Denoising and Guided Upsampling of Monte Carlo Path Traced Low Resolution Renderings

Kadir Cenk Alpay
Middle East Technical University
Ankara, Turkey
kadircenk@ceng.metu.edu.tr

Ahmet Oğuz Akyüz
Middle East Technical University
Ankara, Turkey
akyuz@ceng.metu.edu.tr

Figure 1: Our framework enables Monte Carlo denoisers to perform robustly when the samples per pixel (SPP) count is decreased harshly. (a) Result of the baseline approach with harshly decreased SPP. (b) Result of the proposed framework. (c) Ground-truth. Scene's renderer file courtesy of [Bitterli 2016].

ABSTRACT

Monte Carlo path tracing generates renderings by estimating the rendering equation using the Monte Carlo method. Studies focus on rendering a noisy image at the original resolution with a low sample per pixel count to decrease the rendering time. Image-space denoising is then applied to produce a visually appealing output. However, denoising process cannot handle the high variance of the noisy image accurately if the sample count is reduced harshly to finish the rendering in a shorter time. We propose a framework that renders the image at a reduced resolution to cast more samples than the harshly lowered sample count in the same time budget. The image is then robustly denoised, and the denoised result is upsampled using original resolution G-buffer of the scene as guidance.

CCS CONCEPTS

• **Computing methodologies → Ray tracing**.

KEYWORDS

Monte Carlo rendering, denoising, guided upsampling, neural nets

ACM Reference Format:
Kadir Cenk Alpay and Ahmet Oğuz Akyüz. 2022. Denoising and Guided Upsampling of Monte Carlo Path Traced Low Resolution Renderings. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '22 Posters), August 07-11, 2022.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3532719.3543250

SIGGRAPH '22 Posters, August 07-11, 2022, Vancouver, BC, Canada

ACM ISBN 978-1-4503-9361-4/22/08.
https://doi.org/10.1145/3532719.3543250

1 INTRODUCTION

Monte Carlo (MC) path tracing enables the rendering of a 3D scene into a 2D image with physically accurate and unbiased light transport calculations, in which the global illumination detail of the scene is included naturally, by estimating the integral of the rendering equation using the Monte Carlo method [Kajiya 1986]. However, it requires an extensive amount of ray samples to be cast into the scene to render a low variance (i.e., visually noise-free) image. Casting a ray sample requires time-consuming computations to be done during the ray's recursive travel in the scene. Due to this nature, MC path tracing requires an excessively high amount of ray samples and a massive time budget to render a low-variance image.

Recent surveys detail the ideas proposed with the aim of decreasing the time cost of MC path tracing without sacrificing the quality of the output image [Huo and Yoon 2021; Zwicker et al. 2015]. One idea is to render the scene with a reduced amount of ray samples to generate a noisy image and then apply a post-process denoising operation on this image using the G-buffer of the scene to produce a denoised image as if the rendering is done with the high sample count. With this approach, a comparable quality image is produced, and the rendering time is reduced vastly as denoising finishes in a shorter time than rendering additional samples. However, this approach fails to produce good quality denoised images when the number of ray samples is decreased harshly to decrease the time cost more. Lowering the sample count below a limit causes the denoisers fail to handle the high variance. Therefore, the accuracy of denoising decreases, resulting in a poor quality denoised output.

In this work, we try to overcome this problem by proposing a framework that is composed of rendering at a lower resolution with higher sample count, denoising, and guided upsampling steps. As the result, the accuracy of the denoiser is recovered, and a good quality denoised output is produced in the time cost of producing at the original resolution with the harshly decreased sample count.

Table 1: Notations and their descriptions.

R_{HR}	Original resolution: The required output resolution and the resolution used in the baseline approach.
R_{LR}	Reduced resolution employed in our approach.
$f_{\downarrow}(R, x)$	Function that returns LR found by dividing the width and height of the input resolution R by the amount x. For instance: $640x360 = f_{\downarrow}(1280x720, 2)$
SPP_{GT}	Sample count used to render an MC image in HR with sufficiently low noise that is visually not distracting.
SPP_{BL}	Sample count of baseline to render an HR image.
SPP_{OUR}	Sample count of our approach to render an LR image.
t_{GT}	Time cost of rendering the HR image with SPP_{GT}.
t_{BL}	Time cost of rendering the HR image with SPP_{BL}.
t_{OUR}	Time cost of rendering the LR image with SPP_{OUR}.

2 APPROACH

Notations are listed in Table 1. The proposed framework is shown in Figure 2. Scene features are the G-buffer elements: Normal, albedo, and depth maps. Diffuse and specular decomposition of [Bako et al. 2017] is employed in our framework, where both components are processed separately first and merged at the end. HR scene features are rendered in a separate rendering pass whose time cost is assumed to be negligibly low, as expensive light calculations such as global illumination are not employed during that pass.

The baseline approach renders scenes at a decreased sample count, SPP_{BL}. This results in a noisy image, but the time cost of the rendering process becomes t_{BL} such that:

$$SPP_{BL} \ll SPP_{GT} \tag{1}$$

$$t_{BL} \ll t_{GT} \tag{2}$$

t_{GT} is the time cost of producing a low-variance and visually noise-free image by rendering at an excessively high sample count, SPP_{GT}, with no denoising being required. The noisy image rendered with SPP_{BL} is then denoised by a MC denoiser, resulting in a final image at a quality comparable to rendering with SPP_{GT}, with the benefit of a largely decreased time cost, t_{BL}. However, if SPP_{BL} is decreased below a limit, the capacity of the denoiser cannot handle the high variance of the samples correctly. Therefore, artifacts start to occur in the denoised image (Figure 1a).

In our framework, we propose to reduce the resolution of this noisy image to be rendered to a lower resolution R_{LR}, with a rate of x on both dimensions of its original resolution R_{HR}, such that:

$$R_{LR} = f_{\downarrow}(R_{HR}, x) \tag{3}$$

Reducing the resolution speeds up the rendering by the rate of x^2, as the pixel count decreases by this amount. We use this extra time budget to cast more ray samples during the rendering of the LR image; however, the amount of ray samples per pixel still being less than SPP_{GT}. We call this new sample count SPP_{OUR}, where:

$$(SPP_{BL} \cdot x^2) = SPP_{OUR} \ll SPP_{GT} \tag{4}$$

Rendering at R_{LR} with SPP_{OUR} results in a noisy LR image with a lower variance, in time cost of t_{OUR} that is similar to the time cost t_{BL} of the baseline, and much smaller than t_{GT}, shown as:

$$t_{BL} \approx t_{OUR} \ll t_{GT} \tag{5}$$

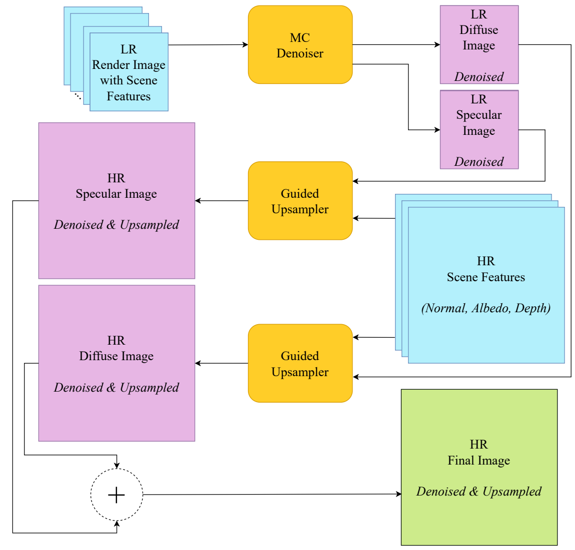

Figure 2: The flow of the proposed framework.

This approach allows us to apply denoising accurately to the LR image rendered with SPP_{OUR}, as it has a higher sample count and thus lower variance than an image rendered with an SPP_{BL} that is below the aforementioned limit of the denoiser capacity. The denoised LR diffuse and specular images are then upsampled to HR using separately-rendered HR scene features as guidance, as they include the edges and texture details of the scene in HR. Both images are then added pixel-wise to produce the final output in HR.

3 CONCLUSION AND FUTURE WORK

We proposed a framework for the denoising and guided upsampling of Monte Carlo images rendered at a reduced resolution with higher number of samples per pixel compared to being rendered at the original resolution using the original harshly reduced sample count in the same time budget. The framework enables MC denoisers to perform accurately on the low resolution but lower variance image. Guided upsampling is then employed to bring the denoised image to original resolution using original resolution G-buffer as guidance. We show the framework can be employed in time-constrained rendering cases where the reduced sample count of the original resolution image is too low to be denoised robustly by an MC denoiser. Future work will examine an end-to-end network that would unify the two steps that are currently executed sequentially.

REFERENCES

Steve Bako, Thijs Vogels, Brian McWilliams, Mark Meyer, Jan Novák, Alex Harvill, Pradeep Sen, Tony Derose, and Fabrice Rousselle. 2017. Kernel-predicting convolutional networks for denoising Monte Carlo renderings. *ACM Trans. Graph.* 36, 4 (2017), 97–1.

Benedikt Bitterli. 2016. Rendering resources. https://benedikt-bitterli.me/resources/.

Yuchi Huo and Sung-eui Yoon. 2021. A survey on deep learning-based Monte Carlo denoising. *Computational Visual Media* 7, 2 (2021), 169–185.

James T Kajiya. 1986. The rendering equation. In *Proceedings of the 13th annual conference on Computer graphics and interactive techniques*. 143–150.

Matthias Zwicker, Wojciech Jarosz, Jaakko Lehtinen, Bochang Moon, Ravi Ramamoorthi, Fabrice Rousselle, Pradeep Sen, Cyril Soler, and S-E Yoon. 2015. Recent advances in adaptive sampling and reconstruction for Monte Carlo rendering. In *Computer graphics forum*, Vol. 34. Wiley Online Library, 667–681.