

# Highlight-Aware Two-Stream Network for Single-Image SVBRDF Acquisition<sup>[1]</sup>复现报告

刘运衡

## 摘要

本报告以复现面向单张图像输入的高光感知双流网络为主题，对所复现方法的相关领域研究进展、基本方法等进行了简述，并对该方法的主要贡献及模型设计进行了说明。随后说明了复现的内容、创新点与实现的复现效果，最后对本次复现的不足与需改进之处进行了阐述。

**关键词：**材质获取；可微渲染；计算机图形学

## 1 引言

由真实拍摄的二维图像重建物体表面的反射特性（Spatially-Varying Bidirectional Reflectance Distribution Function, SVBRDF）是一个具有一定重要性的任务，并有着较为广泛的应用价值。相关方法能够应用于产品设计、视觉效果制作、虚拟现实设计及文化遗产修复等多种场景。然而，不同的材料参数的组合可能导致相同的辐射度观测值说明该任务具有病态性，并为相关方法的设计提出了挑战。部分早期的方法需要使用昂贵的硬件设备来对 SVBRDF 的参数空间进行详尽的采样，但繁琐和艰巨的过程往往限制了其适用性。本次复现拟通过对<sup>[1]</sup>论文所提出的方法进行复现，实现一个基于深度神经网络的新型单图像 SVBRDF 重建模型。通过该模型降低对 SVBRDF 重建过程的限制，使得非专业的用户也能够简单地获取较高质量的 SVBRDF，从而为相关的图形学应用提供便利。

## 2 相关工作

近年来有多篇论文尝试使用深度学习方法完成 SVBRDF 重建任务，这些方法都致力于重建出更为准确、真实的 SVBRDF 材质，但是侧重点各异。从方法原理上大致可以分为直接回归的方式和基于先验进行优化的方式两类。

### 2.1 直接回归的方法

直接回归的方法整体框架通常较为清晰，即输入参考图像，通过自编码器等模型输出材质贴图，并在数据集上进行端到端训练。本报告所复现的论文也属于这一类别。这类方法的优点在于模型训练、测试较为方便且通常不需要额外标定相机和光照参数。

<sup>[2]</sup>使用 U-Net 自编码器结构将单张参考图像映射为材质贴图，并加入与自编码器并行的全局特征分支，使用求和或特征拼接与相应的自编码器层互相交换特征以增强模型对全局特征的把握。该文章创新地提出了渲染损失：通过比较输出材质贴图和其真实值在多种光照下的渲染结果引导模型训练，提升了输出材质贴图的视觉效果，该渲染损失也被后续方法广泛应用。

### 2.2 基于优化的方法

基于优化的方法常需要学习一个能表达材质贴图特性的模型作为先验，然后利用该模型对输出图像或隐空间中与输出对应的隐向量使用反向梯度传播进行更新、迭代优化，使得输出能在满足先验的同时尽量接近输入的参考图像。

<sup>[3]</sup>提出训练一个将材质贴图压缩到低维隐空间再进行重建的自编码器，以该自编码器的隐空间作为先验用于输出材质贴图的优化，并进一步通过引入额外的平滑损失函数使隐空间满足一定的正则化性质。在优化时以单图像 SVBRDF 获取模型<sup>[2]</sup>的输出作为初始化进行隐编码优化，最后使用了逐像素的反向传播优化作为后处理来提升细节效果。<sup>[4]</sup>使用 styleGAN2<sup>[5]</sup> 模型代替<sup>[3]</sup>中的自编码器进行先验的学习。通过最小化渲染损失的方式反向传播梯度以交替的方式优化风格隐向量和随机隐向量，获取该材质在 styleGAN2 隐空间中的位置，从而得到符合模型表达的先验的材质贴图。

### 3 本文方法

本文复现的方法使用单张图像作为输入，通过模型直接输出材质贴图，不需要标定相机及场景参数或进行额外的优化或后处理。同时，提出的高光感知（Highlight-Aware, HA）卷积模块及双流模型架构能够更好地处理高光部分信息缺失造成输出材质贴图在高光部分产生瑕疵的问题，从而达到较好的 SVBRDF 材质重建效果。以下对该论文方法的主要创新点进行简要说明。

#### 3.1 HA 卷积模块

该论文受图像修复领域中部分方法的思路启发，将高光区域的信息恢复与图像修复任务类比，提出了高光感知卷积模块。类似图像修复领域中的门控卷积（Gated Convolution）<sup>[6]</sup>将图像缺失区域与剩余区域分开进行特征提取并逐步更新缺失区域的遮罩，HA 卷积也将高光区域与曝光正常区域分开提取，但区别于门控卷积将已知的遮罩作为初始化：HA 卷积通过一个卷积层自动地学习得到图像中的高光部分，使得 HA 卷积更能适应图像中比例和颜色各异的高光区域，其结构如图 1所示。

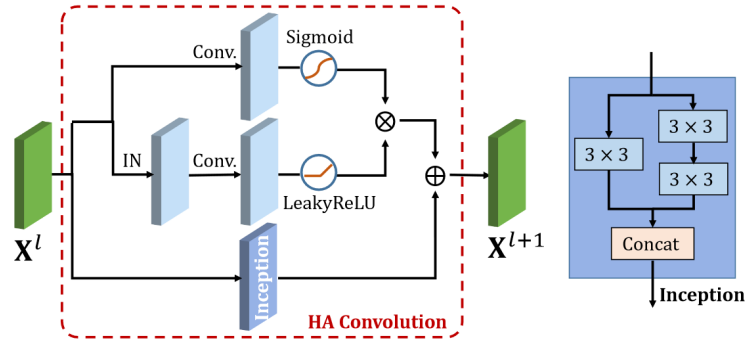


图 1: HA 卷积结构

#### 3.2 模型架构及损失函数设计

模型整体架构如图 2所示。在特征提取方面，该论文设计了分别使用 HA 卷积和普通卷积的双分支网络结构来进行特征提取以获取更为丰富的特征；在特征聚合及材质贴图重建方面，区别于大部分方法使用单个生成器生成多个通道作为不同材质贴图，本文对不同材质贴图分别使用了不同的生成器，并引入了基于注意力的特征选择模块（Attention-based feature selection, AFS）对不同材质贴图所需的特征进行动态选择提升了特征融合的效果。

对于损失函数，该论文在以往重建损失和重渲染损失的基础上，参考图像修复的经典方法增加了局部、全局对抗损失<sup>[7]</sup>。通过局部判别器确保修复的高光区域视觉上的真实性，再通过全局判别器确保修复的部分和图像整体的一致性，提升了输出材质贴图高频细节的质量。

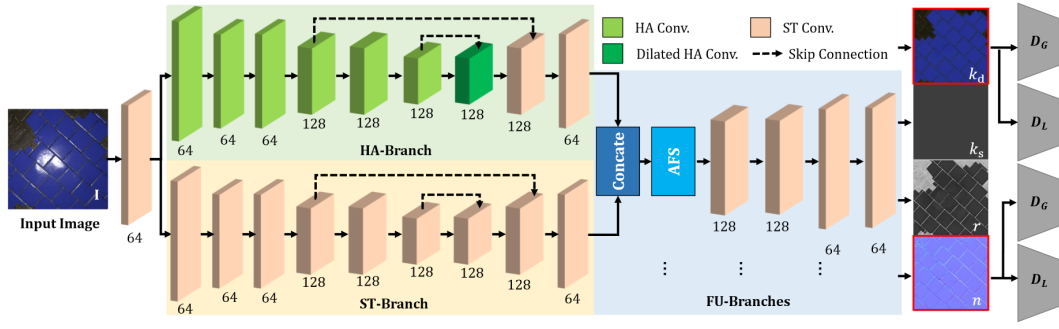


图 2: 模型整体架构

## 4 复现细节

### 4.1 与已有开源代码对比

本次复现论文有基于 tensorflow 框架的网络结构开源代码<sup>[8]</sup>，非官方实现有基于 tensorflow 的完整实现<sup>[9]</sup>，及基于 pytorch 框架的实现<sup>[10]</sup>但缺失部分模块及训练、测试代码。复现项目开始时现有开源项目均暂未提供基于 pytorch 框架的完整实现，也未包含预训练参数。

本次复现基于 pytorch 框架，主要复现内容包括，实现基于 Cook-Torrance BRDF 模型的可微渲染并用于计算损失函数；实现模型训练、测试及可视化等。在本复现项目中数据处理及模型结构均参考了<sup>[10]</sup>，对于渲染损失的实现则参考了<sup>[9]</sup>。

### 4.2 使用说明

本次复现实现了模型的训练、测试及对重渲染结果的可视化。通过运行 main.py 文件可以启动模型训练，部分训练参数可通过 main.py 开头定义的变量进行修改，支持使用 DataParallel 及 DistributedDataParallel 进行多卡训练。test.py 则提供了测试模型的代码实现，使用与修改方式与训练类似。visualize.py 提供了使用复现的可微渲染模块在随机光照条件下渲染图像并保存至指定目录的功能，使用时需将待渲染的材质贴图整理为与数据集相同的格式。此渲染仅用作模型训练和调试，在视觉效果可能较为粗糙。

### 4.3 创新点

本次复现在原方法的基础上增加了能够调整材质贴图通道数的功能。具体而言，法线贴图可以通过将三维法线分量转为二维的角度值减少通道数，粗糙度贴图也可仅用单通道图像表示，希望通过减少通道可以轻微减小模型的内存消耗并方便模型的训练。但目前实验中使用该功能并未获得较好的效果，仍有待进一步改进。

## 5 实验结果分析

该方法使用的数据集<sup>[2]</sup>体量较大，模型训练需要较强的算力及较长时间，由于个人能力有限没有对模型进行较为精细调整优化，导致最终的结果可能仍有待提升，但总体而言仍达到了一定的材质贴图重建效果。复现出的模型在虚拟生成的测试集图像上的效果如图 3 所示。可以看出模型输出的法线贴图较好地捕捉了输入图像的几何信息，对漫反射贴图、粗糙度贴图及镜面反射贴图的预测也与真实值相近。如图 4，对于使用手机随意拍摄的真实图像输入，该模型也表现出了一定的材质重建准确度。

此外，通过模型获得材质贴图后可以使用项目中复现的渲染模块或其他渲染器（如 Eevee 等）在

不同于输入的光照、视角条件下进行渲染。在部分真实数据上的重渲染效果如图 5所示，渲染图像具备一定的真实感说明了复现模型的有效性。

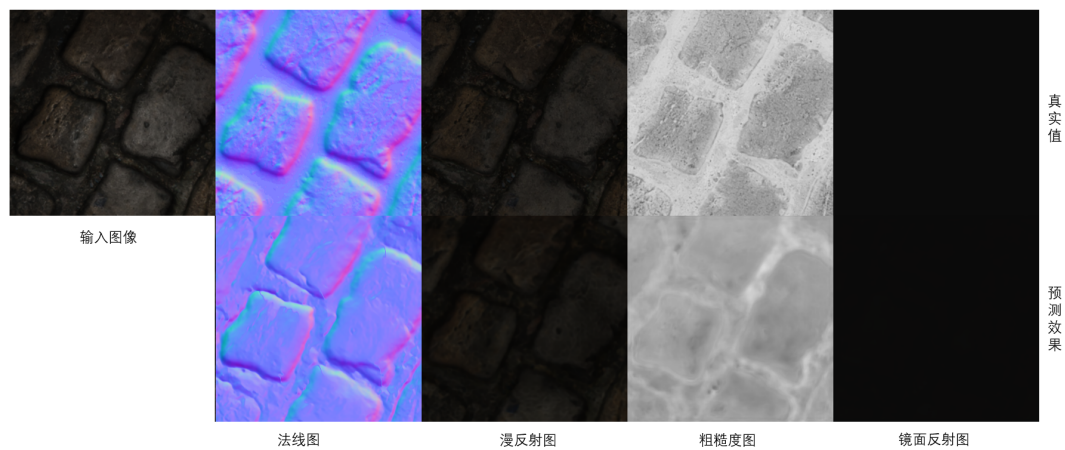


图 3: 测试集虚拟图像实验结果示意

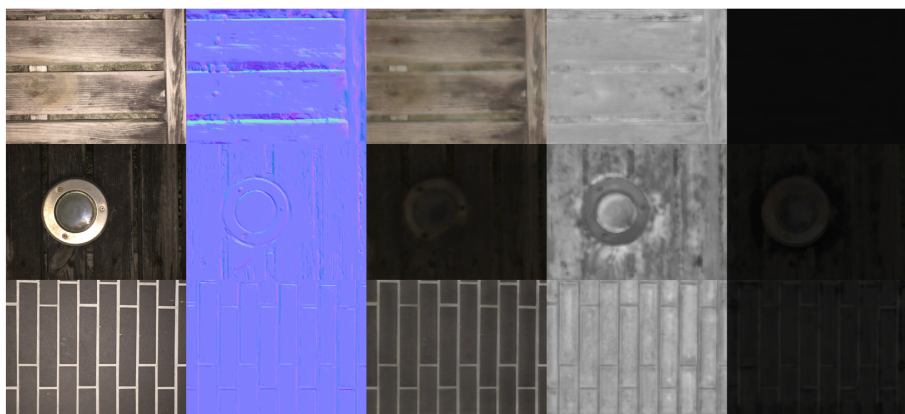


图 4: 真实图像上的实验结果示意



图 5: 重渲染结果

## 6 总结与展望

通过本次复现了解了部分基于深度学习的 SVBRDF 获取方法，并一定程度提高了个人编程复现能力。本次复现的代码基于 pytorch 框架，希望能够为相关领域研究提供便利。

作为缺陷，当前版本在代码的可扩展性上还有待加强，由于直接使用 pytorch 提供的接口实现对多 GPU 的支持导致代码量增加，可读性和可维护性下降。后续考虑使用 pytorch-lightning 等框架进一步封装模型，简化训练及测试代码。此外，当前复现结果由于个人能力及硬件条件有限未达到最佳的训练效果，需要考虑通过进一步的模型修正及参数调整提升效果。

## 参考文献

- [1] GUO J, LAI S, TAO C, et al. Highlight-aware two-stream network for single-image SVBRDF acquisition [J/OL]. ACM Transactions on Graphics, 2021, 40(4): 1-14 [2022-10-11]. <https://dl.acm.org/doi/10.1145/3450626.3459854>. DOI: 10.1145/3450626.3459854.
- [2] DESCHAINTE V, AITTALA M, DURAND F, et al. Single-image svbrdf capture with a rendering-aware deep network[J]. ACM Transactions on Graphics (ToG), 2018, 37(4): 1-15.
- [3] GAO D, LI X, DONG Y, et al. Deep inverse rendering for high-resolution SVBRDF estimation from an arbitrary number of images.[J]. ACM Trans. Graph., 2019, 38(4): 134-1.
- [4] GUO Y, SMITH C, HAŠAN M, et al. Materialgan: reflectance capture using a generative svbrdf model [J]. arXiv preprint arXiv:2010.00114, 2020.
- [5] 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. 2020: 8110-8119.
- [6] YU J, LIN Z, YANG J, et al. Free-form image inpainting with gated convolution[C] // Proceedings of the IEEE/CVF international conference on computer vision. 2019: 4471-4480.
- [7] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics (ToG), 2017, 36(4): 1-14.
- [8] Leisuzz, <https://github.com/leisuzz/Highlight-Aware-Two-Stream-Network-for-Single-Image-SVBRDF-Acquisition>.
- [9] Happyvictor008, <https://github.com/happyvictor008/Highlight-Aware-Two-Stream-Network-for-Single-Image-SVBRDF-Acquisition>.
- [10] Wenyihan4396, [https://github.com/wenyihan4396/Two-Stream-BRDF-Network\\_Reproduction](https://github.com/wenyihan4396/Two-Stream-BRDF-Network_Reproduction).