

一种基于边缘特征的无人机视觉定位方法

刘春江¹, 张鹏宇²

(1 四川航天系统工程研究所 成都 610100;

2 北京遥测技术研究所 北京 100076)

摘要: 针对无人机在无导航信号环境中低空动态飞行时, 视觉同时定位与建图 (Simultaneous Localization And Mapping, SLAM) 定位方法在图像快速变化和过曝光时稳定性差的问题, 本文提出一种基于边缘特征的无人机视觉定位方法, 通过对传统特征提取算法的数据进行降维, 生成边缘特征, 利用卷积神经网络进行连续关键帧之间边缘特征匹配, 得到边缘特征重投影误差函数, 最终通过非线性优化完成位姿估计。实验结果表明: 在数据集下, 与最新的 ORB-SLAM3 (支持视觉、视觉加惯导、混合地图的 SLAM 系统) 算法相比, 新方法定位时间缩短 31%, 在低纹理场景中, 定位精度提高 15.04%。飞行实验结果表明无人机定位的准确性和稳定性得到了显著提高。

关键词: 无人机; 视觉定位; SLAM; 边缘特征; 卷积神经网络

中图分类号: V279; TP391 文献标志码: A 文章编号: 2095-1000(2024)06-0093-06

DOI: 10.12347/j.ycyk.20240122001

引用格式: 刘春江, 张鹏宇. 一种基于边缘特征的无人机视觉定位方法[J]. 遥测遥控, 2024, 45(6): 93-98.

A Visual Localization Method for Unmanned Aerial Vehicles Based on Edge Features

LIU Chunjiang¹, ZHANG Pengyu²

(1. Sichuan Aerospace Systems Engineering Research Institute, Chengdu 610100, China;

2. Beijing Research Institute of Telemetry, Beijing 100076, China)

Abstract: Aiming at the issue of poor stability of visual localization and mapping (SLAM) methods during dynamic low-altitude flight of unmanned aerial vehicles (UAVs) in the absence of navigation signals, this paper proposes a UAV visual localization method based on edge features, which generates the edge features by downsizing the traditional feature extraction algorithm and finally completes the position estimation by nonlinear optimization. A convolutional neural network is employed to match edge features between consecutive key frames, yielding an edge feature reprojection error function, and finally the position estimation is completed by nonlinear optimization. The experimental results demonstrate that compared to the state-of-the-art ORB-SLAM3 algorithm, the proposed method reduces localization time by 31% on the dataset and improves localization accuracy by 15.04% in low-texture scenes. Flight experiments further indicate a significant enhancement in the accuracy and stability of UAV localization.

Keywords: UAV; Visual localization; SLAM; Edge feature; Convolutional neural network

Citation: LIU Chunjiang, ZHANG Pengyu. A Visual Localization Method for Unmanned Aerial Vehicles Based on Edge Features[J]. Journal of Telemetry, Tracking and Command, 2024, 45(6): 93-98.

0 引言

随着无人机相关技术的进步和应用场景的日渐丰富, 视觉 SLAM^[1,2] 技术被认为是实现无人机自主作业的核心技术。它的出现使得无人机在无导航信号的环境中, 也能通过搭载的相应传感器获

取自身状态和周围的环境信息, 从而实现地图构建和环境认知, 自主完成特定任务。

无人机常用的自主定位方法为视觉定位法: 通过相机获取环境信息, 经过前后帧的图像在三维空间中的位移得到自身位姿的变化, 现阶段其主要算法为 SLAM 算法, 其中包括特征法和直接

法^[3-6]。与直接法相比, 特征法更具鲁棒性。点特征是最常见的特征类型, 但常用的点特征算法, 如 Shi-Tomasi(一种改进的角点检测算法)算法、FAST^[7](用于特征点检测的算法)算法和 SIFT(尺度不变特征转换)算法^[8]运行时间较长, 不能够满足无人机在快速飞行过程的实时定位; 同时, 在走廊、操场或平原等纹理较差的低空场景中难以检测到足够数量的点特征, 导致无人机处在纹理不佳、光照变化等场景下时难以保持定位稳定性。

针对上述问题, 本文提出一种基于边缘特征的 SLAM 算法。采用 Canny^[9](边缘检测)算法对每帧图像进行边缘特征提取, 减少位置估计过程中的不确定性和模糊性, 同时利用卷积神经网络匹配关键边缘特征, 降低特征匹配过程中的运算量来提高系统的实时性。最终, 利用非线性优化算法求解重投影误差函数完成定位。本文旨在提高 SLAM 系统在低空动态环境下的稳定性和精度。

1 方法框架

本文主要研究图像处理算法和重投影误差算法, 采用非线性优化框架。如图 1 所示的橙色部分和绿色部分。该算法由 Vins-Mono^[10]算法改进而得。本文研究的重点是边缘特征提取、匹配以及重投影误差算法:

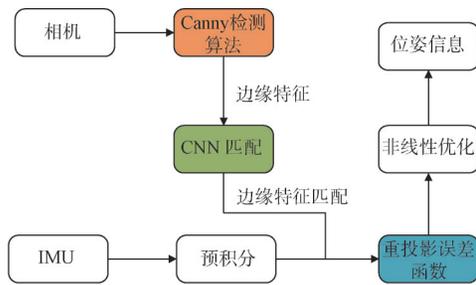


图 1 算法框图
Fig.1 Algorithm block diagram

Canny 算法通过提取稳定的边缘特征, 可以很好地处理平移、旋转和动态视角的情况。其次, 能够对连续帧的边缘特征进行追踪, 为特征匹配增加约束条件。在特征匹配方面, 采用基于深度卷积神经网络的匹配算法来匹配边缘特征, 能够大幅减少匹配所需的时间。在本文中, 使用非线性优化算法求解边缘特征的重投影误差函数完成定位。

2 本文方法

2.1 图像处理

图像处理流程图如图 2 所示。



图 2 图像处理算法
Fig.2 Edge feature processing algorithm

2.1.1 边缘特征提取

考虑像素点及像素点条带周围的空间临近度和灰度相似关系, 对采集到的图像用 Canny 边缘检测算法进行处理。边缘特征提取主要是通过高斯滤波和图像梯度来实现的, 如下式所示。

$$g_{\sigma}(m, n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{m^2+n^2}{2\sigma^2} \cdot f(m, n)} \tag{1}$$

其中 (m, n) 为像素点, 其灰度值(这里只考虑二值图)为 $f(m, n)$ 。

使用 Sobel(用于边缘检测的离散微分算子)算子或其他类似的边缘检测算子计算图像在 x 和 y 方向上的梯度。这些梯度反映图像中强度的变化率, 并且可以用于检测边缘。具体公式如下:

$$G(m, n) = \sqrt{g_x(m, n)^2 + g_y(m, n)^2} \tag{2}$$

$$\theta = \arctan \frac{g_y(m, n)}{g_x(m, n)}$$

图 3(a)原图经过边缘检测后能够清晰地提取出图像中的边缘特征, 极大地压制了噪声(图 3b)。



图 3 原图(a)和 Canny 边缘检测结果(b)

Fig.3 Original image (a) and Canny edge detection results (b)

利用边缘特征的本质目的, 是要找到某像素与其相邻像素的灰度差值, 并放大这种差值, 从而用于图像增强, 本文对比 Canny(罗伯特交叉)算子和 Laplacian(拉普拉斯)算子以及 Robertes cross 三类方法的基本指标, 性能结果如表 1 所示。

表1 3种算子的性能对比

Table 1 Performance comparison of 3 operators

| 方法 | 运行环境 | 耗时/ ms | 算子 数量 | 细节 损失 |
|------------------|----------------|-----------|----------|----------|
| Canny算子 | i7-4720HQ+8 GB | 0.927 | 261 | 小 |
| Laplacian算子 | i7-4720HQ+8 GB | 1.64 | 249 | 大 |
| Robertes cross算子 | i7-4720HQ+8 GB | 5.79 | 284 | 小 |

2.1.2 特征匹配

前后两帧的边缘特征在线性上不是理想状态。因此，为了得到理想的匹配基线，本文基于卷积神经网络技术^[11,12]对无人机图像进行特征匹配，自动学习前后两帧图像的匹配模式，实现无人机图像间的自动匹配流程，得到更为丰富和准确率更高的匹配点对和基线。

基于卷积神经网络从前后两帧成对的图像块中学习一个相似函数(similarity function)，该网络直接将成对的图像叠起来，形成两通道的图像进行训练，称为两通道神经网络。其结构包括一个普通的卷积神经网络，最后再加一个决策层(decision layer)，网络结构如图4所示。

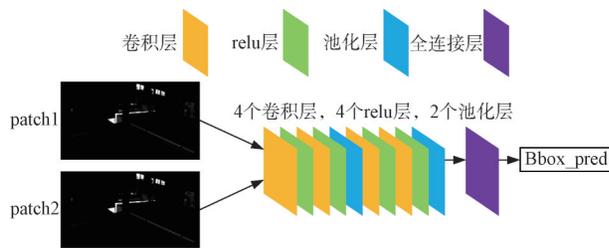


图4 两通道神经网络结构

Fig.4 Two channels neural network structure

经过对数据迭代训练分析，在进行网络迭代训练时，设置迭代次数为250次，训练过程网络的loss曲线如图5所示。

根据图中数据变化趋势分析，曲线在0~50次

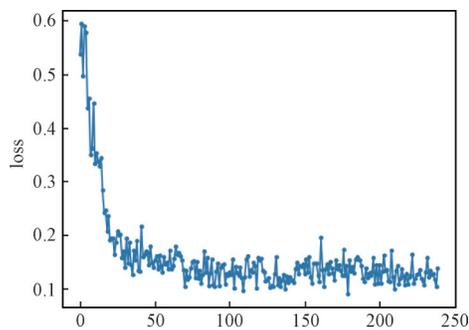


图5 两通道神经网络loss曲线

Fig.5 Two channels neural network loss curves

迭代训练中下降幅度最大，因为该网络结构将单通道问题转换为多通道问题，联合处理两个图像使得训练更快。随着迭代次数增加，loss曲线趋于平稳。同时该网络结构由一系列的卷积层、ReLU(激活函数)层、池化层和全连接层组成，结构简单，减少了两图像特征的匹配时间。并使用双通道图像进行推理，网络有更大的自由度去捕获两图像之间的特征，增加特征匹配的准确度。

两通道神经网络将前后两帧的两个灰度化单通道patch(补丁)拼接成一张两通道的patch，直接送入网络的前项模块，由一系列卷积、ReLU层和max-pooling(最大池化)层组成。这部分的输出作为后项模块的输入，由全连接线性决策层组成，最后输出关键帧的匹配图像，如图6所示。

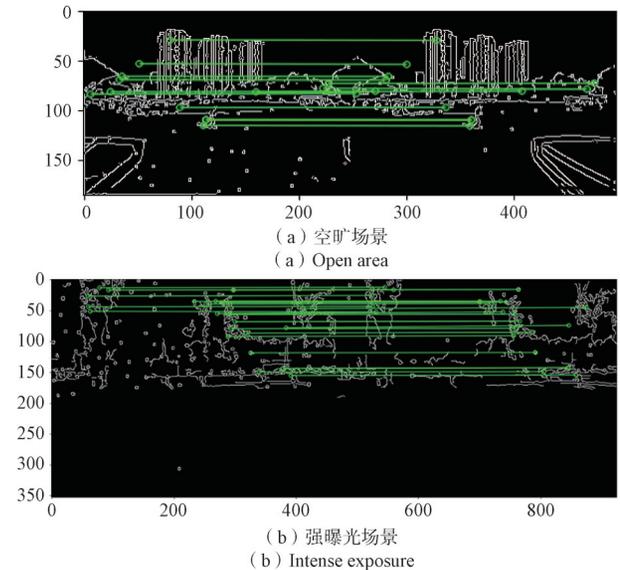


图6 关键帧的两通道神经网络匹配

Fig.6 Two channels neural network matching of keyframes

2.2 位姿估计

相机提取的边缘特征，它的直线边缘被投影到图像中作为曲线轮廓。假设一组 N 条直线，在物体坐标系中由它们的方向向量定义与一组弯曲的图像轮廓线 L_k 进行匹配^[13]，考虑 L_k 上任意一点 M_{k0} ，后一边缘上任意一点 M_{ki} 在目标帧中表示为：

$$M_{ki} = M_{k0} + \sigma_{ki} L_k \quad (3)$$

因此，对于曲线上的每个像素，可以写出如下的投影方程^[14]：

$$sm_{ki} = K [R_\delta R_i T + \delta T_i] (M_{ki} + \sigma_{ki} L_k) \quad (4)$$

这意味着轮廓的每个像素都会产生一对约束：

$$u_{ki} = au \frac{R_{1i}(M_{ki} + \sigma_{ki}L_k) + T_{xi}}{R_{3i}(M_{ki} + \sigma_{ki}L_k) + T_{zi}} + u_0$$

$$v_{ki} = au \frac{R_{2i}(M_{ki} + \sigma_{ki}L_k) + T_{yi}}{R_{3i}(M_{ki} + \sigma_{ki}L_k) + T_{zi}} + v_0$$
(5)

很明显, 将三维直线边缘与图像曲线匹配并不能得知每个轮廓像素对应的三维边缘点是什么。尤其是 σ_{ki} 的值是未知的。因此, 式(5)可以表示为:

$$u_{ki} \triangleq x_{uki}(R, T, \Omega, a, V, S) \quad (6)$$

其中, Ω 是所有参数 σ_{ki} 的向量。对 n 条直线边缘匹配的图像曲线每条观测 m 个像素 $[\hat{u}_{ki}, \hat{v}_{ki}]$, 并将其与式(4)的理论投影进行比较, 可以得到一个表示重投影误差的 $n \times m$ 方程:

$$\varepsilon_{ki} = \begin{bmatrix} \hat{u}_{ki} - \zeta_{uki}(R, T, \Omega, a, V, S) \\ \hat{v}_{ki} - \zeta_{vki}(R, T, \Omega, a, V, S) \end{bmatrix} \quad (7)$$

对于位姿参数和速度参数 R, T, Ω, a, V 以及对应的变量 S , 可以表示为最小二乘意义上的代价函数:

$$\varepsilon = \sum_{k=1}^n \sum_{i=1}^m [\hat{u}_{ki} - \zeta_{uki}(R, T, \Omega, a, V, S)]^2 + [\hat{v}_{ki} - \zeta_{vki}(R, T, \Omega, a, V, S)]^2 \quad (8)$$

该代价函数使用 Levenberg Marquardt 算法^[15]求解。

3 实验结果与分析

实验在妙算主机上实现。具体配置如下: CPU 为 Intel®Core™i7-8559U 处理器, 主频为 2.7 GHz~4.5 GHz, 内存为 8 G, 无 GPU, 系统版本为 Ubuntu 18.04 LTS 64 位。实验采用公共数据集 EuRoC^[16]对算法进行评估, 数据集包含快速运动及昏暗场景等复杂条件下的传感器数据, 能够比较完整地模拟无人机飞行环境, 可以全面评估算法的性能。

为了验证本文算法的实时性, 选取 EuRoC 数据集中 6 个具有代表性的序列进行实验, 用每帧图像的平均处理时间来评估本文算法与最新的 ORB-SLAM3 算法的图像处理速度, 对比结果如图 7 所示。

从图 7 可以看出, 本文算法在 6 个数据集中的图像处理耗时相较于 ORB-SLAM3 降低了 31%, 相比 ORB-SLAM3 在实时性上有明显提升。在序列 V1_01 中, ORB-SLAM3 耗时略小于本文算法, 这是由于图像特征清晰, 因此在画面平稳的状态下

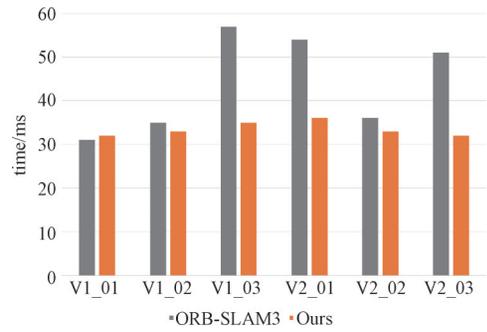


图 7 本文算法与 ORB-SLAM3 的每帧图像平均处理时间
Fig.7 The average processing time per frame of the algorithm in this paper compared with ORB-SLAM3

图像处理的时间差距不明显; 在 V1_03 和 V2_03 这一类光照变化明显、图像纹理较少的数据集中, 本文算法相较于 ORB-SLAM3 耗时更少。

绝对轨迹误差 (Absolute Trajectory Error, ATE) 是通过描述位姿估计和真实位姿之间的差异, 直观表达算法的全局性和精度。ORB-SLAM3 和本文算法的绝对轨迹误差对比结果如表 2 所示。

表 2 ORB-SLAM3 和本文算法的绝对轨迹误差对比
Table 2 Comparison of absolute trajectory errors between ORB-SLAM3 and the proposed method

| Serial | 单位:m | |
|-----------------|-----------|-------|
| | ORB-SLAM3 | Ours |
| MH_01_easy | 0.121 | 0.145 |
| MH_02_easy | 0.169 | 0.150 |
| MH_03_medium | 0.235 | 0.218 |
| MH_04_difficult | 0.358 | 0.224 |
| MH_05_difficult | 0.263 | 0.257 |
| V1_01_easy | 0.125 | 0.085 |
| V1_03_difficult | 0.185 | 0.148 |
| V2_01_easy | 0.138 | 0.128 |
| V2_02_medium | 0.162 | 0.150 |
| V2_03_difficult | 0.196 | 0.183 |

从上表中绝对轨迹误差可以得出: 针对纹理缺失、快速运动的场景, 本文算法均表现出优异的精确性, 平均定位精度相较于 ORB-SLAM3 提高 15.04%。在 MH_01_easy 数据集中也是由于图像清晰稳定的原因导致本文算法误差略高于 ORB-SLAM3 算法。

飞行实验。实验场地为西安某公园, 相对空旷, 周围建筑较少, 特征点提取困难, 较大的光照强度和快速飞行对算法稳定性的要求更为严格。为了验证本文算法与 ORB-SLAM3 算法在低空高速

环境中的性能,飞机飞行高度设置为10 m,飞行速度为15 m/s。四旋翼无人机配备了微型主机和D453i相机(相机帧速率30 Hz,IMU(惯性测量单元)帧速率200 Hz),实验采集设备如图8所示。



图8 实验采集设备

Fig.8 Experimental collection equipment

实验方案:如图9所示。分别针对空旷场景和强曝光场景下两种算法的轨迹精度,进行两次飞行实验。飞机上配备差分RTK(实时动态测量),误差精度在厘米级,可作为飞机真实轨迹,启动定位算法后,控制无人机飞行按照既定航迹飞行。同时实时定位和绘制其轨迹,最终比较两种算法在飞行轨迹上与真实轨迹的差别。

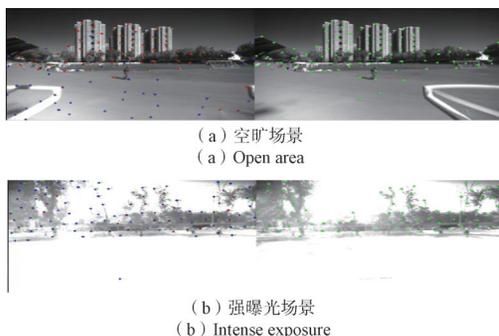


图9 典型飞行实验场景

Fig.9 Typical Flight Experiment Scenarios

两种算法的实际运行轨迹如图10所示,ORB-SLAM与本文算法在飞行过程中的实际指标如表3所示。

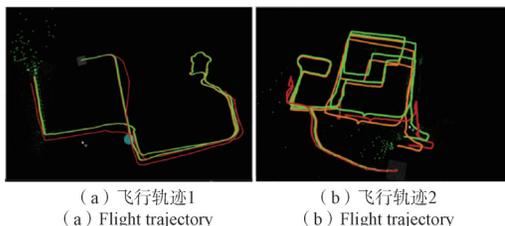


图10 本文算法与ORB-SLAM3的飞行定位轨迹

Fig.10 This article compares the algorithm with the flight positioning trajectories of ORB-SLAM3

表3 2种算法的飞行实际指标

| 算法 | 匹配特征数量 | 匹配平均耗时/ms | 绝对误差/m |
|-----------|--------|-----------|--------|
| 原始轨迹 | 0 | 0 | 0 |
| 本文算法 | 150 | 48.21 | 0.47 |
| ORB-SLAM3 | 150 | 62.93 | 3.851 |

图9中红色轨迹为ORB-SLAM3算法定位结果,绿色轨迹为本文算法定位结果,橙色轨迹为实际飞行轨迹。从图9(a)可知,在区域空旷、特征纹理较少的场景下,本文算法定位精度更高且基本与实际飞行轨迹吻合。从图9(b)可以看出,ORB-SLAM3算法在飞行不久后发散,这是由于光照变换较大无法实时通过图像处理获取有效特征点,导致定位失效。

在表3可以得到,实际飞行过程中特征数量设置为150时,本文算法平均匹配时间为ORB-SLAM3算法的76.6%;在过度曝光和无人机高速飞行下能够全程保持定位稳定,定位精度为ORB-SLAM3算法的8倍。

综上,通过以上实验对比可以发现,本文算法降低了图像算法运算时间,在快速运动及过度曝光等复杂条件下均保持良好的位姿估计精度与鲁棒性。

4 结束语

本文提出了一种基于边缘特征的无人机视觉惯性定位方法,为有效降低边缘特征提取难度,在图像特征提取算法中采取Canny边缘检测算法;针对长时间运行时的定位耗时长的问题,通过卷积神经网络的匹配算法,降低图像处理耗时,显著提高了定位精度;最后对边缘特征构建特征重投影误差函数,提供了更精确的位姿估计。实验结果表明,本文算法在动态环境下定位精度更高,相较于传统算法实时性更好,定位轨迹更平滑。未来的工作将进一步研究提升卷积神经网络对图像数据处理的实时性,推动视觉定位算法在无人机平台的轻量化部署。

参考文献

- [1] CHEN W, WANG X, GAO S, et al. Overview of multi-robot collaborative SLAM from the perspective of data fusion[J]. Machines, 2023, 11(6): 653.
- [2] 邹斌,林思阳,尹智帅.基于YOLOv3和视觉SLAM的

- 语义地图构建[J]. 激光与光电子学进展, 2020, 57(20): 116-122.
- ZOU B, LIN S Y, YIN Z S. Semantic map based on YOLOv3 and visual SLAM[J]. Laser & Optoelectronics Progress, 2020, 57(20): 116-122.
- [3] JULIER S J, UHLMANN J K. Unscented filtering and nonlinear estimation[J]. Proceedings of the IEEE, 2004, 92(3): 401-422
- [4] QIN T, LI P, SHEN S. Vins-mono: A robust and versatile monocular visual-inertial state estimator[J]. IEEE Transactions on Robotics, 2018, 34(4): 1004-1020.
- [5] MUR ARTAL R, TARDÓS J D. Orb-slam2: An open-source slam system for monocular, stereo, and RGB-D CA-meras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.
- [6] 林晓华,王茂森,戴劲松. 基于 ROS 的无人机占据地图构建方案设计[J]. 兵器装备工程学报, 2021, 42(5): 215-220.
- LIN Xiaohua, WANG Maosen, DAI Jinsong. Design of UAV occupied map construction scheme based on ROS [J]. Journal of Ordnance Equipment Engineering, 2021, 42(5): 215-220.
- [7] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[C]//Computer Vision-ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9. Berlin Heidelberg, Springer: 2006: 430-443.
- [8] ZHOU H, YUAN Y, SHI C. Object tracking using SIFT features and mean shift[J]. Computer Vision and Image Understanding, 2009, 113(3): 345-352.
- [9] 段锁林, 殷聪聪, 李大伟. 改进的自适应 Canny 边缘检测算法[J]. 计算机工程与设计, 2018, 39(6): 1645-1652.
- DUAN Suolin, YIN Congcong, LI Dawei. Improved adaptive Canny edge detection algorithm[J]. Computer Engineering and Design, 2018, 39(6): 1645-1652.
- [10] 王瑞, 王茂森, 戴劲松, 等. 基于改进视觉惯性里程计的无人机系统设计[J]. 兵器装备工程学报, 2023, 44(4): 275-281.
- WANG Rui, WANG Maosen, DAI Jinsong, et al. UAV system design based on an improved visual inertial odometry[J]. Journal of Ordnance Equipment Engineering, 2023, 44(4): 275-281.
- [11] 张晏涵, 张尹馨, 黄战华, 等. 基于改进 DFM 的密集特征匹配算法[J]. 激光与光电子学进展, 2024, 61(8): 0815001.
- ZHANG Yanhan, ZHANG Yinxin, HUANG Zhanhua, et al. Dense feature matching based on improved DFM algorithm[J]. Laser & Optoelectronics Progress, 2024, 61(8): 0815001.
- [12] 王志明, 张航. 融合多层卷积神经网络特征的快速图像检索方法[J]. 计算机辅助设计与图形学学报, 2019, 31(8): 1410-1416.
- WANG Zhiming, ZHANG Hang. A Fast image retrieval method integrating multi-layer convolutional neural network features[J]. Journal of Computer Aided Design and Graphics, 2019, 31(8): 1410-1416.
- [13] 宁小娟, 李洁茹, 高凡, 等. 基于最佳几何约束和 RANSAC 的特征匹配算法[J]. 系统仿真学报, 2022, 34(4): 727-734.
- NING Xiaojuan, LI Jieru, GAO Fan, et al. Feature matching algorithm based on optimal geometric constraints and RANSAC[J]. Journal of System Simulation, 2022, 34(4): 727-734.
- [14] SIE N J, SRIGRAROM S, HUANG S N. Field test validations of vision-based multi-camera multi-drone tracking and 3d localizing with concurrent camera pose estimation[C]//2021 6th International Conference on Control and Robotics Engineering (ICCRE), Beijing, China, 2021: 139-144.
- [15] FAN J, PAN J. On the convergence rate of the inexact Levenberg-Marquardt method[J]. Journal of Industrial & Management Optimization, 2017, 7(1): 199-210. DOI: 10.1137/0111030.
- [16] BURRI M, NIKOLIC J, GOHL P, et al. The EuRoC micro aerial vehicle datasets[J]. The International Journal of Robotics Research, 2016, 35(10): 1157-1163.

[作者简介]

刘春江 1993年生, 硕士, 工程师。

张鹏宇 1996年生, 硕士, 助理工程师。

(本文编辑: 潘三英)

(英文编辑: 赵尹默)