https://doi.org/10.2352/CIC.2025.33.1.18

Learning-Free Cross-Sensor Color Constancy Using Optimal Nonsingular Matrices

***Liangwei Chen*[1], *Minchen Wei*[1*], *Ming Ronnier Luo*[2]**
[1] ***Color Imaging and Illumination Laboratory, The Hong Kong Polytechnic University, Kowloon, Hong Kong***
[2] ***State Key Laboratory of Extreme Photonics and Instrumentation, Zhejiang University, Hangzhou, China.***
**** minchen.wei@polyu.edu.hk***

Abstract

Automatic white balance (AWB) plays a crucial role in digital imaging, with modern learning-based methods achieving better performance. These methods, however, require extensive training data captured by a specific sensor, which cannot be directly deployed other sensors due to the different spectral sensitivity functions. This paper presents a novel cross-sensor adaptation method based on 3×3 color transformation matrices. By leveraging least-squares optimization and a Mahalanobis distance strategy, our approach constructs sensor-specific mapping matrices using 24-patch ColorChecker data. The results derived using the NUS dataset demonstrate that the proposed method has much smaller angular errors without requiring additional data collection or complicated network tuning.

Introduction

In the field of image signal processor (ISP) pipeline, automatic white balance (AWB) plays an important role in reproducing consistent and accurate color appearance across various lighting conditions. The existing methods can be broadly categorized into two groups: statistics- and learning-based methods.

Statistics-based methods, such as the gray-World [1] and white patch methods [2], or other more sophisticated yet reliable methods (e.g., grey indexing [3] and PCA-based algorithms [4]) estimate the illuminant by analyzing the global or local statistical properties of the images. These methods are computationally efficient and do not require training data, but their performance is significantly affected by the scenes and are poor for many challenging scenes. In contrast, learning-based methods [5, 6] adopt a data-driven approach to model the complicated relationship between images and illuminants, allowing an easy adaptation to various scenes. Though these learning-based methods significantly outperform the statistics-based methods, the complexity of the network often hinders practical deployment. With the rapid advancement of neural networks, Liu et al. [8] recently introduced Kolmogorov-Arnold Networks (KANs), a novel structure that is fundamentally different from multi-layer perceptrons (MLPs) and demonstrates superior accuracy and interpretability. Chen et al.[9] adopted the KAN framework for estimating the illuminant of pure color images, leading to better performance than the MLP-based method (i.e., PCC method [7]).

Cross-sensor performance, however, remains a key challenge among the various learning-based AWB methods. This is mainly because the RGB signals captured by a camera are sigificantly affected by the spectral sensitivity functions of the camera sensor, which vary a lot. Figure 1 shows the spectral sensitivity functions

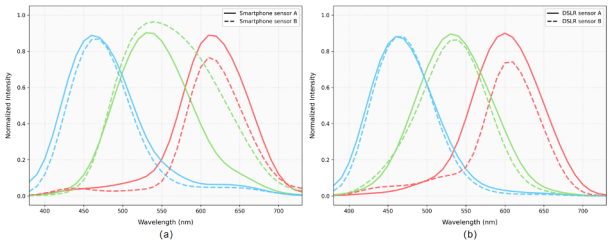

Figure 1. *Spectral sensitivity functions of camera sensors (a) two smartphone cameras; (b) two DSLR cameras.*

of two smartphone camera sensors and two DSLR camera sensors from different manufacturers. Due to the inherent characteristics of electronic devices, even two sensors of the exact same model may still have slight differences. A sensor S_s with the spectral sensitivity functions $\phi_s(\lambda)$ and a sensor S_t with $\phi_t(\lambda)$ will produce different signals Y_s and Y_t under a same scene.

$$Y_s = \int L(\lambda)\phi_s(\lambda)d\lambda, \quad Y_t = \int L(\lambda)\phi_t(\lambda)d\lambda \tag{1}$$

In such a case, if the model $f_s(Y)$ is trained using sensor S_s, it cannot be applied to S_s directly.

To address this issue, existing studies have proposed cross-sensor adaptation strategies, such as model-re-training-free (MRTF) methods that develop universal models or employ few-shot fine-tuning [10], and data-re-collection-free (DRCF) methods that focus on individual sensors [11]. However, these approaches typically require additional target-domain data or complex network adjustments. Recently, a dual-mapping strategy [12] was proposed, which adopts transformation matrices to reconstruct the images captured by sensors different from the training dataset. This approach effectively reduces sensor-induced domain shifts and provides an efficient solution for cross-sensor illumination estimation. Relying on D65 only may not be enough in practice.

In this study, we propose a novel cross-sensor color transformation method based on nonsingular 3×3 matrices derived from a 24-patch ColorChecker. It results in good performance without requiring extensive additional data collection. Furthermore, we systematically investigate: (1) Which illumination conditions produce the most effective transformation matrices when using ColorChecker measurements, and (2) Whether applying white balance before deriving the transformation matrix improves the results.

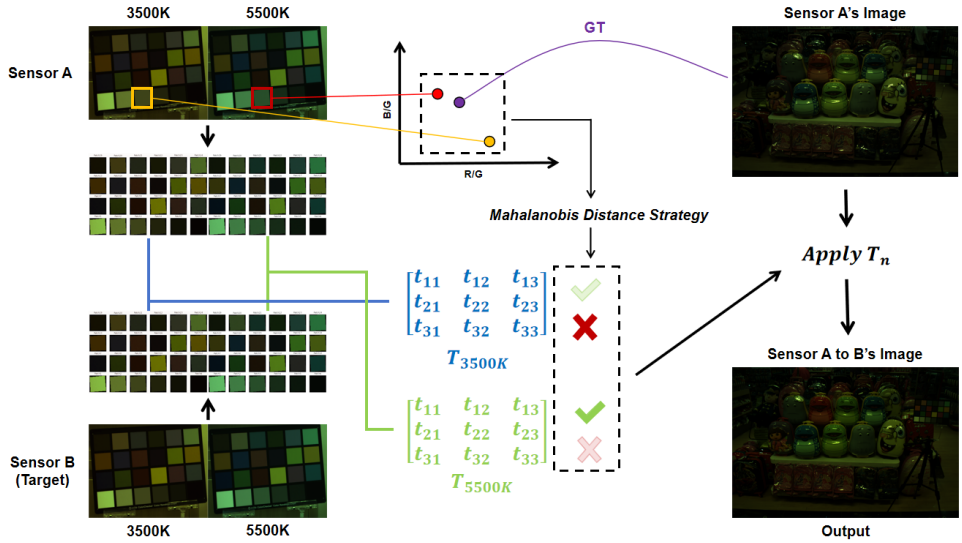

Figure 2. *Workflow of the proposed cross-sensor color transformation method. The process includes: (1) capturing ColorChecker images under different illuminations, (2) extracting and normalizing patch values, (3) calculating the transformation matrix T via a least squares method, and (4) applying T to map colors between sensors. The Mahalanobis distance selects the optimal matrix for an accurate adaptation.*

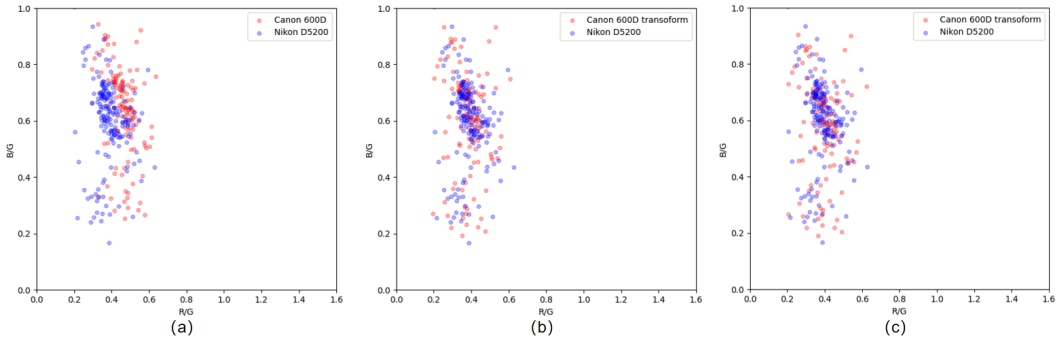

Figure 3. *Comparison of the image pixels of a same scene in the R/G-B/G spaces, with the images resampled and downsampled by 100 x for visualization. (a) Direct comparison between Canon 600D and Kikon D5200 images, (b) results using the proposed method, and (c) results of the baseline method.*

Proposed Method

The Least Squares Method

To transform the dataset captured by Sensor A to Sensor B's color space, the procedure begins by capturing the images of a ColorChecker under 3500 and 5500 K illumination conditions using the two sensors, with the RAW RGB values of the 24 patches $A \in \mathbb{R}^{3\times 24}$ and $B \in \mathbb{R}^{3\times 24}$ extracted. The constrained least squares minimization can then be used to derive the optimal transformation matrix $T \in \mathbb{R}^{3\times 3}$:

$$\min_{T} \|T\hat{A} - \hat{B}\|_F^2 \tag{2}$$

where $\hat{A} = A/(2^{16}-1)$ and $\hat{B} = B/(2^{16}-1)$ are the normalized input matrices, since the pipeline takes 16-bit images as the input. The closed-form solution is derived through the normal equation:

$$T = \hat{B}\hat{A}^T(\hat{A}\hat{A}^T)^{-1} \tag{3}$$

This formulation yields a full transformation matrix accounting for the cross-channel interactions:

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \tag{4}$$

The residual error is characterized using the Frobenius norm:

$$\mathscr{E} = \|T\hat{A} - \hat{B}\|_F = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{N} |(T\hat{A} - \hat{B})_{ij}|^2} \tag{5}$$

Mahalanobis Distance Strategy

Two precomputed transformation matrices, T_{3500K} and T_{5500K}, are derived via the least squares optimization under their respective correlated color temperatures (CCTs). The 21st patch on the ColorChecker was used as the neutral white for avoiding saturation [4], with the RGB values (R_w, G_w, B_w) transformed into

chromaticities; the ground truth illuminant (R_μ, G_μ, B_μ) are also transformed into chromaticities:

$$(r,b) = \left(\frac{R}{G}, \frac{B}{G}\right), \tag{6}$$

For selecting the matrix, we used a **weighted Mahalanobis distance metric** in the chromaticity space, accounting for the non-uniform sensitivity of the human visusal system to chromatic shifts. For a pair of chromaticity vectors $\mathbf{c}_1 = (r_1, b_1)$ and $\mathbf{c}_2 = (r_2, b_2)$, the distance metric is defined as:

$$D(\mathbf{c}_1, \mathbf{c}_2) = \sqrt{(\mathbf{c}_1 - \mathbf{c}_2)^\top \mathbf{\Sigma}^{-1} (\mathbf{c}_1 - \mathbf{c}_2)} \tag{7}$$

where $\mathbf{\Sigma}$ is the covariance matrix empirically estimated from the training set to weight chromaticity deviations:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_r^2 & \rho\sigma_r\sigma_b \\ \rho\sigma_r\sigma_b & \sigma_b^2 \end{bmatrix} \tag{8}$$

Here, σ_r^2 and σ_b^2 are the variances of r and b in the dataset, and ρ is the correlation coefficient. Then, the transformation matrix is selected:

$$\mathbf{T}^* = \underset{k \in \{3500K, 5500K\}}{\arg\min} \mathscr{D}(\mathbf{c}_w, \mathbf{c}_k) \tag{9}$$

The selected matrix is applied to the linearized image $\mathbf{I}$: $\mathbf{I}_{\text{transformed}} = \mathbf{I} \cdot \mathbf{T}^\top$.

Figure 2 illustrates the workflow of the proposed method, with Figure 3 showing examples of transforming the NUS-Canon600D dataset to the NikonD5200 color space using the proposed method in the r-b space and the visulaization are shown with 100 times downsampled images.

Experiment and Results

Angular Error

The angular error was adopted as the metric for characterizing the performance. Given the ground truth illuminant $\boldsymbol{\ell} = [\ell_R, \ell_G, \ell_B]^\top$ and predicted illuminant $\hat{\boldsymbol{\ell}} = [\hat{\ell}_R, \hat{\ell}_G, \hat{\ell}_B]^\top$, the angular error $\theta \in [0, \pi]$ can be calculated as:

$$\theta = \cos^{-1}\left(\frac{\langle \boldsymbol{\ell}, \hat{\boldsymbol{\ell}} \rangle}{\|\boldsymbol{\ell}\|_2 \cdot \|\hat{\boldsymbol{\ell}}\|_2}\right) \tag{10}$$

An additional ℓ_2 regularization is added into the loss function:

$$\mathscr{L}(\theta; \mathbf{W}) = \underbrace{\frac{1}{N}\sum_{i=1}^{N} \theta_i}_{\text{Angular term}} + \underbrace{\lambda \|\mathbf{W}\|_2}_{\text{Regularization}} \tag{11}$$

where N is the batch size, $\mathbf{W}$ is the trainable parameters to avoid overfitting, and the regularization strength is set to $\lambda = 10^{-5}$.

Table 1. Summary of images in the NUS dataset

NUS Cameras	Real-world images No.	Laboratory images No.
Canon EOS-1Ds Mark III	259	105
Canon EOS 600D	201	105
Fujifilm X-M1	196	105
Nikon D5200	200	105
Olympus E-PL6	208	105
Panasonic Lumix DMC-GX1	203	105
Samsung NX2000	202	105
Sony SLT-A57	268	105
Nikon D40	117	104

Dataset

The experiment was performed on the NUS dataset [4], which is widely used in various studies related to cross-sensor issues. The dataset includes images captured by nine cameras for a wide range of indoor and outdoor scenes, with the number of the scenes captured by each camera summarized in Table 1. We carefully selected the pairs of the images for the same scene, and identified 176 image pairs captured by Canon 600D and Nikon D5200.

Model Testing

Three models were considered in this study. The Cross-Camera Convolutional Color Constancy (C5) model [5] pioneers the cross-sensor adaptation through a hypernetwork framework, with the dual-path U-Net dynamically generating histogram filters conditioned on auxiliary test images for addressing the different spectral sensitivity functions across cameras. The Pure Color Constancy (PCC) model [7] adopts a five-layer MLP to efficiently estimate the illuminant using four chromaticitiy features (i.e., max/mean/brightest/darkest pixels). The Large Size Color Constancy (LSCC) [9] adopts the Kolmogorov-Arnold Networks (KANs) based on the PCC model, replacing the fixed activation functions with learnable spline functions for enhancing interpretability and reducing the number of parameters.

Results

Given the 176 pairs of images captured by Canon 600D (denoted as $\mathscr{C}ori$) and Nikon D5200 ($\mathscr{N}ori$), a total of **176 transformation matrices** $\{\mathbf{T}_i\}_{i=1}^{176}$ were derived to map C_{ori} to Nikon's color space, which was labeled as the baseline dataset $C2N_{\text{baseline}}$. The ground truth $\mathscr{C}_{ori}$ was projected into the r-b space. For illuminant estimation, we analyzed the gray card data under the 16 controlled conditions with different CCTs (i.e., 2800,3000,3500 ... 10000 K) and selected five representative matrices $\{\mathbf{T}_{m,n}^{(k)}\}_{k=1}^{5}$ corresponding to the five illuminants (i.e., 3000, 3500, 4800, 5500, 6000 K). With these, the optimized datasets were derived as:

$$C2N_{\text{optm}} = \{\mathbf{T}_{m,n}^{(k)} \cdot \mathbf{Y}_i \mid \mathbf{Y}_i \in \mathscr{C}_{ori}\}, \quad m \in \{5,3,2,1\} \tag{12}$$

where m denotes the number of the representative matrices used, with $C2N_{\text{opti3}}$ using the 6000, 4800, and 3000 K matrices, $C2N_{\text{opti2}}$ using the 5500 and 3500 K matrices, and $C2N_{\text{opti1}}$ using the 5500 K matrix.

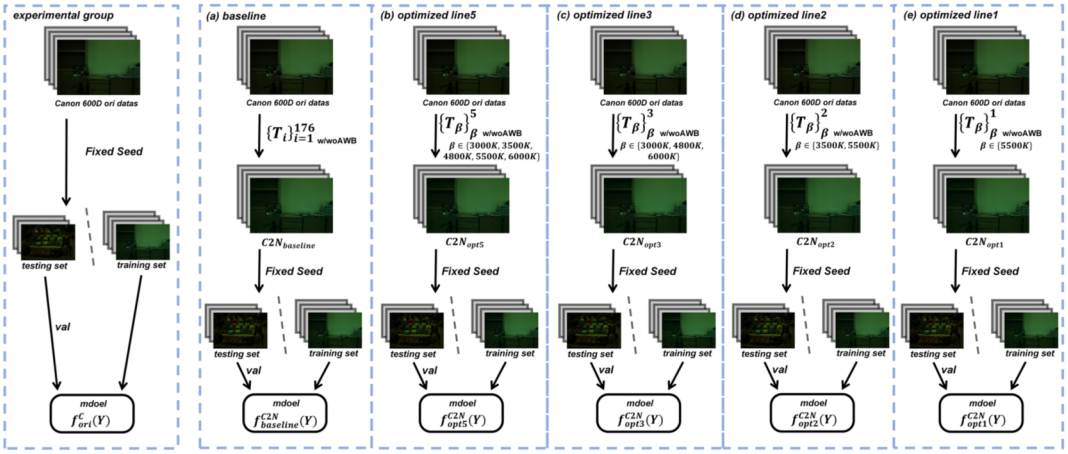

Figure 4. *Flowchart of the experiment. The experimental group (serving as control) was trained directly on the $\mathscr{C}ori$ training set. (a) Baseline group: images were first transformed using their respective sensor-specific T matrices, then partitioned into training sets for model training. (b)-(e) Experimental groups: images were transformed using 5, 3, 2, and 1 "representative matrices" respectively, then partitioned into training sets for model training.*

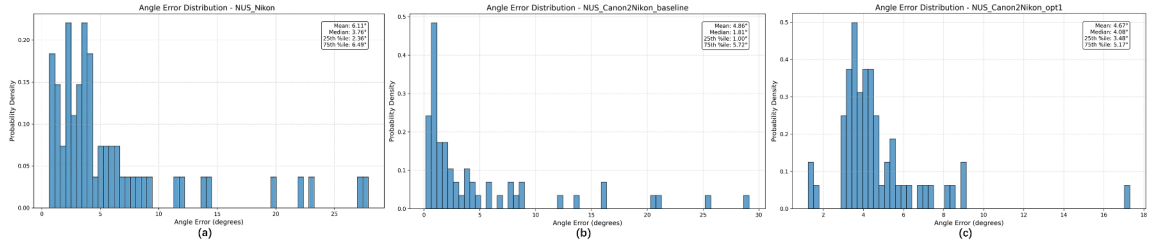

Figure 5. *Summary of the angular error distributions for applying the C5 model on the dataset without AWB. (a) Direct apply $f^{C}_{\text{ori}}(Y)$ on $\mathscr{N}ori$; (b) baseline results applying $f^{C2N}_{\text{baseline}}(Y)$ on $\mathscr{N}ori$; (c) applying $f^{C2N}_{\text{opt1}}(Y)$ on $\mathscr{N}ori$.*

Figure 4 shows the experiment workflow. We trained two baseline models: $f^{C}_{\text{ori}}(Y)$ and $f^{N}_{\text{ori}}(Y)$ on $\mathscr{C}_{\text{ori}}$ and $\mathscr{N}_{\text{ori}}$, respectively. For the transformed datasets, the models were trained as $f^{C}_{\text{method}}(Y), \text{method} \in \{\text{base}, \text{opt5}, \text{opt3}, \text{opt2}, \text{opt1}\}$. During the training process, we fixed the random seed to ensure a consistent split between the training and testing sets, and used the testing set in $\mathscr{N}_{\text{ori}}$ to evaluate all the models, with Table 2 summarizing the angular errors. To investigate the impact of white balance, the experiment was also carried out by performing the white balance on each image using the ground truth illuminant, with the results summarized in Table 3. Table 4 summarizes the training and testing sets used in the experiments.

In comparison to directly apply the model to a different sensor, $f^{C2N}_{\text{baseline}}(Y)$ was able to reduce the angular errors by 20-30%. Among the various configurations, $f^{C2N}_{\text{opt1}}(Y)$ or $f^{C2N}_{\text{opt2}}(Y)$ outperformed others, suggesting that excessive matrix diversification may not necessarily improve the performance. Regarding the effect of AWB, the matrices derived from the AWB-applied data generally resulted in smaller angular errors, especially for the LSCC method. Figure 5 shows the angular errors derived from the C5 model (without AWB) as an example, suggesting the the representative matrix was effectively in reducing the angular errors, especially to the images with larger angular errors.

Conclusion

We propose a lightweight adaptation framework based on full-rank color transformation matrices to address the weaknesses of learning-based AWB methods in cross-sensor applications. The proposed method, combining with the least squares optimization with Mahalanobis distance-based matrix selection, was found effective in reducing the angular errors caused by the difference of the sensor spectral sensitivity functions, based on the NUS dataset. It enhances the efficiency when applying learning-based AWB methods to different sensors.

References

[1] Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1), 1-26.

[2] Land, E. H. (1977). The retinex theory of color vision. *Scientific American*, 237(6), 108-129.

[3] Qian, Y., Kamarainen, J.K., Nikkanen, J. and Matas, J., 2019. On finding gray pixels. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8062-8070).

[4] Cheng, D., Prasad, D.K. and Brown, M.S., 2014. Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution. *JOSA A*, 31(5), pp.1049-1058.

[5] Afifi, M., Barron, J. T., LeGendre, C., Tsai, Y. T., & Bleibel, F. (2021). Cross-camera convolutional color constancy. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp.1981-1990).

Table 2. Summary of the angular errors on Nikon D5200 sensor, when the different models were trained in different ways as summarized in 4 without performing AWB.

	C5							PCC							LSCC						
	Mean	Median	Time an	Best 25%	Worst 25%	Worst 5%	Best	Mean	Median	Time an	Best 25%	Worst 25%	Worst 5%	Best	Mean	Median	Time an	Best 25%	Worst 25%	Worst 5%	Best
(a)	2.89	1.79	1.94	0.69	6.08	14.39	0.38	2.49	1.44	1.67	0.40	6.40	13.73	0.09	2.49	1.47	1.69	0.51	6.15	12.55	0.20
(b)	3.27	1.85	2.15	0.54	8.25	19.81	0.10	3.05	2.29	2.41	0.80	6.84	14.21	0.13	2.92	1.45	1.92	0.57	6.30	14.29	0.08
(c)	6.11	3.76	4.09	1.43	14.66	27.45	0.66	5.53	3.21	3.67	2.11	12.90	24.69	1.42	5.62	3.54	3.88	2.21	12.86	24.86	1.12
(d)	4.86	1.81	2.59	0.60	13.89	27.29	0.18	4.65	2.18	2.66	0.67	12.69	25.11	0.24	4.74	2.05	2.60	0.58	12.95	24.98	0.25
(e)	4.87	4.15	4.21	2.84	8.20	15.54	1.84	4.72	4.05	4.08	2.69	8.02	15.50	0.98	4.98	4.20	4.33	2.86	8.39	14.95	1.96
(f)	5.21	4.16	4.30	3.14	9.01	18.34	1.53	5.13	4.20	4.40	3.02	8.77	**14.16**	**0.54**	5.06	4.29	4.31	2.96	8.52	14.92	1.78
(g)	5.34	4.41	4.60	2.92	9.47	16.54	1.59	**4.23**	**3.30**	**3.50**	**2.20**	**7.76**	14.45	0.60	**4.07**	**3.28**	**3.41**	**2.95**	7.55	**14.05**	1.00
(h)	**4.67**	**4.08**	**4.20**	**2.83**	**7.51**	**13.15**	**1.25**	4.87	4.09	4.18	2.80	8.34	14.63	1.22	4.51	3.70	3.84	2.87	**7.48**	14.07	**0.99**

Table 3. Summary of the angular errors on Nikon D5200 sensor, when the different models were trained in different ways as summarized in 4 with AWB performed.

	C5							PCC							LSCC						
	Mean	Median	Time an	Best 25%	Worst 25%	Worst 5%	Best	Mean	Median	Time an	Best 25%	Worst 25%	Worst 5%	Best	Mean	Median	Time an	Best 25%	Worst 25%	Worst 5%	Best
(a)	2.69	1.79	1.94	0.69	6.08	14.39	0.38	2.49	1.44	1.57	0.40	6.40	13.73	0.09	2.49	1.47	1.59	0.51	6.15	12.55	0.20
(b)	3.27	1.85	2.15	0.54	8.25	19.81	0.10	3.05	2.29	2.41	0.80	6.84	14.21	0.13	2.92	1.45	1.92	0.57	6.30	14.29	0.08
(c)	6.11	3.76	4.09	1.43	14.66	27.45	0.66	5.53	3.21	3.67	2.11	12.90	24.69	1.42	5.62	3.54	3.88	2.21	12.86	24.86	1.12
(d)	4.78	4.17	4.29	2.28	8.41	17.01	0.56	4.04	2.77	3.11	0.95	9.24	17.31	0.42	3.96	2.59	2.96	0.90	9.40	17.02	0.20
(e)	6.93	6.27	6.47	2.96	12.00	20.91	0.83	4.13	**3.06**	3.16	0.89	**9.15**	**16.86**	0.37	4.04	2.93	**3.14**	0.99	9.24	16.48	0.52
(f)	6.38	5.64	5.87	2.09	11.70	14.85	0.53	4.58	3.14	3.51	1.46	9.76	18.01	0.39	4.06	3.16	3.23	1.05	**8.79**	**16.41**	**0.13**
(g)	**4.48**	**3.43**	**3.44**	**1.12**	9.75	25.74	**0.41**	**4.08**	3.18	**3.40**	**0.79**	9.36	16.98	**0.08**	**3.96**	**2.49**	**2.79**	**0.93**	9.04	16.58	0.38
(h)	4.49	3.49	3.78	1.93	**8.65**	**14.46**	1.00	4.42	3.21	3.42	0.89	10.11	18.10	0.17	4.25	2.68	3.23	0.95	9.92	17.52	0.48

Table 4. Summary of the training and testing sets used for Tables 2 and 3

Group	Training data	Testing data
(a)	NUS_Canon_train (117)	NUS_Canon_test (59)
(b)	NUS_Nikon_train (117)	NUS_Nikon_test (59)
(c)	NUS_Canon_train (117)	NUS_Nikon_test (59)
(d)	NUS_Canon2Nikon_train (117) (using all matrix)	NUS_Nikon_test (59)
(e)	NUS_Canon2Nikon_opt5 (117) (using 5 matrix)	NUS_Nikon_test (59)
(f)	NUS_Canon2Nikon_opt3 (117) (using 3 matrix)	NUS_Nikon_test (59)
(g)	NUS_Canon2Nikon_opt2 (117) (using 2 matrix)	NUS_Nikon_test (59)
(h)	NUS_Canon2Nikon_opt1 (117) (using 1 matrix)	NUS_Nikon_test (59)

[6] Y. Hu, B. Wang, and S. Lin, "FC4: fully convolutional color constancy with confidence-weighted pooling," in *IEEE Conference on Computer Vision and Pattern Recognition (2017)*, pp. 330–339.

[7] Yue, S., & Wei, M. (2023). Color constancy from a pure color view. *JOSA A*, 40(3), 602-610.

[8] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., & Tegmark, M. (2024). Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*.

[9] Chen, L., Luo, M.R. and Wei, M., 2024, October. Large Size of Color Constancy: Enhancing Pure Color Image Illuminant Estimation with Kolmogorov-Arnold Networks. In *Color and Imaging Conference* (Vol. 32, pp. 95-100). Society for Imaging Science and Technology.

[10] S. Bianco and C. Cusano, "Quasi-unsupervised color constancy," in *IEEE Conference on Computer Vision and Pattern Recognition* (2019), pp. 12212–12221.

[11] D. Hernandez-Juarez, S. Parisot, B. Busam, et al., "A multihypothesis approach to color constancy," in *IEEE Conference on Computer Vision and Pattern Recognition* (2020), pp. 2270–2280.

[12] Yue, S. and Wei, M., 2024. Effective cross-sensor color constancy using a dual-mapping strategy. *JOSA A*, 41(2), pp.329-337.

Author Biography

Liangwei Chen is a Master student at Zhejiang University, and he was an exchange student at the Hong Kong Polytechnic University from February to August in 2025. His work mainly focuses on color constancy and image transform.

Minchen Wei is a professor at the Hong Kong Polytechnic University. His research mainly focus on color science and imaging science.

Ming Ronnier Luo is a professor at Zhejiang University. His research mainly focuses on color science and imaging science.