在相机对应的相机坐标下的第二深度坐标值。

[0136] 在一些可选的实施例中,还包括:

[0137] 逆投影模块85,用于基于第二深度坐标值和相机的内参矩阵,对第一平面坐标值进行从图像坐标系到局部三维坐标系的逆投影变换;确定目标物体在局部三维坐标系下的第二平面坐标值。

[0138] 在一些可选的实施例中,还包括:

[0139] 坐标修正模块86,用于基于特征图,确定目标物体在图像坐标系下的中心偏移值;基于中心偏移值对第二平面坐标值进行修正,得到修正后的第三平面坐标值。

[0140] 示例性电子设备

[0141] 下面,参考图9来描述根据本公开实施例的电子设备。该电子设备可以是第一设备100和第二设备200中的任一个或两者、或与它们独立的单机设备,该单机设备可以与第一设备和第二设备进行通信,以从它们接收所采集到的输入信号。

[0142] 图9图示了根据本公开实施例的电子设备的框图。

[0143] 如图9所示,电子设备90包括一个或多个处理器91和存储器92。

[0144] 处理器91可以是中央处理单元(CPU)或者具有数据处理能力和/或指令执行能力的其他形式的处理单元,并且可以控制电子设备90中的其他组件以执行期望的功能。

[0145] 存储器92可以包括一个或多个计算机程序产品,所述计算机程序产品可以包括各种形式的计算机可读存储介质,例如易失性存储器和/或非易失性存储器。所述易失性存储器例如可以包括随机存取存储器(RAM)和/或高速缓冲存储器(cache)等。所述非易失性存储器例如可以包括只读存储器(ROM)、硬盘、闪存等。在所述计算机可读存储介质上可以存储一个或多个计算机程序指令,处理器91可以运行所述程序指令,以实现上文所述的本公开的各个实施例的目标物体的深度估计方法以及/或者其他期望的功能。在所述计算机可读存储介质中还可以存储诸如输入信号、信号分量、噪声分量等各种内容。

[0146] 在一个示例中,电子设备90还可以包括:输入装置93和输出装置94,这些组件通过总线系统和/或其他形式的连接机构(未示出)互连。

[0147] 例如,在该电子设备是第一设备100或第二设备200时,该输入装置93可以是上述的麦克风或麦克风阵列,用于捕捉声源的输入信号。在该电子设备是单机设备时,该输入装置93可以是通信网络连接,用于从第一设备100和第二设备200接收所采集的输入信号。

[0148] 此外,该输入装置93还可以包括例如键盘、鼠标等等。

[0149] 该输出装置94可以向外部输出各种信息,包括确定出的距离信息、方向信息等。该输出装置94可以包括例如显示器、扬声器、打印机、以及通信网络及其所连接的远程输出设备等等。

[0150] 当然,为了简化,图9中仅示出了该电子设备90中与本公开有关的组件中的一些,省略了诸如总线、输入/输出接口等等的组件。除此之外,根据具体应用情况,电子设备90还可以包括任何其他适当的组件。

[0151] 示例性计算机程序产品和计算机可读存储介质

[0152] 除了上述方法和设备以外,本公开的实施例还可以是计算机程序产品,其包括计算机程序指令,所述计算机程序指令在被处理器运行时使得所述处理器执行本说明书上述“示例性方法”部分中描述的根据本公开各种实施例的目标物体的深度估计方法中的步骤。

[0153] 所述计算机程序产品可以以一种或多种程序设计语言的任意组合来编写用于执行本公开实施例操作的程序代码,所述程序设计语言包括面向对象的程序设计语言,诸如Java、C++等,还包括常规的过程式程序设计语言,诸如“C”语言或类似的设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备

或服务器上执行。

[0154] 此外,本公开的实施例还可以是计算机可读存储介质,其上存储有计算机程序指令,所述计算机程序指令在被处理器运行时使得所述处理器执行本说明书上述“示范性方法”部分中描述的根据本公开各种实施例的目标物体的深度估计方法中的步骤。

[0155] 所述计算机可读存储介质可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以包括但不限于电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。可读存储介质的更具体的例子(非穷举的列表)包括:具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

[0156] 以上结合具体实施例描述了本公开的基本原理,但是,需要指出的是,在本公开中提及的优点、优势、效果等仅是示例而非限制,不能认为这些优点、优势、效果等是本公开的各个实施例必须具备的。另外,上述公开的具体细节仅是为了示例的作用和便于理解的作用,而非限制,上述细节并不限制本公开为必须采用上述具体的细节来实现。

[0157] 本说明书中各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其它实施例的不同之处,各个实施例之间相同或相似的部分相互参见即可。对于系统实施例而言,由于其与方法实施例基本对应,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0158] 本公开中涉及的器件、装置、设备、系统的方框图仅作为例示性的例子并且不意图要求或暗示必须按照方框图示出的方式进行连接、布置、配置。如本领域技术人员将认识到的,可以按任意方式连接、布置、配置这些器件、装置、设备、系统。诸如“包括”、“包含”、“具有”等等的词语是开放性词汇,指“包括但不限于”,且可与其互换使用。这里所使用的词汇“或”和“和”指词汇“和/或”,且可与其互换使用,除非上下文明确指示不是如此。这里所使用的词汇“诸如”指词组“诸如但不限于”,且可与其互换使用。

[0159] 可能以许多方式来实现本公开的方法和装置。例如,可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本公开的方法和装置。用于所述方法的步骤的上述顺序仅是为了进行说明,本公开的方法的步骤不限于以上具体描述的顺序,除非以其它方式特别说明。此外,在一些实施例中,还可将本公开实施为记录在记录介质中的程序,这些程序包括用于实现根据本公开的方法的机器可读指令。因而,本公开还覆盖存储用于执行根据本公开的方法的程序的记录介质。

[0160] 还需要指出的是,在本公开的装置、设备和方法中,各部件或各步骤是可以分解和/或重新组合的。这些分解和/或重新组合应视为本公开的等效方案。

[0161] 提供所公开的方面的以上描述以使本领域的任何技术人员能够做出或者使用本公开。对这些方面的各种修改对于本领域技术人员而言是非常显而易见的,并且在此定义的一般原理可以应用于其他方面而不脱离本公开的范围。因此,本公开不意图被限制到在此示出的方面,而是按照与在此公开的原理和新颖的特征一致的最宽范围。

[0162] 为了例示和描述的目的已经给出了以上描述。此外,此描述不意图将本公开的实施例限制到在此公开的形式。尽管以上已经讨论了多个示例方面和实施例,但是本领域技术人员将认识到其某些变型、修改、改变、添加和子组合。

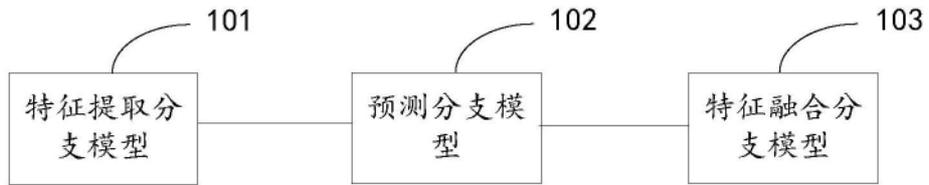


图1a

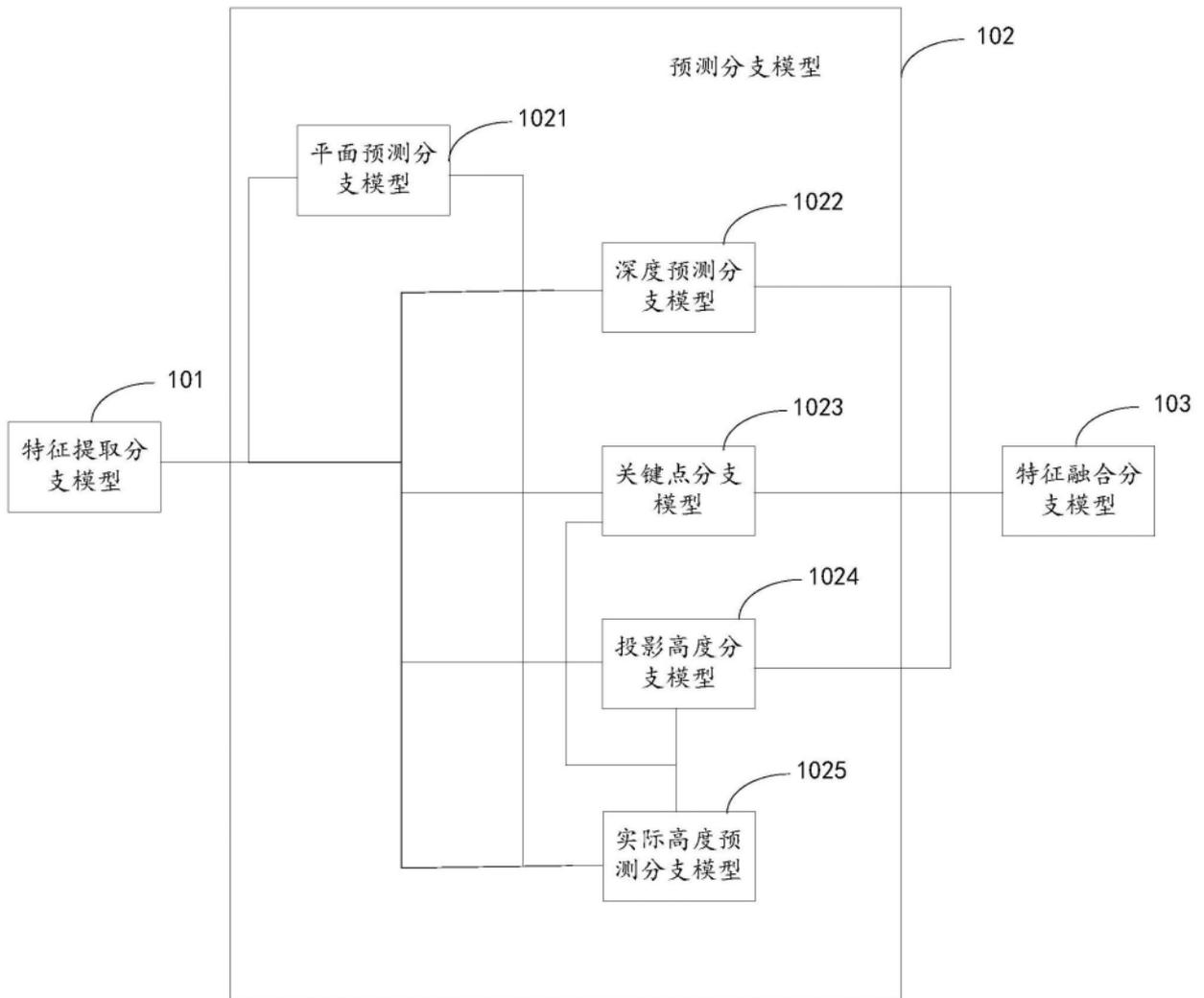


图1b

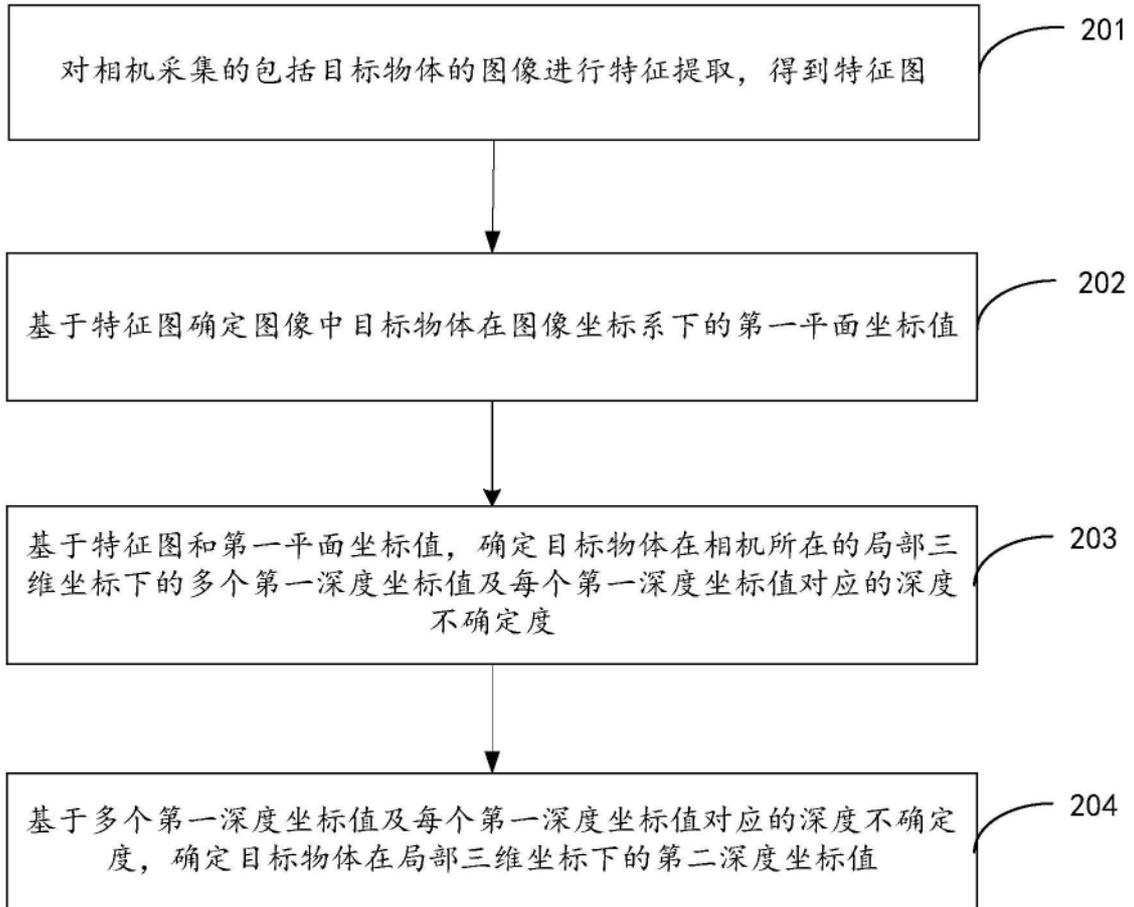


图2

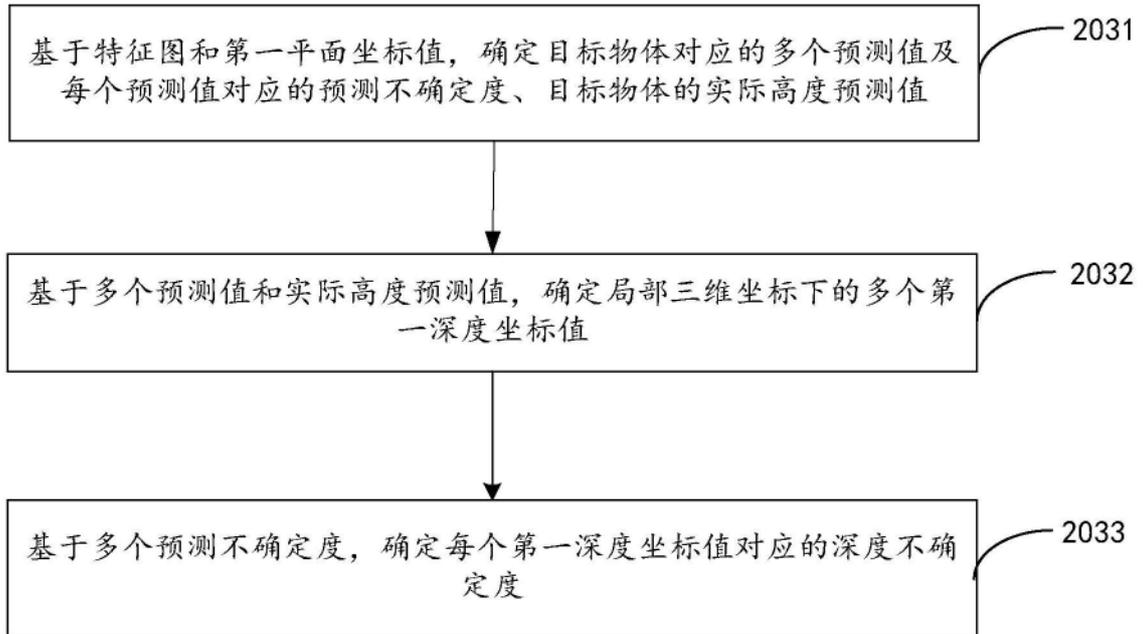


图3

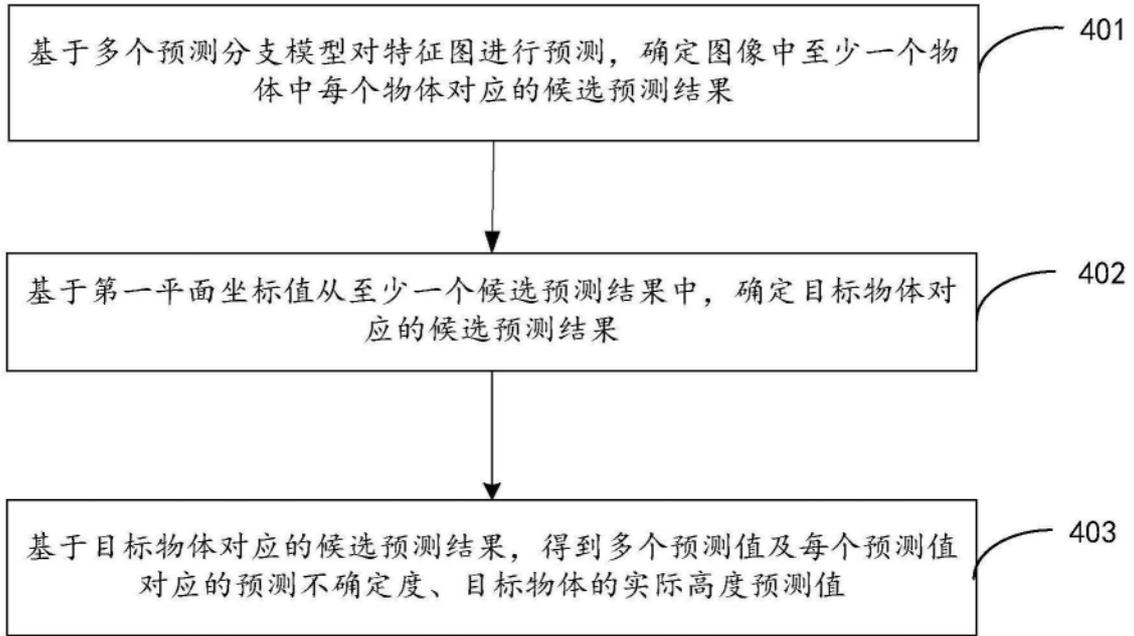


图4

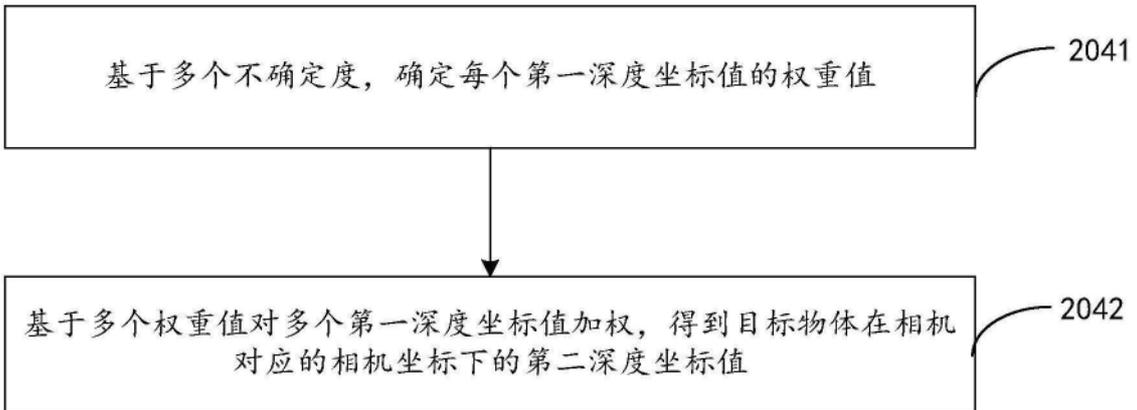


图5

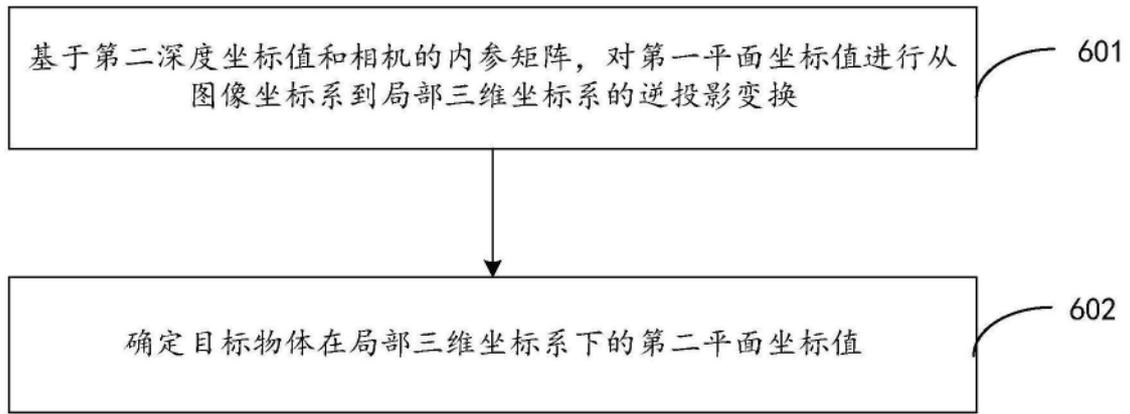


图6

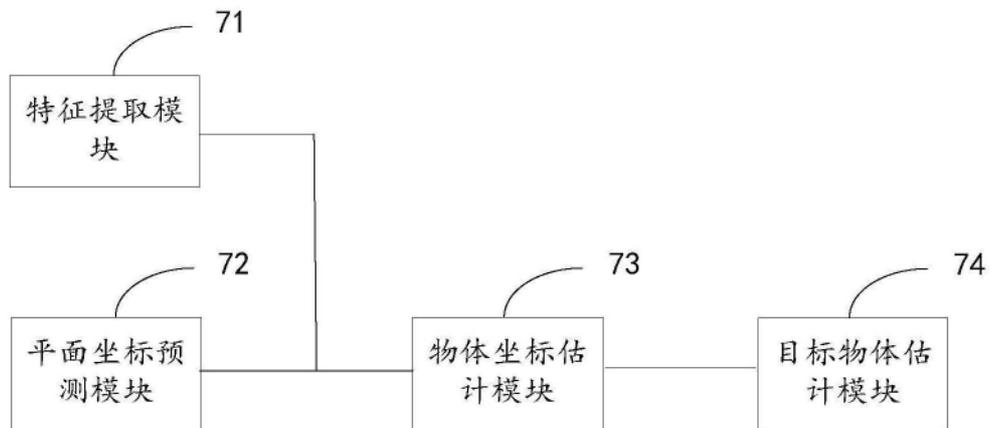


图7

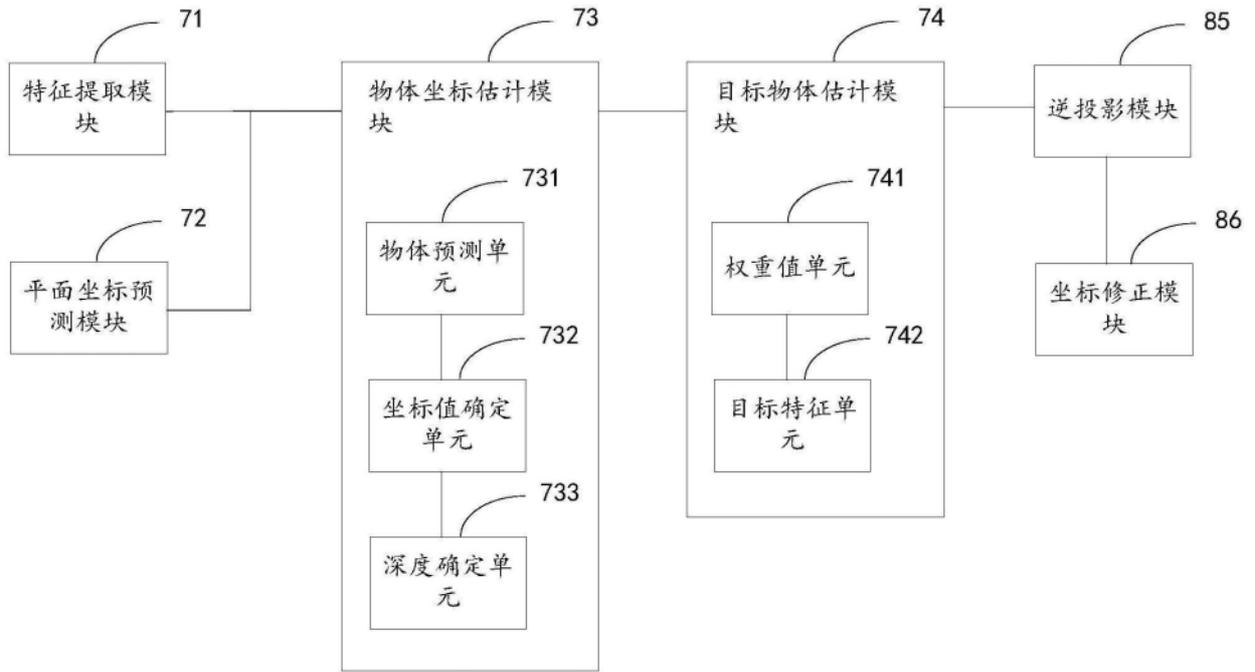


图8

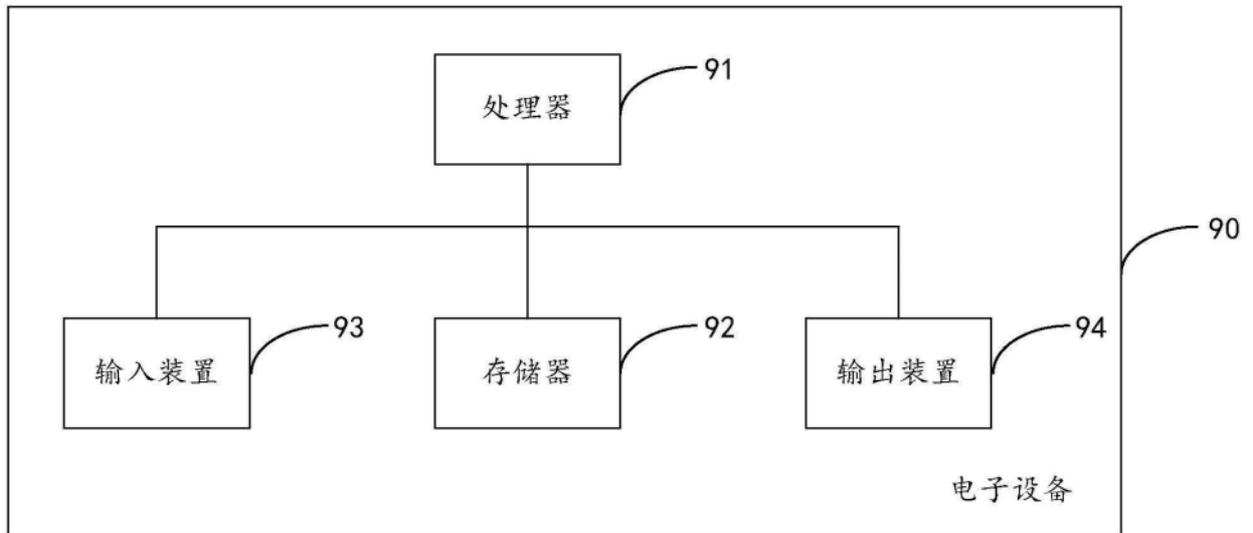


图9