



## 融合递归自编解码器的蒙卡画面重构降噪方法

卢娟, 陈纯毅

引用本文:

卢娟,陈纯毅. 融合递归自编解码器的蒙卡画面重构降噪方法[J]. 智能系统学报, 2023, 18(3): 459–467.

LU Juan, CHEN Chunyi. A denoised method by fusing recursive auto-encoder decoder for Monte Carlo rendering image reconstruction[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(3): 459–467.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202112011>

## 您可能感兴趣的其他文章

### 采用编码输入的生成对抗网络故障检测方法及应用

Fault detection method and its application using GAN with an encoded input

智能系统学报. 2022, 17(3): 496–505 <https://dx.doi.org/10.11992/tis.202102003>

### 利用残差密集网络的运动模糊复原方法

Image restoration with residual dense network

智能系统学报. 2021, 16(3): 442–448 <https://dx.doi.org/10.11992/tis.201912002>

### 空洞卷积与注意力融合的对抗式图像阴影去除算法

An antagonistic image shadow removal algorithm based on dilated convolution and attention mechanism

智能系统学报. 2021, 16(6): 1081–1089 <https://dx.doi.org/10.11992/tis.202011022>

### 基于风格转换的无监督聚类行人重识别

Clustering approach based on style transfer for unsupervised person re-identification

智能系统学报. 2021, 16(1): 48–56 <https://dx.doi.org/10.11992/tis.202012014>

### 基于生成式对抗网络的道路交通模糊图像增强

Enhancement of blurred road-traffic images based on generative adversarial network

智能系统学报. 2020, 15(3): 491–498 <https://dx.doi.org/10.11992/tis.201903041>

### 基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks

智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

DOI: 10.11992/tis.202112011

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20230323.1334.004.html>

# 融合递归自编解码器的蒙卡画面重构降噪方法

卢娟, 陈纯毅

(长春理工大学 计算机科学技术学院, 吉林 长春 130022)

**摘要:** 针对生成对抗模型降噪结果出现的伪影和模糊问题, 提出一种基于对抗生成递归自编解码器的蒙卡渲染画面降噪方法。在模型上, 设计多尺度卷积编码结构, 以多尺度残差自编码模型为生成器, 通过组合连接实现不同层次特征提取, 融合不同感受野的特征信息。以设计的递归残差网络模型为判别器, 判断蒙卡渲染画面真伪, 在对抗中提高网络性能。以端到端的方式将辅助信息特征、含有噪声的蒙卡画面和 8 192 采样率下的画面输入到融合递归自编解码器中进行降噪处理。实验表明, 该方法在测试场景下的平均峰值信噪比为 32.44 dB, 比生成对抗网络方法和残差网络方法分别提升 4.80% 和 3.13%; 平均结构相似性为 0.92, 比 2 种已有的算法分别提高 2.54% 和 1.01%。

**关键词:** 递归残差网络; 多尺度卷积; 生成对抗网络; 蒙特卡罗渲染画面; 图像处理; 图像去噪; 深度学习; 自编码网络

中图分类号: TP183 文献标志码: A 文章编号: 1673-4785(2023)03-0459-09

中文引用格式: 卢娟, 陈纯毅. 融合递归自编解码器的蒙卡画面重构降噪方法 [J]. 智能系统学报, 2023, 18(3): 459-467.

英文引用格式: LU Juan, CHEN Chunyi. A denoised method by fusing recursive auto-encoder decoder for Monte Carlo rendering image reconstruction[J]. CAAI transactions on intelligent systems, 2023, 18(3): 459-467.

## A denoised method by fusing recursive auto-encoder decoder for Monte Carlo rendering image reconstruction

LU Juan, CHEN Chunyi

(School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China)

**Abstract:** Aiming at the problem of artifacts and blurring in the noise reduction results of generative adversarial model, a noise reduction method for Monte Carlo rendering image based on generative adversarial recurrent auto-encoder is proposed in this paper. On the model, the multi-scale convolutional coding structure is designed, and the multi-scale residual auto-encoder model is used as the generator. The feature extraction at different levels is realized by combination connection, fusing the feature information of different receptive fields. The designed recursive residual network model is used as the discriminator to judge the authenticity of the Monte Carlo rendering picture and improve the network performance in the confrontation. In an end-to-end way, the auxiliary information features, the noise-containing Monte Carlo pictures and the pictures with the sampling rate of 8 192 are input into the recursive auto-encoder decoder for noise reduction. Experiments show that the average peak signal-to-noise ratio of this method is 32.44 dB in the test scenario, which is 4.80% and 3.13% higher than the generative adversarial network method and the residual network method, respectively. The average structural similarity is 0.92, which is 2.54 % and 1.01 % higher than the existing two algorithms, respectively.

**Keywords:** recursive residual network; multi-scale convolution; generative adversarial networks; Monte Carlo rendering screen; image processing; image denoising; deep learning; auto-encoder network

收稿日期: 2021-12-06. 网络出版日期: 2023-03-24.

基金项目: 吉林省科技发展计划项目 (20190302113GX).

通信作者: 陈纯毅. E-mail: [2306915844@qq.com](mailto:2306915844@qq.com).

随着图形渲染软件硬件的发展, 光线追踪技术逐渐在计算机辅助设计、动画电影制作、图形

可视化显示等领域崭露头角。针对蒙特卡罗三维渲染画面噪声的解决方法有 2 种:第 1 种是通过重要性采样对蒙特卡罗样本进行近似,加快蒙特卡罗收敛速率,如 Lehtinen 等<sup>[1]</sup>提出利用图像梯度的 Metropolis 光线跟踪算法,通过直接计算图像梯度值,求解 Poisson 方程来重建图像。第 2 种方法是对蒙特卡罗渲染的画面进行重构降噪,如 Rousselle 等<sup>[2]</sup>提出基于非局部均值滤波(non-local means, NLM)的去噪算法,该方法利用真实图像和噪声图像滤波系数的相关性,进行自适应滤波。

基于像素的蒙特卡罗深度学习降噪方法主要分为基于预测滤波参数的降噪方法、基于内核预测的降噪方法和基于像素值预测的降噪方法。Kalantari 等<sup>[3]</sup>利用滤波器的权重和蒙卡渲染画面噪声之间的关系,设计的一种基于滤波核的网络模型,该模型基于多层感知机模型(multi-layer perceptron, MLP)方法对画面进行降噪,当训练完成后会根据已有的场景滤波参数来对其他场景进行降噪。该方法使用固定的滤波器作为后端处理,但也继承了固有滤波器带来的限制。Meng 等<sup>[4]</sup>采用参数估计网络来解决路径跟踪的噪声伪影问题,该方法包括采样和重建阶段,重建阶段采用改进的 MLP 网络,利用提取的特征预测最终图像各向异性滤波器的最佳重建参数。

Chaitanya 等<sup>[5]</sup>提出循环自编码器的降噪方法,该网络结构为对称循环自编码网络结构,同时采用跳跃结构连接解码层,目的是为保留更多高频特征,该网络模型主要降噪对象为交互渲染场景画面,同时利用粗糙度特征进行降噪;但该方法主要针对时间序列关联度较高的图像降噪效果显著。谢川等<sup>[6]</sup>研究了深度自编码网络降噪方法,该网络模型使用对称跳跃连接的自编码器作为生成模型,并使用深度卷积网络作为判别模型来处理低采样下渲染的噪声数据,但该方法不针对运动模糊场景。Vogels 等<sup>[7]</sup>提出基于核预测的卷积网络模型,该方法利用源编码器和特征提取空间对辅助信息进行预处理,该方法使用源感知编码器提取低级特征,以及使用空间和时间模块提取抽象的高级特征。Wong 等<sup>[8]</sup>提出基于残差网络去除蒙卡画面中的噪声,该方法通过加深网络深度的方式对蒙卡画面进行降噪,可能会丢失过多的细节,导致画面过度平滑,同时也会出现拟合和训练困难等问题。Xu 等<sup>[9]</sup>提出利用生成对抗网络模型进行降噪,该方法将含有噪声的 RGB 图像分离为镜面反射图像(diffuse)和高光图

像(specular),同时输入辅助信息作为网络训练的数据,增加编解码器进行预处理,最后通过判别网络对编码网络的参数进行监督和训练获得重构后的噪声图像,该方法首次将分离的镜面和漫反射特征进行单独处理。

针对 U-Net 自编码网络模型处理 RGB 和辅助信息时会出现伪影和过模糊现象;同时端到端结构提取的图像特征会随着网络深度的变化而发生变化,导致提取特征的能力受限。本文提出一种基于对抗生成递归自编码器的蒙卡渲染画面降噪方法,在结果上能明显优于端到端的对抗生成模型,在降噪能力上也明显高于 U-Net 模型。

## 1 递归自动编解码网络模型

### 1.1 算法原理

机器学习降噪方法是将含有噪声的画面和参考画面输入到网络中学习,从而得到最优网络模型进行去噪,其原理表示为<sup>[10-11]</sup>

$$\theta' = \arg \min_{\theta} \frac{1}{N} \sum_{n=1}^N \ell(r^n, g(X^n, \theta))$$

式中: $\theta$ 是降噪网络模型权重, $(X^1, r^1), (X^2, r^2), \dots, (X^N, r^N)$ 表示是  $N$  个训练样本, $r^n$ 表示参考画面;通过模型预测得到  $g(X^n, \theta)$ ,  $\ell(\cdot)$ 表示损失函数。网络模型中使用的辅助信息包含深度、法向量和反照率等。本文网络模型为降噪自编码网络和递归网络模型,自编码网络采用编码函数和解码函数映射的方法得到初始的降噪画面,利用蒙卡噪声画面和辅助信息训练网络模型。如图 1 所示,上述框架第 1 部分采用 U-Net 网络,框架第 2 部分采用深度递归级联卷积神经网络。将含有噪声的渲染画面和辅助信息作为模型输入,利用编码函数提取高层特征,解码函数将特征图映射至输出层,得到输出重构画面。将上述结果和噪声图像进行融合,将其输入到递归级联卷积神经网络模型,进行训练。基于 U-Net 编解码模型中存在如下问题:

1) 当训练样本不足或场景信息丢失较严重时,模型无法根据特征信息进行画面还原,画面重构效果较差。

2) 在进行深层次特征提取后采样时,容易丢失边缘等重要特征,采样过程会导致信息的完整性低等问题,重构效果不明显。

3) 在解码层使用上采样还原图像时,在还原画面时会出现细节的缺失,会引起还原的画面信息丢失,导致画面出现模糊等问题。

4) 深度残差网络中参数过多会导致拟合困难, 同时也会降噪后出现过平滑等问题。

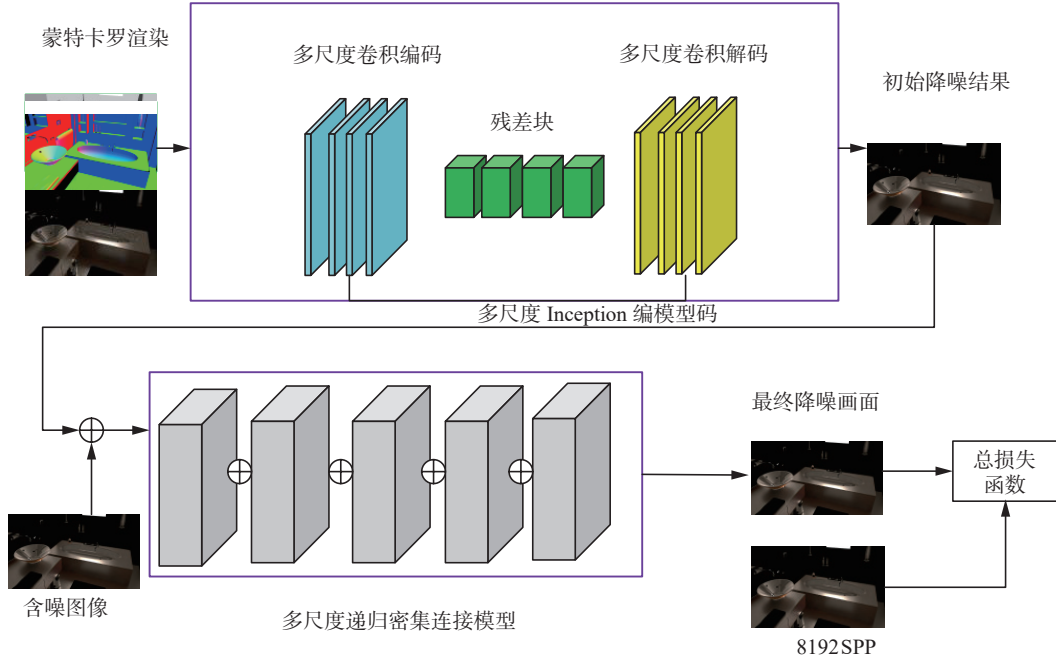


图 1 融合递归自编码模型框架

Fig. 1 Frame of fusion recursive auto-encoder model

本文针对上述问题进行多尺度卷积, 将特征信息进行融合, 提升画面信息质量; 针对池化过程出现的细节损失问题, 采用联合最大池化层 (maxpool) 和平均池化 (avgpool) 的方法, 进行空间维度的压缩, 并采用全连接层将 2 个池化层进行连接, 通过该方法来调节权重; 在解码层使用反卷积代替上采样, 减少图像还原时的细节丢失, 同时增加 skip 跳跃连接的方式来保留画面细节。将深度残差网络模型改为递归残差模型来提升网络表达能力。

## 1.2 递归网络模型

在深度残差网络模型<sup>[12]</sup>中, 往往采用加深网络深度的方法解决降噪问题, 该方法网络参数和网络表达能力难以平衡, 但引入递归结构可以很好的解决这个问题。本文改进深度残差网络, 改进后的递归结构模包含多个卷积层和 LReLU 激活函数, 在递归模型重复调用的次数设为  $T$ , 若将一个递归结构表示为映射函数  $G_M(\cdot)$ , 第  $t$  次到第  $t+1$  次的递归表示:

$$F_{t+1} = G_M(F_t)$$

式中:  $F_t$  表示第  $t$  次输入递归结构的特征图, 将重复  $T$  次的递归结构展开, 使多个网络层共享权值, 能够获得较好的网络表达能力, 同时减少网络参数的数量。

整个模型是由多个模块级联而成, 每个模块包含一个递归卷积神经单元 (recursive CNN unit,

RCU)。如图 2(a) 所示, 在级联卷积神经网络之间采集到输入的特征图后, 更新每个卷积神经网络单元的输出特征。每个卷积层的卷积核大小为  $3 \times 3$ , 在每个 CNN 单元, 每层的输出通道数为 64, 网络使用 LeakyReLU 激活函数, 进行非线性激活。递归残差块结构如图 2(b) 所示。递归模型的收敛速率较快, 经过  $T$  次递归操作, 对含有噪声的画面进行降噪处理, 输出降噪后的画面。

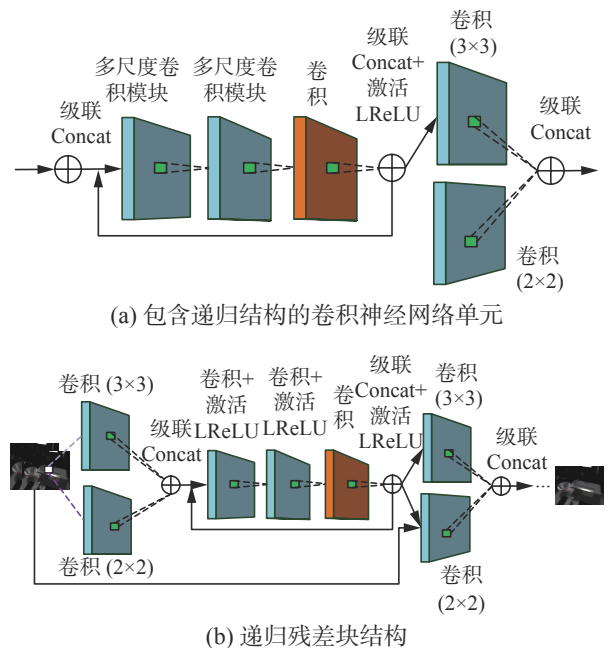


图 2 递归网络结构

Fig. 2 Recursive network structure



### 1.3 自动编解码模型

为提高 U-Net 自编码网络模型泛化能力, 将 U-Net 模型分为编码层、残差块层和解码层<sup>[13-14]</sup>。网络具有的优点是: 网络结构能对不同的蒙特卡罗画面具有普适性, 对蒙特卡罗渲染画面还原度高。

#### 1) 编码层 Encoder

①如图 3(a) 所示, 第 1 层编码器由 2 个不同尺度卷积层, 1 个池化层、标准化层和 1 个 ReLU 激活层构成。第 1 层卷积层为  $3 \times 3 \times 32$ , 第 2 层卷积层为  $1 \times 1 \times 64$ , 采用多尺度卷积操作, 有助于多个卷积核提取图像不同尺寸的信息, 将提取的信息进行融合, 可获得更好的表征。池化层采用联合池化和平均池化的方式, 进行空间维度压缩, 并采用全连接层将 2 个池化层进行连接, 池化层大小为  $2 \times 2$  和卷积层  $2 \times 2 \times 64$ , stride 设置为 1。每层后加入批归一化, 使用激活层采用 ReLU 函数防止梯度消失, 最后由 Concat 层进行连接。经过编码块输出的第一层通道数为 64。

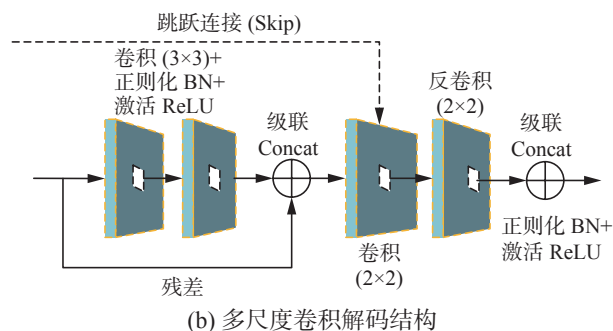
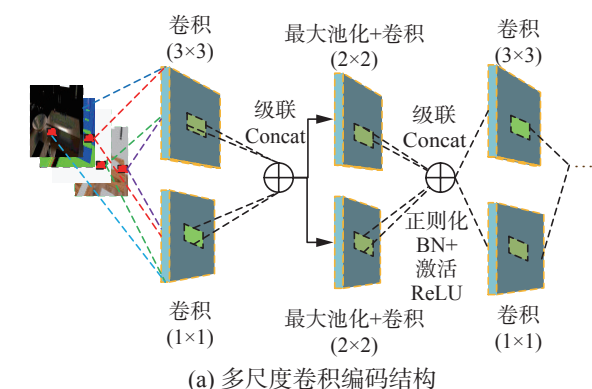


图 3 递归残差块结构

Fig. 3 Recursive residual block structure

②第 2~4 层编码层均设有 2 个不同尺度的卷积层、1 个池化层、1 个批归一化层和 ReLU 激活层。第 2 个编码块的卷积层分别为  $3 \times 3 \times 64$  和  $1 \times 1 \times 128$ , 经过第 2 个编码块的通道数为 128。第 3 个编码块的卷积层分别为  $3 \times 3 \times 128$  和  $1 \times 1 \times 256$ , 第 3 层编码块的总通道数为 256。第 4 层编码块

为  $3 \times 3 \times 256$  和  $1 \times 1 \times 512$ , 总通道数为 512。池化层的大小为  $3 \times 3$ , 滑动步长设为 1。

#### 2) 残差结构 (resblock)<sup>[15]</sup>

如图 3(b) 所示, 在自编码器中增加残差块进行权重共享, 每个残差块由 2 个卷积层为  $3 \times 3 \times 512$ 、正则化和 ReLU 激活层组成。每个残差块输出的通道数为 512, 滑动步长设置为 1。每层后加入批归一化, 使用激活层采用 ReLU 函数防止梯度消失, 最后由 Concat 层进行连接。

#### 3) 解码层 Decoder

①解码器对画面进行还原, 通过 skip 的方式进行特征信息的保留。解码层往往采用反卷积的方式对空间维度进行还原, 采用 skip 跳跃连接的方式进行权重信息共享, 提升网络细节保留的能力, 卷积层设定为  $2 \times 2 \times 512$ , 滑动步长设定为 2, 通道数为 512, 加入实例归一化和 ReLU 激活函数。该方法能够有效的进行特征信息的融合, 提升降噪的质量。

②在第 2~4 层的解码块中, 反卷积大小分别为  $2 \times 2 \times 256$ 、 $2 \times 2 \times 128$  和  $2 \times 2 \times 64$ , 最后使用 SoftMax 进行输出。使用反卷积实现上采样层, 使用上采样层将解码层的图像进行还原, 将图像恢复到原尺寸。

自编码网络因输入的噪声带有不同特性, 特征提取时容易忽略浅层特征; 利用噪声画面和辅助信息重构画面时会导致降噪效果较差。综上所述, 在 U-Net 自编码网络模型上, 增加多尺度卷积和联合最大池化和平均池化来保证信息的完整度, 在反卷积前增加残差块进行参数共享。在解码层使用反卷积代替上采样, 减少图像还原时的细节丢失。

### 1.4 辅助特征信息

在使用蒙特卡罗路径追踪算法渲染三维场景时, 需要从相机出发记录每一条光线。当渲染时记录光线第 1 次与三维模型相交的信息时, 将该信息保存到几何缓冲区 (G-buffer), 其保存的信息为相交点所在的材质等信息 (法向量、反射率、深度等信息)。

如图 4 所示, 本文参考文献 [9] 的辅助信息选择, 采用反射率 (3 维), 深度 (1 维) 和法向量 (3 维), 加上画面本身颜色 (3 维) 总共 10 维的输入特征。对于深度信息, 将其归一化为  $[0, 1]$ , 通过输出结果证明加入辅助信息对重构画面的有效性, 场景 bathroom、classroom、cars、room3 和 staircase 的辅助信息<sup>[16-17]</sup>。

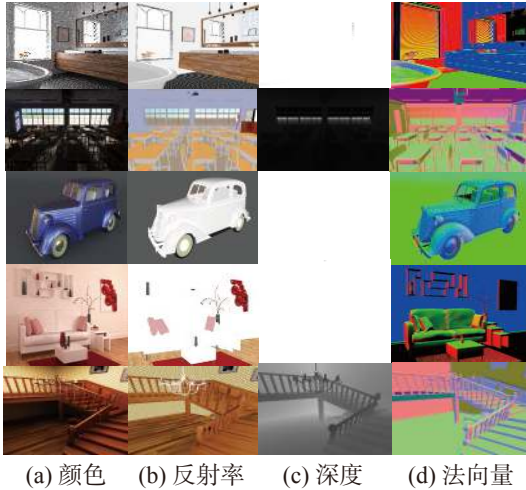


图4 辅助信息示例

Fig. 4 Exemplary diagram of auxiliary information

### 1.5 损失函数

网络模型的输出结果应和相应的参考图像相似。本文采用自编码器常用损失函数<sup>[18]</sup>,同时加入梯度函数<sup>[19]</sup>以及损失函数 $L_{GAN}$ <sup>[20-21]</sup>:

$$\min_G \max_D L_{GAN}(G, D)$$

优化过程会求解 $D$ 的参数值,使 $L_{GAN}$ 达到最大,然后固定 $D$ 求解 $G$ 使得 $L_{GAN}$ 达到最小,其公式为<sup>[21]</sup>

$$G^* = \arg \min_G \max_D L_{GAN}(G, D)$$

式中: $G^*$ 表示生成器具有降噪图像的模型参数,当 $D$ 最大时,会使得 $G^*$ 具有合理的模型参数。 $L_{GAN}(G, D)$ 可写为<sup>[20]</sup>

$$L_{GAN}(G, D) = E_{X \sim P_X} [\ln D(X)] + E_{r \sim P_r} [\ln(1 - D(G(r)))]$$

式中: $E_{X \sim P_X}$ 表示参考图像样本的期望; $E_{r \sim P_r}$ 指生成图像样本的期望。总损失函数 $L_{total}$ 为

$$L_{total} = \lambda_1 L_H + \lambda_2 L_t + \lambda_3 L_{GAN}$$

其中, $L_H$ 为对称绝对比误差损失, $L_t$ 为像素损失, $\lambda_1$ 、 $\lambda_2$ 和 $\lambda_3$ 为平衡参数。通过上述损失函数使整个网络协同工作,提升最后降噪的质量。

## 2 实验结果与分析

本实验分别在8个场景数据集上进行训练和测试,验证集为公开的迪士尼场景以及PBRT渲染数据集。

### 2.1 数据集及实验环境

训练数据集为迪士尼数据集,使用了8个场景,分别是bathroom、car、classroom、house、dining-room、bedroom、spaceship和staircase。如图5所示,该数据集共含有1482幅多通道高分辨渲染图,每幅画面的分辨率均为1280×720,其中含有噪声的渲染画面的像素采样数分别为256、512、1024和4096,参考画面的像素采样数为8192。



图5 数据集

Fig. 5 Datasets

本实验选用car、classroom、staircase、dining、bedroom和spaceship共6个场景作为训练数据集,参考画面的采样数为8192。测试数据集为公开数据集中的bathroom、house和PBRT渲染场景中的classroom、room3和car2。由于每个测试场景的像素采样数不同,本文将对公开数据集中的bathroom、house(采样数为512)和PBRT渲染的场景画面进行分析。实验环境如表1所示。

表1 实验环境

Table 1 Experimental environment

| 实验环境   | 参数                                    |
|--------|---------------------------------------|
| 设备型号   | Think Station P920                    |
| 操作系统   | Ubuntu 18.06版本                        |
| CPU处理器 | Intel Xeon Gold 5120T CPU 2.2 GHz x56 |
| GPU    | NVIDIA Quadro P5000                   |
| 机器学习环境 | Pytorch                               |
| CUDA版本 | V10.0                                 |
| 语言环境   | Python, C++                           |

### 2.2 评价指标及参数设置

本文用峰值信噪比(peak signal to noise ratio, PSNR)<sup>[22]</sup>和结构相似度(structural similarity, SSIM)<sup>[23]</sup>作为蒙特卡罗渲染降噪结果的客观评价量化指标。PSNR对噪声敏感,它是基于均方误差的重构误差指标。一般地,当PSNR大于28时,画面质量差别不会太明显,在35~40时,人眼几乎分辨不出明显差异。SSIM以每个像素为中心的小块为单位进行对比,SSIM的取值范围为[0,1],SSIM值越高表示降噪画面质量越高。PSNR定义为<sup>[22]</sup>

$$R_{PSN} = 10 \lg \left( \frac{(2^n - 1)^2}{M} \right)$$



式中:  $n$  为 8,  $M$  为真实图像和降噪后图像均方误差。SSIM 定义为<sup>[23]</sup>

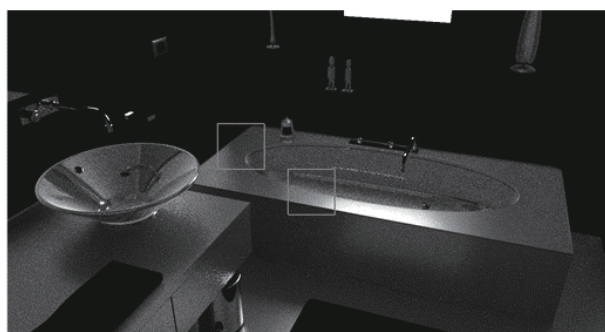
$$R_{SS} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

式中:  $\mu_x$  和  $\mu_y$  表示图像的灰度平均值,  $\sigma_{xy}$  表示 2 幅图像的协方差;  $\sigma_x$  和  $\sigma_y$  表示图像的标准差,  $C_1=6.5025$ ,  $C_2=58.5225$ 。

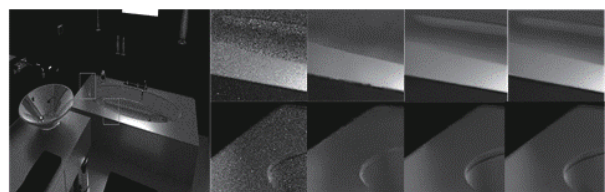
为提高网络的泛化能力,使网络具有更高的鲁棒性,本文采用迪士尼蒙特卡罗渲染数据集进行网络模型训练。本文网络模型将与深度残差网络模型和对抗生成网络模型在视觉效果、PSNR 和 SSIM 3 个方面进行对比。本实验使用的批处理(batch)大小为 5,总共训练迭代次数为 100,学习率为  $10^{-4}$ ,训练时间约为 6 h,在训练过程中针对损失函数来预定义权重系数,预定义权重系数为 1.0。

### 2.3 消融实验

本文对不含递归结构的自编码器网络进行消融实验分析,将含噪声的渲染画面和辅助信息输入到自编码网络中训练。选择的消融实验的场景画面为 bathroom,视觉效果如图 6 所示,不含递归结构的自编码网络降噪效果低于含有递归结构的自编码网络。



(a) 测试场景 (bathroom)



(a) 场景 (b) MC 渲染 (c) DAE 模型 (d) 本文 (e) 参考  
(b) 消融实验 (bathroom)



(a) 场景 (b) 残差网络 (c) 本文  
(c) 消融实验 (bathroom)

图 6 消融实验结果

Fig. 6 Results of ablation experiments

在降噪速度上,不含递归结构的自编码网络处理速度较快,算法的复杂度较低。不含递归结构和含递归结构的 PSNR 分别为 24.553 9 dB 和 27.913 2 dB, SSIM 值分别为 0.796 6 和 0.861 5。就计算效率而言,含递归的网络模型时间复杂度更高。

### 2.4 收敛性

深度学习对噪声画面进行降噪时,需要考虑训练模型是否存在欠拟合或过拟合等问题,而误差曲线能够更好地反应这一问题。图 7 给出了递归网络模型进行训练时的 PSNR 随迭代次数变化曲线,随着 epoch 数的增加,训练误差和验证误差均收敛,当迭代到 100 时,PSNR 收敛为 29.5 dB 左右,说明模型训练结果合理。

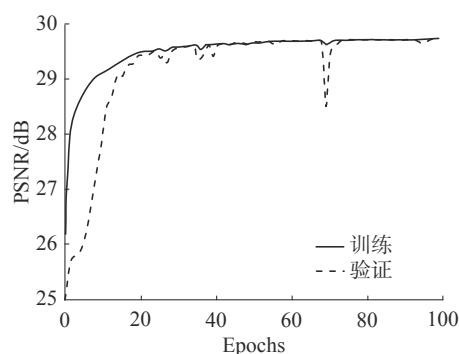


图 7 收敛曲线

Fig. 7 Diagram of convergence curves

### 2.5 视觉效果和评价指标

为更加客观地分析本文模型对比其他模型的降噪效果,本文选取对比的算法为基于残差网络模型和基于 U-Net 为基础的网络模型,分别为文献[8]和文献[9],选用场景为 classroom、room3、house、和 car2,降噪后的视觉效果为图 8 所示。文献[8]在重构时,池化和上采样过程会造成像素等重要信息丢失,导致画面产生伪影,而文献[9]对画面进行全局信息的保留,忽略了局部细节,本方法主要是在画面重构时减少重构时的像素损失和对噪声画面进行重构,在细节保留上能够比文献[8]和文献[9]效果好,同时加速网络收敛,得到更高的视觉效果。

本文选取的测试场景分别为迪士尼数据和 PBRT 渲染的画面。其中,表 2 为降噪后蒙特卡罗渲染画面的 PSNR 值,表 3 为降噪后的蒙特卡罗渲染的 SSIM 值,根据表 2 和表 3 的 PSNR 和 SSIM 进行分析得出,降噪后画面的 PSNR 和 SSIM 值相对于文献[8]和文献[9]得到一定提升,说明本文的降噪算法效果优于文献[8]和文献[9]的方法。



图 8 降噪后视觉效果

Fig. 8 Visual effect after performing denoising

表 2 不同方法和场景的平均 PSNR 评价指标

Table 2 Average PSNR evaluation index value under different methods and scenes

| 场景        | 文献[8]    | 文献[9]    | 本文              |
|-----------|----------|----------|-----------------|
| staircase | 30.995 5 | 32.431 9 | <b>32.674 3</b> |
| bathroom  | 25.208 7 | 25.562 7 | <b>27.913 2</b> |
| room3     | 28.375 3 | 28.720 5 | <b>30.528 1</b> |
| spaceship | 34.673 7 | 34.947 5 | <b>36.463 1</b> |
| classroom | 28.395 0 | 28.808 5 | <b>29.297 7</b> |
| car2      | 35.165 4 | 35.534 5 | <b>35.896 3</b> |
| house     | 33.876 1 | 34.178 0 | <b>34.311 7</b> |
| 平均        | 30.955 7 | 31.454 8 | <b>32.440 6</b> |

图 8 给出了文献 [8]、文献 [9] 的方法和本文的降噪算法计算结果的视觉差异。残差网络方法没有考虑场景细节信息, 导致场景过于模糊。例如, 对比的方法在“room3”场景的枕头和相框处呈现平

滑效果, 产生伪影, 图像出现模糊, 同时也出现锯齿现象。相比而言, 本文方法能够得到更清晰的图像。实验结果证明, 本文方法的主观视觉效果比其他方法能保留的细节更多, 视觉效果也更好。

表 3 不同方法和场景的平均 SSIM 评价指标

Table 3 Average SSIM evaluation index value under different methods and scenes

| 场景        | 文献[8]   | 文献[9]   | 本文             |
|-----------|---------|---------|----------------|
| staircase | 0.910 0 | 0.920 8 | <b>0.931 0</b> |
| bathroom  | 0.852 2 | 0.857 9 | <b>0.861 5</b> |
| room3     | 0.848 1 | 0.886 3 | <b>0.907 4</b> |
| spaceship | 0.938 8 | 0.957 7 | <b>0.962 6</b> |
| classroom | 0.855 6 | 0.867 6 | <b>0.892 4</b> |
| car2      | 0.950 9 | 0.953 7 | <b>0.958 8</b> |
| house     | 0.941 9 | 0.948 7 | <b>0.943 5</b> |
| 平均        | 0.899 6 | 0.913 2 | <b>0.922 5</b> |



### 3 结束语

本文使用自编码网络引入递归残差网络结构对蒙特卡罗渲染画面进行降噪,将深度、法向量以及反射率等辅助信息作为数据集,得到初始的重构图像。将获得的初始图像与噪声图像融合,融合后的图像在递归网络模型中进行降噪。本文在递归网络模型中减少了近一半的网络参数,与深度残差网络相比提升了网络降噪的速率。未来将考虑结合视觉感知来进行蒙特卡罗渲染画面的降噪,并设计局部视觉感知的网络,进一步研究视觉感知差异对降噪的效果。

### 参考文献:

- [1] LEHTINEN J, KARRAS T, LAINE S, et al. Gradient-domain metropolis light transport[J]. *ACM transactions on graphics*, 2013, 32(4): 1–12.
- [2] ROUSSELLE F, KNAUS C, ZWICKER M. Adaptive rendering with non-local means filtering[J]. *ACM transactions on graphics*, 2012, 31(6): 1–11.
- [3] KALANTARI N K, BAKO S, SEN P. A machine learning approach for filtering Monte Carlo noise[J]. *ACM transactions on graphics*, 2015, 34(4): 1–12.
- [4] MENG Xiaoxu, ZHENG Quan, Varshney A, et al. Real-time Monte Carlo denoising with the neural bilateral grid[C]//Eurographics Symposium on Rendering 2020. London: ESRL, 2020: 13–24.
- [5] CHAITANYA C R A, KAPLANYAN A S, SCHIED C, et al. Interactive reconstruction of Monte Carlo image sequences using a recurrent denoising autoencoder[J]. *ACM transactions on graphics*, 2017, 36(4): 1–12.
- [6] 谢川, 王勇超, 林志洁, 等. 基于对抗生成网络的蒙特卡罗噪声去除算法 [J]. *模式识别与人工智能*, 2018, 31(11): 1047–1060.  
XIE Chuan, WANG Yongchao, LIN Zhijie, et al. Monte Carlo noise removal algorithm based on adversarial generative network[J]. *Pattern recognition and artificial intelligence*, 2018, 31(11): 1047–1060.
- [7] VOGELS T, NOVÁK J, ROUSSELLE F, et al. Kerneel predicting convolutional neural networks for denoising [P]. US20180293711: 15/814190, 2018-11-10.
- [8] WONG K M, WONG T T. Deep residual learning for denoising Monte Carlo renderings[J]. *Computational visual media*, 2019, 5(3): 239–255.
- [9] XU Bing, ZHANG Junfei, WANG Rui, et al. Adversarial Monte Carlo denoising with conditioned auxiliary feature modulation[J]. *ACM transactions on graphics*, 2019, 38(6): 1–12.
- [10] VICINI D, ADLER D, NOVÁK J, et al. Denoising deep Monte Carlo renderings[J]. *Computer graphics forum*, 2019, 38(1): 316–327.
- [11] HUO Yuchi, YOON S E. A survey on deep learning-based Monte Carlo denoising[J]. *Computational visual media*, 2021, 7(2): 169–185.
- [12] 王婕, 罗静蕊, 岳广德. 一种改进的多尺度融合并行稠密残差去噪网络 [J]. *小型微型计算机系统*, 2021, 42(4): 798–804.  
WANG Jie, LUO Jingrui, YUE Guangde. Improved parallel residual dense denoising network based on multi-scale fusion[J]. *Journal of Chinese computer systems*, 2021, 42(4): 798–804.
- [13] RONNEBERGER O. Invited talk: U-net convolutional networks for biomedical image segmentation[C]//Bildverarbeitung für die Medizin 2017. Berlin: Springer Vieweg, 2017: 3–3.
- [14] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-ResNet and the impact of residual connections on learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. San Francisco: AAAI Press, 2017, 31(1): 11231.
- [15] WANG Tianyang, SUN Mingxuan, HU Kaoning. Dilated deep residual network for image denoising[C]//2017 IEEE 29th International Conference on Tools with Artificial Intelligence. Boston: IEEE, 2018: 1272–1279.
- [16] BENEDIKT Bitterli. Rendering resources[EB/OL]. [2021–12–06]. <https://benedikt-bitterli.me/resources/>.
- [17] LIN Weiheng, WANG Beibei, WANG Lu, et al. A detail preserving neural network model for Monte Carlo denoising[J]. *Computational visual media*, 2020, 6(2): 157–168.
- [18] SIMONYAN K, ZISSERMAN A. Two-stream convolutional networks for action recognition in videos[EB/OL]. (2014–06–09) [2021–12–06]. <https://arxiv.org/abs/1406.2199>.
- [19] KREINOVICH V, NGUYEN H, OUNCHAROEN R. How to estimate forecasting quality: a system-motivated derivation of symmetric mean absolute percentage error and other similar characteristics[EB/OL]. [2021–12–06]. <http://www.cs.utep.edu/vladik/2014/tr14-53.pdf>.
- [20] ISOLA P, ZHU Junyan, ZHOU Tinghui, et al. Image-to-image translation with conditional adversarial networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 5967–5976.
- [21] 许家荣. 核预测与 GAN 特征保持的蒙特卡罗渲染图像去噪 [D]. 广州: 华南理工大学, 2020.

XU Jiarong. Denoising of Monte Carlo renderings based on kernel prediction and feature-preserved GAN[D]. Guangzhou: South China University of Technology, 2020.

- [22] 佟雨兵, 张其善, 祁云平. 基于 PSNR 与 SSIM 联合的图像质量评价模型 [J]. 中国图象图形学报, 2006, 11(12): 1758–1763.

TONG Yubing, ZHANG Qishan, QI Yunping. Image quality assessing by combining PSNR and SSIM[J]. Journal of image and graphics, 2006, 11(12): 1758–1763.

- [23] 林湛黔, 张梦凯, 吴成茂. 利用残差密集网络的运动模糊复原方法 [J]. 智能系统学报, 2021, 16(3): 442–448.

LIN Zhenxian, ZHANG Mengkai, WU Chengmao. Image restoration with residual dense network[J]. CCAI transactions on intelligent systems, 2021, 16(3): 442–448.

#### 作者简介:



卢娟, 硕士研究生, 主要研究方向为深度学习、图像降噪。



陈纯毅, 教授, 博士生导师, 博士, 中国图象图形学学会虚拟现实专委会委员、中国计算机学会会员。主要研究方向为虚拟现实、特种电影和光电系统仿真; 承担过国家自然科学基金、国家重点研发计划、国家科技支撑计划、863 计划、973 计划等国家高层次项目, 发表学术论文 60 余篇。

## 2023 中国人工智能大会 Chinese Congress on Artificial Intelligence 2023

八闽大地, 群贤毕至; 数智交融, 造福于民。7 月 22—24 日, 由中国人工智能学会主办的 2023 中国人工智能大会 (Chinese Congress on Artificial Intelligence 2023, CCAI 2023) 将在福建福州举行。

以“数智领航·共筑未来”为主题, 大会将邀请海内外学者专家、业界大咖、科研人员同台交流, 展现人工智能关键领域的学术研究和科技创新成果, 引领我国数智科技创新趋势, 助力福州数字经济高质量发展。

作为我国最早发起举办的国际化人工智能会议, 中国人工智能大会自 2015 年创办以来连续举办 8 届, 每届参会人数达千人, 已成为我国人工智能领域规格最高、规模最大、影响力最强的学术交流活动之一, 已在北京、杭州、深圳、青岛、南京、成都和长沙等城市落地。

戴琼海院士、周志华院士、杨强院士等 3 位享誉国际的学者担任本届大会主席, 并将携手多位中外院士在内的 100 多位知名科学家、创业家、教育家共聚一堂, 在主旨报告、尖峰对话、专题论坛等系列活动上, 围绕人工智能理论研究、技术创新、行业应用、社会治理、人才培养、生态建设议题进行成果分享和观点碰撞。

在数字中国建设大潮下, 福州坚持数智引领, 正在抢抓新一轮数字经济先机。以本届大会为汇智聚才平台, 福州不仅将对外生动展示这座现代化都市的发展活力, 而且将向内汲取全国, 乃至全球的创新智慧。

数字时代, 智能科技已成为造福于民的关键因素, 当数智之光映射在有福之州, 将激发我们更深刻的思考和追寻科技进步的终极含义。

盛夏七月, 2023 中国人工智能大会携手福州, 邀您一起以智“福”人!

官方网址: <https://ccai.ccai.cn/>