

# 基于树形结构编码器的点云补全算法

汤继武<sup>1</sup>, 谭台哲<sup>1,2</sup>

<sup>1</sup>广东工业大学, 计算机学院, 广东 广州

<sup>2</sup>河源市湾区数字经济技术创新中心, 广东 河源

收稿日期: 2022年3月7日; 录用日期: 2022年4月6日; 发布日期: 2022年4月13日

## 摘要

点云补全是目前计算机三维视觉领域较为重要的一个方向, 目前的深度学习算法采用的是编码器-解码器结构, 常用的编码器难以提取出精细的局部特征。本文基于生成对抗网络, 提出一种带有树形结构编码器的点云补全算法。树状卷积结构可以提取更为精细的点云特征向量, 并提高算法的计算效率。最后利用特征金字塔模型来生成点云的缺失部分。实验结果表明, 基于该网络结构补全的点云数据具有有效性, 并且补全精度相对于PCN算法有一定精度提高。

## 关键词

点云补全, 图卷积, 树形编码器, 生成对抗网络, 计算机视觉

# Point Cloud Completion Algorithm Based on Tree Structure Encoder

Jiwu Tang<sup>1</sup>, Taizhe Tan<sup>1,2</sup>

<sup>1</sup>School of Computer Science, Guangdong University of Technology, Guangzhou Guangdong

<sup>2</sup>Bay Area Digital Economy Technology Innovation Center, Heyuan Guangdong

Received: Mar. 7<sup>th</sup>, 2022; accepted: Apr. 6<sup>th</sup>, 2022; published: Apr. 13<sup>th</sup>, 2022

## Abstract

Point cloud completion is an important direction in the field of computer 3D vision. The current deep learning algorithm uses an encoder-decoder structure. It is difficult for commonly used encoders to extract fine local features. This paper proposes a method based on generative confrontation network, a neural network structure with a tree structure encoder to automatically repair the shape of the 3D point cloud, and the point cloud is better extracted by tree convolution to ex-

tract the feature vector of the point cloud and higher computational efficiency, and finally it uses the feature pyramid model to generate the missing parts of the point cloud. The experimental results show that the point cloud data completed based on the network structure is effective, and the completion accuracy is improved to a certain extent compared with the PCN algorithm.

## Keywords

Point Cloud Completion, Graph Convolution, Tree Encoder, Generative Adversarial Network, Computer Vision

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

三维视觉在实际生产生活中有着广泛的应用, 相关技术也随着深度学习的发展得到了进一步的推进。在各类 3D 数据中, 点云由于其数据量小, 表现更完美, 受到研究人员的热捧。现实世界中的点云数据通常由激光扫描仪、立体相机或低成本的 RGB-D 扫描仪等传感器采集而成。但在实际应用中由于采集设备性能和环境的限制, 真实扫描到的点云数据通常会因为遮挡、视点不同和噪声干扰等问题, 造成数据的不完整, 限制了点云数据在深度学习领域中的运用。例如, 在对点云数据进行分类的任务中, 若点云数据不完整, 则会极大地影响到最终的分类效果。在点云数据不完整的情况下, 分类效果不理想。因此如何从不完整的点云数据将其重新补充为完整的点云数据成为一个实际应用中不可缺少的任务。现有的点云补全网络发展迅速, 大多数用于形状补全的工作是以全监督的方式以部分 - 完成配对成训练数据对算法模型进行训练。虽然他们在域内数据上获得了很好的结果, 但这些方法推广到域外数据以部分 - 完成配对成训练数据对算法模型进行训练具有挑战性。尽管他们在自己的数据集上表现很优异, 但是推广到其它的数据时, 模型的泛华能力就会大打折扣。

考虑到真实世界数据的复杂性, 这些模型往往不能获得令人十分满意的结果, 对局部细节的补全效果不够理想。

目前主流算法使用体素和距离字段来描述三维对象, 以此完成三维形状补全任务。基于三维对象, 使用三维卷积来进行完成点云补全和其他任务, 并且在三维补全任务上取得了重大成功, 像 3D-ED-GAN 和 3D-RecGAN, 虽然实现了三维数据的补全, 但是使用三维卷积的方法受限于分辨率, 随着分辨率的增加, 内存负担和计算成本成指数增加, 虽然基于稀疏矩阵的方法又缓解这一问题, 但是对于细节信息丢失严重。不同于上述这些方法, 非结构化的点云数据因为内存消耗小, 对细节表示能力更强, 所以现有的研究通常使用点云数据来表示三维对象。但是由于点云数据的无序性质和非结构化性质, 它不适用于现有的卷积操作。Pointnet [1]通过直接处理点云数据, 解决了点云的无序性和旋转性问题, 这对点云的深度学习有着至关重要的影响, 并且激发了大量研究的灵感。PCN [2]是第一个专注于点云补全任务的深度学习网络模型, 它提出了编码器 - 解码器框架, 编码器主要由 PointNet++ [3]组成。Pointnet++将点云分解成多个块, 使用 PointNet 提取它们的特征向量, 并将这些向量通过 MLP 生成全局特征向量, 提取到的全局特征向量考虑到了点云的局部特性。PCN 使用的解码器由 FoldingNet [4]组成, 可以生成详细的点云和粗略的点云。由 TopNet 提出一种新的解码器, 可以通过在根树架构中隐式建模点云结构来生成结构

化点云。根树允许建模结构, 包括点集上的任何拓扑, 使解码器成为更通用的结构解码器, 以上两种方法都是将点云数据通过一个全局的最大池化层得到的特征向量, 再根据生成的特征向量, 来进行点云补全任务。然而, 由最大池化成生成的特征向量通常缺失精细的局部特征, 并且在模型考量中没有参考原始点云数据的信息, 从而造成补全完整后的点云缺失精细局部细节, 和原始的点云数据有一定的偏差。PF-Net [5]将不完整的点云数据输入, 只补全缺失的点云部分, 并且提出多分辨率的编码器和解码器结构, 以此更加有效的表示点云的局部信息, 完成点云补全的任务。因此, 考虑如何更有效率的生成具有细粒度的局部细节的特征变量来补全不完整的点云, 并且更好的生成点云的平均密度是本文主要的工作。

本文引入了一种简单而有效的采样方法, 在物体表面上, 通过采样降低计算性能, 并惩罚点与点之间的平均距离的方差, 这样可以在训练 GAN 时增强对象的一致性, 用更少的点来表述点云, 并且这样我们补全的点云保持了形状的合理性和多样性。通过树形卷积网络更有效地增加了模型对于点云局部特征信息的提取能力。通过多分辨率特征提取和金字塔结构点云生成, 构建了新颖的端到端点云补全网络, 进一步提升了网络的性能, 并且在 ShapeNet 数据集上补全效果更好。

## 2. 相关工作

### 2.1. 树形结构编码器(TreeGCN)

TreeGCN [6]是一种新的树状结构图卷积网络, 没有使用关于对象模型的先验知识, 保留了每个点的前置信息, 利用这些信息通过图卷积来提取新的点。TreeGCN 提出一个分支过程和一个循环单元, 这两个结构使的 TreeGCN 能够产生更精确的点云对象, 并且使模型的泛化性得到提高。此外, TreeGCN 中使用特征的前置信息比使用传统 GCN 中的相邻特征更具有计算效率。

### 2.2. 生成对抗网络模型

GAN 主要应用于生成式任务中, 原理如图 1 所示。它的核心是生成器 G 和鉴别器 D 之间的博弈。D 是一个鉴别器, 判别一张图片是不是“真实的”。它的输入参数是点云采样数据, 输出  $D(x)$ 代表  $x$  为真实点云的概率, 如果为 1, 就代表 100%是真实的点云, 而输出为 0, 就代表不可能是真实的点云。用生成器生成更加真实的点云数据去欺骗鉴别器的鉴别, 而鉴别器就要正确区分真实点云和虚假点云, 通过一代代的对抗使生成器生成的图像更加逼真, 鉴别器的鉴别能力也越来越强, 从而从样本的维度解决了衡量两个分布差异的问题。

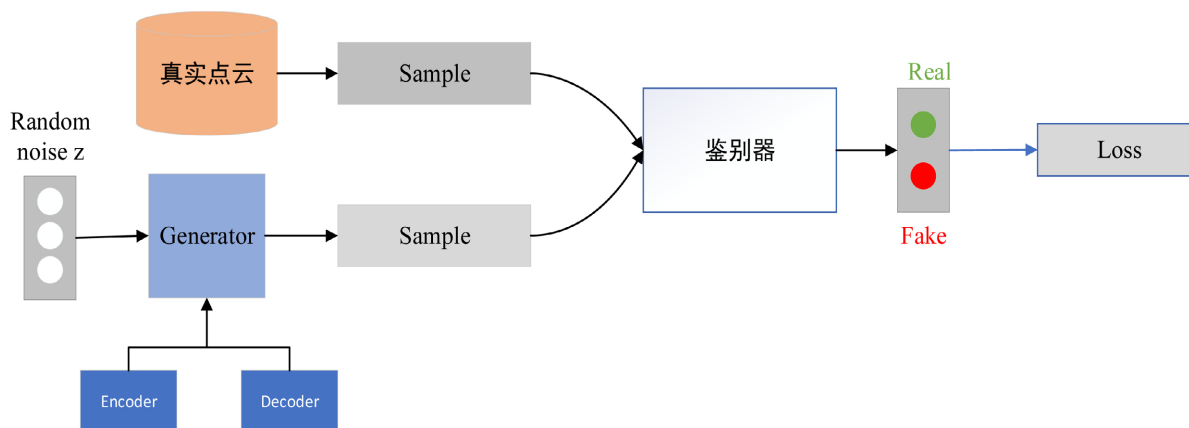


Figure 1. GAN schematic diagram

图 1. GAN 原理图

### 2.3. 网络模型

图 2 是本文提出的基于树形卷积生成器的多尺度点云补全网络模型整体框架。本文采取 GAN 对抗生成的编码器-解码器结构, 由树形卷积多分辨率生成器 TMRG, 金字塔解码器 PPD 和鉴别器 D 组成, 首先输入一个完整点云, 通过迭代最远点采样方法进行多次最远点采样, 选取代表点来更好的表示原始点云, 并且通过控制采样的点数来获取各个分辨率下的点云采样数据, 降低模型计算的复杂度和更好的提取其全局特征, 将采样后的低分辨率点云通过 TMMG 生成其潜在的特征向量  $V_1, V_2, V_3, V_4$  然后将多分辨率下的多个特征向量连接起来, 通过 MLP 将潜在特征映射到最终的特征向量。PPD 采取的是由 PFNET 提供的金字塔解码器, 全连接解码器擅长于预测点云的全局几何形状。然而, 全链接解码器会导致局部几何信息的损失, 因为它只使用最后一层来预测形状。以前的工作结合了全连接解码器与基于折叠的解码器, 以加强预测形状的局部几何。但是如果原始曲面相对复杂, 基于折叠的解码器不擅长保留原始详细几何信息。为了克服这一局限性, 我们将 PPD 设计为一个基于特征点的层次结构, 解码器以最终的特征向量  $V$  作为输入, 旨在输出缺失的点云, 最终使用全连接层来预测点云的全局几何形状。

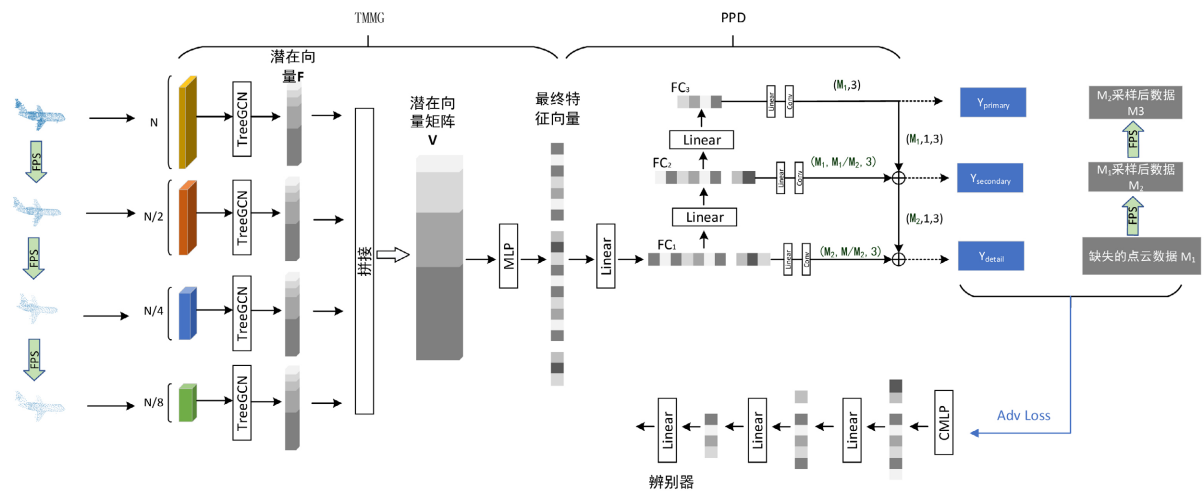


Figure 2. Point cloud completion network model of base TreeGCN

图 2. 基 TreeGCN 的点云补全网络模型

先将 TMRG 提取的特征向量  $V$  通过全连接层生成  $F_{C1}$ , 然后将  $F_{C1}$  依次通过全连接层生成  $F_{C2}$ ,  $F_{C3}$  再与  $F_{C1}$ ,  $F_{C2}$ ,  $F_{C3}$  通过全连接层和卷积网络生成最初预测的缺失点云  $Y_1, Y_2, Y_3, Y_4$ ,  $Y_{21}$  由  $F_{C2}$  和  $Y_2$  产生,  $Y_{31}$  由  $Y_{21}$  和  $F_{C3}$  合并生成, 最终迭代操作多次, 生成预测的缺失点云。将预测的缺失点云和真实数据计算损失, 其对应的真实点云送入鉴别器  $D$ , 经特征提取器计算后输出特征向量, 随后通过连续的全连接层  $F_C$ , 最终结果为真实(fake)或者虚假(real)的二值输出。

### 2.4. 模型损失

模型的损失函数包括两部分: 生成和对抗性损失。生成损失由编码器和解码器产生组成, 用来衡量生成点云和真实点云之间的差异程度。对抗性损失试图通过优化解码器和编码器来使补全看起来更贴近真实点云。常用的用来衡量点云之间的度量指标有倒角距离 CD (Chamfer Distance) 和 EMD [7] 距离。CD 测量预测点云  $S_1$  和真实点云  $S_2$  之间的平均最近点距离, 距离越小, 说明生成的点云越真实。

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

生成误差  $L_{com}$  由  $d_{CD1}, d_{CD2}, d_{CD3}$  组成,  $d_{CD1}, d_{CD2}, d_{CD3}$  分别为不同分辨率下的平均最近点距离,  $\alpha$  是加权值和超参数, 我们通过经验来设定, 通过设置  $\alpha$  来确定各个分辨率误差比重,  $Y_{detail}, Y_{primary}, Y_{secondary}$  分别为在不同分辨率下生成的点云,  $Y_{gt}, Y'_{gt}, Y''_{gt}$  为缺失点的 FPS 采样。

$$L_{com} = d_{CD1}(Y_{detail}, Y_{gt}) + \alpha d_{CD2}(Y_{primary}, Y'_{gt}) + 2\alpha d_{CD3}(Y_{secondary}, Y''_{gt}) \quad (2)$$

对抗损失: 我们的生成器模型为  $F$ , 鉴别器为  $D$ ,  $F$  将输入点云  $X$  提取特性向量后, 根据特征向量生成缺失点云  $Y$ , 然后通过鉴别器  $D$  试图去区分补全的缺失点和真实的缺失点差异性。

$$L_{adv} = \sum_{1 \leq i \leq S} \log(D(y_i)) + \sum_{1 \leq j \leq S} \log(1 - D(F(x_j))) \quad (3)$$

$F$  和  $D$  在训练时交替使用 Adam 进行联合优,  $\lambda_{com}$  和  $\lambda_{adv}$  为完成损失和对抗性损失的权重,  $L_{adv}$  为对抗损失, 我们的网络的联合损失被定义为:

$$L = \lambda_{com} L_{com} + \lambda_{adv} L_{adv} \quad (4)$$

并满足以下条件:

$$\lambda_{com} + \lambda_{adv} = 1 \quad (5)$$

### 3. 实验结果与分析

#### 3.1. 实验环境和数据集

本文选取的数据集为 ShapeNet, 数据集的 80% 作为训练集, 20% 作为测试集, 选取的类别为椅子, 飞机, 桌子 3 个常见类别的数据单独训练其对应的网络模型, 然后在 ShapeNet 测试集上对所训练的网络模型进行评估比较。每个点云采样 2048 个点进行训练, 然后在 ShapeNet 测试集上对所训练的网络模型进行评估比较。缺失的点云数据是通过随机选择一个视角作为中心, 并从中去除一定半径内的点来生成的。我们控制半径来得到不同数量的缺失点, 设置缺失率为 25% 的原始数据的点云进行训练和测试。

本文在 PyTorch 上构建了网络。所有这三个构建块都是通过交替使用 ADAM 优化器进行训练。初始学习率为 0.0001, 批处理大小为 64。在 TMRG 和鉴别器中使用批处理归一化(BN)和 RELU 激活单元, 在 PPD 中只使用了 RELU 激活单元。在 TMRG 中, 设置默认采取 4 个不同分辨率下的点云。在 PPD 中, 根据每个形状的点数为设置  $M1 = 64, M2 = 128$ 。只改变  $M$  来控制最终预测的大小。硬件配置为 Inter Xeon E5 2680 v3 (3.3 GHz) 处理器, GPU 为 NVIDIA GeForce GTX3090Ti、RAM64 GB。软件配置采用 Pytorch 深度学习框架, 操作系统为 64 位 Ubuntu16.04LTS。

#### 3.2. 实验结果分析

本文对比了 PCN 网络模型、折叠网络(FoldingNet)和 PSG [8] 网络模型, 实验结果如表 1 所示。通过实验结果可发现在基于倒角距离和推土机距离(更小说明更接近真实值点云)的对比下, 提出的方法生成的点云有具更高的保真度。

**Table 1.** Error rate of completion result based on CD  
**表 1.** 基于 CD 误差的补全结果的错误率

	飞机	椅子	桌子
PCN	0.0203	0.0152	<b>0.0178</b>
FoldingNet	0.0299	0.0311	0.0268
PSG	0.0215	0.0210	0.0191
Ours	<b>0.0154</b>	<b>0.0168</b>	0.0204



表 1 给出了在不同方法下对椅子、飞机、桌子等点云模型的补全效果的 CD 误差统计, 从表 1 中可以看出, 在针对具有相同缺失点云的修复补全任务中, 本文提出的网络模型维持了较好的效果。相比于 FoldingNet 和 PCN, 本文模型在修复补全缺失点云情况下与真实点云更加接近, CD 值更低, 补全效果更好。并且, 在 ShapeNet 数据集上, CD 的平均值要比 EMD 小。原因为 CD 是通过一对一来计算点云之间的距离, 而 EMD 是通过一对多计算; CD 更加注重点云的整体性, 而 EMD 更加注重局部细节, 所以通过 EMD 计算点云损失要更大(见表 2)。

**Table 2.** Error rate of completion result based on EMD

**表 2.** 基于 EMD 误差的补全结果的错误率

	飞机	椅子	桌子
PCN	0.0487	<b>0.0508</b>	0.0483
FoldingNet	0.0881	0.0895	0.0803
PSG	0.0842	0.0826	0.0824
Ours	<b>0.0321</b>	0.0513	<b>0.0329</b>

#### 4. 总结

本文针对三维点云形状的修复补全, 提出了一种有效的点云形状补全神经网络。该网络直接以离散点云数据作为输入, 通过生成对抗网络框架实现三维形状的修复补全。该网络的生成器采用自编码器架构, 由于其结合了输入点云数据的结构信息, 能够生成保持输入点云的精细结构特征的修复点云; 判别器则通过优化减少生成点云数据和真实点云数据之间的距离并用以指导生成器的不断优化, 最终实现输入点云形状的保结构修复补全。实验表明本文提出的点云形状补全网络针对稠密点云数据集和稀疏点云数据集均取得了较好的修复补全效果, 在修复形状的同时很好地保持了原有形状的精细结构信息[9], 并且对于不同程度的模型缺失, 具有很好的鲁棒性。

#### 参考文献

- [1] Qi, C.R., Su, H., Mo, K., et al. (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, 22-25 July 2017, 652-660.
- [2] Yuan, W., Khot, T., Held, D., et al. (2018) PCN: Point Completion Network. 2018 *International Conference on 3D Vision (3DV)*, Verona, 5-8 September 2018, 728-737. <https://doi.org/10.1109/3DV.2018.00088>
- [3] Charles, R.Q., Li, Y., Su, H. and Guibas, L.J. (2017) Point Net++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Long Beach, 3 June 2017, 5105-5114.
- [4] Yang, Y.Q., Feng, C., Shen, Y., et al. (2020) Folding Net: Point Cloud Auto-Encoder via Deep Grid Deformation. *IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake, 18-23 June 2018, 206-215.
- [5] Huang, Z., Yu, Y., Xu, J., et al. (2020) PF-Net: Point Fractal Network for 3D Point Cloud Completion. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, 13-19 June 2020, 7659-7667. <https://doi.org/10.1109/CVPR42600.2020.00768>
- [6] Shu, D.W., Park, S.W. and Kwon, J. (2019) 3D Point Cloud Generative Adversarial Network Based on Tree Structured Graph Convolutions. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 3858-3867. <https://doi.org/10.1109/ICCV.2019.00396>
- [7] Achlioptas, P., Diamanti, O., Mitliagkas, I. and Guibas, L.J. (2018) Learning Representations and Generative Models for 3D Point Clouds. *International Conference on Machine Learning*, Monterey, California, 6-10 August 2018, 40-49.
- [8] Fan, H.Q., Su, H. and Guibas, L. (2020) A Point Set Generation Network for 3D Object Reconstruction from a Single Image. *IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, 22-25 July 2017, 605-613.
- [9] 刘心溥, 马燕新, 许可, 万建伟, 郭裕兰. 嵌入 Transformer 结构的多尺度点云补全[J]. *中国图象图形学报*, 2022, 27(2): 538-549.