Computational Flash Photography through Intrinsics - Supplementary Material

Extended Results and Comparisons

Sepideh Sarajian Maralan Chris Careaga Yağız Aksoy

Simon Fraser University

1. Illumination Editing

The results of our decomposition network, unmixed flash, and ambient illuminations, can be used to interactively alter the two illuminations independently, as Figure 3 demonstrates.

2. Additional Comparisons

In this section, we compare each of the prior methods to our own and highlight their individual differences.

2.1. IAN [7]

As illustrated in Figure. 1, IAN [7] produces facial artifacts during the decomposition task (row two and three). We also observe that the decomposition and the generation is not done completely in some cases as shown in row one and four. Compared to their decomposition, the generation portraits appear more realistic, though the generation results can still look blurry in some areas (row five and six). For better visualization we utilize the ratio image high resolution approach from Aksoy et al. [1].

2.2. OIDDR [6]

In OIDDR [6] we focus on comparison of shading estimations. As seen in the first three rows in Figure 2, the shading predicted for the generation task by this method is not usable for editing. Compared to the generation, the decomposition task's predicted shadings are more accurate, but some areas still have incorrect specularities. Row three and four are not from our test set.

2.3. DeepFlash [2]

We trained [2]'s encoder-decoder based on VGG-16 [4] with their bilateral filtering for both tasks but only the decomposition model converged which is also the task they tackled in their paper. Figure [2] contains three comparisons between our method and theirs. As seen in this figure, this method struggles with images outside of the training domain.

2.4. FAID [1] and Pix2PixHD [5]

FAID [1] and Pix2PixHD are both generative adverserial networks. Pix2PixHD has a coarse-to-fine generator and multi-scale discriminators that operate at different image scales while FAID utilizes the pix2pix [3] architecture with the use of ratio images to make the network's job easier. Both of these methods produce artifacts. While FAID performs better than Pix2PixHD on the test set, it still produces a lot of artifacts for the in-the-wild dataset. The results for both generation and decomposition are shown in Figure. 5.

References

[1] Yağız Aksoy, Changil Kim, Petr Kellnhofer, Sylvain Paris, Mohamed Elgharib, Marc Pollefeys, and Wojciech Matusik. A dataset of flash and ambient illumination pairs from the crowd. In *Proc. ECCV*, 2018. 1, 4

[2] Nicola Capece, Francesco Banterle, Paolo Cignoni, Fabio Ganovelli, Roberto Scopigno, and Ugo Erra. DeepFlash: Turning a flash selfie into a studio portrait. *Signal Processing: Image Communication*, 2019. 1, 4

[3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proc. CVPR*, 2017. 1

[4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *Proc. ICLR*, 2015. 1

[5] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional GANs. *Proc. CVPR*, 2018. 1, 4

[6] Amirsaeed Yazdani, Tiantong Guo, and Vishal Monga. Physically inspired dense fusion networks for relighting. *Proc. CVPR Workshops*, 2021. 1, 3

[7] Zuo-Liang Zhu, Zhen Li, Rui-Xun Zhang, Chun-Le Guo, and Ming-Ming Cheng. Designing an illumination-aware network for deep image relighting. *IEEE Trans. Image Process.*, 2022. 1, 2

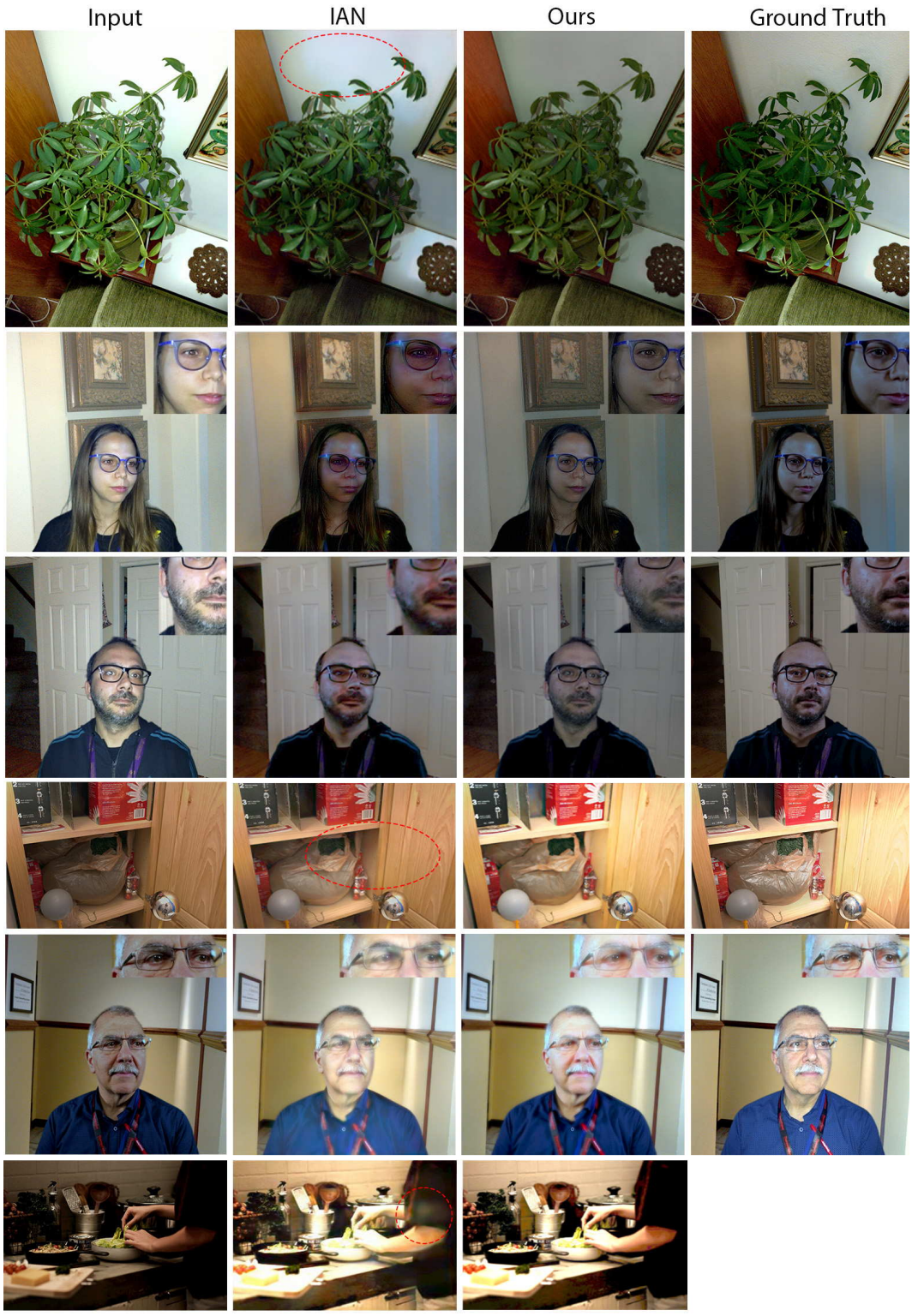

Figure 1. Comparison between our method to IAN [7] in decomposition (row one to three) and generation task (row four to six). Row six is not from our test set.

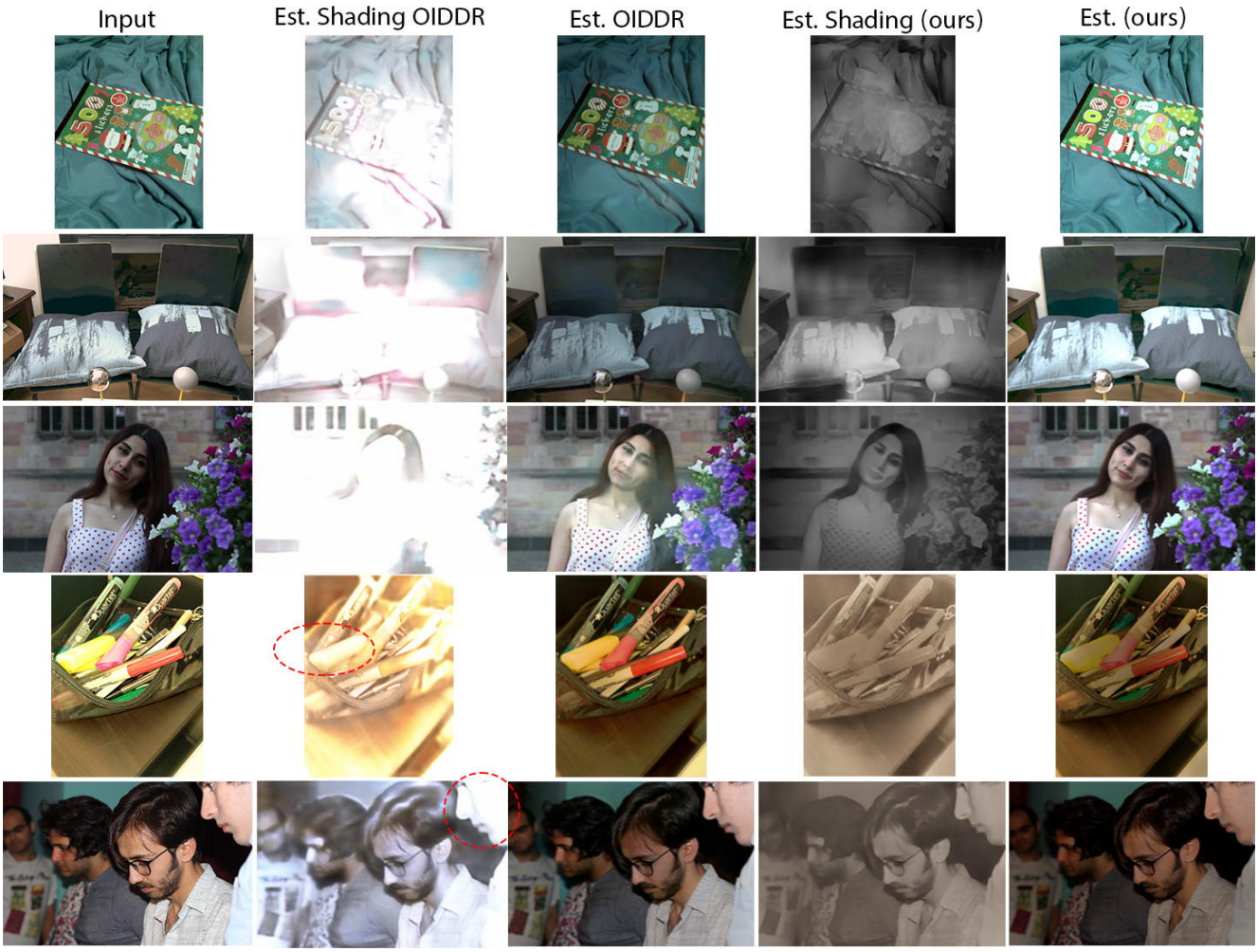

Figure 2. Comparison between our method to OIDDR [6]'s in generation task (row one to three) and decomposition task (row four and five) for both the shading and final illumination. Row three and five are not from our test set.

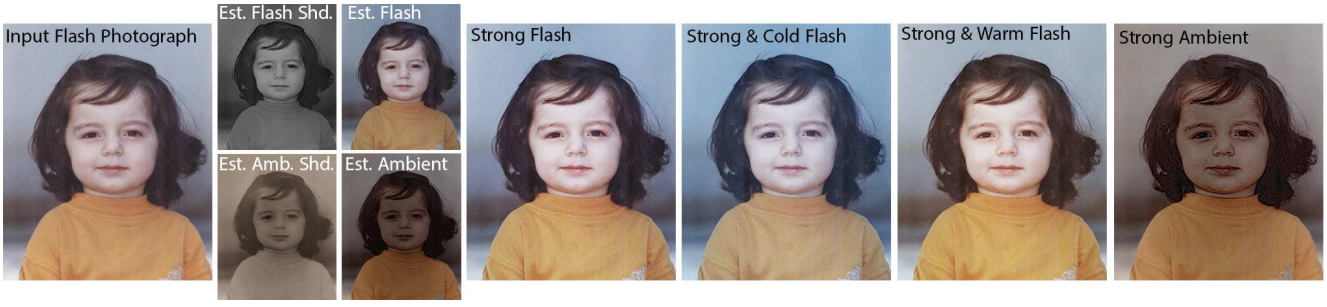

Figure 3. The separate flash and ambient shadings we estimate for a given flash photograph can be edited separately to render the same scene under varying illuminations.

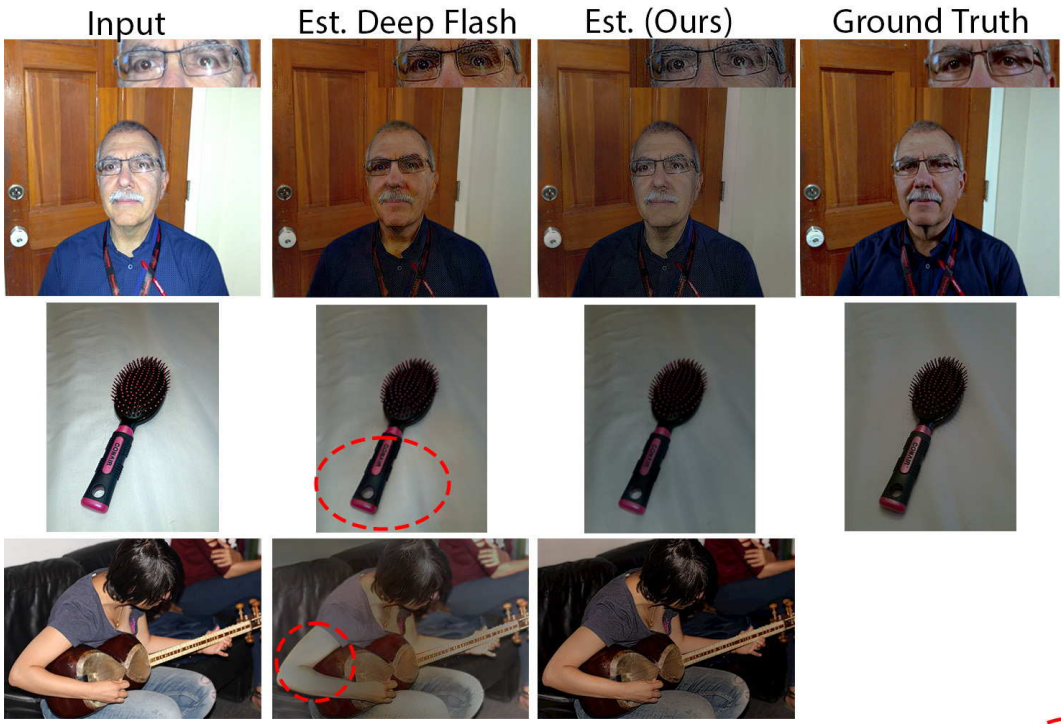

Figure 4. Comparison between our method to Deep Flash [2]'s in the decomposition task. The last row is not from the test set.

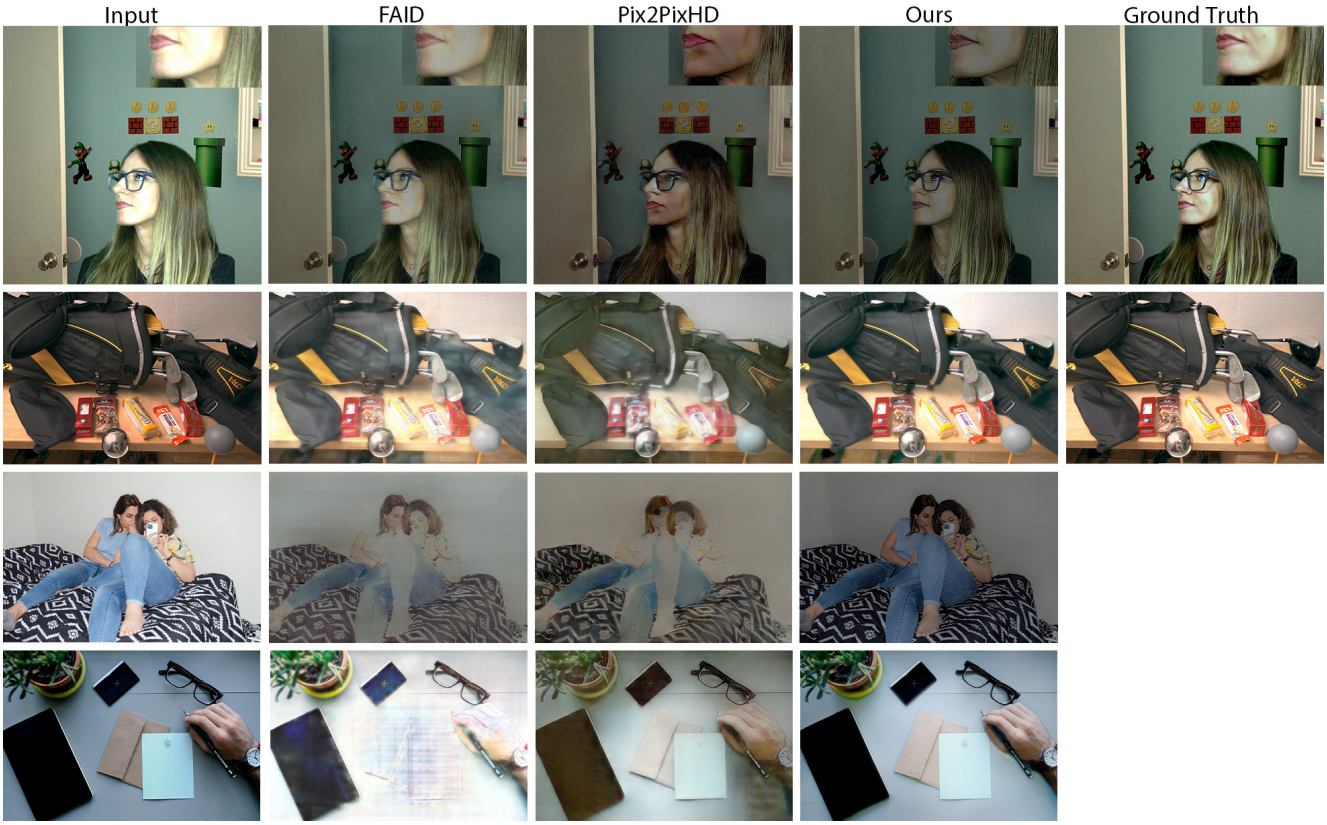

Figure 5. Comparison between our method to Pix2PixHD [5] and FAID [1] in decomposition (row one and three) and generation task (row two and four). Row three and four are not from the test set and we do not have the ground truth for them.