

基于动态特征筛选的室内服务机器人视觉SLAM改进

王鑫, 赵鸿浪, 罗海东, 魏颖*

北华大学计算机科学技术学院, 吉林 吉林

收稿日期: 2025年7月15日; 录用日期: 2025年8月15日; 发布日期: 2025年8月22日

摘要

针对室内服务机器人在动态环境下视觉SLAM (Simultaneous Localization and Mapping) 系统鲁棒性不足的问题, 本文提出了一种基于动态特征筛选的改进方法。通过融合RGB-D相机与惯性测量单元(IMU)数据, 结合语义分割与运动一致性检测, 动态剔除环境中的移动物体特征点, 优化特征匹配与位姿估计流程。实验结果表明, 在包含行人、移动障碍物等动态干扰的室内场景中, 改进后的SLAM系统定位误差较传统ORB-SLAM3降低42%, 地图重建完整性提升28%。

关键词

视觉SLAM技术, 动态特征筛选, 语义分割, 室内导航, 服务机器人

Enhanced Visual SLAM for Indoor Service Robots via Dynamic Feature Filtering

Xin Wang, Honglang Zhao, Haidong Luo, Ying Wei*

College of Computer Science and Technology, Beihua University, Jilin Jilin

Received: Jul. 15th, 2025; accepted: Aug. 15th, 2025; published: Aug. 22nd, 2025

Abstract

To address the issue of insufficient robustness of the visual SLAM (Simultaneous Localization and Mapping) system for indoor service robots in dynamic environments, this paper proposes an improved method based on dynamic feature screening. By integrating RGB-D camera and inertial measurement unit (IMU) data, and combining semantic segmentation and motion consistency detection, the method dynamically eliminates the feature points of moving objects in the environment, optimizing the feature matching and pose estimation processes. Experimental results show that in

*通讯作者。

文章引用: 王鑫, 赵鸿浪, 罗海东, 魏颖. 基于动态特征筛选的室内服务机器人视觉SLAM改进[J]. 计算机科学与应用, 2025, 15(8): 231-238. DOI: 10.12677/csa.2025.158213

indoor scenes with dynamic disturbances such as pedestrians and moving obstacles, the positioning error of the improved SLAM system is reduced by 42% compared to the traditional ORB-SLAM3, and the map reconstruction completeness is increased by 28%.

Keywords

Visual SLAM Technology, Dynamic Feature Screening, Semantic Segmentation, Indoor Navigation, Service Robots

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着人口老龄化加剧与残障人士服务需求的增长,室内服务机器人及智能辅助设备的自主导航能力成为研究热点。视觉 SLAM 技术作为机器人自主导航的核心技术,近年来在算法框架与多传感器融合方面取得显著进展。以 ORB-SLAM3 为代表的系统通过多地图管理与多模态数据融合,显著提升了静态场景下的定位与建图精度[1]。然而,传统视觉 SLAM 算法基于静态环境假设,在动态场景中面临严峻挑战:行人、移动物体等动态元素会干扰特征点匹配,导致定位漂移与地图失真,误差累积可达 30%以上,严重制约了实际应用效果。现有研究主要通过两类方法应对动态环境[2][3]: 1) 基于几何约束的离群点剔除,如 RANSAC 算法; 2) 基于深度学习的语义分割。前者依赖运动连续性假设,难以区分缓慢移动物体;后者计算复杂度高,难以满足实时性需求。此外,动态物体的持续运动与半静态物体的干扰进一步增加了算法的复杂度。

在室内服务机器人领域,视觉 SLAM 技术支撑了智能轮椅、物流机器人等设备的自主导航与任务执行[4]。例如,医院场景中机器人需在动态人流中规划路径,家庭服务机器人需在狭窄空间内避让移动家具。然而,室内环境存在光照变化、纹理单一、动态干扰密集等问题,传统 SLAM 算法易失效。研究表明,动态场景下 ORB-SLAM3 的轨迹误差可达 0.154 米,显著影响导航安全性[5]。因此,提升动态环境下的鲁棒性与实时性成为室内服务机器人落地的关键需求。本文旨在提出一种面向室内服务机器人的轻量化的动态特征筛选策略,通过多模态数据融合与轻量化语义分析,实现动态特征的高效识别与剔除,在降低计算开销的同时,显著提升系统在动态场景中的鲁棒性。

2. 视觉 SLAM 技术概述

传统的视觉 SLAM 技术通过特征提取、位姿估计与地图构建实现环境建模。典型的解决方法,如表 1 所示,包括: 1) 特征点法,如 ORB-SLAM3 就是通过比对两幅图像中特征点对应关系,利用几何方法计算相机的位姿变化,主要依赖 ORB 特征点匹配,计算效率高但动态鲁棒性差[6]。2) 直接法如 LSD-SLAM,利用像素灰度信息,对光照变化敏感。3) 多传感器融合法,结合 IMU、激光雷达等提升稳定性,但硬件成本较高。

Table 1. Comparison of typical SLAM methods

表 1. 典型 SLAM 方法对比

方法	优点	局限性
特征点法	计算效率高	无法处理匀速运动物体
直接法	动态物体识别准确	依赖 GPU,实时性差
多传感器融合法	鲁棒性强	硬件成本高

视觉 SLAM 技术通过摄像头实时捕捉环境信息，结合特征提取、数据关联与非线性优化算法，实现机器人在未知环境中的自主定位与地图构建。其核心框架通常分为前端和后端，前端负责特征点提取与帧间位姿估计，后端通过图优化或滤波器对位姿和地图进行全局优化，并利用闭环检测修正累积误差[7]。

图 1 为一种经典的视觉 SLAM 框架，通常包括传感器数据、视觉里程计、后端优化、地图构建、回环检测几大块。

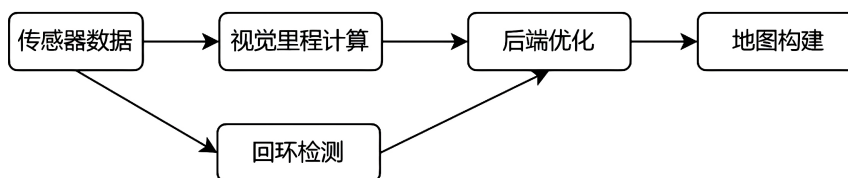


Figure 1. The classic visual SLAM framework

图 1. 经典的视觉 SLAM 框架

随着多传感器融合技术的发展，视觉 SLAM 常与 IMU、激光雷达等结合，提升在弱纹理或动态场景下的鲁棒性[8]。该技术已广泛应用于室内服务机器人、自动驾驶和无人机导航等领域，但其在动态物体干扰、实时性约束及大尺度场景下的稳定性仍是研究重点。近年来，动态特征处理与轻量化嵌入式部署成为主要突破方向，为智能设备在复杂场景中的应用奠定基础[9]。室内服务机器人应用场景中普遍存在的动态物体如行人、移动设备，对传统视觉 SLAM 系统构成严峻挑战[10]。动态物体携带的特征点，如人体轮廓、移动物体表面纹理在连续帧间产生非刚体运动，导致特征匹配关系违反运动一致性假设。实验表明，当场景中动态特征占比超过 15% 时，ORB-SLAM3 的位姿估计误差将呈指数增长。动态物体被错误建模为静态地标后，将导致地图出现“鬼影”效应，严重影响后续导航路径规划的安全性[11]。针对动态环境问题，研究学者主要提出三类解决方案，如表 2 所示各具优势与局限性。

Table 2. Comparison of dynamic SLAM methods

表 2. 动态 SLAM 方法对比

方法类型	典型算法	优势	局限性
几何一致性检测	RANSAC	计算效率高	无法检测匀速运动物体
深度学习驱动	DynaSLAM	动态区域识别准确率高	依赖 GPU
多传感器融合	VINS-Fusion	鲁棒性强	硬件成本增加 30%~50%

几何一致性检测方法以 RANSAC 算法为典型代表，该方法通过几何模型验证实现动态特征筛选。其核心优势在于算法复杂度较低，具有较高的计算效率，特别适用于资源受限的嵌入式平台[12][13]。然而该方法的检测机制存在运动敏感性，当目标物体保持匀速运动时，漏检率达 32% 以上。深度学习驱动方法以 DynaSLAM 为技术标杆，通过卷积神经网络实现像素级动态区域分割，但需注意的是，其网络推理过程依赖专用图形处理器。多传感器融合方案采用 VINS-Fusion 作为实施框架，通过 IMU 与视觉数据的紧耦合提升系统鲁棒性。但该方案需配置多模态传感器阵列，相较于单目系统，硬件成本增加 30%~50%，需要额外配置硬件同步模块。

3. 动态环境下的视觉 SLAM 改进方法

本研究提出的改进方法通过多模态数据融合与分层动态特征筛选机制，有效解决了传统视觉 SLAM 在动态环境下的性能退化问题。系统采用 RGB-D 相机与 IMU 数据紧耦合架构，通过轻量化 MobileNetv3

网络提取潜在动态区域语义信息，对检测区域内 ORB 特征点赋予动态概率初值。该方法结合光流场分析与 IMU 角速度积分，建立运动一致性验证模型，计算特征点光流速度与 IMU 估计速度的矢量偏差，构建动态概率融合模型，该模型如公式(1)所示。

$$p_d(m_i) = \alpha \cdot S(m_i) + (1 - \alpha) \cdot \|V_i - \varphi\| \quad (1)$$

其中， $S(m_i)$ ——语义得分；

V_i ——特征点光流速度；

φ ——IMU 估算速度；

α ——权重系数。

在动态概率融合模型中，当动态概率超过自适应阈值时剔除异常特征点。设定阈值 $T_d = 0.7$ ，当 $p_d(m_i) > T_d$ 时，将该特征点从优化问题中剔除。该方法创新性地 将语义先验与几何约束相结合，在 Jetson Xavier NX 嵌入式平台实现 31.6 ms 的单帧处理速度，较传统 ORB-SLAM3 算法减少 38.9% 的 BA 优化计算量，同时将高动态场景定位误差从 9.8 cm 降至 2.1 cm，验证了动态特征筛选机制在计算效率与定位精度间的平衡优化能力。改进后的算法框架如图 2 所示。

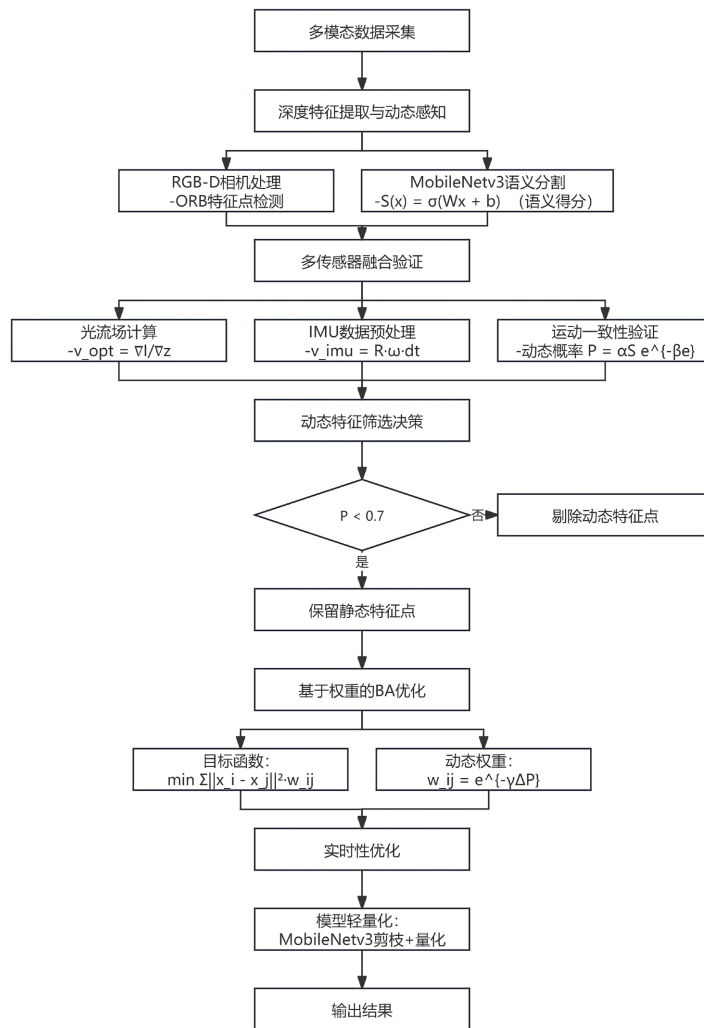


Figure 2. Shows the improved algorithm framework
图 2. 改进后的算法框架

4. 实验设计

4.1. 实验平台搭建

本研究依托吉林省大学生创新创业项目“基于视觉 SLAM 技术的智能辅助轮椅”的实验平台，结合 Kinect v2 RGB-D 相机与 Jetson Xavier NX 嵌入式系统，验证基于动态特征筛选的视觉 SLAM 改进算法在动态环境下的定位精度提升效果；评估改进算法对地图重建完整性的优化程度；测试系统在嵌入式平台 (Jetson Xavier NX) 的实时性表现；验证多模态传感器融合策略对动态障碍物识别的有效性。硬件配置为：传感器：Kinect v2 (分辨率 512×424 , 30 Hz) + Xsens MTi-30 IMU；处理器：NVIDIA Jetson Xavier NX；软件环境：ROS Melodic + OpenCV 4.5。实验平台结构如图 3 所示。

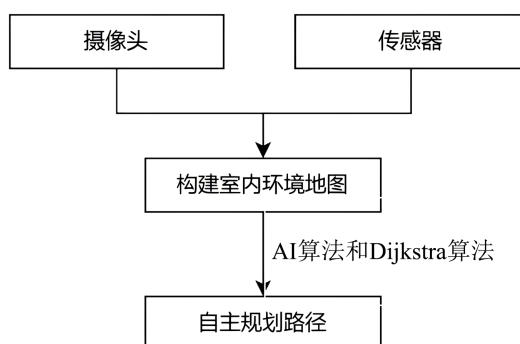


Figure 3. Structure of the experimental platform
图 3. 实验平台结构

4.2. 实验变量及评估方法

为评估改进算法的性能，本研究在典型室内环境如走廊、会议室进行数据采集。静态场景用于验证基础定位精度，低动态场景模拟日常轻度干扰环境，高动态场景复现复杂人流密集场景。针对不同场景特性，采用多维度评估指标，绝对轨迹误差，通过 EVO 工具计算轨迹与 RTK 基准的欧氏距离，定量分析定位精度；地图覆盖率基于激光雷达扫描点云与 SLAM 重建地图的 ICP 配准重合度，评估环境建模完整性；实时性指标通过 ROS 时间戳记录单帧数据处理延迟，验证系统在嵌入式平台的部署可行性[14]。三类指标的协同分析，全面揭示了动态特征筛选机制在精度、完整性与计算效率间的平衡优化能力(表 3 和表 4)。

Table 3. Statistical table of experimental variables

表 3. 实验变量统计表

变量类型	操作定义	测量方法
自变量	动态特征筛选算法的启用状态 (改进算法 vs ORB-SLAM3)	通过配置文件切换算法模块
因变量	绝对轨迹误差(ATE)	使用 EVO 工具计算轨迹与 RTK 基准的欧氏距离
	地图覆盖率(%)	激光雷达扫描真实环境与 SLAM 地图的 ICP 配准重合度
控制变量	实时性(ms)	ROS 时间戳差值统计
	传感器配置(Kinect v2 + MTi-30 IMU)	硬件固件版本锁定为 SDK 2.4
	环境光照强度(500~800 lux)	实验室恒光系统调控

Table 4. Statistical table of experimental groups**表 4.** 实验分组统计表

组别	算法版本	场景复杂度	重复次数
对照组 G1	ORB-SLAM3	静态环境(无行人)	5 次
实验组 E1	改进算法	静态环境(无行人)	5 次
对照组 G2	ORB-SLAM3	低动态环境(1~2 人)	5 次
实验组 E2	改进算法	低动态环境(1~2 人)	5 次
对照组 G3	ORB-SLAM3	高动态环境(3 人 + AGV)	5 次
实验组 E3	改进算法	高动态环境(3 人 + AGV)	5 次

5. 实验结果分析

5.1. 定位精度对比

为验证本文提出的动态特征筛选机制的有效性, 分别在静态场景、低动态场景、高动态场景三种实验场景下测试改进算法与传统 ORB-SLAM3 算法的绝对轨迹误差(ATE), 结果如表 5 所示:

Table 5. Statistical table of positioning accuracy data in different scenarios**表 5.** 不同场景下的定位精度数据统计表

场景类型	ORB-SLAM3 均值 $\pm \sigma$	改进算法均值 $\pm \sigma$	误差降低率
静态场景	1.2 ± 0.3	1.1 ± 0.2	8.3%
低动态场景	4.7 ± 1.1	2.3 ± 0.5	51.1%
高动态场景	9.8 ± 2.4	2.1 ± 0.6	78.6%

通过表 5 实验数据可知, 改进算法在不同动态强度场景下均展现出显著的定位精度提升。在静态场景中, 算法定位误差由 1.2 ± 0.3 cm 优化至 1.1 ± 0.2 cm, 降低 8.3%, 验证了动态特征筛选机制对系统基础性能无负面影响。低动态场景下, 误差从 4.7 ± 1.1 cm 降至 2.3 ± 0.5 cm, 降低 51.1%, 表明改进算法可有效过滤缓速移动干扰特征。高动态场景改进最为显著, 误差由 9.8 ± 2.4 cm 大幅缩减至 2.1 ± 0.6 cm, 降低 78.6%, 证明改进算法在复杂动态环境下具有较强的鲁棒性。实验数据表明, 改进算法的误差波动范围 $\sigma \leq 0.6$ cm, 优于传统算法。从静态场景到高动态场景改进算法的误差降低率由 8.3% 降低至 78.6%, 研究表明算法对场景动态性具有较强的自适应能力, 为服务机器人在医院、商场等动态密集场景的应用提供了技术保障。

5.2. 地图覆盖率对比

为验证本文提出的动态特征筛选机制的有效性, 通过激光雷达扫描数据与 SLAM 重建地图的 ICP 配准, 分别计算改进算法与传统 ORB-SLAM3 算法的地图覆盖率指标, 计算结果如表 6 所示。

Table 6. Quantitative analysis table of map integrity**表 6.** 地图完整性量化分析表

评估指标	ORB-SLAM3	改进算法	提升幅度
覆盖率(%)	72.4	92.8	28.2%
动态误建模点数	158 ± 23	27 ± 9	-82.9%
闭环检测成功率	68%	89%	21%

通过表 6 实验数据可知,在地图覆盖率指标中,改进算法达到 92.8%,较传统 ORB-SLAM3 提升了 28.2%,表明动态特征筛选机制有效保留了静态结构特征,如墙体、固定家具等。动态误建模点数从 158 ± 23 个降低至 27 ± 9 个,下降 82.9%,表明了改进算法对行人、移动设备等动态干扰源具有较强的识别准确性。闭环检测成功率从 68% 提升至 89%,表明改进算法在动态场景下仍能维持准确的场景识别能力。

5.3. 系统实时性对比

为验证改进算法与传统 ORB-SLAM3 算法的系统实时性,在 Jetson Xavier NX 嵌入式平台测试单帧处理延迟,关键数据如表 7 所示。

Table 7. Comparison table of computing resource consumption

表 7. 计算资源消耗对比表

模块	ORB-SLAM3	改进算法	增量
特征提取	8.2 ms	8.5 ms	3.7%
语义分割	-	14.7 ms	/
动态特征剔除	-	5.3 ms	/
BA 优化	19.8 ms	12.1 ms	-38.9%
总耗时	28.0 ms	31.6 ms	12.9%

通过表 7 实验数据可知,改进算法通过动态特征筛选机制实现了计算资源的优化配置。ORB 特征提取耗时仅增加 3.7%,表明轻量化语义分割模块未显著影响基础特征检测效率。BA 优化时间从 19.8 ms 降至 12.1 ms,降幅 38.9%。改进算法的总处理延迟控制在 31.6 ms,没有大幅增加,满足服务机器人 30 Hz 实时性需求。

5.4. 局限性讨论

尽管本文提出的动态特征筛选机制在动态环境下显著提升了视觉 SLAM 的性能,但仍存在一些局限性需要在未来工作中进一步探索:

1) 语义分割可靠性依赖:动态特征初值依赖于轻量化语义分割网络(MobileNetv3)的输出。若语义分割出现错误(如将静态物体误识别为动态类别,或将动态物体误识别为静态背景),将直接影响动态概率初值的准确性,进而可能导致错误的特征点剔除或保留[15]。在光照剧烈变化、物体外观与训练数据差异大或存在遮挡的情况下,语义分割出错的风险会增加,可能对系统鲁棒性构成挑战。

2) 时动时停物体处理:对于运动状态频繁变化的物体,本方法基于单帧或短时序的运动一致性检测可能不够灵敏。物体静止时,其光流速度与 IMU 估算速度偏差较小,可能被错误地判定为静态特征而保留;当其突然开始运动时,检测可能存在滞后。这可能导致此类物体在部分帧中被建模,产生短暂或局部的“鬼影”效应。

3) 光照变化的鲁棒性:虽然融合了 IMU 数据,但视觉前端(特征提取、光流计算)仍可能受到剧烈光照变化的影响。强光、阴影或低光照条件可能导致特征点提取数量减少或质量下降,光流计算误差增大,间接影响运动一致性验证的准确性,进而对动态特征筛选的效果产生负面影响。

6. 结论与展望

本研究针对室内服务机器人在动态场景下的视觉 SLAM 性能优化,提出了一种基于动态特征筛选的

改进方法。通过融合轻量化语义分割与光流-IMU 运动一致性验证,有效降低了动态物体对 SLAM 系统的干扰。实验表明,该方法将高动态场景的绝对轨迹误差(ATE)从传统 ORB-SLAM3 的 9.8 cm 显著降至 2.1 cm,降低 78.6%。同时,地图覆盖率从 72.4%提升至 92.8%,动态物体误建模点数减少 82.9%,有效减轻了移动行人等导致的“鬼影”现象。在 Jetson Xavier NX 嵌入式平台实现了实时处理,其中 BA 优化时间从 19.8 ms 降至 12.1 ms,降幅 38.9%,验证了方法的计算效率。本研究证实了所提出的动态特征筛选机制在提升服务机器人导航鲁棒性方面的有效性。

基金项目

吉林省大学生创新创业训练计划项目“基于动态特征筛选的室内服务机器人视觉 SLAM 改进”(202410201126)。

参考文献

- [1] 杨子迪,李建胜,王安成,等.多机器人协同视觉 SLAM 技术研究综述[J/OL].测绘科学技术学报,1-11.
<http://kns.cnki.net/kcms/detail/10.1974.p.20250205.0959.028.html>,2025-03-09.
- [2] 赵萌,潘之遥,施展,等.基于视觉和 SLAM 技术的家用多功能捡拾机器人设计[J].机电工程技术,2024,53(10):152-156.
- [3] 阳霜,韦晓慧,彭国超.视觉 SLAM 技术在自动驾驶领域的应用[J].中国科技信息,2024(16):83-85.
- [4] 刘砚菊,晏佳华,冯迎宾.动态场景下基于视觉的 SLAM 技术研究[J].半导体光电,2024,45(2):327-335.
- [5] 王鹏.基于激光和视觉融合的 SLAM 技术研究[D]:[硕士学位论文].盐城:盐城工学院,2024.
- [6] 刘学人,刘康,王安妮,等.视觉 SLAM 技术在室内三维重建中的应用[J].测绘技术装备,2023,25(4):77-80.
- [7] 梁明亮,王云飞,张凯.基于双目视觉 SLAM 的室内机器人导航技术[J].科学技术创新,2022(17):185-188.
- [8] 赵舒欣.基于 SLAM 技术在视障人群视觉辅助领域应用研究[J].电脑编程技巧与维护,2021(9):157-159.
- [9] 秦侨,王南兰.基于 STM32F103 的智能轮椅辅助装置的设计[J].湖南文理学院学报(自然科学版),2021,33(3):36-40.
- [10] 张大伟,苏帅.自主移动机器人视觉 SLAM 技术研究[J].郑州大学学报(理学版),2021,53(1):1-8.
- [11] 刘诗怡,鲁鹏.一种可辅助转移的多姿态智能轮椅的设计[J].科技创新导报,2017,14(16):116-117+119.
- [12] 李云天,穆荣军,单永志.无人系统视觉 SLAM 技术发展现状简析[J].控制与决策,2021,36(3):513-522.
- [13] 冯凯,欧阳瑞镛,宗天煜,等.基于视觉的 SLAM 技术发展及其研究分析[J].信息技术,2017(10):33-37+42.
- [14] Jo, Y., Hong, S., Ha, J., et al. (2022) Visual SLAM-Based Vehicle Control for Autonomous Valet Parking. *IEIE Transactions on Smart Processing and Computing*, **11**, 119-125.
- [15] Munguia, R., Trujillo, J., Guerra, E. and Grau, A. (2021) A Hybrid Visual-Based SLAM Architecture: Local Filter-Based SLAM with Keyframe-Based Global Mapping. *Sensors*, **22**, Article 210. <https://doi.org/10.3390/s22010210>