

同濟大學

TONGJI UNIVERSITY

毕业设计（论文）

课题名称 基于雷视融合的塔吊吊装过程动态避障

副标题 与预警方法研究

学 院 土木工程学院

专 业 土木工程

学生姓名 陈宇凡

学 号 2330894

指导教师 卢昱杰教授

日 期 2026 年 2 月 8 日

基于雷视融合的塔吊吊装过程动态避障

摘要

塔式起重机（塔吊）承担了建筑施工中大量的垂直与水平运输任务，其作业过程具有载荷大、作业半径长、人员与设备交叉作业频繁等特点，安全风险长期处于高位。传统塔吊安全管理多依赖司机视野、指挥信号及固定式监控，难以对吊钩端附近的临近人员、临时障碍物以及动态变化的施工环境形成稳定、定量、可闭环的安全感知。

针对上述问题，本文面向“吊钩端”这一关键危险源位置，构建双激光雷达与工业相机协同的轻量化感知系统，研究雷视融合条件下的空间定位、危险区域检测以及分级预警方法。首先，建立多传感器一体化安装与数据链路，完成相机内参、双雷达外参以及雷达—相机外参标定，并给出时间同步与延迟补偿策略；其次，提出双雷达点云实时融合与坐标统一方法，引入LIO-SAM实现吊钩端位姿估计与静态地图构建，在此基础上通过ROI裁剪、时空对齐和静动态分离实现目标生成与跟踪；最后，针对高风险动态目标（如作业工人）引入视觉侧识别结果，设计雷视一致性关联与置信度融合机制，构建基于距离、相对速度与时间到碰撞（TTC）的风险判定模型，实现静态障碍与动态目标的分级预警与策略输出。

通过典型施工工况的对比实验，验证了所构建系统在定位精度、目标关联一致性以及实时性与稳定性方面的有效性，可为塔吊吊装过程的动态避障与风险预警提供工程化参考。

关键词：雷视融合，塔吊吊装，吊钩端感知，动态避障，风险预警

Research on Dynamic Obstacle Avoidance and Early Warning Method for Tower Crane Hoisting Process Based on Radar-Vision Fusion

ABSTRACT

Tower cranes play a critical role in construction logistics, yet hoisting operations involve high loads and complex, dynamic environments, where workers and equipment often coexist within the working radius. Conventional safety approaches mainly rely on the operator's experience, manual signaling, or fixed surveillance, which lack reliable and quantitative perception around the hook end, leading to insufficient collision avoidance capability and delayed risk response.

This thesis focuses on hook-end safety perception and proposes a dynamic obstacle avoidance and early warning method based on radar-vision fusion. A compact perception platform integrating dual LiDARs and an industrial camera is designed for hook-end deployment. A complete multi-sensor calibration pipeline is established, including camera intrinsics, LiDAR-LiDAR extrinsics, and LiDAR-camera extrinsics, together with time synchronization and delay compensation. For online perception, a real-time dual-LiDAR point cloud fusion method and unified coordinate transformation are developed. Hook-end localization and mapping are achieved using LIO-SAM, enabling motion compensation and consistent point cloud generation. On this basis, ROI cropping, spatio-temporal alignment, static-dynamic separation, and tracking are applied to produce reliable obstacle hypotheses. For high-risk dynamic objects (e.g., workers), vision-side detection results are fused with LiDAR observations via bidirectional consistency association and confidence fusion. A risk assessment model combining distance, relative velocity, and time-to-collision (TTC) is then proposed to drive a graded warning state machine.

Experiments under representative hoisting scenarios demonstrate the effectiveness of the proposed system in localization performance, fusion consistency, and real-time stability, providing practical support for intelligent safety management in hoisting operations.

Keywords: Radar-vision fusion, tower crane hoisting, hook-end perception, dynamic obstacle avoidance, risk early warning

目 录

1	引言	3
1.1	研究背景	3
1.2	研究目的与意义	3
1.2.1	研究目的	3
1.2.2	研究意义	4
1.3	国内外研究现状	4
1.3.1	吊装感知硬件布置方案研究	4
1.3.2	施工现场环境感知方法研究	6
1.3.3	碰撞检测算法与预警模型研究	11
1.3.4	文献总结	12
1.4	主要研究内容	13
1.5	技术路线	13
2	吊钩端传感器系统构建与配准	15
2.1	系统总体架构与硬件选型布置	15
2.1.1	传感设备选型	15
2.1.2	安装布局优化	15
2.1.3	吊钩端一体化箱体设计（SolidWorks）	16
2.2	多传感器外参标定	17
2.2.1	相机内参标定	17
2.2.2	双雷达外参自动标定与精配准	19
2.2.3	雷达-相机外参标定	20
2.3	数据链路与接口规范	22
2.3.1	通信链路搭建	22
2.3.2	时间同步与延迟补偿	23
2.3.3	坐标系与 TF 接口规范	24
2.4	本章小结	25
3	基于融合点云的定位与空间危险区域检测方法	26
3.1	双雷达点云实时融合方法	26
3.1.1	原始点云数据特征	26
3.1.2	外参矩阵应用与坐标统一	27
3.1.3	点云实时融合与发布	29
3.2	基于融合点云的 SLAM 定位与静态地图构建（LIO-SAM）	30
3.2.1	算法选择理由与系统输入输出	30
3.2.2	LIO-SAM 流程	31
3.2.3	位姿输出与点云去畸变/配准	32
3.3	面向目标提取的点云标准化与静态目标生成	33
3.3.1	近场关注区域 ROI 裁剪与点云预处理	33
3.3.2	基于位姿的时空对齐与局部滑窗表示	34

3.3.3	静态点分离与动态点云簇提取	35
3.3.4	目标短时跟踪与状态量估计	35
3.4	本章小结	36
4	基于雷视融合的避障预警策略	38
4.1	视觉侧高风险目标识别	38
4.1.1	危险目标（车辆）检测模型与部署	38
4.1.2	检测结果时序稳定化处理	38
4.1.3	点云动态目标融合接口与数据结构定义	39
4.2	雷视融合策略	40
4.2.1	融合总体流程与设计原则	40
4.2.2	双相一致性关联策略	40
4.2.3	轨迹级绑定与遮挡鲁棒融合	41
4.2.4	置信度融合与语义标签输出	42
4.3	风险判定与分级预警策略	42
4.3.1	静态障碍：基于距离/侵入的阈值预警	42
4.3.2	动态目标：基于相对速度/轨迹预测的碰撞风险	43
4.3.3	分级预警状态机	44
5	实验验证与系统评估	45
5.1	实验平台与测试场景设置	45
5.1.1	吊钩端实验平台构成	45
5.1.2	典型测试工况设计	45
5.2	多传感器定位性能实验	45
5.2.1	位姿解算精度验证	45
5.2.2	吊钩动态摆动场景分析	45
5.3	雷视协同感知效果分析	46
5.3.1	危险目标检测准确性评估（车辆）	46
5.3.2	雷视关联一致性分析	46
5.4	危险区域检测效果验证	46
5.4.1	静态障碍场景测试	46
5.4.2	动态目标接近场景测试（车辆）	46
5.5	系统实时性与稳定性验证	46
5.5.1	各模块耗时分析	46
5.5.2	工程应用可行性分析	47
6	结论与展望	48
6.1	主要研究结论	48
6.2	创新点总结	48
6.3	存在的不足与未来展望	48
	参考文献	49

1 引言

1.1 研究背景

建筑业作为关系国计民生的基础性产业，是国民经济的重要组成部分。据国家统计局发布的《中国统计年鉴 2023》数据显示，2022 年建筑业总产值为 83383.1 亿元，占国内生产总值的 6.9%^[1]。在建筑施工作业中，塔式起重机是建筑施工现场中使用最广泛的垂直运输设备，具有起重高度高、作业范围广、工作效率高等优势^[2-4]，其吊装作业承担着钢筋笼、模板、物料等关键构件的吊运任务。然而，塔吊施工环境普遍存在空间狭窄、视野受限、动态干扰因素复杂等特点，加之运行高度大、工作半径长，一旦发生碰撞或人员误入吊装路径，其后果往往极为严重。根据统计，2014 年至 2022 年间全国起重机事故数量不低于 220 起；其中，吊物伤人与物体碰撞类事故累计超过 75 起，占比显著且发生频率较高^[5]。因此，如何实现吊装过程的实时环境感知与风险预警，是当前智慧工地安全管理中亟需攻克的重要问题之一。

近年来，随着建筑智能化与智能建造理念的提出，基于传感器的数据感知技术得到了快速发展，为施工过程数据采集的实时性和自动化需求提供了潜在的解决方案^[6]。大量研究集中于塔吊塔身或塔臂端的防碰撞系统，例如对多塔协同作业的塔臂干涉检测、塔机与周边建筑结构的距离监测等。然而，这类研究大多基于固定位置的传感器布置，主要关注塔臂或塔身层面的宏观防碰撞问题，而对实际吊装环节中最危险的部位塔吊吊钩端关注显著不足。

吊钩端是吊装过程中运动最频繁、最接近人员、车辆及构件的部位，其作业特点具有以下显著特征：（1）空间运动高度不确定：吊钩随吊臂回转、钢丝绳伸缩及外力扰动持续发生三维空间运动，摆幅可达数米，传统固定式传感器难以覆盖其动态空间范围。（2）近场作业风险极高：施工人员在吊运构件下方或附近进行配合作业、车辆/设备在近场区域穿行等均可能导致危险目标无意靠近吊钩端，叠加吊物摆动与构件碰撞等因素，易造成重大安全事故。（3）环境遮挡复杂且光照变化剧烈：钢筋笼、模板体系等大型构件会造成相机视野遮挡，而工地环境存在尘土、夜间照明等不利因素，对视觉系统效果影响显著。（4）传统方法难以实现实时、可落地的近场避障：现有塔吊安全系统普遍依赖塔身安装的超声、雷达或视觉设备，其检测范围难以覆盖吊钩周围 3~5 米的高危近场区域，使得关键危险点无法及时识别。

随着轻量化激光雷达和高性能嵌入式计算平台的快速发展，为塔吊吊钩端构建实时空间感知系统提供了技术基础。激光雷达具备高精度距离测量能力，能够在强光、弱光、粉尘等复杂工况中保持稳定输出，并可实现高达 360° 的水平视场（FoV），从而覆盖吊钩周围的全方位近场空间。相比之下，视觉传感器具有丰富的语义理解能力，特别是在识别施工人员、车辆等关键危险目标方面具有显著优势。因此，构建“雷达为主、视觉为辅”的雷视融合感知架构能够充分结合两类传感器的互补特性，为实现塔吊吊钩端的实时、可靠避障预警系统创造了可行条件。

基于上述背景，本研究面向吊钩端高风险作业场景，选用激光雷达与工业相机设计并开发集成式吊钩端感知系统，以实现对近场三维环境的高精度感知与危险目标识别，减少施工现场吊装作业中的碰撞风险、保障人员生命安全，也为建筑施工装备的数字化、智能化升级提供了一种可落地、可推广的技术路径，具有重要的工程应用价值和现实意义。

1.2 研究目的与意义

1.2.1 研究目的

针对塔吊吊装作业中吊钩端近场空间运动不确定、风险集中、环境干扰复杂等核心痛点，以及现有安全感知系统对吊钩端覆盖不足、单一传感器鲁棒性差、预警精度与实时性难以兼顾的问

题，本研究旨在实现以下目标：

- 1.设计并构建一套可工程化落地的吊钩端集成式感知系统，通过激光雷达与工业相机的协同部署，突破传统传感器在覆盖范围、环境适应性上的局限，实现对吊钩周围 3~5 米高危险区域的全方位、高精度感知。
- 2.提出基于雷视融合的多源数据处理与信息融合方法，整合激光雷达的三维几何测距优势与视觉传感器的语义识别能力，实现静态障碍精准建模、动态危险目标（实验中以车辆为代表，可扩展至施工人员等）可靠识别与轨迹跟踪。
- 3.建立兼顾空间距离与时间裕度的动态风险评估模型，形成“感知 - 识别 - 预判 - 预警”的闭环机制，实现对碰撞风险的分级预警与动态避障辅助决策，提升预警的提前量与准确性，降低误报、漏报率。
- 4.验证系统在复杂施工工况下的稳定性、实时性与工程适用性，为塔吊吊装作业安全防护提供可推广、可复用的技术方案。

1.2.2 研究意义

建筑业塔吊吊装事故频发，给人员生命安全与工程财产造成严重损失。本研究聚焦吊钩端这一最高风险区域，通过构建实时感知与预警系统，可直接减少人员误入、吊物碰撞等恶性事故发生概率，为施工现场安全管理提供技术保障。同时，系统采用轻量化硬件集成与高效算法设计，兼顾部署成本与运行稳定性，适配施工场地粉尘、震动、光照多变等恶劣环境，具备较强的工程落地性，可快速推广至各类建筑施工场景，推动塔吊装备的智能化升级，助力智慧工地建设。

现有塔吊安全感知研究多集中于塔身、塔臂等宏观层面，对吊钩端动态近场感知的研究较为匮乏。本研究提出“雷达为主、视觉为辅”的雷视融合架构，完善了多传感器在高危作业端的部署与标定方法，丰富了动态复杂场景下近场感知的理论体系。同时，研究建立的静态目标分离、轨迹预测与分级预警模型，为多源数据融合在工程装备安全防护中的应用提供了新的思路，填补了吊钩端高精度、高鲁棒性感知技术的研究空白，为同类大型工程机械的安全智能化提供了理论参考与技术借鉴。

随着智能建造理念的深入推进，施工装备的数字化、智能化成为行业发展必然趋势。本研究通过传感器集成、数据融合与智能算法的深度结合，实现了吊装作业从“被动防护”向“主动预警”的转变，契合建筑行业安全升级与效率提升的核心需求。研究成果不仅可应用于塔吊设备，还可迁移至履带吊、汽车吊等其他起重机械，为工程机械行业的智能化转型提供典型示范，推动建筑业向安全、高效、智能的方向高质量发展。

1.3 国内外研究现状

随着智能建造与施工现场数字化水平的不断提升，工程机械与大型吊装设备的安全感知与智能防护逐渐成为国内外学术界和工程界的研究热点。围绕施工设备运行安全，已有研究主要集中在作业环境信息获取以及碰撞风险识别与预警等方面，形成了从感知、理解到决策的多层次研究体系。然而，由于施工现场环境复杂、目标类型多样、作业过程动态性强，现有研究在感知对象、传感器布置位置以及安全防护粒度等方面仍存在一定局限，尤其在塔吊等大型起重设备的吊钩端动态作业区域，相关研究尚不充分。

1.3.1 吊装感知硬件布置方案研究

现有塔吊安全感知研究多聚焦塔机本体（塔身、塔臂、小车等）的宏观防碰撞与运行状态监测，而对吊钩端这一近场高风险区域的直接感知研究相对不足。但已有工作在传感器选型、布设

优化、数据链路组织与风险触发逻辑等方面形成的工程经验，对吊钩端场景具有重要参考价值。

吊装作业是施工现场最常见的高危作业之一。为防范作业过程中可能对人员或建筑构件造成的损害，众多学者致力于通过智能感知技术与传感器部署来提升作业安全性。例如，在塔吊吊臂、小车、驾驶室及施工现场周边等位置，布设摄像机^[7]、惯性测量单元（Inertial Measurement Unit, IMU）^[8,9]、射频识别（RFID）^[10]、超宽带（Ultra Wide Band, UWB）^[6]、超声波传感器^[11]、高度传感器^[12]、风速传感器^[13]等多种传感器，以实时监测作业状态、识别周边障碍物与人员，从而实现风险预警。

值得注意的是，传感器布设方案（含布设位置、传感器选型及集成方式）与传感精度直接决定环境感知效果，而感知数据作为后续风险识别、动态避障等算法判断的核心输入依据，其质量优劣直接影响决策的准确性与可靠性。不同布设位置、不同类型传感器所获取的数据，在感知范围、实时性、抗干扰性等方面存在显著差异，因此科学优化传感器布设方案，是提升吊装作业智能防护水平的关键前提。

在实际施工现场，多台塔吊交叉作业的场景普遍存在，作业空间交错复杂，极易引发塔吊间结构干涉、吊物与周边物体碰撞等叠加风险。早期塔吊防碰撞研究多依赖塔机自身运行参数进行间接判断，其中部分方案的传感器布设虽未直接聚焦塔臂，但为后续塔臂布设方案提供了基础思路。例如，周飞虎^[14]在塔机回转机构、变幅机构与起升机构等关键部位布设角度、幅度与高度传感器，采集塔机姿态与运行状态数据，通过几何关系推算吊钩位置及构件相对关系，将潜在碰撞转化为构件间几何距离与阈值的比较以触发预警。

进一步地，Zhong 等^[15]从工程应用角度出发，构建了一套基于无线传感器网络（WSN）与物联网（IoT）的塔吊群安全管理系统。该系统在塔吊关键结构部位布设多类型传感器节点，包括塔臂、小车及起升机构等位置，用于采集塔吊运行状态与作业参数，并通过无线传感器网络实现多塔设备之间的数据传输与集中管理。传感器采集的数据主要包括塔臂回转角度、小车位置、吊钩高度等信息，系统基于这些参数在统一坐标系下对塔吊结构的空間关系进行计算，从而判断塔吊之间是否存在潜在的干涉或碰撞风险。当相关参数超过预设安全阈值时，系统通过物联网平台向操作人员发送预警信息，实现对塔吊群作业安全状态的实时监控。

针对塔吊间碰撞预警，Hwang 等^[6]在实验室环境下搭建了两个缩尺塔吊模型，在每台塔吊吊臂前端和后端各布置 1 个 UWB 标签，并在实验场地角点固定布设 UWB 接收器，通过 AoA 与 TDoA 相结合的方式获取标签的空间位置信息。基于塔吊轴向旋转的运动特性，该方法采用标签间距离作为等效几何约束，将塔吊结构之间的安全距离映射为标签距离阈值，从而实现塔吊间潜在碰撞的实时预警。但该方案的传感器仍仅布设于塔吊结构本体，未在吊钩或吊物上增设独立感知单元，因此无法覆盖吊物与环境障碍物之间的碰撞风险，感知范围存在明显局限。

为直接感知吊臂周边障碍物，赵宇^[16]提出了一种基于超声波传感与无线组网的塔吊防碰撞方案，沿吊臂方向在吊臂关键位置均匀安装多个节点式超声波探测器阵列，用于实时感知吊臂周围障碍物的距离信息。由于单个超声波传感器视场角有限，该方法通过阵列化布设扩大探测范围，在吊臂周围形成连续的线性保护区域，可直接实现吊臂与障碍物之间的碰撞判断。同时，依靠部署在塔机回转机构和起升机构等关键部位的高度、角度传感器来获取吊物高度及塔身回转角度等参数，融合后推算出吊装物体的空间位置，从而间接实现吊装物体与障碍物之间的碰撞预警，但这样的方式仍存在因吊索与吊物摆动导致的定位偏差问题。

Zhou 等^[17]将 GNSS 接收机固定于塔吊小车的支撑框架顶部，实时获取各塔吊塔臂端部在统一坐标系下的空间位置信息，从而实现对多塔空间关系的持续监测。基于 GNSS 解算得到的平面坐标数据，系统将塔臂简化为几何线段，并通过计算不同塔吊塔臂之间的相对距离来评估潜在的碰撞风险，当距离小于预设安全阈值时触发相应的碰撞预警。该方法通过引入 GNSS 定位传感器

提升了多塔作业环境下的空间感知能力，但其传感器布设仍主要集中于塔吊结构本体，防碰撞对象以塔臂之间的结构干涉为主，未对吊钩或吊装物体进行直接感知。

张知田等^[18]提出面向“塔吊—工人”空间交互的危险场景自动检测框架，通过工人位置与吊钩（等效吊物）位置的同步获取实现预警。硬件布设上，该方法在塔吊驾驶室下方架设远视角摄像头以捕捉作业面内工人分布，同时在吊臂滑动小车下方安装垂直向下摄像头以估计吊钩/吊物的水平摆动偏移，并在塔吊主体布设角度与测距传感器获取吊臂旋转/下压角及小车、吊钩关键距离参数，通过局域网络实现数据实时传输。基于多源数据融合，方法先由图像估计吊钩偏移，再结合角度/测距数据计算吊钩三维坐标，并将工人与吊钩位置映射至 BIM 模型；在判别层面，以吊钩垂直投影点为中心构建 10 m 圆形半径的动态危险区域，当工人进入该区域即触发危险场景预警。

上述方法中，传感器布设集中于塔机本体吊臂、小车等平台，工程部署难度低、成本可控，但因未直接感知周边环境，对周围障碍物及吊装作业的直接风险缺乏有效感知能力。且它们大都通过传感数据组合联立，以计算出如吊钩、吊臂的特定位置，并与障碍物计算相对距离。总体而言，这类方案对吊钩端近场障碍与人员的直接几何感知能力不足，且吊索摆动、遮挡与传感器视场限制会进一步削弱对端部高风险区域的覆盖。因此，部分工作进一步将传感设备（或可被可靠观测的靶标）直接部署在吊钩端/吊钩块/吊物，以更直接获取端部真实运动状态并提升预警可靠性。

Lee 等^[19]在塔吊盲吊导航系统中，将吊钩块作为关键观测对象：在变幅臂端安装激光传感器，并在吊钩块安装反射板作为测距靶标，利用激光束反射测得臂端到吊钩块的实际垂直距离。考虑到吊钩块剧烈摆动会导致“错失反射板”而产生误测，系统进一步结合吊钩块运动速度与方向对异常数据进行滤除，同时以放绳长度编码器作为备份测量源，最终将多源位置与视频信息发送至驾驶室内的导航服务器并在 BIM 模型中实时显示吊物位置与周边环境关系，主要用于提升盲吊场景下的态势感知与导航辅助。

在此基础上，为更直接表征“吊钩端摆动/旋转”等动态风险，Fang 等^[20]提出将无线 IMU 刚性安装在吊物或吊钩端载荷上，由 IMU 测得吊物的欧拉旋转角，在已知吊绳长度并近似刚性的前提下，可将姿态角进一步转换为吊物在三维空间中的相对位置，从而实现了对吊物摆动状态的在线监测，并与编码器等机构传感数据结合用于安全辅助与告警。

不过，相关综述也指出这类“端部直测”方案往往要求在吊物/吊钩块/吊钩上增加额外硬件，带来电池维护、现场粗放工况下的损伤风险，以及金属环境对无线链路的干扰等工程约束。

进一步地，Ku 等^[19]在机器人化塔吊系统中将“吊钩块”视为末端执行器，直接在吊钩块安装 3D LiDAR 获取近场点云以感知周围结构与潜在障碍，并将点云处理流水线部署在吊钩块上的工业计算机中实时提取路径方向上的障碍高度等关键信息，处理结果再无线回传至驾驶室内主机用于决策与避障执行。

除此之外，由于摄像机轻便的特性，许多研究者也尝试将摄像机安置于吊钩端，对下方拍摄感知，实现如施工危险对象识别^[21]、工人与负载碰撞预警^[22]等塔吊安全监测与预防研究。

1.3.2 施工现场环境感知方法研究

施工现场环境感知是塔吊安全预警与避障控制的基础，其核心在于在复杂、动态且强不确定的工况下持续获取“设备自身状态—作业空间几何—人员与障碍物信息”，并据此进行风险评估与提示。现有研究与工程应用中，塔吊相关感知传感器大体可按信息来源划分为三类：一类用于获取塔吊本体或吊钩端的运动状态与空间位姿，为后续感知与预警提供统一参考；第二类用于感知作业环境中的几何距离与空间结构，支撑安全距离判断与危险区域入侵检测；第三类用于获取目标的语义或身份信息，尤其是对危险目标（如工人、车辆等）的识别与管理。不同传感器在测

量维度、精度、部署成本与环境鲁棒性方面存在显著差异，也决定了其在塔吊避障系统中常见的组合方式与融合层级选择。早期研究主要探索不同类型传感器在感知问题上的应用，随着研究不断深入以及感知判断要求的不断提升，多源传感器融合方式为提高系统鲁棒性、准确性提供了良好的解决办法。

（1）传感技术。

施工现场常见的感知信息源可概括为“设备自身状态/位姿感知”“外部环境几何测距与空间结构感知”“目标语义与身份识别”三类。不同传感器在测量维度、精度与环境鲁棒性方面差异显著，工程实现通常采用多源组合以兼顾覆盖范围、成本与可维护性。

编码器（包含角度、高度等状态传感器）常用于获取塔吊自身的状态，通过获取回转半径和回转角度等信息，能够步步推算出小车位置、吊钩高度以及结构关键部位的运动边界信息。该类传感器可靠性高、成本低、工程布置成熟，能够为塔吊运动学模型、作业包络计算与禁行区生成提供直接输入，是典型的“间接感知”手段。然而其测量对象主要是设备自身状态，难以感知外部环境中临时出现的障碍物、人员靠近以及吊物周边的局部复杂结构，因此通常需要与外部环境传感器结合才能形成完整的避障能力。

惯性测量单元（IMU）能够提供角速度与加速度信息，在吊钩端强动态条件下用于运动状态估计、姿态变化监测以及对扫描类传感器（如 LiDAR）进行运动畸变补偿。对于吊钩端避障场景，IMU 的价值主要体现在两方面：其一，用于描述吊钩摆动、旋转等快速姿态变化，为点云稳定化与坐标系一致性提供支撑；其二，与其他传感器（如 LiDAR、视觉、GNSS）融合形成更稳定的位姿估计。但 IMU 存在累计漂移，单独使用难以提供长期稳定的位置与姿态基准，通常需要与外部定位或地图约束结合。

GNSS（及 RTK）常用于塔吊场景中的宏观定位与设备状态监测，例如塔吊基座或关键构件的绝对位置获取、施工区域级的空间约束建立等。RTK 具备较高的平面定位精度，适用于开阔环境下的绝对坐标基准构建，也便于与 BIM/施工组织规划进行空间对齐。然而在城市密集施工环境中，GNSS 易受遮挡与多径效应影响，且对吊钩端等局部高动态部位的精细定位支撑有限；此外，GNSS 本身无法直接感知障碍物几何边界，因此更适合作为全局参考或与 IMU、LiDAR 等共同构建多传感器定位框架。

UWB 定位系统在施工安全管理领域应用较多，典型方式是在工人佩戴标签、现场布置基站，从而实现人员定位、电子围栏与区域入侵告警。UWB 的优势在于可对人员提供显式定位信息，便于实现“人”的风险管理与统计，并且可与塔吊作业区域规则进行结合。但 UWB 系统部署依赖基站布设与标定，定位精度与稳定性受环境遮挡、金属结构、多径效应影响明显；同时 UWB 难以描述障碍物的几何形状与边界，通常更适合作为人员高危提示或辅助决策信息，而非替代基于几何测距的避障感知。

RFID 技术在塔吊作业中更多用于构件、吊具或物料的身份识别与流程管理，例如对吊装构件进行唯一标识、实现装卸与流转记录、辅助吊装任务调度与追溯。在安全预警层面，RFID 可用于识别“正在吊装的对象是什么”、与作业权限或工序规则联动，从而在管理层面增强安全性。但 RFID 通常难以提供稳定连续的空间位置与距离信息，其通信距离、读写可靠性也受金属遮挡与现场电磁环境影响，因此更适合作为“身份/状态信息源”融入系统，而非用于实时几何避障判定。

视觉作为一种信息丰富且价格便宜的信息获取方式，在语义理解方面具有明显优势，能够直观区分不同目标类别，且硬件成本较低、部署灵活。由于塔吊操作员的特殊工作位置导致不可避免存在实现盲区，早期的研究者们就已经尝试在塔吊上部署视觉系统^[23]以显示吊装的负载和下方的工作区域，让在驾驶室的操作员能够直观看到吊钩端的情况，减少视觉盲区。如今，随着计算机技术发展，通过结合传统图像处理或深度学习算法，能够实现施工对象的目标检测，在

此基础上可以实现吊物空间姿态测量^[24]、生产效率统计^[25]、工人运动轨迹监测^[22]等任务。然而，视觉感知方法对光照条件和环境遮挡较为敏感，在夜间作业、强光或复杂遮挡条件下感知性能容易下降，同时难以直接获取精确的空间距离信息，在安全距离判断等应用中存在一定局限。与普通相机相比，深度相机及结构光相机能够直接提供局部深度信息，适合近距离空间测量与区域入侵检测，但在强光环境、远距离以及室外复杂工况下稳定性和测距范围会受到限制，其适用边界需要结合具体作业场景评估。

毫米波雷达具备全天候工作能力，对雨雾粉尘等恶劣工况鲁棒，且对运动目标具有较强敏感性，在夜间或能见度较低场景中具有优势。其不足在于角分辨率与空间细节表达能力通常不如 LiDAR，难以精确刻画复杂障碍物的几何边界，因此在塔吊避障中更常用于冗余检测、近距离存在性判断或对运动目标的补充提示，而非单独承担精细空间建模任务。

超声波传感器成本低、实现简单，适合极近距离防撞与边界触发（例如小范围防护或特定方向的距离阈值检测）。但其测距范围有限，受传播介质、风噪与环境干扰影响较大，且难以覆盖吊钩端大范围三维空间，因此更多作为近距冗余或辅助触发传感器，与其他三维传感器配合使用。

激光雷达（LiDAR）能够输出三维点云并提供稳定的测距能力，是施工现场空间几何感知的关键传感器之一。对于塔吊避障而言，LiDAR 可直接支持危险区域入侵判定、安全距离阈值触发、障碍物轮廓提取等任务，并且对光照变化不敏感，具备较高工程鲁棒性。但 LiDAR 对语义类别理解较弱，对“人”与“非人”的区分通常需要结合视觉或学习型算法；此外，金属反射、玻璃等材质以及遮挡引起的点云缺失、离群点等问题，需要配合滤波、聚类与一致性策略提高稳定性。因此在实际系统中，LiDAR 常作为“测距与空间判定主传感器”，与其他语义或状态传感器形成互补。

（2）多传感器融合感知

在复杂施工现场，仅依赖单一传感器往往难以同时满足安全预警所需的**空间几何精度、语义理解能力与环境鲁棒性**。因此，多传感器融合逐渐成为施工环境感知的重要技术路线。一般而言，多传感器融合是指在统一的时间与空间参考框架下，对来自不同传感器的信息进行对齐、关联与组合，以获得更完整、更可靠的环境状态估计或风险判定结果。其本质目标可概括为两类：其一，通过不同模态信息的**互补性**提升感知维度（例如视觉提供语义、测距传感器提供几何距离）；其二，通过信息的**冗余性**增强系统可靠性与容错能力，从而降低施工现场不确定因素带来的误报与漏报。

在塔吊避障与安全管理领域，多传感器融合研究大体可以分为两类路径：一类面向**塔吊/吊钩端状态与位姿获取**，以编码器、限位、GNSS/RTK、IMU 等为主要信息源，通过融合提升位姿连续性与全局参考稳定性，为作业包络计算、限制区生成与塔群协同防碰撞提供支撑。

许多研究者^[8, 20, 26-29]明确了测量起重机操作的多个关键运动：吊臂回转角、提升绳伸长量和吊臂伸长量，并采用编码器、倾角传感器进行测量获取，以捕捉塔吊的运动状态。这种方式主要聚焦于布设于塔吊不同位置的传感器，根据每部分的参数通过建立塔吊的运动方程得到塔吊自身的状态，在原理上属于数据的拼接。仅依靠这些数据，在现实摆动的情况下无法准确估计吊钩和吊物的状态，这也是值得关注的高风险区域。对此，有通过线性激光装置测量塔吊小车到吊钩端的距离，再结合旋转编码器获取的旋转角度来得到吊物所在的位置^[29]；也有在吊钩上侧金属壁安装 IMU，并结合起重机的传感数据以捕捉吊钩的运动姿态（图 1）^[20]。而 Price 等^[8]考虑到吊物冲击较大时可能损坏传感器，采用了视觉系统来定位吊物的偏移和旋转，但是缺乏吊物的深度信息，仍通过传感获取的缆绳长度作为吊物高度的间接估计。

由于施工现场是一个动态复杂环境，必须将起吊现场环境数据集成到起吊辅助框架中，才能起到实际的安全防护作用，同时，由于起重机特有的运动属性，其覆盖范围不能仅考虑二维，而

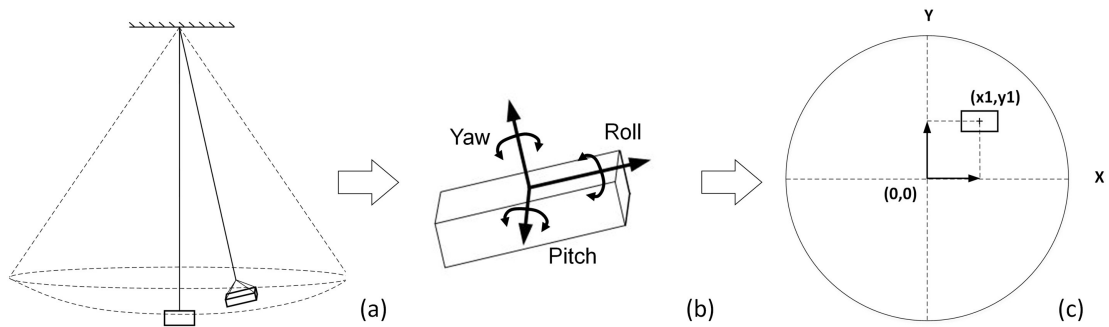


图 1.1 Transforming the angular measurements to absolute positions

是要实现三维感知。而三维激光扫描技术为上述命题提供了一种解决方案。依托地面激光扫描仪或者无人机摄影测量技术^[30]，可以获取工地范围内的 3D 点云数据，每个扫描点都包含了三维坐标信息及 RGB 颜色数据，可以表征场景中物体的几何形状。扫描后的点云经过点云降采样、分割、聚类 and 方向估计等步骤，对场景中各种对象形成定向边界框，形成一张静态地图并集成到安全预警系统中，随周期性的扫描进行更新^[8,20]。但是这样的方式仍存在分辨率低和更新频率低的问题。同时，点云更新和扫描范围也是一个问题，每台塔吊有自己的工作范围，大面积的场景建模会扫描许多冗余点云，为场景点云更新和算法判断带来处理性能压力。

另一类路径面向外部环境与危险目标风险识别，更强调对障碍物几何距离与目标语义信息的联合获取。典型做法包括以激光雷达/毫米波雷达提供可靠测距与空间入侵判定，以视觉方法提供危险目标（如工人、车辆等）的语义识别与可视化解释，并在决策层进行风险分级预警。在实际研究中，以避障为最终研究目的的研究，通常使用多个传感器，在实现塔身姿态的内部感知基础上，完成外界状态的外部感知。

在施工现场环境感知中，视觉传感器因能提供目标语义信息与可视化解释，成为多传感器融合系统的核心感知单元，相关研究围绕技术优化与工程适配展开了大量探索。

Yang 等^[31] 采用 Mask R-CNN 对人员与障碍物进行语义分割，通过像素 - 实际距离转换模型，联合输出人员语义状态与安全距离，计算误差控制在 3% 以内；Golcarenarenji 等^[21] 设计的 CraneNet 深度学习模型，基于单目视觉实现 50 米范围内人员检测（语义识别）与远距离测量（几何距离），准确率达 92.59%，且在嵌入式设备上实现 19 FPS 的实时性，适用于起重机驾驶室等空间受限场景。这类方案的优势在于部署成本低、语义信息提取能力强，但测距精度受图像分辨率与特征提取质量影响，弱光环境下性能易下降。双目视觉方案则通过立体匹配直接获取障碍物距离，几何精度高于单目，同时结合 CNN 实现人员 / 设备语义分类，但其测量范围受基线长度限制，近距离遮挡时匹配困难，设备体积也大于单目视觉，更适用于固定监测点的近距离作业场景。

视觉与其他传感器的融合方案进一步提升了联合感知的鲁棒性与精度。视觉与激光雷达融合方案中，视觉负责人员、设备、障碍物的语义识别，激光雷达提供高精度几何距离与三维点云，通过特征级融合将距离特征融入视觉检测网络，有效弥补视觉测距精度不足与激光雷达语义信息匮乏的缺陷，Price 等^[32] 基于该方案实现盲吊场景的障碍物距离与人员语义联合预警，适配高精度要求的复杂环境；视觉 + RFID/IMU 融合方案则利用 RFID/IMU 的精准定位能力（几何距离）与视觉的行为语义提取优势，通过决策级融合实现人员动态风险预警，Zhang 等^[33] 采用 FairMOT 算法跟踪人员与吊具，结合 Transformer 模型预测轨迹，联合人员位置信息与移动语义，建立多等级碰撞风险预警规则，提升人员密集场景的风险识别可靠性。此外，聚焦人员风险识别的专项融合方案成为研究焦点，Shapira 等^[23] 的视觉系统通过人员检测与距离判断，在缩短起重机作业周期的同时降低安全风险。

（3）雷视融合感知

在智能交通、自动驾驶、智慧工地等复杂环境感知场景中，精准、实时、鲁棒地获取目标信息是系统安全运行的核心前提。雷达与视觉传感器作为核心感知载体各有优劣且高度互补：视觉传感器语义信息提取能力强、成本低、部署灵活，可精准识别目标类别、外观特征及场景细节，但易受光照、天气及遮挡影响，远距离测距测速精度不足；雷达具备高精度测距测速、抗干扰能力强等优势，激光雷达能输出精准三维点云，毫米波雷达适应恶劣天气且动态跟踪效果好，但激光雷达语义信息匮乏、毫米波雷达分辨率有限。雷视融合通过数据级、特征级与决策级等融合策略整合二者信息，可借助视觉语义优势弥补雷达短板，同时依托雷达空间感知能力修正视觉偏差，从而提升复杂场景下的感知鲁棒性与可解释性。

从融合策略与信息处理层级出发，现有研究将雷视融合划分为数据级融合、特征级融合、目标属性级融合与决策级融合四类^[34]。

数据级融合强调在原始数据层面实现信息联合，对多源测距数据进行坐标统一与点云拼接、对图像与深度/点云进行时空对齐后再进行联合建模。该策略的关键在于统一不同模态数据的表达尺度与坐标基准，从而在最底层最大化保留信息量，并为后续处理提供更一致的输入表示。Zhou 和 Omar^[35] 采用像素级融合红外图像与 RGB 图像，采用自适应加权平均算法处理非饱和像素数据，针对饱和像素区域（眩光场景）实施主成分分析（PCA）算法，将两种输入的图像融合为单一输入，实现图像增强，以更鲁棒地检测实际场景中出现的人员和障碍。由于雷达和图像在水平与垂直分辨率上远低于光学图像^[34]，因此，需要在时空间上分别进行对齐，而非图像间的归一化像素级融合。Kim 等^[36] 综合使用 GPS、毫米波雷达、LiDAR 与视觉构建环境表示，通过多帧 LiDAR 观测累积生成网格地图，并以栅格“障碍观测计数”超过阈值触发风险预警，同时用雷达候选目标与视觉识别目标交叉验证并更新静态地图与安全区域。Lekic 与 Babic^[37] 利用 GAN 将雷达信息生成环境图像并与光学图像融合，Ouyang 等^[38] 则使用条件 GAN 在图像监督下由点云重建语义场景图像，旨在缓解点云与相机直接融合带来的计算负担，并在 KITTI 数据集上验证实时检测有效性。

特征级融合通常先从各传感器数据中提取如几何特征、运动特征或深度网络特征的中间表示，再在特征空间完成联合建模，能够在信息利用与系统复杂度之间取得一定平衡，尤其在学习型感知框架中较为常见。传统机器学习时期往往需要先提取 HOG、GLCM 等人工特征，再用 SVM/AdaBoost 等分类器完成识别；而深度学习兴起后，特征融合更多被嵌入神经网络结构中，以端到端方式自动学习跨模态互补表示^[34]。在雷视融合任务中，特征级融合常用于道路/可行驶区域检测等场景，例如 Caltagirone 等^[39] 基于全卷积网络实现 LiDAR 与相机融合的道路检测。

更贴近工程避障预警需求的是目标属性级融合，各传感器先输出目标候选，再进行融合以降低单一传感器误报与漏报，该层级的抽象程度介于数据级融合与特征级融合之间，能够在保持计算可控的同时显著提升系统稳定性。一个与“吊钩端避障可视化”高度一致的范式是“LiDAR 生成 ROI → 图像中识别/解释 ROI”，如 Wu 等^[40] 利用 LiDAR 的距离与角度信息在图像中生成 ROI，并结合 3D 点云形状信息进一步验证目标以降低误报、缓解遮挡带来的相机漏检。同时，该方法通过匹配 LiDAR 与图像端的目标列表来提升检测速度并获得较高的行人检测精度。Han 等^[41] 从系统实现角度强调实时性与工程落地，为“点云候选 + 视觉解释”的在线管线提供了工程参考。类似地，Zhong 等^[42] 利用毫米波雷达运动信息提供图像 ROI，并在 ROI 上用 CNN 识别目标，同时合并雷达与相机的目标列表以提升鲁棒性。

决策级融合则由各传感器独立完成检测、识别或风险判定，再通过规则、置信度或逻辑推理进行融合输出，如采用贝叶斯推理、D-S 证据理论、模糊推理等。对工程避障而言，Wei 等^[43] 直接面向工业场景的实时避碰系统，体现了“检测结果融合 → 避障决策”的应用导向。De Silva 等^[44] 则强调在移动机器人平台上实现鲁棒融合。

综上所述，雷视融合研究的核心共识在于：通过融合激光雷达的高精度三维几何与测距能力以及相机的纹理与语义表达能力，可以在复杂动态环境中显著提升目标感知的完整性、鲁棒性与可解释性，并降低单一传感器在遮挡、光照变化或传感噪声下造成的误检与漏检风险。现有工作虽然在融合深度与实现形式上差异较大，但普遍强调“可靠的时空对齐与外参标定”是雷视融合有效性的基础，并在此之上形成从端到端学习到工程化组合的多样路线。进一步而言，面向安全预警与避障等工程任务，雷视融合的研究趋势并非一味追求更深的网络耦合，而是更强调在实时性、系统复杂度与安全可靠之间取得平衡：一方面利用几何信息实现稳定的距离与区域风险触发，另一方面用视觉语义增强对高风险目标的识别与可视化解释，从而构建“可落地、可维护、可扩展”的感知—预警闭环。

1.3.3 碰撞检测算法与预警模型研究

关于塔吊碰撞的检测与约束建模，既有研究首先在塔吊吊装规划与虚拟仿真场景中形成了较为系统的方法体系^[45]。在该场景下，研究人员多采用虚拟环境构建手段以模拟实际吊装场景，并以碰撞检测作为 4D 路径规划的约束条件^[46]，如通过为塔吊构建定向包围盒（Oriented bounding box, OBB）^[47]，采用离散碰撞检测（Discrete Collision Detection, DCD）^[48, 49] 或最小距离计算^[50] 等方式判断碰撞是否发生。Lai 和 Kang^[51] 中以简化外界边界（如球体、圆柱体等）近似现场机械与结构构件，从而显著降低实时碰撞检测的计算负担。这些方法在虚拟仿真环境下具有可行性，因为此时障碍物信息通常是已知且假设完全可观测的。但在实际场景中，周围环境的信息需要通过传感器实时获取，传感器的观测范围、精度以及数据延迟等问题，单纯依赖离线/静态几何模型的碰撞约束往往难以直接迁移。

针对施工过程中的实时安全辅助，研究逐渐从“是否碰撞”的二值判别，转向“碰撞风险强度”与“分级预警”建模。一类典型思路是构建安全包络或风险区域，采用最小距离/侵入判别触发告警^[52]。Wang 等^[53] 将风险区域与目标简化为 AABB，并通过进入/逼近判别实现安全预警逻辑，体现了“几何包络 + 阈值规则”在实时系统中的工程可实现性。Yang 等^[54] 使用 Mask R-CNN 检测危险源并根据照片与实际的像素转换关系换算计算工人与危险源的安全距离。

更进一步的研究强调在作业过程中持续评估风险并向操作者提供主动式辅助，面向吊装作业的实时主动安全辅助框架的核心在于动态工作空间建模与风险提示机制，而非仅在碰撞发生时被动报警^[20]。Fang 等^[9] 进一步从人机协同角度讨论了安全风险呈现与操作者态势感知评估，为预警模型如何“可解释地”服务现场作业提供了参考。

当障碍物与塔吊/吊载存在相对运动时，仅以“当前距离是否小于阈值”进行判断可能出现两类问题：其一，两个目标距离尚大但存在快速相向趋势，风险会在短时间内迅速上升；其二，阈值设置过大又会带来高误报，影响作业效率。因此，很多研究引入时间域 surrogate risk 指标，将风险刻画从“空间距离”推进到“时间余量”，典型如 Time-to-Collision (TTC) 等。虽然 TTC 最初广泛用于交通安全 surrogate measures 的研究语境，Nadimi 等^[55] 对 TTC 等指标的适用性与局限进行了系统讨论，为“时间裕度型预警阈值如何设定与解释”提供了可借鉴的建模思路。在吊装场景下，将 TTC 类思想与相对速度/加速度估计结合，可形成“距离—速度联合预警”的框架，使预警更关注“是否会在未来短时窗内进入危险区”，而非仅关注当前几何接触。

根据 Kim 等^[56] 的研究，相较于仅预防邻近事故，预测潜在的碰撞事故能够更为有效地避免危险。围绕预测驱动的碰撞预警，研究通常先对作业对象进行短时轨迹预测，再对未来时窗内的碰撞可能性进行判别或分级。张冬^[57] 将碰撞预警方式区分为两种形式：针对塔吊与静止建筑物的碰撞，考虑为基于距离的预警；而针对塔吊与塔吊之间的碰撞，在距离模型的基础上结合速度模型来进行综合判断。进一步地，轨迹预测方法本身也形成了从统计模型到深度学习模型的谱系，

例如通过隐马尔可夫模型（Hidden Markov Model, HMM）^[58]、长短期记忆人工神经网络（Long Short-Term Memory, LSTM）^[59]、深度神经网络（Deep Neural Network, DNN）^[60] 等对人员或障碍物运动进行建模与预测。在获得塔吊与障碍物的未来轨迹后，可通过综合分析双方的相对运动趋势来进行风险判断，例如采用基于碰撞锥^[61, 62]、基于约束优化等方法^[63]。

除风险度量模型外，动态场景下的碰撞检测还面临“实时性—精度/漏检风险”的工程折中问题。为此，相关研究在算法实现上通常采用“分层剔除 + 精细检测”的策略。在吊装路径规划与动态环境中，Dutta 等^[48] 提出了近实时重规划模块，并在决策支持部分采用多层级 OBB 来进行风险相关的判别与触发，体现了“层次包围体 + 决策逻辑”在动态场景中的可扩展性。Zhu 等^[64] 强调为提高碰撞检测效率与精度，需要融合空间划分与层次包围体等策略，并讨论了连续碰撞检测（CCD）在吊装扫掠体场景下的重要性。同样地，Lin 等^[65] 提出基于点云的碰撞检测方法并将其嵌入规划流程，反映出“直接用现场环境状态驱动碰撞判别”的趋势。

在更一般的动态避碰理论中，碰撞风险还可通过“相对速度可达集合”进行刻画，例如碰撞锥（collision cone）与速度障碍（velocity obstacle, VO）范式。Fiorini and Shiller 早在 1998 年便提出了 VO 框架以在动态环境中选择规避机动^[66]，其思想可用于将塔吊端执行体与动态障碍的相对运动关系显式化。Van den Berg 等^[67] 将该思想扩展到多主体实时避碰，为复杂多障碍、多作业体并行施工条件下的风险判别提供了更系统的理论工具。进一步地，安全关键建模常将避碰问题表述为“满足安全集约束”的控制问题，控制障碍函数（CBF）成为近年来重要方向之一。Jian 等^[68] 提出了 D-CBF，将障碍物预测与安全约束结合，用以保证动态避障的可行性与安全性，这为“预警模型如何与后续控制/干预策略耦合”提供了可迁移的建模范式。

综上，塔吊碰撞检测与预警模型研究呈现出从“仿真/规划约束下的几何碰撞判别”向“面向施工过程的分级预警”再到“时间裕度与预测驱动风险评估”的演进趋势。一方面，距离/包络等几何判别仍是预警触发的基础，但在动态场景下需结合时间裕度指标与轨迹预测以提升提前量并降低误报；另一方面，为实现在线运行，碰撞检测算法本身也趋向采用层次包围体、空间划分与连续碰撞检测等策略来兼顾效率与漏检风险。上述研究为后续构建面向塔吊作业的实时碰撞预警系统提供了可复用的方法基础，同时也提示预警模型需要在提前量、误报/漏报、实时性与鲁棒性之间进行系统权衡。

1.3.4 文献总结

综合国内外研究现状，围绕塔吊吊装安全的感知硬件布置、环境感知方法、碰撞检测与预警模型三大核心方向，现有研究已形成一定的技术积累，但仍存在以下关键不足，为本文研究提供了明确的切入点：

- 1.感知硬件部署聚焦不足：**现有传感器布置多集中于塔吊塔身、塔臂或施工现场固定位置，以宏观防碰撞为目标，对吊钩端这一最高风险区域的直接感知研究匮乏。虽有少数研究尝试在吊钩端部署传感设备，但存在硬件维护困难、抗干扰能力弱、感知范围有限等问题，难以满足近场全方位、高精度感知需求。
- 2.环境感知方法存在局限：**单一传感器感知存在明显短板（如视觉受光照影响、雷达语义匮乏），多传感器融合成为主流趋势，但现有融合方案或过于依赖复杂深度学习模型导致实时性不足，或仅停留在数据拼接层面未能充分发挥互补优势。针对施工场景的雷视融合研究，尚未形成兼顾几何精度、语义识别与工程鲁棒性的成熟方案，尤其缺乏针对吊钩端动态特性的适配优化。
- 3.预警模型适配性不足：**现有碰撞检测算法多源于虚拟仿真或路径规划场景，依赖静态几何模型，难以应对施工现场动态干扰多、障碍物复杂的实际情况；预警机制多基于单一

空间距离阈值，未充分结合目标运动趋势与时间裕度，易出现误报或预警滞后问题。同时，对危险目标（如施工人员、车辆等）的语义关联不足，难以实现精准分级预警。

1.4 主要研究内容

围绕塔吊吊装作业中吊钩端近场风险高、作业视野受限、环境遮挡与光照变化显著等问题，本文以“雷达提供三维几何测距与空间约束、视觉提供语义识别与解释”为总体思路，构建面向工程落地的吊钩端多传感器感知与分级预警方法。全文研究内容按照“系统构建与标定—融合点云生成与定位建图—目标提取与跟踪—雷视融合与风险判定—实验评估”的主线展开，各部分之间的逻辑关系如图1.2所示。

本文围绕“基于雷视融合的塔吊吊装过程动态避障与预警方法研究”这一主题，面向塔吊吊钩端近场作业过程中“空间狭窄、遮挡复杂、目标动态性强、操作视野受限”等典型特征，构建以激光雷达空间测距为主、视觉语义识别为辅的在线碰撞风险判定与分级预警方法体系。具体研究内容包括：

（1）吊钩端传感器系统构建

针对吊钩动态作业特性，完成了多传感器硬件集成与一体化防护设计，包括双激光雷达与工业相机选型与布局优化。基于 ROS 构建通信链路，实现多传感器标定（内参、外参）、时间同步与延迟补偿，建立统一坐标系与 TF 体系，为后续处理提供可靠数据基础。

（2）基于融合点云的定位与空间危险区域检测方法

基于双雷达融合点云，实现吊钩端精准定位与危险区域感知。采用 LIO-SAM 完成 SLAM 定位与静态地图构建，通过点云预处理与时序分析分离静态点云，结合聚类与卡尔曼滤波跟踪动态目标。构建“吊物尺寸+圆柱并集”动态危险区域模型，通过交集检测与连贯性验证输出危险区域信息。

（3）基于雷视融合的避障预警策略

融合激光雷达几何信息与视觉语义检测，提出分级预警方法。通过视觉检测时序稳定化处理，设计双向一致性融合策略实现目标轨迹级关联。针对静态障碍采用距离阈值预警，对动态目标构建“距离-时间”双维度风险评估模型，结合 TTC 与轨迹预测实现三级分级预警，提升预警准确率与鲁棒性。

在上述研究内容的支撑下，本文的组织结构如下：第 1 章阐述研究背景、意义与国内外研究现状；第 2 章给出吊钩端多传感器系统的总体架构、硬件选型与布置方案，并完成相机、双雷达及雷达—相机外参标定与数据链路规范；第 3 章提出双雷达点云实时融合方法，并引入 LIO-SAM 完成吊钩端位姿估计与静态地图构建，在此基础上实现 ROI 预处理、静态分离、动态目标聚类与跟踪；第 4 章针对高风险动态目标引入视觉检测结果，提出雷视双向一致性关联与置信度融合方法，并构建结合距离与时间裕度的分级预警策略；第 5 章搭建实验平台并设计典型工况，对系统定位、协同感知与预警效果进行验证与评估；第 6 章总结全文工作并给出展望。

1.5 技术路线

本文技术路线可概括为：首先在吊钩端集成双激光雷达与工业相机，建立供电、通信与 ROS 数据链路，完成多传感器标定与时间同步，形成统一坐标系与 TF 关系；其次，基于双雷达外参实现点云坐标统一与实时融合，并以融合点云与 IMU 数据为输入，通过 LIO-SAM 完成吊钩端位姿估计与静态地图构建；随后，基于位姿将连续帧点云进行时空对齐，在近场 ROI 内执行滤波、下采样与离群点抑制，并结合静态地图完成静动态分离与动态点云簇提取，进一步通过短时跟踪得到目标位置、速度等状态量；最后，引入视觉侧危险目标检测结果（实验中以车辆为代表），进行

装

订

线

雷视双向一致性关联与轨迹级语义增强，构建基于距离阈值与 TTC 等时间裕度指标的风险评估与分级预警状态机，输出可解释的预警等级与风险目标信息，为现场作业提供主动式安全辅助。

技术路线图：系统构建与标定 → 融合点云生成与定位建图 → 目标提取与跟踪
→ 雷视融合与风险判定 → 实验评估

图 1.2 本文技术路线示意图

装

订

线

2 吊钩端传感器系统构建与配准

本章围绕吊钩端近场危险区域感知的工程化实现，给出多传感器系统的总体架构、硬件集成与标定配准方案，并形成可供后续算法模块直接调用的数据接口与坐标系规范。由于后续的点云融合、定位建图与雷视关联均依赖传感器外参稳定性与跨模态时空一致性，因此本章的核心目标是将“可部署”的传感器平台转化为“可计算”的统一观测空间。

具体而言，第 2.1 节给出吊钩端硬件选型与布置原则，说明在有限空间与强振动工况下如何实现视场覆盖与结构可靠性；第 2.2 节分别建立相机内参、双雷达外参与雷达-相机外参，确保点云与图像的几何一致性；第 2.3 节进一步规范通信链路、时间同步与 TF 接口，以减少链路抖动与时延带来的融合误差，为第 3 章的融合点云定位与检测算法提供稳定输入。

2.1 系统总体架构与硬件选型布置

2.1.1 传感设备选型

吊钩端感知系统需要在有限安装空间内覆盖吊钩周围近场危险区域，并在粉尘、强光/弱光、遮挡与振动等工况下保持稳定输出。为兼顾几何测距精度、环境鲁棒性与语义识别能力，本文采用“双 LiDAR + 工业相机”的异构配置：

双激光雷达（Livox Mid-360）作为近场空间几何感知主传感器。LiDAR 具备 360° 水平覆盖能力，对光照变化不敏感，可在尘土等不利条件下保持稳定测距输出。采用双雷达布置可通过不同俯仰角与基线配置实现盲区互补，并提升关键方向的点云密度与冗余度。

工业相机（海康威视 MV-CS016-10UC V5，镜头 MVL-HF(X)28M-6MPE，1/1.8”、600 万像素）用于提取危险目标（以车辆为代表）的语义信息。视觉信息为预警提供可解释的目标类别与可视化依据，与 LiDAR 的几何测距结果形成互补，使风险判定既具备空间约束也具备语义可解释性。

系统的 IMU 信息来源于 Mid-360 内置 IMU，用于 LIO-SAM 中的短时运动约束与点云去畸变，从而在吊钩端摆动与回转工况下维持位姿估计与点云时序对齐的稳定性。

为便于工程表达，表 2.1 给出本系统关键传感器及其在方法体系中的作用说明。

表 2.1 吊钩端感知系统关键组件与作用

组件	作用	关键说明
LiDAR-1（Mid-360）	近场几何感知	作为主测距源，提供近场点云，用于 ROI 建模、距离阈值判定与目标几何状态估计
LiDAR-2（Mid-360）	补盲与冗余	与 LiDAR-1 形成盲区互补，提高关键方向点云覆盖与稳定性，支撑融合点云生成
工业相机 (MV-CS016-10UC V5)	语义识别	检测危险目标（以车辆为代表）并输出 2D 检测框与置信度，用于雷视关联与语义增强
镜头（MVL- HF(X)28M-6MPE）	成像质量	满足作业距离下的分辨率需求，提升弱光/逆光条件下可用性
IMU（内置）	运动约束	为位姿估计提供短时高频运动信息，辅助点云去畸变与定位稳定性

2.1.2 安装布局优化

双 LiDAR 的布置遵循“盲区互补 + 结构遮挡规避 + 振动冲击可控”的原则。考虑吊钩端作业时吊钩块与吊具、钢丝绳及吊物本体会对传感器视场造成遮挡，本文采用两台 LiDAR 在箱体两侧/上下错位布置，并设置不同俯仰角：一台侧重覆盖吊钩下方与近距离区域，另一台侧重覆盖

侧向与远端区域，从而在吊钩摆动与回转工况下仍能保持对 3~5 m 近场范围的连续观测。

在具体布置角度上，两台 Mid-360 采用**左右对称**安装，并将扫描平面相对于水平面设置为约 37° **俯仰角**，以增强对吊钩下方近场区域的覆盖并降低吊钩块/吊具对水平视场的遮挡影响。对称布置有助于在左右遮挡不均或局部反射异常时提供冗余观测，从而提升融合点云的连续性与稳定性。

工业相机与 LiDAR 共同面向主要作业区域布置，视场方向尽量与 LiDAR 的高密度点云覆盖区重叠，以提高“**点云候选—图像验证**”的有效关联率。安装结构上通过刚性支架与减振连接件降低高频振动对外参稳定性的影响，并通过限位设计避免传感器在摆动极值位置与吊钩结构发生碰撞。

相机采用**垂直向下**的安装方式，对吊钩下方作业面进行俯视拍摄。一方面该视角与“危险目标（如车辆/人员等）接近吊钩端”的风险模式更一致，便于在遮挡较少的区域获得目标图像证据；另一方面也有利于在预警输出中提供直观可解释的可视化依据。为避免镜头被吊具遮挡或被扬尘污染，相机安装位置与窗口开孔在结构设计阶段进行约束，并预留必要的防护与清洁维护空间。

从工程约束角度看，上述布置主要服务于四类目标：**视场覆盖**（双雷达不同俯仰方向的覆盖互补以减小遮挡盲区）；**外参稳定**（刚性安装基准 + 防松限位抑制振动引起的微小位姿漂移）；**安装安全与可靠性**（在摆动极值与冲击条件下避免与吊钩结构干涉并降低设备损伤风险）；**供电与走线**（尽量缩短外露线缆长度，线束固定与应力释放满足吊钩端长期振动工况）。

2.1.3 吊钩端一体化箱体设计（SolidWorks）

为满足施工现场工程化部署需求，吊钩端感知系统采用一体化箱体进行集成，重点解决**振动冲击、防尘防水、线缆可靠性、供电与通信稳定性**等问题。箱体结构上为传感器提供刚性安装基座并设置防松与限位结构，降低长期运行导致的相对位姿漂移；外壳采用密封与防尘设计，关键接口使用防水接头与应力释放结构，减少线缆在摆动与摩擦环境下的疲劳破坏风险。

在系统集成方面，将传感器、供电转换模块与计算/通信模块集中布置，减少外露走线长度；网络侧通过统一交换/桥接方式管理相机与雷达的数据流，算法侧基于 ROS 建立标准话题发布与 TF 广播，实现数据采集、融合、定位、识别与预警模块的解耦与可维护性。箱体内部预留散热与维护空间，便于现场更换传感器或调整安装角度。

为实现上述集成目标，本文基于 SolidWorks 完成吊钩下方传感器平台的结构与装配验证，其设计流程可概括为：**需求定义（视场覆盖、抗振防护、供电与走线、维护更换）→ 总体方案（上/下盖两体式箱体与传感器布置）→ 细化建模（安装基准、窗口开孔、接口位置、紧固与限位结构）→ 装配与干涉检查（摆动极值与可维护空间）→ 布局与可靠性校核（重心、线束应力释放、散热通道）**。该流程保证结构与算法所需“**外参稳定、时空对齐一致性**”在工程层面能够长期维持。

在结构实现上，吊钩端平台采用**上盖 + 下盖**的两体式箱体方案：下盖承担传感器与支架的主要安装基准，上盖用于防护与密封，并与下盖形成封闭腔体以容纳电源与计算单元。箱体内部集成**电池、NVIDIA Jetson 边缘计算模块、走线与固定件**等部件，所有供电与信号线缆均在箱体内部完成布线与应力释放，外部仅保留必要的充电/调试接口，从而降低施工现场线缆外露导致的钩挂、磨损与进水风险。

内部布局遵循“**重心靠近吊点、线束最短路径、热源隔离与可维护**”的原则：电池与边缘计算模块布置在箱体内部，尽量靠近结构中心区域，以减小摆动时的惯性矩并降低偏载；线束采用分区走线与扎带/卡扣固定，关键连接处增加余量并布置应力释放结构，避免长期振动导致接插件

松动；对 Jetson 等发热部件预留散热通道与维护空间，以满足长时间在线运行的稳定性要求。该一体化平台使得传感器布置、外参稳定性与工程防护在结构层面协同设计，为后续定位建图、雷视关联与预警算法提供可靠的数据输入基础。

为说明一体化盒体的结构构型与关键约束，本文在本小节给出平台外观、装配关系与内部布局的示意图位置（见图2.1-图2.3）。当前以占位框表示，后续将以实际渲染图或工程图替换，以完整呈现结构设计细节。

SolidWorks 吊钩下方传感器平台整体外观渲染图（占位）

图 2.1 吊钩下方一体化传感器平台整体外观示意（SolidWorks）

平台上盖、下盖与主要部件装配关系示意图（占位）

图 2.2 平台上盖、下盖与主要部件装配关系示意（SolidWorks）

为进一步说明盒体内部的空间布局与走线组织方式，本节同样预留“内部布局示意图”位置，如图2.3所示。该示意图可包含电池与 Jetson 的安装位置、线束走向与应力释放点、传感器接口位置以及散热与维护开口等关键信息，用于说明结构空间约束与可维护性设计。

盒体内部布局与走线组织示意图（占位）

图 2.3 盒体内部布局与走线示意（SolidWorks）

2.2 多传感器外参标定

2.2.1 相机内参标定

相机内参标定用于获取焦距、主点与畸变参数，保证图像去畸变与点云到图像投影的几何一致性。本文采用平面棋盘格标定板采集多组不同距离、不同姿态的标定图像，基于张正友标定法求解相机内参矩阵 \mathbf{K} 及径向/切向畸变系数。标定完成后，使用重投影误差对标定质量进行评价，并将去畸变后的图像作为后续危险目标检测（以车辆为代表）与雷视关联的输入。

相机内参标定以针孔成像模型描述三维点到像素点的投影关系。设空间点在相机坐标系下为 $\mathbf{P}_c = [X \ Y \ Z]^T$ ，其在归一化成像平面上的坐标为 $\mathbf{p}_n = [x \ y]^T$ ，则

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (2.1)$$

其中 (u, v) 为像素坐标， s 为尺度因子，相机内参矩阵 \mathbf{K} 定义为

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (2.2)$$

式中 f_x, f_y 为像素单位的等效焦距， (c_x, c_y) 为主点坐标。若以棋盘格平面作为世界坐标系 $\{w\}$ 的

参考，则相机外参（标定时每张图像对应一组）可写为旋转矩阵 \mathbf{R} 与平移向量 \mathbf{t} ，有

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K} [\mathbf{R} | \mathbf{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (2.3)$$

张正友标定法利用平面标定板在不同姿态下产生的多组单应性（Homography）约束，从而在无需精密三维标定场的条件下同时估计 \mathbf{K} 与畸变参数，工程上具有实现简单、精度可靠等优点。

实际镜头通常存在畸变，本文采用径向畸变与切向畸变的组合模型。令 $r^2 = x^2 + y^2$ ，则畸变后的归一化坐标 (x_d, y_d) 可表示为

$$\begin{cases} x_d = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 xy + p_2(r^2 + 2x^2), \\ y_d = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2y^2) + 2p_2 xy, \end{cases} \quad (2.4)$$

其中 k_1, k_2, k_3 为径向畸变系数， p_1, p_2 为切向畸变系数。将 (x_d, y_d) 代入式公式 (2.1) 即可得到带畸变的像素坐标预测值。

在实现层面，本文相机内参标定使用 MATLAB（Computer Vision Toolbox）完成，采用棋盘格标定板进行数据采集与参数求解。为便于论文复现与工程落地，标定流程按“数据采集—角点提取—参数估计—误差评估—参数导出”组织：选用平面棋盘格标定板并记录方格边长 L （mm），在固定焦距与曝光等关键参数的前提下采集不少于 15 ~ 25 张多姿态图像，覆盖不同距离、不同俯仰/偏航角以及不同画面位置，避免样本长期集中在图像中心或姿态变化过小。图2.4为采集方式示意。

相机内参标定的数据采集示意图（占位）

图 2.4 相机内参标定的数据采集示意（MATLAB）

角点提取阶段在 MATLAB 中使用 cameraCalibrator 应用（或调用 detectCheckerboardPoints）自动检测棋盘角点，核验每张图像角点是否完整、是否存在误检，并剔除模糊、强反光或角点缺失严重的样本。角点检测效果与样本覆盖范围示意图2.5。

棋盘格角点检测与样本覆盖示意图（占位）

图 2.5 棋盘格角点检测与样本覆盖示意

参数求解阶段在 MATLAB 中设置方格边长 L ，调用 estimateCameraParameters 估计 \mathbf{K} 与畸变系数，并自动给出每张图像的外参。完成标定后，可对原图进行去畸变（undistortImage），检查直线结构（如门框、墙角、标定板边缘）是否得到明显矫正，以验证畸变模型合理性。

标定质量评价采用重投影误差衡量标定精度。设第 i 张图像上第 j 个角点的观测像素为 \mathbf{p}_{ij} ，模型预测像素为 $\hat{\mathbf{p}}_{ij}$ ，则单点重投影误差为 $\|\mathbf{p}_{ij} - \hat{\mathbf{p}}_{ij}\|_2$ ，总体均方根误差（RMSE）定义为

$$e_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{n_i} \|\mathbf{p}_{ij} - \hat{\mathbf{p}}_{ij}\|_2^2}, \quad (2.5)$$

其中 M 为图像数量， n_i 为第 i 张图像可用角点数， $N = \sum_i n_i$ 为总角点数。MATLAB 通常同时给出每张图像的重投影误差分布，便于发现离群样本并迭代剔除。误差统计与可视化结果在文中

预留，如图2.6所示。

相机标定重投影误差可视化结果（占位）

图 2.6 相机内参标定的重投影误差评价示意

最终标定输出包括内参矩阵 \mathbf{K} 、畸变系数 $\{k_1, k_2, k_3, p_1, p_2\}$ 以及每张标定图像对应的外参 $\{\mathbf{R}_i, \mathbf{t}_i\}$ 。本文将 \mathbf{K} 与畸变参数用于图像去畸变与“点云到图像”的投影计算，为后续雷达外参标定与跨模态关联提供统一的几何基础。

2.2.2 双雷达外参自动标定与精配准

双雷达外参标定目标是求解两台 Mid-360 坐标系之间的刚体变换 \mathbf{T}_2^1 ，以便将两路点云统一到同一参考系。本文采用“粗配准 + 自动精配准”的两阶段流程：首先依据箱体结构尺寸与安装姿态给出初值（粗配准），确保两路点云在同一量级的空间范围内重叠；随后采集同步点云并进行自动优化，对外参进行精配准，使得两路点云在环境结构上的几何一致性最优。

本研究的自动精配准部分参考 Livox 开源的多雷达外参自动标定工具实现思路^①，其核心思想可概括为“基准雷达先建图，目标雷达逐帧与地图配准并迭代外参，最后用一致性方法得到稳定外参”。该类方法对外参初值不要求十分精确，但需要大致对齐；同时对数据同步与运动畸变较为敏感，因此采集时应尽量慢速运动、避免急转弯并保证两雷达帧间时间对齐。

记基准雷达坐标系为 $\{l_1\}$ ，目标雷达坐标系为 $\{l_2\}$ ，待求外参为

$$\mathbf{T}_{l_2}^{l_1} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \in SE(3), \quad (2.6)$$

其中 $\mathbf{R} \in SO(3)$ ， $\mathbf{t} \in \mathbb{R}^3$ 。自动标定采用“先建图后配准”范式：

该范式包含“建图约束”与“迭代配准”两类关键环节：在固定参考系 $\{w\}$ 下用基准雷达的多帧点云构建局部/全局地图 \mathcal{M} ，以获得更稳定的环境几何表示；同时将目标雷达的每帧点云 $\mathcal{P}_k^{(2)}$ 通过当前外参估计变换到基准雷达/世界坐标下，并与地图 \mathcal{M} 进行配准，得到增量变换 $\Delta\mathbf{T}_k$ 用于修正外参。工程上常用 ICP 及其鲁棒变体进行点云配准；参考 Livox 工具实现，本文采用广义 ICP（Generalized ICP, GICP）思想，在点到点最近邻基础上引入局部协方差近似点到平面的约束，以提升稀疏或噪声场景下的稳定性，其典型优化形式为

$$\min_{\mathbf{T}} \sum_i (\mathbf{p}_i - \mathbf{T}\mathbf{q}_i)^\top \mathbf{C}_i^{-1} (\mathbf{p}_i - \mathbf{T}\mathbf{q}_i), \quad (2.7)$$

其中 \mathbf{q}_i 为源点云（目标雷达帧）中的点， \mathbf{p}_i 为地图中与其对应的匹配点（或局部子地图点）， \mathbf{C}_i 为由两侧局部几何估计得到的协方差（权重）矩阵。优化得到的配准变换用于更新外参估计，例如

$$\mathbf{T}_{l_2}^{l_1} \leftarrow \Delta\mathbf{T}_k \mathbf{T}_{l_2}^{l_1}. \quad (2.8)$$

为避免错误匹配导致外参发散，工程实现通常设置配准质量门控（如 fitness score 阈值、最大对应距离、迭代次数上限等），对不可信帧跳过更新或回退到初始猜测。逐帧配准会得到外参估计序列 $\{\mathbf{T}_k\}$ ，其受动态点、遮挡、运动畸变与局部几何退化影响可能出现离群值；参考 Livox 工具的做法，可在平移/欧拉角参数空间对该序列进行稳健拟合与一致性筛选，仅保留配准质量较高

① Livox Automatic Calibration Tools: https://github.com/Livox-SDK/Livox_automatic_calibration

的帧，并通过共识机制（如 RANSAC）抑制离群点，最终得到稳定外参 $\hat{\mathbf{T}}$ 。该步骤利用“外参为常量”的刚性约束，使估计结果在长序列数据上更稳定、可复现。

结合本文吊钩端系统特点与 Livox 工具的要求，双雷达自动标定在工程上可抽象为“数据采集与帧对齐—粗外参初始化—基准建图—目标逐帧配准迭代—一致性筛选与输出”的离线流水线。数据采集应选择结构稳定且环境几何丰富（存在平面、立柱、墙角等）的场地，并尽可能保证两雷达时间对齐；移动采集时宜慢速运动以降低单帧运动畸变，必要时在离线阶段剔除动态目标或仅选取静态时间段。本文采用 ROS 录制两台 LiDAR 的 rosbag 数据，并在离线阶段按固定时间片（例如每 100 ms）聚合为点云帧导出为 .pcd，同时确保两路点云在帧序号与时间戳层面尽可能一一对应并保持统一的读取顺序与目录组织方式（例如分别存放于 Base_LiDAR_Frames/ 与 Target_LiDAR_Frames/）。数据整理与帧对齐的示意位置在文中预留，如图 2.7 所示。

双雷达标定数据集整理与帧对齐示意图（占位）

图 2.7 双雷达外参自动标定的数据准备与分帧同步示意

粗配准外参 \mathbf{T}_0 可由吊钩端盒体的结构设计模型（安装位姿与孔位基准）计算得到，并在工程实现中写入配置文件（如 Init_Matrix.txt）作为初始猜测参与逐帧配准更新。该初值不必十分精确，但应保证两路点云在同一场景下具有足够重叠，以提高配准收敛到正确解的概率。随后，以基准雷达帧序列构建地图 \mathcal{M} 作为后续配准参考；为提升建图质量，可对点云进行体素下采样、离群点剔除等预处理，并在必要时对帧间运动畸变进行校正。在外参迭代阶段，将目标雷达每帧点云在当前外参下变换后与地图（或局部子地图/近邻点集合）进行 GICP 配准，得到增量修正 $\Delta \mathbf{T}_k$ 并更新外参，同时记录每帧配准质量指标（如 fitness score）以支撑一致性筛选。标定过程界面及前后点云叠加效果在文中预留，如图 2.8 所示。

双雷达点云配准叠加效果（标定前/后对比，占位）

图 2.8 双雷达自动标定的点云配准效果示意（前后对比）

完成逐帧配准后，对外参估计序列进行稳健筛选与拟合，输出最终外参 $\hat{\mathbf{T}}_{l_2}^{l_1}$ （可用旋转/平移、欧拉角或四元数表示）。同时可输出标定质量统计（例如有效帧比例、fitness score 分布）用于可解释性验证；相关统计图在文中预留，如图 2.9 所示。

自动标定质量评估与收敛统计结果（占位）

图 2.9 自动标定过程的质量评估与收敛示意

为确保自动标定结果可靠，工程部署需保证双雷达数据尽可能同步且帧对应关系稳定，采集运动尽量缓慢并减少急转弯以降低运动畸变带来的配准偏差，环境结构保持足够丰富以避免几何退化，同时粗外参需实现“大致对齐”以提高 GICP 收敛到正确极值的概率。最终得到的 $\hat{\mathbf{T}}_{l_2}^{l_1}$ 将用于第 3 章的点云实时融合与坐标统一。

2.2.3 雷达-相机外参标定

雷达-相机外参标定用于建立点云到图像的投影关系，是第 4 章雷视关联的基础。本文采用“同一标定目标在雷达坐标系下的 3D 坐标 + 在相机像素平面下的 2D 坐标”构建约束，通过最小化重投影误差求解外参。该流程参考 Livox 开源工具^①的思路：使用标定板角点作为跨模态对应特征，并通过非线性优化获得稳定外参；标定完成后再通过点云投影或点云着色等方式进行结果验证。

① Livox Camera-LiDAR Calibration: https://github.com/Livox-SDK/livox_camera_lidar_calibration

外参求解可表述为 3D-2D 重投影误差最小化问题。设 LiDAR 坐标系为 $\{l\}$ ，相机坐标系为 $\{c\}$ ，待求外参为

$$\mathbf{T}_l^c = \begin{bmatrix} \mathbf{R}_l^c & \mathbf{t}_l^c \\ \mathbf{0}^\top & 1 \end{bmatrix} \in SE(3). \quad (2.9)$$

对于 LiDAR 中观测到的角点三维坐标 $\mathbf{P}_l = [X_l \ Y_l \ Z_l]^\top$ ，其在相机坐标系下为

$$\mathbf{P}_c = \mathbf{R}_l^c \mathbf{P}_l + \mathbf{t}_l^c = [X_c \ Y_c \ Z_c]^\top. \quad (2.10)$$

结合 2.2.1 节求得的相机内参 \mathbf{K} （以及畸变模型），像素坐标预测值可写为

$$\hat{u} = f_x \frac{X_c}{Z_c} + c_x, \quad \hat{v} = f_y \frac{Y_c}{Z_c} + c_y, \quad (2.11)$$

其中 (\hat{u}, \hat{v}) 为由外参 + 内参得到的理论投影点。若以观测像素为 (u, v) ，则单个对应点的重投影残差为

$$\mathbf{r} = \begin{bmatrix} \hat{u} - u \\ \hat{v} - v \end{bmatrix}. \quad (2.12)$$

对多个标定姿态与多组角点对应，构建总目标函数

$$\min_{\mathbf{R}_l^c, \mathbf{t}_l^c} \sum_k \sum_i \|\mathbf{r}_{k,i}\|_2^2, \quad (2.13)$$

并采用非线性最小二乘（如 Ceres Solver 的 Levenberg-Marquardt）求解。为避免欧拉角奇异与保证旋转约束，工程实现中常用四元数 \mathbf{q} 表示旋转，并在优化中施加单位四元数局部参数化。

Livox 工具中提供两种求解模式：**getExt1** 仅优化外参（固定内参），适用于相机内参可信的情况；**getExt2** 可在数据量更充分时同时优化内参中关键分量与外参，用于当重投影误差长期偏大且怀疑内参存在偏差时的补救。

结合本文“先 rosbag 录制、再离线处理”的工程实践，雷达-相机外参标定可概括为：在相对开阔、背景干扰较少的场地布置大尺寸平面标定板（例如 $1\text{ m} \times 1.5\text{ m}$ ），采集至少 10 组不同距离与姿态的数据，并保持相机与雷达到标定板距离在数米量级（例如约 3 m）以获得更稳定的角点定位精度；对每组姿态保存清晰图像并记录四角点像素坐标 (u, v) ，同时采用统一且一致的角点编号顺序（如从左上角开始按顺时针/逆时针固定编号）以避免对应关系错误；将包含标定板的 LiDAR 数据离线转换为 PCD（或直接从 rosbag 读取点云帧），在点云中提取标定板四角点三维坐标并与图像侧编号保持一致，从而构造稳定的 3D-2D 对应。随后，将结构设计或安装测量得到的外参初值写入配置文件，优先使用 getExt1（固定内参，仅优化外参）求解 \mathbf{T}_l^c ，并通过迭代代价（cost）与最终重投影误差对异常数据进行剔除后再优化。最后，将求得外参用于点云投影到图像以验证对齐一致性，也可进一步进行点云着色并在可视化工具中检查融合效果。文中相应示意位置如图 2.10-图 2.13 所示。

雷达-相机外参标定的场景与标定板摆放示意图（占位）

图 2.10 雷达-相机外参标定的场景与标定板摆放示意

相机侧角点标注与像素坐标读取示意图（占位）

图 2.11 相机侧角点像素坐标获取示意

雷达侧标定板角点三维坐标提取示意图（占位）

图 2.12 雷达侧角点三维坐标获取示意

点云投影到图像的对齐验证效果（占位）

图 2.13 雷达点云投影到图像的外参验证示意

通过上述流程获得的 \mathbf{T}_l^c 将在后续雷视融合中用于 3D 点云投影门控（3D→2D）与一致性验证（2D→3D），保证雷达几何信息与视觉语义信息在同一观测空间中可对齐与可解释。

2.3 数据链路与接口规范

2.3.1 通信链路搭建

为保证吊钩端多传感器数据的稳定传输与可控时延，本文采用“独立接入、边缘侧统一采集与转发”的数据链路组织方式：双 LiDAR 通过以太网接入边缘计算单元（Jetson/工控机等），工业相机通过 USB3.0 接入同一计算单元；计算单元在本地完成驱动接入、数据时间戳标注与必要缓存，再通过 ROS 话题发布到系统内部总线供定位、检测与融合模块订阅。

吊钩端多传感器通信链路与数据流拓扑示意图（占位）

图 2.14 吊钩端多传感器通信链路与数据流拓扑示意

在工程实现中，链路稳定性通常受制于**瞬时带宽**、**CPU 解码/拷贝开销**与**缓存积压**。为降低丢包与突发拥塞带来的随机延迟，边缘计算侧应优先保证网络地址与子网划分清晰可控（以太网 LiDAR 固定 IP，必要时独立子网以减少广播与地址冲突），并通过合理设置网卡接收缓冲与驱动线程调度降低丢包概率，同时避免过大的用户态缓存导致端到端延迟不可控。链路容量可用吞吐核算进行约束：以话题 i 为例，其平均吞吐为

$$\bar{B}_i = f_i S_i \quad (2.14)$$

其中 f_i 为发布频率， S_i 为单帧消息大小；系统总吞吐 $\bar{B} = \sum_i \bar{B}_i$ 应留出裕量以覆盖突发峰值与协议开销。

在双 LiDAR 以太网直连边缘计算单元的场景中，常见工程约束是**两台雷达默认处于同一网段**（如 192.168.2.x），且雷达侧 Host IP 固定指向 192.168.2.1。在不修改雷达侧网络配置的前提下，若主机通过两块物理网卡分别直连两台雷达，应避免在两块网卡上同时配置同一个 Host IP，否则会造成地址绑定与二层转发的不确定性并引发丢包。本文采用**软件桥接（bridge）**将两块物理口接入同一二层广播域：创建桥接接口 br0 并将**唯一的** 192.168.2.1/24 配置在 br0 上，使主机对外呈现为单一逻辑接口。工程落地时需保证 IP 仅配置在 br0、链路不引入额外路由/网关路径、网络管理服务不自动重写地址/路由，并在日志中记录链路状态与丢包统计以支撑上层延迟补偿与告警。

ROS1 话题通信采用发布/订阅机制，若生产端速率高于消费端处理能力，队列积压会直接转化为延迟增长并破坏融合一致性。因此本文对关键话题采用“**小队列 + 可丢帧**”策略，使系统尽可能运行在低延迟稳态。设订阅端的平均处理时延为 T_p 、队列长度为 Q ，则在持续拥塞时队列引入的额外排队时延近似为

$$T_q \approx (Q - 1) T_p \quad (2.15)$$

因此对点云、图像等大消息应避免设置过大的 Q 。

为统一工程接口并降低模块耦合，本文将关键数据流抽象为“原始采集话题”和“处理后输出话题”。建议的最小接口集合如表2.2所示（命名可按实际工程替换，原则是保持 `frame_id` 与 TF 一致）。

表 2.2 吊钩端系统建议的最小 ROS 话题接口集合

数据	消息类型	说明
点云（LiDAR1/2）	<code>sensor_msgs/PointCloud2</code>	含时间戳与 <code>frame_id</code> ；必要时包含点时间字段用于去畸变
IMU（内置）	<code>sensor_msgs/Imu</code>	角速度/加速度；时间戳用于 LIO 预积分与去畸变
相机图像	<code>sensor_msgs/Image</code>	原始或压缩传输；与相机内参 <code>CameraInfo</code> 配套
相机内参	<code>sensor_msgs/CameraInfo</code>	由相机内参标定结果写入并随图像发布
TF/静态外参	<code>tf2_msgs/TFMessage</code>	运行时发布静态/动态 TF，保证跨模块坐标统一

2.3.2 时间同步与延迟补偿

多传感器系统中，采样频率、驱动缓存、USB/以太网传输与解码会引入不同程度的时间偏差与延迟。若直接以到达时刻进行融合，容易产生目标关联漂移、点云投影错位与速度估计失真。为此，本文以 **ROS 时间戳** 为统一时间基准，结合“同步对齐 + 延迟估计 + 融合门控”的策略实现跨传感器时间一致性。

多传感器时间同步与延迟补偿时序示意图（占位）

图 2.15 多传感器时间同步与延迟补偿时序示意

针对以太网 LiDAR，除基于主机时间戳的“软件对齐”外，硬件/网络层的统一时钟能够从源头降低跨设备时间漂移。Livox 官方文档给出了多种时间同步方式及其适用场景（见https://livox-wiki-cn.readthedocs.io/zh-cn/latest/tutorials/new_product/common/time_sync.html），其中在无 GPS 与 PPS 硬件接入的情况下，PTP/gPTP 通过网络中单一 master 时钟设备，基于 IEEE 1588v2 或 IEEE 802.1AS（gPTP）实现 LiDAR 与其他设备间的时间同步，外部硬件要求较低且适合双 LiDAR 同网段部署；当系统中其他传感器已由 GPS 信号同步时，也可将 PPS 秒脉冲与 GPS 时间信息（GPRMC/GNRMC）硬件接入 LiDAR，使其统一到 GPS 时钟基准。

在 PTP v2.0 中，Livox 设备作为 slave，与 master 通过 `Sync`、`Follow_Up`、`Delay_Req`、`Delay_Resp` 报文交互获得 t_1, t_2, t_3, t_4 ，据此估计链路传输时延与主从时钟偏移。根据文档给出的计算式，可写为

$$\text{Delay} = \frac{(t_4 - t_1) - (t_3 - t_2)}{2} \quad (2.16)$$

$$\text{Offset} = (t_2 - t_1) - \text{Delay} = \frac{(t_2 - t_1) + (t_3 - t_4)}{2} \quad (2.17)$$

当 Offset 收敛并维持在较小范围内时，LiDAR 点云时间戳即可视为与 master 时钟对齐。

时间同步工程化落地应避免在同一二层网络中出现多个 PTP master（文档指出 PTP 同步优先级最高，且不支持多主时钟环境）；若使用主机作为 PTP master，可检查网卡的软件/硬件时间戳能力（硬件时间戳通常更有利于同步精度与稳定性）。在 Linux 上可采用 `linuxptp`（`ptp4l`、`phc2sys` 等）建立并维护 PTP 同步关系，并将网卡硬件时钟（PHC）与系统 `CLOCK_REALTIME` 对齐，使 LiDAR 时间与主机系统时间保持一致；正文不展开命令细节，工程上以同步状态与长期漂移统计作为主要验收依据。同步状态既可通过抓包确认网络中存在 `Sync/Follow_Up` 报文来判断 master 是否工作，也可结合设备数据包头中的 `timestamp_type` 字段验证同步类型（文档说明 1 表示 PTP、2 表示 GPS，时间戳单位为 ns）。

需要强调的是：时间同步并不能替代上层的“时间戳对齐与延迟补偿”。本文在完成 LiDAR 侧的 PTP/GPS 同步（或无法同步时使用主机时间）后，仍在 ROS 融合侧通过最近邻/插值与固定延迟补偿来保证跨传感器数据在同一参考时刻对齐。

设 LiDAR 点云帧时间戳为 t_k^l ，相机图像时间戳为 t_m^c ，IMU 测量时间戳序列为 t_n^i 。融合与关联通常需要将不同频率数据对齐到同一参考时刻 t ：对给定 t ，可采用最近邻对齐选择满足 $|t^s - t|$ 最小的传感器测量作为对齐样本（ $s \in \{l, c, i\}$ ）；对于高频 IMU，当需要 t 时刻的测量值时，可在相邻两帧 (t_n^i, \mathbf{z}_n^i) 与 $(t_{n+1}^i, \mathbf{z}_{n+1}^i)$ 之间进行线性插值

$$\mathbf{z}^i(t) = \alpha \mathbf{z}_n^i + (1 - \alpha) \mathbf{z}_{n+1}^i, \quad \alpha = \frac{t_{n+1}^i - t}{t_{n+1}^i - t_n^i} \quad (2.18)$$

对于 LiDAR 点云，若驱动提供点时间（每点相对帧起始的偏移），则可在定位模块中结合 IMU 进行去畸变（deskew），以减小吊钩摆动与回转造成的运动畸变。

对相机与 LiDAR 等异构链路而言，整体时延可粗略分解为

$$\Delta = \Delta_{\text{acq}} + \Delta_{\text{tx}} + \Delta_{\text{decode}} + \Delta_{\text{proc}} \quad (2.19)$$

其中 Δ_{acq} 为曝光/扫描与驱动采集引入的延迟， Δ_{tx} 为传输延迟， Δ_{decode} 为解码与拷贝开销， Δ_{proc} 为消息进入 ROS 后的调度与回调处理开销。本文不在文中给出具体数值，而是将其作为可标定/可测量的系统参数，通过日志时间戳与统计分析得到估计 $\hat{\Delta}$ ，并在融合侧进行补偿：

$$\tilde{t} = t - \hat{\Delta} \quad (2.20)$$

其中 t 为原始时间戳， \tilde{t} 为补偿后的有效时间戳。

为避免跨帧错配，本文在融合关联中引入时间窗门控：对任意待关联的两条观测，要求其时间差满足

$$|t_a - t_b| \leq \tau \quad (2.21)$$

其中 τ 为时间窗阈值，由系统的稳态时延与允许的运动误差共同决定。工程上， τ 过大将引入误关联，过小则会造成有效数据被拒绝；因此需结合驱动发布频率、链路拥塞与处理耗时进行调参。

2.3.3 坐标系与 TF 接口规范

跨模块的空间一致性依赖于统一的坐标系定义与 TF 发布规范。本文以吊钩端参考坐标系 $\{h\}$ 作为系统统一空间参考（可对应 ROS 中的 `base_link/hook_link`），并为两台 LiDAR 坐标系 $\{l_1\}$ 、 $\{l_2\}$ 与相机坐标系 $\{c\}$ 建立刚体变换关系。外参标定得到的 $\mathbf{T}_{l_2}^{l_1}$ 、 \mathbf{T}_l^c 等变换在运行时通过 TF 发布，确保定位、点云融合、投影门控与距离计算在同一空间参考下进行。

吊钩端多传感器 TF 树结构示意图（占位）

图 2.16 吊钩端多传感器 TF 树与坐标系关系示意

为避免坐标系混用导致的投影翻转与距离计算错误，本文建议所有发布的 `PointCloud2` 与 `Image` 消息正确填写 `header.frame_id` 并与 TF 树中的坐标系名称完全一致；所有坐标系采用右手坐标系、长度单位统一为 `m`，旋转使用四元数或旋转矩阵表示以避免链路中多次欧拉角转换带来的歧义；相机坐标系 $\{c\}$ 与投影模型保持一致（通常 x 指向右、 y 指向下、 z 指向前），若驱动默认坐标系不同，应在 TF 中显式给出转换。

在运行时，外参应视为静态变换：

$$\mathbf{T}_h^{l_1}, \mathbf{T}_h^{l_2}, \mathbf{T}_h^c \text{ 为常量，且由标定结果给出} \quad (2.22)$$

动态位姿（例如 $\mathbf{T}_{\text{map}}^h(t)$ ）由定位模块（LIO-SAM 等）输出，并以一定频率发布。为了保证模块解耦与可维护性，本文建议：静态外参宜使用静态 TF 发布者（如 `static_transform_publisher` 或 `StaticTransformBroadcaster`）发布，以避免重复发布造成 TF 树抖动；定位模块仅负责发布动态 TF（如 `map→odom→hook_link`），感知与融合模块通过 TF 查询获得变换而不自行维护外参；同时应尽量保持变换链路简洁，融合时优先通过一条明确链路完成坐标变换，例如

$$\mathbf{p}^h = \mathbf{T}_l^h \mathbf{p}^l, \quad \mathbf{p}^c = \mathbf{T}_h^c \mathbf{p}^h \quad (2.23)$$

以减少多次组合带来的数值误差与实现复杂度。

统一的坐标系与 TF 接口规范使得定位、目标提取、雷视融合与预警模块能够在同一空间参考下协同工作，并为后续增加其他传感器或替换算法提供清晰的接口边界。

2.4 本章小结

本章面向吊钩端近场风险感知的工程部署需求，完成了多传感器系统的硬件集成与外参配准，并给出了通信链路、时间同步与 TF 接口的实现约束。通过相机内参、双雷达外参以及雷达-相机外参的统一建模与验证，系统获得了可用于跨模态融合的几何一致性；通过链路与时间戳规范化以及关键延迟补偿策略，系统具备了可控时延与可重复的时空对齐能力。上述工作为第 3 章的融合点云定位建图与危险区域检测提供了稳定、统一且可维护的数据输入基础。

3 基于融合点云的定位与空间危险区域检测方法

本章以“融合点云驱动的定位建图与目标生成”为核心，给出双雷达点云实时融合、基于LIO-SAM的位姿估计与静态地图构建，以及面向危险区域检测的点云标准化与静态目标生成方法。需要强调的是，点云预处理可分为“定位前端所需的轻量预处理”和“检测/风险判定所需的任务预处理”两类：定位建图侧更关注特征保真与时序一致性，通常仅进行必要的时间对齐、去畸变与基础滤波；而危险区域检测侧则会进一步执行ROI裁剪、下采样、离群点抑制与静态分离等任务相关处理。因此，本文采用“先以融合点云+IMU完成LIO-SAM定位，再在位姿约束下对点云做任务预处理”的流程，以兼顾定位稳定性与后续检测的鲁棒性。

3.1 双雷达点云实时融合方法

3.1.1 原始点云数据特征

双雷达点云融合的首要前提是对原始观测进行“可用性刻画”，即明确哪些数据特性会直接影响后续的外参变换、时间对齐与定位估计。本系统的原始点云由两台Livox Mid-360提供，工作场景为吊钩端近场风险感知，关心的是几米尺度范围内的结构轮廓与动态目标。因此，点云在近场具有较高的有效点密度与较宽的观测覆盖，这为后续的欧式聚类、最小距离计算与目标跟踪提供了几何分辨率基础；但这一优势仅在时间与坐标一致的条件下成立，否则密集观测会转化为更明显的“重复边界”“拖影”与伪动态。

吊钩端平台具有摆动、回转与加减速等典型强动态特征。扫描式LiDAR在一个扫描周期内会经历持续运动，使得同一帧点云天然包含帧内畸变；与此同时，吊钩块、吊具与被吊物会产生显著遮挡与强反射，导致局部点云缺失或出现离群点。上述事实意味着：定位侧若在输入端进行过强的几何滤波，可能破坏真实结构边界并削弱特征约束；而完全不处理则会放大配准残差中的噪声成分，进而影响位姿输出的稳定性。因此，本章采用“定位前端轻量预处理+位姿约束后任务预处理”的分层策略，仅将对定位必需的对齐与格式统一前置，把ROI裁剪、离群点抑制与静态分离等任务相关步骤后置。

为便于后续算法描述，记第 i 台雷达（ $i \in \{1, 2\}$ ）在其自身坐标系 $\{l_i\}$ 下的原始点云为

$$\mathcal{P}_{l_i}(t) = \left\{ (\mathbf{p}_j^{l_i}, \tau_j^{(i)}) \right\}_{j=1}^{M_i}, \quad (3.1)$$

其中 $\mathbf{p}_j^{l_i} = [x_j \ y_j \ z_j]^T$ 为点坐标， $\tau_j^{(i)}$ 为该点在扫描周期内的采样时间戳（相对或绝对时间均可，只要在系统内一致）。当平台存在角速度/线速度扰动时，同一帧内不同 τ 的点实际对应不同的传感器位姿，因此“帧内去畸变”与“两路点云的时间对齐”是双雷达融合与后续定位稳定性的前置条件。

双雷达异步采样可抽象为两组离散时间序列 $\{t_k^{(1)}\}$ 与 $\{t_m^{(2)}\}$ 。在实时系统中通常采用最近邻对齐：对参考时刻 $t = t_k^{(1)}$ ，选取

$$m^*(k) = \arg \min_m \left| t_k^{(1)} - t_m^{(2)} \right|, \quad (3.2)$$

若 $\left| t_k^{(1)} - t_{m^*}^{(2)} \right| \leq \varepsilon_t$ （ ε_t 为允许的最大时间差阈值），则认为两帧可用于融合；否则丢弃或等待后续匹配。该对齐策略的证据基础来自工程约束：系统必须在有限队列与有限延迟下输出稳定频率的融合点云，因此无法对两路数据做长时间缓冲与全局插值。其解释在于，最近邻匹配将“对齐误差”显式限制在 ε_t 内，使得后续IMU去畸变与配准残差的误差项具有可控上界。实际部署中

ε_t 的取值应同时满足输出实时性与融合误差可接受性，取值依据建议在实验中以“位姿抖动/地图清晰度/动态分离误检率”为指标进行验证【待补充： ε_t 的最终取值与验证方法】。本节对原始点云特性的形式化描述为下一节外参矩阵应用与坐标统一提供了必要的时间与空间前提。

3.1.2 外参矩阵应用与坐标统一

双雷达融合要将来自不同传感器坐标系的点统一到同一参考系下，而这一过程的核心是外参齐次变换矩阵的正确应用。设有两个右手直角坐标系 $\{a\}$ 与 $\{b\}$ ，外参标定给出的几何关系可表示为从 $\{a\}$ 到 $\{b\}$ 的刚体变换，其在齐次坐标下写为

$$\mathbf{T}_a^b = \begin{bmatrix} \mathbf{R}_a^b & \mathbf{t}_a^b \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad (3.3)$$

其中 $\mathbf{R}_a^b \in SO(3)$ 为旋转矩阵， $\mathbf{t}_a^b \in \mathbb{R}^3$ 为平移向量。为保证推导与工程实现一致，本章默认旋转采用绕当前自身坐标轴的内旋约定；在该约定下，若旋转按时间顺序依次执行，则总旋转等于各旋转矩阵按顺序右乘，即

$$\mathbf{R} = \mathbf{R}_1 \mathbf{R}_2 \cdots \mathbf{R}_n \quad (3.4)$$

该表达的意义在于将“标定得到的欧拉角/旋转序列”直接映射为可在程序中复现的矩阵乘法，避免因旋转顺序或坐标系约定不一致导致的系统性误差。

以绕 z 轴与 y 轴旋转为例，对应旋转矩阵可写为

$$\mathbf{R}_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.5)$$

$$\mathbf{R}_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (3.6)$$

在本系统的外参求解与使用中，旋转序列可按“先绕自身 z 轴旋转 180° ，再绕自身 y 轴旋转 -30° ，随后再次绕自身 y 轴旋转 -30° ”来表示。由于后两次旋转绕同一轴，可合并为单次 -60° 旋转，从而总旋转写为

$$\mathbf{R}_a^b = \mathbf{R}_z(180^\circ) \mathbf{R}_y(-60^\circ). \quad (3.7)$$

该化简的证据来自旋转群的封闭性：同轴旋转可通过角度相加合并。其工程解释在于减少计算与配置复杂度，并降低由于重复写入/读取外参导致的出错概率。

$$\mathbf{R}_z(180^\circ) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.8)$$

$$\mathbf{R}_y(-60^\circ) = \begin{bmatrix} \cos 60^\circ & 0 & -\sin 60^\circ \\ 0 & 1 & 0 \\ \sin 60^\circ & 0 & \cos 60^\circ \end{bmatrix} \quad (3.9)$$

$$\mathbf{R}_a^b = \begin{bmatrix} -\cos 60^\circ & 0 & \sin 60^\circ \\ 0 & -1 & 0 \\ \sin 60^\circ & 0 & \cos 60^\circ \end{bmatrix} \quad (3.10)$$

在完成前两步旋转后,沿当前自身坐标系的 x 轴负方向平移:169.204874mm(177.140239mm),记为 d 。

在自身坐标系下的平移向量为:

$$\mathbf{t}_{\text{local}} = \begin{bmatrix} -d \\ 0 \\ 0 \end{bmatrix} \quad (d > 0) \quad (3.11)$$

若平移向量在旋转后的局部坐标系中定义,则需通过旋转矩阵转换到参考坐标系:

$$\mathbf{t}_a^b = \mathbf{R}_a^b \mathbf{t}_{\text{local}} \quad (3.12)$$

其中:

$$\mathbf{t}_a^b = \begin{bmatrix} d \cos 60^\circ \\ 0 \\ -d \sin 60^\circ \end{bmatrix} \quad (3.13)$$

据此可得到坐标系 $\{a\}$ 到 $\{b\}$ 的齐次变换矩阵表达:

$$\mathbf{T}_a^b = \begin{bmatrix} \mathbf{R}_a^b & \mathbf{t}_a^b \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad (3.14)$$

将上述旋转与平移展开,可写为:

$$\mathbf{T}_a^b = \begin{bmatrix} -\cos 60^\circ & 0 & \sin 60^\circ & d \cos 60^\circ \\ 0 & -1 & 0 & 0 \\ \sin 60^\circ & 0 & \cos 60^\circ & -d \sin 60^\circ \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.15)$$

在系统实现中,上述齐次变换矩阵对应于双雷达外参(或雷达到吊钩端参考系外参)的具体数值。为保证工程可维护性与可复现性,本文将外参作为参数化配置项保存于融合模块的配置文件中,记为 $\mathbf{T}_{l_2}^{l_1}$ (将 2 号雷达点变换至 1 号雷达坐标系),并在运行时直接用于坐标统一与点云融合。

若系统同时维护吊钩端参考系 $\{h\}$ 与两雷达坐标系 $\{l_1\}, \{l_2\}$ 的外参链路,则应满足一致性

关系

$$\mathbf{T}_{l_2}^{l_1} = (\mathbf{T}_{l_1}^h)^{-1} \mathbf{T}_{l_2}^h, \quad (3.16)$$

其中 $\mathbf{T}_{l_i}^h$ 表示从 $\{l_i\}$ 到 $\{h\}$ 的齐次变换。本文在后续融合与定位流程中固定选择 $\{l_1\}$ 或 $\{h\}$ 作为融合参考系，并通过 TF 统一发布坐标变换关系，使得定位、目标生成与后续雷视融合均在同一空间参考下运行，从而降低实现复杂度并减少因多次链式组合带来的数值累积误差。

需要强调的是，外参误差会直接投影到融合点云的几何一致性上。即便平移误差较小，旋转误差也会随距离放大并表现为点云边界错位，从而影响后续特征提取与配准残差的稳定性。因此，外参在投入融合前应进行可视化或定量一致性检查，例如在静止场景下观察两雷达对同一结构的重叠程度，或统计重叠区域的最近邻距离分布【待补充：外参验证方法与阈值】。该结论为下一小节点云实时融合的滤波策略与参数上界提供依据。

3.1.3 点云实时融合与发布

在获得双雷达外参 $\mathbf{T}_{l_2}^{l_1}$ 后，系统以 $\{l_1\}$ 或吊钩端参考系 $\{h\}$ 作为融合参考坐标系。运行时，对两路点云按时间戳进行对齐，并将 $\mathcal{P}_{l_2}(t)$ 通过 $\mathbf{T}_{l_2}^{l_1}$ 变换后与 $\mathcal{P}_{l_1}(t)$ 合并得到融合点云 $\mathcal{P}_f(t)$ 。为降低重复观测与噪声影响，融合后对点云进行体素下采样与离群点抑制，并可按需求剔除吊钩结构的固定点云区域。

融合点云的实时发布不仅是“把两帧点云拼接在一起”，更关键的是在延迟受控的条件下维持时间语义一致、坐标语义一致与噪声语义可控。系统实现上，驱动层使用支持多话题发布的模式同时接入两台 Mid-360，并分别输出各自点云流 $\mathcal{P}_{l_1}(t)$ 与 $\mathcal{P}_{l_2}(t)$ ；融合层由 livox_merge 节点完成对齐、变换与合并，并发布融合点云 $\mathcal{P}_f(t)$ 供定位与后续任务使用【待补充：实际 ROS 话题名与帧名】。

时间对齐策略沿用上一节的最近邻匹配思想，其论点是将跨雷达的异步误差显式限制在 ε_t 内，从而避免“融合帧时间含混”导致的配准残差漂移。其证据来自系统实时性约束：融合节点必须在有限缓存与稳定频率下输出，因此将队列长度与最大等待时间设置为上限，并在超过上限时选择丢弃过期帧。该策略的解释在于将时延控制转化为参数化约束，使得融合输出可被后续定位模块视为近似同步观测。实际部署时， ε_t 与队列上限应结合输出频率与网络抖动统计确定【待补充：队列长度与延迟上限配置】。

在时间语义确定后，坐标统一通过外参矩阵直接完成。对匹配到的帧对 (t, t') ，将 $\mathcal{P}_{l_2}(t')$ 通过 $\mathbf{T}_{l_2}^{l_1}$ 变换到 $\{l_1\}$ 并与 $\mathcal{P}_{l_1}(t)$ 合并，可写为

$$\mathcal{P}_f(t) = \mathcal{P}_{l_1}(t) \cup \mathbf{T}_{l_2}^{l_1} \mathcal{P}_{l_2}(t'). \quad (3.17)$$

该表达的证据基础是刚体运动假设：两雷达在同一刚性平台上，其相对位姿在短时内可视为常量。其解释在于将多传感器问题降维为“单参考系下的点集合合并”，并把不确定性集中在外参误差与时间对齐误差上，便于后续通过实验做误差归因。

融合后的轻量滤波用于控制噪声与点数规模，但其边界条件是不能破坏定位侧特征结构。体素下采样与离群点剔除在此处仅用于抑制重复观测与随机噪声，滤波强度应以“不改变主要结构边界”为准【待补充：体素分辨率与离群点阈值】。对于吊钩自身结构引起的固定反射，可通过空间掩膜进行剔除，其作用是减少后续动态分离中“自体伪动态”的误判。

此外，为降低吊钩端姿态倾斜对点云空间一致性的影响，融合节点引入基于 IMU 重力方向的自动调平（仅校正横滚/俯仰，保留航向）。设单位重力方向在当前时刻估计为 $\hat{\mathbf{g}}(t)$ ，目标是构造旋转 $\mathbf{R}_{\text{level}}(t)$ 使得 $\mathbf{R}_{\text{level}}(t)\hat{\mathbf{g}}(t) = [0 \ 0 \ 1]^T$ 。在工程实现中可由 IMU 姿态解算得到 roll/pitch，并

令

$$\mathbf{R}_{\text{level}}(t) = \mathbf{R}_y(-\text{pitch}(t)) \mathbf{R}_x(-\text{roll}(t)). \quad (3.18)$$

对点云做 $\mathbf{R}_{\text{level}}$ 旋转后发布，可使近场结构在短时滑窗累积中更稳定，从而提升后续静态分离与聚类的可重复性。

自动调平的论点在于减少由横滚/俯仰引入的系统性倾斜，使后续 ROI 定义与距离阈值在空间上具有更一致的物理含义。其证据来自吊钩端近场任务的特点：关注域通常以“水平距离 + 高度区间”定义，若点云坐标系随姿态摆动发生倾斜，则同一物体在不同帧中的 z 分布会产生非物理波动，进而影响裁剪与聚类。其解释是通过将重力方向对齐，消除了姿态变化在点云坐标中的主要低频分量。该步骤的小结是：融合发布阶段已同时控制了时间语义、空间语义与姿态语义，为下一节将融合点云作为单输入进行 LIO-SAM 定位提供了稳定接口。

3.2 基于融合点云的 SLAM 定位与静态地图构建（LIO-SAM）

3.2.1 算法选择理由与系统输入输出

定位模块在本系统中承担“为风险检测提供稳定空间参考”的角色，因此其首要目标是输出连续、可用于跨帧对齐的位姿序列 $\mathbf{T}_w^h(t)$ 。吊钩端平台同时存在短时高频摆动与长时低频漂移，单独依赖 LiDAR 里程计在快速摆动与点云畸变条件下容易出现配准残差波动，单独依赖 IMU 又会因积分累积产生不可忽略的漂移。基于上述事实，本章选择融合几何约束与惯性约束的 LIO-SAM 作为定位建图框架，其论点是让 LiDAR 提供结构一致性，让 IMU 提供高频运动先验与去畸变支持，从而在强动态平台上获得更稳定的位姿输出。

系统输入由两部分构成：一是融合点云 $\mathcal{P}_f(t)$ ，二是 IMU 的角速度与加速度测量及其时间戳。系统输出包括里程计/优化位姿 $\mathbf{T}_w^h(t)$ 、关键帧点云与静态地图。其证据基础是后续模块的接口需求：目标生成与风险判定需要把连续帧点云变换到统一参考系，并在局部时间窗内估计背景与动态目标，因此必须有可查询、可插值的位姿输出。其解释在于，位姿不仅是“定位结果”，也是后续数据标准化的约束条件；位姿稳定性直接决定静态分离阈值是否可迁移、聚类是否可重复。

需要说明的是，定位模块的输入接口通常假设单一 LiDAR 点云流与单一 IMU 流。为满足该结构约束，系统采用双雷达前置融合为单点云流的适配策略，即将两雷达点云在同一参考系下对齐并合并为 $\mathcal{P}_f(t)$ 后再输入定位模块。该设计不引入多雷达联合优化的额外状态量，其证据是工程部署的可维护性需求：保持定位模块接口不变可以降低系统集成风险并便于参数复用。其解释在于将多雷达带来的主要不确定性显式集中为外参误差、时间对齐误差与扫描结构参数适配等因素，从而便于后续通过实验进行误差归因与参数整定【待补充：定位输出稳定性评价指标与对比设置】。

除算法本体外，系统输入输出还隐含了对坐标系与时间基准的约束。其论点是：位姿输出只有在“帧定义一致、时间戳一致”的前提下才具有可复用性，否则相同的数值会在不同模块中对应不同的物理含义。其证据来自工程集成事实：融合点云通常在雷达参考系或吊钩端参考系下发布，而定位模块内部还会维护世界系（地图系）与里程计系等，若缺少明确的 TF 链路，则后续 ROI 裁剪与动态判别会出现阈值随帧漂移的问题。其解释在于，点云处理中的距离阈值、窗口长度与门控半径均依赖空间尺度，而空间尺度的稳定性依赖于坐标系定义的稳定性【待补充：系统采用的 world/odom/base 帧命名与 TF 发布关系】。

时间基准同样关键。其论点是：IMU 与融合点云的时间对齐质量会直接影响去畸变与前端初值，进而影响位姿连续性。其证据是去畸变模型需要用到 $\mathbf{T}(t \leftarrow \tau)$ ，该相对位姿来自 IMU 在扫描周期内的连续积分；若 IMU 时间戳相对点云存在偏移，则会把真实运动误当作噪声注入点云

几何残差。其解释是系统应在驱动与融合阶段就保证时间戳单调、跨传感器对齐误差可控，并在定位侧以消息队列与插值查询的方式获得与点云帧一致的 IMU 预积分区间【待补充：时间同步方式与允许偏移上限】。

本节小结是明确定位模块的输入输出、接口约束以及坐标系与时间基准要求，为下一节 LIO-SAM 的流程展开提供边界条件。

3.2.2 LIO-SAM 流程

LIO-SAM 将激光雷达与惯导的融合过程组织为“前端增量估计 + 后端平滑优化”的两级结构，其核心思想是把高频运动信息交由 IMU 约束提供，把低频几何一致性交由 LiDAR 约束校正，并在统一的因子图中实现一致性融合^[69]。其论点是：在扫描式 LiDAR 的帧内运动畸变与强动态平台摆动同时存在的条件下，仅依赖几何配准会导致残差具有结构性偏差，而仅依赖惯导积分又会出现不可控漂移；因此需要一个既能利用连续时间运动先验、又能利用环境结构约束的紧耦合框架。其证据是该框架将去畸变、初值提供、偏置估计与地图约束串联为闭环信息流，使误差不再单向累积；其解释在于每一步输出都为下一步提供可验证的约束，从而把“稳定位姿输出”落到可优化的残差形式。小结而言，本节关注的不是某个模块的实现细枝末节，而是该流程在系统集成中的输入输出边界与误差传递链路。

从观测层面看，LIO-SAM 先用 IMU 预积分在扫描周期内构造连续运动先验，并将其用于点云去畸变与雷达里程计初值^[69]。其论点是：点级时间戳差异是帧内畸变的根源，必须在进入几何配准之前显式处理。设一次扫描周期为 $[t, t + \Delta T]$ ，对任意点采样时刻 $\tau \in [t, t + \Delta T]$ ，由 IMU 约束获得传感器从 τ 到 t 的相对位姿变换 $\mathbf{T}(t \leftarrow \tau)$ ，则点的去畸变表达为

$$\tilde{\mathbf{p}}(t) = \mathbf{T}(t \leftarrow \tau) \mathbf{p}(\tau). \quad (3.19)$$

其证据在于该变换把“同一帧点云对应同一参考时刻位姿”的假设恢复为近似成立，使帧间几何残差更接近真实静态结构；其解释是 IMU 的连续时间信息被注入到离散点云观测中，从而把由平台自运动造成的系统误差分离出去。小结而言，去畸变既提升了前端配准的可观测性，也为后端优化提供更一致的观测基础，相关实现细节与参数适配将在下一小节进一步讨论【待补充：去畸变查询区间的实现方式】。

在几何约束构造方面，前端通常对去畸变后的点云进行几何特征抽取（如边缘/平面）或直接构造 scan-to-map 的匹配残差，从而形成 LiDAR 侧的增量约束^[69]。其论点是：相较于全点云匹配，稀疏且结构稳定的特征能够在有限算力下提供更可靠的几何约束。其证据是边缘与平面等局部结构在不同视角下具有更高的重复观测概率，并且其残差对小幅位姿扰动更敏感；其解释在于该敏感性提高了位姿估计的数值条件性，使得 IMU 初值能被几何约束快速拉回一致解。小结而言，前端输出的 LiDAR 里程计不仅用于连续位姿估计，也会反向参与惯导偏置的估计与校正，从而抑制由偏置漂移导致的去畸变误差反馈【待补充：前端残差的构造形式与收敛判据】。

后端部分以因子图形式联合优化位姿、速度与 IMU 偏置等状态量，并通过增量平滑方法在滑动窗口内维持实时性^[69]。其论点是：把约束表达为因子并显式指定噪声模型，能够将“融合权重”从经验调参提升为可解释的统计建模问题。其证据来自工程需求：前端增量估计在长时间运行中难以避免漂移，而后端优化能够在窗口内对状态序列施加一致性约束，并通过边缘化控制计算规模；其解释在于图结构的稀疏性允许在保留必要历史信息的同时丢弃冗余变量，避免计算量随时间线性增长。以关键帧时刻 k 的状态 \mathbf{x}_k 表示位姿、速度与偏置等量，则优化目标可抽象写

为

$$\min_{\{\mathbf{x}_k\}} \sum_k \|\mathbf{r}_{\text{imu}}(\mathbf{x}_{k-1}, \mathbf{x}_k)\|_{\mathbf{W}_{\text{imu}}}^2 + \sum_k \|\mathbf{r}_{\text{lidar}}(\mathbf{x}_k)\|_{\mathbf{W}_{\text{lidar}}}^2, \quad (3.20)$$

其中 \mathbf{r}_{imu} 由 IMU 预积分形成, $\mathbf{r}_{\text{lidar}}$ 由特征匹配或 scan-to-map 约束形成。其证据是“IMU 高频稳定、LiDAR 低频校正”的互补关系;其解释是权重矩阵对应噪声方差的倒数,决定了两类约束在优化中的相对影响,因此参数整定应围绕噪声一致性展开【待补充:IMU 噪声/偏置与匹配噪声模型的设置依据】。小结而言,后端的主要价值在于把短时的几何信息与惯性信息融合为一致的状态估计,从而提升位姿序列的可复用性。

为满足在线运行的实时性, LIO-SAM 在地图管理上采用关键帧选择与固定规模局部子地图(子关键帧集合)等策略,并将回环相关机制视为可选增强^[69]。其论点是:若把每一帧都与全局地图匹配,计算量与误匹配风险都会随运行时间累积,而局部匹配能在保持几何约束密度的同时显著降低计算开销。其证据是局部滑窗内的结构相对稳定且重叠度高,匹配更容易收敛;其解释在于该策略把优化约束限制在“可观测且可实时处理”的范围内,并通过关键帧稀疏化减少冗余信息。结合吊钩端近场短时运行场景,本文更强调里程计连续性与局部一致性,因此回环检测与闭环优化作为可选项:在具备重复轨迹与足够视野覆盖时可用于抑制累计漂移,但在强动态、遮挡频繁或任务时间较短时应谨慎开启以避免错误闭环带来的突变【待补充:是否启用回环的判据与对位姿连续性的影响评估】。

本节小结是: LIO-SAM 通过“IMU 约束用于高频运动建模与去畸变, LiDAR 约束用于几何校正,因子图用于一致性平滑”的流程输出连续位姿 $\mathbf{T}_w^h(t)$ 。该输出将作为后续点云标准化与目标生成的空间参考,并为下一小节围绕去畸变/配准与扫描结构参数适配的工程讨论奠定边界条件。

3.2.3 位姿输出与点云去畸变/配准

位姿输出与点云去畸变/配准的关系可以理解为“位姿是约束,去畸变是将约束落到点级别”。其论点是:在强动态平台上,如果不对帧内采样时间差做处理,则同一帧点云不满足刚体假设,配准残差会出现结构性偏差,最终表现为位姿抖动与地图模糊。其证据来自扫描式 LiDAR 的工作机制:扫描周期 ΔT 内平台持续运动,点云中的每个点对应不同的采样时刻 τ 。其解释在于,去畸变通过 IMU 约束恢复“同一帧对应同一参考时刻位姿”的近似,使得跨帧配准建立在更一致的几何观测上。

当 LIO-SAM 输出 $\mathbf{T}_w^h(t)$ 后,后续模块可以把 ROI 点云变换到统一参考系或局部参考系,从而实现滑窗累积、静动态分离与目标跟踪。其论点是:将时间序列观测变为同一坐标系下的空间比较问题;其证据是动态目标的定义本质来自跨帧位移,而位姿对齐提供了计算位移的共同参考。其解释是把“平台自运动”从观测中剥离,使剩余变化更接近外界目标运动。小结而言,定位输出不仅服务于建图,更为本章后半部分的目标生成提供了关键约束。

在双雷达前置融合场景下,还需要对点云的扫描结构参数进行一致性适配。定位前端的特征提取往往依赖垂直方向的分组(scan ring)组织邻域搜索与曲率统计,可用垂直通道数 n_{scan} 抽象表征。其论点是:融合点云的扫描分组规模必须与算法的 n_{scan} 假设一致,否则会导致 ring 分配错误或特征统计偏置,使残差不再反映真实几何误差。其证据体现在初始化阶段:当参数不匹配时,配准残差会呈现非随机偏置并易诱发短时漂移。其解释是错误的分组会把本应相邻的点视为远邻,破坏曲率估计与特征匹配的局部性。

基于上述原因,系统遵循“融合点云的 n_{scan} 等于参与融合的各雷达 n_{scan} 之和”的配置原则。在双雷达配置下,将 n_{scan} 从单雷达常用的 4 调整为 8,以匹配融合输入的等效扫描结构。该

装

订

线

设置的有效性应通过对 ring 分布与特征提取稳定性进行验证，例如统计每个 ring 的点数分布是否均衡，以及短时位姿输出的抖动是否收敛【待补充：验证指标与对比设置】。本节结论为下一节点云标准化与目标生成中的阈值设计提供了误差来源解释。

3.3 面向目标提取的点云标准化与静态目标生成

本节聚焦于“目标生成”：在位姿约束下对点云进行标准化与静态分离，输出静态障碍表征与动态簇候选及其状态量，不在本节进行风险分级与预警触发。

3.3.1 近场关注区域 ROI 裁剪与点云预处理

近场关注区域的 ROI 裁剪与点云预处理的目标是将“定位输出的空间参考”转化为“对风险任务有意义的点集”。其论点是：吊钩端风险检测关注的是有限空间内的可碰撞对象，若直接在全量融合点云上执行分离与聚类，会引入大量与任务无关的远场点并显著增加计算负担。其证据来自工程约束：预警刷新率需要稳定，算法必须在固定算力下在线运行，因此必须控制点数规模与噪声水平。其解释在于 ROI 将计算域限定在风险相关的几米尺度范围，使后续阈值（距离、持续帧数、聚类半径）具有更一致的物理含义。

ROI 的定义依赖于吊钩端参考系 $\{h\}$ ，其优势是与吊钩端运动共同移动，能够在平台摆动时保持关注域相对稳定。为使 ROI 具备可复现的数学描述，本文以 $\{h\}$ 下的点 $\mathbf{p}^h = [x \ y \ z]^T$ 定义近场关注域 Ω 。当选用圆柱形 ROI 时，可写为

$$\Omega = \left\{ \mathbf{p}^h \mid r_{\min} \leq \sqrt{x^2 + y^2} \leq r_{\max}, z_{\min} \leq z \leq z_{\max} \right\}. \quad (3.21)$$

其证据基础是作业安全尺度通常可由“水平距离 + 有效高度区间”描述；其解释是把风险相关空间用简洁约束表达，便于在程序中稳定复现。 r_{\max} 、 z_{\min} 与 z_{\max} 的最终取值需结合吊装工况与现场布置确定【待补充：ROI 范围的取值依据与最终数值】。

对融合点云 $\mathcal{P}_f(t)$ 的裁剪记为 $\mathcal{P}_\Omega(t) = \mathcal{F}_\Omega(\mathcal{P}_f(t))$ 。在 ROI 内的预处理遵循“轻量、在线、可解释”的原则，其论点是在不破坏几何边界的前提下降低随机噪声与点数规模。体素下采样与离群点抑制可抽象为算子 $\mathcal{F}_{\text{voxel}}(\cdot)$ 与 $\mathcal{F}_{\text{sor}}(\cdot)$ ，得到

$$\hat{\mathcal{P}}_\Omega(t) = \mathcal{F}_{\text{sor}}(\mathcal{F}_{\text{voxel}}(\mathcal{P}_\Omega(t))). \quad (3.22)$$

其证据是点云噪声与重复观测会导致聚类边界不稳定，进而影响跟踪与风险判定；其解释是通过控制点密度与异常点比例，使后续阈值更稳定。体素尺寸、离群点统计邻域大小与标准差阈值应结合点云密度与实时性约束整定【待补充：下采样与离群点参数】。

通过将该步骤放在位姿约束之后，可以在不破坏定位前端特征的前提下，显著降低后续静态分离与聚类的计算量并提高阈值判别的稳定性。

在工程实现中，ROI 裁剪还承担“把空间先验显式化”的作用。其论点是：在吊钩端场景中有一部分几何结构是长期随平台运动的自体结构（如吊具、钢丝绳局部），若不做剔除会在动态判别中形成稳定的伪前景。其证据是这类结构在点云中具有高反射与高密度特征，并且在平台摆动时相对传感器产生显著运动，但它们并非风险目标。其解释在于，结合安装几何可在 $\{h\}$ 下定义一个固定空间掩膜 Ω_{self} ，并在 ROI 内进一步执行 $\mathcal{P}_\Omega(t) \leftarrow \mathcal{P}_\Omega(t) \setminus \Omega_{\text{self}}$ ，从而把后续的动态分离重点聚焦在外界目标上【待补充：自体掩膜的几何范围与确定依据】。

同时，ROI 的垂直边界对姿态变化敏感，这也是上一章引入自动调平的原因之一。其论点是：若不校正横滚/俯仰，则同一物体在不同帧中的 z 分布会随姿态摆动产生系统性漂移，使得

z_{\min}, z_{\max} 的物理含义不再稳定。其证据来自近场判别阈值的经验性：垂直裁剪往往依据“吊钩附近有效高度”设定，而该高度应与重力方向一致；其解释是调平把重力方向对齐后，ROI 在 z 方向的约束更接近“真实高度区间”，从而减少裁剪抖动与边界截断导致的聚类不稳定【待补充：调平前后 ROI 点数波动对比】。

本节小结是：ROI 裁剪与轻量预处理将风险关注空间显式化并控制点数规模，同时通过自体掩膜与调平约束减少伪动态来源，为下一节的位姿对齐与滑窗背景建模提供更稳定的输入。

3.3.2 基于位姿的时空对齐与局部滑窗表示

时空对齐与局部滑窗表示的目标是构造一个对当前时刻有效、对噪声与稀疏更鲁棒的近场背景参考。其论点是：单帧 ROI 点云可能由于遮挡或采样稀疏导致结构不完整，而在位姿约束下做短时累积可以增强静态结构表达，从而为后续静动态分离提供更稳定的“背景”。其证据来自近场场景的观测特点：吊钩端附近存在反射不均与遮挡，静态结构在单帧中可能呈现断裂；其解释是滑窗累积将多个时间片的观测叠加到统一参考系，使静态结构以更高密度被观测到。

利用 LIO-SAM 输出的位姿 $\mathbf{T}_w^h(t)$ ，可将时刻 t 的 ROI 点云 $\mathcal{P}_\Omega(t)$ 变换到统一参考系实现跨帧对齐。为降低全局漂移对近场阈值的影响，本文采用局部参考系滑窗：以当前时刻 t 的位姿 $\mathbf{T}_w^h(t)$ 为基准定义局部系 $\{\ell(t)\}$ ，并将历史帧点云映射到该局部系。记 $\mathbf{T}_\ell^h(t) = \mathbf{I}$ ，则对任意历史时刻 $t - k$ 的点云有

$$\mathcal{P}_\Omega^\ell(t - k) = (\mathbf{T}_w^h(t))^{-1} \mathbf{T}_w^h(t - k) \mathcal{P}_\Omega(t - k). \quad (3.23)$$

据此定义滑窗背景表示为

$$\mathcal{B}(t) = \bigcup_{k=1}^N \mathcal{P}_\Omega^\ell(t - k). \quad (3.24)$$

该表示的证据在于局部系内短时间段有位姿误差更容易被认为是“近似刚体”，从而减少背景对齐误差；其解释是把长期漂移从背景建模中剥离，使背景更服务于短时风险判断。小结而言， $\mathcal{B}(t)$ 提供了一个短时有效的背景点集。

滑窗长度 N 的选择需要在稳定性与实时性之间权衡。其论点是： N 过小会导致背景不完整、对噪声敏感， N 过大则引入过时信息并增加计算开销与拖影风险。其证据是动态目标在滑窗中会留下轨迹残留，若窗口过长则可能被误认为静态结构；其解释是窗口长度等效于背景的时间跨度 $T_{\text{win}} \approx N/f$ （ f 为融合点云输出频率）。实际部署中 N 的取值建议通过对“静态背景完整度”“动态误检率”“处理耗时”进行综合对比确定【待补充： N 的最终取值与依据】。本节结论为下一节静动态点分离提供了背景定义与误差来源。

值得注意的是，时空对齐不仅依赖于位姿序列本身，还依赖于“位姿查询时刻”的一致性。其论点是：若点云帧时间戳与位姿更新时间戳不完全一致，则直接使用最近一帧位姿会引入额外对齐误差，并在背景中形成结构性重影。其证据来自系统的异步链路：点云融合、IMU 预积分与后端优化往往以不同频率更新；其解释是应基于点云时间戳 t 对位姿 $\mathbf{T}_w^h(t)$ 做插值或外推查询，并保证查询策略在全系统中一致，从而使 \mathcal{P}_Ω^ℓ 的构造具有可复现的时间语义【待补充：位姿插值方式与时间对齐误差统计】。

局部滑窗表示的误差来源可以归纳为定位噪声、去畸变残差与传感器测量噪声三部分。其论点是： $\mathcal{B}(t)$ 的“背景”并非绝对静止真值，而是这些误差叠加下的近似，因此后续阈值 δ_s 与持续性门限需要与误差上界相匹配。其证据是即便在静止场景下，点到背景的最近邻距离 $d(\mathbf{p}, \mathcal{B})$ 也不会收敛到 0，而会呈现与误差相关的分布；其解释是用统计量（分位数或均值方差）描述该分布，可为动态判别阈值提供更可解释的选取依据【待补充：静止背景下 $d(\mathbf{p}, \mathcal{B})$ 的统计结果】。本节小结是：通过局部系累积得到的 $\mathcal{B}(t)$ 提供了“短时稳定的背景近似”，并把误差以可量化形式

传递给下一节的静动态判别。

3.3.3 静动态点分离与动态点云簇提取

静动态点分离的任务是从 ROI 点云中提取“相对平台发生运动的点”，并将其组织为可跟踪的动态目标候选。其论点是：风险检测关注的不是全空间的点云变化，而是在位姿约束下仍然无法被静态背景解释的变化部分。其证据来自系统结构：前面已通过定位把平台自运动从观测中剥离，因此剩余变化更可能对应外界动态或噪声；其解释在于，静动态分离将“结构一致性”作为判别依据，使后续聚类与跟踪面对更干净的前景点集。

在得到背景表示 $\mathcal{B}(t)$ 后，可以通过点到背景的一致性度量进行判别。为使判别规则明确可复现，本文在局部参考系 $\{\ell(t)\}$ 下定义点到背景的最近邻距离

$$d(\mathbf{p}, \mathcal{B}(t)) = \min_{\mathbf{q} \in \mathcal{B}(t)} \|\mathbf{p} - \mathbf{q}\|_2. \quad (3.25)$$

对当前帧点 $\mathbf{p} \in \mathcal{P}_\Omega^\ell(t)$ ，若 $d(\mathbf{p}, \mathcal{B}(t)) \leq \delta_s$ ，则认为其与背景结构一致并判为静态；反之判为前景候选。其证据基础是静态结构在短时滑窗内应保持空间一致，而动态目标会产生明显位移；其解释在于 δ_s 同时吸收了传感器测量噪声、对齐误差与残余去畸变误差。实际部署中可将 δ_s 理解为误差上界的经验取值，并通过统计静止场景的 $d(\mathbf{p}, \mathcal{B})$ 分布来确定阈值【待补充： δ_s 的取值与统计方法】。

仅用单帧一致性容易受遮挡与噪声影响，因此还需引入时序一致性门限。其论点是动态目标应在时间上呈现持续存在或持续位移，而随机噪声往往是瞬时的。其证据在于离群点与遮挡边缘通常只在少数帧出现；其解释是通过要求前景候选在连续 m 帧内被重复观测，并且其局部位移幅值超过 δ_m ，从而把“瞬时异常”与“真实运动”区分开。参数 m 与 δ_m 应结合输出频率与目标速度范围确定【待补充： m, δ_m 取值依据】。

得到动态点集后，采用基于欧氏距离的聚类将其划分为若干簇 $\{\mathcal{C}_u\}$ 。对每个簇计算质心与包围盒等几何属性，例如质心

$$\mathbf{c}_u = \frac{1}{|\mathcal{C}_u|} \sum_{\mathbf{p} \in \mathcal{C}_u} \mathbf{p}, \quad (3.26)$$

并由各坐标轴方向的最小/最大值构造轴对齐包围盒，为后续短时跟踪与风险评估提供统一的目标观测量。

为降低误检，动态簇还需满足最小点数与最小持续帧数等一致性约束，其论点是避免将稀疏噪声簇误认为真实目标。其证据是点云在强反射与遮挡条件下会产生局部稀疏点团；其解释是通过“点数阈值+持续性阈值”对候选做二次筛选。对吊钩自身结构引起的固定反射区域，可通过空间屏蔽减少自体误检【待补充：屏蔽区域定义与依据】。本节小结是输出一组具有几何属性的动态簇候选，为下一节短时跟踪与状态估计提供观测输入。

3.3.4 目标短时跟踪与状态量估计

目标短时跟踪的目的不是构建长期语义轨迹，而是为风险判定提供短时可用的速度与运动趋势估计。其论点是：仅依靠单帧动态簇无法稳定计算相对速度与 TTC，而跨帧关联能够把离散观测转换为连续状态量。其证据来自点云聚类的不稳定性：同一目标在相邻帧中的点数与边界会波动，若不引入状态模型就会造成速度估计剧烈抖动；其解释是通过滤波与门控将观测噪声吸收到协方差中，从而输出更平滑、更可用的速度。

具体地，设目标状态向量为

$$\mathbf{x}(t) = [x \ y \ z \ \dot{x} \ \dot{y} \ \dot{z}]^T, \quad (3.27)$$

在采样周期 Δt 下采用匀速（CV）离散模型

$$\mathbf{x}(t + \Delta t) = \mathbf{F} \mathbf{x}(t) + \mathbf{w}(t), \quad \mathbf{F} = \begin{bmatrix} \mathbf{I}_3 & \Delta t \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}, \quad (3.28)$$

观测取聚类质心 $\mathbf{z}(t) = \mathbf{H} \mathbf{x}(t) + \mathbf{v}(t)$ ，其中 $\mathbf{H} = [\mathbf{I}_3 \ \mathbf{0}_3]$ 。卡尔曼滤波通过“预测—更新”迭代得到对 $\mathbf{p}(t) = [x \ y \ z]^T$ 与 $\mathbf{v}(t) = [\dot{x} \ \dot{y} \ \dot{z}]^T$ 的最小均方差估计。

数据关联采用门控策略以抑制误匹配。其论点是：预测状态给出了下一时刻的合理位置范围，观测若偏离过大则更可能来自不同目标或噪声簇。其证据是近场场景中可能存在多目标与遮挡，聚类结果会出现分裂/合并；其解释是门控将“可匹配区域”显式参数化，从而减少跨目标跳变。对预测位置 $\hat{\mathbf{p}}(t)$ 与当前观测 $\mathbf{c}_u(t)$ ，若其欧氏距离 $\|\mathbf{c}_u(t) - \hat{\mathbf{p}}(t)\|_2 \leq \delta_a$ （或马氏距离小于阈值），则允许匹配并更新对应轨迹；否则创建新轨迹或将其暂存为待确认目标。门控阈值 δ_a 的取值应结合采样周期、目标最大速度与定位误差确定【待补充： δ_a 的取值依据】。

为了使跟踪输出能够稳定服务于第 4 章风险指标计算，轨迹管理还应引入“确认/删除”的持续性规则。其论点是避免单帧噪声簇频繁生成轨迹并污染风险评估；其证据是点云动态分离在阈值附近会出现抖动；其解释是对轨迹设置最小连续命中帧数与最大允许丢失帧数，从而在遮挡条件下保持短时连续性，同时在噪声条件下快速清除伪目标【待补充：轨迹确认与删除规则参数】。

从风险计算角度看，跟踪输出的关键在于“速度是否可用”。其论点是：近场预警更关心目标沿危险方向的接近趋势，而不是全局三维速度的每个分量都精确。其证据在于 TTC 等指标通常基于距离变化率构造；其解释是可以利用滤波后的状态估计得到接近速度，例如对吊钩端参考系下的相对位置 $\mathbf{p}(t)$ 与速度 $\mathbf{v}(t)$ ，其沿连线方向的接近速度可写为 $v_{\text{rel}}(t) = -\mathbf{p}(t)^T \mathbf{v}(t) / \|\mathbf{p}(t)\|$ ，该量对横向噪声更不敏感，因此更适合作为工程上稳定的风险输入【待补充：速度抖动统计与平滑效果对比】。

另一方面，聚类观测的不稳定会引发轨迹的分裂与合并，这是近场点云跟踪的主要挑战之一。其论点是：若只依赖最近邻匹配，轨迹在遮挡边缘或点数骤变时容易发生 ID 跳变，从而使速度估计出现尖峰。其证据是动态簇的质心会随可见部分变化而偏移；其解释是在门控的基础上应结合包围盒尺寸、点数规模与短时运动连续性共同约束匹配质量，并在观测异常时选择“降低更新权重或暂缓更新”，以把观测波动吸收到协方差增长中而不是直接映射为速度突变【待补充：观测异常判别规则与滤波参数设置】。通过上述处理，短时跟踪能够在有限计算量下输出更稳定的状态量，为后续分级预警状态机提供可靠输入。

本节小结是：通过“聚类观测—门控关联—滤波估计—轨迹管理”的链路，可输出稳定的目标位置与速度等状态量，并以可参数化的方式把观测不确定性传递给后续相对速度与 TTC 的风险计算。

3.4 本章小结

本章围绕“为吊钩端风险检测提供稳定、可复用的三维空间参考”这一目标，建立了从融合点云输入到动态目标状态量输出的完整点云侧处理链路。其论点是：风险预警并不直接依赖某一帧点云的瞬时形态，而依赖在统一时空语义下可重复的空间度量（距离、速度、时间裕度），因此必须首先解决多传感器观测的一致性与平台自运动带来的非刚体效应。其证据来自系统集成事实：双雷达异步、外参误差、姿态摆动与扫描畸变会以不同方式注入后续阈值判别；其解释在于

本章通过在输入端显式约束时间与坐标语义，把不确定性集中并可追踪地传递到后续模块。小结而言，本章给出了后续预警策略可依赖的“统一参考系下的几何与运动输入”。

在融合点云构建方面，本章提出双雷达点云的时间对齐、坐标统一与轻量滤波发布流程，并引入基于 IMU 重力方向的自动调平以提升近场结构的短时一致性。其论点是：只有在同一参考系与可控延迟条件下形成的融合点云，才能作为定位与目标生成的稳定接口。其证据是外参误差与时间偏移会直接表现为点云边界错位与结构性重影；其解释在于通过对齐与调平把系统性误差抑制在可接受范围，为后续特征提取、配准与阈值设计提供更稳定的几何观测。小结而言，该部分完成了多雷达输入到单点云流的工程化适配。

在定位与静态地图构建方面，本章基于 LIO-SAM 形成“IMU 约束去畸变与初值 + LiDAR 几何约束校正 + 因子图一致性平滑”的位姿输出机制，从而为跨帧点云对齐提供连续位姿序列。其论点是：位姿不仅是定位结果，更是后续点云标准化与动态判别的约束条件；其证据是时空对齐、滑窗背景与动态分离均以位姿查询为前提；其解释在于通过将平台自运动从观测中剥离，后续模块才可以把“变化”主要归因于外界目标运动。小结而言，该部分确立了本章后半段目标生成的时空基准。

在目标生成与跟踪方面，本章在位姿约束下定义 ROI 裁剪与预处理，构建局部参考系滑窗背景表示，并基于点到背景的一致性度量实现静态动态分离；进一步通过聚类与短时跟踪输出目标的几何属性与速度等状态量。其论点是：通过“背景近似 + 时序一致性”的组合，可以在不引入复杂语义建模的前提下得到可用于风险计算的动态目标候选。其证据是单帧聚类对稀疏与遮挡敏感，而滑窗背景与门控滤波可把观测不确定性吸收到统计量与协方差中；其解释在于最终输出的状态量具备可解释的误差来源，便于在第 4 章将其转化为距离、相对速度与 TTC 等风险指标。小结而言，本章完成了从融合点云到“静态障碍表征 + 动态目标状态量”的关键中间层。

综上，本章为下一章的雷视融合与分级预警提供了两类基础输入：一是位姿与静态地图，用于定义稳定的空间尺度与安全边界；二是动态目标候选及其状态量，用于构造运动趋势与时间裕度相关的风险判定指标。基于这些输入，第 4 章将进一步引入视觉侧高风险目标语义，并给出面向在线部署的融合关联与预警策略。

4 基于雷视融合的避障预警策略

4.1 视觉侧高风险目标识别

4.1.1 危险目标（车辆）检测模型与部署

吊钩端近场的预警目标并不局限于作业人员。对于实验室缩尺平台而言，能够可控复现且具有代表性的高风险动态目标主要是**移动车辆**（例如遥控小车）。本节将视觉侧任务统一表述为**危险目标检测**：在图像序列中识别危险目标类别并给出二维位置，用于为后续雷视关联提供语义先验。其核心论点是：在近场避障中，LiDAR 更适合承担几何测距与距离判定，而视觉侧更适合提供“目标是否属于预警关注对象”的语义信息，从而在不引入复杂三维语义重建的前提下显著降低误报来源。其依据在于：仅依赖点云动态簇难以区分“真实危险目标”与“非危险动态物体/噪声”（如局部反射、摆动绳索、偶发遮挡边缘），而车辆具有更稳定的外观与尺度特征，适合通过学习型检测模型以较低成本获得语义判别。小结而言，本节将视觉输出限定为“可解释的 2D 候选集合”，以服务于第 4.2 节的可落地关联策略。

在模型选择上，本文采用 YOLOv8 系列检测器作为危险目标（车辆）检测模型，并将其部署在边缘计算单元，以满足在线预警的实时性约束。需要强调的是，选择 YOLOv8 的出发点并非追求极限精度，而是追求**可部署、可复现、可在遮挡与光照变化下保持稳定输出**的工程综合性能：其单阶段结构便于在较高帧率下运行，且可通过量化/加速推理框架获得较稳定的端到端时延。对于本研究的实验平台，摄像机采样频率约为 20 FPS；推理时延的均值与抖动将直接影响后续时间窗门控的设置，因此需要在系统评估阶段对“采集—推理—发布”的链路时延进行统计【待补充：YOLO 推理时延均值/方差或 95% 分位】。

为便于与点云模块融合，视觉检测模块对每帧图像输出时间戳、检测框与检测置信度，并在时序稳定化后给出轨迹标识 $track_id$ 。其中检测框定义为

$$bbox = (u_{\min}, v_{\min}, u_{\max}, v_{\max}). \quad (4.1)$$

此外，为适配“危险目标”这一更一般表述，视觉侧同时输出类别标识 $c \in \mathcal{C}$ 与类别置信度 s_{det} 。在当前缩尺实验中，类别集合可简化为 $\mathcal{C} = \{\text{vehicle}\}$ ，即只对车辆进行检测；在真实工况扩展到人员/机械时，只需在保持接口不变的前提下扩充 \mathcal{C} 即可【待补充：扩展类别集合的定义与标注口径】。上述输出通过 ROS 话题发布，作为雷视融合模块的语义输入。

4.1.2 检测结果时序稳定化处理

单帧目标检测在近场遮挡、运动模糊与光照变化条件下容易产生抖动（框位置漂移、间歇漏检、置信度尖峰），而雷视关联与风险状态机对输入的连续性更为敏感。为此，本文对危险目标检测输出进行时序稳定化，使其从“逐帧检测”转化为“具备短时记忆的 2D 轨迹候选集合”。其论点是：关联与预警的稳定性首先取决于目标身份（track）在短时尺度上的一致性；其依据是多目标情况下任意一次 ID 切换都可能导致 3D 绑定漂移，进而把速度估计的尖峰误传递到 TTC；其解释是通过帧间匹配、丢检容错与置信度滞回，可把观测噪声吸收到轨迹级状态而不是直接触发预警跳变。小结而言，本节输出的并非“更准的一帧”，而是“更稳的一段时间”。

首先进行帧间匹配与轨迹分配。对时刻 t 的检测集合 $\mathcal{D}(t)$ 与上一帧集合 $\mathcal{D}(t - \Delta t)$ ，本文以检测框重叠度（IoU）与中心距离共同构造匹配代价，并在门控约束下完成一对一分配。以第 i 个

上一帧检测框与第 j 个当前帧检测框为例，可定义

$$\text{IoU}_{ij} = \frac{|bbox_i \cap bbox_j|}{|bbox_i \cup bbox_j|}, \quad d_{ij} = \|\mathbf{c}(bbox_i) - \mathbf{c}(bbox_j)\|_2, \quad (4.2)$$

其中 $\mathbf{c}(\cdot)$ 表示检测框中心。综合代价可写为

$$C_{ij} = \lambda(1 - \text{IoU}_{ij}) + (1 - \lambda) \frac{d_{ij}}{d_0}, \quad \lambda \in [0, 1], \quad (4.3)$$

并设置门控条件 $\text{IoU}_{ij} \geq T_{\text{iou}}$ 且 $d_{ij} \leq T_{\text{dist}}$ ，以剔除显然不可能的匹配对。随后可采用匈牙利算法进行全局最优分配，亦可在目标数较小的情况下采用贪心近似以降低实现复杂度【待补充：实际采用的分配方式与阈值取值】。完成分配后，对匹配成功的检测继承上一帧的 $track_id$ ；对未匹配的检测创建新轨迹；对未被匹配的旧轨迹进入“暂存”状态等待短时恢复。

其次引入短时丢检容错机制。考虑到遮挡或检测阈值附近的抖动，本文允许轨迹在连续不超过 M 帧未匹配时仍保持存在，并用上一时刻位置或简单的恒速预测补偿其短时缺失。若采用二维运动模型，可对框中心引入 2D 卡尔曼滤波进行预测；若不引入显式滤波，也可仅用最后一次观测位置维持轨迹，但需要配合更严格的解除条件以避免轨迹“僵尸化”【待补充：是否采用 2D 滤波与预测模型】。当丢检超过 M 帧仍未恢复匹配，则判定轨迹消失并释放其 $track_id$ 。

最后对检测置信度进行平滑，并用双阈值滞回实现“确认—消失”的稳定切换。为获得更鲁棒的置信度输出，本文对检测置信度进行滑动平均：

$$\bar{s}_{\text{det}}(t) = \alpha \bar{s}_{\text{det}}(t-1) + (1 - \alpha) s_{\text{det}}(t), \quad \alpha \in [0, 1] \quad (4.4)$$

并采用双阈值滞回策略实现稳定的“确认—消失”切换：

$$\bar{s}_{\text{det}}(t) > T_{\text{on}} \Rightarrow \text{确认目标为可信危险目标（车辆）}, \quad \bar{s}_{\text{det}}(t) < T_{\text{off}} \Rightarrow \text{降级/移除目标}, \quad T_{\text{on}} > T_{\text{off}}. \quad (4.5)$$

在“安全优先”的工程偏好下， T_{on} 可适当取低以降低漏检风险，同时通过更强的滞回间隔 $T_{\text{on}} - T_{\text{off}}$ 与最小持续时间约束抑制误报抖动【待补充：阈值与最小持续时间参数】。通过上述处理，视觉侧最终输出连续、鲁棒的危险目标候选集合，为后续雷视目标关联提供可靠输入。

4.1.3 点云动态目标融合接口与数据结构定义

为使视觉信息能够以统一、可复现的形式参与雷视融合，有必要对视觉侧输出进行接口化建模，并明确其时间语义与坐标语义。对每一帧图像，危险目标检测模块输出候选集合 $\mathcal{D}(t)$ ，其中每个候选至少包含检测框坐标 $bbox = (u_{\min}, v_{\min}, u_{\max}, v_{\max})$ 、检测置信度 s_{det} 、类别标识 $c \in \mathcal{C}$ 以及由时序稳定化模块分配的 $track_id$ 。从工程实现角度，建议该消息同时携带图像时间戳 t 、消息发布的 $frame_id$ （相机坐标系或其光学坐标系标识）与相机内参版本号，避免在多次标定或多相机拓展时出现“数据可用但语义不一致”的隐性错误【待补充：实际消息字段定义】。

融合模块以时间戳为索引，将图像检测集合与点云动态目标集合进行匹配。需要注意的是，“同一时刻”在异步传感器与学习推理链路下通常不能严格成立：相机、LiDAR、推理与点云处理会引入不同的延迟与抖动。因此本文采用缓冲队列 + 延迟补偿 + 时间窗门控的方式实现可落地同步：对每路消息维护有限长度的时间序列缓存，使用估计的端到端延迟 $\hat{\Delta}$ 将检测时间戳修正为 $t' = t - \hat{\Delta}$ ，并在另一模态缓存中检索满足 $|t' - t_{\text{lidar}}| \leq \tau$ 的最近邻数据作为关联候选【待补充： $\hat{\Delta}$ 的估计方法与 τ 的取值依据】。该策略能够在 20 FPS/20 Hz 的系统频率下保证在线关联的可重

复性，并显式暴露“时间不一致”这一误差源。

在空间语义上，外参矩阵与投影模型直接采用第 2 章标定结果。融合时，点云侧 3D 动态目标可表示为三维点集或 3D 包围盒，并通过外参变换与相机内参投影映射到图像平面，从而使视觉检测框自然地成为“空间门控区域”。因此，该接口设计保证了视觉信息以“门控区域 + 语义置信度 + 轨迹 ID”的形式参与融合：视觉负责提出“哪些目标值得关注”，点云负责提供“目标在哪里、离危险源多远、运动趋势如何”。小结而言，本节把雷视融合从概念层面的“可投影”落实为工程层面的“可同步、可关联”。

4.2 雷视融合策略

本节面向工程落地，强调从“点云可投影到图像”进一步走向“点云与图像目标可稳定关联”。为降低遮挡、误检与漏检对预警的影响，本文采用双向一致性关联与轨迹级绑定：以点云动态簇为主候选，并由视觉检测提供语义确认；同时通过 2D→3D 门控与 3D→2D 反向验证抑制错配，并在短时遮挡条件下维持轨迹连续性。

4.2.1 融合总体流程与设计原则

本研究的雷视融合以工程可落地为导向，遵循“LiDAR 主几何、视觉供语义与验证”的设计原则：以点云动态簇为主要候选目标，利用其三维位置、几何范围与短时速度完成距离计算与风险触发；视觉检测作为语义先验与一致性验证，用于确认动态目标是否属于预警关注的危险目标（当前为车辆），从而降低误报。其核心论点是：对于吊钩端近场避障，风险判定本质上依赖几何距离与运动趋势（相对速度/TTC），因此三维几何测距应由 LiDAR 主导；而语义的主要作用是抑制“几何上是动态、但并非危险目标”的干扰源。其证据来自系统约束：在 20 Hz 级别的在线处理链路中，引入端到端 3D 语义重建或特征级深度融合将显著提高训练数据需求与算力开销，并带来更复杂的失效边界；其解释在于目标级融合以外参投影为几何纽带，能够用较小的实现代价获得足够的语义增益。小结而言，本节的融合策略选择了“可解释、可复现、可部署”的路径。

从流程角度看，融合模块的输入包括：点云侧 3D 动态目标轨迹（位置 $\mathbf{p}(t)$ 、速度 $\mathbf{v}(t)$ 、几何范围等）与视觉侧 2D 危险目标轨迹（ $bbox, \bar{s}_{det}, track_id$ 等），以及由第 2 章给出的标定参数与 TF 变换链。输出包括：每条 3D 轨迹的语义标签与融合置信度，以及用于第 4.3 节风险计算的“已确认危险目标集合”。为保障链路可实现，本文在设计上强调三点边界条件：其一，时间同步不追求严格同时，而是通过 $\hat{\Delta}$ 与时间窗 τ 实现可控的不一致；其二，空间关联不追求单帧完美，而是通过跨帧一致性与轨迹级绑定提升鲁棒性；其三，任何一侧感知退化时应具备降级路径，例如视觉暂时失效时仍可输出“未知动态目标”的几何风险提示，点云暂时失效时则避免输出不可信的距离/TTC 并仅给出视觉提示【待补充：降级策略与输出规则】。

4.2.2 双相一致性关联策略

双向一致性关联包含两个互补步骤：

首先进行 2D→3D 门控关联。对每个视觉检测框 $bbox$ ，利用雷达—相机外参 \mathbf{T}_l^c 与投影模型，将同一时刻点云动态簇的三维点/包围盒投影到图像平面；若投影点落入 $bbox$ 或与 $bbox$ 存在足够重叠（例如投影点集与检测框的覆盖率超过阈值），则认为该动态簇满足“视觉门控”，形成候选关联对。实际实现中， \mathbf{T}_l^c 可由 TF 在运行时查询得到，或用等价链路 $\mathbf{T}_h^c = \mathbf{T}_l^c \mathbf{T}_h^l$ 先将目标状态从吊钩端参考系变换到相机坐标系后再投影，从而保证投影门控与风险计算在坐标语义上自治【待补充：投影时采用的坐标系链与 TF 名称】。该步骤以视觉框作为空间先验，能够有效剔除大量与危险目标无关的动态簇（如摆动吊索、局部反射噪声）。

随后进行 **3D→2D 反向一致性验证**。对每个点云动态簇，除投影落入框的几何条件外，还要求满足时间窗一致性与几何合理性：例如动态簇在相邻帧的投影位置变化与 2D track 的运动趋势相符，且动态簇的尺度/高度与车辆的先验范围不矛盾。若仅满足单帧投影重叠但跨帧不一致，则降级为“unknown”或拒绝关联。该反向验证可以抑制由偶然重叠、遮挡边缘与误检导致的错误绑定。

为使上述策略在多目标场景中可复现，需要进一步明确门控指标、时间一致性与歧义分配规则。首先，在几何门控上，若直接使用“投影点是否落入框内”在点云稀疏或遮挡条件下容易产生漏关联，因此本文建议同时引入**最小投影点数或投影面积约束**：例如要求落入框内的投影点数 $n_{in} \geq n_{min}$ ，或将 3D 包围盒八顶点投影形成多边形并与 $bbox$ 计算 IoU，从而降低“少量偶然点”造成的误绑定【待补充：覆盖率/IoU 指标定义与阈值】。其次，在时间一致性上，候选关联必须满足 $|t'_{cam} - t_{lidar}| \leq \tau$ ，其中 t'_{cam} 为延迟补偿后的视觉时间戳，这使得时间同步误差以参数形式进入关联门控，而非隐式地污染几何一致性。

最后，在多目标歧义分配上，一帧中可能出现“一框对应多个动态簇”或“一簇落入多个框”的情况。为避免绑定抖动，本文将关联视为二分图匹配问题：对每个候选对构造关联得分（由几何重叠、时间一致性与轨迹一致性共同决定），再进行一对一分配。该分配可采用匈牙利算法得到全局最优，也可在目标数较小时用贪心近似得到近似最优解【待补充：关联得分函数与实际分配方法】。通过“候选生成—反向验证—全局分配”的组合，双相一致性关联从直觉描述提升为可实现的工程流程。

4.2.3 轨迹级绑定与遮挡鲁棒融合

在线运行时，视觉侧输出的 2D 目标轨迹（2D track）与点云侧输出的 3D 动态簇轨迹（3D track）均具有短时丢检的可能。为降低遮挡带来的频繁切换，本文采用**轨迹级绑定策略**：一旦某个 2D track 与某个 3D track 通过双向一致性验证形成稳定关联，则在后续若干帧内维持绑定关系，并以门控半径/投影区域进行快速匹配更新。

当视觉侧发生短时丢检但点云侧仍保持连续跟踪时，系统保留该目标的**车辆语义状态**，并以 3D 轨迹预测更新其位置与速度；当点云侧因稀疏或遮挡出现短时失配但视觉仍连续时，可暂时以视觉轨迹维持目标存在并等待点云重新捕获。若连续超过预设帧数仍无法恢复一致性，则解除绑定并将目标降级。该策略在保证误报可控的前提下，提高了**危险目标语义**在遮挡与复杂背景条件下的连续性。

为增强可复现性，轨迹级绑定可形式化为一个轻量状态机。设某条 3D 轨迹 \mathcal{T}^{3D} 与某条 2D 轨迹 \mathcal{T}^{2D} 的候选绑定关系为 B ，则 B 的状态可包括：未绑定（Unbound）、候选绑定（Tentative）、稳定绑定（Bound）与丢失保持（Hold）。进入 Tentative 的条件为连续 K 帧通过双向一致性验证；进入 Bound 的条件为在 Tentative 阶段累计通过帧数达到阈值且关联得分稳定；进入 Hold 的条件为在 Bound 状态下出现短时丢检但仍满足时间窗与空间门控；退出并解除绑定的条件为连续超过 L 帧无法满足一致性，或关联得分持续低于阈值。上述 K, L 的设置体现“安全优先”的策略： K 不能过小以避免偶然重叠触发错误绑定，而 L 也不能过大以避免错误绑定长期污染风险计算【待补充： K, L 取值与依据】。

需要强调的是，绑定的目标是保持“语义连续性”，而不是在任何情况下都强行维持同一绑定对。因此，当检测到明显的 ID 切换风险（例如多目标交汇导致关联得分突变）时，系统应选择暂时降级为 unknown 并重新进入 Tentative，而非把不确定性直接传递给 TTC 与预警状态机。小结而言，轨迹级绑定通过显式的状态约束把遮挡与漏检的影响限制在短时范围内，为后续置信度融合与风险评估提供更稳定的输入。

4.2.4 置信度融合与语义标签输出

融合模块输出目标语义标签及其置信度分数。考虑到本文缩尺实验中危险目标类别集合为 $\mathcal{C} = \{\text{vehicle}\}$ ，因此语义标签可简化为

$$y \in \{\text{vehicle}, \text{unknown}\}, \quad (4.6)$$

其中 **unknown** 表示“动态目标存在但未能可靠确认其为车辆”。该设计体现了工程上的保守原则：在无法可靠获得语义时，系统仍可在几何层面计算距离风险，但不应以不可靠语义触发更强的危险等级。

置信度分数由多源信息共同决定。视觉侧 \bar{s}_{det} 提供语义置信度，点云侧可提供动态性与稳定性指标（如跨帧一致性、点数规模、速度估计稳定性等）作为几何可信度 s_{lidar} 。为避免“权重拍脑袋”导致的不可复现，本文给出一种简单且可落地的线性融合形式：

$$s_{\text{fuse}}(t) = w_v \bar{s}_{\text{det}}(t) + w_l s_{\text{lidar}}(t), \quad w_v + w_l = 1, w_v, w_l \in [0, 1]. \quad (4.7)$$

其中 s_{lidar} 可由若干归一化指标组合得到，例如点数稳定性 s_n 、轨迹连续命中率 s_{trk} 与速度平滑性 s_v 等【待补充： s_{lidar} 的具体构造与归一化方式】。当 s_{fuse} 超过阈值 S_{on} 时输出为 **vehicle**；当低于 S_{off} 时输出为 **unknown**，并引入滞回以抑制阈值附近抖动（ $S_{\text{on}} > S_{\text{off}}$ ）。当出现遮挡或短时丢检时，轨迹级绑定可在有限帧数内维持目标状态并降低频繁切换。小结而言，本节输出把“语义是否可信”显式量化并可参数化，从而为风险判定提供可控输入。

4.3 风险判定与分级预警策略

风险判定采用**距离阈值 + 运动趋势（相对速度/TTC）**的组合策略：对静态障碍强调可控的最小安全距离与危险区域侵入判定；对动态目标强调提前量，通过相对接近速度与 **TTC** 刻画短时未来风险，并结合滞回与最小持续时间输出稳定的分级预警。

4.3.1 静态障碍：基于距离/侵入的阈值预警

静态障碍的风险判定目标是给出一个可解释、可标定、对噪声鲁棒的“近场侵入”触发机制：只要吊钩端附近出现固定构件、脚手架边缘或堆料区边界等静态结构，其与吊钩端的几何裕度一旦低于安全阈值就应触发提醒。与动态目标不同，静态障碍不强调远期预测，而强调对“是否进入危险区域”的稳定判断，并把误报抖动控制在可接受范围内。相关研究普遍采用“最小距离/碰撞几何 + 阈值触发”的思路来量化吊装场景中的碰撞风险与安全距离设置^[27, 50]。

(1) 指标定义与侵入判据。以吊钩端参考点为原点建立吊钩端坐标系 $\{h\}$ ，将融合后的近场点云（经坐标统一与时序对齐）裁剪得到 ROI 点集 $\mathcal{P}(t) = \{\mathbf{x}_i(t) \in \mathbb{R}^3\}$ 。定义 ROI 内到原点的最近点距离为

$$d_{\text{raw}}(t) = \min_{\mathbf{x} \in \mathcal{P}(t)} \|\mathbf{x}\|. \quad (4.8)$$

为把吊钩端本体尺寸与安装偏差纳入判定，将吊钩端危险源近似为等效球（或圆柱）包络，设其等效半径为 r_h ，则**最小安全距离**可定义为

$$d_{\text{min}}(t) = d_{\text{raw}}(t) - r_h. \quad (4.9)$$

当 $d_{\text{min}}(t)$ 较小时，表示静态结构已进入“吊钩端安全区”附近。等价地，也可将危险区域定义为

球形区域 $\mathcal{Z}(D) = \{\mathbf{x} \mid \|\mathbf{x}\| < r_h + D\}$ ，则“侵入”可用指示量表示为 $I(t; D) = \mathbf{1}\{d_{\min}(t) < D\}$ 。该写法把“侵入”与“距离阈值”统一为同一几何条件，便于后续状态机复现。

(2) 从静态地图与点云到 $d_{\min}(t)$ 的计算流程。本系统在前述定位建图基础上获得吊钩端位姿并构建静态地图，从而将点云统一到一致参考系并抑制摆动导致的空间漂移；同时，在线融合点云用于反映临时出现但运动缓慢的近场静态结构。实践中，静态障碍距离计算通常包含：**坐标变换与运动补偿**（把点云转到 $\{h\}$ 并与时间戳对齐）、**近场 ROI 裁剪**（仅保留吊钩端一定半径/扇区内点）、**离群点抑制**（去除单帧孤立点或远离主结构的噪点）以及**最近距离求解**（输出 $d_{\text{raw}}(t)$ ）。上述流程与基于环境点云进行吊装路径规划/防碰撞监测的思路一致：用点云表达几何约束，再将风险量化为可计算的距离裕度^[26, 65]。

需要注意的是，直接取“绝对最小值”会对少量噪点高度敏感。为提高鲁棒性，工程上可将 $d_{\text{raw}}(t)$ 替换为**分位数最小距离**（例如取距离序列的 q 分位），或要求“最小距离由不少于 n_{\min} 个点共同支持”，从而把偶发离群点造成的误触发概率显式压低【待补充：采用分位数/点数支持的具体设置】。

(3) 阈值设定：吊钩端尺寸+安全裕度的可标定方法。触发阈值采用进入/退出滞回对 $(D_{\text{on}}, D_{\text{off}})$ （满足 $D_{\text{on}} < D_{\text{off}}$ ）。其中 D_{on} 表示允许的最小安全裕度下限， D_{off} 表示解除预警所需的恢复裕度。为避免“拍脑袋”，本文建议将阈值分解为“结构尺寸上界+误差裕度”的组合： r_h 由吊钩端几何包络、传感器安装偏差与姿态估计误差给出保守上界； D_{on} 与 D_{off} 则由作业摆动引起的点云抖动幅度、点云稀疏带来的距离估计偏差以及对误报容忍度共同决定，并通过静态场景实验日志进行迭代标定，使误报率与响应延迟满足工程需求【待补充： $r_h, D_{\text{on}}, D_{\text{off}}$ 的取值、标定步骤与最终数值】。

(4) 稳定性机制：滞回+最小持续时间。为抑制阈值附近的频繁翻转，预警触发采用滞回：当 $d_{\min}(t) < D_{\text{on}}$ 时进入预警，当 $d_{\min}(t) > D_{\text{off}}$ 时退出预警。进一步地，为与第 4.3.3 节中的输出稳定化策略一致，静态障碍同样引入最小持续时间 Δt_{\min} ：只有当触发条件连续满足超过 Δt_{\min} 才更新状态，从而把短时离群点或瞬时遮挡造成的尖峰吸收在状态机内部。该组合机制使得预警输出既保守（安全优先），又能避免高频误报影响作业可用性。

(5) 与实验验证的闭环。基于上述定义，静态障碍预警的可复现评价可直接落到四类量化指标：距离估计的抖动幅度、滞回进入/退出的稳定性、误报率以及触发响应延迟；相应的测试工况与记录方式在第 5.4.1 节中给出。通过“指标定义—阈值标定—实验验证”的闭环，静态障碍预警从经验规则提升为可参数化、可复现实验支撑的工程策略。

4.3.2 动态目标：基于相对速度/轨迹预测的碰撞风险

对于动态危险目标（本文为车辆），仅以当前距离阈值触发容易出现“预警不提前”或“阈值附近抖动”的问题。其原因在于：当目标以一定速度接近危险源时，距离阈值只能在接近到很近时才触发，提前量不足；当目标横向通过或短时远离时，距离在阈值附近反复穿越会引发误报。为此，本文引入相对运动信息，构建距离—时间裕度联合风险评估指标。其论点是：近场预警需要显式刻画“接近趋势”而不仅是“当前接近程度”；其依据是 TTC 等指标能够在距离尚未过小前提前触发风险关注；其解释是将几何距离与接近速度结合，可在保证安全性的同时减少无效报警。小结而言，本节给出一组可解释的动态风险量。

设目标在吊钩端参考系下的位置与速度分别为 $\mathbf{p}(t)$ 、 $\mathbf{v}(t)$ ，则目标到吊钩端的距离 $d(t) = \|\mathbf{p}(t)\|$ 。为降低横向噪声对速度估计的影响，定义沿连线方向的接近速度

$$v_{\text{rel}}(t) = -\frac{\mathbf{p}(t)^T \mathbf{v}(t)}{\|\mathbf{p}(t)\|}, \quad (4.10)$$

当 $v_{\text{rel}}(t) > 0$ 表示目标在接近；当 $v_{\text{rel}}(t) \leq 0$ 表示目标远离或主要为横向运动，此时不应由 TTC 触发更高风险等级，可将 TTC 视为 $+\infty$ 或仅依据距离阈值判定。时间到碰撞（Time-to-Collision, TTC）可近似为

$$\text{TTC}(t) = \frac{d(t)}{v_{\text{rel}}(t) + \epsilon}, \quad (4.11)$$

其中 ϵ 为数值稳定项。系统结合 $d(t)$ 与 $\text{TTC}(t)$ 进行风险分级：距离用于约束近场侵入，TTC 用于提供提前量。

在几何距离的定义上，若仅使用质心点距离，可能低估具有体积的车辆与危险源（吊钩端）之间的真实碰撞风险。为增强保守性，可引入安全半径或包围体最小距离：例如将危险源近似为半径 r_h 的球体、车辆近似为半径 r_o 的球体，则有效距离可取 $d_e(t) = \max(d(t) - r_h - r_o, 0)$ 并用于 TTC 计算【待补充：车辆尺寸与安全半径设定依据】。此外，由于速度估计可能因点云稀疏与遮挡产生抖动，建议在使用 $\mathbf{v}(t)$ 前引入滤波/门控（例如第 3 章的跟踪协方差约束或速度平滑）以避免 TTC 因分母抖动产生尖峰。通过这些处理，动态风险指标能够在保持可解释性的同时更贴近工程安全需求。

4.3.3 分级预警状态机

预警输出采用三级状态机：**Normal—Warning—Danger**。其论点是：在在线系统中，风险量 $d(t)$ 与 $\text{TTC}(t)$ 都存在测量噪声与短时异常（遮挡、离群点、关联短时失败），直接用单阈值触发会导致输出频繁跳变，既降低操作可用性，也可能造成“报警疲劳”。其依据是工程经验与系统约束：预警输出应当在“安全优先”的前提下尽量稳定；其解释在于通过滞回与最小持续时间，可把噪声抖动吸收到状态机内部而不是暴露给最终提示。小结而言，本节给出一个可参数化、可复现的分级预警输出层。

状态转移同时受距离阈值与 TTC 阈值约束。设静态/动态统一的关注距离阈值为 $D_{\text{on}}, D_{\text{off}}$ （满足 $D_{\text{on}} < D_{\text{off}}$ ），动态风险的时间阈值为 $T_{\text{warn}}, T_{\text{danger}}$ （满足 $T_{\text{danger}} < T_{\text{warn}}$ ），并引入最小持续时间 Δt_{min} 。一个可实现的规则示例如下：当 $d(t) < D_{\text{on}}$ 或 $(v_{\text{rel}}(t) > 0 \text{ 且 } \text{TTC}(t) < T_{\text{warn}})$ 持续超过 Δt_{min} 时进入 **Warning**；当 $d(t) < D_{\text{danger}}$ 或 $(v_{\text{rel}}(t) > 0 \text{ 且 } \text{TTC}(t) < T_{\text{danger}})$ 持续超过 Δt_{min} 时进入 **Danger**；当 $d(t) > D_{\text{off}}$ 且 $(\text{TTC}(t) > T_{\text{warn}} \text{ 或 } v_{\text{rel}}(t) \leq 0)$ 持续超过 Δt_{min} 时退出到更低等级【待补充：具体阈值表与持续时间设置】。

当同一时刻存在多个危险目标时，系统应遵循“取最危险者”的合成原则：以所有目标的风险等级最大值作为系统输出，并记录触发该等级的目标 ID 以便日志复现与错误分析。对于语义为 **unknown** 的动态目标，系统可选择仅输出几何层面的距离提示而不进入 **Danger**，或采用更保守策略直接按动态目标处理，具体取决于误报容忍度与实验验证结果【待补充：**unknown** 的策略选择】。通过上述设计，预警状态机能够在保证安全性的同时降低误报对作业效率的影响。

5 实验验证与系统评估

5.1 实验平台与测试场景设置

5.1.1 吊钩端实验平台构成

实验平台由吊钩端一体化感知箱体与边缘计算单元构成。吊钩端箱体集成双 Livox Mid-360、海康工业相机（MV-CS016-10UC V5）及必要的供电转换与防护结构；边缘计算单元运行 ROS1（Noetic）并负责数据采集、点云融合、LIO-SAM 定位建图、视觉危险目标检测（YOLOv8，实验中以车辆为代表）、雷视关联与风险评估。系统通过统一时间戳与 TF 管理不同传感器坐标关系，并将预警结果以 ROS 话题形式发布，便于记录日志与离线复现。

为保证实验可复现与可对比，平台在每次实验前进行外参复核（或加载已标定外参），并记录关键运行参数（ROI 范围、滑窗帧数 N 、聚类阈值、跟踪门控阈值、预警阈值等），确保不同工况下的结果具备可比性。

5.1.2 典型测试工况设计

本文围绕吊钩端近场风险形成机制设计典型测试工况，覆盖静态障碍临近、动态目标（车辆）接近、吊钩端摆动与回转引起的视场遮挡等场景。测试工况强调“可复现、可量化、可对比”，在保证安全前提下，设置不同距离、不同接近速度与不同遮挡程度，以便分别评估定位稳定性、目标提取能力、雷视关联一致性与预警提前量。

5.2 多传感器定位性能实验

5.2.1 位姿解算精度验证

位姿精度验证关注“短时稳定性”与“累计漂移”两类指标。由于吊钩端存在摆动与振动，本实验将验证重点放在：（1）里程计输出的连续性（是否出现跳变/丢失）、（2）姿态估计对摆动的响应一致性（滚转/俯仰变化趋势是否合理）、（3）位置估计的重复性（在相同轨迹/相同工况下的回到原点误差）以及（4）静态地图的结构清晰度（墙面/柱体是否成面、是否出现明显重影）。

实验记录 LIO-SAM 输出位姿 $T_w^b(t)$ 与关键帧地图，必要时引入外部基准（如人工测距标定、已知尺寸结构或其他定位系统）对尺度与漂移进行对照。评价指标可采用相邻帧位姿增量统计、闭环回到原点误差、以及地图中典型结构的点到面残差分布等。

5.2.2 吊钩动态摆动场景分析

摆动场景用于检验“强动态条件下点云去畸变与定位/跟踪管线”的鲁棒性。实验通过控制吊钩产生不同幅度与频率的摆动（含回转/加减速过程），对比开启/关闭 IMU 约束（或不同去畸变策略）时的地图重影程度、动态簇稳定性与预警输出抖动情况。重点观察：（1）点云边界是否出现拉伸/弯曲畸变，（2）静态动态分离是否将摆动引起的背景误差误判为动态，（3）动态目标跟踪是否发生频繁丢失或 ID 切换。

该实验为后续阈值设定提供依据：例如滑窗帧数 N 、动态一致性最小持续帧数、以及状态机进入/退出滞回参数等。

5.3 雷视协同感知效果分析

5.3.1 危险目标检测准确性评估（车辆）

危险目标检测准确性评估采用离线标注或抽样人工核验的方式，对 YOLOv8 输出进行统计（实验对象为车辆）。评价指标包括 Precision/Recall、 F_1 值，以及在不同光照（强光/背光/夜间补光）、不同遮挡程度（被构件遮挡、目标局部出画等）下的分组表现。考虑到预警系统更关注“漏检导致的安全风险”，实验同时统计漏检率与连续漏检帧数分布，并评估时序稳定化（滑动平均 + 滞回）对误检/漏检的影响。

5.3.2 雷视关联一致性分析

雷视关联一致性用于验证“双向一致性关联 + 轨迹级绑定”能否稳定输出“车辆”语义。实验在车辆接近、横向穿越与短时遮挡等场景下，统计：（1）关联成功率（满足一致性验证并绑定成功的比例）、（2）错误关联率（将非车辆动态簇误绑定为车辆）、（3）ID 稳定性（绑定关系维持时长与切换次数）以及（4）遮挡鲁棒性（短时丢检情况下语义保持能力）。

同时对比仅使用“3D→2D 投影重叠”的单向关联与本文“双向一致性验证”的差异，以说明反向验证对抑制偶然重叠错配的作用。

5.4 危险区域检测效果验证

5.4.1 静态障碍场景测试

静态障碍测试围绕“最小安全距离 $d_{\min}(t)$ 估计”与“侵入触发稳定性”展开。实验在吊钩端近场布置典型静态结构（如脚手架边缘、堆料区边界或固定构件），并设置不同的最近距离与不同运动方式（直线靠近、侧向掠过、摆动接近）。记录系统输出的 $d_{\min}(t)$ 曲线与预警状态机输出，评估：（1）距离估计的单调性与抖动幅度、（2）进入/退出阈值的滞回效果、（3）误报率（未侵入时触发）与（4）响应延迟（从进入阈值到预警发布的时间）。

5.4.2 动态目标接近场景测试（车辆）

动态目标（车辆）接近测试用于验证“距离 + TTC”的提前量优势。实验设计包含车辆向吊钩端接近、与吊钩端相向运动、以及车辆横穿近场区域等典型工况，并覆盖不同接近速度与不同遮挡条件。系统在雷视融合后输出目标的三维位置 $\mathbf{p}(t)$ 、速度 $\mathbf{v}(t)$ 与语义标签，并基于 $d(t)$ 与 $TTC(t)$ 输出分级预警。评价重点包括：（1）预警提前量（危险发生前的时间裕度）、（2）误报/漏报（尤其是漏报对安全性的影响）、（3）预警等级切换稳定性（是否频繁抖动）以及（4）在短时遮挡/丢检时的连续性。

由于车辆运动/控制仍存在一定随机性（如速度波动、路径偏差等），实验可采用多次重复统计并给出箱线图或分位数描述，以体现不同工况下系统表现的稳定区间。

5.5 系统实时性与稳定性验证

5.5.1 各模块耗时分析

系统实时性通过对关键模块的耗时进行统计评估。本文建议分别记录并分析：点云融合与预处理耗时、LIO-SAM 更新耗时、静态分离与聚类耗时、跟踪与状态量估计耗时、视觉检测耗时、雷视关联耗时、风险评估与状态机耗时，以及端到端延迟（从传感器采样到预警发布）。由于本文当前阶段尚未形成完整的数值统计结果，本节先给出评估口径与记录方式，待采集实验日

志后补充具体耗时数据与瓶颈分析。

5.5.2 工程应用可行性分析

工程可行性主要从安装维护、链路稳定性与长期运行可靠性三个方面分析：其一，吊钩端一体化箱体结构应支持快速安装与外参复用，并具备一定防尘防水与抗振能力；其二，传感器数据链路需在摇摆与遮挡环境下保持可接受的丢包率与延迟；其三，系统应具备异常恢复能力（短时丢帧、短时丢检、时间戳漂移等），并通过滞回与时序稳定化降低误报对作业的干扰。

装

订

线

6 结论与展望

6.1 主要研究结论

本文面向塔吊吊装作业的吊钩端近场高风险区域，构建了双 LiDAR（Livox Mid-360）与工业相机（海康威视 MV-CS016-10UC V5）协同的吊钩端感知系统，建立了多传感器标定、时间同步与坐标系规范；提出双雷达点云实时融合方法，并引入 LIO-SAM 完成吊钩端位姿估计与静态地图构建；在位姿约束下实现近场 ROI 预处理、静动态分离、动态簇提取与短时跟踪，为风险判定提供稳定的几何与运动状态量；进一步结合 YOLOv8 危险目标检测结果（实验中以车辆为代表），设计雷视双向一致性关联与轨迹级语义增强策略，构建基于距离与 TTC 的分级预警状态机，实现对静态障碍与动态目标接近风险的主动提示。上述方法为吊钩端近场安全防护提供了可工程化落地的技术路径。

6.2 创新点总结

本文创新点主要体现在：（1）面向吊钩端近场风险的双 LiDAR + 工业相机一体化感知系统集成与接口规范；（2）融合点云驱动的定位建图与目标生成流程，结合位姿约束实现稳定的静动态分离与动态目标跟踪；（3）面向危险目标（以车辆为代表）的雷视双向一致性关联与置信度融合方法，以及结合距离与时间裕度（TTC）的分级预警策略。

6.3 存在的不足与未来展望

本文仍存在以下不足与改进空间：其一，吊钩端强振动与复杂遮挡可能导致外参微小漂移与点云稀疏，后续可引入在线外参自校准与更鲁棒的去畸变策略；其二，视觉检测在极端光照与严重遮挡条件下仍可能出现漏检，后续可结合多尺度训练、轻量化跟踪与小目标增强策略提升稳定性；其三，预警阈值与状态机参数对工况具有一定依赖性，后续可通过数据驱动方式进行自适应阈值学习，并探索与塔吊控制/指挥系统的闭环联动，实现从“预警提示”到“安全辅助决策/约束”的进一步落地。

参考文献

- [1] 中华人民共和国国家统计局. 2023 中国统计年鉴 [M]. 北京: 中国统计出版, 2023.[M].
- [2] 肖智珺. 塔式起重机智能路径规划研究 [D]. 太原科技大学,2023.[J].,
- [3] SHAPIRA A, LYACHIN B. Identification and Analysis of Factors Affecting Safety on Construction Sites with Tower Cranes[J]. Journal of Construction Engineering and Management, 2009, 135(1): 24-33. 10.1061/(ASCE)0733-9364(2009)135:1(24).
- [4] SHAPIRA A, LUCKO G, SCHEXNAYDER C J. Cranes for Building Construction Projects[J]. Journal of Construction Engineering and Management, 2007, 133(9): 690-700. 10.1061/(ASCE)0733-9364(2007)133:9(690).
- [5] 张达. 2014 年—2022 年塔式起重机安全事故案例统计分析[J]. 建筑安全, 2025, 40(2): 81-85.
- [6] HWANG S. Ultra-wide band technology experiments for real-time prevention of tower crane collisions[J]. Automation in Construction, 2012, 22: 545-553. 10.1016/j.autcon.2011.11.015.
- [7] ALI A H, ZAYED T, WANG R D, et al. Tower crane safety technologies: A synthesis of academic research and industry insights[J]. Automation in Construction, 2024, 163: 105429. 10.1016/j.autcon.2024.105429.
- [8] PRICE L C, CHEN J, PARK J, et al. Multisensor-driven real-time crane monitoring system for blind lift operations: Lessons learned from a case study[J]. Automation in Construction, 2021, 124: 103552. 10.1016/j.autcon.2021.103552.
- [9] FANG Y, CHO Y K, DURSO F, et al. Assessment of operator's situation awareness for smart operation of mobile cranes[J]. Automation in Construction, 2018, 85: 65-75. 10.1016/j.autcon.2017.10.007.
- [10] CHAE S, YOSHIDA T. Application of RFID technology to prevention of collision accident with heavy equipment[J]. Automation in Construction, 2010, 19(3): 368-374. 10.1016/j.autcon.2009.12.008.
- [11] ZHANG C, HAMMAD A, RODRIGUEZ S. Crane Pose Estimation Using UWB Real-Time Location System[J]. Journal of Computing in Civil Engineering, 2012, 26(5): 625-637. 10.1061/(ASCE)CP.1943-5487.0000172.
- [12] 郁志明. 塔吊安全监控系统的设计与研究[D/OL]. 东北大学. 2019. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD%5C&dbname=CMFD201901%5C&filename=1019029559.nh>.
- [13] 张伟, 廖阳新, 蒋灵, 等. 基于物联网的塔式起重机安全监控系统[J]. 中国安全科学学报, 2021, 31(2): 55-62. 10.16265/j.cnki.issn1003-3033.2021.02.008.
- [14] 周飞虎. 塔式起重机防碰撞技术与安全监控系统研究[D/OL]. 长安大学. 2016. <https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD%5C&dbname=CMFD201601%5C&filename=1015802742.nh>.
- [15] ZHONG D, LV H, HAN J, et al. A Practical Application Combining Wireless Sensor Networks and Internet of Things: Safety Management System for Tower Crane Groups[J]. Sensors (Basel, Switzerland), 2014, 14(8): 13794-13814. 10.3390/s140813794.

- [16] 赵宇. 塔吊群作业无线组网技术与防碰撞算法研究[D/OL]. 哈尔滨理工大学. 2017. <https://kn.s.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD%5C&dbname=CMFD201701%5C&filename=1016072789.nh>.
- [17] OFGEOMATRICS S, URBAN SPATIAL INFORMATION B U O C E, ARCHITECTURE C M W I W S O C, Beijing, et al. Research on Intelligent Anti-collision Monitoring for Construction Tower Crane Group Based on GNSS Sensors[J]. International Journal of Computer and Communication Engineering, 2019, 8(4): 169-177. 10.17706/IJCCE.2019.8.4.169-177.
- [18] 张知田, 王园园, 罗柱邦, 等. 塔吊与工人空间交互下危险场景自动检测[J]. 清华大学学报（自然科学版）, 2024, 64(2): 198-204. 10.16511/j.cnki.qhdxxb.2023.22.047.
- [19] KU T K X, ZUO B, ANG W T. Robotic tower cranes with hardware-in-the-loop: Enhancing construction safety and efficiency[J]. Automation in Construction, 2024, 168: 105765. 10.1016/j.autcon.2024.105765.
- [20] FANG Y, CHO Y K, CHEN J. A framework for real-time pro-active safety assistance for mobile crane lifting operations[J]. Automation in Construction, 2016, 72: 367-379. 10.1016/j.autcon.2016.08.025.
- [21] GOLCARENARENJI G, MARTINEZ-ALPISTE I, WANG Q, et al. Machine-learning-based top-view safety monitoring of ground workforce on complex industrial sites[J]. Neural Computing and Applications, 2022, 34(6): 4207-4220. 10.1007/s00521-021-06489-3.
- [22] ZHANG M, GE S. Vision and Trajectory-Based Dynamic Collision Prewarning Mechanism for Tower Cranes[J]. Journal of Construction Engineering and Management, 2022, 148(7): 04022057. 10.1061/(ASCE)CO.1943-7862.0002309.
- [23] SHAPIRA A, ROSENFELD Y, MIZRAHI I. Vision System for Tower Cranes[J]. Journal of Construction Engineering and Management, 2008, 134(5): 320-332. 10.1061/(ASCE)0733-9364(2008)134:5(320).
- [24] XING-YU F, DAN N, QI L, et al. Position-pose measurement of crane sway based on monocular vision[J]. The Journal of Engineering, 2019, 2019(22): 8330-8334. 10.1049/joe.2019.1072.
- [25] JEONG I, HWANG J, KIM J, et al. Vision-Based Productivity Monitoring of Tower Crane Operations during Curtain Wall Installation Using a Database-Free Approach[J]. Journal of Computing in Civil Engineering, 2023, 37(4): 04023015. 10.1061/JCCEE5.CPENG-5105.
- [26] SLEIMAN J P, ZANKOUL E, KHOURY H, et al. Sensor-Based Planning Tool for Tower Crane Anti-Collision Monitoring on Construction Sites[J]., 2016: 2624-2632. 10.1061/9780784479827.261.
- [27] YONG Y P, LEE S J, CHANG Y H, et al. Object Detection and Distance Measurement Algorithm for Collision Avoidance of Precast Concrete Installation during Crane Lifting Process[J]. Buildings, 2023, 13(10): 2551. 10.3390/buildings13102551.
- [28] FANG Y, CHEN J, CHO Y K, et al. Vision-based load sway monitoring to improve crane safety in blind lifts[J]. Journal of Structural Integrity and Maintenance, 2018, 3(4): 233-242. 10.1080/24705314.2018.1531348.
- [29] LEE G, KIM H H, LEE C J, et al. A laser-technology-based lifting-path tracking system for a robotic tower crane[J]. Automation in Construction, 2009, 18(7): 865-874. 10.1016/j.autcon.2009.03.011.

- [30] SIEBERT S, TEIZER J. Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system[J]. Automation in Construction, 2014, 41: 1-14. 10.1016/j.autcon.2014.01.004.
- [31] YANG Z, YUAN Y, ZHANG M, et al. Safety Distance Identification for Crane Drivers Based on Mask R-CNN[J]. Sensors, 2019, 19(12): 2789. 10.3390/s19122789.
- [32] PRICE L C, CHEN J, PARK J, et al. Multisensor-driven real-time crane monitoring system for blind lift operations: Lessons learned from a case study[J]. Automation in Construction, 2021, 124: 103552. 10.1016/j.autcon.2021.103552.
- [33] ZHANG M, GE S. Vision and Trajectory-Based Dynamic Collision Prewarning Mechanism for Tower Cranes[J]. Journal of Construction Engineering and Management, 2022, 148(7): 04022057. 10.1061/(ASCE)CO.1943-7862.0002309.
- [34] WANG Z, WU Y, NIU Q. Multi-Sensor Fusion in Automated Driving: A Survey[J]. IEEE Access, 2020, 8: 2847-2868. 10.1109/ACCESS.2019.2962554.
- [35] ZHOU Y, OMAR M. Pixel-Level Fusion for Infrared and Visible Acquisitions[J]. International Journal of Optomechatronics, 2009, 3(1): 41-53. 10.1080/15599610902717835.
- [36] KIM B, KIM D, PARK S, et al. Automated Complex Urban Driving based on Enhanced Environment Representation with GPS/map, Radar, Lidar and Vision[J]. IFAC-PapersOnLine, 2016, 49(11): 190-195. 10.1016/j.ifacol.2016.08.029.
- [37] LEKIC V, BABIC Z. Automotive radar and camera fusion using Generative Adversarial Networks[J]. Computer Vision and Image Understanding, 2019, 184: 1-8. 10.1016/j.cviu.2019.04.002.
- [38] OUYANG Z, LIU Y, ZHANG C, et al. A cGANs-Based Scene Reconstruction Model Using Lidar Point Cloud[C/OL]//. 2017: 1107-1114. <https://ieeexplore.ieee.org/abstract/document/8367397>. 10.1109/ISPA/IUCC.2017.00167.
- [39] CALTAGIRONE L, BELLONE M, SVENSSON L, et al. LIDAR-camera fusion for road detection using fully convolutional neural networks[J]. Robotics and Autonomous Systems, 2019, 111: 125-131. 10.1016/j.robot.2018.11.002.
- [40] WU T E, TSAI C C, GUO J I. LiDAR/camera sensor fusion technology for pedestrian detection[C/OL]//. 2017: 1675-1678. <https://ieeexplore.ieee.org/abstract/document/8282301>. 10.1109/APSIPA.2017.8282301.
- [41] HAN X, LU J, TAI Y, et al. A real-time LIDAR and vision based pedestrian detection system for unmanned ground vehicles[C/OL]//. 2015: 635-639. <https://ieeexplore.ieee.org/abstract/document/7486580>. 10.1109/ACPR.2015.7486580.
- [42] ZHONG Z, LIU S, MATHEW M, et al. Camera Radar Fusion for Increased Reliability in ADAS Applications[J]. Electronic Imaging, 2018, 30: 1-4. 10.2352/ISSN.2470-1173.2018.17.AVM-258.
- [43] WEI P, CAGLE L, REZA T, et al. LiDAR and Camera Detection Fusion in a Real-Time Industrial Multi-Sensor Collision Avoidance System[J]. Electronics, 2018, 7(6): 84. 10.3390/electronics7060084.
- [44] DE SILVA V, ROCHE J, KONDOZ A. Robust Fusion of LiDAR and Wide-Angle Camera Data for Autonomous Mobile Robots[J]. Sensors, 2018, 18(8): 2730. 10.3390/s18082730.

- [45] ZHANG Z, PAN W. Lift planning and optimization in construction: A thirty-year review[J]. Automation in Construction, 2020, 118: 103271. 10.1016/j.autcon.2020.103271.
- [46] CHI H L, KANG S C. A physics-based simulation approach for cooperative erection activities[J]. Automation in Construction, 2010, 19(6): 750-761. 10.1016/j.autcon.2010.03.004.
- [47] KUCHKUDA R. An Introduction to Ray Tracing[C]//EARNshaw R A. Berlin, Heidelberg: Springer, 1988: 1039-1060. 10.1007/978-3-642-83539-1_44.
- [48] DUTTA S, CAI Y, HUANG L, et al. Automatic re-planning of lifting paths for robotized tower cranes in dynamic BIM environments[J]. Automation in Construction, 2020, 110: 102998. 10.1016/j.autcon.2019.102998.
- [49] ZHU A, ZHANG Z, PAN W. Crane-lift path planning for high-rise modular integrated construction through metaheuristic optimization and virtual prototyping[J]. Automation in Construction, 2022, 141: 104434. 10.1016/j.autcon.2022.104434.
- [50] YANG B, ZHANG H, SHEN Z. Minimum Distance Calculation Method for Collision Issues in Lifting Construction Scenarios[J]. Journal of Computing in Civil Engineering, 2024, 38(6): 04024041. 10.1061/JCCEE5.CPENG-5914.
- [51] LAI K C, KANG S C. Collision detection strategies for virtual construction simulation[J]. Automation in Construction, 2009, 18(6): 724-736. 10.1016/j.autcon.2009.02.006.
- [52] WU H, TAO J, LI X, et al. A location based service approach for collision warning systems in concrete dam construction[J]. Safety Science, 2013, 51(1): 338-346. 10.1016/j.ssci.2012.08.006.
- [53] WANG C C, WANG M, SUN J, et al. A Safety Warning Algorithm Based on Axis Aligned Bounding Box Method to Prevent Onsite Accidents of Mobile Construction Machineries[J]. Sensors, 2021, 21(21): 7075. 10.3390/s21217075.
- [54] YANG Z, YUAN Y, ZHANG M, et al. Safety Distance Identification for Crane Drivers Based on Mask R-CNN[J]. Sensors, 2019, 19(12): 2789. 10.3390/s19122789.
- [55] NADIMI N, BEHBAHANI H, SHAHBAZI H. Calibration and validation of a new time-based surrogate safety measure using fuzzy inference system[J]. Journal of Traffic and Transportation Engineering (English Edition), 2016, 3(1): 51-58. 10.1016/j.jtte.2015.09.004.
- [56] KIM D, LIU M, LEE S, et al. Trajectory Prediction of Mobile Construction Resources Toward Proactive Struck-by Hazard Detection[C/OL]//. Banff, AB, Canada, 2019. http://www.iaarc.org/publications/2019%5C_proceedings%5C_of%5C_the%5C_36th%5C_isarc/trajectory%5C_prediction%5C_of%5C_mobile%5C_construction%5C_resources%5C_toward%5C_pro%5C_active%5C_struck%5C_by%5C_hazard%5C_detection.html. 10.22260/ISARC2019/0131.
- [57] 张冬. 塔式起重机智能监控系统研究与开发[D/OL]. 上海交通大学. 2020. <https://doi.org/10.27307/d.cnki.gsjtu.2016.001713>.
- [58] RASHID K, DATTA S, BEHZADAN A. Coupling risk attitude and motion data mining in a pre-emptive construction safety framework[M]. 2017.
- [59] TANG S, GOLPARVAR-FARD M, NAPHADE M, et al. Video-Based Motion Trajectory Forecasting Method for Proactive Construction Safety Monitoring Systems[J]. Journal of Computing in Civil Engineering, 2020, 34(6): 04020041. 10.1061/(ASCE)CP.1943-5487.0000923.

- [60] KIM D, LEE S, KAMAT V R. Proximity Prediction of Mobile Objects to Prevent Contact-Driven Accidents in Co-Robotic Construction[J]. Journal of Computing in Civil Engineering, 2020, 34(4): 04020022. 10.1061/(ASCE)CP.1943-5487.0000899.
- [61] SUNKARA V, CHAKRAVARTHY A. Collision avoidance laws for objects with arbitrary shapes[C/OL]// . 2016: 5158-5164. <https://ieeexplore.ieee.org/abstract/document/7799058>. 10.1109/CDC.2016.7799058.
- [62] PARK J, CHO N. Collision Avoidance of Hexacopter UAV Based on LiDAR Data in Dynamic Environment[J]. Remote Sensing, 2020, 12(6): 975. 10.3390/rs12060975.
- [63] ZHANG X, LINIGER A, BORRELLI F. Optimization-Based Collision Avoidance[J]. IEEE Transactions on Control Systems Technology, 2021, 29(3): 972-983. 10.1109/TCST.2019.2949540.
- [64] ZHU A, ZHANG Z, PAN W. Developing a fast and accurate collision detection strategy for crane-lift path planning in high-rise modular integrated construction[J]. Advanced Engineering Informatics, 2024, 61: 102509. 10.1016/j.aei.2024.102509.
- [65] LIN X, HAN Y, GUO H, et al. Lift path planning for tower cranes based on environmental point clouds[J]. Automation in Construction, 2023, 155: 105046. 10.1016/j.autcon.2023.105046.
- [66] FIORINI P, SHILLER Z. Motion Planning in Dynamic Environments Using Velocity Obstacles[J]. The International Journal of Robotics Research, 1998, 17(7): 760-772. 10.1177/027836499801700706.
- [67] BERG J, LIN M, MANOCHA D. Reciprocal Velocity Obstacles for real-time multi-agent navigation[C/OL]// . 2008: 1928-1935. <https://ieeexplore.ieee.org/abstract/document/4543489>. 10.1109/ROBOT.2008.4543489.
- [68] JIAN Z, YAN Z, LEI X, et al. Dynamic Control Barrier Function-based Model Predictive Control to Safety-Critical Obstacle-Avoidance of Mobile Robot[EB/OL]. 2022. <https://arxiv.org/abs/2209.08539v1>.
- [69] SHAN T, ENGLLOT B J, MEYERS D, et al. LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping[J/OL]. CoRR, 2020, abs/2007.00258. arXiv: 2007.00258. <https://arxiv.org/abs/2007.00258>. 10.48550/arXiv.2007.00258.