

DOI:10.3969/j.issn.1003-5060.2024.05.010

基于物体单视图的隐式曲面重建

邢燕¹, 牛赛虎¹, 洪沛霖², 檀结庆¹

(1. 合肥工业大学 数学学院, 安徽 合肥 230601; 2. 安徽中医药大学 医药信息工程学院, 安徽 合肥 230012)

摘要: 基于隐式曲面的三维重建方法在保真度、灵活性和压缩能力方面提供了良好的权衡。文章利用隐式曲面网络学习物体形状的三维表面, 首先利用视觉几何群(visual geometry group-16, VGG-16)网络从图像中提取全局特征, 对建模空间中的每个采样点从 VGG-16 网络中获取局部特征; 其次对每个采样点利用多层感知器(multi-layer perceptron, MLP)进行位置编码得到点特征; 然后将全局特征和局部特征分别与点特征串联起来送入 2 个解码器中, 获得隐式场中采样点的符号距离函数(signed distance function, SDF)的大小与符号, 并最终得到物体的隐式曲面。文中所提出的方法应用于 ShapeNet 数据集上进行三维对象重建任务, 定性和定量评估均优于现有方法, 特别是对于具有孔洞和薄结构的复杂拓扑物体。

关键词: 三维重建; 全局特征; 局部特征; 深度学习; 符号距离函数(SDF)

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1003-5060(2024)05-0642-07

Implicit surface reconstruction based on object single view

XING Yan¹, NIU Saihu¹, HONG Peilin², TAN Jieqing¹

(1. School of Mathematics, Hefei University of Technology, Hefei 230601, China; 2. School of Medical Information Engineering, Anhui University of Chinese Medicine, Hefei 230012, China)

Abstract: The three-dimensional(3D) reconstruction method based on implicit surface provides a good trade-off in terms of fidelity, flexibility and compression ability. In this paper, the implicit surface network is used to learn the 3D surface of an object shape. Firstly, the global feature is extracted from the image by using visual geometry group-16 (VGG-16) network, and the local feature is obtained from the VGG-16 network for each sampling point in the modeling space. Then, each sampling point is encoded by a multi-layer perceptron(MLP) to obtain the point feature. Furthermore, the global feature and the local feature are concatenated with the point feature respectively and fed into the two decoders respectively to obtain the magnitude and sign of the signed distance function(SDF) of the sampling point in the implicit field. Finally, the implicit surface of the object is obtained. The proposed method has performed 3D object reconstruction task on ShapeNet datasets with both qualitative and quantitative evaluations superior to state-of-the-art methods, especially for complex topological objects with holes and thin structures.

Key words: three-dimensional(3D) reconstruction; global feature; local feature; deep learning; signed distance function(SDF)

基于单视图图像的三维重建是一个具有挑战性的任务。随着机器学习的发展, 利用神经网络学习三维重建受到了学者们的广泛关注。现

有的方法通过体素^[1-3]、点云^[4-6]和网格^[7-11]等表示获得了不同程度的成功。体素可以看作二维像素在三维空间上的扩展, 易于在神经网络上编码。

收稿日期: 2023-02-20; **修回日期:** 2023-04-02

基金项目: 国家自然科学基金资助项目(62172135); 合肥工业大学校级教研资助项目(KCSZ2022034)和安徽中医药大学教研重点资助项目(2020xjyy_zd005)

作者简介: 邢燕(1977—), 女, 安徽合肥人, 博士, 合肥工业大学副教授, 硕士生导师;
檀结庆(1962—), 男, 安徽望江人, 博士, 合肥工业大学教授, 博士生导师。

文献[1]率先提出通过递归神经网络从单视图重建三维体素形状,然而当给定同一组不同视角的输入图像时,该方法无法产生收敛的重建结果;文献[2]在文献[1]的基础上提出感知融合模块和细化模块使重建结果更加真实;文献[3]在文献[2]的基础上,增加多尺度连接和通道注意力机制,保留多尺度信息,强化重点特征学习的同时提出一种融合型损失函数,减小不平衡数据与类间差异对重建效果的影响,虽然体素重建能获得一个较好的结果,但是对于高分辨率三维物体重建内存消耗过大;文献[4-5]提出通过深度学习从单视图重建三维点云;文献[6]使用点云重建豆株形状,点云重建虽然节省了内存,但是难以重建表面细节,而且点与点之间没有连接性。为了克服上述缺点,有些学者提出用网格表示三维形状,网格是一种高效且连续的曲面表示,它将孤立的点(即网格顶点)与网格边和面连接起来。

学习曲面网格重建的挑战在于学习网格顶点之间的连接关系。为了解决这一问题,文献[7-8]提出从一个固定拓扑的网格模板出发,逐渐变形为对应物体的三维网格形式,虽然取得良好的结果,但是重建结果仅限于与模板模型相同的拓扑结构,当目标对象具有不同的拓扑时,导致较大的重建误差。为了解决这一问题,文献[9]提出从一组参考模型中寻找初始网格结构,虽在一定程度上优于单个模板变形结果,但在复杂拓扑的曲面网格重构中仍存在局限性;文献[10]提出通过对变形网格表面进行裁剪细化以达到重建复杂拓扑的结果,虽然能够重建复杂拓扑结构,但是裁剪阈值的选取影响拓扑重建的效果;文献[11]在文献[10]的基础上提出使用自适应阈值策略优化网格拓扑裁剪,相比于固定阈值,自适应阈值策略具有更好的泛化重建能力,但是最终得到的仍是一个不封闭的网格。

对于所有的网格变形方法,在变形过程中动态学习顶点连接关系是一项相当具有挑战性的任务,因此上述方法对薄结构的重建效果较差。为了解决上述体素、点云和网格中的限制,另外一种三维形状表示——隐式三维曲面表示被提出。隐式曲面表示^[12-16]在空间上与点云表示具有相似的优势,且能重建出具有复杂拓扑结构的连续表面。文献[12-14]通过二值占用网格或者连续的符号距离函数生成隐式三维曲面,虽然能重建一个相对较好的拓扑结果,但是对于细节丰富的区域,重建结果较差。其主要原因是二值占用网格对高分

辨率难以处理,虽然连续的符号距离函数表示更好一些,但是仅靠提取的全局特征重建的细节区域还不够。文献[15]提出增加局部特征信息,改进了细节区域的重建,文献[16]通过改进文献[15]的局部特征提取方式,使重建结果进一步提升。

虽然现有的方法已经能从单视图重建物体的全局三维形状,但是对于细节区域重建效果并不理想,如孔洞和薄结构等。这种细粒度的细节在三维空间中只占据一小部分,在定量度量上舍去它们影响并不大,但是在定性比较上可以很明显地看到它们与对应真实形状之间的差别,导致视觉效果不足。为了解决这个问题,本文引入了一个局部特征提取模块。利用局部特征预测查询点的符号距离函数(signed distance function, SDF)的绝对值,降低全局特征中无关特征对查询点的绝对距离影响,用全局特征预测其 SDF 符号,然后将绝对距离与符号相乘得到对应点的 SDF 值。这种将局部特征与全局特征结合求 SDF 值的方法,显著提高了生成的三维形状中细粒度细节的重建质量,重建出现有的一些方法无法生成的孔洞或者薄结构。

1 方 法

本文使用 SDF 建模三维表面。SDF 是空间中的一个有限区域上确定一个点到物体边界的有符号的距离,点在物体内部为负,外部为正,位于边界上时为 0。给定空间中的一个查询点 $x \in \mathbf{R}^3$,将其作为 SDF 的输入,可得一个实数 $y = f(x) \in \mathbf{R}$ 。 y 的绝对值大小表示查询点到物体表面的最短距离, y 的符号表示点在物体表面的内部(负),还是外部(正), $y=0$ 表示查询点在物体表面上,如图 1 所示。

令 S 表示物体的三维表面,它是符号距离函数的零水平集。

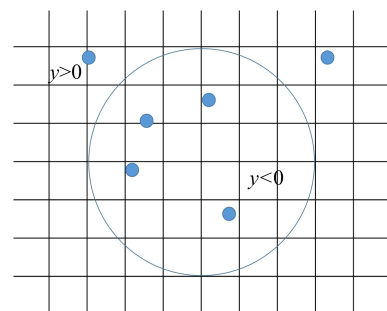


图 1 符号距离函数的符号说明

本文的任务是输入一张彩色图像,根据深度神经网络从图像中提取到的特征,预测空间中任意点的 SDF 值,从而重建图像中物体的三维表面,力求忠实地还原物体的三维形状。本文以单幅图像作为输入,预测空间中任意点的 SDF 值。由于隐式场的连续性,本文方法的表面重建是平滑的,理想情况下可获得无限的分辨率。本文发现局部特征对于重建细节区域有很好的效果,因此采取局部特征与全局特征相结合的 SDF 深度学习方法在获得了查询点符号距离函数值后,使用文献[17]的方法识别等值面并提取形状的三维网格。

本文方法的流程图如图 2 所示。输入一张彩色图像,先用卷积神经网络提取它的特征向量,再将查询点升维得到的点特征分别与图像的全局特征向量和局部特征向量连接,然后用双流的隐式曲面网络分别进行解码并结合,最后预测出查询点的 SDF 值。与文献[15]的 DISN 方法不同,其用全局特征和局部特征分别学习 1 个 SDF 值,然后将 2 个 SDF 值相加得到预测的 SDF 值;本文方法通过全局特征学习符号,局部特征学习点到表面的距离,然后将空间中对对应点的符号与距离相乘,得到预测的 SDF 值。

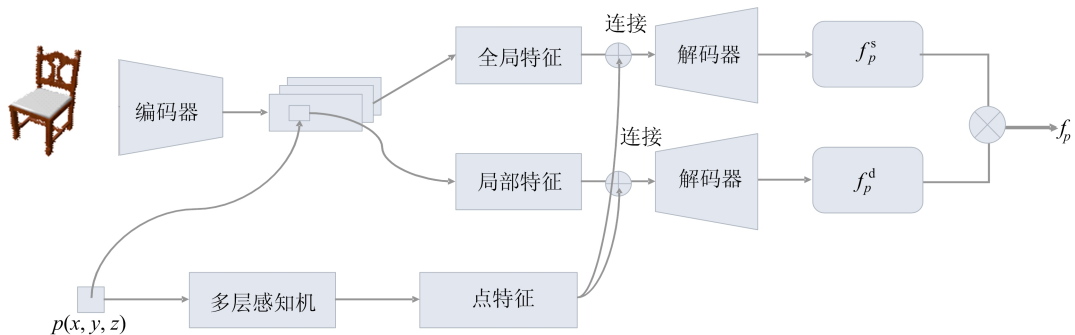


图 2 单视图隐式曲面重建网络的管线图

1.1 局部预测点到表面距离

对于物体的三维表面 S , 本文将表面表示为符号距离函数(SDF) f 的零集, 即

$$S = \{x \in \mathbf{R}^3 \mid f(x) = 0\} \quad (1)$$

文献[12-14]的研究结果表明, 这种连续场的隐式表示特别适用于神经网络。神经网络可以被训练成函数, 将图像作为输入, 并输出 SDF 的近似值:

$$f(x) \approx f_p(x) = F_\theta(x \mid z) \quad (2)$$

$$z = E_\varphi(I) \quad (3)$$

其中: I 为输入图像; z 为通过编码器 E 对图像 I 进行编码所得到的结果, 即图像中物体的潜在描述; F 为以潜在向量 z 为条件的神经网络实现; 解码器网络 F 和编码器 E 分别由 θ 和 φ 参数化。要预测空间中一个查询点 x 到物体三维表面的绝对距离 $f^d(x)$, 只需要查询点对应的局部特征:

$$z_x^l = E_\varphi^l(I) \quad (4)$$

$$f_p^d(x) = F_{\theta_l}^d(x \mid z_x^l) \quad (5)$$

其中: z_x^l 为查询点 x 在图像中对应局部位置的编码; φ 为编码器的参数; $F_{\theta_l}^d$ 为预测查询点绝对距离的局部解码器; θ_l 为局部解码器的参数。利用局部特征估计查询点的 SDF 绝对距离比利用全局特征估计查询点的绝对距离更准确, 这是由于

局部编码 z_x^l 比全局编码 z^G 更集中地表示查询点周围的信息。本文将在 2.2 节描述局部特征的提取。

1.2 全局预测查询点 SDF 符号

相比于提取局部特征预测点到表面的距离, 需要一个局部特征提取器, 采用全局特征预测点的符号。为了预测空间中查询点 x 处的符号 $f^s(x)$, 本文需要物体的全局信息。由于查询点在物体内部还是外部根据局部信息是不易判断的, 本文用卷积神经网络从输入图像提取全局特征:

$$z^G = E^G G_\varphi(I) \quad (6)$$

$$f_p^s(x) = \text{sgn}(F_{\theta_g}^s(x \mid z^G)) \quad (7)$$

其中: z^G 为图像中物体全局信息的编码; I 为输入图像; φ 为编码器的参数; $F_{\theta_g}^s$ 为预测查询点 SDF 符号的全局解码器; θ_g 为全局解码器的参数。

从而得出本文方法的最终公式为:

$$f(x) \approx f_p(x) = f_p^d(x) * f_p^s(x) \quad (8)$$

在训练过程中, 本文的数据集给出了真实的 SDF, 本文网络预测的 $f_p(x)$ 目标就是逼近真实的 $f(x)$ 。这与之前的三维重建方法常用预测形状与真实形状之间的 Chamfer 距离和 EMD 损失不同, 本文使用预测的 SDF 值与真实值之间的 L_1 损失。

2 网络模型和训练

2.1 模型流程

本文所提出的模型的流程图如图 2 所示。首先输入 1 张二维图像,经过编码器提取出其全局特征,然后根据本文的局部特征提取模块提取出查询点投影的局部特征。对于查询点的位置,本文使用一个多层感知机将给定的点位置映射到一个更高维的特征空间,然后将点的特征与全局特征和局部特征连接送入 2 个解码器中,解码出对应点的 SDF 绝对距离与 SDF 符号,将其相乘得到本文预测的 SDF 值。

2.2 全局特征和局部特征提取

提取图像全局特征的方法已经发展的相当成熟,本文使用现有的视觉几何群(visual geometry group-16, VGG-16)网络对输入图像提取全局特征。单个潜在编码容量有限,使得全局特征在重建过程中虽然能描述物体的大致形状,但是对于细节丰富的区域,无法进行高质量的重建。因此本文使用一个局部特征提取模块,提取查询点附近信息,以提高对物体细节的重建。本文特征提取过程如图 3 所示。

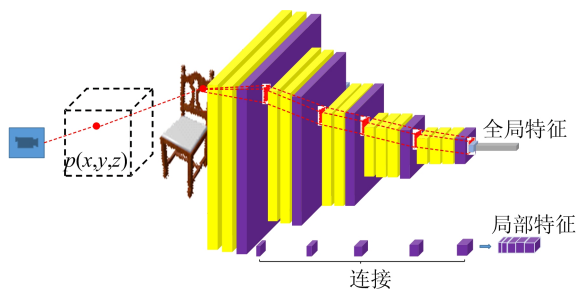


图 3 图像特征提取

采用的特征提取网络是 VGG-16 网络,它含有多个 CNN 层,每层将输出一个特征图,浅层特征图反映了局部特征,深层特征图对应抽象的语义特征。局部特征提取过程如下:首先将查询点 $p \in \mathbf{R}^3$ 投影到图像平面上得到点 $q \in \mathbf{R}^2$,然后由浅到深抽取多个卷积层输出的特征图上点 q 位置对应的特征,并将它们连接起来,得到该点对应的局部特征。由于后面层的特征图在尺寸上比前面层的特征图要小,本文通过双线性插值将其调整为原图大小,并在点 q 位置处提取调整后的特征。

2.3 点特征提取和特征融合

对于查询点,本文没有直接把它三维坐标和全局特征、局部特征进行连接去预测 SDF 值,

而是首先使用一个多层感知器(multi-layer perceptron, MLP)进行位置编码,将点位置映射到一个更高维的特征空间。因为低维信息经过神经网络往往只能得到低频信号,所以本文采用 MLP 对其升维,再与全局和局部特征结合,重建出细节丰富的三维形状。

文献[14]提出的 DeepSDF 网络将查询点的位置与从深度图像或点云提取的形状嵌入连接起来,并使用自动解码器获得相应的 SDF 值,其自动解码器的结构需要优化每个输入的形状嵌入。本文以前馈的方式把点的位置与从图像中提取的形状特征直接连接送入解码器网络,会存在收敛问题。

2.4 损失函数

将图 2 中符号函数值 $f_p^s(\mathbf{x})$ 和距离函数值 $f_p^d(\cdot)$ 相乘得到本文所需要的符号距离函数值 $f_p(\mathbf{x})$ 。SDF 是一个连续函数,因此本文可以提取空间中的任意等值面,但是为了确保本文的网络训练集中在零等值面附近,损失函数中本文使用截断符号距离函数。损失函数定义为:

$$L_{\text{TSDf}} = \sum_{\mathbf{x} \in \mathbf{R}^3} w_x \| f_p(\mathbf{x}) - f(\mathbf{x}) \| \quad (9)$$

其中: $\| \cdot \|$ 为 L_1 范数; w_x 为权重,即

$$w_x = \begin{cases} 1, & |f(\mathbf{x})| < t; \\ 0, & \text{其他。} \end{cases}$$

当空间中查询点的 SDF 绝对值大于一定阈值 t 时,本文使用权重 0。

最终本文的总损失为:

$$L = L_{\text{TSDf}} \quad (10)$$

2.5 表面重建

本文在空间中定义了一个稠密的三维网格,并预测网格中点的 SDF 值,其零等值面就是本文求得的隐式曲面。为了生成图像中对应物体的三维网格表面,本文可以对其零等值面使用文献[17]方法获得。

3 实 验

本节对单视图三维重建进行实验,给出了与文献[7-9, 12-13, 15]方法定量和定性的比较结果,并进行了消融研究,结果表明本文特征融合的方法对于具有丰富细节对象重建的重要性。

3.1 数据集

对于 SDF 预测,本文采用文献[7-9, 12]的设置,并使用 ShapeNet 核心数据集[18],其中包括 13 个对象类别以及 1 个给定的训练/测试分割方

式训练和测试本文的方法。对于二维图像,本文使用文献[1]提供的渲染视图作为输入,本文在所有类别上训练单个网络,并给出由该网络生成的测试结果。

3.2 数据准备和实施细节

本文使用的 ShapeNet 数据集 SDF 真值根据文献[15]的提议,使用文献[19-20]方法生成,是分辨率为 256^3 的 SDF 网格。文献[1]中的每张图片视为不同相机视角下的渲染图。

本文使用 VGG-16 作为图像编码器,当训练 SDF 预测网络时,本文使用真实摄像机参数提取局部特征。尽管本文能够通过连续采样点并回归其 SDF 值生成具有任意分辨率的符号距离场,然而实验中本文由于只关注零等值面附近的点,训练过程中使用蒙特卡罗抽样选择了高斯分布 $N(0, 0.1)$ 下的 2 048 个网格点。本文选择 $t=0.3$ 作为损失函数的参数。实验中优化算法为 Adam 算法,学习率为 1×10^{-4} ,批处理大小为 16。

3.3 评价指标

对于定量评价,本文使用 3 种常用的度量计算重建网格对象与其真值网格之间的差异:

① Chamfer距离(Chamfer distance, CD);② 搬土

距离(earth mover's distance, EMD);③ 体素化网格上的交并比(intersection over union, IoU)。

3.4 本文方法与现有方法的比较

本节比较了本文的单视图重建方法与 Pixel2Mesh^[7]、AtlasNet^[8]、3DN^[9]、OccNet^[12]、ImNet^[13]、3D CNN^[15] 和 DISN^[15] 等方法。Pixel2Mesh 和 AtlasNet 从一个二维图像中生成一个固定的拓扑网格。3DN 使给定的源网格发生变形,以重建目标模型。OccNet 和 ImNet 都可以预测 SDF 的符号重建三维形状。3D CNN 和 DISN 都是运用 VGG-16 网络编码图像特征,其中 3D CNN 使用形状嵌入生成一个固定的三维体积,而 DISN 使用形状嵌入预测 SDF 值重建三维形状。本文给出这些方法的结果,并与之进行对比。

AtlasNet、Pixel2Mesh 和 3DN 使用显式曲面生成,而 3D CNN、ImNet、OccNet 使用与本文相同的表面重建方法从重建隐式曲面输出重建网格。

3.4.1 定性比较

本文方法与其他重建方法的定性比较如图 4 所示。

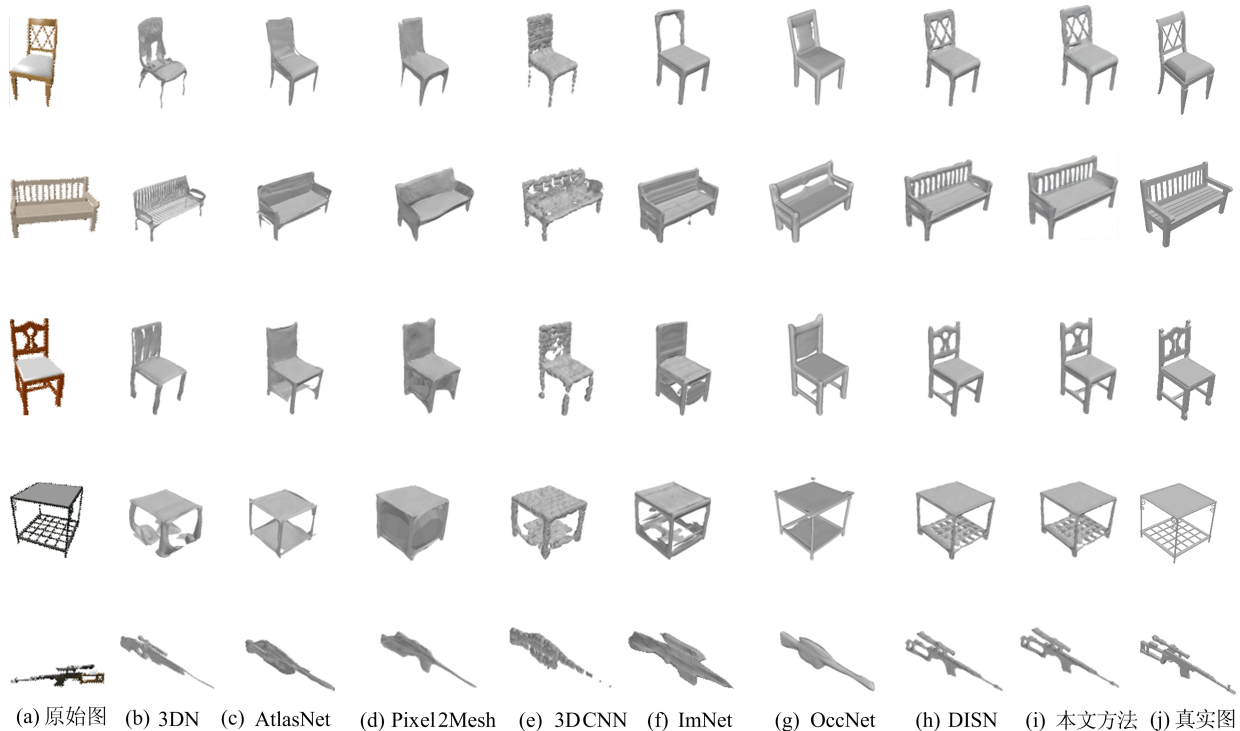


图 4 本文方法和其他方法的定性比较

图 4 结果表明隐式曲面重建对于拓扑结构复杂的物体比显示曲面重建效果更好,这是由于显

示表面重建多从一个固定拓扑的模板出发变形得到,而 SDF 是一个连续函数,可以构建任意精度

的拓扑结构。与 3D CNN 的比较表明,本文方法产生的表面更平滑,这是由于 SDF 是一个关于点位置的连续函数,深度网络很难用形状嵌入来近似整体体积。本文方法在恢复形状细节方面优于 ImNet 和 OccNet,本文方法能准确地重建椅子背部的孔洞细节、桌子下面的网状隔板和步枪的枪托孔洞,而 ImNet 和 OccNet 则无法捕捉到这些细节。相比于 DISN,本文方法在一些细节区域重建更加平滑,如椅子和桌子的腿、步枪枪身等。

3.4.2 定量比较

本文方法与其他方法的定量比较见表 1 所列。本文方法在 EMD 指标上优于所有其他重建方法,其中 DISN 与本文方法最为接近;在 IoU 指标上,DISN 最好,本文方法与 DISN 不相上下,均优于其他重建方法;在 CD 指标上 3DN 最好,其次 DISN,然后是本文的方法,这可能是由于 3DN 的输入是一个与目标网格相似的待变形网格,比本文的二维图像输入具有更多的信息。

表 1 本文方法和其他方法的定量比较

指标	方法	飞机	长凳	盒子	汽车	椅子	显示器	灯	扬声器	步枪	沙发	桌子	手机	船	均值
EMD	AtlasNet	3.39	3.22	3.36	3.72	3.86	3.12	5.29	3.75	3.35	3.14	3.98	3.19	4.39	3.67
	Pixel2Mesh	2.98	2.58	3.44	3.43	3.52	2.92	5.15	3.56	3.04	2.70	3.52	2.66	2.94	3.34
	3DN	3.30	2.98	3.21	3.28	4.45	3.91	3.99	4.47	2.78	3.31	3.94	2.70	3.92	3.56
	ImNet	2.90	2.80	3.14	2.73	3.01	2.81	5.85	3.80	2.65	2.71	3.39	2.14	2.75	3.13
	3D CNN	3.36	2.90	3.06	2.52	3.01	2.85	4.73	3.35	2.71	2.60	3.09	2.10	2.67	3.00
	DISN	2.67	2.48	3.04	2.67	2.67	2.73	4.38	3.47	2.30	2.62	3.11	2.06	2.77	2.84
	本文方法	2.59	2.40	3.10	2.7	2.60	2.75	4.09	3.48	2.34	2.55	3.05	2.08	2.69	2.80
CD	AtlasNet	5.98	6.98	13.76	17.04	13.21	7.18	38.21	15.96	4.59	8.29	18.08	6.35	15.85	13.19
	Pixel2Mesh	6.10	6.20	12.11	13.45	11.13	6.39	31.41	14.52	4.51	6.54	15.61	6.04	12.66	11.28
	3DN	6.75	7.96	8.34	7.09	17.53	8.35	12.79	17.28	3.26	8.27	14.05	5.18	10.20	9.77
	ImNet	12.65	15.10	11.39	8.86	11.27	13.77	63.84	21.83	8.73	10.30	17.82	7.06	13.25	16.61
	3D CNN	10.47	10.94	10.40	5.26	11.15	11.78	35.97	17.97	6.80	9.76	13.35	6.30	9.80	12.30
	DISN	9.96	8.98	10.19	5.39	7.71	10.23	25.79	17.90	5.58	9.16	13.59	6.40	11.91	10.98
	本文方法	10.08	8.96	9.95	5.67	7.76	10.65	25.57	18.65	5.48	9.97	13.26	6.65	11.75	11.11
IoU	AtlasNet	39.2	34.2	20.7	22.0	25.7	36.4	21.3	23.2	45.3	27.9	23.3	42.5	28.1	30.0
	Pixel2Mesh	51.5	40.7	43.4	50.1	40.2	55.9	29.1	52.3	50.9	60.0	31.2	69.4	40.1	47.3
	3DN	54.3	39.8	49.4	59.4	34.4	47.2	35.4	45.3	57.6	60.7	31.3	71.4	46.4	48.7
	ImNet	55.4	49.5	51.5	74.5	52.2	56.2	29.6	52.6	52.3	64.1	45.0	70.9	56.6	54.6
	3D CNN	50.6	44.3	52.3	76.9	52.6	51.5	36.2	58.0	50.5	67.2	50.3	70.9	57.4	55.3
	OccNet	54.7	45.2	73.2	73.1	50.2	47.9	37.0	65.3	45.8	67.1	50.6	70.9	52.1	56.4
	DISN	57.5	52.9	52.3	74.3	54.3	56.4	34.7	54.9	59.2	65.9	47.9	72.9	55.9	57.0
本文方法	57.9	52.6	49.8	75.5	52.9	56.6	34.6	56.3	58.0	67.3	50.4	70.9	57.4	56.9	

3.5 消融研究

为了分析局部特征提取的影响,本文在 ShapeNet“椅子”类别上进行了消融研究,其原因是椅子的拓扑结构具有多样性,结果如图 5 所示。

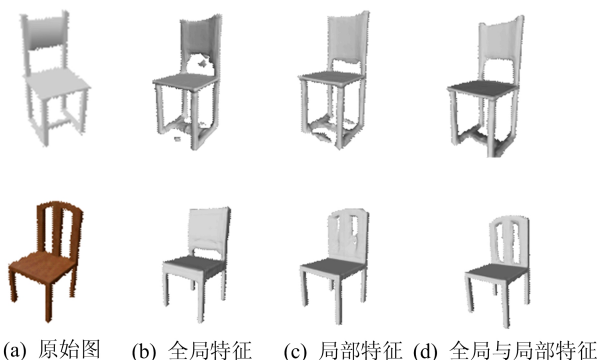


图 5 消融研究

实验结果表明:单流的全局特征重建能够恢复出物体的大致形状,但细节可能严重缺失;单流的局部特征重建可能会引入伪影,细节也会有所不足;而局部特征和全局特征相结合的双流模型获得了令人满意的结果。

4 结 论

本文提出了一种从单视图重建物体三维表面的深度学习方法。针对输入的图像,本文的网络根据空间中给定的查询点,预测该点的 SDF 值。通过改变采样点的数目,本文方法可生成任意分辨率的三维曲面。此外,本文方法不需要一个固定的拓扑假设即可生成与真实形状相同的拓扑结构。相比于用单个潜在向量编码全局特征的方法,本文方法将全局特征和局部特征相结合,分别

编码了 2 个潜在向量,有效地提高了重建的保真度。局部特征的添加,使得本文方法对孔洞或薄结构等细节丰富区域有了更好的视觉重建结果。定性和定量的实验表明,本文方法与现有方法相比具有竞争力。

虽然本文方法能从单视图重建出令人满意的结果,但是也存在着限制,如训练过程是用合成的图像数据来完成的,实验中处理的是清晰背景的图片。本文未来的工作将是增加真实世界的图片进行训练,通过纹理预测增加物体三维表面重建的真实性。

[参 考 文 献]

- [1] CHRISTOPHER B C, DANFEI X, JUNY G, et al. 3D-R2N2: a unified approach for single and multi-view 3D object reconstruction [C]//Proceedings of the European Conference on Computer Vision. [S. l. : s. n.], 2016: 628-644.
- [2] XIE H Z, YAO H X, SUN X S, et al. Pix2Vox: context-aware 3D reconstruction from single and multi-view images [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S. l.]; IEEE, 2019: 2690-2698.
- [3] 何鑫睿, 李秀梅, 孙军梅, 等. 基于改进 Pix2Vox 的单图像三维重建网络 [J]. 计算机辅助设计与图形学学报, 2022, 34(3): 364-372.
- [4] FAN H, SU H, GUIBAS L J. A point set generation network for 3D object reconstruction from a single image [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2017: 605-613.
- [5] ACHLIOPTAS P, DIAMANTI O, MITLIAGKAS I, et al. Learning representations and generative models for 3D point clouds [C]//International Conference on Machine Learning. [S. l. : s. n.], 2018: 40-49.
- [6] 傅汝佳, 洗楚华, 李桂清, 等. 面向表型精确鉴定的豆株快速三维重建 [J]. 浙江大学学报(理学版), 2021, 48(5): 531-539.
- [7] WANG N Y, ZHANG Y D, LI Z W, et al. Pixel2mesh: generating 3D mesh models from single RGB images [C]//Proceedings of the European Conference on Computer Vision. [S. l. : s. n.], 2018: 52-67.
- [8] GROUEIX T, FISHER M, KIM V G, et al. A papier-mache approach to learning 3D surface generation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2018: 216-224.
- [9] WANG W, CEYLAN D, MECH R, et al. 3DN: 3D deformation network [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2019: 1038-1046.
- [10] PAN J, HAN X, CHEN W, et al. Deep mesh reconstruction from single RGB images via topology modification networks [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S. l.]; IEEE, 2019: 9964-9973.
- [11] NIE Y, HAN X, GUO S, et al. Total 3D understanding: joint layout, object pose and mesh reconstruction for indoor scenes from a single image [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2020: 55-64.
- [12] MESCHEDER L, OECHSLE M, NIEMEYER M, et al. Occupancy networks: learning 3D reconstruction in function space [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2019: 4460-4470.
- [13] CHEN Z, ZHANG H. Learning implicit fields for generative shape modeling [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2019: 5939-5948.
- [14] PARK J J, FLORENCE P, STRAUB J, et al. DeepSDF: learning continuous signed distance functions for shape representation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]; IEEE, 2019: 165-174.
- [15] XU Q, WANG W, CEYLAN D, et al. DISN: deep implicit surface network for high-quality single-view 3D reconstruction [J]. Advances in Neural Information Processing Systems, 2019, 32: 492-502.
- [16] 王钊. 基于单视图三维重建研究 [D]. 成都: 电子科技大学, 2020.
- [17] LORENSEN W E, CLINE H E. Marching cubes: A high resolution 3D surface construction algorithm [J]. ACM SigGraph Computer Graphics, 1987, 21(4): 163-169.
- [18] CHANG A X, FUNKHOUSER T, GUIBAS L, et al. ShapeNet: an information-rich 3D model repository [EB/OL]. [2024-04-28]. <http://arXiv/pdf/1512.03012.pdf>.
- [19] XU H Y, BARBIC J. Signed distance fields for polygon soup meshes [C]//Proceedings of Graphics Interface 2014. Boca Raton: CRC Press Inc., 2020: 35-41.
- [20] SIN F S, SCHROEDER D, BARBIC J. Vega: nonlinear FEM deformable object simulator [C]//Computer Graphics Forum. [S. l. : s. n.], 2013: 36-48.

(责任编辑 李 凯)