Unrolled Primal-Dual Networks for Lensless Cameras: Supplementary Material

Oliver Kingshott,[1] Nick Antipa,[2], Emrah Bostan[3] and Kaan Akşit[1, *]

[1] *University College London, London, United Kingdom*
[2] *University of California San Diego, San Diego, United States of America*
[3] *ams OSRAM, Lausanne, Switzerland*
[*] *k.aksit@ucl.ac.uk*

1. Detailed Experimental Results

We evaluate our model performance in detail against work from [1] in Figure 1.

We thank [2] for providing sample reconstructions from their model, as their code and weights are not yet publically available. We evaluate our test reconstructions against sample reconstructions from their model in Figure 2.

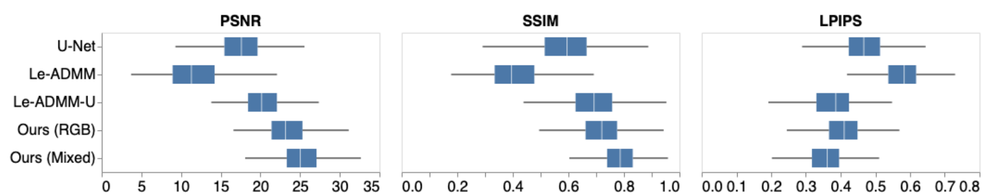

Fig. 1. Comparison of our proposed models trained using the DiffuserCam MirFlickr dataset [1]. We evaluate our performance on the test set containing 1000 images. Our higher PSNR performance suggests that our model is able to extract more information from each lensless measurement. Our model performs less consistently well on perceptual image quality metrics such as SSIM and LPIPS, which we chose not to optimize for.

2. Lensless Camera Prototype

Figure 3 shows photographs of our lensless camera and the display used to capture paired training examples. Figures 4 and 5 illustrate our how our calibration software captures paired training examples. Figure 6 demonstrates test reconstructions from our camera.

3. Learned models from Zero Initialization

Our proposed work initializes θ_k with the point spread function measured by shining a point light source along the optical axis of the lensless camera. To show that our model is able to estimate a physical PSF from data, we show the values of θ_k after training when θ_k is zero-initialized in Figure 7.

References

1. K. Monakhova, J. Yurtsever, G. Kuo, N. Antipa, K. Yanny, and L. Waller, “Learned reconstructions for practical mask-based lensless imaging,” Opt. Express **27**, 28075 (2019).
2. T. Zeng and E. Y. Lam, “Robust reconstruction with deep learning to handle model mismatch in lensless imaging,” IEEE Trans. on Comput. Imaging **7**, 1080–1092 (2021).
3. A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” SIAM J. on Imaging Sci. **2**, 183–202 (2009).
4. R. O. Duda and P. E. Hart, “Use of the hough transformation to detect lines and curves in pictures,” Commun. ACM **15**, 11–15 (1972).

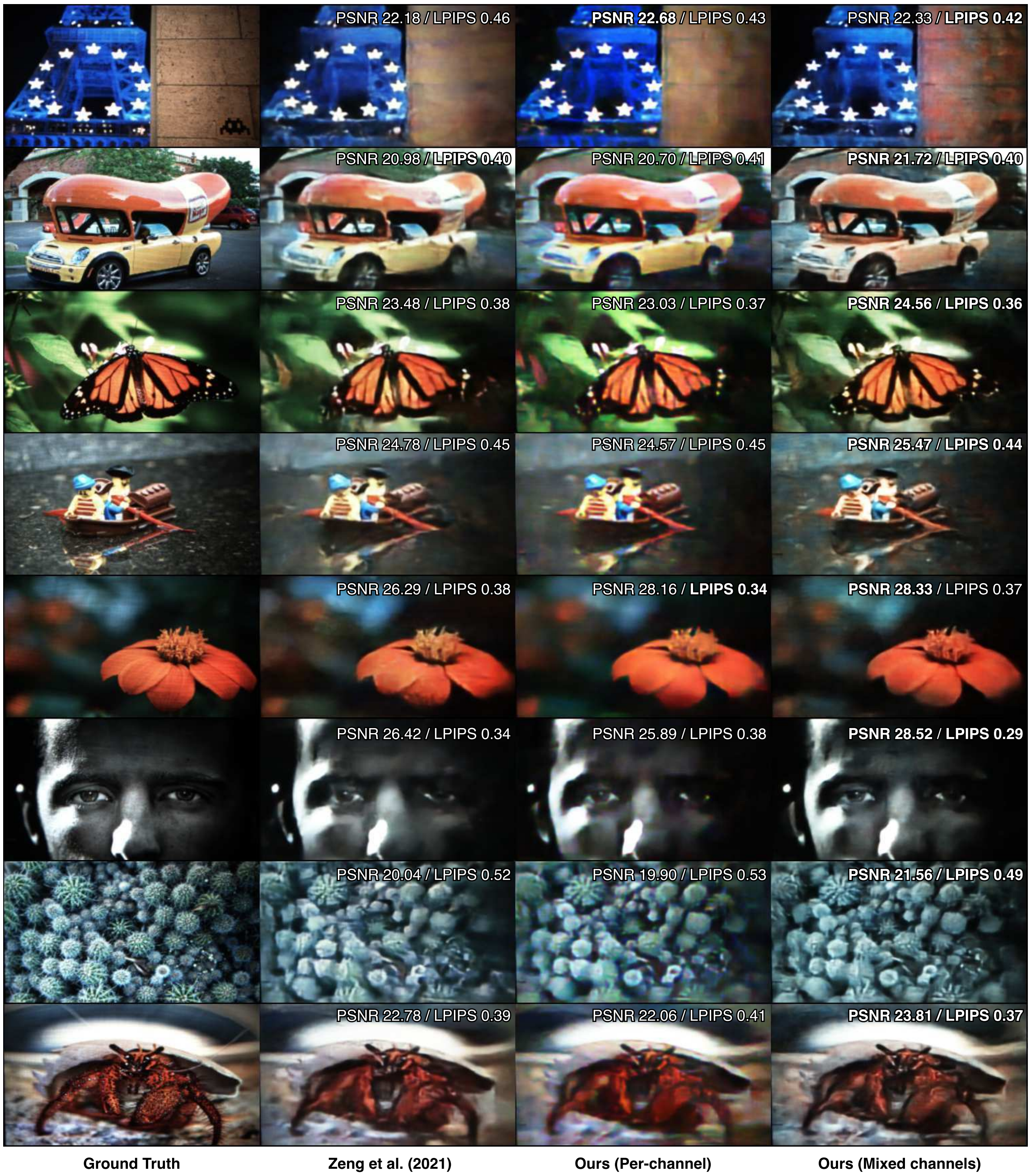

Fig. 2. Comparison of our proposed models against sample reconstructions provided by [2]. Our results are similar in terms of reconstruction quality, with our mixed-channel model yielding a slightly higher PSNR in most cases. Close inspection suggests that our model suffers from fewer artifacts as a result of more closely emulating the physics of light transport. However, we believe that a more detailed comparison is needed in order to establish which model offers superior performance in terms of reconstruction speed and quality.

Fig. 3. Photos of our lensless camera prototype. The camera is aimed at the display to capture paired training examples. The paired training examples are used to train our learned image reconstruction model.

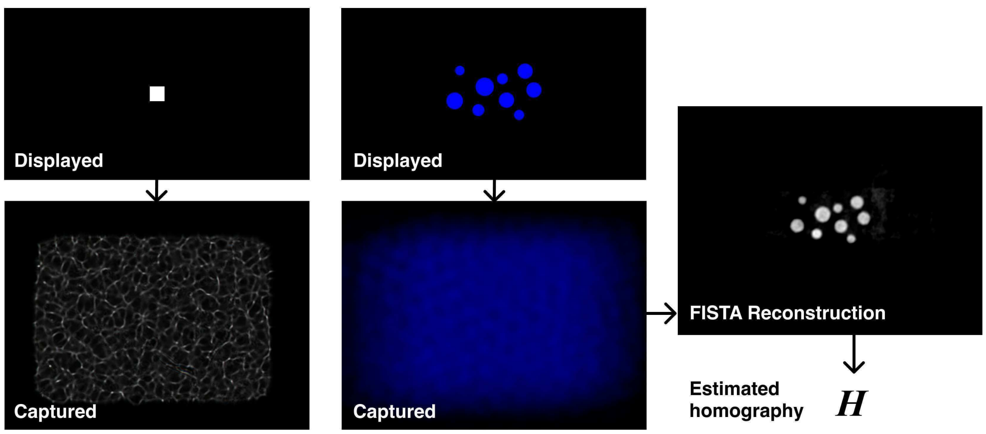

Fig. 4. Illustration of our homography estimation routine. The on-axis PSF is recorded by illuminating a point light source in the center of the display. An image of circles is then shown on the display and is reconstructed using the recorded PSF and FISTA [3]. The centers of these circles are then found using a Hough transform [4], and the difference in circle centers between the source image and the captured image are used to estimate the homography $\mathbf{H}$.

5. E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops,* (2017).

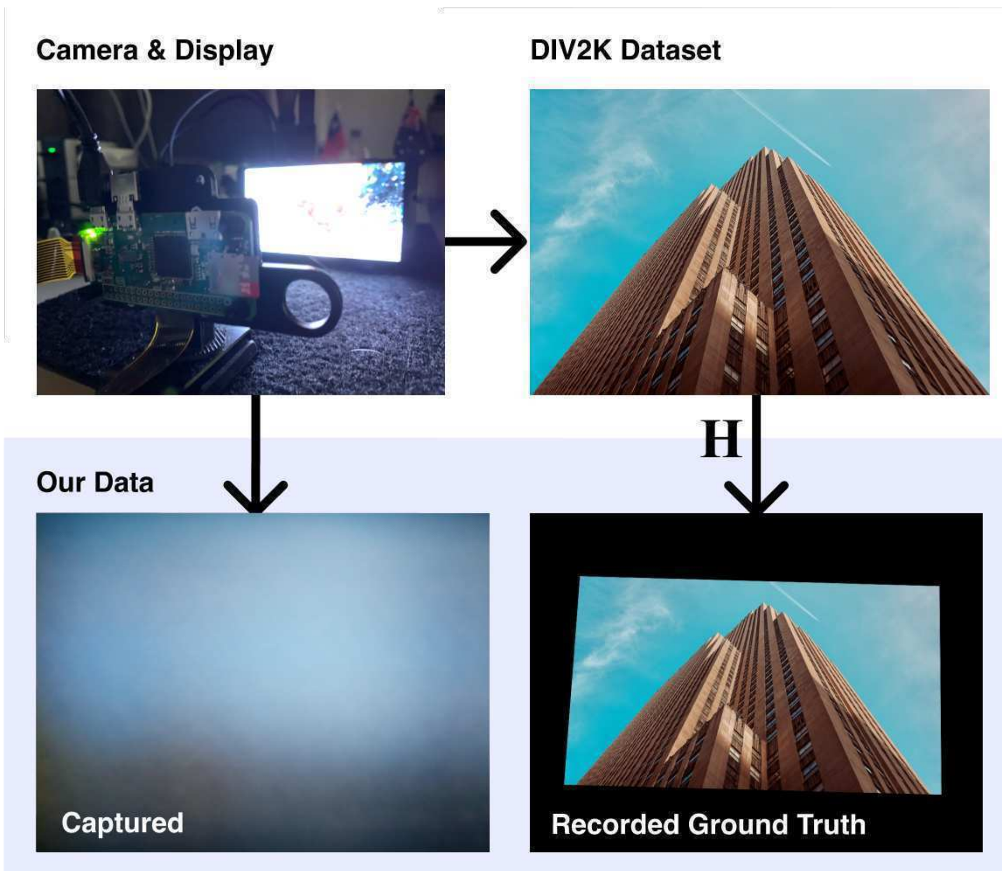

Fig. 5. Illustration showing the flow of data from our calibration software. Images are shown on the display and are recorded by the camera as lensless measurements. The estimated homography **H** is applied to each ground truth image to match the lensless camera's view of the display. We use the Div2K dataset [5] to capture paired training examples.

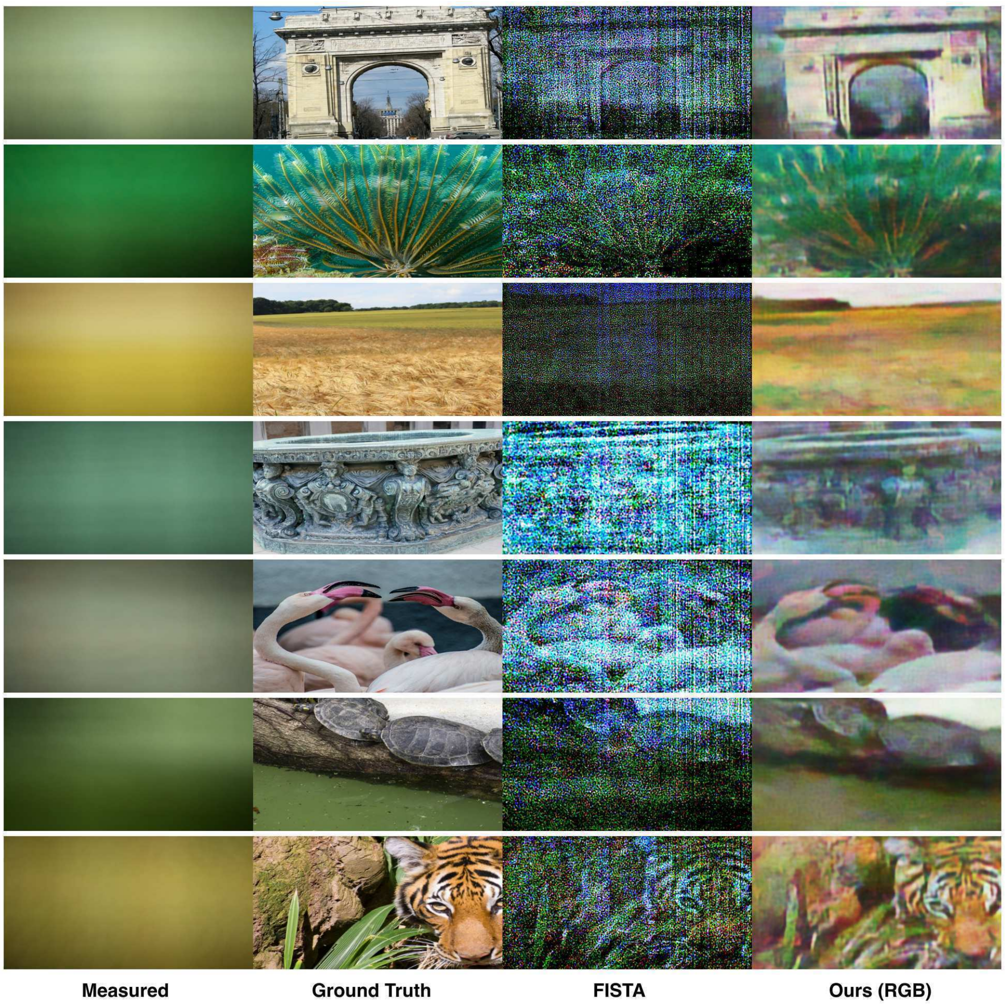

Fig. 6. Test measurements reconstructions from our prototype camera. We compare the results of unregularized FISTA [3] against our RGB model trained with paired training examples. Our unrolled model is able to reconstruct usable images at least one order of magnitude faster. The amount of noise in the unsupervised reconstructions suggests that our camera has a poor signal to noise ratio, this is likely due to the characteristics of the diffuser and the imaging sensor.

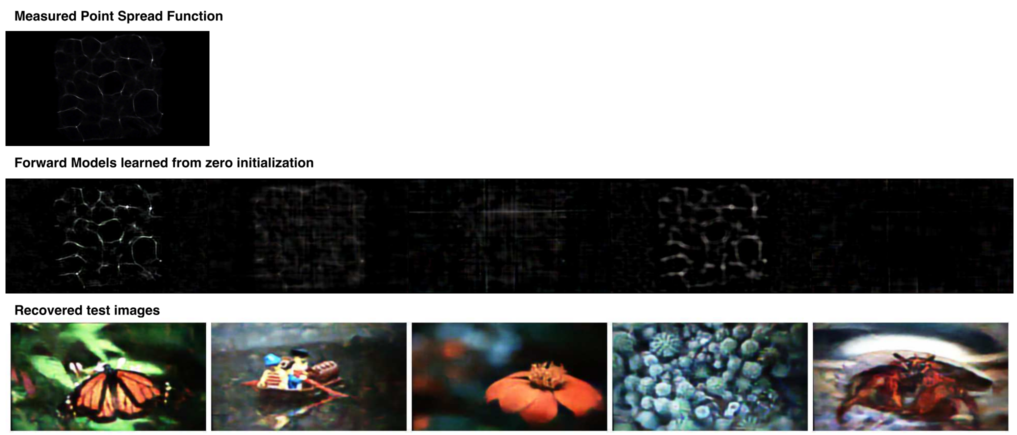

Fig. 7. Results of training our RGB model with zero initialization of θ_k. The top row shows the experimentally measured PSF from the DiffuserCam MiRFlickr dataset by [1]. The second row shows $\theta_k^{1\dots5}$. The third row shows test images recovered using these models, which have a slightly higher test PSNR than initialization with the calibrated PSF (22.22dB). We can see that θ_k^1 resembles a strong similarity to the measured PSF, while the other values $\theta_k^{2\dots5}$ are less physically interpretable.