Supplementary Material
Spatially Aware Metadata for Raw Reconstruction

Abhijith Punnappurath Michael S. Brown
Samsung AI Center – Toronto
{abhijith.p,michael.b1}@samsung.com

This supplementary material provides additional results and details that could not be included in the main paper due to space constraints.

Raw reconstruction We first provide additional results for raw reconstruction in Fig. S1. This figure extends Fig. 3 of our main paper. It can be observed from the per-pixel error maps that our proposed method produces more accurate raw reconstruction results than competing approaches.

In our experiments on raw recovery in Section 4.1 of our main paper, we had used seven of the total of nine cameras in the NUS dataset [3]. Our approach requires the demosaiced raw-RGB image, which we obtained using the software ISP of [4], since the hardware ISPs of these cameras are inaccessible. Our method assumes that the demosaiced raw-RGB image and its rendered sRGB version are spatially aligned. We found that the hardware ISPs of the Samsung NX2000 and the Fujifilm X-M1 cameras from the NUS dataset apply different lens correction and demosaicing routines compared to the software ISP of [4]. This leads to a spatial misalignment between the sRGB images (produced by the cameras' hardware ISP) and the demosaiced raw-RGB images (obtained using the software ISP of [4]). Therefore, we had omitted images from the Samsung NX2000 and the Fujifilm X-M1 cameras for this experiment. We would like to highlight that this spatial misalignment is an outcome of the hardware ISPs being inaccessible, and is not to be construed as our proposed method's lack of generality.

Re-rendering to sRGB We also present additional results for sRGB re-rendering in Fig. S2. This figure extends Fig. 4 of the main paper. We compare against the UPI [2] and Cycle [6] methods that model the ISP in both the forward and reverse directions. It can be seen from the results that we outperform both UPI and Cycle.

Exposure correction In Section 4.4 of the main paper, we had demonstrated our method's usefulness for exposure correction. Table S1 provides quantitative results on the Samsung NX2000 camera from the NUS dataset [3] which we had used as our test set. Comparisons are provided against the 'Auto tone' feature in Photoshop, and the recent exposure correction method DIEC [7] that handles both underexposed and overexposed images. We selected the DIEC algorithm in particular because most existing exposure correction methods are designed solely for correcting either underexposure or overexposure errors, with the vast majority of them, including recent deep learning methods, focusing on the underexposure scenario [7].

Table S1. Quantitative exposure correction results on the Samsung NX2000 camera images from the NUS dataset [3] averaged over -2 EV (underexposure error) and +1 EV (overexposure error).

Method	PSNR (dB) / SSIM			
	Mean	**Median**	**Worst 25%**	**Best 25%**
Auto tone	16.43	15.93	13.01	20.83
Photoshop	0.874	0.906	0.741	0.950
DIEC [7]	18.79	18.73	14.40	23.39
	0.880	0.893	0.790	0.941
Ours	**41.99**	**42.58**	**37.26**	**45.92**
	0.995	**0.996**	**0.991**	**0.998**

Motion deblurring We had demonstrated our method's applicability to the problem of motion deblurring in Section 4.5 of our main paper. Table S2 quantitatively illustrates the advantage of deblurring in the raw-RGB space using our proposed approach versus directly in sRGB.

Ablation study As explained in the last paragraph of Section 3 of our main paper, our proposed spatially aware raw reconstruction pipeline is implemented patch-wise. In the main paper, we had presented a comparison against a 'W/O spatial' variant of our method that used all N corresponding raw-sRGB samples at once to compute a *global*

Table S2. Quantitative motion deblurring results on the Samsung NX2000 camera images from the NUS dataset [3].

Method	PSNR (dB) / SSIM			
	Mean	**Median**	**Worst 25%**	**Best 25%**
sRGB	24.77	24.35	19.81	30.49
	0.779	0.792	0.647	0.893
Ours	**27.10**	**26.57**	**22.10**	**32.78**
	0.903	**0.937**	**0.778**	**0.972**

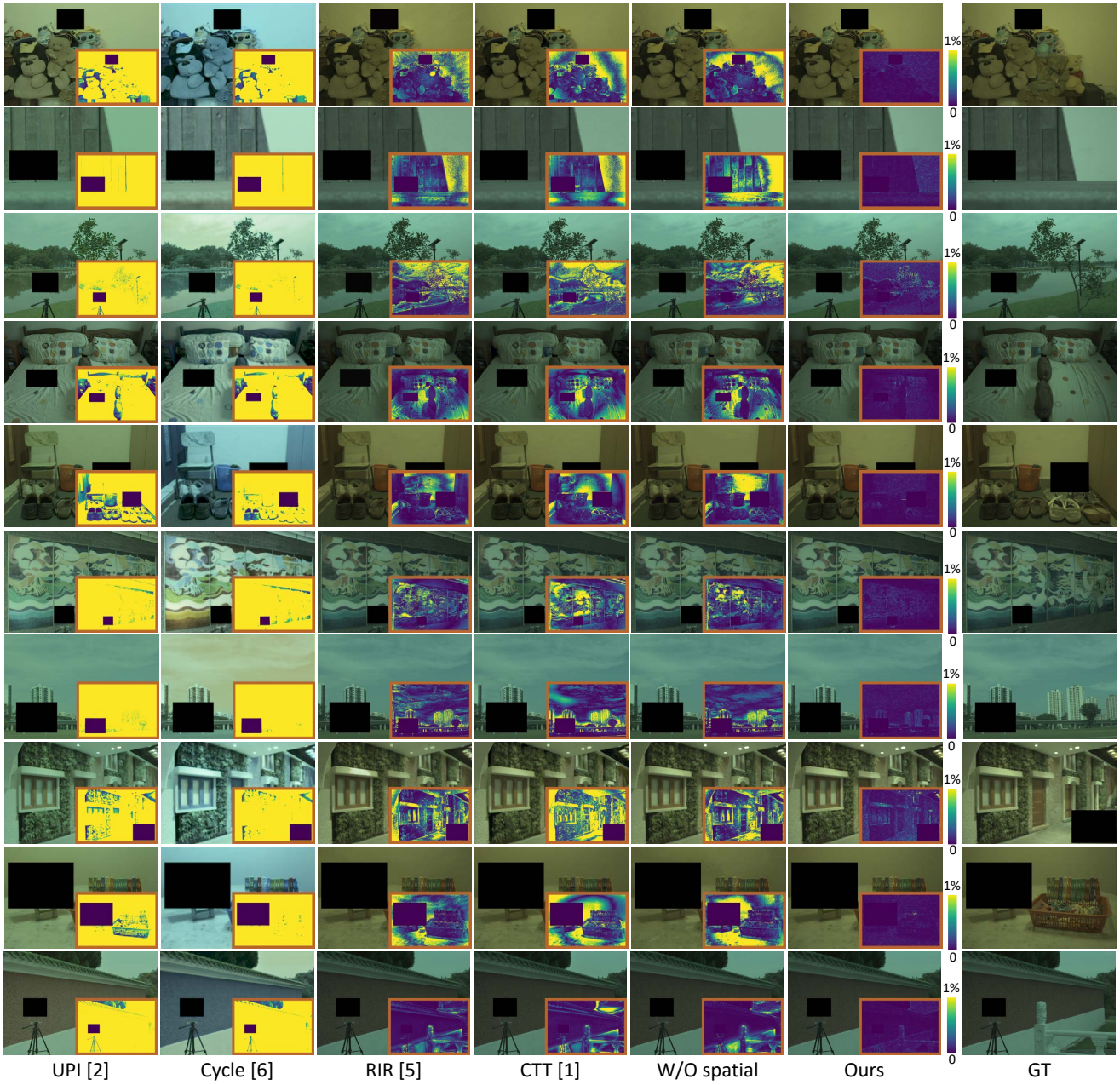

Figure S1. Raw reconstruction. This figure extends Fig. 3 of our main paper. Results of the competing approaches of [2, 6, 5, 1] are presented. A variant of our method, denoted 'W/O spatial', that constructs a global mapping function by considering only pixel RGB values and ignoring spatial information, is also provided for comparison. Our results and the ground truth (GT) are shown in the last two columns, respectively. The corresponding per-pixel error maps are shown as insets. Note that a gamma function has been applied to all raw-RGB images for better visualization.

mapping function, similar to existing methods. The mapping function considers only the RGB intensity values of the samples, but ignores spatial information. In particular, $\mathbf{s}_i = (s_{R_i}, s_{G_i}, s_{B_i})$, $1 \leq i \leq N$, and the last two coordinates (s_{X_i}, s_{Y_i}) are discarded during the interpolation process. Here, we provide another baseline comparison that uses only the RGB intensity values similar to W/O spatial. However, unlike W/O spatial, the interpolation is performed patch-wise similar to our proposed method. This lends a spatial component to this variant because although only RGB intensity values are used, the mapping function itself varies per patch. We denote this baseline as 'W/O spa-

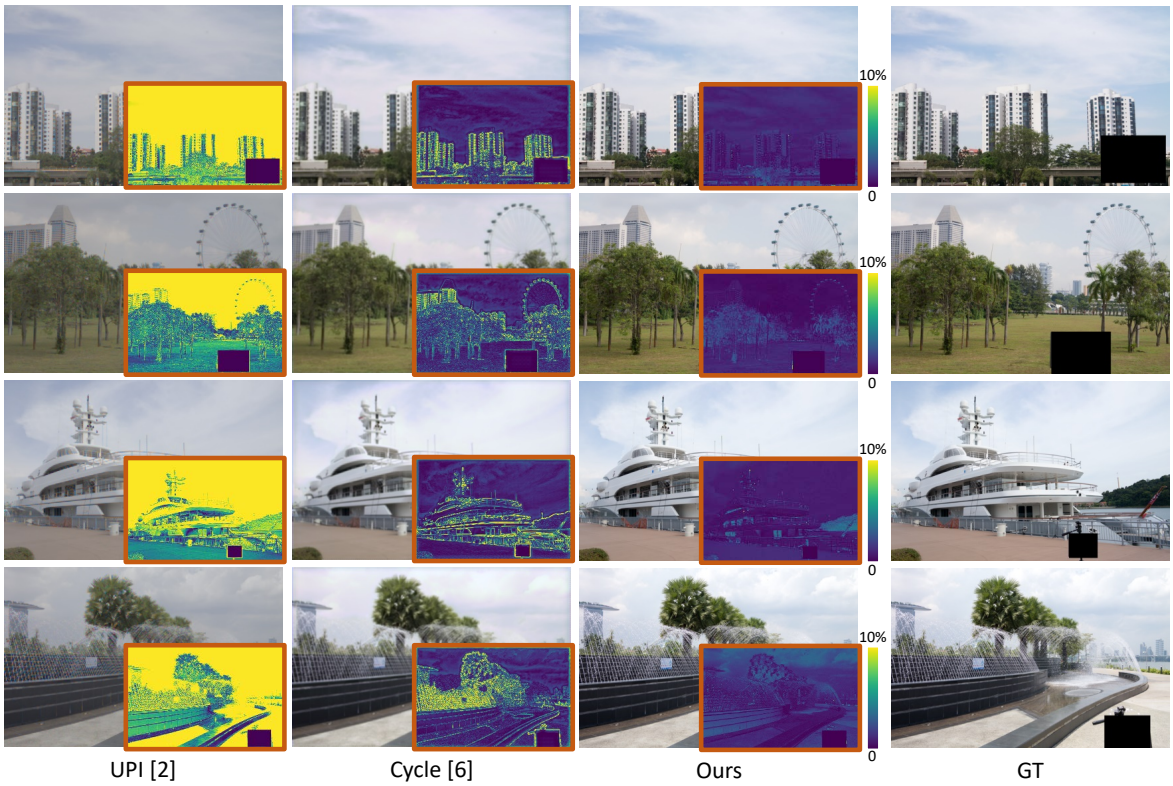

Figure S2. Re-rendering to sRGB. This figure extends Fig. 4 of our main paper. Results of [2, 6] and our method are provided. The ground truth is presented in the last column. The corresponding per-pixel error maps are shown as insets.

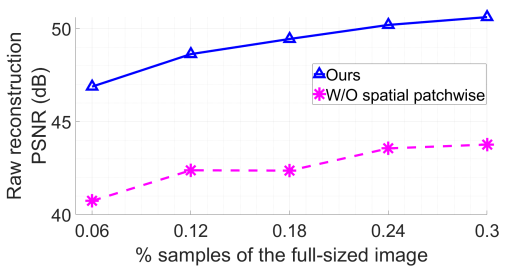

Figure S3. Comparison between our proposed method and a W/O spatial variant run patchwise.

tial patchwise'. A performance evaluation between our proposed method and the W/O spatial patchwise variant run on the 200 images from the Canon 600D camera is presented in the plot of Fig. S3. It can be observed that our proposed method's raw reconstruction accuracy is significantly higher validating the importance of including the spatial coordinates into the interpolation framework.

References

[1] Mahmoud Afifi, Abhijith Punnappurath, Abdelrahman Abdelhamed, Hakki Can Karaimer, Abdullah Abuolaim, , and Michael S. Brown. Color temperature tuning: Allowing accurate post-capture white-balance editing. In *Color Imaging Conference*, 2019.

[2] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T. Barron. Unprocessing images for learned raw denoising. In *Computer Vision and Pattern Recognition*, 2019.

[3] Dongliang Cheng, Dilip K. Prasad, and Michael S. Brown. Illuminant estimation for color constancy: Why spatial-domain methods work and the role of the color distribution. *Journal of the Optical Society of America A*, 31(5):1049–1058, 2014.

[4] Hakki Can Karaimer and Michael S. Brown. A software platform for manipulating the camera imaging pipeline. In *European Conference on Computer Vision*, 2016.

[5] Rang M. H. Nguyen and Michael S. Brown. RAW image reconstruction using a self-contained sRGB-JPEG image with only 64 KB overhead. In *Computer Vision and Pattern Recognition*, 2016.

[6] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. CycleISP: Real image restoration via improved data synthesis. In *Computer Vision and Pattern Recognition*, 2020.

[7] Qing Zhang, Yongwei Nie, and Wei-Shi Zheng. Dual illumination estimation for robust exposure correction. *Computer Graphics Forum*, 38(7):243–252, 2019.