

局部-全局动态图学习与互补融合的点云配准方法^{*}

邱巧燕¹, 叶海良¹, 曹飞龙¹, 吕科²

¹(中国计量大学 理学院, 浙江 杭州 310018)

²(中国科学院大学 工程科学学院, 北京 100049)

通信作者: 曹飞龙, E-mail: caofeilong88@zjnu.edu.cn



摘 要: 现有基于深度学习的点云配准方法主要聚焦于特征提取和特征匹配方面的研究, 然而, 其在特征提取阶段对局部和全局图结构的挖掘尚不充分, 同时在匹配过程中对差异信息的探索也较为有限. 为此, 提出了一种局部-全局动态图学习与互补融合的点云配准方法. 具体而言, 动态偏移的局部图学习模块通过构造包含几何和语义信息的代理点来刻画特征空间中潜在的图结构, 从而获得更具判别性的局部特征. 其次, 设计了动态关注的全局图学习模块, 根据点之间的相互关系自适应地调整关注权重, 有效地捕获了点云中的长程依赖关系. 为了进一步提高两个点云之间的对应关系, 构造了注意力驱动的互补融合模块, 根据交叉注意力机制来挖掘相似信息和差异信息, 并利用自注意力机制优化特征之间的关联性. 实验结果表明, 该方法在公开数据集上实现了最优的配准效果, 并具备良好的计算效率.

关键词: 深度学习; 点云配准; 图神经网络; 特征提取; 特征匹配

中图法分类号: TP181

中文引用格式: 邱巧燕, 叶海良, 曹飞龙, 吕科. 局部-全局动态图学习与互补融合的点云配准方法. 软件学报, 2025, 36(12): 5739–5754. <http://www.jos.org.cn/1000-9825/7416.htm>

英文引用格式: Qiu QY, Ye HL, Cao FL, Lyu K. Point Cloud Registration Method Based on Local-global Dynamic Graph Learning and Complementary Fusion. Ruan Jian Xue Bao/Journal of Software, 2025, 36(12): 5739–5754 (in Chinese). <http://www.jos.org.cn/1000-9825/7416.htm>

Point Cloud Registration Method Based on Local-global Dynamic Graph Learning and Complementary Fusion

QIU Qiao-Yan¹, YE Hai-Liang¹, CAO Fei-Long¹, LYU Ke²

¹(College of Sciences, China Jiliang University, Hangzhou 310018, China)

²(School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Existing deep learning-based point cloud registration methods primarily focus on feature extraction and feature matching. However, the exploration of local and global graph structures during the feature extraction stage remains insufficient, and the investigation of difference information during the matching process is also limited. To address these issues, this study proposes a point cloud registration method based on local-global dynamic graph learning and complementary fusion. Specifically, the dynamic offset-based local graph learning module characterizes the underlying graph structure in the feature space by constructing proxy points that contain both geometric and semantic information, leading to more discriminative local features. In addition, a dynamic attention-based global graph learning module is designed, which adaptively adjusts attention weights based on the relationships between points, effectively capturing long-range dependencies in the point cloud. To further enhance the correspondence between the two point clouds, the attention-driven complementary fusion module utilizes the cross-attention mechanism to extract similar and distinctive information, while applying the self-attention mechanism to refine the relationships between features. Experimental results demonstrate that the proposed method achieves optimal registration performance on public datasets while maintaining acceptable computational efficiency.

* 基金项目: 国家自然科学基金 (62032022, 62176244)

收稿时间: 2024-12-13; 修改时间: 2025-01-10; 采用时间: 2025-02-18; jos 在线出版时间: 2025-09-03

CNKI 网络首发时间: 2025-09-04

Key words: deep learning; point cloud registration; graph neural network (GNN); feature extraction; feature matching

点云配准是三维视觉领域的核心任务之一,广泛应用于姿态估计^[1,2]、自动驾驶^[3,4]和即时定位与地图构建^[5,6]等前沿技术中.其主要目标是通过求解最优变换参数,将来自不同设备或视角的点云对齐至同一坐标系,从而实现点与点之间的精确对应.然而,点云数据常常受到噪声干扰、物体遮挡及部分点缺失等因素的影响,这使得如何提取具有判别性的特征并进行高效匹配成为亟待解决的关键问题.

早期的点云配准方法主要依靠优化算法,通过反复迭代计算变换参数,最小化对应点之间的几何误差来完成点云的对齐.其中,迭代最近点^[7]及其变体^[8,9]是最具代表性的配准方法.但它们高度依赖初始位姿,且容易陷入局部最优解.此外,基于正态分布变换^[10]和四点全等集^[11]的方法也是经典的点云配准方法.虽然这些方法在处理噪声和部分重叠的点云时能够取得较好的配准效果,但仍面临计算复杂度高和缺乏特征学习的挑战.

近年来,深度学习技术的迅速发展为点云配准领域带来了诸多突破. Aoki 等人^[12]开创性地将 PointNet 网络^[13]与经典的 LK 算法相结合,提出了首个端到端的点云配准框架.针对部分到部分点云配准任务, Xu 等人^[14]提出了一种基于全局特征的重叠区域估计方法,减少了非重叠区域的干扰.然而,这些基于 PointNet 的方法主要侧重于全局特征的提取,未能充分捕捉局部细节信息.为此, Lu 等人^[15]引入了 PointNet++ 网络^[16],旨在更有效地获取点云的局部语义特征. Yew 等人^[17]设计了一种融合几何属性的混合特征提取网络,以获得更精确的初始匹配. Li 等人^[18]将局部几何特征和距离特征整合到匹配过程中,增强了识别正确对应点对的能力.随后, Mei 等人^[19]根据点与点之间的位置关系自适应地构造卷积核来捕获局部相关性.此外,一些研究^[20-22]致力于将几何结构描述纳入特征学习过程中,进一步提升对应关系的鲁棒性.尽管上述方法在局部特征提取方面取得了一定成效,但仅依赖局部特征进行匹配容易受到局部结构相似性的影响,从而导致对应关系的不唯一性.因此,在局部特征的基础上引入全局特征约束,以确保两个点云的整体对应关系,对于提高匹配一致性至关重要.

随着图神经网络的兴起,基于图的点云配准方法取得了显著的进展. Wang 等人^[23]通过共享的动态图神经网络提取局部特征,并结合 Transformer 模块^[24]建模长程依赖关系,从而生成更具判别性的点特征.随后, Huang 等人^[25]在获得超点的特征后,利用图神经网络对每个点云的上下文信息进行编码,以量化点云之间的重叠概率. Fischer 等人^[26]基于图神经网络架构,借助多头注意力实现上下文信息的聚合.在此基础上, Shi 等人^[27]构建了一种带有注意力机制的动态多路图用于整合全局信息,同时结合局部邻域信息来增强特征表示.为了建立鲁棒的点对应关系, Kadam 等人^[28]基于每个点的局部属性采用多跳学习策略,提取从局部到全局的层次化特征. Chen 等人^[29]提出了一种点云结构提取器,旨在建模全局关系并识别相邻结构特征,从而增强配准模型对噪声的鲁棒性.最近, Jiang 等人^[30]通过引入全局图卷积将局部特征映射为全局特征,有效地捕捉了点云的上下文关系.然而,这些方法提取的特征往往局限于单个点云,忽略了点云之间的交互关系.因此,必须引入特征交互机制,使得两个点云能够相互影响和更新,从而提升匹配的准确性.

为了实现更精确的对齐,部分研究者提出了高效的特征匹配方法. Wu 等人^[31]选择与另一点云相似的局部特征和全局特征来更新自身点云的特征,进一步提高了配准精度. Xu 等人^[32]设计了一个包含旋转和平移的双分支结构,并在多个阶段引入逐点特征交互模块,以促进信息的深度融合.随后, Fu 等人^[33]将软对应关系作为辅助信息,利用交叉图卷积对另一点云的特征进行聚合.为了更有效地感知两个点云之间的相似点对,一些研究^[34-37]采用了自注意力和交叉注意力机制来实现点云之间的上下文交互,在一定程度上增强了两个点云的关联性.

尽管近年来基于深度学习的点云配准方法在多个方面都取得了重要突破,但仍然面临一些局限性.

- 大多数现有的局部特征提取方法倾向于直接聚合几何信息和语义特征,未能充分利用几何信息来指导语义特征的动态调整,容易导致特征表示对局部几何结构的感知能力不足.此外,仅依赖坐标空间的邻近关系来构建图结构,忽视了特征空间中潜在的关联,从而限制了特征的表示能力.

- 一些方法在学习全局特征时,往往采用全连接的方式聚合所有点的信息,无法区分点对之间的相关性,从而导致低相关性信息的过度引入.

- 在特征匹配阶段,许多方法主要侧重两个点云相似信息的交互,忽略了差异信息的提取与利用.虽然相似信

息对于匹配过程至关重要, 但差异信息同样不容忽视, 因为它能够揭示两个点云在位姿、形状或细节方面的独特性。

针对上述问题, 本文提出了局部-全局动态图学习与互补融合网络 (local-global dynamic graph learning and complementary fusion network, LGCNet), 旨在更高效地应对点云配准任务。该方法致力于在特征提取阶段生成具有局部和全局视野的高判别性的特征描述符, 以便在后续的匹配阶段可以融合两个点云之间的相似信息与差异信息来提高点云配准的准确性。具体而言, 在动态偏移的局部图学习模块中, 根据动态偏移机制在每个点的局部邻域中生成富含几何特征和语义信息的代理点, 并基于代理点之间的图结构通过图卷积操作来实现局部特征的增强表达。在动态关注的全局图学习模块中, 综合考虑特征空间和几何距离来构建全局图的邻接矩阵, 再结合动态关注机制自适应地调整各点之间的关联权重, 从而实现信息的长程依赖并抑制低相关点的影响。为了进一步提升两个点云在匹配过程中的交互, 设计了一个注意力驱动的互补融合模块, 其新颖之处在于能够采用交叉注意力机制从差异性和相似性两个角度来关注点云之间的相互作用, 同时引入自注意力机制来加强互补融合后的特征在全局范围内的一致性。本文的主要贡献如下。

- 提出了一种基于局部-全局动态图学习与互补融合的点云配准方法, 有效提升了点云配准的性能。
- 设计了动态偏移的局部图学习模块, 通过动态偏移机制生成代理点并构建局部图, 实现了更加精细的局部特征表示。
- 设计了动态关注的全局图学习模块, 利用动态关注机制自适应地优化全局图结构, 获得了更具判别性的特征。
- 构建了注意力驱动的互补融合模块, 根据交叉注意力机制从相似性和差异性两个方面挖掘点云之间的交互, 进一步增强了对应关系的准确性。

本文第 1 节详细介绍所提出的局部-全局动态图学习与互补融合网络 (LGCNet)。第 2 节通过实验验证 LGCNet 的有效性。第 3 节总结全文, 并展望未来研究方向。

1 局部-全局动态图学习与互补融合网络

1.1 网络的整体框架

本文设计了一个面向点云配准的局部-全局动态图学习与互补融合网络 (LGCNet), 整体框架如后文图 1 所示。对于输入的源点云 $X \in \mathbb{R}^{N \times 3}$ 和目标点云 $Y \in \mathbb{R}^{M \times 3}$, 其中, N 和 M 分别表示点云中的点数量, 配准的目的在于找到使两个点云对齐的变换参数 $\{R, t\}$ 。通常, 基于端到端学习的配准流程包括特征提取、特征匹配和变换估计等阶段^[38]。具体而言, 在特征提取阶段, 本文提出了一种层次化的特征提取架构, 通过堆叠动态偏移的局部图学习模块和动态关注的全局图学习模块, 实现了局部与全局特征的协同增强。其中, 动态偏移的局部图学习模块有效保证了对局部细节的精准描述; 动态关注的全局图学习模块则能够捕捉点云中的长程依赖关系, 强化全局特征的建模。在特征匹配阶段, 注意力驱动的互补融合模块引入交叉注意力机制和自注意力机制, 促进了两个点云之间信息的有效交互与增强。最后, 计算两个点云之间的对应矩阵, 并利用奇异值分解 (SVD) 算法求解最终的变换矩阵, 从而完成点云配准任务。

1.2 动态偏移的局部图学习模块

针对点云配准任务, 局部特征的有效提取对于获得正确的初始对应关系至关重要。点云的坐标包含了丰富的空间结构信息, 但现有的方法大多将预定义的几何信息与语义特征进行简单的加权聚合, 忽略了几何坐标对语义特征的潜在影响。特别是在局部结构复杂的情况下, 生成的特征表示往往无法有效适应空间结构, 缺乏对局部几何的精准感知能力。此外, 特征空间的高阶关系能够揭示点与点之间的深层次联系, 但一些方法无法充分挖掘这些潜在的关联, 导致图结构的构建过于依赖几何邻近性, 从而限制了特征的表示能力。为此, 本文提出了一个动态偏移的局部图学习模块, 通过构造代理点来引导图结构朝着更加适应几何特性和特征分布的方向发展, 从而获得更加判别性的局部表示。图 2 展示了动态偏移的局部图学习模块的详细过程。

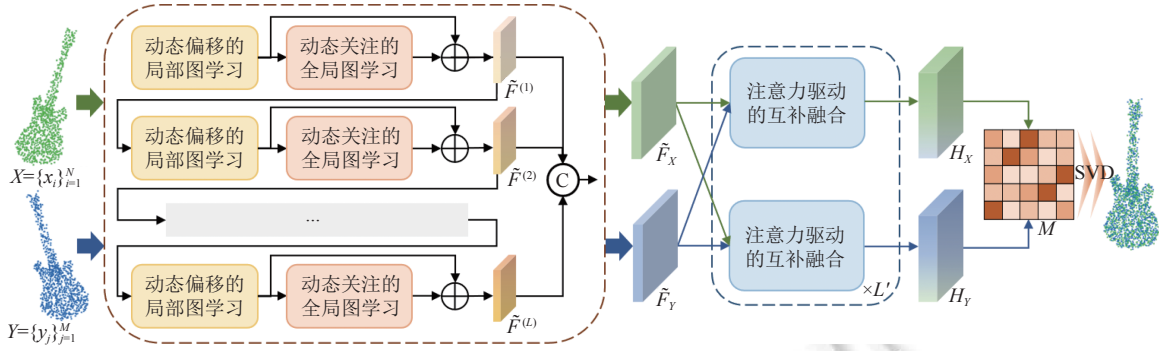


图1 局部-全局动态图学习与互补融合网络 (LGCNet) 的整体框架

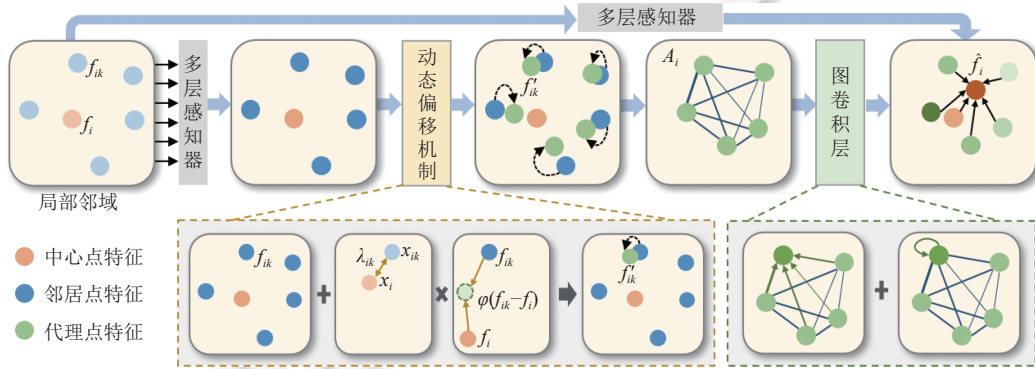


图2 动态偏移的局部图学习模块示意图

以源点云为例, 对于点云中的每个点 x_i , 在坐标空间中选择 K 个最近的相邻点 $\{x_{ik}\}_{k=1}^K$ 构造其局部邻域图, 同时将局部图中的点集定义为 $X_i = \{x_i, x_{i1}, x_{i2}, \dots, x_{iK}\}$, 所对应的特征表示为 $F_i = \{f_i, f_{i1}, f_{i2}, \dots, f_{iK}\}$. 为了更准确地捕捉局部邻域内的结构差异, 根据邻居点与中心点间的坐标关系, 利用高斯核函数为每个邻居点 x_{ik} 生成偏移因子 λ_{ik} :

$$\lambda_{ik} = \exp\left(-\frac{\|x_i - x_{ik}\|^2}{\sigma}\right) \quad (1)$$

其中, $\|\cdot\|$ 为欧几里得范数, σ 是高斯核参数. 通过这种方式, 偏移因子 λ_{ik} 能够动态地控制每个邻居点特征偏移的程度. 具体而言, 靠近中心点的邻居点特征会进行更大程度的调整, 从而更精确地反映局部几何结构; 而远离中心点的邻居点特征调整幅度相对较小, 有助于确保其在全局结构中的稳定性. 接着, 根据得到的偏移因子 λ_{ik} 和语义特征的差异对邻居点的原始特征 f_{ik} 进行动态调整, 生成代理点特征 f'_{ik} :

$$f'_{ik} = f_{ik} + \lambda_{ik} \cdot \varphi(f_{ik} - f_i) \quad (2)$$

其中, $\varphi(\cdot)$ 表示一个带有非线性激活函数的多层感知器. 这一过程利用动态偏移机制实现了局部特征的精细化调整, 有效提高了特征的代表能力.

为了进一步揭示点与点之间更深层次的依赖关系, 针对所构造的代理点, 编码其相对关系以构建局部动态图的邻接矩阵 $A_i \in \mathbb{R}^{K \times K}$:

$$a_{ik} = \text{Sigmoid}(((x_i - x_{ik}) \parallel (\varphi(f_i) - \varphi(f_{ik}))) W) \quad (3)$$

$$A_i(k, k') = \text{Softmax}(a_{ik} \cdot a_{ik'}^T), k, k' = 1, 2, \dots, K \quad (4)$$

其中, $W \in \mathbb{R}^{(3+K) \times 1}$ 是一个可学习的参数, $\parallel$ 表示沿着通道维度的拼接操作, $\varphi(\cdot)$ 为两个非共享的线性变换操作. $\text{Sigmoid}(\cdot)$ 函数将输入映射到 $(0, 1)$ 区间, 当值越大时相关性越强, 反之则相关性越弱. 由于文献 [31, 39] 认为边特征更能揭示局部结构信息, 于是本文进一步强化了边特征在代理点特征更新中的作用. 具体地, 利用图卷积操作实

现依赖邻接矩阵的边特征聚合和独立的边特征自更新, 得到增强的代理点特征表示 $\hat{f}_{ik}$:

$$\hat{f}_{ik} = \sum_{k'=1}^K A_i(k, k') \cdot \zeta(g_1(\varphi(f'_{ik'} - f_i))) + \zeta(g_2(\varphi(f'_{ik} - f_i))) \quad (5)$$

其中, g_1 和 g_2 分别表示两个独立的全连接层, $\zeta(\cdot)$ 是 $ReLU$ 激活函数. 最后, 通过相关性系数自适应地聚合代理点特征来重建中心点的特征 f'_i . 为了防止信息丢失, 使用残差连接与中心点原始特征相加, 从而得到最终的局部特征描述符 $\hat{f}_i$. 该过程可以表示为:

$$f'_i = \sum_{k=1}^K a_{ik} \cdot \hat{f}_{ik} \quad (6)$$

$$\hat{f}_i = \varphi(f_i) + f'_i \quad (7)$$

其中, $\varphi(\cdot)$ 是一个带有非线性激活函数的多层感知器.

总体而言, 动态偏移的局部图学习模块通过基于距离的动态偏移机制来构造具有几何特性和语义信息的代理点, 从特征空间的潜在图结构中挖掘丰富的局部信息, 显著增强了特征表示的判别性.

1.3 动态关注的全局图学习模块

在特征提取的过程中, 动态偏移的局部图学习模块主要聚焦于局部邻域的特征表示, 实现了对局部细节的精准捕捉. 与已有的一些方法类似, 这种方式通常缺乏全局视角, 忽略了点云中远距离点之间的长程依赖关系. 因此, 本文提出了动态关注的全局图学习模块, 自适应地选择学习全局上下文信息, 增强特征的表达能力. 图 3 展示了动态关注的全局图学习模块的具体架构.

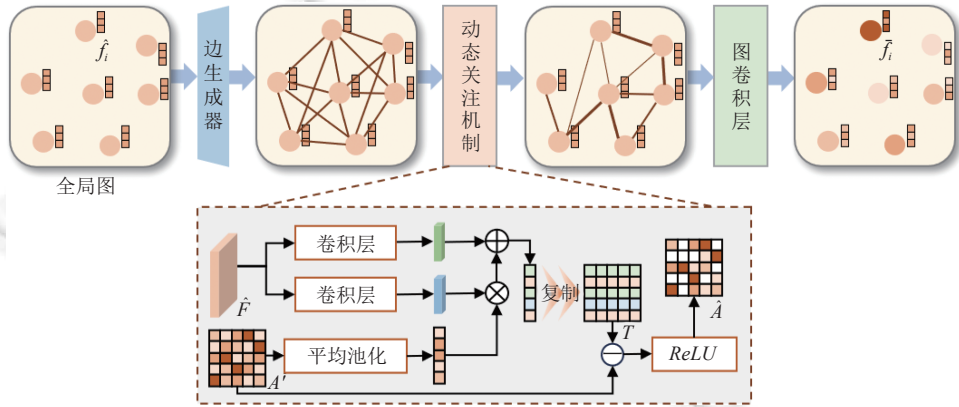


图 3 动态关注的全局图学习模块示意图

将每个点云中的所有点构造为一个加权全局图, 图中的每个节点表示点云中的每个点, 每条边的权重表示两个点之间的相关性. 具体而言, 为了更加全面捕捉点与点之间的相关性, 综合考虑逐点的初始坐标和局部特征来构造全局邻接矩阵 $A' = (a'_{ij}) \in \mathbb{R}^{N \times N}$:

$$a'_{ij} = \text{Softmax} \left(\exp \left(- \frac{\alpha \cdot \|x_i - x_j\|^2 + (1 - \alpha) \cdot \|\hat{f}_i - \hat{f}_j\|^2}{\sigma} \right) \right) \quad (8)$$

其中, σ 是高斯核参数, α 是一个可学习参数. 之后, 引入特征学习到的阈值自适应地对邻接矩阵中的低相关性连接进行筛选和去除, 使其更加聚焦于重要的特征. 具体过程如下:

$$T_i = \frac{\phi_1(\hat{f}_i)}{N} \sum_{j=1}^N a'_{ij} + \phi_2(\hat{f}_i) \quad (9)$$

$$\hat{a}_{ij} = \text{ReLU}(a'_{ij} - T_i), j = 1, 2, \dots, N \quad (10)$$

其中, $\hat{A} = (\hat{a}_{ij})$, $\phi_1(\cdot)$ 和 $\phi_2(\cdot)$ 是两个独立的卷积层. 对于生成的全局图, 采用图卷积操作将全局上下文信息融合到每个点的特征表示中:

$$\tilde{f}_i = \sum_{j=1}^N \hat{a}_{ij} \cdot \zeta(g_1(\hat{f}_j)) + \zeta(g_2(\hat{f}_i)) \quad (11)$$

在获得每个点的局部特征 $\hat{f}_i$ 和全局特征 $\tilde{f}_i$ 之后, 通过自适应加性融合得到最终的逐点特征 $\tilde{f}_i$, 表示为:

$$\tilde{f}_i = \beta \cdot \hat{f}_i + (1 - \beta) \cdot \tilde{f}_i \quad (12)$$

其中, β 是可学习参数.

为了避免信息丢失和过平滑问题, 使用跳跃连接将前一层输入特征与当前层的输入特征进行连接. 同时, 将每一层的输出特征按通道拼接和非线性变换 $\psi(\cdot)$ 的方式进行整合, 得到最终提取到的源点云特征可以表示为:

$$\tilde{F}_X = \psi(\tilde{F}_X^{(1)}, \tilde{F}_X^{(2)}, \dots, \tilde{F}_X^{(L)}) \in \mathbb{R}^{N \times D} \quad (13)$$

其中, $\tilde{F}_X^{(l)} = \{\tilde{f}_{x_1}^{(l)}, \tilde{f}_{x_2}^{(l)}, \dots, \tilde{f}_{x_N}^{(l)}\}$, $l = 1, 2, \dots, L$ 是每一层局部-全局提取到的源点云特征. 同样, 每一层提取到的目标点云的特征为 $\tilde{F}_Y^{(l)} = \{\tilde{f}_{y_1}^{(l)}, \tilde{f}_{y_2}^{(l)}, \dots, \tilde{f}_{y_M}^{(l)}\}$, $l = 1, 2, \dots, L$, 最终整合得到的特征为 $\tilde{F}_Y = \psi(\tilde{F}_Y^{(1)}, \tilde{F}_Y^{(2)}, \dots, \tilde{F}_Y^{(L)}) \in \mathbb{R}^{M \times D}$.

通过这种方式, 动态关注的全局图学习模块能够在关注局部细节的基础上兼顾全局上下文信息, 增强点云的整体表达与点对之间的全局关联性, 从而有效提升点云配准的性能.

1.4 注意力驱动的互补融合模块

由于单一的特征提取无法充分确保两个点云之间的准确对齐, 尤其是在处理复杂结构或部分重叠的点云时, 特征交互变得尤为关键. 在之前的一些研究中, 注意力机制被广泛应用于特征匹配阶段, 通过促进点云之间信息的深度交互来优化配准效果. 然而, 这些方法仅依赖于提取两个点云之间的相似信息, 忽视了差异信息的相互感知. 因此, 本文设计了一个注意力驱动的互补融合模块, 通过引入交叉注意力机制来自适应地挖掘与融合不同点云的差异信息和相似信息, 并结合自注意力机制来进一步加强点云之间的特征交互. 该模块包括以下 4 个阶段, 以源点云为例, 其详细架构如图 4 所示.

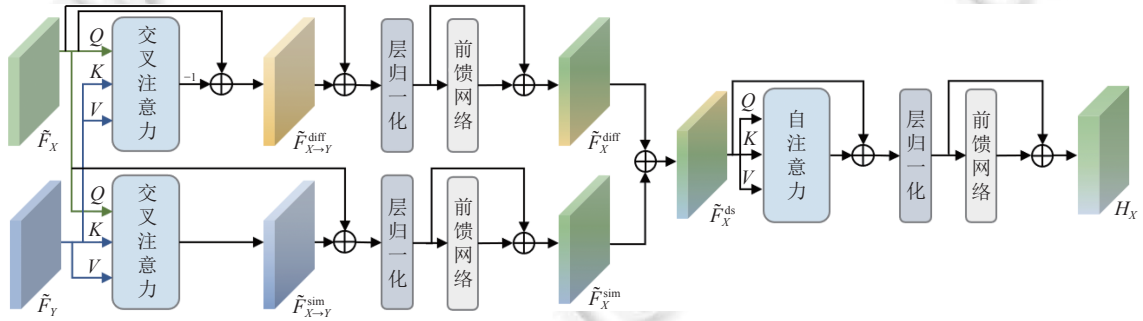


图 4 注意力驱动的互补融合模块示意图

(1) 基于交叉注意力的相似信息提取. 针对源点云和目标点云, 首先利用交叉注意力机制计算两个点云之间的相似性, 并提取其相似性信息 $\tilde{F}_{X \rightarrow Y}^{\text{sim}}$:

$$\tilde{F}_{X \rightarrow Y}^{\text{sim}} = \text{Softmax} \left(\frac{(\tilde{F}_X W_1)(\tilde{F}_Y W_2)^T}{\sqrt{d}} \right) (\tilde{F}_Y W_3) \quad (14)$$

其中, W_1 , W_2 和 W_3 分别是特征映射的权重, d 表示特征的维度. 接着, 将相似信息与源点云的原始信息共同注入前馈网络中, 以充分利用相似信息对原始特征进行增强. 该过程可以表示为:

$$\tilde{F}_X^{\text{sim}} = \text{LN}(\tilde{F}_X + \tilde{F}_{X \rightarrow Y}^{\text{sim}}) + \text{FFN}(\text{LN}(\tilde{F}_X + \tilde{F}_{X \rightarrow Y}^{\text{sim}})) \quad (15)$$

其中, $\text{LN}(\cdot)$ 表示层归一化, $\text{FFN}(\cdot)$ 表示一个具有两层全连接层的前馈神经网络.

(2) 基于交叉注意力的差异信息提取. 为了弥补仅依赖相似信息进行特征匹配的不足, 该模块进一步引入差异信息的挖掘. 在通过交叉注意力得到两点云之间相似信息的基础上, 将目标点云的原始特征与提取到的相似特征进行逐元素相减, 得到目标点云中与源点云无关的差异信息 $\tilde{F}_{X \rightarrow Y}^{\text{diff}}$. 同样, 这些差异特征也会与源点云的原始特征进行融合关注, 以捕捉两个点云之间的特有的结构差异与姿态变化. 公式表示如下:

$$\tilde{F}_{X \rightarrow Y}^{\text{diff}} = \tilde{F}_Y - \tilde{F}_{X \rightarrow Y}^{\text{sim}} \quad (16)$$

$$\tilde{F}_X^{\text{diff}} = \text{LN}(\tilde{F}_X + \tilde{F}_{X \rightarrow Y}^{\text{diff}}) + \text{FFN}(\text{LN}(\tilde{F}_X + \tilde{F}_{X \rightarrow Y}^{\text{diff}})) \quad (17)$$

类似地, $\text{LN}(\cdot)$ 为层归一化操作, $\text{FFN}(\cdot)$ 是一个两层前馈神经网络.

(3) 相似-差异特征融合. 将前两个阶段得到的相似增强信息 $\tilde{F}_X^{\text{sim}}$ 和差异增强信息 $\tilde{F}_X^{\text{diff}}$ 进行加权融合, 以确保融合后的特征 $\tilde{F}_X^{\text{ds}}$ 不仅包含点云之间的共性, 还保留了其独特性:

$$\tilde{F}_X^{\text{ds}} = \eta \cdot \tilde{F}_X^{\text{sim}} + (1 - \eta) \cdot \tilde{F}_X^{\text{diff}} \quad (18)$$

其中, η 是可学习的参数.

(4) 基于自注意力的交互增强. 在完成特征融合之后, 应用自注意力机制对点云内的互补特征进一步优化, 使得不同点云之间的互补信息能够在消息传递过程中得到有效的交互和强化:

$$\tilde{F}_X^{\text{self}} = \text{Softmax}\left(\frac{(\tilde{F}_X^{\text{ds}} W_1)(\tilde{F}_X^{\text{ds}} W_2)^T}{\sqrt{d}}\right)(\tilde{F}_X^{\text{ds}} W_3) \quad (19)$$

$$H_X = \text{LN}(\tilde{F}_X^{\text{ds}} + \tilde{F}_X^{\text{self}}) + \text{FFN}(\text{LN}(\tilde{F}_X^{\text{ds}} + \tilde{F}_X^{\text{self}})) \quad (20)$$

通过上述 4 个步骤, 每层交互后得到的源点云可表示为 $H_X^{(l')} \in \mathbb{R}^{N \times D}$, 其中, $l' = 1, 2, \dots, L'$. 类似地, 计算目标点云和源点云之间的相似信息和差异信息并进行融合, 然后对目标点云中的信息进行交互增强, 从而得到目标点云的特征表示为 $H_Y^{(l')} \in \mathbb{R}^{M \times D}$, 其中, $l' = 1, 2, \dots, L'$.

经由上述 4 个阶段的协同作用, 注意力驱动的互补融合模块能够充分挖掘两个点云之间的相似信息和差异信息, 同时在全局范围内增强特征的融合与交互, 从而在匹配阶段实现更加精确的点云对齐.

1.5 变换估计与损失函数

在获得源点云和目标点云的特征后, 采用文献 [33] 中的方法来计算匹配矩阵 M :

$$M = \text{Sinkhorn}(H_X W H_Y^T) \quad (21)$$

其中, W 是一个可学习参数矩阵, $\text{Sinkhorn}(\cdot)$ 表示 Sinkhorn 算法. 之后, 再使用 SVD 算法来求解变换参数 $\{R, t\}$.

在训练过程中, 本文利用真实匹配矩阵 $\hat{M} = (\hat{M}_{ij})_{N \times M}$ 与估计匹配矩阵 $M = (M_{ij})_{N \times M}$ 之间的交叉熵损失作为损失函数, 具体可表示为:

$$\text{Loss} = - \sum_i^N \sum_j^M (\hat{M}_{ij} \log M_{ij} + (1 - \hat{M}_{ij}) \log(1 - M_{ij})) \quad (22)$$

2 实验分析

2.1 数据集

为了全面评估所提出的方法在点云配准任务中的性能, 本文在 3 个公共数据集上进行了实验.

(1) ModelNet40^[40] 是 3D 点云配准领域的经典基准数据集, 包含 40 个不同类别的 12 311 个高质量 CAD 模型. 其中, 训练集和测试集分别由 9 843 个和 2 468 个点云样本组成.

(2) ShapeNet Part 数据集^[41] 涵盖 16 个对象类别, 共包含 16 881 个样本, 每个样本由 2 000 多个点组成. 训练集

包含 14 007 个点云数据, 而测试集则包含 2 874 个点云数据.

(3) FAUST 数据集^[42]包含 10 个不同人体的 3D 模型, 每个受试者在 30 种不同姿势下进行扫描, 每种姿势的点云大约由 80 000 个点组成.

针对上述数据集, 本文采用与文献 [17,33] 类似的方式, 将训练集的 80% 作为实际训练集, 剩余的 20% 作为验证集. 对于每个点云, 随机选择 1 024 个点并打乱其顺序来生成源点云. 接着, 在 $[0^\circ, 45^\circ]$ 范围内随机生成欧拉角旋转, 并沿着每个轴在 $[-0.5, 0.5]$ 之间生成平移变换. 随后, 将这些变换作用于源点云, 并打乱顺序以生成对应的目标点云. 为了模拟现实环境中的点云数据, 对每个点云添加由 $N(0, 0.01^2)$ 分布产生的高斯噪声. 在此基础上, 从源点云和目标点云中分别去除 30% 的点, 以模拟部分遮挡的点云场景. 此外, 在未知形状的情况下, 使用所有类别的样本进行训练和测试; 在未知类别的情况下, 仅选择前 20 个类别的样本进行训练, 并使用剩余 20 个类别的样本进行测试.

2.2 实验设置

所有实验均在配置单个 NVIDIA RTX 2080Ti GPU 的设备上使用 PyTorch 框架实现的. 对于完整点云配准, 使用 Adam 优化器对网络进行 50 次训练. 在训练过程中, 将初始学习率设置为 1×10^{-3} , 衰减率为 0.2, 并逐步下降至 4×10^{-5} . 对于部分点云配准, 训练回合数设为 100 轮, 并在第 40 轮、60 轮以及 80 轮时进行学习率调整.

表 1 提供了本文所提出模型的详细设置. 具体来说, 选择点云的坐标作为输入的点特征, 并将特征提取阶段的模块堆叠层数设置为 $L = 4$, 匹配阶段的模块堆叠层数设置为 $L' = 2$. 在动态偏移的局部图学习模块中, 每个点的邻域大小为 $K = 24$, 所生成的特征维数为 $C = 128$. 在动态关注的全局图学习模块中, 高斯核超参数设置为 $\sigma = 0.8$, 所生成的特征维数为 $C = 128$. 因此, 特征匹配阶段的特征维度为 $D = 512$.

表 1 LGCNet 的详细设置

模块名称	输入维度	操作	输出维度
输入层	$N \times 3$	—	—
动态偏移的局部图学习 ($L = 1$)	$N \times 3$	图卷积 (128) & MLP (128)	$N \times 128$
动态关注的全局图学习 ($L = 1$)	$N \times 128$	图卷积 (128)	$N \times 128$
动态偏移的局部图学习 ($L = 2$)	$N \times 128$	图卷积 (128) & MLP (128)	$N \times 128$
动态关注的全局图学习 ($L = 2$)	$N \times 128$	图卷积 (128)	$N \times 128$
动态偏移的局部图学习 ($L = 3$)	$N \times 256$	跳跃连接 & 图卷积 (128) & MLP (128)	$N \times 128$
动态关注的全局图学习 ($L = 3$)	$N \times 128$	图卷积 (128)	$N \times 128$
动态偏移的局部图学习 ($L = 4$)	$N \times 256$	跳跃连接 & 图卷积 (128) & MLP (128)	$N \times 128$
动态关注的全局图学习 ($L = 4$)	$N \times 128$	图卷积 (128)	$N \times 128$
整合层	$N \times 512$	跳跃连接 & MLP (512)	$N \times 512$
注意力驱动的互补融合 ($L' = 1$)	$N \times 512$	注意力机制 (512) & 前馈网络 (512)	$N \times 512$
注意力驱动的互补融合 ($L' = 2$)	$N \times 512$	注意力机制 (512) & 前馈网络 (512)	$N \times 512$
变换估计	$N \times 512$	SVD	$3 \times 3, 3 \times 1$

2.3 性能评估

为了验证所提出方法的配准性能, 本文将其与以下基于深度学习的端到端的点云配准方法进行比较: DCP^[23]、IDAM-GNN^[18]、RPM-Net^[17]、DeepBBS^[43]、PREDATOR^[25]、FNet^[32]、RegTR^[34]、OGMM^[35]、RGM^[33]、GeoTransformer^[20]、IFNet^[22]和 GNN-GSSC^[44]. 为确保对比实验的公平性, 在相同的环境下对所有模型采用相同的数据处理方式进行重新训练和测试. 同时, 与文献 [23,33] 一样, 本文针对旋转矩阵和平移向量分别计算平均绝对误差 (MAE)、平均各项同性误差 (MIE) 和均方根误差 (RMSE), 以衡量模型的性能. 这些误差度量的值越小, 表示配准的效果越好.

表 2 和表 3 分别展示了在未知形状模式下完整点云和部分点云的配准结果. 在表中所示的结果中, 加粗表示最优, 下划线表示次优, 下同. 对于带有高斯噪声的完整点云, 本文方法在所有评估指标上均展现出最佳性能. 与 DeepBBS 相比, 尽管在平移指标上的表现相当, 但在旋转指标上优势更加显著. 这主要归因于本文方法在特征提取阶段有效地学习了局部和全局的图结构, 从而在配准过程中更准确地捕获旋转变换的细微差异. 对于部分点云配准, 本文所提出的方法在所有评价指标上都取得了优异的表现. 同时, 从图 5 的可视化结果来看, 预测点云在整体结构和细节上都与目标点云高度接近, 这表明了该方法能够有效获取高判别性的特征, 精准对齐部分重叠的区域, 从而实现高质量的配准. 总体来说, 本文方法对处理未知形状点云的配准任务中展现出明显的优势.

表 2 完整点云在未知形状上的定量对比

方法	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
DCP	2.3913	0.0059	4.6454	0.0117	3.7204	0.0084
IDAM-GNN	0.9177	0.0043	1.8219	0.0086	2.6800	0.0094
RPM-Net	0.3051	0.0028	0.5835	0.0058	0.8430	0.0064
DeepBBS	0.2951	<u>0.0006</u>	0.5826	<u>0.0011</u>	2.1727	<u>0.0009</u>
PREDATOR	0.8879	0.0077	1.7173	0.0159	2.2186	0.0164
FINet	2.6161	0.0288	4.9926	0.0582	3.9298	0.0385
RegTR	0.4580	0.0035	0.9130	0.0075	1.5475	0.0105
OGMM	1.3503	0.0083	2.6833	0.0174	5.5197	0.0349
RGM	0.0772	0.0007	0.1437	0.0014	0.3411	0.0019
GeoTransformer	0.3166	0.0030	0.6264	0.0062	0.3718	0.0036
IFNet	0.1713	0.0007	0.3764	0.0014	1.9055	0.0059
GNN-GSSC	<u>0.0756</u>	<u>0.0006</u>	<u>0.1423</u>	0.0014	<u>0.3185</u>	0.0025
本文方法	0.0601	0.0006	0.1104	0.0011	0.0971	0.0008

表 3 部分点云在未知形状上的定量对比

方法	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
DCP	5.0422	0.0607	9.7294	0.1204	6.9151	0.0785
IDAM-GNN	8.3604	0.0882	15.8371	0.1750	14.4293	0.1388
RPM-Net	0.7693	0.0070	1.5085	0.0149	1.9153	0.0156
DeepBBS	1.5563	0.0164	3.1409	0.0327	5.1110	0.0447
PREDATOR	1.0662	0.0097	2.0629	0.0203	2.4908	0.0239
FINet	2.8004	0.0278	5.4249	0.0585	4.6450	0.0468
RegTR	0.8094	0.0063	1.6013	0.0133	2.1321	<u>0.0154</u>
OGMM	4.8156	0.0993	9.6630	0.2319	8.1634	0.1662
RGM	0.5395	0.0042	1.0564	0.0090	3.9282	0.0207
GeoTransformer	0.6618	0.0071	1.3123	0.0151	<u>0.7757</u>	0.0087
IFNet	4.0893	0.0293	7.9760	0.0580	9.2278	0.0452
GNN-GSSC	<u>0.3527</u>	<u>0.0026</u>	<u>0.6502</u>	<u>0.0055</u>	2.8846	0.0156
本文方法	0.1479	0.0014	0.2772	0.0027	0.2757	0.0057

在未知类别模式下, 完整点云和部分点云的配准结果分别如表 4 和表 5 所示. 针对带噪声的完整点云和部分点云, 本文方法在所有评价指标上均表现出最佳性能. 这不仅得益于特征提取阶段对局部和全局图结构的有效学习, 还归功于特征匹配阶段对两个点云之间的相似信息和差异信息的充分挖掘与融合. 同样, 图 5 的可视化结果也反映了本文方法在处理复杂场景时具有较好的视觉效果, 能够准确识别出点云之间的对应关系.

为了进一步验证本文所提出方法的有效性, 在 ShapeNet Part 和 FAUST 数据集进行了泛化实验, 结果如表 6 所示. 可以发现, 本文方法在 ShapeNet Part 数据集上达到最好的效果, 在 FAUST 数据集上实现次优的泛化性能. 尽管性能略低于 GNN-GSSC, 但显著降低了模型的复杂度和参数量, 这表明了本文方法在保持低计算成本的同时, 仍然具有在处理未知点云数据方面的卓越泛化能力.

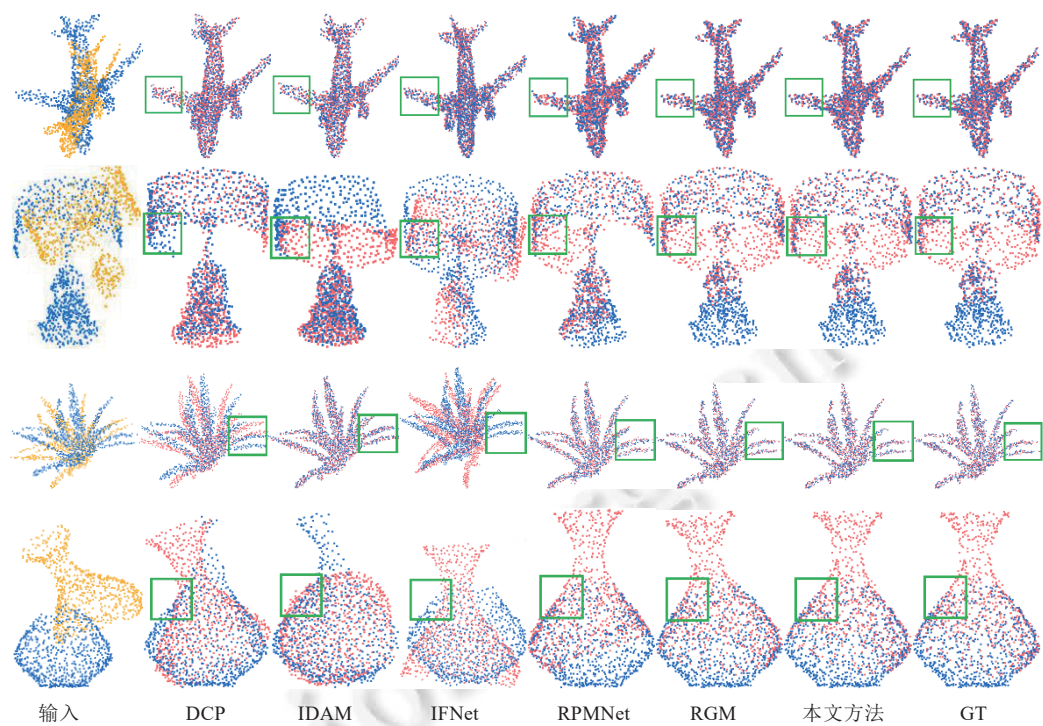


图 5 不同方法在 ModelNet40 数据集上的视觉效果比较

表 4 完整点云在未知类别上的定量对比

方法	MAE (<i>R</i>)	MAE (<i>t</i>)	MIE (<i>R</i>)	MIE (<i>t</i>)	RMSE (<i>R</i>)	RMSE (<i>t</i>)
DCP	3.0008	0.0122	5.9035	0.0242	4.8209	0.0171
IDAM-GNN	0.7911	0.0040	1.5492	0.0081	2.1472	0.0074
RPM-Net	0.3063	0.0027	0.5837	0.0055	0.8343	0.0045
DeepBBS	0.5870	0.0008	1.2009	0.0016	4.3374	0.0052
PREDATOR	0.7602	0.0069	1.4492	0.0140	1.5162	0.0119
FINet	0.9931	0.0095	2.0343	0.0200	2.8182	0.0204
RegTR	0.6336	0.0044	1.2593	0.0094	2.6207	0.0147
OGMM	1.4310	0.0087	2.9116	0.0191	7.3972	0.0437
RGM	0.0796	0.0007	0.1454	0.0014	0.3057	0.0026
GeoTransformer	0.5818	0.0045	1.1160	0.0091	0.6659	0.0053
IFNet	0.1952	0.0010	0.4200	0.0020	1.7149	0.0076
GNN-GSSC	<u>0.0684</u>	<u>0.0006</u>	<u>0.1265</u>	<u>0.0012</u>	<u>0.2440</u>	<u>0.0015</u>
本文方法	0.0533	0.0005	0.0967	0.0010	0.0748	0.0005

表 5 部分点云在未知类别上的定量对比

方法	MAE (<i>R</i>)	MAE (<i>t</i>)	MIE (<i>R</i>)	MIE (<i>t</i>)	RMSE (<i>R</i>)	RMSE (<i>t</i>)
DCP	6.3763	0.0667	12.5257	0.1330	8.6278	0.0852
IDAM-GNN	7.2938	0.0924	13.8058	0.1827	11.8599	0.1347
RPM-Net	0.8204	0.0080	1.5723	0.0167	1.9560	0.0175
DeepBBS	2.2858	0.0248	4.5133	0.0493	7.9844	0.0617
PREDATOR	1.1024	0.0110	2.1001	0.0232	2.6823	0.0304
FINet	3.7333	0.0406	7.2068	0.0853	5.7400	0.0632
RegTR	0.9310	0.0081	1.7959	0.0171	2.2922	0.0206
OGMM	7.2901	0.0861	14.4825	0.1894	12.3177	0.1375

表 5 部分点云在未知类别上的定量对比 (续)

方法	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
RGM	0.9374	0.0076	1.7366	0.0159	3.4673	0.0291
GeoTransformer	1.0748	0.0107	2.0803	0.0228	<u>1.2355</u>	0.0132
IFNet	3.8955	0.0302	7.5285	0.0595	8.7858	0.0463
GNN-GSSC	<u>0.2799</u>	<u>0.0024</u>	<u>0.5466</u>	<u>0.0048</u>	1.3519	<u>0.0108</u>
本文方法	0.1388	0.0012	0.2611	0.0024	0.4557	0.0027

表 6 在 ShapeNet Part 和 FAUST 数据集上的泛化对比

数据集	方法	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
ShapeNet Part	DCP	9.1382	0.0847	18.0784	0.1693	13.2157	0.1084
	IDAM-GNN	7.6985	0.0813	14.7016	0.1626	12.5218	0.1176
	RPM-Net	2.9505	0.0264	5.7783	0.0527	5.5754	0.0420
	DeepBBS	4.3841	0.0412	8.5682	0.0824	13.7924	0.0857
	PREDATOR	1.2682	0.0092	2.5553	0.0183	3.9301	0.0354
	FINet	3.5014	0.0329	6.8544	0.0680	5.2333	0.0473
	RegTR	0.6017	0.0050	1.1881	0.0105	1.8163	0.0122
	OGMM	1.4527	0.0089	2.8275	0.0188	6.5023	0.0362
	RGM	1.8351	0.0145	3.5402	0.0312	8.0900	0.0506
	GeoTransformer	0.5662	0.0054	1.0955	0.0113	<u>0.6522</u>	<u>0.0067</u>
	IFNet	1.6007	0.0112	3.0485	0.0221	6.9232	0.0335
	GNN-GSSC	<u>0.2064</u>	<u>0.0015</u>	<u>0.4110</u>	<u>0.0033</u>	1.3374	0.0098
	本文方法	0.1467	0.0013	0.2722	0.0026	0.2806	0.0035
FAUST	DCP	9.2065	0.0855	18.6825	0.1687	12.3396	0.1084
	IDAM-GNN	7.9601	0.1048	15.0701	0.2048	11.1200	0.1489
	RPM-Net	3.8772	0.0255	7.5757	0.0499	5.8126	0.0367
	DeepBBS	5.2267	0.0461	9.7946	0.0916	13.6302	0.0895
	PREDATOR	1.9391	0.0173	3.6243	0.0362	3.8604	0.0493
	FINet	3.5709	0.0332	6.9948	0.0685	5.3451	0.0474
	RegTR	0.7143	0.0057	1.3673	0.0116	0.9825	0.0079
	OGMM	0.5311	0.0035	1.0597	0.0070	6.3270	0.0310
	RGM	1.1356	0.0085	2.1849	0.0167	3.1746	0.0233
	GeoTransformer	0.4455	0.0042	0.8784	0.0096	0.5204	0.0053
	IFNet	3.1007	0.1170	4.2095	0.2359	5.4971	0.1687
	GNN-GSSC	0.1024	0.0008	0.1902	0.0016	0.1398	0.0011
	本文方法	<u>0.2303</u>	<u>0.0018</u>	<u>0.4285</u>	<u>0.0036</u>	<u>0.3176</u>	<u>0.0026</u>

2.4 消融实验

为了验证所提出 3 个模块的有效性, 本文以 ModelNet40 数据集为例, 针对未知类别的部分点云配准进行如下消融实验: (I) 移除动态偏移的局部图学习模块; (II) 移除动态关注的全局图学习模块; (III) 移除注意力驱动的互补融合模块. 此外, 为验证注意力驱动的互补融合模块中差异信息的重要性, 以及差异信息与相似信息协同作用的有效性, 还设计了以下 3 个消融实验: (IV) 移除基于交叉注意力的差异信息提取步骤; (V) 在模型 IV 的基础上, 将源点云和目标点云进行拼接融合; (VI) 在模型 IV 的基础上, 将源点云和目标点云进行加性融合.

实验结果如表 7 所示. 可以发现, 在模型 I 中, 移除动态偏移的局部图学习模块, 性能显著下降. 这表明该模块在特征提取阶段发挥了关键作用, 尤其是在捕获点云的局部细节方面. 相比于固定图结构的特征学习方式, 该模块通过动态偏移策略使点云特征能够在局部范围内进行自适应调整, 从而更有效地捕获关键几何信息. 这一策略不仅提升了配准精度, 还为点云处理领域提供了一种新的思路, 即通过局部自适应学习来增强特征表征能力. 在模型 II 中, 去除动态关注的全局图学习模块, 模型性能略微降低, 说明该模块也发挥了重要的作用. 在模型 III 中, 删

除注意力驱动的互补融合模块, 配准效果有所下降, 说明点云之间的交互对于特征匹配至关重要. 此外, 模型 IV 说明了差异信息能够有效捕获点云之间的细微差异并缓解匹配模糊问题. 模型 V 和模型 VI 的结果都出现了明显下降, 由此可见, 单独使用源点云和目标点云的融合方法无法充分挖掘差异信息和相似信息之间的互补性, 限制了配准效果的提升. 综上所述, 本文所设计的模块都发挥了重要作用, 有利于提高点云配准的准确性.

表 7 不同消融实验在 ModelNet40 数据集上的定量比较

模型	MAE (<i>R</i>)	MAE (<i>t</i>)	MIE (<i>R</i>)	MIE (<i>t</i>)	RMSE (<i>R</i>)	RMSE (<i>t</i>)
I	1.0429	0.0095	1.9317	0.0198	4.7818	0.0388
II	0.1664	0.0017	0.3151	0.0034	0.5128	0.0084
III	0.3210	0.0029	0.6015	0.0060	1.4448	0.0125
IV	0.1936	0.0016	0.3675	0.0033	1.1596	0.0096
V	0.3267	0.0028	0.6408	0.0057	1.4802	0.0130
VI	0.2889	0.0025	0.5476	0.0053	1.1552	0.0110
本文方法	0.1388	0.0012	0.2611	0.0024	0.4557	0.0027

2.5 超参数讨论

本节主要讨论不同超参数的设置对模型性能的影响, 包括动态偏移的局部图学习模块中的最近邻数量 K , 动态偏移的局部图学习和动态关注的全局图学习模块中的高斯核参数 σ , 特征提取阶段局部-全局图学习模块堆叠的层数 L 以及特征匹配阶段注意力驱动的互补融合模块堆叠的层数 L' .

(1) 最近邻数量 K . 本文分别对 K 为 16、24 和 32 的情况进行了实验, 结果如表 8 和图 6(a) 所示. 当 K 较小时, 局部图中的信息相对有限, 可能导致特征提取能力不足. 当 K 较大时, 部分点包含较多与中心点无关的信息, 聚合后容易造成特征描述的模糊性. 因此, 本文经验性地设置 $K = 24$.

表 8 最近邻数量 K 的影响

K	MAE (<i>R</i>)	MAE (<i>t</i>)	MIE (<i>R</i>)	MIE (<i>t</i>)	RMSE (<i>R</i>)	RMSE (<i>t</i>)
16	0.1502	0.0096	0.2755	0.0029	0.8184	0.0096
24	0.1388	0.0012	0.2611	0.0024	0.4557	0.0027
32	0.3606	0.0031	0.6712	0.0064	2.7045	0.0214

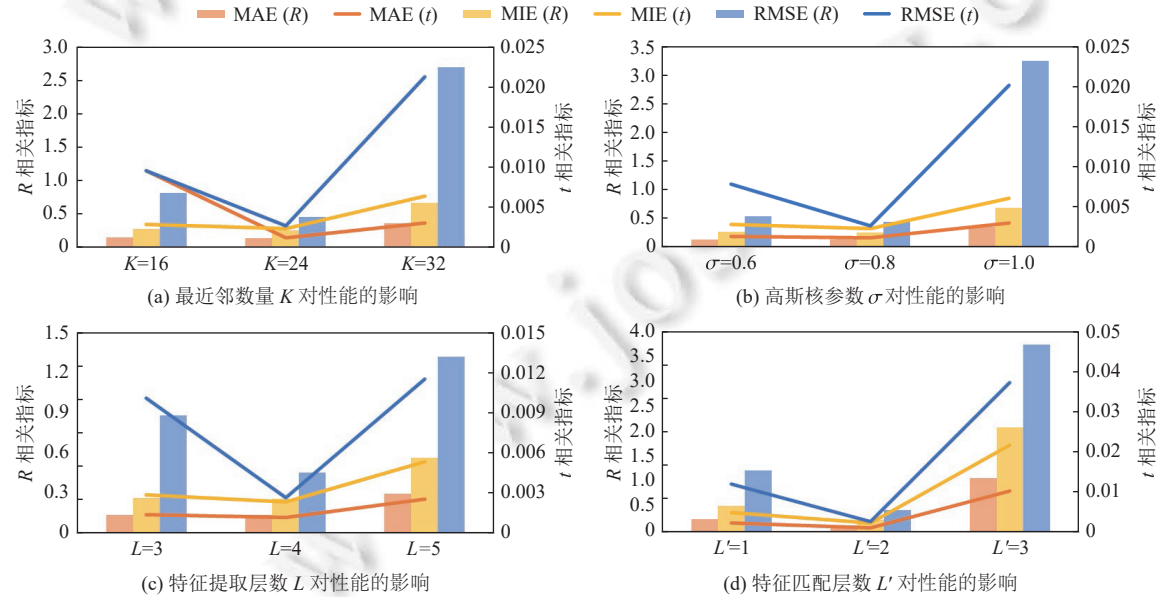


图 6 不同超参数在 ModelNet40 数据集上的影响

(2) 高斯核参数 σ . 本文经验性地将高斯核参数 σ 设置为 0.6、0.8 和 1 进行实验分析. 从表 9 和图 6(b) 中可以看出, 当 σ 较小时, 全局图的构造过于关注少数强相关的点对, 导致信息覆盖范围不足, 限制了全局特征的表达能力. 当 σ 较大时, 大量低相关性的点对被赋予较高的权重, 反而引入了较多噪声. 因此, 本文将高斯核参数 σ 设置为 0.8.

表 9 高斯核参数 σ 的影响

σ	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
0.6	0.1407	0.0014	0.2760	0.0029	0.5467	0.0080
0.8	0.1388	0.0012	0.2611	0.0024	0.4557	0.0027
1	0.3655	0.0031	0.6987	0.0062	3.2791	0.0204

(3) 特征提取层数 L . 在保持其他超参数不变的情况下, 将特征提取层数 L 从 3 层逐渐增加至 5 层. 如表 10 和图 6(c) 所示, 当 $L = 3$ 时, 模型已经能够取得较好的配准效果. 当 $L = 4$ 时, 结果进一步提升, 说明增加层数有助于捕获更加丰富的特征表示. 然而, 当 $L = 5$ 时, 模型性能出现下降, 可能是由于堆叠层数过多导致的特征过平滑问题. 综上所述, 本文选择层数 L 为 4 以实现最佳性能.

表 10 特征提取层数 L 的影响

L	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
3	0.1415	0.0014	0.2678	0.0029	0.8850	0.0102
4	0.1388	0.0012	0.2611	0.0024	0.4557	0.0027
5	0.3003	0.0026	0.5677	0.0054	1.3265	0.0116

(4) 特征匹配层数 L' . 针对注意力驱动的互补融合模块堆叠的层数 L' 分别为 1、2 和 3 的情况, 如表 11 和图 6(d) 所示. 实验结果表明, 当层数 $L' = 2$ 时, 模型性能达到最佳. 当 $L' = 1$ 时, 模型在挖掘与融合差异信息和相似信息方面能力不足, 限制了性能的进一步提升. 而当层数增加至 3 层时, 由于模型过于复杂, 容易引入信息冗余和过拟合等问题, 进而对配准结果产生负面影响. 综合考虑, 本文选择将层数 L' 设置为 2 层.

表 11 特征匹配层数 L' 的影响

L'	MAE (R)	MAE (t)	MIE (R)	MIE (t)	RMSE (R)	RMSE (t)
1	0.2789	0.0025	0.5368	0.0050	1.2572	0.0123
2	0.1388	0.0012	0.2611	0.0024	0.4557	0.0027
3	1.1041	0.0106	2.1341	0.0222	3.8298	0.0382

2.6 计算复杂度分析

本文在 ModelNet40 数据集上对所提出的方法与现有方法进行了计算复杂度的全面比较, 包括参数量、FLOPs、训练时间、测试时间等方面. 从表 12 可以看出, 本文方法的参数量和 FLOPs 略高于 DCP、DeepBBS 和 RegTR, 但显著低于 RGM 和 GNN-GSSC. 同时, 训练时间和测试时间均保持在可接受的范围内, 但在性能上取得了显著提升. 综上所述, 本文所提出的方法在性能和计算效率之间达到了良好的平衡.

表 12 不同方法的计算复杂度

方法	参数量 (G)	FLOPs (M)	训练时间 (s)	测试时间 (s)
DCP	10.91	5.34	0.0528	0.0127
IDAM-GNN	1.55	0.09	0.0221	0.0096
RPM-Net	8.48	0.91	0.0519	0.0443
DeepBBS	13.18	5.67	0.0783	0.1533
PREDATOR	—	11.01	0.0638	0.0530
FINet	15.85	14.22	0.0829	0.0517
RegTR	25.1	5.48	0.0957	0.0418

表 12 不同方法的计算复杂度 (续)

方法	参数量 (G)	FLOPs (M)	训练时间 (s)	测试时间 (s)
OGMM	10.09	4.58	0.052 5	0.027 6
RGM	51.83	24.46	0.102 1	0.043 2
GeoTransformer	—	4.49	0.069 4	0.032 8
IFNet	32.63	5.50	0.096 9	0.065 5
GNN-GSSC	53.98	19.20	0.103 8	0.044 8
本文方法	26.22	13.21	0.109 8	0.037 1

3 总 结

本文提出了一种局部-全局动态图学习与互补融合的点云配准方法, 通过结合动态偏移的局部图学习、动态关注的全局图学习和注意力驱动的互补融合模块, 有效地提升了点云配准的精度. 在特征提取阶段, 根据动态偏移得到的代理点来学习特征空间的图结构, 获得了更具判别性和细节增强的局部特征表示. 同时, 利用特征学习的自适应阈值动态地构建全局图, 进一步实现了点之间的长程依赖关系. 在特征匹配阶段, 注意力驱动的互补融合模块通过交叉注意力机制提取相似和差异信息, 并采用自注意力机制进一步优化了特征的关联性. 实验结果表明, 本文方法取得了最优的配准效果, 同时保持了较高的计算效率. 在未来的研究中, 可以针对不同场景进一步优化算法, 以提高其在复杂环境中的鲁棒性, 特别是对高噪声与部分遮挡点云的处理.

References:

[1] Dang Z, Wang LZ, Guo Y, Salzmann M. Match normalization: Learning-based point cloud registration for 6D object pose estimation in the real world. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2024, 46(6): 4489–4503. [doi: [10.1109/TPAMI.2024.3355198](https://doi.org/10.1109/TPAMI.2024.3355198)]

[2] Jiang HB, Salzmann M, Dang Z, Xie J, Yang J. SE(3) diffusion model-based point cloud registration for robust 6D object pose estimation. In: *Proc. of the 37th Int'l Conf. on Neural Information Processing Systems*. New Orleans: Curran Associates Inc., 2023. 21285–21297.

[3] Shi CH, Chen XYL, Lu HM, Deng WB, Xiao JH, Dai B. RDMNet: Reliable dense matching based point cloud registration for autonomous driving. *IEEE Trans. on Intelligent Transportation Systems*, 2023, 24(10): 11372–11383. [doi: [10.1109/TITS.2023.3286464](https://doi.org/10.1109/TITS.2023.3286464)]

[4] Zhou XY, Lin ZW, Shan XJ, Wang YT, Sun DQ, Yang MH. DrivingGaussian: Composite Gaussian splatting for surrounding dynamic autonomous driving scenes. In: *Proc. of the 2024 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024. 21634–21643. [doi: [10.1109/CVPR52733.2024.02044](https://doi.org/10.1109/CVPR52733.2024.02044)]

[5] Li GF, Fan HW, Jiang GZ, Jiang D, Liu YT, Tao B, Yun JT. RGBD-SLAM based on object detection with two-stream YOLOv4-MobileNetv3 in autonomous driving. *IEEE Trans. on Intelligent Transportation Systems*, 2024, 25(3): 2847–2857. [doi: [10.1109/TITS.2023.3284228](https://doi.org/10.1109/TITS.2023.3284228)]

[6] Wang YN, Tian YB, Chen JW, Xu K, Ding XL. A survey of visual SLAM in dynamic environment: The evolution from geometric to semantic approaches. *IEEE Trans. on Instrumentation and Measurement*, 2024, 73: 2523221. [doi: [10.1109/TIM.2024.3420374](https://doi.org/10.1109/TIM.2024.3420374)]

[7] Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis & Machine Intelligence*, 1992, 14(2): 239–256. [doi: [10.1109/34.121791](https://doi.org/10.1109/34.121791)]

[8] Segal A, Haehnel D, Thrun S. Generalized-ICP. In: *Proc. of Robotics: Science and Systems*. 2009. [doi: [10.15607/RSS.2009.V.021](https://doi.org/10.15607/RSS.2009.V.021)]

[9] Yang JL, Li HD, Jia YD. Go-ICP: Solving 3D registration efficiently and globally optimally. In: *Proc. of the 2013 IEEE Int'l Conf. on Computer Vision*. Sydney: IEEE, 2013. 1457–1464. [doi: [10.1109/ICCV.2013.184](https://doi.org/10.1109/ICCV.2013.184)]

[10] Magnusson M, Lilienthal A, Duckett T. Scan registration for autonomous mining vehicles using 3D-NDT. *Journal of Field Robotics*, 2007, 24(10): 803–827. [doi: [10.1002/rob.20204](https://doi.org/10.1002/rob.20204)]

[11] Aiger D, Mitra NJ, Cohen-Or D. 4-points congruent sets for robust pairwise surface registration. *ACM Trans. on Graphics*, 2008, 27(3): 1–10. [doi: [10.1145/1360612.1360684](https://doi.org/10.1145/1360612.1360684)]

[12] Aoki Y, Goforth H, Srivatsan RA, Lucey S. PointNetLK: Robust & efficient point cloud registration using PointNet. In: *Proc. of the 2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 7156–7165. [doi: [10.1109/CVPR.2019.00733](https://doi.org/10.1109/CVPR.2019.00733)]

[13] Qi CR, Su H, Mo KC, Guibas LJ. PointNet: Deep learning on point sets for 3D classification and segmentation. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 652–660. [doi: [10.1109/CVPR.2017.16](https://doi.org/10.1109/CVPR.2017.16)]

- [14] Xu H, Liu SC, Wang GF, Liu GH, Zeng B. OMNet: Learning overlapping mask for partial-to-partial point cloud registration. In: Proc. of the 2021 IEEE/CVF Int'l Conf. on Computer Vision. Montreal: IEEE, 2021. 3112–3121. [doi: [10.1109/ICCV48922.2021.00312](https://doi.org/10.1109/ICCV48922.2021.00312)]
- [15] Lu WX, Wan GW, Zhou Y, Fu XY, Yuan PF, Song SY. DeepVCP: An end-to-end deep neural network for point cloud registration. In: Proc. of the 2019 IEEE/CVF Int'l Conf. on Computer Vision. Seoul: IEEE, 2019. 12–21. [doi: [10.1109/ICCV.2019.00010](https://doi.org/10.1109/ICCV.2019.00010)]
- [16] Qi CR, Yi L, Su H, Guibas LJ. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 5105–5114.
- [17] Yew ZJ, Lee GH. RPM-Net: Robust point matching using learned features. In: Proc. of the 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 11821–11830. [doi: [10.1109/CVPR42600.2020.01184](https://doi.org/10.1109/CVPR42600.2020.01184)]
- [18] Li JH, Zhang CH, Xu ZY, Zhou HN, Zhang C. Iterative distance-aware similarity matrix convolution with mutual-supervised point elimination for efficient point cloud registration. In: Proc. of the 16th European Conf. on Computer Vision. Glasgow: Springer, 2020. 378–394. [doi: [10.1007/978-3-030-58586-0_23](https://doi.org/10.1007/978-3-030-58586-0_23)]
- [19] Mei Q, Wang FX, Tong C, Zhang JQ, Jiang B, Xiao J. PACNet: A high-precision point cloud registration network based on deep learning. In: Proc. of the 13th Int'l Conf. on Wireless Communications and Signal Processing. Changsha: IEEE, 2021. 1–5. [doi: [10.1109/WCSP52459.2021.9613579](https://doi.org/10.1109/WCSP52459.2021.9613579)]
- [20] Qin Z, Yu H, Wang CJ, Guo YL, Peng YX, Ilic S, Hu DW, Xu K. GeoTransformer: Fast and robust point cloud registration with geometric Transformer. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2023, 45(8): 9806–9821. [doi: [10.1109/TPAMI.2023.3259038](https://doi.org/10.1109/TPAMI.2023.3259038)]
- [21] Slimani K, Achard C, Tamadazte B. RoCNet++: Triangle-based descriptor for accurate and robust point cloud registration. Pattern Recognition, 2024, 147: 110108. [doi: [10.1016/j.patcog.2023.110108](https://doi.org/10.1016/j.patcog.2023.110108)]
- [22] Xie YF, Wang BY, Li SQ, Zhu JH. Iterative feedback network for unsupervised point cloud registration. IEEE Robotics and Automation Letters, 2024, 9(3): 2327–2334. [doi: [10.1109/LRA.2024.3355784](https://doi.org/10.1109/LRA.2024.3355784)]
- [23] Wang Y, Solomon J. Deep closest point: Learning representations for point cloud registration. In: Proc. of the 2019 IEEE/CVF Int'l Conf. on Computer Vision. Seoul: IEEE, 2019. 3522–3531. [doi: [10.1109/ICCV.2019.00362](https://doi.org/10.1109/ICCV.2019.00362)]
- [24] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- [25] Huang SY, Gojcic Z, Usvyatsov M, Wieser A, Schindler K. PREDATOR: Registration of 3D point clouds with low overlap. In: Proc. of the 2021 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 4265–4274. [doi: [10.1109/CVPR46437.2021.00425](https://doi.org/10.1109/CVPR46437.2021.00425)]
- [26] Fischer K, Simon M, Ölsner F, Milz S, Groß HM, Mäder P. StickyPillars: Robust and efficient feature matching on point clouds using graph neural networks. In: Proc. of the 2021 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 313–323. [doi: [10.1109/CVPR46437.2021.00038](https://doi.org/10.1109/CVPR46437.2021.00038)]
- [27] Shi CH, Chen XYL, Huang KH, Xiao JH, Lu HM, Stachniss C. Keypoint matching for point cloud registration using multiplex dynamic graph attention networks. IEEE Robotics and Automation Letters, 2021, 6(4): 8221–8228. [doi: [10.1109/LRA.2021.3097275](https://doi.org/10.1109/LRA.2021.3097275)]
- [28] Kadam P, Zhang M, Liu S, Kuo CCJ. R-PointHop: A green, accurate, and unsupervised point cloud registration method. IEEE Trans. on Image Processing, 2022, 31: 2710–2725. [doi: [10.1109/TIP.2022.3160609](https://doi.org/10.1109/TIP.2022.3160609)]
- [29] Chen GY, Wang ML, Zhang QX, Yuan L, Yue YF. Full Transformer framework for robust point cloud registration with deep information interaction. IEEE Trans. on Neural Networks and Learning Systems, 2024, 35(10): 13368–13382. [doi: [10.1109/TNNLS.2023.3267333](https://doi.org/10.1109/TNNLS.2023.3267333)]
- [30] Jiang YN, Zhou BT, Liu XY, Li QY, Cheng C. GTINet: Global topology-aware interactions for unsupervised point cloud registration. IEEE Trans. on Circuits and Systems for Video Technology, 2024, 34(7): 6363–6375. [doi: [10.1109/TCSVT.2024.3367529](https://doi.org/10.1109/TCSVT.2024.3367529)]
- [31] Wu BL, Ma J, Chen GJ, An P. Feature interactive representation for point cloud registration. In: Proc. of the 2021 IEEE/CVF Int'l Conf. on Computer Vision. Montreal: IEEE, 2021. 5510–5519. [doi: [10.1109/ICCV48922.2021.00548](https://doi.org/10.1109/ICCV48922.2021.00548)]
- [32] Xu H, Ye NJ, Liu GH, Zeng B, Liu SC. FINet: Dual branches feature interaction for partial-to-partial point cloud registration. In: Proc. of the 36th AAAI Conf. on Artificial Intelligence. Virtually: AAAI, 2022. 2848–2856. [doi: [10.1609/aaai.v36i3.20189](https://doi.org/10.1609/aaai.v36i3.20189)]
- [33] Fu KX, Luo JZ, Luo XY, Liu SL, Zhang CX, Wang MN. Robust point cloud registration framework based on deep graph matching. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2023, 45(5): 6183–6195. [doi: [10.1109/TPAMI.2022.3204713](https://doi.org/10.1109/TPAMI.2022.3204713)]
- [34] Yew ZJ, Lee GH. REGTR: End-to-end point cloud correspondences with Transformers. In: Proc. of the 2022 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 6667–6676. [doi: [10.1109/CVPR52688.2022.00656](https://doi.org/10.1109/CVPR52688.2022.00656)]
- [35] Mei GF, Poiesi F, Saltori C, Zhang J, Ricci E, Sebe N. Overlap-guided Gaussian mixture models for point cloud registration. In: Proc. of the 2023 IEEE/CVF Winter Conf. on Applications of Computer Vision. Waikoloa: IEEE, 2023. 4500–4509. [doi: [10.1109/WACV56688.2023.00449](https://doi.org/10.1109/WACV56688.2023.00449)]

- [36] Fung S, Pan W, Liu X, Yearwood J, Dazeley R, Lu XQ. TopFormer: Topology-aware Transformer for point cloud registration. In: Proc. of the 12th Int'l Conf. on Computational Visual Media. Wellington: Springer, 2024. 112–128. [doi: [10.1007/978-981-97-2095-8_7](https://doi.org/10.1007/978-981-97-2095-8_7)]
- [37] Yu H, Hou J, Qin Z, Saleh M, Shugurov I, Wang K, Busam B, Ilie S. RIGA: Rotation-invariant and globally-aware descriptors for point cloud registration. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2024, 46(5): 3796–3812. [doi: [10.1109/TPAMI.2023.3349199](https://doi.org/10.1109/TPAMI.2023.3349199)]
- [38] Li JW, Zhan JW. Review on 3D point cloud registration method. Journal of Image and Graphics, 2022, 27(2): 349–367 (in Chinese with English abstract). [doi: [10.11834/jig.210243](https://doi.org/10.11834/jig.210243)]
- [39] Wang Y, Sun YB, Liu ZW, Sarma SE, Bronstein MM, Solomon JM. Dynamic graph CNN for learning on point clouds. ACM Trans. on Graphics, 2019, 38(5): 146. [doi: [10.1145/3326362](https://doi.org/10.1145/3326362)]
- [40] Wu ZR, Song SR, Khosla A, Yu F, Zhang LG, Tang XO, Xiao JX. 3D ShapeNets: A deep representation for volumetric shapes. In: Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 1912–1920. [doi: [10.1109/CVPR.2015.7298801](https://doi.org/10.1109/CVPR.2015.7298801)]
- [41] Yi L, Kim VG, Ceylan D, Shen IC, Yan MY, Su H, Lu CW, Huang QX, Sheffer A, Guibas L. A scalable active framework for region annotation in 3D shape collections. ACM Trans. on Graphics, 2016, 35(6): 210. [doi: [10.1145/2980179.2980238](https://doi.org/10.1145/2980179.2980238)]
- [42] Bogo F, Romero J, Loper M, Black MJ. FAUST: Dataset and evaluation for 3D mesh registration. In: Proc. of the 2014 IEEE Conf. on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 3794–3801. [doi: [10.1109/CVPR.2014.491](https://doi.org/10.1109/CVPR.2014.491)]
- [43] Hezroni I, Drory A, Giryas R, Avidan S. DeepBBS: Deep best buddies for point cloud registration. In: Proc. of the 2021 Int'l Conf. on 3D Vision. London: IEEE, 2021. 342–351. [doi: [10.1109/3DV53792.2021.00044](https://doi.org/10.1109/3DV53792.2021.00044)]
- [44] Cao FL, Wang LP, Ye HL. SharpGConv: A novel graph method with plug-and-play sharpening convolution for point cloud registration. IEEE Trans. on Circuits and Systems for Video Technology, 2024, 34(8): 7095–7105. [doi: [10.1109/TCSVT.2024.3369468](https://doi.org/10.1109/TCSVT.2024.3369468)]

附中文参考文献:

- [38] 李建微, 占家旺. 三维点云配准方法研究进展. 中国图象图形学报, 2022, 27(2): 349–367. [doi: [10.11834/jig.210243](https://doi.org/10.11834/jig.210243)]



邱巧燕(1999—), 女, 硕士生, 主要研究领域为深度学习, 点云分析.



曹飞龙(1965—), 男, 博士, 教授, 博士生导师, 主要研究领域为深度学习, 点云分析.



叶海良(1990—), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为深度学习, 点云分析.



吕科(1971—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为计算机视觉, 计算机图形学.