Efficient Illuminant Estimation for Color Constancy Using Grey Pixels

Kai-Fu Yang Shao-Bing Gao Yong-Jie Li
University of Electronic Science and Technology of China, Chengdu, China
{yang_kf,gao_shaobing}@163.com, liyj@uestc.edu.cn

Abstract

Illuminant estimation is a key step for computational color constancy. Instead of using the grey world or grey edge assumptions, we propose in this paper a novel method for illuminant estimation by using the information of grey pixels detected in a given color-biased image. The underlying hypothesis is that most of the natural images include some detectable pixels that are at least approximately grey, which can be reliably utilized for illuminant estimation. We first validate our assumption through comprehensive statistical evaluation on diverse collection of datasets and then put forward a novel grey pixel detection method based on the illuminant-invariant measure (IIM) in three logarithmic color channels. Then the light source color of a scene can be easily estimated from the detected grey pixels. Experimental results on four benchmark datasets (three recorded under single illuminant and one under multiple illuminants) show that the proposed method outperforms most of the state-of-the-art color constancy approaches with the inherent merit of low computational cost.

1. Introduction

Efficiently removing the color cast triggered by light source, i.e., color constancy, is necessary for color feature extraction in both computer vision systems and biological visual systems[19, 23, 24, 29]. As a typical ill-posed inverse problem [33], the majority of the existing color constancy models are generally realized by first estimating the color of light source from the given color-biased image and then transforming the color-biased image to a canonical image rendered under the white light source through the process of chromatic adaption [30].

One common way to estimate the illuminant from a recorded image primarily relies on certain knowledge or constraints imposed on the distribution of the reflectance that reflects a kind of intrinsic properties of the scene. Typical methods include the Grey-world-based models [23, 9, 37, 17, 44], which usually take the average or maximum response of each (preprocessed) color channel of an RGB image as the estimated illuminant components by assuming that certain kind of reflectance distribution in a scene is achromatic. For example, the Grey-edge [44] assumes that the average reflectance difference in separate color channel of a scene is normally achromatic.

Instead of hypothesizing the reflectance distribution of a scene with an oversimplified statistics (e.g., mean), more sophisticated assumptions about the reflectance distribution of a scene have been further introduced. Bayesian based algorithms consider the reflectance as the random variable of normal distribution [8, 26]. Gamut-mapping based methods assume that the observable reflectance distribution under certain illuminant in real world is limited [2, 29, 18]. Compared to the simple-hypothesis based models, the obvious advantage of these models is that they are able to capture the more detailed structure information of surface reflectance and further improve the performance of illuminant estimation, but most of these models are database-oriented and normally depend on an extensive training phase [30].

Other models attempt to build more universal illuminant estimation algorithms through machine learning techniques [28]. Some of them are based on middle- or high-level information and may have a physical interpretation [45, 7]. In contrast, others are mainly based on empirical observation by using scene descriptors to decide which algorithm is likely to work better for specific scene category (e.g., indoor or outdoor) [43, 5, 6]. On the other hand, some biologically-inspired models have brought the recent advances in visual color constancy closer to the computational level [24, 25, 15, 13, 41, 34].

Another line of illuminant estimation is to exploit certain photometric indicators in images which may encode the information about the illuminant. The accurate illuminant estimation can be obtained once the indicators are recovered. Some studies used the presence of highlight or bright pixels in images to estimate the illuminant [42, 38, 35, 14, 11]. A recent work improves the performance of Grey-Edge (GE) by explicitly measuring the specular edges in images [31]. However, these models generally suffer from the difficulty of retrieving the specular reflection [30] and pixel clipping [21]. In addition, some methods attempt to identify the in-

trinsic grey surfaces in an image, since these surfaces contain strong cues for illuminant estimation. Xiong *et al*. suggests to project the image pixels on so-called LSI space to identify the grey surfaces [47].

Moreover, all of methods mentioned above mainly assume that the illuminant is uniform across the scene. However, this assumption is often violated in real-world scenes. There are some authors trying to extend existing algorithms to multi-illuminant scenes by applying color constancy to image patches rather than the entire image [4, 32, 43]. For example, Beigpour *et al*. formulates multi-illuminant estimation as an energy minimization task within a conditional random field over a set of local illuminant estimates [4].

In this work, we propose a simple color constancy model based on another simple hypothesis. We argue that there are some grey (or approximately grey) pixels widely appearing in natural scenes, which can be utilized to reliably estimate the illuminant. After comprehensively validating this hypothesis, we then develop a simple framework to detect the grey pixels from the color-biased images and then estimate the illuminant using these pixels. In addition, we further extend our method to the situation of multiple illuminants by a simple yet efficient strategy.

The main contributions of this paper can be summarized as several aspects. (1) For the first time we comprehensively validated that almost all natural images contain approximately grey pixels, which can be utilized for efficient illuminant estimation. (2) We proposed an efficient photometric-based strategy for automatically detecting the grey pixels from color-biased images. (3) The experiments reveal the feasibility and effectiveness of the proposed method when facing scenes either with single or multiple illuminants. Taken together, our method could be easily implanted in camera's white-balance pipeline due to its good performance and simple implementation with inherent low computation cost.

2. The Grey-Pixel Based Framework

2.1. Basic Hypothesis and Formulation

We argue that most of the natural images contains some intrinsic grey pixels (or at least very approximate to grey) when under a white light source. Once these grey pixels are efficiently retrieved, they can be used for efficient light source color estimation. However, the challenging problem is how to find the grey pixels efficiently from RGB values of original color-biased images since the changing external illuminant usually alters their color appearance.

Studies on color-based recognition seek for color descriptors independent of the color, intensity, and geometry of scene light source [22, 27]. Note that these descriptors may be roughly attributed as one of a certain category of color constancy algorithms, but they are not aimed at recovering the real color appearance of scenes [30]. Motivated by this, we consider to define a certain illuminant-invariant measure (IIM) to identify the grey pixels in images.

If we define the IIM based on the spatial relationship of pixels within local image patches, they can be used in the both situations of uniform and non-uniform illuminant. For example, the standard deviation of local patches calculated in logarithmic space can be used as an IIM (see more details later). However, the true grey pixels being isolated or within a uniform region will have little chance to be detected, since no reliable local difference can be used for IIM computation. Therefore, in order to accurately find the grey pixels, we need to further define the detectable grey pixels by adding following constraints on IIM

- *(a) The grey pixels should be located in small grey patches with the minimum size of* 3×3 *pixels.*
- *(b) These small grey patches are not uniform, i.e., the contrasts of grey patches are non-zero.*

Based on the *constraint (a)*, we can compute IIM in each color channel of the color-biased images. Because IIM is required to be independent of illuminant, the grey pixels can be detected by searching for the points with equal IIM across three color channels. Meanwhile, *constraint (b)* excludes the points with zero IIM. The mathematic derivation is described as follows.

The captured image values $I^i(x,y)$ with $i \in \{r,g,b\}$ can be normally expressed as the product of the illuminant $C(x,y)$ and surface reflectance $R(x,y)$ [33, 30, 16]

$$I^i(x,y) = C^i(x,y) \cdot R^i(x,y), \quad i \in \{r,g,b\} \tag{1}$$

With logarithmic transform, we have

$$\begin{aligned} I^i_{\log}(x,y) &= \log(I^i(x,y)) \\ &= \log(C^i(x,y)) + \log(R^i(x,y)) \end{aligned} \tag{2}$$

We reasonably assume that the illuminant $C(x,y)$ is uniform within small local patches (at least with a size of 3×3 pixels). Thus, it is obvious that in logarithmic space, any measure defined as the difference between neighboring pixels is independent of illuminant. More generally, we denote the local IIM as $\Delta I_{\log}$, which can be, for example, the standard deviation in a local region (a kind of local contrast). The IIM defined by local standard deviation is a function of $(I^i_{\log}(x_m,y_m) - \bar{I}^i_{\log})$, i.e.,

$$\begin{aligned} \Delta I^i_{\log}(x,y) &= f(I^i_{\log}(x_m,y_m) - \bar{I}^i_{\log}) \\ &= f(R^i_{\log}(x_m,y_m) - \bar{R}^i_{\log}) \end{aligned} \tag{3}$$

where (x_m,y_m) is a pixel within the small patch centered at (x,y). $\bar{I}^i_{\log}$ and $\bar{R}^i_{\log}$ are respectively the mean of $I^i_{\log}$ and $R^i_{\log}$ over the patch. Note that it is a general choice to define the contrast using standard deviation, and other types

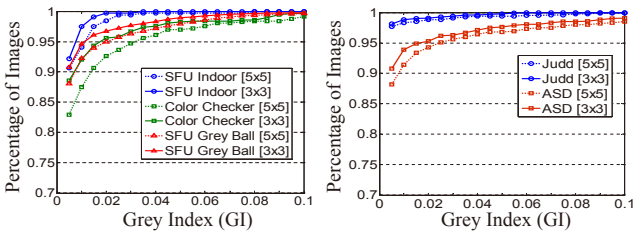

Figure 1. The percentage of images containing detectable grey pixels defined at various Grey Index (GI) level. Left: on the ground-truth images of three benchmark datasets [3, 12, 39]. Right: on two irrelevant natural datasets [1, 36].

of IIMs could be also adopted, such as gradient, center-surround difference, etc. in logarithmic space.

Furthermore, the three components of local $\Delta I_{\log}$ with $i \in \{r, g, b\}$ should be equal to each other for each pixel within a local patch that is grey. In line with *constraint (b)*, we finally identify a pixel as grey when it meets

$$\Delta I^r_{\log}(x,y) = \Delta I^g_{\log}(x,y) = \Delta I^b_{\log}(x,y) \neq 0 \qquad (4)$$

Once we find these grey pixels, we can easily extract the illuminant of a color-biased image from them. However, it is worthy to note that Equation 4 is just a necessary condition for pixel (x, y) being grey. This means that some pixels meeting Equation 4 may be colored. In this work, in order to obtain more robustly the grey pixels, we further exclude the pixels with quite low luminance in original space (RGB space) and relatively isolated in spatial location. More details can be found in Section 2.3.

2.2. Hypothesis Validation

In order to verify whether most of the natural images contain more or less grey pixels, we evaluate the possibility of detectable grey pixels in ground-truth images, which are produced by transforming the raw color-biased images into the canonical images under white light source based on the ground-truth illuminant provided in the benchmark datasets [3, 12, 39]. In addition, we also evaluate our hypothesis on two irrelevant natural datasets collected for other applications, e.g., saliency region detection [1, 36].These datasets include thousands of images captured by different cameras in various natural scenes, presuming that these images have been well white-balanced by cameras themselves.

To measure how close a pixel approximates to grey, we define a *Grey Index (GI)* for each pixel, which is computed as follows. For each pixel (x, y) , we first compute the local mean and local contrast (defined by standard deviation in this paper) within a local window with the size of $\eta \times \eta$ pixels in each channel. Then, we denote the locally averaged maps as $GT^i_\eta(x, y)$ and the local contrast maps as $CT^i_\eta(x, y)$ with $i \in \{r, g, b\}$. Then GI is computed as

$$GI(x,y) = \sqrt{\frac{1}{3}\sum_{i\in\{r,g,b\}} \frac{\left(GT^i_\eta(x,y) - \overline{GT}_\eta(x,y)\right)^2}{\overline{GT}_\eta(x,y)}} \qquad (5)$$

where $\overline{GT}_\eta(x,y) = \frac{1}{3}\left(GT^r_\eta + GT^g_\eta + GT^b_\eta\right)(x,y)$. It is obvious that a pixel with lower GI has higher probability of being grey.

The detectable grey points should have very low GI, while CT^r_η, CT^g_η, and CT^b_η are non-zero. Figure 1 shows the percentage of images containing detectable grey pixels on three benchmark datasets and two irrelevant natural datasets. Note that the gray ball in SFU Gray Ball dataset and the color board in Color-Checker dataset were masked out in this experiments. This figure shows that more than 95% of images in each dataset contain detectable grey pixels with GI lower than 0.02 and almost all images contain detectable grey pixels with GI lower than 0.1. This strongly supports the proposed hypothesis that most natural scenes under white light source indeed contain detectable pixels that are approximately grey.

The next question is whether these grey pixels with low GIs could be utilized as the accurate indicators for illuminant estimation. To answer it, we first estimated the illuminant by summing the detected pixels in a color-biased image, which have low GI values computed from its corresponding ground-truth image. Then, the *angular error* is computed to measure the error between the estimated and the ground-truth illuminants [30]. Figure 2 shows the angular errors evaluated on three benchmark datasets [3, 12, 39] when using the detected grey pixels for illuminant estimation. It is obvious that the median and mean angular errors on the three datasets are smaller than or around 1.0 at various GI levels, which are significantly better than all existing color constancy models as far as our knowledge. This indicates that perfect illuminant estimation can be obtained based on the grey pixels of a color-biased image if these grey pixels could be correctly detected out.

So far, the above analysis has grounded our hypothesis that most natural images always include some detectable grey (or approximately grey) pixels, which can be reliably utilized for accurate illuminant estimation.

2.3. Algorithm Implementation

As analyzed above, for a grey point in a scene under white light source, the local means, as well as the local contrasts, are equal for the three channels. However, for color-biased images, only the constraint of equal local contrasts in three channels is available.

Figure 3 shows the flowchart of the proposed algorithm. For an input color-biased image $I(x, y)$, we first transform the color channels (I^r, I^g, I^b) into logarithmic space (I^r_{log}, I^g_{log}, I^b_{log}). Then the IIM defined by local contrast (standard

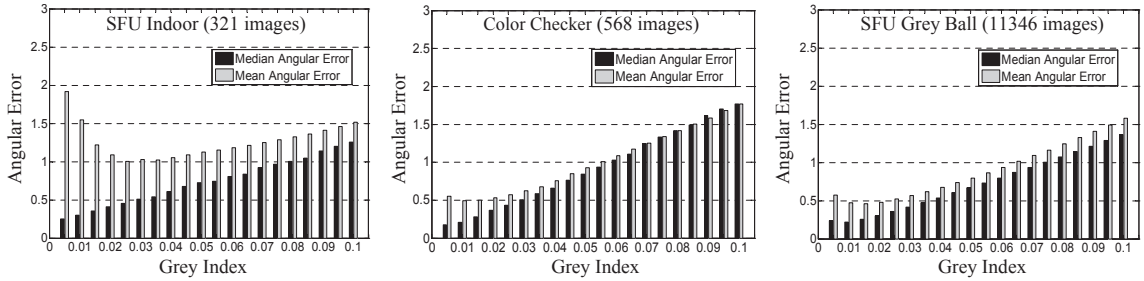

Figure 2. The angular error statistics (median and mean) of illuminant estimation on three benchmark datasets when using the retrieved grey pixels of color-biased images, which are defined by different GI levels in the corresponding ground truth images.

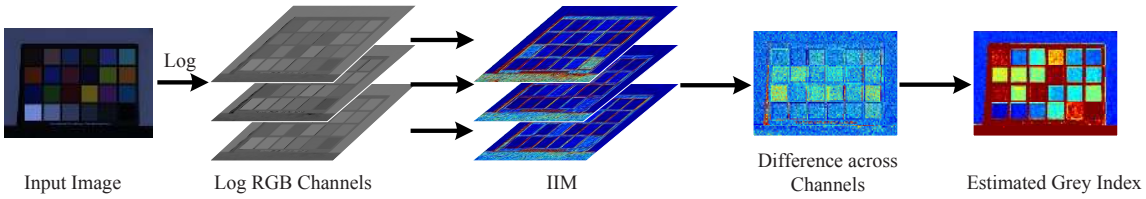

Figure 3. The flowchart of the proposed algorithm.

deviation) within the patch of $\eta \times \eta$ pixels, denoted by SD^r_η, SD^g_η, and SD^b_η, are computed for each pixels. In the work we always set $\eta = 3$. Finally, similar with the definition of GI in Equation 5, the relative standard deviation of local contrasts in three channels are computed as

$$P(x,y) = \sqrt{\frac{1}{3}\sum_{i\in\{r,g,b\}} \frac{\left(SD^i_\eta(x,y) - \overline{SD}_\eta(x,y)\right)^2}{\overline{SD}_\eta(x,y)}} \qquad (6)$$

where $\overline{SD}_\eta(x,y) = \frac{1}{3}\left(SD^r_\eta + SD^g_\eta + SD^b_\eta\right)(x,y)$. Based on the analysis described in Section 2.1, $SD^i_\eta(x,y)$ is an IIM. $P(x,y)$ should be close to 0, if the point (x,y) is approximately grey. We exclude the points with $SD^r_\eta = SD^g_\eta = SD^b_\eta = 0$ based on the *constraint (b)*.

As mentioned before, not all pixels meeting Equation 4 are grey. We further exclude some undesired points that usually are of low luminance (e.g., dark pixels) in original RGB space or are isolated in spatial locations. Therefore, we weaken the influence of these dark or isolated pixels by

$$GI^*(x,y) = AF\left\{\frac{P(x,y)}{L(x,y)}\right\} \qquad (7)$$

where $L = (I^r + I^g + I^b)/3$ denotes the luminance of each pixel in the input image. $AF\{\}$ indicates a averaging filter that is applied within the local neighborhood of 7×7 pixels. Finally, $GI^*(x,y)$ is considered as the final GI map of a given color-biased image. The pixels with lower magnitudes in $GI^*(x,y)$ have higher possibility of being grey.

In Figure 4, the first column shows the input color-biased images; the second column shows the grey index maps $GI(x,y)$ by Equation 5 for the ground-truth images, and the last column shows the grey index maps $GI^*(x,y)$ for

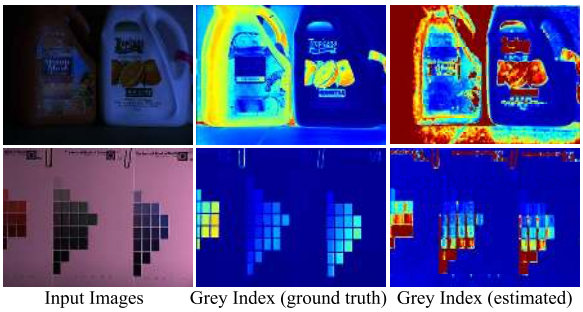

Figure 4. The heat map of grey index: The darker the blue, the greater the probability of being grey pixels.

the color-biased images with Equation 6∼7. From this figure, the grey index maps evaluated for color-biased images are very similar to that computed from the ground-truth images, i.e., the GIs that indicate highly possible grey pixels in ground truth maps (dark blue) are accurately retrieved in the GI maps (dark blue) of the color-biased images.

For an image of the same field of natural scene view, the absolute number of grey pixels varies but its percentage almost keeps steady when the image is zoomed in or out. Therefore, we sort the points in ascending order of GI values in $GI^*(x,y)$, and choose the top $n\%$ pixels with the lowest GIs (denoted by the set GP_n) as the final grey pixels retrieved for illuminant estimation. Then the illuminant components e_i are simply computed as

$$e_i = \frac{1}{N}\sum_{(x,y)\in GP_n} I^i(x,y), \quad i \in \{r,g,b\} \qquad (8)$$

where N is the number of the grey points in GP_n. It is worthy to note that when $n\% = 100\%$, i.e., all pixels in

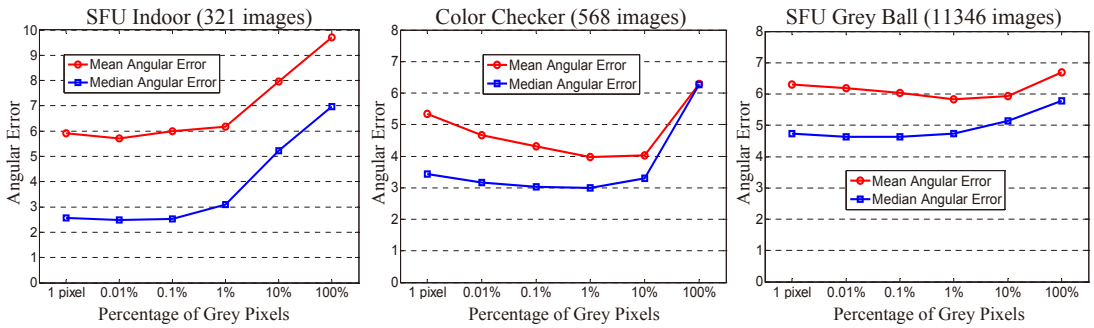

Figure 5. The influence of parameter $n\%$ on the illuminant estimation performance of the proposed method $GP(std)$.

input image are used as grey pixels, our method is reduced to the classical Grey-World.

2.4. Multi-illuminant Estimation

When a scene contains more than one light source, we need to compute the illuminant for each pixel. Our method still works in extracting grey points, because we evaluate the grey index map $GI^*(x,y)$ based on small patches (e.g., 3×3 pixels), in which illuminant can be reasonably considered as being uniform. We first detect the grey pixels with the same steps as the situation of single illuminant (Equation 6∼7), then cluster these retrieved grey pixels into M groups using a simple *K-means* based on their spatial locations. Next, the illuminant for each group ($\xi_k^i,\ i \in \{r,g,b\},\ k = 1,2,..,M$) is computed by Equation 8 using the corresponding grey pixels. Meanwhile, we can obtain the centroids (x_k, y_k) for each group of grey pixels. For any point (x,y), its illuminant can be evaluated by

$$e_i(x,y) = \frac{1}{\mu}\sum_{k=1}^{M} w_k \xi_k^i, \quad i \in \{r,g,b\} \tag{9}$$

$$w_k = \exp\left(-\frac{D_k}{2{\sigma_d}^2}\right), \ k = 1,2,...M \tag{10}$$

where $\mu = \sum_k w_k$ is a normalization coefficient, and $D_k = \|(x,y) - (x_k,y_k))\|/\sqrt{W^2+H^2}$ denotes the normalized spatial distance between pixel (x,y) and the centroid of k^{th} group of grey pixels. W and H denote respectively the width and height of the input image. σ_d controls the weighting sensitivity. In this work, we always set $\sigma_d = 0.2$.

3. The Experiments

We evaluated the proposed method on three benchmark datasets with single illuminant [3, 12, 39] and one dataset with multiple illuminants [4]. The corresponding ground-truth illuminant is also provided for each image.

The proposed approach can be roughly classified as a low-level based method and we focused on the comparison of it with other low-level based methods: Inverse-Intensity Chromaticity Space) (IICS) [42], Grey World (GW) [9], White Patch (WP) [37], Shade of Grey (SoG) [17], General Grey World (GGW) [30], first order Grey Edge (GE1), second order Grey Edge (GE2) [44], Local Surface Reflectance Statistics (LSRS) [23], and Random Sample Consensus (RANSAC) [20]. In addition, the available state-of-the-art learning-based methods are also included: pixel-based Gamut Mapping (GM(pixel)) [18], edge-based Gamut Mapping (GM(edge)) [29], Spatio-Spectral statistics (SS-ML) [10], Weighted Grey Edge (WGE) [31], Regression (SVR) [46], Natural Image Statistics (NIS) [28], Exemplar-based method (Exemplar) [43], Bayesian [26], Thin-plate Spline Interpolation(TPS) [40], and GSI [47]. Finally, a recent biologically inspired method (DO) [24] is also considered. The performances of the most existing methods considered are directly from [29], and the details of which can be download from `http://colorconstancy.com/`, except that the results of SS-ML, GSI, LSRS, TPS, RANSAC, and DO are from Ref [10], [47], [23], [40], [20] and [24], respectively.

3.1. Parameter setting and analysis

In order to demonstrate that other IIMs may also work well for grey pixels detection in our system, we also tested to use the local edge information as a kind of IIM by replacing the standard deviation (SD_η^i) defined local contrast with local gradient in each logarithmic color channel. In the following, we will denote the model with local contrast defined IIM as $GP(std)$, and the model with local gradient defined IIM as $GP(edge)$. Note that in $GP(edge)$, we detect grey pixels by searching the pixels with equal edge response across three logarithmic channels. This is definitely different from the popular Grey-Edge, which assumes that average edge responses in separate color channels are equal for illuminant estimate.

The only one free parameter in our model is $n\%$, which determines the number of detected grey pixels used for illuminant estimation. Figure 5 exhibits the influence of the parameter $n\%$ on the illuminant estimation of $GP(std)$ on three single-illuminant datasets which show that the proposed method is fairly robust for a wide range of $n\%$. For

Figure 6. Results of several algorithms on the selected images of the SFU indoor dataset (the first row) and the color-checker dataset (the second to fourth rows). The angular error is indicated on the lower left corner of each image.

example, the proposed method obtains very good performance on the three datasets when the parameter $n\%$ is within the range of 0.01% to 1.0%. Moreover, it is clear that our method performs better than that when using all pixels as grey points ($n\% = 100\%$), i.e., the results of the Grey-World method. As a whole, the proposed method arrives at best performance with $n\% = 0.01\%$ on these three datasets, which corresponds to about hundreds pixels used as grey pixels. In the following experiments, we always set $n\% = 0.01\%$ for all the three datasets tested here. Note that when only one detected grey pixel is used for estimating illuminant, our method still provides good performance. The reason is that the averaging filter applied in Equation 7 can efficiently remove noises when computing $GI^*(x, y)$.

3.2. Single-illuminant Dataset

Table 1 lists the angular error statistics of various models on two datasets with linear images. The SFU indoor dataset contains 321 linear images captured in laboratory under 11 different lights [3]. The Color-checker dataset includes 568 high dynamic linear natural images including both of indoor and outdoor scenes [39, 26]. The color board in each image of color-checker dataset was masked out during illuminant estimation, as done in many other studies.

It can be seen from Table 1 that the proposed methods perform best among almost all low-level based models. For example, our models outperform the best low-level based method (GE2) on the SFU indoor dataset by increasing the median measure with 15% ($GP(edge)$) and 7% ($GP(std)$). Similar conclusion can be drawn that our methods significantly outperform the best low-level based method (GGW) on the color-checker dataset by increasing the median measure with 11% ($GP(edge)$) and 9% ($GP(std)$). Moreover, our methods almost arrives at the best performance of state-of-the-art learning-based method (e.g., GM(pixel) and SVR) on the SFU indoor dataset.

Dataset (linear)		SFU Indoor		Color Checker	
Method		Median	Mean	Median	Mean
Do nothing (DN)		15.6	17.3	13.6	13.7
Simple hypothesis (low-level)	IICS	8.2	15.5	13.6	13.6
	GW	7.0	9.8	6.3	6.4
	WP	6.5	9.1	5.7	7.5
	SoG	3.7	6.4	4.0	4.9
	GGW	3.3	5.4	3.5	4.7
	GE1	3.2	5.6	4.5	5.3
	GE2	2.7	5.2	4.4	5.1
	LSRS	2.4	5.7	2.6	3.4
	RANSAC	-	-	2.3	3.2
Biological	DO	2.4	4.8	2.6	4.0
Learning -Based	GM(pixel)	2.3	3.7	2.3	4.2
	GM(edge)	2.3	3.9	5.0	6.5
	SS-ML	3.5	5.6	3.0	3.7
	WGE	2.4	5.6	-	-
	SVR	2.2	-	6.7	8.1
	TPS	2.4	-	2.8	-
	GSI	3.9	-	-	-
	Bayesian	-	-	3.5	4.8
	NIS	-	-	3.1	4.2
	Exemplar	-	-	2.3	2.9
Proposed	**GP(std)**	**2.5**	**5.7**	**3.2**	**4.7**
	GP(edge)	**2.3**	**5.3**	**3.1**	**4.6**

Table 1. Performance of various methods on two linear dataset, i.e., SFU indoor (321 images) and Color Checker (568 images).

For the checker-color dataset, our methods ($GP(std)$ and $GP(edge)$) arrives at or beyond the performance of the most learning-based methods (e.g., NIS, SS-ML, Bayesian, GM(edge), SVR). However, compared to the complex implementation of the learning-based GM(pixel) and Exemplar, the proposed models are quite simpler. Testing on this dataset (568 images with the size of about 1500×2000 pixels), our method took mean computation time of about 0.88

Method	Median	Mean	Best-25%	Worst-25%
DN	6.7	8.3	1.0	18.8
IICS	5.6	6.6	1.8	13.3
GW	7.0	7.9	2.2	15.2
WP	5.3	6.8	1.2	14.7
SoG	5.3	6.1	1.8	11.9
GGW	5.3	6.1	1.8	11.9
GE1	4.7	5.9	1.6	11.9
GE2	4.8	6.1	1.6	12.4
LSRS	5.1	6.0	-	11.9
GSI	5.5	-	-	-
TPS	4.6	-	-	-
GM(pixel)	5.8	7.1	1.7	14.7
GM(edge)	5.8	6.8	1.9	13.5
NIS	3.9	5.2	1.2	11.1
Exemplar	3.4	4.4	1.0	9.4
GP(std)	**4.6**	**6.2**	**1.0**	**14.0**
GP(edge)	**4.6**	**6.1**	**1.1**	**13.6**

Table 2. Performance of various methods on non-linear dataset, i.e., SFU Grey Ball dataset (11346 images).

second for $GP(edge)$ and 2.33 second for $GP(std)$ on each image with un-optimized MATLAB codes. The computer used here is Inter i7, 3.5GHZ with 16.0G RAM.

Figure 6 shows the results on example images from the SFU indoor dataset and the color-checker dataset.

We further evaluated our methods on the SFU grey ball dataset [12], which contains 11346 non-linear images. This dataset has been processed with known and unknown post-processing in camera, which makes it quite difficult to obtain accurate illuminant estimate [23]. For unbiased evaluation, we masked the grey ball attached on the lower right corner of each image before the experiments.

Table 2 lists the results on this dataset. Our methods perform best among all low-level based models and also outperform most of the learning-based methods (e.g., GM(edge)). Although NIS and Exemplar perform clearly better than our method, they rely on extensive training phase. Moreover, we notice that most of the learning-based algorithms perform not so well on this dataset as that on the two datasets evaluated above. One possible reason is that the images in this dataset are of very low quality, which cannot provide enough reflectance samples for these complex reflectance hypothesis based methods.

It should be pointed out that if we did not mask out the color-checker patch in each image of the color checker dataset and the grey-ball patch in the SFU grey ball dataset, the median angular errors of our $GP(std)$ on these two datasets could be significantly reduced from 3.2 to 2.5 and from 4.6 to 3.1. This further suggests that our model can achieve perfect illuminant estimation once there are reliable grey (or approximately grey) pixels available in the scenes.

Some authors have suggested that the bright pixels of a scene could be also used for accurate illuminant estimation

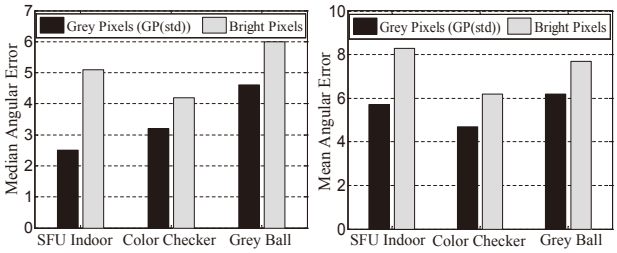

Figure 7. Mean and median error comparisons when our model uses the grey and bright pixels for illuminant estimation.

[35]. Here we also directly took the top 0.01% bright pixels (same number as the grey pixels used above) to estimate illuminant. Figure 7 shows the comparison between the results with grey pixels and with bright pixels on the three datasets. It is clear that the performance based on the grey pixels is better than that based on the bright pixels, which indicates that the good performance of our method benefits mainly from the detected grey points instead of bright pixels. Note that ref. [35] provides better performance by using bright pixels than we re-implemented here, which perhaps results from some careful pre-processing (e.g., clipped pixels). In contrast, we have found that the clipped pixels have little influence on the performance of our model, because the clipped pixels are of quite small amount among the detected grey pixels.

3.3. Multi-illuminant Dataset

Our method was further evaluated on a recent dataset with multiple illuminants [4], which includes 78 high quality images (58 laboratory images and 20 real-world images) captured under two dominant illuminants; the corresponding pixel-wise ground-truth illuminants were also recorded.

Since each image has two dominant illuminants, we set the parameter of cluster number as $M = 2$. Slightly different from the situation with single illuminant, we set $n\% = 10\%$ to involve more grey points, since grey points will be divided into different groups for multiple illuminant estimation. Figure 8 shows examples of non-uniform illuminant evaluation. The estimated illuminant maps are very similar to the ground-truth and the color cast resulting from multiple illuminants is almost completely removed.

Table 3 lists the quantitative comparison. The results of the two recent methods compared here are directly from [4]. Roughly, these two algorithms run the grey-world based algorithms on local patches independently and then combine the resulting estimates. Table 3 indicates that our model obtains clearly better performance than these two methods. For example, our method outperforms the Gijsenij *et al.* [32] by increasing the median measure with around 40% on laboratory images and 14% on real-world images, and achieves slightly better performance than MIRF [4].

Table 3 also reports the results with $M > 2$ (i.e., the

Figure 8. The results of the proposed method on several images with multiple illuminants (the images in the first and second columns are from [4], and in the third and forth columns are from [32]). From first to third rows: the original color-biased images, the pixel-wise ground-truth illuminant, the ground-truth image. The fourth row lists the pixel-wise illuminants estimated by the proposed method and the fifth row lists the corrected images.

Multi-illuminant	Laboratory(58)		Real-world(20)	
Method	Median	Mean	Median	Mean
DN	10.5	10.6	8.8	8.9
Gijsenij *et al.*	4.2	4.8	3.8	4.2
MIRF	2.6	2.6	3.3	4.1
GP(std)$M = 2$)	**2.53**	**3.13**	**3.28**	**5.74**
GP(std)($M = 4$)	**2.27**	**2.94**	**3.44**	**5.59**
GP(std)($M = 6$)	**2.20**	**2.88**	**3.51**	**5.68**
GP(std)($M = 8$)	**2.21**	**2.86**	**3.56**	**5.69**
GP(std)($M = 10$)	**2.21**	**2.85**	**3.53**	**5.70**

Table 3. Performance of various methods on multi-illuminant dataset. Note that the angular errors of our method were averaged over 20 run times, since the output of *K-means* clustering in Section 2.4 may slightly vary with random initialization.

cluster number is larger than the illuminant number), which shows that the performance was further improved when the cluster number was moderately increased. This is because that though the dataset is presumed to contain two dominant illuminants, there are actually some image regions lighted by the mixture of two illuminants.

3.4. The limitation and future work

Based on the hypothesis proposed in Section 2.1, our method may exhibit poor performance on the images that contain quite few detectable grey pixels. Figure 9 lists two failure examples from the SFU indoor dataset. Although images without detectable grey (or approximately grey) pixels are quite singular, especially in natural scenes (as indicated by Figure 1), this issue is still worthy to be re-thought.

Some authors have suggested that color contrast information is also important for color constancy, especially the

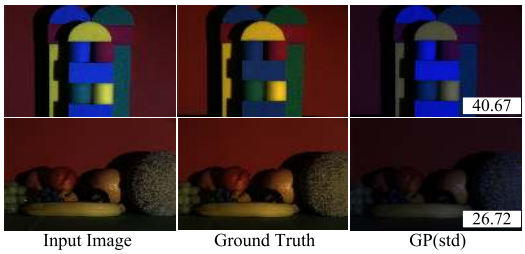

Figure 9. Two failure examples with quite high angular errors.

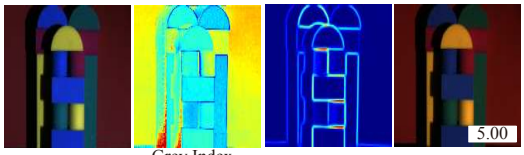

Figure 10. Example of one "blocks1" image (left) without well-marked grey patches. Instead, some grey pixels are present on borders of opponent color surfaces (middle left). When selecting the pixels with strong color edge responses (middle right) as grey pixels, the angular error was reduced remarkably (right).

color contrast between opponent colors [24]. We conducted a simple test on the special "blocks1" images (11 images) from the SFU indoor dataset, an example of which is shown in Figure 10. From the GI map of the ground-truth images, we found that the grey pixels indeed present on the boundaries of different color surfaces. Therefore, we simply extracted the pure color boundaries with the *Color-Opponent* edge detector proposed in [48], and then we selected the pixels with the top 0.01% edge responses for illuminant estimation. The mean angular error of these 11 images was decreased significantly from 35.96 ($GP(std)$) to 9.24. This suggests that integrating more low-level features would improve the ability of our method to deal with special or complex scenes, which is a future direction for our study.

4. Conclusion

We developed a new framework for illuminant estimation by retrieving grey (or approximately grey) pixels (GP) from color-biased images based on the photometric property of images. The comprehensive experiments on multiple collections of diverse datasets indicate that most of the natural images in real world contain approximately grey pixels (defined by IIM), which could be utilized for efficient illuminant estimation with very simple implementation. Extensive evaluations on benchmark datasets indicate that the proposed method with grey pixels can obviously outperform the models with the grey-world and grey-edge information. Our model can also obtain quite competitive results in comparison to the state-of-the-art models, in the both situations of uniform and non-uniform illuminant.

Acknowledgements

This work was supported by the 973 Project under Grant 2013CB329401, the NSFC under Grant 61375115, 91420105, and the Doctoral Support Program of UESTC.

References

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1597–1604. IEEE, 2009.

[2] K. Barnard. Improvements to gamut mapping colour constancy algorithms. In *European Conference on Computer Vision*, pages 390–403. Springer, 2000.

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

[4] S. Beigpour, C. Riess, J. van de Weijer, and E. Angelopoulou. Multi-illuminant estimation with conditional random fields. *Image Processing, IEEE Transactions on*, 23(1):83–96, 2013.

[5] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini. Improving color constancy using indoor–outdoor image classification. *Image Processing, IEEE Transactions on*, 17(12):2381–2392, 2008.

[6] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini. Automatic color constancy algorithm selection and combination. *Pattern recognition*, 43(3):695–705, 2010.

[7] S. Bianco and R. Schettini. Adaptive color constancy using faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(8):1505–1518, 2014.

[8] D. H. Brainard and W. T. Freeman. Bayesian color constancy. *Journal of Optical Society of America A: Optics, Image Science, and Vision*, 14(7):1393–1411, 1997.

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

[10] A. Chakrabarti, K. Hirakawa, and T. Zickler. Color constancy with spatio-spectral statistics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(8):1509–1519, 2012.

[11] 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. *Journal of Optical Society of America A: Optics, Image Science, and Vision*, 31(5):1049–1058, 2014.

[12] F. Ciurea and B. Funt. A large image database for color constancy research. In *Color and Imaging Conference*, volume 2003, pages 160–164. Society for Imaging Science and Technology, 2003.

[13] S. M. Courtney, L. H. Finkel, and G. Buchsbaum. A multistage neural network for color constancy and color induction. *Neural Networks, IEEE Transactions on*, 6(4):972–985, 1995.

[14] M. S. Drew, H. R. V. Joze, and G. D. Finlayson. The zeta-image, illuminant estimation, and specularity manipulation. *Computer Vision and Image Understanding*, 127:1–13, 2014.

[15] P. A. Dufort and C. J. Lumsden. Color categorization and color constancy in a neural network model of v4. *Biological Cybernetics*, 65(4):293–303, 1991.

[16] G. D. Finlayson, M. S. Drew, and B. V. Funt. Color constancy: generalized diagonal transforms suffice. *Journal of Optical Society of America A: Optics, Image Science, and Vision*, 11(11):3011–3019, 1994.

[17] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41. Society for Imaging Science and Technology, 2004.

[18] D. A. Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–35, 1990.

[19] D. H. Foster. Color constancy. *Vision research*, 51(7):674–700, 2011.

[20] B. Funt and M. Mosny. Removing outliers in illumination estimation. In *Color and Imaging Conference*, volume 2012, pages 105–110. Society for Imaging Science and Technology, 2012.

[21] B. Funt and L. Shi. The rehabilitation of maxRGB. In *Color and Imaging Conference*, volume 2010, pages 256–259. Society for Imaging Science and Technology, 2010.

[22] B. V. Funt and G. D. Finlayson. Color constant color indexing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(5):522–529, 1995.

[23] S. Gao, W. Han, K. Yang, C. Li, and Y. Li. Efficient color constancy with local surface reflectance statistics. In *European Conference on Computer Vision*, pages 158–173. Springer, 2014.

[24] S. Gao, K. Yang, C. Li, and Y. Li. A color constancy model with double-opponency mechanisms. In *IEEE International Conference on Computer Vision*, pages 929–936. IEEE, 2013.

[25] S.-B. Gao, K.-F. Yang, C.-Y. Li, and Y.-J. Li. Color constancy using double-opponency. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, in press, 2015.

[26] 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*, pages 1–8. IEEE, 2008.

[27] T. Gevers and A. W. Smeulders. Color-based object recognition. *Pattern recognition*, 32(3):453–464, 1999.

[28] A. Gijsenij and T. Gevers. Color constancy using natural image statistics and scene semantics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(4):687–698, 2011.

[29] A. Gijsenij, T. Gevers, and J. Van De Weijer. Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision*, 86(2-3):127–139, 2010.

[30] A. Gijsenij, T. Gevers, and J. Van De Weijer. Computational color constancy: Survey and experiments. *Image Processing, IEEE Transactions on*, 20(9):2475–2489, 2011.

[31] A. Gijsenij, T. Gevers, and J. Van De Weijer. Improving color constancy by photometric edge weighting. *Pattern Analysis*

and Machine Intelligence, IEEE Transactions on, 34(5):918–929, 2012.

[32] A. Gijsenij, R. Lu, and T. Gevers. Color constancy for multiple light sources. *Image Processing, IEEE Transactions on*, 21(2):697–707, 2012.

[33] S. D. Hordley. Scene illuminant estimation: past, present, and future. *Color Research & Application*, 31(4):303–314, 2006.

[34] C.-H. Huang and C.-T. Lin. Bio-inspired computer fovea model based on hexagonal-type cellular neural network. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 54(1):35–47, 2007.

[35] 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*, volume 2012, pages 41–46. Society for Imaging Science and Technology, 2012.

[36] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *IEEE International Conference on Computer Vision*, pages 2106–2113. IEEE, 2009.

[37] E. H. Land and J. McCann. Lightness and retinex theory. *Journal of Optical Society of America*, 61(1):1–11, 1971.

[38] H.-C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of Optical Society of America A: Optics, Image Science, and Vision*, 3(10):1694–1699, 1986.

[39] 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/`.

[40] L. Shi, W. Xiong, and B. Funt. Illumination estimation via thin-plate spline interpolation. *Journal of Optical Society of America A: Optics, Image Science, and Vision*, 28(5):940–948, 2011.

[41] H. Spitzer and S. Semo. Color constancy: a biological model and its application for still and video images. *Pattern Recognition*, 35(8):1645–1659, 2002.

[42] R. T. Tan, K. Nishino, and K. Ikeuchi. Color constancy through inverse-intensity chromaticity space. *Journal of Optical Society of America A: Optics, Image Science, and Vision*, 21(3):321–334, 2004.

[43] H. Vaezi Joze and M. Drew. Exemplar-based colour constancy and multiple illumination. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(5):860–873, 2014.

[44] J. Van De Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *Image Processing, IEEE Transactions on*, 16(9):2207–2214, 2007.

[45] J. Van De Weijer, C. Schmid, and J. Verbeek. Using high-level visual information for color constancy. In *IEEE International Conference on Computer Vision*, pages 1–8. IEEE, 2007.

[46] W. Xiong and B. Funt. Estimating illumination chromaticity via support vector regression. *Journal of Imaging Science and Technology*, 50(4):341–348, 2006.

[47] W. Xiong, B. Funt, L. Shi, S.-S. Kim, B.-H. Kang, S.-D. Lee, and C.-Y. Kim. Automatic white balancing via gray surface identification. In *Color and Imaging Conference*, volume 2007, pages 143–146. Society for Imaging Science and Technology, 2007.

[48] K. Yang, S. Gao, C. Li, and Y. Li. Efficient color boundary detection with color-opponent mechanisms. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2810–2817. IEEE, 2013.