Attentive Illumination Decomposition Model for Multi-Illuminant White Balancing

Supplementary Material

A. More Qualitative Results

More qualitative results are shown in this section. We include the prediction results of all the slots that could not be included in the paper due to spatial constraints. As noted in the paper, we used seven slots for the multi-illuminant WB experiment on the LSMI [2] and MIIW [3] dataset and five slots for single-illuminant WB on the NUS-8 dataset [1]. The LSMI galaxy subset results are presented in Fig. 2 and 3, while Fig. 4, 5, and Fig. 6, 7 display the Nikon and Sony results, respectively. The angular errors between the ground truths and the predicted results are indicated in degrees at the upper right corner of the illumination maps. Fig. 9 shows the test result of NUS-8. Fig. 8 shows a visualization of the prediction results for MIIW dataset. Each row in the figure shows the results for a 1,2,3-illuminant scene, respectively. All visualizations demonstrate that each slot is effectively activated by the color gamut assigned to it through centroid-matching loss, while the remaining slots remain inactive. We emphasize that the AID framework demonstrates consistent and robust predictions of illuminant chromaticity and weight maps across the LSMI, MIIW, and NUS-8 datasets.

B. Slot Evolution through Iterative Calibraion

We also visualize the gradual evolution results of every slot and weight map. As shown in Fig. 10, the chromaticity and weight map of each slot become more accurate as the iteration progresses.

C. Importance of Centroid-Matching Loss

As demonstrated in the paper, our centroid-matching loss guides the GRU to update slots more accurately. Here, we visualize the test result of AID trained without centroid-matching loss $\mathcal{L}_{centroid}$ to clarify the effect of $\mathcal{L}_{centroid}$. Fig. 11 shows that without the matching loss $\mathcal{L}_{centroid}$, every slot tries to engage and evolve its own chromaticity to approximate the ground truth illumination, resulting in conflicts among the slots. On the other hand, by using the matching loss, each slot specializes in a specific chromaticity range, and the slots that do not match the GT illuminant chromaticity are deactivated, resulting in more accurate illumination predictions for the scene.

D. More Implementation Details

For all experiments in the main paper, we use square images with a size of 256 for both training and testing.

AID Framework We only describe two components, $\mathcal{L}_{mixed}$ and $\mathcal{L}_{centroid}$, that constitute the loss term in the main paper due to the spatial constraint. The total training loss is the sum of these two terms, and it is trained through Eq. (1).

$$\mathcal{L}_{mixed} + \mathcal{L}_{centroid} \tag{1}$$

Experiment on NUS-8 To maintain consistency with previous works, we employ the same experimental setup, which involves conducting three-fold cross-validation for each camera and measuring the mean, median, tri-mean, best 25%, and worst 25% MAE in degrees. In line with the methodology of previous works, we report the geometric mean of each metric obtained from the eight cameras in the main paper. To augment the training data, we randomly sample ground truth illumination from a pool for each training fold and apply it to the scene during training.

Multi-Illumination In the Wild The Multi-Illumination In the Wild dataset comprises images with 25 different lighting settings for each scene, encompassing over a thousand diverse scenes. We generated train, validation, and test sets using a methodology similar to that employed in the original paper [3]. For each given scene, we selected one to three images to create 1- to 3-illuminant images, and the weight map α was generated using the G channel of the sampled images. Illuminant chromaticity was also randomly sampled from 1 to 3, corresponding to the number of sampled images, and multiplied with the weight map to create the ground truth mixed illumination map. Fig. 1 shows generated multi-illuminant scenes, using MIIW dataset. The MIIW dataset includes canonical objects such as gray and chrome balls within the scenes. These objects were masked during training to ensure that the model did not use them as part of its learning process.

E. Failure Cases

Despite the strong performance of AID, we can observe some failure cases. Fig. 12(a) illustrates that the model fails

[White Balanced GT] [RGB weightmap] [Tinted Image]

Figure 1. MIIW dataset visualization. Each row shows 1-, 2-, and 3-illuminant scene.

to capture each GT chromaticity for a multi-illuminant scenario. In (b), the model activates only one slot, although there are two illuminants in the scene. Lastly, (c) shows the opposite case of case 2: activates multiple slots when there is only one illuminant in a scene. We also provide the illumination plots on the rightmost of the figure to easily demonstrate the above cases.

F. Demo Video

We made a demo video to clearly show one of our contributions, which is an ability to controllable multi-illuminant white balance. Since AID generates a fully decomposed illumination map consisting of the chromaticity of each illuminant and a weight map, interactive WB modification is possible even in a multi-illuminant scenario. Please refer to the link for the video.

References

[1] 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 (JOSA A)*, 31(5):1049–1058, 2014. 1

[2] Dongyoung Kim, Jinwoo Kim, Seonghyeon Nam, Dongwoo Lee, Yeonkyung Lee, Nahyup Kang, Hyong-Euk Lee, ByungIn Yoo, Jae-Joon Han, and Seon Joo Kim. Large scale multi-illuminant (lsmi) dataset for developing white balance algorithm under mixed illumination. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2410–2419, 2021. 1

[3] Lukas Murmann, Michael Gharbi, Miika Aittala, and Fredo Durand. A dataset of multi-illumination images in the wild. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4080–4089, 2019. 1

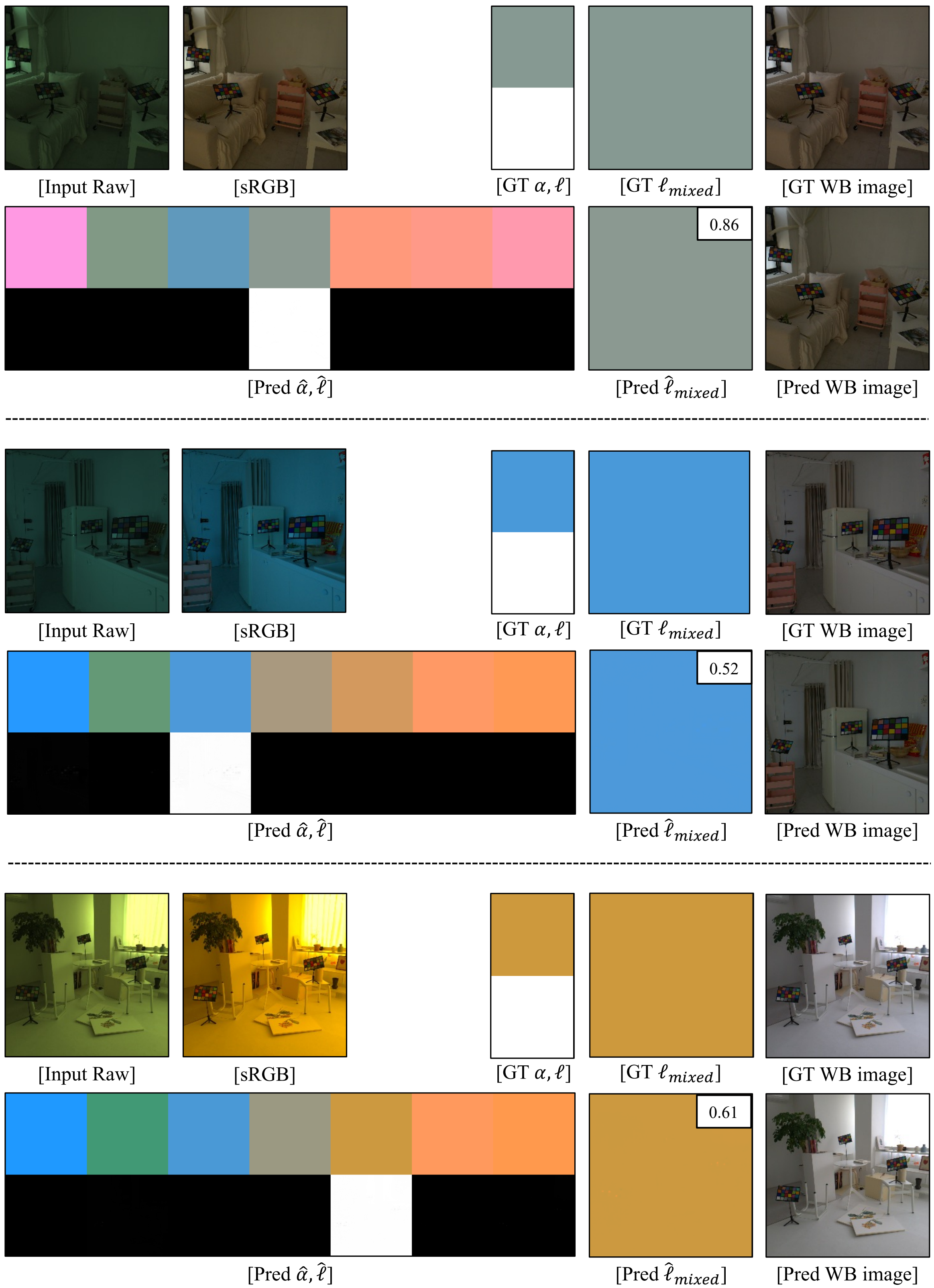


Figure 2. Single-illuminant WB result of LSMI Galaxy test set.

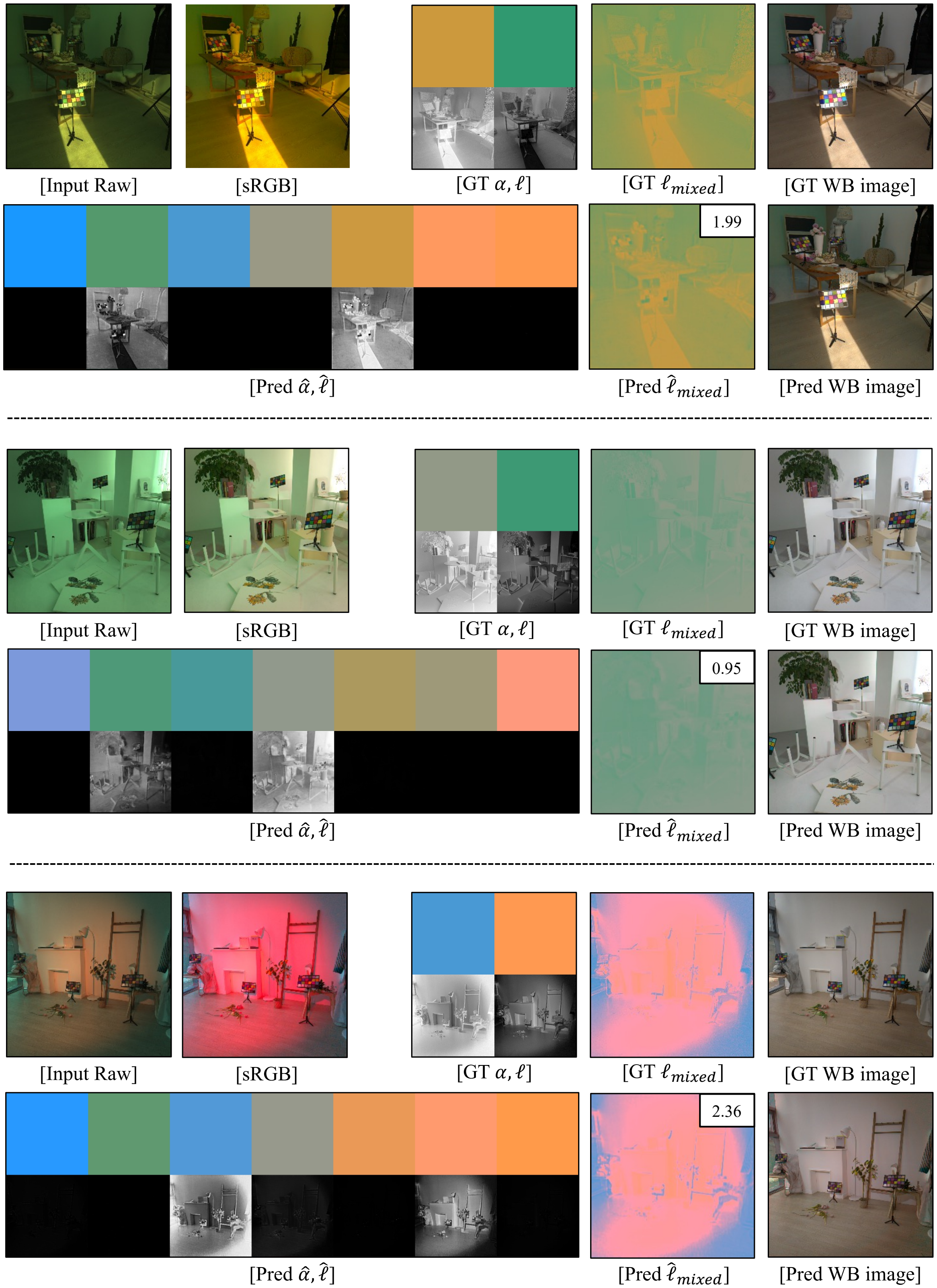


Figure 3. Multi-illuminant WB result of LSMI Galaxy test set.

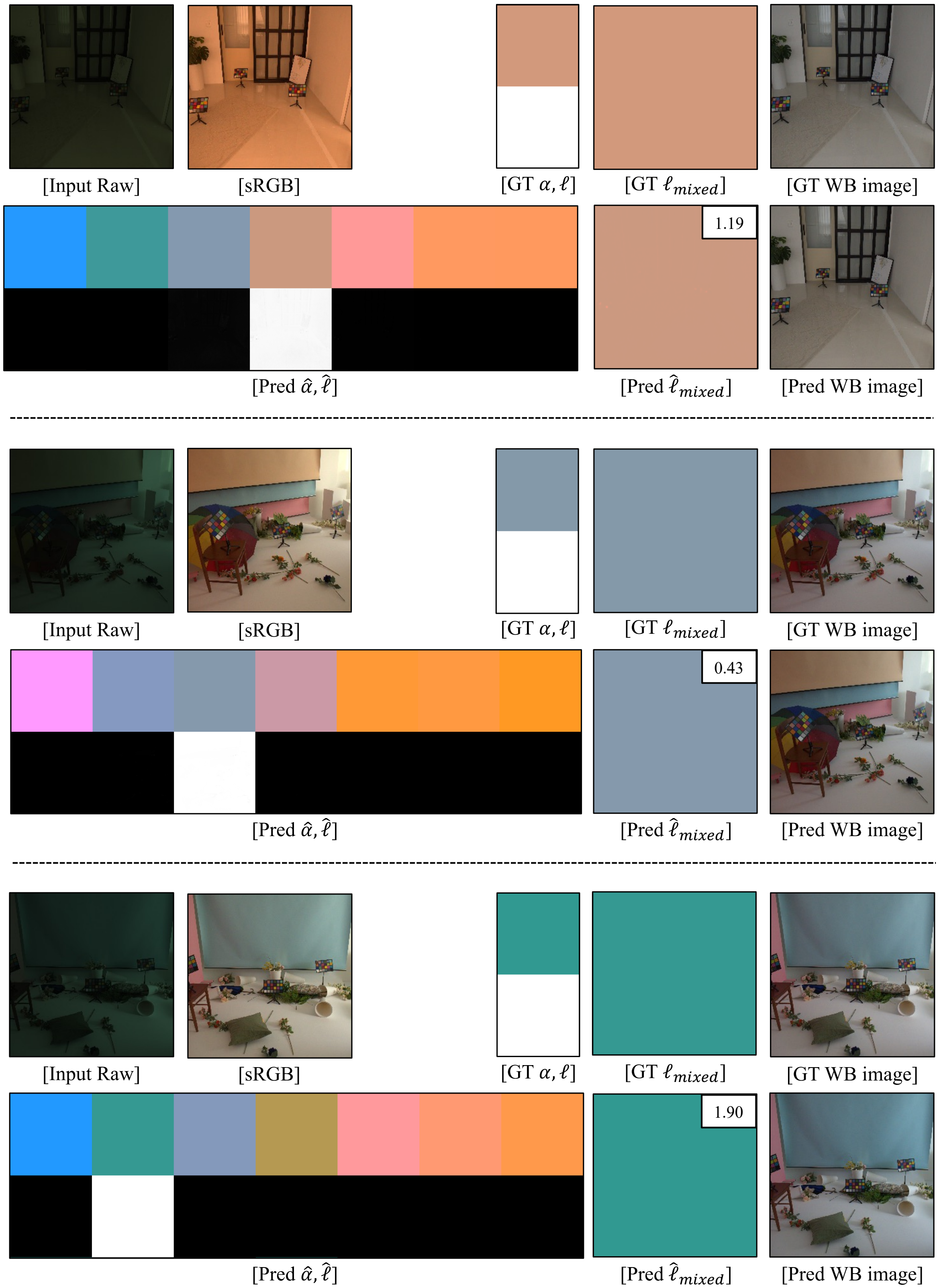


Figure 4. Single-illuminant WB result of LSMI Nikon test set.

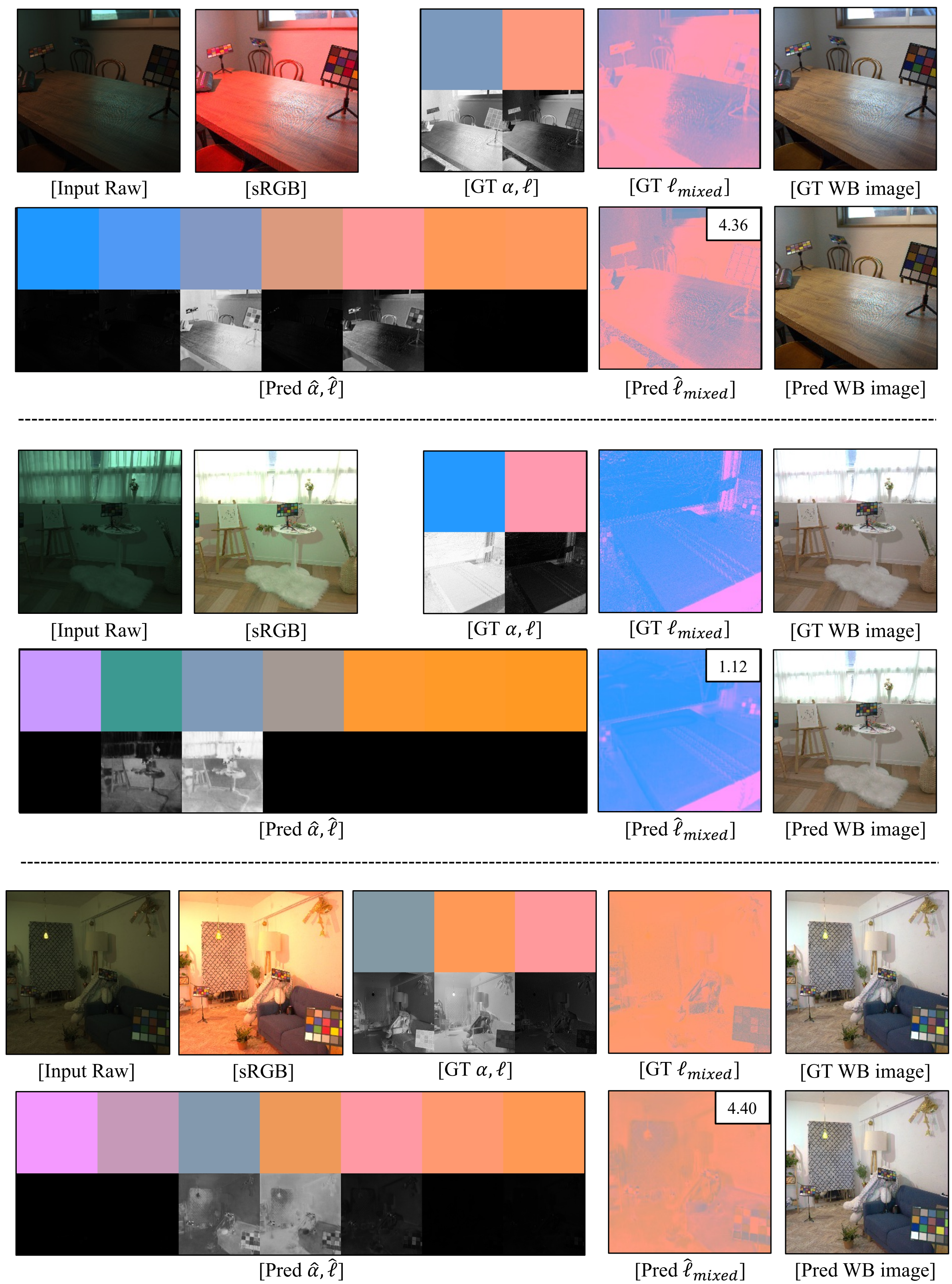


Figure 5. Multi-illuminant WB result of LSMI Nikon test set.

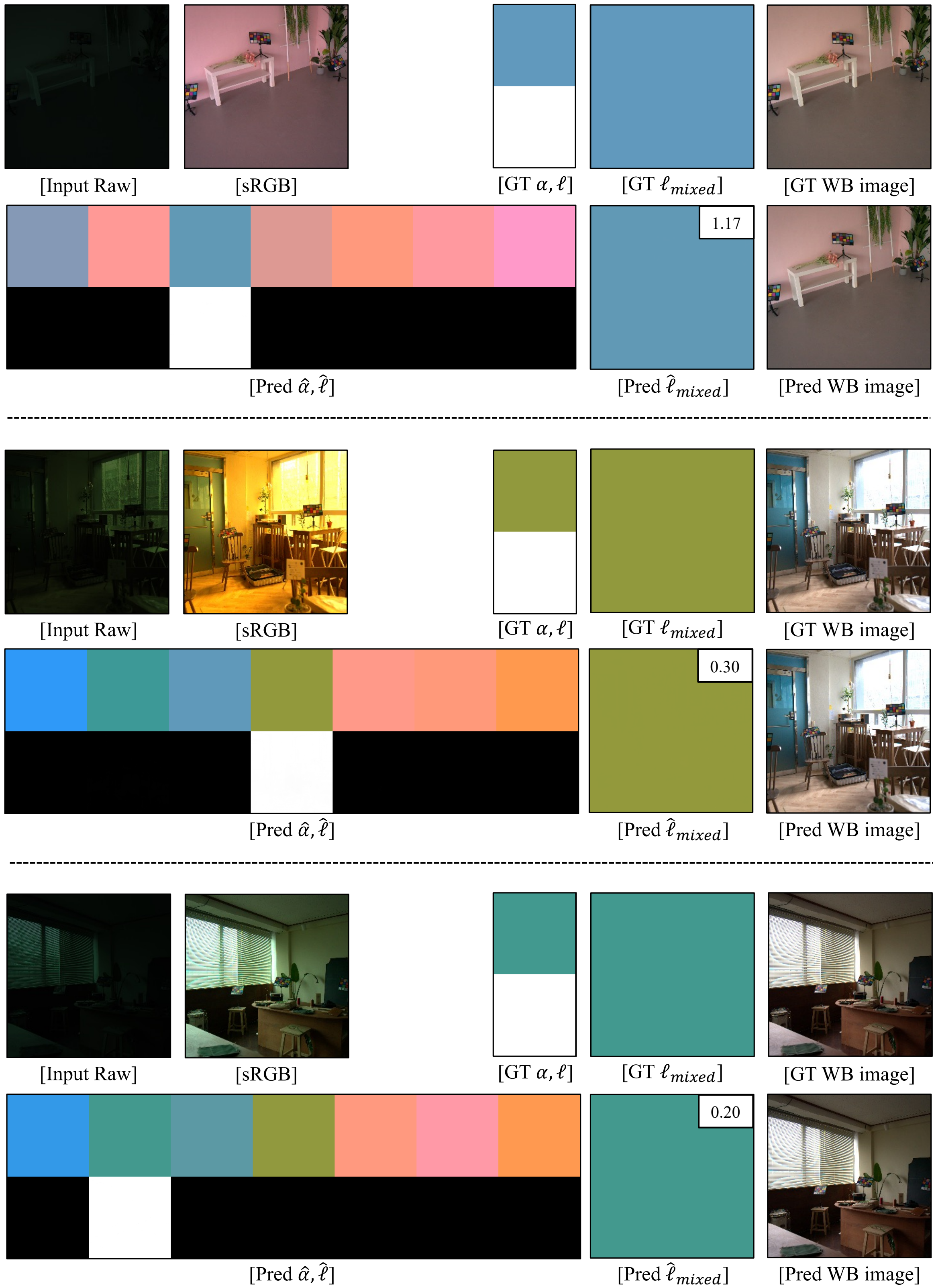


Figure 6. Single-illuminant WB result of LSMI Sony test set.

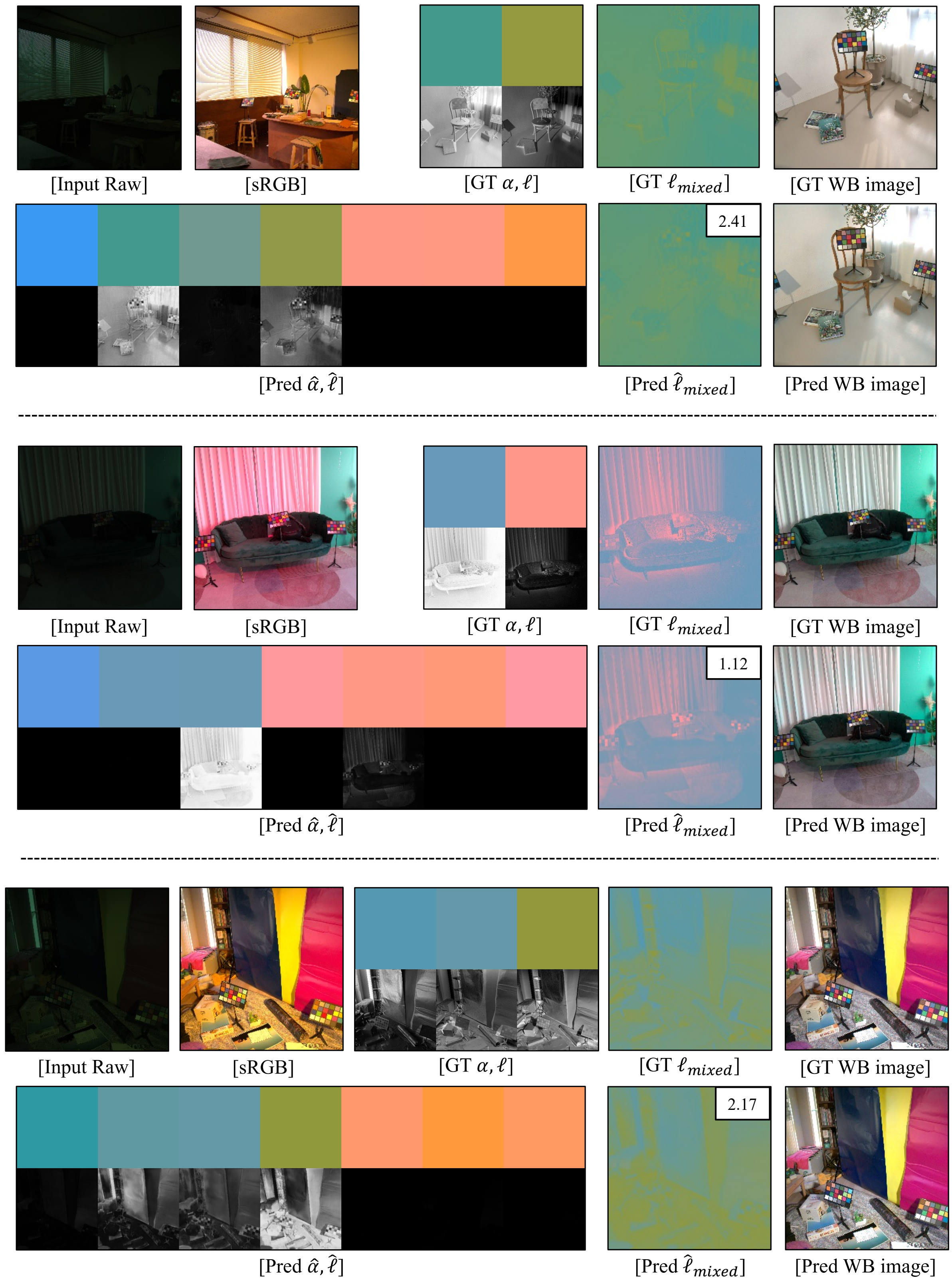


Figure 7. Multi-illuminant WB result of LSMI Sony test set.

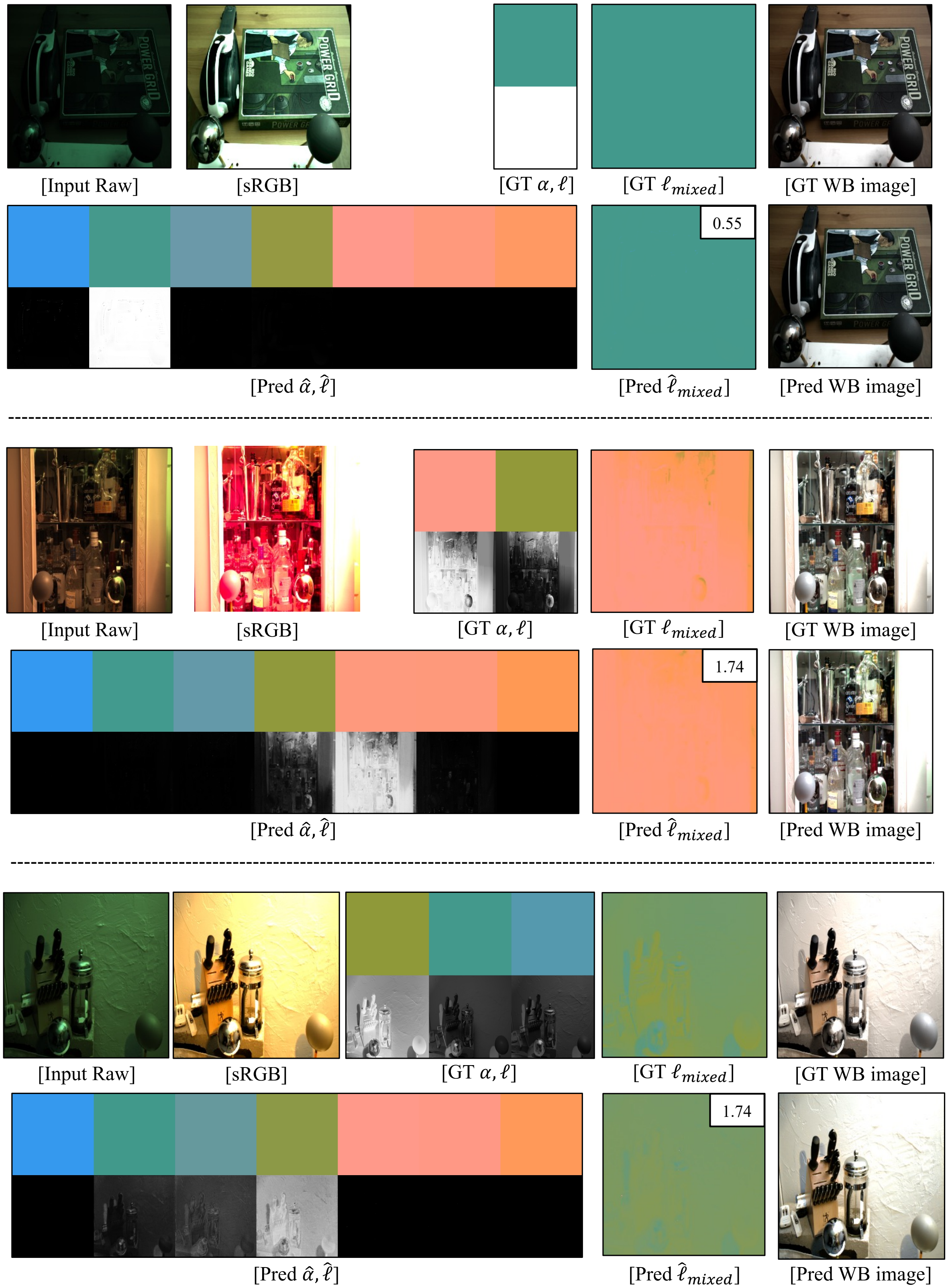


Figure 8. WB results of MIIW dataset (1,2,3 illuminant scene) using AID framework.

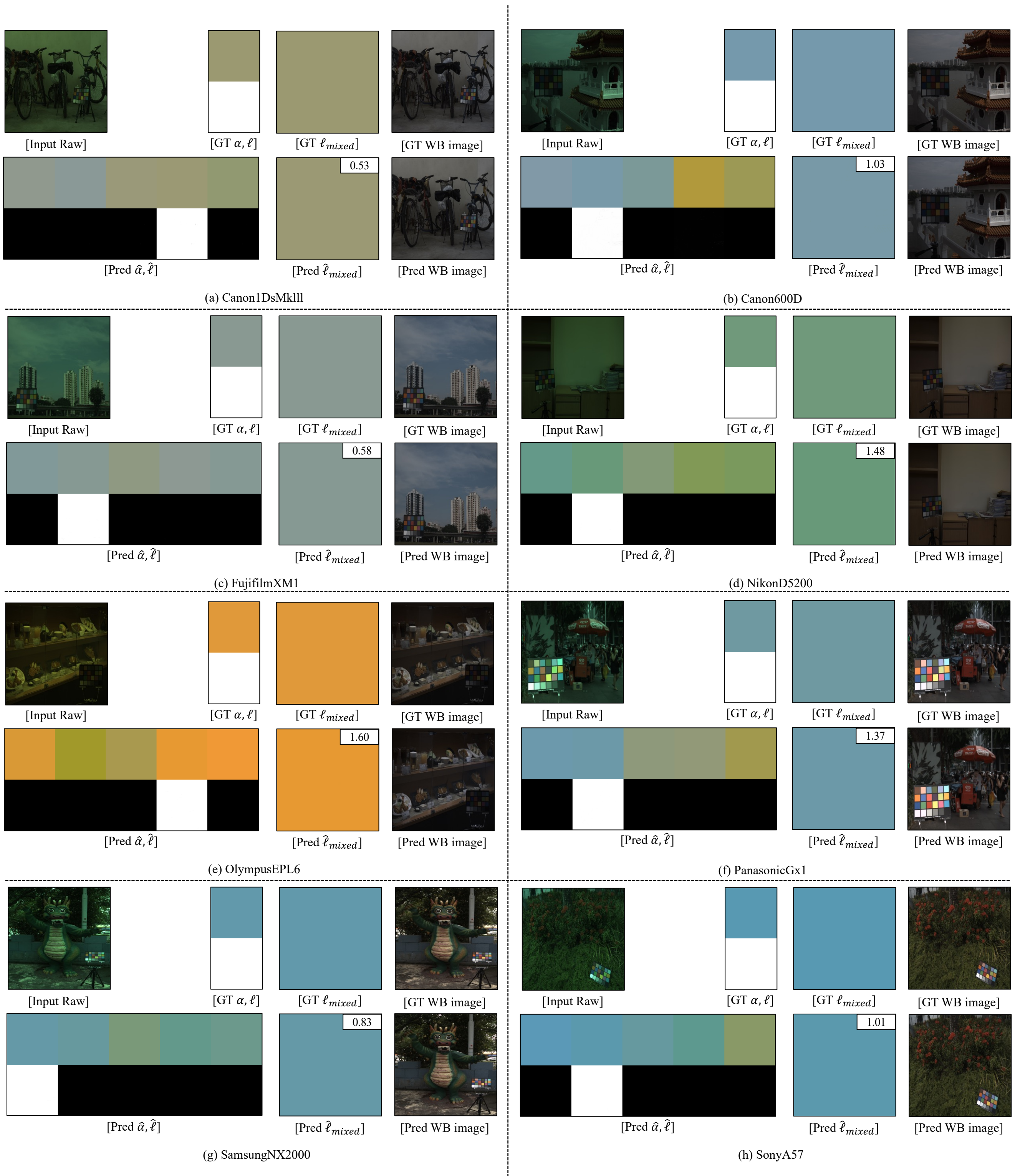


Figure 9. WB results of NUS-8 dataset using AID framework (5 slots are used).

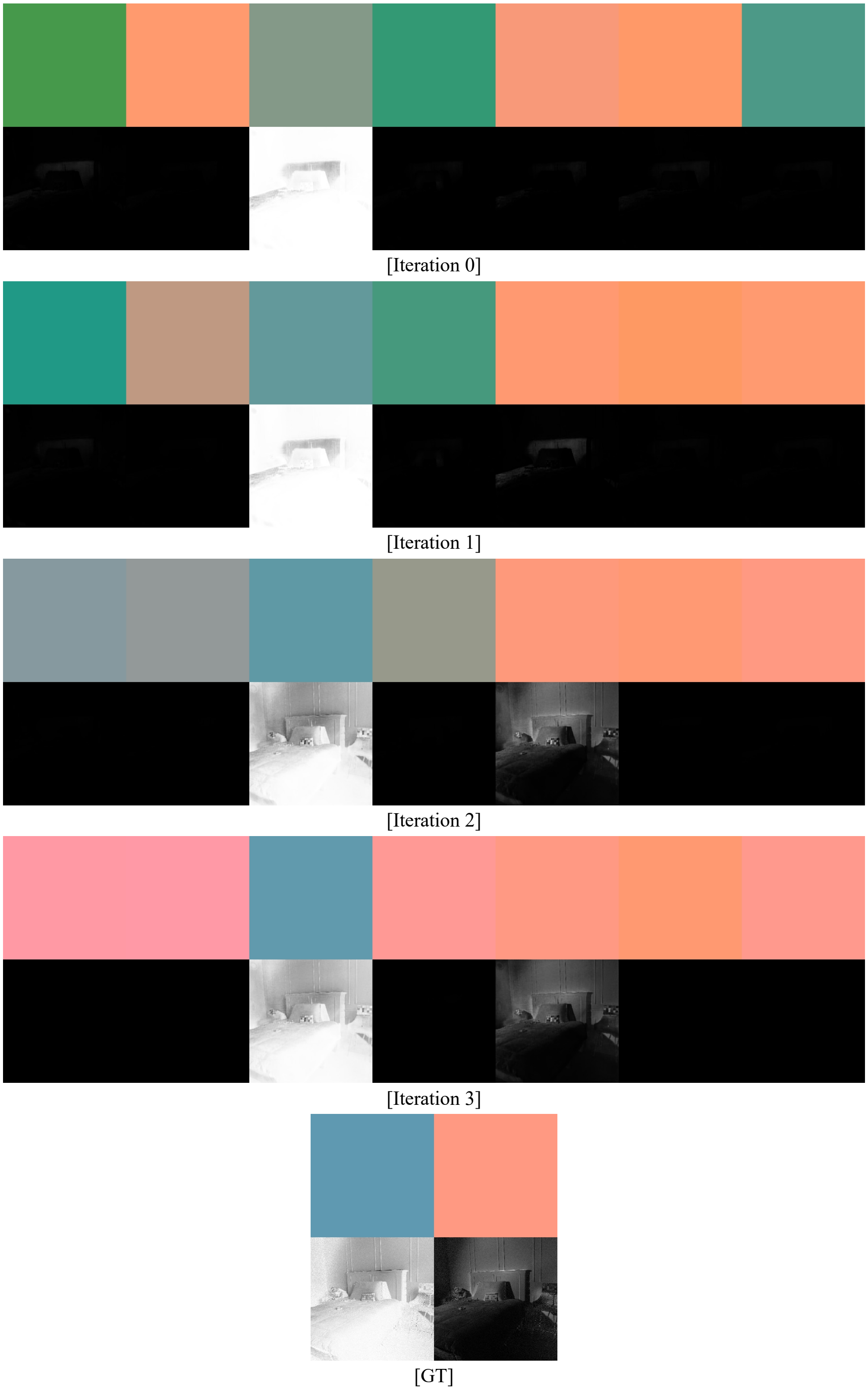


Figure 10. Slot evolution visualization with respect to iterative calibration process.

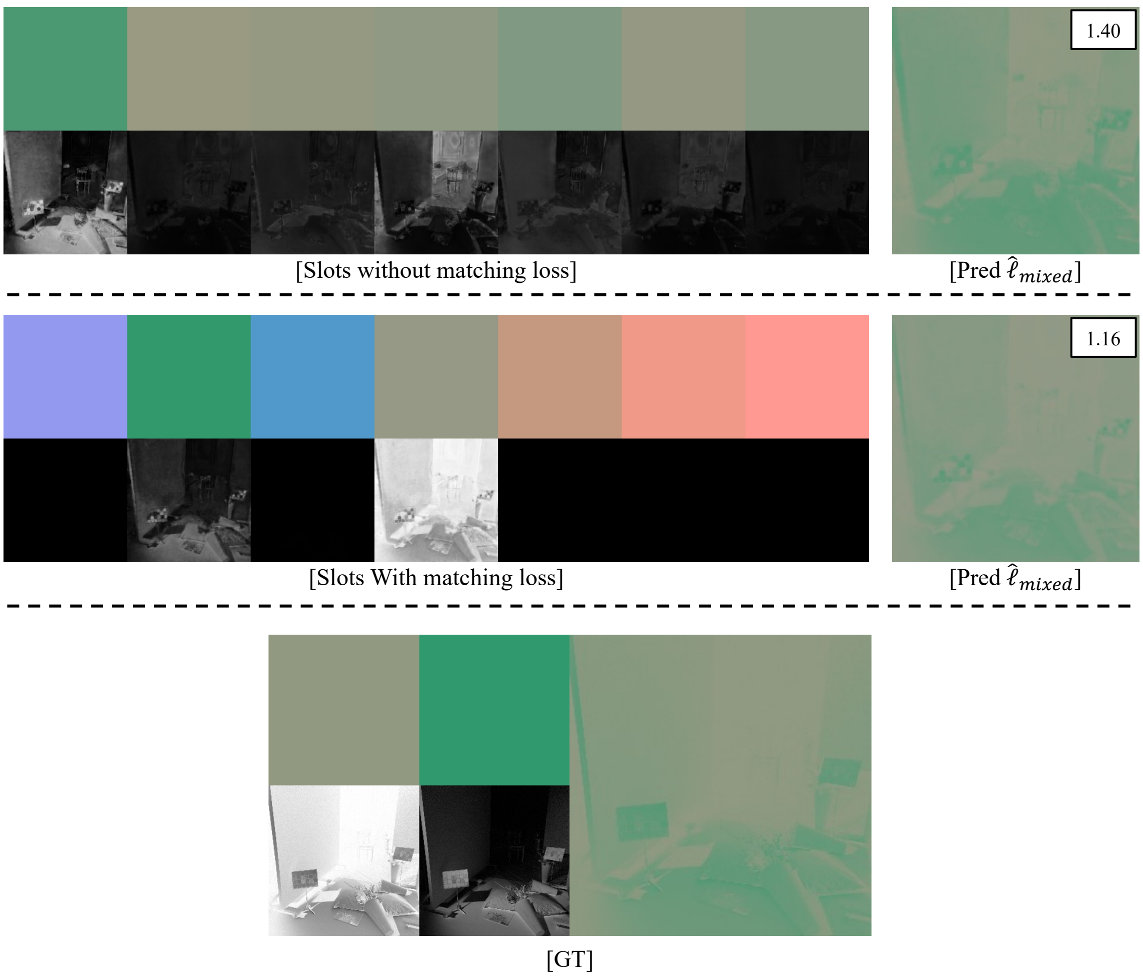
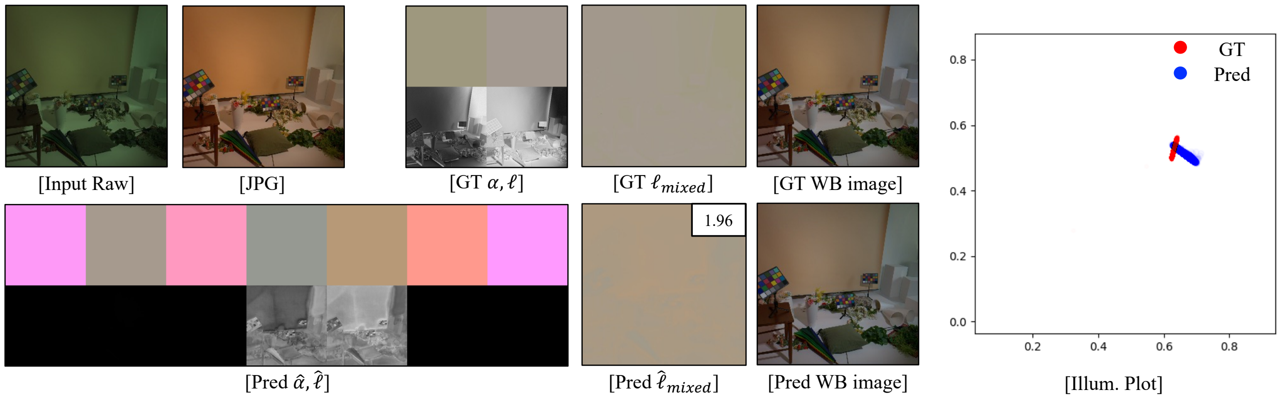
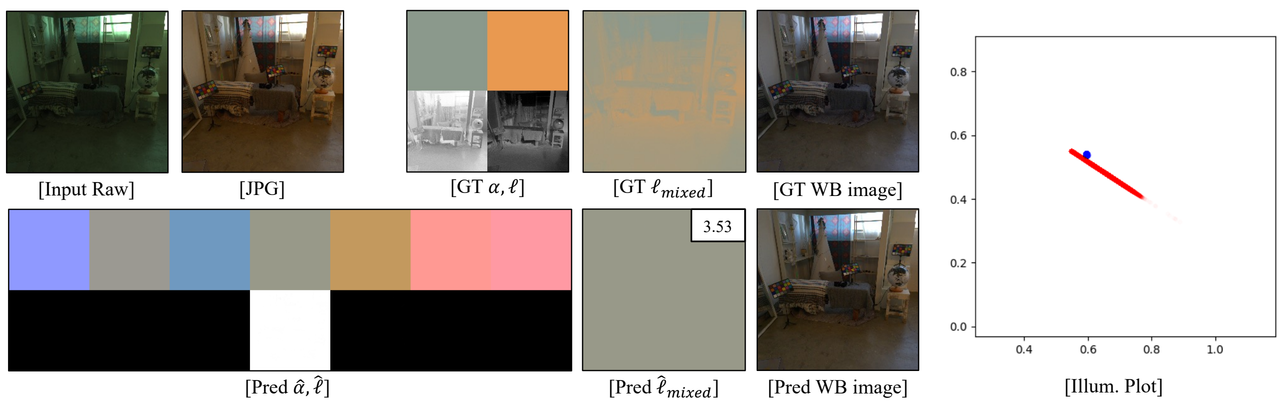


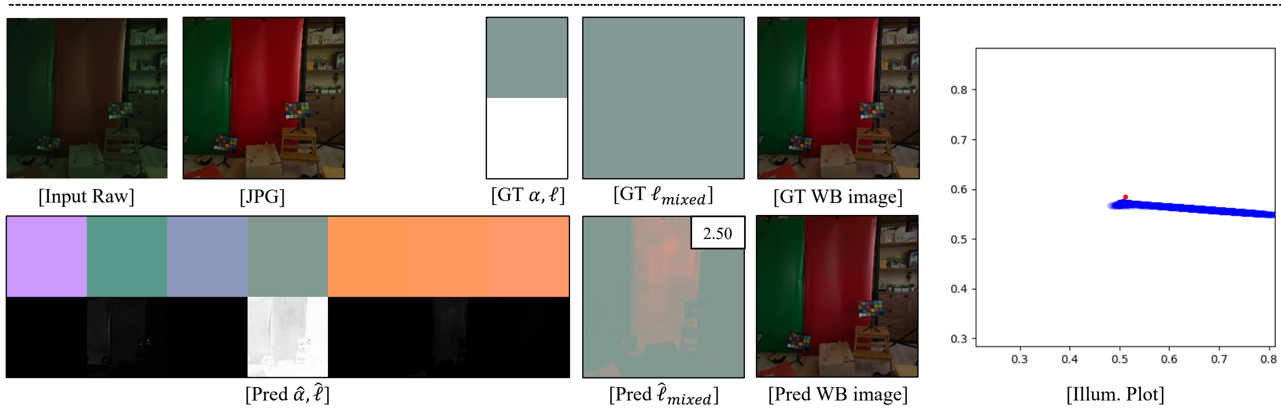
Figure 11. Without the matching loss $\mathcal{L}_{centroid}$, each slot evolves its own chromaticity to approximate the ground truth illumination, resulting in conflicts among the slots.



(a) Failure case 1



(b) Failure case 2



(c) Failure case 3

Figure 12. Failure cases of AID framework.