

基于LV-DOT算法增强传感器融合与追踪精度研究

郑 妮, 杨 旗*

沈阳理工大学机械工程学院, 辽宁 沈阳

收稿日期: 2025年12月7日; 录用日期: 2025年12月30日; 发布日期: 2026年1月7日

摘要

为了实现对动态目标进行高精度检测与稳定追踪, 本文将现有的基于LiDAR和视觉传感器相融合的LV-DOT框架进行改进优化, 核心创新涵盖两条: 一是提出基于形态跟大小一致性的跨模态检测结果校验举措, 在数据融合阶段审视LiDAR点云与视觉检测结果的几何形态和尺度特性一致性, 减小错误检测和遗漏检测的发生概率; 二是设置基于运动方向的误差补救机制, 采用动态校验的方式来确认障碍物运动轨迹连续性与方向一致, 抑制传感器噪声、环境遮挡或瞬时观测偏差引发的轨迹漂移, 提升追踪的稳健水平。通过实验结果表明, 所提方法在检测精度、召回率及实时处理效率方面均有显著改善, 此优化框架不仅增强了多传感器融合在动态场景下的稳健性与稳定性, 也给移动机器人检测、追踪动态障碍物提供了可行的技术办法。

关键词

多传感器融合, 目标跟踪, 误差修正, 跨模态验证

Research on Enhancing Sensor Fusion and Tracking Accuracy Based on the LV-DOT Algorithm

Ni Zheng, Qi Yang*

School of Mechanical Engineering, Shenyang Ligong University, Shenyang Liaoning

Received: December 7, 2025; accepted: December 30, 2025; published: January 7, 2026

Abstract

To achieve high-precision detection and stable tracking of dynamic targets, this paper improves the

*通讯作者。

existing LiDAR-Vision ss-modal detection validation method based on morphological and size consistency is proposed, which examines the geometric and scale agreement between LiDAR point clouds and visual detection results during the data fusion stage, thereby reducing false and missed detections; second, a motion-direction-based error remediation mechanism is introduced, which employs dynamic verification to ensure the continuity and directional consistency of obstacle motion trajectories, mitigating trajectory drift caused by sensor noise, environmental occlusion, or transient observation deviations, thus enhancing tracking robustness. Experimental results demonstrate that the proposed approach significantly improves detection accuracy, recall rate, and real-time processing efficiency. The optimized framework not only strengthens the robustness and stability of multi-sensor fusion in dynamic scenarios, but also provides a viable technical solution for mobile robots in detecting and tracking dynamic obstacles.

Keywords

Multi-Sensor Fusion, Target Tracking, Error Correction, Cross-Modal Verification

Copyright © 2026 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

动态障碍物感知与追踪是实现移动机器人自主安全导航的核心问题，其对环境感知能力，尤其是动态障碍物感知能力的要求日益苛刻。单一传感器受限于其物理特性，难以在复杂的真实世界场景中提供全面、可靠的感知信息。因此，融合 LiDAR 的精确几何感知能力与摄像头的丰富纹理语义信息的多传感器融合技术，已成为环境感知领域的主流解决方案。LV-DOT 作为一种典型的 LiDAR-视觉融合追踪框架 [1]，展现了良好的应用前景，但其在实际部署中仍存在精度与鲁棒性的瓶颈。本文深入分析了 LV-DOT 框架的固有缺陷，并从运动规律与几何一致性两个维度提出优化方案，以期推动相关技术向更高可靠性发展。

2. 问题研究与分析

2.1. 单一传感器的局限性

动态障碍物的感知是自主机器人在复杂环境中实现安全导航的关键[2]。传统的 LiDAR 或摄像头传感器各有其优缺点。LiDAR 传感器广泛的 360 度探测范围，在距离测量方面表现优秀，能提供精确的深度信息，受光照条件影响小，但在物体分类和纹理识别上存在不足，且在动态环境中数据稀疏，更新频率较低，难以检测小物体或远距离物体；而摄像头则提供丰富的视觉特征(纹理、颜色)和高频数据，能够有效识别静态物体，但深度感知能力受限，受光照条件影响较大，背景杂乱时可能产生噪[3]。

因此，传感器融合方法对于提高障碍物检测的鲁棒性和精度至关重要，尤其是在动态和复杂的环境下。

2.2. 传感器融合的挑战

传统的 LiDAR 与视觉的融合方法主要借助交并比和置信度加权匹配，此种方法在复杂场景时面临显著局限[4]：当有一个传感器的置信度显著比另一个传感器高时，融合结果会过度依赖高置信度的数据起

始源，要是该数据存在瞬间误差，就容易造成虚检现象；当环境造成双模态数据质量同时有所下降，低置信度会让系统变得过于保守，引起漏检问题或目标特征模糊化。是没有有效评估数据质量和逻辑一致性[5]，本文根据此条件提出双重优化途径：引入运动方向契合性校验，通过核对预测轨迹和观测数据的运动矢量，删掉不符合运动规律的异常检测情形；采用几何形态与尺度的相似性综合评定，保障关联目标在物理维度里合理契合。此策略能有效达成过拟合与欠拟合的平衡，增进复杂动态场景下检测的抗干扰性。

3. 提出 LV-DOT 算法优化

3.1. 融合模块中形态与大小一致性检查

因 LiDAR 和视觉传感器的特性不一样，其生成的障碍物边界框，频繁会展现出尺寸上的不一致现象，例如：LiDAR 传感器大多时候提供的是稀疏的点云数据，引起在远距离时障碍物尺寸估计结果不大，视觉传感器虽能提供相对密集的数据，但由于透视失真以及遮挡效应，或许会高估物体的大小，这些不一致情况或许会造成误报或者漏报[6]，本文在 LiDAR-视觉融合模块里采用了形态与大小一致性检查，保障在数据融合阶段，LiDAR 与视觉所生成的障碍物检测结果在形态和大小方面保持一致。

核心思想：计算每个障碍物的边界框尺寸(宽度、高度和深度或 2D 的宽度和高度)，并进行一致性检查。如果 LiDAR 和视觉生成的尺寸或形态差异超过阈值，则认为是错误的检测结果，丢弃或修正。

数学公式：

1) 尺寸一致性检查：

定义 LiDAR 检测框的尺寸为 (W_L, H_L) ，视觉检测框的尺寸为 (W_V, H_V) 。计算宽度和高度的相对误差：

$$\epsilon_W = \frac{|W_L - W_V|}{\max(W_L, W_V)} \quad (1)$$

$$\epsilon_H = \frac{|H_L - H_V|}{\max(H_L, H_V)} \quad (2)$$

一致性判断：若 $\max(\epsilon_W, \epsilon_H) > \alpha$ ，则视为无效匹配，本文设置尺寸一致性阈值 $\alpha = 0.3$

2) IoU (交并比)一致性：

交并比(IoU)是衡量两个检测框空间重叠度的常用指标。计算 LiDAR 检测框 B_L 和视觉检测框 B_V 的 IoU

$$\text{IoU} = \frac{\text{Area}(B_L \cap B_V)}{\text{Area}(B_L \cup B_V)} \quad (3)$$

一致性判断： $\text{IoU} < \beta$ 表明两个检测框空间位置差异过大，不宜进行融合。本文设置 IoU 一致性阈值 $\beta = 0.5$ 。

3.2. 追踪模块中基于运动方向的误差修正

动态障碍物的运动预测是追踪模块里的一个关键问题，精准的运动预测对障碍物追踪十分关键，目前采用的方法主要依靠卡尔曼滤波等基础追踪算法来估计障碍物的位置、速度以及加速度[6]。但这些方法在验证运动预测与实际检测结果的一致性方面存在不足，为了应对运动预测方面难题，在追踪模块中引入了依据运动方向的误差修正机制，该机制保障检测到的障碍物运动方向与预期运动轨迹相匹配。

核心思想：使用卡尔曼滤波预测障碍物的运动轨迹，并估算其位置、速度和加速度。比较 LiDAR 和

视觉数据的速度和运动方向，如果两者的检测结果偏离预期的运动轨迹，说明存在融合误差，应该丢弃该检测结果。

数学公式：

1) 卡尔曼滤波预测：

状态向量：定义在时刻 t 的状态向量为 $\mathbf{x}_t = [p_x, p_y, v_x, v_y]^T$ ，其中 p_x, p_y 表示障碍物在二维平面上的位置， v_x, v_y 表示其速度。

状态预测方程：

$$\hat{\mathbf{x}}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{w}_t \quad (4)$$

其中， $\hat{\mathbf{x}}_t$ 是时刻 t 的先验状态估计， \mathbf{F} 是状态转移矩阵(对于匀速模型， $\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$)， Δt 为时间间隔)， \mathbf{w}_t 是过程噪声，假设为高斯白噪声。

观测方程：

$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{v}_t \quad (5)$$

其中， \mathbf{z}_t 是传感器观测值(如直接测量的位置)， \mathbf{H} 是观测矩阵(通常 $\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$)， \mathbf{v}_t 是观测噪声。

2) 运动方向一致性：

分别从 LiDAR 和视觉的观测数据中计算出障碍物在时刻 t 的瞬时速度矢量 $\mathbf{v}_{\text{lidar}}$ 和 $\mathbf{v}_{\text{vision}}$ ；从卡尔曼滤波器中获取当前追踪轨迹的预测速度矢量 \mathbf{v}_{pred} ；计算各观测速度与预测速度的方向夹角余弦值，以评估一致性：

$$\cos(\theta_{\text{lidar}}) = \frac{\mathbf{v}_{\text{pred}} \cdot \mathbf{v}_{\text{lidar}}}{\|\mathbf{v}_{\text{pred}}\| \|\mathbf{v}_{\text{lidar}}\|} \quad (6)$$

$$\cos(\theta_{\text{vision}}) = \frac{\mathbf{v}_{\text{pred}} \cdot \mathbf{v}_{\text{vision}}}{\|\mathbf{v}_{\text{pred}}\| \|\mathbf{v}_{\text{vision}}\|} \quad (7)$$

一致性判断：若 $\cos(\theta_{\text{lidar}}) < \gamma$ 或 $\cos(\theta_{\text{vision}}) < \gamma$ ，则认为该观测数据方向偏离预期轨迹。本文设置运动方向一致性阈值 $\gamma = \cos(\theta_{\min}) = 0.866$ (对应约 30 度夹角)。

在 Matlab 仿真环境中运行修改前后的 LV-DOT 算法，结果如图 1 与图 2 表示，从两幅图中可以看到改进后(LV-DOTG)的融合算法跟踪轨迹与真实轨迹更加吻合与贴近。改进优化 LV-DOT 算法的思维流程如图 3 所示。

4. 实验结果分析

4.1. 实验设置

本文在多种室内环境下对优化后的 LV-DOTG 算法进行了测试，使用一架配备 Intel Realsense D415 RGB-D 相机和 Livox Mid-360 LiDAR 传感器的移动机器人小车。机器人被放置在具有静态和动态障碍物的复杂环境中，进行实时障碍物检测与追踪检测效果如图 4 所示。从图 4 可以看到改进算法(LV-DOTG)在复杂 S 弯及曲线行走过程中能够持续保持对移动目标的检测与跟踪，算法表现效果较为良好。

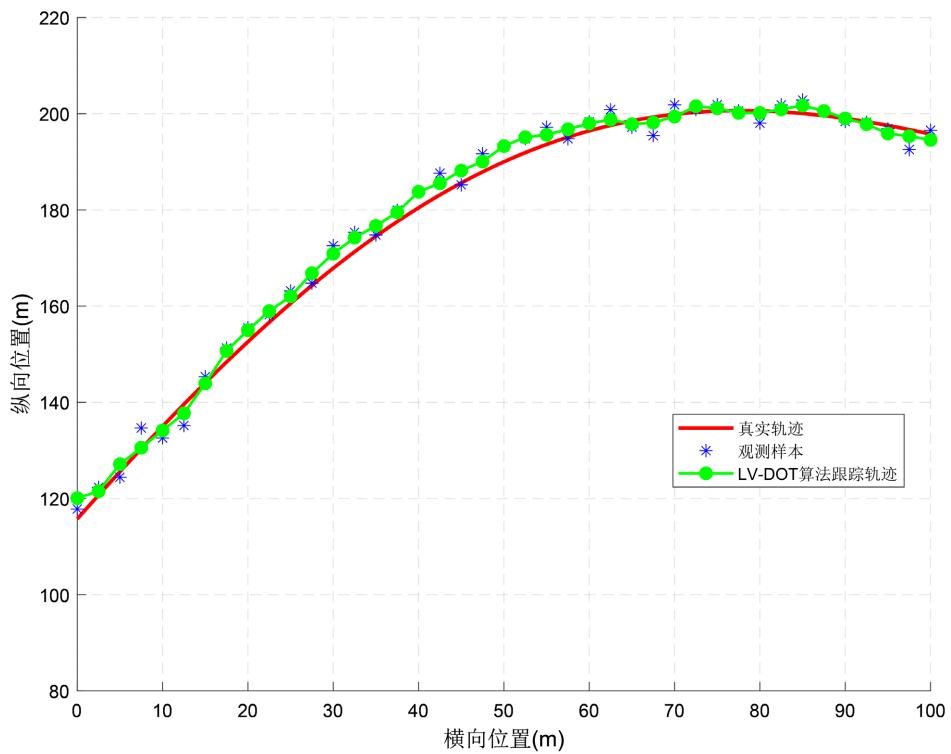


Figure 1. LV-DOT trajectory

图 1. LV-DOT 跟踪轨迹图

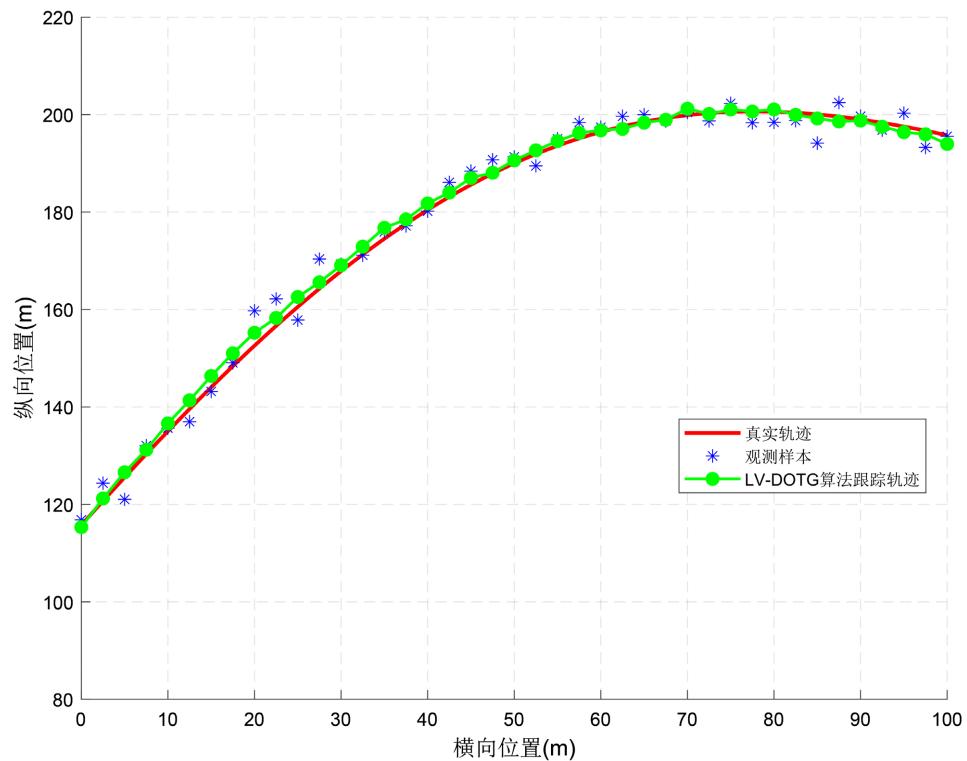


Figure 2. LV-DOTG trajectory

图 2. LV-DOTG 跟踪轨迹图

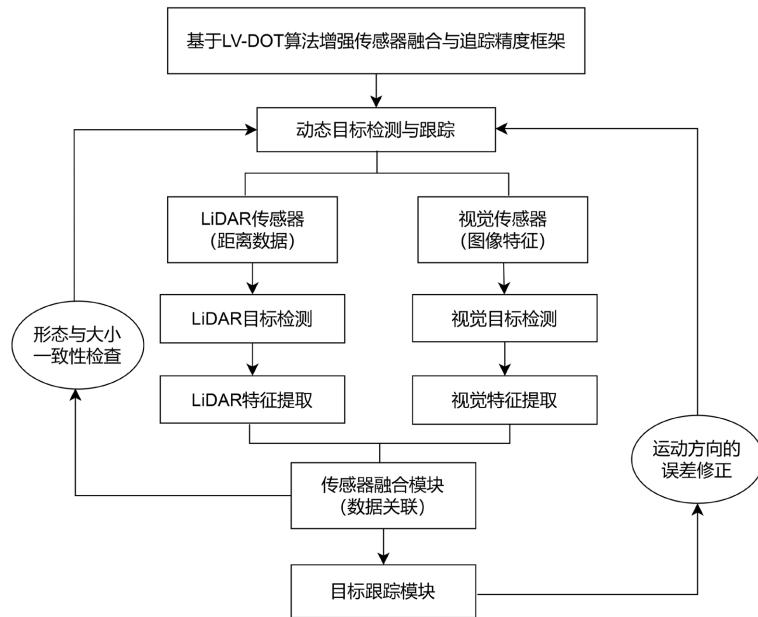


Figure 3. Optimized LV-DOT flowchart
图 3. 优化 LV-DOT 算法框架流程图

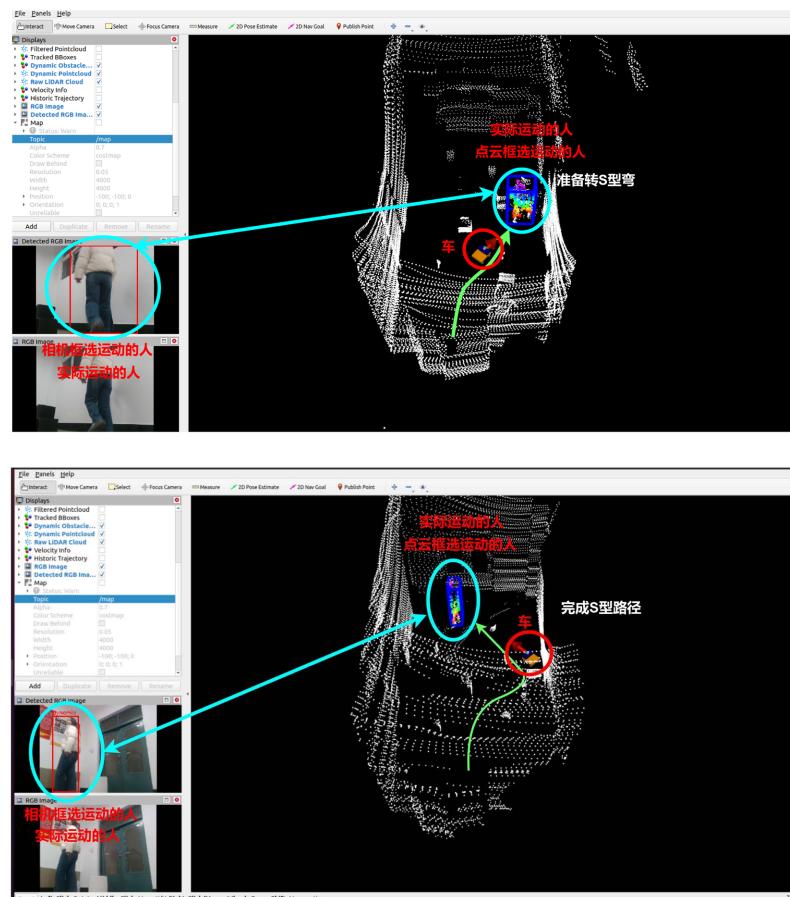


Figure 4. LV-DOTG algorithm test
图 4. LV-DOTG 算法实验检测图

4.2. 评估指标

在实现机器人目标检测与跟踪工作任务中，衡量其检测与追踪的精准程度通常有如下这些基本原则 [7]：

1) 交并比(IoU): 交并比(IoU)即两个相交矩形框的面积与二维目标检测中两个相交矩形框合并的面积之比，以及两个相交矩形框的体积与三维目标检测中两个相交矩形框合并的体积之比。交并比越大，融合就越精确[8]。

2) 精度(Precision): 精确率为正确预测为正例的数量占所有预测为正例的数量的比例，通常用以下公式表示：

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (8)$$

3) 召回率(Recall): 召回率表示正确预测为正例的数量占所有实际正例的比例，用以下公式表示：

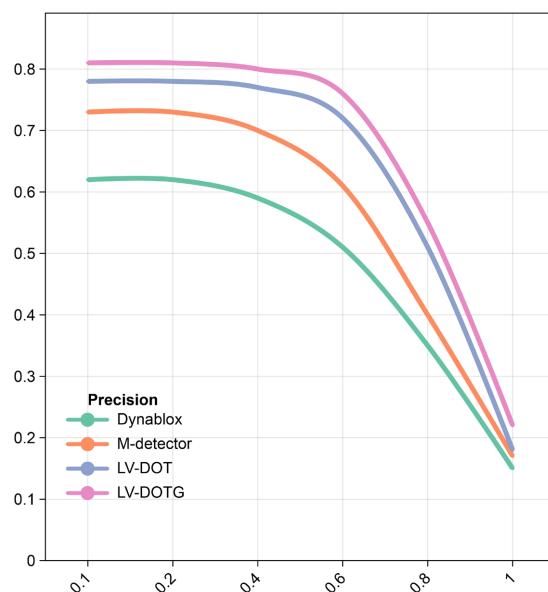
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (9)$$

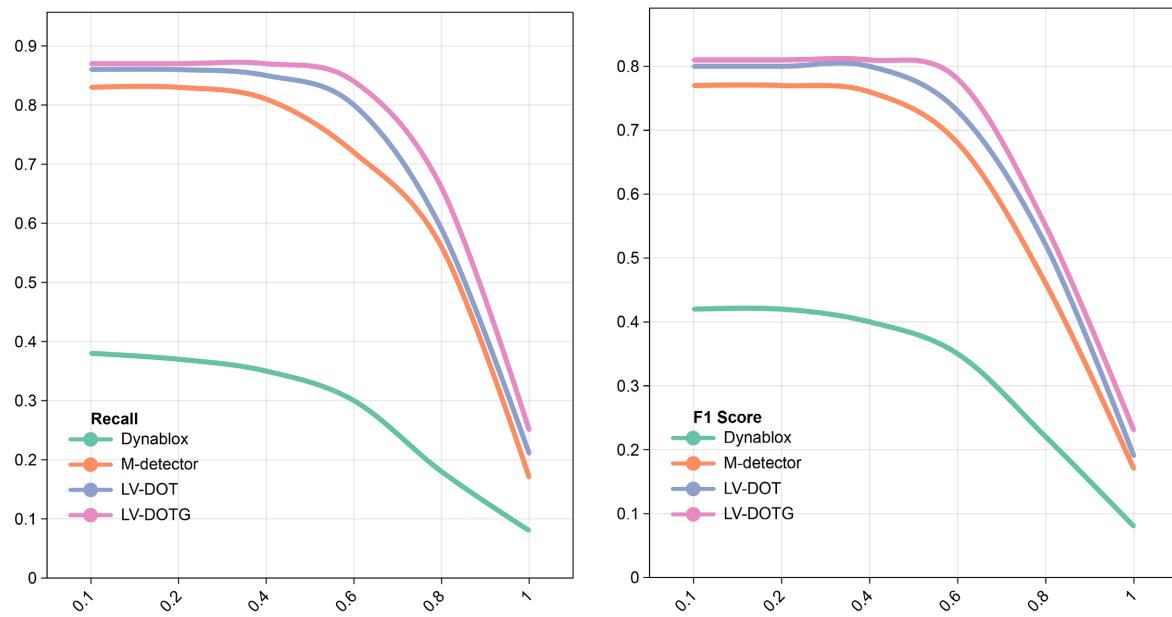
其中 TP 是被正确预测为正例的数量，FP 是被错误预测为正例的负例数量，FN 是被错误预测为负例的实际正例数量。

4) F1 分数: F1 分数是精确率和召回率的调和平均数，它是一个综合指标，用于同时衡量一个模型的“查得准”和“查得全”的能力，用以下公式表示：

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

实验中将 Dynablock、M-detector 和 LV-DOT 算法选为基准方法。针对每种算法在实验过程完成的效果，测量并绘制了从 0.1 到 1 范围内各交并比阈值对应的精度、召回率和 F1 分数曲线。如图 5 所示，从图中可以看出改进后的方法(LV-DOTG)在所有交并比阈值下均保持最高 F1 分数，显著优于基准方法。各方法的平均位置误差均约为 0.1 米。改进后的算法与原 LV-DOT 方法相比，两者在召回率方面均达到极高且相近的水平，表明动态障碍物检测几乎不会出现漏检情况。不过，原 LV-DOT 方法精度略低，实验显示它偶尔会将静止障碍物误判为动态障碍物。相比之下，改进后的 LV-DOTG 算法显著提升了动态障碍物识别的准确率。



**Figure 5.** IoU threshold evaluation curve**图 5.** IoU 阈值评估曲线

在实验过程中记录了算法在机器人机载计算机上白天与晚上的运行时间, 如下表 1 所示该表给出了 LV-DOT 算法改进前后的平均运行时间和标准差。改进前 LV-DOT 算法在白天与晚上的平均运行时间分别为 0.86 毫秒和 1.12 毫秒。而改进后的 LV-DOT 算法具有极高的运行效率, 在白天与晚上的平均运行时间分别仅为 0.43 毫秒和 0.69 毫秒, 整个系统在实际机器人上实现了实时性能, 满足了现实环境中动态障碍物感知与导航的关键需求。

Table 1. Algorithm tracking results: Comparisons
表 1. 算法跟踪结果对比

方法	白天		夜晚	
	平均值	标准差	平均值	标准差
LV-DOT	0.86 ms	0.48 ms	1.12 ms	0.81 ms
LV-DOTG	0.43 ms	0.16 ms	0.69 ms	0.42 ms

4.3. 讨论

基于运动方向的误差修正模块为本次算法优化的核心创新点之一, 常规手段在应对动态障碍物之际, 老是由于运动预测不精确, 引发跟踪中断或丢失, 该优化算法通过深度验证目标的连续移动方向, 打造了更为稳健的运动轨迹预测模型。这个模块实时分析相邻帧彼此之间障碍物的位移矢量, 并跟历史运动趋势开展一致性比对, 进而能有效区分真实动态障碍物与因传感器噪声以及光照变化产生的虚假目标, 这一改进明显提升了系统对快速移动对象的持续跟踪能力, 致使整体召回率实现显著提升[9]。

形态与大小一致性的检查从空间维度为多传感器的融合给予了关键保障, 在原始流程阶段, 源于激光雷达与摄像头的原始检测框, 常因感知原理有别, 存在细微的形状及尺度差异, 直接将其融合, 易产生畸变或错误的包围盒[10], 此算法采用的形态及大小一致性核查办法, 会于融合前对源自不同传感器的候选边界框进行严苛的几何属性校验, 涉及长宽的比例数值、体积的比例关系以及空间重叠度的指标,

此步骤尤其适合解决传感器初步感知不一致的情形，该机制有力提高了复杂场景下障碍物检测的完整性及准确性，减小了早期融合错误所引发的后续感知链失效风险。

4.4. 局限性分析与未来展望

尽管所提方法取得了良好效果，但仍存在以下局限性，有待未来研究改进：

对高机动与非常规目标适应性有限：基于固定阈值的运动方向校验，在目标进行合法急转弯等剧烈机动时可能导致误拒；同时，依赖规则边界框的形态检查，对长条形、非刚性或严重遮挡目标的泛化能力不足。

密集场景关联与极端条件鲁棒性有待提升：当前数据关联策略在目标高度密集时易产生歧义；且系统对传感器单模态极端失效(如摄像头致盲、LiDAR 大雨衰减)或标定严重漂移的容错能力较弱。

未来工作将重点探索：采用自适应阈值或更复杂运动模型(如 CTRV)以提升机动适应性，并引入分割掩码等特征增强形状泛化能力；融合外观 Re-ID 特征与全局优化关联以改善密集场景性能；设计传感器故障检测与降级至鲁棒单模态追踪的策略，提升系统整体可靠性。

5. 结语

本文提出了一种针对 LV-DOT 算法实施优化的方法 LV-DOTG，通过引入基于运动方向的误差优化和形态跟大小的一致性审查，明显提升了动态障碍物检测跟追踪的精准程度与稳固性，实验结果说明，优化后的算法在复杂环境里展现出更高的准确性与实时性，未来工作会对传感器融合技术做进一步改进，探究减少误报问题和增强追踪稳定性的手段。

参考文献

- [1] Xu, Z., Shen, H., Han, X., et al. (2025) LV-DOT: LiDAR-Visual Dynamic Obstacle Detection and Tracking for Autonomous Robot Navigation. <https://arxiv.org/abs/2502.20607>
- [2] Choi, J.D. and Kim, M.Y. (2023) A Sensor Fusion System with Thermal Infrared Camera and LiDAR for Autonomous Vehicles and Deep Learning Based Object Detection. *ICT Express*, **9**, 222-227. <https://doi.org/10.1016/j.icte.2021.12.016>
- [3] Zhou, Y., Sun, P., Zhang, Y., et al. (2020) End-to-End Multi-View Fusion for 3d Object Detection in Lidar Point Clouds. 2020 Conference on Robot Learning, Cambridge, 16-18 November 2020, 923-932.
- [4] Laddha, A., Gautam, S., Palombo, S., Pandey, S. and Vallespi-Gonzalez, C. (2021) MVFusenet: Improving End-to-End Object Detection and Motion Forecasting through Multi-View Fusion of Lidar Data. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Nashville, 19-25 June 2021, 2859-2868. <https://doi.org/10.1109/cvprw53098.2021.00321>
- [5] Gao, H., Cheng, B., Wang, J., Li, K., Zhao, J. and Li, D. (2018) Object Classification Using CNN-Based Fusion of Vision and LIDAR in Autonomous Vehicle Environment. *IEEE Transactions on Industrial Informatics*, **14**, 4224-4231. <https://doi.org/10.1109/tni.2018.2822828>
- [6] Wang, L., Zhang, X., Song, Z., Bi, J., Zhang, G., Wei, H., et al. (2023) Multi-Modal 3D Object Detection in Autonomous Driving: A Survey and Taxonomy. *IEEE Transactions on Intelligent Vehicles*, **8**, 3781-3798. <https://doi.org/10.1109/tiv.2023.3264658>
- [7] Yu, X., Salimpour, S., Queralta, J.P. and Westerlund, T. (2023) General-Purpose Deep Learning Detection and Segmentation Models for Images from a Lidar-Based Camera Sensor. *Sensors*, **23**, Article 2936. <https://doi.org/10.3390/s23062936>
- [8] 陈德灯, 陆芸, 黄志杰, 等. 基于 YOLO 算法的自动传感器融合技术在自动驾驶中的应用研究[J]. 汽车维修与保养, 2025(9): 129.
- [9] 马祖怀, 周军香, 刘闯. 基于单片机的多传感器融合智能室内环境检测控制系统设计[J]. 造纸装备及材料, 2025, 54(8): 46-48.
- [10] 戴哲文. 城市道路自动驾驶多传感器特征融合研究[J]. 汽车知识, 2025, 25(9): 171-173.