

基于逆渲染的单张图像高效材质编辑

谢坤亮¹, 易任娇¹, 周海芳¹, 朱晨阳¹, 刘豫皖², 徐凯^{1*}

¹国防科技大学计算机学院, 湖南 长沙 410005;

²东部战区联合参谋部战勤局, 江苏 南京 210000

摘要 基于图像的材质编辑作为一个图像逆渲染问题,对增强现实与交互设计领域都十分必要。提出了一种将单张图像中的物体材质编辑成一系列材质特征差距较大的新材质的方法,包含了高光分离、本征图像分解、高光编辑三个子任务。使用参数化材质模型,合成了一个大规模包含多种光照和材质光泽度属性的数据集,并利用深度卷积网络进行源材质到目标材质的转换。实验证明了三个子任务在多种定性、定量分析比较下的有效性,在合成和真实两类测试图片上展示了材质编辑效果。创新性提出的一种对二维图像高光层进行直接编辑的材质转换方法支持多种常见材质(塑料、木、石头、金属等),在合成或真实图片上可以实现高效、逼真的材质编辑结果。

关键词 机器视觉; 逆向渲染; 深度学习; 增强现实; 材质编辑

中图分类号 TP399 文献标志码 A

DOI: 10.3788/LOP202259.1415014

Efficient Material Editing of Single Image Based on Inverse Rendering

Xie Kunliang¹, Yi Renjiao¹, Zhou Haifang¹, Zhu Chenyang¹, Liu Yuwan², Xu Kai^{1*}

¹School of Computer Science, National University of Defense Technology, Changsha 410005, Hunan, China;

²Dongbu Zhanqu Zhanqinju, Nanjing 210000, Jiangsu, China

Abstract As an inverse rendering problem, image-based material editing is essential for augmented reality and interaction design. Herein, we propose a method to edit the object material in a single image and convert it into a series of new materials with widely varying material characteristics. This approach involves specular highlight separation, intrinsic image decomposition, and specular highlight editing. Using a parametric material model, we synthesized a large-scale dataset of objects under various illumination conditions and material shininess parameters. Based on this dataset, we converted the source material into the target material using deep convolutional network. Experiments illustrate that the three parts of the approach are effective on various qualitative and quantitative evaluations, showing the material editing effect on the synthetic image and real test image. This novel material-editing method based on directly editing the specular highlight layer of a single image supports a variety of materials, such as plastic, wood, stone, and metal and efficiently produces realistic results for both synthetic and real pictures.

Key words machine vision; inverse rendering; deep learning; augmented reality; material editing

1 引言

近年来,逆向渲染(即对真实世界的图像输入反推场景中的三维信息,如材质、几何、光照等问题)成为了计算机视觉和图形学研究的热点,这一课题在增强现实、虚拟现实等领域都有大量应用需求。恢复出的三维信息用来对输入图像进行场景编辑,如材质编辑、光照编辑等。其中,二维图像的材质编辑在产品设

计领域具有极大的应用价值,如何使得不了解计算机视觉、图形学的设计师也能够轻松地实现基于图像的高效材质编辑,是本文的研究目的。

较早的研究将材质编辑作为一个图像滤波问题,Khan等^[1]提出的一种简单的启发式方法估计近似的场景几何与光照信息,实现材质编辑。Boyadzhiiev等^[2]利用多个图像滤波器来改变材质性质,如光泽度等。这些方法在对物体的本征性质具有明确先验的情

收稿日期: 2021-12-24; 修回日期: 2022-02-05; 录用日期: 2022-02-21

基金项目: 国家自然科学基金青年基金(62002375, 62002376)、国家重点研发计划(2018AAA0102200)、湖南省自然科学基金重点项目(62132021)、湖南省自然科学基金青年基金(2021JJ40696)、国防科技大学科研计划(ZK19-30)

通信作者: kevin.kai.xu@gmail.com

况下能够提供较为真实的数据效果。在仅有图像输入的情况下, Liu 等^[3]和 Meka 等^[4]提出一种利用深度神经网络在合成数据上进行有监督学习的方法, 文献^[3]完成了物体间的材质迁移, 文献^[4]用于实时的材质估计任务, 然而与本文方法相比, 它们无法支持对普通用户友好的、可交互的参数化材质编辑。Shi 等^[5]将单幅图像分解成漫反射反照率层、漫反射光影层、高光层, 以支持对非朗伯物体的本征图像分解, 通过改变漫反射反照率层或者高光层可以做到一定的材质编辑效果; 然而文献^[5]只是泛泛介绍了方法对基于图像的材质编辑问题的可行性, 缺少如何对高光层材质参数进行有效控制的实验和讨论。本文在文献^[5]的基础上进一步对该方法进行分析和创新, 不仅拓展了漫反射层的材质编辑技术方案, 并且实现了对高光层材质参数的分层调节, 支持基于图像的交互式材质编辑。

所提方法针对二维图像材质编辑任务, 解决了高光分离、本征图像分解和高光编辑几个子问题, 高光编辑模块是本文的创新点之一。对于高光分离和本征图像分解, 前人已有不少相应的研究成果。首先, 高光分离问题有较长的研究历史, 早期方法基于各种先验约束来求解^[6-8]; 与人工定义先验约束相比, 近期的研究多集中在利用深度卷积网络自动从训练数据中学习先验知识^[9-10]。其次, 本征图像分解是计算机视觉的基础问题之一, 较早的方式通过人工定义先验项来解决, 先验项用来判断梯度变化是由材质反射率引起的还是由光影变化引起的^[11], 或者用来加强哑光反射率的稀疏性质^[12]; 此外也有通过输入额外数据信息的方式来提供更多的约束^[13-14]。近年来在该问题的研究上深度学习成为主流, 主要可分为基于合成数据集的有监督学习方法^[5, 15-17]和定义无监督误差后的无监督学习方法^[18-19]。

本文采取有监督学习来解决高光分离和本征图像分解两个子任务。对漫反射层和高光层分别进行处理, 利用参数化材质模型, 实现对二维图像中源材质到目标材质的高效转换。所提方法的主要贡献是: 生成了一个大规模不同光照、不同材质光泽度的高光反射数据集, 该数据集可用于其他逆向渲染任务, 如材质估计等; 提出了一种新颖的基于高光层直接转换的材质编辑方法, 避免了重渲染步骤的时间与资源需求, 实现了高效材质编辑; 为非计算机专业用户提供交互友好的用户界面, 通过简单操作即可得到多样化的材质编辑结果。

2 所提方法内容

物体在真实环境下的成像结果由物体材质、几何形状、光照条件、相机参数 4 个因素共同决定。为了高效又不失真实感地实现物体的材质编辑, 需要在保持其他场景因素不变的基础上, 使用参数化材质模型对输入图片进行材质分解和重新组合。所提方法面向二

维图像的处理过程, 在这个过程中相机参数不会发生变化, 因此需要分解出物体材质、几何形状、场景光照条件这些数据信息; 鉴于这些数据信息的高维度性和非线性等特征, 提出了一种基于卷积神经网络的单张图片材质编辑方法。方法主要流程是: 使用基于物理的成像模型, 通过高光分离网络分离出图像高光层和漫反射层; 然后利用本征图像分解网络获取图像漫反射层的材质信息, 同时使用高光编辑模块对图像高光层直接进行材质编辑; 结合参数化材质模型, 分别实现了对漫反射层和高光层材质信息的调节和转换。

2.1 成像模型

给定一张输入图像, 分解出的材质、几何、光照三方面信息是基于物理的成像模型。根据渲染方程^[20], 假设物体不是光源, 其表面上一点在观察方向的辐射度 $L(x, \mathbf{v})$ 等于该点处入射光辐射度 $L(x, \mathbf{l})$ 乘以物体材质的双向反射分布函数 (BRDF) $f(\mathbf{l}, \mathbf{v})$ 后在半球上 (以表面法线 \mathbf{n} 为中心) 的积分, 表达式为

$$L(x, \mathbf{v}) = \int_{\Omega} f(\mathbf{l}, \mathbf{v}) L(x, \mathbf{l}) \cos \theta_i d\omega_i, \quad (1)$$

式中: x 是物体表面上一点; \mathbf{v} 是观察方向; \mathbf{l} 是光线入射方向; θ_i 是 x 点处的法线 \mathbf{n} 和 \mathbf{l} 的夹角 (大于 90° 时以 90° 计算)。BRDF 与光线入射方向和相机观察方向有关, 用来反映物体表面的材质属性。首先关注 BRDF, 它分为测量式和参数式两类; 测量式 BRDF 是相应设备实际测量后得到的离散的真实材质数据; 而参数式 BRDF, 也称为参数化材质模型, 是通过定义若干参数后拟合真实材质数据的函数。目前已提出的参数式 BRDF 通常由相似的结构组成:

$$f = k_d \cdot f_d(\mathbf{l}, \mathbf{v}) + k_s \cdot f_s(\mathbf{l}, \mathbf{v}), \quad (2)$$

式中: f_d 为漫反射项; f_s 为光泽反射项; k_d 和 k_s 分别表示漫反射反照率和光泽反射反照率。联合式 (1)、(2), 可得

$$L(x, \mathbf{v}) = k_d \int_{\Omega} f_d(\mathbf{l}, \mathbf{v}) L(x, \mathbf{l}) \cos \theta_i d\omega_i + k_s \int_{\Omega} f_s(\mathbf{l}, \mathbf{v}) L(x, \mathbf{l}) \cos \theta_i d\omega_i = k_d s_d + k_s s_s, \quad (3)$$

式中: s_d 表示漫反射光影项; s_s 表示光泽反射光影项。由此, 输入图像 \mathbf{I} 可以视为漫反射层和光泽反射层的线性组合:

$$\mathbf{I} = \mathbf{A} \odot \mathbf{S} + \mathbf{R} = \mathbf{D} + \mathbf{R}, \quad (4)$$

式中: \mathbf{A} 表示漫反射反照率图, 其中每一个像素点的值代表不同的 k_d , 仅与物体的材质信息有关; \mathbf{S} 为漫反射光影图, 其中每一像素点的值为 s_d , 表示物体在对应光照条件下发生漫反射后光线和物体几何形状作用的结果; \odot 表示矩阵点乘; $\mathbf{A} \odot \mathbf{S}$ 整体记为 \mathbf{D} , 表示漫反射层 (也称为哑光反射层); \mathbf{R} 为高光反射层, 表示物体所受光照条件下发生光泽反射后, 光线、物体形状、表面材质综合作用的结果, 其中每一个像素点的值为 $k_s s_s$ 。所提材质编辑方法针对的是与材质有关的 \mathbf{A} 和 \mathbf{R} 。整体思路是: 将输入图像的高光层分离开, 分别得到漫反

射层 D 和高光层 R ; 执行本征图像分解, 将 D 分解为 A 和 S ; 然后调节材质参数得到新漫反射反照率图 A' , 再由高光编辑模块得到新的高光层 R' ; 最后输出的新材质图像 I' 等于新漫反射反照率图 A' 和新高光层 R' 与漫反射光影图 S 的重新组合,

$$I' = A' \odot S + R'. \quad (5)$$

2.2 参数化材质模型

选用的参数式材质模型是 Blinn-Phong BRDF^[21], 记为 f_b , 具体模型为

$$\begin{cases} f_b = k_d + k_s \cdot (\cos \theta_h)^\theta \\ \mathbf{h} = \frac{\mathbf{l} + \mathbf{v}}{|\mathbf{l} + \mathbf{v}|} \end{cases}, \quad (6)$$

式中: \mathbf{h} 为 \mathbf{l} 、 \mathbf{v} 之间的半角向量; θ_h 是 \mathbf{h} 和法线 \mathbf{n} 的夹角 (大于 90° 时取 90°)。该模型中的材质参数包括漫反射反照率 k_d 、光泽反射反照率 k_s 和物体表面光泽度 θ 。

由式 (4)、(6) 可知, 不同材质的漫反射层对应不同的 k_d , 保持漫反射光影图 S 不变, 源材质和新材质漫反射反照率图 A 之间的过渡是线性的, 只需调节 k_d 即可; 然而, 不同材质的高光层 R 有不同的 k_s 和 θ 属性, 光泽度 θ 的改变导致高光层的变化是非线性的。为了实现对 k_s 和 θ 的调节, 设计了一种新颖的高光编辑模块, 同

时基于该材质模型, 还合成了一个大规模包含多种光照和材质光泽度属性的数据集, 用于高光编辑模块的网络训练。

2.3 网络架构

整体网络架构如图 1 所示, 其中包含高光分离网络 (Specular-Net)、本征图像分解网络 (IID-Net) 和一个高光编辑模块 (specular layer editing module), 该模块由 10 个分层级的光泽度调节网络 (Shininess-Net) 组成。方法整体处理流程是: 1) 将图片输入高光分离网络, 获取漫反射层 (diffuse layer) 和高光层 (specular layer); 2) 将漫反射层输入本征图像分解网络, 得到漫反射反照率图和漫反射光影图, 对漫反射反照率图进行调节或者纹理替换 (提供用户界面), 得到新材质的漫反射反照率图; 3) 将高光层输入高光编辑模块, 模块中 10 个分层级的光泽度调节网络 (hierarchical Shininess-Net) 对应的光泽度调节目标如图 1 所示, 每一级光泽度都对应一个网络模型, 根据调节目标, 由其中一个模型负责输出新的高光层; 4) 将原漫反射光影图、新材质漫反射反照率图、新材质高光层按照式 (5) 重新组合 (image layers recombination), 输出最终结果。

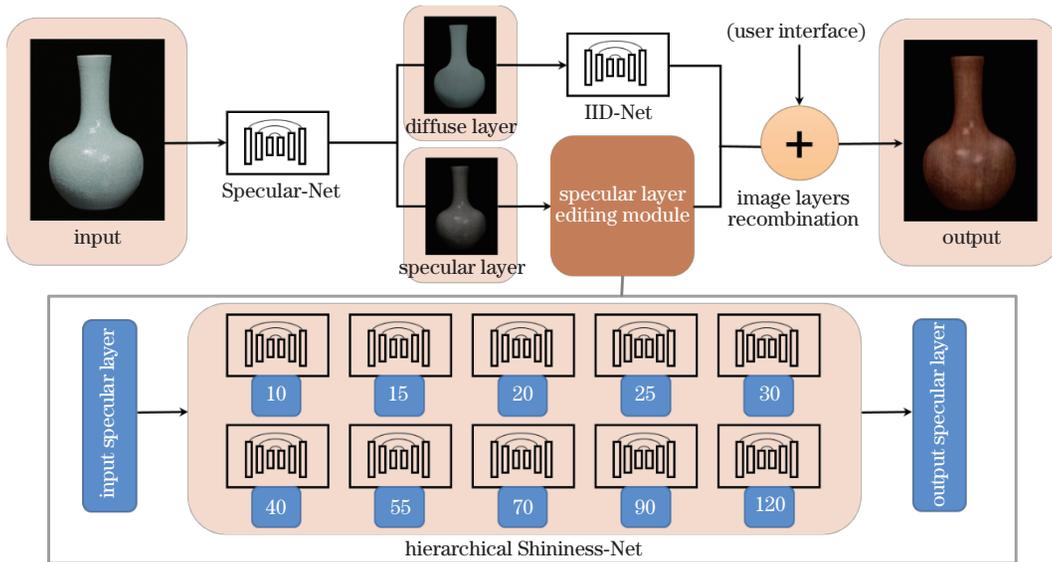


图 1 所提神经网络架构

Fig. 1 Architecture of proposed neural network

所提网络中的高光分离网络、本征图像分解网络、光泽度调节网络的结构均基于 Cycle-GAN^[22] 中的生成器, 该生成器网络是 U-Net。和普通 Encoder-Decoder 结构的网络相比, U-Net 加入了跳跃连接 (skip connection), 有利于提升生成图片的细节, 并广泛用于图像生成任务中, 同样适用于三个子任务。

1) 高光编辑模块。设计该模块的目的是实现对高光层的直接编辑, 从而完成由源材质到目标材质的高效转换。然而原图中光照条件未知, 源材质高光层参数未知, 因此高光编辑模块不仅需要能处理不同

的光照条件, 而且需要能接收不同高光层参数的源材质图片作为输入, 然后输出调节后的目标材质图片。为了应对这一难题, 合成了一个大规模包含多种光照和材质光泽度属性的数据集 (Hierarchical Shininess), 用于训练分层级的光泽度调节网络, 使之学习到不同光照条件和不同材质光泽度的数据特征。

具体来说, 采用将光泽度划分等级的方法, 以覆盖常见材质的光泽度变化范围。对光泽度等级的划分基于 OpenGL 中 GL_SHININESS 材质参数的有效范围, 有效范围是 0~128。利用材质模型渲染数据集时发

现:在 0~30 范围内,光泽度变化步长小于 5 时,渲染出的图片之间光泽度差异很小,随着光泽度的增大,需要使用更大的步长,才能使相邻层级的图片之间保持比较合适的光泽度差异。为尽可能覆盖大的光泽度转换范围,同时兼顾不同层级的光泽度图像之间均匀的递进变化,最终将光泽度等级设定为 10, 15, 20, 25, 30, 40, 55, 70, 90, 120 这 10 个级别。

为了实现对高光层参数 k_s 和 ρ 的调节,首先对高光分离网络输出的高光层执行归一化操作,同一种材质 k_s 一般为常量,归一化操作后近似得到了 $k_s=1$ 时的高光层,可记为 \hat{R} ;接着将 \hat{R} 输入高光编辑模块,模块中光泽度调节网络有从低到高十个等级,每一个等级都对应一个网络模型,值得注意的是它们隐式地包含了对源材质光泽度大小的估计。例如目标材质的光泽度为 55,将源材质的 \hat{R} 输入光泽度等于 55 的对应网络,网络会自动根据源材质的数据特征输出光泽度为 55 的目标材质 \hat{R} ,完成高光层光泽度的非线性转换。后续对 k_s 的调节同漫反射层参数 k_d 类似,属于线性转换。目标材质 \hat{R} 乘以目标材质的 k_s ,即可得到目标材质高光层。

2) 漫反射层参数调节与纹理替换。根据使用的参数化材质模型,漫反射层材质参数的调节对应于漫反射反照率 k_d 的改变。经过本征图像分解网络的处理后,漫反射层 D 分解为漫反射反照率图 A 和漫反射光

影图 S ,保持漫反射光影图不变,只对漫反射反照率图进行调节。具体来说:如果目标材质属于匀质材质,则漫反射反照率图中每一个像素都具有相同的漫反射反照率 k_d ,此时可以通过调节 k_d 的大小改变漫反射层的整体材质效果;如果目标材质是不均匀的,随着空间变化的,此时需要对漫反射反照率图中每一个像素使用不同的材质参数。采用纹理替换的方法改变每一个像素的漫反射反照率,把漫反射反照率图 A 替换成目标材质的纹理图片,然后结合原漫反射光影图 S ,得到目标材质的漫反射层。使用纹理替换,进一步扩大了所提方法可编辑的材质范围和增强了材质编辑的真实感。

2.4 训练细节

说明高光分离网络、本征图像分解网络、光泽度调节网络的训练细节,首先它们的训练数据和测试数据之比分别是 8:2、9:1、8:2。超参数设置如表 1 所示。批次大小定为 64,优化算法采用了默认参数的 Adam 算法,初始学习率是 1×10^{-4} ,收敛时的迭代次数分别为 50×10^3 、 28×10^3 、 35×10^3 。三个网络都是有监督训练,损失函数采用 MSE loss,每 10 个 epoch 监测一次损失值变化情况,若未下降,自动调整学习率为原来的 1/10,学习率下限设置为 1×10^{-7} 。训练的显卡配置为 Nvidia GTX1080Ti,高光分离网络训练时间是 10 h,本征图像分解网络训练时间是 6 h,每一个光泽度调节网络训练时间是 7 h。

表 1 网络训练超参数设置
Table 1 Hyperparameter settings in network training

Network	Batch size	Optimizer	Initial learning rate	Number of iterations
Specular-Net	64	Adam	1×10^{-4}	50×10^3
IID-Net	64	Adam	1×10^{-4}	28×10^3
Shininess-Net	64	Adam	1×10^{-4}	35×10^3

训练光泽度调节网络时,采用合成的 Hierarchical Shininess 数据集,其中包括 10 种不同光泽度等级的图片,把其中 1 种作为监督数据(Ground Truth, GT),另外 9 种作为输入数据,依次类推,通过这种方式可以训练得到 10 个光泽度调节网络模型;同时根据数据集的渲染机制(见第 3 节),采用了有针对性的数据增强措施,用于渲染的光源不仅包含白光,而且包含有颜色光,因此合成的图片中,图片颜色即光源颜色。对网络每批次输入图片的 RGB 通道以随机的方式重新排列组合,Ground Truth 的 RGB 三个通道也使用相同的重

排顺序,使得训练数据的颜色分布更加多样化,达到数据增强的目的。

3 数据集

所提方法使用的数据集如表 2 所示。高光分离网络使用的是 LIME 数据集,由 Meka 等^[4]基于 Blinn-Phong 模型^[21]合成,大小为 8.5×10^4 ;数据集提供了输入图片、漫反射层、高光层、法线图(normal map)等,高光分离网络使用与高光分离任务相关的部分进行训练。本征图像分解网络使用本征图像数据集

表 2 网络训练所使用的数据集
Table 2 Datasets used by proposed networks

Network	Dataset	Source	Size
Specular-Net	LIME	Meka et al. ^[4]	8.5×10^4
IID-Net	ShapeNet Intrinsic	Shi et al. ^[5]	20×10^4
Shininess-Net	Hierarchical Shininess	Ours	140×10^4

ShapeNet Intrinsic,是由 Shi 等^[5]基于 ShapeNet 数据集^[23]合成的,大小为 20×10^4 。光泽度调节网络使用本文合成的大规模数据集 Hierarchical Shininess,大小为 140×10^4 。

选取了 LIME 数据集提供的法线图 17500 张,其中包含 40 多个不同形状的物体。然而相比于 LIME 数据集,本文没有采用低动态范围图(LDR)作光源,而是采用反映真实场景光照信息的高动态范围图(HDR)作光源,使得渲染结果更具真实感。为了使 Hierarchical Shininess 数据集包含多种不同的光照条件,从 HDRIHAVEN 网站^[24](CC0 许可协议)收集到

56 个室内外不同场景的 HDR 图像,作为数据合成时的光源。将 HDR 图像的光照分布使用二阶球谐函数^[25-26]表示,然后在单位球坐标系下均匀采样 64 个点,每一个点近似为平行光源;利用归一化的 Blinn-Phong BRDF^[27]进行不同光泽度高光层的渲染。

Hierarchical Shininess 数据集中部分数据如图 2 所示。第一列是法线图,第二列到最后列是合成的高光层,从左到右光泽度依次增加。数据集包含了不同光照下具有不同光泽度材质参数的高光层。本课题组将公开该数据集,可供相关研究人员将其用于其他任务,如材质估计等。

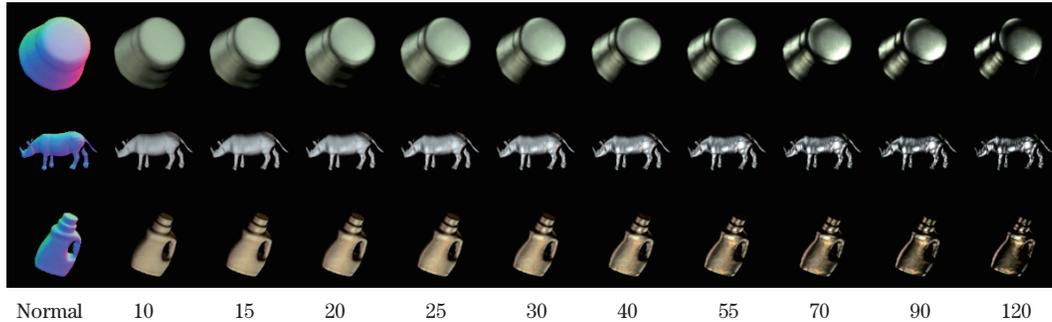


图 2 Hierarchical Shininess 数据集展示

Fig. 2 Picture display in the Hierarchical Shininess dataset

4 实验结果

4.1 高光分离

高光分离的对比方法分别是: Shen13^[28]、Shi17^[5]、Souza18^[9]、Yamamoto19^[29]、Fu21^[10]。采用的量化指标为尺度不变的均方误差(SMSE)、局部均方误差(LMSE)、结构相似性(DSSIM),具体可参考文献[15]。对分离高光后的漫反射层进行量化分析和视觉对比。

表 3 展示了所提方法和对比方法在三个指标上的对比。请注意 SHIQ 数据集来自 Fu21^[10],在其测试集上随机选取了 200 张;ShapeNet Intrinsic 数据集来自 Shi17^[5],其测试集和训练集具体如何划分是未知的,从数据集中随机选取了 600 张;实拍图片数据集来自 Yi20^[19],该数据集提供了 20 张实际拍摄的存在高光的

图片和利用偏振片消除高光的图片。从表 3 可以看出: Fu21^[10]在 SHIQ 数据集上的表现是最好的;所提方法在 ShapeNet Intrinsic 数据集上不仅优于 Shi17^[5]这个本身就基于 ShapeNet Intrinsic 数据集训练的方法,而且在 SMSE 和 DSSIM 两个指标上超过了当前的最新方法 Fu21^[10];对于 20 张实拍图片,所提方法同样在 SMSE 和 DSSIM 两个指标上表现最优。

图 3 是所提方法与对比方法在实拍图片上的效果对比,可以发现:在高光较强造成颜色饱和的区域,所提方法能够得到很好的结果,而 Shen13^[28]、Shi17^[5]、Souza18^[9]、Yamamoto19^[29]这些方法倾向于在饱和区域过度估计高光强度,导致得到的哑光图片上原高光区域出现灰暗斑块。与所提方法相比, Fu21^[10]输出的图片存在残留的高光区域,如图 3 第二、第四行图片。

表 3 SHIQ 数据集、ShapeNet Intrinsic 数据集和实拍图片数据集上的量化评估

Table 3 Quantitative comparisons on the SHIQ dataset, ShapeNet Intrinsic dataset, and some real pictures

Method	SHIQ dataset			ShapeNet Intrinsic dataset			Real pictures		
	SMSE	LMSE	DSSIM	SMSE	LMSE	DSSIM	SMSE	LMSE	DSSIM
Shen13 ^[28]	0.0587	0.333	0.2002	0.1286	0.2433	0.207	0.0224	0.167	0.1745
Shi17 ^[5]	0.0305	0.1789	0.2199	0.0133	0.1056	0.1576	0.0167	0.2497	0.2087
Souza18 ^[9]	0.0607	0.349	0.2062	0.1258	0.2391	0.2103	0.009	0.1178	0.1639
Yamamoto19 ^[29]	0.0594	0.3364	0.2011	0.1286	0.2418	0.206	0.0225	0.1667	0.1743
Fu21 ^[10]	0.0003	0.0019	0.0088	0.0107	0.0332	0.1034	0.0056	0.1035	0.1414
Ours	0.0108	0.0485	0.1765	0.0078	0.0449	0.0927	0.0056	0.1448	0.1384

Notes: bold font is the best value.

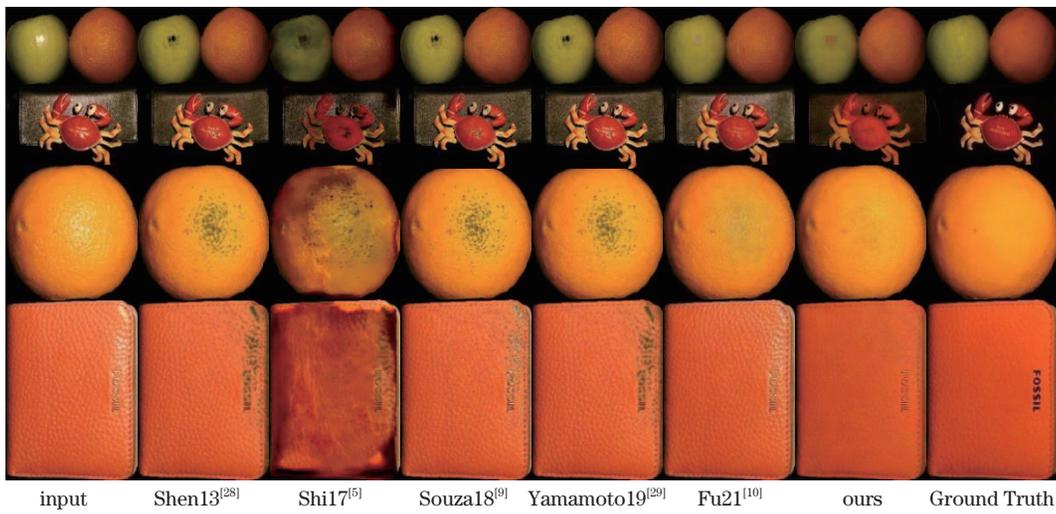


图 3 实拍图片上的高光分离效果对比

Fig. 3 Visual comparisons of highlight separation on real pictures

4.2 本征图像分解

本征图像分解中,对比方法分别是 SIRFS15^[30]、DI15^[15]、Shi17^[5]、Yi20^[19]、SMCH21^[31]。使用 MIT 本征数据集^[32]进行测试,同样采用 SMSE、LMSE、DSSIM 指标进行量化分析,结果如表 4 所示。

从表 4 中可以发现:SIRFS15^[30]有最好的分解结

果,这是由于该算法使用了纯朗伯表面反射和白光照明设置的先验,比较适合 MIT 本征数据集;所提方法通过网络训练漫反射光影图(shading),而漫反射反照率图(albedo)根据本征图像分解方程 $A \odot S = D$ 获取,为保证方程恒成立,对两项分别计算误差时会有一定损失。

表 4 在 MIT 本征数据集上的测试对比

Table 4 Quantitative comparisons on the MIT intrinsics dataset

Method	SMSE		LMSE		DSSIM	
	albedo	shading	albedo	shading	albedo	shading
SIRFS15 ^[30]	0.0147	0.0083	0.0416	0.0168	0.1238	0.0985
DI15 ^[15]	0.0277	0.0154	0.0585	0.0295	0.1526	0.1328
Shi17 ^[5]	0.0278	0.0126	0.0503	0.0240	0.1465	0.1200
Yi20 ^[19]	0.0274	0.0145	0.0476	0.0284	-	-
SMCH21 ^[31]	0.0225	0.0146	0.0484	0.0278	0.1499	0.1912
Ours	0.0273	0.0129	0.0386	0.0335	0.1484	0.1420

Notes: bold font is the best value, '-' indicates that it is not available.

图 4 展示了所提方法与对比方法在 MIT 本征数据集上的输出效果对比。第一列为测试输入,第 2~7 列中,奇数行展示了各个方法输出的 albedo,偶数行展示

了各个方法输出的 shading。通过数值和视觉效果上的对比发现:所提本征图像分解网络取得了与对比方法相当的实验结果。

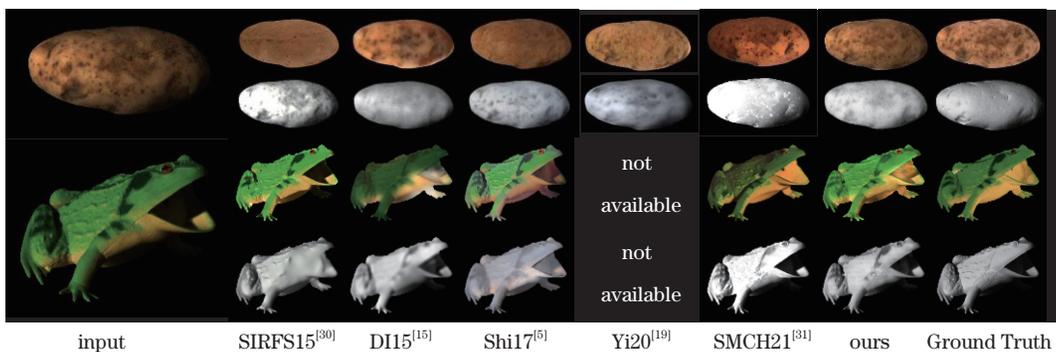


图 4 在 MIT 本征图像分解数据集上测试

Fig. 4 Visual comparisons on the MIT intrinsics dataset

4.3 材质编辑

利用经高光分离网络对输入图片进行处理后,将得到的高光层继续输入高光编辑模块,调节材质

光泽度。为了评价高光编辑模块的输出效果,图 5 展示了合成的真实高光层(GT)和网络输出(output)的效果对比。

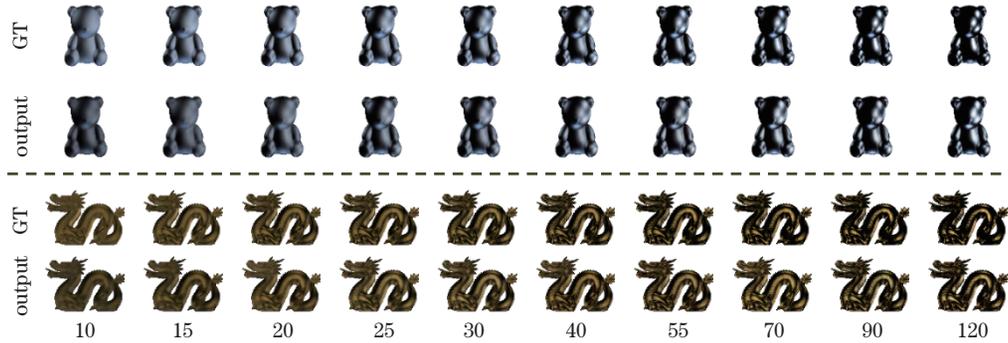


图 5 真实合成的高光层同网络输出结果之间的对比

Fig. 5 Visual comparisons on specular layer between GT images and network output images

如图 5 所示,第 1 行、第 2 行分别是小熊的真实高光层和网络输出的结果;第 3 行、第 4 行分别是斯坦福龙的真实高光层和网络输出的结果。从第 1 列到第 10 列,光泽度依次增加,高光分布越来越集中。请注意:小熊和龙的高光层数据是独立于 Hierarchical Shininess 数据集的,其几何形状与数据集图片完全不同,并且使用新的 HDR 图作为光源。在测试时以光泽度等于 28 的合成高光层作为网络输入,然后输出 10 到 120 这十种不同光泽度的网络预测图。从视觉效果上看,预测图和

真实合成的高光层非常接近,这在一定程度上证明了所提高光编辑模块的泛化能力。

根据式(4)将不同材质光泽度的高光层与漫反射层叠加起来,输出完整的图片。其真实合成效果同所提方法输出结果之间的对比如图 6 所示,第 1 行和第 3 行是真实合成图片(GT),第 2 行和第 4 行是所提方法得到的结果(output);从第 1 列到第 5 列,材质光泽度分别是 15, 25, 40, 70, 120。经过对比发现,本文的材质光泽度调节效果与真实合成的图片十分接近。

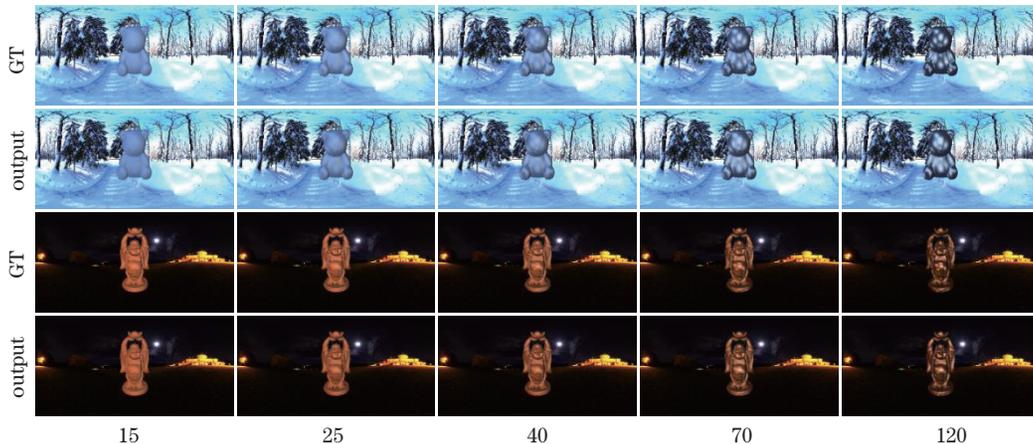


图 6 不同材质光泽度的高光层与漫反射层叠加后真实合成效果和所提方法输出结果之间的对比

Fig. 6 Visual comparisons between the synthetic images (GT) after the combination of specular layer and diffuse layer with different material shininess parameters and the output results of the proposed method

要完全改变一个物体的材质,除了调整材质表面光泽度 ρ ,还需要调整漫反射反照率 k_d 和光泽反射反照率 k_s 。应用高光分离、本征图像分解、高光编辑三个子任务,呈现整个方法流程的材质编辑效果。

图 7 展示了对真实图片的材质编辑效果,输入图片来源于网络。(a)~(h)对应各自的输入图像(input)进行材质编辑后的结果。其中(a)(b)表示将绿色陶瓷花瓶调节成黄色和蓝色,并且增加了光泽度;(c)(d)展示使用纹理贴图(i)(j),将青色陶瓷花瓶编辑成裂纹釉

和红木材质,使用纹理图片,相当于改变了原图中每一个像素的漫反射反照率,可获得更多样化的材质编辑结果;(e)(f)展示将木质兔子编辑成蓝色塑料和玉石材质;(g)(h)分别表示对石头青蛙贴上纹理(k)和调节光泽反射率的 RGB 三通道大小,得到了黄色石头材质和金属材质。从图 7 可以看出,所提方法在真实图像上实现了逼真的材质编辑效果,能够实现多样化材质间的转换。

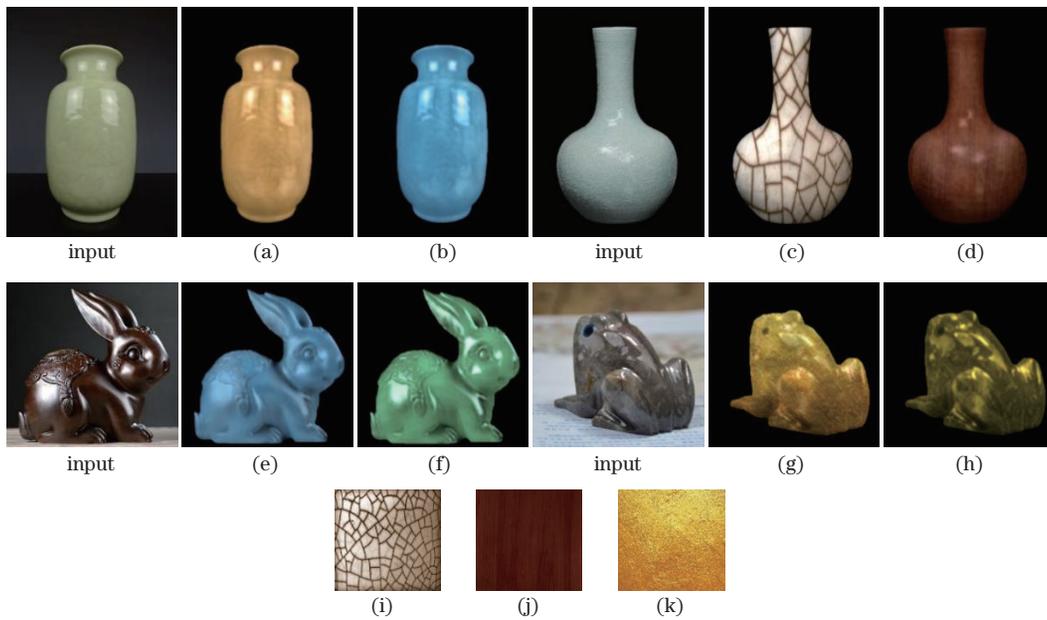


图 7 真实图片的材质编辑

Fig. 7 Material editing of real pictures

5 讨 论

所提方法利用参数化材质模型,并基于物理的图像形成过程,通过深度卷积网络获取到图像材质、几何、光照三个方面的信息,其中几何与光照信息由漫反射光影图和高光反射层隐式表示;需要改变的材质信息包括漫反射反照率、光泽反射反照率及光泽度三个方面,针对它们进行调节,使得所提方法具有编辑高效、逼真、交互性好等优点。

所提方法也有一定的局限性:1)不支持光线次表面散射、折射等复杂材质交互效果,比如不支持对液体、透明或半透明物体进行材质编辑;2)不能利用图片输入作为目标材质进行材质迁移;3)高光分离和本征图像分解作为所提方法中的两个子任务,二者本身仍然存在未解决的问题和挑战。

实现对艺术家友好的、具有照片级真实感的、支持复杂几何和材质类型的二维图片材质编辑是长期目标,这在多个领域具有广阔的应用前景。

6 结 论

提出一种面向单张图像的材质编辑方法,并将其拓展到实际应用,具有编辑高效、逼真、交互友好的优势。利用参数化的材质模型,完成了高光分离、本征图像分解和高光编辑 3 个子任务。包含的主要创新点是:通过合成一个大规模的数据集,训练出分层级的光泽度调节网络,由它们组成高光编辑模块实现对高光层的直接编辑。与对三维模型进行材质编辑和重渲染相比,所提方法避免了对新材质重渲染的耗时步骤,实现了未知光照下的高效编辑。所提方法不仅可以快速展现结果,而且能有效降低非专业用户的学习成本。

针对所提方法存在的局限性,接下来的工作将考虑实现多张图像之间的材质迁移并拓展可以支持的材质种类。

参 考 文 献

- [1] Khan E A, Reinhard E, Fleming R W, et al. Image-based material editing[J]. ACM Transactions on Graphics, 2006, 25(3): 654-663.
- [2] Boyadzhiev I, Bala K, Paris S, et al. Band-sifting decomposition for image-based material editing[J]. ACM Transactions on Graphics, 2015, 34(5): 163.
- [3] Liu G L, Ceylan D, Yumer E, et al. Material editing using a physically based rendering network[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2280-2288.
- [4] Meka A, Maximov M, Zollhöfer M, et al. LIME: live intrinsic material estimation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 6315-6324.
- [5] Shi J, Dong Y, Su H, et al. Learning non-Lambertian object intrinsics across ShapeNet categories[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5844-5853.
- [6] Tan P, Lin S, Quan L, et al. Highlight removal by illumination-constrained inpainting[C]//Proceedings Ninth IEEE International Conference on Computer Vision, October 13-16, 2003, Nice, France. New York: IEEE Press, 2003: 164-169.
- [7] Kim H, Jin H L, Hadap S, et al. Specular reflection separation using dark channel prior[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition,

- June 23-28, 2013, Portland, OR, USA. New York: IEEE Press, 2013: 1460-1467.
- [8] Guo J, Zhou Z J, Wang L M. Single image highlight removal with a sparse and low-rank reflection model[M]// Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11208: 282-298.
- [9] Souza A C S, Macedo M C F, Nascimento V P, et al. Real-time high-quality specular highlight removal using efficient pixel clustering[C]//2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), October 29-November 1, 2018, Parana, Brazil. New York: IEEE Press, 2018: 56-63.
- [10] Fu G, Zhang Q, Zhu L, et al. A multi-task network for joint specular highlight detection and removal[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 7748-7757.
- [11] Funt B V, Drew M S, Brockington M. Recovering shading from color images[M]//Sandini G. Computer vision-ECCV'92. Lecture Notes in Computer Science. Heidelberg: Springer, 1992, 588: 124-132.
- [12] Shen L, Yeo C. Intrinsic images decomposition using a local and global sparse representation of reflectance[C]// Conference on Computer Vision and Pattern Recognition 2011, June 20-25, 2011, Colorado Springs, CO, USA. New York: IEEE Press, 2011: 697-704.
- [13] Weiss Y. Deriving intrinsic images from image sequences [C]//Proceedings Eighth IEEE International Conference on Computer Vision, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE Press, 2001: 68-75.
- [14] Lee K J, Zhao Q, Tong X, et al. Estimation of intrinsic image sequences from image+depth video[M]// Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7577: 327-340.
- [15] Narihira T, Maire M, Yu S X. Direct intrinsics: learning albedo-shading decomposition by convolutional regression [C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 2992.
- [16] Li Z Q, Snavely N. CGIntrinsics: better intrinsic image decomposition through physically-based rendering[M]// Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11207: 381-399.
- [17] Bell S, Bala K, Snavely N. Intrinsic images in the wild [J]. ACM Transactions on Graphics, 2014, 33(4): 1-12.
- [18] Fan Q N, Yang J L, Hua G, et al. Revisiting deep intrinsic image decompositions[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 8944-8952.
- [19] Yi R J, Tan P, Lin S. Leveraging multi-view image sets for unsupervised intrinsic image decomposition and highlight separation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12685-12692.
- [20] Kajiya J T. The rendering equation[J]. ACM Computer Graphics, 1986, 20(4): 143-150.
- [21] Blinn J F. Models of light reflection for computer synthesized pictures[J]. ACM Computer Graphics, 1977, 11(2): 192-198.
- [22] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.
- [23] Chang A X, Funkhouser T, Guibas L, et al. ShapeNet: an information-rich 3D model repository[EB/OL]. (2015-12-09)[2021-02-06]. <https://arxiv.org/abs/1512.03012>.
- [24] HDRIHAVEN. All HDRIs[EB/OL]. [2021-05-06]. <https://hdrihaven.com/hdri/>.
- [25] Ramamoorthi R, Hanrahan P. An efficient representation for irradiance environment maps[C]//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, August 12-17, 2001, Los Angeles, California, USA. New York: ACM Press, 2001: 497-500.
- [26] RobinGreen. Spherical harmonic lighting: the gritty details[EB/OL]. [2021-04-21]. <http://www.research.scea.com/gdc2003/spherical-harmonic-lighting.html>.
- [27] Akiene-Möller T, Haines E, Hoffmann N. Real-time rendering[M]. 3rd ed. Boca Raton: CRC Press, 2008: 270-276.
- [28] Shen H L, Zheng Z H. Real-time highlight removal using intensity ratio[J]. Applied Optics, 2013, 52(19): 4483-4493.
- [29] Yamamoto T, Nakazawa A. General improvement method of specular component separation using high-emphasis filter and similarity function[J]. ITE Transactions on Media Technology and Applications, 2019, 7(2): 92-102.
- [30] Barron J T, ShapeMalik J. Illumination, and reflectance from shading[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(8): 1670-1687.
- [31] Lichy D, Wu J Y, Sengupta S, et al. Shape and material capture at home[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 6119-6129.
- [32] Grosse R, Johnson M K, Adelson E H, et al. Ground truth dataset and baseline evaluations for intrinsic image algorithms[C]//2009 IEEE 12th International Conference on Computer Vision, September 29-October 2, 2009, Kyoto, Japan. New York: IEEE Press, 2009: 2335-2342.