

面向室内空间智能的三维场景图表达与应用

汤圣君^{1,2}, 杜思齐², 王伟玺^{1,2}, 郭仁忠^{1,2}

1. 深圳大学亚热带建筑与城市科学全国重点实验室, 广东 深圳 518061;

2. 深圳大学建筑与城市规划学院, 广东 深圳 518061

收稿日期: 2023-10-20; 修回日期: 2024-06-12

中图分类号: P208

文献标识码: A

文章编号: 1001-1595(2024)07-1355-16

基金项目: 广东省自然科学基金(2024A1515030061); 深圳市科技计划项目(KJZD20230923115508017; JCYJ20210324093012033);

亚热带建筑与城市科学全国重点实验室自主研究课题(2023ZB18)

第一作者简介: 汤圣君(1991—), 男, 博士, 副研究员, 主要研究方向为城市三维要素结构化重建、多传感器融合测图等。

E-mail: shengjuntang@szu.edu.cn

通信作者: 王伟玺 E-mail: wangwx@szu.edu.cn

摘要: 现有室内三维场景表达聚焦于对象化的描述方法, 其要素表达还停留在对象级语义理解层面, 欠缺对室内场景复杂关系信息的显示表达。面向室内空间智能任务需求, 亟需一个能够完整、准确描述室内要素几何、语义及关系, 且具备语义检索和分析推理能力的结构化模型支撑。基于三维场景图基础理论, 本文创新性地提出面向室内空间智能的三维场景图表达模型, 系统性介绍了室内三维场景图要素层级化组织、几何表达、语义描述、关系描述方法, 建立了一个可统一描述室内要素几何、语义及关系的室内三维场景图概念模型。同时, 该图模型可以与现有的三维场景表达方法进行融合表达, 具有良好的数据兼容性。最终, 本文基于公开的 IFC 模型构建了完整且具有多层级关系信息的三维场景图模型, 并且结合大语言模型通过复杂场景检索、拓扑分析等应用对该模型的应用能力、潜力及局限性进行了系统性探讨和分析。结果表明, 室内三维场景图模型具备复杂计算和分析能力, 可直接与大语言模型集成, 通过简单的自然语言提示实现复杂的场景分析应用。

关键词: 室内建模; 图模型; IFC; CityGML; 大语言模型

近年来, 我国大力推行实景三维中国的建设, 自然资源部在 2021 年 9 月印发的《自然资源三维立体时空数据库主数据库设计方案》明确, 实景三维中国的三维数据构成全国统一的三维空间框架, 方案中明确规定了部件级实景三维建模包含室内要素的详细信息, 从室外建模到室内精细化重建是城市实景三维建模的趋势也是支撑精细化城市管理的重要基础。应用方面, 室内场景结构化模型是机器人智能任务规划和执行的核心基础, 在机器人智能任务中起到人-机-环境交互的桥梁作用。它是将现实物理世界中的物体形态、属性、功能及相互关

系映射到数字空间, 形成几何、语义和关系的准确描述, 该模型用于更加有效地理解、规划、执行室内智能任务^[1-3]。现有的室内三维结构化模型, 如 IFC (industry foundation classes)^[4]、OGC CityGML^[5]、IndoorGML^[6]等主要用于室内空间数据管理和交互, 其模型构建方式是将语义描述挂接到要素几何中, 并且通过面向对象关系表达对要素之间的关系进行间接描述。然而, 现有标准室内三维模型是以要素几何为基础载体, 单体模型数据量大, 在实际应用过程中一般需要借助专业的可视化和分析软件才能进行分析应用, 在轻量化集成方面存在不

足。同时,现有室内三维模型标准在细粒度要素表达方面还存在不足,且关系表达局限于层级和包含关系的表达,难以支撑深层次的语义关联分析应用。可以说,室内结构化建模并不是所有信息都要事无巨细地实景三维表达,基于“按需、动态”的智能化建构图理念,模型要能反映事物的本质特征,把那些在分析计算中并无积极作用的细节和机制隔离出来^[7-8]。因此,面向轻量化数据集成及深层次关联分析应用需求,亟需一个能够统一描述室内要素几何、语义及关系,且具备轻量化集成能力及深层次关联分析能力的结构化模型。

随着三维传感器技术发展,我国在智慧城市建设过程中采集了大量的室内三维场景数据^[9-11],形式可分为多视图像、三维点云、三角网格或者体素,但这些非结构化数据描述方式难以直接用于机器人智能任务规划^[12-13]。为了更高层次地理解室内场景,一些研究学者借助深度学习方法构建了多种任务感知框架,包括三维目标检测^[14-16]、实例化分割^[17-19]、三维形状预测^[20]、语义分割和分类等^[21-24]。特别是文献^[25—26]提出的全景分割方法可以同时像素进行分类并赋予实例标签。随着深度神经网络技术的进步,以对象为中心的语义理解任务和基于现有模型的各种商业化应用都取得了重要突破^[27-28]。然而,这些方法仅对场景要素进行实例分割和分类,还停留在对象级语义理解层面,并未系统性考虑室内场景的结构化描述方法。此外,一些研究学者致力于构建室内结构化三维模型,并探索了多种方法,包括室内平面图恢复^[29-31]、矢量模型构建^[32-36]、CityGML^[5] 标准化模型构建^[36-37] 及 IFC 模型重建^[38] 等。这些方法可以自动提取建筑结构要素(如墙面、天花板、梁、柱等)或恢复楼层空间布局,可用于火灾疏散^[39-42]、室内导航^[43-45] 等位置服务应用。但是,这类方法主要关注建筑物结构要素的提取和重建,缺乏对室内细粒度要素信息和深层次要素关系的理解和表达。可以说,测绘科学和人工智能学科在各自领域为室内场景语义理解、结构化表达与重建等提供了重要的原理和方法借鉴,但这些研究在场景要素统一组织、深层次关系信息描述,以及支撑深度计算分析能力等方面还存在很大的不足,尤其是现有室内结构化模型采用层级组织方式,其要素关系通过继承或聚合等方式隐式表达,难以直接支撑室内复杂智能应用中语义检索和分析推理的需求。

近年来,场景图已经成为编码图像语义和上下文信息的重要表达方式。场景图通常由相互连接

的节点组成的稠密语义网络进行表达。它在图像描述、视觉问答和图像生成或编辑方面具有优越的分析能力^[46]。与此相比,三维场景图可以更好地抽象表达三维场景语义信息。它的“节点”可以表示三维场景中的要素,而“边”则用来表示这些要素之间的相互关系。与现有的三维数据表达方法相比,三维场景图具有许多优势:①它是一种高度可读性和更加紧凑的场景表达方式,能够描述室内空间要素的几何、语义,以及它们之间的关系,是对室内场景的一种高维度抽象,可满足轻量化场景信息集成应用并与现有的大语言模型进行集成交互;②图模型显示表达了场景中各类要素详细语义及关系信息,可直接支撑更深层次的场景分析和计算需求;③三维场景图是将各类要素抽象为独立的节点并赋予特定的语义信息,进而通过关系描述对各个节点进行关联。因此,场景图中的节点可与现有的室内三维结构化模型中要素模型进行映射与关联,实现不同数据结构的有效融合,该模型具有良好的数据兼容性。可以看出,三维场景图在场景几何、语义和关系的融合表达,以及支撑场景语义分析和推理方面具有天然的优势,可为室内三维场景理解和结构化表达提供更加深入的视角,具备支撑轻量化信息集成及深层次场景分析任务的潜力。到目前为止,室内三维场景图尚未有统一的定义和表示方法^[47]。

本文将围绕室内三维场景图要素分类与组织、要素几何与语义提取、要素关系重建等方面展开研究,提出一种面向室内要素几何、语义和关系为一体的三维场景图表达与结构化重建方法,主要包含以下 3 点贡献。

(1) 在理论层面,系统性研究“室内三维场景图”中要素分类与组织、几何和语义表达及关系描述,建立完整、统一的“室内三维场景图”模型,这将在概念层面实现室内三维场景结构化表达从对象级提升到知识级。

(2) 在方法层面,基于室内要素分割、语义分类、关系推理等多层次算法,提出室内“实体要素”与“空间要素”一体的多层级室内三维场景图重建方法。

(3) 基于自然语言提示与大语言模型探讨室内三维场景图的智能化应用场景,建立基于自然语言的室内三维场景计算和分析逻辑框架。

1 室内三维场景图表达框架

1.1 三维场景图表达方法

采用三维场景图表达方法,可将现实室内三维

场景 S 转化为场景图模型,其场景图 T_S 是由多个带有几何 G_S 和语义信息 A_S 的要素节点 F_S 和各节点之间关系表达的边集合 R_S 组成。

$F_S = \{f_{S,1}, \dots, f_{S,i}, \dots, f_{S,n}\}$ 表示图模型中的要素集合, $f_{S,i}$ 则表示第 i 个要素节点,这些节点可以是不同粒度的实体要素(在物理空间中真实存在实体)或者具有一定语义内涵的非实体要素。同时每个节点可通过几何和语义信息进行描述。 $G_S = \{g_{S,1}, g_{S,2}, \dots, g_{S,i}, \dots, g_{S,n}\}$ 表示图模型中所有要素的几何信息集合, $g_{S,i}$ 则表示第 i 个要素的几何表达,该几何表达可以是二维的点、线、面或者三维的几何描述。 $A_S = \{a_{S,f_1,1}, a_{S,f_1,2}, \dots, a_{S,f_1,k_1}, a_{S,f_2,1}, a_{S,f_2,2}, \dots, a_{S,f_2,k_2}, \dots, a_{S,f_i,1}, a_{S,f_i,2}, \dots, a_{S,f_i,k_i}, \dots, a_{S,f_n,1}, a_{S,f_n,2}, \dots, a_{S,f_n,k_n}\}$ 表示图模型中所有要素的语义信息集合, $a_{S,f_i,j}$ 则表示要素 $f_{S,i}$ 的第 j 个语义信息, $k_i \geq 0$ 且 $j \in \{1, 2, \dots, k_i\}$ 。 $R_S = \{r_{S,1 \rightarrow 2}, r_{S,1 \rightarrow 3}, r_{S,n-1 \rightarrow n}\}$ 表示三维图模型中的关系集合, $r_{S,i \rightarrow j}$ 则表示由第 i 个要素 $f_{S,i}$ 与第 j 个要素 $f_{S,j}$ 组成的关系三元组 $\text{Tri}_{S,i \rightarrow j} = \langle \text{sub}_{S,i}, r_{S,i \rightarrow j}, \text{obj}_{S,j} \rangle$, $\text{sub}_{S,i}$ 与 $\text{obj}_{S,j}$ 分别表述关系表达中的主体和客体。如“桌子”为 $\text{sub}_{S,i}$, “水杯”为 $\text{obj}_{S,j}$, 桌子和水杯关系为 $r_{S,i \rightarrow j}$ “在上面”, 基于该三元组可涵盖水杯在桌子上面的自然语言描述信息, 在三维场景图中, 一个室内要素可以和多个不同的要素组成关

系三元组,该三元组能够准确表达不同要素之间的关联信息,可支撑室内场景各类分析应用。

1.2 室内三维场景图表达模型

室内三维场景图模型是需要将现实物理世界中的物体形态、属性、功能及相互关系映射到数字空间,形成全要素几何、语义和关系的准确描述。一个可靠的室内三维场景图模型应该能够准确地表征客观环境,满足室内复杂任务执行时场景理解和分析推理需求。该模型应该具备以下特点。

- (1) 准确性:能够精确描述场景中的要素及其关系信息。
- (2) 适用性:能够适用于多种类型的复杂机器人智能任务。
- (3) 易用性:使用过程简单,能够直接支撑计算和分析应用。
- (4) 可扩展性:具备多尺度场景描述能力,且支持动态扩展。

以室内智能化任务需求为牵引,顾及准确性、适用性、易用性及可扩展性的基本要求,本文分别从要素分类及结构组织层、几何和语义描述层及关系抽象层入手,构建室内三维场景图概念模型,如图 1 所示。

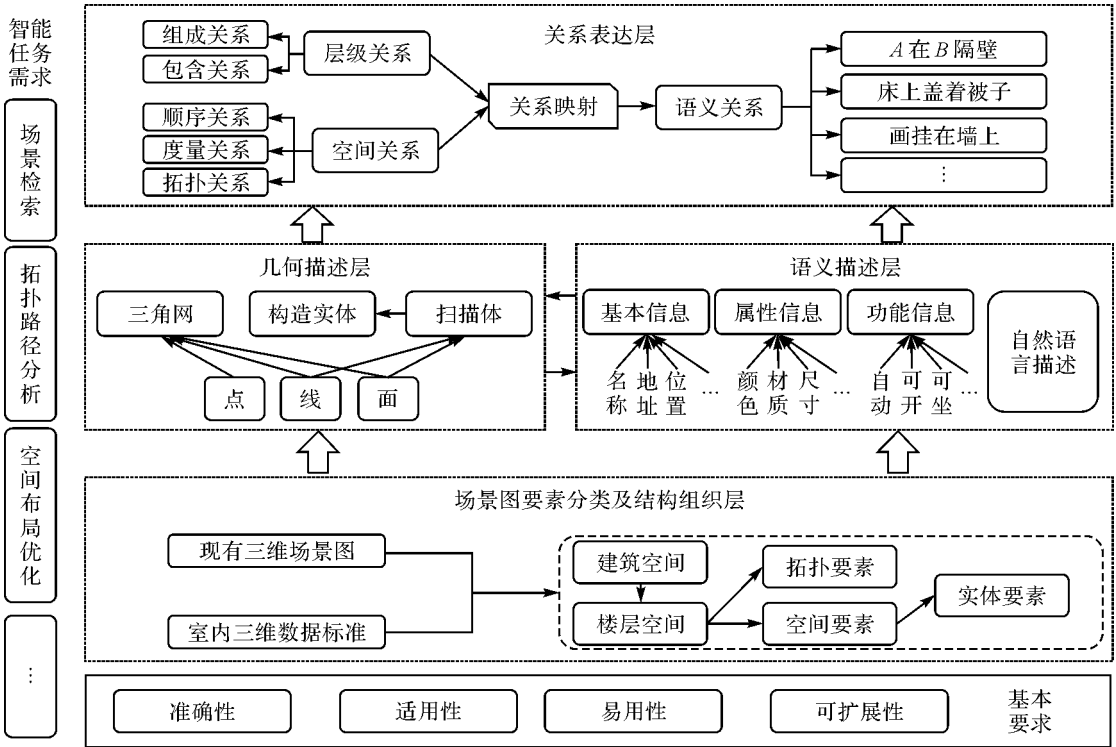


图 1 室内三维场景图概念模型

Fig. 1 Conceptual model of indoor 3D scene graph

2 室内三维场景图要素组织

2.1 场景要素分类及结构组织层

在 OGC CityGML^[25] 和 ISO 19164^[37] 等室内三维数据标准中对室内空间进行了层级化的划分,如楼栋中包含楼层、楼层中包含空间等,但这些标准未明确定义在不同层级中需要包含哪些要素且缺乏对细粒度实体要素的信息表达。在本研究中,以机器人、人、环境交互关系为基本依据,参考现有室内三维数据标准的层级化要素组织逻辑,对三维场景图要素进行层级化划分并对不同层级下要素节点明确定义,建立面向三维场景图构建的室内三维要素分类与层级化组织框架。如图 2 所示,本框架将室内三维场景图要素按照如下层级规则进行组织:复杂建筑要素→单体建筑要素→建筑楼层要素→功能空间要素、连接空间要素→构造实体要素、附属实体要素。这样的层级化组织有助于准确和系统地表达室内三维场景图中的要素信息。需要注意的是,考虑到综合体建筑要素可能包含多个单体建筑要素,而不同的单体建筑要素之间可能会采用室外通道进行关联。因此,本文采用室外空间要素对该类信息进行定义,该要素可表达连接各个单体建筑要素的通道或者空中花园等。

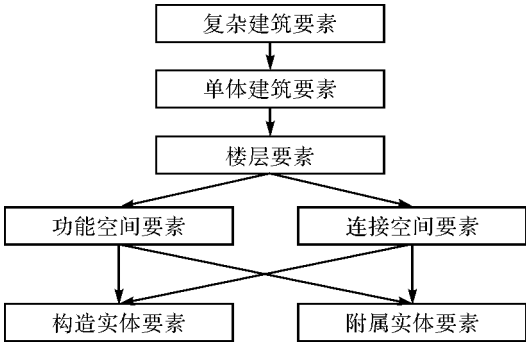


图 2 室内三维场景图要素层级化组织
Fig. 2 Hierarchical organization of indoor 3D scene graph

基于以上分类体系,对场景图要素的层级以及对应的描述信息进行定义,具体见表 1。

2.2 几何描述层

精确的几何表达是室内空间分析的基本前提,现有三维场景图模型一般采用不规则三角网或者三维坐标直接描述场景图中的实体要素的几何信息,未对要素进行规则化建模,上述类型的几何描述方法仅能恢复要素之间简单的空间关系,无法支撑要素之间深层次关系计算,以及要素语义的精准计算(如,“房间是由哪几面墙构成”“墙面的精确长

宽及厚度”)。本文将基于场景图要素分类与层级化组织概念框架,根据原始数据类型及要素特点对各类要素节点的几何进行清晰的定义。

表 1 各层级要素类型描述

Tab. 1 Description of element types at each hierarchical level

级别	描述
复杂建筑要素	用于描述一个建筑群,一座建筑可以跨越几个相连或不相连的建筑,组成一个建筑群,它是由一个以上的独立建筑组成的。
单体建筑要素	用于描述一个单体建筑,一个单体建筑至少包含一个楼层要素,建筑中不同楼层可能具有特定用途或所有权。
楼层要素	用于描述单个楼层空间,每个楼层由各种室内空间要素和室内实体要素组成。
功能空间要素	用于描述具有室内空间功能属性的结构信息,可以包含房间、阳台、走廊等功能空间。
连接空间要素	用于描述连接不同楼层的要素信息,可以包含楼梯、电梯等。
构造实体要素	用于描述建筑的固有组成部分的要素信息,包含墙、楼板、梁、柱、门、窗等要素。
附属实体要素	用于描述附属于建筑物内特定空间的要素信息,包含家具、设施、传感器等要素。

随着三维传感器技术的不断发展,获取室内三维场景数据的方法日益多样化且便利化。感知平台获取的原始三维数据可通过三维点云、三角网格等非结构化数据表达方法进行存储。基于以上非结构化数据,采用要素分割、语义和关系提取等方法可实现室内三维场景图模型的智能化构建。因此,在几何表达层面,所有类型的要素都可直接基于点云数据进行表达,进而根据要素特点和场景建模能力,可不断细化为不规则三角网、构造实体几何以及扫描体等几何描述。如表 2 所示,由于所有要素都具有占据空间,因此都可以采用包围盒描述其空间占据信息,每类要素的几何表达方式的详细定义如下。

(1) 复杂建筑要素:复杂建筑要素包含要素信息繁多,类型多样,其包含的要素中可能存在不同类型的几何表达方式。在对复杂建筑要素的几何表达时,需要综合考虑建筑物内部其他要素的几何表达类型,并进行数据集成和融合,形成完整的复杂建筑要素几何。因此,复杂建筑要素的几何表达形式可覆盖三维点云、不规则三角网、构造实体几何及扫描体等多类型表达方式。

(2) 单体建筑要素:与复杂建筑要素几何表达类似,单体建筑要素包含楼层要素、空间要素及实体要素,其表达方式具有多样性,可覆盖三维点云,不规则三角网、构造实体几何及扫描体等多类型表达方式。

(3) 楼层要素:单个楼层中包含空间要素和实体要素,而空间要素表达可采用三维点云、构造实体几何及扫描体等多类型几何表达,实体要素则涉及三维点云和不规则三角网两种类型的几何表达,因此楼层要素的表达可覆盖三维点云、不规则三角网、构造实体几何及扫描体等多类型表达方式。

(4) 空间要素(功能空间要素、构造实体要素):空间要素类型一般具有规则形状,如房间、电梯等,可基于原始点云自动化构建规则模型,可形成构造实体几何、扫描体等规则描述。一般可采用三维点

云、构造实体几何及扫描体进行几何表达。

(5) 构造实体要素:构造实体包含墙面、梁等建筑规则构件,基于点云数据可自动化构建规则模型,一般采用三维点云、构造实体几何、扫描体等进行几何表达。

(6) 附属实体要素:附属实体要素包含各种类型家具、传感设施等信息,难以直接通过点云进行规则化重建,一般采用三维点云或者不规则三角网进行几何表达。

表 2 室内三维场景图要素几何表达

Tab. 2 Geometric representation of indoor 3D features

室内要素节点类型	三维点云	不规则三角网	构造实体几何	扫描体	包围盒	示例
复杂建筑要素	✓	✓	✓	✓	✓	
单体建筑要素	✓	✓	✓	✓	✓	
楼层要素	✓	✓	✓	✓	✓	
功能空间要素	✓		✓	✓	✓	
连接空间要素	✓		✓	✓	✓	
构造实体要素	✓		✓	✓	✓	
附属实体要素	✓	✓			✓	

2.3 语义描述层

三维场景图要素的语义描述是指对通过基本信息(如名称、位置、地址)、属性信息(如颜色、材质)及功能信息(如可坐、可打开)等对场景图中的不同类型节点进行详细的信息描述。现有的三维场景图模型皆采用统一、一致的属性对节点的语义进行表达,忽略了不同类型、不同层级节点在语义描述上的差异,同时对不同属性的数据类型未进行明确的定义,使得三维场景图模型难以在计算机进行统一的结构化存储。本文将充分结合现有室内三维数据标准,包括 OGC CityGML^[25] 和 ISO 19164^[37],并根据三维场景图表达需求,分析要素与要素之间的关联关系,以及各类要素在属性、功能等方面的异同点,对各层级要素的语义进行详细定义。

首先,针对场景图中的所有要素类型对其基本语义进行定义(表 3),要素基本信息包含描述、类别、权属信息,可作为所有场景图要素的基本信息描述。表 4 对复杂建筑要素语义进行了明确定义,在基本信息的基础上,添加了地址及单体建筑数量语义描述,用于有效表达复杂建筑的语义属性。表 5 对单体建筑要素语义进行了明确定义,包含基本信息、地址、地点、使用方法、楼层数、完成日期、访问权限等信息。表 6 对楼层要素语义进行了明确定义,包含楼层编号、高度、地点、楼层功能、访问权限。表 7 则对功能空间与连接空间要素语义进行了明确定义,功能空间与连接空间在语义表达上具有高度相似性,都可通过楼层编号、访问权限、空间状态、面积、体积等信息进行语义描述。与功能空间不同的是,连接空间可采用可通行人数、是否无障

碍两个语义进行补充描述。表 8 对构造实体要素和附属实体要素语义进行了明确定义,其中灰底为附属实体要素特有属性,包含基本信息、楼层编号、位置、长度、宽度、高度、是否可移动、功能等语义属性,其中“是否可移动”及“功能”属于附属实体要素的特定语义描述。

表 3 要素基本信息描述

Tab. 3 Description of fundamental attributes of elements		
序号	要素基本信息 类型	定义
1	描述	对该要素的基本情况 进行文字描述
2	类别	描述该要素的类别
3	权属	描述该要素的权属关系

表 4 复杂建筑要素语义定义

Tab. 4 Semantic definition of complex building elements		
序号	复杂建筑要素 语义类型	定义
1	基本信息	一个建筑群的基本信息,包括名称和所有权
2	地址	一个建筑群的地址
3	单体建筑数量	组成一个建筑群的单个建筑物的数量

表 5 单体建筑要素语义定义

Tab. 5 Semantic definition of single building elements		
序号	单体建筑要素 语义类型	定义
1	基本信息	一个建筑群的基本信息,包括名称和所有权
2	地址	一个建筑群的地址
3	地点	组成一个建筑群的单个建筑物的数量
4	使用方法	组成复杂的各个建筑物的信息
5	楼层数	建筑物内的层数
6	完成日期	建筑物建成的日期
7	访问权限	进入建筑物的限制

表 6 楼层要素语义定义

Tab. 6 Semantic definition of floor elements		
序号	楼层要素语义 类型	定义
1	基本信息	一个楼层的基本信息,包括名称和所有权
2	楼层编号	一层楼的编号或名称
3	高度	一层楼的高度
4	地点	层楼的位置
5	楼层功能	楼层的功能属性
6	访问权限	进入建筑物的限制

表 7 功能空间与连接空间要素语义定义

Tab. 7 Semantic definition of space elements		
序号	空间要素语义 类型	定义
1	基本信息	一个建筑群的基本信息,包括名称和所有权
2	楼层编号	空间要素所在楼层的编号或名称
3	访问权限	访问一个空间要素的限制
4	空间状态	一个室内空间的空间状态
5	面积	室内空间的面积
6	体积	一个空间要素的体积
7	可通行人数	连接空间可通行人数
8	是否无障碍	是否属于无障碍通道或者有无障碍通行设施

表 8 实体要素语义定义

Tab. 8 Semantic definition of entity elements		
序号	实体要素语义 类型	定义
1	基本信息	一个建筑群的基本信息,包括名称和所有权
2	楼层编号	空间要素所在楼层的编号或名称
3	位置	实体要素的位置信息
4	长度	实体要素的长度
5	宽度	实体要素的宽度
6	高度	实体要素的高度
7	是否可移动	描述要素是否具有可移动的能力
8	功能	要素的功能属性

2.4 关系表达层

在三维场景图中,一个节点被看作是一个具有几何和语义表达的要素对象,而要素之间的关系被称为边,对三维场景中要素之间关系的准确、完整的定义是支撑跨节点智能分析的重要基础。现有的三维场景图关系描述主要包含动作关系(如挂在、倚靠等)、支撑关系、相邻关系、比较关系及遮挡关系等,现有研究并未系统性解析三维空间中需要包含哪些关系信息、不同要素之间关系表达的异同。本文室内三维场景要素分类规则以及不同要素几何和语义表达方式,对场景中相同层级要素之间以及不同层级要素之间的关系进行明确定义,依据三维空间分析及场景图理论,可以将其场景图中要素关系分成层级关系、空间关系和语义关系。

2.4.1 层级及数量关系

该关系描述不同层级要素之间的组成和包含关系,如图 3 所示,一个建筑楼层要素是由多个室内空间要素组成,一个室内空间要素可能包含多个附属实体要素。在三维场景图计算过程中层级关系可以支撑要素归属推理的能力,如分析“建筑物 A 中房间 B 内椅

子的个数”。因此在描述层级关系的同时,还内置了要素之间的数量关系,如房间中有多少个椅子、多少个桌

子等,数量关系信息将有助于提升算法对三维场景图中要素分布以及空间中要素组成情况的准确认知。

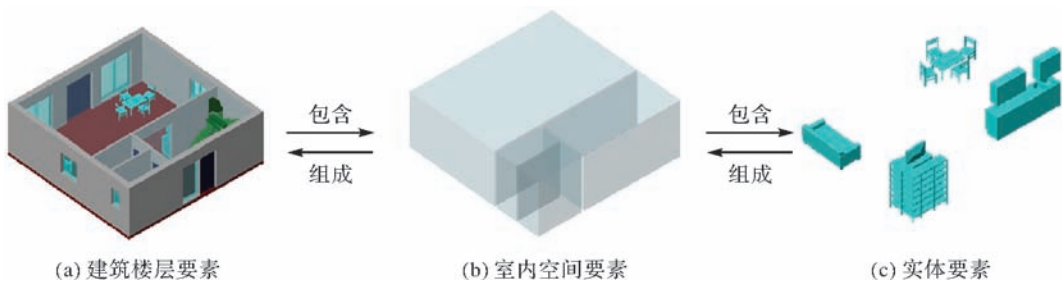


图 3 三维场景图层级关系描述示例

Fig. 3 Inheritance relationship of 3D scene graph

2.4.2 空间关系

依据三维空间分析理论,空间关系是指在三维空间中物体之间的相对位置关系,涉及物体的位置、方向、距离、连通性等方面,可将其分为顺序关系、拓扑关系和度量关系。

(1) 顺序关系。本文将采用二维空间方位关系描述模型之一最小边界矩形模型来描述三维空间方向模型。在描述三维空间的方向的时候,可用以下方

向元素来进行组合描述:上、下、左、右、前、后、同位,或者上、下、东、西、南、北、同位。根据模型所处坐标系不同,可分别选用不同组别的三维空间方向元素来描述三维空间物体的方向。也就是说相对二维空间方向模型,三维空间模型的方向元素增加了上、下两组主方向。如图 4 所示,绿色圆球表示要素质心所在位置,黄色圆球表示方位信息,根据三维空间方向元素进行组合,最终可以组成 27 个方向空间。

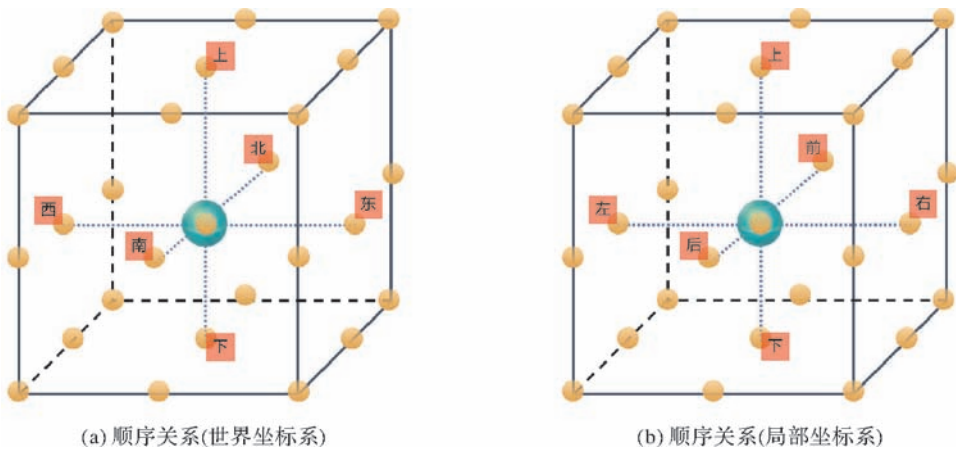


图 4 顺序关系描述

Fig. 4 Direction relationship representation

(2) 拓扑关系。该关系是最基本的空间关系,它体现了空间实体在空间上的一种不依赖于几何形变的内在联系。基于三维空间拓扑分析理论,在三维空间中所有要素几何都可以通过体对象进行表达,即由一组邻接的面对象集合所形成的多面体。文献[48]基于九交关系编码规则对二维点线、二维点线与三维体,以及三维体与三维体之间的空间关系进行了明确定义。根据该定义,三维体与三维体之间的空间关系与二维空间面面相交的空间关系一致,包含 8 种类型的空间关系。如图 5 所示,

三维空间体对象的拓扑关系可以包含相离、相接、进入、相等、包含于、包含、穿越和被穿越,由于相等关系主要表达同一个要素之间的关系,不涉及不同类型要素的关系表达,在本表达模型中不考虑。图 5 表达了三维空间中 7 种类型的拓扑关系示例,图 5(a)中床要素与楼板要素属于相接关系,椅子与床属于相离关系;图 5(b)中楼梯要素与房间要素之间属于进入关系;图 5(c)中窗户要素包含于墙面要素,墙面要素包含窗户要素;图 5(d)中楼梯要素穿越楼层要素,楼层要素则被穿越。

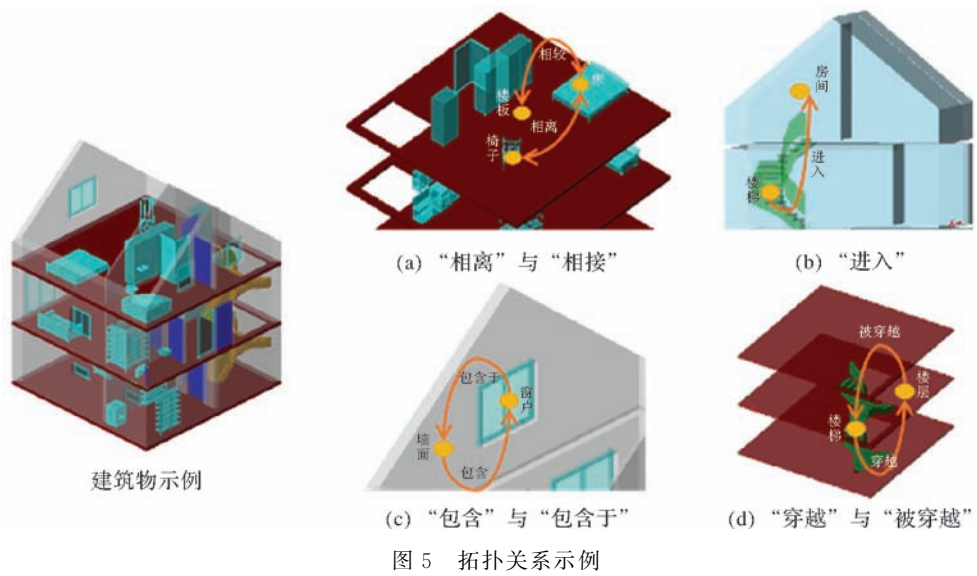


图 5 拓扑关系示例

Fig. 5 Topology relationship representation

(3) 度量关系。该关系是在欧氏空间(Euclidean space)和度量空间(metric space)上进行的操作,它是三维空间数据量化描述的基础。在三维场景中,可以将其分为距离关系和比较关系,距离关系主要描述三维空间中两个要素节点之间的绝对空间距离,如图 6(a)所示,椅子与床的距离为

2.2 m,该关系的主体节点和客体节点在要素组织框架中可以是任意层级中的要素节点;比较关系则是对同类型的要素属性进行比较产生的关系,如图 6(b)所示,楼层 1、楼层 2 和楼层 3 面积相同,该关系的主体节点客体节点在要素组织框架中一般是同层级中的要素节点。

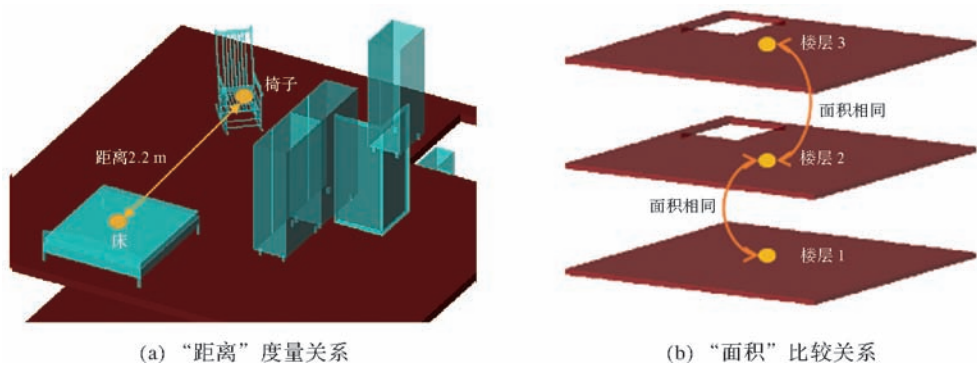


图 6 度量关系

Fig. 6 Metric relationship representation

3 试验分析

3.1 数据介绍与预处理

为验证本文提出的室内三维场景图表达模型的有效性,本文选取国际标准模型 IFC 数据^[49]作为试验数据,该数据来自于 Karlsruhe Institute of Technology (KIT),提供了多个 IFC 标准数据模型,本文以 Jasmin-Sun-105 模型作为基础进行室内三维场景图的构建,并在此基础上基于自然语言进行场景图智能化分析应用。如图 7 所示,Jasmin-

Sun-105 模型具有 3 层建筑结构,包含建筑物、楼层、空间、墙面、门、窗、楼板、屋顶及家具等各种类型的要素信息,且不同的要素信息之间通过对应的关系进行连接。

为了构建室内三维场景图模型,本文基于 IFCOpenShell 开源工具对该模型中要素的几何、语义和属性信息进行提取,且根据各个要素之间的几何结构恢复不同要素之间的层级关系与空间关系。在此基础上,基于 Cypher 语言构建室内三维场景图模型,实现层次化室内三维场景图模型的结构化

重建。

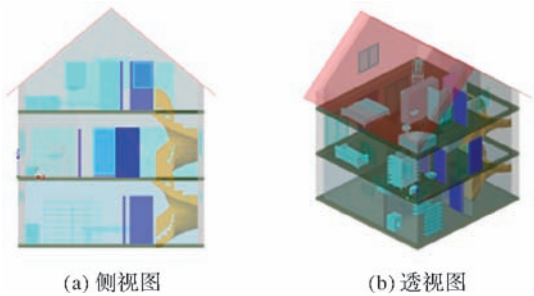


图 7 室内三维模型
Fig. 7 Indoor 3D model

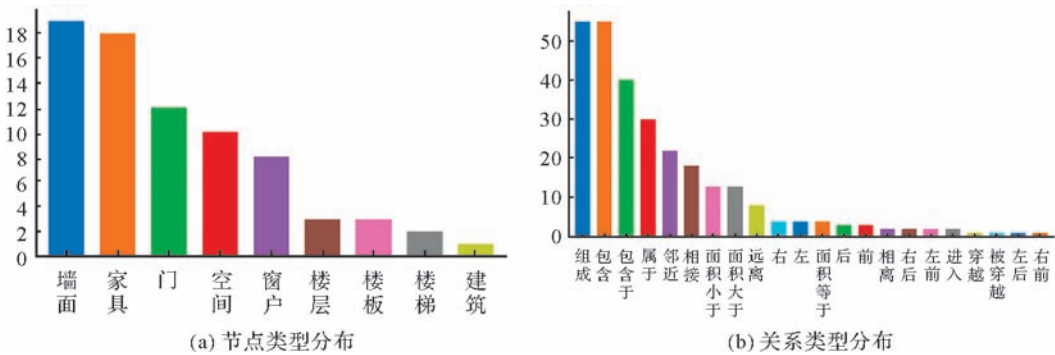


图 8 节点与关系类型分布
Fig. 8 Nodes and edges distribution

如图 9 所示,展示了从建筑→楼层→空间之间的层次化场景图模型,建筑与楼层之间通过“属于”关系进行关联,楼层与空间之间通过“组成”和“包含”关系进行关联,不同的楼层之间共享楼板组件。基于该层次化的描述方式,将墙面、门、窗、家具、楼梯等要素与楼层、空间、要素进行关联,形成完整的

场景图模型。图 10 展示了单个房间的节点与关系图,Keller-2 房间中包含了 4 个墙面,两个门及 3 个家具的信息,各个要素与房间节点之间存在“包含”“组成”等关系,门与房间之间关系采用“包含”和“包含于”进行描述。

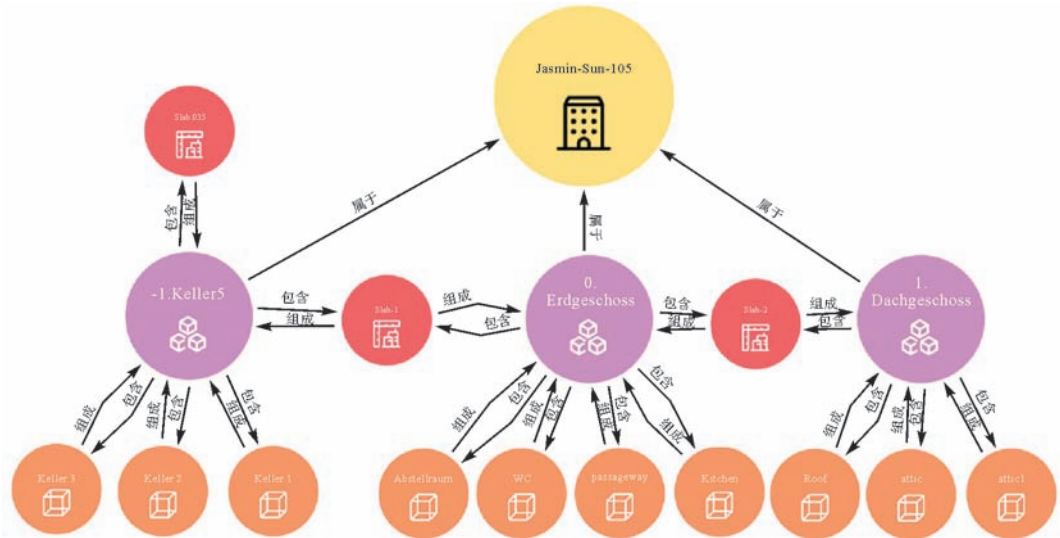


图 9 场景图层次化结构
Fig. 9 Hierarchical 3D scene graph

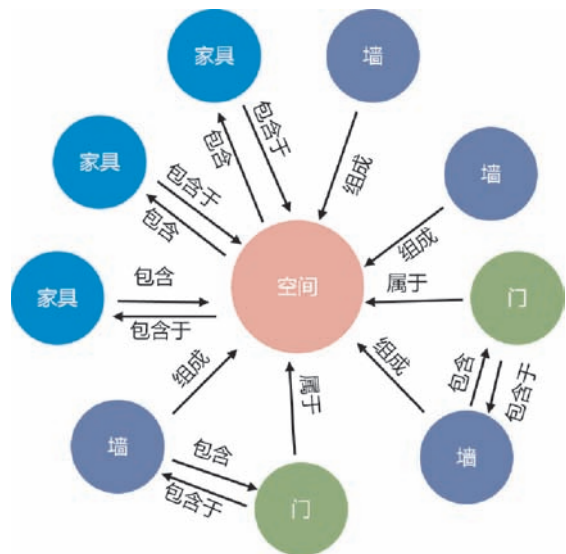


图 10 单个房间节点及关系

Fig. 10 Nodes and edges of single room

3.3 三维场景图分析应用

为了实现良好的数据交互,本文基于 GLM-4 接口开发了 GraphChat 图模型交互框架。基于以上场景图模型,基于该自然语言交互框架进行场景检索、拓扑分析的图应用探索。具体应用过程中,首先将原始场景图中的节点与边信息导入语言模型,进而基于原始的关系信息构建原始关系与自然语言描述之间的一一映射,在此基础上采用自然语

言提问方式进行场景检索、拓扑路径分析等应用,以验证本文提出的室内三维场景图的应用能力。

3.3.1 GraphChat 图模型交互框架

为实现大语言模型与三维场景图的交互与智能分析,研究构建了大语言模型与三维场景图交互分析框架 ChatGraph。该框架具备自主任务规划、任务拆解、连续执行的能力,并能在任务执行过程中对结果提供实时反馈及对后续任务进行修正。ChatGraph 改进了大语言模型在图分析应用中对人类抽象指令的理解、任务执行控制,以及与图数据的交互等关键问题。在任务规划阶段,设计了以引导模型实现自主任务规划与任务关联性分析。随后,为解决复杂任务执行过程中,流程控制与结果的实时反馈的重要问题,提出了一套基于计算图组织的任务执行框架。该框架基于 LLM 任务拆解结果,生成计算图以控制任务分析流程。基于计算图,ChatGraph 能够有序地在图分析工具与 LLM 间传递信息,执行复杂分析任务。最后,为解决 LLM 与 Graph 数据的交互,针对 Neo4j 图数据库与传统关系型数据库分别构建了对应的图分析工具,并使用基于思维链进行函数调用的 Agent 实现语言模型到图数据的交互。ChatGraph 的整体框架如图 11 所示。

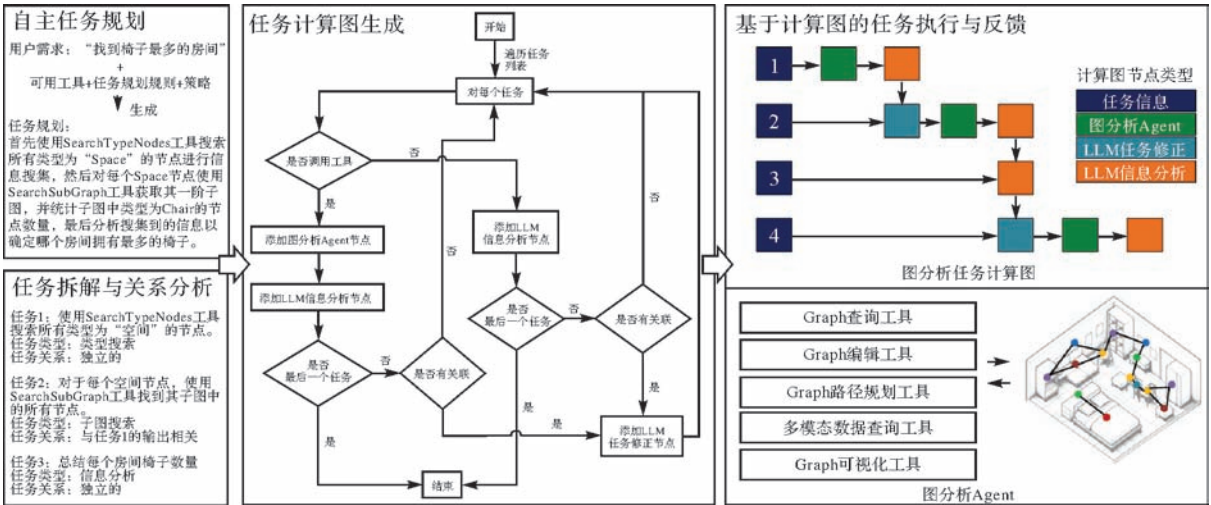


图 11 ChatGraph 的整体框架

Fig. 11 The framework of ChatGraph

(1) 自主任务规划:在任务规划阶段,研究模拟了人类通过逐步搜集信息对场景进行感知的思路,提出了 TPSP(tools, plan and solve prompting)任务规划模板。该模板在输入用户需求的同时向

LLM 提供 3 种信息:① LLM 可调用的工具;② 任务规划的基本规则,要求 LLM 基于能够使用的工具,以先搜集信息,再对信息分析的基本思路生成任务规划;③ 任务执行策略,要求 LLM 使用最少的工具

实现任务,同时要求其合并连续的分析任务。随后 LLM 将基于上述信息生成任务执行的总体规划。

(2) 任务拆解:在获得总体任务规划后,该模块将任务规划拆解为更小、更易管理的子任务。研究通过构建小样本提示词,引导模型思考并输出各个子任务的执行细节,包括任务编号、任务类型、任务内容,以及任务与其他任务的关联性,用于后续计算图生成。

(3) 任务计算图生成:对于具有关联性的任务,后一个任务通常需要前一个任务所收集的信息以明确其执行对象,这需要在连续任务执行阶段及时分析信息,处理其中需要修正的任务。针对这一需求,研究使用计算图规范任务执行流程。ChatGraph 计算图是根据任务拆解结果创建的静态图结构,包含任务信息节点和 3 类任务执行节点。计算图的边表示信息传递方向。在创建时,算法基

于任务拆解得到的任务类型和关联性需求,按顺序向图中添加 Agent,信息分析或任务修正节点,确保每个任务得到适当的工具分配和执行时序,使得在连续任务执行过程中的信息流管理更加高效和准确。

(4) 基于计算图的任务连续执行与反馈:在任务执行阶段,ChatGraph 根据计算图依次调用工具与在不同模块间传递任务信息,对执行结果进行分析或做出反馈。

3.3.2 场景检索分析

场景检索分析是图模型基础分析能力,为了验证本文创建的三维场景图的场景检索能力,本文基于自然语言提问的方式对不同类型的目标进行信息检索。如图 12 所示,本文采用大语言模型加载生成的室内三维场景图并且进行了 4 组场景检索的试验。

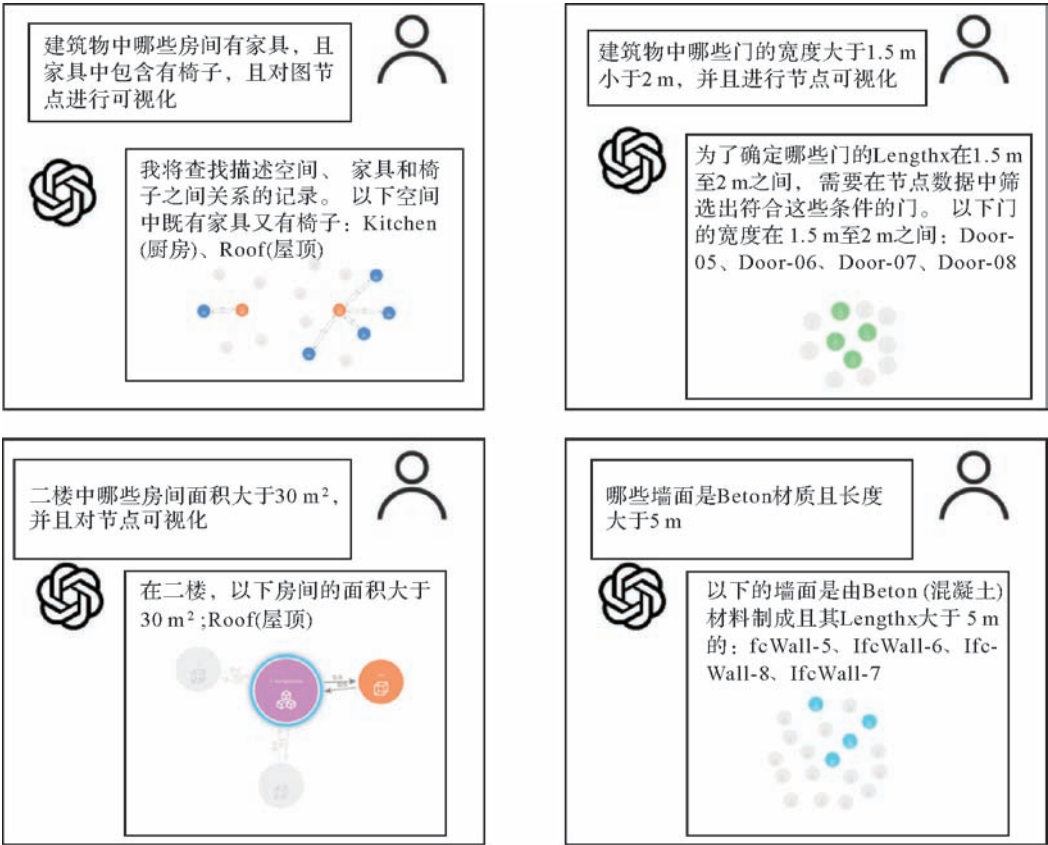


图 12 场景检索分析案例
Fig. 12 Samples of scene retrieval

(1) 基于包含关系检索:采用自然语言提问哪些房间中的家具中包含椅子。GPT 根据场景图检索得到厨房(Kitchen)与屋顶(Roof)空间满足该条件并进行了可视化展示空间节点与椅子节点的关系。

(2) 基于要素尺寸检索:采用自然语言提问哪

些门宽度大于 1.5 m 且小于 2 m。GPT 根据场景图检索到 4 扇门满足该条件且进行了可视化展示。说明通过自然语言的表达可准确实现要素尺寸的计算和分析。

(3) 基于模糊语句查询:采用自然语言提问二楼

哪些房间面积大于 30 m²。然而,在场景图中实际不包含二楼的属性信息,GPT 根据不同楼层之间楼板的高度进行了区分,且正确检索得到“1. Dachgeschoss”楼层中的 Roof 空间面积大于 30 m²。

(4) 组合属性查询:采用自然语言提问哪些墙面材质为 beton 且长度大于 5 m。GPT 正确查找出 4 扇墙面满足该条件,且将 beton 翻译为混凝土,说明通过 GPT 的场景图检索既可以实现组合属性的检索分析,同时可适应多语言环境检索。

基于以上分析可知,通过自然语言的提示可直接对室内三维场景图模型中要素信息及其关系进行检索分析,同时大语言模型在提示词不够明确的

情况下也具有一定的信息纠错能力且能够适应多语言环境的场景检索。

3.3.3 拓扑路径分析

拓扑路径分析是三维空间分析的重要应用之一,是进行路径导航和最短路径分析的重要前提。传统拓扑路径分析需要单独构建室内空间的路径导航地图实现。本文基于三维场景图中的节点语义及关系信息,基于自然语言提问的方式探讨室内三维场景图模型在拓扑路径分析方面的应用能力。如图 13 所示,展示了不同的自然语言提示词进行拓扑路径分析的结果。

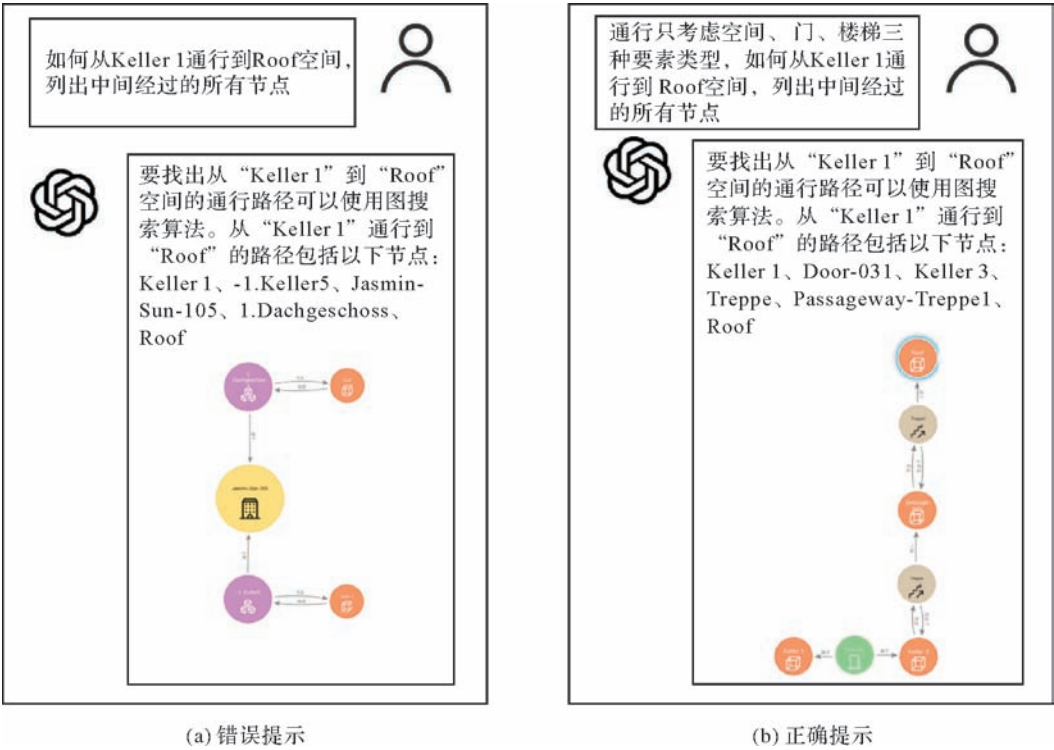


图 13 拓扑路径分析案例
Fig. 13 Samples of topology analysis

(1) 错误提示:如图 13(a)所示,基于已有的室内场景图模型,希望通过自然语言提问找到如何从 Keller 1 空间通行到 Roof 空间中,并且列出经过的所有节点信息。基于该提示,GPT 对图模型进行分析并且进行节点搜索,得到路径节点为:Keller 1(空间)、-1. Keller5(楼层)、Jasmin-Sun-105(建筑)、1. Dachgeschoss(楼层)、Roof(空间)。从节点可视化结果来看,从 Keller 1 到 Roof 确实可通过其余 3 个节点相连并且顺利抵达。然而,GPT 模型并未考虑室内空间行走过程中的规则信息,而是简单地查找最短的路径,导致其获取的路径产生错误。

(2) 正确提示:如图 13(b)所示,在试验一的基础上,在提问时嵌入了一定的知识进行导航路径搜索的约束信息。在进行拓扑路径分析提问之前,提示 GPT 检索过程中只考虑空间、门、楼梯 3 种要素类型。在该提示下,GPT 模型可学习到不同空间仅仅可通过门或者楼梯进行通行,并且找到正确的导航路径信息,包括 Keller 1、Door-031、Keller 3、Treppe、Passageway-Treppe1、Roof 节点。

基于以上分析可知,在进行复杂的拓扑分析时,由于场景中除了导航要素信息还包含大量的其他要素类型,GPT 无法直接区分哪些要素可直接用

于路径导航。因此,在分析过程中,有必要嵌入有效的知识以提升语言模型对问题的理解,并获取准确的分析结果。

4 结论与讨论

本文深入分析现有室内三维场景表达模型及方法在场景计算和分析过程存在的问题,首次提出面向室内空间智能的三维场景图表达框架,旨在建立一种可直接应用于场景检索、空间分析等智能化应用的结构化室内模型。本文对场景图要素层级化组织、几何表达和语义描述、关系表达进行了详细论述。在几何表达层面,根据各个层级要素的特性及其建模方式,选取了 5 种典型的三维数据表达结构对其进行明确定义;在语义表达层面,对每一类要素所需的语义属性进行了明确抽象与定义,可为场景图要素语义的定义提供明确的规范;在关系描述方面,采用了层级和空间关系对要素的三维关系进行了详细描述和定义。在此基础上,本文建立

了可统一表达要素几何、语义及关系的室内三维场景图概念模型,将为室内三维场景图的构建提供清晰的要素组织和表达框架。

为验证本文提出的室内三维场景图在室内智能化应用中的潜力,本文采用公开数据 IFC 模型为基础,构建完整的室内三维场景图模型,并且结合大语言模型通过自然语言提示的方式对场景图在场景检索、室内拓扑分析方面的能力进行了验证和探讨。试验表明,基于自然语言提示,本文提出的室内三维场景图模型可满足复杂场景要素检索需求以及空间分析需求。室内三维场景图模型的建立主要基于原始 IFC 模型中各种类型要素之间的几何关系构建,关系表达主要以空间关系位置,较少嵌入自然语义信息,如“桌子上放了个茶杯”。这类自然语义信息通过简单的空间分析难以获得,在后期工作中,笔者将探索基于图神经网络的学习方法,进一步提升室内要素关系的预测能力,有望进一步提升室内三维场景图的计算和分析能力。

参考文献

[1] 应申,朱利平,李霖,等. 基于室内空间特征的室内地图表达[J]. 导航定位学报, 2015, 3(4): 74-78, 91.
YING Shen, ZHU Liping, LI Lin, et al. Representation of indoor maps with the analysis of indoor space[J]. Journal of Navigation and Positioning, 2015, 3(4): 74-78, 91.

[2] 高伟,吴毅红,申抒含,等. 视觉主导的多传感器融合机器地图构建与定位技术[M]//卫星导航定位与北斗系统应用. 北京:测绘出版社,2019: 138-143.
GAO Wei, WU Yihong, SHEN Shuhan, et al. Vision led multi-sensor fusion machine map construction and positioning technology [M]//Satellite navigation positioning and Beidou system application. Beijing: Surveying and Mapping Press, 2019: 138-143

[3] 徐凯,胡瑞珍,杨鑫. 几何引导的主动式三维感知与交互[J]. 图学学报, 2022, 43(6): 1049-1056.
XU Kai, HU Ruizhen, YANG Xin. Geometry-guided active 3D perception and interaction[J]. Journal of Graphics, 2022, 43(6): 1049-1056.

[4] LAAKSO M, KIVINIEMI A. The IFC standard:a review of history, development, and standardization[J]. Electronic Journal of Information Technology in Construction, 2012, 17: 134-161.

[5] GRÖGER G, PLÜMER L. CityGML: interoperable semantic 3D city models[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2012, 71: 12-33.

[6] LI K J, CONTI G, KONSTANTINIDIS E, et al. OGC IndoorGML: a standard approach for indoor maps[M]//Geographical and fingerprinting data to create systems for indoor positioning and indoor/outdoor navigation. Amsterdam; Elsevier, 2019: 187-207.

[7] 朱庆,张利国,丁雨淋,等. 从实景三维建模到数字孪生建模[J]. 测绘学报, 2022, 51(6): 1040-1049. DOI: 10.11947/j. AGCS. 2022. 20210640.
ZHU Qing, ZHANG Liguu, DING Yulin, et al. From real 3D modeling to digital twin modeling[J]. Acta Geodaetica et Cartographica Sinica, 2022, 51(6): 1040-1049. DOI: 10.11947/j. AGCS. 2022. 20210640.

[8] 游天,周成虎,陈曦. 室内地图表示方法研究与实践[J]. 测绘科学技术学报, 2014, 31(6): 635-640.
YOU Tian, ZHOU Chenghu, CHEN Xi. The research and practice of indoor map representation[J]. Journal of Geomatics Science and Technology, 2014, 31(6): 635-640.

[9] 危双丰,刘振彬,赵江洪,等. SLAM 室内三维重建技术综述[J]. 测绘科学, 2018, 43(7): 15-26.
WEI Shuangfeng, LIU Zhenbin, ZHAO Jianghong, et al. A review of indoor 3D reconstruction with SLAM[J]. Science of Surveying and Mapping, 2018, 43(7): 15-26.

[10] 崔杨. 移动 LiDAR 点云室内三维结构化重建方法和关键技术研究[J]. 测绘学报, 2021, 50(7): 990. DOI: 10.11947/j. AGCS. 2021. 20200592.

CUI Yang. Research on methodology and the key technology of indoor 3D structured reconstruction from mobile LiDAR point clouds[J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(7): 990. DOI: 10.11947/j. AGCS. 2021. 20200592.

[11] 付燕平, 严庆安, 廖杰, 等. 基于彩色图像引导的 RGB-D 相机追踪与三维重建[J]. *武汉大学学报(工学版)*, 2022, 55(1): 92-100.

FU Yanping, YAN Qing'an, LIAO Jie, et al. RGB-D camera tracking and 3D reconstruction via color image guiding[J]. *Engineering Journal of Wuhan University*, 2022, 55(1): 92-100.

[12] WALD J, DHAMO H, NAVAB N, et al. Learning 3D semantic scene graphs from 3D indoor reconstructions[C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle:IEEE, 2020: 3961-3970.

[13] 王玮琦, 游雄, 杨剑, 等. 一种改进匹配点对选取策略的 ElasticFusion 室内三维重建算法[J]. *武汉大学学报(信息科学版)*, 2020, 45(9): 1469-1477.

WANG Weiqi, YOU Xiong, YANG Jian, et al. Elastic fusion for indoor 3D reconstruction with an improved matching points selection strategy[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(9): 1469-1477.

[14] HOU Ji, DAI A, NIEBNER M. 3D-SIS: 3D semantic instance segmentation of RGB-D scans[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach:IEEE, 2019: 4421-4430.

[15] SU Hang, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3D shape recognition[C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago:IEEE, 2015: 945-953.

[16] SONG Shuran, LICHTENBERG S P, XIAO Jianxiang. SUN RGB-D: a RGB-D scene understanding benchmark suite[C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston:IEEE, 2015: 567-576.

[17] NAJIBI M, LAI Guangda, KUNDU A, et al. DOPS: learning to detect 3D objects and predict their 3D shapes[C]//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle:IEEE, 2020:11913-11922.

[18] 熊汉江, 郑先伟, 丁友丽, 等. 基于 2D-3D 语义传递的室内三维点云模型语义分割[J]. *武汉大学学报(信息科学版)*, 2018, 43(12): 2303-2309.

XIONG Hanjiang, ZHENG Xianwei, DING Youli, et al. Semantic segmentation of indoor 3D point cloud model based on 2D-3D semantic transfer[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 2303-2309.

[19] YI Li, ZHAO Wang, WANG He, et al. GSPN: generative shape proposal network for 3D instance segmentation in point cloud[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach:IEEE, 2019: 3947-3956.

[20] LIU Yunze, YI Li, ZHANG Shanghang, et al. P4Contrast: contrastive learning with pairs of point-pixel pairs for RGB-D scene understanding[EB/OL]. [2023-12-01]. <http://arxiv.org/pdf/1907.11692>.

[21] CHARLES R Q, HAO Su, MO Kaichun, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu:IEEE, 2017: 652-660.

[22] 邓林涛, 方志军. 基于特征负反馈卷积的点云分析方法[J]. *激光与光电子学进展*, 2022, 59(12): 1210006.

DENG Lintao, FANG Zhijun. Point cloud analysis method based on feature negative feedback convolution[J]. *Laser & Optoelectronics Progress*, 2022, 59(12): 1210006.

[23] CHEN X, WU H, LICHTI D, et al. Extraction of indoor objects based on the exponential function density clustering model[J]. *Information Sciences*, 2022, 607: 1111-1135.

[24] NING Xiaojuan, WANG Yinghui, HAO Wen, et al. Structure-based object classification and recognition for 3D scenes in point clouds [C]//*Proceedings of 2014 International Conference on Virtual Reality and Visualization*. Shenyang: IEEE, 2014: 166-173.

[25] KIRILLOV A, HE Kaiming, GIRSHICK R, et al. Panoptic segmentation[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 9404-9413.

[26] REN Dayong, WU Zhengyi, LI Jiawei, et al. Point attention network for point cloud semantic segmentation[J]. *Science China Information Sciences*, 2022, 65(9): 192104.

[27] CHEN Kai, OUYANG W, LOY C C, et al. Hybrid task cascade for instance segmentation[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach:IEEE, 2019: 4974-4983.

[28] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]//*Proceedings of the 18th International Conference of Medical Image Computing and Computer*. Cham: Springer, 2015: 234-241.

[29] FANG Hao, LAFARGE F, PAN Cihui, et al. Floorplan generation from 3D point clouds: a space partitioning approach[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 175: 44-55.

[30] LIU Chen, WU Jiajun, KOHLI P, et al. Raster-to-vector: revisiting floorplan transformation[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017: 2195-2203.

[31] 张蔚, 王腾, 卢政达, 等. 基于 CNN 的住宅平面图元素识别与布局语义分析[J]. *中国体视学与图像分析*, 2020, 25(2): 174-182.

ZHANG Wei, WANG Teng, LU Zhengda, et al. Floor plan recognition and semantic layout analysis based on a convolutional neural network[J]. *Chinese Journal of Stereology and Image Analysis*, 2020, 25(2): 174-182.

[32] FANG Hao, PAN Cihui, HUANG Hui. Structure-aware indoor scene reconstruction via two levels of abstraction[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 178: 155-170.

[33] HAN Jiali, RONG Mengqi, JIANG Hanqing, et al. Vectorized indoor surface reconstruction from 3D point cloud with multistep 2D op-

timization[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 177: 57-74.

[34] TRAN H, KHOSHELHAM K. Procedural reconstruction of 3D indoor models from lidar data using reversible jump Markov chain Monte Carlo[J]. Remote Sensing, 2020, 12(5): 838.

[35] HAN Jiali, LIU Yuzhou, RONG Mengqi, et al. FloorUSG: indoor floorplan reconstruction by unifying 2D semantics and 3D geometry [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 196: 490-501.

[36] JAYARAJ P, RAMIYA A M. D citygml building modelling from LiDAR point cloud data[J]. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2018, 425: 175-180.

[37] 王敬淳, 汤圣君, 王伟玺, 等. 室内三维点云空间自动划分与规则化方法[J]. 测绘通报, 2022(8): 93-97, 138.
WANG Jingchun, TANG Shengjun, WANG Weixi, et al. Automatic division and regularization of indoor 3D point cloud space[J]. Bulletin of Surveying and Mapping, 2022(8): 93-97, 138.

[38] TANG Shengjun, LI Xiaoming, ZHENG Xianwei, et al. BIM generation from 3D point clouds by combining 3D deep learning and improved morphological approach[J]. Automation in Construction, 2022, 141: 104422.

[39] 丁雨淋, 何小波, 朱庆, 等. 实时威胁态势感知的室内火灾疏散路径动态优化方法[J]. 测绘学报, 2016, 45(12): 1464-1475. DOI: 10.11947/j. AGCS. 2016. 20160053.
DING Yulin, HE Xiaobo, ZHU Qing, et al. A dynamic optimization method of indoor fire evacuation route based on realtime situation awareness[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(12): 1464-1475. DOI: 10.11947/j. AGCS. 2016. 20160053.

[40] 邓中亮, 王小恒. 一种用于大型建筑火灾中的应急疏散算法[J]. 软件, 2011, 32(2): 112-114.
DENG Zhongliang, WANG Xiaoheng. An emergency evacuation algorithm in large building fire[J]. Computer Engineering & Software, 2011, 32(2): 112-114.

[41] 朱庆, 胡明远, 许伟平, 等. 面向火灾动态疏散的三维建筑信息模型[J]. 武汉大学学报(信息科学版), 2014, 39(7): 762-766, 872.
ZHU Qing, HU Mingyuan, XU Weiping, et al. 3D building information model for facilitating dynamic analysis of indoor fire emergency [J]. Geomatics and Information Science of Wuhan University, 2014, 39(7): 762-766, 872.

[42] 韩李涛, 郭欢, 张海思. 一种多出口室内应急疏散路径规划算法[J]. 测绘科学, 2018, 43(12): 105-110.
HAN Litao, GUO Huan, ZHANG Haisi. An algorithm for route planning applied in multi-exit indoor emergency evacuation[J]. Science of Surveying and Mapping, 2018, 43(12): 105-110.

[43] 李清泉, 周宝定, 马威, 等. GIS辅助的室内定位技术研究进展[J]. 测绘学报, 2019, 48(12): 1498-1506. DOI: 10.11947/j. AGCS. 2019. 20190455.
LI Qingquan, ZHOU Baoding, MA Wei, et al. Research process of GIS-aided indoor localization[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(12): 1498-1506. DOI: 10.11947/j. AGCS. 2019. 20190455.

[44] 刘涛, 张星, 李清泉, 等. 顾及地标可视性的室内导航路径优化算法[J]. 武汉大学学报(信息科学版), 2017, 42(1): 43-48.
LIU Tao, ZHANG Xing, LI Qingquan, et al. An indoor pedestrian route planning algorithm based on landmark visibility[J]. Geomatics and Information Science of Wuhan University, 2017, 42(1): 43-48.

[45] 陈锐志, 陈亮. 基于智能手机的室内定位技术的发展现状和挑战[J]. 测绘学报, 2017, 46(10): 1316-1326. DOI: 10.11947/j. AGCS. 2017. 20170383.
CHEN Ruizhi, CHEN Liang. Indoor positioning with smartphones: the state-of-the-art and the challenges[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(10): 1316-1326. DOI: 10.11947/j. AGCS. 2017. 20170383.

[46] WALD J, NAVAB N, TOMBARI F. Learning 3D semantic scene graphs with instance embeddings[J]. International Journal of Computer Vision, 2022, 130(3): 630-651.

[47] ZHU G, ZHANG L, JIANG Y, et al. Scene graph generation: a comprehensive survey[EB/OL]. [2024-01-25]. <http://arxiv.org/abs/2201.00443>.

[48] ZLATANOVA S. On 3D topological relationships[C]//Proceedings of 2000 International Workshop on Database & Expert Systems Applications. London:IEEE, 2000.

[49] IFC Wiki. KIT IFC Examples[EB/OL]. [2024-02-15]. https://www.ifcwiki.org/index.php?title=KIT_IFC_Examples.

(责任编辑:张艳玲)

3D scene graph representation and application for intelligent indoor spaces

TANG Shengjun^{1,2}, DU Siqu², WANG Weixi^{1,2}, GUO Renzhong^{1,2}

- 1. *State Key Laboratory of Subtropical Building and Urban Science, Shenzhen University, Shenzhen 518061, China;*
- 2. *School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518061, China*

Abstract: Existing methods for indoor 3D scene representation focus on object-oriented descriptions, with element representations limited to object-level semantic understanding. These methods lack the ability to express complex relational information within indoor scenes. Addressing the demands of intelligent indoor space tasks, there is a critical need for a structured model that can comprehensively and accurately describe the geometry, semantics, and relationships of indoor elements, while also supporting semantic retrieval and analytical reasoning. Based on the fundamental theory of 3D scene graphs, this paper innovatively proposes a 3D scene graph representation model tailored for intelligent indoor spaces. It systematically introduces the hierarchical organization, geometric representation, semantic description, and relational description methods of indoor 3D scene graphs. A conceptual model is established that uniformly describes the geometry, semantics, and relationships of indoor elements. Additionally, this graph model is compatible with existing 3D scene representation methods, ensuring good data compatibility. Finally, a comprehensive multi-level relational 3D scene graph model is constructed based on the publicly available IFC model. This model’s application capabilities, potential, and limitations are systematically explored and analyzed through applications such as complex scene retrieval and topological analysis, in conjunction with large language models. The results demonstrate that the indoor 3D scene graph model possesses complex computation and analysis capabilities, can be directly integrated with large language models, and enables complex scene analysis applications through simple natural language prompts.

Key words: indoor modeling; graph model; IFC; CityGML; large language model

Foundation support: The Natural Science Foundation of Guangdong Province (No. 2024A1515030061); Research Project of Shenzhen Science and Technology Innovation Committee (Nos. KJZD20230923115508017; JCYJ20210324093012033); The Research Project of State Key Laboratory of Subtropical Building and Urban Science (No. 2023ZB18)

First author: TANG Shengjun (1991—), male, PhD, associate researcher, majors in urban 3D element structured reconstruction and multi-sensor fusion mapping.

E-mail: shengjuntang@szu.edu.cn

Corresponding author: WANG Weixi

E-mail: wangwx@szu.edu.cn