Multi-Scale Color Constancy based on Salient Varying Local Spatial Statistics

Oguzhan Ulucan ⓘ, Diclehan Ulucan ⓘ and Marc Ebner ⓘ

Computer Science, University of Greifswald, Walther-Rathenau-Str. 47, Greifswald, 17489, Germany.

*Corresponding author(s). E-mail(s): oguzhan.ulucan@uni-greifswald.de;
Contributing authors: diclehan.ulucan@uni-greifswald.de; marc.ebner@uni-greifswald.de;

Abstract

The human visual system unconsciously determines the color of the objects by "*discounting*" the effects of the illumination, whereas machine vision systems have difficulty performing this task. Color constancy algorithms assist computer vision pipelines by removing the effects of the illuminant, which in the end enables these pipelines to perform better on high-level vision tasks based on the color features of the scene. Due to its benefits, numerous color constancy algorithms have been developed, and existing techniques have been improved. Combining different strategies and investigating new methods might help us design simple yet effective algorithms. Thereupon, we present a color constancy algorithm based on the outcomes of our previous works. Our algorithm is built upon the biological findings that the human visual system might be discounting the illuminant based on the highest luminance patches and space-average color. We find the illuminant estimate based on the idea that if the world is gray on average, the deviation of the brightest pixels from the achromatic value should be caused by the illuminant. Our approach utilizes multi-scale operations by only considering the salient pixels. It relies on varying surface orientations by adopting a block-based approach. We show that our strategy outperforms learning-free algorithms and provides competitive results compared to the learning-based methods. Moreover, we demonstrate that using parts of our strategy can significantly improve the performance of several learning-free methods. We also briefly present an approach to transform our global color constancy method into a multi-illuminant color constancy approach.

Keywords: Computational color constancy, illumination estimation, white-balancing, scale-space

1 Introduction

If the human visual system had not evolved to *discount* the effects of illumination, color which is one of the most significant features we use to give meaning to our surroundings, would lose its importance. Semir Zeki, a well-known neuroscientist, emphasized the importance of identifying colors irrespective of the type of illuminant present in the scene by stating that without this ability, objects could no longer be reliably identified by their color [1]. The ability to perceive colors as constant regardless of the type of illuminant is called *color constancy*, and it is performed unconsciously by the human visual system [2]. Even though many studies have been conducted to understand this phenomenon, it is still unknown how the brain arrives at the color constant descriptors of the scene [3]. Investigating how we perform color constancy might be the key to unraveling how visual color processing works in the visual cortex, and the outcomes can help us to design more robust artificial systems [3, 4]. Therefore, color constancy is an

1

important field not only for neuroscientists to understand how the human visual system works but also for computer scientists to mimic our capabilities to create algorithms that perform better in scene understanding.

We can explain computational color constancy by making use of the following image formation model. An image captured by a device is an integrated signal of three key components which can be formulated as follows [3];

$$I_i(x,y) = \int R(x,y,\lambda) \cdot L(x,y,\lambda) \cdot S_i(\lambda)d\lambda \quad (1)$$

where, I_i presents a pixel measured at spatial location (x,y), R denotes the reflectance of objects, L is the wavelength distribution of the illumination, S_i is the capturing device's sensor function with $i \in \{$red, green, blue$\}$, and λ is the wavelength of the visible spectrum.

Without the knowledge of the measuring device and the type of light source illuminating the scene, it is very challenging for machine vision systems to identify the true colors of the objects. Therefore, in the field of computational color constancy, researchers aim at developing algorithms that discount the illumination conditions to assist various computer vision tasks. Since color constancy is an ill-posed problem, frequently, the assumption is made that the sensors' responses are narrow-band and the illumination is uniform throughout the entire scene. It is worth mentioning that while the relaxation of the problem helps us to estimate the illuminant of the scene, most scenes are not illuminated by uniform light sources [5, 6].

The formation of the measured data I can then be simplified as follows [3];

$$I(x,y) = R(x,y) \cdot \mathbf{L}. \quad (2)$$

With this simplification, researchers widely accept that after estimating $\mathbf{L}$, a white-balanced image I_{wb} can be obtained from color cast image I by using a 3×3 diagonal matrix as follows [7, 8];

$$I_{wb} = \begin{bmatrix} L_{est_g}/L_{est_r} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & L_{est_g}/L_{est_b} \end{bmatrix} \cdot I \quad (3)$$

where, $\mathbf{L}_{est} = [L_{est_r}, L_{est_g}, L_{est_b}]$ is the estimated color vector of the illuminant, and r, g and b represent the red, green, and blue color channels, respectively.

In the field of computational color constancy, numerous methods have been developed. Two algorithms, i.e. the max-RGB and gray world algorithm, inspired by the human visual system have been proposed some time ago, yet they are still used as building blocks of different methods [3, 9]. In one of our previous studies, we also used the assumptions of the max-RGB and gray world algorithm to design a color constancy algorithm. In our method, we further assumed that if the scene is achromatic on average, the shift of the highest luminance pixels away from an achromatic value should be the result of external factors, and this shift should be in the direction of the global illumination source [10]. We observed that our learning-free method provides satisfying results in estimating the illuminant even when lights beyond the standard illuminants are present in the scene, while learning-based methods face a challenge in discounting these lights since they are seldomly included in existing benchmarks. On the other hand, we realized that there are some pixels that reduce the performance of our algorithm, which also coincides with the observations reported in several studies [11–15]. In order to minimize the negative impact of these pixels, we adjusted our method so that only the blocks containing salient image elements, i.e. whitest pixels, are used [16]. In the same study, we showed that block-based computations and salient pixels increase the performance of the max-RGB and gray world algorithms. By taking inspiration from our previous investigations, recently, we further modified our method with a scale-space approach and demonstrated that like block-based operations, multi-scale computations can also improve the effectiveness of several color constancy algorithms [17].

In this paper, we extend our previous works by adding new discussions and experiments on new datasets, while also modifying our algorithm for mixed illumination conditions with a simple yet effective approach. First of all, we investigate the contribution of the stages of our method to the illumination estimation accuracy and analyze how the best-performing stages of our method contribute to the efficiencies of the learning-free color constancy algorithms. With this investigation, we not only introduce a color constancy algorithm but also provide a comprehensive analysis on how scale-space computations as well as block-based operations by only considering the salient pixels, and their combination affect the performance of existing methods. As a result, we demonstrate a simple yet effective approach to modify

the algorithms so that their accuracy in estimating the illuminant improves significantly. Moreover, we propose an approach to transform our global color constancy method into an algorithm that can achieve color constancy for mixed illumination conditions without utilizing any prior information about the scene.

Overall, we can summarize our contributions as follows:

- We propose a multi-scale block-based color constancy algorithm that takes advantage of scale-space and the varying local spatial statistics, while only considering the informative image elements to estimate the illuminant.
- We demonstrate that the efficiency of several learning-free color constancy algorithms can be improved by using different parts of our algorithm.
- We show that with a simple modification, the proposed technique designed for global color constancy can be converted into an algorithm for mixed illumination conditions which does not require any information about the scene, i.e. the number of the light sources.

This paper is organized as follows. We provide a brief literature review in Section 2. We detail the proposed method in Section 3. We present our experimental setup in Section 4, and we discuss our results in Section 5. Lastly, we give a brief summary of the study in Section 6.

2 Related Work

Numerous algorithms have been proposed to overcome the ill-posed nature of color constancy. In this section, we provide a brief review of global and multi-illuminant color constancy algorithms, which we utilize for comparison in our experiments. We simply group these methods into two categories as traditional algorithms and learning-based algorithms. While the former estimates the illuminant purely from image statistics, the latter extracts features from large-scale datasets to discount the effects of the illuminant. We would like to note that providing a full literature survey is outside of our scope, but research dedicated to this aim can be found in the following studies [3, 18, 19].

2.1 Traditional Algorithms

There are two well-known traditional color constancy methods, i.e. the gray world and the max-RGB. The gray world assumption was formalized by Buchsbaum in 1980. The gray world method depends on the assumption that on average the world is gray [20]. To estimate the color vector of the light source, the gray world method takes the mean of each color channel individually and outputs a vector formed by these average values as its illuminant estimate. The max-RGB method is based on the Retinex algorithm proposed by Land in 1971 [21]. To find an illuminant estimate, the max-RGB algorithm finds the maximum response of each color channel separately to form the color vector of the illuminant. These two methods establish the foundations of many color constancy approaches due to their simplicity and effectiveness. For instance, the shades of gray algorithm supposes that the mean of pixels raised to a certain power is gray [22]. The gray edge method and weighted gray edge algorithm assume that the mean of the high-frequency components of the image is achromatic [23–25]. The bright pixels method stresses the importance of the bright image elements [11]. The mean-shifted gray pixels method detects the gray pixels and uses them to find the illuminant [13]. The gray pixels color constancy algorithm detects the gray pixels in a scene by using a grayness measure and utilizes them to estimate the illuminant [14].

There are also other traditional methods that depend on different approaches. For instance, the local surface reflectance statistics approach considers the feedback modulation mechanism in the eye and the linear image formation model [26]. The principal component analysis based color constancy method uses only certain pixels, which have the largest gradient in the data matrix [12]. The double-opponency-based color constancy method is based on the physiological findings on color processing [27]. The biologically inspired color constancy method is based on the hierarchical structure of the human visual system and mimics our sensation on color illusions [9].

Alongside the algorithms assuming the illuminant is uniform throughout the scene, there are also several methods trying to solve the problem of color constancy for spatially varying illumination conditions. One early algorithm called local space average color is introduced by Ebner [28], who computationally modeled the biological finding that the human visual system might be discounting the illuminant based on local space average color. Gijsenij *et al.* introduced a block-based method that can be applied to global color constancy algorithms in a local manner [6]. The conditional random fields based method modifies existing

global color constancy methods to apply them for mixed illuminant conditions [29]. The retinal inspired color constancy model investigates the color processing mechanisms in a certain level of the retina [30]. The color constancy weighting factors method divides the image into regions and checks if a region contains sufficient information for color constancy by utilizing the normalized average absolute difference of each area as a measure of confidence [31]. The color constancy for image of non-uniformly lit scenes divides the image into regions and estimates the illuminant by using the regions containing sufficient color variation [32]. The color constancy adjustment based on texture of image method takes advantage of textures to detect image elements that have sufficient color variation and utilizes these to find the illuminant estimate [33]. The visual mechanism based color constancy with the bottom-up method mimics the bottom-up mechanisms of the human visual system [34]. The N-white balancing algorithm finds the number of white points in the image to discount the illuminant of the scene [35].

2.2 Learning-based Algorithms

The convolutional color constancy study is one of the earliest color constancy methods based on convolutional neural networks (CNNs) [36]. In this study, Barron transformed the operation of illumination estimation to an object detection application by formulating color constancy as a 2D spatial localization task. The deep specialized network for illuminant estimation is based on a convolutional network architecture and it is sensitive to diverse local regions of the scene [37]. The fast Fourier color constancy method carries out computations in the frequency domain and transforms the light source estimation task into a spatial localization operation on a torus [38]. The quasi-unsupervised color constancy algorithm relies on the detection of gray pixels without using a huge amount of labeled data during its training phase [39]. The color constancy convolutional autoencoder method utilizes convolutional autoencoders and unsupervised pre-training to estimate the illuminant of the scenes [40]. The sensor-independent color constancy model maps the input to a sensor-independent space by using an image-specific matrix [41]. The bag of color features color constancy method utilizes convolutional neural networks and it is based on bag-of-features pooling [42]. The cross-camera convolutional color constancy algorithm is trained on images captured with several different cameras, and during inference, additional unlabeled images are given as input so that the model can calibrate itself to the spectral properties of the testing set [43]. The combining color constancy algorithms model efficiently merges color constancy algorithms from the motivation that not all algorithms perform well on all scenes since the performance of algorithms is sensitive to the content of the input scene [44]. The one-net: convolutional color constancy simplified model does not utilize pre-trained layers and large kernels to show that complex models are not necessary to achieve high performance [45]. Alongside single illuminant cases, learning-based models are also being used in mixed-illumination conditions. Bianco *et al.* introduced a CNN-based color constancy method which has a detector that finds the number of light sources in the scene [46]. The physics-driven and generative adversarial networks (GAN) based color constancy method transforms the illumination estimation task into an image-to-image domain translation problem [47].

Apart from algorithms aiming at estimating the illuminant to perform color constancy, there also exist models that correct improperly white-balanced images without explicitly estimating the color vector of the light source. For instance, the KNN white-balance method computes a nonlinear color mapping function for correcting the colors of the image [48]. The extension of the KNN white-balance method is the interactive white-balance which relates the nonlinear color mapping functions directly to the colors chosen by the users to allow interactive white-balance adjustment [49]. The deep white-balance algorithm realistically edits the white-balance of an sRGB image by mapping the input to two white-balance settings by using a deep neural network [50]. The auto-white balance for mixed scenes renders the input scene with a small number of predefined white-balance settings through which it forms weight maps that are used to fuse the rendered scenes for generating a white-balanced output [51]. The style white-balance improves the contemporary auto-white balance models for single and mixed-illuminant inputs by modeling the illumination as style factor [52].

While the learning-based algorithms usually outperform the traditional methods on well-known benchmarks, their performance may decrease when they face images captured with unknown hardware specifications and/or the test samples contain lighting conditions different from their training set. These observations have been reported in several studies,
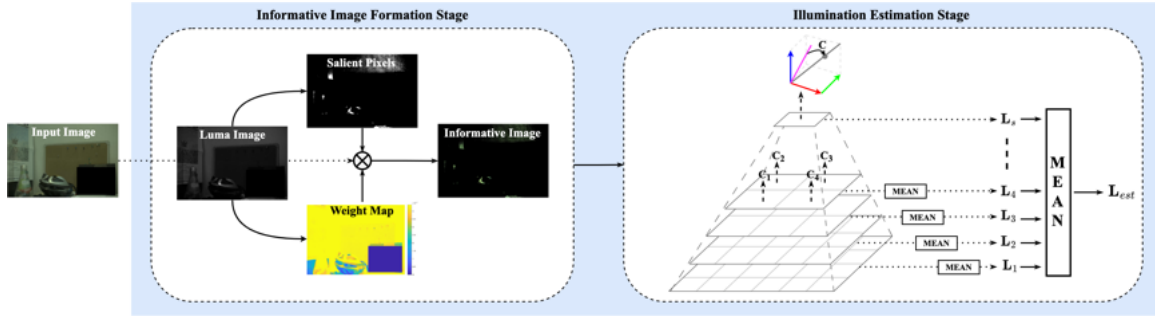
4

Fig. 1 The flowchart of our proposed illumination estimation algorithm. We use the luma image to obtain the salient pixels and their weights to form our informative image. Subsequently, we estimate the illuminant by carrying out computations in scale-space. We compute the scaling vectors $\mathbf{C}$ to find the deviation of the brightest pixels from the gray value. Then, we take the mean of the scaling vectors at each level to find an illuminant estimate $\mathbf{L}_s$ for the corresponding scale s. Lastly, we average all the illuminant estimates to find the color vector of the global light source $\mathbf{L}_{est}$ illuminating the scene.

where researchers shared their concerns about the problems in learning-based methods [14, 53–55]. Also, in our recent study, we have explicitly shown the performance decrease of learning-based algorithms, when they process images containing unique illumination conditions that are seldomly considered in available benchmarks [10]. We can briefly summarize the reasons behind the performance drop as follows. Most public benchmarks are formed with similar hardware, out-of-ordinary lights are usually not considered during dataset creation, and learning-based algorithms suppose that their train and test sets are similar in one way or another [10, 54].

3 Proposed Method

In this section, we detail our method (Fig. 1). We build our algorithm upon two assumptions, which have a correspondence in the human visual system. Since the human visual system might be discounting the illuminant of the scene based on space-average color and highest luminance patches [5, 20, 21, 56–60], we assume that on average, the world is gray, and there are several bright pixels somewhere in the scene. We form our main idea around these assumptions and assume that the deviation of the brightest pixels from the achromatic value should be caused by the light source (Fig. 2). We note that both the bright pixels and the achromatic value might change throughout the scene due to the varying local surface orientations. Therefore, in our method, we rely on a block-based approach to ensure that our algorithm is sensitive to local spatial information, which is usually neglected in other studies operating at the image level. Moreover, as stated in many color constancy studies, not all
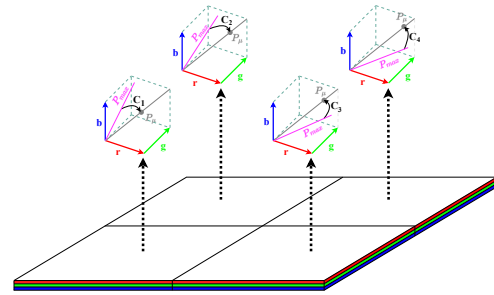

Fig. 2 The proposed method is based on two assumptions: *(i)* the world is gray on average and *(ii)* there are several bright pixels somewhere in the scene. Our aim is to find the color vector of the light source by finding the deviation of the brightest pixels $\mathbf{P}_{max}$ from the gray value P_μ by using the scaling vectors $\mathbf{C}$. Since the local surface orientations vary throughout the image, we calculate the scaling vectors for each non-overlapping block. Finally, we estimate the color of the global light source by taking the mean of all scaling vectors.

image elements are informative for estimating the illuminant. Therefore, we utilize only the image elements that we extract from the brightest pixels in the scene since pixels having the highest luminance might be useful for human color constancy [57–60]. Moreover, we adaptively weigh these salient pixels since their contribution to the task of color constancy could vary throughout the scene, i.e. they may not have an equal contribution to the illumination estimation task due to changing local statistics of the scene. Furthermore, we carry out our computations in multiple scales since the importance of utilizing the scale-space, especially for the tasks taking advantage of the color feature is demonstrated in many studies due to its sensitivity to the low-level features of images [61–65].

In our method, first of all, we apply a gamma correction in case an sRGB image is provided as

5

input. We carry out this operation to obtain the linear relationship between the pixels. Moreover, we do not consider the under- and over-saturated pixels, i.e., approximately the top 5% and the bottom 2% of image elements, in our calculations to reduce possible noise. Then, we focus on the extraction of the most informative pixels, i.e. salient regions, in the scene. In several studies, it is pointed out that not all pixels in an image are useful for the task of color constancy [9, 11, 12, 14, 15]. For instance, dominant sky regions have a tendency to bias the estimates of the light source, hence they should be handled separately. In order to find the regions containing only the informative pixels, we make use of the brightest pixels in the scene since it is known that the human visual system might be discounting the illuminant based on the highest luminance in the scene [57, 58]. To determine the salient pixels, we benefit from the black-white opponent channel O_{BW} of the opponent color space. We compute O_{BW} as follows [3];

$$O_{BW} = \frac{r + g + b}{\sqrt{3}} \qquad (4)$$

Subsequently, we form a binary saliency map $\mathcal{S}$, by selecting the informative pixels that correspond to the top 3.8% brightest pixels in the black-white opponent channel O_{BW}. The selection of the brightest pixels is described in detail in Subsection 5.1.

For a scene, all the pixels highlighted by the saliency map may not have equal brightness. Thus, using the binary saliency map directly may not be effective while estimating the illuminant. Thereupon, we adaptively weight the pixels in the saliency map by forming a map $\mathcal{W}$ from the black-white opponent channel O_{BW} by fitting the pixels into a Gaussian function (Eqn. 5). This weight map allows us to weaken the contribution of the darker image elements while giving more attention to the pixels having the highest luminance.

$$\mathcal{W}(x,y) = 1 - \frac{1}{2\pi\sigma^2} \cdot exp\left(-\frac{(O_{BW}(x,y) - \mu)^2}{2\sigma^2}\right) \quad (5)$$

where, σ and μ represent the standard deviation and the mean of O_{BW}, respectively.

Subsequently, by using the saliency and weight maps, we create an informative image $\mathcal{I}$, where the salient regions are adaptively weighted as follows;

$$\mathcal{I} = I(x,y) \cdot \mathcal{W}(x,y) \cdot \mathcal{S}(x,y). \qquad (6)$$
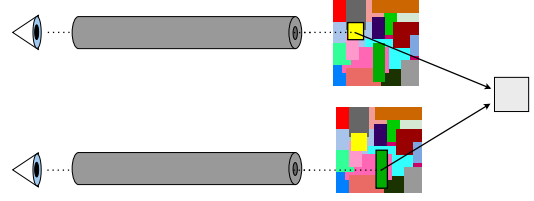

Fig. 3 The observers viewed only the center of the colored patches of the Mondrian through an adjustable aperture. Even though the colors of the patches are yellow and green, the observers perceived them as grayish-white. When the Mondrian is viewed as a whole, the observers identify the actual reflectance of the patches.

Afterwards, we carry out our computations in scale-space, where we can highlight the low-level features of images. We obtain representations of the images at different levels, while we determine the number of scales $\mathcal{L}$ adaptively based on the image resolution, and it can be calculated as $\mathcal{L} = \lfloor log(min(h,w))/log(2) \rfloor$, where, h and w are the width and height of the image. Then, to respect the local surface orientations, we divide the image into non-overlapping blocks which consist of m number of pixels (Fig.1). The parameter m depends on the image resolution at the scale of interest and it is calculated as $m = \sqrt{(h \cdot w)/\eta}$, where, η is the controlling parameter of m which is taken as 120 based on practical experiments (the process of determining this parameter is provided in Section 5.1).

It is worth mentioning that we divide the image into non-overlapping blocks only at levels higher than half of the number of possible scales that can be obtained in the pyramid. The reason behind this is that at coarser scales the blocks would contain a small number of pixels which would violate one of our algorithm's assumptions, i.e. the world is achromatic on average, since the gray world assumption requires a sufficient number of distinct colors to be present in the scene [3]. This requirement also coincides with the mechanisms of the human visual system which are explicitly demonstrated in Land's study [56]. In the experiments, Land showed that two color patches with different colors, i.e. yellow and green, taken from a Mondrian image are perceived as grayish-white only when their centers are viewed through viewing tubes in void mode, i.e. patches are viewed so that they are isolated from their local neighbors (Fig. 3) [3].

After we divide the images in scale-space into blocks, we determine the descriptors to estimate the color vector of the light source in every block. Note that for simplicity, we refer to each image in coarser scales that is not divided into blocks as a block. As

we mentioned before, our algorithm builds upon the assumptions of the gray world method and the max-RGB. For each block P we assume that there is a unique achromatic value, which is our first descriptor that is used to find the illuminant estimate. We compute this gray value P_μ by taking the average over all channels within the block of interest instead of taking the mean of all pixels in the image directly. We calculate a particular gray value for each block to take the varying local spatial statistics into account. For the same reason, we determine our second descriptor by taking the maximum response of each block channel individually and represent them by $\mathbf{P}_{max} = [P_{r,max}, P_{g,max}, P_{b,max}]$.

We calculate the color vector of the light source of each block in every scale by using both of our assumptions. Based on our main idea we compute how much the brightest values $\mathbf{P}_{max}$ deviate from the gray value P_μ. We assume that if the world is achromatic on average, then the summation of the intensity values of $\mathbf{P}_{max}$ should equal P_μ. However, if there is a shift away from the achromatic value this deviation should be in the direction of the color vector of the light source. We can find this deviation by using a vector $\mathbf{C}_P = [c_r, c_g, c_b]$, where each element of $\mathbf{C}_P$ scales the intensities of $\mathbf{P}_{max}$ such that they sum to P_μ as follows;

$$P_{r,max} \cdot c_r + P_{g,max} \cdot c_g + P_{b,max} \cdot c_b = P_\mu. \quad (7)$$

We convert Eqn. 7 into an optimization problem to find $\mathbf{C}_P$ as follows;

$$\mathbf{C}_P = \arg\min_{\mathbf{C}_P} \|P_{max}\, \mathbf{C}_P - P_\mu\|_2$$
$$with\ \forall c \in \mathbf{C}_P : c > 0. \quad (8)$$

Note that we do not only minimize the norm of this optimization problem but also the Euclidean norm of $\mathbf{C}_P$. In other words, if we have multiple solutions for this problem, we take the solution that minimizes the norm of $\mathbf{C}_P$.

After we obtain the $\mathbf{C}_P$ values for each block, we calculate the illuminant estimate for each scale by averaging the $\mathbf{C}_P$ as follows;

$$\mathbf{L}_s = \frac{1}{n} \sum_{k=1}^{n} \mathbf{C}_{P_k} \quad (9)$$

where, $\mathbf{L}_s$ is the color vector of the light source at a certain scale s, and n is the number of blocks. In case

we obtain $\mathbf{C}_P$ in coarser scales where we do not divide the image into blocks, we directly take the deviation obtained from Eqn. 8 as our illuminant estimate for that scale.

Since we assume that the scenes are uniformly illuminated, we linearly combine the estimations from each level to obtain a single illuminant estimate $\mathbf{L}_{est}$ for the given image as follows;

$$\mathbf{L}_{est} = \frac{1}{\mathcal{L}} \sum_{k=1}^{\mathcal{L}} \mathbf{L}_{s_k} \quad (10)$$

where, $\mathbf{L}_{est}$ is then converted into a unit vector.

Lastly, we can discount the illuminant to obtain a white-balanced image I_{wb} by using Eqn. 3.

3.1 Application to Multi-illuminant Color Constancy

A common drawback in several multi-illuminant color constancy methods is the requirement of prior information related to the number of clusters or segments, which depends on the number of illuminants present in the scene [66]. We argue that as we cannot know the type of light source and/or type of the capturing device, we can also not be sure about the number of lights illuminating the scene. Therefore, algorithms that utilize the number of illuminants as prior information will face a significant challenge in discounting the illuminant when this parameter is not provided correctly. Consequently, we cannot obtain a robust algorithm that achieves color constancy for mixed-illumination conditions effectively if we cannot design an algorithm free of prior information.

As aforementioned, we develop our method for the cases where we have a uniformly illuminated scene. However, we can modify our method with a simple yet effective approach so that it can provide pixel-wise estimates of the illuminant for the scenes illuminated by varying illumination conditions. In this subsection, we explain our modification that can transform our global color constancy method into a multi-illuminant color constancy approach, which does not require any prior information about the scene.

Our first adjustment is for our informative image formation stage (Fig. 1). For scenes illuminated by multiple light sources, we use all pixels that are close to white, i.e. pixels having the highest luminance, instead of using the top brightest image elements. We need this modification since in multi-illuminant color constancy, we need to take more spatial information

into account. If we utilize only a small number of pixels, we cannot provide accurate pixel-wise illuminant estimates since we would lose the local relationship between the neighboring pixels, which is an important cue for mixed illumination conditions [9, 61]. As a result, we are using all the pixels, which are closest to white. We can explain the idea of using the whitest pixels from two different perspectives, *(i)* digital photography, and *(ii)* human color constancy. In digital photography, we know that we can easily determine the color vector of the illumination by using the white pixels in the image instead of utilizing the colored ones [9]. For instance, let us assume that we capture the picture of a room illuminated with yellow light. The room has white walls and contains different objects having distinct colors. We can determine the color vector of the light source easier from the white walls rather than the objects since the capturing device will measure the light reflected from the white walls as yellow. Furthermore, from the findings on human color constancy, we know that the areas having the highest luminance might be used by the human visual system to discount the effects of the light source [3, 9, 57, 59, 60].

We determine the image elements closest to white by using a simple yet effective approach [9]. In order to find such image elements, we form a temporary color vector of the light source by taking the mean of each color channel individually, in other words, we apply the gray-world algorithm since we are assuming that world is gray on average. Then, we obtain a temporary white-balanced image I_{temp} by scaling the input image according to this temporary color vector of the light source by Eqn. 3. Afterwards, we create a pixel-wise whiteness map $\mathcal{W}$ by calculating the pixel-wise distance between the white vector $\mathbf{w} = [1\ 1\ 1]$ and the I_{temp} as follows;

$$\mathcal{W}(x,y) = cos^{-1}\left(\frac{\mathbf{w}\cdot I_{temp}}{\|\mathbf{w}\|\cdot\|I_{temp}\|}\right). \quad (11)$$

Since the contribution of all spatial locations differ, we obtain a certainty map $\mathcal{C}$, by fitting $\mathcal{W}$ into a Gaussian function as follows;

$$\mathcal{C}(x,y) = \frac{1}{2\pi\sigma_\mathcal{W}^2}\cdot exp\left(-\frac{(\mathcal{W}(x,y)-\mu_\mathcal{W})^2}{2\sigma_\mathcal{W}^2}\right) \quad (12)$$

where, $\mu_\mathcal{W}$ and $\sigma_\mathcal{W}$ are the mean and the standard deviation of $\mathcal{W}$, respectively.

By using $\mathcal{C}$, we form our informative image (Eqn. 13) which we use for our multi-scale block-based computations.

$$\mathcal{I} = I(x,y) \cdot \mathcal{C}(x,y) \quad (13)$$

After obtaining our informative image, we follow similar operations to our global color constancy approach, i.e. we carry out block-based operations in scale-space. However, contrary to the single-illuminant case, for the multi-illuminant scenario, we do not consider the coarser scales of the pyramid where the locality starts to degrade since locality is an important cue for multi-illuminant color constancy [67]. We use half of the number of possible scales that can be reached in a pyramid since we observed in our experiments that the performance generally starts to degrade afterwards. Moreover, we do not take the mean of the computed deviation $\mathbf{C}_P$, but we place $\mathbf{C}_P$ into the center of the corresponding block to obtain a sparsely populated image $\mathcal{I}_{C_P}$, i.e. an image only containing the estimated deviations in the center of each block. Then, in order to fill the missing pixels in $\mathcal{I}_{C_P}$, i.e. to obtain a dense image for a specific scale, where for every spatial location a pixel-wise estimate $L_s(x,y)$ is present, we carry out an interpolation between the neighboring center pixels by convolving $\mathcal{I}_{C_P}$ with a Gaussian kernel (Eqn. 14). We follow this approach to obtain smooth transitions between adjacent blocks instead of assuming that all the pixels in a block have the same deviation, which would result in sharp changes between adjacent blocks.

$$L_s(x,y) = \mathcal{I}_{C_P} * \frac{1}{2\pi\sigma^2} \cdot exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (14)$$

where, $*$ denotes the convolution operation. It is important to note that the scaling factor σ, i.e. controlling parameter of the Gaussian kernel, should be large enough to ensure that at least two neighboring deviations are inside of the kernel. This parameter is practically determined and it is calculated as $\sigma = 0.1\beta$, where $\beta = (min(h,w)/2)$.

After obtaining L_s for each scale, in order to find our pixel-wise estimates $L_{est}(x,y)$, we process every L_s as follows. We first upsample the coarsest scale so that it matches the size of the consecutive finer level. Then, we linearly combine the upsampled image with the one on the finer scale. Afterwards, we upsample the resulting image to linearly combine it with its

Fig. 4 (Left-to-right) The RECommended ColorChecker, the INTEL-TAU, the NUS-8, the MIMO, and the Mixed-Illuminant Test Set datasets, respectively.

consecutive finer level. We carry out this operation until the finest scale is reached. The resulting image represents our pixel-wise illumination estimate L_{est}.

4 Experimental Setup

In this section, we demonstrate our experimental setup. In the following Subsection 4.1, we introduce the datasets that we utilize to compare our algorithm's performance with the studies briefly explained in Section 2, and the initial and previous versions of our method, i.e. block-based color constancy, and block-based color constancy with salient pixels, respectively [10, 16]. In Subsection 4.2, we explain the error metrics that we used to report the statistical results. We would like to note that we performed the experiments on an Intel i7 CPU @ 2.7 GHz Quad-Core 16GB RAM machine using MATLAB R2021a.

4.1 Datasets

In order to investigate the contribution of the stages of the proposed algorithm to the performance, and benchmark our method and the modified algorithms, we carry out comprehensive evaluations on 3 well-known global color constancy datasets namely, the RECommended ColorChecker [68], INTEL-TAU [69], and NUS-8 [12] datasets. While in our previous works', we used the RECommended ColorChecker and INTEL-TAU datasets, in this study we extend our discussion by also utilizing the well-known NUS-8 dataset. Moreover, to provide an analysis on multi-illuminant cases, we utilize 2 benchmarks, the Multiple Illuminant and Multiple Object (MIMO) Dataset [29], and the Mixed-Illuminant Test Set which is recently created by Afifi *et al.* [51] (Fig. 4). In order to use these datasets, we mask out the calibration objects, i.e. color checker, and if necessary we subtract the black level from the original images. Also, we clip the under- and over-saturated pixels to prevent the contribution of the noisy pixels since it is known that these image elements negatively affect the performance of color constancy methods. In the following, we briefly introduce these benchmarks.

The RECommended ColorChecker Dataset

The RECommended ColorChecker dataset is the updated version of the Gehler-Shi dataset [70]. In this modified version, the researchers provide accurate ground truths for each scene to solve the problems in the Gehler-Shi dataset. The dataset contains 568 scenes, captured under single dominant illumination. The scenes contain 254 indoor, 85 outdoor, and 229 close-up images, which are taken by two different capturing devices namely, Canon 1D and Canon 5D [54].

The INTEL-TAU Dataset

The INTEL-TAU dataset is one of the largest color constancy benchmarks containing 7022 images. The 1466 indoor, 2327 outdoor and 3229 close-up scenes are captured with different devices namely, Nikon D810, Canon 5DSR, and Sony IMX135 [54]. The images are captured under one light source. All the images in INTEL-TAU are already processed, i.e. images have a linear response and their black level is calibrated. Moreover, contrary to other datasets, all the sensitive data in the INTEL-TAU dataset is handled, i.e. the faces and license plates are masked out. To utilize this dataset, we used the images belonging to the sets of "*field1*" and "*field3*" since the calibration object is unmasked in other sets, i.e. "*lab printouts*", and "*lab realscene*", and their masks are not provided.

The NUS-8 Dataset

The NUS-8 dataset is another publicly available global color constancy benchmark, which contains a total of 1736 raw images. The images containing 415 indoor, 279 outdoor and 1159 close-up scenes are captured with 8 different cameras namely, Canon EOS-1Ds Mark III, Canon EOS 600D, Fujifilm X-M1, Nikon D5200, Olympus E-PL6, Panasonic Lumix DMC-GX1, Samsung NX2000, and Sony SLT-A57 [54].

The MIMO Dataset

The MIMO dataset is one of the well-known benchmarks in the field of color constancy for mixed illumination conditions. We evaluate our illumination estimation strategy on this dataset since many algorithms have already been tested on this benchmark [54]. The MIMO dataset contains a total of 78 linear images for two different sets; *(i)* the real-world scenes containing 20 complex scenes, and *(ii)* 58 laboratory scenes consisting of simple scenes.

The Mixed-Illuminant Test Set

The Mixed-Illuminant Test Set is a publicly available benchmark. This recently created dataset is rendered by computer graphics, hence, the ground truths are not biased by camera sensor specifications. The synthetic dataset contains a total of 150 images with 30 varying scenes. Each scene is rendered with 5 different mixed illumination conditions at different color temperatures, and for each scene, the ground truth white-balanced image is provided.

4.2 Error Metric

To present statistical results, we adopt the well-known error metric, the angular error, between the color vector of the estimated illuminant $\mathbf{L}_{est}$ and the ground truth $\mathbf{L}_{gt}$. The angular error between two vectors can be calculated as follows;

$$\varepsilon(\mathbf{L}_{est}, \mathbf{L}_{gt}) = cos^{-1}\left(\frac{\mathbf{L}_{est} \cdot \mathbf{L}_{gt}}{\|\mathbf{L}_{est}\| \cdot \|\mathbf{L}_{gt}\|}\right). \qquad (15)$$

While we report the mean, the median, the mean of the best 25%, and the mean of the worst 25% of the angular error for uniform illuminant cases, we analyze the mean and the median pixel-wise angular error for the mixed-illuminant cases.

5 Experimental Discussion

In this section, we provide a detailed experimental discussion while presenting both statistical and visual analyses on all datasets. Firstly, in Subsection 5.1, we demonstrate the parameter selection process that we followed during our algorithm design. Then, in Subsection 5.2, we investigate the contributions of the stages of our algorithm to the performance in detail. After analyzing the stages of our method, in Subsection 5.3, we discuss our outcomes on single-illuminant color constancy by adopting 3 datasets, while we also investigate the effects of modifying existing color constancy algorithms by using our observations in the ablation study. Lastly, in Subsection 5.4, we provide our results for the application to multi-illuminant color constancy.

As a final note, the detailed ablation study, the modification of existing algorithms by using the best-performing stages of our method, and the discussions on the NUS-8, MIMO, and Mixed-Illuminant Test Set datasets are all extensions to our previous works' experimental discussions.

5.1 Discussion on Parameter Selection

As we mentioned in Section 3, our method depends on two parameters, i.e. the parameter that controls the size of the non-overlapping blocks, and the percentage of the brightest pixels which we utilize to form our informative image. In order to analyze their effects on performance, we investigate how the efficiency of our method changes for global color constancy with the consideration of their different combinations. For this analysis, we create a subset containing random samples from the global color constancy datasets. In order to determine the best-performing combination, we check the parameter combination having the lowest mean angular error (Table 1).

As aforementioned, our algorithm is built upon the gray world assumption. Thus a sufficient number of pixels has to fall into each non-overlapping block since the gray world assumption is only valid when there is an adequate number of distinct colors in the image, i.e. block. In our algorithm, both of the parameters (number of brightest pixels and number of pixels per block) affect the number of pixels that we use to estimate the illuminant in each block. While the top brightest pixels are related to the image statistics, the number of image elements in each non-overlapping block is dependent on the image size. As shown in Table 1, the performance of our algorithm increases when a sufficient number of blocks having an adequate number of image elements is taken into account. We can explain this observation with the fact that the possibility of changing surface orientations increases as we select a sufficient number of blocks with an adequate number of pixels. Moreover, since the chance of obtaining uniform colored areas decreases as we choose blocks with an appropriate size, we can satisfy the assumptions of the gray world by choosing blocks

Table 1 Selecting the best parameter combination on a subset containing random samples.

	The top % of brightest pixels							
	3%	3.2%	3.4%	3.6%	3.8%	4%	4.2%	4.4%
40	3.49	3.48	3.47	3.48	3.48	3.49	3.48	3.48
60	3.44	3.43	3.43	3.42	3.42	3.45	3.46	3.46
80	3.43	3.42	3.43	3.44	3.44	3.46	3.45	3.46
100	3.47	3.45	3.45	3.45	3.44	3.46	3.47	3.48
120	3.39	3.37	3.37	3.37	**3.36**	3.38	3.40	3.42
140	3.43	3.42	3.41	3.42	3.42	3.43	3.43	3.44
160	3.45	3.44	3.43	3.43	3.43	3.44	3.44	3.44
180	3.43	3.41	3.40	3.40	3.40	3.41	3.42	3.43

The parameter η (row labels, leftmost column).

Table 2 Ablation study on the steps of the proposed method.

	Mean	Median
Baseline	9.10	8.16
Informative Image	9.04	8.03
Scale-Space	3.77	2.92
Blocks	3.80	3.07
Scale-Space, Informative Image	4.03	3.36
Blocks, Informative Image	3.57	2.76
Blocks, Scale-Space	3.80	2.85
Scale-Space, Blocks, Informative Image	3.16	2.22

with a sufficient number of pixels. Due to similar reasons, the number of the brightest pixels falling into a block has to satisfy our assumptions.

As seen in Table 1, the best combination is obtained by choosing the top 3.8% brightest pixels, and the controlling parameter of the block size as 120.

5.2 Ablation Study

We conduct an ablation study on a single dataset, which we obtain by combining all benchmarks, to analyze the contributions of each step of our method to the performance of color constancy. In Table 2, we provide the results of our investigation, where *baseline* refers to solving only Eqn. 8 without considering the informative image formation step and without carrying out block-based computations in scale-space. The outcomes we provide alongside the baseline correspond to solving Eqn. 8 by considering only particular stages of our proposed approach, such as solving Eqn. 8 by using only the informative image or only carrying out the computations in scale-space. In order to analyze how the steps of the algorithm affect the performance of our method, we divide our investigation into three parts. First, we analyze their

contribution alone, then we investigate their dual combinations, and lastly, we present the results of our proposed technique. For each part, we choose the best-performing strategy alongside the proposed method which we will use later to modify several learning-free color constancy algorithms. We observe that each component of our algorithm contributes to the performance noticeably. We see that solving Eqn. 8 by utilizing the informative image slightly increases the performance of our method. However, as presented in Table 2, using either the blocks or the scale-space increases the performance substantially. When we check the angular errors, carrying the computations into scale-space rather than utilizing blocks results in a slightly better performance. The reason behind this can be explained as follows. While we reduce the image from the finer scale to the consecutive coarser scales, we actually apply a local averaging between the pixels while reducing the image size, which acts like a block-based operation, especially for the coarser scales. Therefore, due to its sensitivity to low-frequency components of the image, i.e., colors, and the partial block-based operations in particular at the coarser scales, the scale-space approach performs slightly better than estimating the illuminant only on the finest scale by using block-based operations. After investigating the steps utilized in our algorithm individually, we analyze their dual combinations. We notice that the efficiency of our method increases significantly when we consider using blocks and the informative image rather than the scale-space and the informative image. The reason behind this performance difference can be explained by the contribution of the varying local statistics. When we use the block-based approach together with the informative image, we give more importance to local statistics than using the informative image with scale-space since when we divide the image into non-overlapping blocks, we take more varying local estimations into account than in the scale-space, i.e., the number of blocks is higher than the number of scales for an image, thus, more information is taken into account in block-based operations combined with the informative image rather than considering the scale space with the informative image. Yet, we obtain the best outcomes when we combine all three steps in our algorithm since the locality is respected the most when all stages are considered and only the informative regions are taken into account.

5.3 Results of Single-Illuminant Color Constancy

In order to present the performance of our method and to analyze the effectiveness of carrying out the computations of several learning-free color constancy algorithms by using the highlighted steps in the ablation study, we make a comprehensive comparison with numerous color constancy algorithms. For each benchmark, we obtain the results of the methods by either running their codes without any optimization or by making use of the reported outcomes of their works and recent publications, which are considered to be up-to-date and comprehensive [14, 41, 43]. While discussing the experimental outcomes, we first focus on the results of our proposed method. Then, we analyze the effects of modifying existing color constancy methods with our approach.

We provide statistical analyses for global color constancy in Table 3, Table 4, and Table 5. On the RECommended ColorChecker dataset (Table 3), the first noticeable outcome is that we obtain the lowest mean and the mean of the worst 25% of the angular error among the traditional algorithms, while we achieve competitive results compared to the learning-based models. Furthermore, we observe that the extensions we made to the former version of our algorithm improved the performance of our method in all metrics. On the INTEL-TAU dataset (Table 4), again we obtain the lowest mean angular error among the learning-free algorithms, while we outperform 4 of the learning-based methods. On the NUS-8 dataset (Table 5), we achieve the lowest mean and the mean of the worst 25% of the angular error among the learning-free methods. Compared to the proposed algorithm's previous version, the improvement in our best and worst cases leads to a significant decrease in the mean angular error.

We provide visual comparisons in Fig.5 and Fig.6 by using random samples from the benchmarks. It is known that in color constancy, scenes containing a limited number of distinct colors are challenging for the algorithms. In our visual results, we observe that even for these scenes our angular error is less than 5 (second and third rows of Fig.5). Also, in Fig.6 where we provide an analysis by taking random samples among the worst cases of our previous version, block-based color constancy with salient pixels, we can see that our worst cases significantly improved. Yet, scenes containing uniformly distributed colors and a

Table 3 Statistical results on the RECommended ColorChecker dataset. The best results are highlighted.

	Algorithms	Mean	Med.	B.25%	W.25%
Learning-based	Quasi-U CC [39]	3.46	2.23	-	-
	COCOA [44]	2.64	1.86	-	-
	C3AE [40]	2.10	1.90	0.80	**4.00**
	CCC [36]	**1.95**	**1.22**	**0.35**	4.76
	SIIE [41]	2.77	1.93	0.55	6.53
	FFCC [38]	2.95	2.19	0.57	6.75
	C5 [43]	2.50	1.99	0.53	5.46
Traditional-based	max-RGB [21]	7.78	5.43	1.49	17.47
	GW [20]	4.71	3.54	0.93	10.44
	SoG [22]	4.09	2.42	0.56	10.28
	1st - GE [24]	5.79	3.68	0.93	14.17
	2nd - GE [24]	6.09	3.97	1.03	14.70
	wGE [25]	6.08	3.93	0.78	15.57
	DOCC [27]	7.23	4.26	0.79	18.04
	PCA-CC [12]	4.11	2.52	0.53	10.19
	LSRS [26]	4.03	3.06	1.39	8.17
	MSGP [13]	3.81	2.96	0.76	8.35
	GI [14]	3.19	1.90	0.44	8.02
	BIO-CC [9]	4.40	3.30	0.86	9.84
	BB-CC [10]	3.82	3.16	1.45	7.38
	BB-CC w/ SP [16]	3.48	2.71	1.06	7.37
	Proposed	3.16	2.16	0.62	7.32
Scale-Space	max-RGB	3.87	2.68	0.80	8.98
	GW	3.92	2.74	0.57	9.20
	SoG	3.61	2.26	0.47	9.13
	1st - GE	3.53	2.22	0.50	8.83
	2nd - GE	3.51	2.19	0.54	8.76
	wGE	3.86	2.14	0.50	10.35
	DOCC	3.60	2.30	0.54	8.85
Blocks, Informative Image	max-RGB	3.46	2.69	1.02	7.11
	GW	3.39	2.46	0.57	7.82
	SoG	3.35	2.45	0.69	7.48
	1st - GE	3.31	2.34	0.59	7.57
	2nd - GE	3.31	2.37	0.61	7.56
	wGE	3.30	2.37	0.58	7.55
	DOCC	3.30	2.36	0.60	7.54
Scale-Space, Blocks Informative Image	max-RGB	3.72	2.59	0.71	8.67
	GW	3.76	2.54	0.68	8.86
	SoG	3.73	2.56	0.67	8.76
	1st - GE	3.87	2.73	0.72	9.05
	2nd - GE	3.87	2.73	0.72	9.05
	wGE	3.87	2.72	0.71	9.05
	DOCC	3.87	2.73	0.73	9.02

limited number of bright pixels are still challenging (last row of Fig.6).

After we investigate the outcomes of our algorithm, we analyze how our modifications affect the existing color constancy methods' performance. As

Table 4 Statistical results on the INTEL-TAU dataset. The best results are highlighted.

	Algorithms	Mean	Med.	B.25%	W.25%
Learning-based	Quasi-U CC [39]	3.12	2.19	0.60	7.28
	One-Net CCC [45]	3.30	3.20	1.10	5.90
	C3AE [40]	3.40	2.70	0.90	7.00
	BoCF CC [42]	2.90	2.40	0.90	6.10
	SIIE [41]	3.42	2.42	0.73	7.80
	FFCC [38]	3.42	2.38	0.70	7.96
	C5 [43]	**2.52**	**1.70**	**0.52**	**5.96**
Traditional-based	max-RGB [21]	10.49	11.14	1.70	19.24
	GW [20]	4.90	3.85	0.93	10.59
	SoG [22]	5.13	3.72	0.86	11.77
	1st - GE [24]	5.89	4.07	0.94	13.79
	2nd - GE [24]	6.09	4.25	1.00	14.13
	wGE [25]	5.99	3.63	0.80	14.89
	DOCC [27]	7.18	4.66	0.80	16.97
	PCA-CC [12]	4.46	3.03	0.68	10.64
	LSRS [26]	4.16	3.42	0.98	8.60
	MSGP [13]	3.57	2.56	0.64	8.23
	GI [14]	3.32	2.18	0.56	8.03
	BIO-CC [9]	4.14	3.05	0.76	9.42
	BB-CC [10]	4.29	3.61	1.19	8.52
	BB-CC w/ SP [16]	3.37	2.63	0.79	7.25
	Proposed	3.23	2.23	0.59	7.47
Scale-Space	max-RGB	4.47	3.28	0.98	9.95
	GW	4.39	3.40	0.78	9.65
	SoG	4.30	3.15	0.76	9.80
	1st - GE	4.27	3.03	0.75	9.96
	2nd - GE	4.10	2.74	0.71	9.75
	wGE	4.18	2.24	0.55	11.09
	DOCC	4.01	2.91	0.69	9.29
Blocks, Informative Image	max-RGB	3.94	3.06	0.90	8.48
	GW	3.66	2.53	0.64	8.51
	SoG	3.66	2.65	0.69	8.33
	1st - GE	3.62	2.56	0.65	8.36
	2nd - GE	3.63	2.57	0.66	8.33
	wGE	3.60	2.53	0.64	8.32
	DOCC	3.57	2.51	0.64	8.26
Scale-Space, Blocks Informative Image	max-RGB	4.01	2.83	0.71	9.35
	GW	4.04	2.82	0.70	9.44
	SoG	4.02	2.82	0.69	9.40
	1st - GE	4.11	2.90	0.71	9.63
	2nd - GE	4.10	2.90	0.71	9.61
	wGE	4.11	2.91	0.71	9.63
	DOCC	4.10	2.87	0.70	9.59

Table 5 Statistical results on the NUS-8 dataset. The best results are highlighted.

	Algorithms	Mean	Med.	B.25%	W.25%
Learning-based	Quasi-U CC [39]	3.00	2.25	-	-
	One-Net CCC [45]	2.16	1.57	0.54	4.76
	CCC [36]	2.38	1.69	**0.45**	5.85
	DSNIE [37]	2.24	1.46	0.48	5.28
	SIIE [41]	2.05	1.50	0.52	4.48
	FFCC [38]	2.87	2.14	0.71	6.23
	C5 [43]	**1.77**	**1.37**	0.48	**3.75**
Traditional-based	max-RGB [21]	10.34	9.90	1.72	19.82
	GW [20]	4.13	3.12	0.86	9.17
	SoG [22]	3.51	2.77	0.85	7.57
	1st - GE [24]	3.40	2.55	0.84	7.47
	2nd - GE [24]	3.66	2.70	0.94	8.05
	wGE [25]	3.09	2.23	0.71	7.00
	DOCC [27]	3.87	2.42	0.77	9.52
	PCA-CC [12]	3.09	2.20	0.65	7.08
	LSRS [26]	3.85	2.98	1.26	7.92
	MSGP [13]	3.03	2.08	0.62	6.99
	GI [14]	3.18	2.22	0.61	7.45
	BP [11]	3.17	2.41	0.69	7.02
	BB-CC [10]	3.78	3.07	1.32	7.47
	BB-CC w/ SP [16]	3.27	2.52	1.06	6.79
	Proposed	2.93	2.21	0.74	6.37
Scale-Space	max-RGB	3.35	2.70	1.01	6.81
	GW	3.65	2.58	0.74	8.31
	SoG	3.35	2.53	0.73	7.38
	1st - GE	3.23	2.50	0.75	7.09
	2nd - GE	3.01	2.16	0.66	6.80
	wGE	2.82	1.95	0.60	6.54
	DOCC	3.05	2.29	0.68	6.73
Blocks, Informative Image	max-RGB	3.58	3.01	1.11	7.03
	GW	3.02	2.13	0.65	6.90
	SoG	3.02	2.25	0.72	6.62
	1st - GE	2.94	2.14	0.65	6.64
	2nd - GE	2.95	2.15	0.66	6.63
	wGE	2.91	2.10	0.64	6.57
	DOCC	2.87	2.08	0.63	6.47
Scale-Space, Blocks Informative Image	max-RGB	3.47	2.57	0.71	7.82
	GW	3.53	2.58	0.70	8.02
	SoG	3.49	2.57	0.69	7.93
	1st - GE	3.57	2.61	0.71	8.14
	2nd - GE	3.56	2.60	0.71	8.13
	wGE	3.57	2.63	0.71	8.13
	DOCC	3.56	2.58	0.70	8.11

aforementioned, we modify the learning-free methods by exchanging Eqn. 8 with their computations on estimating the illuminant. In particular, we investigate the effects of modifying several color constancy algorithms through our observations in the ablation study (Table 2). 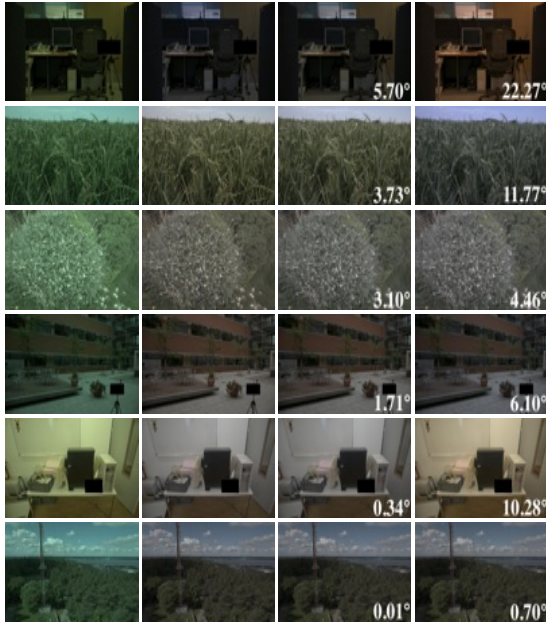As aforementioned, we select the best-performing steps of our algorithm. It is worth mentioning that we do not modify traditional algorithms that require information from parts, which are discarded in our approach. For instance, principal component analysis based color constancy [12]
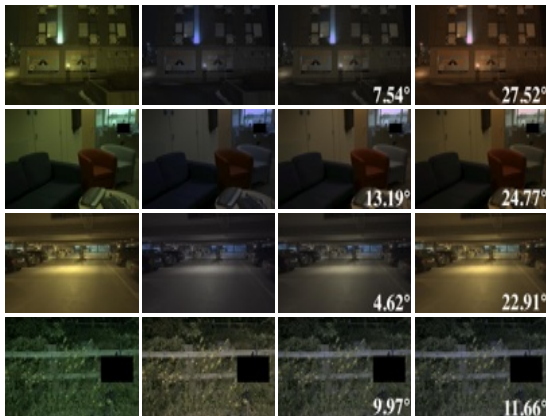
Fig. 5 Visual comparison on random samples. (Left-to-right) Input image, ground truth, proposed algorithm, and color constancy method (top-to-bottom) max-RGB [21], wGE [25], GI [14], GW [20], DOCC [27], and PCA-CC [12]. The angular error is provided on the bottom-right side of the image.

Fig. 6 The visual comparison of the proposed method for random samples taken among the worst cases in its previous version block-based color constancy with salient pixels. (Left-to-right) Input image, ground truth, proposed method, and former version [16]. The angular error is provided on the bottom-right side of the image.

needs information from both the brightest and darkest regions of the image, thus our strategy, which does not consider the darkest pixels, is not suitable to modify this algorithm. Furthermore, we do not apply our approach to learning-based methods, since they have a fixed input size requirement, while we use blocks with varying sizes in our approach, and resizing these

blocks so that they meet the input requirements would distort the image.

As shown in Table 3, Table 4, and Table 5, the modified algorithms achieve lower mean angular error compared to their original versions, while they also outperform several other color constancy algorithms. We observe that all three steps of our method, which are used to modify the algorithms increase the performance significantly, while the highest performance increase is usually obtained by utilizing the blocks and the informative image. We can explain this outcome by two facts; *(i)* all pixels are not informative for color constancy, and *(ii)* taking varying local spatial information into account allows us to highlight local features that might not be noticed while operating on global scale. In short, the noteworthy result is that by making slight modifications, existing simple yet effective methods can be improved so that they can compete with the state-of-the-art algorithms or even outperform them. For instance, the original version of the max-RGB has a mean angular error of 7.78 on the RECommended ColorChecker dataset, however, when we modify this algorithm by using the blocks and informative image, its mean angular error reduces to 3.46. Also, when the weighted gray edge method is applied in scale-space, it outperforms most of the state-of-the-art algorithms on the NUS-8 Dataset as shown in Table 5. Moreover, when we modify the algorithms, the worst cases reduce significantly on all benchmarks. Since in color constancy it is known that it is important to improve the algorithms' performance for the worst cases, this is a valuable outcome.

5.4 Results of Multi-Illuminant Color Constancy

For mixed illumination conditions, we provide statistical results on the MIMO dataset, and on the Mixed-Illuminant Test Set. We report the results of existing methods by either running their codes or making use of already published works considered up-to-date and comprehensive. As aforementioned, the MIMO dataset contains two sets, i.e. the "Real-World" set and the "Laboratory" set, and the Real-World set includes images that are closer to the scenes we observe in our daily lives. Thus, compared to the Laboratory set it contains more complex scenes, which makes it more challenging than the Laboratory set [14]. The Mixed-Illuminant Test Set is a recent dataset and it contains synthetic images rendered with computer graphics. This benchmark includes images with different room

14

Table 6 Statistical results on the MIMO dataset. The best results are highlighted.

Algorithms		Real-World Mean	Med.	Laboratory Mean	Med.
Single-Illuminant	Do Nothing	8.8	8.9	10.6	10.5
	max-RGB [21]	6.8	5.7	7.8	7.6
	GW [20]	5.3	4.3	3.5	2.9
	SoG [22]	6.2	3.7	4.9	4.6
	1st - GE [24]	8.0	4.7	4.3	4.1
	wGE [25]	7.9	4.1	4.4	4.0
	DOCC [27]	7.9	5.0	4.6	4.4
	PCA-CC [12]	7.7	3.5	4.1	3.8
	LSRS [26]	4.9	3.8	3.9	3.5
	MSGP [13]	5.8	5.0	13.3	12.6
	BIO-CC [9]	5.0	4.3	4.2	4.1
	C3AE [40]	12.4	12.3	13.9	14.1
	SIIE [41]	5.9	5.1	9.0	9.0
	C5 [43]	11.9	13.0	7.0	7.1
Multi-Illuminant	LSAC [28]	4.9	4.2	2.7	2.5
	Gijsenij *et al.* w/ max-RGB [6]	4.2	3.8	5.1	4.2
	Gijsenij *et al.* w/ GW [6]	4.4	4.3	6.4	5.9
	Gijsenij *et al.* w/ 1st - GE [6]	9.1	9.2	4.8	4.2
	CRF w/ max-RGB [29]	4.1	3.3	3.0	2.8
	CRF w/ GW [29]	3.7	3.4	3.1	2.8
	CRF w/ 1st - GE [29]	4.0	3.4	2.7	2.6
	N-WB w/ max-RGB [35]	4.1	3.4	2.6	2.2
	N-WB w/ GW [35]	4.6	4.5	3.7	3.1
	N-WB w/ SoG [35]	4.2	3.8	2.8	2.3
	N-WB w/ 1st - GE [35]	4.7	3.6	2.5	2.2
	VM-CC w/ Bottom-Up [34]	5.0	4.0	3.7	3.4
	RM-CC [30]	5.2	4.3	3.2	2.7
	CCWF [31]	3.8	3.8	**1.6**	**1.5**
	CCATI [33]	3.8	3.8	2.6	2.6
	CCAFIS [32]	4.2	4.3	2.1	2.7
	GI (M=4) [14]	3.9	3.4	2.7	2.2
	GI (M=6) [14]	3.9	3.4	2.6	2.1
	CNNs-based CC [46]	**3.3**	3.1	2.3	2.2
	GAN-based CC [47]	3.5	2.9	-	-
	Proposed	3.7	**2.8**	2.8	2.5

715
716
717
718
719
720
721
722
723
724
725
726
727
728

Table 7 Statistical results on the Mixed-Illuminant Test Set. The best results are highlighted.

Algorithms	Mean	Med.
MSGP [13]	19.7	17.2
GI [14]	6.4	5.7
LSAC [28]	4.7	4.5
KNN White-Balance [48]	5.8	5.8
Interactive White-Balance [49]	5.8	5.6
Deep White-Balance [50]	**4.5**	4.2
Auto White-Balance for Mixed-Scenes [51]	4.7	**4.1**
Style White-Balance [52]	5.1	4.9
Proposed	4.6	4.4

layouts that are illuminated under varying mixed illumination conditions. Thus, in our experiments, we evaluate our approach not only for the real-world scenes but also for synthetically created "real-world-like" challenging images. In this section, first, we provide statistical results for the MIMO dataset and we discuss the outcomes. Then, we report our results for the Mixed-Illuminant Test Set.

According to the statistical analysis on the MIMO dataset (Table 6), our approach can provide pixel-wise estimates for mixed illumination conditions and it surpasses several methods, which are specifically designed to transform global color constancy algorithms into multi-illuminant cases. Also, while the statistical results seem competitive on both sets it is worth pointing out that our algorithm neither requires any prior information about the scene, i.e. the number of illuminants or the number of image segments/clusters nor it is trained using the illuminants from the MIMO dataset such as the GAN-based Color Constancy. We provide our pixel-wise estimations for both sets in Fig. 7.

Among all traditional algorithms, our method obtains the best angular error on average on the Real-World set. Compared to the learning-based techniques, we provide the best median angular error together with GAN-based CC, while the best mean angular error is obtained by CNNs-based CC. In the Laboratory set, we provide competitive results. Our performance is higher than in the Real-World set, which arises due to the complexity difference between these sets.

For the Mixed-Illuminant Test Set (Table 7), the proposed illumination estimation strategy obtains the second-best mean and third-best median angular error. It is important to stress that our approach is a learning-free method, hence, it is independent of data. The high-cost training phases are not incurred, which we see as our main advantage among the state-of-the-art works, i.e., Auto White-Balance for Mixed-Scenes [51], Style White-Balance [52].

As a final note, we would like to highlight the advantages and limitations of our algorithm which we addressed throughout our paper. Our algorithm is learning-free, thus we have lower computational costs since we do not have a training phase. Also, our method is easy to implement, and we utilize only

729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
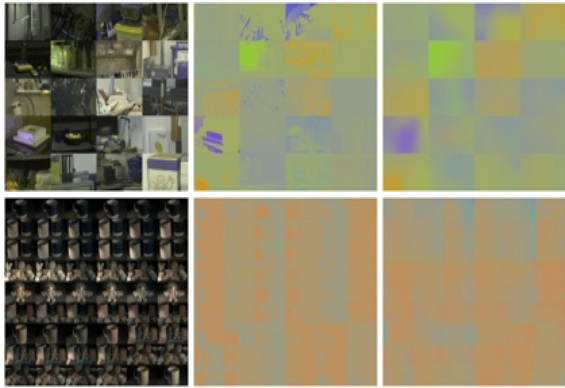756
757
758
759
760
761
762
763
764
765

15

Fig. 7 Results on MIMO dataset. (Left-to-right) Input scenes, pixel-wise ground truths, and pixel-wise estimation of the proposed method.

two parameters which is considerably lower compared to learning-based methods. Moreover, as stated in other color constancy studies not all pixels are informative, therefore we only consider the salient regions by utilizing the bright pixels which enable us to reduce the impact of the non-informative image elements. Furthermore, we carry out our computations using non-overlapping blocks that allow us to take the varying local statistics of the scenes into account which might not be possible while operating on a global scale. Therefore, modifying the algorithms by using the salient regions and blocks improves their performance on average significantly. Also, we use scale-space computations which are sensitive to the low-level features of images, to highlight the color features that can be missed while operating only on a single scale. On the other hand, in the field of color constancy, it is well-known that the methods utilizing statistical properties, in particular, traditional algorithms relying on the gray world assumption, have difficulty in estimating the color vector of the light source in uniformly colored images, i.e., scenes containing dominant grass and sky regions, since when there are a limited number of distinct colors, the gray world algorithm is not valid. To tackle this problem, most studies, and also our method, try to guide their approaches by only considering specific regions or pixels. While highlighting the color features and taking local statistics only in the salient regions into account allows us to improve the efficiency of our traditional algorithm, scenes containing large uniformly colored areas are more challenging than other images.

6 Conclusion

Color is an important feature not only for humans but also for various computer vision pipelines to perform accurate high-level vision tasks, i.e. object recognition and image dehazing. Due to the importance it holds, computational color constancy has been an attractive field of study, and researchers in this domain have developed many successful color constancy algorithms. Yet, the aim of researchers is not only to develop new techniques to find the color vector of the light source but also to improve existing methods by combining various strategies since this might help us to design simple yet efficient methods. From this motivation, we develop a computational color constancy algorithm based on the observation that space-average color and highest luminance patches carry significant cues for human color constancy. We estimate the color vector of the light source by assuming that on average, the world is achromatic and there are several bright image elements somewhere in the scene. We further assume that if the scene is gray on average, the shift of the brightest pixels from the achromatic value should be caused by the light source. We carry out our computations in scale-space where we find the estimations for each non-overlapping block individually by only considering the salient regions of the scene. Thereby, we take into account that surface orientations might vary throughout the scene, and not every image element is informative for performing color constancy. According to the experiments, the proposed algorithm achieves better performance than existing learning-free algorithms, while providing competitive results with learning-based methods. Furthermore, we demonstrate that the performance of several learning-free algorithms can be significantly improved by using particular steps of our algorithm. Lastly, we propose an approach that can convert our global color constancy algorithm into a method that is free from prior information about the scene for mixed illumination conditions which obtains competitive results compared to the state-of-the-art.

Declarations

Conflict of interest. The authors state that there are no conflicts of interest.

Data availability. The datasets analyzed in this study are publicly available.

16

Acknowledgement

References

[1] Zeki, S.: A Vision of the Brain. Blackwell Science, Oxford (1993)

[2] Ebner, M.: A communication-based model of consciousness. J. Artif. Intell. Consciousness **9**, 193–226 (2022)

[3] Ebner, M.: Color Constancy, 1st ed. Wiley Publishing, Hoboken, NJ, USA (2007)

[4] Gomez-Villa, A., Martin, A., Vazquez-Corral, J., Bertalmío, M.: Convolutional neural networks can be deceived by visual illusions. In: Gupta, A., Hoiem, D., Hua, G., Tu, Z. (eds.) Conf. Comput. Vision Pattern Recognit., Long Beach, CA, USA, pp. 12309–12317 (2019). IEEE/CVF

[5] Ebner, M.: Color constancy based on local space average color. Mach. Vision Appl. **20**(5), 283–301 (2009)

[6] Gijsenij, A., Lu, R., Gevers, T.: Color constancy for multiple light sources. IEEE Trans. Image Process. **21**(2), 697–707 (2011)

[7] Finlayson, G.D., Drew, M.S., Funt, B.V.: Color constancy: generalized diagonal transforms suffice. J. Opt. Soc. America A **11**, 3011–3019 (1994)

[8] Lo, Y.-C., Chang, C.-C., Chiu, H.-C., Huang, Y.-H., Chen, C.-P., Chang, Y.-L., Jou, K.: Clcc: Contrastive learning for color constancy. In: Conf. Comput. Vision Pattern Recognit., Virtual, pp. 8053–8063 (2021). IEEE/CVF

[9] Ulucan, O., Ulucan, D., Ebner, M.: BIO-CC: Biologically inspired color constancy. In: Hicks, Y., Mustafa, A., Bui, T., A., N., Angelova, A., Wannous, H., Xiao, B., Bianco, S., Wiles, O., Feng, Z., Ciocca, G., Ray, N., Dokania, P., Fang, Y., Aviles-Rivero, A.I. (eds.) Brit. Mach. Vision Conf., London, UK, p. (2022). BMVA Press

[10] Ulucan, O., Ulucan, D., Ebner, M.: Color constancy beyond standard illuminants. In: Berthoumieu, Y., Frossard, P. (eds.) Int. Conf. Image Process., Bordeaux, France, pp. 2826–2830 (2022). IEEE

[11] Joze, H.R.V., Drew, M.S., Finlayson, G.D., Rey, P.A.T.: The role of bright pixels in illumination estimation. In: Gille, J., Fredembach, C. (eds.) Color Imag. Conf., Los Angeles, CA, USA, pp. 41–46 (2012). Society for Imaging Science and Technology

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

[13] Qian, Y., Pertuz, S., Nikkanen, J., Kämäräinen, J.-K., Matas, J.: Revisiting gray pixel for statistical illumination estimation. arXiv preprint arXiv:1803.08326 (2018)

[14] Qian, Y., Kämäräinen, J.-K., Nikkanen, J., Matas, J.: On finding gray pixels. In: Gupta, A., Hoiem, D., Hua, G., Tu, Z. (eds.) Conf. Comput. Vision Pattern Recognit., Long Beach, CA, USA, pp. 8062–8070 (2019). IEEE/CVF

[15] Ono, T., Kondo, Y., Sun, L., Kurita, T., Moriuchi, Y.: Degree-of-linear-polarization-based color constancy. In: Dana, K., Hua, G., Roth, S., Samaras, D., Singh, R. (eds.) Conf. Comput. Vision Pattern Recognit., New Orleans, LA, USA, pp. 19740–19749 (2022). IEEE/CVF

[16] Ulucan, O., Ulucan, D., Ebner, M.: Block-based color constancy: The deviation of salient pixels. In: Narayanan, S., Kotropoulos, C., Ma, K., Petropulu, A. (eds.) Int. Conf. Acoust. Speech Signal Process., Rhodes Island, Greece, pp. 1–5 (2023). IEEE

[17] Ulucan, O., Ulucan, D., Ebner, M.: Multi-Scale Block-based Color Constancy. will appear in Eur. Signal Process. Conf. (2023)

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

[19] Das, P., Baslamisli, A.S., Liu, Y., Karaoglu, S., Gevers, T.: Color constancy by GANs: An experimental survey. arXiv preprint arXiv:1812.03085 (2018)

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

[21] Land, E.H., McCann, J.J.: Lightness and retinex theory. J. Opt. Soc. America A **61**, 1–11 (1971)

[22] Finlayson, G.D., Trezzi, E.: Shades of gray and colour constancy. In: x (ed.) Color Imag. Conf., Scottsdale, AZ, USA, pp. 37–41 (2004). Society for Imaging Science and Technology

[23] Gijsenij, A., Gevers, T., Van De Weijer, J.: Physics-based edge evaluation for improved color constancy. In: Essa, I., Kang, S.B., Pollefeys, M. (eds.) Conf. Comput. Vision Pattern Recognit., Miami, FL, USA, pp. 581–588 (2009). IEEE

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

[25] Gijsenij, A., Gevers, T., Van De Weijer, J.: Improving color constancy by photometric edge weighting. IEEE Trans. Pattern Anal. Mach. Intell. **34**, 918–929 (2011)

[26] Gao, S., Han, W., Yang, K., Li, C., Li, Y.: Efficient color constancy with local surface reflectance statistics. In: Fleet, D., Pajdla, T., Schiele, T. B. Tuytelaars (eds.) Eur. Conf. Comput. Vision, Zurich, Switzerland, pp. 158–173 (2014). Springer

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

[28] Ebner, M.: Combining white-patch retinex and the gray world assumption to achieve color constancy for multiple illuminants. Pattern Recognit. Proc. 25th DAGM Symp., 60–67 (2003)

[29] Beigpour, S., Riess, C., Van De Weijer, J., Angelopoulou, E.: Multi-illuminant estimation with conditional random fields. IEEE Trans. Image Process. **23**, 83–96 (2013)

[30] Zhang, X.-S., Gao, S.-B., Li, R.-X., Du, X.-Y., Li, C.-Y., Li, Y.-J.: A retinal mechanism inspired color constancy model. IEEE Trans. Image Process. **25**(3), 1219–1232 (2016)

[31] Hussain, M.A., Akbari, A.S.: Color constancy algorithm for mixed-illuminant scene images. IEEE Access **6**, 8964–8976 (2018)

[32] Hussain, M.A., Akbari, A.S., Mporas, I.: Colour constancy for image of non-uniformly lit scenes. Sensors **19**, 2242 (2019)

[33] Hussain, M.A., Akbari, A.S., Halpin, E.A.: Color constancy for uniform and non-uniform illuminant using image texture. IEEE Access **7**, 72964–72978 (2019)

[34] Gao, S.-B., Ren, Y.-Z., Zhang, M., Li, Y.-J.: Combining bottom-up and top-down visual mechanisms for color constancy under varying illumination. IEEE Trans. Image Process. **28**(9), 4387–4400 (2019)

[35] Akazawa, T., Kinoshita, Y., Shiota, S., Kiya, H.: N-white balancing: White balancing for multiple illuminants including non-uniform illumination. IEEE Access **10**, 89051–89062 (2022)

[36] Barron, J.T.: Convolutional color constancy. In: Ikeuchi, K., Schnörr, C., Sivic, J., Vidal, R. (eds.) IEEE Int. Conf. Comput. Vision, Santiago, Chile, pp. 379–387 (2015). IEEE/CVF

[37] Shi, W., Loy, C.C., Tang, X.: Deep specialized network for illuminant estimation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) Eur. Conf. Comput. Vision, Amsterdam, The Netherlands, pp. 371–387 (2016). Springer

[38] Barron, J.T., Tsai, Y.-T.: Fast fourier color constancy. In: Liu, Y., Rehg, J.M., Taylor, C.J., Wu,

Y. (eds.) Conf. Comput. Vision Pattern Recognit., Honolulu, HI, USA, pp. 886–894 (2017). IEEE/CVF

[39] Bianco, S., Cusano, C.: Quasi-unsupervised color constancy. In: Gupta, A., Hoiem, D., Hua, G., Tu, Z. (eds.) Conf. Comput. Vision Pattern Recognit., Long Beach, CA, USA, pp. 12212–12221 (2019). IEEE/CVF

[40] Laakom, F., Raitoharju, J., Iosifidis, A., Nikkanen, J., Gabbouj, M.: Color constancy convolutional autoencoder. In: Hou, Z., Hussain, A., Yang, C., Zeng, Z., Zhang, Y. (eds.) Symp. Ser. Comput. Intell., Xiamen, China, pp. 1085–1090 (2019). IEEE

[41] Afifi, M., Brown, M.S.: Sensor-independent illumination estimation for DNN models. Brit. Mach. Vision Conf. (2019)

[42] Laakom, F., Passalis, N., Raitoharju, J., Nikkanen, J., Tefas, A., Iosifidis, A., Gabbouj, M.: Bag of color features for color constancy. IEEE Trans. Image Process. **29**, 7722–7734 (2020)

[43] Afifi, M., Barron, J.T., LeGendre, C., Tsai, Y.-T., Bleibel, F.: Cross-camera convolutional color constancy. In: Damen, D., Hassner, T., Pal, C., Sato, Y. (eds.) Int. Conf. Comput. Vision, Virtual, pp. 1981–1990 (2021). IEEE/CVF

[44] Zini, S., Buzzelli, M., Bianco, S., Schettini, R.: COCOA: Combining color constancy algorithms for images and videos. IEEE Trans. Comput. Imag. **8**, 795–807 (2022)

[45] Domislović, I., Vršnak, D., Subašić, M., Lončarić, S.: One-net: Convolutional color constancy simplified. Pattern Recognit. Letters **159**, 31–37 (2022)

[46] Bianco, S., Cusano, C., Schettini, R.: Single and multiple illuminant estimation using convolutional neural networks. IEEE Trans. Image Process. **26**(9), 4347–4362 (2017)

[47] Das, P., Liu, Y., Karaoglu, S., Gevers, T.: Generative models for multi-illumination color constancy. In: Forsyth, D., Gkioxari, G., Tuytelaars, T., Yang, R., Yu, J. (eds.) Conf. Comput. Vision Pattern Recognit., Montreal, BC, Canada, pp. 1194–1203 (2021). IEEE/CVF

[48] Afifi, M., Price, B., Cohen, S., Brown, M.S.: When color constancy goes wrong: Correcting improperly white-balanced images. In: Gupta, A., Hoiem, D., Hua, G., Tu, Z. (eds.) Conf. Comput. Vision Pattern Recognit., Long Beach, CA, USA, pp. 1535–1544 (2019). IEEE/CVF

[49] Afifi, M., Brown, M.S.: Interactive white balancing for camera-rendered images. arXiv preprint arXiv:2009.12632 (2020)

[50] Afifi, M., Brown, M.S.: Deep white-balance editing. In: Liu, C., Mori, G., Saenko, K., Savarese, S. (eds.) Conf. Comput. Vision Pattern Recognit., Virtual, pp. 1397–1406 (2020). IEEE/CVF

[51] Afifi, M., Brubaker, M.A., Brown, M.S.: Auto white-balance correction for mixed-illuminant scenes. In: Anand, S., Farrell, R., Souvenir, R., Zhao, C. (eds.) Winter Conf. Appl. Comput. Vision, Waikoloa, HI, USA, pp. 1210–1219 (2022). IEEE/CVF

[52] Kınlı, F., Yılmaz, D., Özcan, B., Kıraç, F.: Modeling the lighting in scenes as style for auto white-balance correction. In: Crandall, D., Gong, B., Lee, Y.J., Souvenir, R., Yu, S. (eds.) Winter Conf. Appl. Comput. Vision, Waikoloa, HI, USA, pp. 4903–4913 (2023). IEEE/CVF

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

[54] Buzzelli, M., Zini, S., Bianco, S., Ciocca, G., Schettini, R., Tchobanou, M.K.: Analysis of biases in automatic white balance datasets and methods. Color Res. Appl. **48**(1), 40–62 (2023)

[55] Ulucan, D., Ulucan, O., Ebner, M.: CC-NORD: A Camera-Invariant Global Color Constancy Dataset. will appear in Eur. Signal Process. Conf. (2023)

[56] Land, E.H.: The retinex theory of colour vision. Proc. Roy. Institution Gr. Britain **47**, 23–58 (1974)

[57] Linnell, K.J., Foster, D.H.: Space-average scene

19

colour used to extract illuminant information. John Dalton's Colour Vision Legacy, 501–509 (1997)

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

[59] Uchikawa, K., Fukuda, K., Kitazawa, Y., MacLeod, D.I.: Estimating illuminant color based on luminance balance of surfaces. J. Opt. Soc. America A **29**(2), 133–143 (2012)

[60] Morimoto, T., Kusuyama, T., Fukuda, K., Uchikawa, K.: Human color constancy based on the geometry of color distributions. J. Vision **21**(3), 7–7 (2021)

[61] Geusebroek, J.-M., Van Den Boomgaard, R., Smeulders, A.W., Dev, A.: Color and scale: The spatial structure of color images. In: Vernon, D. (ed.) Eur. Conf. Comput. Vision, Dublin, Ireland, pp. 331–341 (2000). Springer

[62] Ebner, M., Tischler, G., Albert, J.: Integrating color constancy into JPEG2000. IEEE Trans. Image Process. **16**, 2697–2706 (2007)

[63] Li, B., Xu, D., Lee, M.H., Feng, S.-H.: A multi-scale adaptive grey world algorithm. Trans. Inf. Syst. **90**(7), 1121–1124 (2007)

[64] Ulucan, D., Ulucan, O., Ebner, M.: Intrinsic image decomposition: Challenges and new perspectives. In: Imai, F., Distante, C., Battiato, S. (eds.) Int. Conf. Image Process. Vision Eng., Prague, Czech Republic, pp. 57–64 (2023). INSTICC

[65] Ulucan, D., Ulucan, O., Ebner, M.: Multi-scale surface normal estimation from depth maps. In: Imai, F., Distante, C., Battiato, S. (eds.) Int. Conf. Image Process. Vision Eng., Prague, Czech Republic, pp. 47–56 (2023). INSTICC

[66] Wang, F., Wang, W., Wu, D., Gao, G.: Color constancy via multi-scale region-weighed network guided by semantics. Frontiers Neurorobotics **16**, 841426 (2022)

[67] Ebner, M., Hansen, J.: Depth map color constancy. Bio-Algorithms Med-Syst. **9**, 167–177 (2013)

[68] Hemrit, G., Finlayson, G.D., Gijsenij, A., Gehler, P., Bianco, S., Funt, B., Drew, M., Shi, L.: Rehabilitating the colorchecker dataset for illuminant estimation. In: Bonnier, N., Vanrell, M. (eds.) Color Imag. Conf., Vancouver, BC, Canada, pp. 350–353 (2018). Society for Imaging Science and Technology

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

[70] Gehler, P., Rother, C., Blake, A., Minka, T., Sharp, T.: Bayesian color constancy revisited. In: Boyer, K., Shah, M., T., S.-M. (eds.) Conf. on Comput. Vision Pattern Recognit., pp. 1–8 (2008). IEEE