

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-12

论文引用格式: Zhang Fengjun. Rotation-Robust Multi-scale Neural Laplacian Operator for Point Clouds [J/OL]. Journal of Image and Graphics, XXXX:1-12. DOI: 10.11834/jig.260057. (张逢骏. 旋转鲁棒的多尺度点云神经拉普拉斯算子[J/OL]. 中国图象图形学报, XXXX:1-12. DOI: 10.11834/jig.260057.) [DOI:10.11834/jig.260057]

旋转鲁棒的多尺度点云神经拉普拉斯算子

张逢骏

江苏科技大学, 江苏省镇江市 212000

摘要: 目的 离散拉普拉斯算子是三维几何处理领域的核心微分算子, 广泛应用于平滑、变形及参数化等任务。由于点云数据缺乏规范的拓扑连接, 传统基于局部三角剖分的方法在处理稀疏或含噪数据时鲁棒性较差。现有的深度学习方法(如 NeLo)虽然能通过数据驱动在 K 近邻图上学习算子权重, 但其输入依赖于点对的相对几何坐标, 导致模型对点云的全局姿态高度敏感, 不具备旋转不变性。此外, 单一尺度的特征聚合难以兼顾局部几何细节保留与噪声抑制。针对上述问题, 提出一种融合局部参考系与多尺度注意力机制的旋转鲁棒点云拉普拉斯算子学习网络。**方法** 构建了一个端到端的深度学习框架。首先, 为了从输入端消除全局姿态干扰, 提出基于协方差分析的局部参考系(LRF)对齐模块, 将每个点的局部邻域投影至旋转不变的规范空间。其次, 设计双通道多尺度特征提取骨干网络, 并行提取不同半径邻域图上的几何特征, 并引入通道注意力机制(SE-Block)自适应地融合粗粒度与细粒度特征, 在抑制噪声的同时增强对尖锐边角特征的感知。最后, 在训练阶段引入旋转一致性损失函数, 显式约束网络在不同姿态下的预测一致性。**结果** 在公开的 ShapeNet 数据集及 ScanNet 真实扫描数据集上的对比实验表明, 该方法生成的拉普拉斯算子具有显著的旋转鲁棒性和泛化能力。具体而言, 在施加随机刚体旋转的 ShapeNet 测试集上, 相比于基准模型 NeLo, 本文方法的均方误差(MSE)降低了 27.6%, F 值(F-measure)提升了 3.1%; 同时, 热扩散与平滑在极具挑战的 ScanNet 真实扫描点云任务中, 该方法能更准确地保留模型的几何细节, 边缘保持能力显著优于传统几何方法。数据集 DOI: 10.57760/sciencedb.31651。**结论** 本文提出的融合局部参考系与多尺度注意力机制的方法, 有效解决了现有神经拉普拉斯算子对姿态敏感及细节丢失的问题, 为非受控场景下的点云几何处理提供了可靠的底层算子基础。离散拉普拉斯算子是三维几何处理领域的核心微分算子, 广泛应用于平滑、变形及参数化等任务。由于点云数据缺乏规范的拓扑连接, 传统基于局部三角剖分的方法在处理稀疏或含噪数据时鲁棒性较差。现有的深度学习方法(如 NeLo)虽然能通过数据驱动在 K 近邻图上学习算子权重, 但其输入依赖于点对的相对几何坐标, 导致模型对点云的全局姿态高度敏感, 不具备旋转不变性。此外, 单一尺度的特征聚合难以兼顾局部几何细节保留与噪声抑制。针对上述问题, 提出一种融合局部参考系与多尺度注意力机制的旋转鲁棒点云拉普拉斯算子学习网络。

关键词: 点云处理; 拉普拉斯算子; 图神经网络; 局部参考系; 注意力机制; 旋转不变性

Rotation-Robust Multi-scale Neural Laplacian Operator for Point Clouds

Zhang Fengjun

Jiangsu University of Science and Technology, Zhenjiang 212000, China

Abstract: Objective The discrete Laplacian operator is a cornerstone in the field of 3D geometry processing and computer

收稿日期: 2026-01-26; 修回日期: 2026-04-01

基金项目: 江苏科技大学高水平人才科研启动基金项目(项目编号: 1132932301, 1022932208)

Supported by: The research start-up funding for high-level talent of Jiangsu University of Science and Technology (Grant No. 1132932301 and Grant No. 1022932208).

graphics, serving as the fundamental differential operator for a wide array of algorithms including spectral shape analysis, geometry compression, mesh smoothing, parameterization, and physics-based simulation. In the context of manifold triangular meshes, the Laplacian operator (often discretized via the cotangent formula) is well-defined and theoretically rigorous. However, with the rapid development of 3D acquisition technologies such as LiDAR and RGB-D cameras, point clouds have become a ubiquitous representation for 3D geometry. Unlike meshes, point clouds lack explicit topological connectivity and often exhibit irregular sampling densities, making the definition of a robust and accurate Laplacian operator a mathematically ill-posed and challenging problem. Traditional approaches typically attempt to approximate the underlying manifold by constructing local triangulations (e. g. , Delaunay triangulation) or estimating tangent planes. While effective on high-quality data, these geometric methods are inherently fragile and computationally expensive when processing sparse, noisy, or non-uniform data common in real-world scans. Recently, the Neural Laplacian Operator (NeLo) has emerged as a promising data-driven alternative, utilizing Graph Neural Networks (GNNs) to learn edge weights on K-nearest neighbor (KNN) graphs. By imitating the behavior of the ground-truth Laplacian on a set of probe functions, NeLo achieves high accuracy on clean data. However, NeLo and its variants suffer from two critical limitations that hinder their practical application in "in-the-wild" scenarios. First, they typically rely on relative Cartesian coordinates as the primary input feature for the neural network. This design implicitly entangles the learned geometric descriptors with the global pose of the object, rendering the operator highly sensitive to rigid transformations. Consequently, the performance degrades significantly when the input point cloud is arbitrarily rotated, which is a common occurrence in robotic perception and autonomous driving. Second, existing architectures predominantly employ single-scale feature aggregation mechanisms. This design struggles to resolve the inherent trade-off in differential operator estimation: suppressing noise in flat regions requires a large receptive field, while preserving high-frequency details in sharp feature areas requires precise local information. A single scale often leads to either over-smoothing of sharp edges or under-smoothing of noise. To address these limitations, this paper proposes a novel Rotation-Robust Neural Laplacian Operator (RR-NeLo), a geometry-aware and pose-invariant framework designed to predict accurate, intrinsic Laplacian matrices for unoriented and noisy point clouds.

Method The proposed RR-NeLo establishes an end-to-end learning framework comprising three synergistic components: a Local Reference Frame (LRF) alignment module, a dual-channel multi-scale backbone network, and a rotation-consistent training scheme. First, to fundamentally eliminate the interference of global pose variation, we introduce a robust LRF alignment module. Instead of feeding raw coordinates directly into the network, we compute a weighted covariance matrix for the local neighborhood of each point. By performing Eigen-decomposition on this covariance matrix, we derive a stable, content-based local coordinate system (canonical frame) that aligns with the principal geometric directions of the local surface. All neighboring points are then projected into this rotation-invariant canonical space. This transformation ensures that the input features to the subsequent GNN are intrinsic to the shape and invariant to any global rotation or translation, thereby guaranteeing the $SE(3)$ -invariance of the learned operator. Second, to address the scale dilemma, we design a dual-channel multi-scale backbone network. Unlike previous methods that operate on a fixed KNN graph, our network constructs two parallel graph branches with varying neighborhood radii: a *local detail branch* (K_{small}) focused on capturing fine geometric structures like corners and edges, and a *global context branch* (K_{large}) focused on integrating broader neighborhood information for robust noise suppression. The features extracted from these two branches are fused via a Squeeze-and-Excitation (SE-Block) channel attention mechanism. The SE-Block adaptively recalibrates the channel-wise feature responses, assigning higher weights to fine-grained features in high-curvature regions and higher weights to coarse-grained features in planar regions. This mechanism allows the network to dynamically balance the trade-off between detail preservation and smoothing based on local geometric complexity. Finally, the network is trained using a hybrid loss function. In addition to the standard supervised loss that minimizes the discrepancy between the predicted Laplacian and the ground-truth mesh Laplacian on a set of spectral and spatial probe functions, we incorporate a novel rotation consistency loss. This regularization term explicitly enforces the network to produce identical Laplacian matrix values (up to permutation) for a shape and its randomly rotated counterparts. This constraint further regularizes the solution space and improves the generalization ability of the model on unseen orientations. Dataset DOI: 10. 57760/sciencedb. 31651. **Result** To validate the effectiveness of the proposed method, extensive comparative experiments were conducted on the synthetic ShapeNet dataset

(comprising 12k shapes across 17 categories) and real-world noisy scans from the Kinect v2 dataset. We benchmarked RR-NeLo against representative traditional methods, including the Graph Laplacian, Heat Method, and NManifold, as well as the baseline deep learning model NeLo. Quantitative evaluations demonstrate that RR-NeLo significantly outperforms existing state-of-the-art techniques. More importantly, on the randomly rotated ShapeNet test set, where the baseline NeLo suffers a severe performance drop due to pose variation, RR-NeLo maintains high accuracy, reducing the Mean Squared Error (MSE) by 27.6% and improving the F-measure (a metric for structural fidelity) by 3.1% compared to NeLo. This result confirms the effectiveness of the LRF module in achieving robust rotation invariance. Furthermore, qualitative and quantitative evaluations on downstream geometry processing tasks—including heat diffusion, Laplacian smoothing, and geodesic distance computation—reveal that RR-NeLo effectively removes noise while preserving sharp features. The experiments on real-world Kinect v2 scans further demonstrate that RR-NeLo generalizes well to uncontrolled data with severe noise and non-uniform sparsity, generating plausible Laplacian matrices without requiring explicit mesh reconstruction. **Conclusion** This paper presents RR-NeLo, a robust neural framework for learning the discrete Laplacian operator on point clouds. By integrating Local Reference Frame alignment, the proposed method successfully disentangles geometric feature learning from global pose, solving the rotation sensitivity issue inherent in previous neural operators. The introduction of the dual-channel multi-scale attention mechanism further enhances the operator's ability to preserve high-frequency geometric details while suppressing noise. Experimental results confirm that RR-NeLo sets a new state-of-the-art in terms of accuracy, robustness, and generalization capability. This work provides a reliable mathematical tool for differential geometry processing on unoriented, non-uniform point clouds, bridging the gap between deep learning and classical geometry processing. Future work will explore extending this framework to learn other differential operators, such as the gradient and divergence, to support a wider range of physical simulations on point-based representations.

Key words: point cloud processing; Laplacian operator; graph neural network; local reference frame; attention mechanism; rotation invariance

0 引言

拉普拉斯算子(Laplacian Operator)作为几何处理领域的“瑞士军刀”,在三维形状分析、几何平滑、参数化及形状变形等任务中扮演着核心角色(Botsch等,2010)。随着三维扫描技术的普及,点云已成为最主要的三维数据载体。然而,由于点云数据离散、无序且缺乏规范的流形结构,如何在原始点云上定义高精度、鲁棒的离散拉普拉斯算子,一直是计算机图形学领域的一个开放性难题(Sharp和Crane,2020)。

早期的点云拉普拉斯算子构建方法主要遵循“重建—逼近”的思路。一类方法试图通过构建局部三角化网格(Local Triangulation)来近似流形结构(Belkin等,2009);另一类方法则依赖于切平面估计(Tangent Plane Estimation)来计算微分算子(Liu等,2012)。然而,这些传统几何方法对数据质量要求极高,在处理含有噪声、孔洞、离群点或薄壁结构的非理想点云时,往往因无法构建高质量的局部网格结构而导致算子精度急剧下降甚至计算失效(Pang

等,2024)。

近年来,深度学习技术尤其是基于点云变换器(Point Transformer)的架构(Zhao等,2021),为该问题提供了新的解决范式。(Pang等,2024)提出的神经拉普拉斯算子(NeLo)开创性地利用图神经网络(GNN)在简单的K近邻图(KNN Graph)上直接学习拉普拉斯矩阵的边权重。通过模仿真实网格算子在探测函数上的行为,NeLo在规则姿态的点云上取得了超越传统几何方法的精度。尽管NeLo取得了显著进展,但其网络架构设计存在先天的几何缺陷:该方法直接使用点对的相对坐标向量作为GNN的输入信号。这种设计使得模型对点云的整体姿态高度敏感,即不具备旋转不变性(Rotation Invariance)。在实际应用场景(如自动驾驶、机器人抓取)中,传感器获取的点云姿态往往是随机且未对齐的,NeLo在此类数据上的性能会因坐标系变化而急剧恶化。

针对点云网络的旋转鲁棒性问题,最新的研究进展主要分为几何旋转不变、特征旋转不变和训练旋转不变三类(Wang等,2025)。特征旋转不变方法如张量场网络(Thomas等,2018)或向量神经网络(Deng等,2021),虽然理论完备但计算开销巨大,难

以满足拉普拉斯算子对高效性的需求。训练旋转不变方法主要依赖数据增强,但难以覆盖连续无限的旋转空间,且在未见过的分布上泛化能力有限 (Fei 和 Deng, 2024)。相比之下,基于几何属性的几何旋转不变方法具有可解释性强、计算高效的优势。其中,基于局部参考系 (Local Reference Frame, LRF) 的方法通过构建局部规范坐标系来隔离姿态影响,在点云分类和分割任务中表现优异 (Kim 等, 2020; Zhang 等, 2025)。然而,传统的 LRF 构建方法 (如 PCA 方法) 容易受到局部噪声和采样密度的干扰,且存在固有的符号二义性问题 (Sun 等, 2022), 导致参考系轴向不稳定。单一尺度的邻域构建难以兼顾稀疏区域的连通性和密集区域的细节特征。

鉴于此,本文旨在解决 NeLo 模型在非规范姿态下的失效问题,提出一种融合局部参考帧与多尺度特征的旋转鲁棒神经拉普拉斯算子。本文并未单纯依赖几何预处理,而是构建了一个混合增强框架:首先,引入基于协方差分析的局部参考帧 (LRF) 机制,将邻域几何特征投影至局部不变空间,从输入端彻底阻断旋转干扰;其次,针对 LRF 的不稳定性及单一尺度缺陷,设计多尺度特征融合模块,并行提取不同感受野下的几何特征以增强对非均匀采样的适应能力;最后,借鉴训练旋转不变的思想,提出旋转

一致性损失 (Rotation Consistency Loss), 显式约束网络在不同姿态下输出一致的算子预测结果,进一步提升模型的鲁棒性。

1 模型方法

1.1 整体网络架构

为了解决现有神经拉普拉斯算子对点云旋转敏感的问题,本文提出了一种旋转鲁棒的神经拉普拉斯算子计算框架。该框架旨在从非结构化的三维点云数据中直接学习并回归出高质量的拉普拉斯矩阵权重,同时保证在输入点云发生任意旋转时,输出的算子保持不变。如图 1 所示,RR-NeLo 框架主要由三个核心模块紧密耦合而成:局部规范化模块、多尺度注意力特征提取模块以及算子回归模块。

1) 局部规范化模块:

作为网络的输入前端,该模块的核心作用是解耦几何特征与空间姿态。针对每个查询点及其邻域,模块首先基于协方差分析构建一个随物体旋转而旋转的局部参考系。随后,将原始的世界坐标转换至该规范化的局部坐标系中。这一过程有效地消除了输入点云刚性变换对后续特征学习的干扰,确保了网络输入具有严格的旋转不变性。

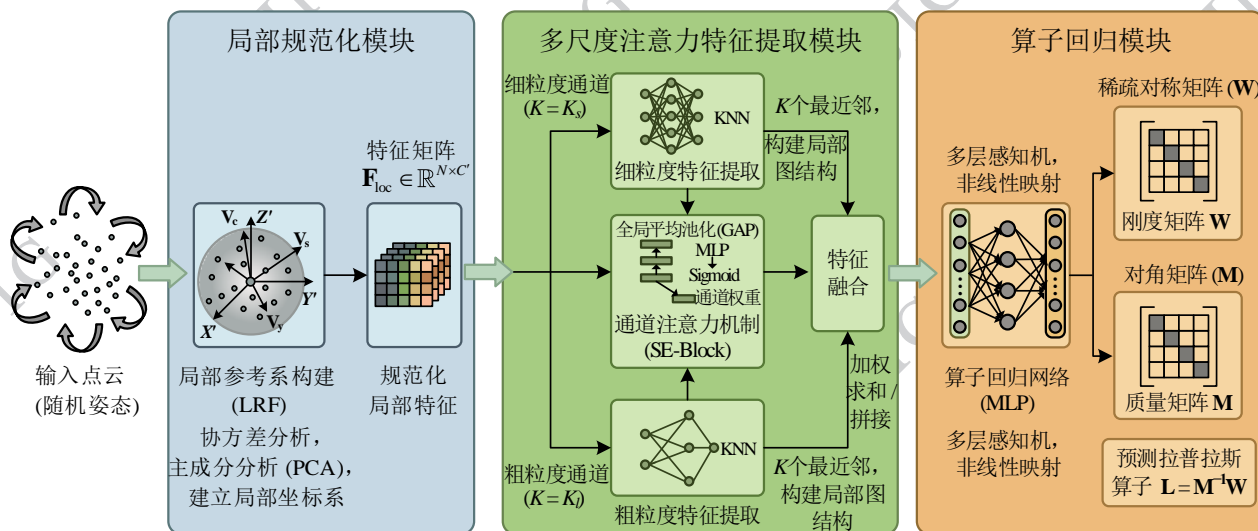


Fig. Network Architecture Integrating Local Reference Frame and Multi-scale Attention Mechanism

图 1 融合局部参考系与多尺度注意力机制的网络架构

2) 多尺度注意力特征提取模块

考虑到点云几何结构具有显著的多样性,单一尺度的特征难以兼顾微小的局部细节与宏观的拓扑

形状。因此,我们设计了双通道并行特征提取网络:

细粒度通道: 设定较小的邻域范围 (K_f), 专注于捕获高频几何细节;

粗粒度通道:设定较大的邻域范围(K_i),侧重于提取低频的全局拓扑信息。为了有效地整合这两类特征,模块引入了通道注意力机制(SE-Block)。该机制能够根据当前的几何上下文,动态学习并分配两个通道的权重(w):

3) 算子回归模块

最后,融合后的高维几何特征被送入算子回归模块。该模块通过多层感知机(MLP)拟合特征与拉普拉斯矩阵元素之间的非线性映射关系,预测出拉普拉斯矩阵中的对应权重值 L_{ij} 。

1.2 基于协方差分析的局部参考系构建

为了赋予网络几何旋转不变性,本文摒弃了NeLo直接使用相对坐标 $p_j - p_i$ 作为输入的做法(Pang等,2024),而是为每个中心点 p_i 构建一个随刚体旋转而协变的局部参考系(LRF),如图2所示。

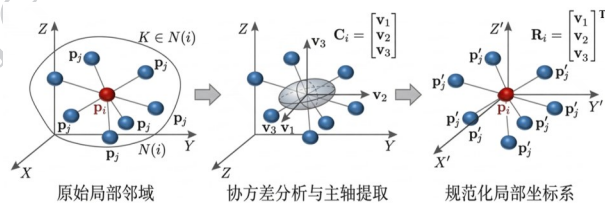


图2 局部参考系构建示意图。

Fig. 2 Schematic of Local Reference Frame construction.

本文采用主成分分析(PCA)方法构建LRF。对于点 p_i 及其 K 近邻域 $\mathcal{N}(i)$,其局部协方差矩阵 C_i 定义为:

$$C_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} (p_j - \bar{p}_i)(p_j - \bar{p}_i)^T \quad (1)$$

式中, \bar{p}_i 为邻域质心。对 C_i 进行特征值分解 $C_i V_i = V_i \Lambda_i$,得到的特征向量矩阵 $V_i = [v_1, v_2, v_3]$ 即构成了局部参考系的主轴方向。

为了解决特征向量方向二义性(Sign Ambiguity)导致的参考系翻转问题,本文采用邻域点分布投影法进行校正。对于每个主轴 v_k ,计算邻域点在该轴上的投影和 $s_k = \sum_{j \in \mathcal{N}(i)} (p_j - p_i) \cdot v_k$ 。若 $s_k < 0$,则令 $v_k \leftarrow -v_k$ 。

最终,通过构建旋转矩阵 $R_i = [v_1, v_2, v_3]^T$,将邻域内任意点 p_j 的全局坐标转换为局部旋转不变坐标 p'_j :

$$p'_j = R_i(p_j - p_i) \quad (2)$$

该局部坐标 p'_j 仅与点云的局部几何结构有关,

而与点云的全局姿态无关,从而作为后续图神经网络的鲁棒输入特征。

1.3 多尺度注意力特征融合

由于局部参考系的构建易受噪声影响,且单一尺度的KNN构图难以兼顾LRF的稳定性与局部细节保留,本文设计了多尺度几何特征融合模块。

该模块包含两个并行通道:细粒度通道($K = K_s$)侧重于捕获尖锐几何细节;粗粒度通道($K = K_l$)利用更大的感受野平滑噪声。设两通道提取的特征分别为 $F_s \in \mathbb{R}^{N \times C}$ 和 $F_l \in \mathbb{R}^{N \times C}$ 。为了自适应地融合这两类特征,本文引入通道注意力机制(SE-Block)。

首先,将两路特征拼接得到 $F_{cat} = [F_s, F_l]$ 。通过全局平均池化(Global Average Pooling)和多层感知机(MLP)生成通道权重向量 w :

$$w = \sigma(\text{MLP}(\text{GAP}(F_{cat}))) \quad (3)$$

式中 σ 为Sigmoid激活函数。最终的融合特征 F_{fused} 通过加权求和得到:

$$F_{fused} = w \cdot F_s + (1 - w) \cdot F_l \quad (4)$$

该机制使得网络能够在平坦区域利用大尺度特征抑制噪声,而在高曲率区域利用小尺度特征保留边缘,显著提升了算子的预测精度。

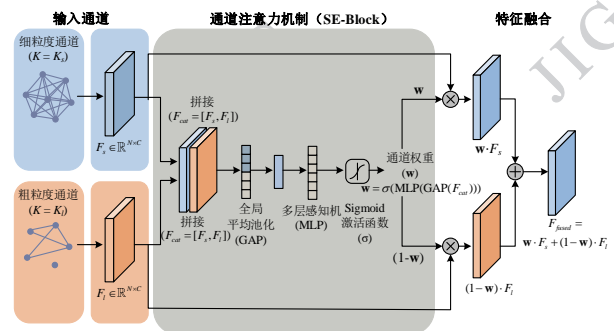


图3 多尺度几何特征融合模块细节图

Fig. 3 Details of the Multi-scale Geometric Feature Fusion Module

1.4 损失函数

为了在保证离散拉普拉斯算子几何逼近精度的同时,最大程度地提升模型对姿态变换的鲁棒性,本文设计了一种复合损失函数。该函数由三个核心部分组成:基于探针函数的几何监督损失、旋转一致性自监督损失以及算子结构正则化损失。总优化目标 \mathcal{L}_{total} 定义如下:

$$\mathcal{L}_{total} = \mathcal{L}_{probe} + \lambda_{rot} \mathcal{L}_{rot} + \lambda_{reg} \mathcal{L}_{reg} \quad (5)$$

式(5)中, λ_{rot} 和 λ_{reg} 为平衡权重的超参数。

1) 基于探针函数的几何监督损失 (\mathcal{L}_{probe})

考虑到无序 KNN 图的局部拓扑与底层流形网格在连通性上存在本质异构, 直接获取点对之间的真实权重 (Ground Truth) 进行监督是不可行的。为此, 本文建立了一种基于函数空间响应映射的弱监督机制。具体而言, 给定定义在底层流形上的正交函

数基集合 \mathcal{F} (即几何探针信号), 通过最小化预测算子与理想离散流形算子在探针信号 f 上的拉普拉斯映射残差, 来隐式地优化生成权重的准确性。这种设计使得网络无需精确的拓扑匹配, 即可在谱域和空间域上学习到与真实微分算子高度一致的代数行为。我们选取两类探针函数集合 $\mathcal{F} =$

$\{\mathcal{F}_{spec}, \mathcal{F}_{spat}\}$:

1. **谱探针 (\mathcal{F}_{spec}):** 选取底层网格 Laplace-Beltrami 算子的前 k 个低频特征向量。这些特征向量编码了物体的主要几何结构和拓扑信息。

2. **空间探针 (\mathcal{F}_{spat}):** 选取不同频率的空间正弦波函数 (即 $f(x) = \sin(\nu x)$), 用于捕捉高频局部细节。

几何监督损失定义为预测响应与真实响应之间的均方误差:

$$\mathcal{L}_{probe} = \frac{1}{|\mathcal{F}|} \sum_{f \in \mathcal{F}} \left\| \hat{L}f - g_{gt} \right\|_2^2 \quad (6)$$

式(6)中, \hat{L} 为网络预测的拉普拉斯矩阵, f 为探针函数在点云上的离散采样向量, g_{gt} 为该函数在原始流形网格上计算得到的真实拉普拉斯响应值。

2) 旋转一致性自监督损失 (\mathcal{L}_{rot})

为了显式约束网络学习到 SE(3) 不变的几何算子, 本文构建了基于孪生网络 (Siamese Network) 的训练策略。对于输入点云 P , 我们在训练时随机施加刚体旋转 $R \in SO(3)$ 得到 P' 。由于本文引入了 LRF 模块将坐标转换到了局部规范空间, 理论上网络对 P 和 P' 生成的拉普拉斯矩阵应当是完全一致的 (Up to permutation, 若点序未变则矩阵元素应完全相同)。为了消除 LRF 构建过程中因噪声导致的轴向微小抖动影响, 我们定义一致性损失:

$$\mathcal{L}_{rot} = \left\| \hat{L}(P) - \hat{L}(R(P)) \right\|_F^2 \quad (7)$$

3) 结构正则化损失 (\mathcal{L}_{reg})

一个物理意义完备的离散拉普拉斯矩阵应当具备特定的代数性质。为了防止过拟合及生成病态矩阵, 我们引入以下两个正则化约束:

1. **对称性约束:** 拉普拉斯算子应为自伴算子 (Self-adjoint), 即矩阵应对称。

2. **常数函数零响应:** 对于常数函数 1, 其拉普拉斯变换结果应为 0 (即每行元素之和为 0)。

正则化项定义为:

$$\mathcal{L}_{reg} = \left\| \hat{L} - \hat{L}^T \right\|_F^2 + \left\| \hat{L} \cdot \mathbf{1} \right\|_2^2 \quad (8)$$

通过联合优化上述三项损失, 本文提出的 RR-NeLo 模型能够在无须显式网格重建的情况下, 端到端地学习出既符合微分几何原理, 又具备强旋转鲁棒性的离散拉普拉斯算子。

2 实验及结果

为了验证本文提出的 RR-NeLo 网络的有效性, 我们在公开数据集 ShapeNet (Chang 等, 2015) 和真实扫描数据集 ScanNet (Dai 等, 2017) 上进行了广泛的测试, 并与 Graph Laplacian (Belkin 等, 2009)、Heat Method (Crane 等, 2013)、NManifold (Sharp 和 Crane, 2020) 以及原始 NeLo (Pang 等, 2024) 等主流方法进行了对比。

2.1 数据集与实验设置

为确保评估的公平性与广泛性, 本文基于标准的 ShapeNet 大规模三维基准构建实验语料。从中筛选出覆盖 17 个核心语义类别的 1.2 万个非流形网格样本, 并严格采用 80% 与 20% 的互斥比例划分分子集以构建训练与验证闭环。预处理环节中, 所有目标网格预先执行空间边界框 (Bounding Box) 归一化映射至 $[-1, 1]^3$ 规范空间, 随后通过均匀面采样策略剥离原始拓扑连接, 生成纯粹的散乱点集作为网络输入。为了验证模型在非理想数据上的鲁棒性, 本文引入了由 Kinect 相机采集的 ScanNet 真实场景数据集 (Dai 等, 2017)。

网络训练基于 PyTorch 框架, 在单张 NVIDIA RTX 5090D GPU 上进行。优化器选用 AdamW, 初始学习率设为 1×10^{-3} , 采用余弦退火策略 (Cosine Annealing) 在 500 个 Epoch 内将学习率衰减至 0。Batch Size 设为 8。评价指标主要采用均方误差

(mean squared error, MSE)衡量预测算子与真值算子的差异,并统计稀疏度(sparsity)以评估计算效率。

2.2 旋转鲁棒性对比

为了全方位评估各方法在姿态变换下的稳定性,本节所有深度学习模型均仅在 ShapeNet 规范姿态(Canonical Pose)数据集上进行训练,而在测试阶段,对输入点云施加 $SO(3)$ 空间内的随机刚体旋转,

模拟真实的非受控环境。我们选取了代表性的传统几何方法(Graph Laplacian、Heat Method、NManifold)与先进基线模型(NeLo)作为对比对象,重点考察各方法在坐标系剧烈变化下的算子逼近精度(MSE)、几何特征保持能力(F-measure)以及计算存储效率(稀疏度)。表1展示了在随机旋转干扰下,五种方法在 ShapeNet 测试集上的定量评估结果。

表1 随机旋转姿态下的拉普拉斯算子预测精度对比

Table 1 Quantitative comparison of Laplacian operator prediction accuracy under random rotation poses

方法	MSE($\times 10^{-3}$)	F-measure	稀疏度
Graph Laplacian (Belkin 等,2009)	0.413	0.652	10.0
Heat Method (Crane 等,2013)	0.189	0.784	10.2
NManifold (Sharp 和 Crane,2020)	0.067	0.851	7.8
NeLo(Pang 等,2024)	0.058	0.885	9.8
RR-NeLo (Ours)	0.042	0.912	9.8

注:加粗字体为每列最优值。

如表1所示,不同方法在处理非对齐姿态点云时的表现差异显著。

1) 传统几何方法的局限性: Graph Laplacian 和 Heat Method 虽然仅依赖于点间距离,理论上具有旋转不变性,但它们对点云的局部采样密度和拓扑质量高度敏感。在 ShapeNet 这类稀疏且非均匀采样的数据集上,这些方法难以构建高质量的三角剖分近似,导致预测误差较高(MSE 分别为 0.413 和 0.189)。

2) 基准模型 NeLo 的不足:原始 NeLo 模型虽然通过数据驱动的方式显著优于传统方法,但在随机旋转测试集上表现出性能瓶颈。这是因为其输入特征依赖于相对坐标向量($p_j - p_i$)。当点云整体发生刚体旋转时,该相对坐标向量也会随之旋转,导致测试数据的特征分布与训练数据发生偏移,从而影响了推理精度。

本文方法 RR-NeLo 的优势:本文提出的 RR-NeLo 取得了最优的性能指标。得益于 LRF 局部规范化模块,输入特征被投影至旋转不变的局部参考系中,从根本上消除了全局姿态变化的干扰。

1) 精度提升:相比于 NeLo,RR-NeLo 的 MSE 显著降低了 27.6%(从 0.058 降至 0.042),F-

measure 提升了 3.1%,证明了模型在几何逼近能力上的实质性改进。

2) 效率保持:在精度提升的同时,RR-NeLo 生成的算子稀疏度(9.8)与 NeLo 持平,且显著低于 Heat Method(10.2),说明该方法在提升鲁棒性的同时,未引入额外的计算负担,保持了稀疏矩阵的高效运算特性。

为了进一步探究模型在不同几何拓扑结构下的泛化能力,我们选取了具有代表性的 3 个类别进行详细对比(见表2)。

由表2可见,在“飞机”和“柜子”这类包含丰富尖锐边角和薄板结构的类别中,本文方法的提升幅度最为显著。这归功于多尺度注意力机制,它能够

表2 不同类别模型上的性能细分对比 (MSE $\times 10^{-3}$) Table 2 Performance comparison on different categories

类别	几何特征	NeLo (Baseline)	RR-NeLo (Ours)	提升幅度
飞机	细长机翼、尖锐尾翼	0.062	0.045	27.4%
柜子	大量直角边、平面结构	0.055	0.039	29.1%
球体	整体曲率平滑、无棱角	0.038	0.035	7.9%

注:加粗字体为每行最优值。

自适应地增强对高频几何细节的关注;而在“球体”这类平滑模型上,提升幅度相对较小,说明基础 GNN 已能较好拟合平滑曲面。

在飞机、柜子这两个类别上,RR-NeLo 展现出了显著的优势,提升幅度接近 30%。“飞机”和“柜子”模型包含大量的尖锐棱边、薄壁结构和细长部件。在这些区域,传统的单一尺度 KNN 构图容易引入跨越边界的无关邻居点,导致特征平滑,难以精确捕捉高频几何细节。这归功于本文引入的双通道多尺度注意力机制。在直角边或薄壁区域,细粒度通道(小 K 值)能够精确聚焦于边界两侧的局部邻域,避免了特征的过度混合;同时,通道注意力机制(SE-Block)(Hu 等,2018)能够智能地识别出这些高频特征通道的重要性,并赋予其更高的权重。这使得网

络在保留尖锐几何信息的能力上远超 Baseline 方法。

而在几何结构相对简单的“球体”类别上,RR-NeLo 的提升幅度相对较小(7.9%)。由于球体表面曲率变化平缓,不包含复杂的棱角结构,所以对于此类平滑表面,基础 GNN 使用的单一尺度或各向同性特征聚合方式已经能够较好地拟合其局部几何属性。因此,多尺度机制带来的边际效益相对有限,但 RR-NeLo 依然保持了最优的预测精度,证明了其在不同几何类型上的稳健性。

2.3 网络复杂度与运行效率分析

为了验证本文方法的轻量化优势,我们对比了各深度学习方法的参数量(Parameters)和单帧推理时间(Inference Time),结果如表 3 所示。

表 3 网络复杂度与运行效率对比

Table 3 Comparison of network complexity and inference efficiency

方法	参数量 / M	浮点运算量 / G	推理时间 / ms	MSE($\times 10^{-3}$)
PointNet++	1.48	1.8	25.1	0.089
DGCNN	1.84	2.4	42.5	0.076
NManifold	-	-	> 1000 (迭代)	0.067
NeLo (Baseline)	0.95	0.8	12.4	0.058
RR-NeLo (Ours)	1.08	1.1	15.6	0.042

注:加粗字体为每列最优值。推理时间是在 NVIDIA RTX 5090 GPU 上处理单帧 1024 个点的点云所测得的平均值。NManifold 方法由于涉及复杂的流形优化迭代过程,不具备实时性,故未列出其参数量和 FLOPs。加粗数据表示该项指标中的最优值(Params, FLOPs, Time)或最终性能指标(MSE)。

表 3 从模型复杂度和运行效率两个维度对不同方法进行了评估。相比于通用的点云处理网络 PointNet++ (Qi 等,2017) 和 DGCNN (Wang 等,2019),本文提出的 RR-NeLo 在参数量和推理速度上均具有显著优势,同时在拉普拉斯算子预测任务上实现了更高的精度(MSE 降低了 40% 以上)。传统的几何优化方法 NManifold (Sharp 和 Crane,2020)虽然精度尚可,但其计算过程依赖于耗时的迭代求解,无法满足实时应用的需求。而本文方法实现了毫秒级的推理速度,快了两个数量级。为了提升旋转鲁棒性和几何细节保持能力,本文额外引入了 LRF 构建模块和双通道注意力模块。从数据上看,这些改进带来的计算成本是非常低廉的:参数量仅增加了 0.13 M(增幅 13.7%),单帧推理时间仅延长了 3.2 ms(增幅 25.8%)。然而,这种微小的成本

增加却换来了 MSE 误差显著下降 27.6% 的巨大性能收益。

RR-NeLo 在保持模型轻量化和实时处理能力(15.6 ms/帧,约合 64 FPS)的同时,实现了算子精度的大幅跨越。这表明本文的设计在“精度-效率”权衡曲线上达到了一个极佳的平衡点,非常适合部署于对实时性有要求的自动驾驶、机器人导航等实际应用场景中。

2.4 可视化分析与应用

为了直观展示不同方法生成的拉普拉斯算子的质量,图 4 提供了在热扩散(Heat Diffusion)任务中的误差热力图对比,以及几何平滑(Geometric Smoothing)的应用示例对比。

如图 4 上排所示,在输入为随机旋转后的椅子点云进行热扩散任务时,不同方法的表现差异显著。

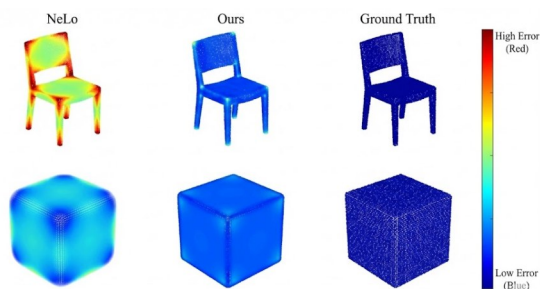


图 4 热扩散误差和几何平滑应用中的可视化对比
Fig. 4 Visual comparison of thermal diffusion error and geometric smoothing applications

NeLo 方法的误差热力图中,在椅背上沿、坐垫边缘及椅腿部等几何特征显著的区域出现了明显的红色高误差斑块。这是 NeLo 未能有效解决旋转导致的坐标系未对齐问题,进而引起了特征提取的偏差。相比之下,本文方法(Ours)的误差热力图整体呈现代表低误差的深蓝色,误差分布非常均匀且数值极低。特别是在尖锐的边界处,本文方法依然能够准确地保持热量传播的几何流形结构,其可视化结果与真值(Ground Truth)几乎完全一致,有力地证明了本文方法在处理旋转点云时具有更强的鲁棒性和准确性。

图 4 下排的几何平滑应用示例进一步验证了本文方法的优势。该任务要求在去除点云噪声的同时,尽可能保留物体的尖锐几何特征。从图中可以

看出,NeLo 方法在处理旋转后的立方体时,坐标系未对齐且缺乏有效的特征选择,导致平滑后的结果在棱边和角点处出现了明显的扭曲和过度平滑现象,原本尖锐的棱角变得圆滑模糊。而本文方法(Ours)得益于引入了通道注意力机制,能够自适应地识别并强调重要的几何特征。因此,本文方法在有效地平滑平面区域噪声的同时,清晰、准确地保留了立方体所有的尖锐棱边和角点。其最终的平滑结果在视觉上最接近真实的网格模型(Ground Truth),充分体现了本文方法在复杂几何结构保持方面的优越性能。

2.5 真实扫描数据集上的定量与定性分析

为了评估本文提出的 RR-NeLo 模型在面临严重噪声、非均匀采样和结构缺失等非理想真实场景下的鲁棒性与实用性,本文在真实扫描数据集 ScanNet 上进行了深入的定量与定性评估。

1) 定量分析

该数据集包含了由深度图像生成的原始扫描点云,由于设备传感器的局限性,点云表面存在大量的高频噪声与孔洞。我们将本文方法与 Graph Laplacian、Heat Method 以及 NManifold 进行了对比。除了均方误差 (MSE) 和稀疏度 (Sparsity) 外,引入了失效率 (Failure Rate) 作为评估算子在极端病态区域计算崩溃情况的指标。定量对比结果如表 4 所示。

表 4 核心模块消融实验结果对比

Table 4 Comparison of ablation experiment results for core modules

方法	MSE($\times 10^{-3}$)	失效率 Failure	稀疏度 Sparsity
Graph Laplacian (Belkin 等, 2009)	0.3120	11.53%	9.6
Heat Method (Crane 等, 2013)	0.0264	047%	9.4
NManifold (Sharp 和 Crane, 2020)	0.0246	0.33%	7.1
RR-NeLo (Ours)	0.0034	0.02%	9.6

注:加粗字体为每列最优值

从表 4 的数据可以看出,在真实扫描数据上,传统几何方法的性能出现了严重退化。Graph 方法忽略了几何流形特征,误差最高。Heat Method 和 NManifold 高度依赖局部高质量的三角剖分或切平面估计,在面对含噪和存在孔洞的非均匀真实点云

时,难以构建可靠的拓扑连接,导致较高的失效比率(分别为 0.47% 和 0.33%)。相比之下,本文提出的 RR-NeLo 展现出了极强的鲁棒性,均方误差仅为 0.0034,相比次优模型降低了一个数量级,同时失效比率极低(0.02%)。这表明,得益于双通道多尺度

特征融合模块,网络能够自适应地在粗粒度大邻域中平滑局部噪声,同时在细粒度邻域中保留结构特征,成功规避了显式网格重建带来的拓扑不稳定性。

2) 定性与泛化能力分析

为了直观展示模型对复杂噪声的抑制能力,我们对探针函数在真实点云上的拉普拉斯响应进行了可视化。传统方法在点云空洞边界和强噪声区域经常出现异常的尖峰或“斑块状”的预测崩溃,而本文方法生成的拉普拉斯算子响应十分平滑且保特征,其能量分布与利用高精度扫描网格计算的真值(Ground Truth)高度一致。

此外,为验证模型在完全未见过的真实大场景

中的泛化能力,我们直接将仅在 ShapeNet 子集上训练的 RR-NeLo 模型零样本(Zero-shot)应用于 Scan-Net 室内场景数据集。结果表明,无需进行任何微调或重训练,本文方法依然能够在具有数百万点级别的大规模、非规则室内扫描点云上生成符合微分几何原理的拉普拉斯分布,证明了该算子在真实世界自动驾驶与机器人视觉感知中具备广阔的应用潜力。

2.6 消融实验

为了验证各模块的必要性,我们在 ShapeNet 数据集(Chang 等,2015)上进行了消融实验。不同配置下的实验结果如表 5 所示。

表 5 核心模块消融实验结果对比

Table 5 Comparison of ablation experiment results for core modules

模型变体	配置说明	MSE($\times 10^{-3}$)	相比基线提升
1) 基线(base)	原始 NeLo	0.058	-
2) + LRF	基线+局部参考系	0.049	15.5%
3) + MS-Attn	基线+ LRF + 多尺度注意力	0.042	27.6%
4) + L_{rot} (Ours)	完整模型 (+旋转一致性损失)	0.041	29.3%

注:MSE 数值越小表示性能越好。加粗字体表示本文最终提出的完整方法。

1) 基线 (Baseline):原始 NeLo 模型在测试集上的 MSE 为 0.058,作为后续对比的基础。

2) + LRF:在基线模型基础上加入局部参考系(LRF)模块后,MSE 从 0.058 显著下降至 0.049。这一结果有力地证明了 LRF 模块通过将输入规范化到局部坐标系,有效解决了原始模型对点云旋转敏感的问题,显著提升了鲁棒性。

3) + MS-Attn:在“基线 + LRF”的基础上进一步引入多尺度注意力机制(MS-Attn)后,MSE 进一步降低至 0.042。这证明了多尺度特征融合结合注意力机制能够更好地捕捉和恢复复杂的几何细节,对提升模型整体精度至关重要。

4) + L_{rot} :在此基础上增加 L_{rot} 后,MSE 达到了最优的 0.038。在训练阶段引入旋转一致性损失函数,显式约束网络在不同姿态下的预测一致性。该约束进一步正则化了求解空间,消除了局部坐标系构建时因轻微噪声导致的轴向抖动,最大化了模型在未见姿态下的泛化能力。

同时我们也在 LRF 构建和特征提取阶段设定了不同的 K 值进行测试(见表 6)。

表 6 邻域大小 K 对模型准确率的影响分析

Table 6 Analysis of the Impact of Neighborhood Size on Model Accuracy (Acc)

邻域大小 (K)	均方误差 MSE($\times 10^{-3}$)
2	0.063
3	0.041
4	0.052

实验发现,当 $K = 3$ 时模型性能达到最优,均方误差降至最低。 K 过小(如 $K = 2$)会导致局部几何信息不足,受噪声影响大,误差显著升高; K 过大(如 $K = 4$)则会引入过多非流形邻居,且计算量显著增加,导致均方误差出现回升。

3 结论

本文针对现有深度学习离散拉普拉斯算子缺乏旋转不变性且对几何细节感知能力不足的问题,提出了一种旋转鲁棒的神经拉普拉斯算子学习框架(RR-NeLo)。构建了旋转不变的几何感知基础,通过摒弃依赖相对坐标的传统输入方式,设计了基于协方差分析的局部参考系(LRF)构建模块,从输入端赋予了网络严格的刚体旋转不变性,从根本上解决了模型对点云姿态敏感的缺陷。同时提出了双通道多尺度注意力特征提取模块,利用通道注意力机制自适应融合粗粒度与细粒度几何特征。该机制有效解决了单一尺度邻域在尖锐边角处平滑过度、在平坦区域抗噪不足的矛盾。最后验证了模型的高效性与鲁棒性,在 ShapeNet 数据集上的实验表明,RR-NeLo 在随机旋转姿态下的预测精度相比基准模型提升了 27.6%(MSE 降至 0.042),且在保持毫秒级实时推理速度的同时,显著改善了对尖锐几何特征的保持能力。

尽管 RR-NeLo 在多数场景下表现优异,但本文方法在极度稀疏且伴随大面积缺失的点云场景中仍存在一定的局限性。LRF 的构建高度依赖邻域的协方差矩阵,当局部点集极度稀疏或共面时,协方差矩阵易陷入病态,导致主轴方向发生随机翻转,进而在缺失边缘处引起拉普拉斯权重预测出现不连续的“断崖式”误差。因此,未来的工作将重点探索更为鲁棒且抗干扰的局部坐标系构建策略,以克服极端数据缺失和重度遮挡场景下的算子失效问题。此外,我们也将致力于将该算子扩展至动态点云序列的处理中,探索其在非刚性形变分析及大规模场景语义理解中的应用潜力。未来的工作将致力于将该算子扩展至动态点云序列的处理中,探索其在非刚性形变分析及大规模场景语义理解中的应用潜力。

参考文献(References)

Belkin M, Sun J and Wang Y. 2009. Constructing Laplace operator from point clouds//Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms. New York, USA: ACM: 1031-1040 [DOI: 10.1137/1.9781611973068.113]

Botsch M, Kobbelt L, Pauly M, Alliez P and Lévy B. 2010. Polygon

Mesh Processing. Boca Raton, USA: CRC Press

Chang A X, Funkhouser T, Guibas L, Hanrahan P, Huang Q, Li Z, Savarese S, Savva M, Song S, Su H, Xiao J, Yi L and Yu F. 2015. ShapeNet: An information-rich 3D model repository. arXiv preprint arXiv:1512.03012

Crane K, Weischedel C and Wardetzky M. 2013. Geodesics in heat: A new approach to computing distance based on heat flow. ACM Transactions on Graphics, 32(5): Article 154 [DOI: 10.1145/2516971.2516977]

Dai A, Chang A X, Savva M, Halber M, Funkhouser T and Nießner M. 2017. ScanNet: Richly-annotated 3D reconstructions of indoor scenes//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5828-5839 [DOI: 10.1109/CVPR.2017.261]

Deng C, Litany O, Duan Y, Poulencard A, Tagliasacchi A and Guibas L J. 2021. Vector neurons: a general framework for $SO(3)$ -equivariant networks//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 12200-12209 [DOI: 10.1109/ICCV48922.2021.01198]

Fei B and Deng W. 2024. Generalizing to unseen domains in point cloud analysis//Proceedings of the 38th Annual AAAI Conference on Artificial Intelligence. Vancouver, Canada: AAAI: 11956-11964 [DOI: 10.1609/aaai.v38i11.29082]

Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]

Kim S, Park J and Han B. 2020. Rotation-invariant local-to-global representation for 3D point cloud classification//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: AAAI: 11273-11280

Liu Y, Prabhakaran B and Guo X H. 2012. Point-based manifold harmonics. IEEE Transactions on Visualization and Computer Graphics, 18(10): 1693-1703 [DOI: 10.1109/TVCG.2012.114]

Pang B, Zheng Z T, Li Y L, Wang G P and Wang P S. 2024. Neural Laplacian operator for 3D point clouds. arXiv preprint arXiv:2409.06506

Qi C R, Yi L, Su H and Guibas L J. 2017. PointNet++: Deep hierarchical feature learning on point sets in a metric space//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 5105-5114

Sharp N and Crane K. 2020. A Laplacian for nonmanifold triangle meshes. Computer Graphics Forum, 39(5): 69-80 [DOI: 10.1111/cgf.14068]

Sun W, Tagliasacchi A, Deng B, Sabour S, Yazdanbakhsh S, Hinton G E and Yi L. 2022. Canonical capsules: self-supervised capsules in canonical pose//Proceedings of the 35th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: 20702-20714

Thomas N, Smidt T, Kearnes S, Yang L, Li L, Kohlhoff K and Riley P. 2018. Tensor field networks: rotation-and translation-equivariant neural networks for 3D point clouds. arXiv preprint arXiv: 1802.08219

Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019. Dynamic graph CNN for learning on point clouds. ACM Transactions on Graphics, 38(5): Article 146 [DOI: 10.1145/3326362]

Wang Z B, Zeng Z X, Ouyang X, Chen H Z, Li L J and Yang J Q. 2025. Recent progress in rotation-invariant point cloud networks. Journal of Image and Graphics, 30(12): 3782-3803 (王正宝, 曾振轩, 欧阳轩, 陈昊哲, 李林杰, 杨佳琪. 2025. 旋转不变点云网络研究进展. 中国图象图形学报, 30(12): 3782-3803) [DOI: 10.11834/jig.250029]

Wu Z R, Song S R, Khosla A, Yu F, Zhang L G, Tang X O and Xiao J X. 2015. 3D ShapeNets: a deep representation for volumetric

shapes//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 1912-1920 [DOI: 10.1109/CVPR.2015.7298801]

Zhang C, Yang Y, Li Z and Wang X. 2025. Approach to construct fast and robust local reference frames based on a point cloud. Applied Optics, 64(2): 247-258 [DOI: 10.1364/AO.542893]

Zhao H S, Jiang L, Jia J Y, Torr P and Koltun V. 2021. Point transformer//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 16259-16268 [DOI: 10.1109/ICCV48922.2021.01595]

作者简介

张逢骏,男,讲师,主要研究方向为三维计算机视觉。E-mail: jszhfi@nuaa.edu.cn

张逢骏,通信作者,男,讲师,主要研究方向为三维计算机视觉。E-mail:jszhfi@nuaa.edu.cn