

面向智能驾驶的自适应多尺度三维点云配准网络

汪霖, 韩栋, 韩景怡, 陈圣梅, 刘成
(西北大学电子信息学院, 710127, 西安)

摘要: 针对智能驾驶场景中, 激光雷达点云稀疏、特征模糊以及低重叠引发关键点误匹配的问题, 提出了一种面向智能驾驶的自适应多尺度三维点云配准网络。首先, 该网络利用自适应多尺度特征提取模块获取采样点的多层级特征, 并结合自适应邻域稳定器实现关键点的精确定位, 从而解决由于特征模糊而导致的匹配难题。其次, 通过融合多尺度特征匹配一致性与几何一致性校验, 提升待配准点对的内点置信度。最后, 基于该置信度, 采用由粗到精的迭代优化策略, 完成三维点云的精确配准。在 KITTI 数据集上的测试结果表明: 该网络的相对平移误差和相对旋转误差分别低至 0.061 8 m 和 0.21°, 配准召回率高达 99.8%。在 nuScenes 数据集上的测试结果表明: 在帧间距离分别为 10、20、30 m 时, 配准召回率分别达到 90.1%、83.3% 和 78.8%。上述结果验证了该网络能够有效抑制复杂道路场景下的误匹配, 实现高精度、高鲁棒性的大规模三维点云配准。该研究适配了动态复杂路况下的实际应用需求, 可为自动驾驶场景下的精准定位、环境感知提供可靠的技术支撑。

关键词: 智能驾驶; 三维点云配准; 自适应; 多尺度特征

中图分类号: TP29 **文献标志码:** A

DOI: 10.7652/xjtub202610003 **文章编号:** 0253-987X(2026)10-0034-11

Adaptive Multi-Scale Three-Dimensional Point Cloud Registration Network for Autonomous Driving

WANG Lin, HAN Dong, HAN Jingyi, CHEN Shengmei, LIU Cheng
(School of Electronic Information, Northwest University, Xi'an 710127, China)

Abstract: An adaptive multi-scale three-dimensional point cloud registration network for autonomous driving is proposed to address the issues of sparse light detection and ranging point clouds, ambiguous features, and keypoint mismatches caused by low overlap in autonomous driving scenarios. First, multi-level features of sampled points are extracted via an adaptive multi-scale feature extraction module, and precise keypoint localization is achieved by incorporating an adaptive neighborhood stabilizer, thereby resolving the matching difficulties caused by feature ambiguity. Second, the inlier confidence of the point pairs to be registered is enhanced by integrating multi-scale feature matching consistency with geometric consistency verification. Finally, based on this confidence, a coarse-to-fine iterative optimization strategy is adopted to accomplish accurate three-dimensional point cloud registration. Test results on the KITTI dataset indicate that a relative translation error as low as 0.061 8 m and a relative rotation error as low as 0.21° are achieved by the network, with a registration recall of up to 99.8%. Furthermore, test results on the nuScenes dataset show that registration recalls of 90.1%, 83.3%, and 78.8% are achieved at inter-frame

收稿日期: 2026-04-21。 作者简介: 汪霖(1983—), 男, 副教授, 博士生导师; 刘成(通信作者), 男, 高级工程师。

基金项目: 陕西省重点研发计划资助项目(2024GX-YBXM-149)。

网络出版时间: 2026-06-24

网络出版地址: <https://link.cnki.net/urlid/61.1069.t.20260623.1712.012>

distances of 10, 20, and 30 m, respectively. It is validated by the aforementioned results that mismatches in complex road scenarios are effectively suppressed by the proposed network, and high-precision, highly robust large-scale three-dimensional point cloud registration is realized. The practical application requirements under dynamic and complex traffic conditions are met by this research, and reliable technical support is provided for precise localization and environmental perception in autonomous driving scenarios.

Keywords: intelligent driving; three-dimensional point cloud registration; adaptive; multi-scale features

自动驾驶技术的迭代升级,推动了汽车行业的智能化转型。智能驾驶感知体系融合了车载相机、毫米波雷达、超声波雷达及激光雷达等多类传感器,其中,视觉感知与超声感知分别在纹理识别、近距离测距方面具备优势,但易受极端光照、恶劣天气与探测距离的约束,环境适应性有限。激光雷达三维感知可输出高精度、高密度的显性三维几何结构信息,并凭借主动探测、不受光照干扰的特性,成为实现全天候、高鲁棒性智能驾驶感知的核心器件,同时也是多传感器融合架构中空间对齐的绝对基准。自动驾驶系统依托连续采集的激光雷达三维点云,构建时空一致的三维环境表征,核心关键在于将多帧点云统一至同一坐标系下,并完成精准对齐与融合,以保障三维环境重建的可靠性。三维点云配准不仅可实现连续点云扫描,扫描与局部地图间的相对位姿估计,还是多传感器外参标定、高精地图构建及即时定位与建图的前置基础,其配准精度直接决定多帧点云融合的空间一致性,进而影响三维重建与高精地图的最终质量^[1-5]。因此,充分发挥激光雷达点云的几何优势,实现高精度、高鲁棒性的三维点云配准,是自动驾驶场景下高质量建图、多源传感器融合与精准稳定定位的核心前提^[6-8]。

然而,相较于室内与静态场景,真实道路环境下的三维点云配准存在诸多难点与挑战。车辆行驶引发的高速运动与视角剧变,会造成相邻帧点云重叠区域的不稳定,同时导致远距离点云稀疏、几何结构残缺。道路场景中,车辆、行人等动态交通参与者产生的非刚体点,易引发匹配偏差,显著限制了传统基于最近邻搜索或单尺度特征配准方法的鲁棒性与稳定性。此外,直线路段、护栏、隧道等场景中广泛存在的几何退化结构,进一步提升了复杂道路场景的点云配准难度^[9-12]。

传统迭代最近点(ICP)方法^[13]对初值敏感,极易出现关键点误匹配、收敛速度慢甚至不收敛的问题,难以满足自动驾驶实时、精准的配准需求。而基于深度学习的三维点云配准方法通过学习判别性特

征在一定程度上提升了匹配精度。Wang 等^[14]提出了基于深度最近点的三维点云配准方法,该方法通过深度特征提取、注意力匹配和可微的奇异值分解求解刚体变换矩阵。为了减弱对初始位姿的依赖,Yew 等^[15]提出了一种三维点云配准网络 RPM-Net,该网络通过利用 Sinkhorn 层与退火策略学习软对应关系,提取空间坐标与局部几何的混合特征,削弱了传统最近邻匹配对初始位姿的敏感性。但该网络在具有大量离散点和复杂结构的配准中缺乏强有力的约束。为此,Fu 等^[16]提出了 RGM 网络,将三维点云转换为图结构,通过增强自身点及其邻域特征的关系,提升在离群点干扰下,点对应关系的质量。为进一步增强高离群点条件下的配准稳定性,Bai 等^[17]提出了 PointDSC 网络,通过引入深度空间一致性和可微谱匹配机制,对候选对应增强鲁棒筛选。随后,Yew 等^[18]提出了 REGTR 网络,利用了 Transformer 的自注意力与交叉注意力机制可直接预测出高质量的对应关系,并且取得了无需最近邻与随机抽样一致算法后处理的突破。不过,REGTR 对刚体不变的几何约束相对不足,因此后续研究逐渐开始将距离、角度等几何关系编码融进网络表示中。Qin 等^[19]提出了 GeoTransformer,通过将点对距离和三元组角度等几何关系引入 Transformer 框架,增强了特征表示的刚体变换不变性,但其模型的复杂度和计算代价较高。然而,以上方法往往依赖固定尺度的特征表示,难以同时兼顾近距离密集区域与远距离稀疏区域的结构差异,以及在复杂自动驾驶场景中的实际应用效果。

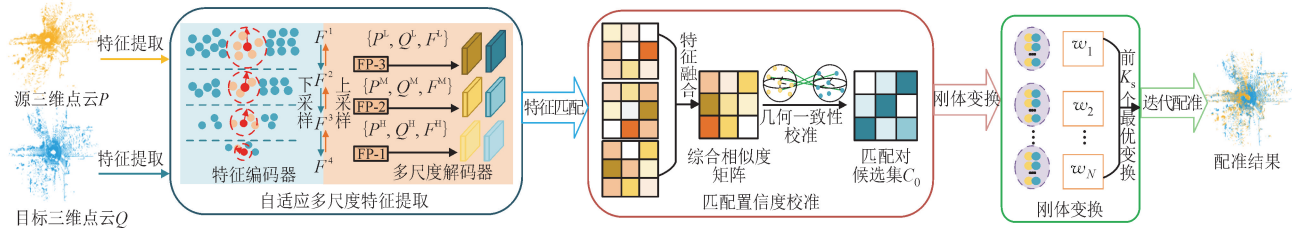
综上所述,在自动驾驶场景中,现有方法因关键点特征模糊、误匹配频发而导致配准精度下降,无法满足自动驾驶实时感知与定位需求。为了解决上述问题,本文提出了一种面向智能驾驶的自适应多尺度三维点云配准网络(AMS-Net)。该网络通过自适应提取关键点多尺度特征,基于匹配一致性形成点对之间的可靠性约束。引入自适应邻域稳定器增

强对关键点选择的稳定性。基于几何一致性校验实现对内点的精确提取,并采用置信度迭代进行精配准,有效缓解因特征模糊导致的误匹配问题,提高了配准精度,为自动驾驶场景下的精准定位、环境感知提供可靠的技术支撑,适配动态复杂路况下的实际应用需求。

1 AMS-Net 网络结构

本文提出的面向智能驾驶的自适应多尺度三维

点云配准网络整体模型结构如图 1 所示。该网络由自适应多尺度特征提取、匹配置信度校准以及刚体变换估计这 3 个核心模块协同作用,通过学习三维点云的多尺度特征,实现高效的外点滤除和精细准配准。具体而言,在特征提取阶段,设计了增强特征区分度的自适应多尺度特征提取模块,以分层特征抽样框架为基础,借助自适应邻域稳定器(ANS)检测关键点,并提取其邻域显著特征,有效缓解复杂道路大规模场景下,局部特征模糊与全局信息缺失的问题。



$F^1 \sim F^4$ —1~4 级下采样编码器提取的层级特征;FP-1~FP-3—特征传播模块; P^L, Q^L, F^L —低尺度下提取的源、目标点云及其特征; P^M, Q^M, F^M —中尺度下提取的源、目标点云及其特征; P^H, Q^H, F^H —高尺度下提取的源、目标点云及其特征; $w_1 \sim w_N$ —匹配点对的匹配置信度; K_s —粗配准阶段筛选的高置信度匹配点对数量。

图 1 AMS-Net 整体模型结构

Fig. 1 Overall model structure of AMS-Net

在特征匹配阶段,构建了匹配置信度校准模块,以不同尺度特征下的匹配一致性为原则构建匹配点对,结合多尺度特征相似度与刚体变换不变的空间几何约束,实现对高置信度内点的加权筛选与异常匹配的有效排除。在刚体变换估计阶段,利用匹配点对的几何一致性约束和匹配置信度逐步精细化配准迭代过程,进而实现高精度三维点云配准。

1.1 自适应多尺度特征提取模块

给定源三维点云 $P = \{p_i\}_{i=1}^{N_p} \in \mathbf{R}^{N_p \times 3}$ 和目标三维点云 $Q = \{q_j\}_{j=1}^{N_q} \in \mathbf{R}^{N_q \times 3}$, 其中, p_i, q_j 分别为源和目标三维点云中第 i 、第 j 个点, N_p, N_q 分别为源和目标三维点云的个数。以源三维点云和目标三维点云作为输入, AMS-Net 构建了一个自适应多尺度特征提取模块, 精准提取并强化不同尺度下的特征信息。自适应多尺度特征提取模块结构如图 2 所示。

此模块基于 PointNet++^[20] 的层次化集合抽样(SA)思想, 构建了经典的编码和解码对称结构。其中, 特征编码器采用 ANS 精确采样点的特征信息, 并将原始三维点云映射到高维特征空间, 自适应地进行分层下采样。而多尺度解码器则基于 PointNet++ 的特征传播(FP)模块构建多尺度的特征提取的上采样分支。二者协同作用, 为后续的三维点云配准提供了丰富的特征支持。

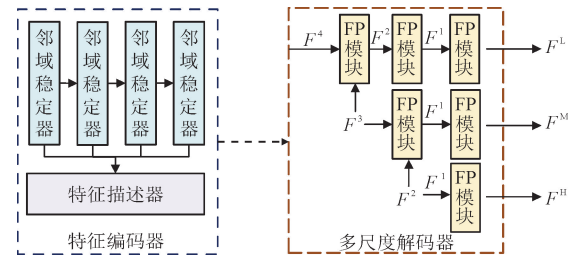


图 2 多尺度特征提取模块结构

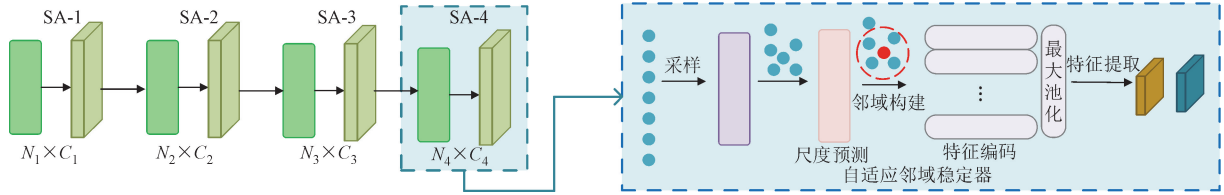
Fig. 2 Structure of multi-scale feature extraction module

特征编码器采用 4 级级联的下采样模块扩大感受野, 编码三维点云在不同尺度下的特征信息。随着网络深度的增加, 采样点的感受野可随之持续扩大, 不仅可以捕获三维点云的精细局部细节, 还可以融合更加完整的全局信息。因此, AMS-Net 采用了 4 级级联的下采样模块构建该编码器, 其中每一级均包含关键点采样、局部邻域聚合和特征提取的完整流程。一级下采样筛选初始关键点, 以保留点云的主体结构, 并生成初始匹配候选点, 后三级则基于初筛关键点, 逐层构建多尺度邻域特征。与此同时, 各层下采样均保留了层级特征与关键点坐标, 以作为解码器的输入并用于上采样及匹配一致性校验。经 4 层迭代下采样, 网络感受野逐层扩张, 从而充分挖掘点云深层结构特征。

多尺度解码器通过自适应邻域半径反距离加权插值,同时采用并行分支结构将各级下采样子集恢复至原始点云规模。为了能输出关键点的多尺度特征,AMS-Net 设置了 3 条解码支路分别对应邻域半径为 0.5~2.0 m 的低尺度局部细节、邻域半径为 2.0~6.0 m 的中尺度结构信息、邻域半径为 6.0~15.0 m 的高尺度全局特征。以低、中、高这 3 个尺度的稀疏特征作为输入,输出多特征 F^L 、 F^M 和 F^H 。定义级联的特征传播模块 $F_P(F^1, F^2)$ 表达式为

$$F_P(F^1, F^2) = \text{MLP}(\text{cat}(\text{Up}(F^2), F^1)) \quad (1)$$

式中: $\text{cat}(\cdot)$ 为特征拼接操作; $\text{Up}(\cdot)$ 为反距离加权插值上采样; $\text{MLP}(\cdot)$ 为多层感知机,在具体网络结构中,该操作由一系列级联的 1×1 一维卷积 (Conv1D)、批归一化和 ReLU 非线性激活函数实现。该映射函数独立作用于每个采样点,旨在将拼接后的高维特征在通道维度上进行深度融合与自适应降维,从而提取更具判别力的局部上下文信息。



SA-1~SA-4—1~4 级分层下采样模块; $N_1 \sim N_4$ —各级下采样后点云的点数; $C_1 \sim C_4$ —各级特征通道数。

图 3 自适应邻域稳定器结构

Fig. 3 Structure of adaptive neighborhood stabilizer

在特征提取网络中对输入三维点云进行 4 次下采样。其中,第 l 层的输入为上一层下采样得到的点集与其特征 $\{\mathbf{P}^{l-1}, \mathbf{F}^{l-1}\}$, 具体定义为

$$\begin{cases} \mathbf{P}^{l-1} = \{\mathbf{P}_j^{l-1}\}_{j=1}^{N_{l-1}} \\ \mathbf{F}^{l-1} \in \mathbf{R}^{N_{l-1} \times C_{l-1}} \end{cases} \quad (2)$$

式中: $\mathbf{P}^{l-1}$ 为点集; $\mathbf{F}^{l-1}$ 为特征; C_{l-1} 为上一层点特征的维度; N_{l-1} 为上一层下采样后点的个数; $\mathbf{P}_j^{l-1}$ 为 $l-1$ 级点集中的第 j 个点,用于参与当前层级的邻域构建与特征聚合。

首先,编码器 ANS 使用 FPS 从 $\mathbf{P}^{l-1}$ 中选取 N_l 个代表点作为本层的中心点集合,利用基于特征的插值算子从上一层特征中获得中心点的邻域特征表达式为

$$\mathbf{g}_i^l = \text{Interp}(\mathbf{p}_i^l, \mathbf{P}^{l-1}, \mathbf{F}^{l-1}) \in \mathbf{R}^{C_{l-1}} \quad (3)$$

式中: $\mathbf{g}_i^l$ 为第 l 级第 i 个中心点的上下文特征向量; $\text{Interp}(\cdot)$ 为特征插值算子,用于从上层点云特征中插值计算当前中心点的邻域特征。

其次,通过尺度预测网络对点级邻域半径进行

应降维,从而提取更具判别力的局部上下文信息。

1.2 自适应邻域稳定器

特征编码器可通过分层下采样,完成点云多尺度特征提取,而下采样策略的稳定性直接影响关键点选取质量与最终配准性能。PointNet++ 所采用的最远点采样 (FPS) 具备良好的空间覆盖能力,能够保障采样点的分布离散性。然而, FPS 对初始点的选取与局部点的分布较为敏感,在三维点云部分重叠、无效区域较多的复杂配准场景下,单一使用 FPS 选取关键点,易出现采样分布不均、有效对应点缺失等问题,增加了有效匹配点对的构建难度。

为了解决上述问题,AMS-Net 在特征编码器的每层中都内置了一个自适应邻域稳定器,结构如图 3 所示。在保留 FPS 空间均匀优势的基础上,引入自适应机制对局部邻域构建与特征提取过程进行优化,以提高关键点选择的稳定性和配准鲁棒性。

回归。该网络由一个线性映射层与非线性激活函数组成,表达式为

$$\begin{cases} \tilde{\sigma}_i^l = \mathbf{a}^T \mathbf{g}_i^l + \beta \\ \sigma_i^l = \text{softplus}(\tilde{\sigma}_i^l) + \epsilon \end{cases} \quad (4)$$

式中: $\tilde{\sigma}_i^l$ 为第 l 级第 i 个中心点的初始邻域半径; $\mathbf{a}$ 和 β 分别为线性映射层的可学习权重向量与偏置; σ_i^l 为第 l 级第 i 个中心点的最终自适应邻域半径; $\text{softplus}(\cdot)$ 为非线性激活函数; $\epsilon > 0$ 为常数。

最后,通过上述步骤,不同空间位置的点可根据其局部结构复杂度自适应选择感受野大小,这样可避免固定尺度在稀疏或密集区域下带来的特征退化问题。以点 $\mathbf{p}_i^l$ 为圆心、 σ_i^l 为半径,在上一层点集上构建尺度自适应邻域,表达式为

$$\zeta_i^l = \{\mathbf{P}_j^{l-1} \in \mathbf{P}^{l-1} \mid \|\mathbf{P}_j^{l-1} - \mathbf{p}_i^l\|_2 \leq \sigma_i^l\} \quad (5)$$

式中: ζ_i^l 为第 l 级第 i 个中心点的自适应邻域点集合; $\|\cdot\|_2$ 为欧氏范数; j 表示第 j 个中心点。

为了实现对局部信息的高效编码,将提取到的邻域特征输入至专门设计的特征描述器中,结构如

图 4 所示。该描述器由 3 层 MLP 与最大池化层级联构成,负责输出重定位关键点的初步特征描述子。以重定位的采样点为中心,以自适应的邻域半径重新构建每个关键的三维点云特征。最后,邻域特征被输入描述子网络生成采样点的描述子。

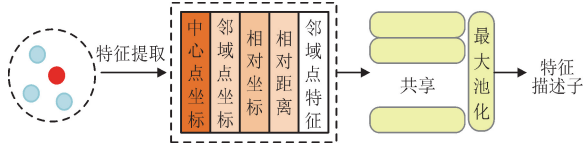


图 4 特征描述器结构

Fig. 4 Structure of feature descriptors

综上所述,自适应多尺度特征提取模块通过学习每个采样点的局部结构复杂度,自适应回归邻域半径,从而实现对不同密度区域的特征调整。同时,结合自适应邻域稳定器,以确保关键点选择的稳定性,从而在多尺度特征下实现高质量特征的自适应提取。

1.3 基于几何一致性的置信度校准模块

经过自适应多尺度特征网络后,可获取到源三维点云与目标三维点云关键点在低、中、高 3 个尺度下的特征描述子。为了降低弱内点在特征匹配阶段被误删除的风险,AMS-Net 不采用根据硬阈值匹配点对的传统方法,而是基于多尺度特征相似性构建软对应集合,并考虑几何约束,实现对匹配点的精确筛选。

在特征匹配的初始阶段,一般的多尺度网络往往在特征层将不同尺度直接拼接成一个单一的全尺度表示,并据此进行一次性相似度计算,这种处理方式对三维点云密度不均和复杂道路大规模场景极为敏感。因此,AMS-Net 首先保留了各尺度独立的判别能力,并分别计算源三维点云与目标三维点云关键点在单一尺度下的相似度,表达式为

$$S_{ij}^r = \left\langle \frac{\mathbf{d}_i^p}{\|\mathbf{d}_i^p\|_2}, \frac{\mathbf{d}_j^q}{\|\mathbf{d}_j^q\|_2} \right\rangle \quad (6)$$

式中: $\mathbf{d}_i^p$ 、 $\mathbf{d}_j^q$ 为源点云与目标点云第 i 、 j 个关键点的特征描述子; S_{ij}^r 为单一尺度下源三维点云中的第 i 个关键点和目标三维点云的第 j 个关键点之间的特征相似度; 上标 r 可取 L、M、H, 为低、中、高 3 个任意尺度; $\langle \cdot \rangle$ 表示计算 2 个向量的点积。

为了在保留各尺度独立判别能力的同时,实现跨尺度信息的互补,基于正确匹配对在任意尺度下均应具备特征一致性的基本原理,将低、中、高 3 个尺度下的相似性矩阵 $\mathbf{S}^L$ 、 $\mathbf{S}^M$ 和 $\mathbf{S}^H$ 进行融合叠加,得

到跨尺度综合相似度矩阵 $\mathbf{S}$ 。对于具体的匹配点对,其在矩阵 $\mathbf{S}$ 中对应的综合相似度 S_{ij} , 表达式为

$$S_{ij} = S_{ij}^L + S_{ij}^M + S_{ij}^H \quad (7)$$

式中: S_{ij}^L 、 S_{ij}^M 、 S_{ij}^H 分别为 3 个尺度下相似性矩阵中的元素; $S_{ij} \in [0, 3]$ 为构建初始候选对应集合 $C_0 = \{(\mathbf{x}_i, \mathbf{y}_i^*) \mid i = 1, \dots, N\}$ 中的综合相似度。其中, N 为候选点对的个数, $\mathbf{x}_i$ 是源三维点云 P 中被筛选出来的关键点, $\mathbf{y}_i^*$ 是目标三维点云 Q 中与 $\mathbf{x}_i$ 相对应的候选匹配点。理想情况下,真实匹配点对在多个尺度下均应表现出较高的特征相似度,因此其综合置信度通常显著高于伪匹配点对。

多尺度特征融合虽能刻画语义相似性,但未考虑几何结构的刚性约束,在重复结构场景下,仍然容易将特征相似但几何错位的伪内点误判为有效匹配。因此,在多尺度特征融合的基础上,引入几何一致性校准(GCC)机制,通过量化匹配对的邻域几何结构一致性,可对综合特征置信度进行修正,形成语义相似和几何一致的双维度判别依据。GCC 机制模型结构如图 5 所示。

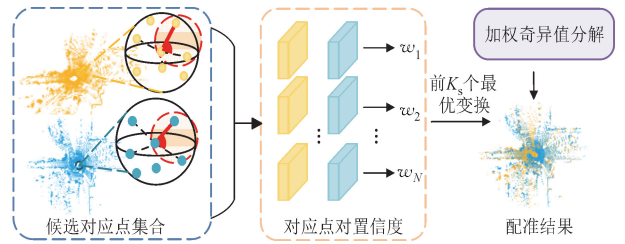


图 5 GCC 机制模型结构

Fig. 5 Model structure of the GCC mechanism

由于刚体的变换不变性,真实匹配点对在其邻域结构中应保持稳定的几何关系。因此,GCC 通过比较候选匹配对其邻域点之间的距离关系差异,构建几何一致性度量,并将其与多尺度特征相似度进行融合,从而形成更加可靠的匹配置信度估计。为了提升匹配点对的置信度,AMS-Net 基于候选匹配集合 C_0 为每个匹配点对构建局部邻域图结构,该结构包含边长和角度两种几何描述。在此基础上,依据真实匹配点对在刚体变换下应保持邻域内距离和角度的关系稳定性,通过比较源三维点云与目标三维点云邻域结构的差异,计算几何一致性度量,并使用多层感知器融合特征,从而实现匹配点对置信度的校准。

在候选集合 C_0 内,对于第 i 个候选匹配点对 $(\mathbf{x}_i, \mathbf{y}_i^*)$,分别在源点云与目标点云选取相同 K 个最近点,以构建具有刚体变换不变性的局部几何约

束。具体而言,针对邻域 $N(\mathbf{x}_i)$ 和 $N(\mathbf{y}_i^*)$ 构造邻域图结构,并由边 $\mathbf{e}_{i,k}^x, \mathbf{e}_{i,k}^{y^*}$ 和角 $a_{r,s}^x, a_{r,s}^{y^*}$ 两种几何表示,表达式为

$$\mathbf{e}_{i,k}^x = \mathbf{x}_i - \mathbf{x}_k; a_{r,s}^x = \angle(\mathbf{e}_{i,r}^x, \mathbf{e}_{i,s}^x) \quad (8)$$

$$\mathbf{e}_{i,k}^{y^*} = \mathbf{y}_i^* - \mathbf{y}_k^*; a_{r,s}^{y^*} = \angle(\mathbf{e}_{i,r}^{y^*}, \mathbf{e}_{i,s}^{y^*}) \quad (9)$$

式中: $\mathbf{x}_k$ 和 $\mathbf{y}_k^*$ 分别为中心点 $\mathbf{x}_i$ 和 $\mathbf{y}_i^*$ 在其局部邻域 $N(\mathbf{x}_i)$ 和 $N(\mathbf{y}_i^*)$ 的第 k 个邻近点, $i \neq k, k = 1, \dots, K$; r, s 为邻域内近邻点的索引, $r, s \in \{1, \dots, K\}$ 。

在获取基础的几何结构表示后,为了更敏锐地捕捉上、下文信息并放大真实匹配与伪匹配间的几何差异,利用多层感知器 f_θ 进行特征融合,表达式为

$$\Delta_{i,k} = f_\theta(\text{concat}(\mathbf{e}_{i,k}^x, a_{r,s}^x)) - f_\theta(\text{concat}(\mathbf{e}_{i,k}^{y^*}, a_{r,s}^{y^*})) \quad (10)$$

$$\delta_{i,k} = \text{softmax}(f_\mu(\Delta_{i,k})) \quad (11)$$

式中: $\delta_{i,k}$ 为第 k 组邻域对应的几何一致性归一化权重; f_μ 为可学习单层线性映射层,用于对几何差值做维度缩放; $\Delta_{i,k}$ 为第 k 组邻域的几何特征差值; $\text{concat}(\cdot)$ 为特征拼接操作; $\text{softmax}(\cdot)$ 为归一化指数激活函数。

在获得邻域几何结构表示后,通过多层感知器对几何一致性特征进行融合。通过加权聚合邻域几何一致性信息,得到候选匹配点对的内点置信度 w_i , 表达式为

$$w_i = 1 - \tanh\left(\left|\varphi\left(\sum_{k=1}^K \delta_{i,k} \Delta_{i,k}\right)\right|\right) \quad (12)$$

式中: $\tanh(\cdot)$ 为双曲正切非线性激活函数,用于将无界的几何结构差异平滑地压缩至 $[0, 1]$ 区间内; $\varphi(\cdot)$ 是网络可学习的无偏置单层线性映射函数,形式为 $\varphi(\mathbf{x}) = \mathbf{v}^T \mathbf{x}$, $\mathbf{v}$ 为权重向量,该函数仅做维度变换,将高维特征空间投影至一维标量空间。

由于输入的几何差值 $\Delta_{i,k}$ 是基于刚体变换不变的边与夹角特征计算得到的,因此当候选匹配点对 $(\mathbf{x}_i, \mathbf{y}_i^*)$ 为真实的内点时,其局部几何结构呈现高度的一致性,即 $\sum_{k=1}^K \delta_{i,k} \Delta_{i,k} \rightarrow 0$ 。此时,无偏置线性映射输出标量 $\varphi(0) = 0$,从而得到理想的最高置信度 $w_i \rightarrow 1$ 。反之,若匹配对为伪匹配时,显著的几何邻域结构差异将导致特征描述子的偏离,进而使得置信度 w_i 被抑制并趋近于 0。

1.4 刚体变换估计

在获得候选匹配点及其内点置信度之后,采用由粗到精的配准策略,以估计三维点云之间的刚体变换。传统由粗到精的配准方法仅通过两次独立变

换估计不同,与其不同的是,AMS-Net 在两个阶段均利用内点置信度进行加权优化,从而使匹配权重能够随着配准状态逐步优化。

在配准时,首先根据候选对应的内点置信度 w_i 对集合 C_0 中的候选匹配点降序排序,选取置信度排名靠前的高质量点对构建子集,表达式为

$$C_h = \{(\mathbf{x}_h, \mathbf{y}_h^*) | h \in \text{TopK}_{K_s}(w_i), i = 1, \dots, N_c\} \quad (13)$$

式中: C_h 为筛选得到的高置信匹配点对集合; N_c 为初始候选匹配集合 C_0 内匹配点对总数量; $\text{TopK}_{K_s}(\cdot)$ 为序列筛选算子,用于提取置信度最高的前 K_s 组匹配点对。若第 i 组匹配点对被选中为子集 C_h 中的第 h 个关键点对,则其对应的置信度权重同步映射为 w_h 。

基于该子集构建加权最小二乘优化目标,通过加权奇异值分解求解源三维点云与目标三维点云之间的初始刚体变换,表达式为

$$(\mathbf{R}, \mathbf{t}) = \arg \min_{\mathbf{R}, \mathbf{t}} \sum_{(\mathbf{x}_h, \mathbf{y}_h^*) \in C_h} w_h \|\mathbf{R}\mathbf{x}_h + \mathbf{t} - \mathbf{y}_h^*\|_2^2 \quad (14)$$

式中: $\mathbf{R}, \mathbf{t}$ 分别为旋转矩阵和平移向量; w_h 为该匹配点对对应的内点置信度; $\mathbf{y}_h^*$ 为源关键点 $\mathbf{x}_h$ 对应的目标匹配点。

随后,利用该初始变换,对源三维点云的关键点进行精配准,在坐标空间中重新构建局部匹配关系,并结合更新后的几何一致性重新评估匹配点的置信度,从而获得更加可靠的匹配权重。通过这种置信度递进优化策略,使匹配点权重能够随配准状态不断优化,进一步提高最终配准精度。

1.5 损失函数

在网络的训练阶段,AMS-Net 采用端到端的有监督学习方式,模型利用数据集提供的真实位姿判定候选匹配点对的真实归属,并以此构建损失函数。为了有效优化三维点云的配准精度,AMS-Net 设计了一种综合考虑特征匹配一致性、几何一致性和外点剔除的损失函数。所有损失的计算均基于初筛出的候选匹配点对集合 C_0 。

特征匹配损失函数用于衡量源三维点云与目标三维点云在特征空间中的相似性。给定候选匹配点对 $(\mathbf{x}_i, \mathbf{y}_i^*)$, 其特征描述符分别为 $\mathbf{d}(\mathbf{x}_i)$ 和 $\mathbf{d}(\mathbf{y}_i^*)$, 特征匹配损失函数 L_{feature} 的表达式为

$$L_{\text{feature}} = \frac{1}{N_1} \sum_{i=1}^{N_1} I_i^{\text{gt}} \|\mathbf{d}(\mathbf{x}_i) - \mathbf{d}(\mathbf{y}_i^*)\|_2^2 \quad (15)$$

式中: N_1 为匹配点对的数量; $I_i^{\text{gt}} \in \{0, 1\}$ 为第 i 个匹

配点对的真实监督标签,表示其是否为内点。

几何一致性损失函数的作用是保证匹配对在刚体变换下的几何关系的一致性。对于候选匹配集合中的真实点对,其与各自局部邻域点之间应当保持稳定的相对欧氏距离。几何一致性损失函数 L_{geometry} 的表达式为

$$L_{\text{geometry}} = \frac{1}{N_1} \sum_{i=1}^{N_1} I_i^{\text{gt}} \| \mathbf{R}_{\text{gt}} \mathbf{x}_i + \mathbf{t}_{\text{gt}} - \mathbf{y}_i^* \|_2^2 \quad (16)$$

式中: $\mathbf{R}_{\text{gt}}$ 和 $\mathbf{t}_{\text{gt}}$ 分别为源点云与目标点云之间的真实旋转矩阵和平移向量。

外点剔除损失函数利用真实标签对网络预测的匹配置信度进行监督。通过迫使网络对真实内点输出高置信度、对外点输出低置信度,从而在后续优化过程中自适应降低外点权重,减少误匹配对最终配准结果的影响。外点剔除损失函数 L_{outlier} 采用二元交叉熵形式,表达式为

$$L_{\text{outlier}} = -\frac{1}{N_1} \sum_{i=1}^{N_1} [I_i^{\text{gt}} \ln \omega_i + (1 - I_i^{\text{gt}}) \ln(1 - \omega_i)] \quad (17)$$

式中: ω_i 为网络预测输出的第 i 个候选匹配点对的内点置信度。

通过加权外点剔除损失函数,模型能够有效剔除误匹配点对,从而优化三维点云的配准结果。

最终,在总损失函数 L_{total} 中,为特征匹配损失函数、几何一致性损失函数、外点剔除损失函数分别赋予权重系数 λ_1 、 λ_2 、 λ_3 , 表达式为

$$L_{\text{total}} = \lambda_1 L_{\text{feature}} + \lambda_2 L_{\text{geometry}} + \lambda_3 L_{\text{outlier}} \quad (18)$$

其中,本文以特征匹配损失函数为优化基础,将其权重 λ_1 设为 1.0。为兼顾刚体变换约束与外点抑制效果,通过 KITTI 数据集上的预实验与消融对比,设 λ_2 为 0.5, λ_3 为 0.2。

2 实验结果及分析

2.1 实验环境与参数设置

本文实验所使用的深度学习框架为 PyTorch 1.8.1,运行环境为基于 Python 3.8 和 Ubuntu 18.04 的系统,并且配置 CUDA 11.1。硬件方面,采用 12 个虚拟 CPU 的 Intel(R) Xeon(R) Silver 4214R@2.40 GHz,配备 90 GB 系统内存,且使用单张 NVIDIA RTX 3080 Ti 显卡,具有 12 GB 显存。在优化方面,选择了 Adam 优化器,并设置初始学习率为 1×10^{-4} ,批次量为 2,训练轮次数设为 50,同时使用的权重衰减为 1×10^{-4} 。

实验所用的数据集为 KITTI 数据集^[21]。使用

序列 0~5 进行训练,序列 6~7 进行验证,序列 8~10 进行测试。在 KITTI 数据集上,由于该数据集主要来源于真实道路环境中的车载激光雷达扫描,三维点云具有尺度大、稀疏性强以及运动变化明显等特点,因此,本文采用的评价指标为平移误差(RTE)、旋转误差(RRE)和配准召回率(RR)。为提高位姿标注精度,真值位姿先通过 ICP 进行精化,并仅保留帧间距离不小于 10 m 的样本对用于评估。

平移误差 R_{RET} 是用于衡量估计平移向量与真实平移向量之间的差异,通常定义为两者的欧氏距离,表达式为

$$R_{\text{RTE}} = \|\hat{\mathbf{t}} - \mathbf{t}\|_2 \quad (19)$$

式中: $\hat{\mathbf{t}}$ 为估计平移向量; $\mathbf{t}$ 为真实平移向量。

相对旋转误差 R_{RRE} 用于衡量估计旋转矩阵与真实旋转矩阵之间的偏差,其表达式为

$$R_{\text{RRE}} = \arccos\left(\frac{\text{trace}(\hat{\mathbf{R}}^T \mathbf{R}) - 1}{2}\right) \quad (20)$$

式中: $\text{trace}(\cdot)$ 为矩阵迹运算; $\hat{\mathbf{R}}$ 为估计旋转矩阵, $\mathbf{R}$ 为真实旋转矩阵。

配准召回率 R_{RR} 定义为同时满足平移误差和旋转误差阈值约束的样本占全部测试样本的比例。假设测试集中共有 M 对三维点云,平移误差阈值为 τ_{tr} ,旋转误差阈值为 τ_{rot} ,则 R_{RR} 表达式为

$$R_{\text{RR}} = \frac{1}{M} \sum_{i=1}^M 1(R_{\text{RTE}_i} < \tau_{\text{tr}}, R_{\text{RRE}_i} < \tau_{\text{rot}}) \quad (21)$$

式中: $1(\cdot)$ 为指示函数,当括号内条件成立时取值为 1,否则取值为 0; R_{RTE_i} 、 R_{RRE_i} 分别为第 i 组样本的相对平移误差与相对旋转误差。

2.2 消融实验及分析

为了验证 AMS-Net 模型中所引入模块的贡献,在 KITTI 数据集上进行了 AMS-Net 的消融实验。实验结果如表 1 所示。本实验在原始的全局尺度模型(BM)上进行粗配准,依次融合自适应多尺度特征提取模块(AMS)、自适应邻域稳定器(ANS)、基于几何一致性的置信度校准模块(GCC)和基于置信度递进的精配准策略(CPFR)。

由表 1 可知,在 BM 的基础上,随着 AMS、ANS、GCC 和 CPFR 模块的逐步引入,AMS-Net 在 KITTI 数据集上的配准性能呈现出稳定且持续的提升。在特征模糊、几何信息较少的区域的情景中,基线模型 BM 的 R_{RTE} 、 R_{RRE} 和 R_{RR} 分别为 0.088 2 m、0.55°和 89.4%,这是由于单一尺度不足以提供充足的特征信息。当加入所提出的 AMS 多尺度结构

后,能够提供更具有判别力的信息。而再加入 ANS 则使固定采样和固定邻域带来的局部结构退化问题得到缓解,这也是 R_{RR} 提升最为明显的重要原因。随

后,进一步引入 GCC,错误匹配被更进一步抑制。而最后在加入 CPFR 后,模型获得了最优结果, R_{RTE} 降至 0.061 8 m, R_{RRE} 降至 0.21°, R_{RR} 提升至 99.8%。

表 1 KITTI 数据集的 AMS-Net 消融实验

Table 1 Ablation experiments of AMS-Net on the KITTI dataset

是否加入模块					R_{RTE}/m	$R_{RRE}/(^{\circ})$	$R_{RR}/\%$
BM	AMS	ANS	GCC	CPFR			
是	否	否	否	否	0.088 2	0.55	89.4
是	是	否	否	否	0.073 4	0.42	91.8
是	是	是	否	否	0.070 7	0.36	96.4
是	是	是	是	否	0.068 4	0.28	98.3
是	是	是	是	是	0.061 8	0.21	99.8

为了进一步验证自适应多尺度特征提取模块中,网络深度对最终点云配准性能的具体影响,本文设计了针对网络层级的消融实验。在保持其他网络结构与配置完全一致的前提下,实验分别测试了网络层级从 1 级逐步递增至 4 级时模型的表现变化。具体的量化评估指标如表 2 所示。

表 2 网络层级对配准性能的影响

Table 2 Impact of network hierarchy on registration performance

网络层级	R_{RTE}/m	$R_{RRE}/(^{\circ})$	$R_{RR}/\%$
1 级	0.065 9	0.28	98.4
2 级	0.063 2	0.24	99.4
3 级	0.062 0	0.21	99.7
4 级	0.061 8	0.21	99.8

由表 2 可知,网络层级由 1 级增至 4 级时,配准性能稳步提升。 R_{RTE} 由 0.065 9 m 降至 0.061 8 m, R_{RRE} 由 0.28° 降至 0.21°, R_{RR} 由 98.4% 提升至 99.8%。1 级提升至 3 级网络的指标改善幅度明显,这说明通过增加网络层级提取多尺度特征,能有效提升配准精度与匹配召回率。而 3 级至 4 级的网络性能提升幅度趋于平缓,原因在于下采样点数的持续减少,可提取的三维点云特征趋于匮乏。综合权衡配准精度与模型复杂度,4 级网络取得最优实验性能,从而验证了 AMS-Net 网络层级设计的有效性与合理性。

2.3 对比实验及分析

为了验证所提方法的有效性,AMS-Net 在大规模场景的 KITTI 数据上进行配准,并与 FCGF^[22]、CoFiNet^[23]、PREDATOR^[24]、PointDSC^[17]、GeoTransformer^[19]、TurboReg^[25] 以及 FLAT^[11] 多

种经典及先进的方法进行对比。实验结果如表 3 所示。

表 3 KITTI 数据集的对比实验

Table 3 Comparative experiments of the KITTI dataset

方法	R_{RTE}/m	$R_{RRE}/(^{\circ})$	$R_{RR}/\%$
FCGF	0.095 4	0.25	96.6
CoFiNet	0.082 4	0.41	99.8
PREDATOR	0.068 7	0.27	99.8
PointDSC	0.083 5	0.38	98.2
GeoTransformer	0.062 0	0.23	99.8
TurboReg	0.079 1	0.39	98.6
FLAT	0.069 0	0.24	99.8
本文方法	0.061 8	0.21	99.8

由表 3 可知,AMS-Net 方法的 R_{RTE} 为 0.061 8 m, R_{RRE} 为 0.21°, R_{RR} 达到 99.8%,综合指标最优。与等召回率同高的算法相比,本文方法的平移、旋转误差均更低,其余基线方法或位姿误差偏高,或召回率不足,仅依靠高匹配召回难以保证精准配准。这表明本文所设计的自适应多尺度特征架构可在维持高匹配召回率的同时,有效削减位姿估计偏差,兼顾匹配完备性与配准精度。

为进一步验证本文所提网络在真实车载场景的配准性能与泛化能力,本文将 nuScenes^[26] 数据集作为测试样本,对三维点云帧间距离分别为 10、20 和 30 m 的数据进行配准。该数据集采用 32 线激光雷达采集,传感器规格、路况场景与原 KITTI 数据集存在明显差异,实验结果如表 4 所示。相较于 KITTI 数据集的 64 线激光雷达,nuScenes 数据集使用 32 线激光雷达采集数据,点云密度更低,且城市道路对模型的泛化能力提出了更高要求。实验结果表明:随着

帧间跨度的增大,点云重叠率和密度急剧下降,从而导致所有基准算法的性能均出现不同程度的衰退。在10 m 的距离下,本文方法的 R_{RTE} 为 0.32 m、 R_{RRE} 为 0.48° 、 R_{RR} 为 90.1%。在 20 m 的距离下,本文方法的 R_{RTE} 为 0.48 m、 R_{RRE} 为 0.75° 、 R_{RR} 为 83.3%。当跨度

增加至 30 m 的极端稀疏场景时,依然保持了 R_{RR} 为 78.8%的高配准率。可见,本文方法在 3 种帧间距离下均取得了较优的配准效果,由此证明该模型能够有效适配不同规格的车载激光雷达数据,在异构三维点云场景下依然保持稳定的配准精度。

表 4 nuScenes 数据集的对比实验

Table 4 Comparative experiments on the nuScenes dataset

方法	10 m			20 m			30 m		
	R_{RTE}/m	$R_{RRE}/(^{\circ})$	$R_{RR}/\%$	R_{RTE}/m	$R_{RRE}/(^{\circ})$	$R_{RR}/\%$	R_{RTE}/m	$R_{RRE}/(^{\circ})$	$R_{RR}/\%$
FCGF	0.35	0.51	77.5	0.75	1.85	62.8	0.78	2.04	50.7
CoFiNet	0.49	0.84	57.3	0.64	1.04	53.0	0.60	1.68	66.2
PREDATOR	0.35	0.57	79.5	0.57	1.88	50.4	0.67	2.16	56.4
PointDSC	0.40	0.67	75.7	0.74	1.76	70.2	0.86	1.92	59.4
GeoTransformer	0.35	0.51	89.9	0.49	0.73	76.7	0.50	1.09	69.0
TurboReg	0.46	0.78	89.1	0.84	0.97	83.6	0.79	1.26	68.2
FLAT	0.47	0.94	80.3	0.57	1.69	66.5	0.76	1.88	58.4
本文方法	0.32	0.48	90.1	0.48	0.75	83.3	0.67	1.03	78.8

2.4 可视化对比及分析

为进一步直观展示在真实三维点云场景下的配准效果,给出不同方法在 KITTI 和 nuScenes 数据集上配准结果的细节差异。

图 6 为不同方法在 KITTI 数据集上的对比实验配准结果可视化,图中从上至下依次展示了 3 个不同路况场景下的配准效果。现有对比算法虽可初步恢复帧间点云空间的对应关系,但均存在如红色框标注的结构偏移与紫色框标注的局部错位问题。尤其在场景边缘、稀疏区域及几何复杂区域,源三维点云与目标三维点云重合度较差,这表明现有方法在复杂干扰场景下仍存在显著误匹配缺陷。相比之下,针对三维点云分布不均、细长结构及低重叠比例等难点场景,本文方法仍可保持稳定的配准一致性,

在多组可视化样例中均展现出更优的配准效果,帧间点云重叠度更高,主体结构边缘对齐更规整,局部细节匹配精度更佳,从而有效验证了该方法在部分重叠点云配准任务中的有效性与优越性。

图 7 为不同方法在 nuScenes 数据集上的对比实验配准结果可视化,图中从上至下依次展示了帧间距离分别为 10、20 和 30 m 的条件下的配准效果。在间隔 10 m 的场景下,对比方法均存在紫色框标注的局部错位问题,但随着帧间距离的增大,对比方法出现了明显的对齐偏差,红色框标注的结构偏移问题显著增多。而本文算法在所有距离下均实现了源三维点云与目标三维点云的高度对齐,道路边缘、建筑轮廓等关键结构无明显偏差,从而验证了其配准精度与鲁棒性。

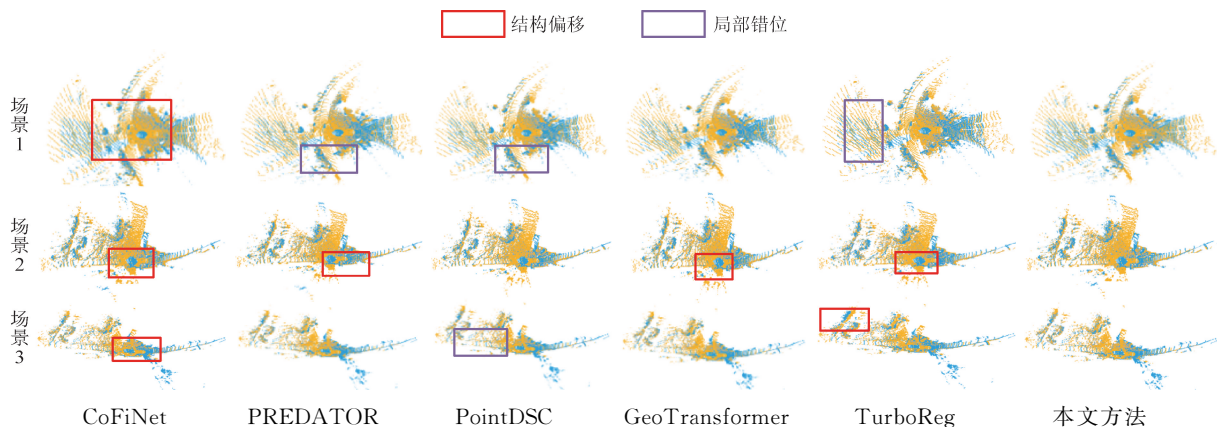


图 6 KITTI 数据集上的配准结果对比

Fig. 6 Comparison of registration results on the KITTI dataset

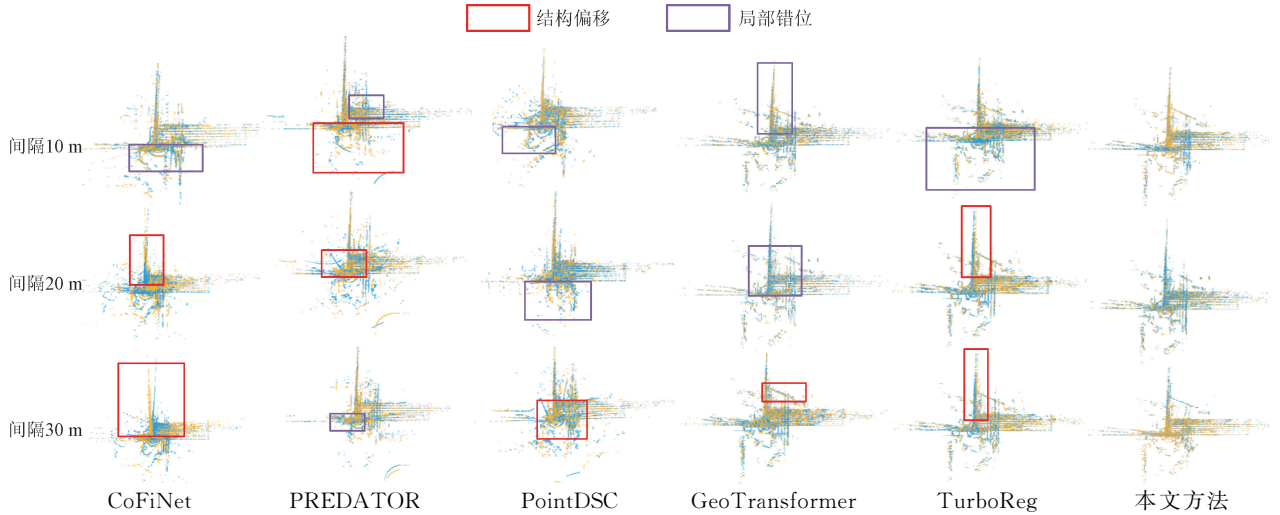


图 7 nuScenes 数据集上的配准结果对比

Fig. 7 Comparison of registration results on the nuScenes dataset

3 结 论

本文面向自动驾驶激光雷达点云配准任务,针对复杂动态道路场景中,三维点云稀疏、特征模糊以及帧间重叠度低导致关键点误匹配的问题,提出了面向智能驾驶的自适应多尺度三维点云配准网络。该网络通过多尺度特征匹配一致性剔除外点,利用自适应多尺度特征提取与邻域稳定器降低特征模糊场景下的误匹配,通过融合特征与几何约束,双重筛选可靠内点,依据内点置信度由粗到精的迭代完成高精度配准。在 KITTI 数据集上,配准误差 R_{RTE} 为 0.061 8 m, R_{RRE} 为 0.21° , R_{RR} 为 99.8%。在 nuScenes 数据集的跨大帧间距场景下,间隔 10、20 和 30 m 的 R_{RTE} 分别为 0.32 m、0.48 m 和 0.67 m, R_{RRE} 分别为 0.48° 、 0.75° 和 1.03° , R_{RR} 分别稳定至 90.1%、83.3% 和 78.8%。实验结果证明,所提三维点云配准网络 AMS-Net 能够有效挖掘点云多尺度特征并抑制帧间误匹配,在真实车载异构点云环境中具备优异的配准精度与场景泛化鲁棒性。

参考文献:

- [1] Yu Zhao, Zhang Peize, Shi Jing. Transformation of industrial robotics with natural language models: recent progress and future prospects [J]. Robotics and Computer-Integrated Manufacturing, 2026, 97: 103113.
- [2] 周翔, 宋林桦, 夏海峰, 等. 个性化热舒适系统在电动汽车中的应用与展望 [J]. 西安交通大学学报, 2026, 60(1): 1-13.

Zhou Xiang, Song Linye, Xia Haifeng, et al. Application and prospects of personalized thermal comfort systems in electric vehicles [J]. Journal of Xi'an Jiaotong University, 2026, 60(1): 1-13.

- [3] Song C H, Blukis V, Tremblay J, et al. Robospacial: teaching spatial understanding to 2D and 3D vision-language models for robotics [C]//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2025: 15768-15780.
- [4] Wang Zhigang, Su Yifei, Li Chenhui, et al. Open-vocabulary octree-graph for 3D scene understanding [C]//2025 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2025: 7037-7047.
- [5] 马庆禄, 白锋, 张杰, 等. 基于激光雷达点云地图的车辆定位与导航 [J]. 光学精密工程, 2024, 32(16): 2537-2549.
Ma Qinglu, Bai Feng, Zhang Jie, et al. Vehicle localization and navigation method based on LiDAR point cloud map [J]. Optics and Precision Engineering, 2024, 32(16): 2537-2549.
- [6] Afzal Maken F, Muthu S, Nguyen C, et al. Improving 3D reconstruction through RGB-D sensor noise modeling [J]. Sensors, 2025, 25(3): 950.
- [7] Tian Yifei, Li Xiangyu, Yin Jieming. From pseudo-to non-correspondences: robust point cloud registration via thickness-guided self-correction [J]. Computers & Graphics, 2026, 134: 104528.
- [8] Ma Kaifeng, Yan Fengtao, Li Shiming, et al. Low-overlap registration of multi-source LiDAR point

- clouds in urban scenes through dual-stage feature pruning and progressive hierarchical methods [J]. *Remote Sensing*, 2025, 17(17): 2938.
- [9] Kashef Tabrizian S, Terryn S, Vanderborght B. Toward autonomous self-healing in soft robotics: a review and perspective for future research [J]. *Advanced Intelligent Systems*, 2025, 7(8): 2400790.
- [10] 王丞, 田暄, 郭瑞, 等. 自适应 Harris 角点提取的点云粗配准算法 [J]. *西安交通大学学报*, 2022, 56(3): 33-44.
- Wang Cheng, Tian Xuan, Guo Rui, et al. Coarse point cloud registration based on adaptive Harris corner extraction [J]. *Journal of Xi'an Jiaotong University*, 2022, 56(3): 33-44.
- [11] Wang Weijie, Mei Guofeng, Zhang Jian, et al. Fully-geometric cross-attention for point cloud registration [C]//2025 International Conference on 3D Vision (3DV). Piscataway, NJ, USA: IEEE, 2025: 347-356.
- [12] Xia Yaqi, Xia Yan, Li Wei, et al. ASFM-Net: asymmetrical Siamese feature matching network for point completion [C]//Proceedings of the 29th ACM International Conference on Multimedia. New York, NY, USA: ACM, 2021: 1938-1947.
- [13] Besl P J, McKay N D. A method for registration of 3-D shapes [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 239-256.
- [14] Wang Yue, Solomon J. Deep closest point: learning representations for point cloud registration [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2019: 3522-3531.
- [15] Yew Z J, Lee G H. RPM-Net: robust point matching using learned features [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 11821-11830.
- [16] Fu Kexue, Liu Shaolei, Luo Xiaoyuan, et al. Robust point cloud registration framework based on deep graph matching [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 8889-8898.
- [17] Bai Xuyang, Luo Zixin, Zhou Lei, et al. PointDSC: robust point cloud registration using deep spatial consistency [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 15854-15864.
- [18] Yew Z J, Lee G H. REGTR: end-to-end point cloud correspondences with transformers [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 6667-6676.
- [19] Qin Zheng, Yu Hao, Wang C, et al. Geometric transformer for fast and robust point cloud registration [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 11133-11142.
- [20] Qi C R, Yi Li, Su Hao, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 5105-5114.
- [21] Geiger A. Are we ready for autonomous driving? The KITTI vision benchmark suite [C]//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2012: 3354-3361.
- [22] Choy C, Park J, Koltun V, et al. Fully convolutional geometric features [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2019: 8957-8965.
- [23] Yu H, Li F, Saleh M, et al. CoFiNet: reliable coarse-to-fine correspondences for robust point cloud registration [J]. *Advances in Neural Information Processing Systems*, 2021, 34: 23872-23884.
- [24] Huang Shengyu, Gojcic Z, Usvyatsov M, et al. PREDATOR: registration of 3D point clouds with low overlap [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 4265-4274.
- [25] Yan Shaocheng, Shi Pengcheng, Zhao Zhenjun, et al. Turboreg: turboclique for robust and efficient point cloud registration [C]//2025 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2025: 26371-26381.
- [26] 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 (CVPR). Piscataway, NJ, USA: IEEE, 2020: 11618-11628.

(编辑 刘懿莹)