

基于超点图的点云实例分割方法

王志成, 余朝晖, 卫刚, 孙雨

(同济大学 CAD 研究中心, 上海 201804)

摘要: 提出一种基于超点图的点云实例分割(ISPG)方法。基于超点图结构提取点云对象相邻点之间的关联性特征,并且将传感器扫描的场景划分为均匀的几何元素,用来表示同属性的点云类,再由一个图卷积网络实现实例分割。结果表明:IoU 阈值为 0.5 的情况下,该方法在斯坦福大型三维(3D)室内空间数据集 S3DIS 上精度达到了 48.9%。

关键词: 点云;超点图;实例分割;图卷积网络
中图分类号: TP391.4

文献标志码: A

Point Cloud Instance Segmentation Method Based on Superpoint Graph

WANG Zhicheng, YU Zhaohui, WEI Gang, SUN Yu
(CAD Research Center, Tongji University, Shanghai 201804, China)

Abstract: A point cloud instance segmentation method based on superpoint graph (ISPG) is proposed in this paper. The correlation between the adjacent points of point cloud objects is extracted based on superpoint graph structure. Then, the scene scanned by a sensor is divided into uniform geometric elements to represent the point cloud classes with same attributes. Finally, a graph convolution network is used to implement instance segmentation. It is shown that the method has an accuracy of 48.9% on Stanford's large 3D indoor spatial dataset S3DIS when the IoU threshold is 0.5.

Key words: point cloud; superpoint graph; instance segmentation; graph convolutional neural network

点云是一种重要的三维数据表示形式,数据保存时每一个点由坐标加上该位置上的值构成。因为点云的存储形式具有不规则、无序性以及数据抽象和难以可视化等特点,因此点云处理面临着巨大的挑

战。传统方法在处理点云数据时存在着精度低、算法泛化能力差等问题。随着深度学习在二维图像任务如识别、检测以及分割上取得巨大突破,学者们开始探索将神经网络运用到点云处理上并取得重大的进展。

将深度学习算法应用到点云处理上的最直接方法是先将数据转换为体素化模型^[1],也就是三维结构化数据,然后直接用三维滤波器来训练一个卷积神经网络,如应用于自动驾驶的全卷积网络模型^[2]。然而,该方法使体积数据变得庞大,训练数据以及卷积核由二维变成三维,这意味着需要大量的计算,现阶段的计算能力难以达到要求。还有一些方法采用多视图数据,比较经典的就是基于多视角的三维目标检测网络^[3]。Charles 等^[4]提出了第一个直接处理点云数据的神经网络算法 PointNet,将点云数据投影到二维平面。该方法不直接处理三维的点云数据,而是先将点云投影到某些特定视角。许多研究对 PointNet 特征提取方式进行改进,如 PointCNN^[5],通过一种自定义的 X-Conv 操作,使常规的 Convolution 也能处理点云。Jiang 等^[6]提出了 PointSIFT 模型,对于每一个点云能端对端地输出一个表征向量,该向量编码了各个方向的信息,同时自适应地选择合适的表征尺度。目前,对于深度学习在点云数据实例分割中的运用已经有了进展,相似度分组网络(SGPN)^[7]模型是第一个直接学习三维点云特征的实例分割网络框架。使用 PointNet 提取特征,通过一个单独网络去预测点组推荐并为各个推荐组分配相应标签;采用深度相似矩阵表征特征空间中每对点之间的相似性,进而产生点分组建议。

现阶段点云实例分割并不能达到理想的效果,各种算法还有待提升。提出了一种基于超点图的点云实例分割(ISPG)方法,旨在提高实例分割任务

收稿日期: 2020-04-11

基金项目: 上海市科委项目(16111107602)

第一作者: 王志成(1975—),男,副研究员,工学博士,主要研究方向为计算机视觉、模式识别。

E-mail: zhichengwang@tongji.edu.cn

通信作者: 余朝晖(1994—),男,硕士生,主要研究方向为目标检测与识别。E-mail: yuzhaohui94@163.com



论文
拓展
介绍

的精度和效率,并且在实验中验证方法的有效性。

1 基于超点图的点云实例分割方法

为了处理无序的、不规则的点云数据,提出将大场景三维点云用一系列相互联系的简单形状构成超点的表达方式,这种表达方式与影像分割中的超像素类似,由一种称为超点图(SPG)^[8]的属性有向图结构描述。超点图的节点表示简单的结构,而富有边特征的边描述了它们的连接关系。

超点图的表达方式有以下几个显著优点:

(1)超点图将目标的分割部分作为一个整体,而不是对单个点或者体素进行分类,部分点的整体一超点具备比单个点或者体素更加显著的特征,有利于后续的实例分割。

(2)超点图可以详细描述相邻目标之间的位置关系,这对于上下文分类至关重要,如车辆通常是在道路上、天花板由墙面包围等。

(3)超点图的大小由场景中简单结构的数量决定,相比于处理之前的点集或者体素,超点图在数据量上要小得多。因此,可以对长距离的交互进行建模,而这在没有强假设的成对连接中是很难实现的。

1.1 超点划分

算法将点云划分成几何上简单且有意义的形状,称之为超点。算法的目标不是检索单个对象,如汽车或椅子,而是将对象分解为简单的几何部分。对于聚类的结果,不仅要求在几何上是简单的,还期望在语义上是同质的,也就是每个几何形状不覆盖不同类的对象。在超点图的网络中,将整个点云作为算法的输入,从点云分区可以很容易地计算出超点图。为了提高效率,这里使用了全局能量模型。

输入点云 C 是 n 个点的集合,每个点由包含 (x, y, z) 的位置信息和其他观察值定义,如颜色和强度等。对于每个点,将一组几何特征 ρ 作为局部特征。在本研究中,使用了文献[9]中提出的三个特征描述符,包括线性、平面性和反射强度,以及文献[8]中引入的垂直度特征。特征描述符是由邻域定义的,对于每个点的邻域,一组点的坐标的协方差矩阵的特征值分别为 $\lambda_1, \lambda_2, \lambda_3$, 规定 $\lambda_1 \geq \lambda_2 \geq \lambda_3$ 。特征描述符的计算式如下所示:

$$\rho_1 = \frac{\lambda_1 - \lambda_2}{\lambda_1}, \rho_p = \frac{\lambda_2 - \lambda_3}{\lambda_1}, \rho_s = \frac{\lambda_3}{\lambda_1} \quad (1)$$

式中: ρ_1 表示线性度; ρ_p 表示平面度; ρ_s 表示散射值。线性度描述了邻域连续程度,平面度评估了邻域点对于平面的拟合程度,散射值反映了各向同性和曲度。在实验中,垂直范围被证明对区分道路和墙壁,以及具有相似尺寸的电线和电线杆至关重要。因此,还使用了垂直范围分量 η 这个描述子。设 u_1, u_2, u_3 是分别与 $\lambda_1, \lambda_2, \lambda_3$ 对应的三个特征量。将垂直范围分量定义为特征向量各维度绝对值的加权之和的近似,权重为对应的特征向量的值,计算式如下所示:

$$\eta = \sum_{i=1}^3 \sum_{k=1}^3 \lambda_k |u_{ki}|, \text{约束 } \|u_k\| = 1 \quad (2)$$

式中: u_k 表示单位化的特征向量。点云的空间结构可以由无向图 $G_u = (V, E)$ 表示,其中节点 V 代表点云的点,边 E 代表节点间的邻接关系。全局能量模型^[9]由距离每个点最近的 n 个点定义,使用无向图 $G_u = (V, E_m)$, E_m 表示最近的 n 个点的平均距离,算法的 n 取 10。该几何分区的问题被定义为以下能量泛函的极值问题:

$$\arg \min_{g \in \mathbb{R}^{V \times 4}} \sum_{j \in V} \|g_j - f_j\|^2 + \mu \sum_{a, b \in E} \sigma(g_a - g_b) \quad (3)$$

式中: g 表示拟合既定特征; f 表示点的几何特征,包含四个维度 $(\rho_1, \rho_p, \rho_s, \eta)$; μ 表示正则化系数,作用是保证分割的形状足够简单; σ 为激活函数。该方法不仅可以达到有效逼近 10-cut 算法的效果,还具有自适应的效果。在几何形状变得复杂时,超点中包含的点的个数将下降,反之将上升。

1.2 超点图定义

在第 1.1 节中得到超点后,使用有向图 $G_o = (S, E, \omega)$ 表示整个点云,并命名为超点图。 S 是所有超点的集合, $S = \{S_1, S_2, S_3, \dots\}$, 其中 S_i 表示一个超点的点集; E 表示超点之间的邻接关系,命名为超边, $E = \{E_1, E_2, E_3, \dots\}$, 其中 E_i 表示超边; ω 表示编码超点图边界的权重。 E 表示无向图中所有边的集合,如果图上有一条边将超点连接在一起,则将两个超点定义为相邻,表达式如下所示:

$$\begin{cases} E = \{(s, t) \in S^2 \mid \exists (i, j) \in E \cap (s, t)\} \\ \omega_{s,t} = |E \cap (s, t)|, \forall (s, t) \in S^2 \end{cases} \quad (4)$$

式中: $\omega_{s,t}$ 表示 (s, t) 连接的个数。

超边示意图如图 1 所示。点之间的边用实线表示,部分超点分割之间的邻接用虚线表示。

超边还可以通过比较相邻超点的形状和大小来定义特征。用 $|S|$ 表示超点 S 包含的点数。此外,还

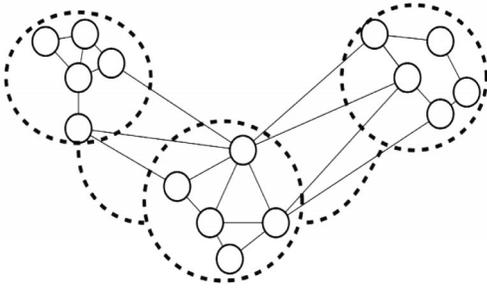


图 1 超边

Fig.1 Superedge

定义了长度 $L = \lambda_1$ 、表面积 $A = \lambda_1 \lambda_2$ 和体积 $v = \lambda_1 \lambda_2 \lambda_3$ 等形状特征。 $\lambda_1, \lambda_2, \lambda_3$ 是每个超点中点的坐标的协方差矩阵的特征值,按递减顺序排序。表 1 描述了超边特征, p 表示一个超点。

表 1 超边描述子及定义

Tab.1 Superedge descriptors and definitions

特征名称	描述子
平均偏移	\bar{p}
标准差偏移	δ_p
质心偏移	$\bar{p}_i - \bar{p}_j$
长度比	$\log L(s) / \log L(t)$
表面积比	$\log A(s) / \log A(t)$
体积比	$\log v(s) / \log v(t)$
点数比	$\log s / \log t $

1.3 超点嵌入

将超点转化为固定维度的向量表示。每个超点的转化都是独立进行的,超点所需要的上下文信息从第 1.4 节中获取。

选取 PointNet 作为超点嵌入算法模型。PointNet 网络简单,并且具有很好的鲁棒性和较高的效率,在实验中表现出不俗的性能,具体流程如图 2 所示。在 PointNet 中,输入点首先由空间变换网络 (STN)^[10] 实现对齐,然后由多层感知器 (MLP) 独立处理,最后用最大池化提取到全局特征。图 2 中, n_p 表示每次输入的节点数, d_p 表示特征维度, d_s 表示输出的特征维度。

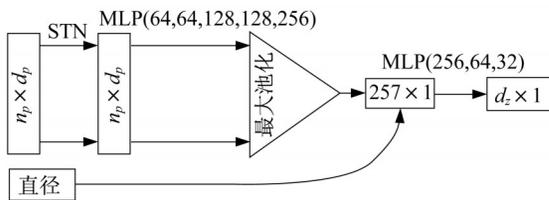


图 2 超点嵌入

Fig.2 Superpoint embedding

在超点的处理中,使用更为简洁的 PointNet 完成嵌入,一定程度上节省了在门控循环单元 (GPU)

上评估时的内存占用。具体做法:首先对超点进行采样,训练时点的个数小于 40 的超点将影响训练效果,因此对此类超点的分类将依据上下文信息进行判断。

为了使 PointNet 学习不同形状的空间分布,在处理之前需要将每个超点都重新缩放到单位球面范围内。因此,每个点由归一化后的位置信息 p'_i 、观测信息 o_j 以及几何特征 f_j 描述。此外,原始矩阵的缩放尺度将作为一个额外特征连接到最大池化后得到全局信息。

1.4 上下文语义中的分割

Pipeline 的最后一步是根据嵌入和超点图中的局部特征对超点进行分割。基于之前建立的超点图模型,采用图卷积来做最后的处理。

根据超边的信息改善超点的嵌入。具体来说,每个超点 s_i 都在 GRU 中保留其隐藏状态。隐藏状态初始化为嵌入 $x_i^{(0)}$,之后保留的是每轮迭代的结果。在迭代中,一个 GRU 单元 (见图 3) 将其隐藏状态 $h_i^{(t)}$ 和传入信息 $x_i^{(t)}$ 作为输入,并计算一个新的隐藏状态 $h_i^{(t+1)}$ 。对超点 s_i ,传入信息 $x_i^{(t)}$ 是和超点 s_i 邻接的超点的隐藏状态 $h_j^{(t)}$ 的加权和。

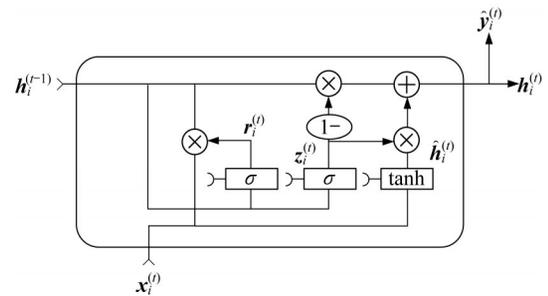


图 3 GRU 模块

Fig.3 GRU module

GRU 由一系列线性以及非线性计算组成,GRU 的前向传播过程如下所示:

$$h_i^{(t)} = (1 - z_i^{(t)}) \odot h_i^{(t-1)} + f_i^{(t)} \odot \hat{h}_i^{(t)} \quad (5)$$

$$\hat{h}_i^{(t)} = \tanh(W_h x_i^{(t)} + U_h (f_i^{(t)} \odot h_i^{(t-1)} + b_h)) \quad (6)$$

$$z_i^{(t)} = \sigma(W_z x_i^{(t)} + U_z h_i^{(t)} + b_z) \quad (7)$$

$$r_i^{(t)} = \sigma(W_r x_i^{(t)} + U_r h_i^{(t)} + b_r) \quad (8)$$

式中: $x_i^{(t)}$ 为输入; $h_i^{(t)}$ 为输出; \odot 为矩阵点乘运算; $\hat{h}_i^{(t)}$ 为隐藏状态; $\sigma(\cdot)$ 为 sigmoid 函数; $z_i^{(t)}$ 为重置门输出矩阵; $r_i^{(t)}$ 为更新门输出矩阵; W, U 和 b 为 GRU 中共享的可训练参数。式 (7) 和式 (8) 中出现了 GRU 的两个组件,即重置门 $z_i^{(t)}$ 和更新门 $r_i^{(t)}$ 。为了提高训练的稳定性,对式 (7) 和式 (8) 在线性变换后采用归一化操作,定义 $\rho(y) := (y - \bar{y}) / (\sigma(y) + \epsilon)$, ϵ 是一个常数。

GRU具有基于隐藏状态对输入进行加权的能力。例如,如果GRU的类别状态非常确定,GRU就会自动忽略上下文,或者只关注特定的特性通道。在输入之前乘上隐藏状态的权重,通过设置这些权重可以控制消息 $x_i^{(t)}$ 。

边缘条件卷积在模型中起着至关重要的作用,通过多层感知机处理连续序列的特征 F_{ji} 动态生成过滤权重。多层感知机对权重矩阵进行回归,这一操作需要对每个边向量进行矩阵向量乘法,带来大量的额外计算量,增加了运行时间、访存以及参数量。为了减少计算量,只对特定于边缘的权重向量进行回归,并进行逐元素乘法,然后与GRU更新规则结合使用,使得通道混合被推迟。通过简化,边缘非特定方式的性能已达到与边缘特定方式相似的水平。尽管感知机没有明确S型激活函数,但逐元素乘法可以选择学习来充当边缘条件门。受到DenseNet^[11]的启发,将所有迭代过程中的隐藏状态串联起来,并进行线性变换,以生成稠密特征 y_i 。多层感知机在时间迭代中是共享的,定义为 W_o ,表达式如下所示:

$$h_i^{(1)} = z_i, y_i = W_o(h_i^{(1)}, \dots, h_i^{(t+1)})^T \quad (9)$$

2 算法流程

ISPG模型如图4所示。算法分为以下三个步骤:

(1)按照几何形状均匀分区

将点云划分为几何上简单且有意义的形状,称之为超点。这一步骤完全是无监督的,分割出来的简单形状可以是大型的形状如道路或墙壁,也可以是更小的组件,如汽车或椅子的零件。

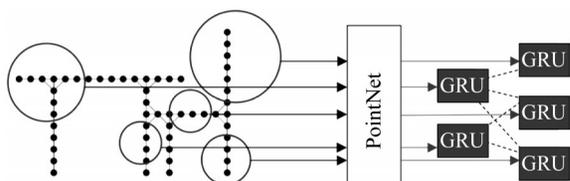


图4 ISPG模型
Fig.4 ISPG model

(2)超点的嵌入

超点图的每个节点对应整体点云的一小部分,也就是对应几何形状简单的基元。假设超点在语义上是齐次的,将这些超点基元降采样到最多数百个点来表示。小尺寸点云可使用PointNet方法处理。

(3)邻域点云的分割

超点图的数量级比建立在原始点云上的任何图

都小,基于图卷积的深度学习算法使用超点的边缘特征对这些节点进行分割,从而实现语义的分割。

3 实验

基于ScanNet^[12]数据集、斯坦福大型三维(3D)室内空间数据集(S3DIS^[11])以及KITTI数据集评估算法。此外,评估指标是基于原始点云计算的,而不是基于超点。使用tensorflow神经网络框架和单个Nvidia GTX1080 Ti GPU进行实验。

3.1 超点质量

从图5的可视化结果可以看出,全局能量模型可以很好地对稠密特征和稀疏特征的物体进行自适应分割。几何形状比较复杂的物体,如桌子、书架等(样本1和样本3),在较小的尺度内被分割成较多超点;墙、地板和天花板等物体(样本2),在特征空间内的距离很近,因此被分割成较为完整的超点。

3.2 对比实验设置

将分割结果与以下两种方法进行比较:

(1)簇分割(Seg-Cluster)模型

该模型在与SGPN^[12]模型相关的论文中首次提出,采取一种贪婪策略对点云进行简单分割。使用PointNet网络执行语义分割,然后选择所有点作为种子。从种子点开始,利用广度优先搜索具有相同标签的邻居。如果找到的群集超过200个点,就将其视为有效群组,然后采用分组合并算法合并有效群组。

(2)SGPN(PointNet)模型

SGPN模型中使用单个网络来预测点云的聚类区域及每个区域对应的语义类别。因此,该方法可以直接提取实例分割的结果。

3.3 ScanNet数据集测试结果

在ScanNet数据集上对18种常见对象类别进行评估和比较。原始点云作为输入,目标是分割场景中属于18个对象类别的每个实例及其语义标签。IoU(α_{IoU})阈值为0.5的平均精度(β_{AP})用作评估指标, α_{IoU} 和 β_{AP} 的计算式如下所示:

$$\alpha_{IoU} = \frac{N_{TP}}{N_{FP} + N_{TP} + N_{FN}}$$

$$\beta_{AP} = \frac{N_{TP}}{N_{TP} + N_{FP}}$$

式中: N_{TP} 、 N_{FP} 、 N_{FN} 分别表示预测正确、预测错误、没有预测到的物体数目。

如表2所示,ISPG模型的平均精度在大多数对

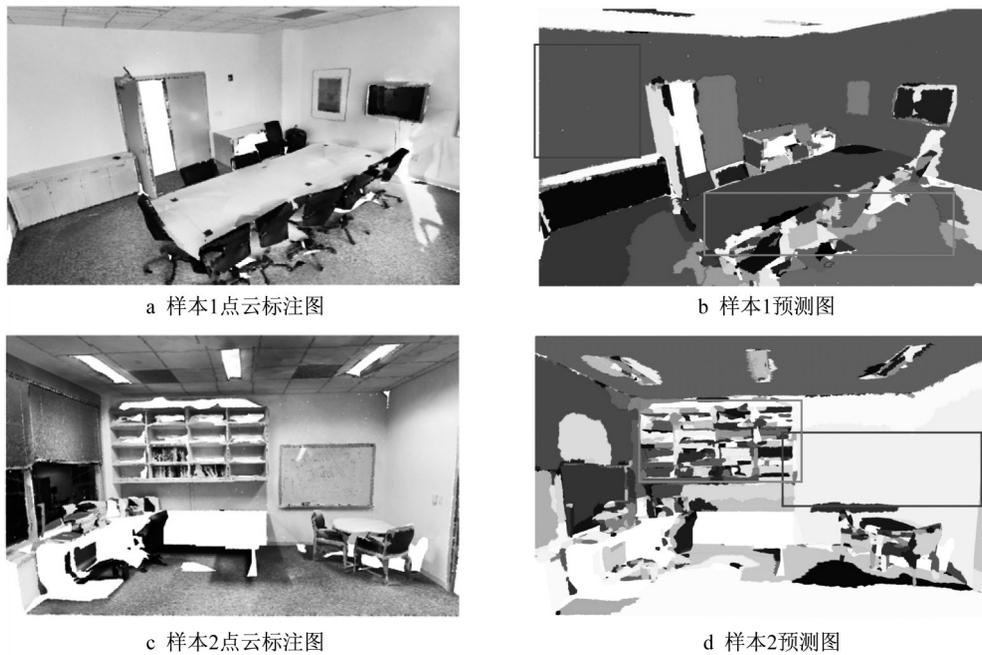


图5 全局能量模型生成超点

Fig.5 Superpoints produced by global energy model

象类别上的表现都超过了SGPN模型。由于ISPG模型只需要捕获相对简单的对象分布就可以生成非常高质量的候选群组,因此可以针对几何变化较小的类别获得很高的精度。对于在分割过程中需要稠密信息的类别(如门、窗),SGPN模型无法获得良好的评分,因为它们的相似性度量无法有效地被编码;ISPG模型在这些类别上获得了更好的结果,这表明在生成候选群组时产生的特征更有利于分类。

图6为SGPN模型与ISPG模型之间的定性比

较。SGPN模型需要在学习的相似性度量空间中划出清晰的边界来区分对象实例,可以观察到许多对象包括一个对象的部分或多个对象的预测。与SGPN模型相比,ISPG模型生成的对象具有更高的针对性,从而带来了更好的细分质量。当存在大量背景点时,SGPN模型很难学习良好的相似性度量标准,因为其仅专注于前景点的语义和相似性的学习,并且在训练过程中忽略了背景点,因此增加了基于背景的假阳性预测。

表2 ScanNet测试集IoU阈值为0.5的平均精度

Tab.2 Average accuracy of ScanNet dataset with 0.5 IoU threshold

模型	平均精度					
	橱柜	床	椅子	沙发	桌子	门
Seg-Cluster	4.70	0.20	0.20	10.70	2.00	3.10
SGPN(PointNet)	6.00	36.10	25.70	33.50	16.10	7.90
ISPG	34.80	40.50	58.90	39.60	27.50	28.30
模型	平均精度					
	窗	书架	画	柜台	书桌	窗帘
Seg-Cluster	0.40	0	18.40	0.10	0	2.00
SGPN(PointNet)	12.20	13.90	1.30	2.60	0	6.20
ISPG	24.50	31.10	2.80	5.40	12.60	6.80
模型	平均精度					
	冰箱	浴帘	马桶	盥洗盆	浴缸	其他
Seg-Cluster	6.50	0	10.90	1.40	33.30	2.10
SGPN(PointNet)	2.60	0	16.10	10.40	19.40	3.80
ISPG	21.90	21.40	82.10	33.10	50.00	29.00

3.4 斯坦福数据集测试结果

斯坦福大型三维室内空间数据集S3DIS^[1]包含从6个区域和13个对象扫描得到的270个室内场景。使用203个场景进行训练,其余67个场景进行

评估。对于斯坦福大型三维室内空间数据集(S3DIS),将每个房间分成面积为1 m×1 m的块,并且将每个点表示为九维(9D)矢量(XYZ,RGB和归一化的空间坐标)。在训练过程中为每个模块采样

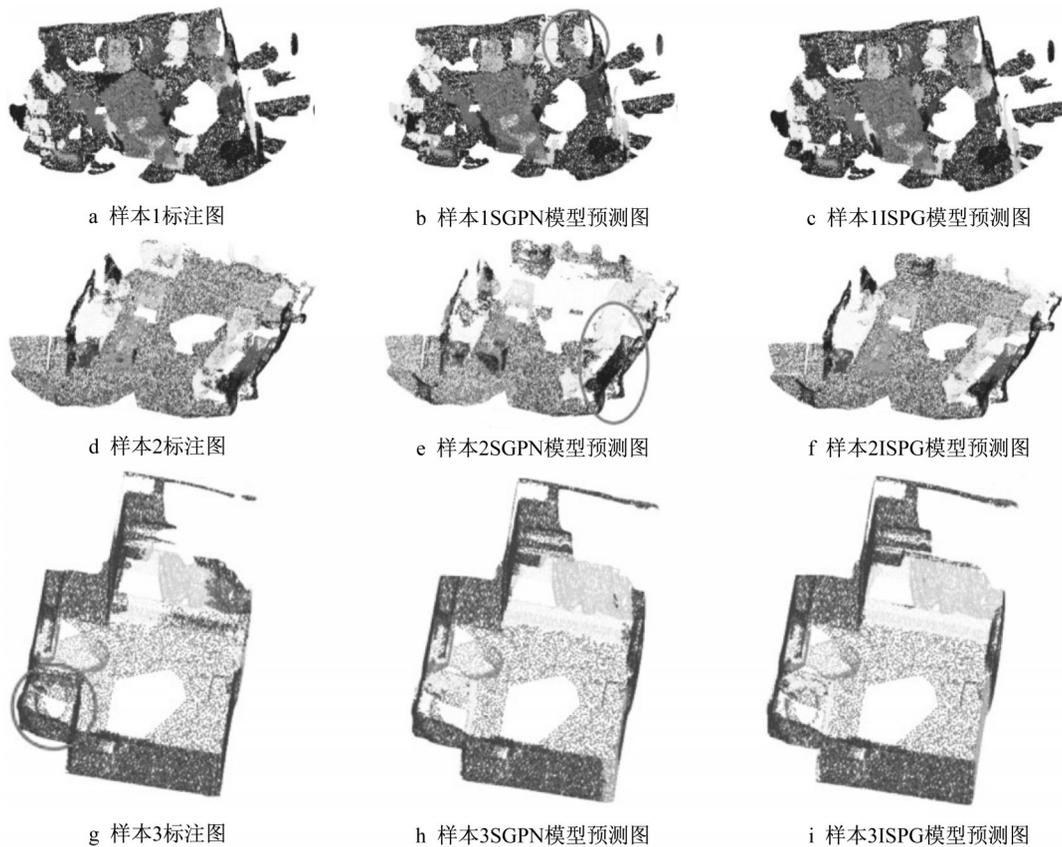


图6 ScanNet数据集上ISPG与SGPN模型可视化结果

Fig.6 Visualization results of ISPG and SGPN model on ScanNet dataset

了4 096个点,所有点都用于测试。

图7显示了使用ISPG模型在S3DIS上的实例分割结果。由于对象实例是无序的,因此同一组中点的颜色不一定与真实情况中对应点的颜色相同。为可视化实例类,还添加了语义分割结果。ISPG模型在各种房间类型上均表现出色。ISPG模型的测试结果与SGPN模型相比,很少出现边缘分割混乱情况,这说明ISPG模型的测试效果有了进一步的提升。

计算每个预测的和真实的标签之间的点的 α_{IoU} 。若 α_{IoU} 大于阈值,则将检测到的实例视为真阳性。进一步计算平均精度以进行实例分割性能评估,表3

为IoU阈值0.5时每个类别的平均精度。

3.5 KITTI数据集测试结果

KITTI数据集^[13]包含市区、乡村和高速公路等场景,每张图像中包含丰富的行为和车辆的信息以及一些遮挡。整个数据集由389对立体图像、光流图39.2 km视觉测距序列以及超过20万个三维标注物体的图像组成,以10 Hz的频率采样及同步。

对于该数据集,ISPG模型在小物体类别上获得了可观的收益,可以通过分区算法检测。此外,在执行空间规则化时,小物体通常与道路合并,可成为之后算法的一个改进方向。

KITTI数据集上的可视化效果如图8所示。可

表3 S3DIS测试集IoU阈值为0.5的平均精度

Tab.3 Average accuracy of S3DIS dataset with 0.5 IoU threshold

模型	平均精度					
	天花板	地板	墙	柱	梁	窗
Seg-Cluster	43.58	35.52	16.64	12.59	15.90	23.86
SGPN(PointNet)	79.44	66.29	88.77	77.98	60.71	66.62
ISPG	84.33	73.60	89.73	81.64	60.14	67.01
模型	平均精度					
	门	桌子	椅子	沙发	书架	木板
Seg-Cluster	15.75	22.63	10.33	3.92	43.33	10.71
SGPN(PointNet)	56.75	46.90	40.77	6.38	6.38	11.05
ISPG	32.80	42.30	42.97	50.10	60.73	78.83

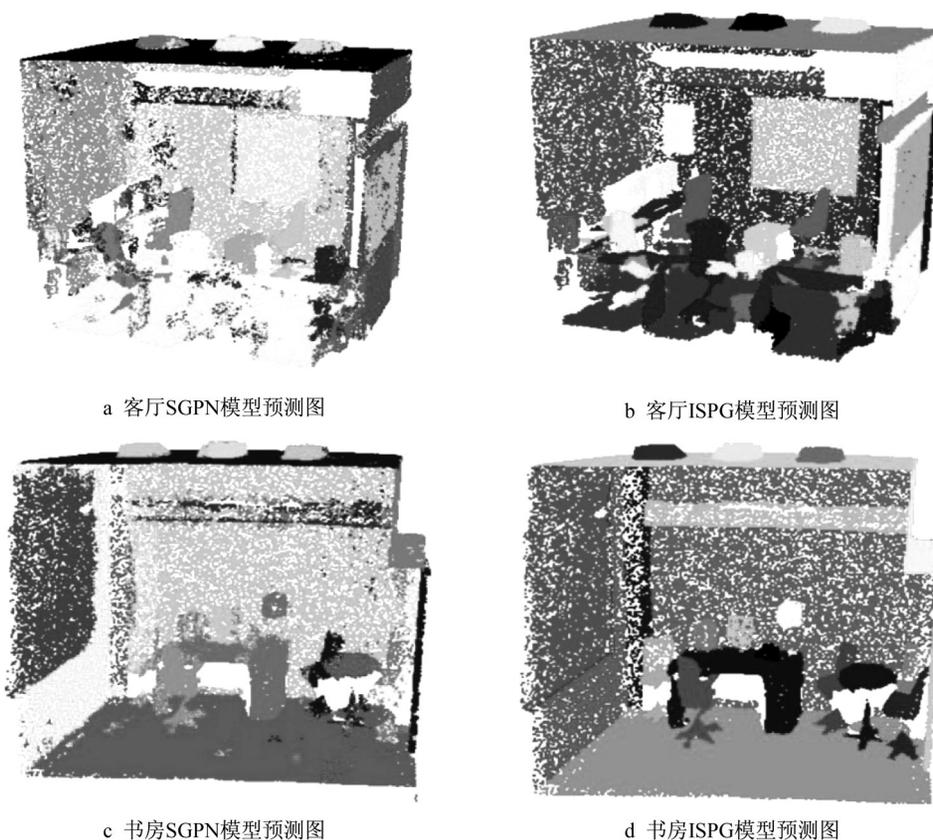


图7 S3DIS数据集上SGPN和ISPG模型可视化结果

Fig.7 Visualization results of SGPN and ISPG model on S3DIS dataset

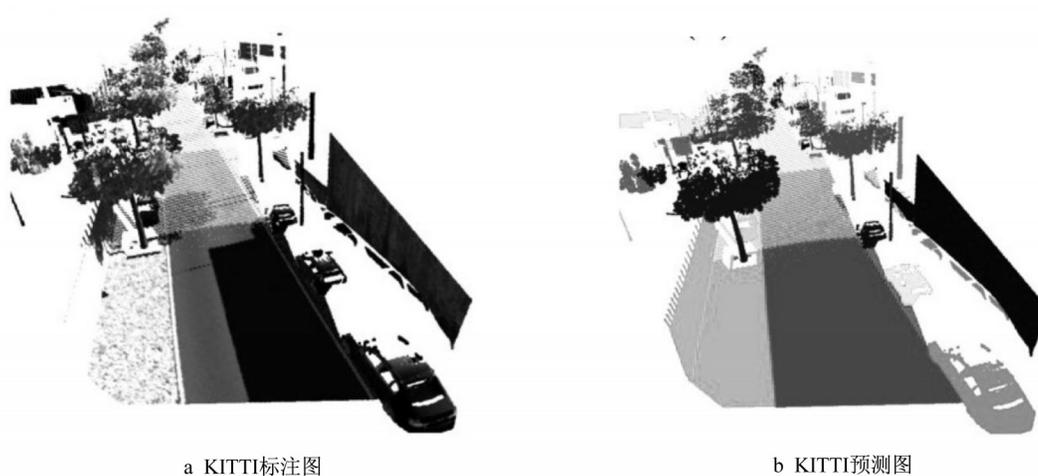


图8 KITTI数据集上ISPG模型可视化结果

Fig.8 Visualization results of ISPG model on KITTI dataset

可以看出,虽然KITTI数据集与前两个数据集差异巨大,但是ISPG模型依旧起到了很好的效果,说明模型具有很好的泛化能力。

4 结语

提出了基于超点图的点云实例分割模型。首先

将点云分割为超点,然后构建超点嵌入与边特征,最后通过上下文信息对超点进行分割。超点图的提出使得卷积运算不再运行在点上,而是运行在超点上,并通过丰富的边特征来表征上下文的信息,在保证分割效果的同时,加快了运行速度。在多个数据集上对该算法进行了全面的验证,取得了不错的实例分割效果。

参考文献:

- [1] ARMINI I, SENER O, ZAMIR A R, *et al.* 3D semantic parsing of large-scale indoor space [C]//2016 IEEE Conference on Computer and Pattern Recognition. Las Vegas: IEEE, 2016: 1534-1543.
- [2] RETHAGE D, WALD J, STURM J, *et al.* Fully-convolutional point networks for large-scale point clouds [C]// 2018 European Conference on Computer Vision. Munich: Springer, 2018: 625-640.
- [3] CHEN Xiaozhi, MA Huimin, WAN Ji, *et al.* Multi-view 3D object detection network for autonomous driving [C]// 2017 IEEE Conference on Computer and Pattern Recognition. Honolulu: IEEE, 2017: 6526-6534.
- [4] CHARLES R Q, HAO Su, KAICHUN Mo, *et al.* PointNet: deep learning on point sets for 3D classification and segmentation [C]// 2017 IEEE Conference on Computer and Pattern Recognition. Honolulu: IEEE, 2017: 77-85.
- [5] LI Yangyan, BU Rui, SUN Mingchao, *et al.* PointCNN: convolution on X-transformed points [C]// 32nd Conference on Neural Information Processing Systems. Montreal: [s.n.], 2018: 820-830.
- [6] JIANG Mingyang, WU Yiran, LU Cewu. PointSIFT: a SIFT-like network module for 3D point cloud semantic segmentation[EB/OL].[2019-06-20].<https://deepai.org/publication/pointsift-a-sift-like-network-module-for-3d-point-cloud-semantic-segmentation>.
- [7] WANG Weiyue, YU Ronald, HUANG Qiangui, *et al.* SGPN: similarity group proposal network for 3D point cloud instance segmentation [C]// 2018 IEEE Conference on Computer and Pattern Recognition. Salt Lake City: IEEE, 2017: 2569-2578.
- [8] GUINARD S, LANDRIEU L. Weakly supervised segmentation-aided classification of urban scenes from 3d LiDAR point clouds [C]// International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Hannover: [s.n.], 2017: 151-157.
- [9] DEMANTKE J, MALLET C, DAVID N, *et al.* Dimensionality based scale selection in 3D lidar point clouds [C]// International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Calgary: [s.n.], 2011: 97-102.
- [10] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks [C]//Advances in Neural Information Processing Systems. Montreal:[s.n.], 2015: 2017-2025.
- [11] HUANG Gao, LIU Zhuang, MAATEN V D, *et al.* Densely connected convolutional networks [C]// 2017 IEEE Conference on Computer and Pattern Recognition. Honolulu: IEEE, 2017: 2261-2269.
- [12] DAI A, CHANG A X, SAVVA M, *et al.* Scannet: richly-annotated 3d reconstructions of indoor scenes [C]// 2017 IEEE Conference on Computer and Pattern Recognition. Honolulu: IEEE, 2017: 2432-2443.
- [13] KIM Y. Convolutional neural networks for sentence classification [C]// 2014 Conference on Empirical Methods in Natural Language Processing. Doha: Association for Computational Linguistics, 2014: 1746-1751.

(上接第1376页)

- performance cycloid inspired wearable electromagnetic energy harvester for scavenging human motion energy [J]. Applied Energy, 2019, 256:113987.
- [9] WU N, JIANG H, LI W, *et al.* Output enhanced compact multilayer flexible nanogenerator for self-powered wireless remote system[J]. Journal of Materials Chemistry A, 2017, 5: 12787.
- [10] WANG B, ZHONG J, ZHONG Q, *et al.* Sandwiched composite fluorocarbon film for flexible electret generator [J]. Advanced Electronic Materials, 2016, 2(4):1500408-1.
- [11] SAKANE Y, SUZUKI Y, KASAGI N. Development of high-performance perfluorinated polymer electret for micro seismic power generation [C]//The Proceedings of the National Symposium on Power and Energy Systems. Sapporo: The Japan Society of Mechanical Engineers, 2008, 13:629-630.
- [12] MADINEI H, KHODAPARAST H H, ADHIKARI S, *et al.* Adaptive tuned piezoelectric MEMS vibration energy harvester using an electrostatic device[J]. The European Physical Journal Special Topics, 2015, 224(14/15):2703.
- [13] OXAAL J, FOSTER D, HELLA M, *et al.* Investigation of gap-closing interdigitated capacitors for electrostatic vibration energy harvesting [J]. Journal of Micromechanics and Microengineering, 2015, 25(10):105010.
- [14] ZHANG X, HILLENBRAND J, SESSLER G M, *et al.* Fluoroethylenpropylene ferroelectrets with patterned microstructure and high, thermally stable piezoelectricity [J]. Applied Physics A, 2012, 107(3):621.
- [15] ZHANG X, CAO G, SUN Z, *et al.* Fabrication of fluoropolymer piezoelectrets by using rigid template: structure and thermal stability[J]. Journal of Applied Physics, 2010, 108(6):064113.1.
- [16] ZHANG X, SESSLER G M, MA X, *et al.* Broad bandwidth vibration energy harvester based on thermally stable wavy fluorinated ethylene propylene electret films with negative charges [J]. Journal of Micromechanics and Microengineering, 2018, 28:065012.
- [17] ZHUKOV S, EDERGOY D, FEDOSOV S, *et al.* Analytical prediction of the piezoelectric d33 response of fluoropolymer arrays with tubular air channels[J]. Scientific Reports, 2018, 8(1):4597.
- [18] ZHANG X, WU L, SESSLER G M. Energy harvesting from vibration with cross-linked polypropylene piezoelectrets [J]. AIP Advances, 2015, 5(7):077185.
- [19] ZHANG X, SESSLER G M, WANG Y. Fluoroethylenpropylene ferroelectret films with cross-tunnel structure for piezoelectric transducers and micro energy