



计算机科学

COMPUTER SCIENCE

基于点云的室内结构三维重建综述

任飞, 常青玲, 刘兴林, 杨鑫, 李敏华, 崔岩

引用本文

任飞, 常青玲, 刘兴林, 杨鑫, 李敏华, 崔岩 [基于点云的室内结构三维重建综述](#) [J]. 计算机科学, 2022, 49(11A): 211000176-11.

REN Fei, CHANG Qing-ling, LIU Xing-lin, YANG Xin, LI Ming-hua, CUI Yan. [Overview of 3D Reconstruction of Indoor Structures Based on Point Clouds](#) [J]. Computer Science, 2022, 49(11A): 211000176-11.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于点云数据的交通环境下单阶段三维目标检测方法](#)

Single-stage 3D Object Detector in Traffic Environment Based on Point Cloud Data
计算机科学, 2022, 49(11A): 210900079-6. <https://doi.org/10.11896/jsjcx.210900079>

[基于PCPNET的点云特征线提取算法](#)

Point Cloud Feature Line Extraction Algorithm Based on PCPNET
计算机科学, 2022, 49(11A): 210800017-6. <https://doi.org/10.11896/jsjcx.210800017>

[基于边缘引导的自校正皮肤检测](#)

Edge Guided Self-correction Skin Detection
计算机科学, 2022, 49(11): 141-147. <https://doi.org/10.11896/jsjcx.220600012>

[基于可变形图卷积的点云表征学习](#)

Deformable Graph Convolutional Networks Based Point Cloud Representation Learning
计算机科学, 2022, 49(8): 273-278. <https://doi.org/10.11896/jsjcx.210900023>

[三维激光雷达点云空间多特征分割](#)

Spatial Multi-feature Segmentation of 3D Lidar Point Cloud
计算机科学, 2022, 49(8): 143-149. <https://doi.org/10.11896/jsjcx.210300275>

基于点云的室内结构三维重建综述

任飞¹ 常青玲¹ 刘兴林¹ 杨鑫¹ 李敏华¹ 崔岩^{1,2}

1 五邑大学智能制造学部 广东 江门 529000

2 四维时代 广东 珠海 519000

(oncecooo@gmail.com)

摘要 室内结构三维重建本质上是一个还原室内布局的多任务问题,可以进一步对墙体细节和家具进行重建和语义分割。主要介绍基于点云数据的室内结构三维重建。首先概述了室内结构三维重建常用的数据集;然后对基于点云的室内结构三维重建的主要方法展开叙述和讨论,并分析总结了3种类型重建方法的优缺点;最后对当前室内结构三维重建研究所面临的困难和挑战进行阐述,并对未来的研究趋势做出展望。可以得出,目前大部分重建模型所针对的场景和完成任务的多样性较为贫乏,利用不同角度的冗余信息共同优化的多任务协调方案在室内结构重建中具有较大潜力。此外,模型对于室内外环境的无缝融合以及实现内外建筑的充分表现仍需要进行改善。

关键词: 室内结构; 三维重建; 布局估计; 多任务; 点云

中图法分类号 TP391

Overview of 3D Reconstruction of Indoor Structures Based on Point Clouds

REN Fei¹, CHANG Qing-ling¹, LIU Xing-lin¹, YANG Xin¹, LI Ming-hua¹ and CUI Yan^{1,2}

1 Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, Guangdong 529000, China

2 4Dage, Zhuhai, Guangdong 519000, China

Abstract The 3D reconstruction of indoor structure is essentially a multi-task problem of restoring the indoor layout, which can further reconstruct and semantically segment wall details and furniture. This paper mainly introduces the 3D reconstruction of indoor structure based on point cloud data. Firstly, the data set commonly used for 3D reconstruction of indoor structure is summarized, and then the main methods of 3D reconstruction of indoor structure based on point cloud are described and discussed, and the advantages and disadvantages of the three types of reconstruction methods are analyzed and summarized. Finally, the difficulties and challenges faced by the current 3D reconstruction research of indoor structures are explained, and the future research trends are prospected. It can be concluded that the diversity of scenes and tasks completed by most reconstruction models at present is relatively poor, and the multi-task coordination scheme that uses redundant information from different angles to jointly optimize has great potential in the reconstruction of indoor structures. In addition, the model still needs to be improved for the seamless integration of the indoor and outdoor environments and the full performance of the interior and exterior buildings.

Keywords Indoor structure, 3D reconstruction, Layout estimation, Multi-task, Point cloud

1 引言

三维重建作为计算机视觉领域的基础性任务,受到了广泛关注。近些年,随着房地产与室内勘测需求的增加,室内结构三维重建得到了迅猛的发展。该技术侧重于还原室内的布局和特征细节,主要用于房地产销售、火灾场景修复、虚拟博物馆等。它不仅需要宏观上重构室内布局,还需要恢复室内(如天花板、标签、墙壁等)永久性结构细节,并通过语义分割、路径规划等任务来丰富室内表达^[1-4],因此室内结构三维重建是一个多任务问题^[5-10],常见的做法是基于点云数据进行室内结构的三维重建。点云是一种基于笛卡尔坐标系的数据存储格式,每一个点都表示物体表面与采样仪器的相对距离,而点云数据常常伴随着缺失和漂移,增加了重建难度,因此利用点云进行重建时往往需要进行补全和正则化操作^[11-13]。此外,场景点云数据一般都含有数百万个点,点的互相关操作

需要花费大量时间,并且难以定义点集间的空间关系。为了解决这些问题,从2003年^[14]至今,研究人员做了大量的工作,分别提出了不同的解决方法。概括来讲,这些方法大致分为基于二维轮廓平面的三维重建、基于网格单元的三维重建、基于邻接图和基于学习的三维重建方法。其中,基于二维轮廓平面的方法将三维点云数据映射到二维平面上,用二维点的高度直方图峰值来代替墙体的拟合,降低了重建难度和时间消耗,适用于长廊等规则建筑体的重建^[5-9],但这些方法大多数仅恢复墙体结构,对室内空间划分和其他永久性结构(如门窗)重建都比较吃力;基于网格单元的方法通过拟合平面建立丰富的候选网格,然后选择合并网格单元重建室内结构分布;该方法适合于满足曼哈顿世界约束^[13]的室内结构体,它省略了层间匹配连接步骤^[16-19],有利于室内结构的语义划分,但无法处理曲折多变的墙体结构。基于邻接图的方法可以一定程度上重构出室内家具,它通过邻接图建立了永久性

建筑结构与室内物体间的层次包含关系,可以逐渐细化并约束室内结构的表达^[20-23],改进了常用的只是把除墙体等永久性结构外的点云作为杂波项省略来提高拟合效率的重建方法^[19,24-26],但对对象结点的连接及关系分类需要消耗大量时间。为了提高点云重建的速度和鲁棒性,基于学习的方法在小批量的点云数据上进行训练,而在大数据集上进行测试,通过神经网络实现端到端的重建过程^[12,19,25,27-28]。为了更好地理解基于点云的室内三维重建方法,也为初学者提供入门资料,本文对2009年至今的相关工作进行综述,在概述常用点云数据集的基础上,重点介绍基于点云数据的室内结构三维重建方法,并将相关方法进行对比分析。此外,结合相关研究进展,对室内结构三维重建存在的问题进行分析总结,并对未来的发展趋势进行展望。本文主要贡献如下:

(1)梳理了2009年至今约80篇基于点云的室内结构三

维重建的文献,为研究者及初学者提供参考依据;

(2)分别介绍了室内结构三维重建方法,并进行了详细的对比分析,同时分析了各算法适用的场景,加深对相关算法的理解;

(3)对现存问题的剖析和未来趋势的展望,则为后续研究提供了一定方向性指引。

2 相关数据集及对比分析

2.1 室内结构三维重建常用数据集

点云数据集根据执行任务的差异具有不同的标签信息,常见的任务有室内结构的几何、语义重建,分别用于提取室内结构和语义划分。本章概述了几个室内结构三维重建常用的数据集,并根据各数据集的规模及适用场景进行对比分析,如表1所列。

表1 基于点云的室内结构重建数据集

Table 1 Data set of indoor structure reconstruction based on point cloud

数据集	数据类型	数据介绍及规模	任务类别
MiMAP ^[29]	RGB 图像+点云	包括室内 LiDAR SLAM、BIM 特征提取、室内定位 3 个数据集	SLAM 与室内定位等
ISPRS ^[30]	RGB 图像+点云	提供了城市分类和三维建筑重建测试项目的两个最先进的机载图像数据集	三维重建+语义分割
SUNCG ^[31]	RGB-D 图像	包含 45622 个不同的场景,49884 层有效楼层,包含 404058 个房间和 5697217 个对象实例,来自 2644 个独特的对象网格,涵盖 84 个类别	三维重建+语义分割
SUN RGB-D ^[32]	RGB-D 图像	10335 张不同场景的室内图片,146617 个 2D 多边形标注(应该指的是 2D 分割),和 58657 个 3D 边框	布局估计+语义分割+三维物体识别
ScanNet ^[33]	RGB 图像+点云	1513 个采集场景数据,共 21 个类别的对象,其中 1201 个场景用于训练,312 个场景用于测试	三维重建+语义分割
NYU ^[34]	RGB-D 图像	NYU depth v1:64 个不同的室内场景,7 个场景级类,2347 个密集标记帧 NYU depth v2:1449 张 RGB-D 图像,描述了 26 个场景级别的 464 个室内场景的 31 个详细的注释	深度估计+语义分割
Stanford 2D-3D-Semantics ^[35]	RGB-D 图像	70496 个常规 RGB 和 1413 个全景 RGB 图像,以及它们相应的深度、表面法线、语义注释、全局 XYZ OpenEXR 格式和相机内参数数据	三维重建+语义分割
Matterport3D ^[36]	RGB-D 图像	90 个建筑规模场景的 194400 个 RGB-D 图像的 10800 个全景视图	表面重建+相机姿态+语义分割

(1)MiMAP^[29]。MiMAP 是第一个将基于光探测与测距(LiDAR)的 SLAM(Simultaneous Localization and Mapping)、建筑信息模型(BIM)特征提取、基于多传感室内测绘与定位的多任务联系起来的数据集。该基准包括 3 个数据集:室内激光扫描数据集(5 个场景)、BIM 特征提取数据集(3 个场景)和室内定位数据集(5 个场景)。数据集中的每个场景都包含来自多光束激光扫描仪的点云以及基于各种复杂室内环境的智能手机传感器的记录,拥有图像及点云信息,包括相机参数与位姿。

(2)ISPRS^[30]。ISPRS 数据集包含 5 个复杂室内环境中的点云数据,分别用 Viametris iMS3D,ZEB REVO,TLS Leica C10,UVigo Backpack,ZEB1 传感器捕捉获得。每个场景具有超过 1000 万个点云数据,部分包含颜色、轨迹信息,并对杂乱程度进行了评价。可用于三维重建、室内轨迹生成等。

(3)SUNCG^[31]。SUNCG 包含了 45622 个不同的场景,带有通过手工创建的逼真的房间和家具布局。有 49884 个有效楼层,包含 404058 个房间和 5697217 个对象实例,来自 2644 个独特的对象网格,涵盖 84 个类别。

(4)SUN RGB-D^[32]。SUN RGB-D 包含 10335 幅 RGB-D 图像,来自 RealSense 3D,Xtion LIVE PRO,Kinect V1,V2 共

4 个不同的传感器,带有 2D 语义信息、物体对象的 3D 框、房间布局等注释。包括 146617 个 2D 语义填充和 64595 个 3D 边界框,具有精确的物体方向,以及每个图像的 3D 房间布局和场景类别。对于这个数据集,一半的图像用于训练,另一半用于测试,以评估场景分类、语义分割、场景理解等任务。

(5)ScanNet^[33]。ScanNet 是一个大型 RGB-D 视频数据集,包含 250 万数据帧,横跨 1513 个场景。场景具有 707 个不同的空间,包括小空间(如壁橱、浴室和杂物间等)和大空间(如教室、公寓和图书馆等),用 3D 摄像机姿态、表面重建和语义分割进行注释。该数据集的 1205 个场景用于训练,另外 312 个场景用于测试。

(6)NYU^[34]。NYU 数据集包括 NYU depth v1 和 v2。NYU depth v1 有 64 个不同的室内场景,7 种场景类型,其中 60%用于训练,40%用于测试。NYU depth v2 包含 1449 张 RGB-D 图像,描述了 26 个场景级别的 464 个不同的室内场景,每幅图像具有 31 个详细的注释。该数据集中 795 张图像用于训练,664 张图像用于测试。

(7)Stanford 2D-3D-Semantics^[35]。Stanford 2D-3D-Semantics 数据集提供了一个大型室内空间数据集,通过实例级语义和几何注释捕捉 2D,2.5D 和 3D 数据。该数据集收集在

6个大型室内区域,包括3个不同的教育用途和3个办公用途的建筑,覆盖面积超过6000平方米,包含70000张RGB图像,以及相应的深度、表面法线、语义注释、全局XYZ图像和相机信息。

(8)Matterport3D^[36]。Matterport3D为室内环境提供了一个大型和多样化的RGB-D数据集。它包含了194400张RGB-D图像中的10800张全景图,描述了90座建筑的室内场景。不同于其他数据集,它包含每个像素点的深度和颜色,可用于表面重建、相机姿态估计及2D和3D语义分割注释等任务。精准的全局对齐与多样的全景视图覆盖整个建筑,实现关键点匹配、视图重叠预测、颜色正常预测、语义分割和区域分类。

2.2 室内结构三维重建常用数据集对比分析

SUNCG^[31], SUN RGB-D^[32], NYU^[34], Stanford 2D-3D Semantics^[35]和Matterport3D^[36]室内建筑数据集主要包含图片色彩和深度信息,以及场景语义分布。深度图是与原图像分辨率相同的深度表示,可以转化为稀疏的点云表示。由于图像深度较易获取,这些数据集都拥有很大的数据规模,主要应用于室内三维重建以及语义分割。MiMAP^[29], ISPRS^[30]和ScanNet^[33]数据集以RGB图像与场景点云为输入源,由于处理的是整个场景的稠密点云信息,因此具有更规范的重建表达。

3 基于点云的室内结构三维重建方法及其对比分析

3.1 基于点云的室内结构三维重建方法

基于点云的室内三维重建方法从点云数据中恢复室内结构的墙体、地板与天花板等室内结构模型,高质量的模型表示意味着在一定点云数量下生成更多的平面并形成与真实场景相匹配的墙体。根据重建模型构造步骤的不同,本节把基于点云的室内三维重建方法分为4类:基于二维轮廓平面的方法、基于网格单元的方法、基于邻接图的方法和基于学习的方法。在概述了每类方法的基础上,对其中的典型方法进行比较分析,以加深对相关重建方法的理解。

3.1.1 基于二维轮廓平面的室内三维重建方法

(1)方法介绍

基于二维轮廓平面的方法将三维点数据映射到二维空间^[5,37],构造出室内结构的二维俯视图轮廓图,通过墙体高度拉伸轮廓图生成重建模型,即先舍弃墙体的高度信息,把重建的重心放在室内环境的二维边缘上,以降低重建难度,同时节约重建时间。

Okorn等^[5]第一次提出了自动提取用于拉伸重建的二维轮廓图方法。首先把三维点空间映射到二维平面上,根据室内建筑的垂直高度(每10cm为界限)生成有关于垂直高度区间内水平点云数量的高度直方图,并分别依据高度直方图顶部和底部(分别代表天花板和地面的水平点云区间)的局部最大值移除天花板和地面点云,减少了生成室内结构模型的干扰。根据移除后的点云分布提取出室内结构轮廓,并用霍夫变换^[38]平滑边缘线段,确保生成水密性的二维轮廓平面。但二维轮廓生成过于依赖点云分布,在嘈杂点云数据下精度较低,并且不能生成门窗等永久性结构。映射后二维点云的分布杂乱,提取生成的二维轮廓并不能准确表达室内俯视轮廓结构。为了提高点云提取的准确性,Turner等^[6]将映射后的

二维点云划分为多个子区域单元。如图1所示,点集(蓝色点)在 $x-y$ 平面上被网格划分(黑色线)。对于每个网格单元,计算区域点的邻域(红色的圆形区域),并将这些点的中间位置(绿色点)作为壁面样本。以采样点替代原点云数据进行边缘拟合,一定程度上减少了数据的嘈杂度。但采样点间只有部分的邻域交集,不能体现采样点间的相关性,生成轮廓并不能完全符合室内结构模型生成的标准。在此研究上,为了减少点云映射分布数据的杂乱并保证生成二维轮廓的水密性,Turner等^[39]将投影点存储在四叉树结构中进行分类,以点云的垂直密度为基准代替邻域分布采样来提取数据样本,并增加了房间的语义信息。为了避免映射后二维点云分布的杂乱并提高采集数据的质量,Stojanovic等^[7]将三维点云数据分割为3个垂直层,通过轴对齐边界框(axis-aligned bounding box)选择点云密度最大的垂直层来生成平面图,减少了轮廓图生成的点云杂波项。

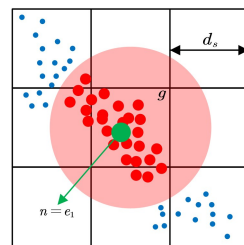


图1 Turner等^[6]提出的二维轮廓生成方法(电子版为彩图)

Fig. 1 Two-dimensional contour generation method proposed by Turner et al. ^[6]

除了对生成二维轮廓平面的点云数据进行处理加工外,还能够通过增加轮廓边缘的互相关性优化生成模型质量。Oesau等^[40]根据水平方向的点云密度把多层的室内点云结构分割为多个轮廓平面,通过多尺度线拟合方法为每个代表局部墙体方向的点生成线假设,可以对多层建筑进行重建,打破了单层建筑结构重建的局限。为了提高生成轮廓的自由度,Monszpart等^[41]提出了从丰富的候选面中选择平面的方法(RAP),它不强调构建水密性结构,也不受曼哈顿世界约束。由于初始点云的噪声影响,拟合的平面不能准确地映射到真实空间中,RAP通过旋转一定角度的初始平面生成多个候选平面,提供了丰富的先验信息。为了提高稀疏点云的重建质量,Lin等^[42]提出了一种对齐多个局部点云数据的方法,在不需要特征匹配的情况下将局部布局组合成一个具有闭环的全局对齐布局。通过平面的墙体数量和凸包边的数量以及作为复杂项生成最优的全局轮廓,保证了缺失点云和多段点云的重建效果。一般地,点云数据在墙角位置并不能稠密体现,导致不能生成水密性二维轮廓。因此许多学者针对未连接边间的组合问题,提出了解决方案。Previtali等^[43]将二维轮廓分解为多个相邻的垂直方格单元,将每个方格单元进行室内外的二元标记,通过最小化能量方程来对缺失墙体进行补充,但这种方法丢失了大量的边缘轮廓信息。为了保证还原轮廓平面的多样性,Wang等^[44]拟合出线性边缘,通过一维均值位移聚类算法优化候选线段,生成了非正交化的多边形标记单元。Bassier等^[45]提出了一种无监督的边缘重建方法。如图2所示,首先匹配每个二维平面端点的相邻墙面,然后确定一组潜在的连接方式(交叉连接、正交连接、混合连接和直接连接),根据成本函数确定一组唯一的连接,并可以创建

每个网格单元与轨迹之间的能见度差,平滑项定义为相邻网格面上的结构点所占面积的比例。为了加强墙体的表达细节,Murali等^[48]提出了3种判断网格间的连接方式:断层连接、门连接和合并连接。若两相邻网格的覆盖面积小于其公共面的10%,则定义为断层连接;将公共面中存在2m×1.8m大小长方形的连接定义为门连接;上述两种都不属于的则定义为合并连接。通过网格的连接方式进一步分配语义信息。但该方法仅针对单层室内结构,不适合复杂结构。基于网格的室内结构重建需要通过语义信息对室内结构进行分布。除了上述通过点云拟合的网格单元进行语义划分外,还能够通过拟合平面的选择来最优化模型生成。Ochmann等^[49]将优化方法的解空间局限在代表建筑墙面的实体点云上,增加了一致性的体量表达,可以在不丢失初始点云信息的情况下进行语义划分。如图3(c)示,采用基于点云相互可见性的无监督聚类方法^[50],自动去除离群点并分割房间。提取标记点云的水平与垂直表面(见图3(d)),并生成大量的候选平面(见图3(e))。

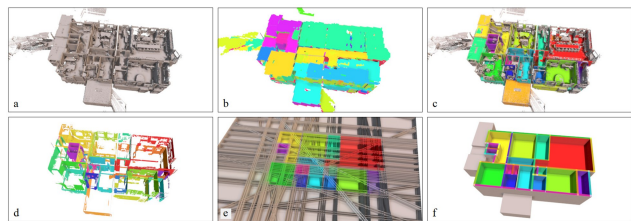


图3 Ochmann等^[49]的重建模型

Fig. 3 Reconstruction model of Ochmann et al. ^[49]

为了增加网格单元互相关性,额外增加了网格单元连接的边缘信息。如图4所示,对于两个相邻的单元格,将它们的邻近关系分为6种,增加了模型的置信度;同时把最优化表述为一个整数线性规划问题,具有唯一的最优解。该方法通过点云聚类分类出空间的标签信息,通过点云附带标签对拟合的室内结构进行分类划分,减少了网格单元合并的运算周期。但该方法过于依赖于点云的聚类分类,在网格合并时未进行多重检测。

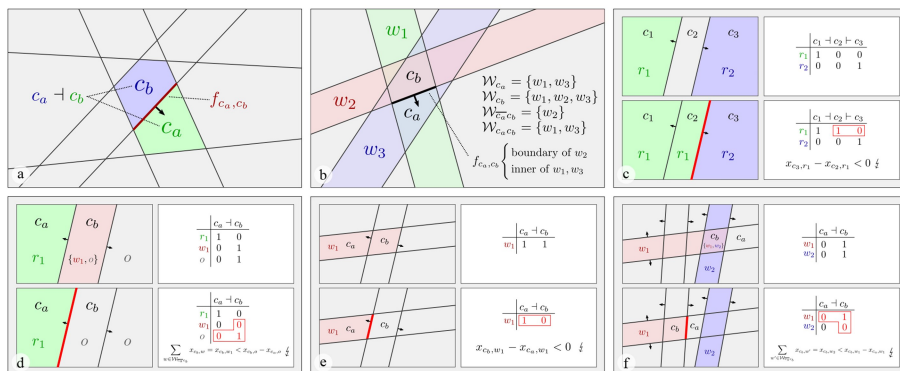


图4 Ochmann等^[49]的约束墙体方法

Fig. 4 Restraining wall method of Ochmann et al.

基于网格单元的室内重建方法解决了基于二维轮廓平面方法不能处理复杂室内建筑的问题,能够对多层复杂室内建筑进行重建,同时进行高精度的室内划分,但不能重建室内物体对象,只能还原建筑结构外形。

(2)方法对比与分析

基于网格单元的室内三维重建方法的对比分析如表3所列。

表3 基于三维网格单元方法的比较

Table 3 Comparison of methods based on 3D grid cell

	数量/万	生成样本		时间/s
		房间	墙体	
Nan等 ^[17]	18.6	—	26	3.36
Coudron等 ^[18]	—	15	—	27
Cui等 ^[24]	897	—	—	65
Ochmann等 ^[49]	645	49	63	219.7

Nan等^[17]以数据融合度、模型复杂度和点覆盖率为约束,通过最小化能量方程来最优化重建模型。输入18.6万个点云数据,生成了28个平面线段和1778个候选平面,最终模型具有26个平面。初始提取、候选生成和平面选择优化共用了3.36s。Coudron^[18]等额外引入了判断相邻平面是否为同一平面的能量项,重建15个房间模型共消耗27s。Cui等^[24]以网格单元标记和相邻单元成对代价为能量项,拟合重建897万个点云数据,共花费了65s。Ochmann等^[49]将最优化

重建模型的解空间限定在实体点云上,初始了6种平面和空间的包含关系。在645万点云数据下生成了49个房间标签,63个墙体单元,共花费了219.7s。可以得出结论:通过最小化能量方程的重建方法会比平面选择与合并的方法更节约时间,而约束项越多的能量方程可以生成更复杂的重建模型,同样也会消耗更多的时间。

3.1.3 基于邻接图的室内三维重建方法

(1)方法介绍

室内空间也可以通过物体标签间的关系进行定义和表示。通过建立室内物体(如墙体、窗户、门等)标签的拓扑关系对室内结构进行重建。室内环境的重建不单是几何重构和语义分割,还需要建立拓扑关系^[43],建立室内信息的具体分布。拓扑表示提供了关于几何对象的有用信息,如建筑物中的门的具体位置,从而便于对室内路线进行规划。基于图的方法是一种拓扑表示,指房间、走廊、和门等物体间的排列关系,通常提供以下信息:哪个房间与哪个房间相邻?哪个门连接哪个房间^[51]?基于邻接图的方法把墙体、天花板等室内结构通过邻接图的方式相关联,简化了重建步骤并节约了保存空间。其中,邻接图的结点表示为室内物体类别,邻接边表示为类别间的包含关系。一般地,自取一点作为邻接图的顶点,通过判断顶点周围邻接物的连接关系来生成整体邻接图,除了抽象名词“房间”“场景”等,其他结点元素都能与室内物体相关联。

Ikehata 等^[20]定义结点由3种类型的边连接。首先,当结构元素有一个共同的表面边界(如相邻的墙体)时,这些结点由无向边连接(图5中的紫线),确保几何形状与边界完全匹配,并保证重建结构的水密性;其次,虚线表示没有边界约束的连接关系(如房间中的对象);最后,有向边表示子节点上的结构元素构成其父节点上元素细节。如在图5中,房间结点包含4个墙体结点。带门的墙体结点在其结点的中间连接包含一个门结点,代表相邻边的门。而父结点的边界约束也必须由子结点继承,以保证水密性。

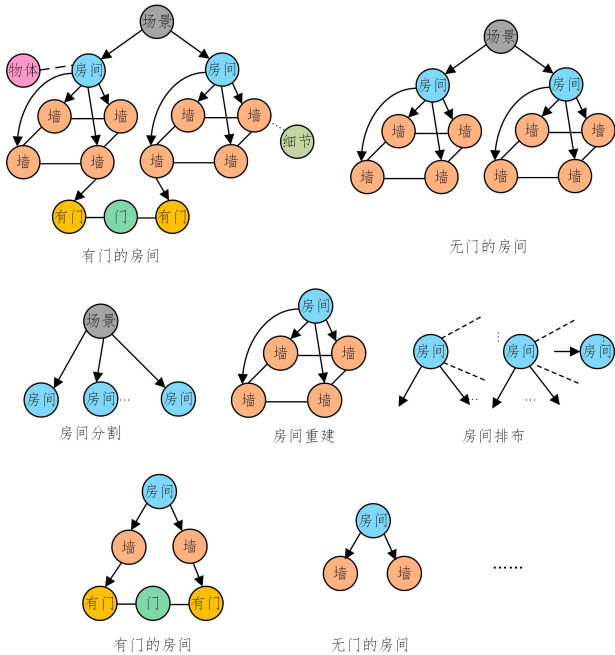


图5 Ikehata 等^[20]的构建邻接图方法(电子版为彩图)

Fig. 5 Method of constructing adjacency graph of Ikehata et al. ^[20]

Mura 等^[21]设定了6种相邻平面的交接方式,用于判定检测邻接图结构化路径,可以对除墙体外的其他平面进行邻接图构造。Guo 等^[52]针对房间家具等细小物体进行了邻接图构建,建立了庞大的邻接图分支,但结点间的连接关系较弱,丢失了生成模型的唯一性。在构建室内结构模型时,纯拓扑结构并不能完整地表达平面间的相接关系,且不宜于后期的调整和优化。Nikoohemat 等^[47]不需要进行空间细化,引入空间封闭的方法处理3维空间提取和非曼哈顿世界建筑,建立了墙体与地板和天花板的连接,保证了水密性建筑的生成。Bassier 等^[22]建立了基于最近邻间欧氏距离的全连接图的表示,并通过3个约束公式来削减连接表示,提高了平面选择的准确度。但邻接图用于处理生成平面的匹配关系,放弃了邻接图结点属性规划,不利于室内空间的语义划分。Nikoohemat^[1]也建立了平面的全连接图表示,可以对楼梯等斜体平面进行构建划分。Lim 等^[2]根据平面的从属关系,建立了一个可以验证水密性的邻接图表示,从而能够解释丢失平面在内嵌空间的相对位置。邻接图可以对候选平面进行选择优化,并可以根据连接关系进行语义划分,但缺少了房间内部结构的语义划分。Nikoohemat 等^[23]初始化可交互的模型类别,通过类别间的包含关系建立一个可以隐藏显示的多边邻接图,如图6所示,几乎所有室内实例都可以连接到房间。实线边表示两个实例必须连接,虚线边表示两个实例之间的可选连接,保留了室内结构的分布表达。

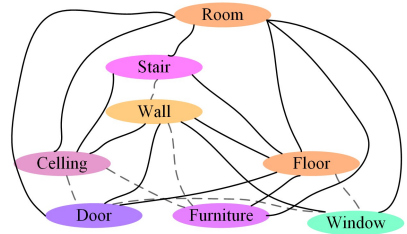


图6 Nikoohemat 等^[23]的结构邻接图

Fig. 6 Adjacency graph of structure of Nikoohemat et al. ^[23]

(2)方法对比与分析

基于邻接图的室内三维重建方法对比如表4所列。Ikehata 等^[20]在曼哈顿假设下预设了3种邻接图边连接方式,分别定义了包含、连接与共有关系,保证了邻接平面间的关系属性。在输入点云为102.2百万点云数量下,生成了6个房间和88个房间样本,初始化、房间分割、重建及优化共耗费了3120s。Mura 等^[21]设定了包含非连接关系的6种平面邻接属性,能够有效处理拟合错位平面的从属关系,在27.8百万数量点云下生成了4958个二进制空间划分单元,生成了6个房间和53个墙体,空间划分、房间检测和重建及优化步骤共用了79.9s。Bassier 等^[22]建立了基于最近邻间欧氏距离的全连接图的表示,将邻接图部署到平面选择上,放弃了结点属性规划的约束条件,构造了丰富的候选平面连接。在115.9百万点云数据下,生成了21.1万三维网状物(meshes),75个墙体,消耗时间5.5s。Lim 等^[2]根据平面的从属包含关系建立了邻接图,嵌入平面的邻接作为其从属平面结点的子图。输入数据为99百万点云数据,建立了42个结点,生成了252个墙体样本。可以得出结论:赋予对象结点属性的邻接图重建有利于对室内语义空间进行划分,但相对来说会更耗费时间。同时,平面关系约束条件约严谨,生成结点会越丰富,同时也会节省重建及优化模型的时间。

表4 基于图的方法比较

	点云数量/ 百万	结点/ 网格	生成样本		时间/s
			房间	墙体	
Ikehata 等 ^[20]	102.2	1603	6	88	3120
Mura 等 ^[21]	27.8	4958	6	53	79.9
Bassier 等 ^[22]	115.9	—	—	75	5.5
Nikoohemat 等 ^[1]	2.5	—	4	330	204
Lim 等 ^[2]	99	42	—	252	—

3.1.4 基于学习的室内三维重建方法

(1)方法介绍

传统方法将重建任务转化为其他问题,注重的是重建的视觉效果和重建平面的数量与复杂度,没有一致性的度量参考。基于学习的方法通过神经网络提取点云数据的特征,以获得高层语义表达。深度学习是一种自学习特征表达方法,相较于传统依赖先验信息的表达方法,深度网络可以有效处理稀疏和缺失点云数据,具有更高的效率^[25,31,53-58]。与图像数据不同,点云数据是一种三维的数据格式,很难从中提取出连续的语义特征,处理过程会更加复杂和困难。Qi 等^[59]率先以纯点云数据为网络输入,通过线性回归的方式对小批量点云数据进行分类和语义分割,模型复杂度随数据的复杂化呈线性增长。该结构由分类和分割两个子网络组成。分类子网络通过对每个点的多层感知器序列来生成点特征,然后

通过最大化池层生成描述原始输入的全局特征,再通过分割子网络将全局特征与分类网络提取的逐点特征连接起来,并应用另一个多层感知序列来生成每个点的最终输出分数。基于学习的方法通常用于点云数据的几何和语义信息的提取,具有较强的泛化能力和较大的拓展空间。由于计算能力限制,目前基于学习的点云室内三维重建方法并不能满足室内建筑全局提取,主要应用于边框提取与点云语义分割上。

Song 等^[31]从 RGB-D 图像中提取出场景局部的连续的体素表达,使用一个基于扩张的三维全局模块,扩展了信息采集域,有效提高了处理稀疏点云数据的精度与速度。但由于没有利用任何颜色信息,类似窗口等出现点云漂移的对象很难进行重建。Tchapmi 等^[58]对三维点云数据体素化,并通过三维全卷积神经网络生成粗糙的下采样体素语义标签,减少了数据处理堆积,再用三线性插值法将体素转换回原始的三维点表示,在恢复原点云数据的同时使点云呈线性分布。Dai 等^[56]提出了一种可以处理大场景点云数据的三维全卷积神经网络(见图 7),对输入数据的空间范围具有不变性。

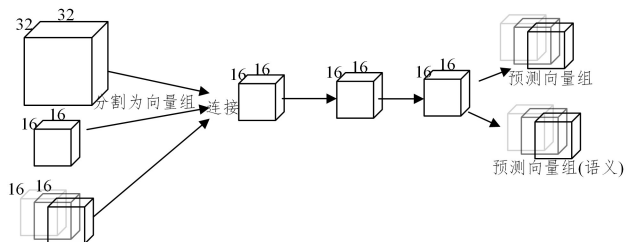


图 7 Dai 等^[56]提出的网络框架

Fig. 7 Network framework proposed by Dai et al. ^[56]

为了提高生成点云的分辨率,使用了从大到小的 3 种体素表达层次结构。大的体素表达用于提供场景的空间背景信息,而小的体素表达用于捕获局部细节。但体素表达仍不够理想,并且不能对室内永久性结构和室内其他对象进行区分。基于此,Coudron 等^[25]简化了网络,提出了基于曼哈顿假设非符号距离空间的数据表示,提取了室内建筑的永久性结构表示。

室内点云数据常伴随丢失与漂移,重建可能会出现错位和变形,因此点云补全任务也十分重要^[13,60]。Yu 等^[60]根据点云数据的稀疏性和不规则性,提出了一种数据驱动的点云上采样技术,输出更密集的点云数据。首先从已知的三维点云数据中提取不同尺度和分布的点集,然后通过分层特征学习和多层特征聚合将原始三维点云坐标映射到特征空间。并通过对点云数据进行重塑来增值点云数据,避免了反卷积或全连接层进行上采样的处理复杂度。

(2)方法对比与分析

基于学习的点云室内三维重建方法比较如表 5 所列。Song 等^[31]从 RGB-D 图像中提取出场景局部的连续的体素表达,在数据集 SUNCG 和 NYU 下生成了 30.5%室内对象。Tchapmi 等^[58]通过点云体素化减少了数据堆积,在 NYUV2 数据集下生成了约 56.4%室内对象。Dai 等^[56]通过建立三维卷积网络模型,可以处理大型场景数据集,在 SUNCG 数据集中生成了 68.6%对象。Coudron 等^[25]提出了非符号距离场的空间表示,在 SUNCG 数据集下生成了 58.84%的物体对象。可以得出结论:利用三维卷积网络能够更适配地重建出室内语义场景,但由于直接在三维数据上进行遍历计算,三维卷积网络处理需要远大于普通卷积操作的时间损耗。

表 5 基于学习的方法比较

Table 5 Comparison of methods based on learning

(单位:%)

方法	场景完成度			语义场景完成度				数据集
	Prec.	Recall	IOU	天花板	地板	墙体	均值	
Song 等 ^[31]	59.30	92.90	56.60	15.10	94.60	24.70	30.50	SUNCG+NYU
Tchapmi 等 ^[58]	—	—	43.45	69.16	95.21	78.89	56.40	NYUV2
Dai 等 ^[56]	—	—	—	97.80	94.30	48.30	68.60	SUNCG
Coudron 等 ^[25]	88.41	94.03	83.71	84.07	85.07	68.94	58.84	SUNCG

3.2 基于点云的室内三维重建方法对比与分析

如表 6 所列,我们对分析基于点云的室内三维重建方

法的原理和核心及难点,并对这些问题的优势和不足进行了讨论。

表 6 点云室内结构重建方法

Table 6 Point cloud indoor structure reconstruction method

方法	原理概述	核心及难点	优势与不足
基于二维轮廓平面的方法	映射室内结构模型的二维平面,通过点云的高度直方图建立三维模型	二维平面的水密性结构构造以及墙体高度一致性	重建速度快,可直接生成 CAD 表示;墙体、地板与天花板间的连接细节不能很好地重建
基于网格单元的方法	拟合点云平面,构造丰富的候选网格,选择并合并网格实现与真实室内模型的拟合	缺失点云的平面拟合和对网格单元的选择与合并	可以在重建过程中对室内结构的房间语义信息进行划分,并且可以对多层建筑进行重建构造;时间复杂度较高,对缺失点云数据不能很好地重建
基于图的方法	构造室内物体间的关系连接图,通过邻接图间的连接映射出环境的包含关系	邻接图结点间的包含关系以及对应于室内结构的复杂程度	易于保存与传输,可以直观地建成室内环境的复杂性;生成模型的精细度不一
基于学习的方法	通过神经网络对点云数据进行回归处理,得到场景语义或几何信息	通过多层神经网络实现点云到重建模型的端到端过程	能够得到连续的语义和几何信息,场景空间的泛化性较高;对细小的三维对象处理能力不足,处理规模有限

基于二维轮廓平面的方法仅针对于室内永久性结构的

重建,重建速度较快,适合于单层大场景点云数据规模的

重建,并且具有符合 BIM^[61]的重建思路。基于网格单元的方法更适合于多层建筑的重建,在室内语义划分的同时也能很好地构造建筑几何信息,但重建周期较长,对缺失点云不能很好地重建。基于图的方法构造室内物体对象间的邻接关系来进行重建,很好地还原了室内语义划分和结构分布,但重建周期很长。基于学习的方法通过以基数小的高质量点云数据训练,应用于大型的场景点云数据,重建速度很快,但一般需要色彩等额外环境属性。

4 基于点云的室内结构三维重建存在的问题与未来发展趋势

本节对基于点云的室内结构三维重建存在的问题进行总结,并对该技术的未来发展趋势进行展望。

4.1 基于点云的室内结构三维重建存在的问题

近年来,室内环境三维重建越来越受关注,在众多领域都得到了广泛的应用,如应急避险、建筑维护、导航引导等。但现有的基于点云的室内三维重建方法在速度、精度和鲁棒性等方面仍然面临着巨大挑战。

(1)室内环境复杂多样性问题。室内结构的复杂多样性会对室内三维重建造成巨大困难,常规方法只能在建立室内永久性结构的同时尽量还原室内物体对象模型。深度神经网络方法尽管具有极强的特征提取能力和语义分割能力,但都是基于小批量高精度数据集训练得到的,对于整体室内结构模型重建依旧存在很大困难。

(2)室内结构强约束问题。由于人造室内建筑大多符合曼哈顿世界假设,一般的重建模型都会以此为约束来保证重建水密性结构。对于曲面或多边的墙体结构不能做出适配的重建优化。

(3)点云数据缺失和漂移问题。点云数据一般提取于激光扫描仪或深度图,会受到不同程度的数据缺失和漂移,这将导致重建平面丢失,更严重的会使模型变形扭曲。虽然目前存在着对点云数据的补全和优化,但一般的重建方法都依赖于原有的数据支持,缺少对点云数据的完整性分析。

(4)缺少室内外合一场景数据集。目前点云数据集都是在室内或室外环境下捕捉采集的,无法对室内外场景进行一致性重建生成,导致训练的重建模型泛化能力较差。

4.2 基于点云的室内结构三维重建发展展望

基于点云的室内结构三维重建的重要性和应用前景已得到广泛认可,大量的研究人员针对该领域存在的问题展开一系列研究。基于此,本文认为基于点云的室内结构三维重建的未来发展趋势主要有以下几点。

(1)非曼哈顿世界约束。现有方法的主要局限性通常与强曼哈顿约束有关^[62-66],即假设建筑平面是相互正交的。目前,有些研究是基于亚特兰大世界^[67-68],增加了生成模型的兼容性。然而,在目前的室内环境中,存在着具有复杂几何结构的室内空间(如圆柱墙、球形天花板、室内布局或其他非平面结构^[69-70])。因此,恢复任意几何形状的完整室内三维模型是未来研究的一个必要课题。

(2)多任务协同优化。单个任务的目标通常不能生成足够的信息来可靠地实现理想的需求。多任务协同优化从不同角度提供冗余互补信息,以辅助任务完成。例如,对于协同估计房间布局 and 检测 3D 目标^[71-72],房间布局估计为目标检测

提供了空间约束信息,而目标检测也为房间布局估计提供了遮挡信息,两者相辅相成。另一个例子是联合优化房间布局 and 语义分割任务^[73]。因此,多任务协同优化与深度学习技术的融合也将是未来室内环境建模与应用的重点。

(3)室内外空间无缝融合。为了对室内外环境同时进行开发勘测并充分表现室内和外部建筑,模型的室内外环境重建一致性和连贯性十分重要^[74-76]。然而,由于实体之间的物理结构、空间关系等方面的差异,以及室内外点云数据的差距,无缝融合具有较大困难^[57,77]。而且,在大多数情况下,由于漂移或缺少匹配的特征,从外部到内部对整个场景进行完整重建也是不准确的^[78-79]。因此,在没有或很少数据重叠的情况下,将室内和室外模型无缝拼接会比较活跃的方向。

(4)具有不同细节层次的室内模型自动重建。到目前为止,建筑的结构建模只关注于主要的结构元素(如墙壁、地板、天花板和门窗)^[80-81]。为了促进特殊场景需求的发展,应该为场景定义和建立更详细的室内三维模型,其中包括描述室内家具的空间结构,并建模室内家具^[82],或建立室内的出入路径^[83-84]。这些应用可能需要具有不同精度级别的内部结构。在未来的工作中,需要探索更精细的几何特征,以生成几何准确、语义丰富的室内空间信息模型。

结束语 本文首先概要介绍了几种常用点云室内三维重建数据集并进行对比分析,重点对基于点云的室内结构重建方法进行分析综述,并针对数据概况对重建方法进行说明。点云数据拥有物体表面的几何信息,对空间内物体的距离有了基本的定义,同时可以通过添加颜色、语义等额外信息加强模型的可拓展性。针对现有基于点云的室内结构重建方法,将其划分为 4 类分别介绍。基于轮廓的方法能够快速进行室内布局估计,但缺少了室内环境物体的关联;基于三维轮廓的方法可以标记房间标签,对室内结构进行划分;而基于图的方法可以加大室内结构的重建规模,进一步约束并丰富室内环境表示;基于学习的方法减少了重建时间且提高了模型的鲁棒性。本文最后对当前研究中面临的难点和未来发展趋势进行了概述,相信室内结构重建研究在未来仍有着巨大的发展和应用潜力。

参 考 文 献

- [1] NIKOOHEMAT S, DIAKITÉ A, ZLATANOVA S, et al. Indoor 3d reconstruction from point clouds for optimal routing in complex buildings to support disaster management [J]. *Automation in Construction*, 2020, 113: 103109.
- [2] LIM G, OH Y, KIM D, et al. Modeling of architectural components for large-scale indoor spaces from point cloud measurements [J]. *IEEE Robotics and Automation Letters*, 2020, 5(3): 3830-3837.
- [3] CHANG K H, CHENG C Y, LUO J, et al. Building-gan: Graph-conditioned architectural volumetric design generation [C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.
- [4] NAUATA N, CHANG K H, CHENG C Y, et al. House-gan: Relational generative adversarial networks for graph-constrained house layout generation [C] // *Proceedings of the European Conference on Computer Vision*. Springer, 2020.
- [5] PREVITALI M, BARAZZETTI L, BRUMANA R, et al. To-

- wards automatic indoor reconstruction of cluttered building rooms from point clouds [C]//ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2014.
- [6] TURNER E, ZAKHOR A. Watertight as-built architectural floor plans generated from laser range data[C]//Proceedings of the 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission. IEEE, 2012.
- [7] STOJANOVIC V, TRAPP M, RICHTER R, et al. Generation of approximate 2d and 3d floor plans from 3d point clouds[C]//Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019). 2018.
- [8] WU Y, SHANG J, XUE F. Regard; Symmetry-based coarse registration of smartphone's colorful point clouds with cad drawings for low-cost digital twin buildings [J]. Remote Sensing, 2021, 13(10):1882.
- [9] CAI Y, FAN L. An efficient approach to automatic construction of 3d watertight geometry of buildings using point clouds [J]. Remote Sensing, 2021, 13(10):1947.
- [10] LI M, NAN L. Feature-preserving 3d mesh simplification for urban buildings [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 173:135-150.
- [11] DAI A, RUIZHONGTAI Q C, NIEßNER M. Shape completion using 3d-encoder-predictor cnns and shape synthesis [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [12] HAN X, LI Z, HUANG H, 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. 2017.
- [13] RIEGLER G, OSMAN ULUSOY A, GEIGER A. Octnet: Learning deep 3d representations at high resolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [14] HÄHNEL D, BURGARD W, THRUN S. Learning compact 3d models of indoor and outdoor environments with a mobile robot [J]. Robotics and Autonomous Systems, 2003, 44(1):15-27.
- [15] COUGHLAN J M, YUILLE A L. Manhattan world: Orientation and outlier detection by bayesian inference [J]. Neural Computation, 2003, 15(5):1063-1088.
- [16] LI M, WONKA P, NAN L. Manhattan-world urban reconstruction from point clouds[C]//Proceedings of the ECCV. 2016.
- [17] NAN L, WONKA P. Polyfit: Polygonal surface reconstruction from point clouds[C]//Proceedings of the ICCV 2017. 2017.
- [18] COUDRON I, PUTTEMANS S, GOEDEMÉ T. Polygonal reconstruction of building interiors from cluttered pointclouds [C]//Proceedings of the European Conference on Computer Vision. 2018.
- [19] ZHU L, SHEN S, GAO X, et al. Urban scene vectorized modeling based on contour deformation [J]. ISPRS International Journal of Geo-Information, 2020, 9(3):162.
- [20] IKEHATA S, YANG H, FURUKAWA Y. Structured indoor modeling[C]// Proceedings of the IEEE International Conference on Computer Vision. 2015.
- [21] Piecewise-planar reconstruction of multi-room interiors with arbitrary wall arrangements [J]. Computer Graphics Forum, 2016, 35(7):179-88.
- [22] BASSIER M, VERGAUWEN M. Clustering of wall geometry from unstructured point clouds using conditional random fields [J]. Remote Sensing, 2019, 11(13):1586.
- [23] NIKOOHEMAT S, DIAKITÉ A, LEHTOLA V, et al. Consistency grammar for 3d indoor model checking [J]. Transactions in GIS, 2020, 25(1):189-212.
- [24] CUI Y, LI Q, YANG B, et al. Automatic 3-d reconstruction of indoor environment with mobile laser scanning point clouds [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, PP(99):1-14.
- [25] COUDRON I, PUTTEMANS S, GOEDEMÉ T, et al. Semantic extraction of permanent structures for the reconstruction of building interiors from point clouds [J]. Sensors, 2020, 20(23):6916.
- [26] LIM G, DOH N. Automatic reconstruction of multi-level indoor spaces from point cloud and trajectory [J]. Sensors, 2021, 21(10):3493.
- [27] LIU Y, D'ARONCO S, SCHINDLER K, et al. Pc2wf: 3d wire-frame reconstruction from raw point clouds [J]. arXiv: 2103.02766, 2021.
- [28] WU Y, XUE F. Floorpp-net: Reconstructing floor plans using point pillars for scan-to-bim [J]. arXiv:2106.10635, 2021.
- [29] WANG C, DAI Y, EL-SHEIMY N, et al. Progress on isprs benchmark on multisensory indoor mapping and positioning [J]. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci., 2019, 42:1709-1713.
- [30] KHOSHELHAM K, VILARINO L D, PETER M, et al. The isprs benchmark on indoor modelling [J]. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci., 2017, 42(2):W7.
- [31] SONG S, YU F, ZENG A, et al. Semantic scene completion from a single depth image[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [32] SONG S, LICHTENBERG S P, XIAO J. Sun rgb-d: A rgb-d scene understanding benchmark suite[C]//IEEE Conference on Computer Vision & Pattern Recognition. 2015.
- [33] DAI A, CHANG A X, SAVVA M, et al. Scannet: Richly-annotated 3d reconstructions of indoor scenes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.
- [34] SILBERMAN N, FERGUS R. Indoor scene segmentation using a structured light sensor[C]//2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops). 2015.
- [35] ARMENI I, SAX S, ZAMIR A R, et al. Joint 2d-3d-semantic data for indoor scene understanding [J]. arXiv:1702.01105, 2017.
- [36] CHANG A, DAI A, FUNKHOUSER T, et al. Matterport3d: Learning from rgb-d data in indoor environments [C]//2017 International Conference on 3D Vision(3DV). 2017.
- [37] OCHMANN S, VOCK R, WESSEL R, et al. Automatic reconstruction of parametric building models from indoor point clouds [J]. Computers & Graphics, 2016, 54:94-103.
- [38] BALLARD D H. Generalizing the hough transform to detect arbitrary shapes [J]. Pattern Recognition, 1981, 13(2):111-122.
- [39] TURNER E, ZAKHOR A. Floor plan generation and room labeling of indoor environments from laser range data[C]//Proce-

- dings of the International Conference on Computer Graphics Theory & Applications, 2015.
- [40] OESAU S, LAFARGE F, ALLIEZ P. Indoor scene reconstruction using feature sensitive primitive extraction and graph-cut [J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2014, 90(APR.): 68-82.
- [41] MONSZPART A, MELLADO N, BROSTOW G J, et al. Rapter: Rebuilding man-made scenes with regular arrangements of planes [J]. *ACM Trans Graph*, 2015, 34(4): 1-12.
- [42] LIN C, LI C, WANG W. Floorplan-jigsaw: Jointly estimating scene layout and aligning partial scans [C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
- [43] PREVITALI M, DÍAZ-VILARIO L, SCAIONI M. Indoor building reconstruction from occluded point clouds using graph-cut and ray-tracing [J]. *Applied Sciences*, 2018, 8(9): 1529.
- [44] WANG R, XIE L, CHEN D. Modeling indoor spaces using decomposition and reconstruction of structural elements [J]. *Photogrammetric Engineering & Remote Sensing*, 2017, 83(12): 827-841.
- [45] BASSIER M, VERGAUWEN M. Topology reconstruction of bim wall objects from point cloud data [J]. *Remote Sensing*, 2020, 12(11): 1800.
- [46] SCHINDLER G, DELLAERT F. Atlanta world: An expectation maximization framework for simultaneous low-level edge grouping and camera calibration in complex man-made environments [C]// *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004)*. IEEE, 2004.
- [47] NIKOOHEMAT S, DIAKITÉ A, ZLATANOVA S, et al. Indoor 3d modeling and flexible space subdivision from point clouds [J]. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2019, 4: 285-292.
- [48] MURALI S, SPECIALE P, OSWALD M R, et al. Indoor scan2bim: Building information models of house interiors [C]// *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2017.
- [49] OCHMANN S, VOCK R, KLEIN R. Automatic reconstruction of fully volumetric 3d building models from point clouds [J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2019, 151: 251-262.
- [50] AGGARWAL C C, WANG H. A survey of clustering algorithms for graph data [M]// *Managing and Mining Graph Data*. Springer, 2010: 275-301.
- [51] TRAN H, KHOSHELHAM K, KEALY A, et al. Extracting topological relations between indoor spaces from point clouds [J]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, 4: 401-406.
- [52] GUO L, FAN G, SHENG W. Dual graphical models for relational modeling of indoor object categories [C]// *2019. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2019.
- [53] YU L, LI X, FU C W, et al. Ec-net: An edge-aware point set consolidation network [C]// *European Conference on Computer Vision*. 2018.
- [54] WANG X, XU Y, XU K, et al. Pie-net: Parametric inference of point cloud edges [J]. *arXiv:2007.04883*, 2020.
- [55] HÜBNER P, WEINMANN M, WURSTHORN S, et al. Pose normalization of indoor mapping datasets partially compliant to the manhattan world assumption [J]. *arXiv:2107.07778*, 2021.
- [56] DAI A, RITCHIE D, BOKELOH M, et al. Scancomplete: Large-scale scene completion and semantic segmentation for 3d scans [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [57] LAWIN F J, DANELLJAN M, TOSTEBERG P, et al. Deep projective 3d semantic segmentation [C]// *Proceedings of the International Conference on Computer Analysis of Images and Patterns*. Springer, 2017.
- [58] TCHAPMI L, CHOY C, ARMENI I, et al. Segcloud: Semantic segmentation of 3d point clouds [C]// *Proceedings of the 2017 International Conference on 3D Vision (3DV)*. 2017.
- [59] QI C R, SU H, MO K, 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*. 2017.
- [60] YU L, LI X, FU C W, et al. Pu-net: Point cloud upsampling network [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [61] VOLK R, STENGEL J, SCHULTMANN F. Building information modeling (bim) for existing buildings—literature review and future needs [J]. *Automation in Construction*, 2014, 38: 109-127.
- [62] FURUKAWA Y, CURLESS B, SEITZ S M, et al. Manhattan-world stereo [C]// *Proceedings/CVPR*. 2009.
- [63] STRAUB J, ROSMAN G, FREIFELD O, et al. A mixture of manhattan frames: Beyond the manhattan world [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
- [64] PINTORE G, PINTUS R, GANOVELLI F, et al. Recovering 3d existing-conditions of indoor structures from spherical images [J]. *Computers & Graphics*, 2018(10): 77.
- [65] PINTORE G, GANOVELLI F, VILLANUEVA A J, et al. Automatic modeling of cluttered multi-room floor plans from panoramic images [C]// *Proceedings of the Computer Graphics Forum*. 2019.
- [66] COUGHLAN J M, YUILLE A L. Manhattan world: Compass direction from a single image by bayesian inference [C]// *Proceedings of the Seventh IEEE International Conference on Computer Vision*. IEEE, 1999.
- [67] LIU Y, CHEN G, KNOLL A. Globally optimal vertical direction estimation in atlanta world [J]. *arXiv:1904.12717*, 2019.
- [68] JOO K, OH T H, KWEON I S, et al. Globally optimal vertical direction estimation in atlanta world [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(4): 1949-1962.
- [69] ZOU C, COLBURN A, SHAN Q, et al. Layoutnet: Reconstructing the 3d room layout from a single rgb image [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision & Pattern Recognition*. 2018.
- [70] MURA C, MATTAUSCH O, PAJAROLA R. Reconstructing complex indoor environments with arbitrary wall orientations [C]// *Proceedings of the Proc Eurographics Posters*. 2014.

- [71] SCHWING A G, FIDLER S, POLLEFEYS M, et al. Box in the box: Joint 3d layout and object reasoning from single images [C]// Proceedings of the IEEE International Conference on Computer Vision. 2013.
- [72] YANG Y, SHI J, LIU R, et al. Automatic 3d indoor scene modeling from single panorama [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018.
- [73] ZHANG W, ZHANG W, GU J. Edge-semantic learning strategy for layout estimation in indoor environment [J]. IEEE Trans. Cybern. , 2020, 50(6): 2730-2739.
- [74] TANG L, LI L, YING S, et al. A full level-of-detail specification for 3d building models combining indoor and outdoor scenes [J]. International Journal of Geo-Information, 2018, 7(11): 419.
- [75] LANG A H, VORA S, CAESAR H, et al. Pointpillars: Fast encoders for object detection from point clouds [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.
- [76] ZHANG C, XU S, JIANG T, et al. Integrating normal vector features into an atrous convolution residual network for lidar point cloud classification [J]. Remote Sensing, 2021, 13(17): 3427.
- [77] VANCLOOSTER A, NICO V, MAEYER P D. Integrating indoor and outdoor spaces for pedestrian navigation guidance: A review [J]. Transactions in GIS, 2016, 20(4): 491-525.
- [78] CHEN S, TIAN D, FENG C, et al. Fast resampling of three-dimensional point clouds via graphs [J]. IEEE Transactions on Signal Processing, 2017, 66(3): 666-681.
- [79] KOCH T. Automatic alignment of indoor and outdoor building models using 3d line segments [C]// Proceedings of the Computer Vision & Pattern Recognition Workshops. 2016.
- [80] WANG C, HOU S, WEN C, et al. Semantic line framework based indoor building modeling using backpacked laser scanning point cloud [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 143(SEP.): 150-166.
- [81] HONG S, JUNG J, KIM S, et al. Semi-automated approach to indoor mapping for 3d as-built building information modeling [J]. Computers Environment and Urban Systems, 2015, 51: 34-46.
- [82] GALLIANI S, LASINGER K, SCHINDLER K. Massively parallel multiview stereopsis by surface normal diffusion [C]// Proceedings of the IEEE International Conference on Computer Vision. 2015.
- [83] KONTARINIS A, ZEITOUNI K, MARINICA C, et al. Towards a semantic indoor trajectory model: Application to museum visits [J]. Geoinformatica, 2021, 25(2): 311-352.
- [84] ZHU Q, LI Y, XIONG Q, et al. Indoor multi-dimensional location gml and its application for ubiquitous indoor location services [J]. ISPRS International Journal of Geo-Information, 2016, 5(12): 220.



REN Fei, born in 1998, postgraduate. His main research interests include 3D reconstruction and deep learning.



CUI Yan, born in 1984, professor. His main research interests include computer vision and computer graphic.