https://doi.org/10.2352/CIC.2025.33.1.16

Integration of RGB Image and ALS for Color Constancy

Ruikai He, Minchen Wei*, The Hong Kong Polytechnic University
minchen.wei@polyu.edu.hk

Abstract

Illuminant estimation plays a key role in computational color constancy, which typically relies on RGB images alone. In recent years, smartphones are commonly equipped with an ambient light sensor (ALS) next to the rear camera, which records low-resolution spectral information. This paper proposes several methods to estimate the illuminant using both RGB images and ALS signals, which are based on the state-of-the-art learning-based methods without significantly increasing computational cost. The proposed methods are found to result in better performance, in comparison to those only rely on RGB images.

Introduction

Computational color constancy aims to adjust the colors of images by removing the color cast of the illumination, simulating the relative color constancy perceived by the human beings due to the chromatic adaptation mechanism in the human visual system [1]. It requires accurate estimation of the illuminant, which remains a challenge especially for many corner cases [2], [3], [4]. Many statistics- and learning-based methods have been developed to estimate the illuminant based on the RGB images [5], [6], and some methods also use the image's metadata [7], [8]. These methods effectively extract the relevant features from RGB images. For example, FC4 [9] uses the backbone similar to AlexNet [10], FFCC [11] and C5 [4] was inspired by CCC [12], and lightweight PCC [6] that was specifically designed for pure color image dataset. Recently, methods have been proposed to integrate additional metadata. For example, CCMNet encodes predefined illuminant colors using color correction matrices (CCMs) alongside images into several uv-histograms, enabling the network to adopt the unseen cameras [8]. Apart from CCMs, contextual metadata (i.e., timestamp and geolocation) has also been considered as the additional input to a network, which can help to reduce the possible range of illuminant colors [7].

In recent years, smartphones have been equipped with ambient light sensors (ALS) to capture the light level and color of the ambient light, which are commonly used to adjust color of the display (e.g., True Tone display) [13]. When an ALS has multiple channels, the information it captures can provide some useful information about the spectral composition of the ambient light. Ilaria et al attempted to concatenate the multispectral information with fully connected image data to estimate the illuminant [2], but it was performed based on simulation [14].

In this study, we aim to investigate whether the ALS signals can be useful in illuminant estimation, and what is the effective way to use it for illuminant estimation. Instead of proposing new methods, PCC, UCC, C5 are fine tuned to take ALS signals as additional input.

The investigations were performed on NUS-8 dataset [15] and also PolyU Pure Color dataset [6].We also collected a new dataset with both RAW images and ALS signals captured at the same time.

Building upon PCC [6], C5 [4], and RACC for color constancy (U-Net [16] for color constancy), we proposed new frameworks via fine-tuning, which allows for the fusion of ALS information and RGB images without significantly increasing computational cost. This approach facilitates our exploration of how ALS information can be leveraged to develop more effective and practical solutions for illuminant estimation.

For the existing datasets, whether consisting of RGB [15], [17], [18] or multispectral images [14], none contains ALS information necessary for our study. we collected new datasets using Xiaomi15Ultra to address this issue, and further details regarding the new datasets can be found in Experiment and Results.

Proposed fine-tuning methods

Fine tuning PCC is relatively straightforward, since it is a multilayer perceptron network. Since both ALS signals and RGB images are used as the input, the number of the neurons need to be increased, with the two vectors concatenated. The numbers of the hidden layers and neurons in each layer are correspondingly adjusted.

The tuning of RACC and C5 follows a similar approach, as the two networks share similar structures. The ALS signals are integrated after the image latent feature getting downsampled. For C5, a two-step deconvoluntion is employed to decode the ALS signals into the same shape as the encoded latent image feature, with a matrix addition to integrate the two. For RACC, a two-step deconvolution is used to decode the ALS signals, which is then concatenated with the encoded latent image feature.

Such fine-tuning allows flexible adjustments of the inputs without significantly increasing the computational costs.

Experiment and result

We used a smartphone for data collection, using the main camera to capture RAW images and the integrated 13-channel ALS to capture the ALS signals. For each scene, an image with a gray card placed in the image was used to capture the ground-truth illuminant, and an image without the gray card was used in the experiment. The data collection was performed in three stages, with 486 pure color images and ALS signals collected in Stage 1, 210 normal scene images and ALS signals collected in Stage 2, and a total of 8,484 images and ALS signals collected in Stage 3. All the images were down-sampled to a resolution of 64 × 48.

We performed the training on the three datasets, with a three-folded cross validation using the three fine-tuned networks, and calculated the angular error between the ground-truth and estimated illuminants. For the fine-tuned PCC, we used three sets of input, including only RGB images (ORGB), only ALS signals (OALS), and RGB images and ALS signals (DUAL). For the fine-tuned RACC and C5, only ORGB and DUAL were used as the input. Figure 3 summarizes the statistics of the angular errors. It is can be observed that the integration of the ALS signals can generally reduce the angular errors in comparison to only using the RGB images as the input, especially to the worst 5% cases, though the performance varies with the methods and datasets. For the PCC method, it is obvious that only using the ALS signals resulted in smaller angular errors than only using the RGB images, especially to the pure color images. Such results clearly suggest the effectiveness of using ALS signals in illuminant estimation.

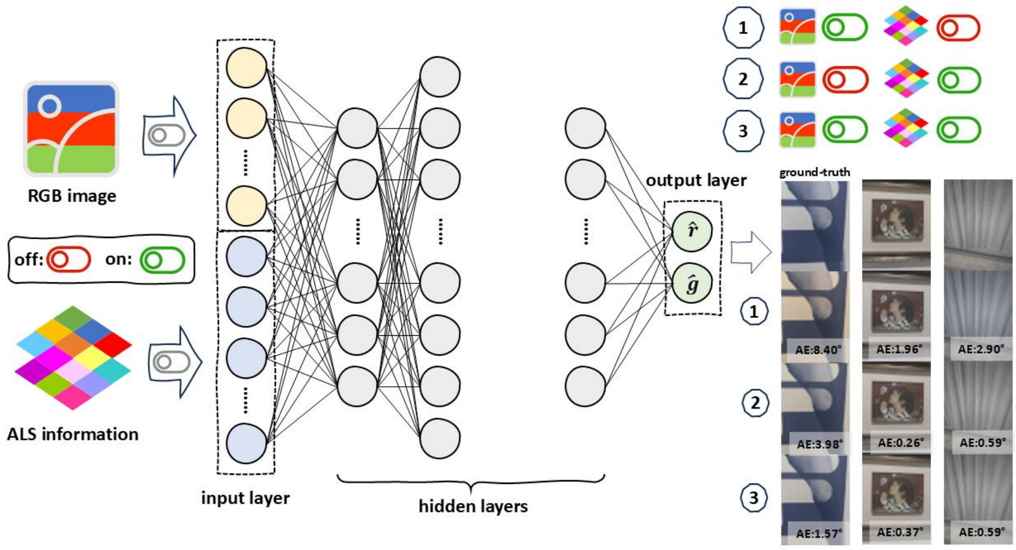


Fig. 1 Architecture of the fine-tuned PCC method.

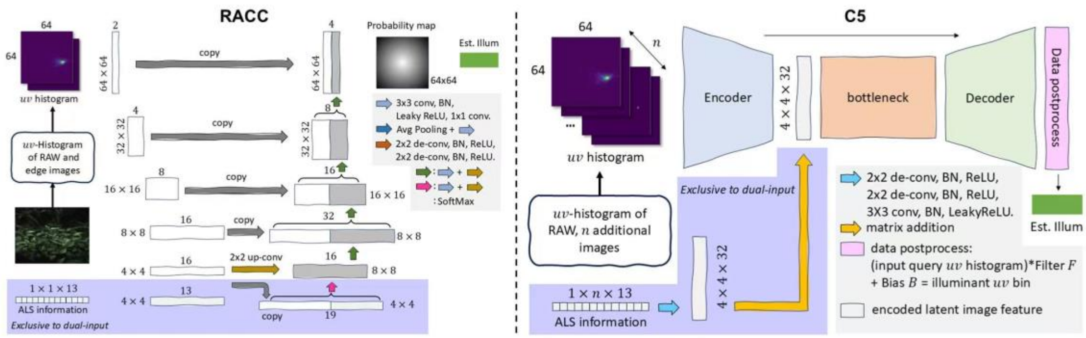
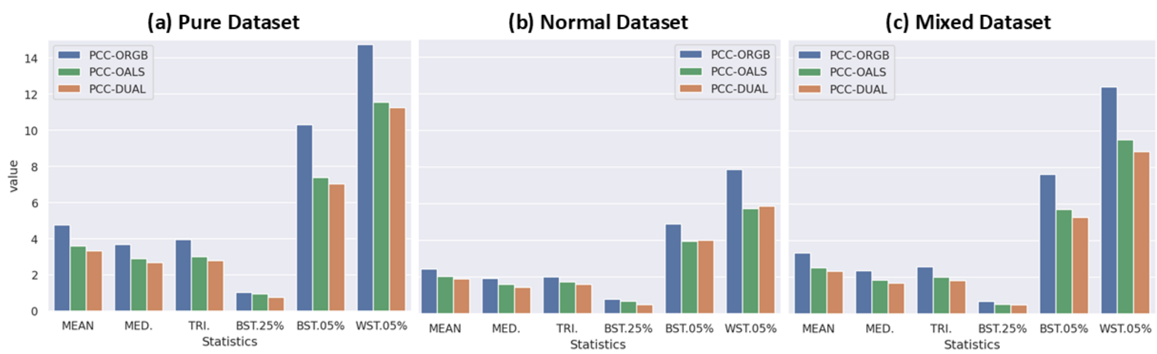


Fig. 2 Architecture of the fine-tuned RACC and C methods.

Ablation studies were then conducted to evaluate the contributions of the different inputs. For the methods trained using both RGB images and ALS signals, we evaluated the contributions by setting one of the two inputs into zero, and then compared the results to DUAL, as summarized in Table 1. We can find that setting one input to zero significantly increased the angular errors, and we did not find consistent results in terms of the importance of the two inputs. Figure 4 shows the subjective results of the various methods.



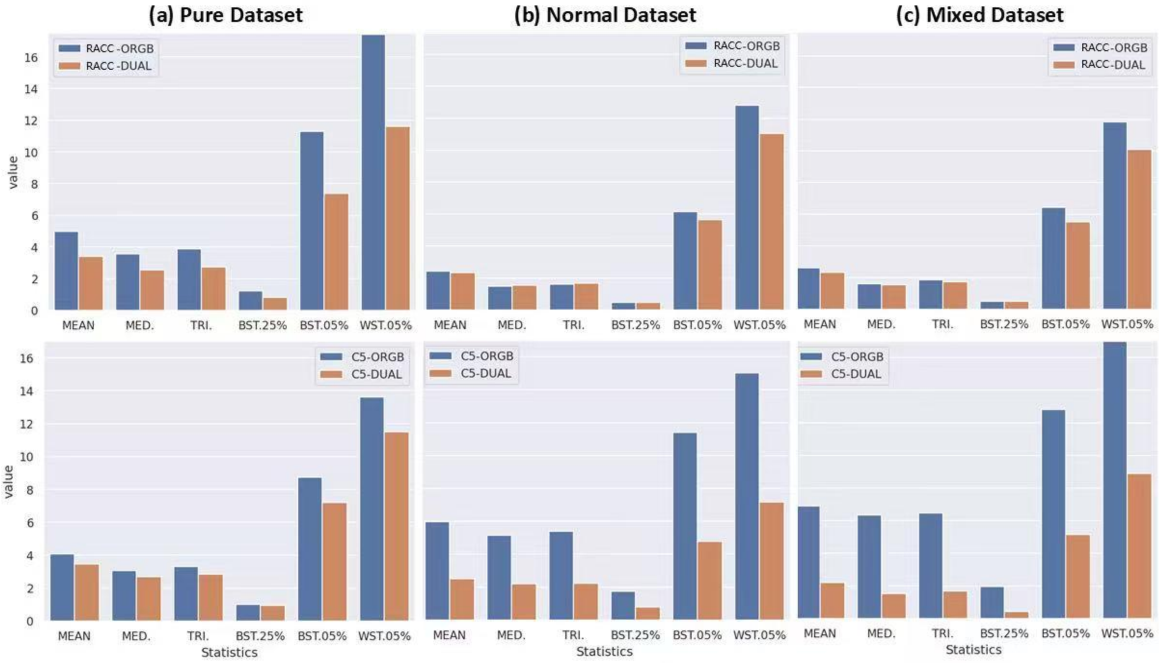


Fig. 3 Summary of the angular errors of the fine-tuned methods on the three datasets.

Table 1 Summary of the angular errors of the fine-tuned methods on the three datasets for the ablation study.

	MEAN.			MED.			BST.25%			WST.25%		
	ucc	C5	PCC	ucc	C5	PCC	ucc	C5	PCC	ucc	C5	PCC
Pure Dataset												
DUAL	3.41	3.45	3.33	2.54	2.69	2.69	0.81	7.20	0.76	7.37	11.49	7.04
ALS set to 0	7.52	5.87	8.69	4.88	4.91	6.32	1.34	1.37	1.53	18.01	11.58	18.09
RGB set to 0	5.76	20.06	4.93	4.41	20.69	4.24	1.47	15.73	1.30	12.69	25.37	9.72
Normal Dataset												
DUAL	2.34	2.52	1.90	1.53	2.20	1.43	0.47	0.80	0.49	5.64	4.77	4.00
ALS set to 0	2.75	5.55	3.58	1.56	4.94	3.12	0.45	2.41	1.35	7.07	9.64	6.50
RGB set to 0	30.83	19.22	12.33	33.35	19.26	13.46	20.23	16.89	5.40	37.34	21.59	18.50
Mix Dataset												
DUAL	2.33	2.21	2.31	1.54	1.54	1.66	0.49	0.43	0.47	5.52	5.14	5.27
ALS set to 0	6.89	8.57	6.71	4.02	7.76	6.20	0.84	2.33	2.01	17.27	15.62	12.35
RGB set to 0	4.10	19.74	7.68	2.99	19.61	7.61	0.86	17.96	2.36	9.35	21.76	13.43

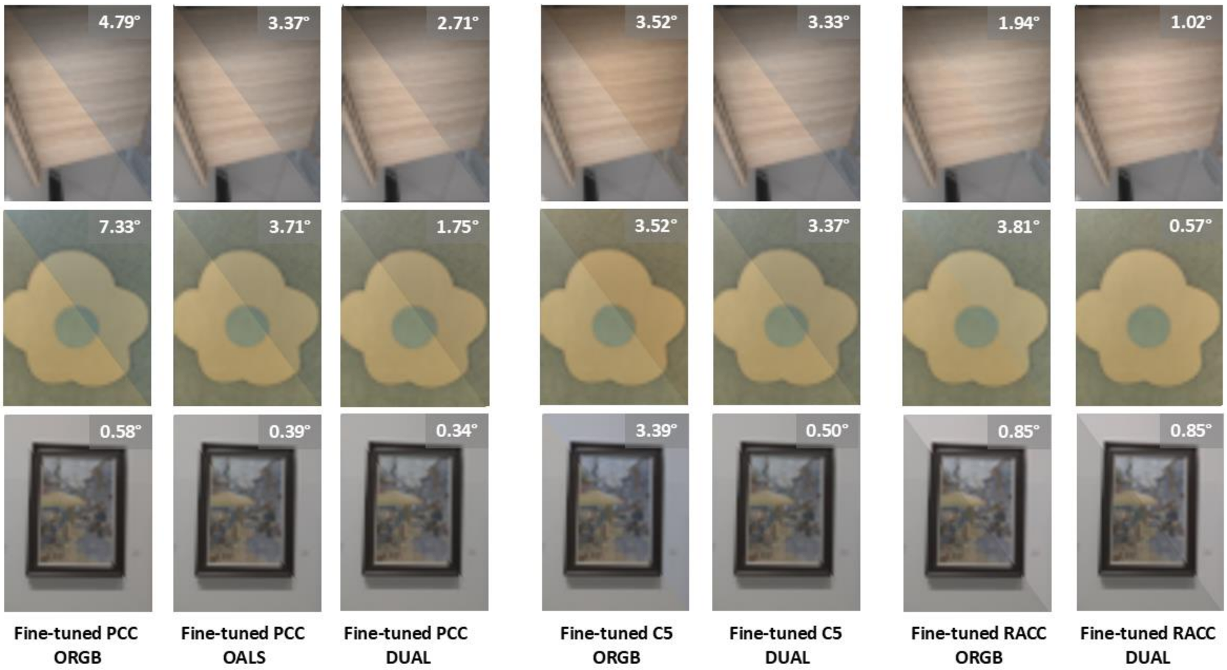


Fig. 4. Illustration of the fine-tuned methods.

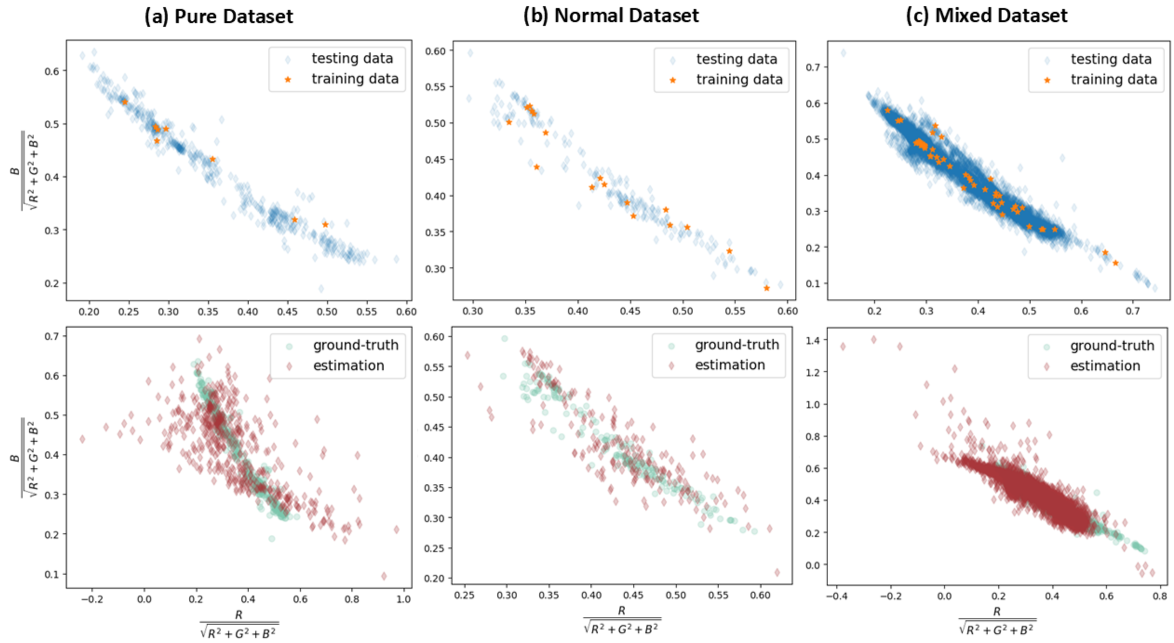


Fig. 5. Illustration of the distribution of the training and testing datasets, and the distribution of the ground-truth and estimated illuminants using the linear-transformation of the ALS signals of the testing datasets.

Table 2. Comparison of the performance, in terms of the angular error, of using linear-transformation and the fine-tuned PCC method for illuminant estimation using the ALS signals as the input.

	MEAN.	MED.	TRI.	BST.25%	WST,25%
Pure Dataset					
PCC (OALS)	3.58	2.89	2.99	0.95	7.37
Linear transformation	7.10	5.28	5.68	1.76	15.63
Normal Dataset					
PCC (OALS)	2.01	1.60	1.72	0.67	3.96
Linear transformation	2.43	1.88	2.03	0.66	5.11
Mixed Dataset					
PCC (OALS)	2.52	1.83	1.99	0.51	5.70
Linear transformation	4.04	3.09	3.30	1.05	8.68

Linear transformation or CNN for ALS to RGB?

Though the learning-based methods reported above can generally produce good results. We were wondering whether it is possible to adopt a linear transformation to convert the ALS signals to RGB, and how would be the performance in comparison to the PCC-OALS method. A K-Means method [19] was used to selects several cluster centers, with the ALS signals and ground-truth RGB signals used as a training set to derive the conversion matrix $M_{13\times3}$. The hyperparameter, the number of clusters, is determined by Gap-statistic [19]. Figure 5 shows the distributions of the training and testing sets and the performance of the illuminant estimation through the linear transformation using the ALS signals, with Table 2 summarizing the angular errors derived from the linear transformation and the PCC-OALS methods. It can be observed that the learning-based method can consistently produce better results.

Conclusion

In this work, we aimed to investigate whether the ALS signals can help the illuminant estimation. We fine-tuned several learning-based illuminant estimation methods, by integrating the ALS signals as the inputs, and tested the methods on a dataset. We found that the ALS signals can definitely improve the accuracy of the illuminant estimation. Also, we found that when ALS signals or RGB image information was missing, the performance of the methods was significantly reduced, which suggests the importance of considering the robustness of the methods. Last but not the least, we compared the performance of learning-based method and linear transformation method for only using the ALS signals as the input, and found that the learning-based method can result in much better performance.

References

[1] D. G. Curry, G. L. Martinsen, and D. G. Hopper, “Capability of the human visual system,” in *Cockpit Displays X*, SPIE, Sept. 2003, pp. 58–69.

[2] I. Erba, M. Buzzelli, J.-B. Thomas, et al., “Improving RGB illuminant estimation exploiting spectral average radiance,” *J. Opt. Soc. Am. A,* 41(3), 516-526 (2024).

[3] D. Cheng, A. Kamel, B. Price, et al., “Two illuminant estimation and user correction preference,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 469-477.

[4] M. Afifi, J. T. Barron, C. LeGendre, et al., “Cross-camera convolutional color constancy,”, *arXiv*: arXiv:2011.11890 (2022).

[5] A. Gijsenij, T. Gevers, and J. van de Weijer, “Computational color constancy: survey and experiments,” *IEEE* Trans Image Process. 20, 2475-2489 (2011).

[6] S. Yue and M. Wei, “Color constancy from a pure color view,” *J. Opt. Soc. Am. A,* 40, 602-610 (2023).

[7] M. Afifi, L. Zhao, A. Punnappurath, et al.,, “Time-aware auto white balance in mobile photography,”, arXiv:2504.05623. (2025).

[8] D. Kim, M. Afifi, D. Kim, et al., “CCMNet: leveraging calibrated color correction matrices for cross-camera color constancy,”, arXiv:2504.07959. (2025).

[9] Y. Hu, B. Wang, and S. Lin, “FC4: fully convolutional color constancy with confidence-weighted pooling,” in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017), pp. 6950-6958.

[10] F. N. Iandola, S. Han, M. W. Moskewicz, et al., “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,”, *arXiv*: arXiv:1602.07360. (2016).

[11] J. T. Barron and Y.-T. Tsai, “Fast Fourier color constancy,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017) pp. 6950-6958.

[12] J. T. Barron, “Convolutional color constancy,”, arXiv:1507.00410. (2015).

[13] P.-C. Hung and Z. Zhang, “Electronic devices with color compensation,” U.S. patent US20210287586A1 (19 April 2019).

[14] R. M. H. Nguyen, D. K. Prasad, and M. S. Brown, “Training-based spectral reconstruction from a single RGB image,” in 2014 European Conference on Computer Vision (ECCV) (2014), pp. 186-201.

[15] D. Cheng, D. K. Prasad, and M. S. Brown, “Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution,” *J. Opt. Soc. Am. A,* 31(5), 1049-1058 (2014).

[16] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: convolutional networks for biomedical image segmentation,”, arXiv:1505.04597. (2015).

[17] F. Laakom, J. Raitoharju, A. Iosifidis, et al., “INTEL-TAU: a color constancy dataset,” arXiv:1910.10404. (2020).

[18] G. Hemrit *et al.*, “Rehabilitating the colorchecker dataset for illuminant estimation,” arXiv: 1805.12262 (2018).

[19] E. Schubert, “Stop using the elbow criterion for k-means and how to choose the number of clusters instead,”, arXiv: 2212.12189. (2023).

Author Biography

Ruikai He is a PhD student at The Hong Kong Polytechnic University. His research mainly focuses on color consistency and color constancy.

Minchen Wei is a professor at The Hong Kong Polytechnic University. He is a Vice President of the International Commission on Illumination (CIE), and editor-in-chief of Color Research and Applicatoin. His research mainly focuses on color and imaging science.