



(12)发明专利申请

(10)申请公布号 CN 111724439 A

(43)申请公布日 2020.09.29

(21)申请号 201911200881.0

(22)申请日 2019.11.29

(71)申请人 中国科学院上海微系统与信息技术研究所

地址 200050 上海市长宁区长宁路865号

(72)发明人 姜昊辰 张晓林 李嘉茂 刘衍青 朱冬晨 彭镜铨

(74)专利代理机构 广州三环专利商标代理有限公司 44202

代理人 郝传鑫 贾允

(51)Int.Cl.

G06T 7/73(2017.01)

G06K 9/00(2006.01)

G06K 9/62(2006.01)

B25J 9/16(2006.01)

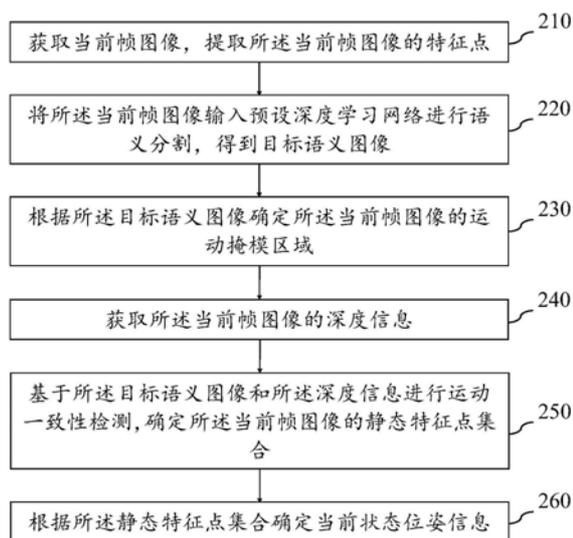
权利要求书2页 说明书9页 附图5页

(54)发明名称

一种动态场景下的视觉定位方法及装置

(57)摘要

本发明涉及机器人导航定位技术领域,具体是一种动态场景下的视觉定位方法及装置,所述方法包括:获取当前帧图像,提取所述当前帧图像的特征点;将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像;根据所述目标语义图像确定所述当前帧图像的运动掩模区域;获取所述当前帧图像的深度信息;基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合;根据所述静态特征点集合确定当前状态位姿信息。本发明通过语义分割结果和深度信息进行运动一致性检测,确定图像的静态特征点集合,能够有效提高在动态环境中对位姿估计的精确性和鲁棒性。



1. 一种动态场景下的视觉定位方法,其特征在于,包括:
 - 获取当前帧图像,提取所述当前帧图像的特征点;
 - 将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像;
 - 根据所述目标语义图像确定所述当前帧图像的运动掩模区域;
 - 获取所述当前帧图像的深度信息;
 - 基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合;
 - 根据所述静态特征点集合确定当前状态位姿信息。
2. 根据权利要求1所述的方法,其特征在于,所述基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合包括:
 - 根据所述目标语义图像确定所述当前帧图像的背景区域;
 - 根据所述当前帧图像的背景区域的特征点确定第一位姿信息;
 - 基于所述第一位姿信息和所述深度信息确定所述运动掩模区域的特征点的类型,所述类型包括动态特征点和静态特征点;
 - 剔除所述运动掩模区域的特征点中的动态特征点,保留静态特征点;
 - 根据所述静态特征点和所述背景区域的特征点生成静态特征点集合。
3. 根据权利要求2所述的方法,其特征在于,所述基于所述第一位姿信息和所述深度信息确定所述运动掩模区域的特征点的类型包括:
 - 根据所述第一位姿信息和所述深度信息计算所述运动掩模区域的特征点的运动得分;
 - 当所述运动得分小于预设阈值时,判定所述特征点为静态特征点;
 - 当所述运动得分大于或等于预设阈值时,判定所述特征点为动态特征点。
4. 根据权利要求3所述的方法,其特征在于,所述根据所述第一位姿信息和所述深度信息计算所述运动掩模区域的特征点的运动得分包括:
 - 获取第一参考帧图像的运动掩模区域的特征点,将所述第一参考帧图像的运动掩模区域的特征点与所述当前帧图像的运动掩模区域的特征点进行匹配,得到匹配点对;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像;
 - 对所述匹配点对进行筛选,去除误匹配的匹配点对;
 - 根据所述第一位姿信息和所述深度信息计算筛选后的匹配点对之间的距离,将所述距离作为所述运动掩模区域的特征点的运动得分。
5. 根据权利要求1或2所述的方法,其特征在于,所述根据所述目标语义图像确定所述当前帧图像的运动掩模区域之后,还包括:
 - 获取第一参考帧图像的第一运动掩模区域和第二参考帧图像的第二运动掩模区域;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像,所述第二参考帧图像为所述第一参考帧图像的前一帧图像;
 - 根据所述第一运动掩模区域和所述当前帧图像的运动掩模区域判断所述当前帧图像是否发生漏检;
 - 如果所述当前帧图像发生漏检,根据所述第一运动掩模区域和所述第二运动掩模区域确定第一目标运动掩模区域;
 - 将所述当前帧图像的运动掩模区域替换为所述第一目标运动掩模区域。

6. 根据权利要求5所述的方法,其特征在于,所述方法还包括:

如果所述当前帧图像未发生漏检,根据所述第一运动掩模区域和所述当前帧图像的运动掩模区域确定第二目标运动掩模区域;

将所述当前帧图像的运动掩模区域替换为所述第二目标运动掩模区域。

7. 根据权利要求1或2所述的方法,其特征在于,所述获取所述当前帧图像的深度信息之后,还包括:

利用预设形态学方法对所述深度信息进行修复。

8. 根据权利要求1或2所述的方法,其特征在于,所述根据所述静态特征点集合确定当前状态位姿信息包括:

在产生新的关键帧时,建立所述静态特征点集合中的特征点、关键帧和地图点之间的数据关联;

根据所述静态特征点集合中的特征点确定第二位姿信息;

根据所述第二位姿信息和所述数据关联进行位姿优化,确定当前状态位姿信息。

9. 根据权利要求8所述的方法,其特征在于,所述方法还包括:

基于所述目标语义图像、所述当前状态位姿信息及所述静态特征点集合生成静态物体点云数据;

根据所述静态物体点云数据进行静态物体点云地图的稠密重建。

10. 一种动态场景下的视觉定位装置,其特征在于,包括:

第一获取模块,用于获取当前帧图像,提取所述当前帧图像的特征点;

语义分割模块,用于将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像;

确定模块,用于根据所述目标语义图像确定所述当前帧图像的运动掩模区域;

第二获取模块,用于获取所述当前帧图像的深度信息;

检测模块,用于基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合;

定位模块,用于根据所述静态特征点集合确定当前状态位姿信息。

一种动态场景下的视觉定位方法及装置

技术领域

[0001] 本发明涉及机器人导航定位技术领域,特别涉及一种动态场景下的视觉定位方法及装置。

背景技术

[0002] 随着人工智能技术的发展,越来越多的智能移动机器人出现在生产生活中的各个场景。从工业机器人到家政服务机器人,从无人机到水下探测机器人,智能化的一个重要条件就是能够自主运动,即实现机器人的自主导航。智能移动机器人要实现各种环境下的自主化移动,需要解决两个基本的问题分别是定位与建图,而其中核心就是同时定位与建图(Simultaneous Localization and Mapping,SLAM)技术。

[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。现有的视觉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,也已经在某些环境下取得了较好的试验结果。但是现有的基于点特征的视觉SLAM技术都是基于静态环境假设的,而对于大多数的实际场景,绝对的静态场景是不存在的,因此在位于动态环境时其位姿估计的准确性会急剧下降甚至无法工作。与此同时,由于未对运动物体加以判断,在重建稠密点云地图的时候会导致“伪迹”的出现,从而造成对环境的错误感知。

发明内容

[0004] 针对现有技术的上述问题,本发明的目的在于提供一种动态场景下的视觉定位方法及装置,能够提高在动态环境中对位姿估计的精确性和鲁棒性。

[0005] 为了解决上述问题,本发明提供一种动态场景下的视觉定位方法,包括:

[0006] 获取当前帧图像,提取所述当前帧图像的特征点;

[0007] 将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像;

[0008] 根据所述目标语义图像确定所述当前帧图像的运动掩模区域;

[0009] 获取所述当前帧图像的深度信息;

[0010] 基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合;

[0011] 根据所述静态特征点集合确定当前状态位姿信息。

[0012] 进一步地,所述基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合包括:

[0013] 根据所述目标语义图像确定所述当前帧图像的背景区域;

[0014] 根据所述当前帧图像的背景区域的特征点确定第一位姿信息;

[0015] 基于所述第一位姿信息和所述深度信息确定所述运动掩模区域的特征点的类型,所述类型包括动态特征点和静态特征点;

- [0016] 剔除所述运动掩模区域的特征点中的动态特征点,保留静态特征点;
- [0017] 根据所述静态特征点和所述背景区域的特征点生成静态特征点集合。
- [0018] 进一步地,所述基于所述第一位姿信息和所述深度信息确定所述运动掩模区域的特征点的类型包括:
- [0019] 根据所述第一位姿信息和所述深度信息计算所述运动掩模区域的特征点的运动得分;
- [0020] 当所述运动得分小于预设阈值时,判定所述特征点为静态特征点;
- [0021] 当所述运动得分大于或等于预设阈值时,判定所述特征点为动态特征点。
- [0022] 具体地,所述根据所述第一位姿信息和所述深度信息计算所述运动掩模区域的特征点的运动得分包括:
- [0023] 获取第一参考帧图像的运动掩模区域的特征点,将所述第一参考帧图像的运动掩模区域的特征点与所述当前帧图像的运动掩模区域的特征点进行匹配,得到匹配点对;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像;
- [0024] 对所述匹配点对进行筛选,去除误匹配的匹配点对;
- [0025] 根据所述第一位姿信息和所述深度信息计算筛选后的匹配点对之间的距离,将所述距离作为所述运动掩模区域的特征点的运动得分。
- [0026] 进一步地,所述根据所述目标语义图像确定所述当前帧图像的运动掩模区域之后,还包括:
- [0027] 获取第一参考帧图像的第一运动掩模区域和第二参考帧图像的第二运动掩模区域;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像,所述第二参考帧图像为所述第一参考帧图像的前一帧图像;
- [0028] 根据所述第一运动掩模区域和所述当前帧图像的运动掩模区域判断所述当前帧图像是否发生漏检;
- [0029] 如果所述当前帧图像发生漏检,根据所述第一运动掩模区域和所述第二运动掩模区域确定第一目标运动掩模区域;
- [0030] 将所述当前帧图像的运动掩模区域替换为所述第一目标运动掩模区域。
- [0031] 优选地,所述方法还包括:
- [0032] 如果所述当前帧图像未发生漏检,根据所述第一运动掩模区域和所述当前帧图像的运动掩模区域确定第二目标运动掩模区域;
- [0033] 将所述当前帧图像的运动掩模区域替换为所述第二目标运动掩模区域。
- [0034] 进一步地,所述获取所述当前帧图像的深度信息之后,还包括:
- [0035] 利用预设形态学方法对所述深度信息进行修复。
- [0036] 进一步地,所述根据所述静态特征点集合确定当前状态位姿信息包括:
- [0037] 在产生新的关键帧时,建立所述静态特征点集合中的特征点、关键帧和地图点之间的数据关联;
- [0038] 根据所述静态特征点集合中的特征点确定第二位姿信息;
- [0039] 根据所述第二位姿信息和所述数据关联进行位姿优化,确定当前状态位姿信息。
- [0040] 进一步地,所述方法还包括:
- [0041] 基于所述目标语义图像、所述当前状态位姿信息及所述静态特征点集合生成静态

物体点云数据；

[0042] 根据所述静态物体点云数据进行静态物体点云地图的稠密重建。

[0043] 本发明另一方面保护一种动态场景下的视觉定位装置,包括:

[0044] 第一获取模块,用于获取当前帧图像,提取所述当前帧图像的特征点;

[0045] 语义分割模块,用于将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像;

[0046] 确定模块,用于根据所述目标语义图像确定所述当前帧图像的运动掩模区域;

[0047] 第二获取模块,用于获取所述当前帧图像的深度信息;

[0048] 检测模块,用于基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合;

[0049] 定位模块,用于根据所述静态特征点集合确定当前状态位姿信息。

[0050] 由于上述技术方案,本发明具有以下有益效果:

[0051] 本发明的动态场景下的视觉定位方法通过对图像进行语义分割,并根据语义分割结果和深度信息进行运动一致性检测,剔除运动掩模区域的动态特征点,将运动掩模中的静态特征点加入静态特征点集合,并利用所述静态特征点集合实现位姿优化,并重建出稠密的静态物体点云地图。能够提高在动态环境中对位姿估计的精确性和鲁棒性,提高所述静态物体点云地图的准确度,从而提高对环境的感知的准确度。

[0052] 另外,本发明的动态场景下的视觉定位方法基于运动连续性假设,对相邻帧图像的语义分割结果进行融合,补偿漏分割的情况,能够进一步提高本方法的准确性。

附图说明

[0053] 为了更清楚地说明本发明的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单的介绍。显而易见地,下面描述中的附图仅仅是本发明的一些实施例,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它附图。

[0054] 图1是本发明一个实施例提供的动态场景下的视觉定位系统的结构示意图;

[0055] 图2是本发明一个实施例提供的动态场景下的视觉定位方法的流程图;

[0056] 图3是本发明另一个实施例提供的动态场景下的视觉定位方法的流程图;

[0057] 图4是本发明另一个实施例提供的动态场景下的视觉定位方法的流程图;

[0058] 图5是本发明另一个实施例提供的动态场景下的视觉定位方法的流程图;

[0059] 图6A是本发明一个实施例提供的ORB-SLAM2系统的测试结果示意图;

[0060] 图6B是本发明一个实施例提供的DS-SLAM系统的测试结果示意图;

[0061] 图6C是本发明一个实施例提供的动态场景下的视觉定位方法的测试结果示意图;

[0062] 图7是本发明另一个实施例提供的动态场景下的视觉定位装置的结构示意图。

具体实施方式

[0063] 下面将结合本发明实施例中的附图,对本发明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发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通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,都属于本发明保护的范围。

[0064] 为了使本发明实施例公开的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本发明实施例进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本发明实施例,并不用于限定本发明实施例。首先,本发明实施例对下述概念进行解释:

[0065] ORB-SLAM2系统:一个开源的针对单目、双目和RGBD相机的SLAM系统。ORB-SLAM系统是一个基于特征点的实时单目SLAM系统,在大规模的、小规模、室内室外的环境都可以运行。该系统对剧烈运动也很鲁棒,支持宽基线的闭环检测和重定位,包括全自动初始化。ORB-SLAM2系统在ORB-SLAM系统的基础上,还支持标定后的双目相机和RGBD相机。

[0066] RGBD:Red Green Blue Deep,三通道彩色图片及深度信息。

[0067] DS-SLAM系统:一个针对动态场景的语义视觉SLAM系统。DS-SLAM系统基于ORB-SLAM2系统提出,将语义分割网络与移动一致性检查方法相结合,减少了动态对象的影响,从而在动态环境中大大提高了定位精度。

[0068] 参考说明书附图1,本实施例提供一种动态场景下的视觉定位系统,所述系统可以包括语义分割与补偿线程110、跟踪估计线程120、局部地图线程130、回环检测线程140和稠密点云建图线程150。

[0069] 所述语义分割与补偿线程110通过语义分割等神经网络方法,输出对于输入图片的逐像素的语义分类结果,基于运动连续性假设,对于相邻帧的语义分割结果融合,补偿漏分割的情况。

[0070] 所述跟踪估计线程120,通过在当前帧图像上提取特征点,集合语义信息,计算静态背景的相对运动关系,接着计算运动掩模区域的特征点匹配关系,将不满足匹配关系的特征点去掉,通过运动一致性检测和深度图加权的方式结合相对运动关系判断特征点的动静情况,剔除动态特征点,并更新静态特征点集。

[0071] 所述局部地图线程130,通过对多个关键帧和特征点运用局部滑窗的方式,基于共视图进行局部位姿调整,是在跟踪估计线程120的基础上对于位姿的二次优化过程。

[0072] 所述回环检测线程140,通过对每一帧的情况和所有关键帧比较,发现有相似情况时对所有关键帧进行一次全局优化和筛选。

[0073] 所述稠密点云建图线程150,通过上几个线程的操作,将动态的物体区域去除,并根据估计出的鲁棒稳定的相对位姿,使用ICP等方法对点云进行拼接,从而得到静态物体的稠密点云地图。

[0074] 参考说明书附图2,本实施例提供一种动态场景下的视觉定位方法,所述方法可以包括以下步骤:

[0075] S210:获取当前帧图像,提取所述当前帧图像的特征点。

[0076] 本发明实施例中,可以通过跟踪估计线程使用ORB特征提取方法提取所述当前帧图像的特征点,采用ORB描述子的特征描述方法。在一些可能的实施例中,也可以采用其他特征描述方法,本发明对此不做限定。

[0077] S220:将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像。

[0078] 本发明实施例中,可以通过语义分割与补偿线程对所述当前帧图像进行语义分

割。所述预设深度学习网络可以包括ENet语义分割网络,所述ENet语义分割网络是一个比较常用的分割网络,它拥有简单的网络结构、快速的运行时间和不多的变量,能够被应用于实时图像分割和移动端设备。

[0079] 实际应用中,可以首先按照三通道彩色图像序列进行标识映射处理,构成训练集对应训练标签;采用非对称网络结构,便于解码端对编码端进行上采样微调;将网络中的线性整流函数(reLU)层替换为参数修正线性单元(PReLU),增加特征图额外参数;把瓶颈结构(bottleneck)中的卷积层替换为空洞卷积并进行串联,增大感受野;使用空间随机丢弃(Dropout)处理,防止过拟合。

[0080] 在一些可能的实施例中,也可以采用其他语义分割网络,本发明对此不做限定。

[0081] S230:根据所述目标语义图像确定所述当前帧图像的运动掩模区域。

[0082] 本发明实施例中,可以通过语义分割与补偿线程确定所述运动掩模区域。所述运动掩模区域可以包括潜在运动物体区域,所述潜在运动物体可以包括人、动物等。

[0083] 在一个可能的实施例中,如图3所示,所述根据所述目标语义图像确定所述当前帧图像的运动掩模区域之后,还可以包括:

[0084] S310:获取第一参考帧图像的第一运动掩模区域和第二参考帧图像的第二运动掩模区域;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像,所述第二参考帧图像为所述第一参考帧图像的前一帧图像。

[0085] 本发明实施例中,所述第一参考帧图像的第一运动掩模区域和第二参考帧图像的第二运动掩模区域可以通过预设深度学习网络进行语义分割确定。

[0086] S320:根据所述第一运动掩模区域和所述当前帧图像的运动掩模区域判断所述当前帧图像是否发生漏检。

[0087] 本发明实施例中,可以通过计算所述当前帧图像的运动掩模与所述第一参考帧图像的运动掩模的交并比,根据所述交并比确定所述当前帧图像是否发生漏检。

[0088] S330:如果所述当前帧图像发生漏检,根据所述第一运动掩模区域和所述第二运动掩模区域确定第一目标运动掩模区域。

[0089] 本发明实施例中,如果所述当前帧图像发生漏检,可以将所述第一参考帧图像的第一运动掩模区域和第二参考帧图像的第二运动掩模区域逐像素投影到当前帧当中得到第一目标运动掩模区域。

[0090] 具体地,假定 S_{t_1} 为第一参考帧图像的运动掩模, S_{t_2} 为第二参考帧图像的运动掩模,假定 S_t 为当前帧图像的运动掩模,则所述当前帧图像的运动掩模与所述第一参考帧图像的运动掩模的交并比计作 D_{iou} ,则 D_{iou} 可以定义为:

$$[0091] \quad D_{iou} = \frac{S_{t_1} \cap S_{t_2}}{S_{t_1} \cup S_{t_2}}$$

[0092] 假设 ε_{iou} 是 D_{iou} 的阈值,若计算得到的 D_{iou} 值小于 ε_{iou} ,则说明当前帧图像语义发生漏检。这时将 S_{t_1} 和 S_{t_2} 的运动掩模逐像素投影到当前帧当中,计算新的语义检测结果,记为 S_{iou} ,则 S_{iou} 可定义为:

$$[0093] \quad S_{iou} = S_{t_1} \cup S_{t_2}$$

[0094] 这里取交集是由于后面的运动一致性检测系统会将满足条件的特征点重新放回

到静态点集中,因此这里的扩大不会影响系统的精度。

[0095] S340:将所述当前帧图像的运动掩模区域替换为所述第一目标运动掩模区域。

[0096] 本发明实施例中,确定所述当前帧图像的运动掩模区域后,可以通过语义分割与补偿线程对当前帧图像是否漏分割进行检测,当判定发生漏检时,利用所述当前帧图像的前两帧图像的语义分割结果融合进行补偿。

[0097] 在另一个可能的实施例中,如图3所示,所述方法还可以包括:

[0098] S350:如果所述当前帧图像未发生漏检,根据所述第一运动掩模区域和所述当前帧图像的运动掩模区域确定第二目标运动掩模区域。

[0099] 本发明实施例中,如果所述当前帧图像未发生漏检,还可以将所述第一参考帧图像的第一运动掩模区域逐像素投影到当前帧当中,与所述当前帧图像的运动掩模区域合并得到第二目标运动掩模区域。

[0100] S360:将所述当前帧图像的运动掩模区域替换为所述第二目标运动掩模区域。

[0101] 本发明实施例中,当判定没有发生漏检时,可以利用所述当前帧图像的前一帧图像的语义分割结果进行补偿,进一步减少发生漏检的可能性。

[0102] S240:获取所述当前帧图像的深度信息。

[0103] 在一个可能的实施例中,所述获取所述当前帧图像的深度信息之后,还可以包括:

[0104] 利用预设形态学方法对所述深度信息进行修复。

[0105] 本发明实施例中,可以通过跟踪估计线程采用膨胀操作和插值等方法对所述当前帧图像的深度信息进行修复。

[0106] S250:基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合。

[0107] 本发明实施例中,可以通过跟踪估计线程进行运动一致性检测,剔除当前帧图像中的运动物体的动态特征点,得到当前帧图像的静态特征点集合。

[0108] 在一个可能的实施例中,如图4所示,所述基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合可以包括:

[0109] S410:根据所述目标语义图像确定所述当前帧图像的背景区域。

[0110] 本发明实施例中,所述背景区域可以包括静态物体区域,所述静态物体可以包括道路、树木和建筑物等。

[0111] S420:根据所述当前帧图像的背景区域的特征点确定第一位姿信息。

[0112] 本发明实施例中,所述根据所述当前帧图像的背景区域的特征点确定第一位姿信息可以包括:获取第一参考帧图像的背景区域的特征点,将所述第一参考帧图像的背景区域的特征点与所述当前帧图像的背景区域的特征点进行匹配,得到匹配点对;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像;通过距离约束和随机抽样一致(Random sample consensus, RANSAC)算法对所述匹配点对进行筛选,去除误匹配的匹配点对;通过归一化八点法利用筛选后的匹配点对计算所述第一位姿信息。

[0113] 具体地,假设所述第一参考帧图像的背景区域的特征点集合为 B_{i-1} ,所述当前帧图像的背景区域的特征点集合为 B_i ,将集合 B_{i-1} 和集合 B_i 中的特征点进行匹配,得到匹配点对集合,通过计算描述子间的距离,并采用RANSAC算法对所述匹配点对集合中的匹配点对进行二次筛选得到稳定的点集对应关系。所述第一位姿信息的计算方法可以包括:

[0114] 从背景区域的特征点集合 B_i 和 B_{i-1} 中确定匹配点对集合,从所述匹配点对集合中选择8个匹配点对,使用归一化八点法估算出基础矩阵 F_i ;

[0115] 计算所述匹配点对集合中的其余匹配点对到其对应极线的距离 d_n ,若 $d_n < d$ 则该点为内点,反之为外点,记下符合该条件的内点的个数为 m_i ,其中, d 为预设距离阈值;

[0116] 迭代 S 次,或者某次得到内点的数目 m_i 占集合总比例大于等于预设比例(例如95%),则停止迭代。选择 m_i 最大的基础矩阵 F_i 作为所述第一位姿。

[0117] S430:基于所述第一位姿信息和所述深度信息确定所述运动掩模区域的特征点的类型,所述类型包括动态特征点和静态特征点。

[0118] 在一个可能的实施例中,如图5所示,所述基于所述第一位姿信息和所述深度信息确定所述运动掩模区域的特征点的类型可以包括:

[0119] S431:根据所述第一位姿信息和所述深度信息计算所述运动掩模区域的特征点的运动得分。

[0120] S432:当所述运动得分小于预设阈值时,判定所述特征点为静态特征点。

[0121] S433:当所述运动得分大于或等于预设阈值时,判定所述特征点为动态特征点。

[0122] 在另一个可能的实施例中,所述根据所述第一位姿信息和所述深度信息计算所述运动掩模区域的特征点的运动得分可以包括:

[0123] 获取第一参考帧图像的运动掩模区域的特征点,将所述第一参考帧图像的运动掩模区域的特征点与所述当前帧图像的运动掩模区域的特征点进行匹配,得到匹配点对;其中,所述第一参考帧图像为所述当前帧图像的前一帧图像;

[0124] 对所述匹配点对进行筛选,去除误匹配的匹配点对;

[0125] 根据所述第一位姿信息和所述深度信息计算筛选后的匹配点对之间的距离,将所述距离作为所述运动掩模区域的特征点的运动得分。

[0126] 本发明实施例中,可以使用包括极线约束和深度约束等约束方式计算所述运动掩模区域的特征点的运动得分。在一些可能的实施例中,也可以采用其他约束方式,本发明对此不做限定。

[0127] 具体地,当获取到匹配点对之后,可以采用距离约束和RANSAC算法对上述的匹配点对进行筛选,得到经过筛选的匹配点对。

[0128] 假定 p_1 和 p_2 为运动掩模的特征点的一组匹配点对, p_1 为所述第一参考帧图像的运动掩模区域的特征点, p_2 为所述当前帧图像的运动掩模区域的特征点, d_1 和 d_2 为 p_1 和 p_2 在深度图上的对应的深度值。由多视角几何知识得到 p_1 和 p_2 应满足极线约束,其表达式如下:

$$[0129] \quad p_1^T F p_2 = 0$$

[0130] 其中,所述 F 为通过背景区域的特征点计算得到的基础矩阵。由于物体发生运动则其实际值不为0,因此得到误差距离。同时将第一参考帧图像特征点通过透视变换以及位姿变换并重投影可以得到当前的深度值,并将该深度值的差作为距离误差,其具体公式如下:

$$[0131] \quad D = \alpha \times p_1^T F p_2 + \beta \times (d(\pi(F\pi(p_1))) - d_2)$$

[0132] 其中, D 表示距离, ε 为对应的阈值,若计算的 D 值大于或等于 ε ,则判断该特征点 p_2 为动态特征点,若 D 值小于 ε 则判断该特征点 p_2 为静态特征点。

[0133] S440:剔除所述运动掩模区域的特征点中的动态特征点,保留静态特征点。

[0134] 本发明实施例中,可以将所述运动掩模区域的特征点计为一个集合,当判断特征点为动态特征点时,将所述特征点从所述集合中删除,当判断特征点为静态特征点时,则保留。

[0135] S450:根据所述静态特征点和所述背景区域的特征点生成静态特征点集合。

[0136] 本发明实施例中,可以遍历所有运动掩模区域的特征点,根据保留的静态特征点和所述背景区域的特征点生成静态特征点集合。

[0137] S260:根据所述静态特征点集合确定当前状态位姿信息。

[0138] 在一个可能的实施例中,所述根据所述静态特征点集合确定当前状态位姿信息可以包括:

[0139] 在产生新的关键帧时,建立所述静态特征点集合中的特征点、关键帧和地图点之间的数据关联;

[0140] 根据所述静态特征点集合中的特征点确定第二位姿信息;

[0141] 根据所述第二位姿信息和所述数据关联进行位姿优化,确定当前状态位姿信息。

[0142] 本发明实施例中,可以通过跟踪估计线程建立静态特征点集合中的特征点、关键帧和地图点之间的数据关联;通过局部地图线程进行位姿优化,确定当前状态位姿信息。具体地,可以根据所述静态特征点集合中的特征点以及与所述特征点匹配的第一参考帧图像中的特征点组成的匹配点对,采用归一化八点法的方式计算基础矩阵,得到第二位姿信息。可以将所述第二位姿信息作为初值利用局部捆集技术进行位姿优化,确定当前状态位姿信息。

[0143] 在一个可能的实施例中,所述方法还可以包括:

[0144] 基于所述目标语义图像、所述当前状态位姿信息及所述静态特征点集合生成静态物体点云数据;

[0145] 根据所述静态物体点云数据进行静态物体点云地图的稠密重建。

[0146] 本发明实施例中,可以通过稠密点云建图线程进行静态物体点云地图的稠密重建。

[0147] 在一个可能的实施例中,可以通过局部地图线程对所有关键帧进行局部优化和关键帧筛选,包括对关键帧和地图点进行对应的增添和删除,可以使得系统的估计更加精准和稳定,所述语义分割与补偿线程和局部地图线程可以进行信息交互,一起进行后端位姿的优化。可以通过回环检测线程对优化后的关键帧进行全局调整和回环检测。可以通过稠密点云建图线程进行位姿估计与动态区域判断,将静态背景部分生成局部点云并完成整体的拼接与更新;所述跟踪估计线程和所述语义分割与补偿线程和稠密点云建图线程可以进行信息交互,通过语义分割结果,将检测出的运动物体的部分的像素剔除,并结合位姿优化得到的精准位姿,通过ICP方法实现对静态物体的点云地图的稠密重建。

[0148] 在一个具体的实施例中,通过国际上开源的数据集,对本发明实施例的实际效果进行了验证。该数据集为慕尼黑工业大学(Technical University of Munich,TUM)开源的RGBD室内数据集,采用8台高速相机(100hz)拍摄,并通过运动捕捉系统提供其真实运动轨迹。该数据集包含原图和对应的深度图,并给出两者之间的对齐脚本,以及用以评估SLAM系统位姿估计准确性的评估脚本。本发明和ORB-SLAM2系统以及DS-SLAM系统进行对比,比较指标分别为绝对轨迹误差(Absolute Trajectory Error,ATE)和相对位置误差(Relative

Pose Error, RPE)。测试结果如图6A至图6C所示,其中,左图为其实际轨迹的对比结果,包含真值轨迹、轨迹轨迹以及两者之间的误差线段,右图是其PRE随时间变化的曲线,描述系统的相对稳定性。由结果可以看出本发明实施例的方法在效果上优于ORB-SLAM2系统和DS-SLAM系统,并且由于采用ENet网络因此系统可以达到22FPS,完全满足实时性的要求。

[0149] 综上所述,本发明的动态场景下的视觉定位方法具有以下有益效果:

[0150] 本发明的动态场景下的视觉定位方法通过对图像进行语义分割,并根据语义分割结果和深度信息进行运动一致性检测,剔除运动掩模区域的动态特征点,将运动掩模中的静态特征点加入静态特征点集合,并利用所述静态特征点集合实现位姿优化,并重建出稠密的静态物体点云地图。能够提高在动态环境中对位姿估计的精确性和鲁棒性,提高所述静态物体点云地图的准确度,从而提高对环境的感知的准确度。

[0151] 另外,本发明的动态场景下的视觉定位方法基于运动连续性假设,对相邻帧图像的语义分割结果进行融合,补偿漏分割的情况,能够进一步提高本方法的准确性。

[0152] 参考说明书附图7,本实施例提供一种动态场景下的视觉定位装置700,所述装置700可以包括:

[0153] 第一获取模块710,用于获取当前帧图像,提取所述当前帧图像的特征点;

[0154] 语义分割模块720,用于将所述当前帧图像输入预设深度学习网络进行语义分割,得到目标语义图像;

[0155] 确定模块730,用于根据所述目标语义图像确定所述当前帧图像的运动掩模区域;

[0156] 第二获取模块740,用于获取所述当前帧图像的深度信息;

[0157] 检测模块750,用于基于所述目标语义图像和所述深度信息进行运动一致性检测,确定所述当前帧图像的静态特征点集合;

[0158] 定位模块760,用于根据所述静态特征点集合确定当前状态位姿信息。

[0159] 需要说明的是,上述实施例提供的装置,在实现其功能时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将设备的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。

[0160] 上述说明已经充分揭露了本发明的具体实施方式。需要指出的是,熟悉该领域的技术人员对本发明的具体实施方式所做的任何改动均不脱离本发明的权利要求书的范围。相应地,本发明的权利要求的范围也并不仅仅局限于前述具体实施方式。

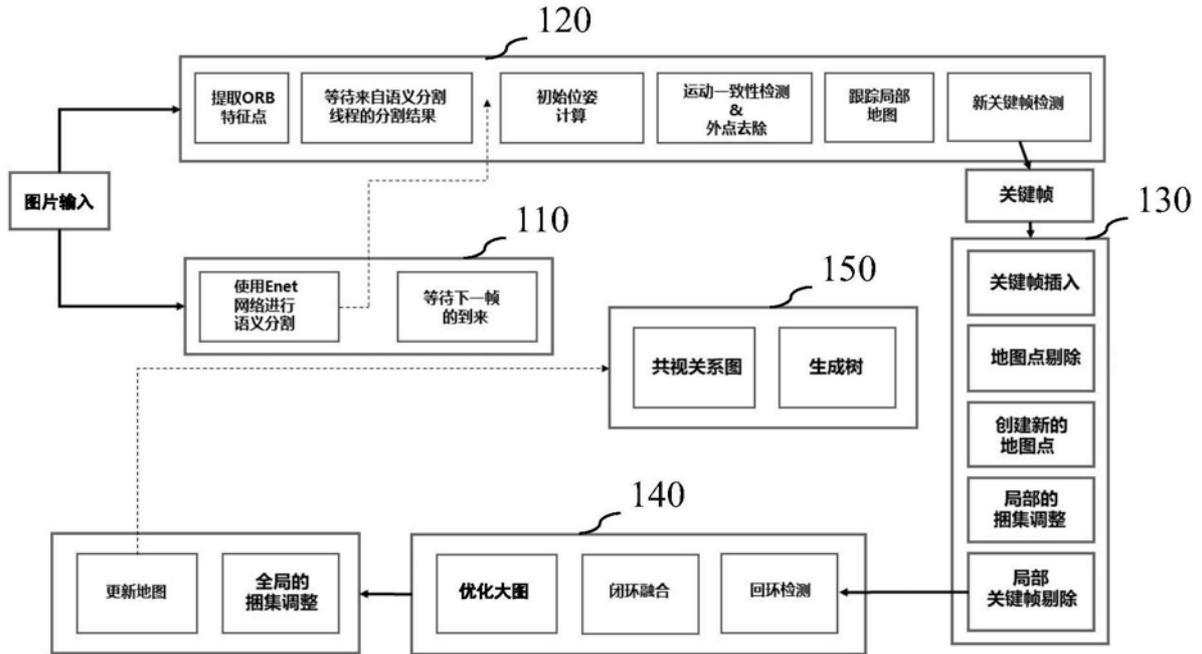


图1

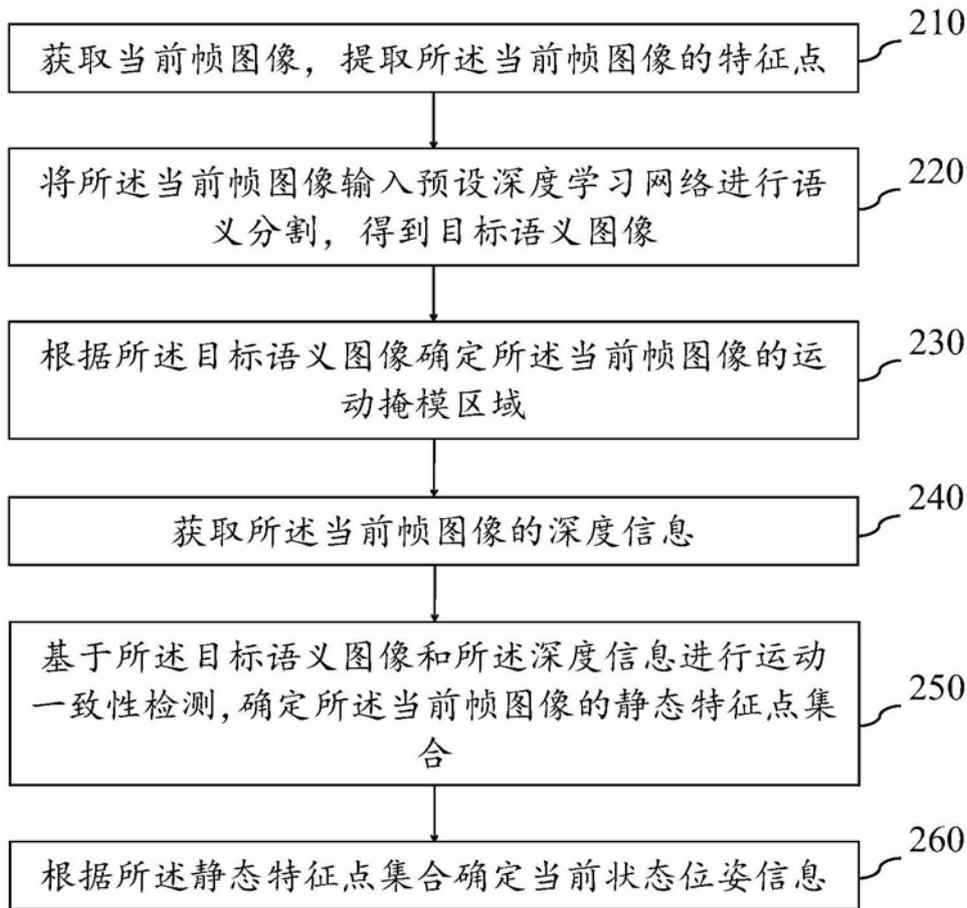


图2

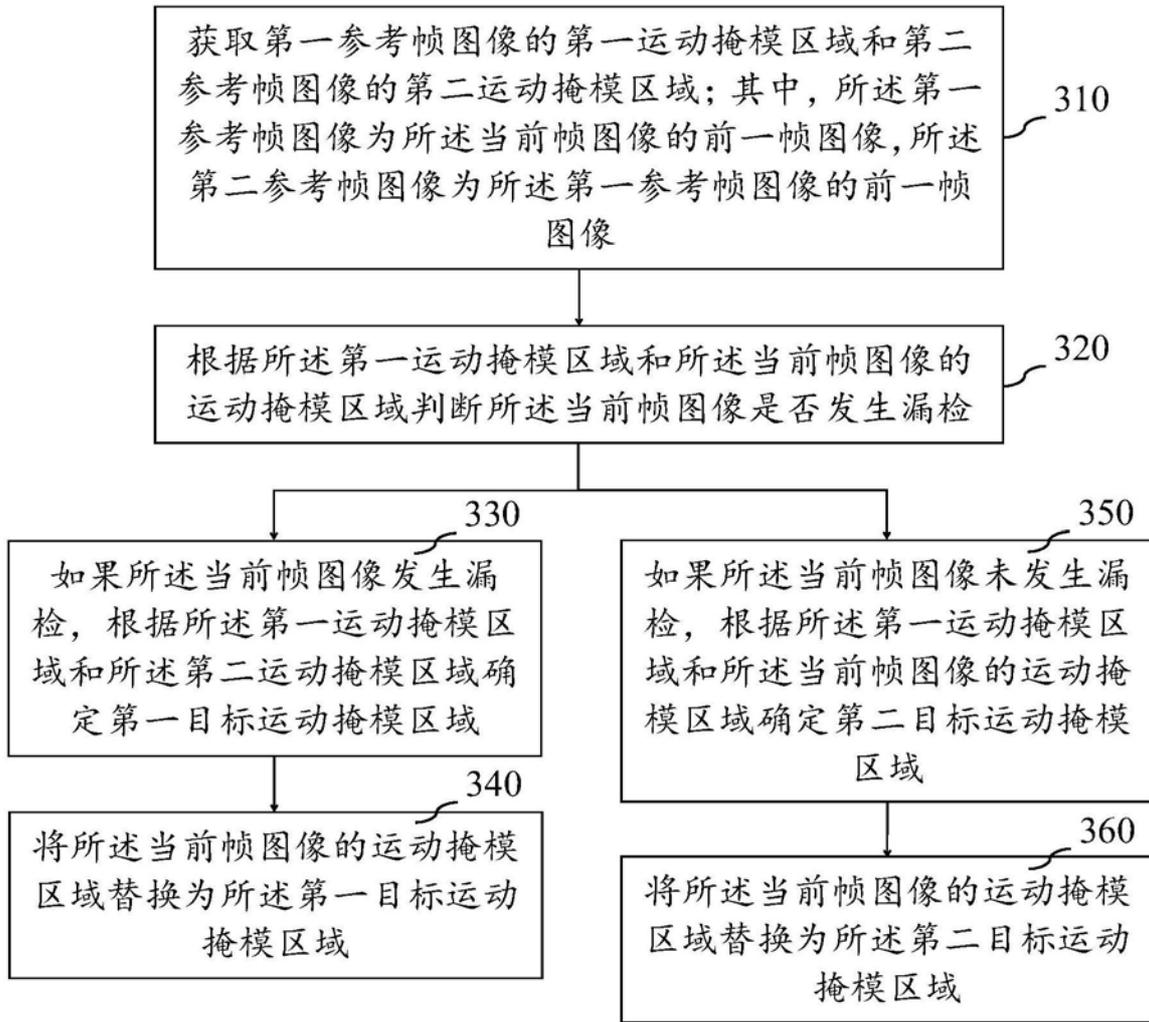


图3

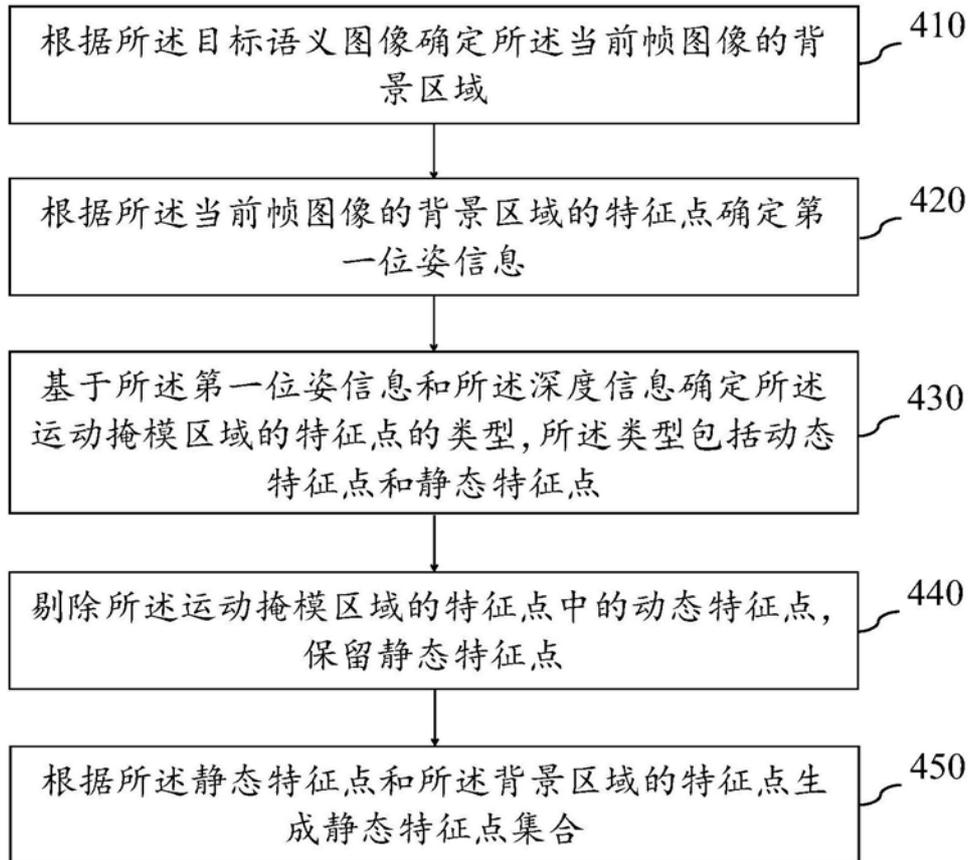


图4

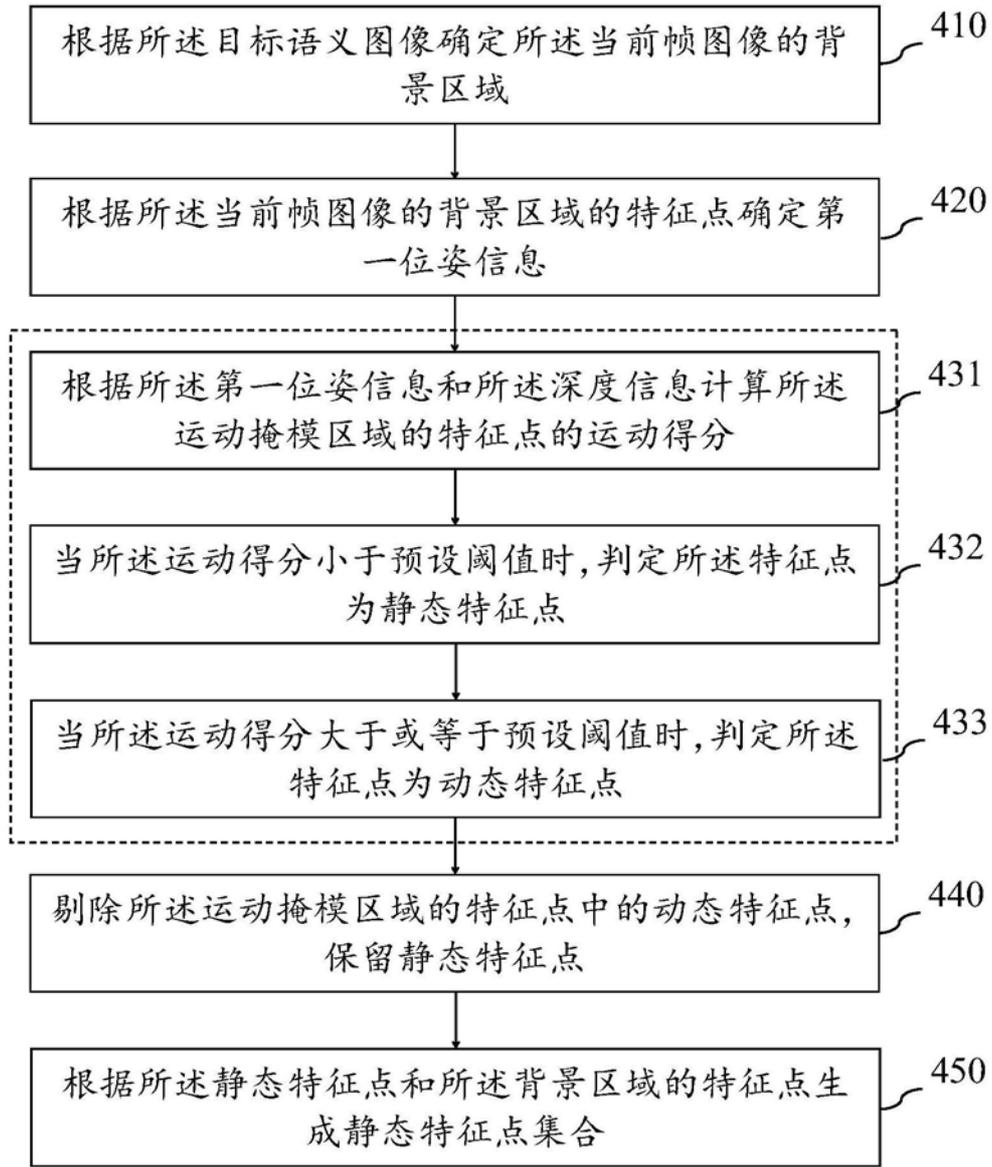


图5

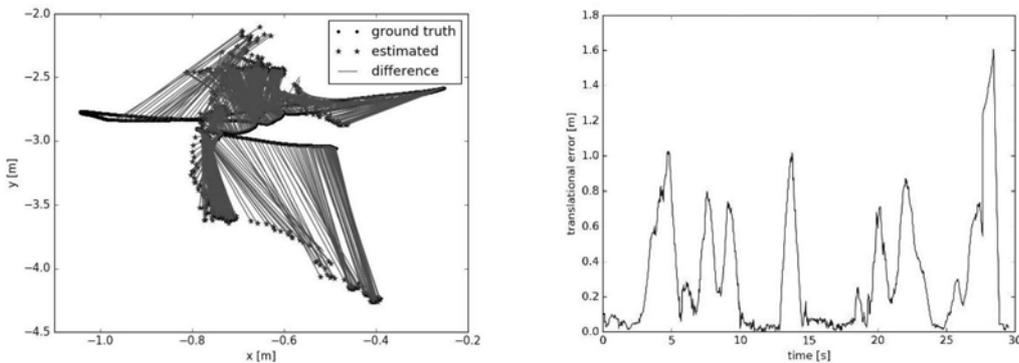


图6A

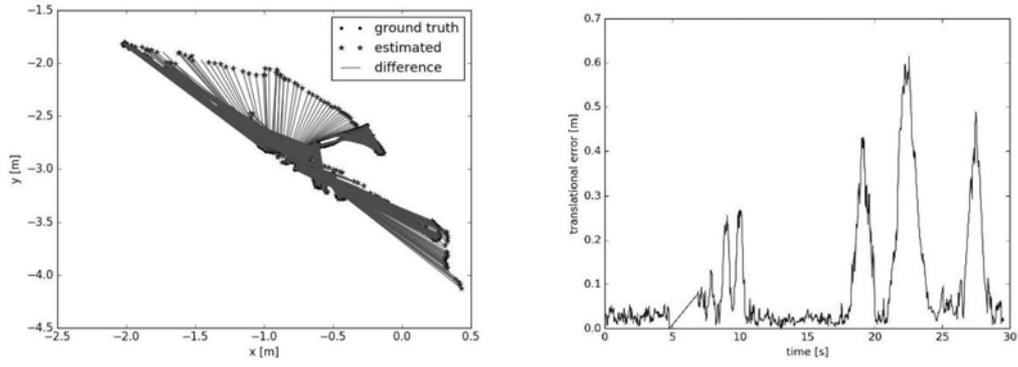


图6B

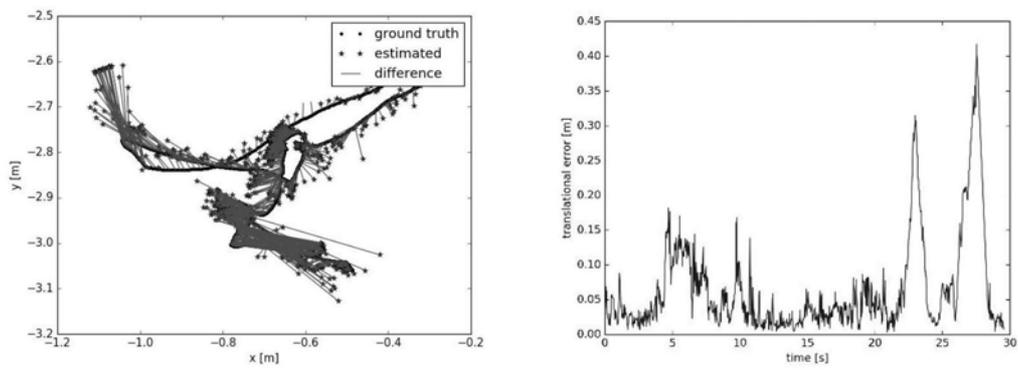


图6C

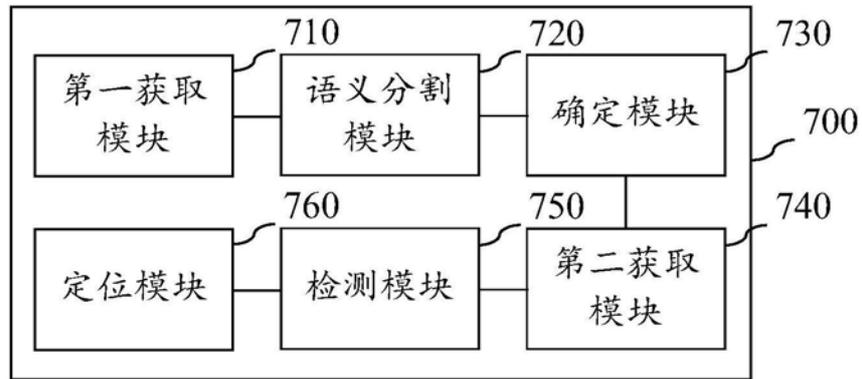


图7