A Color Constancy Model with Double-Opponency Mechanisms

Shaobing Gao[1], Kaifu Yang[1], Chaoyi Li[1,2], Yongjie Li[1]

[1]University of Electronic Science and Technology of China, Chengdu, China

[2]Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

{gao_shaobing, yang_kf}@163.com, cyli@sibs.ac.cn, liyj@uestc.edu.cn

Abstract

The double-opponent color-sensitive cells in the primary visual cortex (V1) of the human visual system (HVS) have long been recognized as the physiological basis of color constancy. We introduce a new color constancy model by imitating the functional properties of the HVS from the retina to the double-opponent cells in V1. The idea behind the model originates from the observation that the color distribution of the responses of double-opponent cells to the input color-biased images coincides well with the light source direction. Then the true illuminant color of a scene is easily estimated by searching for the maxima of the separate RGB channels of the responses of double-opponent cells in the RGB space. Our systematical experimental evaluations on two commonly used image datasets show that the proposed model can produce competitive results in comparison to the complex state-of-the-art approaches, but with a simple implementation and without the need for training.

1. Introduction

Color constancy is one of the amazing abilities of perceptual constancy of the human visual system, which enables the perceived color of objects largely constant as the light source color changes [7]. In contrast, captured with regular digital cameras or videos, the physical color of scenes may be shifted by the varying external illuminant. Figure 1 shows the shift of color distribution between the canonical and color-biased images of the same scene. One of the important problems in computer vision, especially for the robust color-based systems, is to extract reliable color features that are invariant to the changes in external lighting. A general solution is to estimate the scene illuminant and then remove the illuminant from color-biased images to get the so-called canonical images under a white light source [17].

Based on the two steps mentioned above, many methods exist. For example, the well-known grey-world theory [4] assumes that the average reflectance of a scene is nor-
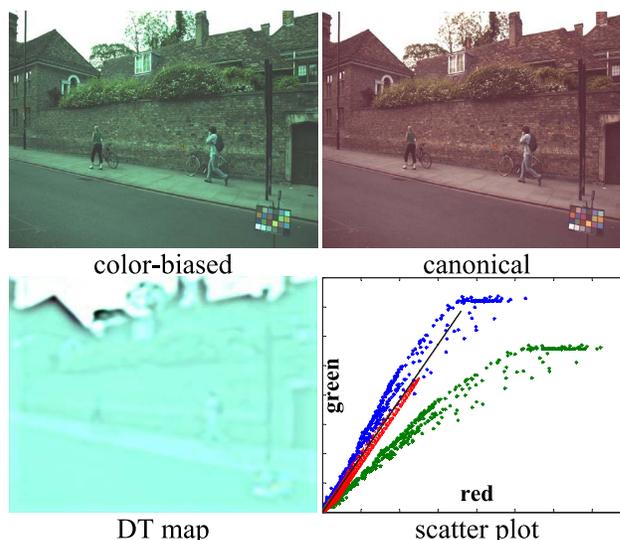

Figure 1. The blue and green crosses in the scatter plot denote the color distributions of the color-biased and canonical images[24], respectively. The red crosses in the scatter plot show the color distribution of the responses of the double-opponent cells in our model (i.e., the DT map, in the RGB space). The true illuminant is shown as a black solid line.

mally achromatic, based on which the illuminant components could be estimated simply by computing the mean in each color channel (i.e., R, G, and B) of the input image. Some other typical grey-world assumption based methods include white patch and max-RGB [19], shade of grey [10], grey-edge [27], etc. However, the reflectance distribution of a natural scene does not always perfectly satisfy the grey-world assumption. For example, the average reflectance is not achromatic in an image containing large uniformly colored surfaces, for which both the grey-world and grey-edge models would fail in accurately estimating the illuminant.

In order to reinforce the simple assumption about the reflectance distribution, a group of learning-based color constancy models introduce priori information about illuminant and employ statistical computation to estimate the illuminant. Typical examples include the gamut mapping [11, 16],
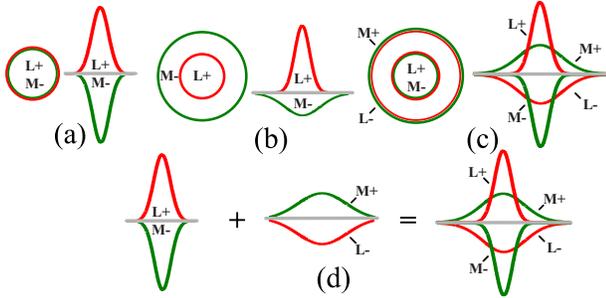
Figure 2. The receptive field (RF) of red-green single-opponent Type II cells in LGN with color-opponent centre-only RF (a) and Type I cells with color-opponent centre-surround RF (b). The R-F of red-green double-opponent cells in the primary visual cortex (V1) (c) can be computationally constructed using two single-opponent Type II cells with different RF scales and opposite signs (d). In the expression of "A+" or "B-", the sign "+" and "-" denote the excitation and inhibition, respectively. Adapted from [7].

Bayesian color constancy [14], regression based [12], statistics based [15, 6], and high-level visual information based [28, 3]. However, most learning-based models need to be correctly trained before the illuminant can be estimated, which is a crucial distinction that may partially determine the suitability of a color constancy model for applicability to real-world systems [17].

Another line of research, not mentioned too much in this paper, is based on the dichromatic reflection model of image formation [26, 20]. These physics-based methods assume non-Lambertian reflectance and explicitly use the presence of specular highlights, and the estimate of the illuminant color can be simply obtained as a solution to a set of equations. However, they generally suffer from the difficulty of retrieving the specular reflections [17].

In this paper, we propose a visual system based color constancy model, and the idea behind the model originates from the computational discoveries that the responses of double-opponent (DO) cells to the color-biased images provide clear information about the scene illuminant. From Figure 1 we can find that the color distribution of the responses of DO cells to a color-biased image coincides well with the direction of light source color (see the supplemental materials for more examples[1]). Based on this substantial observation, we directly use the MAX mechanism to extract the true illuminant from the responses of a group of DO cells. Evaluation of our model on two typical datasets commonly used in the field of computational color constancy demonstrate our model's competitive results to the state-of-the-art approaches.

The rest of this paper is organized as follows. We begin with an overview of the color processing mechanisms in the visual system in Section 2. Then, in Section 3, we describe

in details our double-opponency based model. In Section 4 the proposed model was experimentally verified with two image datasets. Finally, in Section 5, some concluding remarks and future directions are presented.

2. Color processing in the visual system

Color processing in the visual system progresses through a series of hierarchical steps [7]: after the light absorption by cone photoreceptors, cone activities are compared by cone-opponent retinal ganglion cells; these color signals are transmitted via the lateral geniculate nucleus (LGN) to the primary visual cortex (V1) and then higher cortical areas. In the following we give a summary to the color processing mechanisms in the early stages of the visual system.

Cone photoreceptor layer of retina: The first stage of color processing of our visual system takes place in the photoreceptor layer of the retina. There are two types of photoreceptors: rods and cones, and cones are responsible for color vision. Based on the spectral sensitivities, cones can be classified into short-wavelength cone (S-cone), medium-wavelength cone (M-cone) and long-wavelength cone (L-cone), which respond preferably to the blue (B), green (G) and red (R) colors, respectively. The color information into the eyes is first coded in a trichromatic way via L-, M- and S-cones in the retina, and then propagated in the way of color opponency via single-opponent and double-opponent neurons at the levels of retinal ganglion layer, LGN and V1 (and the higher cortical areas).

Single-opponent cells: Most color-sensitive cells at the levels of ganglion layer and LGN are single-opponent cells that code the color information within their receptive fields (RFs) in the ways of red-green, blue-yellow, and black-white opponency. Figure 2(a) and (b) show the RF structures of the red-on/green-off single-opponent cells of type II and I, respectively. Type II cells (Figure 2(a)) have center-only color opponent RF and respond well to uniformly colored areas [23]. In contrast, type I cells (Figure 2(b)) have center-surround RF structure with color opponency and respond well to color contrast [23].

Double-opponent cells: Many experiments [21, 25, 13] have revealed that the double-opponent (DO) cells exist widely in V1. Based on the spatial structure of RF, DO cells could be classified into two types: DO cells with concentric RFs (Figure 2(c)) and DO cells with oriented RFs (not shown here). In particular, the RFs of the first type (Figure 2(c)) have concentrically organized center-surround structure and are both spectrally and spatially opponent. These properties of such type of DO cells make them the reasonable physiological building blocks of color constancy as well as the phenomenon of color contrast [7, 13]. In this study we use the DO cells with concentric RFs to build color constancy model. It has been physiologically found that the majority of DO cells, of the both types, receive unbalanced

cone input [23]. Hence, the DO cells used in this work have unbalanced center-surround RF structure, and DO cells with such RF property may respond to both the spatial color contrasts (i.e., color edges) and the color regions [7].

3. Double-opponency based color constancy

Figure 3 shows the flowchart of our double-opponency based color constancy model. The different layers in this bottom-up hierarchical framework mainly correspond to the color processing strategies involved in the HVS from the retina to V1. The key idea behind our model is to estimate the illuminant from the responses of the double-opponent cells.

Cone layer: The input color image is separated into three channels: red (R), green (G), blue (B), which are denoted by $r(x, y)$, $g(x, y)$, and $b(x, y)$ and sent into the L-, M-, and S-cones, respectively. Also, an extra yellow (Y) channel, given by $y(x, y) = (r(x, y) + g(x, y))/2$, is constructed for the computation of blue-yellow (B-Y) opponency. In addition, a luminance (i.e., R+B+G) channel, denoted by $l(x, y) = r(x, y) + g(x, y) + b(x, y)$, is constructed for the computation of black-white opponency of luminance.

Retinal ganglion/LGN layer: The retinal ganglion cells receive the outputs from the cones via horizontal cells and bipolar cells, and then send signals to LGN. Generally, ganglion cells and LGN cells have similar RF properties. Here we implement the processing of ganglion layer and LGN into a single step for simplicity. In this study we only consider the type II single-opponent (SO) cells (Figure 2(a)), which are used to computationally construct the RFs of double-opponent (DO) cells (Figure 2(c) and (d)). The signals in this layer are first transformed from the RGB space to the single-opponent space according to [9]

$$
\begin{aligned}
O_{rg}(x, y) &= (r(x, y) - g(x, y))/\sqrt{2} \\
O_{gr}(x, y) &= -O_{rg}(x, y) \\
O_{yb}(x, y) &= (y(x, y) - 2b(x, y))/\sqrt{6} \\
O_{by}(x, y) &= -O_{yb}(x, y) \\
O_{l_+}(x, y) &= l(x, y)/\sqrt{3} \\
O_{l_-}(x, y) &= -O_{l_+}(x, y)
\end{aligned} \quad (1)
$$

The RF spatial structure of each component of a type II single-opponent (SO) cell could be described using a two-dimensional (2D) Gaussian function [7] written as

$$
RF(x, y; \sigma) = \frac{1}{2\pi\sigma^2} exp(-\frac{x^2 + y^2}{2\sigma^2}) \quad (2)
$$

where the standard deviate σ controls the scale (i.e., the size) of RF. Taking a SO cell of type II (Figure 2(a)) with red-on/green-off (R+G-, or L+M-) opponency as example, its response is computed as

$$
SO_{r+g-}(x, y; \sigma) = O_{rg}(x, y) * RF(x, y; \sigma) \quad (3)
$$

where $*$ denotes the convolution. Similarly, we can compute $SO_{g+r-}(x, y)$ for M+L- (or G+R-) SO cells, and $SO_{b+y-}(x, y)$ and $SO_{y+b-}(x, y)$ for the blue-yellow opponency, and $SO_{l+}(x, y)$ and $SO_{l-}(x, y)$ for the response of brightness-sensitive cells. Note that in the expression of "A+B-", the sign "+" and "-" denote the excitation and inhibition, respectively.

V1 layer: It is not yet fully understood how LGN projects to V1 to form the concentrically organized receptive fields (RFs) of DO cells that are both chromatically and spatially opponent. Here, we adopt the viewpoint of physiological experiments [21], i.e., the RF of a DO cell of L+M-/M+L- shown in Figure 2(c) could be constructed using the outputs from two SO cells of type II with different scales: one red-on/green-off SO cell with smaller RF scale and another green-on/red-off SO cell with larger RF scale (Figure 2(d)). Thus, the response of a DO cell can be computed as

$$
\begin{aligned}
DO_{rg}(x, y) &= SO_{r+g-}(x, y; \sigma) + k \cdot SO_{g+r-}(x, y; \lambda\sigma) \\
DO_{by}(x, y) &= SO_{b+y-}(x, y; \sigma) + k \cdot SO_{y+b-}(x, y; \lambda\sigma) \\
DO_{l}(x, y) &= SO_{l+}(x, y; \sigma) + k \cdot SO_{l-}(x, y; \lambda\sigma)
\end{aligned} \quad (4)
$$

where σ and $\lambda\sigma$ define respectively the scales of the RF center and its surround of a DO cell. We set $\lambda = 3$ based on the physiological finding that the size of receptive field (RF) surround is roughly 3 times (in diameter) larger than that of RF center [22]. k is a relative cone weight that controls the contribution of RF surround. $k \neq 1$ implies that DO cells receive unbalanced cone inputs, and hence, have unbalanced center-surround structures [23].

Higher visual cortex: It is as yet unknown which level of the human visual system finally realizes color constancy. What is known, however, is that color constant cells have been found at the level of visual area V4 [7, 13]. The cells of V4 normally have a very large receptive field, which may endow the V4 cells with ability to extract light source color based on global statistics. Along this line, we first transform the output of DO cells from the double-opponent space to the RGB space according to [9]

$$
\begin{pmatrix} DT_r(x, y) \\ DT_g(x, y) \\ DT_b(x, y) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}^{-1} \begin{pmatrix} DO_{rg}(x, y) \\ DO_{by}(x, y) \\ DO_{l}(x, y) \end{pmatrix} \quad (5)
$$

Note that the transformation of DO space to RGB by Equation 5 is just for the comparison of the estimated and real light source colors in RGB space.

Taking a color-biased image as shown in Figure 1 as input, we compute the output of DO cells using Equations (1)~(5). We can find from the scatter plot of Figure 1 that the distribution of the responses of DO cells (i.e., the DT maps given by Equation(5)) fits closely to the true illuminant, which holds true for the most of color-biased images in the two datasets [24, 1] used in this study (see the
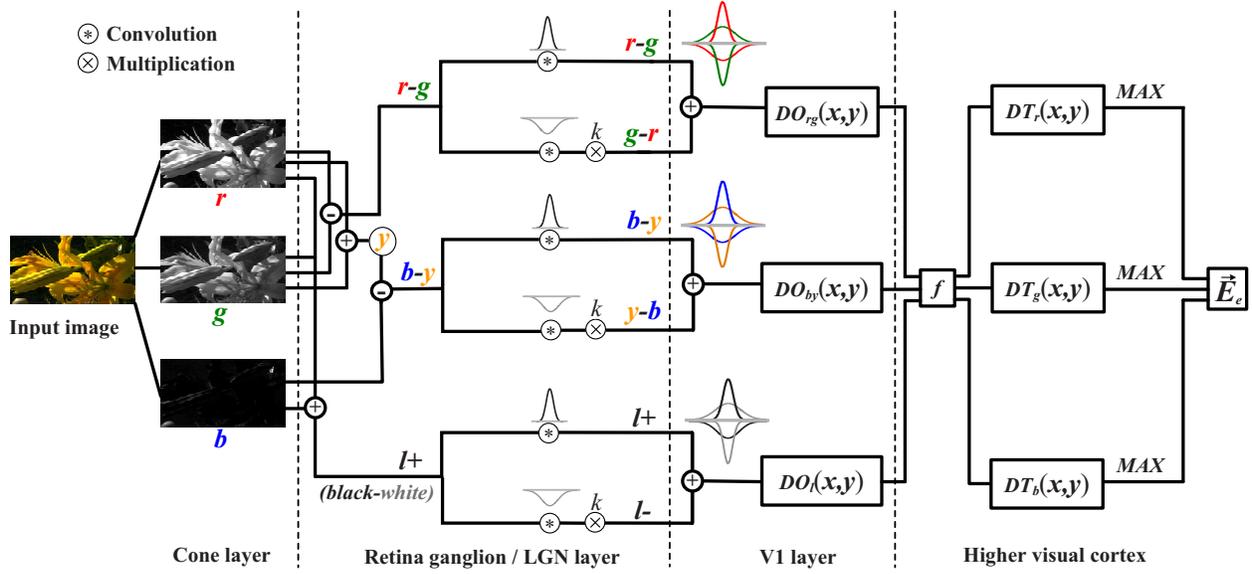
Figure 3. The flowchart of our double-opponency based color constancy model. The symbol f indicates the linear transform defined by Equation (5). r, g, b, y, and l denote respectively the components of the red, green, blue, yellow and luminance channels. *MAX* denotes the mechanism of MAX.

supplemental materials for more examples). Based on this substantial phenomenon, we speculate that V4 cells might adopt certain mechanisms to compute an accurate illuminant estimate of the scene using the color distribution of the DO cells' output in the RGB space.

We assume that the scene is illuminated by a single light source, and the color of the illuminant, $\vec{E}_e = (e_r, e_g, e_b)$, is estimated using a canonical neural computation mechanism of MAX [5], according to

$$e_i = \max_{(x,y)}\big(DT_i(x,y)\big)/coef, \quad i \in \{r,g,b\}$$
$$coef = \sum_{i \in \{r,g,b\}} \max_{(x,y)}\big(DT_i(x,y)\big) \tag{6}$$

4. Experimental results

The proposed model was compared with multiple methods on two typical datasets, i.e., Gehler-Shi dataset [24] and SFU lab dataset [1]. The existing methods considered for comparison in this study are classified into three types according to [17], including: (1) dichromatic reflection model based: inverse-intensity chromaticity space (IICS) [26]; (2) low-level statistics based: Grey World (GW) [4], White Patch (WP) [19], 2nd-Grey Edge (GE2) [27], Shades of Grey (SG) [10], general Grey World (GG); (3) learning based: Bayesian [14], Regression (SVR) [12], automatic color constancy algorithm selection (CART-AAS) [2], using natural image statistics (CCNIS) [15], spatio-spectral statistics (SS) [6], pixel-based gamut mapping (GM(pixel)) [11], edge-based gamut mapping (GM(edge)) [16]. Roughly, the proposed model could be classified into the low-level

statistics based ones. The frequently used angular error ε is chosen as the error metric:

$$\varepsilon = cos^{-1}\Big(\big(\vec{E}_e \cdot \vec{E}_t\big)/\big(\|\vec{E}_e\| \cdot \|\vec{E}_t\|\big)\Big) \tag{7}$$

where $\vec{E}_e$ and $\vec{E}_t$ are the estimated and the true light source colors, respectively, and $\vec{E}_e \cdot \vec{E}_t$ is their dot product. $\|\cdot\|$ denotes the Euclidean norm. Besides the primarily used median angular error, we also reported the measures of mean, trimean, worst-25%, best-25%, and maximum angular er-
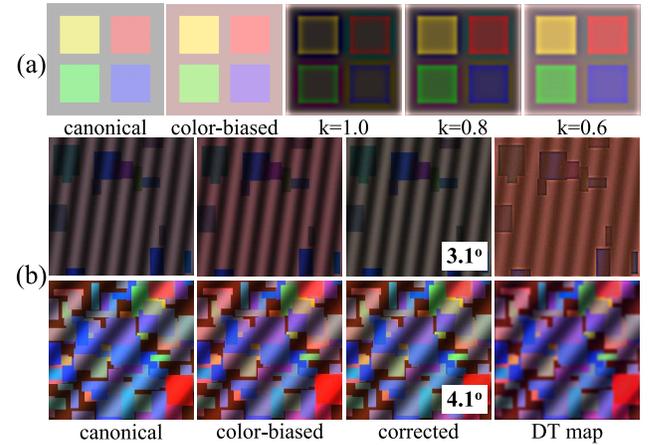

Figure 4. (a) A synthetic image rendered under a reddish light and the DT maps (given by Equation (5)) with different k values. (b) Mondrian images with less (first row) and more (second row) color blocks. The DT maps and the images corrected with our model are also shown (the angular error is shown in the bottom right corner). $k = 0.9$ and $k = 0.3$ are used for the DT map computation of the first and second rows. The mondrian images are generated using the surface reflectance spectra combined with illuminant spectra from [1].

932

rors for more comprehensive comparison.

The proposed model has two free parameters, i.e., the scale of receptive field (σ) (in Equations (2)$\sim$(4)) and the cone weight (k) (in Equation (4)). Figure 5 shows the influences of different parameter values on the model's performance in terms of median angular error, according to which we set $\sigma = 2.5$ for the both datasets used in this work, and set $k = 0.2$ for the Gehler-Shi dataset and $k = 0.9$ for the SFU lab dataset.

Figure 4(a) shows that DO cells with balanced cone inputs (i.e., $k = 1$) only respond to the color contrast, and in contrast, DO cells with unbalanced cone inputs (i.e., $k \neq 1$) only respond to both the color contrast and color regions [7], and with a higher k, the color edges are enhanced much more than the color regions. Figure 4(b) indicates that if fewer objects and hence fewer edges are available in an image, enhancing edges would contribute much more than enhancing color regions to the illuminant estimation. This holds true for the parameter setting of k for the two image datasets mentioned above, i.e., compared to the Gehler-Shi real-world dataset (with $k = 0.2$), the images of the SFU lab (indoor) dataset normally contain fewer edges [15, 6], and hence, our model with a higher k ($k = 0.9$) could provide a better result, as demonstrated below.

4.1. Real-World Image Set

Gehler-Shi dataset [24] contains 568 high dynamic range linear images, including a variety of indoor and outdoor scenes, captured using a high-quality digital SLR camera in RAW format and therefore free of any color correction. In this study, the color-checker patch in each image used for computing ground truth illuminant was masked out in order to fully evaluate the performance of a specific model. The results of multiple methods are listed in Table 1, which reports the measures for the entire dataset, as well as separately for the indoor and outdoor sets. It can be seen from Table 1 that the performance of the proposed method almost arrives at or beyond the best performance (in terms of median and mean angular errors) of the state-of-the-art learning-based algorithms (e.g., the SS, CCNIS, GM(pixel)) on the indoor, outdoor and entire datasets. Although the median angular error of the best algorithm (GM(pixel)) is slightly lower than our method, the robustness (indicated by the measure of worst-25%) of our method is much better than that of GM(pixel). Though gamut-based models have potentially high accuracy, they require appropriate preprocessing and learning from training data [17]. In addition, compared to the complex implementation of gamut-based models, our model is quite simpler. Figure 7 shows the median angular errors over all 568 images of the dataset for different methods, the error bars in which indicate a confidence interval of 95% [15]. This figure further shows the good performance of our method.

Figure 6 show examples of both indoor and outdoor images corrected with the illuminant estimates of various methods. From Figure 6 we can find that the performances of several learning-based algorithms, e.g., Bayesian, spatio-spectral (SS), GM(pixel), are very bad on, especially on the two indoor images, and in contrast, the quite simpler algorithms (e.g., Grey World (GW), Grey-Edge (GE2) and the proposed) achieve good performance. The reason may be that the learning-based algorithms are seriously dependent on the illuminant priori used for training. However, both the two indoor images in Figure 6 contain self-emitting light source (e.g., the LCD and halogen lamp), which may not satisfy the illuminant priori assumed by these methods.

Figure 6 also indicates that both the learning-based and the proposed algorithms perform well on outdoor scenes with more uniformly colored large surfaces or with fewer edges. However, the grey-world based methods (e.g., Grey World, Grey Edge) obtain poor performance on these images, where the mean of reflectance is obviously not achromatic. More performance analysis is provided in the supplemental materials.

4.2. Laboratory Image Set

We further test our model on the SFU lab dataset [1], which contains 321 available images of 31 different objects
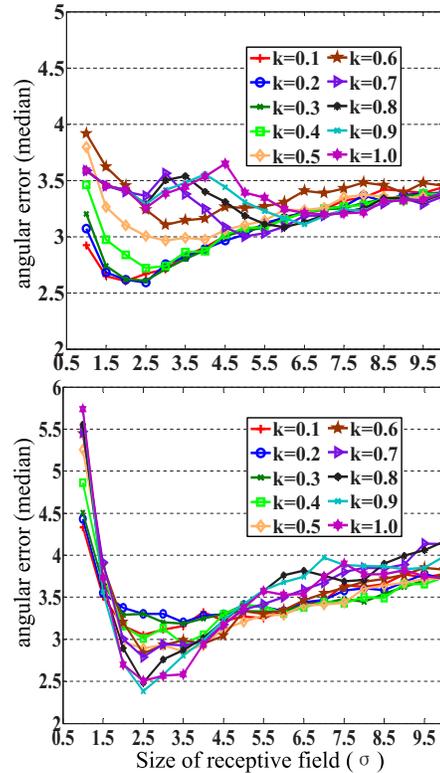

Figure 5. The influence of receptive field size (σ) and the cone weight (k) of our model on the measure of median angular error. **Top**: Gehler-Shi dataset, **Bottom**: SFU lab database. In this study we set $k = 0.2$ for the top and $k = 0.9$ for the bottom.

Methods		All images (568)			Indoor (246)			Outdoor (322)		
		Median	Mean	Worst-25%	Median	Mean	Worst-25%	Median	Mean	Worst-25%
Do nothing (DN)		13.55°	13.65°	17.15°	13.37°	13.85°	19.02°	13.61°	13.50°	15.54°
Physics-based	IICS	13.56°	13.61°	17.95°	13.09°	13.35°	19.45°	13.68°	13.81°	16.68°
(Static) low-level statistics-based	GW	6.28°	6.35°	10.58°	5.67°	5.85°	10.42°	6.60°	6.73°	10.63°
	WP	5.68°	7.54°	16.12°	8.00°	8.78°	17.05°	4.22°	6.60°	15.25°
	GE2	4.50°	5.12°	9.25°	4.64°	5.07°	8.96°	4.20°	5.16°	9.41°
	SG	4.00°	4.92°	10.19°	4.59°	5.39°	10.69°	3.52°	4.56°	9.70°
	GG	3.45°	4.66°	10.18°	3.97°	5.11°	10.70°	3.14°	4.32°	9.62°
Learning-based	SVR	6.72°	8.08°	14.89°	7.84°	9.26°	16.79°	6.08°	7.18°	12.55°
	Bayesian	3.46°	4.82°	10.48°	5.93°	6.52°	11.91°	2.44°	3.52°	7.81°
	CART-AAS	3.34°	4.49°	10.10°	4.26°	5.23°	10.86°	2.73°	3.92°	9.15°
	CCNIS	3.13°	4.19°	9.21°	3.86°	4.83°	9.60°	2.76°	3.70°	8.50°
	SS	3.09°	3.99°	8.47°	4.07°	5.06°	10.00°	2.57°	3.17°	6.48°
	GM(pixel)	2.44°	4.20°	11.15°	4.43°	5.61°	12.57°	1.61°	3.12°	8.87°
	GM(edge)	5.60°	6.71°	13.46°	7.34°	7.92°	14.45°	4.34°	5.79°	11.90°
Proposed		2.60°	4.03°	9.35°	3.95°	5.01°	10.86°	2.06°	3.13°	7.46°

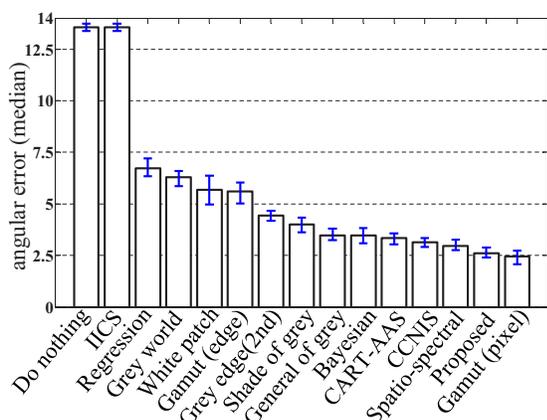Table 1. Performance of various methods on the Gehler-Shi dataset.

Figure 7. Median angular errors for different algorithms on Gehler-Shi dataset plotted with a 95% confidence interval.
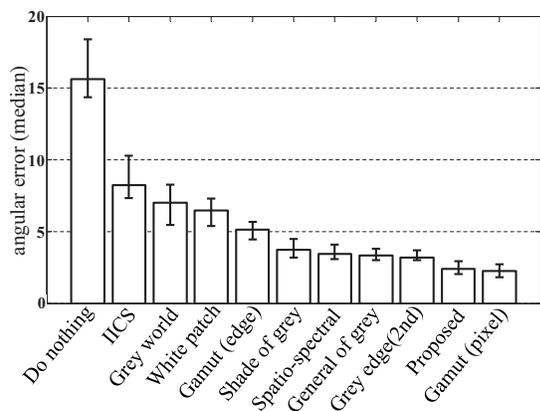

Figure 8. Median angular errors of different algorithms on SFU lab dataset plotted with a 95% confidence interval.

Methods	Median	Mean	Trimean	Max	Best-25%	Worst-25%
DN	15.60°	17.27°	16.90°	36.98°	3.57°	32.42°
IICS	8.23°	15.52°	11.89°	80.88°	2.21°	40.01°
GW	7.00°	9.78°	8.09°	37.30°	0.89°	23.33°
WP	6.47°	9.08°	7.58°	36.21°	1.83°	20.88°
GM(edge)	5.14°	6.53°	5.93°	30.01°	1.80°	13.27°
SG	3.74°	6.38°	4.97°	29.60°	0.58°	16.40°
SS	3.45°	5.63°	4.78°	21.56°	1.23°	12.84°
GG	3.32°	5.41°	4.08°	28.92°	0.49°	13.66°
GE2	3.18°	5.58°	4.13°	31.55°	1.05°	13.96°
GM(pixel)	2.26°	3.69°	2.65°	27.09°	0.45°	9.26°
Proposed	2.38°	4.82°	3.50°	24.90°	0.63°	12.73°

Table 2. Performance of various methods on the SFU lab dataset.

dataset, but worse than GM(pixel). However, there is almost no significant difference (Figure 8) between our model and GM(pixel) in terms of the measure of median angular error. More performance analysis is provided in the supplemental materials.

5. Conclusion and Future Work

We proposed a physiologically based color constancy model, which is inspired by the physiological research of color constancy in the human and primate visual system. The hierarchical steps of the proposed model correspond to the color processing mechanisms involved in the HVS from the retina to the primary visual cortex (V1). We found that the responses of double-opponent (DO) cells in V1 to the color-biased images contain the exact information about the scene illuminant; especially, the max responses of the DO cells could be utilized as the estimated illuminant. Systematical evaluation on two typical datasets validates the efficiency of the proposed model.

Though DO cells have been modeled by many researchers, even for the purpose of color constancy [9], our model differs entirely from them in how to utilize the output of DO cells. Dufort and Lumsden (1991) proposed a neural

captured with calibrated camera under 11 different lights in laboratory. Table 2 reports the angular error statistics of different algorithms on this dataset. Figure 8 shows the median angular errors over all of the 321 images in the dataset for different methods. We find that our model performs better than both statistics- and learning-based algorithms on this
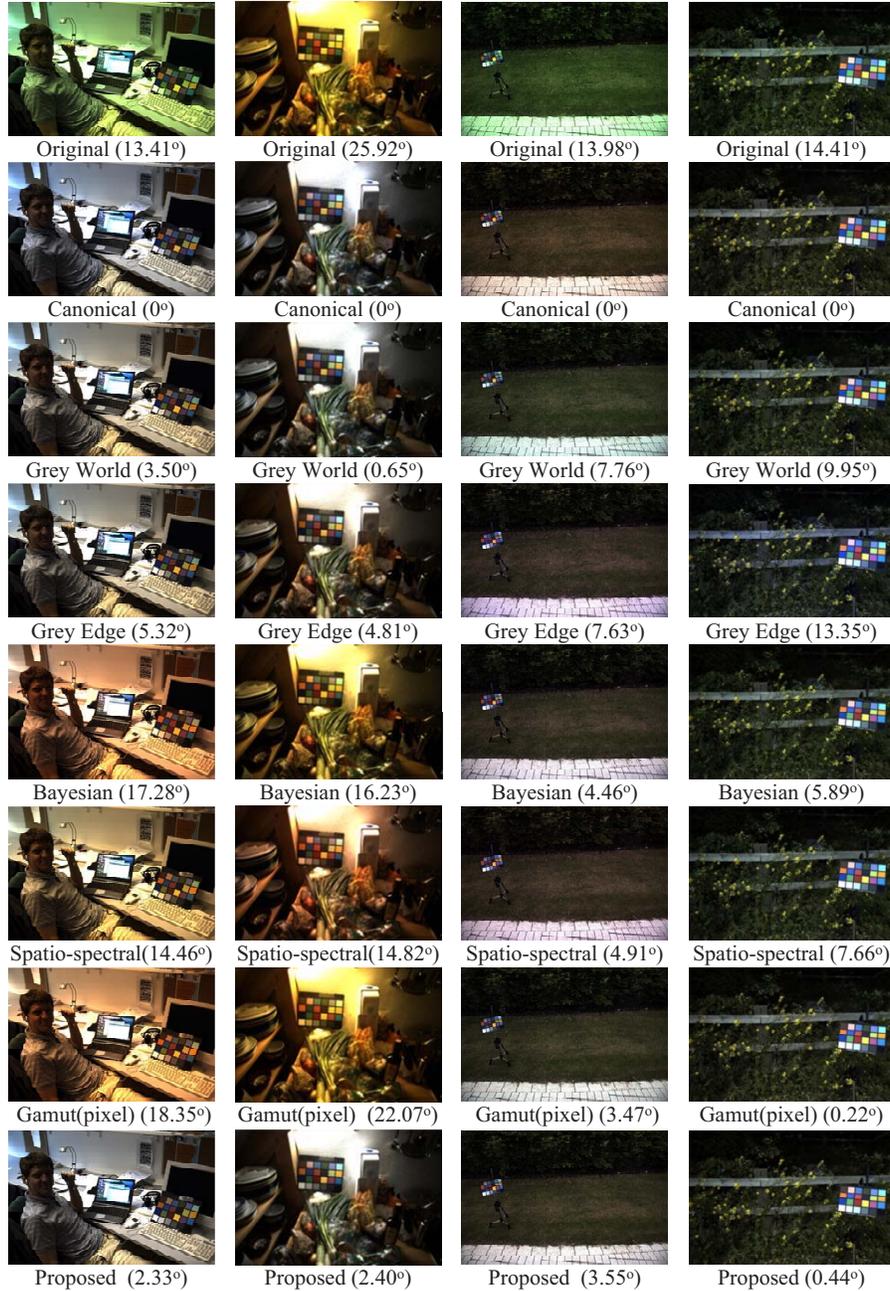
Figure 6. Some examples of indoor and outdoor images from Gehler-Shi dataset corrected with multiple methods.

network based on DO cells for color constancy [8]. In their model, color constancy is achieved by using the output from the DO cells as input for a neural network of four neurons in V4. As for the popular max-RGB, the light source color is estimated from the maximum response of the separate color channels [19]. Gray-edge hypothesis assumes that the average edge difference of a scene is gray [27].

Different from the above-mentioned typical methods, of which the assumptions are based on the color distribution (e.g., the max or average pixel values) in the original (or preprocessed) RGB images, our model computes the light source color by searching the maximum from the separate RGB channels, which are transformed from the DO responses obtained in the double-opponent space. In addition, our DO cells receive unbalanced cone inputs, which provides a flexible way to utilize the information of scenes. Because of the imbalance, these DO cells would respond best to (and enhance greatly) the edges defined by the chromatic and luminance differences [7, 13]. As most edges in the real world do combine luminance and chromatic differences [29], such cells would be extremely useful for the analysis of natural scenes. Furthermore, the unbalanced

center-surround RF structure endows the DO cells with a band-pass and partially low-pass spatial frequency tuning property [23], which ensures that the DO cells can also transfer partially the smoothed low-frequency global (possibly the light source) color information while smoothing and suppressing the color of local surfaces. These features endow our DO-based model with larger chance to get an accurate illuminant estimate by utilizing more effective local and global information of scenes.

Since it is generally accepted that color constancy is one of the distinct capabilities of the human visual system that no existing artificial intelligent system can compare with [9], we believe that building a physiologically based color constancy model is not only helpful for the insight understanding of the human visual system, but also quite useful to develop efficient color constancy algorithms for engineering applications.

As an important future direction, it is necessary to incorporate the non-linear properties of the early visual stages into the model, especially the non-linearity at the level of V1 [7, 13]. For example, the non-linear color-luminance interactions between the blue-yellow and luminance channels, which may reveal a possible neural correlate of stable perception of color constancy [18].

Acknowledgements

References

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

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

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

[4] G. Buchsbaum. A spatial processor model for object colour perception. *J Franklin Inst*, 310(1):1–26, 1980.

[5] M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *NAT REV NEUROSCI*, 13(1):51–62, 2011.

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

[7] B. R. Conway, S. Chatterjee, G. D. Field, G. D. Horwitz, E. N. Johnson, K. Koida, and K. Mancuso. Advances in color science: from retina to behavior. *J Neurosci*, 30(45):14955–14963, 2010.

[8] 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.

[9] M. Ebner. *Color constancy*, volume 6. 2007.

[10] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *CIC*, volume 2004, pages 37–41, 2004.

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

[12] B. Funt and W. Xiong. Estimating illumination chromaticity via support vector regression. In *CIC*, volume 2004, pages 47–52, 2004.

[13] K. R. Gegenfurtner. Cortical mechanisms of colour vision. *NAT REV NEUROSCI*, 4(7):563–572, 2003.

[14] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp. Bayesian color constancy revisited. In *CVPR*, 2008.

[15] A. Gijsenij and T. Gevers. Color constancy using natural image statistics and scene semantics. *IEEE PAMI*, 33(4):687–698, 2011.

[16] A. Gijsenij, T. Gevers, and J. Van De Weijer. Generalized gamut mapping using image derivative structures for color constancy. *IJCV*, 86(2-3), 2010.

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

[18] G. D. Horwitz, E. Chichilnisky, and T. D. Albright. Blue-yellow signals are enhanced by spatiotemporal luminance contrast in macaque v1. *J Neurophysiol*, 93(4):2263–2278, 2005.

[19] E. H. Land. *The retinex theory of color vision*. 1977.

[20] H.-C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *JOSA A*, 3(10):1694–1699, 1986.

[21] M. S. Livingstone and D. H. Hubel. Anatomy and physiology of a color system in the primate visual cortex. *J Neurosci*, 4(1):309–356, 1984.

[22] R. W. Rodieck. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision research*, 5(12):583–601, 1965.

[23] R. Shapley and M. J. Hawken. Color in the cortex: single- and double-opponent cells. *Vision research*, 51(7):701–717, 2011.

[24] 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/.

[25] S. G. Solomon and P. Lennie. The machinery of colour vision. *NAT REV NEUROSCI*, 8(4):276–286, 2007.

[26] R. T. Tan, K. Nishino, and K. Ikeuchi. Color constancy through inverse-intensity chromaticity space. *JOSA A*, 21(3):321–334, 2004.

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

[28] J. Van De Weijer, C. Schmid, and J. Verbeek. Using high-level visual information for color constancy. In *ICCV*, 2007.

[29] K. Yang, S. Gao, C. Li, and Y. Li. Efficient color boundary detection with color-opponent mechanisms. In *CVPR*, 2013.