BIO-CC: Biologically inspired color constancy

Oguzhan Ulucan
oguzhan.ulucan@uni-greifswald.de
Diclehan Ulucan
diclehan.karakaya@uni-greifswald.de
Marc Ebner
marc.ebner@uni-greifswald.de

Institut für Mathematik und Informatik
Universität Greifswald
Greifswald, Germany

Abstract

The human visual system unconsciously discounts the illumination of the scene. This is called color constancy. Color constancy and color illusion perception are two phenomena that might help us reveal unknown mechanisms of human perception. A computational color constancy model based on biological findings, which can be fooled by color illusions similar to the human visual system might be the key to figure out how the brain arrives at a color constant descriptor. Therefore, in this study, a color constancy model following the hierarchical structure of the color perception mechanism of the human vision system is proposed. This system can also reproduce the response of the human visual system to color assimilation illusions. Focal and peripheral vision, the retinotopy structure, the response of the double-opponent cells, and the saliency map of the visual cortex are taken into account during modelling. Even though the algorithm is not fully optimized yet, according to the experimental results the proposed method demonstrates competitive performance on a well-known color constancy benchmark.

1 Introduction

There are many different phenomena that occur during visual processing. Understanding why color illusions show up in certain situations might be the key to unravelling how visual color processing works. Understanding the underlying mechanisms of the visual cortex can help us to explain other processes in the brain. Among all the visual attributes, color vision might be the *easiest* to understand as stated by Semir Zeki [33]. One of the interesting aspects of color processing is the difference between the colors perceived by the human visual system and the actual physical reflectance in certain situations. The human brain tries to estimate reflectance. This is what we perceive as the "color" of an object. The perceived color remains constant regardless of the illumination that is present in a scene [7]. This phenomenon is called *color constancy*. It is performed unconsciously by the human visual system. Reproducing the same behavior in an artificial system is a challenging task.

The image of the scene I at position (x, y) can be modeled as [7],

$$I(x,y) = \int_w R(x,y,\lambda)E(x,y,\lambda)S(\lambda)d\lambda \tag{1}$$

where, $R(x,y,\lambda)$ is the reflectance of the object, $E(x,y,\lambda)$ is the irradiance falling onto the object, $S(\lambda)$ is the sensor response characteristics of the capturing device for red, green and blue sensor spectral sensitivities, and λ is the wavelength of the visible spectrum w.

The color vector $\mathbf{L}$ of the global light source illuminating the scene can be defined as,

$$\mathbf{L} = [l_R\ l_G\ l_B]^T = \int_w E(x,y,\lambda)S(\lambda)d\lambda. \tag{2}$$

Computational modelling of color constancy is important to understand how the human visual system works [8]. Since its first description by Monge in 1789, color constancy has been extensively studied [25]. In the 19th century, Helmholtz described the phenomenon as *"discounting the illuminant"*, which is an expression still widely used in the field of color constancy [31]. Later on in the 20th century, Land made a great contribution to color constancy with his famous experiments on color vision and natural images [21, 22]. Both in the 20th and 21st centuries with the improvements in computer vision, numerous color constancy algorithms have been proposed.

One of the milestone studies of color constancy is carried out by Land and is known as the white patch Retinex algorithm [23]. The algorithm identifies the maximum response of each image channel separately to find the illuminant of the scene. The gray world assumption introduced by Buchsbaum is another well-known color constancy method [3]. Like the Retinex method, the gray world algorithm also has a correspondence in the human visual system, since illumination might be discounted based on the space-average color during color perception [6, 24]. The gray world method estimates the color constant descriptor of the scene by assuming that the average reflectance of the scene is achromatic. The shades of gray algorithm proposed by Finlayson and Trezzi makes the assumption that the mean of pixels raised to a certain power is gray [9]. The gray-edge hypothesis and weighted gray-edge algorithm build up on the gray world assumption and underline the fact that the gradient information of the scene is an important cue for illumination estimation [13, 29]. Principal component analysis (PCA) based color constancy shows the effectiveness of increasing the gradient information by shuffling the sub-blocks of the image and the contribution of the brightest and darkest pixels for illumination estimation [5]. The local surface reflectance statistics based color constancy method relies on the linear image formation model and makes use of the biological findings about the feedback modulation mechanism in the eye [10]. The physiologically inspired color constancy method proposed by Gao *et al.* benefits from physiological findings on color information processing of the human visual system [11]. The algorithm is based on the observation that the color of the light source coincides with the color distribution of the double-opponent cells.

Not only traditional approaches, but also machine learning algorithms are applied to address the problem of color constancy [14]. These algorithms mostly outperform the traditional methods on well-known benchmarks, however it is known that these approaches are highly dependent on the training data and take the camera sensitivity specifications as prior information. Therefore, in real-world applications where these algorithms have to remove a color cast, which has a spectral distribution that deviates significantly from the ones seen during training, and operate with unknown camera specifics they face a significant challenge [12]. Thus, traditional algorithms generally outperform learning-based approaches

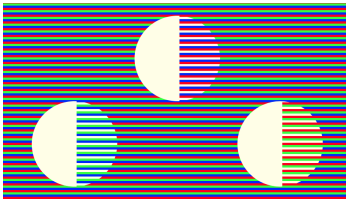

Figure 1: Color assimilation illusion [1].

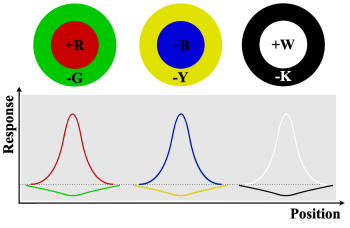

Figure 2: Center-surround opponent cells.

when images with unknown statistical distributions and unknown hardware specifications are of interest.

As detailed above, there are numerous color constancy algorithms which try to obtain a canonical image from a color biased scene. Most of these methods only aim at estimating the illuminant without giving any explanation of how the algorithm corresponds to the human visual system. Moreover, these algorithms do not consider how the human visual system interprets color illusions, such as color assimilation illusion where the perceived color of the target shifts towards that of its local neighbours (Fig. 1) [4]. However, alongside color constancy, color illusions can also reveal the mechanisms of color perception [2]. Therefore, investigating both phenomena together can help the researchers in the field of computer vision to design models that are one step closer to mimic the human visual system [15]. In other words, a computational model inspired by the biological mechanisms of the human visual system may help us to understand how the brain processes visual information. Therefore, in this paper, a novel color constancy algorithm based on the biological findings of color perception, which provides not only a global estimate of the illumination for color constancy, but also reproduces the behaviour of the human visual system on color assimilation illusions by making use of pixel-wise estimations is proposed. This is the first color constancy study that also considers the color illusion perception phenomenon. Furthermore, to the best of available knowledge, it is the first study that uses the focal and peripheral vision, and the retinotopy structure during modeling.

This paper is organized as follows. Section 2 gives a brief summary of the color perception process in the visual system. Section 3 introduces the proposed method. Section 4 presents the experimental results. Section 5 provides a brief conclusion and future directions.

2 Color Perception Mechanism

Color has always been an important que. Due to its importance, the color perception mechanism, the processing of raw data reaching the eyes, the procedure of arriving at the color sensation and the derivation of how a color constant descriptor might be computed, has been studied for centuries. One of the remarkable advances in the field of color vision is made by Thomas Young who equated the wavelengths with colors and described the photoreceptors as the particles in the retina [33]. Unfortunately, the importance of Young's findings were not realized until Helmholtz created the *trichromatic color theory* based on the observations of his color matching experiments and Young's work. However, this theory could not explain the effects of color blindness and afterimages, therefore Hering proposed the *opponent color theory*, which explains the colors as opponent pairs. Decades later, Hurvich

and Jameson [17], combined these two theories to form a theory that strongly coincides with the findings in the human visual system. The retina processes the incoming signals as in the trichromatic theory, while the receptive fields of the retinal ganglion layer, lateral geniculate nucleus (LGN), primary visual cortex (V1) and higher visual cortex (e.g. V2 and V4) are stimulated by the opponent colors [11].

Although there are many gaps to be filled to fully comprehend the biological color perception mechanisms, it is known that the color vision process starts at the earlier stages of the visual system [7]. The reflected light from the objects hitting the retinal receptors start the process of color perception. These signals are hierarchically processed by retinal ganglion cells, LGN, V1 and in the end V4 [33].

Vision starts when the light reflected from the objects enters the eye. There are two functions in eye sight; the focal vision focuses on the salient regions, and the peripheral vision is associated with the attentions outside the point of fixation [33]. While the image in the focal vision is clear, the image in the peripheral vision is locally disordered, i.e. blurred [18].

In the retinal layer, the photoreceptors responsible for color vision, i.e. cones, are stimulated by three different wavelengths. According to their spectral sensitivity, these cones can be grouped into three as, long-wavelength (L-cone), medium-wavelength (M-cone) and short-wavelength (S-cone) cones [33]. These cones are associated with red, green and blue colors, respectively. As in the trichromatic color theory, the color information is processed by L-, M- and S-cones in the retina. The photoreceptors are non-uniformly distributed in the retina. Neighboring photoreceptors receiving closely spaced points in the visual space fall onto the retina as an image. This structure is known as *retinotopy* and one can associate it with the pixel-wise neighboring relationships in images [26].

The information output by the cones is fed through several retinal cells, such as amacrine cells, to the retinal ganglion and afterwards to the LGN layers, where the retinotopy structure is well preserved [26]. The color sensitive cells in these layers are stimulated by single-opponent colors and the color information is coded as red-green (R-G), blue-yellow (B-Y) and white-black (W-K) opponency within their receptive fields (Fig. 2) [7]. These receptive fields have diverse structures [11], with different sizes, i.e. the size of the surround function is approximately 2 to 5 (in diameter) times larger than the size of the center function [34]. In the single-opponent cells, there might be a $+Red - Green$ structure in one cell, while in another cell this structure might be reversed [27]. The signs "+" and "−" are representing the excitation (on) and inhibition (off) of the cells. It is also worth to mention here that, the cells of ganglion and LGN have similar receptive field properties, which respond to color contrast and uniformly colored areas.

The output of the LGN layer reaches V1, where the complete map of the visual field covered by the eyes, i.e. "pixel-wise neighboring relationship", is sustained [26, 33]. In V1, the color sensitive cells are present widely as a combination of the single-opponent cells, namely double-opponent cells, with a center-surround structure [11]. Furthermore, maximum-like thresholding operations are performed in V1 and it also behaves as a gateway that distributes the incoming signals to the corresponding higher cortical areas, particularly the color information is segregated to V4 [20, 33]. Moreover, it is indicated that V1 guides the visual attention by creating a saliency map of the visual field [35].

Semir Zeki discovered that color information is mainly processed in V4, where also color constant cells have been observed [33]. Another key finding in V4 for understanding the color constancy mechanism is that the retinotopy structure preserved in the retina, LGN and V1 starts to degrade in this area and the receptive field size of V4 cells are larger compared to the previous layers.

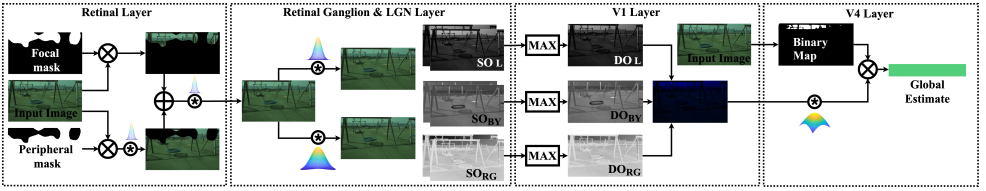

Figure 3: Flowchart of Bio-CC.

Despite all the known facts about the color perception mechanism, it is still a mystery how the brain arrives at color constant descriptors, and how the color illusions fool the human visual system [8].

3 The Proposed Method

As discussed in Sec. 1, there are numerous studies, which rely on different approaches to provide a solution to the ill-posed nature of color constancy. The studies building upon the findings of the human visual system, i.e white patch Retinex and gray world, still have a great impact in the field of color constancy and many researchers use these methods as building blocks of their algorithms. Thereupon, in this study a biologically inspired color constancy model is proposed (Fig. 3).

The proposed algorithm has a hierarchical order as the human visual system. The main idea is to find the color constant descriptor of the scene by taking into account the color vision mechanism of the human visual system; the focal and peripheral vision, retinotopy structure of the layers, the response of the double-opponent cells, and the saliency maps created in the visual cortex, which are discussed in Sec. 2. To the best of available knowledge, this is the first computational color constancy study, which takes the focal and peripheral vision, and the retinotopy structure of the layers into account.

The proposed algorithm takes in a linearized RGB image in order to obtain a linear relationship between pixels. It is assumed that each pixel in each image channel corresponds to a photoreceptor of the retina hence, since each pixel in the image has neighboring pixels, the retinotopy structure of the adjacent photoreceptors is simulated.

After the cones receive the input signals, to obtain an output similar to the focal and peripheral vision, the points of interest in the input image are detected by making use of salient regions. Since this study does not aim to develop a salient region detection method, the technique proposed by Hou *et al.* is used to determine the salient areas [16]. The method relies on the computation of the discrete cosine transform coefficients and detects the points of interest similar to the human visual system. In the proposed method, the detected salient regions correspond to the area observed by the focal vision (m), whereas the remaining parts are associated with the peripheral vision ($\bar{m}$). The areas of the image overlapping with the regions observed by the peripheral vision are blurred via convolving them with a Gaussian kernel, since peripheral areas are seen more blurry than the regions perceived by the focal vision as stated in Sec. 2. After simulating the focal and peripheral vision, the image is convolved with the same kernel again since it is known that the image falling onto the retina is blurred [30]. This process can be represented as follows,

$$\widetilde{I}(x,y) = (I(x,y)\cdot m + (I(x,y) * \mathcal{G}(x,y;\sigma))\cdot \bar{m}) * \mathcal{G}(x,y;\sigma) \tag{3}$$

where, $*$ denotes the convolution operation, and $\mathcal{G}(x,y;\sigma)$ is the 2D Gaussian kernel simulating the receptive field function and it is taken as $\frac{1}{2\pi\sigma^2} exp(-\frac{x^2+y^2}{2\sigma^2})$. To preserve the retinotopy structure, the size and the scaling factor σ of the kernel are taken as 3×3 and 1, respectively.

Afterwards, for the sake of simplicity the output of the cones is directly sent to the retinal ganglion layer without considering the sub-layers. Moreover, since the receptive field properties of the retinal ganglion and the LGN layer are similar, they are considered as a single layer. In this stage, the single-opponent colors are created by relying on the center-surround receptive field structure of the single-opponent cells as in Fig. 2. The incoming signals from the retinal layer are convolved with two different Gaussian functions with distinct kernel sizes and scaling factors. Since the receptive fields approximately double in each layer, while the scaling factor of the center function is taken as twice the σ value of the retinal layer, the σ value of the surround function is chosen as 3 times the σ value of the center function based on biological findings [34]. Moreover, the kernel size of the surround function is taken slightly larger than the kernel size of the center function to simulate the center-surround structure. Hence, the responses, which will be used to create the single-opponent cells are formed for the center (c) and surround (s) signals individually as follows,

$$\widetilde{I}_c(x,y) = \widetilde{I}(x,y) * \mathcal{G}(x,y;\sigma_c) \qquad \widetilde{I}_s(x,y) = \widetilde{I}(x,y) * \mathcal{G}(x,y;\sigma_s). \tag{4}$$

Subsequently, these two signals are combined to form the single-opponent cells for every spatial location (x,y) similar to the study of Ebner as follows [7],

$$SO_{RG} = \frac{(\widetilde{I}_{c_R} - \widetilde{I}_{s_G})}{\sqrt{2}} \quad SO_{YB} = \frac{(\widetilde{I}_{c_Y} - \sqrt{2}\widetilde{I}_{s_B})}{\sqrt{6}} \quad SO_{\mathcal{L}+} = \frac{(\widetilde{I}_{c_R} + \widetilde{I}_{c_G} + \widetilde{I}_{c_B})}{\sqrt{3}} \tag{5}$$

where, SO_{RG} is the red-on/green-off, SO_{YB} is the yellow-on/blue-off and $SO_{\mathcal{L}+}$ is the brightness-sensitive cells. Similar calculations are carried out for SO_{GR}, SO_{BY}, and $SO_{\mathcal{L}-}$. The yellow channel is constructed by combining the red and green channels. The single-opponent cell computations above correspond to the difference of Gaussians, which is known to be associated with the receptive fields of ganglion cells [32].

Then, the formed single-opponent signals are fed to V1, where the double-opponent cells are widely present and pooling operations are observed as mentioned in Sec. 2. Thereupon, in this layer the double-opponent cells (DO) are formed by combining the corresponding single-opponent cells via taking the absolute maximum response for every pixel individually as follows,

$$DO_{RG}(x,y) = max(SO_{RG}(x,y), SO_{GR}(x,y)) \tag{6}$$

where, max represents the operation of selecting the absolute maximum value between channels. For DO_{YB} and $DO_{\mathcal{L}}$ the same operation is performed with corresponding cells.

After the double-opponent signals are formed, they are fed to the last layer, where the color constant descriptor of the scene is estimated. As stated in Sec. 2, in V4 the retinotopy structure is degraded and the receptive fields in V4 are larger than in any other layer. As aforementioned, the retinotopy structure can be explained as the preservation of the local spatial information. The degradation of this structure can be simulated by assigning weights to pixels outside the local neighborhood. These weights can be controlled with the scaling

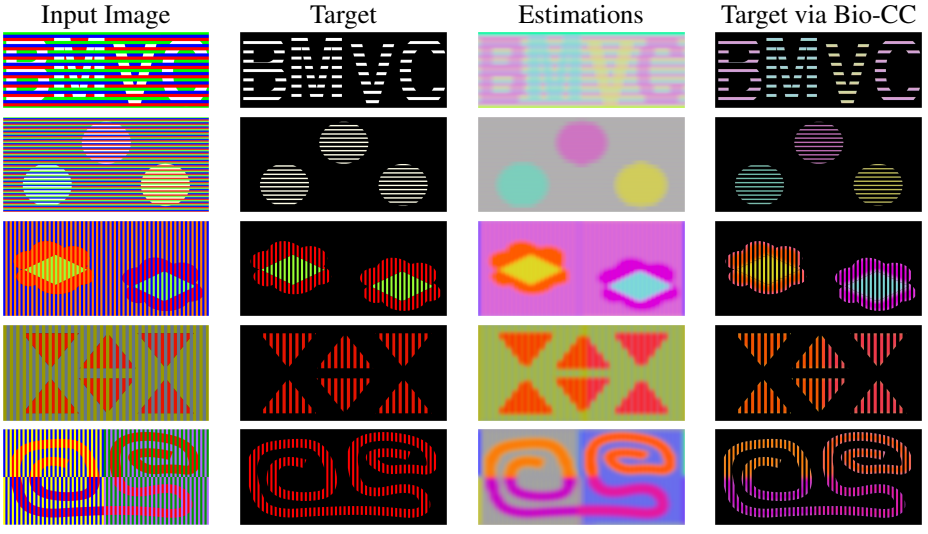

Figure 4: The response of Bio-CC on color assimilation illusions. Although, the target has the same color, due to the context its colors is shifted towards its neighbors' colors.

factor of the receptive field function and increasing it leads to the consideration of distant pixels. Hence, to simulate the retinotopy degradation and the characteristics of the receptive fields in V4, the scaling factor and the size of the receptive field function are increased. Then, the double-opponent image is convolved with this receptive field function to find the pixel-wise estimations in double-opponent space as follows,

$$L_{DO}(x,y) = DO(x,y) * \mathcal{G}(x,y;\sigma) \tag{7}$$

where, L_{DO} is the pixel-wise estimations in double-opponent space, the size of the kernel is approximately 8% of the image and σ is 32 times larger than the σ in the retinal layer.

Afterwards, $L_{DO}(x,y)$ is converted from double-opponent color space to RGB color space by using the computations in the study of Ebner and L_{RGB} is formed [7].

The pixel-wise estimations in L_{RGB} represent the behaviour of the human visual system on color assimilation illusions. Examples of processing color assimilation illusions via Bio-CC are given in Fig. 4. Although the true reflectance values of the target region are the same, due to the context they are perceived by the human observer as if they have different colors. After removing the context, it is clear that the reflectance values of the targets are identical.

In order to provide a global illumination estimate to discount the illuminant in the scenes, a saliency map obtained from the pixels closest to white are used to determine the most informative pixels. The biological reason using these pixels is the observation that the human visual system might be discounting the illuminant by taking the regions having the highest luminance in a scene into account [6]. This observation has also a correspondence in digital photography and can be explained by a simple example [7]. Let us assume that a picture of a room illuminated with blue light, which has white walls and contains colored objects is captured. The illumination value of the light source can be estimated easier from the white wall rather than the colored objects, since the white wall illuminated with blue light will be measured by the camera sensors as blue. Hence, even if the scenes are very complex, the light

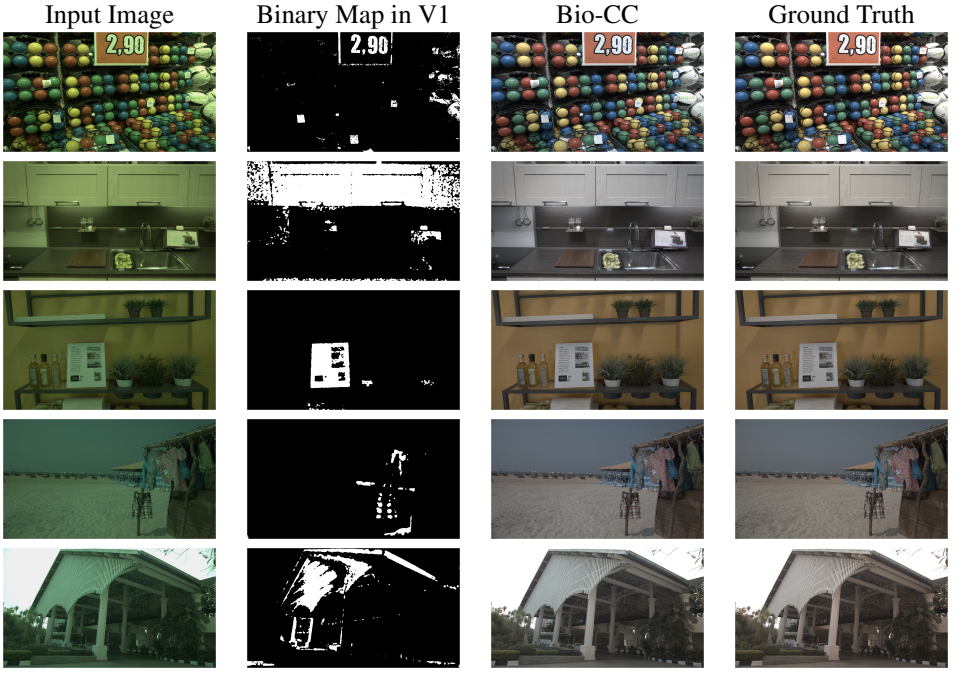

Figure 5: The result of Bio-CC. Gamma correction is applied for better visualization.

vector can be approximated by using the pixels closest to white. In order to find these pixels, a simple yet effective approach is followed in the V1 layer of this study. Firstly, 2% of the darkest and the brightest pixels are clipped, since over- and under-exposed pixels in scenes cause noise while estimating the illumination. Then, a temporary illumination estimate $\mathbf{L}_{temp}$ is found by taking the mean values of each color channel of the clipped image separately. Subsequently, a temporary white balanced image I_{temp} is obtained by scaling its RGB values according to $\mathbf{L}_{temp}$. Afterwards, the pixels closer to white are determined by checking the angular error between the pixels of I_{temp} and the white vector, $[1,1,1]$. The pixels having an angular error less than 5 are considered as the pixels closest to white. A binary map is extracted by using the spatial locations of these pixels. Then, the pixels of L_{RGB} coinciding with this map are averaged in order to obtain the global illuminant estimate of the scene (Fig. 5). Subsequently, the canonical image is obtained by dividing the intensity values of the color-casted image into the global illuminant estimates of BIO-CC.

Readers may refer to the supplementary material for more visual results of Bio-CC and for explanations about the opponent color space conversions, the parameter selection process of the V4 layer, and the angular error computation.

4 Experiments and Discussions

Bio-CC is compared with the following methods, gray world (GW) [3], white patch Retinex (max-RGB) [23], shades of gray (SoG) [9], 1^{st} and 2^{nd} order gray-edge (GE) [29], weighted gray-edge (WGE) [13], double-opponent cells based color constancy (DOCC) [11], PCA

Figure 6: Comparison of Bio-CC with best performing methods. From left-to-right: input image, LSRS, PCA-CC, Bio-CC, and ground truth. The angular error of each method is provided on the bottom-right side of the image. Gamma correction is applied for better visualization. Note that, more examples are provided in the supplementary material.

	Camera-Invariant					Canon					Nikon					Sony					
	ΔE	Mean	Median	B-25%	W-25%	ΔE	Mean	Median	B-25%	W-25%	ΔE	Mean	Median	B-25%	W-25%	ΔE	Mean	Median	B-25%	W-25%	run time
GW	4.22	4.91	3.88	0.95	10.59	3.93	4.52	3.55	0.85	9.83	4.15	5.29	4.19	1.03	11.49	4.62	4.77	3.84	0.98	10.06	0.14
max-RGB	10.15	11.01	13.16	1.81	19.44	11.54	13.41	17.64	2.37	20.99	8.74	10.02	11.31	1.56	17.65	10.69	9.98	11.52	1.83	16.76	**0.13**
SoG	5.30	5.51	4.16	0.97	12.29	5.82	6.16	4.26	1.05	14.30	4.64	5.17	3.82	0.96	11.46	5.71	5.33	4.41	0.92	11.34	0.23
1st GE	5.80	6.09	4.23	0.96	14.26	6.34	6.93	4.34	0.94	16.87	4.91	5.47	3.65	0.87	12.87	6.49	6.14	4.79	1.13	13.35	0.38
2nd GE	6.09	6.41	4.49	1.04	14.73	6.70	7.33	4.76	1.04	17.29	5.20	5.79	4.02	0.93	13.38	6.69	6.34	5.01	1.20	13.71	0.42
WGE	5.64	6.00	3.64	0.81	14.90	6.18	6.86	3.55	0.79	17.81	4.72	5.29	3.19	0.72	13.16	6.36	6.13	4.41	0.98	14.12	2.63
DOCC	6.65	7.19	4.67	0.81	16.98	7.13	8.20	5.00	0.78	19.73	5.29	6.07	3.62	0.72	15.03	8.05	7.72	6.27	1.00	16.26	0.43
PCA-CC	4.14	4.47	**3.03**	**0.69**	10.64	4.45	4.81	3.11	**0.71**	11.87	3.52	**4.09**	**2.76**	**0.67**	**9.82**	4.67	4.65	**3.42**	**0.71**	10.50	0.16
LSRS	3.82	4.17	3.42	0.98	**8.61**	3.69	3.94	3.08	1.01	8.16	3.63	4.33	3.59	0.99	8.97	**4.18**	**4.17**	3.54	0.94	**8.55**	0.14
Bio-CC	**3.55**	**4.14**	3.05	0.76	9.42	**3.17**	**3.68**	**2.85**	0.75	**8.00**	**3.27**	4.22	2.92	0.77	9.88	4.32	4.49	3.45	0.73	10.00	1.61

Table 1: Statistical results of the methods. For each metric the best result is highlighted. Moreover, the average run time (in seconds) of each method is provided in the last column.

based color constancy (PCA-CC) [5], and color constancy with local surface reflectance estimation (LSRS) [10], which are all briefly explained in Section 1. The experiments are conducted on an Intel i7 CPU @ 2.7 GHz Quad-Core 16 GB RAM machine. The codes of the algorithms employed for comparison are reached from the official webpages of the authors. No optimization is carried out on the methods and they are used in their default settings.

Bio-CC is benchmarked on the INTEL-TAU Dataset, which is one of the largest color constancy datasets containing diverse illumination types [19]. This dataset contains 7022 images, which are taken with 3 different cameras, Canon 5DSR and Nikon D810 DSLR, and Mobile Sony IMX135. All the images contain one dominant illumination color in the scene. Since images are captured with 3 different cameras whose camera sensitivities differ, 4 distinct test sets are created from the INTEL-TAU Dataset. While one test set, namely *Camera-Invariant*, contains all the images in the dataset, the other 3 test sets, namely *Canon, Nikon, and Sony*, contain images captured with corresponding devices.

During benchmarking, two well-known metrics in the field of color constancy, ΔE 2000 and angular error, are employed [7]. In color constancy, it is important to maintain the color information after obtaining the canonical image. ΔE provides analysis demonstrating the color difference between the ground truth and color corrected image [28]. Scores closer to 0 indicate a better outcome, where the scores in range $[0,4)$ can be interpreted as an unnoticeable difference between the ground truth and output image is present for a human observer. The ΔE scores and the mean, the median, the mean of the best 25% and the mean of the worst 25% of the angular error are reported in Table 1.

As it can be observed in Table 1, Bio-CC presents state-of-the-art performance and outputs the best mean angular error among the algorithms for the entire dataset. The median angular error of Bio-CC is less than the mean angular error for all test sets, which demon-

strates that Bio-CC tends to produce more outcomes closer to the best cases rather than the worst cases. Moreover, apart from the mean angular error, in color constancy studies it is important to reduce the worst 25% of the angular error, since it is desired to obtain results, where the worst cases do not deviate significantly from the mean angular error. Bio-CC produces consistent results for each set in terms of worst 25% of the angular error and outperforms most of the algorithms in Table 1 in this metric. In addition to the angular error, Bio-CC also provides the best *Delta*E scores in three of the test sets. ΔE scores mostly less than 4 are produced, which indicates that most of the color difference between the ground truth and output images is unnoticeable.

Alongside the statistical analysis, visual comparisons are also provided in order to show the performance of Bio-CC (Fig. 6). In indoor scenes Bio-CC tends to produce low angular errors and the color biased images are successfully white-balanced. In outdoor scenes again visually pleasing images are produced by Bio-CC, yet the angular error is higher compared to indoor scenes in case a sky region is present in the image. Such regions are a well-known challenge in color constancy studies and as future work of this study it is considered to modify the saliency map obtained in the V1 layer to reduce the contribution of these pixels to the global illuminant estimate.

5 Conclusion

Color vision might be the *simplest* visual attribute that may help us to understand the underlying mechanisms of the brain. A computational color constancy model based on biological findings, which can also reproduce the behaviour of the human visual system on color illusions might help us to unravel the mechanism of color vision. Therefore, in this study a simple yet effective learning-free color constancy model Bio-CC is developed. This model relies on the hierarchical structure of the human visual system. Bio-CC aims at finding the color constant descriptor of the scene by taking into account the focal and peripheral vision, retinotopy structure of the layers, the response of the double-opponent cells, and saliency maps formed in the visual cortex. The focal and peripheral vision, and retinotopy structure are used for the first time in the field of computational color constancy. While pixel-wise estimates of the image are used for the color illusions, a map created in the V1 layer by using the pixels closest to white is employed to estimate the global illuminant of the scene. Bio-CC presents state-of-the-art performance, while surpassing most of the existing color constancy methods.

To the best of available knowledge, this is the first study introducing a computational model that is able to both discount the illuminants in scenes and replicate the behaviour of the human visual system on color assimilation illusions. As future work, the proposed method will be modified for different color illusions and multi-illuminant scenes.

6 Acknowledgement

The authors would like to thank Michael Bach and Akiyoshi Kitaoka for giving permission to use their color illusions.

References

[1] M. Bach. Color assimilation illusions, Last accessed: 26.07.2022. michaelbach.de/ot.

[2] M. Bach and C. M. Poloschek. Optical illusions. *Adv. Clin. Neuroscience Rehabil.*, 6 (2):20–21, 2006.

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

[4] X. Cerda-Company, X. Otazu, N. Sallent, and C. A. Parraga. The effect of luminance differences on color assimilation. *J. Vision*, 18(11):10–10, 2018.

[5] D. Cheng, D. K. Prasad, and M. S. Brown. Illuminant estimation for color constancy: Why spatial-domain methods work and the role of the color distribution. *J. Opt. Soc. Amer. A*, 31(5):1049–1058, 2014.

[6] M. Ebner. A parallel algorithm for color constancy. *J. Parallel Distrib Comput.*, 64(1): 79–88, 2004.

[7] M. Ebner. *Color Constancy, 1st ed.* Wiley Publishing, ISBN: 0470058299, 2007.

[8] M. Ebner. How does the brain arrive at a color constant descriptor? In *Proc. Int. Symp. Brain Vision Artif. Intell.*, 2007.

[9] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Proc. Color Imag. Conf.*, 2004.

[10] S. Gao, W. Han, K. Yang, C. Li, and Y. Li. Efficient color constancy with local surface reflectance statistics. In *Proc. Eur. Conf. Comput. Vision*, 2014.

[11] S.-B. Gao, K.-F. Yang, C.-Y. Li, and Y.-J. Li. Color constancy using double-opponency. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(10):1973–1985, 2015.

[12] S.-B. Gao, M. Zhang, C.-Y. Li, and Y.-J. Li. Improving color constancy by discounting the variation of camera spectral sensitivity. *J. Opt. Soc. Amer. A*, 34(8):1448–1462, 2017.

[13] A. Gijsenij, T. Gevers, and J. Van De Weijer. Physics-based edge evaluation for improved color constancy. In *Proc. Comput. Vision Pattern Recognit.*, 2009.

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

[15] A. Gomez-Villa, A. Martin, J. Vazquez-Corral, and M. Bertalmío. Convolutional neural networks can be deceived by visual illusions. In *Proc. Comput. Vision Pattern Recognit.*, pages 12309–12317, 2019.

[16] X. Hou, J. Harel, and C. Koch. Image signature: Highlighting sparse salient regions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(1):194–201, 2011.

[17] L. M. Hurvich and D. Jameson. An opponent-process theory of color vision. *Psychological Rev.*, 64(6p1):384, 1957.

[18] J. J. Koenderink and A. J. van Doorn. Blur and disorder. *J. Vis. Commun. Image Representation*, 11(2):237–244, 2000.

[19] F. Laakom, J. Raitoharju, J. Nikkanen, A. Iosifidis, and M. Gabbouj. Intel-tau: A color constancy dataset. *IEEE Access*, 9(-):39560–39567, 2021.

[20] I. Lampl, D. Ferster, T. Poggio, and M. Riesenhuber. Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *J. Neurophysiology*, 92(5):2704–2713, 2004.

[21] E. H. Land. Experiments in color vision. *Scientific Amer.*, 200(5):84–99, 1959.

[22] E. H. Land. The retinex. *Amer. Scientist*, 52(2):247–264, 1964.

[23] E. H. Land. The retinex theory of color vision. *Scientific Amer.*, 237(6):108–129, 1977.

[24] K. J. Linnell and D. H. Foster. Space-average scene colour used to extract illuminant information. In *Proc. John Dalton's Colour Vision Legacy*, 1997.

[25] J. Mollon. Monge: the verriest lecture, lyon, july 2005. *Vis. Neuroscience*, 23(3-4): 297–309, 2006.

[26] M. G. P. Rosa. Visual maps in the adult primate cerebral cortex: some implications for brain development and evolution. *Brazilian J. Med. Biol. Res.*, 35(-):1485–1498, 2002.

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

[28] G. Sharma, W. Wu, and E. N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Res. Appl.*, 30(1):21–30, 2005.

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

[30] D. Vishwanath and E. Blaser. Retinal blur and the perception of egocentric distance. *J. Vision*, 10(10):26–26, 2010.

[31] H. von Helmholtz and J. P. C. Southall. Helmholtz's treatise on physiological optics. *Opt. Soc. Amer.*, 1(-), 1924.

[32] R. Young. The gaussian derivative model for spatial vision. I- retinal mechanisms. *Spatial Vision*, 2(4):273–293, 1987.

[33] S. Zeki. *A Vision of the Brain*. Blackwell Science, ISBN: 0632030545, 1993.

[34] X.-S. Zhang, S.-B. Gao, C.-Y. Li, and Y.-J. Li. A retina inspired model for enhancing visibility of hazy images. *Frontiers Comput. Neuroscience*, 9(-):151, 2015.

[35] L. Zhaoping. The V1 hypothesis-creating a bottom-up saliency map for preattentive selection and segregation. *Understanding Vision: Theory Models Data*, -(-):189–314, 2014.