

# HumanMaterial : 通过渐进训练从单张图像进行人材质估计

Yu Jiang, Jiahao Xia, Jiongming Qin, Yusen Wang, Tuo Cao, and Chunxia Xiao \* *Senior Member, IEEE,*

**Abstract**—基于物理渲染的全身人类逆向渲染旨在获取高质量的材质, 这有助于在任意光照下实现照片级真实感渲染。该任务需要估计多个材质贴图, 通常依赖于渲染结果的约束。材质贴图缺乏约束使得逆向渲染成为一个不适宜问题。以前的工作通过构建材质数据集进行训练来缓解这一问题, 但是它们简化的材质数据和渲染方程导致了渲染结果真实感有限, 尤其是皮肤部分。为进一步缓解这一问题, 我们基于扫描的真实数据和统计材质数据构建了一个更高质量的数据集 (OpenHumanBRDF)。除了法线、漫反射反射率、粗糙度、镜面反射率之外, 我们还生成了位移和次表面散射, 以增强渲染结果的真实感, 特别是对于皮肤。随着需要预测的材质任务的增多, 像以前的工作中那样使用端到端模型难以平衡各种材质贴图之间的重要性, 导致模型欠拟合。因此, 我们设计了一个具有渐进训练策略的模型 (HumanMaterial), 以充分利用材质贴图的监督信息, 提升材质估计性能。HumanMaterial 首先通过三个先验模型获得初始材质结果, 然后通过一个微调模型对结果进行优化。先验模型估计不同的材质贴图, 每个贴图对渲染结果的意义不同。因此, 我们设计了一个 Controlled PBR Rendering (CPR) 损失, 在先验模型训练期间增强待优化材质的重要性。对 OpenHumanBRDF 数据集和真实数据的广泛实验表明, 我们的方法实现了最先进的性能。

**Index Terms**—Reflectance Modeling, Physically Based Rendering, Deep Learning

## I. 引言

基于物理的渲染 (PBR) 是一种计算机图形学中的光线追踪渲染技术, 通过利用物理上准确的材质定义 (即 PBR 材质) 来实现任意光照下的照片级真实渲染。这种技术广泛应用于电子游戏、虚拟现实、3D 打印和动画电影中。

逆向渲染是一种广泛用于获取人类 PBR 材料的技术, 其核心在于从渲染结果推断对象的材质贴图。不同组合的 PBR 材料可能会产生类似的渲染结果, 而其中一些组合甚至可能明显错误。此外, 输入图像数据缺乏关键信息 (例如只有单一图像) 以及 PBR 渲染模型的简化会加剧材质估计中的模糊性。上述问题使得人类逆向渲染成为一个不适宜的问题。

解决该问题的有效方法是使用监督信息来约束估计的材料。以前的工作 [1]–[5] 已经尝试构建人类材质数据集来缓解这一任务。在这些工作中, TR [1] 基于真实捕获数据构建了一个高质量的数据集, 但材质贴图类型的简化和渲染模型的简化限制了其应用。另一方面, FBHR [2] 使用

This work is partially supported by the National Natural Science Foundation of China (No. 62372336).

Yu Jiang, Jiongming Qin, Yusen Wang, Tuo Cao, and Chunxia Xiao are with the School of Computer Science, Wuhan University, Wuhan, China (e-mail: jiangyu1181@whu.edu.cn, jiongming@whu.edu.cn, wangyusen@whu.edu.cn, maplect@whu.edu.cn, cxxiao@whu.edu.cn).

Jiahao Xia is with the Faculty of Engineering and IT, University of Technology Sydney Ultimo, NSW, 2007, Australia (e-mail: jiahao.xia-1@uts.edu.au)

\* Chunxia Xiao is the corresponding author.

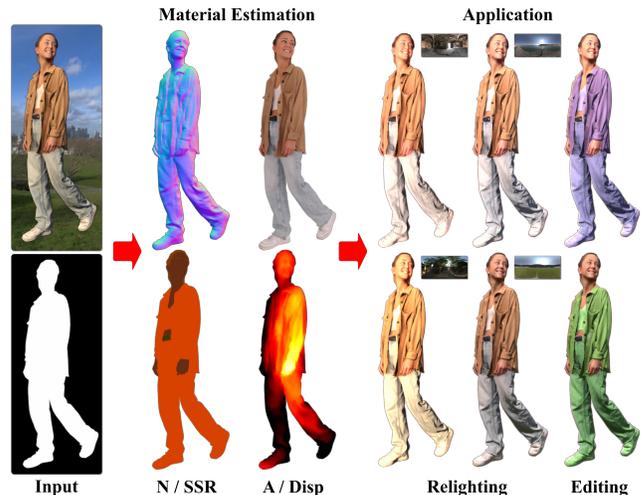


Fig. 1. HumanMaterial. 我们的方法旨在从单个人物图像和前景蒙版中估计 PBR 材料。它能够估计包括“正常”(N)、“漫反射反射率”(A)、“次表面散射、镜面反射率和粗糙度”(SSR) 以及“位移”(Disp) 在内的 PBR 材料。利用这些估计结果, 我们的方法可以在新的环境映射下实现诸如重新光照和材质编辑等逼真的应用。

渲染引擎构建了合成数据。更多类型的 PBR 材质贴图扩展了应用, 但渲染结果的真实感有限。上述数据集难以满足更高质量 PBR 渲染的要求, 并且很难表现真实的皮肤。

为了进一步缓解这个病态的问题, 我们制作了一个更高质量的人体 PBR 材质数据集。各种材质的统计数据在 [6] 中展示, 其中与人体相关的有头发、皮肤、织物和皮革。在此基础上, 我们手动设计了材质参数, 并根据统计数据在 Blender [7] 中生成 OpenHumanBRDF。该数据集包括 147 个人体模型, 覆盖六个种族。除了广泛使用的材质贴图 (法线、漫反射反射率、粗糙度和镜面反射率), 我们还生成了置换和次表面散射数据。置换贴图显示表面, 并在没有完整几何信息时有助于制作更详细的渲染结果。次表面散射贴图可以通过使用高级 PBR 着色器增强皮肤的现实感。这使得呈现的皮肤视觉效果更真实, 更接近真实世界。

为了实现逼真的渲染效果, 我们建立了更多的材质数据并采用了高级的 PBR 着色器, 这导致了模型需要估计的材质贴图数量的增加。如图 1 所示, 该模型需要估计六个材质贴图, 从而进行重光照和材质编辑等应用。先前使用端到端模型同时估计所有材质贴图的工作会加剧模型训练的难度, 并限制材质估计的性能。因此, 为了充分利用材质贴图的监督信息, 我们设计了 HumanMaterial, 采用渐进式训练策略。

HumanMaterial 首先使用三个先验模型估计六个材质图, 然后使用微调模型联合优化所有材质。独立训练的先

验模型可以提高材质估计的性能。微调操作通过满足各种光照约束的 PBR (物理渲染) 增强了材质图之间的连接, 并有助于获得更具说服力的材质图。

基于材质贴图之间的关联性, 我们精心设计了每个先验模型估计的目标贴图。具体来说, Geometry Prior Model 估计法线和位移, Albedo Prior Model 估计漫反射反照率, RSS Prior Model 估计粗糙度、镜面反照率、次表面散射。此外, 每个材质贴图对渲染结果的影响程度不同。为了增加其影响, 我们设计了一个 Controlled PBR Rendering 损失函数。这个渲染损失的关键在于, 在合理范围内仔细设置非优化材质的值, 而不仅仅是使用 GT 数据。通过这样做, 它有效减轻了非优化参数在各种光照条件下对渲染结果的影响。这种方法还使训练过程更加稳定, 并且估计结果更加准确。这使得在训练过程中估计的材质贴图更容易被优化。

总结来说, 我们做出了以下主要技术贡献。

- 生成一个高质量的人体 PBR 材料数据集 (OpenHumanBRDF), 该数据集具有精心设计的参数设置和逼真的皮肤表现。
- 提出一种从单个人体图像进行材料估计的方法 (HumanMaterial), 该方法构建三个先验模型以获得初始材料, 经过微调后可以估计出更高质量的 PBR 材料。
- 设计一个 Controlled PBR Rendering 损失函数, 可以增强材料在训练期间被优化的重要性, 从而提高先验模型的性能。

我们在自有数据集和真实数据上的实验表明, 我们的方法能够估计出高质量的材料, 实现逼真的重新照明, 并达到最新的技术水平。

## II. 相关工作

逆向渲染是计算机图形学中备受推崇的研究课题。在之前的工作中, 一些方法基于简化的 PBR 模型 [1], [8], [9] 进行外观分解。其他方法则使用基于神经网络的渲染来实现重光照 [1], [2], [8], [10]。为了在复杂光照条件下实现更稳健的重光照和材料编辑, 一些工作专注于估计 PBR 材料 [5], [11], [12]。这些方法可以根据输入数据进行分类, 包括基于单张图像 [13], [14]、多张图像 [15]–[19] 和基于视频的方法 [20]。

### A. 外观分解

单张图像的人体外观分解。对于单张人体图像的再光照, 一种有效的方法是使用经验模型对外观进行分解。例如, Lagunas *et al.* [2] 提出了一种数据驱动的方法, 将单张全身图像分解为漫反射、镜面反射和与光有关的残余项。Pandey *et al.* [1] 将外观分解为漫反射和镜面反射成分, 并训练一个神经渲染器以实现再光照。PhorHum [8] 是一个端到端可训练的深度神经网络, 能够通过单目 RGB 图像进行逼真 3D 人类重建。多图像外观分解。基于多图像和神经辐射场可以实现无监督的外观表示。基于一个数据驱动的 BRDF 先验和 NeRF, NeRFactor [21] 从多视角图像中恢复物体的形状和空间变化的反射率。PhysSG [15] 使用球面高斯混合来表示镜面 BRDF 和环境光照, 并可以从多图像中重建自然的几何、材质和光照。上述方法能够获得令人印象深刻的再光照效果。由于难以获得 PBR 材质, 这限制了在复杂光照下的渲染逼真性, 也使得实现类级材质编辑变得困难。

### B. PBR 材质估计与重建

单张图像的材料估计。单张图像的材料估计比外观分解更具挑战性, 因为病态的 PBR 问题。在生产了近似平面的 PBR 材料数据集 [13] 之后, 已经发表了一些重要的研究 [22], [23]。RADN [13] 制作了一个大型的合成近平面表面材料数据集, 并使用 unet 从单张近平面拍摄中恢复材料。Guo *et al.* [14] 设计了一个高光感知卷积模块, 以缓解由于过曝区域造成的材料估计中的伪影问题。基于预训练网络, Guo *et al.* [24] 和 Guo *et al.* [25] 通过在潜在空间中优化网络来减少重建结果与外观 (输入) 之间的错误。多图像材质重建。许多研究 [16], [26]–[30] 已尝试从多张图像中重建 PBR 材质。例如, Nvdiffric [31] 利用基于坐标的网络来紧凑地表示体积纹理, 使用可微分的四面体行进法在表面上实现基于梯度的优化。镜面反射是视点相关的, 破坏了多视图一致性, 这使得重建反射物体的几何和材质具有挑战性 [32]。为了解决这一挑战, NeRO [12] 提出了一种两步法, 首先重建几何, 然后恢复环境光和材质。多图像人体材质重建。最近, 有一些研究 [33]–[35] 尝试估计与人相关的 PBR 材质。Relit-NeuLF [36] 使用双平面光场表示来参数化每条光线在一个 4D 坐标系统中, 实现高效学习和推理, 然后以自监督的方式从多视图图像中恢复面部材料。在使用廉价的偏振箱捕捉具有交叉偏振和平行偏振光的面部图像后, Azinovi *et al.* [37] 使用运动结构重建面部的显式表面网格, 然后在可微分渲染中利用相机和光源的共定位, 使用合成分析方法优化面部材质。虽然基于可微分逆向渲染的方法在静态物体上取得了成功, 但一些令人印象深刻的工作 [20], [38], [39] 试图在人动态材质估计上开展研究。

以上方法在材质重建方面取得了显著的成果, 而对于具有多样材质类别的人体图像, 缺乏类别引导的材质区分可能会降低准确性。此外, 这可能导致渲染结果的逼真度差异不足, 尤其是在织物、皮革和皮肤方面。得益于我们的具有分类的 PBR 材质数据集, 我们的方法能够估计更准确的 PBR 材质, 实现逼真的光照和材质编辑。

TABLE I

与以往数据集的比较 (FBHR [2]、AFHIR [3] 和 TR [1])。“OPEN”表示数据集是开源的。“TYPE”表示数据类型, 而“SCAN”表示扫描自真实数据。“RENDER”表示数据支持的渲染器类型。“GEO”表示数据集中的几何数据。“SKIN”表示数据集是否增强了皮肤的真实感。

	Open	Type	Render	Geo	Skin
FBHR	/	Synthetic	Neural	/	/
AFHIR	/	Synthetic	PBR+Neural	Depth	/
TR	/	Scan	Neural	/	/
Ours	✓	Scan	PBR	Surface	✓

## III. OPENHUMANBRDF 数据集

以往的研究已经建立了一些材料数据集, 如表格 I 所示, FBHR [2] 和 AFHIR [3] 都建立了一个用于从单张图像估计人类 PBR 材料的数据集。然而, 这些数据是合成的, 与真实场景有一个固定的差距。TR [1] 建立了一个用于肖像重光的真实数据集, 其中材料数据包括法线和漫反射反照率, 但不支持 PBR 渲染。需要注意的是, 这些数据集都不是开源的, 我们将在本文被接受后发布我们自己的数据集。

在工业中生成高质量的 PBR 材质数据集是重要且具有商业价值的。PBR 反向渲染的病态问题通常要求有经验的艺术家进行繁琐的调整。此外，人体材料复杂，不同个体的外观差异很大。所有这些因素导致了高质量开源人体 PBR 材质数据集的缺乏。

基于大量的实验数据，RTR [6] 计算出了真实世界中典型材料的镜面反射率的数值范围，涉及人类相关材料的有头发、皮肤、织物和皮革。在大多数情况下，它们也是全身的主要材料。为了使人类 PBR 材料数据的生产可行并简化过程，我们做出一个合理的假设，即全身主要包含四种材料类别，即头发、皮肤、织物和皮革。参考 RTR [6] 中这些材料镜面反射率的统计范围，我们为每种材料设置了一个值。此外，基于不同材料粗糙度的大致大小关系（从小到大：皮革、皮肤、头发、织物），我们在 Blender 中调整了每种材料的粗糙度值，以确保在不同光照下的真实外观。最终设置如表 ?? 所示。此外，我们收集了 1092 张 HDR 环境贴图（来自 [40] 的 782 张真实世界贴图，以及我们创建的 310 张合成贴图），它们用于在 Blender 中渲染外观。

在确认材料的参数设置后，我们从 RenderPeople [41] 收集了 147 个高质量的人体模型，这些原始数据是从现实世界中扫描的。为了涵盖广泛的人类多样性，我们制定了三个标准。(1) 我们选择的模型在男性和女性之间保持 50% 的平衡。(2) 所有人体模型来自六个种族，包括亚洲人、黑人、印度人、中东人、白人和西班牙裔/拉丁裔，每个种族占 16.7%。(3) 所涵盖的年龄范围是所有年龄段中最大的，为 19-50 岁。

基于基础数据和合理的设置，我们在 Blender 中构建了 OpenHumanBRDF 数据集。(1) 我们标记了每个人体模型的头发、皮肤、织物和皮革区域。(2) 根据上述设置，我们手动调整每个类别的粗糙度、镜面反射率和次表面散射。(3) 我们为每个人体烘焙材质 UV 贴图，包括法线、漫反射反照率、粗糙度、镜面反照率、次表面散射和位移贴图。(4) 我们固定相机，旋转人体模型，并在两个环境贴图下渲染外观。

通过以上处理，我们可以获得合理的材质图，与仅使用原始收集数据的结果相比，渲染出更加真实的外观。正如图 ?? 所示，(a) 中的织物（裙子）有不合理的高光，使其看起来像塑料且不真实。我们可以发现 (b) 看起来更加真实，证明了我们的 PBR 材料设置是准确的。此外，通过对比 (a)、(b) 和 (c)，我们可以看到材料中的次表面散射增强了皮肤的真实感，特别是在鼻子和手的部位。

我们生成了 147 组数据，其中 127 组用于训练，20 组用于测试。对于训练数据，我们为每个模型渲染 100 组图像。因此，训练集总共包含 12,700 组数据。在渲染图像集时，初始视角为相机直接面向模型。然后，我们围绕模型均匀旋转相机一圈，渲染 100 次以获得多视角渲染结果。对于测试数据，我们在初始视角下为每个模型渲染 10 组图像。因此，测试集总共有 200 组数据。图 ?? 中展示了三个样本，图像集的分辨率是  $512 \times 512$  或  $4096 \times 2048$ （仅环境贴图）。每个集合包括前景遮罩、PBR 材料（法线、漫反射反照率、粗糙度、镜面反照率、次表面散射和位移）以及五个在不同环境贴图下的重新光照结果。位移数据呈现了从几何表面到世界中心平面的距离。需要注意的是，材料类别的边界可能是模糊的，例如头发和皮肤，这可能导致细微的分类错误。这也是为什么我们不简单地将材料估计视作分类任务的原因。

给定一张单一的人类图像，我们的任务是估计 PBR 材

质，包括法线、漫反射反照率、粗糙度、镜面反照率、位移和次表面散射。从单张图像中估计六种材质图是一项具有挑战性的任务。为了充分利用材质监督信息，并使模型的训练更容易拟合，如图 ?? 所示，我们提出了 HumanMaterial，并采用逐步训练策略来处理这一任务。HumanMaterial 首先使用三个先验模型来估计六种材质先验。每个模型分别估计一种、两种和三种材质。随后，使用联合优化来优化结果。

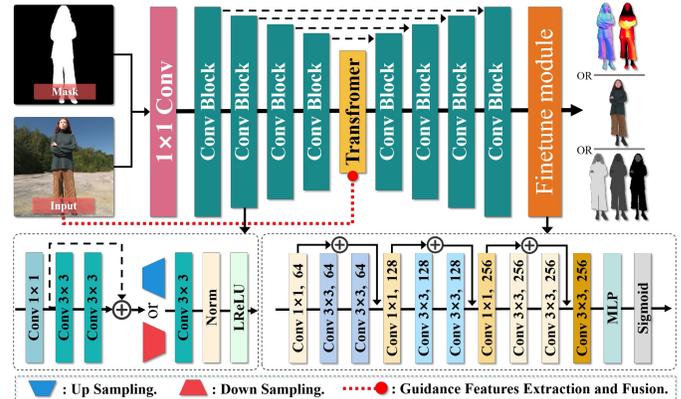


Fig. 2. 先验模型的结构。三个先验模型的网路相似，它们之间的区别仅在于输出的材料地图。每个卷积块将执行一次上采样或下采样。类似于图 ??，我们也使用指导编码器从“输入”中提取指导特征并进行融合。

HumanMaterial 包含两个阶段：(1) 以单张全身图像作为输入，使用 [42] 提取“前景”（如果未提供“掩膜”），然后使用三个先验模型来估计材料先验（见 Sec. ??），这些模型是在我们的 OpenHumanBRDF 数据集上训练的。(2) 基于三个预训练的先验模型，我们设计了一个引导编码器，用于从估计的先验材料中提取引导特征。随后，我们设计了一个 Finetuning Model (FTM) 来优化先验。该模型以前景图为输入，并利用引导特征来指导解码过程。在联合优化后，FTM 估计最终结果。

为了实现从单张图像中估计六个材质图，并使其更具实用性，一个直接且有效的方法是使用多个模型进行材质估计。这可以在一定程度上简化估计过程，同时提高准确性和可靠性。此外，当硬件资源有限时，独立地训练三个具有不同功能的模型可以更好地利用资源，相较于将所有功能集成到一个模型中。

根据材料贴图之间的相关性，我们精心设计了每个先验模型的目标估计图。具体来说，法线和位移信息与几何表面密切相关。因此，我们创建一个 Geometry Prior Model (GPM) 来估计它们。在数据构建过程中，我们根据材料类别设置 R、S 和 SSS 的值。基于此，我们只需使用一个 RSS Prior Model (RPM) 即可有效地估计它们。值得注意的是，在材料类别的边界处存在类别混合。因此，我们不能简单地将 R、S 和 SSS 的估计视为分类任务。最后，我们设计了一个 Albedo Prior Model (APM) 来准确估计剩余的漫反射反照率。

如图 2 所示，三个先验模型在网络结构上具有相似性。它们每个的输入都是“输入”图像和“遮罩”，但这三个模型之间的主要区别在于网络输出的通道。具体来说，GPM 的输出为四个通道，包括三个通道的法线和一个通道的位移。APM 的输出是三通道的漫反射反照率。RPM 的输出也是三个通道，包含三个单独的通道，即粗糙度、镜面反照率和次表面散射。

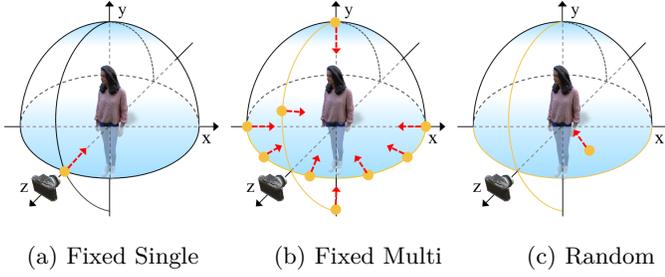


Fig. 3. 用于 PBR 渲染的三种光照类型。“固定单点”指的是在固定位置的单点光源。“固定多点”指的是在固定位置的多点光源。“随机”指的是在正半球上随机位置的单点光源。

先验模型呈现了一个编码器-解码器结构。首先，输入图像进行  $1 \times 1$  卷积操作。然后，通过四个卷积块进行下采样，在从“输入”中提取指导特征后，使用一个变换器块进行指导特征的融合和潜在特征编码，接着它通过一个四层的卷积块进行上采样。在上采样过程中，通过跳跃连接来融合相同尺度的特征。最后，通过微调模块输出目标图像。下采样和上采样过程中通道的数量依次为 64, 128, 256, 512, 512, 512, 256, 128 和 64。卷积块的结构和微调模块的结构分别在图 2 的左下角和右下角展示。卷积块包含  $1 \times 1$  和  $3 \times 3$  卷积操作，如上采样和下采样、规范化和激活函数。微调模块包含三组卷积和残差连接操作。经过多层感知器 (MLP) 和 sigmoid 激活函数后，可以获得目标图像。

#### A. Finetuning Model

训练三个独立的先验模型来估计多种材料是一种有效的策略，并且减少了只使用单一模型的任务难度。但这也削弱了不同材料之间的联系，使得难以保证物理合理性。基于先验模型估计的材料先验似乎具有适当的性能。将其作为初始材料，然后通过联合优化进一步增强材料组合的合理性是一个合理的微调策略。因此，我们构建了一个 Finetuning Model 来处理这个任务。

如图 ?? 所示，FTM 包含一个潜编码器、四个材料 (Nor, Disp, Diff 和 RSS) 解码器，以及四个指导编码器。FTM 的输入是“前景”和三个先验模型的估计先验。潜编码器将“前景”编码为潜在特征，并作为四个材料编码器的输入。同时，指导编码器用于将先验编码为指导特征，也作为四个材料编码器的输入。通过融合潜在特征和指导特征然后进行解码，可以获得最终的材料。

潜在编码器和四个材料解码器的结构与先验模型相同。引导编码器由九个卷积块组成，其通道数依次为 64, 64, 64, 128, 128, 128, 256, 256 和 256。这个卷积块的结构与先验模型中的卷积块结构一致。

为了在模型训练过程中计算渲染损失，我们采用了一种基于物理的渲染着色器。这个着色器可以有效地模拟光线的物理行为。它是在渲染方程的基础上实现的，该方程考虑了光线在表面上的相互作用和反射。我们通过以下公式计算表面反射：其中， $L$  代表渲染结果， $x$  是几何表面上的一个点， $L_i$  是光强。 $materials$  包含法线 ( $n$ )、漫反射反照率 ( $d$ )、粗糙度 ( $r$ )、镜面反照率 ( $s$ )、次表面散射 ( $sss$ ) 和位移 ( $disp$ )。需要注意的是，位移表示当前相机下表面点  $x$  到世界中心的距离。 $w_i$  和  $w_o$  分别代表光线的入射方向和出射方向。 $f_r$  是 BSDF，它定量描述表面

反射率如何随着入射光的不同方向变化，我们采用 Disney BSDF 模型 [43] 来进行计算。

#### B. 损失函数

我们的模型由三个先验模型和一个微调模型组成。在训练过程中，这些模型的损失主要包括基于像素的 L1 损失和渲染损失。由于每个模型的任务不同，我们根据估计的对象为每个模型设计了一种 Controlled PBR Rendering (CPR) 损失。因此，我们的方法是在一个联合损失函数  $\mathcal{L}_{total}$  上进行训练的，其可以通过以下公式计算：

$$\mathcal{L}_{total} = \underbrace{\mathcal{L}_p^g + \mathcal{L}_{cpr}^g}_{Geometry} + \underbrace{\mathcal{L}_p^a + \mathcal{L}_{cpr}^a}_{Albedo} + \underbrace{\mathcal{L}_r^r + \mathcal{L}_r^t}_{RSS} + \underbrace{\mathcal{L}_p^m + \mathcal{L}_r^m}_{Finetune}, \quad (1)$$

其中， $\mathcal{L}_p$  表示基于像素的 L1 损失， $\mathcal{L}_{cpr}$  表示 CPR 损失，而  $\mathcal{L}_r$  表示重光照损失。

像素级的 L1 损失  $\mathcal{L}_p$  可以通过以下公式计算：

$$\mathcal{L}_p = \sum_i^{all} (\|pred - gt\|_1), \quad (2)$$

其中  $all$  取决于每个模型的估计目标，例如，GPM 的目标是法线和位移，而 APM 的目标是漫反射反照率。

单张图像的材质估计是一个病态问题，验证材质准确性的有效方法是计算在各种光照条件下的渲染损失。得益于我们的材质数据，我们可以在训练过程中生成在任意光照下渲染的 GT 和估计材质。渲染所用的光照越多样，材质估计就越稳健。

因此，我们设计了一种多照明渲染损失用于训练。它包含  $M$  (设置为 37) 种用于 PBR 渲染的照明，其中 36 种为固定且不同位置的点光源 (如图 3 (b) 所示)，1 种为随机位置的点光源 (如图 3 (c) 所示)，具有随机光强 (3.0-8.0)。在图 3 (b) 中，在  $XOZ$  和  $YOZ$  平面内，相邻点光源之间的角间隔为  $10^\circ$ 。这个角度间隔足以在大多数场景中准确描绘照明方向。基于 GT 和估计材质，我们在这些照明下渲染外观并计算损失，多照明渲染损失 ( $\mathcal{L}_r$ ) 按以下公式计算：

$$\mathcal{L}_r = \sum_i^M \|(\mathcal{R}(gt, light_i), \mathcal{R}(pred, light_i))\|_1, \quad (3)$$

其中  $pred$  为估计材质， $gt$  为 GT 材质， $\mathcal{R}$  为 PBR 着色器， $light_i$  为渲染用的照明。

损失函数的目的是确保每种材料特性能够独立优化，同时保持物理合理性。关键思想是固定不相关的参数 (可以使用 GT 数据或将其设置为物理上合理的值以尽量减少干扰)，从而使待优化材质在渲染过程中发挥最显著的作用。此外，在设置不相关参数时，应考虑材料的真实性，并避免设置极端值，以降低模型的泛化能力。

具体而言，(1) 对于 Geometry Prior Model，我们需要使渲染结果中的高光更容易出现。这可以增强法线和位移对渲染结果细节的影响。因此，我们将粗糙度设置为 0.2，并设置中等强度的镜面反射反照率 (0.5)。(2) 对于 Albedo Prior Model，减少高光的出现可以增强漫反射反照率对渲染结果的重要性。因此，我们将粗糙度设置为 0.8，并将镜面反照率设置为 0.03，尽量减少镜面反射在物理合理范围内的影响。使用次表面散射和法线的 GT 数据消除几何干扰。(3) 对于 RSS Prior Model，提升性能的关键是

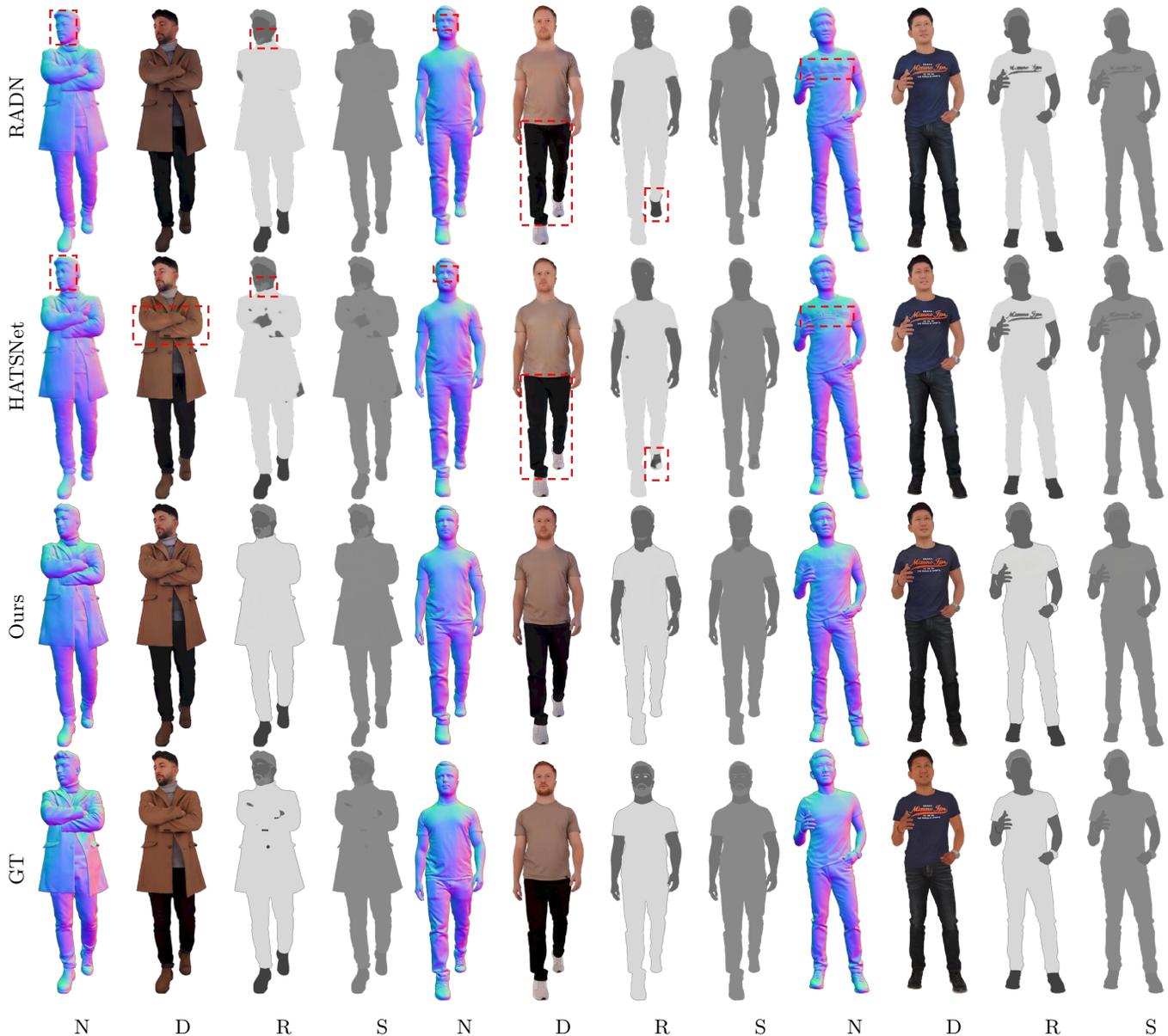


Fig. 4. 在 OpenHumanBRDF 数据集上的 PBR 材质估计中，与之前工作的性能对比 (RADN [13] 和 HATSNet [14])。图 ?? 中展示了“输入”、估计的“Disp”和“SSS”以及“我们的方法”。“N”表示法线，“D”表示漫反射反射率，“R”表示粗糙度，“S”表示镜面反射率。重光照的结果如图 5 所示。更多细节请参见第 ?? 节。

增强高光对渲染结果的影响。基于此，我们认为直接使用 GT 数据（法线、位移和漫反射反射率）可以在高光中提供多样性，是最佳选择。这样， $\mathcal{L}_{cpr}$  可以通过

$$\mathcal{L}_{cpr} = \sum_i^M \|(\mathcal{R}(pred, cm', light_i) - \mathcal{R}(gt, cm', light_i))\|_1, \quad (4)$$

定义，其中  $\mathcal{R}$  是 PBR 着色器， $M$ （设为 37）意味着 PBR 渲染的照明种类， $light_i$  是用于渲染的照明，而  $cm'$  则表示受控材料。需要注意的是，HDR 环境图在光源的数量、强度和颜色方面比点光更为复杂。为了避免环境图对渲染结果造成超过预估材料的影响，我们仅计算在  $M$  照明下的渲染损失。 $M - 1$  是在固定位置下的点光源，其中一个是在随机位置下的点光源，具有随机光强度。

在微调模型训练期间，理论上需要加载先验模型，并将预测的材料作为材料模型的输入。然而，由于 GPU 的图形内存有限，我们无法在训练时同时加载先验模型和材料模型。因此，为了使训练可行，在训练材料模型时，我们直接使用 GT 材料作为输入。在测试时，先验模型预测的材料被用作材料模型的输入。

三个先验模型和一个材料模型是独立训练的。所有这些模型都是在单个具有 24GB 内存的 NVIDIA RTX 3090 GPU 上以批量大小为 1 训练最多 100 个 epoch。完成这样的训练过程对于每个先验模型大约需要两天，对于 Finetuning Model 需要三天。我们使用 TensorFlow [44] 实现我们的方法。我们采用 AdamW 优化器进行损失优化，学习率设置为  $1e-4$ 。

我们首先进行定性和定量实验，以验证我们的方法在 OpenHumanBRDF 数据集和真实数据上的性能。然后，



Fig. 5. 估计材料的重新照明结果。估计的材料图来自图 4（左边的样本来自 OpenHumanBRDF 数据集）和图 ??（右边的样本是实际数据上的结果）。请参见附加视频中的动态光照下的渲染结果。

我们进行消融研究，以验证训练策略和各个组件的有效性。

为了验证我们方法的有效性，我们从两个方面与相关工作进行比较：PBR 材质估计和重新光照。

**PBR 材料估计。** RADN [13] 和 HATNet [14] 从近似平面图像中估计材料，两种方法都可以估计高质量的 PBR 材料，包括法线，漫反射反照率，粗糙度和镜面反照率。我们从头开始在 OpenHumanBRDF 上训练它们。SL [5] 的目标与我们的工作类似，并通过将物理引导的架构与一个预训练框架结合起来执行人类再光照。通过自监督的预训练策略，SL 可以估计 PBR 材料，包括法线，漫反射反照率，粗糙度和镜面反照率。尽管作者没有发布他们的数据集和代码，他们提供了一个用于推理的网站。为了公平比较，我们只在真实数据上将我们的方法与 SL 进行比较。

**重光。**我们比较了重光性能，使用了 FBHR [2] 和 TR [1]。FBHR 将人体外观分解为反射率和依赖于光的残余项，即便它只估算单一材质（漫反射反照率），也实现了令人印象深刻的重光效果。TR 在深度学习框架中引入了每像素光照表示，并明确建模了外观的漫反射和镜面反射组件，产出令人信服的重光肖像。基于我们的数据集，我们可以生成 FBHR 和 TR 所需的训练数据，并且我们从头开始训练了它们。

为了定量评估我们的方法和对比方法的材料估计性能，我们计算了模型估计的材料图与 GT 数据之间的 PSNR（表 II）。此外，我们还计算了在各种照明条件下渲染结

果的 PSNR，例如点光源、真实环境图和合成环境图。

为了展示我们方法（HumanMaterial）的有效性，我们在 OpenHumanBRDF 数据集和真实数据上进行了对比实验。用于材料估计的输入图像如图 ?? 所示。

如表 II 所示，我们的方法在“材质”和“重新照明”两方面几乎都达到了最佳和具有竞争力的性能。其主要原因是我们方法的渐进训练策略可以在优化过程中减少各种材质的干扰，从而更容易实现模型拟合。材质估计越准确，重新照明结果的性能就越高。基于神经着色器方法的在真实光照下的重新照明方法（FBHR 和 TR）相比基于 PBR 方法（RADN 和 HATNet）的“重新照明”性能更好，但在合成光照下没有重现这一结果。我们的结果在“真实”环境下的“重新照明”性能具有与神经着色器方法竞争的优势，但在“合成”环境下，我们的方法表现更优。这是因为神经着色器方法依赖训练数据中的光照种类，重新照明效果容易过拟合。当在训练数据中没有新的或相似的光照进行重新照明时，这些方法在实现准确的重新照明上会遇到困难。基于 PBR 的方法可以估计出适合渲染的材质，并能在任意光照下进行稳健的渲染。

如图 4 和图 ?? 所示，得益于逐步训练策略，与对比方法相比，我们估计的法线包含更多细节，漫反射反照率的色调也更加准确。SL 没有在我们的数据集上训练，阴影或高光被烘焙在估计的漫反射反照率中（图 ??），而在我们数据集上训练的其他方法的烘焙阴影更少。这证明我们



Fig. 6. 与之前工作的性能比较 (FBHR [2], TR [1]) 在 OpenHumanBRDF 数据集和真实数据上的重光照。左边三个是 OpenHumanBRDF 数据集上的结果, 右边三个是真实数据上的结果。更多细节见第 ?? 节。

的数据集接近真实场景, 并且估计出更高质量的漫反射反照率是有效的。此外, 我们展示了图 5 中材料的重新照明结果。与其他方法相比, 我们的方法可以估计次表面散射 (图 ??), 这有助于在皮肤区域产生逼真的高光。

为了评估神经着色方法的重新照明性能, 我们在 OpenHumanBRDF 数据集和真实数据上与 FBHR 和 TR 进行了对比实验 (图 6)。当新的光照条件复杂时, 基于神经渲染的方法难以实现准确的重新照明结果。复杂光照场景主要表现为: (1) 输入图像的光照与新光照不同; (2) 环境图中光源之间的差异 (颜色或强度) 很大。在复杂的新光照条件下, FBHR 和 TR 产生了一些伪影, 并且难以实现逼真的重新照明结果。由于估算出的准确可渲染的 PBR 材质贴图, 我们的方法在复杂光照下实现了准确的重新照明, 并产生了更逼真的皮肤反射效果。

我们在 OpenHumanBRDF 数据集和真实数据上进行了广泛的消融实验, 以验证我们的训练策略和各个组件的有效性。具体来说, 我们展示了微调、CPR 损失、引导特征融合、多光源渲染、次表面散射效果以及位移效果的影响, 并进行了六种不同设置的实验: (1) “w/o Guid” 表示微调模型不使用指导图进行解码, 而是直接使用 Finetuning Model 来估计最终的材质图。(2) “w/o P-Guid” 表示微调模型使用输入图像作为指导图, 并直接使用 Finetuning Model 来估计最终的材质图。(3) “w/o Opt-C” 表示在先验模型训练过程中不使用 CPR 损失, 并直接通过先验模

型估计材质结果。(4) “w/o Opt” 表示直接使用先验模型的材质估计结果而不进行微调。(5) “w/o SSS” 表示在重光照过程中不使用次表面散射。(6) “w/o Disp” 表示在重光照过程中不使用位移, 也就是说将表面几何形状近似为平面。

如表 II 和图 7 所示, 通过仔细观察, 我们发现完整模型 (“Full”) 的结果最接近 GT。基于设计良好的结构, “w/o Guid.” 取得了令人印象深刻的表现。当未使用先验模型的初始材料 (“w/o P-Guid.”) 时, 我们发现直接使用输入图像指导解码过程也提高了性能。这就是为什么先验模型使用引导特征融合操作的原因。基于上述结论, “w/o Opt.” 和 “Full” 的结果表明微调操作可以进一步提升材料估计的性能。

在计算 GPM 和 APM 的渲染损失时, 使用 CPR 的 “w/o Opt” 在 “N”, “Disp” 和 “D” 方面比 “w/o Opt-C” 具有更高的准确性。这表明 CPR 减少了非优化材料对渲染结果的干扰, 从而增强了材料估计的鲁棒性。

次表面散射可以使皮肤的反射更加真实, 而位移能够正确地表示几何表面, 有助于实现细节渲染。因此, 如表 II 和图 ?? 所示, “Full” 的重新照明性能优于 “w/o SSS” 和 “w/o Disp”, 并且皮肤区域更加真实。

TABLE II

材料和渲染结果误差 (以 PSNR 衡量) 在 OPENHUMANBRDF 数据集上的对比, 涉及之前的工作 (RADN [13], HATSNET [14], FBHR [2], 和 TR [1])。"N" 代表法线, "D" 代表漫反射反照率, "R" 代表粗糙度, "S" 代表镜面反照率, "SSS" 代表次表面散射。"Disp" 代表位移。"POINT" 表示在如图 3 所示的固定单点光源下的再光照。"REAL" 表示在四个真实 HDR 环境图下的再光照。"SYN" 表示在合成 HDR 环境图下的再光照。"RELIT" 表示再光照结果。最佳得分以粗体显示, 第二佳得分以下划线表示。

Method	Material							Relighting				Mean			
	N	D	R	S	SSS	Disp	Point	Real				Syn	Material	Relit.	Total
								Real_1	Real_2	Real_3	Real_4				
RADN	20.3	26.0	22.4	38.1	/	/	21.4	21.5	21.9	21.7	21.8	22.6	26.7	21.9	24.3
HATSNET	<u>20.5</u>	<u>26.2</u>	<u>22.9</u>	38.4	/	/	<u>21.9</u>	21.7	21.9	21.8	22.0	<u>22.8</u>	<u>27.0</u>	22.0	<u>24.5</u>
FBHR	/	25.7	/	/	/	/	/	21.9	22.7	<u>22.4</u>	<u>22.1</u>	20.5	25.7	21.9	23.8
TR	19.5	25.9	/	/	/	/	/	<u>22.0</u>	23.1	22.6	22.3	20.7	22.7	<u>22.1</u>	22.4
Ours	21.2	27.1	24.1	39.5	41.6	28.8	23.2	22.4	<u>22.9</u>	<u>22.5</u>	22.3	24.5	30.4	22.9	26.7

### C. 应用

HumanMaterial 能够从单个全身图像中有效估计出 PBR 材质。这些估计出的材质可直接用于 Blender 渲染, 并可应用于多种应用场景, 如重新照明和材质编辑。

材质编辑。如图 8 所示, 我们进行了一个类别级别的外观编辑。由于我们的方法可以通过分类估计 PBR 材料, 我们可以从估计得到的粗糙度、镜面反射反照率或次表面散射中提取头发、皮肤、织物和皮革的蒙版。然后, 我们可以轻松地编辑这些材质, 例如将衣物从织物改为闪亮的皮革, 改变衣物和头发的颜色。我们在附带的视频中提供了编辑演示。

### D. 局限性

我们的模型在处理强光照的样本时表现出困难。正如图 9 所示, 高光和阴影可能会融入到漫反射反照率中。造成这种情况的原因可能是输入是一张单独的图像, 只有单一视角和单一类型的照明。这使得分离材质和照明变得极其困难。一个稳健的光感知模型或对输入图像的处理 (如去除高光和阴影) 是解决这一问题的可行方法。真实世界中与人体相关的材料相当复杂, 我们的数据集考虑了四种最相关的材料, 这限制了模型的鲁棒性, 并使得难以处理复合材料, 例如有灰尘的织物。

我们提出了 HumanMaterial, 这是一种从单张图像进行全身 PBR 材料估计的方法。我们的方法主要包括以下核心点: (a) 构建一个高质量的人体 PBR 材料数据集, 该数据集具有逼真的皮肤表现。(b) 提出一个设计良好的模型 (HumanMaterial), 采用渐进训练策略, 包括两个阶段, 从先验模型中获得初始材料, 通过联合优化微调以获得最终材料。(c) 基于设计良好的损失, 我们在 OpenHumanBRDF 数据集上训练模型 (HumanMaterial), 以估计可渲染的 PBR 材料, 并使用 Blender 实现逼真的重新照明和材质编辑。在 OpenHumanBRDF 数据集和真实数据上的大量实验表明了我们方法的优越性能。

未来工作。由于模型参数和计算资源的限制, 我们估计材质的 PBR 分辨率为  $512 \times 512$ 。未来, 我们将探索构建一个更高效的网络, 并设计一个有效的策略以生成更高分辨率的结果。

### REFERENCES

- [1] R. Pandey, S. O. Escolano, C. Legendre, C. Haene, S. Bouaziz, C. Rhemann, P. Debevec, and S. Fanello, "Total relighting: Learning to relight portraits for background replacement," vol. 40, no. 4, August 2021.
- [2] M. Lagunas, X. Sun, J. Yang, R. Villegas, J. Zhang, Z. Shu, B. Masia, and D. Gutierrez, "Single-image full-body human relighting," in *Eurographics Symposium on Rendering (EGSR)*. The Eurographics Association, 2021.
- [3] D. Tajima, Y. Kanamori, and Y. Endo, "All-frequency full-body human image relighting," *arXiv preprint arXiv:2411.00356*, 2024.
- [4] L. Zhang, A. Rao, and M. Agrawala, "Scaling in-the-wild training for diffusion-based illumination harmonization and editing by imposing consistent light transport," in *The Thirteenth International Conference on Learning Representations*, 2025.
- [5] H. Kim, M. Jang, W. Yoon, J. Lee, D. Na, S. Woo, and B. Ai, "Switchlight: Co-design of physics-driven architecture and pre-training framework for human portrait relighting," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 25 096–25 106, Jan 2024.
- [6] T. Akenine-Mo and N. Hoffman, "Real-time rendering," 2018.
- [7] Blender Online Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. [Online]. Available: <http://www.blender.org>
- [8] T. Alldieck, M. Zanfir, and C. Sminchisescu, "Photorealistic monocular 3d reconstruction of humans wearing clothing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.
- [9] Y. Liu, Y. Li, S. You, and F. Lu, "Unsupervised learning for intrinsic image decomposition from a single image," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [10] T. Sun, J. T. Barron, Y.-T. Tsai, Z. Xu, X. Yu, G. Fyffe, C. Rhemann, J. Busch, P. Debevec, and R. Ramamoorthi, "Single image portrait relighting," *ACM Transactions on Graphics (TOG)*, p. 1–12, Aug 2019.
- [11] Z. Li, Z. Xu, R. Ramamoorthi, K. Sunkavalli, and M. Chandraker, "Learning to reconstruct shape and spatially-varying reflectance from a single image," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, dec 2018.
- [12] Y. Liu, P. Wang, C. Lin, X. Long, J. Wang, L. Liu, T. Komura, and W. Wang, "Nero: Neural geometry and brdf reconstruction of reflective objects from multiview images," *ACM Transactions on Graphics (TOG)*, vol. 42, no. 4, jul 2023.
- [13] V. Deschaintre, M. Aittala, F. Durand, G. Drettakis, and A. Bousseau, "Single-image svbrdf capture with a rendering-aware deep network," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, jul 2018.
- [14] J. Guo, S. Lai, C. Tao, Y. Cai, L. Wang, Y. Guo, and L.-Q. Yan, "Highlight-aware two-stream network for single-image svbrdf acquisition," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, jul 2021.
- [15] K. Zhang, F. Luan, Q. Wang, K. Bala, and N. Snavely, "PhysSG: Inverse rendering with spherical gaussians for physics-based material editing and relighting," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 5453–5462.

- [16] Y. Yao, J. Zhang, J. Liu, Y. Qu, T. Fang, D. McKinnon, Y. Tsin, and L. Quan, "Neilf: Neural incident light field for physically-based material estimation," in *European Conference on Computer Vision (ECCV)*, 2022.
- [17] X. Ma, Y. Yu, H. Wu, and K. Zhou, "Efficient reflectance capture with a deep gated mixture-of-experts," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 30, no. 7, pp. 4246–4256, 2024.
- [18] E. Miandji, T. Tongbuasirilai, S. Hajisharif, B. Kavoosighafi, and J. Unger, "Frost-brdf: A fast and robust optimal sampling technique for brdf acquisition," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 30, no. 7, pp. 4390–4402, 2024.
- [19] J. Li, Q. Deng, H. Ling, and B. Huang, "Dpcs: Path tracing-based differentiable projector-camera systems," *IEEE Transac-*

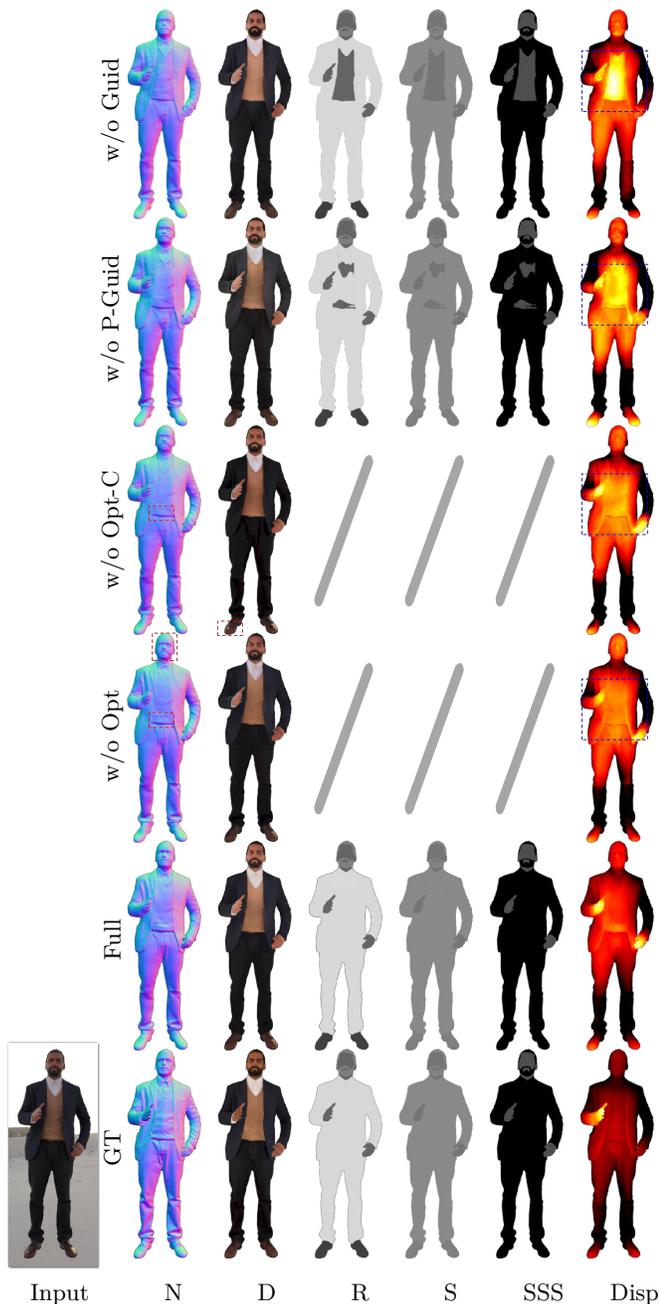


Fig. 7. 在 OpenHumanBRDF 数据集上的消融研究。“N”、“D”、“R”、“S”、“SSS”和“Disp”分别代表法线、漫反射反照率、粗糙度、镜面反照率、次表面散射和位移。更多细节见第 ?? 节。

- tions on Visualization and Computer Graphics (TVCG)*, vol. 31, no. 5, pp. 3666–3676, 2025.
- [20] W. Sun, Y. Che, H. Huang, and Y. Guo, "Neural reconstruction of relightable human model from monocular video," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 397–407.
- [21] X. Zhang, P. P. Srinivasan, B. Deng, P. Debevec, W. T. Freeman, and J. T. Barron, "Nerfactor: Neural factorization of shape and reflectance under an unknown illumination," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 6, pp. 1–18, 2021.
- [22] X. Li, Y. Dong, P. Peers, and X. Tong, "Modeling surface appearance from a single photograph using self-augmented convolutional neural networks," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, jul 2017.
- [23] X. Zhou and N. K. Kalantari, "Look-ahead training with learned reflectance loss for single-image svbrdf estimation," *ACM Transactions on Graphics (TOG)*, vol. 41, no. 6, nov 2022.
- [24] D. Gao, X. Li, Y. Dong, P. Peers, K. Xu, and X. Tong, "Deep inverse rendering for high-resolution svbrdf estimation from an arbitrary number of images," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 4, jul 2019.
- [25] Y. Guo, C. Smith, M. Hašan, K. Sunkavalli, and S. Zhao, "Materialgan: Reflectance capture using a generative svbrdf model," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 6, nov 2020.
- [26] Y. Zhang, J. Sun, X. He, H. Fu, R. Jia, and X. Zhou, "Modeling indirect illumination for inverse rendering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New York, NY, USA: Association for Computing Machinery, 2022, pp. 18 643–18 652.
- [27] F. Paraperas Papantoniou, A. Lattas, S. Moschoglou, and S. Zafeiriou, "Relightify: Relightable 3d faces from a single image via diffusion models," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.
- [28] H. Jin, I. Liu, P. Xu, X. Zhang, S. Han, S. Bi, X. Zhou, Z. Xu, and H. Su, "Tensor: Tensorial inverse rendering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.
- [29] A. Chen, Z. Xu, A. Geiger, J. Yu, and H. Su, "Tensorf: Tensorial radiance fields," in *European Conference on Computer Vision (ECCV)*, 2022.
- [30] M. Boss, R. Braun, V. Jampani, J. T. Barron, C. Liu, and H. P. Lensch, "Nerf: Neural reflectance decomposition from image collections," in *IEEE International Conference on Computer Vision (ICCV)*, 2021.
- [31] J. Munkberg, J. Hasselgren, T. Shen, J. Gao, W. Chen,

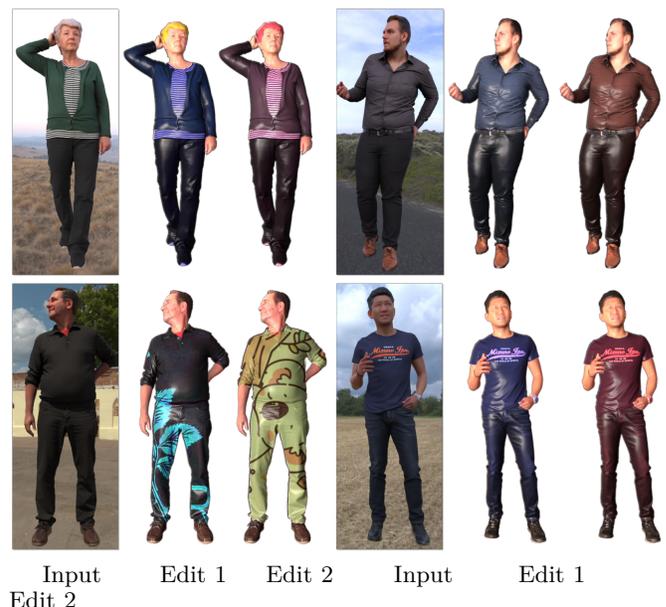


Fig. 8. 材料编辑。请查看附带视频中的动态编辑过程。



Fig. 9. 限制样本。“N”表示正常。“D”表示漫反射率。“Disp”表示位移。“Relit”表示在新环境贴图下面的重照明外观。“SSR”表示次表面散射、镜面反射率和粗糙度的结果，这些在通道轴下连接起来。

- A. Evans, T. Müller, and S. Fidler, “Extracting Triangular 3D Models, Materials, and Lighting From Images,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 8280–8290.
- [32] Y. Fan, I. Skorokhodov, O. Voynov, S. Ignatyev, E. Burnaev, P. Wonka, and Y. Wang, “Factored-neus: Reconstructing surfaces, illumination, and materials of possibly glossy objects,” *arXiv preprint arXiv:2305.17929*, 2023.
- [33] A. Lattas, S. Moschoglou, S. Ploumpis, B. Gecer, A. Ghosh, and S. Zafeiriou, “Avatarme++: Facial shape and brdf inference with photorealistic rendering-aware gans,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 9269–9284, 2021.
- [34] A. Lattas, S. Moschoglou, B. Gecer, S. Ploumpis, V. Triantafyllou, A. Ghosh, and S. Zafeiriou, “Avatarme: Realistically renderable 3d facial reconstruction in-the-wild,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 760–769.
- [35] Z. Shu, S. Hadap, E. Shechtman, K. Sunkavalli, S. Paris, and D. Samaras, “Portrait lighting transfer using a mass transport approach,” *ACM Transactions on Graphics (TOG)*, p. 1, Jul 2017.
- [36] Z. Li, L. Song, C. Liu, J. Yuan, and Y. Xu, “Relit-neulf: Efficient novel view synthesis with neural 4d light field,” in *Proceedings of the 31st ACM International Conference on Multimedia*, 2023.
- [37] D. Azinović, O. Maury, C. Hery, M. Nießner, and J. Thies, “High-res facial appearance capture from polarized smartphone images,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2023.
- [38] Z. Chen and Z. Liu, “Relighting4d: Neural relightable human from videos,” in *European Conference on Computer Vision*. Springer, 2022, pp. 606–623.
- [39] Z. Xu, S. Peng, C. Geng, L. Mou, Z. Yan, J. Sun, H. Bao, and X. Zhou, “Relightable and animatable neural avatar from sparse-view video,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 990–1000.
- [40] PolyHaven, “HDR environment images,” <https://polyhaven.com/hdris>.
- [41] RenderPeople Team, “Renderpeople,” <https://renderpeople.com>, 2021.

- [42] D. Gatis, 2020. [Online]. Available: <https://github.com/danielgatis/rembg>
- [43] B. Burley and W. D. A. Studios, “Physically-based shading at disney,” in *Acm Siggraph*, vol. 2012. vol. 2012, 2012, pp. 1–7.
- [44] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, “Tensorflow: a system for large-scale machine learning,” ser. OSDI’16. USA: USENIX Association, 2016, p. 265–283.