Point Cloud Color Constancy – Supplement Material

Figure 1. AWB-biased images and their corrected version with our labeled ground-truth illuminations from DIODE datasets [9] and NYU-v2 [7], respectively.

A. Dataset Preprocessing

Introduced in paper Section 4, we use three public RGBD datasets and collect DepthAWB dataset for point cloud color constancy. We will detail the procedure of labeling the datasets and collecting our dataset.

A.1. Incorrect White Balance

In camera ISP auto white balance (AWB) is a fundamental processing step, where the incorrect white balance images may exist due to fragile AWB algorithm or improper manual setting. Figure 1 shows the AWB-biased samples we found in the NYU-v2 dataset [7] and DIODE dataset [9], these images are linearized sRGB image.

A.2. Public Datasets Labeling

Cheng *et al.* [2] identified 66 images with two illuminations from Gehler-Shi dataset [3, 6] , and relabeled two illumination vectors by finding neutral objects in the scene. We use the same idea to label the dominant illumination for NYU-v2 [7], DIODE [9] and ETH3d [5] datasets. Given a raw image **I**, we first locate several areas under the cast of the dominant light, e.g. a piece of printed paper under the sunlight outside the window, and a gray wall illuminated by the ceiling light in the room. Then we visually check whether the area is saturated or underexposed, and select the well-exposed part. As the final step we average the rgb instensities of the selected area $\mathbf{I}_{sel}$, yielding the global illumination **E**:

$$\mathbf{E} = \frac{1}{N} \sum_{R,G,B} \mathbf{I}_{sel}. \tag{1}$$

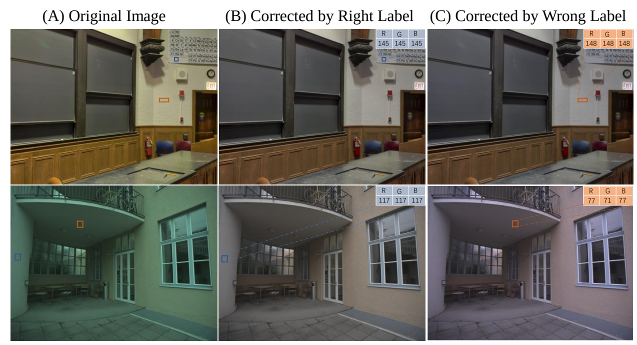

Figure 2. Illustration of the procedure of labeling the illumination by finding canonical neutral surface. We mark the blue and red boxes in column (A) indicating the right and wrong area of ground truth illumination calculation, respectively. Column (B) presents images corrected with right label. Column (C) shows images with wrong label. $\gamma = 2.2$ is applied for visualization.

Figure 2 shows the resulting corrected images if we label correctly or with some mistake.

A.3. Dataset Collection

Our DepthAWB dataset contains of three types data, 1) RGB images by GH5s camera, 2) depth maps by Intel Realsense L515 LiDAR camera, and 3) illumination ground-truth by labeling from Xrite Color Checker.

Images from our dataset are mainly collected under indoor scenarios, considering the following reasons: 1) limited working range of the depth camera, similar as NYU-v2 [7] and SUN RGB-D [8], and 2) comparing to the outdoor scenarios, the indoor scenes have more artificial and mixed illumination which makes the illumination estimation more challenging.

We collect our data in several different places to ensure diverse scenes and wide illumination distribution. We are interested in challenging scenes, *e.g.*, shopping mall,

Figure 3. Selected scenes from DepthAWB dataset

Figure 4. Our data capture setup.

large pure color area, and the scene containing illumination-like misleading color. We also pay attention to areas under mixed-illumination but have one dominant illumination. Figure 3 shows several selected scenes from our dataset.

Our data capture setup is shown in Figure 4, the Intel Realsense L515 depth camera is set on the top of Panasonic Lumix GH5s camera.

We fix the camera focal length during shooting, the camera aperture is set as $f/8.0$, and we use aperture priority mode to yield a clear vision and less noise. RGB camera is placed on a tripod and ToF camera keeps the rigid connection with RGB camera.

We create a script to collect RGB images and depth maps from two cameras simultaneously, with a single click on the laptop keyboard.

A reliable illumination label lays on putting the color checker in the right position. In each scene, we put the color checker under the dominant illumination, and make sure the color checker reflects the dominant illumination into the camera. For the similar scenes under same dominant illumination, we direct use the first labeled illumination. We took two images at one single scene, one image with color checker is for the illumination labeling, and another without color checker is for the training and testing period.

We obtain the ground-truth illumination using the right exposed grayscale color on color checker.

A.4. Illumination Distribution

Theoretically, computational color constancy is to learn a mapping pattern from images in wild to illumination in wild. Therefore, a fair and complete illumination distribution is needed. We compare our collection and relabeled datasets to two widely used color constancy datasets on illumination distribution. Figure 4. in main paper shows the illumination distribution of each datasets, similar as the NUS-600D color constancy datasets, the illumination distribution of our datasets are also following the blackbody radiation curve. And Our datasets contain a relatively large range of illumination.

B. Camera Calibration

We present the camera intrics we used or calibrated in three different datasets.

B.1. DepthAWB Dataset

DepthAWB dataset RGB camera (Under (480,640)):

$$\begin{bmatrix} 464.0010 & 0 & 319.4235 \\ 0 & 463.1813 & 241.4676 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

DepthAWB dataset RGB camera (Under (240,320)):

$$\begin{bmatrix} 232.1495 & 0 & 157.8825 \\ 0 & 232.0269 & 123.0501 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

DepthAWB dataset Depth camera (Under (480,640)):

$$\begin{bmatrix} 470.2773 & 0 & 295.0742 \\ 0 & 470.2187 & 242.5917 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

DepthAWB dataset Depth camera (Under (240,320)):

$$\begin{bmatrix} 233.5576 & 0 & 148.4298 \\ 0 & 233.6542 & 125.9854 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

Rotation:

$$\begin{bmatrix} 0.9999 & 0.0096 & -0.0103 \\ -0.0097 & 0.9999 & -0.0063 \\ 0.0102 & 0.0064 & 0.9999 \end{bmatrix} \tag{6}$$

Shift:

$$\begin{bmatrix} 0.8710 & -105.6946 & -90.1409 \end{bmatrix} \tag{7}$$

B.2. NYU-v2 Dataset

NYU dataset RGB camera (Under (480,640)):

$$\begin{bmatrix} 525.0 & 0 & 319.5 \\ 0 & 525.0 & 329.5 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

B.3. DIODE Dataset

DIODE dataset RGB camera (Under (480,640)):

$$\begin{bmatrix} 866.81 & 0 & 512.0 \\ 0 & 927.06 & 384.0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

B.4. ETH3D Dataset

ETH3D presents different RGB camera intrics on different scene, where we only use a common setting in our point cloud building, since we only need the relative distance between pixels.

ETH3D dataset RGB camera (Under (4032,6048)):

$$\begin{bmatrix} 3406.79 & 0 & 3040.861 \\ 0 & 3404.57 & 2014.4 \\ 0 & 0 & 1 \end{bmatrix} \tag{10}$$

C. More Visual Results

We present the visual results of global and local illumination estimation.

C.1. Global Illumination Estimation

We show more visual comparison results with two state-of-the-art color constancy methods in Figure 5. As introduced in Section 5 of main paper, our PCCC performs well in majority scenes, yet it also has some limitation in solving huge gray-like area (row 1&2 of Figure 5). Our method outperforms other methods on other scenes, including, dominating pure-color surface (row 3-5 of Figure 5).

C.2. Local Illumination Estimation

Unlike thumbnail size point clouds we used in global illumination estimation task, we feed full-scale point cloud to the learned PCCC to obtain higher resolution illumination map.

Pixel-wise illumination correction results are achieved by the pixel-wise map and presented in Figure 6. We select the multi-illuminant scenes from our DepthAWB dataset and NYU-v2 dataset [7]. The second row of Figure 6 shows images recovered by the illumination map. In column 1, our method corrected the illuminations from outside and the windows, separately. In column 4, our method also recovered the outdoor illumination, while keeping the indoor light the same.

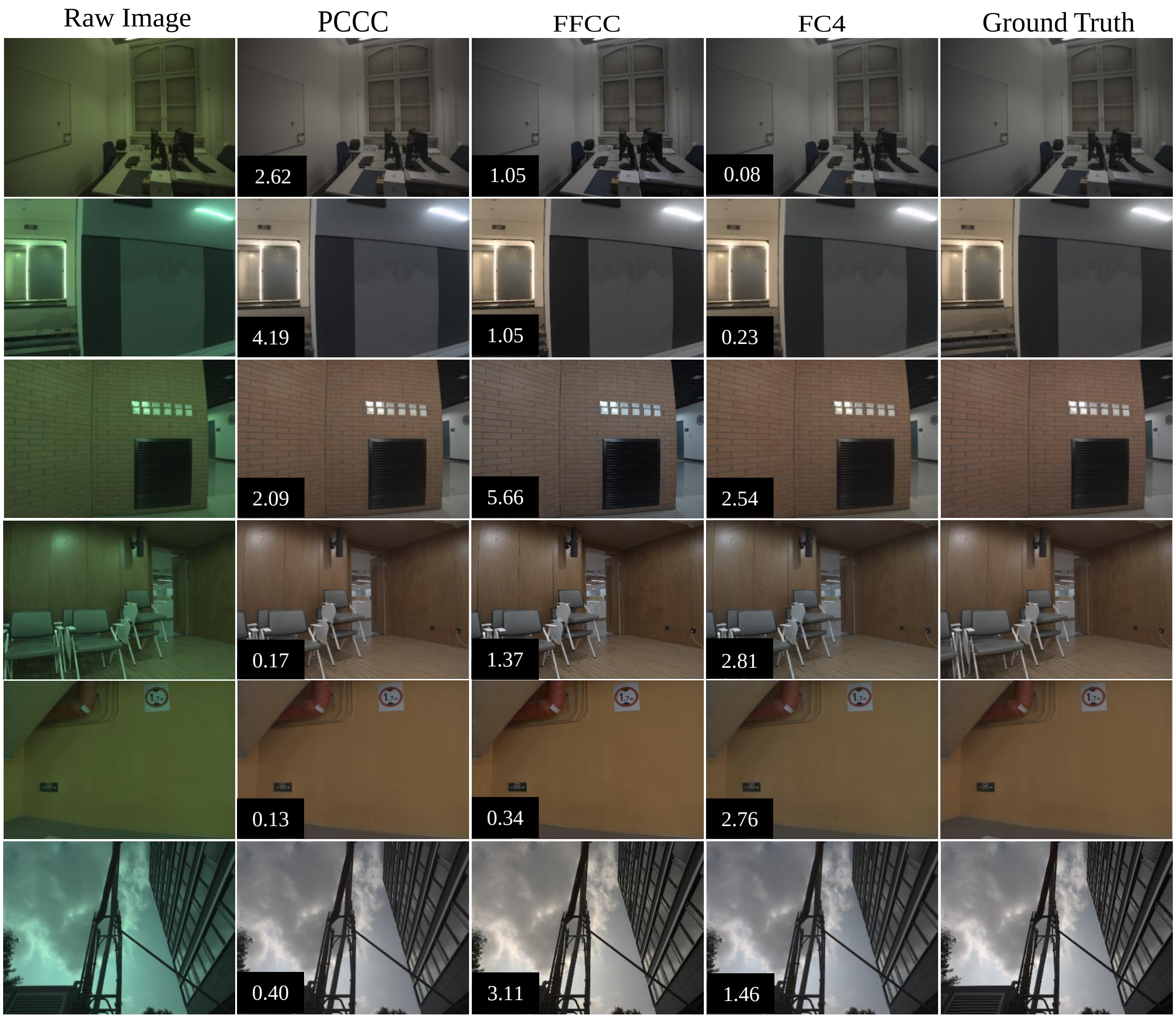

Figure 5. More visual results of our methods comparing with FC4 [4] and FFCC [1]. The angular error of each image is listed in the black box. We present our failure (angular error > 2) and well (angular error < 0.5) estimation results.

Figure 6. More visual results of our methods on pixel-wise illumination estimation, we select dual-illuminant scenes from DepthAWB and NYU-v2 datasets, the first row is the original images, and the second row shows images recovered by our pixel-wise illumination map.

References

[1] J. T. Barron and Y.-T. Tsai. Fast fourier color constancy. In *CVPR*, 2017. 4

[2] D. Cheng, A. Kamel, B. Price, S. Cohen, and M. S. Brown. Two illuminant estimation and user correction preference. In *CVPR*, 2016. 1

[3] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp. Bayesian color constancy revisited. In *CVPR*, 2008. 1

[4] Y. Hu, B. Wang, and S. Lin. FC4: Fully convolutional color constancy with confidence-weighted pooling. In *CVPR*, 2017. 4

[5] T. Schps, J. L. Schnberger, S. Galliani, T. Sattler, and A. Geiger. A multi-view stereo benchmark with high-resolution images and multi-camera videos. In *CVPR*, 2017. 1

[6] L. Shi and B. Funt. Re-processed version of the gehler color constancy dataset of 568 images. *accessed from* `http://www.cs.sfu.ca/~colour/data/`, 2010. 1

[7] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, pages 746–760. Springer, 2012. 1, 3

[8] S. Song, S. P. Lichtenberg, and J. Xiao. Sun rgb-d: A rgb-d scene understanding benchmark suite. In *CVPR*, pages 567–576, 2015. 1

[9] I. Vasiljevic, N. Kolkin, S. Zhang, R. Luo, H. Wang, F. Z. Dai, A. F. Daniele, M. Mostajabi, S. Basart, M. R. Walter, et al. Diode: A dense indoor and outdoor depth dataset. *arXiv preprint arXiv:1908.00463*, 2019. 1