

Learning Camera-Agnostic White-Balance Preferences

Supplementary Material

Luxi Zhao Mahmoud Afifi Michael S. Brown

AI Center-Toronto, Samsung Electronics

{lucy.zhao, m.afifi1, michael.bl}@samsung.com

This supplementary material provides additional qualitative results, along with details about the dataset presented in the main paper.

1. Data

In the main paper, we mentioned that we use the in-camera AWB-estimated illuminant as the ground truth. Here, we provide further explanation for this choice. To ensure the aesthetic consistency of the ground-truth illuminants across all cameras, we extract the in-camera AWB estimated illuminant from the raw metadata as our ground truth. This decision is based on the observation that AWB algorithms from the same manufacturer are typically calibrated to produce a consistent visual style, especially when the cameras are part of the same smartphone or belong to consecutive smartphone generations. Two cameras from the same manufacturer capturing the same scene under identical illumination may yield different raw illuminant values due to differences in their sensor spectral responses and other camera-specific characteristics. However, once transformed into the sensor-agnostic CIE XYZ color space, these illuminants should converge to similar values. Significant discrepancies in this space would primarily reflect differences in the cameras’ AWB biases rather than sensor characteristics.

We demonstrate that the cameras in our dataset exhibit consistent AWB aesthetic styles by capturing a Macbeth color checker under 12 different illuminations, with correlated color temperatures ranging from 2000K to 7500K. We capture the color checker using all four sensors involved in our training and evaluation: S24U-W, S25U-W, S25U-T, and S25U-ST. As shown in Table 1, the three S25U cameras show relatively high angular errors with respect to S24U-W (the training sensor) when compared in raw space. However, after converting the illuminants to the CIE XYZ color space, the angular errors are significantly reduced. This supports our observation that the in-camera AWB outputs are aligned in aesthetic style once sensor-specific differences are removed.

Table 1. Average angular errors of S25U cameras against S24U-W in raw v.s. CIE XYZ, computed from a color chart captured under 12 illuminations ranging from 2000K to 7500K.

Color space	S25U-W	S25U-T	S25U-ST
Camera’s raw	2.57	2.51	6.13
CIE XYZ	0.66	0.24	0.53

Figure 1 further illustrates the consistency in aesthetic styles across the cameras. To remove the effects of differing fields of view between cameras, we present results using a color checker in Fig. 2. We visualize the angular errors of the color checker, between S24U-W and the three S25U cameras, under the D50 illuminant in both raw (no white balance applied) and CIE XYZ (with white-balance applied under the respective camera-estimated illuminant). Angular errors reduced significantly from raw to CIE XYZ.

The same trend can be observed from Fig. 3, another visualization of the same color chart used in Fig. 2. The left side of Fig. 3 shows the color values in raw RGB. This is a reproduction of Fig. 2 in the main paper, with the horizontal and vertical axes scaled to be the same as the right side of Fig. 3, which plots the color values when converted to CIE XYZ. The different colors in raw RGB converge to the same value in XYZ because the latter is a sensor-agnostic space, and the white balance algorithms onboard these smartphones are configured to produce the same aesthetic styles.

2. Qualitative Results

In the main paper, we presented qualitative results for our proposed mapping applied on cross-camera methods C5 [1] and gray world [2]. In this section, we present more results for our mapping based on C4 [4] in Fig. 4, and weighted gray-edge (wGE) [3] in Fig. 5. As shown in the figures, our results match more closely with the ground-truth’s aesthetic style than the cross-camera method’s estimated neutral illuminant.

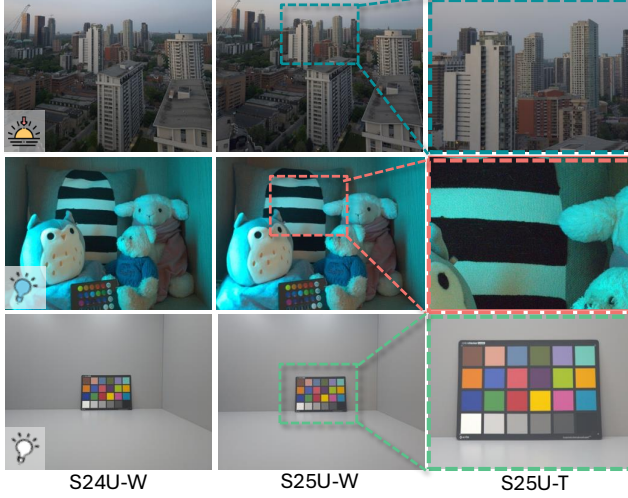


Figure 1. Scenes captured by different cameras from the same manufacturer (Samsung Galaxy series) often exhibit consistent white balance biases. The images are visualized by first white-balancing the raw using the respective camera-estimated illuminants, then converted to sRGB with no additional photofinishing.

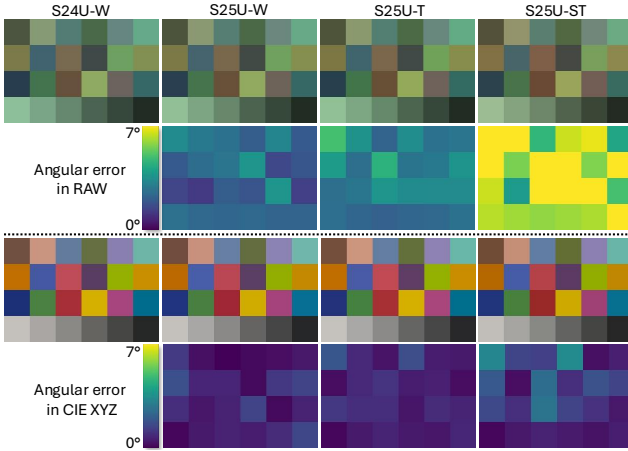


Figure 2. Top section: color chart captured by four cameras under the same lighting condition, shown in raw (gamma applied only to aid visualization). Angular error is computed between the three S25U cameras against S24U-W in raw. Bottom section: Angular error is computed between the three S25U cameras against S24U-W after white-balancing the color chart with the respective camera's estimated illuminant and converting to CIE XYZ (XYZ2sRGB is applied in the 3rd row, only to aid visualization).

References

- [1] Mahmoud Afifi, Jonathan T Barron, Chloe LeGendre, Yun-Ta Tsai, and Francois Bleibel. Cross-camera convolutional color constancy. In *ICCV*, 2021. 1
- [2] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1): 1–26, 1980. 1
- [3] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Improving color constancy by photometric edge weighting.

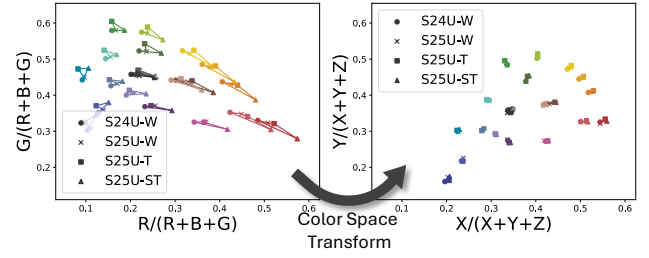


Figure 3. Each point in the figure represents a raw color from a Macbeth color checker. For the same object under identical illumination, different sensors produce different raw values due to variations in spectral response and other camera characteristics. After converting to CIE XYZ space, the illuminant values from different cameras converge to the same point.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(5):918–929, 2011. 1, 3

- [4] Huanglin Yu, Ke Chen, Kaiqi Wang, Yanlin Qian, Zhaoxiang Zhang, and Kui Jia. Cascading convolutional color constancy. In *AAAI*, 2020. 1, 3

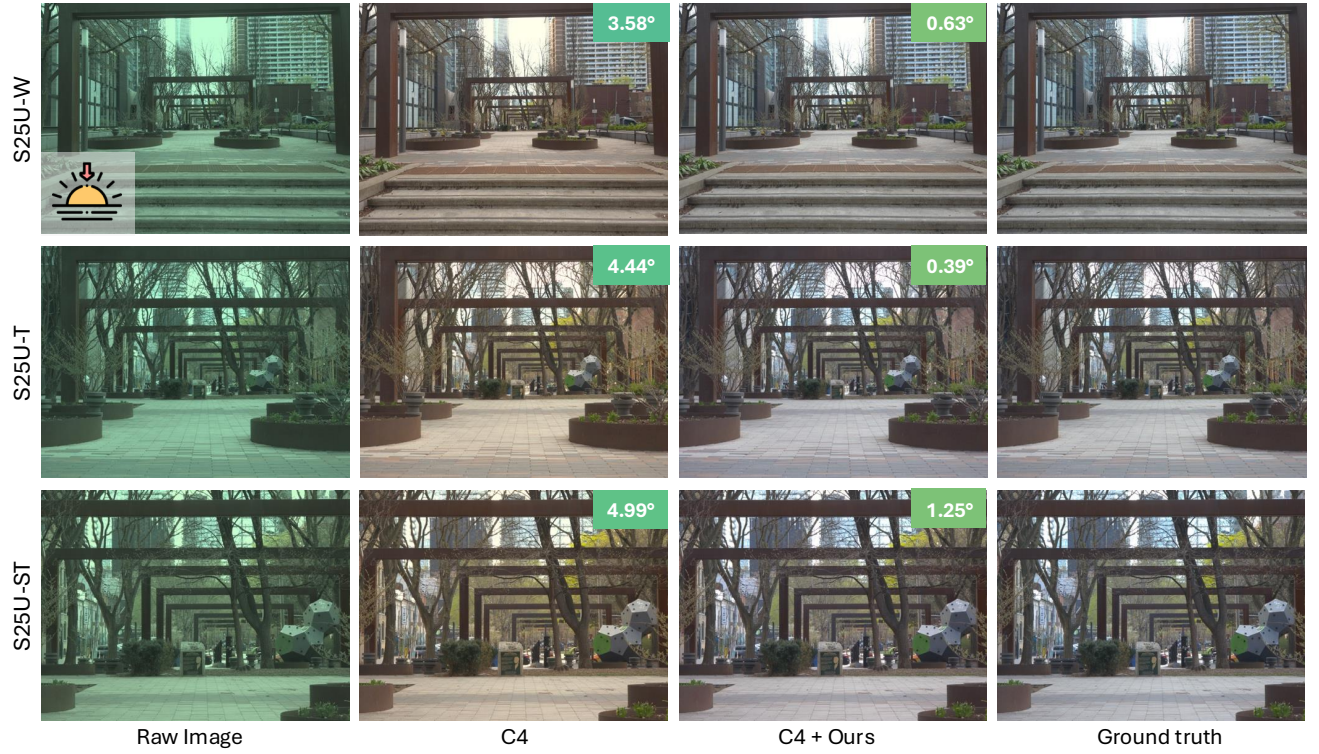


Figure 4. Qualitative results. Shown is a sunset scene captured by the S25U wide, telephoto, and super-telephoto sensors, white-balanced using C4 [4] predicted illuminant, C4 corrected with our mapping, and the ground truth aesthetically-preferred illuminant.



Figure 5. Qualitative results. Shown is an indoor scene illuminated by an orange LED light, captured by the S25U wide, telephoto, and super-telephoto sensors, white-balanced using wGE [3] predicted illuminant, wGE corrected with our mapping, and the ground truth aesthetically-preferred illuminant.