

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2024)12-3739-17

论文引用格式: Yang J and Sun H W. 2024. Edge feature enhancement and hierarchical attention fusion for low-overlap point cloud registration. Journal of Image and Graphics, 29(12):3739-3755(杨军, 孙鸿炜. 2024. 边缘特征增强与层次注意力融合的低重叠点云配准. 中国图象图形学报, 29(12):3739-3755)[DOI:10.11834/jig.230871]

边缘特征增强与层次注意力融合的低重叠点云配准

杨军^{1,2*}, 孙鸿炜¹

1. 兰州交通大学电子与信息工程学院, 兰州 730070; 2. 兰州交通大学测绘与地理信息学院, 兰州 730070

摘要: 目的 针对目前基于深度学习的低重叠度点云配准方法在学习全局点云场景后进行特征匹配时, 忽略局部特征间作用的问题, 提出了一种结合边缘特征增强的层次注意力点云配准方法。方法 首先, 利用边缘自适应核点卷积(edge adaptive kernel point convolution, EAKPConv)模块提取源点云、目标点云特征, 增强边缘特征识别能力。然后, 利用局部空间差异注意模块(local spatial contrast attention module, LSCAM)聚合局部空间差异捕捉点云的几何细节, 利用序列相似度关联模块(sequential similarity association module, SSAM)计算量化两点云间的相似分数, 并利用相似分数引导局部匹配。最后, 通过LSCAM模块与SSAM模块结合的层次化注意力融合模块(hierarchical attention fusion module, HAFM)整合局部、全局特征, 实现全局匹配。结果 在室内场景点云配准数据集3DMatch和三维模型数据集ModelNet-40上进行了对比实验, 本算法在3DMatch和3DLoMatch上的配准召回率分别达到93.2%和67.3%; 在ModelNet-40和ModelLoNet-40上取得了最低的旋转误差(分别为1.417和3.141)以及平移误差(分别为0.013 91和0.072)。此外, 本文算法在推理效率上比REGTR算法减少了10 ms左右。结论 本文算法通过自底向上的层次化处理方式显著提升了有限重叠场景点云的配准精度, 同时降低了推理时间。

关键词: 三维点云配准; 低重叠度点云; 边缘特征; 层次注意力; 局部相似匹配

Edge feature enhancement and hierarchical attention fusion for low-overlap point cloud registration

Yang Jun^{1,2*}, Sun Hongwei¹

1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;

2. Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract: Objective Low overlap point cloud registration presents a significant obstacle in the realm of computer vision, specifically in the context of deep-learning-based approaches. After acquiring knowledge from global point cloud scenes for feature matching, these deep-learning-based methods often fail to consider the interactions among local features, thus greatly impeding the efficiency of registration in settings where local feature interactions are vital for establishing precise alignment. The intricate interplay among local characteristics, which is crucial for accurately identifying and aligning partially intersecting point clouds, is also inadequately represented. This lack of consideration not only affects the reliability of point cloud registration in situations with limited overlap but also restricts the use of deep learning methods in varied and intricate settings. Therefore, techniques that include the comprehension of local feature interactions into the deep learning

收稿日期: 2023-12-29; 修回日期: 2024-04-03; 预印本日期: 2024-04-10

* 通信作者: 杨军 yangj@mail.lzjtu.cn

基金项目: 国家自然科学基金项目(42261067)

Supported by: National Natural Science Foundation of China(42261067)

framework are crucial for point cloud registration, especially in situations with limited overlap. **Method** The present study introduces a novel technique for aligning point clouds with low overlap. This technique uses the edge adaptive KPConv (EAKPConv) module to enhance the identification of edge characteristics. The integration of local and global features is effectively accomplished by the combination of the hierarchical attention fusion module (HAFM) and the local spatial comparison attention module (LSCAM). LSCAM exploits the capacity of the attention mechanism to consolidate information, thus enabling the model to prioritize those connections with target nodes and to position itself near the clustered center of mass. In this way, the model can flexibly capture complex details of the point cloud. The SSAM system utilizes a hierarchical architecture, in which each tier of local matching modules applies its own similarity metric to quantify the similarities among point clouds. The local features are subsequently modified and transmitted to the subsequent layer of attention modules to establish a hierarchical structure. This structure also allows the model to collect and merge the inputs from local matches at different scales and levels of complexity, thereby forming global feature correspondences. In this model, the multilayer perceptron (MLP) is used to accurately find the ideal correspondences and successfully complete the alignment procedure. **Result** This work provides empirical evidence supporting the improved efficacy of the proposed algorithm as demonstrated by its consistent performance across multiple datasets. Notably, this algorithm achieved impressive registration recall rates of 93.2% and 67.3% on the 3DMatch and 3DLoMatch datasets, respectively. In the experimental evaluation conducted on the ModelNet-40 and ModelLoNet-40 datasets, this algorithm achieved minimal rotational errors of 1.417 degrees and 3.141 degrees, respectively, and recorded translational errors of 0.013 91 and 0.072. These outcomes highlight the effectiveness of this algorithm in point cloud registration and demonstrate its capability to accurately align point clouds with low rotational and translational discrepancies. These results also point to a significant enhancement in the accuracy of the proposed algorithm compared with the REGTR approach. Specifically, in contrast to REGTR, the proposed algorithm achieved significantly reduced inference times of 27.205 ms and 30.991 ms on the 3DMatch and ModelNet-40 datasets, respectively. The findings of this study emphasize the performance of the proposed algorithm in effectively addressing the challenging issue of disregarding features in point cloud registration tasks with minimal overlap. **Conclusion** This article presents a novel point cloud matching technique that combines edge improvement with hierarchical attention. This technique integrates polynomial kernel functions into the EAKPConv framework to improve the identification of edge features in point clouds and uses HAFM to extract specific local information. The module improves feature matching by using the similarities in edge features. This approach successfully achieves a harmonious combination of local and global feature matching, hence enhancing the comprehension of point cloud data. Implementing a hierarchical analysis technique greatly increases the registration accuracy by accurately matching local-global information. Furthermore, increasing the cross-entropy loss function enhances the accuracy of local matching and reduces misalignments. This study assesses the performance of the proposed algorithm on the ModelNet-40, ModelLoNet-40, 3DMatch, and 3DLoMatch datasets, and results indicate that this algorithm substantially enhances registration accuracy, particularly in difficult situations with limited data overlap. This algorithm also exhibits superior registration efficiency compared with standard approaches.

Key words: 3D point cloud registration; low-overlap point cloud; edge features; hierarchical attention; local similar matching

0 引言

三维点云配准任务是将采集时间、方向不同的两帧点云进行变换,得到两者最优匹配的过程。它是计算机视觉领域的一项基本任务,可以应用于点云分类、分割、三维重建等其他高级视觉任务,并在自动驾驶、虚拟现实和机器人控制等领域有巨大应用前景。随着深度学习在点云分类、语义分割以及

目标检测等研究中获得成功,涌现出一些基于深度学习的配准方法。基于学习的配准方法中,如3DRegNet(Pais等,2020)、ReAgent(Bauer等,2021)等已在高重叠点云上取得了准确的匹配结果,但在有限重叠点云配准上表现不佳。

本文发现由于源点云、目标点云的重叠区域更小,更难找到正确的特征匹配,因此导致低重叠度点云配准会出现更高的错配率。目前的配准方法中,如PRNet(partial registration network)(Wang和Solo-

mon, 2019a)、HOUV (hybrid optimization with unconstrained variables) (Yan 等, 2023) 等大都通过提取源点云、目标点云的重叠区域, 匹配重叠区域或将重叠区域放大再做全局匹配。但随着重叠区域减小, 这种通过重叠区域引导点云配准方法的匹配成功率显著降低。

针对上述问题, 本文提出一种基于增强边缘特征相似度指标优化的特征匹配方法, 解决传统方法在低重叠场景中通过重叠区域引导配准的局限性。本文方法采用从局部到全局的层次化解码器重构点云, 确保特征的完整性及配准的准确性。此外, 本文设计的层次注意力融合模块 (hierarchical attention fusion module, HAFM) 减少两次注意力计算的时间复杂度, 在配准效率上相比已有方法有所提升, 一定程度上解决了点云配准的效率问题。

本文主要创新点及贡献有: 1) 设计了一个多项式核函数计算权重矩阵, 并应用到核点卷积 (kernel point convolution, KPConv) (Thomas 等, 2019) 中, 以增强后续操作对边缘的感知能力。2) 构建一种局部到全局的 HAFM, 该模块首先通过局部注意力模块动态捕获源目标点云的复杂局部特征, 通过相似度优化完成局部相似区域匹配, 并得到全局尺度特征。3) 为改善局部匹配效果, 改进了交叉熵损失函数。通过估计局部对应与真实对应的关系更新参数, 以避免错误的局部对齐。

1 相关研究工作

三维点云配准方法主要分为基于对应点匹配的配准方法和基于全局特征的配准方法 (武越 等, 2023; 秦红星 等, 2022)。此外, 根据是否使用深度卷积神经网络 (deep convolutional neural network, DCNN), 又可划分为传统方法和基于深度学习的方法。

1.1 基于对应点匹配的配准方法

传统方法方面, 迭代最近点算法 (iterative closest point, ICP) (Besl 和 McKay, 1992) 作为最著名、应用最广泛的配准方法, 首先利用初始点对或上一次迭代点对, 将点云变换得到一个临时点云簇, 然后比较目标点云和临时点云簇, 找出目标点云中每个点在临时点云簇中的最邻近点, 并通过奇异值分解 (singular value decomposition, SVD) 求解最优变换或通

过对应点间的距离判断配准误差进行迭代拟合 (杨军 等, 2018)。然而, ICP 算法是一种局部配准算法, 很容易陷入局部极小值。为提高配准的准确性, Rusu 等人 (2009) 设计了快速点特征直方图描述子 (fast point feature histograms, FPFH), 从点云的局部邻域中选择任意点对计算它们的特征描述, 然后比较点对特征描述与预设区间得到的对应特征直方图。相比于点特征直方图 (point feature histograms, PFH) (Rusu 等, 2008) 算法, FPFH 的时间复杂度更低, 但 FPFH 未计算邻近点所有组合, 可能因此丢失局部几何特征。

在深度学习方面, Wang 和 Solomon (2019b) 提出的深度最近点算法 (deep closest points, DCP) 利用共享的动态图卷积神经网络 (dynamic graph CNN, DGCNN) 提取特征, 并且通过 Transformer 交叉学习点云模型间的条件残差以获得两个点云模型的对应关系 (Vaswani 等, 2017)。该方法通过自注意力软指针 (soft pointer) 预测对应关系, 与当时传统方法相比有较好的配准准确率和速度, 但其对重叠点云的配准效果较差。Bai 等人 (2020) 将 KPConv 作为主干网络提出了局部特征描述子算法 (description of 3D local features, D3Feat), 提升了特征点检测和局部几何特征的检测能力, 对室内点云的匹配效果较好。SIRA-PCR (sim-to-real adaptation for 3D point cloud registration) (Chen 等, 2023) 是首次尝试将模拟数据转换到真实场景下应用, 制作了 3D-FRONT 数据集, 提出了自适应重采样模拟真实数据的特性, 通过理解和模仿真实世界数据的内在结构和分布, 保证模型面对真实世界数据时的有效性和可靠性, 但匹配准确度略逊色于其他特征提取匹配方法。

上述方法依赖于匹配点的精确度, 不利于三维点云的精细化匹配。此外, 借助 Transformer 的方法对算力需求较高, 随着输入点云数据量的增加, 算法的推理时间将显著增加。

1.2 基于全局特征的配准方法

在 ICP 算法提出后, Yang 等人 (2016) 在此基础上提出了一种基于全局最优的迭代最近点算法 (globally optimal iterative closest point, Go-ICP)。该方法将局部 ICP 算法集成到分支界定算法 (branch-and-bound, BnB) (Ross 和 Soland, 1975) 中, 解决了原有 ICP 算法求解过程中容易陷入局部最优的问题,

保证算法能获得全局最优解,但是此方法极大地降低了配准速度。在改进算法的效率方面,Zhou等人(2016)提出的快速全局配准算法(fast global registration, FGR)通过构造目标函数将非凸问题转化为凸优化问题,并使用优化迭代技巧提升效率。然而,该算法的目标是找到误差最小的解,而最小误差点不一定是最优解。

Aoki等人(2019)利用三维分类和分割任务中的经典主干网络PointNet(Qi等,2017),提出了基于全局特征配准方法的PointNetLK(PointNet Lucas & Kanade),开创性地将PointNet和LK(Lucas & Kanade)算法相结合,设计了一种可学习的循环深度神经网络进行点云配准。该算法针对不同场景数据都具有较好的鲁棒性。Sarode等人(2019)提出了一种无监督点云配准网络(point cloud registration network, PCRNet),将三维旋转表示为四元数,对有噪声的重叠点云配准效果优于PointNetLK,并且训练和配准速度更快。然而,该网络对输入点云进行标准化处理,在点云密度不均匀或存在噪声扰动情况下性能会有所下降。Huang等人(2020)提出了具备解码器的特征度量配准算法(feature-metric registration, FMR),将配准问题从此前的最小化点对之间的投影误差转化为最小化特征差异,并通过解码器还原点云达到去除冗余的作用,保证特征的完整性。然而,该方法使用RANSAC(random sample consensus)算法来估计变换参数,可能会受到点云采样质量的影响导致配准结果不稳定。Yang等人(2024)通过点对间的兼容性约束构建图,并采用互投票机制有效地增强了对应关系之间的相互验证,提高了3D点云匹配的准确性和鲁棒性。

该类方法解决了局部特征描述符表达全局信息时代价过高的问题,在重叠场景的点云数据集中有较好的效果,但在低重叠场景的点云数据中未考虑非重叠特征的提取问题,对配准精度有一定的影响。

2 边缘特征增强与层次注意力融合的配准方法

2.1 问题定义及方法概述

设源点云为 $P = \{p_i \in \mathbb{R}^3 | i = 1, \dots, N\}$,目标点云为 $Q = \{q_i \in \mathbb{R}^3 | i = 1, \dots, M\}$ 。其中, N, M 分别为源点云、目标点云的数量。点云配准的目标是求解从 P 到 Q 的三维刚性变换矩阵 $T = \{R, t\}$ 。其中,包含三维旋转矩阵 $R \in SO(3)$ 、三维平移矩阵 $t \in \mathbb{R}^3$, $SO(3)$ 表示旋转矩阵集合。该变换可通过式(1)进行求解,具体为

$$\arg \min_{R, t} \sum_{(p_i, q_i) \in C^*} \|Rp_i + t - q_i\|_2^2 \quad (1)$$

式中, C^* 是从 P 到 Q 的真实测地距离。

本文算法的网络结构如图1所示。首先利用边缘自适应核点卷积(edge adaptive KPConv, EAKPConv)同时提取源点云 P 和目标点云 Q 的特征,得到稀疏化特征表示 P_s 和 Q_s 。然后,利用已处理的稀疏化特征,动态捕获与目标节点在空间中的位置关系,通过模型在不同尺度上累积独立相似性度量的局部特征,并最终将这些局部特征逐层叠加整合到全局特征 f_{GP}, f_{GQ} 的对应关系中。最后,由多层感知机(multi-layer perceptron, MLP)还原点云三维位置编码 $\hat{P}$ 和 $\hat{Q}$,并计算输出刚性变换矩阵。

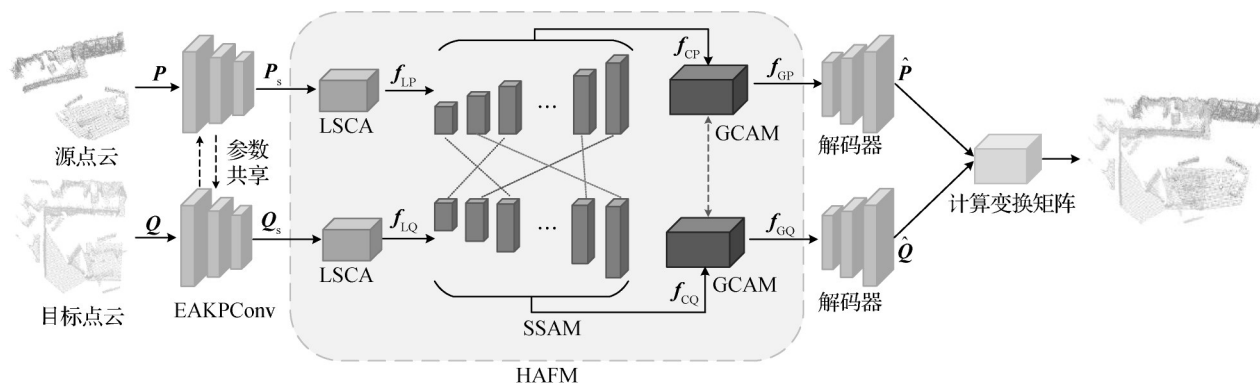


图1 算法网络框架

Fig. 1 Network structure of the proposed method

2.2 EAKPConv 模块

由于KPConv算法此前主要用于提取局部特征,对边缘特征不敏感。受杨军和张琛(2023)的启发,在KPConv的基础上设计了EAKPConv模块,并在EAKPConv中定义了一种多项式核函数替代高斯核函数,实现边缘特征增强。具体操作如下:将输入点云的 N 个点定义为 $X = \{x_i | i \in (1, N)\}$, x_i 表示第 i 个点。以 x_i 为起始点,利用 K 最近邻算法(K-nearest neighbor, KNN)采样得到 x_i 的局部特征图。然后利用核函数计算新的权重矩阵,通过该矩阵对邻域特征进行变换。本文在拉普拉斯核函数的基础上设计了新的多项式核函数,具体为

$$K(x_i, x_j) = \lambda_1 \left(-\frac{\|x_j - x_i\|}{\sigma} \right) + \lambda_2 ((x_i^T x_j) + I) \quad (2)$$

式中, x_i, x_j 分别为关键点及对应邻域点, $\lambda_1(\cdot)$ 项为拉普拉斯核, $\lambda_2(\cdot)$ 项为度量 x_i, x_j 的距离。 $\lambda_1 + \lambda_2 = 1$, λ_1, λ_2 为小于1的正数, σ 表示关键点与局部邻域距离范围, I 是单位矩阵,起到了偏置作用。此外,通过在EAKPConv之间建立残差连接,然后结合ResNet块继续下采样。

2.3 HAFM 模块

对于点云配准任务,利用注意力模块在学习局部上下文特征时,只能学习每个点与其邻域的上下

文特征,容易受到噪声及不同采样密度的影响,并且现有利用注意力的点云配准算法时间复杂度较高。本文局部邻域层次编码策略受到了深层点云注意力网络(point transformer-in-transformer, Point-TNT)、基于注意力的图卷积网络(attention-based graph convolutional network, AGCN)的启发(Berg等,2022;Xie等,2020),并在此基础上设计了HAFM模块。利用注意力机制的聚合能力使模型更加关注与目标节点的联系,并在空间上更接近聚类质心,达到动态捕获复杂点云特征的目的。又利用层次化架构的特征,让每一层局部匹配模块都具备独立的相似性度量来捕获点云之间的局部特征相似性,这些局部特征会被更新并传递到下一层的注意力模块中,形成一个层次结构。这种层次结构使得模型能够在不同的尺度、复杂性级别上累积、综合这些局部匹配的贡献,并最终传导到全局特征对应关系中,该模块结构如图2所示。首先,利用KNN算法构建局部特征图;然后,使用局部空间差异注意模块(local spatial contrast attention module, LSCAM)学习点云局部之间的关系;再次,使用局部匹配模块计算两个点云相似分数,明确两个点云的局部位置编码。利用上层局部注意力模块的输出,逐层将局部特征输入给下层注意力模块;最后,串联各层之间的局部特征,输出带有位置编码及相似匹配加权的全局点云特征。

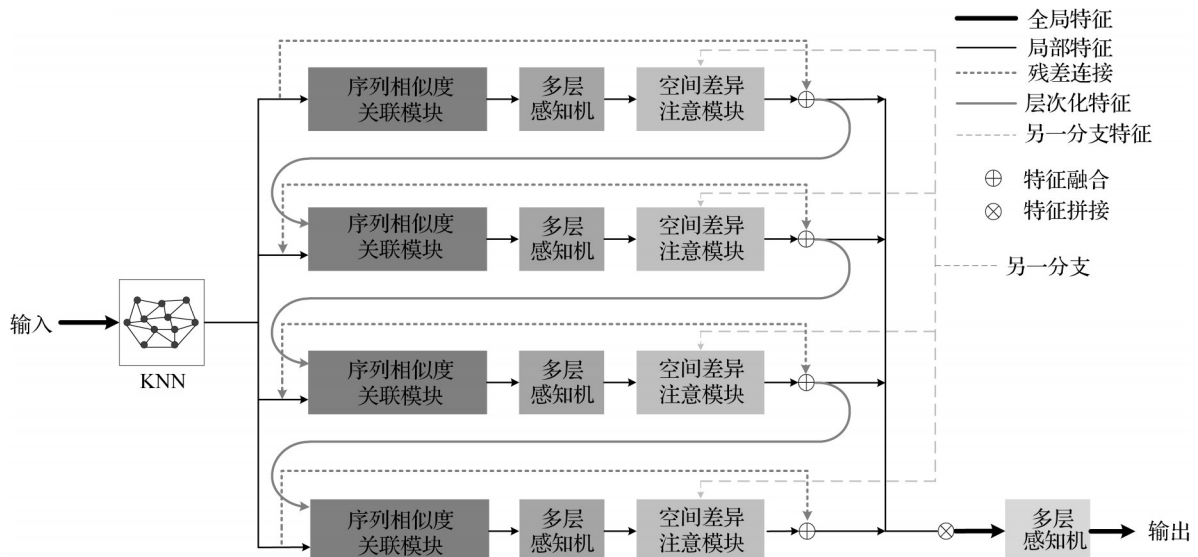


图2 HAFM模块结构图

Fig. 2 HAFM structure diagram

2.3.1 LSCAM 模块

为了精确提取局部几何信息,本文提出了局部

空间差异注意模块(LSCAM),该模块用于强化局部特征之间的空间对比,从而提升了模型对于几何结

构的识别能力。例如,更新节点 i 、与 i 相邻的节点 j 之间的关系,得到节点 i 和节点 j 的聚合特征 f'_i ,具体为

$$f'_i = \alpha_{ij} f_j + f_i \quad (3)$$

式中, f_i 为节点 i 的特征, f_j 为节点 i 中最近邻节点 j 的特征。 α_{ij} 为节点 i 、 j 之间的注意力权重,用于控制节点 j 对节点 i 的影响,其计算为

$$\alpha_{ij} = \frac{f_{tp}}{\sum_{j \in n} f_{ij} f_i} \quad (4)$$

式中, f_{tp} 为节点 i 、近邻节点 j 的相对位置特征, n 为节点 i 的邻域节点集合, f_i 为节点 i 的特征, f_j 是节点 i 中最近邻节点 j 的特征, $\sum_{j \in n} f_{ij} f_i$ 是一个归一化因子。

每个维度的相对位置计算为

$$f_{tp}^{(d)} = f_{ij}^{(d)} - f_i^{(d)} \quad (5)$$

式中, d 为特征值维度索引, $f_{tp}^{(d)}$ 是节点 i 、节点 j 在 d 维特征空间上的差异, $f_{ij}^{(d)}$ 是节点 i 中最近邻节点 j 的 d 维邻域特征, $f_i^{(d)}$ 是节点 i 的 d 维特征。

在局部自注意力的聚合过程中,将动态注意力权重 α_{ij} 引入到特征差异位置计算,可以在空间中更接近相关节点,能提高处理两个点云的局部匹配的性能。在计算局部匹配的过程中,可以提供具有邻域上下文信息的局部特征。特别是对于点密度不均匀以及稀疏的数据,通过特征差异计算能更容易地区分最近邻的局部特征。

2.3.2 SSAM模块

为了增强聚合后源点云、目标点云的全局匹配能力,本文算法通过序列相似度关联模块(sequential similarity association module, SSAM)进行局部匹配加权。首先计算源点云、目标点云之间的相似度矩阵,利用该矩阵确定局部特征中相似性最高的区域,并据此进行最佳匹配预测;然后,使用该相似性分数更新聚合后的局部特征。相似性分数计算为

$$s_{ij} = \frac{f'_i \cdot f'_j{}^T}{\|f'_i\|_2 \cdot \|f'_j\|_2} \quad (6)$$

式中, s_{ij} 表示源点云局部邻域 i 的聚合特征 f'_i 、目标点云局部邻域 j 的聚合特征 f'_j 的相似度矩阵。 $\|f'_i\|_2$ 、 $\|f'_j\|_2$ 分别为 f'_i 、 f'_j 两个局部特征的L2范数,用于对点积进行归一化。同时,利用邻域间局部相似分数构建了相似矩阵 S_{final} ,具体为

$$S_{\text{final}} = \text{Combine}(s^{(1)}, s^{(2)}, \dots, s^{(k)}) \quad (7)$$

式中,每个元素 $s^{(\cdot)}$ 表示局部相似性分数矩阵,Combine表示平均加权。

根据两点云局部邻域特征间相似性矩阵 s_i 为局部邻域特征施加影响。相似矩阵 s_i 用于逐元素地调整权重矩阵 W_i ,计算为

$$f_s^i = f_i + \text{ReLU}(f_i(W_i \odot s_i) + b_i) \quad (8)$$

式中, f_i 表示KNN算法采样的局部邻域特征, $\odot$ 表示逐元素位置点积, $f_i(W_i \odot s_i)$ 为局部邻域特征匹配更新。并且在每层局部自注意力层后将局部匹配特征串联为全局注意力特征 $f_c = [f_s^1, \dots, f_s^k]$ 。在此基础上,将源点云、目标点云的全局特征 f_{cp} 、 f_{cq} 利用与CCNet(criss-cross network)(Huang等, 2023)、REGTR(registration transformer)(Yew和Lee, 2022)类似的交叉注意力机制得到全局条件特征 f_{gp} 、 f_{gq} ,如图3所示。此外,本文的解码器使用3层MLP回归点云三维坐标并协方差矩阵计算变换矩阵。

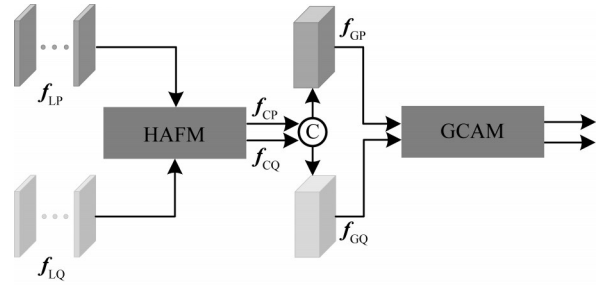


图3 HAFM与GCAM模块连接示意图

Fig. 3 Schematic diagram of HAFM and GCAM modules

2.4 损失函数

为预测局部邻域相似对应并减少因为匹配二义性导致的错误加权,本文借鉴了UTOPIC(uncertainty-aware overlap prediction network)(Chen等, 2022)计算软对应损失,通过预测对应 C 与真实对应 C^* 的局部对应损失 L_c 来更新训练,具体为

$$L_c = - \sum_{i=1}^N \sum_{j=1}^M C_{ij} \log C_{ij}^* + (1 - C_{ij}) \log (1 - C_{ij}^*) \quad (9)$$

$$C_{ij} = \begin{cases} 1 & c_{ij} = c_{ij}^* \\ 0 & \text{其他} \end{cases} \quad (10)$$

式中, i, j 分别为源点云、目标点云节点索引, $c_{ij} \in C$ 表示预测局部对应, $c_{ij}^* \in C$ 表示真实局部对应。在模型训练过程中,组成与分数矩阵同维的真实对应独热矩阵,通过独热矩阵、分数矩阵相对位置点积以矫

正分数矩阵,通过反向传播来更新权重,如图4所示。

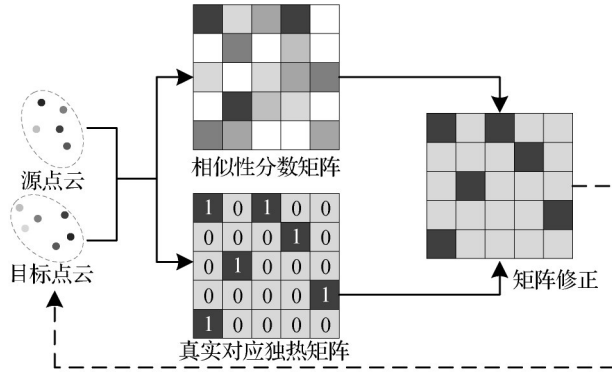


图4 矫正分数矩阵示意图

Fig. 4 Schematic diagram of correction score matrix

此外,本文的损失函数还包括平移损失 L_t 、旋转损失 L_r 以及特征损失 L_f ,利用真实平移、旋转估计预测值,具体为

$$L_t = \|t - t_{\text{real}}\|_2, L_r = \|RR_{\text{real}}^T - I\|_2 \quad (11)$$

式中, R_{real} , t_{real} 为真实旋转与形变, I 为同维度的单位矩阵。

为了引导两点云建立对应关系时考虑几何属性,本文最小化式(2)来最大化正负样本的互信息实现,具体为

$$L_f = -E_{x \in X} \log \frac{f(x, p_x)}{f(x, p_x) + n_x f(x, n_x)} \quad (12)$$

式中, E 表示数学期望, f_x 表示点云解码后的特征, p_x , n_x 分别表示点云解码后目标点云中与源点云匹配、不匹配的关键点。 $f(x, c) = \exp(f_x^T W_f c)$ 是一个双线性插值模型, W_f 是一个可学习的线性变换。总损失函数 L 由以上损失函数线性叠加组成,即 $L = L_t + L_r + \lambda_f L_f + L_c$,其中 λ_f 借鉴了REGTR的参数设置。

3 实验结果与讨论

本文实验的硬件设备基于一台 GeForce RTX 3090的图形工作站,软件环境基于Ubuntu 20.04,使用Pytorch 1.9.1作为深度学习框架,Minkowski Engine稀疏张量自动微分库为训练加速。使用AdamW作为网络的优化器,帮助模型更快逼近最小值。为了验证层次化点云配准算法的有效性,在

3DMatch(Zeng等,2017)、ModelNet-40(Wu等,2015)数据集上训练测试,并在它们的低重叠数据集上测试。

3.1 数据集及评价指标

3DMatch是包含62个室内场景的点云数据集,实验选择其中的54个场景进行训练,在另外8个场景中进行测试。标准的3DMatch忽略了重叠度小于30%的点云匹配,3DLoMatch(Huang等,2021)对3DMatch中的8个场景重新构建,只考虑重叠度在10%~30%的点云匹配的低重叠数据集。使用体素网格对原始点云进行预处理,对两个点云施加刚性扰动及位置抖动。计算两个点云的重叠区域,将两个点云数据、两个点云的重叠掩码、两个点云的对对应矩阵以及随机的变换矩阵组成新的数据用于训练。

ModelNet-40是一个CAD三维模型基准数据集,包含从40个不同类别的12311个模型中采样的合成点云数据。实验使用其中的6314个样本进行训练,另外的1266个样本进行测试。其中,ModelNet-40的成对平均重叠率为73.5%,ModelLoNet-40是通过ModelNet-40生成的低重叠度数据,平均重叠率为53.6%。在预处理中对该数据集的两个点云均匀下采样2048个点并缩放到单位球内,在单位球中找到一个平面,从中下采样717个点并丢弃其余部分。最后计算点云间的重叠部分,组成与3DMatch相同的数据结构用于训练。

针对3DMatch数据集使用相对旋转误差(relative rotation error, RRE)、相对平移误差(relative translation error, RTE)以及配准召回率(registration recall, RR)作为算法的评估指标。相对旋转误差量化了两个点云坐标系之间的角度差,通过衡量两个点云旋转矩阵之间测地距离的差值计算旋转误差。相对平移误差通过测量两点云的欧几里得距离评价点云配准的准确性。具体为

$$f_{\text{RRE}} = \arccos\left(\frac{\text{tr}(R^T R^*) - 1}{2}\right) \quad (13)$$

$$f_{\text{RTE}} = \|t - t^*\|_2 \quad (14)$$

式中, R 为预测旋转矩阵, R^* 为真实旋转矩阵, t 为预测平移矩阵, t^* 为真实平移矩阵。

配准召回率量化了真实对应总数中正确对应的百分比,可以衡量实际任务中点云配准端到端的性

能。本文使用均方根误差 (root mean square error, RMSE) 作为配准召回率的计算方法, $RMSE < 0.2$ 作为所有点对分数的阈值, 均方根误差的计算为

$$f_{RMSE} = \sqrt{\frac{1}{|C^*|} \sum_{(p,q) \in C^*} \|T(p) - q\|_2^2} \quad (15)$$

式中, C^* 为真实点云对应集合, $T(p)$ 为算法预测目标点云。

对于 ModelNet-40 数据集采用相对旋转误差、相对平移误差以及倒角距离 (chamfer distance, CD) 作为算法的评价指标。倒角距离通过从源点云中迭代每个点与其在目标点云中的最近点之间的最小距离来量化源点云与目标点集合之间的相似性, 具体为

$$CD(P, Q) = \frac{1}{|P|} \sum_{p \in P} \min_{q \in Q_{raw}} \|T(p) - q\|_2^2 + \frac{1}{|Q|} \sum_{q \in Q} \min_{p \in P_{raw}} \|q - T(p)\|_2^2 \quad (16)$$

式中, P_{raw} , Q_{raw} 分别为未处理的源点云、目标点云数据, P , Q 为预处理后的点云数据。

3.2 3DMatch 和 3DLoMatch 实验结果分析

本文对 3DMatch 数据集进行了 70 轮次的训练, 批次大小为 2, 初始学习率为 0.000 1, 权重正则化系数为 0.000 1, 每训练 10 个轮次学习率做一次衰减。为了验证本文算法在真实点云场景的有效性, 与其他主流点云配准算法的实验结果进行了对比, 并展示了采样点为 5 000 时的配准实验结果, 如表 1 所示。

表 1 3DMatch 数据集的实验结果对比
Table 1 Comparison of experimental results on the 3DMatch dataset

指标	方法	场景								平均
		厨房	家 1	家 2	酒店 1	酒店 2	酒店 3	书房	实验室	
配准召回率/% ↑	FCGF	98.0	94.3	68.6	96.7	91.0	84.6	76.1	71.1	85.1
	D3Feat	96.0	86.8	67.3	90.7	88.5	80.8	78.2	64.4	81.6
	Predator-5k	97.6	97.2	74.8	98.9	96.2	88.5	85.9	73.3	89.0
	PCAM	96.9	93.4	78.6	96.7	83.3	84.6	81.6	68.9	85.5
	REGTR	97.8	90.6	75.5	97.8	94.9	100.0	88.5	91.1	92.0
	SIRA-PCR	98.8	91.3	74.9	96.3	96.1	98.1	92.1	92.2	92.5
	MV+FCGF	98.3	93.6	78.3	95.6	92.7	87.3	85.9	76.9	91.8
	本文	98.4	92.3	77.4	98.1	94.6	100.0	91.5	93.8	93.2
相对旋转误差/(°) ↓	FCGF	1.767	1.849	2.210	1.867	1.667	2.417	2.024	1.792	1.949
	D3Feat	2.016	2.029	2.425	1.990	1.967	2.400	2.346	2.115	2.161
	Predator-5k	1.861	1.806	2.473	2.045	1.600	2.458	2.067	1.926	2.029
	PCAM	1.965	1.644	2.145	1.874	1.434	1.631	2.250	1.521	1.808
	REGTR	1.729	1.347	1.797	1.639	1.289	1.810	1.570	1.357	1.567
	SIRA-PCR	1.753	1.425	1.841	1.442	1.365	1.884	1.586	1.580	1.610
	MV+FCGF	3.572	2.912	3.260	3.729	3.211	4.18	3.643	3.376	3.805
	本文	1.756	1.125	1.744	1.689	1.318	1.756	1.425	1.410	1.528
相对平移误差/m ↓	FCGF	0.053	0.056	0.071	0.062	0.061	0.055	0.082	0.090	0.066
	D3Feat	0.055	0.065	0.080	0.064	0.078	0.049	0.083	0.064	0.067
	Predator-5k	0.048	0.055	0.070	0.073	0.060	0.065	0.080	0.063	0.064
	PCAM	0.050	0.051	0.060	0.066	0.051	0.058	0.081	0.052	0.059
	REGTR	0.040	0.041	0.057	0.057	0.042	0.039	0.054	0.058	0.049
	SIRA-PCR	0.067	0.046	0.058	0.051	0.047	0.052	0.077	0.052	0.053
	MV+FCGF	0.107	0.093	0.099	0.115	0.103	0.085	0.093	0.095	0.097
	本文	0.040	0.043	0.052	0.055	0.038	0.037	0.046	0.052	0.045

注: 加粗字体表示各指标的各列最优结果, ↑ 表示数值越大性能越好, ↓ 表示数值越小性能越好, MV 为 Mutual Voting。

从表1可以看出,本文算法在3DMatch数据集上获得了93.2%的配准召回率(在酒店3、实验室两个场景中获得最佳配准结果),并在平均相对旋转误差、平均相对平移误差两项指标的测试中,本文算法均为最佳。本文算法与Predator、REGTR算法都采用双流编解码结构,并且都使用了KPconv主干提取点云、Transformer寻找点云对应关系。但与本文算法不同的是,Predator侧重于超点的聚合作用而忽略了局部非超点特征,其通过DGCNN最大化超点特征并多次更新超点,用超点特征代替原有局部特征,导致其丢失了局部拓扑结构及位置信息,因此在其大部分场景中的配准召回率以及所有平移误差、旋转误差的实验结果都低于本文算法。REGTR通过Transformer可以捕获不同位置的依赖关系,并构建全局上下文关联,但忽略了不同层次局部特征的交互作用。而本文算法在此基础上利用不同层次的局部特征构建全局上下文,使用底层Transformer模块学习局部特征,高层模块学习不同层次之间的关联,取得了较好的配准结果。在SIRA-PCR算法的评估中,首先在FlyingShapes数据集上训练SIRA-PCR,然后将训练好的模型迁移到3DMatch上进行微调。此外,为了更好地比较SIRA-PCR与本文算法,在微调阶段采样点选取2 000进行训练。SIRA-PCR采用与本文算法类似的从局部到全局的模型构建思路,采用自适应重采样缩小预训练数据集与真实数据集之间的差距,通过多尺度鉴别器完成对全局特征的调整,这使得SIRA-PCR在不同场景中都有较强的泛化能力。使用与本文算法相近采样点数目,其配准召回率与本文算法相近,并在其中两个场景获得最优结果。本文通过在Swig工具编译Mutual Voting算法,通过编写Swig接口方式接入本文实验环境,利用FCGF(fully convolutional geometric features)(Choy等,2019)作为特征提取器,并在3DMatch数据集进行测试。由于Mutual Voting与FCGF的组合可以有效筛选出正确的匹配点对,但在准确估计匹配对之间精确的空间变换参数存在一定局限性,导致算法的配准召回率结果非常接近最新的端到端学习方法,但旋转误差与其他方法相差较大。

本文将算法在3DMatch数据集上训练的模型直接迁移到低重叠数据集3DLoMatch上进行测试,实验结果如表2所示。可以看出,本文算法的相对旋转误差获得测试的最佳,其余指标接近SOAT(state

of the art)模型。Predator是除本文算法外唯一考虑边缘特征对配准影响的算法,但本文算法使用的多项式核函数相比Predator使用的高斯核函数峰值密度更高,边缘响应更明显,因此在更多的场景中表现出比Predator更好的实验结果。此外,稀疏全卷积结构的FCGF通过使用4种不同的损失函数在两个场景获得了超过大部分算法的实验结果(厨房场景获得了相对旋转误差的最佳,而家2场景的相对平移误差仅次于本文算法和SIRA-PCR的实验结果)。PCAM(product of cross-attention matrices)(Cao等,2021)采用与本文算法相似的不同尺度特征结合的匹配方法,但其获取不同尺度交叉注意力上下文信息的点云对齐方法对初始姿态要求更高并且对噪声更敏感,导致其平均配准召回率结果低于本文算法。

此外,为了更客观地评估本文算法,本实验与最新算法SIRA-PCR和Mutual Voting进行比较。其中,SIRA-PCR算法在3DLoMatch数据上表现出较好的配准结果,使用多尺度鉴别器使算法在重叠度较低的数据中也能获得所有算法的最佳配准召回率,与本文算法相当。而SIRA-PCR在相对旋转误差逊色于本文算法,相对平移误差与本文算法相当。由于Mutual Voting方法侧重于全局对应关系的评分和排序,在旋转误差上与其他方法差距较大,但配准召回率接近先进算法,说明它能够有效地识别和排除错误的对应关系。该算法的互相投票机制强化了对正确对应关系的信任度,在可用的共享重叠区域较小时也可以做出合理的对应判断,它在处理特定问题时相比传统深度学习方法更高效和灵活,不需要从庞大的数据中学习。

图5为本文算法利用Open 3D工具库在3DMatch、3DLoMatch的配准结果可视化。其中,图7(a)(b)为3DMatch的可视化结果,图7(c)~(e)为3DLoMatch的可视化结果。本文算法可以准确地匹配点云的相应区域,在图7(a)(e)场景中的墙壁可以完全对齐,图7(c)(d)场景中的桌椅等小部件物品的边缘基本重合,而且对具有明显起伏、不规则边缘的场景图7(e)也能基本对齐。

3.3 ModelNet-40、ModelLoNet-40实验结果分析

为了验证本文算法在三维模型数据集上的实验效果,在ModelNet-40数据集上对算法进行了训练、测试。实验进行了350轮次,批次大小为4,初始学习率为0.000 1,每训练50轮次学习率衰减一次,权

表2 3DLoMatch数据集的实验结果对比
Table 2 Comparison of experimental results on the 3DLoMatch dataset

指标	方法	场景								平均
		厨房	家 1	家 2	酒店 1	酒店 2	酒店 3	书房	实验室	
配准召回率/% ↑	FCGF	60.8	42.2	53.6	53.1	38.0	26.8	16.1	30.4	40.1
	D3Feat	49.7	37.2	47.3	47.8	36.5	31.7	15.7	31.9	37.2
	Predator-5k	71.5	58.2	60.8	77.5	64.2	61.0	45.8	39.1	59.8
	PCAM	71.1	56.7	60.8	70.8	59.9	41.5	36.4	42.0	54.9
	REGTR	66.2	58.5	64.9	72.7	61.3	70.7	53.0	71.0	64.8
	SIRA-PCR	71.9	60.8	65.3	78.3	65.7	74.1	53.8	70.6	67.5
	MV+FCGF	60.4	57.7	53.3	52.9	54.0	52.8	49.3	47.1	60.2
	本文	70.8	61.5	67.5	75.6	63.8	74.4	52.3	72.4	67.3
相对旋转误差/(°) ↓	FCGF	2.904	3.229	3.277	2.768	2.801	2.822	3.372	4.006	3.147
	D3Feat	3.226	3.492	3.373	3.330	3.165	2.972	3.708	3.619	3.361
	Predator-5k	3.079	2.637	3.220	2.694	2.907	3.390	3.046	3.412	3.048
	PCAM	3.501	3.518	3.571	3.649	3.197	3.278	4.148	3.368	3.529
	REGTR	3.366	2.446	3.244	2.732	2.439	2.919	3.044	2.428	2.827
	SIRA-PCR	3.326	2.717	2.932	2.861	2.520	3.026	3.757	2.470	2.948
	MV+FCGF	4.102	4.534	4.710	5.170	4.807	4.592	4.474	4.316	5.153
	本文	3.187	2.182	3.069	2.957	2.727	2.724	3.100	2.486	2.804
相对平移误差/m ↓	FCGF	0.084	0.097	0.076	0.101	0.084	0.077	0.144	0.140	0.100
	D3Feat	0.088	0.101	0.086	0.099	0.092	0.075	0.146	0.135	0.103
	Predator-5k	0.081	0.080	0.084	0.099	0.096	0.077	0.101	0.130	0.093
	PCAM	0.088	0.108	0.090	0.112	0.098	0.068	0.117	0.110	0.099
	REGTR	0.079	0.064	0.078	0.094	0.074	0.060	0.093	0.077	0.077
	SIRA-PCR	0.069	0.073	0.071	0.097	0.079	0.061	0.106	0.073	0.079
	MV+FCGF	0.093	0.098	0.108	0.132	0.104	0.092	0.099	0.091	0.114
	本文	0.078	0.071	0.067	0.095	0.076	0.058	0.101	0.089	0.079

注:加粗字体表示各指标的各列最优结果,↑表示数值越大性能越好,↓表示数值越小性能越好,MV为Mutual Voting。

重衰变为0.000 5,实验结果如表3所示。

由于ModelNet-40相比3DMatch数据集更稀疏且点分布不均匀,因此在ModelNet-40数据集上进行测试更具挑战性。从表3可以看出,本文算法在ModelNet-40的倒角误差逊色于REGTR,在ModelLoNet的平移误差略逊色于SIRA-PCR,其他评价指标均高于其他算法。Predator使用RANSAC获得最终的对应点集,相比REGTR及本文算法的端到端配准模型,其收敛到全局最优匹配结果的能力较差。此外,将ModelNet-40数据集的训练模型迁移到ModelLoNet-40上进行测试,本文算法在3种评价指标中均为最佳,表明本文算法在稀疏的低重叠度点云数据上也具备较好的匹配结果。使用交叉注意力方法的Predator、REGTR的相对旋转误差优于使用特征

融合方法的OMNet(Xu等,2021)、RPM-Net(Yew和Lee,2020)。OMNet利用重叠遮罩寻找点云重叠特征的方法以及RPM-Net利用退火算法计算匹配矩阵的方法比DCP算法具有更小的倒角误差。本文利用层次化注意力方法在不同层次做局部特征选择相比REGTR获得了更好的测试结果,同时实验也验证了本文局部特征聚合的方法可以很好地适应稀疏点云数据。为了在ModelNet-40、ModelLoNet-40数据集上评估SIRA-PCR,使用预训练模型在ModelNet-40上微调SIRA-PCR,其中采样点数目设置为1 000。由于ModelNet-40数据的点密度较稀疏,本文算法除了在ModelLoNet-40的相对平移误差外,其他指标均优于SIRA-PCR,说明本文算法在处理低密度点云数据配准方面具有优越性。Mutual Voting(MV)在

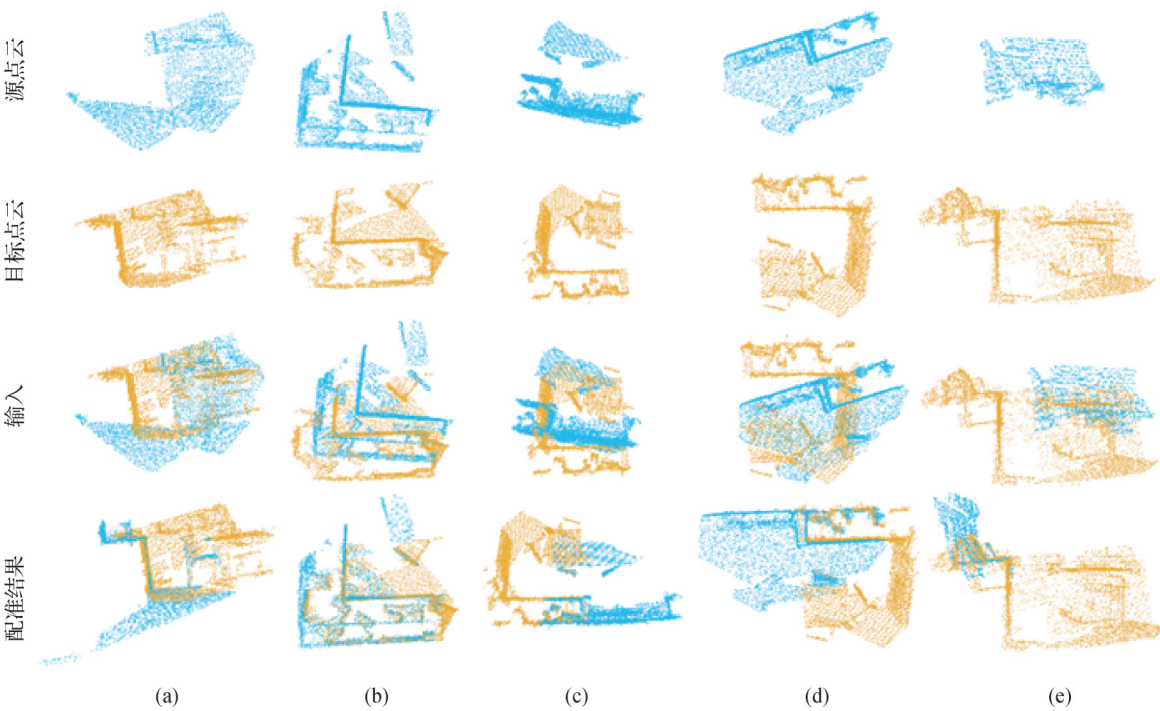


图 5 3DMatch、3DLoMatch 数据集的配准可视化结果

Fig. 5 Visualization of registration results on the 3DMatch and 3DLoMatch datasets

表 3 ModelNet-40、ModelLoNet-40 数据集的实验结果对比

Table 3 Comparison of experimental results on the ModelNet-40 and ModelLoNet-40 datasets

方法	ModelNet-40			ModelLoNet-40		
	RRE /(°)	RTE /m	CD/m	RRE /(°)	RTE /m	CD/m
PointNetLK	29.725	0.297	0.023 50	48.567	0.507	0.036 7
OMNet	2.947	0.032	0.001 50	6.517	0.129	0.007 4
DCP-v2	11.975	0.171	0.011 70	16.501	0.300	0.026 8
RPM-Net	1.712	0.018	0.000 85	7.342	0.124	0.005 0
Predator	1.739	0.019	0.000 89	5.235	0.132	0.008 3
REGTR	1.473	0.014	0.000 78	3.930	0.087	0.003 7
SIRA-PCR	1.431	0.016	0.000 80	3.376	0.071	0.003 5
MV+FCGF	4.319	0.162	0.017 70	21.617	0.252	0.023 8
本文	1.417	0.014*	0.000 81	3.141	0.072	0.003 4

注:加粗字体表示各列最优结果,标*实验结果为0.013 91,在表中保留3位小数。

ModelNet-40 和 ModelLoNet-40 数据集的评估与 3DMatch 数据集处理方式相同,相比端到端深度学习算法, Mutual Voting(MV)可能受到了图处理机制的限制,在捕获和学习数据分布上存在不足。此外,对于这类由 CAD 生成的点云模型, Mutual Voting(MV)

学习的特征和投票机制可能未充分捕捉到足够的空间变换信息,导致其存在较大误差。

此外,在三维模型数据集 ModelNet-40、ModelLoNet-40 上的配准结果如图 6 所示。其中,图 6(a)(b)为 ModelNet-40 的可视化结果,图 6(c)—(e)为 ModelLoNet-40 的可视化结果。虽然 ModelNet-40 相比 3DMatch 的点云数据更稀疏,但本文算法能较准确地匹配两个点云,如图 6(c)的桌腿部分以及图 6(d)车辆的轮胎部分。

3.4 消融实验结果分析

3.4.1 模块贡献度分析

为了验证本文算法各个模块之间的有效性,在室内场景、三维模型数据集上进行测试。本文做的 4 组消融实验(序号 2—5)以及两组对比标准实验(分别为序号 1、6)的结果如表 4 所示。对核函数 kf、层次注意力模块 ha、局部匹配模块 lm、以及局部匹配损失函数 ml,根据它们之间的关联性进行组合测试。为了更好地分析算法在 ModelNet-40、ModelLoNet-40 数据集上的实验结果,在表 4 中的结果比表 3 中多保留了一位小数。当只保留多项式核函数时,由于其衰减是线性的,而且局部聚合能力弱的特点使其在两类数据集上的结果低于序号 1 的基准模型。在序号为 2 的消融实验中,将多项式

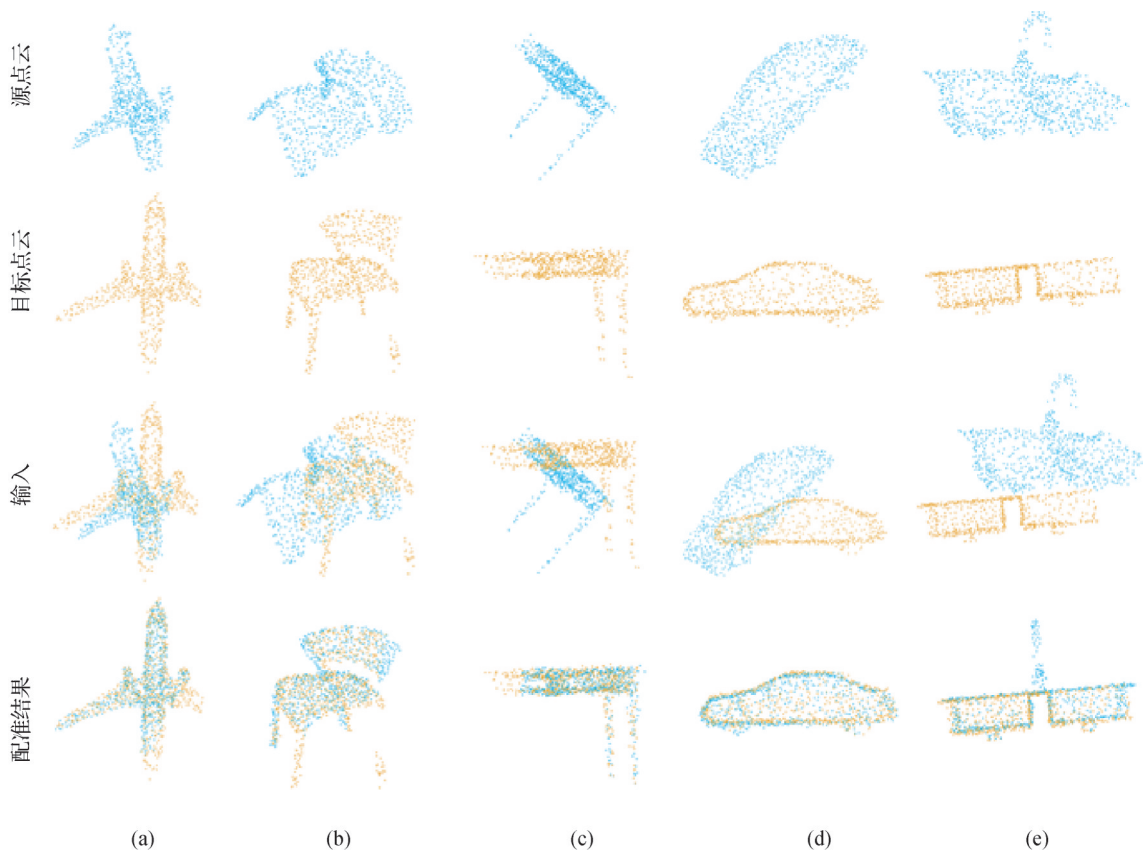


图6 ModelNet-40、ModelLoNet-40数据集的配准可视化结果

Fig. 6 Visualization of registration results on the ModelNet-40 and ModelLoNet-40 datasets

核函数放在第1、3层的EAKPConv中(ModelNet-40的训练模型只放置在第1层)。在序号为3的实验中结合层次注意力、序号为4的实验中结合局部匹配的方法,表明局部注意力预测点对之间的关

系是有效的。在4号实验的基础上,5号实验加入局部匹配损失函数后,配准召回率、倒角误差均有提升,表明本文算法改进的局部匹配损失起到了作用。

表4 本文算法消融实验结果

Table 4 Ablation experiment results of the proposed algorithm

序号	kf.	ha.	lm.	ml.	配准召回率/% ↑		倒角距离/m ↓	
					3DMatch	3DLoMatch	ModelNet-40	ModelLoNet-40
1	-	-	-	-	91.9	61.6	0.000 983	0.004 83
2	√	-	-	-	92.2	62.7	0.000 952	0.004 13
3	√	√	-	-	92.5	65.9	0.000 848	0.003 62
4	-	√	√	-	92.6	66.7	0.000 841	0.003 58
5	-	√	√	√	92.9	67.0	0.000 817	0.003 51
6	√	√	√	√	93.2	67.3	0.000 805	0.003 44

注:“-”表示该组件未加入实验,“√”表示该组件已加入实验;↑表示数值越大性能越好,↓表示数值越小性能越好。

3.4.2 模块可解释性分析

为了评估不同模块在本文算法中的作用,随机选取3个点云场景模型计算热力值,并根据热力值

可视化,如图7所示。其中,EAKPConv采用关键点的特征激活作为热力值,红色圆形区域是由关键点计算得到的近邻区域的可视化结果。HAFM模块和

GCAM 模块将注意力权重作为热力值,在局部注意力模块中提取每层注意力权重,将它们缩放到0~1之间。为了更直接地比较不同模块的作用,本实验将3个模块的热力值均归一化到0~1之间。图中颜色越接近黄色,热力值越趋近1,表示算法对该区域特征的表征越显著,反之对该区域特征的表征越弱。可以看出,EAKPConv 中的3个场景的热力值基本覆盖了关键点近邻区域,并且热力值为黄色的区域基本覆盖了场景的所有区域,说明本文算法采样覆盖

范围是完整且全面的。其次,合并局部注意力权重的 HAFM 模块相比 EAKPConv 所显示的高热力值区域大幅减少,并且3组场景的热力值过渡更自然,说明本文 HAFM 模块将权重聚合在局部特征匹配上。GCAM 模块相比 HAFM 模块高热力值区域更聚焦,并且高亮区域与配准的重叠区域基本重合,如图7重叠部分所示,验证了本文算法在匹配过程中权重更聚焦于重叠区域,因此能得到较好的配准结果。

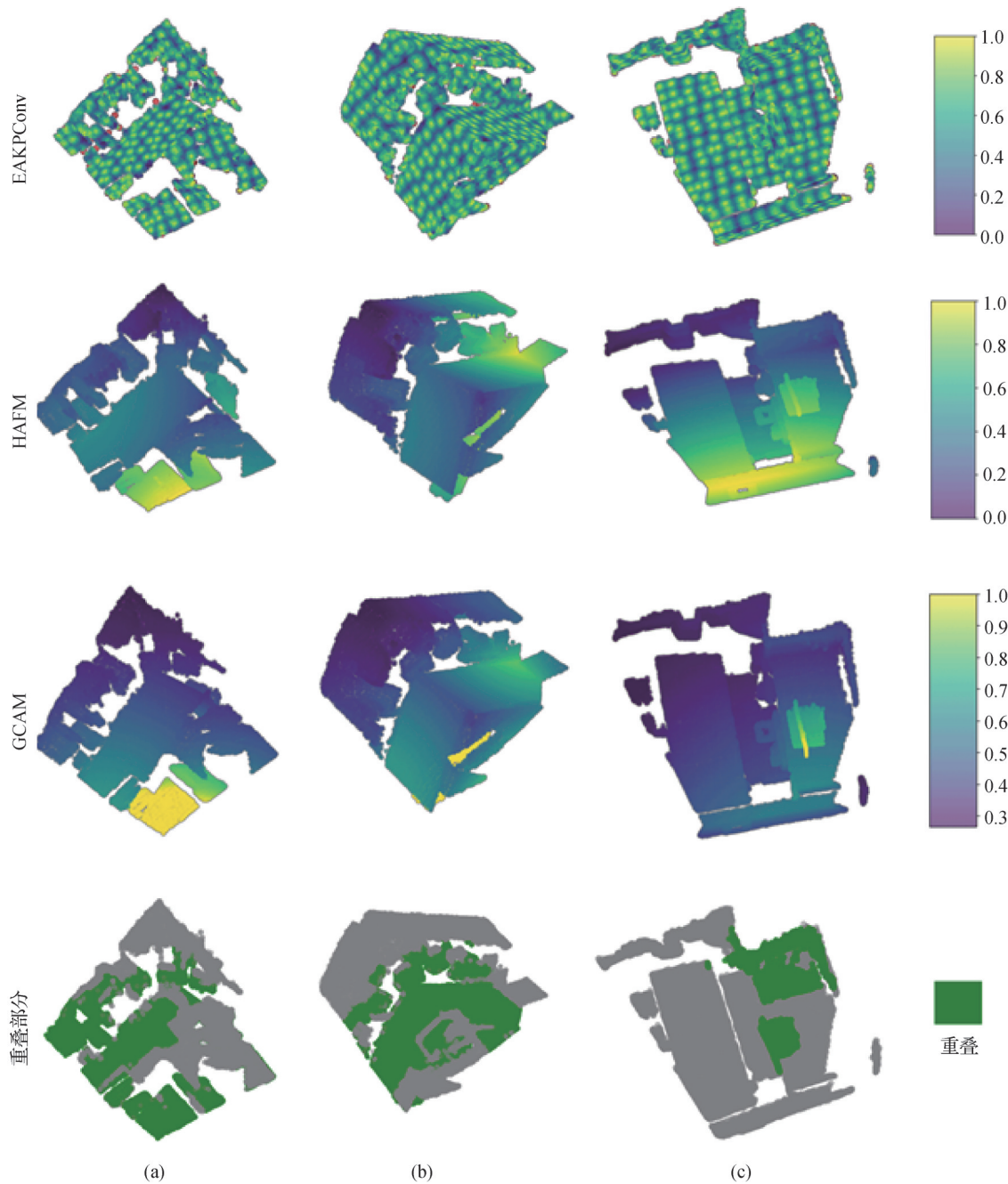


图7 各模块权重可视化结果

Fig 7 Visualization results of weights for each module

3.4.3 重叠率分析

为了验证本文算法在不同重叠率点云对之间的配准有效性,随机选取了3组点云对进行匹配并可视化配准结果,如图8所示。实验中选取重叠率分别为60.3%、36.9%、17.5%的点云对进行配准对比。可以看出,不同重叠率的点云对都能较准确地完成匹配。此外,本实验从两点云上选取各自的关键点,计算关键点之间的对应关系(correspondences),并用

绿色线段表示。为了更好地观察两点云的对应结果,已将源点云向右平移。不同重叠率点云模型的对应点都可以较好地完成对应点匹配。最困难的配准是(c)组实验,从图中可以看到,由于重叠率降低,导致选择的关键点相比(a)、(b)组大幅减少,但本文算法依然可以准确匹配两个模型间的关键点。实验结果表明,本文算法对不同重叠率点云的配准具有较强的一致性。

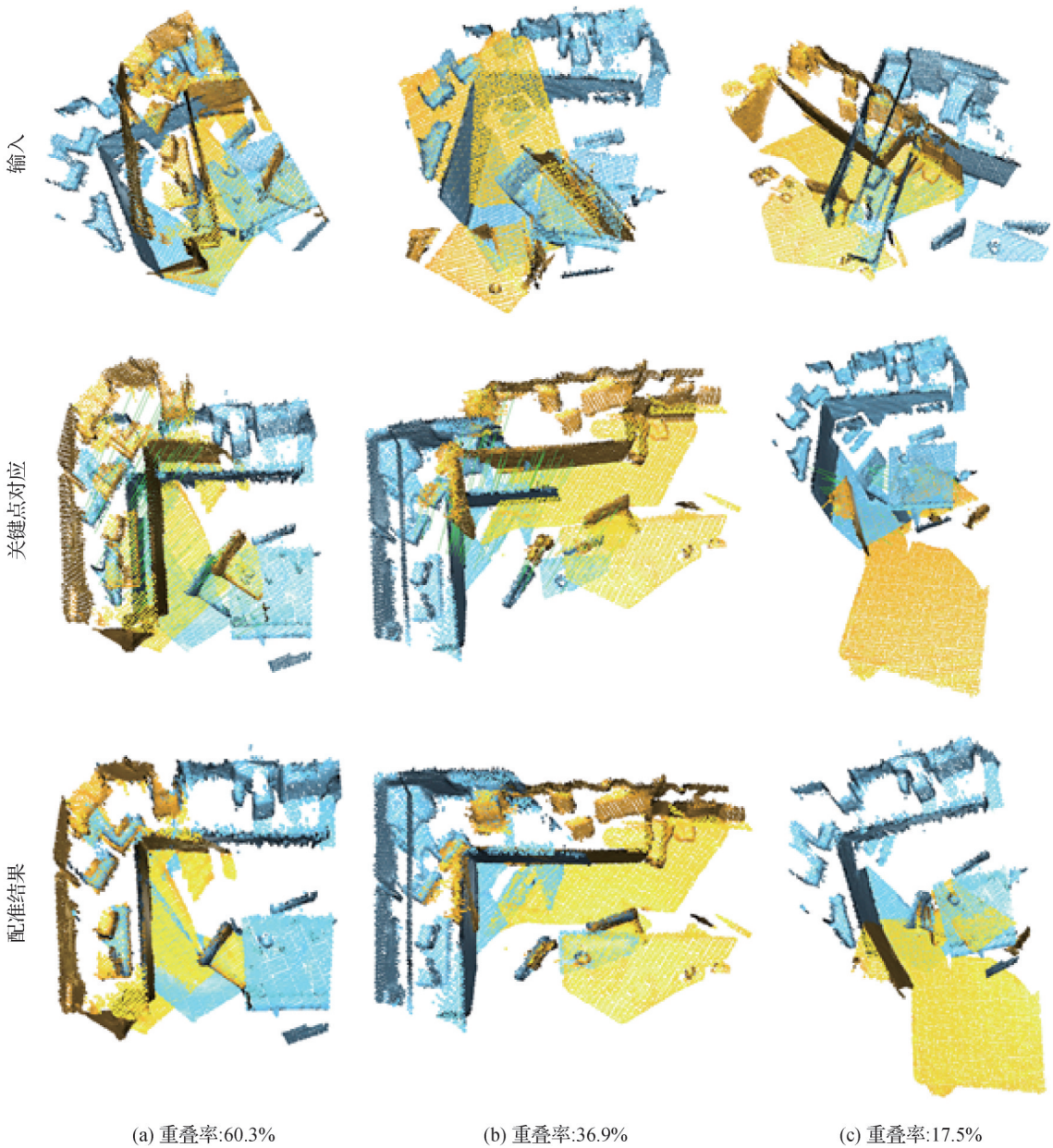


图8 不同重叠率点云配准结果可视化

Fig. 8 Visualization of point cloud registration results with different overlap rates
((a) overlap: 60.3%; (b) overlap: 36.9%; (c) overlap: 17.5%)

3.5 推理效率评估

本文比较了REGTR及本文方法的平均推理时

间,如表5所示。传统注意力计算的时间复杂度为 $O(n^2)$,而本文改进的局部注意力、全局注意力计算

的时间复杂度分别为 $O(m^2)$ 、 $O(l^2)$, 其中 $m^2 + l^2 < n^2$, m, l 分别表示局部邻域维度、全局特征维度。本文算法在局部注意力阶段更关注局部区域, 使算法可以大幅减小计算维度。相比传统注意力 $O(n^2)$ 的时间复杂度, 本文将局部处理的复杂度降至 $O(m^2)$ 。到全局注意力阶段, 将使用局部特征聚合进行全局配准, 由于 HAFM 模块局部层次化处理, 聚合全局特征计算的复杂度 $O(l^2)$ 也小于 $O(n^2)$ 。此外, 在 HAFM 模块中的局部特征部分是并行处理的, 也可以进一步优化推理时间。因此, 在两个数据集的测试结果中, 本文算法的推理所用时间均比 REGTR 更少, 表明本文所使用的层次注意力操作策略的效率高于 REFTR 的传统注意力操作方式。

表 5 本文算法与 REGTR 推理时间对比
Table 5 Comparison of inference time between the proposed algorithm and REGTR

方法	推理时间/ms	
	3DMatch 数据集	ModelNet-40 数据集
REGTR	38.428	40.015
本文	27.205	30.991

4 结 论

针对三维点云配准任务中忽略局部邻域特征的问题, 通过增强对远距离局部边缘结构的构建能力, 同时利用层次注意力模块执行局部匹配, 将局部特征串联得到两个点云的条件特征并预测点云间对应关系。本文在不同数据集验证了算法的改进效果, 在室内场景 3DMatch、3DLoMatch 上分别取得 93.2%、67.3% 的配准召回率。本文在 ModelNet-40 和 ModelLoNet-40 的相对旋转误差、相对平移误差分别为 1.417、0.013 91 和 3.141、0.072。消融实验证明本文所提出的层次注意力模块获得比标准注意力模块更好的配准结果, 而且推理时间大幅减小。然而, 本文算法在训练中使用了标注的变换矩阵, 这对算法应用造成了限制。因此, 计划未来在现有算法的基础上对点云无监督或弱监督匹配进行深入研究。

参考文献 (References)

Aoki Y, Goforth H, Srivatsan R A and Lucey S. 2019. PointNetLK;

robust and efficient point cloud registration using PointNet//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 7156-7165 [DOI: 10.1109/CVPR.2019.00733]
Bai X Y, Luo Z X, Zhou L, Fu H B, Quan L and Tai C L. 2020. D3Feat: joint learning of dense detection and description of 3D local features//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 6358-6366 [DOI: 10.1109/CVPR42600.2020.00639]
Bauer D, Patten T and Vincze M. 2021. ReAgent: point cloud registration using imitation and reinforcement learning//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 14581-14589 [DOI: 10.1109/cvpr46437.2021.01435]
Berg A, Oskarsson M and O' Connor M. 2022. Points to patches: enabling the use of self-attention for 3D shape recognition//Proceedings of the 26th International Conference on Pattern Recognition (ICPR). Montreal, Canada: IEEE: 528-534 [DOI: 10.1109/ICPR56361.2022.9956172]
Besl P J and McKay N D. 1992. A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2): 239-256 [DOI: 10.1109/34.121791]
Cao A Q, Puy G, Boulch A and Marlet R. 2021. PCAM: product of cross-attention matrices for rigid registration of point clouds//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE: 13209-13218 [DOI: 10.1109/ICCV48922.2021.01298]
Chen S Y, Xu H, Li R, Liu G H, Fu C W and Liu S C. 2023. SIRA-PCR: sim-to-real adaptation for 3D point cloud registration//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE: 14348-14359 [DOI: 10.1109/ICCV51070.2023.01324]
Chen Z L, Chen H H, Gong L N, Yan X F, Wang J, Guo Y W, Qin J and Wei M Q. 2022. UTOPIc: uncertainty-aware overlap prediction network for partial point cloud registration. Computer Graphics Forum, 41(7): 87-98 [DOI: doi.org/10.1111/cgf.14659]
Choy C, Park J and Koltun V. 2019. Fully convolutional geometric features//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE: 8957-8965 [DOI: 10.1109/ICCV.2019.00905]
Huang S Y, Gojcic Z, Usvyatsov M, Wieser A and Schindler K. 2021. PREDATOR: registration of 3D point clouds with low overlap//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE: 4265-4274 [DOI: 10.1109/CVPR46437.2021.00425]
Huang X S, Mei G F and Zhang J. 2020. Feature-metric registration: a fast semi-supervised approach for robust point cloud registration

- without correspondences//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA; IEEE: 11363-11371 [DOI: 10.1109/CVPR42600.2020.01138]
- Huang Z L, Wang X G, Wei Y C, Huang L C, Shi H, Liu W Y and Huang T S. 2023. CCNet: criss-cross attention for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45 (6) : 6896-6908 [DOI: 10.1109/TPAMI.2020.3007032]
- Pais G D, Ramalingam S, Govindu V M, Nascimento J C, Chellappa R and Miraldo P. 2020. 3DRegNet: a deep neural network for 3D point registration//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA; IEEE: 7191-7201 [DOI: 10.1109/cvpr42600.2020.00722]
- Qi C R, Su H, Mo K C and Guibas L J. 2017. PointNet: deep learning on point sets for 3D classification and segmentation//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- Qin H X, Liu Z T and Tan B Y. 2022. Review on deep learning rigid point cloud registration. *Journal of Image and Graphics*, 27 (2) : 329-348 (秦红星, 刘镇涛, 谭博元. 2022. 深度学习刚性点云配准前沿进展. *中国图象图形学报*, 27(2): 329-348) [DOI: 10.11834/jig.210556]
- Ross G T and Soland R M. 1975. A branch and bound algorithm for the generalized assignment problem. *Mathematical Programming*, 8 (1) : 91-103 [DOI: 10.1007/BF01580430]
- Rusu R B, Blodow N and Beetz M. 2009. Fast point feature histograms (FPFH) for 3D registration//Proceedings of 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan; IEEE: 3212-3217 [DOI: 10.1109/ROBOT.2009.5152473]
- Rusu R B, Marton Z C, Blodow N and Beetz M. 2008. Persistent point feature histograms for 3D point clouds//Proceedings of the 10th International Conference on Intelligent Autonomous Systems (IAS-10). Baden-Baden, Germany; IOS: 119-128 [DOI: 10.3233/978-1-58603-887-8-119]
- Sarode V, Li X Q, Goforth H, Aoki Y, Srivatsan R A, Lucey S and Choset H. 2019. PCRNet: point cloud registration network using PointNet encoding [EB/OL]. [2024-03-05]. <https://arxiv.org/pdf/1908.07906.pdf>
- Thomas H, Qi C R, Deschaud J E, Marcotegui B, Goulette F and Guibas L. 2019. KPConv: flexible and deformable convolution for point clouds//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South); IEEE: 6410-6419 [DOI: 10.1109/ICCV.2019.00651]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA; Curran Associates Inc.: 6000-6010
- Wang Y and Solomon J. 2019a. PRNet: self-supervised learning for partial-to-partial registration//Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, Canada; Curran Associates Inc.: 8814-8826
- Wang Y and Solomon J. 2019b. Deep closest point: learning representations for point cloud registration//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South); IEEE: 3522-3531 [DOI: 10.1109/ICCV.2019.00362]
- 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): #146 [DOI: 10.1145/3326362]
- Wu Y, Yuan Y Z, Xiang B H, Sheng J L, Lei J Y, Hu C Y, Gong M G, Ma W P and Miao Q G. 2023. Overview of the computational intelligence method in 3D point cloud registration. *Journal of Image and Graphics*, 28(9): 2763-2787 (武越, 苑咏哲, 向本华, 绳金龙, 雷佳熠, 胡聪颖, 公茂果, 马文萍, 苗启广. 2023. 三维点云配准中的计算智能方法综述. *中国图象图形学报*, 28(9): 2763-2787) [DOI: 10.11834/jig.220727]
- 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 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA; IEEE: 1912-1920 [DOI: 10.1109/CVPR.2015.7298801]
- Xie Z Y, Chen J Z and Peng B. 2020. Point clouds learning with attention-based graph convolution networks. *Neurocomputing*, 402: 245-255 [DOI: 10.1016/j.neucom.2020.03.086]
- Xu H, Liu S X, Wang G F, Liu G H and Zeng B. 2021. OMNet: learning overlapping mask for partial-to-partial point cloud registration//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada; IEEE: 3112-3121 [DOI: 10.1109/ICCV48922.2021.00312]
- Yan Y J, An J Y, Zhao J and Shen F R. 2023. Hybrid optimization with unconstrained variables on partial point cloud registration. *Pattern Recognition*, 136: #109267 [DOI: 10.1016/j.patcog.2022.109267]
- Yang J and Zhang C. 2023. Semantic segmentation of 3D point cloud by fusing dual attention mechanism and dynamic graph convolution neural network (杨军, 张琛. 2023. 融合双注意力机制和动态图卷积神经网络的三维点云语义分割) [EB/OL]. [2023-12-12]. <https://bxb.buaa.edu.cn/bhbk/cn/article/doi/10.13700/j.bh.1001-5965.2022.0775>
- Yang J, Zhang Y and Huang L. 2018. Research on 3D model registration by improved ICP algorithm. *Journal of Frontiers of Computer*

- Science and Technology, 12(1): 153-162 (杨军, 张瑶, 黄亮. 2018. 改进的 ICP 算法在三维模型配准中的研究. 计算机科学与探索, 12(1): 153-162) [DOI: 10.3778/j.issn.1673-9418.1610033]
- Yang J L, Li H D, Campbell D and Jia Y D. 2016. Go-ICP: a globally optimal solution to 3D ICP point-set registration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(11): 2241-2254 [DOI: 10.1109/TPAMI.2015.2513405]
- Yang J Q, Zhang X Y, Fan S C, Ren C L and Zhang Y N. 2024. Mutual voting for ranking 3D correspondences. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(6): 4041-4057 [DOI: 10.1109/TPAMI.2023.3268297]
- Yew Z J and Lee G H. 2020. RPM-Net: robust point matching using learned features//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 11821-11830 [DOI: 10.1109/CVPR42600.2020.01184]
- Yew Z J and Lee G H. 2022. REGTR: end-to-end point cloud correspondences with transformers//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 6667-6676 [DOI: 10.1109/CVPR52688.2022.00656]
- Zeng A, Song S R, Nießner M, Fisher M, Xiao J X and Funkhouser T. 2017. 3DMatch: learning local geometric descriptors from RGB-D reconstructions//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 199-208 [DOI: 10.1109/CVPR.2017.29]
- Zhou Q Y, Park J and Koltun V. 2016. Fast global registration//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherlands: Springer: 766-782 [DOI: 10.1007/978-3-319-46475-6_47]

作者简介

杨军,男,教授,博士生导师,主要研究方向为深度学习、三维模型空间分析、遥感大数据智能解译和地理信息系统。

E-mail: yangj@mail.lzjtu.cn

孙鸿炜,男,硕士研究生,主要研究方向为计算机视觉和模式识别。E-mail: sunhongwei1020@yeah.net