文章编号:1674-2974(2024)08-0185-13 DOI:10.16339/j.cnki.hdxbzkb.2024189

地下停车场环境融合语义特征的视觉惯性定位方法

秦兆博[1,2],李琦[1],邢喆[1],高铭[1,2†],谢国涛[1,2],王晓伟[1,2]
(1. 湖南大学 机械与运载工程学院, 湖南 长沙 410082;
2. 湖南大学 无锡智能控制研究院, 江苏 无锡 214115)

摘 要:针对地下停车场环境GPS信号差、光线暗、特征少、纹理弱等带来的定位问题,提出了一种融合语义信息的视觉惯性定位算法.该算法首先通过视觉里程计和IMU预积分进行视觉惯性信息的融合.同时,利用4个鱼眼摄像头输入图像构建全景环视图像,并采用语义分割算法提取停车场环境语义信息.然后,根据视觉惯性紧耦合位姿完成逆投影变换,获得语义特征投影地图并采用回环检测和位姿图优化方式减小累积误差,完成全局位姿图优化,实现较高精度的定位效果.最后,通过Gazebo仿真与实车测试对该算法进行了验证.结果表明,本文算法能充分利用环境语义信息构建较为完整的语义地图,且基于重复定位误差对比,相较于ORB-SLAM3提高了车辆定位精度.

关键词:自主代客泊车;视觉惯性定位;多鱼眼全景环视;语义地图
中图分类号:TP311 **文献标志码**:A

A Visual Inertial Localization Method Integrating Semantic Features in Underground Parking Environment

QIN Zhaobo[1,2], LI Qi[1], XING Zhe[1], GAO Ming[1,2†], XIE Guotao[1,2], WANG Xiaowei[1,2]
[1. College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China;
2. Wuxi Intelligent Control of Research Institute, Hunan University, Wuxi 214115, China]

Abstract: A visual inertial localization algorithm integrating semantic information is proposed to address the positioning problems caused by poor GPS signals, dim lighting, limited features, and weak textures in underground parking lots. Firstly, this algorithm fuses visual inertial information through visual odometry and IMU pre-integration. Simultaneously, a panoramic surround view image is constructed using four fisheye cameras, and semantic segmentation algorithms are employed to extract semantic information from the parking environment. Then, the semantic feature projection map is obtained through inverse projection transformation based on the tightly coupled visual inertial pose. Additionally, loop detection and pose graph optimization are employed to reduce accumulated errors and achieve global pose graph optimization, thereby achieving higher localization accuracy. This paper verifies the proposed algorithm through Gazebo simulation and real vehicle testing. The results indicate that

* **收稿日期**:2023-06-09
基金项目:国家重点研发计划项目(2021YFB2501803),National Key Research and Development Program of China(2021YFB2501803);湖南省青年科技创新人才资助项目(2022RC1033),Science and Technology Innovation Program of Hunan Province (2022RC1033)
作者简介:秦兆博 (1990—),男,黑龙江大庆人,湖南大学副研究员,博士
† 通信联系人,E-mail: lizab314@163.com

this algorithm can fully utilize the semantic information of the environment to construct a complete semantic map and achieve higher vehicle localization accuracy than ORB-SLAM3 based on repeated localization error comparisons.

Key words: autonomous valet parking; visual inertial localization; multi-fisheye panoramic surround view; semantic mapping

近年来,随着汽车行业的蓬勃发展以及国内人均汽车保有量的逐渐增多,泊车已成为广大司机面临的严峻问题.自主代客泊车(Autonomous Valet Parking, AVP)技术可以充分利用有限的停车资源,提高停车位利用率,缓解城市停车难问题.建图定位是AVP系统不可或缺的核心模块,且其应用的场景大多数为地下停车场环境,考虑环境的独特性,卫星导航信号弱,同时定位与建图(Simultaneous Localization and Mapping, SLAM)成为解决地下停车场环境建图定位问题的有效方案.其中,基于相机传感器的视觉SLAM技术具有成本低廉、适用范围广泛 、功耗低等特点,受到了广泛关注.

传统视觉SLAM依赖相机的图像捕捉,在高速运动或快速旋转的情况下容易失效,而惯性测量单元(Inertial Measurement Unit, IMU)能高频获得车辆当前运动状态,不受周围环境的影响,能有效弥补相机在地下停车场环境的局限性,提高SLAM系统的稳定性和鲁棒性[1].Mur-Artal等人[2-3]提出了ORB-SLAM和ORB-SLAM2系统,将特征点逆深度优化、g2o第三方库图优化和基于DBow2词袋模型的回环检测融入算法中,使得ORB-SLAM成为首个完整的视觉SLAM解决方案.Campos等[4]扩展了ORB-SLAM相关工作[5],在纯视觉SLAM基础上提出ORB-SLAM3.该系统支持视觉与惯性融合定位,将IMU信息加入算法中耦合求解,改善了定位算法性能.Qin等[6]提出了一种基于单目相机和惯性测量单元(IMU)的视觉惯性紧耦合SLAM系统:VINS-mono,不仅仅使用视觉信息进行建图和定位,还利用IMU提供的加速度计和陀螺仪数据进行运动估计.

此外,随着深度学习技术的不断发展[7-8],融合深度学习的语义SLAM成为热门研究领域.相较于传统视觉SLAM方法利用环境中点、线、面等低维度几何特征信息进行定位建图的思路,语义SLAM通过深度学习提取环境中的语义信息,以获得高维度的复杂环境特征,进而可以在定位精度、鲁棒性、地图可视化等方面提高SLAM系统的整体性能[9].因此,结合深度学习的语义SLAM定位技术被广泛应用于AVP系统中.

Hu等[10]提出了利用道路标记语义分割来解决长期而高精度的视觉定位问题.基于道路标志分割的语义数据关联,设计了点云融合和环路检测策略来提高语义地图构建的性能,并提出了一种结合语义和几何推理的点云配准算法,以提高定位精度.2020年,Qin等[11]针对自主代客泊车的定位问题,提出了一种基于停车场地面语义特征的定位方法,通过改进语义分割Unet模型进行停车场中车位线、车道线 、箭头符号等语义标志物的分割,然后构建停车场语义地图,再进行基于先验地图的匹配定位,最后通过扩展卡尔曼滤波(Extended Kalman Filter, EKF)将视觉定位结果与里程计融合,输出较为稳定的定位结果.2021年,Zhang等[12]构建停车场高清矢量图进行定位,并提出新的数据关联方法来匹配环视图像和高清矢量图,并通过矢量图中独特道路标志组合进行封闭形式重定位,在多层停车场中实现厘米级定位精度.Xiang等[13]提出了利用鸟瞰图像的混合边缘信息的语义SLAM框架,从鸟瞰图和自由空间轮廓的合成图中提取有用的边缘信息用于SLAM任务.Huang等[14]利用停车场地面语义特征,提出了一种基于两步直接方法的视觉定位算法.

综上所述,视觉SLAM技术已在AVP系统中得到应用,但面向地下停车场环境时,现有方法仍具有以下问题:1)在地下停车场等室内场景下无法接收GPS信号,依赖相机的视觉SLAM定位存在光线较暗、特征少、纹理较弱等问题,导致定位效果差,不稳定.2)面向地下停车场环境的建图定位方案大多通过激光雷达等高精建图定位方式[15-16]离线构建停车场环境先验地图,再基于先验地图匹配语义信息完成车辆在场景中的定位.对于陌生停车场场景,无法预先得到准确的地图信息,而对于不同停车场先验地图构建的准确与否直接影响定位的精确性与鲁棒

性的高低.

针对上述问题,本文提出了一种融合语义信息的视觉惯性定位系统.该系统首先通过视觉里程计和IMU预积分进行视觉惯性信息的融合,并完成视觉惯性初始化.同时,利用4个鱼眼摄像头输入图像构建全景环视图像,并采用语义分割算法提取停车场环境语义信息.然后根据视觉惯性紧耦合位姿完成逆投影变换,获得语义特征投影地图.最后采用回环检测和位姿图优化方式减小累积误差,完成全局位姿图优化,实现更高精度的定位效果.本文的主要工作如下.

1)提出了一种融合语义特征的地下停车场视觉惯性定位算法框架,即融合停车场语义特征与视觉惯性位姿估计构建停车场全局语义地图,再通过地图匹配与回环检测进行位姿图优化实现更高精度的定位.

2)研究了一种基于多鱼眼摄像头的全景环视图构建方法.本文研究了鱼眼摄像头广角相机理论和联合标定过程,完成了鱼眼摄像头畸变矫正、逆投影变换、图像拼接融合等过程,最终构建了以车辆为中心的全景环视图像,为利用停车场全景环视下的语义信息提供支持.

3)基于Gazebo仿真环境和地下停车场实车场景进行算法试验.为了验证本文所研究的面向停车场环境融合语义特征的视觉惯性SLAM定位算法的可行性,分别基于软件仿真环境和自动驾驶实验测试平台车进行实验测试.

1　系统设计

在面向地下停车场环境时,由于环境中光照弱以及纹理特征少,传统的单目视觉SLAM方案无法提供稳定精确的定位信息.本文引入IMU惯性传感单元,通过角速度与加速度等IMU信息的引入更准确快速地恢复单目视觉里程计中模糊的尺度信息,进而减少因光照不足、环境单一无纹理以及运动模糊造成的运动估计误差,提高定位跟踪的性能.此外,在视觉摄像头图像信息与IMU惯导信息的基础上,本文还通过多目鱼眼摄像头融合拼接构建车辆环视图,再通过提取环视语义信息,采用匹配对齐的方式辅助回环检测,同时结合语义信息进行全局定位建图.总体架构如图1所示.

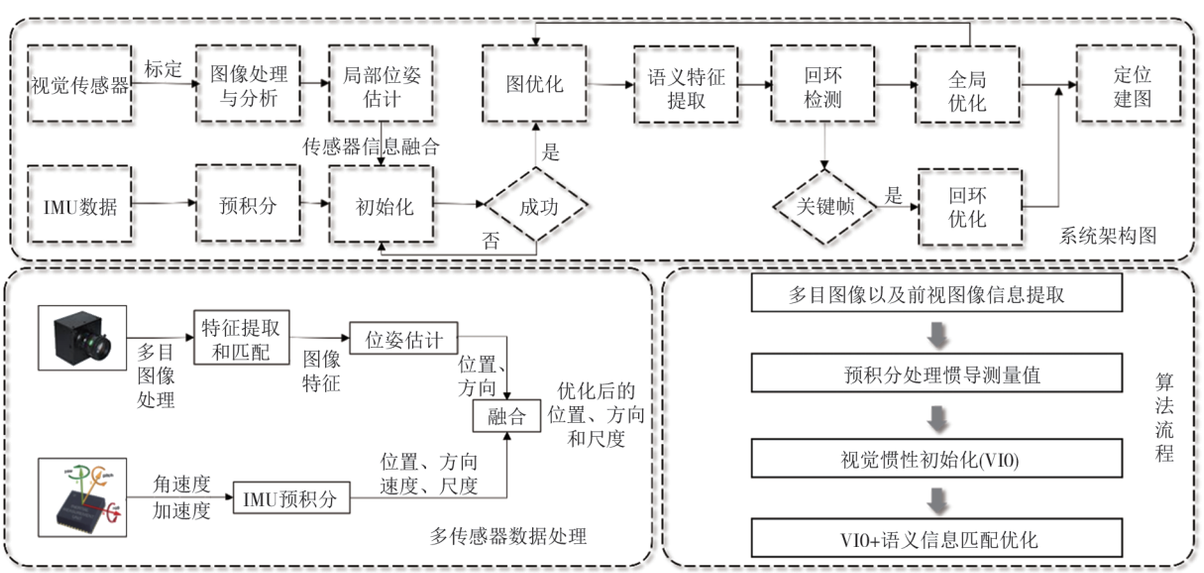


图 1　系统结构图

Fig.1　System structure diagram

2　传感器信息融合

2.1　IMU数据处理

IMU通过加速度计和陀螺仪测量得到在自身刚体坐标系下的加速度和角速度的值,通过积分获得两图像帧之间的旋转和位移的变换.而在SLAM后端非线性优化的时候,需要优化每个图像帧对应的位姿,每次状态变量优化迭代更新都需要对两帧图像之间的IMU数据重新积分,提高了计算复杂度,为了保证实时性与优化求解的简化,可采取预积分策略[17-18].

IMU积分模型转化为预积分模型为:

$$\alpha_{b_{k+1}}^{b_k}=\iint_{t_k}^{t_{k+1}}R_t^{b_k}\left(\tilde{a}_t-b_{at}-\eta_a\right)\mathrm{d}t^2 \tag{1}$$

$$\beta_{b_{k+1}}^{b_k}=\int_{t_k}^{t_{k+1}}R_t^{b_k}\left(\tilde{a}_t-b_{at}-\eta_a\right)\mathrm{d}t \tag{2}$$

$$\gamma_{b_{k+1}}^{b_k}=\int_{t_k}^{t_{k+1}}\frac{1}{2}\Omega\left(\tilde{\omega}_t-b_{\omega t}-\eta_\omega\right)q_t^{b_k}\mathrm{d}t \tag{3}$$

式中:$\alpha_{b_{k+1}}^{b_k}$、$\beta_{b_{k+1}}^{b_k}$、$\gamma_{b_{k+1}}^{b_k}$分别代表第k到第k+1时刻IMU数据积分得到的第k+1时刻的位置、速度和旋转,$R_t^{b_k}$、$q_t^{b_k}$为IMU在t时刻相对k时刻的旋转外参;$\tilde{a}_t$、$\tilde{\omega}_t$分别是两个时刻加速度、角速度的平均值;b_{at}、$b_{\omega t}$分别是加速度计、陀螺仪的零偏(bias);η_a、η_ω分别是加速度计、陀螺仪的噪声.

2.2 视觉惯性紧耦合优化

首先对参考坐标系进行定义,将第一帧图像对应的相机坐标系固定为世界坐标系,后续时刻相机姿态和地图点位置相对世界坐标系的关系均以这一帧坐标系为基准进行估计,如图2所示.

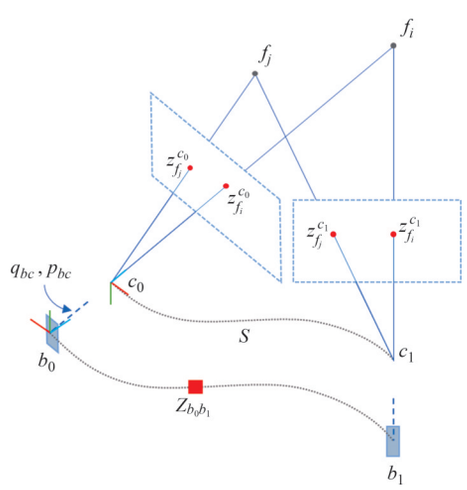


图 2 相机IMU数据关联图

Fig. 2 Camera and IMU datas association diagram

图2中:(q_{bc}, p_{bc})表示相机到IMU坐标系的外参;b_k、c_k分别表示k时刻的IMU坐标系和相机坐标系;b_0、c_0分别表示初始时刻IMU坐标系和相机坐标系;S表示尺度深度;f_i、f_j表示观测到的地图点;$Z_{b_0b_1}$是指b_0到b_1的真实尺度.

根据数据关联对状态变量构建最小二乘问题进行非线性优化求解,主要包括两部分残差约束.由上文视觉特征和IMU数据处理可知,将优化变量分为如下三类:

$$\Lambda=\left[x_0, x_1, \cdots, x_n, x_c^b, \lambda_0, \lambda_1, \cdots, \lambda_m\right];$$

$$X_k=\left[P_{b_k}^{\mathrm{w}}, v_{b_k}^{\mathrm{w}}, q_{b_k}^{\mathrm{w}}, b_a, b_g\right], k\in[0, n];$$

$$x_c^b=\left[P_b^{\mathrm{w}}, q_b^{\mathrm{w}}\right] \tag{4}$$

式中:X_k代表第k帧图像帧对应的IMU状态变量,其中上标w表示世界坐标系下,主要为位置、速度、姿态以及两项偏置;n和m分别表示窗口内待优化的关键帧数和特征点数;λ_m表示第m个特征的第一帧观测逆深度;x_c^b表示相机坐标系与IMU坐标系之间的外参.

对于上述状态变量的优化主要通过构建残差约束的形式进行计算,即基于相机模型的重投影约束与基于IMU测量模型的预积分约束.

2.2.1 视觉重投影约束

在SLAM系统中,对于三维地图点集的某点P,假定其在第i帧首次被相机图像采集,并完成了初始化过程求出当前帧下的逆深度,又在之后的第j帧的图像下产生一个观测的像点.基于相机模型与相机外参的矩阵变换,可将该点P从第i帧投影到第j帧产生一个投影点.若为理想条件,投影点与实际像点应完全重合为一点,但实际像点与投影点之间会有一定距离,即为视觉重投影误差.视觉重投影误差构建过程如下.

将变换矩阵拆分为旋转矩阵和平移向量,则有下式:

$$\begin{aligned}\boldsymbol{P}^{c_j}=&\left(\boldsymbol{R}_c^b\right)^{\mathrm{T}}\left(\boldsymbol{R}_{b_j}^{\mathrm{w}}\right)^{\mathrm{T}}\boldsymbol{R}_{b_i}^{\mathrm{w}}\boldsymbol{R}_c^b\boldsymbol{P}^{c_i}+\\&\left(\boldsymbol{R}_c^b\right)^{\mathrm{T}}\left(\left(\boldsymbol{R}_{b_j}^{\mathrm{w}}\right)^{\mathrm{T}}\left(\left(\boldsymbol{R}_{b_i}^{\mathrm{w}}p_c^b+p_{b_i}^{\mathrm{w}}\right)-p_{b_j}^{\mathrm{w}}\right)-p_c^b\right)\end{aligned} \tag{5}$$

即P^{c_j}为从第i帧相机坐标系投影到第j帧相机坐标系的投影点坐标,将其简化表示为$[x_{c_j}, y_{c_j}, z_{c_j}]^{\mathrm{T}}$,而由图像观测可知该点在第$j$帧的相机归一化坐标表示为$[u_{c_j}, v_{c_j}, 1]^{\mathrm{T}}$,则视觉重投影残差可表示为:

$$\boldsymbol{e}_c=\begin{bmatrix}\dfrac{x_{c_j}}{z_{c_j}}-u_{c_j}\\[2ex]\dfrac{y_{c_j}}{z_{c_j}}-v_{c_j}\end{bmatrix} \tag{6}$$

对式(5)中各变量进行优化求解,迭代直到重投影误差最小时得到各变量的最优解.

2.2.2 IMU预积分约束

根据IMU误差模型以及预积分理论,可知位置P、速度v、旋转q、加速度计偏置b_a和陀螺仪偏置b_g等状态更新公式为:

$$\begin{bmatrix} P^{w}_{b_{k+1}} \\ v^{w}_{b_{k+1}} \\ q^{w}_{b_{k+1}} \\ b^{a}_{k+1} \\ b^{\omega}_{k+1} \end{bmatrix} = \begin{bmatrix} P^{w}_{b_k} + v^{w}_{b_k}\Delta t - \frac{1}{2} g_w \Delta t^2 + q^{w}_{b_k} \alpha^{b_k}_{b_{k+1}} \\ v^{w}_{b_k} - g_w \Delta t + q^{w}_{b_k} \beta^{b_k}_{b_{k+1}} \\ q^{w}_{b_k} \otimes \gamma^{b_k}_{b_{k+1}} \\ b^{a}_{k} \\ b^{\omega}_{k} \end{bmatrix} \tag{7}$$

则预积分测量残差为:

$$\boldsymbol{e}_b = \begin{bmatrix} R^{b_k}_{w}\left(P^{w}_{b_{k+1}} - P^{w}_{b_k} - v^{w}_{b_k}\Delta t + \frac{1}{2} g_w \Delta t^2 \right) - \alpha^{b_k}_{b_{k+1}} \\ R^{b_k}_{w}\left(v^{w}_{b_{k+1}} - v^{w}_{b_k} + g_w \Delta t \right) - \beta^{b_k}_{b_{k+1}} \\ 2\left[\left(q^{w}_{b_k} \right)^{-1} \otimes q^{w}_{b_{k+1}} \otimes \left(\gamma^{b_k}_{b_{k+1}} \right)^{-1} \right]_{xyz} \\ b^{a}_{k+1} - b^{a}_{k} \\ b^{\omega}_{k+1} - b^{\omega}_{k} \end{bmatrix} \tag{8}$$

IMU 的约束残差依赖预积分的模型构建,IMU 预积分约束能为视觉补充较好的观测数据.最后根据视觉重投影误差和预积分测量残差构建紧耦合优化目标函数如下:

$$\min_{\Lambda} \left\{ \sum_{k \in B} \left\| e_b \right\|^2_{P^{b_k}_{b_{k+1}}} + \sum_{k \in C} \left\| e_c \right\|^2_{P^{c_j}_{l}} \right\} \tag{9}$$

综上所述,构建残差约束函数,包括视觉重投影误差函数、IMU 预积分误差函数组合即构成了整体视觉惯性紧耦合的代价函数,通过求解最小二乘问题获得各状态变量最优值,即作为视觉惯性里程计位姿输出,为后续基于语义特征的建图定位提供基础.

3　语义特征提取

3.1　鱼眼相机标定畸变矫正

本文采用 Kannala 等[19]提出的经典鱼眼相机模型,其将鱼眼相机成像规律描述为单位球面投影模型,同时考虑成像过程中的畸变并引入模型参数进行畸变矫正.鱼眼相机成像过程是已知入射角 θ 求出射角 θ_d,而其畸变矫正则是已知矫正后的像点位置 (x', y') 求实际入射角 θ.已知相机参数,可根据 (x', y') 以及相机焦距求得 θ_d 的值,因此,畸变矫正实质是求解关于 θ 的一元高次方程:

$$\theta_d = \theta\left(1 + k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + k_4\theta^8\right) \tag{10}$$

式中:k_1、k_2、k_3、k_4 为畸变系数,可根据相机标定结果得到.

以牛顿迭代法求解该一元高次方程:

$$f(\theta) = \theta\left(1 + k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + k_4\theta^8\right) - \theta_d \tag{11}$$

采用式(11)的投影模型构建重投影误差函数,再根据多个角点之间的对应关系,采用非线性优化的方法,从图像中矫正畸变.

3.2　AVM 环视图像生成

全景环视系统(Around View Monitor, AVM)的构建首先需要对畸变矫正后的鱼眼图像进行逆投影变换,将其转变为俯视图,再进行图像的拼接,生成全景鸟瞰图.本文通过在汽车周围四个角各放置一块棋盘格标定板,再根据逆投影变换(Inverse Perspective Mapping, IPM)原理,计算从相机到地面的转换矩阵.通过转换矩阵即可根据图像中某像素点坐标恢复其在摄像机坐标系投影到 Z=0 平面(即地面)的真实坐标 $(X, Y, 0)$.显然该地平面假设路面平齐,在遇到车辆颠簸以及斜坡不平路面时,假设不成立,逆投影变换便无法使用.但本文面向停车场环境,路面基本平整,可以实现此逆投影变换,具体过程如下.

假设世界坐标系下某点 $P_w(X_w, Y_w, Z_w)$,投影之后的相机坐标系下的坐标为 $P_c(X_c, Y_c, Z_c)$,对应像素坐标系下的点齐次坐标为 $P'(u, v, 1)$,从世界坐标系到像素坐标系之间的变换关系为:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{u} \\ \bar{v} \\ Z_c \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{O}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{12}$$

令 $\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{O}^{\mathrm{T}} & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$,根据俯视鸟瞰图为地面模型坐标系下,即 $Z_w = 0$,则将式(12)转化为:

$$\begin{bmatrix} \bar{u} \\ \bar{v} \\ Z_c \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix} \tag{13}$$

令

$$\boldsymbol{H} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix} \tag{14}$$

取 $\boldsymbol{H}_d = \boldsymbol{H}^{-1}$,则 H 即为从地面世界坐标转化为像

素坐标的投影变换矩阵,也称为单应性矩阵,$\boldsymbol{H}_d$表示逆投影变换矩阵,即若已知$\boldsymbol{H}_d$可以根据像素坐标求解对应地面坐标系下真实坐标,进而转化为俯视鸟瞰图.接下来对这个过程进行推导.记

$$\boldsymbol{H}_d=\begin{bmatrix}H_{d11} & H_{d12} & H_{d13}\\ H_{d21} & H_{d22} & H_{d23}\\ H_{d31} & H_{d32} & 1\end{bmatrix}=\boldsymbol{H}^{-1} \tag{15}$$

由$\bar{u}=uZ_c,\bar{v}=vZ_c$,有:

$$\begin{bmatrix}X_w\\ Y_w\\ 1\end{bmatrix}=\begin{bmatrix}H_{d11} & H_{d12} & H_{d13}\\ H_{d21} & H_{d22} & H_{d23}\\ H_{d31} & H_{d32} & 1\end{bmatrix}\begin{bmatrix}\bar{u}\\ \bar{v}\\ Z_c\end{bmatrix}=\begin{bmatrix}H_{d11} & H_{d12} & H_{d13}\\ H_{d21} & H_{d22} & H_{d23}\\ H_{d31} & H_{d32} & 1\end{bmatrix}\begin{bmatrix}uZ_c\\ vZ_c\\ Z_c\end{bmatrix} \tag{16}$$

对应求解,则有:

$$\begin{cases}X_w=\dfrac{H_{d11}u+H_{d12}v+H_{d13}}{H_{d31}u+H_{d32}v+H_{d33}}\\ Y_w=\dfrac{H_{d21}u+H_{d22}v+H_{d23}}{H_{d31}u+H_{d32}v+H_{d33}}\end{cases} \tag{17}$$

因此,可将相机坐标系下深度系数Z_c消去,得到真实世界坐标系下坐标($Z_c=0$)为$P_w(X_w,Y_w,0)$.

综上所述,通过布置棋盘格标定板,提取对应角点,进而构建重投影误差函数的形式进行投影变换矩阵$\boldsymbol{H}$和逆投影变换矩阵$\boldsymbol{H}_d$的求解.根据上述鱼眼相机标定以及逆投影变换过程,可推导鱼眼相机原始图像某像素坐标u_f与对应地面全景环视图像像素坐标u_G之间变换公式为:

$$u_f=K_fD_f\left(\boldsymbol{H}\boldsymbol{K}_G^{-1}u_G\right) \tag{18}$$

式中:K_f、D_f分别代表鱼眼相机的内参和畸变系数;$\boldsymbol{H}$代表从地面坐标系到鱼眼相机坐标系的投影变换矩阵,由上文知,$\boldsymbol{H}_d=\boldsymbol{H}^{-1}$,$\boldsymbol{H}_d$代表从鱼眼相机坐标系逆投影到地面真实坐标系下的变换矩阵,也即逆投影变换矩阵;$\boldsymbol{K}_G$代表地面真实坐标系投影到全景环视图像坐标系的转换矩阵,也可描述为垂直正对地面且相距地面一定高度的虚拟相机对地面采集图像生成全景环视图.

在完成图像逆投影变换之后,可得到四个方向的鸟瞰俯视图,需要对俯视图像进行拼接融合,最终构建出环视全景图,步骤如下.

1)对4个俯视图重合部分提取ORB(Oriented Fast and Rotated BRIEF)特征[20]进行粗匹配,利用随机抽样一致算法(Random Sample Consensus,RANSAC)剔除误匹配点,并拟合出单应性矩阵初始值.

2)对匹配的对应点重合,保留图像中未重合部分,完成图像拼接,再对单应性矩阵的初值以及重投影误差函数构建最小二乘问题,采用非线性优化方法进行非线性迭代最小求解.

3)对拼接后得到的俯视图,相邻两图拼接处存在拼接缝,会导致图像在经过拼接缝时明显跳跃.采用加权平均融合的方式对重叠区做平滑处理来消除拼接缝,即对重合区域两边界与拼接缝距离按比例设置权值,使其过渡平滑且亮度跳跃较小,达到最佳融合效果.

综上所述,本文最终构建出实际停车场场景全景环视图如图3所示.图像中心为实验测试车辆区域,以车辆图形覆盖表示,图像重叠拼接区域放置的四块矩形标定板清晰可见,结构完整,无明显图像畸变和亮度跳跃.本文构建的全景环视图像能获得车辆周围完整地面图像信息,为后续提取地面语义特征提供了帮助.

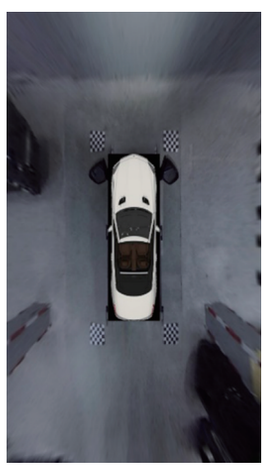


图3 全景环视图像生成

Fig.3 Panoramic surrourd image generation

3.3 语义特征分割

为了满足SLAM实时建图定位对高效率的需求,采用轻量化语义分割模型(如SegNet[21],ERF-Net[22])进行数据集的训练和语义分割的实现.这些模型通过对传统网络模型做出结构性优化和提升,改变模型变量的数值,减少网络的计算量,在一定程度上提高模型算法的推导速率,兼顾了实时性和分割精度的性能需求,达到SLAM系统实时性要求.

在深度学习语义分割中,除了分割模型的选取,不同的数据集同样会对分割效果产生影响,需要使用合适的数据集进行分割网络的模型训练,以获得

较好的分割效果.面向停车场环境的全景环视图像公开语义数据集很少,Wu等[23]借助同济智能电动汽车(TiEV)构建了面向停车场环境的全景环视图像分割数据集(Panoramic Surround View Dataset, PSV dataset).

PSV数据集共包含4 239张停车场场景全景环视图像,其中按6∶1∶3的比例分为训练集、验证集和测试集.每个图像集都包含了一定数量不同类型停车场采集的图像,例如室内、室外停车场,强光或阴影条件下的停车场图像.而真实标签分别对这些图像中的车位和车道线做了标注,对五个类别路标(停车位、白色实线、白色虚线、黄色实线、黄色虚线)用不同颜色的像素标签进行区分,在语义标签中将虚线进行了连接补全以实线表示.

本文采用PSV数据集对SegNet网络进行训练,其中训练集2 543张,测试集1 271张,验证集425张图像,同时在PSV测试集上验证,将原图与分割结果图进行对比观察,测试结果如图4所示.

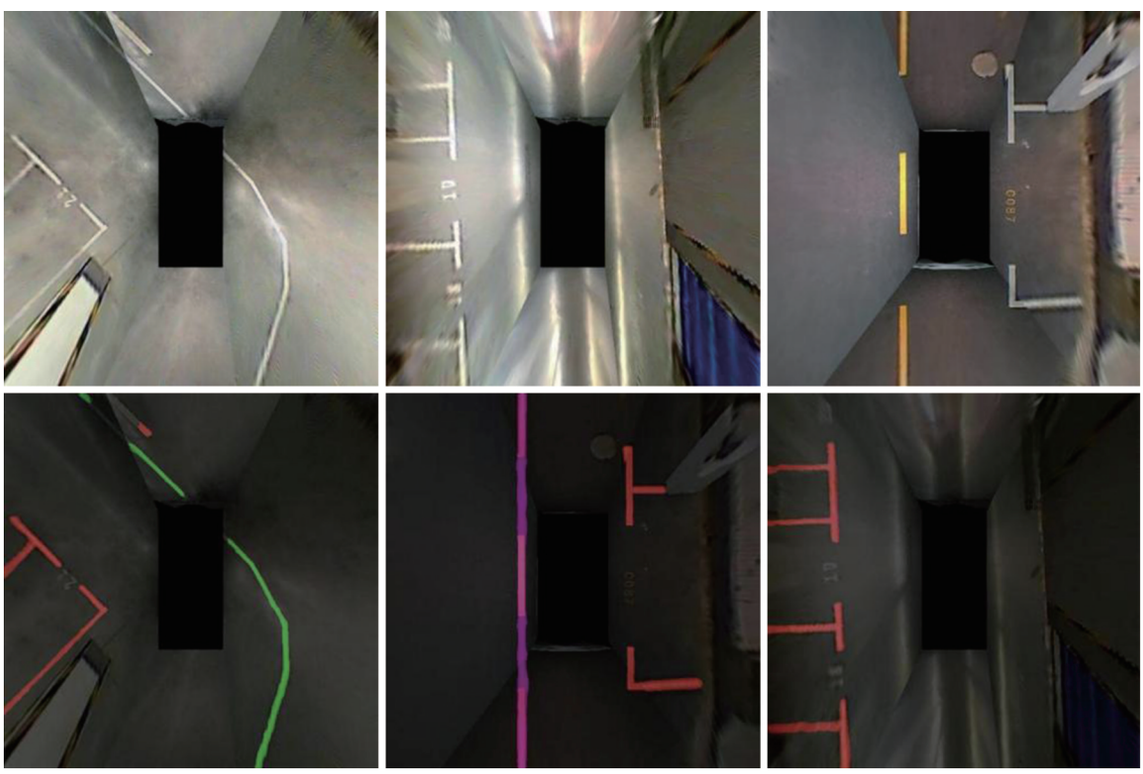


图4　语义分割效果图

Fig.4　Semantic segmentation rendering

从图4可知,SegNet网络对PSV环视图的语义分割在不同的道路环境下均取得了不错的分割效果.因此,可采用语义分割方法提取环视图像语义信息,获得像素级的语义特征.

4　融合语义特征的局部建图定位算法

4.1　语义地图构建

基于停车场环视图像完成语义特征提取,再将语义特征投影到汽车中心坐标系下,根据视觉惯性紧耦合恢复车辆坐标系与世界坐标系之间的位姿信息,将语义特征构建的每一帧局部地图投影并添加到世界坐标系下,从而构建全局地图.

在语义分割提取像素及语义特征坐标后,需要将其从环视图像素坐标系转换到世界坐标系,首先需要将环视图像中像素点的坐标转换到地面坐标系(即Z方向数值为0的车体坐标系)下,由上文对环视图像处理可知,可将环视图像视为垂直正对地面的针孔相机,地面为相机成像平面,环视图像中像素点坐标$\boldsymbol{u}_{\mathrm{G}}$转换为地面坐标系下坐标$\boldsymbol{P}_{\mathrm{G}}$有:

$$\boldsymbol{P}_{\mathrm{G}}=\boldsymbol{K}_{\mathrm{G}}^{-1}\boldsymbol{u}_{\mathrm{G}} \tag{19}$$

式中:$\boldsymbol{K}_{\mathrm{G}}$是从地面坐标系($Z$=0)变换到环视图像坐标系的转换矩阵,即语义特征点在地面坐标系(Z=0时的车辆坐标系)下的坐标为:

$$\boldsymbol{P}_{\mathrm{G}}=\begin{bmatrix}u\\v\\0\end{bmatrix} \tag{20}$$

再由视觉惯性里程计得到位姿$[R_k \quad t_k]$,可将车辆坐标系下的点恢复到世界坐标系下,令

$$\boldsymbol{P}_{\mathrm{w}}=\begin{bmatrix}x_{\mathrm{w}}\\y_{\mathrm{w}}\\z_{\mathrm{w}}\end{bmatrix}=\boldsymbol{R}_k\begin{bmatrix}u\\v\\0\end{bmatrix}+t_k \tag{21}$$

式中:$\boldsymbol{P}_{\mathrm{w}}$代表$\boldsymbol{P}_{\mathrm{G}}$在世界坐标系下的坐标.

根据上述分析,可将分割出的语义像素点投影到三维世界坐标系下,而随着车辆运动,位姿通过视觉惯性里程计更新,环视图中不同像素点的集合在世界坐标系下即构成了局部地图.

4.2 回环检测与位姿图优化

根据上文分析,视觉惯性里程计恢复的位姿随着时间慢慢漂移,则会导致经由车辆位姿转换到世界坐标系下的局部地图同样会产生漂移,导致局部地图不连续且出现重影情况.因此,采用回环检测消除漂移,即在局部建图过程中,每隔一段时间就将当前局部地图与之前局部地图进行比较,采用迭代最近点(Iterative Closest Point, ICP)对两项局部地图点云进行匹配.ICP通过迭代优化,不断寻找对应匹配点,再最小化匹配点之间的欧式距离误差.若能匹配成功,则两者之间由于位姿漂移会有一个相对位姿,这个相对位姿用于全局位姿优化来矫正累计漂移.

通过ICP求解推导最新局部地图与周围局部地图之间的相对变换,在获得相对位姿之后采用位姿图优化对地图进行全局优化,将局部地图位姿作为优化变量即顶点,构建两种形式的边,包括二元边和一元边.二元边由两个局部地图相对位姿与回环检测求出的相对位姿作差的残差构建,一元边则是局部地图位姿与车辆视觉惯性里程计位姿差,防止优化后的位姿与里程计相差太大.

假设在该位姿图中,局部地图的位姿为顶点,采用三轴旋转:$\boldsymbol{r}=[r_x \quad r_y \quad r_z]^{\mathrm{T}}$以及三轴平移:$\boldsymbol{t}=[t_x \quad t_y \quad t_z]^{\mathrm{T}}$来表示变量,则该位姿图优化的方程为

$$X=\arg\min_X\sum_k\left\|f\left(r_{k+1},t_{k+1},r_k,t_k\right)-z^c_{k,k+1}\right\|^2+\sum_{i,j\in\varphi}\left\|f\left(r_i,t_i,r_j,t_j\right)-z^l_{i,j}\right\|^2 \tag{22}$$

式中:$X=\left[r_0,t_0,\cdots,r_k,t_k\right]^{\mathrm{T}}$表示所有局部地图的位姿;$z^c_{k,k+1}$表示里程计相邻帧第$k$帧与第$k$+1帧之间的相对位姿;$\varphi$表示检测的回环匹配对;$z^l_{i,j}$代表ICP回环求解的回环帧第$i$帧和第$j$帧之间的相对位姿差;函数$f(\cdot)$表示局部地图相对位姿差.

4.3 基于先验地图的定位

结合语义地图构建的结果,保存全局一致的地图,基于停车场环境下的先验地图,车辆行驶时可以进行定位.定位环节首先通过视觉惯性初始化以里程计输出作为定位的初始值,对环视图像下检测语义特征,同样投影到车辆坐标系下,而此时可通过语义特征点与先验地图特征进行ICP匹配估计车辆当前姿态,推导过程如下.

假设定位过程时,鱼眼摄像头采集获得的全景环视图中分割出的车位线某特征点坐标为$u_{\mathrm{G}}(u_k,v_k)$,转换为齐次坐标形式为$[u_k \quad v_k \quad 1]^{\mathrm{T}}$,则将其投影到车辆坐标系下的齐次坐标形式$\boldsymbol{P}_{\mathrm{G}}$表示为:

$$\boldsymbol{P}_{\mathrm{G}}=\begin{bmatrix}u\\v\\1\end{bmatrix}=\boldsymbol{K}_{\mathrm{G}}\begin{bmatrix}u_k\\v_k\\1\end{bmatrix} \tag{23}$$

显然该特征点在车辆坐标系下Z轴坐标应为0,则该特征点在车辆坐标系下坐标为$\boldsymbol{P}_{\mathrm{G}}=[u \quad v \quad 0]^{\mathrm{T}}$,通过ICP匹配的方式将其与全局地图下的最近点进行匹配计算,估计当前车辆位姿,公式表示为:

$$\boldsymbol{r}^*,\boldsymbol{t}^*=\arg\min_{r,t}\sum\nolimits_{k\in S}\left\|R(r)\begin{bmatrix}u\\v\\0\end{bmatrix}+t-\begin{bmatrix}x_k\\y_k\\z_k\end{bmatrix}\right\|^2 \tag{24}$$

式中:$\boldsymbol{r},\boldsymbol{t}$分别表示当前帧的三维旋转和平移向量;$\boldsymbol{r}^*,\boldsymbol{t}^*$为达到最优匹配时的三维旋转和平移向量;$S$代表所有特征点的集合;$\begin{bmatrix}u\\v\\0\end{bmatrix}$表示车辆坐标系下的当前语义特征点坐标;$\begin{bmatrix}x_k\\y_k\\z_k\end{bmatrix}$表示该语义特征在全局先验地图中的最近迭代点.

综上所述,本文通过视觉惯性融合语义特征的建图定位,增加视觉惯性SLAM定位系统的鲁棒性,使其在地下停车场环境下也能稳定可靠地运行,获得良好的定位效果.

5　试验与分析

为了验证本文所研究的地下停车场环境融合语义特征的视觉惯性SLAM定位算法的可行性,分别基于仿真环境和自动驾驶实验测试平台车以及基于地下停车场的数据进行测试验证.首先通过仿真软件构建地下停车场地面模型,在仿真条件下测试算法的可行性,再通过地下停车场真实环境数据,验证本文定位系统在地下停车场环境下的定位准确性和鲁棒性.

5.1　仿真实验

5.1.1　仿真环境搭建

本文使用开源机器人仿真器Gazebo建立了地下停车场仿真模拟环境,建立了搭载多个传感器的轮式机器人模型和地下停车场地面环境模型,以验证算法的有效性.在模型构建过程中,利用轮式移动机器人配置多个相机,采用IPM逆投影变换等技术,建立地面图像与机器人载体坐标之间的变换关系,获得地面特征点在相机坐标系下的投影,进而构建出全景环视投影地图.除此之外,本文还建立地下停车场基本语义特征模型,包括停车位、车道线等地面模型,模拟真实场景地面环境.

仿真实验应用RVIZ平台展现ROS主题消息以及Gazebo仿真软件算法实现可视化效果.使用RVIZ工具,用户不仅可以发布特定的控制命令,还可以通过可视化反馈对目标进行操控.构建停车场仿真环境如图5所示.

5.1.2　基于先验地图的定位试验及结果分析

本文通过仿真实验构建了全局一致的停车场语义地图,在仿真环境下,仿真机器人基于先验地图实现自身的定位,预先构建的语义地图加载如图6所示.通过先验地图,机器人能迅速恢复自身当前位姿,实现快速重定位,同时在运动过程中,基于摄像头语义点云特征感知提取,通过ICP匹配提取特征与先验地图信息,求解每一帧下的位姿,实现定位过程,同时在经过曾经到过的地方时可以进行回环检测,实现全局位姿图优化,减小累计误差.

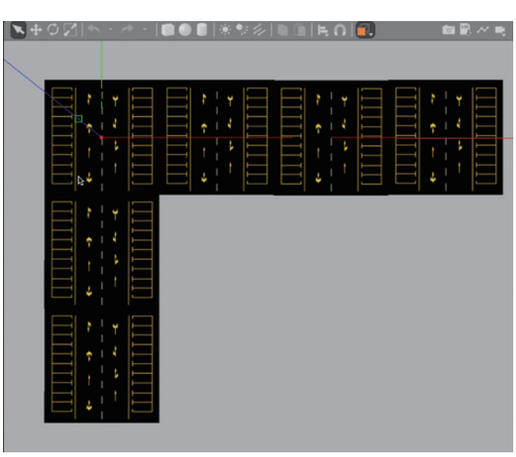


图5　停车场地面仿真环境

Fig.5　Parking lot ground simulation environment

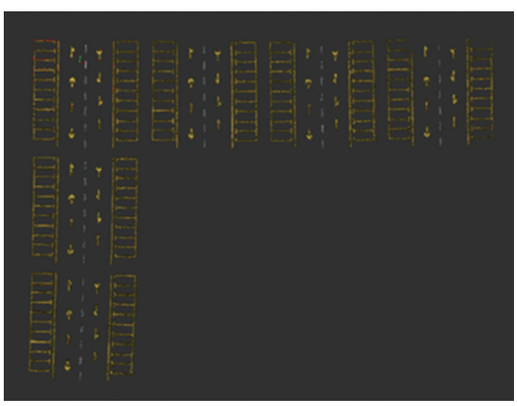


图6　先验地图构建结果

Fig.6　Prior map construction results

在仿真环境下,Gazebo仿真软件可实时获得机器人真实位姿作为真值,与本文采用的定位算法求解的机器人位姿对比,评估算法的误差及可靠性,定位轨迹对比图如图7所示.蓝色曲线代表移动机器人在仿真环境中的真实轨迹,黑色曲线代表视觉里程计Odometry运动推算输出的轨迹结果,红色曲线与绿色曲线分别代表仅采用ICP匹配的纯SLAM定位轨迹和基于先验地图的位姿图优化后的定位轨迹,显然绿色曲线要更为接近真实轨迹,定位效果最好.

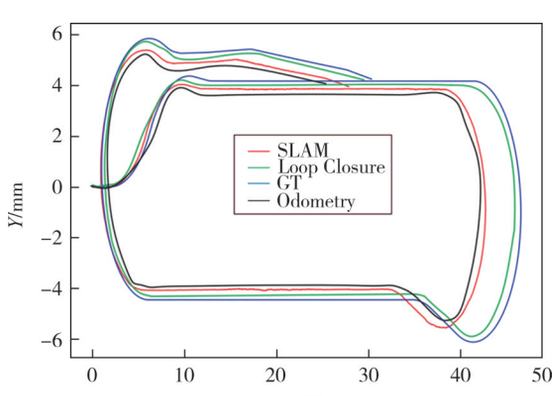


图7　不同条件定位轨迹图

Fig.7　Localization trajectories under different conditions

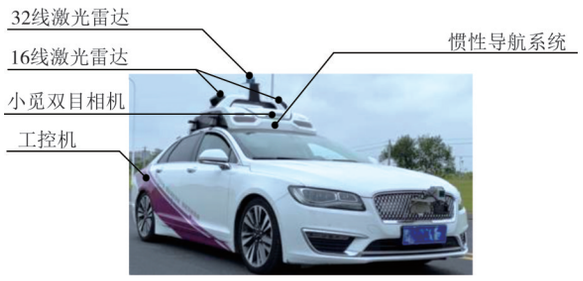
由图7可知,经过回环检测位姿图优化的SLAM算法可以有效地提高定位精度.通过对比定位轨迹和真实轨迹,可以发现,在运动初期,三者定位轨迹与真实轨迹贴合度较高.然而,随着里程数的增加,尺度不稳定问题逐渐显现,导致SLAM算法的定位误差逐渐增大,特别是经过旋转后,在图7中表现为Odometry定位轨迹以及仅基于ICP匹配的SLAM定位轨迹在x方向上和真实轨迹存在较大偏差,定位误差较大.而基于先验地图通过回环检测和全局位姿图优化,对错误尺度下的定位轨迹进行修正,可以恢复地图尺度并获得更高精度的Loop Closure轨迹,误差减小.因此,基于先验地图融合语义特征回环检测的SLAM算法具有更高的定位精度,表明可以有效利用停车场环境下常见的语义标识符等特征实现基于语义特征的建图定位.

5.2 实车场景实验

本文通过实车实验研究面向地下停车场环境融合语义特征的视觉惯性SLAM定位算法,使用智能网联车辆平台验证算法的可行性和定位效果.该平台通过林肯MKZ汽车改装而成,配备了ARK-3500嵌入式工控机作为自动驾驶算法的处理器与计算单元,如图8(a)所示.实验环境选择某规则化地下停车场,如图8(b)所示.该停车场规模较大,结构标准,停车位布局规整,因此选择在该停车场环境下测试系统轨迹定位精度,验证本文所研究的视觉惯性融合定位算法在真实场景下的定位效果.

考虑地下停车场场景无法获得车辆真实位姿的问题,无法将本文算法定位性能与真值进行任意时刻的绝对误差对比,需要选择合适的方式对算法性能进行侧面评估.本文通过测量车辆两次经过同一位置时的重复定位误差来分析定位效果,例如选取车辆出发点附近作为标记点,从起点出发,在实验场景中行驶一圈后停在起点位置,此位置也设为终点,分析算法输出的定位轨迹是否能在起点和终点闭合以及起点和终点定位坐标的偏差大小来评估算法输出轨迹的误差情况.此外,ORB-SLAM3[4]是目前性能表现相对最好的视觉SLAM算法,本文选择将ORB-SLAM3作为对比算法进行比较分析.如图9所示,将车辆行驶路径上不同位置的三个点作为重复定位误差的测试点.

驾驶车辆沿该路径行驶两周,即车辆分别通过各标记点两次,在算法输出的定位结果中找到并记录每次经过标记点的坐标,得到三组前后经过坐标点的坐标值,然后通过计算相同标记点前后坐标差值,即可得到三轴重复定位误差及绝对重复误差距离.



(a)智能网联平台车



(b)地下停车场实验环境

图8 实验条件配置

Fig.8 Experimental conditions and configurations

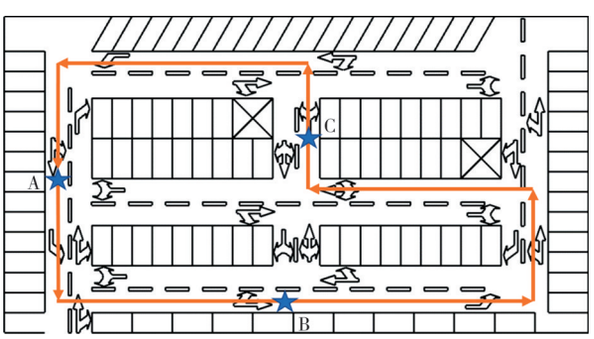
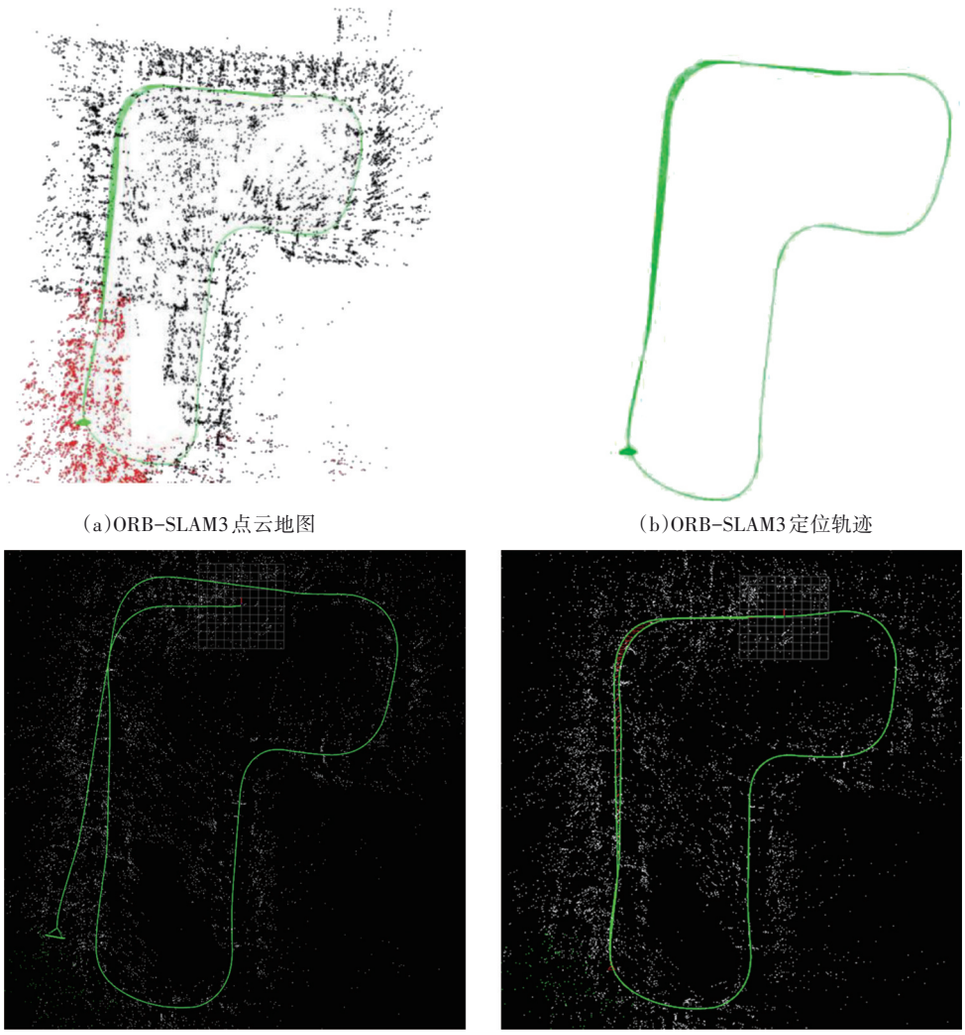


图9 行驶路线及标记点位置

Fig.9 Driving route and location of marking points

图10展示了本文算法和ORB-SLAM3在该地下停车场行驶轨迹所输出的定位轨迹图.图10(a)与图10(b)分别表示ORB-SLAM3算法所构建的点云地图和定位轨迹输出结果,从图中可看出,ORB-SLAM3算法在经过回环检测后能较好地回到起点位置形成闭环,但后续轨迹仍存在一定误差.而图10(c)与图10(d)分别表示未进行回环检测和检测到回环并进行全局位姿图优化后的结果.显然,本文算法在行驶最后阶段不进行回环检测优化时的误差较大,起点和终点位置无法闭合,且有较大漂移.在添加回环检测线程后,在运行轨迹的末段形成闭合回路,通过回环检测与全局位姿图优化减小了位姿累积漂移.



(a)ORB-SLAM3点云地图　　(b)ORB-SLAM3定位轨迹

(c)本文算法未检测回环时定位轨迹　　(d)全局位姿图优化后定位轨迹

图 10　本文算法和 ORB-SLAM3 在该地下停车场行驶轨迹所输出的定位轨迹图

Fig.10　The positioning trajectory map output by the algorithm and ORB-SLAM3 in the underground parking lot

此外,在相同环境工况条件下分别运行ORB-SLAM3算法和本文算法驾驶车辆经过三个标记点,测量计算各个点的重复定位误差,得到对应重复定位误差,结果如表1所示.

表 1　重复定位误差结果

Tab.1　Repeated positioning error results

算法	位置	坐标/m	各轴误差(Δx, Δy, Δz)/m	绝对误差/m
ORB-SLAM3	A	(0.036 1,0.046 9,0.024 3) (-0.051 5,0.162 3,0.126 3)	(0.087 6,-0.115 4,-0.102 0)	0.177 1
	B	(-24.197 2,18.411 9,0.101 2) (-24.367 5,18.246 0,0.212 6)	(0.170 3,0.165 9,-0.111 4)	0.262 6
	C	(-16.732 4,-4.177 8,0.109 8) (-16.861 3,-4.365 6,0.203 3)	(0.128 9,0.187 8,-0.093 5)	0.246 2
本文算法	A	(0.027 1,0.078 9,0.035 6) (-0.065 3,0.136 8,0.126 3)	(0.092 4,-0.057 9.-0.090 7)	0.141 8
	B	(-24.329 2,18.307 7,0.087 5) (-24.187 5,18.426 0,0.137 3)	(-0.141 7,-0.118 3,-0.049 8)	0.191 2
	C	(-16.820 7,-4.220 3,0.090 2) (-16.693 1,-4.066 3,0.153 6)	(-0.1276,-0.1540,-0.0634)	0.209 8

由表1可知,车辆行驶经过标记点A时在ORB-SLAM3算法下两次输出的坐标分别为(0.036 1,0.046 9,0.024 3)、(-0.051 5,0.162 3,0.126 3),其对应的重复定位坐标误差为(0.087 6,-0.115 4,-0.102 0),绝对距离相差0.177 1 m.而本文研究的视觉惯性算法获得的A点的前后坐标分别为(0.027 1,0.078 9,0.035 6)、(-0.065 3,0.136 8,0.126 3),重复定位三轴误差为(0.092 4,-0.057 9.-0.090 7),相差0.141 8 m.同理,对B、C两个标记点进行相同分析过程,可知ORB-SLAM3在经过B、C两点时的距离误差分别为0.262 6 m、0.246 2 m,而通过本文算法求得的B、C两点时的距离误差分别为0.191 2 m、0.209 8 m.

综上所述,综合三个标记点前后坐标信息,可以计算得到ORB-SLAM3的重复定位误差平均为22.9 cm,而本文算法计算得到的重复定位误差平均为18.1 cm,在一定程度上优于前者的定位精度,体现了本文算法可以有效提升停车场环境下自动驾驶车辆自动泊车应用场景的定位精度和可靠性.

6 结论

本文针对地下停车场环境的定位问题构建了融合语义信息的视觉惯性定位系统.该系统使用视觉里程计和IMU预积分进行基于视觉惯性信息融合的位姿估计.同时,采用4个鱼眼摄像头构建全景环视图像,并基于轻量化语义分割模型SegNet分割提取停车场环境中的语义信息.结合视觉惯性位姿估计将语义分割特征点投影在世界坐标系下从而构建局部语义地图,再采用ICP算法完成回环检测与全局位姿图优化获得全局一致地图.最后基于先验地图实现更高精度的定位.未来工作会针对实车构建语义地图的实时性和非标准停车场环境两方面进行进一步研究,提高定位系统的实时性和鲁棒性.

参考文献

[1] 孙永全,田红丽. 视觉惯性SLAM综述[J]. 计算机应用研究,2019,36(12):3530-3533, 3552.
SUN Y Q, TIAN H L. Overview of visual inertial SLAM [J]. Application Research of Computers, 2019, 36(12): 3530-3533, 3552.(in Chinese)

[2] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.

[3] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras [J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.

[4] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: an accurate open-source library for visual, visual-inertial, and multimap SLAM [J]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.

[5] 韩彦峰,唐超超,肖科. 基于改进ORB-SLAM2算法的RGB-D稠密地图构建[J]. 湖南大学学报(自然科学版),2023,50(2):52-62.
HAN Y F, TANG C C, XIAO K. RGB-D dense map construction based on improved ORB-SLAM2 algorithm[J]. Journal of Hunan University (Natural Sciences), 2023, 50(2): 52-62.(in Chinese)

[6] QIN T, LI P L, SHEN S J. VINS-mono: a robust and versatile monocular visual-inertial state estimator[J]. IEEE Transactions on Robotics, 2018, 34(4): 1004-1020.

[7] 王军,夏利民. 基于深度学习特征的异常行为检测[J]. 湖南大学学报(自然科学版),2017,44(10):130-138.
WANG J, XIA L M. Abnormal behavior detection based on deep-learned features [J]. Journal of Hunan University (Natural Sciences), 2017, 44(10): 130-138.(in Chinese)

[8] 秦晓辉,黄启东,常灯祥,等. 基于改进YOLOv5的露天矿山目标检测方法[J]. 湖南大学学报(自然科学版),2023,50(2):23-30.
QIN X H, HUANG Q D, CHANG D X, et al. Object detection method in open-pit mine based on improved YOLOv5 [J]. Journal of Hunan University (Natural Sciences), 2023, 50(2): 23-30.(in Chinese)

[9] 李小倩, 何伟, 朱世强, 等. 基于环境语义信息的同步定位与地图构建方法综述[J]. 工程科学学报,2021,43(6):754-767.
LI X Q, HE W, ZHU S Q, et al. Survey of simultaneous localization and mapping based on environmental semantic information[J]. Chinese Journal of Engineering, 2021, 43(6): 754-767.(in Chinese)

[10] HU J X, YANG M, XU H Q, et al. Mapping and localization using semantic road marking with Centimeter-level accuracy in indoor parking lots [C]//2019 IEEE Intelligent Transportation Systems Conference (ITSC). Auckland, New Zealand: IEEE, 2019: 4068-4073.

[11] QIN T, CHEN T Q, CHEN Y L, et al. AVP-SLAM: semantic visual mapping and localization for autonomous vehicles in the parking lot [C]//2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Las Vegas, NV, USA: IEEE, 2020: 5939-5945.

[12] ZHANG C, LIU H, XIE Z J, et al. AVP-loc: surround view localization and relocalization based on HD vector map for automated valet parking [C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Prague, Czech Republic: IEEE, 2021:5552-5559.

[13] XIANG Z Z, BAO A B, SU J B. Hybrid bird's-eye edge based semantic visual SLAM for automated valet parking [C]//2021 IEEE International Conference on Robotics and Automation (ICRA). ACM, 2021:11546-11552.

[14] HUANG T F, CHEN Z H, ZHAO J Q, et al. Semantic visual localization and mapping in the parking lot using direct method [C]//2021 5th CAA International Conference on Vehicular Control and Intelligence (CVCI). Tianjin, China: IEEE, 2021: 1-6.

[15] CHANG D X, ZHOU Y S, HU M J, et al. Robust accurate LiDAR-GNSS/IMU self-calibration based on iterative refinement [J]. IEEE Sensors Journal, 2023, 23(5):5188-5199.

[16] CHANG D X, ZHANG R B, HUANG S J, et al. WiCRF: weighted bimodal constrained LiDAR odometry and mapping with robust features[J]. IEEE Robotics and Automation Letters, 2023, 8(3):1423-1430.

[17] FORSTER C, CARLONE L, DELLAERT F, et al. IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation[C]//Robotics: Science and Systems XI. 2015.

[18] FORSTER C, CARLONE L, DELLAERT F, et al. On-manifold preintegration for real-time visual: inertial odometry[J]. IEEE Transactions on Robotics, 2017, 33(1):1-21.

[19] KANNALA J, BRANDT S S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(8):1335-1340.

[20] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF [C]//2011 International Conference on Computer Vision. Barcelona, Spain: IEEE, 2011: 2564-2571.

[21] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12):2481-2495.

[22] ROMERA E, ÁLVAREZ J M, BERGASA L M, et al. ERFNet: efficient residual factorized ConvNet for real-time semantic segmentation[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(1): 263-272.

[23] WU Y, YANG T, ZHAO J Q, et al. VH-HFCN based parking slot and lane markings segmentation on panoramic surround view[C]// 2018 IEEE Intelligent Vehicles Symposium (Ⅵ). Changshu, China: IEEE, 2018: 1767-1772.