 DOI: https://doi.org/10.12349/iser.v7i4.10048

Methods for 3D Environment Perception and Real-time Visualization for Autonomous Navigation of Unmanned Aerial Vehicles

Yi Wang Bowen Yang Min Zhong Guanyu Jia Xuan Wang*

Sichuan Communication Research and Planning Design Co., Ltd., Chengdu, Sichuan, 610000, China

Abstract

Focusing on the characteristics of unmanned aerial vehicles such as low payload and high real-time requirements, this paper conducts research on building a 3D environment visual perception framework based on visual multi-object recognition and semantic cognition. By introducing an attention mechanism model with bird's-eye view perspective and a lightweight expression mechanism, the accurate representation of target spatial positions and dynamic changes in complex scenes is achieved. Through the construction of a hierarchical rendering and data compression mechanism, the real-time computing efficiency is further improved. The research results show that this method has good real-time performance and system scalability while ensuring accuracy, and can provide technical support for applications such as unmanned aerial vehicle inspection and disaster monitoring.

Keywords

autonomous navigation of unmanned aerial vehicles; 3D environment perception; multi-modal fusion; BEV representation

面向无人机自主导航的三维环境感知与实时可视化方法

王燚 杨博文 钟敏 贾桄宇 王旋*

四川通信科研规划设计有限责任公司，中国 · 四川 成都 610000

摘 要

围绕无人机低载荷、实时性要求高等特点，本文围绕基于视觉多目标识别及语义认知构建三维环境视觉感知框架开展研究，通过引入鸟瞰视角的注意力机制模型与轻量化表达机制，实现复杂场景下目标空间位置与动态变化的准确表征。并通过构建分层渲染与数据压缩机制，进一步提升实时运算效率。研究结果表明，该方法在保证精度的同时具备良好的实时性能与系统可扩展性，可为无人机巡检、灾害监测等应用提供技术支撑。

关键词

无人机自主导航；三维环境感知；多模态融合；BEV表示

1 引言

随着无人机技术在巡检、应急救援与环境监测等领域的广泛应用，其自主导航能力逐渐成为研究重点。复杂环境中存在动态障碍物、多尺度结构以及光照变化等问题，对环境感知提出了更高要求。传统二维视觉方法难以提供完整空间信息，三维环境感知技术通过融合多源传感器数据，将环境信息以空间结构形式表达，为路径规划与避障提供可靠依据。在现有研究中，双目视觉传感器具有一定的轻量化深度估计能力，但仅支持目标某个方向的环境感知，全向环境感知能力有限。基于此，本文围绕无人机应用场景，构建三维环境感知与实时可视化一体化方法，重点解决感知精度、计算效率与系统集成之间的平衡问题。

【作者简介】王燚（1975–），男，中国四川成都人，博士，高级工程师，从事信息与通信工程研究。

【通讯作者】王旋（1984–），男，中国四川成都人，博士，高级工程师，从事无线通信与信息技术，智能网络系统研究。

2 无人机三维视觉环境感知技术体系构建

已有的双目视觉深度估计通过两个摄像头的视觉差做景深估计，并基于此将多个估计值进行拼接融合生成空间位置感知结果，在融合处理单个感知源时因为单个数据源的误差累计导致整体空间感知误差放大。为此，业界提出一种采用鸟瞰图表示与注意力机制相结合的联合空间感知处理技术架构，将多个原始景深图生成伪云图，按照空间映射形成全局鸟瞰伪像点云进行降采样与滤波处理，从而减少数据规模并保持关键结构信息；通过体素网格规则划分空间结构，使数据适配三维卷积网络；鸟瞰图则将生成的三维信息投影

到二维平面，降低计算复杂度。在无人机应用场景中，为满足实时性要求，需要进一步对模型进行轻量化设计。通过减少用于特征提取的网络层数、采用稀疏卷积以及引入量化技术，可以在保证精度的同时降低注意力权值参数计算开销，使模型适配嵌入式平台运行。

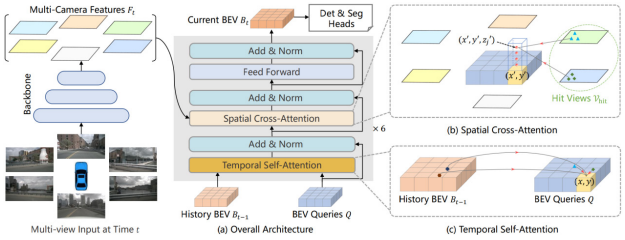


图 1. 基于 BEV 三维视觉空间决策技术架构

3 三维目标检测与动态环境建模方法

3.1 三维目标检测模型设计

三维目标检测的关键在于从多源融合数据中准确恢复目标的空间几何属性，并以统一参数形式进行表达。通常以七维参数描述目标状态，包括三维位置坐标、长宽高以及朝向角度，使检测结果能够直接服务于空间决策与路径规划。在模型构建方面，深度学习方法已成为主流路径，通过构建端到端网络，实现特征提取与边界回归的统一优化，以提升对小目标与远距离目标的识别能力。同时，通过结合鸟瞰视角表达，可进一步强化空间结构信息，使模型在复杂环境中仍能保持较高的检测稳定性与精度。

3.2 多模态融合检测策略

复杂环境下单一传感器难以全面获取有效信息，图像数据在纹理表达方面具有优势，而点云数据在空间几何描述方面更为精确，多模态融合因此成为提升检测性能的重要路径。在融合过程中，通过构建统一特征空间，实现不同模态信息的协同表达，使语义信息与空间信息形成互补关系。在无人机动态场景中，特征级融合更具适应性，可在不同环境条件下保持较高鲁棒性。进一步结合注意力机制与跨模态对齐方法，可增强关键区域信息权重，使系统在复杂背景中仍能准确识别目标，提升整体感知能力。

3.3 动态目标跟踪与轨迹预测

无人机在实际运行过程中面临持续变化的动态环境，仅依赖单帧检测结果难以支撑稳定导航。通过引入多目标跟踪机制，可对目标进行时间维度上的连续建模，实现状态更新与身份保持。在跟踪过程中，通常结合检测结果与运动模型，对目标位置进行滤波估计，从而减少检测噪声带来的波动影响。同时，通过建立目标关联策略，实现不同帧之间的目标匹配，保证轨迹连续性。轨迹预测是在跟踪基础上的进一步扩展，通过分析历史运动数据，结合运动学模型或学习型预测方法，对目标未来位置进行估计。

4 实时三维可视化方法与系统实现

4.1 多层级可视化表达机制

三维环境数据具有体量大、结构复杂及语义层次多样等特征，单一展示方式难以全面反映环境信息。基于此，多层级可视化表达机制通过对数据进行结构化分解，将几何信息、语义信息与决策信息分别映射至不同层级，实现信息的有序呈现。在基础层，系统以点云或网格形式展示环境的空间结构，突出地形轮廓与障碍物分布，为整体认知提供几何基础。在语义层，通过融合目标检测与分类结果，对关键目标进行标注，使环境信息具备可理解性。在决策层，将路径规划结果、避障策略等信息叠加于空间结构之上，体现系统运行状态与决策逻辑。各层之间通过统一坐标系进行关联，既保证信息表达的一致性，又避免数据冗余。该机制能够在复杂场景中实现由“结构—语义—决策”的逐级递进表达，使系统输出更加清晰、直观，有助于提升人机交互效率与系统调试能力。

4.2 高效渲染与数据压缩策略

实时可视化对系统计算能力与数据处理效率提出了较高要求，在多视觉多目标联合处理场景下，处理过程易出现延迟与资源占用过高的问题。为此，需要从数据组织与计算流程两个方面进行优化。在数据层面，通过对空间结构进行分块处理与层次化管理，仅对注意力特征表达数据进行处理，实现对关键点提取与数据进行压缩，在保证主要结构特征不丢失的前提下降低数据规模，在无人机实时导航的应用中能够显著提升系统响应速度，使可视化过程在高频数据更新条件下仍保持稳定流畅。

4.3 BEV 可视化与空间决策支持

鸟瞰视角（BEV）在三维环境表达中具有结构简洁、信息集中等优势，能够有效降低空间理解难度。通过将三维点云数据投影至二维平面坐标系，可将复杂空间关系转化为平面分布形式，从而直观呈现目标位置、运动轨迹与环境结构。在该视角下，目标检测结果通常以边界框形式进行表达，并结合类别信息与置信度进行标注，使环境态势更加明确。同时，通过叠加路径规划结果与安全边界信息，可以构建面向决策的可视化界面，为无人机导航提供直观依据。在工程实践中，BEV 可视化不仅用于实时监控，还可作为系统调试与性能评估的重要工具，通过对历史数据进行回放与分析，能够识别系统在复杂场景中的表现差异。相较于传统三维渲染方式，BEV 表达在计算复杂度与信息表达效率之间取得较好平衡，为无人机自主导航系统提供了可靠的空间决策支持手段。

5 算法验证与性能优化

在无人机自主导航的实测过程中，采用了 NVIDIA Jetson Orin nano 作为无人机载三维环境感知计算平台，在

Ubuntu 操作系统及 TensorRT 优化后的 CUDA 系统环境下运行经文中方案改进后的 BEV 算法，采用 KTTI 上提供的标准 6 幅三维车载数据集进行训练，并对训练结果进行测试，测试结果表明 BEV 方案能够很好的适应无人机自主导航所需的轻量化视觉三维环境感知应用场景。

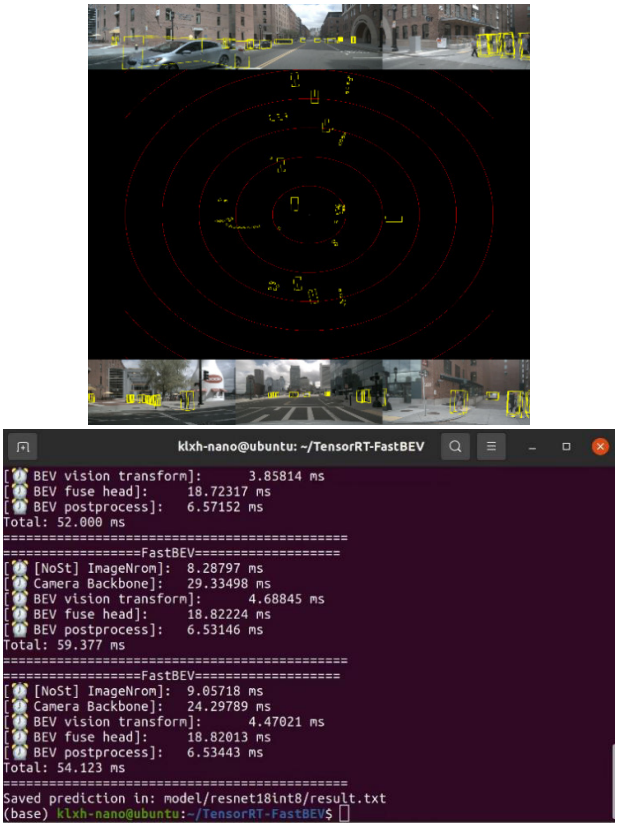


图 2. 基于 NVIDIA Jetson OrinNano 算力平台的性能测试

6 结语

围绕无人机自主导航需求，本文构建了三维环境感知与实时可视化一体化方法，从多源数据融合、目标检测、动态建模及可视化表达等方面进行了系统分析与设计。通过引入深度学习模型与 BEV 表达机制，实现了对复杂环境的高效感知与直观表达。在系统实现层面，通过嵌入式优化与模块化架构设计，提升了整体运行效率与可扩展性。未来研究可进一步结合 Transformer 结构与自监督学习方法，提升模型对复杂场景的适应能力，并探索更高效的数据融合策略，以实现更高精度与更强鲁棒性的无人机自主导航系统。

参考文献

[1] 张政.无人机自主导航视觉SLAM和路径规划研究[D].西安理工大学,2025.

[2] 杨轶.拒止环境下基于双目相机的无人机自主导航系统设计[D].南京理工大学,2023.

[3] 肖斌.城市低空环境下的无人机三维环境感知与避障方法研究[D].东北大学,2017.

[4] AgiBot World Colosseo: A Large-scale Manipulation Platform for Scalable and Intelligent Embodied Systems[J]. https://arxiv.org/abs/2503. 06669

[5] End-to-End Autonomous Driving: Challenges and Frontiers[J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, Vol.46(12), 2024

[6] BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.47(3),2025.