

融合语义分割的葡萄果园机器人稠密地图构建方法

冯桑 张禧龙 杨润彬 陈彦阳 黄晓涛

(广东工业大学机电工程学院, 广州 510006)

摘要: 针对葡萄果园作业机器人存在的定位精度低、果实识别不准确及地图质量不理想等问题,提出一种融合语义分割的稠密建图算法 PDS-SLAM。基于 ORB-SLAM3 框架,通过改进特征点提取策略,结合角点数量自适应调节 FAST 阈值,并改进二叉树算法,提高特征点分布均匀性,提升定位精度;在 PIDNet 基础上融合 DSA 模块提出 PDSNet,改善对果实的空间感知能力,提高果实识别效果;引入稠密建图线程与二叉树线程,通过点云恢复算法得到局部点云,利用统计离群点滤波与半径滤波优化局部点云,并结合语义掩膜对葡萄点云进行语义标注生成语义地图,最后转换为二叉树地图。在 EuRoC 数据集实验中,PDS-SLAM 绝对轨迹误差比 ORB-SLAM3 降低 27.3%,ORB 特征点匹配数量平均提升 15.5%;在自建数据集上,PDSNet 在速度 126.92 f/s 下 IoU 达到 78.9%。研究结果表明,PDS-SLAM 可提升果园机器人定位和感知能力,为果园机器人导航与作业提供支持。

关键词: ORB-SLAM3; 葡萄果园机器人; 语义分割; 语义地图

中图分类号: TP242.6; S24 文献标识码: A 文章编号: 1000-1298(2026)06-0036-09

OSID:



Dense Mapping Method for Grape Orchard Robots Integrating Semantic Segmentation

FENG Sang ZHANG Xilong YANG Runbin CHEN Yanyang HUANG Xiaotao

(School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Aiming to address low localization accuracy, unreliable fruit recognition, and poor map quality in vineyard robots, PDS-SLAM, a dense mapping algorithm that integrated semantic segmentation was proposed. Built on ORB-SLAM3, each image was partitioned during feature extraction; the regional FAST threshold was adaptively adjusted according to regional corner counts; and quadtree uniformization method with minimum distance was applied, which improved spatial uniformity and matching robustness of feature points, thereby enhancing localization accuracy. A network, PDSNet, was proposed by integrating a DSA module into PIDNet, enhancing spatial perception of grape clusters and improving fruit recognition. A dense mapping thread and an octree thread were introduced; images were projected to recover local dense point clouds via a point cloud recovery algorithm; statistical outlier filter and radius filter were applied to remove aberrant points; semantic masks were used to annotate grape clusters, yielding a dense semantic map that was finally converted into an octomap. In experiments on the EuRoC dataset and a self-collected dataset, a 27.3% reduction in absolute trajectory error (ATE) on the MH03 sequence relative to ORB-SLAM3 and a 15.5% average increase in matched ORB features were achieved, indicating improved localization accuracy. PDSNet achieved an IoU of 78.9% for grape segmentation at 126.92 f/s. The results demonstrated that PDS-SLAM enhanced localization perception and produced dense semantic maps and octree maps, supporting autonomous navigation and precision operations for orchard robots.

Key words: ORB-SLAM3; grape orchard robot; semantic segmentation; semantic map

0 引言

果园非结构化环境下的智能化管理依赖于同步

定位与建图 (Simultaneous localization and mapping, SLAM) 技术提供的实时定位与环境理解能力^[1]。

在产量估算方面,LIU 等^[2]开发了视觉-激光雷达融

合系统 ORB - Livox,用于果实的实时动态检测与定位;YUAN 等^[3]通过改进 VINS - RGBD 算法,在草莓园中实现了高精度的定位与建图以支撑产量评估。在自主采摘任务中,视觉系统是关键^[4]。XIONG 等^[5]则构建三维语义地图以提供环境先验信息。针对精准喷药需求,相关研究利用 LiDAR SLAM 结合 IMU 或改进 ORB - SLAM2 算法,为喷药机器人构建了全自主导航系统及稠密环境地图,以实现果园环境下的精准作业^[6-8]。

为提升 SLAM 系统在果园复杂环境下的性能,研究者从前端特征提取与多源信息融合等方面进行了优化。LI 等^[9]将点线特征引入前端匹配,缓解了低纹理对追踪的干扰;沈跃等^[10]提出激光雷达与 IMU 紧耦合算法,通过在线偏置估计保证了定位全局一致性;HONG 等^[11]进一步利用改进 VGICP 与去畸变算法减少了长时间运行的误差积累。针对动态环境与光照变化,薛金林等^[12]结合目标检测剔除动态干扰并利用平面特征优化匹配;姜路等^[13]引入自适应阈值策略提升了特征提取稳定性。然而,上述研究主要集中于几何层面,生成的地图缺乏语义信息,且在纹理稀疏果园中仍面临定位瓶颈。

近年来,深度学习开始被引入 SLAM 框架以构建语义地图。WANG 等^[14]将 SCTNet - B 嵌入 ORB - SLAM2,生成了辅助导航的语义八叉树地图;张鹏飞等^[15]利用贝叶斯滤波融合多帧语义预测,降低了单帧分割误差;林义忠等^[16]融合 YOLO v7 - mask 对关键帧进行分割,赋予了地图语义信息;LV 等^[17]集成 Mask R - CNN 以剔除农业场景中的动态对象干扰。尽管深度学习增强了语义感知能力^[18],但高性能网络计算量大,难以兼顾实时性与精度,且现有方法普遍缺乏对点云的后处理,导致地图中噪点较多,难以直接用于作业。

针对果园 SLAM 面临的定位精度低、感知与实时性矛盾及地图质量差等问题,本文提出一种融合语义分割的稠密建图方法 PDS - SLAM。基于 ORB - SLAM3 框架,通过改进自适应 FAST 阈值与四叉树算法优化特征分布,提升定位精度;提出轻量化 PDSNet 兼顾分割精度与速度;并结合点云滤波构建高质量语义八叉树地图,以满足实时导航需求。

1 实验平台

选用 R550 Plus 型阿克曼机器人(轮趣科技(东莞)有限公司)作为果园机器人移动平台,如图 1 所示,外观尺寸为 0.776 m × 0.671 m × 0.319 m,质量为 35.6 kg,额定载荷能力为 45 kg,最大运行速度可

达 1.65 m/s,具备 25°爬坡能力,具有较强的地形适应性与灵活的机动性能,适用于果园环境下移动作业任务。



图 1 移动机器人实验平台

Fig. 1 Mobile robot experimental platform

系统采用 ZED2 型双目立体视觉相机(Stereolabs 公司),集成双广角镜头、RGB 摄像头、六轴 IMU 及气压计,可实现 0.2 ~ 20 m 的高精度深度感知能力,支持最高分辨率 2 208 像素 × 1 242 像素与帧率 100 f/s。

数据处理平台采用联想 Y9000 系列便携式计算机,配备 i5 - 12500H 处理器、16 GB 内存以及 8 GB 显存的 RTX3060 独立显卡,运行 Ubuntu 20.04 操作系统,用于完成传感器数据采集、语义分割、点云处理与目标定位等关键任务。

2 数据采集

实验场景位于广东省广州市某葡萄园,葡萄种植密度高、枝叶与果实遮挡严重,符合果园机器人在葡萄果实检测与定位建图任务中所面临的作业环境,如图 2 所示。使用移动机器人实验平台对葡萄园区域进行连续图像采集,相机分辨率为 640 像素 × 360 像素,帧率为 30 f/s。从原始图像序列中随机选取 940 幅,采用 LabelMe 工具进行葡萄类别标注。将数据集按比例 8:1:1 划分为训练集、验证集和测试集,用于语义分割算法训练与性能评估。



图 2 实验场景

Fig. 2 Experimental scenarios

3 PDS - SLAM 框架

果园场景中存在低纹理区域,且受光照变化影响,使得固定 FAST 阈值角点提取策略无法提取足量特征点,且会出现分布不均现象,导致定位精度下降。传统的 ORB - SLAM3 的稀疏地图由离散的三维特征点与关键帧位姿构成,缺少对环境几何特性和语义信息的表达,仅能满足果园机器人的定位需

求,缺少对导航和作业的支持。

融合语义分割的建图算法 PDS-SLAM 结构如图 3 所示,在 ORB-SLAM3 基础上,在追踪线程中加入分区自适应 FAST 阈值机制,通过分区提取初始 FAST 角点,对不满足数量要求的分区选用降阈

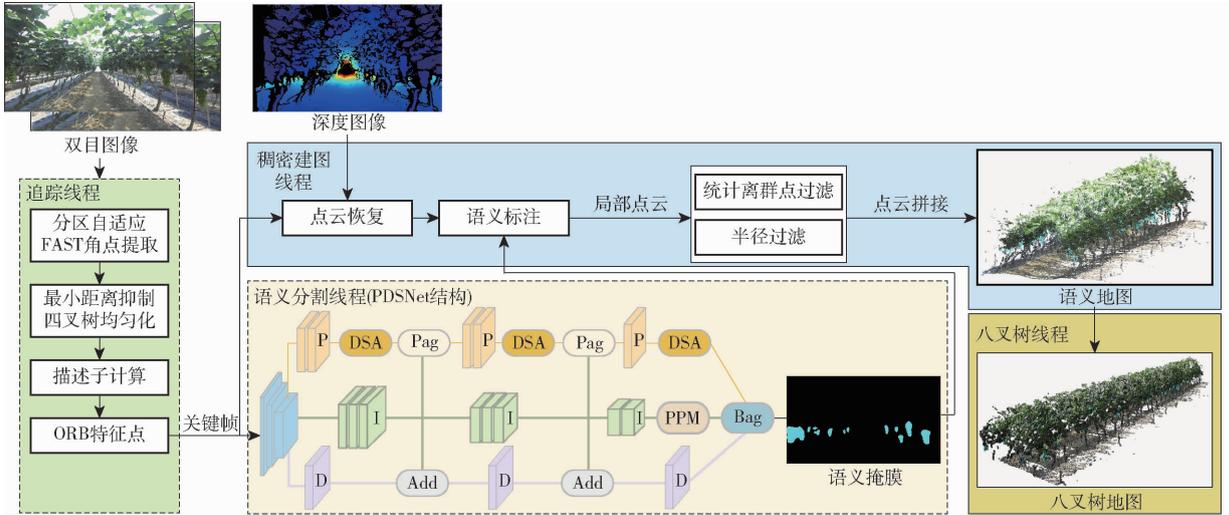


图 3 PDS-SLAM 算法结构

Fig. 3 Structure of PDS-SLAM algorithm

在 PIDNet 基础上提出 PDSNet,引入 DSA 模块,在保证实时性前提下,提高了算法空间感知能力,在对关键帧进行语义分割后,将语义掩膜输入稠密建图线程,用于构建语义地图。引入稠密建图线程,通过点云恢复算法,将二维像素点反投影到三维空间中,根据二维像素点在语义掩膜中的标签,对葡萄果实进行语义标注,最后通过八叉树线程将稠密点云地图转换成八叉树地图,用于果园机器人导航和作业。

3.1 分区角点提取和均匀化策略改进

ORB 特征点^[19]由 FAST 角点^[20]与旋转 BRIEF 描述子组成,通常在多尺度图像金字塔上以固定的 FAST 阈值进行角点检测,然后估计主方向并生成二进制描述,具备较高的实时性与旋转不变性^[21]。但在葡萄园场景中,叶片与果实常呈均匀或重复纹理,且光照不均会使局部对比度降低,固定的 FAST 阈值难以有效提取角点,进而造成特征点空间分布不均并降低匹配稳定性,增加误匹配与跟踪漂移风险,最终影响位姿估计与三维重建质量。

为解决该问题,提出一种分区自适应 FAST 阈值的角点提取算法,对角点不足的区域调整 FAST 阈值提取足够数量的角点,保证特征点数量,从而保障特征匹配和追踪质量,该部分算法基本流程如图 4 所示。

首先将第 l 层金字塔图像 I_l ,划分为 r 行 c 列个网格单元,在每个网格 $C_{r,c}$ 内,以初始 FAST 阈值 τ_{ini}

序列的 FAST 阈值提取角点,使得每个分区都满足角点提取要求,提高了低纹理区域的特征点数量;同时,在原四叉树均匀化策略中引入最小距离抑制机制,可有效防止局部特征点聚集,提高特征点均匀性和匹配数量,提高算法定位精度。



图 4 特征点提取流程

Fig. 4 Feature point extraction workflow

进行角点检测,得到初次检测角点候选集 $K_{r,c}^{(1)}$,并提取该网格内检测到的特征点数量 $n_{r,c}^{(1)}$,过程为

$$K_{r,c}^{(1)} = \text{FAST}(C_{r,c}, \tau_{ini}) \quad (1)$$

$$n_{r,c}^{(1)} = \text{num}(K_{r,c}^{(1)}) \quad (2)$$

检查网格内特征点数量,当 $n_{r,c}^{(1)} \geq k_{\text{target}}$ 时,直接保留响应值前 k_{max} 个特征点;当 $n_{r,c}^{(1)} < k_{\text{target}}$ 时,对该网格采用 CLANE 增强局部对比度。然后采用降阈序列的 FAST 阈值 τ_t 进行角点提取,得到第 t 次角点提取的结果 $K_{r,c}^{(t)}$ 。每次补点后,若 $\text{num}(K_{r,c}^{(t)}) \geq k_{\text{target}}$,则停止补点。设定目标每格特征点数 k_{target} 为 4,最大每格特征点数为 6,具体过程为

$$\begin{cases} K_{r,c} = \text{TopK}(K_{r,c}^{(0)}, k_{\text{max}}) & (n_{r,c}^{(0)} \geq k_{\text{target}}) \\ K_{r,c}^{(t)} = \text{FAST}(C_{r,c}, \tau_t) & (n_{r,c}^{(0)} < k_{\text{target}}) \end{cases} \quad (3)$$

其中

$$\tau_t = \tau_{\text{max}} - \frac{t}{T}(\tau_{\text{max}} - \tau_{\text{min}}) \quad (4)$$

$(t = 1, 2, \dots, T)$

式中 τ_t ——第 t 次迭代时使用的 FAST 检测阈值

$\tau_{\text{max}}, \tau_{\text{min}}$ ——最大、最小 FAST 阈值

T ——降阈序列值个数,取 3

在 FAST 角点提取完成后,利用四叉树算法实现均匀化分布。然而,传统四叉树仅依赖空间划分,

易在局部区域保留过于接近的点,导致分布不均。因此,在二叉树的基础上引入最小距离抑制机制,通过设定尺度自适应的最小距离间隔,有效剔除相邻角点,从而在保证全局均匀性的同时进一步提升局部分布合理性,具体距离计算式为

$$d_l = \max(d_{\min}, d_0 s_l^{-1}) \quad (5)$$

式中 d_0 ——初始抑制距离

d_{\min} ——最小抑制距离

s_l ——从原图到第 l 层的缩放因子

3.2 语义分割算法

PIDNet^[22]是一种面向实时语义分割的多分支网络,借鉴PID控制算法,设计3个不同的网络分支,在保持较高推理效率的同时兼顾全局语义与局部细节,通过分工协作的思想实现精度与速度的平衡。其中,P分支侧重保留高分辨率的细节与边界信息,为主干网络提供几何信息特征,是缓解小目标欠分割与边界模糊的关键。

然而葡萄果园中存在大量遮挡、形状不规则的现象,使得固定采样卷积难以精准感知目标的结构和空间。针对这个问题,本文提出PDSNet,在PIDNet的P分支中引入DSA(Deformable spatial attention)模块^[23],PDSNet结构如图3所示。

DSA模块融合了单轴可变形卷积(DSCN)结合空间注意力机制,通过单轴可变形卷积和注意力机制,有效提取不规则目标区域的特征信息,提升空间分支对各向异性结构的特征感知能力,DSA模块如图5所示。

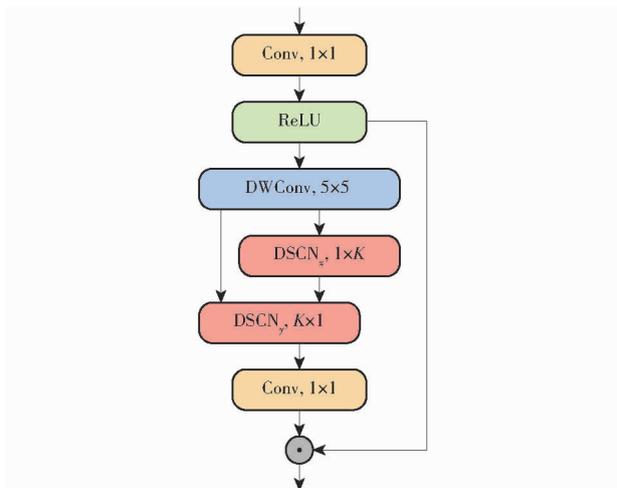


图5 DSA模块

Fig.5 DSA module

首先,输入特征图 x 经过 1 个 1×1 卷积进行通道变换和 ReLU 激活函数,过程为

$$x' = \text{ReLU}(\text{Conv}_{1 \times 1}(x)) \quad (6)$$

式中 $\text{Conv}_{1 \times 1}$ —— 1×1 卷积

x' ——输出特征图, $x \in \mathbf{R}^{C \times H \times W}$

随后,通过 2 个分别沿水平方向与垂直方向的双轴可变形卷积(DSCN-X与DSCN-Y),并设计残差路径以解决单轴依赖带来的特征丢失问题。

通过 5×5 的深度卷积提取局部空间上下文,计算式为

$$x'_c = \text{DWConv}_{5 \times 5}(x') \quad (7)$$

式中 x'_c ——局部空间上下文特征图

单轴可变形卷积计算式为

$$x_{\text{attn-x}} = \text{DSCN}_x(x'_c) \quad (8)$$

$$x_{\text{attn-y}} = \text{DSCN}_y(x_{\text{attn-x}}, x'_c) \quad (9)$$

式中 $x_{\text{attn-x}}$ ——水平方向注意力特征图

$x_{\text{attn-y}}$ ——垂直方向注意力特征图

最后通过 1×1 卷积压缩并生成空间注意力图为

$$a_{\text{atten}} = \text{Conv}_{1 \times 1}(x_{\text{attn-y}}) \quad (10)$$

将该空间注意力图应用于中间特征表示 x' , 完成注意力增强得到模块的输出特征图 x_{output} , 即

$$x_{\text{output}} = x' \odot a_{\text{atten}} \quad (11)$$

式中 \odot ——逐元素乘法

3.3 稠密地图和语义地图构建

在ORB-SLAM3中引入建图线程,构建稠密地图,满足果园机器人导航与作业需求。该线程利用优质关键帧,首先通过深度图像过滤无效深度像素点,再通过点云恢复算法,将像素坐标系中的有效像素点 $d(u, v)$ 反投影到相机坐标系中的三维点得到局部点云。局部点云坐标计算过程为

$$\begin{cases} z_c = d \\ y_c = \frac{v - c_y}{f_y} z_c \\ x_c = \frac{u - c_x}{f_x} z_c \end{cases} \quad (12)$$

式中 (x_c, y_c, z_c) ——相机坐标系下点云坐标

(c_x, c_y) ——相机主点坐标

f_x, f_y ——相机焦距

d ——该像素对应深度

局部点云根据相机位姿进行坐标变换,将其从相机坐标系转移到世界坐标系,拼接得到全局点云地图,计算过程为

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \mathbf{R}_{wc} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} + \mathbf{t}_{wc} \quad (13)$$

式中 \mathbf{R}_{wc} —— 3×3 旋转矩阵

\mathbf{t}_{wc} —— 3×1 平移向量

(x_w, y_w, z_w) ——世界坐标系下点云坐标

为使生成的点云地图不仅具有几何结构信息,还能够融合语义信息,在点云上色阶段,从语义分割掩膜中获取该像素坐标对应的语义标签。根据语义

标签判断该像素是否属于葡萄类别,是则赋予该点云青色,否则保留原像素颜色,过程为

$$C(u, v) = \begin{cases} (0, 255, 255) & (\text{label}(u, v) = \text{"grape"}) \\ I_{RGB}(u, v) & (\text{label}(u, v) = \text{"other"}) \end{cases} \quad (14)$$

式中 $I_{RGB}(u, v)$ ——该像素对应的原始颜色信息
稠密地图和语义地图可提供几何结构信息和语义信息,但地图占用内存较大,无法满足果园机器人实时导航和作业需要。为了优化内存使用并提高计算效率,系统引入八叉树线程,将稠密点云地图转换为八叉树地图。

八叉树地图是一种基于三维栅格的空间数据结构,通过递归划分三维空间,能够在保持地图结构完整性的同时,降低地图的内存占用^[24-25]。八叉树地图可以通过调整体素的大小来控制地图的分辨率,根据实际需要动态调整存储的精度,平衡地图细节和效率,适用于果园机器人导航和作业需求。

3.4 地图滤波优化

葡萄植株间距较密、枝叶结构复杂,且受自然光照与风力扰动影响下,稠密地图中的植株边缘及支架附近出现的离群点和空间噪声点云,为解决上述问题,引入统计离群点滤波。

通过统计点云的邻域点,检测与周围点云密度不符的异常点,对于每个点 p_i ,计算每个点 p_j 与其 k 个邻域点平均距离 μ_i ,即

$$\mu_i = \frac{1}{k} \sum_{j=1}^k \|p_i - p_j\| \quad (15)$$

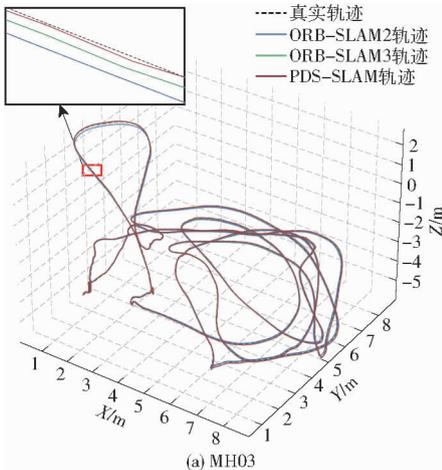
当某个具体点平均距离不符合整体点云均值 $\bar{\mu}$ 加上 λ 倍标准差时,该点被视为离群点,计算式为

$$\mu_i > \bar{\mu} + \lambda\sigma \quad (16)$$

式中 $\bar{\mu}$ ——全局点云平均距离

σ ——整体标准差 λ ——容忍倍数

在剔除统计离群点后,为进一步移除空间噪声



点云,还引入半径离群点滤波。统计每个点在固定半径 r 范围内的邻点数量,若该数量小于最小邻点数量 N_{\min} ,则认为该点为离群点,即

$$N_r(p_i) < N_{\min} \quad (17)$$

式中 r ——搜索半径

N_{\min} ——最小邻点数量

$N_r(p_i)$ ——点 p_i 在半径 r 内邻点数量

3.5 评价指标

为了验证所提算法综合性能,从定位精度、语义分割效果及地图存储成本3个维度选取评价指标:
①定位精度:采用绝对轨迹误差(Absolute trajectory error, ATE)衡量估计轨迹与真实轨迹的差异,并计算其最大值、平均误差(ME)、误差中值、最小值、均方根误差(RMSE)、方差(Var)及标准差(SD)。
②语义分割性能:选取交并比(IoU)、像素准确率(Pa)、F1值评估分割精度,并使用浮点运算量、参数量及帧率评估算法实时性与复杂度。
③地图存储成本:统计不同地图构建方式下的内存占用量(MB),以评估八叉树地图的存储优化效果。

4 结果与分析

4.1 基于EuRoC数据集实验

EuRoC数据集^[26]包含多个用于评估视觉与惯性传感器融合算法的标准测试序列。使用EuRoC数据集中MH03和MH04序列评估所提算法的定位精度。

采用EVO工具进行误差评估。绝对轨迹误差(ATE)衡量估计轨迹与地面真值轨迹之间的差异,通过计算每一时刻估计位置与真实位置距离评估误差累积,即

$$ATE_i = \|P_i - \hat{P}_i\| \quad (18)$$

式中 P_i ——真实轨迹 \hat{P}_i ——估计轨迹

图6为ORB-SLAM2、ORB-SLAM3和PDS-SLAM

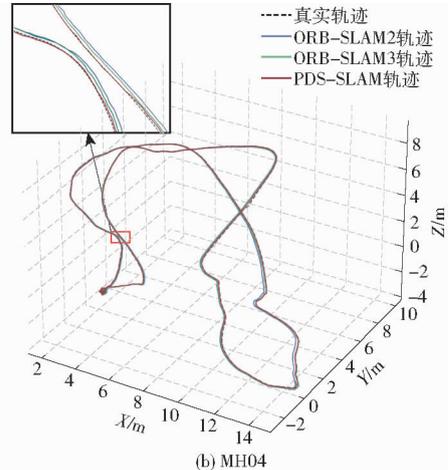


图6 MH03和MH04序列下轨迹对比

Fig. 6 Trajectory comparison on MH03 and MH04 sequence

SLAM(本文算法)分别在 MH03 和 MH04 序列下的轨迹对比。从图 6 可见,PDS-SLAM 的估计轨迹相较 ORB-SLAM2 和 ORB-SLAM3 更接近真实轨

迹。图 7 为 2 个序列下绝对轨迹误差分布,从图 7 可以看出,PDS-SLAM 绝对轨迹误差优于 ORB-SLAM2 和 ORB-SLAM3。

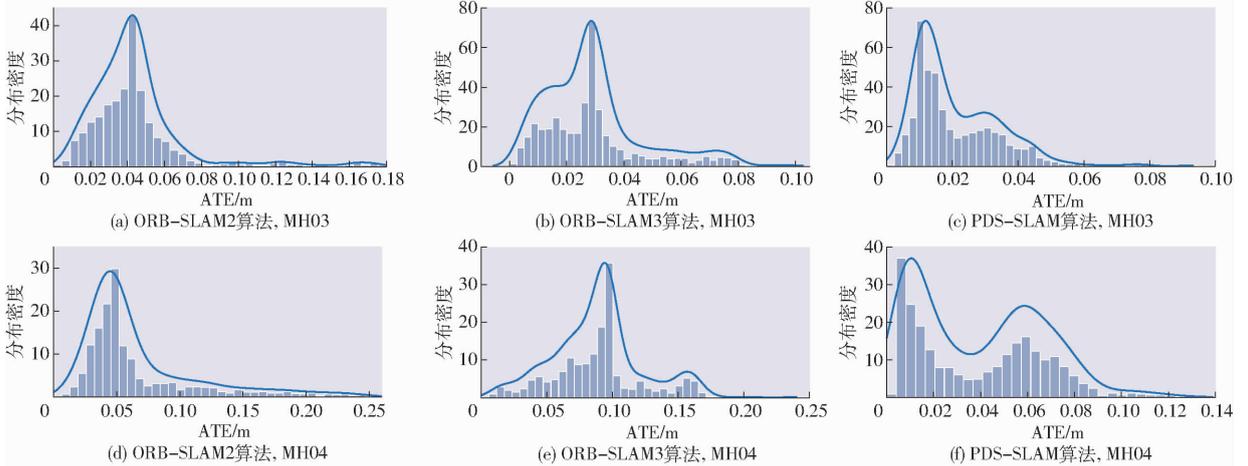


图 7 MH03 和 MH04 序列下绝对轨迹误差曲线

Fig. 7 Absolute trajectory error curves on MH03 and MH04 sequence

表 1 为 MH03 和 MH04 序列下绝对轨迹误差。从表 1 可以看出,PDS-SLAM(本文算法)相比于 ORB-SLAM2 和 ORB-SLAM3,平均误差降低 45.7% 和 32.4%, RMSE 平均降低 40.9% 和 27.3%。综上,可以得出 PDS-SLAM 定位精度优于 ORB-SLAM3。

表 1 MH03 与 MH04 序列绝对轨迹误差

Tab. 1 ATE on MH03 and MH04 sequence

参数	MH03 序列			MH04 序列		
	ORB-SLAM2	ORB-SLAM3	本文算法	ORB-SLAM2	ORB-SLAM3	本文算法
最大值/m	0.179 3	0.102 5	0.097 4	0.244 3	0.236 8	0.132 6
平均误差/m	0.044 3	0.029 0	0.021 1	0.079 2	0.060 7	0.039 5
误差中值/m	0.041 9	0.027 8	0.016 1	0.091 9	0.049 0	0.037 8
最小值/m	0.002 8	0.000 4	0.001 5	0.007 4	0.009 9	0.002 3
均方根误差/m	0.051 9	0.033 9	0.025 0	0.085 8	0.067 6	0.048 8
方差/ m^2	5.084 1	3.036 1	1.655 7	12.155 9	9.030 6	4.697 5
标准差/m	0.027 0	0.017 7	0.013 5	0.034 9	0.029 8	0.028 5

为验证本文算法在葡萄果园场景中的定位效果,在果园中选择 1 条直线路径进行定位实验,并与不同算法对比,结果如图 8 所示。从图 8 可知,PDS-SLAM 在运行过程中累计误差较小,更贴近真实轨迹。

4.2 特征点均匀化效果对比

为验证 PDS-SLAM 在特征点提取的效果,在公开数据集和自建数据集中随机挑选 20 幅图像,进行特征点匹配实验,统计特征点匹配数量。如图 9 所示,PDS-SLAM 算法在公开数据集和自建数据集中,能够有效提高低纹理区域特征点提取数量,并且缓解特征点局部密集现象,提高特征点均匀化程度。

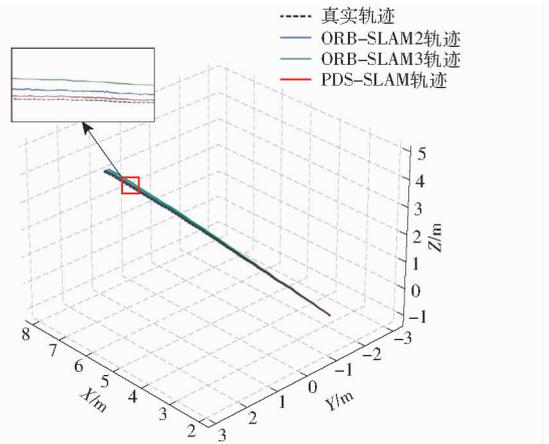


图 8 自建数据集轨迹对比

Fig. 8 Trajectory comparison on self-collected dataset

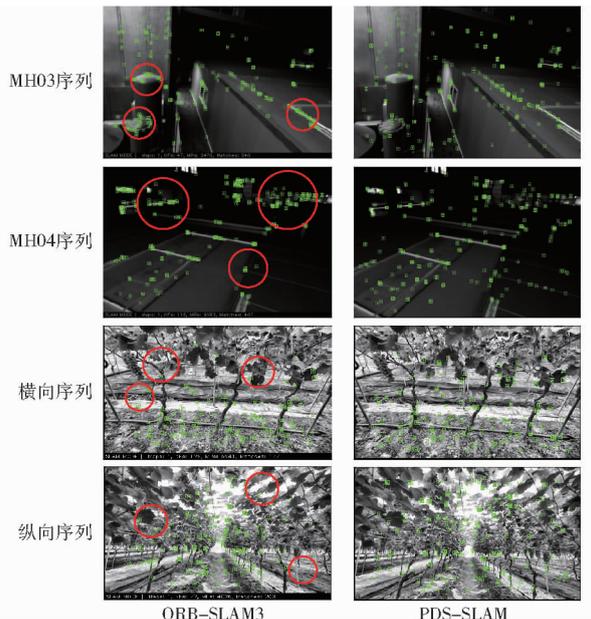


图 9 特征点均匀化效果

Fig. 9 Uniformization effect of feature points

统计实验结果如图 10 所示。由图 10 可知,在相同场景下,PDS-SLAM 特征点匹配数量均高于 ORB-SLAM3,其 ORB 特征点匹配数量平均提升 15.5%,提高了特征点提取和匹配效果。

4.3 语义分割算法性能

不同算法训练及测试均在同一环境下运行,实验环境为 Pytorch 1.11,CUDA 10.1,训练优化器为 SGD,初始学习率为 0.01,权重衰退系数为 0.005,批次大小为 8,迭代次数为 60 000。为全面评估所提出的 PDSNet 算法对葡萄果实的语义分割性能,基于自建数据集,选取 SegFormer-B0、BiSeNetV2、与 PIDNet 作为对比算法,并在统一训练配置下进行实验。

结果如表 2 所示,在自建测试集上,PDSNet 取得最佳 IoU,达到 78.9%,相比原算法提高 2.29 个

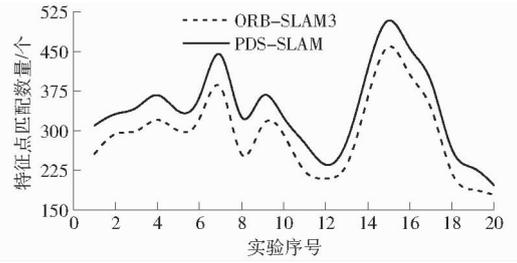


图 10 特征点匹配数量结果

Fig. 10 Number of feature point matching results

百分点,虽 PDSNet 计算量与参数量高于其他算法,但在相同硬件与输入分辨率下仍达到 126.92 f/s,可以满足果园机器人对果实检测实时性要求。为直观体现算法分割效果,选取 2 幅测试集图像对不同算法的分割结果进行可视化,如图 11 所示。

表 2 不同分割算法在测试集上的实验结果

Tab. 2 Different segmentation algorithms on test set

模型	IoU/%	Pa/%	F1 值/%	浮点运算量	参数量	帧率/(f·s ⁻¹)
SegFormer-B0	74.93	95.25	85.67	6.860×10^9	3.72×10^6	177.70
BiSeNetV2	75.77	95.27	86.21	1.082×10^{10}	3.34×10^6	223.04
PIDNet	76.61	95.30	86.76	1.985×10^{10}	2.88×10^7	162.16
PDSNet	78.90	95.36	88.21	3.079×10^{10}	3.08×10^7	126.92

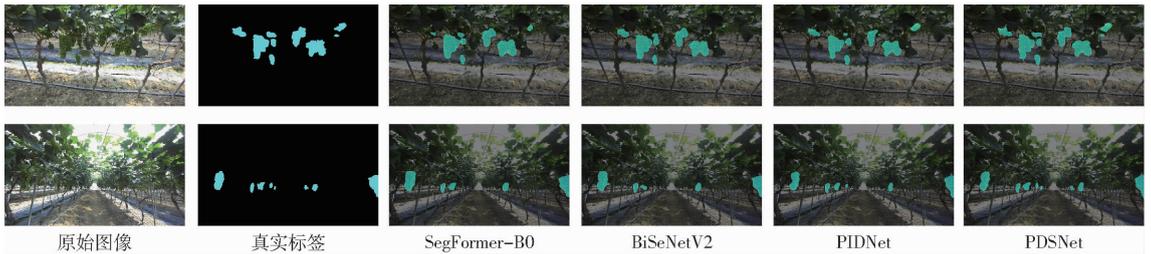


图 11 不同分割算法分割效果

Fig. 11 Segmentation effect of different segmentation algorithms

从图 11 可以看到,SegFormer-B0 对远处小尺度和枝叶较多处的果实出现漏检问题;BiSeNetV2 对远处果实有一定感知能力,但仍存在漏检现象,且在树枝遮挡情况下无法准确分割被遮挡果实;PIDNet 对远处小尺度果实存在漏检,同时对果实分割不完整;PDSNet 分割效果最好,对小尺度果实不存在漏检现象,对枝叶遮挡果实能够准确完整分割。总体而言,PDSNet 分割效果优于其他算法,满足果园机器人果实识别要求。

4.4 地图构建与滤波优化

在葡萄果园中,以横向视角对单列葡萄架和以纵向视角对双列葡萄架分别进行稠密建图实验,得到稠密地图如图 12 所示,其中图 12a、12c 为稠密点云地图,图 12b、12d 为结合语义掩膜对葡萄果实进行语义标注后的语义地图(其中红框区域为局部细节)。从图 12 可以看出,算法能够根据真实场景图

像信息,实现稠密建图,并且融合语义分割信息,将稠密地图中的葡萄点云进行语义标注,得到包含葡萄语义信息的语义地图。

以横向序列为例,滤波优化后的稠密地图效果,如图 13(图中红框为同一处局部区域的滤波前后细节)所示,图 13a、13b 中上下稠密地图分别对应滤波优化前后整体地图。从图 13 可看出,双重滤波机制能够有效去除稠密地图和语义地图中的离群点和空间点云噪声,有效提升稠密地图质量,减少杂乱点云对果园机器人干扰。

综上,本文算法能够融合语义信息,构建可用于果园机器人作业的语义地图,为果园机器人实现喷药、采摘等工作任务提供语义地图信息,同时融合双重滤波机制,能有效减少地图中的离群点和空间点云噪声,提高稠密地图质量,减少无效点云干扰。

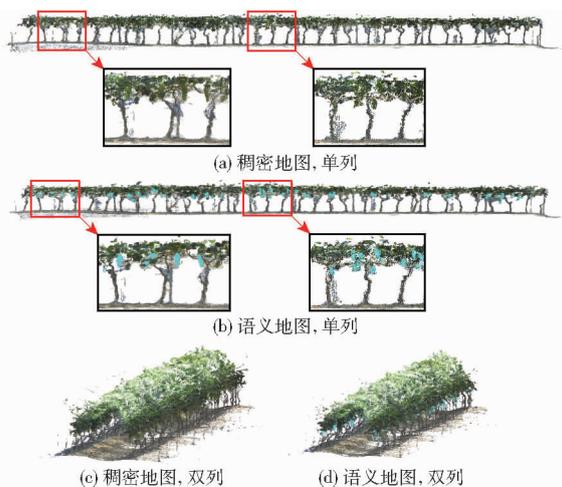


图 12 单双列葡萄稠密建图效果

Fig. 12 Dense mapping of single and double row grape trellis

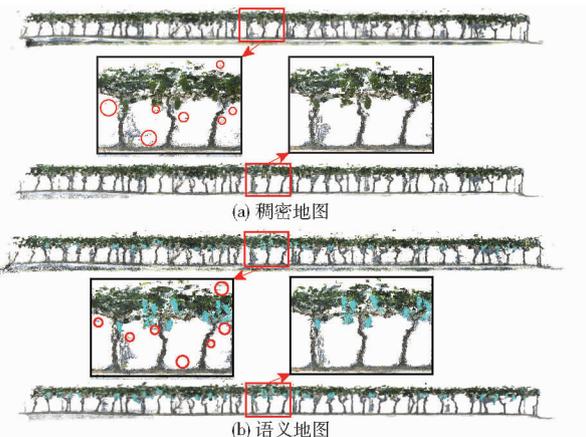


图 13 稠密地图优化效果

Fig. 13 Effects of dense map optimization

4.5 八叉树地图实验

为满足果园机器人导航和作业需要,将稠密地图转换为八叉树地图,如图 14(图中红框区域为局部细节)所示,不同地图处理后内存成本如表 3 所示,双重滤波机制在剔除稠密地图离群点和空间噪声点云后,能够在一定程度上减少地图的内存占用量,而将稠密地图转换成八叉树地图,能够大幅减少地图大小,八叉树地图相比语义地图,内存占用量平均减少 85.9%,便于果园机器人进行地图检索。

5 结论

(1) 提出了一种融合语义分割的稠密建图方法

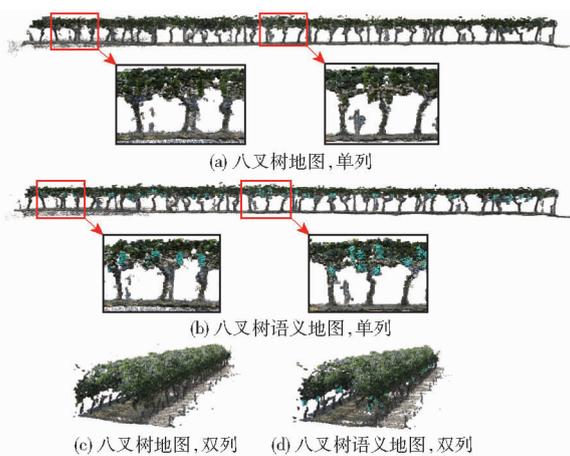


图 14 单双列葡萄八叉树地图

Fig. 14 Octomap of single and double row grape trellis

表 3 不同地图内存成本对比

Tab. 3 Storage cost comparison of different maps

地图	横向序列	纵向序列	平均
语义地图/MB	20.5	36.6	
滤波后语义地图/MB	13.1	29.1	
八叉树地图/MB	1.8	7.1	
内存占用量提升/%	91.2	80.6	85.9

PDS-SLAM,设计了一种适用于果园非结构化环境的前端鲁棒定位方法,通过分区自适应 FAST 阈值角点提取策略与引入最小距离抑制的四叉树均匀化算法,增强了在弱纹理场景下特征点提取能力与均匀性,有效降低了轨迹误差,提升了定位精度。实验结果表明,本文方法相比 ORB-SLAM3 在定位精度上有显著提升。

(2)设计了一种轻量化语义分割网络 PDSNet,针对葡萄果实密集、遮挡、形态不规则的分割难点,通过在 PIDNet 的空间分支中嵌入 DSA 模块,有效增强了算法的特征感知能力。PDSNet 在自建数据集上,处理速度为 126.92 f/s 时 IoU 达 78.9%。

(3)构建了从特征点提取到高质量语义点云地图的建图算法,不仅通过语义掩膜实现了对葡萄果实点云的精确标注,并集成了统计与半径双重滤波机制,有效剔除了环境噪声与离群点,提升了地图质量;最终生成八叉树地图,大幅降低内存占用量,为果园机器人进行自主导航和精准作业提供了实时性的解决方案。

参 考 文 献

[1] WANG R, CHEN L, HUANG Z, et al. A review on the high-efficiency detection and precision positioning technology application of agricultural robots[J]. Processes, 2024, 12(9): 16-24.

[2] LIU T, KANG H, CHEN C. ORB-Livox: a real-time dynamic system for fruit detection and localization[J]. Computers and Electronics in Agriculture, 2023, 209(13): 107834.

[3] YUAN Q, WANG P, LUO W, et al. Simultaneous localization and mapping system for agricultural yield estimation based on improved VINS-RGBD: a case study of a strawberry field[J]. Agriculture, 2024, 14(5): 34-38.

[4] MONTOYA-CAVERO L E, DE LEÓN TORRES R D, GÓMEZ-ESPINOSA A, et al. Vision systems for harvesting robots:

- produce detection and localization[J]. *Computers and Electronics in Agriculture*, 2022, 192(8): 106562.
- [5] XIONG J, LIANG J, ZHUANG Y, et al. Real-time localization and 3D semantic map reconstruction for unstructured citrus orchards[J]. *Computers and Electronics in Agriculture*, 2023, 213(36): 108217.
- [6] JIANG S, QI P, HAN L, et al. Navigation system for orchard spraying robot based on 3D LiDAR SLAM with NDT_ICP point cloud registration[J]. *Computers and Electronics in Agriculture*, 2024, 220(5): 108870.
- [7] LI Y, LI Y, NIE J, et al. Navigation of the spraying robot in jujube orchard[J]. *Alexandria Engineering Journal*, 2025, 126(6): 320–340.
- [8] 丛佩超, 崔利营, 万现全, 等. 基于改进 ORB-SLAM2 的果园喷药机器人定位与稠密建图算法[J]. *农业机械学报*, 2023, 54(7): 45–55.
CONG Peichao, CUI Liying, WAN Xianquan, et al. Localization and dense mapping algorithm for orchard spraying robot based on improved ORB-SLAM2[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(7): 45–55. (in Chinese)
- [9] LI Q, WANG X, WU T, et al. Point-line feature fusion based field real-time RGB-D SLAM[J]. *Computers & Graphics*, 2022, 107(5): 10–19.
- [10] 沈跃, 肖鑫桦, 刘慧, 等. 果园机器人 LiDAR/IMU 紧耦合实时定位与建图方法[J]. *农业机械学报*, 2023, 54(11): 20–28, 48.
SHEN Yue, XIAO Xinhua, LIU Hui, et al. Real-time localization and mapping method for agricultural robot in orchards based on LiDAR/IMU tight-coupling[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(11): 20–28, 48. (in Chinese)
- [11] HONG Y, MA R, LI C, et al. Three-dimensional localization and mapping of multiagricultural scenes via hierarchically-coupled LiDAR-inertial odometry[J]. *Computers and Electronics in Agriculture*, 2024, 227(26): 109487.
- [12] 薛金林, 褚阳阳, 宋悦, 等. 基于自适应阈值 ORB 特征提取的果园双目稠密地图构建[J]. *农业机械学报*, 2024, 55(6): 42–51, 59.
XUE Jinlin, CHU Yangyang, SONG Yue, et al. Construction of binocular dense map of orchard based on adaptive threshold ORB feature extraction[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(6): 42–51, 59. (in Chinese)
- [13] 娄路, 张忍, 李一天, 等. 基于平面匹配与目标检测的视觉 SLAM 算法[J]. *计算机工程与设计*, 2024, 45(4): 1240–1247.
LOU Lu, ZHANG Ren, LI Yitian, et al. Visual SLAM algorithm based on plane matching and object detection[J]. *Computer Engineering and Design*, 2024, 45(4): 1240–1247. (in Chinese)
- [14] WANG P, LUO W, LIU J, et al. Real-time semantic SLAM-based 3D reconstruction robot for greenhouse vegetables[J]. *Computers and Electronics in Agriculture*, 2025, 237(7): 110582.
- [15] 张鹏飞, 李宏伟, 赵亚帅, 等. 基于改进 SLAM 框架的动态场景三维语义地图构建方法研究[J]. *计算机应用与软件*, 2024, 41(11): 234–240.
ZHANG Pengfei, LI Hongwei, ZHAO Yashuai, et al. Construction method of 3D semantic map of dynamic scene based on improved SLAM framework[J]. *Computer Applications and Software*, 2024, 41(11): 234–240. (in Chinese)
- [16] 林义忠, 谢震鹏. 基于 YOLOv7-mask 和 ORB-SLAM2 的语义八叉树地图构建[J]. *机床与液压*, 2023, 51(23): 86–91.
LIN Yizhong, XIE Zhenpeng. Semantic octree map establishment based on YOLOv7-mask and ORB-SLAM2[J]. *Machine Tool & Hydraulics*, 2023, 51(23): 86–91. (in Chinese)
- [17] LV J, YAO B, GUO H, et al. MOLO-SLAM: a semantic SLAM for accurate removal of dynamic objects in agricultural environments[J]. *Agriculture*, 2024, 14(6): 7–10.
- [18] DIAO Z, CHEN L, YANG Y, et al. Localization technologies for smart agriculture and precision farming: a review[J]. *Computers and Electronics in Agriculture*, 2025, 236(9): 110464.
- [19] 高翔, 张涛, 刘毅, 等. 视觉 SLAM 十四讲 [M]. 北京: 电子工业出版社, 2019.
- [20] ROSTEN E, DRUMMOND T. Machine learning for high-speed corner detection[C]// *Computer Vision-ECCV 2006*, 2006: 430–443.
- [21] 李旭, 阳奥凯, 刘青, 等. 基于 ORB-SLAM2 的温室移动机器人定位研究[J]. *农业机械学报*, 2024, 55(增刊1): 317–324, 345.
LI Xu, YANG Aokai, LIU Qing, et al. Greenhouse mobile robot localization based on ORB-SLAM2[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(Supp.1): 317–324, 345. (in Chinese)
- [22] XU J, XIONG Z, BHATTACHARYYA S P. Pidnet: a real-time semantic segmentation network inspired by PID controllers[C]// *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023: 19529–19539.
- [23] YU Z, ZHANG X, ZHAO L, et al. Dsan: exploring the relationship between deformable convolution and spatial attention[J/OL]. *Techrxiv*, 2024. DOI:10.36227/techrxiv.171328860.02479778/v1.
- [24] HORNING A, WURM K M, BENNEWITZ M, et al. Octomap: an efficient probabilistic 3D mapping framework based on octrees[J]. *Autonomous Robots*, 2013, 34(3): 189–206.
- [25] 鲍秀兰, 包有刚, 马萧杰, 等. 自然环境下柑橘采摘机器人避障规划研究[J]. *农业机械学报*, 2025, 56(2): 420–428.
BAO Xiulan, BAO Yougang, MA Xiaojie, et al. Obstacle avoidance planning of citrus picking robot in natural environment[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(2): 420–428. (in Chinese)
- [26] MICHAEL B, JANOSCH N, PASCAL G, et al. The euroc micro aerial vehicle datasets[J]. *The International Journal of Robotics Research*, 2016, 35(10): 1157–1163.