

DOI: 10.13718/j.cnki.xdzk.2025.05.021

汪伟, 郑丹丹. 基于深度学习的手绘草图三维建模自动化算法研究和应用 [J]. 西南大学学报(自然科学版), 2025, 47(5): 242-252.

基于深度学习的手绘草图三维建模 自动化算法研究和应用

汪伟¹, 郑丹丹²

1. 武汉工商学院 艺术与设计学院, 武汉 430065; 2. 中国地质大学(武汉) 艺术与传媒学院, 武汉 430074

摘要: 与基于图像的建模方法相比, 基于草图的三维(3D)模型设计为用户提供了更直观的生成方式, 可满足手绘草图的设计愿景。然而, 由于手绘草图中的信息通常较为稀疏且模糊, 从中创建准确的 3D 形状设计成为一项充满挑战的任务。为了解决这一问题, 提出一种基于深度学习的手绘草图 3D 建模自动化算法并应用于汽车模型设计。该系统充分利用用户提供的手绘草图输入, 通过深度学习技术估计目标 3D 模型的无符号距离字段(UDF)。采用行进立方体算法(Marching Cubes), 系统从估计的 UDF 字段中提取汽车模型网格。为增加用户的定制灵活性, 还引入了模型编辑功能, 允许用户对生成的网格进行灵活修改。通过对汽车模型生成和编辑的定量评估, 并与当前领先方法的比较表明: 该系统在生成高质量汽车模型方面具有显著优势, 仅使用单一草图就成功生成了不同类型的汽车模型, 且这些模型具有复杂的局部特征。与其他机器学习结构相比, 此方法展现出更高的性能水平, 为智能处理手绘草图提供了创新思路和强大工具, 突显了深度学习在手绘草图 3D 建模中的关键作用。

关键词: 手绘草图; 三维建模; 三维汽车模型; 深度学习;

自动化

中图分类号: TP391

文献标志码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2025)05-0242-11

Research and Application of a Deep Learning-Based Automated 3D Modeling Algorithm for Hand-Drawn Sketches

WANG Wei¹, ZHENG Dandan²

1. School of Art & Design, Wuhan Technology and Business University, Wuhan 430065, China;

2. School of Arts and Communication, China University of Geosciences (Wuhan), Wuhan 430074, China

Abstract: Compared to image-based modeling approaches, sketch-based 3D model design provides users

收稿日期: 2024-03-18

基金项目: 国家重点研发计划项目(2023YFC3709900); 湖北省教育厅教育科学规划项目(2019GB136); 武汉工商学院校级科学研究项目(A2020013)。

作者简介: 汪伟, 硕士, 讲师, 主要从事数字媒体艺术及城市街道景观设计研究。

with a more intuitive way to realize their design visions captured in hand-drawn sketches. However, due to the typically sparse and ambiguous nature of hand-drawn sketches, accurately creating 3D shape designs from them poses a challenging task. To address this challenge, this study proposed a deep learning-based automation algorithm for the three-dimensional modeling of hand-drawn sketches, and applied it specifically to the design of automotive models. The system maximizes the utilization of user-provided hand-drawn sketch inputs by employing deep learning techniques to estimate the unsigned distance field (UDF) of the target 3D model. Utilizing the Marching Cubes algorithm, the system intelligently extracts the grid of the automotive model from the estimated UDF field. To enhance user customization flexibility, model editing functionality is introduced, allowing users to make flexible modifications to the generated grid. Through quantitative evaluations of automotive model generation and editing, comparisons with current state-of-the-art methods, the system's significant advantages in producing high-quality automotive models is revealed. The system successfully generated diverse types of automotive models using only a single sketch input. These models exhibited intricate local features. In comparison to other machine learning structures, this approach demonstrates superior performance levels. This research contributes innovative perspectives and powerful tools for intelligently handling hand-drawn sketches, emphasizing the crucial role of deep learning in three-dimensional modeling of hand-drawn sketches.

Key words: hand-drawn sketches; 3D modeling; 3D car models; deep learning; automation

手绘草图是一种普遍的交流和艺术方式,它跨越障碍,将人类社会联系在一起,是人类最原始、最直接的表达方式之一,无需特定的工具和技能,只需一支笔、一张纸就能将脑海中的想法和情感生动地呈现出来。在远古时期,我们的祖先就已经开始在洞穴墙壁和陶器上绘制各种图形和符号,用以记录生活、表达情感和传递信息。随着人类社会的发展,手绘草图逐渐演变成一种重要的沟通工具,被广泛应用于各个领域。无论是科学家、工程师、建筑师,还是艺术家、设计师、教师等都可以通过手绘草图来表达自己的想法和创意。手绘草图具有很强的直观性和表现力,能够将复杂的概念和抽象的想法变得具体而形象。它能够跨越语言和文化障碍,让人们在没有文字的情况下也能够理解和交流。特别是在现代社会,随着全球化的推进,手绘草图作为一种国际语言越来越受到人们的重视。与专业素描、法医素描、漫画、技术绘画和油画等其他相关表现形式不同,手绘草图不需要训练,也不需要特殊设备,因此不受年龄、种族、语言、地域或国界的限制,可以被视为大脑内部对世界的表述,无论是感知的、还是想象的。

自由手绘草图尽管高度简洁和抽象,但却能说明问题,在交流和设计等各种场景中非常有用。因此,手绘草图在计算机视觉和模式识别^[1]、计算机图形学^[2]、人机交互^[3]、机器人学^[4]和认知科学^[5]等领域被广泛研究。其中,早期的研究可追溯到 20 世纪 60、70 年代。近年来,基于草图的方法已被应用于内容创作支持的各个领域,例如图像生成^[6]、3D 模型重建^[7]和动画控制^[8]。3D 汽车模型在娱乐、视觉效果和汽车设计等多个领域被大量使用。由于生成与现实世界相似特征模型的过程通常既耗时又费力,因此从单个图像输入重建 3D 模型的自动方法可以作为有效的解决方案。

尽管在相关领域进行了大量研究,但由于多种原因,复杂 3D 汽车模型的高质量重建仍然具有挑战性。在动画、游戏和汽车制造行业,设计过程通常涉及汽车概念图,而概念图通常是在预定义视角上以草图形式表现出来,因此以往使用照片输入的方法在设计新形状或修改现有设计时并不实用。对基于深度学习的方法来讲,不同类型汽车(如 SUV 和卡车)具有明显不同的特征,使得单一神经网络难以学习复杂的数据分布。汽车建模细节是一个明显的问题,因为汽车既有共同特征,如车轮的位置,也有独特

部分,如尾翼和车顶形状。因此,构建一个能够充分表达常见汽车特征和独特汽车特征的多样化汽车子空间是一项挑战。

近年来,触摸屏设备普及使草图创建变得更加便捷,进而推动了面向草图的应用程序流行。同时,深度学习不断进步推动了手绘草图在 3D 汽车模型设计领域的研究和应用,而基于深度学习的方法成功实现了从单一简单草图生成更简洁、更准确模型的目标。文献[9]尝试通过将模板网格变形为目标形状的方式从草图中生成 3D 模型,这是一种常见的从单一视角图像还原 3D 模型的方法;而文献[10]则利用隐式函数学习方法,通过学习符号距离函数(SDF)来生成目标形状。然而,文献[9]可能受到模板模型拓扑结构的限制;而文献[10]则仅限于水密的 3D 网格设计。

本文的目标是从用户手绘草图中生成非水密的 3D 汽车模型。为了解决这个问题,采用了无符号距离函数隐式场,通过系统从手绘草图生成 3D 汽车模型,进而生成非水密网格的汽车 3D 模型,这是一种在处理非水密网格时具有卓越性能的代表方式。使用无符号距离函数(UDF)能够更加灵活地捕捉和呈现手绘草图中的复杂形状,而无需受限于特定模板的拓扑结构。UDF 的引入使得本文系统能够更加自由地生成非水密网格的汽车 3D 模型,更好地满足用户对手绘草图的设计需求。通过应用 UDF,可成功克服文献[9]和文献[10]模型中存在的限制,从而提升本文模型的灵活性和适用性。本文的创新之处在于:

1) 提出了一种新颖的基于草图的 3D 汽车模型生成系统。相较于基于图像的建模,该方法更注重用户通过手绘草图传达设计意图,使生成的 3D 模型更加贴合用户需求。

2) 利用深度学习技术估计手绘草图的无符号距离字段(UDF),并通过 Marching Cubes 算法从估计的字段中提取模型网格。UDF 的使用对复杂形状建模更为灵活,尤其适用于处理非水密网格。

3) 引入了深度学习驱动的自动编码器结构,系统不仅生成汽车模型,还提供了用户友好的编辑功能。用户可以通过编辑生成的网格实现对模型的个性化定制,为创造性和设计灵活性提供了更多的空间。

1 相关工作

1.1 基于图像的 3D 形状建模

基于图像的 3D 形状建模文献很多,主要可分为单视图和多视图 3D 形状建模,现有方法可以根据所使用的 3D 标识进一步分类。基于体素的模型是研究最广泛的模型,其中卷积神经网络(CNN)已被成功应用于学习给定 2D 视图图像的 3D 对象重建的概率潜在空间。然而,这些方法往往受到计算效率限制,经常导致产生的体素数量不满足。为了缓解这种情况,研究人员提出了基于八叉树的网络和点云模型来生成更高分辨率的模型。目前,基于多边形网格的方法由于同时编码了几何(点云)和拓扑(表面连通性)线索,获得了广泛关注。然而,这些方法仍然受到模板形状拓扑限制,难以准确地重建亏格大于 0 的形状。一些学者已经提出将表面法线和深度图的多个视图作为监督信号来训练 3D 建模系统。使用额外的融合步骤将其合并到点云中,然后使用预测的轮廓去除异常值。

1.2 基于草图的 3D 形状建模

由于两个主要因素,从人体草图估计 3D 形状比从单个图像估计更具挑战性:① 模糊性,即多个 3D 形状可能投影到同一张图纸上;② 失真,即草图和 3D 形状之间存在错位。早期的研究尝试通过利用几何属性来推断 3D 形状,即主要通过推断局部几何属性来预测 3D 形状。手工设计规则通常用于将轮廓拉伸或抬高成光滑的 3D 表面。同时,一些方法采用多面体脚手架、横截面线和曲率流线等工具来创建自由曲面。这些早期方法试图应对模糊性和失真这两个挑战,从而提高依据人体草图估计 3D 形状的准确性。

近年来, 深度学习方法被用于基于草图的建模任务。蔡兴泉等^[11]建议首先将输入草图转换为深度/法线图, 并将其融合构建完整的 3D 模型。但是, 所有这些方法都是基于整体草图执行的, 导致缺少对象的细节。Li 等^[12]利用深度神经网络从 2D 草图中预测人脸的潜在代码, 生成详细的人脸模型。蔺琛皓等^[13]则训练网络来预测程序模型参数, 从草图中生成详细的形状。这些方法可以生成复杂的高分辨率模型, 但仅适用于通过编程方式生成的形状。为了处理更复杂的形状, Li 等^[14]设计了一个端到端的卷积神经网络, 给定输入草图且基于体积的表示形式生成相应的 3D 形状, 但是这种方法仅限于低分辨率 3D 网格。Liu 等^[15]提出了一个基于生成对抗网络(GAN)框架的方法用于生成法线图, 并通过结合用户特定法线的交互过程来提高生成结果的质量。但该方法的实现成本相对较高, 因为它需要额外的点掩模和用户指南输入。Liu 等^[16]采用对抗性训练将合成图像和来自 3D 形状库的参考 3D 形状作为源, 形成联合特征, 尽管取得了令人满意的结果, 但该方法严重依赖画廊中的模板形状, 在很大程度上限制了其普适性。Ma 等^[17]基于图像空间轮廓渲染得到的合成草图进行训练, 致力于 3D 形状建模, 通过单视图 CNN 和更新程序 CNN 训练, 成功地从草图生成了占用体素网格。然而, 该方法在重建薄结构方面存在一定的局限性。

2 方法

本文基于草图的汽车模型生成框架利用自动编码器结构神经网络从用户手绘草图中重建 3D 汽车形状模型。图 1 概述了该框架从输入草图生成及编辑 3D 模型的过程, 目的是通过等值面提取和 UDF 的三角化网格生成, 从输入草图重建 3D 模型。基于汽车形状模型抽取隐式 UDF 字段, 构建学习数据集, 可用于绘制草图以及查看和编辑生成的模型。

该系统由两个网络组成: ① 将草图图像编码为潜码 K_c 的编码器; ② 根据潜向量 K 生成 UDF 的解码器。解码器根据 K_c 预测所有 $u_x \in A$ 的无符号距离 d_x^p , 并输出 UDF。然后, 使用行进立方体算法获得三维模型网格, 同时用户还可以编辑生成的模型。

2.1 无符号距离字段

将 UDF 用于形状表面表示, 设 3D 模型所在的三维空间为 Ω 。对于任何坐标 u , UDF $\Phi(u)$ 的计算如公式(1)所示。

$$\Phi(u) = d(u, \partial\Omega) \quad u \in \Omega \tag{1}$$

式中: $\partial\Omega$ 表示 Ω 的边界; $d(u, \partial\Omega)$ 表述坐标 u 到三维模型边界 $\partial\Omega$ 的最小距离; 坐标 u 到 $\partial\Omega$ 的距离 d 可通过公式(2)计算; Ω 表示三维模型占据的空间区域。

$$d(u, \partial\Omega) = \min_{r \in \partial\Omega} |u - r| \tag{2}$$

式中: r 表示边界 $\partial\Omega$ 上的任意点; $|u - r|$ 表示坐标 u 与点 r 之间的欧几里得距离。由于形状表面可以用零水平集 $UDF=0$ 表示, 因此本文可以采用诸如行进立方体等网格重建算法, 以生成网格。

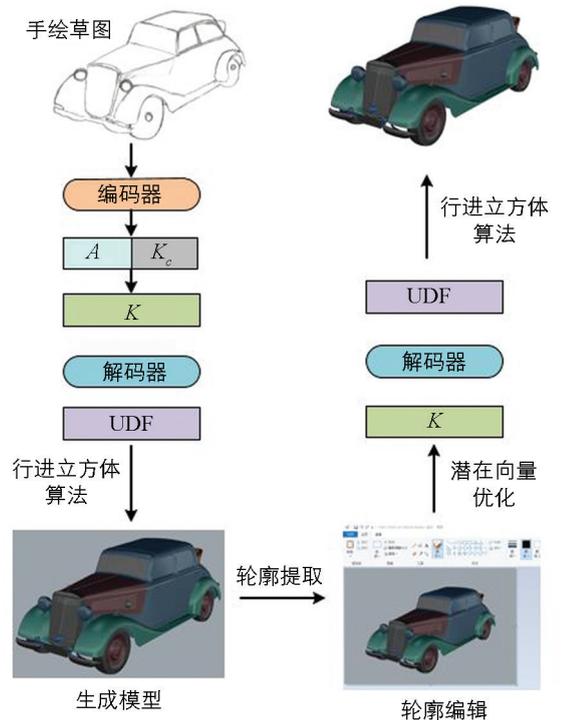


图 1 本文设计生成的系统框架

2.2 数据集构建

基于合成方法降低了获取昂贵成对 2D 和 3D 训练样本的成本,同时生成数据库可以很容易地扩展为更大的尺寸和规模。本文使用条件 GAN 生成轮廓草图,并从 3D 网格创建点云表示。通过等高线草图注释,使不具备三维建模技术的用户也可以轻松地使用该方法。通过特征保留数据增强技术,创建逻辑正确的汽车网格。

本文的汽车网格来源于 ShapeNet,因为该数据库提供了各种型号汽车的不同变化情况,有效地训练深度学习系统需要具备足够规模的数据库。采用文献[18]中提及的方法对汽车网格进行增强,该方法可以改变汽车形状并同时保留局部特征,核心在于对网格进行体素化,并通过增强(即缩放)体素网格进行插值来改变顶点位置。本文系统将网格分辨率设置为 $5 \times 10 \times 15$,随后以 $\pm 20\%$ 、 $\pm 15\%$ 、 $\pm 10\%$ 和 $\pm 5\%$ 的缩放率分别应用于高度和长度方向。通过对 7 028 个汽车网格的应用增强方法,成功地生成了 56 224 个不同的网格,如图 2 所示。



图 2 不同的生成结果

为了从草图生成 UDF,需要草图和 UDF 的成对数据。

(1) UDF 数据进行采样

本文系统旨在学习网格的 UDF 场分布,因此需要从给定网格中对 UDF 场进行采样。UDF 字段表示一组采样顶点以及从顶点到网格表面的无符号距离。将网格数据归一化为最大坐标值 1,使其适合训练。由于采样是从曲面外部和内部执行的,因此调整比例拟合任意半径略小于 1 的曲面,以便从曲面外部距离模型中心最远处及调整后的网格上执行采样。根据以往经验,本文中的量表设置为 0.8。

对于每个 W_x 网格,采样顶点数设置为 T ,距离 d_x^p 从 T 个 3D 顶点坐标 u_x 到 W_x 计算曲面。采样次数没有限制,但考虑到数据处理速度,将 T 设置为 120 000,以便在尽可能短的时间内获得尽可能多的样本。采样数据由曲面上、曲面附近和网格周围边界框中的随机点组成,采样集中在网格表面上。在实现中,从表面 0.05 以内的坐标中获取 48 000 个样本;从表面 0.3 以内的坐标中获取 32 000 个样本。此外,还有 24 000 个随机抽样来自 W_x 曲面,16 000 个在边长为 2(范围 $[-1, 1]$)的边界框内采样。本文中 d_x^p 使用 *KDTree* 计算。 W_x 网格不需要特殊处理。

(2) 草图—网格配对数据

本文使用的数据集不提供成对的草图图像,要获得草图需要首先为每个模型生成渲染图像,然后使用渲染图像生成草图图像。考虑到手绘草图的稀疏特征,本文从渲染深度图中以不同角度提取目标网格的二值化轮廓作为草图图像。深度图可以删除不必要的细节生成更高质量的轮廓。对于每个网格 W_x ,本文从多个角度渲染深度图。通过这种方法获得的网格数据与草图数据配对,如图 3 所示。

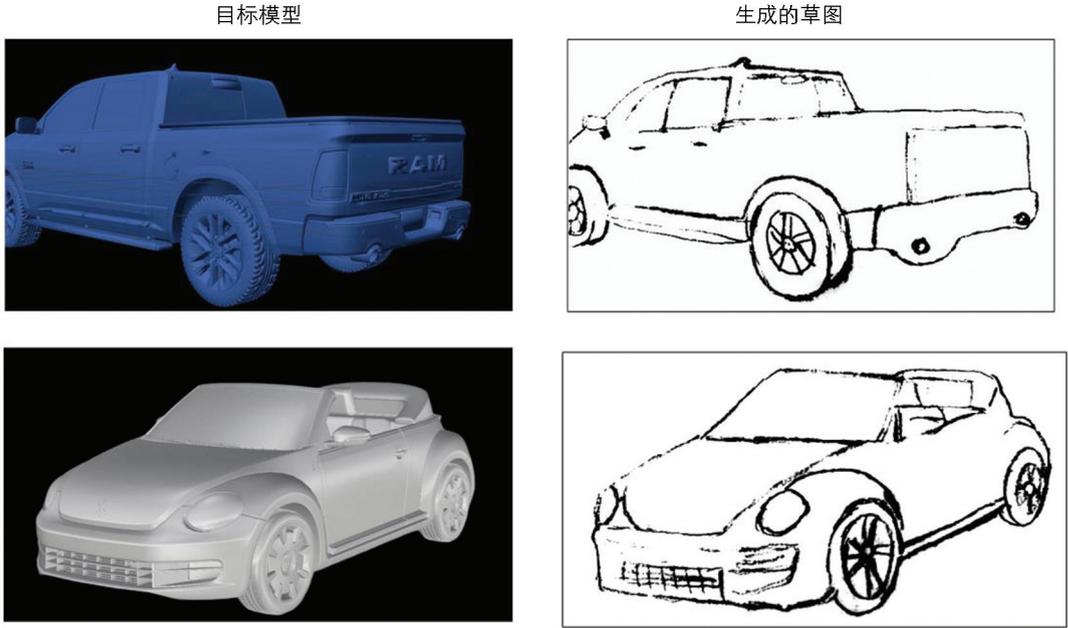


图 3 生成草图—网格数据配对

2.3 隐式曲面表示学习

在这项研究中, 本文使用深度学习的自动编码器结构进行 UDF 训练。设 E 是将输入草图编码为潜在向量的编码器, D 是生成 UDF 数据的解码器, S 是输入草图。编码器 E 将输入草图 S 编码为潜在代码 K_c 。网格是在三维空间 $[-1, 1]^3$ 中对三维形状进行离散法网格采样的结果, 用 A 表示。将潜码 K_c 与 A 串联形成潜向量 K , 由解码器 D 生成 UDF, 过程如公式(3)所示。解码器可预测所有输入坐标 $u_x \in A$ 的无符号距离 d_x^p 。

$$UDF = D(E(S), A) \tag{3}$$

2.3.1 目标函数

网络的学习目标是为给定的 S 生成最合适的 UDF。学习中的优化目标如公式(4)所示, 利用 $L1$ 距离, 其中 df_{an}^p 为基础真值。为了减少分散在网络中样本坐标的影响, df_{an}^p 和 $D(K)$ 在 δ 处设置阈值以去除离散值, 称为夹紧函数。此外, 本文采用几何正则化损失来改进网络学习, 如公式(5)所示。正则化损失是一种广泛应用技术, 在机器学习和深度学习领域具有实际效果和价值。例如, 正则化损失在图像分类任务中的应用, 通过对神经网络权重进行正则化, 可以提高模型对新样本的泛化能力, 从而避免对训练数据过拟合。在自然语言处理中, 正则化损失有助于防止循环神经网络(RNN)模型在处理长序列时产生梯度爆炸或梯度消失问题, 提高其在不同文本数据上的性能。对输入潜在向量 K 进行归一化, 并对归一化结果进行优化。归一化损失如公式(6)所示。 γ 和 λ 分别表示归一化因子和 UDF 值的标量。基于上述内容, 网络损失函数由公式(7)计算。

$$L(D(K), df_{an}^p) = | clamp(D(K), \delta) clamp(df_{an}^p, \delta) | \tag{4}$$

$$L_{reg_geo} = \frac{1}{T} \sum \exp(-\gamma \cdot D(K)) \tag{5}$$

$$L(K) = \lambda \| K \|_2 \tag{6}$$

$$L = L(D(K), df_{an}^p) + L(K) + L_{reg_geo} \tag{7}$$

式中: δ 表示夹紧阈值, 用于抑制离散点的影响; γ 表示几何正则化强度系数; T 表示采样点总数; λ 表示归一化损失权重。

2.3.2 生成模型优化

在对生成结果的优化和编辑中,本文参考了 Hertz 等^[19]的方法。将草图结果和倒角距离最小化,以优化作为解码器输入的潜在向量,并将倒角距离作为优化的限制约束。倒角距离是在 2D 中实现的,2D 倒角损失 L_{CD} 由公式(8)计算。

$$L_{CD} = \sum_{u_u \in S_u} \min_{u_s | F_s(u_s)=0} \|u_u u_s\|^2 + \sum_{u_s | F_s(u_s)=0} \min_{u_u \in S_u} \|u_u - u_s\|^2 \quad (8)$$

式中: S_u 表示 3D 网格的投影轮廓点集; F_s 表示草图区域的指示函数; $F_s(u_s)=0$ 表示点 u_s 在草图区域内; u_u 表示投影轮廓上的点; u_s 表示草图区域上的点; $u_u - u_s$ 表示两点间的欧几里得距离平方。

由于生成结果和输入之间的损失,对生成结果的优化本质上是反向传播梯度值。由于 UDF 数据和 3D 模型都是不可微分的,因此优化目标是解码器输入的潜在向量。由于潜在向量作为网络学习过程的一部分是可微分的,因此可以进行优化。这种优化允许对网络输出结果进行调优。具体来讲,如果生成的 3D 网格为 W ,其表面顶点的采样为 Q ,则使用公式(9)计算潜在向量 K 的梯度 $\frac{\partial L_{CD}}{\partial K}$ 。计算出的梯度通过反向传播过程优化潜在向量。

$$\frac{\partial L_{CD}}{\partial K} = \sum_{q \in Q} \frac{\partial L_{CD}}{\partial q} \nabla D(q, K) \frac{\partial D}{\partial K}(q, K) \quad (9)$$

式中: $\nabla D(q, K)$ 表示解码器对坐标 q 和潜在向量 K 的梯度。通过优化的潜在向量,利用解码器生成 UDF 数据,从而实现对 3D 模型的重建。

2.3.3 网络培训

本文系统的深度学习部分通过 PyTorch 和 PyTorch 3D 实现。在 UDF 生成阶段,使用自动编码器采用行进立方体算法进行实现,解码器由 9 个感知器层组成。MLP 层的隐藏维度数为 512,激活函数使用 ReLU。为了减少激活函数导致的详细信息损失,对输入的 3D 坐标进行傅里叶位置编码。编码器使用 ResNet 将草图图像编码为潜在向量。在训练过程中,损失函数由 UDF 导致的 L_1 损失[公式(4)]和输入归一化损失[公式(6)]的总和构成。散射去除因子 δ 设置为 0.1,归一化强度 λ 为 10^{-4} 。网络训练流程如图 4 所示。

批量大小为 16, epoch 数为 2 000。本文采用的优化器为 Adam,并根据公式(10)中后续训练周期数来设置学习率。 lr_{init} 为初始学习率,初始值设置为 0.5×10^{-3} 。 ω 为 0.5, γ 为 500。在编码器参数设置方面,设 $\alpha=1$;在解码器参数设置方面,设 $\alpha=0.1$ 。训练过程使用两个 NVIDIA GeForce RTX 3090 进行。每个 epoch 的训练时间约为 10 s,完整训练时间约为 5.5 h。其中,学习率的计算公式为:

$$Learning_rate = \alpha \times lr_{init} \times (\omega^{\lfloor \frac{epoch}{\gamma} \rfloor}) \quad (10)$$

式中: ω 表示衰减因子; $epoch$ 表示当前训练轮数; γ 表示衰减步长; α 表示编码器和解码器的权重系数。

3 实验结果与分析

在实验部分: ① 将本文方法与从图像和草图中重建 3D 形状的最新工作进行定性和定量比较; ② 对本

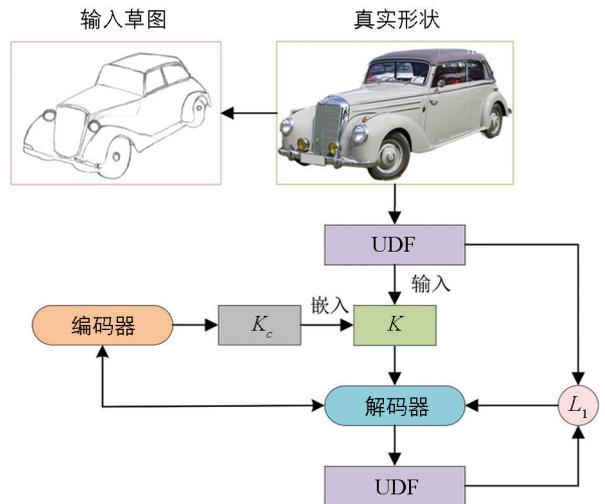


图 4 网络的训练流程

文方法每个组成部分的重要性进行评估; ③ 演示将本文方法应用于草图实例级 3D 形状检索的过程。

3.1 数据拆分和实施细节

随机将 ProSketch 3D 数据集分成 450 个用于训练的 3D 形状和 50 个用于测试的 3D 形状。由于 ProSketch 3D 数据集的大小有限, 本文使用 ShapeNet 数据集中 6 278 个 3D 汽车生成草图来预训练 benw 网络和所有竞争对手。对于每个 3D 形状, 使用非照片级真实感渲染生成相应的正面和侧面正交投影, 预训练后进一步在 ShapeNet 数据集训练分区上对网络进行训练。

对比分析本文系统与其他 3 种方法(文献[15-17]), 由于原始实现中使用的自定义修改版(Caffe 库版本)已过时, 本文在 PyTorch 中重新实现了它们的工作, 采用 Adam 优化器将学习率设置为 0.000 5, 并将批量大小设定为 64。

3.2 结果和讨论

3.2.1 定性结果

首先将本文方法生成的单个视点与文献[15]生成的视点进行比较(图 5), 这两种方法每个视图都是通过融合每个视图的法线图、深度图和前景掩膜生成的。从图 5 中可以看出, 本文方法生成的视图具有更精细的细节和更清晰的轮廓。

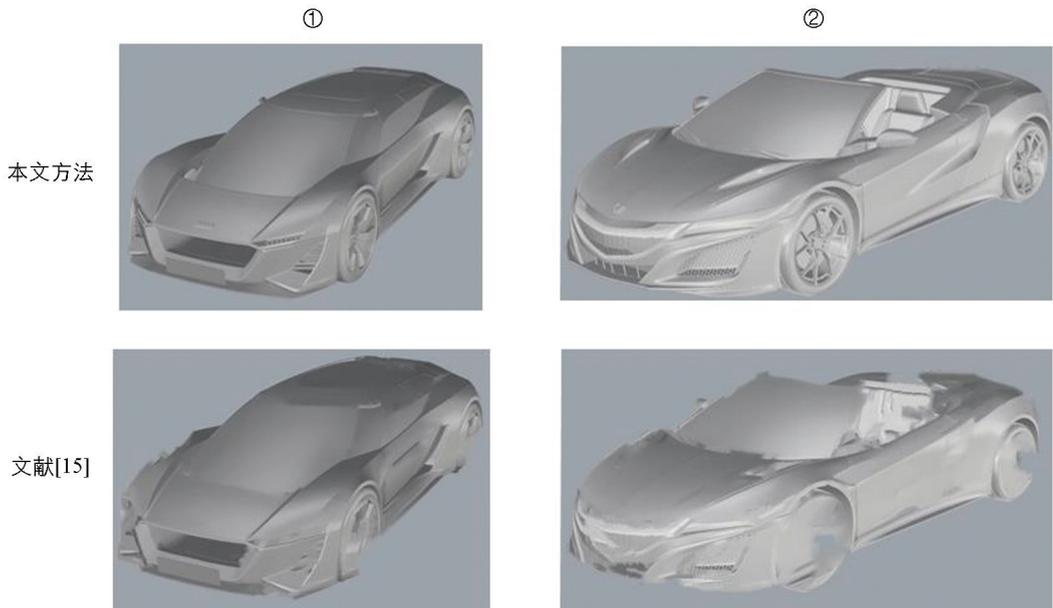


图 5 生成视图比较

将本文方法生成的 3D 模型与文献[16-17]生成的模型进行比较(图 6)。从图 6 中可以明显看出, 本文系统生成的 3D 形状不仅具备更为丰富的细节, 而且其轮廓也更加清晰。相较于其他几种方法, 本文系统生成的 3D 形状在质量和精准度上都有显著优势。使用本文方法进行 3D 形状重建, 可以更加精确地呈现出原始草图的各个部分, 同时也能更好地捕捉到那些精细的局部细节。

3.2.2 定量结果

本文使用 5 个指标来定量评估生成的 3D 汽车形状质量: ① 法线角度距离(NAD)和真实三维形状之间每对最近点的平均法向角距离。② 豪斯多夫距离(HD)测量重建形状上每个表面点到参考形状上最近表面点的最大距离。③ 深度误差(DE)是生成的深度图和重建的深度图之间的欧几里得像素距离。④ 掩膜误差(ME)通过测量预测掩模和真实掩模之间逐像素平均误差来计算。⑤ 倒角距离(CD)评估的是生成形状上每个点(每个形状随机抽取 5 000 个点)到真实形状上最近表面点的平均距离。从表 1 可知, 所有指标(数值越小性能越优)均显示本文方法优于现有对比方法。

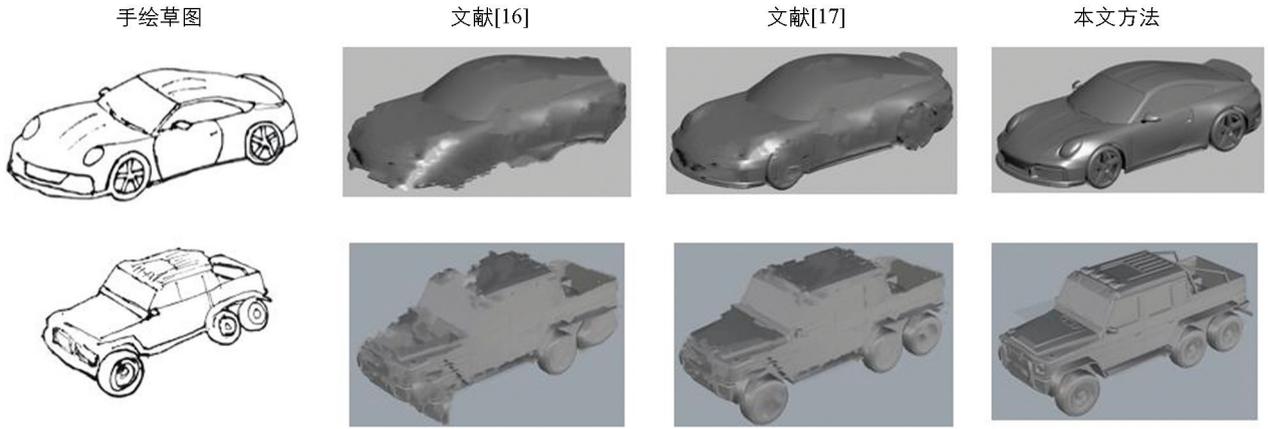


图 6 生成的整个 3D 形状比较

表 1 不同方法定量比较结果

指标	方法			
	文献[15]	文献[16]	文献[17]	本文
NAD	35.975	34.347	33.853	33.014
HD	0.528	0.367	0.453	0.197
DE	0.034	0.023	0.016	0.009
ME	0.375	0.246	0.214	0.109
CD	0.088	0.068	0.053	0.029

3.2.3 使用合成数据进行训练

专业艺术家的草图虽然比新手草图更准确,但与简单的合成草图相比仍然会出现严重的失真。为了评估真实人类草图在训练阶段的重要性,本文比较了对 ProSketch 3D 数据集训练分集进行微调所获得的结果,以及对训练分集中 3D 图形合成草图进行微调所获得的结果,如图 7 所示。由图 7 可知,仅仅使用简单合成草图进行训练,即使是在 ProSketch 3D 数据集中精心创建的草图,也不能很好地推广到人类草图中。结果表明,仅使用合成草图进行训练,不如使用专业草图进行训练效果好。



图 7 仅基于合成数据训练与本文系统训练的比较

3.3 3D 形状检索和识别

对于检索,本文利用编码器作为查询草图和 3D 形状的特征描述符,并计算该特征空间中的欧几里得距离。在实验中,测试集中的所有汽车都被用作查询草图,图库集包含了来自 ShapeNetCore 的所有 6 778 个 3D 汽车形状,对比结果如图 8 所示。

从图 8 中可以看出,检索到的排名靠前的 3D 形状在视觉上与 ground-truth 三维形状高度相似。排名靠前的结果与真实形状非常相关,且与自动生成的结果表现出较高的相似性(真值除外)。

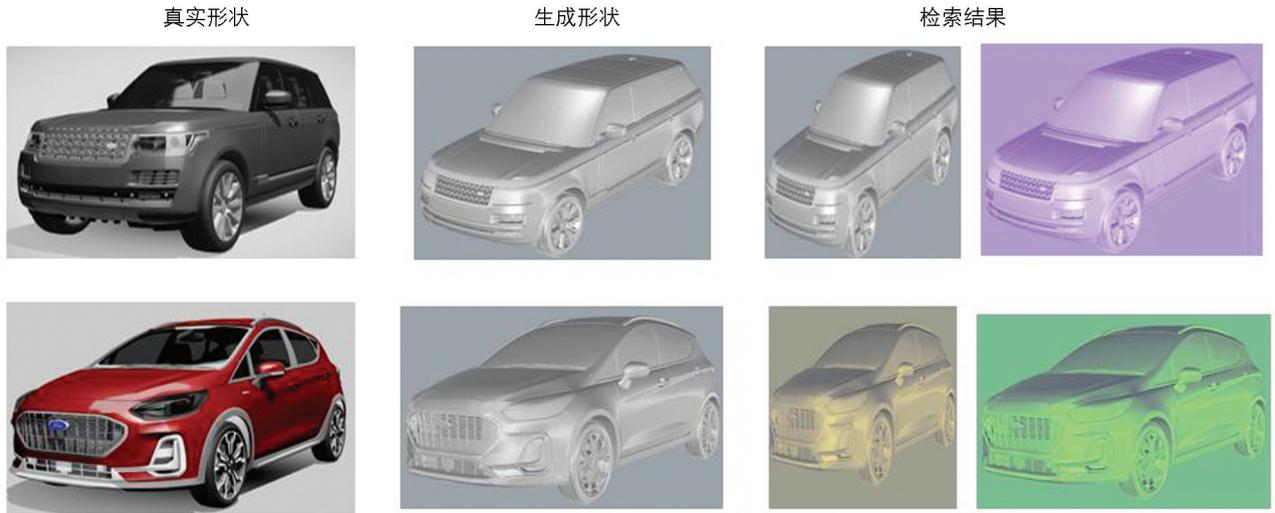


图 8 基于草图的 3D 形状检索示例

在识别方面, 本文使用 PointNet++ 来测量生成的 3D 汽车模型的识别程度。识别准确率如图 9 所示。由图 9 可知, 与其他 3 种对比方法相比, 本文方法生成的数据更容易被正确分类, 表明其具有优势。

4 总结

在电脑游戏、视觉效果和汽车设计等领域, 3D 汽车模型得到广泛应用, 因此以最低劳动力成本生产这类模型变得越来越重要。

为了解决这一问题, 本文提出一种新颖系统, 可使用单个 2D 手绘草图重建详细的 3D 汽车形状, 且设计生成的系统利用自动编码器结构学习 UDF 隐式表示, 并输出 UDF 数据。网格通过行进立方体方法从 UDF 中提取, 同时该系统还允许普通用户绘制草图并编辑生成 3D 汽车模型。当用户编辑生成模型轮廓时, 基于优化潜向量编码器实现模型编辑, 并通过评估实验, 验证本文系统在从手绘草图生成 3D 汽车模型方面的有效性。虽然本文方法是产品设计中通用的方法, 但为了实现这一目标需要构建更多的通用数据集, 包含各种产品类型的配对草图图像和 3D 模型。该数据集将涵盖家居用品、电子产品、交通工具等多种不同类型的产品, 未来将进一步深入探索更多的方向, 例如可以利用宜家数据集从草图图像中生成 3D 家具模型, 该数据集包含大量家具设计草图及相应的 3D 模型, 可供训练和测试框架使用。通过使用该数据集, 能够提高家具设计领域内的模型准确性和效率, 从而进一步推动产品设计发展。除了家具设计外, 还可以探索服装设计、工业设计等其他领域的应用。

参考文献:

- [1] 李雄, 苏建宁, 张志鹏. 基于深度学习的产品概念草图生成设计研究 [J]. 机械工程学报, 2023, 59(11): 16-30.
- [2] 孟志刚, 廖帅元. 基于深度学习的手绘图形识别的研究 [J]. 计算机与数字工程, 2023, 51(4): 949-953.
- [3] 李洋, 康文惠, 黄进, 等. 基于笔式交互的人机混合 CAD 概念设计 [J]. 计算机集成制造系统, 2022, 28(7): 2050-2063.

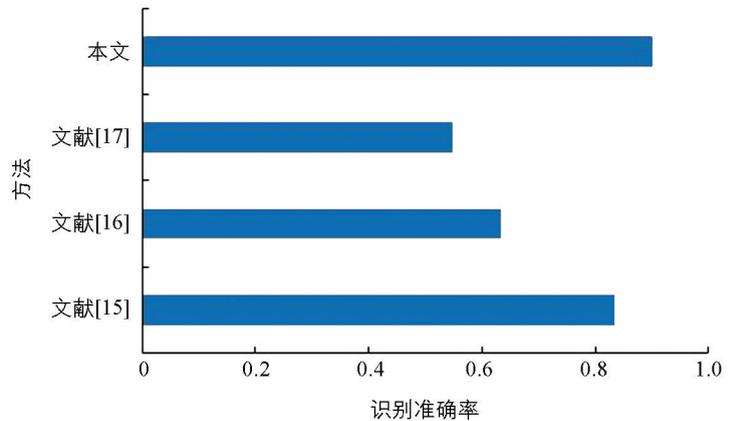


图 9 不同方法生成的 3D 形状的识别准确率

- [4] 李萌, 查思雨, 宫未, 等. 面向儿童学习的智能家居人机交互技术综述 [J]. 计算机辅助设计与图形学学报, 2023, 35(2): 248-261.
- [5] 周荣庭, 余悦玫, 李瑾. 科普文创产品的设计与营销策略 [J]. 自然科学博物馆研究, 2022, 7(5): 49-55.
- [6] XU P, HOSPEDALES T M, YIN Q, et al. Deep Learning for Free-Hand Sketch: A Survey [J]. IEEE Trans Pattern Anal Mach Intell, 2023, 45(1): 285-312.
- [7] OKURA F. 3D Modeling and Reconstruction of Plants and Trees: A Cross-Cutting Review across Computer Graphics, Vision, and Plant Phenotyping [J]. Breeding Science, 2022, 72(1): 31-47.
- [8] BRODT K, BESSMELTSEV M. Sketch2Pose [J]. ACM Transactions on Graphics, 2022, 41(4): 1-15.
- [9] CHEN T R, CAO R L, LI Z J, et al. Deep3DSketch-Im: Rapid High-Fidelity AI 3D Model Generation by Single Free-hand Sketches [J]. Frontiers of Information Technology & Electronic Engineering, 2024, 25(1): 149-159.
- [10] CHEN T R, CAO R L, LI Z J, et al. Deep3DSketch-Im: Rapid High-Fidelity AI 3D Model Generation by Single Free-hand Sketches [J]. Frontiers of Information Technology & Electronic Engineering, 2024, 25(1): 149-159.
- [11] 蔡兴泉, 李治均, 奚梦瑶, 等. 基于神经网络的手绘服饰图纹上色及风格迁移 [J]. 系统仿真学报, 2023, 35(3): 604-615.
- [12] LI M Y, CHEN W, WANG S S, et al. Survey on 3D Reconstruction Methods Based on Visual Deep Learning [J]. Journal of Frontiers of Computer Science & Technology, 2023, 17(2): 279.
- [13] 蔺琛皓, 沈超, 邓静怡, 等. 虚假数字人脸内容生成与检测技术 [J]. 计算机学报, 2023, 46(3): 469-498.
- [14] LI Y X, QIAN J M, FENG S J, et al. Deep-Learning-Enabled Dual-Frequency Composite Fringe Projection Profilometry for Single-Shot Absolute 3D Shape Measurement [J]. Opto-Electronic Advances, 2022, 5(5): 210021.
- [15] LIU C, WANG D L, ZHANG H, et al. Using Simulated Training Data of Voxel-Level Generative Models to Improve 3D Neuron Reconstruction [J]. IEEE Transactions on Medical Imaging, 2022, 41(12): 3624-3635.
- [16] LIU M T, ZHU A H, MAITI P, et al. Style Transfer Generative Adversarial Networks to Harmonize Multisite MRI to a Single Reference Image to Avoid Overcorrection [J]. Human Brain Mapping, 2023, 44(14): 4875-4892.
- [17] MA Z P, ZHOU J, MA J L, et al. A Novel 3D Shape Recognition Method Based on Double-Channel Attention Residual Network [J]. Multimedia Tools and Applications, 2022, 81(22): 32519-32548.
- [18] AYUBI J, AMIRANI M C, VALIZADEH M. A New Seam Carving Method for Image Resizing Based on Entropy Energy and Lyapunov Exponent [J]. Multimedia Tools and Applications, 2023, 82(13): 19417-19440.
- [19] HERTZ A, PEREL O, GIRYES R, et al. Spaghetti [J]. ACM Transactions on Graphics, 2022, 41(4): 1-20.

责任编辑 夏娟

周梦媛