

基于形状先验的隐式三维建模算法研究

倪天杰, 何良华

同济大学计算机科学与技术系, 上海

收稿日期: 2023年10月1日; 录用日期: 2023年10月28日; 发布日期: 2023年11月3日

摘要

近年来, 隐式神经三维建模算法是计算机视觉领域的热门研究方向, 特别是基于符号距离函数(SDF)的方法, 然而目前的模型往往泛化性不强。由此问题, 该文利用动态图卷积网络对点云进行形状编码, 引入了形状先验假设, 提出了一种结合物体个性特征和其所属类别共性信息相结合的隐式三维建模算法, 提升了模型的准确性与泛化性。在ShapeNetV2数据集上, 相较于现有算法取得了更好的效果, 表明了本方法在三维隐式建模问题上的优越性。

关键词

隐式神经场, 深度学习, 三维建模, 点云卷积, 符号距离函数

Research on Implicit 3D Modeling Algorithm Based on Shape Prior

Tianjie Ni, Lianghua He

Department of Computer Science and Technology, Tongji University, Shanghai

Received: Oct. 1st, 2023; accepted: Oct. 28th, 2023; published: Nov. 3rd, 2023

Abstract

In recent years, implicit neural 3D modeling algorithm has become a popular research direction in the field of computer vision, especially the method based on signal distance function (SDF). However, this kind of model is often not strong generalization. To solve this problem, this paper uses the dynamic graph convolutional network to encode the shape of objects, introduces the shape prior hypothesis, and proposes an implicit 3D modeling algorithm that combines the individual characteristics of objects and the generic information of their categories, which improves the accuracy and generalization of the model. On ShapeNetV2 dataset, better results are obtained compared with the existing algorithms, which shows the superiority of the proposed method in three-dimensional implicit modeling.

Keywords

Implicit Neural Field, Deep Learning, 3D Modeling, Point Cloud Convolution, Signal Distance Function

Copyright © 2023 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

点云, 作为一种三维显示表征方式, 相较于体素以及网格而言, 是一种轻量化的结构, 但是损失了拓扑结构的信息, 观察起来并不直观, 为此, 有许多传统算法实现了从点云到表面网格的转化, 目前, 也有越来越多的深度学习算法来完成该任务, 并且大多数任务最终模型的输出并不是直接的表面, 而是依靠符号距离函数, 隐式的生成三维模型。特别地, 基于神经辐射场的三维重建算法近几年相当火热, 其依靠多视图以及相机位姿的输入, 生成相应的光线, 并在光线上采样点的方式, 本质仍然是基于点的方式隐式生成表面。

在基于深度学习的算法中, 大多算法忽略了模型的泛化性, 它在未知的形状上的表现不佳。这是因为, 他们忽略了形状之间的联系, 对于属于同一种类的形状而言, 基本上其集合结构应该是相似的, 特别是对于室内物体而言, 它们都是人造的, 人类也是基于经验制造的该类结构性的物体以满足相应的功能。

为此, 本文提出了一种形状的假设, 该假设结合了同一类别物体的共性信息和其中实例的个性特征, 使得同一类别的实例的拓扑结构之间产生了联系, 并在该假设下, 建立了相应的模型, 该模型采用流行的隐式表面重建的思路, 通过该模型可以学习空间中任意一点(在边界内)的 SDF 值, 并理论上可以完成任意精度的高质量三维表面重建。我们在 ShapeNetV2 数据集上进行了相关三组类别的实验, 并验证了形状假设的可行性和效果。

2. 相关工作

2.1. 隐式表面重建

基于三维隐式表示的深度学习算法在近几年层出不穷, 隐式表示的表征能力相较于显示方法更好, 这是因为隐式表征理论上可以显示任意精度的三维模型, 而显式方法受存储的限制而扩展精度。[1]提出了 IM-NET 的自编码器, 需要点坐标和形状的特征向量的输入, 采用解码器和对抗生成后输出该点的占有率。[2]提出了一种基于多视图和体渲染的方以彩色图作为监督学习三维模型, 其中[3]建立了 SDF 值与体渲染之间的联系可提取出高质量的表面提取网络。上述方法都是基于 2D 进行三维隐式曲面估计的方法, 但本质上仍然是基于点的方式。此外, 还有一类直接的方法, [4]直接使用多层感知机预测每个点的 SDF 值, 并提出了一种自解码器的架构, 使得其具有了表示一类形状的能力。[5]提出了一种表面生成框架 AtlasNet, 它将 3D 形状表示为参数表面元素的集合, 直接地推断出形状的表面表示。[6]提出了一类基于占有率的网络 Occupancy Network, 并引入了多分辨率等值面提取算法, 以更好地提取表面。[7]提出了局部隐式网格表示, 具有可扩展性和通用性, 训练一个自动编码器来学习在该尺寸下嵌入 3D 形状的局部组件。

2.2. 基于深度学习的点云特征提取

因为点云的无序、无连接、不规则的性质以至于无法像像素那样, 将 2D 卷积直接应用于 3D。为此, 对于点云的深度学习也是具有挑战性的一个任务。[8]提出了 PointNet 保证了输入点的排列不变性, 并且其非常的高效, 随后[9]改善了 PointNet 无法捕捉局部结构的问题, 可以学习多尺度结合的点云特征。[10]提出了一种边卷积模块, 包含了点云的邻域信息并且可以堆叠学习全局形状属性, 在点云分类和分割任务上取得了优秀的性能。

3. 方法

3.1. 总览

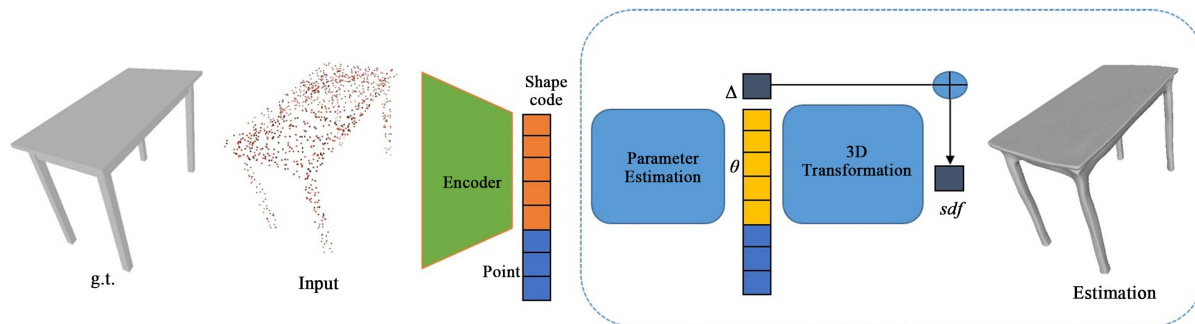


Figure 1. The whole algorithm framework of this paper

图 1. 本文算法整体框架

隐式神经三维表面重建的基本思想是用神经网络拟合一个连续函数 f_θ , 使其满足等式(1)。即可以通过

$$f_\theta(x) \approx sdf(x), \mathbb{R}^3 \rightarrow \mathbb{R} \quad (1)$$

过该模型, 获取三维空间中任意一点的符号距离函数(SDF), 其定义为该点到离其最近的表面的有向距离, 符号由该点在表面的内部还是外部决定, 内部为负, 外部为正。需要注意的是, 在实际应用中, 一般会为距离设定阈值 δ , 以此增强表面细节的刻画并减少无意义的计算, 在本算法中, 设定阈值 $\delta = 0.1$, 单位为米。最后, 可以利用光线投射进行二维渲染[11]或者 Marching Cubes [12]抽取 SDF 值为零的等值面, 本文使用后两者获取三维表面网格。

然而, 为每一个物体训练单独的网络不仅花销很大且是不合理的。直觉上, 同一个类别的物体, 几何结构基本上是类似, 只是每个具体的物体有些独自的特征, 二者结合获得最终的形状。这一点, 是本文形状先验假设开展的必要条件, 事实上, 生活中的物体大多都是结构化的物体, 特别是室内的物体。为此, 本文希望构建一个针对类别的隐式神经场, 以提升模型的泛化性, 受到[13]的启发, 提出了如等式(2)的形状先验的等式:

$$S^i = T(\bar{S}, \theta^i) + \Delta^i \quad (2)$$

假设有某一类别的形状集合 $S = \{S^i\}, i = 1, 2, \dots, n$, \bar{S} 为该类别所有形状的共性特征, T 是变换函数, 在[13]中将其定义为一个三维的薄板样条插值(TPS), 该方法由于核函数选取和锚点个数的限制, 对于三维变换的表征能力有限而本文采用深度学习的方法, 提升其表征的能力, 其中 θ^i 为变换函数的参数。此外, 为了捕获无法从共性信息中获得的物体个性特征, 加入调整项 Δ^i 来改善。

在这一形状先验假设的指导下, 我们的模型结构如图 1 所示, 采用 Encoder-Decoder 架构。输入为某

一类别 S 的实例点云 $S^i = \{x_j^i | x_j^i \in \mathbb{R}^3\}$, 在编码器阶段采用动态图卷积提取点云的形状编码 z_i , 在解码器阶段, 又分为两部分, 在三维变换参数估计网络中, 将提取的形状编码 z_i 和点云中的一点 x_j^i 作为输入, 输出该点的变换参数 θ_j^i 和改进参数 Δ_j^i , 而后在三维变换网络中, 将 θ_j^i 和 x_j^i 连接后作为输入, 并输出该点的符号距离值 s_j^i 并加上 Δ_j^i 作为最终的输出, 在训练完成后, 通过 Marching Cubes (移动立方体) 在提取出实例 S^i 的表面 $\partial\Omega^i$ 。

3.2. 编码器

在编码阶段, 对于输入的点云, 采用动态图卷积网络(DGCNN)提取形状的特征。其结构如图 2 所示。

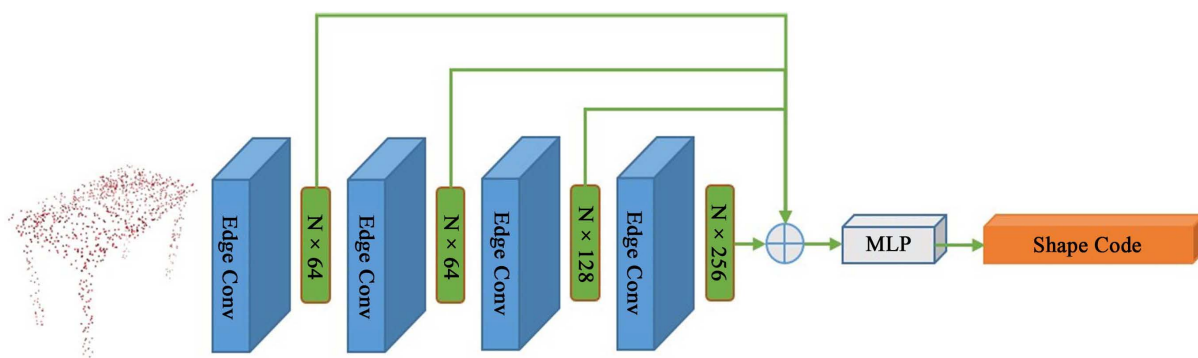


Figure 2. The network of encoder

图 2. 编码器网络

对于输入的形状点云 $S^i = \{x_j^i | x_j^i \in \mathbb{R}^3\}$, 在经过四层图边卷积层(Edge Convolution)后, 通过最大池化层和全连接层, 得到其形状编码 z_i 以供后续模块使用。DGCNN 在提取点云的特征时, 不仅考虑全局的全局特征, 并且融合了点云的局部信息, 使得其在点云分类任务和点云分割任务上都相较于别的点云特征提取算法有了很大的提升, 因此本文考虑使用其作为编码器, 生成形状编码供后续模块使用。

在 DGCNN 中, 最为核心的是边卷积模块, 其流程如图 3 所示。对于任意一点 $x_i \in \mathbb{R}^f$, 首先根据该点构建 k 近邻图, 其边的定义如等式(3), 可以看到通过 x_i 保持了点的全局排列不变性, 并通过

$$e_{ij} = h_{\theta}(x_i, x_j - x_i) \quad (3)$$

$x_j - x_i$ 保证了点的局部特征。 x_i 的 m 通道的计算公式如等式(4)所示, 其本质是在对 x_i 进行 k 近搜

$$x_{im} = \max_{(i,j) \in \delta} \text{ReLU}(\theta_m(x_j - x_i), \Phi_m x_i) \quad (4)$$

索后的全连接的计算, 其动态特性是指在每一层边卷积层前重新构图, 使得算法具有非局部扩散的性质。

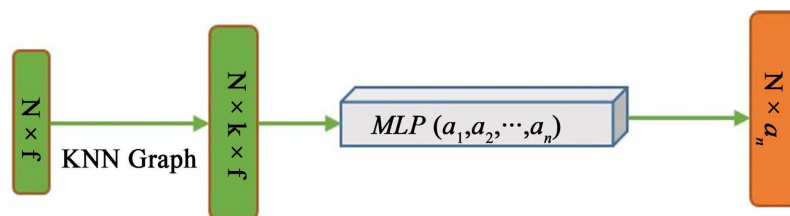


Figure 3. Edge convolution

图 3. 边卷积模块

3.3. 基于形状先验的解码器

在解码器模块, 其结构如图 4 所示, 基于形状先验等式(2), 我们将其分为两个部分, 第一个为形变参数估计网络 E , 对于形状 S^i 每一个输入的点 x_j^i , 我们估计其的形变参数 θ_j^i 以及位移 Δ_j^i , 在[13]的工作中, 其采用传统的三维薄板样条插值来定义三维变换, 在计算变换参数时, 首先要进行锚点的匹配算法, 随后根据匹配的锚点, 在迭代地估计参数, 因此其计算成本相当巨大, 此外, 由于锚点选取的多少以及核函数的选取的优劣, 有可能影响重建的效果。而在本文中, 使用神经网络实现端到端的参数估计。对于输入的每一个点都会计算相应的变换, 不再受锚点的限制。此外, 受到[14]的启发, 我们引入了超网络模型(Hyper Network) H , 以接收由编码器输出的形状编码 z_i , 初始化参数估计网络的权重参数, 以此更好地利用形状编码的形状特征的信息。因此, 在形变参数估计网络存在两个相应的函数如等式(5)所示。

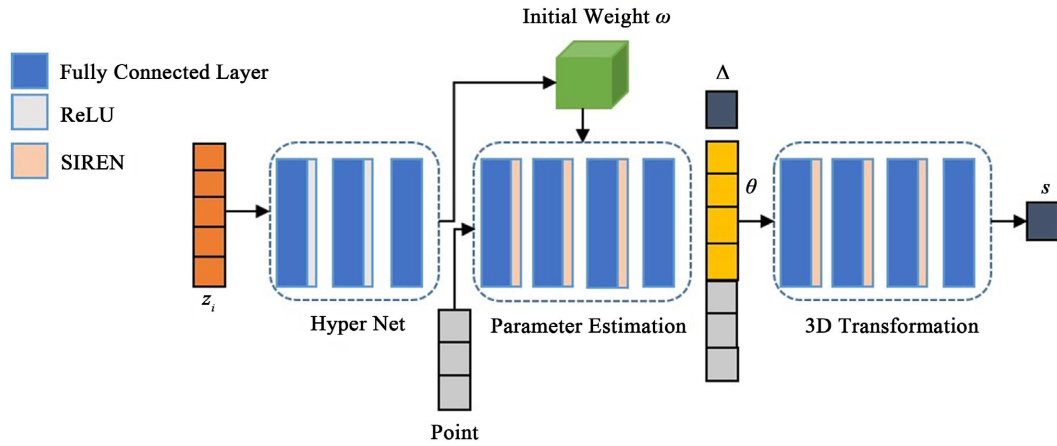


Figure 4. The network of decoder
图 4. 解码器网络

$$\begin{aligned}
 H: z^i \in \mathbb{R}^d &\rightarrow \omega^i \in \mathbb{R}^m \\
 E_\omega: x_j^i \in \mathbb{R}^3 &\rightarrow (\theta_j^i, \Delta_j^i) \in \mathbb{R}^{f+1}
 \end{aligned}
 \tag{5}$$

在完成上述的计算后, 我们设计了第二部分的网络, 将 x_j^i 与 θ_j^i 合并连接, 作为变换网络 T 的输入, 这一部分与[4]类以, 输出该点未加以改善的符号距离函数的值 s_j^i , T 的映射为:

$$T: (\theta_j^i, x_j^i) \in \mathbb{R}^3 \rightarrow s_j^i \in \mathbb{R}^{f+1}
 \tag{6}$$

这里需要注意两点, 第一是由超网络 H 输出的参数是为变换网络 T 进行初始化的, 因为这一部分变换描述的是从物体的共性信息带有物体参数 θ_j^i 而进行的变换, 也就是基于平均形状的转变, 第二点, 也是最为重要的一点是, 该网络的输出并不是点 x_j^i 对应的最终的符号距离的值, 正如在形状先验(2)中描述的那样, 物体的个性特征是无法从变换网络 T 中经过三维变换转化而来的, 因此用 Δ_j^i 加以修正, 所以点 x_j^i 的最终符号距离值我们将其定义为 $s_j^i + \Delta_j^i$ 。

综上所述, 结合 DGCNN 编码器和基于形状先验的解码器模块, 最终的模型形式为: 对于输入的特定的类别的点云 $S^i = \{x_j^i | x_j^i \in \mathbb{R}^3\} \in S$, 隐式神经重建网络 f 为:

$$\begin{aligned}
 f: S^i \in \mathbb{R}^{n \times 3} &\rightarrow \partial \Omega^i \\
 f(S^i) &= \bigcup_{x_j^i \in S^i} \left\{ T \left(x_j^i, E_{\theta_j^i}^{H(D(S^i))} \right) + E_{\Delta_j^i}^{H(D(S^i))} \right\}
 \end{aligned}
 \tag{7}$$

3.4. 损失函数

记输入的特定类别的点云 S^i , ∂S^i 为 S^i 表面上的点, SDF 的损失函数为:

$$L_{sdf} = \sum_{x_j^i \in S^i} \left[\left(f(x_j^i) - \overline{s_j^i} \right) + \left(\left\| \nabla f(x_j^i) \right\|_2 - 1 \right) \right] + \sum_{x_j^i \in \partial S^i} \left(1 - \left\langle \nabla f(x_j^i), \overline{n_j^i} \right\rangle \right) \quad (8)$$

对于左边求和的第一项, $\overline{s_j^i}$ 为 x_j^i 的 SDF 的真值, 第二项中 ∇ 为空间梯度, 对于 $f(x_j^i)$ 梯度, 易从反向传播过程中计算, 这一项是由 Eikonal 方程导出, 是基于 SDF 的隐式神经场中常用的损失函数。第二个求和项是用于约束三维物体表面上的物体的法线, $\overline{n_j^i}$ 为表面点切平面法线的真值。此外, 也使用正则化损失来约束学习的形状编码, 最终的损失函数定义为:

$$L_{reg} = \sum_i \|z_i\|_2^2 \quad (9)$$

$$Loss = \omega_1 L_{sdf} + \omega_2 L_{reg}$$

4. 实验

4.1. 数据集与预处理

本文选取 ShapeNetV2 数据集, 其中包含许多室内常见种类的三维模型(OBJ 格式), 选取其中的 Chair (03001627), Table (04379243)以及 sofa (04256520)分别训练三个模型。但是, 其中的三维模型尺寸规模不一, 且没有相对应的点云格式以及需要监督的 SDF 和法向量的真值信息。为此, 我们首先将所有模型归一化在半径为 $1/1.03$ 的球体内[4]。对于表面上的点, 使用 100 个虚拟相机在单位球体上定期采样, 为每一个归一化模型渲染 100 张虚拟图像, 通过从这些虚拟图像中反向投影深度像素来获得表面点及其 SDF 和法向量的真值。对于自由空间上的点, 在 $[-1,1]^3$ 的立方体中均匀采样, 并通过同样的方法计算相应的 SDF 的值。样本数据如图 5 所示。

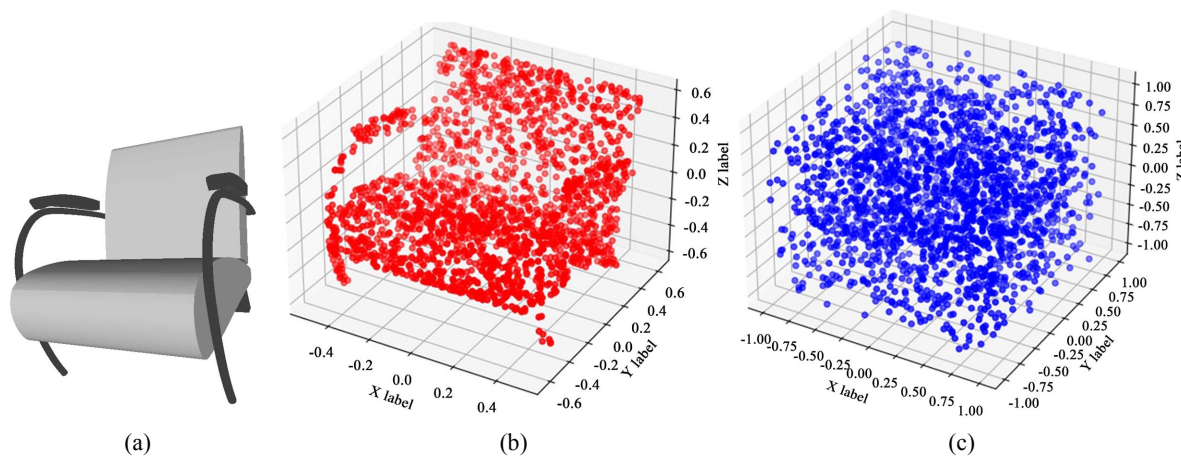


Figure 5. Presentation of source data and preprocessed data

图 5. 源数据与预处理后的数据展示

4.2. 训练细节

在编码器阶段, 我们均匀采样点云表面 5 K 个点, 采用四层边卷积, 使用的 MLP (多层感知机)的输出维度分别为 (64, 64, 128, 256), 每个点选取的近邻点数(K)为 20, 为了抽取多尺度的特征, 将边卷积每一层的输出进行拼接, 得到一个 512 维的向量, 随后使用一层输出为 1024 维的多层感知机聚合特征, 最

后经过二层(512,256)的全连接层获得 256 维的形状编码, 最后两层使用 Droupout (随机失活), 比例为 0.5。此外, 所有的全连接层使用 LeakyRelu 激活和 Batch Normalization (批标准化)。

在解码器阶段, 超网 H 采用 $(256, 512, 512, R^m)$ 的网络, 激活函数使用 Relu, 输出作为参数估计网络 E 的权重, 对于 E , 使用四层(128,128,128,4)全连接网络, 对于变换网络 T , 使用与 E 一致的网络结构, 不过最后的输出是一维, 这两个网络都使用 SIREN 激活[15]。

本文选取的三个数据集的大小都在 3 K ~ 4 K 左右, 在解码器阶段, 参照[4], 训练的关键点是平衡训练集和测试集的正样本和负样本, 具体地说, 选取一半 SDF 值为负的点, 另一半为正, 此外, 降采样 16,384 个自由空间上的点, 以及 16,384 个表面上的点用于一个批次的训练。使用 Adam 优化器, 初始学习率为 0.0001。模型收敛性如图 6 所示。

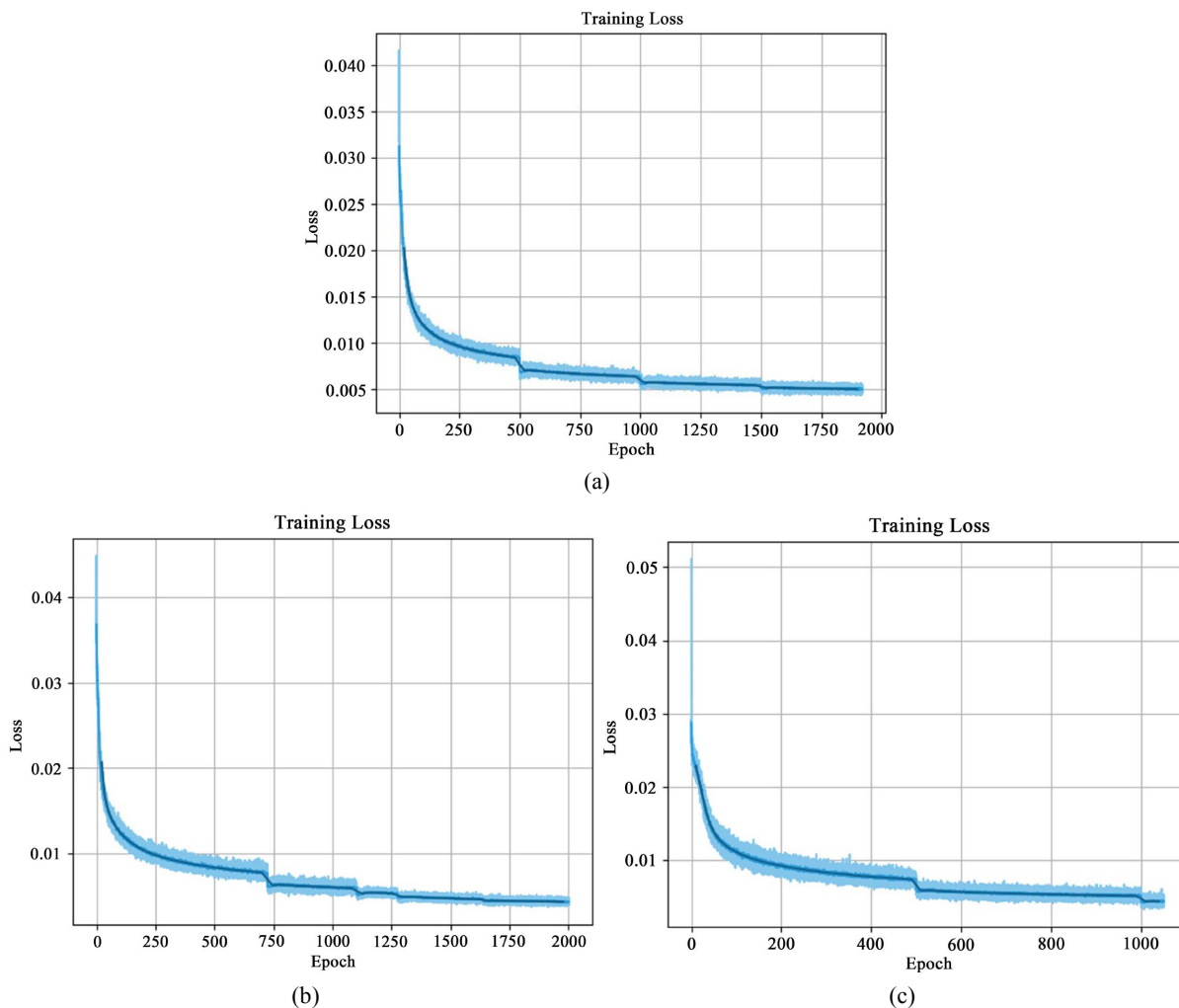


Figure 6. Chair, sofa, table training loss trend
图 6. 椅子、沙发、桌子训练误差趋势

4.3. 衡量指标

4.3.1. 倒角距离

倒角距离(Chamfer Distance)是三维建模任务中最为常见的指标, 给定两个点集 P_1 和 P_2 , 该距离是每

个点到另一个点集中最近点的最近邻距离之和, 定义如下:

$$d_{CD}(P_1, P_2) = \sum_{x \in P_1} \min_{y \in P_2} \|x - y\|_2 + \sum_{y \in P_2} \min_{x \in P_1} \|x - y\|_2 \quad (10)$$

在我们的实验中, 我们在真值表面和由算法得到的表面上, 随机选取了 10 K 个点。

4.3.2. 推土距离

推土距离(Earth Mover's Distance)在三维建模任务中同样是一个常见的评价指标, 其定义如下:

$$d_{EMD}(P_1, P_2) = \sum_{\phi: P_1 \rightarrow P_2} \|x - \phi(x)\|_2 \quad (11)$$

其中 ϕ 是一个双射, 也就意味着该指标旨在寻找 P_1 和 P_2 的一一映射, 再进行对比, 而不是像倒角距离的任意比较, 在计算时, 随机选取表面上 8 K 个点进行 ϕ 的估计与推土距离的计算。

4.3.3. 网格精度

参考[4], 网格精度(Mesh Accuracy)定义为 90%的测试点位于真值表面的半径为 d 的球体内的最小距离 d 。具体地, 我们在测试样本上随机选取 1 K 个点, 并计算相应其到真值表面的最小距离。

4.4. 结果分析

我们选取了 DeepSDF [4]和 AtlasNet [5]网络作为比较对象, 它们都是基于点云的隐式神经三维建模算法, 我们以相同的数据集和样本数量和其源码中给出的超参数设定进行了模型的训练, 使用 200 个样本作为测试样本, 最终在三个类别上的指标如表 1 和表 2 所示。

Table 1. Comparison of average chamfer distance, earth movers' distance and mesh accuracy of test samples

表 1. 测试样本的倒角距离、推土距离、网格精度平均值比较

方法	类别	倒角距离($\times 1$ K)	推土距离	网格精度
AtlasNet	Chair	0.953	0.095	0.041
DeepSDF	Chair	0.523	0.049	0.019
Ours	Chair	0.355	0.036	0.006
AtlasNet	Table	1.131	0.121	0.052
DeepSDF	Table	0.901	0.045	0.024
Ours	Table	0.699	0.038	0.011
AtlasNet	Sofa	1.389	0.177	0.077
DeepSDF	Sofa	0.508	0.057	0.021
Ours	Sofa	0.268	0.037	0.009

Table 2. Comparison of the median of chamfer distance and earth movers' distance of test samples

表 2. 测试样本的倒角距离、推土距离中位数比较

方法	类别	倒角距离($\times 1$ K)	推土距离
AtlasNet	Chair	0.731	0.072
DeepSDF	Chair	0.363	0.038
Ours	Chair	0.209	0.029

Continued

AtlasNet	Table	0.883	0.095
DeepSDF	Table	0.751	0.033
Ours	Table	0.319	0.020
AtlasNet	Sofa	0.972	0.136
DeepSDF	Sofa	0.412	0.041
Ours	Sofa	0.197	0.024

从这些指标中可以看到, 无论是平均值还是中位数, 我们的模型, 相较于 AtlasNet 和 DeepSDF 在每个类别上都拥有着很大的提升。直观地, 从图 7 中可以看出, 我们的方法不仅在三维模型的可视效果、平滑度上都拥有着很好的表现, 而且对于未见过的物体的形状的完成度上也更好, 以此证明了模型的泛化性。特别地, AtlasNet 在对于沙发的重建上, 出现了大量的存在空洞的例子, 从指标分析中, 我们也可以看出该方法相较于我们的方法和 DeepSDF 都有着一定的差距。

5. 结束语

本文提出了一种基于形状先验的隐式神经网络, 联系了类别中各个样本的个性特征和类别整体的共性信息的方法, 并且在实验中证明了该方法可以完成高质量的表面生成任务, 体现了算法的有效性和优越性。

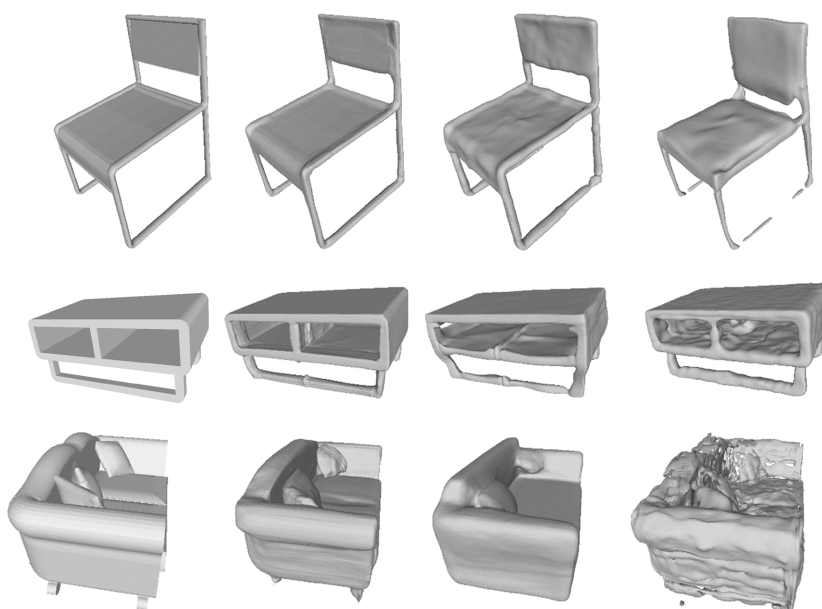


Figure 7. Comparison of experimental results (from left to right, the true value, the proposed algorithm, DeepSDF, and Atlasnet)

图 7. 实验结果比较(从左至右分别是真值、本文提出的算法、DeepSDF 和 Atlasnet)

接下来进一步的工作, 首先, 本文的算法在获得高精度三维表面网格的同时, 训练的耗时比较长, 因为都是全连接层的缘故, 所以考虑进行模型的剪枝或者使用知识蒸馏的方法, 获得一个参数量少一点的模型; 其次, 我们考虑可以从 RGB 图和深度图片作为源数据, 实现从图片到模型的端到端的网络, 其

中的算法, 可以沿用本文的思路, 难点是, 点云到世界坐标系的配准, 以及考虑到图片的像素有限, 有效的三维信息可能较少。

基金项目

自然科学基金联合基金(项目号: U18092006); 自然科学基金(项目号: 62171323); 科技部重点研发项目(项目号: 2020YFA0711400); 上海市重大科技项目(项目号: 2021SHZDZX0100); 长江学者奖励计划; 中央高校基本科研业务费等项目的支持。

参考文献

- [1] Chen, Z.Q. and Hao, Z. (2019) Learning Implicit Fields for Generative Shape Modeling. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 5939-5948. <https://doi.org/10.1109/CVPR.2019.00609>
- [2] Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., et al. (2020) NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M., Eds., *Computer Vision—ECCV 2020. Lecture Notes in Computer Science*, Springer, Cham, 405-421. https://doi.org/10.1007/978-3-030-58452-8_24
- [3] Wang, P., Liu, L.J., Liu, Y., Theobalt, C., Komura, T. and Wang, W.P. (2021) NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-View Reconstruction. *35th Conference on Neural Information Processing Systems (NeurIPS)*, 6-14 December 2021, 123-131.
- [4] Park, J.J., Florence, P., Straub, J., Newcombe, R. and Lovegrove, S. (2019) Deepsdf: Learning Continuous Signed Distance Functions for Shape Representation. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 165-174. <https://doi.org/10.1109/CVPR.2019.00025>
- [5] Groueix, T., Fisher, M., Kim, V.G., Russell, B.C. and Aubry, M. (2018) A Papier-Mache Approach to Learning 3D Surface Generation. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-23 June 2018, 216-214. <https://doi.org/10.1109/CVPR.2018.00030>
- [6] Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S. and Geiger, A. (2019) Occupancy Networks: Learning 3D Reconstruction in Function Space. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 4455-4465. <https://doi.org/10.1109/CVPR.2019.00459>
- [7] Jiang, C., Sud, A., Makadia, A., Huang, J., Nießner, M. and Funkhouser, T. (2020) Local Implicit Grid Representations for 3D Scenes. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 6000-6009. <https://doi.org/10.1109/CVPR42600.2020.00604>
- [8] Charles, R.Q., Su, H., Kaichun, M. and Guibas, L.J. (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. 2017 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 77-85. <https://doi.org/10.1109/CVPR.2017.16>
- [9] Qi, C.R., Yi, L., Su, H. and Guibas, L.J. (2019) Pointnet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. 2017 *31st Conference on Neural Information Processing Systems (NeurIPS)*, Long Beach, 4-9 December 2017, 5105-5114.
- [10] Wang, Y., Sun, Y.b., Liu, Z.w., Sarma, S.E., Bronstein, M.M. and Solomon, J.M. (2019) Dynamic Graph CNN for Learning on Point Clouds. *ACM Transactions on Graphics (TOG)*, **38**, 1-12. <https://doi.org/10.1145/3326362>
- [11] Newcombe, R.A., et al. (2011) Kinect Fusion: Real-Time Dense Surface Mapping and Tracking. 2011 *10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Basel, 26-29 October 2011, 127-136. <https://doi.org/10.1109/ISMAR.2011.6092378>
- [12] Lorensen, W.E. and Cline, H.E. (1987) Marching Cubes: A High Resolution 3d Surface Construction Algorithm. *ACM SIGGRAPH Computer Graphics*, **21**, 163-169. <https://doi.org/10.1145/37402.37422>
- [13] Bao, S.Y., Chandraker, M., Lin Y. and Savarese, S. (2013) Dense Object Reconstruction with Semantic Priors. 2013 *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 23-28 June 2013, 1264-1271. <https://doi.org/10.1109/CVPR.2013.167>
- [14] Tancik, M., et al. (2021) Learned Initializations for Optimizing Coordinate-Based Neural Representations. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 2845-2854. <https://doi.org/10.1109/CVPR46437.2021.00287>
- [15] Vincent, S., et al. (2020) Implicit Neural Representations with Periodic Activation Functions. 2020 *34th Conference on Neural Information Processing Systems (NeurIPS)*, 4-9 December 2020, 7462-7473.