引文格式:沈燕芬,张小红,朱锋. ORB-SLAM2 车载双目视觉的自主导航性能评估[J]. 导航定位学报,2018,6(2);29-35. (SHEN Yanfen, ZHANG Xiaohong, ZHU Feng. Performance evaluation on autonomous navigation of ORB-SLAM2 vehicle stereo vision[J]. Journal of Navigation and Positioning,2018,6(2);29-35.) DOI;10.16547/j. cnki. 10-1096. 20180206.

ORB-SLAM2 车载双目视觉的自主导航性能评估

沈燕芬1,张小红1,2,朱 锋1

(1. 武汉大学 测绘学院, 武汉 430079; 2. 地球空间信息技术协同创新中心, 武汉 430079)

摘要:针对全球定位系统用于车载导航在复杂环境下定位精度较低、低成本惯导漂移明显等问题,给出 $ORB ext{-}SLAM2$ 用于车载双目视觉自主导航的原理和方法,介绍特征匹配与提取及位姿估计优化部分的关键技术,评估采用不同轨迹、不同运动状态的车载数据的性能,重点分析 $ORB ext{-}SLAM2$ 整体的定位定姿精度、累积误差随距离的变化以及回环检测对车载定位定姿的影响。实验结果表明:持续的动态物体干扰会对 $ORB ext{-}SLAM2$ 整体的导航精度有所影响;无回环探测时,位置误差随行驶距离呈现明显的 1~%线性发散,成功的回环检测能够有效抑制位置姿态误差的发散并提高其精度,定位精度保持在 2~m 左右。

关键词: ORB-SLAM2; 自主导航; 性能评估; 双目视觉

中图分类号: P228 文献标志码: A 文章编号: 2095-4999(2018)02-0029-07

Performance evaluation on autonomous navigation of ORB-SLAM2 vehicle stereo vision

SHEN Yanfen¹, ZHANG Xiaohong^{1,2}, ZHU Feng¹

- (1. School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China;
- 2. Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China)

Abstract: Aiming at the problems that the positioning accuracy is low and the drift of low-cost inertial navigation is obvious in the vehicle navigation with GPS under complex environment, the paper gave the principle and methods of autonomous navigation with vehicle stereo vision by ORB-SLAM2, introduced the key technology of feature matching and extraction, pose estimation and optimazation, and evaluated the vehicle-data performance with different traces and motion states, among which the accuracy of positioning and attitude, the cumulative errors varied with the distances as well as the influence of loop detection for ORB-SLAM2 were mainly focused. Experimental result showed that the navation accuracy would be slightly affected by continuously dynamic objects in view of the cameras; when the loop detection was failed, the position errors would show a linear divergence of 1 % with driving distances, and the error divergence of positions and attitudes could be effectively suppressed and the accuracy of those could be improved by successful loop detection; moreover, the accuracy of positioning would keep around 2 m.

Keywords: ORB-SLAM2; autonomous navigation; performance evaluation; stereo vision

0 引言

近几年来,随着无人机、无人驾驶汽车、智能移动机器人相关技术的飞速发展,视觉导航作为一种非传统的定位技术被广泛应用。文献 [1-2] 提到:视觉导航仅利用单一相机传感器完成,精度稳定且无需场景和运动的先验信息,可避免由

于传感器精度降低或低成本惯性测量单元(inertial measurement unit, IMU)漂移严重引起的误差发散问题;其被动无源的特点,也弥补了全球卫星导航系统(global navigation satellite system, GNSS)在复杂城市或室内环境下信号易受遮挡而无法定位的缺陷。

自 1982 年文献 [3] 首次提出了一套完整的基

收稿日期: 2017-10-27

第一作者简介:沈燕芬(1994--),女,浙江嘉兴人,硕士研究生,研究方向为基于视觉导航的多源信息组合导航。

于计算机视觉的理论框架以来,视觉导航领域的 相关算法和应用层出不穷。视觉导航算法根据其 对图像信息的利用程度可分为特征法和直接法。 直接法直接对图像所有像素点进行操作,可充分 利用环境几何信息,其代表作为文献[4]的基于 直接法的大范围单目即时定位和地图构建方法 (large-scale direct monocular simultaneous localization and mapping, LSD-SLAM), 在精确计算自 身位姿的前提下,还可构建出大规模的半稠密地 图。而单目定位构图(monocular simultaneous localization and mapping, MonoSLAM)、平行跟 踪与地图构建(parallel tracking and mapping, PTAM) 等算法则采用了对图像关键信息进行检测 提取的特征法。MonoSLAM 基于扩展卡尔曼滤波 器实现,将相机位姿及地图点 3D 坐标作为状态向 量进行观测更新,算法计算复杂度较高,且线性 化也造成了一定误差。文献[5]提出基于关键帧 提取及光束法平差优化的 PTAM 算法,首次将定 位与构图分为 2 个独立工作的线程。2014 年文 献「6]提出的半稠密单目视觉里程计算法(semi-direct monocular visual odometry, SVO)则融合了特 征法与直接法,首先采用加速分割检测特征(features from accelerated segment test, FAST)角点 跟踪关键点,再根据关键点的周围信息估计帧间 位姿,其实时性高、算法简洁,非常适用于平台 受限的场合,如无人机、手持增强现实(augmented reality, AR)及虚拟现实(virtual reality, VR) 设备等,该法对位姿估计精度依赖性强,易出现 丢失和重定位的情况。为达到实时性的自主导航 定位要求, 文献 [7] 将特征信息在用于位姿估计 后不进行保留,后续优化仅依据位姿关系,减轻 了存储计算负担,其基于光束法平差(bundle adjustment, BA)的实时视觉定位构图算法成为了在 大范围内进行定位构图的开创性作品。文献[8] 的基于外观的实时定位构图方案(real-time appearance-based mapping, RTAB-MAP)则从内 存管理机制角度出发、将定位点根据被访问的频 率及存在的时间长度设置权重,以此为依据进行 分类与修剪,从而提高了回环检测的效率。

文献 [9-10] 提出的基于旋转不变特征点的实时定位构图系统 (oriented brief-SLAM, ORB-SLAM)被公认为现有综合性能最好的视觉导航算法: 在整体架构上延续了 PTAM 的多线程思想;但改进了 PTAM 需手动进行初始化的缺陷,扩展了应用环境的多样性和系统的交互性。ORB-SLAM

实现了基于单目相机的实时算法,而 ORB-SLAM2 在此基础上扩展为双目、RGBD 相机,并被广泛 应用于采集车、无人机及室内外机器人等设备。

本文采用不同采集环境的车载数据对双目视 觉自主导航性能进行评估,重点分析其整体定位 定姿精度、累积误差随距离的变化情况以及回环 检测对定位结果的影响。

1 利用 ORB-SLAM2 进行位姿估计

ORB-SLAM2 整个系统分为位置跟踪、局部地图优化、回环检测 3 个并行处理的线程,其整体架构如图 1 所示。

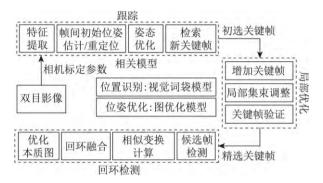


图 1 ORB-SLAM2 算法整体架构

跟踪线程主要负责每帧影像的位姿估计。经 畸变矫正的双目序列影像和相机的标定参数输入 至系统后,采用具有方向信息的多尺度 FAST 算 子对影像进行特征提取,并用 256 位的具有方向信 息的二元加速鲁棒特征(oriented binary robust independent elementary features, rBRIEF)二进制描 述符表示特征点。根据小孔成像原理以及双目相 机高度一致、光轴平行的特点,左右影像匹配后 可重构出每对匹配点在相机坐标系中的坐标。将 第一帧影像的左相机坐标系作为系统参考坐标系, 通过序列影像的帧间匹配可将特征点的世界坐标 进行传递和扩展,在帧间进行 3D-3D 对应估计可 获得参考坐标系与每帧影像的相对位姿关系。设 第 i 帧中某匹配特征点在该帧相机坐标系中的坐标 为(x,y,z),与i-1帧进行匹配可得该点在系统 参考坐标系中的坐标为(X,Y,Z),即

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R}_{Re}^{i} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \mathbf{t}_{Re}^{i}$$
 (1)

式中 R_{Re} 、 t_{Re} 为系统参考坐标系相对于第 i 帧相机 坐标系的旋转矩阵和平移向量。由于误匹配点的存在使得直接线性求解变得十分不可靠,故通常 先采用运动模型或关键帧模型获得初始的位姿矩

阵,再基于图优化模型最小化 3D—2D 重投影误差函数迭代求解以得到更精确的位姿,其误差函数为

$$\begin{aligned} & \{\boldsymbol{R}_{\mathrm{Re}}^{i}, t_{\mathrm{Re}}^{i}\} = \underset{\boldsymbol{R}_{\mathrm{Re}}^{i}, t_{\mathrm{Re}}^{i}}{\operatorname{sign}} \sum_{j=1}^{N} \rho(\mathrm{E}(i,j)) \\ & \mathrm{E}(i,j) = \| (u_{j}, v_{j}) - \operatorname{project}(\boldsymbol{R}_{\mathrm{Re}}^{i} \boldsymbol{X}^{j} + \boldsymbol{t}_{\mathrm{Re}}^{i}) \|_{\Sigma}^{2} \end{aligned}$$

式中: (u_j, v_j) 为第 j 个特征点的像素坐标;N 为匹配到的特征点总数; ρ 为 Huber 损失函数用于外点的剔除;project()表示 3D-2D 重投影函数; \mathbf{X}^j 为第 j 个匹配点在系统参考坐标系中的 x 坐标; $\mathbf{\Sigma}$ 为和每个特征点的尺度相关的信息矩阵。

在对每帧影像进行以上操作的同时,局部地图优化线程存储关键点和关键帧,实时地优化局部若干帧的位姿及相应 3D 点的位置,对误差进行修正;而回环检测线程则通过回环探测和融合实现全局位姿优化。

通过位姿估计得到帧间的旋转平移关系后,可以初始位置为坐标原点,仅考虑车辆在系统参考坐标系的运动情况,反演出载体的行驶轨迹为

$$\mathbf{P}_{0} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^{T} \\
\mathbf{P}_{i} = \mathbf{R}_{Re}^{i} \cdot (-\mathbf{t}_{Re}^{i}), 1 \leqslant i \leqslant m$$
(3)

式中: P_0 为初始时刻相机在系统参考坐标系的位置坐标; P_i 表示第i 帧影像的相机拍摄中心在系统参考坐标系下的位置坐标;m 为总帧数。若已知初始坐标原点的绝对位置,则可通过坐标变换递推得车辆在世界坐标系下的绝对行驶轨迹,实现车载双目视觉导航。

2 ORB-SLAM2 位姿估计的关键技术

ORB-SLAM2 可有效对目标物体进行跟踪,即使丢帧也可快速重定位,能够实现可靠的自主导航;其关键技术集中在特征提取与匹配和位姿精度优化2部分。文献 [11] 提到,特征提取与匹配的准确度保证了每帧左右及帧间影像重构的三维点的准确度,从而决定了位姿估计的整体精度,故提高特征提取与匹配的正确率是算法提升自主导航鲁棒性的有力措施;而多次的位姿精度优化则是在此基础上的双重保障。

1)特征提取与匹配:特征在提取时候做了网格处理,从而避免了特征点选取的不均匀性,在左右影像间进行特征匹配时相继用到了带状搜索、绝对差值和 SAD(sum of absolute differences,一种图像匹配算法)滑窗以及抛物线拟合以确保匹配点达到亚像素级的精度;帧间匹配,在缺乏先验

信息时,抛弃了传统的高维图像特征描述子比对,转而使用词袋模型(bag of words, BoW)对描述子进行分类,通过类别节点的设置大大简化了运算,反之则通过重构特征点、三角化或相似变换约束缩小搜索范围以加速匹配,并对特征点加入了正向景深、视差距离、反投影误差、尺度及旋转一致性的严格判定。

2) 位姿精度优化:除单帧位姿优化外,系统还将在关键帧插入后对局部窗口帧及地图点进行局部光束法平差(local bundle adjustment, Local BA),在成功进行回环修正后对所有关键帧及地图点进行全局光束法平差(full bundle adjustment, Full BA)。Local BA 对由关键帧和地图点的连接关系进行优化,构造而成的共视图(covisibility graph, CG)图 2 所示。

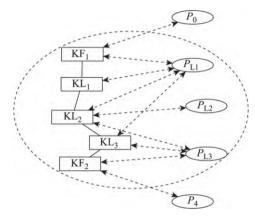


图 2 Local BA 优化共视

如图所示,邻接共视帧 KL_1 、 KL_2 、 KL_3 的位 姿及其共视点 P_{L1} 、 P_{L2} , P_{L3} 的位置在优化中被修正,而包含共视点但非邻接关键帧 KF_1 、 KF_2 的位 姿则在优化中被设为固定,只参与计算,不进行优化。若帧间特征点的匹配关系(用 χ 表示)已知,则其优化函数可表示为

$$\begin{cases}
X_{i \in P_L, l \in KL}^i, R_l, t_l \\
i \in P_L, l \in KL
\end{cases} = \underset{\mathbf{X}^i, \mathbf{R}_l, t_l}{\operatorname{argmin}} \sum_{k \in KL \cup KF} \sum_{j \in \chi} \rho(E(k, j)) \\
E(k, j) = \| (u_j, v_j) - \operatorname{project}(\mathbf{R}_k \mathbf{X}^j + \mathbf{t}_k) \|_{\mathbf{\Sigma}}^2$$
(4)

式中:R、t分别为对应帧的旋转矩阵和平移向量;下标 1 表示对应的共视帧;下标 k 表示共视帧与关键帧的交集。当通过一系列自适应阈值筛选得到可靠的候选回环帧并进行回环修正后,系统会另开辟线程执行 Full BA。Full BA是 Local BA 的扩展,除初始帧外,所有关键帧和地图点均被优化;故可有效解决因误差累积导致的尺度漂移问题,大大提高系统的定位测姿精度,提升双目自主导航的可靠性与稳定性。

3 实验与结果分析

为了对基于 ORB-SLAM2 的车载双目视觉自主导航进行全面的性能评估,本节将采用不同道路采集环境的数据对其算法的整体定位测姿性能及其累积误差随载体行驶距离的变化、回环检测对自主导航结果的影响等方面做细致分析。

本次分析采用的数据为卡尔斯鲁厄工学院和丰田技术学院(Karlsruhe Institute of Technology and Toyota Technological Institute,KITTI)联合提供的数据集的序列影像,其采集环境主要为高速公路、公路及居民街区。实验设备包含 1 套OXTS RT 3003 的 GPS/IMU 组合导航系统、4 个灰点公司的 Flea 2 系列相机、4 个埃蒙德的 $4 \sim 8$ mm变焦镜头,分析数据仅涉及其中 2 个灰色相机。相机与地平面平行安装,采样帧率为 10~Hz,其实验设备和不同环境的采集影像如图 3~M 所示。

3.1 整体精度分析

以差分 GPS/IMU 的组合定位结果为参考值,解算轨迹与参考值的偏差值大小,结果由不同灰度值表示,如图 4 所示,灰度越深代表误差越大。



(a) 实验采集设备



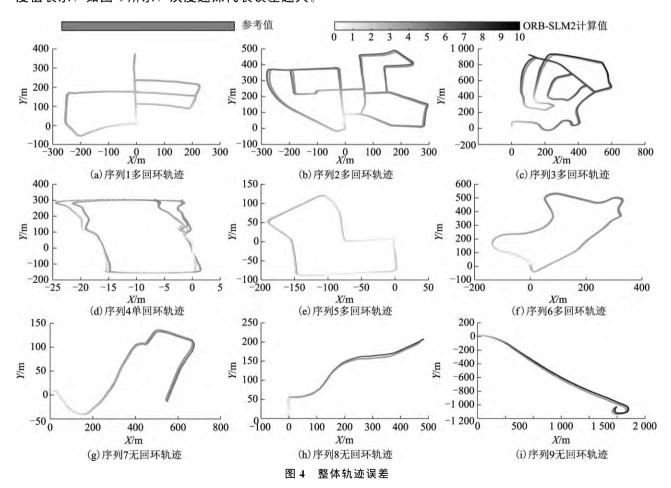






(b) 采集环境分别为高速公路、公路及居民街区

图 3 实验装备与数据采集



由图 4 可知,所有轨迹均与真值重合度较高,整体定位精度约为 1.02%,即行驶 100 m 偏离真值 1 m,角度误差为 0.0019 deg/m,各方向定位精度在 $4\sim5$ m 左右,定姿精度在 $0.6\sim0.7$ deg 之间,如表 1 所示。

表 1 ORB-SLAM2 的整体精度统计

影像序列	差相对		位置均方根差/m			角度均方根差/deg		
彩冰炉外				У	z	航向	横滚	俯仰
多闭环序列	0. 562	0,0007	4. 49	5, 83	3 81	0.86	0, 63	0. 65
单闭环序列	0. 513	0.0014	1, 61	2 22	0.93	0.49	0, 5	0. 32
无闭环序列	1. 970	0.0036	10, 53	7. 19	6, 52	0, 45	0, 99	0. 81
平均	1. 020	0.0019	5, 54	5, 08	3, 76	0, 60	0.71	0, 60

由表 1 可知:回环轨迹的整体定位定姿精度显著优于无回环轨迹;单闭环轨迹在回环融合附近的误差修正效果明显,而无回环轨迹随着行驶距离的增加误差明显逐渐发散。对多回环影像序列 2、序列 3 的异常轨迹分析发现,大转弯、照片过曝或阴影遮挡会影响定位测姿精度,而影像中的在相机视野内持续运动物体会严重干扰位姿估计,如图 5 所示,3 号轨迹中异常轨迹段中,有一辆摩托跟随行驶了 500 多米。



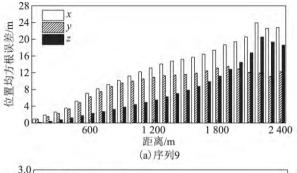
图 5 3 号序列影像中特殊场景分析

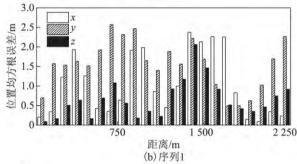
3.2 累积误差随距离变化的影响分析

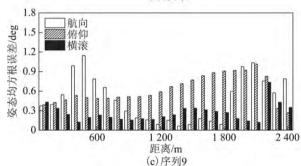
增量式的位姿估计会导致绝对位置误差不断累 积,位姿愈发发散,从而最终限制系统的整体精度。

选取无回环的 9 号高速公路影像(全长 2 451 m)及多回环的 1 号居民街区影像(全长 2 204 m)进行处理,并统计其累积误差(滑窗宽度为 100 m,统

计窗口内为位置误差及姿态估计误差的均方根误差),如图 6 所示。







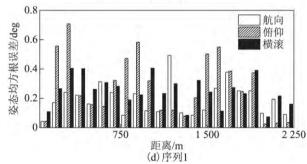


图 6 位置、姿态累积误差随距离的变化

由图 6 可知: 高速公路环境下的行驶路线,各方向位置精度随距离增长呈明显线性发散趋势,行驶到 2 500 m 左右时,侧向精度将近达到 25 m,而居民街区轨迹由于多次的回环修正,各方向精度均始终保持在 2 m 左右;姿态精度在 2 种采集模式下无明显规律可循,均始终保持在 1 deg 以内,原因在于相对姿态误差在局部的位姿优化中有很大程度的修正,而回环影像的优势主要在于可修正由误差积累所导致的尺度漂移的问题。故无回环轨迹各方向的位置精度随距离呈约 1 % 的线性发散,而经过回环修正则可使精度始终保持在 2 m,

提高了车载自主导航的鲁棒性。

3.3 回环修正对结果的影响分析

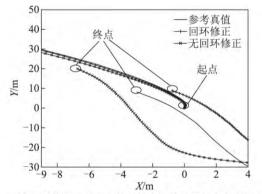
为进一步定量分析回环修正对定位测姿精度的改善情况,本小节选取 4 组长距离回环序列影像进行分析,对有回环修正和无回环修正 2 种模式下的定位测姿精度分别进行统计,并重点对 6 号序列影像进行细致分析,如表 2 所示。

表 2 回环修正结果精度统计

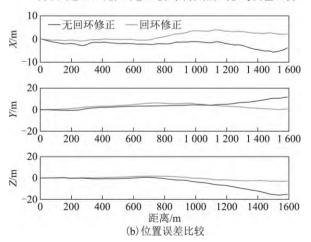
影像序列	位置制	精度/m	角度精度 /deg			
	回环修正	无回环修正	回环修正	无回环修正		
回环 1	2, 626 5	4. 725 1	0. 563 2	1. 212 8		
回环 2	1. 073 4	1. 908 4	0. 475 6	0. 591 8		
回环 3	14. 211 2	15. 823 7	1. 253 5	4. 052 8		
回环 4	4. 622 4	8. 717 8	0. 683 5	1. 680 9		

由表可知,针对不同序列影像,闭环检测平均约可改善 36 %的位置精度和 50 %的姿态估计精度。

进一步分析比较 6 号影像轨迹(如图 7 所示)可



(a) 回环修正、无回环修正模式计算结果与参考真值比较



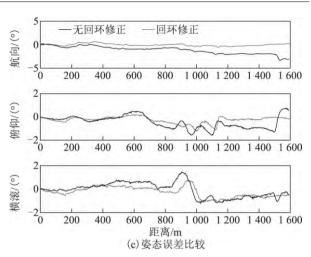


图 7 6 号影像轨迹比较

以发现:没有进行回环修正的轨迹尺度漂移明显,各方向的位置逐渐偏离参考值,误差逐渐增大,姿态误差也存在发散趋势,但幅度较小;反之,成功的回环修正可使得起点与终点高度重合,明显修正位置误差,起到抑制误差发散、优化整体位姿精度的作用。

4 结束语

本文介绍了利用 ORB-SLAM2 车载双目视觉自 主导航的基本位姿估计原理及其关键技术,重点对 ORB-SLAM2 的自主导航性能进行评估,分析了其 整体定位测姿精度、累积误差随行驶距离的变化情 况以及回环修正对结果的具体影响。对一定量序列 影像数据分析得出:该算法的整体定位相对精度 约为 1.02 %, 定姿相对精度约为 0.001 9 deg/m; 在无回环探测时,位置误差随行驶距离呈现明显 的 1 %线性发散;成功闭环探测约可提高 36 %的 位置精度和 50 %的姿态精度,且无发散情况,定 位精度始终保持在 2 m 左右。故该算法存在和其 他导航系统或传感器,如 GNSS、IMU 等进行组 合导航的理论基础。目前视觉导航发展的难点和 热点为高精度、高效率及多传感器融合,故在接 下来的研究中,将着重对视觉导航和 IMU 及 GNSS 组合系统的相对姿态标定以及组合导航方案 进行深入研究,望实现一套完整可靠、低成本、 高精度的双目视觉辅助惯导/GNSS 的组合导航实 时系统。

参考文献

- [1] SCARAMUZZA D, FRAUNDORFER F. Visual odometry: Part I: the first 30 years and fundamentals[J]. IEEE Robotics & Automation Magazine, 2011, 18(4): 80-92.
- [2] 权美香,朴松昊,李国. 视觉 SLAM 综述[J]. 智能系统学报,2016,11(6):768-776.

- [3] 卢维. 高精度实时视觉定位的关键技术研究[D]. 杭州:浙江大学,2015.
- [4] ENGEL J, SCHÖPS T, CREMERS D. LSD-SLAM: large-scale direct monocular SLAM[C]//Anon. Proceedings of European Conference on Computer Vision. Zurich: Springer, 2014:834-849.
- [5] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces [EB/OL]. [2017-02-19]. http://www.robots.ox.ac.uk/~gk/publications/KleinMurray2007ISMAR.pdf.
- [6] FORSTER C, PIZZOLI M, SCARAUZZA D. SVO: fast semi-direct monocular visual odometry [EB/OL]. [2017-02-19]. http://rpg. ifi. uzh. ch/docs/ICRA14_Forster. pdf.
- [7] KONOLIGE K, AGRAWAL M. FrameSLAM: from bundle adjustment to real-time visual mapping[J]. IEEE Transactions on Robotics, 2008, 24(5):1066-1077.
- [8] LABBÉ M, MICHAUD F. Appearance-based loop closure detection for online large-scale and long-term operation[J]. IEEE Transactions on Robotics, 2013, 29(3); 734-745.
- [9] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM; a versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2017, 31(5):1147-1163.
- [10] 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, 2016, 33(5):1255-1262.
- [11] FRAUNDORFER F, SCARAMUZZA D. Visual odometry: Part II: matching, robustness, optimization, and applications [J]. IEEE Robotics & Automation Magazine, 2012, 19(2): 78-90.

(上接第22页)

- [4] MILLER R. A new strapdown attitude algorithm [J]. Journal of Guidance, Control, and Dynamics, 1983, 6 (4): 287-291.
- [5] PARK C G, KIM K J, LEE J G, et al. Formalized approach to obtaining optimal coefficients for coning algorithms [J]. Journal of Guidance, Control, and Dynamics, 1999, 22(1):165-168.
- [6] SAVAGE P G. Strapdown inertial navigation integration algorithm design part 1:attitude algorithms[J] Journal of Guidance, Control, and Dynamics, 1998, 21(1):19-28.
- [7] ROSCOE K M. Equivalency between strapdown inertial navigation coning and sculling integrals/algorithms[J]. Journal of Guidance, Control, and Dynamics, 2001, 24(2):201-205.
- [8] 严恭敏,严卫生,徐德民. 经典圆锥误差补偿算法中剩余误差估计的局限性研究[J]. 中国惯性技术学报,2008,16(4): 379-385.
- [9] WANG M S, WU W Q, WANG J L, et al. High-order attitude compensation in coning and rotation coexisting environment [J]. IEEE Transactions on Aerospace and Electronic Systems, 2015, 51(2):1178-1190.
- [10] 严恭敏,翁浚,杨小康,等.基于毕卡迭代的捷联姿态更新精确数值解法[J].宇航学报,2017,38(12):1307-1313.
- [11] 武元新,郁文贤. 一种基于函数迭代积分的刚体姿态解算方法:201710273489.3[P].2017-11-10.
- [12] 严恭敏,杨小康,翁浚,等. 一种求解姿态不可交换误差补偿系数的通用方法[J]. 宇航学报,2017,38(7):723-727.
- [13] 李中华. Su-27 飞机眼镜蛇机动及其战术意义[J]. 飞行力学,2000,18(1):54-57.