Learning to Synthesize a 4D RGBD Light Field from a Single Image
Supplementary Material

Pratul P. Srinivasan[1], Tongzhou Wang[1], Ashwin Sreelal[1], Ravi Ramamoorthi[2], Ren Ng[1]

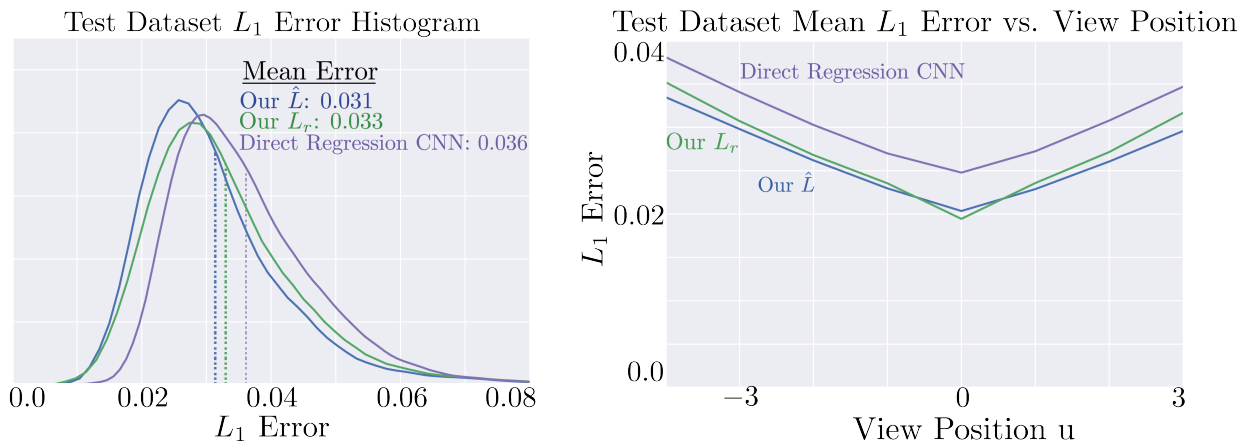[1]University of California, Berkeley   [2]University of California, San Diego

Figure 1. To quantitatively validate our results, we visualize histograms of the L_1 errors on the testing dataset for the outermost views of our predicted light fields $\hat{L}$, our Lambertian light fields L_r, and the light fields predicted by a CNN that directly regresses from an input image to an output light field. Our predicted light fields and Lambertian light fields both have lower errors than those of the direct regression CNN. We also compute the mean L_1 errors as a function of view position u, and demonstrate that our algorithm consistently outperforms the direct regression CNN.
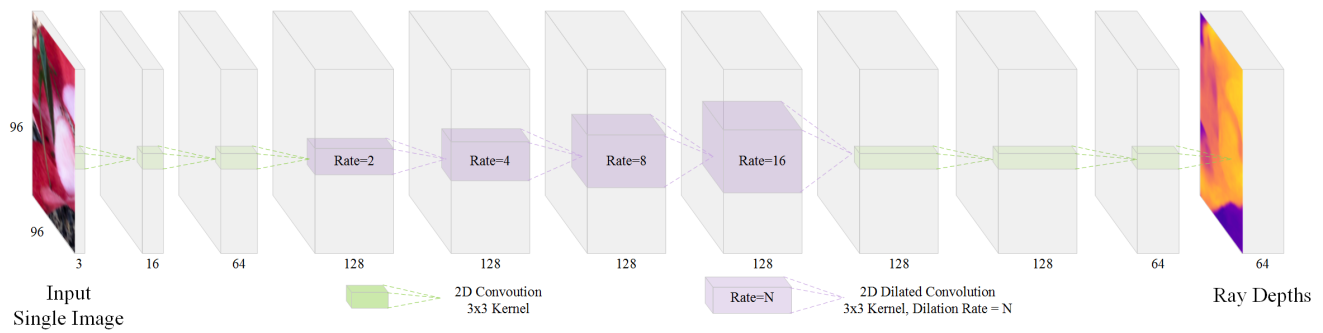
1

Figure 2. The convolutional neural network to estimate 4D ray depths from the input 2D image consists of 10 convolutional layers. We use dilated convolutions to enable each predicted ray depth to have access to the entire input image without the resolution loss caused by spatial downsampling or pooling. All filters are 3x3, and we use batch normalization and exponential linear unit activation functions for each layer except the last layer. The last layer is followed by a scaled tanh activation function to constrain the possible disparities.
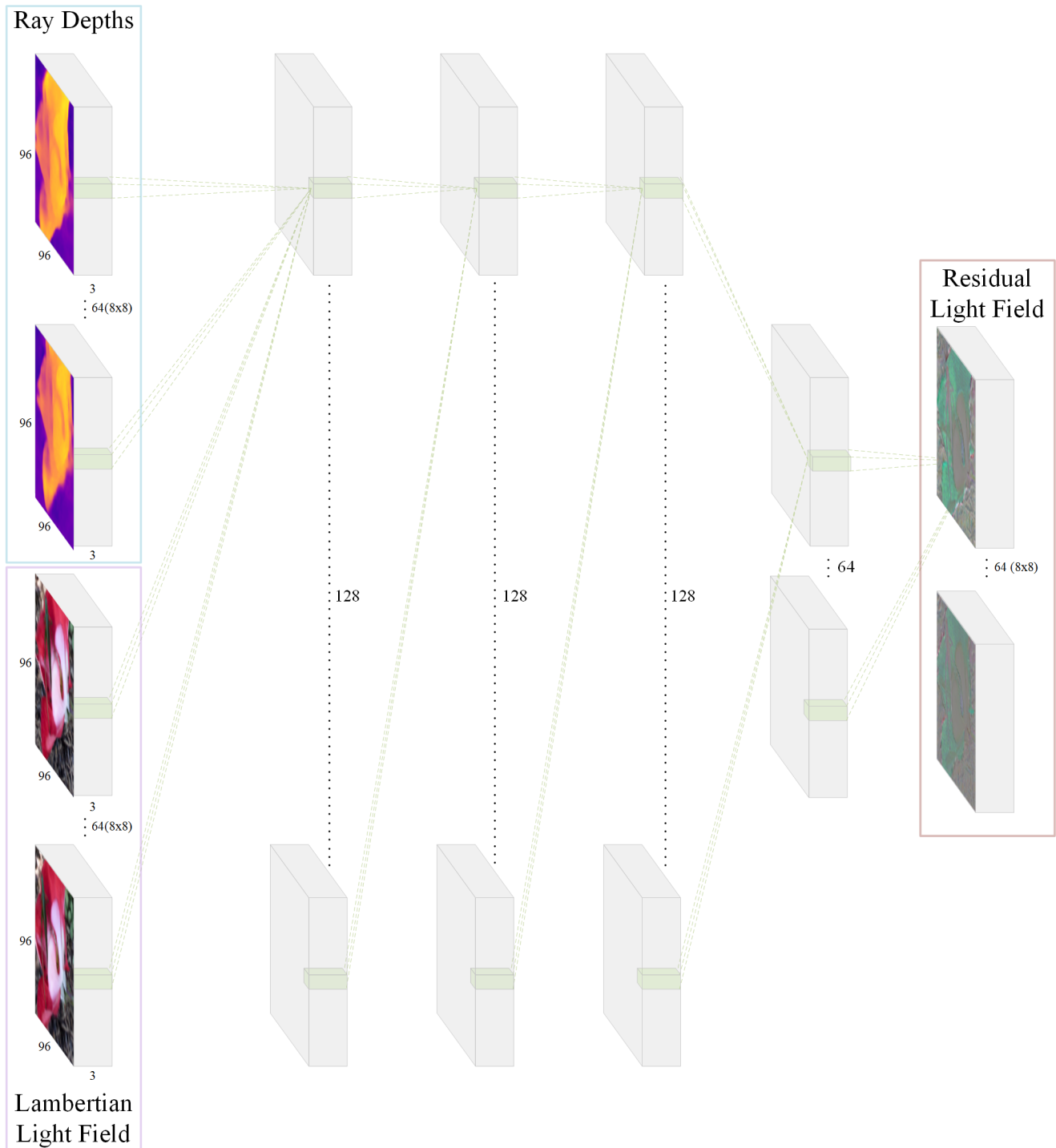
Figure 3. The convolutional neural network to predict occluded rays and non-Lambertian effects is structured as a residual block, and its output is added to the Lambertian light field to estimate our final predicted light field. This ensures that decreases in the loss are driven by correctly predicting occluded rays and non-Lambertian effects. We use 3D convolutions of size 3x3x3 that are full in the angular dimension, so each filter has access to every viewpoint. As in the depth estimation network, each convolution is followed by an exponential linear unit activation function and batch normalization, except for the last layer, which is followed by a tanh activation function to constrain the predicted light field values.