

基于深度学习的实时点云修补算法

乔瑶瑶, 张兰兰

(黄河交通学院智能工程学院, 河南 焦作 454950)

✉ 1157764920@qq.com; 2089208988@qq.com



摘要: 传统点云修补方法对于特征不明显的点云修补效果不佳, 为此提出了一种基于深度学习的点云修补算法, 算法框架包括自动编码器、对抗生成网络及强化学习模块。该算法首先在潜在空间表征上训练对抗生成网络, 然后引入自动编码器网络, 自动编码器网络中的编码器可以将输入点云压缩成潜在空间表征, 最后引入强化学习模块, 为对抗生成网络的生成器选择合适的向量合成完整点云的潜在空间表征。实验表明, 该算法可以修补特征不明显的点云, 并能在毫秒级别内较为有效地完成点云缺失区域修补。

关键词: 点云修补; 深度学习; 对抗生成网络

中图分类号: TP391 **文献标识码:** A

Real-Time Point Cloud Repair Algorithm Based on Deep Learning

QIAO Yaoyao, ZHANG Lanlan

(College of Intelligent Engineering, Huanghe Jiaotong University, Jiaozuo 454950, China)

✉ 1157764920@qq.com; 2089208988@qq.com

Abstract: Traditional point cloud repair methods have poor performance in repairing point clouds with unclear features. Therefore, this paper proposes a point cloud repair algorithm based on deep learning, which includes an automatic encoder, adversarial generation network, and reinforcement learning module. This algorithm first trains the Generative Adversarial Network (GAN) on potential spatial representations. Then, an automatic encoder network is introduced, the encoder of which can compress the input point cloud into potential spatial representations. Finally, a reinforcement learning module is introduced to select suitable vectors for the generator of the GAN to synthesize the potential spatial representations of the complete point cloud. Experiments have shown that the proposed algorithm can repair point clouds with unclear features and effectively repair missing areas of point clouds within milliseconds.

Keywords: point cloud repair; deep learning; GAN

1 引言(Introduction)

修补是点云预处理中的重要环节和基础, 其主要是通过一定的先验信息对缺失点云进行修补。传统修补方法的先验信息为物体的基础结构信息, 比如对称性信息和语义类信息等, 因此传统方法在处理结构特征不明显的点云时效果不佳, 而基于深度学习的方法在三维点云处理方面取得显著成效。PointNet^[1]首次将深度学习方法应用到非结构化点云上; 王

春香等^[2]利用 GA-BP 遗传算法优化神经网络对三维点云进行修补; 吕富强等^[3]利用随机森林算法对地表点云空洞进行修补; 张艺真等^[4]通过训练 RBF 神经网络对点云数据进行修补。然而, 上述方法仅从单一的全局形状表征预测整个点云, 大多面临着结构细节信息丢失的问题。为解决这一问题, 本文采用基于深度学习的 RL-GAN 网络对缺失点云进行修补。RL-GAN 网络是一种基于强化学习(Reinforcement Learning,

RL)控制的生成对抗网络(Generative Adversarial Networks, GAN), 可以从不完整点云数据中预测完整点云信息。利用强化学习实现点云修补, 可以节约时间, 实现点云的实时修补, 并且对于大区域的缺失具有鲁棒性。

2 基于深度学习的实时点云修补方法(Real-time point cloud repair method based on deep learning)

基于深度学习的实时空洞修补算法的网络, 它由三个基本模块组成, 包括自编码器(Autoencoder, AE)、GAN 和强化学习模块。每一个模块都是一个深度神经网络, 需要单独训练。第一阶段, 自编码器使用完整点云作为数据集进行训练, 经过训练的编码器用于提取训练集中的每个点云的全局特征向量(Global Feature Vector, GFV)。第二阶段, 使用提取到的 GFV 训练 GAN。第三阶段, 将强化学习模块中的强化学习智能体(Reinforcement Learning-agent, RL-agent)与预先训练好的自编码器和 GAN 进行联合训练, 训练后的 RL-agent 将为 GAN 的生成器选择合适的种子向量, 使得生成器可以快速生成完整点云的 GFV。该网络的前向传播如图 1 所示。

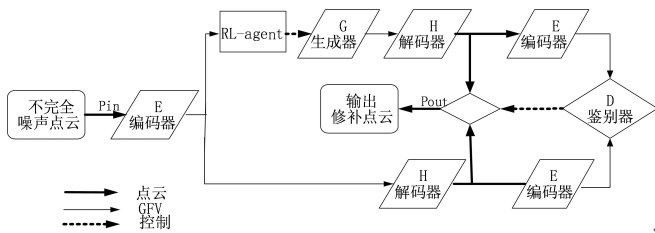


图 1 网络的前向传播

Fig. 1 Forward propagation of the network

2.1 自动编码器

自编码器由编码器和解码器组成。编码器将复杂的输入转换为潜在空间表征, 解码器将潜在空间表征还原到原始维度。自编码器是通过反向传播减少输入和输出点云之间的距离进行的训练。这里的点云距离可以采用推土机距离(Earth Mover's Distance, EMD)^[5], 也可以采用倒角距离(Chamfer Distance, CD)。EMD 用于衡量在某一特征空间下两个多维分布之间的不同。这里使用 EMD 作为损失函数, 它是基于两个点云 S_1 和 S_2 之间的距离表示的:

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

其中, φ 是 S_1 到 S_2 的双映射。

2.2 GAN

GAN 网络主要由生成器和鉴别器构成, 生成器将一个噪声包装成一个逼真的样本, 鉴别器判断送入的样本是否为真实样本, 在这个不断迭代的过程中, 鉴别器对样本的判别能力不断上升, 生成器的生成能力也不断上升, 最终两者的能力达到平衡。生成器的期望是将所生成的数据送入鉴别器后, 鉴别器能将其判别为真实数据, 鉴别器的期望是能将所有的生成数据和原始数据区分开, 二者在博弈的过程中共同提升性能。

GAN 的训练模式一般会先固定生成器, 迭代多次训练鉴

别器, 然后固定鉴别器训练生成器, 两者依次交替, 使用梯度下降的方法进行更新。GAN 的训练可能表现出发散、不稳定性和模式崩溃等现象。为了解决这些问题, ACHLIOPTAS 等^[6]提出在潜在空间表征上训练 GAN, 他们的研究也表明在潜在空间表征上训练 GAN 比在原始点云上训练得到的结果更加稳定。GURUMURTHY 等^[7]也采用了该方法训练 GAN 网络。因此, 本文采用该方法在 GFV 上训练 GAN 网络。

本文使用的生成器和鉴别器是 SAGAN^[8]网络的生成器和鉴别器, SAGAN 的损失函数借鉴的是 WGAN-GP^[9]的损失函数。为了解决原始 GAN 网络训练困难、生成样本缺乏多样性及损失函数无法引导训练过程等问题, 对改进的生成对抗网络(Wasserstein Generative Adversarial Networks-Gradient Penalty, WGAN-GP)做了一些更改, 对于生成器和鉴别器的损失函数不再取对数, 并且为鉴别器损失函数加上梯度惩罚项。鉴别器和生成器的损失函数分别如下:

$$L_D = E_{x \sim p_{data}} D(x) - E_{z \sim p_z} D(G(z)) + d_{gp}$$

$$L_G = -E_{z \sim p_z} D(G(z))$$

其中, p_{data} 是输入点云经过编码器编码生成的 GFV 的样本分布, $D(x)$ 是以 x 为输入和输出的鉴别器, 并且 x 是在 $[0, 1]$ 之间的标量。 $G(x)$ 是生成器, 能够将噪声 z 包装成一个逼真的样本, d_{gp} 是梯度惩罚项。

2.3 强化学习

在典型的强化学习框架中, 智能体与环境一直处于互动状态, 智能体发送动作至环境, 环境返回状态和奖励。在每个时刻, 智能体会接收到来自环境的状态, 基于这个状态, 智能体会依据一定的策略做出相应的动作, 然后环境会依据一定的状态转移概率转移到下一个状态, 与此同时, 环境会根据此时状态的好坏反馈给智能体一个奖励。智能体可以根据环境的反馈调整其策略, 然后继续在环境中探索, 最终学习到一个能够获得最多奖励的最优策略。

这里训练一个基于演员-评论家的强化学习网络, 该网络可以从连续的动作空间中学习到策略, 然后使用学习到的策略控制 GAN 完成点云修补工作。基于演员-评论家的网络架构由智能体和环境组成, 通过动作和奖励完成智能体 and 环境的交互。这里的环境是一个由自编码器和 GAN 组成的点云修补框架, 动作是 GAN 生成器的输入, 观测到的状态就是从不完整点云中编码获得的全局特征向量。假设环境具有马尔可夫性, 并且是完全可预测的, 也就是说最近的观测结果足以定义状态。智能体采取行动为生成器选择正确的种子, 然后合成的全局特征向量通过解码器获得完整的点云形状。强化学习控制的生成对抗网络(A Reinforcement Learning Agent Controlled GAN Network, RL-GAN-NET^[10])用到的是值优化和策略优化相结合的深度确定性策略梯度(DDPG)算法。本文采用双延迟深度确定性策略梯度算法(Twin Delayed Deep Deterministic Policy Gradient Algorithm, TD3)作为强化学习框架。

训练强化学习智能体的一个主要任务是建立正确的奖励函数。奖励函数通过对损失函数取相反数获得, 也就是说,

最大化奖励函数就等价于最小化损失函数，奖励函数 re 由 re_emd 、 re_D 、 re_G 构成，分别如下：

$$re_emd = -d_{EMD}, re_D = -L_D, re_G = -L_G$$
$$re = \alpha_1 re_emd + \alpha_2 re_D + \alpha_3 re_G$$

其中， α_1 、 α_2 、 α_3 是每个损失函数的权重， d_{EMD} 、 L_D 、 L_G 分别是自编码器 AE、鉴别器 D 、生成器 G 的损失函数。

3 实验(Experiment)

3.1 实验数据集和网络实现细节

(1)数据集。使用 ShapeNet 数据集作为实验的数据集，每个类别包含不同的模型网格。所有的模型被平移至以原点为中心的位置，并且将模型进行缩放，使得其边框对角线的长度为单位长度。之后对模型进行均匀采样，生成包含 16 384 个点的点云数据作为完整点云用于训练自编码器，并且通过自编码器编码的全局特征变量用于训练 GAN 网络。为了使输入的分布更接近真实世界的传感器数据，使用反向投影的深度图像作为不完整点云，得到的不完整点云作为测试集，并在完整点云数据集上训练强化学习智能体。

(2)网络实现细节。自编码器由编码器和解码器组成。编码器将输入点云转化成全局特征向量，解码器将全局特征向量转换成点云模型。编码器由 5 个一维卷积层组成，分别为 64、128、128、256、128 通道；解码器由 3 个全连接层组成，分别为 256、256、6 144 通道。每一层后面都跟着 ReLu 激活单元。通过减小输入点云与输出点云之间的 EMD 距离训练自编码器。

GAN 网络是由自编码器的编码器、生成器和鉴别器构成。采用的是自注意的生成对抗网络的主要结构。生成器和鉴别器均由 4 个卷积层、1 个一维转换卷积层及 1 个自注意机制网络组成。使用 WGAN-GP 对抗损失函数训练 GAN 网络，生成器每更新 1 次，则鉴别器更新 5 次。使用 Adam 作为优化器，其中生成器和鉴别器的学习率均设置为 0.000 01，批处理设置为 50。

强化学习模块由智能体和环境构成。使用基于演员-评论家的体系结构实现对 GAN 的连续控制，并且使用 TD3 算法训练。演员网络由 3 个全连接层组成，前两层使用 ReLu 激活，最后一层使用 tanh 激活。评论家网络由 3 层全连接层组成，前两层使用 ReLu 激活。

3.2 实验结果

使用来自 ShapeNet 的数据集测试所改进算法的有效性。实验网络输入的是缺失点云，期望得到的是高质量和高分辨率的补全点云。图 2 展示了本文方法在 ShapeNet 数据集上的修补结果，展示的分别是飞机、橱柜、汽车、椅子、灯具、沙发、桌子和船舶类缺失点云的修补结果。从图 2 可以看到，修补后的结果和实际点云可能不是每个细节都是对齐的，但是在语义上是合理的，并且不存在明显可见的空洞，修补结果达到了预期。图 3 展示的是桌子点云缺失不同部位和缺失程度不同的点云修补结果，修补结果良好，语义合理，没有明显可见的空洞。

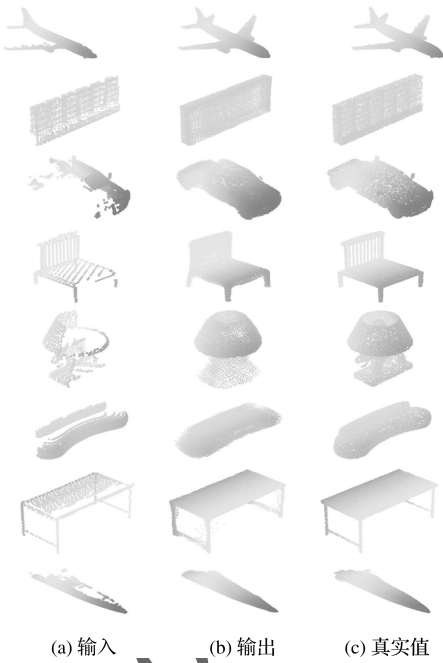


图 2 ShapeNet 数据集实验结果

Fig. 2 Experimental results of ShapeNet dataset

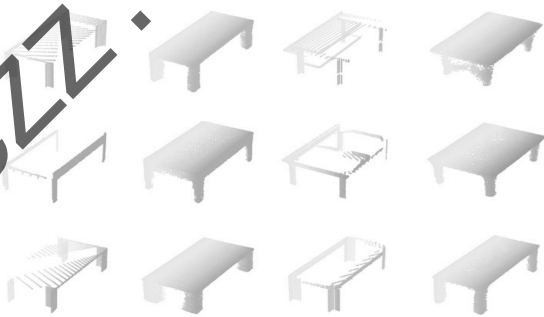


图 3 不同缺失程度的点云修补结果

Fig. 3 Point cloud repair results with different missing degrees

3.3 实验分析

(1)修补时间。在 ShapeNet 数据集上测试点云修补完成的时间。使用训练好的模型进行点云修补测试，输入缺失点云，输出完整的点云。每一类修补所需要的平均时间如表 1 所示。从表 1 中可以看出，本文所提方法可以做到实时的点云修补。

表 1 点云修补完成时间

Tab.1 Completion time of point cloud repair								
类别	飞机	橱柜	汽车	椅子	灯具	沙发	桌子	船舶
平均时间/ms	8.76	7.65	7.71	7.89	8.34	7.47	7.48	7.47

(2)相似性度量。点云修补的一个挑战是如何与真实数据进行比较。现有的相似性度量主要包括倒角距离 CD 和推土机距离 EMD ，使用这两个度量作为评估指标。对于两个点云， CD 测量一个点云中每个点到另一个点云中最近的点之间的平均距离， EMD 要求两个点云的大小一致。

将本文方法与以下方法进行比较。3D 编码器预测网络 (3D-Encoder Predictor Network, 3D-EPN^[11]) 是大型合成数据集端到端的训练，是体素修补方法的代表。为了进行比较，

将 3D-EPN 的距离场输出转换为点云,提取等值面,在生成的网格上均匀采样 16 384 个点。融合卷积自编码器(Fusion Convolutional Autoencoder, FCAE)使用直观的自编码器,编码器与 PointNet 中的相似,解码器是一个全连接层,使用倒角距离作为该方法的损失函数。三维形状曲面深度学习生成网络(AtlasNet^[12])使用类似的编码器,每次前向传播能输出 2 500 个点,将多次传递的生成点进行组合。点云补全网络(Point Completion Network, PCN)也使用自编码器完成点云的修补,它使用 PointNet 的改进版本作为编码器,并以由粗到细的方式生成点云,从中随机抽取 16 384 个点进行比较。

定量结果见表 2 和表 3,从中可以看出,本文方法在 CD 和 EMD 两个方面都优于 3D-EPN、FCAE、AtlasNet 和 PCN 方法。EMD 在点云比较方面更具有判别性和说服力。在所有类别中,本文方法的 EMD 是最低的,表明了该方法的优越性。EMD 在不同的对象类别中是不同的,这表明完成不同类别的困难是不同的。具体来说,在数据集中有更多数量的飞机和汽车,由于它们的结构相对简单和稳定,因此它们更容易完成。相比之下,在数据集中的各种灯是相对孤立的,因此更难完成。就 CD 而言,不同方法和对象类别之间的差异相对较小。虽然本文方法使用 EMD 作为损失函数,但是在 CD 方面也是优于其他方法的。

表2 ShapeNet 数据集结果定量比较: CD × 100

Tab.2 Quantitative comparison of ShapeNet dataset results: CD × 100

方法	飞机	橱柜	汽车	椅子	灯具	沙发	桌子	船舶	平均值
3D-EPN	1.31	2.17	2.02	1.86	2.56	2.1	2.15	1.83	2
AtlasNet	0.85	2.49	1.42	1.58	1.82	2.67	1.46	2.3	1.82
FCAE	0.7	1.4	1.1	1.29	1.72	1.37	1.16	1.33	1.26
PCN	0.66	1.35	1.1	1.41	1.46	1.36	1.14	1.23	1.21
本文方法	0.30	0.77	0.84	0.78	1.16	1.00	0.50	0.74	0.76

表3 ShapeNet 数据集结果定量比较: EMD × 100

Tab.3 Quantitative comparison of ShapeNet dataset results: EMD × 100

方法	飞机	橱柜	汽车	椅子	灯具	沙发	桌子	船舶	平均值
3D-EPN	6.28	7.92	8.93	7.78	10.9	8.23	8.23	8.36	8.33
AtlasNet	3.27	8.91	4.2	5.03	16.71	6.97	5.07	8.11	6.53
FCAE	4.1	11.2	7	7.64	14.64	7	7.77	7.22	8.32
PCN	3.44	8.79	4.44	6.89	15.45	6.28	6.84	6.56	7.34
本文方法	2.31	7.42	4.6	4.71	13.22	4.89	4.32	4.20	5.71

4 结论(Conclusion)

本文提出一种基于深度学习的实时点云修补算法,其主要目的是从不完整点云中生成完整点云,实现对不完整点云的实时修补,其中强化学习可以为 GAN 网络提供快速而稳定的控制,通过控制 GAN 将输入点云数据转换为具有高保真度的完整点云。为了能够实现实时的点云修补,一方面使用强化学习智能体控制生成器,另一方面该网络删除了成本高且复杂的优化过程,将损失函数转化为奖励函数。该网络可以较为稳定地补全含有大面积缺失区域的点云。在数据集 ShapeNet 上进行实验,实验表明本文所提方法在点云修补上可以实现不完整点云的实时修补,可以将部分点云数据转换成符合语义的完整点云。

参考文献(References)

[1] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation[C]// IEEE. Proceedings of the IEEE conference on computer vision and pattern recognition. Piscataway: The Institute of Electrical and Electronics Engineers, 2017:652-660.

[2] 王春香,郝林文,王耀,等.基于 GA-BP 神经网络的散乱点云孔洞自动修补[J].工程设计学报,2021,28(2):155-162.

[3] 吕富强,唐诗华,李灏杨,等.基于随机森林算法的地表点云孔洞修补[J].无线电工程,2022,52(9):1589-1593.

[4] 张艺真,孙志毅,柏艳红,等.基于 RBF 神经网络的三维点云数据孔洞修补[J].太原科技大学学报,2022,43(1):23-28.

[5] FAN H, SU H, GUIBAS L J. A point set generation network for 3d object reconstruction from a single image[C]// IEEE. Proceedings of the IEEE conference on computer vision and pattern recognition. Piscataway: The Institute of Electrical and Electronics Engineers, 2017:605-613.

[6] ACHLIOPTAS P, DIAMANTI O, MITLIAGKAS I, et al. Learning representations and generative models for 3d point clouds[C]// IMLS. International conference on machine learning. New York: PMLR, 2018:40-49.

[7] GURUMURTHY S, AGRAWAL S. High fidelity semantic shape completion for point clouds using latent optimization[C]// IEEE. 2019 IEEE Winter Conference on Applications of Computer Vision. Piscataway: The Institute of Electrical and Electronics Engineers, 2019:1099-1108.

[8] ZHANG H, GOODFELLOW I, METAXAS D, et al. Self-attention generative adversarial networks[C]// IMLS. International conference on machine learning. New York: PMLR, 2019:7354-7363.

[9] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans[C]// NIPS. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). New York: Curran Associates Inc., 2017:5769-5779.

[10] SARMA M, LEE H J, KIM Y M. RL-GAN-net: A reinforcement learning agent controlled GAN network for real-time point cloud shape completion[C]// IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: The Institute of Electrical and Electronics Engineers, 2019: 5898-5907.

[11] DAI A, RUIZHONGTAI QI C, NIEBNER M. Shape completion using 3d-encoder-predictor cnns and shape synthesis[C]// IEEE. Proceedings of the IEEE conference on computer vision and pattern recognition. Piscataway: The Institute of Electrical and Electronics Engineers, 2017:5868-5877.

[12] GROUEIX T, FISHER M, KIM V G, et al. A papier-mache approach to learning 3d surface generation[C]// IEEE. Proceedings of the IEEE conference on computer vision and pattern recognition. Piscataway: The Institute of Electrical and Electronics Engineers, 2018:216-224.

作者简介:

乔瑶瑶(1994-),女,硕士,助教.研究领域:计算机视觉.
张兰兰(1989-),女,硕士,助教.研究领域:图形图像处理与模式识别.