

基于指数函数密度聚类模型的室内物体提取

陈西江^{1,2,4}, 林嘉颖^{2*}, 韩贤权³, 王昊骏²¹ 武昌理工学院人工智能学院, 湖北 武汉 430223;² 武汉理工大学安全科学与应急管理学院, 湖北 武汉 430070;³ 长江科学院, 湖北 武汉 430010;⁴ 湖北中图勘测规划设计有限公司, 湖北 武汉 430070

摘要 室内点云包括墙壁、天花板、地板和房间内的许多物体。室内墙壁、天花板、地板等多种物体的提取是室内导航、设施管理、建筑改造等多种应用的关键。针对此, 本文根据截止距离内的局部密度, 采用指数函数构建密度聚类模型。首先, 利用边界点云与室内点云之间的距离构造墙壁面密度聚类的约束条件; 其次, 根据 z 值的指数函数和局部密度模型提取天花板和地板; 然后, 构建不同对象的局部密度模型, 根据局部密度的大小确定约束距离; 最后, 根据点到聚类中心(聚类中心被认为是局部密度与约束距离乘积异常大的点)的距离, 得到每个点归属于聚类中心的聚类; 最后, 通过判断相邻聚类之间的距离实现室内物体点云的提取, 从而完成对不同类型场景室内点云的提取, 并将本文方法与快速搜索密度峰值聚类(CFDP)算法、基于密度的空间聚类(DPC)算法进行比较。对比结果表明, DPC 方法不如本文方法但优于 CFDP 方法。计算了本文方法提取不同类型室内点云的精确度、召回率和 F1 分数, 结果表明, 物体间的靠近程度会影响本文方法的性能。

关键词 测量; 目标提取; 点云; 密度聚类; 激光扫描

中图分类号 TP274

文献标志码 A

DOI: 10.3788/CJL202249.1104001

1 引言

聚类分析是数据挖掘中最常用的技术之一, 用于发现数据集中的未知对象类。聚类分析已被广泛应用于模式识别^[1]、异常检测^[2]等领域。聚类不仅可以用于对数据进行分类, 还可以通过识别特征和检测异常^[3]来洞察数据。因此, 聚类可以用来观察现有数据源中数据点之间的未知关系。数据的分布性质不同, 聚类算法的性能也会有所不同。

聚类方法主要包括基于分区的聚类方法、基于层次的聚类方法和基于密度的噪声应用空间聚类方法(DBSCAN)。基于分区的聚类方法将数据库 D 中的 n 个对象划分为 k 个簇, 该方法从数据库 D 的初始划分开始, 使用迭代控制策略优化目标函数。每个簇用数据集重心(k-means 算法, 模糊 C-means

算法)^[4]或簇中的一个对象(k-medoid 算法)^[5]表示。基于分区的聚类方法不需要遵循层次结构, 将所有聚类作为数据的一个分区, 根据适当的距离便可以覆盖聚类全局形状的所有信息。该算法的优点是简单高效, 具有较低的时间复杂度和空间复杂度, 缺点是当数据量较大时, 聚类结果容易陷入局部最优。针对此问题, 文献[6]提出了一种基于网格的 k-means 快速聚类算法。该算法通过对海量数据进行网格划分, 并基于每个网格的质心实现聚类, 从而缩减欧氏距离的计算次数, 实现海量数据的快速聚类。k-means 算法对不均衡数据聚类时容易产生“均匀效应”。为此, 文献[7]利用待定聚类数据集的调整机制, 结合近邻思想对对象所属类别进行重新划分, 避免了“均匀效应”对聚类的影响。噪声是影响聚类的主要因素, 特别是对基于隶属度的模糊

收稿日期: 2021-10-08; 修回日期: 2021-10-28; 录用日期: 2021-11-15

基金项目: 国家自然科学基金(42171428)、湖北省安全生产专项资金科技项目(SJIX 20211006)、重庆市技术创新与应用发展专项面上项目(cstc2019jsex-msxmX0051)、长江科学院开放研究基金资助项目(CKWV2019758/KY)、国家大学生创新创业训练计划(312040000217)

通信作者: *1729959215@qq.com

C-means 算法的聚类影响较大。为了抑制噪声对模糊 C-means 聚类算法的影响,文献[8]利用 K 最近邻方法来自适应调整松弛参数,结合隶属度的稀疏性,达到了抑制噪声的目的。文献[9]在超像素分区的基础上比较了区域间光谱特征的相似性,提出了融合光谱特征的模糊 C 均值聚类算法,并采用该算法实现了对高分辨率遥感影像的准确聚类。

层次聚类(HCA)是一种聚类分析算法,它使用树状结构来描述簇之间的关系。层次聚类算法用于寻找集群层次,不需要预先确定集群数量,也不存在局部最小问题。层次聚类算法通常分为两类:凝聚集群和分裂集群。凝聚集群采用的是一种“自下而上”的策略,即每次观察都从自己的集群开始,随着层次的上升,一对集群融合在一起。相反,分裂集群采用的是“自上而下”的策略,所有的观察都从一个集群开始,随着层级的下降,分裂不断进行。分级聚类结果通常以树状图的形式显示,将聚类描述为树状数据结构^[10]的节点。目前,研究人员已经提出了许多基于层次的聚类方法,如 BIRCH^[11]、CURE^[12]等。层次聚类在图像及点云数据中的应用较少,不过还是能检索到利用层次聚类实现图像分割的公开报道,如黄亮等^[13]利用层次聚类对超像素进行合并,在像素合并的基础上实现了高分辨率遥感影像的精确分割。

基于密度的空间聚类(DPC)采用密度阈值来检测聚类边界。具体来说,DPC 算法^[14]使用邻域半径 Eps 和邻域内点的最小数目 MinPts 定义了一个密度阈值,然后在此基础上定义并检测核心点、边界点和噪声点。该算法的优点是可以处理数据集中的大量噪声。基于密度的聚类算法对噪声不敏感,可以找到任意形状的聚类。然而,基于密度的聚类对邻域距离和最小值等参数较为敏感,其鲁棒性较差。鉴于此,文献[15]提出了具有一定鲁棒性的密度引力聚类算法,该算法的优点是针对不同局部密度确定初始簇,然后对初始簇进行合并,避免了聚类较为分散的现象。但是该算法的核心是基于密度的聚类,存在参数对邻域距离和最小值等敏感的问题。针对聚类对邻域距离较为敏感的问题,文献[6]构建了基于基尼系数的自适应截止距离调节的密度峰值聚类算法,该算法通过基尼系数结合簇中心权重自适应调节截止距离,从而避免了算法对截止距离敏感的问题。

快速搜索密度峰值聚类(CFDP)^[16]算法与基于

密度的空间聚类的集群原理一样,能够自动找到正确的集群数量。实验表明,该算法具有良好的性能。但是,该聚类算法的缺点是局部密度、相邻点到聚类中心的距离和截止距离等参数在很大程度上依赖于主观经验^[17]。为了克服 CFDP 的局限性,研究人员提出了一种基于局部引力和距离的聚类算法,该算法不仅考虑了截止距离的选择,还考虑了簇的密度估计^[18]。为了解决截止距离确定的难题,文献[19]利用 K 最近邻算法计算每个点的截止距离和局部密度。另外,CFDP 算法没有考虑稀疏区域和稠密区域聚类中心的差异,因此忽略了稀疏区域的聚类中心。针对这一问题,文献[20]将点的局部密度定义为邻近区域中点的数量,从而将稀疏区域和稠密区域的聚类中心同等对待。以上聚类算法主要通过设计新的局部密度估计来改进原有的 CFDP 算法,密度对聚类结果有较大影响。文献[14]在聚类中心识别中仅利用相对密度作为局部密度,减小了密度峰值聚类算法对参数的依赖。该方法主要适用于聚类特征明显的数据集,但是很难对分布均匀的二维或三维数据进行聚类。针对这一情况,本课题组利用指数函数构建了适用于室内点云目标提取的聚类算法。该聚类算法优化了截止距离和局部密度,并给出了均匀化聚类中心。本课题组构建的这一聚类算法可以提取出房间内的墙壁、天花板、地板和物体。

2 本文方法

为了建立一种适用于室内目标提取的聚类算法,本文提出了一种用于室内点云提取的基于指数函数的聚类方法。所提方法的总体框架如图 1 所示。

2.1 截止距离的确定

假设角度分辨率为 α ,扫描距离为 D ,相邻点个数为 k ,如图 2 所示。图 2 中的右图显示了点 O 及其相邻点。

根据角度分辨率和扫描距离,计算点 O 与其最近邻点之间的距离。计算公式为

$$d_{\min} = D \sin(\alpha), \quad (1)$$

式中: α 为角度分辨率; D 为扫描距离。

根据相邻点的个数,计算相邻区域周边点到点 O 的最近距离。计算公式为

$$d'_{\min} = D \sin\left(\frac{\sqrt{k+1}-1}{2}\alpha\right), \quad (2)$$

式中: k 是点 O 的相邻点的个数。同时,计算得到相邻区域周边点到点 O 的最远距离。计算公式为

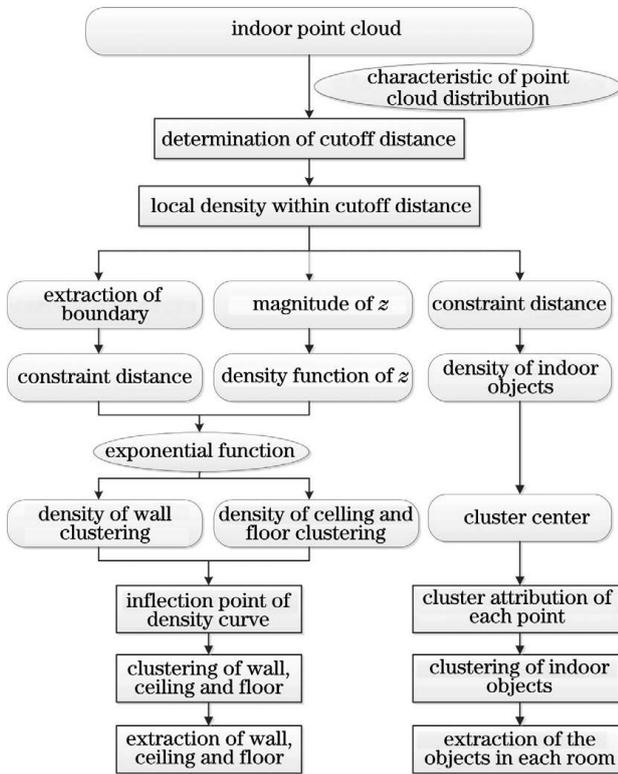


图 1 所提室内目标点云提取方法的框架

Fig. 1 Framework of proposed method for the extraction of indoor objects point cloud

$$\tilde{d}_c = \sqrt{2} D \sin\left(\frac{\sqrt{k+1}-1}{2}\alpha\right). \quad (3)$$

本文将最远距离称为截止距离。

2.2 截止距离内的局部密度

给定一含有 N 个数据点的数据集: $S = \{P_i\}_{i=1}^N$ 。对于数据集 S 中的任何数据点 P_i , d_{ij} 表示点 P_i 与截止距离内的任何其他点之间的距离, k_i 表示截止距离内点的个数。所有距离的平均值被认为是局部密度的一个指标, 即

$$\bar{d}_i = \frac{1}{k_i} \sum_{i=1}^k |d_{ij}|. \quad (4)$$

通过考虑点的距离和截止距离内点的个数, 得到初始局部密度模型为

$$\rho_i = \exp(-\tilde{d}_i) + \exp\left(-\frac{1}{k_i}\right). \quad (5)$$

对于室内物体, 本文不仅要考虑点的距离和数量, 还要考虑点的分散程度。对于数据集 $S = \{P_i\}_{i=1}^N$, 在截止距离内点 P_0 的相邻点为 $[P_1 P_2 \dots P_k]$ 。如图 3 所示, 以点 P_0 为起点, 相邻点 $[P_1 P_2 \dots P_k]$ 为终点, 定义点 P_0 与相邻点之间的向量为 $\mathbf{V}_{P_0 P_j}$ ($j=1, 2, \dots, k$)。将 $\mathbf{V}_{P_0 P_1}$ 定义为起始方向, 由此可以得到起始方向与其他矢量的夹角, 即

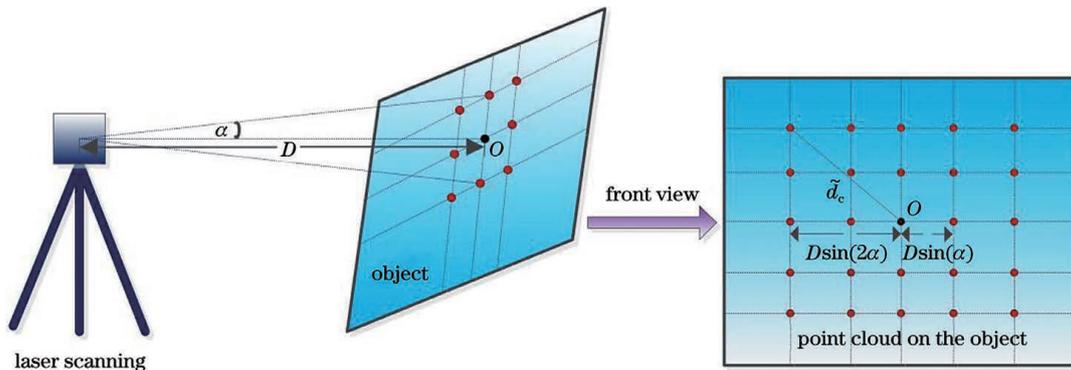


图 2 点 O 及其相邻点

Fig. 2 Point O and its adjacent points

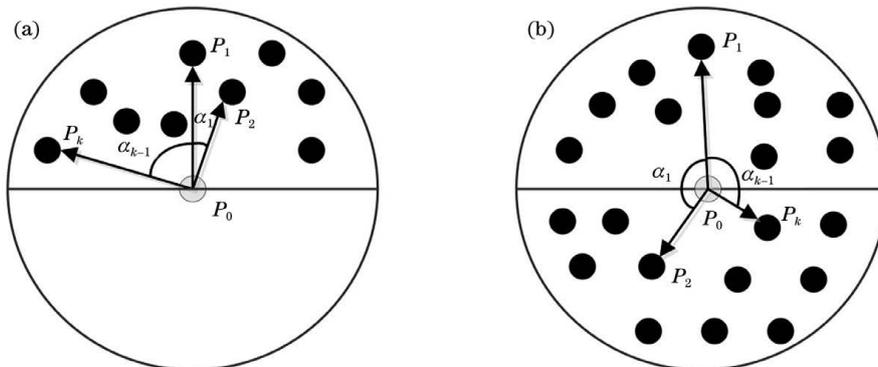


图 3 相邻点分布。(a)邻域偏向一边;(b)邻域分布均匀

Fig. 3 Distribution of adjacent points. (a) Neighbors bias to one side; (b) neighbors distribute evenly

$$\alpha_j = \arccos\left(\frac{\mathbf{V}_{\rho_0 \rho_1} \mathbf{V}_{\rho_0 \rho_j}}{|\mathbf{V}_{\rho_0 \rho_1}| |\mathbf{V}_{\rho_0 \rho_j}|}\right), (j = 1, 2, \dots, k)。$$

(6)

从图 3(a)可以看出,当邻域偏向一侧时,起始方向与其他向量的夹角均为锐角。如图 3(b)所示,如果有锐角和钝角,则相邻点围绕点 P_0 。尽管图 3(a)中点 P_0 的局部密度与图 3(b)中点 P_0 的局部密度相似,但图 3(a)中的点 P_0 不能被识别为聚类中心。为了准确地确定聚类中心,利用向量之间的夹角来确定距离的正负,即

$$\begin{cases} d_{ij} \text{ is positive, } \alpha_j < 90^\circ \\ d_{ij} \text{ is negative, } \alpha_j \geq 90^\circ \end{cases} \quad (7)$$

考虑到距离和方向,距离 d_{ij} 平均值的计算公式为

$$\mu_i = \frac{\sum_{j=1}^k d_{ij}}{k_i}。 \quad (8)$$

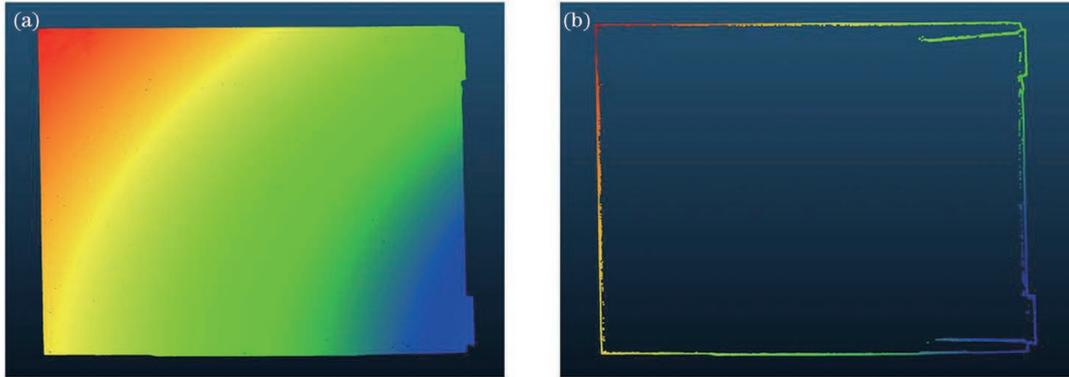


图 4 墙体边界的提取。(a)墙体的原始点云;(b)墙体边界

Fig. 4 Extraction of wall boundary. (a) Original point cloud of wall; (b) wall boundary

其次,计算点云中各点与边界点之间的距离,以最小距离 δ_i 作为判断墙体的约束因子。根据截止距离和距离约束下的局部密度,得到了墙体提取的综合密度函数为

$$\rho_i^w = \rho_i + \frac{1}{\exp(\delta_i)}。 \quad (11)$$

最后,依据 ρ_i^w 得到了墙壁的面密度分布。如图 5 所示,本文采用次优特征选择方法确定拐点^[22]。如图 6 所示,拐点右侧对应墙体点云。

2.3.2 地板和天花板的提取

地板和天花板聚类的主要约束条件是 z 值的大小,因此,本文用指数函数来构造 z 值的密度函数。构造的密度函数为

$$\tilde{\rho}_i = \begin{cases} \exp\left(-\frac{1}{z_i}\right), \text{ used for clustering ceiling} \\ \exp(-z_i), \text{ used for clustering floor} \end{cases}。 \quad (12)$$

同时,将距离 d_{ij} 作为误差,利用标准差来描述点 P_0 的相邻点数据集的离散度。标准差的计算公式为

$$\sigma_i = \sqrt{\frac{1}{k_i - 1} \sum_{j=1}^k d_{ij}^2}。 \quad (9)$$

最后,通过考虑截止距离内点的分布和距离,得到室内目标点云的局部密度模型,即

$$\rho_i^o = \rho_i + \exp(-\sigma_i) + \exp(-|\mu_i|)。 \quad (10)$$

2.3 截止距离的确定

2.3.1 墙壁的提取

由于外墙均在边界处,因此本文以点云到边界的距离作为判断外墙面的约束条件,如图 4 所示。首先,将点云投影到 $x-y$ 平面上,根据文献[21]所提算法提取边界点。为了保证边界提取的精度,避免错误的边界对墙壁提取造成影响,将文献[21]中由投影面点集法向量构成的方位角设置得尽可能大,本文将其设置为 130° 。

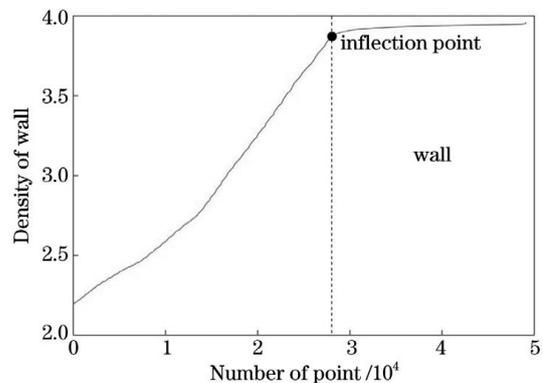


图 5 墙壁面密度函数的幅值分布

Fig. 5 Magnitude distribution of density function of wall

根据截止距离内的局部密度和 z 值的密度函数,得到了天花板和地板聚类的密度函数,即

$$\rho_i^{\text{fc}} = \begin{cases} \rho_i + \exp\left(-\frac{1}{z_i}\right), \text{ ceiling clustering} \\ \rho_i + \exp(-z_i), \text{ floor clustering} \end{cases}。 \quad (13)$$

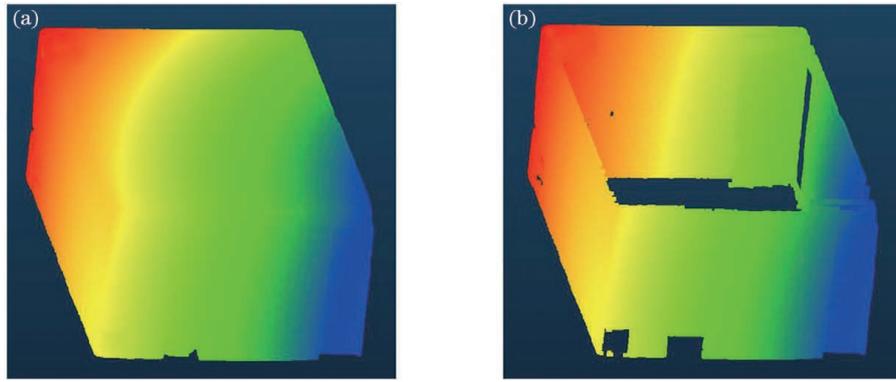


图 6 根据墙体点云密度簇提取墙体。(a)原始点云;(b)提取的墙体点云

Fig. 6 Extraction of wall according to density cluster of wall point cloud. (a) Original point cloud; (b) extracted wall point cloud

根据式(13)得到了地面和天花板的密度分布,如图 7 所示。

由图 7 可知,地面和天花板密度曲线上存在不止一个拐点。利用次优特征选择方法确定地面和

天花板密度曲线的最高拐点,并根据地面和天花板密度曲线的最高拐点提取地面和天花板点云,最高拐点右侧对应着地面和天花板点云,如图 8 所示。

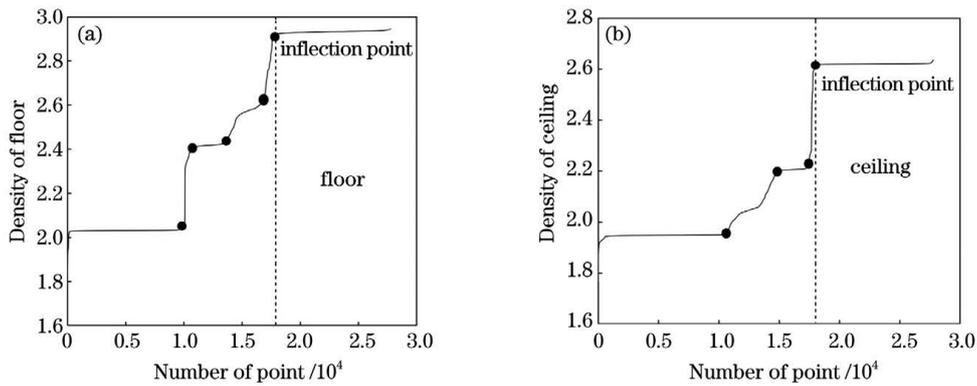


图 7 密度曲线。(a)地板的密度曲线;(b)天花板的密度曲线

Fig. 7 Density curves. (a) Density curve of floor; (b) density curve of ceiling

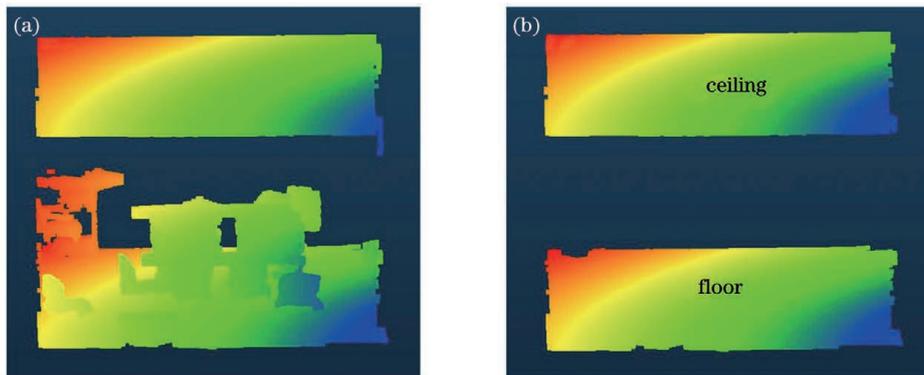


图 8 地面和天花板点云的提取结果。(a)去除墙壁后的天花板和地面原始点云;(b)提取的天花板和地面点云

Fig. 8 Extracted ceiling and floor point cloud. (a) Original point clouds of ceiling and floor after removing walls; (b) extracted ceiling and floor point clouds

2.3.3 室内物体的提取

室内物体提取的关键是室内点云的聚类,室内点云聚类的前提是确定聚类中心,聚类中心主要通

过约束距离和局部密度来确定。

1) 约束距离 δ_i

约束距离 δ_i 根据截止距离内局部密度的大小

来确定。当点 P_i 为最大密度时, δ_i 为该点与任意其他点之间的最大距离, 否则, δ_i 为点 P_i 与任意其他密度更高的点之间的最小距离, 即

$$\delta_i = \begin{cases} \min \{d_{ij}\}, \rho_i \neq \max(\rho_j) \\ \max \{d_{ij}\}, \rho_i = \max(\rho_j) \end{cases}, \quad (14)$$

式中: d_{ij} 为点 P_i 与其他点之间的距离。

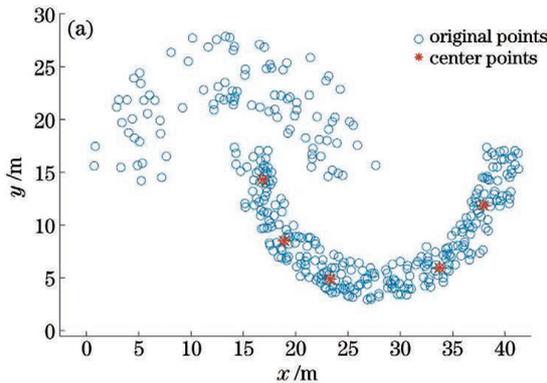
假设点 P_i 和 P_m 具有相同的最大局部密度, 即

$$\rho_i^\circ = \rho_m^\circ = \max(\rho_j^\circ), \quad (15)$$

依据这一前提条件[式(15)], 计算点 P_i 与点 P_m 之间的距离 d_{im} , 得到距离 d_{im} 与截止距离 \tilde{d}_c 之间的差值为

$$\Delta d = d_{im} - \tilde{d}_c. \quad (16)$$

如果 $\Delta d \leq 0$, 则点 P_i 与 P_m 属于同一对象, 否则, 点 P_i 和 P_m 属于两个不同的对象。



2) 聚类中心的确定

考虑到局部密度 ρ_i° 和约束距离 δ_i 的尺度不同, 将 ρ_i° 和 δ_i 标准化到 $0 \sim 1$ 范围内, 得到标准化后的局部密度 $\tilde{\rho}_i$ 和约束距离 $\tilde{\delta}_i$ 。

根据公式

$$\tilde{\gamma}_i = \tilde{\rho}_i \cdot \tilde{\delta}_i \quad (17)$$

得到聚类中心, 聚类中心为 $\tilde{\gamma}_i$ 值异常大的点。

假设由 $\tilde{\gamma}_k$ 和 $\tilde{\gamma}_m$ 确定的聚类中心是 $C_k(x_k, y_k, z_k)$ 和 $C_m(x_m, y_m, z_m)$, 并且这两个聚类中心之间的距离为 $d_{C_{km}}$ 。如果 $d_{C_{km}} \leq \tilde{d}_c$, 则将这两个集群视为一个集群, 并将它们合并在一起, 否则, 认为它们属于两个不同的集群。

如图 9 所示, 为了验证本文方法确定聚类中心的优点, 使用 Jain 数据^[23]进行评价, 并将评价结果与 CFDP 方法的评价结果^[16]进行对比, 对比结果如图 9 所示。

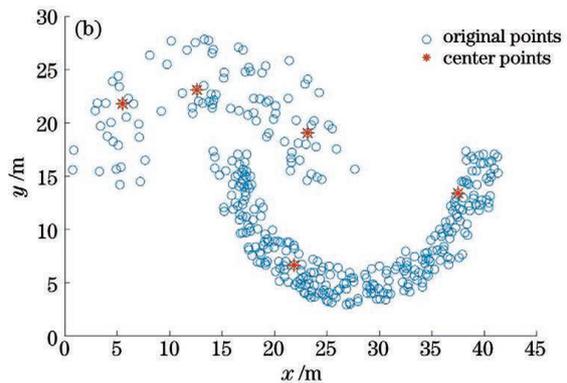


图 9 不同方法确定的聚类中心。(a) CFDP 方法确定的中心点; (b) 本文方法确定的中心点

Fig. 9 Cluster centers determined by different methods. (a) Center points determined by CFDP method; (b) center points determined by proposed method

由图 9(a)可知, CFDP 方法确定了 5 个聚类中心, 而且这 5 个中心集中在一个高密度区域, 其中的两个中心几乎非常接近另外两个中心。造成这种现象的原因是 CFDP 方法只考虑了截止距离内的密度。因此, 稀疏点区域没有分配到中心点, 簇中心分布非常不均匀。与此相反, 本文方法确定的中心点较为均匀, 中心点几乎覆盖了所有原始点区域, 如图 9(b)所示。

3) 基于聚类中心的点云聚类

根据聚类中心确定方法得到聚类中心 $[C_1 C_2 \dots C_m]$ 和其他数据点 $[C_{m+1} C_{m+2} \dots C_N]$ 。对于非中心数据点 $C_i, m < i \leq N$, 计算非中心数据点与其中心数据点 $C_j (1 \leq j < i)$ 之间的距离。根据最小距离, 逐个确定其他点的关联簇。计算点 C_{m+1}

与点 C_j 之间的距离, 记为 $d_{m+1, j}$ 。若 $d_{m+1, 2}$ 为最小距离, 则表示该点 C_2 和 C_{m+1} 属于同一簇。若 $d_{m+1, m+2}$ 为最小距离, 则表示该点 C_2 与 C_{m+1} 属于同一簇。同理, 可以逐个对其他点进行聚类分配。根据所有点的聚类分配, 可以得到每个聚类的点。

本文使用 K 最邻近点搜索算法(KNNS)来确定每个簇的最近邻簇, 并得到相邻簇之间的最近邻距离。如果最近邻距离小于截止距离, 则相邻的两个聚类属于同一类, 并将它们合并在一起。最终, 房间中的每个物体都可以被提取出来。

3 比较分析

采用 FARO Focus X330 HDR 扫描仪采集不同的室内点云场景(公寓、卧室、会议室、大堂、阁

楼),扫描完成后启动配套软件 SCENE,打开点云配准功能,输入扫描标靶中心点数据,利用基于标靶的配准模块完成相邻测站扫描点云数据的配准,完成配准后将数据导出,如图 10 所示。该扫描仪利用反射目标上的激光来获取地面点云,扫描频率为 122.00~976.00 Hz,扫描范围为 0.6~330 m,激光

波长为 1550 nm,扫描速度为 970000 point/s,扫描视场角为 $360^{\circ} \times 300^{\circ}$ 。由于是室内扫描,所以测距精度达到了 0.1 mm,点云中每个点的点位精度为 2 mm,点与点之间的间隔为 3 mm。在扫描过程中,点云难免不会出现漏扫,进而造成了一定的点云空洞。



图 10 不同的室内场景(公寓、卧室、会议室、大堂、阁楼)

Fig. 10 Different indoor scenes (apartment, bedroom, boardroom, lobby, and loft)

为了对室内目标进行提取,首先提取墙面、天花板和地面。根据 2.3.1 节及 2.3.2 节提出的方法对

公寓和卧室的墙面、天花板和地面进行提取,提取结果如图 11 所示。

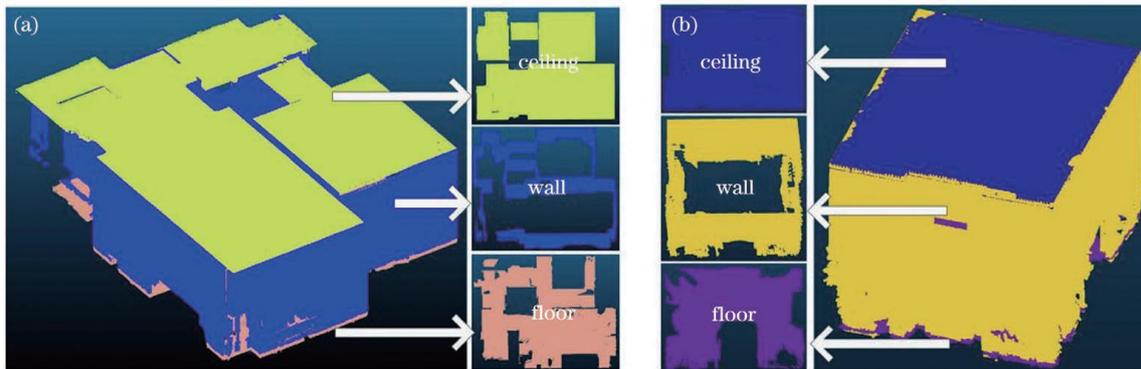


图 11 不同颜色显示的公寓和卧室的墙面、天花板和地面提取结果。(a)公寓墙面、天花板和地面的提取结果;(b)卧室墙面、天花板和地面的提取结果

Fig. 11 Different colors show extracted walls, ceilings, and floors of apartment and bedroom. (a) Extracted walls, ceilings, and floors of apartment; (b) extracted wall, ceiling, and floor of bedroom

图 11(a)所示的公寓包含了不同的空间格局,且不同空间格局的天花板的高度不同,但从其颜色可以看出公寓上部天花板基本都被提取出来,同时侧墙面也被较为准确地提取出来,底部的地面点云基本上也被提取出来。对于图 11(b)所示的空间格局较为简单的卧室,虽然上部天花板侧面包含了少许墙面点云,但其总体被准确地提取出来,同样,侧墙面和底部地面也被准确提取出来。为了准确地对墙面、天花板和地面的提取效果进行评价,将提取的

墙面、天花板、地面点云与实际点云进行比较,从而得到提取效果,如表 1 所示。

由表 1 可以明显看出,墙面、天花板和地面点云的提取率基本在 95%左右。对于公寓,提取的墙面、天花板和地面点云数量要少于实际点云数量,存在点云漏提取的现象。对于卧室,提取的天花板和地面点云的数量也少于实际点云数量,也存在点云漏提取的现象。而提取的墙面的点云数量多于实际点云数量,存在过提取现象。造成这种现象的主

表 1 墙面、天花板和地面的提取效果

Table 1 Extraction effects of wall, ceiling, and floor

Item		Actual number of points	Number of points extracted	Extraction ratio / %
Apartment	Ceiling	165723	163220	98.50
	Wall	431545	427853	99.10
	Floor	149421	147339	98.60
Bedroom	Ceiling	39944	38170	95.50
	Wall	180153	182244	98.80
	Floor	32983	32102	97.30

要原因是天花板和地面的一部分点云被认为是墙面点云。

对于公寓来说,房间里有椅子、桌子、沙发等。

从图 12 中可以看出,CFDP、DPC 和本文方法分别提取了公寓内的 7 个、13 个和 14 个物体。这三种方法都成功提取了盒子,且提取结果与房间角落中

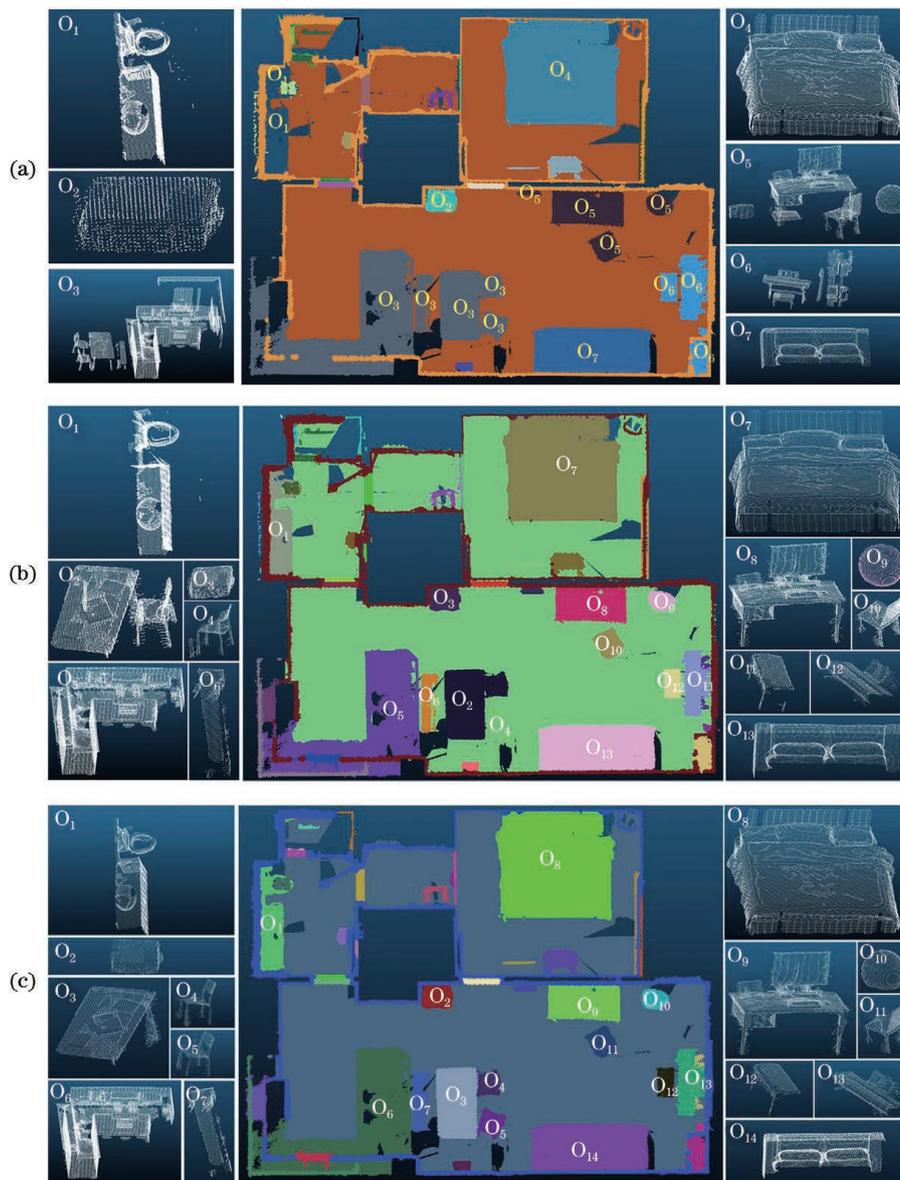


图 12 采用不同的方法提取公寓中的不同对象。(a) CFDP 方法;(b) DPC 方法;(c)本文方法

Fig. 12 Different objects in the apartment are extracted by different methods. (a) CFDP method; (b) DPC method; (c) proposed method

的实际盒子基本相同,如图 12(a)、(c)中的 O_2 以及图 12(b)中的 O_3 所示。床和沙发也被成功提取,如图 12(a)中的 O_4 和 O_7 、图 12(b)中的 O_7 和 O_3 及图 12(c)中的 O_8 和 O_{14} 所示。洗漱台和马桶合并在一起被提取出来,如图 12(a)中的 O_1 所示。但是,床上的枕头或沙发上的坐垫因紧靠在床上或沙发上而未能将它们分开。同时,CFDP 方法没有将椅子、盒子、桌子和球分开,而且未能单独提取厨房台,如图 12(a)中的 O_5 和图 12(a)中的 O_3 所示。然而,DPC 和本文方法都成功地分离了它们,并获得了几乎相同的提取效果,如图 12(b)中的 O_8 、 O_9 、

O_{10} 以及图 12(c)中的 O_9 、 O_{10} 、 O_{11} 所示。造成这种现象的一个可能原因是,相邻目标物距离较近,影响了 CFDP 方法的分割效果。DPC 和本文方法都成功地提取了低矮的长凳,如图 12(b)中的 O_6 和图 12(c)中的 O_7 所示;DPC 方法提取的桌椅合并在一起,如图 12(b)中 O_2 所示,而本文方法可以成功地将它们分离开来,并以不同的颜色显示,如图 12(c)中 O_3 、 O_4 和 O_5 所示。

对于卧室,三种方法都能准确地提取出椅子,如图 13 中的 O_1 所示,也能准确地提取靠墙的柜子,如图 13(a)中的 O_3 和图 13(b)、(c)中的 O_4 所示,



图 13 采用不同的方法提取卧室中的不同物体。(a) CFDP 方法;(b) DPC 方法;(c)本文方法

Fig. 13 Different objects in the bedroom are extracted by different methods. (a) CFDP method; (b) DPC method; (c) proposed method

还可以提取吊灯,如图 13(a)中的 O_2 和图 13(b)、(c)中 O_3 所示。三种方法都无法对床进行提取,CFDP 方法的提取结果中包括 2 个床头柜和 1 把椅子,如图 13(a)中的 O_4 所示,DPC 方法与本文方法将床头柜和床一并提取了且提取效果相同,如图 13(b)、(c)中的 O_5 所示(产生这种提取结果的原因是床头柜和床靠得太近)。同时,DPC 和本文方法都成功地提取出了靠近床的物体(椅子和床头柜)如图 13(b)、(c)中的 O_6 和 O_7 所示。CFDP 和 DPC 方法将梳妆台和椅子一起提取出来,如图 13(a)中的 O_5 及图 13(b)中的 O_8 所示,而本文方法则成功地分离了这两个物体,如图 13(c)中的 O_8 和 O_9 所示。造成这种现象的原因是梳妆台和椅子之间存在少量噪声,CFDP 和 DPC 方法受噪声的影响较大,而本文方法不受少量噪声点的影响。

对于阁楼,CFDP、DPC 和本文方法分别提取了 6、10 和 11 个对象,包括沙发、椅子、楼梯等。CFDP

方法成功地提取了阁楼右侧的三个目标,如图 14(a)中的 O_3 、 O_4 、 O_5 所示。DPC 和本文方法也对这三个目标进行了提取。CFDP 方法根据聚类结果准确地提取到了楼梯,如图 14(a)中的 O_6 所示;DPC 和本文方法无法提取出整个楼梯,特别是 DPC 方法将楼梯分成了两部分,如图 14(b)中的 O_9 、 O_{10} 所示,而本文方法可将大部分楼梯提取出来,如图 14(c)中的 O_{11} 所示。CFDP 方法将沙发和沙发左侧的咖啡桌一并提取出来,如图 14(a)的中 O_2 所示;DPC 和本文方法都准确地提取出了沙发和沙发左侧的咖啡桌,且提取效果相同,如图 14(b) 中的 O_4 、 O_5 和图 14(c)中的 O_5 、 O_6 所示。DPC 和本文方法都未能将沙发上的抱枕和靠垫分开。CFDP 方法将厨台和三把转椅一并提取出来,如图 14(a)中 O_1 所示;而 DPC 方法较为准确地提取出了厨台,但提取出的厨台中包含了一把椅子,如图 14(b)中的 O_1 所示;本文方法可以准确提取出厨台,且不包含

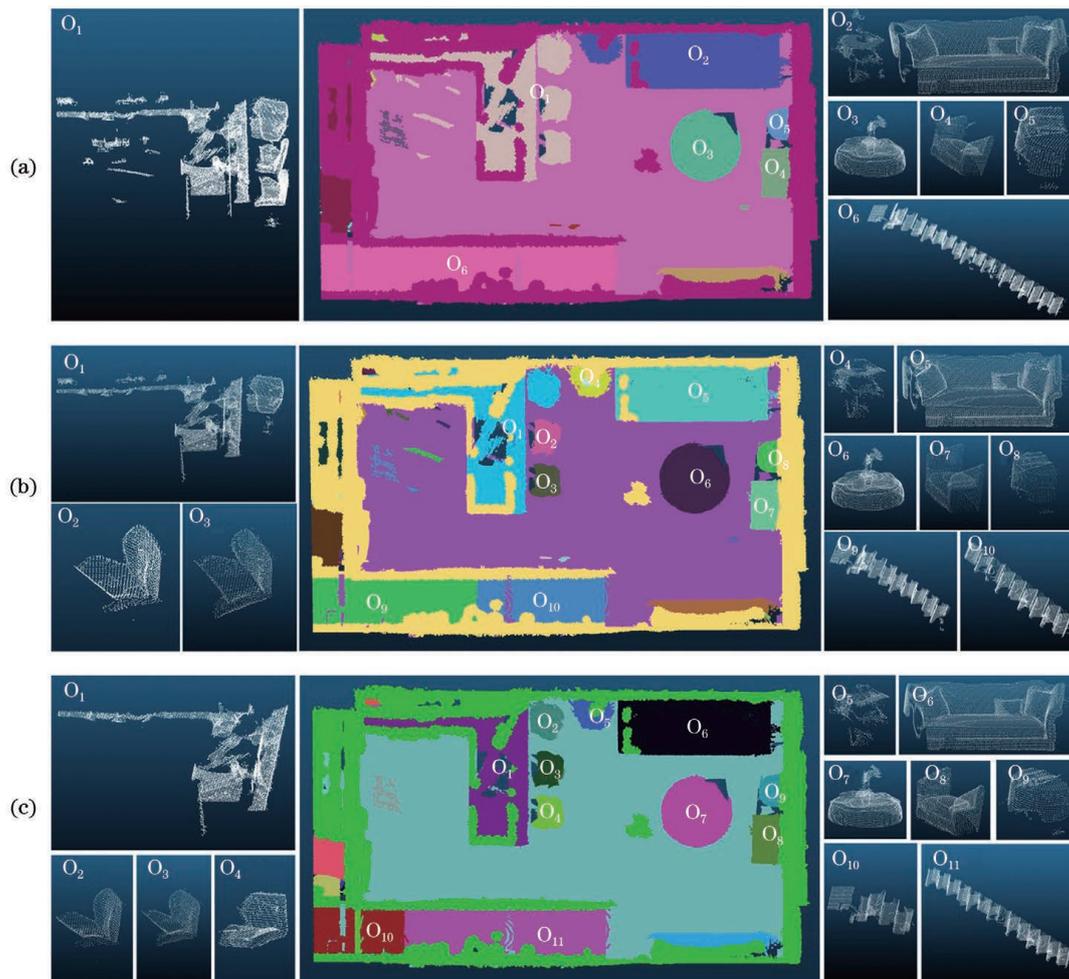


图 14 采用不同方法提取阁楼中的不同对象。(a) CFDP 方法;(b) DPC 方法;(c) 本文方法

Fig. 14 Different objects in the loft are extracted by different methods. (a) CFDP method; (b) DPC method; (c) proposed method

其他实际物体,如图 14(c)中的 O_1 所示。另外,这三种方法都有同样的缺点,即:如果一些物体在其他物体之上,则无法将它们分开。例如,床上的枕头与床合并在了一起,桌上的杯子与桌子合并在了一起,桌上的书与桌子合并在了一起,等等。

如图 15 所示,会议室和大厅中的相邻物体比公寓、卧室和阁楼中相邻物体靠得更近,尤其是会议桌和椅子。对于一些相邻但不紧挨的物体,这三种方法都可以进行分离并提取。同时,如图 15 中的黑色虚线六边形所示,这三种方法不仅可以将房间内的物体分开,也可以将天花板上的灯分

开。同时,如图 15 中黄色虚线圆角矩形所示,这三种方法都无法提取出其他物体(桌子、沙发等)上的一些物体,而是将这些物体聚在了一起。本文方法成功分离了两个相邻物体,但 CFDP 和 DPC 方法没有达到这种提取效果,如图 15(h)中的黄色虚线矩形所示。CFDP 和 DPC 方法完全不能提取会议桌附近的转椅,如图 15(a)中的黑色虚线矩形所示,而本文方法可以提取出会议桌附近的部分转椅,这部分转椅与会议桌的颜色有所不同。出现这一现象的原因是会议桌和旋椅彼此之间靠得很近或有一些噪声点。

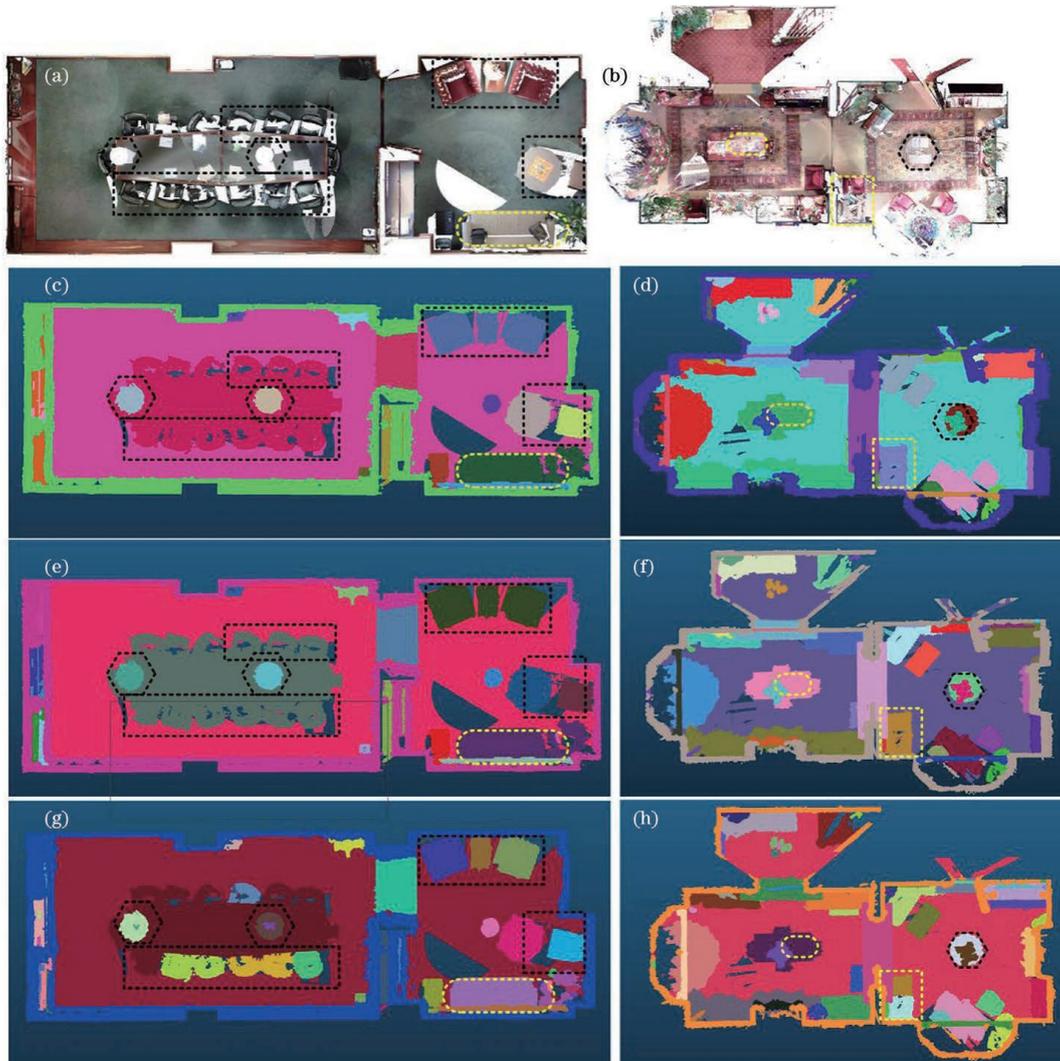


图 15 CFDP、DPC 和本文方法提取并以不同颜色显示的不同对象。(a)(b)会议室和大厅;(c)(d) CFDP 聚类结果;
(e)(f) DPC 聚类结果;(g)(h)本文方法的聚类结果

Fig. 15 Different objects extracted by CFDP, DPC, and proposed method, and displayed with different colors.

(a)(b) Boardroom and lobby; (c)(d) CFDP clustering results; (e)(f) DPC clustering results; (g)(h) clustering results of proposed method

由以上分析可知,本文方法提取出了室内大部分邻近物体的点云,而且可以分离出一些相互靠近

的物体,但 DPC 和 CFDP 方法无法实现这样的提取效果。对于某些相邻物体,本文方法的提取效果与

DPC 方法相似,但如果相邻物体之间存在少量噪声点时,本文方法优于 DPC 方法。总体来看,本文方法的提取性能优于 DPC 和 CFDP 方法。

4 性能分析

为了评价本文方法的性能,将本文方法提取的物

体与真实情况进行了比较。从两个方面对本文方法的性能进行评价:一是可视化的性能评价,二是定量的性能评价。采用三种类型的室内点云对提取结果进行评价。第一种类型是房间内的所有物体几乎都是独立的,第二种类型是只有少数物体靠在一起,最后一种类型是许多物体相互靠在一起,如图 16 所示。

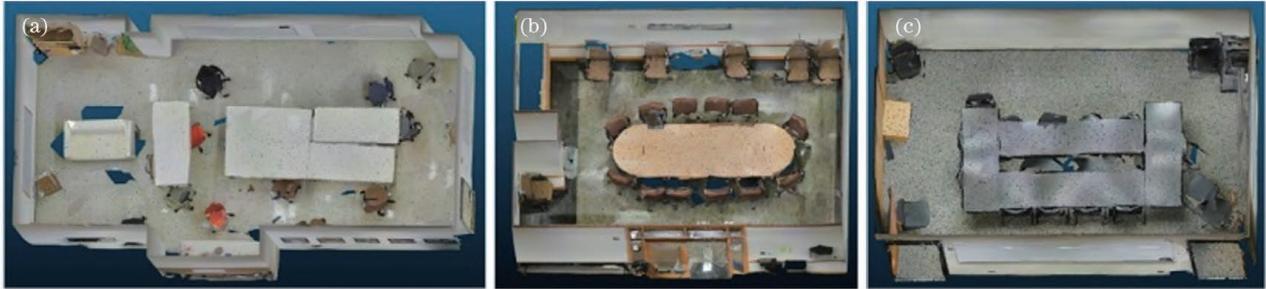


图 16 室内点云的三种类型。(a)房间中几乎所有的物体都是独立的,不相互靠近;(b)房间中只有几个物体相互靠近;(c)房间中有多个物体相互靠近

Fig. 16 Three types of indoor point clouds. (a) Almost all the objects in the room are independent and not close against together; (b) a few objects in the room close against together; (c) many objects in the room close against together

图 16 显示了三种类型的会议室。图 16 (a) 所示为长方形会议室,会议室中有木板、书柜、椅子、杂物、沙发和桌子,从图中可以很清楚地看到几乎所有物体都是分开的。图 16 (b) 为另一间会议室,内有书柜、椅子、杂物和桌子,房间内只有几件东西是相互靠近的,比如相邻两把椅子靠在一起。图 16 (c) 所示的会议室内主要包括椅子、杂物、沙发和桌子,桌子周围的椅子几乎都靠在一起,所有的桌子彼此紧靠在一起。

4.1 可视化的性能评估

首先,根据密度聚类提取墙壁、天花板、地板的点云,如图 17 所示。可以清楚地看到,通过密度聚类得到的天花板和地板几乎与实际相同,但墙上的一些板子没有被提取出来,如图 17 中的虚线矩形所示。图 17(a)、(c) 中的虚线椭圆显示了提取的门,但是如图 17(d)、(f) 中的椭圆所示,本文方法将门和墙合并在一起。导致这一现象的原因是门上点的密度与墙几乎相同。

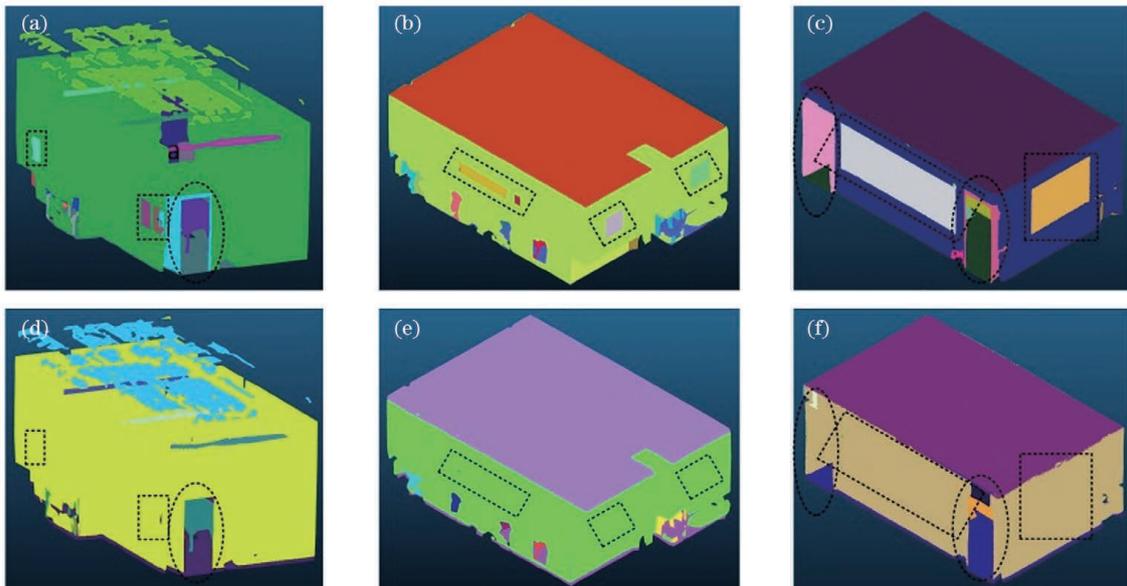


图 17 墙壁、天花板、地板的提取结果。(a)~(c)实际的墙壁、天花板、地板;(d)~(f)采用本文方法提取的墙壁、天花板、地板

Fig. 17 Extraction results of wall, ceiling, and floor. (a)~(c) Actual wall, ceiling, and floor; (d)~(f) extracted wall, ceiling, and floor by proposed method

其次,用不同的颜色显示本文方法提取的不同物体,如图 18 所示。在图 18(a)所示的房间场景中,几乎所有的物体都是独立的,彼此不靠近。比较图 18(a)、(b)可以得出,通过本文方法提取的不同对象几乎与真实物体一样。聚类方法很难将靠近墙壁的物体分离,所以它们常常聚集在一起,如图 18(b)中的黄色虚线矩形所示;同时,聚类算法也

无法提取桌面上的一些杂物,如图 18(b)中的虚线椭圆所示。对于第二种类型的房间,围着桌子的一对转椅和靠墙的一对转椅没有分开,主要是这些相邻的转椅相互之间靠得非常近。本文方法无法提取墙壁上的一些物体,如图 18(d)中的虚线椭圆所示。当房间中的某些物体靠得很近时,提取性能变差,如图 18(f)所示。

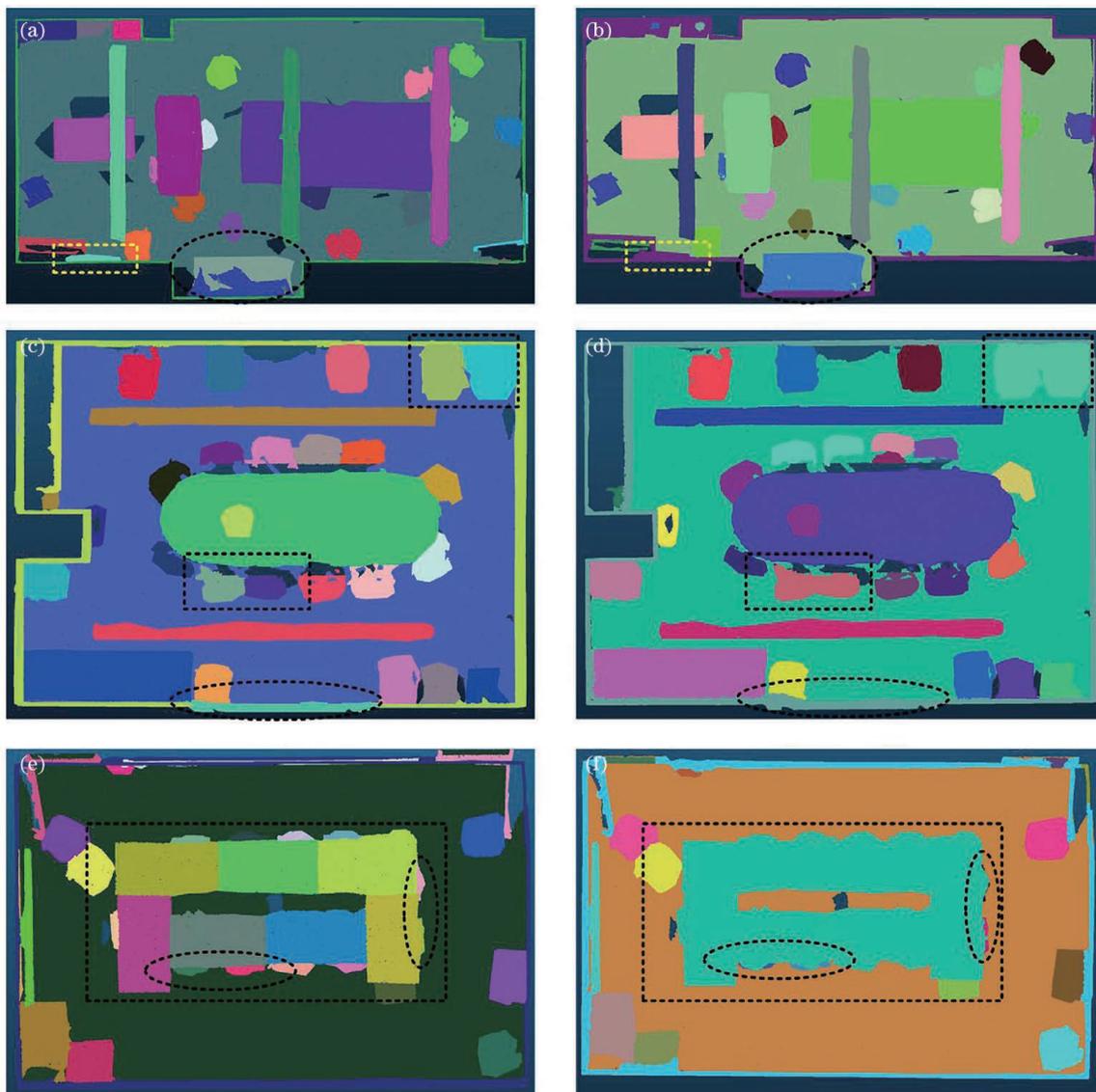


图 18 本文方法对三种类型室内点云的提取结果。(a)(c)(e)实际物体;(b)(d)(f)本文方法提取的不同物体

Fig. 18 Extracted three types of indoor point clouds by proposed method. (a)(c)(e) Actual objects; (b)(d)(f) different objects extracted by proposed method

图 18(e)中的虚线矩形包括 7 张桌子和 11 把转椅,经过聚类后只显示了 1 张桌子和 4 把转椅,特别是相邻的桌子合并在一起,椅子和桌子合并在一起,如图 18(f)所示。这表明,当物体非常靠近时,本文方法不能实现准确提取。但是,本文方法可以将几个靠近桌子的转椅分开,如图 18(f)中的虚线

椭圆所示。产生这种结果的原因是相邻物体之间并不是完全靠近。因此,本文方法的提取性能也会受到相邻物体靠近程度的影响。

4.2 性能定量评价

根据提取结果,得到了本文方法提取的三种类型房间中的对象数,如表 2 所示。

表 2 本文方法提取的三种类型房间的对象数量
Table 2 The number of objects extracted by proposed method in these three types of rooms

Type of room	The number of objects	
	Actual	Extracted
The first type of room	28	28
The second type of room	35	30
The third type of room	30	18

由表 2 可知,本文方法提取的第一类和第二类房间的对象数量最多,提取的第三类房间的数量最

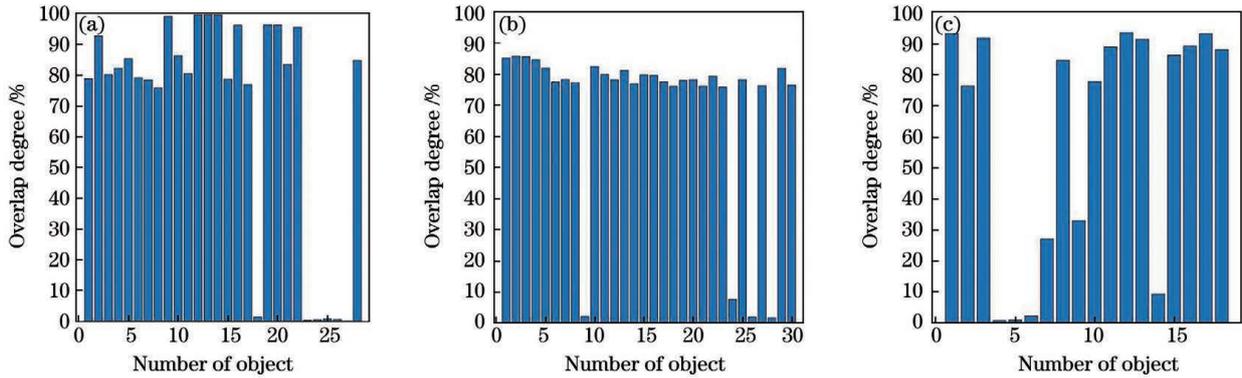


图 19 不同物体的重叠程度。(a)第一类房间;(b)第二类房间;(c)第三类房间

Fig. 19 Overlap degree of different objects. (a) The first type of room; (b) the second type of room; (c) the third type of room

从图 19 中可以清楚地看到,被提取目标的重叠程度大部分大于 75%,可以认为提取准确。准确提取的对象数量如表 3 所示。

表 3 本文方法提取的对象数量

Table 3 The number of objects extracted by proposed method

Type of room	Actual number of objects	The number of accurately extracted objects
The first type of room	28	22
The second type of room	35	26
The third type of room	30	12

对于第一类房间,本文方法准确提取了 22 个物体,提取率为 78.6%;同样,第二类和第三类房间的提取率分别为 74.3%和 40%。可见,在第一类和第二类房间中,大部分物体被提取出来,而在第三类房间中,仅有 40%的物体被提取出来。主要原因是第一类和第二类房间中几乎所有的物体都不相互靠近,但在第三类房间中,大多数物体都相互靠近。

少。使用重叠度(OD, f_{OD})来比较本文方法提取的目标与实际目标的差别,其计算公式为

$$f_{OD} = \frac{\sum_{i=1}^m \delta(r_i, s_i)}{m}, \quad (18)$$

式中: r_i, s_i 分别为实际对象点和提取的对象点; m 为每个实际对象点的数量; δ 为指标函数,其计算公式为

$$\delta(r_i, s_i) = \begin{cases} 1, & \text{if } r_i = s_i \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

根据式(18)得到了三种类型室内点云中每个提取对象的重叠程度,如图 19 所示。

为了全面准确地评估提取效果,本文采用精确度、召回率和 F1 分数这三种评价参数来评价目标的提取效果。 N_{TP} 表示准确提取的物体个数, N_{FP} 为不同对象聚在一起的个数, N_{FN} 为一个对象被分为两个或两个以上不同对象的个数。根据提取结果最终确定 N_{TP}, N_{FP}, N_{FN} ,如表 4 所示。

表 4 三种类型房间的 N_{TP}, N_{FP} 和 N_{FN}

Table 4 N_{TP}, N_{FP} , and N_{FN} of three types of rooms

Type of room	N_{TP}	N_{FP}	N_{FN}
The first type of room	22	2	4
The second type of room	26	5	4
The third type of room	12	13	5

利用 N_{TP}, N_{FP}, N_{FN} 确定精确度(P)、召回率(R)和 F1 分数(F_1),公式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (20)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (21)$$

$$F_1 = \frac{2P \cdot R}{P + R} \quad (22)$$

精确度越高,说明提取精度越高,而更高的召回率或 F1 分数则意味着更好的提取性能。根据 N_{TP} 、 N_{FP} 、 N_{FN} 计算出了不同类型房间的三种评价指标,如表 5 所示。

表 5 本文方法提取不同类型房间的精确度、召回率和 F1 分数

Table 5 Precision, recall rate, and F1-score of extraction of three types of rooms by proposed method

Type of room	Precision	Recall rate	F1-score
The first type of room	0.91	0.85	0.88
The second type of room	0.84	0.87	0.86
The third type of room	0.48	0.71	0.57

本文方法提取第一类房间的精确度略高于第二类房间,且对第一、第二类房间的提取精确度都比第三类房间高得多。本文方法提取第一类房间的召回率略低于第二类房间,但高于第三类房间,这是因为召回率只与 N_{TP} 和 N_{FN} 有关。本文方法提取第一类房间的 F1 分数与第二类房间的 F1 分数基本相

同,且均高于第三类房间。可见,本文方法对第一、第二类房间的提取质量优于对第三类房间的提取质量。造成这种结果的原因是在第三类房间里有许多物体彼此靠近。因此,本文方法适用于对第一类和第二类房间内目标的提取,但不适用于对第三类房间内目标的提取。本文方法的提取性能与相邻目标的邻近程度有关。

5 实例分析

本文以 MATLAB 和 CloudCompare 为应用平台进行实例分析,硬件配置为 Intel 酷睿 i7-5500U,内存 2.40 GHz,4 GB。利用本文方法对包括 10 间办公室、2 间休息室、3 个走廊、2 间盥洗室和 2 间储物室在内的大场景室内点云进行提取。10 间办公室的提取结果如图 20 所示。从图 20 中的不同颜色可以看出一些物体被清晰地提取出来,比如桌子、椅子等,但有些物体未被提取出来,而是与其他物体合并在一起。

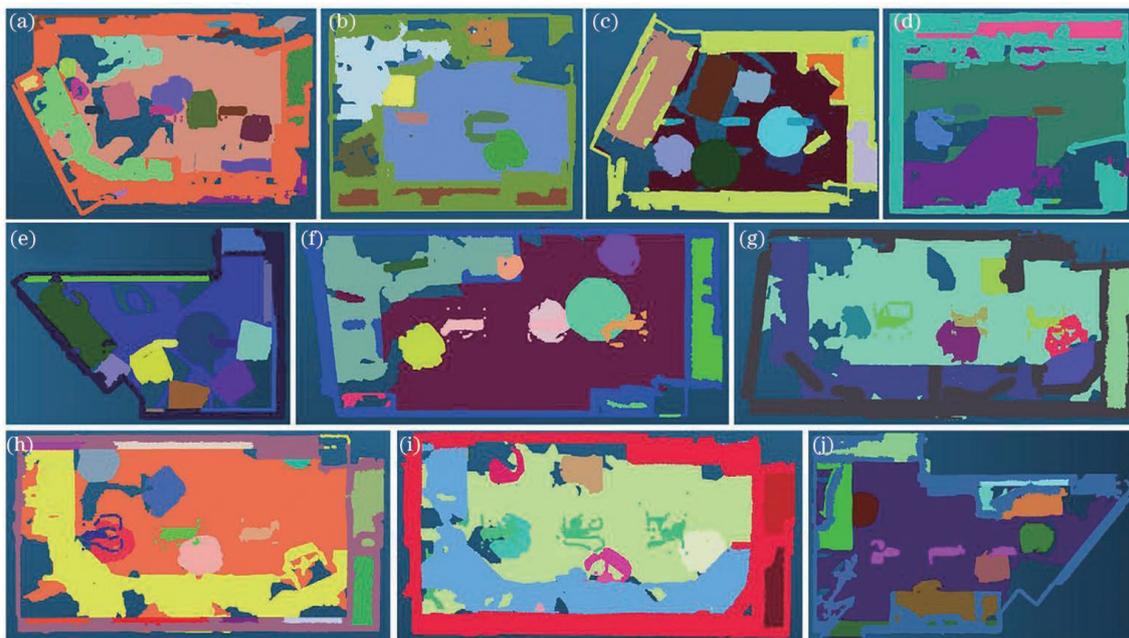


图 20 10 间办公室内不同对象的提取结果

Fig. 20 Extraction results of different objects in ten offices

采用本文方法提取了 2 间休息室中的不同物体,并用不同的颜色显示,如图 21 所示。每间休息室两侧都标记着不同的物品,包括书桌、椅子、沙发、花盆等。很明显,几乎所有的物体都被提取出来。其中,3 张小圆桌、2 把圆凳、2 张沙发和 1 张支架桌被准确提取出来,但是其他 4 把圆形座椅和另一张长桌合并在了一起。第 2 间休息室内几乎所有的物

体都被提取出来,并显示了不同的颜色,如图 21(b) 所示。其中,8 把椅子,3 张小圆桌和 1 个花盆被提取出来,同时获得了每个物体的完整部分,包括桌腿、桌面、靠背、椅底、花盆和花等。

图 22 为本文方法对 3 个走廊的提取结果。由图中的不同颜色可以看出,提取结果中只有墙壁和地板。这 3 个走廊都是空的,没有放置任何物品。

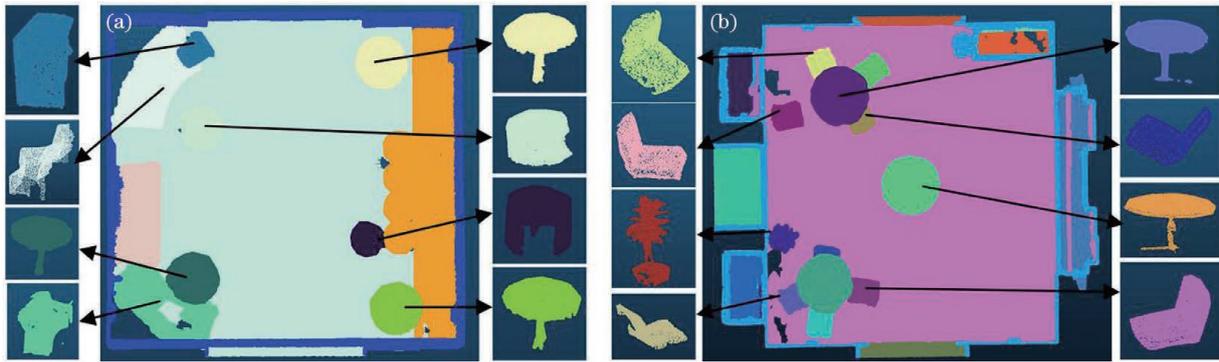


图 21 本文方法提取的 2 间休息室中的不同物体。(a) 第 1 间休息室的提取结果;(b) 第 2 间休息室的提取结果
 Fig. 21 Different objects in two lounges are extracted by proposed method. (a) Extraction result of the first lounge;
 (b) extraction result of the second lounge

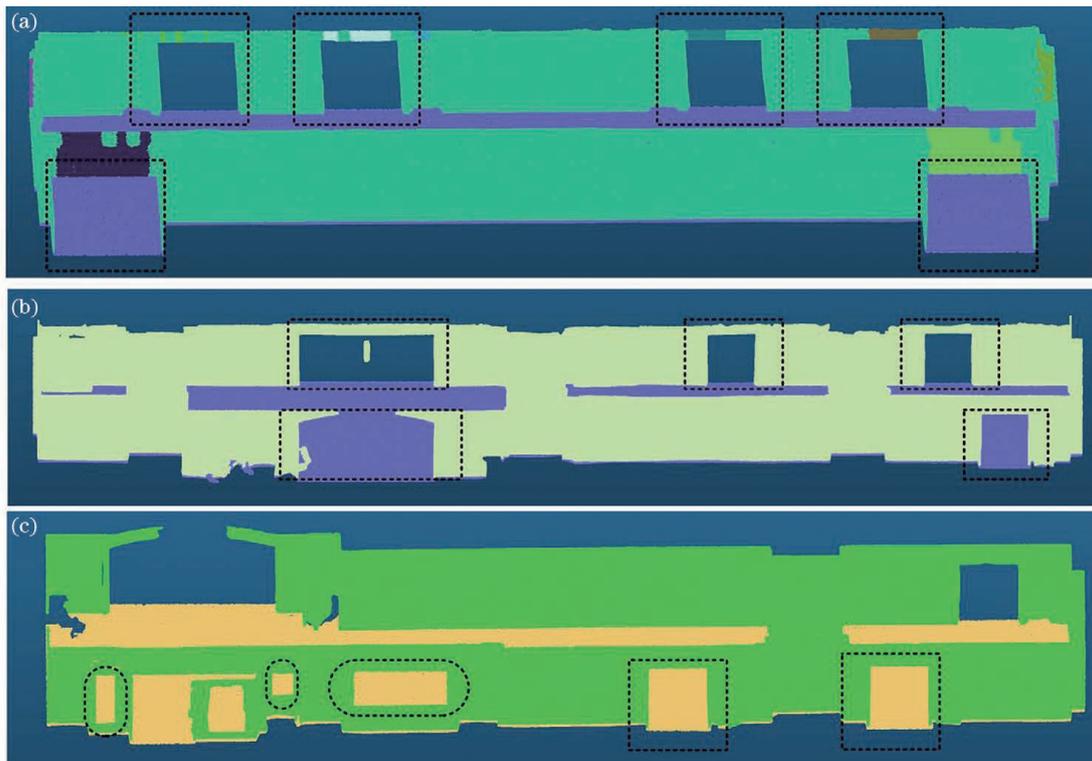


图 22 本文方法提取的 3 个走廊。(a) 第 1 个走廊;(b) 第 2 个走廊;(c) 第 3 个走廊
 Fig. 22 Extraction results of three hallways by proposed method. (a) The first hallway; (b) the second hallway;
 (c) the third hallway

图 22 中的虚线矩形显示了 3 个走廊的门,第 1 个、第 2 个和第 3 个走廊分别包括 6 扇门、5 扇门和 5 扇门,如图 22(c)中的虚线圆角矩形所示。

由于盥洗室和储物室的一些角落无法扫描,因此,点云中有漏洞,这些漏洞会对数据的聚类产生影响。图 23 所示为两间盥洗室和两间储物室的提取

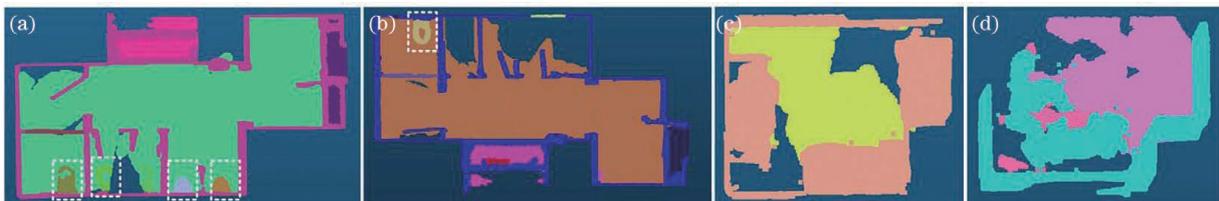


图 23 本文方法提取的两间盥洗室和两间储物室。(a) 盥洗室 1;(b) 盥洗室 2;(c) 储物室 1;(d) 储物室 2
 Fig. 23 Extraction results of two water closets and two storage rooms. (a) Water closet 1; (b) water closet 2;
 (c) storage room 1; (d) storage room 2

结果。从图 23 中的不同颜色可以看出,两间盥洗室和两间储物室的墙面和地板被准确地分离和提取。盥洗室内没有其他物体,马桶被准确提取出来,如图 23(b)中的虚线矩形所示。从两间储物室的聚类结果可以看出,由于储物室点云存在一些漏洞,因此两间储物室的提取结果会出现混乱,有些靠墙的物体无法被提取出来。

6 结 论

本文提出了基于指数函数的室内点云密度聚类模型。首先,根据点云的距离和角度分辨率构造截止距离函数模型。其次,分析点数、距离均值和距离标准差,用这些参数构建基于指数函数的局部密度模型。最后,根据点云与边界之间的距离得到判断墙体的约束距离密度。同理,根据 z 值的幅值分布和指数函数构造 z 值密度函数。结合局部密度,得到提取墙壁、天花板和地板的密度聚类模型。对于室内物体,根据截止距离内局部密度的大小确定约束距离。同时,根据约束距离与局部密度乘积的大小确定聚类中心。最后,根据每个点的聚类属性对室内目标进行聚类。根据上述密度聚类模型,可以提取出房间的墙壁、天花板、地板,以及房间内的物体。

将本文方法与其他聚类算法在不同室内场景下的聚类结果进行比较,结果表明,本文方法提取的目标数多于 CFDP 和 DPC 方法。同时,当相邻目标间存在少量噪声点时,本文方法的提取效果优于 CFDP 和 DPC 方法。此外,使用精确度、召回率和 F1 分数来评价本文方法的目标提取性能,本文方法的性能随着房间类型的不同而不同。提取结果表明,本文方法更适合用于提取物体不相邻类型房间内的目标,且本文方法的提取性能与相邻物体的靠近程度有关。

针对本文方法的缺点,未来拟将工作重点集中在提取相互靠近的物体上,同时实现物体上一些小物品的提取,如桌子上的书或杯子等小目标的提取。

参 考 文 献

- [1] 王祝,王智,张旭,等. 基于二维激光雷达的自适应阈值聚类分割方法[J]. 中国激光, 2021, 48(16): 1610005.
Wang Z, Wang Z, Zhang X, et al. Adaptive threshold clustering segmentation method based on two-dimensional lidar[J]. Chinese Journal of Lasers, 2021, 48(16): 1610005.
- [2] 李维刚,梅洋,樊响,等. 基于车载激光点云的铁路轨道检测[J]. 中国激光, 2022, 49(4): 0410002.
Li W G, Mei Y, Fan X, et al. Railway track detection based on vehicle laser point cloud [J]. Chinese Journal of Lasers, 2022, 49(4): 0410002.
- [3] Jain A K. Data clustering: 50 years beyond K-means [J]. Pattern Recognition Letters, 2010, 31(8): 651-666.
- [4] 郭佳,韩李涛,孙宪龙,等. 自动确定聚类中心的比较密度峰值聚类算法[J]. 计算机应用, 2021, 41(3): 738-744.
Guo J, Han L T, Sun X L, et al. Comparative density peaks clustering algorithm with automatic determination of clustering center [J]. Journal of Computer Applications, 2021, 41(3): 738-744.
- [5] Vinod H D. Integer programming and the theory of grouping [J]. Journal of the American Statistical Association, 1969, 64(326): 506-519.
- [6] 唐菓,邓兴升,王清阳. 基于密度聚类的点云滤波算法研究[J/OL]. 激光与光电子学进展: 1-17 [2021-11-03]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20210716.1534.012.html>.
Tang G, Deng X S, Wang Q Y. Research on point cloud filtering algorithm based on density clustering [J/OL]. Laser & Optoelectronics Progress: 1-17 [2021-11-03]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20210716.1534.012.html>.
- [7] 武森,汪玉枝,高晓楠. 基于近邻的不均衡数据聚类算法[J]. 工程科学学报, 2020, 42(9): 1209-1219.
Wu S, Wang Y Z, Gao X N. Clustering algorithm for imbalanced data based on nearest neighbor [J]. Chinese Journal of Engineering, 2020, 42(9): 1209-1219.
- [8] 高云龙,王志豪,潘金艳,等. 基于自适应松弛的鲁棒模糊 C 均值聚类算法[J]. 电子与信息学报, 2020, 42(7): 1774-1781.
Gao Y L, Wang Z H, Pan J Y, et al. Robust fuzzy C-means based on adaptive relaxation[J]. Journal of Electronics & Information Technology, 2020, 42(7): 1774-1781.
- [9] 黄亮,姚丙秀,陈朋弟,等. 高分辨率遥感影像超像素的模糊聚类分割法[J]. 测绘学报, 2020, 49(5): 589-597.
Huang L, Yao B X, Chen P D, et al. Superpixel segmentation method of high-resolution remote sensing image based on fuzzy clustering [J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(5): 589-597.
- [10] Chakraborty S, Paul D, Das S. Hierarchical clustering with optimal transport [J]. Statistics & Probability Letters, 2020, 163: 108781.

- [11] Lorbeer B, Kosareva A, Deva B, et al. Variations on the clustering algorithm BIRCH [J]. *Big Data Research*, 2018, 11: 44-53.
- [12] Naresh V R K, Gope D, Lipasti M H. The cure: cluster communication using registers [J]. *ACM Transactions on Embedded Computing Systems*, 2017, 16(5s): 1-19.
- [13] 黄亮, 姚丙秀, 陈朋弟, 等. 融合层次聚类的高分辨率遥感影像超像素分割方法 [J]. *红外与毫米波学报*, 2020, 39(2): 263-272.
Huang L, Yao B X, Chen P D, et al. Superpixel segmentation method of high resolution remote sensing images based on hierarchical clustering [J]. *Journal of Infrared and Millimeter Waves*, 2020, 39(2): 263-272.
- [14] Hou J, Zhang A H, Qi N M. Density peak clustering based on relative density relationship [J]. *Pattern Recognition*, 2020, 108: 107554.
- [15] 温晓芳, 杨志翀, 陈梅. 数据点的密度引力聚类新算法 [J]. *计算机科学与探索*, 2018, 12(12): 1996-2006.
Wen X F, Yang Z C, Chen M. Density attraction clustering algorithm between data points [J]. *Journal of Frontiers of Computer Science and Technology*, 2018, 12(12): 1996-2006.
- [16] Rodriguez A, Laio A. Clustering by fast search and find of density peaks [J]. *Science*, 2014, 344(6191): 1492-1496.
- [17] Liu R, Wang H, Yu X M. Shared-nearest-neighbor-based clustering by fast search and find of density peaks [J]. *Information Sciences*, 2018, 450: 200-226.
- [18] 杜洁, 马燕, 黄慧. 基于局部引力和距离的聚类算法 [J/OL]. *计算机应用*: 1-9 [2021-07-24]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210722.1504.008.html>.
Du J, Ma Y, Huang H. Clustering algorithm based on local gravity and distance [J/OL]. *Journal of Computer Applications*: 1-9 [2021-07-24]. <http://kns.cnki.net/kcms/detail/51.1307.TP.20210722.1504.008.html>.
- [19] Liu Y H, Ma Z M, Yu F. Adaptive density peak clustering based on K-nearest neighbors with aggregating strategy [J]. *Knowledge-Based Systems*, 2017, 133: 208-220.
- [20] Guo Z S, Huang T Y, Cai Z L, et al. A new local density for density peak clustering [M] // Phung D, Tseng V S, Webb G I, et al. *Advances in knowledge discovery and data mining. Lecture notes in computer science*. Cham: Springer, 2018, 10939: 426-438.
- [21] Chen X J, Yu K G. Feature line generation and regularization from point clouds [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(12): 9779-9790.
- [22] Işık Ş. Dominant point detection based on suboptimal feature selection methods [J]. *Expert Systems With Applications*, 2020, 161: 113741.
- [23] Jain A, Law M. *Data clustering: a user's dilemma* [M] // Pal S K, Bandyopadhyay S, Biswas S. *Pattern recognition and machine intelligence. Lecture notes in computer science*. Heidelberg: Springer, 2005, 3776: 1-10.

Extraction of Indoor Objects Based on Exponential Function Density Clustering Model

Chen Xijiang^{1,2,4}, Lin Jiaying^{2*}, Han Xianquan³, Wang Haojun²

¹ School of Artificial Intelligence, Wuchang University of Technology, Wuhan 430223, Hubei, China;

² School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan 430070, Hubei, China;

³ Yangtze River Scientific Research Institute, Wuhan 430010, Hubei, China;

⁴ Hubei Zhongtu Brands Company Limited, Wuhan 430070, Hubei, China

Abstract

Objective Indoor point clouds include walls, ceilings, floors, and many objects in a room. The extraction of a room's walls, ceilings, floor, and many objects is critical for many applications, including object identification for indoor navigation, facility management, and reconstruction of construction. Given this, this study proposes using an exponential function to construct a density clustering model according to the local density within a cutoff distance. First, the distance between the boundary and an indoor point cloud is used to construct the constraint condition of

wall density clustering. Then, a room's ceiling and floor are extracted according to the exponential function of the z value and local density model. Next, the local density model of different objects is constructed, and the constraint distance is determined according to the size of the local density. Simultaneously, cluster centers are recognized as points for which the product value of the local density and constraint distance is abnormally large. Finally, the cluster of each point is determined by the distance between the point and cluster center. Then, indoor object extraction is achieved by judging the distance between neighboring clusters. We conduct the extraction of indoor point clouds of different scenes and compare our method with the CFDP and DPC algorithms. Comparison results show that the DPC method is inferior to the proposed method but superior to the CFDP method. In addition, the extraction accuracy, recall, and F1-measure of different types of indoor point clouds are calculated in terms of the matching rate, true positive, false positive, and false negative. This study's findings show that the performance of the proposed method is affected by the degree of closeness between objects.

Methods For an indoor point cloud, the exponential function is used to construct the density cluster model of walls, ceilings, floors, and objects in a room. We can extract them based on their density clusters. Before the determination of the density cluster, the cutoff distance and local density within the cutoff distance are determined according to the relationship between angular resolutions and the scan distance. Concerning wall extraction, the comprehensive density function for wall extraction is obtained according to the local density within the cutoff distance and distance constraints. A wall's point cloud can be extracted according to the inflection point of the wall density. Concerning floor and ceil extraction, the exponential function is used to construct the density function of the z value, and then, the distributions of the density functions of ceilings and floors are obtained. According to the inflection point of ceiling and floor density, the ceiling and floor's point clouds are extracted. Concerning indoor object extraction, the prerequisite of indoor point cloud clustering is the determination of the cluster center. The cluster center can be determined by the constraint distance and local density. First, the constraint distance is determined according to the magnitude of the local density within the cutoff distance. Second, the clustering center is determined according to the product of the local density and constraint distance. Third, point cloud clustering is performed according to the distance between each point and clustering centers. Finally, we perform the extraction of indoor objects according to the aggregation of adjacent point cloud clusters.

Results and Discussions Table 1 shows that the extraction ratios of point clouds on the wall, ceiling, and ground are $\sim 95\%$. For the apartment, the number of wall, ceiling, and ground points is less than the actual number, indicating that there is an omission extraction phenomenon. For the bedroom, the ceiling and ground points extracted are less than the actual points, indicating that there is an omission extraction phenomenon. However, points extracted from the wall are larger than the actual points, so there is an over-extraction phenomenon. The main reason for this phenomenon is that parts of the point clouds from the ceiling and ground are considered point clouds from the wall. As presented in Table 3, for the first type of room, the proposed method accurately extracts 22 objects, and the extraction ratio is 78.6%. Similarly, the extraction ratios in the second and third types of rooms are 74.3% and 40%, respectively. Therefore, most of the objects in the first and second types of rooms are extracted, and no more than half of the objects in the third type of room are accurately extracted using the proposed method. The reason for this result is that almost all objects in the first and second types of rooms are not close to each other, but most of the objects in the third type of room are close to each other. Table 5 shows that the precision of the first type of room is slightly higher than that of the second type of room, but they are all significantly higher than that of the third type of room. The recall of the first type of room is slightly less than that of the second type of room but greater than that of the third type of room. The F1-score of the first type of room is almost the same as that of the second, and they are greater than that of the third type of room. The extraction quality of the first and second types of rooms is superior to that of the third type of room.

Conclusions This study presents a clustering model of indoor point cloud density based on an exponential function. First, the cutoff distance function model is developed according to the distance and angular resolution of point clouds. Second, the local density model based on the exponential function is constructed by analyzing the number of points and distance mean and standard deviation. Third, according to the distance between the point cloud and boundary, the constraint distance density of judging a wall is obtained. Similarly, the density function of the z value is constructed according to the amplitude distribution and the exponential function of the z value. Combined with the local density, the density clustering model of walls, ceilings, and floors is obtained. For indoor objects, the constraint distance is determined according to the local density within the cutoff distance. In addition, the clustering

center can be determined according to the product of the constraint distance and local density. Finally, indoor targets are clustered according to the clustering attribute of each point. Based on the density clustering model, walls, ceilings, floors, and objects in the room can be extracted. The proposed method is compared with other clustering algorithms in different indoor scenarios, and the results show that the number of objects extracted using the proposed method is greater than that extracted using the CFDP and DPC methods. In addition, when there are a few noise points between adjacent targets, the extraction effect of the proposed method is better than that of the CFDP and DPC methods. Furthermore, accuracy, recall, and F1-score are used to evaluate the object extraction performance of the proposed method, which varies with types of rooms. The results show that the proposed method is more suitable for rooms with non-adjacent objects, and its performance is related to the closeness of adjacent objects. Given the shortcomings of the proposed method, future research work will focus on the extraction of objects close to each other. In addition, a future clustering algorithm can accurately extract some small items on other objects, such as books or cups on a table.

Key words measurement; objects extraction; point cloud; density clustering; laser scanning