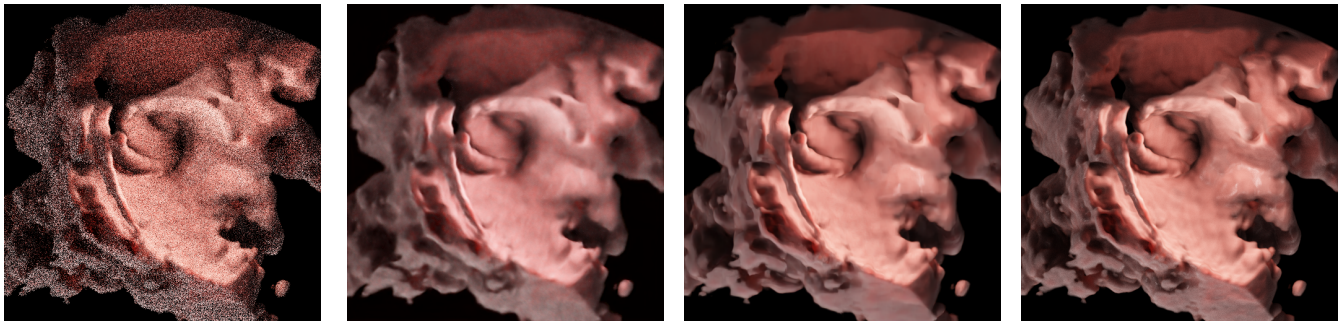
Rendering of 4D Ultrasound Data with Denoised Monte Carlo Path Tracing

Kaloian Petkov
Siemens Healthineers
Princeton, NJ

Daphne Yu
Siemens Healthineers
Princeton, NJ

Helene Houle
Siemens Healthineers
Mountain View, CA



(a) 16 spp (b) 16 spp (denoised) (c) 200 spp (denoised) (d) Reference Image

Figure 1: 4D cardiac ultrasound data rendered with Monte Carlo path tracing. When using a recurrent denoising autoencoder trained on a large collection of partially rendered images, interactive rendering at 16 spp (Fig. 1(b)) produces noise-free images with acceptable spatial and temporal artifacts. High-quality clips of an entire heartbeat can be produced in 1-2 min (Fig. 1(c)) with an order of magnitude reduction in the number of samples compared to the reference image (Fig. 1(d)).

ABSTRACT

We present a rendering system for 4D ultrasound data based on Monte Carlo path tracing, where a recurrent denoising autoencoder is trained on a large collection of images to produce noise-free images with a reduced number of samples per pixel. While the diagnostic value of photorealistic shading for 3D medical imaging has not been established definitively, the enhanced shape and depth perception allow for a more complete understanding of the data in a variety of scenarios. The dynamic nature of ultrasound data typically limits the global illumination effects that can be rendered interactively, but we demonstrated that AI-based denoising together with Monte Carlo path tracing can be used both for interactive workflows and for rendering an entire heartbeat sequence at high quality in about a minute, while also allowing for complex lighting environments. Specifically, our contribution is a model compatible with the NVIDIA OptiX interactive denoiser, which has been trained on ultrasound-specific rendering presets and data.

KEYWORDS

path tracing, denoising, ultrasound, 4D, volume rendering

SIGGRAPH '20 Posters, August 17, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7973-1/20/08.
https://doi.org/10.1145/3388770.3407425

ACM Reference Format:
Kaloian Petkov, Daphne Yu, and Helene Houle. 2020. Rendering of 4D Ultrasound Data with Denoised Monte Carlo Path Tracing. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '20 Posters), August 17, 2020*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3388770.3407425

1 INTRODUCTION AND MOTIVATION

3D acquisition is available across a variety of medical imaging modalities, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and ultrasound. Together with the increasing processing power of modern computers, 3D volume reconstruction and volume rendering play an important role in clinical practice, educational settings and patient communication, where the spatial understand of the images can be improved compared to the traditional 2D rendering of the data. Spatial reasoning can be improved further by photorealistic lighting of the data, and while a variety of global illumination techniques exist for medical volume rendering, we focus specifically on Monte Carlo path tracing [Dappa et al. 2016].

As a realtime acquisition modality, 3D ultrasound presents several unique challenges in deploying a Monte Carlo path tracer to the ultrasound clinical workflow. While the creative flexibility of the lighting design is significantly expanded compared to existing realtime global illumination volume rendering techniques, the path tracing requires many samples per pixel to compute a noise-free result. In the following, we explore an AI-based image denoising approach trained on our volume renderer, with ultrasound-specific rendering presets and a large collection of ultrasound data.

2 TECHNICAL APPROACH

Our Monte Carlo denoiser is based on a recurrent denoising autoencoder [Chaitanya et al. 2017], which takes as input a noisy final color image together with a set of auxiliary buffers generated during path tracing. The auxiliary buffers contain an average of the voxel data properties at the first light bounce location along the primary viewing ray and are updated together with the noisy color image. We compute averaged view-space normals from the volume data gradient while the material color is derived from the voxel classification as specified by a transfer function. Our input features for the initial experiments are chosen to be compatible with the NVIDIA OptiX library, which we use to evaluate the effectiveness of our volume data-specific training against the generic denoising model provided in that library.

The training data is derived from individual medical data review workflows, where image generation uses a single classification preset and anatomy-specific data windowing and camera configurations. We randomly perturb the data windowing, camera, clipping plane orientation and lightprobe orientation parameters during rendering.

3 RESULTS AND FUTURE WORK

Our rendering system uses a custom interactive OpenGL-based volume path tracer inspired by Kroes et al. [Kroes et al. 2012] with support for the majority of 3D imaging medical data formats. For ultrasound data, we specifically implement realtime scan conversion for the acoustic-space data, as well as modality-specific data prefiltering. The denoising work we present here leverages NVIDIA OptiX and we focus on comparing the generic denoising model that ships with the library against a model trained from our own data.

Our training dataset comprises 172 scenes with randomly chosen viewing, lighting and volume data combinations. The volume data was arbitrarily selected from 12,061 3D images over 321 4D datasets with a single transfer function preset, although in the future, we plan to train with the entire dataset and multiple transfer functions. The lighting environment consisted of a single randomly oriented HDR lightprobe, as well as a small interior spherical light source attached to the origin of the volume data in acoustic space, corresponding to the center of the physical ultrasound probe. We generated individual noisy final color, normal and albedo images at 16 spp increments, which resulted in 10664 partially rendered images for a total of 400GB of image data.

Our initial experiments demonstrate a 14% RMSE reduction for the 16 spp image in Fig. 1(b) for our model compared to the generic model shipping with the NVIDIA OptiX library, while the error is similar at 64 spp and above. More importantly, our trained model avoids high frequency artifacts at low spp, as demonstrated in Fig. 2(c). Such artifacts may be mistaken for anatomical structures in clinical workflows. For the target case of 4D cardiac ultrasound rendering, we demonstrate a system that can rendering a video clip of a single heartbeat in about a minute for the majority of cases, which is an order of magnitude improvement and was deemed acceptable in initial interviews with our clinical collaborators.

In the future, we plan to scale up our model training to include multiple transfer functions, a wider variety of lighting conditions and broader array of datasets that contain medical devices (both imaged physical devices and depth-merged surface models), as well as non-ultrasound volume data. In addition, the performance of the system may be improved further by implementing denoised 1 spp global illumination together with traditional raycasting and using auxiliary images that capture more of the volumetric properties of the data.

(a) No denoising (b) Default model (c) Custom model

Figure 2: Cropped region from Fig. 1. All images are rendered at 16 spp. The custom model (Fig. 2(c)) avoids the high frequency artifacts introduced by the generic denoising model (Fig. 2(b)) at the cost of a slight reduction in sharpness.

ACKNOWLEDGMENTS

- Data courtesy of the Piedmont Heart Institute, Atlanta, GA.
- The authors have a contractual relationship with Siemens Healthineers and have received financial compensation.
- The products/features mentioned in the paper might not be approved in some countries or, for other reasons, not yet commercially available. Furthermore, they might still be under development and not commercially available yet. Due to regulatory reasons their future availability cannot be guaranteed.
- Not all products, services or offers referenced in this paper are approved or offered in every market and approved labeling and instructions may vary from one country to another.
- The paper may include discussions of device features and applications that are not FDA cleared in the United States.

REFERENCES

Chakravarty R. Alla Chaitanya, Anton S. Kaplanyan, Christoph Schied, Marco Salvi, Aaron Lefohn, Derek Nowrouzezahrai, and Timo Aila. 2017. Interactive Reconstruction of Monte Carlo Image Sequences Using a Recurrent Denoising Autoencoder. *ACM Trans. Graph.* 36, 4 (2017). https://doi.org/10.1145/3072959.3073601

Evelyn Dappa, Kai Higashigaito, Jürgen Fornaro, Sebastian Leschka, Simon Wildermuth, and Hatem Alkadhi. 2016. Cinematic rendering — an alternative to volume rendering for 3D computed tomography imaging. *Insights into Imaging* 7, 6 (2016), 849–856. https://doi.org/10.1007/s13244-016-0518-1

Thomas Kroes, Frits H. Post, and Charl P. Botha. 2012. Exposure Render: An Interactive Photo-Realistic Volume Rendering Framework. *PLOS ONE* 7, 7 (2012), 1–10. https://doi.org/10.1371/journal.pone.0038586