Novel View Direct/Indirect Image Synthesis

Xiaonong Sun, Anagh Malik
University of Toronto

Motivation

- When we take a conventional image of a scene, by only summing up the photons over the exposure time we discard information on the photon's arrival trajectory and the paths taken.
- Emerging sensors allow new imaging techniques which enable viewing a trillion frames per second, down to individual single photon arrivals.
- One new imaging technique enabled is transient imaging, which could allow distinguishing between the direct light (which contains photons which have bounced off at most 1 surface) and indirect light (which has bounced off at least 2 surfaces).
- This opens potential applications like capturing images outside the sensor's direct line of sight, seeing behind walls and objects, and capturing depth information
- Capturing this image data is however expensive.
- Recent developments in 3D Computer Vision have however led to algorithms for Novel View Synthesis, which can recover indirect/direct light images from a small number of multi-view observations.

Related Work

Scene Reconstruction.

- KinectFusion [1] is a variant of RGB-D reliant method that combines multiple depth measurements fused using a signed distance function (SDF) with real-time tracking to reconstruct objects in real-time.
- Single View MPI is a method that relies on just RGB images [2], which learns to generate multi-plane images given one or more images with known viewpoints.
- Coordinate-based multi layer perceptrons (MLP) is a popular representation of the 3D scene [3]. MLP takes an input 3D location in the model space and outputs occupancy, density or colour.

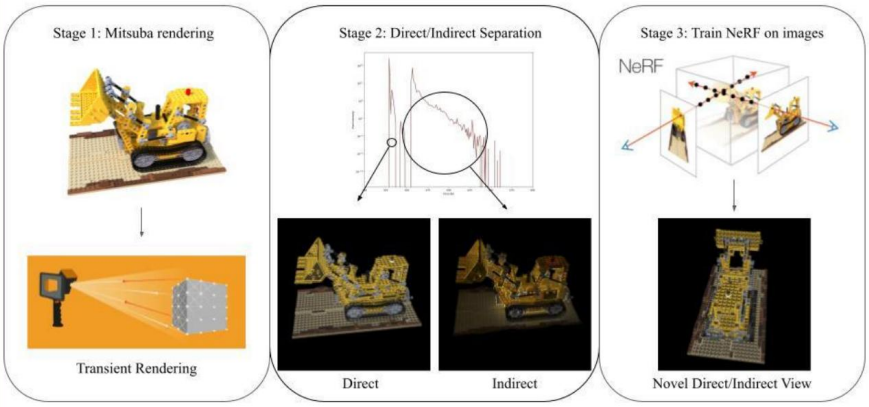
Direct/Indirect Light Separation

- [4] proposes to use high frequency light sources and multi-view images to get light separation. However reconstructing novel indirect/direct light views has not yet been attempted.

References

[1] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davi- son, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In 2011 10th IEEE In- ternational Symposium on Mixed and Augmented Reality, pages 127–136, 2011. [2] R. Tucker and N. Snavely. Single-view view synthesis with multiplane images. In The IEEE Conference on Computer Vision and Pattern Recog- nition (CVPR), June 2020. [3] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In ECCV, 2020 [4] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast sep- aration of direct and global components of a scene using high frequency illumination. ACM Trans. Graph., 25(3):935–944, jul 2006.

Method

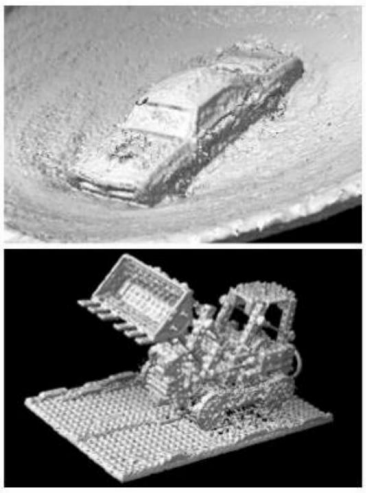


Results

Novel View Synthesis Quality

Ours | Ground Truth

Unsupervised Geometry



Scene	Component	PSNR	SSIM
Lego	Direct	33.64	0.98
Lego	Indirect	28.50	0.95
Car	Direct	38.54	0.98
Car	Indirect	33.01	0.90

Conclusion & Future Work

- NeRF's simple rendering equation is able to represent complex lighting effects, such as indirect reflections.
- Nonetheless results are better for the reconstruction of the scene using the direct component of light. This is probably due to the image formation process being closer to the rendering equation.
- We are able to recover geometry from the images. It would be interesting to see how the reconstructed geometry compares to the geometry reconstructed with normal images.
- It would also be interesting to see these reconstruction for more complicated surface, for example with sub-surface scattering.