

Real-time Monte Carlo Denoising with Weight Sharing Kernel Prediction Network

Dezeming Family

2023 年 7 月 3 日

正常字体：表示论文的基本内容解释。

粗体：表示需要特别注意的内容。

红色字体：表示容易理解错误或者混淆的内容。

蓝色字体：表示额外增加的一些注释。

绿色字体：表示额外举的一些例子。

目录

一 基本介绍	1
二 相关工作	1
三 问题描述和网络架构	2
四 局限性和问题	2
参考文献	2

一 基本介绍

实时蒙特卡罗去噪的目的是在严格的时间预算内，在低像素样本 (spp) 的情况下去除严重的噪声。最近，核预测方法使用神经网络来预测每个像素的滤波核，并显示出去除蒙特卡罗噪声的巨大潜力。然而，繁重的计算开销阻碍了这些方法的实时应用。

本文扩展了核预测方法，并提出了一种在实时帧率下对非常低 spp (例如, 1-spp) 的蒙特卡罗路径跟踪图像进行去噪的新方法。我们不是使用神经网络来直接预测内核映射 (kernel map)，即每个每像素滤波内核的完整权重，而是预测内核映射的编码，然后是具有展开操作的高效解码器，以实现滤波内核的高质量重建。内核映射编码产生了内核映射的紧凑单通道表示，这可以显著降低内核预测网络的吞吐量。此外，我们采用了可扩展的内核融合模块 (scalable kernel fusion module) 来提高去噪质量。

所提出的方法保留了核预测方法的去噪质量，同时对 1-spp 噪声输入的去噪时间大致减半。此外，与最近的基于神经双边网格的实时去噪器相比，我们的方法得益于基于核的重建的高度并行性，并在同等时间产生更好的去噪结果。

我们的方法的核心观点在于引入核构建模块来减少神经网络的吞吐量，我们不是直接预测每像素的滤波权重，而是将神经网络视为编码器，以输出滤波权重的紧凑单通道格式，表示为重要性图 (importance maps)。然后，我们将重要性映射提供给手工设计的解码器，以构建完整的内核映射。在实践中，我们利用展开 (unfolding) 操作和归一化步骤来构建滤波器内核构建模块，并使用端到端的训练方式使网络与这种手工制作的解码器协同工作。作为最后一步，我们使用构建的滤波内核来滤波有噪声的图像，这可以有效地实现屏幕空间后处理。我们的方法得益于基于核的重建的高度并行性，并且易于集成到现有的蒙特卡罗去噪模块中。

我们设计的解码架构能够构建特定大小的滤波内核。我们通过内核融合模块进一步提高了解码器的容量。具体地说，改进的解码器从一组重要性图中构造滤波核，然后将核独立地应用于有噪声的输入，并将去噪结果与学习的权重相结合。实际上，我们构造不同大小的滤波核以适应不同的噪声频率。此外，我们在去噪阶段之前添加了一个时间累积操作。输入图像的有效 spp 计数的增加可以帮助稳定连续帧的去噪过程并抑制时间伪影。综合实验结果表明，该方法在实时帧率下对 1-spp 帧具有良好的去噪效果。我们的方法可以保持核预测方法的去噪质量，同时将其去噪时间大致减半。此外，与最近的基于神经双边网格的实时去噪器相比，我们的方法得益于基于核的重建的高并行性，并在同等时间产生更好的去噪结果。总之，我们的方法做出了以下贡献：

- 一种新的内核预测架构，使用内核映射的紧凑表示和可扩展解码器进行滤波内核重建，显著降低了神经网络的吞吐量和内存需求。
- 与最先进的实时内核预测蒙特卡罗去噪方法相比，仅使用大约一半的去噪时间即可实现可比的渲染质量。

二 相关工作

Bako 的 kernel prediction 去噪方法，缺点是预测大的滤波内核是耗时和消耗内存的，而使用小尺寸的内核会导致质量下降。

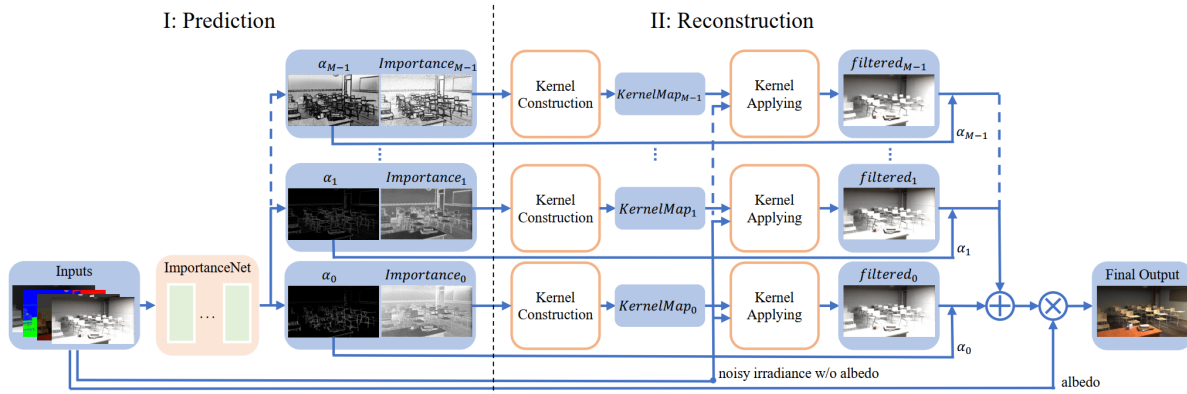
Vogels 的多分辨率滤波来近似大滤波核，考虑到在缩小分辨率下使用小尺寸内核的滤波对应于在原始分辨率的大感受野中的滤波。他们还使用特定任务模块分解去噪管道，独立提取源感知 (source-aware) 的和时空的信息。

Dahlberg 等人将这种模块化分层核预测方法部署到现有的商业蒙特卡罗路径跟踪引擎中，并展示了其在对故事片制作进行去噪方面的实用能力和灵活性。然而，最初的内核预测方法主要是为 16-64 spp 的离线应用程序设计的，并且在几秒钟内运行。Hasselgren 等人 and Munkberg 等人分别使用分层核预测架构对重新采样的蒙特卡罗图像和样本飞溅层进行去噪，并实现了交互速度。此外，Thomas 等人还利用具有特征提取网络的分层架构，该网络对量化误差具有弹性，以探索用于图像重建的轻量化网络的可行性。

与直接使用内核预测架构的方法不同，我们的方法通过对内核映射的编码进行操作，将其扩展到具有 1-spp 输入的实时去噪，以减少神经网络推理开销。

三 问题描述和网络架构

对于实时应用来说，直接预测每像素的大尺寸滤波核权重在网络推理中花费了太多时间。一种解决方案是修剪 (prune) 层次内核预测架构 [1] 以用于实时应用。然而，该解决方案的质量严重下降。我们观察到，原始核预测网络的总去噪时间很难缩短，主要是因为最后一个卷积层的高通量。例如，在输出层中有 169 个卷积核用于预测 13×13 大小的滤波核，这对于实时架构来说太重了。鉴于此，我们建议预测内核映射的编码紧凑表示，然后使用高效且可扩展的解码器来重建完整的内核映射。通过这种方式，我们可以避免预测过滤窗口大小为 k 的重 $k \times k$ 信道内核映射。相反，我们只需要预测单个信道内核映射编码，这大大降低了网络吞吐量。有了一个高效且可扩展的解码器，我们可以使用轻量级的神经网络编码器架构，使我们的去噪管道以实时速度运行，同时保持去噪质量。



ImportanceNet 可以有两种设置，要么是 6 层 RepVGG，要么是 3 层 RepVGG。网络会预测不同尺度的 M 个重要性图，这 M 个重要性图被用来预测核。关于核的编解码，其实就是一堆 CNN 卷积，比较简单。

四 局限性和问题

时序稳定性确实不太行，因为跟以前一样，就是先把时序帧通过重投影来获取，然后融合到当前帧，然后再对当前帧做去噪。

由于核映射相当于压缩预测核，所以不太能用来预测大核（比如 21×21 的大核）。

参考文献

[1] VOGELS T., ROUSSELLE F., MCWILLIAMS B., RÖTHLIN G., HARVILL A., ADLER D., MEYER M., NOVÁK J.: Denoising with kernel prediction and asymmetric loss functions.