

基于多分支结构的点云补全网络

罗开乾^{1,2}, 朱江平^{1,2*}, 周佩^{1,2}, 段智涓^{1,2}, 荆海龙²

¹ 四川大学计算机学院, 四川 成都 610065

² 四川大学视觉合成图形图像技术重点学科实验室, 四川 成都 610065

摘要 点云是一种重要的三维表达方式,在计算机视觉和机器人领域都有着广泛的应用。由于真实应用场景中存在遮挡和采样不均匀等情况,传感器采集的目标物体点云形状往往是不完整的。为了提取点云的特征和补全目标点云,提出了一种基于多分支结构的点云补全网络。编码器从输入信息中提取局部特征和全局特征,解码器中的多分支结构将提取的特征转换成点云,以得到目标物体完整的点云形状。在 ShapeNet 和 KITTI 数据集以及不同残缺比例、不同几何形状的情况下进行实验,结果表明,本方法可以很好地补充目标缺失的点云,得到完整、直观、真实的点云模型。

关键词 图像处理; 形状补全; 深度卷积网络; 多分支结构

中图分类号 TP242 **文献标志码** A

doi: 10.3788/LOP57.241019

Point Cloud Completion Network Based on Multibranch Structure

Luo Kaiqian^{1,2}, Zhu Jiangping^{1,2*}, Zhou Pei^{1,2}, Duan Zhijuan^{1,2}, Jing Hailong²

¹ College of Computer Science, Sichuan University, Chengdu, Sichuan 610065, China;

² National Key Laboratory of Fundamental Science on Synthetic Vision,
Sichuan University, Chengdu, Sichuan 610065, China

Abstract Point cloud is an important three-dimensional expression, and it has a wide range of applications in computer vision and robotics. Due to occlusion and uneven sampling in real application scenarios, the shape of the target object point cloud collected by the sensor is often incomplete. To achieve the point cloud of feature extraction and shape completion, a new point cloud completion network based on the multibranch structure is proposed in this paper. The encoder is primarily responsible for extracting the global and local features from the input information, and the multibranch structure in the decoder is responsible for converting the features to point clouds to obtain the complete point cloud shape of the object. Experiments are conducted using the ShapeNet and KITTI data sets, with different incomplete proportions and geometric shapes. Results show that the method can well supplement the missing point cloud of the target and obtain a complete, intuitive, and true point cloud model.

Key words image processing; shape completion; deep convolution network; multibranches structure

OCIS codes 100.6890; 100.2960; 110.6880

1 引言

点云是一种能描述三维(3D)物体和场景的数据^[1-2],广泛应用于机器视觉、机器人等领域。传感器在真实场景中采集 3D 点云数据时,受相机分辨

率、遮挡等因素的影响,采集到的点云是稀疏、残缺的,不能描述物体真实的几何信息。因此,对缺失大量信息的点云进行补全具有重要意义。

近年来,对 3D 形状补全的方法主要有基于几何的方法、基于对齐的方法和基于学习的方法。基

收稿日期: 2020-05-06; **修回日期:** 2020-05-30; **录用日期:** 2020-06-24

基金项目: 国家自然科学基金(61901287)、四川省重点研发专项(2020YFG0112,20ZDYF3197)、四川省重大科技专项(2019ZDZX0039,2018GZDZX0029)

* **E-mail:** zjp16@scu.edu.cn

于几何的方法利用物体的几何属性,如直线、平面补全物体的三维形状,包括通过插值思想填充不完整区域的曲面重建法^[3-4]、通过识别具有对称性和重复性结构完成补全的对称法^[5-6]等。这些方法假设输入信息具有完整的结构,缺失的部分可以直观推断,对原始数据的要求较高,不适用于真实场景中收集的残缺点云。基于对齐的方法将输入点云与数据库中的完整结构进行匹配,从而获得完整的物体形状。包括在数据库中直接检索对象整体结构特征的直接检索法^[7-8]、通过检索对象几何结构特征并结合输入结构的结构重组法^[9-10]等。这类方法在推理和迭代过程中的优化代价高,不适用于实时任务,对含噪声的数据实用性较差。基于学习的方法通过构造参数化模型学习输入点云的特征^[11-15],然后将其映射到输出点云的完整形状,该方法的泛化程度^[16-17]远远优于前两种方法。

点云补全网络(PCN)^[18]利用目前最先进的点云处理神经网络 PointNet^[19],直接从点云中提取特征,然后利用解码器补全点云的形状,提取点云特征。但在提取特征的过程中,PCN忽略了对局部结构特征的提取。针对该问题,本文提出一种新的端到端深度神经网络(DNN),其结构类似编码-解码

框架。网络的编码器分别提取输入点云的局部特征和全局特征,解码器是一个由多分支结构网络组成的生成器,结合从编码器中提取的点云特征,可完成对点云形状的补全任务。这种网络结构结合点云的局部特征和全局特征学习输入点云的信息,使DNN对点云特征的学习效果更好。

2 基本原理

实验使用的多分支结构三维点云补全网络如图1所示,其中,Conv为卷积操作,S和G分别表示采样和分组操作,MLP为多层感知机,[]中的数字为特征向量的维度,DNN的输入为原始的残缺点云。首先,用特征提取块提取点云的局部特征和全局特征;然后,用解码器将两者结合起来以补全点云。编码器是一个分层网络,由4个特征提取块提取原始输入点云的特征。解码器由多分支网络构成,用于生成点云形状。解码器的第一部分是通过4个特征提取块提取的全局特征,再通过MLP对全局特征进行特征变换,得到初步点云形状;第二部分是编码器中4个特征提取块提取的局部特征通过反卷积运算得到的全局特征,与第一部分生成的点云进行联合。最后,将两个分支相结合以获得完整的点云。

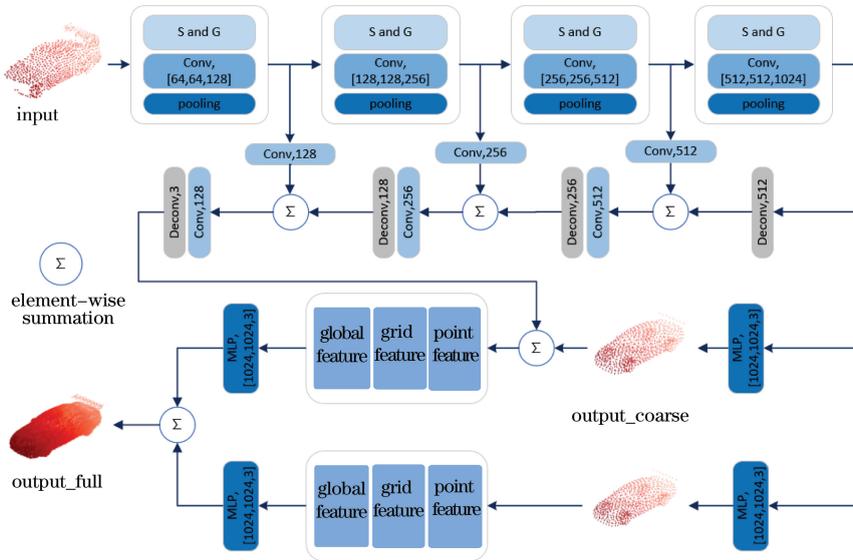


图1 多分支结构的三维点云补全网络

Fig. 1 Three-dimensional point cloud completion network with multibranch structure

2.1 特征提取的编码器

编码器由4个特征提取块组成,主要负责提取输入点云的全局特征和局部特征。编码器通过特征提取块提取全局特征,并保留提取的点云局部特征,然后将这两个特征传递到解码器的不同分支结构中。

输入的残缺点云通过编码器中4层不同的

DNN,学习全局特征向量 $\mathbf{v}' \in \mathbf{R}^k$ ($k=1024$ 为全局特征向量的维度)以及4个局部特征 $\mathbf{v}_i \in \mathbf{R}^{k_i}$ 。其中, i 为4个特征提取块, k_i 为每个特征提取块对应的特征向量维度, $k_1 \sim k_4$ 分别为3、128、256、512。每个块有三个子任务,包括对点云进行下采样、分组以及通过卷积神经网络(CNN)提取点云特征并对特

征进行最大池化。特征的下采样、分组任务是对点云的局部结构进行采样,然后通过不同的尺度进行分组,从而提取点云的局部特征,该方法可使网络结构对点云的局部特征更敏感;多层 CNN 可以提取采样点云和分组点云的特征;最大池化操作是一个对称函数,用来提取点云特征中信息量最多的特征。最后将提取的所有点特征信息进行汇总,得到全局特征 v' 和局部特征 v_i 。分别由 4 个块处理点云的局部特征和全局特征,并将其作为解码器中不同分支的输入。

2.2 多分支的解码器

解码器将编码器中提取到的全局特征和局部特征结合起来,生成一组完整的点云。多分支通过反卷积操作结合局部特征,再通过 MLP 运算结合全局特征。然后将全局特征与文献[20]中提出的二维网格点相结合,输出补全的点云,从而形成完整的多分支解码器网络结构。

解码器主要由两部分分支组成,第一部分分支由编码器中 4 个特征提取块提取的局部特征并通过反卷积运算得到,提取的 4 个局部特征的维度分别为 512、256、128、3,得到最终的点云特征维度为 (1024,3),这部分粗略点云 1 和标准有效值的差构成损失 L_1 。粗略点云 1 与第一部分分支融合,再与全局特征和网格点相结合并通过 MLP,输出第一部分的完整点云;第二部分分支是全局特征 v' 通过 MLP 线性变换生成的粗略点云 2,粗略点云 2 与标准有效值的差构成损失 L_2 。粗略点云 2 再结合网格点,通过全连接网络生成完整点云的第二部分。

结合两部分点云的全局特征、局部特征以及二维网格点,生成一个完整的点云。点云补全的本质是一个粗略点云通过 MLP 变换后得到完整点云的过程,最终完整点云的数量比粗略点云多出了二维网格点的数量,完整点云与标准有效值的差构成了损失 L_3 。

2.3 损失函数

本方法用倒角距离(CD)和搬土距离(EMD)描述输出点云 S_1 和标准点云 S_2 之间的距离损失函数,实验使用的损失函数是这两种距离函数的线性组合。点云 S_1 与 S_2 之间的平均最近点距离 d_{CD} 可表示为

$$d_{CD}(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2, \quad (1)$$

式中, $S_1, S_2 \in \mathbf{R}^3$, x, y 分别为点云集合 S_1, S_2 中的

点, $\|\cdot\|_2$ 为二范数。 d_{EMD} 可表示为

$$d_{EMD}(S_1, S_2) = \min_{\varphi: S_1 \rightarrow S_2} \frac{1}{|S_1|} \sum_{x \in S_1} \|x - \varphi(x)\|_2, \quad (2)$$

式中, φ 为一个 S_1 到 S_2 之间的映射,可用 $1+\epsilon$ 作为迭代参数。实验使用的损失函数可表示为

$$L(Y_{pre}, Y_{GT}) = \sum_{p=1,2,3} \alpha_p L_p, \quad (3)$$

式中, α_i 为权重, L_1, L_2 分别为第一部分、第二部分输出的预测值 Y_{pre-c1}, Y_{pre-c2} 与下采样标准数据 Y_{GT} 之间的距离, L_3 为完整输出的预测值 Y_{pre-f} 与标准数据 Y_{GT} 之间的距离。可以发现, L_i 是利用 CD 或 EMD 计算预测点与标准数据的距离,然后用 L 描述全局损失函数。

3 数据集及网络训练

实验用 ShapeNet 数据集中合成对象的 CAD 模型创建数据集,以训练本方法中的深度神经网络。该数据集不仅包含对象的完整点云模型,还包含了目标物体的残缺点云,共有飞机、橱柜、汽车、椅子、灯具、沙发、桌子、容器等 8 种实物模型,总数据量为 32913 个。这些点云模型在网格表面上均匀采样 16384 个点,作为对象的完整点云。残缺点云模型是将目标点云的 2.5D 深度图像反投影到 3D 图像中生成的部分点云,使数据更接近传感器采集的真实点云数据,残缺输入点云为 1024 个点。

ShapeNet 数据集中的复合点云模型具有多种物体完整的 3D 模型,可作为训练数据集,且某些对象在实际应用中无法获得,如飞机、船只。ScanNet^[21] 或 S3DIS^[22] 数据集也包含物体较好的 3D 模型,但这些 3D 模型中的缺失区域没有完整的测试样本数据,因此,无法作为实验的训练数据。

实验中的验证集包含 100 个模型,测试集包含 150 个模型,其余数据作为网络模型的训练数据集。用 Adam^[23] 作为 DNN 训练的优化器。设置网络的初始学习率为 0.0001,周期为 50,批次数为 24,每迭代 40000 次后,学习率降低 0.5。实验环境:CPU 为 Intel i7-7820, GPU 为 NVIDIA GTX 1080Ti,操作系统为 Ubuntu16.04,深度学习环境为 CUDA9.0+cuDNN7.1.4+Tensorflow 1.10。训练过程中, GPU 的运行效率为 94%~98%, GPU 的缓存使用率为 91%~95%。

4 实验结果

实验分为两个部分,第一部分在 ShapeNet 数

据集中展示点云形状补全后的结果,并将本方法与其他方法进行对比;第二部分展示了本方法在KITTI数据集中真实点云上的补全效果。

4.1 ShapeNet 点云模型的补全结果

用本方法对 ShapeNet 数据集中点云缺失的模型进行补充测试,并与其他方法进行对比。图 2 为本方法在 ShapeNet 数据集中对不同形状物体的点云补全效果,通过将目标物体的 2.5D 深度图像反投影到 3D 环境中获得输入对象点云,粗略输出是一个对象的完整点云模型,包含 1024 个点,这些点由局部特征与全局特征相结合生成,完整输出是包含 16384 个点的完整对象点云模型。

表 1 为多分支结构点云补全网络与其他点云补全网络在不同物体点云数据上补全效果的对比,包括折叠网络(FN)^[20]、点云生成(PSG)网络^[24]、PCN。而其他点云形状补全方法的前提是假设输入的先验信息超过了残缺点云的基本信息,因此,不能直接与本方法进行对比。可以发现,相比其他 3 种具有相似 DNN 结构的方法,本方法对一些具有凸外形结构物体的残缺点云补全效果更好。

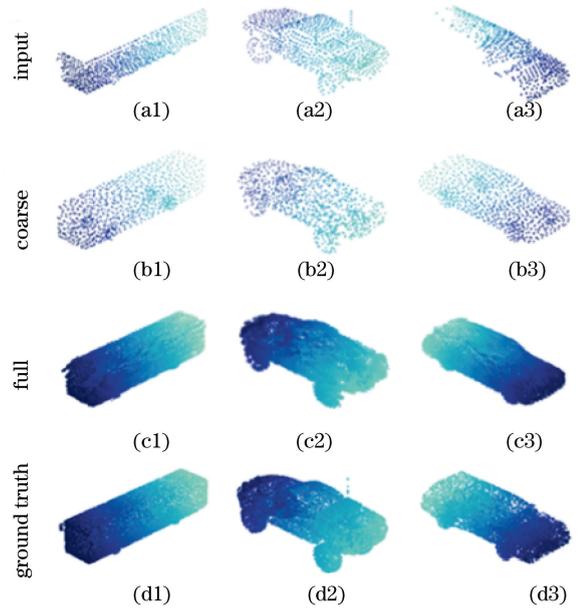


图 2 本方法的补全效果。(a) 原始点云; (b) 稀疏补全后的点云;(c) 完整补全后的点云;(d) 标准对比值
Fig. 2 Completion effect of our method, (a) Original point cloud; (b) point cloud after sparse completion; (c) point cloud after complete completion; (d) standard comparison value

表 1 不同方法的点云补全结果

Table 1 Point cloud completion results of different methods

Input	PSG	FN	PCN	Ours	GT

4.2 KITTI 模型的补全结果

KITTI 数据集中的图像数据来源于实际场景中的激光雷达扫描,即一系列用 Velodyne 扫描的点云场景,包含许多点云残缺的目标。对于 KITTI 数据集中的每个点云场景,首先提取标记为汽车的边界框内的点云,得到一个包含 2381 个目标的残缺点云集;然后获取每个目标的残缺点云模型,完成 KITTI 数据与 ShapeNet 数据的转换,再通过本方

法中的 DNN 进行训练,完成残缺车辆点云模型的补全;最后,通过旋转平移变换,生成与原始点云场景重合的汽车点云模型,完成点云的配准。与 ShapeNet 数据集中 2.5D 图像反投影的点云模型不同,激光雷达扫描的点云非常稀疏,如表 2 所示。可以发现,与 ShapeNet 数据集相比,部分点云包含的数据点非常少,合成物包含的数据点不超过 1000 个点。实验使用的模型可将这些不完整的点云映射为

具有完整形状的点云模型,并通过局部旋转和平移变换,将生成的完整汽车点云模型与原始汽车点云对齐,原因是本方法很好地结合了KITTI数据集和残缺的目标点云模型补全目标。

表2 KITTI数据集中本方法的补全效果

Table 2 Completion effect of our method in KITTI data set

Input	Output	Registration
		
		
		

5 分析与讨论

为了对不同方法在不同情况下的补全效果进行对比分析,首先对上述几种点云补全方法与本方法的补全效果进行定量的对比分析;其次分析了本方法对不同残缺比例点云的补全情况;最后讨论了本方法在不同外表形状下对残缺点云进行补全时产生的问题。

表3 不同方法补全结果的错误率(CD)

Table 3 Error rates of different methods of completion results (CD)

Method	Chair	Lamp	Airplane	Car	Table	Sofa	Bed	Bus	Average
PSG	0.0211	0.023	0.0215	0.0188	0.0191	0.0223	0.0201	0.0209	0.0203
FN	0.0291	0.0312	0.0299	0.0265	0.0260	0.0291	0.0267	0.0257	0.0274
PCN	0.0172	0.0201	0.0203	0.0152	0.0191	0.0192	0.0169	0.0178	0.0182
Ours	0.0171	0.0212	0.0189	0.0155	0.0178	0.0188	0.0168	0.0175	0.0177

表4 不同方法补全结果的错误率(EMD)

Table 4 Error rates of different methods of completion results (EMD)

Method	Chair	Lamp	Airplane	Car	Table	Sofa	Bed	Bus	Average
PSG	0.0821	0.0903	0.0892	0.0752	0.0821	0.0812	0.0767	0.0771	0.0781
FN	0.0891	0.0998	0.0981	0.0892	0.0903	0.0912	0.0899	0.0909	0.0921
PCN	0.0508	0.0512	0.0482	0.0471	0.0481	0.0483	0.0472	0.0469	0.0489
Ours	0.0491	0.0492	0.0497	0.0461	0.0469	0.0477	0.0478	0.0452	0.0474

5.3 不同几何形状物体的补全效果

本方法用MLP和二维网格相结合的方法补全点云,二维网格的本质是通过一个MLP函数对二

5.1 不同方法的补全效果

不同方法得到不同物体点云补全结果的CD和EMD如表3和表4所示,可以发现,对椅子、汽车、巴士、床等几何形状较为简单的物体,本方法的CD和EMD的效果都更好(CD和EMD的值均是越小越好)。

从表3和表4可以看出,相比FN、PCN和PSG方法,本方法对不同对象的补全效果更好,且CD的平均值比EMD好。图3为不同方法在ShapeNet数据集上的CD和EMD,可以看出,不同实例的CD均小于EMD,原因是EMD通过一对一的方式计算两个点云之间的距离,而CD是通过一对多的方式计算点集之间的距离。因此,通过CD计算得到的点云物体更重点云的整体结构,而通过EMD计算得到的点云更重点云的局部细节。理论上可通过EMD得到更多的计算细节,但计算损失的代价要大得多,整体形状的表现也要差得多。

5.2 不同残缺比例点云的补全效果

在残缺比例分别为25%、50%、75%的点云上,对比不同方法的补全效果,结果如图4所示。可以看出,EMD随输入点云残缺比例的增大会增加得更快,且在残缺比例较小时,本方法与PCN方法对点云的补全效果相近;相比PSG和FN方法,在同等情况下,本方法的补全效果也更好。

维网格进行映射,使二维网格通过旋转、扭曲等形变得到点云形状,无法有效解决二维网格在变换过程中产生的扭曲形变。ShapeNet数据集包含各种物

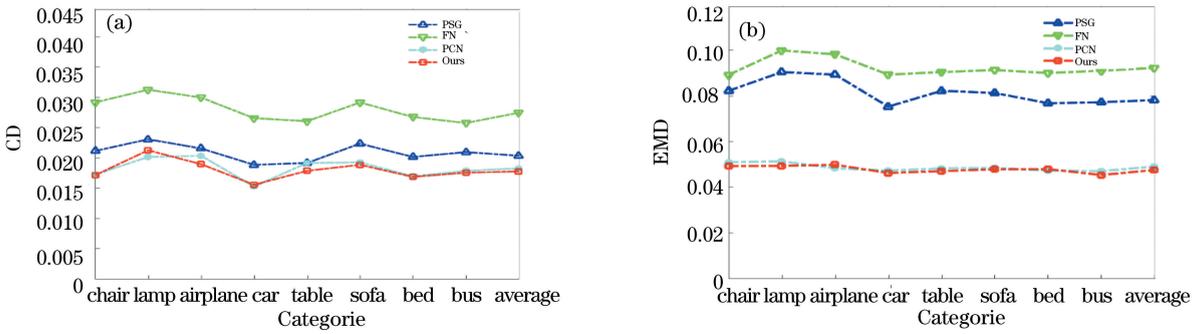


图 3 ShapeNet 数据集中不同物体的 CD 和 EMD。(a) CD; (b) EMD

Fig. 3 CD and EMD of different objects in the ShapeNet data set. (a) CD; (b) EMD

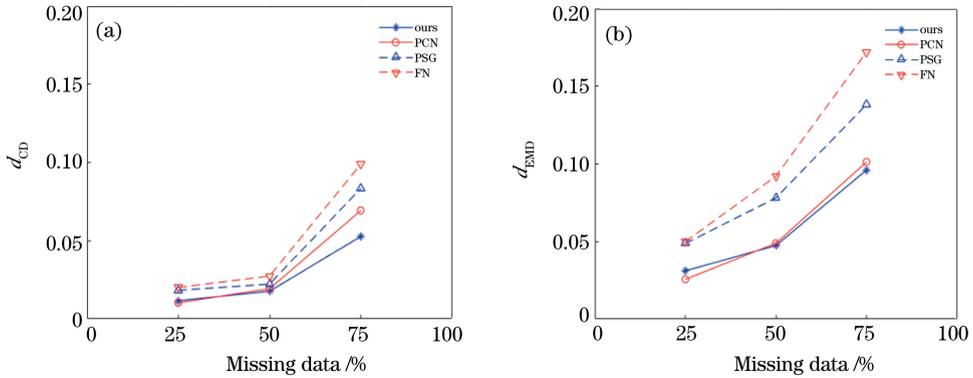


图 4 不同方法在不同残缺比例点云上的补全效果。(a) CD; (b) EMD

Fig. 4 Completion effects of different methods on point clouds with different incomplete ratios. (a) CD; (b) EMD

体的三维点云,有表面结构简单的,也有复杂的。本方法对外形复杂的凹面结构、镂空结构物体的补全效果如表 5 所示。通过多分支结构中的 MLP 结合二维网格的特征变换,可提高本方法对内部实体或外形凸面结构物体的补全效果,如表 6 所示。

表 5 本方法对凹面结构物体的点云补全效果

Table 5 Point cloud completion effect of our method on concave structure objects

	Table lamp	Ceiling lamp1	Ceiling lamp2
Input			
Pred			
GT			

表 6 本方法对凸面结构物体的点云补全效果

Table 6 Point cloud completion effect of our method on convex structure objects

	Car1	Car2	Car3	Car4
Input				
Pred_1				
Pred_2				
GT				

6 结 论

针对真实应用场景中存在的遮挡和采样不均匀导致传感器采集的目标物体点云形状不完整问题,提出了一种基于多分支结构编码-解码的点云形状补全网络。将采集的点云直接作为原始输入,不需要转换成其他形式的数据。网络的编码器部分可直

接从原始点云信息中提取特征,在特征提取过程中,编码器学习点云的全局特征和局部特征。解码器对编码器提取的特征进行解码,从而生成完整的点云。用EMD和CD修正生成的点云形状,得到目标点云的完整形状。在ShapeNet和KITTI数据集上的实验结果表明,本方法可以很好地补全对象缺失的点云形状,且在不同物体、不同残缺比例以及不同几何集中物体的补全效果较好。

参 考 文 献

- [1] 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.
- [2] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? the KITTI vision benchmark suite [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 3354-3361.
- [3] Berger M, Tagliasacchi A, Seversky L, et al. State of the art in surface reconstruction from point clouds [C] // Eurographics 2014 - State of the Art Reports, April 7-11, 2014, Strasbourg, France. Girona: ViRVIG, 2014: 161-185.
- [4] Davis J, Marschner S R, Garr M, et al. Filling holes in complex surfaces using volumetric diffusion [C] // Proceedings of First International Symposium on 3D Data Processing Visualization and Transmission, June 19-21, 2002, Padova, Italy. New York: IEEE, 2002: 428-441.
- [5] Mitra N J, Guibas L J, Pauly M, et al. Partial and approximate symmetry detection for 3D geometry [J]. ACM Transactions on Graphics, 2006, 25(3): 560-568.
- [6] Mitra N J, Pauly M, Wand M, et al. Symmetry in 3D geometry: extraction and applications [J]. Computer Graphics Forum, 2013, 32(6): 1-23.
- [7] Han F, Zhu S C. Bottom-up/top-down image parsing with attribute grammar [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(1): 59-73.
- [8] Li Y Y, Dai A, Guibas L, et al. Database-assisted object retrieval for real-time 3D reconstruction [J]. Computer Graphics Forum, 2015, 34(2): 435-446.
- [9] Felzenszwalb P F, Girshick R B, McAllester D, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [10] Gupta S, Arbeláez P, Girshick R, et al. Aligning 3D models to RGB-D images of cluttered scenes [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 4731-4740.
- [11] Liu J, Bai D. 3D point cloud registration algorithm based on feature matching [J]. Acta Optica Sinica, 2018, 38(12): 1215005.
刘剑, 白迪. 基于特征匹配的三维点云配准算法 [J]. 光学学报, 2018, 38(12): 1215005.
- [12] Zhang K, Qiao S Q, Zhou W Z. Point cloud segmentation based on three-dimensional shape matching [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121011.
张坤, 乔世权, 周万珍. 基于三维形状匹配的点云分割 [J]. 激光与光电子学进展, 2018, 55(12): 121011.
- [13] Liu M, Shu Q, Yang Y X, et al. Three-dimensional point cloud registration based on independent component analysis [J]. Laser & Optoelectronics Progress, 2019, 56(1): 011203.
刘鸣, 舒勤, 杨赞秀, 等. 基于独立成分分析的三维点云配准算法 [J]. 激光与光电子学进展, 2019, 56(1): 011203.
- [14] Tang Z R, Liu M Z, Jiang Y, et al. Point cloud registration algorithm based on canonical correlation analysis [J]. Chinese Journal of Lasers, 2019, 46(4): 0404006.
唐志荣, 刘明哲, 蒋悦, 等. 基于典型相关分析的点云配准算法 [J]. 中国激光, 2019, 46(4): 0404006.
- [15] 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.
- [16] Dai A, Qi C R, Nießner M. Shape completion using 3D-encoder-predictor CNNs and shape synthesis [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6545-6554.
- [17] Song S R, Yu F, Zeng A, et al. Semantic scene completion from a single depth image [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu,

- HI, USA. New York: IEEE, 2017: 190-198.
- [18] Yuan W T, Khot T, Held D, et al. PCN: point completion network[C]//2018 International Conference on 3D Vision (3DV), September 5-8, 2018, Verona, Italy. New York: IEEE, 2018: 728-737.
- [19] Charles R Q, Hao S, 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: 77-85.
- [20] Yang Y Q, Feng C, Shen Y, et al. FoldingNet: point cloud auto-encoder via deep grid deformation [EB/OL]. [2020-04-30]. <https://arxiv.org/abs/1712.07262>.
- [21] Dai A, Chang A X, Savva M, et al. ScanNet: richly-annotated 3D reconstructions of indoor scenes [C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 22-25, 2017, Honolulu, Hawaii. New York: IEEE, 2017: 2432-2443.
- [22] Armeni I, Sax S, Zamir A R, et al. Joint 2D-3D-semantic data for indoor scene understanding [EB/OL]. [2020-04-23]. <https://arxiv.org/abs/1702.01105>.
- [23] Kingma D P, Ba J. Adam: a method for stochastic optimization [EB/OL]. [2020-04-20]. <https://arxiv.org/abs/1412.6980>.
- [24] Fan H Q, Su H, Guibas L. A point set generation network for 3D object reconstruction from a single image [EB/OL]. [2020-04-25]. <https://arxiv.org/abs/1612.00603>.