



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
06
VOL.26

ISSN1006-8961
CN11-3758/TB



中国图象图形学学会成立三十周年

图像图形学 发展 年度报告



第26卷第6期（总第302期）
2021年6月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元

序言王耀南



生物特征识别学科发展报告
(第1254页)

图像处理与通信技术

视频处理与压缩技术

贾川民, 马海川, 杨文瀚, 任文琦, 潘金山, 刘东, 刘家瑛, 马思伟 1179

面向体验质量的多媒体计算通信

陶晓明, 杨铀, 徐迈, 段一平, 黄丹蓝, 刘文予 1201

数字媒体取证技术综述

李晓龙, 俞能海, 张新鹏, 张卫明, 李斌, 卢伟, 王伟, 刘晓龙 1216

面向智慧城市的交通视频结构化分析前沿进展

赵耀, 田永鸿, 党建武, 付树军, 王恒友, 万军, 安高云, 杜卓然, 廖理心, 韦世奎 1227

生物特征识别学科发展报告

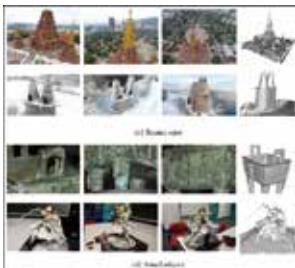
孙哲南, 赫然, 王亮, 阚美娜, 冯建江, 郑方, 郑伟诗, 左旺孟, 康文雄, 邓伟洪, 张杰, 韩琥, 山世光, 王云龙, 茹一伟, 朱宇豪, 刘云帆, 何勇 1254

自然场景文本检测与识别的深度学习方法

刘崇宇, 陈晓雪, 罗灿杰, 金连文, 薛洋, 刘禹良 1330

基于深度学习的跨模态检索综述

尹奇跃, 黄岩, 张俊格, 吴书, 王亮 1368



大规模室外图像3维重建技术
研究进展(第1429页)

三维视觉和图形技术

三维视觉前沿进展

龙霄潇, 程新景, 朱昊, 张朋举, 刘浩敏, 李俊, 郑林涛, 胡庆拥, 刘浩, 曹汛, 杨睿刚, 吴毅红, 章国锋, 刘焯斌, 徐凯, 郭裕兰, 陈宝权 1389

大规模室外图像3维重建技术研究进展

颜深, 张茂军, 樊亚春, 谭小慧, 刘煜, 彭杨, 刘宇翔 1429

视觉传感成像技术与数据处理进展

王程, 陈峰, 汶德胜, 雷浩, 宋宗玺, 赵航芳 1450

视觉—惯性导航定位技术研究进展

司书斌, 赵大伟, 徐婉莹, 张勇刚, 戴斌 1470

三维视觉测量技术及应用进展

张宗华, 刘巍, 刘国栋, 宋丽梅, 屈玉福, 李旭东, 魏振忠 1483

虚实融合场景中的深度感知研究综述

平佳敏, 刘越, 翁冬冬 1503

可微绘制技术研究进展

许威威, 周漾, 吴鸿智, 过洁 1521

沉浸式立体显示技术在临床医学领域中的应用

郜永航, 石俊生 1536



虚实融合场景中的深度感知
研究综述(第1503页)

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Overview of biometrics research (P1254)



Progress in the large-scale outdoor image 3D reconstruction (P1429)



Review of depth perception in virtual and real fusion environment (P1503)

Image Processing & Communication Technology

Video processing and compression technologies

Jia Chuanmin, Ma Haichuan, Yang Wenhan, Ren Wenqi, Pan Jinshan, Liu Dong, Liu Jiaying, Ma Siwei 1179

Multimedia computing communications

Tao Xiaoming, Yang You, Xu Mai, Duan Yiping, Huang Danlan, Liu Wenyu 1201

Overview of digital media forensics technology

Li Xiaolong, Yu Nenghai, Zhang Xinpeng, Zhang Weiming, Li Bin, Lu Wei, Wang Wei, Liu Xiaolong 1216

Frontiers of transportation video structural analysis in the smart city

Zhao Yao, Tian Yonghong, Dang Jianwu, Fu Shujun, Wang Hengyou, Wan Jun, An Gaoyun, Du Zhuoran, Liao Lixin, Wei Shikui 1227

Overview of biometrics research

Sun Zhenan, He Ran, Wang Liang, Kan Meina, Feng Jianjiang, Zheng Fang, Zheng Weishi, Zuo Wangmeng, Kang Wenxiong, Deng Weihong, Zhang Jie, Han Hu, Shan Shiguang, Wang Yunlong, Ru Yiwei, Zhu Yuhao, Liu Yunfan, He Yong 1254

Deep learning methods for scene text detection and recognition

Liu Chongyu, Chen Xiaoxue, Luo Canjie, Jin Lianwen, Xue Yang, Liu Yuliang 1330

3D Vision & Graphics Technology

Survey on deep learning based cross-modal retrieval

Yin Qiyue, Huang Yan, Zhang Junge, Wu Shu, Wang Liang 1368

Recent progress in 3D vision

Long Xiaoxiao, Cheng Xinjing, Zhu Hao, Zhang Pengju, Liu Haomin, Li Jun, Zheng Lintao, Hu Qingyong, Liu Hao, Cao Xun, Yang Ruigang, Wu Yihong, Zhang Guofeng, Liu Yebin, Xu Kai, Guo Yulan, Chen Baoquan 1389

Progress in the large-scale outdoor image 3D reconstruction

Yan Shen, Zhang Maojun, Fan Yachun, Tan Xiaohui, Liu Yu, Peng Yang, Liu Yuxiang 1429

Review on imaging and data processing of visual sensing

Wang Cheng, Chen Feng, Wen Desheng, Lei Hao, Song Zongxi, Zhao Hangfang 1450

Review on visual-inertial navigation and positioning technology

Si Shubin, Zhao Dawei, Xu Wanying, Zhang Yonggang, Dai Bin 1470

Overview of the development and application of 3D vision measurement technology

Zhang Zonghua, Liu Wei, Liu Guodong, Song Limei, Qu Yufu, Li Xudong, Wei Zhenzhong 1483

Review of depth perception in virtual and real fusion environment

Ping Jiamin, Liu Yue, Weng Dongdong 1503

Differential rendering: a survey

Xu Weiwei, Zhou Yang, Wu Hongzhi, Guo Jie 1521

Application of immersive 3D imaging technology in the clinic medical field

Tai Yonghang, Shi Junsheng 1536

中图法分类号: TP37 文献标识码: A 文章编号: 1006-8961(2021)06-1521-15

论文引用格式: Xu W W, Zhou Y, Wu H Z and Guo J. 2021. Differential rendering: a survey. Journal of Image and Graphics, 26(06): 1521-1535 (许威威, 周漾, 吴鸿智, 过洁. 2021. 可微绘制技术研究进展. 中国图象图形学报, 26(06): 1521-1535) [DOI:10.11834/jig.200853]

可微绘制技术研究进展

许威威¹, 周漾², 吴鸿智¹, 过洁³

1. 浙江大学 CAD&CG 国家重点实验室, 杭州 310058; 2. 深圳大学, 深圳 518057; 3. 南京大学, 南京 210093

摘要: 可微绘制技术是当前虚拟现实、计算机图形学与计算机视觉领域研究的热点,其目标是改造计算机图形学中以光栅化或光线跟踪算法为主的真实感绘制流程,支持梯度信息回传以计算由输出图像的变化导致的输入几何、材质属性变化,通过与优化及深度学习技术等相结合支持从数据中学习绘制模型和逆向推理,是可微学习技术在计算机图形学绘制技术中的应用的具体体现,在增强/虚拟现实内容生成、三维重建、表观采集建模和逆向光学设计等领域中有广泛的应用前景。本文对可微绘制当前的发展状况进行调研,重点对该技术在真实感绘制、3 维重建和表观采集建模中的研究和应用情况进行综述,并对可微绘制技术发展趋势进行展望,以期推动可微技术在学术界和产业界的进一步发展。

关键词: 可微绘制; 3 维重建; 表观采集建模; 基于图像的绘制; 表达学习; 深度学习

Differential rendering: a survey

Xu Weiwei¹, Zhou Yang², Wu Hongzhi¹, Guo Jie³

1. State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058, China;
2. Shenzhen University, Shenzhen 518057, China; 3. Nanjing University, Nanjing 210093, China

Abstract: Differential rendering is currently a research focus in virtual reality, computer graphics, and computer vision. Its goal is to reform the rendering pipeline in computer graphics to support gradient backpropagation such that the change in the output image can be related to the change in input geometry or materials. The development of differential rendering technique is highly related to the deep learning, since neural networks are usually represented as computational graphs to support gradient backpropagation using the chain rule. Thus, the gradient backpropagation is the key to convert a computational procedure into a learnable process, which can significantly generalize the deep learning technique to a wide range of applications. Differential rendering follows this trend to integrate gradient backpropagation into rendering pipeline. It can significantly facilitate the gradient computation through auto-differential techniques. In fact, the derivatives of rendering results regarding to the mesh vertex have already computed in variational 3D reconstruction and shape from shading. However, differential rendering integrates the derivative computation into global rendering pipelines and neural networks. Therefore, the rendering process can be directly integrated into optimization or neural network training to approximate rendering pipeline or inverse graphics reasoning; it has wide applications in content creation in augmented/virtual reality, 3D reconstruction, appearance modeling, and inverse design. The advantage of differential rendering over traditional rendering pipeline is that it allows to train neural networks to approximate the forward rendering pipeline. Once trained, the rendering results can be obtained through forward inference of the network, a much faster procedure in many situations. Moreover, the gradient

收稿日期: 2020-12-31; 修回日期: 2021-03-13; 预印本日期: 2021-03-20

基金项目: 国家自然科学基金项目(61732016, 61772457, 61972194, U2001206)

Supported by: National Natural Science Foundation of China(61732016, 61772457, 61972194, U2001206)

information provided by differential rendering is helpful to improve the efficiency of the global rendering. For instance, the first- and second-order gradients can be used to guide the sampling process in Monte Carlo rendering. Another advantage of differential rendering is that it can directly be used in view interpolation or view synthesis through captured images, which traditional rendering pipeline needs geometry, appearance and lighting information simultaneously to render an image at specified viewpoints. In the application of differential rendering to view synthesis or image-based rendering, the implicit representation of a 3D scene is usually inferred from the captured images directly via deep neural networks supervised by differential rendering loss. Such a process falls into the category of self-supervised learning because ground truth 3D data are not provided during training. It bypasses the expensive multi-view 3D reconstruction and thus significantly simplifies the view synthesis procedure. Numerous representations, such as neural texture, neural volume, and neural implicit function, are proposed to handle freeview point rendering of a 3D scene. However, the training and rendering cost of these methods is still expensive. Thus, reducing their computational cost forms a new research direction. Differential rendering also enables the end-to-end inference of spatially variant bidirectional reflectance distribution function (BRDF) material properties from capture images. The BRDF parameters can be derived from a single image after training the deep neural network on a large amount of data by representing the material properties in a latent space. Moreover, with a differentiable pipeline, the layout of the light sources and projection patterns of dedicated appearance acquisition equipment can be optimized. The recent development of differential rendering, including its application in realistic rendering, 3D reconstruction, and appearance modeling is comprehensively surveyed. We expect this study to further boost the research on differential rendering in academia and industry.

Key words: differential rendering; 3D reconstruction; appearance modeling; image-based rendering; representation learning; deep learning

0 引言

可微绘制技术是当前虚拟现实、计算机图形学与计算机视觉领域研究的热点,其目标是改造计算机图形学中以光栅化或光线跟踪算法为主的真实感绘制流程以支持梯度信息回传,提供图像变化与输入的几何、材质属性变化的相关性,与优化及深度学习技术等相结合支持逆向推理。因此,可微绘制的发展可将计算机图形学的绘制技术演化成为视觉知识推理引擎,在增强/虚拟现实内容生成、3维重建、表观采集建模和逆向光学设计等领域中有广泛的应用前景。

可微绘制的提出与当前深度学习技术的发展密切相关。由于神经网络在计算机中被建模为支持梯度回传的计算图,以便于以随机梯度下降算法进行神经网络参数学习。因此,支持梯度回传是将传统计算过程改造成为数据驱动的参数学习过程的关键,并形成了由美国纽约大学 Yann LeCun 教授等最早提出的可微学习概念。可微绘制技术则是可微学习这一概念在计算机图形学绘制技术中的体现,并得到了迅速发展。事实上,3维重建技术,例如,基于着色信息的形状恢复(shape from shading)

(Zhang 等,1999)和基于变分的3维重建(variational 3D reconstruction)(Vu 等,2012),经常需要计算绘制图像与所拍摄图像之间的差别对所推断的3维几何或材质信息的导数,以从图像中恢复3维结构。该导数计算可以认为是可微绘制技术的最初形式。当前的可微绘制技术则更多地结合了全局光照绘制算法和深度神经网络技术,大幅拓展了可微绘制技术的应用范围。

可微绘制技术在降低绘制算法对输入的要求、提高绘制速度和处理噪声与非朗伯材质物体能力上具有特定的优势。计算机图形学中的真实感绘制技术需用户输入3维几何、反射材料属性和光照信息。虽然在数字化设计场景中应用广泛,但是在直接基于图像等传感器输入进行视点绘制上效率较低。通过可微绘制中的梯度回传技术,使得用户可以训练神经网络从图像等输入中直接推导绘制所需信息,进行基于图像的高真实感绘制,可大大简化得到高真实感图像的流程。由于神经网络在提取高识别度的多尺度图像特征上具有优势,对处理包含非朗伯体的图像推断3维几何信息上也显现了优势。同时,给定新视点,绘制流程仅需所训练的神经网络的一次前向推理,在绘制速度上也具有较大的优势。

本文力图对可微绘制当前的发展状况进行全面调研,重点对该技术在正向真实感绘制、逆向3维重建和表观采集建模中的研究和应用情况进行综述,进一步提高国内相关科研人员对该前沿方向的关注度,投入这一方向的原理和应用的研究,产生在学术界和工业界有引领作用的科研成果。

1 国际研究现状

可微绘制技术围绕如何集成自动微分技术、逆向绘制的目标图像与绘制图像差别对3维几何表达的导数计算以及基于自动导数回传的神经网络训练展开。

1.1 基于自动微分的正向可微绘制技术

在计算机图形学领域,图像合成技术,即绘制,旨在将显式或隐式表达的3维场景转化为2维的图片。传统的绘制包括基于光栅化的绘制、基于物理的真实感绘制以及基于图像的绘制(image-based rendering, IBR)。目前,已经有一些工作(如OpenDR(Loper和Black,2014)、Soft Rasterizer(Liu等,2019a)和RenderNet(Nguyen-Phuoc等,2018)等)将光栅化绘制管线改造成支持绘制过程的梯度回传,以服务于3维物体重建等场景。

基于物理的真实感绘制的核心即为求解绘制方程(the rendering equation)(Kajiya,1986)。作为一个高度复杂的积分方程,目前最流行的求解方法为基于蒙特卡罗(Monte Carlo,MC)采样的数值积分方法(Veach,1997),如路径跟踪、双向路径跟踪和Metropolis路径跟踪等。一般而言,这类方法所生成的图片质量取决于随机采样路径的选取效果。为了尽可能避免盲目随机采样造成的高方差(表现为高噪声),当前有很多研究工作聚焦在如何“智能”地选取随机采样路径上。可微绘制所提供的3维场景导数信息能够给蒙特卡罗采样路径的选择提供有益帮助。

美国麻省理工大学(Massachusetts Institute of Technology, MIT)的Li等人(2015)最早提出用可微绘制中的自动微分技术获取的场景一阶导数和二阶导数来估计绘制点所在3维空间的局部形态。这些局部信息有助于后续采样点的生成,提高后续采样对积分的贡献,如图1所示。

通常情况下,估计二阶导数及对其进行操作都

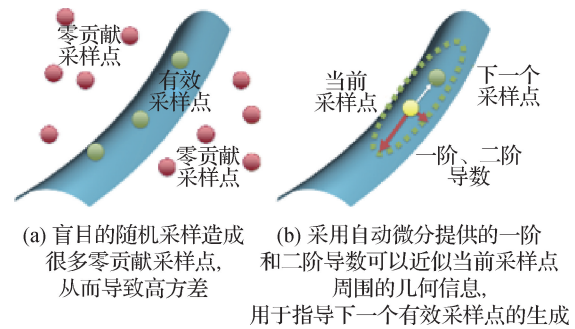


图1 基于梯度信息的采样(Li等,2015)

Fig. 1 Gradient-based sampling(Li et al., 2015)

((a) random sampling results in a large number of useless points, which leads to high variance; (b) first and second derivative from automatic differentiation can approximate the geometric information around sampling points, which can effectively guide the sampling of next points)

比较耗时,会影响3维空间中随机路径搜索的性能。针对这个问题,Luan等人(2020)提出了一种只利用一阶导数的高效采样方法。这个方法受Langevin扩散过程的启发,在采样状态的演化过程中只用到了—阶导数。

以上两种导数驱动的采样方式一般用在MCMC(Markov Chain Monte carlo)技术中,其好处在于自动化和通用性。但是,它们要求场景中的每条采样路径都能计算导数,而这对存在大量不连续点的3维场景而言存在一定的困难。在基于物理的可微绘制技术中,解决这个问题目前有两种常见策略。第1种策略尝试找到所有的不连续边界,并对这些边界进行特定的显式积分来获取绘制方程的导数(Li等,2018a; Zhang等,2019; Li等,2018b)。第2种策略通过重参数化或者散度定理将边界采样转化为面采样,从而回避边界的不连续性对导数计算的影响(Loubet等,2019; Zhang等,2020; Bangaru等,2020)。随着基于物理的可微绘制技术的发展,越来越多的真实感绘制算法或者绘制引擎支持自动求导(Nimier-David等,2019,2020),这给开发更高效的MC采样技术提供了新的途径和更广阔空间。

可微绘制也可以用于基于图像的绘制(Shum等,2007)任务中。与传统的图形绘制不同,基于图像的绘制不需要任何显式的3维场景信息(几何、材质和光照等),而是直接通过不同视点采样的2维图像序列隐式编码场景,从而支持生成该场景在任何视点下的图像。在基于图像的绘制任务中,可微

绘制通常被封装为一个可回传导数的独立的层。下面将展开对该方向的典型工作的探讨。

美国伯克利大学(University of California(UC) Berkeley)的 Mildenhall 等人(2020)提出了 NeRF(neural radiance field),一种采用神经网络建模的隐式3维场景。为了便于从任意视点绘制场景并支持导数的计算,作者开发了一个可微的体绘制方法(如图2所示):将整个场景类比于一个3维介质,通过沿着相机光线随机采样场景中的点并借助介质绘制方法获取该光线的能量并投影到2维图像空

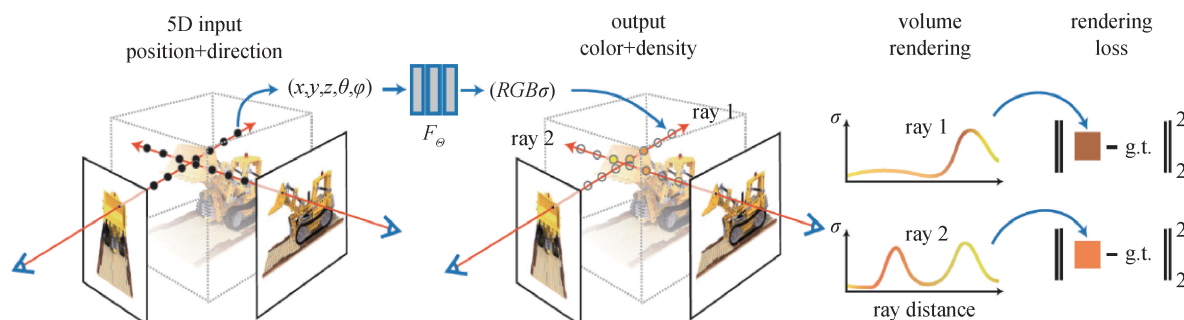


图2 NeRF中所用的可微的体绘制方法(Mildenhall等,2020)

Fig.2 The differentiable volume rendering method used in NeRF(Mildenhall et al.,2020)

德国慕尼黑工业大学的 Thies 等人(2019)开发了一套延迟神经绘制(deferred neural rendering)方法,在2维图像的基础上引入了一些不完整的3维几何信息,从而不仅能够支持针对静态场景的新视点生成,也能实现动态场景的重绘制,具有更高的灵活性。类似于传统的延迟绘制技术,该技术同样依赖于一个G缓存,称为Neural Texture。不同的是,这个缓存纹理和绘制引擎一起都是可学习的。类似的工作还有DeepVoxels(Sitzmann等,2019a)。不过DeepVoxels

学习的是用体素表达的场景。

类似于DeepVoxels,Lombardi等人(2019)提出的Neural Volumes试图从多视角视频中恢复场景的3维体素表达,如图3所示。Neural Volumes中每个体素包含空间中某点的RGB颜色和 α 通道,以及一个形变场,通过可微分光线步进(ray marching)方法将体素绘制成图片,从而通过像素误差驱动网络学习。由于形变场的引入,该方法能用较小的体素分辨率高质量地表达动态物体,并进行视角插值。

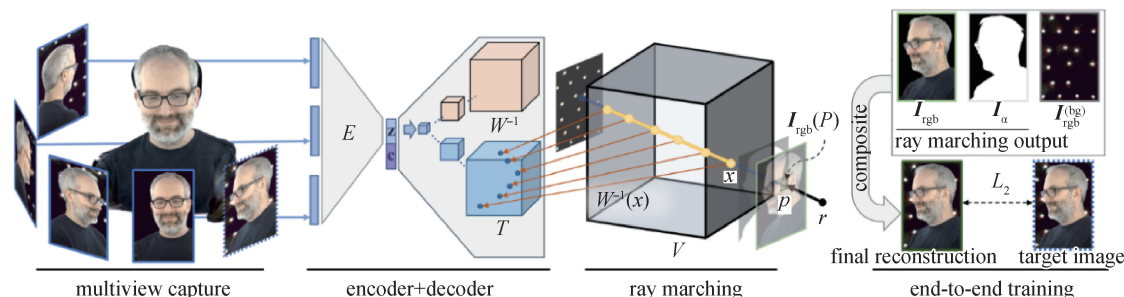


图3 基于体素表达的可微分绘制3维重建与视角插值(Lombardi等,2019a)

Fig.3 Neural volume for 3D reconstruction and view synthesis(Lombardi et al.,2019a)

值得注意的是,目前在基于图像的绘制中所使用的可微绘制都是为特定任务或场景表达定制的,不具有通用性。考虑到性能因素,也没有用到上面

提到的基于物理的可微绘制技术(Li等,2018a;Zhang等,2019,2020;Loubet等,2019;Bangaru等,2020;Nimier-David等,2019,2020)。

1.2 基于可微绘制的逆向3维重建

创建或重建3维内容,是计算机图形学的主要任务之一。然而3维监督获取代价高昂,相比之下,2维数据触手可得。可微分绘制构建了3维场景与2维图像之间的桥梁,为2维反向推理3维提供了一套可行的解决方案。

在计算机图形学中,绘制是依据3维场景描述,如几何信息、材质属性、光照分布和相机位姿等,生成3维场景对应2维图像的过程。而可微绘制将绘制算法解释为函数,该函数将输入场景描述转换成输出图像,并对函数进行微分获得图像相对于场景描述的梯度信息,从而通过反向传播梯度实现场景参数的优化。相比传统的基于图像的3维重建方法主要依赖2维特征点/线,或其他几何和表面材质先验,这些方法很难利用图像上的所有像素信息,而实际上每个像素都与场景参数相关,这为可微绘制超越传统方法带来了可能。同时伴随着神经网络3维数据学习的发展,近几年基于可微绘制的3维重建也迅速发展,涌现出大量工作。

本文根据3维场景的表达,将这些工作分为两大类:基于显式(explicit)和基于隐式(implicit)表达的方法。基于显式表达的方法主要包括基于网格(mesh)、基于体素(voxel)和基于点云(point cloud)的方法。基于隐式表达的方法主要包括符号距离函数(signed distance function)和空间占用函数(occupancy function)两种。

1.2.1 基于网格的方法

网格是3维模型最常用的表达。当使用网格模型进行可微绘制时,最大的问题在于相对于场景的几何参数(如网格的顶点坐标)求导时,物体的边界会引入不连续性,如果不注意处理,在涉及遮挡、阴影和反射等现象时会得到不正确的梯度。为了处理这类问题,许多工作采用各种算法计算相应的梯度。第1个通用可微绘制系统OpenDR(Loper和Black,2014)在梯度计算中采用了一阶泰勒展开近似。Soft Rasterizer(Liu等,2019a)近似绘制的光栅化阶段,将原本基于Z缓冲区的三角形选择更改为所有相关三角形颜色的概率加权,因此投影到某个像素上的每个三角形都有一定的概率为其颜色做出贡献,这使得计算过程是可微分的。Chen等人(2019a)提出仿射插值(可微)的方法来进行光栅化,与OpenGL保持一致。Kato等人(2018)则直接

使用神经网络近似绘制函数。这些工作可用于纹理估计、多视图模型形变和刚性或非刚性(如人体)姿态估计,或与神经网络结合后用于单幅图像3维重建等。最近Li等人(2018a)根据链式法则,直接推导了绘制像素与几何模型顶点之间的解析梯度传导,从而可以将图像的变化(如进行各种滤波操作)回传(back propagation)进行修改几何模型,实现表面平滑、几何纹理迁移等表面编辑。

上述工作通常考虑简化的绘制模型,假设物体为漫反射表面,且忽略间接光照。为了将可微绘制更广泛地应用在基于真实照片的3维重建中,需要考虑更真实的光学现象(Zhao等,2020)。Mitsuba 2(Nimier-David等,2019)是一个开源可微分渲染器,具有利用GPU(graphics processing unit)加速计算的功能,可用于多种用途,包括计算焦散设计(computational caustics design)、非均匀参与介质(heterogeneous participating media)的体密度重建和3D打印半透明板设计等。Loubet等人(2019)采用重新参数化技巧处理不连续性。该方法已被集成到Mitsuba 2中,提供相对于几何参数求导的功能。与Li等人(2018a)方法相比,该方法在含有复杂模型的场景中有更高的效率,并展示了图像中阴影线索对几何形状和纹理的重建的重要性。

最近Lyu等人(2020)利用可微分光线折射追踪方法重建透明物体。光线穿过透明物体时会发生折射,折射方向反映了物体的几何形状,因而其损失函数是基于光线的方向,而不是通常的基于像素颜色定义的。但这也要求透明物体具有光滑的表面。与传统的透明物体重建方法(Wu等,2018)相比,基于可微绘制的优化方法可迅速收敛到更好结果,重建出更多透明物体的几何细节,结果如图4。

除了物体表面的折射和反射现象,另一类重要的光学现象涉及体绘制:光可以在由微小颗粒组成的参与介质中被吸收和散射。辐射传递理论可用来建模光的体积散射。Zhang等人(2019)提出可微分的辐射转移理论,该理论展示了如何计算辐射传递方程(radiative transfer equation, RTE)的各个组成部分相对于场景中任意可微变化的导数。该方法可以准确处理许多光传输现象,比如非均匀介质、体积吸收和散射,以及各向异性相位函数(anisotropic phase functions)。它可应用于参与介质或参与介质中物体的3维重建。在绘制方程中,相对于材质和

光源参数的梯度可以通过自动微分来计算。但是,由于内存占用量大,因此应用仅限于简单的场景。为了解决这个问题,Nimier-David 等人(2020)提出了一种称为“辐射反向传播”的梯度计算方法。在他们的方中,绘制时不存储计算图,相反,在反向传播期间,相机会投射“携带”梯度的光线,光线与物体相交后梯度会传播到物体表面。由于没有考虑由物体边界的变化导致的梯度,这种方法难以用于优化物体的形状。

1.2.2 基于体素的方法

体素类似像素,是将空间进行均匀划分来表达3维信息,其规则的结构天然契合3D卷积神经网络。Yan 等人(2016)提出沿光线将体素投影成2维剪影,使用多视角剪影误差来驱动3维体素的学习,

从而实现基于单张照片重建模型3维体素表达。类似还有 Rezende 等人(2016)、Zhu 等人(2017)、Wu 等人(2017)的工作,但这些工作均不考虑颜色、纹理和光照,只能称之为可微分投影。Tulsiani 等人(2017)通过跟踪光线与3维体素的相交情况,定义了一个新颖的可微分的光线一致性(differentiable ray consistency)损失,相比投影的方法,能处理更多的任务,例如重建带颜色3维体素。Nguyen-Phuoc 等人(2018)提出了第1个基于3维体素的可微绘制网络 RenderNet,这是一个可用像素误差进行训练的全卷积网络,能用于新视角合成、纹理编辑、场景重光照以及基于单张图片的3维重建等。基于体素方法的主要缺点是空间浪费较大,重建分辨率仍然受到显存限制。

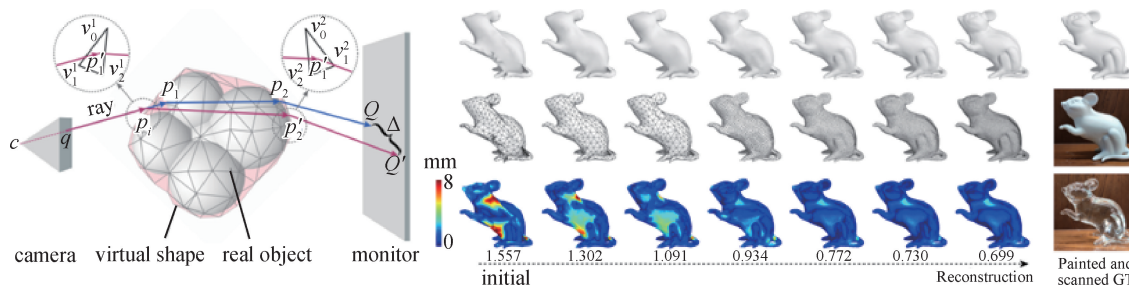


图4 基于网格表达的可微绘制用于透明物体表面重建(Lyu 等, 2020)

Fig. 4 Mesh-based differentiable rendering for the reconstruction of transparent surfaces(Lyu et al. , 2020)

1.2.3 基于点云的方法

点云是物体表面的采样,是3维场景的一种高效表达。然而由于离散、稀疏等特点,基于点云的方法很难直接通过绘制去计算像素误差。如图5所示,Lin 等人(2018)提出一种从密集点云合成新视角深度图像的可微分伪渲染器(pseudo-renderer)。该渲染器通过增加投影分辨率,并采用基于深度的最大池化操作(max-pooling)来处理碰撞和可见性问题。然而该方法受点云数量多少影响,为此,Insafutdinov 和 Dosovitskiy (2018)提出在每个点上附加一个高斯密度函数,再进行体素化投影,而 Navaneet 等人(2019)则在投影后的每个2维投影点上附加一个高斯函数,以此来生成点云绘制图像。Surface splatting (Zwicker 等,2001)是一种经典的点云绘制技术。最近 Wang 等人(2019)提出第1个可微分的 Surface splatting 绘制技术,可应用于基于点云的表面编辑、点云优化和逆向绘制等几何处理,相比基于网格的方法仅更新顶点位置或深度,可微分 Surface

splatting 技术能同时更新法向量以及模型剪影,并能处理较大的拓扑变化。这一方法的主要缺陷是,绘制时间开销较大,导致算法整体优化时间较长。对此,Han 等人(2020)最近再次提出仅用投影点与多视角2维轮廓的匹配程度来驱动3维点云的恢复。其创新在于提出了两项约束,一是投影点必须落在轮廓线区域内;二是投影点之间需互相远离。使用投影点省去了烦琐的绘制过程,从结果看,其仍然可较好地由2维约束推理3维结构。

1.2.4 基于隐式表达的方法

上述3种基于显式表达的可微绘制方法虽然取得了很大进展,但受限于显式表达本身,每种方法都存在一定的局限性,例如基于网格的方法往往需要可形变的网格模板或者一系列的3维团块,最终预测结果容易出现自相交和非封闭的情况,对于较大的拓扑变化也难以处理;基于体素的方法则受制于体素分辨率;点云虽然存储高效,但缺乏拓扑结构与连接性。

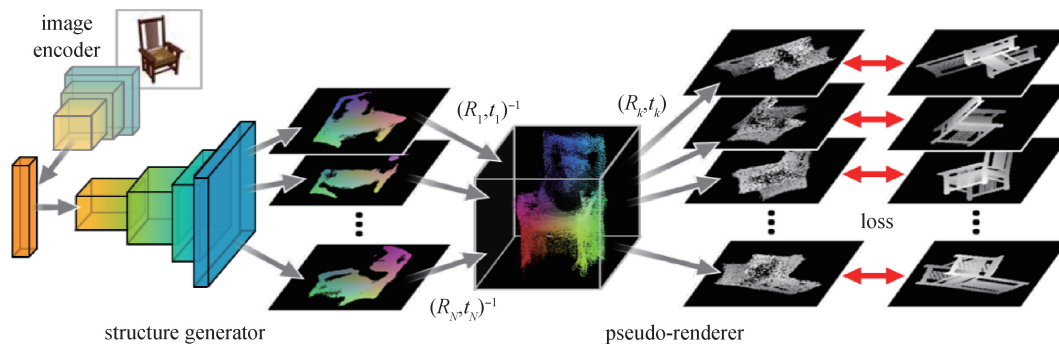


图5 基于可微绘制的单视角点云生成(Lin 等, 2018)

Fig. 5 Single-view point cloud generation based on differentiable rendering(Lin et al. , 2018)

随着研究深入,3 维隐式表达被发现非常适用于神经网络训练(Park 等, 2019; Mescheder 等, 2019; Chen 和 Zhang, 2019b), 因此催发出了一系列基于隐式表达的相关研究。3 维隐式表达可以粗略地分为符号距离场(signed distance fields, SDFs)和空间占用场(occupancy field), 二者都是对3 维空间中值面的连续表达。不同于显式表达, 隐式表达是连续的, 且其不受制于分辨率的性质, 使得神经网络可以处理任意分辨率和拓扑的水密形状数据。

为进行多视角图像3 维重建, Liu 等人(2019b)首次尝试结合神经网络隐式表达与可微绘制, 流程如图6 所示。首先将图像编码特征和3 维空间点联合输入多层感知机(multilayer perceptron, MLP)去预测场景的空间占用场, 然后提出了一种基于空间占用场的光线探测方式: 先在空间中采样一些稀疏的锚点(anchor points), 根据预测的占用概率为每个点附加一个球形支撑范围, 从而与光线相交, 最后采用沿光线的最大池化(max-pooling)方法得到物体在该视角的投影。

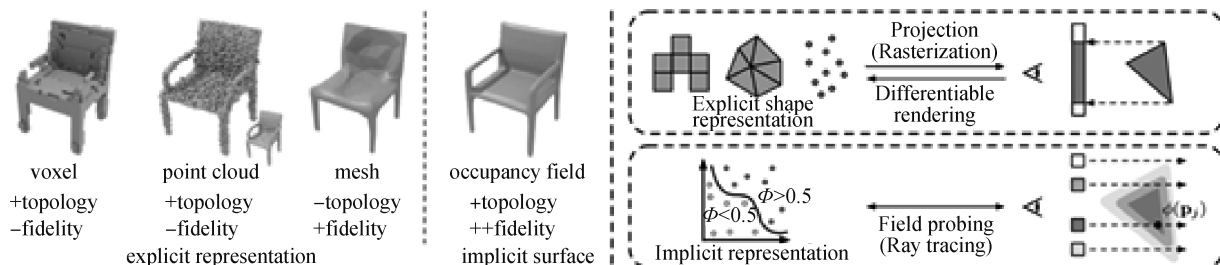


图6 基于空间占用场的可微绘制3 维重建(Liu 等, 2019b)

Fig. 6 3D reconstruction using occupancy field(Liu et al. , 2019b)

随后深度符号距离场(SDF)被引入可微绘制的3 维几何推理。受经典的SDF 函数绘制方法 Sphere tracing (Hart, 1996) 启发, Liu 等人(2020b) 与 Jiang 等人(2020) 分别提出了不同的基于深度符号距离场的可微绘制方法, 不同的是 Liu 等人(2020b) 选择绘制法向量、深度和剪影图, 而 Jiang 等人(2020) 仅绘制 Shading 结果。两个工作都展示出了不错的多视角甚至单视角3 维重建结果。

然而, 上述几种方法并没有考虑物体表面材质。Sitzmann 等人(2019b) 提出 SRN(scene representation network) 网络, 算法核心是通过全连接网络, 即多层感知机将场景几何与表面材质进行编码。具体

来说, 他们将空间中的3 维点坐标通过 MLP 转换为特征向量, 该特征向量编码了点最近表面的距离和表面材质颜色等。为绘制图片, 他们从像素出发, 沿光线采样3 维点, 然后将该点对应的特征向量输入长短期记忆网络(long short-term memory, LSTM) 去预测步进距离, 直至光线与物体表面相交。这个过程便是一个可微分光线步进(ray-marching)。相交点最后通过 1×1 卷积核转换为 RGB 颜色。该方法获得了较好的绘制效果和3 维结构。但是为了计算梯度, 必须存储训练期间的中间结果, 消耗了大量的显存。

DVR(differentiable volumetric rendering) (Niemi-

eyer 等,2020)使用多层感知机估计 3 维目标的空间占用场,然后同样利用 Ray-marching 的方法找到光线与表面的交点,最后利用共享参数的感知机对交点集进行绘制得到结果图像。与 SRN 网络不同,DVR 根据隐函数求导法则推导出了绘制目标函数相对网络参数梯度的代数解,使训练不需要存储中间参数,直接运算即可将参数回传,大大节省了显存需要。

不同于 DVR, IDR (implicit differentiable renderer) (Yariv 等,2020) 预测有符号距离场(SDF),并分离了几何预测与绘制操作(如图 7 所示)。IDR 利用隐函数求导法则以及有符号距离场的法线性质,进一步探究了隐式 3 维结构,图像与神经网络之

间的关系,显式推导出表面点与神经网络梯度和相机结构之间的表达,从而直接求得当前符号距离场与光线的表面交点,使整个过程可微分。IDR 再利用神经绘制器对前一步表面交点进行 BRDF (bidirectional reflectance distribution function) 估计,得到对应纹理颜色。IDR 的实验表明,该神经绘制器实际编码了物体表面的材质及整个场景光照信息。由于几何与绘制的分离,该方法还能实现不同场景间的表面纹理迁移,如图 8 所示。

Lin 等人(2020)针对单幅图像重建,通过将单幅图像中的物体轮廓进行距离变换,从而为深度符号距离场(其本质上是水平集)提供额外的监督信息,大大提高单幅图像 3 维重建效果。

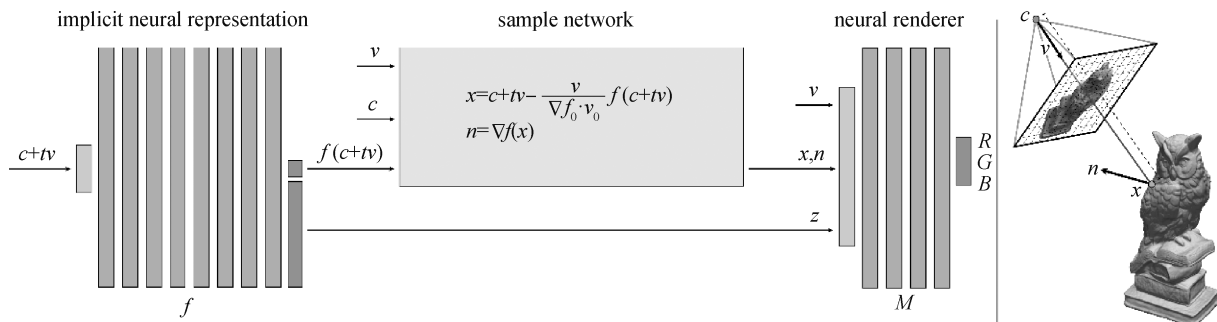


图 7 基于隐式表达网络与绘制网络的 3 维重建(Yariv 等, 2020)

Fig. 7 3D reconstruction using neural implicit function and rendering(Yariv et al., 2020)



图 8 基于 IDR 的表面纹理迁移(Yariv 等, 2020)

Fig. 8 Surface texture transfer based on IDR(Yariv et al., 2020)

1.3 基于可微绘制的逆向表观采集建模

表观采集建模的对象是随着空间位置、视角以及光照方向等因素变化而变化的高维表观函数,其常见的表达形式有 6 维的随空间变化而变化的双向反射分布函数 (spatially-varying bidirectional reflectance distribution function, SVBRDF) 和 6 维的双向纹理函数 (bidirectional texture function, BTF)。高维表观函数能够在虚拟世界中高真实感地呈现在任意观

察和光照条件下的原始物理表观。表观采集建模一般可以分为两个阶段:首先在采集阶段,获取物理对象在不同条件的表观测量数据;随后在重建阶段,将这些测量数据进行处理后得到表观的数字化表达。这里的根本性挑战在于维度上的不匹配:常见图像传感器被设计成只能对 2 维空间域进行有效直接采样,而表观采集建模需要以不同条件下拍摄的 2 维图像作为输入,恢复出 6 维甚至更高维度的数字化

表观结果。维度的不匹配导致了通常情况下输入数据中的有效信息量严重不足,使得复杂高维表观建模困难重重。

针对以上的根本性挑战,传统研究分别在采集和重建阶段做了大量的工作,试图改进建模效率和质量:在采集阶段,对表观所在的高维物理域进行采样,获取尽可能丰富的测量信息;在重建阶段,对采样样本进行有效的计算处理,最终转换成高维数字化表观表达形式。然而受限于相关理论、技术与工具,传统表观采集建模研究大都基于手工、临时(Ad-hoc)推导的先验、模型和算法,其假设条件在实际应用中往往难以满足,在质量、效率和鲁棒性上有着种种不足。因此与相对成熟的几何采集建模相比,高维表观采集建模的实用性/商用推广程度较低,难以满足文物数字化、电子商务虚拟展示等实际应用中的迫切需求。

随着深度学习的快速发展,表观采集建模方向的从业者开始思考并推进与人工智能的深度结合,涌现出一批优秀的研究工作及商业产品,形成了可微分表观采集建模的系列成果。与传统工作不同的是,深度学习的引入能够从大规模标注数据中自动学习出鲁棒的对应关系,这使得真正意义上的、数据驱动的表现采集建模成为可能。通过深度学习的最新技术和相关可微分工具,人们对采集/重建过程的部分甚至全部流程进行了端到端的自动设计与联合优化,与传统方法相比显著提升了表观采集建模的质量和效率。微软亚洲研究院董悦在其综述文章(Dong,2019)中,对截止2019年的基于深度学习的表观建模研究工作进行了充分的梳理、归纳和总结。国际学术界的可微分采集建模研究大致可以分为两类,以下分别进行简要介绍。

第1类研究是基于深度学习的表观建模,其输入数据和传统相关研究保持不变,主要聚焦于如何利用深度学习来自动得到更高效普适的先验、模型和算法,从而提升表观重建的各类指标(例如质量、效率等)。由于物理采集没有得到优化,此类工作的输入信息量有限,往往需要增加额外的假设来使得表观估计变得可行。

韩国三星电子的Kim等人(2017)提出了利用神经网络从多视角观测数据中估计匀质BRDF表观。法国蔚蓝海岸大学的Deschaintre等人(2018)提出了一种以闪光灯条件下的单张照片作为输入的

建模方法,通过大规模的过程式表观数据来训练深度神经网络,最终可以使用网络来估计出与输入图像对应的SVBRDF。相同的作者们在随后的工作中提出了一个与光照和视角信息无关的SVBRDF估计网络,能够从每一幅输入图像中自动抽取隐式特征向量,并通过最大池化来汇总,以支持任意数量的输入图像(Deschaintre等,2019)。使用在未知环境光照和靠近相机的闪光灯条件下的单幅图像作为输入,美国加州大学圣地亚哥分校的Li等人(2018c)提出一个回归网络来估计随空间变化而变化的反射率、粗糙度、法向量变化以及由自动分类出的表观类型所确定的单一匀质高光反射稀疏,并且进一步把神经网络拓展到可以同时估计随空间变化而变化的表观以及3维几何(Li等,2018c)。德国马克斯普朗克研究所(MPI)的Meka等人(2019)利用神经网络,从两幅在彩色梯度光照下的输入图像来估计出人脸上的各向同性反射表观。

第2类研究是表观物理采集与计算重建的联合优化,在第1类研究的基础上更进一步,采集时获取能更有效刻画物理表观特征的信息,并基于此数据生成相应的高效处理方法,实现从硬件端到软件端的、更一般意义上的端到端表观采集重建。美国加州大学圣地亚哥分校的Xu等人(2018)提出了能生成在新光照条件下场景表观的重光照方法。其主要贡献是优化了点光源的物理采样方式,并联合训练了采样网络与重建网络。这样做既能使得重建网络能基于采样网络的输入数据生成高质量的表观结果,又能让采样网络为重建网络提供优化过的输入数据。

除了学术界,国际产业界受到实际应用对表观数据强烈需求的驱动,推出了一系列专业级高维表观采集装备以及配套的集成解决方案。美国XRite公司发布了TAC7专业材质扫描仪(<https://www.xrite.com/categories/appearance/tac7>),用来获取平面样本的各向同性表观数据。单台扫描仪售价高达近10万美金,拥有32个呈圆弧分布的独立LED光源,通过机械转盘对样本进行快速旋转来获取不同视角和光照条件下的表观信息,总采集时间大约在1—2h。除了TAC7外,美国MURA公司的PBR纹理扫描仪(<https://www.muravision.com/>)和法国达索集团的DeltaTex2扫描仪也是具有类似功能的竞争产品。

2 国内研究进展

国内有一些工作尝试借助可微绘制提升 IBR 任务的性能。其中比较有代表性的是上海科技大学的虞晶怡团队开发的神经网络绘制引擎 (relightable

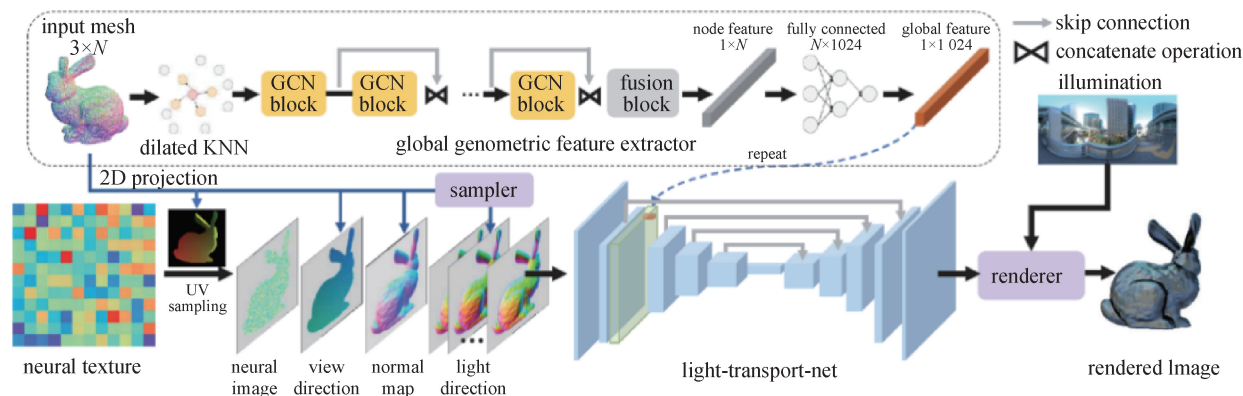


图9 RNR 流程图 (Zhang 等, 2020)

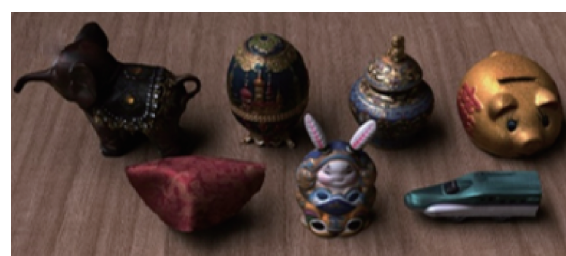
Fig. 9 RNR pipeline (Zhang et al., 2020)

国内学术界在可微分采集建模研究方面起步较早,紧跟并且部分引领了国际学术界的主流趋势,在部分方向上做出了具有创新性、指标国际领先的研究成果。中国科学技术大学与微软亚洲研究院的 Li 等人(2017)率先提出了一种基于自增强学习的建模方法,能从环境光照下的单张输入图片恢复出带单一高光反射分量的表观与法向量分布。清华大学与微软亚洲研究院的 Gao 等人(2019)提出了一种支持任意张数输入图像的神经重光照 (neural relighting),只需要使用两个相机(其中一个带闪光灯)从不同角度拍摄输入图像,就能通过训练获得复杂表观的神经网络表达形式,可以真实感地重现毛发、玉石等已有方法难以处理的材质外观。国内对利用图形学绘制生成对抗样本也进行了研究(冷凯轩,2020)。

浙江大学的 Kang 等人(2018)针对已有研究手工设计采集/重建过程,建模效率不高,无法处理近场布局等复杂情况的难点,提出了可微分采集建模架构,揭示了表观建模和跨领域编/解码之间的联系,把采集/重建过程完整映射到神经网络,自动学习了采集阶段

neural renderer, RNR) (Chen 等,2020),能够同时支持新视点合成和重光照。与在其他 IBR 中使用的可微绘制过程不同,RNR 考虑了环境光照、物体本征属性以及光传输过程对绘制效果的影响,所以更具“物理真实性”,因此无论是新视点合成还是重光照效果都比之前的方法好,方法流程如图9。

的最优光照,实现了目标驱动的、从硬件端到软件端的全自动建模过程优化,提高了整体效率。Kang 等人(2019)还将可微分架构推广到联合采集建模,使用自动学习到的光照来采集图像,并用神经网络在图像域逐像素精确分解了表观与几何信息,实现了复杂各向异性表观与几何的高精度联合建模。与已有工作相比,整体采集效率提高了1个数量级,重建结果如图10所示。



(a) 作为代表性成果被选为 ACM Transactions on Graphics 的期刊封底图(Kang 等, 2019)



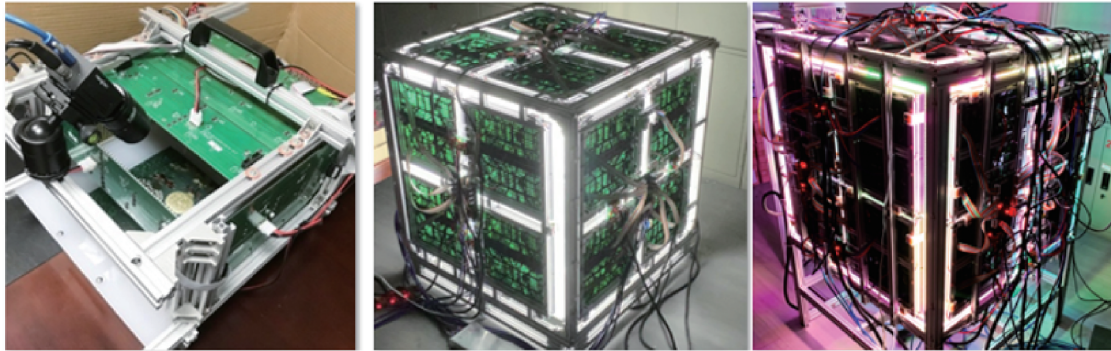
(b) Gao 等人(2020)论文中的结果

图10 复杂表观的可微分采集建模结果及对比

Fig. 10 Results of differentiable sampling in complex surface ((a) Kang et al. (2019) is selected as back cover of ACM Transactions on Graphics; (b) result of Gao et al. (2020))

浙江大学 CAD&CG 国家重点实验室还从底层 PCB (printed circuit board) 电路板开始,研制了多套具有自主知识产权的高密度光源采集装备,其关键指标达到了国际领先:在 0.5 m^3 体积内分布了 7 万多个独立控制的高亮 LED 光源,光照投射速度

4.8 万帧/s,总光通量近 6 万流明,如图 11 所示。国内产业界在表观采集装备上有着很好的发展潜力,已有多家科技创业公司布局相关专业装备研发,预期在工业检测、电子商务和影视游戏制作等领域将获得大规模的落地应用。



(a) 平面表观采集装备 (Kang 等, 2018)

(b) 非平面表观与几何联合采集装备 (Kang 等, 2019)

图 11 浙江大学 CAD&CG 国家重点实验室所研制的具有自主知识产权的高密度光源采集装备

Fig. 11 The acquisition equipment with high-density light sources developed in CAD&CG State Key Laboratory in Zhejiang University ((a) the appearance acquisition equipment for planar objects (Kang et al., 2018); (b) joint appearance and geometry acquisition equipment for 3D objects (Kang et al., 2019))

3 国内外研究进展比较

近年来国际上涌现出大量的可微绘制的理论创新和应用工作,目前国内在可微绘制的研究工作基本以应用研究为主,包括训练神经网络进行绘制、采用可微绘制进行透明物体的 3 维重建。在可微分采集建模的研究进展上,国外的主要优势和特点在于建模方向上布局全面,训练数据的积累丰富,并且产学研之间紧密联动。以美国加州大学圣地亚哥分校为例,Ravi Ramamoorthi 教授团队在光场、表观和几何等各类图形对象的可微分建模方面发表了一系列高水平论文 (Mildenhall 等, 2020; Zhang 等, 2019)。另外,众所周知大规模高质量训练数据对于基于机器学习的可微分方法而言至关重要。而表观数据由于其获取困难,相关的高质量数据库屈指可数。其中值得一提的是美国 Adobe 公司的 Adobe Stock 数据库,通过构建良性循环的付费社区生态,拥有了大量由全世界美工创造的高质量 3 维图形数据(包括高维表观),巧妙地绕开了表观采集能力不足所导致的困难。由于 Adobe 公司和多所一流大学之间有长期的科研合作,Stock 数据库被用于训练多篇顶尖可微分建模文章中所提出的神经网络,获得了良好

的效果,推进了相关研究的发展。相比之下,国内的主要优势和特点是布局可微分表观采集建模较早,并且在高密度采集装备指标等方面达到了国际领先。中国科学技术大学、微软亚洲研究院等单位率先提出了利用单幅输入图像来估计高维表观的可微分方法;浙江大学率先提出了可微分表观采集建模/表观与几何联合采集建模架构,在软硬件结合方面有着自己的特色。

国内与国外的主要差距在于产学研用的推进还不够深入,在可微绘制的工具、平台上工作较少,也缺乏理论方面的工作。随着“破四唯”等新研究成果评价导向的树立,这样的差距有望在未来能够缩小,甚至国内可以赶上国际先进水平。

4 发展趋势与展望

可微绘制以梯度回传作为工作基础,结合自动微分技术所发展的算法和平台能够有效提高正向绘制和逆向重建算法的开发效率,是当前计算机图形学和计算机视觉领域的热点研究方向之一。可以预见,可微绘制技术的发展对基于图像的 3 维信息推断等逆向问题将产生重要影响,对解决高沉浸感虚拟现实 3 维内容创建的效率问题具有现实的产业意

义。当然,作为新型工具,可微绘制还有很多值得研究的课题,距离最终运用可微绘制解决实际绘制、3 维重建等问题还有很多要做的工作:

1) 常见的损失函数描述绘制图像与目标图像的差异,但在什么特征空间描述比较符合人的感知? 使用什么距离函数? 是否可以用图像之外的数据衡量差异? 深度神经网络提供的多尺度特征在度量图像差异时可以同时考虑到底层和高层特征,模拟人的感知,可为解决这些问题提供一个新的角度。

2) 如何更好地避免局部极小值? 目前可微绘制的应用通常需要提供较好的初始值,否则在优化过程中难以得到理想的结果。支持全局光照,高光、透明材质的反射、折射等复杂现象的高真实感的渲染函数)可以提供更多的信息,更不易受到局部最小值的影响。但是,高真实感的渲染函数复杂度高,对微分计算及计算效率都提出了更多的挑战。

3) 如何拓展可微绘制技术处理不可微分的变化,如模型拓扑结构的改变? 如何发展几何的表达形式以高效地集成绘制流水线? 当前基于隐式函数的几何表达受到了广泛的关注。该表达不受拓扑变化和体表达分辨率限制,并与深度神经网络表现出的优良的函数逼近性能较好地结合,是一个值得关注的几何表达研究方向。

4) 如何更好地将可微绘制集成到机器学习管线中? 基于局部光照的可微绘制比较简单,但无法生成包含阴影、反射和折射的逼真的图像。基于全局光照的可微绘制可以生成带软影等全局光照效果的高真实感图像,但计算复杂度高,对于训练神经网络来说速度上过于缓慢。可以考虑引入游戏引擎中的实时绘制方法,在较短时间内绘制高度逼真的图像。例如,集成快速但复杂的实时绘制方法(如阴影贴图和环境贴图)到可微绘制中也是值得探索的研究方向。

参考文献 (References)

Bangaru S P, Li T M and Durand F. 2020. Unbiased warped-area sampling for differentiable rendering. *ACM Transactions on Graphics*, 39(6): #245 [DOI: 10.1145/3414685.3417833]

Bemana M, Myszkowski K, Seidel H P and Ritschel T. 2020. X-Fields: implicit neural view-, light- and time-image interpolation. *ACM Transactions on Graphics*, 39(6): #257 [DOI: 10.1145/3414685.3417827]

Chen W Z, Gao J, Ling H, Smith E J, Lehtinen J, Jacobson A and Fidler S. 2019. Learning to predict 3D objects with an interpolation-based differentiable renderer//*Neural Information Processing Systems*. Vancouver, Canada: [s. n.]: #8953765

Chen Z, Chen A P, Zhang G L, Wang C Y, Ji Y, Kutulakos K N and Yu J Y. 2020. A neural rendering framework for free-viewpoint relighting//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, USA: IEEE: 5598-5609 [DOI: 10.1109/CVPR42600.2020.00564]

Chen Z Q and Zhang H. 2019. Learning implicit fields for generative shape modeling//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, USA: IEEE: 5932-5941 [DOI: 10.1109/CVPR.2019.00609]

Deschaintre V, Aittala M, Durand F, Drettakis G and Bousseau A. 2018. Single-image SVBRDF capture with a rendering-aware deep network. *ACM Transactions on Graphics*, 37(4): #128 [DOI: 10.1145/3197517.3201378]

Deschaintre V, Aittala M, Durand F, Drettakis G and Bousseau A. 2019. Flexible SVBRDF capture with a multi-image deep network. *Computer Graphics Forum*, 38(4): 1-13 [DOI: 10.1111/cgf.13765]

Dong Y. 2019. Deep appearance modeling: a survey. *Visual Informatics*, 3(2): 59-68 [DOI: 10.1016/j.visinf.2019.07.003]

Gao D, Chen G J, Dong Y, Peers P, Xu K and Tong X. 2020. Deferred neural lighting: free-viewpoint relighting from unstructured photographs. *ACM Transactions on Graphics*, 39(6): #258 [DOI: 10.1145/3414685.3417767]

Gao D, Li X, Dong Y, Peers P, Xu K and Tong X. 2019. Deep inverse rendering for high-resolution SVBRDF estimation from an arbitrary number of images. *ACM Transactions on Graphics*, 38(4): #134 [DOI: 10.1145/3306346.3323042]

Han Z Z, Chen C, Liu Y S and Zwicker M. 2020. DRWR: a differentiable renderer without rendering for unsupervised 3D structure learning from silhouette images//*Proceedings of the 37th International Conference on Machine Learning (ICML)*. Virtual: [s. n.]

Hart J C. 1996. Sphere tracing: a geometric method for the antialiased ray tracing of implicit surfaces. *The Visual Computer*, 12(10): 527-545 [DOI: 10.1007/s003710050084]

Insafutdinov E and Dosovitskiy A. 2018. Unsupervised learning of shape and pose with differentiable point clouds//*Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montréal, Canada: Curran Associates Inc.: 2807-2817

Jiang Y, Ji D T, Han Z Z and Zwicker M. 2020. SDFDiff: differentiable rendering of signed distance fields for 3D shape optimization//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 1248-1258 [DOI: 10.1109/CVPR42600.2020.00133]

Kajiya J T. 1986. The rendering equation//*Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Tech-*

- niques. *Virtual Event: Association for Computing Machinery*; 143-150 [DOI: 10.1145/15922.15902]
- Kang K Z, Chen Z M, Wang J P, Zhou K and Wu H Z. 2018. Efficient reflectance capture using an autoencoder. *ACM Transactions on Graphics*, 37(4); #127 [DOI: 10.1145/3197517.3201279]
- Kang K Z, Xie C H, He C G, Yi M Q, Gu M Y, Chen Z M, Zhou K and Wu H Z. 2019. Learning efficient illumination multiplexing for joint capture of reflectance and shape. *ACM Transactions on Graphics*, 38(6); #165 [DOI: 10.1145/3355089.3356492]
- Kato H, Ushiku Y and Harada T. 2018. Neural 3D mesh renderer//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA; IEEE: 3907-3916 [DOI: 10.1109/CVPR.2018.00411]
- Kim K, Gu J W, Tyree S, Molchanov P, Nießner M and Kautz J. 2017. A lightweight approach for on-the-fly reflectance estimation//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy; IEEE: 20-28 [DOI: 10.1109/ICCV.2017.12]
- Leng K X. 2020. Research on Generation of Adversarial Examples Based on Graphics. Chengdu; University of Electronic Science and Technology of China (冷凯轩. 2020. 基于图形的对抗样本生成技术研究. 成都: 电子科技大学)
- Li T M, Aittala M, Durand F and Lehtinen J. 2018a. Differentiable Monte Carlo ray tracing through edge sampling. *ACM Transactions on Graphics*, 37(6); #222 [DOI: 10.1145/3272127.3275109]
- Li T M, Lehtinen J, Ramamoorthi R, Jakob W and Durand F. 2015. Anisotropic Gaussian mutations for metropolis light transport through Hessian-Hamiltonian dynamics. *ACM Transactions on Graphics*, 34(6); #209 [DOI: 10.1145/2816795.2818084]
- Li X, Dong Y, Peers P and Tong X. 2017. Modeling surface appearance from a single photograph using self-augmented convolutional neural networks. *ACM Transactions on Graphics*, 36(4); #45 [DOI: 10.1145/3072959.3073641]
- Li Z Q, Sunkavalli K and Chandraker M. 2018b. Materials for masses; SVBRDF acquisition with a single mobile phone image//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany; Springer: 74-90 [DOI: 10.1007/978-3-030-01219-9_5]
- Li Z Q, Xu Z X, Ramamoorthi R, Sunkavalli K and Chandraker M. 2018c. Learning to reconstruct shape and spatially-varying reflectance from a single image. *ACM Transactions on Graphics*, 37(6); #269 [DOI: 10.1145/3272127.3275055]
- Lin C H, Kong C and Lucey S. 2018. Learning efficient point cloud generation for dense 3D object reconstruction//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA; AAAI
- Lin H L, Wang C Y and Lucey S. 2020. SDF-SRN: learning signed distance 3D object reconstruction from static images//Proceedings of the 34th Annual Conference on Neural Information Processing Systems. Vancouver, Canada; [s. n.]
- Liu L J, Gu J T, Lin K Z, Chua T S and Theobalt C. 2020a. Neural sparse voxel fields//Proceedings of the 34th Conference on Neural Information Processing Systems. Vancouver, Canada; [s. n.]
- Liu S C, Chen W K, Li T Y and Li H. 2019a. Soft rasterizer: a differentiable renderer for image-based 3D reasoning//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South); IEEE: 7707-7716 [DOI: 10.1109/ICCV.2019.00780]
- Liu S C, Saito S, Chen W K and Li H. 2019b. Learning to infer implicit surfaces without 3D supervision//Proceedings of the 33rd Conference on Information Processing Systems. Vancouver, Canada; [s. n.]
- Liu S H, Zhang Y D, Peng S Y, Shi B X, Pollefeys M and Cui Z P. 2020b. DIST: rendering deep implicit signed distance function with differentiable sphere tracing//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE: 2016-2025 [DOI: 10.1109/CVPR42600.2020.00209]
- Lombardi S, Simon T, Saragih J, Schwartz G, Lehrmann A and Sheikh Y. 2019. Neural volumes: learning dynamic renderable volumes from images. *ACM Transactions on Graphics*, 38(4); #65 [DOI: 10.1145/3306346.3323020]
- Loper M M and Black M J. 2014. OpenDR: an approximate differentiable renderer//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland; Springer: 154-169 [DOI: 10.1007/978-3-319-10584-0_11]
- Loubet G, Holzschuch N and Jakob. 2019. Reparameterizing discontinuous integrands for differentiable rendering. *ACM Transactions on Graphics*, 38(6); #14228 [DOI: 10.1145/3355089.3356510]
- Luan F J, Zhao S, Bala K and Gkioulekas I. 2020. Langevin Monte Carlo rendering with gradient-based adaptation. *ACM Transactions on Graphics*, 39(4); #140 [DOI: 10.1145/3386569.3392382]
- Lyu J H, Wu B J, Lischinski D, Cohen-Or D and Huang H. 2020. Differentiable refraction-tracing for mesh reconstruction of transparent objects. *ACM Transactions on Graphics*, 39(6); #195 [DOI: 10.1145/3414685.3417815]
- Meka A, Häne C, Pandey R, Zollhöfer M, Fanello S, Fyffe G, Kowdle A, Yu X M, Busch J, Dourgarian J, Denny P, Bouaziz S, Lincoln P, Whalen M, Harvey G, Taylor J, Izadi S, Tagliasacchi A, Debevec P, Theobalt C, Valentin J and Rhemann C. 2019. Deep reflectance fields: high-quality facial reflectance field inference from color gradient illumination. *ACM Transactions on Graphics*, 38(4); #77 [DOI: 10.1145/3306346.3323027]
- Mescheder L, Oechsle M, Niemeyer M, Nowozin S and Geiger A. 2019. Occupancy networks: learning 3D reconstruction in function space//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE: 4455-4465 [DOI: 10.1109/CVPR.2019.00459]
- Mildenhall B, Srinivasan P P, Tancik M, Barron J T, Ramamoorthi R and Ng R. 2020. NeRF: representing scenes as neural radiance

- fields for view synthesis//Proceedings of the 16th European Conference on Computer Vision. Glasgow, Scotland: Springer; 405-421 [DOI: 10.1007/978-3-030-58452-8_24]
- Navaneet K L, Mandikal P, Agarwal M and Babu R V. 2019. CAPNet; continuous approximation projection for 3D point cloud reconstruction using 2D supervision//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Hawaii, USA; AAAI; 8819-8826
- Nguyen-Phuoc T, Li C, Balaban S and Yang Y L. 2018. RenderNet; a deep convolutional network for differentiable rendering from 3D shapes//Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS). Montréal, Canada; [s. n.]
- Niemeyer M, Mescheder L, Oechsle M and Geiger A. 2020. Differentiable volumetric rendering; learning implicit 3D representations without 3d supervision//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE; 3501-3512 [DOI: 10.1109/CVPR42600.2020.00356]
- Nimier-David M, Speierer S, Ruiz B and Jakob W. 2020. Radiative backpropagation; an adjoint method for lightning-fast differentiable rendering. *ACM Transactions on Graphics*, 39(4): #146 [DOI: 10.1145/3386569.3392406]
- Nimier-David M, Vicini D, Zeltner T and Jakob W. 2019. Mitsuba 2; a retargetable forward and inverse renderer. *ACM Transactions on Graphics*, 38(6): #203 [DOI: 10.1145/3355089.3356498]
- Park J J, Florence P, Straub J, Newcombe R and Lovegrove S. 2019. DeepSDF; learning continuous signed distance functions for shape representation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE; 165-174 [DOI: 10.1109/CVPR.2019.00025]
- Rezende D J, Eslami S M A, Mohamed S, Battaglia P, Jaderberg M and Heess N. 2016. Unsupervised learning of 3D structure from images//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain; Curran Associates Inc.; 5003-5011
- Shum H Y, Chan S C and Kang S B. 2007. *Image-Based Rendering*. Boston: Springer [DOI: 10.1007/978-0-387-32668-9]
- Sitzmann V, Thies J, Heide F, Nießner M, Wetzstein G and Zollhöfer M. 2019a. DeepVoxels; learning persistent 3D feature embeddings//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE; 2432-2441 [DOI: 10.1109/CVPR.2019.00254]
- Sitzmann V, Zollhöfer M and Wetzstein G. 2019b. Scene representation networks; continuous 3D-structure-aware neural scene representations//Proceedings of the 33rd Conference on Neural Information Processing Systems. Vancouver, Canada; [s. n.]
- Thies J, Zollhöfer M and Nießner M. 2019. Deferred neural rendering; image synthesis using neural textures. *ACM Transactions on Graphics*, 38(4): #66 [DOI: 10.1145/3306346.3323035]
- Tulsiani S, Zhou T H, Efros A A and Malik J. 2017. Multi-view supervision for single-view reconstruction via differentiable ray consistency//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA; IEEE; 209-217 [DOI: 10.1109/CVPR.2017.30]
- Veach E. 1997. *Robust Monte Carlo Methods for Light Transport Simulation*. Stanford; Stanford University
- Vu H H, Labatut P, Pons J P and Keriven R. 2012. High accuracy and visibility-consistent dense multiview stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5): 889-901 [DOI: 10.1109/TPAMI.2011.172]
- Wang Y F, Serena F, Wu S H, Öztireli C and Sorkine-Hornung O. 2019. Differentiable surface splatting for point-based geometry processing. *ACM Transactions on Graphics*, 38(6): #230 [DOI: 10.1145/3355089.3356513]
- Wu B J, Zhou Y, Qian Y M, Cong M L and Huang H. 2018. Full 3D reconstruction of transparent objects. *ACM Transactions on Graphics*, 37(4): #103 [DOI: 10.1145/3197517.3201286]
- Wu J J, Wang Y F, Xue T F, Sun X Y, Freeman W T and Tenenbaum J B. 2017. MarrNet; 3D shape reconstruction via 2.5D sketches//Proceedings of the 31st International Conference on Neural Information Processing Systems. Los Angeles, USA; Curran Associates Inc.; 540-550
- Xu Z X, Sunkavalli K, Hadap S and Ramamoorthi R. 2018. Deep image-based relighting from optimal sparse samples. *ACM Transactions on Graphics*, 37(4): #126 [DOI: 10.1145/3197517.3201313]
- Yan X C, Yang J M, Yumer E, Guo Y J and Lee H. 2016. Perspective transformer nets; learning single-view 3D object reconstruction without 3D supervision//Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona, Spain; Curran Associates Inc.; 1704-1712
- Yariv L, Kasten Y, Moran D, Galun M, Atzmon M, Basri R and Lipman Y. 2020. Multiview neural surface reconstruction with implicit lighting and material [EB/OL]. [2020-12-01]. <https://arxiv.org/pdf/2003.09852v2.pdf>
- Zhang C, Miller B, Yan K, Gkioulekas I and Zhao S. 2020. Path-space differentiable rendering. *ACM Transactions on Graphics*, 39(4): #143 [DOI: 10.1145/3386569.3392383]
- Zhang C, Wu L F, Zheng C X, Gkioulekas I, Ramamoorthi R and Zhao S. 2019. A differential theory of radiative transfer. *ACM Transactions on Graphics*, 38(6): #227 [DOI: 10.1145/3355089.3356522]
- Zhang R, Tsai P S, Cryer J E and Shah M. 1999. Shape-from-shading; a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8): 690-706 [DOI: 10.1109/34.784284]
- Zhao S, Jakob W and Li T M. 2020. Physics-based differentiable rendering; from theory to implementation//ACM SIGGRAPH 2020 Course. Virtual Event: Association for Computing Machinery; 14 [DOI: 10.1145/3388769.3407454]
- Zhu R, Galoogahi H K, Wang C Y and Lucey S. 2017. Rethinking

reprojection; closing the loop for pose-aware shape reconstruction from a single image//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy; IEEE; 57-65 [DOI: 10.1109/ICCV.2017.16]

Zwicker M, Pfister H, van Baar J and Gross M. 2001. Surface splatting//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. Virtual Event; Association for Computing Machinery; 371-378 [DOI: 10.1145/383259.383300]

作者简介



许威威, 1975年生, 男, 研究员, 主要研究方向为3维重建、物理仿真和虚拟现实。

E-mail: xww@cad.zju.edu.cn

周漾, 男, 助理教授, 主要研究方向为3维重建、几何分析、纹理合成、图像处理。E-mail: zhouyangvcc@szu.edu.cn

吴鸿智, 男, 副教授, 主要研究方向为高密度采集装备与可微分建模。E-mail: hwu@acm.org

过洁, 男, 副研究员, 主要研究方向为真实感绘制、实时绘制、复杂材质分析建模以及虚拟现实技术。

E-mail: guojie@nju.edu.cn