



(12) 发明专利

(10) 授权公告号 CN 114140527 B

(45) 授权公告日 2024. 09. 10

(21) 申请号 202111373890.7
 (22) 申请日 2021.11.19
 (65) 同一申请的已公布的文献号
 申请公布号 CN 114140527 A
 (43) 申请公布日 2022.03.04
 (73) 专利权人 苏州科技大学
 地址 215009 江苏省苏州市高新区科锐路1号
 (72) 发明人 沈晔湖 李星 卢金斌 王其聪
 赵冲 蒋全胜 朱其新 谢鸥
 牛福洲 牛雪梅 付贵忠
 (74) 专利代理机构 苏州今迈知识产权代理事务所(特殊普通合伙) 32524
 专利代理师 张佩璇

(51) Int. Cl.
 G06T 7/73 (2017.01)
 G06T 7/215 (2017.01)
 G06T 7/80 (2017.01)
 G06N 3/08 (2023.01)
 G06N 3/0464 (2023.01)
 G06V 10/74 (2022.01)
 G06V 10/82 (2022.01)
 G06V 10/40 (2022.01)
 G01C 21/00 (2006.01)
 G06V 10/26 (2022.01)

(56) 对比文件

沈晔湖等. 基于几何-语义联合约束的动态环境视觉SLAM 算法. 数据采集与处理. 2022, 第37卷(第3期), 全文.

审查员 王焱

权利要求书2页 说明书7页 附图1页

(54) 发明名称

一种基于语义分割的动态环境双目视觉SLAM方法

(57) 摘要

本发明涉及一种基于语义分割的动态环境双目视觉SLAM方法, 包括如下步骤: 获取物体的语义掩膜, 所述语义掩膜通过深度学习网络生成; 采用双目相机获取多帧连续的双目图像; 提取每帧双目图像上的特征点, 匹配相邻帧双目图像上的特征点; 剔除位于语义掩膜上的特征点, 并根据剩余的特征点计算相机位姿; 基于所述相机位姿分离所述双目图像上的动态物体和静态物体; 基于分离后的静态物体重新计算相机位姿; 基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图。本发明使用双目相机, 将经过语义信息分割后的图像作为引导, 能够识别场景中的动态、静态物体, 并实现地图的构建, 本方法运算简单、成本低, 且能够应用于大多数实际场景中。



1. 一种基于语义分割的动态环境双目视觉SLAM方法,其特征在于,包括如下步骤:
 - 获取物体的语义掩膜,所述语义掩膜通过深度学习网络生成;
 - 采用双目相机获取多帧连续的双目图像;
 - 提取每帧所述双目图像上的特征点,匹配相邻帧双目图像上的特征点;
 - 剔除位于所述语义掩膜上的特征点,并根据剩余的特征点计算相机位姿;
 - 基于所述相机位姿分离所述双目图像上的动态物体和静态物体;
 - 基于分离后的动态物体估计动态物体的运动参数;
 - 基于分离后的静态物体重新计算相机位姿;
 - 基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图;所述基于所述相机位姿分离所述双目图像上的动态物体和静态物体;基于分离后的动态物体估计动态物体的运动参数的方法包括:
 - 分离动态物体:基于所述相机位姿,和相邻帧双目图像与所述语义掩膜的位置关系,计算该语义掩膜对应的物体的运动概率,若所述运动概率大于第一阈值,则判断该语义掩膜对应的物体为动态物体;
 - 动态物体匹配:针对所述动态物体,计算该动态物体对应的语义掩膜在相邻帧双目图像的hu矩、中心点欧氏距离以及直方图分布,基于所述hu矩、中心点欧氏距离和直方图分布计算相邻帧双目图像中所述动态物体匹配的概率,若概率大于第二阈值,则相邻帧双目图像中的两个动态物体为同一物体;以及
 - 动态物体运动估计:通过所述动态物体匹配完成连续帧之间动态物体的关联,通过PnP算法估计所述动态物体的运动参数;所述基于分离后的静态物体重新计算相机位姿的方法包括:剔除位于所述动态物体对应的语义掩膜上的特征点,根据剩余的特征点,采用PnP算法更新所述相机位姿;
 - 所述基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图的方法包括:
 - 基于更新后的相机位姿和位于所述静态物体上的特征点确定多个关键帧;
 - 匹配多个所述关键帧上的特征点,剔除不匹配的特征点;
 - 检验匹配后的特征点是否满足对极几何约束,剔除不满足的特征点;
 - 检查剩余特征点正向景深、视差、反投影误差和尺度是否一致,剔除不一致的特征点,基于剩余特征点生成地图点;
 - 基于所述地图点构建所述静态地图。
2. 如权利要求1所述的基于语义分割的动态环境双目视觉SLAM方法,其特征在于,用于生成所述语义掩膜的深度学习网络为Mask R-CNN模型。
 3. 如权利要求1所述的基于语义分割的动态环境双目视觉SLAM方法,其特征在于,所述提取每帧所述双目图像上的特征点,匹配相邻帧双目图像上的特征点的方法包括:
 - 采用ORB方法提取所述特征点;
 - 获取每个特征点分别在各帧双目图像上的描述子,计算一个特征点在位于相邻两帧双目图像上的两个描述子之间的汉明距离,汉明距离最小的两个特征点组成一组匹配的特征点。
 4. 如权利要求1所述的基于语义分割的动态环境双目视觉SLAM方法,其特征在于,判断

所述特征点是否位于所述语义掩膜上的方法包括:所述语义掩膜至少包括物体的边框,所述特征点的坐标位于所述边框范围内,则该特征点位于该语义掩膜上。

5.如权利要求1所述的基于语义分割的动态环境双目视觉SLAM方法,其特征在于,所述根据剩余的特征点计算相机位姿的方法包括:采用PnP算法求解所述相机位姿。

6.如权利要求1所述的基于语义分割的动态环境双目视觉SLAM方法,其特征在于,所述分离动态物体的步骤包括:

基于所述相机位姿,计算上一帧的语义掩膜在当前帧对应的位置;

利用视差图计算投影以后所述语义掩膜上所有特征点的三维坐标,所述视差图通过所述双目图像计算所得;

计算上一帧与当前帧对应特征点在x,y,z三个方向上的误差,其中误差的最大值作为该特征点的误差值;

将所述误差值转换为与该特征点所在的语义掩膜对应的物体的运动概率,基于所述运动概率判断与该语义掩膜对应的物体是否为动态物体。

7.如权利要求1所述的基于语义分割的动态环境双目视觉SLAM方法,其特征在于,在构建所述静态地图前,还包括对生成的地图点通过光束法平差进行优化的步骤。

一种基于语义分割的动态环境双目视觉SLAM方法

技术领域

[0001] 本发明涉及视觉空间定位技术领域,尤其涉及一种基于语义分割的动态环境双目视觉SLAM方法。

背景技术

[0002] 随着计算机技术和人工智能的发展,智能自主移动机器人成为机器人领域的一个重要研究方向和研究热点。随着移动机器人逐渐智能化,移动机器人对自身的定位以及环境地图的要求越来越高。目前,智能移动机器人在已知环境中完成自身定位和地图构建已经有了一些实用的应用,但是在未知环境中仍有很多挑战。在这样的环境中完成定位和地图构建的技术被称为SLAM(Simultaneous Localization and Mapping),即同步定位和地图构建,SLAM的目标是使机器人能够在未知环境的移动过程中,完成自身定位和增量式地图构建。

[0003] 传统的SLAM算法主要依赖于稳定性较好的距离传感器,如激光雷达。然而激光雷达获得的距离数据非常稀疏,这就造成SLAM构建得到的环境地图仅包含极少量的路标点。这个地图仅能被用来提高机器人的定位精度,而无法用于路径规划等机器人导航的其它领域。此外激光雷达高昂的价格、较大的体积重量以及耗电量限制了其在某些领域的应用。相机虽然能在一定程度上克服激光雷达在价格、体积、质量以及耗电量上的劣势,同时相机能够获取丰富的信息,但是相机也存在一些问题,例如对光线变化敏感,运算复杂度高等。目前还有多传感器融合的SLAM算法,虽然能够有效缓解单一传感器自身不足带来的问题,但是也进一步增加了成本和算法的复杂程度。

[0004] 现有的视觉SLAM算法大多基于环境静态假设,即场景是静态的,不存在相对运动的物体。但是在实际室外场景中大量存在行人、车辆等动态物体,从而限制了基于上述假设的SLAM系统在实际场景中运用。针对动态环境下视觉SLAM算法的定位精度和稳定性下降的问题,现有的算法使用了一些基于概率统计或者几何约束的算法,减少了动态物体对视觉SLAM算法精度和稳定性的影响。例如当场景中存在少量动态物体时,可以使用RANSAC(Random Sample Consensus)等概率算法来剔除动态物体。但是当场景中大量动态物体时,上述算法将无法正常区分动态物体。而另一些算法使用光流法来区分动态物体,在存在大量动态物体的场景中,使用光流法确实能够区分动态物体,但是由于计算稠密光流的过程较为耗时,会降低SLAM算法的执行效率。

[0005] 因此,如何提供一种运算简单、成本低,且能够应用于多数实际场景中的基于语义分割的动态环境双目视觉SLAM方法是本领域技术人员亟待解决的一个技术问题。

发明内容

[0006] 本发明提供一种基于语义分割的动态环境双目视觉SLAM方法,以解决上述技术问题。

[0007] 为解决上述技术问题,本发明提供一种基于语义分割的动态环境双目视觉SLAM方

法,包括如下步骤:

[0008] 获取物体的语义掩膜,所述语义掩膜通过深度学习网络生成;

[0009] 采用双目相机获取多帧连续的双目图像;

[0010] 提取每帧所述双目图像上的特征点,匹配相邻帧双目图像上的特征点;

[0011] 剔除位于所述语义掩膜上的特征点,并根据剩余的特征点计算相机位姿;

[0012] 基于所述相机位姿分离所述双目图像上的动态物体和静态物体;

[0013] 基于分离后的动态物体估计动态物体的运动参数;

[0014] 基于分离后的静态物体重新计算相机位姿;

[0015] 基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图。

[0016] 较佳地,用于生成所述语义掩膜的深度学习网络为Mask R-CNN模型。

[0017] 较佳地,所述提取每帧所述双目图像上的特征点,匹配相邻帧双目图像上的特征点的方法包括:

[0018] 采用ORB方法提取所述特征点;

[0019] 获取每个特征点分别在各帧双目图像上的描述子,计算一个特征点在位于相邻两帧双目图像上的两个描述子之间的汉明距离,汉明距离最小的两个特征点组成一组匹配的特征点。

[0020] 较佳地,判断所述特征点是否位于所述语义掩膜上的方法包括:所述语义掩膜至少包括物体的边框,所述特征点的坐标位于所述边框范围内,则该特征点位于该语义掩膜上。

[0021] 较佳地,所述根据剩余的特征点计算相机位姿的方法包括:采用PnP算法求解所述相机位姿。

[0022] 较佳地,所述基于所述相机位姿分离所述双目图像上的动态物体和静态物体;基于分离后的动态物体估计动态物体的运动参数的方法包括:

[0023] 分离动态物体:基于所述相机位姿,和相邻帧双目图像与所述语义掩膜的位置关系,计算该语义掩膜对应的物体的运动概率,若所述运动概率大于第一阈值,则判断该语义掩膜对应的物体为动态物体;

[0024] 动态物体匹配:针对所述动态物体,计算该动态物体对应的语义掩膜在相邻帧双目图像的hu矩、中心点欧氏距离以及直方图分布,基于所述hu矩、中心点欧氏距离和直方图分布计算相邻帧双目图像中所述动态物体匹配的概率,若概率大于第二阈值,则相邻帧双目图像中的两个动态物体为同一物体;以及

[0025] 动态物体运动估计:通过所述动态物体匹配完成连续帧之间动态物体的关联,通过PnP算法估计所述动态物体的运动参数。

[0026] 较佳地,所述分离动态物体的步骤包括:

[0027] 基于所述相机位姿,计算上一帧的语义掩膜在当前帧对应的位置;

[0028] 利用视差图计算投影以后所述语义掩膜上所有特征点的三维坐标,所述视差图通过所述双目图像计算所得;

[0029] 计算上一帧与当前帧对应特征点在x,y,z三个方向上的误差,其中误差的最大值作为该特征点的误差值;

[0030] 将所述误差值转换为与该特征点所在的语义掩膜对应的物体的运动概率,基于所

述运动概率判断与该语义掩膜对应的物体是否为动态物体。

[0031] 较佳地,所述基于分离后的静态物体重新计算相机位姿的方法包括:剔除位于所述动态物体对应的语义掩膜上的特征点,根据剩余的特征点,采用PnP算法更新所述相机位姿。

[0032] 较佳地,所述基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图的方法包括:

[0033] 基于更新后的相机位姿和位于所述静态物体上的特征点确定多个关键帧;

[0034] 匹配多个所述关键帧上的特征点,剔除不匹配的特征点;

[0035] 检验匹配后的特征点是否满足对极几何约束,剔除不满足的特征点;

[0036] 检查剩余特征点正向景深、视差、反投影误差和尺度是否一致,剔除不一致的特征点,基于剩余特征点生成地图点;

[0037] 基于所述地图点构建所述静态地图。

[0038] 较佳地,在构建所述静态地图前,还包括对生成的地图点通过光束法平差进行优化的步骤。

[0039] 与现有技术相比,本发明提供的基于语义分割的动态环境双目视觉SLAM方法使用双目相机,将经过语义信息分割后的图像作为引导,能够识别场景中的动态、静态物体,并实现地图的构建,本方法运算简单、成本低,且能够应用于大多数实际场景中。

附图说明

[0040] 图1为本发明一具体实施方式中基于语义分割的动态环境双目视觉SLAM方法的流程示意图;

[0041] 图2为本发明一具体实施方式中分离动态物体的流程示意图。

具体实施方式

[0042] 为了更详尽的表述上述发明的技术方案,以下列举出具体的实施例来证明技术效果;需要强调的是,这些实施例用于说明本发明而限于限制本发明的范围。

[0043] 本发明提供的基于语义分割的动态环境双目视觉SLAM方法,如图1所示,包括如下步骤:

[0044] 获取物体的语义掩膜,所述语义掩膜通过深度学习网络生成,本实施例中,用于生成所述语义掩膜的深度学习网络为Mask R-CNN模型,从而实现高质量的语义分割。

[0045] 采用双目相机获取多帧连续的双目图像,从所述双目图像中可以获取二维图像像素点的第三维深度信息,当然,关于所述双目相机的内参以及外参,主要包括:相机焦距 f 、相机光心 (u, v) 、相机镜头径向畸变系数 k_{c1} 和 k_{c2} 等,上述参数可通过张正友标定法标定获得。

[0046] 提取每帧所述双目图像上的特征点,匹配相邻帧双目图像上的特征点。具体方法包括:

[0047] 采用ORB(英文全称:Oriented Fast and Rotated Brief)方法提取所述特征点;

[0048] 获取每个特征点分别在各帧双目图像上的描述子,计算一个特征点在位于相邻两帧双目图像上的两个描述子之间的汉明距离,汉明距离最小的两个特征点组成一组匹配的

特征点。

[0049] 剔除位于所述语义掩膜上的特征点,并根据剩余的特征点计算相机位姿。判断所述特征点是否位于所述语义掩膜上的方法包括:所述语义掩膜至少包括物体的边框,所述特征点的坐标位于所述边框范围内,则该特征点位于该语义掩膜上;若不位于所述边框范围内,则该特征点不位于该语义掩膜上。所述根据剩余的特征点计算相机位姿的方法包括:采用PnP(英文全称Perspective-n-Point)算法求解所述相机位姿,构建重投影误差并优化如下式(1)所示:

$$[0050] \quad E(T) = \arg \min_T \frac{1}{2} \sum_{i=1}^n \left\| \mathbf{u}_i - \frac{1}{s_i} \mathbf{K} \mathbf{T} \mathbf{p}_i \right\|_2^2 \quad (1)$$

[0051] 通过最小化重投影误差,得到最优解,即为所需的相机位姿。

[0052] 基于所述相机位姿分离所述双目图像上的动态物体和静态物体,具体方法包括:

[0053] 分离动态物体:基于所述相机位姿,和相邻帧双目图像与所述语义掩膜的位置关系,计算该语义掩膜对应的物体的运动概率,若所述运动概率大于第一阈值,则判断该语义掩膜对应的物体为动态物体。具体步骤如图2所示,包括:

[0054] 基于所述相机位姿,计算上一帧的语义掩膜在当前帧对应的位置;

[0055] 利用视差图计算投影以后所述语义掩膜上所有特征点的三维坐标,所述视差图通过所述双目图像计算所得,具体可以采用ELAS(英文全称:Efficient Large Scale Stereo Matching)算法计算视差图;

[0056] 计算上一帧与当前帧对应特征点在x,y,z三个方向上的误差,其中误差的最大值作为该特征点的误差值;

[0057] 将所述误差值转换为与该特征点所在的语义掩膜对应的物体的运动概率,基于所述运动概率判断与该语义掩膜对应的物体是否为动态物体。

[0058] 由相机成像原理可知,三维坐标系和像素(二维)坐标系之间转换关系以及深度和视差转换为:

$$[0059] \quad Z = \frac{bf}{d} \quad (2)$$

$$[0060] \quad Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3)$$

[0061] 记t-1帧第j个语义掩膜在像素坐标系上的坐标集合为 \mathbf{M}_j^{t-1} ,通过公式(2)和公式(3),得到该时刻该语义掩膜的三维坐标集合 \mathbf{P}_j^{t-1} 。

[0062] 通过公式(4)得到运动过后三维点集合 ${}^{pro}\mathbf{P}_j^t$,

$$[0063] \quad {}^{pro}\mathbf{P}_j^t = \mathbf{R}\mathbf{P}_j^{t-1} + \mathbf{t} \quad (4)$$

[0064] 通过公式(3)得到 ${}^{pro}\mathbf{P}_j^t$ 转换到像素坐标系下的集合 ${}^{pro}\mathbf{M}_j^t$,然后利用 ${}^{pro}\mathbf{M}_j^t$ 和视差图通过公式(2)和公式(3)计算得到 ${}^{com}\mathbf{P}_j^t$ 。

[0065] 记 ${}^{pro}i\mathbf{p}_j^t$ 为 ${}^{pro}\mathbf{P}_j^t$ 中第 i 个点, ${}^{com}i\mathbf{p}_j^t$ 为 ${}^{pro}\mathbf{M}_j^t$ 第 i 个点, 计算两点之间误差 Δi 为:

$$[0066] \quad \Delta_j^i = \begin{bmatrix} \Delta X_j^i \\ \Delta Y_j^i \\ \Delta Z_j^i \end{bmatrix} = {}^{pro}i\mathbf{p}_j^t - {}^{com}i\mathbf{p}_j^t \quad (5)$$

[0067] 则该特征点对应的物体的误差为:

$$[0068] \quad \Delta_j = \frac{\sum_{i=1}^n \text{Max}(\|\Delta X_j^i\|, \|\Delta Y_j^i\|, \|\Delta Z_j^i\|)}{n} \quad (6)$$

[0069] 计算的运动概率 $S(\Delta_j)$ 即为:

$$[0070] \quad S(\Delta_j) = \frac{1}{1 + e^{-a(\Delta_j - b)}} \quad (7)$$

[0071] 动态物体匹配: 针对所述动态物体, 计算该动态物体对应的语义掩膜在相邻帧双目图像的hu矩(即图像矩)、中心点欧氏距离以及直方图分布, 基于所述hu矩、中心点欧氏距离和直方图分布计算相邻帧双目图像中所述动态物体匹配的概率, 若概率大于第二阈值, 则相邻帧双目图像中的两个动态物体为同一物体。具体地, 图像的hu矩是一种具有平移、旋转和尺度不变性的图像特征。

[0072] 图像和普通矩计算公式如下:

$$[0073] \quad m_{pq} = \sum_{x=1}^C \sum_{y=1}^R x^p y^q f(x, y), p, q = 0, 1, 2, \dots \quad (8)$$

[0074] 计算hu矩需要计算中心距, 首先计算质心坐标:

$$[0075] \quad x_0 = \frac{m_{10}}{m_{00}}$$

$$[0076] \quad y_0 = \frac{m_{01}}{m_{00}} \quad (9)$$

[0077] 然后构造中心矩:

$$[0078] \quad \eta_{pq} = \frac{u_{pq}}{u_{00}r} \left(r = \frac{q+p}{2} \right) \quad (10)$$

[0079] 接着对中心距进行归一化:

$$[0080] \quad \eta_{pq} = \frac{u_{pq}}{u_{00}r} \left(r = \frac{q+p}{2} \right) \quad (11)$$

[0081] 通过中心矩构造hu矩, hu矩具有7个不变矩, 具体公式如下:

$$[0082] \quad \Phi_1 = \eta_{20} + \eta_{02}$$

$$[0083] \quad \Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$[0084] \quad \Phi_3 = (\eta_{20} - 3\eta_{12})^2 + 3(\eta_{21} - \eta_{03})^2$$

$$[0085] \quad \Phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$[0086] \quad \Phi_5 = (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})] [3(\eta_{30} + \eta_{12})^2 -$$

$$\begin{aligned}
& (\eta_{21} + \eta_{03})^2 \\
[0087] \quad & \Phi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11} (\eta_{30} + \eta_{12}) (\eta_{21} + \eta_{03}) \\
[0088] \quad & \Phi_7 = (3\eta_{21} - \eta_{03}) (\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30}) (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (12)
\end{aligned}$$

[0089] 记 H_j^{t-1} 为第 $t-1$ 帧的 j 个语义掩膜的hu矩,两个语义掩膜之间hu矩的距离为:

$$[0090] \quad G_{ij} = \frac{H_j^{t-1} \cdot H_i^t}{\|H_j^{t-1}\| \|H_i^t\|} \quad (13)$$

[0091] 计算每个语义掩膜的中心位置,然后计算前后帧之间各个语义掩膜的中心点位置的欧氏距离,记为:

$$[0092] \quad E_{ij} = \|O_j^{t-1} - O_i^t\| \quad (14)$$

[0093] 计算该语义掩膜的直方图分布,然后归一化,记为 $R_j^{t-1}(x)$,然后计算前后帧不同语义掩膜的K1散度(英文全称:Kullback-Leibler divergence,又称相对熵:relative entropy)。

$$[0094] \quad KL_{ij} = KL(R_j^{t-1} \| R_i^t) = \sum_{x \in X} R_j^{t-1}(x) \log \frac{R_j^{t-1}(x)}{R_i^t(x)} \quad (15)$$

[0095] 联合hu矩、欧氏距离和直方图,估计匹配概率:

$$[0096] \quad B_{ij} = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(G_{ij}-\mu_1)^2}{2\sigma_1^2}} * \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(E_{ij}-\mu_2)^2}{2\sigma_2^2}} * \frac{1}{\sqrt{2\pi}\sigma_3} e^{-\frac{(KL_{ij}-\mu_3)^2}{2\sigma_3^2}} \quad (16)$$

[0097] 基于分离后的动态物体估计动态物体的运动参数的方法包括:动态物体运动估计:通过所述动态物体匹配完成连续帧之间动态物体的关联,通过PnP算法估计所述动态物体的运动参数。

[0098] 基于分离后的静态物体重新计算相机位姿,具体的方法包括:剔除位于所述动态物体对应的语义掩膜上的特征点,根据剩余的特征点,采用PnP算法更新所述相机位姿,具体计算方法可参考前述首次计算相机位姿的方法。

[0099] 基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图,具体方法包括:

[0100] 基于更新后的相机位姿和位于所述静态物体上的特征点确定多个关键帧;

[0101] 匹配多个所述关键帧上的特征点,然后将匹配到的特征点进行三角化,对于没有匹配上的点,再与其他关键帧中未被匹配的特征点进行匹配,直至找到全部的匹配的特征点,剔除不匹配的特征点;

[0102] 检验匹配后的特征点是否满足对极几何约束,剔除不满足的特征点;

[0103] 检查剩余特征点正向景深、视差、反投影误差和尺度是否一致,剔除不一致的特征点,基于剩余特征点生成地图点;

[0104] 基于所述地图点构建所述静态地图。

[0105] 较佳地,在构建所述静态地图前,还包括对生成的地图点通过光束法平差(BA,英文全称:bundle adjustment)进行优化的步骤。

[0106] 上述方法通过对双目图像的处理,识别出双目图像中存在的动态物体,并且估算出相机位姿以及动态物体的位姿,构建环境地图,满足移动机器人对三维地图的需求。

[0107] 综上所述,本发明提供的基于语义分割的动态环境双目视觉SLAM方法,包括如下步骤:获取物体的语义掩膜,所述语义掩膜通过深度学习网络生成;采用双目相机获取多帧连续的双目图像;提取每帧所述双目图像上的特征点,匹配相邻帧双目图像上的特征点;剔除位于所述语义掩膜上的特征点,并根据剩余的特征点计算相机位姿;基于所述相机位姿分离所述双目图像上的动态物体和静态物体;基于分离后的动态物体估计动态物体的运动参数;基于分离后的静态物体重新计算相机位姿;基于更新后的相机位姿以及位于所述静态物体上的特征点构建静态地图。本发明使用双目相机,将经过语义信息分割后的图像作为引导,能够识别场景中的动态、静态物体,并实现地图的构建,本方法运算简单、成本低,且能够应用于大多数实际场景中。

[0108] 显然,本领域的技术人员可以对发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包括这些改动和变型在内。

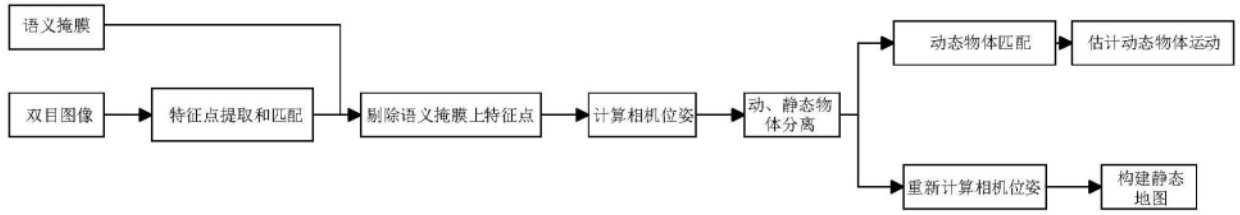


图1

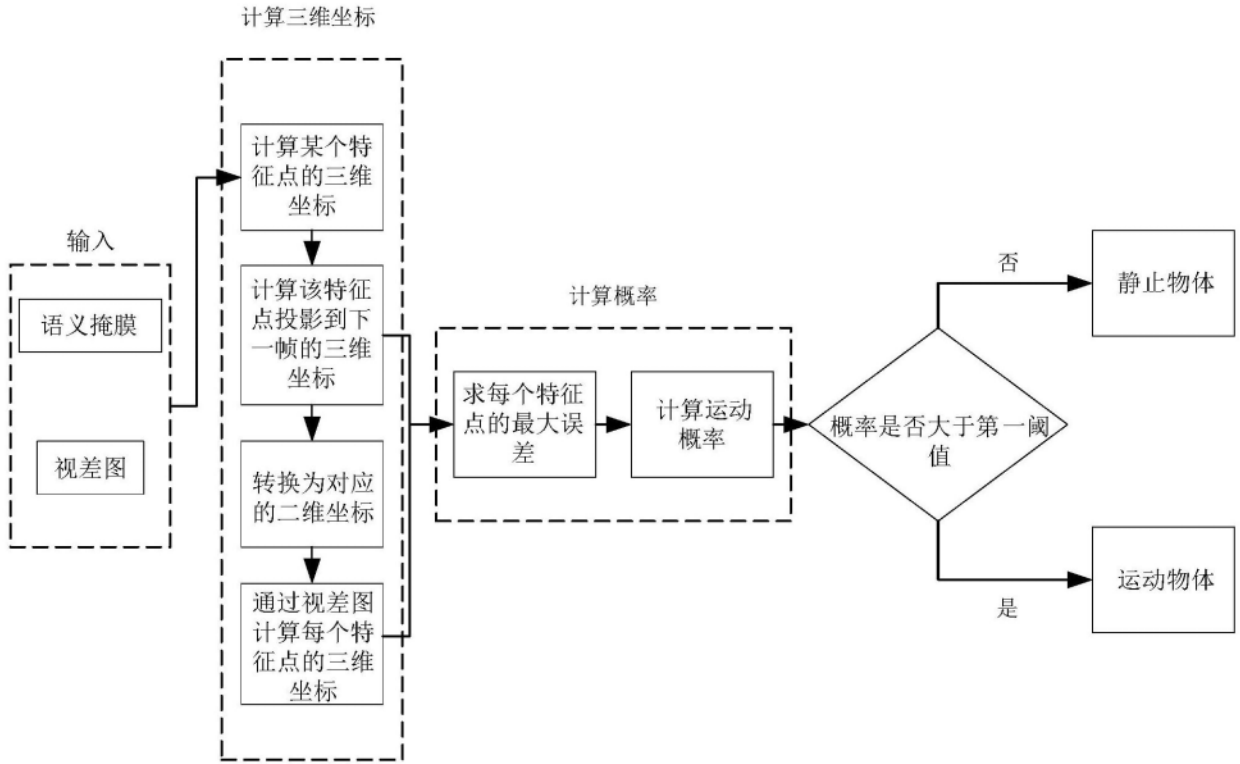


图2