



# Single Pixel Spectral Color Constancy

Samu Koskinen<sup>1,2</sup> · Erman Acar<sup>1</sup> · Joni-Kristian Kämäräinen<sup>1,2</sup>

Received: 31 July 2022 / Accepted: 26 July 2023  
© The Author(s) 2023

## Abstract

Color constancy is still one of the biggest challenges in camera color processing. Convolutional neural networks have been able to improve the situation but there are still problems in many conditions, especially in scenes where a single color is dominating. In this work, we approach the problem from a slightly different setting. What if we could have some other information than the raw RGB image data. What kind of information would help to bring significant improvements while still be feasible in a mobile device. These questions sparked an idea for a novel approach for computational color constancy. Instead of raw RGB images used by the existing algorithms to estimate the scene white points, our approach is based on the scene's average color spectra-single pixel spectral measurement. We show that as few as 10–14 spectral channels are sufficient. Notably, the sensor output has five orders of magnitude less data than in raw RGB images of a 10MPix camera. The spectral sensor captures the “spectral fingerprints” of different light sources and the illuminant white point can be accurately estimated by a standard regressor. The regressor can be trained with generated measurements using the existing RGB color constancy datasets. For this purpose, we propose a spectral data generation pipeline that can be used if the dataset camera model is known and thus its spectral characterization can be obtained. To verify the results with real data, we collected a real spectral dataset with a commercial spectrometer. On all datasets the proposed Single Pixel Spectral Color Constancy obtains the highest accuracy in the both single and cross-dataset experiments. The method is particularly effective for the difficult scenes for which the average improvements are 40–70% compared to state-of-the-arts. The approach can be extended to multi-illuminant case for which the experimental results also provide promising results.

**Keywords** Color constancy · Spectral sensing · Illuminant estimation

## 1 Introduction

A well working color constancy (CC) algorithm is a key component in the camera color processing pipelines. Color constancy is obtained by algorithms that estimate the illuminant white point from captured images. There are static methods that are based on physical or statistical properties of scenes (Yang et al. , 2015; Qian et al. , 2019) and learning-based methods that learn white point mapping from training

data (Barron, 2015; Barron & Tsai, 2017; Hu et al. , 2017). While the color constancy has been studied for a long time, the problem is not fully solved. Even the best algorithms may fail, for example, when the scene is dominated by a single color.

In this work, we propose a novel approach for computational color constancy. In our approach we replace the raw RGB images used by the existing methods with *average color spectra* of captured scenes. Such spectral sensors are already available in the high-end mobile phones. For example, Huawei P40 Pro is equipped with an 8-channel average spectral sensor. It is noteworthy that average spectral measurements completely lack the spatial dimension, but the spectral domain information captures *spectral fingerprints* of illuminants and thus the illuminant white point can be estimated by a simple regression.

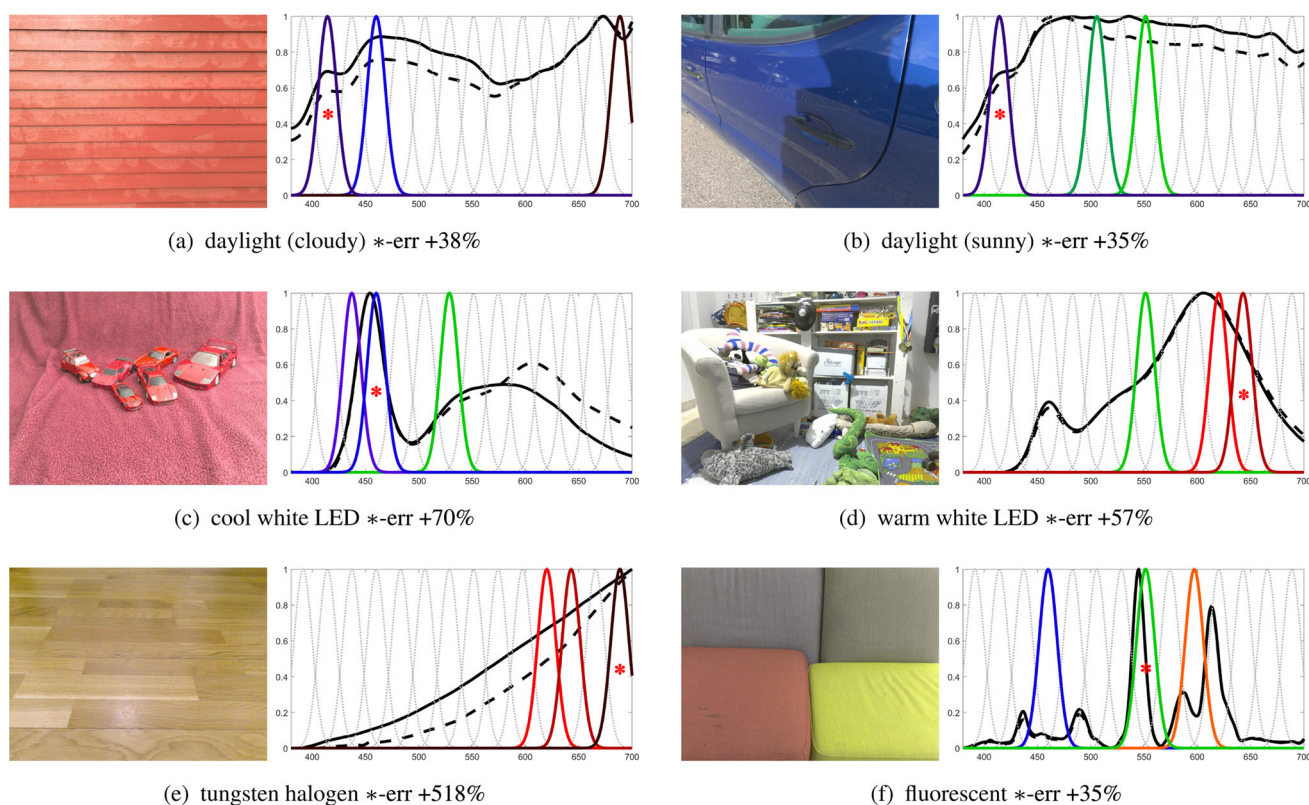
The core idea of spectral fingerprints is illustrated in Fig. 1. The typical light sources such as a daylight, fluorescent, LED and tungsten are recognizable by the shapes of their power

Communicated by Stuart James.

✉ Samu Koskinen  
samu.koskinen@tuni.fi  
Erman Acar  
erman.acar@huawei.com  
Joni-Kristian Kämäräinen  
joni.kamarainen@tuni.fi

<sup>1</sup> Huawei Technologies Oy (Finland) Co. Ltd, Tampere, Finland

<sup>2</sup> Vision Group, Tampere University, Tampere, Finland



**Fig. 1** Real spectral dataset examples. Solid black line denotes the light source power spectrum (ground truth) and the dashed is the measured average reflected spectrum. Gray dotted lines are the 14 spectral channels used in our experiments. For each image the three most important

channels found by leave-one-out are colored with their corresponding wavelength and the most important denoted by an asterisk (percentage numbers denote the increase in angular error if this is removed compared to the second most important channel)

spectra.<sup>1</sup> The claim can be validated by taking a spectral white point regressor trained with all channels and testing it on unseen images and switching off each channel one by one, i.e. a spectral channel was set to zero for the separate runs without re-training the model. After running all the combinations, we compared the results to the reference where all channels were used normally. The increase of the error when a channel was set to zero will indicate the importance of the channel for the given test image. The most important channel(s) should be characteristic for each light source. The results for the MLP regressor in Sect. 3 are shown in Fig. 1 for various scenes and light sources. The reflected spectra are not that different from the ground truth illuminant spectrums even though the scenes are very chromatic in several of the illustrations. For example, for the both daylight cases the most important channel is the same around 415nm even though the color content of the scenes are very different. That

wavelength contains a characteristic bump of the daylight spectrum. For tungsten halogen the important wavelengths are in the near-infrared region. That is characteristic to tungsten sources which have substantial amount of IR energy as compared to their visible light region. The LED case illustrates the practicality of the illuminant fingerprints to identify specific spectral peaks. The blue die peak in the cool white LED case is captured with the most important channel and the other important channels record more information about the blue peak and the phosphor bump. Other channels are clearly less meaningful in the LED case. The warm white LED has so much more power in the yellow area that the important channels are focused there. It is a similar and equally intuitive case with the fluorescent spectrum too.

Our main contribution is

- (1) the novel approach for computational color constancy using average color spectrum. In addition, we propose
- (2) a method to generate spectral data from the existing tristimulus (RGB) color constancy datasets for training purposes and
- (3) simulation based analysis of optimal spectral sensor design.

<sup>1</sup> The power spectra of CIE standard illuminants and various white LEDs can be found from technical notes and other publications (Dietz, 2011; Kokka et al. , 2018).

In all experiments our method obtains lower average angular error than the existing RGB based methods and it is noteworthy that the results are better in cross-dataset experiments where our method is trained with generated data but tested with real data.

This work is an extended version of our recent paper (Koskinen et al., 2021). This work addresses the important additional research and design questions that were not addressed in the preliminary work. As the first extension, (4) we study the performance upper bound if the practical 14 channel sensor is replaced with a dedicated 65 channel sensor that corresponds to commercial spectrometers. As the second extension, (5) we study the difficult but important multi-illuminant case which is assumed to be difficult for our single pixel sensor without spatial information. Finally, this work includes (6) a more detailed description of image data augmentation for training spectral color constancy with limited samples, and additional visualizations and illustrations of the approach and its experimental results.

## 2 Related Work

Color constancy algorithms estimate the illuminant  $L$  in order to recover the scene  $R$  under the white light. In the conventional setting  $L$  is estimated from the raw RGB image  $I$ . The existing algorithms can be divided into *learning-free* (static) and *learning-based* methods. Classical learning-free methods are based on image statistics in the RGB color space in order to find the illuminant white point. The most common such algorithm is a gray world algorithm (Buchsbbaum, 1980) which assumes that the image chromaticity is gray on average. That assumption works in scenarios where there are a lot of color variation in the scenery. Extended versions of the grey world algorithm are a max RGB (Barnard et al., 2002) and a gray edge (van de Weijer & Gevers, 2005) algorithms. They assume that it is more likely to find achromatic content in an image if you only consider certain areas of the image, like regions near edges (gray edge) and around the maximum value (max RGB) of the image. The updated versions of the algorithms can also apply weights for each pixel based on its spatial statistics like the pixel's gradient or relative brightness. The classical methods can work well in fairly many cases, but they are very inefficient in the challenging conditions, such as when the scenery is dominated by a single chromatic color.

In the recent evaluations on multiple datasets (Qian et al., 2019; Keshav & GVSL, 2019) the best performing learning-free algorithms are *Gray Index* (GI) (Qian et al., 2019), *Local Surface Reflectance Statistics* (LSRS) (Gao et al., 2014), and Cheng et al. (2014) and the best performing learning-based are *Decoupled Semantic Context and Color Correlation* (DSCCC) (Keshav & GVSL, 2019), *Fast*

*Fourier Color Constancy* (FFCC) (Barron & Tsai, 2017) and *Fully Convolutional Color Constancy with Confidence* (FC<sup>4</sup>) (Keshav & GVSL, 2019). The best method varies between the datasets and depending on whether the evaluation is single or cross-dataset, but in overall the differences are small.

There are a few works that study color constancy for (multi)spectral images. For example, Gevers et al. (2000) use spectral sensing for color constancy assuming that a white reference is available in the scene. Chakrabarti et al. (2011) model color constancy via spatio-spectral statistics similar to conventional RGB white balance algorithms. Khan et al. (2017) also extend traditional color constancy algorithms to multispectral images with varying spectral resolutions. These works assume that a full spatial spectral image is available, but compact high resolution spectral cameras are difficult to manufacture. Work done by Chen (2017) studies how the Corrected-Moments algorithm (Finlayson, 2013) can be extended and improved when applied for multispectral images. Spectral sharpening by Finlayson et al. (1994) aims to improve color constancy with the help of spectral sensing. Hui et al. have studied an illuminant source separation task for which they utilize spectral data (Hui et al., 2018, 2019). Their training data generation in the former paper is physics based and use pre-defined databases for illuminant and reflectance spectra. They also weight their spectral estimation according to a camera spectral response.

Research on spectral measurements is timely as new technological advances make it possible to manufacture miniaturized multispectral sensors. The recent works of Jensen (2020) and Wang et al. (2019) investigate practical implementations of portable spectral sensors.

## 3 Methods

Spectral sensors can be expressed mathematically in a similar way as the RGB sensors of digital cameras. Formation of a raw RGB image  $I$  of a scene  $R$  with the camera  $C$  of known spectral sensitivities  $S_{i=R,G,B}$  and under a global illumination  $L$  can be expressed as (von Kries, 1970)

$$I_i(x, y) = \int L(\lambda) S_i(x, y, \lambda) R(x, y, \lambda) d\lambda, \quad (1)$$

$$i \in \{R, G, B\}$$

where  $S_i(x, y, \lambda)$  denote the spectral sensitivity of the Red, Green and Blue elements:  $i = \{R, G, B\}$ .  $\lambda$  is the spectral wavelength that for human perceivable colors is 380–700 nanometers (nm). Below 380nm is the ultra-violet band and above 700nm is the infra-red band.

The RGB sensors are designed to capture photographs that match the color sensitive cells of the human visual system

(HVS) (Palmer, 1999). However, for accurate color measurements the HVS-inspired wide-band RGB sensors  $C = C^{RGB}$  produce various problems such as the metamerism. The problems can be largely avoided by spectral imaging with a spectral camera  $C^{spec}$  that has multiple narrowband sensor elements  $S_{i=1,\dots,N}$ . Manufacturing of a spectral camera with a high spatial resolution is difficult as it requires a mechanical filter wheel or a large number of photo receptors for each band (Nathan & Michael, 2013; Gao & Wang, 2016).

### 3.1 Average Spectral Measurement

In this work, we omit the spatial dimension for color constancy. In that case, a spectral camera is not needed. Average spectrum can be measured by a point sensor that needs

1. A wide angle lens or a diffuser that covers the scene on the image plane  $(x, y)$  of Eq. 1 and
2.  $N$  narrowband spectral sensor elements  $S_i$  behind the lens.

The sensor  $S_i$  response is

$$\bar{I}_i = \int_x \int_y I_i(x, y) = \int L(\lambda) S_i(\lambda) R(\lambda) d\lambda . \quad (2)$$

The average spectral measurement of a scene  $R$  and under the illumination  $L$  is stored as a vector  $\mathbf{s} = (\bar{I}_1, \bar{I}_2, \dots, \bar{I}_N)$ . The color constancy problem is to obtain the illuminant  $L$  using the spectral response vector  $\mathbf{s}$ . In our simulations,  $\mathbf{s}$  of only  $N=14$  elements provides good accuracy. This means that sufficient information is available in five orders of magnitude ( $10^5 \times$ ) less data than in a 10MPix camera image.

The field of view (FOV) of the sensor should be as wide as possible in order to integrate and average the changes in the surrounding scenery. This helps to reduce small chromatic objects strongly affecting the shape of the reflected spectrum in a same way as a classic gray world (Buchsbaum, 1980) color constancy algorithm works. The field of view should be at least on a same level as the camera's FOV.

### 3.2 Sensor Design

The physical design has restrictions due to the optics, electronics and material properties (Hamamatsu, 2019), but for simulation purposes the sensor responses  $S_i$  can be approximated by a Gaussian function,  $Gauss(\mu, \sigma)$ , with the maximum at 1.0 i.e. perfect quantum efficiency at the peak wavelength. The Gaussian filter response  $S_i$  is defined by the central wavelength  $\mu_i$  and bandwidth  $\sigma_i$

$$S_i(\lambda) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\lambda - \mu_i}{\sigma_i}\right)^2\right) . \quad (3)$$

The Gaussian spectral shape is a fair assumption also for a practical implementation (Jensen, 2020; Wang et al. , 2019).

Our objective is to find the optimal spectral sensor for color constancy so that it can be implemented in a miniaturized hardware. The number of channels were experimentally tested for  $N=4, 6, \dots, 16$ . The central bandwidths, Gaussian peaks, were adjusted to uniformly cover the visible spectrum ranging from 380nm to 700nm. This range covers the core of the CIE photopic luminosity function (Guild & Petavel, 1931). The channel bandwidth was defined by the *full width at half maximum (FWHM)* and the FWHM bandwidths of 10nm, 20nm and 30nm were tested. These bandwidths were selected to match the capabilities provided by the current technologies. These settings provide 21 different configurations evaluated in Sect. 5.1.

### 65 & 3 channel reference sensors

In addition to finding the best practical spectral sensor design for mobile use, we included to our experiments a "high quality reference sensor" that mimics the best available scientific spectrometers. For that purpose we defined a sensor with 5nm wide (FWHM) channels with 5nm intervals resulting to 65 channels in the same 380–700nm range. This setting is similar to a Konica-Minolta CL-70F spectrometer for the given spectral range. The 65 channel version is considered as an upper bound performance target for the more practical designs in both theoretical and real world use cases.

Some experiments were also done with a 3 channel sensor that used a spectral response of a Huawei Mate 20 Pro as the channels. The 3 channel "RGB" sensor would act similarly as a normal mobile camera that is downscaled to a single pixel. While the shape of the channels are very different to the other Gaussian shaped designs, this simulated sensor gives us a lower bound of the performance opposite to the 65 channel design.

### 3.3 White Point Regression

The spectral sensor produces a measurement vector  $\mathbf{s} = (\bar{I}_1, \bar{I}_2, \dots, \bar{I}_N)$  from (2) using the Gaussian responses  $S_i$  (Sect. 3.2). Color constancy corresponds to an estimation of the global ambient scene illumination  $L = \hat{\ell} \approx \ell$  (Finlayson et al. , 2001). The estimated *white point* is used to normalize the image colors so that achromatic regions appear gray. The white point estimation is defined as a regression problem  $\ell = (\ell_R, \ell_G, \ell_B)^T = f(\mathbf{s}_{N \times 1})$ , where  $\ell$  is the illuminant white point in RGB and  $f(\cdot)$  is a regression function that maps the spectral measurement  $\mathbf{s}$  to a white point estimate of  $L$ .

For  $f$  we tested a number of popular regression methods: Kernel Ridge regression (KR) (Murphy, 2012), Random Forest regression (RF) (Breiman, 2001), and Multilayer Per-

ceptron (MLP) (Geoffrey, 1989). The Scikit-Learn Python library was used for KR and RF. The methods' hyperparameters were optimized by grid search and cross-validation on the training data and for each sensor configuration separately. MLP was implemented using TensorFlow. MLP has three fully connected hidden layers of sizes 512-1024-512 and the standard Adam optimizer was used. In our experiments the differences between KR, RF and MLP regressors were small and thus any of them is a feasible choice.

## 4 Data

### 4.1 Generated Spectral Data

In order to train the white point regressors in Sect. 3.3 we need spectral color constancy training data. It would be straightforward to convert existing spectral image datasets (Parkkinen et al., 1988; Westland et al., 2000; Kerekes et al., 2008) for our purposes, but they are too small and do not contain natural scenes. Alternatively, spectral training data can be generated from the existing color constancy datasets using one of the RGB-to-Spectral conversion methods (Kawakami et al., 2011; Arad & Ben-Shahar, 2016; Jia et al., 2017). The recent Cube+ dataset (Banić and Lončarić, 2017) fits to our purposes. For spectral approximation we adopt parts of our recent Sensor-to-Sensor Transfer (SST) model (Koskinen et al., 2020). The original model is designed for RGB-to-RGB conversion between two different RGB sensors and therefore we need to adapt it for RGB-to-Spectral conversion using the following spectral processing steps (Fig. 2):

1. Illuminant spectrum estimation:  $\ell$  to  $\hat{L}'_{spec}$ ,
2. Raw to spectral image transform:  $I_{raw}$  to  $\hat{R}'_{spec}$ ,
3. Spectral image refinement:  $\hat{R}'_{spec}$  to  $\hat{R}'_{spec}$ ,
4. Sensor sampling of the average reflected illuminant:  $\hat{R}'_{spec} \cdot \hat{L}'_{spec}$  to  $s$ .

#### Illuminant spectrum estimation

$\hat{L}'_{spec}$  estimation is made by finding the closest matching spectrum from an existing database and then refining it to perfectly match the ground truth RGB tristimulus white points in Cube+. For this purpose, we gathered an illuminant database of 100 spectra. Most illuminants were picked from the CIE standard illuminants (International Organization for Standardization, 2006). The standard does not contain modern LEDs and therefore 13 different LED spectra were measured and added. It does provide an equation to calculate different daylight spectra as the function of a correlated color temperature:  $L(\lambda) = L_0(\lambda) + M_1L_1(\lambda) + M_2L_2(\lambda)$ .  $L_i$  are predefined illuminant characteristics vectors and  $M_i$  are coef-

ficients depending on the selected white point. We selected 70 different daylight illuminants ranging from 2500K to 9400K to cover various conditions from sunsets to cloudy days. The standard also provides typical fluorescent spectra and we selected 8 of those. Finally, we also added 9 tungsten halogen spectra ranging from 2200K to 3250K by using the Planck's law.

As in Eq. 1 the image  $I$  is formed according to (von Kries, 1970):

$$I_i(x, y) = \int L(\lambda)S_i(\lambda)R(x, y, \lambda)d\lambda, \quad (4)$$

$$i \in \{R, G, B\}.$$

Now that we are only comparing illuminant spectra and the Cube+ ground truth white points, we can set the reflectance spectrum  $R$  to a perfect white and thus effectively omit it from the equation. For the same reason, the spatial information  $(x, y)$  can be removed. We obtained the camera model used in the Cube+ and measured the sensor response spectra  $S_i$  using Labsphere QES-1000. For spectral matching the image term  $I_i$  is replaced with the ground truth illuminant white point  $\ell$ . Therefore, we need to find the illuminant  $L_d$  from our database that minimizes the equation

$$\hat{L}'_{spec} = \arg \min_{L_d} \left\| \int L_d(\lambda)S_i(\lambda)d\lambda - \ell \right\|^2, \quad (5)$$

$$i \in \{R, G, B\}.$$

$\hat{L}'_{spec}$  is the best match within the 100 illuminants. Since our database contains real illuminant spectra, the best matching illuminant has the natural shape of the corresponding white point. The found spectrum has also similar tristimulus response, but needs fine-tuning. To keep the spectral shape and naturalness intact, refining is done by linearly adjusting the red and blue parts of the spectrum from the pivot point of 530nm. The pivot point is selected to be in the middle of a typical green channel response. The refining is done iteratively until a perfect tristimulus match is achieved for  $\hat{L}'_{spec}$  by utilizing the equation ( $\hat{L}'_{spec}^{(0)} = \hat{L}'_{spec}$ )

$$\hat{L}'_{spec}{}^{(t+1)}(\lambda) = \hat{L}'_{spec}{}^{(t)}(\lambda)w(\lambda), \quad (6)$$

where  $w$  is the weight vector having a value of 1 at 530nm.

#### Raw to spectral image transform

After estimating the illuminant spectrum  $\hat{L}'_{spec} \approx L$ , the only unknown is the scene reflectance spectrum  $R$  in Eq. 4. The same approach from Section 4.1 can be used for reflectance spectrum estimation. The only difference is that the illuminant database is replaced with a database of natural reflectance spectra. The Munsell Glossy dataset (Orava,

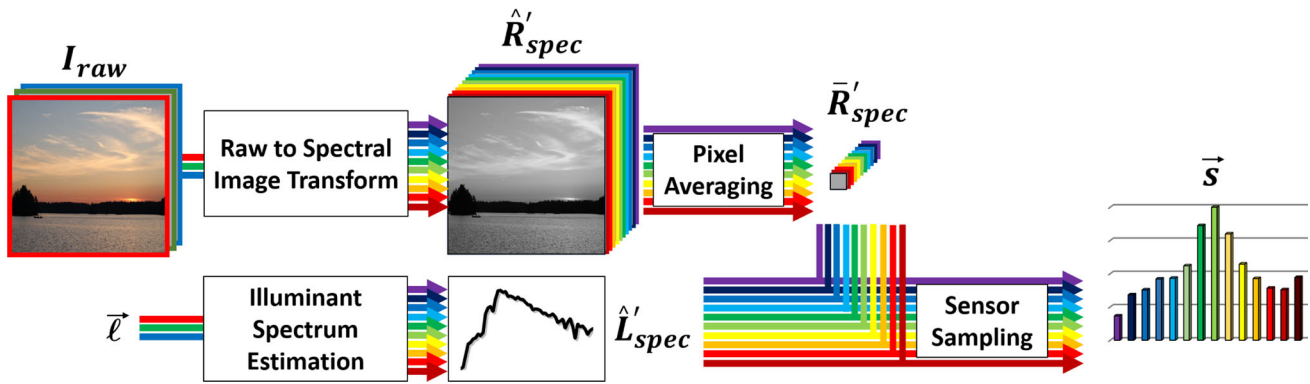


Fig. 2 The RGB-to-Spectral conversion model used to generate spectral training data (Sect. 4.1)

1995) is suitable for our purposes. The spectra are well spread over the gamut and the shapes are smooth in nature. Another change we did for the reflectance spectrum estimation is that the matching is made in the CIE  $L^*a^*b^*$  color space (International Organization for Standardization, 2008) where the luminance component  $L^*$  can be omitted. The matching is done in a 2D space using the Euclidean distance. We use  $k$  nearest neighbors and the weighted sum of their Munsell spectra to replace the RGB values of each location  $(x, y)$  with a spectral vector. The results were not very sensitive to selection of  $k$  and thus  $k$  was set to 2 in

$$\begin{aligned} \hat{R}_{spec}(x, y) &= \sum_k w_k R_{Munsell}^k \\ \{w_k\} &= \arg \min_{\{w_k\}} \|I_{raw,i}(x, y) \\ &\quad - \sum_k w_k \int \hat{L}'_{spec}(\lambda) S_i(\lambda) R_{Munsell}^k(\lambda) d\lambda\|_{a,b}^2. \end{aligned} \quad (7)$$

### Spectral image refinement

The spectral image refinement is required to perfectly match the Cube+ image RGB values. We normalized the camera spectral responses  $S_i$  so that the sum of the color channels ( $i \in \{R, G, B\}$ ) for each wavelength is one. The normalized curves  $\bar{S}_i$  are utilized as weighting functions for the iteration process ( $\hat{R}'_{spec}^{(0)} = \hat{R}_{spec}$ )

$$\begin{aligned} \hat{R}'_{spec}{}^{(t+1)}(x, y, \lambda) &= \hat{R}'_{spec}{}^{(t)}(x, y, \lambda) + \\ &\quad \left( \frac{e_i + \epsilon}{\hat{e}_i} - 1 \right) \cdot \left( \hat{R}'_{spec}{}^{(t)}(x, y, \lambda) \cdot \bar{S}_i(\lambda) \right), \end{aligned} \quad (8)$$

where the color channel specific (RGB) variables are  $\hat{e}_i$  for the estimate and  $e_i$  for the target. Iteration is finished when the spectrum matches the raw tristimulus values, i.e.  $\hat{e}_i = e_i$ . We use  $\epsilon = 10^{-6}$  to make sure the spectra are always positive. The raw input image  $I_{raw}$  contains the target values and the estimates are calculated using Eq. 4 by placing  $L = \hat{L}'_{spec}$ ,

$S = S_i$  (measured Cube+ camera spectral characterization curves) and  $R = \hat{R}'_{spec}$ .

### Sensor sampling

In the final step the estimated scene reflectance spectra and the estimated light source spectra are used to construct the spectral sensor response. First, the image spectra are averaged  $\hat{R}'_{spec} \rightarrow \bar{R}'_{spec}$ . The spectral response  $S$  now corresponds to the wide angle multi-channel sensor in Sect. 3.2 and in the following the index  $i$  refers to the channel number. The final sensor response  $s$  is computed from

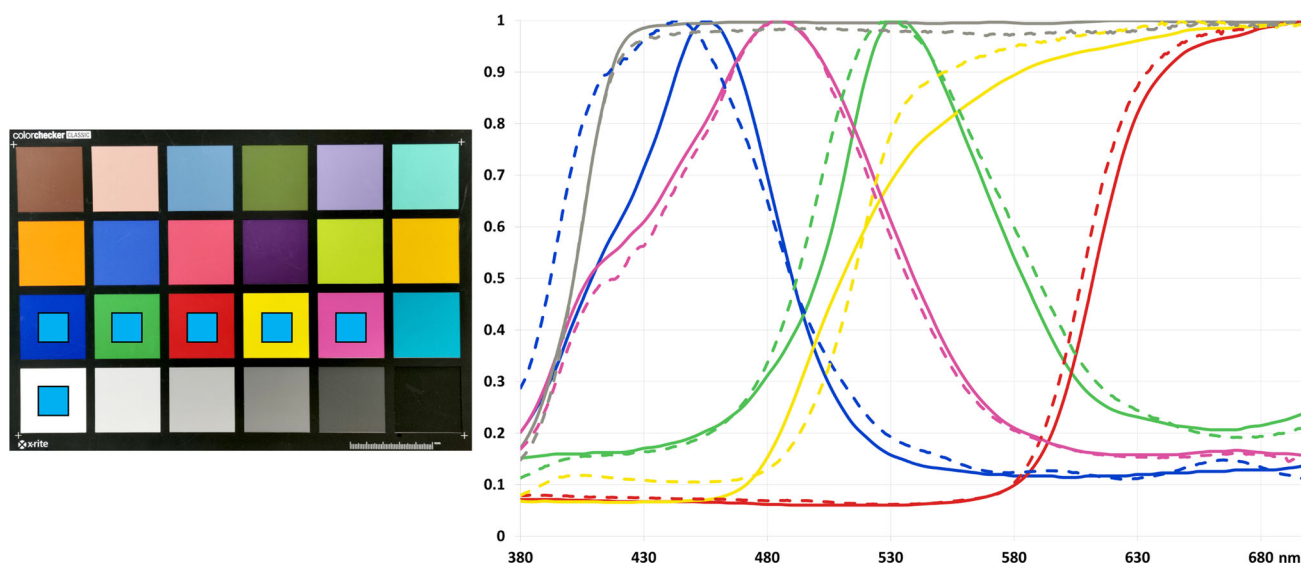
$$s = \int \hat{L}'_{spec}(\lambda) S_i(\lambda) \bar{R}'_{spec}(\lambda) d\lambda. \quad (9)$$

### Data augmentation

During the preliminary experiments it was noticed that more training samples were needed for the MLP method than the 1657 vectors from averaged images provided by the Cube+ dataset. To generate more data, the spectral images were split to 12 equal sized sub-images for which Eq. 9 was computed separately. Since the illuminant spectra is the same for all pixels, the augmentation expanded the number of different natural surfaces. This way the Cube+ dataset produced 19,884 spectral sensor vectors and white point ground truths. It is noteworthy, that the amount of data is still vastly less than typically used for conventional RGB image color constancy algorithms.

### Noise model

For more realistic results we added noise to the generated training samples. The noise gives benefit to wider channels with better signal-to-noise levels. The computational spectral sensor channels were defined to have a 100% peak quantum efficiency. We empirically set a very low light condition



**Fig. 3** Visualized accuracy of the RGB-to-Spectral conversion. The spectral accuracy is shown for the ColorChecker patches indicated with cyan squares. The ground truths (solid lines) and the estimates (dashed

lines) are plotted on the right with colors corresponding to the patches (note that there is no visible difference between the colors of the solid and dashed lines). The spectra were normalized to the peak wavelength

where the amount of photons to the most sensitive sensor channel is 20 times the FWHM width  $W$  of the channel (in nm). So in effect we assume the same exposure time for each sensor design. We only calculated the photon noise and disregarded the less significant noise sources, such as a read-out noise and ADC noise as those depend heavily on the hardware design which is not known. The photon noise is signal dependent Poisson distributed noise. The strength of the noise can be modeled as a noise which standard deviation grows with a square root of the signal level (Foi et al. , 2008; Hasinoff, 2014). Therefore, an equation  $\mathbf{s} = \mathbf{s} + \sqrt{20W\mathbf{s}}X$  was used to add noise to the sensor response  $\mathbf{s}$  for which most sensitive channel is normalized to one.  $X$  is a random sample from the normal distribution  $\mathcal{N}(\mu, \rho^2) = \mathcal{N}(0, 20W)$ .

### Transform accuracy verification

In order to verify the accuracy of the used RGB-to-Spectral conversion, we measured the spectral reflectances of the color patches of an X-Rite ColorChecker with a Photo Research PR-670 spectrometer. The spectra were then converted to RGB values using Eq. 1, where the camera spectral sensitivities were from a Huawei Mate 20 Pro and illuminant was set to an illuminant E. The RGB values were then transformed back to spectral values using the proposed RGB-to-Spectral conversion and compared to the original measured ground truth spectra. Any visible errors in the spectral domain are metameric as the differences in the RGB values are negligible. The results are shown in Fig. 3 for the challenging saturated content. The average spectra of a scene is typically

much less saturated and thus easier for the estimation as indicated by the plotted white patch accuracy.

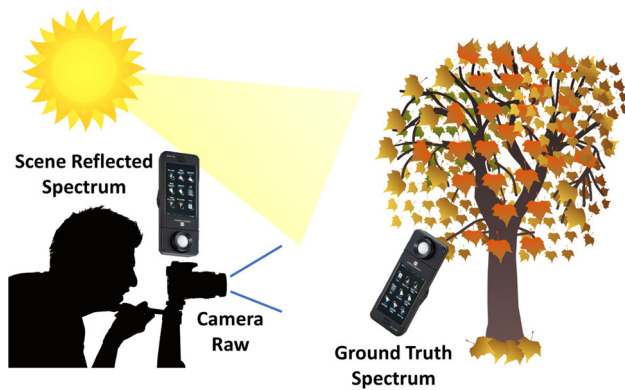
### 4.2 Multi-Illuminant Data

Multi-illuminant color constancy is a complex and largely unsolved problem. However, we wanted to study whether the spectral sensor can be helpful for the multi-illuminant case despite that it completely lacks the spatial information. To succeed in the multi-illuminant case the spectral method should detect multiple illuminant spectral fingerprints simultaneously.

The multi-illuminant data was generated by adapting the processing pipeline in Sect. 4.1 so that each image was re-illuminated by a mixture of two random illuminants. In specific, we replaced the estimated ground truth illuminant spectrum  $\hat{L}'_{spec}$  by a mixture of two randomly selected illuminant spectra. Dominant illuminant was selected and its intensity was randomly selected from (50%, 90%]. Then another secondary illuminant was randomly selected and added with intensity of at least 10%. The illuminants were picked from the set of 100 light source spectra used for the illuminant spectrum estimation in Sect. 4.1. Similar data augmentation to the single illuminant case was applied and thus resulting to the total of 83,000 spectral samples.

### 4.3 Real Spectral Color Constancy Data

To validate the results with real data, we collected a spectral color constancy dataset. Each sample contains a raw



**Fig. 4** The setup used to capture the real spectral color constancy dataset

image captured with a Huawei Mate 20 Pro mobile phone and two spectral measurements by a Konica Minolta CL-70F spectrometer. The first spectral measurement represents the average spectrum of the scene reflected illuminant and the second the ground truth illuminant. The first measurement was made by placing Konica Minolta next to the phone and pointing it towards the scene. The second measurement was made by placing the spectrometer to the scene to measure the ground truth illumination falling on the area. The data gathering setup is illustrated in Fig. 4. The ground truth white points were calculated using the illuminant spectrum, the camera spectral response and a perfect white reflectance spectrum in Eq. 4.

The dataset consists of 235 raw images with their corresponding spectral measurements. The dataset was purposely made difficult for color constancy by including scenes that are dominated by a few chromatic colors and often without any clear gray areas. These cases are challenging also to spectral color constancy as the illuminant spectrum and the reflected spectrum are clearly different (the solid and dashed lines in Fig. 1). Examples from the dataset are shown in Figs. 1 and 6.

## 5 Experiments

### 5.1 Sensor Design

We tested the 21+2 sensor configurations in Sect. 3.2: 7 different filter configurations from  $N = 4$  to  $N = 16$  and 3 different filter bandwidths from 10nm to 30nm. In addition, we had a 65 channel reference design that was a target for the other configurations and a 3 channel design that gave understanding about the lower bound performance. The evaluations were made with the generated Cube+ spectral images (Sect. 4.1) and with the real spectral data (Sect. 4.3). All results are average numbers from 3-fold cross-validation and the experiments were carried out with noise-free and

noise added measurements. The noisy measurements reflect better the performance in realistic low light conditions and demonstrate the difference between the narrow (10nm) and wide (30nm) band sensors. The performance measure in all our experiments is the mean angular error between the ground truth white point  $\ell$  and the estimated white point  $\ell'$  (Finlayson et al., 2017)

$$err = \cos^{-1} \left( \frac{\ell \cdot \ell'}{\|\ell\| \cdot \|\ell'\|} \right). \quad (10)$$

Results are shown in Fig. 5 and provide two expected findings:

1. Adding more channels systematically improves the results until they saturate at  $N \geq 10$ .
2. Wider filters are more robust to low light and noisy scenes (Cube+).

The average error with the real data ( $\approx 2.4^\circ$ ) is clearly worse than with the generated Cube+ ( $\approx 0.5^\circ$  for clean and  $\approx 1.0^\circ$  for noisy) which can be explained by the fact that the real dataset is more challenging. However, both results are well below  $3.0^\circ$  that is the generally used just noticeable difference of human color perception.

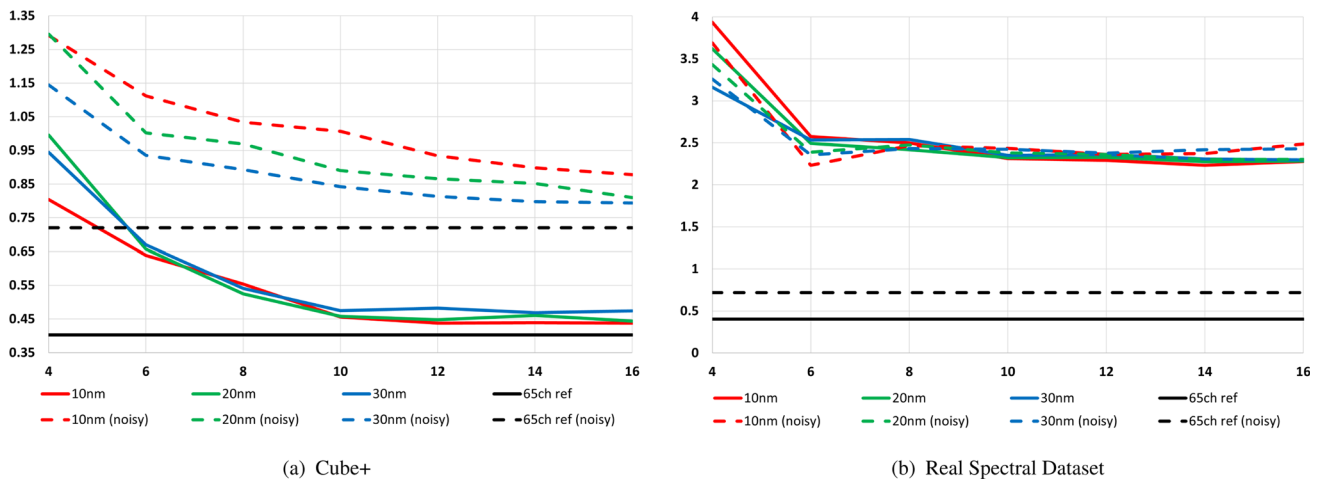
The results with our spectral dataset are clearly worse than with Cube+ and there is no significant difference between the clean and noisy results. The main reason for this is that our scenes are more difficult (often only a few dominating colors) for color constancy, there are much fewer scenes and spectra were measured using a real spectrometer. Based on the noise-free and noisy results and trying to keep the design feasible, we selected the option with 14 channels having 20nm width as the best miniaturized sensor design for the remaining experiments in addition to the 65 channel reference design representing a high-end spectrometer.

Supplementary tests in addition to the Gaussian shaped sensors were carried out with the 3 channel design that represented a typical pixel of a mobile camera. While the "RGB" sensor does not see any spatial information, and thus cannot perform as well as the real mobile cameras, it gives a reliable lower bound for the multi-channel sensors. We conducted the evaluation using the Cube+ dataset as its image count was high enough to give very stable results. The accuracy of the "RGB" sensor dropped 54% on average and 64% on the 95<sup>th</sup> percentile compared to the Gaussian (10nm) 4 channel sensor. The results are in line with the expectations when looking at the results for the Gaussian shaped sensors in Fig. 5.

### 5.2 Method Comparison

We compared the spectral color constancy with the settings  $N = 14$  and sensor bandwidth 20nm against three SotA





**Fig. 5** Results for the Gaussian shaped sensors using the MLP white point regressor. Y-axis is the mean angular error from a 3-fold cross-validation and x-axis is the number of channels except with the black lines that show the reference 65 channel sensor as the target performance

**Table 1** Comparison of the proposed spectral (both 14 channel and 65 channel options) and SotA color constancy methods in the 3-fold cross-validation. The numbers are angular errors (lower is better)

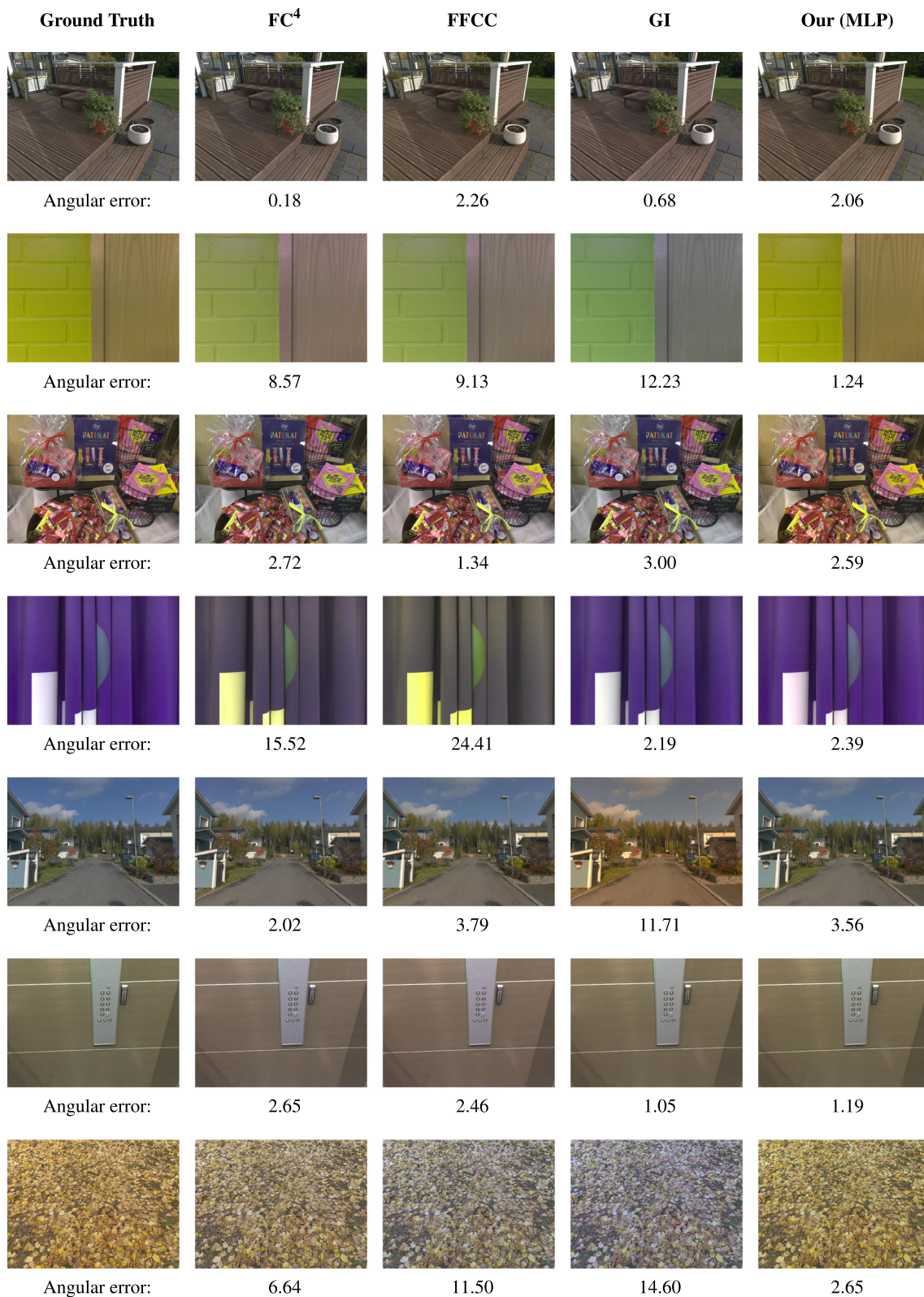
Method	Cube+			Real Spectral Dataset		
	Mean	Med	95 <sup>th</sup>	Mean	Med	95 <sup>th</sup>
GI	2.871	1.664	9.299	7.575	6.512	21.420
FFCC	1.297	0.676	6.420	3.811	2.655	15.288
FC <sup>4</sup>	1.762	1.297	4.608	3.645	2.828	10.282
Spectral, 14ch (KR)	0.658	0.475	1.640	<u>2.215</u>	1.593	<u>6.244</u>
Spectral, 14ch (RF)	0.511	0.307	1.541	2.296	1.773	6.567
Spectral, 14ch (MLP)	<u>0.461</u>	0.312	<u>1.202</u>	2.274	<u>1.584</u>	6.456
Spectral, 65ch (KR)	0.658	0.439	1.759	2.274	1.698	6.411
Spectral, 65ch (RF)	0.491	<b>0.291</b>	1.437	2.240	<u>1.564</u>	6.703
Spectral, 65ch (MLP)	<b>0.403</b>	<u>0.292</u>	<b>1.108</b>	<b>1.111</b>	<b>0.526</b>	<b>4.436</b>

Particularly note the remarkable improvements on the most difficult cases (95<sup>th</sup> percentile)

**Table 2** Angular errors for the cross-dataset experiment

Method	NUS			Intel-TUT			Shi-Gehler			Real Spectral Dataset		
	Mean	Med	95 <sup>th</sup>	Mean	Med	95 <sup>th</sup>	Mean	Med	95 <sup>th</sup>	Mean	Med	95 <sup>th</sup>
GI	2.962	2.103	7.932	3.966	1.923	13.778	3.077	2.168	9.123	7.575	6.512	21.420
FFCC	2.474	1.900	8.457	3.296	2.178	13.044	2.145	1.359	9.048	6.143	4.008	22.333
FC <sup>4</sup>	2.227	1.749	5.581	2.919	2.082	8.416	2.417	1.537	8.331	4.795	<b>3.267</b>	14.597
Spectral, 14ch (KR)	1.446	0.921	3.344	<b>1.583</b>	1.156	<b>4.213</b>	1.206	0.737	<b>3.966</b>	5.043	4.405	10.239
Spectral, 14ch (RF)	1.461	1.012	4.179	2.330	1.397	7.663	1.284	0.549	5.614	<u>3.961</u>	3.327	<u>9.056</u>
Spectral, 14ch (MLP)	<b>1.009</b>	0.792	<b>2.585</b>	<u>1.649</u>	<b>1.044</b>	<u>5.219</u>	<b>1.008</b>	<u>0.491</u>	<u>4.367</u>	4.589	3.937	10.628
Spectral, 65ch (KR)	1.101	0.814	3.062	2.572	1.321	9.514	1.332	0.659	4.981	4.887	4.313	10.677
Spectral, 65ch (RF)	1.339	<u>0.766</u>	5.219	2.547	1.202	9.114	1.486	<b>0.482</b>	7.098	<b>3.906</b>	<u>3.294</u>	<b>8.635</b>
Spectral, 65ch (MLP)	<u>1.028</u>	<b>0.746</b>	<u>2.976</u>	2.162	<u>1.111</u>	6.702	<u>1.112</u>	0.518	5.126	4.444	3.932	9.674

All methods except GI (a static method not requiring training) are trained using the Cube+ images  
 The results with the spectral CC algorithms are shown using the selected 14 channel sensor design with 20nm FWHM and the 65 channel reference sensor with 5nm FWHM



**Fig. 6** Visualized errors for the tested algorithms from the Real Spectral Dataset. The images include also a static color space transform and an sRGB gamma for displaying purposes. The average results of the algorithms are similar as shown statistically for the given dataset in Table 2

methods: Grayness Index (GI) (Qian et al. , 2019), Fast Fourier Color Constancy (FFCC) (Barron & Tsai, 2017) and Fully Convolutional with Confidence (FC<sup>4</sup>) (Hu et al. , 2017). GI is a static method that does not need training data, but it is competitive against the learning-based methods and particularly effective in cross-dataset evaluations. FFCC and FC<sup>4</sup> are SotA learning-based methods, but with an important difference: FFCC omits the spatial dimension and uses image RGB distributions while FC<sup>4</sup> directly uses the RGB images.

We repeated the 3-fold cross-validation of the previous experiment with the generated Cube+ and the Real Spectral Dataset. The results in Table 1 provide two important findings:

1. All variants of spectral color constancy outperform the SotA RGB methods on both datasets.
2. The spectral method is particularly effective on the most difficult scenes (95<sup>th</sup> percentile) for which it obtains remarkable improvements of 39% to 74% even with the 14 channel configuration

### 5.3 Cross-dataset Evaluation

The cross-dataset evaluations are important as the methods are not allowed to use training data from the tested datasets and therefore the results reflect better the practical performance. For the cross-dataset evaluations all methods were trained with the Cube+ images. From the popular color constancy benchmarks we selected those where we were able to find the same camera model and measure its spectral response. The selected test datasets were Intel-TUT (Aytekin et al. , 2017), NUS (Cheng et al. , 2014) and Shi-Gehler (Hemrit et al. , 2018), with 142, 197 and 482 images, in addition to our own collected Real Spectral Dataset with 235 images.

The results are shown in Table 2 and visualized in Fig. 6. The spectral color constancy method achieved superior or on par accuracy on all four datasets. Similar to the previous experiment, the performance was particularly good for the most difficult images (95<sup>th</sup> percentile) where the spectral method achieved notable improvements of 38–54% with the 14 channel design.

### 5.4 Multi-Illuminant Case

The MLP network created for the single illuminant color constancy (3 outputs) was modified to produce two white points and their relative intensities (6+2 outputs). The output of the “dual-MLP” setting can be expressed as a weighted sum of the two white points  $w'_1 \ell'_1 + w'_2 \ell'_2$ , where  $\ell'_i$  are the estimated white points and  $w'_i$  their weights. For simplicity, the second weight could be defined as  $w'_2 = 1.0 - w'_1$ , but we did not find much difference in the results of the two and therefore used two outputs. Dual-MLP is able to estimate the

**Table 3** Weighted angular errors for the generated images containing random mixtures of two illuminants (3-fold cross-validation)

<i>Estimator</i>	14 channels			65 channels		
	Mean	Med	95 <sup>th</sup>	Mean	Med	95 <sup>th</sup>
MLP	5.811	4.767	14.059	5.712	4.661	13.975
Dual-MLP	4.535	3.548	11.443	4.423	3.404	11.430

illuminant(s) for the both single and dual illuminant case. In the correctly estimated single illuminant instance, the other weight is evaluated to be 0. It should be noted that in practice two illuminants are often also spatially separated, for example, consider an image captured indoors in an office that includes window viewing outdoors. However, since the single pixel sensor has no spatial information the weights represent the spatial extent of the two lights.

During the experiments, it was noted that the regressor detected well the two illuminants but was less successful detecting the mixing weights. Therefore we used the following compound error that uses the ground truth mixing weights to measure how well the correct illuminants were detected (for the single illuminant MLP  $\ell'_1 = \ell'_2$ )

$$err_{dual} = w_1 \cdot err(\ell_1, \ell'_1) + w_2 \cdot err(\ell_2, \ell'_2) ,$$

where  $err()$  is the angular error in Eq. 10. Note that since the order of the two white points is arbitrary, the white points were swapped and the minimum recorded as the error.

The results for the dual-MLP are shown in Table 3. The numbers are clearly worse than in the single illuminant experiments and that demonstrates the difficulty of multi-illuminant color constancy. However, the dual-MLP architecture obtains systematically 15–25% better accuracies than the single illuminant MLP indicating that the single pixel spectral measurement can detect the spectral fingerprints of multiple illuminants. It should be noted that these results are promising but only preliminary as for the practical multi-illuminant color constancy also the spatial segments of the different illuminants should be estimated.

## 6 Conclusions

We introduced a new approach for computational color constancy. Instead of the conventional procedure of using RGB images, our approach uses average color spectra sampled from the visible part of the electromagnetic spectrum. The spectral color constancy achieved the highest accuracy with clear margins to SotA RGB methods. In particular, remarkable improvement of over 50% in the challenging cross-dataset evaluations was achieved in the most difficult

cases using a design that is practical for mobile devices. It also proved that the data generation method was effective as the results with the generated training data and tested on real measured data still achieved superior results. In addition, we showed that a single pixel spectral sensor is able to detect multiple illuminants from a single global measurement. We conclude that the spectral dimension is more important than the spatial dimension for estimating the illuminant white points.

**Acknowledgements** The authors are employed by Huawei Technologies Oy (Finland) Co. Ltd and Tampere University. No other funding was received for conducting this study.

**Funding** Open access funding provided by Tampere University including Tampere University Hospital, Tampere University of Applied Sciences (TUNI).

**Data Availability** Most of the datasets analysed during the current study are derived from the following public domain resources:

- Cube+ (Banić and Lončarić, 2017)  
– <https://doi.org/10.48550/arXiv.1712.00436>  
– [https://ferhr.sharepoint.com/:f:/s/imageprocessinggroup/Eno1wDbdYiFNkwo6-1689-sBsViYV\\_KHRVAdulvM5jjMzQ?e=2MXqWg](https://ferhr.sharepoint.com/:f:/s/imageprocessinggroup/Eno1wDbdYiFNkwo6-1689-sBsViYV_KHRVAdulvM5jjMzQ?e=2MXqWg)
- Intel-TUT (Aytekin et al., 2017)  
– <https://doi.org/10.1109/TIP.2017.2764264>  
– <https://etsin.fairdata.fi/dataset/8155b8d2-4947-436f-b323-756f95a7058e>
- NUS (Cheng et al., 2014)  
– <https://doi.org/10.1364/JOSAA.31.001049>  
– [https://cvil.eecs.yorku.ca/projects/public\\_html/illuminant/illuminant.html](https://cvil.eecs.yorku.ca/projects/public_html/illuminant/illuminant.html)
- Shi-Gehler (Hemrit et al., 2018)  
– <https://doi.org/10.2352/ISSN.2169-2629.2018.26.350>  
– [https://www2.cs.sfu.ca/protect/unhbox\voidb@x\penalty\M\}color/data/shi\\_gehler/](https://www2.cs.sfu.ca/protect/unhbox\voidb@x\penalty\M\}color/data/shi_gehler/)

Our collected spectral dataset and codes were created while working for a commercial company (Huawei Technologies Oy (Finland) Co. Ltd) and cannot be made public for business reasons.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arad, B., & Ben-Shahar, O. (2016). High-resolution hyperspectral imaging via matrix factorization. *ECCV*. <https://doi.org/10.1109/CVPR.2011.5995457>
- Aytekin, Ç., Nikkanen, J., & Gabbouj, M. (2017). INTEL-TUT dataset for camera invariant color constancy research. *CoRR*, abs/1703.09778, doi10.1109/TIP.2017.2764264.
- Banić, N. & Lončarić, S. (2017). Unsupervised learning for color constancy. *CoRR*, doi10.48550/arXiv:1712.00436.
- Barnard, K., Cardei, V., & Funt, B. (2002). A comparison of computational color constancy algorithms-part I: Methodology and experiments with synthesized data. *IEEE Signal Processing Society*. <https://doi.org/10.1109/TIP.2002.802531>
- Barron, J. T. (2015). Convolutional color constancy. *ICCV*. <https://doi.org/10.1109/ICCV.2015.51>
- Barron, J. T. & Tsai, Y. (2017). Fast fourier color constancy. In *2017 IEEE conference on computer vision and pattern recognition (CVPR)*. 10.1109/CVPR.2017.735.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1), 1–26. [https://doi.org/10.1016/0016-0032\(80\)90058-7](https://doi.org/10.1016/0016-0032(80)90058-7)
- Chakrabarti, A., Hirakawa, K., & Zickler, T. (2011). Color constancy with spatio-temporal statistics. *IEEE PAMI*, 34(8), 58. <https://doi.org/10.1109/TPAMI.2011.252>
- Chen, X. (2017). Color Constancy for RGB and Multispectral Images. Thesis (School of Computing Science): Simon Fraser University.
- Cheng, D., Prasad, D. K., & Brown, M. S. (2014). Illuminant estimation for color constancy: Why spatial-domain methods work and the role of the color distribution. *Journal of the Optical Society of America A*, 31(5), 552. <https://doi.org/10.1364/JOSAA.31.001049>
- Dietz, C. (2011). C&A Application Note No.1: Light sources and illuminants. Technical report, Konica Minolta.
- Finlayson, G., Hordley, S., & Hubel, P. M. (2001). Color by correlation: A simple, unifying framework for color constancy. *IEEE PAMI*, 32(11), 552.
- Finlayson, G. D. (2013). Corrected-moment illuminant estimation. In *ICCV* (pp. 1904–1911). 10.1109/ICCV.2013.239.
- Finlayson, G. D., Drew, M. S., & Funt, B. V. (1994). Spectral sharpening: Sensor transformations for improved color constancy. *Journal of the Optical Society of America A*, 11(5), 1553–1563. <https://doi.org/10.1364/JOSAA.11.001553>
- Finlayson, G. D., Zakizadeh, R., & Gijsenij, A. (2017). The reproduction angular error for evaluating the performance of illuminant estimation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(7), 52. <https://doi.org/10.1109/TPAMI.2016.2582171>
- Foi, A., Trimeche, M., Katkovnik, V., & Egiazarian, K. (2008). Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10), 1737–1754. <https://doi.org/10.1109/TIP.2008.2001399>
- Gao, L., & Wang, L. V. (2016). A review of snapshot multidimensional optical imaging: Measuring photon tags in parallel. *Physics Reports*, 616(9), 1–37. <https://doi.org/10.1016/j.physrep.2015.12.004>
- Gao, S., Han, W., Yang, K., Li, C., & Li, Y. (2014). Efficient color constancy with local surface reflectance statistics. In *ECCV* (pp. 158–173). 10.1007/978-3-319-10605-2\_11.
- Geoffrey, E. H. (1989). Connectionist learning procedures.
- Gevers, T., Stokman, H., & van de Weijer, J. (2000). Color constancy from hyper-spectral data. *BMVC*. <https://doi.org/10.5244/C.14.30>
- Guild, J. & Petavel, J. E. (1931). The colorimetric properties of the spectrum. In *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* (vol. 230, pp. 681–693), 10.1098/rsta.1932.0005.
- Hamamatsu (2019). Image Sensors: Selection guide-November 2019. Hamamatsu Photonics K.K.
- Hasinoff, S. W. (2014). Photon, poisson noise. Springer (pp. 608–610). 10.1007/978-0-387-31439-6\_482.

- Hemrit, G., et al. (2018). Rehabilitating the colorchecker dataset for illuminant estimation. In *Color and Imaging Conference*. 10.2352/ISSN.2169-2629.2018.26.350.
- Hu, Y., Wang, B., & Lin, S. (2017). FC4: Fully convolutional color constancy with confidence-weighted pooling. *CVPR*. <https://doi.org/10.1109/CVPR.2017.43>
- Hui, Z., Chakrabarti, A., Sunkavalli, K., & Sankaranarayanan, A. C. (2019). Learning to separate multiple illuminants in a single image. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. 10.1109/CVPR.2019.00390.
- Hui, Z., Sunkavalli, K., Hadap, S., & Sankaranarayanan, A. C. (2018). Illuminant spectra-based source separation using flash photography. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*. 10.1109/CVPR.2018.00650.
- International Organization for Standardization (2006). ISO 11664-2 Standard - Colorimetry Part 2: Standard Illuminants for Colorimetry.
- International Organization for Standardization (2008). ISO 11664-4 Standard - Colorimetry Part 4: CIE 1976 L\*a\*b\* Colour Space.
- Jensen, K. (2020). Chip-scale spectral sensing: understanding the new uses for ultra-precise light-source measurement. Technical report, ams AG.
- Jia, Y., Zheng, Y., Gu, L., Subpa-Asa, A., Lam, A., Sato, Y., & Sato, I. (2017). From RGB to spectrum for natural scenes via manifold-based mapping. *ICCV*. <https://doi.org/10.1109/ICCV.2017.504>
- Kawakami, R., Matsushita, Y., Wright, J., Ben-Ezra, M., Tai, Y.-W., & Ikeuchi, K. (2011). High-resolution hyperspectral imaging via matrix factorization. *CVPR*. <https://doi.org/10.1109/CVPR.2011.5995457>
- Kerekes, J. P., Strackerjan, K., & Salvaggio, C. (2008). Spectral reflectance and emissivity of man-made surfaces contaminated with environmental effects. *Optical Engineering*, 47(10), 1–10.
- Keshav, V. & GVSL, T. P. (2019). Decoupling semantic context and color correlation with multi-class cross branch regularization. In *ICME* (pp. 1492–1497). 10.1109/ICME.2019.00258.
- Khan, H. A., Thomas, J.-B., Hardeberg, J. Y., & Laligant, O. (2017). Illuminant estimation in multispectral imaging. *Journal of the Optical Society of America A*, 34(7), 1085–1098. <https://doi.org/10.1364/JOSAA.34.001085>
- Kokka, A., et al. (2018). Development of white LED illuminants for colorimetry and recommendation of white LED reference spectrum for photometry. *Metrologia*, 55(4), 526–534. <https://doi.org/10.1088/1681-7575/aacae7>
- Koskinen, S., Acar, E., & Kämäräinen, J.-K. (2021). Single pixel spectral color constancy. In *BMVC*.
- Koskinen, S., Yang, D., & Kämäräinen, J.-K. (2020). Cross-dataset color constancy revisited using sensor-to-sensor transfer. In *BMVC*.
- Murphy, K. (2012). *Machine learning: A probabilistic perspective* (Vol. 58). The MIT Press.
- Nathan, A. H., & Michael, W. K. (2013). Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9), 1–23. <https://doi.org/10.1117/1.OE.52.9.090901>
- Orava, J. (1995). The reflectance spectra of 1600 glossy Munsell color chips. <https://sites.uef.fi/spectral/databases-software/spectral-database/>. Accessed 22 July 2022.
- Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. The MIT Press.
- Parkkinen, J., Jaaskelainen, T., & Kuittinen, M. (1988). Spectral representation of color images. In *IEEE 9th International Conference on Pattern Recognition* (vol. 2, pp. 933–935).
- Qian, Y., Kämäräinen, J.-K., Nikkanen, J., & Matas, J. (2019). On finding gray pixels. In *CVPR* (pp. 8054–8062). 10.1109/CVPR.2019.00825.
- van de Weijer, J., & Gevers, T. (2005). Color constancy based on the Grey-edge hypothesis. *ICIP*. <https://doi.org/10.1109/ICIP.2005.1530157>
- von Kries, J. (1970). Influence of adaptation on the effects produced by luminous stimuli. *Source of Color Science*.
- Wang, Z., et al. (2019). Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nature Communications*, 10(1020), 560. <https://doi.org/10.1038/s41467-019-08994-5>
- Westland, S., Shaw, A., & Owens, H. (2000). Colour statistics of natural and man-made surfaces. *Sensor Review*, 20(1), 50–55. <https://doi.org/10.1108/02602280010311392>
- Yang, K.-F., Gao, S.-B., & Li, Y.-J. (2015). Efficient illuminant estimation for color constancy using grey pixels. *CVPR*. <https://doi.org/10.1109/CVPR.2015.7298838>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.