

点云稀疏编码三维模型簇协同分割

杨军^{1*}, 李东浩²

¹兰州交通大学电子与信息工程学院, 甘肃 兰州 730070;

²兰州交通大学自动化与电气工程学院, 甘肃 兰州 730070

摘要 为了在函数空间内将多个三维模型进行关联,并在整个模型簇上进行协同分割,提出了一种基于点云稀疏编码的三维模型簇协同分割方法。首先,提取点云数据特征,将三维信息转换至特征空间;其次,用深度学习网络将特征向量分解成基向量,并构建字典矩阵及稀疏向量;最后,对测试数据进行稀疏表示,并确定点云模型中每个点所属的类别,将同类点划分到同一区域以得到协同分割结果。实验结果表明,算法在 ShapeNet Parts 数据集上的分割准确率达到 85.7%。所构建的协同分割算法能够有效地计算模型簇的关联结构,与当前主流分割算法相比,分割效果和准确率均得到提升。

关键词 机器视觉; 协同分割; 模型簇; 稀疏编码; 基向量; 字典

中图分类号 TP391 **文献标志码** A

doi: 10.3788/LOP57.201510

Co-Segmentation of 3D Model Clusters Based on Point Cloud Sparse Coding

Yang Jun^{1*}, Li Donghao²

¹School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China;

²School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China

Abstract Aiming at the problems of the co-analysis of multiple 3D models in the function space and the co-segmentation of the whole model cluster, we propose a co-segmentation method based on point cloud sparse coding. First, the point cloud feature is extracted and the 3D information is transformed into the feature space. Second, the dictionary matrix and sparse vectors are constructed after the decomposition of the feature vectors into the base vectors by the deep learning network. Finally, the test data is represented by dictionary sparseness and the category of each point in the point cloud model is determined. To get the co-segmentation result, the homogeneous points are divided into the same region. The experimental results show that the segmentation accuracy on ShapeNet Parts dataset obtained using the proposed algorithm is 85.7%. Compared to the current mainstream algorithms used for segmentation, the proposed algorithm can not only compute the relational structure of model clusters more effectively, but also improve the segmentation accuracy and effect.

Key words machine vision; co-segmentation; shape clusters; sparse coding; base vector; dictionary

OCIS codes 150.0155; 100.6890; 100.4996

1 引言

三维模型分割是数字几何处理与建模中的基础研究问题,在计算机图形学与计算机视觉领域有广泛应用。现有的三维模型分割方法主要是基于部件的分割,即将一个完整的三维目标分解成若干语义

部件^[1]。随着深度学习技术的发展,许多前沿方法已被广泛应用于三维模型分割领域。例如文献[2]提出同步光谱卷积神经网络(Synchronized Spectral Convolutional Neural Network, SSCNN),该方法在由拉普拉斯特征基张成的谱域空间中,通过参数化卷积核来实现权值共享。文献[3]提出 PointCNN

收稿日期: 2020-02-02; 修回日期: 2020-03-05; 录用日期: 2020-03-09

基金项目: 国家自然科学基金(61862039)

* E-mail: yangj@mail.lzjtu.cn

架构,可将卷积神经网络结构用于处理无序的点云数据,能有效避免点的扰动,提高了模型处理效果。文献[4]提出模型全卷积网络(Shape Fully Convolutional Networks, SFCN),该框架能将三维模型表示成图结构,因此可使用卷积和池化操作直接处理三维数据。目前,基于深度学习的分割方法需要大量的带有标签的数据作为网络训练的数据集,然而三维模型数据标记是一项费时费力的工作。

三维模型簇的协同分割旨在以无监督方式同时分割一组有共同特征且拓扑结构相似的模型集合。集合的共同特征与模型所属的类别相关。对协同性的分析是实现模型簇协同分割的重点步骤。为实现模型簇之间的协同分析,需要在整个模型簇的函数空间中寻找模型共有的潜在特征。然而,三维模型不同于二维图像,每个三维模型都有其独立的函数空间。因此,将这些函数空间连接起来进行协同分析是非常困难的。传统方法需利用流形谐波的思想来解决这一问题。首先,对源模型与目标模型进行拉普拉斯特征分解,分别计算出两组模型的基函数;然后,将两组基函数关联起来,并转换基函数以同步模型信息。然而,在处理过程中,先计算对应关系再进行分割将会耗费巨大的计算成本。

近年来,稀疏编码(Sparse Coding)在计算机视觉和机器学习领域都取得了一定成果。稀疏编码提供一种新的信号描述方式,相比于传统采样理论,可以从更少的采样数据中重建出原始信号,从而减少采样数据,只保留包含重要信息的采样数据^[5]。

受上述方法启发,本文提出一种基于点云稀疏编码的三维模型簇协同分割网络。深度学习网络先将每个三维模型的函数空间分解成一个基函数集,集合中的每个基函数都反映了模型的点特征。网络训练所生成的基函数可表示该类模型的共同特征。这些基函数的线性组合不仅能重构样本信息,而且能够传递一致性因素,从而实现模型间的协同分析。

2 相关研究工作

三维模型分割的目的是将一个完整的三维模型分割成若干语义部件,其在模型的理解和分析等方面起着重要作用。目前三维模型分割方法可分为两类:基于几何的方法和基于学习的方法。

基于几何的方法是根据三维模型表面的几何性质,并利用一些分解技术将完整模型分割成部件。常见的基于几何的方法有分层聚类分割(Hierarchical Clustering, HC)^[1]、区域增长法分割

(Region Growing, RG)^[6]、基于分水岭的分割(Watershed Segmentation, WS)^[7]、迭代聚类分割(Iterative Clustering, IC)^[8]及基于边界的分割(Boundary Segmentation, BS)^[9]等。这些方法通常适用于聚合三维模型,模型表面的网格面有相似的几何性质,且分割的效果由特征描述符来控制。然而,这些描述符提取的特征属于几何层面的低级特征,无法提取具有语义性质的高级特征,而且描述符的设计要依靠经验丰富的专家,所以该类方法的适用场景受到限制。

随着深度学习技术在三维模型分析和处理领域的应用,涌现出很多基于学习的模型分割技术。基于学习的分割技术又可分为基于特征的分割和基于数据的分割。

基于特征的分割算法从原始输入数据中提取低级或中级特征,并将这些特征输入到算法中进行学习,而不是将原始数据本身作为学习过程的输入。如在文献[10]中,算法先对三维网格模型进行分割,并提取低级特征;然后将低级特征输入到自编码器中进行学习,其输出是已训练好的高级特征;最后用高级特征进行聚类运算以得到分割结果。该方法的分割效果相比于传统的基于几何的方法有所改善,但是提取低级特征所用到的特征描述符依赖于丰富的先验知识。

基于数据的分割是指将原始数据本身直接输入到算法中,算法在学习时会自动从原始数据中提取三维模型的特征信息,无需再用描述符提取特征。由于三维模型特征高度不规则,因此无法与二维图像一样直接将原始数据输入到学习算法中。文献[11]提出了三维模型体素化的方法,将由三角网格或其他边界形式组成的模型转化为离散的体素集合,这样便可以使用三维卷积神经网络处理体素模型。文献[12]提出了投影卷积神经网络(Projective Convolutional Neural Network, PCNN),对三维模型进行二维化表示,即在多个视角下对三维模型进行拍照渲染,得到该模型的多视角位姿图像,并将这些位姿图像直接输入到已有的图像处理框架中以分割图像内容,如VGG^[13]、AlexNet^[14]等,最后将特征反投影到三维模型表面以得到分割结果。文献[15]改进了投影卷积神经网络,用多视图卷积神经网络(Multi-view Convolutional Neural Network, MVCNN)从对应关系中学习局部描述符。该描述符是从大量数据中拟合得到的,因此泛化能力强,相比于手工设计的描述符,其不受模型类别的限制,应

用场景更加广泛。近年来由于深度传感器及三维扫描仪的普及,三维点云数据的获取越来越容易^[16],很多应用开始侧重于点云数据,如三维点云场景的分类^[17]。而在模型分割领域,点云数据也具有一定优势,如文献[18]提出可直接处理三维点云数据的深度学习网络架构。该方法取消了三维模型数据变换过程,所以可保留更多的三维特征,在点云分割方面取得了很好的效果。文献[19]提出了基于平均曲率的局部特征权值,并通过局部特征权值和曲率极值法对特征点进行双重检测,该方法既能够提取点云模型中的显著特征和尖锐特征,又能够提取特征强度变化的曲线特征。

以上方法都是将分割表示成有监督的特征聚类问题,并且只针对单体模型进行分割,然而对于复杂的模型部件,单体模型无法提供更多的几何信息,而且监督分割方法需要大量的手动标记分割数据,这限制了该类方法的应用。文献[20]首次提出了协同分割的思想,旨在对一组具有某些共同特征的模型进行一致分割,主要探究整个模型集合的结构理解。此方法利用模型排列方式,在模型间建立对应关系并传播相关性信息以实现协同分割。文献[21-24]都将协同分割看作是描述符空间内的聚类问题。这些方法先对模型进行过分割,然后对过分割的部件行特征进行镶嵌和聚类处理,得到一致分割的结果。虽然以上方法都用无监督的方式进行训练,但分析过程均采用手工设计的描述符和基于启发式的聚类方法,当模型结构过于复杂时通常会出现不准确的分割结果。在文献[25]中,研究者提出一种标签驱动协同分割方法,主要用于分割出用户指定标记的模型区域,可以很精准地分割出结果,但是,该方法需要使用者明确指定出特殊标记,适用场景有限。

综上所述,有监督协同分割方法虽然能够获得较好的分割效果,但训练过程过度依赖一致分割数据集,而无监督协同分割中的部分技术仍需要通过手动设计描述符来提取模型特征,或通过模型间构建对应关系来寻找相关性,增加了计算成本和算法的复杂性。因此本文在文献[18]的基础上,结合稀疏编码的思想,通过建立三维模型字典矩阵来稀疏表示原始数据,从而保留模型的关键信息,去除冗余特征,改善模型的分割效果。

3 稀疏编码

点云模型分割可被看成是逐点分类问题,将三维模型某一部件的所有采样点划分到同一类别即可

得到该部件的分割结果,而且协同分割的关键任务在于挖掘模型集合的特征相关性。因此,本文提出用稀疏编码来解决上述问题,在保留模型关键特征的同时滤除冗余特征,并通过训练字典以学习相似特征的共性,最后通过稀疏编码分类过程对特征进行分类。

稀疏表示分类方法的关键在于字典 \mathbf{D} 的选取,字典所包含的判别性越强,分类效果就会越好。假设训练样本有 m 个类别,记为 $A = \{A_1, A_2, \dots, A_m\}$,其中, A 为训练集合, A_m 为训练集中第 m 类样本。在学习时,针对 m 类数据的每个类别分别训练一个子字典,最后将各个子字典组合在一起以生成一个全局字典。学习后得到的字典 $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_i, \dots, \mathbf{D}_m]$,其中 \mathbf{D}_i 表示第 i 个类别的子字典。该类方法对应的目标函数为

$$\underset{\mathbf{D}, \boldsymbol{\alpha}}{\operatorname{argmin}} \left(\sum_{i=1}^m \|\mathbf{x}_i - \mathbf{D}_i \boldsymbol{\alpha}_i\|_2^2 \right) + \lambda_1 \|\boldsymbol{\alpha}_i\|_1 + \lambda_2 \|\mathbf{g}_D\|_1, \quad (1)$$

式中:下标 1 和 2 表示 L_1 和 L_2 范数; \mathbf{g}_D 表示在字典学习过程中对字典 \mathbf{D} 所添加的约束,其可增强字典的判别性;参数 λ_1 和 λ_2 为正则化超参数,以防止训练过程中的网络过拟合; \mathbf{x}_i 是输入的第 i 类样本数据; $\boldsymbol{\alpha}_i$ 为需要计算的第 i 类稀疏矩阵; $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_i, \dots, \boldsymbol{\alpha}_m]$ 。在分类阶段,利用重建误差的分类策略实现分类,将稀疏分解后的系数作为特征输入到相应的分类器中,即可实现样本数据的分类^[5]。

4 点云稀疏编码的协同分割网络

每个三维模型都有其独立的函数空间,为关联相互独立的函数以实现协同分析,本文将点云数据的特征提取框架和稀疏编码相结合,提出了基于点云稀疏编码的协同分割网络,以进一步提升模型分割性能。

4.1 网络结构

本文提出的网络结构分为训练阶段和测试阶段。训练阶段主要用于构建模型字典,并针对每一类模型来分别训练网络结构。如图 1 所示,首先提取每一类点云模型的特征,将三维信息映射到特征空间,得到该类模型集中每个模型的特征向量;然后,将同一类模型的特征向量训练成一个子字典,该子字典可作为此类模型的特征集,最后将所有子字典融合成一个全局字典。测试阶段则需要用全局字典重构出模型,先提取测试模型的特征,获取特征向

量;然后用测试模型的特征和全局字典共同计算出测试模型的稀疏向量,再用该向量和全局字典重构出测试模型;最后,通过损失函数计算重构模型与测试模型的残差值,用残差结果实现逐点分类并得到分割结果。

文中构建的深度学习网络是对同一类模型的多个样本同时进行学习,所以网络将模型分解成基向

量后,每个基向量都包含了该类模型的拓扑信息。携带拓扑信息的基向量不再依赖函数映射来完成信息的传递与共享。构建字典时,基向量会在字典中进行线性组合,字典在训练过程中会不断调整基向量的组合并更新优化向量内系数。基向量的组合优化过程就是模型信息融合与共享过程,网络协同分析的特性也由此体现。

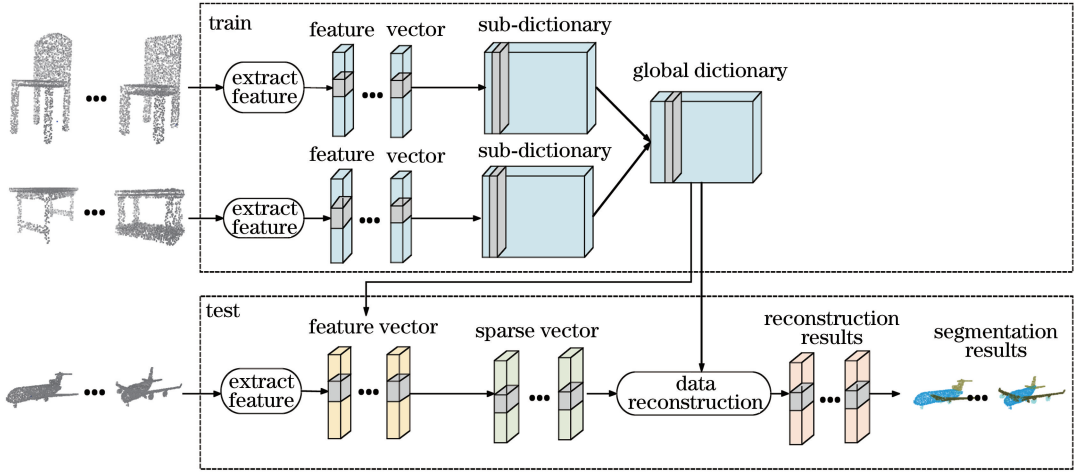


图1 点云稀疏编码协同分割网络结构

Fig. 1 Structure of co-segmentation network for point cloud sparse coding

4.2 构建模型字典

点云模型由一系列无序点组成,包含三维模型丰富的几何信息,因此在提取特征时可直接对点集进行采样。文献[18]提出将点集中的每个点映射到函数空间,使得每个点都被看作为逐点函数,逐点函数的组合便可表达语义信息。因此本文将采用文献[18]中的特征提取策略。

为了使构建的字典矩阵更有判别性,本文在构造过程中没有直接对点云数据进行稀疏编码,而是先提取点云特征,并将点集映射到特征空间中,再对特征进行稀疏编码。给定一个三维模型集合 $X = \{X_1, X_2, \dots, X_i, \dots, X_m\}$, 假设该集合内有 m 个模型类别,每个模型表面有 n 个采样点,经特征提取后,所有点集都被编码了特征信息,其中每类模型都有一个与之对应的特征函数 f_i 。为学习特征之间的相关性,需构建一个函数字典矩阵 $D(X, \theta)$, 其中 θ 为神经网络的权重值。字典中的词条称之为原子,在字典矩阵 $D(X, \theta)$ 中以列向量表示,其中原子数量为 k ,原子的维度大小即采样点个数为 n 。如图2所示,左边的模型是将点云数据映射到特征空间后所生成的特征模型,以特征数据作为输入来训练字典。右边的向量 α_i 是 f_i 的稀疏向量。因为 f_i 包含的是稠密数据,其中部分特征在分割时是冗

余特征,所以要对 f_i 进行稀疏化处理。令 $f_i = D_i \alpha_i$,即用字典和稀疏向量的乘积逼近输入数据。但在实际编码时,无法使 f_i 和 $D_i \alpha_i$ 相等,因此将问题转化为

$$L[D(X, \theta), f] = \min_{\alpha, D} \left[\sum_{i=1}^m \|D_i(X_i, \theta)\alpha_i - f_i\|_2^2 \right] + \lambda \|\alpha_i\|_1, \quad (2)$$

式中: $L[D(X, \theta), f]$ 为损失函数; f 为特征函数; X_i 表示第*i*类输入模型(包含*n*个采样点); f_i 是 X_i 的特征函数; λ 为正 regularization 超参数。将字典矩阵嵌入到损失函数中,训练时同步更新网络权重和字典系数。为了使字典与稀疏向量的乘积逼近函数 f_i ,要降低 f_i 和 $D_i \alpha_i$ 之间的差值。在(2)式中,采用 L_1 范数对参数 α_i 进行约束,以获得稀疏解并确保字典的稀疏性。

对(2)式损失函数的优化过程可分为内部优化和外部优化。在损失函数内部,优化稀疏向量 α ;在损失函数外部,更新网络自身权重值来获取最佳的字典矩阵。由于存在约束条件,优化损失函数不会有解析解,因此无法用梯度下降法计算(2)式中 $L[D(X, \theta), f]$ 的梯度值。本文将采用交替迭代优化算法,交替更改字典 $D(X, \theta)$ 和稀疏向量 α 的值,从

而使目标函数达到最小。每次迭代分两步:1)在损失函数内部,固定字典 $\mathbf{D}(X, \theta)$, 优化向量 α ; 2)在损失函数外部,固定向量 α , 优化字典 $\mathbf{D}(X, \theta)$ 。不断迭代上述过程,直至目标函数得以收敛。参数优化过程为

$$\alpha^t = \underset{\alpha}{\operatorname{argmin}} \sum_{i=1}^m (\| \mathbf{D}_i^t \alpha_i - f_i \|_2^2), \text{ s. t. } C(\alpha), \quad (3)$$

$$\theta^{t+1} = \theta^t - \eta \nabla L[\mathbf{D}(X, \theta^t), f, \alpha^t], \quad (4)$$

式中: α^t 为 t 时刻的稀疏矩阵; \mathbf{D}_i^t 为 t 时刻第 i 类字典矩阵; θ^t 为 t 时刻的网络参数; $C(\alpha)$ 为稀疏向量 α 的约束条件; η 为学习率; ∇ 为那勃勒算符。

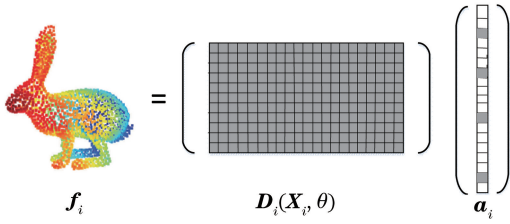


图2 模型字典矩阵的构建过程

Fig. 2 Process for construction of model dictionary matrix

4.3 协同分割约束条件

在构造字典时,为了使字典具有更强的判别性,字典要更加完备,原子数目要更多。当对输入数据进行稀疏表示时,每次只抽取包含当前重构数据关键特征的原子,其余原子被滤除。为满足分割需求,要对字典和稀疏向量加以约束。

训练时,针对每类模型均学习一个子字典,且字典中的原子均是一个归一化向量。因此要将 $\mathbf{D}(X, \theta)$ 和 α 的值都约束在 $[0, 1]$ 范围内,对 $\mathbf{D}(X, \theta)$ 中每一列都增加约束条件。约束条件为

$$\begin{cases} 0 \leq |\alpha| \leq 1 \\ 0 \leq |\mathbf{D}(X, \theta)| \leq 1 \\ \sum_j D(X, \theta)_{s,j} = 1 \text{ for all } s \end{cases}, \quad (5)$$

式中: $D(X, \theta)_{s,j}$ 是字典矩阵中第 s 行第 j 列元素。第一项关于 α 的约束用于内部优化;第二项和第三项约束用于外部优化,在更新网络权重的同时对 α 和 $\mathbf{D}(X, \theta)$ 进行约束。

4.4 计算残差

当字典构造完成后,类别信息已被编码进入字典,字典有较强的判别性。假设给定测试数据 y , 先对 y 提取特征得到 h , 可利用(4)式计算得到 h 的稀疏向量 α' , 进而计算二者的残差值为

$$\begin{cases} e_i = \| h - \mathbf{D}\alpha'_i \|_2 \\ \alpha' = [\alpha'_1, \alpha'_2, \dots, \alpha'_m] \end{cases}, \quad (6)$$

式中: e_i 为残差值; α'_i 是字典中第 i 类样本的稀疏向量。根据残差 e_i 确定分类策略:将测试样本归为具有最小重构误差的那一类。这样模型中的每个点都被归到相应的部件类别中,当所有点归类完成后,模型的分割结果也相应产生。

5 实验结果与分析

本实验的软件环境为 Linux Ubuntu 16.04 操作系统,硬件为 Intel Xeon E5-2620 v4 CPU 和 NVIDIA Quadro M4000 GPU(8G 显存)处理器,采用的深度学习框架为 Tensorflow,基于该框架的深度学习运算平台为 CUDA 8.0, GPU 加速库为 cuDNN 6.0。

5.1 实验数据集

本研究采用 ShapeNet Parts^[26] 数据集,该数据集包含 16881 个三维模型,16 个类别,共 50 个带有标签的部件,大多数模型都被标记为 2~5 个部分,且标签都被注释在三维模型的采样点上。

5.2 实验参数设置

为了提升分割准确率,从模型表面均匀采样 2048 个点用于训练和测试实验。本文使用自适应矩估计优化器(Adaptive Moment Estimation, Adam)来优化神经网络模型。设置初始学习率为 0.001,以较大学习率快速得到一个最优解。将动量设为 0.9,避免网络更新参数时陷入局部最优。为了防止训练时网络中间层数据分布不一致,在网络每一层输入前插入一个批归一化层(Batch Normalization),批归一化处理的指数衰减系数设置为 0.5,使损失函数得以快速收敛。在本实验中,当训练批尺寸(Batch Size)设为 32 时,训练误差对网络梯度值的影响最小。

5.3 评估标准

三维模型部件分割的准确率以点集上的交并比(Intersection over Union, IoU)来度量,其为模型真实分割部件和预测分割部件的交集与并集之比。对于整个模型簇,每一类模型的 IoU 值为

$$I_{\text{IoU}} = \frac{p_{uu}}{p_{uv} + p_{vu} - p_{uu}}, \quad (7)$$

式中: u 表示真实值; v 表示预测值; p_{uu} 表示本属于 u 且预测为 u 的概率; p_{uv} 表示本属于 u 却预测为 v 的概率; p_{vu} 表示本属于 v 却预测为 u 的概率。对于某类模型中的每个部件,先计算部件的 IoU,再对该类别中的所有部件的 IoU 进行平均化处理以得到一个完整模型的 IoU 值。

获得每类模型的 IoU 值后,为了评估模型的整体分割效果,需采用平均交并比(mean Intersection over Union, mIoU)来进行评判:

$$m_{\text{IoU}} = \frac{1}{m+1} \sum_{u=0}^m \frac{p_{uu}}{\sum_{v=0}^m p_{uv} + \sum_{v=0}^m (p_{vu} - p_{uu})}, \quad (8)$$

式中: m 表示共有 m 个类别。mIoU 值越大,表明真实值与测试值越接近,分割效果越好。

在数据集中,由于每类模型的部件数量均不一致,因此,需从定量的角度继续对分割准确率进行评判,故引入部件平均交并比(part averaged IoU, pIoU):

$$p_{\text{IoU}} = \frac{1}{\sum_{u=0}^m \sum_{v=0}^m p_{uv}} \times \left[\sum_{u=0}^m \frac{p_{uu} \sum_{v=0}^m p_{uv}}{\sum_{v=0}^m p_{uv} + \sum_{v=0}^m (p_{vu} - p_{uu})} \right]. \quad (9)$$

(9)式表示根据每一类模型出现的频率对各类的 IoU 进行加权求和,本实验中的频率即为每类模型的数量。该指标更侧重于对小部件的预测结果,pIoU 值越高,表明分割准确率越高。

5.4 分割准确率的对比

本文所提算法对三维点云模型簇的协同分割结果如图 3 所示,其中不同颜色的点云表示不同的分割部件。可以看出,每一类模型的分割结果是一致的。例如吉他模型,图 3 中的六把吉他都被分割为三个部件:琴头、琴身和琴颈,而且每把吉他中对应部件的颜色一致。本算法在实现点云模型簇协同分割方面取得了较好的效果。

算法在 ShapeNet Parts 数据集上与文献[18]、文献[25]、文献[27]的对比实验结果如表 1 所示。本算法用稀疏编码提高逐点分类能力,相比于文献[18]将局部特征与全局特征相融合的分类策略,本算法的 pIoU 提升了 2.1%,mIoU 提升了 0.6%。同时本算法构建的字典基函数中被编码了一致性信息,相比于文献[25]采用平均池化策略获取模型一致性信息,pIoU 提高了 2.6%,mIoU 提高了 1.1%。文献[27]利用多尺度分组和多分辨率分组提高局部特征的提取能力,而本文则使用字典的分类能力提升点特征的判别性,使 pIoU 提高了 0.6%,但 mIoU 降低了 1.3%,这是因为文献[27]中背包、帽子和摩托的分割准确率均高于本算法,其余类别则与本文接近,因此平均分割准确率较高。

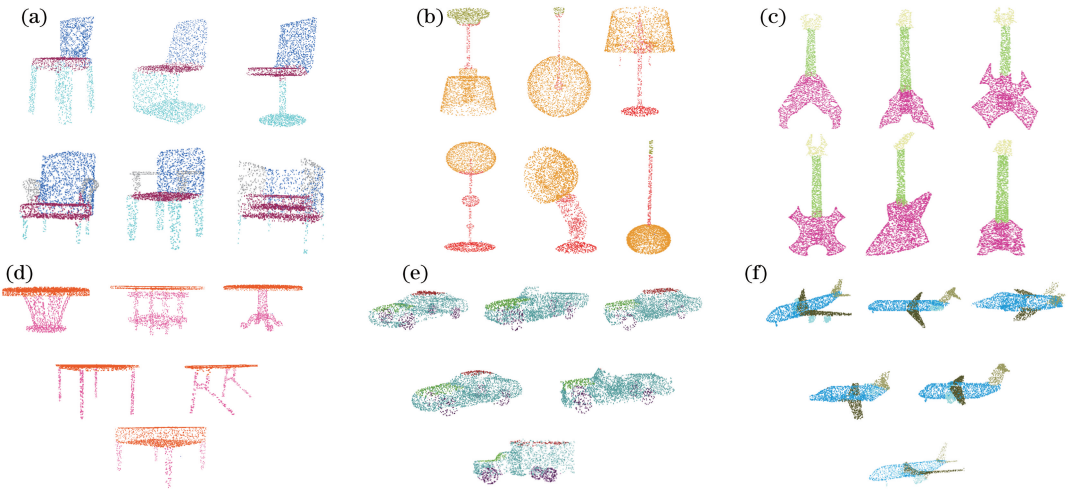


图 3 点云模型的分割结果。(a)椅子;(b)台灯;(c)吉他;(d)桌子;(e)汽车;(f)飞机

Fig. 3 Segmentation results of point cloud models. (a) Chair; (b) lamp; (c) guitar; (d) table; (e) car; (f) plane

表 1 不同算法的分割准确率

Table 1 Segmentation accuracy of each algorithm

Algorithm	mIoU	pIoU / %
PointNet ^[18]	80.4	83.6
Tags2Parts ^[25]	79.5	83.1
PointNet++ ^[27]	81.9	85.1
Ours	80.6	85.7

不同算法在各类模型上的分割准确率如表 2 所示。与文献[18]相比,本文算法对滑板和桌子模型的分割效果得到显著提升,IoU 分别提高了 7.3% 和 9.4%。文献[25]的方法侧重于分割用户指定标记的模型区域,属于一种强监督学习算法,因此对细小部件的分割准确率高,如摩托车的准确率高于本

算法 4.3%。文献[27]可根据点云不同的密度采用不同的分辨率提取特征,其在点云密度变化明显的

模型上的分割准确率高於本算法,如手提包和帽子模型,分别高於本算法 6.7%和 9.9%。

表 2 不同算法在各类模型上的 pIoU

Table 2 pIoU of each algorithm on various models

Algorithm	Airplane	Bag	Cap	Car	Chair	Earphone	Guitar	Knife
PointNet ^[18]	83.2	80.4	86.1	76.6	89.5	70.8	91.2	85.3
Tags2Parts ^[25]	81.4	83.2	75.3	74.8	89.0	71.7	89.2	84.2
PointNet++ ^[27]	82.4	79.0	87.7	77.3	90.8	71.8	91.0	85.9
Ours	81.9	72.3	77.8	76.7	88.4	76.6	91.1	85.4
Algorithm	Lamp	Laptop	Motorbike	Mug	Pistol	Rocket	Skateboard	Table
PointNet ^[18]	80.4	95.3	64.5	91.8	81.3	61.3	72.8	80.4
Tags2Parts ^[25]	75.9	93.7	69.3	94.2	81.4	54.3	73.4	80.9
PointNet++ ^[27]	83.7	95.3	71.6	94.1	81.3	58.7	76.4	82.6
Ours	77.7	95.5	65.0	91.4	80.6	58.9	80.1	89.8

5.5 字典原子重构准确率分析

本算法在所有模型上的 pIoU 为 85.7%,标签索引与字典原子在所有分割部件上的 pIoU 为 78.1,相差 7.6%。为了判断字典矩阵中原子重构出模型信息的准确度,本文建立了标签索引和原子索引在每类模型上的 pIoU,如表 3 所示。表 3 中第一行表示本算法(标签)在每类模型上的 pIoU,第二行表示原子

索引在每类模型上的 pIoU,若二者越接近,则表明原子重构模型信息的准确度越高。可以看出,在大多数类别中,上下两行的数值非常接近,表明原子能够准确地表达模型信息。所有模型中只有灯具类模型的 pIoU 为 41.1%,低于其他类别。原因是算法将台灯的底座和支架分割为一体,因此在计算灯具相同部件原子索引的准确度时出现了误差,致使 pIoU 过低。

表 3 原子索引和标签在每类模型上的 pIoU

Table 3 pIoU of atomic index and label on each model

Method	Airplane	Bag	Cap	Car	Chair	Earphone	Guitar	Knife
Label	81.9	72.3	77.8	76.7	88.4	76.6	91.1	85.4
Atomic index	80.4	56.1	62.6	75.8	87.8	61.9	91.1	84.5
Method	Lamp	Laptop	Motorbike	Mug	Pistol	Rocket	Skateboard	Table
Label	77.7	95.5	65.0	91.4	80.6	58.9	80.1	89.8
Atomic index	41.1	94.6	63.2	90.3	80.4	50.3	76.5	78.0

5.6 分割效率的对比

在测试时,以每个模型的分割时间作为分割效率的评判依据。表 4 为四种算法在测试阶段中分割一个模型的时间对比,可以看出,本文的分割效率低于文献[18],高于其余两个算法。由于本文在分割时用稀疏编码对点云特征进行进一步的筛选,故计算量增加,分割效率低于文献[18]。文献[25]侧重于强监督,分割时标签的分配增加了运算时间。而文献[27]采用多尺度分组策略提升局部特征的提取效果,所以需要消耗更多的计算成本。

表 4 不同算法的分割时间

Table 4 Segmentation time of each algorithm

Algorithm	Time /ms
PointNet ^[18]	25.3
Tags2Parts ^[25]	71.4
PointNet++ ^[27]	163.2
Ours	40.7

6 结 论

提出一种基于深度稀疏编码的学习方法,用于解决三维模型协同分割问题。首先,对三维点云模型提取特征,将点云数据映射到函数空间,把每个模型表示为特征函数的形式。然后,对这些特征进行稀疏编码,并计算出字典矩阵和稀疏向量。最后,用字典和稀疏向量重构测试数据,计算测试数据与重构数据间的残差,并利用残差实现点集分类,得出分割结果。在 ShapeNet Parts 数据集上进行了对比实验。实验结果证明,所提的协同分割算法在模型分割方面取得了良好的效果。但所提算法也存在需要改进的地方:一方面,局部特征的数量对字典判别性有重要影响,但提高局部特征的提取能力会耗费大量的计算成本,因此可采用更高效的特征提取框架以改善分割效果;另一方面,构建的模型字典无法更深层次地挖掘模型的细粒度特征,致使部分模型

的分割准确率不高。如何通过不同的体系和层次结构增强神经网络将函数空间分解为原子的能力,这将是今后继续研究的重点。

参 考 文 献

- [1] Zhang J Y, Zheng J M, Wu C L, et al. Variational mesh decomposition [J]. *ACM Transactions on Graphics*, 2012, 31(3): 1-14.
- [2] Yi L, Su H, Guo X W, et al. SyncSpecCNN: synchronized spectral CNN for 3D shape segmentation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6584-6592.
- [3] Li Y, Bu R, Sun M, et al. Pointcnn: convolution on X-transformed points [C] // *Advances in Neural Information Processing Systems*, December 2-8, 2018, Montréal, Canada. New York, USA: NIPS, 2018: 820-830.
- [4] Wang P Y, Gan Y, Shui P P, et al. 3D shape segmentation via shape fully convolutional networks [J]. *Computers & Graphics*, 2018, 70: 128-139.
- [5] Baraniuk R G, Candes E, Elad M, et al. Applications of sparse representation and compressive sensing[J]. *Proceedings of the IEEE*, 2010, 98(6): 906-909.
- [6] Xiao D, Lin H W, Xian C H, et al. CAD mesh model segmentation by clustering[J]. *Computers & Graphics*, 2011, 35(3): 685-691.
- [7] Benjamin W, Polk A W, Vishwanathan S V N, et al. Heat walk: robust salient segmentation of non-rigid shapes [J]. *Computer Graphics Forum*, 2011, 30(7): 2097-2106.
- [8] Yan D M, Wang W P, Liu Y, et al. Variational mesh segmentation via quadric surface fitting [J]. *Computer-Aided Design*, 2012, 44(11): 1072-1082.
- [9] Rodrigues R S V, Morgado J F M, Gomes A J P. A contour-based segmentation algorithm for triangle meshes in 3D space [J]. *Computers & Graphics*, 2015, 49: 24-35.
- [10] Shu Z Y, Qi C W, Xin S Q, et al. Unsupervised 3D shape segmentation and co-segmentation via deep learning [J]. *Computer Aided Geometric Design*, 2016, 43: 39-52.
- [11] Wu Z R, Song S R, Khosla A, et al. 3D ShapeNets: a deep representation for volumetric shapes[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 1912-1920.
- [12] Kalogerakis E, Averkiou M, Maji S, et al. 3D shape segmentation with projective convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6630-6639.
- [13] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [C] // *International Conference on Learning Representations*, May 7-9, 2015, San Diego, CA, USA. New York: IEEE, 2015: 1409-1418.
- [14] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [15] Huang H B, Kalogerakis E, Chaudhuri S, et al. Learning local shape descriptors from part correspondences with multiview convolutional networks[J]. *ACM Transactions on Graphics*, 2018, 37(1): 1-14.
- [16] Li Y, Tong G F, Yang J C, et al. 3D point cloud scene data acquisition and its key technologies for scene understanding [J]. *Laser & Optoelectronics Progress*, 2019, 56(4): 040002.
李勇, 佟国峰, 杨景超, 等. 三维点云场景数据获取及其场景理解关键技术综述[J]. *激光与光电子学进展*, 2019, 56(4): 040002.
- [17] Tong G F, Du X C, Li Y, et al. Three-dimensional point cloud classification of large outdoor scenes based on vertical slice sampling and centroid distance histograms[J]. *Chinese Journal of Lasers*, 2018, 45(10): 1004001.
佟国峰, 杜宪策, 李勇, 等. 基于切片采样和质心距直方图特征的室外大场景三维点云分类[J]. *中国激光*, 2018, 45(10): 1004001.
- [18] Charles R Q, Su H, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 652-660.
- [19] Wang X H, Wu L S, Chen H W, et al. Feature line extraction from a point cloud based on region clustering segmentation [J]. *Acta Optica Sinica*, 2018, 38(11): 1110001.
王晓辉, 吴禄慎, 陈华伟, 等. 基于区域聚类分割的点云特征线提取[J]. *光学学报*, 2018, 38(11):

- 1110001.
- [20] Golovinskiy A, Funkhouser T. Consistent segmentation of 3D models [J]. *Computers & Graphics*, 2009, 33(3): 262-269.
- [21] Yang J, Tian Z H, Li L J, et al. Segmentation of 3D geometric models based on mesh Laplace [J]. *Computer Science*, 2015, 42(5): 295-299.
杨军, 田振华, 李龙杰, 等. 基于网格 Laplace 的三维几何模型分割 [J]. *计算机科学*, 2015, 42(5): 295-299.
- [22] Wang Y H, Asafi S, van Kaick O, et al. Active co-analysis of a set of shapes [J]. *ACM Transactions on Graphics*, 2012, 31(6): 1-10.
- [23] Sidi O, van Kaick O, Kleiman Y, et al. Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering [J]. *ACM Transactions on Graphics*, 2011, 30(6): 1-10.
- [24] Yang J, Zhang P. Three-dimensional shape segmentation by combining topological persistence and heat diffusion theory [J]. *Journal of Image and Graphics*, 2018, 23(6): 887-895.
杨军, 张鹏. 结合拓扑持续性和热扩散理论的 3 维模型分割 [J]. *中国图象图形学报*, 2018, 23(6): 887-895.
- [25] Muralikrishnan S, Kim V G, Chaudhuri S. Tags2Parts: discovering semantic regions from shape tags [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 2926-2935.
- [26] Yi L, Kim V G, Ceylan D, et al. A scalable active framework for region annotation in 3D shape collections [J]. *ACM Transactions on Graphics*, 2016, 35(6): 1-12.
- [27] Qi C R, Yi L, Su H, et al. PointNet++: deep hierarchical feature learning on point sets in a metric space [C] // Advances in Neural Information Processing Systems, December 4-9, 2017, Long Beach, CA, USA. New York: NIPS, 2017: 5099-5108.