

高度真实感表面材质逆向恢复方法综述

邢佑鑫¹⁾, 欧城睿²⁾, 徐延宁¹⁾, 王璐^{1)*}

¹⁾ (山东大学软件学院 济南 250101)

²⁾ (山东大学机电与信息工程学院 威海 264209)

(luwang_hcivr@sdu.edu.cn)

摘要: 大规模复杂三维场景的真实感渲染是三维动画、数字人、虚拟现实、虚拟仿真、数字孪生、视频游戏等领域中的关键技术之一, 而具备高度真实感的表面材质是三维渲染的重要基石, 其参数贴图的准确性和精细程度决定了渲染结果的逼真程度. 文中根据使用设备和材质表示形式的不同, 将相关工作划分为精密仪器测量、过程式节点图材质制作和轻量级设备恢复 3 大类, 总结了表面材质逆向恢复方法的评估指标和能力. 首先介绍了基于精密仪器的测量方法, 对专业设备与材质表示模型发展的相关性进行分析; 然后简述了基于过程式节点图的材质重建和生成方法, 对相关软件的特点进行总结, 重点分析了基于风格损失优化的过程式材质逆向方法的发展现状及面临的问题; 再回顾了基于轻量级设备的材质高效重建和生成方法, 阐述了轻量级设备采集的标准, 并从提取材质隐空间特征、估计材质自相似性、抑制高光对材质恢复的影响、材质自监督训练设计、多模态引导材质生成等方面, 对比和分析了多种深度学习方法的优劣; 最后对高度真实感表面材质逆向恢复的研究现状进行了归纳和总结, 指出了现阶段表面材质逆向恢复任务仍面临采集条件严苛、清漆和闪烁等复杂材质恢复困难、网络训练成本高且极易受到强高光影响的问题; 未来将围绕特殊类别材质建模、强高光的软遮罩预测与区域重绘、三维物体材质提取以及精准材质编辑等方向展开.

关键词: 真实感渲染; 材质逆向; 材质生成; 过程式材质模型

中图分类号: TP391.41 DOI: 10.3724/SP.J.1089.2024-00724

A Survey of Inverse Recovery Methods of Highly Realistic Surface Materials

Xing Youxin¹⁾, Ou Chengrui²⁾, Xu Yanning¹⁾, and Wang Lu^{1)*}

¹⁾ (School of Software, Shandong University, Jinan 250101)

²⁾ (School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai 264209)

Abstract: Realistic rendering of large-scale complex 3D scenes is one of the key technologies in the fields of 3D animation, digital humans, virtual reality, virtual simulation, digital twins, video games, etc. Surface materials with a high degree of realism are the cornerstones of 3D rendering, and the accuracy and fineness of their parametric maps determine the degree of realism of the rendering results. In this paper, according to the different equipment and material representations, the related work is classified into three categories: precision equipment measurement, procedural node graph material creation, and lightweight equipment-based recovery. The evaluation indexes and capabilities of surface material inverse recovery methods are also summarized. Firstly, this paper

收稿日期: 2024-12-01; 修回日期: 2025-04-29. 基金项目: 国家重点研发计划(2022YFB3303200); 国家自然科学基金(62272275); 泰山学者工程资助(tsqn202312231). 邢佑鑫(1997—), 男, 博士研究生, CCF 学生会员, 主要研究方向为真实感渲染、实时渲染、材质表现建模、材质逆向恢复; 欧城睿(2001—), 男, 硕士研究生, 主要研究方向为真实感渲染、材质逆向恢复; 徐延宁(1975—), 男, 博士, 副教授, 硕士生导师, CCF 高级会员, 主要研究方向为人机交互与虚拟现实、真实感材质建模与物理渲染、三维 GIS 与 BIM; 王璐(1981—), 女, 博士, 教授, 博士生导师, CCF 专业会员, 论文通信作者, 主要研究方向为真实感渲染、实时渲染、材质表现建模、高性能渲染.

introduces the measurement methods based on precision equipment and analyzes the correlation between professional equipment and the development of material representation models. Then, material reconstruction and generation methods based on procedural node graphs are briefly described, the characteristics of related software are summarized, and the development status and problems of procedural material inversion methods based on style loss optimization are highlighted. In addition, this paper reviews the efficient reconstruction and generation of materials based on lightweight equipment, describes the criteria for lightweight equipment acquisition, and compares and analyzes the advantages and disadvantages of various deep learning methods in terms of extracting latent spatial features of materials, estimating self-similarity of materials, suppressing the effect of highlights on material recovering, designing self-supervised training of materials, and generating materials by multimodal guided generation. Finally, the state of the art in inverse recovery of highly realistic surface materials is summarized. This paper pointed out that the current surface material recovery task still faces the problems of harsh acquisition conditions, difficulties in recovering complex materials such as clear coatings and glints, high cost of network training, and susceptibility to strong highlights. Future work will focus on the modeling of special classes of materials, soft-mask prediction, and region redrawing for strong highlights, material extraction from 3D objects, and accurate material editing.

Key words: realistic rendering; material recovery; material generation; procedural material model

1 本文概述

高质量的表面数字材质是实现三维场景高度真实感渲染的关键。在渲染领域,通常使用双向反射分布函数(bidirectional reflectance distribution function, BRDF)^[1]定义物体表面的性质,这些性质决定了材质与光线交互的结果。此外,由于一些材质在空间上表现出明显的变化,因此采用空间变化 BRDF(spatially varying BRDF, SVBRDF)^[2]对这类材质进行建模。近年来,材质相关研究在物体表面外观建模方面取得了进步,但由于材质本身结构的复杂性以及光线、阴影的干扰,其逆向恢复仍是一个具有挑战性的研究领域。早期的方法直接测量材质的 SVBRDF/BRDF 物理数值并应用到渲染过程中。随着材质模型的发展,研究人员提出一些经验模型(如 Phong^[3], Blinn-Phong^[4], Ward^[5]等模型),以及基于物理的 BRDF 模型(如 Cook-Torrance^[6], SpongeCake^[7]等模型),并参与到表面材质逆向恢复任务中。这些模型通过一组材质参数控制物体表面材质的渲染结果,在实际渲染过程中更加灵活可控。此外,基于物理的 BRDF 遵循能量守恒等特性,在视觉效果上比经验模型更逼真。以 Cook-Torrance 模型为例,常见的材质参数包括漫反射反照率、法线、粗糙度、镜面反射反照率等,这类材质模型同样支持广泛的材质类

别,如金属、塑料、皮革、木材、纸张、地面、砖墙等。本文主要对通用材质的逆向恢复技术进行综述,不包含特定领域内的材质,如人脸、布料、车漆等特殊材质,并着重关注二维平面上的 SVBRDF/BRDF 反射率和 SVBRDF/BRDF maps,不涵盖 BTF(bidirectional texture function)的相关内容。逆向恢复技术包含重建和生成 2 类:重建针对输入目标材质图像获取精准的材质反射率参数,而生成则是在多模态条件输入下获取多种新材料质的反射率参数。

使用昂贵且精密的硬件设备获得高精度的材质时,对采集环境和材质样本的尺寸有极高的要求,同时,逐一采集样本会耗费大量的时间。一些专业工作者使用 Substance Designer^[8]等专业的过程式材质制作软件同样可以创建丰富的材质(如图 1 所示),但要求从业人员具有丰富的专业知识和经验。针对表面材质参数获取代价昂贵的问题,在数字材质恢复领域中出现了一系列重建和生成的方法。

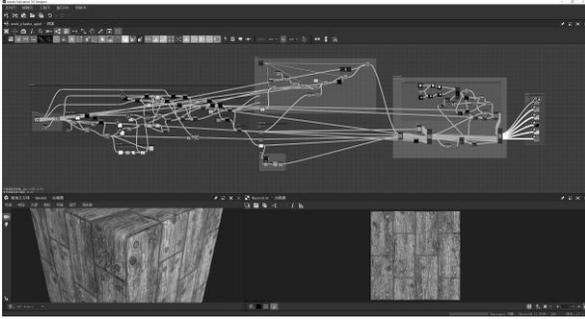


图1 过程式节点图示例

本文对高度真实感表面材质逆向恢复方法进行综述。(1) 对材质逆向恢复的理论背景进行概述并对相关方法进行分类;(2) 详细介绍基于精密仪器测量方法的发展过程;(3) 介绍基于过程式节点图的材质正向制作以及逆向重建和生成方法;(4) 分别介绍基于轻量级设备的材质重建方法和生成方法;(5) 对高度真实感材质逆向恢复方法进行总结,并阐述该领域的发展趋势,为研究人员顺利开展后续研究工作提供帮助。

2 背景及分类

2.1 背景介绍

真实感表面材质的反射率通过 BRDF/SVBRDF 表示。BRDF 是关于参数 $(\theta_1, \phi_1, \theta_0, \phi_0)$ 的四维函数,其中, θ_1 和 ϕ_1 表示入射方向的仰角和方位角, θ_0 和 ϕ_0 表示出射方向的仰角和方位角。因此,入射光方向和光线的反射方向(观察方向)分别被定义为 $\omega_1 = (\theta_1, \phi_1)$ 和 $\omega_0 = (\theta_0, \phi_0)$ 。BRDF 表示为

$$f(\omega_1, \omega_0, \lambda) = \frac{dL_o(\omega_0, \lambda)}{L_i(\omega_1, \lambda) \cos \theta_1},$$

其中, L_i 和 L_o 分别表示入射和反射的辐射率, λ 表示波长。BRDF 分为漫反射项和镜面反射项。漫反射项描述光与物体表面作用时引起的无方向性反射,通常使用朗伯定律^[9]近似。镜面反射项则高度依赖于入射光和反射光的方向,通常在反射角

度附近集中反射,包含法线分布函数、几何项和非涅尔项。其中,法线分布函数描述微表面法线的分布情况,表示表面光滑程度对镜面反射的影响,如 GGX 分布函数^[10];几何项描述表面微结构对光的遮蔽和散射效应的影响,如 Smith 联合遮蔽阴影函数^[11];菲涅尔项则表示入射光与表面法线的角度对反射率的影响^[12]。通常,光在表面的反射强度随着入射角的增大而增大,特别是当入射角接近切线时。此外 BRDF 包含各向同性和各向异性 2 种类别,各向同性 BRDF 在方位角变化时保持不变。

SVBRDF 在 BRDF 的基础上额外地增加了 2 个维度,是关于参数 $(x, y, \theta_1, \phi_1, \theta_0, \phi_0)$ 的 6 维函数,其中, x, y 表示物体表面的位置。SVBRDF 假设物体表面上每个点的 BRDF 各自独立,因此其逆向恢复比 BRDF 恢复的难度更高。

2.2 表面材质逆向恢复方法分类

由于材质反射率或反射率参数图在物理上与人射光方向和观察方向密切相关,因此使用复杂设备测量表面材质的方法最先被提出,并在实践中得到了应用。按照测量过程的特点,这类方法大致划分为密集型直接测量方法、结合材质分析与表示的测量方法,以及结合深度学习的测量方法。此外,随着分形几何和 Perlin 噪声等数学理论的成熟,出现了一系列基于过程式节点图的材质制作软件。早期,主要由专业人员通过正向的过程式节点图流程创作表面材质;现阶段,逐步出现了基于可微代理的自动优化方法,允许根据条件图逆向重建或者生成大量的多样化数字材质。随着电子设备的快速迭代和发展,使用智能手机、平板电脑等设备拍摄照片成为更加便捷的选择,越来越多的方法基于这些轻量级设备采集的材质照片作为输入,并配合深度神经网络以数据驱动的方式重建或生成材质。现有方法的分类如图 2 所示。

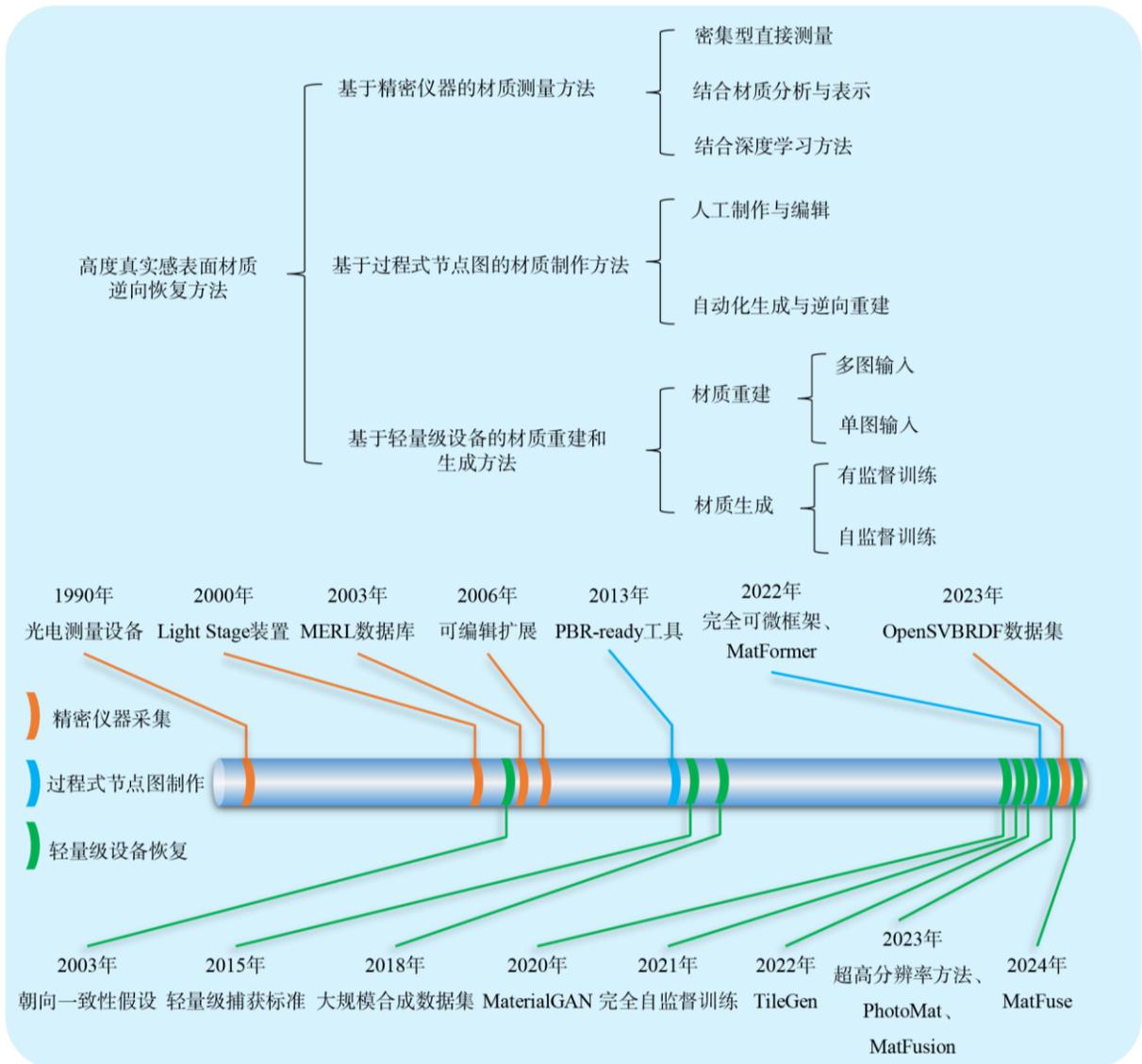


图 2 材质逆向恢复方法分类及关键方法时间轴

图 2 中，划分的 3 大类方法均包含数据驱动式方法，为了便于拆分和表述，本文按照采集工具和材质表示形式的不同，将这些方法划分为基于精密仪器的材质测量方法、基于过程式节点图的材质制作方法，以及基于轻量级设备的材质重建和生

成方法，其中，轻量级设备特指智能手机、平板电脑等易得的成像设备。此外，本文从材质的数据特征、拟合方法、材质模型表示等角度，对上述 3 大类方法的相关性和不同之处进行分析和总结，如表 1 所示。

表 1 分类方法的异同点

指标类型	精密仪器测量方法	过程式节点图制作方法	轻量级设备恢复方法
数据特征	SVBRDF/BRDF	SVBRDF/BRDF	SVBRDF/BRDF
拟合方法	直接估计、聚类优化、神经网络	聚类优化、神经网络	神经网络
材质模型表示	Phong ^[3] , Blin-Phong ^[4] , Ward ^[5] , Cook-Torrance ^[6]	过程式节点图、Cook-Torrance ^[6]	Phong ^[3] , Blin-Phong ^[4] , Ward ^[5] , Cook-Torrance ^[6] , SpongeCake ^[7]
时间周期	数小时	数小时	训练数天，推理数分钟
精度	高	中	中
应用特性	需要复杂的专业设备；适合构建大规模、超高精度、高分辨率的材质数据库	需要具有经验丰富的专业人员辅助编辑；适合生成后期可编辑且无限分辨率的材质	需要轻量级别的采集工具，如手机、平板电脑等；适合逆向重建或生成高分辨率、强多样性的材质

下文，将详细介绍这些方法采用的技术手

段，并提供便于理解的示例和结果。

2.3 材质重建和生成的评价指标

对于材质重建, 传统的定量评价方式是对估计的材质和材质的真实值, 使用均方误差(mean squared error, MSE)、均方根误差(root mean squared error, RMSE)等指标逐像素比较差异; 一些方法使用结构相似度度量指标(structural similarity index measure, SSIM)^[13]对图像的亮度、对比度和结构进行评估, 或使用峰值信噪比(peak signal-to-noise ratio, PSNR)^[14]评价图像的重建质量; 还有的方法使用学习感知图像块相似度(learned perceptual image patch similarity, LPIPS)^[15]指标, 基于感知损失度量 2 幅图像之间的差别. 与上述传统的指标相比, LPIPS 指标更符合人类的视觉感知情况. 材质生成任务通过计算生成材质的 CLIP-IQA^[16]和 FID 分数^[17]来评估不同生成数据集的质量和多样性.

为了更好地对比材质重建和生成的方法的能力, 本文以定性的形式细化材质重建和生成的指标和配置条件:

(1) 输入数量. 重建材质时输入材质图像的数量, 输入数量越多, 该方法恢复的效果越好, 但相应的采集过程、处理过程越复杂.

(2) 光照条件. 重建材质时输入材质图像的光照类型, 常见的光照条件有手机闪光灯、LED 灯、环境光照、LCD 屏幕光照等.

(3) 几何条件. 重建或生成材质图像时额外输入的三维几何形状.

(4) 分辨率. 重建或生成材质图像的分辨率大小.

(5) 多样性. 重建或生成不同类别材质的丰富程度, 以及相同材质类别下材质细节内容的丰富程度.

(6) 材质性质. 材质是否具有自相似性, 即稳定材质, 是指由随机的重复元素或结构组成的材质, 由于这类材质不具备全局的纹理图案, 因此其恢复难度相对较低.

(7) 处理高光的能力. 在重建或生成的材质反射率参数图中剔除额外高光的能力.

(8) 编辑能力. 重建或生成的材质是否支持进一步的编辑, 如插值生成新材质、修改生成内容等.

(9) 训练/计算/推理时间. 重建或生成材质方法的性能.

3 基于精密仪器的材质测量方法

在过去的几十年中, 虽然已有一些较为成熟的真实感材质模型表示方法, 但仍无法轻松地在真实感渲染过程中重现现实世界的外观, 原因是缺失基于物理的真实感材质数据库. 为了构建精准的真实感材质数据库, 一些研究者提出使用精密的采集装置来精准地测量真实世界的材质反射率或参数图, 但由于材质的测量受到采集密度、采集设备、光照配置和重建模型等方面的影响, 因此该类测量任务具有较高的复杂度.

本文将材质测量方法分为密集型直接测量方法、结合材质分析与表示的测量方法, 以及结合深度学习的测量方法 3 类. 选择其中较为关键的代表性材质测量设备和方法进行详细介绍, 并给出便于理解的基本步骤来说明该方法的过程.

3.1 密集型直接测量方法

早期, 该类方法采用光电测量设备^[18-19]直接采集 BRDF. Murray-Coleman 等^[18]设计了经典的采集装置, 并根据 BRDF 的定义提供 4 个自由度的采集方向, 该装置包括: 一个 MR16 白炽灯、硅光电二极管型反射率和透射率探测器、电机驱动的底部转环和灯架等, 如图 3 所示.

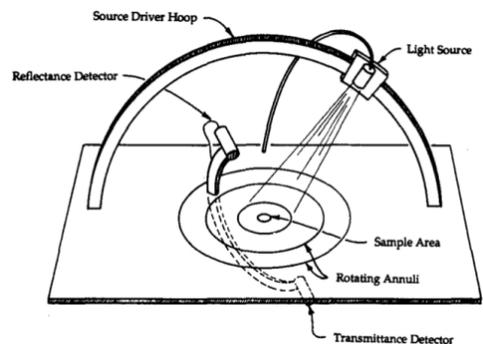


图 3 BRDF 光电探测采集设备^[18]

随着采集密度的增加, 直接测量方法的存储和时间代价也随之增加. 基于对常见材质的观察, 大部分建筑材料的 BRDF 是各向同性的, 即在许多角度范围内 BRDF 的变化非常缓慢, 因此这类方法中通常选用更少的数据值描述这部分的 BRDF, 从而更好地权衡采集分辨率和存储之间的关系. 为了加速对金属等各向异性材质的测量速度, 在传统的光电探测设备基础上, Ward^[5]首次采用搭配鱼眼镜头的 CCD 相机捕获材质表面光线反射的情况, 有效地简化了原始的机械结构. 此外, 由于早期阶段的硬件设备条件有限, 因此材质反

射率采集的精度也在一定程度上受到了限制。

随着工业生产力的进步,采集装置也得到不断地发展并逐渐成熟。其中,LED 灯取代白炽灯,数码相机取代角光谱反射计和 CCD 相机,串行测量变为并行测量。如图 4a 所示,Gu 等^[20]的设备使用 20 面体结构的圆顶作为骨架,并在骨架之上布置了多台高分辨率数码相机和上百个白色 LED 光源,实际采集时对摄像机和光源使用定制的控制

光源圆顶的特殊结构设计,这类装置可以在同一时间并行捕获多组反射率,因此该方法的实用性和效率显著提升;该类装置结构稳定,有效地避免了由于机械转动而引起的测量误差;这类方法还通过数码相机拍摄一段时间内高分辨率、高动态范围的材质变化,并支持在渲染引擎中使用测量的 SVBRDF 反射率重新渲染物体干燥、生锈的变化过程,如图 4b 所示。目前,已知角分辨率最好的 BRDF 测量装置是由 Obein 等^[21]在 2015 年提出的,角分辨率为 0.022°。

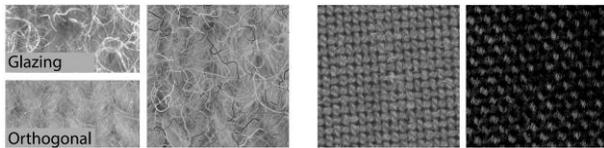


a. SVBRDF 采集设备

b. 金属茶壶锈蚀效果示意图

图 4 Gu 等^[20]设计的时间变化 SVBRDF 采集设备及恢复效果

早期大部分测量工作集中在单点或单一尺度上,忽略了各向异性透射效果,因此无法捕获布料材质的参数,尤其是微观尺度上的飞线,如图 5 所示。



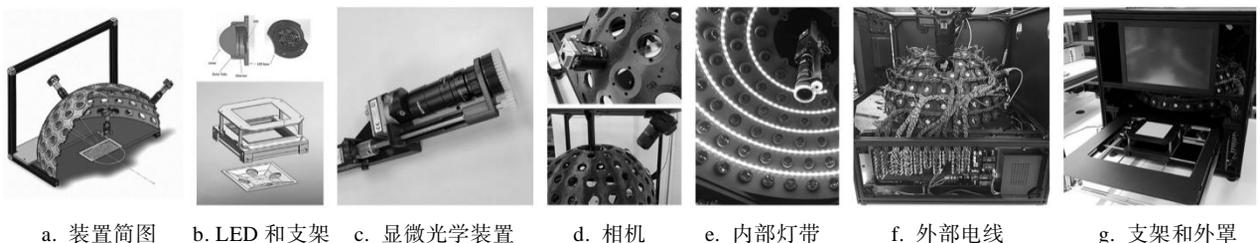
a. 飞线恢复效果

b. 纤维级别恢复效果

图 5 基于图像的飞线测量及纤维级别的估计结果^[22]

得益于现代工业相机和镜头的发展,基于小型高分辨率工业相机和液态变焦镜头对的微观光学系统^[22]被设计并得到应用,该系统要求自身体积足够小,以尽量避免对测量材质产生遮挡和阴影,同时具备很强的放大能力。与传统镜头相比,液态镜头更符合该场景的应用需求,它具有小型

化、快速对焦、变焦范围广、无机械运动部件、适应性强等优点,是一种新型的光学元件。安装液态镜头的测量设备中,无论相机与采集材质的距离如何,都可以在毫秒级实现稳定的电子对焦,而不影响成像质量。此外,得益于三维打印技术的进步,设备可以被更灵活、更精准地设计并制作出来。如图 6a 所示的半封闭式圆顶装置中,在保持自身结构稳定的同时,可以尽量避免外界环境光对内部材质的测量产生影响。在上述设备的基础上,现有的方法可以成功地捕获光纤级别(3~20 μm)的微观细节,以及约 11 cm×11 cm 的中尺度纹理。图 6b 中,顶部是校准的 LED 设计图,底部是主支架和背光支架;图 6e 是圆顶内部提供漫反射光的 LED 灯带;图 6f 是圆顶外部的电线。



a. 装置简图

b. LED 和支架

c. 显微光学装置

d. 相机

e. 内部灯带

f. 外部电线

g. 支架和外罩

图 6 Garcés 等^[22]设计的各向异性反射率材质采集装置

尽管如此, 测量现实世界中的所有材质显然是不现实的, 会耗费大量的时间成本并付出高昂的存储代价, 因此这类方法仅适合对代表性材质进行精确测量。

3.2 结合材质分析与表示的测量方法

与直接测量方法主要针对采集装置的硬件和设计进行优化不同, 当前讨论的材质反射率测量任务尽管也使用专业的材质测量设备, 但是重点关注材质反射率模型表示方法的革新, 目的是降低材质采集的密度与成本并提高测量效率和精度。

通常, 这类任务对采集的 BRDF 数据使用主成分分析法(principal component analysis, PCA)^[23] 匹配相应的线性变换, 在保留 BRDF 反射率大部分有效信息的前提下将数据从高维映射到低维^[24]。与传统的 BRDF 反射率测量方法相比, 这类方法节省了存储空间, 同时支持重建并衍生出大量 BRDF 进行高度真实感渲染。此外, 通过对自然界中的材质观察发现, 材质反射率可以被建模为一组小基函数的线性组合, 这些基函数可以通过对不同材质进行大量且密集的采样并分析其反射率函数得到。因此, Matusik^[25]提出基于小波分析的 BRDF 反射率函数表示方法, 首先通过分析常见的表面反射率函数确定函数域内每个点的局部信号变化, 然后利用小波分析推导出 BRDF 的公共小波基和相应的非均匀采样模式, 最后在均匀空间中重建 BRDF。该方法不再盲目地关注于所有的材质, 而是针对相对少量材质(100 个)的测量数据进行分析 and 插值, 显著地缩短了各向同性反射率的测量时间; 同时, 基于线性和非线性插值和外推的方法构建多样且可靠的 MERL(mitsubishi electric research laboratories)数据库^[26], 该数据库因高质量的数据和开放的访问政策, 逐渐成为全球范围内研究人员广泛使用的基准数据集。

与材质 BRDF 反射率相比, SVBRDF 反射率不仅描述光线在某个特定点上的反射情况, 还考虑了表面不同位置的反射特性差异, 因此更能表现现实世界材质的全局纹理变化, 更加全面地描述物体的反射特性; 但是其测量难度也随之增加, 需要更复杂的测量设备和数据处理算法。为了提升测量准确度, 一类方法利用采集区域像素的自相似性先验将反射性质相同的像素聚类到一起,

并对该组像素统一使用 Levenberg-Marquardt 方法迭代估计反射率的参数图^[27]。与逐像素估计相比, 这种聚类估计的方式极大地避免了因为部分像素点采集数据异常带来的失败风险, 因此在确保精度的基础上提高测量方法的鲁棒性, 同时降低了反射率估计所需的时间。基于聚类的 Levenberg-Marquardt 方法虽然对一般情况能得到较好结果, 但对于具有强烈高光的金属、塑料、光滑地板等材质, 无法较好地恢复其高光特性; 此外, 该类方法没有考虑到布料材质透光的问题, 采集建模后会出现较大误差。

使用传统的反射率模型建模采集的数据时, 难以表示复杂的材质^[28]。非参数的表示形式提供了更通用, 更精准的表示能力, 且 Lawrence 等^[2]通过逆向着色树框架扩展了这种表示形式的可编辑功能, 在树形表示结构中, 叶子节点为采样的一维曲线和二维图像, 表示直观的概念(如镜面高光形状或纹理图), 以相乘再相加的形式组合成中间节点。这种表示形式比参数模型更准确, 比其他非参数方法更直观, 非常适合交互式渲染和编辑。该方法的结构图如图 7 所示。

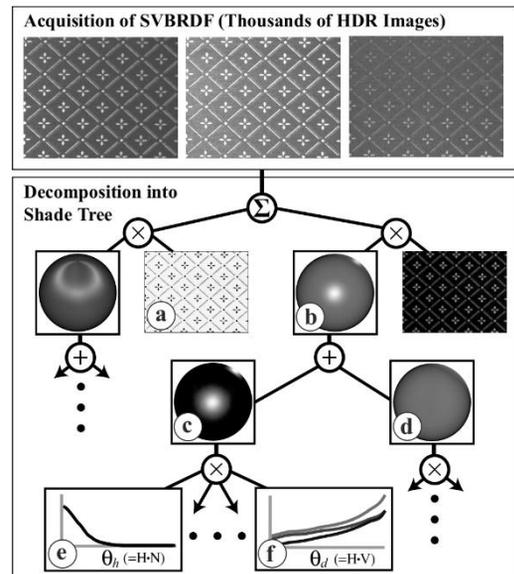


图 7 着色树框架^[2]

另一类方法不直接测量材质的 BRDF 或 SVBRDF, 而是测量材质表面的非局部反射场。Debevec 等^[29]提出 Light Stage 装置, 如图 8 所示。非局部反射场能够模拟复杂的光传播现象, 如光的多次反射或散射, 尤其适用于那些材质具有复杂表面特性(如皮肤、布料等)时的反射行为。而传统的模型(如 BRDF 或 SVBRDF)无法完全捕捉这些现象。由于本文重点关注 BRDF/SVBRDF 的测

量, 因此这部分内容详见文献[30].



图 8 Light Stage 装置^[29]

本文发现, 材质表示和材质反射率测量的发展是相互促进的关系. 材质表示方法的理论先于测量方法被提出. 随着测量工作的推进, 已有的

材质表示方法在存储上和精度上不能更好地表达测量结果, 因此更多用于特定任务的材质表示方法被配套提出, 如基于基函数线性组合的材质反射率表示方法^[25]等, 这类方法的显著特点是高压缩比和快速卷积还原, 能够在保证数据准确率的基础上大大缩减存储; 但是, 这类材质表示方法往往受限于特定的数据集分布, 不具备普适性. 为了更好地适配材质表示模型, 研究人员设计了更多实用的测量设备.

3.3 结合深度学习的测量方法

当前, 研究人员提出一系列基于神经网络优化和辅助拟合 SVBRDF 参数图的方法^[31-33], 首先通过训练非线性编码器和线性解码器预测准确的 Lumitexel 数值, 然后借助神经网络强大的表示能力, 逐点根据不同光照模式下的光照效果解码对应的 Lumitexel, 最后借助 Lumitexel 拟合 SVBRDF 参数图. 其中, OpenSVBRDF^[33]是这类方法的典型代表, 该方法使用的测量设备如图 9 所示. 其中, 图 9a 中包含位于顶部和侧面的 2 台相机.

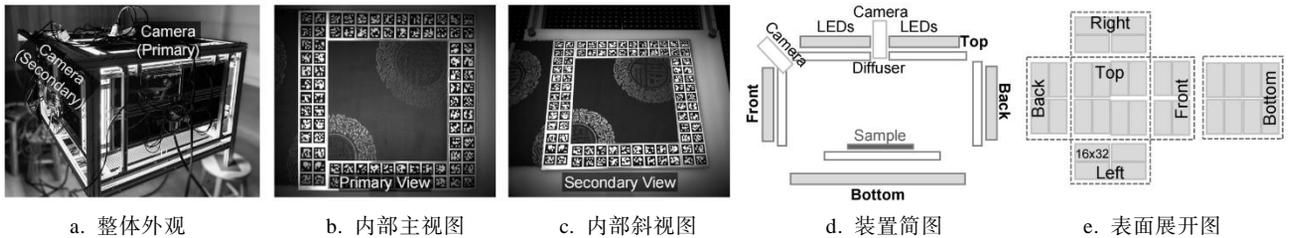


图 9 OpenSVBRDF 使用的测量设备^[33]

OpenSVBRDF 方法^[33]的基本步骤如下:

Step1. 初步训练预测 lumitexel 的网络, 以 64 个光照模式和 2 组不同视角下的 Lumitexels 的点积作为输入, 以主视角下的 Lumitexels 的预测值作为输出. 其中, 训练数据集是合成数据. 训练流程图如图 10a

所示.

Step2. 继续微调预测 lumitexel 的网络, 微调过程采用真实数据进行优化. 微调流程图如图 10b 所示.

Step3. 用渲染损失优化并拟合 SVBRDF 参数图.

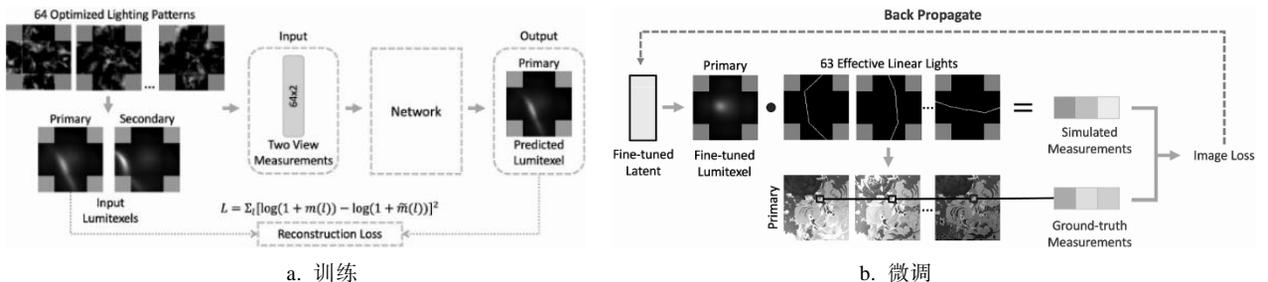


图 10 Lumitexel 预测网络训练及微调流程^[33]

由于整个网络微调 and 拟合过程是针对每个

材质且逐像素点进行的, 因此必然消耗大量的训

练和推理时间;但是,借助两阶段的 Lumitexel 训练和微调过程,该方法将每像素点处 BRDF 的丰富先验压缩到 Lumitexel 的神经表示中,同时在微调阶段提供大量采集的现实材质照片提高真实感,所以在 SVBRDF 参数图拟合时能准确地预测包含强高光材质的参数值.该方法提供了目前第一个大规模的空间变化各向异性测量材质数据库,涵盖木材、织物、金属等 9 大类别,包含共计 1 000 组高质量的近平面 SVBRDF 参数图,是材质测量领域的突破性方法.除了采集准确率的提升,这一阶段的测量方法在采集范围和采集尺度上也有重大突破.

3.4 讨论

测量材质反射率和参数图不仅需要专业、高精度的装置,还需要配套的材质分析和降维方法,同时利用材质反射率的分布规律减少数据采集时间.在设备方面,应该尽可能地限制可移动件的数量,减少震动和硬件故障发生的记录,最终实现稳健、高质量、高精度和高效的数字材质采集和重建.

4 基于过程式节点图模型的材质制作

分形几何理论支持生成具有自相似性的复杂图案,为计算机模拟自然界中的纹理提供了数学基础. Perlin 噪声能够生成具有随机性和连续

性的噪声图案,常用于模拟自然界的云、山脉等,同样为过程式材质生成提供了数学基础.过程式材质模型是一种基于单向图状数据结构的材质生成模型,与传统方法测量得到的材质模型不同,过程式材质模型通过一系列数学公式、算法和随机过程,以单向图的形式动态地生成各种各样的材质,如图 1 所示.单向图中的每个网络节点表示一种图像操作类型(滤波、变换等),这种方式生成的材质在内存占用上非常紧凑且与分辨率无关,并且提供显式的编辑操作(如节点拖拽和连接等),可以非常容易地修改材质模型的参数.

本节首先介绍人工正向设计创造过程式材质模型的方法,然后介绍自动化生成和逆向恢复过程式材质模型的技术.

4.1 人工制作与编辑

大部分过程式材质模型是由熟练的艺术家使用专业软件手工创作的,这些软件提供了丰富的节点、数学计算工具和不同材质类别的工作流程,辅助艺术家们高效地创建各种复杂的材质.常见的专业软件有 Substance Designer^[8], Houdini^[34], Unreal Engine Material Editor^[35], Unity Shader Graph^[36]等,其中,Substance Designer 是专用于过程式材质建模的软件;其他则是专业渲染或游戏引擎软件,但包含了过程式材质建模的功能.本文对上述软件的特点进行总结,如表 2 所示.

表 2 过程式材质建模相关软件的特点

名称	特点
Substance Designer ^[8]	内置丰富的节点库,涵盖各种纹理生成、颜色调整、噪声生成等功能;支持同众多三维软件无缝集成
Houdini ^[34]	强大的物理模拟,可以结合物理模拟生成更加真实的材质,如液体、布料等;支持创建自定义节点扩展软件功能
Unreal Engine Material Editor ^[35]	支持实时预览 shader 效果;与 Unreal 引擎深度集成;节点功能丰富,种类多
Unity Shader Graph ^[36]	支持实时预览 shader 效果;与 Unity 引擎深度集成;节点相对简单,易于上手

过程式材质创建的每个步骤对应于节点图中每个节点的参数设置和与其他节点的拖拽式连线.注意:任意节点上参数的修改都可以显式地影响生成的材质参数图的结果,意味着整个过程以实时的状态进行,编辑者在高度熟悉节点的功能和基本工作流之后,可以创建出高度逼真的材质.然而,熟悉海量节点并了解材质的基本属性不是一件容易的事情,具有非常高昂的学习成本.

为此,下面介绍自动化生成和逆向恢复过程式材质模型的相关技术.

4.2 自动化生成与逆向重建

4.2.1 过程式材质节点图的自动化生成

通过分析过程式节点图的特点,定义节点图为有向无环图 $g=(N,E)$,它由节点 $N=\{n_1,n_2,\dots\}$ 和边 $E=\{e_1,e_2,\dots\}$ 组成.给定一组节点参数,节点图输出 SVBRDF 参数图,如漫反

射反照率、粗糙度等. 节点图自动化生成材质面临的难点如下:

- (1) 材质节点图的结构是无规则的, 且允许有任意数量的节点;
- (2) 每个节点的类型和参数数量是任意的;
- (3) 节点的输入槽和输出槽的参数是有语义的, 不同节点的输入槽和输出槽随意搭配可能导致材质生成无效;
- (4) 节点和节点之间的连线数量是任意的, 且节点图中可以包含节点构成的长链, 长链上的远程节点之间允许长距离连接.

上述难点意味着节点图中有大量各种类型的节点, 且节点之间允许存在长距离的依赖关系. 传统的序列生成和图生成方法, 如图神经网络^[37]和循环神经网络(recurrent neural network, RNN)^[38]等, 不再适合处理这类任务. 受到自然语言处理领域中长序列理解任务的启发, Devlin 等^[39]提出基于 Transformer 网络的过程式节点图生成方法, 通过注意力机制关注节点图中不同节点之间的依赖关系, 确保图中所有部分都可以获得相同的关注度. 这类方法的典型代表是 MatFormer^[40], 由于注意力机制的内存和计算代价与序列长度相关, 因此为了避免这种昂贵的代价, 该方法将节点图的生成拆分成 3 个步骤进行优化, 依次分别生成节点序列、每个节点的参数序列和节点与节点之间输入输出槽的有向边序列, 避免了直接处理整个节点图, 如图 11 所示.

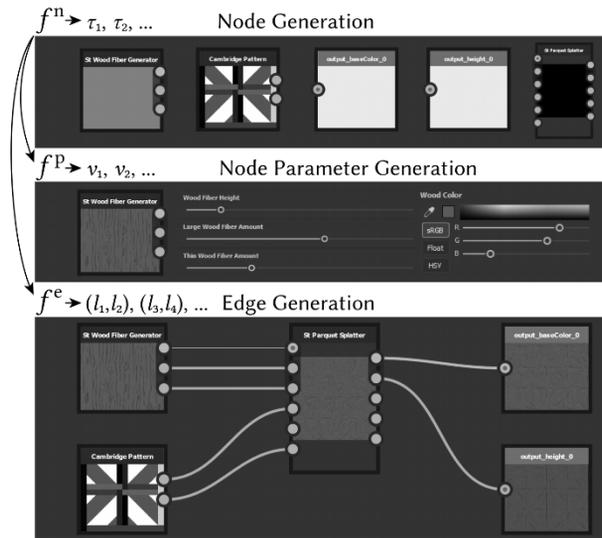


图 11 MatFormer 的材质图表示^[40]

MatFormer^[40]是首个过程式节点图的自动化生成方法, 经过训练后, 其网络既支持无条件的材质节点图生成, 也支持在材质节点图的初步框

架基础上补全缺失部分, 如图 12 所示. 本文发现, 这类方法生成的材质具有高度自相似性, 即生成的材质有着大量相似且重复的结构. 但是, 该方法目前无法处理对于带有语义的全局图案的材质. 后续或许可以通过结合大语言模型和扩散模型提供精准的全局纹理描述和生成. 此外, 该方法不支持以目标图像为引导生成过程式材质节点图并匹配目标材质图像. 为此, Hu 等^[41]在 MatFormer^[40]基础上引入基于 CLIP 的编码器, 允许在多模态的提示条件(文本或图像)下生成多个节点图或补全, 进一步降低了用户设计的门槛; Li 等^[42]则在合成图像和真实图像上, 使用外观相似性激励函数微调网络进行图像条件引导, 采用强化学习的框架, 提高了过程式材质节点图参数生成器的预测精度. 但是, 这类方法并没有完全解决过程式材质节点图的可微计算问题, 而是将生成的过程进行拆分, 使得序列(节点序列、参数序列、边序列)的优化是可微的.

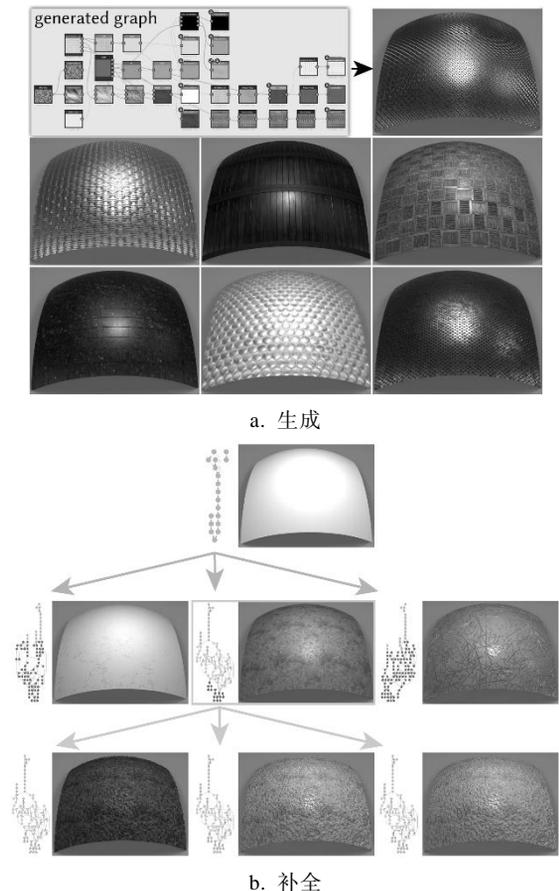


图 12 过程式材质节点图自动化生成及补全^[40]

4.2.2 过程式材质节点图的逆向恢复

早期的逆向恢复方法以手机等轻量级设备拍摄的材质照片作为输入, 输出对应的过程式材

质节点图和相应的材质参数图. 研究人员依托开源社区, 收集并整理了大量的真实世界的材质照片和高质量的过程式材质节点图, 为了将二者关联, 通常使用 K -means 聚类方法^[43], 按照不同的材质风格将真实世界的材质照片和过程式材质节点图划分为若干子类并进行粗略的对应, 方便根据输入图像选择合适的节点图^[44]; 同时, 还训练卷积神经网络 (convolutional neural network, CNN), 根据输入的材质照片估计过程式材质节点图的参数, 生成最终的过程式材质结果. 由于仅仅是分类后对材质照片和过程式材质节点图进行简单的对应, 因此生成的材质结果与输入的目标图片仍有差距. 为了缩减误差, 常用一个额外的风格迁移步骤在像素级别进行微调, 但是过程式材质节点图会丧失编辑功能.

另一类方法采用基于神经网络风格损失优化的材质逆向技术, 分为节点图参数优化和节点图选择^[45]2 部分. 首先使用预训练好的 VGG-19 网络^[46], 针对目标材质图像计算纹理描述符 (用 Gram

矩阵表示); 然后输入到全连接层网络中估计过程式材质节点图的参数, 并用可微的参数图估计器计算材质参数图 (漫反射反照率、法线、粗糙度等); 最后执行可微渲染步骤得到渲染的材质结果, 并与输入材质照片计算风格损失. 其中, 参数图估计器代表一系列可微操作, 包括像素值操作 (颜色、对比度编辑) 和图像处理 (滤波、扭曲、混合等), 这些操作由节点参数控制 (如滤波核的大小、混合操作的不透明度等), 并且与 CNN 的卷积、Point-wise 操作 (ReLU, Sigmoids 等激活函数) 对应, 可以方便地转化为现代深度学习框架处理. 但是, 添加离散噪声和生成结构化纹理的操作不包含在可微步骤内, 这部分参数通常被设置为固定值. 上述过程中, 经过初步训练后得到节点图参数估计网络的初始参数, 再继续通过渲染结果的风格损失和参数图的 L_1 损失, 从预先构建的大量节点图池中选取最优的节点图, 对应的网络管线如图 13 所示.

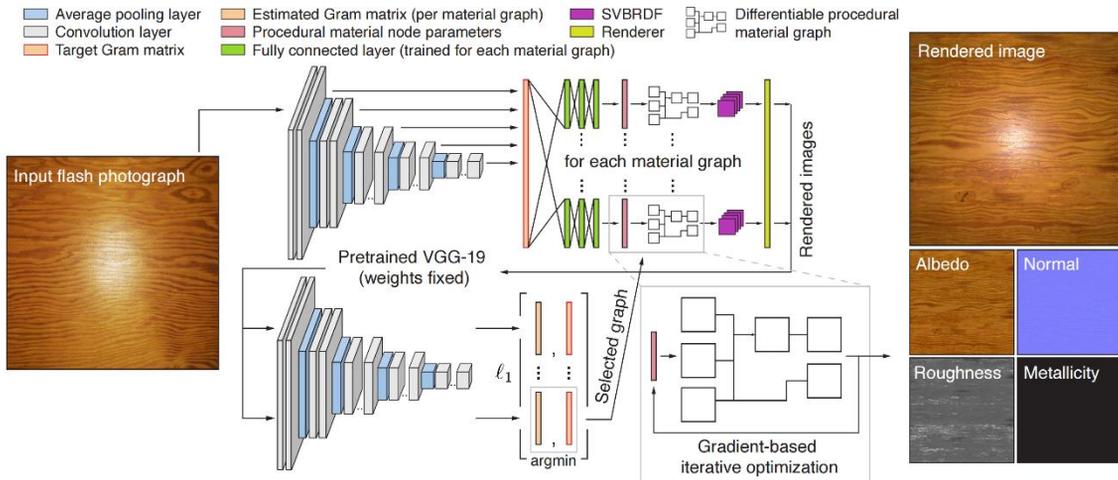


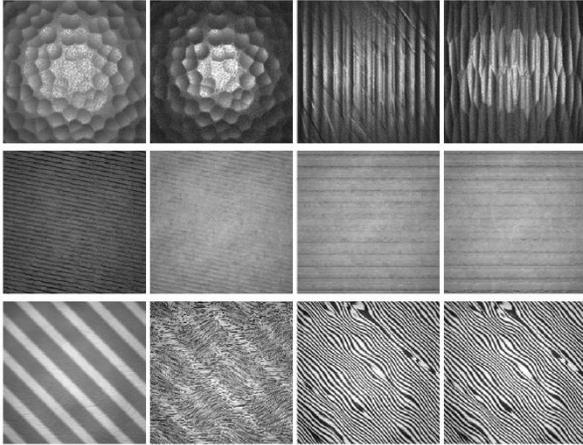
图 13 基于神经网络风格损失优化的材质逆向方法流程^[45]

上述方法是在已有的大量预定义的节点图的基础上选取合适的节点图, 需要大规模数据集做支撑. 为了避免这个问题, Hu 等^[47]提出一种基于半自动化材质层级拆分的方法, 可以在简单的人工标注下, 自动地以频谱感知的匹配算法将每个材质图片或材质参数图拆分为树状的子材质; 使用逐步滤波策略针对子材质构建多层的噪声模型, 捕获不同尺度的局部变化; 最后对连续参数和离散参数分别采用 Powell 方法^[48]和高斯过程的贝叶斯优化方法^[49]进行优化.

与聚类匹配的过程式材质逆向方法相比, 基于神经网络优化的方法保留了可编辑性, 但本质上二者并没有解决整个节点图的可微计算问题, 而是按照节点类型将节点拆分成可微和不可微部分并分别优化, 同时严重依赖于预先设定的节点图框架或者人工的引导, 因此无法优化多类型材质的全部特性.

基于可微代理的方法^[50]将大部分节点的可微性进行统一. 给定一个任意的、不可微的二维图像生成器 $G(\theta)$, 其中, θ 表示过程参数, 通过

可微代理可以将该不可微的图像生成器转化为可微的图像生成器。使用 CNN 近似 $G(\theta)$ ，并通过训练再现 $G(\theta)$ 的行为。通过多阶段可微优化，基于可微代理的方法提供了一个比以往工作更通用的材质外观匹配优化解决方案，支持恢复更贴近于目标输入的结果，如图 14 所示。



a. 初始化 b. MATch^[45] c. Hu 等^[50]方法 d. 目标图

图 14 2 种方法的结果对比

4.3 讨论

无论是节点图的生成还是逆向恢复，本质上均是在寻找合适的方法，将异构有向图优化困难的问题拆解为部分参数可微或整个节点图可微的问题。然而，尽管大部分节点可以转化为可微节点，但是有一些复杂节点，如 FX-map，由于具有难以区分的离散和随机行为，很难做到可微，因此后续仍需要进一步的研究以解决这类问题。

5 基于轻量级设备的数据驱动式材质重建和生成

基于过程式节点图模型的材质创作方法保留了节点图的结构和节点参数，可以方便地进行编辑和扩展并支持超高的分辨率。另外一种数据驱动式的材质恢复方法中，放弃了中间的过程式节点图结构，以轻量级设备(手机、平板电脑等)拍摄的材质照片作为输入，直接在像素级别生成相应的材质反射率或参数图。尽管这类方法丧失了过程式节点图的强大控制能力和高分辨率，但是不会受到不可微节点的限制，且提高了方法的易用性，即使是行业之外的人士也能轻松使用。本文详细调研了目前最先进的基于轻量级设备的数据驱动式材质恢复技术。首先以轻量级设备采集为起点；然后分别介绍以多图输入和单图输入

作为条件的材质参数逆向重建方法，以及以生成网络为核心的材质参数生成方法。对各类方法使用关键技术和网络结构进行分析，选取其中较为关键或者新颖的方法进行详细介绍，并对相应方法的优缺点进行总结。

5.1 轻量级设备采集标准

为了用轻量级设备实现丰富的反射率测量，Aittala 等^[51]首次规范了采用手机闪光灯捕获材质照片的配置标准，其示意图如图 15 所示。在拍摄过程中，手机位于被拍摄材质平面上方固定位置并与之保持平行。由于闪光灯与手机的相机镜头十分靠近，因此在每个采集图像像素位置处的照明方向都近似与相机的观察方向一致。为了减少饱和和高光对输入图像的影响，还有的方法额外提供一张相同条件但不开闪光灯的拍摄照片作为补充。除了上述采集方式外，还有一类方法支持 in-the-wild 采集，即在任意环境光照下使用手机拍摄，这种条件输入下恢复材质的难度更高，但同时更能体现恢复方法的强大能力。

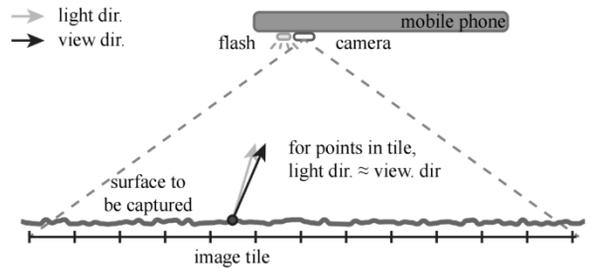


图 15 Aittala 等^[51]的拍摄配置

5.2 材质重建

轻量级设备的引入简化了材质采集过程，但同时给材质恢复带来挑战。现有的方法通常使用环境光贴图模拟自然光照或室内光照，使用点光源模拟手机闪光灯简化材质采集过程中的光照条件限制。还有的方法假设采集物体的几何模型已知来提供法线的先验。基于此，研究者在若干特定的假设下提出传统的直接估计或优化方法，以及基于深度神经网络的直接估计或优化方法来逆向恢复材质反射率或反射率参数图，不再使用昂贵的采集设备获取材质属性，或者避免由专业艺术家耗费精力创作材质。根据材质逆向过程中需要输入图像的数量，本文将非过程式的材质重建方法划分为多图输入的方法和单幅图像输入的方法 2 大类，分别进行介绍。

5.2.1 多图输入

(1) 传统的材质逆向重建方法

传统方法通常在考虑对象的几何形状(平面或非平面)和照明条件(受控制或不受控制)下, 使用优化拟合数据到 SVBRDF/BRDF 模型. 为了简化采集和建模, 避免复杂的形状估计, SVBRDF/BRDF 的研究通常集中于平面几何形状.

a. 镜面反射量和漫反射量的拆分. 早期的方法^[52]主要关注镜面反射量和漫反射量的恢复, 结合光谱分析对人眼视网膜感知颜色的 3 种原色刺激量进行估计, 并推导出相应的颜色分布模型用于拆分镜面反射和漫反射. 随着硬件的发展, Ghosh 等^[53]直接借助线性偏振器^[54]拆分漫反射和镜面反射, 并用数据拟合的方法建模 BRDF, 降低了采集输入图像的数量; Riviere 等^[55]则利用液晶屏幕的固有偏振分离漫反射和镜面反射率, 并将输入图像数量减少到了 1~2 幅, 且该方法支持在广泛类别的平面材质样本上实现高质量的反射率恢复和中观尺度的结构重建.

b. 特殊类型材质的恢复. Aittala 等^[51]引入稳定类型材质的假设, 并规定这些材质由材质表面随机重复分布的相似元素或结构组成, 这类材质也被称作自相似材质, 与通用材质相比, 稳定材质的恢复难度降低; 还提出采用包含闪光灯以及不含闪光灯拍摄的图像对来对材质进行参数恢复, 并在大量数据集上验证其有效性, 但对每个数据集的处理时间均需要 3 h. Hertzmann 等^[56]引入朝向一致性的假设, 认为在几何已知且同一材质的条件下, 2 个具有相同表面方向的点在图像中必会有相同或相似的外观, 该方法在镜面材质和天鹅绒材质上获得了可观的效果, 并支持对同一物体的不同材质属性进行分类.

c. 材质恢复与图像数量. 材质反射率与入射光方向和观察方向密切相关, 因此越提供多张不同光源和视角的输入材质照片, 越有利于从材质照片中分析并解耦出材质的反射率或参数图. Hui 等^[57]在 250 幅输入图像条件下构建了一个 BRDF 字典, 通过解决线性逆向问题, 根据给定法线恢复每个像素位置处的 BRDF, 但该方法产生的 SVBRDF 可能有噪声且输入图像必须要求点光源照明; Palma 等^[58]将输入图像数量扩展到了视频级别, 采用 Phong 模型估计材质的镜面分量, 重建近似的 SVBRDF, 但是在恢复质量上与真实世界物体仍存在差异, 且需要已知几何条件; 为了提升恢复质量, Xia 等^[59]首先在不同视角下

对物体(或旋转物体)拍摄一段视频作为输入, 然后根据随时间变化的辐射率迭代式地恢复表面法线分布, 并进一步求解反射率, 该方法不再假设几何形状已知, 并将视频拍摄条件扩展到了未知的光照条件下.

d. 运动先验. 还有的方法根据物体或相机的少量运动信息解耦物体形状和各向同性的 BRDF 反射率^[60], 并根据运动先验推理出对形状、反照率、BRDF 和方向光的联合估计.

上述传统方法受到硬件算力的限制, 以恢复 SVBRDF/BRDF 反射率为目标. 随着现代 GPU 的发展, 出现了基于神经网络推理的材质逆向重建方法, 该领域的研究方法逐渐从恢复 SVBRDF/BRDF 反射率过渡到恢复 SVBRDF/BRDF maps.

(2) 基于神经网络推理的材质逆向重建方法

与传统方法相比, 基于神经网络的方法适应能力更强, 支持自动特征提取和建模非线性关系, 且特别擅长处理大规模数据并通过正则化等方法避免过拟合. 这类方法一般选用不同类型的神经网络, 从大规模训练数据中学习材质反射率的分布, 并根据输入条件还原对应的材质反射率. 如根据基于卷积的神经网络对输入的二维图像进行编码, 并使用多层感知机(multilayer perceptron, MLP)从随机观测的样本数据中学习多个观测方向上的反射率特征. 对输入图像进行编码, 一方面是为了压缩数据, 减少存储和计算量; 另一方面是为了将像素的信息转移到具备高度表达能力的特征空间中, 并得到特征向量. 这个空间也被称作隐空间, 获取到的特征向量也叫隐向量. 隐向量通常作为材质的神经表示, 具有丰富的信息. 通过对隐空间的神经材质优化, 可以有效地在后续的解码步骤中获取准确的材质反射率. 此外, 一些方法中, 隐空间的神经材质还允许通过插值获得类别更丰富的新材质.

a. 直接估计方法. Kim 等^[61]提出一种轻量级的方法, 直接从输入的 RGB 图像中估计三维物体的表面反射率, 采用级联的 3 个网络分别估计形状、光照和 SVBRDF, 并在训练过程中使用级联网络的早期值作为后期网络的输入, 指导网络进行更好的预测; 还提供了第 1 个大规模合成数据集 SynBRDF, 确保每个场景的推理时间小于 90 ms; Deschaintre 等^[62]提出了另一种形式的网络设计方法, 在深度神经网络中引入一个与顺序无

关的融合层,从每张输入照片中提取关键信息,实现了任意光线和观察方向条件下的材质恢复。

由于物体的几何先验包含材质反射率中的法线信息,因此一部分方法假设几何条件已知^[63],并采用分阶段联合优化的方法,基于 CNN 分别对 SVBRDF maps 和几何形状进行学习,最终在移动设备上实现了材质捕获和估计。

另一类方法实现了在不校准相机和光照时提升恢复结果的鲁棒性. Ye 等^[64]使用神经网络从输入的视频帧序列中提取反射率参数图,采用运动线索对齐网络将编码后的中间神经反射特征在各个关键帧之间对齐,更好地解码反射率;该方法无需仔细校准相机和光照参数,但是仅限于恢复各向同性的表面反射率参数图. 还有一种方法以双分支网络的形式,通过门控递归单元^[65]融合不同图像中提取的特征,解决了不校准相机和光照时产生的材质恢复质量下降问题;该方法在材质逆向恢复过程中获得了干净和高质量的结果,同时仅需要输入多张照片。

b. 神经网络优化方法. 将空间变化的外观参数嵌入隐空间中进行优化,不仅可以支持任意数量的输入照片,还支持逆向恢复高分辨率(1 024×1 024)的材质参数图^[66],因此隐空间的设计和优化尤为重要. Deschaintre 等^[67]的方法基于已训练的深度网络对输入样本进行微调,并通过样本应用随机几何变换以及拼接等增强方法避

免网络过拟合;该方法可以实现将若干个样本的外观转移到目标图像上,并提取出近似的 SVBRDF 参数图,但是无法区分具有相似颜色和纹理的不同材质,同时难以处理包含强光的材质。

还有一类方法对更深层次的隐空间进行探索,其中的代表方法是 Guo 等^[68]提出的 MaterialGAN. 该方法基于 StyleGAN2^[69],通过一系列的卷积和上采样操作对输入图像进行编码得到隐向量 z ;然后通过全连接层的非线性映射网络将其转化为中间隐向量 w . 并将其通过可学习的仿射变换与控制风格向量的均值和方差相关联,引入 w 可以更好地避免材质属性之间的耦合;在此基础上,进一步将 w 复制若干次后,得到更深层次的隐向量 $w+$;经过优化后, $w+$ 的隐空间 $W+$ 成为比 w 的隐空间 W 更强大的扩展隐空间,并可以与高斯噪声的隐空间 N 结合,组成更具表达性的新的隐空间 $W+N$. MaterialGAN^[68]允许从 3~7 张闪光灯下拍摄的照片中恢复 SVBRDF 参数图. 在结果方面,当前的方法取得了超越前人的材质参数恢复结果,如图 16 所示;同时支持材质的高级语义编辑,如不同材质之间的插值等,但受限于网络全连接层的设计和生成对抗网络(generative adversarial network, GAN)模式崩塌的问题,该方法仅限于 256×256 分辨率下的材质恢复。

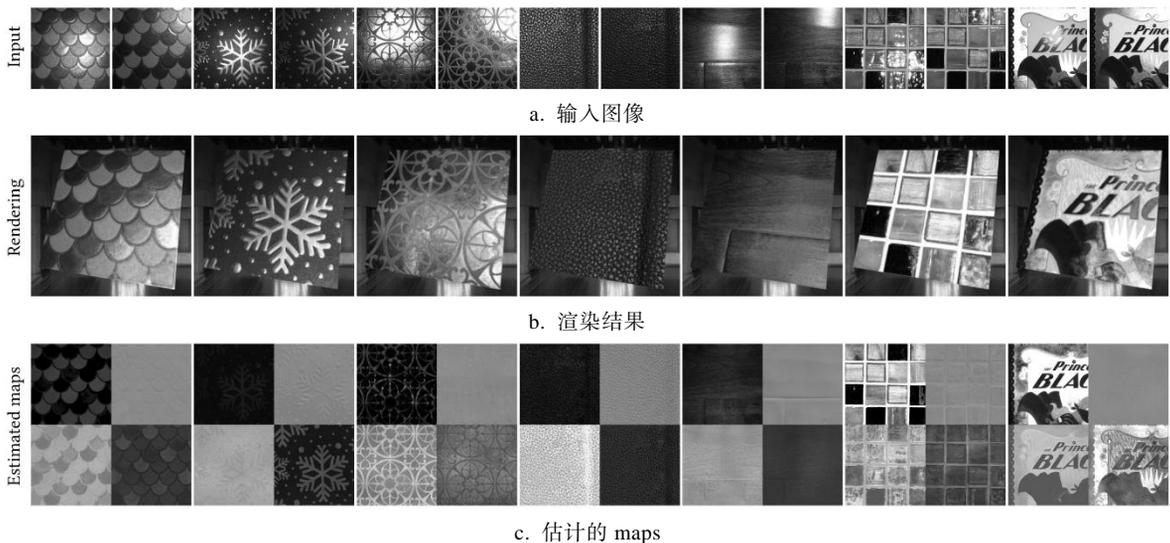


图 16 MaterialGAN 的多图输入材质逆向恢复结果^[68]

在神经网络方法中,训练网络时使用合理的损失函数非常关键. 现有的方法中通常使用

SVBRDF maps 损失(reflectance 损失)和 rendering 损失联合优化网络. 但是在训练过程中,对渲染

结果或任何一个参数图的错误估计都会被反向传播影响其他参数图. Saryazdi 等^[70]提出一种新型的损失函数设计策略, 解除了材质参数图之间以及材质参数图和高光之间过分耦合的问题, 在使用材质参数图通过可微渲染计算渲染结果时, 该损失函数仅对材质的其中一个参数图使用估计值进行可微计算, 对剩余几个参数图使用真实值; 该方法降低了不同视线和光线下采集的输入照片的数量, 与 Deschaintre 等^[62]的方法相比, 提高了重建的材质参数图的准确性.

5.2.2 单图输入

与多图输入的材质逆向重建相比, 单图输入的条件更加苛刻, 因此传统的方法难以通过有限的信息和先验恢复出合理的材质反射率. 现有的单图输入材质逆向恢复问题通常基于数据驱动的神经网络解决. 本文按照神经网络的类别, 分别对单图输入的逆向恢复方法进行归纳和总结. 3类神经网络分别为 CNN, GAN 以及扩散模型. 同时, 对标志性的方法列出相应的示例, 方便读者理解.

(1) 基于 CNN 的材质逆向重建方法

CNN 是深度学习模型中的一种, 独特之处在于其结构设计, 包括卷积层、池化层、激活函数层、全连接层等. 一方面, 能通过神经元的局部连接提取局部特征; 另一方面, 采用同一卷积核对图像操作, 确保权值共享, 并且通过 ReLU, Sigmoid 等激活函数为网络引入非线性, 提高模型的表达能力. CNN 具有自动特征提取、平移不变性、局部感知、层次化特征学习等优势, 目前已在图像识别、自然语言处理、语音识别、医学影像等多个领域取得了显著的成果. CNN 在单图输入的材质逆向重建领域同样具有典型的应用. 本文将单图输入材质逆向恢复任务的管线简单归纳为: 输入一张某种光照条件下拍摄的材质照片, 由 CNN 找到一组合适的 SVBRDF 材质参数图, 该参数图经过正向可微渲染, 产生与输入照片相似的逼真结果.

通过对材质观察, Aittala 等^[71]对自相似性材质(如皮革、地面、塑料、纸巾、木材等)进行研究, 由于这类材质结构有着整体随机但局部相似的视觉效果, 因此他们大胆假设这类材质对应的 SVBRDF 参数图也具有自相似性(具有相同的局部均值和方差). 该方法采用深度 CNN 估计出一

组小分辨率的 SVBRDF 参数图块, 用于表示未知的材质; 然后将该组参数图块在不同观测角度和光照配置下的渲染结果, 与输入图像中的不同位置处的图像块进行比较, 并在傅里叶域中通过统计图像之间的差异驱动优化 SVBRDF 参数图. 本质上, 该方法的优化过程基于神经网络的纹理统计而非逐像素比较, 因此可以成功地恢复自相似材质的 SVBRDF 参数图; 但是, 对于具有高度规则结构的材质(如砖墙等), 该方法无法恢复其精准排列的结构. 此外, 一些具有强各向异性的材质(如布料)也超出了该方法的处理范围.

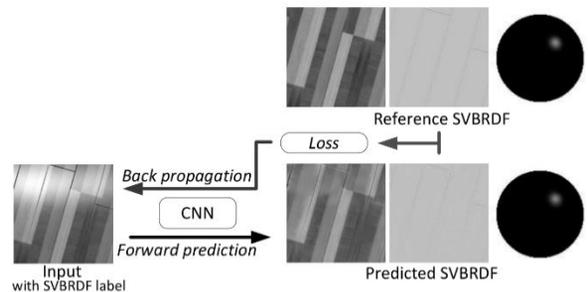
同样是针对自相似材质, Henzler 等^[72]提出一种完全自监督的训练方法, 利用卷积编码器将所有的闪光灯图像嵌入隐空间中, 并使用 CNN 将随机空间场转换为 SVBRDF 参数图的场; 然后根据输入图像的视觉特征匹配对应的 SVBRDF 参数图. 该方法支持生成多样且无限大的材质参数图, 并允许插值且无需合成数据, 但是仅支持恢复自相似的材料, 因此仍然无法避免 Aittala 等^[71]的方法面临的问题.

虽然自相似性的假设成功地简化了材质逆向问题, 但是也限制了材质的全局结构和类别, 因此大部分研究者不再继续使用这一假设, 而是转向在训练策略和网络模块设计上寻求突破. 为了减少所需的标注训练数据量, Li 等^[73], Ye 等^[74]和孙彬^[75]提出自增强的策略, 将大量未标注样本嵌入神经网络来配合训练, 从而在一定程度上避免了繁重的数据采集工作. 本文以 Li 等^[73]提出的方法为例, 简述具体的训练过程, 如图 17 所示. 步骤如下:

Step1. 使用 SVBRDF maps 和相应的渲染图像, 以有监督的形式在 CNN 上进行训练.

Step2. 对初步训练的网络输入真实照片得到估计的 SVBRDF maps, 并采用可微渲染得到新视角下渲染的照片.

Step3. 用估计的 SVBRDF maps 和渲染图作为新的训练数据集, 再次微调 CNN.



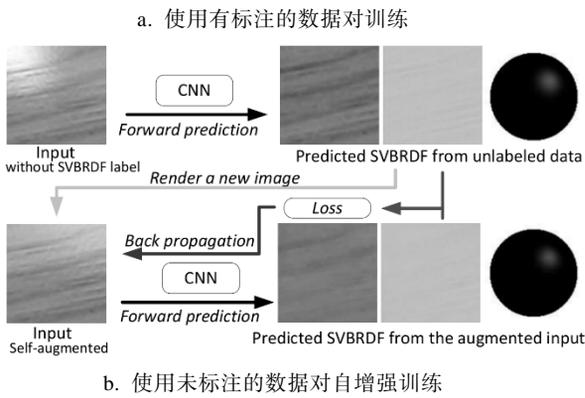


图 17 自增强 CNN 训练过程示意^[73]

这种方法在一定程度上缓解了标注数据集(材质的 SVBRDF maps 和相应的渲染图)不足的问题, 并支持任意自然光照条件下的材质参数估计. 但是本文观察到, 由于中间网络预测的 SVBRDF maps 与真实值相比仍有一部分误差, 因此经过大量微调后反而会丧失数据的真实感; 此外, 该方法采用各向同性的镜面反射反照率和粗糙度参数, 仅支持还原各向同性的高光.

为了彻底解决训练数据紧缺的问题, Deschaintre 等^[76]整理了由艺术家们创建的过程式材质数据集, 并通过旋转、裁剪、混合等增强方法进一步扩大数据集的规模, 最终构建了包含约 20 万组 SVBRDF maps 和相应渲染图(分辨率为

256×256)的材质数据集, 为后续大量的研究方法提供了数据保障. 在大规模数据驱动下, 训练深度 CNN(U-Net^[77])估计材质参数图仍面临全局特征损失的问题, 这是因为 U-Net 采用沙漏形状的多尺度设计, 尽管可以在长距离的像素区域内保证合理的空间关系, 但是在实际下采样和上采样过程中缺少保留长距离特征细节的手段, 所以仍会损失部分全局特征细节. Deschaintre 等^[76]在每个 U-Net 卷积层的基础上, 额外添加了用于从像素中提取全局特征的网络模块, 并在训练时将提取的全局特征细节注入回 U-Net 中, 同时引入渲染损失评估再现的材质外观和真实值之间的误差. Deschaintre 等^[76]提出的网络结构如图 18 所示, 其中, 黑色图块表示原始的 U-Net, 绿色图块表示额外添加的全局特征提取模块. 该方法适用于闪光灯照片输入, 其结果优于 Aittalta 等^[71]和 Li 等^[73]的工作, 但仅支持 256×256 分辨率的图像输入, 提高图像分辨率会增加内存占用并阻碍网络收敛. 因此, 该方法仍有改进空间, 如先用小分辨率图像训练网络, 再在大分辨率图像上进一步微调或额外增加一个预训练的超分网络提升图像分辨率.

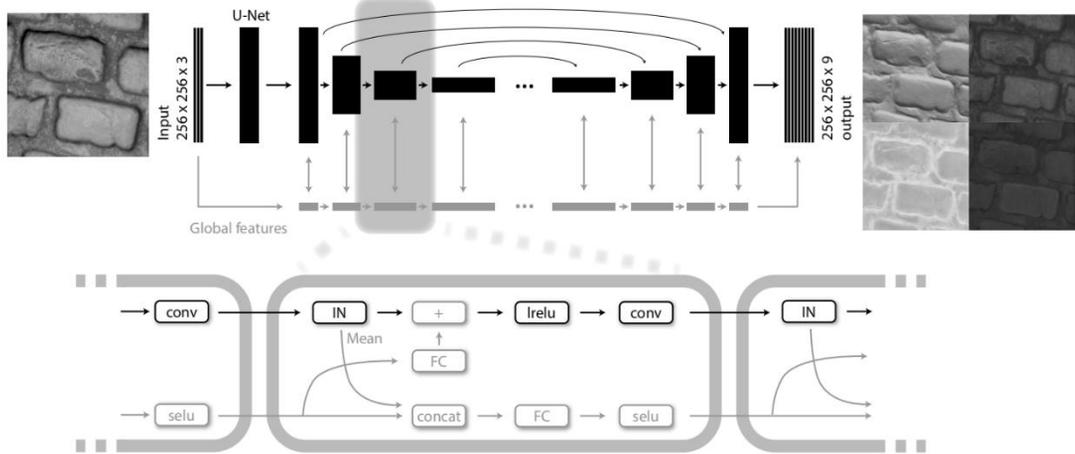


图 18 Deschaintre 等^[76]提出的网络结构示意图

上述方法中, U-Net 包含一个编码器和一个解码器, 其中, 编码器用于将输入的闪光灯照片压缩到隐空间得到特征向量, 解码器用于将特征向量恢复为 SVBRDF maps. 由于单个材质的 SVBRDF maps 在分布上并不完全相关, 因此一些方法采用多个解码器解码 SVBRDF maps. Li 等^[78]对每个 SVBRDF 参数图提供一个解码器, 额外添加了一个与卷积网络并行训练的材质分类器

用于判定材质的类别, 并对不同类别的材质激活不同的网络通道进行训练; 此外, 针对解码的 SVBRDF 参数图的不同种类提供不同的优化策略, 以法线图为例, 对图中的法线按照与 z 轴的夹角不同进行分组, 并提供了不同的训练权重, 期望在训练过程中能确保法线图避免过度平滑. Li 等^[78]方法的核心思想是, 通过区分不同材质类别、同一材质不同参数图的类别, 使用不同的网络通

道和训练策略进行学习, 减轻共用整个网络的负担, 因此该方法在成功分类时能够很好地解耦出材质的不同参数图, 但是如果推理过程中分类失败, 会严重影响到后续的材质参数估计。

除了上述按照类别拆分网络通道的方法外, 还有的方法对材质本身进行拆分和重表示。Wang 等^[79]将材质统一用一组基材质和权重的线性混合表示, 其中, 基材质进一步拆分为各像素数值一致的全局基和各像素数值各不相同的局部基。该方法使用神经网络预测材质的全局基、局部基和权重, 并通过渲染损失进行优化; 在训练过程中, 额外引入变化-一致性损失避免全局基、局部基和权重的耦合; 其关键是通过拆分材质的方式, 将材质中低频的信息和高频的信息分开学习, 在一定程度上避免材质参数图之间的过分耦合。该方法面临的主要问题是, 无法区分特殊材质中的纹理、高频阴影和光照, 最终导致解耦失败, 如印有三维图案的塑料包装表面等。

对于部分闪光灯照明的材质, 标准卷积难以从高光覆盖的饱和像素中学习有效的特征^[76,78]。为此, Guo 等^[80]提出高光感知 (highlight-aware, HA) 的卷积, 旨在从不饱和的内容中提取出有意义的特征并且抑制强高光污染的影响; 网络的编码器部分采用双流的架构, 由标准卷积学习材质参数的基本信息, HA 卷积自动地从输入图像中学习一个与过曝区域相关的软掩模, 用于缓解高光带来的影响; 2 个不同的卷积层通过注意力机制进一步融合。该方法采用一种通用的训练策略, 先在大规模数据集上训练材质估计网络, 再在小规模测试图像上优化网络参数来进一步提升网络性能。但是, 在处理包含大范围饱和和高光的输入图像时, HA 卷积的主要问题是会恢复出偏暗的漫反射结果, 同时重建出来的法线贴图在过曝区域会缺少细节, 意味着该方法可以用来缓解较为明显的高光带来的影响, 对于非常强烈的高光仍无法处理, 这也是当前所有单图输入的方法面临的重要难题。

除了对网络结构进行重新设计外, 还有一类方法借鉴其他领域的训练策略。Zhou 等^[81]受到元学习^[82]方法的启发, 在训练过程(采用 rendering 损失和 reflectance 损失)中加入测试时间优化(采用 rendering 损失)来训练网络, 并在测试时间优化步骤中, 提供一个额外的辅助网络估计单幅图

像的反射率参数图, 并以此作为参考值计算材质参数图的损失(伪 reflectance 损失)。该方法的本质是先通过优化手段对网络权重进行初始化, 为了避免过拟合到输入图像, 再进行标准的网络训练, 引导恢复的材质反射率逼近真实值, 因此比直接估计再优化的方法能产生更好的结果, 同时在速度上快 2 个数量级。

与上述采用前馈神经网络从训练数据分布中学习先验预测 SVBRDF maps 的方法不同, Luo 等^[83]使用 RNN 迭代更新给定重建似然梯度下的反射率参数, 结合学习到的先验和似然度量提供 SVBRDF 参数图的最大后验估计。本质上, 该方法学习的是预测材质反射率图的梯度, 而非是直接预测材质反射率图, 因此在实践中, 这种优化方案支持产生比前馈方法更准确的结果; 但是同样地, 该方法无法处理包含强烈高光的输入图片。

还有一些方法将卷积网络逐渐从低分辨率输出扩展到超高分辨率输出。针对自然光照条件下捕获的照片输入, Martin 等^[84]通过训练 2 个级联的 U-Net 推理空间变化的反射率参数图, 并在梯度域中插值来恢复高分辨率输出 (2 048×2 048); 由于设计的限制, 该方法不适用于从金属物体上提取 SVBRDF 参数图。同样是支持高分辨率输出, Guo 等^[85]采用端到端的框架恢复高质量超高分辨率 (4 096×4 096) 的 SVBRDF 材质参数图, 在隐式神经反射模型的基础上, 采用分治的解决方案和一个全局特征提取器, 解决卷积网络难以支持高分辨率输入图像训练的问题, 该方法获得了与输入图像匹配的局部空间细节和一致的全局结构。

(2) 基于 GAN 的材质逆向重建方法

GAN 是一种深度学习模型, 由负责生成新数据样本的生成器和负责判断输入数据属于真实值或生成值的判别器组成。在训练过程中, 这 2 个网络不断地进行对抗, 生成器逐渐学习生成更加逼真的数据, 而判别器也逐渐提高其判断的能力。与 CNN 相比, GAN 具有更强大的生成能力, 可以生成高质量、多样的数据, 如人脸图像、艺术作品、音乐等, 在图像生成、艺术创作、数据增强等领域展现了巨大的潜力。然而, GAN 的训练过程比较复杂, 存在不稳定等问题。下文介绍基于 GAN 的材质逆向重建方法。

Zhao 等^[86]提出一个自监督的 GAN, 仅使用

真实数据(材质的闪光灯照片)作为输入训练网络,并从单幅图像中捕获 SVBRDF 参数图. 考虑到相同材质的不同参数图之间的关联性,该方法提供了一个双流的生成器,分 2 组输出 SVBRDF 参数图的预测结果(其中,法线和粗糙度一组,漫反射和镜面反射一组);通过正则化计算获取猜测的漫反射参数图并用于计算漫反射损失,提升材质参数预测的准确程度. 该方法支持将具有重复结构的低分辨率输入图像,生成超高分辨率(4 096×4 096)的高质量 SVBRDF 参数图,但对于每个输入均需要约 3 h 的时间代价. 随后, Wen 等^[87]对 Zhao 等^[86]的网络框架进行改进,在微调 GAN 结构的同时,针对自相似材质的稳定性特征引入傅里叶损失,在频率域内约束 SVBRDF 参数图的特征,使得学习到的分布更加均匀,缓解了不均匀高光对材质逆向恢复的影响;此外,引入两阶段的训练策略减少了原网络的训练时间;但是,这种方法同样继承了自相似材质带来的缺点. 上述 2 种方法猜测的漫反射参数图并未对强高光输入保持鲁棒性.

针对更为广泛的材质类别, Zhou 等^[88]提出一种基于 GAN 的单幅图像 SVBRDF 参数图估计方法,在合成数据集和真实数据集上混合训练,并使用一组全连接层从编码器提取的特征中估计输入光的位置;每个 SVBRDF 参数图分别由各自独立的解码器和鉴别器处理. 由于采用了对抗性损失训练网络,该方法在某些情况下产生了输入图像中不存在的细节和结构. 同样是在合成数据和真实数据集上进行混合训练, Vecchio 等^[89]提出一种基于分块的 GAN SurfacerNet,并根据单幅图像恢复高质量高分辨率(2 048×2 048)的材质参数图.

与上述生成方法相比,一些研究者通过引入辅助条件更好地恢复材质参数. 如 Zhou 等^[90]提出一种特定类别 SVBRDF 参数图生成模型 TileGen,支持根据模式条件图片生成支持平铺的新材质,或者通过优化得到与目标照片匹配的材质,同时支持材质编辑,其结果图如图 19 所示. 可以看出,该方法具备一定去除高光的能力,但是需要对每类材质准备相应的数据集并训练网络,且仅支持生成 512×512 分辨率的材质参数图.

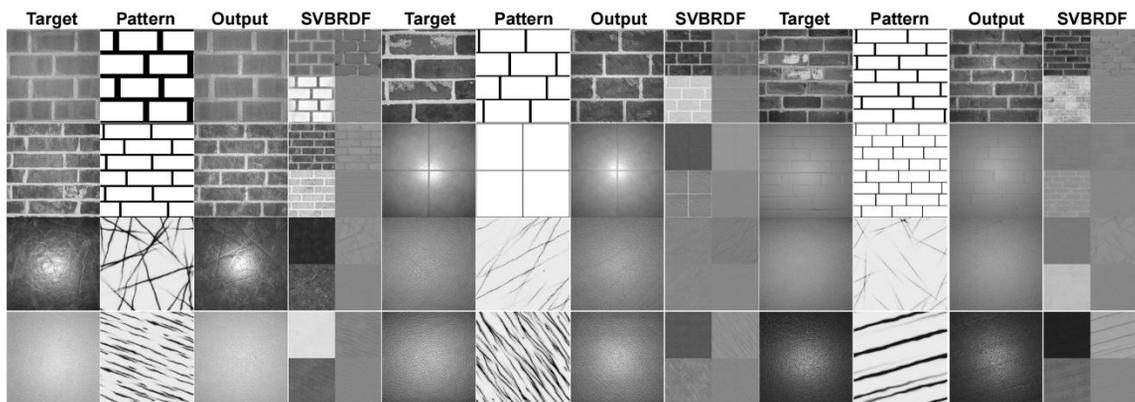


图 19 TileGen 在目标输入和条件图案指导下恢复的材质参数图^[90]

还有一类方法将注意力机制与 GAN 结合起来提高重建质量. 以单张漫反射照明下的图像作为输入, Rodriguez-Pardo 等^[91]采用具有注意力模块的 GAN 提高全局一致性,并提供一种通用的不确定性量化框架用于材质捕获;该方法使用消费级平板扫描仪提供漫反射照明,且目前仅在布料材质上进行了逆向重建的应用. 同样是在 GAN 中嵌入注意力模块, Shi 等^[92]提出一个新的端到端的 GAN,与已有的对每个材质参数使用单独的鉴别器的方法不同,该方法使用一个统一的框架简化网络结构,并采用注意力模块引导网络关注材质参数图中的细节;但与大多数方法一样,

该方法无法处理输入图像中的强烈高光对材质逆向恢复的影响.

针对高光问题,刘克梅^[93]提出一种基于高光消除的单张图片 SVBRDF 参数图重建方法,其中设计了一个基于密集特征融合连接和高光多级识别的高光消除模块,在恢复 SVBRDF 参数图的过程中减少强光区域对材质重建的影响,在一定程度上消除了高光伪影现象.

(3) 基于扩散模型的材质逆向重建方法

近年来,扩散模型成为热门的生成模型,它类似于自然界烟雾等粒子的扩散过程,通过逐渐向数据中添加噪声,然后学习噪声的添加过程从

噪声中恢复出原始数据. 扩散模型具有超越 GAN 的训练稳定性, 并且生成的样本更具多样性, 在图像生成、图像修复、超分辨率、文本生成、音频生成等领域有广泛的应用; 该模型的缺点是计算量较大、训练时间长. 目前, 已有一些方法, 如 LoRA^[94], ControlNet^[95]等, 可以在牺牲部分质量的条件快速微调模型, 得到了不错的收敛结果.

标准的扩散模型支持通过预训练的 CLIP 模型, 将文本或者图像等条件注入扩散模型的交叉注意力模块中, 实现有条件的生成. 但是, 这种方式生成的内容无法做到条件图和生成图在像素上保持一致, 本质上是一种风格的转移. 对于 SVBRDF 材质图逆向重建任务, 希望输入的材质照片和生成的材质参数图之间在像素上有强相关性. 因此, Vecchio 等^[96]用 CLIP 模型将文本和图像作为全局风格条件输入到扩散模型中, 并使用 ControlNet 方法为扩散模型输入相同的图像条件补充局部细节, 最终得到逐像素匹配的 SVBRDF 参数图重建结果. 该方法可以根据未知照明的图像, 输入生成合理、可平铺且高分辨的材质参数图; 以隐式扩散模型^[97]为基础并修改生成过程,

引入噪声翻滚、基于分块的多尺度扩散、基于分块的解码和边界重绘等策略, 支持逆向重建可拼接且任意分辨率的材质参数图, 同时保持了良好的质量.

为了实现更加精简、有效的条件引导, 同样受到 ControlNet 的启发, Sartor 等^[98]提出一个在不同光照条件输入的情况下逆向重建 SVBRDF 参数图的生成式扩散模型 MatFusion, 先在大规模合成 SVBRDF 数据集上训练一个无条件的材质图生成模型; 然后仅修改预训练好扩散模型的输入层通道, 将额外的条件图输入; 最后在不同的光照条件输入下对 SVBRDF 扩散模型进行微调. 该方法提供了 3 种输入条件: 仅有闪光灯照明的单幅图像、在不受控自然光照明下拍摄的单幅图像、有无闪光灯照明的图像对, 它们对应的条件和重建结果图如图 20 所示; 在推理过程中, 通过改变随机种子生成供用户选择的不同材质参数图结果. 与 ControlNet 需要提供额外的网络模块接收条件图像输入不同, 该方法仅对输入头进行修改, 在实现上更简单同时保证了更低的开销, 其代价是在微调网络时一定程度上会污染原始的扩散网络.

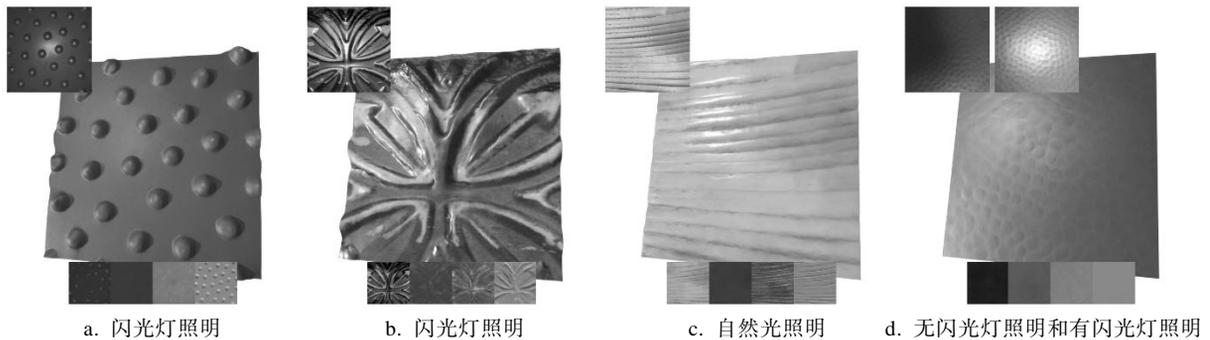


图 20 MatFusion 方法在不同光照条件下恢复的材质参数图^[98]

5.2.3 对比和分析

基于轻量级设备的材质重建方法研究如何从少量特定照明图片中恢复细节丰富、分辨率高且多样性强的材质参数, 为高度真实感渲染提供高质量且丰富的材质数据支撑. 本文将以前述输入的传统材质逆向重建方法和基于深度学习的材质逆向重建方法进行对比, 结果如表 3 和表 4 所示; 表 5 和表 6 所示为单幅图像输入的材质重建方法在多个指标下的表现情况. 其中, 表 3 和表 5 中主要关注重建方法的输入条件和网络特

点; 表 4 和表 6 中主要关注重建材质的特征和该方法的的重建能力; 有查询空间意味着可以对数据进行插值; 表 5 中, 合成数据和真实数据指训练时数据集的构成; 由于所有的统计方法使用的硬件设备不完全统一, 因此以原文献中提到的设备为准; 请注意区分 SVBRDF/BRDF 和 SVBRDF/BRDF maps, 前者是反射率, 后者是反射率的多张参数图. 此外, 由于本文讨论的所有单图输入的材质逆向重建方法的输出均为各向同性 SVBRDF maps, 因此表 6 中省略了该项.

表 3 多图输入的材质重建方法的输入条件和网络特点结果对比

方法类别	方法	输入数量	光照条件	几何条件	网络/方法类型	训练/计算时间	推理性能
------	----	------	------	------	---------	---------	------

传统	Aittala 等 ^[51]	2 张	有闪光灯和无闪光灯	无	参数优化	每个数据集 3 h
传统	Ghosh 等 ^[53]	多张	偏振光	无	参数优化	
传统	Riviere 等 ^[55]	视频	有闪光灯和无闪光灯或 LCD 面板光照	无	参数优化	
传统	Hertzmann 等 ^[56]	多张	未知环境光照	已知	参数优化	每个材质 5 h
传统	Hui 等 ^[57]	250 张	点光源	无	参数优化	
传统	Palma 等 ^[58]	视频	环境光照	已知	参数优化	数 min~数 h
传统	Xia 等 ^[59]	视频	未知静态光照	无	参数优化	10 h
传统	Chandraker 等 ^[60]	多张	方向光	无	参数优化	
深度学习	Kim 等 ^[61]	多张	有闪光灯和无闪光灯, 以及物体前景遮罩图	无	CNN	<90 ms
深度学习	Deschaintre 等 ^[62]	多张	闪光灯且无需校准	无	CNN	7 d
深度学习	Boss 等 ^[63]	2 张	有闪光灯和无闪光灯	无	CNN	700 ms
深度学习	Ye 等 ^[64]	视频	闪光灯且无需校准	无	CNN	
深度学习	Zhu 等 ^[65]	多张	闪光灯且无需校准	无	CNN, RNN	5 d
深度学习	Gao 等 ^[66]	多张	闪光灯	无	CNN	16 h
深度学习	Deschaintre 等 ^[67]	多张	任意环境光照和闪光灯	无	CNN	8 d
深度学习	Guo 等 ^[68]	3~7 张	闪光灯	无	StyleGAN2	5 d
深度学习	Saryazdi 等 ^[70]	多张	闪光灯	无	CNN	3 d
传统	Dong 等 ^[99]	视频	未知环境光照	已知	参数优化	8~10 h
传统	Zhou 等 ^[100]	多张	环境光照	已知	参数优化	20~380 min
传统	Hui 等 ^[101]	多张	闪光灯	无	参数优化	

表 4 多图输入的材质重建方法的重建材质特征及重建能力结果对比

方法类别	方法	恢复类别	分辨率	材质性质	多样性	查询空间	处理高光的能力	编辑能力
传统	Aittala 等 ^[51]	各向异性 SVBRDF maps		自相似	○○○○○	无	○○○○	无
传统	Ghosh 等 ^[53]	各向异性 SVBRDF maps		任意	○○○○	有	○○○○○	有
传统	Riviere 等 ^[55]	各向同性 SVBRDF maps		任意	○○○○○	无	○○○○○	无
传统	Hertzmann 等 ^[56]	各向同性 SVBRDF		任意	○	无	○○	无
传统	Hui 等 ^[57]	各向同性 SVBRDF		任意	○○○	有	○○○	有
传统	Palma 等 ^[58]	各向同性 SVBRDF		任意	○	无	○○	无
传统	Xia 等 ^[59]	各向同性 SVBRDF maps		任意	○○○○○	无	○○○○○	无
传统	Chandraker 等 ^[60]	各向同性 BRDF		任意	○	无	○○○○	无
深度学习	Kim 等 ^[61]	各向同性 BRDF		任意	○	无	○○	无
深度学习	Deschaintre 等 ^[62]	各向同性 SVBRDF maps	256×256	任意	○○○○○	无	○○○○○	无
深度学习	Boss 等 ^[63]	各向同性 SVBRDF maps	256×256	任意	○○○	无	○	无
深度学习	Ye 等 ^[64]	各向同性 SVBRDF maps	512×512	任意	○○○○○	无	○○○○	无
深度学习	Zhu 等 ^[65]	各向同性 SVBRDF maps	256×256	任意	○○○○○	无	○○○○○	无
深度学习	Gao 等 ^[66]	各向同性 SVBRDF maps	1 024×1 024	任意	○○○○○	无	○○○	无
深度学习	Deschaintre 等 ^[67]	各向同性 SVBRDF maps	可平铺至无限大	任意	○○○○○	无	○	无
深度学习	Guo 等 ^[68]	各向同性 SVBRDF maps	256×256	任意	○○○○○	有	○○○○○	有
深度学习	Saryazdi 等 ^[70]	各向同性 SVBRDF maps	256×256	任意	○○○○○	无	○○○○	无
传统	Dong 等 ^[99]	各向同性 SVBRDF	1 024×1 024	任意	○○○○	无	○○○○○	无
传统	Zhou 等 ^[100]	各向同性 SVBRDF		任意	○○○○	有	○○○	有
传统	Hui 等 ^[101]	各向同性 SVBRDF		任意	○○○	有	○○○	有

表 5 单图输入的材质重建方法的输入条件和网络特点结果对比

方法	合成数据	真实数据	光照条件	网络类型	训练时间	推理性能
Aittala 等 ^[71]	√		闪光灯	CNN	每个数据集 2 h	
Henzler 等 ^[72]		√	闪光灯	CNN	4 d	5 min
Li 等 ^[73]	√		未知环境光照	自增强 CNN	38 h	0.3 s
Ye 等 ^[74]	√		未知环境光照	自增强 CNN		

孙彬 ^[75]	√		未知环境光照	自增强 CNN		
Deschaintre 等 ^[76]	√		闪光灯	CNN	7 d	
Li 等 ^[78]	√		闪光灯	CNN		
Wang 等 ^[79]	√		闪光灯且无需校准	CNN	12 h	1~3 s
Guo 等 ^[80]	√		闪光灯	CNN	4 d	0.17 s
Zhou 等 ^[81]	√		闪光灯	CNN	5 d	
Luo 等 ^[83]	√		闪光灯	RNN	21 d	0.1 s
Martin 等 ^[84]	√		环境光照	CNN	5 d	4 s
Guo 等 ^[85]	√		闪光灯	CNN	60 h	10 s
Zhao 等 ^[86]		√	闪光灯	GAN	3 h	
Wen 等 ^[87]		√	闪光灯	GAN	2.5 h	
Zhou 等 ^[88]	√	√	闪光灯	GAN	3 d	
Vecchio 等 ^[89]	√	√	环境光照	GAN		
Zhou 等 ^[90]	√		闪光灯	StyleGAN2	5 d	2 min
Rodriguez-Pardo 等 ^[91]	√		漫射光照	GAN		0.131 s
Shi 等 ^[92]	√		任意未知光照	GAN	7 d	
刘克梅 ^[93]	√		闪光灯	GAN	108 h	0.13 s
Vecchio 等 ^[96]	√		环境光照	扩散模型		
Sartor 等 ^[98]	√		闪光灯或环境光照或 有闪光灯和无闪光灯	扩散模型	357 h	
Li 等 ^[102]	√		室内环境光照	CNN		

表 6 单图输入的材质重建方法的重建材质特征及重建能力结果对比

方法	分辨率	材质性质	多样性	查询空间	处理高光的能力	编辑能力
Aittala 等 ^[71]	256×256	自相似	○○○○	无	○○○	无
Henzler 等 ^[72]	任意大小	自相似	○○○○	有	○○○	有
Li 等 ^[73]	256×256	任意	○○○	无	○	无
Ye 等 ^[74]	256×256	任意	○○○	无	○	无
孙彬 ^[75]	256×256	任意	○○○	无	○	无
Deschaintre 等 ^[76]	256×256	任意	○○○○○	无	○	无
Li 等 ^[78]	256×256	任意	○○○○	无	○	无
Wang 等 ^[79]		任意	○○○○○	无	○	无
Guo 等 ^[80]	1 024×1 024	任意	○○○○○	无	○○○	无
Zhou 等 ^[81]		任意	○○○○○	无	○○○	无
Luo 等 ^[83]		任意	○○○○○	无	○	无
Martin 等 ^[84]	4 096×4 096	任意	○○○○○	无	○	无
Guo 等 ^[85]	4 096×4 096	任意	○○○○○	无	○	无
Zhao 等 ^[86]	4 096×4 096	任意	○○○○	无	○	无
Wen 等 ^[87]	1 024×1 024	自相似	○○○○	无	○	无
Zhou 等 ^[88]	256×256	任意	○○○○○	无	○○	无
Vecchio 等 ^[89]	2 048×2 048	任意	○○○○○	无	○○	无
Zhou 等 ^[90]	512×512	任意	○○○○○	有	○○○○○	有
Rodriguez-Pardo 等 ^[91]	512×512	任意	○○○	无	○	无
Shi 等 ^[92]	256×256	任意	○○○○○	无	○○	无
刘克梅 ^[93]	256×256	任意	○○○○○	无	○○	无
Vecchio 等 ^[96]	768×768	任意	○○○○○	有	○	有
Sartor 等 ^[98]	256×256	任意	○○○○○	无	○○	无
Li 等 ^[102]		任意	○○○○○	有	○	有

在多图输入的材质重建方法中, 传统的材质

逆向重建方法以直接估计和参数优化为主, 同时

假定了大量的先验,如几何形状已知、光照条件已知等.一些方法利用采集视频中物体运动过程中的变化对材质反射率进行提取和分类,其中大部分仅支持各向同性的 SVBRDF/BRDF 反射率恢复,不支持参数图的还原,且处理材质的多样性有限.基于深度学习的多图输入材质逆向重建方法充分利用大规模数据集的丰富先验,支持恢复更多类别的材质属性参数图,并且通过生成可平铺的材质参数图,或者优化材质隐空间的特征向量来提供高分辨率的支持.还有一些方法允许接收任意数量的输入图像,因此可以获得不同光照、视角下的材质先验,在训练过程中扩充网络的感受野从而解码出更高质量的材质参数图,并且消除高光对材质参数图的影响.

在单图输入的材质重建方法中,整体逐渐呈现由深度 CNN 到 GAN,再到扩散模型的发展趋势;此外,深度 CNN 通过引入自增强策略进一步减少了输入数据集的需求.与多图输入的重建方法相比,单图输入的有效信息更少,且大部分方法以闪光照片作为输入,对包含强烈高光的材质输入,极易导致材质参数图之间以及材质参数图和光照之间出现强耦合,从而在恢复出来材质参数图中包含高光伪影的错误结果.为此,研究人员将 HA 模块、高光消除模块、注意力机制等引入深度网络中处理高光的影响.但是,已有的方法对于包含极强高光的材质输入仍无法做到彻底解耦,原因是高光破坏了材质的背景纹理信息,仅从单幅图像输入很难恢复被饱和和高光覆盖的纹理,因此该问题仍是当前面临的一个重要的挑战.

5.3 材质生成

数据驱动式生成方法以生成网络为核心,如 GAN、扩散模型等.本文按照网络训练过程中是否有 SVBRDF 参数图作为 Ground Truth 监督,将生成式方法分为有监督的材质生成方法和自监督(无监督)的材质生成方法,下面介绍各部分的详细内容及优缺点.

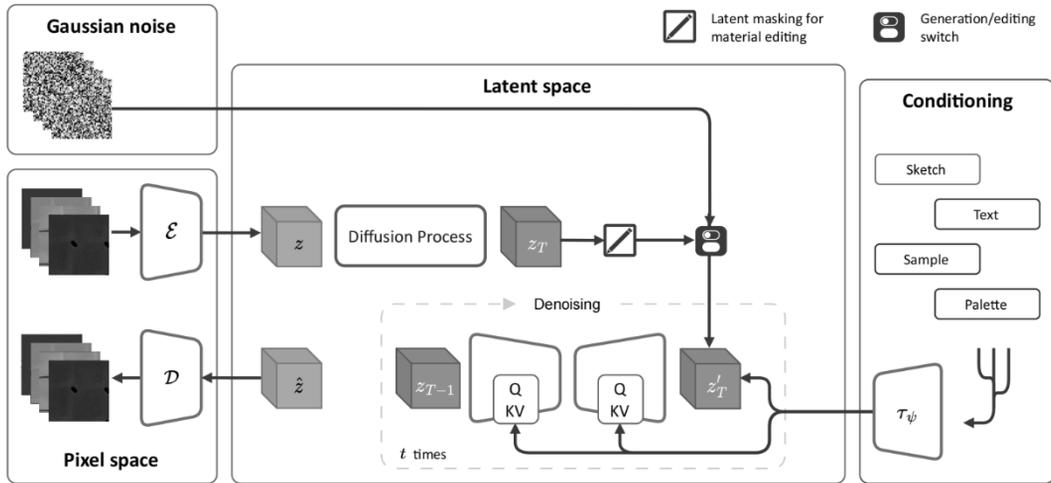
5.3.1 基于有监督训练的生成式方法

有监督式生成方法的显著特征是结合生成式神经网络,以数据驱动的方式充分利用更全面且灵活的先验信息.这类方法以充分学习合成数据集中 SVBRDF maps 的分布(reflectance 损失)为

目标,并尽可能地让估计的参数图渲染出的材质结果逼近用数据集中参数图渲染出的结果(rendering 损失).

一些方法通过引入额外的条件实现生成过程中的编辑和控制. Hu 等^[103]以 MaterialGAN^[68]作为先验,通过 rendering 损失在 MaterialGAN^[68]隐空间中优化材质的外观,并将目标材质中微观尺度和中观尺度的纹理细节迁移到指定的材质上生成新材料;引入标签图和重采样的 Sliced Wasserstein 损失^[103]规范转移的材质范围,消除目标材质中不明确的歧义信息,但受到 StyleGAN2^[69]生成结果的分辨率限制,无法生成超高分辨率的材质参数图;该方法还对条件引导图像的光照条件有要求,仅支持闪光灯或太阳光照明,不支持更复杂的环境光照条件.同样是基于 GAN, Li 等^[104]提出一个多投影 GAN,将生成器生成的三维形状映射为多个不同的二维投影,并通过多个鉴别器判定二维投影是否属于同一个几何指导网络学习并生成新的特定类别的三维形状;该方法适用于从低维训练数据集学习其他高维分布,并被推广到 SVBRDF 参数图的生成任务中.

作为一种生成式模型,扩散模型从一个完全随机的噪声图像出发,通过逐步去噪的过程逐渐生成高质量的图像;此外,通过引入注意力机制,结合 CLIP 模型,扩散模型可以在无条件、文本条件、图像条件等多种形式下,生成多样的、可控的、高质量的结果.在材质参数图生成领域,基于扩散模型的材质生成方法同样被广泛研究. Vecchio 等^[105]提出一种可控的材质参数图生成方法 MatFuse,并支持材质编辑,该方法的管线中集成多种输入条件实现细粒度的生成控制,包括调色板、草图、文本和图片;采用 VQ-GAN^[106]架构,使用 4 种不同的编码器学习材质参数图特定的隐空间先验,并在训练过程中添加 rendering 损失提高恢复性能.同样是受到生成模型的限制,MatFuse 在恢复高分辨率材质时内存需求急剧增加,如处理 512×512 分辨率材质时需要约 18 GB 显存,而处理 768×768 分辨率材质时则需要约 24 GB 显存,严重地限制了高分辨率材质的生成. MatFuse 的结构如图 21 所示.

图 21 MatFuse 的结构^[104]

为了获取更高的生成分辨率, Xin 等^[107]提出一个基于扩散模型的生成框架 DreamPBR, 在生成模型的所有卷积层中开启循环填充的设置, 保证输出图像在保留原有图案的同时支持无缝的图案平铺; 还引入材质超分辨率模块^[108], 将 512×512 的原始输出结果映射到 2048×2048 的高分辨率结果. 该方法中加入一个多模态引导模块, 包括像素对齐引导、样式图像引导和三维形状引导, 增强对扩散模型的控制, 创建特定的 SVBRDF 参数图; 还加入一个 HA 的反照率解码器, 去除生成的 RGB 图像中隐含的高光. 使用 DreamPBR 可以更加自由地创建材质, 并获取对应的高分辨率材质参数图.

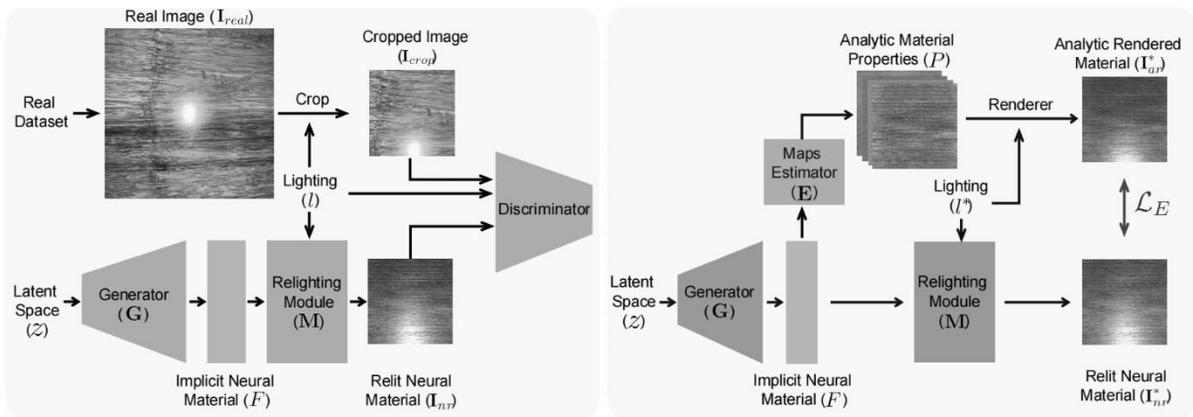
另一类方法则基于神经网络表达的表面材质模型生成新的表面材质. Thies 等^[109]从噪声或者不完整的几何表面出发, 通过高维特征图学习包含更多信息的神经纹理, 并采用延迟神经渲染管线进行解释渲染, 最终生成具有真实感的渲染结果. 该方法支持新视角合成、场景材质编辑、人像动画合成等功能, 且仅需几毫秒便可以实现高分辨率的输出; 但是, 该方法严重依赖于预处理步骤中的几何代理, 如果物体几何太粗糙, 生成的结果更加模糊. Zeltner 等^[110]将神经纹理扩展到多层, 将学习到多层的材质图整合为紧凑且统一的神经表示; 然后输入不同的神经解码器产生反射率值和重要性采样的方向, 并在解码器中引入中尺度先验以及微表面采样分布. 该方法支持各向异性采样, 并结合硬件加速实时生成细粒度的渲染结果, 如图 22 所示; 但是, 该方法不符合能量守恒, 且网络训练过程极易受到超参数的影

响.

图 22 Zeltner 等^[110]神经渲染结果

5.3.2 基于自监督训练的生成式方法

尽管有监督的生成式方法能够在合成数据集上获得参数图分布的先验, 但是其生成结果始终与现实世界的材质有一定视觉上的差距. 为了进一步逼近现实世界的材质质量, Zhou 等^[111]提出了 PhotoMat 方法, 其是首个完全自监督的材质生成器, 仅在闪光灯照明下由手机拍摄的真实照片上进行训练, 无需任何合成数据, 并解码出多类别且超高质量的材质参数图, 可以直接用于三维真实感渲染. 本文经过分析认为, PhotoMat 能够做到真正无监督的关键是, 将可控的含高光材质生成模块与材质参数图的解码过程进行拆分. 该方法分为 2 步进行: 第 1 步训练一个以光源位置为条件的材质图生成器; 第 2 步巧妙地复用已经训练好的生成器在指定光源位置下生成含高光的材质图片, 并用 rendering 损失训练材质参数图解码器. PhotoMat 包含一个基于 StyleGAN2^[69]的生成器和一个 U-Net SVBRDF maps 解码器, 其结构如图 23 所示. 受到 GAN 架构的限制, PhotoMat 在训练高分辨率版本和可平铺版本的网络模型时出现模式崩溃的情况.

图 23 PhotoMat 的结构^[111]

PhotoMat 在生成材质参数图时是完全不可控的, 即给定一组随机种子, 将生成类别完全不可知的材质结果. 因此, Yuan 等^[112]在 PhotoMat 的基础上提出一个新的扩散模型, 通过单幅参考图像引导生成材质; 为了避免训练扩散模型通常需要的大量计算资源和大规模数据集的需求, 该方法复用了 PhotoMat 基于 StyleGAN2 的材质生成器, 并将其作为扩散模型输出的隐向量的解码器; 扩散模型和 PhotoMat 的结合增强了隐向量的约束, 因此可以通过扩散模型的条件输入, 引导 PhotoMat 生成某类别下的材质参数图. 但是, 由于该方法没有深入研究文本提示条件来提供更严格的约束, 因此该引导过程仍是一个粗略的过程; 此外, 受到 PhotoMat 生成器的限制, 该方法

同样具有 GAN 模式崩溃的缺点.

5.3.3 对比和分析

目前, 基于轻量级设备的材质生成方法不满足于生成大量未知类别的材质参数图, 而是在此基础上提供更加可控的生成引导, 实现更具备多样性的生成类别以及更高分辨率的生成结果. 本文对不同材质生成方法进行对比, 结果如表 7 所示. 本文发现, 研究者倾向于选择 GAN 或者扩散模型作为材质参数生成模型的基本网络结构, 且由于 GAN 具有模式崩塌的问题, 扩散模型逐渐成为主流; 与有材质参数图监督的方法相比, 自监督方法的训练步骤更加烦琐且难度更高, 但允许生成更逼近现实世界材质质量的材质生成结果.

表 7 材质生成方法对比

方法	监督方法	输入图像数量	网络类型	训练时间	分辨率	多样性	编辑能力
Hu 等 ^[103]	有监督	2 张	GAN	2~3 d	可平铺至 1 024×1 024	○○○○	有
Li 等 ^[104]	有监督	0 张	GAN	40 h	64×64	○○○	无
Vecchio 等 ^[105]	有监督	1 张	扩散模型		256×256	○○○○	有
Xin 等 ^[107]	有监督	1~3 张	扩散模型	18 d	2 048×2 048	○○○○	有
Thies 等 ^[109]	有监督	1 张	U-Net		2 048×2 048	○○○○	有
Zeltner 等 ^[110]	有监督	0 张	MLP	4~5 h	1 920×1 080	○○○○	无
Zhou 等 ^[111]	自监督	0 张	StyleGAN2		512×512	○○○○	无
Yuan 等 ^[112]	自监督	1 张	扩散模型、 StyleGAN2	2 d	256×256	○○○○	有

6 结 语

6.1 总 结

材质逆向恢复能够为高度真实感的三维渲染提供多类别、细节丰富且高分辨率的材质参数图数据, 是真实感渲染中的重要研究方向. 本文

对计算机图形学渲染领域中高度真实感材质逆向恢复的方法进行归纳, 将它们分为精密仪器测量方法、基于过程式节点图材质制作的方法和轻量级设备恢复方法; 对 3 类方法中的关键技术进行了详细的探索和调研, 并给出其中里程碑式的方法的步骤以方便读者理解; 此外, 还对不同方法的关键性评价指标进行对比, 方便读者了解不同

方法的优劣, 进一步, 指出各类方法的发展趋势和目前仍面临的重要挑战。

在测量装置方面, 现有的材质逆向恢复方法总体上呈现出由复杂且昂贵的专业级别设备, 过渡到普通消费级别的移动设备的趋势。随着目前现代工业的进步, 更加精准的工业设备, 如三维打印装置、机械臂等极大地提升了采集的精度; 而随着采集精度的提升, 与材质反射率相关的数据量也急剧增加, 给数据拟合带来了庞大的计算量, 并对存储空间有了更大的需求。因此, 现阶段精密仪器测量方法的关键是提出复杂度更低的拟合方法和压缩技术, 以去除冗余数据。测量方法也逐渐与深度神经网络融合, 借助网络模型的强大压缩和生成能力优化采集数据的细节, 这类方法的主要应用是构建大规模、高精度、高分辨率的数据集。从易用性角度上, 当前越来越多的研究者倾向于选择轻量级采集装置, 如手机、平板电脑等来捕获真实世界的材质数据。尽管轻量级采集装置在精度上有所损失, 但是具有端端、便捷、易上手、速度快的优势。

在重建任务中, 接收输入图像的数量整体上表现出由视频连续帧到多幅图像, 再到单幅图像的过渡。多图输入时, 物体的运动线索, 如旋转的物体, 为重建方法提供更多的材质与光线交互的先验, 因此这类方法非常容易去除材质中的高光; 大量不同视角、不同光照方向下的图像输入, 以及有闪光灯和没有闪光灯的图像输入, 都有助于重建方法移除材质中高光。单图输入的重建问题非常困难, 一些方法通过 HA 模块或者引入额外的图案条件控制, 抑制输入图像中饱和区域的影响, 但是仍无法处理图像中的强光; 还有部分方法采用纹理合成的方式对饱和区域进行图像合成, 也仍无法复原准确的全局结构; 输入图像的光照条件设置为漫反射光照时, 重建难度会降低, 但相应的在恢复材质质量和多样性上会受到限制, 因为无法捕获准确的镜面效果。因此, 单幅闪光灯图像输入下的内容重建比多图输入的情况更具有研究价值。

在重建和生成的基础网络选择方面, 扩散模型逐渐替代 GAN 成为主流的网络结构。GAN 在训练过程中受到模式崩塌的影响, 生成内容质量下降。扩散模型拥有更稳定的训练效果, 同时支持多种条件引导, 如文本、图像、线稿图、调色

板等。现有方法存在的主要问题是训练缓慢。一些基于微调的方法, 如 LoRA, ControlNet 等以部分质量损失为代价显著地加速微调过程, 是当前研究的趋势。

6.2 未来工作展望

本文期望高度真实感材质逆向恢复能在更加宽松的光照条件和拍摄条件下, 以更可控的方式快速获取超高分辨率、更广泛类别以及更丰富多样细节的材质参数图, 并且保留编辑能力。该领域未来的一些研究方向如下。

(1) 消除高光伪影

一些带有强烈高光材质图像中, 高光位置的大部分像素处于过度饱和状态, 现有的各种单张图片材质逆向重建方法难以完全消除这些高光带来的影响, 并且很容易在恢复的材质图中引入高光伪影错误。由于这些过饱和像素本质上已经不再包含任何有效的纹理信息, 因此仅通过高光抑制模块降低高光强度是远远不够的。基于对大量现有方法的研究, 本文认为, 高光消除任务既需要使用柔和的 HA 模块确定高光的覆盖范围(预先计算软遮罩), 又需要对感知范围内的过饱和区域进行材质重绘, 补全材质的纹理(同时确保重绘内容与材质的整体纹理一致)。

(2) 更宽松的采集条件

采集过程涉及的配置包括相机、光照以及待采集材质的设置。现有方法通常使用手机同时开启闪光灯对平面放置的材质拍摄照片, 要求手机与平面放置的材质平行, 并且保持固定的高度; 但是, 按照这种方式采集大量材质照片时很难保证拍摄稳定, 在后续的材质恢复过程中引入误差。一些方法在拍摄时提供特殊的标记点, 并在后期使用算法校正拍摄的照片, 但是在实际采集一些材质时, 如狭小空间中的材质, 标记方法很难实施; 现实生活中存在大量非平面的材质, 这类材质按照传统的采集设置进行拍摄也会失败; 现有方法在开启手机闪光灯拍摄照片时要求保持黑暗, 这对于材质采集任务来说无疑是增加了难度。因此, 需要一种更加宽松的采集配置, 包括更为宽松的光照条件、被拍摄物体不必局限于平面形状、无需额外的标记、手机与被拍摄物体的距离不必固定等。受到 Gal 等^[113]提出的一种基于扩散模型的 Textual Inversion 方法的启示, 这类方法可以通过对扩散模型进行微调, 实现从一个物体的几

幅拍摄图像中学习一种风格并转移到任意的其他物体上. 本文认为, 可以采用类似的方式将物体的材质在扩散模型的隐空间中学习出来, 并转移到材质参数图中, 借此适当放宽采集条件.

(3) 更广泛的材质类别

现有的材质逆向方法通常采用 Cook-Torrance 模型表示各种类别的材质, 这种方法对于微观尺度的材质表面细节, 尤其是入射光在掠射角度下呈现的反射和阴影等效果无法做到很好地还原, 这是材质模型本身的局限性导致的. 本文认为, 可以引入基于微薄片理论的材质表示模型, 如 SpongeCake 模型, 提供更好的表达和恢复. SpongeCake 模型提供了多层单次散射的解析解, 同时具有朝向参数, 可以非常方便地扩展到任意层数的材质, 如叶片、多层窗帘等, 并且能够表示在空间上具有强方向性的材质, 如丝绸、木材等; 由于微薄片的优势, 该方法同样能够表示掠射角下入射光的反射和阴影效果.

针对带有清漆和微小颗粒反射效果的汽车车漆, 以及有着强各向异性的金属划痕等材质, 仍需要提供一种解析且高效的材质模型进行拟合, 并支持逆向渲染过程中的可微优化. 这类材质的特点是反射光线非常集中, 形成明亮且清晰的高光点; 对视角的变化非常敏感, 材质表面同一位置点随着观察角度的变化而发生明显的变化. 这类材质与其微观结构密切相关, 如金属表面的微小划痕或者凹凸不平的形状, 显著地影响高光的分布和形状.

(4) 更高效的训练过程

在训练阶段, 材质重建和生成任务中使用的神经网络模型严重依赖高性能设备. 由于这些任务通常需要数天甚至几周的时间才能学习到大规模数据集的分布和先验知识, 因此需要研究如何采用更加高效的微调策略, 获取与直接训练近似的性能和质量. 现有的方法中, LoRA 及其衍生工作能够加速扩散模型的微调过程. 本文经过测试发现, 该方法适合小规模数据集上的微调, 当数据集变大或者微调的迭代次数增加时非常容易训练失败. ControlNet 方法也可以加速微调的速度, 但是对于最新的扩散模型(如 FLUX^[114]), 基于 ControlNet 方法微调扩散模型仍需要数十 GB 的显存才能运行, 且由于扩散模型底模的参数数量及其庞大, 实际训练速度并没有得到本质上的提升. 因此, 仍然需要更好的微调策略, 解决参数计算

量大、训练时间长的问题.

(5) 更灵活的编辑方式

现有的方法通常在材质参数的查询空间(如 Look-up 表), 或者在材质神经表示的隐空间中进行插值来实现材质的编辑, 而基于扩散模型的方法的编辑手段更加灵活. 得益于 CLIP 和交叉注意力模块, 该类别的材质逆向方法可以将二维遮罩图像嵌入扩散模型中, 并且支持使用文字提示和图片引导局部重绘的内容. 但是, 仍然需要更加灵活的编辑方法直接对材质参数图进行操作, 如通过简单的文本提示或者图像引导实现: 修改漫反射图中的纹理(同时确保其他参数图对应的像素也得到合理的更新), 改变法线图的凹凸程度和光滑程度, 以及改变相应的反射率图, 影响实际材质反射高光的颜色、形状和大小等.

高度真实感材质逆向恢复的工作任重道远, 本文相信, 经过不懈的努力和创新, 现有的展望终将变为现实.

参考文献(References):

- [1] Nicodemus F E. Directional reflectance and emissivity of an opaque surface[J]. *Applied Optics*, 1965, 4(7): 767-775
- [2] Lawrence J, Ben-Artzi A, DeCoro C, *et al.* Inverse shade trees for non-parametric material representation and editing[J]. *ACM Transactions on Graphics*, 2006, 25(3): 735-745
- [3] Phong B T. Illumination for computer generated pictures[J]. *Communications of the ACM*, 1975, 18(6): 311-317
- [4] Blinn J F. Models of light reflection for computer synthesized pictures[C] // *Proceedings of the 4th Annual Conference on Computer Graphics and Interactive Techniques*. New York: ACM Press, 1977: 192-198
- [5] Ward G J. Measuring and modeling anisotropic reflection[C] // *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques*. New York: ACM Press, 1992: 265-272
- [6] Cook R L, Torrance K E. A reflectance model for computer graphics[J]. *ACM Siggraph Computer Graphics*, 1981, 15(3): 307-316
- [7] Wang B B, Jin W H, Hašan M, *et al.* SpongeCake: a layered microflake surface appearance model[J]. *ACM Transactions on Graphics*, 2023, 42(1): Article No.8
- [8] Adobe. User guide | substance 3D designer[EB/OL]. [2024-12-01]. <https://helpx.adobe.com/substance-3d-designer/home.html>
- [9] Worthing A G. On the deviation from Lambert's cosine law of the emission from tungsten and carbon at glowing temperatures[J]. *Astrophysical Journal*, 1912, 36: Article No.345
- [10] Walter B, Marschner S R, Li H S, *et al.* Microfacet models for refraction through rough surfaces[C] // *Proceedings of the 18th Eurographics Conference on Rendering Techniques*. Goslar: Eurographics Association, 2007: 195-206
- [11] Heitz E. Understanding the masking-shadowing function in microfacet-based BRDFs[J]. *Journal of Computer Graphics Techniques*, 2014, 3(2): 32-91

- [12] Schlick C. An inexpensive BRDF model for physically-based rendering[J]. *Computer Graphics Forum*, 1994, 13(3): 233-246
- [13] Wang Z, Bovik A C, Sheikh H R, *et al.* Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612
- [14] Chervyakov N, Lyakhov P, Nagornov N. Analysis of the quantization noise in discrete wavelet transform filters for 3D medical imaging[J]. *Applied Sciences*, 2020, 10(4): Article No.1223
- [15] Zhang R, Isola P, Efros A A, *et al.* The unreasonable effectiveness of deep features as a perceptual metric[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 586-595
- [16] Wang J Y, Chan K C K, Loy C C. Exploring CLIP for assessing the look and feel of images[C] //Proceedings of the 37th AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2023: 2555-2563
- [17] Heusel M, Ramsauer H, Unterthiner T, *et al.* GANs trained by a two time-scale update rule converge to a local nash equilibrium[C] //Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 6629-6640
- [18] Murray-Coleman J F, Smith A M. The automated measurement of BRDFs and their application to luminaire modeling[J]. *Journal of the Illuminating Engineering Society*, 1990, 19(1): 87-99
- [19] Drolen B L. Bidirectional reflectance and specularity of twelve spacecraft thermal control materials[J]. *Journal of thermophysics and heat transfer*, 1992, 6(4): 672-679
- [20] Gu J W, Tu C I, Ramamoorthi R, *et al.* Time-varying surface appearance: acquisition, modeling and rendering[J]. *ACM Transactions on Graphics*, 2006, 25(3): 762-771
- [21] Obein G, Audenaert J, Ged G, *et al.* Metrological issues related to BRDF measurements around the specular direction in the particular case of glossy surfaces[C] //Proceedings of the Measuring, Modeling, and Reproducing Material Appearance 2015. Bellingham: Society of Photo-Optical Instrumentation Engineers, 2015: Article No.93980D
- [22] Garces E, Arellano V, Rodriguez-Pardo C, *et al.* Towards material digitization with a dual-scale optical system[J]. *ACM Transactions on Graphics*, 2023, 42(4): Article No.152
- [23] Abdi H, Williams L J. Principal component analysis[J]. *WIREs Computational Statistics*, 2010, 2(4): 433-459
- [24] Wang Lu, Li Binghui, Xu Yanning, *et al.* BRDF (Bidirectional Reflectance Distribution Function) reflectance model deriving method based on measured data: CN, 104700445A[P]. 2015-06-10(in Chinese)
(王璐, 李冰辉, 徐延宁, 等. 一种基于测量数据的 BRDF 反射模型衍生方法: 中国, 104700445A[P]. 2015-06-10)
- [25] Matusik W. A data-driven reflectance model[D]. Cambridge: Massachusetts Institute of Technology, 2003
- [26] Matusik W, Pfister H, Brand M, *et al.* Efficient isotropic BRDF measurement[C] //Proceedings of the 14th Eurographics Workshop on Rendering. Goslar: Eurographics Association, 2003: 241-247
- [27] Zhang Jinbo. Material appearance SVBRDF data acquisition and modeling methods[D]. Hangzhou: Zhejiang University, 2020(in Chinese)
(张锦博. 材质表观 SVBRDF 数据采集与建模方法研究[D]. 杭州: 浙江大学, 2020)
- [28] Ngan A, Durand F, Matusik W. Experimental analysis of BRDF models[C] //Proceedings of the 16th Eurographics Conference on Rendering Techniques. Goslar: Eurographics Association, 2005: 117-126
- [29] Debevec P, Hawkins T, Tchou C, *et al.* Acquiring the reflectance field of a human face[C] //Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2000: 145-156
- [30] Hao Conghui, Du Youyang, Wang Lu, *et al.* Survey of digital face rendering and appearance recovery methods[J]. *Journal of Image and Graphics*, 2024, 29(9): 2513-2540(in Chinese)
(郝琮晖, 杜悠扬, 王璐, 等. 数字人脸渲染与外观恢复方法综述[J]. *中国图象图形学报*, 2024, 29(9): 2513-2540)
- [31] Kang K Z, Chen Z M, Wang J P, *et al.* Efficient reflectance capture using an autoencoder[J]. *ACM Transactions on Graphics*, 2018, 37(4): Article No.127
- [32] Ma X H, Kang K Z, Zhu R S, *et al.* Free-form scanning of non-planar appearance with neural trace photography[J]. *ACM Transactions on Graphics*, 2021, 40(4): Article No.124
- [33] Ma X H, Xu X M, Zhang L Y, *et al.* OpenSVBRDF: a database of measured spatially-varying reflectance[J]. *ACM Transactions on Graphics*, 2023, 42(6): Article No.254
- [34] SideFX. Houdini documentation[EB/OL]. [2024-12-01]. <https://www.sidefx.com/docs/>
- [35] Epic Games, Inc. Unreal Engine 5.5 Documentation[EB/OL]. [2024-12-01]. <https://dev.epicgames.com/documentation/en-us/unreal-engine/unreal-engine-5-5-documentation>(in Chinese)
(Epic Games, Inc. 虚幻引擎 5.5 文档[EB/OL]. [2024-12-01]. <https://dev.epicgames.com/documentation/zh-cn/unreal-engine/unreal-engine-5-5-documentation>)
- [36] Unity Technologies. Unity 6.1 User Manual[OL]. [2024-12-01]. <https://docs.unity3d.com/Manual/index.html>(in Chinese)
(Unity Technologies. Unity 用户手册 2020.3 (LTS)[OL]. [2024-12-01]. <https://docs.unity.cn/cn/2020.3/Manual/UnityManual.html>)
- [37] Scarselli F, Gori M, Tsoi A C, *et al.* The graph neural network model[J]. *IEEE Transactions on Neural Networks*, 2009, 20(1): 61-80
- [38] Cho K, van Merriënboer B, Bahdanau D, *et al.* On the properties of neural machine translation: encoder-decoder approaches[C] //Proceedings of the 8th Workshop on Syntax, Semantics and Structure in Statistical Translation. Stroudsburg: Association for Computational Linguistics, 2014: 103-111
- [39] Devlin J, Chang M W, Lee K, *et al.* BERT: pre-training of deep bidirectional transformers for language understanding[C] //Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: Association for Computational Linguistics, 2019: 4171-4186
- [40] Guerrero P, Hašan M, Sunkavalli K, *et al.* MatFormer: a generative model for procedural materials[J]. *ACM Transactions on Graphics*, 2022, 41(4): Article No.46
- [41] Hu Y W, Guerrero P, Hašan M, *et al.* Generating procedural materials from text or image prompts[C] //Proceedings of the ACM SIGGRAPH 2023 Conference Proceedings. New York: ACM Press, 2023: Article No.4
- [42] Li B C, Hu Y W, Guerrero P, *et al.* Procedural material generation with reinforcement learning[J]. *ACM Transactions on*

- Graphics, 2024, 43(6): Article No.280
- [43] MacQueen J. Some methods for classification and analysis of multivariate observations[C] //Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press, 1967: 281-297
- [44] Hu Y W, Dorsey J, Rushmeier H. A novel framework for inverse procedural texture modeling[J]. ACM Transactions on Graphics, 2019, 38(6): Article No.186
- [45] Shi L, Li B C, Hašan M, *et al.* Match: differentiable material graphs for procedural material capture[J]. ACM Transactions on Graphics, 2020, 39(6): Article No.196
- [46] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C] //Proceedings of the 3rd International Conference on Learning Representations. Washington DC: ICLR, 2015: 1-14
- [47] Hu Y W, He C A, Deschaintre V, *et al.* An inverse procedural modeling pipeline for SVBRDF maps[J]. ACM Transactions on Graphics, 2022, 41(2): 1-17
- [48] Virtanen P, Gommers R, Oliphant T E, *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python[J]. Nature Methods, 2020, 17(3): 261-272
- [49] Cui Jiaxu, Yang Bo. Survey on Bayesian optimization methodology and applications[J]. Journal of Software, 2018, 29(10): 3068-3090(in Chinese)
(崔佳旭, 杨博. 贝叶斯优化方法和应用综述[J]. 软件学报, 2018, 29(10): 3068-3090)
- [50] Hu Y W, Guerrero P, Hašan M, *et al.* Node graph optimization using differentiable proxies[C] //Proceedings of the ACM SIGGRAPH 2022 Conference Proceedings. New York: ACM Press, 2022: Article No.5
- [51] Aittala M, Weyrich T, Lehtinen J. Two-shot SVBRDF capture for stationary materials[J]. ACM Transactions on Graphics, 2015, 34(4): Article No.110
- [52] Shafer S A. Using color to separate reflection components[J]. Color Research and Application, 1985, 10(4): 210-218
- [53] Ghosh A, Chen T B, Peers P, *et al.* Estimating specular roughness and anisotropy from second order spherical gradient illumination[J]. Computer Graphics Forum, 2009, 28(4): 1161-1170
- [54] Ma W C, Hawkins T, Peers P, *et al.* Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination[C] //Proceedings of the 18th Eurographics Conference on Rendering Techniques. Goslar: Eurographics Association Press, 2007: 183-194
- [55] Riviere J, Peers P, Ghosh A. Mobile surface reflectometry[J]. Computer Graphics Forum, 2016, 35(1): 191-202
- [56] Hertzmann A, Seitz S M. Shape and materials by example: a photometric stereo approach[C] //Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings. Los Alamitos: IEEE Computer Society Press, 2003: I-1
- [57] Hui Z, Sankaranarayanan A C. A dictionary-based approach for estimating shape and spatially-varying reflectance[C] //Proceedings of the IEEE International Conference on Computational Photography (ICCP). Los Alamitos: IEEE Computer Society Press, 2015: 1-9
- [58] Palma G, Callieri M, Dellepiane M, *et al.* A statistical method for SVBRDF approximation from video sequences in general lighting conditions[J]. Computer Graphics Forum, 2012, 31(4): 1491-1500
- [59] Xia R, Dong Y, Peers P, *et al.* Recovering shape and spatially-varying surface reflectance under unknown illumination[J]. ACM Transactions on Graphics, 2016, 35(6): Article No.187
- [60] Chandraker M. On shape and material recovery from motion[C] //Proceedings of the 13th European Conference on Computer Vision. Heidelberg: Springer, 2014: 202-217
- [61] Kim K, Gu J W, Tyree S, *et al.* A lightweight approach for on-the-fly reflectance estimation[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 20-28
- [62] Deschaintre V, Aittala M, Durand F, *et al.* Flexible SVBRDF capture with a multi-image deep network[J]. Computer Graphics Forum, 2019, 38(4): 1-13
- [63] Boss M, Jampani V, Kim K, *et al.* Two-shot spatially-varying brdf and shape estimation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 3981-3990
- [64] Ye W J, Dong Y, Peers P, *et al.* Deep reflectance scanning: recovering spatially-varying material appearance from a flash-lit video sequence[J]. Computer Graphics Forum, 2021, 40(6): 409-427
- [65] Zhu P F, Lai S C, Chen M F, *et al.* SVBRDF reconstruction by transferring lighting knowledge[J]. Computer Graphics Forum, 2023, 42(7): Article No.e14973
- [66] Gao D, Li X, Dong Y, *et al.* Deep inverse rendering for high-resolution SVBRDF estimation from an arbitrary number of images[J]. ACM Transactions on Graphics, 2019, 38(4): Article No.134
- [67] Deschaintre V, Drettakis G, Bousseau A. Guided fine-tuning for large-scale material transfer[J]. Computer Graphics Forum, 2020, 39(4): 91-105
- [68] Guo Y, Smith C, Hašan M, *et al.* MaterialGAN: reflectance capture using a generative SVBRDF model[J]. ACM Transactions on Graphics, 2020, 39(6): Article No.254
- [69] Karras T, Laine S, Aittala M, *et al.* Analyzing and improving the image quality of StyleGAN[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 8107-8116
- [70] Saryzadi S, Murphy C, Mudur S P. Disentangled rendering loss for supervised material property recovery[C] //Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021) – GRAPP. Setúbal: SciTePress, 2021: 113-121
- [71] Aittala M, Aila T, Lehtinen J. Reflectance modeling by neural texture synthesis[J]. ACM Transactions on Graphics, 2016, 35(4): Article No.65
- [72] Henzler P, Deschaintre V, Mitra N J, *et al.* Generative modeling of BRDF textures from flash images[J]. ACM Transactions on Graphics, 2021, 40(6): Article No.284
- [73] Li X, Dong Y, Peers P, *et al.* Modeling surface appearance from a single photograph using self-augmented convolutional neural networks[J]. ACM Transactions on Graphics, 2017, 36(4): Article No.45
- [74] Ye W J, Li X, Dong Y, *et al.* Single image surface appearance modeling with self-augmented cnns and inexact supervision[J]. Computer Graphics Forum, 2018, 37(7): 201-211
- [75] Sun Bin. Research on key problems of SVBRDF material efficient modeling based on CNN[D]. Qingdao: Shandong University, 2019(in Chinese)
(孙彬. 基于 CNN 的 SVBRDF 材质高效建模关键问题研究[D]. 青岛: 山东大学, 2019)
- [76] Deschaintre V, Aittala M, Durand F, *et al.* Single-image

- SVBRDF capture with a rendering-aware deep network[J]. *ACM Transactions on Graphics*, 2018, 37(4): Article No.128
- [77] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C] //Proceedings of the 18th International Conference on Medical image computing and computer-assisted intervention. Heidelberg: Springer, 2015: 234-241
- [78] Li Z Q, Sunkavalli K, Chandraker M. Materials for masses: SVBRDF acquisition with a single mobile phone image[C] //Proceedings of the 15th European Conference on Computer Vision. Heidelberg: Springer, 2018: 74-90
- [79] Wang L, Zhang L H, Gao F Z, *et al.* DeepBasis: hand-held single-image SVBRDF capture via two-level basis material model[C] //Proceedings of the SIGGRAPH Asia 2023 Conference Papers. New York: ACM Press, 2023: Article No.85
- [80] Guo J, Lai S C, Tao C Z, *et al.* Highlight-aware two-stream network for single-image SVBRDF acquisition[J]. *ACM Transactions on Graphics*, 2021, 40(4): Article No.123
- [81] Zhou X L, Kalantari N K. Look-ahead training with learned reflectance loss for single-image SVBRDF estimation[J]. *ACM Transactions on Graphics*, 2022, 41(6): Article No.266
- [82] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks[C] //Proceedings of the 34th International Conference on Machine Learning. Cambridge: PMLR, 2017: 1126-1135
- [83] Luo X, Scandolo L, Bousseau A, *et al.* Single-image SVBRDF estimation with learned gradient descent[J]. *Computer Graphics Forum*, 2024, 43(2): Article No.e15018
- [84] Martin R, Roullier A, Rouffet R, *et al.* MaterIA: single image high-resolution material capture in the wild[J]. *Computer Graphics Forum*, 2022, 41(2): 163-177
- [85] Guo J, Lai S C, Tu Q H, *et al.* Ultra-high resolution SVBRDF recovery from a single image[J]. *ACM Transactions on Graphics*, 2023, 42(3): Article No.33
- [86] Zhao Y Z, Wang B B, Xu Y N, *et al.* Joint SVBRDF recovery and synthesis from a single image using an unsupervised generative adversarial network[C] //Proceedings of the Eurographics Symposium on Rendering 2020. Goslar: Eurographics Association Press, 2020: 53-66
- [87] Wen T, Wang B B, Zhang L, *et al.* SVBRDF Recovery from a single image with highlights using a pre-trained generative adversarial network[J]. *Computer Graphics Forum*, 2022, 41(6): 110-123
- [88] Zhou X L, Kalantari N K. Adversarial single-image SVBRDF estimation with hybrid training[J]. *Computer Graphics Forum*, 2021, 40(2): 315-325
- [89] Vecchio G, Palazzo S, Spampinato C. SurfacerNet: adversarial SVBRDF estimation from a single image[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2021: 12820-12828
- [90] Zhou X L, Hašan M, Deschaintre V, *et al.* TileGen: tileable, controllable material generation and capture[C] //Proceedings of the SIGGRAPH Asia 2022 Conference Papers. New York: ACM Press, 2022: Article No.34
- [91] Rodriguez-Pardo C, Dominguez-Elvira H, Pascual-Hernandez D, *et al.* UMat: uncertainty-aware single image high resolution material capture[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2023: 5764-5774
- [92] Shi Z Q, Lin X Y, Song Y. An attention-embedded GAN for SVBRDF recovery from a single image[J]. *Computational Visual Media*, 2023, 9(3): 551-561
- [93] Liu Kemei. High-precision SVBRDF recovery based-on image[D]. Jinan: Shandong University, 2023(in Chinese) (刘克梅. 基于图像的高精度 SVBRDF 重建方法[D]. 济南: 山东大学, 2023)
- [94] Hu E J, Shen Y L, Wallis P, *et al.* LoRA: low-rank adaptation of large language models[C] //Proceedings of the 12th International Conference on Learning Representations. Washington DC: ICLR, 2022: 1-13
- [95] Zhang L M, Rao A Y, Agrawala M. Adding conditional control to text-to-image diffusion models[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2023: 3813-3824
- [96] Vecchio G, Martin R, Roullier A, *et al.* ControlMat: a controlled generative approach to material capture[J]. *ACM Transactions on Graphics*, 2023, 43(5): Article No.164
- [97] Rombach R, Blattmann A, Lorenz D, *et al.* High-resolution image synthesis with latent diffusion models[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2022: 10684-10695
- [98] Sartor S, Peers P. MatFusion: a generative diffusion model for SVBRDF capture[C] //Proceedings of the SIGGRAPH Asia 2023 Conference Papers. New York: ACM Press, 2023: Article No.86
- [99] Dong Y, Chen G J, Peers P, *et al.* Appearance-from-motion: Recovering spatially varying surface reflectance under unknown lighting[J]. *ACM Transactions on Graphics*, 2014, 33(6): Article No.193
- [100] Zhou Z M, Chen G J, Dong Y, *et al.* Sparse-as-possible SVBRDF acquisition[J]. *ACM Transactions on Graphics*, 2016, 35(6): Article No.189
- [101] Hui Z, Sunkavalli K, Lee J Y, *et al.* Reflectance capture using univariate sampling of BRDFs[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 5372-5380
- [102] Li Z Q, Shafiei M, Ramamoorthi R, *et al.* Inverse rendering for complex indoor scenes: shape, spatially-varying lighting and SVBRDF from a single image[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 2472-2481
- [103] Hu Y W, Hašan M, Guerrero P, *et al.* Controlling material appearance by examples[J]. *Computer Graphics Forum*, 2022, 41(4): 117-128
- [104] Li X, Dong Y, Peers P, *et al.* Synthesizing 3D shapes from silhouette image collections using multi-projection generative adversarial networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 5530-5539
- [105] Vecchio G, Sortino R, Palazzo S, *et al.* MatFuse: controllable material generation with diffusion models[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2024: 4429-4438
- [106] Esser P, Rombach R, Ommer B. Taming transformers for high-resolution image synthesis[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern

- Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 12868-12878
- [107] Xin L X, Zhang Z, Wei J F, *et al.* DreamPBR: text-driven generation of high-resolution SVBRDF with multi-modal guidance[OL]. [2024-12-01]. <https://arxiv.org/abs/2404.14676>
- [108] Wang X T, Xie L B, Dong C, *et al.* Real-ESRGAN: training real-world blind super-resolution with pure synthetic data[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2021: 1905-1914
- [109] Thies J, Zollhöfer M, Nießner M. Deferred neural rendering: Image synthesis using neural textures[J]. ACM Transactions on Graphics, 2019, 38(4): Article No.66
- [110] Zeltner T, Rousselle F, Weidlich A, *et al.* Real-time neural appearance models[J]. ACM Transactions on Graphics, 2024, 43(3): Article No.33
- [111] Zhou X L, Hašan M, Deschaintre V, *et al.* PhotoMat: a material generator learned from single flash photos[C] //Proceedings of the ACM SIGGRAPH 2023 Conference Proceedings. New York: ACM Press, 2023: Article No.49
- [112] Yuan L, Yan D K, Saito S, *et al.* DiffMat: Latent diffusion models for image-guided material generation[J]. Visual Informatics, 2024, 8(1): 6-14
- [113] Gal R, Alaluf Y, Atzmon Y, *et al.* An image is worth one word: personalizing text-to-image generation using textual inversion[C] //Proceedings of the 11th International Conference on Learning Representations. Washington DC: ICLR, 2023: 1-17
- [114] Black Forest Labs. Introducing FLUX.1 tools[EB/OL]. [2024-12-01]. <https://blackforestlabs.ai/flux-1-tools/>