Illuminant Estimation for Color Constancy: Why spatial domain methods work and the role of the color distribution

Dongliang Cheng,[1] Dilip K. Prasad,[1,*] and Michael S. Brown[1]

[1]*School of Computing, National University of Singapore, Singapore 117417*

compiled: March 22, 2014

Color constancy is a well studied topic in color vision. Methods are generally categorized as: 1) low-level statistical methods; 2) gamut-based methods; 3) and learning-based methods. In this work, we distinguish methods depending on whether they work directly from color values (i.e. color domain) or from values obtained from the image's spatial information (e.g. image gradients/frequencies). We show that spatial information does not provide any additional information that cannot be obtained directly from the color distributions and that the indirect aim of spatial domain methods is to obtain large color differences for estimating the illumination direction. This finding allows us to develop a simple and efficient illumination estimation method that chooses bright and dark pixels using a projection distance in the color distribution and then applies PCA to estimate the illumination direction. Our method gives state-of-the-art results on existing public color constancy data sets as well as on our newly collected data set containing 1736 images from 8 different high-end consumer cameras.

OCIS codes: (040.1490) Cameras; (150.2950) Illumination; (330.1715) Color, rendering and metamerism; (330.1720) Color vision.

http://dx.doi.org/10.1364/XX.99.099999

1. Introduction and related work

An image captured by a camera is an integrated signal resulting from the camera's sensitivity of the spectral scene content and scene illumination. Scene illumination can have a notable effect on the overall RGB values of an image, introducing color casts that are perceptually undesirable and that have adverse effects on subsequent processing such as object recognition and tracking. The human visual system has an innate ability to perceive colors under different illumination in a constant manner [1–7]. This ability is aptly termed *color constancy*. For cameras, however, color changes due to illumination must be corrected in post-processing. The key to camera-based color constancy methods is the estimation of the color of the illumination in a scene which is typically modeled as a direction in the camera's RGB colorspace [8]. Based on the estimated illumination, the colorspace is transformed such that the illumination direction lies along the achromatic line in the color space (i.e. $(\mathrm{R} = \mathrm{G} = \mathrm{B})$). This procedure, often called white-balancing, serves to normalize the lighting condition to an ideal achromatic white light and is a crucial component in consumer cameras and pre-processing for many image processing tasks.

Research in color constancy has a long history spanning several decades [1–3]. A full literature review is outside the scope of this paper, however, extensive surveys can be found here [9–11]. Work pre-dating digital cameras was related primarily to color perception by the human vision system. With the rise of consumer digital cameras, research begun to focus on efficient white-balancing methods that could be performed onboard the camera (e.g. . [12–18]). From these efforts it became widely accepted that a 3×3 diagonal matrix is sufficient to perform white balancing [8] - the challenge lies in estimating the illumination direction in the RGB colorspace for a captured image.

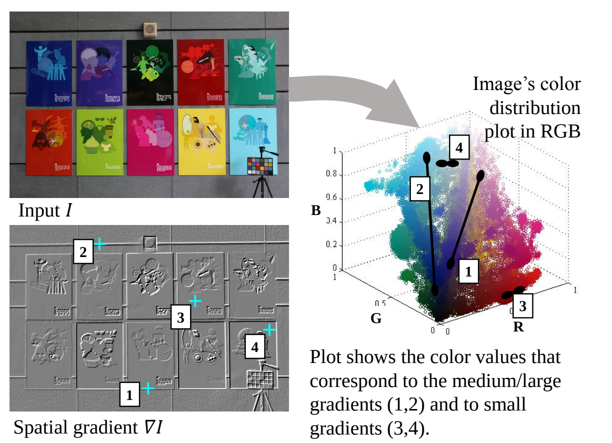

Fig. 1. In the spatial domain methods, gradients serve as a means of computing color differences. Spatial gradients with strong responses can be attributed to scene content whose color values are far apart in the color domain as shown in this figure.

Early work examined statistical properties of the RGB col-

* Corresponding author: brown@comp.nus.edu.sg

orspace (e.g. average RGB value, or max RGB value). These statistics were assumed to provide insight to where the true achromatic values in the scene were and thus insight into the illuminant direction [2, 5, 19]. Statistical approaches are still popular to date [20–23] given their efficiency. In addition to examining the RGB colors directly, spatio-statistical methods were developed (e.g. [3, 21, 24–29]) that used gradient and frequency information (e.g. image derivatives, difference of Gaussians, etc). These methods argued that the illuminant information was correlated with the image's spatial information. Other methods also examined scene content looking for physics-based insight to illumination, such as specularity/highlights [30–34], shadows [1], black-body radiation [35] and inter-reflections [24].

Another popular approach was to consider the finite gamut of the sensor and scenes appearance under different illuminations [36, 37]. This approach later gave rise to machine-learning based methods that use training images to determine both color and spatio-statistic information to help estimate the illuminant (e.g. [27, 28, 38–42]). These gamut-based and learning-based methods often give better performance than statistical and spatio-statical methods, but require significantly more computational power and are not well suited for real-time use.

In this paper, we investigate why the spatial domain methods (i.e. spatio-statistical as well as learning based on spatial information) work and what is their connection to the methods that work directly in the color domain. For this purpose, the color constancy methods are categorized by the type of information they use to estimate illumination, i.e.: (1) methods based on color distribution [2, 20, 22] (2) methods based on spatial information such as image gradients or other spatial differences [21, 25–29].

While the spatial information is known to be important for color constancy in human vision [3], it is intriguing to consider why spatial derivatives might give insight to the scene illumination direction for computational color constancy. While spatial-domain methods clearly show a correlation between spatial changes and illumination direction, the underpinning reason is not clear. Spatial derivatives and their variations (e.g. examining various spatial frequencies) are related to scene albedo change from surface texture and depth discontinuities. More importantly, they are dependent on the spatial relationship of objects in the scene. This makes such approaches sensitive to the scene content. Yet, these methods have seen reasonably good success.

From our analysis, we find that the spatial information serves merely as a means of obtaining samples of color differences in the color domain, and, that the majority of the spatial information is not useful. More specifically, spatial domain methods benefit from the large gradients in the scene which correspond to differences from colors far apart in the color domain (see Fig. 1). This observations lead us to question whether computing this information directly from the color domain might be a better strategy than relying on spatial content. To this end, we introduce a novel illumination estimation method that works from the color domain and selects pixels that describe the illumination well. Our method is simple, efficient, and gives state-of-the-result results. Lastly, as a part of our work, we have produced an image data set of eight current consumer cameras with over 1600 high-quality images where each camera is observing the same scene (see `http://www.comp.nus.edu.sg/~whitebal/illuminant/illuminant.html`. This serves as a useful data set for color constancy research.

The rest of this paper is as follows. Section 2 gives more background on color domain and spatial-domain methods. Section 3 provides analysis into why spatial methods work. Section 4 presents our methods followed by results and a discussion in Sections 5 and 7. Our dataset is discussed in Section 6.

2. Color and spatial domain methods

We discuss color domain and spatial domain methods here. Given the long history of color constancy research, only representative examples are discussed. As previously mentioned, we categorize the approaches based on the information used to estimate the illumination, i.e. RGB values (i.e. color domain) or spatial information. Let an image I be denoted as a collection of vectors $I(x) = [I_R(x) \quad I_G(x) \quad I_B(x)]$, where x indicates the pixels (or corresponding color points in the color domain) and $I_c(x)$ denotes the color value of $c = R, G, B$ color channels.

Color domain approaches Among the methods based on the color domain distribution, the most popular methods are the max-RGB [19, 22] and grey world method [2], along with their variants such as those employing p-norm averages [20]. All these methods are based on statistical hypotheses about the spectral properties of the scene. For example, the *Grey world* method [2] and variants assume that the average of a particular Minkowsky norm of a scene's RGB values is achromatic (in other words a constant for all the three color channels). Thus, performing such a norm average on the color data of an image will estimate the illumination direction. Mathematically, for such approaches, the color constancy matrix $\mathbf{T} = \mathrm{diag}(T)^{-1}$ is given by the illumination direction $T = [t_\mathrm{R} \quad t_\mathrm{G} \quad t_\mathrm{B}]$ which is estimated as:

$$t_c = \frac{\left(\sum_x |I_c(x)|^p\right)^{\frac{1}{p}}}{N}, \tag{1}$$

where $|\circ|$ denotes the absolute value and N is the number of pixels in the image. The max-RGB method is also a subset of this since it considers ($p = \infty$) Minkowsky norm. Here, we note that the average is typically taken on all the pixels (after possibly removing the pixels corresponding to the saturation and dark noise). This is a general approach. However, more specific choice of pixels is also considered at times. For example, pixels corresponding to specularity only may be chosen [31].

Spatial domain methods In the spatial domain methods, a spatial domain operator $f(I)$ is applied on the image I to obtain a transformed image J:

$$J(x) = f(I(x)), \forall x \tag{2}$$

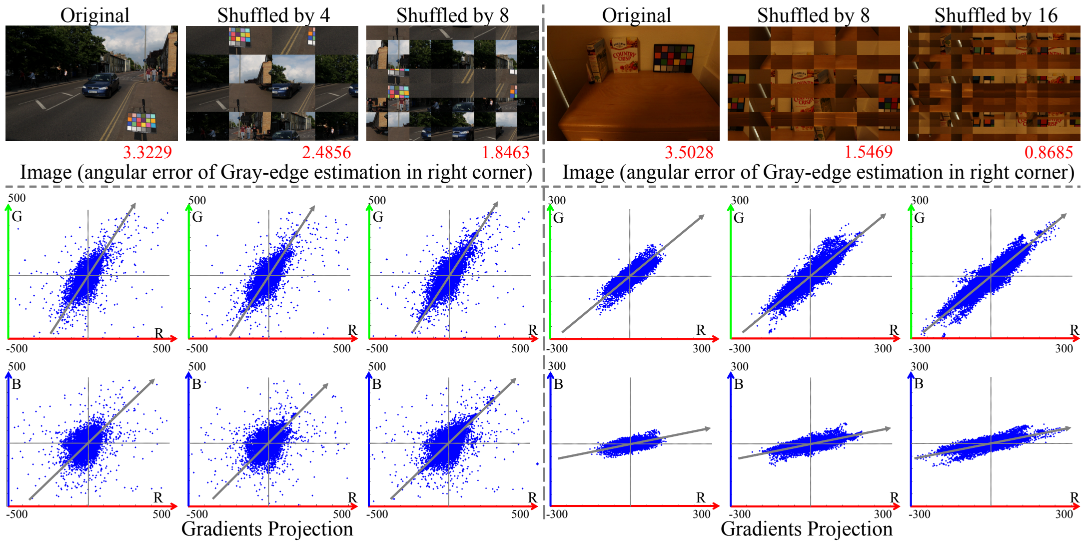

Fig. 2. This figure shows an example images from [40] where synthetic gradients are introduced by shuffling the image by blocks (top row). Note that the scene content and overall color distribution does not change. The gradients of these images projected on different color planes show that introduction of new gradients makes the distribution more elongated and directional. This shuffling actually improves the illumination estimation for a well known spatial technique [21].

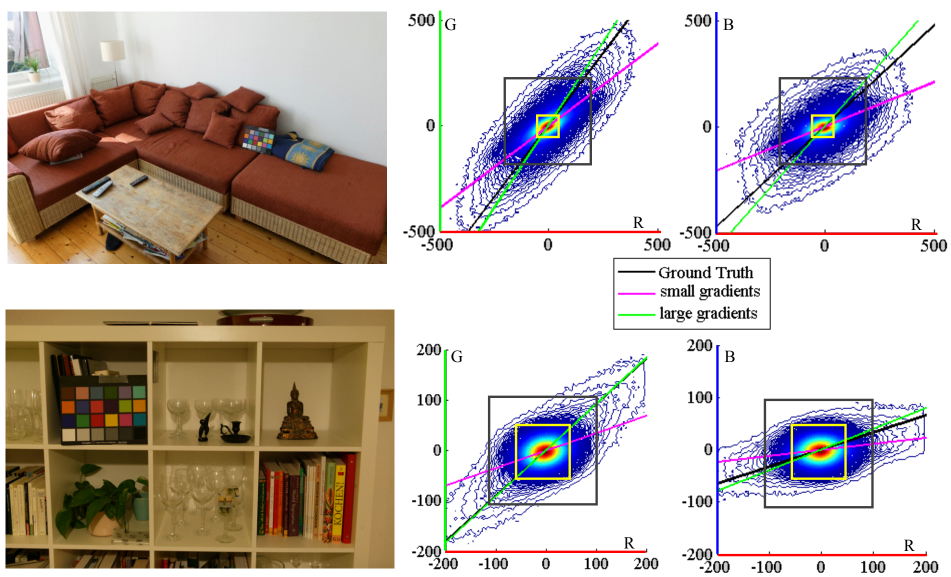

Fig. 3. Probability map of the image(from [40]) gradients and the illumination information in the various regions of this map are shown. The magenta lines show the PCA vector of the pixels inside the yellow box (i.e. illumination information in small gradients) and the green lines show the PCA vector of the pixels outside the grey box (i.e. illumination information in large gradients). Black lines show the ground truth.

These methods operate directly in the transformed image J. For example, the *Grey edge* [21, 28] hypothesizes that the derivatives of an image in the spatial domain represent achromatic color. As with grey world, a pth Minkowsky norm can be used as in eq. (1) to estimate the illumination direction operating on the J instead of I.

An enhanced version of the grey edge method is the weighted grey edge [25, 29] where the edges are classified according to physical properties such as specularity, shadows, etc. The operator $f(I)$ can be represented as a weighted n^{th} order derivative:

$$J(x) = w(x)\nabla^n I(x) \tag{3}$$

where $w(x)$ is the weight given to a pixel based on photometric classifications, such as discussed above.

Other spatial domain methods use operators such as Gaussian filter, discrete cosine transform [27], discrete wavelet transform [26] etc. The idea is to suppress/remove the smooth portion of the data and keep only the spatial high frequency components (equivalent to derivatives) in the image [26].

3. Why spatial domain methods work

As stated in Section 1, our focus is to investigate what makes spatial domain methods work. Work in [39] provided an earlier insight into this by considering the correlation between all pixels in an image with one another. This work argued that spatial domain methods can be thought of as a subset of this exhaustive correlation approach, where only correlation between local neighborhoods of pixels are considered. Here, we provide a much more direct analysis to give insight into why spatial domain methods work. We do this using two experiments to help reveal the relationship of the spatial information to image samples in color domain.

Introducing artificial gradients We first look at synthetically introducing gradients in an image by dividing the image into uniform blocks and randomly shuffling the blocks to create a new image. For this new image, neither the illumination, color distribution, nor the net image content has changed. This new image does have new image gradients due to the boundaries created by the shuffled blocks, but these gradients are artificial and do not represent anything physical about the scene. Such manipulation will have no effect on color domain approaches. However, for spatial domain approaches this has a surprisingly positive effect on the illumination estimation.

Figure 2 shows two examples. The top row shows two images divided into different number of blocks which have been shuffled. The bottom rows shows the ground truth illumination (the grey arrow) and plotted gradients against the R-G and R-B planes. As the number of blocks increases, the number of large gradients increases. These new gradients correspond to large color differences at the edges of the blocks. More importantly, these new gradients are completely artificial and have no physical meaning. It is interesting to see that these new large gradients also appear to be following the direction of the illumination. The addition of these gradients improves the Grey Edge algorithm [21]. This is shown by the angular error from the ground truth which decreases as the shuffling increases. The angular error $\varepsilon_{\mathrm{angle}}(\mathbf{e}_{\mathrm{est}})$ of the estimated illumination direction $\mathbf{e}_{\mathrm{est}}$ from the illumination direction of the ground truth $\mathbf{e}_{\mathrm{gt}}$ is computed as follows:

$$\varepsilon_{\mathrm{angle}}(\mathbf{e}_{\mathrm{est}}) = \cos^{-1}\left(\frac{\mathbf{e}_{\mathrm{est}} \cdot \mathbf{e}_{\mathrm{gt}}}{\|\mathbf{e}_{\mathrm{est}}\|\,\|\mathbf{e}_{\mathrm{gt}}\|}\right). \quad (4)$$

Gradient analysis Our second experiment examines how gradients contribute to illumination estimation. It is well known that natural images have significantly large number of small valued gradients and a sparse number of large gradients [43]. This means that we should expect the majority of the gradients obtained for spatial methods to be small valued. This is shown in Fig.3 on two example images. The gradients probability map shows the relative occurrence of a particular gradient value. Here we have considered the horizontal spatial derivative for simplicity; the vertical derivative shows a similar trend.

The goal here is to investigate the contribution of low valued gradients (i.e. the majority of the gradients) to the illumination estimation. For this we consider the gradients *inside* the yellow boxes shown in Fig. 3. Using the pixels that lie *inside* the yellow box only, we compute and plot the dominant direction in this distribution using principal component analysis (PCA). The result is shown in Fig. 3) using a magenta colored line. The ground truth illuminant is shown as a black line. Further, we consider if the large gradients are helpful in illumination estimation. For this, we consider the pixels *outside* the grey box and compute the PCA of the large gradients in Fig. 3), which is shown using green colored line.

It can be seen that the illumination estimation for small gradients (inside yellow box) have more angular deviation from the ground truth than the pixels with higher gradients (outside grey box). Thus, small gradients can actually bias the solution in an erroneous manner. Thus, removing these small gradients through heuristics is a way to improve the performance of spatial methods. Such heuristics may involve identification of the pixels lying along edges, specularities, or shadows.

Both of these experiments serve to underscore that large color differences are key to illumination estimation. Moreover, our first experiment shows that relying on the scene content to provide these differences may not be the best strategy. Simply by shuffling the image content to introduce artificial gradients, we were able to obtain better results. This begs the question if we can design a method to obtain similar large color differences in the color domain directly and bypass the reliance on the spatial content to give us these differences.

4. Proposed method

Based on our findings in Section 3 we propose a new method that selects colors in the color domain distribution that effectively provide large differences. This is similar to examining large gradients without the reliance on the scene content to guide the selection of the colors. Our method is described in the following.

4.A. Selection of colors

It was empirically shown for the gradient domain in [28] that specular pixels and shadow pixels help in reducing the error of illumination estimation. Such observations were also reported in [30–34]. It is interesting to note that the pixels that lie on the edges of specular and shadow generally represent regions with notable color differences between them in the color domain. In practice, however, selecting specular and shadow pixels only from an image is not straight forward. For example, we have to distinguish between specularities and bright surfaces and shades and dull surfaces. Only advanced image processing and the knowledge of scene and camera's spectral properties could give a good classification for them.

In this sense, if we choose just the bright and dark pixels in the image, we can have the clusters of points with largest color differences between the clusters. Doing so has several advantages. First, we are no more dependent on the scene's actual content and spatial correspondence for color estimation. This is important because sometimes two images of different scenes may result in drastically different estimation

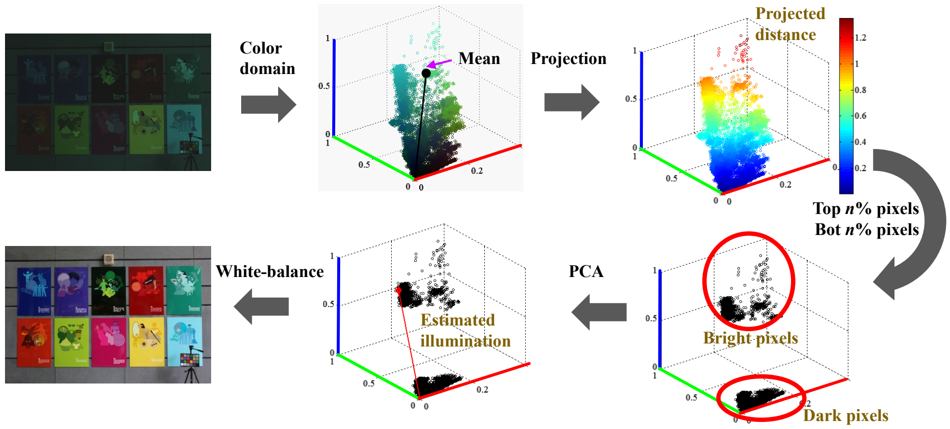

Fig. 4. This figure shows the framework of the proposed method.

Fig. 5. These images(from [40]) of different scenes are taken in the same illumination, but the error in illumination estimation using spatial domain methods is quite different for the two images. The labels in the top corners of the images show the angular errors of the grey edge (GE) and weighted grey edge (WGE) algorithm.

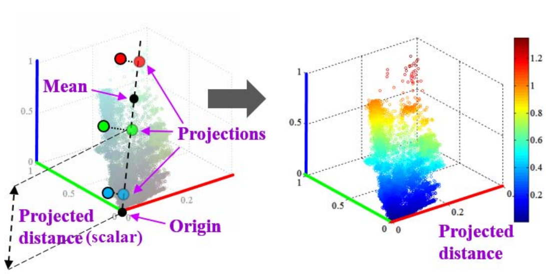

Fig. 6. Illustration of the projection distance used in eq. (5).

using spatial domain methods, see Fig. 5 for example. Second, such approach does away with computationally involved spatial domain processing such as filtering. Third, we do not need to compute photometric pixels having qualities such as specularity or shade. This is quite handy since either the classification of such pixels require sophisticated and advanced processing or very crude approximations are used to classify them with large error probability.

4.B. Our Algorithm

An illustration of the proposed method is shown in Fig. 4. We first compute the projection of all the color points in the color domain on the direction of the mean vector. The projected distances be denoted as d_x, where x is the index of a color point. The term d_x is a scalar distance given as:

$$d_x = \frac{I(x) \cdot I_0}{\|I(x)\| \|I_0\|}, \tag{5}$$

where $\|A\|$ denotes the Euclidean norm of a vector A, $A \cdot B$ represents the vector dot product of the vectors A and B, and the vector I_0 is given as:

$$I_0 = [t_{\mathrm{R}} \quad t_{\mathrm{G}} \quad t_{\mathrm{B}}], \tag{6}$$

where t_c is given by eq. (1) with $p = 1$. Further, $I(x) = [I_R(x) \quad I_G(x) \quad I_B(x)]$ is the vector containing the RGB color values of a color point x. The projection distance is illustrated in Fig. 6. We then sort the color points in the ascending order of the projection distances d_x. Then we choose the top $n\%$ and bottom $n\%$ of color points, thus selecting the color points with largest and smallest projections on the mean vector.

Then we compute the first PCA vector of the data matrix formed using $I(x)$ corresponding to the selected pixels only. This vector is taken as the estimated illumination direction. The effect of control parameter n on the performance of our method is shown in Fig. 7 using the mean and median errors for the Color Checker dataset [40]. It is seen that the median error is the lowest at $n = 3.5\%$. We note that while our method is simple, our results show that it is quite effective in estimating the illumination.

5. Results

We show our results on three data sets. The first is the well established *SFU data set* [44] comprising of 321 images taken in laboratory scenario with controlled scenes and illuminants. The second is the more recent *Color Checker dataset* [40] comprising of 568 images of natural scenarios with natural

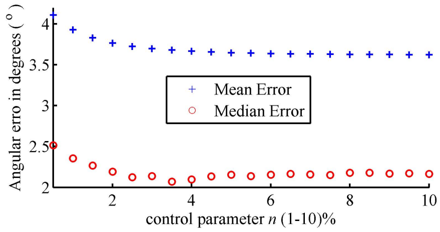

Fig. 7. Effect of the control parameter n on the performance of the proposed method.

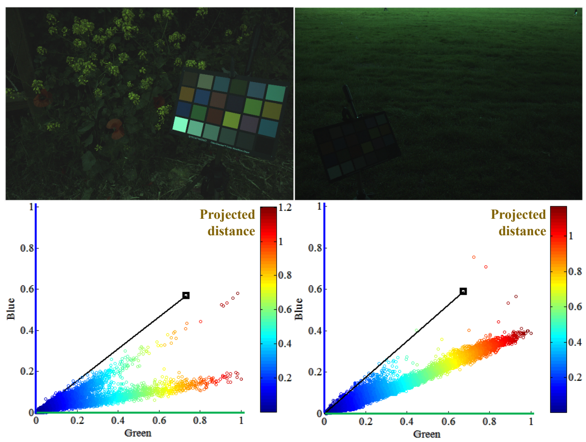

Fig. 8. Strongly axial color distribution causes failure for most methods of color constancy. The black vectors in the bottom row are the actual illumination vectors (i.e. ground truth).

scenes and illuminations. These two data sets are currently the standards used when comparing color constancy algorithms. We have also collected a new data set of 1736 images taken from 8 different cameras. Additional 117 images from Nikon D40 camera and its results for all the methods is made available on our project website. The details of our dataset are presented in section 6.

We compare our results against 14 existing techniques that represent a wide range of color constancy techniques (see Tables 2 and 3). We have used angular error (eq. 4) as the error

Fig. 9. This figure shows some examples from the Color Checker dataset [40] in which the error in illumination estimation is high when only bright pixels (B) are used and reduces significantly when both bright and dark (BD) pixels are used.

Method	SoG	GGW	BP	GE1	GE2	PG	EG	IG
Parameters	p	p, σ	$p, \%$	p, σ	p, σ	σ	σ	σ
CC	4	9,9	2,2	1,6	1,1	4	4	9
SFU	7	10,5	2,0.5	7,4	7,5	4	2	4
Canon1	3	1,9	2,3	3,6	9,9	10	7	9
Canon2	3	3,9	4,3	9,3	3,3	8	10	9
Fuji	3	3,9	4,3	3,3	3,3	10	10	10
Nikon	3	3,9	4,3	3,3	9,3	8	3	8
Oly	9	1,1	2,3	3,1	3,1	9	10	9
Pan	9	1,1	1,5	1,1	3,1	10	10	10
Sam	9	1,1	4,3	1,1	9,3	10	4	10
Sony	3	1,9	2,3	9,9	3,3	7	8	7

Table 1. Control parameters used by various methods. Abbreviations of methods and datasets are the same as Table 2.

metric to evaluate the methods as it is most widely used in evaluating color constancy algorithms [9] and is correlated to the perceptual Euclidean distance [45]. The mean, median, tri-mean, and maximum angular errors of most state of the art methods and our method for various datasets are shown in Table 2. The error for best 25% images and worst 25% images are listed in Table 3. We have used $n = 3.5\%$ for generating our results. The control parameters of the other methods are shown in Table 1. The control parameters have been chosen as recommended in the respective papers and the color constancy website `http://colorconstancy.com/`. With these guidelines, the control parameters producing optimal results (minimum mean errors) were chosen for reporting the results of other methods for our datasets. Results for SFU dataset [44] and ColorChecker dataset [40] are reported as reported in `http://colorconstancy.com/`. Results of few methods for SFU and ColorChecker datasets are kept blank as the data was not reported previously.

The training and test times for our Canon1 dataset are reported in Table 4. All the results were generated on Intel Core i5 @3.2GHz with 4GB RAM using Matlab 2010.

It is seen that our method performs reasonably well for all the datasets in terms of the mean, median, tri-mean errors, and errors for best 25 % images. Our method performs the poorest on the SFU dataset. The reason for the poor performance is discussed in the failure cases below. In all the other datsets, the error metrics of our method compete well against the other methods. Our method often has the least error metric and in other cases error metric quite close to the least value. Often, the methods that perform better than our method for a given metric and dataset are based on machine learning or gamut fitting (collectively called learning based methods). Methods in both these classes use images in the same dataset for 3-fold training and validation before testing is done on the same images. The 3-fold learning is used for maintaining consistency with previously reported results on `http://colorconstancy.com/`. Thus, it is not too surprising that error metrics of these methods are often quite small. Nevertheless, it is not guaranteed that these methods will always result in very small errors since their performance is quite sensitive to the choice of control parameter.

Training also imposes a high computational requirement on learning based methods, as is confirmed in the large training times reported in Table 4. In addition, as noted in Table 4, the test times are also large for such methods. On the other hand, our method takes just a few minutes (including image read time) for the 259 images of dataset Canon1. Thus, it is seen that our method provides a good combination of accuracy and speed and does not need prior learning. It has been observed that training and testing times increases rapidly with the increase of the control parameters σ (most learning methods require higher sigma to obtain better results) and increase in size of the image in the dataset. Further, it was noted in [46, 47] that an angular error of 3° is perceptually acceptable. As noted in our statistics for median and tri-mean errors, the performance of our method is perceptually acceptable for most cases.

Failure cases It is well-known that if the scene is biased to contain shades of only one or two colors, then the projection of the illuminated scene on the camera sensor is strongly biased along one or two directions in the color domain. This makes the illumination estimation to lie along either of these direction or somewhere between them. Two such examples from Color Checker dataset [40] are shown in Fig. 8. In such cases, most methods that effectively use color domain statistics (which includes spatial domain methods), including ours, result in poor estimation of illumination. Most images in the SFU dataset are illuminated using unusual red and blue illuminants. This gives a similar effect as having only one or two colors in the scene and biases the color domain distribution to lie along only one or two directions in the color domain. As a result, many statistical methods, including ours, perform poorly for the SFU dataset as can be observed in Table 2 and Table 3.

Gamut mapping and machine learning methods are expected to perform better, since they do not use single image to estimate the illumination and rather use pre-learnt priors. Indeed this assumes that hopefully diverse set of images were used for training such that large part of the color domain is spanned by the training data and the test image is a subset of the color domain portion used for training.

The role of dark pixels It might be argued that bright pixels may be sufficient for illumination estimation. In Fig. 9, we show some examples in which bright pixels are not sufficient for illumination estimation and using both bright and dark pixels reduces the error significantly. In fact, for the 568 images in the Color Checker dataset [40], the illumination estimation using bright and dark pixels (BD) is better than using bright pixels (B) alone for 220 images.

6. Our Data set

We have captured a new image data set similar in nature to the Color Checker dataset [40], however, with more images and up to date camera models. In addition, our dataset has images of the same scene with the different cameras, something not done in the previous methods and datasets. This gives a way to compare the performance accross different cameras on the same input. Our data set is composed of images from 9 commercial cameras: Canon 1D-

Fig. 10. Examples of images in our dataset are shown here.

S Mark III (*Canon1* in Table 2), Canon 600D (*Canon 2* in Table 2), Fujifilm XM1 (*Fuji* in Table 2), Nikon D5200 (*Nikon* in Table 2), Olympus EPL6 (*Oly* in Table 2), Panasonic GX1 (*Pana* in Table 2), Samsung NX 2000 (*Sam* in Table 2)), and Sony α57 (*Sony* in Table 2). For these cameras, we captured over 200 images each(Canon1(259 images), Canon2(200 images), Fuji(196 images), Nikon(200 images), Oly(208 images), Pana(203 images), Sam(203 images), and Sony(268 images)), such that the scene and illumination is the same for all the 8 cameras. There are slight misalignment issues because the camera positioning cannot be exactly ensured, but these errors are small. The images are taken in natural settings, both indoor and outdoor. For outdoor, the sunny and shade conditions are considered. For indoor, various common commercial lightings are considered (e.g. tungsten, florescent, etc). Example from the data set are shown in Fig. 10. Our complete dataset (around 100GB in compressed state) and our matlab source code are all publicly available at `http://www.comp.nus.edu.sg/~whitebal/illuminant/illuminant.html`.

7. Discussion and Conclusion

We have observed that spatial and gradient domain methods works because of color differences, which can be easily obtained from color domain. We have also seen that not only the bright pixels [31] are important for illumination estimation but also the dark pixels are important for illumination estimation. Our method based on bright and dark pixels chosen using the projection distance in the color domain performs better than most non-machine learning methods for natural images across various consumer cameras. Though a comparison of non-learning based methods with learning based methods is not completely fair, but for the sake of completeness we have compared our method with both non-learning based methods (GW [2], WP [19], SoG [20], GGW [11], BP

	Statistics based methods								Learning based methods						
Method	Our	GW	WP	SoG	GGW	BP	GE1	GE2	PG	EG	IG	BL	ML	GP	NIS
Dataset	Mean angular error (degrees °)														
CC	**3.52**	6.36	7.55	4.93	4.66	–	5.33	5.13	4.20	6.52	4.20	4.82	3.67	3.59	4.19
SFU	6.07	9.78	9.09	6.39	5.41	–	5.58	5.19	3.70	3.92	**3.62**	–	5.63	–	–
Canon1	**2.93**	5.16	7.99	3.81	3.16	3.37	3.45	3.47	6.13	6.07	6.37	3.58	3.58	3.21	4.18
Canon2	2.81	3.89	10.96	3.23	3.24	3.15	3.22	3.21	14.51	15.36	14.46	3.29	2.80	**2.67**	3.43
Fuji	3.15	4.16	10.20	3.56	3.42	3.48	3.13	3.12	8.59	7.76	6.80	3.98	3.12	**2.99**	4.05
Nikon	**2.90**	4.38	11.64	3.45	3.26	3.07	3.37	3.47	10.14	13.00	9.67	3.97	3.22	3.15	4.10
Oly	**2.76**	3.44	9.78	3.16	3.08	2.91	3.02	2.84	6.52	13.20	6.21	3.75	2.92	2.86	3.22
Pan	2.96	3.82	13.41	3.22	3.12	3.05	2.99	2.99	6.00	5.78	5.28	3.41	2.93	**2.85**	3.70
Sam	**2.91**	3.90	11.97	3.17	3.22	3.13	3.09	3.18	7.74	8.06	6.80	3.98	3.11	2.94	3.66
Sony	**2.93**	4.59	9.91	3.67	3.20	3.24	3.35	3.36	5.27	4.40	5.32	3.50	3.24	3.06	3.45
Dataset	Median angular error (degrees °)														
CC	**2.14**	6.28	5.68	4.01	3.48	–	4.52	4.44	2.33	5.04	2.39	3.46	2.96	2.96	3.13
SFU	3.01	7.00	6.48	3.74	3.32	–	3.18	2.74	2.27	2.28	**2.09**	–	3.45	–	–
Canon1	**2.01**	4.15	6.19	2.73	2.35	2.45	2.48	2.44	4.30	4.68	4.72	2.80	2.80	2.67	3.04
Canon2	**1.89**	2.88	12.44	2.58	2.28	2.48	2.07	2.29	14.83	15.92	14.72	2.35	2.32	2.03	2.46
Fuji	2.15	3.30	10.59	2.81	2.60	2.67	**1.99**	2.00	8.87	8.02	5.90	3.20	2.70	2.45	2.95
Nikon	**2.08**	3.39	11.67	2.56	2.31	2.30	2.22	2.19	10.32	12.24	9.24	3.10	2.43	2.26	2.40
Oly	**1.87**	2.58	9.50	2.42	2.15	2.18	2.11	2.18	4.39	8.55	4.11	2.81	2.24	2.21	2.17
Pan	**2.02**	3.06	18.00	2.30	2.23	2.15	2.16	2.04	4.74	4.85	4.23	2.41	2.28	2.22	2.28
Sam	**2.03**	3.00	12.99	2.33	2.57	2.49	2.23	2.32	7.91	6.12	6.37	3.00	2.51	2.29	2.77
Sony	**2.33**	3.46	7.44	2.94	2.56	2.62	2.58	2.70	4.26	3.30	3.81	2.36	2.70	2.58	2.88
Dataset	Tri-mean error (degrees °)														
CC	**2.47**	6.28	6.35	4.23	3.81	–	4.73	4.62	2.91	5.43	2.93	3.88	3.10	3.04	3.45
SFU	3.69	7.60	7.45	4.59	3.78	–	3.74	3.25	**2.53**	2.70	2.38	–	4.33	–	–
Canon1	**2.22**	4.46	6.98	3.06	2.50	2.67	2.74	2.70	4.81	4.87	5.13	2.97	2.97	2.79	3.30
Canon2	**2.12**	3.07	11.40	2.63	2.41	2.47	2.36	2.37	14.78	15.73	14.80	2.40	2.37	2.18	2.72
Fuji	2.41	3.40	10.25	2.93	2.72	2.82	**2.26**	2.27	8.64	7.70	6.19	3.33	2.69	2.55	3.06
Nikon	**2.19**	3.59	11.53	2.74	2.49	2.44	2.52	2.58	10.25	11.75	9.35	3.36	2.59	2.49	2.77
Oly	**2.05**	2.73	9.54	2.59	2.35	2.36	2.26	2.20	4.79	10.88	4.63	3.00	2.34	2.28	2.42
Pan	2.31	3.15	14.98	2.48	2.45	2.30	**2.25**	2.26	4.98	5.09	4.49	2.58	2.44	2.37	2.67
Sam	**2.22**	3.15	12.45	2.45	2.66	2.64	2.32	2.41	7.70	6.56	6.40	3.27	2.63	2.44	2.94
Sony	**2.42**	3.81	8.78	3.03	2.68	2.73	2.76	2.80	4.45	3.45	4.13	2.57	2.82	2.74	2.95
Dataset	Maximum angular error (degrees °)														
CC	28.35	24.83	40.58	22.40	22.04	–	26.35	23.88	23.18	28.99	24.22	24.48	**21.58**	21.64	26.20
SFU	44.00	37.31	36.22	29.60	28.93	–	31.55	26.74	27.10	27.70	27.10	–	**21.56**	–	–
Canon1	16.20	22.37	39.12	15.74	16.72	18.87	17.69	15.73	29.09	33.59	28.96	**13.54**	**13.54**	16.62	21.43
Canon2	17.33	15.93	22.76	**15.08**	18.38	17.56	17.86	17.68	22.54	22.48	22.59	15.60	15.43	15.54	20.16
Fuji	21.16	21.06	25.10	18.55	20.83	21.45	22.79	24.44	21.73	21.89	19.68	18.32	18.75	**15.07**	28.54
Nikon	**15.50**	20.61	53.08	15.53	15.54	15.61	23.57	24.33	33.72	60.87	33.73	17.85	17.65	16.63	56.44
Oly	23.28	16.46	25.11	16.99	22.20	18.11	20.57	19.58	18.85	53.56	34.03	22.22	15.14	**14.21**	16.53
Pan	16.59	16.74	23.89	18.47	17.61	17.97	21.15	20.03	26.91	52.08	24.75	19.51	15.29	**14.54**	21.34
Sam	15.52	17.32	23.99	13.80	**12.41**	14.11	20.90	20.85	18.09	29.40	18.35	18.12	15.76	14.04	15.25
Sony	**12.39**	17.84	39.78	13.79	17.89	12.94	15.04	15.78	50.45	32.70	50.42	18.05	15.63	14.78	12.96

Table 2. Comparison of mean, median, tri-mean, and maximum angular errors of our method with other methods for various datasets is shown here. Abbreviations of methods: Grey World [2] (GW), White Patch [19] (WP), Shades of Grey [20] (SoG), Generalized Grey World [11] (GGW), Bright-Pixels [31] (BP), Grey Edge - 1st order [21] (GE1), Grey Edge - 2nd order [21] (GE2), Pixels based Gamut [36] (PG), Edge based Gamut [36] (EG), Intersection based gamut [36] (IG), Bayesian learning [40] (BL), Spatio-spectral learning [27] (ML), Spatio-spectral learning using Gen-prior [27] (GP), Natural Image Statistics [28] (NIS). Abbreviations of datasets: Color Checker Set of [40] (CC), SFU laboratory dataset of [44].

	Statistics based methods								Learning based methods						
Method	Our	GW	WP	SoG	GGW	BP	GE1	GE2	PG	EG	IG	BL	ML	GP	NIS
Dataset	Error for best 25% images (degrees °)														
CC	**0.50**	2.33	1.45	1.14	1.00	–	1.86	2.11	**0.50**	1.90	0.51	1.26	0.95	0.91	1.00
SFU	0.67	0.89	1.84	0.59	0.49	–	1.05	1.10	**0.46**	0.51	0.50	–	1.23	–	–
Canon1	**0.59**	0.95	1.56	0.66	0.64	0.62	0.81	0.86	1.05	1.38	1.18	0.76	0.76	0.88	0.78
Canon2	**0.55**	0.83	2.03	0.64	0.63	0.67	0.73	0.80	9.98	11.23	10.02	0.69	0.72	0.68	0.78
Fuji	**0.65**	0.91	1.82	0.87	0.73	0.76	0.72	0.70	3.44	2.30	2.18	0.93	0.75	0.81	0.86
Nikon	**0.56**	0.92	1.77	0.72	0.63	0.59	0.79	0.73	4.35	3.92	4.05	0.92	0.91	0.86	0.74
Oly	**0.55**	0.85	1.65	0.76	0.72	0.63	0.65	0.71	1.42	1.55	1.38	0.91	0.86	0.78	0.76
Pan	0.67	0.82	2.25	0.78	0.70	0.66	**0.56**	0.61	2.06	1.76	1.54	0.68	0.84	0.82	0.79
Sam	**0.66**	0.81	2.59	0.78	0.77	0.81	0.71	0.74	2.65	3.00	2.25	0.93	0.80	0.75	0.75
Sony	**0.78**	1.16	1.44	0.98	0.85	0.81	0.79	0.89	1.28	0.99	1.11	**0.78**	0.93	0.87	0.83
Dataset	Error for worst 25% images (degrees °)														
CC	8.74	10.58	16.12	10.20	10.09	–	10.03	9.26	10.72	13.58	10.70	10.49	7.61	**7.43**	9.22
SFU	16.82	23.45	20.97	16.49	13.75	–	14.05	13.51	9.32	9.91	**9.38**	–	12.90	–	–
Canon1	6.82	11.00	16.75	8.52	7.08	7.82	7.69	7.76	14.16	13.35	14.47	7.95	7.95	**6.43**	9.51
Canon2	6.50	8.53	18.75	7.06	7.58	7.22	7.48	7.41	18.45	18.66	18.29	7.93	5.99	**5.77**	7.76
Fuji	7.30	9.04	18.26	7.55	7.62	7.68	7.32	7.23	13.40	13.44	12.51	8.82	6.93	**5.99**	9.37
Nikon	**6.73**	9.69	21.89	7.69	7.53	7.01	8.42	8.21	15.93	24.33	16.18	8.18	6.88	6.90	10.01
Oly	6.31	7.41	18.58	6.78	6.69	6.30	6.88	6.47	15.42	30.21	14.41	8.19	**6.09**	6.14	7.46
Pan	6.66	8.45	20.40	7.12	6.86	6.95	7.03	6.86	12.19	11.38	10.70	8.00	6.07	**5.90**	8.74
Sam	6.48	8.51	20.23	6.92	6.85	6.57	7.00	7.23	13.01	16.27	11.98	8.62	6.46	**6.22**	8.16
Sony	**6.13**	9.85	21.27	7.75	6.68	6.78	7.18	7.14	11.16	9.83	11.93	8.02	6.55	6.17	7.18

Table 3. Comparison of best-25% and worst-25% of our method with other methods for various datasets is shown here.

Method	Our	GW	WP	SoG	GGW	BP	GE1	GE2	PG	EG	IG	BL	ML	GP	NIS
Train(min)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	254	245	251	32.2	133.2	126.9	453.2
Test(min)	9.9	7.8	8.0	14.6	27.3	13.6	29.5	34.6	254	184	235	2316	168.3	61.7	25.2

Table 4. Training and testing time (in minutes) for our Canon1DsMarkIII dataset (trends are similar for the other 8 cameras in our dataset).

[31], GE1 [21], GE2 [21]) as well as learning based methods (PG [36], EG [36], IG [36], BL [40], ML [27], GP [27], NIS [28]). Our method performs better than most non-learning based methods and performs similar or close to the learning based methods in terms of several practically useful error metrics. Further our method is computationally fast and practically more useful than the learning based methods.

We conclude with three highlights of our method. First, instead of using statistical moments such as in eq. (1), we use the first PCA vector for estimating the illumination direction that inherently considers the first and second order moments of the data. Second, instead of using intensity values for determining bright and dark pixels, we use a projection based distance measure to determine the bright and dark pixels. This allows the pixels to be ranked according to their deviation from the statistical mean of the data. Third, unlike other works such as [31], we consider the dark pixels as well for illuminant estimation. Lastly, our dataset of the same scene under same illumination for 8 cameras will be a useful resource for future research in color constancy.

Acknowledgement

This study was funded by Agency for Science, Technology and Research (grant no. 1121202020). We thank our photographer Mr. Looi Wenhe (Russell) for his hardwork and effort in collecting our dataset. We thank Dr. Peter Gehler [40] for giving his permission to use images from his dataset for Figs. 2,3,5,8, and 9.

References

[1] S. M. Newhall, R. W. Burnham, and R. M. Evans, "Color constancy in shadows," J. Opt. Soc. Am. **48**, 976–984 (1958).

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

[3] K. T. Blackwell and G. Buchsbaum, "Quantitative studies of color constancy," J. Opt. Soc. Am. A **5**, 1772–1780 (1988).

[4] J. S. Werner and B. E. Schefrin, "Loci of achromatic points throughout the life span," J. Opt. Soc. Am. A **10**, 1509–1516 (1993).

[5] Q. Zaidi, B. Spehar, and J. DeBonet, "Color constancy in variegated scenes: role of low-level mechanisms in discounting illumination changes," J. Opt. Soc. Am. A **14**, 2608–2621 (1997).

[6] K.-H. Bäuml, "Increments and decrements in color constancy," J. Opt. Soc. Am. A **18**, 2419–2429 (2001).

[7] N. N. Krasilnikov, O. I. Krasilnikova, and Y. E. Shelepin, "Mathematical model of the color constancy of the human visual system," J. Opt. Technol. **69**, 327 (2002).

[8] 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).

[9] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," IEEE Transactions on Image Processing **20** (2011).

[10] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms. i: Methodology and experiments with synthesized data," IEEE Transactions on Image Processing **11**, 972–984 (2002).

[11] K. Barnard, L. Martin, A. Coath, and B. Funt, "A comparison of computational color constancy algorithms. ii. experiments with image data," IEEE Transactions on Image Processing **11**, 985–996 (2002).

[12] C. van Trigt, "Linear models in color constancy theory," J. Opt. Soc. Am. A **24**, 2684–2691 (2007).

[13] S. D. Hordley and G. D. Finlayson, "Reevaluation of color constancy algorithm performance," J. Opt. Soc. Am. A **23**, 1008–1020 (2006).

[14] M. D'Zmura and G. Iverson, "Color constancy. i. basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces," J. Opt. Soc. Am. A **10**, 2148–2165 (1993).

[15] M. D'Zmura and G. Iverson, "Color constancy. ii. results for two-stage linear recovery of spectral descriptions for lights and surfaces," J. Opt. Soc. Am. A **10**, 2166–2180 (1993).

[16] M. D'Zmura and G. Iverson, "Color constancy. iii. general linear recovery of spectral descriptions for lights and surfaces," J. Opt. Soc. Am. A **11**, 2389–2400 (1994).

[17] B. Funt and H. Jiang, "Nondiagonal color correction," in "International Conference on Image Processing," (2003), pp. I–481.

[18] G. Iverson and M. D'Zmura, "Criteria for color constancy in trichromatic bilinear models," J. Opt. Soc. Am. A **11**, 1970–1975 (1994).

[19] D. H. Brainard and B. A. Wandell, "Analysis of the retinex theory of color vision," J. Opt. Soc. Am. A **3**, 1651–1661 (1986).

[20] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in "Color and Imaging Conference," (2004), pp. 37–41.

[21] J. Van De Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," IEEE Transactions on Image Processing **16**, 2207–2214 (2007).

[22] L. Shi and B. Funt, "Maxrgb reconsidered," Journal of Imaging Science and Technology **56**, 20501–1 (2012).

[23] M. P. Lucassen, T. Gevers, A. Gijsenij, and N. Dekker, "Effects of chromatic image statistics on illumination induced color differences," J. Opt. Soc. Am. A **30**, 1871–1884 (2013).

[24] M. S. Drew and B. V. Funt, "Variational approach to interreflection in color images," J. Opt. Soc. Am. A **9**, 1255–1265 (1992).

[25] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor - outdoor image classification," IEEE Transactions on Image Processing **17**, 2381–2392 (2008).

[26] T. Celik and T. Tjahjadi, "Adaptive colour constancy algorithm using discrete wavelet transform," Computer Vision and Image Understanding **116**, 561–571 (2012).

[27] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy with spatio-spectral statistics," IEEE Transactions on Pattern Analysis and Machine Intelligence **34**, 1509–1519 (2012).

[28] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," IEEE Transactions on Pattern Analysis and Machine Intelligence **33** (2011).

[29] A. Gijsenij, T. Gevers, and J. Van De Weijer, "Improving color constancy by photometric edge weighting," IEEE Transactions on Pattern Analysis and Machine Intelligence **34**, 918–929 (2012).

[30] H.-C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," J. Opt. Soc. Am. A **3**, 1694–1699 (1986).

[31] H. R. V. Joze, M. S. Drew, G. D. Finlayson, and P. A. T. Rey, "The role of bright pixels in illumination estimation," in "Color and Imaging Conference," (2012), pp. 41–46.

[32] M. S. Drew, H. R. V. Joze, and G. D. Finlayson, "Specularity, the zeta-image, and information-theoretic illuminant estimation," in "Computer Vision–ECCV 2012. Workshops and Demonstrations," (2012), pp. 411–420.

[33] R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," J. Opt. Soc. Am. A **21**, 321–334 (2004).

[34] F.-J. Chang, S.-C. Pei, and W.-L. Chao, "Color constancy by chromaticity neutralization," J. Opt. Soc. Am. A **29**, 2217–2225 (2012).

[35] R. Kawakami, J. Takamatsu, and K. Ikeuchi, "Color constancy from blackbody illumination," J. Opt. Soc. Am. A **24**, 1886–1893 (2007).

[36] K. Barnard, "Improvements to gamut mapping colour constancy algorithms," in "European Conference on Computer Vision," (2000), pp. 390–403.

[37] D. A. Forsyth, "A novel algorithm for color constancy," International Journal of Computer Vision **5**, 5–35 (1990).

[38] 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).

[39] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: A simple, unifying framework for color constancy," IEEE Transactions on Pattern Analysis and Machine Intelligence **23**, 1209–1221 (2001).

[40] 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," (2008), pp. 1–8.

[41] D. H. Brainard and W. T. Freeman, "Bayesian color constancy," J. Opt. Soc. Am. A **14**, 1393–1411 (1997).

[42] L. Shi, W. Xiong, and B. Funt, "Illumination estimation via thin-plate spline interpolation," J. Opt. Soc. Am. A **28**, 940–948 (2011).

[43] Y. Weiss and W. T. Freeman, "What makes a good model of natural images?" in "IEEE Conference on Computer Vision and Pattern Recognition," (2007), pp. 1–8.

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

[45] A. Gijsenij, T. Gevers, and M. P. Lucassen, "Perceptual analysis of distance measures for color constancy algorithms," J. Opt. Soc. Am. A **26**, 2243–2256 (2009).

[46] G. D. Finlayson, S. D. Hordley, and P. Morovic, "Colour constancy using the chromagenic constraint," in "IEEE Conference on Computer Vision and Pattern Recognition," (2005), pp. 1079–1086.

[47] C. Fredembach and G. D. Finlayson, "The bright-chromagenic algorithm for illuminant estimation," J. Imaging Sci. Technol. **52**, 040906 (2008).