

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2025)07-2408-12

论文引用格式: Wang J, Xiong H R and Huang H. 2025. Application of the guided set diffusion model algorithm to floor plan reconstruction. Journal of Image and Graphics, 30(7):2408-2419(王静, 熊皓然, 黄惠. 2025. 引导集扩散模型在户型平面图重建中的应用. 中国图象图形学报, 30(7):2408-2419)[DOI:10.11834/jig.240508]

引导集扩散模型在户型平面图重建中的应用

王静, 熊皓然, 黄惠*

深圳大学可视计算研究中心, 深圳 518061

摘要: 目的 户型平面图的矢量化是一项关键技术, 用于从户型平面点阵图中提取精确的结构信息, 广泛应用于建筑装修、家居设计以及场景理解等领域。现有方法通常采用两阶段流程: 第1阶段利用深度神经网络提取户型区域的掩膜, 第2阶段通过后处理步骤从掩膜轮廓中提取墙体的矢量信息。然而, 这种方法存在误差累积问题, 后处理算法难以保证鲁棒性。为了解决上述问题, 提出一种基于引导集扩散模型的户型平面图矢量重建算法。方法 该算法通过将目标检测或实例分割方法中获得的粗糙轮廓输入扩散模型, 逐步迭代轮廓点进行重建。此外, 还引入了一种轮廓倾斜度损失函数, 以帮助网络生成更规整的房间布局, 从而进一步提升矢量化结果的准确性。结果 在公开的 CUBICase5K 数据集上, 对提出的算法进行广泛测试。实验结果表明, 在不同的输入条件下, 该算法均能有效优化房间轮廓的精度, 显著提高墙线矢量化提取精度。结论 所提出的基于引导集扩散模型的矢量重建算法, 通过解决传统方法中的误差累积问题, 实现了室内户型平面图中墙体矢量化精度的提升。这一改进为建筑与家居设计等领域的应用提供了更为可靠的技术支持。

关键词: 深度学习; 室内户型图; 生成式重建; 扩散模型; 户型图矢量技术

Application of the guided set diffusion model algorithm to floor plan reconstruction

Wang Jing, Xiong Haoran, Huang Hui*

Visual Computing Research Center, Shenzhen University, Shenzhen 518061, China

Abstract: Objective Indoor floor plan vectorization is a sophisticated technique that aims to extract precise structural information from raster images and convert them into vector representations. This process is essential in fields, such as architectural renovation, interior design, and scene understanding, where the accurate and efficient vectorization of floor plans can considerably enhance the quality and usability of spatial data. The vectorization process has been traditionally performed using a two-stage pipeline. In the first stage, deep neural networks are used to segment the raster image, producing masks that define room regions within the floor plan. These masks serve as the foundation for the subsequent vectorization process. In the second stage, post-processing algorithms are applied to these masks to extract vector information, with focus

收稿日期: 2024-09-02; 修回日期: 2024-11-19; 预印本日期: 2024-11-26

* 通信作者: 黄惠 huihuang@szu.edu.cn

基金项目: 国家自然科学基金项目(U21B2023); 广东省基础与应用基础研究重点项目(2023B1515120026); 广东高等学校创新团队项目(2022KCXTD025); 深圳市科技创新项目(KQTD20210811090044003, RCJC20200714114435012); 深圳大学研究生教育改革项目(SZUGS2022JG01); 卓越研究计划项目(2035)

Supported by: National Natural Science Foundation of China (U21B2023); Guangdong Basic and Applied Basic Research Foundation (2023B1515120026); Department of Education of Guangdong Province Innovation Team (2022KCXTD025)

on elements, such as walls, doors, and other structural components. However, this process also poses challenges. One of the major issues is error accumulation. Inaccuracies in initial mask generation can lead to compounded errors during vectorization. Moreover, post-processing algorithms frequently lack robustness, particularly when dealing with complex or degraded input images, leading to suboptimal vectorization results. **Method** To address these challenges, we propose a novel approach based on diffusion models for the vector reconstruction of indoor floor plans. Diffusion models, which were originally developed for generative tasks, have exhibited considerable promise in producing high-quality output by iteratively refining input data. Our method leverages this capability to enhance the precision of floor plan vectorization. In particular, the algorithm starts with rough masks generated using object detection or instance segmentation models. Although these masks provide a basic outline of room regions, they may lack the accuracy required for precise vectorization. The diffusion model is then employed to iteratively refine the contour points of these rough masks, gradually reconstructing room contours with higher accuracy. This process involves multiple iterations, during which the model adjusts the contour points based on patterns learned from the training data, leading to a more accurate representation of room boundaries. A key innovation of our approach is the introduction of a contour inclination loss function. This loss function is specifically designed to guide the diffusion model in generating more reasonable and structurally sound room layouts. By penalizing unrealistic or impractical contour inclinations, the model is encouraged to produce output that closely resembles real-world room configurations. This process not only improves the visual accuracy of room contours but also enhances the overall quality of the vectorized floor plan. The benefits of our diffusion model-based approach are manifold. First, we can significantly reduce the error accumulation that plagues traditional methods by refining room contours before the vectorization stage. The more accurate the room contours, the more precise the subsequent wall vector extraction process. Second, the use of a diffusion model allows for a more robust handling of complex or noisy input images. In contrast with traditional post-processing algorithms, which may experience difficulty with irregularities in the input data, the diffusion model's iterative refinement process is better equipped to address such challenges, leading to more reliable vectorization results. **Result** We have thoroughly validated our method on the public CubiCase5K dataset, which is a widely used benchmark in the field of floor plan vectorization. The results of our experiments demonstrate that our approach significantly outperforms existing methods in terms of accuracy and robustness. Notably, we observe a marked improvement in the precision of wall vector extraction, which is crucial for applications in architectural renovation and interior design. The diffusion model's ability to produce more accurate room contours directly translates into better vectorized representations of floor plans, making our approach an invaluable tool for professionals in these fields. **Conclusion** In conclusion, the vectorization of indoor floor plans is a critical task with far-reaching applications, and the limitations of traditional methods have highlighted the need for more advanced techniques. Our diffusion model-based algorithm represents a significant step forward, offering a more accurate and reliable solution for the vector reconstruction of indoor floor plans. By refining room contours through iterative adjustments and incorporating a contour inclination loss function, our method not only addresses the issue of error accumulation but also enhances the overall quality of the vectorized output. Furthermore, our method exhibits broad application prospects in the field of interior design. Modern interior designers increasingly rely on digital tools for design and planning, wherein accurate indoor space representation is essential. With our method, designers can quickly generate high-quality room outlines and apply them to various design scenarios, such as furniture placement, lighting design, and space optimization. This efficient and precise vectorization technology not only saves designers significant time but also improves the feasibility and practicality of design solutions. In the field of building renovation, our approach also offers significant advantages. During the renovation of old buildings, redrawing and optimizing the original floor plan are frequently necessary. This process is often constrained by the quality of the original drawings, particularly in older buildings, where the original plans may have already deteriorated or been lost. Through our diffusion model vectorization method, architects can quickly reconstruct digital versions of these old floor plans and use them as a foundation for renovation designs. This step not only improves the efficiency of the renovation process but also helps preserve the historical character of the building. Lastly, our method demonstrates important application potential in the field of architectural scene understanding. With the rise of smart homes and automated building management systems, accurate interior floor plan data are critical for enabling these intelligent features. Our vectorization method can generate highly accurate representations of indoor spaces, providing reli-

able foundational data for intelligent systems. Overall, our approach promises to significantly improve the efficiency and effectiveness of floor plan vectorization, paving the way for more accurate architectural designs and enhanced spatial understanding.

Key words: deep learning; indoor floor plan; generated reconstruction; diffusion model; floorplan vector technology

0 引言

户型平面图是建筑设计的核心元素之一,蕴含丰富的结构和语义信息,对建筑的设计、修改、分析和优化至关重要。然而,常见的户型平面图通常以点阵图格式保存,虽然便于直观查看,但丢失了设计时的原始矢量信息。在家装的二次设计、摆放、三维展现等实际应用时,矢量图都是不可或缺的(于柳和吴晓群,2024;杜韬等,2024)。因此,需要对点阵图进行矢量化处理。户型矢量化提取有助于将纸质图纸或位图格式的户型图转化为更灵活的数字矢量格式,便于编辑、共享和再利用。若手动处理,需要设计师重新描绘墙线、标出门窗、判别房间类型等,这一过程既耗时又耗力。因此,若能实现点阵图格式户型平面图的快速、准确、自动化矢量提取,将极大地降低设计师手动标注的工作量,提高工作效率。

随着人工智能技术的迅猛发展,基于计算机视觉技术的户型平面图矢量化研究取得显著成果(Pizarro等,2022;Zhu等,2020;岳亮等,2024)。这些方法通常是先将平面户型图进行分割,得到每一个房间结构的掩膜,然后通过多个后处理步骤对掩膜进行墙体的矢量化提取。目前,主流研究主要专注于提升房间的分割准确率。尽管现有的户型矢量化工作已取得显著进展,但在保证户型整体的结构完整性方面仍存在不足。容易出现房间、门窗、走廊等细节结构在检测中丢失的情况,导致无法获取完整的矢量信息。此外,由于房间的拓扑关系并非固定,矢量化提取过程中,也难以处理复杂且多变的户型结构。因此,要实现一种精确且具有通用性的平面户型图矢量化技术,首先需要得到更加准确的户型结构轮廓,以降低墙体矢量化提取的难度,增加算法的鲁棒性。基于以上考虑,本文提出一种基于引导集扩散模型的室内户型平面图矢量重建算法,将目标检测或实例分割方法中获得的粗糙轮廓,利用引导集扩散模型逐步优化轮廓点,以重建房间结构,让矢量化过程更加鲁棒,从而适应不同风格的户型

图。这一技术的实现将为虚拟现实和增强现实等先进领域提供有力的数据支持,实现户型平面图矢量化重建的真正落地应用。

1 相关工作

目前,户型矢量化的工作主要采用基于掩膜的方法或基于轮廓的方法。

1.1 基于掩膜的户型矢量化

在计算机视觉领域中,掩膜是一个与原图像相同的矩阵(叶乙轩等,2024),以像素为最小单元进行分类,每个像素点对应一个类别标签。在户型平面图矢量化领域,最早的技术路线是基于掩膜的处理方法——对于给定的平面户型图,首先采用目标检测技术识别各房间的位置与类别,然后通过语义分割或实例分割技术提取掩膜(万骏辉等,2024),再结合相关后处理技术进行墙体矢量的提取。Liu等人(2017)在平面图矢量化过程中首次引入了基于语义分割的方法,利用所得结果分析预测房间的拐点位置,实现高效提取户型平面矢量图。Lyu等人(2021)首先利用YOLOv4(you only look once version 4)目标检测模型获取各房间的位置与区域,然后采用Deeplabv3+(Chen等,2018)语义分割模型获取各房间的语义信息,基于墙体、门、窗大多为垂直或水平的假设,通过后处理得到各房间的轮廓信息并识别门、窗等元素。Chen等人(2022)采用Mask R-CNN(mask region-based convolutional neural network)(He等,2017)架构,根据目标检测所得的各房间边框与分类进行语义分割,将矢量化过程转为优化问题,采用室内房间层级的坐标递减顺序、最小化损失函数得到矢量化的户型结构。这类方法能够较好地获取各房间的语义信息与区域范围,但可能存在边缘不清晰的问题,在从掩膜结果转为精准的矢量轮廓时,由于需要进行后处理,而后处理算法的设计需要引入更多的条件与假设,导致设计难度大,难以在多种风格的数据集上具备鲁棒性与通用性。

1.2 基于轮廓的户型矢量化

RoomFormer(Yue等, 2023)将3D点云的俯视密度图作为户型平面图输入, 利用Transformer(Vaswani等, 2017)网络结构迭代地输出各个房间的轮廓点, 将每个轮廓点顺序连接用以表示每个房间的多边形轮廓。该方法借鉴了实例分割BoundaryFormer(Lazarow等, 2022)中提出的可微光栅化模块计算损失, 从而指导网络训练, 输出每一个实例的多边形轮廓。PolyDiffuse(Chen等, 2023)沿用了RoomFormer中的设定, 将户型平面图重建定义为多边形轮廓重建, 即点集构成轮廓的集合, 并通过其他方法得到的初始轮廓, 利用改进后的去噪扩散概率模型(denoising diffusion probabilistic models, DDPM)(Ho等, 2020)进行加噪和去噪的学习。这类方法由于直接处理多边形轮廓, 因此无需复杂的后处理即可直接得到各个房间的轮廓表示, 能够更好地保证轮廓的封闭性、完整性等, 且提取轮廓准确性也更高。然而RoomFormer如果存在预测点丢失或错误时, 可能会导致房间轮廓出现较大的偏差; 而PolyDiffuse则较为依赖于房间的初始状态, 识别户型的位置、点的数量、轮廓精度均会影响最终重建的结果。

2 基于引导集扩散网络的户型矢量化

由于现有的算法依赖于目标检测或分割的结果, 在提取矢量多边形时, 存在误差累积, 导致后处理方法难以保持鲁棒性。因此, 本文算法以目标检测或实例分割输出的粗糙轮廓作为输入, 利用引导集扩散网络, 将其重建为多个准确的户型多边形, 再利用墙线提取模块进行墙体的矢量提取。本文算法的流程图如图1所示。

2.1 引导集扩散网络

引导集扩散网络相较于标准去噪扩散概率模型, 主要引入了一个集合内的每个元素服从不同高斯噪声分布的假设。如式(1)所示, 一个 $\mathbf{x}$ 元素中的每个点用二维坐标表示, 顺序连接即可得到多边形轮廓。具体为

$$\mathbf{x} = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^i, \dots, \mathbf{x}^N\} \quad (1)$$

因此, 在前向过程中, 引导集扩散网络需要先输入每个元素的初始轮廓, 同时训练一个引导网络; 在反向过程中, 初始房间轮廓在引导网络的指导下, 获得各个元素的不同高斯噪声分布。具体过程参

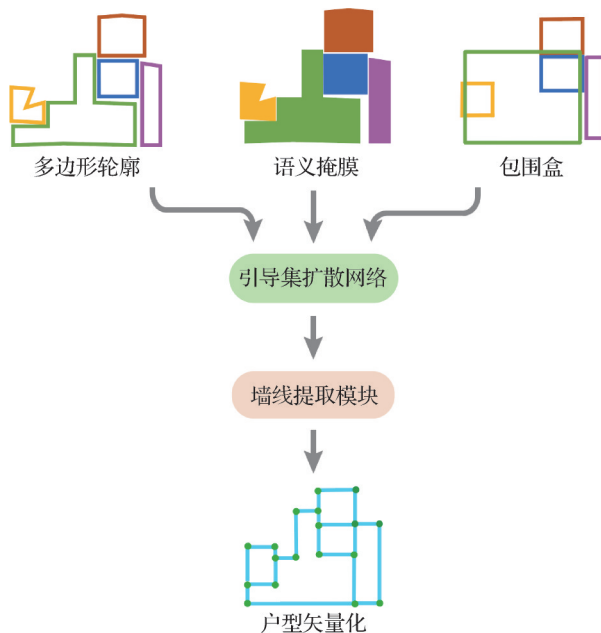


图1 本文算法流程图

Fig. 1 The overview of our method

考图2。

2.2 网络搭建

根据引导集扩散网络公式搭建网络模型, 主要由引导网络模块和去噪网络模块组成。

2.2.1 引导集网络模块

根据公式(2), 该公式描述了在第 t 步中, 第 i 个轮廓点的状态 x_t^i 的分布 q , 它是一个高斯分布 N 。扩散的前向过程仅依赖于均值 $\mu_\phi(x_0, t, i)$ 与方差 $\sigma_\phi^2(x_0, t, i)$, 且每个轮廓元素 x_t^i 均如式(1), I 表示单位矩阵, 网络需满足由这两个参数构成的特定分布 q 。

$$q(x_t^i | x_{t-1}^i, x_0) = N(x_t^i; \sqrt{1 - \beta_t} x_{t-1}^i + \mu_\phi(x_0, t, i), \beta_t \sigma_\phi^2(x_0, t, i) I) \quad (2)$$

因此, 引导网络主要学习这两个参数, 直接将这两个参数映射为 DETR (detection Transformer) 风格的 Transformer (Trans) 解码器模块网络, 同时解耦时间 t 与各个轮廓的高斯噪声, 噪声由参数 β_t 控制以适用于现有的扩散模型架构。解码器网络的初始输入为每个轮廓 x_0 的点坐标, 其特征编码由 X 、 Y 、轮廓坐标3个位置编码拼接构成。具体为

$$\begin{cases} \bar{\mu}_\phi(x_0, 0, i) = 0 \\ \bar{\mu}_\phi(x_0, t, i) = \sqrt{\alpha_t} \bar{\mu}_\phi(x_0, t-1, i) + \mu_\phi(x_0, t, i) \\ \bar{\sigma}_\phi^2(x_0, 0, i) = 0 \\ \bar{\sigma}_\phi^2(x_0, t, i) = \alpha_t \bar{\sigma}_\phi^2(x_0, t-1, i) + (1 - \alpha_t) \sigma_\phi^2(x_0, t, i) \end{cases} \quad (3)$$

$$(4)$$

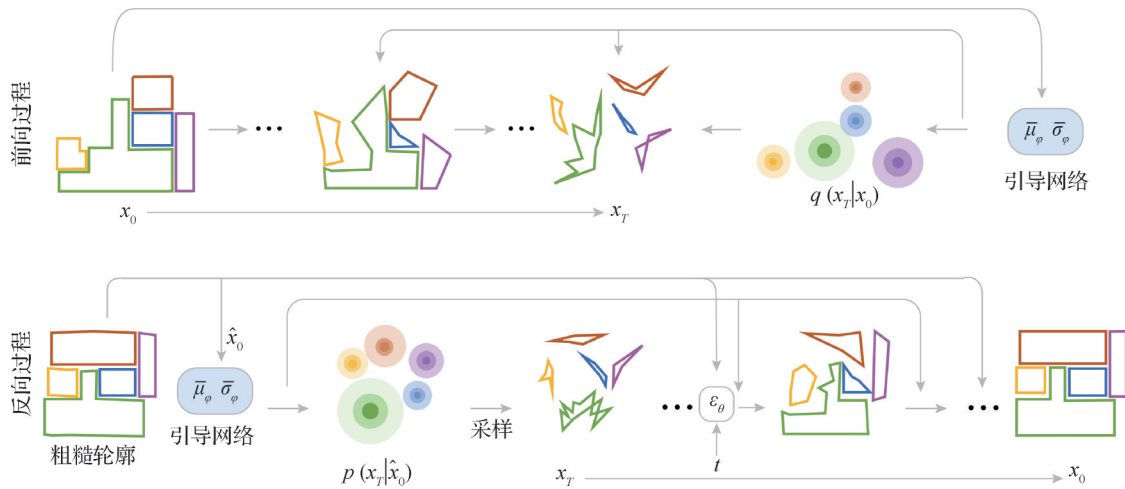


图2 平面户型图矢量化重建流程

Fig. 2 The pipeline of floorplan reconstruction

$$\begin{cases} \bar{\mu}_\phi(\hat{x}_0, T, i) = \text{Trans}_{\bar{\mu}_\phi}(\hat{x}_0, i) \\ \bar{\sigma}_\phi(\hat{x}_0, T, i) = \text{Trans}_{\bar{\sigma}_\phi}(\hat{x}_0, i) \end{cases} \quad (5)$$

$$\bar{\mu}_\phi(\hat{x}_0, t, i) = (1 - \sqrt{\alpha_t}) \text{Trans}_{\bar{\mu}_\phi}(\hat{x}_0, i) \quad (6)$$

$$\bar{\sigma}_\phi(\hat{x}_0, t, i) = (1 - \sqrt{\alpha_t}) \text{Trans}_{\bar{\sigma}_\phi}(\hat{x}_0, i) \quad (7)$$

2.2.2 去噪网络模块

在引导网络输出的两个参数之下,去噪网络还需要学习到噪声 $\varepsilon_\theta(x_i, t, y)$, 该网络借鉴 RoomFormer 中的编码解码器架构, 仅加入了元素内注意力层, 并将元素内注意力添加到每个解码器的全局注意力中。该部分节点输入和引导网络相同。具体为

$$\begin{aligned} x_{t-1}^i &= \frac{1}{\sqrt{\alpha_t}} [x_t^i - \bar{\mu}_\phi(\hat{x}_0, t, i) - \\ &\bar{\sigma}_\phi^2(\hat{x}_0, t, i) \frac{1 - \alpha_t}{1 - \bar{\alpha}_t} \varepsilon_\theta(x_t, t, y)] + \\ &\bar{\mu}_\phi(\hat{x}_0, t-1, i) + \sigma_t z^i \end{aligned} \quad (8)$$

2.3 损失计算

本文的损失计算分为两部分, 第1部分是引导集网络在训练过程的损失函数, 其目标是学习当前多边形轮廓在特定排列情况下各个轮廓的分布, 并与其他排列进行区分, 所以采用一种可以评估不同排列的损失函数方法。具体为

$$\hat{L}_{\text{perm}}(\phi) = \sum_{i=1}^N \left(\max_{l \leq j \leq N, j \neq i} L_{\text{Triplet}}(x_t^i, x_0^i, x_0^j) + \max_{1 \leq j \leq N, j \neq i} L_{\text{Triplet}}(x_0^i, x_t^i, x_t^j) \right) \quad (9)$$

式中, N 为多边形轮廓的数量。三元组损失函数定义为

$$L_{\text{Triplet}}(x_t^i, x_0^i, x_0^j) = \max(0, \alpha + D(x_0^i, x_t^i) - D(x_0^i, x_0^j)) \quad (10)$$

$$L_{\text{Triplet}}(x_0^i, x_t^i, x_t^j) = \max(0, \alpha + D(x_0^i, x_t^i) - D(x_0^i, x_t^j)) \quad (11)$$

式中, D 表示两个多边形轮廓的距离函数。 $L_{\text{Triplet}}(x_t^i, x_0^i, x_0^j)$ 衡量了在时间 t 的元素 x_t^i 、初始元素 x_0^i 和另一个初始元素 x_0^j 之间的三元组损失; $L_{\text{Triplet}}(x_0^i, x_t^i, x_t^j)$ 衡量了初始元素 x_0^i 、在时间 t 的元素 x_t^i 和另一个初始元素 x_t^j 之间的三元组损失, 本文设置 $\alpha = 0.1$ 。该方法从房间的排列 $N!$ 种变为顺序轮廓的排列 $2M$ 种, 距离函数的运算则是类似点对应算法, 考虑哪一种顺序轮廓排列方式情况下各个点之间的距离最短。

第2部分的损失计算是选取轮廓倾斜度损失函数 CI(contour inclination) 进行训练, 验证该重建方法能够使较为粗糙的倾斜的轮廓 c_i 变得更加规整。

轮廓倾斜度指标参考式(12), 是同时考虑长度 l 和角度 a 的一个指标, 当其为0时, 表示轮廓中的边完全水平或竖直, 指标越大轮廓越倾斜。具体为

$$C = \sum_{i=1}^N l(c_i) \times \ln(c_i) \quad (12)$$

计算流程如下: 首先对轮廓进行归一化处理, 然后将轮廓中每条边的长度乘以这条边角度的倾斜度, 最后求和得到该指标。由于轮廓倾斜度指标进行归一化并引入长度信息, 因此该指标具有尺度无关性, 仅与轮廓的斜边角度和斜边长度占比有关。

每条边角度的倾斜度计算式为

$$In(a) = \frac{1}{2} \sin(4 \times a - \frac{\pi}{2}) + \frac{1}{2} \quad (13)$$

角度倾斜度指标函数图像如图3所示。在水平或竖直时,对应的角度倾斜度最低为0,在45°或135°时,角度倾斜度最高为1。

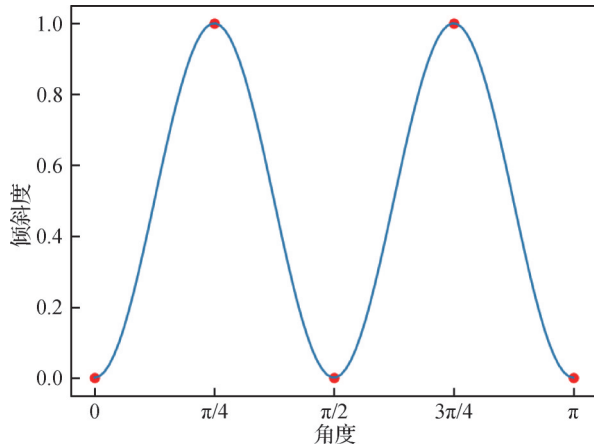


图3 角度倾斜度指标函数图像

Fig. 3 The picture of the angle inclination metric function

2.4 墙线提取算法

墙线提取算法通过合并所有非门窗的房间轮廓,以统一的可平面图(planar graph)来表示墙线,如图4所示。该算法分为3个步骤,即预处理、轮廓合并和后处理。算法步骤如下:

输入:初始轮廓集合;

输出:平面户型图的矢量化表达;

1)包围盒预处理。

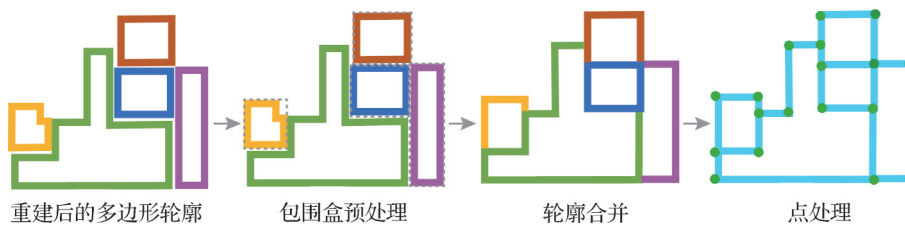


图4 墙线提取算法

Fig. 4 Wall extraction algorithm

3 实验结果

首先介绍公开数据集的统计与处理方法,随后开展两阶段的实验:第1阶段输入平面户型图进行目标检测或实例分割,输出包围盒或粗糙掩膜的轮廓;第2阶段利用引导集扩散模型对第1阶段输出的结果进行重建,并将重建后的轮廓输入到后处理

(1)提取每个轮廓的最小包围盒;

(2)如果轮廓面积占最小包围盒面积的95%以上,将最小包围盒的4个点坐标替换掉原有的轮廓点;

(3)如果任意两边平行且距离小于30像素,则将两边放入并查集中,最后将所有需要合并的边的平行方向轴坐标更改为中线对应轴坐标,对应于提取墙中线;

(4)如果其他轮廓的点与矩形轮廓的点距离小于25像素,则直接移动到矩形轮廓的点位置,或其他轮廓的点距离矩形轮廓的边小于30像素,则投影到矩形轮廓的边位置。

2)轮廓合并。

(1)找到所有平行且距离小于20像素的两边,在并查集中进行合并操作;

(2)根据并查集得到所有的边,根据边更新所有的点坐标;

(3)将相邻的点进行合并表示,得到点的列表,同时更新所有边的表示,去重后得到边的列表,构成可平面图的墙线表示。

3)后处理。

(1)根据点之间的距离进行聚类,用每个簇的中心点替换原本的簇;

(2)计算相邻边的角度,将最偏离水平或垂直的两边的连接点删去,重复该过程直到角度均小于阈值,得到精确的平面户型图的矢量化表达。

算法中以提取墙线,最终实现墙体的矢量结果可视化。

3.1 数据处理

首先介绍使用的CubiCasa5K数据集,然后介绍对该数据集的预处理方法。

3.1.1 CubiCasa5K数据集

CubiCasa5K数据集(Kalervo等,2019)是大规模的较有挑战的数据集,包含上千幅图像和超过80个

物体类别。该数据集跨度广,如图5所示,包含多种手绘风格和CAD(computer-aided design)渲染图,如稠密结构图、高质量结构图和彩色图,适合用于开发与研究更具鲁棒性的网络。该数据集原本标注采用可缩放矢量图形文件SVG(scalable vector graphics),本次数据处理前首先采用数据来源提供的方法处理为掩膜图像,再进行提取轮廓处理,如图6所示。

1)数据处理方面。CubiCasa5K数据集标注分

为以下11类:户外、厨房、客厅、卧室、浴室、玄关、储物间、车库、其他房间、窗户、门。该数据集包含4 200幅训练图像和400幅测试图像,经过数据清洗后,包含4 116幅训练图像和396幅测试图像。

2)数据统计方面。CubiCasa5K训练集图像平均高度840像素、平均宽度940像素,标注方面最大房间数40,最大轮廓点数40。CubiCasa5K测试集图像平均高度708像素、平均宽度811像素,标注方面房间数最多32间,轮廓点数最多36个。

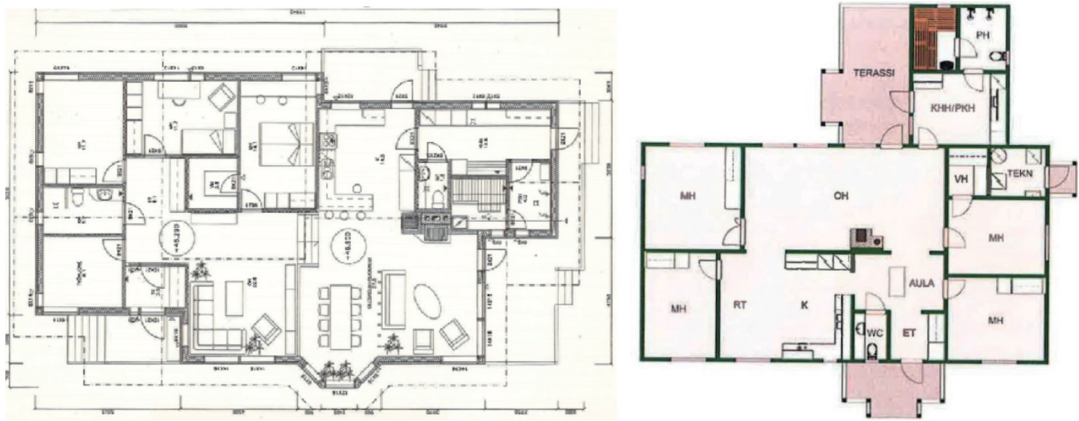


图5 CubiCasa5K数据集的示例图像
Fig. 5 Sample images from the CubiCasa5K dataset

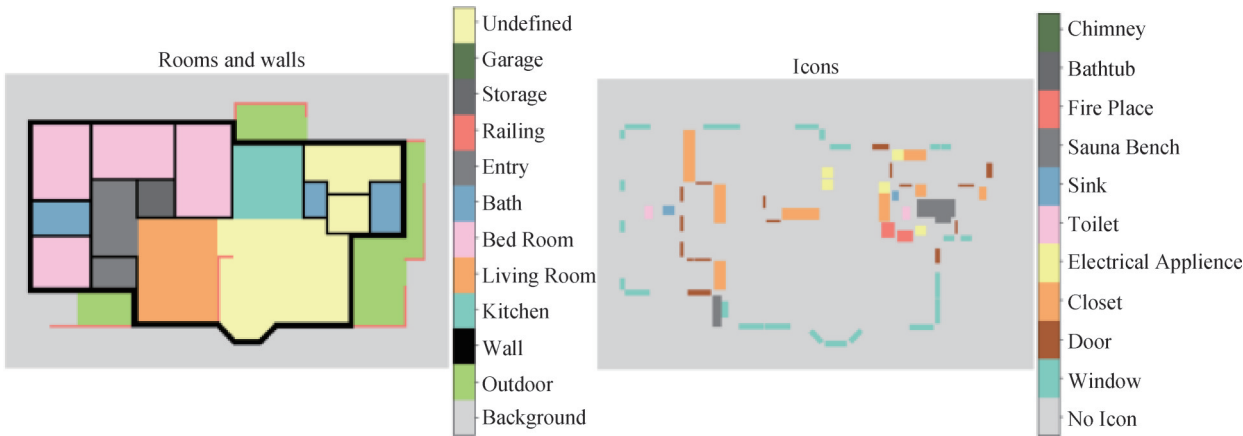


图6 CubiCasa5K数据集对应的掩膜图像
Fig. 6 Corresponding mask images from the CubiCasa5K dataset

3.1.2 数据处理方法

数据处理分为类别处理、格式处理、数据清洗3个步骤。

1)类别处理。主要选取与房间结构相关且本文技术路线关心的类别,如客餐厅、门窗等,而浴缸等房间内元素以及墙体等类别均从标注中删去。

2)格式处理。主要将语义掩膜信息转为矢量轮廓信息,即根据语义掩膜图像和图像处理技术,将每个相同颜色的独立联通区域转为点线表示的最外层轮廓,并最后整理为COCO(common objects in context)格式。具体步骤如下:首先,仅保留当前类别对应的所有房间,并进行二值化;然后,基于OpenCV库中的FindContours函数进行自动提取每个联通分

量的轮廓,再结合 Shapely 库中的 Simplify 函数简化轮廓的点数量,转为轮廓表示;最后,提取最小包围盒、面积等信息。

3)数据清洗。主要保证数据集能够在本技术路线所需的模型上训练、推理。因为数据集已经转为 COCO 格式,且具有包围盒、矢量轮廓等信息,因此兼容所有目标检测和实例分割模型。而本文算法模型对于每幅图像的房间数量、房间轮廓点的数量有限制,需要在训练前将最大房间数、最大轮廓点数量作为超参数进行模型构建,因此需要将具有过多房间、过多轮廓点数量的图像进行删除处理。本文数据清洗选择最大房间数为 40,最多轮廓点数量为 40。

3.2 实验设置

参数设置上,针对目标检测结果进行的后处理分数阈值设置为 0.5,优化器选择 Adam,训练最大轮次为 400,在单机 4 卡上训练。在实例分割模型中,每张卡上的每批图像数量为 1,学习率设置为 0.000 1,数据集的图像尺寸大小设置为最长边 1 536 像素。在大模型分割模型中,图像尺寸大小为 640 × 640 像素,学习率设置为 0.000 2,数据集的最大房间数为 40,轮廓最大边数为 40。

3.3 户型平面图矢量化重建实验结果

3.3.1 定量测试

使用的评测方法来源于 FloorNet (Chen 等, 2018),测算了重建前分重建后的点的位置 (corner)、角度 (angle)、边 (edge) 和房间 (room) 的精确度 (prec.)、召回率 (recall) 和 F1 分数。首先,对比了当前具有影响力的方法 HEAT (holistic edge attention Transformer) (Chen 等, 2022),由于该方法没有形成闭合的多边形,因此只能测算边 (edge) 的

结果。而 PolyDiffuse (Chen 等, 2023) 等其他方法可以重建出闭合多边形,因此测算角度 (angle) 的结果。对比结果如表 1 所示,经过本文方法的迭代优化后,能生成精度、召回率和 F1 分数更高的实验结果,体现出本文算法的优越性。其次,测试本文算法的通用性,分别使用 SAM (segment anything model) 模型 (Kirillov 等, 2023)、目标检测模型 RT-DETR (Zhao 等, 2024) 和实例分割模型 BoundaryFormer (Lazarow 等, 2022) 的实验结果,即语义掩膜、包围盒和多边形轮廓,本文算法都可以利用引导集扩散模型对其进行重建,明确了本文算法的优越性。如表 2 所示,迭代次数从 0~4 代表使用扩散模型进行重建的次数。次数 0 表示 3 种方式未进行重建前的结果。直接使用未经处理的目标检测的包围盒结果进行生成式重建,各项指标都低于分割精确度更高的 SAM 语义掩膜的结果,而 BoundaryFormer 能够获取到房间更为精确的轮廓,因此重建效果比 SAM 的更好。不论是较为粗糙的包围盒结果,还是存在语义歧义的掩膜结果,甚至是多个形状不一的多边形轮廓,本文算法均能利用引导集扩散模型进行有效性重建。

除上述实验,还设计了消融实验,用以证明倾斜度损失函数的有效性,该部分实验结果如表 3 所示,输入 RT-DETR (real-time detection Transformer) 的包围盒结果显示,经过倾斜度损失函数的训练后,相较于原有结果,房间轮廓更加规整,轮廓倾斜度降低,从而有助于提升矢量化算法的鲁棒性。

3.3.2 定性测试

本文对重建后的结果进行了可视化比较。户型图重建后的效果图如图 7 所示,图 7(a)代表原始户型平面点阵图,图 7(b)代表房间轮廓的真实标注,

表 1 与其他方法的矢量化结果对比

Table 1 Comparison of vectorization results with other methods

方法	房间			位置			边			角度		
	prec./%	recall/%	F1/%	prec./%	recall/%	F1/%	prec./%	recall/%	F1/%	prec./%	recall/%	F1/%
Chen 等人 (2022)	74.40	77.00	75.70	81.80	87.10	84.40	68.80	74.90	71.70	—	—	—
本文	76.10	77.10	76.60	86.00	85.40	85.70	73.30	73.90	73.60	—	—	—
Chen 等人 (2023)	77.07	82.50	79.33	69.70	61.78	64.85	—	—	—	65.26	57.83	60.74
本文	77.23	82.51	79.41	72.60	61.22	65.88	—	—	—	69.24	58.50	62.92

注:加粗字体表示各次实验中各列最优结果,“—”表示无相应数据。

表 2 不同输入下重建效果的比较

Table 2 Comparison of reconstruction effects under different inputs

方法	迭代次数	房间			位置			角度		
		prec./%	recall/%	F1/%	prec./%	recall/%	F1/%	prec./%	recall/%	F1/%
Kirillov 等人(2023)	0	61.20	65.55	63.02	54.81	48.53	50.98	50.88	45.13	47.37
Zhao 等人(2024)		58.55	62.84	60.38	9.33	29.10	14.04	0.04	0.12	0.07
Lazarow 等人(2022)		74.59	80.02	76.88	62.39	61.36	60.87	56.16	54.79	54.62
Kirillov 等人(2023)	1	76.27	81.53	78.43	71.22	59.64	64.37	67.24	56.51	60.91
Zhao 等人(2024)		71.50	76.41	73.54	41.72	58.72	48.15	30.63	42.92	35.28
Lazarow 等人(2022)		77.07	82.50	79.33	69.70	61.78	64.85	65.26	57.83	60.74
Kirillov 等人(2023)	2	76.78	82.05	78.94	72.24	59.83	64.91	68.40	56.86	61.61
Zhao 等人(2024)		75.14	80.30	77.28	59.04	60.76	59.19	52.13	53.61	52.26
Lazarow 等人(2022)		77.15	82.49	79.36	71.84	61.50	65.69	68.02	58.33	62.29
Kirillov 等人(2023)	3	77.08	82.33	79.23	72.71	59.96	65.18	68.97	57.08	61.97
Zhao 等人(2024)		75.77	80.86	77.87	65.17	61.09	62.47	60.55	56.84	58.09
Lazarow 等人(2022)		76.99	82.27	79.17	72.19	61.17	65.68	68.68	58.30	62.57
Kirillov 等人(2023)	4	77.14	82.37	79.28	72.82	59.94	65.21	69.07	57.05	61.99
Zhao 等人(2024)		76.20	81.18	78.30	67.71	61.45	63.85	63.49	57.68	59.91
Lazarow 等人(2022)		77.23	82.51	79.41	72.60	61.22	65.88	69.24	58.50	62.92

表 3 倾斜度损失函数的有效性验证

Table 3 Effectiveness evaluation of the orientation loss function

方法	房间			位置			角度		
	prec./%	recall/%	F1/%	prec./%	recall/%	F1/%	prec./%	recall/%	F1/%
未加入 CI	75.89	79.69	77.35	60.00	59.53	58.94	54.03	53.65	53.12
加入 CI	75.77	80.86	77.87	65.17	61.09	62.47	60.55	56.84	58.09

注:加粗字体表示各列最优结果。

图 7(c)代表重建前的结果,图 7(d)代表重建后结果,图 7(e)代表提取矢量化后的结果。根据户型平面图轮廓重建效果图 7(d)可以看出,本文方法能够使得轮廓更为准确规整,还学习到了周围轮廓的信息。在处理相交的房间时,轮廓能够贴合,同时具备识别倾斜轮廓的能力。虽然识别结果与真实标注存在一定偏差,但整体结构更加准确,说明本文的生成式重建方法在轮廓重建方面具有潜力。

在经过本文算法重建后,可以利用墙线提取算法进一步获得墙线矢量图,以便于下游应用扩展。其效果图如图 8 所示,可以看到使用引导集扩散模

型进行重建后的效果优于未重建的结果,从而提高了后续设计、修改、分析和优化的准确度。

本文同时可视化了模型在每次迭代生成过程中的结果,如图 9 所示。观察结果可以发现,对于不同的户型结构,当存在边界模糊的房间区域时,需要迭代多次才能获得较好的结果;而对于房间边界较为清晰的户型,仅需两步便可实现精确重建。如图 9 第 1 列所示,引入轮廓倾斜度损失函数在一定程度上规整了多边形轮廓,但这并不意味着生成的结果都是直角,当出现斜边的墙体时,本文方法同样可以实现精确的重建。

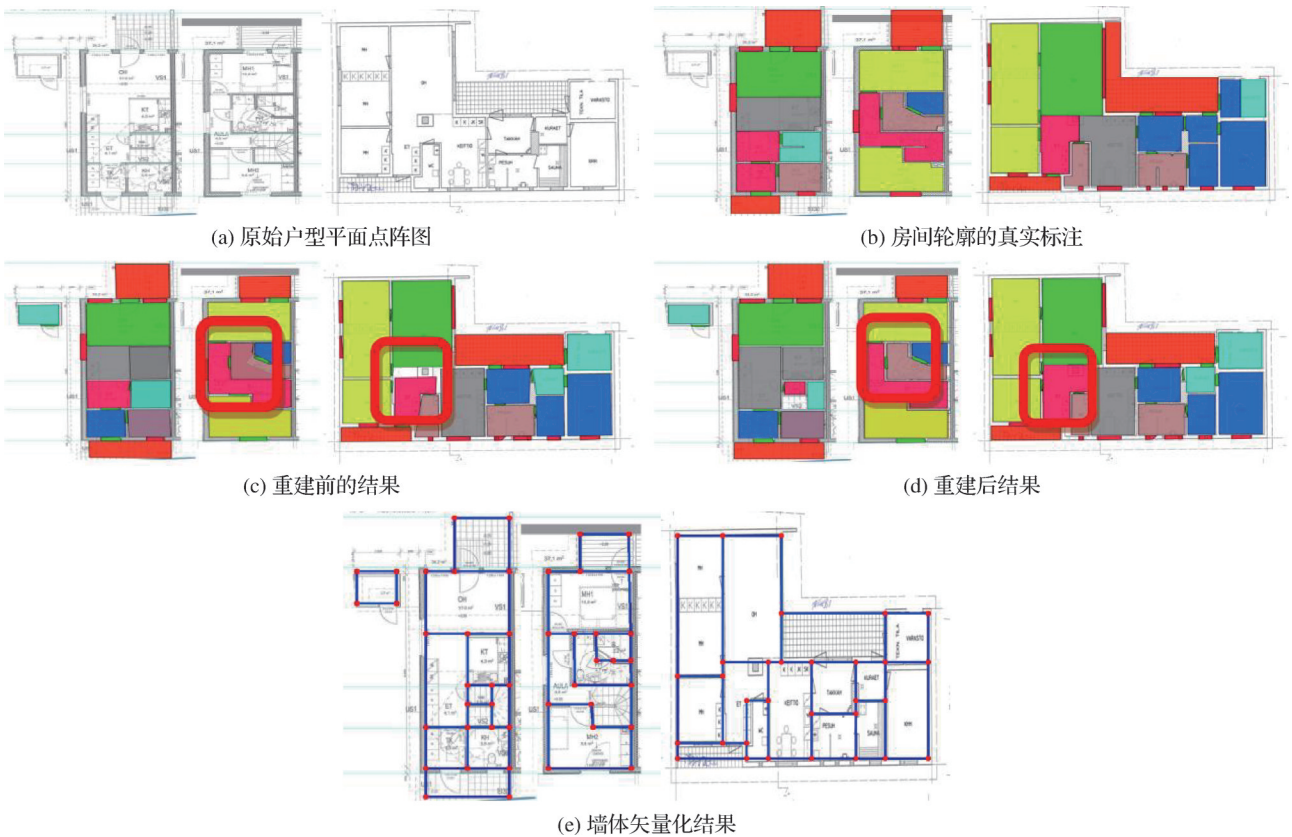


图 7 户型平面图结构重建效果图

Fig. 7 The structural reconstruction of the floor plan layout ((a) original floor plan images; (b) ground truth; (c) network output results; (d) reconstruction results; (e) wall vectorization results)



图 8 提取墙线的矢量化图

Fig. 8 The vectorized diagram of the extracted wall lines((a) the output of SAM; (b) the SAM results of reconstruction; (c) the output of BoundaryFormer; (d) the BoundaryFormer results of reconstruction)

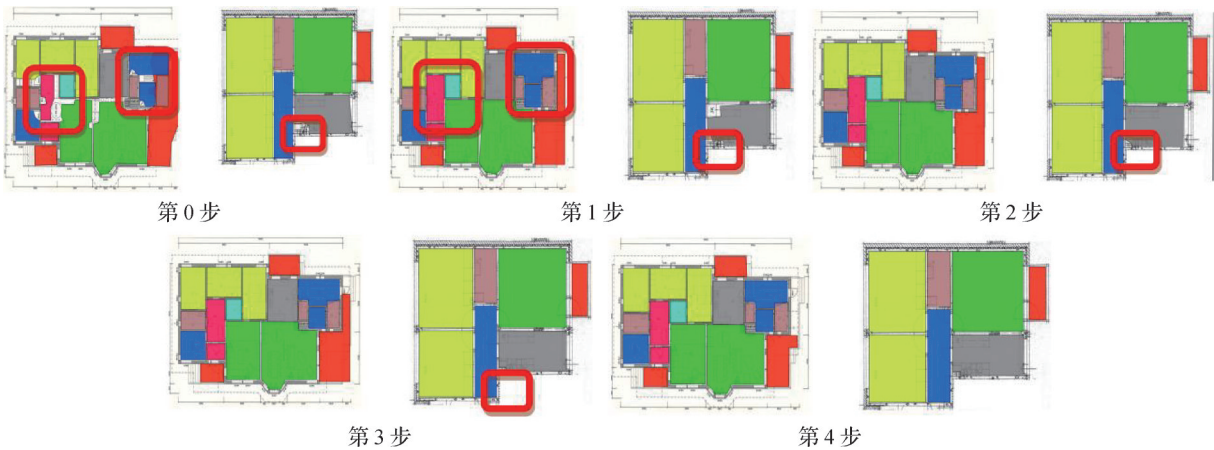


图9 扩散模型迭代步骤可视化效果图

Fig. 9 The visualization of the iterative steps in the diffusion model

4 结 论

在建筑领域相关研究中,本文创新性地基于引导扩散模型,通过逐步迭代的方式细致地重建户型的多边形轮廓。在这一过程中,所提出的轮廓倾斜度损失函数发挥了关键作用。该函数具有独特的优势,它能够有效地促使生成的多边形更为规整,其规整程度相较于传统方法有显著提升。这种规整的多边形为后续墙体的矢量化处理奠定了坚实基础,极大地提高了处理的精确度,使得整个流程更加高效和准确。

然而,通过该方法提取出的墙线结果,实际上是从房间的边界线获得的。这就意味着,当遇到一些特殊情况时,该方法可能会受到限制。例如,如果存在弯曲较大的墙体,本文所采用的算法可能并不适用。在这种情况下,可能会出现墙线提取不准确或无法有效提取的问题,从而影响整个户型重建及后续处理的效果。

因此,为了进一步提升本文方法的效果和适用性,结合墙线检测来对本文方法进行进一步优化,是一个非常值得深入研究的方向。通过深入研究和探索如何更好地结合墙线检测,有望克服当前方法在面对弯曲墙体时的不足,从而使整个方法更加完善和可靠,为建筑户型处理等相关领域提供更强有力的技术支持和解决方案。

参考文献(References)

Chen J C, Deng R Z and Furukawa Y. 2023. PolyDiffuse: polygonal

shape reconstruction via guided set diffusion modelS//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates: 1863-1888 [DOI: 10.5555/3666122.3666212]

Chen J C, Qian Y M and Furukawa Y. 2022. HEAT: holistic edge attention transformer for structured reconstruction//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 3856-3865 [DOI: 10.1109/CVPR52688.2022.00384]

Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 833-851 [DOI: 10.1007/978-3-030-01234-2_49]

Du T, Hu R Z, Liu L B, Yi L and Zhao H. 2024. Research progress in human-like indoor scene interaction. Journal of Image and Graphics, 29(6): 1575-1606 (杜韬, 胡瑞珍, 刘利斌, 弋力, 赵昊. 2024. 室内场景拟人交互研究进展. 中国图象图形学报, 29(6): 1575-1606) [DOI: 10.11834/jig.240004]

He K M, Gkioxari G, Dollár P and Girshick R. 2017. Mask R-CNN//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2980-2988 [DOI: 10.1109/ICCV.2017.322]

Ho J, Jain A and Abbeel P. 2020. Denoising diffusion probabilistic models//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: #574 [DOI: 10.5555/3495724.3496298]

Kalervo A, Ylioinas J, Häikiö M, Karhu A and Kannala J. 2019. Cubi-Casa5K: a dataset and an improved multi-task model for floorplan image analysis//Proceedings of the 21st Scandinavian Conference on Image Analysis. Norrköping, Sweden: Springer: 28-40 [DOI: 10.1007/978-3-030-20205-7_3]

Kirillov A, Mintun E, Ravi N, Mao H Z, Rolland C, Gustafson L, Xiao T T, Whitehead S, Berg A C, Lo W Y, Dollár P and Girshick

- R. 2023. Segment anything//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE: 3992-4003 [DOI: 10.1109/ICCV51070.2023.00371]
- Lazarow J, Xu W J and Tu Z W. 2022. Instance segmentation with mask-supervised polygonal boundary transformers//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 4372-4381 [DOI: 10.1109/CVPR52688.2022.00434]
- Liu C, Wu J J, Kohli P and Furukawa Y. 2017. Raster-to-vector: revisiting floorplan transformation//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2214-2222 [DOI: 10.1109/ICCV.2017.241]
- Lyu X L, Zhao S C, Yu X Y and Zhao B Q. 2021. Residential floor plan recognition and reconstruction//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 16712-16721 [DOI: 10.1109/CVPR46437.2021.01644]
- Pizarro P N, Hitschfeld N, Sipiran I and Saavedra J M. 2022. Automatic floor plan analysis and recognition. Automation in Construction, 140: #104348 [DOI: 10.1016/j.autcon.2022.104348]
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I. 2017. Attention is all you need//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates: 6000-6010 [DOI: 10.5555/3295222.3295349]
- Wan J H, Liu X P, Chen L L, Ao S, Zhang P and Guo Y L. 2024. Geometric attribute-guided 3D semantic instance reconstruction. Journal of Image and Graphics, 29(1): 218-230 (万骏辉, 刘心溥, 陈莉丽, 敖晟, 张鹏, 郭裕兰. 2024. 几何属性引导的三维语义实例重建. 中国图象图形学报, 29(1): 218-230) [DOI: 10.11834/jig.230106]
- Ye Y X, Du X, Chen S, Zhu S Z and Yan Y. 2024. Sparse adversarial patch attack based on QR code mask. Journal of Image and Graphics, 29(7): 1889-1901 (叶乙轩, 杜侠, 陈思, 朱顺慈, 严严. 2024. 二维码掩膜下的稀疏对抗补丁攻击. 中国图象图形学报, 29(7): 1889-1901) [DOI: 10.11834/jig.230453]
- Yu L and Wu X Q. 2024. Survey of texture optimization algorithms for 3D reconstructed scenes. Journal of Image and Graphics, 29(8): 2303-2318 (于柳, 吴晓群. 2024. 三维重建场景的纹理优化算法综述. 中国图象图形学报, 29(8): 2303-2318) [DOI: 10.11834/jig.230478]
- Yue L, Tan H, Huang J K and Zhang S K. 2024. Review of digital 3D indoor scene synthesis. Journal of Image and Graphics, 29(9): 2471-2493 (岳亮, 谈皓, 黄俊凯, 张少魁. 2024. 数字室内三维场景构建综述. 中国图象图形学报, 29(9): 2471-2493) [DOI: 10.11834/jig.230712]
- Yue Y W, Kontogianni T, Schindler K and Engelmann F. 2023. Connecting the dots: floorplan reconstruction using two-level queries//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 845-854 [DOI: 10.1109/CVPR52729.2023.00088]
- Zhao Y A, Lv W Y, Xu S L, Wei J M, Wang G Z, Dang Q Q, Liu Y and Chen J. 2024. DETRs beat YOLOs on real-time object detection//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 16965-16974 [DOI: 10.1109/CVPR52733.2024.01605]
- Zhu R Y, Shen J C, Deng X T, Walldén M and Ino F. 2020. Training strategies for CNN-based models to parse complex floor plans//Proceedings of the 9th International Conference on Software and Computer Applications. Langkawi, Malaysia: Association for Computing Machinery: 11-16 [DOI: 10.1145/3384544.3384566]

作者简介

王静,女,博士研究生,主要研究方向为计算机图形学。

E-mail: wangjinginszu@gmail.com

黄惠,通信作者,女,教授,博士生导师,主要研究方向为计算机图形学。E-mail: huihuang@szu.edu.cn

熊皓然,男,博士研究生,主要研究方向为计算机图形学。

E-mail: xiong2002@gmail.com