



(12)发明专利申请

(10)申请公布号 CN 108596974 A

(43)申请公布日 2018.09.28

(21)申请号 201810298042.6

(22)申请日 2018.04.04

(71)申请人 清华大学

地址 100084 北京市海淀区清华园

(72)发明人 刘辛军 于超 乔飞 谢福贵
王智杰

(74)专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201

代理人 张润

(51)Int.Cl.

G06T 7/73(2017.01)

G06T 17/00(2006.01)

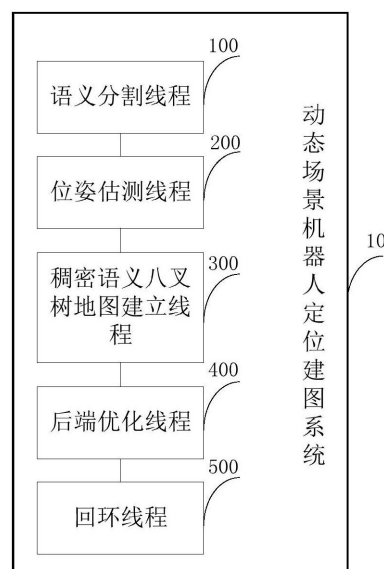
权利要求书1页 说明书9页 附图7页

(54)发明名称

动态场景机器人定位建图系统及方法

(57)摘要

本发明公开了一种动态场景机器人定位建图系统与方法,其中,系统包括:语义分割线程用于根据RGB图像获取语义分割结果,生成带有语义信息的语义图像;位姿估测线程用于根据RGB图像获取ORB特征点,通过运动型一致性检测获得运动区域,结合语义图像的语义信息剔除动态目标上的外点,并通过匹配剩余稳定目标的特征点得到变换矩阵,生成语义图像关键帧;稠密语义八叉树地图建立线程用于根据语义图像关键帧生成带语义标签的3D稠密八叉树地图。该系统通过将语义分割网络与运动一致性检测法相结合,完成对动态物体的剔除,并且构建的语义稠密八叉树地图,能够服务于更高层次的工作,从而有效提高了定位建图的适用性、实用性和鲁棒性。



1. 一种动态场景机器人视觉定位建图系统,其特征在于,包括:

语义分割线程,用于采用深度学习技术实时的获取语义分割结果,以生成带有语义信息的语义图像;

位姿估测线程,用于根据RGB图像获取ORB特征点,通过运动一致性检测获得运动区域,结合语义分割的信息剔除动态区域上的外点,并通过匹配剩余稳定目标的特征点得到变换矩阵,生成语义图像关键帧;

稠密语义八叉树地图线程,用于根据所述语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,完成动态目标的剔除,可用于高级复杂任务;

后端优化线程,用于对所述关键帧进行局部优化和关键帧筛选;

回环线程,用于对局部优化和关键帧筛选的结果进行整体优化和回环检测。

2. 根据权利要求1所述的动态场景机器人定位建图系统,其特征在于,所述语义分割线程与所述位姿估测线程进行数据交流,与运动一致性检测相结合剔除环境中的动态目标。

3. 根据权利要求1所述的动态场景机器人定位建图系统,其特征在于,所述稠密语义八叉树地图建立线程还用于通过概率融合法建立稳定的3D八叉树地图,将动态目标剔除,并与语义信息相融合,该地图可以用于导航以及机器人完成高级复杂任务。

4. 一种动态场景机器人定位建图方法,其特征在于,包括以下步骤:

采用深度学习技术实时的获取语义分割结果,以生成带有语义信息的语义图像;

根据RGB图像获取ORB特征点,通过运动一致性检测获得运动区域,结合语义分割的信息剔除动态区域上的外点,并通过匹配剩余稳定目标的特征点得到变换矩阵,生成语义图像关键帧;

根据所述语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,完成动态目标的剔除,可用于高级复杂任务;

对所述关键帧进行局部优化和关键帧筛选;

对局部优化和关键帧筛选的结果进行整体优化和回环检测。

动态场景机器人定位建图系统及方法

技术领域

[0001] 本发明涉及机器人定位建图技术领域,特别涉及一种动态场景机器人定位建图系统及方法。

背景技术

[0002] 机器人定位建图系统主要依赖SLAM(Simultaneous Localization and Mapping, 实时定位与地图构建)技术,该技术被视作移动智能机器人的一项重要技术,最早于1988年提出。机器人通过传感器获取周围环境的相关数据进行地图构建,然后将当前所感知到的环境特征与地图中的特征进行匹配,从而完成定位。

[0003] 根据传感器类型的不同,SLAM技术主要可以分为激光SLAM和视觉SLAM。由于图像在信息存储上的丰富性,以及图像对于一些更高层次工作(如语义分割与物体检测)的服务型,视觉SLAM技术在近年来被广泛研究。此外,当下的视觉SLAM技术往往都是一个完整的架构,包含了特征提取、回环检测等部分,如ORB-SLAM2, LSD-SLAM等现有的视觉SLAM技术已经在某些环境下取得了较好的试验结果。

[0004] 但在现有的机器人定位建图系统中,所应用的SLAM技术仍然存在着一些问题。一方面是如何让机器人的定位系统能够适应复杂多变的环境,例如环境中存在动态物体;另一方面是如何依靠SLAM技术构建更有价值的地图,提供更多的信息来服务于系统的高层次的应用。

发明内容

[0005] 本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

[0006] 为此,本发明的一个目的在于提出一种动态场景机器人定位建图系统,该系统可以有效提高了定位建图的适用性、实用性和鲁棒性。

[0007] 本发明的另一个目的在于提出一种动态场景机器人定位建图方法。

[0008] 为达到上述目的,本发明一方面实施例提出了一种动态场景机器人定位建图系统,为5线程架构,包括:语义分割线程,采用深度学习技术实时的获取语义分割结果,以生成带有语义信息的语义图像;位姿估测线程,根据RGB图像获取ORB特征点,通过运动一致性检测获得运动区域,结合语义分割的信息剔除动态区域上的外点,并通过匹配剩余稳定目标的特征点得到变换矩阵,生成语义图像关键帧;稠密语义八叉树地图线程,用于根据所述语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,完成动态目标的剔除,可用于高级复杂任务。后端优化线程,用于对所述关键帧进行局部优化和关键帧筛选;回环线程,用于对局部优化和关键帧筛选的结果进行整体优化和回环检测。

[0009] 本发明实施例的动态场景机器人定位建图系统,通过将语义分割网络与运动一致性检测法相结合,并对动态物体的剔除,减少了环境中运动物体对实时定位与地图构建的影响,提高了在动态环境下的定位精度,并且构建的稠密八叉树地图,能够服务于更高层次的工作,从而有效提高了定位建图的适用性、实用性和鲁棒性。

[0010] 另外,根据本发明上述实施例的动态场景机器人定位建图系统还可以具有以下附加的技术特征:

[0011] 进一步地,在本发明的一个实施例中,所述语义分割线程与所述位姿估测线程进行数据交流,与运动一致性检测相结合剔除环境中的动态目标。

[0012] 进一步地,在本发明的一个实施例中,所述稠密语义八叉树地图建立线程还用于通过概率融合法建立稳定的3D八叉树地图,将动态目标剔除,并与语义信息相融合,该地图可以用于导航以及机器人完成高级复杂任务。

[0013] 进一步地,在本发明的一个实施例中,所述位姿估测线程进一步用于:获取当前帧到达匹配特征点的LK光流金字塔;通过多个样本点和RANSAC得到基础矩阵;根据所述基础矩阵得到所述当前帧的极线;判断匹配点到所述极线的距离是否小于预设阈值;如果所述距离大于所述预设阈值,则所述匹配点为移动点。

[0014] 进一步地,在本发明的一个实施例中,所述稠密语义八叉树地图建立线程还用于通过逻辑回归值法构建稳定的3D八叉树地图,并将所述3D八叉树地图与所述语义信息相融合。

[0015] 为达到上述目的,本发明另一方面实施例提出了一种动态场景机器人定位建图方法,包括以下步骤:采用深度学习技术实时的获取语义分割结果,以生成带有语义信息的语义图像;根据RGB图像获取ORB特征点,通过运动一致性检测获得运动区域,结合语义分割的信息剔除动态区域上的外点,并通过匹配剩余稳定目标的特征点得到变换矩阵,生成语义图像关键帧;根据所述语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,完成动态目标的剔除,可用于高级复杂任务;对所述关键帧进行局部优化和关键帧筛选;对局部优化和关键帧筛选的结果进行整体优化和回环检测。

[0016] 本发明实施例的动态场景机器人定位建图方法,通过将语义分割网络与运动一致性检测法相结合,并对动态物体的剔除,减少了环境中运动物体对实时定位与地图构建的影响,提高了在动态环境下的定位精度,并且构建的稠密八叉树地图,能够服务于更高层次的工作,从而有效提高了定位建图的适用性、实用性和鲁棒性。

[0017] 另外,根据本发明上述实施例的动态场景机器人定位建图方法还可以具有以下附加的技术特征:

[0018] 进一步地,在本发明的一个实施例中,所述用于根据RGB图像获取语义分割结果,以生成带有语义信息的语义图像,进一步包括:通过与运动一致性检测相结合的实时语义分割网络剔除面向动态环境下的动态目标。

[0019] 进一步地,在本发明的一个实施例中,所述稠密语义八叉树地图建立线程,用于根据所述语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,进一步包括:通过概率融合法构建稳定的3D八叉树地图,并将所述3D八叉树地图与所述语义信息相融合。

[0020] 进一步地,在本发明的一个实施例中,所述根据RGB图像获取ORB特征点,并根据所述语义图像的语义信息和运动一致性检测剔除动态目标上的特征点,并通过匹配剩余的稳定目标的动态点得到变换矩阵,以得到所述语义图像的关键帧,进一步包括:获取当前帧到达匹配特征点的LK光流金字塔;通过多个样本点和RANSAC得到基础矩阵;根据所述基础矩阵得到所述当前帧的极线;判断匹配点到所述极线的距离是否小于预设阈值;如果所述距离大于所述预设阈值,则所述匹配点为移动点。

[0021] 本发明附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0022] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0023] 图1为根据本发明一个实施例的动态场景机器人定位建图系统的结构示意图;

[0024] 图2为根据本发明一个具体实施例的动态场景机器人定位建图系统的结构示意图;

[0025] 图3为根据本发明一个实施例的动态场景机器人定位建图系统的功能示意图;

[0026] 图4为根据本发明一个实施例的运动一致性检测流程图;

[0027] 图5为根据本发明一个实施例的动态目标剔除流程图;

[0028] 图6为根据本发明一个实施例的数据集测试结果示意图;

[0029] 图7为根据本发明一个实施例的真实环境实验示意图;

[0030] 图8为根据本发明一个实施例的语义八叉树地图构建实验示意图;

[0031] 图9为根据本发明一个实施例的动态场景机器人定位建图方法的流程图。

具体实施方式

[0032] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,旨在用于解释本发明,而不能理解为对本发明的限制。

[0033] 在介绍动态场景机器人定位建图系统及方法之前,先简单介绍一下相关技术的语义分割技术和SLAM技术。

[0034] 在人机交互的过程中,机器人主要完成的工作主要有二:一是理解人的指令,二是对指令作出反应。相关技术的机器人定位建图系统与其依托视觉SLAM技术构建的地图,所提供的信息往往只包含地理信息,提供语义信息对于提高机器人的人机交互性便显得尤为重要。另外,相关技术的系统架构往往是适用于一些特定的环境(如空旷地带等)。但在实际应用中,环境通常是动态多变的,图像中出现的人、车等动态物体都会影响机器人建图的准确性,这对系统的鲁棒性提出了更高的要求。

[0035] 解决上述问题的技术基础是语义分割技术。传统的语义分割技术主要依靠手工标记特征点,这种方法的计算消耗较小,但精度并不能让人满意。随着深度学习的发展,不少基于卷积神经网络的语义分割技术问世,如FCN(Fully Convolutional Networks,全卷积网络)。在实际的机器人应用里,实时指令发布是机器人系统的一个关键因素,因此语义分割网络的实时表现与精度是至关重要的。然而,大部分的语义分割网络并不能满足这些要求。

[0036] 在相关技术中,一种针对语义分割的深度完全卷积神经网络模型——SegNet。SegNet的设计考虑了存储与运算的高效性,所需要的训练集数目也远远小于其他模型。有了语义分割网络,就可以使用语义视觉SLAM技术。语义SLAM技术构建的地图包含了语义信息,能够被用于机器人系统更高层次的工作,也能用于提高动态环境下的系统稳定性。

[0037] 正是基于上述原因,本发明实施例提出了一种动态场景机器人定位建图系统及方法。

[0038] 下面参照附图描述根据本发明实施例提出的动态场景机器人定位建图系统及方法,首先将参照附图描述根据本发明实施例提出的动态场景机器人定位建图系统。

[0039] 图1是本发明一个实施例的动态场景机器人定位建图系统的结构示意图。

[0040] 如图1所示,该动态场景机器人定位建图系统10包括:语义分割线程100、位姿估测线程200和稠密语义八叉树地图建立线程300。

[0041] 其中,语义分割线程100用于根据RGB图像获取语义分割结果,以生成带有语义信息的语义图像。位姿估测线程200用于根据RGB图像获取ORB特征,以得到潜在外值,并根据语义图像的语义信息和潜在外值剔除动态目标上的特征点,并通过匹配剩余的稳定目标的动态点得到变换矩阵,以得到语义图像的关键帧。稠密语义八叉树地图建立线程300用于根据语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图。本发明实施例的系统10通过将语义分割网络与运动一致性检测法相结合,并对动态物体的剔除,并且构建的稠密八叉树地图,能够服务于更高层次的工作,从而有效提高了定位建图的适用性、实用性和鲁棒性。

[0042] 可以理解的是,本发明实施例的系统10基于ORB-SLAM2,为5线程架构,语义分割得到的结果能够作为地图信息的补充,本发明实施例的系统10可以根据语义分割的结果,对符合条件的动态物体进行剔除,保证了建图的准确性。在本发明实施例的语义分割线程100中,包含了一个与运动一致性检测相结合的实时语义分割网络,用于剔除面向动态环境下的一些动态目标,如人等,并据剔除后的结果提升了系统的定位精度;稠密语义八叉树地图线程300利用了概率融合法(log-odds score method)对动态目标进行剔除,并将地图与语义信息相融合。

[0043] 具体而言,如图2和图3所示,本发明实施例利用传感器获取RGB图像,并将RGB图像同时送入语义分割线程100与位姿估测线程200。位姿估测线程200首先进行ORB特征点检测与运动一致性检测,待语义分割线程100产生语义图像后,根据语义图像结果进行外点剔除,从而得到位置模拟结果并生成关键帧。得到语义图像与关键帧后,语义八叉树构建线程300产生3D稠密语义八叉树地图。以上三个线程为并发执行,并辅以ORB-SLAM2中的后端优化线程与回环线程,构成了本发明实施例的基本系统架构与执行流程。

[0044] 进一步地,在本发明的一个实施例中,语义分割线程100还用于通过与运动一致性检测相结合的实时语义分割网络剔除面向动态环境下的动态目标。

[0045] 具体而言,如图3所示,在语义分割线程100中,由于本发明实施例是设计用于真实场景下的,因此需要平衡精度与实时性,通过利用基于caffe的SegNet来做到实时语义分割。SegNet采用PASCAL VOC数据集进行训练,能够对20种类型的数据进行分割。

[0046] 进一步地,在本发明的一个实施例中,位姿估测线程200进一步用于:获取当前帧到达匹配特征点的LK光流金字塔;通过多个样本点和RANSAC得到基础矩阵;根据基础矩阵得到当前帧的极线;判断匹配点到极线的距离是否小于预设阈值;如果距离大于预设阈值,则匹配点为移动点。

[0047] 具体而言,如图4所示,由于运动一致性检测是需要消耗时间的,再加上语义分割的结果可以从其他线程中获取,从而只需要确定哪些分割结果中的哪些关键点是移动的。

如果分割后的一个物体里含有一些动态的关键点,那么这个物体就可以被认为是动态的。本发明实施例中移动关键点的判断是很直接的。具体包括:

[0048] 第一步先计算出当前帧到达匹配特征点的LK光流金字塔(opticalflowpyramid),如果匹配对与图像的边缘太过接近,或者匹配对中央 3×3 大小的图块的像素差过大,那么匹配对将会被忽略。随后需要选取大多数样本点,利用RANSAC找出基础矩阵。然后需要用基础矩阵来计算出当前帧的极线(epipolarline)。最后,需要判断匹配点到它相关的极线的距离是否小于阈值,如果该距离大于阈值,那么该匹配点将被认为是移动的。

[0049] 基础矩阵标记了上一帧中的关键点到与之相关的当前帧的匹配范围,即极线。让 p_1 、 p_2 分别代表上一帧与当前帧的匹配点, P_1 、 P_2 是他们的齐次坐标形式:

[0050] $P_1 = [u_1, v_1, 1], P_2 = [u_2, v_2, 1],$

[0051] $p_1 = [u_1, v_1], p_2 = [u_2, v_2]。$

[0052] 极线被定义为 l_1 ,它的计算方式如下:

[0053]
$$l_1 = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = FP_1 = F \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix},$$

[0054] X 、 Y 、 Z 代表线向量, F 代表基础矩阵。那么匹配点到它相关的极线的距离可以定义为:

[0055]
$$D = \frac{|P_2^T F P_1|}{\sqrt{\|X\|^2 + \|Y\|^2}},$$

[0056] D 代表距离。以下是运动一致性检测的算法与动态点的定义, ε 是阈值。

[0057] 输入:前一帧 F_1 ,前一帧的特征点 P_1 ,当前帧 F_2

[0058] 输出:动态点集 S

[0059] 1.判断当前帧是否为第一帧,如是则跳过。

[0060] 2.获取ORB特征点。

[0061] 3.计算LK光流金字塔

[0062] 4.剔除当前帧特征点的异常值

[0063] 5.利用RANSAC算法计算基础矩阵

[0064] 6.计算到极线的距离

[0065] 7.判断该距离是否小于阈值,是则将当前帧归入动态点集

[0066] 进一步地,如图5所示,由于移动物体(例如人)会产生复杂且易变形的动作,因此运动一致性检测法不容易确定动态物体的完整边界,时间上也会增加不少耗费。在本发明实施例中,有了语义分割网络的存在,物体的轮廓可以被快速确定。本发明实施例可以考虑将运动一致性检测法与语义信息相结合,来构成两种基础语义信息:移动物体或非移动物体。如果物体被判为移动的,那么将移除该物体边界的所有特征点,从而动态物体的剔除便简化了许多。

[0067] 等待语义分割线程100完成的时间将被用来做运动一致性检测,时间的利用率得到大幅提升。在语义分割结果产生之后,如果没有检测到动态物体(人),那么所有的ORB特征点都会被用于匹配上一帧并作定位预测。此外,判断人是否移动依靠的是运动一致性检测的结果。如果人被判定为静止的,那么预测结果将会立刻提供,如果人是移动的,则需要先进行动态目标剔除。

[0068] 进一步地,在本发明的一个实施例中,稠密语义八叉树地图建立线程300还用于通过逻辑回归值法对动态目标进行剔除,并将3D八叉树地图与语义信息相融合。

[0069] 可以理解的是,如图3所示,本发明实施例从视觉跟踪进程中获取新的关键帧,从语义线程获取语义图,并对每一帧关键帧计算本地点云,本地点云将会被用于全局八叉树地图。其中,八叉树地图的构建采用逻辑回归值法构建的八叉树地图能够高效存储,且易于导航。

[0070] 具体而言,本发明实施例将语义分割得到的语义信息融入八叉树地图中。八叉树地图中的每一个三位像素点都由一种特殊的颜色代替,而每一种颜色代表一种语义标签。同时由于八叉树地图存储结构的优势,可以用于大场景建图,实时性较好。

[0071] 本发明实施例是设计面向动态场景的稳定建图,最后留存的地图应该是由全部静态物体组成。但语义分割结果的精度是有限的,在复杂场景或者物体重叠较多的情况下,会存在分割错误或者不能完整分割的情况,同时,被分割物体的边缘通常并不能被完美处理。为了解决该问题,本发明实施例采用逻辑回归来处理这种问题,尽可能地减少动态物体对地图的影响,实现动态建模。逻辑回归值(logoddscore)来描述某一个独立的三维像素点是否被占据。定义 $p \in [0, 1]$ 为每一个三维像素点被占用的概率,定义 $l \in \mathbb{R}$ 为概率的逻辑回归值, l 的计算概率如下:

$$[0072] \quad l = \log it(p) = \log\left(\frac{p}{1-p}\right),$$

[0073] 对上式进行变化可得:

$$[0074] \quad p = \log it^{-1}(l) = \frac{\exp(l)}{\exp(l)+1},$$

[0075] 定义 z_t 为像素点 n 在 t 时刻的观察结果,它在 t 时刻的逻辑回归值是 $L(n|Z_{1:t})$,那么在 $t+1$ 时刻,该像素点的逻辑回归值可由下式得到:

$$[0076] \quad L(n|Z_{1:t+1}) = L(n|Z_{1:t-1}) + L(n|Z_t)。$$

[0077] 当 t 时刻像素点 n 被观察到为占有状态时, $L(n|Z_t)$ 与 τ 相等,否则为0(增量 τ 的大小是预先设定的)。观察该公式可知,当一个像素点被重复观察到为占有状态时,逻辑回归值将会增大。只有当一个像素点的概率 p 大于预先设置的阈值时,该像素点将会被是占有且的且在八叉树地图中可视。换言之,只有当一个像素点被多次观察到时才可以被看作稳定出现,从而本发明实施例可以很好地完成动态环境下的建图。

[0078] 进一步地,在本发明的一个实施例中,本发明实施例的装置10还包括:后端优化线程400和回环线程500。

[0079] 其中,后端优化线程400用于对关键帧进行局部优化和关键帧筛选。回环线程500用于对局部优化和关键帧筛选的结果进行整体优化和回环检测。

[0080] 综上,本发明实施例利用视觉传感器获取RGB图像,并将之送至位姿估测线程与语义分割线程。位姿估测线程首先获取ORB特征点,随后对特征点进行粗略的运动一致性检测,并记录潜在的外值(动态点)。当语义分割的结果产生后,位姿估测线程得到带有语义信息的图像,并根据语义信息与先前得到的潜在外值,剔除动态目标上的特征点。之后,线程将会通过匹配剩余的稳定目标的动态点来计算变换矩阵。后端优化线程与回环线程同ORB-SLAM2相同。稠密语义八叉树地图建立线程用于建立带语义标签的3D八叉树地图,从而使机器人能适应更高层次的工作。

[0081] 在本发明的一个具体实施例中,如图6所示,为了验证本发明实施例的使用效果,分为两步进行验证。首先使用TUM RGB-D进行验证,该数据集提供了多种不同的动态环境,包含了行走、坐等动作以及桌子等静态物体。本发明实施例以ORB-SLAM2为基础,因此将本发明实施例的试验结果与其进行比较,以ATE (Absolute Trajectory Error,绝对轨迹误差)与RPE (Relative Pose Error,相对位置误差)作为主要指标。对照附图5为测试结果,其中上方为ORB-SLAM2的实验结果,下方为本发明实施例的实验结果。其中左边为ATE的测试结果,由真实值 (ground truth)、模拟值 (estimated)、相对误差 (difference) 构成;右图为RPE的测试结果,纵轴代表相对位置误差距离。由结果可见,本发明实施例在两项指标上均有较大幅度的提升,而且在实时性上的表现也更好。

[0082] 为了进一步测试本发明实施例的系统10的鲁棒性和实时表现效果,本发明实施例还利用Kinect V2摄像头 (960×540) 作为图像采集源,在ROS系统下进行了真实环境测试。如图7所示,从上至下每行分别代表ORB特征点检测结果、基于光流法的运动一致性检测、语义分割结果、动态目标剔除后的图像。图中的绿点代表ORB特征点、红线代表光流方向。可见经过本发明实施例的系统10的处理后,人作为动态物体被剔除,且显示效果较好。如图8所示,可见八叉树地图建图效果,红色的像素块代表沙发,蓝色的像素块代表显示器。其中,沙发和显示器均以数据库中特定颜色显示。地图中展示的是概率融合法筛选后的像素块,可见走动的人作为动态物体并不会影响建图效果。

[0083] 综上,本发明实施例阐述了一种实时的、鲁棒性较好的、针对动态环境的语义SLAM技术。本发明实施例系统10基于ROS (Robot Operating System,机器人操作系统) 系统,它包含了五个部分:视觉跟踪、语义分割、本地建图、回环检测以及稠密八叉树地图构建,并通过TUM RGB-D数据集验证,较于ORB-SLAM2在动态环境下的建图精度与鲁棒性有较为显著的提升。此外,本发明实施例的效果也在现实环境下得到了测试,效果较好。具体地,本发明实施例利用TUM数据集的测试结果与ORB-SLAM2相比,ATE与RPE的根均方误差与标准差有90%以上的提升,提升了近一个量级;此外,在真实环境实验中,本发明实施例对于移动物体的剔除做的较好,建图也较精确,验证了系统的可用性和实时性。

[0084] 根据本发明实施例提出的动态场景机器人定位建图系统,通过将语义分割网络与运动一致性检测法相结合,并对动态物体的剔除,减少了环境中运动物体对实时定位与地图构建的影响,提高了在动态环境下的定位精度,并且构建的稠密八叉树地图,能够服务于更高层次的工作,从而有效提高了定位建图的适用性、实用性和鲁棒性。

[0085] 其次参照附图描述根据本发明实施例提出的动态场景机器人定位建图方法的流程图。

[0086] 如图9所示,该动态场景机器人定位建图方法包括以下步骤:

[0087] 在步骤S901中,采用深度学习技术实时的获取语义分割结果,以生成带有语义信息的语义图像。

[0088] 在步骤S902中,根据RGB图像获取ORB特征点,通过运动一致性检测获得运动区域,结合语义分割的信息剔除动态区域上的外点,并通过匹配剩余稳定目标的特征点得到变换矩阵,生成语义图像关键帧。

[0089] 在步骤S903中,根据语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,完成动态目标的剔除,可用于高级复杂任务。

[0090] 在步骤S904中,对关键帧进行局部优化和关键帧筛选。

[0091] 在步骤S905中,对局部优化和关键帧筛选的结果进行整体优化和回环检测。

[0092] 进一步地,在本发明的一个实施例中,用于根据Raw图像获取语义分割结果,以生成带有语义信息的语义图像,进一步包括:通过与运动一致性检测相结合的实时语义分割网络剔除面向动态环境下的动态目标。

[0093] 进一步地,在本发明的一个实施例中,稠密语义八叉树地图建立线程,用于根据语义图像的关键帧生成带语义标签的3D稠密语义八叉树地图,进一步包括:通过概率融合法建立稳定的3D八叉树地图,并将3D八叉树地图与语义信息相融合。

[0094] 进一步地,在本发明的一个实施例中,根据RGB图像获取ORB特征点并根据语义图像的语义信息和运动一致性检测剔除动态目标上的特征点,并通过匹配剩余的稳定目标的动态点得到变换矩阵,以得到语义图像的关键帧,进一步包括:获取当前帧到达匹配特征点的LK光流金字塔;通过多个样本点和RANSAC得到基础矩阵;根据基础矩阵得到当前帧的极线;判断匹配点到极线的距离是否小于预设阈值;如果距离大于预设阈值,则匹配点为移动点。

[0095] 需要说明的是,前述对动态场景机器人定位建图系统实施例的解释说明也适用于该实施例的动态场景机器人定位建图方法,此处不再赘述。

[0096] 根据本发明实施例提出的动态场景机器人定位建图方法,通过将语义分割网络与运动一致性检测法相结合,并对动态物体的剔除,减少了环境中运动物体对实时定位与地图构建的影响,提高了在动态环境下的定位精度,并且构建的稠密八叉树地图,能够服务于更高层次的工作,从而有效提高了定位建图的适用性、实用性和鲁棒性。

[0097] 在本发明的描述中,需要理解的是,术语“中心”、“纵向”、“横向”、“长度”、“宽度”、“厚度”、“上”、“下”、“前”、“后”、“左”、“右”、“竖直”、“水平”、“顶”、“底”“内”、“外”、“顺时针”、“逆时针”、“轴向”、“径向”、“周向”等指示的方位或位置关系为基于附图所示的方位或位置关系,仅是为了便于描述本发明和简化描述,而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作,因此不能理解为对本发明的限制。

[0098] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中,“多个”的含义是至少两个,例如两个,三个等,除非另有明确具体的限定。

[0099] 在本发明中,除非另有明确的规定和限定,术语“安装”、“相连”、“连接”、“固定”等术语应做广义理解,例如,可以是固定连接,也可以是可拆卸连接,或成一体;可以是机械连接,也可以是电连接;可以是直接相连,也可以通过中间媒介间接相连,可以是两个元件内部的连通或两个元件的相互作用关系,除非另有明确的限定。对于本领域的普通技术人员而言,可以根据具体情况理解上述术语在本发明中的具体含义。

[0100] 在本发明中,除非另有明确的规定和限定,第一特征在第二特征“上”或“下”可以是第一和第二特征直接接触,或第一和第二特征通过中间媒介间接接触。而且,第一特征在第二特征“之上”、“上方”和“上面”可是第一特征在第二特征正上方或斜上方,或仅仅表示第一特征水平高度高于第二特征。第一特征在第二特征“之下”、“下方”和“下面”可以是第一特征在第二特征正下方或斜下方,或仅仅表示第一特征水平高度小于第二特征。

[0101] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不必针对的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0102] 尽管上面已经示出和描述了本发明的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本发明的限制,本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

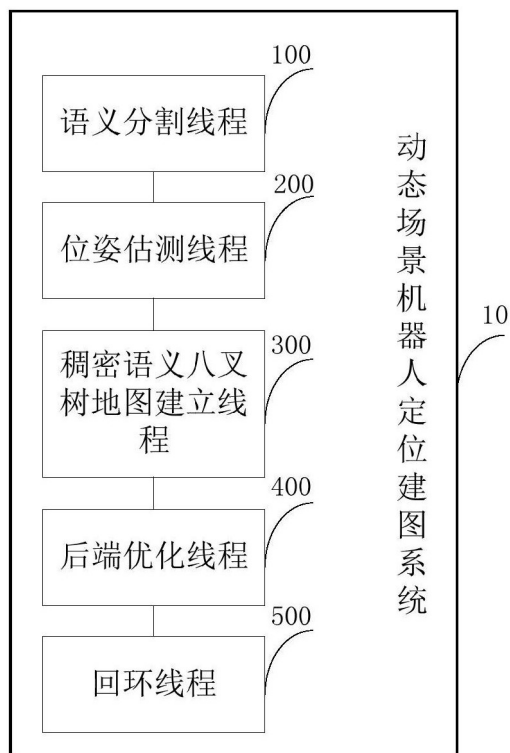


图1

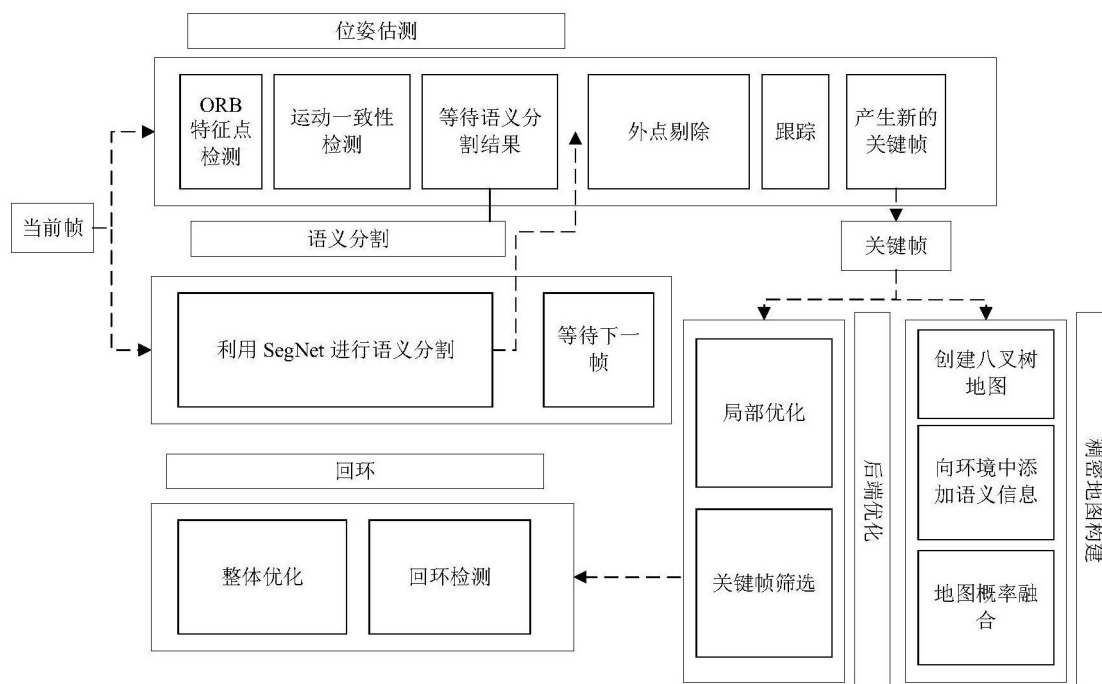


图2

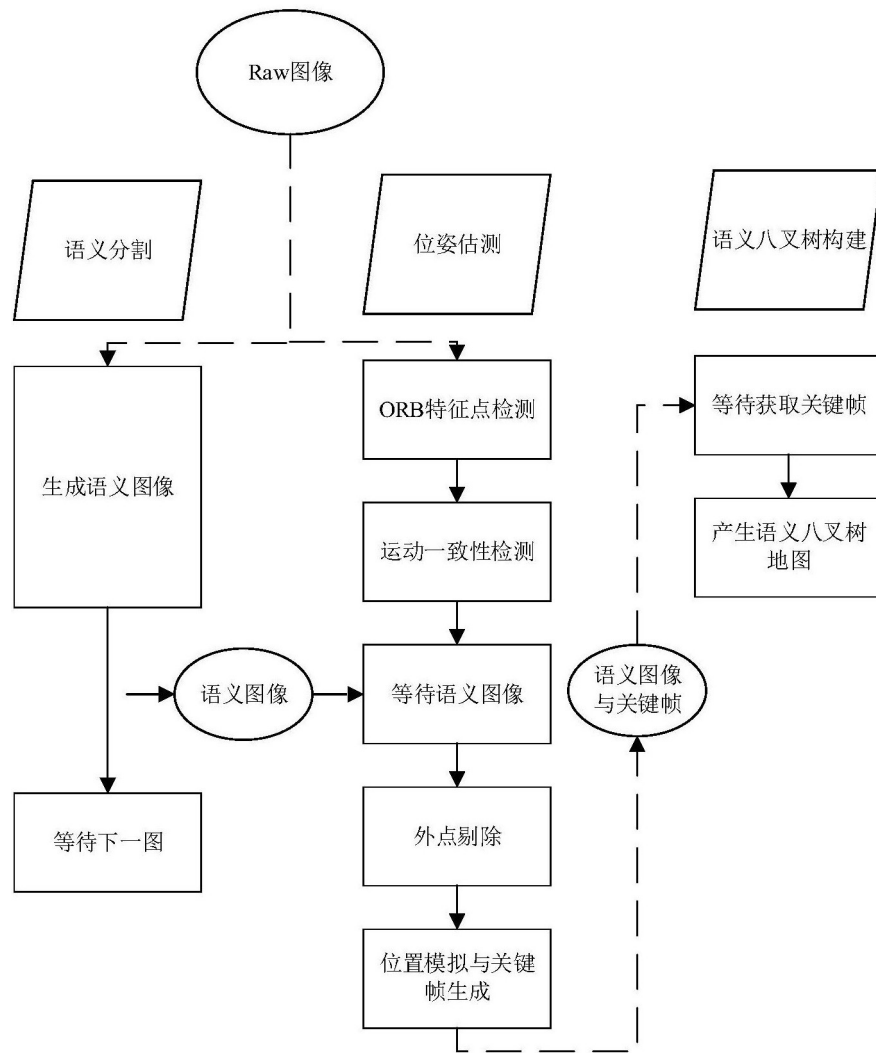


图3

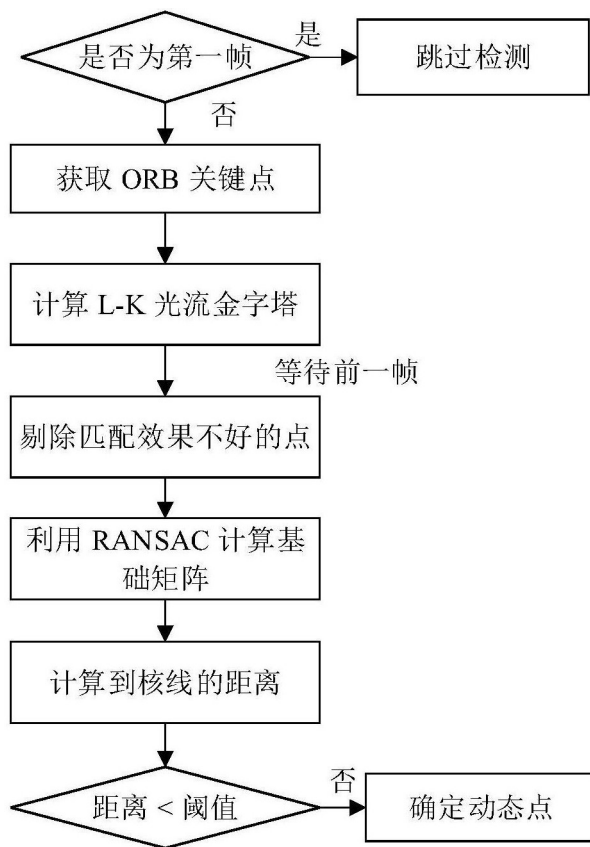


图4

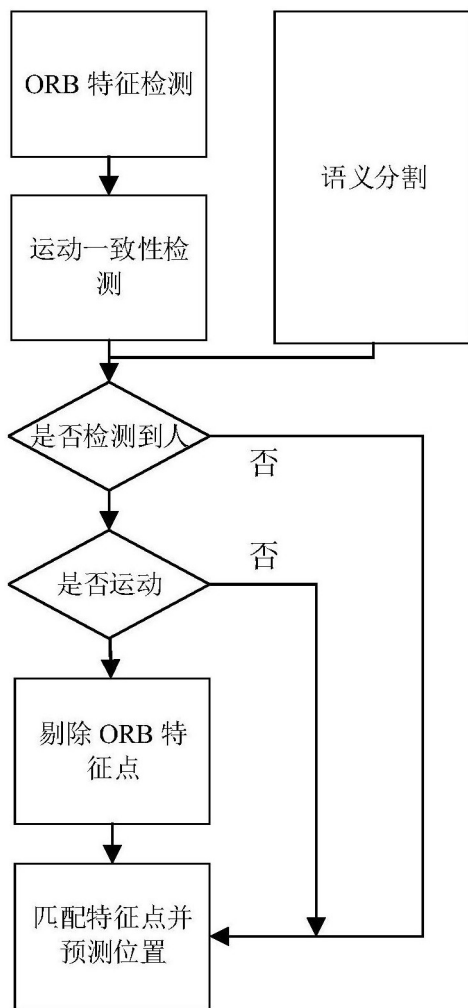


图5

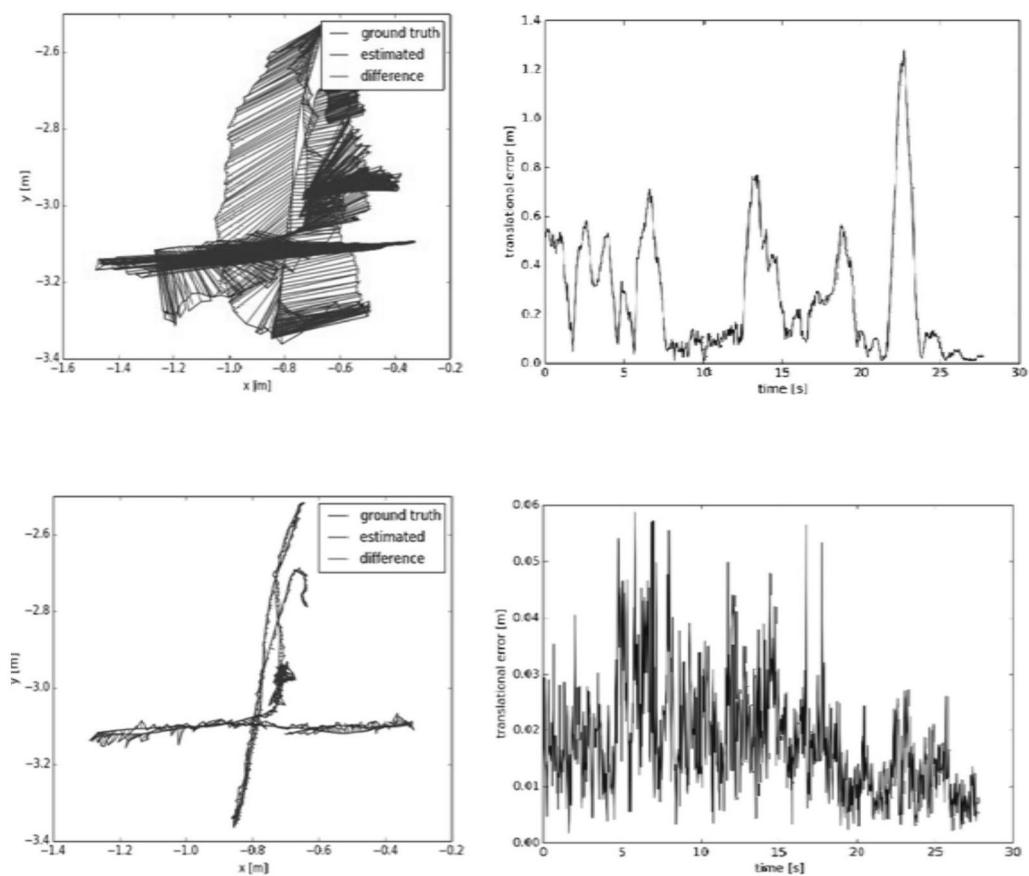


图6

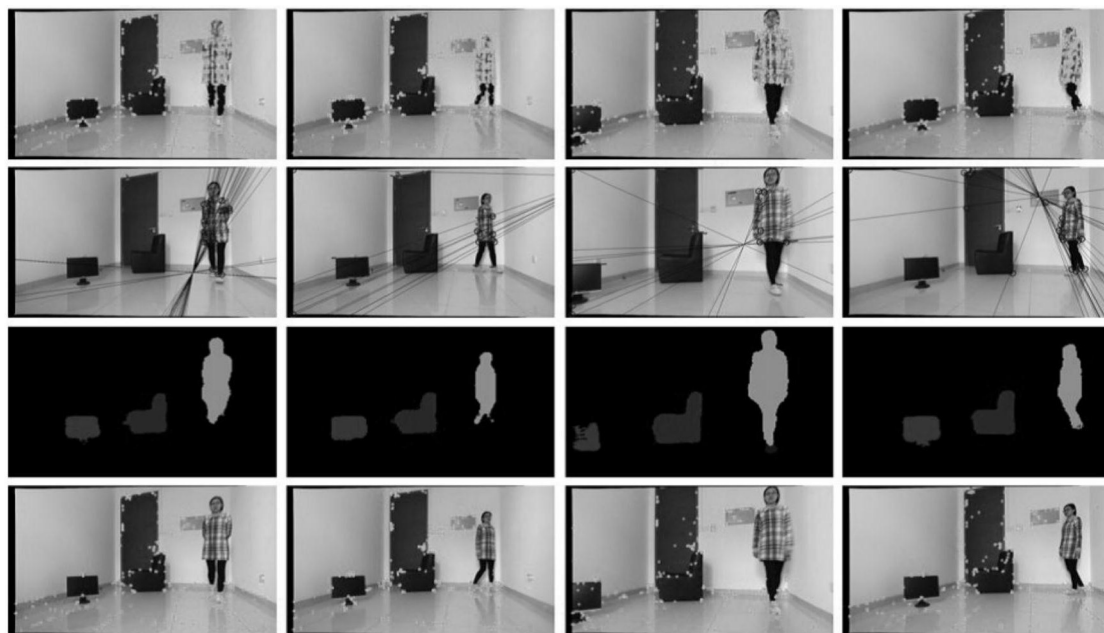


图7



图8

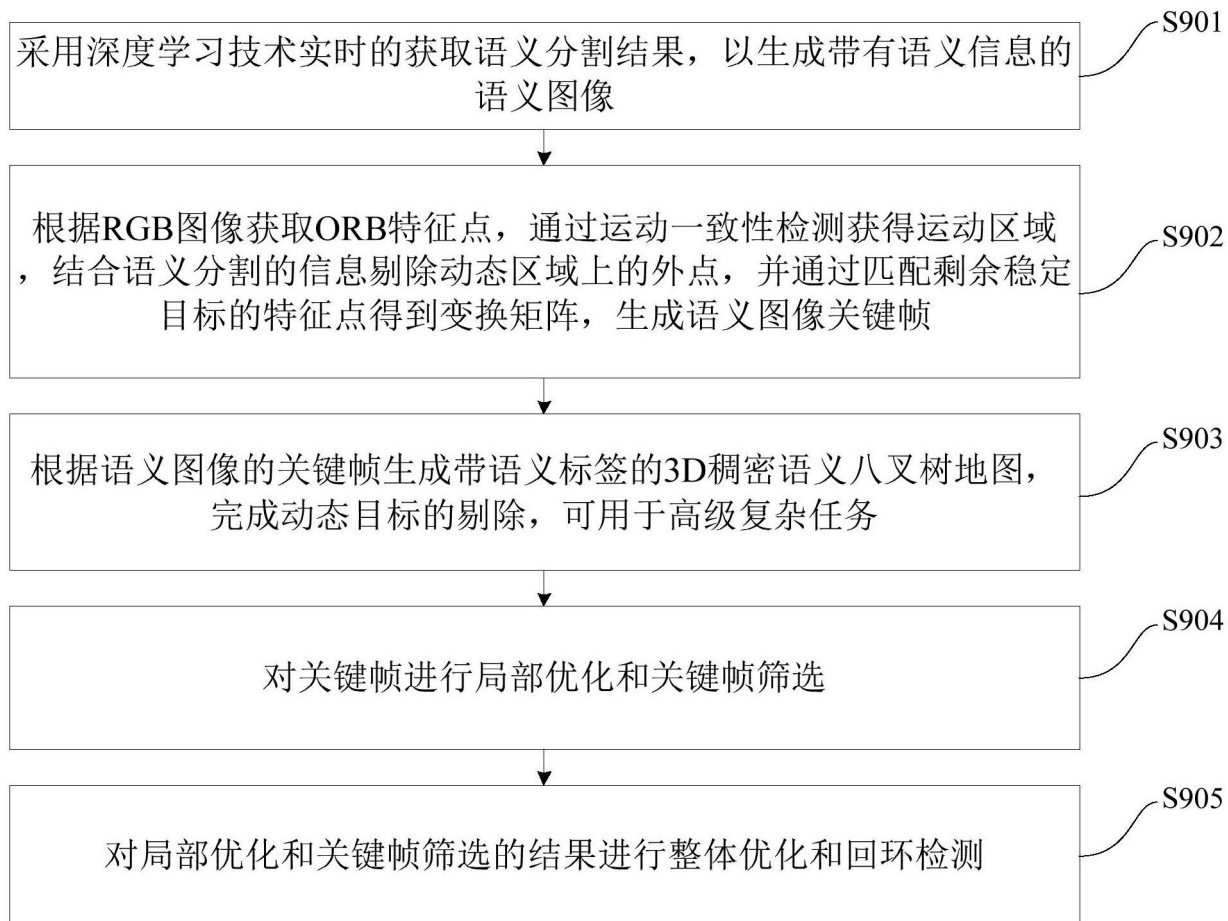


图9