



(12) 发明专利

(10) 授权公告号 CN 113902856 B

(45) 授权公告日 2023. 08. 25

(21) 申请号 202111320552.7

G06T 19/20 (2011.01)

(22) 申请日 2021.11.09

G06N 3/04 (2023.01)

G06N 3/08 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 113902856 A

(56) 对比文件

(43) 申请公布日 2022.01.07

CN 112085840 A, 2020.12.15

CN 113160420 A, 2021.07.23

(73) 专利权人 浙江商汤科技开发有限公司

US 2020082553 A1, 2020.03.12

地址 311215 浙江省杭州市萧山区宁围街道市心北路857号288-8室

审查员 舒瀚

(72) 发明人 段永利 周晓巍 鲍虎军 孙佳明 甄佳楠

(74) 专利代理机构 北京中知恒瑞知识产权代理有限公司 11889

专利代理师 吴迪

(51) Int. Cl.

G06T 17/00 (2006.01)

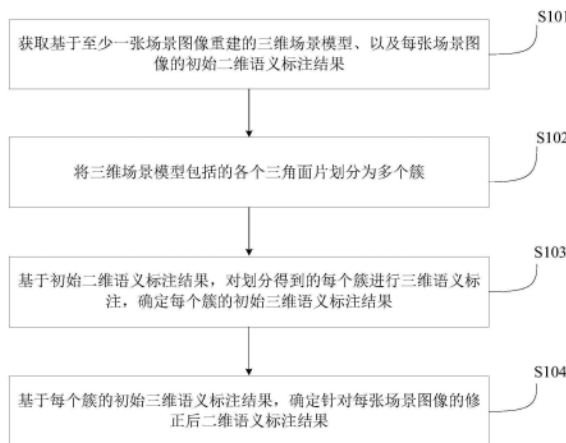
权利要求书4页 说明书15页 附图2页

(54) 发明名称

一种语义标注的方法、装置、电子设备及存储介质

(57) 摘要

本公开提供了一种语义标注的方法、装置、电子设备及存储介质,其中,该方法包括:获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;将三维场景模型包括的各个三角面片划分为多个簇;基于初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个簇的初始三维语义标注结果;基于每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。本公开基于三维空间内簇的初始三维语义标注结果对二维空间内场景图像的初始二维语义标注结果进行了修正,使得修正后二维语义标注结果更为准确,同时,上述修正过程自动完成,使得标注效率更高。



1. 一种语义标注的方法,其特征在于,所述方法包括:

获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

从所述三维场景模型包括的各个三角面片中,随机选取预设数量个三角面片,将选取的每个三角面片的目标向量分别作为待划分的簇的中心向量;确定所述各个三角面片中每个三角面片的所述目标向量,分别与每个所述中心向量之间的距离,并将该三角面片划分至对应的所述距离最小的中心向量所在的簇;确定每个簇对应的新的中心向量,并基于该新的中心向量,返回执行划分簇的步骤,直到满足划分截止条件;

针对每个三角面片,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,确定该三角面片在至少一张场景图像中对应的投影图像区域;以及,基于该三角面片在所述至少一张场景图像中对应的所述投影图像区域的所述初始二维语义标注结果,确定该三角面片的初始三维语义标注结果;针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果;

基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

2. 根据权利要求1所述的方法,其特征在于,所述确定每个簇对应的新的中心向量,包括:

基于每个簇包括的各个三角面片的目标向量,确定各个三角面片的平均向量,将所述平均向量作为该簇对应的新的中心向量。

3. 根据权利要求2所述的方法,其特征在于,所述目标向量包括目标法向量和目标颜色向量,所述中心向量包括中心法向量和中心颜色向量;所述确定所述各个三角面片中每个三角面片的所述目标向量,分别与每个所述中心向量之间的距离,包括:

针对所述各个三角面片中的每个三角面片,基于该三角面片的目标法向量以及每个所述中心向量中的中心法向量,确定该三角面片与每个所述中心向量所在的簇之间的第一距离,以及,基于该三角面片的目标颜色向量以及每个所述中心向量中的中心颜色向量,确定该三角面片与每个所述中心向量所在的簇之间的第二距离;

基于所述第一距离及其对应的权重、以及所述第二距离及其对应的权重,确定该三角面片的所述目标向量,分别与每个所述中心向量之间的距离。

4. 根据权利要求1所述的方法,其特征在于,所述基于该三角面片在所述至少一张场景图像中对应的所述投影图像区域的所述初始二维语义标注结果,确定该三角面片的初始三维语义标注结果,包括:

针对每个所述三角面片,基于该三角面片在所述至少一张场景图像中对应的投影图像区域的初始二维语义标注结果,确定对应各种初始二维语义标注结果的标签数量;

基于对应各种初始二维语义标注结果的标签数量,确定该三角面片的初始三维语义标注结果。

5. 根据权利要求4所述的方法,其特征在于,所述基于对应各种初始二维语义标注结果的标签数量,确定该三角面片的初始三维语义标注结果,包括:

将标签数量最多的一种初始二维语义标注结果,确定为该三角面片的初始三维语义标注结果;

所述针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果,包括:

基于每个簇包括的各个三角面片的初始三维语义标注结果,确定对应各种初始三维语义标注结果的三角面片的数量;

将对应的三角面片的数量最多的一种初始三维语义标注结果,确定该簇的初始三维语义标注结果。

6. 根据权利要求4所述的方法,其特征在于,所述基于对应各种初始二维语义标注结果的标签数量,确定该三角面片的初始三维语义标注结果,包括:

将对应各种初始二维语义标注结果的标签数量与标签总数量进行比值运算,确定该三角面片属于每种初始二维语义标注结果的概率值;

基于属于各种初始二维语义标注结果的概率值,确定该三角面片的初始三维语义标注结果;

所述针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果,包括:

针对划分得到的每个簇,基于该簇包括的每个三角面片对应的属于每种初始二维语义标注结果的概率值,以及针对该三角面片预先分配的指向每种初始二维语义标注结果的权重,确定该簇属于每种初始二维语义标注结果的概率值;

基于该簇属于各种初始二维语义标注结果的概率值,确定该簇的初始三维语义标注结果。

7. 根据权利要求6所述的方法,其特征在于,在确定任一簇属于多种初始二维语义标注结果的概率值大于预设阈值的情况下,所述方法还包括:

针对任一簇,将该簇划分为多个子簇;

所述基于所述初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个所述簇的初始三维语义标注结果,包括:

基于所述初始二维语义标注结果,对划分得到的每个子簇进行三维语义标注,确定每个子簇的初始三维语义标注结果;

所述基于每个所述簇的三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

基于所述每个子簇的初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

8. 根据权利要求1-7任一所述的方法,其特征在于,所述基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

获取修正后三维语义标注结果;所述修正后三维语义标注结果为针对多个簇中至少部分簇的初始三维语义标注结果进行人工修正后的三维语义标注结果;

基于所述修正后三维语义标注结果,以及未经人工修正的其他簇的初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

9. 根据权利要求1-7任一所述的方法,其特征在于,所述基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

基于每个簇的初始三维语义标注结果,确定各个簇在语义标注层面对应的概率分布误

差项;以及基于任意相邻的不属于同一三维语义标注结果的两个所述簇之间的夹角信息,确定每个簇在几何特征层面对应的夹角误差项;

根据所述概率分布误差项和所述夹角误差项,确定能量函数;

确定使得所述能量函数的值最小的情况下,针对每个簇确定的更新后三维语义标注结果;

基于针对每个簇的更新后三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

10. 根据权利要求1-7任一所述的方法,其特征在于,所述基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

针对每个簇,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,从所述至少一张场景图像中确定与该簇包括的每个三角面片对应的场景图像;

针对每张场景图像,基于该张场景图像对应的三角面片所属簇的初始三维语义标注结果,确定针对该张场景图像的修正后二维语义标注结果。

11. 根据权利要求1-7任一所述的方法,其特征在于,所述初始二维语义标注结果为利用初始语义标注神经网络对所述每张场景图像进行语义标注得到的,所述确定针对所述每张场景图像的修正后二维语义标注结果之后,还包括:

基于所述每张场景图像的修正后二维语义标注结果对所述初始语义标注神经网络进行训练,得到训练好的语义标注神经网络。

12. 一种语义标注的装置,其特征在于,所述装置包括:

获取模块,用于获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

划分模块,用于从所述三维场景模型包括的各个三角面片中,随机选取预设数量个三角面片,将选取的每个三角面片的目标向量分别作为待划分的簇的中心向量;确定所述各个三角面片中每个三角面片的所述目标向量,分别与每个所述中心向量之间的距离,并将该三角面片划分至对应的所述距离最小的中心向量所在的簇;确定每个簇对应的新的中心向量,并基于该新的中心向量,返回执行划分簇的步骤,直到满足划分截止条件;

第一标注模块,用于针对每个三角面片,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,确定该三角面片在至少一张场景图像中对应的投影图像区域;以及,基于该三角面片在所述至少一张场景图像中对应的所述投影图像区域的所述初始二维语义标注结果,确定该三角面片的初始三维语义标注结果;针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果;

第二标注模块,用于基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

13. 一种电子设备,其特征在于,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权利要求1至11任一所述的语义标注的方法的步骤。

14. 一种计算机可读存储介质,其特征在于,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权利要求1至11任一所述的语义标注的方法的步骤。

## 一种语义标注的方法、装置、电子设备及存储介质

### 技术领域

[0001] 本公开涉及语义标注技术领域,具体而言,涉及一种语义标注的方法、装置、电子设备及存储介质。

### 背景技术

[0002] 语义分割技术是对图像像素级别的理解,是在图像上对物体进行像素级的分类,即将同属于同一类物体的像素归为一类,使用指定的标签(label)进行标记,被广泛应用于各种技术场景,如自动驾驶、室内导航、虚拟现实、图像识别等。

[0003] 相关技术中,主要是通过预先训练好的图像分割网络对视频图像进行语义分割,在对图像分割网络的训练过程中,需要获取大量带标注的样本图像。现有的语义标注方法主要采用人工方式对二维的视频图像进行逐像素点的标注,标注的效率较低。

### 发明内容

[0004] 本公开实施例至少提供一种语义标注的方法、装置、电子设备及存储介质。

[0005] 第一方面,本公开实施例提供了一种语义标注的方法,所述方法包括:

[0006] 获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

[0007] 将所述三维场景模型包括的各个三角面片划分为多个簇;

[0008] 基于所述初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个所述簇的初始三维语义标注结果;

[0009] 基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

[0010] 采用上述语义标注的方法,在对基于场景图像重建的三维场景模型进行簇的划分的情况下,可以基于每张场景图像对应的初始二维语义标注结果,对划分的每个簇进行三维语义标注,确定对应的初始三维语义标注结果,基于初始三维语义标注结果可以确定针对每张场景图像的修正后二维语义标注结果,也即,本公开实施例是基于三维空间内簇的初始三维语义标注结果对二维空间内场景图像的初始二维语义标注结果进行了修正,由于三维空间内一个簇指向的三角面片可能是多个,而每个三角面片所指向的场景图像可能是多个,基于上述对应关系可以对簇进行聚合性的标注,这将使得修正后二维语义标注结果更为准确,与此同时,上述修正过程自动完成,使得标注效率更高。

[0011] 在一种可能的实施方式中,所述将所述三维场景模型包括的各个三角面片划分为多个簇,包括:

[0012] 从所述三维场景模型包括的各个三角面片中,随机选取预设数量个三角面片,将选取的每个三角面片的目标向量分别作为待划分的簇的中心向量;

[0013] 确定所述各个三角面片中每个三角面片的所述目标向量,分别与每个所述中心向量之间的距离,并将该三角面片划分至对应的所述距离最小的中心向量所在的簇;

[0014] 确定每个簇对应的新的中心向量,并基于该新的中心向量,返回执行划分簇的步骤,直到满足划分截止条件。

[0015] 这里,可以基于三角面片的目标向量与各簇的中心向量实现至少一次簇的划分。每进行一次划分,均可以针对簇确定新的中心向量,从而进行下一次划分。由于划分到一个簇的多个三角面片很大程度上存在相关性,可以利用一个簇的多个三角面片所对应的标注结果来确定簇的初始三维语义标注结果,具有较高的准确性,从而进一步确保了后续修正后二维语义标注结果的准确性。

[0016] 在一种可能的实施方式中,所述确定每个簇对应的新的中心向量,包括:

[0017] 基于每个簇包括的各个三角面片的目标向量,确定各个三角面片的平均向量,将所述平均向量作为该簇对应的新的中心向量。

[0018] 在一种可能的实施方式中,所述目标向量包括目标法向量和目标颜色向量,所述中心向量包括中心法向量和中心颜色向量;所述确定所述各个三角面片中每个三角面片的所述目标向量,分别与每个所述中心向量之间的距离,包括:

[0019] 针对所述各个三角面片中的每个三角面片,基于该三角面片的目标法向量以及每个所述中心向量中的中心法向量,确定该三角面片与每个所述中心向量所在的簇之间的第一距离,以及,基于该三角面片的目标颜色向量以及每个所述中心向量中的中心颜色向量,确定该三角面片与每个所述中心向量所在的簇之间的第二距离;

[0020] 基于所述第一距离及其对应的权重、以及所述第二距离及其对应的权重,确定该三角面片的所述目标向量,分别与每个所述中心向量之间的距离。

[0021] 这里,考虑到三角面片的法向量以及颜色向量对于簇划分的影响程度并不相同,这里,可以基于加权求和方法实现上述有关三角面片的目标向量,分别与每个中心向量之间的距离的确定,适用性更强。

[0022] 在一种可能的实施方式中,所述基于所述初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个所述簇的初始三维语义标注结果,包括:

[0023] 针对每个三角面片,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,确定该三角面片在至少一张场景图像中对应的投影图像区域;以及,

[0024] 基于该三角面片在所述至少一张场景图像中对应的所述投影图像区域的所述初始二维语义标注结果,确定该三角面片的初始三维语义标注结果;

[0025] 针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果。

[0026] 这里,可以基于场景图像的初始二维语义标注结果,共同确定对应三角面片的初始三维语义标注结果,还可以基于一个簇包括的多个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果,可见,有关簇的初始三维语义标注结果的确定依赖于多个三角面片的三维标注结果,而每个三角面片的标注结果又依赖于各张场景图像的二维标注结果,这使得最终所确定的簇的初始三维语义标注结果更为准确。

[0027] 在一种可能的实施方式中,所述基于该三角面片在所述至少一张场景图像中对应的所述投影图像区域的所述初始二维语义标注结果,确定该三角面片的初始三维语义标注结果,包括:

[0028] 针对每个所述三角面片,基于该三角面片在所述至少一张场景图像中对应的投影图像区域的初始二维语义标注结果,确定对应各种初始二维语义标注结果的标签数量;

[0029] 基于对应各种初始二维语义标注结果的标签数量,确定为该三角面片的初始三维语义标注结果。

[0030] 在一种可能的实施方式中,所述基于对应各种初始二维语义标注结果的标签数量,确定该三角面片的初始三维语义标注结果,包括:

[0031] 将标签数量最多的一种初始二维语义标注结果,确定为该三角面片的初始三维语义标注结果;

[0032] 所述针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果,包括:

[0033] 基于每个簇包括的各个三角面片的初始三维语义标注结果,确定对应各种初始三维语义标注结果的三角面片的数量;

[0034] 将对应的三角面片的数量最多的一种初始三维语义标注结果,确定该簇的初始三维语义标注结果。

[0035] 在一种可能的实施方式中,所述基于对应各种初始二维语义标注结果的标签数量,确定该三角面片的初始三维语义标注结果,包括:

[0036] 将对应各种初始二维语义标注结果的标签数量与标签总数量进行比值运算,确定该三角面片属于每种初始二维语义标注结果的概率值;

[0037] 基于属于各种初始二维语义标注结果的概率值,确定该三角面片的初始三维语义标注结果;

[0038] 所述针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果,包括:

[0039] 针对划分得到的每个簇,基于该簇包括的每个三角面片对应的属于每种初始二维语义标注结果的概率值,以及针对该三角面片预先分配的指向每种初始二维语义标注结果的权重,确定该簇属于每种初始二维语义标注结果的概率值;

[0040] 基于该簇属于各种初始二维语义标注结果的概率值,确定该簇的初始三维语义标注结果。

[0041] 在一种可能的实施方式中,在确定任一簇属于多种初始二维语义标注结果的概率值大于预设阈值的情况下,所述方法还包括:

[0042] 针对任一簇,将该簇划分为多个子簇;

[0043] 所述基于所述初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个所述簇的初始三维语义标注结果,包括:

[0044] 基于所述初始二维语义标注结果,对划分得到的每个子簇进行三维语义标注,确定每个子簇的初始三维语义标注结果;

[0045] 所述基于每个所述簇的三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

[0046] 基于所述每个子簇的初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

[0047] 这里,在确定一个簇存在多种语义标注可能性的情况下,这里,可以先针对簇进行



子簇的划分,进而基于划分后的子簇实现语义标注,这将提高标注精度。

[0048] 在一种可能的实施方式中,所述基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

[0049] 获取修正后三维语义标注结果;所述修正后三维语义标注结果为针对所述多个簇中至少部分簇的初始三维语义标注结果进行人工修正后的三维语义标注结果;

[0050] 基于所述修正后三维语义标注结果,以及未经人工修正的其他簇的初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

[0051] 这里,考虑到人工修正对于后续有关修正后二维语义标注结果的准确性的关键影响,这里,可以基于至少部分簇的人工修正的修正后三维语义标注结果来确定针对每张场景图像的修正后二维语义标注结果,进一步提升标注准确性。

[0052] 在一种可能的实施方式中,所述基于所述每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

[0053] 基于每个簇的初始三维语义标注结果,确定各个簇在语义标注层面对应的概率分布误差项;以及基于任意相邻的不属于同一三维语义标注结果的两个所述簇之间的夹角信息,确定每个簇在几何特征层面对应的夹角误差项;

[0054] 根据所述概率分布误差项和所述夹角误差项,确定能量函数;

[0055] 确定使得所述能量函数的值最小的情况下,针对每个簇确定的更新后三维语义标注结果;

[0056] 基于针对每个簇的更新后三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

[0057] 这里,可以基于各个簇指向的概率分布误差项以及相邻簇指向的夹角误差项针对每张场景图像进行二维语义标注结果的修正,自动实现,可以进一步提升标注效率。

[0058] 在一种可能的实施方式中,所述基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果,包括:

[0059] 针对每个簇,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,从所述至少一张场景图像中确定与该簇包括的每个三角面片对应的场景图像;

[0060] 针对每张场景图像,基于该张场景图像对应的三角面片所属簇的初始三维语义标注结果,确定针对该张场景图像的修正后二维语义标注结果。

[0061] 在一种可能的实施方式中,所述初始二维语义标注结果为利用初始语义标注神经网络对所述每张场景图像进行语义标注得到的,所述确定针对所述每张场景图像的修正后二维语义标注结果之后,还包括:

[0062] 基于所述每张场景图像的修正后二维语义标注结果对所述初始语义标注神经网络进行训练,得到训练好的语义标注神经网络。

[0063] 第二方面,本公开实施例还提供了一种语义标注的装置,所述装置包括:

[0064] 获取模块,用于获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

[0065] 划分模块,用于将所述三维场景模型包括的各个三角面片划分为多个簇;

[0066] 第一标注模块,用于基于所述初始二维语义标注结果,对划分得到的每个簇进行

三维语义标注,确定每个所述簇的初始三维语义标注结果;

[0067] 第二标注模块,用于基于每个所述簇的所述初始三维语义标注结果,确定针对所述每张场景图像的修正后二维语义标注结果。

[0068] 第三方面,本公开实施例还提供了一种电子设备,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第一方面及其各种实施方式任一所述的语义标注的方法的步骤。

[0069] 第四方面,本公开实施例还提供了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第一方面及其各种实施方式任一所述的语义标注的方法的步骤。

[0070] 关于上述语义标注的装置、电子设备、及计算机可读存储介质的效果描述参见上述语义标注的方法的说明,这里不再赘述。

[0071] 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

[0072] 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。

[0073] 图1示出了本公开实施例所提供的一种语义标注的方法的流程图;

[0074] 图2示出了本公开实施例所提供的一种语义标注的装置的示意图;

[0075] 图3示出了本公开实施例所提供的一种电子设备的示意图。

#### 具体实施方式

[0076] 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。

[0077] 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定义和解释。

[0078] 本文中术语“和/或”,仅仅是描述一种关联关系,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,本文中术语“至少一种”表示多种中的任意一种或多种中的至少两种的任意组合,例如,包括A、B、C中的至少一种,可以表示包括从A、B和C构成的集合中选择的任意一个或多个元素。

[0079] 经研究发现,相关技术中,主要是通过预先训练好的图像分割网络对视频图像进行语义分割,在对图像分割网络的训练过程中,需要获取大量带标注的样本图像。现有的语义标注方法主要采用人工方式对二维的视频图像进行逐像素点的标注,标注的效率较低。

[0080] 基于上述研究,本公开提供了一种语义标注的方法、装置、电子设备及存储介质,以提升标注效率。

[0081] 为便于对本实施例进行理解,首先对本公开实施例所公开的一种语义标注的方法进行详细介绍,本公开实施例所提供的语义标注的方法的执行主体一般为具有一定计算能力的计算机设备,该计算机设备例如包括:终端设备或服务器或其它处理设备,终端设备可以为用户设备(User Equipment, UE)、移动设备、用户终端、蜂窝电话、无绳电话、个人数字助理(Personal Digital Assistant, PDA)、手持设备、计算设备、车载设备、可穿戴设备等。在一些可能的实现方式中,该语义标注的方法可以通过处理器调用存储器中存储的计算机可读指令的方式来实现。

[0082] 参见图1所示,为本公开实施例提供的语义标注的方法的流程图,方法包括步骤S101~S104,其中:

[0083] S101:获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

[0084] S102:将三维场景模型包括的各个三角面片划分为多个簇;

[0085] S103:基于初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个簇的初始三维语义标注结果;

[0086] S104:基于每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0087] 为了便于理解本公开实施例提供的语义标注的方法,首先对该方法的应用场景进行简单说明。上述语义标注的方法可以应用于任何需要进行语义标注的场景中,例如,可以是自动驾驶场景中对于车辆、行人等的标注,还可以是室内导航中对于障碍物等的标注,还可以是其它场景下的标注,本公开实施例对此不做具体的限制。

[0088] 相关技术中采用的是逐像素点的语义标注方式,该标注方式人工完成。考虑到在进行网络训练的过程中,往往需要大量标注好的样本图像,导致耗费大量的人力物力。

[0089] 正是为了解决上述问题,本公开实施例才提供了一种自动进行语义标注的语义标注的方案,可以在三维空间内实现语义标注,然后基于三维空间与二维空间之间的转换关系,将三维空间内的语义标注结果投影到二维空间内,这可以大大提高标注效率。

[0090] 为了实现在三维空间内的语义标注,这里可以首先获取基于一张或多张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果。

[0091] 其中,上述三维场景模型可以是基于场景图像进行三维重建确定的。这里的三维重建一般可以分为三个步骤:稀疏重建、密集重建、表面重建。其中,稀疏重建用于完成所有相机位姿初始化计算,其输入可以是各张场景图像,输出的可以拍摄场景图像的相机的位姿;密集重建用于在相机位姿已知的前提下,逐像素的计算图像中每一个像素点对应的三维点,得到场景物体表面密集的三维点云,其输入可以是多视角的场景图像和相机位姿,输入可以是稠密点云;表面重建用于完成由稠密点云到三维场景模型的转换。

[0092] 在具体应用中,可以采用colmap或者其它方式实现上述三维重建的过程。

[0093] 另外,本公开实施例中有场景图像的初始二维语义标注结果可以是利用初始语义标注神经网络标注得到的。基于有关场景图像的初始二维语义标注结果、以及三维空间与三维空间之间的坐标转换关系,可以实现针对三维场景模型包括的三角面片的三维语义标注。

[0094] 上述初始语义标注神经网络可以是利用现有数据库中存储的已标注样本图像训练得到的,训练的可以是已标注样本图像与标注信息之间的对应关系。

[0095] 本公开实施例中,三维场景模型可以是由三角面片拼接而成的,也即,三角面片可以是三维场景模型的最小组成单元。在实际应用中,考虑到不同三角面片所具备的语义可以是相同的,还可以是不同的,这里,可以按照语义对三角面片进行簇的划分,以便可以对簇内属于同一语义的三角面片进行统一的语义标注,从而可以在确保了标注准确性的前提下,提升标注效率。

[0096] 本公开实施例中有簇的初始三维语义标注结果可以是基于上述有关场景图像的初始二维语义标注结果来确定的。这主要是考虑到一个簇与多个三角面片之间的对应关系,一个三角面片与多张场景图像之间的对应关系,这样,在确定场景图像的初始二维语义标注结果的情况下,可以基于上述对应关系确定簇的初始三维语义标注结果。

[0097] 本公开实施例中,在确定每个簇的初始三维语义标注结果的情况下,可以基于三维空间与二维空间的坐标转换关系,一次性的确定所有场景图像的修正后二维语义标注结果,相比二维空间下的人工标注方式,很大程度上提高了标注效率。

[0098] 本公开实施例中,基于场景图像的修正后二维语义标注结果可以对初始语义标注神经网络进行训练。由于修正后二维语义标注结果的准确性更高,因而所训练得到的语义标注神经网络的标注准确性更好,便于后续进行各种场景下的应用。

[0099] 考虑到簇的划分对于后续进行语义标注的关键作用,接下来可以对划分簇的方法进行详细描述。上述有关簇的划分方法具体包括如下步骤:

[0100] 步骤一、从三维场景模型包括的各个三角面片中,随机选取预设数量个三角面片,将选取的每个三角面片的目标向量分别作为待划分的簇的中心向量;

[0101] 步骤二、确定各个三角面片中每个三角面片的目标向量,分别与每个中心向量之间的距离,并将该三角面片划分至对应的距离最小的中心向量所在的簇;

[0102] 步骤三、确定每个簇对应的新的中心向量,并基于该新的中心向量,返回执行划分簇的步骤,直到满足划分截止条件。

[0103] 这里的目标向量包括目标法向量和/或目标颜色向量,可以从不同维度表征三角面片的特征。有关目标法向量和目标颜色向量可以是基于三维重建结果确定的。

[0104] 本公开实施例可以基于各簇的中心向量与各三角面片的目标向量之间的距离进行至少一次划分操作,直到满足划分截止条件可以得到最终的划分结果。

[0105] 在每次划分完成后,可以确定簇对应的心的中心向量,并基于这一中心向量进行下一次划分。有关簇的中心向量可以是该簇的各个三角面片的目标向量所确定的平均向量,也即,可以是对簇中所有三角面片的目标向量求取平均值。

[0106] 本公开实施例中的划分截止条件可以是针对一个簇,多次计算的中心向量不再发生变化,还可以是划分次数达到预设次数,还可以是其它划分截止条件,本公开实施例对此不做具体的限制。

[0107] 考虑到法向量和颜色向量对于簇划分的影响程度并不相同,因而,这里可以基于两种向量确定两个距离,并结合加权求和方式确定上述距离。

[0108] 具体可以通过如下步骤来实现:

[0109] 步骤一、针对各个三角面片中的每个三角面片,基于该三角面片的目标法向量以及每个中心向量中的中心法向量,确定该三角面片与每个中心向量所在的簇之间的第一距离,以及,基于该三角面片的目标颜色向量以及每个中心向量中的中心颜色向量,确定该三角面片与每个中心向量所在的簇之间的第二距离;

[0110] 步骤二、基于第一距离及其对应的权重、以及第二距离及其对应的权重,确定该三角面片的目标向量,分别与每个中心向量之间的距离。

[0111] 这里,可以依据不同的应用场景对上述第一距离对应的权重、以及第二距离对应的权重进行设置。例如,针对一些单一色彩的场景而言,颜色向量与中心颜色向量之间的一致性对于簇的划分至关重要,这时可以分配更高的权重。

[0112] 本公开实施例提供的划分方法旨在使得相同语义的三角面片属于同一个簇,不同语义的三角面片属于不同的簇。这里,如一个簇的三角面片的目标向量由 $R_i$ 表示, $R_i$ 包括目标法向量 $n(x)$ 和中心颜色向量 $c(x)$ ,簇的中心向量为 $P_i$ , $P_i$ 包括中心法向量 $n_i$ 和中心颜色向量 $c_i$ 。为了实现上述技术目的,这里是找到一组簇可以使得如下公式所示的误差最小。

$$[0113] \quad E(R, P) = \sum_{i=1 \dots k} E(R_i, P_i)$$

$$[0114] \quad \text{其中,误差函数 } E(R_i, P_i) = \iint_{x \in R_i} \|n(x) - n_i\|^2 + \|c(x) - c_i\|^2 dx。$$

[0115] 本公开实施例中,可以是基于场景图像对应的初始二维语义标注结果确定针对簇的初始三维语义标注结果,具体包括如下步骤:

[0116] 步骤一、针对每个三角面片,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,确定该三角面片在至少一张场景图像中对应的投影图像区域;以及,

[0117] 步骤二、基于该三角面片在至少一张场景图像中对应的投影图像区域的初始二维语义标注结果,确定该三角面片的初始三维语义标注结果;

[0118] 步骤三、针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果。

[0119] 这里,考虑到三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,可以先针对每个三角面片,确定该三角面片在各张场景图像中对应的投影图像区域。

[0120] 这里的投影图像区域可以是根据三角面片的中心坐标、图像位姿、内参等数据投影到场景图像得到的各个像素点的集合。由于各张场景图像的初始二维语义标注结果是预先标注得到的,这样,场景图像对应的投影图像区域的初始二维语义标注结果也可以确定。

[0121] 在针对三角面片确定其投影图像区域的初始二维语义标注结果的情况下,可以确定对应各种初始二维语义标注结果的标签数量,也即,可以得到三角面片分别属于某一标签类的票数,将各标签类的票数归一化可以得到三角面片属于某一标签类的语义概率,这里,可以选取语义概率最大,也即标签数量最多的初始二维语义标注结果,作为三角面片的

初始三维语义标注结果。

[0122] 这里的簇是基于三角面片组合而成的,这里,针对每个簇可以基于簇包括的各个三角面片的初始三维语义标注结果,确定对应各种初始三维语义标注结果的三角面片的数量,三角面片的数量越多的初始三维语义标注结果,对应为簇的初始三维语义标注结果的可能性也越高。这里,可以选取对应的三角面片的数量最多的初始三维语义标注结果,作为簇的初始三维语义标注结果。

[0123] 除此之外,本公开实施例还可以按照如下步骤确定簇的初始三维语义标注结果:

[0124] 步骤一、将对应各种初始二维语义标注结果的标签数量与标签总数量进行比值运算,确定该三角面片属于每种初始二维语义标注结果的概率值;

[0125] 步骤二、基于属于各种初始二维语义标注结果的概率值,确定该三角面片的初始三维语义标注结果;

[0126] 步骤三、针对划分得到的每个簇,基于该簇包括的每个三角面片对应的属于每种初始二维语义标注结果的概率值,以及针对该三角面片预先分配的指向每种初始二维语义标注结果的权重,确定该簇属于每种初始二维语义标注结果的概率值;

[0127] 步骤四、基于该簇属于各种初始二维语义标注结果的概率值,确定该簇的初始三维语义标注结果。

[0128] 这里,可以先对三角面片进行属于每种初始二维语义标注结果的概率值的确定,而后再结合一个簇包含的三角面片实现针对每个簇的初始三维语义标注结果的确定。

[0129] 其中,有关一个三角面片属于每种初始二维语义标注结果的概率值是基于标签数量的比值实现的。这里,可以统计一个三角面片在指向各张各张场景图像中对应的投影图像区域的过程中,所确定的针对每种初始二维语义标注结果的投票数量(对应标签数量),还可以统计所有投票数量(对应标签总数量),通过两个数量的比值可以确定上述概率值。

[0130] 在针对一个簇进行概率计算的过程中,可以基于该簇包括的每个三角面片预先分配的指向每种初始二维语义标注结果的权重来完成。

[0131] 在具体应用中,上述权重可以是结合三角面片的面积来确定,面积越大的三角面片可以对应分配较大的权重,面积越小的三角面片可以对应分配较小的权重,这主要是考虑到大面积三角面片在投票过程往往占据更为关键的作用。这里,通过加权求和即可以确定出一个簇属于各种初始二维语义标注结果的概率值,一个簇的初始三维语义标注结果则可以看作是一个概率分布。

[0132] 在实际应用中,可能会出现无法直接对簇进行语义标注的情况。为了解决这一问题,本公开实施例可以对簇再进行进一步划分,通过对划分的子簇的标注实现最终的标注效果。

[0133] 这里,有关子簇的划分方式参见上述簇的划分方式,在此不再赘述。

[0134] 考虑到人工修正方式对于标签准确性的强影响性,这里,可以先获取针对多个簇中至少部分簇的初始三维语义标注结果进行人工修正的修正后三维语义标注结果,而后基于获取的修正后三维语义标注结果,以及未经人工修正的其他簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0135] 本公开实施例可以基于web的标注器进行人工修正,也即,可以将各个簇的初始三维语义标注结果加载到基于web的标注器中,通过这一标注工具手动修正。

[0136] 例如,可以是在确定串联的多个簇中存在不同语义的目标簇的情况下,可以调取这一目标簇,并在确定针对目标簇标注错误的情况下,手动修正到正确的标注结果。这里,修正的簇所属类别的概率设置为1,其它类别的概率为0。

[0137] 需要说明的是,人工修正标注结果的情况下,有关簇以及簇内的三角面片的语义标注结果也将随之发生变化。

[0138] 本公开实施例中,不仅可以基于人工修正方式进行簇的三维语义标注结果的修正,还可以基于预先定义的能量函数来更新簇的三维语义标注结果,以使得所确定的针对场景图像的修正后二维语义标注结果更为准确。

[0139] 其中,上述三维语义标注结果的更新过程具体包括如下步骤:

[0140] 步骤一、基于每个簇的初始三维语义标注结果,确定各个簇在语义标注层面对应的概率分布误差项;以及基于任意相邻的不属于同一三维语义标注结果的两个簇之间的夹角信息,确定每个簇在几何特征层面对应的夹角误差项;

[0141] 步骤二、根据概率分布误差项和夹角误差项,确定能量函数;

[0142] 步骤三、确定使得能量函数的值最小的情况下,针对每个簇确定的更新后三维语义标注结果。

[0143] 为了便于理解上述更新过程,接下来可以结合如下能量函数公式来具体说明。

$$[0144] \quad E(\bar{x}) = \sum_{f \in F} e_1(f, x_f) + \lambda \sum_{\{f, g\} \in N} e_2(x_f, x_g)$$

$$[0145] \quad \text{其中,} \begin{cases} e_1(f, x_f) = -\log(\max(P(f | x_f), \epsilon)) \\ e_2(x_f, x_g) = \begin{cases} -\log(\theta(f, g) / \pi) & x_f \neq x_g \\ 0 & x_f = x_g \end{cases} \end{cases} \circ$$

[0146] 其中,上述 $e_1(\cdot)$ 和 $e_2(\cdot)$ 分别对应概率分布误差项和夹角误差项,F表示所有簇的集合,N表示相邻簇的集合, $x_f$ 表示簇f所属的三维语义标注结果, $P(f | x_f)$ 表示簇f属于三维语义标注结果 $x_f$ 的概率, $\theta(f, g)$ 表示簇f和簇g的夹角。

[0147] 需要说明的是,针对概率分布误差项而言,这里除了可以考虑簇的初始三维语义标注结果,还可以考虑人工修正后的标注结果,例如可以给予人工修正的标注结果以最大的语义概率。

[0148] 通过最小化上述能量函数,可以确定每个簇的更新后三维语义标注结果。

[0149] 本公开实施例中,可以基于每个簇的更新后三维语义标注结果实现针对场景图像的二维语义标注结果的修正。

[0150] 这里,首先基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,可以针对每个,从至少一张场景图像中确定与该簇包括的每个三角面片对应的场景图像,而后针对每张场景图像,基于该张场景图像对应的三角面片所属簇的更新后三维语义标注结果,确定针对该张场景图像的修正后二维语义标注结果。也即,本公开实施例可以将带有三维语义的三维场景模型渲染到二维,从而得到场景图像对应的二维语义标注,自动实现,效率和准确性均较高。

[0151] 其中,上述由三维到二维的渲染过程可以是基于opengl实现的,还可以是基于其它方式实现的,在此不做具体的限制。

[0152] 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。

[0153] 基于同一发明构思,本公开实施例中还提供了与语义标注的方法对应的语义标注的装置,由于本公开实施例中的装置解决问题的原理与本公开实施例上述语义标注的方法相似,因此装置的实施可以参见方法的实施,重复之处不再赘述。

[0154] 参照图2所示,为本公开实施例提供的一种语义标注的装置的示意图,装置包括:获取模块201、划分模块202、第一标注模块203和第二标注模块204;其中,

[0155] 获取模块201,用于获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

[0156] 划分模块202,用于将三维场景模型包括的各个三角面片划分为多个簇;

[0157] 第一标注模块203,用于基于初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个簇的初始三维语义标注结果;

[0158] 第二标注模块204,用于基于确定的每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0159] 本公开实施例中,在对基于场景图像重建的三维场景模型进行簇的划分的情况下,可以基于每张场景图像对应的初始二维语义标注结果,对划分的每个簇进行三维语义标注,确定对应的初始三维语义标注结果,基于初始三维语义标注结果可以确定针对每张场景图像的修正后二维语义标注结果,也即,本公开实施例是基于三维空间内簇的初始三维语义标注结果对二维空间内场景图像的初始二维语义标注结果进行了修正,由于三维空间内一个簇指向的三角面片可能是多个,而每个三角面片所指向的场景图像可能是多个,基于上述对应关系可以对簇进行聚合性的标注,这将使得修正后二维语义标注结果更为准确,与此同时,上述修正过程自动完成,使得标注效率更高。

[0160] 在一种可能的实施方式中,划分模块202,用于按照如下步骤将三维场景模型包括的各个三角面片划分为多个簇:

[0161] 从三维场景模型包括的各个三角面片中,随机选取预设数量个三角面片,将选取的每个三角面片的目标向量分别作为待划分的簇的中心向量;

[0162] 确定各个三角面片中每个三角面片的目标向量,分别与每个中心向量之间的距离,并将该三角面片划分至对应的距离最小的中心向量所在的簇;

[0163] 确定每个簇对应的新的中心向量,并基于该新的中心向量,返回执行划分簇的步骤,直到满足划分截止条件。

[0164] 在一种可能的实施方式中,划分模块202,用于按照如下步骤确定每个簇对应的新的中心向量:

[0165] 基于每个簇包括的各个三角面片的目标向量,确定各个三角面片的平均向量,将平均向量作为该簇对应的新的中心向量。

[0166] 在一种可能的实施方式中,目标向量包括目标法向量和目标颜色向量,中心向量包括中心法向量和中心颜色向量;划分模块202,用于按照如下步骤确定各个三角面片中每个三角面片的目标向量,分别与每个中心向量之间的距离:

[0167] 针对各个三角面片中的每个三角面片,基于该三角面片的目标法向量以及每个中



心向量中的中心法向量,确定该三角面片与每个中心向量所在的簇之间的第一距离,以及,基于该三角面片的目标颜色向量以及每个中心向量中的中心颜色向量,确定该三角面片与每个中心向量所在的簇之间的第二距离;

[0168] 基于第一距离及其对应的权重、以及第二距离及其对应的权重,确定该三角面片的目标向量,分别与每个中心向量之间的距离。

[0169] 在一种可能的实施方式中,第一标注模块203,用于按照以下步骤基于初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个簇的初始三维语义标注结果:

[0170] 针对每个三角面片,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,确定该三角面片在至少一张场景图像中对应的投影图像区域;以及,

[0171] 基于该三角面片在至少一张场景图像中对应的投影图像区域的初始二维语义标注结果,确定该三角面片的初始三维语义标注结果;

[0172] 针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果。

[0173] 在一种可能的实施方式中,第一标注模块203,用于按照以下步骤基于该三角面片在至少一张场景图像中对应的投影图像区域的初始二维语义标注结果,确定该三角面片的初始三维语义标注结果:

[0174] 针对每个三角面片,基于该三角面片在至少一张场景图像中对应的投影图像区域的初始二维语义标注结果,确定对应各种初始二维语义标注结果的标签数量;

[0175] 基于对应各种初始二维语义标注结果的标签数量,确定为该三角面片的初始三维语义标注结果。

[0176] 在一种可能的实施方式中,第一标注模块203,用于按照以下步骤针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果:

[0177] 将标签数量最多的一种初始二维语义标注结果,确定为该三角面片的初始三维语义标注结果;

[0178] 基于每个簇包括的各个三角面片的初始三维语义标注结果,确定对应各种初始三维语义标注结果的三角面片的数量;

[0179] 将对应的三角面片的数量最多的一种初始三维语义标注结果,确定该簇的初始三维语义标注结果。

[0180] 在一种可能的实施方式中,第一标注模块203,用于按照以下步骤针对划分得到的每个簇,基于该簇包括的每个三角面片的初始三维语义标注结果,确定该簇的初始三维语义标注结果:

[0181] 将对应各种初始二维语义标注结果的标签数量与标签总数量进行比值运算,确定该三角面片属于每种初始二维语义标注结果的概率值;

[0182] 基于属于各种初始二维语义标注结果的概率值,确定该三角面片的初始三维语义标注结果;

[0183] 针对划分得到的每个簇,基于该簇包括的每个三角面片对应的属于每种初始二维

语义标注结果的概率值,以及针对该三角面片预先分配的指向每种初始二维语义标注结果的权重,确定该簇属于每种初始二维语义标注结果的概率值;

[0184] 基于该簇属于各种初始二维语义标注结果的概率值,确定该簇的初始三维语义标注结果。

[0185] 在一种可能的实施方式中,在确定任一簇属于多种初始二维语义标注结果的概率值大于预设阈值的情况下,上述划分模块202还用于:

[0186] 针对任一簇,将该簇划分为多个子簇;

[0187] 第一标注模块203,用于按照以下步骤基于初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个簇的初始三维语义标注结果:

[0188] 基于初始二维语义标注结果,对划分得到的每个子簇进行三维语义标注,确定每个子簇的初始三维语义标注结果;

[0189] 第二标注模块204,用于按照以下步骤基于确定的每个簇的三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果:

[0190] 基于每个子簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0191] 在一种可能的实施方式中,第二标注模块204,用于按照以下步骤基于每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果:

[0192] 获取修正后三维语义标注结果;修正后三维语义标注结果为针对多个簇中至少部分簇的初始三维语义标注结果进行人工修正后的三维语义标注结果;

[0193] 基于的修正后三维语义标注结果,以及未经人工修正的其他簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0194] 在一种可能的实施方式中,第二标注模块204,用于按照以下步骤基于每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果:

[0195] 基于每个簇的初始三维语义标注结果,确定各个簇在语义标注层面对应的概率分布误差项;以及基于任意相邻的不属于同一三维语义标注结果的两个簇之间的夹角信息,确定每个簇在几何特征层面对应的夹角误差项;

[0196] 根据概率分布误差项和夹角误差项,确定能量函数;

[0197] 确定使得能量函数的值最小的情况下,针对每个簇确定的更新后三维语义标注结果;

[0198] 基于针对每个簇的更新后三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0199] 在一种可能的实施方式中,第二标注模块204,用于按照以下步骤基于每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果:

[0200] 针对每个簇,基于三维场景模型对应的第一坐标系以及场景图像对应的第二坐标系之间的转换关系,从至少一张场景图像中确定与该簇包括的每个三角面片对应的场景图像;

[0201] 针对每张场景图像,基于该张场景图像对应的三角面片所属簇的初始三维语义标注结果,确定针对该张场景图像的修正后二维语义标注结果。

[0202] 在一种可能的实施方式中,初始二维语义标注结果为利用初始语义标注神经网络

对每张场景图像进行语义标注得到的,上述装置还包括:

[0203] 训练模块205,用于确定针对每张场景图像的修正后二维语义标注结果之后,基于每张场景图像的修正后二维语义标注结果对初始语义标注神经网络进行训练,得到训练好的语义标注神经网络。

[0204] 关于装置中的各模块的处理流程、以及各模块之间的交互流程的描述可以参照上述方法实施例中的相关说明,这里不再详述。

[0205] 本公开实施例还提供了一种电子设备,如图3所示,为本公开实施例提供的电子设备结构示意图,包括:处理器301、存储器302、和总线303。存储器302存储有处理器301可执行的机器可读指令(比如,图2中的装置中获取模块201、划分模块202、第一标注模块203、第二标注模块204对应的执行指令等),当电子设备运行时,处理器301与存储器302之间通过总线303通信,机器可读指令被处理器301执行时执行如下处理:

[0206] 获取基于至少一张场景图像重建的三维场景模型、以及每张场景图像的初始二维语义标注结果;

[0207] 将三维场景模型包括的各个三角面片划分为多个簇;

[0208] 基于初始二维语义标注结果,对划分得到的每个簇进行三维语义标注,确定每个簇的初始三维语义标注结果;

[0209] 基于每个簇的初始三维语义标注结果,确定针对每张场景图像的修正后二维语义标注结果。

[0210] 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语义标注的方法的步骤。其中,该存储介质可以是易失性或非易失的计算机可读取存储介质。

[0211] 本公开实施例还提供一种计算机程序产品,该计算机程序产品承载有程序代码,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语义标注的方法的步骤,具体可参见上述方法实施例,在此不再赘述。

[0212] 其中,上述计算机程序产品可以具体通过硬件、软件或其结合的方式实现。在一个可选实施例中,所述计算机程序产品具体体现为计算机存储介质,在另一个可选实施例中,计算机程序产品具体体现为软件产品,例如软件开发包(Software Development Kit,SDK)等等。

[0213] 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。

[0214] 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。

[0215] 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。

[0216] 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)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。

[0217] 最后应说明的是:以上所述实施例,仅为本公开的具体实施方式,用以说明本公开的技术方案,而非对其限制,本公开的保护范围并不局限于此,尽管参照前述实施例对本公开进行了详细的说明,本领域的普通技术人员应当理解: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,其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化,或者对其中部分技术特征进行等同替换;而这些修改、变化或者替换,并不使相应技术方案的本质脱离本公开实施例技术方案的精神和范围,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所述以权利要求的保护范围为准。

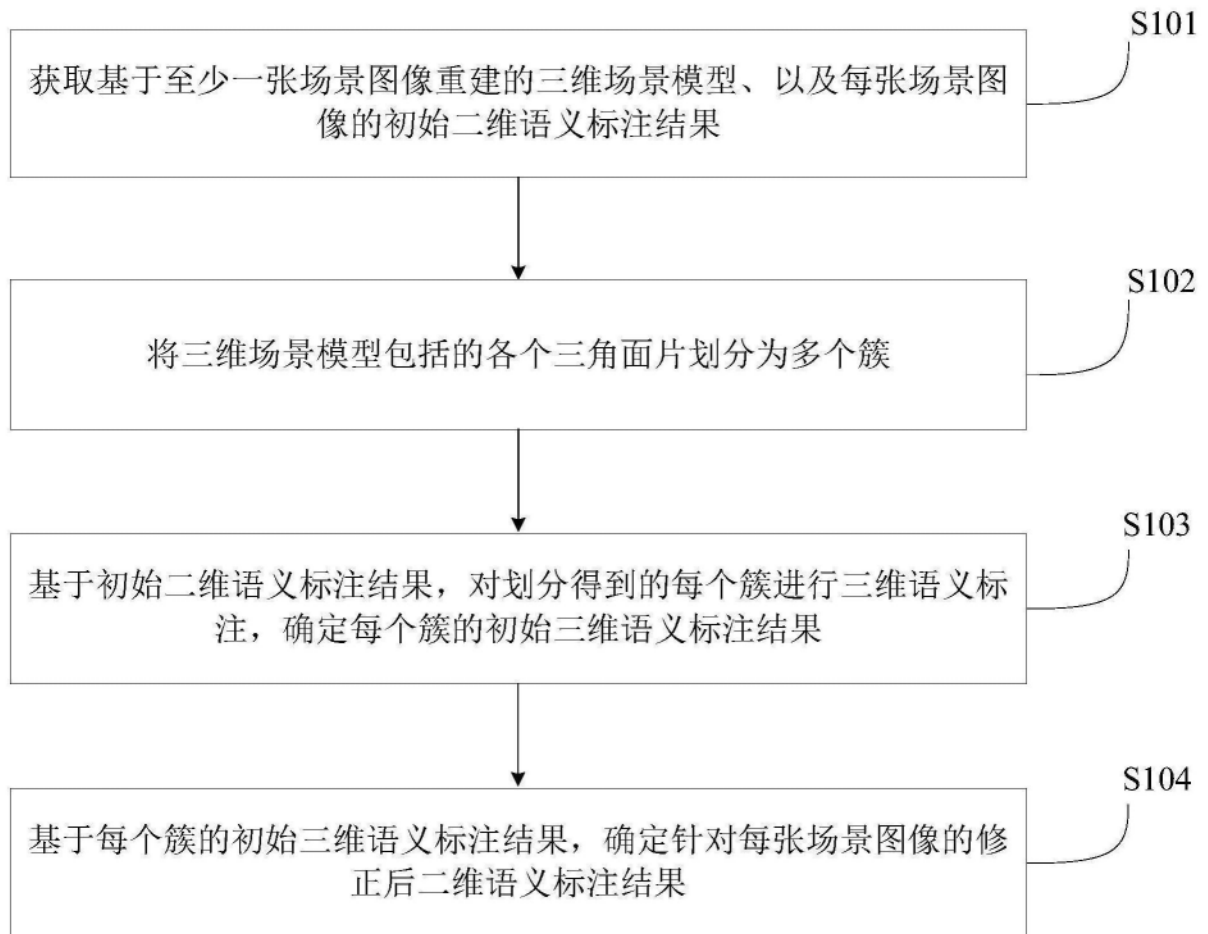


图1

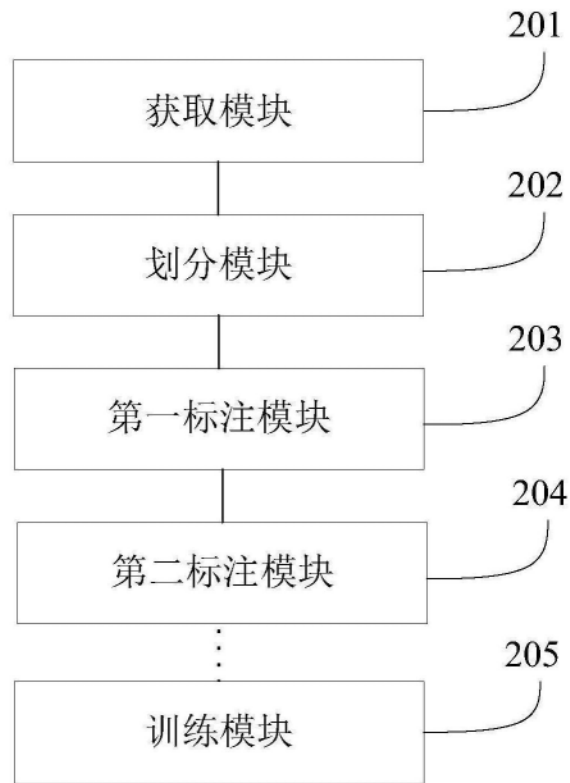


图2

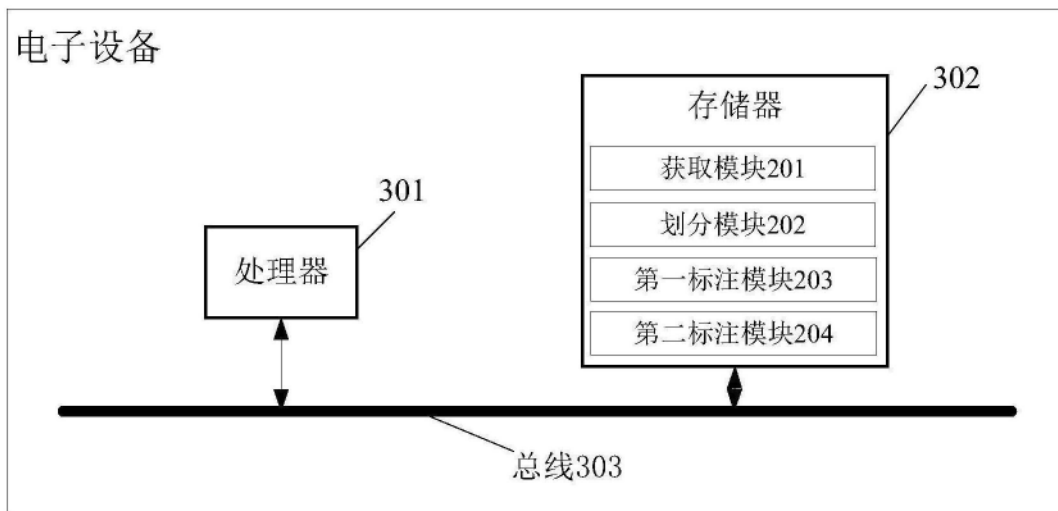


图3