Supplemental Document

Optics EXPRESS

Dense Lissajous sampling and interpolation for dynamic light-transport: supplement

XIAOMENG LIU,[1,*] KRISTOFER HENDERSON,[1] JOSHUA REGO,[2] SUREN JAYASURIYA,[2,3] AND SANJEEV KOPPAL[1]

[1] *Department of Electrical and Computer Engineering, University of Florida, FL 32611-6200, USA*
[2] *School of Electrical, Computer and Energy Engineering, Arizona State University, AZ 85287-5706, USA*
[3] *School of Arts, Media and Engineering, Arizona State University, AZ 85281, USA*
[*] *liuxiaomeng51@ufl.edu*

Supplement DOI: https://doi.org/10.6084/m9.figshare.14625039

Parent Article DOI: https://doi.org/10.1364/OE.425061

Supplemental Material: Dense Light-Transport Sampling and Interpolation for Fast Flying-Spot Photography

In this supplemental material, we present more details about our neural network architecture for light transport interpolation. We also show additional interpolation results for our real experimental data capture at 24 FPS using a Lissajous subsampling pattern.

1. ZSM NETWORK DETAILS

We adopt the network from [1] to use as our network architecture for frame interpolation (which we entitled ZSM). The network architecture is visualized in Figure S1. This network interpolates between any two frames by extracting visual features, then uses deformable convolutions to help temporally interpolate these features, before feeding both the interpolated features and the original frames into a bidirectional LSTM to output the final interpolated frame [1]. We use the same layers and parameters as the original network except we do not include the final upsampling layer in our implementation.

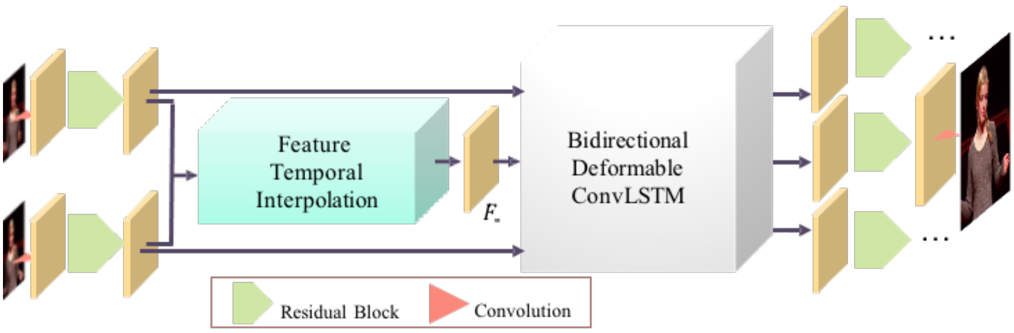

Fig. S1. Network architecture from [1] which we leverage for interpolating between columns of the light transport matrix.

2. INTERPOLATION PROCEDURE AND RESULTS

To interpolate between two column images with indices i and j, we utilize a simple binary search tree procedure to fill in the missing indices between i and j. In Figure S3, we show this procedure in graphical form. The network interpolates the midpoint $k = mid(i, j)$ in the first pass, then $k_1 = mid(i, k), k_2 = mid(k, j)$ in the second pass, and so on until all interpolations are complete. Then all of these interpolated results are compared to ground truth column images with both a MSE and perceptual loss as described in the main paper. Since the flying spot travels in a linear fashion, typically i and j lie on an approximately horizontal line and can be interpolated using our method. Depending on the pattern, e.g. very sparse Lissajous patterns for the real 24 FPS capture, we perform vertical interpolation to fill the gaps.

In Figure S2, we show the progressive iterations of our interpolation procedure for the real data captured with a Lissajous subsampling. As we can see, the vertical interpolation method slowly fills in the gaps in the light transport until the Lambertian backplane is filled. There are some slight errors along the boundary of the plane due to difficulty interpolating along edges/depth discontinuities in the scene.

In Figure S4, we show the results of network's final interpolation for several frames of the resulting video sequence. Note how many of the gaps in the light transport are replaced and the scene exhibits several caustic and global light effects. This result shows the effectiveness of

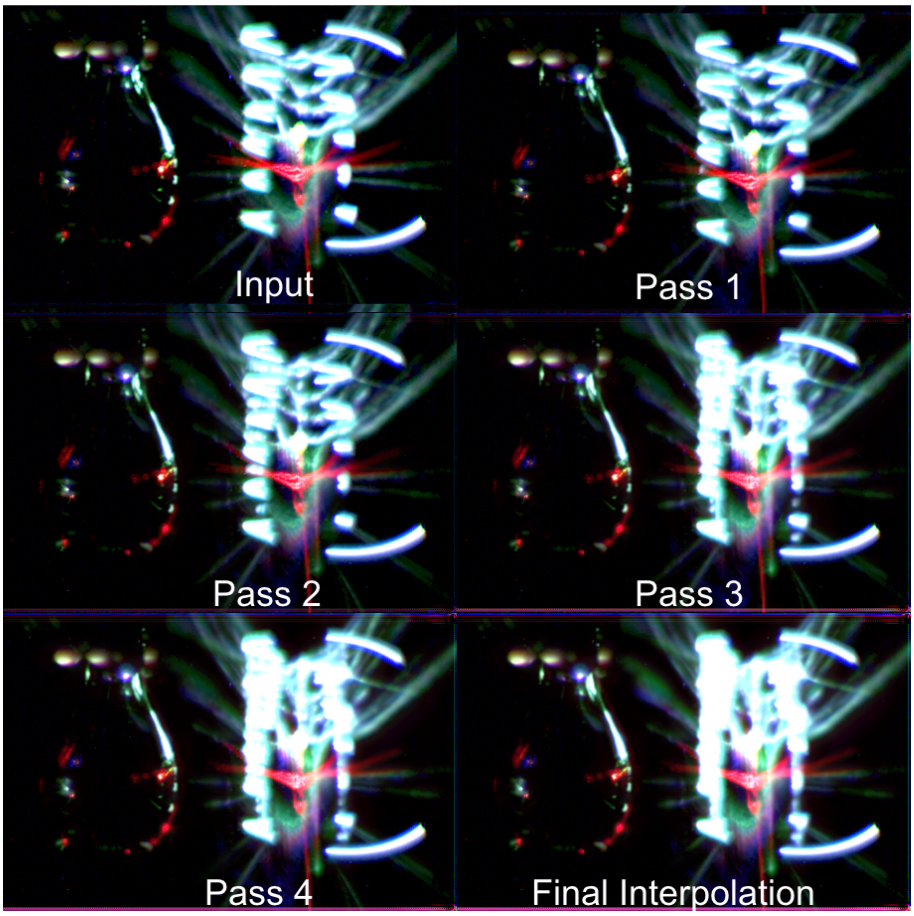

Fig. S2. Interpolation procedure for real LoS data captured by the experimental prototype.

machine learning to help speed up dynamic light transport capture. We encourage the reader to refer to the supplemental video to view the result.

In Figure S5, we compare different subsampling methods for a single frame. The three methods sample 12.5% of the original light transport matrix, with quantitative results for the corresponding interpolated scene of PSNR=25.72dB/SSIM=0.922 for Lissajous sampling, PSNR=26.89dB/SSIM=0.971 for uniform sampling, and PSNR=20.59dB/SSIM=0.886 for random sampling. Note that we only measure these metrics in the center region of the images to avoid any edge effects from the different sampling methods. While these results show that a uniform sampling method would be preferred for better interpolation performance, the Lissajous sampling method is most physically feasible for high-speed patterns for the current hardware setup.

REFERENCES

1. X. Xiang, Y. Tian, Y. Zhang, Y. Fu, J. P. Allebach, and C. Xu, "Zooming slow-mo: Fast and accurate one-stage space-time video super-resolution," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), pp. 3370–3379.

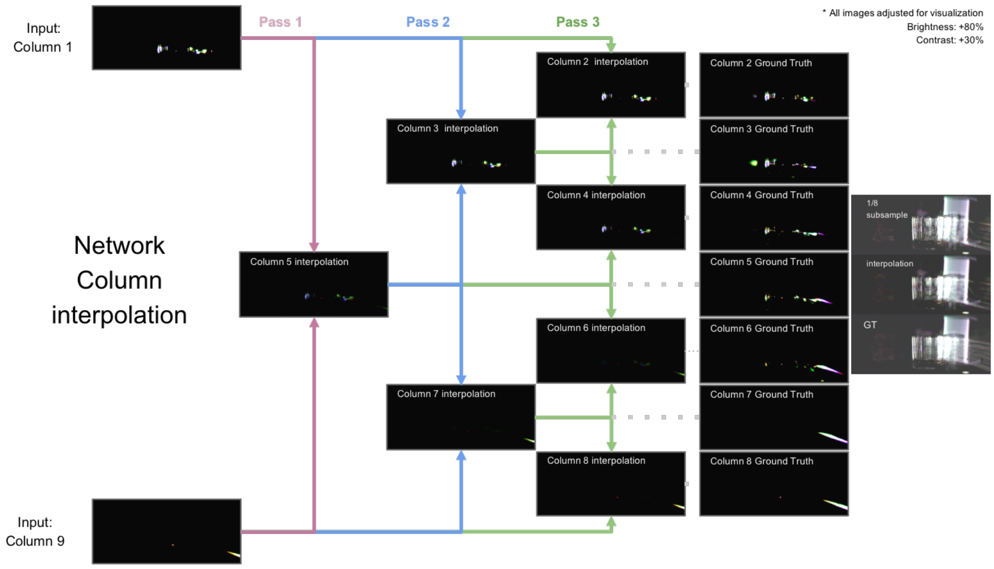

Fig. S3. Network interpolation procedure for column images 2-8 given columns 1 and 9 from a 1/8 uniform subsampling simulation. On the right, the corresponding floodlit from all the columns of the light transport is shown for the subsampled scene, the interpolated scene, and the ground truth. Note for visualization purposes that all images have increased brightness by 80% and contrast by 30%.

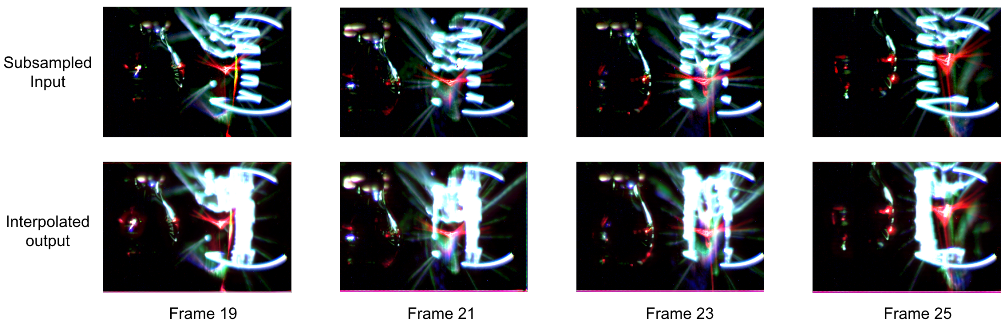

Fig. S4. Network interpolated frames from the real LoS data captured with a Lissajous subsampling at 24 FPS. Please see supplemental video for the entire sequence.

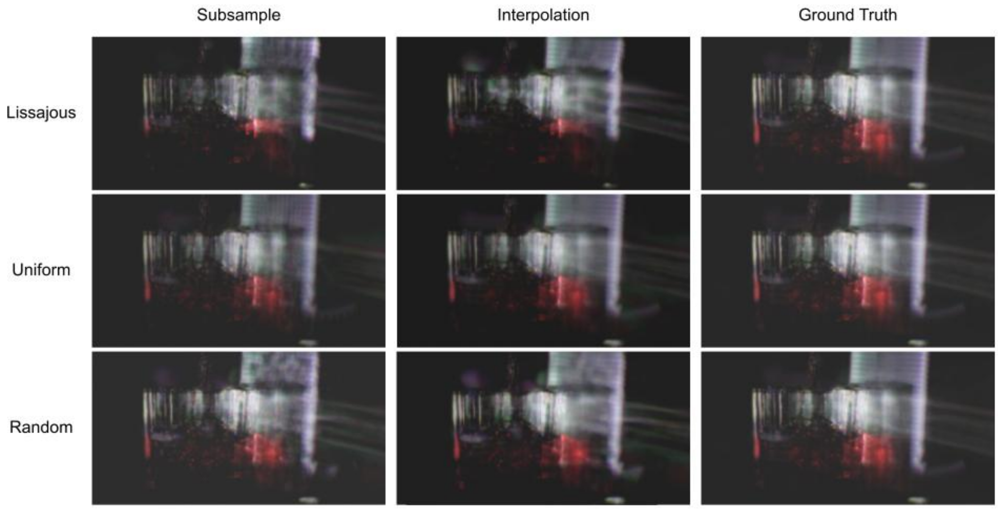

Fig. S5. Comparison of simulated Lissajous subsampling and the resulting interpolation vs uniform and random subsampling methods.