文章编号:1674-2974(2025)04-0034-10 DOI:10.16339/j.cnki.hdxbzkb.2025264

基于显著性峰值与特征对齐的点云配准模型

张富凯[1,2†],张露露[1],张海燕[3],张艳梅[4],袁冠[4],赵珊[1],王登科[2],霍占强[1],陈玮[1]
(1. 河南理工大学 软件学院,河南 焦作 454000;
2. 河南理工大学 河南省瓦斯地质与瓦斯治理重点实验室,河南 焦作 454000;
3. 河南理工大学 计算机科学与技术学院,河南 焦作 454000;
4. 中国矿业大学 计算机科学与技术学院,江苏 徐州 221116)

摘 要:点云配准的核心是估算变换矩阵.两个点云对之间存在部分重叠、高噪声和密度差异,现有方法无法精确解决显著点云对应关系之间特征对齐问题.因此,提出显著性峰值与特征对齐网络(significance peak and feature alignment network, SPFANet),实现由粗到细的点云配准.SPFANet由多显著性峰值检测器、粗配准和细配准三部分组成.首先,多显著性峰值检测器引入一种基于描述符方差和重叠分数重加权峰值损失方法,去除非歧视与非重叠的关键点云;其次,粗配准阶段通过检测互补关键点集来计算粗步的配准方案;最后,细配准阶段引入带有前向后向变换的特征度量框架细化粗配准,完成高效点云配准.在同源3DMatch数据集和跨源3DCSR数据集上的实验验证了SPFANet的有效性.

关键词:点云配准;跨源点云;多显著性峰值检测;特征度量;重加权峰值损失
中图分类号:TP391.41 **文献标志码**:A

Point Cloud Registration Model Based on Significance Peak and Feature Alignment

ZHANG Fukai[1,2†], ZHANG Lulu[1], ZHANG Haiyan[3], ZHANG Yanmei[4], YUAN Guan[4], ZHAO Shan[1], WANG Dengke[2], HUO Zhanqiang[1], CHEN Wei[1]
(1. School of Software, Henan Polytechnic University, Jiaozuo 454000, China;
2. State Key Laboratory Cultivation Base for Gas Geology and Gas Control, Henan Polytechnic University, Jiaozuo 454000, China;
3. School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China;
4. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: The core of point cloud registration is to estimate the transformation matrix. There is partial overlap, high noise, and density difference between two point cloud pairs. The existing methods cannot accurately solve the problem of feature alignment between significant point cloud correspondences. Therefore, a significance peak and feature alignment network (SPFANet) is proposed to achieve point cloud registration from coarse to fine one.

* **收稿日期**:2024-02-18
基金项目:国家自然科学基金资助项目(71774159,52174174), National Natural Science Foundation of China (71774159,52174174);河南省科技攻关项目(252102320210), Henan Province Science and Technology Research (252102320210);河南省高等学校重点科研项目(24B520014), Key Scientific Research Projects of Colleges and Universities in Henan Province (24B520014);河南理工大学青年骨干教师资助计划(2023XQG-14), Henan Polytechnic University Youth Backbone Teacher Support Program (2023XQG-14)
作者简介:张富凯(1986—),男,河南焦作人,河南理工大学副教授,博士(后),硕士生导师
† 通信联系人,E-mail: zhangfukai@hpu.edu.cn

SPFANet consists of three parts: multi-significance peak detector, coarse registration, and fine registration. Firstly, the multi-significance peak detector introduces a re-weighted peak loss method based on descriptor variance and overlap score to remove non discriminatory and non overlapping key point clouds. Secondly, the coarse registration stage detects the complementary key point sets to compute the coarse registration scheme. Finally, the fine registration stage introduces a feature metric framework with a forward-backward transform to refine the coarse registration scheme and achieve efficient point cloud registration. The effectiveness of SPFANet is validated through experiments on the same source 3DMatch dataset and cross-source 3DCSR dataset.

Key words: point cloud registration; cross-source point cloud; multi-significance peak detection; feature metric; re-weighted peak loss

点云配准(point cloud registration, PCR)旨在恢复同一场景中两个点云之间的变换矩阵,对于自动驾驶、机器人和增强现实等计算机视觉应用至关重要.根据获取点云的传感器类型,可将点云划分为同源点云和跨源点云.由于传感器的成像机制不同,跨源点云更复杂,同时存在密度差异和尺度差异等挑战.这些挑战意味着源点云中的点可能不在其对应目标点云的相应位置上,即目标点云不是由源点云直接变换而来.因此,两个点云之间可能不存在完美的一一对应关系,这是配准面临的主要难题.

近年来,多数国内外学者不断尝试解决上述问题,提高配准精度和效率.一些传统配准方法如ICP[1]和Go-ICP[2]依赖于对应搜索和变换估计搜索对应点解决点云配准问题.然而,ICP在计算点间几何投影误差以寻找点对应时,对异常值和噪声敏感.为消除噪声和异常值的影响,一些方法利用高斯模型[3]、图匹配[4]、张量优化[5]、RANSAC[6]及其变体[7]进行点云配准,却无法解决部分重叠和大尺度差异带来的难题.一些方法利用特征描述符搜寻对应关系,如NARF[8]、KPQ[9]以及ISS[10],或通过关键点检测方法建立对应关系,如PFH[11]、FPFH[12]等,从而在外观变化(如昼夜、天气和季节变化)的条件下比传统方法有显著改善.然而,面对部分重叠的点云,特征描述符或关键点的准确性较低.FGR[13]、3DMatch[14]、D3Feat[15]、陆军等人[16]、李雪梅等人[17]、熊丰伟等人[18]利用深度特征来估计准确的对应关系进而计算最佳变换,然而这些方法无法消除部分重叠带来的影响.一些基于特征度量的点云配准方法,如PointNetLK[19]、FMR[20]以及PointNetLK Revisited[21],提议跳过对应搜索,直接将特征学习与逆合成算法相结合,从而最小化点云对特征投影的误差.这类算法虽然对噪声和异常值具有鲁棒性,但在处理部分重叠的点云时依旧会遇到困难.为应对部分重叠带来的挑战,张元等人[22]、阎翔鑫等人[23]、易见兵等人[24]分别从不同角度约束和多尺度融合等方面解决点云配准问题,但无法消除不完美对应的影响.随着深度学习的发展,Predator[25]、IDAM[26]、REGTR[27]以及Geo-Transformer[28]方法试图结合上下文信息和配对点云信息的相关性来消除非重叠区域的影响,但它们仍然需要通过计算对应关系来估计变换,因而无法有效避免真实对应之间不存在完美对应的影响.

综上,点云配准存在难以解决的部分重叠问题以及真实对应之间不存在完美对应的挑战,为此,本文提出显著性峰值与特征对齐网络(significance peak and feature alignment network, SPFANet),着重解决点云对非重叠区域信息冗余以及显著点云特征对齐问题.首先,采用核点卷积骨干网络和Transformer编码结构提取点云内部交互特征,生成多个显著性峰值检测热图,同时为了抑制非重叠区域以及特征区分性较弱的区域产生关键点,引入基于特征方差和重叠分数的重加权峰值损失训练关键点检测器;其次,通过非极大值抑制获得特征热图的候选关键点,采用基于全局特征相关性的粗匹配方法提升配准效率;最后,采用基于前向后向的无监督方式预训练细配准模块,同时通过计算关键点云之间的特征度量误差细化配准方案,降低真实对应之间不存在完美对应的影响,从而提高配准精度.在多个不同数据集上的实验验证了SPFANet的有效性.

1 显著性峰值与特征对齐的点云配准

给定一组点云对 $P \in \mathbb{R}^{N\times 3}$ 和 $Q \in \mathbb{R}^{M\times 3}$,希望找到一个变换矩阵使得 P 和 Q 之间的空间位置差别最小.网络通过最小化位于重叠区域的可辨别性关键点集的特征投影误差获得最佳变换,如下式所示:

$$\operatorname{argmin}\left\| F\left(\tilde{P}\right)-F\left(\boldsymbol{R}\times\tilde{Q}+\boldsymbol{t}^{*}\right)\right\|_2^2 \tag{1}$$

式中:$\tilde{P} \in \mathbb{R}^{\tilde{N}\times 3}$ 和 $\tilde{Q} \in \mathbb{R}^{\tilde{M}\times 3}$ 表示 P 和 Q 的重叠区域,$F\left(\tilde{P}\right) \in \mathbb{R}^{\tilde{N}\times C}$ 和 $F\left(\tilde{Q}\right) \in \mathbb{R}^{\tilde{M}\times C}$ 代表对应的特征;$\boldsymbol{R}$ 是变换矩阵中的旋转矩阵;$\boldsymbol{t}^{*}$ 是变换矩阵中的平移向量.

SPFANet 的整体结构如图 1 所示,点云对首先被送入显著性峰值检测器中,获得多个显著性峰值特征热图,检测位于重叠区域的不同可辨别性关键点集;然后基于全局特征相关性匹配方案进行粗匹配;最后利用特征误差计算和前向后向变换矩阵迭代进行细配准,获得最终的变换矩阵.

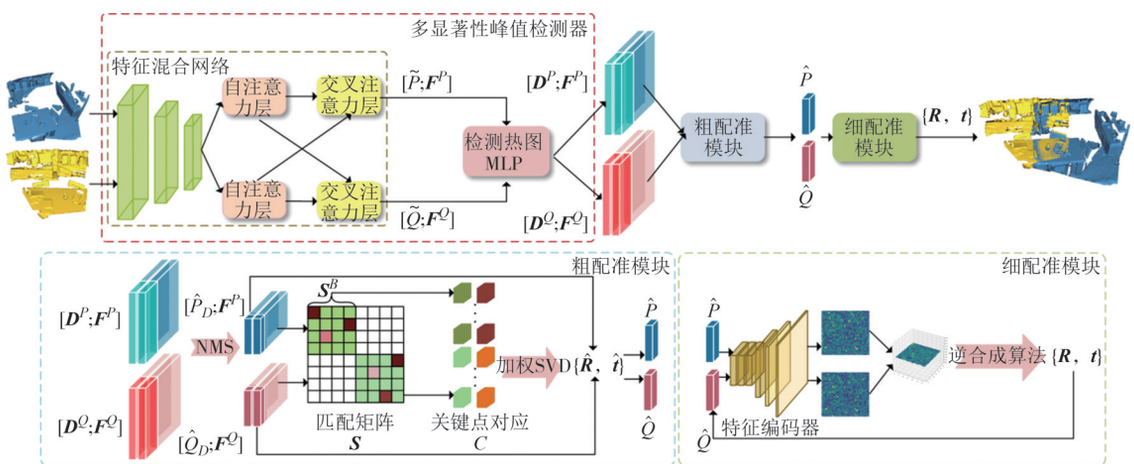


图 1 本文网络结构图

Fig. 1 Structure of the proposed network

1.1 多显著性峰值检测器

多显著性峰值检测器旨在从同一点云中提取不相交的关键点集,降低点云匹配复杂度,包括点云特征提取与特征检测热图生成两部分.

1.1.1 点云特征提取

采用特征混合网络获得输入点云的几何上下文信息以及两个点云之间的混合交互特征.特征提取过程如下:

$$\bar{F}^{P},\tilde{P}=f(P)_{\text{mix}} \tag{2}$$

其中,$f(\cdot)_{\text{mix}}$ 是由一系列类似于 ResNet 的跨步卷积块、Transformer 自注意力模块和交叉注意力模块组成的特征混合网络,获得原始点云的简化版本 $\tilde{P} \in \mathbb{R}^{\tilde{N}\times 3}$ 和 $\tilde{Q} \in \mathbb{R}^{\tilde{M}\times 3}$.$\bar{F}^{P} \in \mathbb{R}^{\tilde{N}\times C}$ 和 $\bar{F}^{Q} \in \mathbb{R}^{\tilde{M}\times C}$ 分别表示 $\dot{P}$ 和 $\dot{Q}$ 的混合特征.

1.1.2 特征检测热图 MLP

获得的点云混合特征被送入检测热图 MLP 中,以生成不同的显著性峰值特征检测热图.检测热图中的每个元素表示在点云中对应点的显著性,即每个点在点云中是否具有判别性以及重叠性.对于两个点云,处于重叠区域中的点有找到匹配对的可能性,而位于判别性区域的点更容易产生正确的对应关系.热图生成步骤如下:

$$\boldsymbol{D}^{P}=\text{MLP}\left(\bar{\boldsymbol{F}}^{P}\right) \tag{3}$$

其中,MLP(·)由三个线性层、两个 ReLU 激活函数以及一个 Sigmoid 层组成,用于生成检测热图 $\boldsymbol{D}^{P} \in \mathbb{R}^{B\times\tilde{N}}$,其中 B 是关键点集的数量,$\tilde{N}$ 是每个检测热图提取的关键点数量.事实上,生成的每个检测热图 $\boldsymbol{D}_b^{P} \in \mathbb{R}^{\tilde{N}}$ 都将用于提取一组关键点,其中 b=1,2,…,B−1.目标点云的检测热图 $\boldsymbol{D}^{Q} \in \mathbb{R}^{B\times\tilde{M}}$ 以相同的方式获得.

1.2 基于候选关键点的点云粗配准模块

1.2.1 候选关键点

得到显著性峰值检测热图之后,需要提取关键点并进行粗匹配.前面提到检测图中每个元素表示对应点的显著性分数,那么候选关键点的选择将在显著性峰值中产生.首先,对于每个热图 $\boldsymbol{D}_b^{P}$,候选关键点 $\dot{P}_B$ 被确立为每个热图局部峰值的点坐标.通过应用局部非极大值抑制(NMS)确立局部峰值:

$$\dot{P}_B=\text{NMS}\left(\boldsymbol{D}^{P},\dot{P}\right) \tag{4}$$

每个热图选择的关键点个数由总关键点个数 K

和检测热图个数B确定,即对于每个显著性热图,仅选择K/B个局部峰值作为候选关键点.通过将关键点与其对应特征进行耦合获取局部特征$\dot{\boldsymbol{F}}_B^P$.目标点云的候选关键点$\dot{\boldsymbol{Q}}_B$及其局部特征$\dot{\boldsymbol{F}}_B^Q$以同样的方式获取.

1.2.2　粗匹配模块

为了找到正确的且无歧义的对应关系,选择一种基于全局特征相关性的匹配方案.值得注意的是,本文的检测器检测一个点云的多个不同关键点集,则每个源关键点集只与特征最相关的对应目标关键点集计算匹配矩阵.如图 1 所示,这种匹配矩阵$\boldsymbol{S}^B \in \mathbb{R}^{(K/B)\times(K/B)}$将每个集合的相应计算复杂度降低到$O\left(K^2/B^2\right)$.

首先,将局部特征$\dot{\boldsymbol{F}}_B^P$和$\dot{\boldsymbol{F}}_B^Q$归一化到一个单位球体上并计算其高斯相关性矩阵$\boldsymbol{S}^B$,其中每个元素值被计算为点云对之间的匹配分数:

$$s_{b_{i,j}} = \exp\left(-\left\| f_{b_{i,j}}^P - f_{b_{i,j}}^Q \right\|_2^2\right) \tag{5}$$

其中,b=1,2,…,B−1,$f_{b_{i,j}}^P$和$f_{b_{i,j}}^Q$分别是点云P和Q的第b个关键点集上的某个关键点的特征.点云的某些点判别性比较低,可能在其他点云中有很多相似的块.

其次,在$\boldsymbol{S}^B$上执行双重归一化操作以进一步抑制歧义匹配:

$$\bar{s}_{b_{i,j}} = \frac{s_{b_{i,j}}}{\sum_{c=1}^{K/B} s_{b_{i,c}}} \cdot \frac{s_{b_{i,j}}}{\sum_{c=1}^{K/B} s_{b_{c,j}}} \tag{6}$$

这种方式可以有效消除错误匹配.

然后,对应关系被选择为所有$\dot{\boldsymbol{S}}^B$中元素最大的N_c个点对,如下:

$$C = \left\{ \left(p_{x_i}, q_{y_i} \right) \middle| \left(x_i - y_i \right) \in \mathrm{top}k_{x,y}\left(\bar{s}_{x,y} \right) \right\} \tag{7}$$

其中,$\bar{s}_{x,y}$是将每个$\dot{\boldsymbol{S}}^B$展平拼接后的每个元素.

最后,使用对应点集C中的点对应关系来求解粗略变换$\left[\hat{\boldsymbol{R}}, \hat{\boldsymbol{t}}\right]$,如下:

$$\hat{\boldsymbol{R}}, \hat{\boldsymbol{t}} = \arg\min_{\boldsymbol{R},\boldsymbol{t}} \sum_i^{N_c} \boldsymbol{w}_i \left\| \boldsymbol{R}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{y}_i \right\|^2 \tag{8}$$

其中,$\left(x_i - y_i\right) \in C$,$\boldsymbol{w}_i \in \mathbb{R}^{N_c}$是每对对应点的显著性分数均值.这可以使用加权 SVD[29]以封闭性形式求解.

得益于高质量的对应关系,这个阶段获得的变换已经比较准确了.得到的关键点$\hat{P}$和$\hat{Q}$以及粗略配准结果$\left[\hat{\boldsymbol{R}}, \hat{\boldsymbol{t}}\right]$被送入细配准阶段对配准结果进行微调.

1.3　基于特征对齐的点云细配准模块

获得新的成对点云之后,使用简单有效的细配准模块细化粗略变换.利用点云整体结构信息,通过迭代逆合成算法(inverse composition, IC)[30]最小化两个点云之间的特征误差以进一步估计变换矩阵.

如 Algorithm 1 所示,细配准模块的编码器能够提取具有旋转注意的独特点云特征.其中,特征编码主干是预训练网络,其参数在整体配准工程中不会更新.

Algorithm 1: Iterative Inverse Composition Algorithm	
Input: Key point sets from two point clouds P and Q; Initial registration $\left[\hat{\boldsymbol{R}}, \hat{\boldsymbol{t}}\right]$; Feature encoder network $f_e(\cdot)$.	
Output: Optimal transformation matrix $\left[\boldsymbol{R}_{\mathrm{est}}, \boldsymbol{t}_{\mathrm{est}}\right]$.	
1	max_iter = 5
2	$\left[\boldsymbol{R}_1, \boldsymbol{t}_1\right] = \left(\hat{\boldsymbol{R}}, \hat{\boldsymbol{t}}\right)$
3	$\hat{Q}_1 = \boldsymbol{R}_1 \hat{P} + \boldsymbol{t}_1$
4	for i = 1 to max_iter **do**
5	$F_i^P = f_e\left(\hat{P}\right)$, $F_i^P = f_e\left(\hat{Q}_i\right)$
6	$\boldsymbol{r} = \left\| f_e\left(\hat{P}\right) - f_e\left(\hat{Q}_i\right) \right\|_2^2$
7	$\zeta = \left[\boldsymbol{R}_i, \boldsymbol{t}_i\right]$
8	$\boldsymbol{J}_i = \left(f_e\left(\boldsymbol{R}_i \hat{P} + \boldsymbol{t}_i\right) - f_e\left(\hat{P}\right) \right) / \zeta$
9	$\Delta\theta = \left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{J}\right)^{-1}\left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{r}\right)$
10	$\theta_{i+1} = \theta_i \Delta\theta$
11	$\hat{Q}_{i+1} = \exp\left(\theta_{i+1}\right)\hat{Q}_i$
12	$\left[\boldsymbol{R}_{i+1}, \boldsymbol{t}_{i+1}\right] = \exp\left(\theta_{i+1}\right)$
13	**end for**
14	$\left[\boldsymbol{R}_{\mathrm{est}}, \boldsymbol{t}_{\mathrm{est}}\right] = \left[\boldsymbol{R}_{\mathrm{max_iter}}, \boldsymbol{t}_{\mathrm{max_iter}}\right]$

1.4　损失函数

SPFANet 由特征损失、检测器损失、相异损失和倒角损失函数组成.

1.4.1　特征损失

为识别有区别的点云特征描述符,采用 InfoNCE 损失函数约束两个点云关键点之间的对应关系.InfoNCE 损失促使两个对应的关键点具有相似的描述符,而不对应的关键点应该具有不相似的描述符.

对于$p_i \in P$与其在Q中的对应关系,源点云的 InfoNCE 损失为:

$$L_{\mathrm{f}} = -E_{p\in P}\left[\log \frac{f\left(p, \Delta p\right)}{f\left(p, \Delta p\right) + \sum_{\Delta n} f\left(p, \Delta n\right)} \right] \tag{9}$$

其中,$f(\cdot,\cdot)$表示一种双线性模型,如下:

$$f\left(p, c\right) = \exp\left(\bar{f}_p^{\mathrm{T}} \boldsymbol{w}_f f_{\mathrm{c}}\right) \tag{10}$$

这里,$\bar{f}_p^{\mathrm{T}}$是点p的特征,Δp和Δn是点p在Q中匹配和不匹配的关键点,其分别由设置为$(m, 2m)$的正边距r_p和负边距r_n确定.m是下采样层中使用的体素距离,所有落在负边界之外的负点都用于Δn.此外,$\boldsymbol{w}_f = \boldsymbol{U}_f + \boldsymbol{U}_f^{\mathrm{T}}$是可学习参数,$\boldsymbol{U}_f$是上三角矩阵,通过这种方式强制$\boldsymbol{w}_f$成为对称函数.

1.4.2 检测器损失

为了鼓励网络产生分布良好的局部显著性峰值的检测热图,同时避免歧视性区域,检测器采用一种重加权峰值损失,定义如下:

$$L_{\mathrm{peaky}}^{P}\boldsymbol{D}_{d_i}^{P} = \boldsymbol{w}_i\left(1 - \left(\max_{\mathrm{patch}\in P_i}\boldsymbol{D}_{d_{\mathrm{patch}}}^{P} - \operatorname*{mean}_{\mathrm{patch}\in P}\boldsymbol{D}_{d_{\mathrm{patch}}}^{P}\right)\right) \tag{11}$$

式中:P_i是以关键点p_i为中心的局部点云块,$\boldsymbol{w}_i$是权重,旨在避免检测器在非重叠性区域以及不具有区分性的区域出现峰值.其定义如下:

$$\boldsymbol{w}_i = \boldsymbol{o}_i \cdot \operatorname*{mean}_{c\in\boldsymbol{C}}\left(\operatorname*{mean}_{\mathrm{patch}\in P_i}F_{\mathrm{patch}}^2 - \left(\operatorname*{mean}_{\mathrm{patch}\in P_i}F_{\mathrm{patch}}\right)^2\right)_c \tag{12}$$

这里,$\boldsymbol{w}_i$由特征局部方差与重叠分数相乘计算得到,局部方差在以关键点p_i为中心的局部点云块上计算获得,在特征通道维度上取平均值.其中,$\boldsymbol{o}_i$是点云的地面真实标签,其定义方式与Predator[25]类似,如下所示:

$$\boldsymbol{o}_i = \begin{cases} 1, \left\| T^*(p_i) - \mathrm{NN}\left(T^*(p_i), Q\right)\right\|_2 < r_o \\ 0, \text{otherwise} \end{cases} \tag{13}$$

式中:$T^*(\cdot)$代表地面真实变换矩阵;$\mathrm{NN}(\cdot)$代表空间最近邻;r_o是重叠阈值.

目标点云Q的峰值损失由同样的方式计算.总的峰值损失计算为:

$$L_{\mathrm{peaky}} = \frac{1}{2}\left(L_{\mathrm{peaky}}^{P} + L_{\mathrm{peaky}}^{Q}\right) \tag{14}$$

1.4.3 相异性损失

为了促进$\boldsymbol{D}$中的每个检测热图都关注不同的区域,避免产生交叉关键点集,提出一种相异性损失,惩罚每对检测图$\left(\boldsymbol{D}_x, \boldsymbol{D}_y\right)$的同位峰值,如下:

$$L_{\mathrm{dissim}}^{P} = \mathrm{mean}\left(\binom{B}{2}^{-1}\Sigma_{(x,y)\in S}\boldsymbol{D}_x \times \boldsymbol{D}_y\right) \tag{15}$$

其中,B是指检测热图的数量,二项式$\binom{B}{2}$是可能的检测器热图组合的数量.同时,目标点云Q的相异性损失L_{dissim}^{Q}可由同样的方式计算获得.总的相异性损失为:

$$L_{\mathrm{dissim}} = \frac{1}{2}\left(L_{\mathrm{dissim}}^{P} + L_{\mathrm{dissim}}^{Q}\right) \tag{16}$$

1.4.4 倒角损失

为了提高细配准的质量,预训练特征编码器十分重要.为此,本文引入一种前向后向旋转模式,以无监督的方式训练该编码器网络.如图2所示,在前向阶段,点云被送入编码器提取特征之后,又被送入解码层还原点云,计算编码-解码前后点云的倒角损失[31].在后向阶段,提取到的点云特征通过逆合成算法估计旋转,将源点云旋转到目标点云之后再以同样的方式旋转回源点云,计算旋转前后的源点云的倒角损失.通过最小化前向和后向阶段的倒角损失,训练编码器关注点云旋转差异.

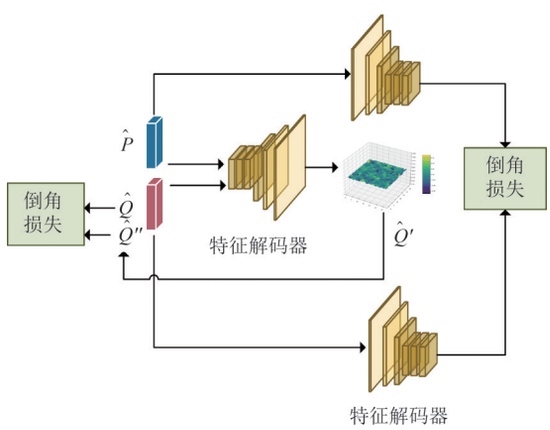


图2 前向后向训练结构图

Fig.2 Structure of forward-backward training

倒角损失是衡量两组点之间相似度的指标,定义如下:

$$\begin{aligned}\mathrm{Cf\ Loss}(p, q) = &\sum_{p\in\boldsymbol{A}}\sum_{i=1}^{N}\min_{q\in S^*}\left\|\varphi_{\theta_i}(p;\ x) - q\right\|_2^2 + \\ &\sum_{q\in S^*}\min_{1\leqslant i\leqslant N}\min_{q\in\boldsymbol{A}}\left\|\varphi_{\theta_i}(p;\ x) - q\right\|_2^2\end{aligned} \tag{17}$$

式中:$p\in\boldsymbol{A}$是在输出点云的单位正方形$[0, 1]^2$中采样的一组点;$\boldsymbol{S}^*$来自源点云;x是点云形状特征;φ_{θ_i}是N个可学习参数.本文联合前向阶段和后向阶段的倒角损失来训练特征编码器,如下:

$$\begin{aligned}L_{\mathrm{Cfl}} = &\frac{1}{\hat{N}}\mathrm{CfLoss}\left(\hat{P}, \hat{P}_{\mathrm{ed}}\right) + \frac{1}{\hat{N}}\mathrm{CfLoss}\left(\hat{Q}, \hat{Q}_{\mathrm{ed}}\right) + \\ &\frac{1}{\hat{N}}\mathrm{CfLoss}\left(\hat{Q}, \hat{Q}_{\mathrm{bc}}\right)\end{aligned} \tag{18}$$

式中:$\hat{P}_{\mathrm{ed}}$和$\hat{Q}_{\mathrm{ed}}$分别是编码-解码后的源点云和目标点云;$\hat{Q}_{\mathrm{bc}}$是目标点云经过旋转变换到源点云又旋转回目标点云之后的形态.

综上所述,SPFANet主要训练多显著性峰值检测器,整体损失函数表达式如下:

$$L = \alpha L_{\mathrm{f}} + \beta L_{\mathrm{peaky}} + \lambda L_{\mathrm{dissim}} \tag{19}$$

其中,α、β和λ是超参数,用于平衡各个损失.多显著性峰值检测器采用两阶段训练.在开始阶段,使用特征损失训练主干网络提取特征,然后在联合训练中使用整体损失L训练整体架构.整个训练过程中局部方差不参与训练,即权重梯度不参与反向传播.细配准模块由倒角损失提前预训练,不参与SPFANet整体训练过程,且其参数在整体训练过程中不会更新.

2　实验及结果分析

2.1　实验数据集

本文在公开同源数据集3DMatch和跨源数据集3DCSR上评估提出的SPFANet方法的有效性.

3DMatch数据集[14]由通过深度传感器获取的62个真实场景组成.其中46个场景用于训练,8个场景用于验证,8个场景用于测试.训练数据按照Predator[25]进行预处理,同时使用重叠度大于30%的3DMatch数据集和重叠度介于10%到30%之间的3DLoMatch数据集进行评估.SPFANet模型训练时采用了点位置抖动、小刚性扰动和点洗牌等数据增强方法.

真实世界跨源点云3DCSR数据集[32]由LiDAR、Kinect和摄像头传感器从众多室内场景中捕捉到的点云组成,包含两组场景,共202对点云.第一组包含19个场景,165对点云由Kinect和LiDAR传感器采集;第二组包含2个场景,37对点云由Kinect和摄像头传感器采集.跨源点云数据集混合了多种噪声、密度差异和部分重叠,给配准过程带来了巨大挑战.同时,使用SPFANet模型配准之前,利用迭代最接近点(ICP)算法进一步细化了3DCSR数据集的变换矩阵真值.

2.2　实验环境

实验在NVIDIA Quadro RTX 8000 GPU上使用Pytorch 3.8完成.

利用带有权重衰减优化器的AdamW[33]算法训练多显著性峰值检测器,初始学习率为10^{-4},且每20个epoch减半,batchsize设置为2,epoch为60.在联合训练中,损失超参数α、β和λ分别设置为1.0、1.0和0.5.细配准模块使用Adam优化器以无监督方式进行预训练,其中学习率为10^{-4},batchsize为16,epoch为200.

2.3　在3DMatch数据集上的点云配准实验

为了评估配准性能,SPFANet分别与基于对应关系的学习算法(如D3Feat[15]、Predator[25]、DHV[34])、直接配准方法(如DGR[35]、REGTR[27]、GeoTransformer[28])、特征度量学习方法(FMR[20]、PointNetLK[19])以及将学习与特征度量相结合的算法OMNet[36]进行比较.所有方法均在同一数据集上完成训练,使用配准召回率(registration recall, RR)作为量化点云配准效率的主要指标,RR为成功配准的点云对[即变换误差的均方根误差(RMSE)小于0.2 m的点云对]占点云对总数的比例.此外,本文还使用相对旋转误差(relative rotation error, RRE)和相对平移误差(relative translation error, RTE)评估配准质量.

表1显示了不同方法在3DMatch和3DLoMatch数据集上的实验结果.可以看到,D3Feat在3DLoMatch数据集上表现不佳,这表明其在重叠度较低的点云对中,难以凭借简单的特征找到位于重叠区域的关键点.由于包含不准确的对应关系,DHV没有取得很好的配准效果.FMR和PointNetLK在3DMatch和3DLoMatch数据集上均表现不佳,表明低重叠场景对之间全局特征的误差过大,无法准确估计变换.当应用于真实场景3DMatch时,OMNet的性能明显下降.

表1　不同方法在3DMatch和3DLoMatch数据集上的配准对比结果

Tab.1　Alignment comparison results of different methods on 3DMatch and 3DLoMatch datasets

配准算法	3DMatch			3DLoMach		
	RRE/(°)	RTE/m	RR/%	RRE/(°)	RTE/m	RR/%
D3Feat[15]	2.161	0.067	81.6	3.361	0.103	37.2
Predator[25]	2.062	0.068	90.6	3.159	0.096	62.5
DHV[34]	2.080	0.061	91.4	2.544	0.078	69.1
DGR[35]	2.103	0.067	85.3	3.954	0.113	48.7
REGTR[27]	1.567	0.049	92.0	2.827	0.077	64.8
GeoTransformer[28]	1.534	**0.045**	92.0	2.764	0.081	75.0
PointNetLK[19]	3.620	0.213	1.6	7.342	0.298	1.2
FMR[20]	4.330	0.130	14.3	7.832	0.185	7.3
OMNet[36]	4.010	0.112	36.4	6.785	0.192	8.6
SPFANet	**1.323**	0.051	**94.6**	**1.721**	**0.072**	**77.2**

从表1看到,SPFANet在平均配准召回率上超过了参与对比的所有方法,在3DMatch和3DLoMatch上分别比最佳结果GeoTransformer提高了2.6个百分点和2.2个百分点,在RRE和RTE方面也取得了更

低的误差数值,这表明即使在低重叠度的情况下,SPFANet在真实世界的配准任务中也非常有效.具体地,SPFANet显著检测到位于重叠区域的可辨别性的部位,如图3所示,其中黄色点云为源点云,蓝色为目标点云,进一步证实了SPFANet的有效性.

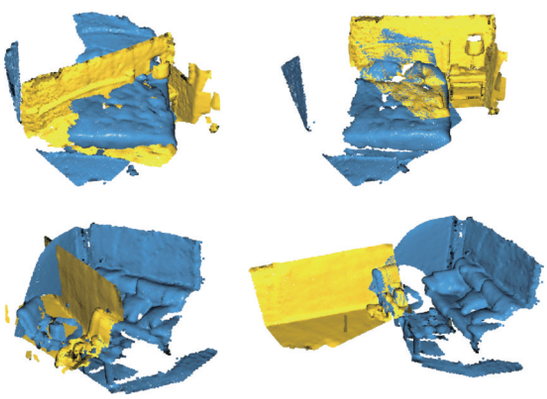


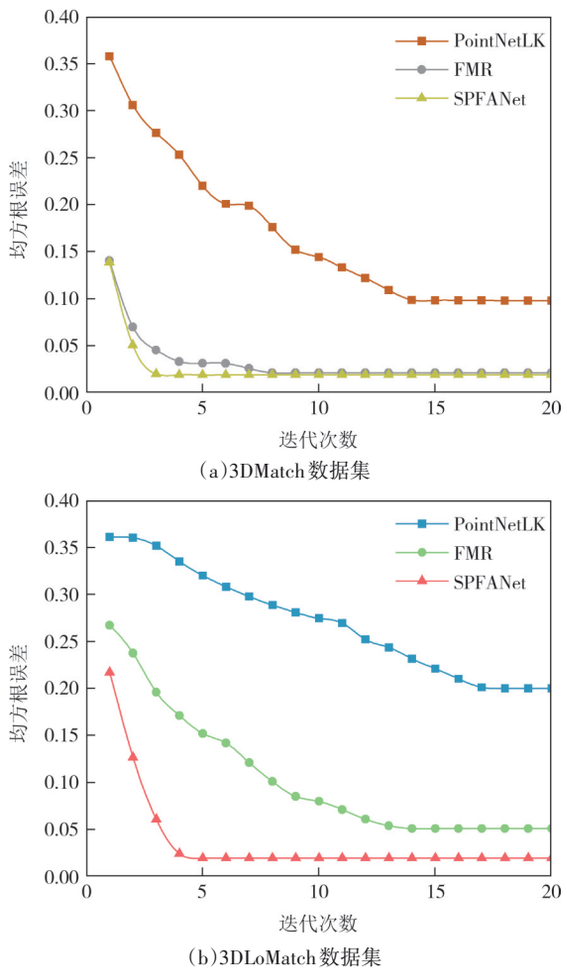
图3 3DMatch和3DLoMatch配准可视化样例

Fig.3 Sample visualization of alignment in 3DMatch and 3DLoMatch

图4(a)和图4(b)分别为SPFANet与FMR算法、PointNetLK算法在3DMatch和3DLoMatch数据集上的配准性能对比实验结果.图4(a)显示,当迭代次数为15时,PointNetLK算法性能达到最优,变换矩阵的均方根误差(RMSE)为0.098;当迭代次数为8时,FMR算法性能达到最优,RMSE为0.021.当迭代次数为5时,SPFANet达到最优,RMSE为0.018.3DLoMatch数据集上PointNetLK基本上不收敛,FMR收敛缓慢,且最优性能的配准效果不理想,如图4(b)所示.可以看到,SPFANet达到了最优的收敛速度与收敛性能.由于FMR和PointNetLK中没有去除非重叠区域点云的影响,从而无法学习到鲁棒的点云特征,导致无法配准低重叠点云.SPFANet考虑到点云非重叠区域的影响,采样显著重叠点,消除冗余信息,并使用独特的前向后向变换方案训练细配准特征编码网络,提取更丰富的特征信息,获得更优越的性能.

2.4 在3DCSR数据集上的点云配准实验

在3DCSR数据集上进行真实跨源点云配准的泛化实验,结果如表2所示.可以看到,GCTR[5]和FGR[13]在找到准确对应关系方面效果不佳.PointNetLK的特征误差优化受到尺度差异、部分重叠以及噪声干扰的严重影响.FMR的配准召回率较低,表明其无法处理部分重叠.最后,尽管DGR实现了次优性能,但由于其无法明确克服离群值的干扰,其整体性能仍然低于SPFANet.很明显,SPFANet的配准



(a)3DMatch数据集

(b)3DLoMatch数据集

图4 三种点云配准方法在3DMatch和3DLoMatch数据集上配准性能对比

Fig.4 Comparison of the alignment performance of the 3 point cloud registration methods on 3DMatch and 3DLoMatch datasets

召回率最高,这说明本文方法足以应对真实世界中跨源点云的挑战,具有良好的泛化能力.

图5(a)为SPFANet和PointNetLK以及FMR算法在3DCSR数据集上的配准性能对比.可以看到,当迭

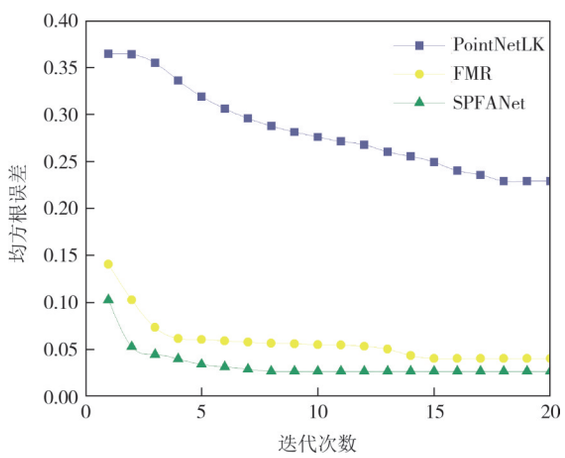
表2 各方法在3DCSR数据集上的配准对比结果

Tab. 2 Alignment comparison results of different methods on 3DCSR datasets

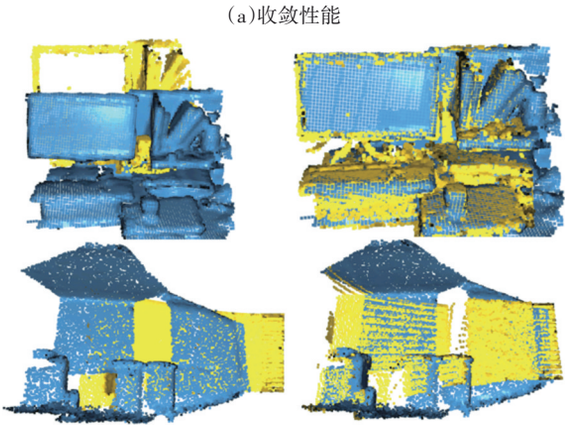
配准算法	RRE/(°)	RTE/m	RR/%
GCTR[5]	7.46	0.17	0.5
RANSAC[37]	8.30	0.13	3.5
JRMPC[38]	8.57	0.71	1.0
ICP[39]	5.71	0.38	24.3
FGR[13]	10.74	0.07	1.5
PointNetLK[19]	12.54	0.09	0.1
FMR[20]	3.22	0.09	24.5
DGR[35]	2.46	**0.04**	59.6
SPFANet	**1.43**	0.08	**68.7**

代次数为22时,PointNetLK算法性能达到最优,迭代次数为19时,FMR算法达到性能最优,而迭代次数为15时,SPFANet达到最优,并且最优性能RMSE在三者中最低.

图5(b)可视化了SPFANet配准结果,其中黄色点云为源点云,蓝色为目标点云.可以清楚看到,点云中存在大量的噪声、缺失值以及部分重叠,但SPFANet对于具有重大挑战的跨源点云能够精确配准.



(a)收敛性能



(b)SPFANet配准结果

图5　三种方法在3DCSR数据集上收敛性能对比及SPFANet配准结果可视化

Fig.5　Comparison of the convergence performance of the three methods on the 3DCSR dataset and visualization of the alignment results of SPFANet

2.5　消融实验

本节将对所介绍方法中每个模块的有效性进行审查,分别对显著性峰值检测器、峰值损失重加权方案、细配准前向后向无监督训练方案进行消融实验,在三个数据集上的实验结果如表3所示.可以看到,在粗配准模块中增加显著性峰值检测器的方法(如表3中2-Dete系列),其配准召回率显然高于没有显著性峰值检测器的方法(如表3中的No-Dete),验证了显著性峰值检测器的有效性,其确实检测到有效的关键点,成功去除了干扰信息,提高了配准成功率.此外,为了验证峰值损失重加权方案的有效性,对比了不使用重加权方案的峰值损失训练的模型(如表3中的2-Dete No-W),其性能明显比采用重加权方案差.为了验证细配准前向后向无监督训练方案的有效性,分别与细配准中使用RANSAC和FMR的方法进行比较,结果表明,本文采用的2-Dete+细配准方案获得最佳性能,比2-Dete+RANSAC和2-Dete+FMR两种方案配准召回率分别高4%和3%,证明本文细配准方案的有效性.

表3　SPFANet在3个数据集上的配准召回率消融实验

Tab.3　Registered recall ablation experiment of SPFANet on 3 datasets

配准算法	3DCSR	3DLoMatch	3DMatch
No- Dete	50.9	51.4	70.7
2-Dete No-W	55.6	58.3	79.6
2-Dete+RANSAC	58.4	60.8	82.2
2-Dete+FMR	62.1	68.6	90.3
2-Dete+细配准	68.7	77.2	94.6

为了验证显著性峰值检测器中超参数检测器个数B的有效性,在三个数据集上分别设置超参数检测器个数B=1,2,4,测试结果如表4所示.可以看到,使用2个检测器的模型提供了最佳的性能,在持续增加B后,配准召回率反而逐渐下降,故SPFANet选取B=2效果最优.

表4　超参数检测器个数B的消融实验结果

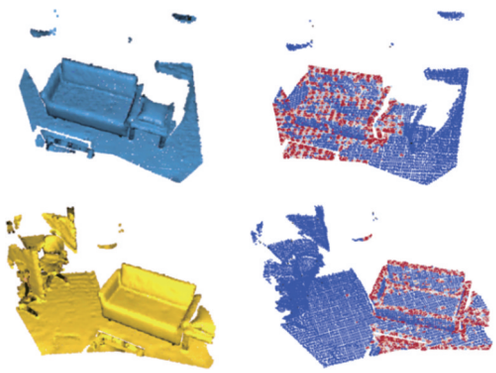
Tab.4　Ablation experiment results with hyperparametric detector number B

配准算法	3DCSR	3DLoMatch	3DMatch
1-Dete	65.3	72.2	90.1
2-Dete	68.7	77.2	94.6
4-Dete	60.3	70.7	88.6

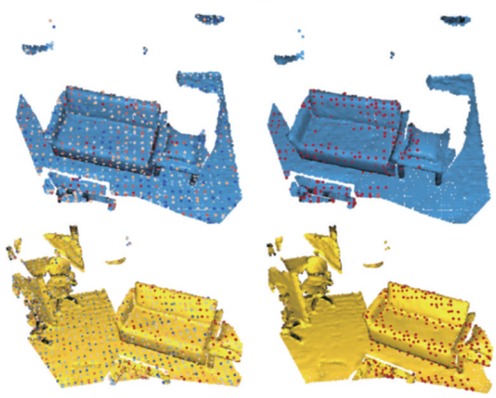
2.6　可视化分析

在低重叠点云数据集3DLoMatch上,损失重加权效果示例如图6所示.图6(a)是源点云和目标点云及其计算的重加权热力图,可以看到,重加权方案使网络更加关注位于重叠区域的显著部位.图6(b)分别是SPFANet使用和不使用重加权方案训练获得的检测热图.可以看到,所使用的重加权方案平滑了点云中非重叠区域的峰值,从而避免了该区域的检测,并且聚集在更具辨别力方向的区域.重加权方案使得网络更加关注重叠区域显著部位,从而鼓励网络学习

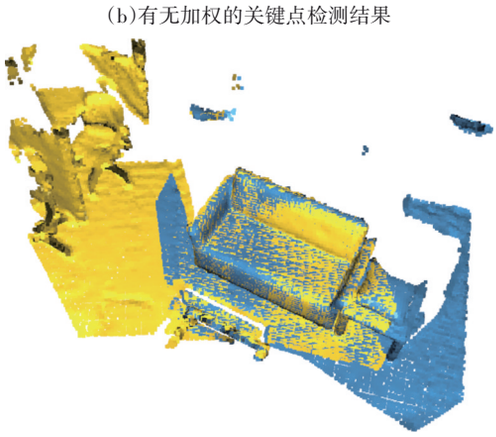
显著部位更细节特征.以图6(c)可以看出两个点云对齐良好.



(a)原始点云与重加权热力图



(b)有无加权的关键点检测结果



(c)配准结果

图6 损失重加权效果示例分析

Fig.6 Example analysis of the effect of loss weight weighting

3 结束语

本文提出了一种结合显著性峰值检测和特征对齐的从粗到细的点云配准网络SPFANet.在网络中,设计了多显著性峰值检测器,结合Transformer自注意力与交叉注意力机制提取多个不相交的关键点集及其特征;同时,设计了一种基于方差和重叠分数的重加权峰值损失抑制方案,避免了非重叠区域与非歧视区域的关键点的检测.在粗配准模块,选择一种基于全局特征相关性的匹配方案,有效抑制了歧义匹配.在细配准模块,提出了一个无监督训练方案,采用前向后向变换方法训练特征度量框架,通过最小化特征度量投影误差,高效地估计点云之间的变换参数,有效提升了真实点云配准的准确率与效率.广泛的实验评估表明,SPFANet能够有效克服各种真实场景中大噪声、密度差异、部分重叠和尺度变化等问题,在同源和跨源数据集上均取得了优秀的结果.但SPFANet并不能解决所有的问题,实验观察到,在超大旋转角度的真实极端场景下,配准失败率会比较高.因此,未来考虑关注点云的旋转属性、引入旋转归一化,以应对高度旋转的点云配准挑战.

参考文献

[1] BESL P J, MCKAY N D. A method for registration of 3-D shapes [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE, 1992: 239-256.

[2] YANG J, LI H, CAMPBELL D, et al. Go-ICP : a globally optimal solution to 3D ICP point-set registration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(11): 2241-2254.

[3] JIAN B, VEMURI B C. Robust point set registration using Gaussian mixture models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1633-1645.

[4] HUANG X S, ZHANG J, FAN L X, et al. A systematic approach for cross-source point cloud registration by preserving macro and micro structures [J]. IEEE Transactions on Image Processing, 2017, 26(7): 3261-3276.

[5] HUANG X S, FAN L X, WU Q, et al. Fast registration for cross-source point clouds by using weak regional affinity and pixel-wise refinement [C]//2019 IEEE International Conference on Multimedia and Expo (ICME). Shanghai, China. IEEE, 2019: 1552-1557.

[6] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6): 381-395.

[7] MELLADO N, AIGER D, MITRA N J, et al. Super 4PCS fast global pointcloud registration via smart indexing [C]//Proceedings of the Symposium on Geometry Processing. ACM, 2014: 205-215.

[8] STEDER B, RUSU R B, KONOLIGE K, et al. NARF: 3D range image features for object recognition [C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS Workshop. 2010: 44.

[9] MIAN A, BENNAMOUN M, OWENS R. On the repeatability and quality of keypoints for local feature-based 3D object retrieval from cluttered scenes [J]. International Journal of Computer Vision, 2010, 89(2): 348-361.

[10] ZHONG Y. Intrinsic shape signatures: a shape descriptor for 3D object recognition [C]//2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops. Kyoto, Japan. IEEE, 2009: 689-696.

[11] RUSU R B, BLODOW N, MARTON Z C, et al. Aligning point cloud views using persistent feature histograms[C]//2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. Nice, France. IEEE, 2008: 3384-3391.

[12] RUSU R B, BLODOW N, BEETZ M, et al. Fast point feature histograms (FPFH) for 3D registration[C]//Proceedings of the 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan. ACM, 2009: 1848-1853.

[13] ZHOU Q Y, PARK J, KOLTON V. Fast global registration[C]// Computer Vision-ECCV 2016: 14th European Conference. Amsterdam, The Netherlands. 2016: 766-782.

[14] ZENG A, SONG S R, NIEßNER M, et al. 3DMatch: learning local geometric descriptors from RGB-D reconstructions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 199-208.

[15] BAI X Y, LUO Z X, ZHOU L, et al. D3Feat: joint learning of dense detection and description of 3D local features [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020: 6358-6366.

[16] 陆军, 陈坤, 范哲君. 点邻域尺度差异描述的点云配准算法[J]. 国防科技大学学报, 2021, 43(3): 128-134.
LU J, CHEN K, FAN Z J. Point cloud registration algorithm based on scale difference descriptor of point neighborhood[J]. Journal of National University of Defense Technology, 2021, 43(3): 128-134. (in Chinese)

[17] 李雪梅, 王春阳, 刘雪莲, 等. 基于超体素双向最近邻距离比的点云配准方法[J]. 吉林大学学报(工学版), 2022, 52(8): 1918-1925.
LI X M, WANG C Y, LIU X L, et al. Point cloud registration method based on supervoxel bidirectional nearest neighbor distance ratio [J]. Journal of Jilin University (Engineering and Technology Edition), 2022, 52(8): 1918-1925. (in Chinese)

[18] 熊丰伟, 庄健, 沈人. 基于邻域曲率的低特征辨识度点云配准方法[J]. 计算机应用研究, 2022, 39(1): 285-290.
XIONG F W, ZHUANG J, SHEN R. Point cloud with low feature recognition registration method based on neighborhood curvature[J]. Application Research of Computers, 2022, 39(1): 285-290. (in Chinese)

[19] AOKI Y, GOFORTH H, SRIVATSAN R A, et al. PointNetLK: robust & efficient point cloud registration using PointNet [C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. IEEE, 2019: 7163-7172.

[20] HUANG X S, MEI G F, ZHANG J. Feature-metric registration: a fast semi-supervised approach for robust point cloud registration without correspondences [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020: 11363-11371.

[21] LI X Q, PONTES J K, LUCEY S. PointNetLK revisited[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA. IEEE, 2021.

[22] 张元, 李晓燕, 韩燮. 一种低重叠率的三维点云配准方法[J]. 激光与光电子学进展, 2021, 58(8): 170-179.
ZHANG Y, LI X Y, HAN X. Three-dimensional point cloud registration method with low overlap rate [J]. Laser & Optoelectronics Progress, 2021, 58(8): 170-179. (in Chinese)

[23] 阎翔鑫, 蒋峥, 刘斌. 基于角度约束的跨源点云配准算法[J]. 激光与光电子学进展, 2023, 60(22): 272-279.
YAN X X, JIANG Z, LIU B. Cross-source point cloud registration algorithm based on angle constraint [J]. Laser & Optoelectronics Progress, 2023, 60(22): 272-279. (in Chinese)

[24] 易见兵, 彭鑫, 曹锋等. 多尺度特征融合的点云配准算法研究[J/OL]. 广西师范大学学报(自然科学版), 2024: 1-12. DOI: 10.16088/j.issn.1001-6600.2023082502.

[25] HUANG S Y, GOJCIC Z, USVYATSOV M, et al. Predator: registration of 3D point clouds with low overlap[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA. IEEE, 2021: 4265-4274.

[26] LI J H, ZHANG C H, XU Z Y, et al. Iterative distance-aware similarity matrix convolution with mutual-supervised point elimination for efficient point cloud registration [C]// Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 378-394.

[27] YEW Z J, LEE G H. REGTR: end-to-end point cloud correspondences with transformers [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA. IEEE, 2022: 6667-6676.

[28] QIN Z, YU H, WANG C, et al. Geometric transformer for fast and robust point cloud registration [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA. IEEE, 2022: 11133-11142.

[29] CHENG X, YU H, ZHANG H, et al. WeightedPose: generalizable cross-pose estimation via weighted SVD[J]. arXiv, 2024, arXiv: 2405.02241.

[30] BAKER S, MATTHEWS I. Lucas-kanade 20 years on: a unifying framework[J]. International Journal of Computer Vision, 2004, 56(3): 221-255.

[31] GROUEIX T, FISHER M, KIM V G, et al. A papier-Mache approach to learning 3D surface generation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 216-224.

[32] HUANG X S, MEI G F, ZHANG J, et al. A comprehensive survey on point cloud registration [EB/OL]. [2024-02-18] https://arxiv.org/abs/2103.02690v2.

[33] LOSHCHILOV I, HUTTER F. Decoupled weight decay regularization [C]//International Conference on Learning Representations, 2019.

[34] LEE J, KIM S, CHO M, et al. Deep Hough voting for robust global registration [C]// Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. IEEE, 2021: 15974-15983.

[35] CHOY C, DONG W, KOLTUN V. Deep global registration[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 2514-2523.

[36] XU H, LIU S C, WANG G F, et al. OMNet: learning overlapping mask for partial-to-partial point cloud registration [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2021: 3112-3121.

[37] MELLADO N, DELLEPIANE M, SCOPIGNO R. Relative scale estimation and 3D registration of multi-modal geometry using growing least squares [J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(9): 2160-2173.

[38] HUANG X S, ZHANG J, WU Q, et al. A coarse-to-fine algorithm for registration in 3D street-view cross-source point clouds [C]//2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA). Gold Coast, QLD, Australia. IEEE, 2016: 1-6.

[39] PENG F R, WU Q, FAN L X, et al. Street view cross-sourced point cloud matching and registration [C]//2014 IEEE International Conference on Image Processing (ICIP). Paris, France. IEEE, 2014: 2026-2030.