Deep Volumetric Ambient Occlusion

Objective:
The authors offered a groundbreaking deep learning-based approach for volumetric ambient occlusion in the context of direct volume rendering in this study. Since Direct volume rendering (DVR) requires additional global information, in the form of the transfer function, authors studied ways for injecting such unstructured global data information into 3D convolutional neural networks in order to calculate volumetric ambient occlusion. The Deep Volumetric Ambient Occlusion (DVAO) method they present can predict per-voxel ambient occlusion in volumetric data sets while taking global information from the transfer function into account.

Contributions:
There are 4 major contributions by the authors of this paper:

1. By leveraging a 3D CNN, they propose DVAO, a unique method for predicting volumetric ambient occlusion during interactive DVR.
2. They compare and contrast several representations and injection methodologies for delivering global unstructured information to the CNN when applied to transfer function data.
3. In a comprehensive review, they show that DVAO's impact extends beyond the structures and modalities encountered during training.
4. They provide guidelines that may be used in a variety of volume visualization learning settings.

VIS designs/techniques used:
Authors use 3D volume rendering and show the results trained with different loss functions.

ML methods used:
The challenge of volumetric ambient occlusion is framed as a problem of supervised learning. To predict volumetric ambient occlusion, they train a 3D convolutional neural network.

Strengths:
Once predictions have been made, this technique may simply be incorporated into current DVR pipelines by adding a simple texture lookup. It's an excellent beginning point for volumetric illumination based on deep learning, and it has a lot of room for improvement with more study. Their explicit TF methods can greatly reduce running time.

Shortcomings:
Because 3D CNNs become too computationally costly at higher resolutions and are rapidly constrained by GPU capacity, the current technique is confined to relatively low-resolution volumes of size 128^3. They can't make predictions for each frame since it would take too long, so they only anticipate a new ambient occlusion volume once the transfer function is modified.

Future work:
In the future, this work can be expanded the applicability to a larger range of GI effects in the future, including higher frequency effects. Splitting volumes into smaller bricks might be a promising path for future research to improve prediction resolution. Because the currently best-performing representation is highly costly during inference, more research on improved transfer function representations is still needed.

Question for discussion:
What effect will including additional unstructured global data like camera parameters have on the speed of the existing model?