



(21) 申请号 201810861715.4

(22) 申请日 2018.08.01

(65) 同一申请的已公布的文献号

申请公布号 CN 109118564 A

(43) 申请公布日 2019.01.01

(73) 专利权人 山东佳音信息科技有限公司

地址 261035 山东省潍坊市奎文区新华路
1567号山东科技职业学院老图书办公
楼4楼402室

(72) 发明人 马燕新 鲁敏 涂兵 郭裕兰

雷印杰

(74) 专利代理机构 南京司南专利代理事务所

(普通合伙) 32431

专利代理师 何毅

(51) Int.Cl.

G06T 15/00 (2011.01)

G06N 3/0464 (2023.01)

(56) 对比文件

CN 106600622 A, 2017.04.26

CN 104700451 A, 2015.06.10

CN 107067470 A, 2017.08.18

CN 105005995 A, 2015.10.28

CN 108154192 A, 2018.06.12

LU Jun. "Point Cloud Registration
Algorithm Based on NDT with Variable Size
Voxel".《Proceedings of the 34th Chinese
Control Conference》.2015,全文.袁华. "基于体素化网格下采样的点云简化
算法研究".《数字视频》.2015,全文.

审查员 齐银凤

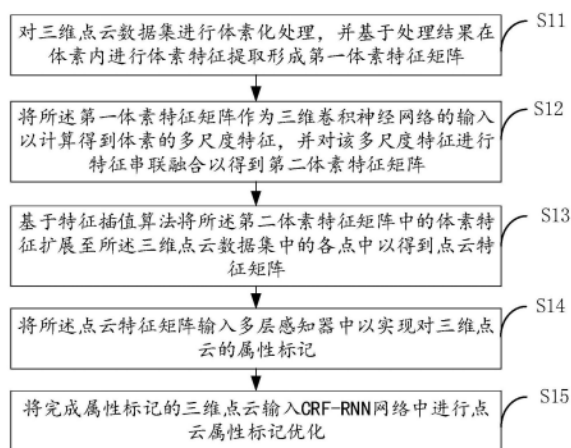
权利要求书2页 说明书15页 附图13页

(54) 发明名称

一种基于融合体素的三维点云标记方法和
装置

(57) 摘要

本发明实施例提供一种基于融合体素的三维点云标记方法和装置。其中,基于融合体素的三维点云标记方法包括对三维点云数据集进行体素化处理,并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵;将第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征,并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵;基于特征插值算法将第二体素特征矩阵中的体素特征扩展至三维点云数据集中的各点中以得到点云特征矩阵;将点云特征矩阵输入多层感知器中以实现三维点云的属性标记。本发明能够实现逐点的较为精细的分类识别,以进一步提升点云标记性能。



1. 一种基于融合体素的三维点云标记方法,其特征在于,所述方法包括:

对三维点云数据集进行体素化处理,并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵;

将所述第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征,并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵;

基于特征插值算法将所述第二体素特征矩阵中的体素特征扩展至所述三维点云数据集中的各点中以得到点云特征矩阵;

将所述点云特征矩阵输入多层感知器中以实现三维点云的属性标记;

基于处理结果在体素内进行点云特征提取形成第一体素特征矩阵的步骤包括:

针对每个体素中的点云计算该点云的中心坐标,并基于该中心坐标对所述体素中的点云数据进行中心归一化处理以得到初始数据矩阵;

将所述初始数据矩阵输入LGAB模块中以实现逐点的局部特征描述,并采用最大值池化对体素内的局部特征集合进行逐点池化操作以得到体素的全局特征并作为第一体素特征矩阵;

基于特征插值算法将所述第二体素特征矩阵中的体素特征扩展至所述三维点云数据集中的各点中的步骤包括:

针对所述三维点云数据集中的每个点分别作为目标点,在由所述第二体素特征矩阵形成的体素空间中搜索与该目标点最近的预设数量的体素作为邻域体素;

对所述邻域体素的体素特征进行扩展计算以得到所述目标点的点云特征。

2. 根据权利要求1所述的基于融合体素的三维点云标记方法,其特征在于,对点云进行体素化处理的步骤包括:

按照预设体素尺寸将点云坐标空间划分为多个体素;

按照所述体素的栅格参数将所述三维点云数据集中的各点归类至对应的体素中;

对归类后的各体素中的点进行采样以使得体素中的点的数量达到第一预设值。

3. 根据权利要求2所述的基于融合体素的三维点云标记方法,其特征在于,对归类后的各体素中的点进行采样的步骤包括:

若体素中包含的点的数量大于第一预设值,则从当前体素中随机采样第一预设值个点以使该体素中的点的数量达到所述第一预设值;

若体素中包含的点的数量小于第一预设值,则从当前体素中随机选取一个或多个点进行复制以使该体素中的点的数量达到所述第一预设值。

4. 根据权利要求1所述的基于融合体素的三维点云标记方法,其特征在于,将所述第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征的步骤包括:

将所述体素特征矩阵转换为4维张量,并将该4维张量分别输入具有不同大小的卷积核的三维卷积神经网络中以计算得到不同尺度下的体素特征;

将所述不同尺度下的体素特征分别输入具有不同大小的卷积核的三维反卷积神经网络中以得到多个不同尺度的体素特征,其中,在进行卷积时的各所述三维卷积神经网络的卷积核大小与进行反卷积时的各所述三维反卷积神经网络的卷积核大小对应相同。

5. 根据权利要求1所述的基于融合体素的三维点云标记方法,其特征在于,所述目标点

p的点云特征 F_p 为 $F_p = \frac{\sum_{j=1}^n w_v^j F_v^j}{\sum_{j=1}^8 w_v^j}$ ，其中， $w_v^j = \frac{1}{d(p, c_j)^2}$ ，表示根据目标点p与第j个邻域体

素中的中心点 c_j 之间的欧式距离得到的权重参数， F_v^j 表示第j个邻域体素的体素特征。

6. 根据权利要求1所述的基于融合体素的三维点云标记方法，其特征在于，所述基于融合体素的三维点云标记方法还包括：

将完成属性标记的三维点云输入CRF-RNN网络中进行点云属性标记优化。

7. 一种基于融合体素的三维点云标记装置，其特征在于，所述装置包括：

体素处理及特征提取模块，用于对三维点云数据集进行体素化处理，并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵；

多尺度体素特征计算模块，用于将所述第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征，并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵；

特征扩展模块，用于基于特征插值算法将所述第二体素特征矩阵中的体素特征扩展至所述三维点云数据集中的各点中以得到点云特征矩阵；

点云标记模块，用于将所述点云特征矩阵输入多层感知器中以实现三维点云的属性标记；

所述体素处理及特征提取模块，具体用于针对每个体素中的点云计算该点云的中心坐标，并基于该中心坐标对所述体素中的点云数据进行中心归一化处理以得到初始数据矩阵；将所述初始数据矩阵输入LGAB模块中以实现逐点的局部特征描述，并采用最大值池化对体素内的局部特征集合进行逐点池化操作以得到体素的全局特征并作为第一体素特征矩阵；

所述特征扩展模块，具体用于针对所述三维点云数据集中的每个点分别作为目标点，在由所述第二体素特征矩阵形成的体素空间中搜索与该目标点最近的预设数量的体素作为邻域体素；对所述邻域体素的体素特征进行扩展计算以得到所述目标点的点云特征。

8. 根据权利要求7所述的基于融合体素的三维点云标记装置，其特征在于，所述体素处理及特征提取模块包括：

体素划分单元，用于按照预设体素尺寸将点云坐标空间划分为多个体素；

点云归类单元，用于按照所述体素的栅格参数将所述三维点云数据集中的各点归类至对应的体素中；

点云采样单元，用于对归类后的各体素中的点进行采样以使得体素中的点的数量达到第一预设值。

一种基于融合体素的三维点云标记方法和装置

技术领域

[0001] 本发明涉及数据标记技术领域,具体而言,涉及一种基于融合体素的三维点云标记方法和装置。

背景技术

[0002] 随着最近几年三维传感器(如LiDAR、Microsoft Kinect、ASUS Xtion等)在各领域(如移动机器人、自动驾驶、遥感解译、虚拟现实、增强现实和战场态势感知等)的广泛应用,三维数据也在显著增长,而三维点云标记作为对三维数据进行处理的重要手段,是指在观测到的场景点云数据中,识别每个点的类别属性,并为每个点分配一个唯一的类别标签,比如建筑、道路或汽车等。但传统的基于体素卷积神经网络的语义标记要求体素内所有点共享同一语义标签,且需要规则化的数据输入,同时得到的标记结果也是体素级别的粗略标记结果。

发明内容

[0003] 有鉴于此,本发明提供一种基于融合体素的三维点云标记方法和装置,以解决上述问题。

[0004] 一方面,本发明较佳实施例提供一种基于融合体素的三维点云标记方法,所述方法包括:

[0005] 对三维点云数据集进行体素化处理,并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵;

[0006] 将所述第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征,并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵;

[0007] 基于特征插值算法将所述第二体素特征矩阵中的体素特征扩展至所述三维点云数据集中的各点中以得到点云特征矩阵;

[0008] 将所述点云特征矩阵输入多层感知器中以实现三维点云的属性标记。

[0009] 另一方面,本发明较佳实施例还提供一种基于融合体素的三维点云标记装置,所述装置包括:

[0010] 体素处理及特征提取模块,用于对三维点云数据集进行体素化处理,并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵;

[0011] 多尺度体素特征计算模块,用于将所述第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征,并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵;

[0012] 特征扩展模块,用于基于特征插值算法将所述第二体素特征矩阵中的体素特征扩展至所述三维点云数据集中的各点中以得到点云特征矩阵;

[0013] 点云标记模块,用于将所述点云特征矩阵输入多层感知器中以实现三维点云的属性标记。

[0014] 与现有技术相比,本发明实施例提供一种基于融合体素的三维点云标记方法和装置,通过在规则化的体素模型上基于体素卷积神经网络构建多尺度空间以提取多尺度体素特征,然后利用特征插值的方式将体素特征扩展到点特征,进而实现逐点的较为精细的分类识别和点云标记。

[0015] 另外,本发明还利用CRF-RNN对完成标记的点云数据进行标记结果优化,以提高标记精度。

附图说明

[0016] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本发明的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0017] 图1为本发明实施例提供的基于融合体素的三维点云标记装置的应用场景示意图。

[0018] 图2为本发明实施例提供的基于融合体素的三维点云标记方法的流程示意图。

[0019] 图3为图2中所示的步骤S11的子流程示意图。

[0020] 图4为本发明实施例提供的基于体素卷积神经网络的三维点云标记方法框架结构示意图。

[0021] 图5为本发明实施例提供的局部与全局特征融合模块的结构示意图。

[0022] 图6为本发明实施例提供的体素特征提取网络结构示意图。

[0023] 图7为图2中所示的步骤S12的子流程示意图。

[0024] 图8为本发明实施例提供的三维卷积操作示意图。

[0025] 图9为本发明实施例提供的卷积与反卷积二维示意图。

[0026] 图10为本发明实施例提供的基于3D CNN的多尺度特征提取结构示意图。

[0027] 图11为本发明实施例提供的基于体素特征的点云特征插值示意图。

[0028] 图12为本发明实施例提供的全连接CRF连接示意图。

[0029] 图13为本发明实施例提供的基于RNN的CRF实现结构示意图。

[0030] 图14为本发明实施例提供的融合体素卷积神经网络与CRF后端优化的三维点云标记网络结构示意图。

[0031] 图15为本发明实施例提供的基于XYZ坐标信息在S3DIS数据集上的标记结果示意图。

[0032] 图16为本发明实施例提供的基于XYZ坐标信息在vKITTI数据集上的标记结果示意图。

[0033] 图17为本发明实施例提供的Point-VoxelNet在两个数据集上的类别混淆矩阵。

[0034] 图18为本发明实施例提供的PVCRF在两个数据集上的类别混淆矩阵。

[0035] 图19为本发明实施例提供的基于融合体素的三维点云标记装置的方框结构示意图。

[0036] 图标:10-电子终端;100-基于融合体素的三维点云标记装置;110-体素处理及特征提取模块;111-体素划分单元;112-点云归类单元;113-点云采样单元;120-多尺度体素

特征计算模块;130-特征扩展模块;140-点云标记模块;200-存储器;300-存储控制器;400-处理器。

具体实施方式

[0037] 经发明人研究发现,在现有的场景点云的全局上下文信息建模上,通常利用图模型(graphical models)的表示能力可以得到解决,如比较常见的方式是将分类器与条件随机场(Conditional Random Fields,CRF)进行组合以估计每个数据点的语义标签。然而,分类器分类识别阶段与CRF优化阶段通常是作为单独的模块独立进行运算的,相互之间没有交互,从而限制了各模块之间的信息交流。

[0038] 其中,对于分类器,三维体素卷积神经网络是一个较好的选择。三维体素卷积神经网络是由二维卷积神经网络扩展而来,其在三维目标分类识别任务中也取得了很好的性能,相对于基于点云的深度神经网络而言,三维体素卷积神经网络也具有网络结构清晰、易于加速实现等优势。但体素神经网络需要规则化的数据输入,且其标记结果也是体素级别的粗略标记。

[0039] 针对现有技术中存在的上述问题,本发明实施例提供一种基于融合体素的三维点云标记方法和装置,通过在规则化的体素模型上基于体素卷积神经网络构建多尺度空间以提取多尺度体素特征,然后利用特征插值的方式将体素特征扩展到点特征,进而实现逐点的较为精细的分类识别,以进一步提升标记性能。为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。通常在此处附图中描述和示出的本发明实施例的组件可以以各种不同的配置来布置和设计。

[0040] 因此,以下对在附图中提供的本发明的实施例的详细描述并非旨在限制要求保护的本发明的范围,而是仅仅表示本发明的选定实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0041] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释。

[0042] 如图1所示,为本发明实施例提供的基于融合体素的三维点云标记装置100的应用场景示意图。其中,电子终端10包括基于融合体素的三维点云标记装置100、存储器200、存储控制器300以及处理器400。其中,电子终端10可以是,但不限于,电脑、移动上网设备(mobile Internet device,MID)等具有处理功能的电子设备,还可以是服务器等。

[0043] 可选地,存储器200、存储控制器300、处理器400各元件相互之间直接或间接地电性连接,以实现数据的传输或交互。例如,这些元件之间通过一条或多条通讯总线或信号线实现电性连接。基于融合体素的三维点云标记装置100包括至少一个可以软件或固件的形式存储于存储器200中或固化在电子终端10的操作系统中的软件功能模块。处理器400在存储控制器300的控制下访问存储器200,以用于执行存储器200中存储的可执行模块,例如基于融合体素的三维点云标记装置100所包括的软件功能模块及计算机程序等。

[0044] 可以理解,图1所示的结构仅为示意,电子终端10还可包括比图1中所示更多或者更少的组件,或者具有与图1所示不同的配置。图1中所示的各组件可以采用硬件、软件或其

组合实现。

[0045] 进一步地,请结合参阅图2,本发明实施例还提供一种基于融合体素的三维点云标记方法,下面结合图2对该基于融合体素的三维点云标记方法进行介绍。

[0046] 步骤S11,对三维点云数据集进行体素化处理,并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵;

[0047] 步骤S12,将第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征,并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵;

[0048] 步骤S13,基于特征插值算法将第二体素特征矩阵中的体素特征扩展至三维点云数据集中的各点中以得到点云特征矩阵;

[0049] 步骤S14,将点云特征矩阵输入多层感知器中以实现三维点云的属性标记。

[0050] 本实施例中,首先对三维点云进行体素化处理,接着对体素内点云进行特征提取,然后将以体素特征为元素的体素模型输入三维卷积神经网络进行多尺度特征提取与融合,接着利用特征插值算法将体素特征扩展到点云特征,实现对三维点云的标记,能够有效提高点云标记精度。

[0051] 详细地,请结合参阅图3,步骤S11中对点云进行体素化处理的过程可通过下述步骤S111-步骤S113实现:

[0052] 步骤S111,按照预设体素尺寸将点云坐标空间划分为多个体素;

[0053] 步骤S112,按照体素的栅格参数将三维点云数据集中的各点归类至对应的体素中;

[0054] 步骤S113,对归类后的各体素中的点进行采样以使得体素中的点的数量达到第一预设值。

[0055] 其中,本实施例在上述步骤S111-步骤S113中引入点云体素化模型以对点云进行体素化处理。具体地,如图4所示,点云体素化即按照给定的体素尺寸将点云坐标空间分割为多个体素。其中,假设输入点云在三个坐标轴即X、Y、Z轴方向上的尺寸分别为W、H、E,每个体素的尺寸为 λ_w 、 λ_h 、 λ_e ,则体素化处理后的模型的尺寸为 $W'=W/\lambda_w$ 、 $H'=H/\lambda_h$ 、 $E'=E/\lambda_e$ 。本实施例中,为了便于后续卷积运算,可以令 W' 、 H' 、 E' 为整数且为2的幂方。

[0056] 在步骤S111中对点云坐标空间实现体素化栅格后,可进一步根据各体素的栅格参数对点云中每个点进行归类,使得每个点归于每一个体素内。但在进行点云归类时,由于进行三维点云数据采集时受测量误差、距离远近、遮挡等因素的影响,采集到的点云往往是不均匀的,如部分区域点云比较集中,部分区域点云比较稀疏。加之,点云数据采集等价于目标表面的采样,因而目标内部是空的,不存在点云数据,从而导致对点云空间进行体素化后,每个体素内部的点云分布是不均匀的,具体如图4所示,其中左下角的体素不包含点云,右上角的体素包含点数较少。因此,为便于后续进行统一的体素特征提取,在进行点云分块后需要从每个体素内采样得到相同数量个点,如第一预设值T(T根据点云分辨率以及存储容量来确定)。

[0057] 应注意的是,在进行采样时,若体素中包含的点的数量大于第一预设值,则从当前体素中随机采样第一预设值个点以使该体素中的点的数量达到第一预设值;若体素中包含的点的数量小于第一预设值,则从当前体素中随机选取一个或多个点进行复制以使该体素中的点的数量达到第一预设值。例如,假设第一预设值为T,对于体素内点的数量超过T的体

素则随机采样T个点即可,对于体素内点数少于T的,则随机复制相应数量的点以得到T个点的集合,进行点云分块与采样后可以获取不超过 $T' \times H' \times E'$ 个包含T个点的体素集合,进而利用体素内的点云数据进行特征学习以获取每个包含点云的体素的有效特征表达。

[0058] 进一步地,上述步骤S11中基于处理结果在体素内进行点云特征提取形成第一体素特征矩阵的步骤包括:针对每个体素中的点云计算该点云的中心坐标,并基于该中心坐标对体素中的点云数据进行中心归一化处理以得到初始数据矩阵;将初始数据矩阵输入LGAB模块中以实现逐点的局部特征描述,并采用最大值池化对体素内的局部特征集合进行逐点池化操作以得到体素的全局特征并作为第一体素特征矩阵。

[0059] 具体地,本发明实施例中采用如图5所示的局部与全局特征融合模块(LGAB)进行堆叠以搭建特征学习网络进行体素特征提取。其中,如图6所示,假设 V_x 为包含T个点的非空体素,即 $V_x = \{p_i = (x_i, y_i, z_i), i=1, 2, 3, \dots, T\}$,那么在将点云数据(初始数据矩阵)输入LGAB模块前进行中心归一化,即首先计算体素内点云的中心坐标 (c_x, c_y, c_z) ,利用中心坐标 (c_x, c_y, c_z) 对点云数据进行中心归一化即得到最终的输入数据初始数据矩阵 $\overline{V}_x = \{\overline{p}_i = (x_i, y_i, z_i, x_i - c_x, y_i - c_y, z_i - c_z)\}$,也就是说,体素特征提取模块的输入为 $T \times 6$ 维的初始数据矩阵。

[0060] 进一步地,利用堆叠的LGAB模块能够得到逐点的局部特征描述,采用最大值池化(Max-Pooling, MP)对体素内逐点特征集合进行池化操作即可得到体素的全局特征。如图6给出了针对非空体素的特征提取示例,为减少参数数量,其余非空体素在进行特征提取时可共享相同网络参数。实际实施时,由于LGAB模块能够很好的融合点云局部邻域共有信息与各自的差异信息,本实施例中利用多个LGAB模块的级联能够有效提取体素内部的点云信息。

[0061] 进一步地,如图7所示,步骤S12中将第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征的过程可通过下述步骤实现。

[0062] 步骤S120,将体素特征矩阵转换为4维张量,并将该4维张量分别输入具有不同大小的卷积核的三维卷积神经网络中以计算得到不同尺度下的体素特征;

[0063] 步骤S121,将不同尺度下的体素特征分别输入具有不同大小的卷积核的三维反卷积神经网络中以得到多个不同尺度的体素特征,其中,在进行卷积时的各三维卷积神经网络的卷积核大小与进行反卷积时的各三维反卷积神经网络的卷积核大小对应相同。

[0064] 详细地,由于空间几何信息是三维目标的重要信息,直接对三维数据进行处理可以提取目标的有效特征描述,因此,本发明借鉴二维卷积神经网络在图像处理中的巨大成功,将二维卷积神经网络扩展到三维卷积神经网络中,即采用三维卷积神经网络对三维数据进行处理时,需要对三维数据进行规则化处理即体素化处理,还需要将二维卷积操作、池化操作等扩展到三维体素数据上。其中,三维卷积公式为:

$$[0065] \quad \hat{y}(m, n, k) = (\hat{x} * \hat{w})(m, n, k) = \sum \sum \sum \hat{x}(i + m, j + n, l + k) \hat{w}(i, j, l) \quad (1)$$

[0066] 式(1)中, $\hat{x}(m, n, k)$ 为输入的三维体素数据, $\hat{w}(m, n, k)$ 为三维卷积核模板, $\hat{y}(m, n, k)$ 为其输出响应。与二维卷积操作类似,三维卷积操作只是将二维卷积核扩展到三维卷积核,包含长宽高三个维度,相应的,如图8所示,其局部感受也由二维平面上的局部邻域转换为三维空间中的局部邻域。实际实施时,由于三维卷积操作会减少三维数据的空间

尺寸,即三维特征图的空间尺寸要小于输入体素数据的尺寸。但对于三维数据标记,需要获取每个数据点的特征信息,即在进行体素特征提取时需要获取每个体素的特征,因而需要将卷积操作后的特征图反向映射回最初的输入体素中。为解决这一问题,本实施例中采用反卷积操作(deconvolution)对得到的三维体素特征图进行处理以获取与输入体素数据同尺寸的特征图。反卷积操作的核心仍然是卷积操作,只是在卷积操作之前对输入特征数据进行边缘补0操作,以保证输出特征图的尺寸要求。例如,图9所示给出了卷积与反卷积操作的二维示例,其中蓝色标记数据为两个操作的输入数据,灰色标记数据为卷积核(两者的尺寸相同),绿色标记数据为两者的响应输出。不难看出,反卷积层的输入为卷积层的输出,反卷积层的输出与卷积层的输入具有相同尺寸。

[0067] 进一步地,基于前述操作对所有非空体素进行处理后,能够获取一系列D维的体素特征。由于每个体素特征与三维空间的体素坐标唯一对应,因而可以将获取的特征表示为4维张量,其尺寸大小为 $W' \times H' \times E' \times D$, (对于空的体素,用D维零矢量作为其特征描述)。将特征表示转换为基于4维张量的特征表示后即可采用三维卷积神经网络(3D CNN)进行进一步的特征优化。考虑到固定尺度(卷积核尺寸)下的特征提取不足以完整表达体素的局部邻域结构信息,本发明中采用多尺度的特征提取与融合,提取更为丰富的局部邻域信息。

[0068] 其中,如图10所示,为具体的三维卷积神经网络结构,即该基于3D CNN的多尺度特征提取主要包含三维卷积操作(Conv3D)、三维反卷积操作(DeConv3D)以及特征串联操作(Concat)。对于输入的 $W' \times H' \times E' \times D$ 维特征,采用三个不同的卷积核分别进行卷积、反卷积操作,记为 $\text{Conv3D}(f_{in}; f_{out}; \text{ker}; \text{st}; \text{pad})$, $\text{DeConv3D}(f_{in}; f_{out}; \text{ker}; \text{st}; \text{pad})$, 其中 f_{in} 、 f_{out} 表示输入输出特征矩阵的维度,ker;st;pad分别表示卷积核尺寸、卷积时卷积核移动步长大小、进行数据扩展时数据填充尺寸,均为三维矢量。为获取不同尺度下的特征,三个卷积核可以为但不限于 $(1;2;2)$; $(2;1;2)$; $(2;2;1)$, 卷积与反卷积操作中采用相同的卷积核,反卷积操作中进行相应的边缘补0操作。此外,卷积层与反卷积层中不仅包含卷积、反卷积操作,每个卷积操作后续还包含归一化层(Batch Normalization layer,BN)以及ReLU激活操作。

[0069] 基于3D CNN的多尺度特征提取沿着三维空间三个相互垂直的方向(即X、Y、Z方向)采用不同的卷积核进行体素特征提取与融合,使得学习到的特征能够包含更多的局部结构信息,实现对点云更为完备的表达。

[0070] 进一步地,在步骤S13中,由于在实现三维点云的标记时,需要获取每个点的特征描述,而前述三维卷积神经网络只能获取每个体素的特征描述,因此,本发明中通过对体素特征进行插值以获取输入点云中每个点的特征描述。如图11所示,对于给定目标点p,找出由第二体素特征矩阵形成的体素空间中最近的预设数量(如8个)个邻域体素,各邻域体素对应的特征描述分别为 $\{F_v^j\}$, 其中 $j=1,2,\dots,8$, 则目标点p的特征描述为:

$$[0071] \quad F_p = \frac{\sum_{j=1}^n w_v^j F_v^j}{\sum_{j=1}^8 w_v^j} \quad (2)$$

[0072] 式(2)中, $w_v^j = \frac{1}{d(p, c_j)^2}$ 表示根据目标点p与第j个邻域体素中的中心点 c_j 之间的欧式距离得到的权重参数, F_v^j 表示第j个邻域体素的体素特征,对三维点云中的各点重复

执行上述过程,即可将第二体素特征矩阵中的体素特征扩展至三维点云数据集中的各点中以得到点云特征矩阵。

[0073] 进一步地,在步骤S14中,将步骤S13得到的各点的点云特征矩阵输入多层感知器(MLP)即可实现点的类别识别,即三维点云标记,其具体网络结构请再次参阅图4。

[0074] 根据实际需求,为了进一步优化上述采用卷积神经网络获取的点云标记结果,提高标记精度,本发明还包括步骤S15。

[0075] 步骤S15,将完成属性标记的三维点云输入CRF-RNN网络中进行点云属性标记优化。

[0076] 具体地,本实施例将基于CNN基本操作实现的FC-CRF嵌入前述基于三维卷积神经网络的点云标记网络中,实现端到端的、融合粗标记与后端优化的三维点云精细标记网络,进一步提高点云标记的准确性,特别是目标边界、轮廓的平滑性。

[0077] 本实施例中,首先对面向三维点云标记的CRF建模,然后基于CNN操作的CRF近似实现,最后融合CRF优化的三维点云标记。

[0078] 传统的语义标记均是建模为逐点的分类识别,有采用局部特征进行分类识别的,也有采用深度神经网络进行分类识别。但是逐点的分类识别通常会带来一些明显不能接受的标记错误,比如对于某个目标内部的部分点可能会被识别为其他类别,这是由于逐点分类识别没有考虑到点与点之间的邻接关系,其只用到了待标记点的部分、小尺寸的邻域信息。如果能够事先对目标结构信息进行建模(比如:所有目标都是连续的,其具有类似特征的相邻点应该标记为同一类目标)并基于建模结果对标记结果进行优化、限制,一些明显的错误便可以得到有效的剔除,进而获得高精度的标记结果。条件随机场(CRF)是对目标连续性以及其邻域结构信息进行建模的有效方法,且已广泛应用于二维图像标记中。其中,条件随机场是在给定一组随机变量的条件下计算另一组输出随机变量的条件概率分布的模型,其主要特点是假设输出变量构成马尔可夫随机场(Markov Random Field, MRF)。

[0079] 详细地,CRF是一种判别式的概率无向图模型,可以对数据中的全局上下文信息、交叉性特征进行建模,是一种能够很好的处理序列数据分割与标记的概率图模型。假设给定随机变量集合 $X = \{X_1, X_2, \dots, X_N\}$ 与 $P = \{P_1, P_2, \dots, P_N\}$,其中, $X_i \in L = \{1_1, 1_1, \dots, 1_M\}$,对于三维点云标记, P 为包含 N 个点的输入点云, P_j 为第 j 个点的观测矢量, X 为输入点云的语义标记结果, X_i 为第 i 个点的语义标签,其取值为 M 个语义标签中的一个,那么对应的CRF模型可以用Gibbs概率分布进行表示,即:

$$[0080] \quad \Phi(X|P) = \frac{1}{Z(P)} \exp \left(- \sum_{o \in O_G} \Lambda(x_o|P) \right) \quad (3)$$

[0081] 式(3)中, G 为在随机变量集 X 上构建的概率无向图, o 为图 G 中的基团,其中, o 中每对节点均是相邻的, o_G 则为 G 中所有基团的集合, $Z(P)$ 为归一化函数, $\Lambda(x_o|P)$ 为基团上的能量函数,又称为势函数。

[0082] 对于任意一种标记结果 $x \in L^N$ 的整体势函数为:

$$[0083] \quad \Lambda(x|P) = \sum_{o \in O_G} \Lambda(x_o|P) \quad (4)$$

[0084] 基于最大后验概率算法进行求解得到最优的标记结果为:

$$[0085] \quad x^* = \arg \max_{x \in L^N} \Phi(X|P) \quad (5)$$

[0086] 从前述最优解求解过程中可以看出,标记结果后验概率的最大化即整体势函数的最小化。不难看出,条件随机场首先通过基团势函数来实现局部上下文信息的建模,然后利用图结构进行上下文信息的传递,进而实现大范围上下文信息的建模。

[0087] 对于全连接CRF模型,图G中每个节点与其余节点均相连,如图12所示,相应的基团 ϕ_0 为包含单个节点或者包含成对节点的基团,因而 x 对应的整体势函数可以表示为:

$$[0088] \quad \Lambda(x) = \sum_i \varphi_u(x_i) + \sum_{i < j} \varphi_p(x_i, x_j) \quad (6)$$

[0089] 其中,为便于描述,在公式(6)中去掉条件后验概率中的条件 P 部分,即有 $\Lambda(x) = \Lambda(x|P)$, $\varphi_u(x_i) = \varphi_u(x_i|P)$ 为单元势函数, $\varphi_p(x_i, x_j|P)$ 为成对二元势函数, $i, j = 1, 2, \dots, N$ 。单元势函数 $\varphi_u(x_i)$ 表示图G中第 i 个节点被标记为 x_i 代价,该函数通常由某判别式分类器的概率输出来定义,这时的估计结果往往包含较多噪点,且分割结果往往在目标边缘不连续。成对二元势函数给出的是同时标记第 i, j 个观测点为 x_i, x_j 的代价,其具有保留相邻观测点标记一致性、减少不一致性的作用,可以提高标记结果的平滑性。

[0090] 采用高斯加权的思想将成对二元势函数定义为:

$$[0091] \quad \varphi_p(x_i, x_j) = \psi(x_i, x_j) \sum_{m=1}^{M_G} w^{(m)} k_G^{(m)}(f_i, f_j) \quad (7)$$

[0092] 式(7)中,函数 $\psi(x_i, x_j)$ 为不同标签之间的一致性计算函数, $w^{(m)}$ 为权值, $k_G^{(m)}$ 为基于高斯核的平滑滤波函数,总共有 M_G 个高斯核函数, f_i, f_j 分别表示观测点 i, j 的特征矢量,且有 $k_G^{(m)}(f_i, f_j) = \exp(-\frac{1}{2}(f_i, f_j)^T \Lambda^m(f_i, f_j))$,每个高斯核函数 $k_G^{(m)}$ 可通过一个对称的、正定矩阵 Λ^m 来定义。至此,面向三维点云标记的全连接CRF建模已完成,下一步就是如何求解得到最优标记结果。

[0093] 基于全连接CRF的三维点云标记优化过程即基于输入点云数据最大化后验概率 $\Phi(X)$ 过程。求解准确的后验概率是比较困难的,且计算量庞大,基于平均场的近似方法可以将后验概率 $\Phi(X)$ 转化为一系列相互独立的边际概率的乘积,即 $\Phi(X) \approx \Theta(X) = \prod_{i=1} \Theta_i(X_i)$,结合式(3)-(7)可以得到任意标记结果 x_i 的边际概率 Θ_i 为:

$$[0094] \quad \Theta_i(x_i = l) = \frac{1}{Z_i} \exp \left(-(\varphi_u(x_i) + \sum_{l' \in L} \psi(l, l') \sum_{m=1}^{M_G} w^{(m)} \sum_{j \neq i} k_G^{(m)}(f_i, f_j) \Theta_j(l')) \right) \quad (8)$$

[0095] 基于式(8)可以构建CRF的迭代推断算法,如算法5.1所示。该迭代算法的收敛性主要通过估计的 Q 与 P 之间的差异性来度量,通过评估算法的收敛性可以得到,在迭代次数为10时估计误差已经很小,证明了算法具有良好的收敛性。

[0096] 算法1:基于平均场近似的CRF迭代推断算法

$$[0097] \quad 1. \text{初始化:对所有节点进行初始化 } \Theta_i(x_i = l) = \frac{1}{Z_i} \exp(-\varphi_u(x_i = l))$$

[0098] 2. while为收敛do

$$[0099] \quad 3. \text{信息传递:计算所有的高斯滤波结果 } \hat{\Theta}_i^{(m)}(l) = \sum_{j \neq i} k_G^{(m)}(f_i, f_j) \Theta_j(l)$$

[0100] 4. 加权滤波: $\tilde{\Theta}_i(l) = \sum_m w^m \hat{\Theta}_i(l)$

[0101] 5. 一致性检测: $\hat{\Theta}_i(l) = \sum_{l' \in L} \psi(l, l') \tilde{\Theta}_i(l')$

[0102] 6. 增加一元势函数: $\check{\Theta}_i(l) = -\varphi_u(l) - \hat{\Theta}_i(l)$

[0103] 7. 归一化: $\Theta_i = \frac{1}{Z_i} \exp(\check{\Theta}_i(l))$

[0104] 8. end while

[0105] 下面对如何利用CNN中的相关操作来实现CRF迭代推断上述算法进行介绍。基于CNN操作对算法进行重构中最大的问题在于能否实现误差的反向传播,即采用BP算法进行参数学习训练。

[0106] (1) 初始化操作

[0107] 算法1中的初始化操作为:

$$[0108] \quad \Theta_i(x_i = l) = \frac{1}{Z_i} \exp(-\varphi_u(x_i = l)) \quad (9)$$

[0109] 其中, $Z_i = \sum_l \exp(-\varphi_u(x_i = l))$, 即对所有的取值可能进行求和。记

$U_i(l) = -\varphi_u(x_i = l)$, 则有 $\Theta_i(l) = \frac{1}{Z_i} \exp(U_i(l))$, $Z_i = \sum_l \exp(U_i(l))$, 可以看出这一操作等价于

在每个场景点上对所有可能标记结果的 $U_i(l)$ 进行基于Softmax函数的激活操作。Softmax函数是CNN网络中常用的激活函数,其不包含任何参数,其误差导数可以进行反向传播,因而也可以利用Back-Propagation(BP)算法进行学习训练。

[0110] (2) 信息传递

[0111] 如算法1所示,CRF中的信息传递即利用 M_g 个高斯滤波器对 Θ_j 进行平滑滤波。高斯滤波器的核函数是根据点云的特征得到的,比如点的坐标信息或颜色、强度信息,表达了每个场景点之间的关联关系。全连接CRF模型中,每个滤波器需要覆盖点云中的所有点,其数据量与计算量均是很大的,因而不能直接进行实现。这里采用基于全自由度多面体晶格方法(permutohedral latTice)来实现快速高斯卷积,其计算量为 $O(N)$, N 为进行滤波的点的数量,相比传统的高斯卷积具有更快的速度以及更好的滤波效果。基于全自由度多面体晶格方法的快速高斯卷积包含四个阶段,多面体晶格构建阶段、扩展(splat)映射、切片(slice)映射以及模糊(blur)阶段等。

[0112] 在反向传播中,当前卷积层的输入(误差导数)为上一层滤波器的输出沿着反方向经过 M_g 个高斯滤波器后的输出结果。基于全自由度多面体晶格方法的高斯卷积中,这一反向传播,可以在保持与正向传播时相同的多面体晶格构建、扩展映射与切片映射的基础上逆转模糊阶段中滤波器的顺序来实现。这一实现方法的计算量仍然为 $O(N)$, 明显减少了计算量,提高了计算效率。

[0113] (3) 加权滤波

[0114] 接下来的计算是对每个语义标签 l 的前述 M_g 个输出结果进行加权求和。在点云标记中,每个语义标签之间是相互独立的,因而这一加权滤波操作可以通过 M_g 个卷积核为 1×1 的卷积操作来实现,其中该卷积操作的输入为包含 M_g 个通道的特征矩阵,输出为包含1个

通道的特征矩阵。在反向传播中,由于这一步操作的输入输出均是已知的,卷积核之间也是相互独立的,因而关于卷积核参数以及关于输入数据的误差导数均是可以计算出来的,进而可以利用BP算法对卷积核参数进行学习训练。

[0115] (4)一致性检测

[0116] 在一致性检测中,利用PoTTs模型对上一步中不同标签的输出结果进行兼容性计算。兼容性计算主要是比较两个相似观测点的标签是否相同,当两个点的语义标签相同时一致性检测为0,当两个点的语义标签不同时引入惩罚项 σ ,其计算如下所示:

$$[0117] \quad \psi(l, l') = \begin{cases} 0, & \text{if } l = l' \\ \sigma, & \text{if } l \neq l' \end{cases} \quad (10)$$

[0118] 相比较于采用固定的惩罚项 σ ,本发明考虑采用基于数据学习到的惩罚值,这是由于不同标签之间的关联度是不一样的,因而在相邻点标记为不同标签对于整个标记结果影响是不一样的。因此,一致性检测也可以看做是一个卷积层,该层的输入输出通道数均为M(标签数),卷积核尺寸为 1×1 ,其学习到的神经元连接权重参数即为转换函数的取值。由于采用了基本卷积操作进行实现,因而这一步也是可以进行反向传播的。

[0119] (5)增加一元势函数

[0120] 将一元势函数 $U_i(l) = -\phi_u(l)$ 与一致性检测中得到的输出结果进行逐元素的组合进而得到完整的势函数结果。增加一元势函数这一步中,不包含任何参数,因而可以简单的将输出中的误差复制到输入端来实现反向传播。

[0121] (6)归一化

[0122] 类似于初始化的过程,归一化步骤中也可以通过基于Softmax函数的激活操作来实现,其反向传播与CNN中的基于Softmax函数的反向传播一致。至此,已经采用CNN网络中的基本操作对算法1中单次迭代的各个步骤进行了实现,对上述步骤进行堆叠即可实现多次迭代的求解算法。

[0123] 基于上述描述,本实施例中首先采用平均场的近似方法对CRF模型进行了近似建模,然后采用CNN中的基本操作对平均场近似方法中的每一步进行了等价实现,即实现了单次迭代的平均场近似算法。对于迭代的平均场近似方法只需要对相关的步骤进行堆叠即可,即采用递归的CNN结构(RNN)即可实现迭代的平均场近似计算,其结构如图13所示,给定输入的点云数据为P,逐点的一元势函数为 $U = U_1(1)$,前一次迭代得到的边际概率为 H_1 ,当前迭代得到的边际概率为 H_2 ,单次平均场近似估计记为 $f_{\Omega}(U, P, H_1)$, Ω 为其参数集合(包含加权滤波以及一致性检测中的所有参数)且 $\Omega = \{w^{(m)}, \psi(l, l')\}$ 。对于 H_1 ,开始迭代时初始化为以U为输入的softmax函数输出,在之后的迭代中等于 H_2 的输出,即有:

$$[0124] \quad H_1(t') = \begin{cases} \text{softmax}(U), & t' = 0 \\ H_2(t' - 1), & 0 < t' \leq T' \end{cases} \quad (11)$$

[0125] 其中, T' 为迭代次数。

[0126] 在得到 H_1 后,基于平均场近似算法进行 H_2 的估计,即有:

$$[0127] \quad H_2(t') = f_{\Omega}(U, P, H_1(t')), 0 < t' \leq T'$$

[0128] 对于输出Y,只在最后一次迭代时输出估计结果,即 $Y = H_2(T')$ 。

[0129] 基于上述分析可以知道,整个网络结构(记为CRF-RNN)中关于参数的误差导数是

可求的,因而其可以用标准的BP算法进行求解,因而也可以嵌入到其他神经网络中进行学习训练。

[0130] 进一步地,基于前述的点云标记网络Point-VoxelNet网络构建了融合三维体素卷积神经网络与CRF后端优化的三维点云标记网络(Point VoxelNet+CRF-RNN,PVCRF),其具体结构如图14所示,其中,对于输入的点云,首先根据输入点云的场景尺寸进行体素化并在体素内随机选取固定数量的点云用于后续的特征提取,接着基于LGAB模块在体素内进行特征提取获取简单的体素特征,在获取所有非空体素特征后(空体素的特征采用补0操作)基于三维卷积神经网络(Conv3D,DeConv3D)进行多尺度的体素特征提取与融合,然后利用插值的方法将多尺度体素特征扩展到所有点上进而获得逐点的点特征,接着将点特征输入多层感知器中以获取初步点云标记结果,最后基于CRF-RNN网络结构进行后端优化。这一网络结构很好的实现了点云标记中分类识别阶段与CRF优化阶段的信息交互,对提高点云标记精度有明显作用。

[0131] 基于对上述基于融合体素的三维点云标记方法的描述,发明人还对该基于融合体素的三维点云标记方法的性能进行了验证,如点集交并比(Intersection over Union, IoU)以及整体准确率(Overall Accuracy, OA)等评价指标对点云标记性能进行评测。

[0132] (1)网络实现与参数设置

[0133] 在点云数据栅格化与采样阶段,需要对不同的数据集进行不同处理。

[0134] • S3DIS:对于S3DIS数据集,其场景沿Z、Y、X轴方向的最大尺寸范围分别为 $E=8\text{m}$, $H=16\text{m}$, $W=50\text{m}$ 。为覆盖全部场景,以 $8\times 16\times 50$ 为整个栅格的尺寸,每个体素的尺寸为 $\lambda_E=0.5\text{m}$ 、 $\lambda_H=0.25\text{m}$ 、 $\lambda_W=0.2\text{m}$,构建的体素模型尺寸为 $E'=16$, $H'=64$, $W'=256$,其中多出来的体素置空。每个体素中选取 $T=32$ 个点。

[0135] • vKITTI:对于vKITTI数据集,其场景沿Z、Y、X轴方向的最大尺寸范围分别为 $E=33\text{m}$, $H=193\text{m}$, $W=148\text{m}$ 。这里设置每个体素的尺寸为 $\lambda_E=2\text{m}$ 、 $\lambda_H=1.6\text{m}$ 、 $\lambda_W=1.2\text{m}$,构建的体素模型尺寸为 $E'=16$, $H'=128$, $W'=128$ 。同样的,每个体素中选取 $T=32$ 个点。

[0136] 对于CRF-RNN网络,为防止过拟合以及梯度消逝等,在训练阶段设置其迭代次数为 $T'=5$,在测试阶段设置 $T'=10$ 。高斯滤波器尺寸与点云数据尺寸一致。

[0137] 本发明采用两步训练的策略,第一步对Point-VoxelNet进行单独训练,第二步对联合Point-VoxelNet与CRF-RNN的网络PVCRF进行微调。在Point-VoxelNet网络,采用动量值为0.9的Adam优化算法进行训练优化,其初始学习率为0.001,训练的batch大小为16。对PVCRF网络,采用动量值为0.6的Adam优化算法进行训练优化,其初始学习率为0.0001,训练的batch大小为16。在训练期间,同样采用早停止策略以获取最优网络参数,最大训练轮数为100,如果网络参数在连续10轮训练后仍然没有更新则停止训练。在测试阶段,同样的,采用6折交叉验证对算法进行验证,训练数据与测试数据的分组情况如表1所示。

[0138] 基于Python语言、采用Tensorflow的深度学习框架来实现提出的三维体素神经网络结构与CRF-RNN网络结构。实验硬件环境为:Intel Core i76700KCPU,48G内存,GTX 1080Ti显卡(支持CUDA 8.0、cuDNN 5.1)。

[0139] (2)定量结果分析

[0140] 基于上述两个数据集比较了深度神经网络模型(Point VoxelNet与PVCRF)在三维点云标记中的应用效果,并与当前较好的标记算法PointNet、MS+CU(2)、SEGCloud、

3DContextNet进行了对比分析。其中,只采用点云的直角坐标信息,即XYZ坐标进行处理。

[0141] 表1

		平均IoU (%)	OA (%)
[0142]	PointNet	40.0	72.1
	MS+CU (2)	43.0	75.4
	Point-VoxelNet	50.3	80.6
	PVCRF	51.8	81.2

[0143] 表2

		平均IoU (%)	OA (%)
[0144]	PointNet	17.9	63.3
	MS+CU (2)	26.4	73.2
	Point-VoxelNet	37.8	82.1
	PVCRF	39.1	82.6

[0145] 表1和表2分别给出了不同网络模型在两个小型数据集上的统计标记结果,其中本实施例提出的PVCRF模型在S3DIS、vKITTI数据集上取得了较优的平均IoU,分别为51.8%、39.1%,在总体标记准确率OA上也取得了较好的结果,分别为81.2%、82.6%,表明PVCRF模型能够较好的获取点云的完备特征表达,证明基于三维体素空间的多尺度特征提取与基于欧式空间的多尺度特征提取具有相当的功能,均能为三维点云标记提供点云场景的细节信息。

[0146] 比较PVCRF与SEGCloud,SEGCloud直接采用二值体素模型进行体素特征学习,而PVCRF采用的是栅格化的点云体素模型,并利用体素内的点云进行体素特征学习。两者均添加了基于全连接CRF模型的后端优化。此外,PVCRF中包含了基于三维体素空间的多尺度特征提取模块,因而PVCRF取得了更高的平均IoU,充分说明利用体素内点云数据进行处理以及多尺度特征提取模块能够提取具有更强表征能力的特征描述。

[0147] 比较Point-VoxelNet与PVCRF模型,PVCRF在平均IoU以及总体准确率OA两方面均取得了较优的性能,这是由于全连接CRF模型可以对更大范围的上下文信息进行建模,对点云的邻接关系具有更强的表征能力。

[0148] 在不同的数据集上,以PointNet作为基准,由于S3DIS数据集中的点云比较集中,点云密度普遍较高,本实施例提出的网络模型PVCRF取得了较大的性能提升,而在vKITTI数据集中,由于其点云分布普遍较为稀疏,因而性能提升则要小一点。

[0149] 不同算法在S3DIS与vKITTI数据集上获得的标记结果以及真实标记结果如图15和图16所示。从左至右依次为:输入彩色点云,基于PointNet、Point-VoxelNet、PVCRF的标记结果,真实标记结果。从图15和图16中的标记结果可以看出,基于PVCRF取得的结果优于PointNet与Point-VoxelNet,标记结果与真实结果更为接近,证明了PVCRF中的多尺度特征

学习与CRF后端优化的有效性。相比较于PointNet与Point-VoxelNet,基于Point-VoxelNet取得的标记结果要优于PointNet,这主要是由于Point-VoxelNet中多尺度特征学习的引入提高了网络模型的特征学习能力。

[0150] 在室内外场景中,对于重叠度比较高或者连接紧密的目标,Point-VoxelNet、PVCRF模型仍然比较难分开,如图15中的墙板(board)与窗户(window)、图16中的道路(road)与地形(terrain)。由于Point-VoxelNet与PVCRF模型中均考虑了多尺度特征提取,对于大于场景内目标尺寸的目标均能很好的标记出来,比如图15中的桌子(table)以及图16中的建筑物(building)等。

[0151] 相比较Point-VoxelNet与PVCRF,由于全连接CRF能够提取场景点云中的大尺度上下文信息,使得标记结果中的细节更为突出,如图15第一行中的椅子(chair)与沙发(sofa),以及图16中的道路(road)与地形(terrain)的分割结果所示。

[0152] (4)类别混淆矩阵分析

[0153] 图17、图18分别给出了Point-VoxelNet与PVCRF两个网络模型在S3DIS、vKITTI数据集上获得的类别混淆矩阵,矩阵网格中的数值为类别标记准确率,网格颜色也代表了准确率的大小。对比两者的结果可以发现,对于室内数据集S3DIS,Point-VoxelNet与PVCRF模型对各类目标的分割精度相当,CRF模型的引入主要是降低了桌子(table)与地板(floor)等之间的混淆。对于室外数据集vKITTI,CRF模型的引入主要是降低了建筑物(building)与货车(van)等之间的混淆度。

[0154] 从图17(a)可以看出,S3DIS数据集中,Point-VoxelNet网络模型对天花板(ceiling)、地板(floor)、门(door)、立柱(column)、横梁(beam)、窗户(window)、书柜(bookcase)、墙板(board)和椅子(chair)的识别准确率均在52%以上,标记准确率较低的有沙发(sofa)、墙壁(wall)和阴影(cluTter),其精度在35%到46%之间,13类目标中最差的是桌子(table),其精度为22%。PVCRF在S3DIS数据集上取得的平均类别准确率的优劣分布情况与Point-VoxelNet网络模型类似,但普遍精度都有所提高,其中桌子的准确率提高到了30%,如图18(a)所示。

[0155] 在vKITTI数据集中也有类似的对比结果,如图17(b)与18(b)所示。这一对比结果证明了CRF模型对场景中目标细节的准确把握与建模。对比图17与图18中两个数据集上不同结果,在vKITTI数据集上的标记精度普遍较低,这主要是由于vKITTI数据集中的点云相对于S3DIS数据集中的点云数据普遍比较稀疏且不均匀,因而点的邻域结构不是那么明显,进而导致信息提取时不充分,不利于实现高精度的点云标记。

[0156] (5)计算时间统计分析

[0157] 在S3DIS数据集上对不同算法的计算效率进行实验分析。S3DIS数据集中共包含272幅独立点云,采用不同算法对每一幅点云进行标记并统计其平均测试时间(只统计神经网络计算耗时,不统计数据准备阶段的耗时),其统计结果如表3。从表3中统计结果可以看出,PointNet的计算时间约为1.8s。本实施例中给出的融合三维体素卷积神经网络与CRF后端优化的网络模型PVCRF计算耗时最大,为4.52s,这是由于CRF的引入明显增大了计算量。

[0158] 表3

算法	PointNet	Point-VoxelNet	PVCRF
[0159]			
时间 (s)	1.75	2.54	4.52

[0160] 最后采用基于卷积神经网络实现的FC-CRF进行进一步优化以获取精细标记结果。

[0161] 进一步地,请结合参阅图19,本发明实施例提供的页面搭建装置100包括体素处理及特征提取模块110、多尺度体素特征计算模块120、特征扩展模块130和点云标记模块140。

[0162] 体素处理及特征提取模块110,用于对三维点云数据集进行体素化处理,并基于处理结果在体素内进行体素特征提取形成第一体素特征矩阵;本实施例中,上述步骤S11可由体素处理及特征提取模块110执行,即关于体素处理及特征提取模块110的具体描述可参考步骤S11,本实施例在此不再赘述。可选地,如图19所示在本实施例中,体素处理及特征提取模块110包括体素划分单元111、点云归类单元112和点云采样单元113。

[0163] 体素划分单元111,用于按照预设体素尺寸将点云坐标空间划分为多个体素;本实施例中,上述步骤S111可由体素划分单元111执行,即关于体素划分单元111的具体描述可参考步骤S111,本实施例在此不再赘述。

[0164] 点云归类单元112,用于按照体素的栅格参数将三维点云数据集中的各点归类至对应的体素中;本实施例中,上述步骤S112可由点云归类单元112执行,即关于点云归类单元112的具体描述可参考步骤S112,本实施例在此不再赘述。

[0165] 点云采样单元113,用于对归类后的各体素中的点进行采样以使得体素中的点的数量达到第一预设值。本实施例中,上述步骤S113可由点云采样单元113执行,即关于点云采样单元113的具体描述可参考步骤S113,本实施例在此不再赘述。

[0166] 多尺度体素特征计算模块120,用于将第一体素特征矩阵作为三维卷积神经网络的输入以计算得到体素的多尺度特征,并对该多尺度特征进行特征串联融合以得到第二体素特征矩阵;本实施例中,上述步骤S12可由多尺度体素特征计算模块120执行,即关于多尺度体素特征计算模块120的具体描述可参考步骤S12,本实施例在此不再赘述。

[0167] 特征扩展模块130,用于基于特征插值算法将第二体素特征矩阵中的体素特征扩展至三维点云数据集中的各点中以得到点云特征矩阵;本实施例中,上述步骤S13可由特征扩展模块130执行,即关于特征扩展模块130的具体描述可参考步骤S13,本实施例在此不再赘述。

[0168] 点云标记模块140,用于将点云特征矩阵输入多层感知器中以实现三维点云的属性标记。本实施例中,上述步骤S14可由点云标记模块140执行,即关于点云标记模块140的具体描述可参考步骤S14,本实施例在此不再赘述。

[0169] 综上,本发明实施例提供了一种基于融合体素的三维点云标记方法和装置,通过在规则化的体素模型上基于体素卷积神经网络构建多尺度空间提取多尺度体素特征,然后利用特征插值的方式将体素特征扩展到点特征,进而实现逐点的较为精细的分类识别和点云标记。

[0170] 在本发明的描述中,术语“设置”、“相连”、“连接”应做广义理解,例如,可以是固定连接,也可以是可拆卸连接,或一体地连接;可以是机械连接,也可以是电连接;可以是直接相连,也可以通过中间媒介间接相连,可以是两个元件内部的连通。对于本领域的普通技术

人员而言,可以具体情况理解上述术语在本发明中的具体含义。在本发明实施例所提供的几个实施例中,应该理解到,所揭露的装置和方法,也可以通过其他方式实现。以上所描述的装置和方法实施例仅仅是示意性的,例如,附图中的流程图和框图显示了根据本发明的预设数量个实施例的装置、方法和计算机程序产品可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分。所述模块、程序段或代码的一部分包含一个或预设数量个用于实现规定的逻辑功能。

[0171] 也应当注意,在有些作为替换的实现方式中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0172] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

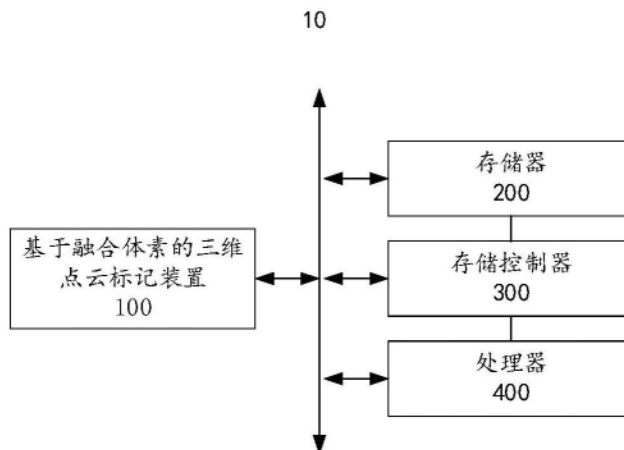


图1

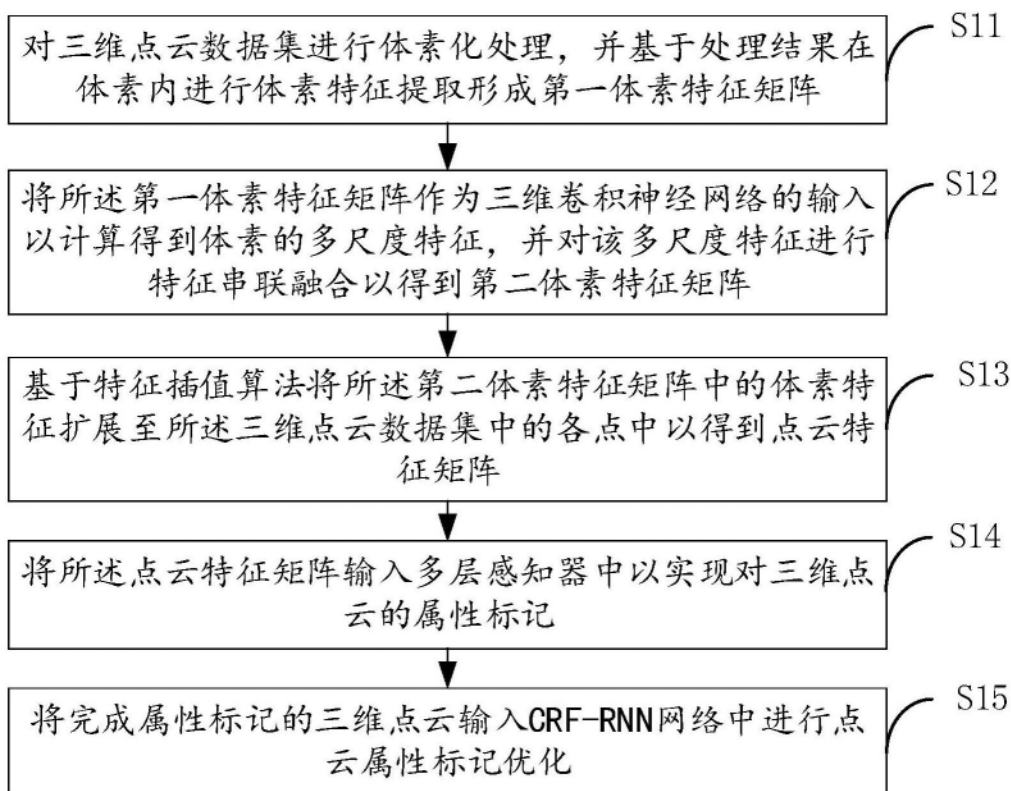


图2

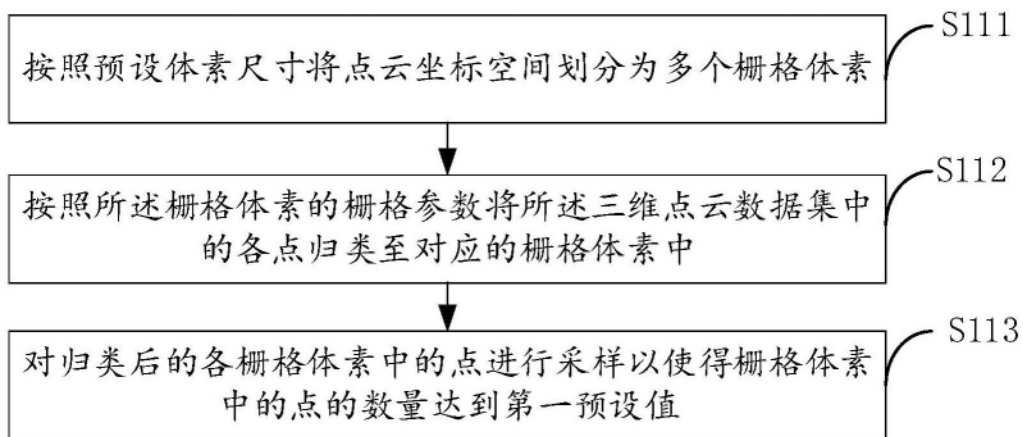


图3

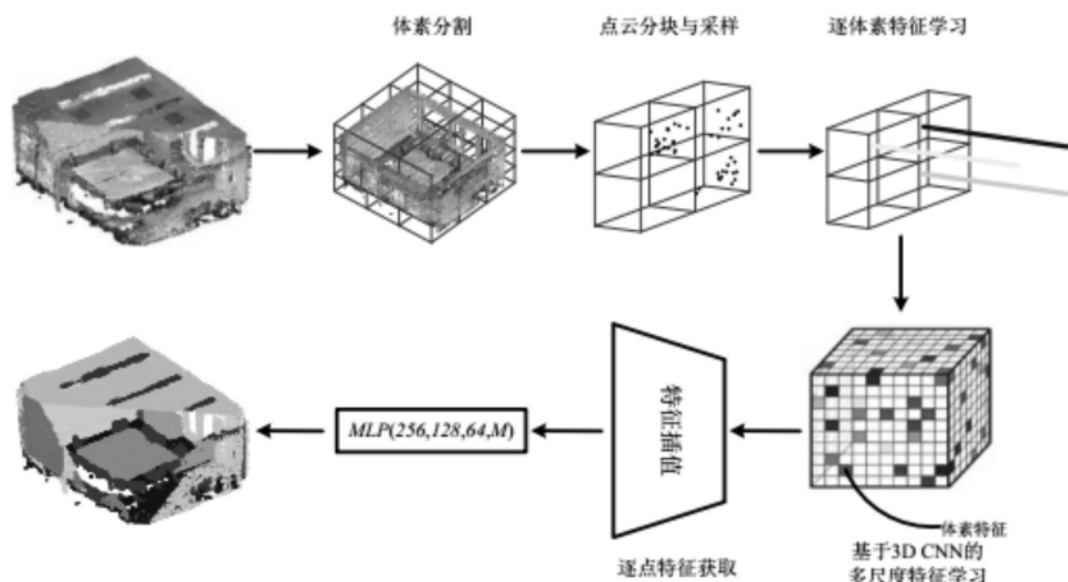


图4

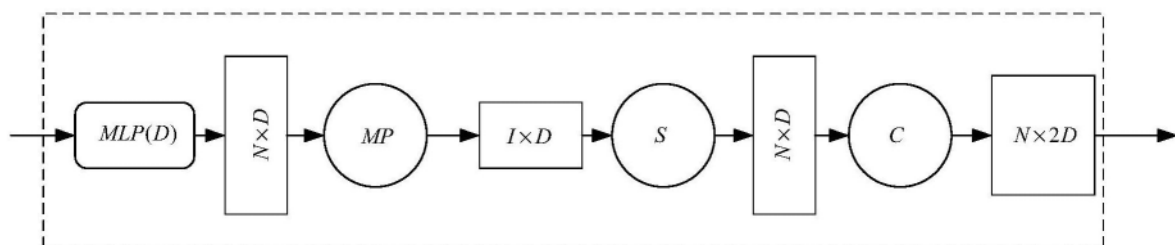


图5

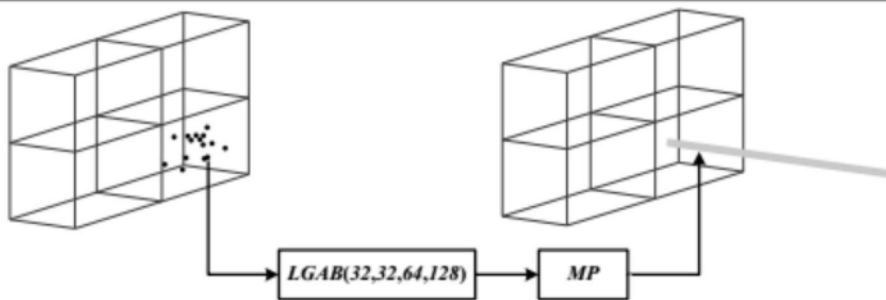


图6

将所述体素特征矩阵转换为4维张量，并将该4维张量分别输入具有不同大小的卷积核的三维卷积神经网络中以计算得到不同尺度下的体素特征

S120

将所述不同尺度下的体素特征分别输入具有不同大小的卷积核的三维反卷积神经网络中以得到多个不同尺度的体素特征，其中，在进行卷积时的各所述三维卷积神经网络的卷积核大小与进行反卷积时的各所述三维反卷积神经网络的卷积核大小对应相同

S121

图7

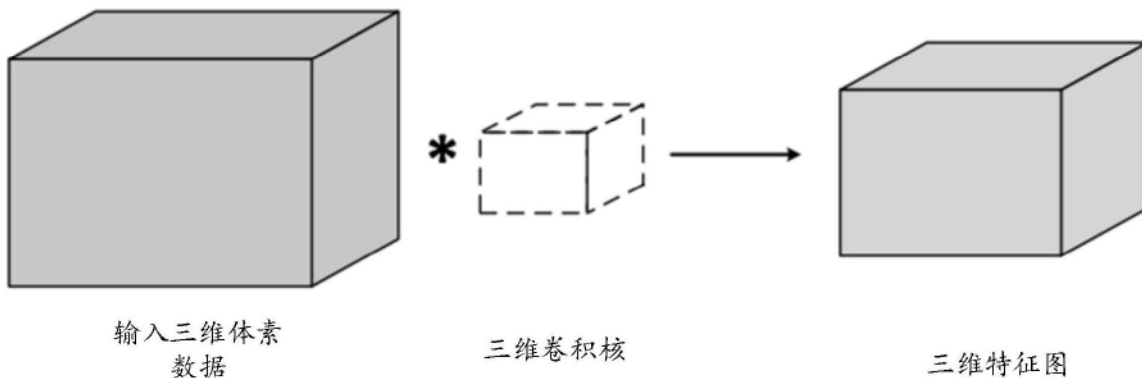
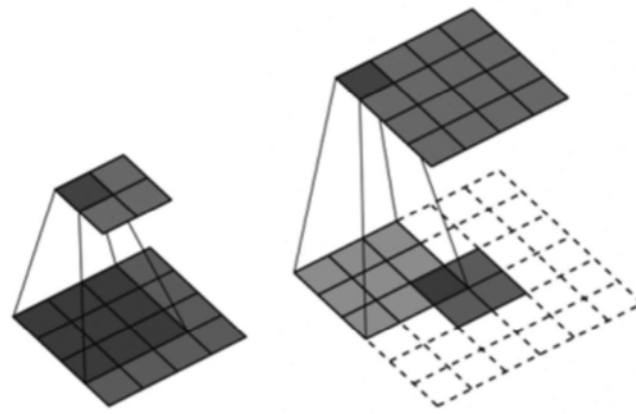


图8



(a) 卷积

(b) 反卷积

图9

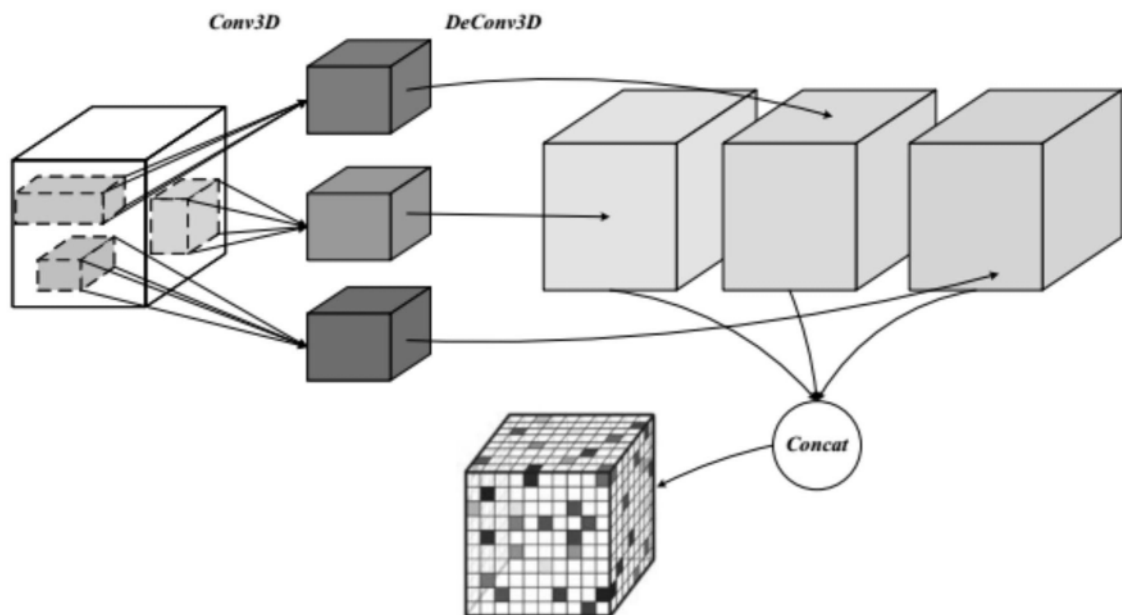


图10

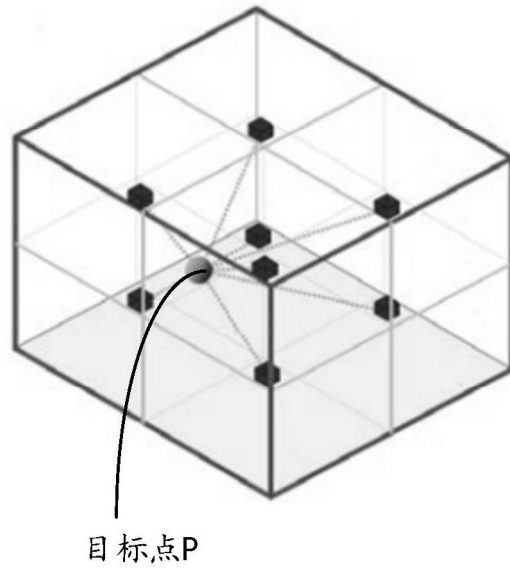


图11

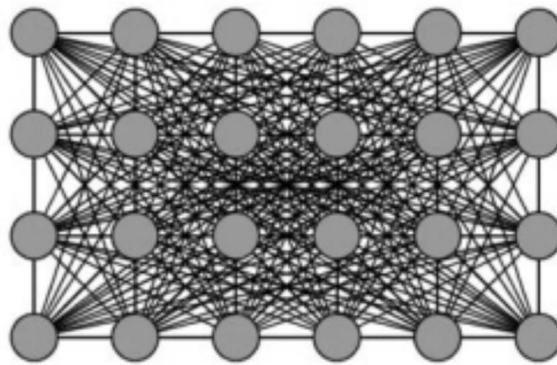


图12

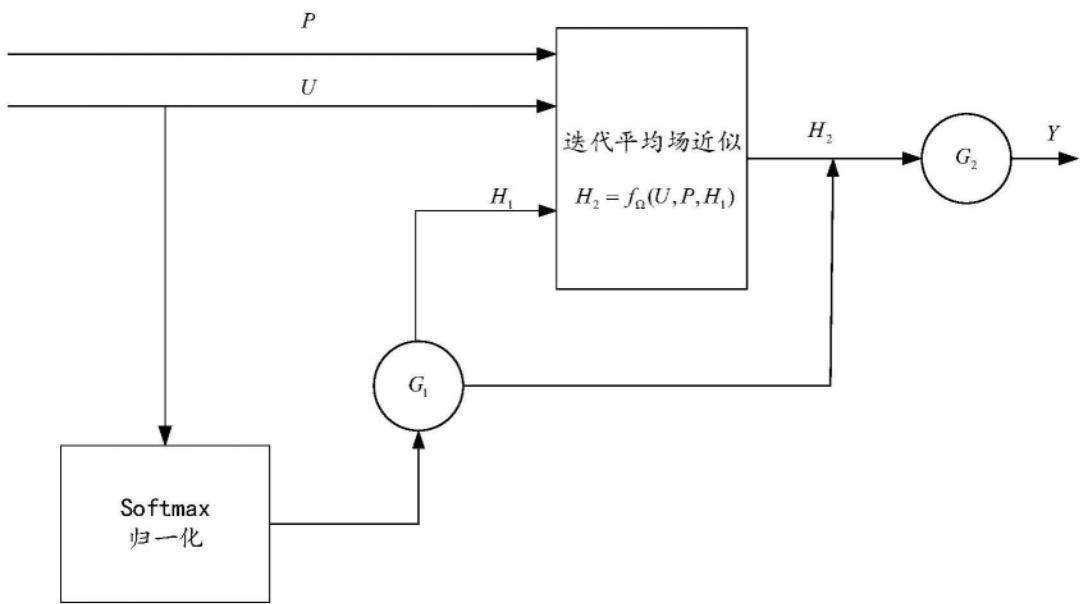


图13

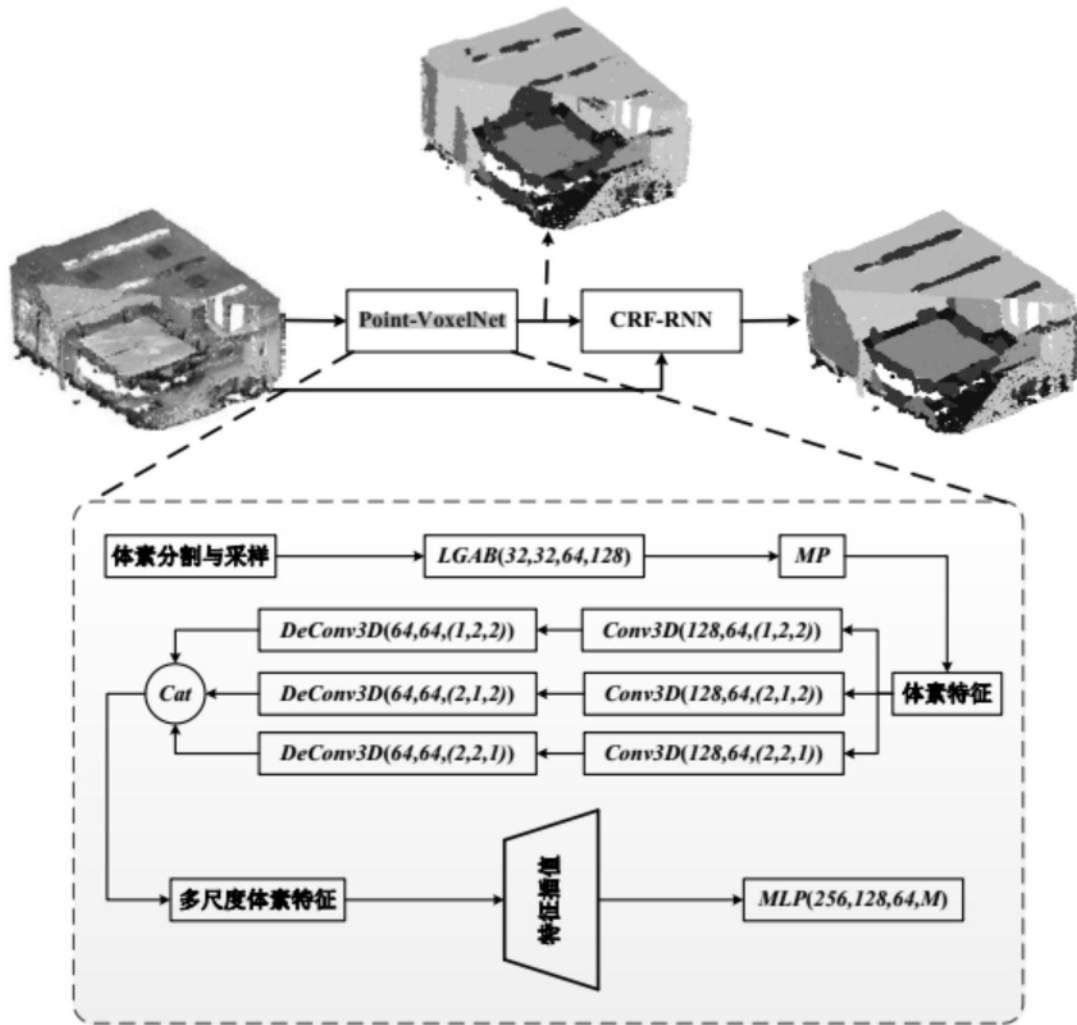


图14

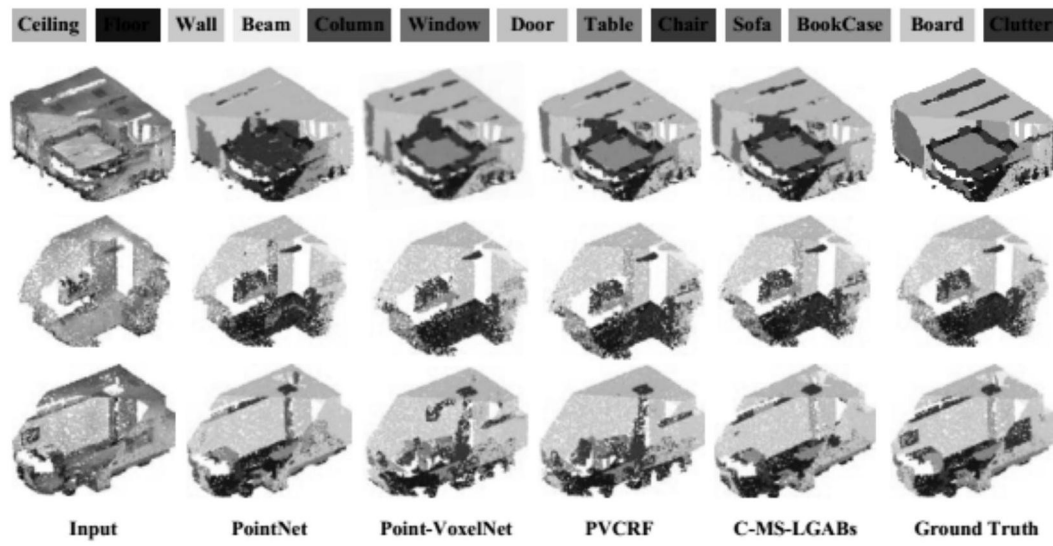


图15

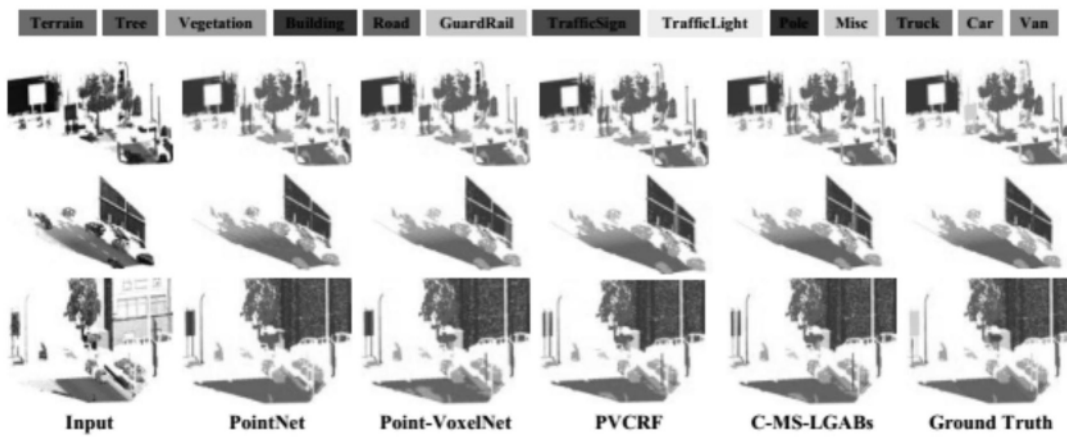
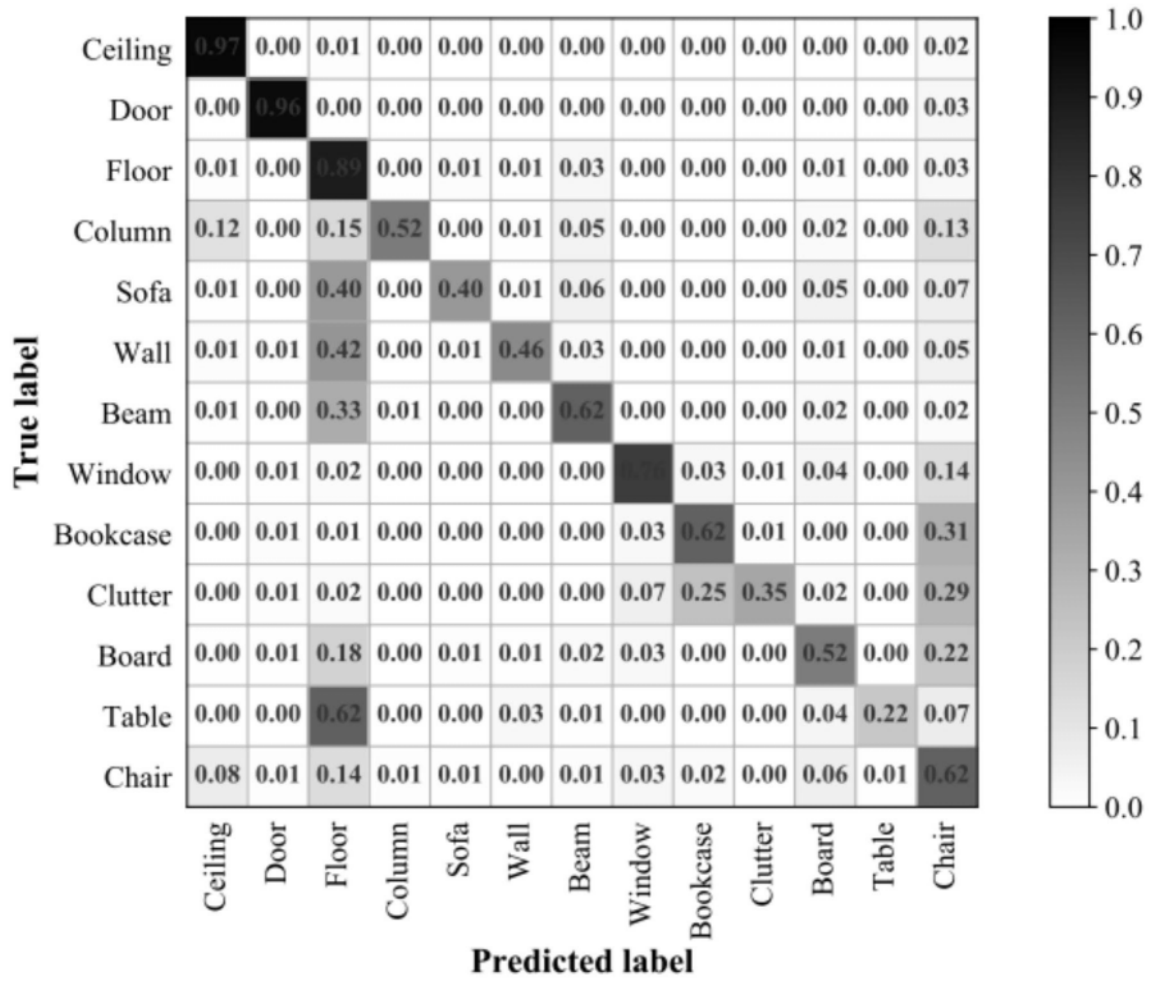
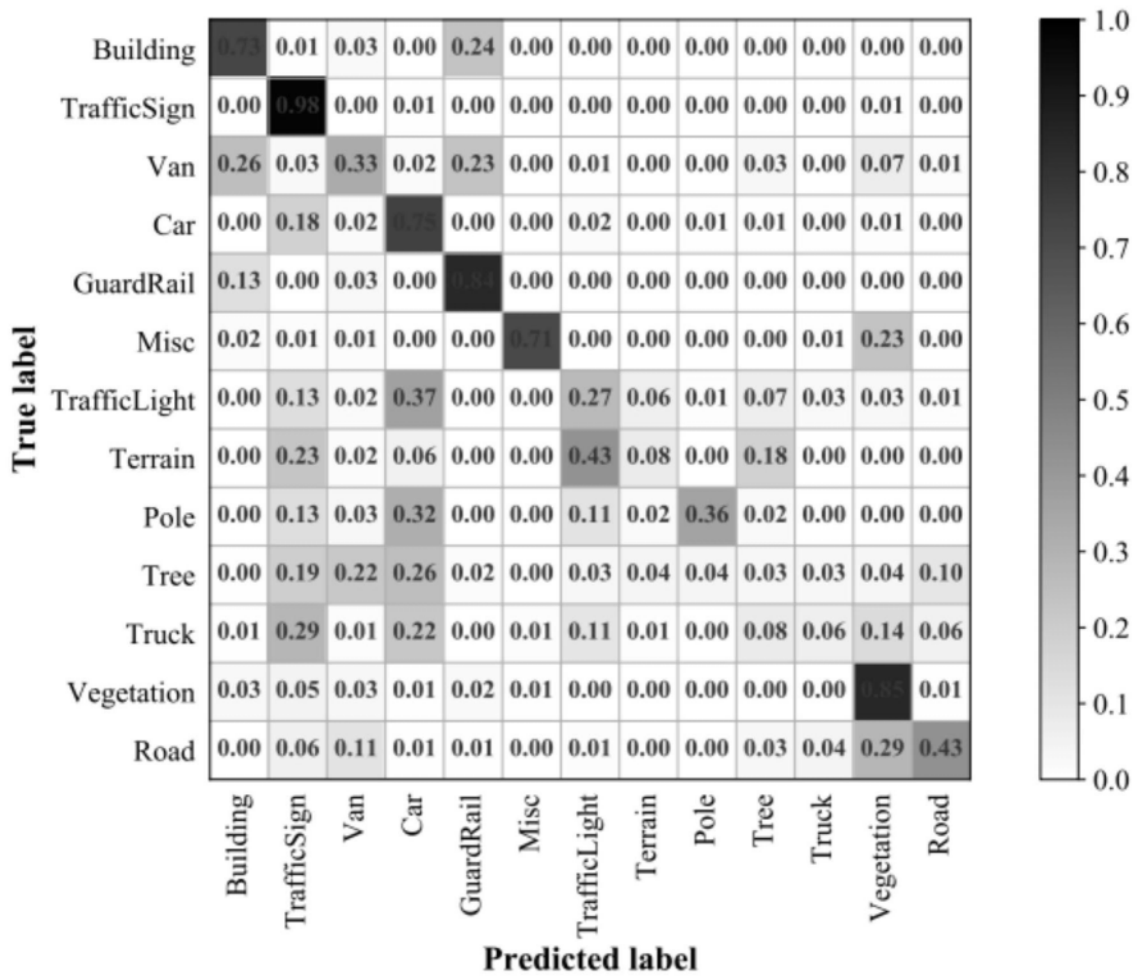


图16

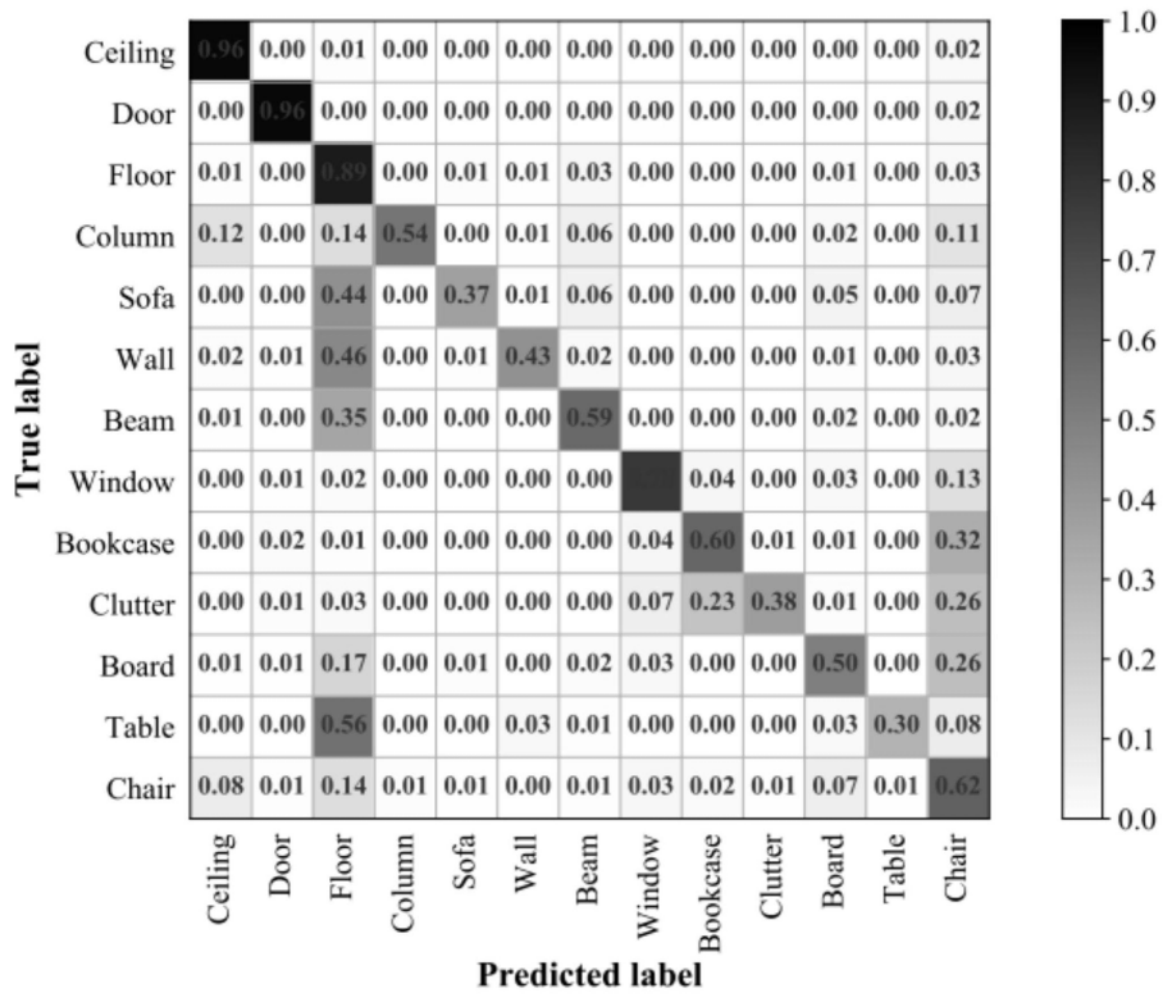


(a) S3DIS

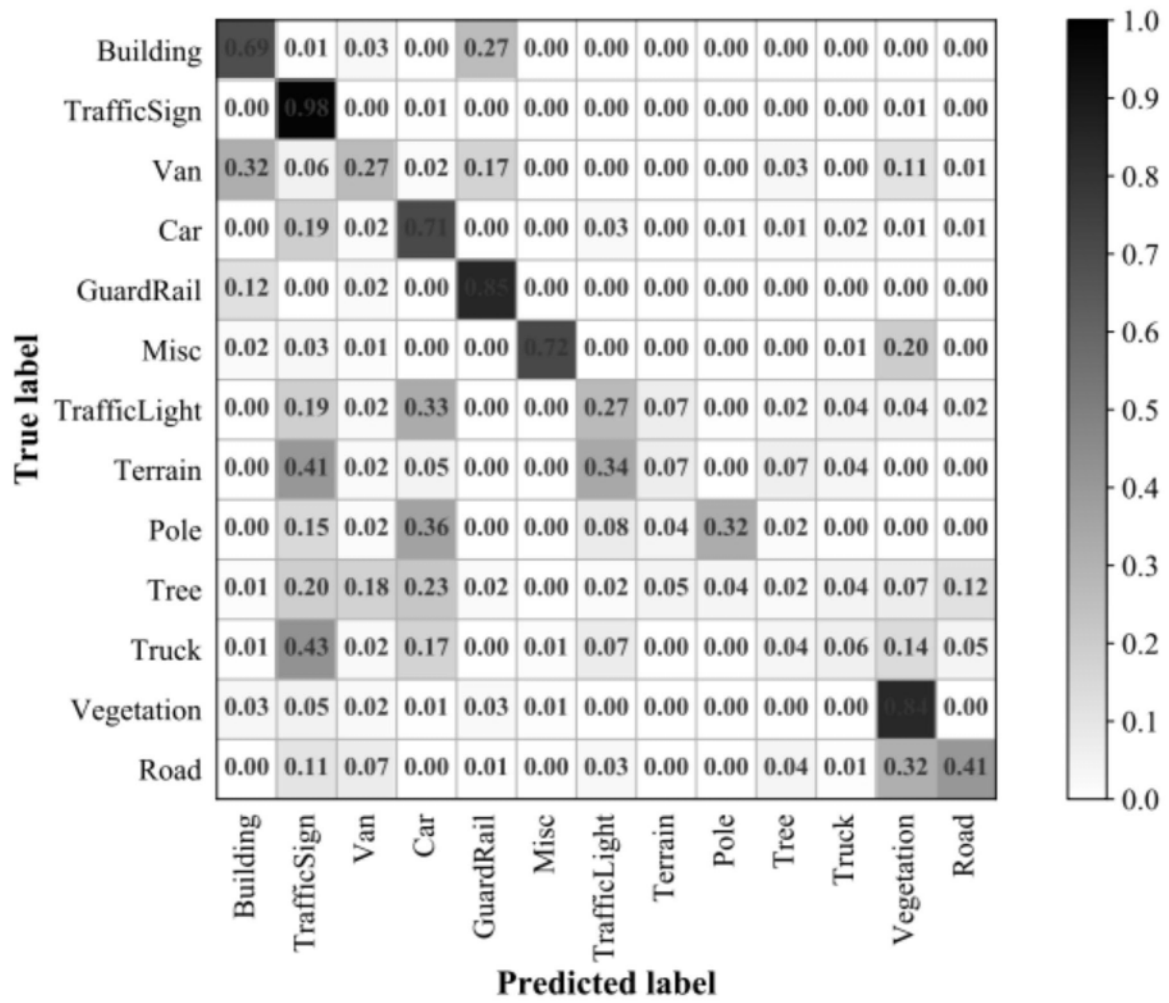


(b) vKITTI

图17



(a) S3DIS



(b) vKITTI

图18

100



图19