

文章编号: 1007-5321(2025)05-0017-08

DOI: 10.13190/j.jbupt.2025-073

面向物联网三维感知的点云到 CAD 模型重建方法

赵帅, 夏镇, 陈俊亮, 程渤, 杜晨阳

(北京邮电大学网络与交换技术全国重点实验室, 北京 100876)

摘要: 随着物联网驱动的自主导航等领域对复杂三维环境理解需求的增长,从点云数据自动重建结构化、可编辑的计算机辅助设计(CAD)模型已成为关键任务。然而,当前研究主要从CAD命令序列、草图与拉伸操作等角度进行重建,普遍面临重建步骤多、平台依赖性强等问题。为此,提出一种高精度几何与拓扑关系的CAD重建方法,即面向物联网三维感知的点云到CAD模型重建方法。首先,将CAD的边界表示解析为几何基元的参数化信息与其拓扑结构,并设计基元变分自编码器(PVAE)与拓扑变分自编码器(TVAE)分别建模其几何特征与拓扑关系;在此基础上,利用PointNet++提取点云数据的多尺度局部特征并融合为全局特征,进而通过拓扑解码器与基元解码器分别预测拓扑树序列和基元参数,实现高精度与更清晰的边界细节的CAD模型重建。为验证方法的有效性,采用倒角距离、边缘倒角距离和结构一致性作为评价指标,在2个数据集上对方法性能进行评估。实验结果表明,所提方法在拓扑结构完整性和几何重建精度方面均优于现有方法,能够实现更高精度的CAD模型重建。

关键词: 物联网三维感知; 点云; 计算机辅助设计; 边界表示法; PointNet++; 变分自编码器

中图分类号: TP391

文献标志码: A

CAD Model Reconstruction Method from Point Clouds for IoT 3D Perception

ZHAO Shuai, XIA Zhen, CHEN Junliang, CHENG Bo, DU Chenyang

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: With the growing demand for understanding complex three dimensions (3D) environments in fields like Internet of things (IoT)-driven autonomous navigation and related applications, the automatic reconstruction of structured, editable computer-aided design (CAD) models from point cloud data has become a critical task. However, current research mainly focuses on reconstructing CAD models from CAD command sequences, sketches, and extrusion operations, and commonly faces challenges such as excessive reconstruction steps and strong platform dependence. To this end, a high-precision geometric and topological CAD reconstruction method is proposed, specifically for IoT 3D perception. First, the boundary representation of CAD is decomposed into parametric information of geometric primitives and their topological structure. Then, a primitive variational autoencoder (PVAE) and a topological variational autoencoder (TVAE) are designed to model their geometric features and topological relationships, respectively. Furthermore, PointNet++ is used to extract multi-scale local features from point cloud data and fuse them into global features. A topological decoder and primitive decoders are then used to predict topological tree sequences and primitive parameters, achieving high-precision CAD model reconstruction with clearer boundary details. To validate the method's effectiveness, metrics such as chamfer distance, edge chamfer distance, and structural consistency are used to evaluate its performance

收稿日期: 2025-07-14

基金项目: 国家自然科学基金项目(62472043, U21A20468)

作者简介: 赵帅(1986—),男,教授,博士生导师,邮箱:zhaoshuai@bupt.edu.cn。

on two datasets. The experimental results show that the proposed model outperforms existing methods in terms of topological integrity and geometric reconstruction accuracy, enabling higher-precision CAD model reconstruction.

Key words: Internet of things three dimensions perception; point cloud; computer-aided design; boundary representation; PointNet ++; variational autoencoder

随着物联网 (IoT, Internet of things) 感知对复杂三维 (3D, three dimensions) 环境理解中的需求日益增加, 将 Photoneo、MechMind 等设备^[1-2]感知到的 3D 数据 (点云) 转换为结构化、可编辑的计算机辅助设计 (CAD, computer-aided design) 模型, 已成为数字孪生^[3]、自主导航^[4]等物联网驱动应用的关键任务。在复杂的物联网应用中, 包括仿真、规划和装配等下游任务都是在 CAD 模型的基础上进行的。因此, 从点云数据到 CAD 模型的重建不仅能够为下游任务提供可验证的 3D 数据, 还能够生成可编辑的 CAD 模型, 为后续应用提供基础支持。

在 CAD 重建研究中, 现有方法主要通过 3 类途径实现: 重构 CAD 设计命令序列^[5-7], 草图与拉伸构成^[8-10], 或者几何参数化信息与拓扑关系^[11-12]。对于 CAD 设计命令序列的重构, 通常将其视为 1 个长序列, 并对每条命令的类型及对应参数进行预测。例如, Ma 等^[5]提出一种基于多模态扩散的方法, 从点云生成 CAD 构造序列, 实现 CAD 重建。然而, 由于同一几何体可能对应多种有效命令序列, 其模型生成结果依赖数据中的规律。此外, 对于复杂 CAD 模型, 过长的命令序列会导致模型训练困难, 从而难以满足创新性设计场景的需求。为克服这一局限, 后续研究转向模拟实际建模过程的分步重构方法。例如, Li 等^[8]提出通过学习 2D 草图与拉伸操作, 从原始几何重建可编辑的 CAD 模型。然而, 这些方法存在显著的软件依赖性。由于不同 CAD 软件在草图设计和拉伸操作上存在建模差异, 跨平台使用时需要适配相应的设计约束逻辑。随后, Liu 等^[12]将拓扑关系与几何参数化信息结合为整体编码, 并支持多模态输入来生成 CAD 模型。然而, 这些重建方法通常将几何参数化信息与拓扑结构解耦处理。在重建拓扑时, 仍需独立建模不同几何基元之间的关联关系, 这导致 CAD 重建过程依赖多步迭代, 从而降低重建效率并可能影响模型一致性。目前, 从点云到 CAD 的重建方法通常需要多步迭代, 这可能导致 CAD 模型的误差随步骤数量增加而累积。此外,

在 CAD 重建中, 还存在几何基元参数精度低以及拓扑结构不完整的问题。

针对上述问题, 提出了面向高精度几何与拓扑重构的点云到 CAD 高效重建模型 (HGTR, efficient reconstruction model from point clouds to CAD for high-precision geometric and topological recovery)。该模型能够同时重建几何基元与拓扑结构, 从而生成具有更高精度和更清晰边界的 CAD 模型。具体而言, 为提高 CAD 重建精度, 首先将 CAD 的边界表示 (BRep, boundary representation) 解析为不同基元 (点、线、环、面) 的参数化信息及其拓扑结构。然后, 针对不同基元的参数化信息, 设计基元变分自编码器 (PVAE, primitive variational autoencoder) 进行训练; 同时, 将拓扑结构构建为拓扑树序列, 用于训练拓扑变分自编码器 (TVAE, topology variational autoencoder); 最后, 使用 PointNet ++^[13]对点云数据进行编码, 提取多尺度局部特征并融合为全局特征。该全局特征随后输入拓扑解码器与基元解码器, 分别预测拓扑结构和基元参数信息, 从而实现高精度 CAD 重建。主要贡献总结如下:

1) 为了能重建出更高精度的几何基元, 将 CAD 模型解析为几何基元的参数化信息及对应的拓扑结构, 并设计不同的变分自编码器;

2) 将拓扑结构构建为拓扑树序列, 设计基于 Transformer 的变分自编码器以获取节点间的关系, 从而实现更完整的拓扑结构重建;

3) 实验结果表明, HGTR 方法在大型 CAD 模型数据集 (ABC, a big CAD model dataset) 和 Fusion 360 数据集上能重建出完整的 CAD 拓扑结构, 并重建出更高精度和边界细节更清晰的 CAD 模型。

1 CAD 重建方法

1.1 方法总述

HGTR 方法由 3 个模块组成: 基于 PointNet ++ 的点云编码模块、拓扑结构重建模块和几何参数化重建模块。该方法基于深度学习, 将传感器感知到的稀疏且不规则点云数据高效重建为具有精确几何

参数和拓扑结构的 CAD 模型,从而实现复杂物体的精确几何和拓扑重建,并满足物联网感知驱动的数字孪生等应用需求。该方法的整体结构如图 1 所示。

字孪生等应用需求。该方法的整体结构如图 1 所示。

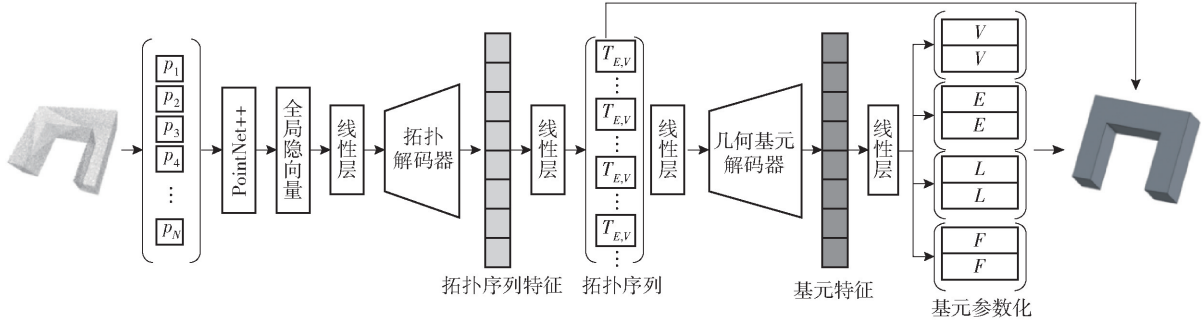


图 1 HGTR 方法结构图

1.2 问题定义

点云到 CAD 模型的高效重建问题旨在将传感器感知采集的三维点云数据转化为具备准确几何基元参数和完整拓扑结构的 CAD 模型。具体而言,给定 1 组由传感器感知的点云集 $P = \{p_1, p_2, \dots, p_N\}$, $p_i \in \mathbb{R}^3$ 作为输入,方法需要重建出基元(点、线、环和面)参数化信息与拓扑结构,并重构出 CAD 的 BRep 信息^[14]。

HGTR 方法通过几何基元参数化和拓扑结构重建相结合的方式能有效降低重建误差。CAD 的 BRep 可解构为几何基元参数化信息和对应的拓扑结构,其中几何基元包括点、线、环和面,具体解构 CAD 得到的参数化信息如下。

1) 点: $V = \{x, y, z\}$ 表示点的参数化信息由三维坐标组成。

2) 线: $E = \{C, o_e, d_e, a_e, x_e, r, r_m, b_e\}$ 表示线的参数化信息,其中 C 为曲线类型, $o_e \in \mathbb{R}^3$ 为线的原点坐标, $d_e \in \mathbb{R}^3$ 为线的方向向量, $a_e \in \mathbb{R}^3$ 为线的轴向量, $x_e \in \mathbb{R}^3$ 定义局部坐标系的 X 轴方向,通常与轴向量 a_e 垂直, r 为半径, r_m 为椭圆的短半轴半径, b_e 为方向布尔值。

3) 环: $L = \{b_e, E_1, E_2, \dots, E_N\}$ 表示环的参数化信息,其中 b_e 为方向布尔值, $\{E_1, E_2, \dots, E_N\}$ 表示环中包含 N 条曲线。

4) 面: $F = \{S, o_f, n_f, a_f, x_f, r, r_m, \alpha, b_f\}$ 表示面的参数化信息,其中 S 为面的类型, $o_f \in \mathbb{R}^3$ 为面的原点坐标, $n_f \in \mathbb{R}^3$ 为面的法向量, $a_f \in \mathbb{R}^3$ 为轴向量的法向量, $x_f \in \mathbb{R}^3$ 为局部坐标系的 X 轴的方向, r 为主半径, r_m 为次半径, α 为半角角度, b_f 为方向布尔值。

CAD 的拓扑结构根据解构到的参数化信息可表示为 $T = \{T_{F,F}, T_{F,L}, T_{L,E}, T_{E,V}\}$,其中 $T_{F,F}$ 为面和

面的关联关系, $T_{F,L}$ 为面和环的关联关系, $T_{L,E}$ 为环和线的关联关系, $T_{E,V}$ 为线和点的关联关系。

1.3 变分自编码器

CAD 几何基元参数化信息重建对精度要求较高。为此,对 BRep 解构出几何基元(点、线、环和面)分别使用不同的基元变分自编码器 PVAE 拟合其数据集中不同基元的数据分布,从而为后续基于点云编码得到的全局特征解码提供支持,图 2 展示了以线为例的 VAE 方法结构。在 CAD 拓扑结构的重建过程中,将不同几何基元之间的关联关系构建为拓扑树序列,并采用拓扑变分自编码器对该序列进行建模与拟合。

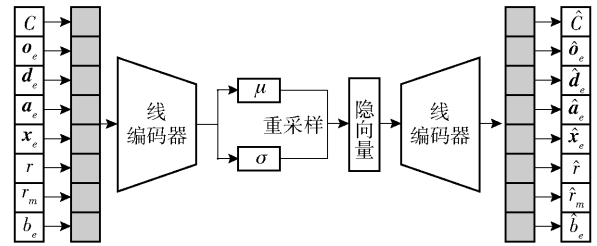


图 2 几何基元(线)的 VAE 方法结构图

1.3.1 基元变分自编码器

基于 Transformer 的 VAE 被用于学习 CAD 几何基元(点、线、环和面)的高效潜在表示。该结构通过自注意力机制建模基元参数间的相关性与层次关系,相较传统多层感知机 (MLP, multilayer perceptron) 或循环神经网络具有更强的高维特征表达能力。结合 VAE 的潜在空间连续建模,可精确重建几何参数并生成合理的新基元,为 CAD 重建提供高质量表示。

1) 基元编码器。给定线的参数化信息 $E_i \in \mathbb{R}^{d_e}$, 首先使用线性层映射其潜在特征 $h_o \in \mathbb{R}^{d_m}$, 随后将

该潜在特征输入至多层标准 Transformer, 利用自注意力机制捕捉线的几何信息在不同维度的相互依赖关系, 最终输出对应的上下文特征 $\mathbf{h}_{\text{enc}} \in \mathbb{R}^{d_{\text{enc}}}$ 为

$$\mathbf{h}_{\text{enc}} = \Phi_{\text{enc}}(\Phi_{\text{MLP}}(E_i)) \quad (1)$$

其中 $\Phi_{\text{enc}}(\cdot)$ 为 Transformer 编码器。基于上下文特征, 分别利用 2 个独立的线性层得到近似分布 $\varphi_{\phi}(z_E | E_i)$ 的均值 $\boldsymbol{\mu}$ 与对数方差 $\ln \sigma^2$ 。

为了使训练过程中实现梯度对采样步骤的反向传播, 不直接从近似分布中直接采样, 而是采用重参数化技巧, 从标准正态分布 $\mathcal{N}(0, I)$ 中采样 1 个随机噪声 $\boldsymbol{\varepsilon}$, 并与均值 $\boldsymbol{\mu}$ 和方差 σ 结合得到隐向量为

$$\mathbf{z}_E = \boldsymbol{\mu} + \sigma \odot \boldsymbol{\varepsilon} \quad (2)$$

其中 \odot 表示逐元素相乘, 通过这种方式隐向量 \mathbf{z}_E 即包含来自编码器确定性信息 ($\boldsymbol{\mu}$ 和 σ) 及随机性 ($\boldsymbol{\varepsilon}$), 并保证整个过程可微, 实现梯度反向传播。

2) 基元解码器。为了精确重构线的参数化信息, 将重采样的隐向量 $\mathbf{z}_E \in \mathbb{R}^{d_E}$ 输入至多层标准 Transformer 结构的解码器, 将其映射回高维特征空间, 以获得上下文特征向量 $\mathbf{h}_{\text{dec}} \in \mathbb{R}^{d_{\text{dec}}}$, 最后使用线性层将其重构为与原始线的参数化信息相同维度的 $\hat{E}_i \in \mathbb{R}^{d_e}$ 为

$$\hat{E}_i = \Phi_{\text{MLP}}(\Phi_{\text{dec}}(\mathbf{z}_E)) \quad (3)$$

1.3.2 拓扑变分自编码器

为建模 CAD 模型中基元间复杂拓扑关系, 将图形式拓扑结构序列化为可学习序列, 用 Transformer 捕捉全局依赖并建模潜在在拓扑约束。基于此构建基于 Transformer 的 VAE 结构, 编码器提取拓扑特征, 潜变量实现连续分布建模, 获得结构一致性且可生成的拓扑表示, 为 CAD 重建提供结构先验。

1) 拓扑编码器。解析 BRep 信息得到 CAD 的拓扑结构, 将其构建为 1 个拓扑树序列 $\{T_N^S\}$, 其中每个元素表示拓扑关系或间隔符。类型包括: $T_{F,F}$, $T_{F,L}$, $T_{L,E}$, $T_{E,V}$ 及几何基元 (V, E, L 和 F)。同时, 引入特殊符号分隔符 (SEP, separator token) 和填充符 (PAD, padding token), $\langle \text{SEP} \rangle$ 用于区分关系类型, $\langle \text{PAD} \rangle$ 用于填充。词表由填充符、拓扑关系和几何基元组成。

为了捕捉拓扑序列位置上的顺序信息, 引入位置编码。对于序列中的第 i 个位置, 嵌入维度索引 k , 其位置编码 $\Phi_{\text{pos}}(i)_k \in \mathbb{R}^{T \times d_m}$ 定义为

$$\Phi_{\text{pos}}(i)_k = \begin{cases} \sin(i/10\,000^{2k/d_m}), & k \text{ 为偶数} \\ \cos(i/10\,000^{2k/d_m}), & k \text{ 为奇数} \end{cases} \quad (4)$$

对输入序列进行嵌入映射, 与位置编码进行逐

元素相加得

$$\mathbf{h}_i = \Phi_{\text{Embed}}(T_N^S) + \Phi_{\text{pos}}(T_N^S) \quad (5)$$

随后通过多头自注意力机制建模拓扑依赖得

$$\mathbf{h}' = \Phi_{\text{MultiAttn}}(\mathbf{h}_i) \quad (6)$$

然后通过前馈神经网络对多头自注意力输出进行非线性变换, 并结合残差连接与层归一化得到编码器输出, 可表示为

$$\mathbf{h}_{\text{enc}} = \Phi_{\text{Norm}}(\mathbf{h}' + \text{ReLU}(\Phi_{\text{MLP}}(\mathbf{h}')))) \quad (7)$$

其中 $\Phi_{\text{Norm}}(\cdot)$ 表示层归一化。为近似分布 $\varphi_{\phi}(z_T | T_N^S)$, 通过 2 个独立的线性层分别处理得到其均值 $\boldsymbol{\mu}$ 与对数方差 $\ln \sigma^2$ 。随后采样其拓扑树序列的隐向量 \mathbf{z}_T 。

2) 拓扑解码器。为实现对拓扑树序列的有效重建, 使用基于 Transformer 架构的解码器处理隐向量 \mathbf{z}_T 生成拓扑树序列表示。解码器由多层 Transformer 与序列映射层组成, 首先将隐向量 \mathbf{z}_T 输入到 Transformer 中通过注意力机制建模拓扑元素间的依赖关系, 然后通过线性层与 softmax 将表示映射至序列空间, 并生成对应的拓扑序列指令概率分布。

1.4 基于 PointNet++ 的点云编码

对于实际场景物体, 通过 Photoneo 或 MechMind 等传感器感知获得点云 $P = \{p_1, p_2, \dots, p_N\}$, $p_i \in \mathbb{R}^3$, 其中 N 表示采样点数。为弥补点云与 CAD 中 BRep 参数化和拓扑结构信息的差异, 采用 PointNet++ 方法对点云编码, 提取其局部细节和全局轮廓的多维度几何特征。

1) 点集抽象模块。首先, 使用最远点采样 (FPS, farthest point sampling) 从 P 筛选出 N' 个质心 $\{c_1, c_2, \dots, c_{N'}\}$, 对每个质心 c_i 通过半径 r 查询得到局部点集 G_i 。为保证平移不变性, 对局部点归一化处理得

$$p'_{i,j} = p_{i,j} - c_i, \forall p_{i,j} \in G_i \quad (8)$$

将归一化后的坐标及对应的附加特征 $g_{i,j}^{L-1}$ (第 $(L-1)$ 层集合抽象模块最后的输出) 进行拼接, 并通过线性层与最大池化层提取该层的局部特征为

$$\mathbf{v}_i = \max_{p_{i,j} \in G_i} \{ \Phi_{\text{MLP}}(p'_{i,j} \| g_{i,j}^{L-1}) \} \quad (9)$$

其中: \mathbf{v}_i 表示聚合到质心 c_i 的局部特征, $\max_{p_{i,j} \in G_i} \{ \cdot \}$ 为最大池化操作, $\|$ 为拼接操作。

2) 全局特征提取。每一层中包含多个点集抽象模块, 每个模块将局部点集的点云聚合到质心, 经过 L 层集合抽象处理后可得到 N_L 个点及其高阶特

征的集合。为了提取输入点云集的全局特征, 首先将每个点通过线性层映射到统一的语义空间, 然后使用最大池化层获取其全局特征为

$$\mathbf{H} = \max_{i \in \{1, 2, \dots, N_L\}} \{ \Phi_{\text{MLP}}(\mathbf{v}_i) \} \quad (10)$$

其中: $\mathbf{H} \in \mathbb{R}^{d_m}$ 为最终经 PointNet++ 编码点云后的全局特征表示, d_m 为重建 CAD 模块的输入维度。

1.5 CAD 重建

为了从点云全局特征 \mathbf{H} 重建 CAD, 使用 2 种解码器分别重建 CAD 的拓扑树序列和几何基元的参数化信息。对于拓扑树序列的重建, 首先通过线性层将点云特征映射至拓扑树序列的语义空间为

$$\hat{\mathbf{z}}_T = \mathbf{H}\mathbf{W} + \mathbf{b} \quad (11)$$

其中 $\hat{\mathbf{z}}_T$ 与 \mathbf{z}_T 的维度保持一致。随后将 $\hat{\mathbf{z}}_T$ 输入拓扑解码器, 通过标准的 Transformer 层获得拓扑树序列潜在特征, 并通过线性层与 softmax 还原至拓扑树序列相应维度, 以得到与对应的拓扑树序列概率分布为

$$\hat{T}_N^S = \text{softmax}(\Phi_{\text{MLP}}(\Phi_{\text{dec}}^T(\hat{\mathbf{z}}_T))) \quad (12)$$

对于几何基元参数的重建, 首先利用不同线性层将点云全局特征映射至不同基元的语义空间, 以获得相应的隐向量 $\hat{\mathbf{z}}_g, g \in \{V, E, L, F\}$, 随后将这些隐向量输入到几何基元解码器中提取潜在特征, 最后通过不同的线性层获得几何基元参数化信息为

$$\hat{\mathbf{X}}_g = \Phi_{\text{MLP}}(\Phi_{\text{dec}}^g(\hat{\mathbf{z}}_g)) \quad (13)$$

其中 $\hat{\mathbf{X}}_g$ 表示重建后的几何基元参数化信息, 与重建后的拓扑结构结合用于重构 CAD。

1.6 训练

训练过程主要分为 2 个阶段, 第 1 个阶段分别训练拓扑结构与几何基元参数化的解码器, 其损失由各自的 VAE 损失决定。以线为例, 通过最大化数据的边际对数似然 $\ln \psi(E_i)$, 即优化其证据下界 (ELBO, evidence lower bound), 以学习最优的编解码器参数 (ϕ, θ) 为

$$\mathcal{L}(\phi, \theta; E_i) = -E_{\mathbf{z} \sim \varphi_\phi(\mathbf{z}|E_i)} [\ln \psi_\theta(E_i|\mathbf{z})] + D_{\text{KL}}[\varphi_\phi(\mathbf{z}|E_i) \parallel \psi(\mathbf{z})] \quad (14)$$

其中: $E[\dots]$ 为重构损失, 其最小化的目标是使重构后的几何参数化信息尽可能接近原始输入; $D_{\text{KL}}[\dots]$ 为 KL 散度, 其目标是约束编码器输出的分布尽可能与先验分布 $\psi(\mathbf{z})$ 相似。

第 2 阶段训练的目标是实现点云到 CAD 的重建, 损失函数由拓扑结构重建损失和几何基元(点、线、环和面)参数重建损失组成。拓扑序列被离散化为“词表索引”, 重建等价于分类任务, 采用交叉

熵 (CE, cross entropy) 损失, 如式 (15a) 所示; 其次, 几何基元(点、边、环和面)参数的重建, 其参数为连续值, 因此采用均方误差 (MSE, mean squared error) 作为损失函数, 如式 (15b) 所示。最终, 整体损失由 2 部分组合而成, 可表示为

$$\mathcal{L}_{\text{topo}} = - \sum_{i=1}^N \sum_{j=1}^M T_{i,j} \ln T'_{i,j} \quad (15a)$$

$$\mathcal{L}_g = \frac{1}{N} \sum_{i \in N} (y_i - \hat{y}_i)^2, g \in \{V, E, L, F\} \quad (15b)$$

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{topo}} + \sum_{g \in \{V, E, L, F\}} \mathcal{L}_g \quad (15c)$$

其中: $\mathcal{L}_{\text{total}}$ 为方法总损失函数, $\mathcal{L}_{\text{topo}}$ 为重建 CAD 拓扑结构的 CE 损失, 其中 $\hat{\mathbf{V}}$ 表示预测的拓扑序列分布, \mathbf{V} 表示真实的拓扑序列。 $\mathcal{L}_g, g \in \{V, E, L, F\}$ 为重建 CAD 几何参数化信息的 MSE 损失, 其中 \hat{y}_i 和 y_i 分别表示预测和真实的几何参数。

2 实验与分析

2.1 实验设置

1) 数据集。为了验证所提方法在 CAD 重建任务中的有效性, 在 ABC 数据集^[15]与 Fusion 360 数据集^[16]进行实验验证。ABC 数据集包括 6 000 个 CAD 模型, 其中训练集数量为 5 000, 测试集数量为 1 000。Fusion 360 数据集包括 3 700 个 CAD 模型, 其中训练集数量为 3 000, 测试集数量为 700。针对 2 个数据集, 移除面数量超过 50 个的 CAD 模型。

2) 评价指标。通过 3 种通用的定量评价指标倒角距离 (CD, chamfer distance)、边缘倒角距离 (ECD, edge chamfer distance)、一致性 (NC, normal consistency) 评价所提方法在 CAD 模型重建任务的效果进行评估。

倒角距离用于衡量预测点集 P 和真实点集 Q 的对称距离, 则 CD 定义为

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

边缘倒角距离用于评估预测边缘点集 P_e 和真实边缘点集 Q_e 的结构特征的还原能力, ECD 定义为

$$D_{\text{ECD}}(P_e, Q_e) = \frac{1}{|P_e|} \sum_{p \in P_e} \min_{q \in Q_e} \|p - q\|_2^2 + \frac{1}{|Q_e|} \sum_{q \in Q_e} \min_{p \in P_e} \|q - p\|_2^2 \quad (17)$$

一致性衡量预测法向量 $\{\hat{n}_i\}_{i=1}^N$ 与真实法向量 $\{n_i\}_{i=1}^N$ 间的一致性, NC 定义为

$$D_{\text{NC}} = \frac{1}{N} \sum_{i=1}^N |\langle \hat{n}_i, n_i \rangle| \quad (18)$$

3) 超参数设置。实验基于 PyTorch 框架实现,并在搭载 NVIDIA Tesla A800 80G 显卡的高性能服务器上训练。TVAE 的编解码器中 Transformer 层数均为 6 层,隐藏层大小为 256;PVAE 的编解码器中 Transformer 层数均为 4 层,隐藏层大小为 128;训练过程中使用 Adam 优化器,初始学习率为 1×10^{-4} ,训练周期为 100 个 Epoch,批量大小 (batch size) 设为 16,输入点云大小为 8 192。

4) 对比方法。为验证 HGTR 方法的重建效果,将在 2 个数据集上与构造实体几何树的无监督发现 (UCSG, unsupervised discovering of constructive solid geometry tree)-Net^[17], CSG-Stump^[18], ExtrudeNet^[19], 素描-拉伸式计算机辅助设计 (SECAD, sketch-extrude computer-aided design)-Net^[8], SfmCAD^[9] 方法进行对比与分析,进一步评估其有效性和可行性。

2.2 实验结果与分析

不同重建方法在数据集 ABC 和 Fusion 360 上的实验结果如表 1 所示,主要从 3 个评价指标 CD、ECD 和 NC 上进行评价,这些指标综合衡量了 CAD 重建模型在几何精度和表面光滑方法的表现。

表 1 2 个数据集上不同方法实验结果

数据集	方法	CD ↓	ECD ↓	NC ↑
ABC	UCSG-Net	1.233	21.786	0.866
	CSG-Stump	0.671	7.751	0.892
	ExtrudeNet	0.519	7.111	0.885
	SECAD-Net	0.506	7.286	0.884
	SfmCAD	0.395	5.038	0.919
	HGTR	0.354	4.867	0.934
Fusion 360	UCSG-Net	2.950	5.277	0.770
	CSG-Stump	2.781	4.590	0.744
	ExtrudeNet	2.263	3.558	0.819
	SECAD-Net	2.052	3.282	0.803
	HGTR	1.965	2.884	0.821

表 1 结果显示 HGTR 方法在 2 个数据集上的 3 个评价指标均取得最优表现。在 ABC 数据集上, HGTR 的 CD 达到 0.354, NC 为 0.934, 相较于性能最好的基准方法 (SfmCAD) 分别提升 10.38% 和

1.63%; 在 Fusion 360 数据集上, HGTR 的 CD 达到 1.965, NC 为 0.821, 相较于性能最好的基准方法 (SECAD-Net) 分别提升 4.24% 和 2.24%; 这些结果表明, HGTR 方法能重建出几何精度更高、边界细节更清晰的 CAD 模型, 同时 NC 指标最优表示该方法在法向一致性上表现最为出色, 能重建出更光滑、连续的表面。相比之下, 早期使用构造实体几何 (CSG, constructive solid geometry) 的 UCSG-Net 和 CSG-Stump 的 CD 和 ECD 上表现较差, 尤其是在 ABC 数据集上, 说明其在复杂结构的 CAD 模型的几何拟合与边界重建能力相对较弱; ExtrudeNet、SECAD-Net 和 SfmCAD 方法相较于 CSG 方法性能有明显提升, 主要原因在于, 这类方法通常基于 CAD 的构建过程, 通过草图绘制与拉伸操作实现 CAD 模型重建, 在几何精度上都有一定程度提升。与上述方法相比, HGTR 通过同时重建几何基元的参数化信息和拓扑结构, 实现了更精细化的 CAD 重建, 因此能够获得更优的性能, 充分证明了 HGTR 在 CAD 重建任务中的有效性。

2.3 消融实验

HGTR 方法采用基于 Transformer 的变分自编码器作为几何参数化信息的编解码器, 同时针对拓扑树序列使用基于 Transformer 的变分自编码器来获取拓扑关系。为了验证所提出的方法中不同模块的作用和有效性, 进行了以下消融实验, 包括使用 PointNet 替换 PointNet++ 编码点云定义为 HGTR + PointNet; 几何基元与拓扑结构的变分自编码器的方法核心结构 Transformer 替换为 MLP 定义为 HGTR + MLP。将上述组合的方法在 ABC 数据集上进行实验, 其实验结果如表 2 所示。实验结果显示, HGTR + PointNet 相较于 HGTR 方法未提取点云数据中的局部特征和多尺度特征, 使其难以直接获得点云中的细节特征, 导致方法性能有所下降。HGTR + MLP 将 Transformer 为基础结构的编码器换成 MLP 使其能以获得节点间的深层次交互关系, 例如面与环的交互关系等, 使方法重建 CAD 的几何精度和边界细节更差。因此, HGTR 各个模块对于重建 CAD 的几

表 2 ABC 数据集上的消融实验结果

方法	CD ↓	ECD ↓	NC ↑
HGTR + PointNet	0.361	5.612	0.864
HGTR + MLP	0.412	6.069	0.822
HGTR	0.354	4.867	0.934

何信息和拓扑关系都有不同程度的影响,验证了其算法的有效性。

2.4 采样方式对算法性能的影响

为了验证点云采样方式对于算法性能的影响,进行不同采样方式实验,包括非均匀采样与均匀采样,其实验结果如表 3 所示。

表 3 ABC 数据集上不同采样方式的实验结果

采样方式	CD ↓	ECD ↓	NC ↑
非均匀采样	0.375	4.753	0.939
HGTR(均匀采样)	0.354	4.867	0.934

实验结果显示,HGTR 方法在均匀采样上的 CD 表现更优,表明其在整体几何轮廓的拟合上更加稳定,能保证全局结构的一致性。而非均匀采样在 ECD 和 NC 上表现更优,表明方法能重建出更好的边界特征。在数字孪生等关注全局一致性的物联网下游任务中,HGTR 采用均匀采样更适合;在装配等关注局部特征的场景中,采用非均匀采样更适合。

2.5 噪声对算法性能的影响

为了评估 HGTR 方法在输入点云存在噪声情况下的鲁棒性,设计不同噪声比例的数据进行实验验证,具体而言,在原始点云中分别引入 1%,3%,5%,8% 和 10% 的高斯噪声,在 ABC 数据集上进行实验,其实验结果如表 4 所示。

表 4 ABC 数据集上不同噪声比例的实验结果

噪声比例/%	CD ↓	ECD ↓	NC ↑
HGTR	0.354	4.867	0.934
1	0.360	4.961	0.929
3	0.374	5.099	0.919
5	0.385	5.154	0.913
8	0.420	5.506	0.907
10	0.482	6.113	0.896

实验结果表明,随着噪声比例增加,HGTR 方法在 CD 和 ECD 指标均有所上升,而 NC 有所下降,表明噪声对算法的几何精度和法向一致性产生了一定影响。具体而言,HGTR 方法在较低比例(1%~5%)的噪声下,性能下降幅度较小,表明 HGTR 有较强的鲁棒性和复杂噪声环境下的适应能力。

2.6 可视化

为了更加直观地展示 HGTR 方法在重建 CAD 任务中的表现,在 ABC 和 Fusion 360 数据集上进行了可视化分析,结果如图 3 所示,对比了点云、重建

CAD 模型和数据集中真实 CAD 模型。结果表明,所提出的 HGTR 方法能够从稀疏的点云输入中有效还原复杂拓扑结构关系和几何基元参数,所生成的 CAD 模型在拓扑结构和几何形状上与真实模型一致。上述结果表明,HGTR 不仅具备出色的全局几何重建能力,还能够对局部细节实现准确捕捉,从而体现出较强的重建性能,为物联网复杂场景下的 CAD 重建提供有效支撑。

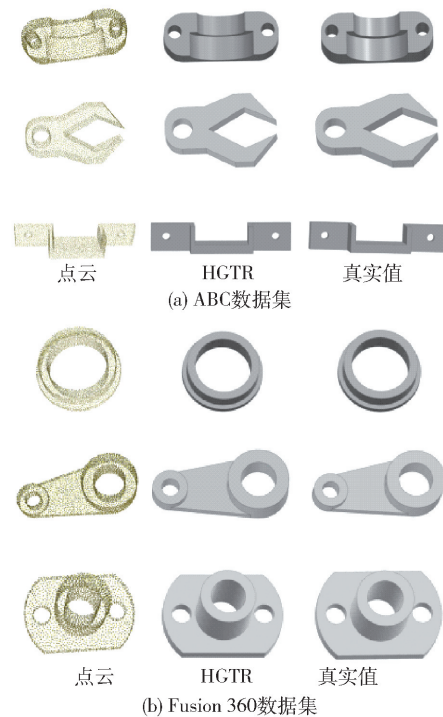


图 3 在 ABC 和 Fusion 360 数据集上 HGTR 方法重建 CAD 可视化结果图

3 结束语

针对物联网驱动的数字孪生与自主导航等复杂三维环境理解需求,提出一种结合 PointNet++ 与 PVAE-TVAE 的点云到 CAD 高精度重建方法。该方法通过协同建模几何基元参数化和拓扑关系,实现复杂物体的几何与拓扑重建,在几何精度和结构完整性上均显著提升。在 2 个公开 CAD 重建数据集上进行系统性评估,在倒角距离、边缘倒角距离和一致性 3 个评价指标上均优于现有方法,验证了所提方法的有效性。但当前方法在几何基元的覆盖范围上仍存在一定局限,尤其对非均匀有理 B 样条(NURBS, non-uniform rational B-splines)曲面等复杂几何的支持不足,未来研究将致力于扩展方法对复杂几何基元重建能力,同时在图编辑距离、流形性

测试等更多指标上,对比更多直接生成 BRep 方法验证该方法的有效性。此外,还将针对更复杂遮挡场景和真实扫描数据开展实验,以提升方法在物联网感知和工程应用中的实用价值。

参考文献:

- [1] KAISER M, BRUSA T, FIECHTER S, et al. Method for parametric evaluation of 3D surface imaging systems for applications with moving objects[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 1-8.
- [2] YANG Y, CUI Z, ZHANG Q, et al. PS6D: Point cloud based symmetry-aware 6D object pose estimation in robot bin-picking[C] // *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2024: 7167-7174.
- [3] WANG F, WU J, YANG Z, et al. Industrial vision inspection using digital twins; Bridging CAD models and realistic scenarios[J]. *Journal of Intelligent Manufacturing*, 2025, 36(7): 4963-4975.
- [4] JIA S, GONG X, LIU F, et al. AI-powered LiDAR point cloud understanding and processing: An updated survey [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(8): 11249-11275.
- [5] MA W, CHEN S, LOU Y, et al. Draw step by step: Reconstructing CAD construction sequences from point clouds via multimodal diffusion[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 27154-27163.
- [6] MA W, XU M, LI X, et al. MultiCAD: Contrastive representation learning for multi-modal 3D computer-aided design models[C] // *Proceedings of the ACM International Conference on Information and Knowledge Management*, 2023: 1766-1776.
- [7] WU R, XIAO C, ZHENG C. DeepCAD: A deep generative network for computer-aided design models[C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 6772-6782.
- [8] LI P, GUO J, ZHANG X, et al. SECAD-Net: Self-supervised CAD reconstruction by learning sketch-extrude operations[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 16816-16826.
- [9] LI P, GUO J, LI H, et al. SfmCAD: Unsupervised CAD reconstruction by learning sketch-based feature modeling operations[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 4671-4680.
- [10] KHAN M S, DUPONT E, ALI S A, et al. CAD-SIG-Net: CAD language inference from point clouds using layer-wise sketch instance guided attention[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 4713-4722.
- [11] GUO H, LIU S, PAN H, et al. ComplexGen: CAD reconstruction by *B*-rep chain complex generation [J]. *ACM Transactions on Graphics*, 2022, 41(4): 1-18.
- [12] LIU Y, XU D, YU X, et al. HoLa: *B*-rep generation using a holistic latent representation [J]. *ACM Transactions on Graphics*, 2025, 44(4): 1-25.
- [13] QI C R, YI L, SU H, et al. PointNet ++: Deep hierarchical feature learning on point sets in a metric space [C] // *Advances in Neural Information Processing Systems*, 2017: 1-10.
- [14] JONES B T, HU M, KODNONGBUA M, et al. Self-supervised representation learning for CAD[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 21327-21336.
- [15] KOCH S, MATVEEV A, JIANG Z, et al. ABC: A big CAD model dataset for geometric deep learning[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 9601-9611.
- [16] WILLIS K D D, PU Y, LUO J, et al. Fusion 360 gallery: A dataset and environment for programmatic CAD construction from human design sequences [J]. *ACM Transactions on Graphics*, 2021, 40(4): 1-24.
- [17] KANIA K, ZIEBA M, KAJDANOWICZ T. UCSG-Net: Unsupervised discovering of constructive solid geometry tree[C] // *Advances in Neural Information Processing Systems*, 2020, 33: 8776-8786.
- [18] REN D, ZHENG J, CAI J, et al. CSG-Stump: A learning friendly CSG-like representation for interpretable shape parsing[C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 12478-12487.
- [19] REN D, ZHENG J, CAI J, et al. ExtrudeNet: Unsupervised inverse sketch-and-extrude for shape parsing [C] // *European Conference on Computer Vision*. Springer, 2022: 482-498.