

浙江大学

硕士研究生读书报告



题目 球体引导的三维形状生成与操作

作者姓名 董爱祁

作者学号 22151031

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 2021 年 12 月 8

SP-GAN: Sphere-Guided 3D Shape Generation and Manipulation

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Li Qilei

By

Dong Aiqi

Zhejiang University, P.R. China

2021.12.8

摘要

论文提出了 **SP-GAN**，一种新的无监督球体引导生成模型，用于合成点云形式的三维形状，将复杂的三维形状生成任务分解成全局形状建模和局部结构调整，简化了学习过程，同时提高了生成形状的质量。该模型在每个球点之间形成隐式的密集联系，从而支持对三维形状的部分编辑，部分形状插值，多部分形状合成等。从实验结果中看，和现有模型相比，**SP-GAN** 能够合成具有更高质量，更多细节和更少噪声的三维形状。

关键词： 三维形状生成，三维形状操作，点云，生成模型

Abstract

The paper present SP-GAN, a new unsupervised sphere guided generative model, which is used to synthesize the 3D shapes in the form of point cloud. The key insight in this paper is to disentangle the complex 3D shape generation task into global shape modeling and local structure adjustment, which simplifies the learning process and improves the quality of the generated shape. The model forms an implicit dense correspondence between the sphere points and points in every generated shape, so as to support partial editing, partial shape interpolation, multi-shape part composition, etc. From the experimental results, compared with the existing models, SP-GAN can synthesize 3D shapes with higher quality, fine details and less noise

Keywords: 3D shape generation, 3D shape manipulation, point clouds, generative model

1 引言

三维形状的物体生成是计算机图形学和计算机视觉领域的重要问题，近年来，很多工作开始使用深度神经网络结合不同的三维表达方式来实现三维形状的生成，如体素，点云，三角网格，隐式曲面表达等，常见的基于点云的生成方式包括自回归模型，基于流的模型，生成对抗网络等，虽然已经取得了实质性的进展，但由于三维空间的形状变化的高度复杂性，生成具有多样性和真实感的 3D 形状仍然具有挑战性。同时，现有的生成模型通常专注于处理细节，缺乏部分感知生成和操作的结构性可控性。

而 SPGAN 提出了新的 GAN 模型，使用球体作为先验指导 3D 形状的生成，能够生成具有多样性，真实性的形状，具有更丰富的细节和更少的噪声，也支持部分感知生成和操作形状。

2 SP-GAN 的算法

2.1 算法概述

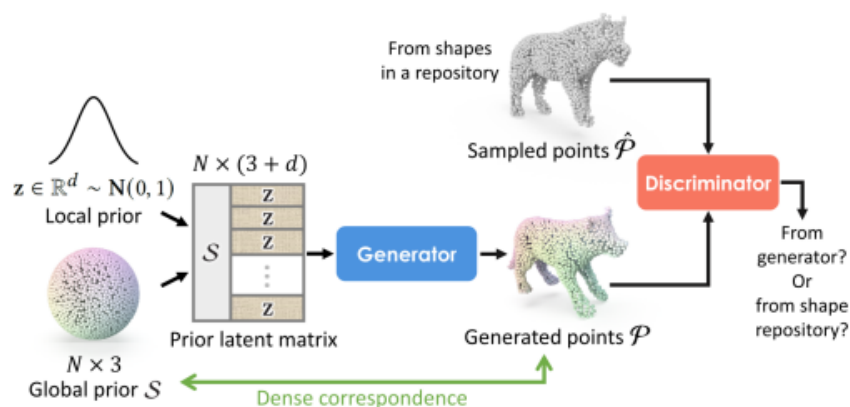
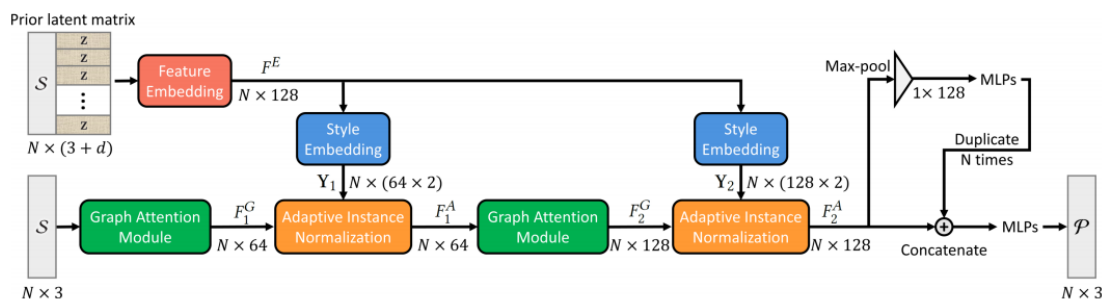


图 1 SP-GAN 概述

SP-GAN 包括两个解耦的输入：一个全局先验 S ，即一个球体 N 个点的三维坐标；一个局部先验 z ，即 d 维的随机潜在向量，其中的每个元素都是从一个标准正态分布中随机抽样得到的。我们将潜在向量 z 打包到 S 的每个点中，从而形成先验潜在矩阵作为生成器的输入，并在生成器中合成点云 $P \in \mathbb{R}^{N \times 3}$ ，并在给定的 3D 库中采样得到另一个点云 $\tilde{P} \in \mathbb{R}^{N \times 3}$ 。在训练过程中，判别器学习区分 P 和 \tilde{P} ，生成器学习生成尽可能类似 $\{\tilde{P}\}$ 的 P 。

2.1.1 生成器 (Generator)



在生成器中， \mathbf{z} 向点特征中引入了不同的局部特征和细节，而 \mathbf{S} 作为特征提取的先验形状，指导生成过程，允许我们使用空间相关性的图卷积来进行特征提取。

首先，生成器使用图形注意模块（Graph Attention Module）从 S 中提取点特征 $F_1^G \in R^{N \times 64}$ ，进行自适应实例归一化处理。同时，在顶部分支中，我们采用非线性特征嵌入（MLP）将先验矩阵转换为高维特征图 $F^E \in R^{N \times 128}$ ，然后通过样式嵌入（MLP）将 F^E 转化成样式 $Y_1 = (Y_1^s, Y_1^b)$ ，其中 $Y_1^s \in R^{N \times 64}$ 控制缩放， $Y_1^b \in R^{N \times 64}$ 控制偏移，自适应实例归一化操作的公式为：

$$F_1^A(i) = Y_1^s(i) \cdot \frac{F_1^G(i) - \mu(F_1^G(i))}{\sigma(F_1^G(i))} + Y_1^b(i), i \in [1, \dots, N]$$

$F_1^A(i)$ 即是归一化后对应的特征向量,通过 $Y_1 = (Y_1^s, Y_1^b)$ 的嵌入,最终的结果具有着更丰富的局部细节。为了进一步丰富特征嵌入,我们再将 F_1^A 输入另一个图形注意模块和自使用实例归一化中,分别得到 F_2^G 和 F_2^A ,然后通过输入 PointNet 中得到 3D 点云 P。

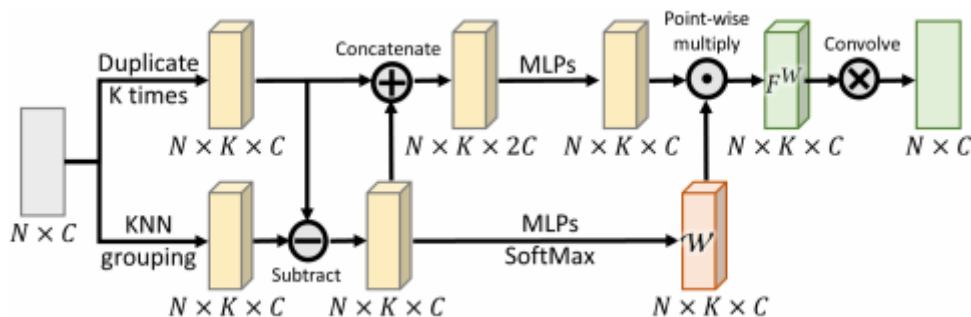


图 3 图形注意模块 (graph attention module) 结构

图形注意模块 (graph attention module) 采用了 DGCNN 模块, 并调整其结构使其进

一步考虑特征空间中 k 邻居之间的关系。回归权值为 $W \in R^{N \times K \times C}$ ，采用点乘法得到加权的相邻特征图 F^w ，最后通过 $1 \times K$ 的卷积核输出 $N \times C$ 的特征图。

2.1.2 判别器 (Discriminator)

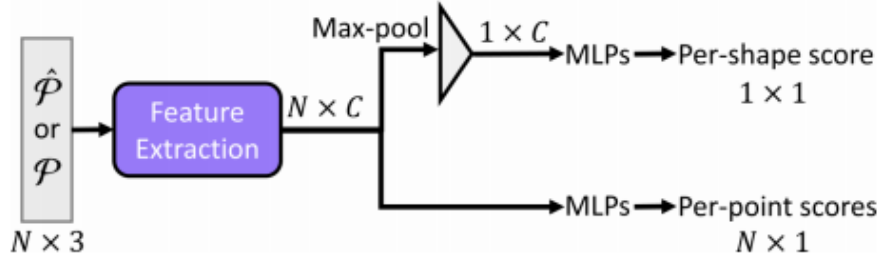


图 4 SP-GAN 判别器结构

传统的判别器仅学习整个形状的特征向量，仅预测形状分数，而 SP-GAN 使用 PointNet 提取特征，额外计算每个输入点的分数，能够更有效的正则化输入点云的全局和局部变化，鼓励生成器同时关注全局结构和局部细节。

2.1.3 损失函数

SP-GAN 为端到端的网络，根据最小二乘法设计了生成器和判别器的损失函数。从判别器结构可知，判别器输出由每个形状的分数和每个点的分数组成，因此，判别器的损失函数为：

$$L_D = L_D^{shape} + \lambda L_D^{point}$$

$$L_D^{shape} = \frac{1}{2} [(D(P) - 0)^2 + (D(\tilde{P}) - 1)^2]$$

$$L_D^{point} = \frac{1}{2N} \sum_{i=1}^N [(D(p_i) - 0)^2 + (D(\tilde{p}_i) - 1)^2]$$

相应的，生成器的损失函数为：

$$L_G = \frac{1}{2} [(D(P) - 1)^2 + \beta \frac{1}{2N} \sum_{i=1}^N [(D(p_i) - 1)^2]]$$

2.2 论文创新点

2.2.1 优势

和其他基于 GAN 生成点云的方式相比，SP-GAN 使用球体 S 作为全局先验模型指导三维形状的生成，采用局部先验 z 为模型提供局部细节。将复杂的三维形状生成任务分解为全局形状建模和局部结构调整两部分，通过这种解耦的方式简化了学习过程，提高了生成形状的质量。

SP-GAN 隐式地嵌入了生成形状之间的密集对应关系，从而促进了生成过程中的可控性，由于这种密集对应关系，能够更好的支持局部操作，如局部编辑，局部插值，多形状部分组合等。

SP-GAN 是无监督的，不需要标签注释，可以直接从随机分布中合成三维形状，具有更强的自由度，生成样本不受部分注释数据或任何预定义的结构关系的限制。

2.2.2 短板

SP-GAN 虽然能够以无监督的方式学习，但仍需要大量的形状进行训练，对于有限的训练样本或者结构复杂的形状，生成的结果仍然较为模糊，噪声较大。

SP-GAN 中采用点云的方式进行表示，虽然形状的生成和操作较为灵活，但不能直接生成表面或拓扑信息，需要后处理来重建表面，而重建表面的过程中可能出现扭曲的边缘和孔。

SP-GAN 虽然支持局部操作，但局部关系在模型中的表示不明确，甚至局部之间也嵌入了一种隐含的关系，使局部操作可能会对未选中局部造成一定影响。

2.3 应用及发展

SP-GAN 主要应用于形状生成，操作和分析，下面将进行具体阐述。

2.3.1 形状生成

SP-GAN 采用了 ShapeNet, SMPL 和 Smal 数据集作为训练数据，在每个形状中均匀采样 N （默认 2048）个点，将其标准化以适应单位球。在生成阶段在标准高斯分布中随机抽取潜在代码，最终生成的图像具有较少的噪声，并覆盖了丰富的全局和局部结构，因此 SP-GAN 可以用于随机生成三维图像。



图 5 SP-GAN 生成形状

2.3.2 形状局部编辑

SP-GAN 在每个生成形状之间建立了紧密的对应关系，我们很容易定位与一个形状中的特定部分相关的局部潜在编码，通过对这段潜在编码的替换即可生成局部变化的形状。

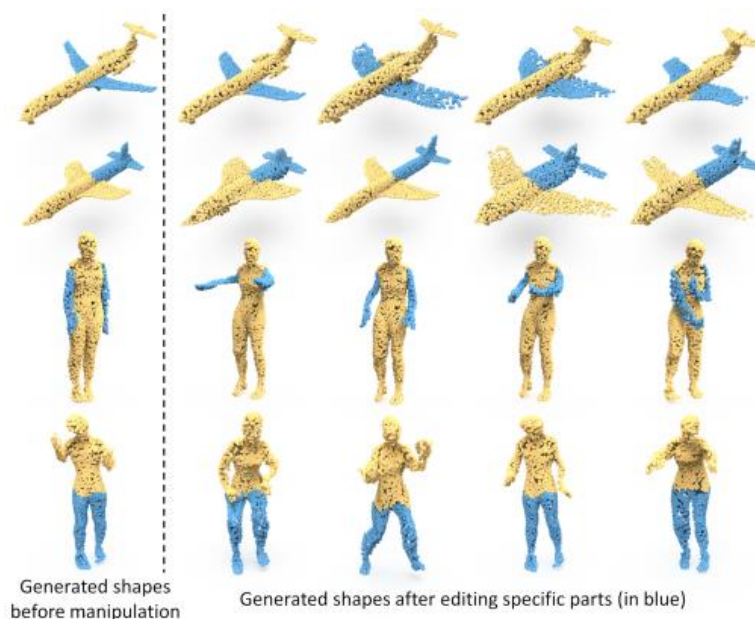


图 5 SP-GAN 局部编辑

2.3.3 形状插值

可以支持全局形状插值和部分形状插值，只需要对其潜在编码 z 进行插值即可得到

中间形状，即 $z_c = (1 - \alpha) \cdot z_a + \alpha \cdot z_b$ ，从实验结果中可以看出中间的生成形状也同样具有很少的噪声。

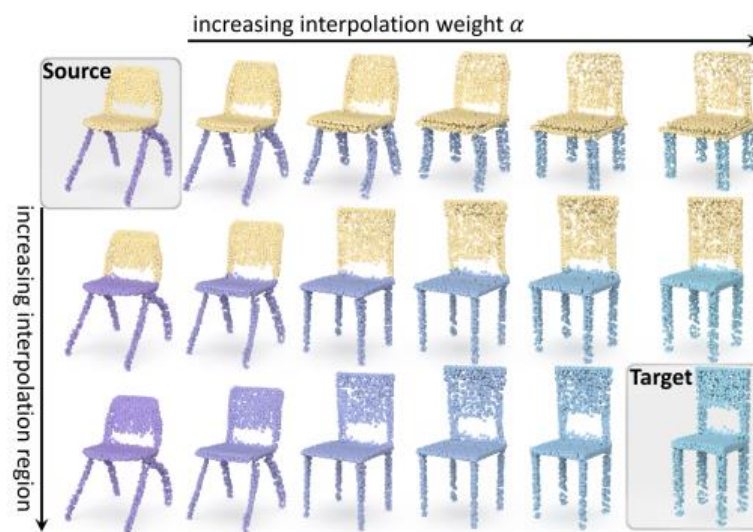


图 6 SP-GAN 局部插值

2.3.4 多形状合成

根据 SP-GAN 的隐式密集对应关系，将不同形状的潜在代码 z 进行组合拼接，用 S 打包一个新的先验潜在矩阵，从而产生新的组合形状。

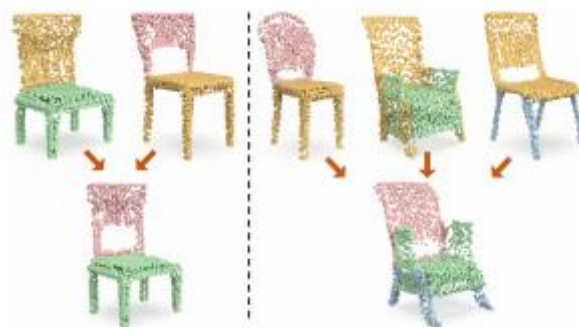


图 6 SP-GAN 多形状合成

2.3.5 形状分割

根据 SP-GAN 的隐式密集对应关系，SP-GAN 可以对一组三维形状进行公共分割，即将形状分割成语义一致的部分，我们只需要手动分割该集中的其中一个形状，通过密集的对对应关系，即可分割该组的其他形状。

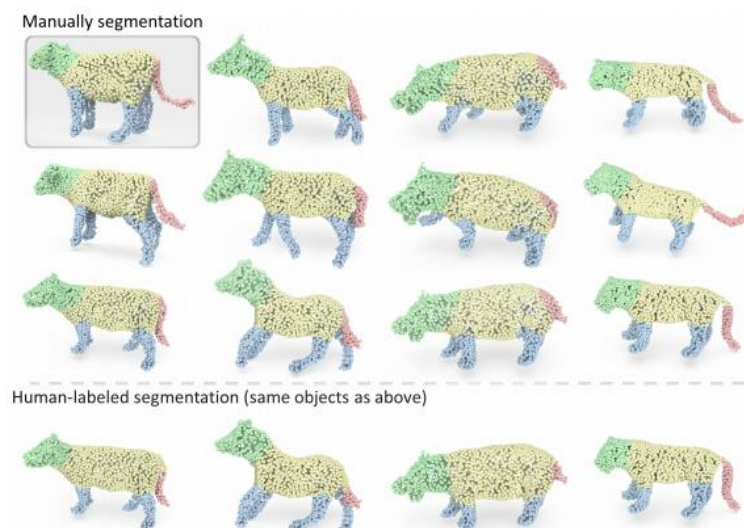


图 7 SP-GAN 形状分割

结合实验数据,SP-GAN 与其他五种最先进的生成模型相比,具有最好的表现能力,最短的训练时间和最少的可学习参数,能够生成多样性强,真实性高,质量好,细节丰富的三维形状,具有广泛的应用前景,未来可以进一步完善交互设计,实现用户指导的自动三维形状生成,尝试在点云的生成过程中探索点法线,优化网格重现。

3 小结

SP-GAN 是一种无条件的生成模型,用于直接生成以点云表示的三维形状,以先验潜在矩阵构造生成器输入,将输入解耦成全局先验和局部先验,建立具有样式嵌入和自适应实例归一化的生成器网络。该方案将三维形状生成任务简化为全局形状建模和局部结构调整,使生成过程能够从共享的全局初始化开始,又能够适应局部空间变化。

同时 SP-GAN 通过隐式的密集对应关系,增加了结构可控性,可以更方便的修改和插值潜在代码从而实现结构感知和形状操作,这是以往的点云生成工作都无法实现的。定量和定性的实验结果都表明,SP-GAN 能够产生多样性强,新的,真实的形状,表现出更精细的细节和更少的噪声,超出了以往工作的生成能力,具有较好的应用前景。

参考文献

- [1]. Ruihui Li, Xiaozhi Li, Ka-hei Hui, et al. SP-GAN:Sphere-Guided 3D Shape Generation and Manipulation[C]. ACM Transactions on Graphics. 2021.
- [2]. Yue Wang, Yongbin Sun, Ziwei Liu, et al. Dynamic graph CNN for learning on point clouds[C]. ACM Transactions on Graphics 38, 5 (2019), 146:1–146:12.
- [3]. Charles R. Qi, Hao Su, Kaichun Mo, and Leonidas J. Guibas. PointNet: Deep learning on point sets for 3D classification and segmentation[C]. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. 652–660.
- [4]. Panos Achlioptas, Olga Diamanti, Ioannis Mitliagkas, and Leonidas J. Guibas. Learning representations and generative models for 3D point clouds[C]. In Proceedings of International Conference on Machine Learning (ICML). 2018.
- [5]. Dong Wook Shu, Sung Woo Park, and Junseok Kwon. 3D point cloud generative adversarial network based on tree structured graph convolutions[C]. In IEEE International Conference on Computer Vision (ICCV). 2019. 3859–3868.
- [6]. Guandao Yang, Xun Huang, Zekun Hao, Ming-Yu Liu, Serge Belongie, and Bharath Hariharan. 2019. PointFlow: 3D point cloud generation with continuous normalizing flows[C]. In IEEE International Conference on Computer Vision (ICCV). 2019. 4541–4550.
- [7]. Le Hui, Rui Xu, Jin Xie, Jianjun Qian, and Jian Yang. Progressive point cloud deconvolution generation network[C]. In European Conference on Computer Vision(ECCV). 2020.
- [8]. Ruojin Cai, Guandao Yang, Hadar Averbuch-Elor, Zekun Hao, Serge Belongie, Noah Snavely, and Bharath Hariharan. Learning gradient fields for shape generation[C]. In European Conference on Computer Vision (ECCV). 2020
- [9]. Angel X. Chang, Thomas Funkhouser, Leonidas J. Guibas, , et al.. ShapeNet: An information-rich 3D model repository[DB/OL]. arXiv preprint arXiv:1512.03012 (2015).
- [10]. Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael

J. Black. SMPL: A skinned multi-person linear model[C]. ACM Transactions on Graphics (SIGGRAPH Asia) 34, 6 (2015), 248:1–248:16

- [11]. Silvia Zuffi, Angjoo Kanazawa, David Jacobs, and Michael J. Black. 3D Menagerie: modeling the 3D shape and pose of animals[C]. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. 5524–5532.