

A Retinal Mechanism Inspired Color Constancy Model

Xian-Shi Zhang, Shao-Bing Gao, Ruo-Xuan Li, Xin-Yu Du, Chao-Yi Li, and Yong-Jie Li, *Member, IEEE*

Abstract—In this paper, we propose a novel model for the computational color constancy, inspired by the amazing ability of the human vision system (HVS) to perceive the color of objects largely constant as the light source color changes. The proposed model imitates the color processing mechanisms in the specific level of the retina, the first stage of the HVS, from the adaptation emerging in the layers of cone photoreceptors and horizontal cells (HCs) to the color-opponent mechanism and disinhibition effect of the non-classical receptive field in the layer of retinal ganglion cells (RGCs). In particular, HC modulation provides a global color correction with cone-specific lateral gain control, and the following RGCs refine the processing with iterative adaptation until all the three opponent channels reach their stable states (i.e., obtain stable outputs). Instead of explicitly estimating the scene illuminant(s), such as most existing algorithms, our model directly removes the effect of scene illuminant. Evaluations on four commonly used color constancy data sets show that the proposed model produces competitive results in comparison with the state-of-the-art methods for the scenes under either single or multiple illuminants. The results indicate that single opponency, especially the disinhibitory effect emerging in the receptive field's subunit-structured surround of RGCs, plays an important role in removing scene illuminant(s) by inherently distinguishing the spatial structures of surfaces from extensive illuminant(s).

Index Terms—Color constancy, retinal ganglion cells, non-classical receptive field, disinhibitory effect.

I. INTRODUCTION

THE LIGHTS entering into our eyes or cameras are normally determined by three factors: the physical properties of the object surfaces, the spectral irradiance of the

illumination, and the spectral sensitivities of the sensors. Color constancy means that the color perception of an object remains largely constant as the light source color changes [1]. Such an ability grants human beings clear evolutionary advantages in adaptation to widely varying external environments and also benefits a variety of computer vision applications [1]–[3]. Mathematically, it is an ill-posed problem, and without certain restrictions on surfaces and (or) illuminants, it is impossible to accurately and robustly recover the reflectance of surfaces from the color-biased images.

Due to the difficulty of this problem, there exists a large body of literature on the computational realization of color constancy. The methods available are mainly divided into two groups, i.e., illuminant estimation based approaches and color invariant approaches [2], [3]. Most of the models are characterized by the estimation of the illuminant. Typical examples of these methods first globally estimate the illuminant that is normally assumed to be uniform across the scene, and then remove it from the color-biased image to obtain the canonical image that seems captured under a white light source (i.e., the canonical illuminant). In general, based on the approaches of the illuminant estimation, these methods could be further broadly divided into three types [3], i.e., physics-based, low-level statistics-based and learning-based. The physics-based methods estimate the illumination on the basis of the physical interaction between the illuminants and the object surfaces. For example, as a typical one, the method of inverse-intensity chromaticity space (IICS) estimates the illumination chromaticity from single-colored and multi-colored surfaces [4].

The low-level statistics-based methods obtain the illuminant estimation based on some statistical assumptions about the distribution of colors in a scene. One of the well-known algorithms of this type is the Grey-World (GW), which is based on the assumption of achromatic average reflectance in a scene under a neutral light source [5]. Another popular algorithm in this group is the White-Patch (WP) with the assumption of perfect-reflectance maximum response in the RGB-channels [6]. Typical extensions of this type include Grey Edge (GE) [7], General Grey World (GG) [7] and Shades of Grey (SG) [8]. Our recent models based on double-opponent mechanism of the visual system [9] and based on local reflectance statistics [10] could be also classified into this group. In general, all these methods are simple in implementation, but most of them are based on specific assumptions, which are difficult to be always satisfied perfectly by the diverse reflectance distribution of real-world scenes.

Manuscript received July 6, 2015; revised December 1, 2015; accepted January 5, 2016. Date of publication January 12, 2016; date of current version January 26, 2016. This work was supported in part by the Major State Basic Research Program under Grant 2013CB329401, in part by the National Natural Science Foundation of China under Grant 61375115, Grant 91420105, and Grant 31300912, and in part by the 111 Project under B12027. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jean-Francois Aujol. (*Corresponding author: Yong-Jie Li.*)

X.-S. Zhang, S.-B. Gao, R.-X. Li, and Y.-J. Li are with the Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: zhangxianshi@163.com; gao_shaobing@163.com; 312365641@qq.com; liyj@uestc.edu.cn).

X.-Y. Du was with the School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China. He is now with the Infrastructure Inspection Institute, China Academy of Railway Sciences, Beijing 100081, China (e-mail: xinyudu@qq.com).

C.-Y. Li is with the Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China, and also with the Center for Life Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China (e-mail: cyli@sibs.ac.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2016.2516953

A learning-based method tries to learn using certain machine learning method to get an appropriate parameter setting on a given image set or a subset including a group of images having similar properties (e.g., indoor or outdoor images). The solutions could be classified into two groups, both aiming to overcome the limitation of the simple assumptions about the reflectance distribution made by the low-level statistics-based methods like GW and GE. (1) These methods introduce more complex statistics and a priori information about the surface reflectance and more sophisticated statistical computation based on learning is employed to estimate the illuminant. Typical examples include Gamut based methods (GM) [11]–[14], Bayesian [15], Regression (SVR) [16], Spatial Correlations (SC) [17] and corrected-moment based model [18]. For example, by assuming that the distribution of RGB color values of an image captured under a canonical illuminant is a limited set, the basic idea of gamut-based methods learn gamuts for different cameras and then use them to constrain the solution space for an input image. (2) These methods use the content of individual images to automatically select and tune the most appropriate (normally simple assumption based) algorithm (or combination of multiple algorithms) for each input image or a certain class of scenes. Typical examples of this group include Classification and Regression Trees (CART) based algorithm selection [19], Natural Image Statistics (NIS) [20], [21], Exemplar [22], and so on. In general, these learning-based methods could yield quite better results compared to those non-learning-based ones. However, they generally suffer from the complication in implementation and the requiring of appropriate preprocessing [1]–[3].

Without explicitly estimating the scene illuminant, color invariant approaches achieve color constancy by utilizing the invariant structures of the target images [3]. More or less, these methods are inspired by the information processing mechanisms of the human visual system (HVS), which has the amazing ability to achieve fairly good color constancy under varying natural illuminants [1]. Conceived as a description of human color perception, the Retinex model, the pioneering work of this type, holds the principle that the spectral properties of any object surface area can be approximated by the ratio of the reflected light in this area to the reflected light in other areas [6], [23], [24]. From this principle, with changes in the way to choose other areas, many extensions of this model have been derived [25]–[27], [64]. Other well-known methods in this type include the models based on retinal mechanisms of adaptation [28], [29] and the models based on neural networks [30], [31], [65], [66]. In general, these HVS based models could produce human-like performance on a variety of psychophysical paradigms designed to test color constancy. However, most of these above models (except for [28], [64]) did not show sufficient color constancy results with real images, so it is unclear whether the models work and how the models perform on various real scenes.

The model of Spitzer and Semo [28] is perhaps the most related one to the proposed. They simulated the color processing from the cone photoreceptors to the ganglion cells. In particular, they simulated the receptive field of the ganglion cells using a three-Gaussian model, which includes

a difference-of-Gaussians (DOG) shaped single-opponent classical receptive field (CRF) and a third broad Gaussian defining a much larger peripheral area that extends far beyond the CRF. Their model performs well in color correction of still images and video sequences when emphasizing a kind of remote adaptation as a gain control occurring in the peripheral area (i.e., the third Gaussian).

Generally speaking, our proposed model belongs to the color invariant based type. Different from most previous studies, however, our neural network model achieves color constancy by directly simulating the underlying mechanisms at the level of retina of the visual systems. It is generally suggested that retinal ganglion cells (RGCs), which respond in a color-opponent fashion within the center-surround structured receptive field (RF) to the activations of different cone classes, are the basis of color processing and the quite suitable building blocks for human color constancy [32]–[34]. More specifically, several physiological findings suggested that the RF surround of RGCs consists of many inhibitory subunits (or subfields), and inhibitory interactions among them lead to a disinhibitory effect, which means a suppression of the center-surround inhibition, in the region of non-classical receptive field (nCRF) [35]–[40]. These findings inspired us to computationally investigate their functional role in color information processing, which is not only an attempt to realize color constancy for computer vision, but also a description about the possible mechanisms of color constancy in biological vision.

The rest of this paper is organized as follows. In section 2, we describe the proposed model in detail. In section 3, we analyze the properties of the proposed model and evaluate it on four commonly used datasets (three single-illuminant datasets and one multi-illuminant dataset). In section 4, we present some concluding remarks and future directions.

II. MODEL

A. General Description

The proposed neural network model follows the color processing mechanisms in the retina (Fig 1). The color-biased image is the network input and the image removed illuminant color is the output.

The red (R), green (G) and blue (B) components of the input color image are sent respectively into long-, medium-, and short-wavelength cone photoreceptors (i.e., L, M, and S cones). The cone activities are adjusted by the feedback modulation from the horizontal cells (HCs), which can receive the original cone signals within relatively large fields. The modulated cone signals are then transmitted to the retinal ganglion cells (RGCs) via several retinal sub-layers like bipolar cells and amacrine cells. In the retinal ganglion layer, the output layer of the retina, color signals are processed by the RGCs with single-opponent receptive field (RF), which receives opponent stimuli in its excitatory center and inhibitory surround from two (or more) different cones. The surround comprises many inhibitory subunits, each of which is first inhibited by its neighboring subunits (i.e., disinhibit), and then inhibits the neuronal response elicited by the stimuli within the RF center. The disinhibition among subunits and the inhibition

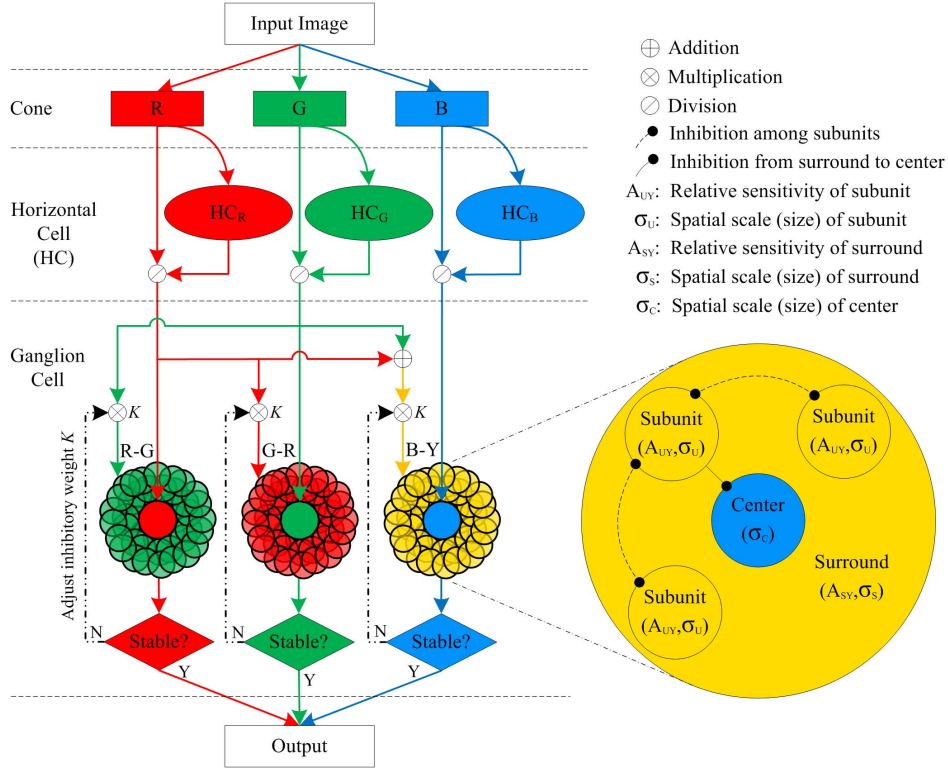


Fig. 1. The network structure of the proposed model. The R, G and B components of the input color image are respectively sent into the corresponding cone types at the level of cone photoreceptor, the outputs of which are modulated by horizontal cells with quite large receptive fields (RFs). The modulated cone signals are then iteratively processed by the single-opponent retinal ganglion cells (RGCs), the RF of which consists of a small excitatory center and a relatively large inhibitory surround (also named the non-classical receptive field, nCRF). The surround is composed of many inhibitory subunits, which first inhibit each other, and then inhibit the center. The iterative processing in RGCs is executed in each opponent channel of R-G, G-R and B-Y till the channel's output is stable. In the expression of "A-B", "A" and "B" denote respectively the signal components received by the excitatory center and the inhibitory surround.

from subunits to RF center vary iteratively, till all the three single-opponent channels reach stable outputs.

B. Horizontal Cell Modulated Cone Activities

Visual processing starts at the photoreceptor layer of the retina. There are two types of photoreceptors, rods and cones. According to the most sensitive wavelength, human cones can be basically classified into three types in terms of the most sensitive wavelength: short-wavelength cones (S-cones), medium-wavelength cones (M-cones) and long-wavelength cones (L-cones) with preferable responses to the blue (B), green (G) and red (R) colors, respectively [41]. In the following, we will use R, G and B to represent L, M and S, respectively.

The color information into eyes is coded in a trichromatic way via the three types of the cones. Given an input image $f_c(x, y)$, $c \in \{R, G, B\}$, the cone activity $F_c(x, y)$ is given by

$$F_c(x, y) = f_c(x, y) \otimes g(x, y; \sigma) \quad (1)$$

where \otimes is a convolution operator, and $g(x, y; \sigma)$ is a two-dimensional (2D) Gaussian function simulating the receptive field (RF) of the cones

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

where σ is the standard deviation controlling the RF size of the cones. In this work, we experimentally set $g(x, y; \sigma)$ as a rectangle template with $\sigma = 3.0$ and a size of 3×3 pixels.

After the light absorption by cones, bipolar cells transmit the cone activities to ganglion cells. Horizontal cells (HCs) provide a lateral modulation to this transmission [42]. HCs are the laterally interconnecting neurons that span widely across multiple cones and summate inputs from them. They help integrate and regulate the inputs from cone cells by measuring the average level of lights falling upon a relatively wide region of the retinal surface. Among several speculated functional roles, HCs have been best imagined as carrying out a global adjusting of the signals for reception by the cells in the inner retina, e.g., ganglion layer [42], and this global modulation by HCs has been suggested to play important role in correcting for the spectral composition of an illuminant [34], [43].

In the proposed model, we imitate this HC modulation as a cone-specific lateral gain control mechanism [43], which is written as

$$I_c = F_c / P_c, \quad c \in \{R, G, B\} \quad (3)$$

with

$$P_c = (\text{mean}(F_c^p))^{1/p} \quad (4)$$

where p is a free parameter emphasizing the bright pixels in an image [8]. P_c simulates the lateral gain control

contributed by HCs. Basically, the HC modulation described by Eqs (3) and (4) is consistent with the von Kries hypothesis (or von Kries adaptation) [34] aiming to account for the role of photoreceptor adaptation in color constancy, which is written as

$$\begin{bmatrix} I_R \\ I_G \\ I_B \end{bmatrix} = \begin{bmatrix} 1/P_R & 0 & 0 \\ 0 & 1/P_G & 0 \\ 0 & 0 & 1/P_B \end{bmatrix} \cdot \begin{bmatrix} F_R \\ F_G \\ F_B \end{bmatrix} \quad (5)$$

In fact, many existing color constancy models, e.g., grey-world [5], are based on the diagonal transform described by Eq (5), with different ways to compute the normalization factors P_c with $c \in \{R, G, B\}$. In particular, we compute P_c using Eq (4), a similar way as that introduced in General Grey World (GG) (i.e., the Grey-World with *pth*-Minkowski norm and zero-order spatial derivative) [7], considering that there is no center-surround mechanism occurring at the levels of cones and HCs for computing spatial image derivatives [33].

Because the strength of the modulatory signal from HCs varies with the relative activities of the different cones, this pathway helps counterbalance the gain of the cones [43]. For example, weakly activated cones receive a stronger HC modulation than strongly activated cones, and consequently, the spectral sensitivity of the cones is dynamically adjusted to the characteristics of the illuminant accordingly [44], [45].

Many studies have suggested that von Kries adaptation accounts only for a minor part of color constancy [11], [12], [32]–[34], mainly due to the independent adaptation among all three cone systems in the diagonal model of von Kries. To account for the contribution of interactions between the cone systems to the color perception [49], the diagonal model has been extended by replacing the diagonal matrix with a full matrix [34] or adding an offset term to the diagonal matrix [12], [46]. Physiologically, our model implements the interactions between the cone systems at the level of retinal ganglion cells, some of which have center-surround and color-opponent receptive fields integrating signals from different cone types.

C. Single-Opponency With Disinhibition

Retinal ganglion cells (RGCs) receive multiple cone signals transmitted via bipolar cells (and other cells) and compare them with the color-opponent mechanism. The receptive field (RF) of most RGCs consists of two regions, i.e., a smaller excitatory (ON) center and a larger inhibitory (OFF) annular surround, and chromatically single-opponent RGCs receive inputs of different cone types within these two different RF regions [33], [40], [49]. In this work we consider three types of single-opponent RGCs [47]–[49]: L-on/M-off, M-on/L-off, and S-on/(L+M)-off, which means that the firing rate of a RGC increases with the activation of one cone type (e.g., L) and decreases with the activation of a different cone type (e.g., M). In the following, we will use R-G, G-R, and B-Y for short to denote respectively the above three single-opponent channels.

The simple center-surround RF structure, also called the classical receptive field (CRF), has conventionally been described by the “difference of Gaussian” (DOG)

model [50], [51], which postulates that a cell’s response is equal to the difference of signals from the excitatory center and the inhibitory surround, described by two Gaussian kernels having different standard deviates. However, according to the physiological experiments [35]–[39], [52], a secondary gentle rise in neuronal responses of some RGCs has been clearly observed when the stimulus was further extended beyond the inhibitory near surround of CRF, indicating a disinhibitory effect contributed by the extensive far surround outside the DOG-shaped CRF, i.e., the activity elicited by the stimulus in this far surround can reduce the inhibitory effect of the near surround. We have speculated [40] that the near and far surrounds could be unified as an extensive disinhibitory region beyond the excitatory RF center, and this disinhibitory region (also called the non-classical receptive field, nCRF) consists of many inhibitory subunits, which first inhibit each other, and then inhibit the RF center. This idea has the following two possible biological plausibilities [53], [54]: (1) RGCs receive excitatory inputs from multiple bipolar cells, which have much smaller RFs than RGCs and therefore constitute many subunits, especially within the extensive RF surround of RGCs; (2) narrow-field amacrine cells covering several bipolar cells help to construct direct interactions among a subset of neighboring subunits via inhibitory amacrine synapses. This work attempts to incorporate this mechanism of disinhibitory effect into the surround of the color-opponent RGCs.

Let $RG_{Sub}(x, y; \sigma_u)$, $GR_{Sub}(x, y; \sigma_u)$, $BY_{Sub}(x, y; \sigma_u)$ denote respectively the responses of the subunits in RF surround after being inhibited by other neighboring subunits in the R-G, G-R and B-Y single-opponent channels. We compute them as

$$\begin{aligned} RG_{Sub}(x, y; \sigma_u) &= MAX[0, I_G(x, y) - A_{UG} \sum_{(p,q) \in Subunit \setminus (x,y)} I_G(p, q) \\ &\quad \times g(|p - x|, |q - y|; \sigma_u)] \\ GR_{Sub}(x, y; \sigma_u) &= MAX[0, I_R(x, y) - A_{UR} \sum_{(p,q) \in Subunit \setminus (x,y)} I_R(p, q) \\ &\quad \times g(|p - x|, |q - y|; \sigma_u)] \\ BY_{Sub}(x, y; \sigma_u) &= MAX[0, I_Y(x, y) - A_{UY} \sum_{(p,q) \in Subunit \setminus (x,y)} I_Y(p, q) \\ &\quad \times g(|p - x|, |q - y|; \sigma_u)] \end{aligned} \quad (6)$$

where $I_Y = (I_R + I_G)/2$, indicating that the yellow (Y) signal received in the surround of B-Y ganglion cells results from the combination of red (R) and green (G) cone signals. A_{UG} , A_{UR} and A_{UY} represent the sensitivities of the subunits in R-G, G-R and B-Y channels, respectively. (p, q) belongs to the point set in a subunit excluding its center. The MAX operator is used to keep the neuronal responses of subunits non-negative.

To compute the responses of the inhibited subunits using formula (6), we first construct a 2D Gaussian template of subunit based on the given standard deviation σ_u . Then we set the value of template center to be zero. By convoluting the

image, e.g., $I_G(x, y)$, with this template, we obtain the inhibition from the subunit surround. Finally, we get the responses of the inhibited subunits $RG_{Sub}(x, y; \sigma_u)$ by subtracting this inhibition (weighted by A_{UG}) from $I_G(x, y)$.

It is interesting to point out that Eq. (6) works like a high pass filter, since Eq. (6) is computationally equivalent to subtracting the Gaussian smoothed stimulus in the surround from the original stimulus in the surround, which makes the high frequency components of the stimulus remained in the surround. The role of high pass behavior of the subunits will be clarified in Section III.

Let $RG_{Sur}(x, y; \sigma_s)$, $GR_{Sur}(x, y; \sigma_s)$ and $BY_{Sur}(x, y; \sigma_s)$ denote the total responses of all the inhibited subunits in the surround in R-G, G-R and B-Y channels. We compute them as

$$\begin{aligned} RG_{Sur}(x, y; \sigma_s) &= A_{SG} \times \sum_{(p,q) \in Surround} RG_{Sub}(p, q; \sigma_u) \\ &\quad \times g(|p-x|, |q-y|; \sigma_s) \\ GR_{Sur}(x, y; \sigma_s) &= A_{SR} \times \sum_{(p,q) \in Surround} GR_{Sub}(p, q; \sigma_u) \\ &\quad \times g(|p-x|, |q-y|; \sigma_s) \\ BY_{Sur}(x, y; \sigma_s) &= A_{SY} \times \sum_{(p,q) \in Surround} BY_{Sub}(p, q; \sigma_u) \\ &\quad \times g(|p-x|, |q-y|; \sigma_s) \end{aligned} \quad (7)$$

where A_{SG} , A_{SR} and A_{SY} are the sensitivities of the surround in the R-G, G-R and B-Y channels. (p, q) belongs to the point set in the inhibitory annular surround.

The final neuronal response is then computed by subtracting the total surround inhibition from the response of excitatory RF center, which is written as

$$\begin{aligned} RG(x, y; \sigma_c) &= MAX[0, \sum_{(p,q) \in Center} I_R(p, q) \\ &\quad \times g(|p-x|, |q-y|; \sigma_c) - RG_{Sur}(x, y; \sigma_s)] \\ GR(x, y; \sigma_c) &= MAX[0, \sum_{(p,q) \in Center} I_G(p, q) \\ &\quad \times g(|p-x|, |q-y|; \sigma_c) - GR_{Sur}(x, y; \sigma_s)] \\ BY(x, y; \sigma_c) &= MAX[0, \sum_{(p,q) \in Center} I_B(p, q) \\ &\quad \times g(|p-x|, |q-y|; \sigma_c) - BY_{Sur}(x, y; \sigma_s)] \end{aligned} \quad (8)$$

where $RG(x, y; \sigma_c)$, $GR(x, y; \sigma_c)$ and $BY(x, y; \sigma_c)$ are the final responses of the single-opponent cells in the R-G, G-R and B-Y channels. (p, q) belongs to the point set in the excitatory RF center. Similar to Eq. (6), here the MAX operator is used to guarantee the neuronal responses to be non-negative.

In Eqs (6)~(8), σ_u , σ_s , and σ_c are the standard deviations of the 2D Gaussian functions describing the subunit, RF surround and RF center, respectively. We experimentally set $\sigma_u = 0.5$, $\sigma_s = 1.5$, and $\sigma_c = 0.5$ in this work. Considering the fact that the Gaussian values would be reduced quite close to zero when its radius is larger than three times of the standard deviation, we roughly set the standard deviation dependant radius of subunit, surround and center as 1, 3, 1 pixel(s), respectively. Partial evidence in support of such setting is

the neurophysiological finding that the size of RF surround is normally 2 to 5 times larger (in diameter) than that of RF center [35]–[39]. Note that such parameter setting will be further experimentally analyzed in Section III.

D. Adaptation of Ganglion Cells

As mentioned above, sufficient studies have indicated that color adaptation starts in the retina [43]–[45]. In particular, RGCs have the capability of adjusting their sensitivities over a few seconds to the chromatic statistics of different visual environments [55]. It has been generally accepted that this adaptation to the statistical properties of visual scenes plays a critical role in emphasizing more relevant visual inputs or increasing the efficiency of color signal transmission from the retina to the brain [55], [56]. It was speculated that this type of retinal adaptation may require inhibitory transmission in the inner retina, e.g., the wide-field transient inhibition from the plentiful plastic synapses of the amacrine cells received by nearly all ganglion cell types [53], [57]. This plasticity, together with other unknown mechanisms, helps adaptively change the spatial (as well as temporal) properties of the ganglion-cell receptive fields in a way that reduces sensitivity to the dominant correlation structure of the visual scenes, e.g., the globally distributed illuminant color, until visual sensitivity is roughly constant.

We formalize the above dynamic adaptation as a simple process in each of the R-G, G-R and B-Y opponent channels. For simplicity, all the three opponent channels share the same surround sensitivity and also share the same subunit sensitivity. This adaptation is described by

$$A_{SR} = A_{SG} = A_{SY} = K \quad (9)$$

$$A_{UR} = A_{UG} = A_{UY} = \alpha A_{SR} \quad (10)$$

where α is a constant controlling the relative sensitivities of subunits compared to the surround, and we experimentally set $\alpha = 1/3$ in this work, since such setting has been used to successfully explain the psychophysical observations like Mach bands [40]. K is a parameter defined as the inhibitory weight, which is automatically determined as follows. K starts from 0.0 and is increased in a step of 0.2 during the adaptation procedure. $K = 0$ corresponds to the state of default receptive fields of RGCs, which are shaped mainly by the fixed excitatory synapses from bipolar cells [53]. The increasing of K , and hence the increasing of the subunit sensitivities (i.e., A_{UR} , A_{UG} , A_{UY}) and surround sensitivities (i.e., A_{SR} , A_{SG} , A_{SY}), corresponds to the increasing of inhibitory amacrine synapses.

We define the spatial average of the recovered surface reflectivity (SPARSUR) in a channel as a measure of the status of this channel. At the end of each iteration, the current status of each channel is separately evaluated by computing SPARSUR. In two successive iterations, if the change of SPARSUR in a channel is small enough, this channel is considered to reach stability, which means a state such that the decreasing of surround inhibition strength due to the increasing disinhibition sensitivity is balanced by the increasing of surround inhibition strength due to the increasing surround

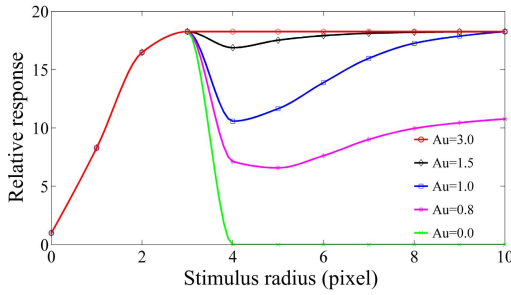


Fig. 2. Response curves as a function of stimulus size (in radius) for different parameter settings. The sizes (in radius) of RF center, RF surround and subunits (i.e., $3\sigma_c$, $3\sigma_s$, and $3\sigma_u$) are 3, 10, 3 pixels, respectively. The sensitivity of RF surround (i.e., A_{SR} , A_{SG} , A_{SY}) is 3.0, and the subunit sensitivity $A_{UR} = A_{UG} = A_{UY} \triangleq A_u$ is set to be 3.0, 1.5, 1.0, 0.8 and 0.0, corresponding to the five curves (from top to bottom).

inhibition sensitivity. For example, if the SPARSUR of the red component suppressed by surrounding green component in the R-G opponent channel changes negligibly, the R-G channel reaches its stable state and the adaptation process of this channel will be finished. The inhibition adaptation continues in other unstable channels until all the three opponent channels reach their stable states. To summarize, without the need of an explicit learning phase, our model processes the individual images with adaptive fine-tuning of the surround and subunit sensitivity values.

With the automatically selected surround and subunit sensitivities mentioned above, $RG(x, y; \sigma_c)$, $GR(x, y; \sigma_c)$ and $BY(x, y; \sigma_c)$ computed by Eq. (8) are respectively the final outputs in R, G, and B channels, which are combined to form the final output image. Note that in general, the range of pixel intensity of the output image does not cover the interval $[0, 255]$. For better visualization, we linearly normalize the final output image to the interval $[0, 255]$ for display in this work.

III. EXPERIMENTS

In this section we first show examples analyzing the properties of the proposed model, and then demonstrate validation tests on three single-illuminant datasets and one multi-illuminant dataset. Note that in the following, we will denote the R-G, G-R, and B-Y opponent channels by R, G, and B channels for short, e.g., here the “R” is the output of the original R component inhibited by the G component in the surround of an R-G ganglion cell.

A. Disinhibitory Effect of the Model Ganglion Cell

1) *Area-Response Properties*: As disinhibition in the RF surround of ganglion cell is caused by the inhibitory interactions among surround subunits, subunit sensitivities (i.e., A_{UR} , A_{UG} and A_{UY}) are the main factors determining the strength of surround disinhibition. With a stimulus pattern of uniform luminance that is large enough to cover both the RF center and its surround of a model ganglion cell, we measured a group of area-response curves by computing the response of the model cell with varying subunit sensitivities while keeping the other parameters fixed (Fig. 2). It is clear

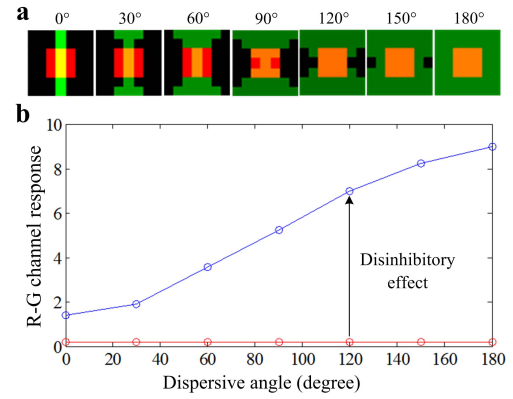


Fig. 3. The effect of various foreground patterns on the responses of a R-G ganglion cell. (a) A group of stimulus patches with a full size of 5×5 pixels (which is enlarged for better visualization). Each patch contains a red block of 3×3 pixels only covering the full RF center and a patterned green foreground across both the RF center and its surround. The green foreground in different patches have the same total light flux but different dispersive angles, from 0° to 180° , as indicated on the top of the panels. (b) Response vs. dispersion curves of a R-G opponent ganglion cell. The red line is for the cell with surround inhibition but without subunit disinhibition. The blue line is for the case that both the subunit disinhibition and surround inhibition work. Here the term “dispersive angle” is defined as the angle within which the green component is distributed between the two opposite sectors of the circle. The minimum dispersive angle is 0° , which means a straight line. The maximum angle is 180° , which means a fully dispersed distribution.

that the neuronal response is gradually increased when the stimulus size is extended within the excitatory RF center (with a radius of 3 pixels in this test). When the stimulus is further extended larger than the RF center, the neuronal response is decreased due to the surround inhibition, and then gradually increased due to the increasing disinhibition effect deduced by the inhibitory interaction among more surround subunits. These area-response curves shown in Fig. 2 are quite consistent with those obtained from our previous physiological studies in the cat ganglion cells [35], [37], [58].

Figure 2 indicates that with appropriate subunit sensitivity A_u (which is adaptively determined by the model), the reproduced disinhibition among subunits has the potential to reduce the inhibition from the extensive surround when being covered by dispersively distributed light source color, and hence, obtain a relatively stronger response to the true color of the center, which helps to weaken the color bias of the center.

2) *Dispersion-Dependent Effect*: Here we demonstrate what would happen when the single-opponent mechanism is introduced into the surround subunit model of ganglion cell, i.e., subunit interaction in RF surround occurs in one color channel (e.g., G) and then the surround inhibits the RF center that receive stimulus in another color channel (e.g., R). We designed a series of stimulus patches (Fig. 3a), each of which contains a red block covering the full RF center and a patterned green foreground across both the RF center and its surround. When the green foreground was spatially distributed with systematically increasing dispersion (quantified by a “dispersive angle” [35], [37], [58]) while keeping the total light flux identical for the green pattern, the responses of a R-G opponent ganglion cell to the patterned stimuli were consistently increased with the increasing dispersive angle of the green foreground (Fig. 3b). This is because more subunits

in the surround were activated and hence stronger disinhibition (and then weaker surround inhibition) was introduced.

We can find that when the green foreground of the stimulus patch is more dispersive, the foreground color seems more likely from external illuminant (e.g., the last patch of Fig. 3a). In this situation, the yellowish appearance of the red block was reduced by the increasing of the R-G neuronal response to the red block covering the RF center, i.e., the yellowish appearance of the red block is removed and the true red color is recovered.

From the point of view of filtering, as mentioned earlier, interaction among subunits shows high pass property, which helps to clean or attenuate the uniform stimulus in the surround (e.g., the last panel of Fig. 3a), which results in the reduced surround inhibition, and hence, the relatively increased the neuronal response to the stimulus in the RF center. In other words, the local patch color in the center (most likely the object surface) that has been attenuated by a large patch of another color (large enough to cover both the center and surround) may be recovered by reducing the surround inhibition due to the high pass property of the subunits.

This example provides a biological explanation how the disinhibition among subunits in the surround contributes to the color constancy at the retinal ganglion level. Note that the yellowish appearance of the red block cannot be reduced when no disinhibition was involved (i.e., $A_{UR} = A_{UG} = A_{UY} = 0$), because the surround inhibition always keeps constant due to the identical total light flux (with any degree of dispersion) in the surround. That is, the opponent RF without disinhibition has difficulty in responding selectively to the true surface color and the light source color.

3) *Disinhibitory Effect on Color Constancy*: Figure 4 shows how the subunit disinhibition plays a role in color constancy in each opponent channel. For a scene under a bluish light source (the first panel in the first row of Fig. 4a), it is obvious that the blue component of the color-biased image is more dominant than the other two channels. With the increase of inhibitory weight K (Fig. 4b), the three channels approach the stable state after different periods of variation. In general, the channels of scenes with stronger color bias require longer adaptation time to reach steady states, e.g., the blue curve in Fig. 4b for the blue channel. This indicates that for a color-biased scene, different opponent ganglion cells with various subunit and surround sensitivities are employed to reduce the different bias of color opponents. To summarize, Figure 4 demonstrates that for a color-biased image with the different degree of color distortion across different color channels, early or late, all the three channels can reach stable after a period of iteration that is automatically terminated by the proposed model.

Given the true illuminant of a scene, we can employ the commonly used *recovery angular error* as the metric for performance evaluation [59], which is defined as the angle between the RGB of the actual measured illuminant color and that estimated one as the recovery error

$$\varepsilon = \cos^{-1} \left[(\vec{e}_t \cdot \vec{e}_e) / (\|\vec{e}_t\| \cdot \|\vec{e}_e\|) \right] \quad (11)$$

where \vec{e}_t and \vec{e}_e are respectively the true and the estimated illuminants containing R, G and B three components.

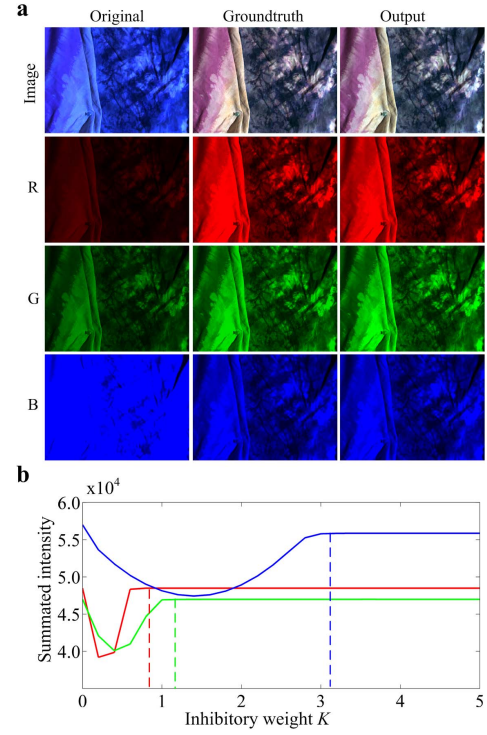


Fig. 4. The disinhibitory effect on color constancy in each channel. (a) The first row: a color-biased image under blue light source, its canonical image under white light source (i.e., the groundtruth) and the output corrected by our model. The second to fourth rows list the corresponding image components in R, G, and B channels, respectively. (b) In each image channel, the inhibitory weight K is iteratively increased (with a step of 0.2) to find the stable state of each channel by judging whether or not the summed intensity over the scene reaches stable in two successive iterations. For the case shown here, the iterations in the channels of R (red curve), G (green curve) and B (blue curve) were terminated when $K = 0.8, 1.2$ and 3.2 , respectively.

Note that because the proposed model does not estimate the illuminant explicitly, we derive the three components of the illuminant estimate (only for computing ε and ε_{REP} for comparison) by computing in each channel the ratio of summed pixel values of the recovered image to that of the original color-biased image.

Figure 5 illustrates that on a synthetic scene comprising red grating (the first panel of Fig. 5a), though the influence of greenish light source could be weakened more or less by the horizontal-cell modulation in the first phase of our model (from the second to third panels of Fig. 5a), the illuminant influence was more greatly reduced by the color-opponent ganglion cells with subunit disinhibition (the last panel of Fig. 5a). Figure 5b indicates that subunit disinhibition (together with surround inhibition) with appropriate inhibitory weights helps obtain smaller angular errors (the blue curve). In contrast, the measure of angular error is difficult to be reduced when the subunit disinhibition or surround inhibition is omitted (the red or green curve). These observations were further validated on a real image shown in Fig. 6, where a bluish scene was recovered as close as to the canonical image by the ganglion cells with both the surround inhibition and subunit disinhibition.

In short, Figure 5 and 6 show that compared to the ganglion cells without subunit disinhibition and (or) surround inhibition,

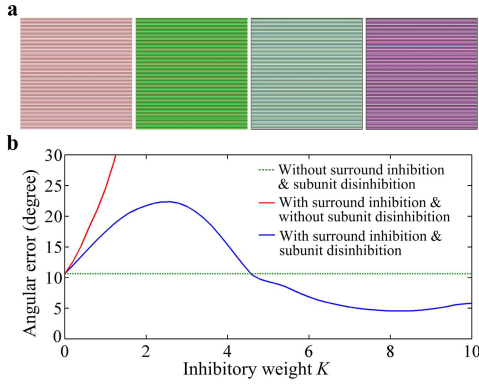


Fig. 5. Illustration of the role of disinhibition in color constancy on a synthetic image. (a) From left to right: a synthetic image comprising red grating under a white light source (i.e., the canonical image), the same scene under a greenish light source, the image corrected only with HC modulation given by Eqs (3) and (4), and the image further corrected by RGCs with disinhibitory RF surround. (b) The relationship between the inhibitory weight K and the angular error. The dotted green curve is with $A_{SR} = A_{SG} = A_{SY} = 0$ (i.e., the surround inhibition does not work, and hence, the subunit disinhibition is also omitted). The red curve is with $A_{UR} = A_{UG} = A_{UY} = 0$ (i.e., the surround inhibition works while the subunit disinhibition is non-effective). The blue curve is for the case that both the subunit disinhibition and surround inhibition work. It is clear that with a suitable value for the inhibitory weight K , the blue curve can reach an angular error that is quite lower than other two curves. In other words, RGCs with both the subunit disinhibition and surround inhibition have chance to produce better performance.

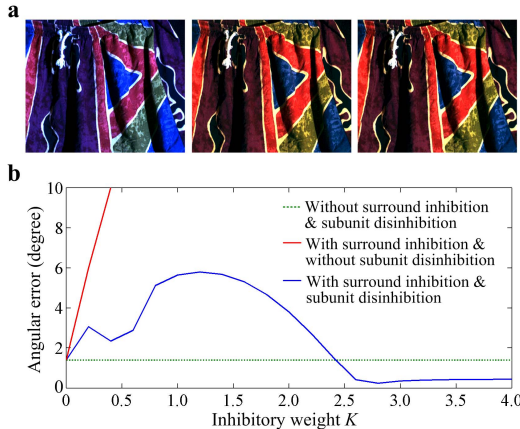


Fig. 6. Illustration of the role of disinhibition in color constancy on a real image. (a) From left to right: the color-biased image in a bluish illuminant (from SFU Lab Dataset [60]), the canonical image from white light source, the corrected output of the color-biased image. (b) The relationship between the inhibitory weight K and the angular error. The meanings of the red, blue and green curves are the same as those of Figure 5.

the cells with both subunit disinhibition and surround inhibition have chance to better remove the influence of light source color (and hence, lower angle error) with appropriate K values.

B. Performance on Three Single-Illuminant Datasets

The proposed model was further evaluated on three commonly used datasets: the SFU Laboratory Dataset [60], the SFU Grey-ball Dataset [61], and the Gehler-Shi Color-checker Dataset [15], [62]. We considered four groups of representative models for comparison, including: (1) physics-based: IICS [4]; (2) low-level statistics-based: GW [5], WP [6], GE [7], SG [8], GG [7]; (3) learning-based: GM (pixel) [11], GM (edge) [12],

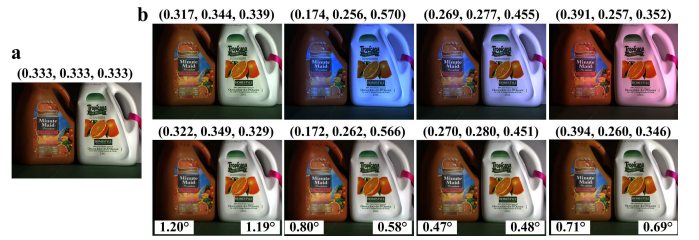


Fig. 7. Some example images of a same scene under various illuminants from the SFU Laboratory Dataset [60]. (a) The canonical scene (from white light source with equal R, G and B components). (b) Top row: four images of the same scene with different light source colors, on the top of which the R, G and B components of the light source are indicated; bottom row: the corresponding results corrected by the proposed model, the estimated illuminant, the reproduction angular error and the conventional angular error are shown on the top, the left and right bottom corners of each image.

CART [19], Bayesian [15], NIS [20], SVR [16], SC [17], Exemplar [22]; (4) color invariant based: RetinexFrankle [25], Retinex99 [26]. Note that from the point of view of the underlying motivation, our model is quite close to the color invariant based models like RetinexFrankle and Retinex99.

The results of the RetinexFrankle and Retinex99 models were reproduced in this study according to the parameters provided in [63], and the results of other methods on the first three single-illuminant datasets are directly cited from the website <http://colorconstancy.com/>.

Considering that the angular error defined by Eq. (11) does not measure how similar the results are to the canonical images, Finlayson and Zakizadeh [71] recently proposed a new measure named *reproduction angular error*, which is defined as the angle between the image RGB of a white surface when the actual and estimated illuminations are removed

$$\varepsilon_{REP} = \cos^{-1} \left[\frac{(\vec{e}_t \cdot \vec{e}_e)}{(\sqrt{3} |\vec{e}_t| |\vec{e}_e|)} \right] \quad (12)$$

It is clear that the two metrics mentioned above emphasize differently on performance evaluation. The angular error focuses on the difference between the estimated illuminant and the real illuminant, while the reproduction angular error measures the discrepancy between the real reflectivity of the object surface and the reproduced one [71], [72]. In the following experiments, we used the both measures defined by Eqs (11) and (12) for performance evaluation in this work.

1) *SFU Laboratory Dataset*: The SFU Laboratory Dataset consists of 31 different scenes, captured with calibrated camera from 11 different light sources for each scene, resulting in a total of 321 images [60]. The images in this dataset can be classified into two subsets: images with minimal specularities (22 scenes, 223 images) and images with dielectric specularities (9 scenes, 98 images). Examples of a same scene from various light sources are shown in Fig.7.

Table I reports the recovery angular error statistics of different methods on the whole dataset. The result of SVR comes from [16]. It is clear that in terms of median error, the proposed model performs as well as the best low-level statistics-based method (GE 2nd), even better than the learning-based SC method, but worse than the best learning-based GM and SVR algorithms.

TABLE I
COMPARISON TO OTHER MODELS ON THE SFU LABORATORY DATASET

Method		Recovery Angular Error		Reproduction Angular Error	
		Median	Mean	Median	Mean
Physics-based	IICS	8.2°	15.5°	9.3°	15.1°
Low-level statistics-based	GW	7.0°	9.8°	7.5°	10.1°
	WP	6.5°	9.1°	7.4°	9.7°
	SG ($p=7$)	3.7°	6.4°	3.9°	6.9°
	GG ($p=10, \sigma=4$)	3.3°	5.4°	3.9°	6.0°
	GE (1 st , $p=7, \sigma=4$)	3.2°	5.6°	3.6°	6.3°
	GE (2 nd , $p=7, \sigma=4$)	2.7°	5.2°	3.3°	5.8°
	GE (2 nd , $p=15, \sigma=10$)	2.8°	5.3°	3.0°	5.8°
Learning-based	SC	3.5°	5.6°	4.1°	6.2°
	GM (pixel, $\sigma=4$)	2.3°	3.7°	2.8°	4.2°
	GM (edge, $\sigma=2$)	2.3°	3.9°	2.7°	4.5°
	SVR	2.2°	-	-	-
Color invariant	Retinex99 ($n=4$)	13.3°	15.2°	13.7°	15.1°
	RetinexFrankle ($n=4$)	13.0°	15.2°	13.7°	15.0°
	Proposed ($p=13$)	2.7°	6.0°	3.0°	6.5°

Table I also reports the reproduction angular errors. Note that the listed results of other models are from [71]. We can find from this table that our model remains its ranking in terms of median error, and this is almost true for other datasets tested below (not shown in this paper due to the space limitation). This is consistent with the finding in [71] and [72] that the results of reproduction and recovery metrics for the same algorithm are very much correlated when the scenes are diverse.

The bottom row of Fig. 7(b) reports the recovered images by our model. It is clear that the different illuminants of the same scene were successfully removed based on the accurate illuminant estimates, which results in quite low recovery and reproduction angular errors and close similarities between the recovered and canonical (ground-truth) images. Note that the values of the two measure types for a same image vary a little, since the measure of reproduction angular error especially emphasizes how well a white surface is reproduced, as mentioned before. It has been recommended to adopt the reproduction error when evaluating on specific images of the same scene captured under different illuminations [72]. However, considering that most of the images of the datasets tested in this work are from diverse scenes under various illuminants, both the two types of angular errors are suitable metrics for evaluating the overall performance of a model on the whole datasets [72], as indicated by Table I.

Though our model performs clearly better than the grey-world based methods like GW, WP, SG and GG on this dataset, the first phase of our model (i.e., horizontal-cell modulation) seems to work similarly to GG. So, we conducted further evaluation to investigate the influence of statistical properties of scenes on the performance of our full model based on the various scenes in this dataset. First, to quantify the deviation of the average reflectance in a scene from the grey-world assumption, we define a measure called *reflectance chromaticity* (RC), which is computed as the angular error between a hypothetical light, whose R, G and B components are the summated intensities of each channel of the true scene (under white light source), and an achromatic light (with equal R, G and B components). It is obvious that a higher RC corresponds to a larger deviation from the grey world assumption. Figure 8 shows the relation between the RCs and the median angular errors of all scenes in the dataset. It is clear that in general, the larger the RC is, the higher the median angular error is (the linear correlation coefficient between the two curves is 0.6, and the p -value for this hypothesis is 0.00045). In other words, our model has a higher chance to obtain better performance on the scenes that match closer the grey-world assumption. This is consistent with most of the grey world assumption based models [2], [3].

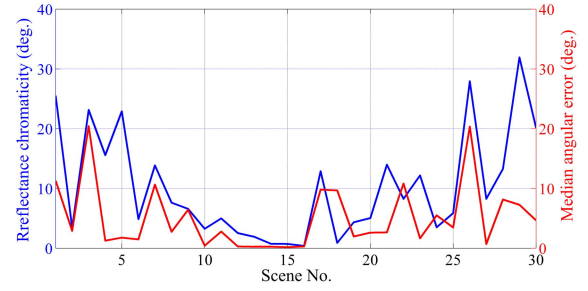


Fig. 8. The relation of median angular error to the reflectance chromaticity of scenes. The blue line shows the average reflectance chromaticity (RC) of each scene. The red line shows the median angular error by our model over the 11 images of each scene. Note that among the total 31 scenes, we have combined the scenes of *apples* and *apples2* into one group, which results in a total of 30 scenes.

TABLE II
COMPARISON TO OTHER MODELS ON THE SFU GREY-BALL DATASET

Method		Median	Mean	Max
Physics-based	IICS	5.6°	6.6°	76.2°
Low-level statistics-based	GW	7.0°	7.9°	48.1°
	WP	5.3°	6.8°	38.7°
	SG ($p=12$)	5.3°	6.1°	41.2°
	GG ($p=12, \sigma=0$)	5.3°	6.1°	41.2°
	GE (2 nd , $p=1, \sigma=2$)	4.9°	6.1°	41.7°
	GE (1 st , $p=1, \sigma=1$)	4.7°	5.9°	41.2°
Learning-based	GM (pixel, $\sigma=5$)	5.8°	7.1°	41.9°
	GM (edge, $\sigma=2$)	5.8°	6.8°	40.3°
	NIS	3.9°	5.2°	44.5°
	Exemplar	3.4°	4.4°	45.6°
Color invariant	RetinexFrankle ($n=4$)	5.5°	6.3°	41.2°
	Retinex99 ($n=4$)	5.4°	6.2°	42.1°
	Proposed ($p=24$)	4.6°	6.3°	35.3°

chromaticity (RC), which is computed as the angular error between a hypothetical light, whose R, G and B components are the summated intensities of each channel of the true scene (under white light source), and an achromatic light (with equal R, G and B components). It is obvious that a higher RC corresponds to a larger deviation from the grey world assumption. Figure 8 shows the relation between the RCs and the median angular errors of all scenes in the dataset. It is clear that in general, the larger the RC is, the higher the median angular error is (the linear correlation coefficient between the two curves is 0.6, and the p -value for this hypothesis is 0.00045). In other words, our model has a higher chance to obtain better performance on the scenes that match closer the grey-world assumption. This is consistent with most of the grey world assumption based models [2], [3].

2) *SFU Grey-Ball Dataset*: The SFU Grey-ball Dataset contains 11346 images, extracted from a video of 2 hours recorded under a large variety of conditions. The ground truth illuminant was captured by a grey ball mounted on the camera [61]. Table II reports the angular error statistics of different methods over the whole dataset. As done with other models, we also masked the grey ball of each scene when running our model.

From Table II, the proposed method obtains the measure of median angular error that is lower (better) than all the low-level based models listed here on this dataset (and is equivalent

TABLE III
COMPARISON TO OTHER MODELS ON THE GEHLER-SHI DATASET

Method		Median	Mean	Max
Physics-based	IICS	13.6°	13.6°	56.7°
	GW	6.3°	6.4°	24.8°
Low-level statistics-based	WP	5.7°	7.6°	40.6°
	GE (1 st , $p=1$, $\sigma=6$)	4.5°	5.3°	26.4°
	GE (2 nd , $p=1$, $\sigma=1$)	4.4°	5.1°	23.9°
	SG ($p=4$)	4.0°	4.9°	22.4°
	GG ($p=9$, $\sigma=9$)	3.5°	4.7°	22.0°
Learning-based	SVR	6.7°	8.1°	32.0°
	GM (edge, $\sigma=4$)	5.0°	6.5°	29.0°
	CART	3.9°	4.5°	22.3°
	Bayesian	3.5°	4.8°	24.5°
	NIS	3.1°	4.2°	26.2°
	SC	3.0°	3.6°	21.6°
	GM (pixel, $\sigma=4$)	2.3°	4.2°	23.2°
	Exemplar	2.3°	2.9°	19.4°
Color invariant	Retinex99 ($n=4$)	17.2°	16.9°	40.1°
	RetinexFrankle ($n=4$)	16.9°	16.5°	40.0°
	Proposed ($p=10$)	2.7°	4.8°	25.4°

to the measure of GE), and also lower than the learning-based GM(pixel) and GM(edge), but higher (worse) than the learning-based NIS and Exemplar. Table II also reveals that with equivalent mean angular error to all other methods, our model achieves the quite lower measure of max angular error.

3) *Gehler-Shi Color Checker Dataset*: This dataset contains 568 images of various indoor and outdoor scenes [62]. The ground truth illuminants were captured by a MacBeth color checker placed in each scene [15]. Because the images in the original dataset were generated from RAW data with automatic settings, they contain clipped pixels and include the effect of the camera's white balance. To avoid these artifacts, Shi reprocessed the raw data and created almost-raw 12-bit Portable Network Graphics images [62]. We use this reprocessed version to evaluate our proposed model.

Table III reports the angular error statistics of different methods over all the 568 images of this dataset. Note that as done with other algorithms, we also masked the color checker of each scene when executing our model. Table III clearly shows that on this dataset, the proposed method performs better than most of the methods compared, only worse than the best learning-based methods, e.g., GM (pixel) and Exemplar.

4) *Performance Dependency on Parameter p* : The first step of the proposed algorithm is to apply a grey world model (with p th-Minkowski norm) to the original image, where p is a free parameter in Eq. (4). From the results in Tables I, II and III, it seems that there is not a unique value of p that gives the best results on different datasets. Here we studied the dependency of the results on different values of p , and Fig. 9 shows such dependency on p on the SFU Lab Dataset and Gehler-Shi Dataset. In general, the measure of median angular error decreases clearly when p is increased from 1 to 6, but changes quite slightly when p is larger than 6. The profile of the median reproduction angular error shows a similar trend. To summarize, though no unique value of p can give the best results on different datasets, the proposed model is not quite sensitive to the parameter p when its value is not too small.

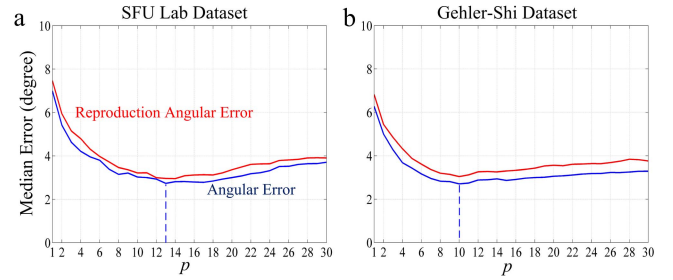


Fig. 9. The dependency of estimation error on the free parameter p in Eq. (4). (a) On the SFU Laboratory Dataset, (b) on the Gehler-Shi Dataset. The blue and red curves are respectively for the median angular error and median reproduction angular error. The vertical dashed lines denote the p values that give the best (lowest) median angular errors, as listed in Table I and III.

TABLE IV
THE PERFORMANCE OF THE PROPOSED MODEL WITH DIFFERENT RF SIZE SETTINGS ON THE GEHLER-SHI DATASET

Radius of Center/Surround /Subunit (in pixel)	Median	Mean	Max
1/3/1	2.7039°	4.8463°	25.3615°
2/6/2	2.7040°	4.8463°	25.3615°
3/9/3	2.7040°	4.8463°	25.3615°
4/12/4	2.7040°	4.8463°	25.3615°
5/15/5	2.7041°	4.8464°	25.3615°
10/30/10	2.7076°	4.8487°	25.3616°

5) *On the Parameter Setting for Receptive Field Size*: As mentioned earlier, we fixed the radius of the subunits, the surround and the center to 1, 3 and 1 pixels, respectively, for the three datasets tested above. To explain in more detail, we have mentioned earlier the physiological plausibility of setting their ratio as 1:3:1. Here we further analyzed the influence of other radius settings with the same ratio on the performance. Table IV compares the performance of the proposed model with different RF size settings on the Gehler-Shi Dataset. It is clear that the measure of angular error varies almost negligibly when the absolute radius values increase. That is why we fixed them to 1, 3 and 1 pixels, with a benefit of computational efficiency when testing on the above three datasets.

6) *Model Performance on Canonical Images*: We also undertook experiments to demonstrate our model's performance on canonical images (i.e., the images captured under white illumination). The canonical pictures were simply obtained by removing the illuminate ground truth from the color-biased images of the SFU Laboratory dataset, SFU Grey-ball dataset and Gehler-Shi dataset. The performance measures on the three datasets are listed in Table V, which reports the measures for the HC modulated cone output and the measures for the final output of the proposed model. Note that the HC modulated cone output computed Eq. (3) is equivalent to the output of GG method. Table V indicates that though the inputs are the canonical images, angular errors are unexpectedly non-zero. In addition, such non-zero angular errors are mainly from the GG-like processing by HC modulation and can not be decreased by the following ganglion cells. Taken together with Table I–III and V, we can find that the ganglion cells in our model work to refine the

TABLE V

THE MODEL PERFORMANCE ON PICTURES UNDER WHITE LIGHT SOURCES FOR THREE SINGLE-ILLUMINANT DATASETS. “BEFORE RGC” DENOTES THE HORIZONTAL MODULATED CONE OUTPUT, AND “FINAL OUTPUT” DENOTES THE OUTPUT OF THE GANGLION CELL

Dataset	Median Recovery Angular Error		Median Reproduction Angular Error	
	Before RGC	Final Output	Before RGC	Final output
SFU Lab	3.0867°	3.0867°	3.1958°	3.1958°
Grey-ball	3.2531°	3.2401°	3.2768°	3.2589°
Gehler-Shi	3.1167°	3.1169°	3.0421°	3.0421°

TABLE VI

COMPARISON ON A MULTI-ILLUMINANT DATASET [70]

Method	Laboratory (58)		Real-world (20)	
	Median	Mean	Median	Mean
DN	10.5°	10.6°	8.9°	8.8°
Gijssenij <i>et al.</i> with GE1	4.2°	4.8°	9.2°	9.1°
Gijssenij <i>et al.</i> with WP	4.2°	5.1°	3.8°	4.2°
MIRF with GE2	2.6°	2.6°	4.5°	4.9°
MIRF with WP	2.8°	3.0°	3.3°	4.1°
Proposed (p=0.7)	2.7°	3.2°	4.3°	5.2°

HC modulated cone output only when the HC modulation decreases, but not increases, the influence of color light source.

C. Performance on a Multi-Illuminant Dataset

Several recent studies focus on the color constancy for the scenes under multiple illuminants [68]–[70]. Here we also validated our model under such condition. The dataset used here contains 78 images under non-uniform illumination [70], including 58 indoor and 20 real-world scenes, accompanied with pixel-wise ground-truth illuminants. Considering that the images in this dataset are partially underexposed and have some black pixels, the size of the RF center is set as 1 pixel to minimize the effect of these black pixels.

Table VI reports the angular error statistics over all the 78 images of this dataset. Figure 10 shows several examples of the proposed method. Note that as mentioned before, our model directly corrects the color-biased images without requiring the explicit procedure of illuminant estimation. The estimated two-dimensional (2D) illuminants presented in Fig. 10 are just for the visualization and the computation of angular error using Eq. (11). In particular, the estimated 2D illuminant distribution for a scene was computed by dividing the original color-biased image by the image corrected by the proposed model, and the angular error for an image was then computed by averaging the angular error at each pixel. Figure 10 demonstrates that our model can correct the color-biased images quite close to the ground-truth by removing the non-uniform illumination.

The results of the two recent methods compared in Table VI are directly from [70]. In particular, the so-called Multi-Illuminant Random Field (MIRF) proposed in [70] estimates illuminants locally (using any simple method like GW or WP) and uses the random field to refine the illuminant color distribution by minimizing an energy function.

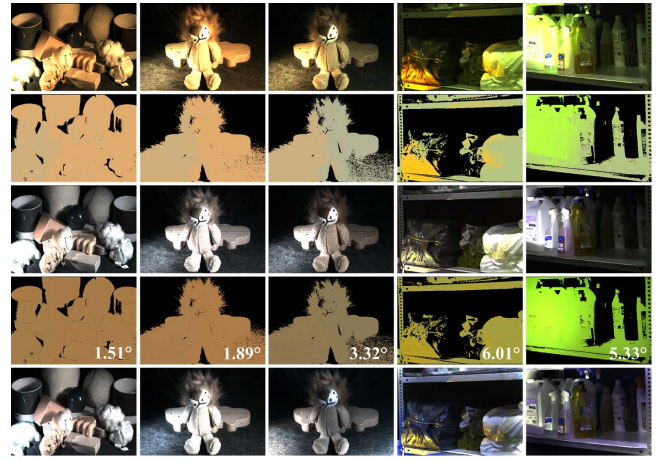


Fig. 10. Example images from the multi-illuminant dataset [70]. The first row: the original color-biased images; the second row: the illuminant ground truth, masked by the masks provided by the dataset; the third row: the images corrected with the illuminant ground truth; the fourth row: the illuminant distribution (indirectly) obtained by the proposed method (the number at the lower-right corner indicates the angular error averaged over the all pixels of the whole scene), also masked by the known masks; the fifth row: the images (directly) corrected by the proposed method. Note that the brightness of the images listed here is linearly increased for better visualization.

The method of Gijssenij *et al.* [68] includes a patch-based illuminant estimation (using, e.g., GW or WP), followed by a combination of local estimates using any clustering algorithm (e.g., k-means) under the assumption that the number of clusters is known (they assume a cluster number of 2 in [68]).

Table VI indicates that compared to the optimization-based Gijssenij’s model [68] and MIRF [70], our model can provide consistently competitive performance for both laboratory and natural multi-illuminant scenes. In general, the performances of MIRF and Gijssenij’s models are local estimator dependent. For example, MIRF obtains the best median measure (2.6°) on the laboratory subset when it uses GE2 as the local estimator, but on the real-world subset, the local estimator WP is required in order to obtain the best median measure (3.4°). As for the model of Gijssenij *et al.* [68], a better median measure (3.8°) than our model (4.3°) is obtained on the real-world subset when it uses WP as the local estimator, and worse median measures are obtained on the laboratory subset by Gijssenij *et al.* [68] with either WP or GE1 as local estimator.

IV. DISCUSSION AND CONCLUSION

In this work, we were concerned with the role of the retinal mechanism in computational color constancy, although the model is incomplete since converging evidence from different lines suggests a multistage mechanism involving from the retina to the higher visual cortexes [1], [28]. Even so, the results indicate that the proposed model arrives at reasonable accuracy. Our model is better than almost all the low-level information based algorithms considered here, and also quite competitive in comparison to the learning-based state-of-the-art approaches. The acceptable results of our model should be attributed to the disinhibitory effect in retinal ganglion cells (RGCs). Because of such effect, the RGC responses to the stimuli with similar intensity but different distribution

are disparate. The more dispersive the stimulus distribution is, the stronger the disinhibition among the subunits is, which then results in weaker surround inhibition and higher neuronal responses in turn [35], [37], [58]. Considering that the object surfaces are in general spatially coherent and the illumination is usually universal, the response sensitivity of RGCs to the object surfaces is higher than the sensitivity to the illumination.

As mentioned in the Introduction, there are several related works that attempt to achieve computational color constancy motivated by the visual mechanisms. Different from the existing models, our model makes a first attempt to introduce the horizontal-cell modulation and subunit-structured RF surround (i.e., the so-called non-classical receptive field, nCRF) for retinal color constancy, both of which are biologically supported. In particular, the subunit-structured surround is a novel mechanism based on the responsive properties of ganglion cells, which has the inherent flexibility in treating the influence of the local reflectances and the more extensive illumination on neuronal responses. Based on the spatial resolution of the disinhibition, the global or extensive low-frequency inputs, which are usually the illuminant components, are reduced due to the cone-specific gain control of horizontal cells and the stronger surround inhibition of ganglion cells. In contrast, the structural inputs, which are usually the object surfaces or edges, are preserved due to the relatively weaker surround suppression. In addition, we designed a novel way to adaptively determine the key parameter of inhibitory weight by finding the stable state of color opponent representation of surfaces in each channel.

Our model does not contain any explicit assumption about the illumination. Ganglion cells in the proposed model process information locally within their RFs (including the RF center and its surround). So our model is inherently suitable for the scenes under multiple illuminants. In addition, our model does not contain any explicit assumption of object surfaces, though the analysis of the SFU Laboratory Dataset preliminarily indicates that our current model performs better for certain scenes, e.g., the scenes that match more closely with the grey-world hypothesis. This is understandable considering that biologically the color constancy ability in the retina is incomplete [1], since there is no feedback modulation from visual cortexes and explicit learning mechanism at the level of retina. Facing scenes that include quite few colors or local structures, neither state-of-the-art computational models [3], [17], [18], [21] nor human observers [1] could reliably recover the true surface colors from changing illuminations.

In conclusion, this work proposed a color constancy model motivated by the retinal mechanisms of the biological visual system, from the cone-specific modulation by HCs to the color-opponent ganglion cells equipped with subunit-structured RF surround. Though the mammalian retina is far more complicated than we have thought and smarter than scientists have believed [42], [67], the simplified retinal network model proposed here provides valuable suggestions about the role of HC modulation and RGC surround disinhibition in color constancy. Additional retinal mechanisms, e.g., temporal and nonlinear properties of neuronal responsiveness, could perhaps be introduced to obtain a more realistic model.

REFERENCES

- [1] D. H. Foster, "Color constancy," *Vis. Res.*, vol. 51, no. 7, pp. 674–700, Apr. 2011.
- [2] S. D. Hordley, "Scene illuminant estimation: Past, present, and future," *Color Res. Appl.*, vol. 31, no. 4, pp. 303–314, Aug. 2006.
- [3] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2475–2489, Sep. 2011.
- [4] R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Amer. A*, vol. 21, no. 3, pp. 321–334, 2004.
- [5] G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Inst.*, vol. 310, no. 1, pp. 1–26, Jul. 1980.
- [6] E. H. Land, "The retinex theory of color vision," *Sci. Amer.*, vol. 237, no. 6, pp. 108–128, 1977.
- [7] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2207–2214, Sep. 2007.
- [8] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proc. Color Imag. Conf.*, 2004, pp. 37–41.
- [9] S. Gao, K. Yang, C. Li, and Y. Li, "A color constancy model with double-opponency mechanisms," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 929–936.
- [10] S. Gao, W. Han, K. Yang, C. Li, and Y. Li, "Efficient color constancy with local surface reflectance statistics," in *Proc. ECCV*, vol. 8690, 2014, pp. 158–173.
- [11] D. A. Forsyth, "A novel algorithm for color constancy," *Int. J. Comput. Vis.*, vol. 5, no. 1, pp. 5–35, Aug. 1990.
- [12] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *Int. J. Comput. Vis.*, vol. 86, no. 2, pp. 127–139, Jan. 2010.
- [13] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1209–1221, Nov. 2001.
- [14] G. D. Finlayson, S. D. Hordley, and I. Tasl, "Gamut constrained illuminant estimation," *Int. J. Comput. Vis.*, vol. 67, no. 1, pp. 93–109, Apr. 2006.
- [15] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, "Bayesian color constancy revisited," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.
- [16] W. Xiong and B. Funt, "Estimating illumination chromaticity via support vector regression," *J. Imag. Sci. Technol.*, vol. 50, no. 4, pp. 341–348, Jul. 2006.
- [17] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy with spatio-spectral statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1509–1519, Aug. 2012.
- [18] G. D. Finlayson, "Corrected-moment illuminant estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 1904–1911.
- [19] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Automatic color constancy algorithm selection and combination," *Pattern Recognit.*, vol. 43, no. 3, pp. 695–705, Mar. 2010.
- [20] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics," in *Proc. CVPR*, Jun. 2007, pp. 1–8.
- [21] A. Gijsenij and T. Gevers, "Color constancy using natural image statistics and scene semantics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 687–698, Apr. 2011.
- [22] H. R. V. Joze and M. S. Drew, "Exemplar-based colour constancy," in *Proc. BMVC*, 2012, pp. 1–12.
- [23] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, pp. 1–11, 1971.
- [24] E. H. Land, "Recent advances in retinex theory," *Vis. Res.*, vol. 26, no. 1, pp. 7–21, 1986.
- [25] J. A. Frankle and J. J. McCann, "Method and apparatus for lightness imaging," U.S. Patent 4 384 336, May 17, 1983.
- [26] J. McCann, "Lessons learned from mondrians applied to real images and color gamuts," in *Proc. Color Imag. Conf.*, 1999, pp. 1–8.
- [27] E. Provenzi, C. Gatta, M. Fierro, and A. Rizzi, "A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1757–1770, Oct. 2008.
- [28] H. Spitzer and S. Semo, "Color constancy: A biological model and its application for still and video images," *Pattern Recognit.*, vol. 35, no. 8, pp. 1645–1659, Aug. 2002.
- [29] H. Spitzer and Y. Barkan, "Computational adaptation model and its predictions for color induction of first and second orders," *Vis. Res.*, vol. 45, no. 27, pp. 3323–3342, Dec. 2005.

- [30] S. M. Courtney, L. H. Finkel, and G. Buchsbaum, "Network simulations of retinal and cortical contributions to color constancy," *Vis. Res.*, vol. 35, no. 3, pp. 413–434, Feb. 1995.
- [31] C.-H. Huang and C.-T. Lin, "Bio-inspired computer fovea model based on hexagonal-type cellular neural network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 1, pp. 35–47, Jan. 2007.
- [32] J. Héroult, "A model of colour processing in the retina of vertebrates: From photoreceptors to colour opposition and colour constancy phenomena," *Neurocomputing*, vol. 12, nos. 2–3, pp. 113–129, Jul. 1996.
- [33] H. Kolb, "How the retina works," *Amer. Sci.*, vol. 91, no. 1, pp. 28–35, 2003.
- [34] M. VanLeeuwen, C. Joselevitch, I. Fahrenfort, and M. Kamermans, "The contribution of the outer retina to color constancy: A general model for color constancy synthesized from primate and fish data," *Vis. Neurosci.*, vol. 24, no. 3, pp. 277–290, 2007.
- [35] C. Y. Li and Z. J. He, "Effects of patterned backgrounds on responses of lateral geniculate neurons in cat," *Experim. Brain Res.*, vol. 67, no. 1, pp. 16–26, Jun. 1987.
- [36] C.-Y. Li, X. Pei, Y.-X. Zhou, and H.-C. von Mitzlaff, "Role of the extensive area outside the X-cell receptive field in brightness information transmission," *Vis. Res.*, vol. 31, no. 9, pp. 1529–1540, 1991.
- [37] C.-Y. Li, Y.-X. Zhou, X. Pei, F.-T. Qiu, C.-Q. Tang, and X.-Z. Xu, "Extensive disinhibitory region beyond the classical receptive field of cat retinal ganglion cells," *Vis. Res.*, vol. 32, no. 2, pp. 219–228, Feb. 1992.
- [38] C.-Y. Li and W. Li, "Extensive integration field beyond the classical receptive field of cat's striate cortical neurons—Classification and tuning properties," *Vis. Res.*, vol. 34, no. 18, pp. 2337–2355, Sep. 1994.
- [39] C.-Y. Li, "Integration fields beyond the classical receptive field: Organization and functional properties," *News Physiol. Sci.*, vol. 11, no. 4, pp. 181–186, Aug. 1996.
- [40] F.-T. Qiu and C.-Y. Li, "Mathematical simulation of disinhibitory properties of concentric receptive field," *Acta Biophys. Sinica*, vol. 11, no. 2, pp. 214–220, 1995.
- [41] H. J. A. Dartnall, J. K. Bowmaker, and J. D. Mollon, "Human visual pigments: Microspectrophotometric results from the eyes of seven persons," *Proc. Roy. Soc. London. B, Biol. Sci.*, vol. 220, no. 1218, pp. 115–130, 1983.
- [42] R. H. Masland, "The neuronal organization of the retina," *Neuron*, vol. 76, no. 2, pp. 266–280, Oct. 2012.
- [43] C. Joselevitch, "Human retinal circuitry and physiology," *Psychol. Neurosci.*, vol. 1, no. 2, pp. 141–165, Jul./Dec. 2008.
- [44] M. VanLeeuwen, I. Fahrenfort, T. Sjoerdsma, R. Numan, and M. Kamermans, "Lateral gain control in the outer retina leads to potentiation of center responses of retinal neurons," *J. Neurosci.*, vol. 29, no. 19, pp. 6358–6366, 2009.
- [45] B. B. Lee, D. M. Dacey, V. C. Smith, and J. Pokorny, "Horizontal cells reveal cone type-specific adaptation in primate retina," *Proc. Nat. Acad. Sci. USA*, vol. 96, no. 25, pp. 14611–14616, 1999.
- [46] G. D. Finlayson, S. D. Hordley, and R. Xu, "Convex programming colour constancy with a diagonal-offset model," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2005, pp. III-948–III-951.
- [47] F. M. De Monasterio, P. Gouras, and D. J. Tolhurst, "Concealed colour opponency in ganglion cells of the rhesus monkey retina," *J. Physiol.*, vol. 251, no. 1, pp. 217–229, Sep. 1975.
- [48] D. M. Dacey and B. B. Lee, "The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type," *Nature*, vol. 367, pp. 731–735, Feb. 1994.
- [49] B. R. Conway *et al.*, "Advances in color science: From retina to behavior," *J. Neurosci.*, vol. 30, no. 45, pp. 14955–14963, 2010.
- [50] R. W. Rodieck and J. Stone, "Response of cat retinal ganglion cells to moving visual patterns," *J. Neurophysiol.*, vol. 28, no. 5, pp. 819–832, 1965.
- [51] E.-C. Christina and J. G. Robson, "The contrast sensitivity of retinal ganglion cells of the cat," *J. Physiol.*, vol. 187, no. 3, pp. 517–552, Dec. 1966.
- [52] I. Hisako and M. J. Wright, "The outer disinhibitory surround of the retinal ganglion cell receptive field," *J. Physiol.*, vol. 226, no. 2, pp. 511–544, Oct. 1972.
- [53] T. Hosoya, S. A. Baccus, and M. Meister, "Dynamic predictive coding by the retina," *Nature*, vol. 436, pp. 71–77, Jul. 2005.
- [54] D. Bölinger and T. Gollisch, "Closed-loop measurements of iso-response stimuli reveal dynamic nonlinear stimulus integration in the retina," *Neuron*, vol. 73, no. 2, pp. 333–346, Jan. 2012.
- [55] G. Vasserman, E. Schneidman, and R. Segev, "Adaptive colour contrast coding in the salamander retina efficiently matches natural scene statistics," *PLoS ONE*, vol. 8, no. 10, p. e79163, 2013.
- [56] G. Buchsbaum and A. Gottschalk, "Trichromacy, opponent colours coding and optimum colour information transmission in the retina," *Proc. Roy. Soc. London. B, Biol. Sci.*, vol. 220, no. 1218, pp. 89–113, 1983.
- [57] S. A. Baccus, "Timing and computation in inner retinal circuitry," *Annu. Rev. Physiol.*, vol. 69, pp. 271–290, Mar. 2007.
- [58] L. Chaoyi and Q. Fangtu, "Simulation of spatial transfer properties of cat retinal ganglion cell," *Acta Biophys. Sinica*, vol. 11, no. 3, pp. 395–400, 1994.
- [59] S. D. Hordley and G. D. Finlayson, "Reevaluation of color constancy algorithm performance," *J. Opt. Soc. Amer. A*, vol. 23, no. 5, pp. 1008–1020, 2006.
- [60] K. Barnard, L. Martin, B. Funt, and A. Coath, "A data set for color research," *Color Res. Appl.*, vol. 27, no. 3, pp. 147–151, Jun. 2002.
- [61] F. Ciurea and B. Funt, "A large image database for color constancy research," in *Proc. Color Imag. Conf.*, Nov. 2003, pp. 160–164.
- [62] L. Shi and B. Funt. (2010). *Re-Processed Version of the Gehler Color Constancy Dataset of 568 Images*. [Online]. Available: <http://www.cs.sfu.ca/~colour/data/>
- [63] B. Funt, J. McCann, and F. Ciurea, "Retinex in MATLAB," *J. Electron. Imag.*, vol. 13, no. 1, pp. 48–57, Jan. 2004.
- [64] A. Moore, J. Allman, and R. M. Goodman, "A real-time neural system for color constancy," *IEEE Trans. Neural Netw.*, vol. 2, no. 2, pp. 237–247, Mar. 1991.
- [65] P. A. Dufort and C. J. Lumsden, "Color categorization and color constancy in a neural network model of V4," *Biological*, vol. 65, no. 4, pp. 293–303, Aug. 1991.
- [66] S. M. Courtney, L. H. Finkel, and G. Buchsbaum, "A multistage neural network for color constancy and color induction," *IEEE Trans. Neural Netw.*, vol. 6, no. 4, pp. 972–985, Jul. 1995.
- [67] T. Gollisch and M. Meister, "Eye smarter than scientists believed: Neural computations in circuits of the retina," *Neuron*, vol. 65, no. 2, pp. 150–164, Jan. 2010.
- [68] A. Gijsenij, R. Lu, and T. Gevers, "Color constancy for multiple light sources," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 697–707, Feb. 2012.
- [69] H. R. V. Joze and M. S. Drew, "Exemplar-based color constancy and multiple illumination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 860–873, May 2013.
- [70] S. Beigpour, C. Riess, J. van de Weijer, and E. Angelopoulou, "Multi-illuminant estimation with conditional random fields," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 83–96, Jan. 2014.
- [71] G. D. Finlayson and R. Zakizadeh, "Reproduction angular error: An improved performance metric for illuminant estimation," in *Proc. BMVC*, 2014, pp. 1–11. [Online]. Available: <http://dx.doi.org/10.5244/C.28.70>
- [72] R. Zakizadeh and G. D. Finlayson, "The correlation of reproduction and recovery angular errors for similar and diverse scenes," in *Proc. Color Imag. Conf.*, 2015, pp. 196–200.



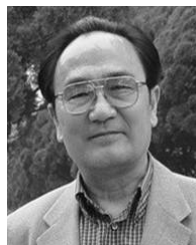
Xian-Shi Zhang received the M.Sc. degree in automation engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2008, where he is currently pursuing the Ph.D. degree in biomedical engineering. His research interests include visual mechanism modeling and image processing.



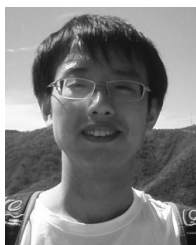
Shao-Bing Gao received the M.Sc. degree in biomedical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2013, where he is currently pursuing the Ph.D. degree in biomedical engineering. His research interests focus on the visual mechanism modeling and biologically inspired image processing.



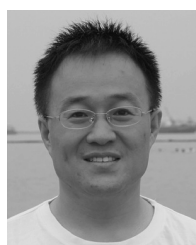
Ruo-Xuan Li received the M.Sc. degree in biomedical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2015, where she is currently pursuing the Ph.D. degree in biomedical engineering. Her research interests focus on the image processing.



Chao-Yi Li received the degrees from Chinese Medical University, Shenyang, China, in 1956, and Fudan University, Shanghai, China, in 1961. He became an Academician of the Chinese Academy of Sciences, Beijing, China, in 1999. He is currently a Professor with the University of Electronic Science and Technology of China, Chengdu, China, and the Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China. His research interest is visual neurophysiology.



Xin-Yu Du received the Ph.D. degree in biomedical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2011. He is currently a Researcher with the Infrastructure Inspection Institute, China Academy of Railway Sciences, Beijing, China. His research interests focus on the image processing.



Yong-Jie Li (M'14) received the Ph.D. degree in biomedical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2004. He is currently a Professor with the Key Laboratory for Neuroinformation, Ministry of Education, School of Life Science and Technology, UESTC. His research interests include visual mechanism modeling, image processing, and intelligent computation.