

面向城市点云的建筑实例分割与结构化重建

吕辰雷 何鑫 谢恒 黄惠

(深圳大学计算机与软件学院可视计算研究中心 广东 深圳 518061)

摘要 作为一类重要的数据资产,大规模城市建筑点云数据具有广泛的应用。然而,城市场景数据规模庞大,需要消耗较大的存储空间,针对数据的模型训练与特征编解码需要较大的计算量;场景缺少语义标注,难以支持高效的数据交互。上述问题制约了城市点云的应用。本文提出一种高效的城市点云建筑实例分割与结构化重建方法 BuildingRecon。该方法通过整合具有语义一致性的初始单体化和基于建筑实例的结构化重建,实现了建筑点云对象的三维单体模型表示。该流程有效减少了对人工标注的需求,降低了数据应用成本,有效提高了效率。点云数据规模得以简化,同时保持可控的几何一致性,保留了必要的拓扑结构,并去除了被噪声点污染的几何细节,进而改善了点云数据的可交互性。实验结果表明,本文方法在点云建筑类别分割标注精度上提升了约35%,在单体实例重建精度上提升了约23%。

关键词 建筑点云;实例分割;结构化重建

中图分类号 TP391

DOI号 10.11897/SP.J.1016.2026.01357

Architectural Instance Segmentation and Structural Reconstruction for Urban Point Clouds

LV Chen-Lei HE Xin XIE Heng HUANG Hui

(Visual Computing Research Center, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong 518061)

Abstract As an important category of data assets, large-scale urban building point cloud data has a wide range of applications. However, urban scene data is massive in scale, requiring substantial storage space. Model training and feature encoding/decoding for such data demand significant computational resources; moreover, the scenes lack semantic annotations, making it difficult to support efficient data interaction. We propose an efficient building instance segmentation and structured reconstruction method, BuildingRecon. By integrating semantically consistent initial individualization and reconstruction based on building instances, it achieves a 3D individual model representation of building point cloud objects. This pipeline effectively reduces the need for annotation, lowers the cost of application, and improves efficiency. The urban point cloud data is simplified while maintaining geometric consistency, preserving necessary topological structures, and removing geometric details contaminated by noisy points, thereby thoroughly enhancing the interactivity of point cloud data. Experimental results show that the proposed method improves the accuracy of point cloud building category segmentation annotation by approximately 35% and

收稿日期:2025-07-03;在线发布日期:2026-02-14。本课题得到国家重点研发计划(2024YFB3908500, 2024YFB3908502)、国家自然科学基金面上项目(62572323)、广东省基础与应用基础研究基金(2023A1515110292, 2023B1515120026)、深圳市高层次人才孔雀团队项目(KQTD20210811090044003)、深圳市科技重大专项(KJZD20240903100022028)、深圳市基础研究专项面上项目(JCYJ20250604181524032)、深圳大学教学改革重点项目(JG2024018)、深圳大学青年教师科研启动经费资助。吕辰雷,博士,讲师,中国计算机学会(CCF)会员,主要研究领域为计算机图形学、几何分析等。E-mail: chenlei.lyu@szu.edu.cn。何鑫,博士研究生,主要研究领域为计算机图形学、几何分析。谢恒,硕士研究生,主要研究领域为人工智能、三维视觉。黄惠(通信作者),博士,教授,中国计算机学会(CCF)杰出会员,主要研究领域为计算机图形学、三维视觉。E-mail: huihuang@szu.edu.cn。

the accuracy of individual instance reconstruction by approximately 23%.

Keywords architectural point clouds; instance segmentation; structural reconstruction

1 引 言

随着三维扫描技术的发展,数字场景重建及其相关应用已成功实现商业化,以满足日常需求。此类应用涉及自动驾驶^[1]、遥感测绘^[2]、建筑设计^[3]等领域。作为一种场景数据,城市场景对应着更广阔的室外区域,并包含多种不同尺度的语义对象。对于城市规划、交通管理^[4]、空间分析等智慧城市管理任务,高质量的数字城市模型,能够提供具有不可替代价值的基础数据底座^[5]。

目前,主流的数字城市场景生成解决方案^[6-7]包括三维激光雷达扫描和多视图图像融合两种技术路线。第一种技术路线直接从城市场景中捕获三维点云数据,能够保留完整的几何结构以支持各种对象的高精度重建。然而,该类方法高度依赖于所选激光雷达设备以及扫描时的环境条件。第二种技术路线利用多视图立体视觉技术从一组图像中重建三维场景,其优势在于数据采集成本显著降低。但该技术路线尚存痛点,如海量数据导致的计算效率下降以及抗噪能力差。

为了实现城市场景的高效三维模型表达,结构化重建方法^[8-9]被提出。此类方法尝试使用一组平面基元来替代复杂的点云或三角网格数据,在保留重要几何特征的同时显著减少了数据量。然而,当前的结构化重建方案仍存在问题:缺乏语义启发以及难以实现基于实例的单体化模型表示。对于

一些涉及场景数据的交互应用,语义对象的单体化计算尤为重要。原始点云数据各类语义对象通常混合在一起,无法准确划分实例数据的精确边界。目前针对该问题,尤其是针对建筑实例的精确单体化重建,主流方法需要相对复杂的特征分析过程,部分技术方案还需要多模态输入进行辅助。这提高了数据处理的计算成本,同时降低了方法的实用性,提高了工程部署难度。因此,目前主流方法在改善城市场景实例数据的可交互性方面存在缺陷。

在本文中,我们提出了一种高效的建筑实例分割与结构化重建方法 BuildingRecon,用于城市点云中的建筑实例建模。该方法基于原始城市点云实现了单体化(从大规模的城市场景建筑群逐步分解到建筑实例对象)与结构化重建,同时保持了语义与几何一致性。其主要包括两个步骤:语义驱动的初始建筑实例自动分割和面向建筑实例的结构化重建。语义驱动的初始建筑实例分割,通过语义分析和基于点的单体分割技术,对原始城市点云中的不同对象进行语义标注,以获得城市建筑实例。该过程在传统神经网络架构的基础上,充分利用空间数据以提高准确性。基于实例的结构化重建则通过建立高质量的多边形网格并实施面片“去粘连”操作(粘连指不同实例或单体对象,在几何上存在的联通关系),进一步提升了建筑实例模型单体边界的准确性。最终,城市点云中的不同建筑实例模型被重建(见图1),兼具良好的几何结构、精确的实例边界以及较低的数据体量。本文主要贡献如下:



图1 面向三维城市点云的建筑实例分割与结构化重建实例(原始点云数据量大,缺少语义标注,难以交互(左图);重建后建筑被精准标注,具有清晰的几何边界以方便交互,同时数据量得到简化(右图))

(1) 本文提出一种针对城市点云的语义驱动初始建筑实例自动分割方法。该方法基于成熟的神经网络架构,通过空间信息嵌入,提高语义分割的精度和鲁棒性。同时,结合点云邻域搜索与空间信息约束,实现城市点云建筑实例的分割,导出对应实际语

义单体对象的点云子集。该方法有效利用了空间数据,显著提升了建筑分割的泛化能力。

(2) 本文提出了一种建筑实例结构化重建方法。该方法利用几何基元提取以及封闭多面体生成,实现对初始建筑实例点云子集的特征结构化表

达。在此基础上,设计基于卷绕数优化的全局定向策略,完成对封闭多面体的定向判定,并驱动一个面片提取步骤。该方法能够重建出简化的多面体网格,同时保持几何一致性。

(3) 本文提出了一种面向结构化模型的去粘连方法。该方法结合初始语义标注和结构化重建全局定向约束,提出粘连面片判定函数,驱动粘连面片的删除,以完成对结构化模型的精准单体重构,获得边界清晰的建筑实例模型。

2 相关工作

现有的解决方案可以归纳为几个部分:面向建筑的结构化重建、建筑 LOD(Level-of-Details 细节层次)建模以及城市场景语义实例重建。

2.1 结构化重建

结构化重建方法基于原始三维点云输入,实现结构化多边形网格重建。Mehra 等^[10]通过使用特征曲线或轮廓作为用于多边形表示的构建块,实现了针对几何建模的抽象表达。Lafarge 等^[11]提出了一种几何结构保持方法,以建立平面组件。Sui 等^[12]设计了一个马尔可夫随机场框架,从三维点云中恢复城市建筑的几何形状。Nan 等^[8]从建筑点云中重建了轻量化的多边形表面。Bauchet 等^[13]设计了一种高效的形状组装机来重建封闭的网格。Fang 等^[14]提出了一种从三维数据中检测平面形状的结构化方法,随后开发了一种混合策略^[15],通过分段连接和切割平面来实现多边形网格重建。Chen 等^[16]提出一种基于学习的建筑点云重建框架,将可学习的隐式表示与显式几何构造相结合。结合语义分析^[17-19]或基于实例的分割框架^[20],这些方法可以间接实现城市建筑的结构化重建。上述方案需要建筑预分割,往往依赖点的法向量辅助,容易受到噪声干扰。本文方法提出的结构化重建,基于面片进行全局定向以导出模型,有效改善了法向量依赖问题。

2.2 建筑 LOD 建模

根据 LOD(细节层次)分级规则^[21-22],一些研究者设计了建筑 LOD 建模方法以实现组件级重建。Tang 等^[23]提出了一种新的上下文感知 LOD 生成方法,用于三维建筑模型。Takikawa 等^[24]利用神经符号距离函数(SDF)实现了 LOD 生成,并通过提出的向量量化自解码器(VQ-AD)方法^[25]以优化三维表示。Bako 等^[26]提出了一个多尺度 LOD 框架,用于

对具有复杂几何和材质的三维环境进行预过滤,同时保留场景的外观。Wysocki 等^[27]提出了一种基于原始点云的 LOD3 建筑重建方法。Schutz 等^[28]提出了一种增量式 LOD 生成方法,能够实时渲染中间三维结果。Zhang 等^[29]提出一种交叉分析方法,以控制 LOD 精细度。Huang 等^[30]建立了一个可语义驱动的部件级 LOD 建模框架。Pan 等^[31]提出一个 LOD 树形部件构建方法来控制细节。上述方案对 LOD 建模的控制需要相对复杂的几何策略,同时对输入数据敏感。相比之下,本文方法利用卷绕数优化,实现了实用的几何全局控制,配合外立面定向辅助,有效改善了结构化重建的实用性。

2.3 城市场景语义实例重建

与组件级别的重建相比,建筑实例重建专注于城市场景内语义对象模型的有效分割与建模。Lin 等^[32]提出了一个系统,用于对三维住宅场景进行语义分解和重建。Li 等^[6]提出了一种全自动的城市场景重建方法,随后作者进一步改善了针对建筑实例的建模精度^[33],用于多视角立体(MVS)建筑。Kelly 等^[7]提出了一种用于城市街区的自动数据融合技术。Niese 等^[34]引入了一种程序化建模方法,用于城市场景中的植被重建。Gui 等^[35]提出了一种基于卫星图像的模型驱动重建方法,用于重建 LOD2 建筑模型。Han 等^[36]设计了一个基于航空影像的大规模城市场景多步骤建模框架。Yang 等^[37]提供了一个大规模城市场景的实例分割和重建基准。Wu 等^[38]提出了一种用于城市场景建筑重建的建筑截面骨架(BBS)方法,该方法包括分段假设与合并。Dong 等^[39]提出一个面向城市场景建筑点云的实例分割与结构化重建方法,可以导出建筑单体模型。一些多模态语义分割方法^[40-42]被广泛用于城市场景语义分割任务。这类方法利用图像数据中的高精度纹理信息提升对语义对象识别精度。但是,这类方法对于图像中的光照敏感,对场景内部几何结构的感知能力不足。

总体而言,上述解决方案主要针对特定的建模对象,需要一些前置条件,如语义标注依赖、模型有效区域标定等,对泛化城市场景建模存在一定的局限性。特别是在基于语义实例的建筑表示与原始输入方面,由于缺少可靠的语义分析和实用的结构化建模方法,已有解决方案的重建性能不够稳定。

3 主体方法

针对已有方案的缺陷,本文提供一个面向城市

点云建筑实例分割和结构化重建的自动化方法 BuildingRecon, 服务于城市场景中的建筑群高效单体重建任务。该方法包含初始建筑实例自动分割和建筑实例结构化重建两个步骤。第一步, 基于已有神经网络结构, 利用空间信息嵌入, 实现语义分割与标注, 同时利用邻域关系探查, 实现初步的建筑实例分割; 第二步, 基于建筑实例点云子集, 提出多边形网格重建方法, 结合去粘连计算, 以输出几何精准、体量简化的建筑实例结构化模型。

3.1 初始建筑实例自动分割

针对城市点云的建筑实例建模, 首先需要获得建筑类别的语义标注。由于原始点云输入中语义对象尺度差异显著且种类繁多(植被、汽车、建筑等), 学习具有泛化能力的语义特征并导出准确的语义标注是一项挑战。本文提出的初始建筑实例自动分割通过两个步骤解决上述问题: 基于体素的空间尺度对齐和基于邻域搜索的建筑实例分割。

基于体素的空间尺度对齐 面向城市点云的语义分割任务, 一个重要的挑战来自输入点云的尺度是未知的, 且差异较大, 即可能是一个街区, 也可能是一个城区。这使得对输入点云的密度优化和特征分析变得极不稳定。传统的点云预处理方法不能解决上述问题。因此, 需要对输入的城市点云进行面向尺度的归一化操作, 来降低尺度差异的影响。为了实现该目标, 本文提出了体素占用率概念, 来驱动一个尺度归一化计算。设 P 表示城市点云, v_p 对应 P 的体素集合。首先, 统计城市场景地面所对应的体素数量 ($|\{v_{base}\}|$), 并将其作为基准。单体素尺度 d 由城市场景中最高的建筑物 d_h 定义, 默认情况下设置 $d_v = d_h$ (默认 $\lambda = 0.005$)。接着, 计算 $|\{v_{base}\}|$ 和 $|\{v_h\}|$ 的乘积 ($|\{v_h\}|$ 为场景纵向(高程)体素的最大数量), 将其视为上述立方体空间的体积。同时, 检查 $\{v_p\}$, 并将包含 P 点集的体素统计到一个 $\{v_p\}$ 的子集 $\{v_o\}$ 中。最后, 计算体素占用率的方法表示为

$$R_p = |\{v_o\}| / (|\{v_{base}\}| \times |\{v_h\}|) \quad (1)$$

其中 R_p 是有效体素占用率。为了实现尺度归一化, 指定了一个标准值 R 。以 R 为优化目标, 迭代地划分城市场景并计算各子区域的相关 R_p 值, 直到 R_p 值的平均值与 R 之间的差异最小化。归一化后, 一个标准分块区域表示为

$$v_N = \operatorname{argmin} |R - \operatorname{avg}\{R_p\}_N| \quad (2)$$

其中, $\{v_N\}$ 是体素的一个子集, 代表一个归一化的子区域, 即 $\{v_N\}$ 根据 v_N 对应整个城市场景的一种划分

方式, $\{R_p\}_N$ 是由 v_N 划分驱动的 R_p 值的集合。每个 v_N 中的子区域被视为语义分析的归一化输入。针对 R 值的优化可被认为是一个体素分割方案的选择过程, 即在一组体素分割方案中, 选择一个最优解。对于小尺度场景, 自然地对应较高的 R 值, 使得不会再对其进行分割。图2展示了一些实例。基于体素的空间尺度对齐, 利用城市点云空间分布的特性, 即通过语义对象纵向高度相对稳定, 通过计算其与地面分布尺度的比值, 以估计输入城市点云的尺度, 导出归一化分块。其优点在于可以不需要语义分析作为辅助。基于标准分块, 我们通过近似内蕴体素简化算法^[43], 为每一个分块指定采样点数, 并移除离群点^[44], 实现归一化驱动的密度优化。

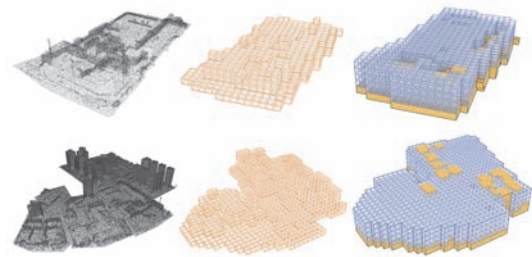


图2 基于体素的城市场景对齐(橙色框表示有效的体素框, 蓝色框表示空体素框。有效的体素占用率在第三列中直接可视化; 第一行: 0.0116, 第二行: 0.0027)

基于点云邻域搜索的建筑实例分割 为了获取建筑语义标注, 我们采用 PointTransformerV3^[45] 提出的序列化注意力网络架构作为主干网络, 实现面向点云的语义分割。为了进一步提升其在城市点云分割任务中的性能, 结合建筑对象的空间分布特性, 本文提出一个高程信息嵌入步骤, 以提升主干网络的分割性能:

$$y_i = \sum_{x_j \in P_v} \rho(\varphi(x_i + h_i)^\top \psi(x_j + h_j) + \delta) \alpha(x_j + h_j) \quad (3)$$

其中, x_i 表示点云中的一点, y_i 表示语义标注, P_v 是一个包含 k 个点的点集 $\{x_i\}$, MLP 映射 ρ , φ , ψ , α 使用 ReLU 激活函数, h_i 和 h_j 为对应点的高程信息, 即基于水平面的 z 坐标数值, 与输入特征结合后输入到注意力层, 并加入位置编码 δ 。这种设计能够显著改善主干网络的空间感知能力, 结合前面的归一化预处理, 能够提升主干网络语义分割的泛化能力, 图3展示了一个对比实例。

基于建筑语义标注, 城市点云中的建筑对象能够被识别。为了获得更精确的建筑实例, 本文采用 k -近邻(KNN)搜索方法来实现初始实例分割。首先, 随机

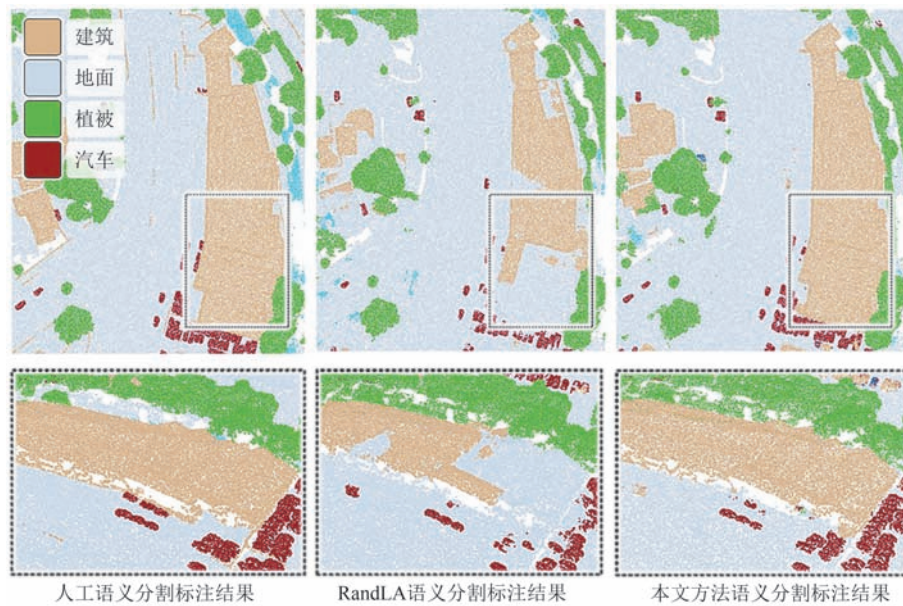


图3 基于低质量点云的建筑语义分割(左图:人工语义标注结果;中图:城市点云分割主流方法RandLA-Net结果;右图:本文方法的结果)

选择一个带有建筑物标签的点作为起始位置。然后,使用KNN获取其具有相同标签的相邻点(根据各向同性假设,理想的点邻域关系指内部的顶点与所有邻居点的距离相等,导出 $k=6$)。每个相邻点被设置为新的起始位置。基于迭代搜索,建立一个具有相同语义标签的点邻接图,直到没有新的起始位置被添加。生成的图可以被视为一个粗略的分段。如果仍然存在带有建筑物标签的点,重复搜索过程,直到所有点都被检查完毕。最终,建筑物实例对应分割后的离散点集,即实现面向城市点云的初始建筑实例自动分割。

3.2 建筑实例结构化重建

基于已分割的建筑实例点集,本文提出一种结构化重建方法,以生成建筑实例多边形网格模型。该方法包含三个部分:凸多面体生成、基于多面体的定向以及基于平面的去粘连,如图4所示。凸多面体生成从实例点集中构建一组封闭的子结构,这些子结构提供基于平面的初始化,同时修复缺失部分。多面体定向用以驱动对封闭多面体的保留与删除判断,并输出具有全局一致性的多边形网格模型。去粘连计算检测和删除网格模型内部错误的相邻平面,来提高建筑实例模型的单体边界精度。

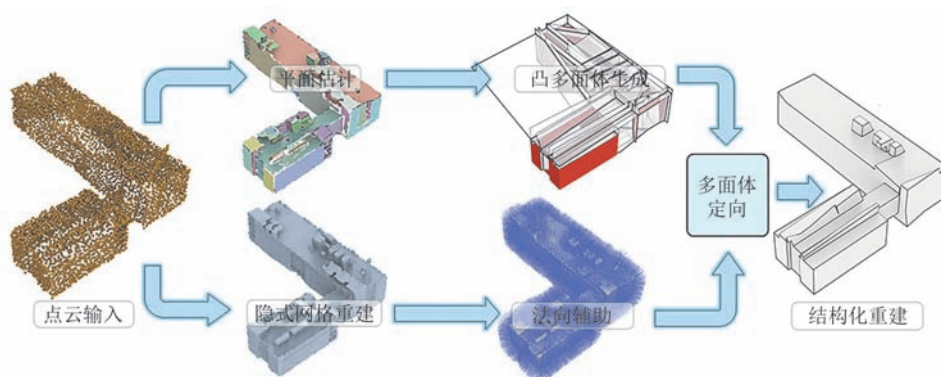


图4 结构化重建的流水线(以分割的建筑点云为输入,通过平面估计获得结构化面片表达,利用面片求交生成凸多面体集合。同时,利用隐式网格估计,实现法向量辅助计算。最终,通过多面体定向,从凸多面体集合导出建筑结构化模型)

凸多面体生成 凸多面体用于表示基于点的建筑物实例基本几何结构,由原始平面提取后^[46]通过外立面定界与扩展平面交叉形成的包围结构组成。

外平面的检测基于以下规则:所有其他平面应位于外平面的同一侧。设 F_0 是在进行平面求交计算时的第一个平面,平面 F_i 为 F_0 的交叉平面,计算平面

对应的 α 形状(点集凸包的推广^[47])以获得 F_0 的边界点集 $\{p\}_b^0$ 。检查 $\{p\}_b^0$ 中的点,并删除满足 $\text{dist}(p_j^0, F_{-i}) < \sigma$ 的点,其中 dist 是 p_j^0 与 F_i 之间的距离, $p_j^0 \in \{p\}_b^0$ 。如果 $\{p\}_b^0$ 中剩余的边界点都位于 F_i 的一侧,则判断整个平面 F_0 位于 F_i 的一侧。该计算过程提高了外平面检测的实用性。在凸多面体空间中,可以感知到上述内部结构信息。随后,采用穷举分割^[8],在外立面包围的区域,执行基于平面扩展的对凸多面体空间求交分割。每个内部平面被扩展并与外立面相交,以导出对应的凸多面体,如图5所示。

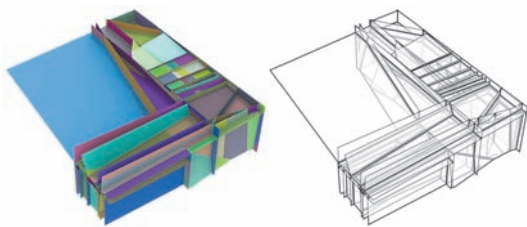


图5 由相交平面生成的凸多面体空间实例

为了确保凸多面体在简化数量的同时更好地对应主体结构,本文引入了一个多面体聚类步骤,该步骤已涵盖在算法1中。对于内部面片集合 $\{F\}_m$ 的搜索,选择面片相交次数最小的一个,进行临近面片聚合。如果有多个平面具有相同的最小值,则选择 α 面积最大的一个进行聚合。聚类过程从凸多面体空间的外部向内部面片搜索,来迭代地进行面片聚合。通过迭代操作,基于一组多面体的聚合结构被生成,同时感知了内部区域,如图6所示。

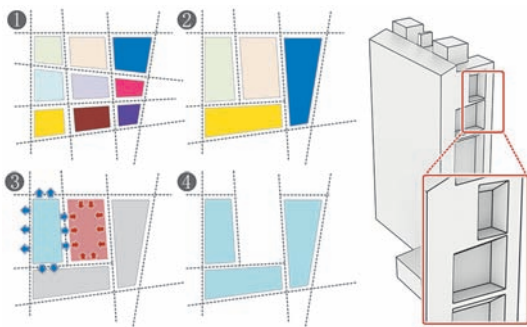


图6 基于多面体定向的一个实例(左图:1是多面体截面示意图;2是多面体聚合结果;3是多面体定向;4是多边形导出;右图:基于多面体定向的结构化模型)

算法1. 多面体生成

输入:内部面片集合 $\{F\}_m$

输出:聚合凸多面体集合 $\{CP\}_c$

1. $\{\text{inc}\}_m$ 收到 \leftarrow 面片相交数计算 $\{F\}_m$
2. while 判断($\{F\}_m$)是否为空

3. 搜索 F_a ,具有最小的inc值
4. 获得相关的子空间 $\{CP\}_a$
5. if($\{CP\}_c$ 包含 $F_j \in \{F\}_m$)
6. 对 F_j 迭代聚类子空间
7. else
8. 合并 $\{CP\}_a$ 到 $\{CP\}_c$
9. 删除 F_a 从 $\{F\}_m$
10. end

基于多面体的定向 经过凸多面体生成,散乱的建筑实例点云子集被表示为封闭的多面体。在此基础上,需要界定多面体的内外定向,完成结构感知,以驱动后续的多边形面片建模。为了提高结构感知的准确性,首先重建一个辅助结构作为定向参考。这里采用IPSR^[48]生成隐式曲面参考结构。隐式曲面对应一个泊松三角网格表示 $M=(V, N)$,其中 V 表示顶点集, N 表示相关的法向量集。借助该参考结构,所有多面体对应的平面可以被初步定向(多面体平面的法线方向应与 M 保持一致)。

基于初始定向的平面,本文利用卷绕数优化来实现具有全局一致性的平面定向。卷绕数优化已广泛用于法线估计和表面重建^[49],其公式化表示:

$$w(q) = \sum_{i=1}^N a_i \frac{(p_i - q) \cdot n_i}{4\pi \|p_i - q\|^3} \quad (4)$$

其中, q 表示多面体的质心, p_i 表示相关多面体的平面 F_i 的中心, a_i 是 F_i 的面积, n_i 是基于平面的法向量,已通过 M 进行了初始定向。上述公式已完成了卷绕数参数的重定义。参数 $w(q)$ 的值表示内部或外部。基于该公式,可以计算所有多面体的卷绕数,进而引导一个迭代定向过程,具体步骤如下:迭代过程中,通过公式(4)为多面体计算卷绕数并标记内外(根据阈值 $\epsilon = 0.5, w(q) > \epsilon: \text{in}$,将其标记为内部);删除标记为外部且位于边界的多面体,然后更新面片法向,进入下一次迭代;所有多面体均标记为内部,结束循环。迭代定向过程如算法2所示。对于最终保留的多面体,其外部轮廓即为重建的多边形网格。得益于质心集的精确确定值,卷绕数优化可以更准确地检查和处理内部结构,实现高效的定向计算。

算法2. 迭代定向过程

输入:多面体集合 $\{CP\}$

输出:内部多面体集合 $\{CP\}_m$

1. while 检查 $\{CP\}$ 中是否存在未标记为内部的多面体
2. 根据公式(4)对每一个多面体 CP_i 进行内外标记
3. if(CP_i 被标记为外部且位于边界)
4. 从 $\{CP\}$ 删除 CP_i
5. else

6. 更新剩余多面体对应的的面片法向
7. end

优化中涉及的质心数量对应于多面体的数量,远少于原始点云,这增强了结构重建的实用性。当确定多面体内外定向后,使用PolyFit^[8]的多边形网格导出,以完成结构化建模。如图6所示,可以看到,通过本文的卷绕数优化方法,更多的细节被保存。IPSR提供的法向,在公式(4)中仅作为一种对外立面的初始化,以提供二值定向参考(判定里或外),不依赖点的法向量,提高了鲁棒性和实用性。

基于平面的去粘连 多边形网格表示的建筑实例,对应了语义标注和结构信息。然而,受限于语义标注的随机误差以及在重建中产生的一些错误拓扑结构,使得建筑实例间存在粘连结构,破坏实例的语义一致性。为了解决这一缺陷,本文提出了一种基于平面的去粘连计算以提高准确性。在重建过程中,位于建筑实例底部的一些结构平面被判断为非流形表面。根据卷绕数优化规则,这些平面对建筑物实例产生了负面影响。去粘连计算检查这些平面并将其从多边形网格中移除。定义 $\{F\}_s$ 为结构化重建后组成多边形网格的面片集合,面片法向指向网格外部,首先检测 $\{F\}_s$ 中是否存在粘连面片,若 $F_i \in \{F\}_s$ 为粘连面片,则需满足:

$$\begin{aligned} \text{dist}(F_i, P_G) < \sigma, \\ \text{acos}\langle n_f, n_G \rangle < \theta_{\text{th}}, \end{aligned} \quad (5)$$

其中 P_G 表示地面, $\text{dist}(F_i, P_G)$ 表示面片 F_i 的中心点到地面的欧式距离, n_f 和 n_G 分别表示 F_i 和 P_G 的法向量, $\text{acos}\langle n_f, n_G \rangle$ 表示两个法向的夹角, $\sigma \in (0, 1)$ 与 θ 为阈值(默认 $\theta_{\text{th}}=30^\circ$)。公式(5)表明,当面片 F_i 距离地面较近且与地面夹角在一定范围之内,该面片可被判定为粘连面片。定义 $\text{Ind}(F_i)$ 为判定函数,若 F_i 为粘连面片, $\text{Ind}(F_i)=-1$,否则 $\text{Ind}(F_i)=1$ 。基于以上定义,根据公式(4)对已经标记为内部的多面体进行二次判断,并对公式(4)中变量进行重定义,其中 q 为标记为内部多面体的质心, p_i, a_i, n_i 则分别为面片 $F_i \in \{F\}_s$ 的中心、面积、法向量,若 $\text{Ind}(F_i)=-1$,则将法向翻转,即 $n_i = \text{Ind}(F_i) \cdot n_i$,然后删除 $w(q) < \epsilon$ 的多面体及相关面片,即可使粘连的建筑实例在几何拓扑层面分离。

该实现同时考虑了几何和语义分析,在建筑实例语义标注的辅助下,实现粘连平面判定,从而获得更准确的建筑实例模型,如图7所示。本文所提出的BuildingRecon流程如算法3所示。得益于语义

一致的实例结构化重建和平面去粘连, BuildingRecon能够为建筑实例构建更简洁高效的三维表示,兼具更好的几何和语义一致性。实验部分将展示BuildingRecon的性能优势。

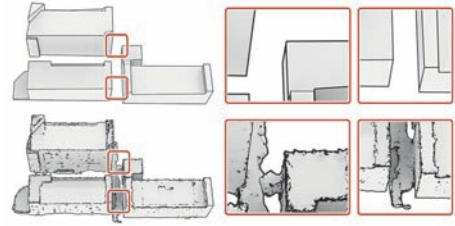


图7 针对相邻建筑实例(基于点的单体化(第一行)会携带错误的连接面,去粘连(第二行)能够获得精准边界)

算法3. BuildingRecon算法流程

输入:城市点云 P

输出:建筑实例模型集合 $\{M\}$

1. 对 P 执行基于体素的空间尺度对齐
2. 使用语义分析建筑点云子集 P_b
3. 使用建筑实例分割获得建筑实例点集 $\{P_i\}_b$
4. while 检查 $(\{P_i\}_b)$ 是否为空
5. 对 P_i 执行多边形生成
6. 多面体的定向生成多边形网格 M_i
7. 对 M_i 执行去粘连计算
8. 从 $\{P_i\}_b$ 删除 P_i
9. 添加 M_i 到 $\{M\}$
10. end

4 实验结果与分析

本文在实际任务中评估了所提BuildingRecon方法的性能。实验机器配置为Intel i9 13900K、128 GB内存、RTX4090,操作系统为Windows11,开发平台为VisualStudio2022。测试点云数据来自SensatUrban^[50]、UrbanBIS^[37]和BuildingNet^[51]。实验包括以下部分:(1)首先介绍选定的测试数据集,并解释了用于城市重建质量评估的相关度量指标;(2)比较不同重建方法的可视化结果和指标,以展示本文所提BuildingRecon方法的优势,包括语义一致性、几何一致性和单体化准确性;(3)对BuildingRecon方法进行性能分析,展示其在实际应用中的样例,并讨论了局限性。

4.1 数据集与度量

实验选择SensatUrban, UrbanBIS以及BuildingNet作为测试数据集。SensatUrban是一个城市规模的点云数据集,数据来自英国三个城市的一些城

区,覆盖7.6平方千米,密度约为480点/m²。该数据集具有13个类别的语义标注。在此数据场景中,建筑物的几何结构相对简单,并且缺乏高层建筑的模型。UrbanBIS是一个新的城市规模数据集,包含六个真实的城市场景。它具有7个类别的语义标签,覆盖10.78平方千米,密度约为250点/m²,六个场景合计包括3370实例建筑数据。为了进行通用的语义分析,本文基于上述两个数据集的语义标注,定义了一组新的语义标注,包括地面、植被、水体、建筑物、汽车和桥梁。本文专注于建筑物类别的识别,其他标签用于辅助该目标。BuildingNet包含1000+带有点云和网格的建筑模型,可用于测试建筑对象结构化重建的实用性。

为了评估建筑实例语义分割与结构化重建方法的质量,本文引入了三个指标,即语义一致性、几何一致性和建筑实例分割的准确性。对于语义一致性,实验直接测量基于点的语义标注准确性以建立定量分析。量化指标包括总体准确率(OA)、平均交并比(mIoU)和建筑物交并比(Building IoU)。实验使用UrbanBIS数据集训练深度网络,并在两个独立的测试数据集上评估模型,以展示其泛化能力。为了衡量几何一致性,实验报告基于点对表面映射计算豪斯多夫距离和平均距离。这些距离明确表示了重建模型与原始点云之间的差异。同时,实验还考虑了重建对数据的简化效率。为了展示建筑实例边界界定的准确性,实验可视化了基于实例的分割结果,展示了组件级别的细节,特别是建筑物底部的区域。

4.2 方法比较

语义一致性面向城市点云的语义分割是一项具有挑战性的任务。城市点云内不同语义对象类别与尺度的差异、点云数据质量的不确定性、点云覆盖范围未知以及非均匀分布等因素增加了语义分割的难度。为了验证本文所提BuildingRecon面向城市点云语义分割的性能,基于数据集建立量化分析,对比方法包括RandLA-Net^[18], KPConv^[52], PointTransformer^[53], OA-CNNs^[54]以及PointTransformerV3^[45]。上述方法已经覆盖主流的城市点云语义分割解决方案。表1报告了语义分割定量分析的结果。本文方法在两个库的场景分割任务中,均获得最优性能。相比之前的主流方法RandLA-Net,本文方案在跨库建筑对象分类精度上,提升了35%。值得一提的是,本文的方法与PointTransformerV3使用相同的主干网络,在对SensatUrban实验数据的交

叉验证中,本文的方法具备明显的优势。其原因在于基于体素的空间尺度对齐以及高程信息嵌入,改善了主干网络在特征训练时的空间感知能力。

表1 城市场景中不同语义分割方法的定量分析

数据集	对比方法	OA	mIoU	Building
UrbanBIS	RandLA-Net	85.49%	53.61%	81.39%
	KPConv	87.68%	54.68%	88.76%
	PointTransformer	89.12%	55.03%	89.62%
	OA-CNNs	91.20%	54.88%	91.97%
	PointTransformerV3	91.22%	58.39%	92.12%
	本文方法	92.03%	60.12%	93.03%
SensatUrban	RandLA-Net	72.12%	36.98%	59.03%
	KPConv	75.68%	34.73%	65.54%
	PointTransformer	77.92%	35.37%	66.87%
	OA-CNNs	71.23%	29.13%	65.76%
	PointTransformerV3	83.11%	40.04%	77.78%
	本文方法	86.12%	42.03%	80.06%

注:本文方法的结果通过初始语义分析获得。即使使用相同的主干网络,本文方法仍提升了性能。

几何一致性 结构化重建应在生成简洁多边形网格的同时,保持与原始点云的几何一致性。为了评估重建的性能,本文将几种主流的结构化重建方法与BuildingRecon进行了比较,包括PolyFit^[8]、KSR^[13]、LowPoly^[55]和RobustLowPoly^[56]。参考LowPoly,从BuildingNet中收集了200个点云进行测试。表2报告了定量结果,包含对上述方法特性的说明。PolyFit方法对某些建筑实例点云无法收敛,原因是平面提取和多面体生成阶段,没有对细碎面片进行控制,缺少高效的全局定向辅助。这降低了其实用性。KSR方法对估计的法向量敏感。LowPoly方法和RobustLowPoly方法需要初始网格来提供基本几何结构,其重建结果缺乏结构一致性。相比之下,BuildingRecon提供了一种更实用的结构化重建方案。尽管使用了基于面的法向量辅助来提高基于多面体定向的准确性和效率,但其不需

表2 不同结构化重建方法的定量分析

对比方法	Dis ^H ↓	Dis ^M ↓	p ^{Avg} ↓	\bar{F} ↓	R ^C ↑	M ^D	N ^D
PolyFit	0.1597	0.0116	556	81.39%	72%	是	否
KSR	0.2637	0.0261	296	88.76%	100%	是	否
LowPoly	0.0948	0.0139	354	89.62%	100%	否	否
R-LowPoly	0.1178	0.0108	316	91.97%	100%	否	否
本文方法	0.0904	0.0120	486	80.06%	100%	是	是

注:指标:豪斯多夫距离Dis^H、平均距离Dis^M、平均点数p^{Avg}、平均面数 \bar{F} 、重建成功率R^C、初始网格独立性M^D和法线独立性N^D。

要在全局定向中使用基于点的法向量来生成多面体。这确保了重建过程中,该方法具备一定程度的

噪声鲁棒性。图8展示了BuildingRecon能够输出更准确的几何细节和更好的结构信息。

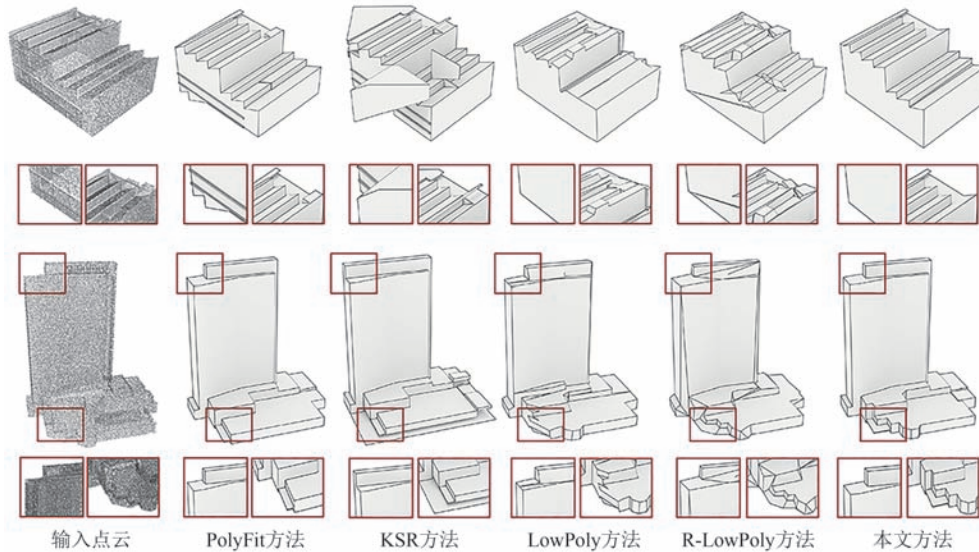


图8 BuildingNet模型中不同结构化重建方法的比较

建筑实例单体化精度 本文提出的基于平面的去粘连计算为解决错误粘连问题提供了解决方案。如图10所示,去粘连计算优化了实例模型的有效边界。图9展示了更直观的对比实例。BuildingRecon针对建筑群建立了更准确的实例边界。表3报告了整体化准确性的定量分析结果,精度平均提升

23%。BuildingRecon生成的整体化模型与真实情况更加一致,弥补了语义分析的不足,提升了整体化的实用性。在图11和12中,本文展示了更多城市建筑集合中的整体化实例。一些相对复杂的建筑结构也能够以精确的边界进行重建,这对3D交互的应用场景具有重要价值。

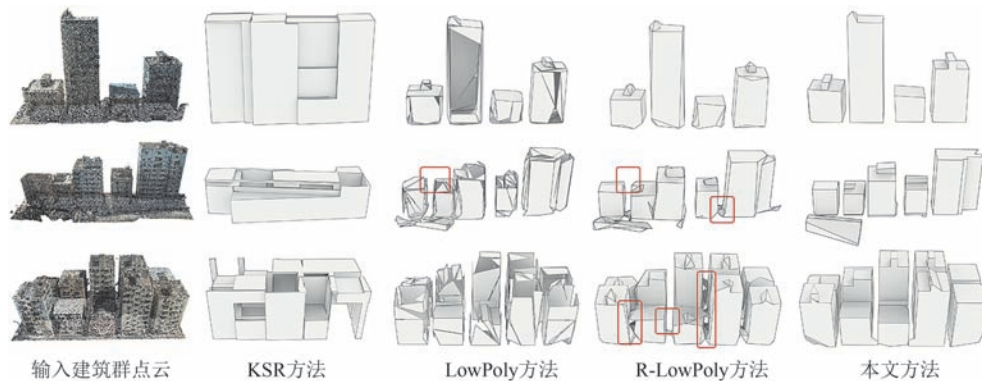


图9 UrbanBIS建筑群中不同结构化重建方法的比较(一些错误粘连的区域用红色线框标记)

表3 UrbanBIS场景实例分割成功百分比

对比方法	Loinghua	Yingrenshi	Wuhu	Yuehai
PolyFit	13%	17%	15%	18%
KSR	53%	51%	54%	53%
LowPoly	50%	55%	61%	57%
R-LowPoly	16%	37%	28%	31%
本文方法	64%	68%	71%	76%

4.3 性能分析

计算性能 本文算法的语义解析过程近似于PointTransformerV3,仅增加高程信息嵌入。算法的核心部分结构化建模,面片抽取计算近似于FPP算法^[46],主要集中在KD-Tree搜索,时间复杂度大致为 $O(n \log n)$,空间复杂度大致为 $O(nk)$, k 为指定的邻点数量,默认为6;凸多面体生成部分,时间复杂度与空间复杂度仅与几何基元数 p 有关,大致为

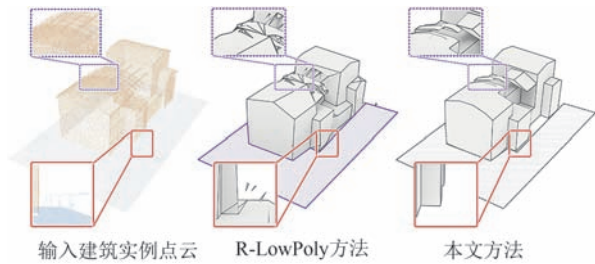


图10 建筑物底部的去粘连实例(本文方法在建筑物周围实现了更清晰简洁的几何(红色实线框)。对于复杂镂空结构(紫色虚线框),本文方法的重构结果存在一定的缺陷。)

$O(p^3)$; 卷绕数优化阶段,时间复杂度 $O(qp)$ 和空间复杂度 $O(q+p)$ 仅与几何基元数 p 和多面体数 q 有关。这里 p 和 q 远远小于原始点云的点数 n 。综上,核心部分的时间复杂度约为 $O(n \log n)$,空间复杂度约为 $O(nk)$ 。

本文比较了不同方法在语义分割和结构化重建任务的时间开销,如表4所示。由于共享相同的主干网络,BuildingRecon的时间开销(每十万点计算时间)与PointTransformerV3相近。受益于体素归一化和空间信息嵌入,本文方法的性能具有一定的优势。对于结构化重建任务,本文方法的计算效率低于KSR,原因在于全局一致优化以及IPSR辅助网格重建,需要额外的计算开销。但是,相比KSR,本文

方法仅需面法向定向辅助,无需点法向,在几何结构的一致性控制方面,具有一定的优势。

表4 语义分割与结构化重建时间对比

方法	建筑实例语义分割			建筑实例结构化重建		
	RandLA	PTV3	BRecon	PolyFit	KSR	BRecon
参数	4.99 M	46.24 M	46.43 M	—	—	—
时间	4.661 s	3.046 s	3.103 s	3127.302 s	7.216 s	87.589 s

注: RandLA: RandLA-Net; PTV3: PointTransformerV3; BRecon: 本文方法。

消融实验 这里主要对基于体素的空间尺度对齐、高程信息嵌入以及结构化重建参数几个部分进行分析,以展示本文方法对于实例分割与建模性能的提升。基于体素的空间尺度对齐能够驱动一个更合理的密度优化计算,相比于传统的随机下采样或K近邻均匀采样,对于点云分布优化的性能提升更明显。高程信息嵌入能够显著地改善对建筑对象的识别精度,因为建筑对象的高度与其他语义对象存在显著的差异。表5展示了对于两个不同的主干网络,在基于SensatUrban的交叉测试中,结合不同模块所带来的性能差异。可以看到,使用基于体素的空间尺度对齐和高程信息嵌入能够提升识别精度。在进行空间尺度对齐时,两个参数对算法会产生影响,包括体素尺度和 k 近邻保留点数。

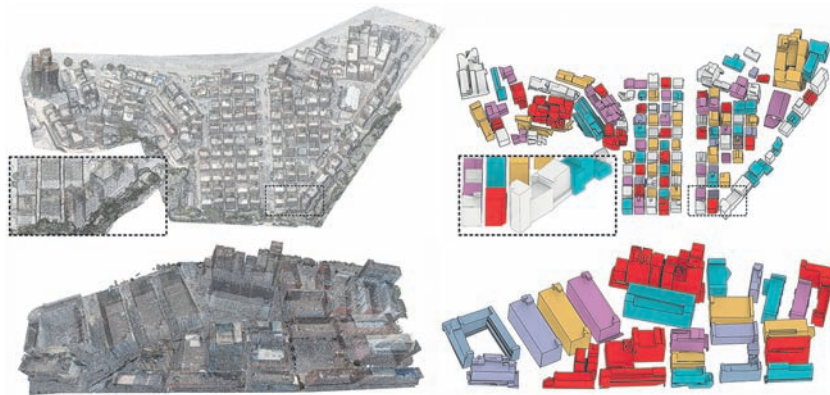


图11 UrbanBIS场景建筑实例分割与结构化重建实例

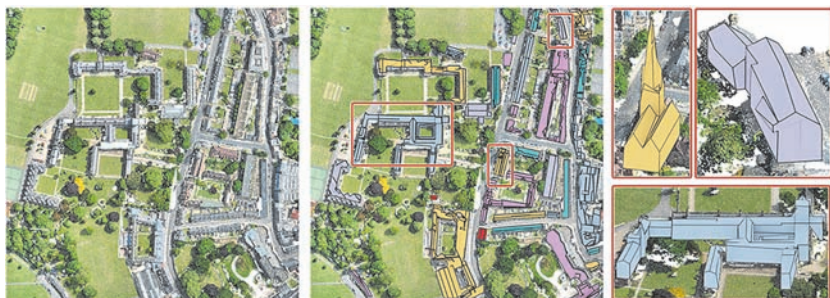


图12 SensatUrban场景建筑实例分割与结构化重建实例

表6对比了不同参数配置下对UrbanBIS数据库语义分割性能的影响。本文的默认参数选择基于最优配置。

表5 组合不同主干网络、基于体素的空间尺度对齐(空间对齐)和高程信息的分割性能对比(RandLA: RandLA-Net; PTV3: PointTransformerV3)

骨干网络	空间对齐	高程信息	OA	Building
RandLA	+	-	80%	62%
	-	+	75%	74%
	+	+	81%	75%
PTV3	+	-	84%	77%
	-	+	83%	79%
	+	+	86%	80%

表6 基于不同参数组合的语义分割性能比较

λ 体素参数	K近邻	OA	mIoU	Building
$\lambda=0.01$	$k=4$	90.68%	57.16%	91.26%
$\lambda=0.005$	$k=6$	92.03%	60.12%	93.03%
$\lambda=0.001$	$k=8$	91.85%	58.63%	92.13%

对于结构化重建,参数 ϵ 在卷绕数优化中,决定了多面体的内部或外部朝向,直接影响最终的重建结果。图13则展示了不同的 ϵ 值对重建结果的影响,可以看出,若 ϵ 较小,重建的网格存在冗余面片;若 ϵ 较大,模型存在明显缺失,本文的默认参数 $\epsilon=0.5$ 为最优选择。图14展示了部分建筑的结构化重建过程,从左至右分别为:输入的建筑点云;从点云中提取的平面结果;平面求交生成凸多面体;基于IPSR^[48]的隐式曲面重建及法向量计算结果;以及结构化重建结果。可以看出,本文重建结果与点云的几何信息一致,结合法向辅助,能够保证全局定向的准确度。

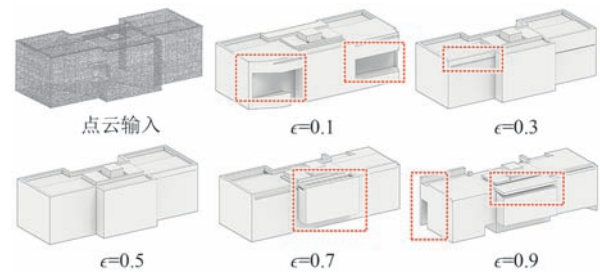


图13 不同 ϵ 取值对结构化重建细节的影响

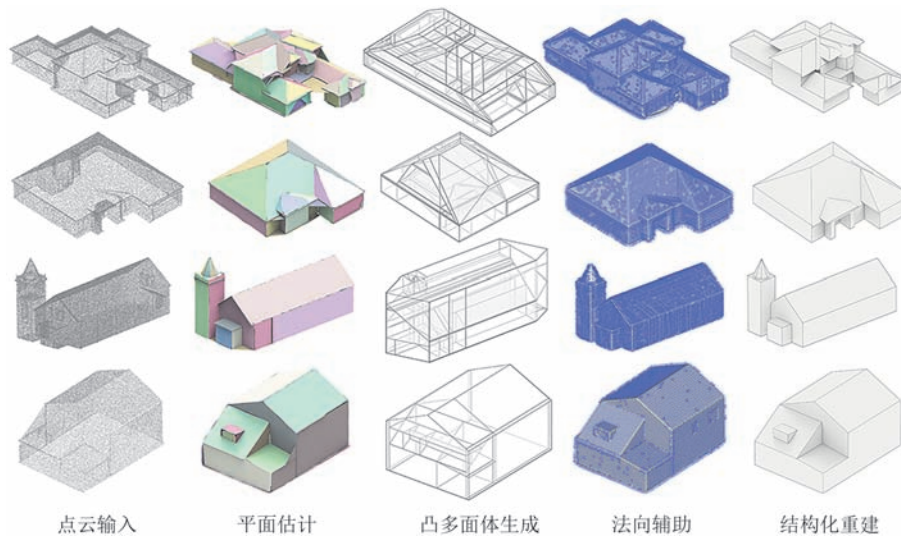


图14 单体建筑的结构化重建过程

局限性 本文方法存在两种依赖关系。第一,基于体素的空间尺度对齐和语义分析的性能取决于建筑物的分布情况。一旦城市场景中像SensatUrban样本那样缺少高层建筑,则分割性能会产生退化,如表1所示。第二,结构分析的准确性依赖于平面检测的精度。如图10所示,对于细碎平面进行删除时,一些有价值的几何细节(建筑顶部镂空结构)也被破坏。尤其对于局部的细碎平面以及曲线变化显著的几何结构,本文方法在对结构信息的一致性控

制方面存在一定的性能退化。

5 结论

本文提出了一种建筑实例分割与结构化重建方法BuildingRecon,用于三维城市建筑实例自动建模。该方法从原始点云输入中提取语义标注,实现自动的建筑实例分割。同时通过平面基元提取以及多面体生成,以支持空间感知。得益于卷绕数优化,

本文实现了多面体的定向,驱动多边形网格重建。结合去粘连计算,BuildingRecon实现了语义和结构分析的一体化策略,完成精确的建筑实例单体建模,有效提升了城市场景建筑建模的实用性。实验表明,BuildingRecon改善了城市点云建筑实例建模计算的整体实用性。

参 考 文 献

- [1] KONG De-Ming, LI Xiao-Wei, YANG Qing-Xin. Multimodal 3D object detection method based on pseudo point cloud feature enhancement. *Chinese Journal of Computers*, 2024, 47(4):759-775 (in Chinese)
(孔德明, 李晓伟, 杨庆鑫. 基于伪点云特征增强的多模态三维目标检测方法. *计算机学报*, 2024, 47(4):759-775)
- [2] XIE Xing-Xing, CHENG Gong, YAO Yan-Qing, et al. Dynamic feature fusion for object detection in remote sensing images. *Chinese Journal of Computers*, 2022, 45(4): 735-747 (in Chinese)
(谢星星, 程琰, 姚艳清, 等. 动态特征融合的遥感图像目标检测. *计算机学报*, 2022, 45(4): 735-747)
- [3] WEN Gong-Jian, WANG Run-Sheng, LUO Yan-Ni. An approach to measuring a city building based on its structure information. *Chinese Journal of Computers*, 2000, 23(11): 1201-1207 (in Chinese)
(文贡坚, 王润生, 罗艳妮. 基于结构信息的城市建筑物测量方法. *计算机学报*, 2000, 23(11): 1201-1207)
- [4] MAO Tian-Lu, WANG Hua, KANG Xing-Chen, et al. Simulating large-scale traffic in complex networks. *Chinese Journal of Computers*, 2017, 40(11): 2466-2477 (in Chinese)
(毛天露, 王华, 康星辰, 等. 复杂路网内大规模车辆运动的仿真. *计算机学报*, 2017, 40(11): 2466-2477)
- [5] SHAN Jie, LI Zhixin, ZHANG Wenyuan, et al. Recent progress in large-scale 3D city modeling. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(12): 1523-1541 (in Chinese)
(单杰, 李志鑫, 张文元. 大规模三维城市建模进展. *测绘学报*, 2019, 48(12): 1523-1541)
- [6] LI M, WONKA P, NAN L. Manhattan-world urban reconstruction from point clouds//*Proceedings of the European Conference on Computer Vision*. Amsterdam, The Netherlands, 2016: 54-69
- [7] KELLY T, FEMIANI J, WONKA P, et al. Bigsur: Large-scale structured urban reconstruction. *ACM Transactions on Graphics*, 2017, 36(6): 204:1-204:16
- [8] NAN L, WONKA P. PolyFit: Polygonal surface reconstruction from point clouds//*Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy, 2017: 2353-2361
- [9] HE X, LV C, HUANG P, et al. WindPoly: Polygonal mesh reconstruction via winding numbers//*Proceedings of the European Conference on Computer Vision*. Milan, Italy, 2024: 294-311
- [10] MEHRA R, ZHOU Q, LONG J, et al. Abstraction of man-made shapes. *ACM Transactions on Graphics*, 2009, 28(5): 10: 1-10:10
- [11] LAFARGE F, ALLIEZ P. Surface reconstruction through point set structuring. *Computer Graphics Forum*, 2013, 32: 225-234
- [12] SUI W, WANG L, FAN B, et al. Layer-wise floorplan extraction for automatic urban building reconstruction. *IEEE Transactions Visualization and Computer Graphics*, 2015, 22(3): 1261-1277
- [13] BAUCHET J P, LAFARGE F. Kinetic shape reconstruction. *ACM Transactions on Graphics*, 2020, 39(5): 156:1-156:14
- [14] FANG H, LAFARGE F, DESBRUN M. Planar shape detection at structural scales//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA, 2018: 2965-2973
- [15] FANG H, LAFARGE F. Connect-and-slice: An hybrid approach for reconstructing 3d object//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Seattle, USA 2020: 13490-13498
- [16] CHEN Z, LEDOUX H, KHADEMI S, et al. Reconstructing compact building models from point clouds using deep implicit fields. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022(194): 58-73
- [17] LANDRIEU L, SIMONOVSKY M. Large-scale point cloud semantic segmentation with superpoint graphs//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA, 2018: 4558-4567
- [18] HU Q, YANG B, XIE L, et al. Randla-net: Efficient semantic segmentation of large-scale point clouds//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Seattle, USA, 2020: 11108-11117
- [19] LI Hai-Sheng, WU Yu-Juan, ZHENG Yan-Ping, et al. A survey of 3D data analysis and understanding based on deep learning. *Chinese Journal of Computers*, 2020, 43(1): 41-63 (in Chinese)
(李海生, 武玉娟, 郑艳萍, 等. 基于深度学习的三维数据分析理解方法研究综述. *计算机学报*, 2020, 43(1): 41-63)
- [20] HAN X, LIU C, ZHOU Y, et al. Whu-urban3d: An urban scene lidar point cloud dataset for semantic instance segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2024, 209: 500-513
- [21] VERDIE Y, LAFARGE F, ALLIEZ P. Lod generation for urban scenes. *ACM Transactions on Graphics*, 2015, 34(30): 1-14
- [22] BILJECKI F, LEDOUX H, STOTER J. An improved lod specification for 3d building models. *Computers, Environment and Urban Systems*, 2016, 59: 25-37
- [23] TANG L, YING S, LI L, et al. An application-driven lod modeling paradigm for 3d building models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 161: 194-207
- [24] TAKIKAWA T, LITALIEN J, YIN K, et al. Neural geometric level of detail: Real-time rendering with implicit 3D shapes//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Online, 2021: 11358-11367
- [25] TAKIKAWA T, EVANS A, TREMBLAY J, et al. Variable bitrate neural fields//*Proceedings of the SIGGRAPH*. New

- York, USA 2022: 41:1-41:9
- [26] BAKO S, SEN P, KAPLANYAN A. Deep appearance prefiltering. *ACM Transactions on Graphics*, 2023, 42: 2: 1-2:23
- [27] WYSOCKI O, XIA Y, WYSOCKI M, et al. Scan2lod3: Reconstructing semantic 3d building models at lod3 using ray casting and bayesian networks//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada, 2023: 6548-6558
- [28] UTZ MCH, HERZBERGER L, WIMMER M. SimLod: simultaneous lod generation and rendering//*Proceedings of the SIGGRAPH*. Los Angeles, USA, 2023, 17:1-17:20
- [29] ZHANG R, PAN S, LV C, et al. Architectural co-lod generation. *ACM Transactions on Graphics*, 2024, 43(6): 193: 1-193:16
- [30] HUANG J, WANG C, LI L, et al. Buildingblock: A hybrid approach for structured building generation//*Proceedings of the SIGGRAPH*. Vancouver, BC, Canada, 2025, 173:1-173:11
- [31] PAN S, ZHANG R, LIU Y, et al. Building lod representation for 3d urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2025, 226: 16-32
- [32] LIN H, GAO J, ZHOU Y, et al. Semantic decomposition and reconstruction of residential scenes from lidar data. *ACM Transactions on Graphics*, 2013, 32: 4:1-4:10
- [33] CHEN J, XU Y, LU S, et al. 3-d instance segmentation of mvs buildings. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-14
- [34] NIESE T, PIRK S, ALBRECHT M, et al. Procedural urban forestry. *ACM Transactions on Graphics*, 2022, 41: 2:1-2:18
- [35] GUI S, QIN R. Automated lod-2 model reconstruction from very-high-resolution satellite-derived digital surface model and orthophoto. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 181: 1-19
- [36] HAN J, ZHU L, GAO X, et al. Urban scene lod vectorized modeling from photogrammetry meshes. *IEEE Transactions on Image Processing*, 2021, 30: 7458-7471
- [37] YANG G, XUE F, ZHANG Q, et al. Urbanbis: A large-scale benchmark for fine-grained urban building instance segmentation//*Proceedings of the SIGGRAPH*. Los Angeles, USA, 2023: 16:1-16:11
- [38] WU Y, XUE F, LI M, et al. A novel building section skeleton for compact 3d reconstruction from point clouds: A study of high-density urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2024, 209: 85-100
- [39] DONG M, XIE K, HUANG H. Coupled study on instance segmentation and structural reconstruction of buildings in 3d urban scenes. *SCIENTIA SINICA Informationis*, 2024, 54(2): 281-300
- [40] HUANG Z, ZHAO Z, LI B, et al. LCPformer: Towards effective 3d point cloud analysis via local context propagation in transformers. *IEEE Transactions Circuits and Systems for Video Technology*, 2023, 33(9): 4985-4996
- [41] ZOU Z, LI Y. Efficient urban-scale point clouds segmentation with bev projection. *arXiv preprint arXiv:2109.09074*, 2021
- [42] YOO S, JEONG Y, JAMEELA M, et al. Human vision based 3d point cloud semantic segmentation of large-scale outdoor scenes//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada, 2023: 6577-6586
- [43] LV C, LIN W, ZHAO B. Approximate intrinsic voxel structure for point cloud simplification. *IEEE Transactions on Image Processing*, 2021, 30: 7241-7255
- [44] LV C, LIN W, ZHAO B. Intrinsic and isotropic resampling for 3d point clouds. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 2022, 45(3): 3274-3291
- [45] WU X, JIANG L, WANG P S, et al. Point transformer v3: Simpler faster stronger//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Seattle, USA, 2024: 4840-4851
- [46] YU M, LAFARGE F. Finding good configurations of planar primitives in unorganized point clouds//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New Orleans, USA 2022: 6367-6376
- [47] EDELSBRUNNER H, KIRKPATRICK D, SEIDEL R. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 2003, 29(4): 551-559
- [48] HOU F, WANG C, WANG W, et al. Iterative poisson surface reconstruction (ipsr) for unoriented points. *ACM Transactions on Graphics*, 2022, 41(4): 128:1-128:13
- [49] XU R, DOU Z, WANG N, et al. Globally consistent normal orientation for point clouds by regularizing the winding-number field. *ACM Transactions on Graphics*, 2023, 42(4): 1-15
- [50] HU Q, YANG B, KHALID S, et al. Towards semantic segmentation of urban-scale 3d point clouds: A dataset, benchmarks and challenges//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Nashville, USA, 2021: 4977-4987
- [51] SELVARAJU P, NABAIL M, LOIZOU M, et al. Buildingnet: Learning to label 3d buildings//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Montreal, Canada, 2021: 10377-10387
- [52] THOMAS H, QI C R, DESCHAUD J E, et al. Kpconv: Flexible and deformable convolution for point clouds//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach, USA, 2019: 6411-6420
- [53] ZHAO H, JIANG L, JIA J, et al. Point transformer//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Montreal, Canada, 2021: 16259-16268
- [54] ENG B, WU X, JIANG L, et al. Oa-cnns: Omni-adaptive sparse cnns for 3d semantic segmentation//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Seattle, USA, 2024: 21305-21315
- [55] GAO X, WU K, PAN Z. Low-poly mesh generation for building models//*Proceedings of the SIGGRAPH*. Vancouver, Canada, 2022: 3:1-3:9
- [56] CHEN Z, PAN Z, WU K, et al. Robust low-poly meshing for general 3d models. *ACM Transactions on Graphics*, 2023, 42(4): 119:1-119:20



LV Chen-Lei, Ph. D., assistant professor. His research interests include digital computer graphic, geometric analysis, 3D vision.

HE Xin, Ph. D. candidate. Her main research interests include computer graphics and geometric analysis

XIE Heng, M. S. candidate. His main research interests include artificial intelligence and 3D vision.

HUANG Hui, Ph. D., professor. Her main research interests include computer graphics and 3D vision.

Background

This research belongs to the field of 3D vision and computer graphics, with a particular focus on semantic analysis and structural modeling of large-scale urban scenes. The data of urban scenes contains large areas and numerous semantic objects, and it faces several limitations in terms of data quality, scale differences, and semantic coupling, which restricts the efficient application of urban data.

In this paper, we propose BuildingRecon, which combines scene data semantic analysis and structured reconstruction techniques to achieve semantic analysis and volumetric modeling of urban data. By embedding elevation information, the semantic perception ability of the semantic analysis network for high-rise buildings is significantly enhanced. Through structured patch extraction, polyhedron intersection, and volumetric reconstruction, our method effectively constructs streamlined main structures for building objects, while correcting erroneous

geometric linkages, resulting in high-quality, low-volume, and clear-boundary structural models. Experimental results show that the proposed algorithm can effectively improve point cloud semantic segmentation accuracy, optimize urban data volume, and obtain more precise building boundaries. This project has good engineering application value.

This research is supported by the National Key R&D Program of China (2024YFB3908500, 2024YFB3908502), NSFC (62572323), Guangdong Basic and Applied Basic Research Foundation (2023A1515110292, 2023B151512002 6), Shen zhen Science and Technology Program (KQTD202108110900-44003, KJZD20240903100022028, JCYJ2025060418152403 2), Key Project of Teaching Reform of Shenzhen University (JG2024018), Scientific Foundation for Youth Scholars of Shenzhen University, and Scientific Development Funds from Shenzhen University.