Supplementary Material–Under-Display Camera Image Restoration with Scattering Effect

Binbin Song[1], Xiangyu Chen[1,2], Shuning Xu[1], and Jiantao Zhou[1†]
[1]State Key Laboratory of Internet of Things for Smart City
Department of Computer and Information Science, University of Macau
[2]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
{yb97426, jtzhou}@umac.mo, {chxy95, rebeccaxu0418}@gmail.com

In this Supplementary Material, we first analyze the statistical characteristics of the whole synthetic dataset in the HSV space in Section 1. Then, the details of the Feature Fusion Block in SRUDC are given in Section 2. More visual comparisons on both the real-world and the synthetic datasets are given in Section 3 and Section 4.

1. Statistical Characteristics of Synthetic Dataset in HSV Space

As mentioned in the main body of our paper, we conduct a statistical analysis of the entire synthetic dataset in the HSV color space to illustrate the superiority of our enhanced image formation pipeline (IFP). Fig. 2 shows the statistical mean values of H, S, and V channels of different datasets, including the real-world dataset [4], the synthetic dataset generated from our enhanced IFP, and the synthetic dataset generated from Feng *et al.* [1]. In each subgraph, the horizontal axis indicates the mean value of H, S, or V channel and the vertical axis indicates the image quantity ratio. The red and blue histograms respectively denote the statistical mean H, S, and V values of the display-free (ground-truth) images and under-display camera (UDC) images.

In the first row of Fig. 2, the statistical mean H and S values of UDC images demonstrate obvious deviations compared with those of display-free images, i.e., the blue line is offset to the left relative to the red line. Such offsets of the mean Hue and saturation are caused by the haziness and contrast distortions in UDC images. In the second row, we can find that the synthetic dataset generated from our enhanced IFP exhibits similar offsets in H- and S-channel distributions. However, in the third row, the blue and red lines of the first two subgraphs almost overlap. This indicates that the IFP proposed by Feng *et al.* [1] cannot simulate such offsets, which are observable in the real-world dataset. Regarding the V-channel distribution, the differences between the display-free and UDC images are small in the synthetic dataset generated by our enhanced IFP, which is consistent with the findings proposed by Luo [4].

†Corresponding author.

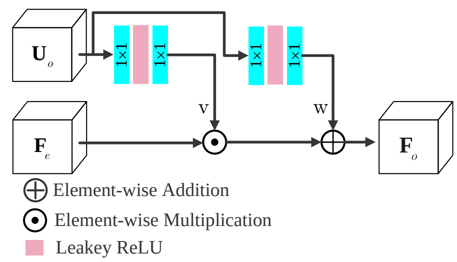


Figure 1. The schematic diagram of FFB.

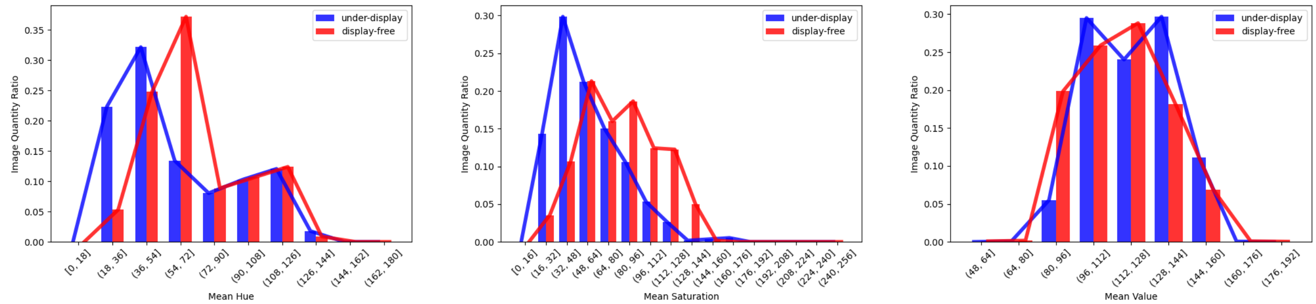
2. Feature Fusion Block of SRUDC

Fig. 1 presents the specific design of our feature fusion block (FFB). In this block, we aim to modulate the encoder features $\mathbf{F}_e$s in the image branch with the global information brought by the features $\mathbf{U}_o$s from the scattering branch. Inspired by the formulation of the UDC scattering model proposed in the main body, we apply the affine transformation in FFB. Specifically, we compute the coefficient $\mathbf{v}$ and the offset $\mathbf{w}$ from $\mathbf{U}_o$ through the combinations of two 1×1 convolution layers and a leaky ReLU function. Eventually, $\mathbf{F}_e$ is recalibrated to be the output feature $\mathbf{F}_o$ by:

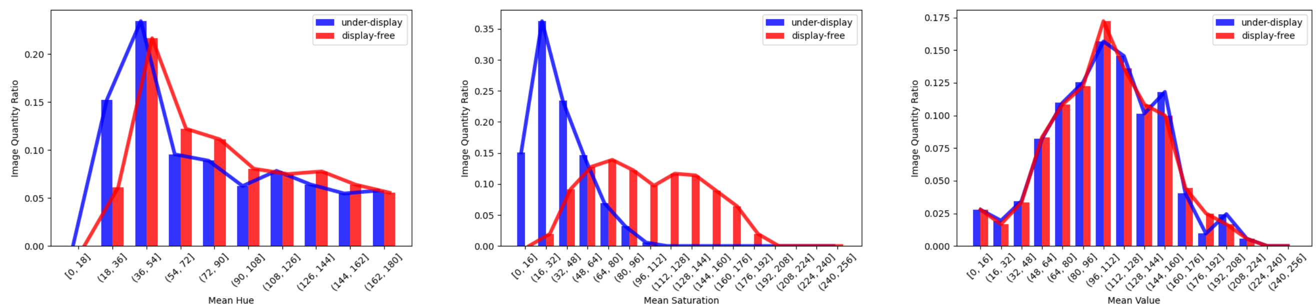
$$\mathbf{F}_o = \mathbf{v} \odot \mathbf{F}_e + \mathbf{w}. \tag{1}$$

3. Visual Comparison on Real-world Dataset

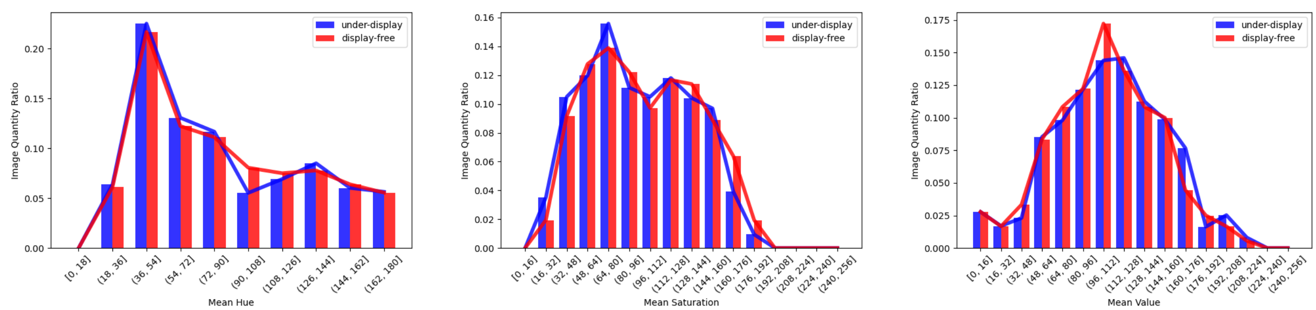
We provide more visual comparisons with state-of-the-art methods, including DISCNet [1], DAGF [5], DWFormer [6], UDCUNet [3], and BNUDC [2], on real-world data in Fig. 3 and Fig. 4. Our proposed method suppresses the haziness and contrast distortion caused by the scattering effect. In addition, our method does not produce color



(a) Statistical mean values of the H, S, and V channels of the real-world dataset captured by Luo *et al.* [4].



(b) Statistical mean values of the H, S, and V channels of the synthetic dataset generated from our enhanced IFP.



(c) Statistical mean values of the H, S, and V channels of the synthetic dataset generated from Feng *et al.*. [1]

Figure 2. The histograms represent statistical mean values of the H, S, and V channels. The horizontal axis of each graph indicates the value of H, S, or V channel, and the vertical axis indicates the image quantity ratio. (a) The statistical mean values of the H, S, and V channels of the real-world dataset [4]. (b) The statistical mean values of the H, S, and V channels of the synthetic dataset generated from our enhanced IFP. (c) The statistical mean values of the H, S, and V channels of the synthetic dataset generated from Feng *et al.* [1].

shifts as other models, e.g., the wrong grass color induced by BNUDC in the second example of Fig. 3 and the wrong ground color caused by DWFormer in the second example of Fig. 4. From the comparison of the last two examples in Fig. 4, we can find that our method generates satisfactory results even on UDC images with strong haziness.

4. Visual Comparison on Simulated Dataset

Fig. 5 presents more visual comparisons with state-of-the-art methods on synthetic UDC images. As can be seen, our method generates sharper edges than the existing SOTA models, e.g., edges of saturation regions in the first and third samples. Compared with the recovered background images of other methods, our results are closer to the ground truths with higher PSNR and SSIM values.

References

[1] Ruicheng Feng, Chongyi Li, Huaijin Chen, Shuai Li, Chen Change Loy, and Jinwei Gu. Removing diffraction image artifacts in under-display camera via dynamic skip connection network. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 662–671, 2021. 1, 2

[2] Jaihyun Koh, Jangho Lee, and Sungroh Yoon. Bnudc: A two-branched deep neural network for restoring images from

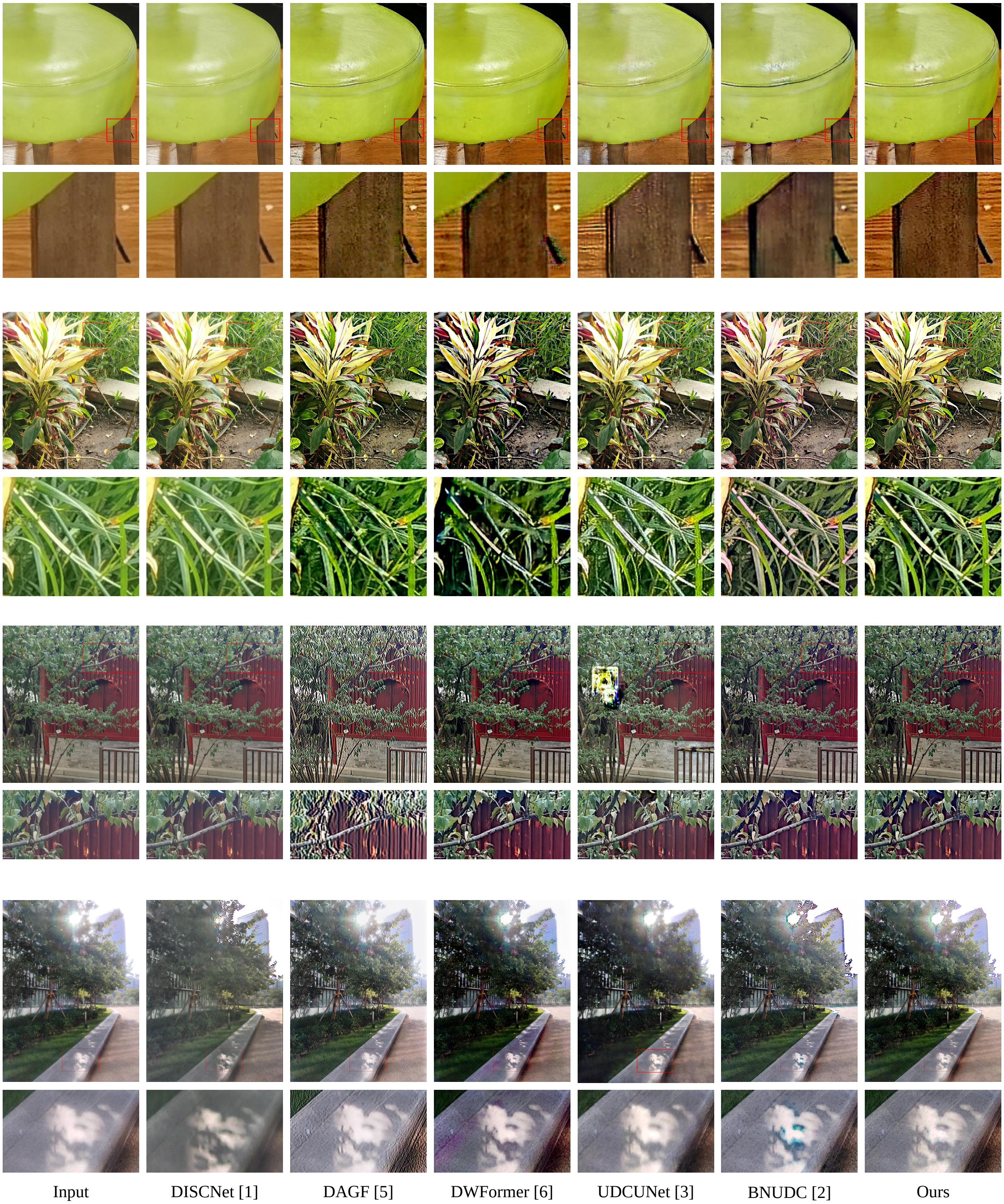
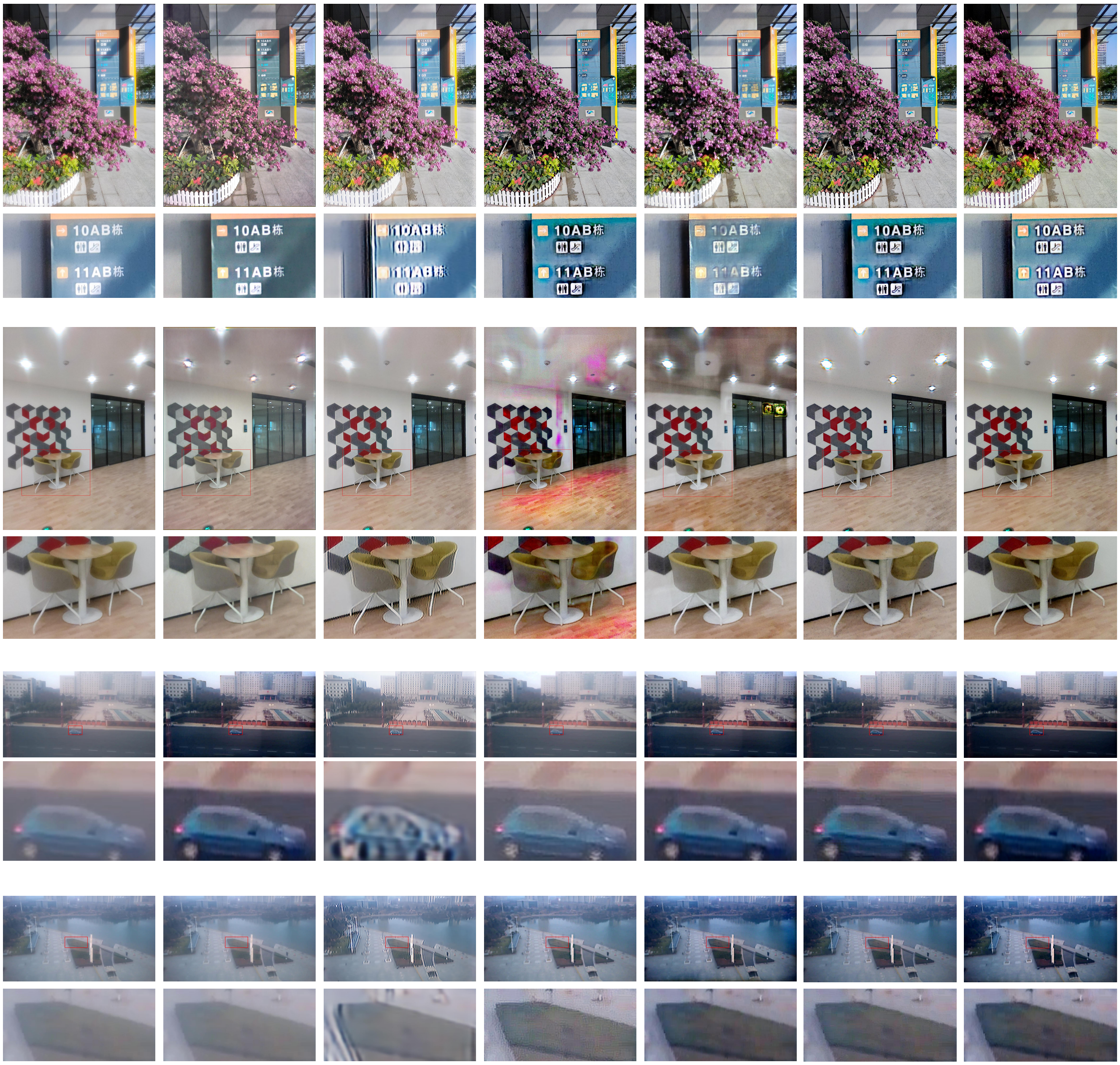


Figure 3. Visual comparison on real-world images.



Input DISCNet [1] DAGF [5] DWFormer [6] UDCUNet [3] BNUDC [2] Ours

Figure 4. Visual comparison on real-world images.

under-display cameras. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1950–1959, 2022. 1

[3] Xina Liu, Jinfan Hu, Xiangyu Chen, and Chao Dong. Udcunet: Under-display camera image restoration via u-shape dynamic network. In *Eur. Conf. Comput. Vis. Worksh.*, pages 113–129, 2022. 1

[4] Jun Luo, Wenqi Ren, Tao Wang, Chongyi Li, and Xiaochun Cao. Under-display camera image enhancement via cascaded curve estimation. *IEEE Transactions on Image Processing*, 31:4856–4868, 2022. 1, 2

[5] Varun Sundar, Sumanth Hegde, Divya Kothandaraman, and Kaushik Mitra. Deep atrous guided filter for image restoration in under display cameras. In *Eur. Conf. Comput. Vis. Worksh.*, pages 379–397, 2020. 1

[6] Yang Zhou, Yuda Song, and Xin Du. Modular degradation simulation and restoration for under-display camera. In *Proceedings of the Asian Conference on Computer Vision*, pages 265–282, 2022. 1

Figure 5. Visual comparison on synthetic images.