Beyond White: Ground Truth Colors for Color Constancy Correction

Dongliang Cheng[1]   Brian Price[2]   Scott Cohen[2]   Michael S. Brown[1]

[1]National University of Singapore

[2]Adobe Research

{dcheng, brown}@comp.nus.edu.sg

{bprice, scohen}@adobe.com

Abstract

A limitation in color constancy research is the inability to establish ground truth colors for evaluating corrected images. Many existing datasets contain images of scenes with a color chart included; however, only the chart's neutral colors (grayscale patches) are used to provide the ground truth for illumination estimation and correction. This is because the corrected neutral colors are known to lie along the achromatic line in the camera's color space (i.e. R=G=B) ; the correct RGB values of the other color patches are not known. As a result, most methods estimate a 3×3 diagonal matrix that ensures only the neutral colors are correct. In this paper, we describe how to overcome this limitation. Specifically, we show that under certain illuminations, a diagonal 3×3 matrix is capable of correcting not only neutral colors, but all the colors in a scene. This finding allows us to find the ground truth RGB values for the color chart in the camera's color space. We show how to use this information to correct all the images in existing datasets to have correct colors. Working from these new color corrected datasets, we describe how to modify existing color constancy algorithms to perform better image correction.

1. Introduction

The goal of computational color constancy is to mimic the human visual system's ability to perceive scene objects as the same color when they are viewed under different illuminations. Cameras do not intrinsically have this ability and color changes due to scene illumination must be corrected. This is a fundamental pre-processing step applied to virtually every image.

Color constancy is typically a two-step procedure: 1) estimate the color of the illumination; 2) apply a transform to remove the effects of the illumination. The majority of published literature addresses step 1. Several datasets have been created to assist in evaluating illumination estimation (e.g. [1, 9, 12, 24, 32]). The basic idea is to place a neu-
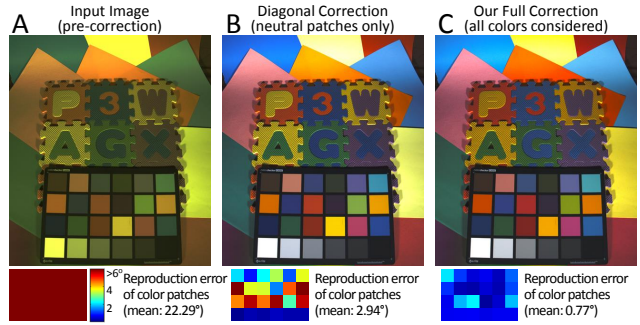

Figure 1. (A) input image before illumination correction. (B) corrected image using a conventional diagonal 3×3 matrix (i.e. white-balancing). (C) corrected image using a full 3×3 matrix estimated from the ground truth colors obtained by our approach. The reproduction angular errors for each 24 color patches are shown below each image as a heat map (red=high error, blue=low error).

tral (white) calibration object in the imaged scene. Under ideal white light, the neutral object should remain achromatic in the camera's color space. A chromatic color cast on the neutral object is considered to be the color of the illumination in the camera's color space. While most methods do not elaborate on image correction, the de facto approach is to compute a 3×3 diagonal matrix to map the estimated illumination RGB values to lie along $R=G=B$. This is effectively known as white-balancing and ensures the neutral colors appear "white" in the corrected image. However, the ability of this diagonal matrix to correct non-neutral colors is ignored (Fig. 1).

This is a significant limitation, because the goal of color constancy is to make *all* colors correct, not just neutral colors. Early color constancy datasets are suitable only for illumination estimation as they only contain a neutral calibration pattern. Newer datasets, such as the widely used Gelher-Shi [24, 32] and the recent NUS dataset [9] include a color rendition chart in every image. However, only the neutral patches on these color charts are used for performance evaluation. The problem is that unlike a neutral material, the ground truth RGB values of the color patches are not known in the camera's color space. While color ren-

dition charts have known mapping values in the CIE XYZ color space, color constancy correction is performed in the camera's color space [8, 29]. Currently, the only way to estimate these colors is with spectral information, including the camera sensor sensitivity functions, spectral reflectances of the patches, and spectra of the illumination. Such spectral data is challenging to obtain, and as a result, most existing color constancy datasets cannot be used to evaluate the performance of color correction.

Contributions This paper makes four contributions towards better image correction for color constancy.

1. We show that a diagonal matrix is able to correct scene colors for certain illuminations (including daylight) well enough to define the ground truth colors for the other illuminations.
2. Based on the findings in 1, we describe a robust method to select the images in the existing color constancy datasets to provide the ground truth colors for the imaged rendition chart. This allows us to re-purpose datasets used for illumination estimation, to also be used for color correction by estimating a full 3×3 color correction matrices for all the images in the dataset.
3. Using the re-purposed datasets from 2, we demonstrate how these full matrices can be immediately used to modify existing color constancy algorithms to produce better color correction results.
4. Finally, we found that existing datasets have a strong bias of images captured in daylight scenes. To create a more uniformly sampled dataset for studying color constancy correction, we have captured an additional 944 images under indoor illuminations to expand the NUS multi-camera dataset.

We believe this work will have significant implications for improving color constancy by allowing the evaluation of color correction algorithms beyond white correction.

2. Related Work

There is a large body of work targeting color constancy, with the vast majority focused on illumination estimation. Representative examples include statistical methods that directly estimate the illumination from an input image's RGB values (e.g. [5, 6, 18, 26, 34, 35]) and learning-based methods that use various features extracted from datasets with ground truth illumination to learn an estimator (e.g. [7, 10, 14, 17, 20, 24, 31, 33]). A full discussion of these methods is outside the scope of this paper, however, more details can be found in the comprehensive survey by Gijsenij et al. [25].

There is significantly less work focused on correcting images. It is generally assumed that the three RGB channels from the camera sensor act as independent gain controls to scene illumination. This is similar to the von Kries hypothesis [36] on human retinal cones. Working from the von Kries assumption, a diagonal 3×3 matrix can be used to correct the three RGB channels by normalizing their individual channel bias. This has long been known to be incorrect [13], but remains the de facto method for image correction.

Early work by Finlayson et al. [15, 16] proposed a method to address this problem with what was termed the *generalized diagonal model*. In their work, a 3×3 spectral sharpening matrix transform, $\mathbf{M}$, was computed to map the sensor's RGB values to an intermediate color space, for which the diagonal correction model works well. Finlayson et al. [16] showed that a two-dimensional linear space of illuminants and a three-dimensional linear space of reflectances (or vice versa) were sufficient to guarantee the generalized diagonal model. Estimating $\mathbf{M}$, however, requires accurate camera responses of known materials under controlled illumination. To achieve this, the camera responses are simulated from spectral data of illumination and reflectance using camera sensitivity functions. Chong et al. [11] later revealed that the generalized diagonal compatibility conditions are impositions only on the sensor measurements, not the physical spectra. They formulated the problem as a rank constraint on an order three measurement tensor to compute the matrix $\mathbf{M}$. Once again, Chong et al. [11] require that the spectral sensitivity of the camera's sensor to be known. The use of this spectral sharpening matrix $\mathbf{M}$ effectively meant the color correction transform was a full 3×3 matrix.

Work in [23, 27] examined the dimensionality of the 9-parameter space of the full 3×3 color correction matrices. Using PCA decomposition, they found that only 3 bases were required to recover the 9 parameters in the full matrix model. The full matrices used in their PCA decomposition were synthetically generated using a known camera sensitivity function and a large database of material spectral reflectances and illumination spectra.

While these methods helped to lay the foundation on how to estimate full 3×3 color correction matrices, the reliance on spectral information makes them impractical. Bianco and Schettini [3] proposed a method to estimate the sharpening matrix without spectral data in an optimization framework that simultaneously estimated the color mapping matrix to a device independent color space. The accuracy of this approach with respect to the camera sensor's color space, however, is unclear.

In the following section, we describe how to estimate the ground truth colors in the camera sensor space directly from camera images.

3. Diagonal Model For Ground Truth Colors

This section performs an analysis which reveals that for certain illuminations, the 3×3 diagonal correction model is
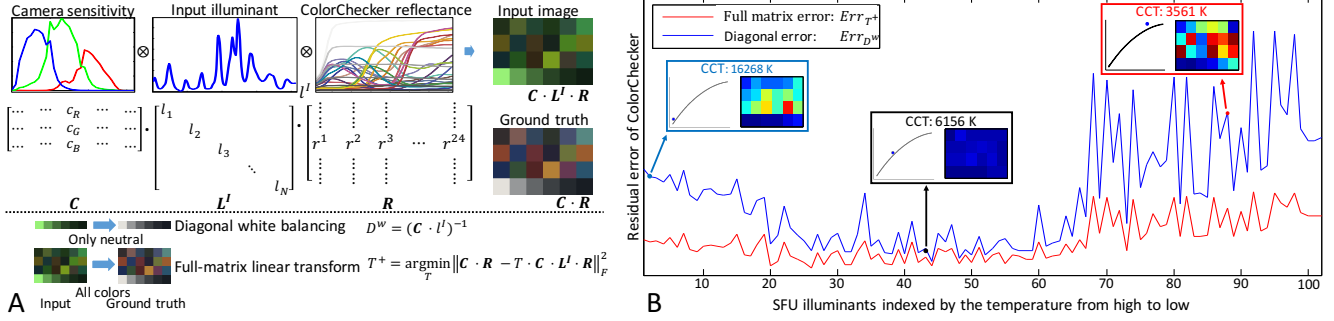
Figure 2. (A) Illustration of the difference between the diagonal white-balancing correction and the full matrix image correction transform. White-balancing only requires the observations of the neutral colors. To estimate the full matrix, the observed color chart and its ideal colors are needed. (B) Shows the residual error comparison of the two different correction models. While the full matrix has consistently lower error, for certain illuminations the error from the diagonal model is close to that from the full matrix. A heatmap visualization of the diagonal matrix errors for each color patch is shown for three illuminates. The chromaticity position of the illuminations with respect to the Plankian color temperature curve and their corresponding correlated color temperature (CCT) are also shown.

useful for full color correction of the scene, and not just neutral colors. This analysis is performed empirically in Sec. 3.1 working from spectral data. Sec. 3.2 shows our mathematical model of the color constancy problem that lends corroborative evidence to our empirical observation.

3.1. Empirical Analysis

Here we show empirically that 3×3 diagonal correction matrices are sufficient to correct the scene's colors for certain illuminations as well as full matrix correction can. Our analysis starts by examining how RGB camera values are formed in the spectral domain. Let $\mathbf{C}$ represent the camera's sensitivity functions that is written as a $3 \times N$ matrix, where N is the number of spectral samples and the rows of $\mathbf{C} = [\mathbf{c_R}; \mathbf{c_G}; \mathbf{c_B}]$ correspond to the R, G, B channels. The camera response for a particular scene material, $\mathbf{r}$ under illumination $\mathbf{l}$ can be obtained by the Lambertian model where the specular reflection is ignored:

$$\boldsymbol{\rho} = \mathbf{C} \cdot diag(\mathbf{l}) \cdot \mathbf{r} = \mathbf{C} \cdot \mathbf{L} \cdot \mathbf{r}, \quad (1)$$

where $\mathbf{l}$ and $\mathbf{r}$ are $N \times 1$ vectors representing the illumination spectra and material spectral reflectance respectively, and $diag(\cdot)$ indicates the operator that creates a diagonal matrix from a vector, i.e. $\mathbf{L}$ is an $N \times N$ illumination matrix with diagonal elements $\mathbf{l}$.

The goal of color constancy is to map an RGB value taken under an unknown illumination, $\boldsymbol{\rho}^I = \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{r}$, to its corresponding color under a canonical illumination, $\boldsymbol{\rho}^C = \mathbf{C} \cdot \mathbf{L^C} \cdot \mathbf{r}$. Although the canonical illumination can be any specific spectra, ideal white light that has equal energy for every wavelength (i.e., the CIE standard illuminant E) is generally chosen. In such a case, $\mathbf{L^C}$ becomes the identity matrix, $\mathbf{I}$, and gives us $\boldsymbol{\rho}^C = \mathbf{C} \cdot \mathbf{r}$. This mapping can be written as:

$$\begin{aligned} \boldsymbol{\rho}^C &= T \cdot \boldsymbol{\rho}^I, \\ \mathbf{C} \cdot \mathbf{r} &= T \cdot \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{r}, \end{aligned} \quad (2)$$
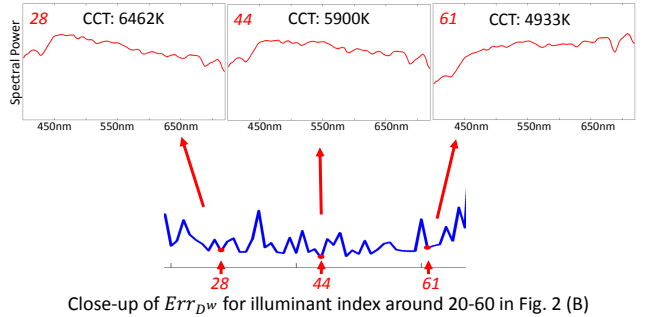
Close-up of Err_{D^w} for illuminant index around 20-60 in Fig. 2 (B)

Figure 3. Spectra (400-720nm) for illuminations on which diagonal white-balancing correction works well. The bottom blue curve corresponds to the blue curve of the diagonal correction error in Fig. 2 (B) for illuminations index around 20-60. The correlated color temperate (CCT) is also shown. These spectra are indicative of broadband sunlight/daylight illumination.

where T is a 3×3 linear matrix that maps $\boldsymbol{\rho}^I$ to $\boldsymbol{\rho}^C$. In general, we have a scene composed of many different materials, and not just one. In this case, if we assume that the scene is illuminated by a single illumination, we have:

$$\mathbf{C} \cdot \mathbf{R} = T \cdot \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{R}, \quad (3)$$

where $\mathbf{R}$ is a matrix of many material reflectances (see Fig. 2 (A)). Due to the metameric nature of Eq.3 an exact solution for T is not possible [21, 30, 37]. We therefore seek a transform T^+ that minimizes the Frobenius norm:

$$T^+ = \operatorname*{argmin}_{T} ||\mathbf{C} \cdot \mathbf{R} - T \cdot \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{R}||_F^2, \quad (4)$$

where $|| \cdot ||_F^2$ indicates the matrix Frobenius norm. A solution to this optimization problem can be obtained using the Moore-Penrose pseudoinverse. Note that, to solve this problem, we need observations of the ideal (ground truth) colors, $\mathbf{C} \cdot \mathbf{R}$, and the input image under the scene illumination, $\mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{R}$.

Given $T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix}$,

off-diagonal-to-diagonal ratio κ of T is:

$$\kappa = \frac{|t_{12}| + |t_{21}| + |t_{13}| + |t_{31}| + |t_{23}| + |t_{32}|}{|t_{11}| + |t_{22}| + |t_{33}|}$$
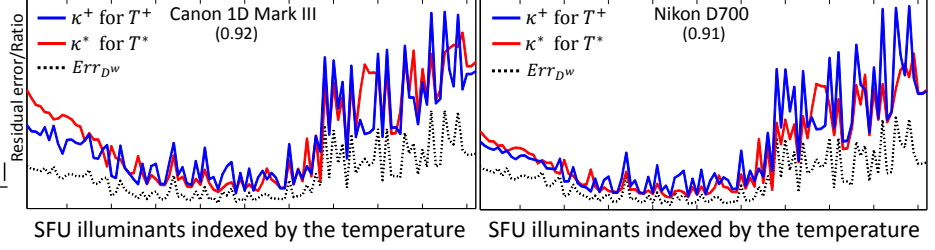
Figure 4. The trend of off-diagonal-to-diagonal ratio of T^* and T^+ for all the illuminations and their correlation. Plots from two specifc cameras are shown here, but all the other cameras share this similar trend: for certain illuminations, the off-diagonal-to-diagonal ratio is low and high correlation can be found from the ratios of two different matrices (correlation coefficients are shown under the camera name).

Let's now consider computing a diagonal, 3×3 correction matrix, D^w, as done by most white-balancing methods. We assume our camera has observed a special neutral **r** that reflects spectral energy at every wavelength equally. This means our camera response is the direct response of the illumination $\mathbf{l}^I$, thus giving us:

$$D^w = diag(\mathbf{C} \cdot \mathbf{l}^I)^{-1}, \qquad (5)$$

where $\mathbf{l}^I$ is the input illumination (i.e., $\mathbf{L}^I = diag(\mathbf{l}^I)$). This only requires the observation of the neutral patches. Fig. 2 (A) illustrates the difference between these methods.

The residual errors for the two solutions over all observed scene materials R can be expressed as the Frobenius norms:

$$\begin{aligned} Err_{T^+} &= ||\mathbf{C} \cdot \mathbf{R} - T^+ \cdot \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{R}||_F^2 \\ Err_{D^w} &= ||\mathbf{C} \cdot \mathbf{R} - D^w \cdot \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{R}||_F^2. \end{aligned} \qquad (6)$$

The question we are interested in is: When does D_w provides a good approximation to T^+? To determine this, we compute the residual errors in Eq. 6 for 28 different cameras using the camera sensitivity functions from [28]. We examined these errors for 101 different real world illuminations captured by [1]. The reflectance materials used were those estimated from the 24 color patches on the Macbeth ColorChecker.

Fig. 2 (B) shows a plot of the residual errors for both T^+ and D^w from two specific cameras (different $\mathbf{C}$ in Equation 6). The horizontal axis is the index of the 101 illuminants. We sort the illuminations by their correlated color temperature in the CIE-xy chromaticity space. We can see that for many illuminations, the errors of these two methods are similar. In particular, for illuminations close to range 6000K, the diagonal D_w is very close to the full matrix T^+. Fig. 3 shows several of the illumination spectra in this range. We note that these spectra resemble those caused by sunlight, including direct daylight and shadows. For other illuminations, especially indoor artificial ones, the correction error from D_w is much larger than that from T^+.

Another useful interpretation of this observation is to examine under what illuminations T^+ becomes more like a diagonal matrix. For this, we can define the off-diagonal-to-diagonal ratio κ of matrix T^+ as:

$$\kappa = \frac{\sum_{i=1}^{3} \sum_{j=1, j \neq i}^{3} |t_{i,j}|}{\sum_{i=1}^{3} |t_{i,i}|}, \qquad (7)$$

where $t_{i,j}$ is the (i,j) element of matrix T and $|\cdot|$ indicates the absolute value. On careful inspection of Eq. 7, we see that κ decreases in value as the diagonal entries in the T matrix become more dominant than the off-diagonal entries of T. When $\kappa = 0$ the matrix T is a diagonal matrix. Fig. 4 plots κ^+ for T^+ against the 101 illuminations for two different cameras, Canon 1D Mark III and Nikon D700. The trend of κ^+ closely follows the observation of the residual errors from diagonal white-balancing correction, Err_{D^w}.

3.2. Mathematical Support for Our Observation

To have further support for this finding, we performed another analysis that does not rely on the scene reflectance **R**. This can be considered as estimating a full matrix that is optimal over all possible reflectance values. In this case, we drop **R** from Eq. 3 to obtain:

$$\mathbf{C} = T \cdot \mathbf{C} \cdot \mathbf{L}^I. \qquad (8)$$

Similar to Eq. 4, the optimal linear transform T^* is the one that minimizes the Frobenius norm of the difference:

$$T^* = \underset{T}{\operatorname{argmin}} ||\mathbf{C} - T \cdot \mathbf{C} \cdot \mathbf{L}^I||_F^2, \qquad (9)$$

and it can also be computed directly from the Moore-Penrose pseudoinverse:

$$T^* = \mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{C}^t \cdot (\mathbf{C} \cdot \mathbf{L}^I \cdot \mathbf{L}^I \cdot \mathbf{C}^t)^{-1}. \qquad (10)$$

Using this T^* that does not rely on any reflectance materials, we plot its corresponding κ^* against the plot for the κ^+ in Fig. 4. We can see that the two plots are highly correlated, providing corroborative evidence to our empirical observation. The overall relationship of T^* to the illumination, L, and camera sensitivities, C, is complex given the number of parameters involved. For the purpose of establishing ground truth colors in existing datasets, we will rely on the use of images captured in daylight illumination as indicated by the experiments in this Section.
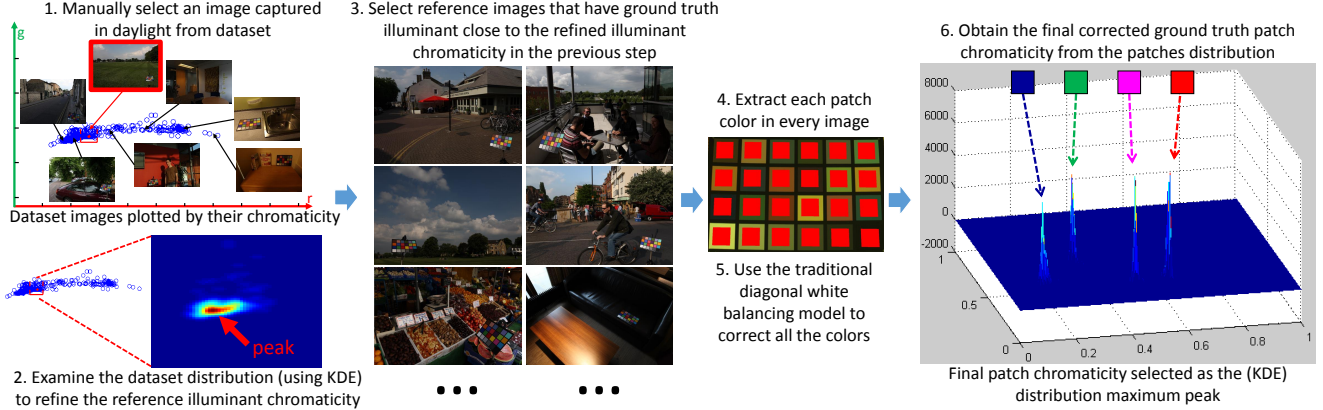
Figure 5. Procedure to calculate the "ground truth" RGB colors for the color chart patches. First, an outdoor image captured under sunlight is manually selected. A kernel density estimation (KDE) method is applied on nearby ground truth illuminations to refine the illumination chromaticity as the peak location of the local illumination chromaticity distribution. Images with illuminations close to this refined reference illumination are selected automatically. Each image in this reference image set is corrected using the diagonal model and each color patch is extracted. KDE is applied to each color patch's corrected colors over the entire set and the KDE peak is selected as the ground truth color.

4. Re-purposing Datasets

Existing color constancy datasets with full color rendition charts in the scene are currently only used for the purpose of illuminant estimation evaluation with the achromatic patches. This is because the ground truth colors of the color patches in the camera's color space are not known. The findings in Sec. 3, however, tell use that under certain illumination the standard diagonal correction matrix is able to correct the scene colors, thus providing a very good approximation of the ground truth colors of the color chart. In this section, we describe how to use the color chart RGB values to re-purpose existing datasets, namely the Gelher-Shi and the NUS datasets, so that they can also be used for the purpose of color correction estimation. We also discuss an appropriate error metric for evaluating color correction as well as our need to augment datasets to have a better balance of indoor and outdoor images.

4.1. Robust Estimation of Patch Colors

The Gelher-Shi and NUS datasets have color rendition charts in every scene. This means there are 24 common materials present in all the images. Here, we describe how to compute the ground truth values of these 24 color patches in the camera sensor's color space.

While we could use a single image captured under daylight to provide the reference colors of the rendition chart, this naive approach risks selecting an image that may possibly be corrupt by factors such as nonuniform illumination and camera noise. Instead, we have devised a robust procedure for selecting the colors. An overview of this procedure is provided in Fig.5. We start with the entire dataset of the images captured from the same camera under different illuminations. The ground truth illuminations for these images

are available from the chart's neutral patches. We manually select an image that is clearly captured in daytime. We then look for a set of images that have similar ground truth illuminations. This is done by performing a 2D kernel density estimation (KDE) [4] on the chromaticity distribution of the ground truth illuminations. We find the peak of the KDE closest to our manually selected image. We then take dataset images whose ground truth illumination chromaticity distance to this KDE peak are smaller than a threshold to form our reference image set. For each image in this reference image set, we correct the image using the diagonal correction matrix based on its ground truth illumination. Note from Fig. 5 that this reference image set may contain a few images which are not outdoor sunlight images. To prevent our ground truth colors from being contaminated by these outliers, we again apply KDE on the corrected chromaticity for each patch and select the peak of the distribution as the ground truth color for each patch. This procedure provides a robust mechanism for finding the ground truth colors for all the patches. When we applied this on the Gehler-Shi dataset (Canon 5D subset), any manually-chosen reference image that was captured in direct sunlight resulted in nearly identical ground truth estimations.

After obtaining the ground truth checker chart colors, we can now compute full matrices to transform all the images in the dataset based on the color checker colors. This can be done using the Moore-Penrose pseudo-inverse similar to Eq. 4. However, as noted by Funt et al. [22], the illumination across the color rendition chart is generally not uniform. As a result, we follow the approach in [22] to minimize the sum of angular error:

$$T = \operatorname*{argmin}_{T} \sum_{i=1}^{24} cos^{-1} \left(\frac{T\boldsymbol{\rho}_i^I \cdot \boldsymbol{\rho}_i^C}{\|T\boldsymbol{\rho}_i^I\| \, \|\boldsymbol{\rho}_i^C\|} \right), \qquad (11)$$
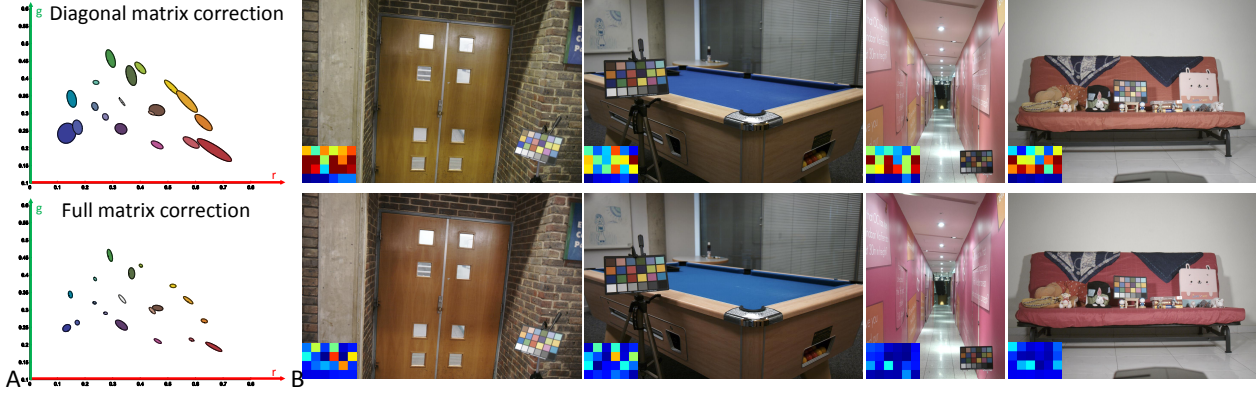
Figure 6. This figure shows the ability of the full matrix to produce better image correction. (A) shows the distribution (modeled by Gaussians) of each color patch in the color checker chart in the entire Gelher-Shi Canon1D dataset after correction using the proposed full matrix and the diagonal matrix. The full matrix correction clearly decreases the variance in the color distributions after correction. (B) shows images (from both Gelher-Shi and NUS datasets) corrected using a diagonal matrix (top) and a full matrix (bottom). The color coded reproduction angular errors for each 24 color patches are also shown (red=high error, blue=low error).

where ρ_i^I is the patch color in this input camera image for patch i and ρ_i^C is the estimated ground truth color for patch i. Fig. 6 (A) shows the ability of the T estimated for each image to provide a better mapping than the traditional diagonal correction. The two plots in Fig. 6 (A) show the distribution of corrected colors of the color patch using the full matrix T and the diagonal matrix. The colors are much more coherent across the entire Gelher-Shi dataset. Fig. 6 (B) shows comparisons of four images selected from the datasets. This is accompanied with a per patch error map which is shown through this paper. The metric used to measure error is described next.

4.2. Correction Error Metric

For illumination estimation, the most common error metric is known as the recovery error, and is computed as the angular error between the estimated illumination and the ground truth illumination in the camera's color space. This is shown in Fig. 7 (A). Note that this can be estimated without correcting the image.

As we are interested in correcting the image, the angular error is computed after correction. This can be defined as:

$$Err_i = cos^{-1}\left(\frac{\rho_i^T \cdot \rho_i^C}{\left\|\rho_i^T\right\|\left\|\rho_i^C\right\|}\right) \quad i = 1..24, \quad (12)$$

where Err_i is the angular error for patch i and ρ_i^T is the color of each patch after correction. Fig. 7 (B)-(C) demonstrates the difference for the neutral color and color patch colors respectively. Interestingly, this approach (termed the reproduction error) was recently advocated by Finlayson and Zakizadeh [19] for illumination estimation as an improved metric. We adopt it here for estimating all the patch colors in the rendition chart.
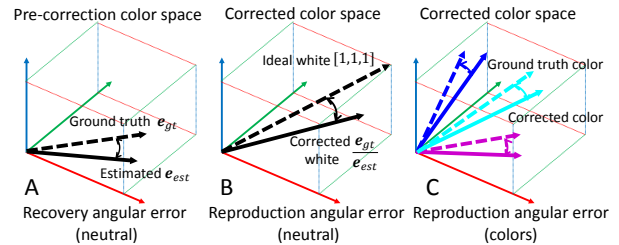

Figure 7. Illustration of recovery angular error (A) and reproduction angular error for neutral (B) and reproduction angular error for non-neutral color (C). Dotted lines represent ground truth colors; solid lines represent estimated or corrected colors.

4.3. Expanding the NUS Dataset

Our analysis of existing illumination correction datasets found that they have significantly more outdoor images where the diagonal matrix works well, than illuminations such as indoor lighting, for which the full matrix is needed for correction. To address this, we have captured additional 944 images to expand the NUS dataset [9]. We use the NUS dataset because it is the newest dataset and has significantly more cameras (e.g. the Gelher-Shi dataset only has two cameras). Using the same cameras used in [9], we captured 18 scenes under 6 different indoor illuminations using each camera. Fig. 8 shows some of the example images. These additional images make the distribution of different illuminations much more uniform.

5. Application to Color Constancy

Here we describe how the re-purposed datasets described in Sec. 3 can be immediately used to improve existing methods. In particular, we show how to modify two specific learning-based methods, the Bayesian method [24, 31] and the Corrected-Moments method [14] to use the full color matrix. To give an insight into the potential of our newly
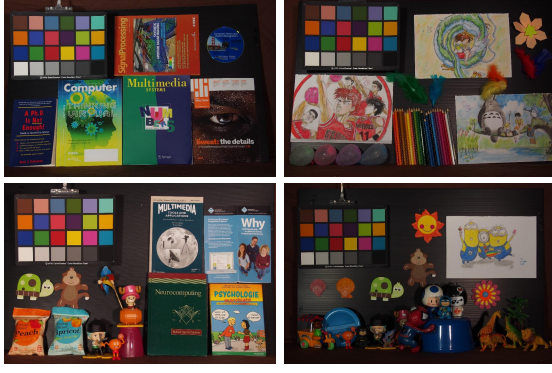
Figure 8. Examples of our newly captured indoor images. Similar to the SFU laboratory image set, for each scene, we capture images under multiple lighting conditions with a Macbeth Color rendition chart in the scene.

computed datasets, we have also implemented an *oracle prediction* method that is used to test our idea beyond the limit of current illumination estimation performance.

Bayesian method The work by Gehler et al. [24] revisited the original Bayesian color constancy method from [31]. The approach begins by correcting all the images in the training set with diagonal white-balancing matrices based on the ground truth illumination color. This is used to build a likelihood probability distribution of the corrected/reflectance colors. Then the prior information of diagonal correction matrices is used to help predict the most possible illumination in the scene within a Bayesian inference framework. We modified this approach by changing the image correction model, as well as the prior information, to be the full matrix correction model. This will effectively output a full matrix transform T by searching for the MAP (maximum a posteriori) of the posterior probability for T:

$$p(T|\boldsymbol{\rho}^I) \propto p(\boldsymbol{\rho}^I|T)p(T). \qquad (13)$$

Corrected-Moments We can also extend a recent method proposed by Finlayson [14] that does not assume any explicit image correction model. This method only requires the original (pre-corrected) input image color/edge moments, denoted by $\mathbf{p}_m$ comprising of m moments. In the training stage, a regression matrix $\mathbf{C}_{m \times 3}$ is learned to map the moments to the final illumination estimation:

$$\mathbf{e}_{est} = \mathbf{p}_m \mathbf{C}_{m \times 3}. \qquad (14)$$

We followed this procedure to estimate the illumination, but replaced the image correction step to use the 3×3 full matrix associated with the image in training-set whose ground truth illumination is closest to $\mathbf{e}_{est}$.

Oracle prediction The use of the Baysesian and Corrected-Moments are intended to show how the new full color datasets can be immediately used to improve color correction based on the existing illumination estimation methods.

We expect, however, continuous improvements in illumination estimation and hope that our datasets will be useful in this effort. We show results using what we term the "oracle method" that assumes an ideal illumination estimation method that can select the image in the training set with the closest illumination in the ground truth dataset to an input test image. We use this oracle method to help reveal the full potential of better color image correction.

Table 1 lists all the results for these three comparison settings using the reproduction error described in Section 4.2. To maximize the performance of the learning-based methods, the results were obtained using a leave-one-out cross validation as performed in [2]. Results are reported on outdoor, indoor, and all the images. For outdoor images, our results are comparable to the existing methods. This is not surprising as Sec. 3 indicates that the current diagonal correction method works well for outdoor images. In addition, since our method attempts to minimize the error across all the color patches and not just neutral, our results on the neutral only patches are not always as good as the diagonal method. However, for indoor illuminations we see significant gains. These gains are more noticeable in the augmented NUS dataset that has a better balance between indoor and outdoor images. Moreover, for the oracle prediction, the full matrix correction wins every camera in the "Color" and "All" categories, which indicates the possible color constancy improvements with better illumination estimation methods in the future. Fig. 9 shows a few examples of subjective comparisons from the Bayesian method.

6. Discussion and Summary

This paper describes how to obtain ground truth colors in camera sensor color space for use in color constancy image correction. To the best of our knowledge, this is the first work to show how to estimate these colors directly from camera images without the need for careful spectral calibration of the camera and imaged materials. Our findings have allowed us to re-purpose existing illumination estimation datasets to be used for evaluating image correction. Our results in Sec. 5 represent that for the first time full matrices can be estimated and evaluated for these datasets. These re-purposed datasets, along with the new indoor images described in Sec. 4 will be made publicly available. Our modifications to existing algorithms have just scratched the surface of the usefulness of these new datasets and we believe this work will have significant implications to researchers in this area who can finally move beyond white-balancing and towards true color constancy.

Acknowledgement

	Outdoor images						Indoor images						All images					
	Neutral		Color		All		Neutral		Color		All		Neutral		Color		All	
	D	T	D	T	D	T	D	T	D	T	D	T	D	T	D	T	D	T
	Bayesian																	
Gehler-Shi Canon 1D (15/71)	**5.53**	6.25	**3.76**	4.32	**4.20**	4.80	**6.16**	6.55	5.91	**5.85**	**5.97**	6.03	**6.05**	6.50	**5.54**	5.58	**5.66**	5.81
Gehler-Shi Canon 5D (307/175)	3.31	**3.20**	2.96	**2.75**	3.04	**2.86**	7.56	**7.00**	6.45	**6.15**	6.73	**6.36**	4.85	**4.58**	4.22	**3.98**	4.38	**4.13**
NUS Canon 1Ds Mark III (197/167)	**4.27**	4.42	**3.46**	3.60	**3.66**	3.80	5.41	**5.30**	5.16	**4.57**	5.22	**4.75**	**4.79**	4.83	4.24	**4.04**	4.38	**4.24**
NUS Canon 600D (145/160)	5.05	**4.79**	4.36	**3.98**	4.53	**4.18**	**4.39**	4.69	4.63	**4.16**	4.57	**4.29**	**4.70**	4.74	4.50	**4.08**	4.55	**4.24**
NUS Fujifilm XM1 (144/157)	**3.52**	3.77	**2.72**	2.91	**2.92**	3.12	**4.05**	4.51	4.34	**3.75**	4.27	**3.94**	**3.80**	4.15	3.57	**3.35**	3.63	**3.55**
NUS Nikon D40 (80/141)	5.16	**4.54**	3.93	**3.56**	4.24	**3.80**	5.97	**5.29**	5.85	**4.59**	5.88	**4.77**	5.68	**5.02**	5.15	**4.22**	5.28	**4.42**
NUS Nikon D5200 (151/154)	**5.07**	5.29	**3.89**	3.99	**4.18**	4.31	**5.06**	5.43	4.89	**4.33**	4.93	**4.61**	**5.06**	5.36	4.39	**4.16**	4.56	**4.46**
NUS Olympus EPL-6 (153/160)	**3.82**	3.91	**3.18**	3.26	**3.34**	3.43	5.16	**4.95**	4.94	**4.78**	4.99	**4.82**	4.50	**4.44**	4.08	**4.04**	4.18	**4.14**
NUS Lumix DMC-GX1 (147/161)	**5.19**	5.23	3.97	**3.86**	4.28	**4.21**	**4.91**	5.49	5.17	**4.65**	5.10	**4.86**	**5.04**	5.36	4.60	**4.28**	4.71	**4.55**
NUS Samsung NX2000 (153/154)	**4.89**	5.15	**3.79**	4.00	**4.06**	4.29	**5.28**	5.44	5.21	**4.61**	5.22	**4.81**	**5.09**	5.29	4.50	**4.31**	4.65	**4.55**
NUS Sony STL-A57 (207/166)	4.22	**4.16**	3.71	**3.62**	3.84	**3.76**	5.60	**5.04**	5.09	**4.21**	5.22	**4.42**	4.83	**4.55**	4.32	**3.89**	4.45	**4.05**
	Corrected-moment																	
Gehler-Shi Canon 1D (15/71)	**3.11**	3.12	**2.05**	2.23	**2.32**	2.46	**3.20**	3.36	4.12	**3.60**	3.89	**3.54**	**3.19**	3.32	3.76	**3.36**	3.62	**3.35**
Gehler-Shi Canon 5D (307/175)	**2.37**	2.39	**2.11**	2.17	**2.17**	2.22	5.15	**5.12**	4.77	**4.34**	4.87	**4.53**	**3.38**	3.38	3.08	**2.95**	3.15	**3.06**
NUS Canon 1Ds Mark III (197/167)	**2.69**	2.70	**2.28**	2.41	**2.38**	2.48	**3.49**	3.59	3.53	**3.09**	3.52	**3.21**	**3.06**	3.11	2.85	**2.72**	2.90	**2.82**
NUS Canon 600D (145/160)	**2.29**	2.29	2.14	**2.03**	2.18	**2.10**	**2.79**	3.12	3.05	**2.72**	2.99	**2.82**	**2.55**	2.72	2.62	**2.39**	2.60	**2.48**
NUS Fujifilm XM1 (144/157)	2.59	**2.55**	**2.08**	2.10	**2.21**	2.21	**3.54**	3.68	3.87	**3.08**	3.79	**3.23**	**3.09**	3.14	3.01	**2.61**	3.03	**2.74**
NUS Nikon D40 (80/141)	**3.18**	3.21	**2.19**	2.37	**2.44**	2.58	3.49	**3.30**	3.77	**3.01**	3.70	**3.08**	3.38	**3.26**	3.20	**2.78**	3.24	**2.90**
NUS Nikon D5200 (151/154)	**2.77**	2.84	2.26	**2.24**	2.39	**2.39**	**3.29**	3.34	3.62	**3.10**	3.54	**3.16**	**3.04**	3.09	2.95	**2.67**	2.97	**2.78**
NUS Olympus EPL-6 (153/160)	**2.21**	2.30	1.98	**1.93**	2.04	**2.02**	**3.33**	3.44	3.66	**3.12**	3.58	**3.20**	**2.79**	2.88	2.84	**2.54**	2.83	**2.62**
NUS Lumix DMC-GX1 (147/161)	**2.50**	2.54	1.89	**1.87**	2.04	**2.04**	**2.95**	3.11	3.68	**2.87**	3.50	**2.93**	**2.73**	2.84	2.83	**2.39**	2.80	**2.50**
NUS Samsung NX2000 (153/154)	**2.66**	2.75	2.26	**2.13**	2.36	**2.28**	3.18	**3.15**	3.67	**2.88**	3.55	**2.94**	**2.92**	2.95	2.97	**2.50**	2.95	**2.61**
NUS Sony STL-A57 (207/166)	2.72	**2.50**	2.65	**2.37**	2.67	**2.40**	3.69	**3.59**	3.46	**3.02**	3.52	**3.16**	3.15	**2.98**	3.01	**2.66**	3.05	**2.74**
	Oracle prediction																	
Gehler-Shi Canon 1D (15/71)	2.16	**1.74**	1.41	**1.03**	1.58	**1.18**	**1.45**	1.68	3.23	**1.88**	2.84	**1.84**	**1.57**	1.69	2.91	**1.73**	2.62	**1.72**
Gehler-Shi Canon 5D (307/175)	**0.42**	0.49	0.98	**0.60**	0.86	**0.57**	**1.04**	1.19	2.72	**1.65**	2.35	**1.55**	**0.64**	0.74	1.61	**0.98**	1.40	**0.93**
NUS Canon 1Ds Mark III (197/167)	**0.54**	0.55	0.90	**0.59**	0.82	**0.58**	0.90	**0.89**	2.48	**1.12**	2.14	**1.07**	0.71	**0.70**	1.62	**0.83**	1.43	**0.80**
NUS Canon 600D (145/160)	**0.57**	0.59	1.24	**0.64**	1.09	**0.63**	**0.66**	0.76	2.18	**1.06**	1.85	**0.99**	**0.61**	0.68	1.73	**0.86**	1.49	**0.82**
NUS Fujifilm XM1 (144/157)	0.72	**0.67**	1.02	**0.61**	0.95	**0.62**	**0.82**	0.97	2.64	**1.10**	2.24	**1.07**	**0.77**	0.83	1.86	**0.87**	1.62	**0.86**
NUS Nikon D40 (80/141)	1.61	**1.58**	1.21	**0.96**	1.29	**1.09**	1.06	**1.05**	2.67	**1.16**	2.32	**1.13**	1.26	**1.24**	2.14	**1.09**	1.95	**1.12**
NUS Nikon D5200 (151/154)	0.67	**0.70**	0.96	**0.70**	0.89	**0.70**	**0.87**	0.98	2.43	**1.13**	2.09	**1.09**	**0.77**	0.84	1.70	**0.92**	1.50	**0.90**
NUS Olympus EPL-6 (153/160)	**0.49**	0.50	0.82	**0.54**	0.75	**0.53**	**1.13**	1.28	2.67	**1.28**	2.34	**1.28**	**0.82**	0.90	1.77	**0.92**	1.56	**0.91**
NUS Lumix DMC-GX1 (147/161)	1.32	**1.10**	1.14	**0.70**	1.18	**0.79**	1.17	**1.05**	2.95	**1.09**	2.57	**1.08**	1.24	**1.07**	2.09	**0.91**	1.90	**0.94**
NUS Samsung NX2000 (153/154)	1.16	**0.95**	1.14	**0.73**	1.14	**0.78**	1.11	**1.08**	2.61	**1.25**	2.28	**1.21**	1.13	**1.01**	1.88	**0.99**	1.72	**1.00**
NUS Sony STL-A57 (207/166)	**0.67**	0.75	1.54	**0.76**	1.35	**0.76**	0.84	**0.82**	2.28	**1.11**	1.97	**1.05**	**0.74**	0.78	1.87	**0.92**	1.62	**0.89**

Table 1. Mean reproduction angular error for different methods with the diagonal correction (indicated as D) and the full matrix correction (indicated as T). Results are summarized for outdoor, indoor and all images. The numbers of outdoor images and indoor images for each camera set are shown after the camera's name. For each category, results are summarized for neutral patches, color (non-neutral) patches and all patches. For each category (e.g. Indoor Images/Color), the minimum error result for D versus T is in bold. The Gehler-Shi dataset is divided into two subsets according to the camera used. For color patches only, our method is consistently better for all indoor image and combined image datasets (highlighted by the red background color), with the exception of the Canon 1D images in the Gehler-Shi, which represents the smallest dataset tested.

Figure 9. Visual comparison of Bayesian method results (from both Gelher-Shi and NUS datasets). The first row shows the result from the diagonal model and the second row shows the results from the modified Bayesian method with full matrix model. The color coded reproduction angular errors for each 24 color patches are shown at the left-bottom of each image (red=high error, blue=low error).

References

[1] K. Barnard, L. Martin, B. Funt, and A. Coath. A data set for color research. *Color Research & Application*, 27(3):147–151, 2002. 1, 4

[2] S. Bianco and R. Schettini. Color constancy using faces. In *CVPR*, 2012. 7

[3] S. Bianco and R. Schettini. Error-tolerant color rendering for digital cameras. *Journal of Mathematical Imaging and Vision*, 50(3):235–245, 2014. 2

[4] Z. Botev, J. Grotowski, D. Kroese, et al. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957, 2010. 5

[5] D. H. Brainard and B. A. Wandell. Analysis of the retinex theory of color vision. *JOSA A*, 3(10):1651–1661, 1986. 2

[6] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of The Franklin Institute*, 310(1):1–26, 1980. 2

[7] A. Chakrabarti, K. Hirakawa, and T. Zickler. Color constancy with spatio-spectral statistics. *TPAMI*, 34(8):1509–1519, 2012. 2

[8] A. Chakrabarti, Y. Xiong, B. Sun, T. Darrell, D. Scharstein, T. Zickler, and K. Saenko. Modeling radiometric uncertainty for vision with tone-mapped color images. *TPAMI*, 36(11):2185–2198, 2014. 2

[9] 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. *JOSA A*, 31(5):1049–1058, 2014. 1, 6

[10] D. Cheng, B. Price, S. Cohen, and M. S. Brown. Effective learning-based illuminant estimation using simple features. In *CVPR*, 2015. 2

[11] H. Y. Chong, S. J. Gortler, and T. Zickler. The von kries hypothesis and a basis for color constancy. In *ICCV*, 2007. 2

[12] F. Ciurea and B. Funt. A large image database for color constancy research. In *Color and Imaging Conference*, 2003. 1

[13] M. D. Fairchild. *Color appearance models, 3rd Edition*. John Wiley & Sons, 2013. 2

[14] G. D. Finlayson. Corrected-moment illuminant estimation. In *ICCV*, 2013. 2, 6, 7

[15] G. D. Finlayson, M. S. Drew, and B. V. Funt. Color constancy: enhancing von kries adaption via sensor transformations. In *IS&T/SPIE Electronic Imaging*, 1993. 2

[16] G. D. Finlayson, M. S. Drew, and B. V. Funt. Diagonal transforms suffice for color constancy. In *ICCV*, 1993. 2

[17] G. D. Finlayson, S. D. Hordley, and P. M. Hubel. Color by correlation: A simple, unifying framework for color constancy. *TPAMI*, 23(11):1209–1221, 2001. 2

[18] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, 2004. 2

[19] G. D. Finlayson and R. Zakizadeh. Reproduction angular error: An improved performance metric for illuminant estimation. In *BMVC*, 2014. 6

[20] D. A. Forsyth. A novel algorithm for color constancy. *IJCV*, 5(1):5–35, 1990. 2

[21] D. H. Foster, K. Amano, S. Nascimento, and M. J. Foster. Frequency of metamerism in natural scenes. *JOSA A*, 23(10):2359–2372, 2006. 3

[22] B. Funt and P. Bastani. Intensity independent rgb-to-xyz colour camera calibration. In *AIC (International Colour Association) Conference*, 2012. 5

[23] B. Funt and H. Jiang. Nondiagonal color correction. In *ICIP*, 2003. 2

[24] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp. Bayesian color constancy revisited. In *CVPR*, 2008. 1, 2, 6, 7

[25] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational color constancy: Survey and experiments. *TIP*, 20(9):2475–2489, 2011. 2

[26] A. Gijsenij, T. Gevers, and J. Van De Weijer. Improving color constancy by photometric edge weighting. *TPAMI*, 34(5):918–929, 2012. 2

[27] C.-c. Huang and D.-K. Huang. A study of non-diagonal models for image white balance. In *IS&T/SPIE Electronic Imaging*, 2013. 2

[28] J. Jiang, D. Liu, J. Gu, and S. Susstrunk. What is the space of spectral sensitivity functions for digital color cameras? In *WACV*, 2013. 4

[29] S. J. Kim, H. T. Lin, Z. Lu, S. Susstrunk, S. Lin, and M. S. Brown. A new in-camera imaging model for color computer vision and its application. *TPAMI*, 34(12):2289–2302, 2012. 2

[30] A. D. Logvinenko, B. Funt, and C. Godau. Metamer mismatching. *TIP*, 23(1):34–43, 2014. 3

[31] C. Rosenberg, A. Ladsariya, and T. Minka. Bayesian color constancy with non-gaussian models. In *NIPS*, 2003. 2, 6, 7

[32] L. Shi and B. Funt. Re-processed version of the gehler color constancy dataset of 568 images. accessed from http://www.cs.sfu.ca/~colour/data/. 1

[33] H. Vaezi Joze and M. Drew. Exemplar-based colour constancy and multiple illumination. *TPAMI*, 36(5):860–873, 2014. 2

[34] J. Van De Weijer and T. Gevers. Color constancy based on the grey-edge hypothesis. In *ICIP*, 2005. 2

[35] J. Van De Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *TIP*, 16(9):2207–2214, 2007. 2

[36] J. Von Kries. Beitrag zur physiologie der gesichtsempfindung. *Arch. Anat. Physiol*, 2:505–524, 1878. 2

[37] G. Wyszecki and W. Stiles. *Color science: concepts and methods, quantitative data, and formulae, 2rd Edition*. John Wiley & Sons, 2000. 3