

文章编号: 2095-2163(2024)03-0123-05

中图分类号: TP391.41

文献标志码: A

停车场环境下的视觉惯性里程计

闫旭东, 钱 莉

(上海工程技术大学 机械与汽车工程学院, 上海 201620)

摘要: 为解决自动驾驶汽车在停车场环境下停车过程中的定位问题, 本文针对停车场环境设计了基于惯性测量单元和环视图像的视觉惯性里程计。利用逆透视变换拼接多个相机传回的图像组成环视图像, 并提取所需的视觉特征, 之后配准估计车辆姿态; 将视觉特征拼接成局部地图后, 再将局部地图与全局地图拼接, 构建停车场的地面点云地图; 融合惯性测量单元测量数据和姿态构建了基于误差状态卡尔曼滤波器的松耦合视觉惯性里程计。在机器人操作系统下实现算法并设计仿真环境进行了在线仿真, 验证了算法可以为自主泊车这种行驶过程中转弯多、行驶距离不大的任务提供较为精确的位置信息。

关键词: 点云地图; 误差状态卡尔曼滤波; 环视图像; 视觉里程计

Visual inertial odometry under parking lot environment

YAN Xudong, QIAN Li

(School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

Abstract: To solve the problem regarding localization of autonomous vehicle under parking lot environment, this work develops an odometry based on Inertial Measurement Unit and surround-view vision. A surround-view vision is achieved using Inverse Perspective Mapping method, at the same time needed visual features are extracted, vehicle pose is then estimated using registration. After the local map is constructed with visual features, it is merged into global map to form a point cloud map for the ground of the parking lot. After that, a loosely coupled visual inertial odometry is developed based on Error State Kalman Filter to fuse measurement of Inertial Measurement Unit and the pose. Finally, the proposed algorithm is implemented and an online simulation is performed under Robotic Operating System environment, it is confirmed that the proposed algorithm can provide the accuracy requirements of automatic parking tasks which has many turns and short driving distance.

Key words: point cloud map; error state Kalman filter; surround-view vision; visual odometry

0 引言

近年来,对智能驾驶的需求持续增加,其中的一个热门方向就是自主泊车。自主泊车指在停车过程中不需要人为干预,汽车自动停入停车位中。自主泊车任务的行驶距离较短,但对精度要求高。为了实现停车过程中精确的车辆定位,智能驾驶汽车都配备了传感器和系列热门技术,如摄像头、全球定位系统(GPS)和惯性测量单元(IMU)等。但在某些环境,如地下停车场,GPS信号不稳定或有较大误差,因此选择不受干扰的IMU和摄像头更为常见,算法的视觉特征使用停车场内部灰度值较高、容易辨别的停车位地面标记和道路交通标志线作为算法的视觉特征。

Im等学者^[1]提出将基于全景影像系统和激光雷达融合的停车场同时定位与建图(SLAM)算法,

将地面标记作为特征,完成了一套可以进行回环检测和停车位识别的系统,融合AVM(Around View Monitor)和激光雷达,利用点云分割算法和区域生长算法等方法提取地面标记点云,接着使用主成分分析算法得出每个点云簇的主方向,并利用点云簇作为特征实现了SLAM算法。Qin等学者^[2]改动并重新训练了一个基于U-NET卷积神经网络,用于对停车场地面图像进行语义分割以完成地面标记的提取,并实现了一个用于代客泊车的SLAM算法。利用停车场中的地面标记作为视觉特征能够减小视觉特征的数量,从而保证了视觉匹配算法的运行速度,而地面标记易识别且不能移动,属于静态特征,特征的位置信息更可靠。Cui等学者^[3]针对室内或地下停车场环境提出了一种使用停车位编号的定位算法OCR-PF,算法由2部分组成,分别是负责提供

作者简介: 闫旭东(1998-),男,硕士研究生,主要研究方向:自动驾驶技术。

通讯作者: 钱 莉(1980-),女,博士,讲师,主要研究方向:导航与定位技术。Email:860212978@qq.com

收稿日期: 2023-03-08

测量的光学字符识别(OCR)和渐进全局定位的粒子滤波,OCR部分使用图像增强和机器学习方法对摄像头所采集的图像进行处理,提取出停车位编号在图像中的位置,此后计算停车位编号中心所处的像素坐标,利用投影法根据摄像头的内参和外参计算摄像头在空间中相对于停车位编号的位置;粒子滤波器部分将OCR获得的位置信息作为测量值输入粒子滤波器中进行更新,实现定位。

1 视觉惯性里程算法

1.1 算法架构

为了降低误差同时保证算法的实时性,本文采用局部地图与当前帧配准的方法,以获得帧间的旋转和平移关系,在设计前端里程计时选择了NDT(Normal Distribution Transform)点云配准算法。NDT相对于ICP(Iterative Closest Point)不容易陷入局部最小值,故对初值的要求更低、且参考帧点云较稀疏时对旋转稳定性更高,因此点云地图可以进行降采样,使得算法运行速度快、且时间稳定^[4-5]。

配准获得当前帧相对于局部地图的旋转和平移信息后,将当前帧融入局部地图并进行体素降采样。当局部地图中的帧数大于10帧时,将局部地图与全局地图配准后拼接,并对全局地图进行更新,而后进行体素降采样,由此获得停车场的点云地图。

IMU的测量值是物体 x 、 y 和 z 轴方向的加速度和角速度。由于惯性传感器相对于相机有采样频率更高且对于高频信号敏感的特性,在松耦合里程计中通常将其测量值作为控制量输入^[6]。将IMU融入视觉里程计的主要作用有2个方面。其一,是为视觉匹配算法提供更好的初值,减少其优化过程中的迭代次数并提升匹配精度^[7];其二,在视觉匹配算法受剧烈运动影响失配或误匹配时,也能维持较准确的位置信息^[8-9]。

本文采用了松耦合的方式融合IMU和视觉里程计数据,以高频率输入的IMU传感器作为控制量,视觉里程计数据作为观测量修正预测值。为了保证系统的稳定性和精度,使用误差状态卡尔曼滤波器(ESKF)方式进行融合。与扩展卡尔曼滤波(EKF)、无迹卡尔曼滤波(UKF)这些直接滤波方法对比,ESKF是一种对系统的误差状态进行维护的间接滤波方法^[10]。相对传统的卡尔曼滤波,由于ESKF所维护误差量通常是小量,从而保证了线性化的合理性和有效性,并且远离奇异值、万向节锁,稳定性高,算法架构如图1所示。

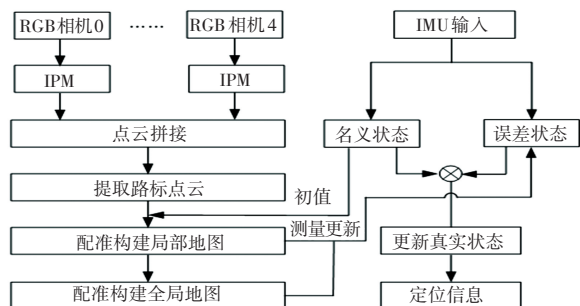


图1 视觉惯性里程算法架构图

Fig. 1 Architecture of the visual inertial odometry algorithm

1.2 多张图像拼接成的环视图

全景影像系统可以最大限度地获取大范围的信息,覆盖车辆周围 360° 的场景^[11]。本文采用5个RGB摄像头采集车辆周边的图像信息,并拼接成环视图,相机帧数为10FPS,同时使用了一个IMU传感器来提供车辆的角速度和加速度信息,相机和IMU的外参和内参以及两者之间坐标转换关系均已提前获得。采用逆透视变换(IPM)方法分别将各摄像头采集到图像的像素投影到车辆坐标系下的地面上,形成多个点云,再将所有相机图片对应的点云拼接组成一个环绕汽车的完整点云,重新将点云投影回二维图像后,提取出属于地面标记的部分。

采用逆投影变换将每个摄像头所采集图像的像素分别投影到车辆坐标系的地平面上(即平面 $z=0$),变换如式(1)所示:

$$\frac{1}{\lambda} \begin{bmatrix} x^v \\ y^v \\ 1 \end{bmatrix} = [R_c \quad t_c]^{-1}_{col:1,2,4} \pi_c^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (1)$$

其中, π_c^{-1} 为摄像头的逆投影模型,作用是将像素坐标投影到摄像头坐标系下的归一化坐标中; $[R_c \quad t_c]$ 为各摄像头相对于汽车中心的外参; $[u \quad v]$ 为像素坐标; $[x^v \quad y^v]$ 为点云在车辆坐标系中的坐标。

得到各个摄像头采集的图像所对应的点云后,再将各摄像头的点云拼接。6个摄像头朝向不同的方向,拼接后获得的点云是环绕汽车的点云。

由于摄像头之间视场不可避免地存在重叠,且图像经过逆投影会产生数量庞大的点,从而使得后续处理复杂化。因此除了将深度太大和太小的点剔除以外,还要使用体素降采样法对点云进行降采样处理。

1.3 停车线视觉特征提取

摄像头收集到的停车场地面图像信息被转换为停车场的地面点云。为了提取地面点云中属于停车位地面标记和道路交通标志线的部分,先丢弃距离较远的点,将RGB点云转换为灰度点云后,将点云

按灰度阈值进行分割^[12]。点云中灰度高的部分保留,丢弃低灰度点,提取出的地面点云如图2所示。

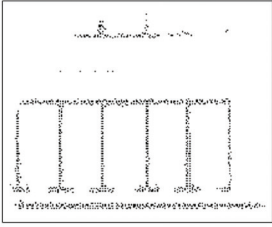


图2 提取出的地面点云

Fig. 2 Extracted point cloud of the ground

1.4 点云配准

将摄像头采集到的图像转化为环绕汽车的当前点云帧。如果是系统启动后采集到的第一帧图像,则将此帧作为局部地图,并将起始位置作为里程计原点。由于直接将当前帧与全局地图做配准会消耗大量计算资源,因此采用关键帧的方法减少与全局地图配准的频率^[13]。利用滤波器的状态值作为点云配准算法的初始值,将其与当前帧配准后拼接。当关键帧大小超过阈值时,将关键帧与全局地图配准并拼接入全局地图,再进行体素下采样。

对经过体素下采样的稀疏点云进行配准时,采用NDT算法作为本文使用的点云配准算法。首先,将参考帧点云分割成多个边长为 L 的栅格,并假设点云分布符合正态分布;计算点的均值 μ_y 和协方差 Σ_y ,并构建点云的概率密度分布函数(PDF);最后通过牛顿法求出置信度 $s(R)$ 极大值和最优变换参数 R ,置信度 $s(R)$ 也称作配准得分(score),其过程可写为式(2)^[14]:

$$s(R) = - \sum_{i=1}^n \left(\frac{1}{(2\pi)^{3/2} \sqrt{|\Sigma_y|}} \right) \exp \left(- \frac{(p_i^{\text{new}} - \mu_y)^T \Sigma_y^{-1} (p_i^{\text{new}} - \mu_y)}{2} \right) \quad (2)$$

其中, p_i^{new} 为当前帧点云中的点。

1.5 数据融合

本文采用松耦合方法融合IMU数据和配准数据,IMU对车辆物体 x 、 y 和 z 轴的加速度 a_x 、 a_y 、 a_z 和角速度 ω_x 、 ω_y 、 ω_z 进行采集。算法在初始化阶段采用静止初始化方法,初始化过程用于求解IMU和全局坐标系的旋转矩阵 R_i^w 和角速度零偏 ω_b ,静止指物体在启动阶段不发生运动^[15]。初始化阶段收集多个连续的IMU测量值,并对加速度和角速度分别求均值,均值得到 a_g 和 ω_b ,并用 a_g 求出重力加速度,进而得出 R_i^w 。

在利用IMU输入作为控制量的算法中,由于零偏和高斯噪声的影响,误差会不断累积,影响系统的

精度^[16]。为了降低干扰,使用误差卡尔曼滤波器(ESKF)。ESKF有3种状态:真实状态(true state) x_t 、名义状态(nominal state) x 和误差状态(error state) δx ,这3种状态的关系,可由式(3)表示为^[17]:

$$x_t = x \oplus \delta x \quad (3)$$

与直接法滤波不同,ESKF中的大信号是由名义状态保存的,而误差状态的值通常较小,工作在原点附近,线性化精度更高^[18-19]。

首先,滤波器用IMU测量值对名义状态进行积分,积分结果为点云地图配准提供了初始值。同时,误差状态也按照滤波器的误差传递方程传播;在随后更新阶段中,ESKF将点云配准得到的旋转和平移作为观测测量,对误差状态进行估计。最后,将误差状态和名义状态结合,得出真实状态。本文所选择的状态有位置 p 、姿态四元数 q 、速度 v 、加速度零偏 a_b 、角速度零偏 ω_b 。名义状态在离散时间下的运动学公式为^[19]:

$$\begin{cases} q \leftarrow q \otimes q \{ (\omega_m - \omega_b) \Delta t \} \\ p \leftarrow p + v \Delta t + \frac{1}{2} (R(a_m - a_b) + g) \Delta t^2 \\ v \leftarrow v + (R(a_m - a_b) + g) \Delta t \\ \omega_b \leftarrow \omega_b \\ a_b \leftarrow a_b \end{cases} \quad (4)$$

其中, g 表示重力加速度;“ \leftarrow ”表示时间更新;姿态与角速度的运算 $q \otimes \omega$ 是四元数乘法一阶泰勒展开得到的四元数积分离散形式;“ $\{ \}$ ”表示旋转向量的四元数形式。

误差状态的形式可以由其定义直接得出,即各状态所对应的误差,分别为姿态误差 $\delta\theta$ 、位置误差 δp 、速度误差 δv 、角速度零偏误差 $\delta\omega_b$ 、加速度零偏误差 δa_b 。误差状态在离散时间的表达如式(5)所示:

$$\begin{cases} \delta\theta \leftarrow \delta\theta - R\delta\omega_b \Delta t - R\omega_n \\ \delta p \leftarrow \delta p + \delta v \Delta t \\ \delta v \leftarrow \delta v + (-[R(a_m - a_b)]_x \delta\theta - R\delta a_b) \Delta t + R a_n \\ \delta\omega_b \leftarrow \delta\omega_b + \omega_{bn} \\ \delta a_b \leftarrow \delta a_b + a_{bn} \end{cases} \quad (5)$$

其中,“ $[]_x$ ”表示向量的反对称矩阵。

与名义状态以IMU的输入为控制量不同,误差状态以系统的高斯白噪声为控制量,可表示为 $n = [\omega_n a_n \omega_{bn} a_{bn}]$ 。用式(4)和式(5)可以得出名义状态和误差状态的状态转移矩阵,并对滤波器进行时间更新,获得先验状态。

当完成点云配准后,得到了摄像头在世界坐标系下的姿态 q_c 和位置 p_c 。由于摄像头所处坐标系

和 IMU 不同,首先将坐标转换到 IMU 坐标系下,此后进入测量更新过程。对误差状态的更新过程和卡尔曼滤波相同,计算卡尔曼增益 \mathbf{K} 、更新误差状态和协方差 δx 和 Σ ,推得的公式为:

$$\begin{cases} \mathbf{K} = \Sigma \mathbf{H}^T (\mathbf{H} \Sigma \mathbf{H}^T + \mathbf{Q}_{ndt})^{-1} \\ \delta x \leftarrow \mathbf{K} (y - h(x)) \\ \Sigma \leftarrow (\mathbf{I} - \mathbf{K} \mathbf{H}) \Sigma \end{cases} \quad (6)$$

其中, \mathbf{Q}_{ndt} 表示点云匹配算法的测量噪声矩阵、 \mathbf{H} 表示测量值相对于 δx 的雅可比矩阵。

利用式(7),将误差状态和名义状态求和,得出真实状态并输出:

$$\begin{cases} q_t = \delta q \otimes q \\ p_t = p + \delta p \\ v_t = v + \delta v \\ \omega_{bt} = \omega_b + \delta \omega_b \\ a_{bt} = a_b + \delta a_b \end{cases} \quad (7)$$

IMU 传感器对运动的测量频率高,对物体短时间的快速运动也能保持较高精度,可以与对低频运动更加精确的视觉传感器形成互补,所估计的轨迹更平滑^[20]。但由于其高频特性,与视觉传感器的频率不一致,需要对数据进行同步,不同传感器的同步方式如图3所示。因此每次算法接收到图像帧时,都应该取与图像帧时间戳最接近的 IMU 测量值所对应的先验状态进行测量更新。又由于帧间点云配准所需的计算时间往往比 IMU 的测量周期长,因此在测量更新后,需要对摄像头时间戳后的 IMU 测量值重新积分,保证算分对任意时刻估计值的合理性^[6]。

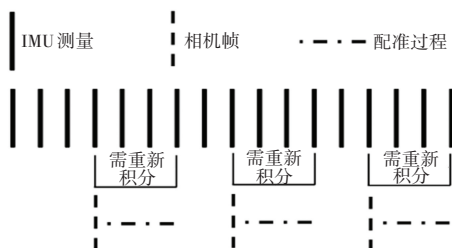


图3 不同传感器同步的方式

Fig. 3 Time synchronization of different sensors

2 仿真结果

在机器人操作系统(ROS)中建立用于仿真的停车场环境和搭载传感器的机器人。实验场景设置为一个停车场模型,地面材质为灰黑色,路面上有黄色以及白色的标志和停车线等特征。场景为方形停车场,场地大小 $90 \times 25 \text{ m}^2$ 。使用一个 TurtleBot 作为仿真平台,TurtleBot 是一个双轮差速模型的移动机器

人,修改其 Unified Robot Description Format (URDF) 模型,使其适合当前实验的大小和动力需求。在机器人上添加研究所需的 RGB 相机和 IMU 传感器,建立相应的配置文件,并添加噪声。摄像头以 72° 均匀摆放,其中一台朝向车辆正前方。相机帧数为 10 FPS,IMU 频率为 300 Hz。

为了验证本文提出的定位算法,使用 Gazebo 仿真环境进行了3次仿真。第一次仿真,行驶距离短,测试将车停入停车位过程中的定位精度,行驶距离 14 m;第二次,在停车场小范围中绕圈行驶,测试定位算法对旋转的鲁棒性,行驶距离 130 m;第三,次增大行驶范围,测试在大场景中的定位精度,行驶距离 117 m。第二次和第三次仿真中行驶过的路径如图4所示,仿真过程中建立的停车场点云地图如图5所示。

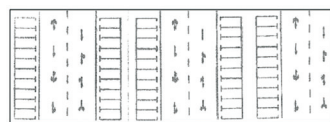
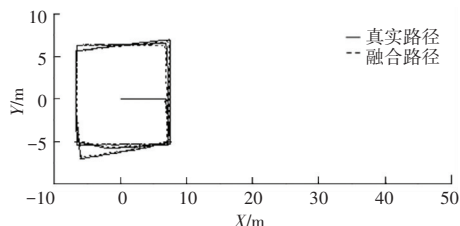
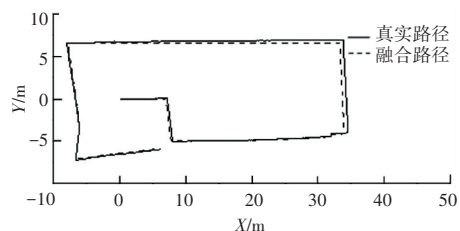


图4 仿真中建立的停车场全局点云地图

Fig. 4 Point cloud map of parking lot constructed in simulation



(a) 第二次仿真路径



(b) 第三次仿真路径

图5 仿真的行驶路径

Fig. 5 Drive path of simulations

计算误差时,为保证真值的实时性,定时请求由 Gazebo 托管的服务端获取位姿真实值,并记录真实值和状态,据此进行数据分析。使用误差均方根、绝对位姿误差最大值和平均值对定位精度进行了评价,结果见表1。

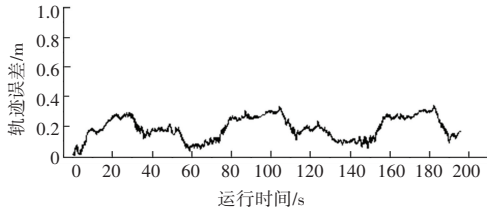
由于第二次和第三次仿真的时间较长,所累计的误差大,2次仿真的绝对轨迹误差 (ATE) - 时间图如图6所示。由图6(a)可知,即使轨迹中旋转较多,但误差并没有持续累积,说明本文提出的方法对旋转的鲁棒性较高。由图6(b)可知,当在大场景中

定位时,误差的值会随距原点距离增加而增大,但当车辆行驶回距原点近的位置时,误差发生回落。

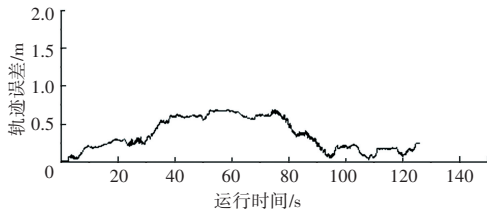
表1 仿真结果绝对轨迹误差

Table 1 Absolute Trajectory Error produced by simulations m

	误差均方根	误差最大值	平均误差
第一次仿真精度	0.108 5	0.231 4	0.075 0
第二次仿真精度	0.201 9	0.343 7	0.186 9
第三次仿真精度	0.413 9	0.696 2	0.353 2



(a) 第二次仿真



(b) 第三次仿真

图6 仿真的绝对姿态误差-时间图

Fig. 6 Absolute Trajectory Error-time of simulations

3 结束语

本文对环视图的拼接、视觉里程计和视觉里程计与IMU测量数据融合的算法做出了讨论,而后仿真测试了视觉惯性里程计的定位精度。

根据仿真结果可知,在停车场环境中使用地面标记和道路交通标志线可以对车辆进行定位,在移动范围不大的情况下可以较好地满足自主泊车任务的定位精度。本文提出的视觉惯性里程计并不会由于轨迹中转弯的数量增多而导致定位精度下降,适合为自主泊车等行驶过程中转弯多、行驶距离不大的任务提供较为精确的位置信息。

参考文献

[1] IM G, KIM M, PARK J. Parking line based SLAM approach using AVM/LiDAR sensor fusion for rapid and accurate loop closing and parking space detection[J]. Sensors, 2019, 19(21): 4811.

[2] QIN Tong, CHEN Tongqing, CHEN Yilun, 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). IEEE, 2020: 5939-5945.

[3] CUI L, RONG C, HUANG J, et al. Monte-Carlo localization in underground parking lots using parking slot numbers[C]//2021

IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 2267-2274.

[4] 李庆玲,翟凯,郭鸿锐,等.一种基于NDT和ICP融合的点云配准算法[J].实验技术与管理,2022,39(11):23-28,39.

[5] Magnusson M, Nuchter A, Lorken C, et al. Evaluation of 3D registration reliability and speed-A comparison of ICP and NDT [C]//2009 IEEE International Conference on Robotics and Automation. Kobe; IEEE, 2009: 3907-3912.

[6] CHEN Shengyang, CHANG C W, WEN C Y. Perception in the dark; development of aToF visual inertial odometry system[J]. Sensors, 2020, 20(5): 1263.

[7] 曾庆化,罗怡雪,孙克诚,等.视觉及其融合惯性的SLAM技术发展综述[J].南京航空航天大学学报,2022,54(6):1007-1020.

[8] 崔文,薛棋文,李庆玲,等.基于三维点云地图和ESKF的无人车融合定位方法[J].工矿自动化,2022,48(9):116-122.

[9] CAMPOS C, ELVIRA R, RODRIGUEZ 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.

[10] ROUMELIOTIS S I, SUKHATME G S, BEKEY G A. Circumventing dynamic modeling: Evaluation of the error-state kalman filter applied to mobile robot localization [C]// Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C). Detroit, USA; IEEE, 1999, 2: 1656-1663.

[11] 周哲,胡剑政,李娜,等.面向智能车的地下停车场环视特征地图构建与定位[J].测绘学报,2021,50(11):1574-1584.

[12] PAN Y, YANG B, LI S, et al. Automatic road markings extraction, classification and vectorization from mobile laser scanning data[J]. Copernicus GmbH.2019.DoI:10.5194/ISPRS-ARCHIVES-XL11-2-W13-1089-2019.

[13] CHRISTIE D, JIANG CANSEN, PAUDEL D, et al. 3D reconstruction of dynamic vehicles using sparse 3D-laser-scanner and 2D image fusion [C]//2016 International Conference on Informatics and Computing (ICIC). Mataram, Indonesia ;IEEE, 2016: 61-65.

[14] BIBER P, STRAßER W. The normal distributions transform: A new approach to laser scan matching [C]//Proceedings of 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453). Las Vegas, USA; IEEE, 2003, 3: 2743-2748.

[15] ZHANG Lintong, WISTH D, CAMURRI M, et al. Balancing the budget: Feature selection and tracking for multi-camera visual-inertial odometry [J]. IEEE Robotics and Automation Letters, 2021, 7(2): 1182-1189.

[16] LIU Xinyu, ZHOU Qingfeng, CHEN Xiang, et al. Bias-error accumulation analysis for inertial navigation methods [J]. IEEE Signal Processing Letters, 2021, 29: 299-303.

[17] 盛树轩,荆崇波,蒋朝阳.视觉与单路侧单元辅助的车辆定位方法[J].汽车工程,2022,44(7):1009-1017.

[18] BRIGADNOV I, LUTONIN A, BOGDANOVA K. Error state extended Kalman filter localization for underground mining environments[J]. Symmetry, 2023, 15(2): 344.

[19] SOLA J. Quaternion kinematics for the error-state Kalman filter [J]. arXiv preprint arXiv:1711.02508,2017.

[20] KUMAR K, VARGHESE A, REDDY P K, et al. An improved tracking using IMU and vision fusion for mobile augmented reality applications[J]. arXiv preprint arXiv:1411.2335, 2014.