

Multi-Spectral Image Color Reproduction

Jiacheng Li^{1*} Chang Chen^{2*} Xue Hu² Fenglong Song² Youliang Yan² Zhiwei Xiong^{1†}
¹University of Science and Technology of China ²Huawei Noah's Ark Lab

jcleee@mail.ustc.edu.cn zwxiong@ustc.edu.cn

{chenchang25, huxue16, songfenglong, yanyouliang}@huawei.com

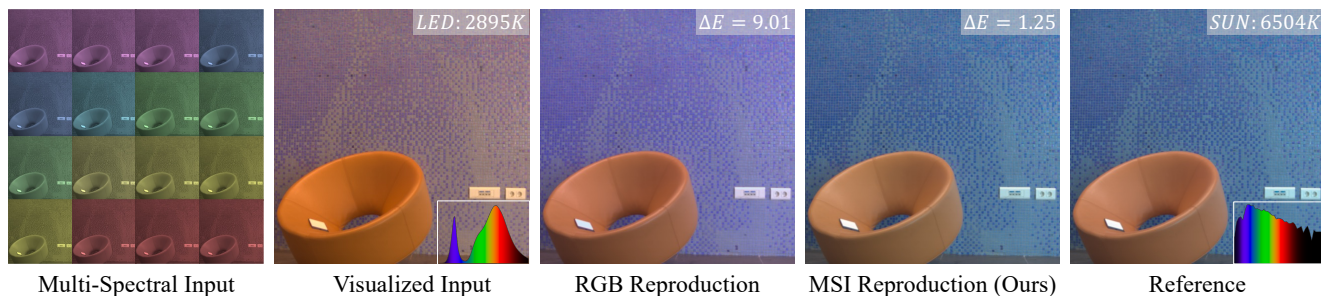


Figure 1. We present a color reproduction system for multi-spectral image (MSI). Our system takes RAW multi-spectral signal as input and produces a standard color representation of that signal for reproducing the color experience of the corresponding scene through any display devices. Here, we show a challenging scene where only a few colors present. The limited spectral resolution of existing RGB cameras leads to incorrect reproduction, while our MSI color reproduction system achieves faithful reproduction. The color differences evaluated with dE2000 (ΔE) are depicted. Images are gamma-corrected and rendered in sRGB for visualization.

Abstract

From camera to screen, researchers have developed a well-established system for capturing and reproducing the color experience of human eyes. In this study, we aim to upgrade this process by transiting from conventional RGB to multi-spectral image (MSI) color reproduction. While MSI offers evident advantages in color matching, we find out it is not trivial to make good use of more spectral information for color constancy. Therefore, we present a regularized color reproduction system that incorporates a spectral prior-guided optimization strategy to establish a sensor-optimized RGB projection for color matching, along with a learning-based chromatic adaptation model for color constancy. Specifically, we define the RGB projection through an end-to-end optimization under the guidance of sensor spectral sensitivities. Subsequently, we devise a chromatic adaptation neural network that estimates the scene illuminance and an illuminance-adaptive matrix for auto white balancing and dynamic color correction, respectively. Comprehensive experiments show the superiority of our system compared to alternative solutions.

*Equal contribution. This work was done when Jiacheng Li was a research intern at Huawei Noah's Ark Lab.

†Corresponding author.

1. Introduction

The objective of color reproduction is to accurately capture and reproduce colors that closely resemble the original colors perceived by human eyes. This pursuit of accurate color reproduction holds great significance across various applications, such as printing, digital imaging, video production, and cinema, where maintaining color fidelity plays a pivotal role in effective visual communication [31, 64, 69].

Beyond matching the color sensitivities of human eyes, accurately reproducing colors necessitates the replication of human's ability to maintain color constancy via chromatic adaptation. From the display perspective, to ensure consistent color reproduction across mediums, device-independent color spaces (e.g., CIE XYZ) and color gamuts (e.g., sRGB) with pre-defined white points (e.g., D65) have been standardized to reproduce human color perception. From the imaging perspective, however, converting RAW signals to standard color representations presents challenges due to varying lighting environments. Additionally, the capture device itself, typically an RGB camera, suffers from a limited spectral resolution, leading to significant spectral information loss during the imaging process and thus unfaithful results in challenging cases, as shown in Fig. 1.

In this study, we embark on an initial exploration aimed at achieving accurate color reproduction through the use of

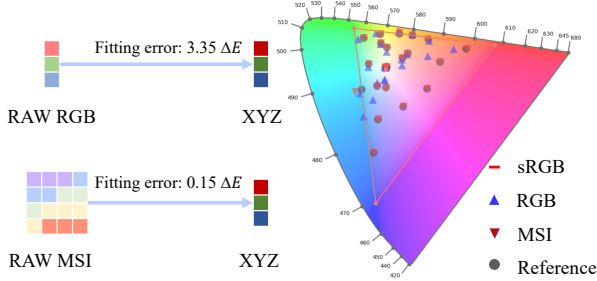


Figure 2. Fitting linear transformation matrices from RAW multi-spectral signals or RAW RGB signals to the CIE XYZ space and the visualization of the projected color checker points in the CIE 1976 UCS chromaticity diagram. The closer to the black reference points, the better.

higher spectral resolution, *i.e.*, multi-spectral image (MSI). Unlike conventional RGB cameras that rely on only three spectral channels, multi-spectral cameras have the ability to capture a broader range of spectral data, enabling more precise and detailed spectral information [17, 45]. To showcase the advantages of MSI in color representation, we present an example of color matching, where we fit linear projection matrices that project RAW MSI or RAW RGB signals to the standard CIE XYZ color space under a standard illuminant. The results shown in Fig. 2 demonstrate the clear advantage of higher spectral resolution in accurately matching the lower-dimensional color space. On the other hand, however, it is not trivial to achieve color constancy through increased spectral resolution, even a scene-adaptive projection matrix is estimated for each MSI. This formulation of direct projection leads to undersaturation, as shown later in Table 1 and Fig. 8.

Thus, the core of an MSI color reproduction system is to determine an optimal and reliable transformation that maps multi-spectral signals to the standard color space, ensuring both accurate color matching and color constancy simultaneously. To accomplish this objective, we design a regularized color pipeline that incorporates a spectral prior-guided optimization strategy to establish a sensor-optimized RGB projection for color matching, in conjunction with a learning-based chromatic adaptation model for color constancy. Beyond fitting the standard CIE XYZ color space, we generalize the principle of color matching and identify a proper RGB subspace for computational color constancy through an end-to-end optimization of our color pipeline under the guidance of sensor spectral sensitivities.

In the identified subspace, we achieve color constancy by introducing a chromatic adaptation neural network that estimates the scene illuminance, as well as an illuminance-adaptive color correction matrix for auto white-balancing and dynamic color correction, respectively. However, obtaining the necessary multi-spectral data to train this

learning-based model can be a time-consuming and costly process. To address this issue, we develop a multi-spectral data augmentation strategy that features a color-to-reflectance retrieval dictionary, enabling the generation of hyper-spectral reflectance data from enormous existing high-quality sRGB images. Consequently, we can pre-train our model on a large-scale synthetic dataset, thereby enhancing its generalization capability and improving its overall robustness.

Our system has been evaluated on both synthetic and real-world datasets, providing comprehensive evidence of its exceptional performance in terms of color accuracy and fidelity, surpassing alternative solutions. Notably, under three representative light conditions, our system achieved less than 1 degree of angular error for most cases, and an average color difference of around 2 ΔE , approaching the discrimination limit of human color perception.

2. Background

2.1. Color Perception in the Human Visual System

Within the human visual system, there exist three distinct types of cones that exhibit sensitivity to different wavelengths of light [35, 62, 70]. Light, serving as an information carrier, can undergo absorption, transmission, or reflection when interacting with objects in a given scene. The color of an object is determined by the wavelengths of light that are reflected, a consequence of the interaction between the object’s reflectance and the illuminant of the scene. Remarkably, humans possess the ability to perceive consistent colors despite variations in illuminance, referred to as color constancy [20, 27, 53]. This ability stems from the human visual system’s capacity to adjust the interpretation of signals from the photoreceptors in the retina through a process known as chromatic adaptation. The von Kries hypothesis [63] is one of the most widely accepted theories in this field, positing that independent gains are applied to the response of human retinal cones. A generalized approach to incorporating the von Kries hypothesis involves employing illuminance-dependent diagonal transformations, integrated within chromatic adaptation transforms (CAT) [24, 52] in color appearance models [21, 54, 60].

2.2. Image Formation Model

Cameras aim to replicate the perception process of the human visual system. In a simplified Lambertian formation model, the captured raw response $\mathbf{I}(\mathbf{x})$ by a camera can be formulated as

$$\mathbf{I}(\mathbf{x}) = \int_{\lambda} e(\lambda) r(\lambda; \mathbf{x}) s_c(\lambda) d\lambda, \quad (1)$$

where the raw response at a pixel location $\mathbf{x}$ is the result of integrating the spectral sensitivity of the sensor $s_c(\lambda)$,

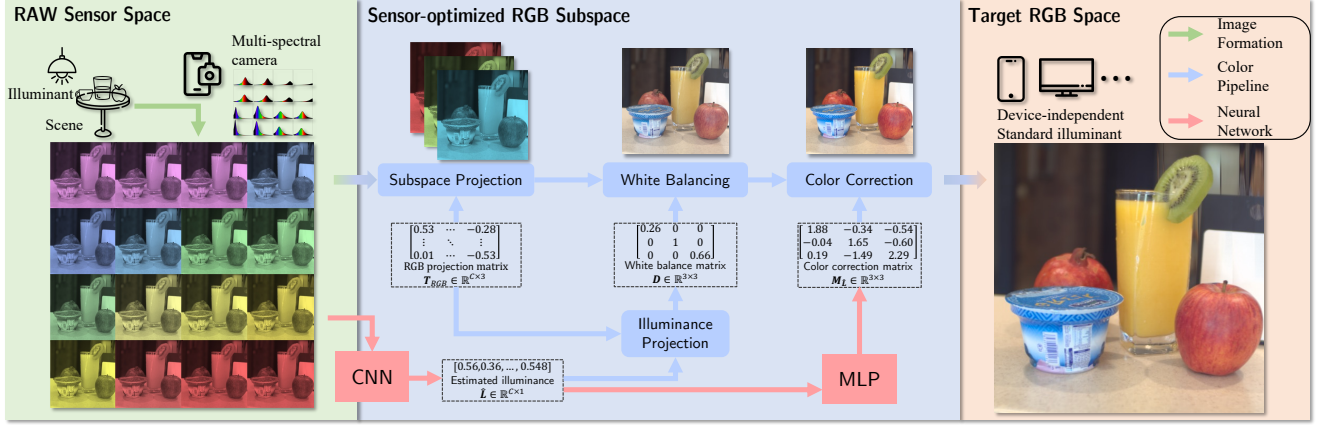


Figure 3. The overview of our proposed color reproduction system. For a scene captured by a multi-spectral camera, we aim to identify an optimal and reliable transformation from the RAW sensor space to target RGB space, where the colors are both device-independent and illuminance-independent. Our regularized pipeline accomplishes this complex transformation in three key stages: RGB subspace projection, white balancing, and color correction, which are implemented through a RGB projection matrix $\mathbf{T}_{RGB}$, a diagonal white balance matrix $\mathbf{D}$, and a color correction matrix $\mathbf{M}_L$, respectively. To achieve scene-specific chromatic adaptation, we devise a neural network, composed of a CNN and an MLP, to estimate the illuminance vector and a residual color correction matrix from the RAW multi-spectral signals. The estimated illuminance vector is further projected to the RGB subspace for auto white balancing, and the residual color correction matrix is incorporated for dynamic color correction. Images are gamma-corrected and rendered in sRGB for visualization.

the spectral reflectance of the object $r(\lambda; \mathbf{x})$, and the spectral power distribution (SPD) $e(\lambda)$ of the scene illuminant. The integration is performed over the wavelength λ for each spectral channel c . This formulation assumes that the scene illuminant is globally uniform across different pixel locations and we define the sensor response of the scene illuminant (referred to as the scene illuminance vector) as $\mathbf{L} = \int_{\lambda} e(\lambda) s_c(\lambda) d\lambda$. In recent years, significant progress has been made in multi-spectral imaging systems to improve the faithful and convenient reconstruction of spectral signals beyond three channels [6, 28, 32, 34, 38, 49, 59, 65, 68]. These advancements have facilitated the potential widespread usage of multi-spectral cameras and motivate us to explore their application in color reproduction.

2.3. Color Reproduction

The primary objective of color reproduction is to map sensor-specific signals to a standardized representation of color that is both *device-independent* and *illuminance-independent*. The former entails using a human-like three-channel RGB space, *i.e.*, CIE XYZ, which was derived from color matching experiments [33, 66]. The latter requires a computational color constancy process that corrects colors to resemble those captured under the reference light of display devices, *e.g.*, 6504K sunlight (D65). Due to manufacturing errors, the camera RAW signal is device-dependent. In practice, a 3×3 color correction matrix, *e.g.*, the one in Fig. 2(a), is calibrated by the manufacturer under standard illuminants for ColorChecker colors. This fixed color correction matrix is used to correct errors, resulting in a

device-independent signal in the CIE XYZ space to achieve color matching. Subsequently, the auto white balancing algorithm in the color pipeline determines the values of the diagonal matrix in the von Kries form for each scene and obtains an illuminance-independent signal to achieve color constancy. Another effective choice of color reproduction pipeline is to conduct color constancy before color matching, which is called wrong von Kries [21]. In summary, the color reproduction problem is defined as a color space conversion problem with linear transformations [8, 50],

$$\mathbf{I}_{RGB} = \mathbf{I}_{RAW} \cdot \mathbf{M}_{CC} \cdot \mathbf{M}_{WB}, \quad (2)$$

where $\mathbf{I}_{RAW} \in \mathbb{R}^{H \times W \times 3}$ refers to the RAW response, $\mathbf{M}_{CC}$ is the color correction matrix, $\mathbf{M}_{WB}$ is the white balancing matrix, and $\mathbf{I}_{RGB} \in \mathbb{R}^{H \times W \times 3}$ represents the color-corrected RGB image. Here, the term “color-corrected” is used to indicate that $\mathbf{I}_{RGB}$ is mapped to the device-independent CIE XYZ space under a standard reference illuminant, such as D65, *i.e.*, complete chromatic adaptation [21]. In this article, we refer to this reference space as the “target RGB space.” After the above linear transformation, the transformed image undergoes gamma correction and gamut mapping, resulting in a display-ready image, *e.g.*, an sRGB image. Please refer to previous works [39, 58] for a comprehensive illustration of color transforms in the image processing pipeline (ISP).

In the context of white balancing, where the goal is to correct achromatic colors (where $R=G=B$), color constancy has been extensively studied [2, 4, 9, 12, 13, 18, 19, 25, 26, 36, 40, 43, 48]. Statistical methods make assumptions in object reflectance across different objects, pro-

ducing poor results when these assumptions are violated [13, 25, 44, 61]. Gamut-based methods have also been proposed [26, 30]. Recent learning-based models have shown promising performance by utilizing annotated data pairs [1, 7, 8, 8, 10, 22, 37, 46, 50, 56, 57, 67, 71, 72]. There is also growing research interest in color constancy beyond achromatic colors [15, 16, 23], *i.e.*, multiplying another 3×3 matrix. Despite the significant progress made in color constancy for RGB cameras, limited efforts have been made in the context of MSIs [14, 41, 42]. Chakrabarti et al. propose utilizing spatio-spectral statistics, similar to RGB color constancy methods [14]. Khan et al. also extend statistical methods to multi-spectral channels [41]. Koskinen et al. focus on illumination estimation using a single-pixel multi-spectral sensor [42]. Other efforts have explored the decomposition of illumination and reflectance in MSIs [47, 73]. Glatt et al. also present a MSI dataset captured with mobile devices for illumination estimation [32]. Distinct from these methods, which only address a part of the color reproduction task, mainly color constancy, we present a complete system to achieve both color matching and color constancy.

3. Multi-Spectral Image Color Reproduction

3.1. Regularized Color Pipeline

Following the formulation of standard RGB color reproduction, the color reproduction task for a multi-spectral image can be formulated as

$$\mathbf{I}_{RGB} = \mathbf{I}_{MSI} \cdot \mathbf{T}, \quad (3)$$

where $\mathbf{I}_{MSI} \in \mathbb{R}^{H \times W \times C}$ refers to the RAW response of the multi-spectral sensor while C denotes the number of spectral channels, $\mathbf{I}_{RGB} \in \mathbb{R}^{H \times W \times 3}$ represents the target RGB image, and $\mathbf{T}$ is a $C \times 3$ matrix.

Different from the standard RGB color pipeline, the introduction of MSI leads to a fundamental challenge to the system design: we cannot define “white” in a multi-channel space. “White” ($R=G=B$) and colors are meaningful *only* in a three-channel space (usually CIE XYZ). To tackle this issue and achieve robust color reproduction, we devise a carefully regularized color pipeline for multi-spectral cameras. As shown in Fig. 3, our color pipeline accomplishes the complex transformation $\mathbf{T}$ in three key stages: RGB subspace projection, white balancing, and color correction. Firstly, we project the multi-spectral RAW response $\mathbf{I}_{MSI}$ onto a three-channel RGB space using a subspace projection matrix $\mathbf{T}_{RGB} \in \mathbb{R}^{C \times 3}$. We assume this RGB subspace meets the von Kries hypothesis. Consequently, we derive a diagonal matrix $\mathbf{D} \in \mathbb{R}^{3 \times 3}$ for white balancing based on the white balance coefficients $\mathbf{W} \in \mathbb{R}^{3 \times 1}$ obtained from the estimated scene illuminance vector $\hat{\mathbf{L}} \in \mathbb{R}^{C \times 1}$. Finally, we further convert the white-balanced image to the target RGB space by applying a color correction matrix $\mathbf{M}_{\hat{\mathbf{L}}}$. In summary, our color pipeline can be formulated as

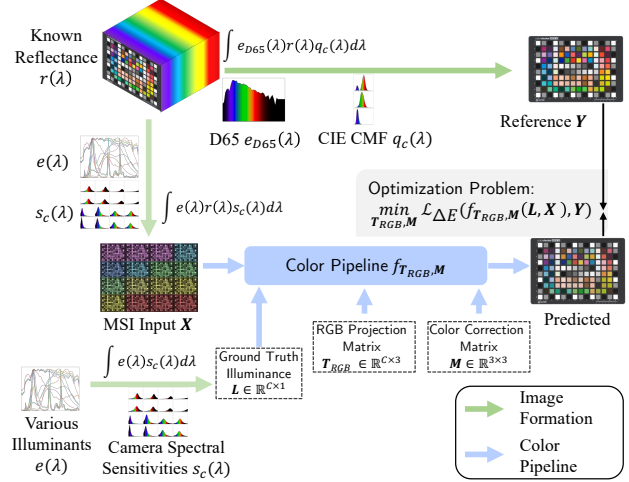


Figure 4. The spectral prior-guided optimization strategy in our system. Guided by the spectral sensitivities of the camera, we construct paired MSI input $\mathbf{X}$, scene illuminance vector $\mathbf{L}$, and the corresponding reference $\mathbf{Y}$ with known reflectance data and illuminants. We formulate an optimization problem to derive a sensor-optimized projection matrix $\mathbf{T}_{RGB}$ and an illuminance-agnostic estimation of the color correction matrix $\mathbf{M}$. Please refer to Algorithm 1 in the supplementary document for more details.

$$\mathbf{I}_{RGB} = \mathbf{I}_{MSI} \cdot \mathbf{T}_{RGB} \cdot \mathbf{D} \cdot \mathbf{M}_{\hat{\mathbf{L}}}, \quad (4)$$

$$\mathbf{D} = \text{diag}(\mathbf{W}), \mathbf{W} = \hat{\mathbf{L}} \cdot \mathbf{T}_{RGB}$$

where $\text{diag}(\cdot)$ is an operation to create an $n \times n$ diagonal matrix from an $n \times 1$ vector.

With this color pipeline, we regularize the composition of the transformation $\mathbf{T}$ following the spirit of the divide-and-conquer paradigm. The primary objective of the RGB subspace projection matrix, $\mathbf{T}_{RGB}$, is to identify an appropriate RGB subspace that facilitates computational color constancy, instead of matching human color sensitivities. Note that this projection should be scene-agnostic, meaning it captures the characteristics of the sensors rather than being specific to a particular scene. Thus, we decompose an illuminance vector $\mathbf{L}$ from the transformation $\mathbf{T}$. This vector is considered a scene-specific variable and plays a crucial role in determining subsequent color constancy transformations. Finally, we partition the complete 3×3 color correction operation by employing a diagonal matrix $\mathbf{D}$ derived from the white balance coefficients, $\mathbf{W}$, and a color correction matrix, $\mathbf{M}$, which is regularized to ensure that the sum of each row is equal to 1. The former corrects achromatic colors, while the latter complements the correction process by addressing the remaining chromatic colors. Please see Fig. 7 and Table 1 for the effectiveness of this design.

3.2. Spectral Prior-guided Optimization

Based on the above formulation of the color pipeline, we present a spectral-prior guided optimization strategy to

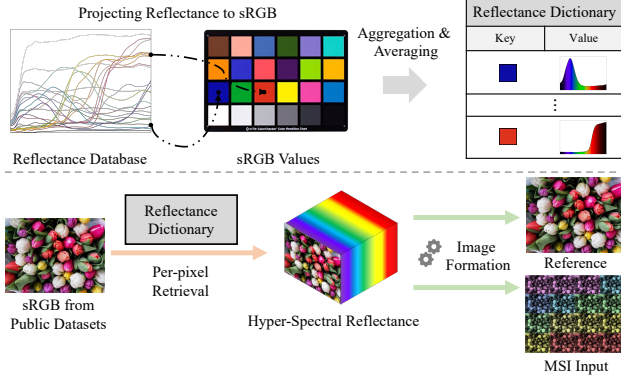


Figure 5. The multi-spectral data augmentation strategy. Upper: We build a mapping dictionary between sRGB value and corresponding hyper-spectral reflectance by projecting reflectance data to the sRGB color gamut. Lower: We query a collection of high-quality sRGB images to acquire a comprehensive dataset of hyper-spectral reflectance. With that, we synthesize MSI inputs and reference images via the image formation model. Please refer to Algorithm 2 in the supplementary document for more details.

identify an appropriate RGB subspace for computational color constancy. As illustrated in Fig. 4, our optimization strategy aims at finding a subspace projection matrix $\mathbf{T}_{RGB}$ and an illuminance-agnostic estimation of $\mathbf{M}$ for color correction. To obtain a robust estimation, we construct pairs of multi-spectral input and reference RGB from various known reflectance database $\mathcal{R}$ and known illuminants $\mathcal{E}$.

$$\mathbf{X}_i = \int_{\lambda} e_i(\lambda) r_i(\lambda) s_c(\lambda) d\lambda, \quad \mathbf{Y}_i = \int_{\lambda} e_{D65}(\lambda) r_i(\lambda) q_c(\lambda) d\lambda, \quad (5)$$

$r_i \in \mathcal{R}, e_i \in \mathcal{E}, i = 1, 2, 3, \dots, N$

where $e_{D65}(\lambda)$ is the SPD of the standard D65 illuminant, s_c the spectral sensitivities of the multi-spectral camera, q_c the CIE 1931 CMF, r_i is randomly sampled from $\mathcal{R}$, e_i from $\mathcal{E}$, and N is the total number of pairs. Besides, the corresponding ground truth illuminance vector can be obtained by $\mathbf{L}_i = \int_{\lambda} e_i(\lambda) s_c(\lambda) d\lambda$. Here, we denote the color pipeline defined in Eq. 4 as $f_{\mathbf{T}_{RGB}, \mathbf{M}}$. The spectral-prior guided optimization problem can then be formulated as

$$\min_{\mathbf{T}_{RGB}, \mathbf{M}} \frac{1}{N} \sum_{i=1}^N \mathcal{L}_{\Delta E}(f_{\mathbf{T}_{RGB}, \mathbf{M}}(\mathbf{X}_i, \mathbf{L}_i), \mathbf{Y}_i), \quad (6)$$

where $\mathcal{L}_{\Delta E}$ is a distance function that measures the color difference, and we adopt the well-recognized dE2000 [51]. With this optimization process, we obtain the optimal $\mathbf{T}_{RGB}$ and the illuminance-agnostic estimation of $\mathbf{M}$, thus defining the appropriate RGB subspace for color constancy. As shown later in Fig. 6, we step up further from the common CIE XYZ choice in the standard RGB pipeline by identifying a *better* sensor-optimized subspace through this end-to-end optimization process.

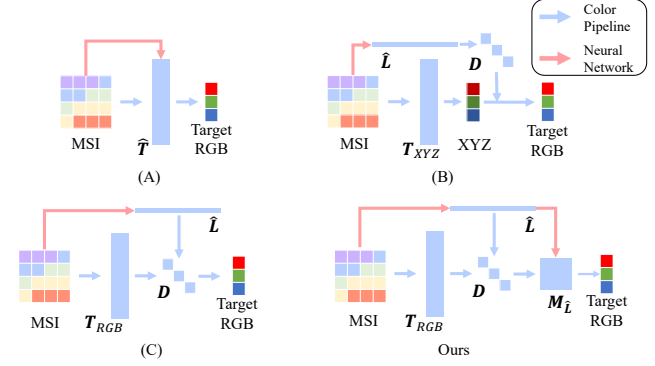


Figure 6. Our color pipeline and variants: (A) direct projection, (B) CIE XYZ color correction, and (C) diagonal-only color correction. $\hat{\mathbf{T}}$, $\mathbf{T}_{XYZ}$, and $\mathbf{T}_{RGB}$ denote projection matrices from MSI to a three-channel space, $\hat{\mathbf{L}}$ the scene illuminance vector, $\mathbf{D}$ the white balance matrix, and $\mathbf{M}_{\hat{\mathbf{L}}}$ the color correction matrix.

3.3. Learning-based Chromatic Adaptation

Having established a sensor-optimized RGB subspace for computational color constancy using $\mathbf{T}_{RGB}$, we now introduce a learning-based chromatic adaptation neural network to correct RGB colors in that subspace to the target RGB. As depicted in Fig. 3, the chromatic adaptation neural network consists of two main components: a convolutional neural network (CNN) and a multi-layer perceptron (MLP) network. We estimate the scene illuminance vector $\hat{\mathbf{L}}$ from the multi-spectral image $\mathbf{I}_{MSI}$, and we predict an illuminance-adaptive residual for the color correction matrix $\mathbf{M}$ from the estimated $\hat{\mathbf{L}}$. To perform white balancing, we project the estimated illuminance vector $\hat{\mathbf{L}}$ onto the RGB subspace using $\mathbf{T}_{RGB}$, resulting in the white balance coefficients $\mathbf{W}$. These coefficients are used to construct a diagonal matrix $\mathbf{D}$ for automatic white balancing. The predicted illuminance-adaptive residual of the color correction matrix is then combined with the illuminance-agnostic estimation to obtain the final color correction matrix $\mathbf{M}_{\hat{\mathbf{L}}}$ for dynamic color correction. For the CNN, we choose the CSPDNet architecture from YOLOv4 [11] for its simplicity. More advanced architectures are also applicable.

The training of the chromatic adaptation network incorporates two loss functions: the angular error (AE) for illumination estimation and the root-mean-square error (RMSE) for the illuminance-adaptive color correction matrix. The AE between the estimated illuminance vector $\hat{\mathbf{L}}$ and ground truth $\mathbf{L}$ is defined as

$$\mathcal{L}_{AE}(\hat{\mathbf{L}}, \mathbf{L}) = \cos^{-1} \left(\frac{\hat{\mathbf{L}}^T \mathbf{L}}{\|\hat{\mathbf{L}}\| \|\mathbf{L}\|} \right). \quad (7)$$

For each illuminant, we optimize a corresponding color correction matrix $\mathbf{M}^{GT}$ by fixing $\mathbf{T}_{RGB}$ in Eq. 6. Thus, the loss function for dynamic color correction can be formulated as

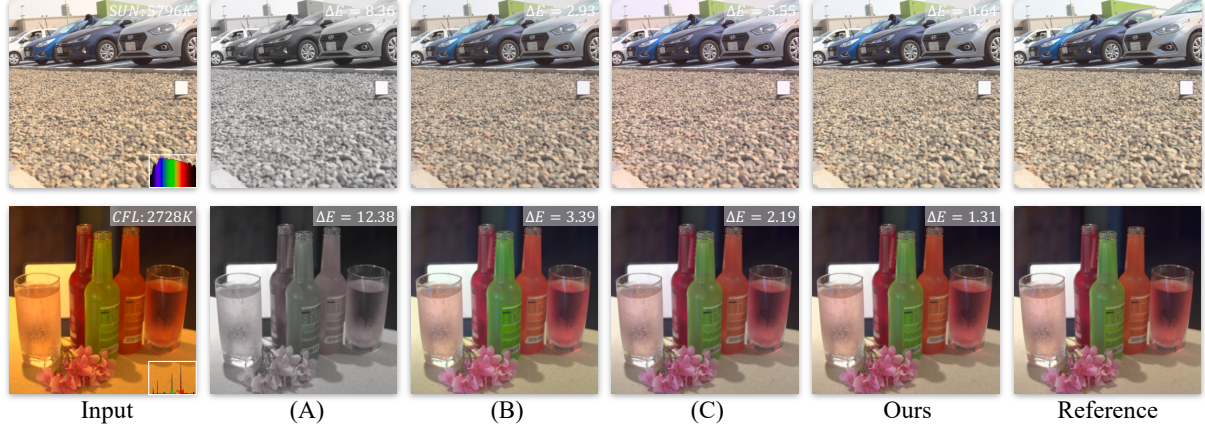


Figure 8. Visual comparison of our color pipeline and variant color pipelines: (A) direct projection, (B) CIE XYZ color correction, and (C) diagonal-only color correction. The MSI inputs are visualized under the scene illuminance via the CIE 1931 color matching function. The color differences evaluated with dE2000 (ΔE) are depicted. Images are gamma-corrected and rendered in sRGB for visualization.

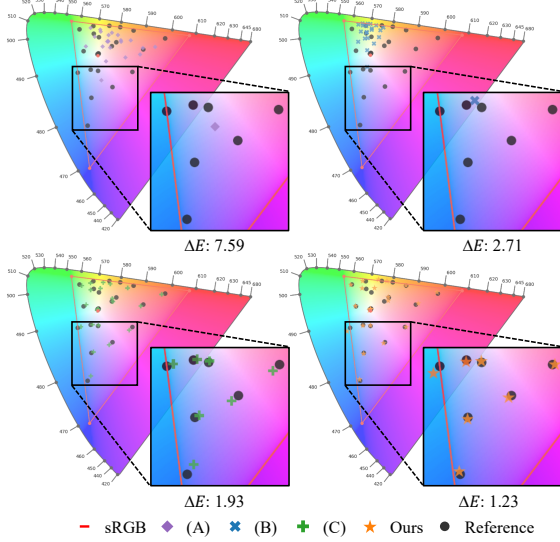


Figure 7. Comparison between our regularized color pipeline and its variants (as described in Fig. 6) in reproducing colors from a standard color checker. We visualize the reproduced color values of different pipelines on the CIE 1976 UCS chromaticity diagram. The closer to the black reference points, the better.

$$\mathcal{L}_{RMSE}(\mathbf{M}_{\hat{\mathbf{L}}}, \mathbf{M}_{\mathbf{L}}) = \sqrt{\|\mathbf{M}_{\hat{\mathbf{L}}} - \mathbf{M}_{\mathbf{L}}\|}. \quad (8)$$

We train the CNN and the MLP in our chromatic adaptation network with the above two losses in an end-to-end manner.

To facilitate the training of our chromatic adaptation network, we develop a data augmentation strategy that utilizes a color-to-reflectance retrieval dictionary to generate hyper-spectral reflectance data from easily available sRGB images, which we use to synthesize rich training pairs of multi-spectral inputs and corresponding canonical RGB representations. As illustrated in Fig. 5, we build a color-to-reflectance dictionary by projecting single-pixel reflectance data to the sRGB color gamut and averaging the reflectance samples that are mapped to the same RGB value.

Table 1. Comparison between our color pipeline and its variants in terms of color reproduction, evaluated with the dE2000 metric on the KAUST dataset. Lower is better. We denote the upper bound results when the scene illuminance is known in gray brackets, which indicate the systematic error of a pipeline. Pipeline (A) does not estimate illuminances, thus its upper bound is left blank.

	CFL	LED	SUN
(A) Direct projection	12.06(-)	12.41(-)	10.83(-)
(B) CIE XYZ color correction	4.46(4.29)	3.06(2.81)	4.05(3.76)
(C) Diagonal-only color correction	3.73(2.63)	2.97(2.27)	3.02(1.19)
Ours	2.64 (2.43)	2.50 (2.05)	1.33 (0.91)

We build a large-scale synthetic multi-spectral reflectance dataset through per-pixel retrieval of this dictionary with high-quality public sRGB datasets, such as DIV2K [3], Multillumin [55], etc. With the synthetic reflectance dataset, we can synthesize the training pairs of multi-spectral image $\mathbf{I}_{MSI}$, illumination vector $\mathbf{L}$, and reference RGB image $\mathbf{I}_{RGB}$ through an image formation process similar to Eq. 5. Here, we make a strong assumption that these sRGB images are well-balanced with the reference D65 illumination. With these synthetic training pairs, we pre-train our chromatic adaptation network to improve its robustness and generalization capacity. Subsequently, we fine-tune the network using a real reflectance dataset, further improving its performance in real-world scenarios. More details about the augmentation process and the synthetic dataset are provided in the supplementary document.

4. Experiments and Results

4.1. Benchmark and Metrics

In addition to the synthetic dataset, we adopt a real-captured hyper-spectral reflectance dataset, KAUST dataset [47], to fine-tune our chromatic adaptation model and evaluate our color reproduction system on a hold-out split of this dataset. Furthermore, we collect a dataset of real-world

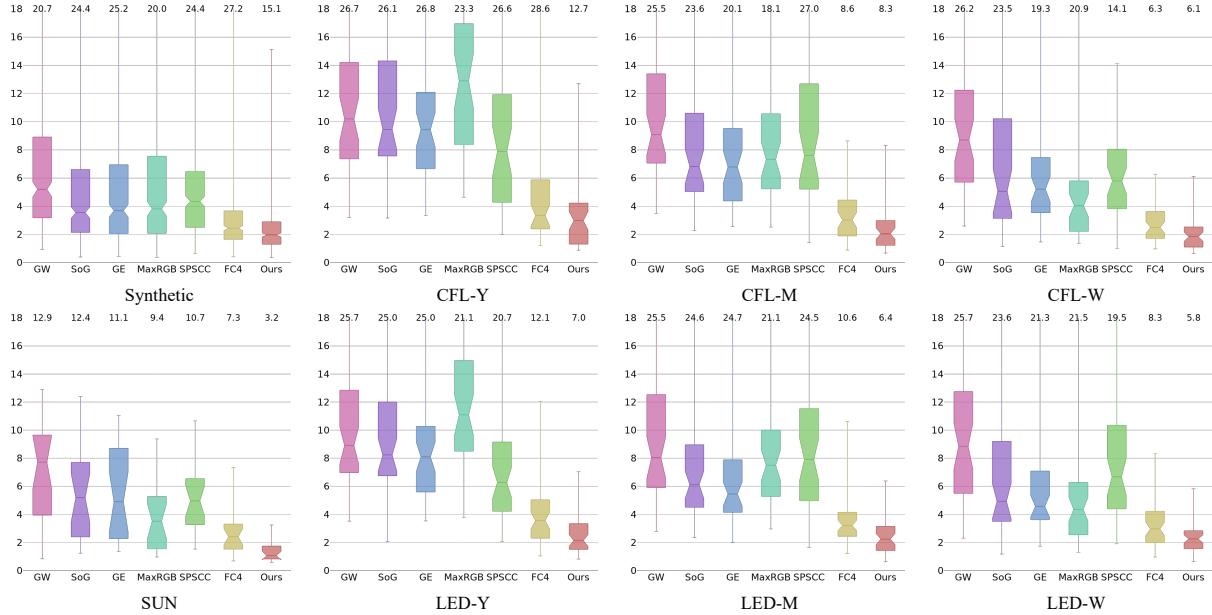


Figure 9. Comparison between our system and other color constancy methods in terms of white balancing performance, evaluated with angular error on both synthetic and real reflectance dataset. Lower is better. We report results across three illuminant types (CFL, LED, and SUN) and different correlated color temperature (CCT) ranges (Y, M, and W). “Y” denotes CCTs less than 3000K, “M” in range from 3000K to 5000K, and “W” above 5000K. The maximum errors of each method are annotated on the top of each box and whisker plot.

illuminants. We select three representative light sources, including sunlight (SUN), compact fluorescent lamp (CFL), and LED lamp (LED), and collect a total 385 illuminants across various temperatures. For the camera spectral sensitivities in MSI image formation, we adopt the response function of a prototype MSFA camera sensor provided by the NTIRE 2022 Spectral Demosaicing Challenge [5]. More details are available in the supplementary document.

To evaluate our color reproduction system, we consider two perspectives. Firstly, we measure the angular error between the predicted white board pixels and the white pixels under a standard reference illuminant [29]. This evaluation serves as an indicator of the accuracy of achromatic colors, *i.e.*, white balancing. Secondly, we assess the color difference between the reproduced image and the reference image using the dE2000 (ΔE) [51]. This evaluation quantifies the color reproduction accuracy across the entire image.

4.2. Comparison with Variant Color Pipelines

Firstly, we evaluate our regularized color pipeline by comparing it with its variants. As illustrated in Fig. 6, we design the following three variants for comparison. (A) *Direct projection*: This pipeline involves estimating an adaptation transformation directly from the multi-spectral image, and applying it to project the RAW MSIs to the target RGB space. This pipeline can be trained with supervision from either the reference RGB images or the projected ground truth illuminance vectors. (B) *CIE XYZ color correction*: In this pipeline, we first project the RAW MSIs to the stan-

dard CIE XYZ space. Then, a chromatic adaptation model is employed to estimate the scene illuminance, which is further utilized for color constancy using a CAT within the CIE XYZ space. (C) *Diagonal-only correction*: This pipeline is similar to ours, except for the omission of the color correction matrix. All these variants undergo a similar optimization process described in Sec. 3.2 to obtain globally optimal estimations for RGB subspace projections. Pipelines (B), (C), and our pipeline share the same chromatic adaptation model for illuminance estimation.

Fig. 7 visualizes the errors in reproducing colors from a standard color checker of these color pipelines in the CIE 1976 UCS chromatic diagram. Pipeline (A) yields a poor performance of 7.59 ΔE , indicating the challenge of finding an optimal RGB projection without chromatic adaptation. Compared to pipeline (B), which has an error of 2.71 ΔE , our regularized pipeline identifies a better RGB subspace than the standard CIE XYZ space for computational color constancy. Our pipeline also outperforms pipeline (C), thanks to the additional color correction matrix, whose influence can not be covered by the projection matrix $\mathbf{T}_{RGB}$, despite both being linear transformations. Notably, training pipeline (A) with supervision from the projected ground truth illuminance vector leads to under-saturated results (see Fig. 8). And training with direct supervision from the target RGB images results in divergence due to the “many-to-one” nature of the color reproduction task, showing the necessity of decomposing the complex linear transformation into different stages and predicting

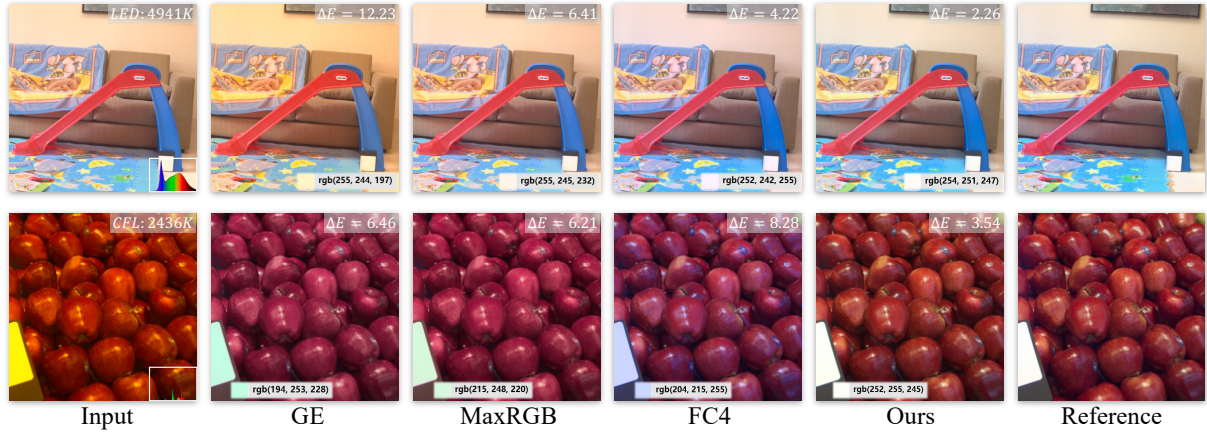


Figure 10. Visual comparison with color constancy methods on the KAUST dataset. The MSI inputs are visualized under the scene illuminance via the CIE 1931 CMF. The color differences evaluated with dE2000 (ΔE) and pixel values on the white board are depicted. Images are gamma-corrected and rendered in sRGB for visualization.

Table 2. Comparison between our system and other chromatic adaptation methods in terms of color reproduction, evaluated with the dE2000 metric on the KAUST dataset. Lower is better.

	CFL			LED			SUN
	Y	M	W	Y	M	W	ALL
GrayWorld [13]	10.93	9.98	9.52	10.04	9.47	9.47	7.31
Shades-of-Gray [25]	11.07	8.43	7.15	9.71	7.71	7.09	5.45
GrayEdge [61]	12.08	9.05	7.40	10.21	7.88	7.09	5.31
MaxRGB [44]	13.21	8.47	4.77	11.49	8.10	5.16	3.79
FC4 (RGB) [37]	6.42	5.30	4.34	4.53	4.19	3.77	4.02
SPSCC [42]	8.85	10.33	6.13	7.17	8.80	7.71	5.26
FC4 (MSI) [37]	4.82	3.40	2.80	3.82	3.80	3.26	2.62
Ours	3.37	2.47	2.08	2.61	2.53	2.36	1.33

coefficients in these transformation matrices with learning-based models. Finally, we present quantitative results on the KAUST dataset in Table 1, as well as a comparison of qualitative results in Fig. 8. Our regularized pipeline achieves the lowest errors as well as the best upper bound.

4.3. Comparison with Color Constancy Methods

We select four statistical methods for comparison, including GreyWorld (GW) [13], Shades-of-Grey (SoG) [25], GreyEdge (GE) [61], and MaxRGB [44]. We optimize a $C \times 3$ transformation that converts MSIs to the standard CIE XYZ space and apply these methods through CATs, *i.e.*, with color pipeline variant (B). Additionally, we compare our approach with SPSCC, a single-pixel multi-spectral color constancy method [42]. Finally, we adapt FC4 [37], a representative learning-based model, to multi-spectral illuminance estimation, making it a stronger baseline than the original RGB version (see Table 2). Both SPSCC and FC4 (MSI) achieve computational color constancy in our sensor-optimized RGB subspace, *i.e.*, with our color pipeline.

Quantitative results on both synthetic and real reflectance datasets in terms of white balancing are reported in Fig. 9. Table 2 lists the quantitative results in terms of color difference. As can be seen, our chromatic adaptation model

shows a significant advantage over existing methods, especially in worst-case scenarios. For example, in the KAUST dataset, our chromatic model achieves less than 1° angular error for the third quartile under all types of illuminants and temperatures, showcasing exceptional robustness to diverse lighting conditions. Moreover, our system outperforms other color constancy methods by more than $1.2 \Delta E$ in terms of color difference, demonstrating superior capabilities in faithful color reproduction. Fig. 10 presents qualitative comparisons of results obtained from different color constancy methods. Particularly, our system performs well in challenging cases with a limited number of colors present in a scene, as depicted in the second case of Fig. 10.

5. Discussion and Conclusion

Although showing superior performance, our system makes certain assumptions and simplifications. Further discussions on exposure quality, demosaicing, lens shading, incomplete chromatic adaptation, and more ablation results are provided in the supplementary document.

To summarize, we conduct an initial investigation into the application of multi-spectral cameras as photography devices, revealing their considerable potential in enhancing color matching and color constancy. The findings presented here not only contribute to the understanding of the advantages associated with multi-spectral cameras but also underscore the importance of a holistic system design for unlocking the full potential of both hardware and software components. We believe our exploration could shed light on this promising direction.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grants 62131003 and 62021001. We acknowledge the support of MindSpore.

References

- [1] Mahmoud Afifi, Jonathan T. Barron, Chloe LeGendre, Yun-Ta Tsai, and Francois Bleibel. Cross-camera convolutional color constancy. In *ICCV*, 2021. 4
- [2] Mahmoud Afifi, Marcus A. Brubaker, and Michael S. Brown. Auto white-balance correction for mixed-illuminant scenes. In *WACV*, 2022. 3
- [3] Eirikur Agustsson and Radu Timofte. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In *CVPR Workshops*, 2017. 6
- [4] Arash Akbarinia and C. Alejandro Párraga. Colour constancy beyond the classical receptive field. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(9):2081–2094, 2018. 3
- [5] Boaz Arad, Radu Timofte, Rony Yahel, Nimrod Morag, and Amir Bernat. NTIRE 2022 spectral demosaicing challenge and data set. In *CVPR*, 2022. 7
- [6] Seung-Hwan Baek, Incheol Kim, Diego Gutierrez, and Min H. Kim. Compact single-shot hyperspectral imaging using a prism. *ACM Trans. Graph.*, 36(6):217:1–217:12, 2017. 3
- [7] Jonathan T. Barron. Convolutional color constancy. In *ICCV*, 2015. 4
- [8] Jonathan T. Barron and Yun-Ta Tsai. Fast fourier color constancy. In *CVPR*, 2017. 3, 4
- [9] Simone Bianco and Claudio Cusano. Quasi-unsupervised color constancy. In *CVPR*, 2019. 3
- [10] Simone Bianco, Claudio Cusano, and Raimondo Schettini. Single and multiple illuminant estimation using convolutional neural networks. *IEEE Trans. Image Process.*, 26(9):4347–4362, 2017. 4
- [11] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *CoRR*, abs/2004.10934, 2020. 5
- [12] Ivaylo Boyadzhiev, Kavita Bala, Sylvain Paris, and Frédo Durand. User-guided white balance for mixed lighting conditions. *ACM Trans. Graph.*, 31(6):200:1–200:10, 2012. 3
- [13] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1):1–26, 1980. 3, 4, 8
- [14] Ayan Chakrabarti, Keigo Hirakawa, and Todd E. Zickler. Color constancy with spatio-spectral statistics. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(8):1509–1519, 2012. 4
- [15] Dongliang Cheng, Brian L. Price, Scott Cohen, and Michael S. Brown. Beyond white: Ground truth colors for color constancy correction. In *ICCV*, 2015. 4
- [16] Hamilton Y. Chong, Steven J. Gortler, and Todd E. Zickler. The von kries hypothesis and a basis for color constancy. In *ICCV*, 2007. 4
- [17] Valerie C Coffey. Hyperspectral imaging for safety and security. *Optics and Photonics News*, 26(10):26–33, 2015. 2
- [18] Partha Das, Yang Liu, Sezer Karaoglu, and Theo Gevers. Generative models for multi-illumination color constancy. In *CVPR Workshops*, 2021. 3
- [19] Marc Ebner. Color constancy using local color shifts. In Tomás Pajdla and Jiri Matas, editors, *ECCV*, 2004. 3
- [20] Marc Ebner. *Color constancy*, volume 7. 2007. 2
- [21] Mark D Fairchild. *Color appearance models*. 2013. 2, 3
- [22] Graham D. Finlayson. Corrected-moment illuminant estimation. In *ICCV*, 2013. 4
- [23] Graham D. Finlayson, Mark S. Drew, and Brian V. Funt. Diagonal transforms suffice for color constancy. In *ICCV*, 1993. 4
- [24] Graham D Finlayson, Mark S Drew, and Brian V Funt. Spectral sharpening: sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11(5):1553–1563, 1994. 2
- [25] Graham D. Finlayson and Elisabetta Trezzi. Shades of gray and colour constancy. In *Color Imag. Conf.*, 2004. 3, 4, 8
- [26] David A. Forsyth. A novel algorithm for color constancy. *Int. J. Comput. Vis.*, 5(1):5–35, 1990. 3, 4
- [27] David H Foster. Color constancy. *Vision research*, 51(7):674–700, 2011. 2
- [28] Michael E Gehm, Renu John, David J Brady, Rebecca M Willett, and Timothy J Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics express*, 15(21):14013–14027, 2007. 3
- [29] Arjan Gijsenij, Theo Gevers, and Marcel P Lucassen. Perceptual analysis of distance measures for color constancy algorithms. *J. Opt. Soc. Am. A*, 26(10):2243–2256, 2009. 7
- [30] Arjan Gijsenij, Theo Gevers, and Joost van de Weijer. Generalized gamut mapping using image derivative structures for color constancy. *Int. J. Comput. Vis.*, 86(2-3):127–139, 2010. 4
- [31] Arjan Gijsenij, Theo Gevers, and Joost van de Weijer. Computational color constancy: Survey and experiments. *IEEE Trans. Image Process.*, 20(9):2475–2489, 2011. 1
- [32] Ortal Glatt, Yotam Ater, Woo-Shik Kim, Shira Werman, Oded Berby, Yael Zini, Shay Zelinger, Sangyoon Lee, Heejin Choi, and Evgeny Soloveichik. Beyond RGB: A real world dataset for multispectral imaging in mobile devices. In *WACV*, 2024. 3, 4
- [33] John Guild. The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 230(681-693):149–187, 1931. 3
- [34] Ralf Habel, Michael W. Kudenov, and Michael Wimmer. Practical spectral photography. *Comput. Graph. Forum*, 31(2pt2):449–458, 2012. 3
- [35] Ewald Hering. Outlines of a theory of the light sense. 1964. 2
- [36] Eugene Hsu, Tom Mertens, Sylvain Paris, Shai Avidan, and Frédo Durand. Light mixture estimation for spatially varying white balance. *ACM Trans. Graph.*, 27(3):70, 2008. 3
- [37] Yuanming Hu, Baoyuan Wang, and Stephen Lin. Fc⁴: Fully convolutional color constancy with confidence-weighted pooling. In *CVPR*, 2017. 4, 8
- [38] Daniel S. Jeon, Seung-Hwan Baek, Shinyoung Yi, Qiang Fu, Xiong Dun, Wolfgang Heidrich, and Min H. Kim. Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Trans. Graph.*, 38(4):117:1–117:13, 2019. 3
- [39] Hakki Can Karaimer and Michael S. Brown. A software platform for manipulating the camera imaging pipeline. In *ECCV*, 2016. 3

- [40] Hakki Can Karaimer and Michael S. Brown. Improving color reproduction accuracy on cameras. In *CVPR*, 2018. 3
- [41] Haris Ahmad Khan, Jean-Baptiste Thomas, Jon Yngve Hardeberg, and Olivier Laligant. Illuminant estimation in multi-spectral imaging. *J. Opt. Soc. Am. A*, 34(7):1085–1098, Jul 2017. 4
- [42] Samu Koskinen, Erman Acar, and Joni-Kristian Kamarainen. Single pixel spectral color constancy. In *BMVC*, 2021. 4, 8
- [43] Firas Laakom, Nikolaos Passalis, Jenni Raitoharju, Jarno Nikkanen, Anastasios Tefas, Alexandros Iosifidis, and Moncef Gabbouj. Bag of color features for color constancy. *IEEE Trans. Image Process.*, 29:7722–7734, 2020. 3
- [44] Edwin H. Land and John J. McCann. Lightness and retinex theory. *J. Opt. Soc. Am.*, 61(1):1–11, Jan 1971. 4, 8
- [45] Richard M Levenson and James R Mansfield. Multispectral imaging in biology and medicine: slices of life. *Cytometry Part A: the journal of the International Society for Analytical Cytology*, 69(8):748–758, 2006. 2
- [46] Bing Li, Haina Qin, Weihua Xiong, Yangxi Li, Songhe Feng, Weiming Hu, and Stephen J. Maybank. Ranking-based color constancy with limited training samples. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(10):12304–12320, 2023. 4
- [47] Yuqi Li, Qiang Fu, and Wolfgang Heidrich. Multispectral illumination estimation using deep unrolling network. In *ICCV*, 2021. 4, 6
- [48] Orly Liba, Kiran Murthy, Yun-Ta Tsai, Tim Brooks, Tianfan Xue, Nikhil Karnad, Qiurui He, Jonathan T. Barron, Dillon Sharlet, Ryan Geiss, Samuel W. Hasinoff, Yael Pritch, and Marc Levoy. Handheld mobile photography in very low light. *ACM Trans. Graph.*, 38(6):164:1–164:16, 2019. 3
- [49] Xing Lin, Yebin Liu, Jiamin Wu, and Qionghai Dai. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. Graph.*, 33(6):233:1–233:11, 2014. 3
- [50] Yi-Chen Lo, Chia-Che Chang, Hsuan-Chao Chiu, Yu-Hao Huang, Chia-Ping Chen, Yu-Lin Chang, and Kevin Jou. CLCC: contrastive learning for color constancy. In *CVPR*, 2021. 3, 4
- [51] M Ronnier Luo, Guihua Cui, and Bryan Rigg. The development of the cie 2000 colour-difference formula: Ciede2000. *Color Research & Application*, 26(5):340–350, 2001. 5, 7
- [52] M Ronnier Luo et al. A review of chromatic adaptation transforms. *Review of Progress in Coloration and Related Topics*, 30:77–92, 2000. 2
- [53] John J McCann, Suzanne P McKee, and Thomas H Taylor. Quantitative studies in retinex theory a comparison between theoretical predictions and observer responses to the “color mondrian” experiments. *Vision research*, 16(5):445–453, 1976. 2
- [54] Nathan Moroney, Mark Fairchild, Robert Hunt, and Changjun Li. The ciecam02 color appearance model. 2002. 2
- [55] Lukas Murmann, Michaël Gharbi, Miika Aittala, and Frédéric Durand. A dataset of multi-illumination images in the wild. In *ICCV*, 2019. 6
- [56] Taishi Ono, Yuhi Kondo, Legong Sun, Teppei Kurita, and Yusuke Moriuchi. Degree-of-linear-polarization-based color constancy. In *CVPR*, 2022. 4
- [57] Jueqin Qiu, Haisong Xu, and Zhengnan Ye. Color constancy by reweighting image feature maps. *IEEE Trans. Image Process.*, 29:5711–5721, 2020. 4
- [58] Rajeev Ramanath, Wesley E. Snyder, Youngjun Yoo, and Mark S. Drew. Color image processing pipeline. *IEEE Signal Process. Mag.*, 22(1):34–43, 2005. 3
- [59] Zhan Shi, Chang Chen, Zhiwei Xiong, Dong Liu, and Feng Wu. HSCNN+: advanced cnn-based hyperspectral recovery from RGB images. In *CVPR Workshops*, 2018. 3
- [60] CIE TC. The cie 1997 interim colour appearance model (simple version), ciecam97s. 1998. 2
- [61] Joost van de Weijer, Theo Gevers, and Arjan Gijsenij. Edge-based color constancy. *IEEE Trans. Image Process.*, 16(9):2207–2214, 2007. 4, 8
- [62] H. von Helmholtz and J.P.C. Southall. *Helmholtz’s Treatise on Physiological Optics*. Number v. 3 in Classics of ophthalmology library. 1925. 2
- [63] Johannes von Kries. Beitrag zur physiologie der gesichtsempfindung. *Arch. Anat. Physiol.*, 2:505–524, 1878. 2
- [64] Brian A Wandell and Louis D Silverstein. Digital color reproduction. *The Science of Color*, 2:281–316, 2003. 1
- [65] Zhu Wang, Soongyu Yi, Ang Chen, Ming Zhou, Ting Shan Luk, Anthony James, John Nogan, Willard Ross, Graham Joe, Alireza Shahsafi, et al. Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nature communications*, 10(1):1020, 2019. 3
- [66] William David Wright. A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4):141, 1929. 3
- [67] Jin Xiao, Shuhang Gu, and Lei Zhang. Multi-domain learning for accurate and few-shot color constancy. In *CVPR*, 2020. 4
- [68] Zhiwei Xiong, Zhan Shi, Huiqun Li, Lizhi Wang, Dong Liu, and Feng Wu. HSCNN: cnn-based hyperspectral image recovery from spectrally undersampled projections. In *CVPR Workshops*, 2017. 3
- [69] Sergej Nikolaevich Yendrikhovskij, FJJ Blommaert, and H De Ridder. Color reproduction and the naturalness constraint. *Color Research & Application*, 24(1):52–67, 1999. 1
- [70] Thomas Young. Ii. the bakerian lecture. on the theory of light and colours. *Philosophical transactions of the Royal Society of London*, (92):12–48, 1802. 2
- [71] Huanglin Yu, Ke Chen, Kaiqi Wang, Yanlin Qian, Zhaoxiang Zhang, and Kui Jia. Cascading convolutional color constancy. In *AAAI*, 2020. 4
- [72] Zichen Zhao, Hai-Miao Hu, Hongda Zhang, Fei Chen, and Qiang Guo. Improving color constancy using chromaticity-line prior. *IEEE Trans. Multimed.*, 25:3642–3656, 2023. 4
- [73] Yinqiang Zheng, Imari Sato, and Yoichi Sato. Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization. In *CVPR*, 2015. 4