



# Improving color constancy by discounting the variation of camera spectral sensitivity

SHAO-BING GAO,<sup>1,2</sup> MING ZHANG,<sup>1</sup> CHAO-YI LI,<sup>1,3</sup> AND YONG-JIE LI<sup>1,\*</sup>

<sup>1</sup>Center for Information in BioMedicine, University of Electronic Science and Technology of China, Chengdu 610054, China

<sup>2</sup>College of Computer Science, Sichuan University, Chengdu 610065, China

<sup>3</sup>Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

\*Corresponding author: liyj@uestc.edu.cn

Received 8 June 2017; revised 9 July 2017; accepted 11 July 2017; posted 11 July 2017 (Doc. ID 297212); published 31 July 2017

It is an ill-posed problem to recover the true scene colors from a color biased image by discounting the effects of scene illuminant and camera spectral sensitivity (CSS) at the same time. Most color constancy (CC) models have been designed to first estimate the illuminant color, which is then removed from the color biased image to obtain an image taken under white light, without the explicit consideration of CSS effect on CC. This paper first studies the CSS effect on illuminant estimation arising in the inter-dataset-based CC (inter-CC), i.e., training a CC model on one dataset and then testing on another dataset captured by a distinct CSS. We show the clear degradation of existing CC models for inter-CC application. Then a simple way is proposed to overcome such degradation by first learning quickly a transform matrix between the two distinct CSSs (CSS-1 and CSS-2). The learned matrix is then used to convert the data (including the illuminant ground truth and the color-biased images) rendered under CSS-1 into CSS-2, and then train and apply the CC model on the color-biased images under CSS-2 without the need of burdensome acquiring of the training set under CSS-2. Extensive experiments on synthetic and real images show that our method can clearly improve the inter-CC performance for traditional CC algorithms. We suggest that, by taking the CSS effect into account, it is more likely to obtain the truly color constant images invariant to the changes of both illuminant and camera sensors. © 2017 Optical Society of America

**OCIS codes:** (100.0100) Image processing; (330.1720) Color vision.

<https://doi.org/10.1364/JOSAA.34.001448>

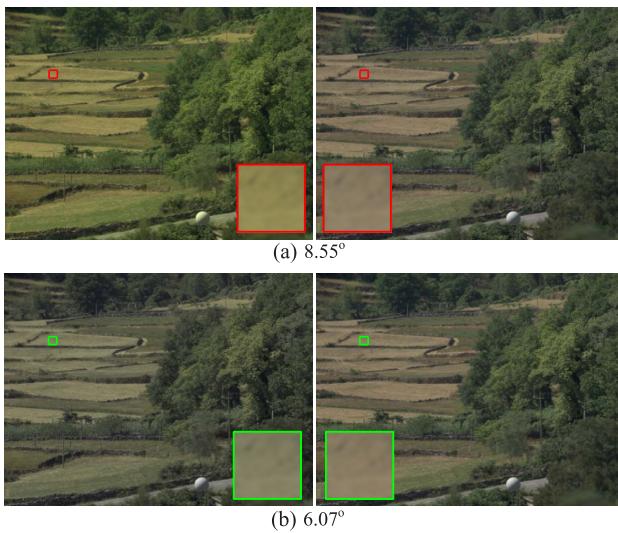
## 1. INTRODUCTION

With the rapid proliferation of digital imaging and videoing, accurate recording of the true color of a scene through the device-captured image is of extreme importance for many practical applications, ranging from color-based object recognition and tracking to quality control of textile and food processing [1–3]. However, these device-captured image colors are always affected by the prevailing changed light source color incident in a scene. Figure 1(a) indicates that the images of the same scene rendered under two different illuminants obviously exhibit various color appearances. Thus, for the sake of maintaining the true color appearance of objects, the color artifacts due to the change of illuminant should be carefully eliminated.

Color constancy (CC) refers to the perceptual constancy of the human visual system, which enables the perceived color of objects in a scene largely constant as the light source color changes [3–5]. Many CC algorithms have been specially designed to imitate this visual attribute by computationally estimating the illuminant and then removing the color cast to discount the bias due to the illuminant (see [3,4,6] for excellent reviews). Recently, the performance of CC on several benchmark datasets

has been significantly progressed, especially by the state-of-the-art CC models that are based on extensive feature extraction and machine learning techniques [7–13]. However, as indicated in Fig. 1, image color is not only influenced by the scene light source color. Actually, during the image acquisition phase, there are three factors influencing the color values that we finally measure at each pixel, i.e., the reflectance of the objects in the scene, the illuminant, and the camera spectral sensitivity (CSS). While CC is always devoted to obtaining a stable color representation of the scene across the changes of the illuminant [Fig. 1(a)], the CSS also affects the color appearance of the scene [Fig. 1(b)].

So far, however, almost all existing benchmark datasets were collected using one camera with a fixed CSS [14,15–19] and many state-of-the-art learning-based CC algorithms implicitly assume that the images in a dataset are captured by the same CSS, and are limited to evaluate their performance on intra-dataset-based CC (intra-CC) [3,7–13,20–23], i.e., training the model on one part of the dataset and testing the trained model on another part of the dataset. Thus, the effect of CSS on CC cannot be deeply probed according to the intra-CC strategy. However, as indicated in Figs. 1(b) and 2, both the



**Fig. 1.** Effects of CSS and illuminant on image color appearance. (a) Images rendered with the same CSS but under different illuminants (left: green light source, right: white light source). (b) Images rendered with different CSSs but under the same illuminant (both are white light sources). Chromatic angular error values between the corresponding patches are given at the bottom of each row.

color distributions of the recorded images and the extracted true illuminants in the datasets rely on CSS.

While an existing intra-CC strategy discounts the color bias induced by the illuminant, the recovered image color is, in fact, the appearance of the combination of the object reflectance and the camera sensor effect [Fig. 1(b)]. As a consequence, the existing CC algorithms may suffer problems when dealing with the inter-dataset-based (inter-CC) application, i.e., training a model on one dataset that was captured by a specific camera and then testing the trained model on another dataset that was captured by another camera with a different CSS.

In this paper, we draw insights into the rich literature on intrinsic image research, that aims to decompose an image into various individual physical characteristics (e.g., reflectance and shade) that are independent of both the illuminant and camera sensor [26–28]. To this end, we primarily focus on studying the effect of the CSS on CC arising in the inter-CC setup.

Specifically, we first point out that the chromatic distributions of both the measured illuminants and recorded images by various CSSs are quite different, even for the same scene under one illuminant, which is one of the significant causes that results in the failure of inter-CC evaluation using current learning-based CC algorithms. Then, in order to overcome this drawback, we propose a simple yet efficient framework that incorporates the information of the CSS into the process of CC. This is the main contribution of this work. In particular, we first learn a transform matrix between the CSS functions of two distinct cameras (CSS-1 and CSS-2). Then, the learned matrix is used to convert the color-biased images and the provided illuminants recorded with CSS-1 into those with CSS-2 before training the model and testing on the image(s) recorded with CSS-2.

Moreover, by taking into account the CSS information, we also demonstrate how to obtain a stable color image

representation of the scene that is almost independent of both the illuminant and camera sensor. This stable color image representation may benefit the further computer vision applications such as intrinsic image decomposition, 3-D view synthesis, physics-based reflectance descriptor, and so on.

Although it is well known that the CSS affects the image formation, how to discount such a prevailing adverse effect in an elegant way and thus improve the performance of other color applications is a very difficult problem (e.g., estimating information from image colors [29]). In fact, some CC research has also been devoted to the effect of imaging sensors on chromatic adaptation in which both the source and destination illuminants are known. For example, the spectral sharpening of sensors attempts to simplify the illuminant change characterization and therefore improve the performance of any CC algorithm that is based on the diagonal-matrix transformation [30,31].

Totally different from those attempts, our aims are to probe into the possible effects of various CSSs on inter-CC performance for traditional learning-based CC models and to propose a feasible solution to solve such a challenging problem. To the best of our knowledge, such an issue has not been explicitly studied before in the area of CC. Our motivation is that although learning-based models work well for intra-CC application, they require the burdensome manual collecting of an extensive training set for each CSS. In contrast, it would be undoubtedly more practically valuable if we could always provide any CSS (i.e., any camera) with high inter-CC performance competitive to that of learning-based intra-CC methods, simply based on the collection of a training set for only one CSS. This is the fundamental advantage of our method.

The rest of this paper is organized as follows. Section 2 formulates the problem. The proposed solution is described in Section 3. Section 4 presents the experiments to validate our theoretical analysis and the proposed method. We conclude in Section 5 by discussing some contributions and limitations of this work.

## 2. COLOR IMAGE FUNDAMENTALS AND COLOR CONSTANCY

In this work, we reasonably assume a single light source color across a scene, as done in many previous studies on CC [3,7,32–35]. Based on the common form of the linear imaging equation, the interaction of surface, light source, and sensor can be indicated by a simple equation written as [3,4]

$$f_c(x) = \int_{\omega} S_c(\lambda) I(\lambda) C(x, \lambda) d\lambda, \quad (1)$$

where the integral is taken over the visible spectrum  $\omega$  and  $c \in \{R, G, B\}$  are sensor channels. Equation (1) states that the captured raw image values  $f(x) = [f_R(x), f_G(x), f_B(x)]^T$  directly depend on the color of the light source  $I(\lambda)$ , the surface reflectance  $C(x, \lambda)$ , and the CSS curve  $S(\lambda) = [S_R(\lambda), S_G(\lambda), S_B(\lambda)]$ , where  $x$  is the spatial coordinate and  $\lambda$  is the wavelength of the light. Note that among the methods aiming at calibrating the CCS effect on the image color, one direct way is to first estimate the CSS of each camera and then use it to correct the image colors among different cameras [25,36]. In contrast, here we are interested in how the CSS affects the performance of CC

on estimating the light source color from the recorded raw image values and how the true image colors can be obtained by discounting both the CSS and illumination effect.

Earlier research has demonstrated that it is sufficient to simulate the color transform of the image induced by the illuminant with a diagonal transform [3,6,30,37,38]. Thus, based on the diagonal transform assumption in CC, Eq. (1) can be simplified as

$$f_c(x) = I_c C_c(x). \quad (2)$$

In general, CC aims at removing the color bias in the captured image  $f_c(x)$  by first estimating the illuminant  $I_c$ ,  $c \in \{R, G, B\}$ , then recovering  $C_c(x)$  by dividing  $f_c(x)$  by  $I_c$ . However, both  $I_c$  and  $C_c(x)$  in Eq. (2) are usually unknown and hence, given only the image values  $f_c(x)$ , estimating  $I_c$  is a typically ill-posed problem that cannot be solved without further constraints [33,39–42].

While many existing algorithms rely on the aforementioned steps to achieve CC, it should be explicitly pointed out that these algorithms are camera dependent (e.g., a fixed camera [8]). Unfortunately, almost all current CC models only focus on developing techniques to eliminate the color bias in images induced by the illuminant but ignoring the fact that the CSS also contributes to the color bias when recording the images. In particular, for almost all the color constancy datasets, the

illuminant ground truth of each image is extracted from a specific local region (e.g., the gray ball or color-checker) of the image recorded by the CSS. That is, the illuminant ground truth is also CSS dependent.

For a clear illustration, we express the sensor responses to both the illuminant and surface, respectively, as [8]

$$I_c = \int_{\omega} S_c(\lambda) I(\lambda) d\lambda, \quad (3)$$

$$C_c(x) = \int_{\omega} S_c(\lambda) C(x, \lambda) d\lambda. \quad (4)$$

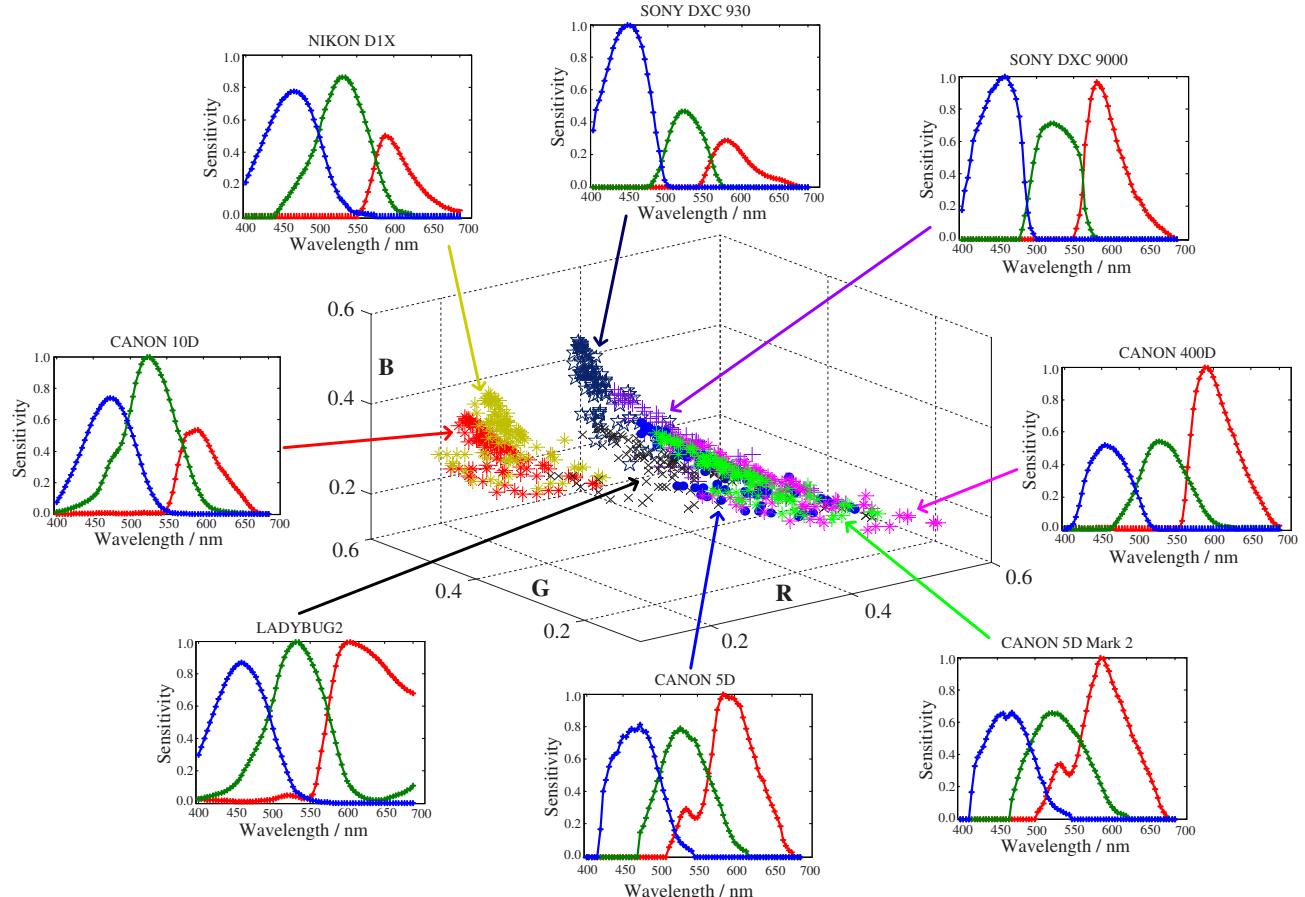
Obviously, both  $I_c$  and  $C_c(x)$  depend on the CSS  $S_c(\lambda)$ . Figure 1(b) intuitively shows that the same scene rendered under a white light source exhibits obvious color difference since two images ( $C_c(x)$ ) are respectively captured by two cameras with different CSSs. Similarly, the responses to illuminants ( $I_c$ ) under diverse CSSs also visually display in various color domains (Fig. 2).

### A. About the Learning-Based CC

We abstract the standard framework of traditional learning-based CC as

$$I_c \approx \Gamma\{K(f_c(x))\}, \quad (5)$$

where  $I_c$  indicates a set of vectors of ground truth illuminants supplied by the dataset.  $K(f_c(x))$  represents a set of vectors of



**Fig. 2.** Chromatic distributions of light source spectra [14] responding to various CSSs [24,25]. Each chromaticity represents one color domain of 108 spectra rendered under a CSS.

certain statistics extracted from the input image  $f_c(x)$ .  $\Gamma$  denotes a certain model that is committed to learn a transformation that can effectively map the extracted features of images  $K(f_c(x))$  to the corresponding illuminant ground truth  $I_c$ . Basically, for the sake of comprehensive consideration of the effectiveness and efficiency, various algorithms train their models using different machine learning techniques  $\Gamma$  and feature descriptors  $K(f_c(x))$  for CC.

For example, the classical gamut mapping method trains a model that is able to correlate the gamut of the input image to the canonical gamut under a white light source of a fixed camera [43–45]. While statistical methods relate the CC as one kind of parameter inference problem by assuming that the reflectance and illuminant meet a specific probability distribution [17,46–48], they need to train a model that can capture the mapping between the preprocessed images and the illuminant ground truth.

A large class of learning-based methods inherently treat CC as a regression problem [8,21,49]. These kinds of methods could be unified under a standard regressive equation like  $I_c \approx K(f_c(x))L$ . Here,  $L$  denotes a regression matrix that needs to be learned. The difference among the regression based models is reflected in the specific choice of the technique of learning a regression matrix  $L$  and the selection of the features  $K(f_c(x))$  adopted by each CC algorithm. Earlier approaches use nonlinear neural networks to learn a regression matrix that can map the binarized chromaticity histogram of the input image to the illuminant [20]. Similar approaches apply support vector regression [19], linear regression [49,50], or thin-plate spline interpolation techniques [22] to learn the regression matrix using almost the same type of input data. More recently, the leading performance is obtained by employing more effective features (color-edge moment [8]), more efficient regression techniques (e.g., regression trees [7]), or even deep learning [12].

Alternatively, if we replace the features  $K(f_c(x))$  with the illuminant estimated by multiple low-level based CC algorithms [5,33,39–42], the regression problem mentioned above is boiled down to the so-called committed-based CC [51–54]. In this protocol,  $L$  is no longer regarded as a feature-mapping matrix, but a weighted matrix that tries to “optimally” fuse the output of multiple CC methods as a single illuminant estimate under a certain rule [e.g., the weights are optimized in the least-mean-square (LMS) sense]. Or, in Refs. [9,10,55–57],  $L$  is taken as a vote matrix that can select the most appropriate CC method or previously stored illuminant for every input image according to the intrinsic properties of natural images, e.g., high-level information [10,51,55,56,58,59].

## B. Intra-CC

The common way to benchmark the performance of a learning-based CC algorithm is to adopt the intra-dataset-based evaluation with the form of so-called  $n$ -folds cross validation [3,8,13,34]. For example, the dataset is first divided into  $n$  parts ( $n$ -folds). Next, by applying the model on the  $n - 1$  parts, the optimal parameters and structures of the model are trained using the corresponding illuminant ground truth. Then, the trained model is tested on the remaining one part of the data. For a complete procedure of  $n$ -folds cross validation, the steps mentioned above are repeated

$n$  times to ensure that each image occurs in the test set only once and all images in the whole dataset are either in the training set or in the test set at the same time. Finally, the measures for each cross validation are averaged as the final metric of the algorithm’s performance. Figure 3(a) summarizes the steps of intra-CC. The logic behind the typical intra-dataset-based evaluation is indeed very suitable to test the performance of a learning-based CC model in the presence of the multiple scenes with diverse illuminants and reflectances in a dataset [3].

Despite the significant advancement in the performance of intra-CC on several benchmark datasets [3,7,8,11,12,22,43,44,53], these methods always implicitly assume that the distribution of the test data should be similar to the distribution of the training data (e.g., both the training images and test images are captured by a fixed camera). However, for practical CC applications this assumption may be readily violated, since various camera trademarks possess quite diverse CSSs (e.g., NIKON and CANON, in Fig. 2), and there is even an apparent difference of CSS among cameras produced by the same manufacturer (e.g., CANON in Fig. 2). Therefore, although the existing state-of-the-art algorithms that are correctly trained can achieve a high performance on intra-CC, they may also suffer serious problems once being applied on inter-CC [7,8,12,21,53].

## C. Problem Formulation for Inter-Dataset-Based CC

Figure 3(b) summarizes the general steps of the traditional inter-CC; we assume that we have trained a CC model  $\Gamma$  on a dataset that is rendered under a certain CSS (we called it CSS-1),

$$I_c^1 \approx \Gamma\{K(f_c^1(x))\}. \quad (6)$$

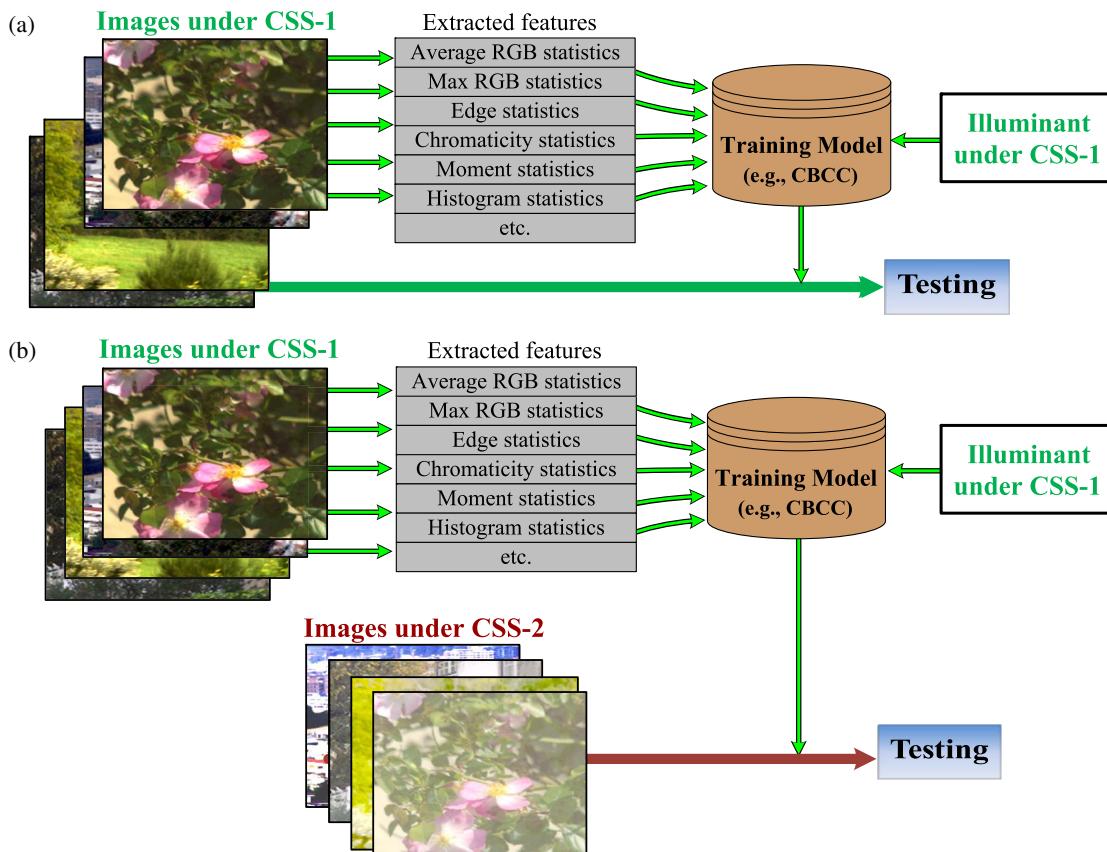
As indicated by Eqs. (3) and (4), both the illuminant  $I_c^1$  and the image  $f_c^1(x)$  are mixed with the information of CSS-1. Similarly, the color domain of the feature statistics  $K(f_c^1(x))$  extracted from  $f_c^1(x)$  is also dependent on CSS-1. Thus, based on these training data, the model  $\Gamma$  just learns a fixed mapping that is only suitable to this fixed CSS-1. In other words, the model  $\Gamma$  will always learn a mapping that attempts to predict the illuminant under a color domain of the corresponding CSS (e.g., a color domain rendered under CSS-1).

Then, to run the inter-CC, we need to test the CC model  $\Gamma$  that is trained on the color domain of CSS-1 on another dataset that is collected by another camera with a distinct CSS (we called it CSS-2),

$$I_c^2 \approx \Gamma\{K(f_c^2(x))\}, \quad (7)$$

where  $I_c^2$ ,  $f_c^2(x)$ , and  $K(f_c^2(x))$  are rendered under CSS-2. If we directly apply the CC model  $\Gamma$  on the dataset that is rendered under CSS-2, the CC model  $\Gamma$  is destined to suffer a serious failure, which is explained below.

While these learning-based CC approaches that are correctly trained can deliver very competitive performances with intra-CC, the training phase undoubtedly relies on the illuminant ground truth supplied by the dataset [3]. As analyzed above, the color domains of the images and illuminants are clearly affected by the changes of CSS [e.g., Figs. 1(b) and 2]. Thus, both  $I_c$  and  $f_c(x)$  are color-domain dependent and the existing learning-based CC models can just learn a fixed model that is only applicable to a fixed camera.



**Fig. 3.** General framework of traditional color constancy (CC) evaluation: (a) intra-dataset-based CC, (b) inter-dataset-based CC.

In other words, the existing learning-based CC algorithms always train a model that tends to predict an illuminant under a specific color domain (e.g., a color domain rendered under a specific CSS). Consequently, once the trained CC model is tested on the images that are captured by another camera with very different CSS, it is destined to suffer failure since they do not consider the effects of CSS during the training.

### 3. IMPROVING THE INTER-DATASET-BASED CC

In the above, we have provided a theoretical analysis of the effect that a CSS has on the chromatic distribution of light source spectra (Fig. 2) and the color appearance of images [Fig. 1(b)], which further indicates that the existing CC models will perform poorly once being tested with inter-CC.

In this section, we propose a method to improve the performance of a CC algorithm on inter-CC evaluation by taking into account the CSS of each camera. Our strategy is straightforward. Specifically, in order to overcome the problem due to the difference between the two CSSs during the process of inter-CC, we first train a sensor transformation that can approximately express the mapping between the two given CSSs, and then we use this mapping to relate the CC model trained on one dataset with CSS-1 to another dataset with CSS-2. The improved framework for CC on inter-CC application is illustrated in Fig. 4.

As indicated by previous work [60], the sensor transformation we are concerned with can be well approximated by a simple

$3 \times 3$  matrix, though such a mapping between two CSSs necessarily loses information [29]. If  $\text{fr}^1(x)$  denotes the response of CSS-1 to a certain reflectance, and  $\text{fr}^2(x)$  denotes the response of CSS-2 to the same reflectance, we define

$$\text{fr}^2(x) = S \cdot \text{fr}^1(x), \quad (8)$$

where  $S$  is a  $3 \times 3$  matrix, which is simply learned by a LMS training technique in this work. In order to learn the matrix  $S$  based on Eq. (8), we always use the 1995 reflectance spectra compiled from the Simon Fraser University (SFU) hyperspectral dataset [14] as the input to produce the responses  $\text{fr}^1(x)$  and  $\text{fr}^2(x)$  based on Eq. (1), which produces the matrix size of  $1995 \times 3$  for both  $\text{fr}^1(x)$  and  $\text{fr}^2(x)$ . The procedure of learning  $S$  is shown in Fig. 5.

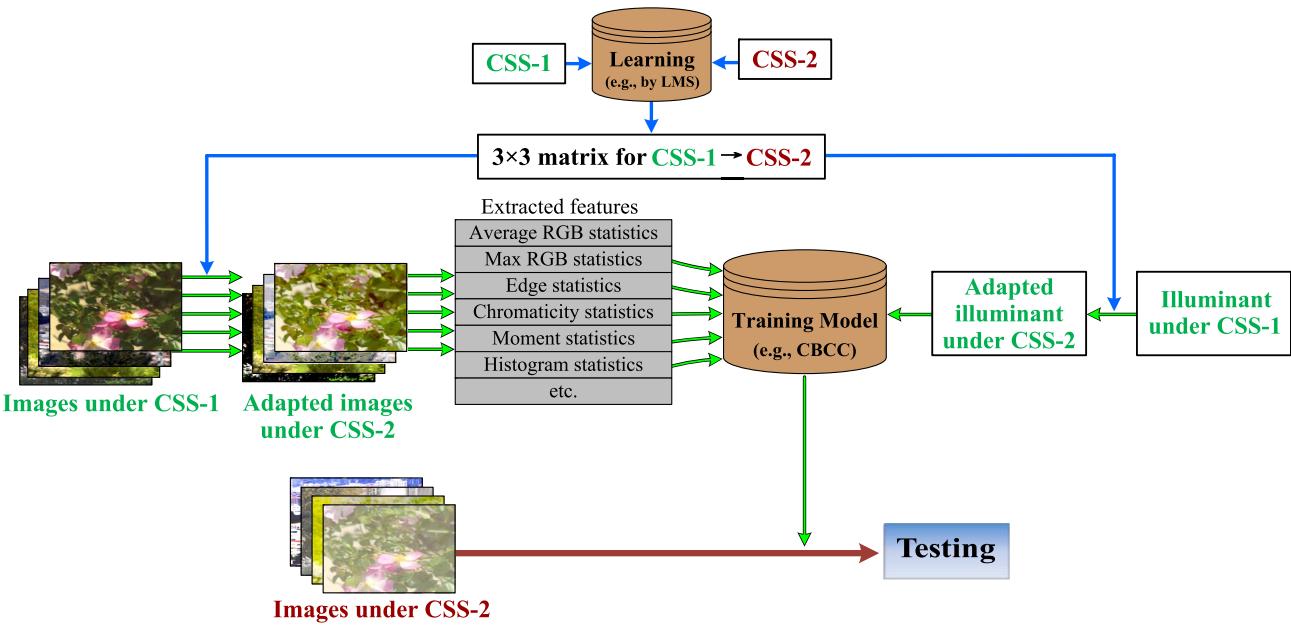
For the inter-CC, this matrix  $S$  is utilized to transform both the image  $f^1(x)$  and illuminant  $I_c^1$  rendered under CSS-1 into CSS-2. After this transformation, the CC model  $\Gamma$  is then trained on this transformed data:

$$S \cdot I_c^1 \approx \Gamma\{K(S \cdot f^1(x))\}. \quad (9)$$

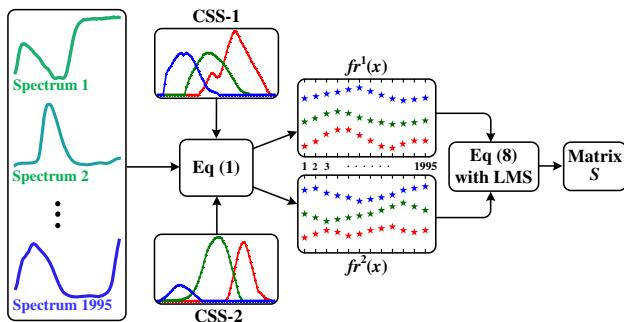
Finally, the trained model is tested on the images rendered under CSS-2 for the inter-CC, i.e., to obtain the illuminant  $I_c^2$  according to

$$I_c^2 \approx \Gamma\{K(f^2(x))\}. \quad (10)$$

It is worth stressing that learning this  $3 \times 3$  matrix  $S$  is very easy and needs only two matrices with the size  $1995 \times 3$  (as indicated by Eq. (8) and Fig. 5), which is normally far less than



**Fig. 4.** General framework of the proposed inter-dataset-based CC improved by taking the CSS difference between datasets into account during the CC model training.



**Fig. 5.** Procedure of learning the transformation matrix  $S$  between CSS-1 and CSS-2 based on the 1995 reflectance spectra.

the number of images required to train an intra-CC model for each camera. Moreover, we have experimentally found that utilizing the same learned  $3 \times 3$  matrix is enough to capture the transformation between two different CSSs for either a Mondrian-like, hyperspectral, or real RGB dataset. Such an attribute is very important since this indicates that the learned CSS adaptation matrix works independently of the scene. This makes our strategy widely applicable, since we can learn a transformation matrix for each camera for later use based on the public surface reflectance samples (e.g., the 1995 reflectance spectra [14] used in this study) and the known CSS curves, which can save much time and effort to prepare the training set for new cameras (e.g., manually label numerous images under various environments and illuminants for each camera).

#### 4. IMPLEMENTATIONS AND EXPERIMENTS

In this section, we first measure the effect of CSS on the inter-CC performance with synthetic Mondrian and hyperspectral

natural images. Then, the proposed method is comprehensively compared and validated on hyperspectral and real images coming from cameras. Finally, we exhibit several examples to show how a stable color representation of an image that is invariant to both the illuminant and CSS could be obtained with the proposed technique.

We select the early committee-based CC model (CBCC) [53] as the representative of the learning-based CC models to demonstrate the CSS effect. Basically, there are mainly two considerations when choosing CBCC as the example. On the one hand, CBCC can be almost taken as the earliest prototype of the regression-based CC models and thus many of the existing state-of-the-art learning-based CC models can be included in its framework [7,8,10,21,49,55,56]. On the other hand, recent work [7,8,61] has indicated that with much simpler implementation, CBCC can actually lead quite a competitive performance in comparison to those more sophisticated learning-based CC methods by incorporating with more effective features (e.g., the color edge moments used in corrected-moment illuminant estimation (CM) [8]). During our implementation, the outputs of the gray-world and gray-edge models were integrated into the framework of CBCC to train the regression model in the LMS sense [8]. The “cross terms” employed in CM [8] that have been demonstrated to be very important for delivering the best CC result were also utilized in our implementation.

We also implement both the CM [8] and spectral sharpening (SS) techniques [31,62] to test their performance on inter-CC, for comparison. CM is one of the state-of-the-art CC algorithms that improves the CC performance by learning a fixed matrix to correct the biased illuminant estimates of some low-level-based CC algorithms. In our implementation, we use 9-edge-moments-based CM. The aim when comparing with the CM is to show if it is more convenient to correct the

CSSs before applying CC algorithms (proposed) than to correct the illuminant estimates after applying CC algorithms (CM).

Spectral sharpening was originally proposed to sharpen the CSS such that each sharpened CSS has its spectral sensitivity concentrated as much as possible within a narrow band of wavelengths. Hypothetically, we expect that the sharpened CSSs could be more similar to each other and thus can improve the inter-CC performance of CBCC. Specifically, for a dataset CSS-1, we first convert the original sensors to its sharpened version via a  $3 \times 3$  matrix multiplication and then learn the CBCC on the sharpened CSS-1. Then, given a second dataset CSS-2, we also perform spectral sharpening on the sensors of CSS-2 and finally test the performance of the learned CBCC of CSS-1 on the sharpened images of CSS-2.

### A. Validation on Mondrian-Like Images

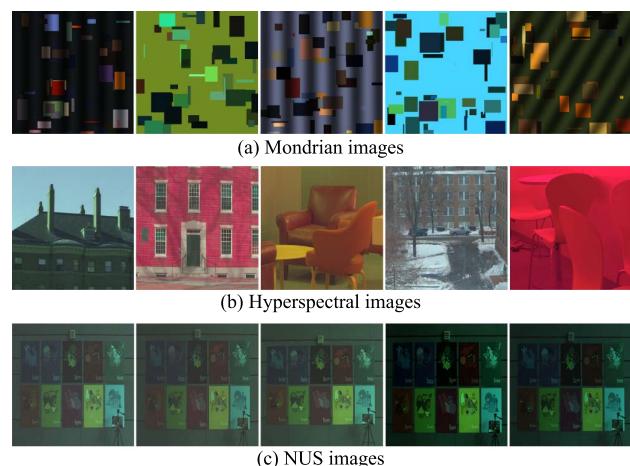
In this experiment we used a dataset containing 1995 spectra of reflectances and 102 spectra of light sources compiled from several sources [14] for the generation of Mondrian-like images. These spectra of reflectances and light sources were carefully collected under both man-made and natural environments. To study the CSS effect, we used a recently published database [24,25], which includes 12 sensors with various CSSs ranging from a common consumer level camera to an industrial camera (several CSSs are shown in Fig. 2). Moreover, the CIE color matching function [63] and standard sRGB function [64] are also included as a kind of CSS during the test of the CSS effect on CC.

We generated the Mondrian-like datasets according to the model of image formation described by Eq. (1). Specifically, the spectra of reflectance and illuminant [14] are first randomly selected and then integrated with the CSS [24] over all the visible spectra to obtain R, G, B values. In practical computation, all spectrum curves are sampled and represented as vectors. Then, using these generated pixel colors, Mondrian-like images are created, and each image randomly contains up to tens of different surfaces, and hence many different transitions. The Gabor grating tuned brightness variations are also added in each Mondrian-like image for simulating as close as possible to the luminance of real-world images [10,34,44,65]. Several examples are shown in Fig. 6(a).

We totally synthesized 14 Mondrian image datasets based on each CSS of 14 sensors [24]. In each dataset, there are totally 510 Mondrian-like images (with a size of  $300 \times 300$  pixels) composited by random choices of illuminants and reflectances by using the aforementioned rules, and hence can simulate the situations of the multiple scenes with diverse illuminants. Thus, each dataset alone could be utilized to test the performance of intra-CC.

Moreover, among 14 Mondrian datasets, the scenes in each dataset are exactly same for all the cameras, but rendered by distinct CSSs. In other words, the only difference among the datasets is the CSS. Therefore, any pair of datasets arbitrarily selected from the Mondrian datasets could be used to accurately measure the CSS effect on the performance of inter-CC. For example, we trained CBCC on one Mondrian dataset that is specifically rendered under a certain CSS (e.g., CANON 5D Mark), and then tested the trained CBCC on another Mondrian dataset that is captured by a different CSS (e.g., NIKON D70).

Figure 7 summarizes the averaged performance of intra-dataset-based cross validation of CBCC on various Mondrian



**Fig. 6.** Examples of (a) synthetic Mondrian-like, (b) hyperspectral, and (c) real camera captured images.

datasets (each dataset is named based on the used camera type). Angular error is usually used to test the accuracy of a CC algorithm [3] by measuring the chromatic difference between the illuminant ground truth and the estimated illuminant by the CC algorithm.

Figure 7 indicates that on all of the Mondrian datasets tested here, CBCC indeed obtains very good intra-CC performance. However, as discussed earlier, such evaluation is only suitable to fairly benchmark the performance of a CC method with the presence of multiple scenes with diverse illuminants. In contrast, for inter-dataset-based evaluation, CBCC seriously decreases performance due to the CSS effects.

From the above evaluation on these 14 Mondrian datasets, we did not observe that there is a CSS that is more effective than others for CC, and for the inter-dataset-based evaluation, we can observe based on the measurement of angular error in Fig. 7 that the more similar the two CSSs are [e.g., the CSS between camera CANON 5D Mark and CANON 5D in Fig. 7(e)], the less impact of CSS on the accuracy of either intra- or inter-dataset-based CBCC. In contrast, the more distinct the two CSSs are [e.g., the CANON 5D Mark and SONY DSC 930 in Fig. 7(a)], the worse the performance of the original inter-dataset-based CBCC is. In short, the performance of the traditional inter-CC relies greatly on the similarity of CSSs among cameras. For example, when applying a model on a dataset that is trained on the images captured by a very distinct CSS [Fig. 7(c)], very bad inter-CC performance is achieved. It is worth noting that for real camera captured images, the degradation caused by CSS is more complicated than the Mondrian situation. We will further discuss this point in the following experiments using real camera captured images.

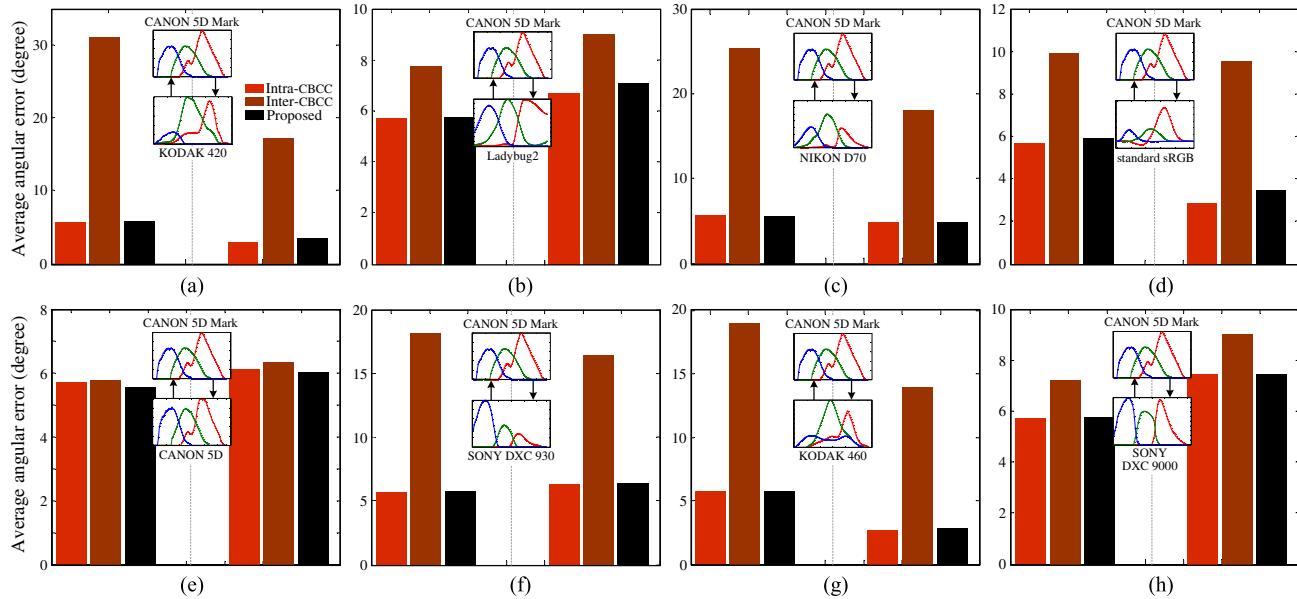
### B. Validation on Hyperspectral Images

While synthesizing Mondrian images according to Eq. (1) is a pretty accurate first-order model of image formation [8,66], it is not able to model other reflective effects (e.g., specular component) and thus may not reflect the real color image formation process. Thus, we also utilized the hyperspectral natural images to measure the CSS effect on the performance of inter-CC. For this experiment, we used a dataset containing 77 high-quality

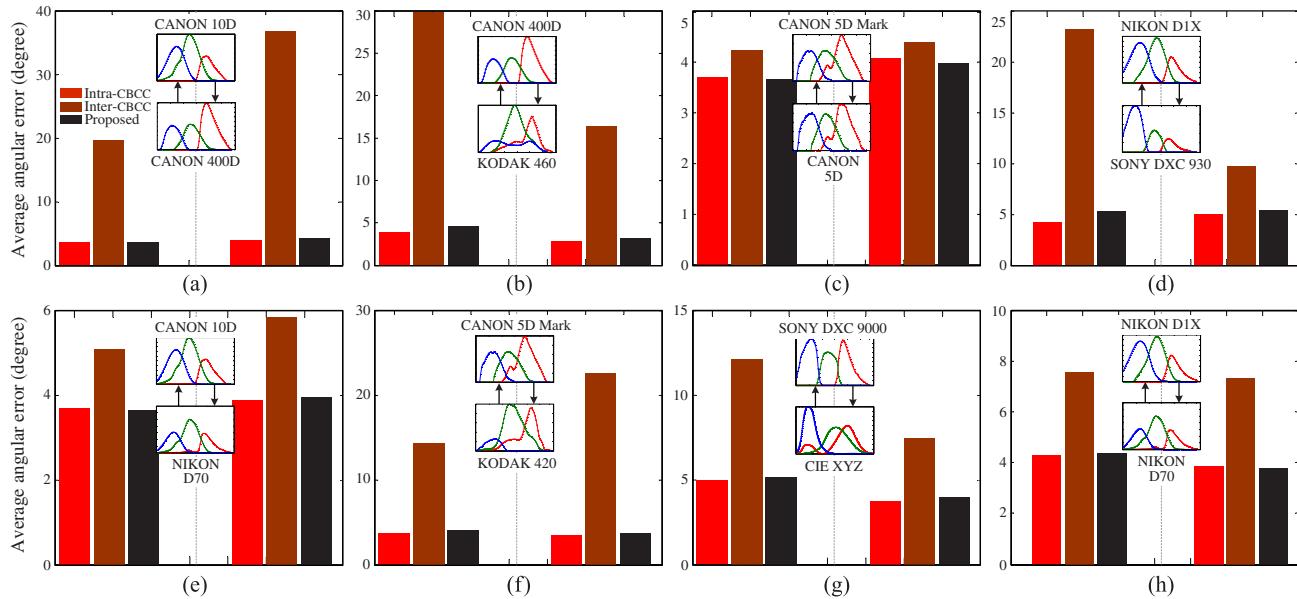
hyperspectral images (with a size of  $1392 \times 1040$  pixels) acquired in real indoor and outdoor scenes [67].

Figure 6(b) shows several examples of hyperspectral images employed in these experiments. The subset of the previous light source spectra dataset that contains 11 both Planckian and non-Planckian [14] illuminants and the 14 CSSs mentioned above were employed to produce 14 hyperspectral datasets. Each

hyperspectral dataset totally contains 847 hyperspectral natural images rendered under multiple illuminants for various investigations of the color appearance of real-world scenes. Similar to the situation of Mondrian scenes, among the 14 hyperspectral datasets, each one possesses identical distributions of reflectance and illuminant but exclusively rendered by distinct CSSs. Figure 8 shows the results of CBCC on both intra- and



**Fig. 7.** Results of CBCC on the 14 Mondrian datasets. Each subgraph contains the results of intra-dataset-based CC (*intra-CBCC*) and inter-dataset-based CC (*inter-CBCC*) without considering the CSS effect, and the improved inter-dataset-based CC by taking the CSS into account (*proposed*) on two datasets (the left three bars are on one dataset and the right on another dataset) with different CSSs (the CSS curves of the top and bottom panels are for the left and right datasets, respectively). In particular, the “inter” bar among the left three is the result of training on the right dataset and testing on the left dataset, and vice versa. The notation of the “intra” bars is similar to that of the “inter” bars. Note that the intra-dataset-based (*intra-CBCC*) results listed here were obtained using three-fold cross-validation, and the reported results were averaged over 100 random repetitions.



**Fig. 8.** Results of CBCC evaluated on the 14 hyperspectral datasets. The conventions are the same as those in Fig. 7.

inter-dataset-based evaluation. The observations on these results are quite consistent with those obtained in the previous experiments with synthetic Mondrian datasets shown in Fig. 7. As a comparison to the experiments shown in Fig. 8, Fig. 9 shows more results with the methods of CM and SS for inter-dataset-based evaluation. Similar to CBCC, the-state-of-the-art CM and early SS also suffer from the drastic decreasing of performance on inter-dataset-based evaluation. The reason why CM fails to achieve a good performance for inter-dataset-based evaluation is that, similar to CBCC, i.e., CM just learns a fixed matrix for a specific camera [8]. In other words, CM is not able to adapt to other cameras. As for the SS method, although the sharpened CSSs are more similar to each other, the actual performance of inter-dataset-based CBCC is surprisingly poor, since we observed that the sharpened CSSs include many negative values, which may finally result in the large angular error. It should be noted that this problem may be alleviated when using the technique of spectral sharpening with only positivity [68].

### C. Improving the Inter-Dataset-Based CC

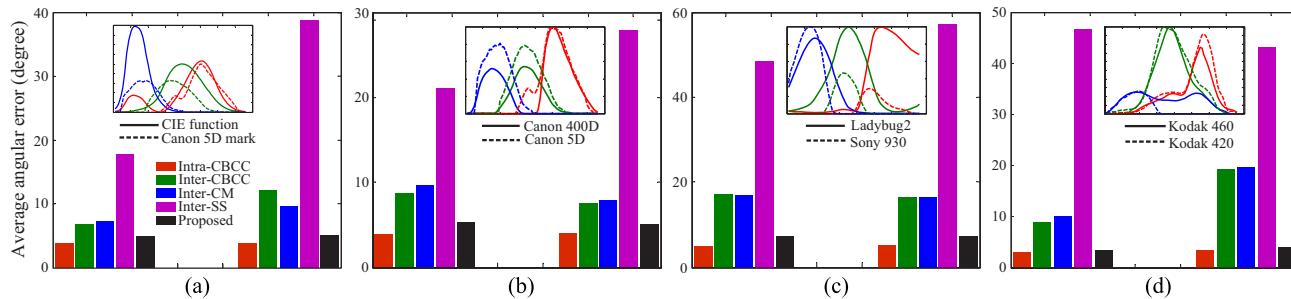
In the above sections, we have provided a systematic analysis of the CSS effect on the color appearance of images (Fig. 2) and the chromatic distribution of light source spectra [Fig. 1(b)], and shown how a CSS can significantly affect the performance of a learning-based CC algorithm on inter-dataset-based evaluation (Figs. 7–9). In this section, we will show the improved performance of inter-CC by the proposed method. The improved performance of CBCC on both Mondrian and hyperspectral datasets is shown in Figs. 7–9, respectively. We can observe that the performance of CBCC on inter-dataset-based

evaluation is greatly boosted by the proposed strategy that includes the CSS effect during the model training. Surprisingly, by employing the proposed strategy, the improved performance of CBCC on inter-dataset-based evaluation almost reaches the same level as that of the CBCC on intra-dataset-based evaluation.

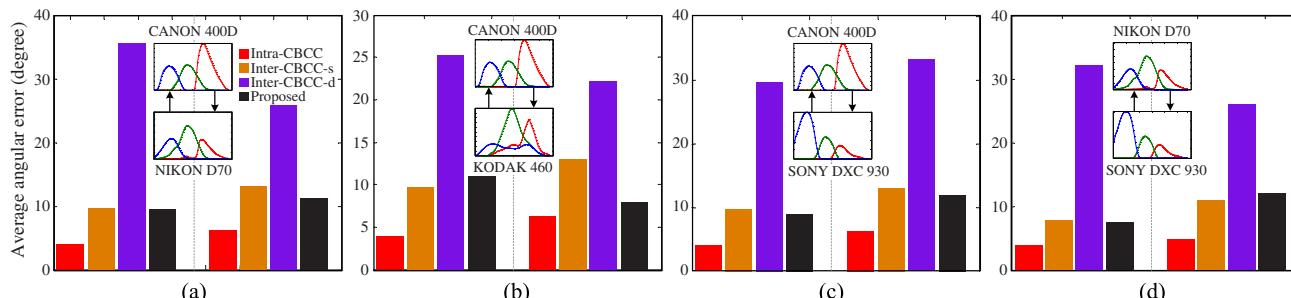
### D. More General Situations

We have shown how to improve the inter-CC by the proposed method on datasets that have exactly the same reflectance distributions but with distinct CSSs (Figs. 7–9). For practical application, however, the general situation is that the training set not only has distinct CSSs but also possesses different reflectance distributions compared to the test set (e.g., the training dataset is captured under an indoor environment, but the test dataset is collected under an outdoor natural environment). In this section, we measure the CSS effect on such a more general situation.

Specifically, we trained the CBCC model on the Mondrian-like dataset and then tested it on the hyperspectral natural dataset. We also trained the CBCC on the hyperspectral dataset and tested it on a real RGB dataset [14,69] (SFU lab dataset) that is captured by SONY DSC 930. In these experiments, we evaluated the CBCC under four situations including intra-dataset-based CC (intra-CBCC), inter-dataset-based CC with same CSSs (labeled as inter-CBCC-s), inter-dataset-based CC with distinct CSSs (labeled as inter-CBCC-d), and finally the improved model (proposed). Figure 10 shows the results of intra-based CBCC on the Mondrian dataset and the results of inter-based CBCC by training on the Mondrian set and testing on the hyperspectral dataset. Similarly, Fig. 11 shows the



**Fig. 9.** Results of CBCC, CM, SS, and the proposed method evaluated on the hyperspectral datasets. The insets show the CSSs of two subsets.



**Fig. 10.** Results of intra-based CBCC on the Mondrian dataset and the results of inter-based CBCC by training on the Mondrian set and testing on the hyperspectral dataset. Each subgraph contains the results of intra-CC (*intra-CBCC*), inter-CC with same CCSs (*inter-CBCC-s*), inter-CC with distinct CCSs (*inter-CBCC-d*), and the improved inter-CC (*proposed*).

results of intra-based CBCC on the hyperspectral dataset and the results of inter-based CBCC by training on the hyperspectral set and testing on the real SFU lab dataset.

In general, the experimental results for each situation behave as expected; CBCC obtains very good CC performance of intra-dataset-based evaluation on all the datasets rendered under any CSSs, illuminants, and scenes. For the situation of inter-CC, intuitively, the greater the difference between the training and the test datasets, the worse the performance CBCC obtains. For example, the performance of inter-based CBCC with the same CSSs (inter-CBCC-s) is worse than its performance on the intra-CC (intra-CBCC), since the two datasets utilized for training and testing have different reflectance distributions (e.g., the scene structure in the hyperspectral dataset is very different from the scene structure in the SFU lab dataset). The worst performance of CBCC is obtained for the inter-CC with distinct CSSs (inter-CBCC-d) because there is not only the huge difference of reflectance distributions but also the huge difference of CCSs between the two datasets. This experiment once again demonstrates the adverse effect of CSS on the performance of a learning-based CC algorithm for the inter-CC application.

It should be noted that the improved performance of the inter-based CBCC with the same CSSs (inter-CBCC-s) does not arrive at the same performance level of the intra-CC, as shown in Figs. 7–9. This is mainly due to the huge difference of the reflectance distribution between the two datasets evaluated here (e.g., the difference between the natural scene in the hyperspectral dataset and the laboratory scene in the SFU lab dataset). Nevertheless, the performance of the inter-based CBCC with different CSSs (inter-CBCC-d) is greatly improved by the proposed strategy.

## E. Validation on Real Camera Captured Images

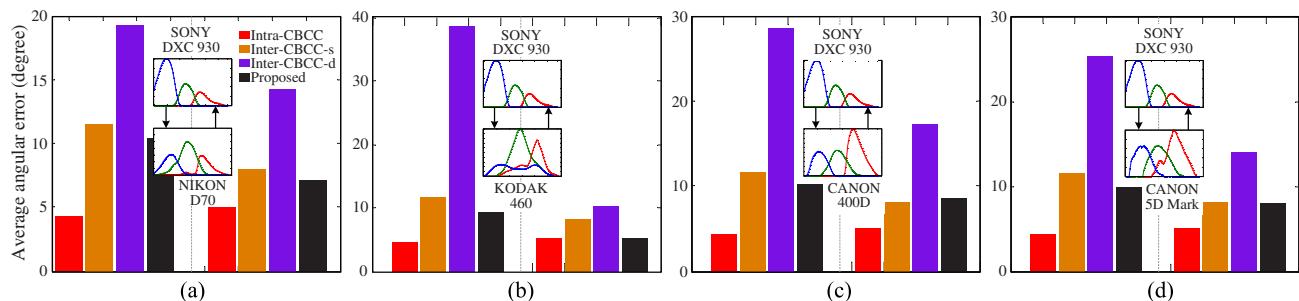
All of the above experiments used the synthetic images or multispectral images, which ignore all the nonlinearities that occur in digital camera pipelines before the RAW image is saved. To validate the applicability of the proposed method in a real-world scenario, we finally tested our proposed method on the National University of Singapore (NUS) dataset [15], which includes 1853 high-quality linear images taken by 9 different cameras in real environments. In addition, this dataset is composed of images of the same scene with the different cameras, such that the scene and illumination are the same for all 9 of the cameras. This makes the dataset suitable for investigating

the CSS effect on inter-CC under a real camera situation. One example of the same scene taken by 5 different cameras is shown in Fig. 6(c). For unbiased evaluation, during experiments we masked out the color checker patch that is originally used to compute the illuminant ground truth of each image. Since SS performed very poorly in previous experiments (Fig. 9), we only compared the improved results of CBCC with CM.

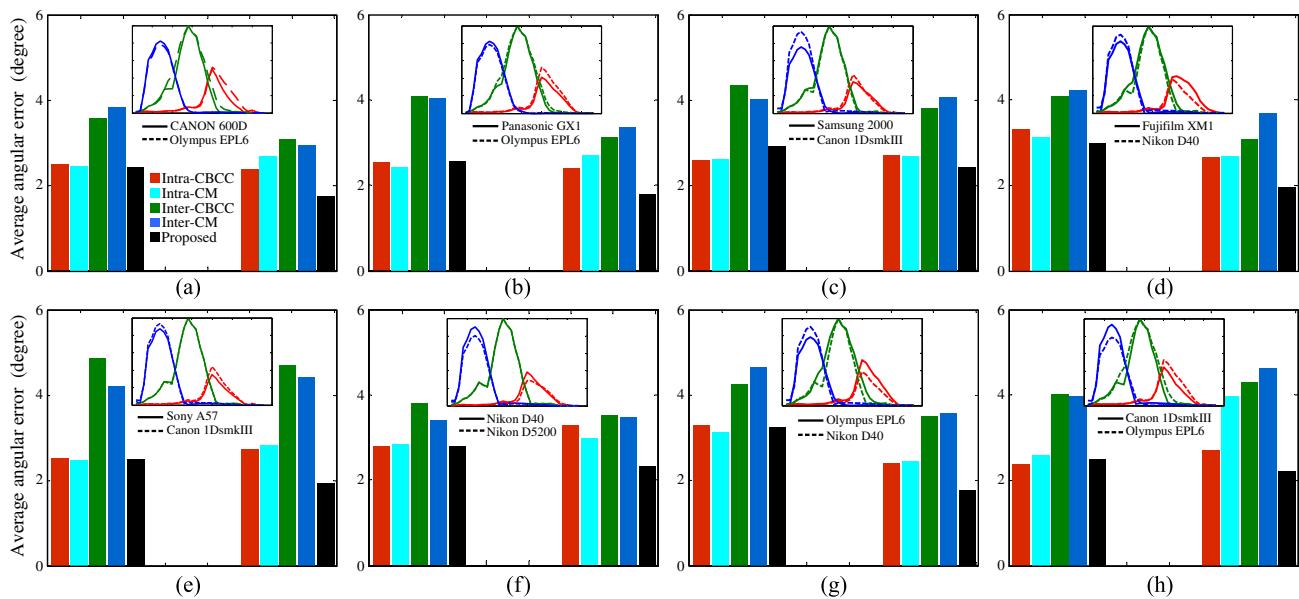
Similar to what we have reported on the synthesized Mondrian-like and hyperspectral images, Fig. 12 shows the results on several NUS subsets captured with various cameras. As expected, CBCC and CM perform very well on intra-dataset-based evaluation. However, once being applied in the inter-CC setup, the performance of both methods is clearly degraded due to the effect of CSS, and in general, the larger the difference between the CSSs is, the worse the inter-CC performs for CBCC and CM.

It should be noted that all the cameras utilized for capturing the NUS dataset have very similar CSSs since they are specifically chosen to satisfy the so-called Luther condition for CSS approximation [70]. Hence, the results also indicate that although such a small difference exists among different CSSs (cameras), which indeed leads to the drastic performance decreasing for the existing learning-based CC models on inter-dataset-based evaluation. Thus, these experiments on the dataset from real cameras further ground our theoretical analysis in Subsection 2.C. In short, in order to develop a CC algorithm with a good generalization ability, we need to reasonably consider the CSS effect.

We noticed that with quite similar CSSs between two cameras for the NUS dataset [e.g., Fig. 12(e)], the inter-CC performance by CBCC and CM are still greatly degraded in comparison to the intra-CC performance. Before a definite reason can be found to explain this phenomenon, we reasonably speculate that based on Eq. (1), besides the CSS, there are other mixed factors, e.g., the distributions of the illuminant and surface reflectance, determining the color appearance of a captured image, and consequently, the inter-CC performance should also be influenced at least by these two factors. Let us take Fig. 12(a) as an example. If we make a mutual change of the training and test sets, the performance of the inter-CBCC varies markedly, as indicated by the green bars on the left and right parts of Fig. 12(a). Similar observations can also be found for the inter-CC evaluations on the Mondrian and hyperspectral datasets (Figs. 7–11). This clearly indicates that besides the difference between the CSSs, the



**Fig. 11.** Results of intra-based CBCC on the hyperspectral dataset and the results of inter-based CBCC by training on the hyperspectral set and testing on the real SFU lab dataset. See similar notations to the bars as indicated in Fig. 10.



**Fig. 12.** Results of CBCC, CM, and the proposed method evaluated on the NUS dataset [15].

inter-CBCC performance is affected determinately by the difference between the training and test sets. This may partially explain the poor inter-CC performance on the two datasets even with the visual similarity of CCS. In fact, such phenomenon further emphasizes the difficulty of color constancy in real applications, where the difference between the inherent features of the training and test images is uncontrollable.

In order to further determine whether the proposed method significantly improves the performance of CBCC and CM on inter-CC, we further utilized the Wilcoxon Signed-Rank Test (WST) to measure the performance difference between the inter-CC (e.g., inter-CBCC and inter-CM) and our method. WST has been recommended as a valuable tool for performance evaluation [3,34,71]. In this study, given any two different algorithms, the WST was run on their angular error distributions on the whole dataset, and its result was used to conclude that at a specific (e.g., 95%) confidence level, the angular error of one algorithm is often lower or higher than that of another algorithm, or there is no significant difference between them.

Figure 13 reports the results of the statistical significance test using WST on all cameras. A sign (+) at location  $(i,j)$  indicates that the average angular error of method  $i$  is significantly lower than that of method  $j$  at the 95% confidence level, and a sign (-) at  $(i,j)$  indicates the opposite situation. A sign (=) means that the average angular errors of the two methods have no significant difference. Figure 13 indicates that the performance of the proposed method exhibits a significant improvement over both CBCC and CM in the most situations of inter-CC evaluation.

#### F. Stable Color Representation of Images

In the previous sections, we have shown the improvement of inter-CC by incorporating CSS information. Here, we show examples of actual improvement on image appearance by further considering the CSS after CC. Even though the two images in Fig. 1(b) have been corrected by CC, they still exhibit obviously

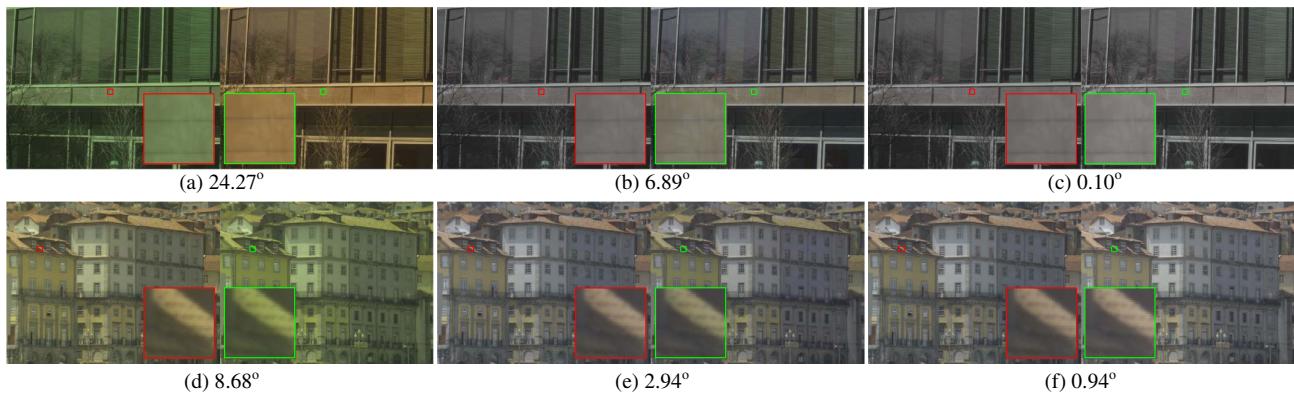
different color appearance due to the difference between CSSs. In order to eliminate the chromatic difference induced by CSSs, the original color-biased images taken by the two CSSs were respectively corrected by a CC algorithm (here by the CC method of gray-edge [41]), then a transformation by the learned matrix was further used to adapt the corrected image rendered under

	1) 1D	2) 600D	3) XM1	4) D40	5) D5200	6) EPL6	7) GX1	8) NX2000	9) A57
Proposed	=	=	-	==	-	-	-	-	-
	=	=	-	=	-	-	-	-	-
	+	+	=	-	-	-	-	-	-
	=	=	+	==	-	-	==	-	-
	+	+	+	+	=	-	-	=	=
	+	+	+	+	+	=	-	-	-
	+	+	+	=	+	+	==	-	-
	+	+	+	+	=	+	+	=	-
	+	+	+	+	=	+	+	+	=

	1) 1D	2) 600D	3) XM1	4) D40	5) D5200	6) EPL6	7) GX1	8) NX2000	9) A57
Proposed	=	=	-	-	-	-	-	-	-
	=	=	-	==	-	-	==	-	-
	+	+	=	-	-	-	-	-	-
	+	=	+	==	-	-	==	-	-
	+	+	+	+	=	-	-	=	=
	+	+	+	+	+	=	-	-	-
	+	=	+	+	+	+	==	-	-
	+	+	+	+	=	+	+	=	=
	+	+	+	+	=	+	+	=	=

**Fig. 13.** WST test between the proposed method and the inter-CBCC (top), and inter-CM (bottom) in the NUS dataset.



**Fig. 14.** Results on two pairs of images from [67,72]. The first column lists two pairs of scenes rendered under different illuminants and CSSs. The second column shows the images after removing the effect of the illuminant. The last column lists the results after further discounting the CSS effect.

one CSS (e.g., KODAK DCS 460) to another CSS (e.g., CANON 5D Mark 2).

Figure 14 shows examples selected from the two hyperspectral datasets [67,72]. It is clear that the colors cast on the image pairs due to the different CSSs and illuminants are well removed after applying a CC algorithm and then eliminating the CSS effect. For example, Fig. 14(a) shows obvious chromatic difference between the two images since they are rendered under distinct illuminants and CSSs. By applying CC on these two images, although the color bias induced by the illuminants is removed [Fig. 14(b)], these two images still present a significant chromatic difference since they are rendered under two different CSSs. However, the chromatic difference of these two images is further eliminated after removing the CSS effect by the proposed method [Fig. 14(c)].

It should be noted that several CC methods have been proposed to impose stabilization on the resulting color-corrected images through specific constraints. For example, in Ref. [57] Sapiro uses the probabilistic Hough transform to introduce the physical constraints that the corrected solution should accomplish; in Ref. [45] Finlayson *et al.* select the illuminant from a predefined subset that better correlates with a set of given reflectances, and in Ref. [59] Vazquez-Corral *et al.* select the illuminant that maximizes the number of psychophysically based focal colors presented in the corrected scene. However, the main aim of those constraints is to improve the accuracy of illuminant estimation instead of explicitly removing the color bias triggered by CSS. In other words, the results of these algorithms is still CSS dependent. It is also worth noting the specific idea of color stabilization recently proposed in the literature [73–75]. Given two images that share some content, this idea aims at converting the colors of one of the images to those of the other. Therefore, when the only difference between the two images is the CSS, a color stabilization method would convert the CSS of the source image to the CSS of the second image. However, it is obvious that discounting both effects of CSS and illumination falls beyond the aim of these color stabilization methods.

## 5. DISCUSSION AND CONCLUSION

In this paper, we demonstrated that CCS is a quite important factor that results in the serious degeneration of learning-based

CC algorithms in inter-dataset-based application due to both the illuminant and image are actually color-domain dependent (camera dependent). It means that the existing CC algorithms only build very limited color-correction models that are only suitable for a specific camera. Thus, once applying the trained CC model on the images collected by a camera with a distinct CSS, the model will undoubtedly suffer serious problems. Obviously, CSS properties should be incorporated into any general account of computational color constancy [29]. We have grounded our theoretical analysis on experiments with various datasets.

We also proposed a simple yet effective way to incorporate the CSS information into the training process through an adapted matrix, which builds a mapping between two CSSs. As a direct consequence, we can model the CSS effect when building a CC model. Comprehensive experiments on synthetic, hyperspectral, and real camera captured datasets have shown how the embedded CSS information can greatly improve the performance of learning-based CC algorithms on inter-dataset-based evaluation. Many existing state-of-the-art learning-based CC methods will benefit from our technique to improve their generalization ability across multiple cameras.

As a practical application, we have also shown examples to demonstrate how a stable color representation of a scene across the changes of illuminants and CSSs could be obtained by the proposed method. This is an important step toward the normalization of color appearance under different acquisitions with different devices.

As a representative example of calibrating the CSS effect, it is a common practice in the application of 3-D view synthesis to calibrate the effect of CSS before integrating the images captured by different cameras from different view angles. However, their primary goal is to minimize the color differences between multi-camera images to achieve inter-camera color consistency, i.e., a point on an object has the same color when being observed in different views. Mostly, this “same color” is not the TRUE color of the point, especially for its common solution, i.e., minimizing the color differences of images between a user-defined reference camera and all other cameras [76,77]. Though, theoretically, its color pattern board-based solution (e.g., [78,79]) can achieve inter-camera color consistency

by realizing color constancy among different views, it is not a practical way to provide a color pattern in various imaging conditions.

In short, if we call the scene color recovered by previous CSS calibration techniques or recovered by previous CC methods as the *relative color*, we can call the color recovered by our proposed technique as the *absolute color*, since it really reflects the constant color of the scene through eliminating both the effects of the CSSs and the illuminant.

The results presented in this work prove very useful for many color-based applications. Imagine we are facing a surveillance task using a large-scale camera network [80] and we want to build a CC model for this camera network [81]. For traditional implementation, we have to first capture a large group of images under diverse illuminants for each camera, and then train the CC model separately for each camera utilizing the corresponding manually labeled images and illuminants. This will not be practical in this large-scale camera network, since manually labeling many images under various illuminants for each camera is very expensive and time consuming. However, by employing the solution proposed in this work, it is enough to train a CC algorithm for any camera with only the information of different CSSs and the images under various illuminants captured by only one camera, since we can adapt a CC algorithm between any two cameras by taking the difference between CSSs into account.

In this work, we assume that the CSS utilized to calculate the adapted matrix between two CSSs is already known. This is practically applicable since such information is mostly publicly available for many camera sensors (e.g., [82]) [26], or we can estimate it using existing CSS estimation algorithms [24,25,36,70,83].

Another point that deserves a brief comment is that all the analysis in this work is true only when we are dealing with RAW images. We think this limitation is acceptable, since as a more logical option color constancy is generally implemented prior to the conversion of the linear raw data to sRGB images [3,84], and other embedded postprocessing steps (e.g., Gamma correction) are not needed to be taken into account.

This work also initiates an open question on how to build a general CC model that could be applied for any cameras and situations. Our proposed method is good at dealing with the failure of the inter-CC situation caused by different CSSs. However, during the experiments we also observed that there are other factors (e.g., scene content) that also affect the performance of inter-CC. Thus, it is very important to employ other useful information while building a CC model with a higher efficiency and universality, which is our future work.

**Funding.** Major State Basic Research Program (2013CB329401); National Science Foundation (NSF) (91420105, 61375115); 111 Project of China (B12027).

**Acknowledgment.** We would like to thank Dr. Dilip Kumar Prasad and Michael S. Brown for providing their CSSs of the NUS dataset and Dr. Javier Vazquez-Corral for sharing the code of spectral sharpening.

## REFERENCES

1. M. Vrhel, E. Saber, and H. J. Trussell, "Color image generation and display technologies," *IEEE Signal Process. Mag.* **22**, 23–33 (2005).
2. B. Funt, K. Barnard, and L. Martin, "Is machine colour constancy good enough?" in *European Conference on Computer Vision (ECCV)* (Springer, 1998), pp. 445–459.
3. A. Gijsenij, T. Gevers, and J. Van De Weijer, "Computational color constancy: Survey and experiments," *IEEE Trans. Image Process.* **20**, 2475–2489 (2011).
4. D. H. Foster, "Color constancy," *Vis. Res.* **51**, 674–700 (2011).
5. S. Gao, K. Yang, C. Li, and Y. Li, "A color constancy model with double-opponency mechanisms," in *IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2013), pp. 929–936.
6. S. D. Hordley, "Scene illuminant estimation: Past, present, and future," *Color Res. Appl.* **31**, 303–314 (2006).
7. D. Cheng, B. Price, S. Cohen, and M. S. Brown, "Effective learning-based illuminant estimation using simple features," in *IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 1000–1008.
8. G. D. Finlayson, "Corrected-moment illuminant estimation," in *IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2013), pp. 1904–1911.
9. H. R. V. Jozé and M. S. Drew, "Exemplar-based color constancy and multiple illumination," *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 860–873 (2014).
10. A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 687–698 (2011).
11. S. Bianco and R. Schettini, "Adaptive color constancy using faces," *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 1505–1518 (2014).
12. S. Bianco, C. Cusano, and R. Schettini, "Color constancy using cnns," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2015), pp. 81–89.
13. B. Li, W. Xiong, W. Hu, B. Funt, and J. Xing, "Multi-cue illumination estimation via a tree-structured group joint sparse representation," *International Journal of Computer Vision* **117**, 21–47 (2016).
14. K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Res. Appl.* **27**, 147–151 (2002).
15. D. Cheng, D. K. Prasad, and M. S. Brown, "Illuminant estimation for color constancy: Why spatial-domain methods work and the role of the color distribution," *J. Opt. Soc. Am. A* **31**, 1049–1058 (2014).
16. F. Ciurea and B. Funt, "A large image database for color constancy research," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 2003), Vol. **2003**, pp. 160–164.
17. P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2008), pp. 1–8.
18. B. Funt and L. Shi, "The rehabilitation of maxrgb," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 2010), Vol. **2010**, pp. 256–259.
19. J. Vazquez-Corral, C. Párraga, R. Baldrich, and M. Vanrell, "Color constancy algorithms: Psychophysical evaluation on a new dataset," *J. Imaging Sci. Technol.* **53**, 31105 (2009).
20. V. C. Cardei, B. Funt, and K. Barnard, "Estimating the scene illumination chromaticity by using a neural network," *J. Opt. Soc. Am. A* **19**, 2374–2386 (2002).
21. B. Funt and W. Xiong, "Estimating illumination chromaticity via support vector regression," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 2004), Vol. **2004**, pp. 47–52.
22. L. Shi, W. Xiong, and B. Funt, "Illumination estimation via thin-plate spline interpolation," *J. Opt. Soc. Am. A* **28**, 940–948 (2011).
23. X.-S. Zhang, S.-B. Gao, R.-X. Li, X.-Y. Du, C.-Y. Li, and Y.-J. Li, "A retinal mechanism inspired color constancy model," *IEEE Trans. Image Process.* **25**, 1219–1232 (2016).
24. H. Zhao, R. Kawakami, R. T. Tan, and K. Ikeuchi, "Estimating basis functions for spectral sensitivity of digital cameras," in *Meeting on Image Recognition and Understanding* (2009), Vol. **2009**.
25. R. Kawakami, H. Zhao, R. T. Tan, and K. Ikeuchi, "Camera spectral sensitivity and white balance estimation from sky images," *Int. J. Comput. Vision* **105**, 187–204 (2013).

26. M. Serra, O. Penacchio, R. Benavente, M. Vanrell, and D. Samaras, "The photometry of intrinsic images," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2014), pp. 1494–1501.
27. H. Barrow and J. Tenenbaum, "Recovering intrinsic scene characteristics," in *Computer Vision Systems*, A. Hanson and E. Riseman, eds. (1978), pp. 3–26.
28. J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 1670–1687 (2015).
29. I. Marin-Franch and D. H. Foster, "Estimating information from image colors: An application to digital cameras and natural scenes," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 78–91 (2013).
30. G. D. Finlayson, M. S. Drew, and B. V. Funt, "Color constancy: Generalized diagonal transforms suffice," *J. Opt. Soc. Am. A* **11**, 3011–3019 (1994).
31. G. D. Finlayson, M. S. Drew, and B. V. Funt, "Spectral sharpening: Sensor transformations for improved color constancy," *J. Opt. Soc. Am. A* **11**, 1553–1563 (1994).
32. S. Gao and Y. Li, "A retinal mechanism based color constancy model," in *Chinese Conference on Pattern Recognition* (Springer, 2012), Vol. **321**, pp. 422–429.
33. S. Gao, W. Han, K. Yang, C. Li, and Y. Li, "Efficient color constancy with local surface reflectance statistics," in *European Conference on Computer Vision (ECCV)* (Springer, 2014), pp. 158–173.
34. S.-B. Gao, K.-F. Yang, C.-Y. Li, and Y.-J. Li, "Color constancy using double-opponency," *IEEE Trans. Pattern Anal. Mach. Intell.* **37**, 1973–1985 (2015).
35. K.-F. Yang, S.-B. Gao, Y.-J. Li, and Y. Li, "Efficient illuminant estimation for color constancy using grey pixels," in *IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 2254–2263.
36. J. Jiang, D. Liu, J. Gu, and S. Susstrunk, "What is the space of spectral sensitivity functions for digital color cameras?" in *IEEE Workshop on Applications of Computer Vision (WACV)* (IEEE, 2013), pp. 168–179.
37. M. Ebner, "Color constancy based on local space average color," *Machine Vis. Appl.* **20**, 283–301 (2009).
38. G. D. Finlayson and S. D. Hordley, "Color constancy at a pixel," *J. Opt. Soc. Am. A* **18**, 253–264 (2001).
39. G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Inst.* **310**, 1–26 (1980).
40. E. H. Land and J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.* **61**, 1–11 (1971).
41. J. Van De Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. Image Process.* **16**, 2207–2214 (2007).
42. G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 2004), Vol. **2004**, pp. 37–41.
43. D. A. Forsyth, "A novel algorithm for color constancy," *Int. J. Comput. Vision* **5**, 5–35 (1990).
44. A. Gijsenij, T. Gevers, and J. Van De Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *Int. J. Comput. Vision* **86**, 127–139 (2010).
45. G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 1209–1221 (2001).
46. D. H. Brainard and W. T. Freeman, "Bayesian color constancy," *J. Opt. Soc. Am. A* **14**, 1393–1411 (1997).
47. A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy with spatio-spectral statistics," *IEEE Trans. Pattern Anal. Mach. Intell.* **34**, 1509–1519 (2012).
48. C. Rosenberg, A. Ladsariya, and T. Minka, "Bayesian color constancy with non-Gaussian models," in *Advances in Neural Information Processing Systems* (2003), pp. 1–8.
49. V. Agarwal, A. V. Gribok, A. Koschan, and M. A. Abidi, "Estimating illumination chromaticity via kernel regression," in *IEEE International Conference on Image Processing* (IEEE, 2006), pp. 981–984.
50. V. Agarwal, A. Gribok, A. Koschan, B. Abidi, and M. Abidi, "Illumination chromaticity estimation using linear learning methods," *J. Pattern Recognit. Res.* **4**, 92–109 (2009).
51. S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Automatic color constancy algorithm selection and combination," *Patter. Recognit.* **43**, 695–705 (2010).
52. G. Schaefer, S. Hordley, and G. Finlayson, "A combined physical and statistical approach to colour constancy," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2005), Vol. **1**, pp. 148–153.
53. V. C. Cardei and B. Funt, "Committee-based color constancy," in *Color and Imaging Conference* (Society for Imaging Science and Technology, 1999), Vol. **1999**, pp. 311–313.
54. S. Bianco, F. Gasparini, and R. Schettini, "Consensus-based framework for illuminant chromaticity estimation," *J. Electron. Imaging* **17**, 023013 (2008).
55. S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor-outdoor image classification," *IEEE Trans. Image Process.* **17**, 2381–2392 (2008).
56. J. Van De Weijer, C. Schmid, and J. Verbeek, "Using high-level visual information for color constancy," in *IEEE 11th International Conference on Computer Vision (ICCV)* (IEEE, 2007), pp. 1–8.
57. G. Sapiro, "Color and illuminant voting," *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 1210–1215 (1999).
58. R. Lu, A. Gijsenij, T. Gevers, V. Nedović, D. Xu, and J.-M. Geusebroek, "Color constancy using 3D scene geometry," in *IEEE 12th International Conference on Computer Vision (ICCV)* (IEEE, 2009), pp. 1749–1756.
59. J. Vazquez-Corral, M. Vanrell, R. Baldrich, and F. Tous, "Color constancy by category correlation," *IEEE Trans. Image Process.* **21**, 1997–2007 (2012).
60. P. M. Hubel, J. Holm, G. D. Finlayson, and M. S. Drew, "Matrix calculations for digital photography," in *Color Imaging Conference* (1997), pp. 105–111.
61. N. Banic and S. Loncaric, "Color dog: guiding the global illumination estimation to better accuracy," in *International Conference on Computer Vision Theory and Applications—Volume 1: VISAPP* (2015), pp. 129–135.
62. J. Vazquez-Corral and M. Bertalmio, "Spectral sharpening of color sensors: Diagonal color constancy and beyond," *Sensors* **14**, 3965–3985 (2014).
63. G. Wyszecki and W. S. Stiles, *Color Science* (Wiley, 1982), Vol. **8**.
64. I. E. Commission, "Multimedia systems and equipment—colour measurement and management—part 2-1: colour management—default rgb colour space—srgb," Technical Report, IEC 61966-2-1 (1999).
65. K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data," *IEEE Trans. Image Process.* **11**, 972–984 (2002).
66. B. A. Wandell, "The synthesis and analysis of color images," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**, 2–13 (1987).
67. A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2011), pp. 193–200.
68. M. S. Drew and G. D. Finlayson, "Spectral sharpening with positivity," *J. Opt. Soc. Am. A* **17**, 1361–1370 (2000).
69. K. Barnard, L. Martin, A. Coath, and B. Fun, "A comparison of computational color constancy algorithms. II: Experiments with image data," *IEEE Trans. Image Process.* **11**, 985–996 (2002).
70. D. K. Prasad, R. Nguyen, and M. S. Brown, "Quick approximation of camera's spectral response from casual lighting," in *IEEE International Conference on Computer Vision Workshops (ICCVW)* (IEEE, 2013), pp. 844–851.
71. S. D. Hordley and G. D. Finlayson, "Reevaluation of color constancy algorithm performance," *J. Opt. Soc. Am. A* **23**, 1008–1020 (2006).
72. S. 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).
73. Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski, "Non-rigid dense correspondence with applications for image enhancement," *ACM Trans. Graph.* **30**, 1 (2011).
74. J. Vazquez-Corral and M. Bertalmio, "Color stabilization along time and across shots of the same scene, for one or several cameras of unknown specifications," *IEEE Trans. Image Process.* **23**, 4564–4575 (2014).
75. O. Frigo, N. Sabater, J. Delon, and P. Hellier, "Motion driven tonal stabilization," *IEEE Trans. Image Process.* **25**, 5455–5468 (2016).
76. F. Shao, G.-Y. Jiang, M. Yu, and Y.-S. Ho, "Fast color correction for multi-view video by modeling spatio-temporal variation," *J. Visual Comm. Image Represent.* **21**, 392–403 (2010).

77. S.-P. Lu, B. Ceulemans, A. Munteanu, and P. Schelkens, "Spatiotemporally consistent color and structure optimization for multiview video color correction," *IEEE Trans. Multimedia* **17**, 577–590 (2015).
78. A. Ilie and G. Welch, "Ensuring color consistency across multiple cameras," in *Tenth IEEE International Conference on Computer Vision* (IEEE, 2005), pp. 1268–1275.
79. G. Unal, A. Yezzi, S. Soatto, and G. Slabaugh, "A variational approach to problems in calibration of multiple cameras," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 1322–1338 (2007).
80. S. Gong, M. Cristani, S. Yan, and C. C. Loy, *Person Re-identification* (Springer, 2014), Vol. 1.
81. J.-P. Renno, D. Makris, T. Ellis, and G. A. Jones, "Application and evaluation of colour constancy in visual surveillance," in *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance* (IEEE, 2005), pp. 301–308.
82. DxOMark, <http://www.dxomark.com/>.
83. K. Barnard and B. Funt, "Camera characterization for color research," *Color Res. Appl.* **27**, 153–164 (2002).
84. R. Ramanath, W. E. Snyder, Y. Yoo, and M. S. Drew, "Color image processing pipeline," *IEEE Signal Process. Mag.* **22**, 34–43 (2005).