

点云融合技术综述：方法、应用与挑战

宋绍京^① 李新建^② 方非易^{*①}

^①(上海第二工业大学计算机与信息工程学院 上海 201209)

^②(上海第二工业大学智能制造与控制工程学院 上海 201209)

摘要：点云融合技术作为3D (Three-Dimensional)数据处理的重要手段，在多个领域展现出巨大的潜力和应用前景。该文系统地综述了点云融合的基础概念、常用技术方法及其应用，深入分析了不同方法的发展现状和未来发展趋势。此外，该文还探讨了点云融合在自动驾驶、建筑和机器人等领域的实际应用及面临的挑战，尤其是在应对噪声、数据稀疏性和密度不均等问题时，如何在保证融合精度的同时平衡其复杂性。通过全面梳理现有研究进展，为未来点云融合技术的发展提供了有力参考，并为进一步提升融合算法的精度、鲁棒性和效率指明了可能的研究方向。

关键词：点云融合；3D数据处理；特征匹配；融合算法；深度学习

中图分类号：TN95

文献标识码：A

文章编号：2095-283X(2025)03-0528-20

DOI: 10.12000/JR24236

CSTR: 32380.14.JR24236

引用格式：宋绍京, 李新建, 方非易. 点云融合技术综述：方法、应用与挑战[J]. 雷达学报(中英文), 2025, 14(3): 528–547. doi: 10.12000/JR24236.

Reference format: SONG Shaojing, LI Xinjian, and FANG Feiyi. A review of point cloud fusion technology: Methods, applications, and challenges[J]. *Journal of Radars*, 2025, 14(3): 528–547. doi: 10.12000/JR24236.

A Review of Point Cloud Fusion Technology: Methods, Applications, and Challenges

SONG Shaojing^① LI Xinjian^② FANG Feiyi^{*①}

^①(School of Computer and Information Engineering, Shanghai Polytechnic University, Shanghai 201209, China)

^②(School of Intelligent Manufacturing and Control Engineering, Shanghai Polytechnic University, Shanghai 201209, China)

Abstract: As an important method of 3D (Three-Dimensional) data processing, point cloud fusion technology has shown great potential and promising applications in many fields. This paper systematically reviews the basic concepts, commonly used techniques, and applications of point cloud fusion and thoroughly analyzes the current status and future development trends of various fusion methods. Additionally, the paper explores the practical applications and challenges of point cloud fusion in fields such as autonomous driving, architecture, and robotics. Special attention is given to balancing algorithmic complexity with fusion accuracy, particularly in addressing issues like noise, data sparsity, and uneven point cloud density. This study serves as a strong reference for the future development of point cloud fusion technology by providing a comprehensive overview of the existing research progress and identifying possible research directions for further improving the accuracy, robustness, and efficiency of fusion algorithms.

Key words: Point cloud fusion; 3D data processing; Feature matching; Fusion algorithm; Deep learning

收稿日期：2024-11-24；改回日期：2025-06-03；网络出版：2025-06-13

*通信作者：方非易 fxfang@sspu.edu.cn

*Corresponding Author: FANG Feiyi, fxfang@sspu.edu.cn

责任主编：陈育伟 Corresponding Editor: CHEN Yuwei

©The Author(s) 2025. This is an open access article under the CC-BY 4.0 License

(<https://creativecommons.org/licenses/by/4.0/>)

1 引言

点云(Point Cloud)是一种用于表达3D (Three-Dimensional)空间物体表面的重要数据形式,通常由大量不规则分布的3D点组成。每个点包含其在3D空间中的位置信息,通常还附加有其他属性,包括强度值等。与其他3D数据表示方法(如体素或网格)相比,点云保留了原始的3D几何信息,表示方式更为直接、简洁,能够在不依赖拓扑结构的情况下捕捉复杂的几何特征^[1]。因此,点云数据为3D计算机视觉领域的3D重建、物体识别和场景理解等任务提供了丰富的几何细节,尤其在需要精确测量物体形状和表面特征的应用中,点云数据具有显著优势,这一优势也使点云成为自动驾驶、机器人导航、增强现实(Augmented Reality, AR)和虚拟现实(Virtual Reality, VR)等领域不可或缺的数据源^[2-4]。然而,点云数据的稀疏性、不规则性和高维性也给数据处理与分析带来了挑战,尤其是在大规模场景下^[5,6]。点云数据的处理方法需要应对计算复杂度高、噪声难以消除以及遮挡处理等技术难点。

在实际应用中,由单一视角或单一传感器生成的点云数据往往难以完整呈现整个场景或目标的几何形态。点云融合技术旨在通过整合来自多个传感器、视角或时间帧的点云数据,生成更加完整、精确的3D表示。这种融合方法不仅能够弥补单一传感器或视角的局限性,还能显著提高目标检测、识别和定位的精度。例如,在自动驾驶领域,激光雷达(Light Detection And Ranging, LiDAR)传感器生成的点云提供了精确的距离信息,而毫米波雷达则能够穿透恶劣天气条件(如雨、雾、雪等),提供稳定的速度和距离测量。通过融合这两类数据,自动驾驶系统能够更好地感知复杂环境,并识别行人、车辆及其他动态物体^[7]。点云融合在机器人领域也发挥了重要作用,尤其是在即时定位与地图构建(Simultaneous Localization And Mapping, SLAM)等任务中^[8]。通过融合多帧点云数据,机器人能够利用不同时间点采集的环境信息,弥补单帧点云中的信息缺失和噪声问题,从而构建更加精确和稳定的3D地图。此外,在AR/VR领域,多视角点云融合能够提升虚拟场景的逼真度。通过融合多帧数据,虚拟物体能够与现实世界更加精准地对齐和互动^[9]。点云融合技术的应用不仅限于此,还在建筑^[10]、医疗成像^[11]和文物保护^[12-15]等多个领域展现了巨大的潜力。

本文旨在对点云融合技术进行全面综述,系统总结其主要方法、研究进展及应用。首先,本文将

介绍点云融合的基本概念及相关数据集;随后,讨论点云融合的分类及常用技术方法,深入分析了不同方法的发展现状和未来发展趋势;最后,本文将探讨点云融合技术的实际应用及其在实践中面临的挑战,为研究人员提供新的思路与方向。

2 点云融合

2.1 点云融合的基本概念

2.1.1 点云数据及点云融合概述

点云作为一种稀疏且无结构的3D数据表示形式,由大量离散的3D点构成,每个点不仅携带其空间坐标,还包含颜色、法向量等附加属性。相较于体素、网格和深度图等其他3D数据形式,点云有其独特优势。与体素网格相比,点云具有更高的存储效率,因其仅记录物体表面的关键点,而非填充整个3D空间;相较于多边形网格,点云不包含物体表面的拓扑结构,使其在处理复杂、不规则的几何形状时更加灵活简便;相比深度图,点云提供了全方位的3D信息,而非只能呈现单视角下的物体表面深度数据。

点云数据的采集通常依赖多种传感器技术,常见设备包括LiDAR、RGB-D相机和结构光设备^[16]。不同传感器生成的点云如图1所示,图1(a)是由深度相机产生的RGB-D点云,图1(b)是由3D扫描仪生成的稠密结构光点云以及图1(c)是由LiDAR生成的稀疏点云。RGB-D相机通过结合红外传感器和RGB摄像头,能够同时捕获彩色图像和深度信息,生成密集点云,广泛应用于室内场景及机器人导航,但在远距离和复杂环境下,其深度精度会受到限制;结构光设备通过投射特定光栅图案并分析其变形来获取物体表面的深度信息,具有较高的精度,但仅适用于短距离的精细物体建模;LiDAR通过发射激光脉冲并接收反射信号来测量物体与传感器之间的距离,生成高精度的3D点云,适用于大范围场景的感知,如自动驾驶,然而,LiDAR生成的点云数据稀疏且缺乏颜色信息。不同传感器技术各有优劣,如何根据具体应用需求选择合适的采集方式,对于点云数据的质量至关重要。

尽管点云能够有效捕获原始空间数据,直观地表达物体的几何特征,但由于其无序性、稀疏性和不规则性,缺乏连通性信息,使得后续处理更具挑战性,特别是在噪声处理和点云密度补全方面仍面临显著挑战^[17,18]。

点云融合是指通过整合来自不同传感器、不同视角或不同时间帧的点云数据,生成更完整、更精确3D点云的技术方法。其核心目标在于克服单一

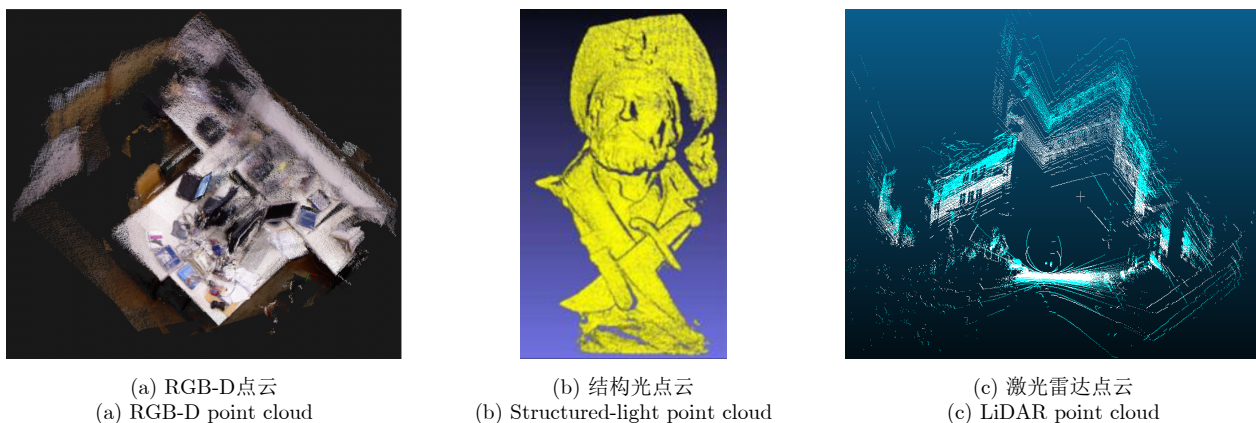


图1 点云数据表示

Fig. 1 Point cloud data representation

传感器或视角的局限性,通过多个数据源提取更多的几何和语义信息,从而提升数据的整体质量^[16]。点云融合通常涵盖多个维度的数据组合,例如多源点云融合(如LiDAR与RGB-D相机的融合)、多帧点云融合(从不同时间帧连续采集的数据融合)以及多视角点云融合(由单个设备从不同视角采集同一场景的数据融合)。通过点云融合,能够实现数据增强,将不同来源的数据整合为更加密集、完整的点云表示,以补充单一数据中的信息缺失。同时,点云融合还能显著提高3D场景的重建精度,因为它能够对不同数据源的冗余信息进行对比优化,从而减少噪声和误差^[19]。总之,点云融合的最终目的是提升点云数据的完整性、精度和可用性,从而在不同场景中提供更加可靠的3D感知能力。

2.1.2 点云融合与点云配准的区别与联系

点云配准(Point Cloud Registration)和点云融合(Point Cloud Fusion)是点云处理中的两个重要概念,尽管它们在处理过程中的作用不同,但它们通常是相互关联的。在点云处理过程中,配准是融合的基础,融合则是在配准的前提下进行数据整合的过程。

点云配准通常指的是将来自不同传感器、不同时间或不同视角的点云数据对齐到同一个坐标系中。这一过程的目标是确保多个点云数据之间的位置和方向关系准确,以便能够进行后续分析与处理。配准过程通常包括粗配准和精配准,粗配准通常依赖于特征匹配,精配准则依赖于优化算法,如迭代最近点算法(Iterative Closest Point, ICP)和正态分布变换算法(Normal Distributions Transform, NDT)。

点云融合则是在完成配准后,将多个点云数据集成,以形成一个完整且连续的三维点云。融合不仅仅是对齐点云数据,更包括对不同来源、不同时

间或不同视角的点云数据进行有效整合,以消除冗余数据,填补缺失部分,并提升最终三维点云的精度与可靠性。

尽管点云配准和点云融合有不同的定义和目标,但它们是紧密联系的:配准为融合提供了空间对齐的基础,融合则依赖于配准的精度来实现数据的综合与优化。换句话说,点云配准是点云融合的前提和基础,点云融合则是配准之后的一步重要处理,目的是构建更高精度的三维点云。

2.2 点云融合的分类

根据数据来源和采集方式,点云融合可分为多源点云融合、多帧点云融合和多视角点云融合。每种融合方式具有其特点与优势,能够在不同程度上提高点云数据的完整性和精度,满足不同应用场景中对环境感知、3D重建以及动态分析的需求。

2.2.1 多源点云融合

多源点云融合是指将来自不同传感器的数据对齐并融合,以实现更全面的场景表达。常见的点云采集传感器包括LiDAR、RGB-D摄像头、毫米波雷达等。每种传感器都有其优势和局限性,例如LiDAR可以精确测量距离,生成高精度的3D点云,但在恶劣天气条件下性能可能受限^[20];RGB-D摄像头能够提供丰富的视觉信息,但缺乏精确的距离测量功能^[21];毫米波雷达能够在恶劣天气中保持良好的探测能力,但其分辨率较低。通过融合这些多源数据,系统能够补偿单一传感器的不足,从而提高整体的感知能力。图2为采用点云融合算法对铁路桥梁的3种测量技术进行融合的示例,铁路桥的点云捕获技术和融合算法的详细信息如表1所示。

2.2.2 多帧点云融合

多帧点云融合是一种通过整合同一传感器在不

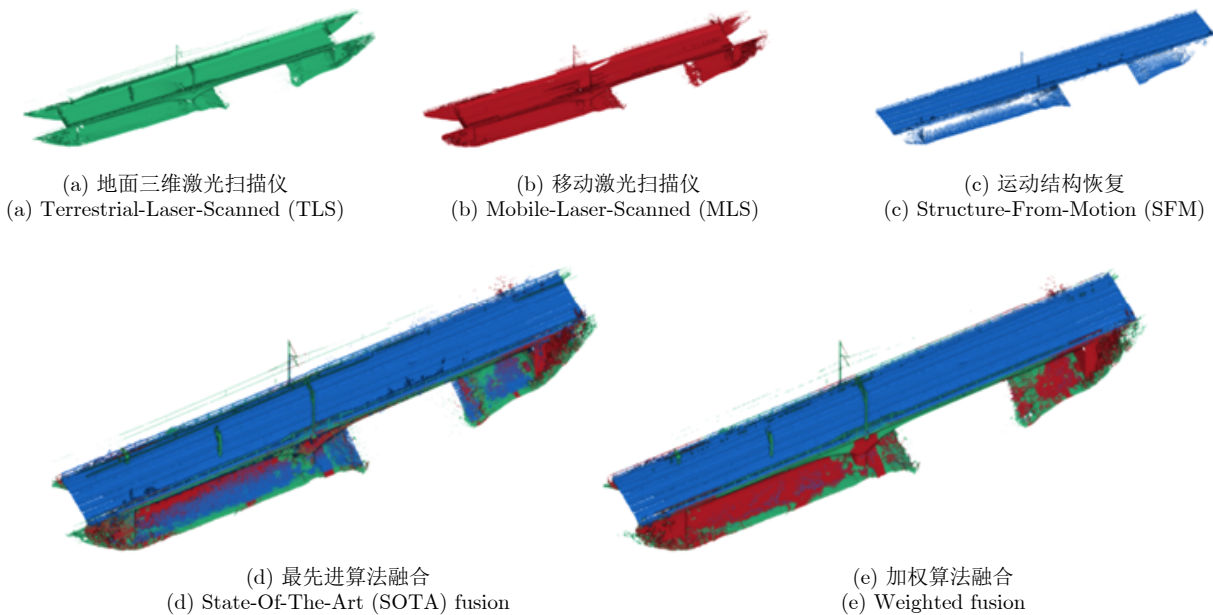


图 2 多源点云融合示例图

Fig. 2 Example of multi-source point cloud fusion

表 1 铁路桥梁点云采集技术及融合算法

Tab. 1 Point cloud from capturing techniques and fusion algorithms of the railway bridge

点云数据	点的数量	精度(m)	平均覆盖率(%)	平均密度(标准差)(点/0.1 m ³)
TLS: Leica P20	12,669,642	0.003	65.64	173 (102)
MLS: Navvis VLX2	18,401,315	0.006	65.30	259 (71)
SfM: DJI P4 + Pix4D	1,322,080	0.050	43.37	20.0 (8.0)
SOTA fusion	32,393,037	0.050	100	372 (146)
Weighted fusion	23,808,697	0.025	100	266 (14)

同时刻采集的点云数据来提高环境感知精度和场景重建效果的方法。与多源点云融合不同，多帧点云融合专注于整合连续采集的点云数据，通常用于解决动态环境感知、场景重建、同步定位与地图构建中的问题。通过在不同时间点采集多帧点云，系统能够利用时间序列信息，生成更加完整和准确的场景表示。这种方法尤其适用于处理动态场景中的移动物体或观察环境的变化。图3为多帧点云融合的示例图，图3中不同颜色表示不同帧的点云数据，黑色虚线框表示发生运动模糊的区域。

2.2.3 多视角点云融合

多视角点云融合是指通过从不同视角采集的点云数据来提升3D场景的重建精度和完整性。每个视角通常只能捕捉到部分场景，存在盲区或遮挡的情况，导致单一视角下的点云数据不够完整。通过融合多个视角的数据，系统能够整合不同角度的信息，生成更全面的场景表示^[22]。这种方法广泛应用

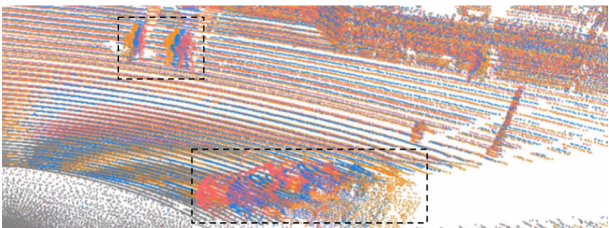


图 3 多帧点云融合示例图

Fig. 3 Example of multi-frame point cloud fusion

于建筑建模、无人机测绘、AR和VR等领域。图4展示了3D重建中的多视角点云融合示例，绿色视角的点云填补了蓝色视角中的盲区，从而使得最终生成的模型更加完整和精确。

2.3 常用点云融合数据集

在点云融合研究中，数据集的选择至关重要。不同的点云融合任务对应不同类型的数据集，每种融合方式的数据集都有其特性，既能反映不同传感

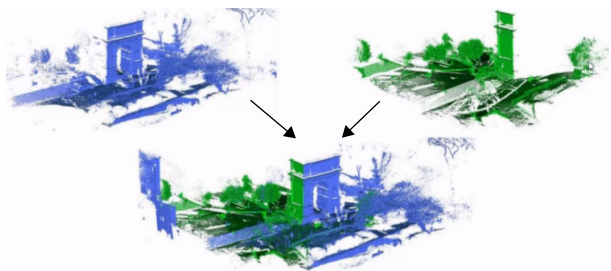


图 4 多视角点云融合示例图

Fig. 4 Example of multi-view point cloud fusion

器数据的融合,也能体现时间或视角上的变化。以KITTI^[23]和nuScenes^[24]为代表的多源点云融合数据集提供了多传感器同步数据,适用于自动驾驶和机器人感知系统的研究,然而,这些数据集的处理难度较高,对存储和计算能力有较高要求;以Oxford RobotCar^[25]为代表的数据集提供了时间序列数据,适合动态场景重建和SLAM等应用,但长时间序列带来了存储与处理的挑战;斯坦福3D扫描模型库^[26]和UWA3M数据集^[27]提供了多视角采集数据,适用于精确的3D重建任务,但通常局限于特定的场景(如室内或固定场景)。表2列举了典型的点云融合相关数据集。

3 点云融合的常用方法

点云融合是将来自不同传感器、不同时间或不同视角的点云数据整合成一个统一、一致的三维点云的过程。本节基于第2节分类方法深入剖析点云融合的方法,为点云融合技术的工程实践提供理论指导与方法论支撑。

3.1 多源点云融合的常用方法

多源点云融合旨在将来自不同传感器或不同位置的原始点云数据整合为统一且鲁棒的三维表示。根据融合阶段的不同,可分为前融合(数据融合)、中间融合(特征融合)和后融合(决策融合)3类方法。图5以时间线形式系统梳理了多源点云融合的3种主要策略,展示了从2017年至今各策略的关键性研究文献及其演进关系。

前融合(数据融合)直接在数据层面操作原始点云,典型方法包括叠加、滤波与加权融合。如图6所示,基于多源点云前融合策略在三维重建初期即整合地面激光扫描(TLS)、无人机航拍(UAS)及索尼地面相机(Sony α7R)的多源点云数据。通过时空同步采集与联合标定,统一全局坐标系,确保

表 2 点云融合相关数据集
Tab. 2 Point cloud fusion related datasets

类型	数据集名称	主要传感器	优势	劣势
多源点云融合	KITTI ^[23]	LiDAR、RGB摄像头、GPS	提供真实驾驶场景,广泛应用于自动驾驶领域	场景复杂性较低,较少覆盖复杂动态场景
	nuScenes ^[24]	LiDAR、雷达、RGB摄像头、IMU	多种传感器数据同步,场景多样,标注丰富	数据量大,处理难度高
	Waymo Open Dataset ^[28]	LiDAR、摄像头	更大规模的数据,覆盖复杂的城市驾驶场景	数据集体量庞大,计算资源需求较高
	A2D2 ^[29]	LiDAR、RGB摄像头、IMU、雷达	提供详细的传感器信息,丰富的驾驶场景	数据集较新,社区支持相对较少
	ApolloScape ^[30]	LiDAR、RGB摄像头、GPS	大规模城市场景,支持3D点云与图像融合	数据质量不及KITTI等较为标准
多帧点云融合	Oxford RobotCar ^[25]	LiDAR、相机、GPS、IMU	涵盖不同天气、光照条件下的长时间序列点云	数据复杂度高,长时间处理需要大量存储资源
	Argoverse ^[31]	LiDAR、RGB摄像头、GPS	提供不同时间点的数据,支持多帧点云融合和动态场景检测	标注信息不如其他自动驾驶数据集丰富
	TUM RGB-D ^[32]	RGB-D相机	提供室内动态场景的连续点云与图像	局限于室内环境,传感器类型单一
	Kaist Urban ^[33]	LiDAR、相机、GPS、IMU	多帧时序数据丰富,场景涵盖日夜不同条件	主要集中于城市场景
多视角点云融合	斯坦福3D扫描模型库 ^[26]	用斯坦福大型雕塑扫描仪、Cyber ware 3030 MS扫描仪	包含9种不同目标的多视图扫描点云数据	无真实纹理或光照变化;数据规模较小,场景复杂度低
	UWA3M 数据集 ^[27]	RGB-D相机	提供物体实例级标注	场景多样性不足(主要为桌面物体)
	Augmented ICL-NUIM 数据集 ^[34]	合成RGB-D数据(模拟Kinect)	大型室内场景数据集,包含Living room 1, Living room 2, Office 1和Office 2共4个场景序列	合成数据与真实传感器数据存在域差异;场景复杂度较低(多为室内简单布局)
	3DMatch ^[35]	RGB-D相机	丰富的室内场景点云数据,适合多视角配准和3D重建	场景局限于室内,缺少室外复杂场景

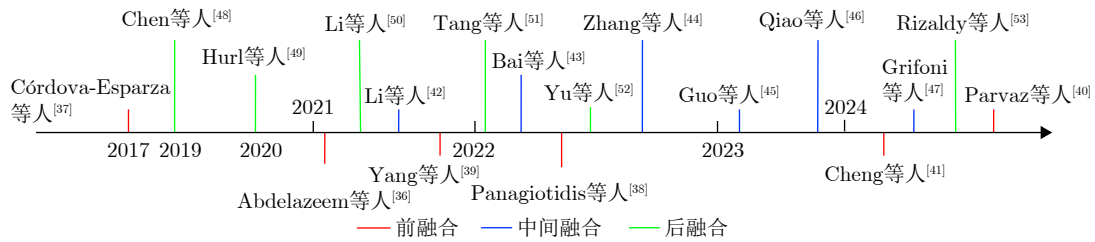
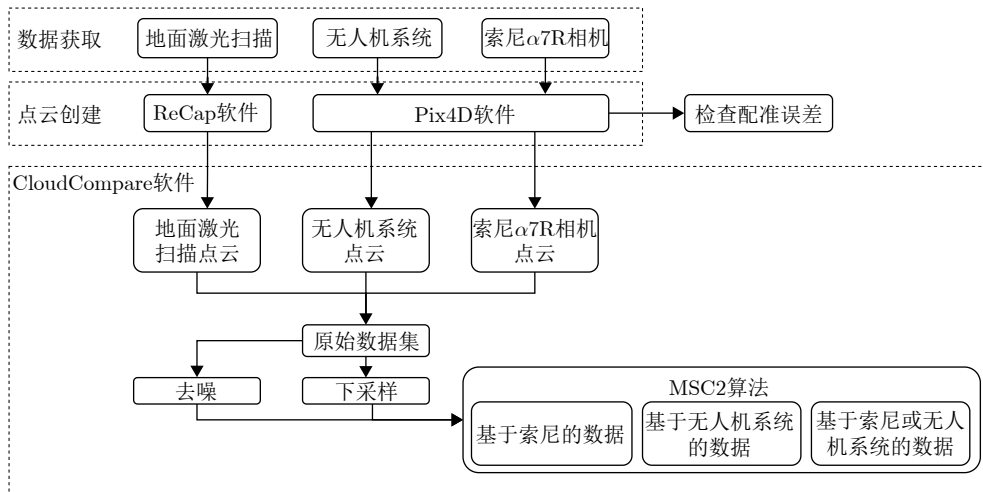


图5 多源点云融合时间线

Fig. 5 Multi-source point cloud fusion timeline

图6 前融合3D模型开发和评估的流程图^[36]Fig. 6 Flowchart of pre-fusion 3D model development and evaluation^[36]

TLS高精度点云、UAS大范围覆盖及Sony细节纹理数据的动态对齐。利用Autodesk ReCap与Pix4D软件协同生成TLS/UAS/Sony复合点云，通过点云直接合并的方法生成完整的3D点云^[36]。近年来，前融合技术经历了从粗粒度叠加到动态精细化的显著演进。早期方法多采用粗粒度叠加策略，如Abdelazeem等人^[36]提出的基于地面控制点的多源点云直接合并技术，该方法通过预先布设控制点实现异构点云对齐，虽能生成完整3D点云，但存在数据冗余度高、对齐精度依赖人工布点、难以适应动态场景等局限性；为克服早期方法的局限性，研究者逐步转向精细化融合框架，Córdova-Esparza等人^[37]突破性地提出多相机联合标定工具，通过优化标定参数实现多源点云无缝叠加，消除了人工布设控制点的需求，显著简化了文化遗产数字化流程，然而，该方法仍属于静态融合范畴，对传感器运动或环境变化的适应性不足；在动态适应性优化阶段，Panagiotidis等人^[38]创新性地引入几何一致性约束，通过无人机与地面LiDAR点云的互补性融合，在森林场景中构建植被三维结构，有效缓解了单传感器因遮挡导致的数据缺失问题，该方法虽提升了复杂场景重建能力，但未考虑点云时空特性差异；为此，Yang等人^[39]设计动态加权融合机制，根据机

载LiDAR与摄影测量点云的密度特征实时调整融合权重，使城市树木参数提取精度提升23.6%，标志着前融合开始具备在线自适应能力；为进一步解决动态采集中的畸变问题，Parvaz等人^[40]提出切片调整策略，通过逐层优化航空点云的空间一致性，成功抑制飞行器运动引起的结构畸变，使城市场景融合误差降低至0.15 m以内，该方法的局限性在于未区分场景语义特征，Cheng等人^[41]的最新研究则通过置信度驱动的全景LiDAR融合框架，筛选高置信度点云优化走廊重建细节，在保证计算效率的同时，还提高了墙面结构完整性。上述技术演进表明，前融合已从简单的空间叠加发展为包含动态权重调节、时空一致性约束、置信度评估的多层次精细化融合体系。

中间融合(特征融合)通过提取并整合多源点云的高层特征，实现鲁棒的跨传感器感知，其相关技术呈现出从动态加权融合到深度语义边缘融合的演进路径。早期研究突破传统手工加权局限，Li等人^[42]于2021年首创拉普拉斯多尺度融合框架，通过频域特征自适应加权将地质灾害监测准确率提升至89.5%，较传统方法提高21.3%，首次验证动态权重分配的技术优势。2022年，Bai等人^[43]引入注意力机制，动态融合车载与路侧LiDAR特征，使3D目标检测

mAP达72.4%，如图7所示，其提出的PillarGrid Network网络框架通过协同变换、柱状体素化和网格级特征拼接实现多源点云的高效对齐与深度融合；其结合路侧与车载服务器的实时交互，在复杂场景(如遮挡、远距目标)中平衡计算效率与检测精度，提升了感知鲁棒性，但其权重计算仍依赖预设空间参数(误差容忍 $\pm 0.5\text{m}$)，暴露出对动态场景适应性不足的缺陷。同期，Zhang等人^[44]构建U-Net深度语义网络，通过端到端学习实现跨传感器特征嵌入，将城市建筑模型提取误差压缩至0.23 m，为语义驱动融合奠定基础。传统方法在轻量级场景中持续发挥作用，如Guo等人^[45]针对林业监测提出高度-密度特征线性加权算法，实现83.2%参数估计精度；深度学习则推动自适应融合技术的快速发展，如Qiao等人^[46]开发的S-AdaFusion模块通过在线学习机制动态调整传感器权重，在复杂交通场景中提升目标跟踪成功率的同时降低误检率。至2024年，相关技术进一步向语义边缘融合深化，Grifoni等人^[47]融合几何与纹理特征提升文化遗产细节还原度至76%。

后融合(决策融合)在感知结果层面整合多传感器的独立输出，强调灵活性与异构兼容性，其进化路径呈现从独立决策、协同优化到反馈增强的阶梯式发展特征。早期独立感知融合阶段采用松耦合架构，如Chen等人^[48]在2019年提出的贝叶斯推理框架，通过独立处理车载LiDAR与摄像头检测结果后执行概率融合，虽然车辆检测精度较高，但存在数据冗余度高、时空对齐误差累积等缺陷。Hurl等人^[49]于2020年构建的TruPercept信任模型，通过量化各传感器检测结果的可信度，使目标关联错误率降低，但未突破“先感知后融合”的串行架构限制。至协同优化阶段，技术转向时空深度耦合。Li等人^[50]在2021年设计非合作目标姿态估计框架，通过时空关联多源点云检测结果优化轨迹预测，将深空目标

姿态估计误差降低30%。Tang等人^[51]于2022年提出多专家学习框架，旨在通过融合不同模型的检测结果提升目标检测的鲁棒性，图8为该框架中的层次化微调流程，首先，针对未标记的点云数据，从源域中提取特征并生成初始检测结果，利用重叠度指标(OIoU)对结果进行分层筛选：在层级0阶段($0.9 < \text{OIoU} < 1.0$)，高重叠度的结果被赋予“真”标签和初始置信度，直接保留为可靠检测；在层级1和层级2阶段($0 < \text{OIoU} \leq 0.9$)，中低重叠度的结果通过多专家学习框架进行特征级和结果级的自适应融合，充分利用不同模型的互补性优化检测精度，最终，所有层级的输出被整合，形成覆盖全重叠范围的融合结果，该框架通过联合优化行人检测框的几何一致性与语义置信度，在KAIST数据集上使漏检率降至6.3%，标志着协同感知从简单结果聚合向深度特征交互演进。同期，Yu等人^[52]开发的TCLF系统通过抽象化检测结果与动态时间补偿，在车路协同场景中平衡实时性与精度需求。当前，后融合研究进入反馈增强阶段，Rizaldy等人^[53]构建维度自适应融合机制，通过动态匹配高光谱与LiDAR点云的维度特征，使农田地物分类效率提升，较传统方法加速至3.2倍，该技术突破性引入反馈控制环路，根据分类置信度实时调整融合策略，形成“感知、融合、评估、策略更新”的闭环架构，但面临反馈延迟导致的决策滞后问题。

多源点云融合技术正加速向全局信息融合与多阶段迭代优化的方向演进。在全局融合层面，通过统一时空基准与语义空间实现多源点云的异构兼容。Panagiotidis等人^[38]在森林监测中通过前融合实现无人机与地面LiDAR的联合标定，中间融合动态加权蓄积量特征，最终通过后融合的林冠密度决策反向优化采样密度，单次迭代周期内重建精度提升12.7%。Cheng等人^[41]进一步引入置信度驱动

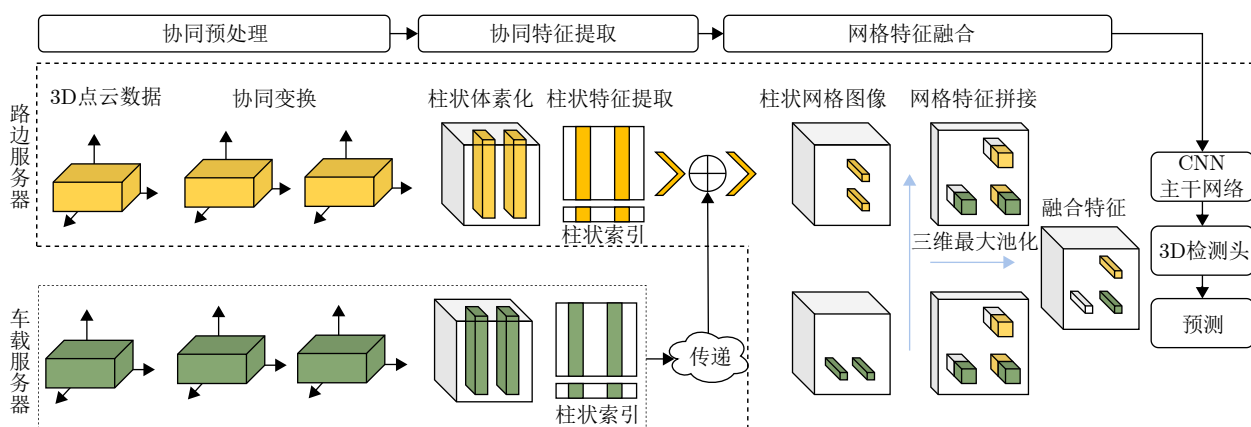


图7 PillarGrid Network^[43]

Fig. 7 PillarGrid Network^[43]

的全景LiDAR融合框架，筛选高置信度点云增强走廊重建细节，同时通过边缘计算加速将迭代延迟压缩至 $5\ \mu\text{s}$ 以内，推动技术向实时动态场景落地。当前，量子时空编码与神经符号系统的结合正成为新热点，通过超图结构建模多模态特征间的复杂关系，在统一嵌入空间中整合LiDAR几何结构、高光谱物质属性与RGB纹理特征，在文化遗产修复等领域实现高细节还原度，这些进展标志着“全域感知到智能演化”新阶段的开启。

3.2 多帧点云融合的常用方法

多帧点云融合是将同一传感器在不同时间获取的点云数据进行融合，通过利用时序信息提升三维感知的精度与鲁棒性，其方法发展经历了从基于运动补偿的早期框架端到端深度学习模型的演进，逐步实现动态场景的高效建模。图9以时间线形式展示了2018年—2024年多帧点云融合领域的研究发展，并标注了关键研究成果的时间分布与贡献者，体现该领域研究的持续演进与技术脉络。

早期研究侧重于通过简单时序建模整合多帧点云。例如，Luo等人^[54]在2018年提出了一种实时端到端框架，通过单一卷积网络联合处理3D检测、跟踪与运动预测，首次将多帧时序信息引入LiDAR

感知任务。2019年，Li等人^[55]提出基于多帧LiDAR点云的障碍物检测方法，通过运动补偿提升动态目标检测的稳定性，为工业机器人避障提供了早期实践基础。2020年，Huang等人^[56]提出基于长短期记忆网络(Long Short Term Memory, LSTM)的时序检测方法，利用循环神经网络捕捉点云序列中的动态特征，如图10所示，该方法通过整合历史帧与当前帧的点云数据，利用LSTM存储时序特征、图卷积建模物体空间关联，并结合自车运动变换对齐坐标系，实现跨帧特征融合；同年，Yin等人^[57]设计图神经网络与时空注意力机制，通过消息传递优化多帧点云的特征交互。这些方法解决了多帧点云“能否融合”的问题，奠定了时序建模的基础，但受限于模型复杂度，难以高效处理长序列点云。尽管如此，这些早期研究为后续高效时序建模方法的开发奠定了重要基础。

近期的多帧点云融合以高效架构与复杂建模为核心，向多模态、自适应融合延伸。2021年，Yang等人^[58]提出多帧注意力网络(3D-MAN)，首次将注意力机制(Transformer)引入点云序列建模，通过自注意力机制捕捉长距离时空依赖，如图11所示，该方法的核心在于记忆库存储历史帧的特征与检测信息，结合多视图对齐与聚合模块，对不同帧的点

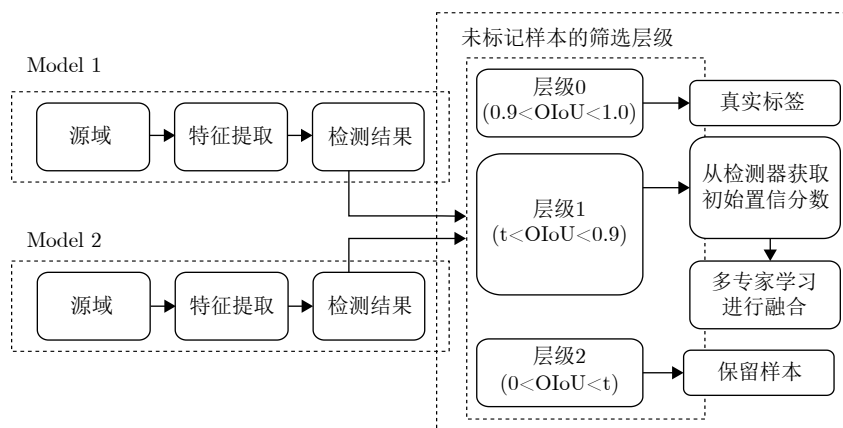


图8 多专家学习框架^[51]

Fig. 8 Multi-expert learning framework^[51]

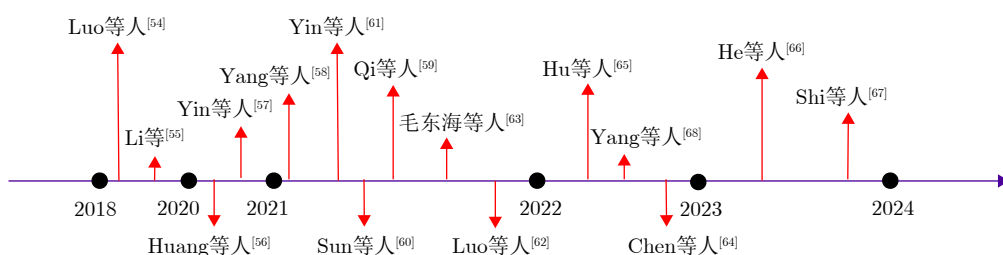
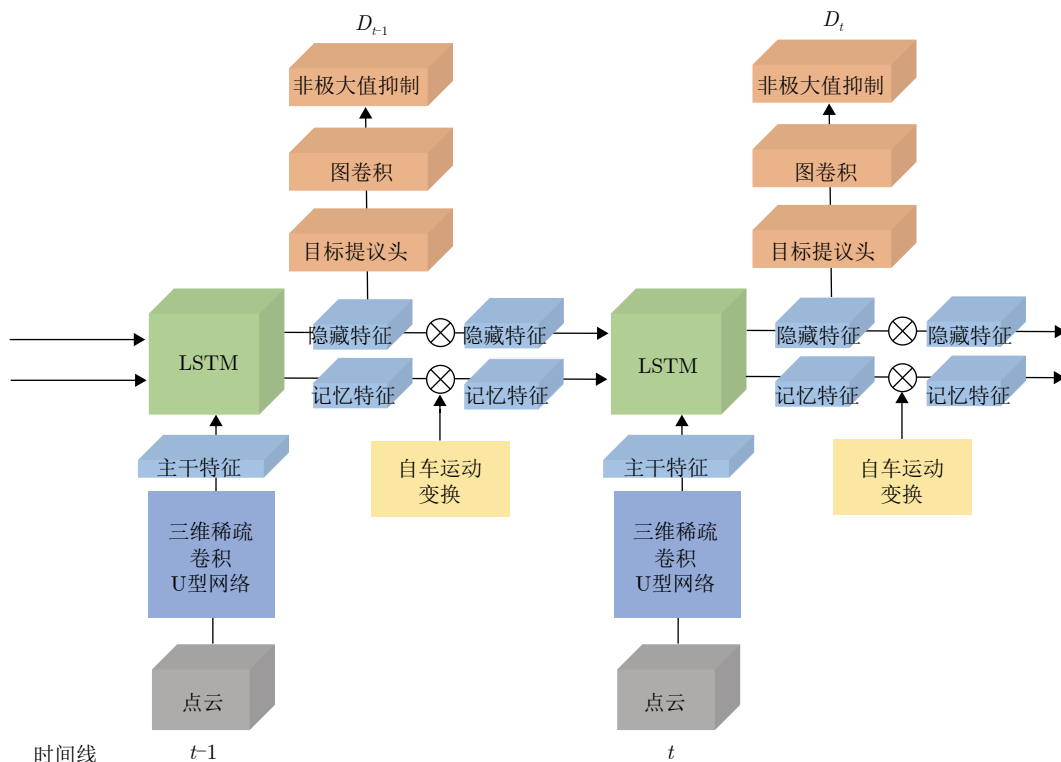
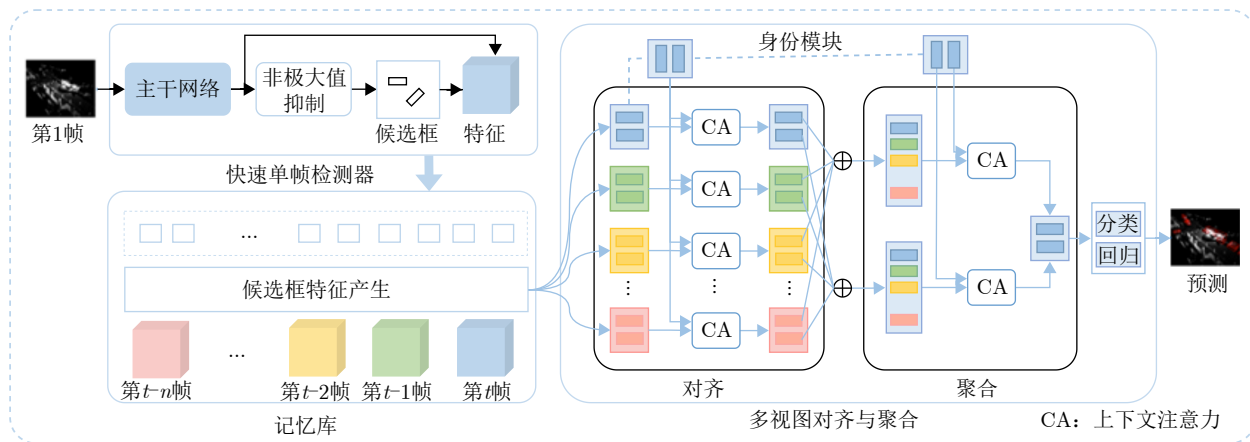


图9 多帧点云融合时间线

Fig. 9 Multi-frame point cloud fusion timeline

图 10 基于LSTM的时序检测方法^[56]Fig. 10 LSTM-based time series detection method^[56]图 11 3D-MAN结构框图^[58]Fig. 11 3D-MAN structure diagram^[58]

云进行坐标变换和特征融合,以解决时空不一致问题,在单帧目标检测与存储中,该方法采用快速单帧检测器,对当前帧点云生成目标检测框,并将其特征存入记忆库,存储多个历史帧的目标信息,然后利用交叉注意力机制,基于目标帧对历史帧特征进行检索和匹配,以保持目标身份的一致性,通过跨帧交叉注意力机制,融合各帧的目标特征,增强小目标、遮挡目标的表征能力,并减少单帧检测误差的累积,最终通过分类、回归输出预测结果,实现高效且稳定的多帧点云融合。

在3D-MAN的工作基础上,Qi等人^[59]提出了一

种离线3D检测框架,通过全局优化长时序点云数据显著提升检测精度。如图12所示,该框架首先利用改进的多帧检测器MVF++对单帧点云进行初步检测,引入时间编码与多帧点云聚合策略,将多帧点云补偿至当前坐标系以融合时空信息,针对单帧检测中因遮挡或稀疏性导致的漏检问题,作者进一步提出对象中心优化策略:通过多目标跟踪关联跨帧检测结果,提取完整的对象轨迹数据,并基于此进行精细化建模,对于静态物体,合并多帧点云生成全局优化框;动态物体则结合点云序列与检测框序列的时空特征进行滑动窗口预测。该方法在

Waymo数据集上实现了对单帧检测器的显著超越，验证了多帧全局融合在离线场景中的关键作用。

多帧点云融合技术逐步摆脱对独立跟踪模块的依赖，从分步优化向端到端联合建模快速演进，诸多学者为动态场景感知提供了更高效的诸多解决方案。Sun等人^[60]提出Range Sparse Net (RSN)，通过稀疏卷积降低多帧点云的计算开销；Yin等人^[61]提出基于中心点的检测与跟踪联合框架，简化多帧数据关联流程；Luo等人^[62]探索轻量化多目标跟踪方法，通过运动预测减少对复杂时序模型的依赖。此外，毛冬海等人^[63]提出LiDAR点云分块滤波方法，结合影像数据优化静态场景的融合效率，为地形测绘提供新思路。Chen等人^[64]在2022年提出MPPNet，通过“代理点”(Proxy Points)实现多帧特征交织，利用可学习模板关联不同帧的关键区域；Hu等人^[65]则挑战传统两阶段检测范式，提出单阶段端到端模型AFDetV2，直接融合多帧点云提升效率。2023年，He等人^[66]提出运动引导序列融合框架，将目标运动估计嵌入特征融合过程，通过运动补偿与自适应加权显著提升动态物体检测精度。Shi等人^[67]针对毫米波雷达提出自主多帧融合方法，通过时序滤波与点云累积增强低分辨率传感器的感知能力；Yang等人^[68]引入场景流(Scene Flow)优化时序点云融合，通过运动场建模增强动

态目标跟踪的鲁棒性，进一步扩展了运动补偿的理论边界。

多帧点云融合技术的发展经历了从基础时序配准到智能动态建模的演进，其核心目标从解决“能否融合”的基础问题，逐步转向“如何提升精度、效率与鲁棒性”的高阶挑战。当前技术趋势呈现两大方向：一是轻量化设计，通过注意力机制与稀疏卷积的融合，动态分配计算资源并降低冗余；二是端到端多任务联合建模，将检测、跟踪与预测整合为统一框架，实现任务间特征共享与协同优化。尤其在动态场景优化中，加权融合策略与不确定性建模成为关键突破点。未来研究可能集中在长时序建模领域，例如利用超长序列Transformer架构捕捉跨时段时空依赖，同时探索动态遮挡的鲁棒性增强方法，从而推动多帧点云融合技术在自动驾驶、工业机器人等领域的深入应用。

3.3 多视角点云融合的常用方法

多视角点云融合是指通过整合来自不同视角的点云数据，生成更完整、更精确的三维场景表示的技术。随着传感器技术和计算能力的提升，多视角点云融合方法逐渐从传统的几何优化向基于深度学习的数据驱动融合方法发展，并在精度、鲁棒性和实时性等方面取得了显著进展。图13以时间线形式

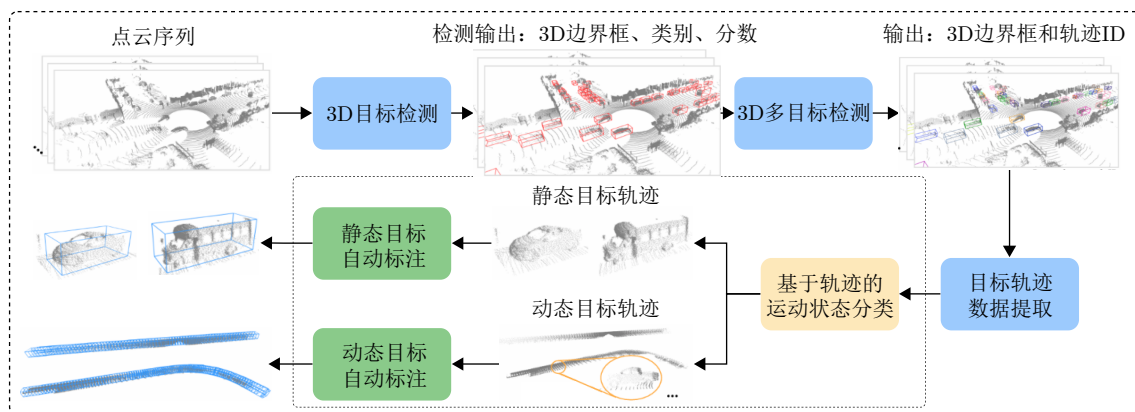


图 12 利用点云序列进行离线三维目标检测的框架^[59]

Fig. 12 Framework for offline 3D object detection using point cloud sequences^[59]

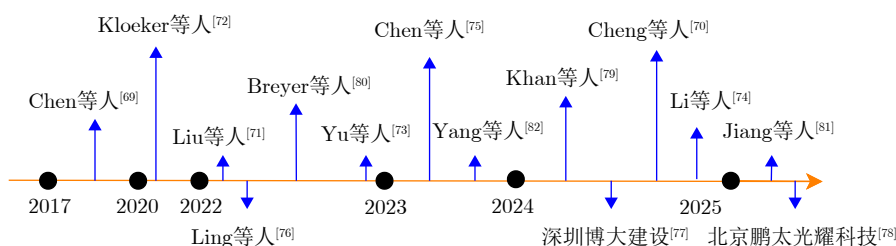


图 13 多视角点云融合时间线

Fig. 13 Multi-view point cloud fusion timeline

梳理了2017年—2025年多视角点云融合领域的研究进展,清晰地展示了该领域的技术演化路径与重要学术贡献。

多视角点云融合方法近年来在理论与应用层面均取得了显著进展。早期的研究主要依赖几何特征匹配与优化算法,例如基于点特征的标定方法通过提取关键点与描述符,实现了不同视角下点云数据的初步对齐。Chen等人^[69]提出了一种基于尺度的主成分分析-迭代最近点算法,以消除两个视点的不同尺度,通过分析不同球面半径上离散采样特征点的持久性,自动提取特征点以实现室外场景的精确重建,图14展示了该方法的主要流程,该方法首先利用尺度主成分分析,在不同球面半径上离散采样,提取稳定的特征点,以减少不同视角下的尺度影响。随后,通过迭代最近点优化刚性变换参数,使点云逐步对齐,并结合误差计算判断收敛,提高对齐精度。Cheng等人^[70]在户外场景三维重建中,结合多视角点云数据,利用几何一致性优化对齐全局精度,并在建筑测绘与城市场景建模领域展现了

良好的应用潜力。然而,上述传统方法在处理复杂遮挡或动态场景时仍面临计算效率低与鲁棒性不足的问题。

随着深度学习在三维视觉领域的范式革新,多视角点云融合已突破传统几何驱动框架,向数据驱动的智能融合体系演进。Liu等人^[71]提出了一种基于概率聚类的多视角点云融合框架,如图15所示。该框架主要包含3个核心阶段:概率估计、概率优化和快速概率聚类,旨在从多个视角的点云数据中自适应挖掘全局一致性的几何信息,并有效抑制噪声干扰,从而提升点云融合的鲁棒性和准确性,相较于传统的几何驱动方法,该方法通过数据驱动的概率优化,更具自适应性和鲁棒性,为机器人巡检等实际应用提供了更高精度的大规模场景重建方案。Kloeker等人^[72]进一步提出均值-方差最小化全局配准方法,结合统计优化策略和深度学习方法,在机器人巡检中实现了高精度的大规模场景重建。Yu等人^[73]提出首个点云预训练的Transformer框架,通过掩码建模学习多视角点云的全局上下文关系,

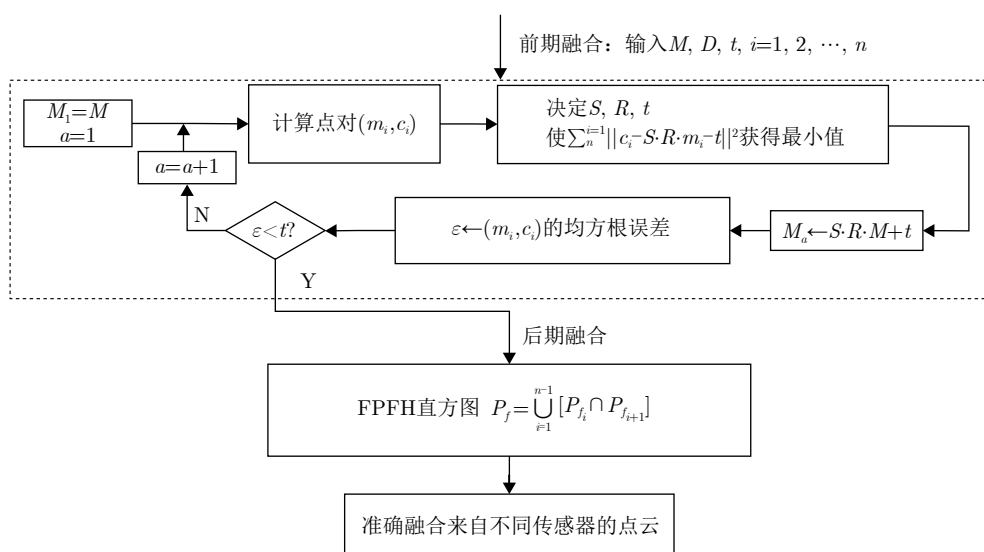


图 14 改进的尺度PCA-ICP算法的实现过程^[69]

Fig. 14 The implementation process of the improved scaled PCA-ICP algorithm^[69]

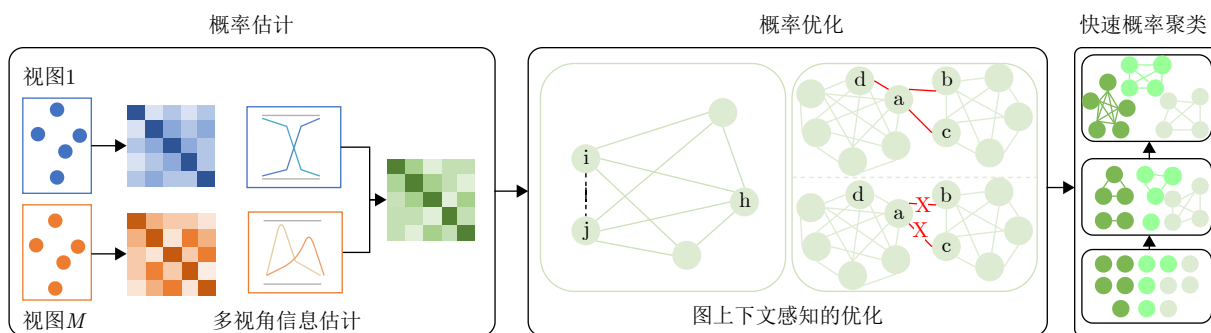


图 15 多视图聚类框架^[71]

Fig. 15 Multi-view clustering framework^[71]

在多视角点云融合任务中表现优异。Li等人^[74]提出的多视图特征聚合模块利用整流线性注意来合并跨视图的全局上下文信息，确保更连贯和语义准确的重建。Chen等人^[75]使用来自3D掩蔽点云的编码标记来生成原始点云和跨各种姿势的多视图深度图像。近期也有研究尝试引入图神经网络，通过图结构来刻画各视角间的关系，并实现跨视角信息的高效传播和融合，这种方法在捕捉局部结构变化及细粒度语义信息上具有优势^[76]。

工业界相关专利技术也体现了从几何特征匹配到数据驱动智能融合的趋势，深圳市博大建设集团的多视角激光点云拼接技术^[77]，如图16所示，该技术通过特征提取算法生成多视角点云的关键特征描述符，并利用最近邻搜索与RANSAC算法实现高效特征匹配，完成初步对齐，针对误差累积问题，引入动态闭环检测机制，通过构建点云配

准边识别重叠区域，结合位姿参数(平移与旋转)的动态优化，逐步减少全局拼接误差，此外，采用ICP变体算法对局部匹配点进行加权优化，抑制噪声干扰，进一步提升融合精度，该方法通过多阶段的智能优化策略，显著降低了因视角差异导致的配准偏差，为复杂场景下的高精度三维重建提供了工业级解决方案。

北京鹏太光耀科技有限公司的全向点云融合方法^[78]则利用多目结构光与特征融合策略，提升了复杂场景下的重建效率，如图17所示，该方法采用环绕式多相机-投影仪阵列，利用实心圆点标定板结合张正友标定法完成设备参数标定，确保多视角数据的空间一致性，通过投射横状条纹与格雷码组合图案，结合格雷码-相移法解码相位信息，生成多视角初始点云，针对噪声问题，采用统计分析滤波去除离群点，并基于高斯滤波与SHOT特征提取生

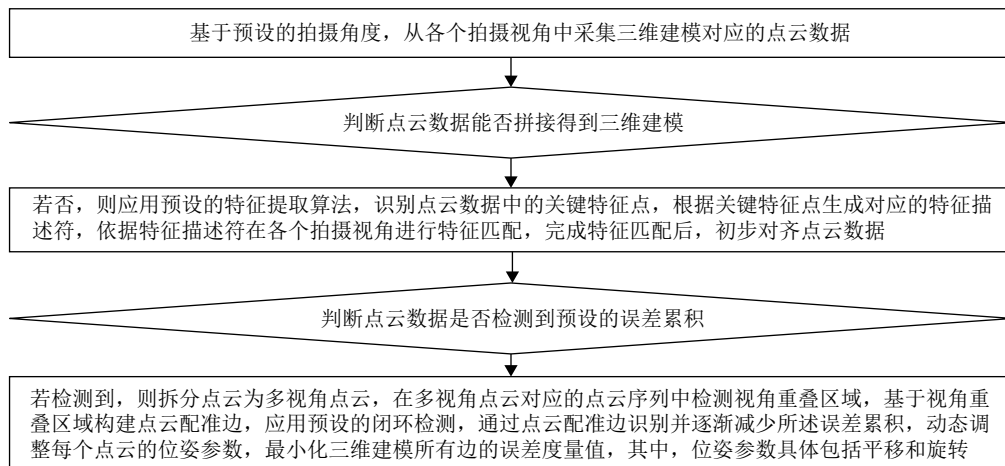


图 16 多视角激光点云拼接技术流程图^[77]

Fig. 16 Multi-view laser point cloud stitching technology flow chart^[77]

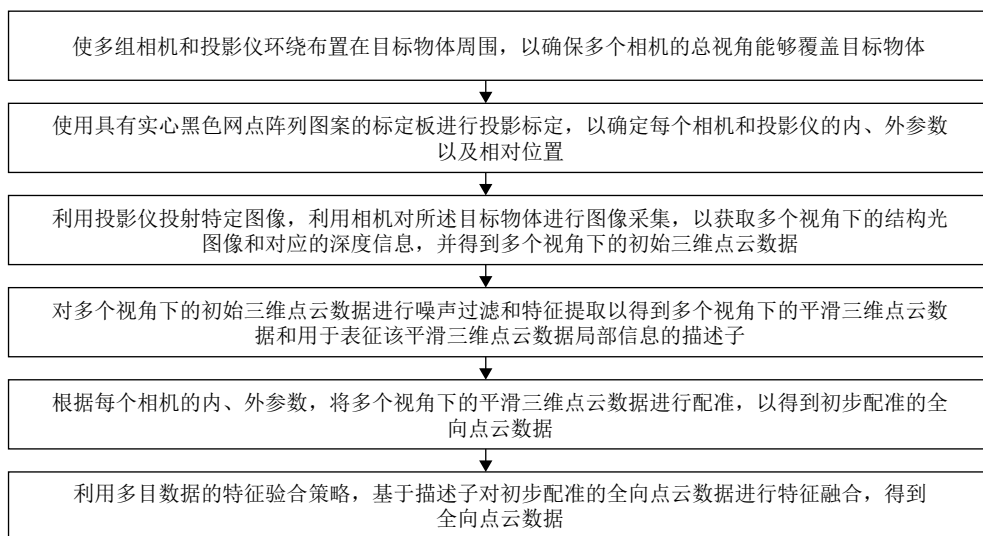


图 17 全向点云融合方法流程图^[78]

Fig. 17 Flowchart of the omnidirectional point cloud fusion method^[78]

成局部描述子, 增强特征区分性, 通过多目特征融合策略匹配高区分性特征点, 结合ICP算法构建全局优化函数, 实现点云合并。该方法通过结构光编码与智能特征融合的双重优化, 显著提升了复杂场景下的重建效率与精度, 为工业级全向三维建模提供了可扩展的解决方案。

面向自动驾驶、智能制造等领域的严苛需求, 研究重心正转向实时闭环优化与跨模态协同感知。Kloeker等人^[72]针对多LiDAR的实时融合问题, 提出无需先验位姿的在线标定方法, 为自动驾驶与智能交通提供了技术支持。Khan等人^[79]在车路协同场景中设计了路侧点云融合框架, 通过多传感器协同增强环境感知能力。此外, Breyer等人^[80]提出了一种基于闭环优化的最优视角规划方法, 通过动态选择最优采集视角, 减少了冗余数据并提升了融合效率。在大规模场景重建中, Jiang等人^[81]结合LiDAR与高斯溅射技术, 实现了高精度、低计算开销的三维建模, 为电力巡检与文化遗产保护等场景提供了新思路。这些进展不仅推动了算法创新, 也通过工业界的专利落地验证了其实际价值。

多视角点云融合目前已从传统几何匹配转向智能化与数据驱动相结合, 实时性与多模态融合正逐渐成为新的焦点。工业界技术进展不断支撑自动驾驶、电力巡检等实际应用场景的技术要求。多视角点云融合未来的发展需突破动态场景自适应建模、轻量化实时算法及跨模态协同优化的瓶颈, 强化复杂遮挡处理能力与工业级闭环系统, 赋能智能驾驶、智慧城市与智能制造的高效三维感知需求。

4 点云融合的应用与未来

4.1 点云融合的应用

点云融合是一项广泛应用于多个领域的重要技术, 可增强3D建模和环境理解。通过集成多源、多帧或多视角的数据, 点云融合提高了数据的分辨率和完整性。这种融合方式使得环境和对象能够得到更加全面的表示, 这对于需要高精度的任务至关重要。本节将详细探讨点云融合在一些重要领域中的应用, 重点讨论在智能驾驶、机器人技术、增强现实(AR)和虚拟现实(VR)以及建筑和工业检测的应用。

在智能驾驶领域, 点云融合在协同感知、实时目标检测与跟踪、数据集驱动研究中发挥着关键作用^[82-84], 诸如通过V2X (Vehicle to Everything)通信实现车-路-云协同感知, 融合多视角LiDAR点云构建全局高精度地图; 使用多帧点云融合预测运动轨迹, 实现实时目标检测与跟踪等。在智能驾驶领

域主流数据集如nuScenes, KITTI等为算法验证提供基准, 结合LiDAR提供的精确深度信息与摄像头捕捉的丰富视觉细节(如颜色和纹理), 从而更准确地构建周围环境的3D表示, 提高在复杂环境中的鲁棒性, 例如在强光、雾霾等恶劣天气条件下, 增强自动驾驶系统的安全性^[85,86]。

在机器人领域, 特别是SLAM系统中, 点云融合得到了广泛应用。通过融合来自不同传感器的3D点云数据, SLAM算法能够提升环境感知的精度和鲁棒性^[87]。这种融合使得机器人能够构建更加精确的3D地图, 从而支持其在动态和未知环境中的自主导航。例如, 在无人仓储机器人中, 点云融合帮助机器人实时检测并避开障碍物, 同时精确规划路径, 确保高效作业。

在AR和VR中, 点云融合对于创建精细且互动性强的虚拟环境至关重要。融合不同的点云数据有助于以更高的精度对虚拟空间进行建模, 使虚拟世界更加逼真。在AR中, 点云融合通过确保虚拟对象与现实世界无缝对齐, 使系统能够准确地将数字内容叠加到现实世界中。在VR中, 点云融合支持创建精细的空间3D模型, 从而增强虚拟体验的沉浸感^[88]。例如, 利用激光扫描和RGB-D数据, VR系统能够动态调整虚拟环境中的物体位置与用户的互动, 提升用户体验。

在建筑和工业检测中, 点云融合可以对建筑物、基础设施和机械进行精确的3D建模^[89,90]。通过集成来自无人机(UAV)、摄影测量和地面激光扫描(TLS)的数据, 建筑师和工程师可以创建高精度的结构模型, 这些模型有助于检测结构问题、规划翻新并确保结构符合设计规范。例如, 在历史建筑保护中, 点云融合技术被用于对复杂结构的数字化重建, 确保每个细节的精确还原。

4.2 点云融合的挑战与展望

点云融合是自动驾驶、机器人和3D建模等领域的关键技术, 通过合并不同传感器、不同时间和不同视角的数据来提供对环境的全面理解。然而, 由于不同传感器的功能性差异、生成点云数据的结构性差异以及实时应用需求方面的差异, 点云融合过程面临着诸多挑战。

4.2.1 数据稀疏性和密度差异

点云融合面临的一个主要挑战是数据稀疏性和密度差异^[91]。不同传感器生成的点云在密度上存在较大差异, 例如, LiDAR生成的点云包含丰富的几何结构, 但较为稀疏, 而摄影测量或RGB-D相机生成的点云则通常具有较高的密度, 提供更多的

纹理细节。由于这些密度差异,点云融合可能导致生成的点云不均匀或不完整,尤其是当过度依赖高密度点云时,稀疏数据可能被忽略,影响最终融合效果。

为了克服这一挑战,未来的点云融合研究应致力于开发更智能的插值算法和自适应融合技术。自适应点云融合可以根据不同来源的点云密度自动调整权重,确保生成的点云既包含足够的几何细节,又不会因为稀疏数据的存在而导致误差^[88,89]。此外,多分辨率表示技术通过在不同尺度上处理点云数据,有效捕捉细节和全局结构,并结合稀疏点云补全算法,帮助填补数据中的稀疏区域,从而确保点云数据在空间上的完整性和一致性^[90]。

4.2.2 数据对齐和配准

点云融合中的核心问题是点云数据的对齐和配准。即使是极小的位置偏差也可能显著降低融合效果,导致几何结构的扭曲或不准确。由于传感器噪声、视角差异及环境变化等因素,点云的对齐与配准尤为复杂,特别是在异构传感器数据之间进行配准时更具挑战性^[91]。

在点云融合中,“对齐”和“配准”通常被用来描述不同的过程。配准是指将多个点云数据集在同一坐标系下对齐的过程,它通常依赖于算法将来自不同传感器或不同视角的数据通过几何变换(如旋转、平移)进行匹配。对齐则是配准的一个重要步骤,通常指的是将点云数据通过特征匹配或迭代优化等方法进行精确对齐,以消除因视角差异、传感器精度和环境变化带来的偏差。点云融合不仅仅是对齐的过程,更是在配准后的基础上,整合来自不同数据源的点云信息,生成更完整、精确的3D表示。因此,尽管对齐和配准是点云融合的前提步骤,但点云融合的目标是通过智能化算法(如深度学习)进一步提高数据的精度和鲁棒性。

为提高点云融合的精度和鲁棒性,未来可以借助深度学习技术开发更智能化的点云融合方法,通过特征提取和几何优化算法来提高融合精度^[92]。同时,改进的ICP算法结合特征匹配和几何优化等技术,能够显著提高融合鲁棒性。在多模态融合中,进一步研究多源数据的时空对齐方法,也有助于提高点云融合的整体准确性^[93]。

4.2.3 计算复杂性与实时性要求

在智能驾驶、无人机导航和实时3D地图生成等应用中,点云融合不仅要求高精度,还要求高实时性^[94]。然而,点云融合过程通常涉及海量数据处理,尤其在处理高分辨率点云时,计算复杂性成为

主要的瓶颈。在这些实时应用场景中,如何在融合多个点云的同时满足时间限制是一项关键挑战。

未来的研究可在提高计算效率方面进一步探索。并行计算、硬件加速(如GPU和FPGA)以及优化的算法结构可以显著减少处理时间^[95,96]。此外,采用分块处理、多线程计算以及基于云计算的分布式系统,将有助于更好地满足实时应用的需求。在保证高速处理的同时,保持融合的几何准确性仍然是需要重点解决的问题^[97]。

4.2.4 噪声与遮挡

点云数据常因传感器误差或环境干扰而受到噪声影响,噪声的存在会对点云融合结果产生负面影响^[98]。此外,由于障碍物或现实场景几何结构的复杂性,数据采集过程中会产生遮挡效应,导致某些区域的点云数据缺失。这些问题给点云融合带来了更高的复杂性,特别是在构建精细化3D模型时,如何处理噪声和遮挡数据仍然是一项重大挑战。

解决这些问题的关键在于开发更先进的去噪和补全算法。未来,基于机器学习的噪声过滤和数据补全方法将成为主流,这些方法能够有效提升点云数据的质量和完整性^[99]。通过深度学习技术,模型能够自适应地识别并修复数据中的缺失部分,同时减少噪声的干扰,从而显著提高点云融合的精度和稳定性^[100]。这将为复杂场景下的高精度三维重建、智能感知和自动化应用提供更强有力的支持。此外,针对遮挡问题的点云融合研究,将有助于增强点云数据的完整性,确保最终生成的点云具备更高的精度^[101]。

5 结论

本文详细探讨了点云融合技术的主要方法及其应用,涵盖了多源、多帧和多视角的点云融合方法。每种方法在不同的应用场景中展现了独特的优势,但同时也面临着计算复杂度、精度、鲁棒性等方面的挑战。通过梳理最新的研究成果和分析各类方法的特点,本文为点云融合领域的进一步发展提供了系统性的参考。虽然现有方法已取得显著进展,但点云数据的稀疏性、不规则性和多源数据融合的复杂性依然是亟待解决的问题。随着计算机视觉、机器学习及传感器技术的持续进步,未来的研究有望在提高融合精度、处理效率和跨领域应用的拓展方面取得更有突破性的成果。

利益冲突 所有作者均声明不存在利益冲突

Conflict of Interests The authors declare that there is no conflict of interests

参考文献

- [1] ABBASI R, BASHIR A K, ALYAMANI H J, *et al.* Lidar point cloud compression, processing and learning for autonomous driving[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(1): 962–979. doi: [10.1109/TITS.2022.3167957](https://doi.org/10.1109/TITS.2022.3167957).
- [2] QURESHI A H, ALALOUL W S, MURTIYOSO A, *et al.* Smart rebar progress monitoring using 3D point cloud model[J]. *Expert Systems with Applications*, 2024, 249: 123562. doi: [10.1016/j.eswa.2024.123562](https://doi.org/10.1016/j.eswa.2024.123562).
- [3] PETKOV P. Object detection and recognition with PointPillars in LiDAR point clouds-comparisons[R]. NSC-614-5849, 2024.
- [4] 龚靖渝, 楼雨京, 柳奉奇, 等. 三维场景点云理解与重建技术[J]. 中国图象图形学报, 2023, 28(6): 1741–1766. doi: [10.11834/jig.230004](https://doi.org/10.11834/jig.230004).
GONG Jingyu, LOU Yujing, LIU Fengqi, *et al.* Scene point cloud understanding and reconstruction technologies in 3D space[J]. *Journal of Image and Graphics*, 2023, 28(6): 1741–1766. doi: [10.11834/jig.230004](https://doi.org/10.11834/jig.230004).
- [5] ZHANG Bo, WANG Haosen, YOU Suilian, *et al.* A small-size 3D object detection network for analyzing the sparsity of raw LiDAR point cloud[J]. *Journal of Russian Laser Research*, 2023, 44(6): 646–655. doi: [10.1007/s10946-023-10173-3](https://doi.org/10.1007/s10946-023-10173-3).
- [6] LI Shaorui, ZHU Zhenchang, DENG Weitang, *et al.* Estimation of aboveground biomass of different vegetation types in mangrove forests based on UAV remote sensing[J]. *Sustainable Horizons*, 2024, 11: 100100. doi: [10.1016/j.horiz.2024.100100](https://doi.org/10.1016/j.horiz.2024.100100).
- [7] 余杭. 基于激光雷达的3D目标检测研究综述[J]. 汽车文摘, 2024(2): 18–27. doi: [10.19822/j.cnki.1671-6329.20230082](https://doi.org/10.19822/j.cnki.1671-6329.20230082).
YU Hang. A review on LiDAR-based 3D target detection research[J]. *Automotive Digest*, 2024(2): 18–27. doi: [10.19822/j.cnki.1671-6329.20230082](https://doi.org/10.19822/j.cnki.1671-6329.20230082).
- [8] 郭迟, 刘阳, 罗亚荣, 等. 图像语义信息在视觉SLAM中的应用研究进展[J]. 测绘学报, 2024, 53(6): 1057–1076. doi: [10.11947/j.AGCS.2024.20230259](https://doi.org/10.11947/j.AGCS.2024.20230259).
GUO Chi, LIU Yang, LUO Yarong, *et al.* Research progress in the application of image semantic information in visual SLAM[J]. *Acta Geodaetica et Cartographica Sinica*, 2024, 53(6): 1057–1076. doi: [10.11947/j.AGCS.2024.20230259](https://doi.org/10.11947/j.AGCS.2024.20230259).
- [9] ZHANG L, VAN OOSTEROM P, and LIU H. Visualization of point cloud models in mobile augmented reality using continuous level of detail method[C]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, London, UK, 2020: 167–170. doi: [10.5194/isprs-archives-XLIV-4-W1-2020-167-2020](https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-167-2020).
- [10] POLIYAPRAM V, WANG Weimin, and NAKAMURA R. A point-wise LiDAR and image multimodal fusion network (PMNet) for aerial point cloud 3D semantic segmentation[J]. *Remote Sensing*, 2019, 11(24): 2961. doi: [10.3390/rs11242961](https://doi.org/10.3390/rs11242961).
- [11] ZENG Tianjiao, ZHANG Wensi, ZHAN Xu, *et al.* A novel multimodal fusion framework based on point cloud registration for near-field 3D SAR perception[J]. *Remote Sensing*, 2024, 16(6): 952. doi: [10.3390/rs16060952](https://doi.org/10.3390/rs16060952).
- [12] CRISAN A, PEPE M, COSTANTINO D, *et al.* From 3D point cloud to an intelligent model set for cultural heritage conservation[J]. *Heritage*, 2024, 7(3): 1419–1437. doi: [10.3390/heritage7030068](https://doi.org/10.3390/heritage7030068).
- [13] KIVILCIM C Ö and DURAN Z. A semi-automated point cloud processing methodology for 3D cultural heritage documentation[C]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Prague, Czech Republic, 2016: 293–296. doi: [10.5194/isprs-archives-XLI-B5-293-2016](https://doi.org/10.5194/isprs-archives-XLI-B5-293-2016).
- [14] WANG Yong, ZHOU Pengbo, GENG Guohua, *et al.* Enhancing point cloud registration with transformer: Cultural heritage protection of the Terracotta Warriors[J]. *Heritage Science*, 2024, 12(1): 314. doi: [10.1186/s40494-024-01425-9](https://doi.org/10.1186/s40494-024-01425-9).
- [15] YANG Su, HOU Miaole, and LI Songnian. Three-dimensional point cloud semantic segmentation for cultural heritage: A comprehensive review[J]. *Remote Sensing*, 2023, 15(3): 548. doi: [10.3390/rs15030548](https://doi.org/10.3390/rs15030548).
- [16] 李佳益, 马智亮, 陈礼杰, 等. 面向施工机器人定位的多模态数据融合方法研究综述[J]. 计算机工程与应用, 2024, 60(15): 11–23. doi: [10.3778/j.issn.1002-8331.2402-0008](https://doi.org/10.3778/j.issn.1002-8331.2402-0008).
LI Jiayi, MA Zhiliang, CHEN Lijie, *et al.* Comprehensive review of multimodal data fusion methods for construction robot localization[J]. *Computer Engineering and Applications*, 2024, 60(15): 11–23. doi: [10.3778/j.issn.1002-8331.2402-0008](https://doi.org/10.3778/j.issn.1002-8331.2402-0008).
- [17] 周燕, 李文俊, 党兆龙, 等. 深度学习的三维模型识别研究综述[J]. 计算机科学与探索, 2024, 18(4): 916–929. doi: [10.3778/j.issn.1673-9418.2309010](https://doi.org/10.3778/j.issn.1673-9418.2309010).
ZHOU Yan, LI Wenjun, DANG Zhaolong, *et al.* Survey of 3D model recognition based on deep learning[J]. *Journal of Frontiers of Computer Science and Technology*, 2024, 18(4): 916–929. doi: [10.3778/j.issn.1673-9418.2309010](https://doi.org/10.3778/j.issn.1673-9418.2309010).
- [18] WANG Zhengren. 3D representation methods: A survey[J]. *arXiv preprint arXiv*, 2410.06475, 2024.
- [19] ZHOU Yin, SUN Pei, ZHANG Yu, *et al.* End-to-end multi-view fusion for 3D object detection in LiDAR point clouds[C]. The 3rd Annual Conference on Robot Learning,

- Osaka, Japan, 2019: 923–932.
- [20] 陈慧娴, 吴一全, 张耀. 基于深度学习的三维点云分析方法研究进展[J]. 仪器仪表学报, 2023, 44(11): 130–158. doi: [10.19650/j.cnki.cjsi.J2311134](https://doi.org/10.19650/j.cnki.cjsi.J2311134).
- CHEN Huixian, WU Yiquan, and ZHANG Yao. Research progress of 3D point cloud analysis methods based on deep learning[J]. *Chinese Journal of Scientific Instrument*, 2023, 44(11): 130–158. doi: [10.19650/j.cnki.cjsi.J2311134](https://doi.org/10.19650/j.cnki.cjsi.J2311134).
- [21] LUO Haojun and WEN C Y. A low-cost relative positioning method for UAV/UGV coordinated heterogeneous system based on visual-lidar fusion[J]. *Aerospace*, 2023, 10(11): 924. doi: [10.3390/aerospace10110924](https://doi.org/10.3390/aerospace10110924).
- [22] LI Yangyan, BU Rui, SUN Mingchao, *et al.* PointCNN: Convolution on X-transformed points[C]. The 32nd International Conference on Neural Information Processing Systems, Montréal Canada, 2018: 828–838.
- [23] GEIGER A, LENZ P, STILLER C, *et al.* Vision meets robotics: The KITTI dataset[J]. *The International Journal of Robotics Research*, 2013, 32(11): 1231–1237. doi: [10.1177/0278364913491297](https://doi.org/10.1177/0278364913491297).
- [24] CAESAR H, BANKITI V, LANG A H, *et al.* nuScenes: A multimodal dataset for autonomous driving[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 11621–11631. DOI: [10.1109/CVPR42600.2020.01164](https://doi.org/10.1109/CVPR42600.2020.01164).
- [25] MADDERN W, PASCOE G, LINEGAR C, *et al.* 1 year, 1000 km: The oxford robotcar dataset[J]. *The International Journal of Robotics Research*, 2017, 36(1): 3–15. doi: [10.1177/0278364916679498](https://doi.org/10.1177/0278364916679498).
- [26] CURLESS B and LEVOY M. A volumetric method for building complex models from range images[C]. The 23rd Annual Conference on Computer Graphics and Interactive Techniques, New Orleans, USA, 1996: 303–312. doi: [10.1145/237170.237269](https://doi.org/10.1145/237170.237269).
- [27] MIAN A S, BENNAMOUN M, and OWENS R A. A novel representation and feature matching algorithm for automatic pairwise registration of range images[J]. *International Journal of Computer Vision*, 2006, 66(1): 19–40. doi: [10.1007/s11263-005-3221-0](https://doi.org/10.1007/s11263-005-3221-0).
- [28] SUN Pei, KRETZSCHMAR H, DOTIWALLA X, *et al.* Scalability in perception for autonomous driving: Waymo open dataset[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 2446–2454. doi: [10.1109/CVPR42600.2020.00252](https://doi.org/10.1109/CVPR42600.2020.00252).
- [29] GEYER J, KASSAHUN Y, MAHMUDI M, *et al.* A2D2: Audi autonomous driving dataset[J]. *arXiv preprint arXiv*, 2004.06320, 2020.
- [30] HUANG Xinyu, CHENG Xinjing, GENG Qichuan, *et al.* The apolloscape dataset for autonomous driving[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, USA, 2018: 954–960. doi: [10.1109/CVPRW.2018.00141](https://doi.org/10.1109/CVPRW.2018.00141).
- [31] CHANG Mingfang, LAMBERT J, SANGKLOY P, *et al.* Argoverse: 3D tracking and forecasting with rich maps[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 8748–8757. doi: [10.1109/CVPR.2019.00895](https://doi.org/10.1109/CVPR.2019.00895).
- [32] STURM J, BURGARD W, and CREMERS D. Evaluating egomotion and structure-from-motion approaches using the TUM RGB-D benchmark[C]. The Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RJS International Conference on Intelligent Robot Systems, Vilamoura, Algarve, 2012: 6.
- [33] CHOI Y, KIM N, HWANG S, *et al.* KAIST multi-spectral day/night data set for autonomous and assisted driving[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(3): 934–948. doi: [10.1109/TITS.2018.2791533](https://doi.org/10.1109/TITS.2018.2791533).
- [34] CHOI S, ZHOU Qianyi, and KOLTUN V. Robust reconstruction of indoor scenes[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 5556–5565. doi: [10.1109/CVPR.2015.7299195](https://doi.org/10.1109/CVPR.2015.7299195).
- [35] ZENG A, SONG S, NIEßNER M, *et al.* 3DMatch: Learning local geometric descriptors from RGB-D reconstructions[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 1802–1811. doi: [10.1109/CVPR.2017.29](https://doi.org/10.1109/CVPR.2017.29).
- [36] ABDELAZEEM M, ELAMIN A, AFIFI A, *et al.* Multi-sensor point cloud data fusion for precise 3D mapping[J]. *The Egyptian Journal of Remote Sensing and Space Science*, 2021, 24(3): 835–844. doi: [10.1016/j.ejrs.2021.06.002](https://doi.org/10.1016/j.ejrs.2021.06.002).
- [37] CÓRDOVA-ESPARZA D M, TERVEN J R, JIMÉNEZ-HERNÁNDEZ H, *et al.* A multiple camera calibration and point cloud fusion tool for Kinect V2[J]. *Science of Computer Programming*, 2017, 143: 1–8. doi: [10.1016/j.scico.2016.11.004](https://doi.org/10.1016/j.scico.2016.11.004).
- [38] PANAGIOTIDIS D, ABDOLLAHNEJAD A, and SLAVÍK M. 3D point cloud fusion from UAV and TLS to assess temperate managed forest structures[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 112: 102917. doi: [10.1016/j.jag.2022.102917](https://doi.org/10.1016/j.jag.2022.102917).
- [39] YANG Weijun, LIU Yang, HE Huagui, *et al.* Airborne LiDAR and photogrammetric point cloud fusion for extraction of urban tree metrics according to street network segmentation[J]. *IEEE Access*, 2021, 9: 97834–97842. doi: [10.1109/ACCESS.2021.3094307](https://doi.org/10.1109/ACCESS.2021.3094307).
- [40] PARVAZ S, TEFERLE F, and NURUNNABI A. Airborne cross-source point clouds fusion by slice-to-slice

- adjustment[C]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Athens, Greece, 2024: 161–168. doi: [10.5194/isprs-annals-X-4-W4-2024-161-2024](https://doi.org/10.5194/isprs-annals-X-4-W4-2024-161-2024).
- [41] CHENG Hongtai and HAN Jiayi. Toward precise dense 3D reconstruction of indoor hallway: A confidence-based panoramic LiDAR point cloud fusion approach[J]. *Industrial Robot*, 2025, 52(1): 116–125. doi: [10.1108/IR-03-2024-0132](https://doi.org/10.1108/IR-03-2024-0132).
- [42] LI Shiming, GE Xuming, HU Han, *et al.* Laplacian fusion approach of multi-source point clouds for detail enhancement[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 171: 385–396. doi: [10.1016/j.isprsjprs.2020.11.021](https://doi.org/10.1016/j.isprsjprs.2020.11.021).
- [43] BAI Zhengwei, WU Guoyuan, BARTH M J, *et al.* PillarGrid: Deep learning-based cooperative perception for 3D object detection from onboard-roadside LiDAR[C]. 2022 IEEE 25th International Conference on Intelligent Transportation Systems, Macau, China, 2022: 1743–1749. doi: [10.1109/ITSC55140.2022.9921947](https://doi.org/10.1109/ITSC55140.2022.9921947).
- [44] ZHANG Pengcheng, HE Huagui, WANG Yun, *et al.* 3D urban buildings extraction based on airborne LiDAR and photogrammetric point cloud fusion according to U-Net deep learning model segmentation[J]. *IEEE Access*, 2022, 10: 20889–20897. doi: [10.1109/ACCESS.2022.3152744](https://doi.org/10.1109/ACCESS.2022.3152744).
- [45] GUO Lijie, WU Yanjie, DENG Lei, *et al.* A feature-level point cloud fusion method for timber volume of forest stands estimation[J]. *Remote Sensing*, 2023, 15(12): 2995. doi: [10.3390/rs15122995](https://doi.org/10.3390/rs15122995).
- [46] QIAO Donghao and ZULKERNINE F. Adaptive feature fusion for cooperative perception using LiDAR point clouds[C]. 2023 IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, USA, 2023: 1186–1195. doi: [10.1109/WACV56688.2023.00124](https://doi.org/10.1109/WACV56688.2023.00124).
- [47] GRIFONI E, VANNINI E, LUNGHI I, *et al.* 3D multi-modal point clouds data fusion for metrological analysis and restoration assessment of a panel painting[J]. *Journal of Cultural Heritage*, 2024, 66: 356–366. doi: [10.1016/j.culher.2023.12.007](https://doi.org/10.1016/j.culher.2023.12.007).
- [48] CHEN Qi, TANG Sihai, YANG Qing, *et al.* Cooper: Cooperative perception for connected autonomous vehicles based on 3D point clouds[C]. 2019 IEEE 39th International Conference on Distributed Computing Systems, Dallas, USA, 2019: 514–524. doi: [10.1109/ICDCS.2019.00058](https://doi.org/10.1109/ICDCS.2019.00058).
- [49] HURL B, COHEN R, CZARNECKI K, *et al.* TruPercept: Trust modelling for autonomous vehicle cooperative perception from synthetic data[C]. 2020 IEEE Intelligent Vehicles Symposium, Las Vegas, USA, 2020: 341–347. doi: [10.1109/IV47402.2020.9304695](https://doi.org/10.1109/IV47402.2020.9304695).
- [50] LI Jie, ZHUANG Yiqi, PENG Qi, *et al.* Pose estimation of non-cooperative space targets based on cross-source point cloud fusion[J]. *Remote Sensing*, 2021, 13(21): 4239. doi: [10.3390/rs13214239](https://doi.org/10.3390/rs13214239).
- [51] TANG Zhiri, HU Ruihan, CHEN Yanhua, *et al.* Multi-expert learning for fusion of pedestrian detection bounding box[J]. *Knowledge-Based Systems*, 2022, 241: 108254. doi: [10.1016/j.knosys.2022.108254](https://doi.org/10.1016/j.knosys.2022.108254).
- [52] YU Haibao, LUO Yizhen, SHU Mao, *et al.* DAIR-V2X: A large-scale dataset for vehicle-infrastructure cooperative 3D object detection[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, 2022: 21361–21370. doi: [10.1109/CVPR52688.2022.02067](https://doi.org/10.1109/CVPR52688.2022.02067).
- [53] RIZALDY A, AFIFI A J, GHAMISI P, *et al.* Dimensional dilemma: Navigating the fusion of hyperspectral and LiDAR point cloud data for optimal precision-2D vs. 3D[C]. 2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 2024: 7729–7733. doi: [10.1109/IGARSS53475.2024.10641140](https://doi.org/10.1109/IGARSS53475.2024.10641140).
- [54] LUO Wenjie, YANG Bin, and URTASUN R. Fast and furious: Real time end-to-end 3D detection, tracking and motion forecasting with a single convolutional net[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 3569–3577. doi: [10.1109/CVPR.2018.00376](https://doi.org/10.1109/CVPR.2018.00376).
- [55] LI Jing, LI Rui, WANG Junzheng, *et al.* Obstacle information detection method based on multiframe three-dimensional lidar point cloud fusion[J]. *Optical Engineering*, 2019, 58(11): 116102. doi: [10.1117/1.OE.58.11.116102](https://doi.org/10.1117/1.OE.58.11.116102).
- [56] HUANG Rui, ZHANG Wanyue, KUNDU A, *et al.* An LSTM approach to temporal 3D object detection in LiDAR point clouds[C]. The 16th European Conference on Computer Vision, Glasgow, UK, 2020: 266–282. doi: [10.1007/978-3-030-58523-5_16](https://doi.org/10.1007/978-3-030-58523-5_16).
- [57] YIN Junbo, SHEN Jianbing, GUAN Chenye, *et al.* LiDAR-based online 3D video object detection with graph-based message passing and spatiotemporal transformer attention[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 11495–11504. doi: [10.1109/CVPR42600.2020.01151](https://doi.org/10.1109/CVPR42600.2020.01151).
- [58] YANG Zetong, ZHOU Yin, CHEN Zhifeng, *et al.* 3D-MAN: 3D multi-frame attention network for object detection[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 1863–1872. doi: [10.1109/CVPR46437.2021.00190](https://doi.org/10.1109/CVPR46437.2021.00190).
- [59] QI C R, ZHOU Yin, NAJIBI M, *et al.* Offboard 3D object detection from point cloud sequences[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition,

- Nashville, USA, 2021: 6134–6144. doi: [10.1109/CVPR46437.2021.00607](https://doi.org/10.1109/CVPR46437.2021.00607).
- [60] SUN Pei, WANG Weiyue, CHAI Yuning, *et al.* RSN: Range sparse net for efficient, accurate LiDAR 3D object detection[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 5725–5734. doi: [10.1109/CVPR46437.2021.00567](https://doi.org/10.1109/CVPR46437.2021.00567).
- [61] YIN Tianwei, ZHOU Xingyi, and KRÄHENBÜHL P. Center-based 3D object detection and tracking[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 11784–11793. doi: [10.1109/CVPR46437.2021.01161](https://doi.org/10.1109/CVPR46437.2021.01161).
- [62] LUO Chenxu, YANG Xiaodong, and YUILLE A. Exploring simple 3D multi-object tracking for autonomous driving[C]. 2021 IEEE/CVF International Conference on Computer Vision, Montreal, Canada, 2021: 10488–10497. doi: [10.1109/ICCV48922.2021.01032](https://doi.org/10.1109/ICCV48922.2021.01032).
- [63] 毛冬海, 李守军, 王锋, 等. LiDAR测量点云融合影像的分块滤波方法[J]. 测绘通报, 2021(10): 67–72, 131. doi: [10.13474/j.cnki.11-2246.2021.307](https://doi.org/10.13474/j.cnki.11-2246.2021.307).
- MAO Donghai, LI Shoujun, WANG Feng, *et al.* Block filtering method for LiDAR point cloud fusion image[J]. *Bulletin of Surveying and Mapping*, 2021(10): 67–72, 131. doi: [10.13474/j.cnki.11-2246.2021.307](https://doi.org/10.13474/j.cnki.11-2246.2021.307).
- [64] CHEN Xuesong, SHI Shaoshuai, ZHU Benjin, *et al.* MPPNet: Multi-frame feature intertwining with proxy points for 3D temporal object detection[C]. The 17th European Conference on Computer Vision, Tel Aviv, Israel, 2022: 680–697. doi: [10.1007/978-3-031-20074-8_39](https://doi.org/10.1007/978-3-031-20074-8_39).
- [65] HU Yihan, DING Zhuangzhuang, GE Runzhou, *et al.* AFDetV2: Rethinking the necessity of the second stage for object detection from point clouds[C]. The 36th AAAI Conference on Artificial Intelligence, 2022: 969–979. doi: [10.1609/aaai.v36i1.19980](https://doi.org/10.1609/aaai.v36i1.19980).
- [66] HE Chenhang, LI Ruihuang, ZHANG Yabin, *et al.* MSF: Motion-guided sequential fusion for efficient 3D object detection from point cloud sequences[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, 2023: 5196–5205. doi: [10.1109/CVPR52729.2023.00503](https://doi.org/10.1109/CVPR52729.2023.00503).
- [67] SHI Lingfeng, LV Yunfeng, YIN Wei, *et al.* Autonomous multiframe point cloud fusion method for mmWave radar[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 4505708. doi: [10.1109/TIM.2023.3302936](https://doi.org/10.1109/TIM.2023.3302936).
- [68] YANG Yanding, JIANG Kun, YANG Diange, *et al.* Temporal point cloud fusion with scene flow for robust 3D object tracking[J]. *IEEE Signal Processing Letters*, 2022, 29: 1579–1583. doi: [10.1109/LSP.2022.3185948](https://doi.org/10.1109/LSP.2022.3185948).
- [69] CHEN Hui, FENG Yan, YANG Jian, *et al.* 3D reconstruction approach for outdoor scene based on multiple point cloud fusion[J]. *Journal of the Indian Society of Remote Sensing*, 2019, 47(10): 1761–1772. doi: [10.1007/s12524-019-01029-y](https://doi.org/10.1007/s12524-019-01029-y).
- [70] CHENG Yuqi, LI Wenlong, JIANG Cheng, *et al.* MVGR: Mean-variance minimization global registration method for multiview point cloud in robot inspection[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 8504915. doi: [10.1109/TIM.2024.3413191](https://doi.org/10.1109/TIM.2024.3413191).
- [71] LIU Junjie, LIU Junlong, YAN Shaotian, *et al.* MPC: Multi-view probabilistic clustering[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, 2022: 9509–9518. doi: [10.1109/CVPR52688.2022.00929](https://doi.org/10.1109/CVPR52688.2022.00929).
- [72] KLOEKER L, KOTULLA C, and ECKSTEIN L. Real-time point cloud fusion of multi-LiDAR infrastructure sensor setups with unknown spatial location and orientation[C]. 2020 IEEE 23rd International Conference on Intelligent Transportation Systems, Rhodes, Greece, 2020: 1–8. doi: [10.1109/ITSC45102.2020.9294312](https://doi.org/10.1109/ITSC45102.2020.9294312).
- [73] YU Xumin, TANG Lulu, RAO Yongming, *et al.* PointBERT: Pre-training 3D point cloud transformers with masked point modeling[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, USA, 2022: 19313–19322. doi: [10.1109/CVPR52688.2022.01871](https://doi.org/10.1109/CVPR52688.2022.01871).
- [74] LI Yao, ZHOU Yong, ZHAO Jiaqi, *et al.* FA-MSVNet: Multi-scale and multi-view feature aggregation methods for stereo 3D reconstruction[J]. *Multimedia Tools and Applications*, 2024: 1–23. doi: [10.1007/s11042-024-20431-4](https://doi.org/10.1007/s11042-024-20431-4).
- [75] CHEN Zhimin, LI Yingwei, JING Longlong, *et al.* Point cloud self-supervised learning via 3D to multi-view masked autoencoder[J]. *arXiv preprint arXiv*, 2311.10887, 2023.
- [76] LING Xiao and QIN Rongjun. A graph-matching approach for cross-view registration of over-view and street-view based point clouds[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 185: 2–15. doi: [10.1016/j.isprsjprs.2021.12.013](https://doi.org/10.1016/j.isprsjprs.2021.12.013).
- [77] 深圳市博大建设集团有限公司. 一种多视角三维激光点云拼接方法及系统[P]. 中国, 202411659468.1, 2024.
- Shenzhen Boda Construction Group Co., Ltd. Multi-view three-dimensional laser point cloud splicing method and system[P]. CN, 202411659468.1, 2024.
- [78] 北京鹏太光耀科技有限公司. 一种基于多目结构光的全向点云融合方法[P]. 中国, 202411346535.4, 2025.
- Beijing Pengtai Guangyao Technology Co., Ltd. Omnidirectional point cloud fusion method based on multi-

- view structured light[P]. CN, 202411346535.4, 2025.
- [79] KHAN K N, KHALID A, TURKAR Y, *et al.* VRF: Vehicle road-side point cloud fusion[C]. The 22nd Annual International Conference on Mobile Systems, Applications and Services, Tokyo, Japan, 2024: 547–560. doi: [10.1145/3643832.3661874](https://doi.org/10.1145/3643832.3661874).
- [80] BREYER M, OTT L, SIEGWART R, *et al.* Closed-loop next-best-view planning for target-driven grasping[C]. 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems, Kyoto, Japan, 2022: 1411–1416. doi: [10.1109/IRoS47612.2022.9981472](https://doi.org/10.1109/IRoS47612.2022.9981472).
- [81] JIANG Changjian, GAO Ruilan, SHAO Kele, *et al.* LI-GS: Gaussian splatting with LiDAR incorporated for accurate large-scale reconstruction[J]. *IEEE Robotics and Automation Letters*, 2025, 10(2): 1864–1871. doi: [10.1109/LRA.2024.3522846](https://doi.org/10.1109/LRA.2024.3522846).
- [82] YANG Shuo, LU Huimin, and LI Jianru. Multifeature fusion-based object detection for intelligent transportation systems[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(1): 1126–1133. doi: [10.1109/TITS.2022.3155488](https://doi.org/10.1109/TITS.2022.3155488).
- [83] MA Nan, XIAO Chuansheng, WANG Mohan, *et al.* A review of point cloud and image cross-modal fusion for self-driving[C]. 2022 18th International Conference on Computational Intelligence and Security, Chengdu, China, 2022: 456–460. doi: [10.1109/CIS58238.2022.00103](https://doi.org/10.1109/CIS58238.2022.00103).
- [84] LIU Zhijian, TANG Haotian, AMINI A, *et al.* BEVFusion: Multi-task multi-sensor fusion with unified bird's-eye view representation[C]. 2023 IEEE International Conference on Robotics and Automation, London, UK, 2023: 2774–2781. doi: [10.1109/ICRA48891.2023.10160968](https://doi.org/10.1109/ICRA48891.2023.10160968).
- [85] KUMAR G A, LEE J H, HWANG J, *et al.* LiDAR and camera fusion approach for object distance estimation in self-driving vehicles[J]. *Symmetry*, 2020, 12(2): 324. doi: [10.3390/sym12020324](https://doi.org/10.3390/sym12020324).
- [86] HUANG Keli, SHI Botian, LI Xiang, *et al.* Multi-modal sensor fusion for auto driving perception: A survey[J]. *arXiv preprint arXiv*, 2202.02703, 2022.
- [87] CAI Yiyi, OU Yang, and QIN Tuanfa. Improving SLAM techniques with integrated multi-sensor fusion for 3D reconstruction[J]. *Sensors*, 2024, 24(7): 2033. doi: [10.3390/s24072033](https://doi.org/10.3390/s24072033).
- [88] RODRIGUEZ-GARCIA B, RAMÍREZ-SANZ J M, MIGUEL-ALONSO I, *et al.* Enhancing learning of 3D model unwrapping through virtual reality serious game: Design and usability validation[J]. *Electronics*, 2024, 13(10): 1972. doi: [10.3390/electronics13101972](https://doi.org/10.3390/electronics13101972).
- [89] KAVALIAUSKAS P, FERNANDEZ J B, MCGUINNESS K, *et al.* Automation of construction progress monitoring by integrating 3D point cloud data with an IFC-based BIM model[J]. *Buildings*, 2022, 12(10): 1754. doi: [10.3390/buildings12101754](https://doi.org/10.3390/buildings12101754).
- [90] XU Yusheng and STILLA U. Toward building and civil infrastructure reconstruction from point clouds: A review on data and key techniques[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 2857–2885. doi: [10.1109/JSTARS.2021.3060568](https://doi.org/10.1109/JSTARS.2021.3060568).
- [91] JHONG S Y, CHEN Y Y, HSIA C H, *et al.* Density-aware and semantic-guided fusion for 3-D object detection using LiDAR-camera sensors[J]. *IEEE Sensors Journal*, 2023, 23(18): 22051–22063. doi: [10.1109/JSEN.2023.3302314](https://doi.org/10.1109/JSEN.2023.3302314).
- [92] LUO Zhipeng, ZHOU Changqing, PAN Liang, *et al.* Exploring point-BEV fusion for 3D point cloud object tracking with transformer[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(9): 5921–5935. doi: [10.1109/TPAMI.2024.3373693](https://doi.org/10.1109/TPAMI.2024.3373693).
- [93] LIU Jianan, ZHAO Qiuchi, XIONG Weiye, *et al.* SMURF: Spatial multi-representation fusion for 3D object detection with 4D imaging radar[J]. *IEEE Transactions on Intelligent Vehicles*, 2024, 9(1): 799–812. doi: [10.1109/TIV.2023.3322729](https://doi.org/10.1109/TIV.2023.3322729).
- [94] HAO Ruidong, WEI Zhongwei, HE Xu, *et al.* Robust point cloud registration network for complex conditions[J]. *Sensors*, 2023, 23(24): 9837. doi: [10.3390/s23249837](https://doi.org/10.3390/s23249837).
- [95] BAYRAK O C, MA Zhenyu, MARIAROSARIA FARELLA E, *et al.* Operationalize large-scale point cloud classification: Potentials and challenges[C]. EGU General Assembly Conference Abstracts 2024, Vienna, Austria, 2024: 10034. doi: [10.5194/egusphere-egu24-10034](https://doi.org/10.5194/egusphere-egu24-10034).
- [96] SHAN Shuo, LI Chenxi, WANG Yiye, *et al.* A deep learning model for multi-modal spatio-temporal irradiance forecast[J]. *Expert Systems with Applications*, 2024, 244: 122925. doi: [10.1016/j.eswa.2023.122925](https://doi.org/10.1016/j.eswa.2023.122925).
- [97] QIAN Jiaming, FENG Shijie, TAO Tianyang, *et al.* Real-time 3D point cloud registration[C]. SPIE 11205, Seventh International Conference on Optical and Photonic Engineering, Phuket, Thailand, 2019: 495–500. doi: [10.1117/12.2547865](https://doi.org/10.1117/12.2547865).
- [98] SHI Fengyuan, ZHENG Xunjiang, JIANG Lihui, *et al.* Fast point cloud registration algorithm using parallel computing strategy[J]. *Journal of Physics: Conference Series*, 2022, 2235(1): 012104. doi: [10.1088/1742-6596/2235/1/012104](https://doi.org/10.1088/1742-6596/2235/1/012104).
- [99] LUO Yukui. Securing FPGA as a shared cloud-computing resource: Threats and mitigations[D]. [Ph.D. dissertation], Northeastern University, 2023.
- [100] WANG Laichao, LU Weiding, TIAN Yuan, *et al.* 6D

object pose estimation with attention aware bi-gated fusion[C]. The 30th International Conference on Neural Information Processing, Changsha, China, 2024: 573–585. doi: [10.1007/978-981-99-8082-6_44](https://doi.org/10.1007/978-981-99-8082-6_44).

[101] HONG Jiaxin, ZHANG Hongbo, LIU Jinghua, *et al.* A transformer-based multi-modal fusion network for 6D pose estimation[J]. *Information Fusion*, 2024, 105: 102227. doi: [10.1016/j.inffus.2024.102227](https://doi.org/10.1016/j.inffus.2024.102227).

作者简介

宋绍京，博士，教授，主要研究方向为机器视觉、自动驾驶、光电信息处理等。

方非易，博士，主要研究方向为图像分割、多模态机器学习、无人驾驶等。

李新建，硕士生，主要研究方向为自动驾驶、点云融合、3D目标检测等。

(责任编辑：张鑫源 高山流水)