基于多尺度特征融合的由粗到精点云形状补全

张德军[1)], 王杨[1)], 谭雪峰[1)], 吴亦奇[1)*], 陈壹林[2)], 何发智[3)]

[1)] (中国地质大学(武汉)计算机学院 武汉 430078)
[2)] (智能机器人湖北省重点实验室(武汉工程大学) 武汉 430205)
[3)] (武汉大学计算机学院 武汉 430072)
(wuyq@cug.edu.cn)

摘 要: 为了以由粗到精的方式实现点云形状补全, 提出一个端到端的两阶段多尺度特征融合网络, 其中的每个阶段都是由一个编码器-解码器构成. 第 1 阶段中, 首先利用点集抽取模块提取残缺点云的全局特征, 在获取不同分辨率点特征的同时能关注更多的局部邻域特征, 然后使用多层感知机作为解码器生成粗糙的点云骨架; 第2阶段中, 利用点云骨架和残缺点云提取多尺度局部特征, 并通过注意力机制与第 1 阶段中的多尺度全局特征相互融合, 使得每个点都包含全局和局部几何信息; 最后将第 2 阶段中的全局特征和多尺度局部特征逐步进行上采样, 并通过多层感知机生成精细的完整点云. 采用倒角距离作为评价标准, 在 ShapeNet, MVP 和 Completion3D 数据集上进行点云补全实验的结果表明, 误差分别比基准网络降低 17.1%, 3.9%和 13.9%, 验证了所提网络的有效性.

关键词: 点云补全; 多尺度特征融合; 由粗到精; 编码器-解码器
中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2024.19871

Coarse-to-Fine Point Cloud Shape Complementation Based on Multi-Scale Feature Fusion

Zhang Dejun[1)], Wang Yang[1)], Tan Xuefeng[1)], Wu Yiqi[1)*], Chen Yilin[2)], and He Fazhi[3)]

[1)] (*School of Computer Science, China University of Geosciences, Wuhan* 430078)
[2)] (*Hubei Key Laboratory of Intelligent Robot (Wuhan Institute of Technology), Wuhan* 430205)
[3)] (*School of Computer, Wuhan University, Wuhan* 430072)

Abstract: To implement the point cloud shape completion in a coarse-to-fine manner, an end-to-end two-stage multi-scale feature fusion network is proposed, in which each stage consists of an encoder-decoder. In the first stage, the set abstraction module extracts the global features of the incomplete point cloud, which can focus on more local neighborhood features while acquiring point features of different resolutions. A decoder built from multilayer perceptrons generates a coarse skeleton. In the second stage, the coarse skeleton and incomplete point cloud are used to learn multi-scale local features. The multi-scale local features are fused with the multi-scale global features of the first stage via an attention mechanism so that each point contains global and local geometric information. Finally, the global features and multi-scale local features are progressively upsampled, and a fine-grained complete point cloud is generated via multilayer perceptrons. The point cloud completion experiments on ShapeNet, MVP, and Completion3D datasets show that the chamfer distance is reduced by 17.1%,

收稿日期: 2022-07-26; 修回日期: 2022-11-01. **基金项目**: 国家自然科学基金(61802355); 智能机器人湖北省重点实验室开放基金(HBIR 202105). **张德军**(1982—), 男, 博士, 副教授, 硕士生导师, CCF 高级会员, 主要研究方向为数据驱动的几何处理与建模、智能点云处理、形状修复; **王杨**(1997—), 男, 硕士研究生, 主要研究方向为智能点云处理; **谭雪峰**(1996—), 男, 硕士研究生, 主要研究方向为智能点云处理; **吴亦奇**(1985—), 男, 博士, 副教授, 硕士生导师, CCF 会员, 论文通信作者, 主要研究方向为人工智能、计算机图形学; **陈壹林**(1989—), 男, 博士, 讲师, 硕士生导师, CCF 会员, 主要研究方向为智能优化、智能点云处理; **何发智**(1968—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为计算机图形学、人工智能、智能计算.

3.9%, and 13.9% compared with the baseline, respectively, demonstrating the effectiveness of the proposed method.

Key words: point cloud completion; multi-scale feature fusion; coarse-to-fine; encoder-decoder

点云数据作为一种最原始的三维数据表征, 由于其具有数据获取容易、能够有效地表示复杂三维形状、无需维护拓扑连接关系等优点, 因此得到广泛应用. 随着三维数据采集设备的日益普及, 在数字几何处理、工业制造、逆向工程等领域中涌现出大量的点云数据. 在采集点云数据的过程中, 受观察角度、视线遮挡和测量不稳定等因素的影响, 实际得到的点云数据存在形状残缺的问题. 因此点云补全一直是点云数据处理中的热门问题.

近些年, 深度学习广泛应用于点云数据处理中, 如点云分类[1-2]和点云分割[3]. 本文重点关注点云补全[4], 以给定的残缺点云生成完整点云. 实际上, 点云补全拥有广泛的应用, 如增强现实[5]和场景理解[6].

点云补全可以分解为2个阶段任务. 第1阶段, 以给定的残缺点云恢复出缺失区域; 第 2 阶段, 既要保留残缺点云的几何形状, 又要对补全后的形状做进一步优化. 为了完成这 2 个阶段的任务, 目前的工作首先从残缺点云中提取全局信息, 然后利用全局信息估计缺失的几何信息[7]. 然而, 编码后的全局形状表示通常会丢失点云的局部几何信息, 而全局形状表示和局部几何信息对点云补全具有较强的互补性.

Peng 等[8]提出 Conformer 网络, 以交互方式融合不同分辨率下的局部特征和全局表示, 在目标检测和实例分割中取得了较好的效果. Guo 等[9]提出一种用于点云学习的框架(point cloud Transformer, PCT), 首先利用 PointNet++[2]从分组的邻近点中提取局部特征, 然后将局部特征输入具有偏移注意力机制的 Transformer 模块中学习全局特征, 并在形状分类、零件分割中取得了较好的性能; 但是, 该网络采用交替学习局部特征和全局特征的方式有待改进. 全局特征和局部特征具有较强的互补关系, 它们的学习过程不应该相互独立, 而是应该密切融合并相互学习.

为了在点云补全中探索全局特征与局部特征之间的互补关系, 以及它们之间的融合关联, 本文提出一种用于点云补全的多尺度特征融合网络(multi-scale feature fusion network, MFF-Net). 该网络由 2 个编码器-解码器子网络构成, 以由粗到精的方式完成点云补全. 第 1 阶段, 首先对输入的残缺点云提取全局特征, 然后使用全局特征生成粗糙的点云骨架; 第 2 阶段, 通过多级融合全局与局部特征逐步精练点云的结构细节, 最终获取的完整点云.

1 相关工作

1.1 三维形状补全

传统的三维形状补全方法首先采用体素或者网格描述三维物体[10], 然后使用三维卷积提取特征, 在形状补全[11]和三维重建[12]中取得了巨大的成功. 但是, 这些方法存在着相同的不足, 即基于体素的方法会消耗大量内存并且无法生成高分辨率的三维形状.

为了提高点云分辨率, 文献[13]利用八叉树结构逐渐体素化特定区域. 由于体素化操作的量化效应, 因此近些年的研究更多专注于网格化的三维重建. 因为现有的网格表示基于模板网格变形的目标网格, 只对特定类型的三维数据有效, 所以研究人员将目光转向使用非结构化的点云数据作为三维对象的表示, 其独特的数据表示形式使得计算内存消耗更少, 但数据量大小相同的情况下, 三维形状的细粒度表示能力更强.

然而, 三维卷积并不适用于处理点云数据, 因为规则的卷积算子不适合计算无序的数据类型. PointNet[1]和 PointNet++[2]是直接处理三维点云的开创性工作, 为点云数据处理提供了新的思路, 并且激发了许多下游任务的研究.

1.2 图网络

将点云中每个点当成图的一个顶点, 基于图的网络找到每个顶点相邻的顶点并生成有向边; 然后对空间中的邻居点进行卷积运算, 使用池化操作聚合每个点的相邻信息, 从而得到图特征信息. PointNet++[2]利用球查询的方法对相邻点进行采样, 获取多尺度局部特征. 文献[14-15]均是采用图卷积核学习局部几何特征, 其通过从局部邻域图中学习边缘特征, 能够有效地提高算法对点云

的学习能力.

1.3　点云补全

点云数据极易获取, 并且具有存储便利、处理数据的计算成本比体素数据更低等优势. 近年来, 面向点云数据的形状补全受到广泛关注. 点云补全的目标是根据局部观测的点云生成完整的三维点云形状. PCN[4], TopNet[16], SA-Net[17]和 ECG[18]都是采用编码器-解码器的框架实现点云补全, 这些方法大都采用 PointNet[1]或 PointNet++[2]的编码器从输入的残缺点云中提取全局特征, 并根据这些特征使用解码器推断出完整的点云形状. PCN[4]从残缺点云中学习全局特征用于生成粗糙的完整点云, 并使用 FlodingNet[7]对其进行上采样, 以获得精细的完整点云. TopNet[16]采用一种树形结构的解码器预测完整的点云形状. SA-Net[17]采用多级 Folding 结构和 skip-attention 修复残缺的点云形状. ECG[18]先利用全局特征生成粗糙的点云形状, 再利用边缘局部特征生成精细的点云形状.

为了保留和恢复局部几何细节, 文献[19]利用局部特征精练粗糙的完整点云. NSFA[20]结合已知特征和预测的缺失特征补全完整的三维点云形状, 该方法假设已知部分和缺失部分的比例约为 1∶1(可观测部分约为整个对象的一半), 这种先验条件在大多数情况下是非常苛刻的. VRCNet[21]专注于对抗学习和变分自动编码器的框架, 以提高生成的完整形状的真实性和一致性. Xie 等[22]提出一种基于样式和对抗性可微分渲染的点云生成器, 用于点云补全. Yu 等[23]将点云补全转换为集合到集合的转换问题, 并提出 Transformer 编码器-解码器体系结构, 能够更好地学习结构知识, 保留详细信息以完成点云补全. 本文提出的 MFF-Net 由 2 个编码器-解码器子网络组成, 以由粗到精的方式逐步补全三维形状. 首先提取全局特征, 然后使用全局特征生成粗糙的点云, 最后结合局部特征和全局特征逐渐精练点云的结构细节.

2　MFF-Net

2.1　概　述

本文将点云 Y 和 X 分别定义为一个完整三维点云物体及其局部观测的残缺点云; X 不是 Y 的子集, 利用一个大规模的合成数据集可以很容易地获取 X 和 Y 的样本. 点云补全是在给定点云 X 的条件下, 预测一个完整的三维点云形状 Y'.

如图 1 所示, 本文使用 2 个阶段的子任务实现点云补全, 即生成点云骨架和精练几何细节. MMF-Net 由 2 个编码器-解码器的子网络构成, 并以由粗到精的方式逐步补全点云形状 Y'.

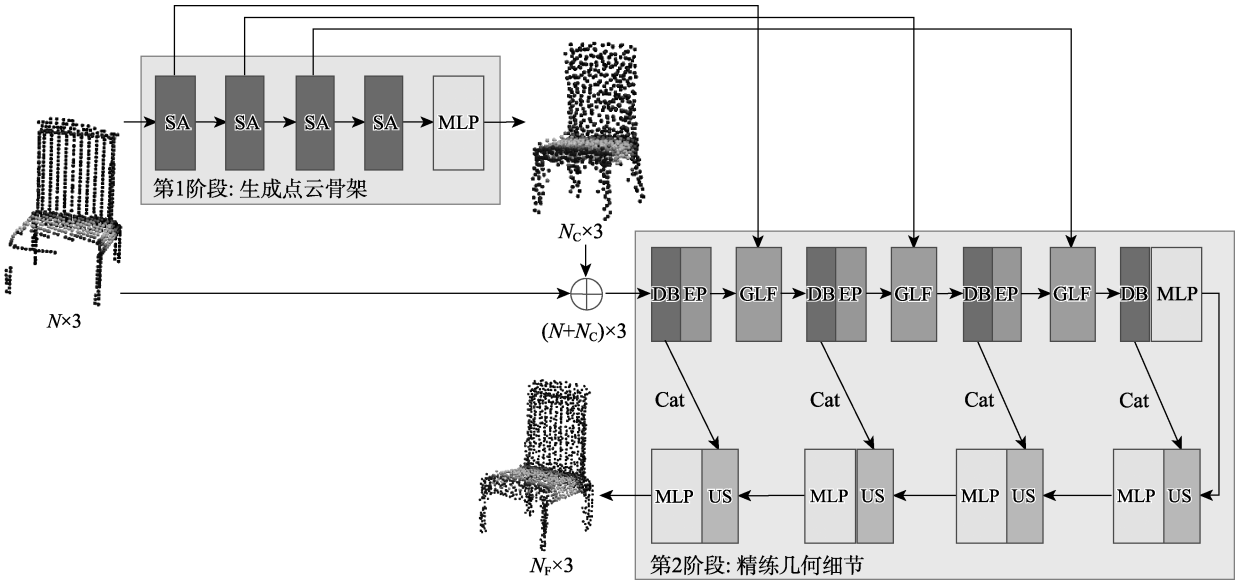


图 1　MFF-Net 结构

第 1 阶段: 生成点云骨架. 首先采用 PointNet++的多级点集抽取模块(set abstraction, SA)提取全局结构信息 F_X^{G}; 然后通过多层感知机(multilayer perceptrons, MLP)补全出一个粗糙的点云 Y_{C} 作为完整点云的骨架.

第 2 阶段: 精练几何细节. 首先利用粗糙点云 Y_{C} 和残缺点云 X 学习多尺度局部几何细节 F_X^{P}. ECG[18]在融合 F_X^{G} 和 F_X^{P} 的过程中, 并未显式地考

虑这 2 种互补性较强的信息之间的关联. 为了既密切融合局部几何特征 F_X^{P} 和全局特征 F_X^{G}, 又保留各自的信息, 本文设计了具有软注意力机制的全局和局部特征融合(global and local feature, GLF)模块, 并将融合之后的特征逐步进行上采样, 最后通过 MLP 生成精细的完整点云 Y'.

2.2 第 1 阶段: 生成点云骨架

文献[4]以由粗到精的方式实现点云补全, 先将学习到的全局特征解码, 生成一个粗糙的补全结果作为整体形状骨架, 再逐步恢复局部几何细节. 在精细化局部几何细节之前需要预测形状骨架, 原因如下: (1) 形状骨架描述的是完整的点云结构, 即恢复局部观测中缺失的区域; (2) 形状骨架可视为自适应三维锚点, 用于挖掘残缺点云中的局部几何特征.

第 1 阶段中, 编码器-解码器的目标是生成一个完整的点云结构, 从残缺点云中恢复完整的三维形状, 同时保留一定的局部几何细节. 本文使用 PointNet++中的 SA, 以多级的分层结构逐步抽象出越来越大的局部区域, 在获取不同分辨率点特征时能够关注更多的局部邻域特征.

本文构建的分层次结构的编码器由一系列 SA 组成, 包含采样层、分组层和 PointNet 层. 采样层对输入点云以最远点采样选择一组点, 作为各自局部邻域的质心; 分组层通过寻找质心周围的邻近点构造局部邻域集; PointNet 层使用 MLP 将局部邻域点(包含质心)编码为特征向量. 最后, 编码器将输入的点云转化成一个 1 024 维的全局特征, 再经过一个由 MLP 构成的解码器生成粗糙的完整点云.

2.3 第 2 阶段: 精练几何细节

第 2 阶段中, 编码器-解码器的目标是基于观测的残缺点云和补全的粗糙点云形状, 进一步生成高质量的完整点云形状. 然而, 在不完整、非均匀分布且稀疏孤立的点云中, 需要有效地提取边缘几何特征, 本文利用粗糙点云的全局形状结构和残缺点云 X 中的局部几何细节来提取边缘几何特征. 点云形状骨架可以被认为是自适应的三维锚点, 有利于从残缺点云 X 中捕获有用的局部边缘特征.

图 1 中, 第 2 阶段的输入包含补全的粗糙点云 Y_{C} 和原始的残缺点云 X, 将它们级联后既具有完整点云的结构骨架, 又保留了输入点云的局部几何细节. 使用具有图网络作用的密集模块(dense block, DB)聚合每个点周围 K 个邻近点的信息, 并与原始点特征级联形成新的局部特征; 再通过边缘保留池化操作(edge-preserved pooling, EP)模块将获取的局部特征下采样. 以上过程把每个点邻域的 K 个点特征聚合, 产生更低分辨率的点特征, 同时保留表征能力更强的局部特征. 然后, 使用 GLF 将聚合生成的局部特征与第 1 个编码器-解码器生成的相同分辨率的全局点特征进行特征融合, 融合后的特征作为下一个 DB 的输入; 重复以上过程, 以获取多分辨率的点特征. 最后, 分别将第 2 个编码器-解码器生成的多分辨率特征与全局特征逐步进行上采样(up-sampling, US)得到的特征进行级联, 并通过 MLP 生成精细的点云.

2.3.1 DB

文献[2]用 2 种策略(球查询或 K-最近邻) 搜索邻近点来提取局部边缘特征. PointNet++使用的球查询算法是在多个具有预定义半径的指定搜索球中选择前 K 个点, 但不能保证在预定义的半径内总是能够选择最接近的 K 个点[15,24]. 受到 EdgeConv[14]的启发, DB 使用 K-最近邻算法, 在特征空间中对每个点选择 K 个最邻近的点构建邻域图, 以此提取边缘特征. 编码的边缘特征 f_p' 可以表示为

$$f_p' = g[H_\Theta(f_p \oplus (f_p - f_q^1)), \cdots, H_\Theta(f_p \oplus (f_p - f_q^K))] \tag{1}$$

其中, f_p 表示质心点 p 的输入特征; f_q^K 表示第 K 近邻的特征; $\oplus$ 表示特征级联; $H_\Theta(\cdot)$ 表示共享权重的 MLP; $g[\cdot]$ 表示最大池化操作.

2.3.2 EP

在输入点密度固定的情况下, 卷积滤波器的感受野依赖构造邻域图时 KNN 的大小. 大的邻域图虽然可以扩大卷积滤波器的感受野, 但同时也增加了网络的复杂度. 为了更加高效地学习点的多尺度边缘特征, 本文采用 EP[15], 对输入点再次采样时, 同时传播所选质心点及其局部邻域点的特征到后续网络层进行处理. 因此, 随着点集变得更稀疏, 其感受野会变得更大. 与 PointNet++[2]仅通过最大池化操作传播局部邻域特征不同, EP 操作能够将重复和关联较少的特征传播到下一层, 定义为

$$f_p' = f_p \oplus g\left[f_q^1, \cdots, f_q^K\right] \tag{2}$$

相应地, 通过级联低级特征和来自前面层次结构的插值特征逐步恢复解码器中的空间信息, 即

$$f_p' = f_p^{\mathrm{E}} \oplus w\left[f_q^1, \cdots, f_q^K\right] \tag{3}$$

其中, f_p^{E} 表示通过跳跃连接直接从编码器传播的对应特征; $w[\cdot]$ 表示反距离加权平均操作; f_q^K 表

示点 p 在其前面层次中的第 K 个最近邻点特征.

2.3.3　GLF

目前, 大多数点云补全方法都未有效地利用全局与局部特征之间的互补性[11]. 受 SKNet[25]的启发, 本文提出 GLF, 其结构如图 2 所示.

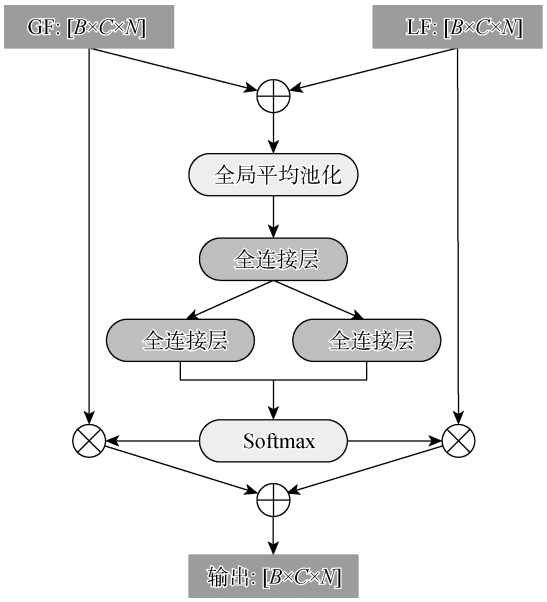


图 2　GLF 结构

首先通过逐元素求和的方式融合全局特征(global feature, GF)和局部特征(local feature, LF)

$$\boldsymbol{U}=\mathrm{GF}+\mathrm{LF} \tag{4}$$

然后使用全局平均池化得到嵌入信息

$$\boldsymbol{S}=F_{\mathrm{GP}}(\boldsymbol{U}) \tag{5}$$

另外, 生成了一个更加紧凑的特征

$$\boldsymbol{Z}=F_{\mathrm{FC}}(\boldsymbol{S}) \tag{6}$$

用来实现自适应的软注意力机制. $\boldsymbol{Z}$ 是通过一个简单的全连接层得到的.

在紧凑特征 $\boldsymbol{Z}$ 的指导下, 跨通道的软注意力向量被用于自适应地选择全局和局部特征中不同重要程度的信息. GF, LF 的软注意力向量 $\boldsymbol{a}_c$, $\boldsymbol{b}_c$ 表示为

$$\begin{cases}\boldsymbol{a}_c=\dfrac{\mathrm{e}^{A_c z}}{\mathrm{e}^{A_c z}+\mathrm{e}^{B_c z}}\\ \boldsymbol{b}_c=\dfrac{\mathrm{e}^{B_c z}}{\mathrm{e}^{A_c z}+\mathrm{e}^{B_c z}}\end{cases} \tag{7}$$

其中, c 表示通道的序号; $A\in\mathbb{R}^{C\times d}$; $B\in\mathbb{R}^{C\times d}$.

通过全局特征和局部特征各自的注意力加权求和, 获得最终的特征

$$\boldsymbol{V}_c=\boldsymbol{a}_c\cdot\mathrm{GF}+\boldsymbol{b}_c\cdot\mathrm{LF} \tag{8}$$

其中, $\boldsymbol{a}_c+\boldsymbol{b}_c=1$. 融合后的特征 $\boldsymbol{V}$ 既包含全局特征的信息, 又包含局部特征的信息.

2.4　损失函数

MFF-Net 包括 2 个阶段, 以由粗到精的方式生成完整点云, 每个阶段都遵循编码器-解码器结构. 第 1 阶段中, 基于全局特征生成粗糙点云 Y_{C}, 被视为三维锚点, 以便于利用残缺点云 X 中的局部几何细节. 第 2 阶段中, 利用残缺点云 X 和补全的粗糙点云, 进一步生成高质量的完整点云形状 Y'. 因此在训练过程中, 本文对粗糙点云 Y_{C} 和完整的三维点云 Y' 都应用联合损失函数.

2.4.1　重构损失

倒角距离(Chamfer distance, CD)和 Earth Mover 距离(earth mover's distance, EMD)被广泛应用于点云补全. 本文选择对称 CD 损失计算点集 Y 和 Y' 的平均最近点距离, 表示为

$$L_{\mathrm{rec}}(Y',Y)=d_{\mathrm{CD}}(Y',Y)=\frac{1}{|Y'|}\sum_{P_x\in Y'}\min_{P_y\in Y}\left\|p_x-p_y\right\|^2+\frac{1}{|Y|}\sum_{P_y\in Y}\min_{P_x\in Y'}\left\|p_x-p_y\right\|^2 \tag{9}$$

其中, p_x 表示预测点集 Y' 中的一个点; p_y 表示最真值点集 Y 中的一个点.

2.4.2　均匀损失

受 PU-GAN[26]的启发, 本文在训练过程中也使用均匀损失生成均匀分布的点云, 公式为

$$\begin{cases}U_{\mathrm{imbalance}}=\dfrac{\left(|S_j|-\hat{n}\right)^2}{\hat{n}}\\ U_{\mathrm{clutter}}=\displaystyle\sum_{k=1}^{|S_j|}\frac{\left(d_{j,K}-\hat{d}\right)^2}{\hat{d}}\\ L_{\mathrm{uni}}(Y')=\displaystyle\sum_{j=1}^{M}U_{\mathrm{imbalance}}(S_j)\cdot U_{\mathrm{clutter}}(S_j)\end{cases} \tag{10}$$

其中, $\hat{n}=\left(rN\cdot r_d^2\right)$; $\hat{d}=\sqrt{2\pi r_d^2\Big/\left(|S_j|\sqrt{3}\right)}$; S_j 表示从预测的 Y' 中以半径为 r_d 的球查询采样得到的一个点子集; rN 表示上采样后的点个数; r_d^2 表示 S_j 中的期望点密度; $U_{\mathrm{imbalance}}$ 测量 $|S_j|$ 和 n 的偏差, 即 S_j 中的预测点数; U_{clutter} 惩罚点到邻居的距离 $d_{j,K}$ 与预期距离 $\hat{d}$ 的偏差.

2.4.3　联合损失

根据引入的损失函数, 本文的联合损失函数定义为

$$L=w_{\mathrm{C}}\left[\lambda_{\mathrm{rec}}L_{\mathrm{rec}}(Y_{\mathrm{C}},Y)+\lambda_{\mathrm{uni}}L_{\mathrm{uni}}(Y_{\mathrm{C}},Y)\right]+w_{\mathrm{F}}\left[\lambda_{\mathrm{rec}}L_{\mathrm{rec}}(Y',Y)+\lambda_{\mathrm{uni}}L_{\mathrm{uni}}(Y',Y)\right] \tag{11}$$

其中，λ_{rec}，λ_{uni}，w_C 和 w_F 分别是重构损失项，均匀损失项，粗预测和精细预测的权值.

3 实验与结果分析

为了评估本文模型的性能，在常用的数据集 ShapeNet, Completion3D 和 MVP 上进行实验.

3.1 数据集

(1) ShapeNet 数据集. PCN[4]提供的点云补全的数据集源于 ShapeNet 数据集. 该数据集由 8 个类别的 30 974 个三维模型组成，在网格表面上均匀采样包含 16 384 个点的真值；部分点云是通过将 2.5D 深度图反向投影到三维生成的. 为了便于比较，本文使用与 PCN 相同的数据集拆分比例.

(2) Completion3D 数据集. 该数据集是一个从 ShapeNet 数据集派生的大规模三维对象数据集. 通过将来自部分视图的 2.5 维深度图像反投影到三维空间中来生成部分三维形状，分别由 28 974 和 800 个样本组成，用于训练和验证；与 PCN 生成的 ShapeNet 数据集不同，真值点云中只有 2 048 个点.

(3) MVP 数据集. 一个高质量的多视图局部点云数据集，有 3 个特点: a. 由不同的部分点云组成; b. 由超过 10 万个高质量的残缺点云和完整点云组成; c. 由 16 个形状类别组成，包括部分形状和完整形状，用于训练和测试.

本文采用与 VRCNet[21]相同的设置，将模型分成训练集(62 400 个形状对)和测试集(41 600 个形状对). 为了公平地比较，本文在 MVP 数据集上使用与 VRCNet[21]相同的训练策略.

3.2 实验参数设置

MFF-Net 的网络训练参数设置如下：迭代次数为 100，批处理大小为 16，优化器为 Adam，初始学习率为 10^{-4}，衰减步骤为 150 000，衰减率为 0.7. 对于实验中比较的网络，本文保持其默认的训练设置，并报告其最佳补全结果，其结果采用 $CD\times10^4$ 度量.

3.3 ShapeNet 数据集的实验结果

为了实现公平的比较, MFF-Net 使用与 TopNet 相同的训练集和测试集，它们是 PCN 创建的数据集的子集. 与 TopNet 保持一致，将不完整的输入点和真值点设置为 2 048. 由于 TopNet 未提供测试集的真值，因此本文从 PCN 提供的 16 384 个点中抽出 2 048 个点，生成真值数据(与 TopNet 的设置保持一致). 首先评估 2 048 个点的补全性能，表 1 所示为 8 个模型对每个类别的补全量化结果. 由于残缺点云是从密集的三维表面上采集的，因此本文还使用不同分辨率(2 048, 4 096, 8 192, 16 384)的残缺点云评估 MFF-Net 的形状补全性能.

表 1 8 种模型在 ShapeNet 数据集上的误差比较

模型	单类								平均
	飞机	柜子	汽车	椅子	台灯	沙发	桌子	船	
AtlasNet[27]	10.37	23.40	13.41	24.16	20.24	20.82	17.52	11.62	17.69
PointNetFCAE[28]	10.31	19.07	11.83	24.69	20.31	20.09	17.57	10.50	16.80
FoldingNet[7]	11.18	20.15	13.25	21.48	18.19	19.09	17.80	10.69	16.48
PCN[4]	5.44	19.88	9.22	17.86	16.29	16.62	13.72	7.95	13.37
TopNet[16]	9.29	18.79	11.57	18.44	14.69	18.63	13.45	8.65	14.19
ECG[18]	2.86	11.98	6.16	9.67	10.04	11.64	9.72	7.02	8.64
VRCNet[21]	4.10	11.86	8.03	13.91	7.42	14.02	7.49	8.11	7.52
MFF-Net	2.53	10.81	6.36	8.37	7.28	9.79	7.06	5.04	7.16

表 2 所示为 7 个模型在不同分辨下的补全结果比较. 可以看出，对于不同分辨率的输入点云，MFF-Net 的补全效果优于对比的模型.

图 3 所示为 ShapeNet 数据集上不同网络及 MFF-Net 不同分辨率(2 048, 4 096, 8 192)的形状补全结果. 可以看出，与输入点云数据对比可以看出, MFF-Net 能够更好地保留残缺点云中的原始几何细节；对于一些较为复杂的点云形状，如台灯、汽车, MFF-Net 仍然能够补全重建出更接近真值的结果. 实验结果表明, MFF-Net 取得了更好的性能，得到更准确的整体形状和更好的局部结构.

3.4 MVP 数据集的实验结果

在 MVP 数据集上，针对不同分辨率(2 048, 4 096, 8 192, 16 384)进行点云补全实验，结果如表 3 所示. 可以看出，MFF- Net 表现良好，优于大多数对比模型；在较低分辨率(2 048, 4 096, 8 192)下，

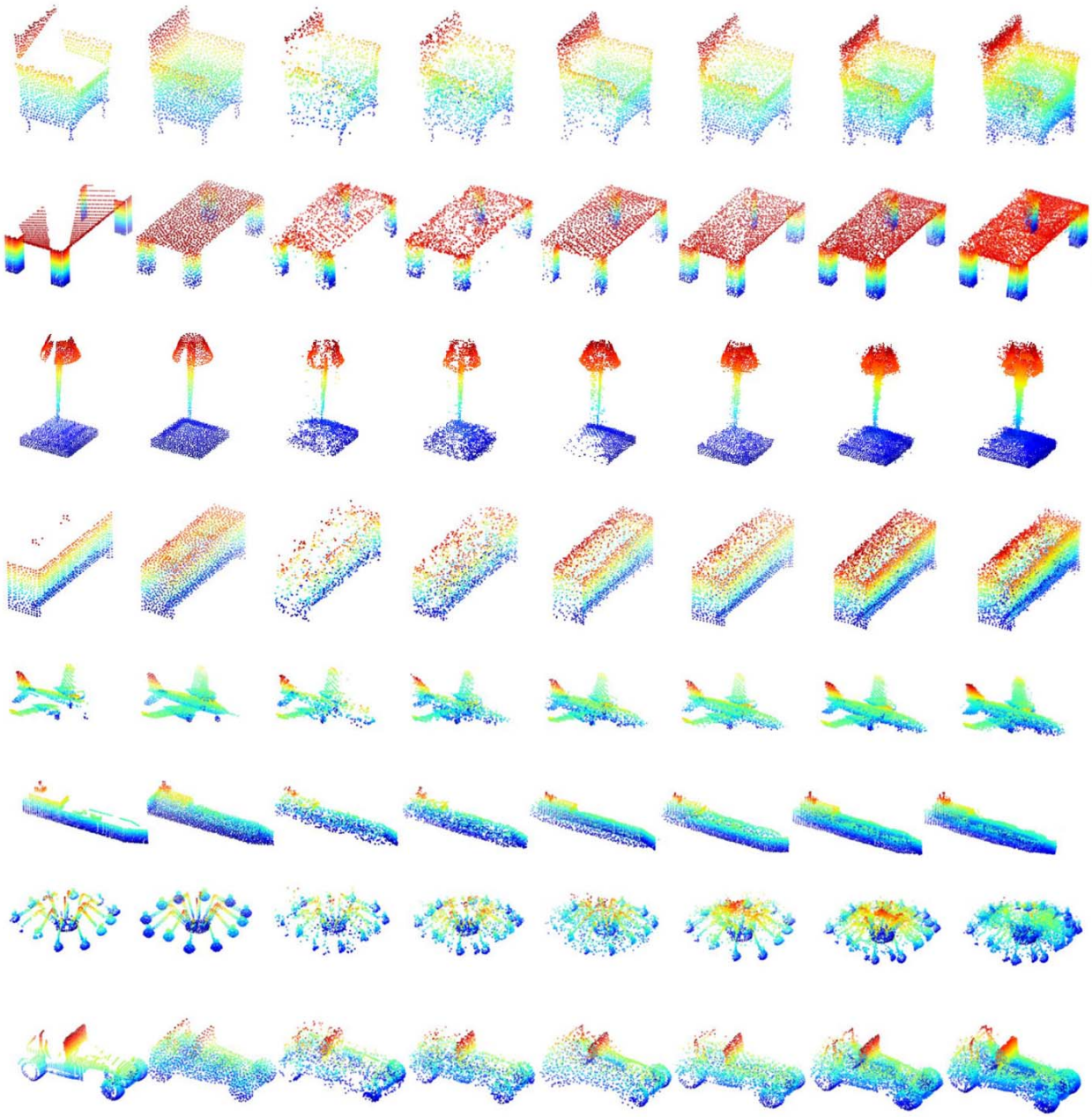
表 2　ShapeNet 数据集上不同分辨率的误差比较

模型	分辨率			
	2 048	4 096	8 192	16 384
AtlasNet[27]	17.69	16.16	15.32	14.85
PointNetFCAE[28]	16.80	14.79	12.57	10.61
FoldingNet[7]	16.48	13.19	12.95	12.26
PCN[4]	13.37	11.39	10.07	8.88
TopNet[16]	14.19	12.65	10.60	8.77
ECG[18]	8.64	8.19	7.53	7.11
MFF-Net	7.15	6.30	6.41	6.12

MFF-Net 的补全效果不如 VRCNet[21]，PMPNet[29]和 SnowflakeNet[30]，但仍优于大多数对比模型；MFF-Net 在较高分辨率(16 384)上的补全效果略差于 PMPNet[29]和 SnowflakeNet[30]，但达到了第 3 名的水平，说明 MFF-Net 在高分辨率补全重建中具有良好的效果.

3.5　Completion3D 数据集的实验结果

在 Completion3D 数据集上验证 MFF-Net 的泛化能力. 结果如表 4 所示. 可以看出，MFF-Net 表现良好，超过了基准网络 ECG[18]和大多数的对比



a. 输入点云 (2 048)　b. 真值 (2 048)　c. PCN[4] (2 048)　d. TopNet[16] (2 048)　e. ECG[18] (2 048)　f. MFF-Net (2 048)　g. MFF-Net (4 096)　h. MFF-Net (8 192)

图 3　ShapeNet 数据集上不同网络及 MFF-Net 不同分辨率的可视化结果

表 3 MVP 数据集上不同分辨率的误差比较

模型	分辨率			
	2 048	4 096	8 192	16 384
TopNet[16]	10.11	8.20	7.00	6.36
MSN[19]	7.90	6.17	5.42	4.90
Cascaded[31]	7.25	5.83	4.90	4.30
ECG[18]	6.64	5.41	4.18	3.58
VRCNet[21]	5.96	4.70	3.64	3.12
PMPNet[29]	6.33	4.63	3.52	2.79
SnowflakeNet[30]	6.06	4.80	3.49	2.75
MFF-Net	6.38	5.10	4.08	3.04

模型, 并在大多数类别中取得了比 ECG 更好的结果, 证明本文模型的改进思路是有效的.

3.6 消融实验

为了进一步验证本文模型中各模块的有效性, 在 MFF-Net 的基础上设计了不同组合模块下的消融实验. 如表 5 所示, A 模型为基准网络, 使用 PCN 的编码器部分产生一个粗糙点云, 再利用粗糙点云和残缺点云通过 DB 和 EP 补全重建完整的点云; B 模型使用 PCN 的编码器部分产生一个粗糙点云, 再利用粗糙点云和残缺点云通过 DB, EP 和 GLF 补全得到完整的点云; C 模型使用 PointNet++中的 SA 作为编码器部分产生一个粗糙点云, 再利用粗糙点云和残缺点云, 并通过 DB 和 EP 补全得到完整的点云; D 模型使用 PointNet++中的 SA 作为编码器部分产生一个粗糙点云, 再利用粗糙点云和残缺点云, 并通过 DB, EP 和 GLF 补全得到完整的点云.

表 4 13 个模型在 Completion3D 数据集上的误差比较

模型	单类								平均
	飞机	柜子	汽车	椅子	台灯	沙发	桌子	船	
FoldingNet[7]	12.83	23.01	14.88	25.69	21.79	21.31	20.71	11.51	19.07
PCN[4]	9.79	22.70	12.43	25.14	22.72	20.26	20.27	11.73	18.22
AtlasNet[27]	10.36	23.40	13.40	24.16	20.24	20.82	17.52	11.62	17.77
TopNet[16]	7.32	18.77	12.88	19.82	14.60	16.29	14.89	8.82	14.25
SoftPoolNet[32]	5.81	24.53	11.35	23.63	18.54	20.34	16.89	7.14	16.15
ECG[18]	4.68	17.12	10.60	14.89	9.36	18.35	11.92	6.94	11.34
SA-Net[17]	5.27	14.45	7.78	13.67	13.53	14.22	11.75	8.84	11.22
GRNet[33]	6.13	16.90	8.27	12.23	10.22	14.93	10.08	5.86	10.64
PMPNet[29]	3.99	14.70	8.55	10.21	9.27	12.43	8.51	5.77	9.23
VRCNet[21]	3.94	10.93	6.44	9.32	8.32	11.35	8.60	5.78	8.12
SnowflakeNet[30]	3.48	11.09	6.90	8.75	8.42	10.15	6.46	5.32	7.60
SoftPool++[34]	4.59	15.85	6.78	11.41	8.82	13.37	9.15	4.93	9.36
MFF-Net	4.56	14.08	8.20	11.66	9.27	15.64	9.89	6.35	9.76

表 5 用于消融实验的不同组合模块

模型	网络结构
A	PCN_encoder+DB+EP
B	PCN_encoder+DB+EP+GLF
C	PointNet++_encoder+DB+EP
D	PointNet++_encoder+DB+EP+GLF

在 ShapeNet 数据集上的消融实验结果如表 6 所示. 可以看出, 与基准模型 A 相比, 模型 B 和 C 中的大部分类别补全效果有所提升, 如台灯、船和桌子这类物体有明显的局部几何细节, MFF-Net 改进的网络对基准模型的补全效果有所提升; 模型 D 综合 PointNet++的 SA 层作为第 1 阶段的特征提取器, 可以有效地提取不同层次的特征, 同时引入融合模块, 两者相结合使得模型的性能大幅度提升.

表 6 在 ShapeNet 数据集上的消融实验结果

模型	单类误差								平均误差
	飞机	柜子	汽车	椅子	台灯	沙发	桌子	船	
A	2.86	11.98	6.16	9.67	10.04	11.64	9.72	7.02	8.64
B	3.10	11.98	6.54	10.36	7.82	12.36	8.86	5.86	8.30
C	2.92	12.05	6.63	10.48	7.76	12.00	8.82	6.06	8.36
D	2.53	10.81	6.36	8.37	7.28	9.79	7.06	5.04	7.16

4 结 语

本文提出一种两阶段网络 MFF-Net, 利用全局特征、局部边缘感知特征和特征融合模块补全点云.第 1 阶段利用多级 SA 和 MLP 从残缺点云中恢复完整的三维形状, 同时保留了点云的局部几何

细节. 第 2 阶段中的 GLF 融合全局特征和密集的局部边缘特征, 使得每个点都包含全局和局部几何信息. 与已有的大多数点云补全方法中的特征融合模块相比, GLF 充分利用全局特征与局部特征之间的优势互补. 在公开数据集上的实验结果表明, 本文模型是有效的.

残缺点云通常包含有用信息和干扰信息, 其中, 结构较为完整的部分被认为是有用信息, 而一些孤立的点则被认为是干扰信息. 由于 MFF-Net 对这些信息采用相同的特征提取方式, 因此受到干扰信息的误导, 影响到算法的整体性能. 如何有差别地利用这些信息, 是未来值得深入研究的问题.

参考文献(References):

[1] Charles R Q, Su H, Kaichun M, *et al*. PointNet: deep learning on point sets for 3D classification and segmentation[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 77-85

[2] Qi C R, Yi L, Su H, *et al*. PointNet++: deep hierarchical feature learning on point sets in a metric space[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 5105-5114

[3] Li Y Y, Bu R, Sun M C, *et al*. PointCNN: convolution on *X*-transformed points[C] //Proceedings of the 32nd International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2018: 828-838

[4] Yuan W T, Khot T, Held D, *et al*. PCN: point completion network[C] //Proceedings of the International Conference on 3D Vision. Los Alamitos: IEEE Computer Society Press, 2018: 728-737

[5] Boud A C, Haniff D J, Baber C, *et al*. Virtual reality and augmented reality as a training tool for assembly tasks[C] //Proceedings of the IEEE International Conference on Information Visualization. Los Alamitos: IEEE Computer Society Press, 1999: 32-36

[6] Dai A, Ritchie D, Bokeloh M, *et al*. ScanComplete: large-scale scene completion and semantic segmentation for 3D scans[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 4578-4587

[7] Yang Y Q, Feng C, Shen Y R, *et al*. FoldingNet: point cloud auto-encoder via deep grid deformation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 206-215

[8] Peng Z L, Huang W, Gu S Z, *et al*. Conformer: local features coupling global representations for visual recognition[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2021: 357-366

[9] Guo M H, Cai J X, Liu Z N, *et al*. PCT: point cloud transformer[J]. Computational Visual Media, 2021, 7(2): 187-199

[10] Dai A, Qi C R, Nießner M. Shape completion using 3D-encoder-predictor CNNs and shape synthesis[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 6545-6554

[11] Han X G, Li Z, Huang H B, *et al*. High-resolution shape completion using deep neural networks for global structure and local geometry inference[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 85-93

[12] Wu Z R, Song S R, Khosla A, *et al*. 3D ShapeNets: a deep representation for volumetric shapes[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2015: 1912-1920

[13] Wang P S, Liu Y, Guo Y X, *et al*. O-CNN: octree-based convolutional neural networks for 3D shape analysis[J]. ACM Transactions on Graphics, 2017, 36(4): Article No.72

[14] Wang Y, Sun Y B, Liu Z W, *et al*. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): Article No.146

[15] Pan L, Chew C M, Lee G H. PointAtrousGraph: deep hierarchical encoder-decoder with point atrous convolution for unorganized 3D points[C] //Proceedings of the IEEE International Conference on Robotics and Automation. Los Alamitos: IEEE Computer Society Press, 2020: 1113-1120

[16] Tchapmi L P, Kosaraju V, Rezatofighi H, *et al*. TopNet: structural point cloud decoder[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 383-392

[17] Wen X, Li T Y, Han Z Z, *et al*. Point cloud completion by skip-attention network with hierarchical folding[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 1936-1945

[18] Pan L. ECG: edge-aware point cloud completion with graph convolution[J]. IEEE Robotics and Automation Letters, 2020, 5(3): 4392-4398

[19] Liu M H, Sheng L, Yang S, *et al*. Morphing and sampling network for dense point cloud completion[C] //Proceedings of the 34th AAAI conference on artificial intelligence. Palo Alto: AAAI Press, 2020: 11596-11603

[20] Zhang W X, Yan Q G, Xiao C X. Detail preserved point cloud completion via separated feature aggregation[C] //Proceedings of the 16th European Conference on Computer Vision. Heidelberg: Springer, 2020: 512-528

[21] Pan L, Chen X Y, Cai Z G, *et al*. Variational relational point completion network[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 8524-8533

[22] Xie C L, Wang C X, Zhang B, *et al*. Style-based point generator with adversarial rendering for point cloud completion[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 4617-4626

[23] Yu X M, Rao Y M, Wang Z Y, *et al*. PoinTr: diverse point cloud completion with geometry-aware transformers[C] //Procee-

dings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2021: 12498-12507

[24] Komarichev A, Zhong Z C, Hua J. A-CNN: annularly convolutional neural networks on point clouds[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 7421-7430

[25] Li X, Wang W H, Hu X L, *et al*. Selective kernel networks[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 510-519

[26] Li R H, Li X Z, Fu C W, *et al*. PU-GAN: a point cloud upsampling adversarial network[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2019: 7203-7212

[27] Vakalopoulou M, Chassagnon G, Bus N, *et al*. AtlasNet: multi-atlas non-linear deep networks for medical image segmentation[C] //Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention. Heidelberg: Springer, 2018: 658-666

[28] Achlioptas P, Diamanti O, Mitliagkas I, *et al*. Learning representations and generative models for 3D point clouds[C] // Proceedings of the 35th International Conference on Learning Representations. Proceedings of the 35th International Conference on Machine Learning. Lille: PMLR Press, 2018: 40-49

[29] Wen X, Xiang P, Han Z Z, *et al*. PMP-net: point cloud completion by learning multi-step point moving paths[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 7443-7452

[30] Xiang P, Wen X, Liu Y S, *et al*. SnowflakeNet: point cloud completion by snowflake point deconvolution with skip-transformer[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2021: 5499-5509

[31] Wang X G, Ang Jr M H, Lee G H. Cascaded refinement network for point cloud completion[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 790-799

[32] Wang Y D, Tan D J, Navab N, *et al*. SoftPoolNet: shape descriptor for point cloud completion and classification[C] // Proceedings of the 16th European Conference on Computer Vision. Heidelberg: Springer, 2020: 70-85

[33] Xie H Z, Yao H X, Zhou S C, *et al*. GRNet: gridding residual network for dense point cloud completion[C] //Proceedings of the 16th European Conference on Computer Vision. Heidelberg: Springer, 2020: 365-381

[34] Wang Y D, Tan D J, Navab N, *et al*. SoftPool++: an encoder–decoder network for point cloud completion[J]. International Journal of Computer Vision, 2022, 130(5): 1145-1164