动态综述

DOI：10.16818/j.issn1001－5868.20251128002

移动机器人基于视觉的导航综述

胡章芳，张　薇*

（重庆邮电大学 电子科学与工程学院，重庆 400065）

摘　要： 视觉导航作为移动机器人自主运行的核心技术支撑，其性能直接决定移动机器人的环境感知精度、定位建图可靠性与路径规划的合理性。文章系统综述移动机器人视觉导航的研究进展，围绕视觉传感器、同步定位与地图构建（Simultaneous Localization and Mapping，SLAM）和路径规划三大核心环节展开分析：在视觉传感器层面，重点探讨单模态、多模态融合视觉传感器和新型视觉传感器的技术特性与适配场景；在SLAM层面，总结传统几何SLAM、多模态融合SLAM以及神经隐式SLAM的技术演进与性能优势；在路径规划层面，重点介绍传统算法与生物启发算法的特点与适用场景。最后，总结当前技术面临的挑战，并对未来研究方向进行展望，为视觉导航技术的进一步发展提供参考。

关键词： 视觉导航；多模态融合；事件相机；神经隐式SLAM；路径规划

中图分类号： TP242.6　**文章编号：** 1001－5868(2026)01－0028－12

A Review of Visual Navigation for Mobile Robots

HU Zhangfang, ZHANG Wei

(**School of Electronic Science and Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065**)

Abstract: Visual navigation has emerged as a key technology that supports the autonomous operation of mobile robots. The performance of visual navigation systems directly determines the accuracy of a robot's perception of its operating environment as well as the reliability of its positioning and mapping capabilities and the rationality of its path planning. In this paper, we systematically review the relevant literature on visual navigation for mobile robots. We explore three core aspects, including visual sensors, simultaneous localization and mapping (SLAM), and path planning. First, we consider the technical characteristics and application scenarios of existing and emerging visual sensors with single- and multimodal fusion. We also summarize the evolution and performance advantages of SLAM methods from classical geometrical techniques through multimodal fusion and neural implicit models. In our review of works on path planning, we introduce the characteristics and application scenarios of traditional and bioinspired algorithms. Finally, we summarize current challenges in visual navigation and suggest some possible avenues for future research. Thus, this review serves as a practical reference for the further development of visual navigation technology.

Key words: visual navigation; multimodal fusion; event camera; neural implicit SLAM; path planning

收稿日期：2025－11－28.

*通信作者：张　薇

0　引言

随着智能制造、服务机器人、自动驾驶等领域的快速发展，移动机器人在复杂未知环境中实现高精度和高鲁棒性自主导航的需求日益迫切，而视觉导航（Visual Navigation）凭借成本低廉、信息丰富和被动感知等优势，已成为实现机器人环境理解、实时定位与路径规划的关键途径[1-2]。完整的视觉导航系统包含三个关键组成部分：视觉传感器模块、同步定位与地图构建（Simultaneous Localization and Mapping，SLAM）模块和路径规划模块。其中，视觉传感器模块负责对外界环境进行数据采集，为系统提供原始感知信息；SLAM模块基于原始感知信息来判断机器人位姿和实时环境建图；路径规划模块根据地图信息生成从起点到目标点的安全、高效运动轨迹。这三个模块相互衔接，构成完整的导航闭环[2]。

近年来，视觉导航技术呈现多维度突破：在传感器层面，从单模态视觉感知向多模态融合及新型事件相机拓展[1]，大幅提升了视觉传感器在极端场景的适应性；SLAM技术从传统几何建模演进至多模态融合与神经隐式表示，实现高精度定位与高保真建图的统一[3-4]；路径规划算法从确定性传统算法向生物启发的全局优化算法延伸，增强动态复杂环境的适配能力[5]。然而，现有技术仍面临诸多问题：多模态感知技术在极端光照、无纹理环境下的融合鲁棒性不足；神经隐式SLAM的高保真重建与实时性、算力需求难以兼顾；路径规划算法难以在追求全局最优解的同时与实际应用需求之间实现高效平衡。

为此，本文围绕视觉导航的视觉传感器、SLAM、路径规划三大核心模块，系统梳理技术演进脉络、关键算法原理、性能对比与应用场景，总结当前研究瓶颈，并展望未来发展方向，旨在为相关领域研究者提供全面的技术参考，推动移动机器人视觉导航技术向高精度、高鲁棒性和低功耗方向迈进。

1　视觉传感器模块

智能机器人强大的自主性取决于选择和融合正确的传感器套件，视觉传感器作为视觉导航系统的感知入口，其技术发展直接决定环境信息采集的精度、鲁棒性与场景适应性。按照感知模态和技术创新性将视觉传感器分为三类：单模态视觉传感器、多模态融合视觉传感器和新型视觉传感器，分别从基础技术特性、导航核心感知性能和导航场景与任务适配三个方面进行描述。

1.1　单模态视觉传感器：单目和多摄像机RGB系统、立体视觉系统、RGB-D

单模态视觉传感器完全依赖于环境反射的可见光或红外光强度信息来形成图像。单目RGB系统结构简单、成本低，但缺乏直接的深度感知[6-7]；多摄像机RGB系统和立体视觉系统通过多视角几何原理计算深度，能提供较为稠密的深度图[8]；RGB-D系统主动投射红外光，直接获取像素级的深度信息，但易受环境光干扰[9]。

单模态视觉传感器在导航感知性能上各具特色：单目RGB系统虽无法直接获取环境深度信息，但通过学习型深度估计或运动恢复结构能实现稠密场景理解与语义定位，为轻量级SLAM系统提供基础[6-7]；多摄像机RGB系统通过多视角图像协同采集与融合，可强化视差计算的鲁棒性与深度估计范围，在大场景覆盖与动态目标跟踪中展现出更优的环境感知完整性；立体视觉系统通过视差计算直接输出度量深度，在室外光照复杂环境中仍能保持稳定的避障与三维重建能力；RGB-D相机则在近距离低纹理环境下提供高精度的密集深度图，极大提升了室内场景中的实时稠密建图与精细操作性能[9-10]。

单模态视觉传感器主要适用于室内结构化环境或光照稳定的室外场景，如仓库AGV导航、服务机器人室内巡弋，以及低速下的无人机避障。由于单模态视觉传感器依赖特征匹配，在缺乏纹理或重复纹理的区域，或者光照剧烈变化的场景中，其定位功能容易失效。

1.2　多模态融合视觉传感器：LiDAR-相机融合系统、视觉-IMU融合系统

多模态融合视觉传感器通过融合视觉传感器与其他互补传感器，以克服单一视觉模态的固有缺陷。LiDAR-相机融合将LiDAR精确且不受光照影响的3D点云与相机丰富的纹理和颜色信息相结合[11]；视觉-IMU融合利用IMU的高频短时运动预测来弥补视觉在快速运动或短暂遮挡时的数据缺失[12]。

多模态融合视觉传感器实现了性能的跃升。LiDAR-相机融合系统能生成带有语义信息的精确3D地图，极大提升了SLAM的精度和场景理解能力[11]。视觉-IMU系统凭借IMU的高频惯性测量，

能有效处理相机快速运动导致的图像模糊，并提供尺度的观测，在剧烈运动和无纹理环境中表现出极强的鲁棒性[12-13]。

LiDAR-相机系统成为自动驾驶汽车和在复杂、大尺度环境中工作机器人的核心感知套件，能够在各种光照条件和速度下实现可靠定位、建图与避障[14]，视觉-IMU系统广泛应用于无人机、AR/VR头显等对重量和成本敏感的设备[13]。

1.3 新型视觉传感器：事件相机

事件相机的像素单元采用独立异步触发机制，仅在感知到亮度变化时输出包含时间戳、坐标及变化极性的事件流，而非像传统相机周期性的采集完整帧图像。这使其具备了微秒级延迟、超过120 dB的高动态范围和极高的时间分辨率[15-16]。

事件相机从根本上解决了运动模糊的问题，使其在高速运动中仍能捕捉清晰的边缘特征。高动态范围特性使得事件相机能在明暗交替的极端光照条件下稳定工作。基于事件的SLAM和光流算法能够利用这些特性，在传统视觉系统失效的高速、高动态范围场景中实现低延迟和高鲁棒性的状态估计[17-18]。

事件相机非常适用于高速无人机在复杂环境下的避障与敏捷飞行、在强光或阴影下穿梭的自动驾驶汽车[18]，以及在光线微弱或频繁闪烁的工业环境中作业的机器人[19-20]。其低功耗特性也使其成为对功耗有严苛要求的长期自主导航系统的理想选择。

1.4 小结

视觉传感器作为视觉导航系统的感知入口，构建了单模态奠基、多模态升级和新型传感器突破的完整技术体系，为后续环境语义处理以及同步定位与地图构建的精准性提供了坚实的数据支撑。典型的视觉传感器对比如表1所示。

表1 典型的视觉传感器对比

传感器类别	具体类型	视场角	时间分辨率	延迟	功耗	典型应用场景
单模态视觉传感器	单目RGB相机	30°～90°	毫秒级(30 fps≈33 ms)	10～30 ms	低(＜1 W)	低成本室内机器人定位、手机AR导航
单模态视觉传感器	多摄像机RGB系统	拼接后可达360°	毫秒级(30 fps≈33 ms)	10～30 ms	中(1～3 W)	移动机器人360°环境感知、全景监控
单模态视觉传感器	立体视觉系统	单个镜头30°～120°，双镜头重叠＞20%	毫秒级(30 fps≈33 ms)	15～40 ms	中(2～5 W)	室外无人机地形测绘、自动驾驶近距离避障
单模态视觉传感器	RGB-D相机	67°～108°	毫秒级(30 fps≈33 ms)	10～30 ms	中高(3～8 W)	室内稠密建图、机器人近距离避障
多模态融合视觉传感器	LiDAR-相机融合系统	LiDAR 360°＋相机30°～120°	激光LiDARHz级＋相机毫秒级(30 fps≈33 ms)	10～50 ms	高(8～15 W)	自动驾驶、城市级SLAM
多模态融合视觉传感器	视觉-IMU融合系统	30°～360°	相机毫秒级(30 fps≈33 ms)＋IMU千赫兹级	5～20 ms	低(＜5 W)	微无人机导航、手持设备定位、动态SLAM
新型视觉传感器	事件相机	～75°	微秒级(1 μs)	＜1 ms	极低	无人机竞速导航、工业高速流水线检测、高动态导航

当前视觉传感技术在极端光照、浓雾、无纹理等复杂环境下，多模态融合鲁棒性显著下降。因此，如何提升多模态融合在极端场景的稳定性，同时降低算力依赖，成为未来研究需要重点关注的方向。

2 SLAM

SLAM是连接视觉感知与路径规划的核心枢纽，其目的是让移动机器人在未知环境中实时自身位姿估计并构建全局一致的环境模型，它的性能直接决定了视觉导航系统的 定位精度与场景适配能力。SLAM技术演进主要围绕三大目标：提升复杂环境下的鲁棒性、优化建图保真度和平衡精度与实时性。这一演进实现了从离散点云建模到连续智能表征的质变[3]，SLAM按照技术形式分为传统几何SLAM、多模态融合SLAM和神经隐式SLAM三类[4]。

2.1 传统几何SLAM

在视觉SLAM的演进过程中，传统方法逐渐从基于滤波的框架过渡到以非线性优化为核心的成熟系统[21-22]。其中，Mur-Artal等提出的ORB-SLAM系统[23]成为标志性成果，该方法通过多线程结构和地图重用机制显著提升系统的鲁棒性与精度。后续Zhang等和Campos等提出改进的ORB-SLAM方法[24-25]。近年来，由Yang等提出UPLP算法[26]，统一融合点、线、面三类几何特征，通过统一的特征表示、关联机制与误差函数，提高了复杂环境下的鲁棒性。

2.2　多模态融合 SLAM

多模态融合 SLAM 分为松耦合与紧耦合两个方面。

在松耦合方面，Zhang 和 Singh 率先提出视觉-LiDAR 松耦合系统[27]，奠定了传感器独立处理和结果融合的基础框架。随着成熟视觉 SLAM 算法的开源，Xu 等推出基于 ORB-SLAM2 扩展的视觉-LiDAR 松耦合方案[28]，此方法在纹理丰富的场景内依赖视觉完成定位建图，当视觉跟踪失效时自动切换至 2D LiDAR SLAM 模式，两类地图信息融合大幅提升室内场景鲁棒性。Zhu 等提出了 CamVox[29]，该方法实现自动标定，增强了在弱纹理环境中的建图能力。Hulchuk 等提出一种基于图的 LiDAR-IMU SLAM 方法[30]，通过融合视觉里程计并引入故障检测机制，提升了在特征稀疏场景中的定位鲁棒性。

在紧耦合方面，Campos 等提出的 ORB-SLAM3 是一个重要进展[25]，该系统支持多种相机配置与视觉-惯性紧耦合，并具备多地图无缝融合能力，显著提升了系统的适用性与鲁棒性。Zheng 等提出 Fast-Livo[31]，该方法是一种紧耦合的稀疏直接激光雷达惯性视觉里程计，具有卓越的实时性能，它将地图点与观测图像像素块结合，并将其投影到新图像上，通过最小化光度误差获得所有状态的姿态估计。针对 Fast-Livo 在动态性适配、精度稳定性、环境鲁棒性等方面的缺陷，Zheng 等提出 Fast-Livo2[32]，通过高效的误差状态迭代卡尔曼滤波（ESIKF）集成了 IMU，LiDAR 和图像数据。Lin 等提出 R3LIVE[33]，该方法包括 LIO 和 VIO 两个子系统，其中 LIO 子系统利用 LiDAR 和惯性传感器构建全局地图的几何结构，VIO 子系统通过降低帧与地图之间的光度误差，直接且高效地融合视觉数据。Lin 与 Zhang 基于 R3LIVE 提出 R3LIVE++[34]，该方法通过视觉-惯性里程计与 LiDAR-惯性里程计的紧耦合，实现了实时高精度的状态估计。Zhou 等提出的 MAst3R-Fusion[35]，借助分层滑动窗口设计、全局回环检测及同构群变换策略，在抑制累积漂移的同时保障定位精度，通过多维度技术手段兼顾了局部性能和全局一致性。

这些方法共同推动了多模态视觉 SLAM 在精度、鲁棒性与实用性方面的显著进展。

2.2.1　典型的多模态融合 SLAM：R3LIVE++

R3LIVE++融合了 LiDAR、惯性和视觉，通过考虑相机的光度定标和相机曝光时间的在线估计，进一步提高了定位和测绘的精度，同时可以动态重建辐射地图。R3LIVE++的系统框图如图 1 所示。

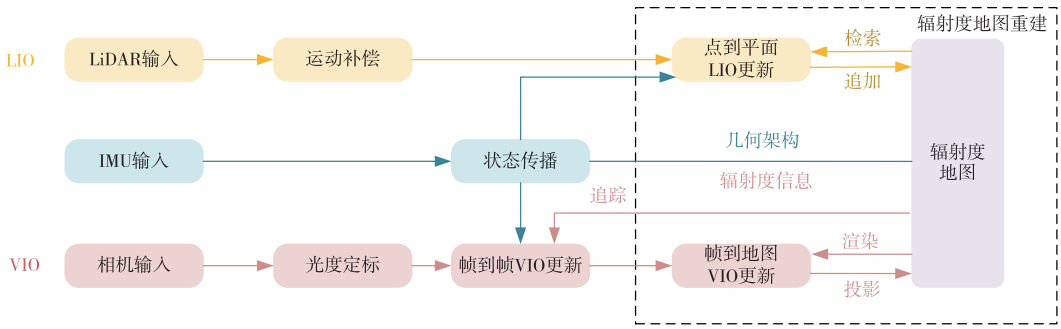


图 1　R3LIVE++的系统框图

R3LIVE++框架包括两个子系统：LIO 子系统和 VIO 子系统。LIO 子系统通过配准每个输入 LiDAR 扫描的点云测量来构建辐射地图的几何结构。VIO 子系统分两步来恢复辐射度地图的信息：帧到帧 VIO 更新通过最小化帧到帧的 PNP（Perspective-n-Point）重投影误差来估计系统状态，而帧到地图 VIO 更新最小化地图点和当前图像之间的辐射误差。这两个子系统紧密耦合在一个流形误差状态迭代卡尔曼滤波框架（ESIKF）内。其中 LiDAR 和相机视觉测量在其各自的数据接收时间上被融合为相同的系统状态。

研究者为测试 R3LIVE++的性能，在公共和私有数据集上对此方法进行广泛的实验。结果表明，该方法不仅显著提升了定位与测绘的精度表现，还在 LiDAR 或相机测量出现退化的高挑战性场景中，展现出极强的鲁棒性，能够稳定输出可靠结果。

2.2.2　典型的多模态融合 SLAM：MAst3R-Fusion

MAst3R-Fusion 是一个多传感器辅助的视觉 SLAM 框架，将前馈点图回归与互补的传感器信息紧密地结合在一起，包括惯性测量和全球导航卫星

系统(GNSS)数据。基于一种分层因子图设计，MAst3R-Fusion允许实时滑动窗口优化和具有主动闭合的全局优化，从而实现实时姿态跟踪、公制尺度结构感知和全局一致映射。MAst3R-Fusion的系统流程图如图2所示。

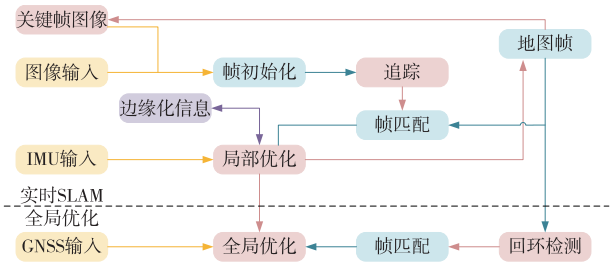


图2 MAst3R-Fusion的系统流程图

系统流程图主要有两部分：实时SLAM和全局优化。在实时SLAM阶段，通过前馈神经网络对输入图像进行处理，恢复点图并实现密集匹配。位姿追踪是通过在连续帧之间对齐点地图来执行的，在固定数量的关键帧上维护滑动窗口因子图，在可见帧之间建立基于相似变换群SIM(3)的视觉约束，同时紧密结合IMU数据。利用概率边际化技术最大化信息利用率，以低漂移的性能实现高精度的位姿估计和度量级结构恢复。在全局优化阶段，实时SLAM存储的信息通过额外的环路闭合和GNSS信息以全局方式进行处理。对于环路闭合，通过基于几何候选来选择的前馈模型实现跨时间的密集关联，全局因子图结合基于SIM(3)的视觉约束、IMU预积分因子和GNSS位置测量，实现了全局一致和无漂移的位姿估计。

研究者为测试MAst3R-Fusion的性能，在公共基准和自我收集的数据集上对此方法进行了评估，实验结果表明，MAst3R-Fusion在实时定位、全局一致性建图和密集三维感知三大核心任务上均表现更优，尤其在大视角闭环鲁棒性及多传感器紧耦合融合方面优势显著。

2.3 神经隐式SLAM

神经隐式表示(Neural Implicit Representation)通过神经网络直接学习场景的几何与外观连续函数，以实现高保真的稠密重建。iMAP[36]作为第一个神经隐式SLAM方法，采用基于多层感知机(Multi-Layer Perceptron，MLP)的表示法进行准实时的跟踪和映射，但重建和相机姿态的结果远远不如传统方法。为提升神经隐式SLAM在场景中的编码效率与重建质量，Zhu等提出NICE-SLAM[37]，该方法使用分层特征网格进行场景编码，在保证高重建质量的同时，增强了系统对大规模场景的可扩展性，显著优化了实用性能。在此基础上，针对神经编码的速度瓶颈，Johari等提出ESLAM[38]，该方法采用了高效的神经编码(Tri-plane与哈希)，大幅提升了编码的速度。随后，Liu等提出DDN-SLAM[39]，该方法融合语义与动态特征点分割，显著提升在动态复杂场景下的定位与建图性能。Li等提出DNS-SLAM[40]，该方法将多视角几何约束与基于图像的特征提取相结合，以改进外观细节，并输出颜色、密度和语义类别信息。Mao和Yu等提出NGEL-SLAM[41]，该方法融合传统特征跟踪模块与神经隐式表征，通过多子图设计、八叉树加速收敛及闭环响应优化，实现全局一致且低延迟的稠密建图与定位。Zhai等提出NIS-SLAM[42]，该方法通过多尺度特征融合与语义一致性优化，填补了传统神经隐式方法在场景理解层面的短板。

2.3.1 典型的神经隐式SLAM：NIS-SLAM

NIS-SLAM旨在构建一个高效的神经隐式语义RGB-D SLAM系统，该方法能够从预训练的2D分割网络中学习3D一致的语义表示，同时实现高精度语义地图构建重建。

NIS-SLAM结合了高频多分辨率四面体特征和低频位置编码作为场景的隐式表示，其中，四面体特征网格通过自适应划分捕捉场景细节，位置编码则建模全局空间信息，二者协同使场景表示既保留毫米级几何精度，又避免传统体素网格的内存爆炸问题。NIS-SLAM引入一种融合策略用于解决二维分割结果在多个视角下的不一致性问题，将历史非关键帧的语义预测概率加权整合至当前关键帧，通过时序一致性约束抑制孤立帧的语义错误，生成全局一致的3D语义场。并在相机位姿优化阶段引入基于置信度的像素采样与渐进优化权重函数，其中，置信度基像素采样根据像素的深度一致性和语义置信度来筛选有效观测点，剔除动态物体或纹理缺失区域的噪声数据。

为验证NIS-SLAM的综合性能，研究团队在Replica，ScanNet和TUM RGB-D等多类数据集上开展了大量实验，部分实验数据如表2。结果表明，NIS-SLAM在定位、重建与语义理解三大核心任务上均表现更优或相当，尤其在复杂场景鲁棒性、表面完整平滑度及多视图语义一致性方面优势显著。

表2　典型的神经隐式SLAM对比

系统	Replica/cm	ScanNet/cm	Depth L1/cm
iMAP	4.43	37.59	5.33
NICE-SLAM	2.08	10.20	3.53
ESLAM	0.62	7.1	1.18
DNS-SLAM	0.45	6.46	3.16
NIS-SLAM	0.46	7.88	1.48

注：Replica是8个场景均值，ScanNet是5个场景均值；Depth L1是深度估计的L1误差，数值越小表示深度重建精度越高。

2.3.2　典型的神经隐式SLAM：DNS-SLAM

DNS-SLAM是一种基于神经隐式表征的RGB-D语义SLAM方法，其核心是解决传统神经隐式SLAM在真实场景中重建过度平滑和噪声深度鲁棒性差的问题，同时能够实现高精度相机定位、稠密几何重建与语义分割的协同优化，并为机器人导航等任务提供语义化三维地图。

DNS-SLAM内容围绕语义、几何和多视图协同优化展开。首先，采用类别专属神经表征，为每个语义类别配备独立小型MLP解码器，结合2D语义先验实现类别级场景分解重建，并支持动态类别初始化，避免不同类别特征干扰。其次，引入多视图一致性约束，从参考帧提取ResNet-18二维特征，通过相机位姿反投影获取当前帧三维点对应的参考帧像素特征并编码为条件输入，多参考帧特征经平均池化聚合后为颜色和语义预测提供丰富外观约束，缓解重建过度平滑等问题。同时，设计自监督轻量粗表征，以精细语义表征的隐式向量为伪标签训练单MLP结构，跟踪时仅用粗表征快速优化相机位姿，从而平衡实时性与精度。最后，引入占用概率近似监督，对截断区域点用高斯分布近似真实占用概率，对远离表面点施加自由空间损失，从而强制占用概率趋近于零，达到提升表面平滑性和细节保留的能力，减少重建伪影。

为验证DNS-SLAM的性能，研究者在Replica，ScanNet和TUM RGB-D等多类数据集上开展了大量实验，部分实验数据如表2。结果表明，DNS-SLAM在定位、重建与语义理解三大核心任务上均表现更优或相当，尤其在噪声深度鲁棒性、重建细节保留及类别级语义分解能力方面优势显著。

2.4　小结

同步定位与地图构建技术已完成从传统几何离散建模到多模态融合优化再到神经隐式连续表征的质变，为路径规划决策提供了兼具几何精度、视觉语义与高保真度的环境模型。表3列举了一些典型SLAM技术的关键特性对比。

但是，目前SLAM技术仍面临在极端环境下多模态融合鲁棒性不足，融合系统定位精度骤降问题；神经隐式SLAM无法平衡精度、实时性和算力，难以满足毫秒级实时导航需求。因此，解决上述挑战仍然是未来需要重点关注的问题。

表3　典型SLAM技术的关键特性对比

系统	年份	传感器套件	EuRoC ATE/m	Loop Closure	Semantic	硬件环境
DS-SLAM	2018	激光LiDAR-惯性-视觉	——	是	是	GPU
Kimera	2020	单目/多目/RGB-D+IMU	——	是	是	CPU
ORB-SLAM3	2021	单目/多目/RGB-D/IMU	0.036	是	否	CPU
R3LIVE	2022	LiDAR+IMU+视觉	0.045	是	是	GPU
FAST-LIVO	2022	LiDAR+IMU+视觉	——	是	否	CPU
PVI-DSO	2023	单目+IMU	0.083	是	否	CPU
UPLP-SLAM	2023	RGB-D	——	是	否	CPU
Eslam	2023	事件相机+IMU+视觉	——	是	是	GPU
DDN-SLAM	2023	RGB-D	——	是	是	GPU
DNS-SLAM	2023	RGB-D	——	是	是	GPU
M-LOAM	2024	LiDAR+IMU+视觉	——	是	否	CPU
R3LIVE++	2024	LiDAR+IMU+视觉	——	否	否	GPU
NGEL-SLAM	2024	RGB	——	是(隐式回环检测)	否	GPU
NIS-SLAM	2024	RGB-D	——	否	是	GPU
MAst3R-Fusion	2025	视觉+IMU+GNSS	——	是	否	GPU

注：EuRoC ATE：欧洲MAV绝对轨迹误差，结果值越小表示定位精度越高；Loop Closure：回环检测，该指标是实现全局定位一致性的核心机制；Semantic：场景语义理解与标注能力。

3　路径规划

路径规划是移动机器人自主导航的核心技术，它基于SLAM构建的环境地图，旨在为移动机器人规划出一条从起点到终点安全且高效率的通行路径。路径规划算法分为两类：传统路径规划算法和生物启发算法，详细分类如图3所示。传统路径规划算法通过确定性搜索或几何计算找到一条从起

点到终点的可行路径。另一类生物启发算法，则通过种群的迭代进化来逼近最优解。与传统的确定性算法不同，生物启发算法提供了一个灵活的框架，可以适应动态的、不可预测的环境，在机器人和自动驾驶车辆等各种应用中具有宝贵的价值。

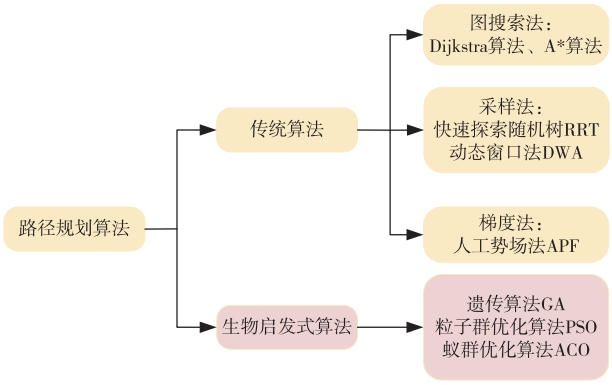


图3 路径规划算法的分类框图

3.1 传统路径规划算法

传统路径规划算法是移动机器人自主导航的基石，这类方法依赖先验的环境模型，并在障碍物与操作约束均为静态或半静态的假设下运行。当环境的拓扑结构、障碍物分布及各类约束能够被预先明确定义时，算法性能达到最优，因此此类算法尤其适用于结构化或可预测的静态环境。传统路径算法按搜索策略分为：图搜索法、采样法和梯度法。

3.1.1 图搜索法：Dijkstra算法、A*算法

（1）Dijkstra算法

Dijkstra算法系统地探索从起始节点开始的所有可能路径，计算到每个相邻节点的累积成本，从而选择总代价最低的路径。Alshammrei等提出一种用于移动机器人的Dijkstra优化变体（MRS）[43]，它在被建模为有向图的环境中生成无碰撞路径。Bhattacharyya和Karmakar将智能能源管理选项与Dijkstra的算法相结合[44]，通过考虑路况和交通参数来降低能源消耗。Garg和Devi提出添加一个自适应惩罚函数[45]，并通过在网络和路线规划等实时应用中缩短计算时间和提升性能来优化寻路，进一步提升算法在实时场景中的适用性。

（2）A*算法

A*算法与Dijkstra算法的不同之处在于前者引入了启发式函数，该函数智能地引导搜索指向目标，A*通过合并实际成本和估计成本，实现了更快的收敛，从而超越了Dijkstra算法。Yonetani等提出神经A*算法[46]，这是一种用于路径规划的数据驱动搜索方法，它对经典的A*算法进行了改进，使其具有可微性，使得神经A可以与卷积编码器连接，产生端到端可训练的神经网络规划器。Fransen和van Eekelen提出一种A*启发式方法[47]，该方法在几何图形中缺乏转向代价，以较少的迭代从而更快找到有效路径。

3.1.2 采样法：快速探索随机树RRT、动态窗口法DWA

（1）快速探索随机树RRT

快速探索随机树RRT是一种基于主动采样的算法，该方法适用于求解具有不同运动约束的路径规划问题，并具有处理多自由度问题的能力。基于RRT算法，Karaman等提出RRT*算法[48]，该算法在保持RRT快速探索空间能力的同时，通过改进树的结构，使路径逐渐收敛到最优解。Wang等提出一种基于改进的RRT*算法路径规划方法[49]，用于解决基于关节结构的移动机器人路径规划问题，并利用矢量化地图和动态约束来实现移动机器人的运动。Li等提出fast RRT*算法[50]，这是一种基于混合采样策略和基于回溯的自适应遗传算法的改进运动规划算法，该方法提升了路径收敛速度和最优性逼近效率，同时适配多自由度移动机器人的路径规划需求。

（2）动态窗口法DWA

动态窗口法DWA是一种利用曲率和速度来控制移动机器人在动态环境中移动的方法。Yang和Teng基于DWA算法提出一种混合算法[51]，通过减少累积转弯角度和平均转弯角度，提高路径平稳性，并为将机器视觉融入复杂环境下的导航奠定了基础。Hou等提出一结合改进的麻雀搜索算法（MISA）和DWA的算法[52]，这种集成使得该方法在静态和动态环境中实现了卓越的路径优化和稳定性。Cao等引入一种新的机器人编队判定方法[53]，优化了动态窗口内的轨迹评价函数，实现了多机器人在未知环境中对动态碍物的高效避障。

3.1.3 梯度法：人工势场法APF

人工势场法APF通过将一般成本函数分解为吸引势和排斥势，在引力和斥力的共同作用下控制移动方向。Luo等提出改进的人工势场算法（IAPF）[54]，通过结合A*算法，成功解决了移动机器人路径规划中的局部极小问题。鉴于传统人工势场法仅依赖目标点的全局吸引子，存在路径不可

预测、易陷入局部极小值的问题，Melchiorre 等在原始算法里引入局部吸引子[55]，将其部署于障碍物周围预设区域，从而提升了障碍物姿态变化的适应性和路径平滑性。

3.2　生物启发算法

生物启发算法的核心在于模拟生物群体行为和物理过程，借助种群的迭代进化来逼近最优解。较于传统算法需要精确建模和确定性求解，启发式算法可以实现在不确定环境中通过自适应或进化式搜索来寻找最优解，这一特性使其在移动机器人、自动驾驶等领域中，实现克服复杂地形约束的同时规避动态障碍物。典型的生物启发算法有遗传算法 GA、粒子群优化算法 PSO 和蚁群优化算法 ACO。

（1）遗传算法 GA

遗传算法 GA 受到自然遗传学和自然选择思想的启发，通过模拟生物的选择、交叉和变异过程来产生最优解。此算法有利于解决具有非线性和多目标的问题，适用于复杂的路径规划任务。GA 算法的软件流程图如图 4 所示。Wang 等在原 GA 算法里通过引入最近邻来提高效率[56]，使得遗传算法路径规划的增强。Ab Wahab 等提出的改进遗传算法使用了基于线性秩次的 CBPRM 方法[57]，显著提高了复杂环境下种群初始化效率和路径优化性能。

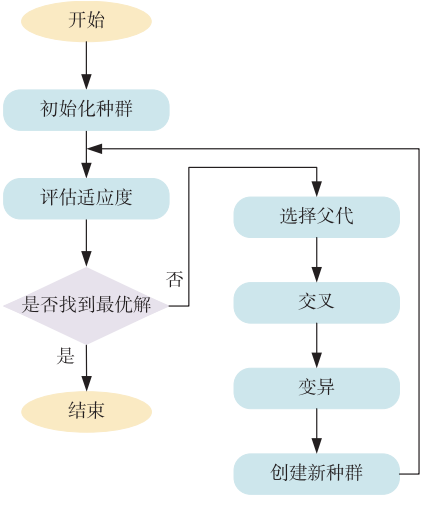


图 4　GA 算法的软件流程图

（2）粒子群优化算法 PSO

粒子群优化算法 PSO 受鸟群和鱼群等群居行为的启发，利用粒子代表潜在解，初始粒子随机分配位置和速度，通过评估适应度更新个体与全局最优，再调整粒子速度和位置，从而平衡解空间探索与已有解优化，此算法主要适用于在高维空间中寻找解。PSO 算法的软件流程图如图 5 所示。Zheng 等通过调整惯性权重参数对粒子群算法进行改进[58]，并将其与人工势场法相结合，显著提高了全局寻优能力，克服了传统方法的局限性。Yang 等提出一种改进的粒子群优化算法（IPSO）[59]，通过增强速度更新和惯性权重，并结合初始化和局部最优避免策略，解决了传统 PSO 算法存在的早熟收敛和局部最优等局限性。Tao 和 Kim 提出一种具有自适应速度调节和随机扰动策略的双种群粒子群算法[60]，以提高路径规划的效率。

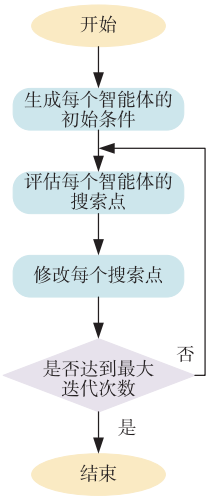


图 5　PSO 算法的软件流程图

（3）蚁群优化算法 ACO

蚁群优化算法 ACO 的灵感源自蚂蚁的群体协作行为，蚂蚁通过信息素痕迹实现群体间的交流，信息素浓度高的路径会吸引更多蚂蚁汇聚，同时，这些蚂蚁又会分泌额外信息素进一步强化这些路径。此方法主要适用于在大型复杂网络中查找最优路径。ACO 算法的软件流程图如图 6 所示。Yang 等在信息素更新和启发式函数方面对传统蚁群算法进行改进[61]，通过与其他方法相结合，蚁群算法具有鲁棒性强、寻优速度快、收敛能力强等特点。针对传统蚂蚁群体搜索效率低以及路径不够平滑而容易落入局部最优的问题，Fang 等提出的结合改进的 Q-Learning 算法[62]，并引入精英蚂蚁搜索策略和自适应信息素波动因子策略。2024 年，Cui 等提出一种多策略自适应蚁群优化算法 MsAACO[63]，通过方向引导机制、自适应启发式函数、确定性状态转移概率规则和非均匀信息素初始

化四种策略相结合，提高了蚁群算法的收敛效率和路径规划性能。

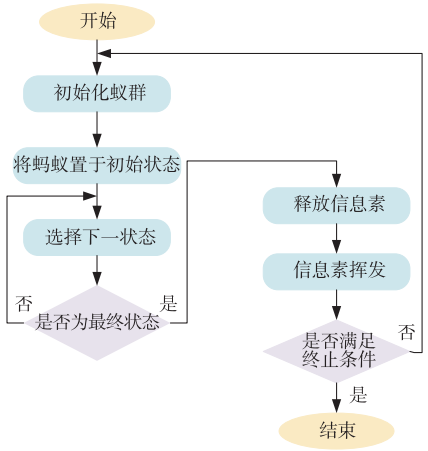


图6 ACO算法的软件流程图

3.3 小结

路径规划作为视觉导航的最终决策环节，已形成传统算法到生物启发优化算法的互补体系，两类算法在不同场景中展现出不同的优势。典型的传统算法在图7环境下的评估如表4所示。

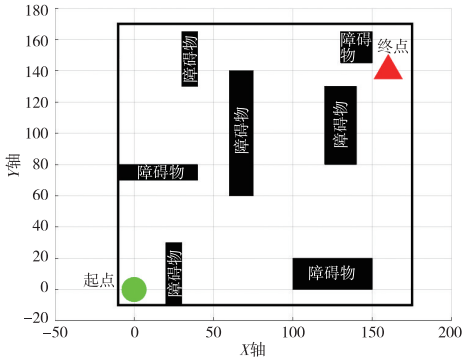


图7 模拟环境配置示意图

表4 典型的传统算法在图7环境下的评估

算法名称	运行时间/s	运行长度/m	转弯次数	平均净空/m	平均曲率	碰撞风险
Dijkstra	205.53	263.99	11	7.451 5	0.022 9	0.000 4
A*	46.38	263.99	18	10.128 1	0.025 4	0.000 1
RRT	675.46	281.53	10	11.530 8	0.023 1	0.047 4
RRT*	641.12	307.14	8	8.825 3	0.016 3	0.097 7
APF	93.92	279.58	11	6.811 9	0.025 4	0.038 8

注：平均净空表示路径与障碍物的平均距离。

当前路径规划技术对动态障碍物的处理仅能依赖实时感知反馈，缺乏轨迹预判能力。在大规模复杂环境中，算法收敛时间过长，难以满足高速导航需求。因此，如何实现动态障碍物的精准预判和快速重规划路径，同时平衡算法精度与收敛效率，是未来研究需重点攻克的难题。

4 总结与展望

本文系统阐述了移动机器人视觉导航在环境感知、同步定位与地图构建和路径规划技术链上的演进与现状。移动机器人视觉导航技术已形成多技术协同、多场景适配的发展态势，各核心模块的技术相互支撑，推动视觉导航从实验室原型走向工业AGV、服务机器人、无人机等多领域工程应用，为机器人自主运行提供了可靠的技术保障。

尽管移动机器人视觉导航技术已取得显著进展，但面对极端环境适应性、动态场景实时性、工程化落地等需求，仍存在诸多待解决的挑战：

（1）如何提升多模态视觉传感器在极端环境下融合的鲁棒性。例如，在强光直射、浓雾、无纹理区域中，跨模态数据常因特征表征不一致而产生严重对齐误差，导致融合系统性能衰退。因此，在保障感知精度的前提下，开发轻量化、自适应的融合框架，仍然是未来需要关注的方向。

（2）如何解决神经隐式SLAM高保真稠密建图能力与系统实时性、内存消耗之间的矛盾。例如，在保障场景重建精度的前提下，神经网络优化过程易产生延迟，且对高性能GPU存在强依赖，这极大限制了其在移动机器人等工业实际场景中的应用。因此，解决这一性能与资源消耗的失衡问题，仍是未来该领域需重点攻克的核心方向。

（3）如何提升动态障碍物轨迹的预判精度，同时强化路径重规划的计算效率。例如，当前路径规划技术对动态障碍物的处理多依赖实时感知反馈，缺乏对运动轨迹的前瞻性预判，难以满足高速运动场景下的安全导航需求。因此，实现动态障碍物运动状态的精准建模与短期预判，同时优化路径重规划的收敛速度，是未来路径规划技术需解决的问题。

（4）如何增强视觉导航系统的跨场景泛化能力。例如，让同一套导航方案既能适配室内结构化环境，又能应对室外动态复杂场景，同时实现感知、建图和决策的协同优化，避免各模块单独优化导致的系统性能脱节，是未来需要解决的问题。

综上所述，本文首先介绍了移动机器人视觉导航系统的核心技术框架，然后详细介绍了视觉传感器模块、SLAM模块及路径规划模块的技术演进、关键方法与典型系统，并对各类技术的性能指标、适用场景进行了分析比较。最后，结合近年来视觉

导航技术的发展趋势，展望未来的发展方向，为移动机器人视觉导航技术的进一步研究与工程化应用提供了参考。

参考文献：

[1] Arafat M Y, Alam M M, Moh S, et al. Vision-based navigation techniques for unmanned aerial vehicles: review and challenges[J]. Drones, 2023, 7(2): 89.

[2] Jin S, Wang X, Meng Q. Spatial memory-augmented visual navigation based on hierarchical deep reinforcement learning in unknown environments[J]. Knowledge-Based Systems, 2024, 285: 111358.

[3] Fan Z, Zhang L, Wang X, et al. LiDAR, IMU, and camera fusion for simultaneous localization and mapping: a systematic review[J]. Artificial Intelligence Review, 2025, 58(6): 174.

[4] Zhang X, Yu R, Ren S. Neural implicit representations for multi-view surface reconstruction: a survey[J]. IEEE Transactions on Visualization and Computer Graphics, 2025, 31(10): 9444-9463.

[5] Ugwoke K C, Nnanna N A, Abdullahi S E. Simulation-based review of classical, heuristic, and metaheuristic path planning algorithms[J]. Scientific Reports, 2025, 15: 12643.

[6] Fields J, Salgian G, Samarasekera S, et al. Monocular structure from motion for near to long ranges[C]//2009 IEEE 12th International Conference on Computer Vision Workshops, 2009: 1702-1709.

[7] Liang H, Ma Z, Zhang Q, et al. Self-supervised object distance estimation using a monocular camera[J]. Sensors, 2022, 22(8): 2936-2936.

[8] Real-Moreno O, Rodríguez-Quiñonez J C, Sergiyenko O, et al. Fast template match algorithm for spatial object detection using a stereo vision system for autonomous navigation[J]. Measurement, 2023, 220: 113299.

[9] Li Y, Li W, Darwish W, et al. Improving plane fitting accuracy with rigorous error models of structured light-based RGB-D sensors[J]. Remote Sensing, 2020, 12(2): 320.

[10] 胡章芳，张　杰，程　亮. 融合特征法与直接法的RGB-D稠密视觉SLAM算法[J]. 半导体光电，2020，41(4)：548-554，559.
Hu Zhangfang, Zhang Jie, Cheng Liang. RGB-D dense visual SLAM algorithm combining feature method and direct method[J]. Semiconductor Optoelectronics, 2020, 41(4): 548-554, 559.

[11] Li X, Xiao Y, Wang B, et al. Automatic targetless LiDAR-camera calibration: a survey[J]. Artificial Intelligence Review, 2023, 56(9): 9949-9987.

[12] Kang J, Kang J. Monocular visual/IMU/GNSS integration system using deep learning-based optical flow for intelligent vehicle localization[J]. Sensors, 2025, 25(19): 6050.

[13] Zhang J, Wang P, Li Y, et al. A deep learning-assisted model-based tightly coupled visual-inertial method for attitude estimation with visual failure conditions[J]. Measurement, 2026, 258: 119442.

[14] Ali M A H, Yusoff N, Abd Razak B, et al. A novel localization algorithm integrated with sensor fusion for position estimation of mobile robot in the road following and round about environment[J]. Engineering Applications of Artificial Intelligence, 2025, 161: 112145.

[15] Gallego G, Delbrück T, Orchard G, et al. Event-based vision: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(1): 154-180.

[16] Gehrig D, Scaramuzza D. Low-latency automotive vision with event cameras[J]. Nature, 2024, 629(8014): 1034-1040.

[17] Furmonas J, Liobe J, Barzdenas V, et al. Analytical review of event-based camera depth estimation methods and systems[J]. Sensors, 2022, 22(3): 1201.

[18] Shariff W, Dilmaghani M S, Kielty P, et al. Event cameras in automotive sensing: a review[J]. IEEE Access, 2024, 12: 51275-51306.

[19] Dold P M, Nadkarni P, et al. Event-based vision in laser welding: an approach for process monitoring[J]. Journal of Laser Applications, 2025, 37: 012040.

[20] Baldini S, Bernardini R, Fusiello A, et al. Measuring vibrations with event cameras[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2024, XLVIII-2/W7-2024: 9-16.

[21] Chen W, Zhou C, Shang G, et al. SLAM overview: from single sensor to heterogeneous fusion[J]. Remote Sensing, 2022, 14(23): 6033.

[22] Smith R C, Cheeseman P. On the representation and estimation of spatial uncertainty[J]. The International Journal of Robotics Research, 1986, 5(4): 56-68.

[23] Mur-Artal R, Montiel J M M, Tardós J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.

[24] Zhang X, Wang W, Qi X, et al. Stereo plane SLAM based on intersecting lines[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021: 6566-6572.

[25] Campos C, Elvira R, Rodríguez 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.

[26] Yang H, Yuan J, Gao Y, et al. UPLP-SLAM: unified point-line-plane feature fusion for RGB-D visual SLAM[J]. Information Fusion, 2023, 96: 51-65.

[27] Zhang J, Singh S. Visual-lidar odometry and mapping: low-drift, robust, and fast[C]//2015 IEEE International Conference on Robotics and Automation (ICRA), 2015: 2174-2181.

[28] Xu Y, Ou Y, Xu T. SLAM of robot based on the fusion of vision and LIDAR[C]//2018 IEEE International Conference on Cyborg and Bionic Systems (CBS), 2018: 121-126.

[29] Zhu Y, Zheng C, Yuan C, et al. CamVox: a low-cost and accurate lidar-assisted visual SLAM system[C]//2021 IEEE International Conference on Robotics and Automation (ICRA), 2021: 5049-5055.

[30] Hulchuk V, Bayer J, Faigl J. Graph-based LiDAR-inertial SLAM enhanced by loosely-coupled visual odometry[C]// 2023 European Conference on Mobile Robots (ECMR). IEEE, 2023: 1-8.

[31] Zheng C, Zhu Q, Xu W, et al. FAST-LIVO: fast and tightly-coupled sparse-direct LiDAR-inertial-visual odometry [C]//2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2022: 4003-4009.

[32] Zheng C, Xu W, Zou Z, et al. FAST-LIVO$_2$: fast, direct LiDAR-inertial-visual odometry[J]. IEEE Transactions on Robotics, 2024, 41: 326-346.

[33] Lin J, Zhang F. R3LIVE: a Robust, Real-time, RGB-colored, LiDAR-Inertial-Visual tightly-coupled state Estimation and mapping package [C]//2022 International Conference on Robotics and Automation (ICRA), 2022: 10672-10678.

[34] Lin J, Zhang F. R^{3}3LIVE: a robust, real-time, radiance reconstruction package with a tightly-coupled LiDAR-inertial-visual state estimator[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 11168-11185.

[35] Zhou Y, Li X, Li S, et al. MASt3R-Fusion: Integrating Feed-Forward Visual Model with IMU, GNSS for High-Functionality SLAM[J]. arXiv preprint arXiv: 2509. 20757, 2025.

[36] Sucar E, Liu S, Ortiz J, et al. iMAP: implicit mapping and positioning in real-time[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 6209-6218.

[37] Zhu Z, Peng S, Larsson V, et al. NICE-SLAM: neural implicit scalable encoding for SLAM[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 12776-12786.

[38] Johari M M, Carta C, Fleuret F. ESLAM: efficient dense SLAM system based on hybrid representation of signed distance fields [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 17408-17419.

[39] Liu J, Li X, Liu Y, et al. RGB-D inertial odometry for a resource-restricted robot in dynamic environments[J]. IEEE Robotics and Automation Letters, 2022, 7(4): 9573-9580.

[40] Li K, Niemeyer M, Navab N, et al. DNS-SLAM: dense neural semantic-informed SLAM [C]//2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2024: 7839-7846.

[41] Mao Y, Yu X, Zhang Z, et al. NGEL-SLAM: neural implicit representation-based global consistent low-latency SLAM system[C]//2024 IEEE International Conference on Robotics and Automation (ICRA), 2024: 6952-6958.

[42] Zhai H, Huang G, Hu Q, et al. NIS-SLAM: neural implicit semantic RGB-D SLAM for 3D consistent scene understanding[J]. IEEE Transactions on Visualization and Computer Graphics, 2024, 30(11): 7129-7139.

[43] Alshammrei S, Boubaker S, Kolsi L. Improved dijkstra algorithm for mobile robot path planning and obstacle avoidance[J]. Computers, Materials and Continua, 2022, 72(3): 5939-5954.

[44] Bhattacharyya S, Karmakar M. Optimal path planning with smart energy management techniques using Dijkstra's algorithm[M]. Singapore: Springer Nature Singapore, 2022: 283-291.

[45] Garg S, Devi B. Shortest path finding using modified Dijkstra's algorithm with adaptive penalty function[C]//2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023: 1-9.

[46] Yonetani R, Taniai T, Barekatain M, et al. Path planning using neural a* search [C]//International Conference on Machine Learning, 2021: 12029-12039.

[47] Fransen K, van Eekelen J. Efficient path planning for automated guided vehicles using A* (Astar) algorithm incorporating turning costs in search heuristic [J]. International Journal of Production Research, 2023, 61(3): 707-725.

[48] Karaman S, Frazzoli E. Sampling-based algorithms for optimal motion planning[J]. The International Journal of Robotics Research, 2011, 30(7): 846-894.

[49] Wang H, Li G, Hou J, et al. A path planning method for underground intelligent vehicles based on an improved RRT* algorithm[J]. Electronics, 2022, 11(3): 294.

[50] Li Q, Wang J, Li H, et al. Fast-RRT*: an improved motion planner for mobile robot in two-dimensional space[J]. IEEJ Transactions on Electrical and Electronic Engineering, 2022, 17(2): 200-208.

[51] Yang H, Teng X. Mobile robot path planning based on enhanced dynamic window approach and improved a algorithm [J]. Journal of Robotics, 2022, 2022: 2183229.

[52] Hou J, Jiang W, Luo Z, et al. Dynamic path planning for mobile robots by integrating improved sparrow search algorithm and dynamic window approach[J]. Actuators, 2024, 13(1): 24.

[53] Cao Y, Mohamad Nor N. An improved dynamic window approach algorithm for dynamic obstacle avoidance in mobile robot formation[J]. Decision Analytics Journal, 2024, 11: 100471.

[54] Luo J, Wang Z X, Pan K L. Reliable path planning algorithm based on improved artificial potential field method[J]. IEEE Access, 2022, 10: 108276-108284.

[55] Melchiorre M, Salamina L, Scimmi L S, et al. Experiments on the artificial potential field with local attractors for mobile robot navigation[J]. Robotics, 2023, 12(3): 81.

[56] Wang J. Intelligent path planning of mobile robot based on genetic algorithm[J]. Journal of Physics: Conference Series, 2023, 2547(1): 012001.

[57] Ab Wahab M N, Nazir A, Khalil A, et al. Improved genetic algorithm for mobile robot path planning in static environments [J]. Expert Systems with Applications, 2024, 249: 123762.

[58] Zheng L, Yu W, Li G, et al. Particle swarm algorithm path-planning method for mobile robots based on artificial potential fields[J]. Sensors, 2023, 23(13): 6082.

[59] Yang Z, Li N, Zhang Y, et al. Mobile robot path planning based on improved particle swarm optimization and improved dynamic window approach[J]. Journal of Robotics, 2023, 2023(1): 6619841.

[60] Tao B, Kim J H. Mobile robot path planning based on bi-population particle swarm optimization with random perturbation strategy[J]. Journal of King Saud University-Computer and Information Sciences, 2024, 36(2): 101974.

[61] Yang L, Fu L, Li P, et al. An effective dynamic path planning approach for mobile robots based on ant colony fusion dynamic windows[J]. Machines, 2022, 10(1): 50.

[62] Fang W, Liao Z, Bai Y. Improved ACO algorithm fused with improved Q-Learning algorithm for Bessel curve global path planning of search and rescue robots[J]. Robotics and Autonomous Systems, 2024, 182: 104822.

[63] Cui J, Wu L, Huang X, et al. Multi-strategy adaptable ant colony optimization algorithm and its application in robot path planning[J]. Knowledge-Based Systems, 2024, 288: 111459.

作者简介：

胡章芳(1969—)，女，重庆市人，教授，硕士生导师，主要研究领域为数字图像处理、微机电系统及智能信号处理；

张　薇(2003—)，女，河南濮阳人，硕士研究生，主要研究方向为基于VSLAM的导航技术。