Check for updates

Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

RGB color constancy using multispectral pixel information

Ilaria Erba, Marco Buzzelli,* and Raimondo Schettini

Department of Informatics Systems and Communication, University of Milano-Bicocca, Milano, Italy
**marco.buzzelli@unimib.it*

Received 21 September 2023; revised 17 November 2023; accepted 14 December 2023; posted 14 December 2023; published 9 January 2024

Multispectral imaging is a technique that captures data across several bands of the light spectrum, and it can be useful in many computer vision fields, including color constancy. We propose a method that exploits multispectral imaging for illuminant estimation, and then applies illuminant correction in the raw RGB domain to achieve computational color constancy. Our proposed method is composed of two steps: first, a selected number of existing camera-independent algorithms for illuminant estimation, originally designed for RGB data, are applied in generalized form to work with multispectral data. We demonstrate that the sole multispectral extension of such algorithms is not sufficient to achieve color constancy, and thus we introduce a second step, in which we re-elaborate the multispectral estimations before conversion into raw RGB with the use of the camera response function. Our results on the NUS dataset show that an improvement of 60% in the color constancy performance, measured in terms of reproduction angular error, can be obtained according to our method when compared to the traditional raw RGB pipeline. © 2024 Optica Publishing Group

https://doi.org/10.1364/JOSAA.506186

1. INTRODUCTION

The human eye has the ability to partially discount for the change in the illumination of a scene, allowing for a coherent perception of color in different lighting conditions: this ability is often referred to as human color constancy [1]. Transferring the same ability to digital camera sensors is often referred to as computational color constancy: this problem has been vastly explored in the raw RGB domain with a variety of approaches [2]; however, a discrepancy between human and computational color constancy still persists [3,4]. Computational color constancy is commonly addressed as a two-stage operation: the former is specialized in estimating the color of the scene illuminant, and the latter corrects the image on the basis of this estimate to generate a new rendering of the scene as if it was taken under a reference illuminant [5]. A large portion of the scientific literature on computational color constancy applies to raw RGB image data [6–10], producing as output RGB illuminant estimations to be typically applied with a von Kries-like transform [11] in the form of a diagonal 3×3 matrix in the device's raw acquisition space. In this work, we address the problem of RGB color constancy by exploiting richer multispectral input image data, motivated by the ever-increasing availability of multispectral imaging devices [12–23], which are applied in a variety of computer vision, remote sensing, and medical imaging applications [24,25]. Multispectral reconstruction methods [26,27] are also gaining traction by the scientific community, further incentivizing working in this domain. However, spectral reconstruction is not yet considered to be a mature field, with several open issues [28,29]. In order to investigate the practical utility of multispectral information, we develop our research under the assumption of data that either come from a multispectral acquisition device, or from a hypothetical perfect multispectral reconstruction method. Khan *et al*. [30,31] demonstrated the advantages of estimating and processing illuminants in the multispectral domain. Taking this work as inspiration and as a starting point, our final color correction takes place in the RGB domain instead, since this is the underlying color model of many viewing devices for end-user consumption, and the von Kries-like transform is among the most basic and widely supported correction models in color imaging pipelines and in color management systems [32,33].

Color constancy is formulated as an ill-posed problem: in its most common RGB configuration, it requires to decompose each input RGB image pixel into the contribution of an RGB illuminant and an RGB of the surface. There is no unique solution to a problem formulated in this way (unless additional constraints are put in place), since different combinations of illuminant and surface would lead to the same observed RGB triplet. The same ambiguity holds true in the case of multispectral data, where illuminant and reflectance must be separated from the observed radiance. In this case, however, the higher cardinality of the input data (N channels, with $N \gg 3$) reduces the number of possible combinations of illuminant and surface that would lead to the same radiance, if one assumes a limited set of plausible illuminant sources. Although such a limit is not explicitly formulated in our work, it may be implicitly modeled

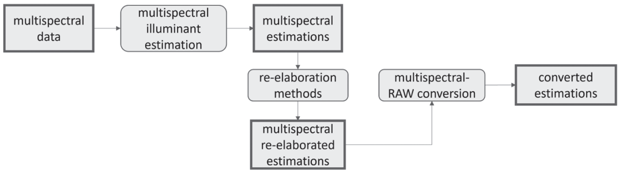


Fig. 1. Pipeline used to retrieve the multispectral estimations and their re-elaborated version starting from the multispectral data provided by the dataset.

by data-driven re-elaboration strategies, which exploit biases from the observed training set. Our work is consequently posed as an investigation to verify whether multispectral information is useful to improve traditional RGB color constancy methods. To this end, we investigate how to extend a selected number of sensor-independent color constancy methods from the RGB domain to an N-dimensional multispectral domain. Showing that this extension is not sufficient to achieve computational color constancy, we then convert the resulting multispectral estimations to the RGB domain: under the assumption that re-elaborating the multispectral illuminant estimation may improve the raw RGB converted result, we investigate several re-elaboration methods. The proposed method for color constancy using multispectral pixel information reaches an improvement of 60% in mean reproduction error on the NUS dataset by Nguyen *et al*. [12], when compared to the corresponding RGB methods. In Fig. 1 the pipeline for the extraction and the re-elaboration of the multispectral and raw RGB estimations is shown.

2. RELATED WORKS

Computational color constancy is a deeply explored field. A large section of the scientific literature approaches the problem in the RGB domain, for which we present here a synthetic general overview. Given our focus on the use of multispectral information, we then specifically address works that connect spectral imaging (including multispectral and hyperspectral) with illuminant estimation.

A. RGB Illuminant Estimation

Over the decades numerous methods [6–10] have been proposed for RGB illuminant estimation; however, no unique solution has been identified. Due to the ill-posed nature of the problem, in fact, color constancy requires the formulation of specific assumptions on the imaged content or the reliance on data-driven biases. To this extent, different methods adopt different strategies. Color constancy methods can be mainly classified into statistical and machine learning methods. Statistical methods [6,8] estimate the illuminant of the image by making assumptions about the color features of the image itself. For example, the max-RGB algorithm assumes the presence of a white surface object in the scene [34]. These methods have fast execution times; however, the resulting performance is heavily dependent from the underlying assumptions. Machine learning algorithms [7,9,10], instead, establish the relationship between the image color distribution and the illuminant through a supervised learning process, meaning they do not need to rely on statistical assumptions and therefore they are more adaptive. However, one of the shortcomings of supervised algorithms is that they are highly dependent on the dataset used to train them [2]. Potentially, they may need to be re-trained on different datasets to overcome data-related bias.

B. Spectral Illuminant Estimation

The use of spectral imaging techniques increased in the last years and their involvement has proven to be beneficial for several fields related to computer vision. Spectral imaging can be categorized in multispectral and hyperspectral imaging. The main difference resides in the spectral resolution; in fact, multispectral imaging captures a small and limited number of spectral bands, while hyperspectral imaging collects the complete and continuous spectrum.

Lenz *et al*. [35] investigated the tasks of illuminant estimation and color correction with the aid of multispectral representation. Specifically, they approximate the spectral description of the scene pixels with a linear combination of bases from a dataset of known spectra. They then characterize the image through the mode of such combination coefficients, which is assumed to represent the global illuminant change.

Li *et al*. [36] proposed an end-to-end unrolling network architecture to estimate both single and multiple illuminants in the input image, casting the problem as a constrained matrix factorization. They also constructed a large spectral image dataset for training and evaluation. Zheng *et al*. [37] modeled the illumination and reflectance spectra separation problem into a low-rank matrix factorization and proved that the illumination–reflectance separation is unique up to an unknown scale under the assumption that reflectance can be approximated with a low-dimensional model. Su *et al*. [38] proposed a general framework to estimate the spectrum of the illumination from specular information in a single hyperspectral image. By utilizing a specular independent subspace they separated the reflectance components and shaped a weighting scheme in order to find specular-contaminated pixels so that the illumination can be directly estimated by factorizing them. Robles-Kelly and Wei [39] presented a convolutional neural network to recover pixel-wise illuminant in multispectral images. The network takes in input a tensor that is constructed by making use of an image patch at different scales in order to allow the network to predict the pixel-wise illuminant using locally supported multiscale information. The relationship between illuminant estimation and multispectral imaging was also explored in a recent paper by Kitanovski *et al*. [40], within the context of reflectance estimation for a spectral filter array camera. The authors investigated the estimation of spectral illuminant from sensor-space illuminant, obtained via measurement or illuminant estimation, as support for the subsequent reflectance estimation. Khan *et al*. [30] investigated the use of illuminant estimation algorithms for multispectral imaging systems to overcome the difficulty in the calibration of multispectral devices. They extended the illuminant estimation algorithms from three channels to N channels. In a subsequent work [31] they then developed a spectral adaptation transform to bring the multispectral image data into a canonical representation. Inspired by the works of Khan *et al*. [30,31], we investigate whether multispectral information may improve the raw RGB color constancy problem. Specifically, we analyze the impact that the multispectral estimation on different wavelengths has on the application to the raw RGB domain. We show that not all the information contained in the multispectral estimations is beneficial for the conversion to the raw RGB domain; therefore we design and evaluate several re-elaboration methods to better fit the correct illuminant.

3. PROPOSED METHOD FOR MULTISPECTRAL-BASED RGB ILLUMINANT ESTIMATION

In this section, we describe our method to exploit multispectral information for raw RGB illuminant estimation. To do so we extend a set of camera-independent color constancy algorithms, originally devised for the raw RGB domain, to the multispectral domain. This has been done consistently with what was proposed by Khan *et al* [30]. Multispectral estimation of the illumination should be then mapped into raw RGB for performing color correction. To this end, a straightforward solution for the multispectral-to-raw conversion consists of exploiting the camera sensitivity function to the multispectral illuminant estimation.

Alternatively, we may assume that an unequal contribution of the estimated multispectral illuminant bands may benefit the eventual raw RGB estimation. This process, which from now on we will refer to as “multispectral illuminant estimations re-elaboration” is discussed mainly in Section 3.B. The concept that is shared among the multispectral illuminant estimations re-elaboration methods is that there exists a mapping (e.g., in the form of weights or biases) that, applied to the multispectral estimation, reduces the distance between the converted multispectral-based raw RGB estimation and the expected raw RGB illuminant.

A. Multispectral Illuminant Estimation Algorithms

In this work, we consider six algorithms belonging to the edge-based color constancy framework (EB), introduced in 2007 by van de Weijer *et al*. [6] as a generalization of multiple algorithms based on low-level image statistics. The general equation for estimation of the illuminant, according to this framework, is

$$\left(\int \left|\frac{\partial^n f^\sigma(x)}{\partial x^n}\right|^p \mathrm{d}x\right)^{\frac{1}{p}} = k e^{n,p,\sigma}. \tag{1}$$

This operation is executed on the separate RGB channels:

$$\begin{aligned}\left(\int \left|\frac{\partial^n f^\sigma(x)}{\partial x^n}\right|^p \mathrm{d}x\right)^{\frac{1}{p}} = \Bigg(&\left(\int \left|\frac{\partial^n R^\sigma(x)}{\partial x^n}\right|^p \mathrm{d}x\right)^{\frac{1}{p}},\\ &\times\left(\int \left|\frac{\partial^n G^\sigma(x)}{\partial x^n}\right|^p \mathrm{d}x\right)^{\frac{1}{p}},\\ &\times\left(\int \left|\frac{\partial^n B^\sigma(x)}{\partial x^n}\right|^p \mathrm{d}x\right)^{\frac{1}{p}}\Bigg).\end{aligned} \tag{2}$$

This framework generates different estimations for the illuminant color, based on three variables.

- n identifies the spatial derivatives order, which is typically set between zero (no derivative, as in the case of the gray world algorithm) and two (for second order derivative).
- p is the Minkowski norm, which determines the relative weights of the multiple measurements from which the final illuminant color is estimated. For example, with $p = 1$, the illuminant is derived by an averaging operation over the derivatives of the channels. For $p = \infty$, the illuminant is computed from the maximum of the derivatives in the scene.
- σ denotes the scale of the local measurements, specifying the intensity of the smoothing operation via Gaussian filtering.

The three parameters of these methods (the spatial derivatives order n, the Minkowski norm p, and standard deviation σ) have been set as proposed in [41]:

- gray world (GW): $n = 0, p = 1, \sigma = 0$;
- white point (WP): $n = 0, p = \infty, \sigma = 0$;
- shades of gray (SoG): $n = 0, p = 4, \sigma = 0$;
- general gray world (GGW): $n = 0, p = 9, \sigma = 9$;
- first order gray edge (GE1): $n = 1, p = 1, \sigma = 6$;
- second order gray edge (GE2): $n = 2, p = 1, \sigma = 1$.

We extend the RGB color constancy algorithms described in Eq. (2) to operate on an arbitrary number of dimensions N, so that they can be applied to multispectral images, producing a multispectral illuminant estimation. In the following, we refer to these algorithms as the spectral counterpart of EB algorithms (e.g., gray world becomes spectral gray world).

B. Multispectral and RGB Illuminant Estimations Re-elaboration

Anticipating the experimental results, we note that the sole multispectral extension of raw RGB illuminant estimation algorithms is not sufficient to achieve color constancy. We hypothesize that the multispectral-to-raw conversion may improve the raw RGB estimation by adopting an unequal contribution from the N multispectral bands. Our approach consists of learning an N-channel modifier to apply to the multispectral illuminant estimation, before converting it through the camera sensitivity function. To verify that the improvement derived from the re-elaboration of the multispectral estimation is due to the multispectral information and not to the re-elaboration method itself, we also apply the re-elaboration methods to the raw RGB estimations. For a clearer understanding of the process steps, we show the baseline of re-elaboration on raw RGB illuminant estimation in Fig. 2.

1. Average Multiplicative Weight

Let $\mathrm{IE}_{\mathrm{RAW}} \in \mathbb{R}^3$ be the illuminant estimation in raw RGB for a single image, and $\mathrm{IE}_{\mathrm{MS}} \in \mathbb{R}^N$ be the illuminant

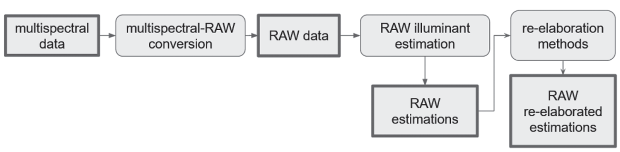


Fig. 2. Pipeline used to retrieve the raw RGB estimations and their re-elaborated version starting from the multispectral data provided by the dataset.

estimation in the multispectral domain. Let $GT_{\{RAW,MS\}}$ be the corresponding ground truth information.

We express the relationship between estimated illuminant (IE) and ground truth illuminant (GT) by means of a multiplicative weight (W):

$$\begin{aligned} IE_{MS} * W_{MS} &= GT_{MS}, \\ IE_{RAW} * W_{RAW} &= GT_{RAW}. \end{aligned} \tag{3}$$

From this relationship, given an image we can define the W factor by dividing the ground truth by the illuminant estimation:

$$\begin{aligned} W_{MS} &= \frac{GT_{MS}}{IE_{MS}}, \\ W_{RAW} &= \frac{GT_{RAW}}{IE_{RAW}}. \end{aligned} \tag{4}$$

Given a training set having cardinality C, we obtain $C \times N$ multispectral weights and $C \times 3$ raw RGB weights, and we subsequently average them along the cardinality dimension. We replicate this process for each one of the six selected color constancy algorithms, estimating in total $6 \times N$ multispectral weights and 6×3 raw RGB weights.

2. Average Additive Bias

The average additive bias method is based on a similar idea to the one of the average multiplicative weight. In this case, instead of a multiplicative weight, we model the re-elaboration through an additive bias (B). The relationship between estimated and ground truth illuminant is expressed as

$$\begin{aligned} IE_{MS} + B_{MS} &= GT_{MS}, \\ IE_{RAW} + B_{RAW} &= GT_{RAW}. \end{aligned} \tag{5}$$

We compute the bias B by simply subtracting the illuminant estimation from the ground truth:

$$\begin{aligned} B_{RAW} &= GT_{RAW} - IE_{RAW}, \\ B_{MS} &= GT_{MS} - IE_{MS}. \end{aligned} \tag{6}$$

As for the previous approach, given a training set of cardinality C, we average the $C \times N$ multispectral biases and the $C \times 3$ raw RGB biases, and we replicate the process for each algorithm, resulting again in $6 \times N$ multispectral biases and 6×3 raw RGB biases. We apply each bias to the testing set estimations of the corresponding algorithm coherently with the relationship expressed in Eq. (6).

3. Optimization-Driven Multiplicative Weight

The search of the weight W from Eq. (3) is here carried out through a direct search method for multidimensional unconstrained minimization [42]. This is obtained by optimizing the end result in terms of raw RGB recovery angular error, between the expected RGB illuminant and the estimated illuminant after the application of the weights:

$$\begin{aligned} W_{MS} &= \operatorname{argmin}_w \{e_{rec}(GT_{RAW}, csf \times (IE_{MS} * w))\}, \\ W_{RAW} &= \operatorname{argmin}_w \{e_{rec}(GT_{RAW}, IE_{RAW} * w\}, \end{aligned} \tag{7}$$

where e_{rec} is the recovery error, *csf* is the camera sensitivity function, w denotes a possible set of weights, and $\times$ is the matrix multiplication. The recovery error [43,44] is defined as

$$e_{rec}(U, V) = \arccos\left(\frac{U \cdot V}{|U||V|}\right), \tag{8}$$

Where "·" indicates the dot product, "||" is the Euclidean norm, and U denotes the actual measured light and V the estimated light by an illuminant estimation algorithm. The direct search method for the multidimensional unconstrained minimization method is known as the Nelder–Mead simplex algorithm [42], for which we use the MATLAB implementation known as "fminsearch." All weights are initialized as ones, indicating no re-elaboration.

4. Feed-Forward Neural Network

For the fourth multispectral illuminant re-elaboration method, we decided to take on a learning-based approach. We designed a topologically simple neural network, based on a feed-forward multilayer perceptron [45], consisting of only three fully connected layers with a sigmoid activation function, as shown in Fig. 3. The first fully connected layer maps the N-band multispectral input to a 60-dimensional latent space, while the last layer maps the result back to N values. Assuming that the input consists of 31-band multispectral data, we have heuristically defined the dimension of the latent space. The multispectral output of the third fully connected layer is converted into raw RGB exploiting the camera sensitivity function. The conversion to raw RGB makes it possible to use as a loss function the recovery error between the raw-converted illuminant estimation and the expected raw illuminant, as defined in Eq. (8).

For sake of comparison, we also apply the same re-elaboration method to the estimated raw RGB illuminants. In this case, the input dimension N is three; we have not modified the dimensionality of the latent space, which is kept to 60. The output of the last fully connected layer is already in raw RGB format, and can be directly used in the computation of the loss function.

4. EXPERIMENTAL SETUP

A. Dataset

The focus of our investigation resides in the use of multispectral imaging to improve raw RGB illuminant estimation; therefore we selected the NUS [12] dataset by Nguyen *et al*. from the

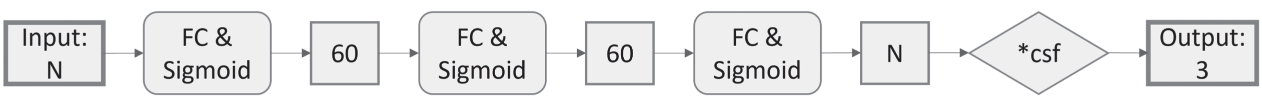


Fig. 3. Feed forward neural network architecture is composed of three hidden layers. The rectangles with round edges indicate the layers used, in this specific case "FC" stands for fully connected layers, which are then followed by a sigmoid activation function. The network re-elaborates the given estimation to better fit the expected illuminant.

National University of Singapore, which contains multispectral data along with the ground truth of their radiance information and the camera sensitivity functions. Having this information it is possible to compute raw RGB data by multiplying the multispectral data by the camera sensitivity functions. The same process of raw RGB computation is also applied to the ground truth multispectral illuminant data. The dataset contains 64 multispectral images along with the corresponding illuminant spectra, which have been acquired using Specim's PFD-CL-65-V10E (400–1000 nm) spectral camera with a Specim OLE23 fore lens. For light sources, the dataset varies from natural sunlight to shade conditions, additionally considering artificial wide-band lights obtained from metal halide lamps of different color temperatures (2500, 3000, 3500, 4300, and 6500 K) and a commercial off-the-shelf LED E400. The subjects of the scenes include both outdoor and indoor images and both natural and man-made objects. Furthermore, a few images of buildings at very long focal lengths were also included. For each spectral image, a total of 31 bands were considered at 400–700 nm, with a spacing of about 10 nm. Of the total 64 multispectral images in the dataset, 24 are reserved for testing and the remaining 40 for training. One or multiple color targets are present in the acquired scenes.

In order to reduce computational complexity, we preprocess the dataset by re-scaling the multispectral images to a tenth of their original size, resulting in 24 testing images of $132 \times W \times 31$, and in 40 training images of $132 \times W \times 31$, where W is the width size that varies between 95 and 237.

5. EXPERIMENTAL RESULTS

This section is divided into four parts. In the first part, we analyze the performance of multispectral extension color constancy algorithms as defined in Section 3.A, compared to the performance of their raw RGB counterpart. The second part is reserved for the assessment of the four re-elaboration techniques extensively discussed in Section 3.B. In order to return a clear idea of how the estimations resulting from these methods and pipelines perform in the color correction step we show some visual examples for a visual comparison. The third part is dedicated to the comparison of the first-section results with two of the main works from the state of the art, namely, those from Khan *et al.* [30] and Robles-Kelly and Wei [39]. In the fourth and final part, we measure the contribution of each input spectral band in our best solution for multispectral-based illuminant estimation, so as to provide a form of model explainability.

To compare the aforementioned methods, we use the reproduction angular error [46], which is defined as follows:

$$e_{\text{rep}}(U, V) = \arccos\left(\frac{\frac{U}{V} \cdot (1, 1, 1)}{|\frac{U}{V}|\sqrt{3}}\right). \tag{9}$$

The rationale is to evaluate the final effect of applying color constancy in the RGB domain, as measured via reproduction error, in addition to the intermediate step of illuminant estimation, as measured via recovery error.

A. Raw RGB versus Multispectral Illuminant Estimation

In this section we assess the performance of multispectral extended color constancy algorithms, compared with the original raw RGB edge-based color constancy algorithms. The results are reported in Table 1. We observe that only three multispectral extended algorithms perform better than their RGB version: spectral white point, spectral first order gray edge, and spectra second order gray edge. While spectral white point mean reproduction error improves only by 0.08° compared to its raw RGB version (corresponding to a 1% improvement), spectral first order gray edge and spectral second order gray edge improve respectively by 0.38° (6%) and 0.65° (9%). Additional statistics are reported in Supplement 1.

A first explanation of why the multispectral extensions of different color constancy algorithms lead to inconsistent variations in performance may be found in the algorithms' linearity. We note, in fact, that the RGB gray world algorithm and the multispectral gray world algorithm produce equivalent results. For the RGB gray world, spectral images are first converted, using linear camera sensitivity functions, to RGB images for the sake of our experiments, and from these, the gray world is applied, which

Table 1. Evaluation of Raw RGB and Multispectral Color Constancy Algorithms Performance Divided by Method[a]

Input	AWB Algorithm	Mean Recovery	Median Recovery	Mean Reproduction	Median Reproduction	% Mean Reproduction Improvement
Multispectral	**GE1**	5.09	4.42	6.38	5.23	6%
Raw RGB		5.46	4.72	6.76	5.67	
Multispectral	**GE2**	5.03	4.49	6.32	5.13	**9%**
Raw RGB		5.55	5.26	6.97	6.29	
Multispectral	**GGW**	3.70	3.15	**4.42**	3.37	−2%
Raw RGB		**3.67**	2.92	4.33	3.09	
Multispectral	**GW**	3.81	**2.55**	**4.42**	**2.91**	0%
Raw RGB		3.81	**2.55**	**4.42**	**2.91**	
Multispectral	**SOG**	3.84	3.15	4.63	3.74	−1%
Raw RGB		3.84	3.16	4.57	3.71	
Multispectral	**WP**	4.68	3.93	5.60	4.99	1%
Raw RGB		4.81	4.29	5.68	5.26	

[a]We report both recovery and reproduction angular errors, expressed in degrees (°). The lower, the better. The last column shows the percentual improvement in mean reproduction angular error of multispectral versus raw RGB algorithms. Best results per metric are highlighted in bold.

consists of equal-weight linear operations. For the multispectral gray world, the two operations are inverted: the algorithm is applied to the multispectral image using equal-weight linear operations, and the result is then brought into RGB using the same linear camera sensitivity functions. The pure linearity of the involved operations leads to equivalent results in the two cases. The other algorithms within the framework, such as the white patch, or the gray edge, display behaviors at a varying degree of non-linearity, which leads to a varying degree of divergence between the RGB and multispectral version.

A second explanation may be formulated as to why some of the multispectral algorithms produce a performance improvement, while others lead to a decrease in the performance. To this extent, we recall that the analyzed RGB color constancy algorithms are based on individual assumptions and hypotheses regarding the input image. For example, the RGB gray world assumes that the average of the observed radiance is achromatic (i.e., all tristimulus responses are on average equivalent), the RGB white patch assumes that the brightest area in the image is a highlight reflecting the scene illumination source, etc. In practice, different images will match such assumptions to different degrees, thus leading to a varying range of performance levels. Similarly, the multispectral extensions of the same algorithms are inherently based on assumptions on the input spectral image, which are different from the assumptions of their RGB counterparts, and as such they lead to a different distribution of errors.

B. Iluminant Estimations Re-elaboration

The illuminant estimation re-elaboration will be assessed method by method, comparing the performance of the re-elaborated raw RGB input and the performance of the re-elaborated multispectral information. As for the previous assessment, the estimations are compared with the reproduction error, and the performance is reported in terms of mean and median errors. Additionally, we show the percent improvement with respect to the traditional raw RGB pipeline from Table 1. The results for all re-elaboration methods are presented in Table 2.

1. Average Multiplicative Weight

All illuminant estimation algorithms benefit from the use of the average multiplicative weight method, both for the multispectral and raw RGB input. Results in Table 2 also show that not only the re-elaboration methods improve the illuminant estimation accuracy but also that the multispectral input improves the performance with respect to the raw RGB input. The best improvement is achieved by multispectral second order gray edge (GE2) but the best performance overall is achieved by multispectral general gray world (GGW) with a 3.69° reproduction error.

2. Average Additive Bias

The only illuminant estimation algorithms that benefit from the average additive bias re-elaboration are: first order gray edge (GE1) and second order gray edge (GE2), even in their multispectral extension.

However, with this re-elaboration, the best performing pipeline is the multispectral general gray world (GGW) with a 5.75° mean reproduction error value, which is still performing worse than the 3.69° achieved by the raw RGB general gray world (which is the best traditional performing algorithm in this investigation).

3. Optimization-Driven Multiplicative Weight

The re-elaboration of the multispectral estimations obtained with the optimization-driven multiplicative weight improves the mean reproduction error value for each algorithm. However, the raw RGB shades of gray (SoG) estimations actually achieve the best result for the optimization-driven multiplicative weight method with a mean reproduction error of 3.5°, which improves the value for that metric with respect to the non-re-elaborated raw RGB estimations by 23%, while the best improvement is achieved by multispectral second order gray edge (GE2) with a 38% improvement compared to the traditional raw RGB estimation.

4. Feed-Forward Neural Network

The performance of the feed-forward re-elaboration method led to more noticeable improvements in processing multispectral information with respect to traditional raw RGB data, as can be easily appreciated from Table 2. The mean reproduction error for the multispectral inputs ranges from 3.15° for the first order gray edge (GE1) down to 2.58° for the spectral white point (WP), achieving the best performance overall in our analysis. The white point (WP) re-elaboration method improves by 55% compared with the traditional raw RGB method for the same illuminant estimation algorithm.

Figure 4 offers a visual representation of the effect of color constancy using our proposed method in different configurations, including raw RGB or multispectral input, and the four re-elaboration methods.

C. Comparison with the State of the Art

Among the methods in the state of the art, Robles-Kelly and Wei [39] and Khan *et al*. [30] are the most similar ones to our proposed method in terms of approach and final goal, i.e., RGB color constancy by exploiting multispectral information.

Robles-Kelly presented a method that employs a convolutional neural network to estimate pixel-wise illuminant in the scene for both trichromatic and spectral images. Khan *et al*. [30] proposed to extend statistical illuminant estimation methods (applied also here, and described in Section 3.A) to N dimensions, and subsequently developed a spectral adaptation transform to bring the multispectral image data into a canonical, or target, multispectral representation [31]. In order to enable a direct comparison with our method it has been necessary to apply a consensus-based strategy for raw RGB illuminant estimation. The strategy consists of:

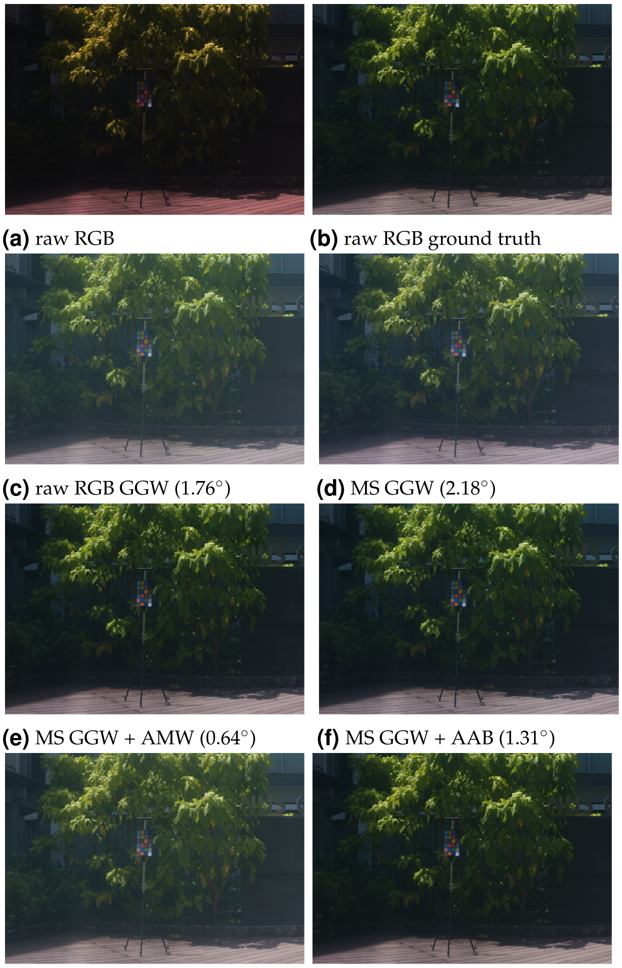
1. converting the input multispectral radiance image into the RGB domain, to obtain a raw RGB image that is not white-balanced;

Table 2. Evaluation of the Re-elaboration of the Multispectral and Raw RGB Illuminant Estimations[a]

Input	Re-elaboration Method	AWB Algorithm	Mean Recovery	Median Recovery	Mean Reproduction	Median Reproduction	% Improvement Mean Reproduction Error
Raw RGB	AAB	GE1	5.28	4.07	6.46	4.75	4%
Raw RGB	AAB	GE2	5.31	4.20	6.49	4.91	7%
Raw RGB	AAB	GGW	5.32	4.39	6.50	5.25	−50%
Raw RGB	AAB	GW	5.30	4.51	6.48	5.50	−47%
Raw RGB	AAB	SOG	5.38	4.24	6.58	5.08	−44%
Raw RGB	AAB	WP	5.32	4.02	6.51	4.82	−15%
Multispectral	AAB	GE1	4.78	3.63	5.91	4.27	13%
Multispectral	AAB	GE2	4.88	3.93	6.01	4.57	14%
Multispectral	AAB	GGW	4.64	3.60	5.75	4.39	−33%
Multispectral	AAB	GW	4.66	3.73	5.77	4.52	−31%
Multispectral	AAB	SOG	4.76	3.58	5.89	4.46	−29%
Multispectral	AAB	WP	4.66	3.35	5.78	4.10	−2%
Raw RGB	AMW	GE1	4.15	3.36	4.79	4.42	29%
Raw RGB	AMW	GE2	4.16	3.98	4.84	4.30	31%
Raw RGB	AMW	GGW	3.42	2.21	3.82	2.55	12%
Raw RGB	AMW	GW	3.61	3.08	3.95	3.31	11%
Raw RGB	AMW	SOG	3.44	2.28	3.84	2.54	16%
Raw RGB	AMW	WP	5.12	4.79	5.57	5.24	2%
Multispectral	AMW	GE1	3.82	2.88	4.51	3.70	33%
Multispectral	AMW	GE2	3.84	3.30	4.54	3.82	35%
Multispectral	AMW	GGW	3.26	2.23	3.69	2.34	15%
Multispectral	AMW	GW	3.62	2.86	4.03	3.28	9%
Multispectral	AMW	SOG	3.28	2.22	3.73	2.60	18%
Multispectral	AMW	WP	4.86	4.44	5.33	5.00	6%
Raw RGB	ODMW	GE1	4.29	3.46	5.25	4.36	22%
Raw RGB	ODMW	GE2	3.95	3.54	4.80	4.16	31%
Raw RGB	ODMW	GGW	3.27	2.87	3.81	3.45	12%
Raw RGB	ODMW	GW	3.71	3.48	4.36	4.08	1%
Raw RGB	ODMW	SOG	3.00	1.94	3.50	2.40	23%
Raw RGB	ODMW	WP	4.22	3.97	4.84	4.46	15%
Multispectral	ODMW	GE1	4.20	3.41	5.15	4.07	24%
Multispectral	ODMW	GE2	3.62	3.04	4.35	3.61	38%
Multispectral	ODMW	GGW	3.15	2.25	3.59	2.31	17%
Multispectral	ODMW	GW	3.52	2.85	3.99	3.17	10%
Multispectral	ODMW	SOG	3.14	2.18	3.65	2.58	20%
Multispectral	ODMW	WP	4.91	4.01	5.61	4.86	1%
Raw RGB	FFNN	GE1	3.93	3.76	5.13	5.40	24%
Raw RGB	FFNN	GE2	3.05	2.16	3.81	3.19	45%
Raw RGB	FFNN	GGW	3.75	3.69	4.48	3.76	−3%
Raw RGB	FFNN	GW	3.44	2.40	4.02	2.68	9%
Raw RGB	FFNN	SOG	3.74	2.79	4.62	3.86	−1%
Raw RGB	FFNN	WP	5.09	2.75	6.20	3.19	−9%
Multispectral	FFNN	GE1	2.48	2.18	3.15	2.63	53%
Multispectral	FFNN	GE2	2.24	1.33	2.77	1.93	**60%**
Multispectral	FFNN	GGW	2.24	1.43	2.71	1.55	37%
Multispectral	FFNN	GW	2.47	1.53	2.91	1.63	34%
Multispectral	FFNN	SOG	2.26	1.11	2.71	1.43	41%
Multispectral	FFNN	WP	**2.13**	**1.04**	**2.58**	**1.28**	55%

[a]All values are expressed in degrees (°); the lower, the better. The last column shows the percentage of improvement of the mean reproduction angular error, between the selected method and the traditional illuminant estimation algorithm for the raw RGB input.

2. dividing the same input multispectral radiance image by the estimated multispectral illuminant to obtain a multispectral reflectance image;
3. multiplying the obtained multispectral reflectance image for the target multispectral illuminant, to obtain a new multispectral radiance image;
4. converting the multispectral image restulting from step 3 into the RGB domain, to obtain a raw RGB image that is white-balanced;
5. obtaining a per-pixel RGB illuminant estimation by dividing the result of point 4 by the result of point 1;



(a) raw RGB **(b)** raw RGB ground truth
(c) raw RGB GGW (1.76°) **(d)** MS GGW (2.18°)
(e) MS GGW + AMW (0.64°) **(f)** MS GGW + AAB (1.31°)
(g) MS GGW + ODMW (0.25°) **(h)** MS GGW + FFNN (0.19°)

Fig. 4. Visual example of some of the most significant methods. (a) Acquired scene; (b) scene corrected with the expected illuminant in raw RGB; (c)–(h) scene corrected with the illuminant estimated via general gray world (GGW) on either raw RGB or multispectral (MS) input, using different re-elaboration methods. For all corrections, angular errors are reported in parentheses (the lower, the better).

6. generating a global raw RGB illuminant estimation of the input image by consensus through average per channel.

In Table 3 a comparison of the results of the previously cited methods is shown against our solution. We consider two baselines of our method without re-elaboration (based on raw RGB and multispectral data), using general gray world as a reference due to its optimal performance as reported in Table 1. We then also consider our best performing configuration, using multispectral white point and FFNN re-elaboration as reported in Table 2. The comparison is performed in terms of mean and median recovery angular error in order to allow for a direct comparison with the results reported by Robles-Kelly *et al.* [39]. The comparison between our own method's baselines and existing methods from the state of the art highlights that a simple approach without re-elaboration achieves similar performance as the solution by Khan *et al*. Additionally, our complete method based on output re-elaboration allows us to achieve superior performance in raw RGB illuminant estimation.

Table 3. Mean and Median Recovery Angular Error (in Degrees °; the Lower, the Better) for Khan's, Robles-Kelly's, and Our Method on NUS Dataset [12][a]

Method	Mean Recovery Angular Error	Median Recovery Angular Error
Robles-Kelly and Wei [39]	12.56	4.62
Khan *et al*. (multispectral GGW) [31]	3.96	2.94
Our baseline (raw GGW)	3.67	2.92
Our baseline (multispectral GGW)	3.70	3.15
Our best method (multispectral WP + FFNN)	**2.13**	**1.04**

[a]In bold the best results.

D. Further Analysis

In order to further study our best model for illuminant estimation from multispectral data, we investigate the relationship between input and output wavelength bands as a form of model explainability. Specifically, we measure the relevance of each input spectral band i by selectively feeding to our trained feedforward neural network FF a set of band-specific impulses (setting band i to one, and all the other bands to zero). We assess the absolute difference in each of the network's output bands j compared to the average network's output A:

$$\text{rel}_i = \sum_j |FF\big(\text{impulse}(i)\big)_j - A_j|, \tag{10}$$

$$A_j = \frac{1}{N}\sum_i FF\big(\text{impulse}(i)\big)_j. \tag{11}$$

The result of this analysis is reported in Fig. 5. By supporting the visualization with the camera sensitivity function curves used in the optimization process, three band clusters emerge, roughly corresponding to the central sections of the camera's color filters, with local minima corresponding to the overlap between two color channels, where information is partially redundant (the estimate on the green channel is partially informed by the information from the blue and red channels). Furthermore, the model presents two outlying peaks and a generally oscillating behavior.

These observations on band relevance could potentially inform the definition of feature reduction techniques for hardware optimization, where fewer and selected wavelength bands are considered in the construction of a multispectral sensor.

6. CONCLUSIONS

We have conducted an investigation to assess whether multispectral information can be beneficial for the raw RGB color constancy problem. We have separated the work into two main steps. (1) In the first step we evaluated the multispectral illuminant estimations and we compared the results with the traditional raw RGB illuminant estimations. (2) For the second step we suggested to re-elaborate multispectral estimations to better fit the expected raw RGB illuminant. To serve this purpose we suggested four re-elaboration methods and we evaluated

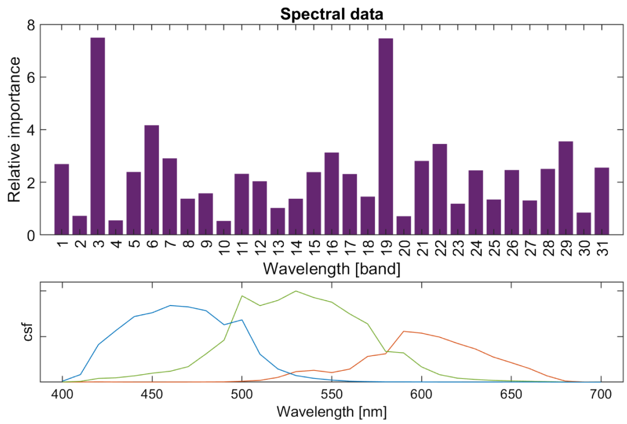


Fig. 5. Relative importance of wavelength bands for our neural model for illuminant estimation from multispectral data. Camera sensitivity functions reported for reference.

them, not only comparing them to the traditional raw RGB approach but also with raw RGB re-elaborated performance.

We proved that multispectral information can be used to improve raw RGB color constancy. In fact, we have shown that some methods (first order gray edge, second order gray edge, and white point) improve with our multispectral-based methods. Results show that re-elaboration methods improve performance both for multispectral and raw RGB illuminant estimation, respectively, with an overall performance increment of 60% and 50%, for the mean reproduction angular error, with respect to the traditional raw RGB pipeline. Of great relevance is the result achieved by the multispectral white point with feed-forward neural network re-elaboration, that achieves a mean recovery error of 2.13°.

In the future, we plan to extend the work to other illuminant estimation algorithms, especially to machine-learning-based algorithms.

Funding. Ministero dell'Università e della Ricerca; National Recovery and Resilience Plan.

Acknowledgment. This work was partially supported by the MUR under the grant "Dipartimenti di Eccellenza 2023-2027" of the Department of Informatics, Systems and Communication of the University of Milano-Bicocca, Italy. This work was partially funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3 - Call for tender No. 341 of 15 March 2022 of Italian Ministry of University and Research funded by the European Union – NextGenerationEU; Award Number: Project code PE00000003, Concession Decree No. 1550 of 11 October 2022 adopted by the Italian Ministry of University and Research, CUP D93C22000890001, Project title "ON Foods - Research and innovation network on food and nutrition Sustainability, Safety and Security – Working ON Foods".

Disclosures. The authors declare no conflicts of interest.

Data availability. The work and results presented in this paper are based on the original NUS dataset by Nguyen *et al*. [12].

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. M. D. Fairchild, *Color Appearance Models* (Wiley, 2013).
2. M. Buzzelli, S. Zini, S. Bianco, *et al.*, "Analysis of biases in automatic white balance datasets and methods," Color Res. Appl. **48**, 40–62 (2023).
3. G. D. Finlayson, "Colour and illumination in computer vision," Interface Focus **8**, 20180008 (2018).
4. M. Ebner, "Color constancy," in *Computer Vision: A Reference Guide* (Springer, 2021), pp. 168–175.
5. S. Bianco, C. Cusano, and R. Schettini, "Single and multiple illuminant estimation using convolutional neural networks," IEEE Trans. Image Process. **26**, 4347–4362 (2017).
6. J. Van De Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," IEEE Trans. Image Process. **16**, 2207–2214 (2007).
7. Y. Hu, B. Wang, and S. Lin, "Fc4: fully convolutional color constancy with confidence-weighted pooling," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4085–4094.
8. 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**, 1049–1058 (2014).
9. M. Afifi and M. S. Brown, "Sensor-independent illumination estimation for DNN models," arXiv, arXiv:1912.06888 (2019).
10. S. Bianco and C. Cusano, "Quasi-unsupervised color constancy," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 12212–12221.
11. J. von Kries, *Theoretische Studien über die Umstimmung des Sehorgans* (Festschrift der Albrecht-Ludwigs-Universität, 1902), pp. 145–158.
12. R. M. Nguyen, D. K. Prasad, and M. S. Brown, "Training-based spectral reconstruction from a single RGB image," in *European Conference on Computer Vision* (Springer, 2014), pp. 186–201.
13. S. M. Nascimento, F. P. Ferreira, and D. H. Foster, "Statistics of spatial cone-excitation ratios in natural scenes," J. Opt. Soc. Am. A **19**, 1484–1490 (2002).
14. S. Hordley, G. Finalyson, and P. Morovic, "A multi-spectral image database and its application to image rendering across illumination," in *International Conference on Image and Graphics* (IEEE, 2004), pp. 394–397.
15. D. H. Foster, K. Amano, S. M. Nascimento, *et al.*, "Frequency of metamerism in natural scenes," J. Opt. Soc. Am. A **23**, 2359–2372 (2006).
16. M. Parmar, F. Imai, S. H. Park, *et al.*, "A database of high dynamic range visible and near-infrared multispectral images," Proc. SPIE **6817**, 68170N (2008).
17. F. Yasuma, T. Mitsunaga, D. Iso, *et al.*, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," IEEE Trans. Image Process. **19**, 2241–2253 (2010).
18. A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Conference on Computer Vision and Pattern Recognition* (IEEE, 2011), pp. 193–200.
19. T. Skauli and J. Farrell, "A collection of hyperspectral images for imaging systems research," Proc. SPIE **8660**, 86600C (2013).
20. S. M. Nascimento, K. Amano, and D. H. Foster, "Spatial distributions of local illumination color in natural scenes," Elsevier Vision Res. **120**, 39–44 (2016).
21. B. Arad, O. Ben-Shahar, and R. Timofte, "Ntire 2018 challenge on spectral reconstruction from RGB images," in *Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, 2018), pp. 929–938.
22. B. Arad, R. Timofte, O. Ben-Shahar, *et al.*, "Ntire 2020 challenge on spectral reconstruction from an RGB image," in *Conference on Computer Vision and Pattern Recognition Workshops* (IEEE/CVF, 2020), pp. 446–447.
23. B. Arad, R. Timofte, R. Yahel, *et al.*, "Ntire 2022 spectral recovery challenge and data set," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 863–881.
24. Y. Manabe, K. Sato, and S. Inokuchi, "An object recognition through continuous spectral images," in *Proceedings of 12th International Conference on Pattern Recognition* (IEEE, 1994), Vol. **1**, pp. 858–860.
25. C. Liu, W. Liu, X. Lu, *et al.*, "Application of multispectral imaging to determine quality attributes and ripeness stage in strawberry fruit," PLoS ONE **9**, e87818 (2014).

26. M. Agarla, S. Bianco, M. Buzzelli, *et al.*, "Fast-n-squeeze: towards real-time spectral reconstruction from RGB images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 1132–1139.
27. Y.-T. Lin and G. D. Finlayson, "A rehabilitation of pixel-based spectral reconstruction from RGB images," Sensors **23**, 4155 (2023).
28. Y.-T. Lin and G. D. Finlayson, "An investigation on worst-case spectral reconstruction from RGB images via radiance Mondrian world assumption," Color Res. Appl. **48**, 230–242 (2023).
29. Y.-T. Lin and G. D. Finlayson, "Investigating the upper-bound performance of sparse-coding-based spectral reconstruction from RGB images," in *Color and Imaging Conference* (2021), Vol. **29**, pp. 19–24.
30. H. A. Khan, J.-B. Thomas, J. Y. Hardeberg, *et al.*, "Illuminant estimation in multispectral imaging," J. Opt. Soc. Am. A **34**, 1085–1098 (2017).
31. H. A. Khan, J.-B. Thomas, J. Y. Hardeberg, *et al.*, "Spectral adaptation transform for multispectral constancy," J. Imaging Sci. Technol. **62**, 20504 (2018).
32. M. Delbracio, D. Kelly, M. S. Brown, *et al.*, "Mobile computational photography: a tour," arXiv, arXiv:2102.09000 (2021).
33. S. Bianco, A. R. Bruna, F. Naccari, *et al.*, "Color correction pipeline optimization for digital cameras," J. Electron. Imaging **22**, 023014 (2013).
34. E. H. Land and J. J. McCann, "Lightness and retinex theory," J. Opt. Soc. Am. A **61**, 1–11 (1971).
35. R. Lenz, P. Meer, and M. Hauta-Kasari, "Spectral-based illumination estimation and color correction," Color Res. Appl. **24**, 98–111 (1999).
36. Y. Li, Q. Fu, and W. Heidrich, "Multispectral illumination estimation using deep unrolling network," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 2672–2681.
37. Y. Zheng, I. Sato, and Y. Sato, "Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 1779–1787.
38. T. Su, Y. Zhou, Y. Yu, *et al.*, "Illumination separation of non-Lambertian scenes from a single hyperspectral image," Opt. Express **26**, 26167–26178 (2018).
39. A. Robles-Kelly and R. Wei, "A convolutional neural network for pixelwise illuminant recovery in colour and spectral images," in *24th International Conference on Pattern Recognition (ICPR)* (IEEE, 2018), pp. 109–114.
40. V. Kitanovski, J.-B. Thomas, and J. Yngve Hardeberg, "Reflectance estimation from snapshot multispectral images captured under unknown illumination," in *Color and Imaging Conference 29* (Society for Imaging Science and Technology, 2021), pp. 264–269.
41. S. Bianco, C. Cusano, and R. Schettini, "Color constancy using CNNS," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2015), pp. 81–89.
42. J. C. Lagarias, J. A. Reeds, M. H. Wright, *et al.*, "Convergence properties of the Nelder–Mead simplex method in low dimensions," SIAM J. Optim. **9**, 112–147 (1998).
43. S. D. Hordley and G. D. Finlayson, "Reevaluation of color constancy algorithm performance," J. Opt. Soc. Am. A **23**, 1008–1020 (2006).
44. A. Gijsenij, T. Gevers, and M. P. Lucassen, "Perceptual analysis of distance measures for color constancy algorithms," J. Opt. Soc. Am. A **26**, 2243–2256 (2009).
45. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
46. G. D. Finlayson and R. Zakizadeh, "Reproduction angular error: an improved performance metric for illuminant estimation," Perception **310**, 1–26 (2014).