

An Improved Pre-Net Network and Its Application in Material Modeling

Jinlian Du, Xiangdong Kong

Beijing University of Technology, Beijing
Email: 975642117@qq.com

Received: Apr. 6th, 2020; accepted: Apr. 21st, 2020; published: Apr. 28th, 2020

Abstract

Material modeling is a hot issue in the research and application of graphics. The purpose is to determine the properties of material appearance such as texture, normal, reflectivity and roughness through modeling, so as to improve the realism of object rendering. In order to solve the problem of traditional material modeling, it takes a lot of manpower, resources and time. Based on the pre-net proposed by Xiao Li [1], this paper studied the over-fitting and prediction accuracy, and proposed a method to improve the model. By adding Dropout layer and Inception structure, the over-fitting phenomenon was reduced and the prediction accuracy of the model was improved. At the same time, a large number of training data needed for the experiment were obtained quickly by using batch rendering technology. Experiments show that this method can make the prediction result of the model more accurate and the generalization ability stronger.

Keywords

CNN, Appearance Modeling, Render Model, Self-Augmentation

一种改进的Pre-Net网络及其在材质建模中的应用

杜金莲, 孔祥栋

北京工业大学, 北京
Email: 975642117@qq.com

收稿日期: 2020年4月6日; 录用日期: 2020年4月21日; 发布日期: 2020年4月28日

摘要

材质建模是图形学研究与应用中的热点问题, 目的是通过建模确定材质表现的纹理、法向、反射率和粗糙度等属性, 从而提高对物体渲染的真实感。为了解决传统的材质建模需要耗费大量人力物力和时间, 目前有人提出利用深度学习的方式实现材质的建模。本文在Xiao Li [1]提出的Pre-Net的基础上, 对过拟合及预测精度问题进行研究, 提出了改进模型的方法, 通过添加Dropout层和Inception结构, 从而减少过拟合现象并提高了模型的预测精度, 同时使用批渲染技术快速大量地获取到实验所需的训练数据。实验表明, 本文方法可以使模型的预测结果更加准确, 泛化能力也更强。

关键词

卷积神经网络, 材质外观, 渲染模型, 自增强

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

材质建模是图形学研究与应用中的热点问题, 是通过某种方法或手段对材质的粗糙度、漫反射率、高光反射率和法向等属性进行建模, 以确定这些属性的值从而提高对物体渲染的真实感。传统的材质建模是通过人工的方式获得, 为了节省成本并提高效率, 目前已经有人使用深度学习的方法通过反复学习基于未知光照环境下的双向散射分布函数材质样本图片重建材质的属性。但利用深度学习进行材质建模的缺点是训练数据十分匮乏, 影响重建的效果, 同时也因为神经网络模型太过复杂, 容易出现过拟合现象。本文基于文献[1]提出的方法对基于深度学习的材质建模方法进行研究, 力图对训练数据匮乏及过拟合问题提出解决方案。论文的主要贡献如下:

1. 提出了一种自动化高效生成所需标记数据集的方法, 通过控制渲染程序批量地渲染了用于训练的标记数据对。
2. 对文献[1]的网络结构进行改进, 通过添加 Dropout 层和 Inception 结构, 减少了过拟合现象的发生并提高了模型的预测精度。
3. 实验并讨论了自增强训练策略对卷积神经网络解决材质建模问题的有效性。

2. 相关工作

2.1. 材质属性获取方法

物体的材质参数建模是一个非常困难的问题, 常见的方法是通过专业仪器多角度的测量获取, 或是进行人工绘制, 费时又费力。近年来出现了很多利用深度学习工具来解决材质建模的方法, 并且效果都比较理想。Dror 等人[3]第一个证明机器可以训练学习算法来对低级别的材料进行图片分类功能, 从那时起利用深度学习工具解决相关问题成为了一个有效的方法。例如图像分解以及反射率和照明度估计, 但是由于材料级别低, 估计值很不理想。Zhu 等人[4]采用了编码器-解码器架构来估计漫反射率和法线贴图。但是, 他们的方法只能恢复材料样本上的均匀镜面反射参数并他们针对不同的材料只训练了一个通用的

网络。Xiao Li 等人[1]提出使用一个卷积神经网络(convolution neural network, CNN)去实现物体的表观建模, 他们只需要一张在未知光照条件下拍摄的平面的材质样本照片, 就可以通过它重建出样本图片中材质的表观特征。他们遇到的最大的障碍在于他们十分缺少标记数据对(未标记的普通图片及对应的表观参数)。

对照明条件的限制更好地显示了反射率属性的估计更依赖于主动照明。Wang [5]等人使用阶梯边缘照明条件来推断空间变化的表面法线, 以及均匀材料的反射率特性。Oxholm 和 Nishino [6]依赖已知的照明来恢复形状并在自然光照下从物体的单张照片获得物体的表面反射率。Aittala 等人[7]利用固定材料的相似性估计空间变化的法线和反射率特性, 从闪光灯下的平面材料样品的一张照片中使用深层纹理合成框架, 而不是依赖主动照明。所有这些方法要么需要主动照明, 要么需要对照明有充分的了解, 这就排除了它们对在非受控捕获条件下获得的输入照片的适用性。

2.2. 训练数据集获方法

许多深度学习方法的共同弱点是都要依赖大型的训练数据集。Wojciech Matusik 等人[8]采样了 100 个不同材质球体的材质属性。他们制作了一个具有两个机械臂的采样所用的机械装置去采样得到这些材质属性特征, 这个机械装置能够自动地在不一样的光照条件和相机视角下, 去拍摄一个球体的模型的照片, 以此来获得所需要的样本数据。但是这种方法不要求制作巨大且昂贵的设备, 还需要手工制作 100 个具有均匀材料的真实模型, 在实际应用中不太可能使用这种耗时耗力的方法。

Jiaping Wang 等人[9]将这个流程优化, 即使用相机固定某一视角从而可以获取大量训练数据(样本图片), 同时在图片数据的每个像素点上采集局部材质特征, 并使用这些材质特征作为寻找材质物体表面相似特征像素点的参照对象。紧接着就是收集包含类似材质属性特征的所有像素点, 最后他们就可以重建出材质图片对应的表观属性。尽管使用了数据之间的相关性减少了必须的图像数据数量, 但是由于一开始的采集部分还是需要使用不少的人力物力, 一组高质量的样本数据仍要花费很长的时间去完成其数据的采样过程。

为了减少需要获取的图像数据量, Zhiming Zhou 等人[10], 提出了一个尽力精简的模型, 使用具备代表性的材质来表现所关联的图片, 从而达到简化网络复杂度的目的。相比与普通的简化网络模型, 他们改进的模型对图片数据数量的需求进一步的缩减, 对于某种材质, 只需几分钟去采集若干张材质样本的照片, 就可以对材质进行高质量的材质表观重建, 但由于模型复杂度低, 训练空间小, 很难训练出有效的模型, 导致预测出的结果非常不理想。

3. 改进的 Pre-Net 网络模型

利用深度学习工具解决图形学领域中的问题已经成为一种常见方法, 构建何种结构的卷积神经网络才能更高效的完成材质建模是本文研究的重点。本文以 Xiao Li 的 Pre-Net 为基础, 仔细分析了模型的重点和难点, 提出了 DISAPre-Net (改变了网络层结构并且经过自增强训练后的 Pre-Net), 改进后 DISAPre-Net 可以更好地应对网络训练中经常出现的过拟合现象同时可以提高模型的预测能力。

3.1. Pre-Net 基础结构

Pre-Net 由两子网络组成, 分别是 Spec-Roughness-Net 和 Albedo-Normal-Net。前者用于估计高光和粗糙度分量, 后者用于估计漫反射和法向量分量。Pre-Net 结构如图 1 所示。

图 2 中蓝色和棕黄色部分是分析子网, 由三部分组成, 分别是卷积层(Convolution)、批量标准化层(Batch Normalization, BN)和激活层(ReLU)。卷积层作为特征提取层, 将每个卷积核进行卷积计算, 并将结果作为下一层的输入, 是提取图片高维特征的重要组成部分。模型在训练过程中不断更新权重参数,

而每一层输入的数据分布和之前相比会有很大差异, 需要模型不断去适应新的数据, 导致训练过程异常艰难。BN 层可以对上一网络层的输入进行标准化处理, 让数据的分布强行回到正太分布, 减少了数据分布对模型的影响。ReLU 激活层可以通过 ReLU 激活函数(非线性函数)增加模型的非线性表达能力。

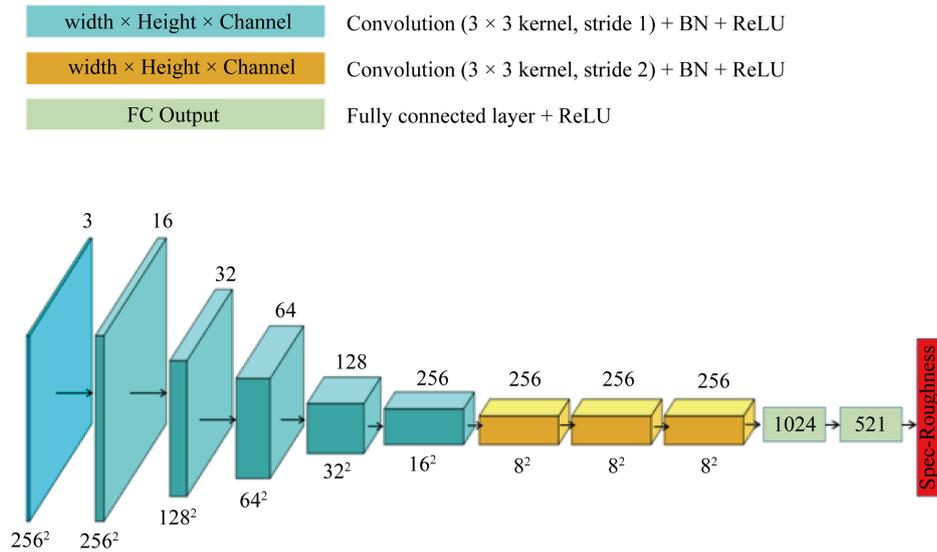


Figure 1. Spec-Roughness-Net
图 1. Spec-Roughness-Net

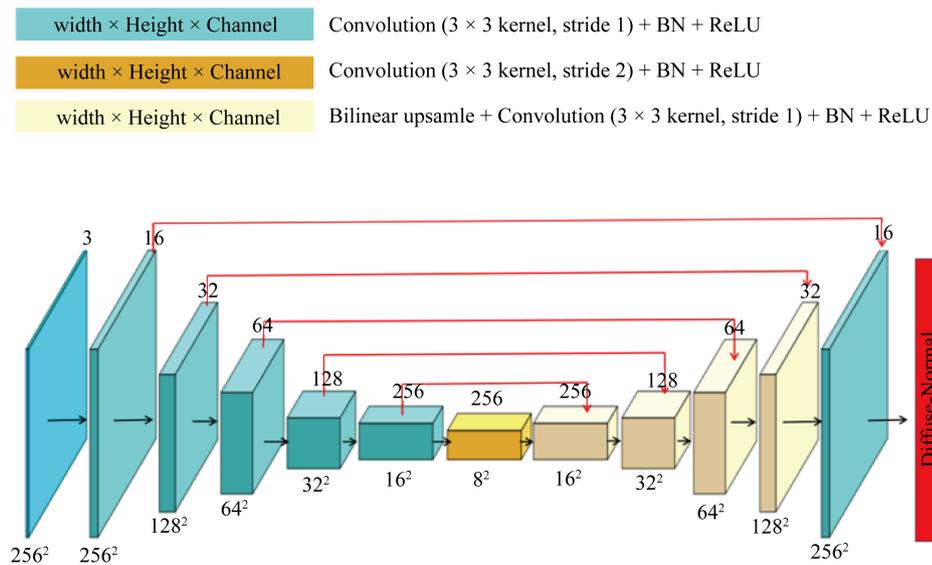


Figure 2. Albedo-Normal-Net
图 2. Albedo-Normal-Net

两子网络中分析子网使用的填充都为 0, 但前者步幅为 2, 后者步幅为 1。选择卷积核的尺寸大小为 3 * 3, 原因是较小的卷积核可以增强网络容量和模型复杂度同时减少了卷积参数的个数。此外, 图中还给出了通过相应的卷积和汇合操作后网络的相关维度。可以形象的看到经过每一次卷积之后图像的大小以及每一次提取的特征数。

分析子网之后跟着两个不同的合成子网。对于 Spec-Roughness-Net 来说, 分析子网后连接了一个具

有 1024 个隐藏变量的全连接层以及六个输出节点(高光反射率和粗糙度对每个颜色通道的预测), 通过这些来合成所需要的合成子网。而对于 Albedo-Normal-Net 来说, 分析子网后面连接了一个双线性上采样层, 用来扩大特征图。在这之后连接了一个卷积层、ReLU 激活层和 BN 层, 共同组成了合成子网。在图 2 中每一层的特征映像和与之对应的子网层次用红色箭头相连, 可以弥补在特征提取过程中丢失的高频像素。

3.2. Pre-Net 改进方法

当卷积神经网络模型太过复杂或训练数据集不匹配模型复杂度时会出现过拟合现象。分析层可以提取高维特征、对参数标准化和增加模型的非线性能力, 在分析层后添加 Dropout 可以屏蔽某些节点, 让经过分析层出来的特征符合样本材质的普遍特征。所以本文在 Pre-Net 网络基础上, 在每个分析层后添加 Dropout 层来解决过拟和问题, 同时为了提高模型的预测精度, 在网络层加入 Inception 结构。

Albedo-Normal-Net 以及 Spec-Roughness-Net 在一定程度上都属于复杂规模的神经网络, 而且训练数据集相对较少, 复杂度不能与模型匹配, 所以在训练过程中模型对表观参数的预测会出现过拟合。过拟合现象是 CNN 训练过程中经常出现的一个问题, 其表现在对训练数据预测较好, 对交叉测试集预测先好后差。Dropout 可以有效解决模型在训练过程中出现的过拟合问题。Dropout 原理如图 3 所示。

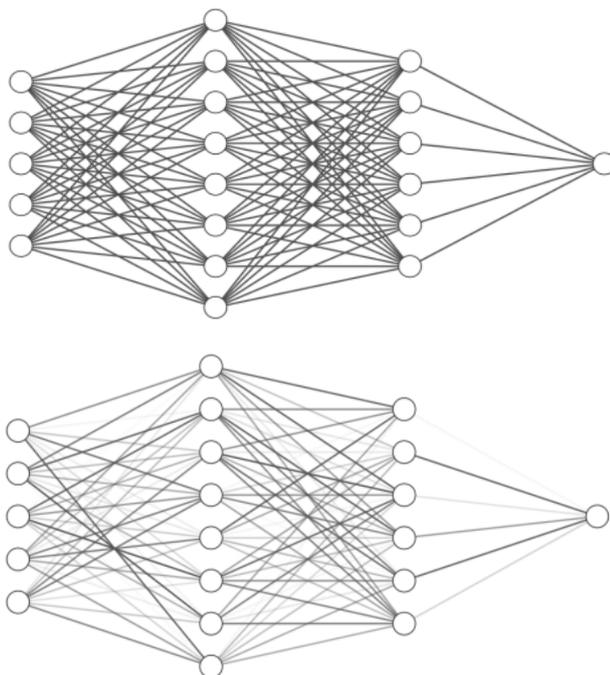


Figure 3. Dropout principle

图 3. Dropout 原理

图 3 上部分是添加 Dropout 之前, 下部分是添加之后。Dropout 的作用主要可以分为两点, 第一点是取平均的作用即选择性的丢弃神经元, 类似训练一些不同的网络。第二点是可以减少神经元之间复杂的共适应关系。简单来说, 模型在前向传播过程中, Dropout 会随机的丢弃神经元, 使得权值的更新不再依赖某些神经元, 即使这些神经元丢弃, 也会在其他神经元中学习到共同的特征。本文使用的 Dropout rate (神经元的丢弃比例) 是 0.5, 模型在进行批训练时, Dropout 程序可以产生不同的网络结构。而当 dropout rate 在 0.5 的时候, 随机生成的网络结构最多。

传统提高模型预测能力的方法主要有两种, 即增加模型的复杂度和扩大训练集。但这两种方法有很

严重的缺陷, 不仅增加了权重参数, 而且非常耗时。“稀疏连接结构”可以从本质上提高网络性能, 而 Inception module 即保持了网络结构的稀疏性, 又能够使用密集矩阵的高效计算。Inception 结构如图 4 所示。

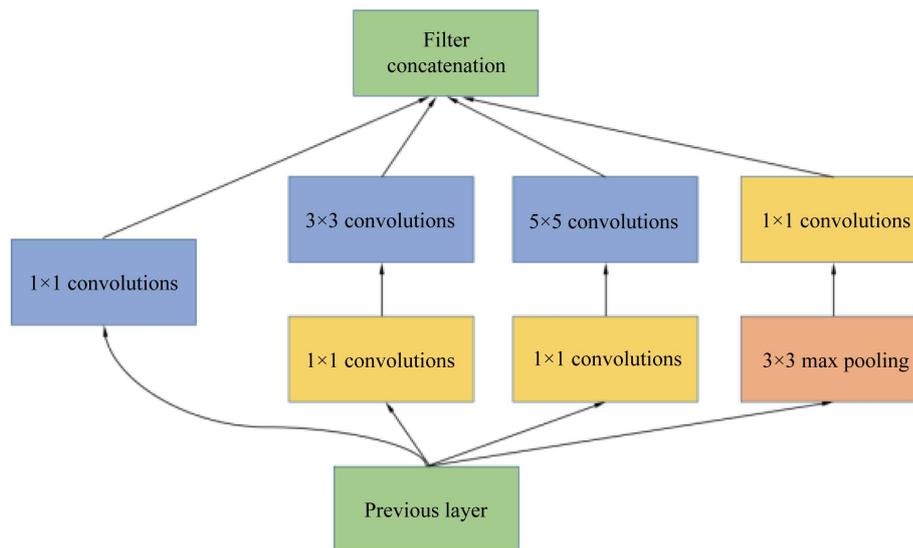


Figure 4. Inception module
图 4. Inception 组件

当加入 Inception 结构后, 可以将单个卷积层变成一个小神经网络, 不同层的特征会被组合到一起, 一般来说同层间的高层特征和底层特征并不能结合, 而 inception 结构恰好解决了这一问题。实现了不同尺度特征融合, 提高了模型的预测精度。

4. 数据渲染及自增强策略

训练模型需要大量的图片数据, 而每一种样本图片都需要随机的未知灯光, 如果手动渲染则会大幅降低渲染效率, 所以本文提出一种批渲染方法, 通过脚本程序来完成渲染任务。同时, 采用了一种利用未标记图片来增加训练数据的训练策略即自增强策略。

4.1. Ward 模型

本文中对 Albedo (漫反射) 和 Specalbedo (高光反射) 建模是由 Ward BRDF (双向反射分布函数) 实现的, 它由两部分组成。第一部分是漫反射部分 ρ_d/π , 由于本文的漫反射分量是基于空间变化的, 所以用 $\rho_d(x)$ 来代替 ρ_d 。另一部则是高斯各向异性反射模型, 这一部分有三个参数 ρ_s , α_x 与 α_y 。其中 ρ_s 表示的是高光反射率, 而 α_x 与 α_y 可以理解为样本材质表面在 x 轴和 y 轴方向上的粗糙度。Ward BRDF 函数模型如公式 (1) 所示:

$$f_r(i, o) = \frac{\rho_s}{4\pi\alpha_x\alpha_y\sqrt{\cos\theta_i\theta_o}} e^{-\tan^2\theta_h\left(\frac{\cos^2\phi_h}{\alpha_x^2} + \frac{\sin^2\phi_h}{\alpha_y^2}\right)} \quad (1)$$

公式中的 θ_i , θ_o 和 θ_h 分别表示了入射向量, 出射向量以及半角向量。其中 θ_h 是入射光向量与出射光向量求和后再标准化, 即 $\theta_h = (\theta_i + \theta_o)/|\theta_i + \theta_o|$ 。计算公式如 (2) 所示:

$$\theta_h = \arctan \sqrt{\frac{-\log u}{\cos^2 \phi_h / \alpha_x^2 + \sin^2 \phi_h / \alpha_y^2}}$$

$$\phi_h = \arctan \left(\frac{\alpha_x}{\alpha_y} \tan(2\pi v) \right) \quad (2)$$

公式(2)中的 u 和 v 表示的是两个完全无关的随机数, 取值范围是 0 到 1。从上式可以看出为了得到 θ_h , 必须先计算出 ϕ_h , 而在计算过程中需要注意一个地方, 那就是 v 和 ϕ_h 必须要在同一象限内。

4.2. Normal 模型

实际生活中, 我们观察到物体表面的凹凸感来源于光线在物体表面上的明暗变化。假如可以通过一幅贴图实现光照的明暗变化, 就可以在视觉效果上模拟出平面物体表面的凹凸感。在计算机中可以通过改变物体的法线朝向, 从而产生这种凹凸感, Normal 函数模型如公式(3):

$$N = \frac{R^2}{\pi(\cos^2 \theta_h (R^2 - 1) + 1)^2} \quad (3)$$

公式(3)中, R 表示样本材质图片表面粗糙度, 粗糙度取值范围是 0 到 1, 当 R 为 1 时表示图片表面粗糙度最高, 即没有任何高光反射。当 R 为 0 是表示图片表面光滑。

4.3. 图片渲染

实验使用 Maya 搭建虚拟 3D 环境, 使用 Mitsuba for Maya 渲染单张图片, 具体步骤如下: 配置渲染设置, 调整场景中的模型、灯光、摄像机等, 调整并赋予模型材质, 最后渲染, 保存图片。由于本文提出的方法需要大量的图片数据, 而且每一张样本图片都需要随机的未知的一种灯光, 都具有不同的材质属性, 每次渲染都需要调整, 这种手动控制渲染图片的过程明显不适合, 效率太低。本文提出使用脚本语言控制 Mitsuba for Maya 完成实验中渲染任务。同时可以利用 Python 一些强大的功能, 比如随机、循环、科学计算等来帮助完成调整材质参数、灯光参数, 以及给予得到未知的或随机的材质参数或灯光参数的能力。

在使用脚本语言后, 可以一次性用写好用来渲染一整个标记训练集的脚本语言文件, 在需要标记训练集时只需要在 Mitsuba for Maya 中运行这个脚本文件即可。而且, 在需求改变后, 只需修改脚本文件内的逻辑。以实验中的一次渲染任务为例, 若想读取一个文件中列出的 1000 组材质参数, 渲染 1000 张未知光照条件下的平面样本图片。首先, 使用 Python 的文件读写库读入已知的 1000 组材质参数; 使用脚本语言配置一些 Mitsuba 必须的渲染设置(如默认渲染器、是否启用全局光照、采样方法、采样深度等等); 然后启动渲染循环, 在每一次循环中, 随机改变光照环境并旋转随机的角度, 赋予材质相应的材质参数; 最后设置本次渲染循环的最终图片的保存位置及命名方式, 保存当前必要信息, 开启本次循环的渲染; 整个过程是通过脚本语言文件内写好的逻辑控制的, 在执行渲染任务时不需要进行人工干预, 高效、快速, 可扩展性强, 可以随时根据实验进度修改渲染逻辑, 既实现根据人工选择的材质集参数渲染初始的训练数据集和未知标记数据, 又实现自动读入网络预测所得参数集去渲染出相应的图片集。

4.4. 自增强训练策略

本文受限于材质规模, 当前可用材质库不具备如此全面的信息, 并且整理后可以用来做训练数据的材质照片远远不够。为了完成文中空间变化的双向反射分布函数网络, 需要借用普通的平面照片, 制作出尽可能多的训练数据, 并且这些训练数据要覆盖此类材质的尽可能多的空间变化特征。在制作训练数

据的过程中, 还需对每种空间变化模拟相机方向和光源方向, 并对每种环境下的数据充分采样, 以此来完成在后期训练完成的模型测试中, 对未知光源环境下的同类材质做表面反射参数的预测。

为了使用未标记(没有表观参数)的平面材质图片作为训练数据, 实验需要先使用有标记的材质图片作为初始训练数据, 网络学习后会得到一个初始的 CNN, 即 Pre-Net。对任一张未标记的平面材质图片, Pre-Net 可以初步预测出材质对应的表观参数。接下来本文将未标记数据和标记数据各自分成多组, 循环作为输入, 用来提升初始网络的预测能力。本文将介绍这个循环改进网络的自增强的训练过程。自增强训练分为两个阶段(如图 5 所示)。

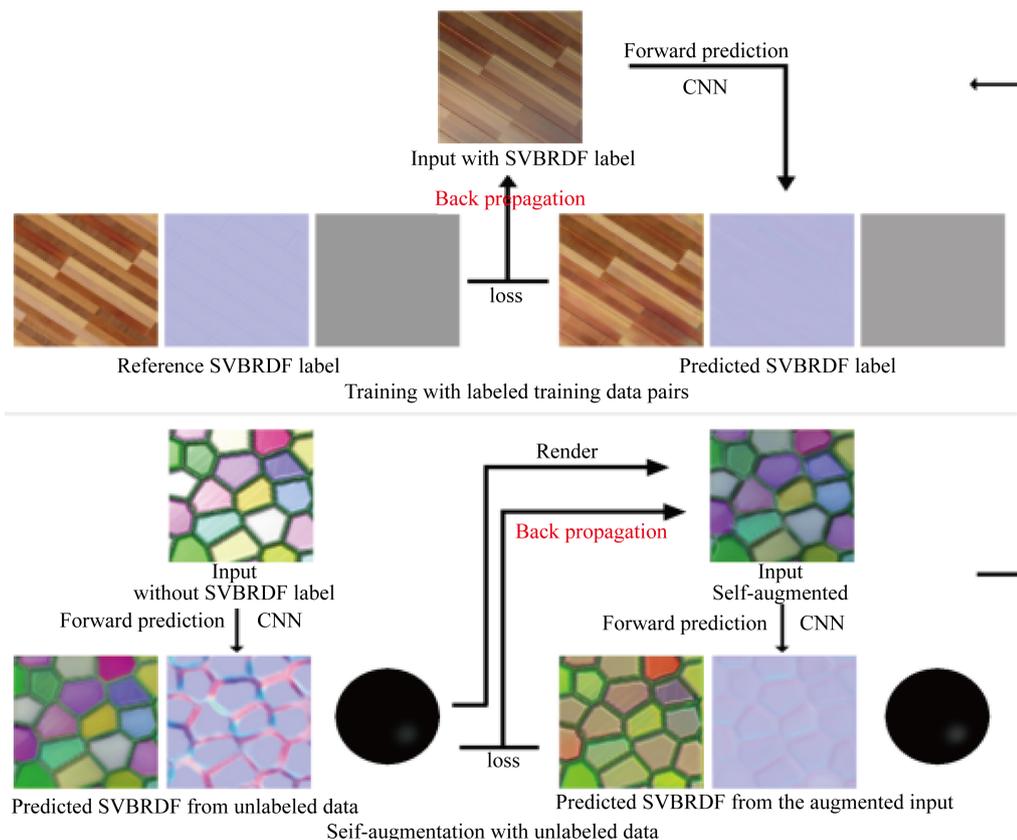


Figure 5. Self-augmented training process

图 5. 自增强训练过程

第一阶段是使用已标记数据集训练 CNN: 使用带有表观参数的平面材料样本中的图片组合为标记数据集对网络进行训练, 得到一个初始的网络(Pre-Net); 该网络对训练图片的预测与其标签的误差被描述为 loss, 根据 loss 的变化, 对超参数进行调整, 或分析得到网络是否收敛或过拟合的信息。

第二阶段是使用未标记数据集训练 Pre-Net: 首先把一批未标记图片送入所得初始网络中进行正向预测, 得到的预测标签; 然后根据预测标签值渲染得到一组新的图片, 与这组预测标签共同组成为一组新的标记数据集, 使用它们对这个初始的网络继续进行训练, 得到新的自增强过的网络(SAPre-Net); 同样, 其对训练图片的预测与其标签的误差被描述为 loss; 最后, 把增强过的网络当作新一轮的初始网络, 把这组渲染的图片当作新一轮的未标记图片, 重复这一阶段的过程, 直至网络收敛。未标记数据为网络上搜集的材质贴图, 材质类型分别是塑料、金属和木头。

本文将经过自增强训练的 Pre-Net 称为 SAPre-Net, 改进后并经过自增强训练的 PreNet-称为 DISAPre-Net。

4.5. 超参选择

超参是卷积神经网络在训练之前需要设置的参数, 而非训练所得, 实验需要找到一组最优超参, 用来提高卷积神经网络的性能和效果。

表中 Learning rate, 是模型训练中重要的超参, 它监督学习并且决定模型的收敛情况(何时收敛或是否能收敛)。本文使用的初始 learning rate 为 0.002, Policy 表示的是学习率策略, Inverse 表示随着迭代次数增加而减少, Gamma 表示学习率每次的变化程度。Weight decay 表示的惩罚权重参数的力度, 从模型的复杂度角度来说, 权重参数更小模型的复杂度也就越低, 对数据的拟合性也更好。momentum 是一个常量。Mini-batch-size 表示选取样本的数量, 如果数量偏大, 会降低模型的泛化能力, 反之, 会导致训练速度过慢。经过实验多次对比, 本文最终选取 Mini-batch-size 为 16。

本文基于初始网络(Pre-Net)提出了自增强网络(SAPre-Net)和改进后的自增强网络(DISAPre-Net)。并对这三种网络进行了多轮的实验对比, 分别进行了实验、训练和调参, 并对得出的结论进行分析对比。此外, 对改进后的 DISAPre-Net, 进行训练集的分类训练, 超参列表如表 1 所示。

Table 1. Hyper-parameter

表 1. 超参数

Learning rate			Weight decay	Momentum	Mini-batch size
Initial	Policy	Gamma	10-4	0.9	16
0.002	Inverse	10-4			

5. 实验结果与分析

5.1. 实验介绍

实验主要探究两个目的: 第一个目的是探究自增强策略的有效性; 第二个目的是探究经过优化后的卷积神经网络能否解决过拟合现象和提高网络模型的预测能力。

本文实验的三种网络分别是 Pre-Net (Xiao Li 提出的基本网络)、SAPre-Net (经过自增强训练的网络)和本文提出的 DISAP-Net (优化后的自增强网络)。通过平面材质参数的预测情况来对比预测模型之间的性能差异。实验步骤为先以 Pre-Net 和 SAPre-Net 的参数提取的准确性证明自增强模式的有效性。之后做 SAPre-Net 和 DISAPre-Net 的参数提取, 验证了 Dropout 层可以解决训练深度学习神经网络时经常会出现的费时(过慢收敛)和过拟合问题。Inception 结构也会提升模型的学习能力, 使预测精度进一步提升。

以上实验都是四个参数同时训练, 本文在后面还会分别对漫反射和法向两个参数单独训练, 判断训练过程中多个参数同时进行是否彼此限制。本文的实验均在 Caffe 环境(深度学习库, 提供了多种网络模型)中完成。实验所用的训练数据集来自于 Sean Bell 等人[2]的在线材质库。

5.2. 三组对比试验

本文将进行三组实验的对比, 第一组实验是 Pre-Net 和 SAPre-Net 之间的对比, 用于验证自增强的有效性; 第二组实验是 SAPre-Net 和 DISAPre-Net 之间的对比, 用于验证后者是否避免了过拟合现象和能否提高模型的预测精度; 第三组实验是 DISAPre-Net 预测所有参数和预测单个参数之间的对比, 用于验证预测单个参数的可视化结果是否更加准确。

实验中的三轮训练均设置了 128 个时期(epoch, 一个 epoch 表示将所有训练数据完整训练一次的过程), 下面展示 Pre-Net、SAPre-Net 和 DISAPre-Net 对材质外观参数的预测以及模型的收敛情况。

图 6~图 8 中, 第三列表示材质样本图片, 前两列与后两列分别表示由材质样本图片提取的漫反射贴图 and 法向贴图, 其中左列为预测值, 右列为参照值。



Figure 6. Pre-Net predicted results

图 6. Pre-Net 预测结果

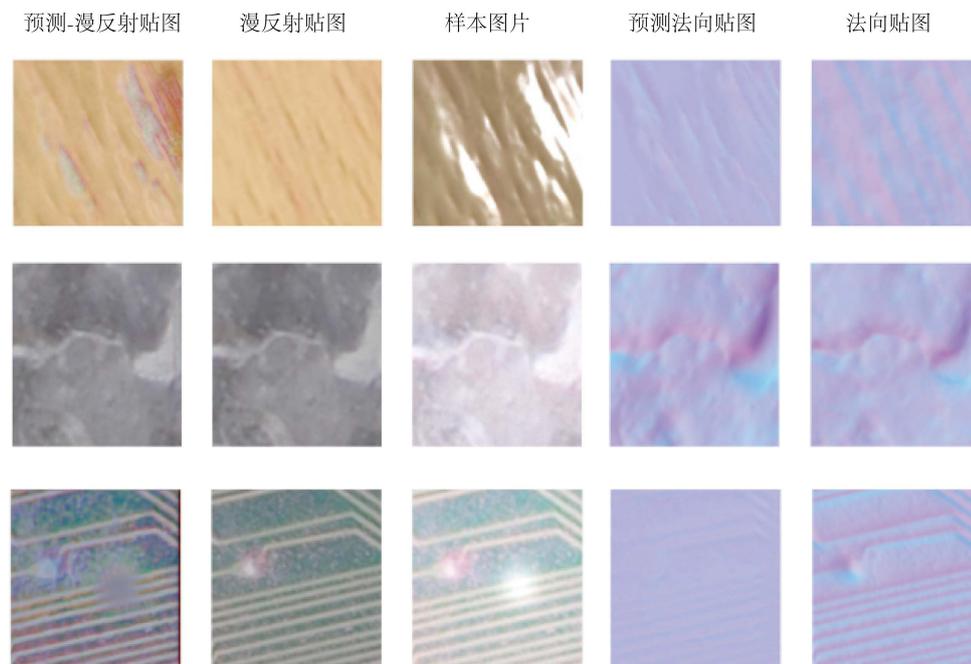


Figure 7. SAPre-Net predicted results

图 7. SAPre-Net 预测结果

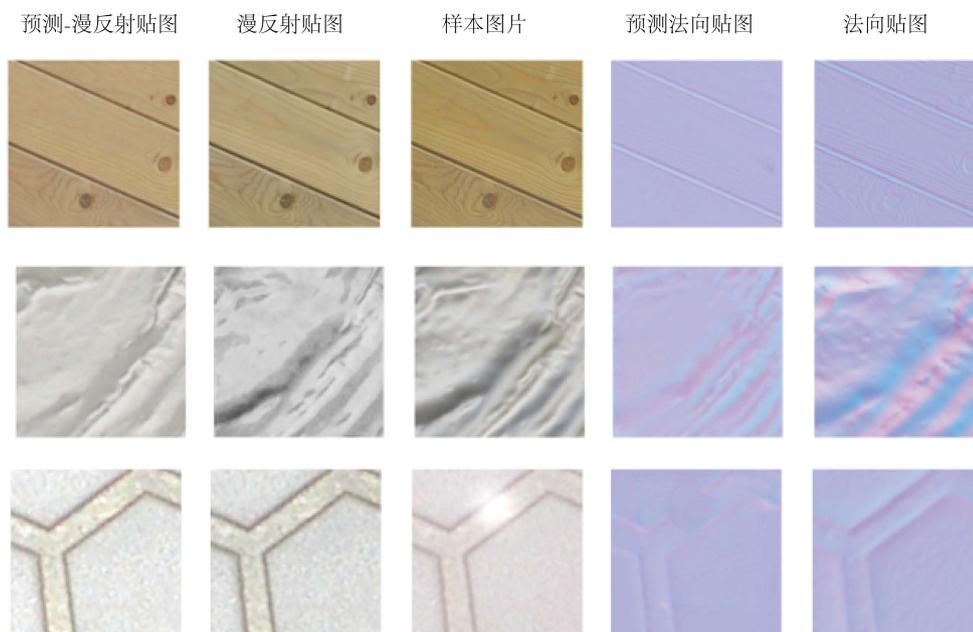


Figure 8. DISAPre-Net predicted results
图 8. DISAPre-Net 预测结果

图 6 可以看出无论是漫反射贴图的纹理形状、颜色还是法向贴图的凹凸情况, 预测结果都与参照值差异较大, 此时的 Pre-Net 不能达到实验对模型的要求。图 7 中是 SAPre-Net 预测出的结果, 对颜色和整体特征的把握有了很大的进步, 纹理的大致信息都较为准确, 但颜色细节没有完全回归。最后图 8 是 DISAPre-Net 预测出的结果, 漫反射的颜色和纹理细节提取都很接近真实值, 足以说明 Inception 和 Dropout 可以实现对模型预测能力的强化。

为了比较模型预测值和实际值, 本文用 MSEloss 函数直观观察三个网络的性能。横坐标表示迭代次数, 纵坐标表示 loss 值即预测值与实际数值的误差, 曲线越平滑幅度越小说明模型的预测能力越好。Pre-Net、SAPre-Net、DISAPre-Net 的 loss 曲线如下:

图 9 为 Pre-Net 的 loss 曲线, 整体来看曲线波动较大, 收敛较慢, 因为第一阶段的训练是为了让模型

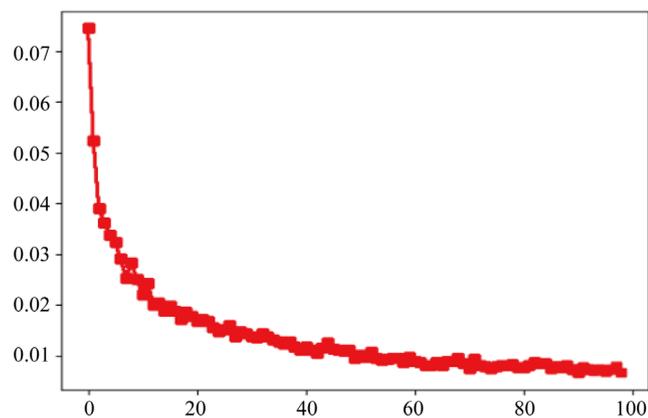


Figure 9. Pre-Net loss curve
图 9. Pre-Net loss 曲线

具备初步的预测能力。实验设置 100 K 次迭代, 迭代完成后由图 6 知网络已经有了一定的参数提取的性能。图 10 为 SAPre-Net 的 loss 曲线, 对比图 11 的 DISAPre-Net, SAPre-Net 也会出现了一定的过拟合现象, 并且相对于添加 Inception 和 Dropout 后的 DISAPre-Net, 可以看出在第 50 K 次迭代的时候, 前者的 loss 值在 0.005, 后者的 loss 在 0.001, 足以说明 DISAPre-Net 在开始阶段就已经表现出很好的收敛情况, 并且曲线整体平滑。这就可以提前结束第一阶段进行第二阶段, 从而缩短整体训练时间。准确来说, 训练时间由 12 小时缩短到了 9 小时。

Loss 曲线结合预测可视化的情况, 也说明了 Inception 和 Dropout 在对模型性能和训练时间上的有效性。

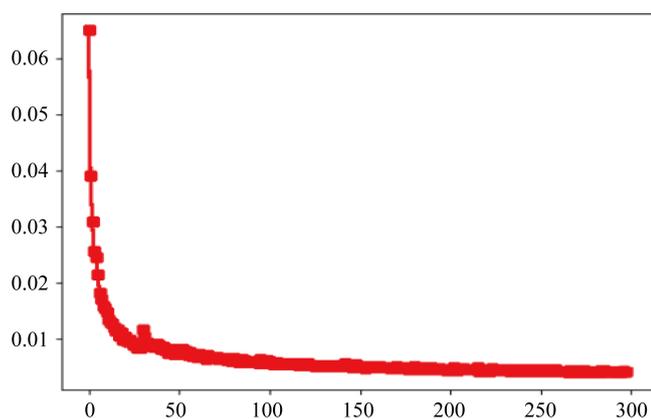


Figure 10. SAPre-Net loss curve

图 10. SAPre-Net loss 曲线

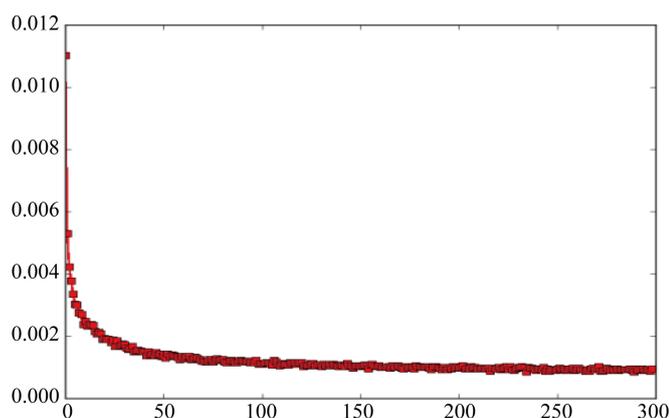


Figure 11. DISAPre-Net loss curve

图 11. DISAPre-Net loss 曲线

5.3. 反射分量拆分训练

实验中默认同时训练四个参数, 本文猜想多参数同时进行训练时是否参数间会有相互影响, 考虑单个参数输入网络学习, 本文尝试单独训练漫反射和法向。DISAPre-Net 同样配置 128 个时期(epoch), 其他参数也保持相同。图 12(a)展示了漫反射的 loss 曲线, 图 12(b)为法向的 loss 曲线。

图 12 中两个训练参数均在 0.001 处收敛。与 DISAPre-Net 的多参数同时训练结果大致相同, 并不能证明多参数同时训练存在相互影响。下面展示预测图:

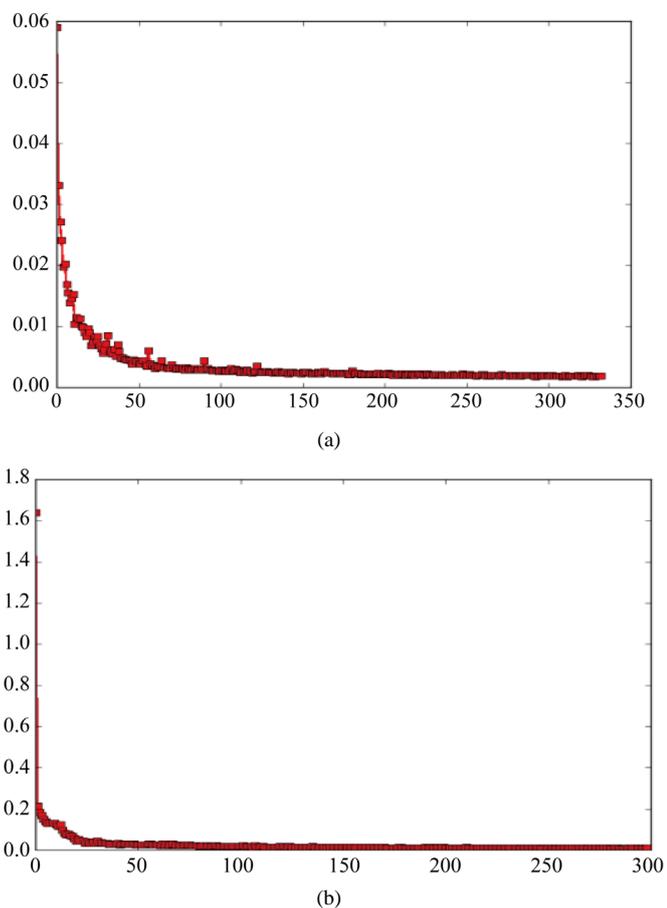


Figure 12. Convergence of loss function in single parameter training
图 12. 单参数训练的 loss 函数收敛情况

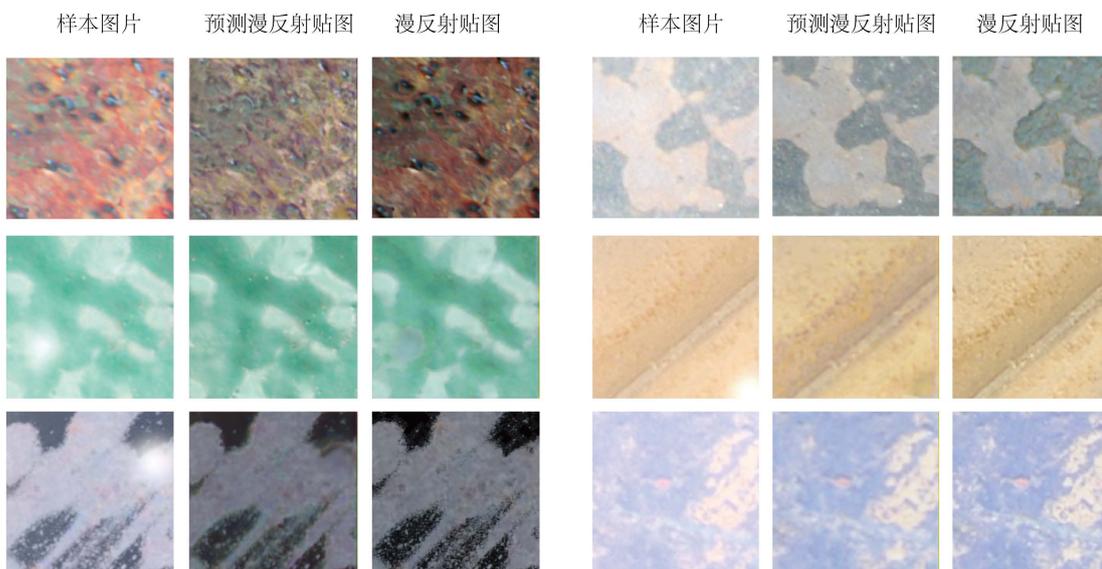


Figure 13. Prediction of diffuse reflection alone
图 13. 单独漫反射预测

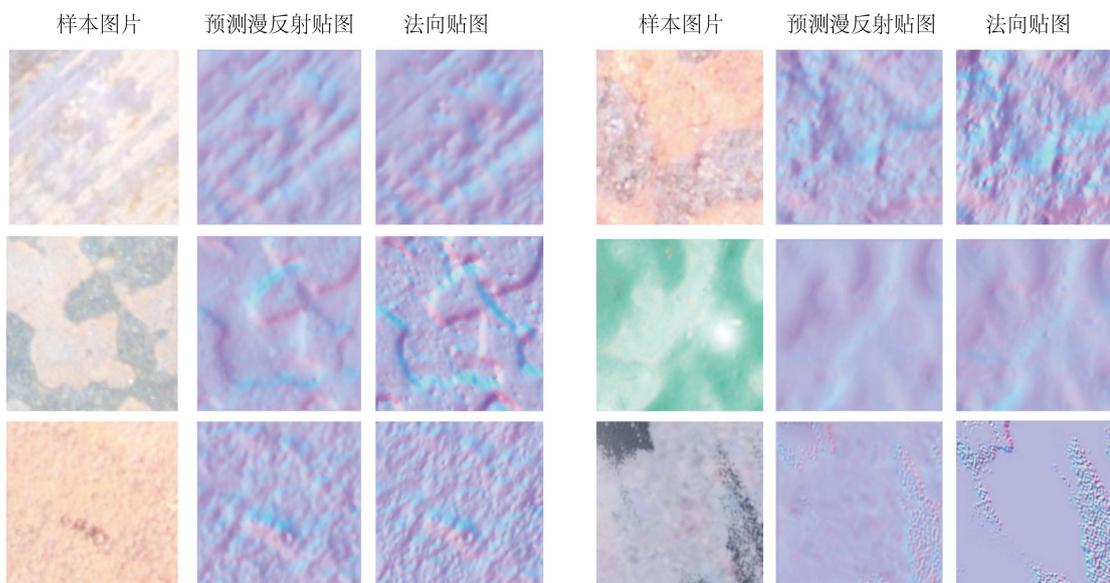


Figure 14. Prediction of normal reflection alone
图 14. 单独的法向预测

根据图 13 和图 14 预测结果和参照值之间的对比, 能够看出几乎所有的特征提取值都与参照值近乎相同。对比四个参数同时训练的参数字提取结果, 虽然 loss 曲线收敛结果貌似并无不同, 但从物理外观上看, 预测的结果与实际基本一致。可以这样认为, 单独训练一个参数, 模型的性能会优于多个参数模型, 但综合来说, 在时间上单独训练一个参数超过了 5 个小时, 四个参数单独训练完毕远超过共同训练的时间。并且第二组对比实验中模型的性能从预测结果来看已经足够好。所以还是建议多参数共同训练。

表 2 中 roughness 列为粗糙度的提取值, rough-ref 列为粗糙度实际值。Spec 列为高光的提取值, Spec-ref 列为高光实际值。从数据上可以看出经过自增强训练后, 网络性能有大幅度提升。并且模型经过 Dropout 和 Inception 优化后预测结果也更加精准。

Table 2. Prediction results of three kinds of networks for specular and roughness
表 2. 三类网络对高光和粗糙度的预测结果

	Roughness	Rough-ref	Spec	Spec-ref
Pre-Net	0.278	0.175	0.347	0.275
SAPre-Net	0.236	0.175	0.316	0.275
DISAPre-Net	0.198	0.175	0.300	0.275

5.4. 反射分 DISAPre-Net 的材质特征提取结果

对于同一类材质, 比如木头, 已经包含了足够多的纹理信息, 所以本文将材质分为木头、金属和塑料三类, 使用最终网络 DISAPre-NET 分别学习并测试。

图 15~图 17 中, 第一列为训练数据, 第二列开始, 上半部分分别为为漫反射、法向、粗糙度和高光的参照值, 下半部分分别是漫反射、法向、粗糙度和高光的预测值。实验中不能排除偶尔相差太大的高光预测的“脏”数据, 总体来说三种材质类型的预测和参照是是比较准确的。

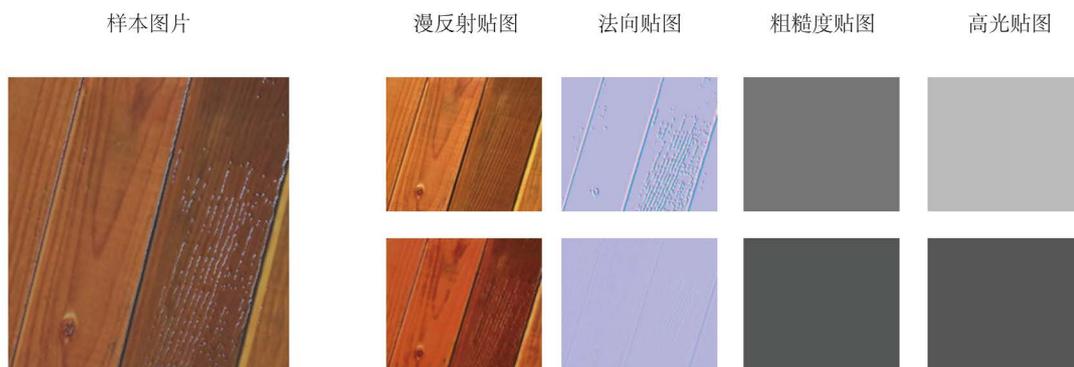


Figure 15. Wood

图 15. 木材

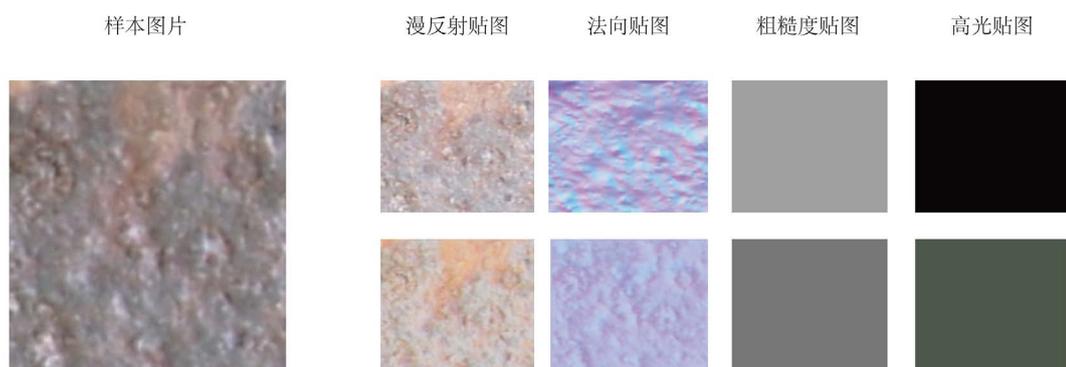


Figure 16. Metal

图 16. 金属

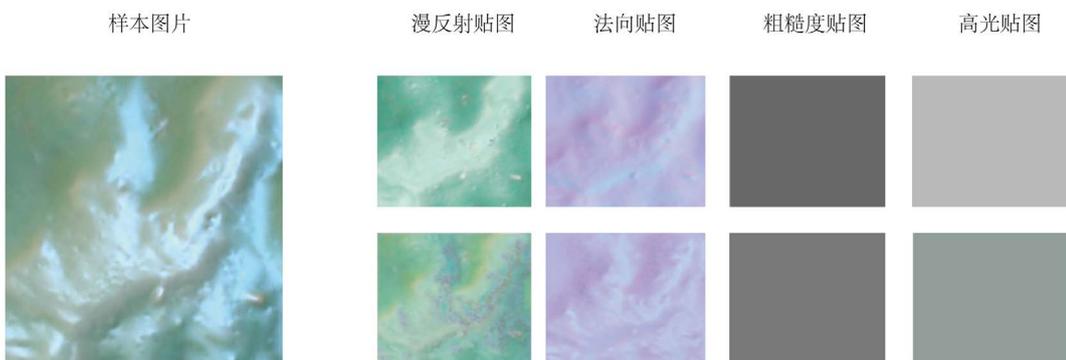


Figure 17. Plastic

图 17. 塑料

6. 结束语

本文在已有材质建模研究方法的基础上提出了一种基于 Pre-Net 的建模研究方法。首先建立渲染环境, 将平面材质图片和对应的表现纹理属性在不同光照环境下渲染出来, 组成新的标记训练对, 用来扩充训练数据集, 保证尽可能多的覆盖材质的空间变化特征。其次搭建一个能从单张未知光照环境下的平面样本材质图片中重建出材质表现属性的卷积神经网络, 该网络是经过 Dropout 和 Inception 优化后可以提高预测精度和避免过拟合问题。最后通过实验对比, 验证了自增强策略是一种有效策略。

参考文献

- [1] Xiao, L., Yue, D., Pieter, P. and Xin, T. (2017) Modeling Surface Appearance from a Single Photograph Using Self-Augmented Convolutional Neural Networks. *ACM Transactions on Graphics*, **36**, 1-45, 11. <https://dl.acm.org/doi/abs/10.1145/3072959.3073641>
<https://doi.org/10.1145/3072959.3073641>
- [2] Sean, B., Paul, U., Noah, S. and Kavita, B. (2013) OpenSurfaces: A Richly Annotated Catalog of Surface Appearance. *ACM Transactions on Graphics*, **32**, 1-111, 17. <https://doi.org/10.1145/2461912.2462002>
- [3] Dror, R.O., Adelson, E.H. and Willsky, A.S. (2001) Recognition of Surface Reflectance Properties from a Single Image under Unknown Real-World Illumination. *Proceedings of IEEE Workshop on Identifying Objects across Variations in Lighting: Psychophysics and Computation*. <http://dspace.mit.edu/bitstream/handle/1721.1/6664/AIM-2001-033.pdf?sequence=2>
- [4] Zhang, R., Zhu, J.-Y., Isola, P., Geng, X.Y., Lin, A.S., Yu, T.H. and Efros, A.A. (2017) Real-Time User-Guided Image Colorization with Learned Deep Priors. *ACM Transactions on Graphics*, **36**, 4. <https://doi.org/10.1145/3072959.3073703>
- [5] Ren, P.R., Wang, J.P., Snyder, J., Tong, X. and Guo, B.N. (2011) Pocket Reflectometry. *ACM Transactions on Graphics*, **30**, 4. <https://dl.acm.org/doi/abs/10.1145/1964921.1964940> <https://doi.org/10.1145/2010324.1964940>
- [6] Oxholm, G. and Nishino, K. (2012) Shape and Reflectance from Natural Illumination. *ECCV*, 528-541. https://doi.org/10.1007/978-3-642-33718-5_38
- [7] Aittala, M., Aila, T. and Lehtinen, J. (2016) Reflectance Modeling by Neural Texture Synthesis. *ACM Transactions on Graphics*, **35**, 1-65, 13. <https://dl.acm.org/doi/abs/10.1145/2897824.2925917>
<https://doi.org/10.1145/2897824.2925917>
- [8] Matusik, W., Pfister, H., Brand, M. and McMillan, L. (2003) Efficient Isotropic BRDF Measurement. In *Proceedings of the 14th Eurographics Workshop on Rendering (EGRW'03)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 241-247. <https://dl.acm.org/doi/abs/10.5555/882404.882439>
- [9] Wang, J.P., Zhao, S., Tong, X., Snyder, J. and Guo, B.N. (2008) Modeling Anisotropic Surface Reflectance with Example-Based Microfacet Synthesis. *ACM SIGGRAPH 2008 Papers (SIGGRAPH'08)*. ACM, New York, NY, Article 41. <https://dl.acm.org/doi/abs/10.1145/1399504.1360640>
<https://doi.org/10.1145/1399504.1360640>
- [10] Zhou, Z.M., Chen, G.J., Dong, Y., Wipf, D., Yu, Y., Snyder, J. and Tong, X. (2016) Sparse-as-Possible SVBRDF Acquisition. *ACM Transactions on Graphics*, **35**, Article 189. <https://doi.org/10.1145/2980179.2980247>
<https://dl.acm.org/doi/abs/10.1145/2980179.2980247>